



HAL
open science

Prise en compte du rôle de la diversité microbienne dans la simulation de la dynamique de la matière organique du sol (MOS) dans un contexte de transition vers l'agro-écologie

Benjamin Louis

► To cite this version:

Benjamin Louis. Prise en compte du rôle de la diversité microbienne dans la simulation de la dynamique de la matière organique du sol (MOS) dans un contexte de transition vers l'agro-écologie. Sciences du Vivant [q-bio]. AGROCAMPUS OUEST, 2016. Français. NNT: . tel-02796104

HAL Id: tel-02796104

<https://hal.inrae.fr/tel-02796104v1>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

AGRO CAMPUS

OUEST



UNIVERSITÉ
EUROPÉENNE
DE BRETAGNE

Benjamin LOUIS • 24 novembre 2016

Thèse AGROCAMPUS OUEST
sous le label de l'Université européenne de Bretagne
pour obtenir le grade de
DOCTEUR D'AGROCAMPUS OUEST
Spécialité Biologie & Agronomie

ÉCOLE DOCTORALE • Vie-Agro-Santé (VAS)

LABORATOIRE D'ACCUEIL • UMR INRA / AGROCAMPUS OUEST
Sol Agro et hydrosystème spatialisation (SAS)

**Prise en compte du rôle de la
diversité microbienne dans la
simulation de la dynamique de la
matière organique du sol (MOS)
dans un contexte de transition
vers l'agro-écologie**

Yannick OUTREMAN

Professeur, AGROCAMPUS OUEST, UMR INRA-UR1-AO IGEPP /
président

Jérôme BALESSENT

Directeur de recherche, INRA Aix-en-Provence / *rapporteur*

David MAKOWSKI

Directeur de recherche, INRA Versailles-Grignon / *rapporteur*

Thomas POMMIER

Chargé de recherche, Université Lyon 1 / *examineur*

Safya MENASSERI

Maître de conférences, AGROCAMPUS OUEST, UMR SAS / *directrice
de thèse*

Philippe LETERME

Professeur, AGROCAMPUS OUEST, UMR SAS / *co-directeur de thèse*

Valérie VIAUD

Chargée de recherche, UMR SAS / *co-encadrante de thèse*

Jean-Luc HARDY

Directeur R&D, Triskalia Landerneau / *membre invité*

Avant-propos

Thèse réalisée à UMR 1069 Sol Agro et Hydrosystème Spatialisation
Agrocampus Ouest, Bâtiment Alain Ruellan (n° 13)
65 rue de Saint-Brieuc CS 84215
F-35042 Rennes cedex

Tél : +33 (0)2 23 48 50 00

Fax : +33 (0)2 23 48 55 10

Web : <http://www6.rennes.inra.fr/umrsas/>

Sous la direction de Safya Menasseri Safya.Menasseri@agrocampus-ouest.fr
Philippe Leterme Philippe.Leterme@agrocampus-ouest.fr

Co-encadrement Valérie Viaud valerie.viaud@rennes.inra.fr

Financement Chaire école-entreprise AEI (<http://www.chaire-aei.fr/>)
Allocation de recherche de la Région Bretagne



Remerciements

Je me suis souvent posé une question au cours de la thèse : "en quoi la diversité est une bonne chose ?". Cette question concernait évidemment la diversité microbienne et son implication dans la dynamique du carbone. Mais elle trouve sa place dans d'autres domaines : après tout, les médecins recommandent de diversifier notre alimentation ; en sport, diversifier son entraînement est nécessaire à une bonne progression ; une musique composée d'une seule note ne présente aucun intérêt. Lorsque le thésard « termine » son travail, le mérite lui revient souvent. Pourtant, c'est grâce à toute une diversité de personnes que ce travail est ce qu'il est. Je voudrais prendre quelques lignes pour tenter de les remercier à la hauteur de ce qu'ils ont apporté.

J'aimerais tout d'abord souligner l'importance de l'encadrement d'un thésard. J'ai eu la chance d'être entouré de trois personnes qui ont cru à ce travail, qui m'ont fait confiance et avec qui ces trois ans de thèse ont été très agréables. Safya, Valérie, Philippe, vous m'avez apporté implication, bienveillance, disponibilité et confiance. Sentir que je pouvais m'exprimer librement dans ce travail tout en sachant que je pouvais compter sur vous a été très agréable. Vous m'avez plusieurs fois poussé dans mes retranchements pour expliquer, convaincre, réfléchir, questionner. Vous avez fait de cette thèse un véritable challenge. Nul doute que vous avez permis à ce travail de tenir debout. C'est avec un grand plaisir que j'ai traversé cette aventure avec vous. Merci d'avoir partagé avec moi vos qualités humaines et scientifiques. Je garderais le souvenir d'une équipe à laquelle je suis fier d'avoir appartenu pour un temps.

Beaucoup vous le diront, travailler au sein de l'UMR SAS est une chance. L'ambiance qui règne dans cette unité permet de s'y sentir très vite à l'aise. C'est une qualité que vous devez conserver ! Merci aux anciens doctorants et autres contractuels : Pauline, Rémi, Alexis, Claude, Guillaume, Mariana, Michael, Thibaud, Pierre-Louis, Olivier, Camille and Co. Merci à Sylvain pour les ressources en magnésium et les discussions avec Alexis dans le bureau. Merci à Michèle pour ton efficacité. Merci à Blandine et Valérie de m'avoir offert du travail en attendant les

financements. Merci à Monique pour ta capacité à trouver n'importe quelle documentation et pour ton dynamisme. Merci à tous ceux avec qui j'ai pu échanger de près ou de loin.

Beaucoup de décisions ont été prises au cours des échanges qu'il y a pu avoir lors des trois comités de thèse. La qualité de ces échanges a été permise par les personnes qui ont accepté de faire partie de ce comité. Merci à Didier Andrivon (UMR IGEPP, INRA, tuteur VAS), Bertrand Pinel (Terrena), Étienne Rivot (UMR ESE, Agrocampus Ouest), Tiphaine Chevallier (UMR Eco&Sols, IRD) , Jean-Luc Hardy (Triskalia), Jean-Luc Le Bénézic (Triskalia), Pierre-Alain Maron (UMR Agroécologie, INRA), Bruno Mary (US AgroImpact, INRA) et Thierry Oboyet (Triskalia). Merci plus particulièrement à Pierre-Alain, référent microbiologiste de cette thèse avec qui les discussions ont été essentielles. Merci de m'avoir fait confiance pour l'utilisation des données DIMIMOS et pour ta pédagogie sur le fonctionnement des micro-organismes du sol. Merci également à Étienne pour tes conseils qui ont facilité mon entrée dans le monde extraordinaire des statistiques bayésiennes.

Qui dit thèse dit financement ! Merci tout d'abord à la chaire AEI de travailler pour un meilleur transfert des connaissances et des besoins entre le monde de la recherche et de l'entreprise, et à ce titre d'avoir participé au financement de cette thèse. Merci également à la région Bretagne.

Au moment où j'écris ces lignes, le travail rapporté dans cette thèse n'a pas encore été évalué. Cette évaluation essentielle est possible par les personnes qui ont accepté d'être membre de mon jury et que j'aimerais remercier. Merci tout d'abord aux rapporteurs, Jérôme Balesdent (UR GSE, INRA) et David Makowski (UMR Agronomie, INRA) qui ont la lourde tâche de lire en détail ce manuscrit. Merci également à Thomas Pommier (Écologie microbienne, Université Lyon 1), Yannick Outreman (UMR 1349 IGEP, Agrocampus Ouest) et Jean-Luc Hardy (Triskalia, Chaire AEI). Je suis sûr que nos échanges seront très intéressants.

Je me dois également de remercier ceux qui m'ont initié à la recherche sur les problématiques liées au sol, du besoin de le comprendre et le protéger. Merci à Dominique Arrouays et à Nicolas Saby (US Infosol, INRA) de m'avoir mis le pied à l'étrier. Mes premiers pas à Orléans étaient vraiment top !

Je crois que ce qui m'a le plus marqué avec la thèse, c'est sa présence chaque jour, chaque heure, chaque minute, comme un bruit de fond qui ne s'éteint jamais. Je dois l'avouer, ce bruit est parfois aussi agréable à écouter qu'une musique de Pink Floyd, mais parfois aussi désagréable

qu'un acouphène. Et dans ces moments, il est bon d'être entouré de personnes dont les qualités étouffent ce bruit.

Il y a beaucoup de personnes que j'aimerais remercier. Je pourrais commencer par les coloc et amis rennais. Ce sont eux que je vois chaque jour et qui me permettent de penser à autres choses, d'échanger, de partager sur tout et rien. Dans l'ordre d'apparition merci à Camille C, Thibault, Lise, Marie, Bigoudie, Mo, Coco, Clémence, Antoine, Camille A, Jonas, Benoit et le p'ti nouveau Michael. Et puis y'a les copains qu'on voit moins, mais avec qui rien ne change. La pote la plus vieille, Claire, l'unique, la fidèle, l'éternelle, l'artiste ! Les gars de la chambre 4, Simon et Math, sans oublier Laura. Mon BS préféré, Benoit, avec qui les soirées sont uniques ! Anne-Choute, le p'ti lutin. Et la clique d'halieutes , Féfé, Schnappy, ainsi que les Badouls et Béchaux ! Y'a les grimpeurs aussi ! Et y'en a pleins, avec qui on s'évade le temps d'une heure, d'une après-midi ou de plusieurs jours au contact de la résine ou encore mieux, du rocher ! Enfin, last but not least, les Chesnoysiens ! Ceux avec qui il est difficile d'être au complet et dont un week-end ne se rate pour rien au monde, pas même une rédaction de thèse. Il y a mon fiston, Luc et la belle future maman, Steph. Il y a la (grosse) déjà maman, Anaïs, et son "ohm", Olive. Il y a les gars du bois : le Gros, Tritri, et le petit, Didi. Il y a le gars de la terre, Vava. Il y a la fille qui n'aimait pas les (mes) pieds, Cha. Il y a la Touz', Fanny. Il y a Polo, juste Polo, et sa belle, Perrine. Refaire le monde, échanger, partager, boire un verre, jouer, rigoler et recommencer. Ce sont des moments précieux. Merci !

Ces remerciements s'achèvent bientôt. Mais il reste des personnes à remercier sans qui je ne serais sûrement pas là. Pour des raisons évidentes, il s'agit bien sûr de mes parents ! Leur dévouement pour le bonheur de leurs enfants est sans limite. Parfois trop, je leur ai souvent dit : "pensez à vous !". Il ne m'écoute pas. J'espère un jour pouvoir leur donner autant qu'ils m'ont donné. Nos chemins de pensée divergent parfois, mais sachez que je suis entièrement reconnaissant de ce que vous m'avez donné, sans rien attendre en retour. Merci de nous soutenir, merci de me soutenir. Bien sûr, un big-up aux frangines aussi. Paupau, ne laisse jamais une personne te faire croire que tu vaux moins qu'elle ! Mathilde, ne te laisse jamais croire que tu vaux moins que les autres ! Big up à M. Vernier également !

Voilà.... J'ai fini....

Ah non ! Il y a toi ! Toi qui j'en suis sûr va lire avec intérêt ces remerciements. Toi qui a partagé cette aventure avec moi. Toi qui partage toutes mes aventures depuis 10 ans maintenant. Toi qui met de la diversité dans toutes mes journées. Et bien, à toi plus particulièrement, merci ! To be continued...

Cette fois c'est bon ! Oui je sais c'est long. C'est long parce que comme toute aventure, elle ne vaut le coup que si elle est accompagnée. Accompagnée de diversité qu'il faut remercier. Parce que si pour les micro-organismes du sol, il reste encore du boulot pour comprendre ce qu'apporte la diversité, pour les relations humaines, j'ai la certitude qu'elle est le remède à beaucoup de maux, qu'elle est ce qui nous fait avancer. Peut-être plus aujourd'hui que jamais. Alors je dédie ce travail à tous ceux qui, comme vous, cultivent et chérissent la diversité.

À eux aussi je dis merci !

Résumé

La gestion de la matière organique du sol (MOS), impactée par les pratiques agricoles et impactant les flux de carbone (C), constitue un levier important de la transition agro-écologique et nécessite de disposer de modèles de prédiction adaptés. Le travail de thèse a pour ambition de proposer et d'évaluer une méthode de prise en compte de la diversité microbienne dans un modèle mécaniste de dynamique du C dans le but qu'il soit utilisé à des fins opérationnelles. Il s'appuie sur 2 jeux de données issus de programmes de recherche acquis à 2 échelles spatiales : l'échelle nationale et l'échelle d'un bassin versant de 5km². Ces jeux de données comportent des mesures de diversité microbienne et des cinétiques de minéralisation du C sur 80 jours en conditions contrôlées, avec et sans apport de paille de blé marquée au ¹³C. Les sols étudiés sont issus de parcelles agricoles occupées par des prairies et des rotations de cultures annuelles.

Un état de l'art a montré que l'approche dominante de la prise en compte explicite et détaillée de la diversité microbienne se faisait au sein de modèles conceptuels de dynamique du C souvent basés sur une multiplication de compartiments microbiens et des flux associés. L'approche que nous proposons est quant à elle basée sur la prise en compte des indices de diversité microbienne (i.e. richesse, équitabilité, indice de Shannon et inverse de l'indice de Simpson) au travers de fonctions de modulation des paramètres des cinétiques d'ordre 1.

Une première étape, basée sur des modèles statistiques additifs généralisés (GAM), a permis d'identifier les indices de diversité pouvant expliquer la dynamique du C avec des importances relatives comparables aux propriétés des sols plus classiques (C/azote, phosphore, pH, texture). Dans cette étape, les résultats ont montré i) une relation négative entre la diversité bactérienne et les flux de minéralisation de la MOS et une relation positive entre la diversité fongique et ii) les flux issus de la minéralisation de la MOS ainsi que iii) les flux issus de la minéralisation du résidu sur la période 7-28 jours.

Les relations observées ont permis de proposer des fonctions de modulation respectivement dépendantes i) de la diversité bactérienne pour moduler le paramètre de rendement d'utilisation

du carbone de la MOS, ii) de la diversité fongique pour moduler la constante de décomposition de la MOS et iii) de la diversité fongique pour moduler la constante de décomposition du sous-pool cellulosique du résidu.

Ces fonctions ont été intégrées dans un modèle générique de dynamique du C avec un nombre de paramètres limité pour faciliter sa calibration. L'ensemble a été calibré à l'aide de méthodes bayésiennes. Globalement, la prise en compte de la fonction de modulation liée à la diversité bactérienne a permis d'améliorer la prédiction des flux issus de la minéralisation de la matière organique du sol, tandis que les fonctions liées à la diversité fongique n'ont permis aucune amélioration.

En conclusion, la méthode d'intégration de la diversité microbienne dans les modèles de dynamique du C proposée ici constitue une approche pertinente pour une représentation parcimonieuse de son influence dans des modèles génériques de dynamique du C. Il s'agit d'un cadre méthodologique complémentaire à la prise en compte explicite et détaillée, dominante dans la littérature, de la diversité microbienne dans ces modèles. Nous pensons que ce travail pourra aider à l'intégration d'information importante comme la diversité microbienne dans des modèles opérationnels de dynamique du C. Pour finir, nous proposons des pistes d'amélioration sous formes d'expérimentation et de modélisation pour la construction des fonctions de modulation liée à la diversité microbienne.

MOTS CLÉS : sol, matière organique, résidu de cultures, diversité bactérienne et fongique, modélisation

Abstract

Soil organic matter (SOM) management is controlled by agro-cultural practices and has a strong impact on carbon (C) fluxes. It represents a critical controlling factor of agro-ecological transition in agriculture and needs to have appropriate predictive models. The goal of this thesis was to suggest and assess a way to integrate microbial diversity in an operational mechanistic model of C dynamic. We used datasets from 2 research programs obtained at two spatial scales: one at national scale and one at landscape scale in a 5 km² catchment. These datasets contain measures of microbial diversity and C mineralisation kinetics from controlled conditions incubations during 80 days, with and without ¹³C labelled wheat straw. Soil was sampled from grasslands and croplands.

A state of the art showed that the main approach to integrate microbial diversity is through new conceptual microbial compartments and associated fluxes. The approach we proposed is based on the construction of modulation functions depending on microbial diversity indices (i.e. richness, evenness, Shannon index and inverse Simpson index) to adjust the value of first order kinetics parameters.

In a first step, based on statistical generalized additive models (GAM), we identified the microbial diversity indices relevant to explain C dynamic with relative importance similar to others classical soil properties (C/Nitrogen, pH, texture, phosphorus). The results showed i) a negative relation between bacterial diversity and SOM mineralisation fluxes and a positive relation between fungal diversity and ii) SOM mineralisation fluxes and iii) residue mineralisation fluxes during the 7-28 days period.

The observed relations allowed us to propose modulation functions respectively depending on i) bacterial diversity to adjust C yield utilisation parameter of SOM kinetic, ii) fungal diversity to adjust decomposition constant parameter of SOM kinetic and iii) fungal diversity to adjust decomposition constant parameter of cellulosic residue sub-pool kinetic.

These functions were integrated in a generic C dynamic model with a limited number of

parameters in order to facilitate the calibration. Both modulation functions and model were calibrated through bayesian inference. Overall, the modulation function depending on bacterial diversity led to an improvement of SOM mineralisation fluxes predictions, while fungal based modulation functions led to no improvement.

To conclude, the proposed method to integrate microbial diversity was a relevant approach towards a parsimonious representation of microbial diversity influence on C dynamic in generic model. It forms a complementary methodological framework to the dominant explicit and detailed way of microbial diversity representation these models. We think this work could help taking into account important information as microbial diversity in more operational C dynamic models. In the end, we propose experiments and modelling perspectives to improve the construction of the modulation functions.

KEY WORDS : soil, organic matter, crop residue, bacterial and fungal diversity, modeling

Table des matières

| | |
|--|-------------|
| Avant-propos | 1 |
| Remerciements | v |
| Résumé | vii |
| Abstract | ix |
| Table des matières | xiii |
| Table des figures | xvi |
| Liste des tableaux | xvii |
| Introduction | 1 |
| Contexte : une transition agro-écologique nécessaire de l’agriculture | 1 |
| La gestion de la dynamique de la matière organique du sol comme levier d’action | 2 |
| Les micro-organismes dans la dynamique de la MOS | 5 |
| Utilité des modèles mathématiques de dynamique de la MO | 6 |
| Objectif général de la thèse | 7 |
| 1 État de l’art | 9 |
| 1.1 La diversité des micro-organismes du sol | 10 |
| 1.1.1 Mesures de la diversité microbienne et mise en relation avec le fonctionne- ment des sols | 10 |
| 1.1.2 Facteurs influençant la diversité microbienne | 13 |
| 1.2 Modéliser la dynamique de la MOS | 15 |
| 1.2.1 Structure générale des modèles de dynamique de MOS | 15 |

| | | |
|----------|--|-----------|
| 1.2.2 | Complexification des modèles et identifiabilité des paramètres | 18 |
| 1.2.3 | Prise en compte des facteurs influençant la dynamique de MOS | 20 |
| 1.3 | Intégration de la diversité microbienne dans les modèles de dynamique de C et N | 20 |
| 1.3.1 | Introduction | 20 |
| 1.3.2 | Role of microbial diversity in soil carbon and nitrogen dynamics | 22 |
| 1.3.3 | Soil microbial diversity in current soil carbon and nitrogen dynamics models | 27 |
| 1.3.4 | What way towards predictive modelling approaches including soil microbial diversity ? | 34 |
| 1.3.5 | Conclusion | 38 |
| 1.4 | Objectifs spécifiques et organisation du travail de thèse | 39 |
| 2 | Matériels disponibles et méthodes générales | 45 |
| 2.1 | Les données disponibles | 45 |
| 2.1.1 | Deux projets ANR comme source des données | 45 |
| 2.1.2 | Propriétés des sols échantillonnés | 47 |
| 2.1.3 | Dynamiques du carbone dans les sols échantillonnés | 54 |
| 2.2 | Méthodes générales d'analyse | 55 |
| 2.2.1 | Lien entre les dynamiques du carbone et les propriétés des sols | 55 |
| 2.2.2 | Intégration et évaluation de la diversité microbienne dans un modèle mé- caniste de dynamique du carbone | 56 |
| 3 | Les indices de diversité microbienne peuvent-ils expliquer la dynamique du C ? | 59 |
| 3.1 | Pertinence des indices de diversité pour expliquer la variabilité de dynamique du C de sols échantillonnés à l'échelle nationale | 60 |
| 3.1.1 | Introduction | 60 |
| 3.1.2 | Materials and methods | 62 |
| 3.1.3 | Results | 69 |
| 3.1.4 | Discussion | 73 |
| 3.1.5 | Conclusion | 80 |
| 3.2 | Contribution temporelle des indices de diversité pour expliquer la variabilité de dynamique du C entre des sols échantillonnés à l'échelle d'un paysage | 81 |
| 3.2.1 | Introduction | 81 |

| | | |
|----------|---|------------|
| 3.2.2 | Matériel et méthodes | 82 |
| 3.2.3 | Résultats | 83 |
| 3.2.4 | Discussion | 90 |
| 3.2.5 | Conclusion | 94 |
| 3.3 | Discussion de l'implication des résultats quant à l'intégration des indices de diversité dans un modèle de dynamique du C | 95 |
| 3.3.1 | Comparaison des résultats DIMIMOS et MOSAIC | 95 |
| 3.3.2 | Discussion sur les conséquences pour l'étape de modélisation | 97 |
| 4 | Intégration des indices de diversité dans un modèle mécaniste de dynamique du C | 103 |
| 4.1 | Introduction | 103 |
| 4.2 | Matériels et méthodes | 105 |
| 4.2.1 | Données utilisées | 105 |
| 4.2.2 | Choix du modèle de dynamique du C | 105 |
| 4.2.3 | Stratégie de calibration | 109 |
| 4.2.4 | Étapes de construction et d'évaluation des fonctions de modulation de la décomposition | 112 |
| 4.3 | Résultats | 125 |
| 4.3.1 | Distributions <i>a priori</i> et paramètres fixes | 125 |
| 4.3.2 | Calibration des fonctions de modulation des paramètres des cinétiques de flux de C | 128 |
| 4.3.3 | Qualités des modèles et évaluation des combinaisons des fonctions de modulation | 131 |
| 4.4 | Discussion | 141 |
| 4.4.1 | Comportement et qualité globale du modèle | 141 |
| 4.4.2 | Impact global de la modulation de la valeur des paramètres | 143 |
| 4.4.3 | Prise en compte spécifique de la diversité microbienne dans le modèle au regard des autres propriétés du sol | 145 |
| 4.4.4 | Retour sur le choix de la méthode pour répondre à nos objectifs | 146 |
| 4.5 | Conclusion | 147 |

| | | |
|----------|---|------------|
| 5 | Discussion générale | 149 |
| 5.1 | Apports et limites du cadre de modélisation de la dynamique du C proposé . . . | 149 |
| 5.1.1 | Cadre de modélisation proposé | 149 |
| 5.1.2 | Discussion sur les approches choisies pour la modélisation | 152 |
| 5.1.3 | Avantages et limites des méthodes utilisées | 157 |
| 5.1.4 | Conclusion sur l'intérêt du cadre de modélisation proposé | 159 |
| 5.2 | Implication des résultats sur l'influence de la diversité microbienne dans la modulation des cinétiques de décomposition et minéralisation du C | 161 |
| 5.2.1 | Cohérence du rôle de la diversité bactérienne dans la minéralisation de la MOS entre résultats, processus et modélisation | 161 |
| 5.2.2 | Un rôle ambivalent de la diversité fongique dans la dynamique du C . . . | 164 |
| 5.3 | Limites et perspectives de recherche | 166 |
| 5.3.1 | Des données riches mais un besoin d'une harmonisation des mesures de diversité | 166 |
| 5.3.2 | Perspectives immédiates de modélisation | 166 |
| 5.3.3 | Mise en place d'une expérimentation pour la confirmation des hypothèses et une meilleure calibration de modèle | 168 |
| 5.3.4 | Quels autres processus peut-on envisager de modéliser pour une meilleure gestion de la MOS? | 171 |
| 5.3.5 | Perspectives d'application pour les transitions agro-écologiques | 175 |
| | Conclusion | 177 |
| | Annexes | 181 |
| A | Information sur les sols de DIMIMOS | 181 |
| B | Information sur les sols de MOSAIC | 183 |
| C | Résultats de la sélection GAM avec les phyla en guise de diversité pour les sols de DIMIMOS | 189 |
| D | Relations entre les covariables et les cinétiques de minéralisation pour MOSAIC | 191 |

| | | |
|----------|---|------------|
| E | Tableaux de références et valeurs associées pour la construction des distributions <i>a priori</i> | 197 |
| E.1 | Paramètres des cinétiques du carbone | 198 |
| E.2 | Paramètres liés au fractionnement Van Soest | 199 |
| E.2.1 | Proportion des fractions Van Soest dans la paille de blé | 199 |
| E.2.2 | Proportion de C dans chaque fraction Van Soest | 199 |
| | Bibliographie | 201 |

Table des figures

| | | |
|-----|--|----|
| 1 | Dynamique simplifiée du carbone de la matière organique (MO) dans le sol. | 3 |
| 1.1 | Modalités de description de la diversité microbienne et possibilité de lien avec le fonctionnement des sols. | 11 |
| 1.2 | Représentation schématique des flux possibles dans un modèle simplifié de dynamique de MO un compartiment substrat (indice S, jaune) et un compartiment microbien (indice B, vert) | 15 |
| 1.3 | Représentation schématique de différentes méthodes de complexification d'un modèle de dynamique de C dans le sol. | 19 |
| 1.4 | Conceptual diagram of the relationship between microbial diversity and soil organic matter (SOM) dynamics. | 27 |
| 1.5 | Diagram of microbial diversity in current carbon and nitrogen dynamics models. | 29 |
| 1.6 | Diagram of model design (solid arrows with numbers) and stages for integrating microbial community descriptors (dashed arrows with letters). | 33 |
| 1.7 | Diagram of a simple model of decomposition of a substrate. | 36 |
| 1.8 | Représentation de la diversité microbienne comme un intermédiaire intéressant pour comprendre le lien entre une situation pédo-climatique associée à des pratiques agricoles et la dynamique de la matière organique du sol (SOM). | 40 |
| 2.1 | Localisation des sols échantillonnés issus du RMQS en France Métropolitaine. | 46 |
| 2.2 | Localisation du bassin versant de Kervidy-Naizin (A) et des sols échantillonnés (B). | 47 |
| 2.3 | Représentation de variabilité des propriétés des sols pour chacune des bases de données (DIMIMOS et MOSAIC) et chacun des modes d'usage (Cultures et Prairies). | 50 |
| 2.4 | Matrice des corrélations de Spearman pour les propriétés des sols de DIMIMOS. | 52 |
| 2.5 | Matrice des corrélations de Spearman pour les propriétés des sols de MOSAIC. | 53 |
| 2.6 | Variabilité du CO ₂ respiré cumulé, marqué ou non à l'isotope C, dans les sols témoins et les sols amendés. | 54 |
| 3.1 | Predictor selection procedure using generalized additive models. | 68 |
| 3.2 | Variability in soil property values in croplands and grasslands. | 70 |
| 3.3 | Variability in mineralisation rate and priming effect (PE) in croplands and grasslands at each sampling time. | 71 |
| 3.4 | Selected soil properties and associated estimated relationships for each mineralisation parameter. | 76 |
| 3.5 | Taux de minéralisation en fonction des périodes différenciées des cinétiques. | 84 |

| | | |
|------|--|-----|
| 3.6 | Variabilité des taux moyens des cinétiques de minéralisation pour les périodes 3-7 jours (bleu), 7-28 jours (orange) et 28-80 (jours). | 85 |
| 3.7 | Importance des covariables sélectionnées par période pour les cinétiques de minéralisation du sol témoins ($R_{s,control}$) et amendé ($R_{s,amended}$), de minéralisation du résidu (R_r) et du priming effect (PE). | 87 |
| 3.8 | Influence des différentes propriétés du sol dans un modèle simple de dynamique du C dans des sols témoins et des sols amendés. | 98 |
| 4.1 | Modèle de dynamique du carbone. | 106 |
| 4.2 | Logarithme de l'indice de colinéarité du modèle en fonction du nombre de paramètres à estimer. | 110 |
| 4.3 | Diagramme des étapes de construction de d'évaluation des fonctions de modulation. | 113 |
| 4.4 | Illustration de courbes de probabilité cumulée du coefficient E de Nash-Sutcliffe représentant la qualité des modèles. | 124 |
| 4.5 | Indices SI de sensibilité globale des différents paramètres du modèle de dynamique du C. | 127 |
| 4.6 | Distributions <i>a priori</i> et <i>a posteriori</i> des paramètres des fonctions de modulation. | 129 |
| 4.7 | Variabilité des fonctions de modulation estimées et fonctions de modulation choisies. | 130 |
| 4.8 | Distributions <i>a priori</i> et <i>a posteriori</i> des paramètres des cinétiques de décomposition pour le modèle appliqué aux sols témoins (A) et aux sols amendés (B) en fonction des différentes modulations prises en compte. | 133 |
| 4.9 | Biais relatif d'ajustement (rouge) et de prédiction (bleu) des flux non marqués par le modèle appliqué aux sols témoins (A) et aux sols amendés (B) et des flux marqués (C). | 135 |
| 4.10 | Probabilités cumulées du coefficient E d'ajustement (calibration) et de prédiction (validation) du modèle sans modulation pour les flux non marqués du modèle appliqué aux sols témoins (A) et aux sols amendés (B) et pour les flux marqués (C) pour chaque date de mesure. | 137 |
| 4.11 | Évolution relative par rapport au modèle de base de l'aire sous la courbe des courbes de probabilités cumulées du coefficient E de Nash-Sutcliffe pour l'ajustement et la prédiction. | 139 |
| 4.12 | Flux mesurés versus simulés pour le modèle appliqué aux sols témoins ou aux sols amendés. | 140 |
| 5.1 | Diagramme de la méthode générale appliquée. | 151 |
| D.1 | Propriétés du sol sélectionnées et relations estimées associées pour les flux non marqués des sols non amendés. | 192 |
| D.2 | Propriétés du sol sélectionnées et relations estimées associées pour les flux non marqués des sols amendés. | 193 |
| D.3 | Propriétés du sol sélectionnées et relations estimées associées pour les flux marqués des sols amendés. | 194 |
| D.4 | Propriétés du sol sélectionnées et relations estimées associées pour le priming effect. | 195 |

Liste des tableaux

| | | |
|-----|--|-----|
| 1.1 | Experimental methods used to create different microbial diversity levels. | 23 |
| 1.2 | Classification of models representing microbial communities according to the class of microbial community (MC) representation and the number of pools associated with the class. | 29 |
| 3.1 | Assessment of models and covariates selected for each mineralisation parameter. . | 72 |
| 3.2 | Évaluation des modèles et des covariables sélectionnées pour chaque cinétique de minéralisation. | 88 |
| 3.3 | Comparaison et interprétation des liens entre covariables et dynamique du C entre DIMIMOS et MOSAIC. | 97 |
| 3.4 | Fonctions de modulation possibles des paramètres d'un modèle de dynamique du C dans le sol. | 99 |
| 4.1 | Fonctions de modulation possibles des paramètres d'un modèle de dynamique du C dans le sol. | 108 |
| 4.2 | Ensemble des modèles calibrés et modulations correspondants des paramètres. . . | 121 |
| 4.3 | Description des paramètres potentiellement présents dans les modèles et distributions <i>a priori</i> associées. | 126 |
| 4.4 | Diagnostic de convergence des modèles utilisés pour calibrer les fonctions de modulation. | 128 |
| 5.1 | Avantages et limites ou contraintes du matériel et des méthodes utilisés. | 155 |
| 5.2 | Avantages et inconvénients des méthodes possibles. | 168 |
| A.1 | Information about soil sample locations in DIMIMOS database. | 182 |
| B.1 | Information sur les sols de MOSAIC. | 183 |
| C.1 | Comparison of quality between microbial diversity index-based models and phyla-based models. | 190 |

Introduction

Contexte : une transition agro-écologique nécessaire de l'agriculture

La deuxième moitié du 20^{ème} siècle a vu se développer une agriculture intensive permettant de répondre à la demande alimentaire d'une population croissante : entre 1960 et 2000, la production globale de céréales est passée de 1 à 2 mégatonnes permettant de nourrir une population passant de 3 à 6 milliards d'individus (Tilman *et al.*, 2002). L'agriculture intensive a grandement contribué à la modification des cycles biogéochimiques du carbone (C), de l'azote (N) et du phosphore (P). Les pertes de C du sol, principalement vers l'atmosphère sous forme de dioxyde de C (CO₂), attribuées aux activités agricoles sont estimées à 55-100 Pg (Lal, 1999) et sont en grande partie attribuables aux changements importants du mode d'usage des sols (Reidsma *et al.*, 2006). Par exemple, le passage d'une forêt à un sol cultivé et d'une prairie à un sol cultivé peuvent être responsables d'une perte jusqu'à 42% et 59% du C du sol, respectivement (Stockmann *et al.*, 2013).

D'autres pratiques agricoles telles que la fertilisation minérale, contribuant à une faible restitution du C dans le sol, et le labour, contribuant à une accélération de la minéralisation du C (Liu *et al.*, 2006) ont également contribué à diminuer la capacité de réservoir de C de la biosphère terrestre, passant de 28,2% en 1980 à 24,2% en 2000. Cette diminution est la résultante d'une dégradation et d'une désertification des sols, que la perte même de C du sol accentue (Lal, 2008). Le dernier rapport du Groupe d'experts Intergouvernemental sur l'Évolution du climat (IPCC, 2014) conclut que l'augmentation des gaz à effets de serre (GES) dans l'atmosphère, sous forme de CO₂ mais également de molécules azotées, pourrait conduire à une augmentation de la température globale terrestre de 1,4 à 5,8°C ayant des conséquences climatiques négatives sur la production agricole (Altieri *et al.*, 2015) : le maïs, le blé et autres productions majeures ont subi une réduction de leur rendement de 40 millions de tonnes par an entre 1981 et 2002

(Lobell *et al.*, 2011). La simulation de production de maïs en Afrique et en Amérique Latine sous différents scénarios climatiques prévoit une diminution de 10% d'ici 2055, entraînant une perte annuelle de 2 milliards \$ (Jones and Thornton, 2003).

Ainsi, pour relever le défi de nourrir une population exponentiellement croissante (United Nations, Department of Economic and Social Affairs, Population Division, 2015) tout en minimisant ses impacts sur l'environnement, de nouvelles formes d'agriculture engagées vers une transition que l'on peut qualifier d'agro-écologique ont émergé ces dernières années (Wezel *et al.*, 2009, 2014; Bonny, 2011; Doré *et al.*, 2011; Altieri *et al.*, 2015). Ces agricultures, aux appellations variées (Agroécologie, Agriculture Écologiquement Intensive,...), ont pour objectif de trouver les pratiques optimales à mettre en place dans le but d'assurer la production agricole nécessaire, tout en valorisant au mieux les processus biologiques et écologiques permettant une réduction des intrants et favorisant la biodiversité (Bonny, 2011; Wezel *et al.*, 2014). Ces pratiques sont variées et concernent le choix et les distributions spatiale et temporelle des variétés cultivées ainsi que les gestions de la fertilisation, de l'irrigation, des adventices, maladies et autres ravageurs.

La gestion de la dynamique de la matière organique du sol comme levier d'action

Avec 2500 Pg de C sur 1 mètre de profondeur, le sol est le troisième réservoir de C sur terre, derrière les réservoirs océanique et géologique (Lal, 2008). Au vu du rôle de l'agriculture sur la capacité du sol à alterner entre source et puits de C, et de ses conséquences sur le climat, il y a un intérêt grandissant sur le potentiel des sols agricoles à stocker du C (Chambers *et al.*, 2016). Le C du sol est majoritairement présent sous forme de matière organique (MO) dont la quantité est la résultante d'une dynamique basée sur des processus biologiques naturels complexes contrôlant le bilan entre les entrées et les sorties de MO (Figure 1).

Dans les sols agricoles, la dynamique de la MO est rythmée par l'apport de MO fraîche (MOF) sous forme de résidus de culture ou de fertilisants organiques, subissant un processus de décomposition (flèche 1) alimentant les micro-organismes en nutriments (flèche 1a) et dont une part est minéralisée (flèche 1b). Ce processus se déroulant à une échelle intra-annuelle constitue la principale source d'entrées de C dans les sols agricoles. Dans le même temps, les micro-organismes trouvent également leur source de C et autres nutriments dans la décomposition/minéralisation, beaucoup plus lente, de la matière organique du sol (MOS) plus stable (flèches 4 - 4a - 4b).

L'importance de ce processus réside dans le fait que la minéralisation d'une partie de la MOS constitue une sortie importante du C du sol sous forme de CO_2 avec la respiration des micro-organismes pour d'autres fonctions que leur croissance (flèche 5). De plus, la dynamique de la MOF dans le sol peut accélérer ou décélérer celle de la MOS par un mécanisme nommé *priming effect* (PE) (Blagodatskaya and Kuzyakov, 2008).

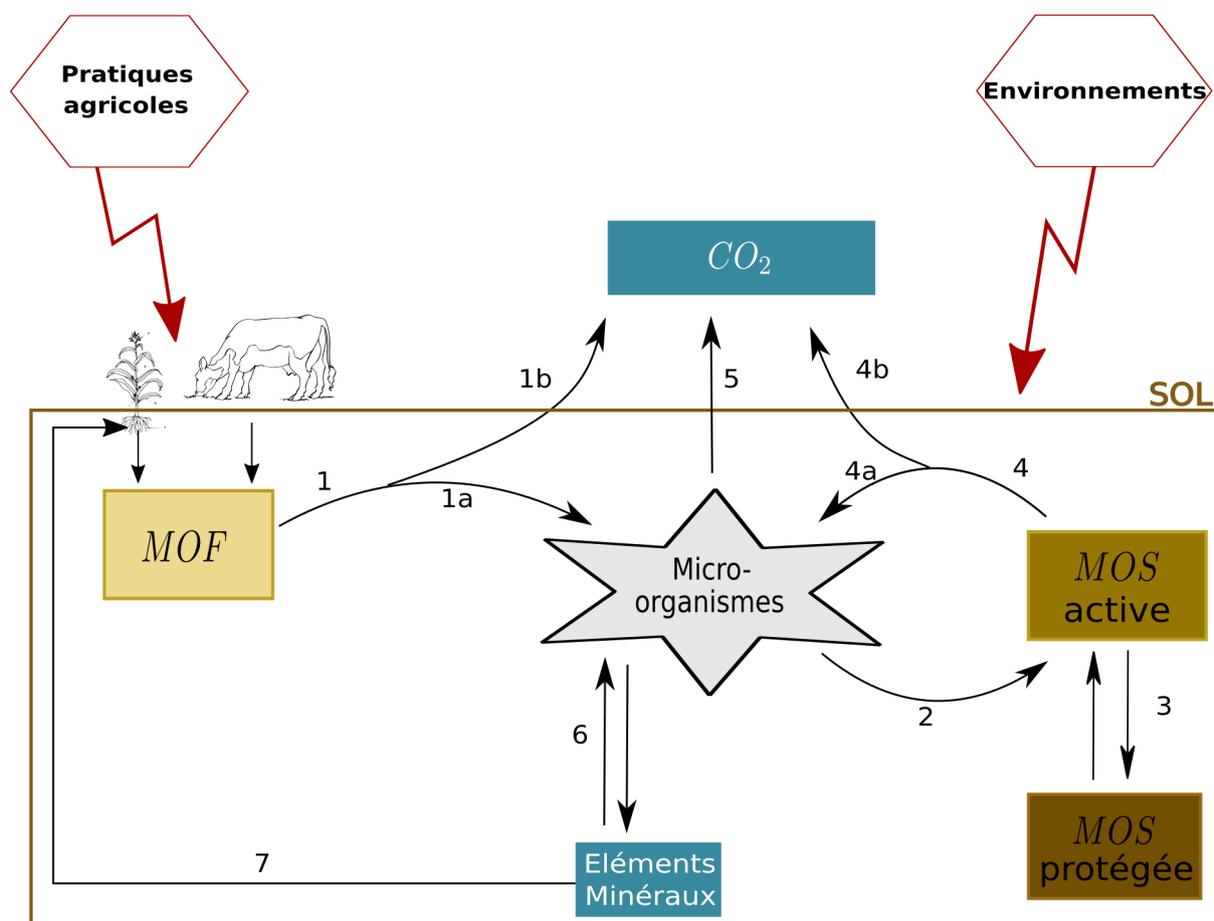


FIGURE 1 – **Dynamique simplifiée de la matière organique (MO) dans le sol.**

Une part des entrées de MO provient des résidus de culture et/ou des effluents organiques sous forme de matière organique fraîche (MOF) qui est décomposée/minéralisée par les micro-organismes du sol (1 - 1a - 1b). Les micro-organismes rejettent des composés carbonés dans le sol (enzymes, exsudats, lyse cellulaire) qui viennent alimenter la MO du sol (MOS) (2). Cette MOS peut être plus ou moins disponible (stabilisation/protection, 3) à la décomposition/minéralisation par les micro-organismes (4 - 4a - 4b). La minéralisation de la MOS constitue une sortie de C vers l'atmosphère sous forme de CO_2 . Les micro-organismes interagissent avec les mêmes éléments minéraux du sol (6) que ceux utilisés par la biomasse végétale (7).

La MOS peut également être protégée dans le sol à travers différents mécanismes de stabilisation (flèches 3) constituant ainsi un pool de MOS ayant une dynamique encore plus lente (MOS protégée). Le pool de MOS est alimenté par différents produits microbiens (e.g. enzymes, exsudats) ainsi que par la mort des micro-organismes (flèche 2). Les quantités de MOS qui transitent

dans les différents pools au cours de sa dynamique sont faibles comparées à la quantité totale de MOS. Les processus de stockage/déstockage de la MOS sont donc des processus très lents et l'étude du bilan net d'évolution de la quantité de MOS n'a réellement d'intérêt qu'à l'échelle inter-annuelle.

Au cours de la dynamique de MO, les micro-organismes interagissent avec les autres éléments minéraux (N, P) (flèches 6) sous des formes également utilisées par la biomasse végétale (flèche 7). L'intérêt pour la production agricole est donc important car la dynamique de la MO dans le sol fournit à court terme les mêmes éléments aux plantes que ceux apportés par les fertilisants minéraux azotés et phosphatés remis en cause aujourd'hui. De plus, à long terme, les éléments organiques stockés sous forme de matière organique du sol plus stable (MOS) contribuent à la qualité globale des sols (Lal, 2008).

Des facteurs pédo-climatiques contrôlent la dynamique de la MOS parmi lesquels on trouve la température et la quantité d'eau présente dans le sol (Davidson *et al.*, 1998; Kätterer *et al.*, 1998), la texture, le pH et d'autres facteurs comme la disponibilité en nutriments (Paul and Clark, 1989; Sylvia *et al.*, 2005). En outre, la dynamique de la MOS est également largement influencée par le mode d'usage des sols, ainsi que les pratiques agricoles telles que le travail du sol, la gestion de la fertilisation, la rotation des cultures (Liu *et al.*, 2007; Stockmann *et al.*, 2013). En Europe, la quantité de sols contenant moins de 3,4% de MOS (ca. 2% C organique) représentent 45% de la surface totale (Rusco *et al.*, 2001). En deçà de ce seuil, la qualité des sols et leur fonctionnement peuvent être significativement altérés.

L'intérêt de la MOS pour la production agricole ne réside pas seulement dans sa capacité à fournir les éléments favorables à la croissance des plantes mais aussi dans son rôle dans d'autres services écosystémiques rendus par les sols : régulation des flux d'eau et de l'érosion (Bakker, 1990 cité dans Liu *et al.*, 2006), filtre des nutriments (Lal, 2008). L'ensemble des effets bénéfiques de la MO sur les sols contribuent grandement à l'augmentation de la production agricole en améliorant les rendements des principales productions mondiales (e.g. 1 Mg C/ha d'augmentation par an pourrait augmenter les rendements de 20–70 kg/ha pour le blé, Lal, 2006). Plus encore, la MOS constitue un stock de C non négligeable dont l'augmentation du stockage constituerait un moyen efficace de lutter contre le réchauffement climatique tout en promouvant la production agricole (Chambers *et al.*, 2016).

Par son influence sur la production agricole, son intérêt dans les impacts environnementaux de l'agriculture et sa sensibilité aux pratiques agricoles, la dynamique de la MO constitue un facteur

clé de la transition agro-écologique des systèmes de production agricole. Une bonne gestion de la dynamique de la MO par le développement de pratiques agricoles durables passe par une bonne compréhension et prédiction de cette dynamique. En particulier, le rythme étant donné par les entrées régulières de MOF à l'échelle intra-annuelle, notamment sous forme de résidus de culture, comprendre la dynamique de décomposition de la MOF constitue une première étape nécessaire. La dynamique des entrées pouvant influencer celle des sorties (i.e. PE), l'une ne peut être évaluée sans l'autre pour avoir une idée du fonctionnement global de la dynamique de la MO d'un sol.

Les micro-organismes dans la dynamique de la MOS

Les micro-organismes du sol sont des acteurs majeurs de la dynamique de la MOS car ils contrôlent la décomposition de la MO par la production d'enzymes extra-cellulaires (Paul and Clark, 1989; Sylvia *et al.*, 2005). Les composés issus de la décomposition seront assimilés par les micro-organismes pour i) leur métabolisme de base, ii) leur croissance ou iii) la production de composés microbiens (e.g. enzymes, polysaccharides...) demandant de l'énergie dont une résultante de production est la respiration (Figure 1, flèches 1b - 4b - 5).

Dans l'étude de la dynamique de la MO du sol, les micro-organismes considérés sont souvent les bactéries et champignons du sol. Ces deux groupes d'individus sont très nombreux : on estime que dans 1 g de sol il y a environ 10^8 bactéries et 10^5 champignons, ce qui correspond à une biomasse de 300-3000 et 500-5000 kg par hectare de sol, respectivement (Sylvia *et al.*, 2005). La biomasse microbienne représente alors aux alentours de 2-4 % du C organique du sol et son intérêt dans la dynamique de la MO est grand car elle peut être un facteur limitant de la décomposition, ou co-limitant avec la quantité de substrat (Fontaine and Barot, 2005; Wutzler and Reichstein, 2008; Manzoni and Porporato, 2009).

En plus de leur grande quantité, la diversité microbienne est également très importante dans les sols. Le nombre d'espèces estimé est de plusieurs millions à la fois pour les bactéries et les champignons. Cette diversité est également difficile à mesurer (Lemanceau *et al.*, 2015) : ce n'est qu'avec les récentes avancées dans les techniques de biologie moléculaire que cette diversité a réellement pu être appréhendée (Decaens, 2010) et que le lien entre diversité microbienne et fonctionnement du sol a pu être étudié (Maron *et al.*, 2011). Bien qu'il reste beaucoup à découvrir, un nombre grandissant de résultats de la littérature montrent que la diversité des micro-organismes affecte le cycle des nutriments du sol (e.g. Baumann *et al.* 2013; Tardy *et al.*

2015; Philippot *et al.* 2013; Juarez *et al.* 2013). Tardy *et al.* (2015) ont même montré que la diversité microbienne pouvait expliquer jusqu'à 77% de la dynamique du C.

La diversité microbienne est influencée par les facteurs pédo-climatiques comme par exemple la texture, le pH (Maron *et al.*, 2011). Cette influence se fait en interaction avec les pratiques agricoles, telles que le labour, la gestion de la fertilisation et des adventices (Doran *et al.*; Feng *et al.*; Berthrong *et al.*; Zhong *et al.*; Lazcano *et al.*; Sul *et al.*), et le mode d'usage des sols (Rampelotto *et al.*, 2013). Le contexte d'agriculture allant vers plus de pratiques agro-écologiques impliquant une diversité de situations plus importantes (e.g. choix des cultures, gestion de la fertilisation) risque donc d'avoir un impact croissant sur la diversité des micro-organismes du sol. Ainsi, la diversité microbienne des sols se place à l'interface entre environnement, pratiques agricoles et dynamique de la MO dans les sols. Inclure le rôle de cette diversité dans la gestion de la dynamique de MO devrait permettre de mieux comprendre et prédire cette dynamique, en particulier dans le contexte de transition agro-écologique actuelle (Gougoulias *et al.*, 2014; Lemanceau *et al.*, 2015).

Utilité des modèles mathématiques de dynamique de la MOS

Les modèles mathématiques de dynamique de la MO constituent des outils intéressants dans la compréhension et la prédiction de cette dynamique sous l'influence des facteurs pédo-climatiques et des pratiques agricoles (Kutsch *et al.*, 2009; Manzoni and Porporato, 2009). Le nombre important de modèles simulant les dynamiques des nutriments du sol témoignent de la complexité de ces dynamiques (Manzoni and Porporato, 2009). La majorité des modèles sont des modèles déterministes à compartiments, où les mécanismes de transformation de la MO sont représentés plus ou moins explicitement (modèles mécanistes). L'intérêt de ces modèles est double : d'une part, ils permettent de tester des hypothèses en formalisant les processus et en comparant les simulations du modèle aux comportements réellement observés ; d'autre part, ils permettent de prédire la dynamique de MO et de pouvoir hiérarchiser le fonctionnement de cette dynamique sous différents scénarios (scénarios climatiques, changements de pratiques agricoles). Les modèles de la deuxième catégorie peuvent être utilisés dans des outils d'aide à la décision pour orienter les pratiques à mettre en place selon des objectifs à atteindre. Pour cela, ces modèles doivent être opérationnels, i.e. pouvant être renseignés et calibrés par des données mesurables.

Bien que les modèles intègrent depuis longtemps la biomasse microbienne (Manzoni and

(Porporato, 2009) plus ou moins explicitement (Wutzler and Reichstein, 2008), la diversité microbienne est quant à elle très peu représentée malgré les récents résultats suggérant son influence. Cela s'explique par la bonne performance des modèles actuels, d'un point de vue conceptuel, sans intégrer le rôle de la diversité, mais aussi par l'absence de lien quantitatif bien démontré entre cette diversité et la dynamique de la MO dans les sols (Maron *et al.*, 2011). Cependant, en raison de sa place entre environnement, pratiques agricoles et dynamique de la MO dans les sols, l'intégration de la diversité microbienne dans les modèles de dynamique de MO pourrait aider au développement de modèles prédictifs permettant la hiérarchisation du fonctionnement des sols sous différents scénarios. Cette hypothèse reste cependant à tester.

Objectif général de la thèse

Dans ce contexte général, le travail de thèse rapporté dans ce manuscrit a pour ambition d'étudier l'intégration de la diversité microbienne dans les modèles de dynamique de la MO dans le sol. L'objectif de modélisation se concentre ici sur l'amélioration d'un modèle prédictif et opérationnel, i.e. qu'il est possible de renseigner et de calibrer avec des données mesurables, plutôt que sur l'utilisation d'un modèle mécaniste complexe. Le C étant l'élément principal de la MO, et par soucis de simplification, on s'intéresse plus spécifiquement à la dynamique du C. En particulier, nous travaillons sur la dynamique intra-annuelle du C dans le sol, lors de la décomposition d'un résidu de culture.

Pour répondre à ces objectifs, une synthèse de la littérature sur la place de la diversité des micro-organismes dans les modèles actuels de dynamique de la MO, présentée dans le premier chapitre, a permis d'identifier les leviers et limites de la prise en compte de la diversité microbienne dans les modèles et de définir une stratégie pour atteindre cet objectif. Cette stratégie ainsi que le matériel disponible sont présentés dans un deuxième chapitre. L'étape consistant en une meilleure compréhension du lien entre diversité microbienne et dynamique de la MO fait l'objet du troisième chapitre et permet d'aider à l'intégration de la diversité microbienne dans un modèle. Cette intégration, ainsi que son évaluation, est présentée dans le chapitre 4. Le cinquième et dernier chapitre sera l'occasion de discuter de l'implication des différents résultats de ce travail de thèse.

Chapitre 1

État de l'art

Ce chapitre est consacré à l'état de l'art concernant le rôle et la place de la diversité des micro-organismes dans le sol et dans les modèles de dynamique de MOS. Les deux premières sections permettent d'introduire brièvement des notions importantes dans la compréhension du travail réalisé. La première apporte une définition de la diversité microbienne et liste les manières de la décrire, en s'attardant sur leurs avantages et inconvénients vis à vis du lien entre diversité microbienne et fonctionnement des sols. On introduira également les facteurs biotiques et abiotiques qui influencent la diversité et expliquent les différences que l'on peut observer sur le terrain. Dans la deuxième section, après avoir introduit la structure générale des modèles compartimentaux de dynamique de la MOS, les différentes complexifications (i.e. ajout de compartiments) de cette structure sont présentées ainsi que leurs intérêts et les conséquences sur l'identifiabilité des paramètres de ces modèles. Ensuite, la prise en compte des propriétés des sols qui modulent la décomposition est examinée.

La troisième section (en anglais) met en lien les deux premières en faisant une synthèse de la littérature sur les modèles qui introduisent une notion de diversité microbienne. Une première sous-partie montre l'importance de la diversité dans la dynamique de la MOS, puis une seconde fait un inventaire critique de la manière dont la diversité est introduite dans les modèles. Enfin, une série de propositions est faite pour aller plus loin dans la prise en compte de cette diversité. Cette partie a fait l'objet d'une publication scientifique sous forme de synthèse bibliographique dans *Environmental Chemistry Letters* (doi: [10.1007/s10311-016-0571-5](https://doi.org/10.1007/s10311-016-0571-5)).

Dans la dernière partie, suite à l'analyse de la littérature, les objectifs spécifiques de la thèse sont présentés, ainsi que l'organisation générale du travail effectué pour répondre à ces objectifs.

1.1 La diversité des micro-organismes du sol

1.1.1 Mesures de la diversité microbienne et mise en relation avec le fonctionnement des sols

Le terme diversité est un terme générique qui désigne la variété d'entités (espèces, entités génétiques, fonctions...) présentes dans un environnement donné. La description de la diversité peut se faire à différents niveaux en lien avec les évolutions technologiques (Figure 1.1).

- Au **niveau taxonomique**, la diversité peut être décrite à différentes résolutions :
 - À l'échelle de la communauté, la mesure de la diversité se fait principalement à l'aide d'*indices de diversité* largement utilisés en écologie (Kennedy and Smith, 1995) tels que la richesse, i.e. le nombre d'entités présentes, ou des indices incorporant également l'abondance relative de ces entités, e.g. indices de Shannon (H'), l'équitabilité (J') ou l'inverse de l'indice de Simpson ($1/D$). Les valeurs de ces indices sont d'autant plus grandes que la richesse est importante et que les entités présentes ont une abondance relative similaire.
 - En augmentant la résolution, la diversité peut être caractérisée par la description de groupes présents à tous les niveaux de la classification taxonomique (du phylum à l'espèce). On parle dans ce cas de *composition* de la communauté. Associée à l'abondance relative de ces groupes, la diversité est également décrite à travers la *structure* de la communauté. *Composition* et *structure* sont des descriptions multi-dimensionnelles de la diversité dans le sens où il faut plusieurs variables pour les mesurer.
- La diversité peut également être décrite au **niveau moléculaire** en ciblant par exemple la diversité des gènes présents dans une communauté.
- Enfin, on peut également mesurer la ***diversité fonctionnelle*** d'une communauté en ciblant les traits fonctionnels présents.

Dans le sol, la diversité des microorganismes est très importante : le nombre d'espèces estimé est de plusieurs millions pour les bactéries comme pour les champignons (Decaens, 2010). La mesure de la diversité microbienne du sol a longtemps souffert de limitations techniques importantes (Lemanceau *et al.*, 2015). Par définition, les micro-organismes ne sont pas visibles à l'œil nu ce qui rend difficile l'appréciation rapide de la diversité. La structure hétérogène du sol en tant qu'habitat des micro-organismes a également longtemps limité leur accès. Pour déterminer la diversité, il fallait alors isoler les micro-organismes dans des milieux de culture appropriés. Les

nouvelles méthodes de mesures ont montré cependant qu'une infime partie des micro-organismes sont isolables dans des milieux de culture (<10%) (Smit *et al.*, 2001; Lemanceau *et al.*, 2015). Ce n'est que récemment, avec l'avancée des techniques de biologie moléculaire, que cette diversité a réellement pu être appréhendée, révélant qu'environ seulement 0,1% des espèces bactériennes et 1% des espèces des champignons sont actuellement décrites (Decaens, 2010).

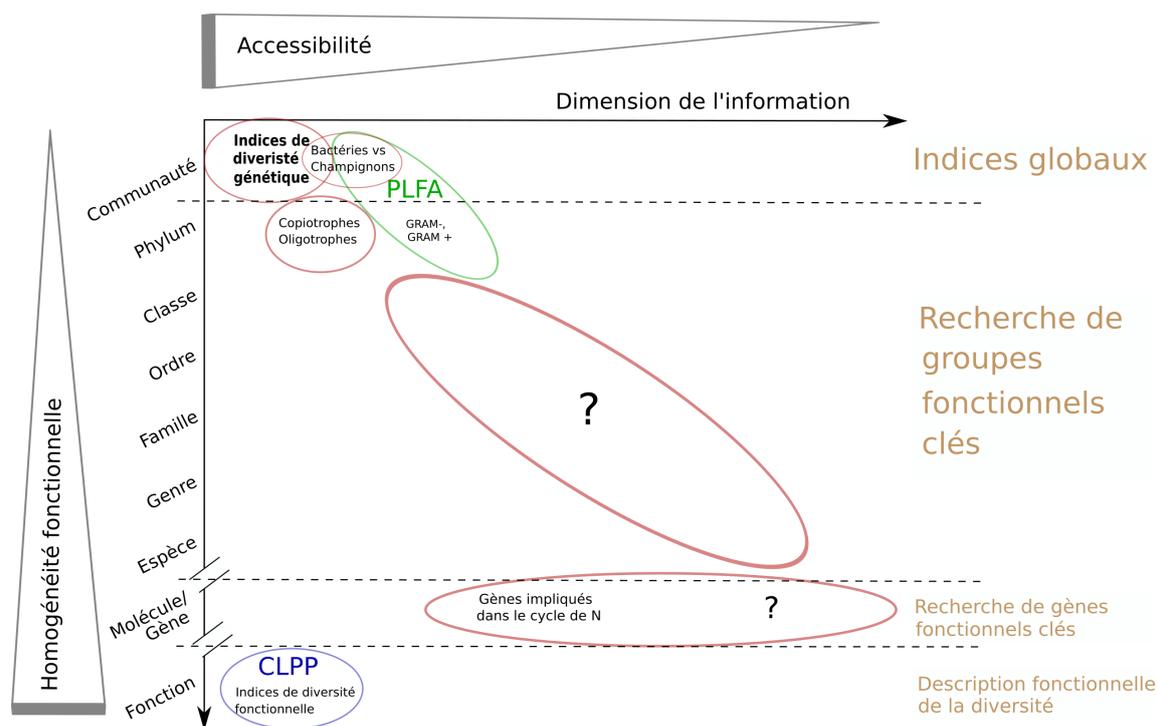


FIGURE 1.1 – Modalités de description de la diversité microbienne et possibilité de lien avec le fonctionnement des sols.

Les principales méthodes pour mesurer la diversité sont la biologie moléculaire (*rouge*), l'analyse des acides gras phospholipidiques (PLFA, *vert*) et les profils physiologiques d'utilisation de substrats (CLPP, *bleu*). L'axe des abscisses représente le nombre de variables nécessaires à l'information (échelle logarithmique), l'axe des ordonnées représente le niveau auquel est décrite la diversité. L'*Accessibilité* représente à la fois la facilité d'analyse de l'information, qui diminue avec le nombre de variables nécessaires, mais aussi le niveau de connaissances actuelles disponible. L'*Homogénéité fonctionnelle* représente la probabilité que des individus issus de groupes formés à cette échelle aient un fonctionnement similaire.

Il existe désormais deux méthodes principales pour mesurer la diversité sans avoir à isoler les micro-organismes en milieu de culture (Hill *et al.*, 2000; Bing-Ru *et al.*, 2006) :

1. L'analyse des acides phospholipidiques (PLFA) présents sur la membrane des cellules microbiennes permet de mesurer la biomasse microbienne dans le sol (Zelles, 1999). En ciblant certains acides phospholipidiques spécifiques, il est possible de connaître l'abondance relative des bactéries et champignons, mais également de certains groupes tels que les bactéries gram négatives et gram positives. Cette méthode permet une mesure de la

- diversité microbienne seulement à basse résolution (Communauté - Phylum, Figure 1.1).
2. Les méthodes de biologie moléculaire ont révolutionné les mesures de la diversité microbienne en permettant des mesures pour une gamme de résolution beaucoup plus importante (Communauté - Espèces) et à d'autres niveaux (Gènes) (Figure 1.1). Brièvement, ces méthodes consistent à prélever les acides nucléiques (ADN, ARN) dans le sol et à analyser l'ensemble de l'information génétique ou certaines séquences génétiques ciblées selon la résolution voulue (Ranjard *et al.*, 2000). Les séquences obtenues peuvent être comparées à des bases de données génétiques afin de relier la diversité taxonomique à la diversité génétique extraite. Cependant, l'inventaire taxonomique des bactéries et champignons est loin d'être complet (Decaens, 2010) et plus on augmente la résolution de description plus il y aura des séquences génétiques appartenant à des groupes non décrits, limitant de plus en plus l'accessibilité à l'information (Figure 1.1).

Le lien entre biodiversité et fonctionnement des écosystèmes a longtemps été mis évidence, en particulier entre diversité des plantes et production de biomasse végétale (Loreau *et al.*, 2001). En raison des limitations des mesures, la diversité microbienne du sol a longtemps été considérée comme fonctionnellement redondante (Nannipieri *et al.*, 2003). Aujourd'hui cette hypothèse est remise en cause et le lien entre diversité microbienne et fonctionnement du sol trouve une place de plus en plus importante dans la littérature scientifique (Andrén and Balandreau, 1999; Torsvik and Øvreås, 2002; Nannipieri *et al.*, 2003; Reed and Martiny, 2007; McGuire and Treseder, 2010; Hättenschwiler *et al.*, 2011; Maron *et al.*, 2011; Graham *et al.*, 2014).

Pour étudier ce lien, la manière la plus logique est de mesurer la diversité fonctionnelle d'une communauté (Figure 1.1). Concernant la dynamique du C, les profils physiologiques d'utilisation de substrats (CLPP, Zak *et al.*, 1994; Hill *et al.*, 2000) consistent à étudier la capacité d'une communauté à décomposer différentes sources de C par colorimétrie. L'intensité des couleurs associée à chaque substrat permet le calcul d'indices de diversité fonctionnelle. Cette méthode, bien qu'utile pour avoir une idée sommaire de la diversité fonctionnelle d'une communauté, a l'inconvénient principal d'être culture-dépendante, utilisant ainsi une partie infime de la communauté microbienne et ne reflétant pas la réalité.

Une avancée majeure de la biologie moléculaire est d'avoir rendu possible le suivi de certains gènes spécifiques impliqués dans une fonction donnée, permettant la comparaison de la présence, l'abondance et la diversité de ces gènes dans différentes communautés microbiennes (Figure 1.1). Cependant, cette méthode n'est possible que pour quelques processus, la plupart impliqués

dans le cycle de l'azote (e.g. [Zehr et al.](#); [Francis et al.](#); [Colloff et al.](#); [Hayatsu et al.](#)), grâce aux connaissances précises des gènes codant pour les enzymes responsables ([Maron et al.](#)). Pour des processus plus complexes impliquant une grande quantité de gènes fonctionnels très peu connus, comme ceux de la dynamique du C, cette méthode qui se concentre sur le suivi de gènes spécifiques est plus problématique.

Une alternative consiste à faire l'hypothèse que la diversité génétique, et donc taxonomique, détermine la diversité fonctionnelle. Faire le lien consiste donc à regrouper des taxons microbiens dans des groupes fonctionnels clés (Figure 1.1). À l'échelle de la communauté, l'information disponible consiste en des indices de diversité génétique. Sous l'hypothèse du lien entre diversité génétique et diversité fonctionnelle, il est possible de mettre en relation quantitativement ces indices avec une mesure du fonctionnement d'un système (e.g. mesure de dynamique du C). Cette méthode peut être facilement mise en place et est un bon moyen de tester le lien entre diversité et fonctionnement. Cependant, ces indices de diversité renferment une information très agrégée et il est fort probable que décrire la diversité à des résolutions plus grandes, où les groupes taxonomiques ont des chances d'être fonctionnellement plus homogènes, permettrait de mieux estimer le lien entre diversité et fonctionnement. [Fierer et al. \(2007\)](#) ont regroupé certains phyla bactériens dans deux groupes aux traits fonctionnels différents : les copiotrophes, compétitifs dans des milieux riches en nutriments, et les oligotrophes, compétitifs dans des milieux pauvres en nutriments. Bien que cette classification soit intéressante, elle concerne des groupes taxonomiques de basse résolution. Définir des groupes clés dans les rangs inférieurs semble être prometteur mais la connaissance acquise devient aussi plus rare et demande d'être augmentée avant de pouvoir le faire.

1.1.2 Facteurs influençant la diversité microbienne

Parmi les facteurs influençant la diversité microbienne, les facteurs pédo-climatiques, comme le type de sol, le pH, la texture ainsi que la teneur en eau et la température sont connus depuis longtemps (e.g. [Paul and Clark, 1989](#); [Sylvia et al., 2005](#); [Maron et al., 2011](#)). [Vries et al. \(2012\)](#) et [Gömöryová et al. \(2013\)](#) soulignent l'influence très importante de facteurs pédo-climatiques (précipitation annuelle moyenne, pH, texture, quantité de C et N) sur la variabilité de la diversité des communautés microbiennes. La dynamique et la diversité de la végétation apparaissent également comme très influentes ([Lauber et al., 2013](#)).

Le mode d'usage des sols est également un facteur prépondérant ([Maron et al., 2011](#)). Des

différences de composition microbienne sont souvent observées entre des sols cultivés et des espaces moins perturbés (e.g. espaces naturels, forêts, prairies naturelles) (Rampelotto *et al.*, 2013). Cette différence est la conséquence de pratiques agricoles hautement influentes sur la diversité microbienne : ces pratiques changent les propriétés physiques, chimiques et biologiques des sols. Des études montrent ainsi l'influence du travail du sol, du type de fertilisation, de la gestion des adventices (Doran *et al.*, 1998; Feng *et al.*, 2003; Zhong *et al.*, 2010; Berthrong *et al.*, 2013; Lazcano *et al.*, 2013; Sul *et al.*, 2013) parfois en interaction avec les facteurs pédo-climatiques (Kaisermann *et al.*, 2013; Lienhard *et al.*, 2013) sur la diversité microbienne.

Les facteurs déterminants de la diversité microbienne dépendent souvent de l'échelle spatiale considérée. À l'échelle globale, Maron *et al.* (2011) rapportent que les paramètres locaux comme le type de sol et le mode d'usage ont une importance plus forte sur la diversité bactérienne que des paramètres globaux (e.g. climat, géomorphologie). À l'échelle de la parcelle, la diversité microbienne est principalement contrôlée par la texture du sol, le type de végétation, le pH et la gestion des intrants organiques. À l'échelle microscopique, la structure et la porosité du sol ont une plus grande influence. Regan *et al.* (2014) ont étudié la distribution spatiale des micro-organismes dans le temps à l'échelle mésoscopique (10-100 cm) et ont montré l'influence des propriétés chimiques du sol. D'autres études ont montré une relation entre diversité microbienne et propriétés du sol, comme la teneur en MOS, le pH, la teneur en $CaCO_3$, la conductivité électrique et la texture pour les communautés bactériennes à l'échelle régionale (Gelsomino and Azzellino, 2011; Van Horn *et al.*, 2013) ou le pH pour les bactéries et la teneur en composés de faible poids moléculaire en général aux échelles de la parcelle et régionale (Sagova-Mareckova *et al.*, 2011; Van Horn *et al.*, 2013). Certains facteurs, comme le pH, semblent déterminants quelle que soit l'échelle : de la parcelle (Maron *et al.*, 2011) à l'échelle continentale (Fierer and Jackson, 2006).

Soumise à la dynamique de ces facteurs (e.g. variations saisonnières du climat, succession de pratiques agricoles), la diversité microbienne suit une dynamique temporelle, peu connue et peu étudiée pour le moment. Quelques études montrent que la diversité n'est pas constante au cours du temps, que ce soit en termes d'indices de diversité, de structure ou de composition (Buckley and Schmidt, 2003; Kivlin and Hawkes, 2016; Lauber *et al.*, 2013). La variabilité temporelle de la diversité ne facilite pas l'analyse du lien entre diversité microbienne et fonctionnement du sol. Si elle n'est pas prise en compte, il faut garder à l'esprit qu'il s'agit d'une simplification.

1.2 Modéliser la dynamique de la MOS

1.2.1 Structure générale des modèles de dynamique de MOS

Il existe un grand nombre de modèles de dynamique de la MO représentant principalement la dynamique de C et/ou N (Manzoni and Porporato, 2009). La grande majorité de ces modèles sont des modèles mécanistes, i.e. représentant explicitement les mécanismes fonctionnels en œuvre lors de la dynamique de C et N, et à compartiments représentant des pools de MO conceptuellement homogènes biochimiquement et/ou fonctionnellement (Halfon, 1979 cité dans Manzoni and Porporato, 2009). Malgré la grande diversité, et ciblant les modèles faisant apparaître un compartiment microbien, une structure générale de modèle simplifié peut être décrite (Manzoni and Porporato, 2009 ; Figure 1.2).

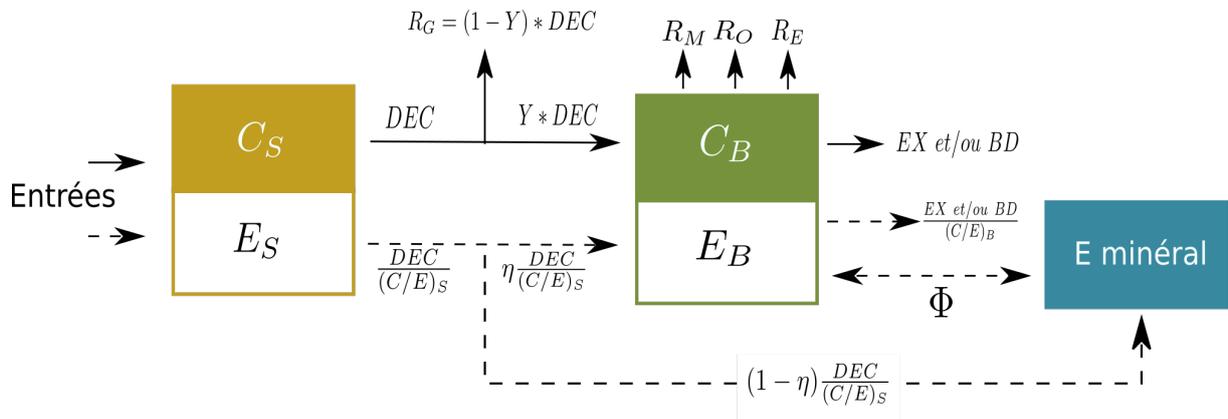


FIGURE 1.2 – Représentation schématique des flux possibles dans un modèle simplifié de dynamique de MO entre un compartiment substrat (indice S , jaune) et un compartiment microbien (indice B , vert) (d'après Manzoni and Porporato, 2009 et Manzoni *et al.*, 2012).

Les flèches pleines et pointillées représentent les flux de carbone C et d'un élément essentiel E (e.g. N, P, S), respectivement, liés par relations stœchiométriques entre la biomasse et le substrat. DEC : flux de décomposition ; EX : excrétion microbienne (e.g. enzymes, polysaccharides) ; BD : mort microbienne ; Φ : flux de minéralisation/assimilation de l'élément minéral (bleu) ; $(C/E)_i$: rapport stœchiométrique entre la teneur en C et en élément E pour le compartiment i ; R_G, R_M, R_E, R_O : respiration microbienne pour la croissance, la maintenance, l'excrétion et le surplus de C microbien, respectivement ; Y, η : paramètres de rendement d'utilisation du C et de E, respectivement.

Dans ce modèle simplifié, deux compartiments sont présents représentant le substrat (S) et la biomasse microbienne (B). Ces compartiments représentent une quantité de C et ou d'élément(s) E (i.e. N, P, S) contenus dans le pool de MO représenté à un instant t . Les modèles simulent des flux entre ces compartiments.

La décomposition du substrat par les micro-organismes est modélisée par une équation (DEC , Figure 1.2) qui représente le taux de C (et E) qui part du substrat à un instant t . Le plus souvent

cette équation s'écrit :

$$DEC = k * f(C_S) * g(C_B) \quad (1.1)$$

où k est un paramètre du modèle appelé constante de décomposition et $f(C_S)$ et $g(C_B)$ sont des fonctions de la quantité de C dans le substrat et dans la biomasse microbienne, respectivement. Généralement, la quantité de C dans le substrat est considérée comme limitante et la décomposition suit une cinétique du premier ordre. Dans ce cas, $f(C_S) = C_S$ et $g(C_B) = 1$. Il existe cependant des formes plus complexes de prise en compte linéaire et non linéaire d'une limitation par la quantité de C du substrat, par la quantité de C microbien ou d'une co-limitation de ces deux quantités. Pour plus d'informations sur les différentes équations possibles et leurs implications, se référer à Wutzler and Reinstein (2008). Lorsqu'il est présent, la quantité de l'élément E qui part du substrat est fonction de la quantité de C et du rapport stœchiométrique entre les quantités de C et de E présents dans le substrat. Une partie Y du C issu de la décomposition sera assimilée par les micro-organismes et l'autre sera minéralisée sous forme de CO_2 et respirée (R_G , Figure 1.2). Ce paramètre Y est souvent appelé rendement d'assimilation du C. Son équivalent pour l'élément E est η dans le modèle présenté ; la partie complémentaire est rejetée dans l'environnement sous forme minérale.

Des nutriments peuvent également sortir du compartiment microbien de diverses manières :

- Le métabolisme de base pour produire l'énergie nécessaire à la maintenance des cellules microbiennes vivantes. Il conduit à la consommation du C organique qui sera minéralisé et respiré (R_M , Figure 1.2).
- Les micro-organismes sécrètent des composés carbonés (EX , Figure 1.2) comme des enzymes et des polysaccharides dont la production demande également de l'énergie (R_E , Figure 1.2).
- Les micro-organismes finissent également par mourir et la lyse cellulaire libère des nutriments dans le milieu (BD , Figure 1.2).

Tous ces flux sont représentés la plupart du temps par des cinétiques proportionnelles à la quantité de C microbien.

Pour assurer le couplage entre le C et l'élément E, deux mécanismes peuvent être modélisés selon le nutriment limitant. Cette limitation est déterminée par un jeu de relations stœchiométriques entre le substrat et les micro-organismes, et impliquant le rapport de ces éléments (Figure

1.2). Si l'élément E est limitant, alors les micro-organismes peuvent i) assimiler l'élément E sous forme minérale présent dans l'environnement ($\Phi < 0$) et/ou ii) rejeter du C supplémentaire pour limiter l'excès (R_O , « *carbone overflow* »). Lorsque le C est limitant, une minéralisation supplémentaire de l'élément E est possible ($\Phi > 0$). Il est évident que tous les flux présentés ici sont très rarement, voir jamais présents en même temps dans les modèles de dynamique de MO. Cela dépend des objectifs du modèle, des données disponibles, mais aussi de l'échelle considérée. Pour plus d'informations, se référer à Manzoni and Porporato (2009).

Suite à l'explication de ce modèle simplifié, on peut définir l'efficacité microbienne d'utilisation du C (*carbon-use efficiency* ou CUE, Manzoni *et al.*, 2012) de cette dynamique. Il s'agit du rapport entre la quantité de C réellement utilisé pour la croissance microbienne et la quantité de C consommé du substrat. Ce paramètre est important car il détermine le devenir du C dans le sol : plus il est grand, plus le C a été converti efficacement en biomasse microbienne et moins le C a été rejeté dans l'atmosphère.

L'estimation de la CUE ainsi que l'identification des facteurs qui l'influencent sont de plus en plus étudiés du fait de l'importance de la compréhension du partitionnement du C par les micro-organismes dans la prédiction du cycle du C et de la gestion durable des agro-écosystèmes (Manzoni *et al.*, 2012; Sinsabaugh *et al.*, 2013; Allison, 2014; Blagodatskaya *et al.*, 2014; Bölscher *et al.*, 2016; Geyer *et al.*, 2016). Geyer *et al.* (2016) proposent un cadre conceptuel pour définir la CUE, selon l'échelle considérée, à l'aide de 3 niveaux emboîtés :

1. Ils définissent d'abord la CUE à l'échelle de population de cellules (CUE_P) comme l'efficacité d'utilisation du C limitée uniquement par les contraintes métaboliques de la population considérée. À cette échelle, on peut séparer par exemple des espèces microbiennes efficaces d'autres espèces moins efficaces de par leur métabolisme intrinsèque.
2. À l'échelle d'une communauté microbienne, d'autres facteurs peuvent influencer la CUE globale de cette communauté. La CUE_C intègre ces facteurs et peut être vue comme une moyenne pondérée des CUE_P des populations présentes dans la communauté affectée par des facteurs environnementaux.
3. Enfin, on peut définir une CUE à l'échelle de l'écosystème (CUE_E) qui inclut le recyclage des communautés microbiennes et constitue un meilleur indicateur de la stabilisation du C dans la biomasse microbienne à long terme.

Geyer *et al.* (2016) définissent ces différentes CUE à partir de la portion de C incorporé dans la cellule. Nous élargissons ici à la portion de C prélevé sur le substrat (i.e. décomposé) comme

dans Manzoni et al. (2012).

Ainsi, en reprenant les termes définis plus haut, les équations permettant de calculer la CUE_C et le CUE_E sont :

$$\begin{aligned} CUE_C &= \frac{DEC - R_G - R_E - R_M - R_O - EX}{DEC} \\ &= Y - \frac{R_E + R_M + R_O + EX}{DEC} \end{aligned} \quad (1.2)$$

$$\begin{aligned} CUE_E &= \frac{DEC - R_G - R_E - R_M - R_O - EX - BD}{DEC} \\ &= Y - \frac{R_E + R_M + R_O + EX + BD}{DEC} \end{aligned} \quad (1.3)$$

La différence entre les deux CUE vient de la prise en compte ou non de la mort des micro-organismes. Geyer et al. (2016) propose que la CUE_C constitue une prise en compte du devenir du C à une échelle de temps qui va jusqu'à la journée tandis que la CUE_E va jusqu'au mois et au-delà, et intègre de fait le recyclage de la nécromasse et des exsudats.

1.2.2 Complexification des modèles et identifiabilité des paramètres

La représentation simplifiée de la figure 1.2 peut être complexifiée pour représenter plus finement les mécanismes de dynamique de MO dans le sol (Figure 1.3). Il y a deux manières de complexifier : i) en complexifiant les équations représentant les flux (e.g. interactions entre variables, non linéarité), non abordé ici (cf. Wutzler and Reichstein, 2008; Manzoni and Porporato, 2009) et ii) en introduisant de nouveaux compartiments décrivant plus précisément les différents pools de MO que l'on trouve dans les sols.

Ces ajouts de compartiments peuvent prendre plusieurs formes - i.e. avec des flux parallèles, en série ou « feedback » (Sierra et al., 2015) - et des buts différents :

- Flux parallèles (e.g. $MOF_i \rightarrow B_j$, Figure 1.3) : souvent utilisés pour modéliser la décomposition de différents pools de MO indépendamment les uns des autres. Dans ce cas, on ajoute des compartiments pour représenter des constantes de cinétique de décomposition différentes (k et Y) selon la récalcitrance des pools de MO. C'est le cas par exemple des modèles CANTIS (Garnier et al., 2001), DAISY (Hansen et al., 1991) ou encore GDM (Moorhead and Sinsabaugh, 2006).
- Flux en série (e.g. $MOS \rightarrow MOD \rightarrow B1$, Figure 1.3) : souvent utilisés pour modéliser la décomposition d'un pool récalcitrant de MO (décomposition lente) alimentant un pool plus facilement décomposable (décomposition rapide) qui est décomposé à son tour. Cette

représentation peut-être utilisée pour introduire un pool de matière organique soluble (MOD , Figure 1.3) qui est alimenté par la décomposition de pools de matière organique du sol et dont le carbone est assimilable par les micro-organismes (e.g. Blagodatsky et al., 2010).

- Flux avec feedback (e.g. $B2 \rightleftharpoons MOS$, Figure 1.3) : souvent utilisés pour modéliser des flux entre deux compartiments qui peuvent aller dans les deux sens comme la décomposition d'un pool de matière organique du sol qui alimente un pool microbien et la mort des micro-organismes de ce dernier qui alimente le pool de matière organique (e.g. CENTURY, Parton *et al.*, 1988 ; RothC, Coleman and Jenkinson, 1996).
- L'ajout de compartiments peut également permettre de différencier des groupes fonctionnels ou taxonomiques de biomasse microbienne. Ce point est détaillé dans la section 1.3.

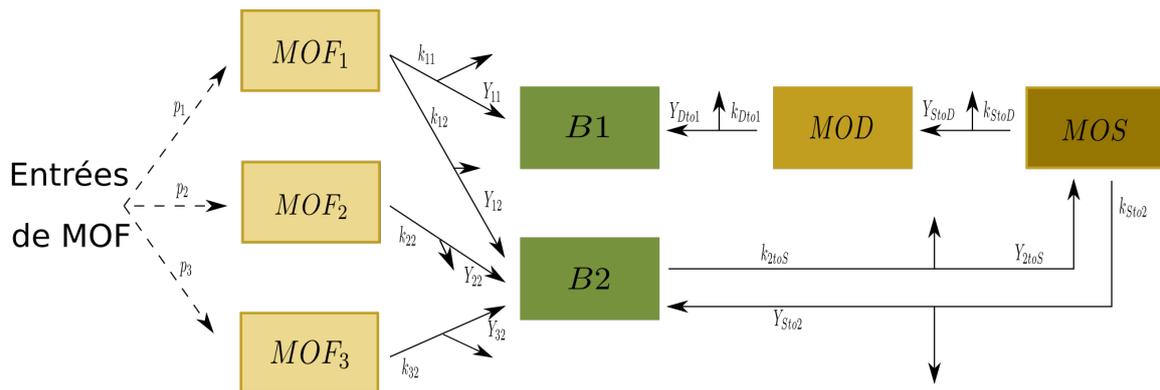


FIGURE 1.3 – Représentation schématique de différentes méthodes de complexification d'un modèle de dynamique de C dans le sol.

Les flèches pointillées représentent la répartition des entrées de MOF dans différents pools. Les flèches pleines représentent des flux de C. MOF : matière organique fraîche ; B : biomasse microbienne ; MOD : matière organique dissoute ; MOS : matière organique du sol ; k_{ij} ou k_{itoj} : constante de sortie de C du compartiment i vers le compartiment j ; Y_{ij} ou Y_{itoj} : rendement de C sortie du compartiment i assimilé par le compartiment j .

Bien que très utiles pour représenter plus finement les mécanismes impliqués dans la dynamique de MO du sol, ces complexifications ont pour conséquences une augmentation exponentielle du nombre de paramètres à estimer. Cela peut poser des problèmes d'identifiabilité des paramètres, i.e. de calibration des paramètres indépendamment les uns des autres, à moins d'avoir une quantité de données très importantes (Sierra *et al.*, 2015) ou d'émettre des hypothèses fortes sur la valeur des paramètres.

1.2.3 Prise en compte des facteurs influençant la dynamique de MOS

Les facteurs de contrôle de la dynamique de la MO dans les sols sont pris en compte comme variables de modulation des valeurs des paramètres de décomposition (paramètre k) du modèle :

$$k = k_0 \times f(T) \times g(W) \times \dots \times g(\theta) \quad (1.4)$$

où k_0 est une valeur de base du paramètre qui est modulé par des fonctions f , g , h dépendantes des facteurs de contrôle T , W , θ . Le plus souvent, ces fonctions sont comprises entre 0 et 1. Les facteurs de contrôle les plus utilisés sont la température et l'humidité (e.g. [Mueller et al., 1998a,b](#); [Garnier et al., 2001](#); [Petersen et al., 2005a,b](#); [Giacomini et al., 2007](#); [Franko et al., 2011](#)) et dans une moindre mesure la disponibilité en azote ([Garnier et al., 2001](#)), la texture (e.g. [Hansen et al., 1991](#); [Franko et al., 2011](#)), ou encore le pH ([Leifeld et al., 2008](#)).

1.3 Intégration de la diversité microbienne dans les modèles de dynamique de C et N

Cette section a fait l'objet d'une publication dans la revue *Environmental Chemistry Letters* et intitulée « *Soil C and N models that integrate microbial diversity* » ([doi:10.1007/s10311-016-0571-5](https://doi.org/10.1007/s10311-016-0571-5))

1.3.1 Introduction

Massive use of mineral fertilisers for agricultural production has altered C, N and P cycles at the global scale. Historic global C loss due to agricultural activities is estimated at 55 to 100 Pg from soil C pool ([Lal, 1999](#)). In Europe, the annual reactive nitrogen surplus of 9.9 Tg contributes to substantial N leaching to water bodies, volatilization as NH_3 and emission of N_2O and NO ([Sutton et al. 2011](#)). Since the 20th century, agriculture intensification has also been a major factor of land use change ([Reidsma et al., 2006](#)), and its impact is particularly noticeable for soil biodiversity ([Creamer et al., 2016](#)). A current major challenge for agriculture therefore is to adapt management practices to preserve ecological processes and reduce inputs and environmental impacts on soil, water and atmosphere.

Management of soil organic matter (SOM) is a key element of agro- ecological transitions because it affects both agricultural production and soil environmental functions and conversely

agricultural management affects SOM (McLauchlan, 2006) : SOM is highly involved in soil chemical quality by supplying nutrients for plant production, in soil physical quality by supporting structural stability and in soil biological quality by supporting biodiversity reserves (Reeves, 1997). In addition to having a direct impact on the soil itself, SOM has implications in pollutant transfers to air and water. It has a strong impact on the local and global C cycles and the interest in the potential of agricultural soils to sequester C is increasing (Chambers *et al.*, 2016). SOM dynamics and stocks partly control the emission and discharge of dissolved nutrients, such as nitrates, and dissolved organic C.

Dynamics of soil C and N are greatly driven by soil microorganisms (Paul and Clark, 1989; Sylvia *et al.*, 2005). New developments in molecular biology methods provided new evidence that non only microbial abundance but also microbial diversity affects C and N cycling in the soil (e.g. Baumann *et al.*, 2013; Juarez *et al.*, 2013) in different experimental and agricultural contexts. Agricultural management practices interact with pedoclimatic factors and can change soil physical properties via soil tillage, soil chemical properties via organic and inorganic inputs and soil biological properties via the type of plant cultivated for instance. All three influence microbial community diversity and differences are often observed in soil microbial abundance and composition between agricultural land and less-disturbed areas such as natural areas, forest or grassland Rampelotto *et al.*, 2013). Other studies emphasise differences in microbial community diversity for contrasting agricultural practices : soil tillage and/or fertilisation and/or weed management (Doran *et al.*, 1998; Feng *et al.*, 2003; Zhong *et al.*, 2010; Berthrong *et al.*, 2013; Lazcano *et al.*, 2013; Sul *et al.*, 2013) in interaction with pedoclimatic factors (Kaisermann *et al.*, 2013; Lienhard *et al.*, 2013). It has been shown that microbial diversity could explain up to 77% of carbon dynamic activities (Tardy *et al.*, 2015). Thus, predictive tools of C and N dynamics which integrate microbial diversity , should make it possible to fully evaluate the impacts of change in agricultural practises induced by agroecological transition.

Modelling of carbon (C) and nitrogen (N) dynamics is an essential tool to predict SOM dynamics in response to pedoclimatic and agricultural systems (Manzoni and Porporato, 2009) to adapt relevant management practises. Today C and N dynamics models are predominantly dynamic deterministic compartment models (Manzoni and Porporato, 2009), in which SOM transformation is assumed to follow first-order kinetics, and where SOM is conceptually divided into several homogeneous pools such as pool of organic matter, microbial biomass or crop residue (Halfon, 1979 cited in Manzoni and Porporato, 2009), each with its unique decomposition rate,

and more or less resistant to decomposition. Many examples of models that integrate microbial biomass are found in the literature, which confirms the interest of including these microorganisms in soil C and N dynamics models (Schimel, 2001; Neill and Gignoux, 2006). However, questions about incorporating information about microbial diversity continue to enliven the debate (e.g. Nannipieri *et al.*, 2003; McGuire and Treseder, 2010; Nielsen *et al.*, 2011; Treseder *et al.*, 2012; Graham *et al.*, 2014). The diversity of soil microorganisms, i.e. number of different species and relative contribution of each species to the community composition, is extensive, with a high level of functional redundancy in C and N transformations (Nannipieri *et al.*, 2003); consequently, it is legitimate to question whether integrating diversity parameters into models could better predict soil C and N dynamics. It has long been assumed that soil microbial diversity has no functional significance because of the high functional redundancy of soil microorganisms (Nannipieri *et al.*, 2003). This assumption was partly justified by the lack of a demonstrated diversity-function relationship, which was mainly attributed to technical limitations, which often restricted analysis of the diversity to indigenous soil microbial communities. This makes difficult to explicitly integrate the microbial community diversity in soil C and N models (Maron *et al.*, 2011). Progress in analytical methods over the two last decades, such as in molecular biology, has generated new methods and has provided the necessary tools to reconsider this assumption (Maron *et al.*, 2007).

In this context, the objective of this paper was to identify the various ways to further develop predictive models of C and N dynamics integrating soil microbial diversity. Focusing on the agricultural management context, we first summarise existing knowledge about the role of microbial diversity in C and N dynamics. We then analyse advantages and limits of how microorganisms have been incorporated into current models. Finally, we discuss methodological approaches towards a better integration of microbial diversity in soil C and N dynamics models.

1.3.2 Role of microbial diversity in soil carbon and nitrogen dynamics

Diversity is a generic term which can be described in different ways. In this review, we refer as taxonomic indexes, the diversity as described by synthetic indexes taking into account a number of entities such as genotypes or species, and their relative distribution in a community : Richness, Evenness, Shannon index... To refer to the identity of the present entities as well as their abundances in the community, we used the common term taxonomic composition. Eventually, diversity describing functional traits of the community is refer as functional diversity. Considering these different ways of describing microbial community diversity, different approaches have been

developed to demonstrate its role in soil carbon and nitrogen dynamics.

1.3.2.1 Influence of reduced soil microbial diversity

Several experimental studies have enabled to assess the impact of soil microbial diversity on C and N dynamics by manipulating diversity and mainly by artificially reducing soil microbial diversity (Nielsen *et al.*, 2011). These studies created gradients of microbial diversity in the soil while keeping a constant biomass, and measured one or several soil functions for each level of diversity. Three main methods exist, each with advantages and disadvantages (Table 1.1).

TABLE 1.1 – Experimental methods used to create different microbial diversity levels.

| Method | References | Principles | Advantages | Disadvantages |
|--------------|---|---|----------------------------|---|
| Construction | Deacon, 1985; Robinson <i>et al.</i> , 1993; Setälä and McLean, 2004; Wohl <i>et al.</i> , 2004; Bell <i>et al.</i> , 2005; Tiunov and Scheu, 2005 | Combinations of different microorganisms previously isolated in culture media | Taxa selection | Culture-dependency Unrealistic diversity levels |
| Destruction | Degens, 1998; Griffiths <i>et al.</i> , 2000, 2004 | Different duration of fumigation events | Realistic diversity levels | Hidden effect (Huston, 1997) : increase in soil nutrients |
| Erosion | Griffiths <i>et al.</i> , 2001, 2004; Wertz <i>et al.</i> , 2006, 2007; Baumann <i>et al.</i> , 2013; Juarez <i>et al.</i> , 2013; Philippot <i>et al.</i> , 2013 | Successive dilutions of a soil suspension | Realistic diversity levels | Hidden effect (Huston, 1997) : non-random taxa selection = elimination of rare taxa |

The “construction” method is performed by creating different microbial consortia (i.e. number of species and/or taxonomic groups) by combining microbial species after isolating them in culture media (Deacon, 1985). This helps understand the link between taxonomic and functional diversity by testing hypotheses of functioning in the presence or absence of specifically chosen microorganisms, such as species that use specific substrates, to study facilitation mechanisms among microorganisms (Wohl *et al.*, 2004). The number of taxa is generally small (<100) compared to the diversity of indigenous communities, and only culturable microorganisms can be isolated, which represent less than 10% of all existing microorganisms (Maron *et al.*, 2011). These taxa often have particular traits such as a fast growth rates, and studies using this method mainly focus on bacterial or fungal taxa, suppressing possible interactions between them (Romani *et al.*, 2006).

The other two methods, called respectively “destruction” and “erosion” methods (Table 1.1), create different levels of microbial diversity by (i) eliminating different quantities of soil microorganisms through different durations of fumigation with chloroform (Jenkinson and Powlson, 1976) or by (ii) inoculating sterile soil with different dilutions of a solution containing microorganisms in suspension (Griffiths *et al.*, 2000). Both approaches provide more realistic diversity levels regarding the number of taxa and coexistence of bacteria and fungi compared to the construction method. However, they also have major disadvantages or “hidden effects” (Huston, 1997) represented by a possible increase in soil nutrients from dead cells in the destruction method, and the non-random selection of taxa due to the preferential elimination of rare ones in the erosion method. Reduction of diversity is assessed by measurement of taxonomic indexes but the effect on taxonomic and functional composition is not controlled.

Most construction method results find a positive relationship between microbial diversity and soil processes involved in soil C and N cycles when there are fewer than 10 species (Deacon, 1985; Robinson *et al.*, 1993; Wohl *et al.*, 2004), but they also reveal functional redundancy (Setälä and McLean, 2004; Bell *et al.*, 2005; Tiunov and Scheu, 2005) in microbial communities when there are more than 10 species (except for Tiunov and Scheu, 2005). The unrealistically low level of diversity obtained from the construction method can result in low niche complementarity (Loreau, 2001). It is likely that certain non-cultivable microorganisms are able to perform functions that the cultivable ones cannot (Hättenschwiler *et al.*, 2011).

Genetic diversity increases functional diversity and helps microorganisms exploit a vast range of organic C compounds (Burns *et al.*, 2013). This unattained resource partitioning could lead to an apparent redundancy in studies using complex substrates (SOM or litter) as a C resource (Setälä and McLean, 2004; Bell *et al.*, 2005; Tiunov and Scheu, 2005), meaning that mineralisation stops because certain compounds cannot decompose or interactions (competition and facilitation) between microorganisms (e.g. bacteria vs. fungi) are not possible.

A hierarchy of functioning is often reported between different levels of diversity in more species-rich communities : the higher the level of diversity (i.e. higher Richness or Shannon index), the higher the process, even if differences are not always statistically significant. We focused on nine studies based on destruction and erosion methods. Four studies (Degens, 1998; Griffiths *et al.*, 2001; Wertz *et al.*, 2006, 2007) found no statistically significant differences in the measured processes between diversity levels and concluded that no relationship exists between microbial diversity and soil processes.

The results of three studies (Griffiths *et al.*, 2000; Baumann *et al.*, 2013; Juarez *et al.*, 2013) indicate a relationship between microbial diversity and soil processes such as C decomposition and denitrification, which are highly redundant functions in the microbial world (Griffiths *et al.*, 2000), and also more specialised processes such as nitrification or methane oxidation (Griffiths *et al.*, 2000). This suggests that decreased diversity initially might not affect soil processes ; however, beyond a given threshold, diversity could become a limiting factor affecting soil processes.

Another hypothesis to explain the complex relationship between diversity and soil processes is that community composition could be a determinant parameter rather than richness per se (Cox *et al.*, 2001; Griffiths *et al.*, 2004; Philippot *et al.*, 2013). In this case, both diversity and community composition could have an impact on soil processes, but the specific effect of each taxa might depend on environmental conditions and biotic interactions between microorganisms (Nielsen *et al.*, 2011) causing an increase (facilitation) or a decrease (inhibition) in functioning. Under this hypothesis, the effect of change in diversity on C and N dynamics may be more variable.

1.3.2.2 Relationship between microbial community composition and soil processes

Recent advances in molecular biology make it possible to track and analyse specific functional genes, which provides the opportunity to compare the presence, abundance and diversity of genes in different communities (Colloff *et al.*, 2008). This method is only possible for a few processes, most of them involved in the N cycle, such as denitrification or nitrification, because of the well-known genetic coding of the enzymes responsible for transforming N (Maron *et al.*, 2007). For other processes, such as C decomposition, that involve highly complex and mostly unknown sets of functional genes, an approach that focuses on functional genes is highly problematic making very difficult so far to describe functional diversity.

An alternative approach is to study the link between taxonomic and functional composition by determining whether taxonomic groups are related to the processes. Relationships between bacterial phyla and functional traits suggest that the phylum level has ecological coherence (Philippot *et al.*, 2010) meaning that members of one phylum share similar life strategy or traits that make them different from members of another phylum. Certain bacterial phyla that only grow and reproduce in nutrient-rich environments (Langer *et al.*, 2004) are classified as “copiotrophs” (Fierer *et al.*, 2007). They also are identified as “r-strategists” or opportunists because they grow rapidly after the addition of easily available C sources in the environment and are highly com-

petitive for this kind of substrate. Other phyla, grown in nutrient-poor environments (Langer *et al.*, 2004), are classified as “oligotrophs” (Fierer *et al.*, 2007). They are mainly “K-strategists” because their metabolism is slower and synthesises more complex enzymes, making possible the decomposition of recalcitrant compounds. These functional traits of phyla have been observed in different studies, such as those studying decomposition of plant residues (Pascault *et al.*, 2010) or comparing it to that of SOM (Pascault *et al.*, 2013) or in a N fertilisation gradient (Fierer *et al.*, 2012).

Even when these classifications are identified, the phylum remains a high taxonomic rank, meaning that members of one phylum can exhibit different physiological traits. Overall, the literature shows the importance of community composition in soil processes. However, knowledge remains limited, and further study is required to understand the influence of microbial composition on C and N dynamics and maybe identify key taxa as indicators of microbial community functional traits (Nannipieri *et al.*, 2003).

The overall view of the cited references bring evidence that soil microbial diversity does not have a neutral effect on the soil processes involved in C and N dynamics. But we are still lacking a full understanding of the underlying changes in soil community functioning with shifts in soil diversity (Nielsen *et al.*, 2011; de Graaff *et al.*, 2015). It seems that a change in taxonomic diversity could lead to a change in the functional traits of the whole community and thus, to a change in soil functioning.

To summarise, we propose a conceptual model (Figure 1.4) that assumes an increase in soil functioning with increased diversity. At low levels of diversity, the redundant relationship found in construction method studies could be an artefact of the method. However, low levels of diversity have been shown to develop under real conditions when a major perturbation occurs, such as fire or severe pollution, and are often associated with a reduction in soil C and N transformations (e.g. Griffiths *et al.*, 2000, 2001, 2004). It is likely that at low microbial diversity soil processes reach a minimum, which lies a certain distance from a potential maximum (Fig 1.4, section A), but grows rapidly with an increase in diversity (Figure 1.4, section B). Conversely, at high levels of diversity, we suggest that a threshold exists at which processes are not affected by an increase in diversity (Figure 1.4, section C) because functional redundancy is reached or because diversity is no longer the determinant factor. This relationship is similar to the well-known mass-ratio hypotheses in plant community diversity (Grime, 1998). On average (Figure 1.4, solid line), an augmentation of dominant species (i.e. augmentation of diversity) controls the ecosystem

functioning. However, the rate at which this plateau is reached (Figure 1.4, section B) is controlled by the functional traits hypothesis (Díaz *et al.*, 2007) which states that ecosystem functioning depends on functional diversity determined by the composition of the community. Besides, this relationship depends on the redundancy in studied processes or whether another parameter becomes more determinant. For instance, the relationship between diversity and processes seems to be expressed more significantly when substrate availability is higher (Philippot *et al.*, 2013; Tardy *et al.*, 2014). This suggests that the general relationship could be “situation-dependent” (e.g. substrate availability-dependent). More knowledge is needed to quantify this relationship and its potential variability in different situations.

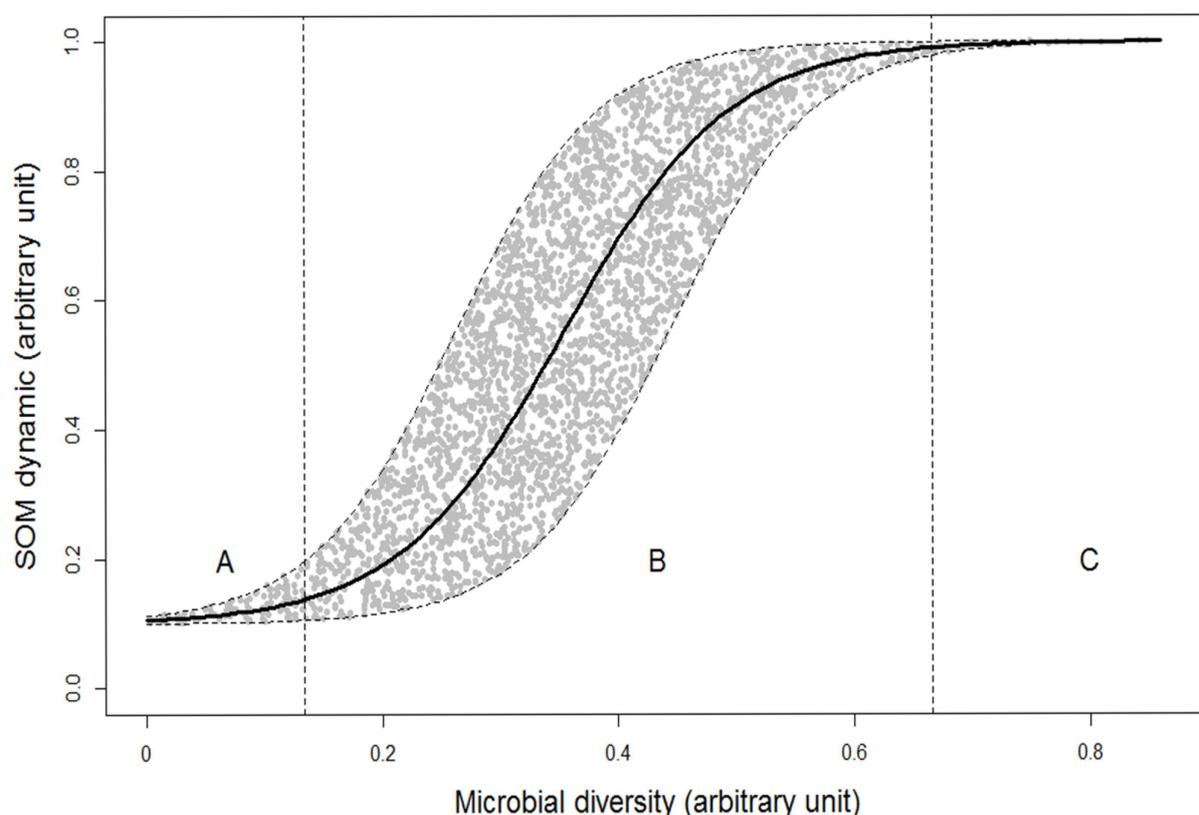


FIGURE 1.4 – **Conceptual diagram of the relationship between microbial diversity and soil organic matter (SOM) dynamics.**

Solid line = general relationship; dashed lines = area of variability in the relationship; grey points = variability of pathways in the relationship; A = area in which the relationship is minimum; B = area of potential highest variability; C = area with no relationship

1.3.3 Soil microbial diversity in current soil carbon and nitrogen dynamics models

Models are characterised by state variables corresponding to C or N compartments and kinetics parameters. According to Manzoni and Porporato (2009) reviewing approximately 250

models of C and N dynamics developed in the last 80 years, thirty percent of the models have no microbial compartment. Thirty percent represent microbial biomass as an additional organic matter compartment. Among these, few models represent explicitly microbial biomass (linearly or non-linearly) in first order kinetics of C and N dynamics (Wutzler and Reichstein, 2008). Forty percent of the models use at least two state variables to represent microorganisms, suggesting a more explicit representation of their diversity. Although models might represent the observed reality, the initialization of such models remains critical as each soil condition requires an optimization analysis to initialize the OM compartments and their decomposition rates. This section focuses on the latter model group plus five other models developed since 2009.

1.3.3.1 Representation of microbial diversity in existing C and N dynamics models

We distinguished two model categories (Table 1.2 and Figure 1.5) : models representing the functional diversity of soil microorganisms (FD), and soil food web models representing taxonomic diversity (TD). The FD models are designed to explicitly model the diversity of substrate decomposition kinetics, by representing different affinities of microbial functional pools for substrates. Defining multiple functional pools enables multiple kinetic components in the model, such as decomposition, growth, and decay rates, depending on the organic matter/microbial pool combination considered (e.g. Hansen *et al.*, 1991; Garnier *et al.*, 2001; Moorhead and Sinsabaugh, 2006; Miki *et al.*, 2010; Perveen *et al.*, 2014).

Functional pools are represented as discrete compartments in 89% of the FD models. In the guild-based model developed by Moorhead and Sinsabaugh (2006), the “opportunistic pool” represents microorganisms that compete highly for easily degradable material and have fast growth and decay rates when the resource is missing. The “decomposer pool” represents microorganisms that degrade cellulosic and lignocellulosic compounds. The “miner pool” represents microorganisms that degrade humified organic matter. In a more recent model, Perveen *et al.* (2014) distinguished two pools of microorganisms to better simulate the priming effect. The “SOM builders” pool is composed of microorganisms that degrade only fresh organic matter. The “SOM decomposers” pool represents microorganisms that degrade fresh organic matter and SOM. More detailed models with a larger number of microbial pools have been developed to represent complex interactions between functional pools or interactions between soil microorganisms and plant diversity (Loreau, 2001). Grant (2001) distinguished nine microbial pools based on nine different metabolisms differing in the source of energy and nutrients mobilised ; this representation enables

different structural and kinetic components in substrate-microbe complexes.

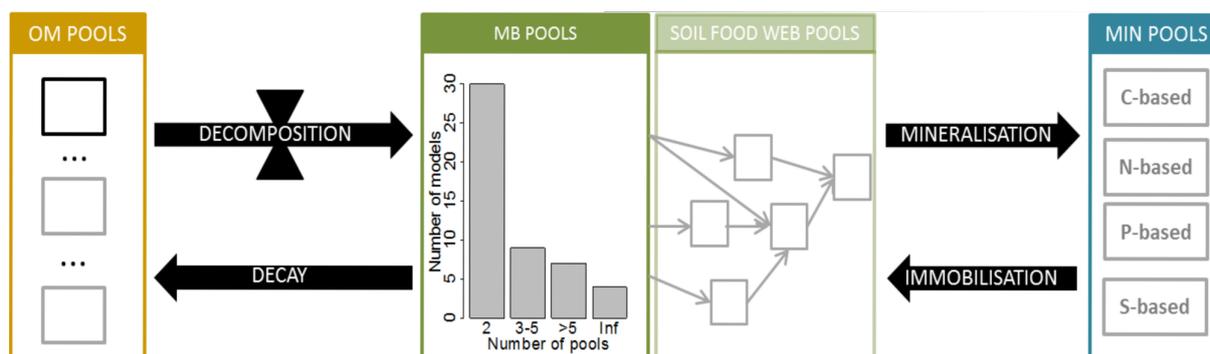


FIGURE 1.5 – **Diagram of microbial diversity in current carbon and nitrogen dynamics models.**

OM = organic matter, MB = microbial biomass, MIN = mineral compounds. Black = pool always encountered in models; grey = pool specific to certain models.

TD models (28% of those listed) distinguish taxonomic groups of microorganisms and focus on interactions along the food web. They are used to analyse contributions of these groups, specifically on N mineralisation (Hunt *et al.*, 1987; Ruiter *et al.*, 1993; Zheng *et al.*, 1999) or C and N dynamics (Hunt *et al.*, 1983; Kuijper *et al.*, 2005). The microbial groups differ in their nutrient ratios, feeding preferences, life spans, assimilation efficiencies, production-to-assimilation ratios and decomposability.

TABLE 1.2: **Classification of models representing microbial communities according to the class of microbial community (MC) representation and the number of pools associated with the class (from Manzoni and Porporato, 2009).**

References in bold were published after Manzoni and Porporato (2009). The class are FD : representation of functional diversity and TD : Soil Food Web models representing taxonomic diversity. Number of pools is equal to the infinite (Inf) for the models presenting a continuous of microbial diversity.

| Model | Reference | MC | Pools |
|---------|----------------------------|----|-------|
| PWNEE | <i>Patten, 1972</i> | TD | 6 |
| PHOENIX | <i>McGill et al., 1981</i> | FD | 2 |
| NCSOIL | <i>Molina et al., 1983</i> | FD | 2 |
| - | <i>Hunt et al., 1983</i> | TD | 4 |
| NCSOIL | <i>Hadas et al., 1987</i> | FD | 2 |
| - | <i>Hunt et al., 1987</i> | TD | 11 |
| - | <i>Leffelaar, 1986</i> | FD | 2 |

Classification of models representing microbial communities (continued)

| Model | Reference | MC | Pools |
|-----------|--|----|-------|
| - | Leffelaar and Wessel, 1988 | FD | 2 |
| - | Robinson <i>et al.</i> , 1989 | TD | 2 |
| DAISY | Hansen <i>et al.</i> , 1991 | FD | 2 |
| GEM | Hunt <i>et al.</i> , 1991 | TD | 5 |
| DNDC | Li <i>et al.</i> , 1992a,b | FD | 4 |
| - | Griffiths and Robinson, 1992 | TD | 2 |
| - | Ruiter <i>et al.</i> , 1993 | TD | 9 |
| Ecosys | Grant <i>et al.</i> , 1993 | FD | 4 |
| - | Kersebaum and Richter, 1994 | FD | 2 |
| Q-model | Bosatta and Agren, 1994 | FD | Inf |
| Q-model | Bosatta and Agren, 1995 | FD | Inf |
| Q-model | Bosatta and Agren, 1996 | FD | Inf |
| - | Zheng <i>et al.</i> , 1997 | TD | 2 |
| NCSOIL | Hadas <i>et al.</i> , 1998 | FD | 2 |
| - | Zheng <i>et al.</i> , 1999 | TD | 2 |
| - | Henriksen and Breland, 1999b | FD | 2 |
| DNDC | Li <i>et al.</i> , 2000; Stange <i>et al.</i> , 2000 | FD | 4 |
| SOILN-NO | Korsaeth <i>et al.</i> , 2001 | FD | 2 |
| CANTIS | Garnier <i>et al.</i> , 2001 | FD | 2 |
| - | Loreau, 2001 | FD | m |
| Ecosys | Grant, 2001 | FD | 9 |
| - | Kravchenko <i>et al.</i> , 2004 | FD | 2 |
| - | Moore <i>et al.</i> , 2004 | TD | 2 |
| - | Foereid and Yearsley, 2004 | TD | 2 |
| - | Long and Or, 2005 | FD | 2 |
| CN-SIM | Petersen <i>et al.</i> , 2005a,b | FD | 2 |
| INDISIM-S | Ginovart <i>et al.</i> , 2005 | FD | 2 |
| - | Kuijper <i>et al.</i> , 2005 | TD | 5 |
| - | Moore <i>et al.</i> , 2005 | TD | 10 |

Classification of models representing microbial communities (continued)

| Model | Reference | MC | Pools |
|-----------------|------------------------------------|-----------|-----------|
| EnzModel | Allison, 2005 | FD | 2 |
| - | Fontaine and Barot, 2005 | FD | 2 |
| - | Raynaud <i>et al.</i> , 2006 | FD | 2 |
| BACWAVE-WEB | Zelenev <i>et al.</i> , 2006 | TD | 5 |
| GDM | Moorhead and Sinsabaugh, 2006 | FD | 3 |
| - | Roy <i>et al.</i> , 2008 | FD | 2 |
| TOUGHREACT-N | Maggi <i>et al.</i> , 2008 | FD | 4 |
| NICA | Ingwersen <i>et al.</i> , 2008 | FD | 2 |
| CEM | d'Annunzio <i>et al.</i> , 2008 | FD | Inf |
| - | Miki <i>et al.</i> , 2010 | FD | 2 |
| DEMENT | Allison, 2012 | FD | NB |
| - | Waring <i>et al.</i>, 2013 | FD | 2 |
| SYMPHONY | Perveen <i>et al.</i>, 2014 | FD | 2 |
| - | Wieder <i>et al.</i>, 2014 | FD | 2 |

1.3.3.2 Advantages and limitations of existing modeling approaches

The models representing functional microbial diversity are mainly theoretical and based on conceptual microbial pools characterized by distinct functional traits. They represent selected groups of microorganisms with distinct functional traits, mainly to test hypotheses on the processes involved in C and N dynamics. The corresponding pools might not be explicitly identified (e.g. Hansen *et al.*, 1991) or distinguished according to known functional traits, as in the distinction between autochthonous microorganisms mainly found in environments rich in recalcitrant material and zymogenous microorganisms found in environments rich in easily decomposable material (Kersebaum and Richter, 1994; Hadas *et al.*, 1998; Garnier *et al.*, 2001).

Few models identify microbial functional pools in explicit taxonomic microbial groups. In these models, bacteria and fungi are often separate pools because their physiological differences induce contrasting C and N stoichiometries, and their relative abundance influences C and N dynamics and net primary production at the ecosystem scale (Waring *et al.*, 2013). All food web models have explicit microbial pools because they study nutrient and energy transfer between

taxonomic groups in an ecosystem. Two groups of microorganisms are found in soil food web models : decomposers, composed of bacteria and fungi, and predators of decomposers, mainly composed of protozoa and nematodes.

Almost half of the models that consider microbial diversity (FD and TD) have not been assessed with empirical datasets, i.e. data from an experiment or from field observations, at least in the reviewed articles. These heuristic models are essentially used to test hypotheses (Figure 1.6, arrows 4a and 5) about the coexistence of species in heterogeneous soil (Long and Or, 2005) or effects on C and N dynamics of species and enzyme diversities (Allison, 2012), rhizodeposition and microbial interactions (Robinson *et al.*, 1989; Griffiths and Robinson, 1992; Foereid and Yearsley, 2004), global change (Kuijper *et al.*, 2005) or anthropogenic and natural disturbances (Moore *et al.*, 2005). Some were only analysed at equilibrium to assess the model's capacity to simulate processes such as the priming effect (Perveen *et al.*, 2014), plant-soil feedback in plant communities (Miki *et al.*, 2010), plant persistence and C storage (Fontaine and Barot, 2005), or the effect of microbial diversity on ecosystems (Loreau, 2001). The main reason not to evaluate such models by comparing predicted and observed data is that they are in essence theoretical; therefore, data are not available to calibrate the parameters (e.g. Moorhead and Sinsabaugh, 2006). At most, the realism of model behaviour is assessed by comparing model outputs to general patterns from the literature or to those observed in real situations. These models are appropriate tools to assess ways to translate a hypotheses or a process into mathematical language, but are less suitable for prediction in the current state of knowledge (Figure 1.6, arrow 4b). The main exceptions are models focusing on N dynamic to whom genes implicated in the different processes are well known and where microbial biomass can be subdivided in measurable functional groups involved in these processes such as nitrifiers and denitrifiers (Lefelaar, 1986; Lefelaar and Wessel, 1988; Li *et al.*, 1992a,b, 2000; Ginovart *et al.*, 2005; Raynaud *et al.*, 2006; Maggi *et al.*, 2008).

Other models were evaluated by comparing their predictions to experimental data (Figure 1.6, arrow 3). However, a limit to these models is that initial values and parameters related to microbial communities are often assumed fixed or calibrated values. For instance, the initial biomass of each model compartment is often non-measurable. When compartments represent known taxonomic groups (e.g. bacteria vs. fungi; most compartments in TD), the initial biomass of each group is measurable, but when compartments represent conceptual functional groups, theoretical values are assumed. In the first case, compartments represent a high level of taxonomic classification, and we question whether this subdivision is always sufficient.

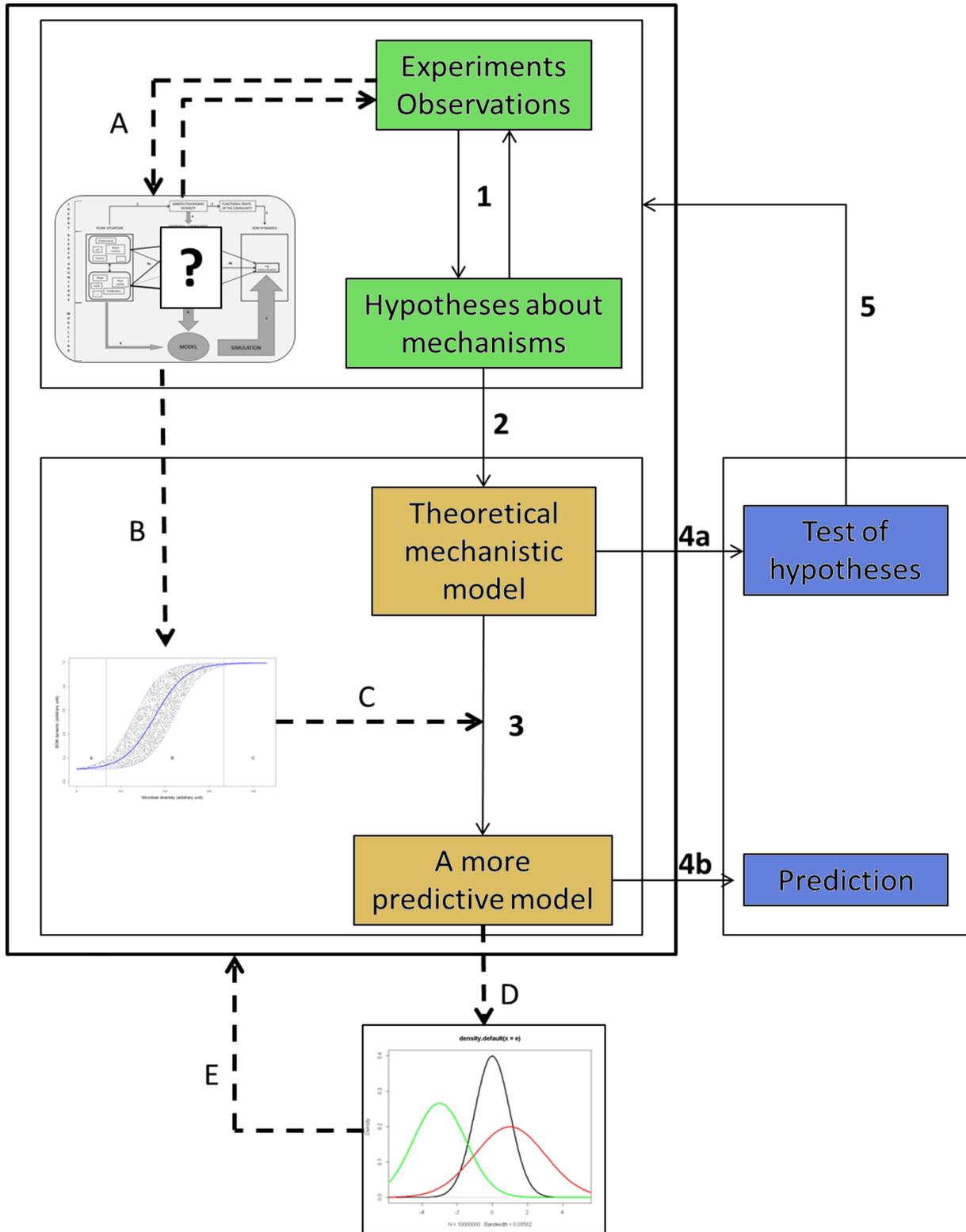


FIGURE 1.6 – Diagram of model design (solid arrows with numbers) and stages for integrating microbial community descriptors (dashed arrows with letters).

Stages correspond to (1) interactions between analysis of experimental/observed data that enable making hypotheses and hypotheses that influence future experiments, (2) translation of hypotheses into mathematical language, (3) model calibration and validation, (4a) simulations for testing hypotheses enabling (5) experiment/hypothesis interactions or (4b) for predictions, (A) search for best microbial community descriptors, (B) statistical learning, (C) coupling mechanistic modelling with statistical modelling, (D) sensitivity and uncertainty analyses, which help (E) in all model design.

This representation is often used at an ecosystem scale, but at smaller scales it may be relevant for better explaining bacterial diversity. In the second case, initial values are most often defined from the literature (Leffelaar and Wessel, 1988; Li *et al.*, 1992b; Henriksen and Breland, 1999b; Korsaeht *et al.*, 2001; Foereid and Yearsley, 2004; Ginovart *et al.*, 2005), but they can also be fitted from a control incubation (Hadas *et al.*, 1998) or estimated during model calibration (Maggi *et al.*, 2008). The main issue of using assumed values compared to real measures is that making correct assumptions is not guaranteed, and the values may not be accurate. Yet, this information can be of great value. For instance, sensitivity analysis showed that the fraction of the late-stage decomposer pool in initial microbial biomass at time $t=0$ in the model of Ingwersen *et al.* (2008) has the highest sensitivity among all fitted parameters. However, authors rarely discussed this aspect, and the way initial values are chosen is not always documented.

Efforts have been made to more explicitly represent microbial communities in C and N dynamics models, but the overview of the existing models shows that they multiply theoretical or conceptual functional compartments of microorganisms. This approach comes with difficulties in getting data to evaluate the models, creating a lack of predictive models.

1.3.4 What way towards predictive modelling approaches including soil microbial diversity ?

1.3.4.1 Getting additionnal datasets to inform and improve explicit models

As mentioned above, a lot of data are available on C and N dynamics and microbial diversity in a lot of context (different soil types, studied functions, used substrates, . . .). Metanalyses could enable to summarize these results and find general relationships between microbial diversity and C and N dynamics. However, considering the differences in the protocols measuring soil microbial diversity, deriving generic quantitative relationships between soil microbial diversity and C, N dynamics suitable to calibrate models, remains challenging (de Graaff *et al.*, 2015). Consequently additional datasets are still necessary (1) to further deepen our knowledge of soil microbial diversity impact on C and N dynamics, (2) to improve and evaluate the current models by testing with numerous quantitative diversity - C,N dynamics relationship, (3) and eventually provide more predictive modelling approaches.

To develop models with a relatively wide validity domain, these datasets must include data on soil microbial communities covering a large range of diversity and presenting enough intermediate levels to capture the likely non-linearity of the relationship (Figure 1.4). The data must

be representative of a range of pedo-climate conditions and agricultural systems. As reviewed in section 2, the aim of experimental studies creating a gradient of diversity is to detect potential difference of functioning when the diversity decreases. Consequently, they work with few levels of diversity (around 5) and seldom if ever build a relationship between diversity descriptors and soil functions. We suggest that the number of levels should be doubled (around 10) and range between a high rate of decrease and a non-decrease of the diversity. In addition, in these incubation experiments determination of microbial diversity should not be restricted to a unique initial point in time (as often done), but should be assessed parallelly to functions measurements since both are known to be strongly dynamic variables.

As microbial diversity is highly impacted by soil physical and chemical properties, climate and land use (Maron *et al.*, 2011), the natural gradients of diversity found in soils could be used to study the relationship between diversity and C and N dynamics (Tardy *et al.*, 2015). The major limit is the unavoidable presence of confusion between factors influencing C and N dynamics. Being capable of determine if different dynamics is a direct consequence of the gradient or an indirect consequence of the reason of the gradient is not that obvious. For instance, Rousk *et al.* (2010) found a strong link between a pH gradient and microbial diversity. In this case, it would be difficult to separate the biological effect (microbial diversity) from the chemical effect (pH) on C and N dynamics. However, using natural gradient, with uncontrolled factors, remains complementary to created gradient where others factors can be controlled.

In most of the available datasets based on incubation under controlled conditions to link soil microbial diversity to C and N dynamics, only the total mineralisation and decomposition fluxes (e.g. total CO₂ emissions) are measured. We argue that the use of isotopic labelling technics and the implementation of experiments, in which soil samples will be incubated with or without added labelled plant residues, are of high interest to better understand the processes driven by microbial diversity and quantify the relationship between microbial diversity and each C and N flux (decomposition, basal respiration, priming effect). These technics combined with microbial diversity characterisation are still used in ongoing research projects (ANR DIMIMOS ANR-08-STRA-06, and ANR MOSAIC ANR-12-AGRO-0005) and should provide key datasets to further develop and parameterize mechanistic models.

1.3.4.2 A complementary modeling approach based on taxonomic indexes

Although studies about the relationship between microbial diversity and SOM dynamics compute synthetic taxonomic diversity indexes (e.g. richness, Shannon index...), none of the reviewed models use those available indexes. This can be explained by the lack of understanding of the relationship between taxonomic diversity and functional traits of a community (van der Heijden and Wagg, 2013). Given the potential and availability of taxonomic indexes, additional effort is needed to better understand this relationship and assess the relevance of using them in predictive models.

As shown above, considering microbial diversity in C and N dynamics models significantly increases their complexity and consequently almost restricts their applicability to theoretical issues. While adding compartments helps represent functional diversity, it is limited by the exponential growth in the number of parameters introduced and the quantity of data needed to calibrate them. The resulting models with many compartments would be conceptually interesting but not highly applicable because it would be challenging to calibrate them. For prediction purposes, alternative modelling approaches based on empirical functions could be tested to account for impacts of microbial diversity in C and N dynamics models, even non-explicitly.

One possible alternative approach would consider soil microbial diversity descriptors in a function that modifies parameters in the first-order kinetics equation describing SOM decomposition/mineralisation : for example, simple representation of decomposition of a substrate by microbial biomass (Figure 1.7). Microbial biomass would be represented by a state variable and its activity by two parameters : decomposition rate (k) and efficiency yield (Y) (Figure 1.7).

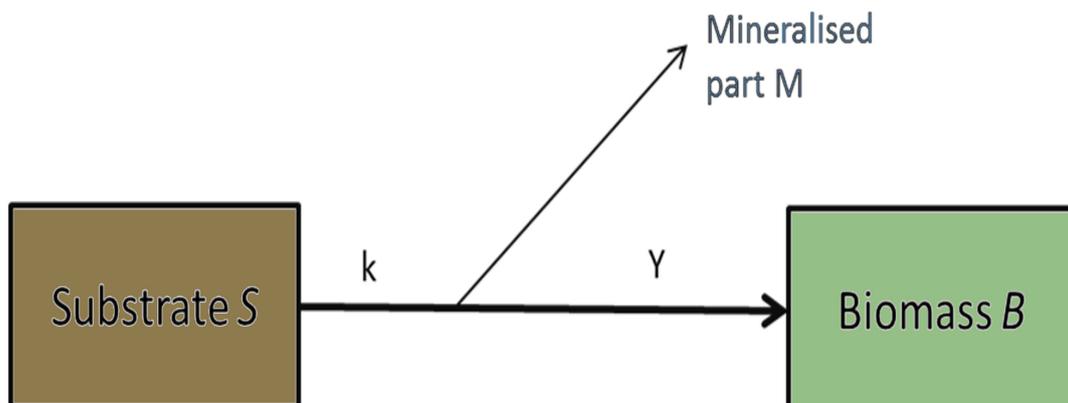


FIGURE 1.7 – Diagram of a simple model of decomposition of a substrate.

The decomposition rate and efficiency yield can be corrected by functions f and g , respectively, depending on groups θ and θ' of key taxonomic indexes, respectively :

$$\begin{array}{ccc} & f(\theta) & \\ k & \longrightarrow & k_{corr} \\ & & \\ & g(\theta') & \\ Y & \longrightarrow & Y_{corr} \end{array}$$

This approach is similar to the traditional method of including the effect of pedoclimatic factors such as temperature, soil moisture or soil texture. Parameters are a function of pedoclimatic factors mainly in a multiplicative form (e.g. Hansen *et al.*, 1991; Garnier *et al.*, 2001). This illustrates why soil microbial communities, represented by the groups of taxonomic indexes, should be quantitatively linked to SOM dynamics, represented by decomposition rate and efficiency yield (Figure 1.6, arrow B).

Since many methods measure descriptors of microbial communities and knowledge about how the latter might influence these parameters is lacking, such empirical functions are difficult to define. The need exists for better integration of empirical knowledge through statistical learning in models (Figure 1.6, arrow C). In particular, statistical tools have three advantages :

1. Many exploratory tools exist (e.g clustering, tree-based method, factorial analyses) to help identify descriptors of microbial communities that control processes of soil C and N dynamics and that are faster than mechanistic modelling and hypothesis testing.
2. Statistical modelling can be useful to link model parameters such as the decomposition rate to microbial community descriptors. Considering current limited knowledge, tools with little a priori (e.g. generalised additive models) are useful to understand complex systems in ecology and help in exploring underlying mechanisms (Guisan *et al.*, 2002).
3. Statistical modelling comes with estimated uncertainties, which could be incorporated into traditional models by developing stochastic models, which have demonstrated their relevance for ecological modelling (Clark, 2007; Katul *et al.*, 2007).

1.3.4.3 Uncertainties and sensitivity : assessing for progressing

Assessing the accuracy of predictions should include comparing model predictions to independent observed data. The relevance of methods used to represent microbial communities more explicitly should be assessed with sensitivity and uncertainty analysis tools (Figure 1.6, arrow

D), which are not used sufficiently in models of C and N dynamics. They have demonstrated their ability to assess the behaviour of complex dynamic system models, such as those simulating ground temperature (Gubler *et al.*, 2013) and soil moisture (Wolf *et al.*, 2008) or in some C and N dynamics models (Hunt *et al.*, 1987; Raynaud *et al.*, 2006; Roy *et al.*, 2008). Sensitivity analysis can improve models and adapt microbial descriptors to include different scales of study (from micro- to global scales), mechanisms or systems (Figure 1.6, arrow E). These tools could be a way to improve the generality of a model (sensu Levins, 1966) by adapting or modifying it according to the situation rather than developing a model applicable to all situations. By combining these tools, experiments would help develop models, and sensitivity analysis would help design experiments.

1.3.5 Conclusion

From our review, the increasing use of advanced tools from microbiology and molecular biology appears to provide evidence of the influence of microbial diversity in SOM dynamics. This diversity is driven by the combination of edaphic factors and agricultural management practices, suggesting that better understanding of its effects on SOM dynamics and its integration into models could increase the accuracy of the latter. Our review shows that a body of modelling studies already confirms that considering microbial diversity in C and N dynamics models should improve the ability to accurately predict SOM dynamics.

Microbial diversity is mainly incorporated by adding new biomass compartments to represent groups of microorganisms. These studies allowed the development of mechanistic models that explicitly represent soil processes. They are needed to test new scientific hypothesis and help understand underlying mechanisms and the impact of the environment and human activity. With certain exceptions, the mechanistic approach seems to be the only one used to model soil C and N dynamics in recent decades. However they are essentially theoretical and are not operational for prospective simulation. Combining the mechanistic approach with statistical tools could facilitate incorporation of explicit taxonomic diversity indexes and improve models that simulate real situations, which should follow a realistic quantitative approach to provide predictions conditioned by real data.

1.4 Objectifs spécifiques et organisation du travail de thèse

L'analyse de la littérature sur le lien entre diversité microbienne et dynamique du C dans les sols et son intégration dans les modèles mathématiques de dynamique de C nous a montré que :

1. La diversité microbienne a une influence non neutre sur la dynamique du C dans le sol (flèches 4b, Figure 1.8) et cette même diversité est influencée par des facteurs pédo-climatiques et des pratiques agricoles (flèches 4a, Figure 1.8). De fait, il semble pertinent d'intégrer cette diversité dans les modèles de dynamique de C (flèche B, Figure 1.8) au même titre que des facteurs plus classiques (flèche A, Figure 1.8). En particulier, le contexte de transition agro-écologique actuel, appelant pour une prise en compte plus importante des processus biologiques, renforce le besoin de modéliser le lien entre diversité et dynamique du C, tous deux largement influencés par les pratiques agricoles.
2. La diversité microbienne peut se mesurer de manières très différentes et il faut trouver un moyen pertinent de décrire cette diversité qui puisse être intégré dans un modèle. Sous l'hypothèse d'un lien entre diversité génétique et diversité fonctionnelle (partie « *THEORY* », Figure 1.8), l'état des connaissances actuelles sur le lien entre diversité génétique et dynamique du C ne nous permet pas d'utiliser la structure ou la composition des communautés microbiennes. A contrario, les indices de diversité pour les bactéries et les champignons semblent être de bons candidats. Toutefois, la pertinence de ces indices à expliquer la variabilité de dynamique de MO reste à tester et à quantifier.
3. Le modèle de dynamique de C dans le sol différencie la MOF (e.g. résidu de culture) de la MOS. L'influence de la diversité peut être dépendante de la source de C considérée. Il semble donc important de tester la pertinence des indices de diversité en fonction de la source de C.
4. De plus, MOF et MOS sont parfois subdivisées en sous-pools représentant les composés plus ou moins récalcitrants les constituants. Ces composés ayant des cinétiques de décomposition plus ou moins rapides, une influence différenciée des diversités bactérienne et fongique sur la dynamique de ces sous-pools devrait conduire à une expression différenciée dans le temps de l'influence des diversités. Il semble alors intéressant d'étudier l'aspect temporel de la pertinence des indices de diversité pour expliquer la dynamique du C.

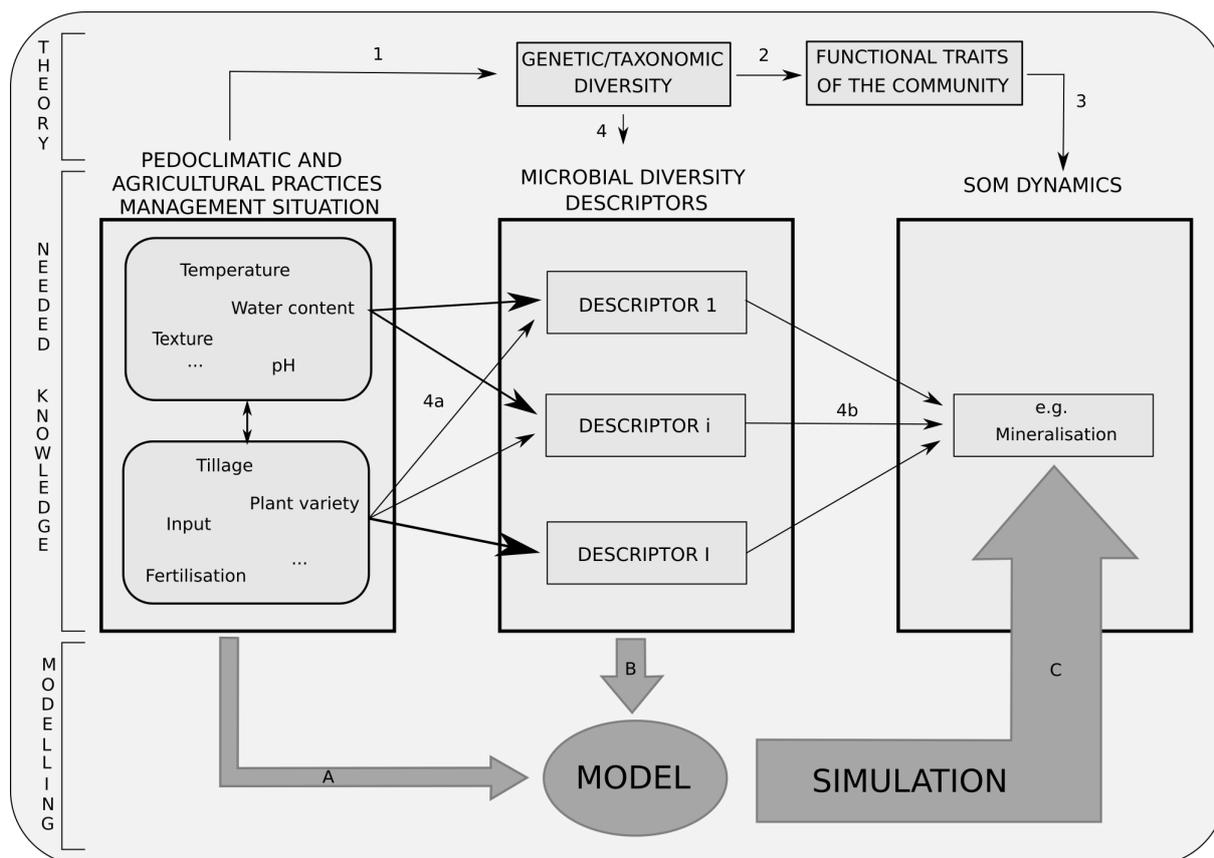


FIGURE 1.8 – Représentation de la diversité microbienne comme un intermédiaire intéressant pour comprendre le lien entre une situation pédo-climatique associée à des pratiques agricoles et la dynamique de la matière organique du sol (SOM).

La diversité microbienne pourrait alors être utilisée dans les modèles pour améliorer leur qualité prédictive. Les flèches fines peuvent être lues comme « a une influence sur ». Les facteurs pédo-climatiques et les pratiques agricoles influencent (1) les diversités génétique et taxonomique, rendant possible (2) un changement des traits fonctionnels de la communauté microbienne et ainsi (3) un changement dans la dynamique de la MOS. Il y a un besoin de trouver (4) des mesures pertinentes de la diversité microbienne qui (4b) influencent la dynamique la MOS et (4a) qui sont influencées par les facteurs pédo-climatiques et les pratiques agricoles. Ces mesures pourraient être incorporées dans (B) les modèles avec (A) des facteurs classiques pour (C) améliorer la précision des simulations de la dynamique de la MOS.

- De même, la diversité est influencée par certains facteurs, en particulier le mode d'usage, qui peuvent être plus ou moins contrastés en fonction de l'échelle d'échantillonnage de sols. En conséquence, la diversité entre sols peut être plus ou moins contrastée selon que l'on compare des sols échantillonnés à des échelles du paysage ou des échelles plus grandes. La comparaison de ces deux modalités d'échantillonnage semble donc pertinente pour déterminer si la dynamique du C est sensible aux variations de diversité microbienne que l'on peut rencontrer entre des parcelles pour ces deux modalités.
- Dans le but de pouvoir produire un modèle prédictif applicable à des jeux de données réels, à l'inverse des modèles théoriques existants actuellement et intégrant la diversité

microbienne, le modèle doit pouvoir utiliser de l'information observable (i.e. pas de compartiments représentant des pools conceptuels de micro-organismes) et être assez simple pour être calibré (i.e. nombre de paramètres limité). Dans ce but, une approche consistant à intégrer la diversité microbienne comme variable de modulation des valeurs des paramètres de décomposition d'un modèle simple semble être pertinente mais reste à confirmer.

Ces différents points nous permettent de définir trois questions de recherche qui structurent ce travail de thèse :

1. Les indices de diversité permettent-ils d'expliquer une partie de la variabilité de dynamique de C observée ?

Cette question sera traitée à l'aide de deux jeux de données d'incubations en laboratoire issus de deux projets de recherche : ANR DIMIMOS (ANR-08-STRA-06) et ANR MOSAIC (ANR-12-AGRO-0005). Nous nous intéressons en particulier à la décomposition/minéralisation du C d'un résidu de culture et de la MOS à court terme dans des sols agricoles. Les incubations ont une durée de 80 jours pendant lesquels des flux de CO₂ ont été mesurés pour caractériser la dynamique du C dans différents sols agricoles amendés avec de la paille de blé ou non. Le carbone de la paille de blé a été marqué par un isotope ¹³C ce qui permet de suivre la décomposition/minéralisation de la MOF et de la MOS indépendamment, ainsi que leur interaction (i.e. le PE). Les sols du jeu de données DIMIMOS ont été prélevés à l'échelle nationale et sont issus du Réseau de Mesure de Qualité des Sols (RMQS, [Arrouays et al., 2002](#)). Ces sols sont associés à deux modes d'usage, grandes cultures et prairies. Les sols du jeu de données MOSAIC ont été prélevés à l'échelle d'un bassin versant de 5 km² dans des exploitations de polyculture-élevage, présentant une certaine variété de systèmes de culture. Avec ce jeu de données, l'aspect temporel de la pertinence des indices de diversité sera abordé. La comparaison des résultats sur les sols issus des deux projets ANR permettra de répondre à la question de la pertinence des indices de diversité entre des sols échantillonnés à grande échelle, où la diversité microbienne a plus de chance d'être très variable, et des sols échantillonnés à l'échelle d'un paysage, pour lesquels la diversité microbienne a des chances d'être moins contrastée. Les bases de données utilisées dans cette thèse sont telles que les sols ne diffèrent pas seulement par leur diversité microbienne mais aussi par d'autres propriétés physiques et chimiques. Une description détaillée des sites et des jeux de données fera

l'objet du chapitre 2 de ce manuscrit. Dans le chapitre 3, une approche statistique innovante d'exploration de données est effectuée sur les deux jeux de données dans le but de répondre à cet objectif spécifique.

2. Quelle(s) fonction(s) de modulation dépendante(s) d'indices de diversité peut-on proposer pour moduler les paramètres de décomposition d'un modèle de dynamique du C ?

L'idée étant d'intégrer la diversité microbienne via des fonctions qui moduleraient la valeur des paramètres de cinétique du C dans un modèle dynamique, cette question concerne à la fois la forme des fonctions proposées, les indices de diversité pertinents à intégrer dans ces fonctions ainsi que les paramètres à moduler dans un modèle dynamique. La réponse à cette question sera possible grâce aux enseignements des résultats du chapitre 3. La sélection d'indices de diversité pour expliquer les flux de CO₂ mesurés selon la source du C (MOF vs MOS) permettra de savoir de quel(s) indice(s) de diversité doivent dépendre les fonctions de modulations et pour quelles cinétiques du modèle (décomposition/minéralisation de la MOF vs décomposition/minéralisation de la MOS). Les relations trouvées dans la méthode de sélection permettront de faire des hypothèses sur les paramètres que peuvent moduler les fonctions de modulation proposées, ainsi que sur la forme de ces fonctions. La réponse à cette question sera donc surtout sous forme d'hypothèse qu'il faudra tester dans une étape de modélisation mécaniste pour répondre à une troisième question :

3. L'intégration de fonctions de modulation de paramètres, dépendantes de la diversité microbienne, permet-elle d'améliorer la qualité d'un modèle de dynamique de C ?

La réponse à cette question fera l'objet du chapitre 4. Il s'agira, dans un modèle mécaniste simple, d'intégrer les fonctions de modulation des paramètres du modèle dépendantes d'indices de diversité microbienne et proposées à l'étape précédente. Ce modèle sera calibrer à partir du jeu de données MOSAIC et le gain prédictif de cette intégration sera évalué à l'aide d'indices de qualité d'ajustement et de prédiction de modèle pertinents. Les fonctions de modulation dépendantes d'indices de diversité microbienne seront intégrées en parallèle d'autres fonctions de modulation dépendantes des autres facteurs influençant la dynamique du C (e.g. pH, texture) afin de déterminer l'apport de la diversité par rapport à ces facteurs plus communément pris en compte dans les modèles. L'idée est de limiter

la complexité des modèles et de déterminer si la prise en compte des facteurs influençant la dynamique du C, en particulier la diversité microbienne, permet d'obtenir des simulations de meilleures qualités qu'un modèle ne les prenant pas en compte, i.e. suffisamment précises et permettant de capturer les différences qui existent entre des sols.

Le chapitre 5 sera l'occasion de discuter des différents résultats de cette thèse et d'envisager des perspectives à ce travail.

Chapitre 2

Matériels disponibles et méthodes générales

La première partie de ce chapitre décrit brièvement les données disponibles. Des sols ont été échantillonnés dans le cadre de deux programmes ANR. Des propriétés physico-chimiques, des données de diversité microbienne ainsi que des dynamiques du C au travers du suivi de flux de CO₂ ont été mesurées sur ces échantillons de sol. Les méthodes d'analyse des échantillons sont parfois décrites plus précisément dans les chapitres suivants. La deuxième partie décrit les méthodes générales utilisées dans les chapitres 3 et 4 pour répondre aux objectifs spécifiques de ce travail de thèse.

2.1 Les données disponibles

2.1.1 Deux projets ANR comme source des données

2.1.1.1 Le projet DIMIMOS

Le projet DIMIMOS (ANR-08-STRA-06) est un projet financé par l'ANR entre 2009 et 2013, et coordonné par l'UMR Agroécologie (P.A. Maron), dont l'objectif était d'évaluer le rôle de la diversité microbienne dans le devenir des MO du sol afin de mieux gérer le cycle biogéochimique du C au sein des agro-écosystèmes. Parmi les expérimentations mises en place dans ce projet, l'une d'elles consistait à évaluer l'aptitude des communautés microbiennes à décomposer la MO en fonction de leur diversité dans une sélection de sols représentatifs de trois modes d'usage (grande culture/prairie/forêt), afin de hiérarchiser les paramètres de régulation des flux de C

dans le sol (diversité microbienne/caractéristiques physicochimiques des sols/mode d'usage...).

Dans cette thèse, seuls les sols sous grandes cultures et sous prairies sont retenus. Les cultures correspondent à des monocultures ou à des rotations pouvant inclure des prairies temporaires de courte durée. Les prairies correspondent à des prairies permanentes ou temporaires de longue durée (>10 ans). Ce jeu de données comprend 10 sols sous cultures et 10 sols sous prairies, échantillonnés sur le territoire métropolitain français (Figure 2.1) au sein du Réseau de Mesure de la Qualité des Sols (RMQS, [Arrouays et al., 2002](#)). Une particularité non négligeable de cette base de données est que les sols ont été stockés sur une période plus ou moins longue selon un protocole précis ([Ratié et al., 2010](#)) avant d'être utilisés pour le projet. Plus d'informations sur les sols étudiés sont disponibles en annexe A.

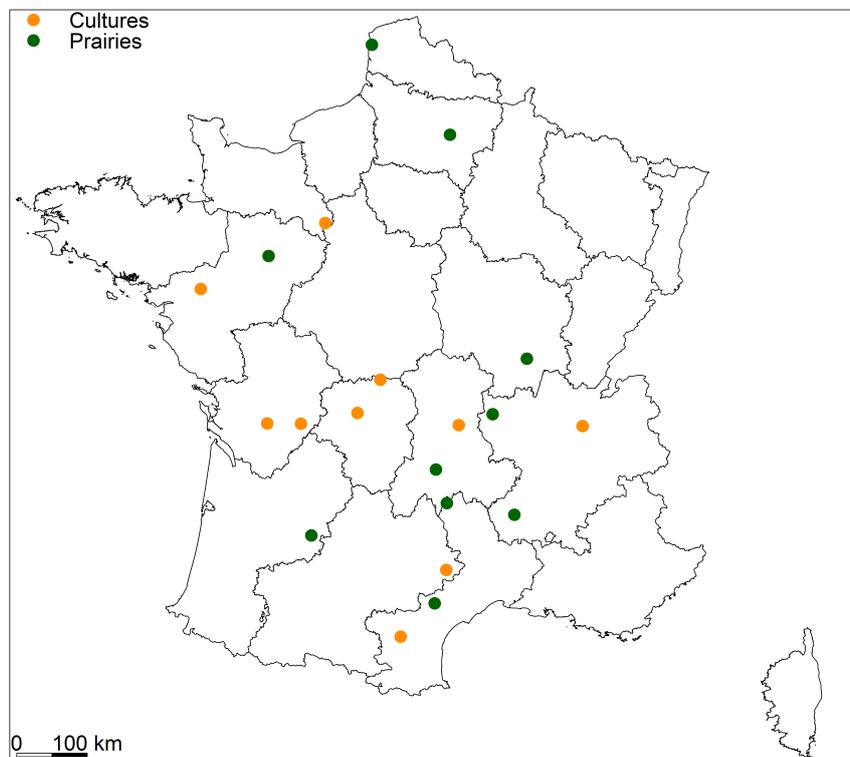


FIGURE 2.1 – Localisation des sols échantillonnés issus du RMQS en France Métropolitaine.

2.1.1.2 Le projet MOSAIC

L'objectif du projet MOSAIC (ANR-12-AGRO-0005), débuté en 2012, est d'étudier la dynamique des MOS (stocks et flux associés) dans un paysage d'agriculture intensive dominée par l'élevage, et les impacts des changements globaux (climat, évolution des systèmes de culture) sur la résilience de la MOS. Le projet s'attache à quantifier et à comprendre l'hétérogénéité des pro-

cessus mis en jeu dans la dynamique de la MOS à l'échelle d'un paysage présentant un gradient de régimes hydriques et de systèmes de culture contrastés en termes de restitutions de MOF dans le sol. En particulier, l'axe 2 vise à mieux comprendre les processus qui contrôlent la dynamique de la MOS dans un gradient de systèmes de culture et de conditions environnementales. C'est dans ce cadre que des sols ont été échantillonnés sur le site de Kervidy-Naizin de l'ORE AgrHys (http://www6.inra.fr/ore_agrhys), bassin versant de 5 km² situé dans le département du Morbihan (56) en Bretagne (Figure 2.2). Après élimination de sols particuliers (e.g. forêts, tourbes), 81 sols sont gardés pour cette thèse, tous sous systèmes cultivés et soumis à des rotations de cultures plus ou moins complexes, comprenant ou pas des prairies de plus ou moins longue durée et une variabilité de gestion de la fertilisation. Plus d'informations sur les sols étudiés sont disponibles en annexe B.

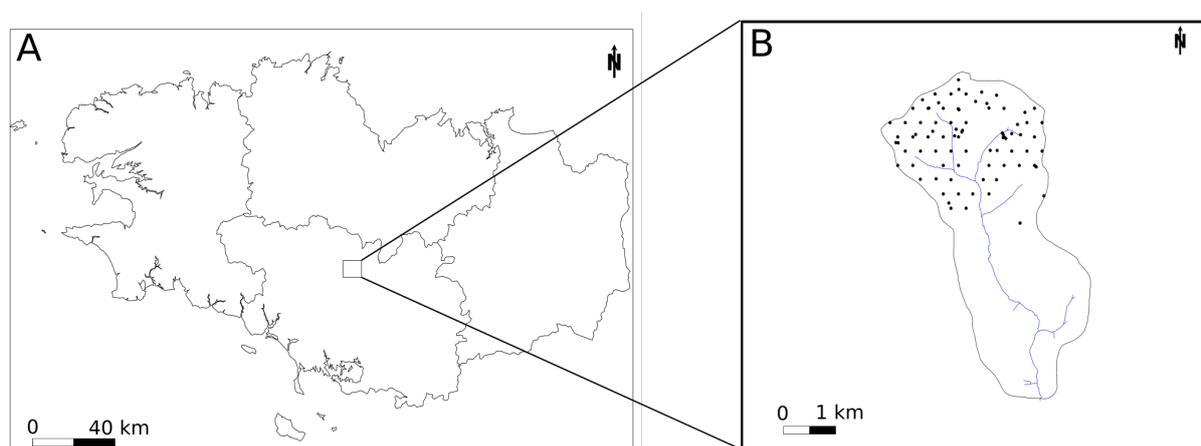


FIGURE 2.2 – Localisation du bassin versant de Kervidy-Naizin (A) et des sols échantillonnés (B).

2.1.2 Propriétés des sols échantillonnés

2.1.2.1 Analyses des échantillons de sol

Aux exceptions près, les mêmes propriétés physiques, chimiques et biologiques des sols échantillonnés ont été mesurées dans les 2 jeux de données. La différence la plus importante se situe dans la période d'échantillonnage et les conditions de stockage. Issus du RMQS, les sols DIMIMOS ont été échantillonnés entre 2003 et 2008 (annexe A), séchés et stockés (Ratié *et al.*, 2010), avant d'être analysés pour le projet DIMIMOS. Les sols MOSAIC ont été échantillonnés en 2013 sur une campagne de 3 journées consécutives. Les analyses ont été effectuées peu de temps après séchage des échantillons pour les propriétés physiques et chimiques et sur les échantillons congelés pour les propriétés biologiques (biomasse moléculaire microbienne, diversité). Cette différence de

protocole rend difficile la comparaison des deux bases, en particulier pour les données biologiques.

Propriétés physiques et chimiques

Les propriétés physiques et chimiques classiques des sols des deux bases de données ont été analysées selon des méthodes standardisées. La texture a été analysée selon 5 fractions granulométriques (NF X 31-107), le carbone organique et l'azote total par combustion sèche (NF ISO 10694) et le pH en suspension dans l'eau (rapport 1 : 5) (NF ISO 10390). Pour la base de données MOSAIC, le phosphore assimilable, mesuré par la méthode Olsen (NF ISO 11263), ainsi que l'azote minéral total, mesuré par extraction au chlorure de potassium (KCl) et spectrophotométrie en flux continu, sont également disponibles.

Propriétés biologiques

Les propriétés biologiques des sols ont été mesurées à partir de l'ADN présent dans ces sols. La procédure est détaillée dans la section 3.1 et dans Tardy *et al.* (2015). Brièvement, l'ADN du sol est extrait et la concentration en ADN est utilisée pour estimer la biomasse moléculaire microbienne totale. Les ADN bactériens et fongiques sont séparés et amplifiés à l'aide de fragment de gènes 16S rARN et 18S rARN respectivement, et par réaction en chaîne par polymérase (PCR). Après pyroséquençage, une analyse bioinformatique sur un nombre de séquences données permet de regrouper les séquences similaires à 95% en unités opérationnelles taxonomiques (OTU). Le nombre d'OTU trouvé est utilisé pour estimer les richesses en bactéries et en champignons dans le sol. Le nombre total d'OTU et l'abondance relative de chaque OTU sont utilisés pour calculer des indices de diversité bactérienne et fongique comme l'équitabilité (J'), l'indice de Shannon (H') et l'inverse de l'indice de Simpson ($1/D$) (Kennedy and Smith, 1995) :

$$J' = \frac{H'}{H'_{max}} \quad (2.1)$$

$$H' = - \sum_{i=1}^N p_i \ln(p_i) \quad (2.2)$$

$$D = \sum_{i=1}^N \frac{n_i(n_i - 1)}{N(N - 1)} \quad (2.3)$$

où H'_{max} est la valeur maximale de l'indice de Shannon, i.e. lorsque toutes les OTU ont la même abondance (i.e. $p_i = 1/N, \forall i$), N est le nombre total d'OTU (i.e. la richesse), p_i est l'abondance relative de l'OTU i (i.e. $p_i = n_i/N$) et n_i est le nombre « d'individus » appartenant à

l'OTU i . L'équitabilité J' est comprise entre 0 et 1 ; 1 représentant une communauté bactérienne ou fongique où toutes les OTU ont la même abondance relative (i.e. il n'y a pas d'OTU dominante dans la communauté). L'indice de Shannon (H') et l'inverse de l'indice de Simpson ($1/D$) prennent tous deux en compte à la fois la richesse et l'équitabilité. La principale différence entre ces indices et la prise en compte des espèces rares qui est importante dans l'indice de Shannon et très peu dans l'indice de Simpson. L'indice de Simpson (D) représente la probabilité que deux individus tirés au hasard dans la communauté appartiennent à la même OTU. La diversité est donc maximale à la valeur 0 et minimale à 1. Pour avoir un indice qui croît lorsque la diversité croît, comme pour J' et H' , l'inverse de l'indice de Simpson est souvent utilisé.

2.1.2.2 Variabilité des propriétés des sols

La figure 2.3 montre la variabilité des sols sous cultures et sous prairies pour DIMIMOS et ceux de MOSAIC (cultures seulement). Globalement, les teneurs en argile (Clay), en limon (Silt), en sable (Sand) et la somme des teneurs en argile et limon fin (Clay + Fine Silt, i.e. somme des fractions les plus fines) dans DIMIMOS couvrent une large variabilité, sans différences importantes entre cultures et prairies. Pour les sols de MOSAIC la variabilité texturale est très réduite (échantillons très localisés). Cette observation est similaire pour le pH : bien que la gamme complète de pH dans MOSAIC soit assez importante (5.5 - 7.5), la majorité des données se trouvent autour de 6, contrairement aux sols de DIMIMOS (6-8). Classiquement, les teneurs en C organique sont plus basses et moins variables dans les sols sous cultures, dont les teneurs sont comparables entre les deux bases de données, que sous prairies, et correspondent à des valeurs classiques observées pour ces modes d'usage ([Gis Sol, 2011](#)). Cette observation est inversée pour le ratio C/N. Les teneurs en P assimilable des sols MOSAIC sont plutôt dans la moitié basse des valeurs françaises en particulier dans une région comme la Bretagne où les valeurs sont généralement plus élevées que la moyenne nationale ([Gis Sol, 2011](#)).

Pour les propriétés biologiques des sols, on peut remarquer que les diversités bactérienne et fongique sont systématiquement plus élevées dans les sols de MOSAIC que dans les sols de DIMIMOS. Cette différence peut avoir deux causes :

1. Un manque de standardisation des méthodes de biologie moléculaire ([Lemanceau *et al.*, 2015](#)). Par exemple, le nombre de séquences génétiques utilisées pour regrouper les OTU n'est pas le même entre DIMIMOS et MOSAIC, ce qui a une conséquence sur les valeurs obtenues.

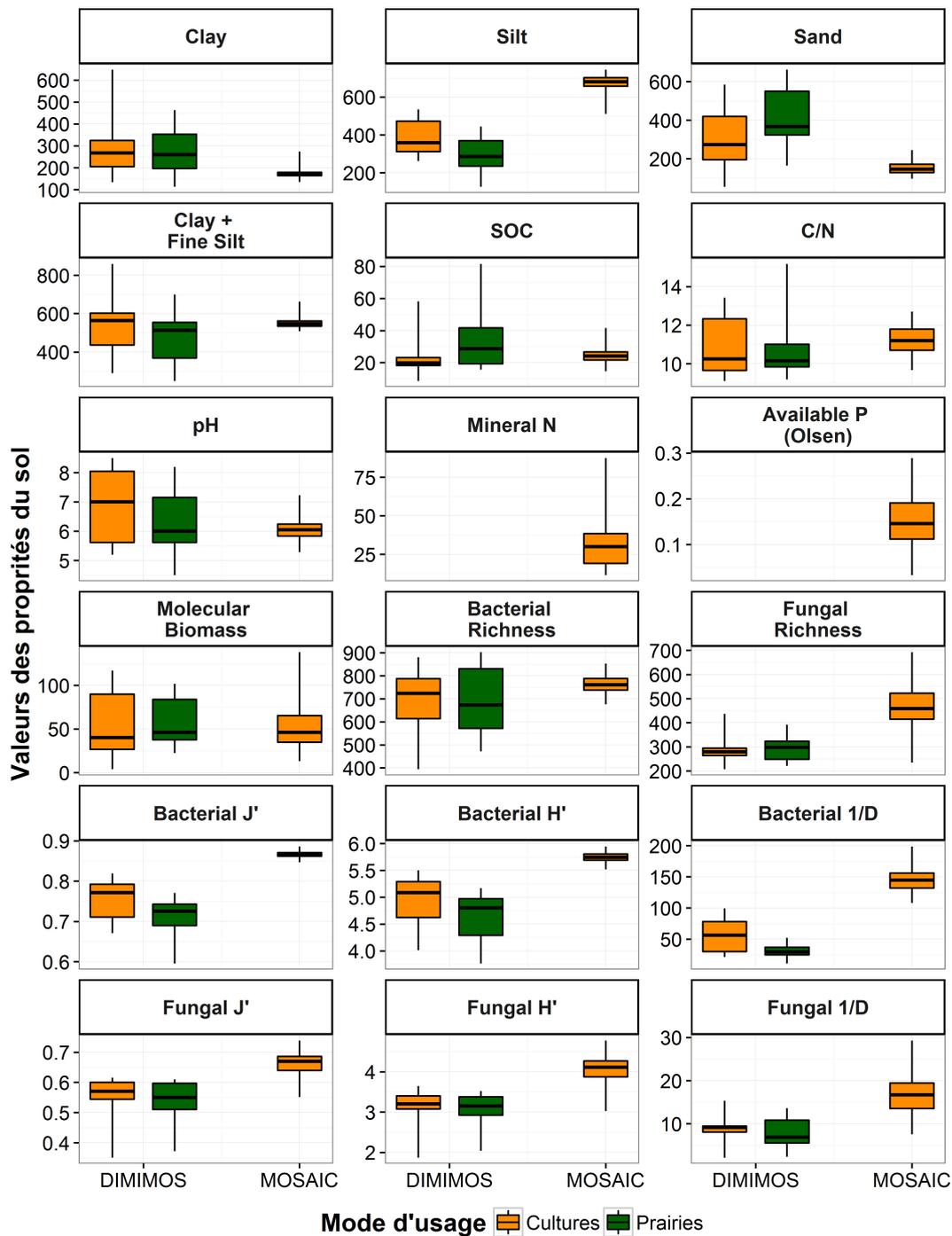


FIGURE 2.3 – Représentation de variabilité des propriétés des sols pour chacune des bases de données (DIMIMOS et MOSAIC) et chacun des modes d'usage (Cultures et Prairies).

Les teneurs en argile (*Clay*), en limon (*Silt*), en Sable (*Sand*), en argile + limon fin (*Clay + Fine Silt*), en carbone organique (*SOC*) et en phosphore assimilable (*Available P (Olsen)*) sont tous exprimées en $[g.kg^{-1}sol]$. L'azote minérale est exprimée en $[mg.L^{-1}]$ La biomasse moléculaire microbienne (*Molecular Biomass*) est exprimée en $[\mu g ADN.g^{-1}sol]$. Les richesses bactérienne et fongique (*Bacterial and Fungal Richness*) sont exprimées en $[nombre\ d'OTU]$ et le *pH* en $[unité\ pH]$. Les variables de diversité sont sans unité. Les segments haut et bas des boîtes de dispersion représentent les valeurs maximales et minimales, respectivement.

2. Une différence de stockage entre les deux bases. Le stockage des sols secs (non congelés) étant long dans DIMIMOS, la ré-humectation de ces sols a pu entraîner la croissance très rapide de micro-organismes opportunistes, i.e. qui croît rapidement dès que les conditions redeviennent favorables (stratégie r). Ces espèces de micro-organismes deviennent donc dominantes en proportion, faisant diminuer la valeur des indices de diversité.

De fait, il est difficile de comparer les valeurs de diversité des deux bases. Cependant, on observe une dispersion des valeurs différentes entre les deux bases avec une diversité bactérienne beaucoup plus variable dans DIMIMOS que dans MOSAIC. Cela peut s'expliquer par la plus grande variabilité des propriétés physiques et chimiques des sols de DIMIMOS (cf. section 1.1.2). Cette différence est moins notable pour la biomasse moléculaire et la diversité en champignons, en particulier si l'on compare les sols sous cultures des deux bases.

2.1.2.3 Corrélation entre les propriétés des sols

Les propriétés des sols n'étant pas manipulées expérimentalement mais issues d'observations, il est intéressant de regarder les corrélations entre ces propriétés afin de discuter d'éventuelles confusions d'effets entre des propriétés fortement corrélées entre elles. Les figures 2.4 et 2.5 montrent les coefficients de corrélation de Spearman (corr) ainsi que les nuages de points associés pour les propriétés des sols prises deux par deux.

Les fractions fines sont corrélées négativement avec le C/N (corr=-0.50 à -0.67). Les indices de diversité bactérienne sont très corrélés positivement entre eux (corr=0.68 à 0.98), tout comme les indices de diversité fongique entre eux (corr=0.33 à 0.97). Les indices de diversité bactérienne sont également moyennement corrélés positivement avec les indices de diversité fongique (corr=0.45 à 0.68). Les indices de diversité bactérienne sont également très corrélés positivement avec le pH (corr=0.52 à 0.77) et moyennement corrélés négativement avec les teneurs en carbone organique (SOC, corr=-0.45 à -0.50). De même, le pH est moyennement corrélé avec les indices de diversité fongique (corr=0 à 0.47).

Pour MOSAIC, les indices de diversité bactérienne sont également très corrélés positivement entre eux (corr=0.45 à 0.94) de même que les indices de diversité fongique entre eux (corr=0.63 à 0.96). Contrairement à DIMIMOS, les autres corrélations, notamment entre les indices des bactéries et des champignons et entre ces indices et le pH ou les teneurs en C, sont plus faibles voire négligeables.

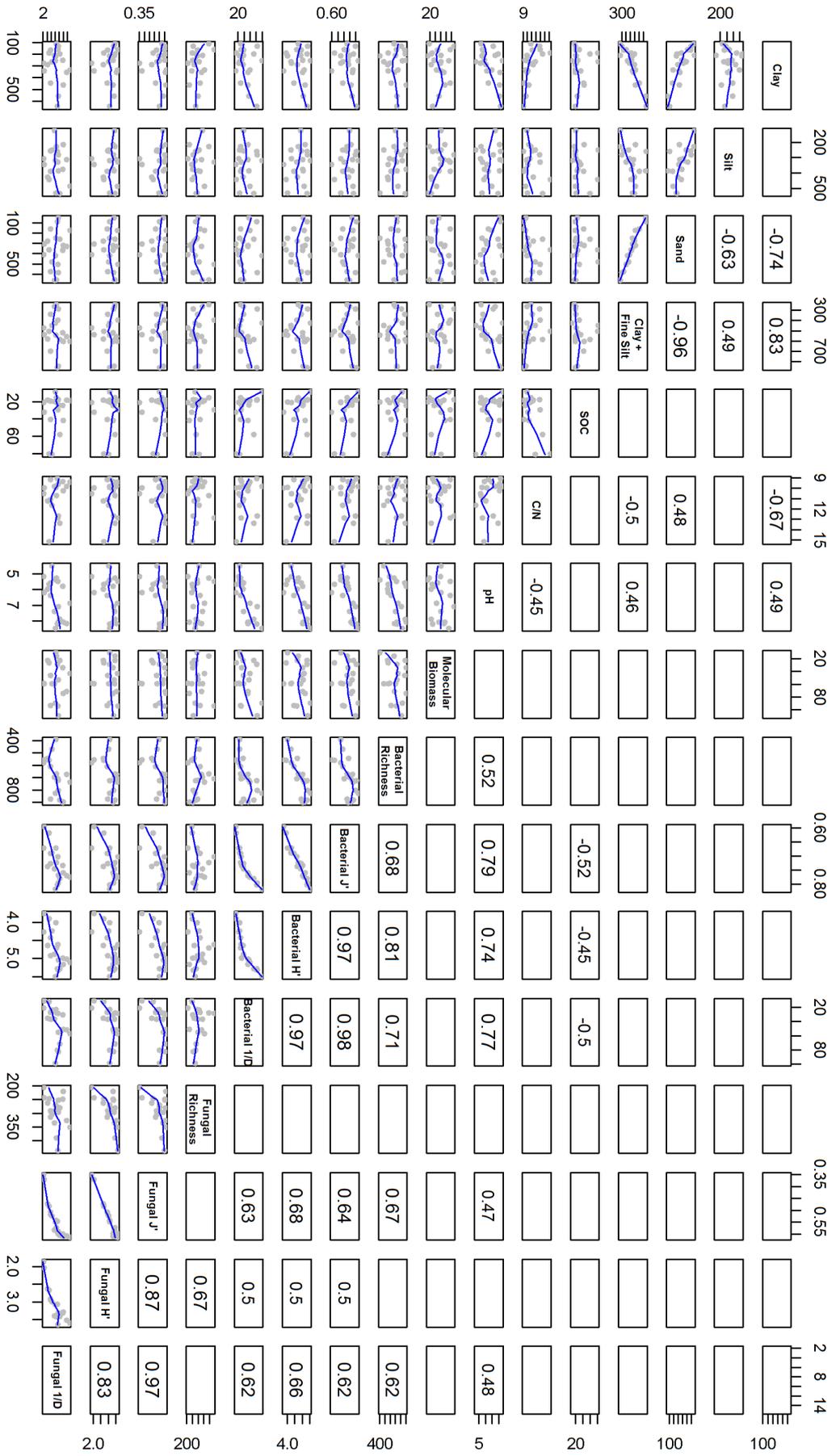


FIGURE 2.4 – **Matrice des corrélations de Spearman pour les propriétés des sols de DIMIMOS.**

Dans la partie supérieure de la matrice les nombres sont les valeurs des coefficients de Spearman. Une case blanche signifie que l'estimation du coefficient de corrélation n'était pas significativement différent de 0 ($p\text{-value} > 0,05$). La taille des coefficients est proportionnelle à leur valeur absolue. Dans la partie inférieure, les nuages de points entre les propriétés des sols sont représentés. La ligne bleue est une estimation de la relation à l'aide de fonctions de lissage.

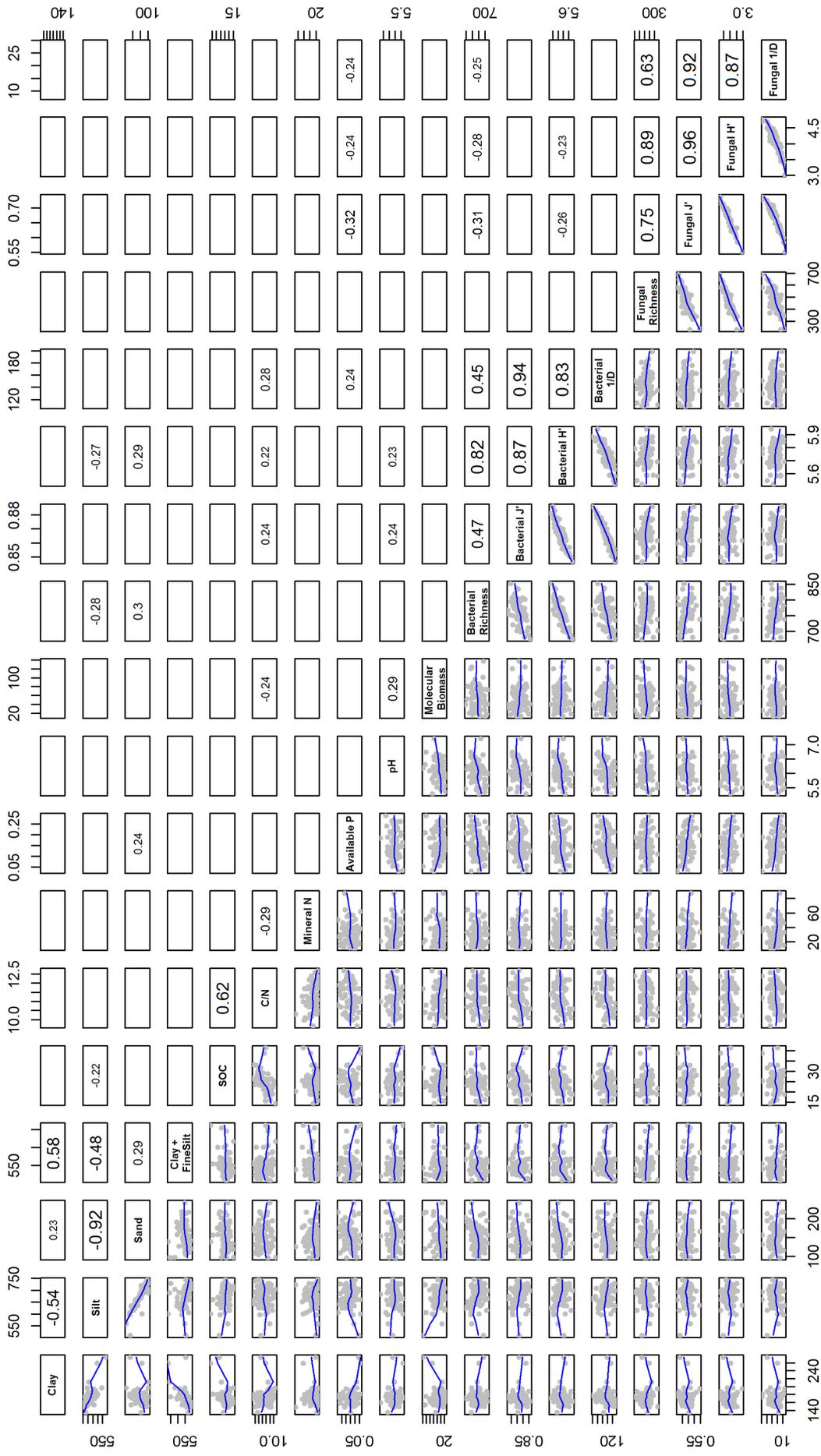


FIGURE 2.5 – Matrice des corrélations de Spearman pour les propriétés des sols de MOSAIC.

Dans la partie supérieure de la matrice les nombres sont les valeurs des coefficients de Spearman. Une case blanche signifie que l'estimation du coefficient de corrélation n'était pas significativement différent de 0 (p -value $> 0,05$). La taille des coefficients est proportionnelle à leur valeur absolue. Dans la partie inférieure, les nuages de points entre les propriétés des sols sont représentés. La ligne bleue est une estimation de la relation à l'aide de fonctions de lissage.

2.1.3 Dynamiques du carbone dans les sols échantillonnés

2.1.3.1 Analyse des dynamiques du carbone dans les sols échantillonnés

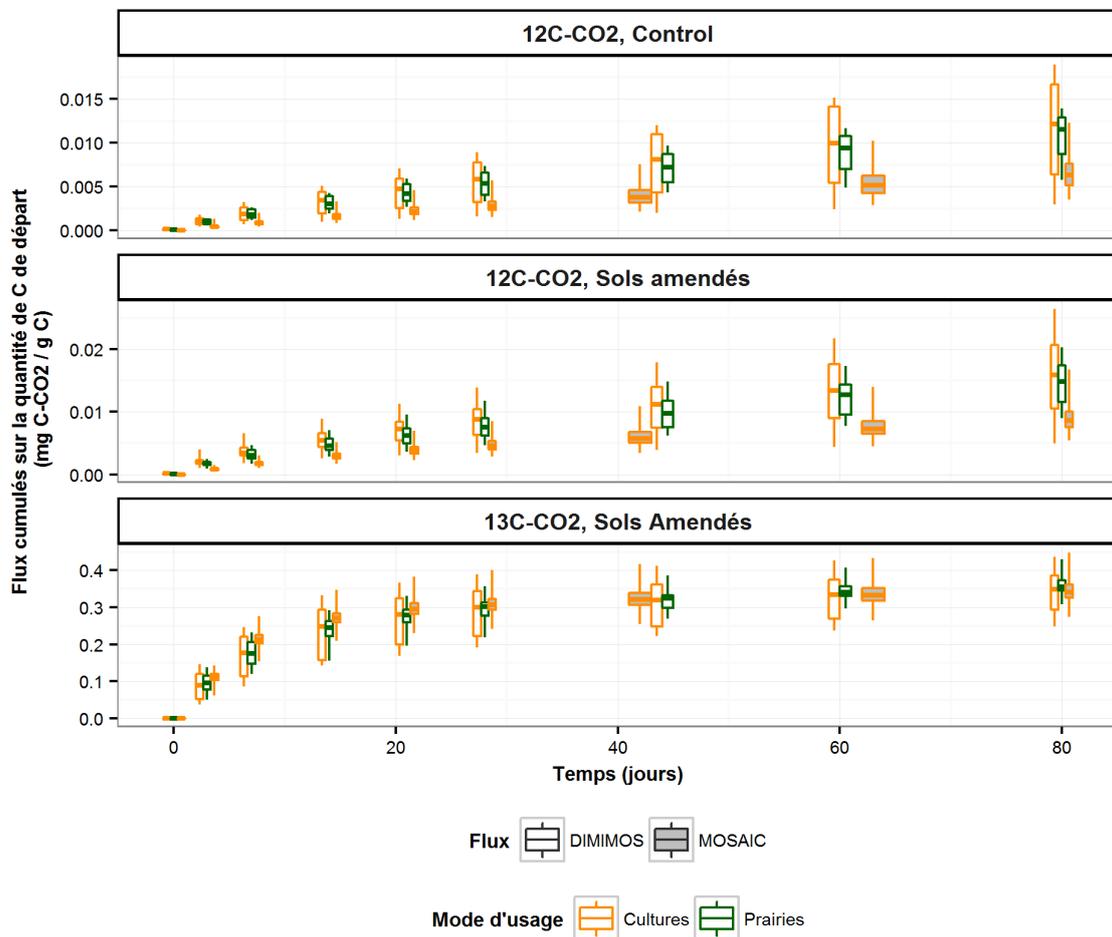


FIGURE 2.6 – Variabilité du CO₂ respiré cumulé, marqué ou non à l'isotope ¹³C, dans les sols témoins et les sols amendés.

Par soucis de comparaison, le CO₂ respiré a été divisé par la quantité de C de départ, i.e. la quantité de C dans la paille de blé pour le ¹³C-CO₂ et la teneur en C organique du sol pour le ¹²C-CO₂.

Les dynamiques du C dans les sols en conditions potentielles de laboratoire sont caractérisées par le suivi de la minéralisation du C organique sous forme de CO₂. La procédure est détaillée dans la section 3.1 et dans Tardy *et al.* (2015). Brièvement, des microcosmes contenant 30g de sol sec sont mis en pré-incubation à 20°C après avoir ajouté de l'eau jusqu'à 60% de la capacité au champ à raison de trois répétitions par sol et par traitement. Les 2 traitements sont le sol seul (control ou témoin) et le sol amendé, après pré-incubation, avec 5 mg/g de sol sec de paille de blé réduit en poudre dont le C a été marqué avec l'isotope ¹³C. L'intérêt d'un tel marquage est de pouvoir suivre indépendamment la cinétique de minéralisation du C du résidu (CO₂ marqué)

et celle de la minéralisation du C de la MOS (CO_2 non marqué). Les sols sont ensuite incubés pendant 80 jours en conditions contrôlées de température (20 °C) et d'humidité (60% de la capacité au champ).

Pour mesurer la minéralisation du C, le CO_2 respiré est alors mesuré à 3, 7, 14, 21, 28, 44, 60 et 80 jours pour DIMIMOS et à 3, 7, 14, 21, 28, 42, 63 et 80 jours pour MOSAIC. L'abondance en isotope ^{13}C est mesurée pour séparer la minéralisation du C du résidu ($^{13}\text{C}\text{-CO}_2$) de la minéralisation du C de la MOS ($^{12}\text{C}\text{-CO}_2$) (Figure 2.6)

2.1.3.2 Variabilité des cinétiques de minéralisation

Les flux de $^{12}\text{C}\text{-CO}_2$ du traitement control sont comparables entre les cultures et les prairies de DIMIMOS (autour de 1% du C organique de départ à 80 jours) mais sont plus faibles dans les sols de MOSAIC (autour de 0.6% du C organique de départ à 80 jours). En présence de résidu, ces différences sont comparables mais les valeurs médianes sont plus élevées (autour de 1.5% à 80 jours pour DIMIMOS et 0.9% à 80 jours pour MOSAIC). La différence de minéralisation entre les sols amendés et les sols témoins s'appelle le *priming effect* (PE).

Les flux liés à la minéralisation du C du résidu (paille de blé) sont plus similaires entre les deux bases. Autour de 35% du carbone du résidu a été minéralisé après 80 jours d'incubation.

2.2 Méthodes générales d'analyse

2.2.1 Lien entre les dynamiques du carbone et les propriétés des sols

Une première étude, présentée dans le chapitre 3, a pour objectif d'étudier le lien entre les dynamiques du C mesurées dans les sols des deux bases et les propriétés physiques, chimiques et biologiques de ces mêmes sols. Une approche statistique, innovante dans le domaine, est utilisée. Cette méthode consiste en une sélection des propriétés des sols qui permettent d'expliquer au mieux les variabilités des dynamiques de C dans des modèles additifs généralisés (GAM, [Hastie and Tibshirani, 1990](#); [Wood, 2011](#)). Ces modèles permettent entre autres de séparer l'effet de chaque covariable (i.e. les propriétés des sols) en offrant la possibilité d'un lien non linéaire entre la variable étudiée (i.e. les dynamiques du C) et ces covariables (cf. section 3.1.2.2). Cette étude, effectuée séparément pour DIMIMOS et MOSAIC, à plusieurs objectifs :

1. Confirmer la pertinence des indices de diversité bactérienne et fongique pour expliquer les dynamiques du C en complément des propriétés chimiques et physiques plus classiques.

Cette pertinence est confirmée si ces indices sont sélectionnés dans les modèles GAM lors de la méthode de sélection. En particulier, la séparation des mesures de minéralisation du C selon sa source (MOS ou MOF) permet l'analyse de la pertinence des indices pour chacune de ces minéralisations, ainsi que pour le PE. De plus :

- (a) la base de données DIMIMOS, en présentant des sols sous cultures et sous prairies, permet la comparaison de la pertinence de ces indices pour expliquer les dynamiques du C dans ces 2 modes d'usage contrastés.
 - (b) la base de données MOSAIC, comprenant plus de sols, permet la comparaison de la pertinence des indices de diversité pour expliquer les dynamiques du C selon la période d'incubation. Cette étude permet d'analyser une éventuelle dépendance au temps de la contribution de la diversité microbienne dans les dynamiques du C.
2. Comparer les résultats des analyses entre les sols de DIMIMOS et les sols de MOSAIC, afin d'évaluer si la pertinence des indices pour expliquer la dynamique du C diffère si la gamme de variabilité de la diversité (et des autres propriétés du sol) est plus restreinte (données MOSAIC sur un territoire peu étendu). Il est en effet possible de ne plus observer de lien entre dynamique du C et diversité microbienne si la variabilité de cette diversité est réduite (Figure 2.3)
 3. Discuter des résultats, avec l'appui de la littérature scientifique, afin de proposer des fonctions de modulation des paramètres d'un modèle mathématique de dynamique du C, dépendantes des propriétés du sol sélectionnées et en particulier des indices de diversité bactérienne et/ou fongique si tel est le cas.

2.2.2 Intégration et évaluation de la diversité microbienne dans un modèle mécaniste de dynamique du carbone

La deuxième étude, objet du chapitre 4, est une étape de modélisation mathématique. Un modèle mathématique simple (i.e. limité en nombre de paramètres) de dynamique du C, adapté à la modélisation des processus de décomposition/minéralisation de la MOF et de la MOS dans le cadre de données d'incubation, est choisit en fonction des résultats des étapes précédentes. Les différentes fonctions de modulation de paramètres proposées seront intégrées dans le modèle et la calibration du modèle est réalisée à l'aide des données MOSAIC. Afin d'utiliser les connaissances de la littérature sur les distributions potentielles des valeurs des paramètres et aider à la calibration du modèle, une calibration bayésienne est effectuée. Les fonctions de modulation, après

calibration de leurs paramètres, seront intégrées selon différentes combinaisons pour comparer le gain de chacune d'entre elles et en compagnie des autres. Pour mesurer le gain de qualité des modèles, 80% des sols de MOSAIC sont utilisées pour la calibration et des indices de qualité d'ajustement sont calculés, et 20% ne sont pas utilisés pour la calibration mais pour le calcul d'indices de qualité prédictive du modèle. L'objectif de cette étape est d'analyser si les fonctions de modulation proposées permettent d'améliorer le modèle. En particulier, on cherche à étudier si les fonctions dépendantes des indices de diversité permettent cette amélioration en plus de celle permise par des propriétés des sols plus classiquement prises en compte dans les modèles mathématiques de dynamique du C.

Chapitre 3

Les indices de diversité microbienne peuvent-ils expliquer la dynamique du C ?

Ce chapitre est destiné à répondre à la première question de recherche de la thèse et en déduire une stratégie d'intégration des indices de diversité microbienne dans les modèles de dynamique de C. Les deux premières parties concernent l'analyse de la pertinence des indices de diversité bactérienne et fongique pour expliquer la dynamique du C selon sa source (MOF, MOS), sur une courte période (80 jours), tout en considérant des covariables classiques (autres propriétés des sols) dont l'influence sur ces dynamiques est connue :

- En particulier, la première section concerne le jeu de données DIMIMOS utilisé pour aborder la question du lien entre diversité et dynamique du C selon le mode d'usage. Cette étude a fait l'objet d'un article publié dans PLoS One ([doi:10.1371/journal.pone.0161251](https://doi.org/10.1371/journal.pone.0161251)).
- Dans la deuxième section, à l'aide du jeu de données MOSAIC, plus conséquent en terme de nombre de sols étudiés, la dépendance temporelle du lien entre diversité et dynamique du C a été étudiée.
- Dans la troisième section, une comparaison détaillée des résultats des deux premières sections nous a permis de mieux cibler les flux (décomposition, assimilation, minéralisation du carbone) dans lesquels les indices de diversité pouvaient être intégrés et de proposer des fonctions de modulation des paramètres de ces flux dépendantes de la diversité microbienne, dans des modèles de dynamique de C.

3.1 Pertinence des indices de diversité pour expliquer la variabilité de dynamique du C de sols échantillonnés à l'échelle nationale

Cette section a fait l'objet d'une publication intitulée « *Microbial Diversity Indexes can Explain Soil Carbon Dynamics as a Function of Carbon Source* » dans la revue *PLoS One* ([doi:10.1371/journal.pone.0161251](https://doi.org/10.1371/journal.pone.0161251)).

3.1.1 Introduction

Among soil biological processes, mineralisation of soil organic matter (SOM) is essential, as SOM is a key component contributing to many functions and services in soil ecosystems (Scharlemann *et al.*, 2014; Banwart *et al.*, 2015). Mathematical models are useful tools to quantitatively describe processes involved in SOM dynamics and help predict the influence of management practices (Kutsch *et al.*, 2009). Many SOM dynamics models have been developed to date, and microorganism biomass is increasingly represented explicitly (Manzoni and Porporato, 2009).

However, microbial diversity is nearly absent in these models despite new evidence of its role in SOM dynamics. Two main reasons have led to this absence : i) from a conceptual perspective, SOM dynamics models performed sufficiently well without needing to consider microbial diversity (Smith *et al.*, 1997) ; and ii) for a long time, technical limitations have hindered demonstration of a relationship between microbial diversity and SOM dynamics, as well as identification of qualitative variables to better describe the microbial pool.

The latest technical advances in molecular biology have made the latter possible, and recent results have shown that microbial diversity can significantly influence transformation of carbon (C) and nitrogen (N) in the soil (Tardy *et al.*, 2015; Philippot *et al.*, 2013; Juarez *et al.*, 2013; Baumann *et al.*, 2013). For instance, Baumann *et al.* (Baumann *et al.*, 2013) found reduced lignin and wheat sugar decomposition, while Philippot *et al.* (Philippot *et al.*, 2013) showed reduced denitrification activity with low microbial diversity.

These results qualitatively demonstrated a microbial diversity-SOM dynamics relationship. However, there is still a need to quantitatively describe this relationship between microbial diversity and SOM dynamics processes and make it possible to integrate it in SOM models. Increasing evidence in the literature indicates that taxonomic and functional compositions of microbial communities are strong drivers of SOM processes (Hooper *et al.*, 2005), and bacterial

phyla have been identified as functional groups (Fierer *et al.*, 2007). However, phylum is a high taxonomic rank, and members of the same phylum can exhibit different functional traits.

Consequently, knowledge about microbial composition and functional traits currently remains limited, and further study is required to understand the relationship between microbial composition and C and N dynamics and identify key taxa as indicators of microbial community functional traits (Nannipieri *et al.*, 2003). Given this lack of knowledge, we argue that it remains difficult to use microbial composition to build the quantitative relationships discussed above, which are necessary to parameterise predictive models. Interesting modelling approaches have integrated microbial diversity by representing different functional groups of microorganisms according to their differing affinities for organic substrates (Moorhead and Sinsabaugh, 2006; Perveen *et al.*, 2014) or enzyme-production strategies (Allison, 2012). However, because of the limited knowledge, the microbial groups in these models are conceptual; thus, it is challenging to validate these models with empirical data (Waring *et al.*, 2013).

In contrast, promising demonstrations of the microbial diversity-SOM dynamics relationship have occurred using microbial diversity indexes (e.g. richness, evenness, Shannon index) to assess microbial diversity (Tardy *et al.*, 2015; Philippot *et al.*, 2013; Juarez *et al.*, 2013; Baumann *et al.*, 2013). This suggests that overall microbial diversity indexes could be relevant covariates of SOM dynamics. Since high-throughput sequencing can now easily quantify microbial diversity in environmental samples, it is possible to assess the increase in accuracy of SOM dynamics models due to integrating microbial diversity via these covariates.

Recently, Tardy *et al.* (2015) showed that microbial diversity explained C dynamic differently according to the quality of C substrates. In particular, they demonstrated that the importance of bacterial vs. fungal diversity may vary with the recalcitrance of C substrates, in agreement with other studies which demonstrated that fungi decompose recalcitrant substrates better than bacteria (de Boer *et al.*, 2005). Additionally, quantitative and qualitative improvement of SOM is generally observed in agroecosystems favouring a fungal-dominated community (Six *et al.*, 2006). Consequently, we hypothesise that model accuracy will improve if the microbial variables distinguish bacterial and fungal diversity, each of which influence dynamics of SOM pools differently as a function of their recalcitrance. Knowing this, it is necessary to go further by building quantitative relationships between C dynamics and bacterial and fungal diversity indexes to investigate the implication of integrating microbial diversity in SOM dynamics models.

The objective of the present study was to statistically investigate the relationships between

bacterial and fungal diversity and decomposition of C pools, characterised by the pattern of CO₂ emissions under controlled conditions. In particular, we aimed to i) confirm the relevance of microbial diversity indexes as covariates to explain variability in CO₂ emissions, ii) build quantitative relationships between microbial diversity and CO₂ emissions and iii) evaluate implications of integrating microbial diversity in SOM dynamics models. Twenty soils with a variety of characteristics and two different land-use histories (cropland and grassland) were incubated with and without addition of ¹³C-labelled wheat residue in soils. ¹³C-labelling allowed us to distinguish and separately analyse the relationship between bacterial and fungal diversity and mineralisation of different sources of OC (soil, wheat residue), as well as the interaction between mineralisation of these two sources (priming effect). Land-use history was taken into account because of the increasing body of evidence that these two land uses differ greatly in their influence on microbial diversity in the soil (Acosta-Martinez *et al.*, 2008; Jangid *et al.*, 2008, 2011; Rampelotto *et al.*, 2013), which has consequences for C cycling (Tardy *et al.*, 2015). Consequently, we hypothesised that studying these two land uses could reveal contrasting diversity-SOM dynamics relationships that are important to understand. Diversity of bacterial and fungal communities was characterised by high-throughput sequencing of ribosomal genes. We applied an innovative data mining approach, based on generalized additive models (GAM) (Wood, 2011; Hastie and Tibshirani, 1990) and a predictive criterion, to select covariates that better explain variability in CO₂ emissions for each C source and explore the relative importance of bacterial and fungal diversity, along with classic soil properties which influence SOM dynamics. The GAM approach provides quantitative relationships useful for integrating microbial diversity in SOM dynamics models.

3.1.2 Materials and methods

3.1.2.1 Data source

Soil samples and analyses

We considered a database of 20 agricultural soils sampled from the French Soil Quality Monitoring Network (RMQS) (Arrouays *et al.*, 2002). The 20 locations of these soils had high variability in soil properties (in particular pH and texture) for two land-use classes : cropland (10 soils) and grassland (10 soils). Croplands corresponded to monoculture systems or crop rotations with or without short-term grasslands. Grasslands corresponded to permanent or long-term grasslands (more than ten years).

Concerning abiotic soil properties, texture was measured by standardized 5-fraction granulometry (NF X 31-107), soil organic carbon (SOC) content was measured by standardized dry combustion (NF ISO 10694) and pH by standardized 1 : 5 water suspension (NF ISO 10390) (Jolivet et al., 2006). These samples were then air-dried and stored in the French national soil sample archive following the protocol described by [Ratié et al. \(2010\)](#).

Annexe A provides more information about the locations and studied soils. The field studies were carried out on private lands where owners gave permission to conduct the soil samplings. Since no protected areas were involved, no specific permissions were required. The field studies did not involve endangered or protected species.

Carbon mineralisation measurements

C mineralisation was measured using microcosm respiration following the procedure of [Tardy et al. \(2015\)](#). Microcosms were established by placing 30 g of equivalent dry soil supplemented with sterile water to attain 60% soil water-holding capacity in 150 ml hermetically sealed plasma flasks. These microcosms were pre-incubated at 20°C for three weeks to avoid the heavy disturbance (overestimation of C mineralisation) caused by adding water after storage. Three replicates of each microcosm were then amended with ¹³C-labelled powder of wheat residues (5 mg g⁻¹ dry weight of soil) while three others were not (control).

Both amended and control microcosms were mechanically mixed. The 120 microcosms (20 locations × 2 treatments (control and amended) × 3 replicates) were incubated in the dark for 80 days under controlled temperature (20 °C) and moisture conditions (60% of water-holding capacity).

Respired CO₂ was measured after 3, 7, 14, 21, 28, 44, 60 and 80 days of incubation in microcosms. The gaseous phase was sampled in 10 ml airtight flasks to measure the CO₂ concentration and in 12 ml airtight flasks to determine the C isotope (¹³C) abundance. The concentrations measured at each sampling date corresponded to the CO₂ accumulated between two sampling dates. ¹³C-labelling of the plant residues allowed separating SOC (R_s) and plant-residue (R_r) mineralisation (μg C-CO₂ g⁻¹ soil) using mass-balance equations :

$$R_s + R_r = R_t \text{ and } R_s * A_s^{13} + R_r * A_r^{13} = R_t * A_s^{13} \quad (3.1)$$

$$R_s = \frac{R_t * A_s^{13} - R_r * A_r^{13}}{A_s^{13}} \quad (3.2)$$

where A_i^{13} , $i = \{s, r, t\}$ is the ^{13}C abundance in soil C, plant residue and total respired CO_2 (R_t) emitted by amended soil, respectively. In this study, the priming effect (PE) was calculated as :

$$PE = \frac{R_{s,amended}}{R_{s,control}} \quad (3.3)$$

where $R_{s,amended}$ and $R_{s,control}$ are SOC mineralisation in amended and control microcosms, respectively. PE was calculated as a ratio, rather than the more traditional difference between $R_{s,amended}$ and $R_{s,control}$, to avoid negative values. This allowed a log-transformation to meet the assumptions of the statistical models used later.

To assess differences in C mineralisation rates between cropland and grassland, a two-way analysis of variance (ANOVA) was performed on each mineralisation kinetic rate ($R_{s,control}$, $R_{s,amended}$, R_r and PE). The model included a fixed factor, “land use”, and a random factor, “time”. Post-hoc analysis, based on Tukey’s honest significant difference test (Tukey, 1949), was performed to assess differences between each level of both factors and their interaction.

Microbial biomass and diversity determination

Following Tardy *et al.* (2015), the microbial community was analysed in each soil just after pre-incubation and before incubation. Microbial DNA was extracted from 1 g of each soil replicate using a slight modification of the ISO-10063 procedure (Plassart *et al.*, 2012). DNA concentrations were determined and used as estimates of molecular microbial biomass (Dequiedt *et al.*, 2011). After DNA purification (MinElute gel extraction kit, Qiagen, Courtaboeuf, France), bacterial and fungal diversity was determined for each soil. For bacteria, a 16S rRNA gene fragment with sequence variability and about 450 bases for 454 pyrosequencing was amplified by PCR using primers F479 and R888. For fungi, an 18S rRNA gene fragment of about 350 bases was amplified using primers FR1 and FF390.

After the pyrosequencing procedure, as described in Tardy *et al.* (2015), bioinformatic analyses were performed on the 16S and 18S rRNA gene sequences to cluster them at 95% sequence similarity into operational taxonomic units (OTU). From OTUs, quantitative diversity indexes were calculated, such as bacterial and fungal richness (number of OTUs), as well as other indexes that consider OTU abundance (Shannon index, H' ; evenness, J' ; and inverse Simpson index, $1/D$). H' increases with an increase in richness and the equity of OTU abundance. J' (range = 0-1) provides information only about the equity of OTU abundance and equals the ratio of H' to its maximum potential value (i.e. if all OTUs had the same abundance). The index

$1/D$ gives the probability that two individuals randomly selected from a sample will belong to the same OTU ; higher values indicate greater diversity. It gives more weight to the more abundant species in a sample. The addition of rare species to a sample causes only small changes in its value, contrary to H' .

3.1.2.2 Statistical selection of soil properties that better predict mineralisation kinetics

We assumed that mineralisation kinetics (R_s , R_r , PE) are influenced by soil properties. The objective was to verify whether microbial diversity parameters are as suitable for predicting C mineralisation as the abiotic soil properties frequently considered in SOM models. Using GAM (Wood, 2011; Hastie and Tibshirani, 1990), soil properties including microbial diversity were statistically selected to explain variability in mineralisation kinetics. Properties were selected for each mineralisation kinetic ($R_{s,control}$, $R_{s,amended}$, R_r and PE) using the *mgcv* package (Wood, 2006) in the free statistical software R (R Core R Core Team, 2015).

Theory of generalized additive models

The structure of a GAM (Hastie and Tibshirani, 1990 ; Wood, 2011) can be written as :

$$g(E(Y_i)) = \alpha + f_1(x_{1i}) + f_2(x_{2i}) + \dots + f_p(x_{pi}) \quad (3.4)$$

where Y_i is a response variable following some exponential family distribution, α is the intercept of the model, and $f_j, j = 1, \dots, p$ are functions of the covariates x_j . This structure allows for nonlinear dependence of the response on the covariates. As functions f_j 's are not necessarily known, they are specified as smooth functions rather than detailed parametric ones. Smooth functions are flexible data-driven functions estimated by semi-parametric methods often using a spline basis. Nevertheless, the f_j 's can also be specified as known functions e.g. identity function for a linear dependence or a power-family function for a polynomial dependence in which case the model becomes a generalized linear model. As for generalized linear models, the link function g allows for a non-normal distribution of $E(Y_i)$.

Stepwise selection of soil properties

The response variable to predict with the GAM referred to C-CO₂ respired at each measurement date : we focused on the rate of respired C-CO₂ ($\Delta R/\Delta t$), defined as the amount of

respired C-CO₂ divided by the number of days between two measurement dates. To avoid heteroscedasticity in models, the logarithm of this rate for the eight measurement dates was used as a response variable for mineralisation kinetics.

The covariates considered for selection with GAM included the abiotic soil properties influencing C mineralisation usually considered in SOM models, as well as microbial diversity indicators. The abiotic variables considered were texture properties (clay, silt, sand and clay + fine silt contents), SOC content, soil C :N ratio and pH. The biological covariates were microbial richness, J', H' and 1/D for both bacteria and fungi. We also included molecular microbial biomass, as it can explain C dynamics greatly, especially interactions between residue and SOC mineralisation (Fontaine and Barot, 2005; Wutzler and Reichstein, 2008). One-way ANOVA was performed to assess the significance of differences between means of soil properties of cropland and grassland.

Relationships between soil properties and the response variable were built with smooth functions estimated by penalised cubic regression splines (s). However, we also allowed f_j to be known parametric functions such as identity (I) or polynomial of degree 2 (poly2) or 3 (poly3). Using these functions reduced the risk of overfitting, a well-known limitation of the GAM approach (Wood, 2006). To assess differences in relationships with cropland and grassland soils, each combination of soil properties and functions f_j interacted with a covariate factor corresponding to “land use”. By doing this, cropland and grassland soils could have different relationships between soil properties and response variables.

Since rates of respired C were time dependent, time (used as a categorical variable) was an obligatory covariate in the GAM. Since measures in mesocosms occurred over time, the covariate “time” represents a temporal pseudo-replication and, consequently, was considered as a random factor. Consequently, the simplest model selected (Figure 3.1 ; step 0, initial model) was :

$$Y = \mu + time \tag{3.5}$$

where Y is the response variable (mineralisation kinetic), μ is the intercept, and time is the random effect of covariate time (in days).

We defined P as all soil properties – which were potential covariates in GAM – and S as the soil properties selected during the selection procedure.

Starting from equation (2), the steps of the selection procedure were (Figure 3.1) :

1. Calculate a mean squared error of prediction (MSEP) of the currently selected model by cross-validation :

$$MSEP = \frac{1}{N} \sum_i (y_{itu} - \hat{y}_{-i,tu})^2 \quad (3.6)$$

where y_{itu} is the value of the response variable for soil i at time t for replicate u , $\hat{y}_{-i,tu}$ is the estimated value of the response variable for soil i at time t for replicate u using the selected model calibrated without values of soil i , and N is the total number of values for the response variable. MSEP is a classic indicator of the predictive quality of a model. Cross-validation estimation of MSEP is known to be less subject to overfitting than a classic goodness-of-fit index.

2. To avoid problems associated with collinearity of covariates in GAM (Guisan *et al.*, 2002), pre-select soil properties based on a variance-inflation factor in linear models between each soil property in P and the soil properties already selected in the model (S) with a threshold of 4 (Mansfield and Helms, 1982).
3. Build all potential improved models by adding one of each combination of pre-selected soil properties in R and associated functions (I, poly2, poly3, and s) or by removing one previously selected soil property. The latter enabled removing previously selected soil properties in case they contained redundant information with newly selected ones.
4. Calculate MSEP for each model by cross-validation.
5. Watch for the model k with the smallest MSEP :
 - (a) If the MSEP of model k is smaller than that of the currently selected model (step 1), model k becomes the new selected model. If model k was better because of a newly added soil property from P , this soil property becomes the one selected and moves to group S (i.e. those previously selected). If model k was better because of the removal of a previously selected soil property, the latter moves from group S to P . The procedure returns to step 1.
 - (b) If the MSEP of model k is greater than that of the currently selected model (step 1), there is no more improvement, and the procedure stops. The current model is the final one, and S contains all of the pre-selected soil properties that best predict the response variable.

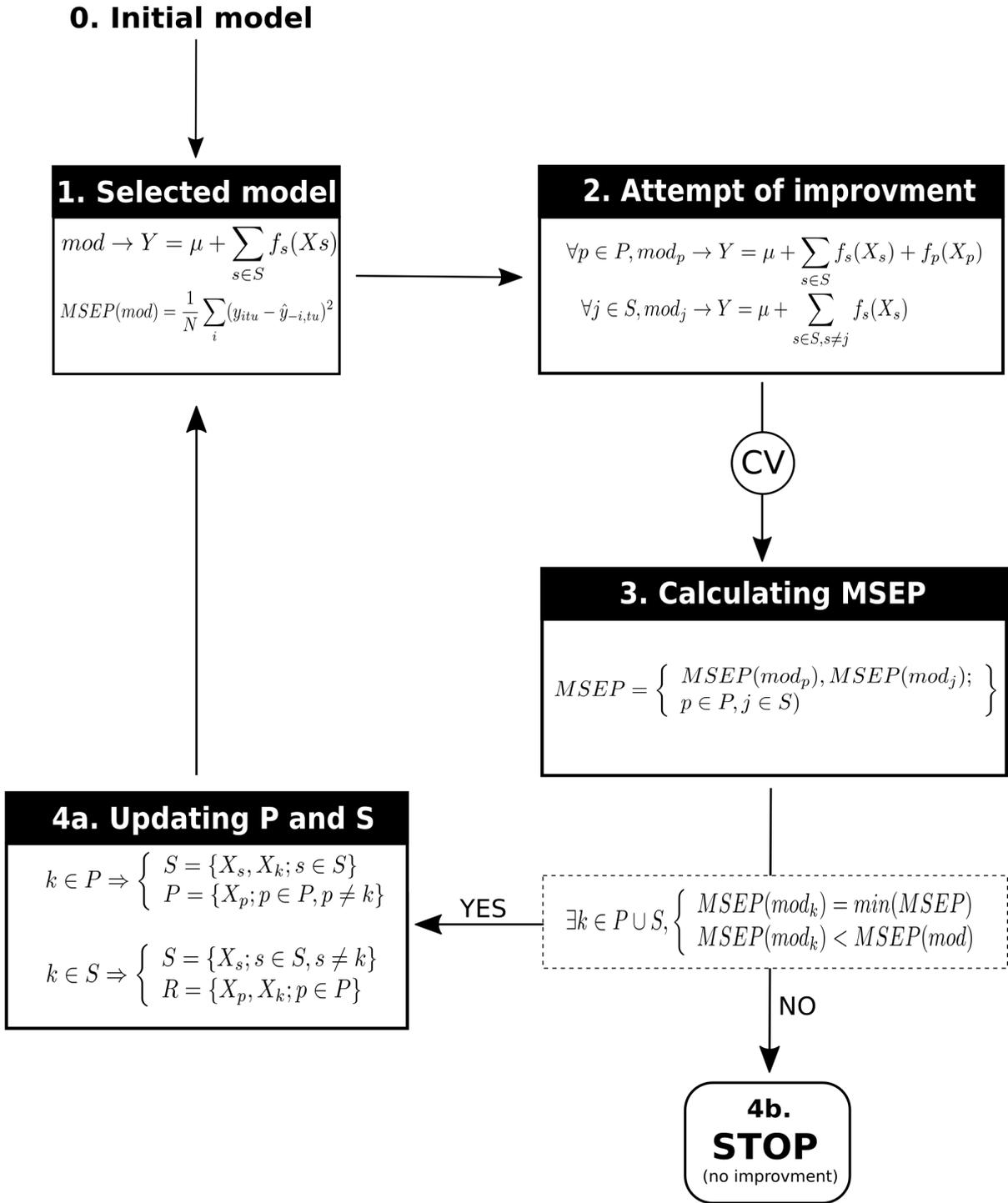


FIGURE 3.1 – Predictor selection procedure using generalized additive models.

P is the set of potential covariates, S is the set of selected covariates, V is a sub-set of P that contains potential covariates not collinear with covariates in S (i.e. already selected), $MSEP()$ indicates the function to calculate $MSEP$, f_j 's are functions (smooth or parametric) to represent relationships between covariates and the response variable (Y) and CV means cross validation.

Assessment of selected models

After the selection procedure, model residuals were graphically checked for identical and independent distribution hypotheses, as GAM can be sensitive to violation of distribution assump-

tions (Wood, 2006). The total percentage of explained deviance (%Dev) was measured to assess the goodness-of-fit of the selected models. To assess the predictive quality of the selected models, the ratio of inter-quartile range to root MSEP (RPIQ) was calculated as the ratio of the inter-quartile range (IQR) of the response variable to the square root of the MSEP calculated according to equation (6). RPIQ, developed by Bellon-Maurel et al. (Bellon-Maurel *et al.*, 2010), represents the degree to which the dispersion of the response variable exceeds the model's prediction error. Higher RPIQ indicates better predictive quality of the model.

Relative importance of selected soil properties

The relative importance of selected soil properties was assessed by estimating the contribution of each selected soil property to the deviance using the R package *vegan* (Oksanen *et al.*, 2016). To achieve this, each component of the linear predictor, i.e. the transformation of each soil property by its I, poly2, poly3 or s function, was extracted from the model, and redundancy analysis was performed between the response variable and these components. Following a variance partitioning approach, the percentage of deviance explained by each component was calculated using the sum of squares from an ANOVA of the redundancy analysis result.

3.1.3 Results

3.1.3.1 Comparison of soil properties and cumulative respiration of cropland and grassland

Except for bacterial 1/D, Bacterial J' and Silt ($P=0.002$, 0.079 and 0.053 , respectively), soil properties did not differ significantly between cropland and grassland soils at the 10% level of significance between cropland and grassland (Figure 3.2). However, the dispersion of SOC content was slightly larger in grassland soils than in cropland soils, as the median in grassland soils was higher than the third quartile in cropland soils.

Regarding soil functioning, ANOVA showed significant differences ($P < 0.05$) between incubation times (Figure 3.3). Significant differences ($P < 0.05$) were also observed between cropland and grassland at each time of incubation, except for the PE, whose difference was significant only at 3 and 7 days ($P = 0.002$ and 0.007 , respectively) and at 14 days at the 10% level of significance ($P = 0.069$).

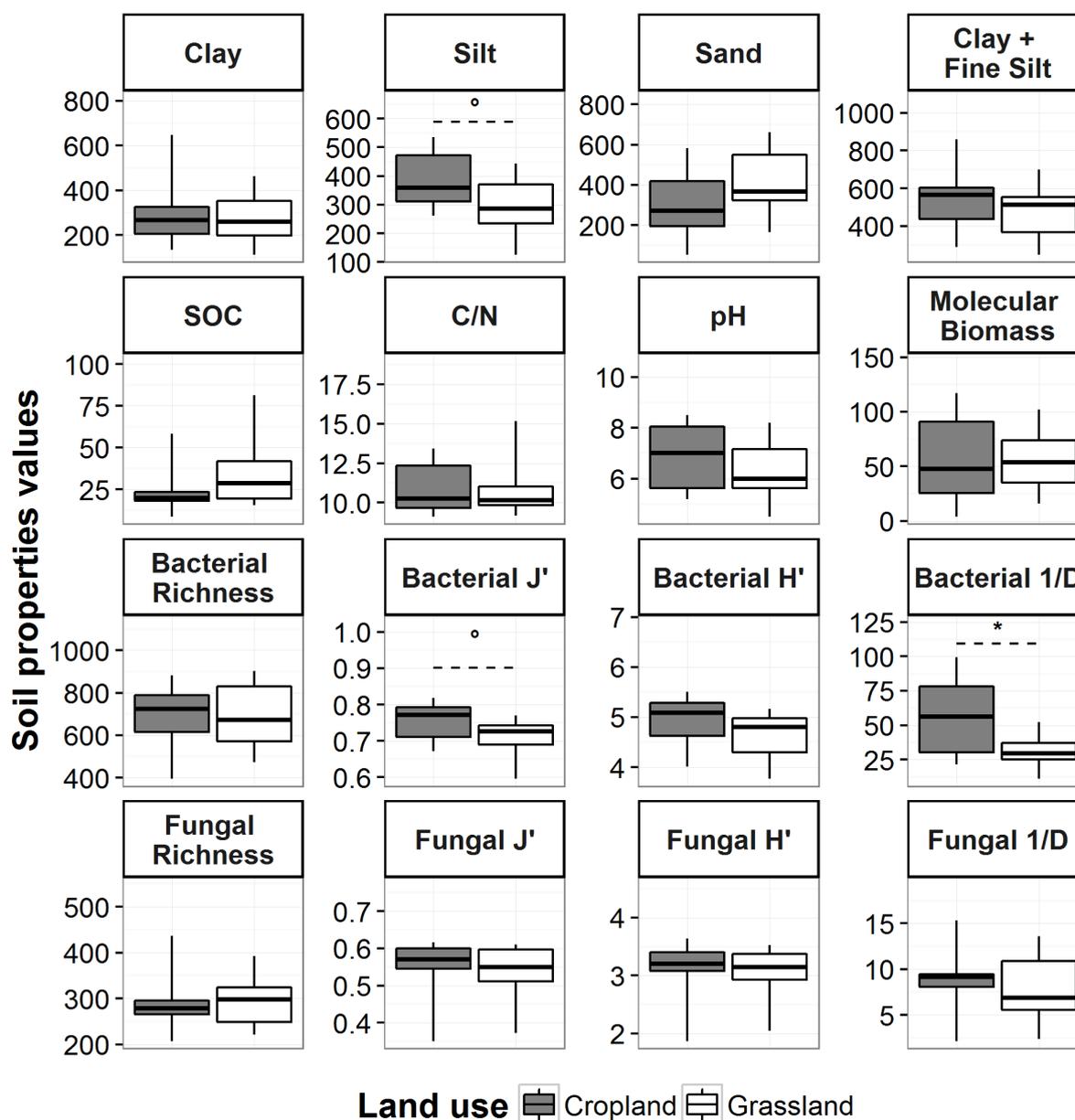


FIGURE 3.2 – Variability in soil property values in croplands and grasslands.

Bottom and top whiskers of boxplots extend to the lowest and highest values, respectively. Units of soil properties are $[g.kg^{-1} \text{ soil}]$ for texture variables and soil organic carbon (SOC) content, $[\mu g \text{ DNA}.g^{-1} \text{ soil}]$ for molecular biomass, [number of operational taxonomic units (OTUs)] for bacterial and fungal richness, [pH unit] for pH and dimensionless for C :N ratio and bacterial and fungal Shannon index (H'), evenness (J') and inverse Simpson index ($1/D$). Dashed lines above boxplots indicate significant differences between cropland and grassland, and the symbol shows the level of significance (***) $P < 0.001$, ** $0.001 < P < 0.01$, * $0.01 < P < 0.05$, ° $0.05 < P < 0.1$).

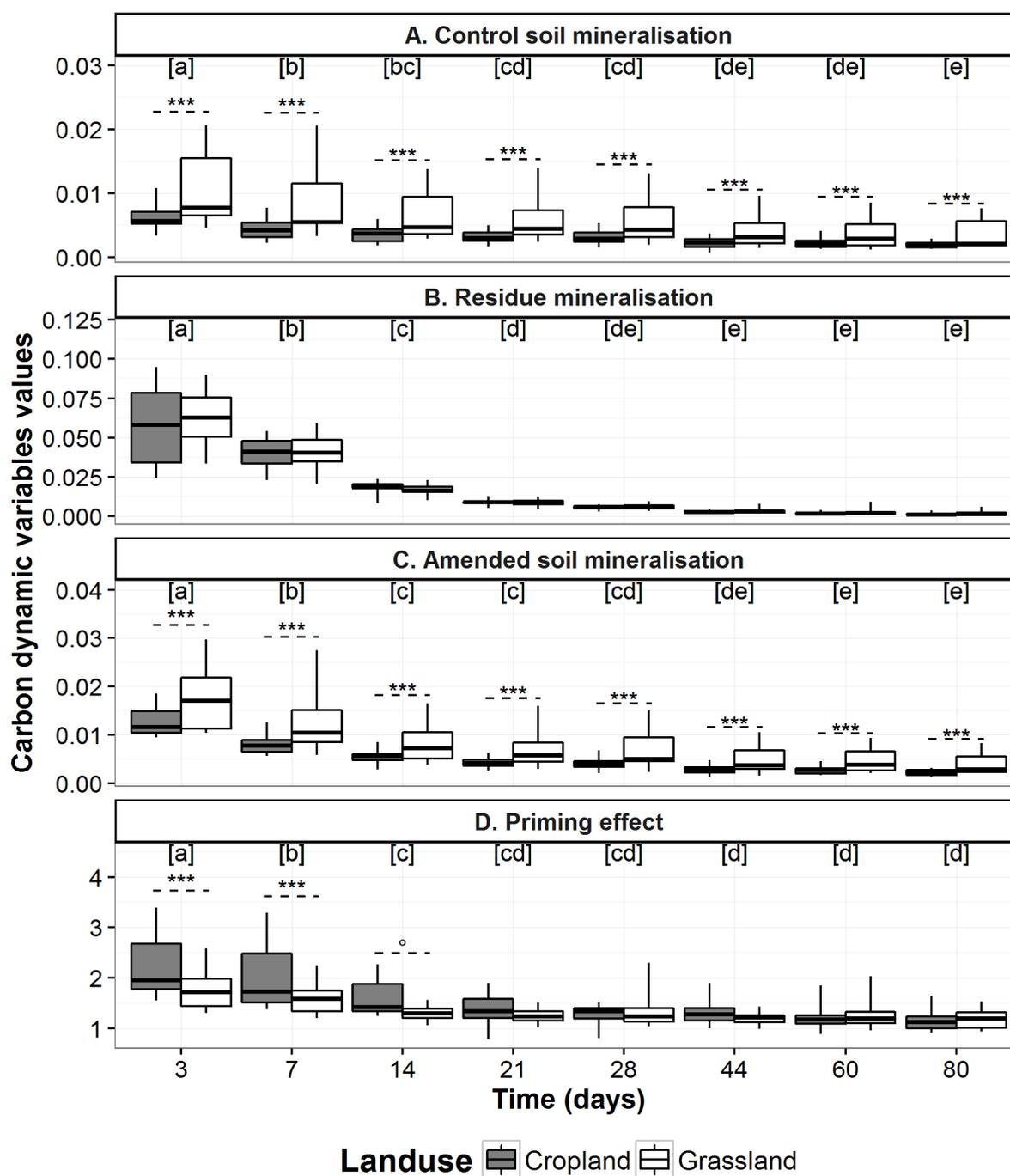


FIGURE 3.3 – Variability in mineralisation rate and priming effect (PE) in croplands and grasslands at each sampling time.

(A) Control-soil mineralisation ($R_{s,control}$, $\text{mg C-CO}_2 \text{ g}^{-1} \text{ soil day}^{-1}$), (B) residue mineralisation (R_r , $\text{mg C-CO}_2 \text{ g}^{-1} \text{ soil day}^{-1}$), (C) amended-soil mineralisation ($R_{s,amended}$, $\text{mg C-CO}_2 \text{ g}^{-1} \text{ soil day}^{-1}$) and (D) PE (dimensionless). Bottom and top whiskers of boxplots extend to the lowest and highest values, respectively. Dashed lines above boxplots indicate significant differences between cropland and grassland, and the symbol shows the level of significance (***) $P < 0.001$, ** $0.001 < P < 0.01$, * $0.01 < P < 0.05$, ° $0.05 < P < 0.1$). Different letters above boxplots indicate significant ($P < 0.05$) differences between incubation times.

3.1.3.2 Quality of selected models

TABLE 3.1: Assessment of models and covariates selected for each mineralisation parameter.

| | df | VarExp ^a | RPIQ | Covariates | Interaction ^b | Relation ^c | Significance ^d | RelImp ^e |
|------------------------------|-------|---------------------|------|-------------------|--------------------------|-----------------------|---------------------------|---------------------|
| R_{s,control} | 18.98 | 93 | 2.9 | Time | NO | RF | *** | 33 |
| | | | | Sand | NO | P3 | *** | 8 |
| | | | | C :N ratio | YES | I | *** | 7 |
| | | | | Bacterial J' | YES | P3 | *** | 44 |
| R_{s,amended} | 13.98 | 90 | 4 | Time | NO | RF | *** | 60 |
| | | | | SOC | YES | I | *** | 13 |
| | | | | pH | NO | I | *** | 12 |
| | | | | Fungal Richness | NO | P2 | *** | 4 |
| | | | | Fungal 1/D | NO | I | *** | <1 |
| R_r | 8.99 | 93 | 6 | Time | NO | RF | *** | 93 |
| | | | | Silt | NO | I | *** | <1 |
| PE | 19.98 | 74 | 2 | Time | NO | RF | *** | 46 |
| | | | | Clay+Fine Silt | YES | P2 | *** | 1 |
| | | | | SOC | NO | I | ns | 1 |
| | | | | C :N ratio | YES | S | ** | 1 |
| | | | | pH | NO | I | ** | 13 |
| | | | | Molecular biomass | NO | I | ** | 3 |
| | | | | Fungal 1/D | NO | S | *** | 9 |

df, degrees of freedom; *RPIQ*, Ratio of inter-quartile range to root Mean Square Error of Prediction; *J'*, evenness; *SOC*, Soil Organic Carbon; *1/D*, inverse Simpson index

^a Pourcentage of explained variance by the model

^b Interaction with land-use?

^c Type of relationship selected : RF, random factor ; I, identity ; P2, polynomial of degree 2 ; P3, polynomial of degree 3 ; s, smooth function

^d Significance of selected covariates based on the Fisher test : *** P<0.001 ; ** 0.001<P<0.01 ; * 0.01<P<0.05 ; ° 0.05<P<0.1 ; ns (not significant) P>0.1

^e Relative importance (%) of selected covariate based on variance partitioning

The selected models were able to explain large percentages of the variability in mineralisation response variables, as %Dev ranged from 74% for the PE model to 93% for the residue and control mineralisation model (Table 3.1). For all models, a large percentage of explained deviance was due to the time covariates, ranging from 33% in control soil mineralisation to 93% in residue mineralisation. Predictive qualities of the models were more variable, ranging from moderate (RPIQ \approx 2) to good (RPIQ \geq 6). RPIQ values increased in the following order : PE (RPIQ=2), control soil mineralisation (RPIQ=2.9), amended soil mineralisation (RPIQ=4), and residue mineralisation (RPIQ=6). Model complexity had the opposite order, as assessed by the total degrees of freedom in the models.

3.1.3.3 Relative importance of selected soil properties in predictions of mineralisation kinetics

All selected soil properties were significant in the models, except SOC content in PE models ($P = 0.45$) (Table 3.1). Among selected abiotic soil properties, texture covariates were selected for three models, but with a high relative importance only for $R_{s,control}$ (8%), which was positively related to sand content (Figure 3.4). SOC content in interaction with land-use history was highly positively related to $R_{s,amended}$ (13%), while the C :N ratio in interaction with land-use history was negatively related to $R_{s,control}$ variability (7%). Soil pH was selected in models for $R_{s,amended}$ and PE, explaining 12% and 13%, respectively ; however, it was negatively related to $R_{s,amended}$ and positively related to PE.

Among selected biological soil properties, molecular microbial biomass was influential only for PE, with a positive relationship and a low relative importance (3%). None of the diversity indexes were selected for R_r . In control incubation, a polynomial relationship with bacterial J' in interaction with land-use history explained 44% of $R_{s,control}$ variability. The relative importance of microbial diversity indexes in other models was lower : fungal 1/D, associated with a polynomial function, explained 9% of PE variability and less than 1% of that in $R_{s,amended}$. Fungal richness was negatively related to $R_{s,control}$ and explained 4% of its variability.

3.1.4 Discussion

3.1.4.1 Quality of the selected models

Except for bacterial 1/D, bacterial J' and silt content, there was as much variability in soil property values and mineralisation response variables within each set of soils (cropland and

grassland) as between the two sets, indicating that differences in the relationships built by the GAM approach for cropland and grassland were due more to differences in processes driving C dynamics between the land-uses than to differences in soil properties. However, the low power of the ANOVA – due to a small dataset ($n=10$ for each set) – decreased its ability to detect significant differences. Grasslands tended to have higher and more variable SOC contents than croplands (Meersmans *et al.*, 2012) and higher values of bacterial diversity indexes (H' , J' , $1/D$). In terms of functioning, PE tended to be slightly higher in croplands, which agrees with a recent study showing an increase in PE intensity with an increase in land-use intensity (Tardy *et al.*, 2015).

Much of the variability was explained by the selected soil properties, as demonstrated by the percentage of deviance explained by each model. Thus, the soil properties contained enough information to explain SOC mineralisation. However, predictive quality depended more on the model. Not surprisingly, predictive quality was linked to model complexity, as assessed by its degrees of freedom (Table 1) : the simpler the model, the higher was its predictive quality. The model for PE was the most complex but had the lowest quality, suggesting that processes involved in PE are also complex or that more relevant covariates, such as mineral nutrient availability (Fontaine *et al.*, 2011; Chena *et al.*, 2013), were not included.

3.1.4.2 Relevance of microbial diversity indexes in explaining SOM dynamics

Microbial diversity indexes were selected for three activities describing SOM dynamics, demonstrating their ability to explain variability in SOM dynamics. The absence of microbial diversity indexes in the residue mineralisation model did not agree with results of a previous study (Tardy *et al.*, 2015), which observed a large importance of fungal richness using the same substrate. Indeed, on complex substrates such as wheat, the influence of microbial diversity may increase (Waldrop and Firestone, 2004).

However, in our study, fungal communities may have been sensitive to the drying and long storage of soils (Kaisermann *et al.*, 2015). Also, Tardy *et al.* (2015) described mineralisation as the area under the curve of CO₂ emissions, which captured mineralisation dynamics throughout the incubation period. As time is no longer needed in this model of mineralisation, other explanatory covariates can be detected more easily. The time variable was able to explain nearly all the variability in the residue mineralisation rate. In other words, variability in the residue mineralisation rate at each sample time was relatively small (coefficient of variation = 4-14%),

likely due to using only one type of residue, and residue quality has been shown as an important driver of residue mineralisation (Nicolardot *et al.*, 2001).

Control soil mineralisation was explained better by bacterial diversity, while amended soil mineralisation and PE were explained better by fungal diversity. This difference can be explained by the observed activation of the fungal community in amended soil, leading to a PE (Blagodatskaya and Kuzyakov, 2008; Fontaine *et al.*, 2011) through successional changes in microbial community structure, with a gradual increase in the *fungal : bacterial* biomass ratio (Lundquist *et al.*, 1999). However, while the relative importance of microbial diversity was high in the control soil (44%), it tended to be lower in amended soil (4%) and for the PE (9%). Thus, while microbial diversity indexes can explain basal respiration, they are less able to explain respiration occurring from soil amended with an organic substrate or the PE, suggesting that both activities may depend more on microbial composition than on taxonomic richness and evenness (Fontaine *et al.*, 2011; Pascault *et al.*, 2013; Tardy *et al.*, 2015).

When diversity indexes are replaced by abundances of bacterial and fungal phyla, and the same selection procedure is applied, model quality (according to Bayesian Information Criterion (BIC)) increases for the amended soil mineralisation and PE models but decreases for the other two (Annexe C). In the PE model, 24% of the variability was explained by phyla abundances, compared to only 9% when microbial diversity indexes are used. This highlights that processes related to PE may be driven more by microbial composition than diversity indexes.

For control soil mineralisation, bacterial evenness was the microbial diversity index selected. As stated by Hooper *et al.* (Hooper *et al.*, 2005), a relationship between microbial diversity and soil processes can be explained by i) an increased likelihood that key species will be present when richness increases or ii) by an increase in positive interactions (complementarity or facilitation) when diversity increases that increases rates of ecosystem processes. While indexes that include richness (richness, H' and $1/D$) can encompass both hypotheses, evenness considers only the relative abundance of each microbial group, regardless of the number of groups, and can capture only the second hypothesis. This suggests that greater equity in the abundance of microbial groups would increase these positive interactions. However, since all bacterial diversity indexes are strongly correlated with each other (Figure 2.4), it is difficult to ensure that one index selected would have the capacity to explain variability in C dynamics better than the others.

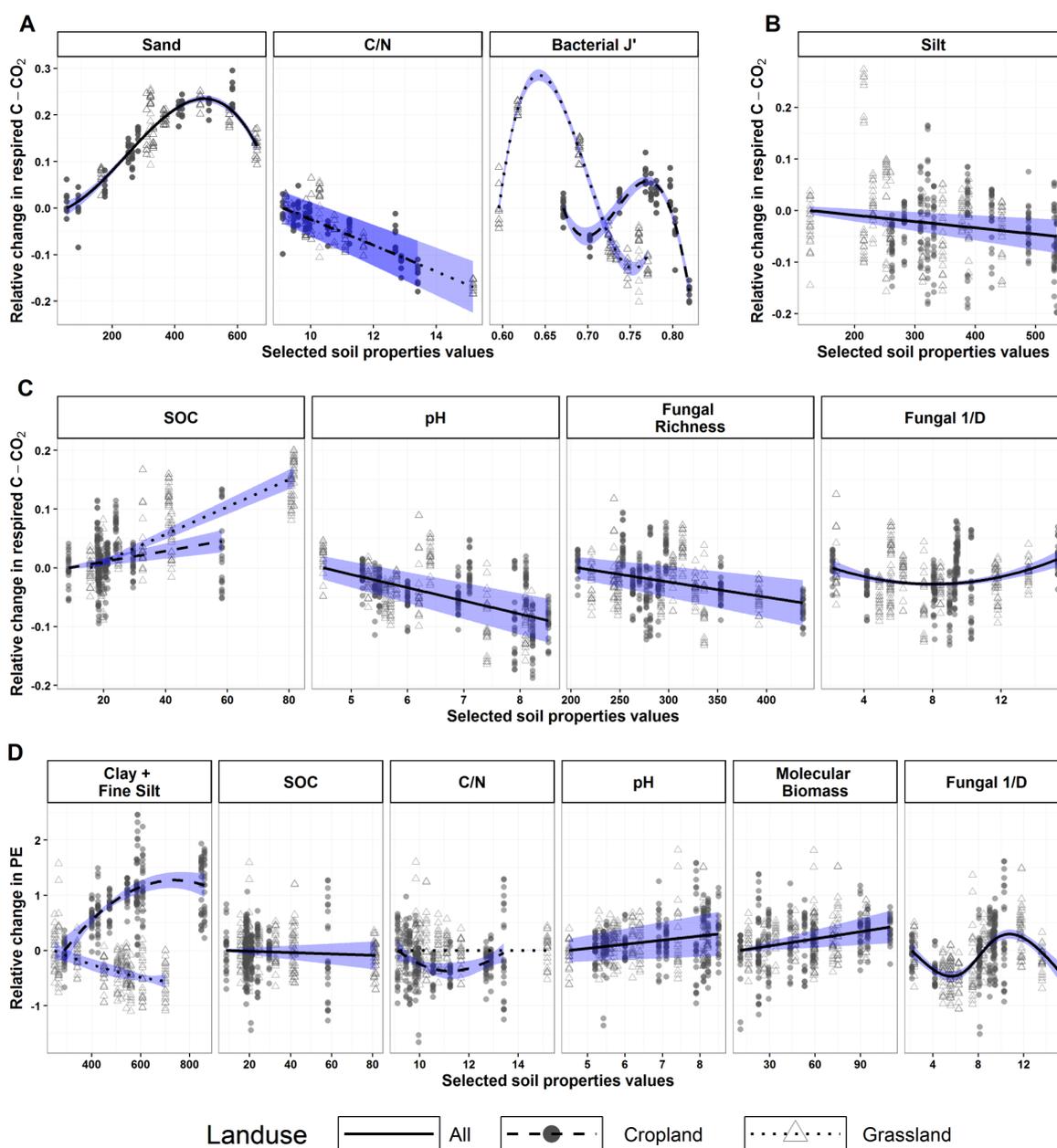


FIGURE 3.4 – Selected soil properties and associated estimated relationships for each mineralisation parameter.

(A) Control soil mineralisation ($R_{s,control}$), (B) residue mineralisation (R_r), (C) amended soil mineralisation ($R_{s,amended}$) and (D) priming effect (PE). The x-axis represents the selected soil property values, and the y-axis (dimensionless) represents the relative change in the mineralisation parameter when selected soil property values vary. Units of soil properties are [$\text{g}\cdot\text{kg}^{-1}$ soil] for texture variables and soil organic carbon (SOC) content, [$\mu\text{g DNA}\cdot\text{g}^{-1}$ soil] for molecular biomass, [number of operational taxonomic units (OTUs)] for fungal richness, [pH unit] for pH and dimensionless for C :N ratio, bacterial evenness (J') and fungal inverse Simpson ($1/D$). To facilitate reading, the relative change in the mineralisation parameter equals zero when the selected soil properties are at their minima. Solid black lines represent estimated relationships between the mineralisation parameter and the selected soil properties, regardless of land use (no interaction). Dashed and dotted black lines represent estimated relationships for cropland and grassland, respectively. Blue areas indicate the 95% confidence interval of the estimated relationships. Dots (cropland) and triangles (grassland) represent partial residuals, i.e. residuals of the models when all selected soil properties are accounted for except for the one considered.

Overall, we showed that microbial diversity indexes can be relevant covariates to include in SOM dynamics models, but that distinctions exist between bacteria and fungi and between different C-substrate pools. Systematically using C-labelled residues and distinguishing between bacterial and fungal diversity in future studies may represent a promising approach to (i) understand the complex relationship between microbial diversity and C mineralisation and (ii) define a suitable strategy for integrating microbial-diversity parameters in SOM modelling. Most current models include at least two compartments to model C dynamics of residues and SOM differently, which improves the prediction of overall C dynamics (Manzoni and Porporato, 2009). Due to differences in physiology, bacterial and fungal communities may influence C dynamics differently, whether individually (Waring *et al.*, 2013) or in interaction (de Boer *et al.*, 2005; Romani *et al.*, 2006).

3.1.4.3 Implications for integrating microbial diversity in SOM dynamics models

Perveen *et al.* (2014) developed a promising SOM dynamics model that represents two functional groups of microorganisms to account for interactions between residue and soil mineralisation (i.e. the PE). One major limitation of this kind of model is the difficulty in identifying groups given the current level of knowledge about the influence of microbial composition on soil processes. Hence, this model remains theoretical until knowledge improves. Since microbial diversity indexes appear to explain soil decomposition/mineralisation well, but composition appears to explain PE better, i) in models integrating processes involved in the PE, microbial composition likely better represents the influence of microbial communities, while ii) in models focusing only on decomposition/mineralisation processes, microbial diversity indexes may represent well the influence of microbial diversity on C dynamics.

Observing relationships between microbial diversity indexes can help to understand the influence of microbial diversity on C dynamics. Interestingly, the relationship between bacterial diversity and control soil mineralisation seemed to depend on land-use history. These relationships for both cropland and grassland were polynomial curves that overlapped little. Grassland soils seemed to be associated with high mineralisation rates and low bacterial evenness, while the opposite was observed for cropland soils. Consequently, land-use history and bacterial evenness appear to have confounding effects. This highlights one limitation of the method : when there is a strong relationship between covariates, it is difficult to know which one influences the other. In this case, it is more likely that land-use management has changed soil properties, such as

bacterial diversity or SOC content, with consequences on C dynamics.

Thus, we assume that a negative relationship exists between bacterial diversity and basal soil mineralisation. This is consistent with the negative relationship between fungal richness and amended-soil mineralisation. Such negative relationships were unexpected, as previous studies demonstrated positive relationships (Baumann *et al.*, 2013; Juarez *et al.*, 2013; Tardy *et al.*, 2014). Nevertheless, the negative relationship does not necessarily contradict these studies, which focused on total flux. Respiration can decrease due to a decrease in the decomposition rate of the substrate or an increase in the substrate-assimilation yield of microorganisms.

Hypotheses exist in the literature about the influence of microbial diversity on C assimilation yield (Six *et al.*, 2006). C assimilation yield represents how microorganisms control the fate of soil C, i.e. the C is either used for bacterial growth and production of microbial products, such as exoenzymes or polysaccharides, or it is mineralised (Schimel and Schaeffer, 2015). As patterns of allocation and production vary among microorganisms, the positive relationship between C-assimilation yield and microbial diversity may be due to the presence of species that efficiently assimilate the C to produce microbial products in species-rich communities. Beyond the individual level, this relationship may also result from an increase in positive interactions between microorganisms that leads to higher C assimilation yield overall : niche partitioning (Hooper *et al.*, 2005), complementarity between enzyme producers and “cheaters” i.e. microorganisms that benefit from decomposition products (Allison, 2005), leading to syntrophic relationships (Burns *et al.*, 2013).

Predictions of SOM dynamics models have been highly sensitive to the C assimilation yield parameter (Six *et al.*, 2006). The hypothesis of increased C assimilation yield with increased microbial diversity is of great interest, as it may help to increase model accuracy by calculating assimilation as a function of soil microbial diversity. Moreover, better knowledge of this relationship could help to understand soil’s response to climate warming, as it depends on C assimilation yield (Allison *et al.*, 2010).

3.1.4.4 Strengths and limitations of the method

As we raised hypotheses from relationships generated by the method, studying relationships between well-known abiotic soil properties and C dynamics was an appropriate way to investigate the method’s ability to build consistent relationships. First, the properties selected were consistent with those the literature. Overall, SOC explained 13% of C mineralisation in amended

soil. It is well established that the amount of SOC drives the amount of soil C which is mineralised (Nicolardot *et al.*, 2001; Fang *et al.*, 2005). SOC content positively affected mineralisation, and the difference between this relationship in croplands and grasslands likely resulted from the difference in the dispersion of SOC content.

In addition to the quantity of C substrates, the quality of SOM, estimated by the C :N ratio, appeared as an important factor controlling soil C mineralisation (Nicolardot *et al.*, 2001). The negative relationship may be related to stoichiometric needs of microorganisms for nutrients and limits to nutrient availability.

Selection of soil pH for amended soil mineralisation and PE confirmed its importance for predicting C mineralisation. Soil pH is a well-known factor influencing C dynamics through its impact on the biomass (Kemmitt *et al.*, 2006), composition (Rousk *et al.*, 2010, 2011), structure and activity (Griffith *et al.*, 1995; Robson, 2012) of soil microbial communities. In other respects, the influence of pH on N dynamics (Parton, 1996; Kader *et al.*, 2013) and other soil chemical properties is also well established (Walse *et al.*, 1998; Kemmitt *et al.*, 2006; Robson, 2012).

Overall, these results demonstrate that the method enabled selecting relationships between mineralisation and certain soil properties already known to influence SOM dynamics. This is a good indicator that the method is able to select drivers of mineralisation and build consistent relationships.

We were unable to interpret some irregular relationships, however, such as the opposite relationships between texture and PE in croplands and grasslands. This difference is likely due to overfitting of data, to which GAM is sensitive, rather than to true biophysical relationships. In a dataset in which little is controlled, confounding among covariates is likely. For example, in the control-soil mineralisation model, one could argue that confounding occurs with SOC content because it is negatively correlated with bacterial evenness (Figure 2.4). Forcing the model to include SOC content produced a model similar to the one for amended-soil mineralisation, but with much lower quality (BIC = -78 instead of -235, results not shown). Following these observations and comments by Burnham and Anderson (Burnham and D.R, 2003), there are three major limitations of the approach used :

1. Results could be in a local minimum : the selection procedure would take too much time to test all possible models and cannot prevent models from falling into local minima. In practice, the model obtained cannot be improved but is not the best of all possible models.
2. There is no such thing as a "best model" : the selection procedure improves models by

optimising a criterion (in this study, the MSEP) at each step. In practice, it is likely that several models have similar criterion values and that the model selected is not significantly better than others. Consequently, models with similar quality may exist.

3. As stated, since relationships were constructed statistically, biophysical interpretations of them may not exist. Consequently, despite the strengths of the method, the relationships identified should be considered with caution. Future studies should confirm or disconfirm these relationships and raised hypotheses should be tested in more experimental studies.

3.1.5 Conclusion

By using an innovative statistical data-mining approach on a dataset combining a broad sample of studied soils with high variability in soil properties and an incubation experiment monitoring C dynamics according to C source, we aimed to i) investigate whether microbial diversity indexes could help explain variability in C dynamics, ii) construct quantitative relationships between these indexes and variables describing C dynamics and iii) assess implications of these relationships for integrating microbial diversity in SOM dynamics models.

Despite some limitations of the approach, by monitoring SOM and residue mineralisation independently, we demonstrated that microbial diversity indexes could constitute good covariates to integrate in SOM dynamics models, depending on the C source and the processes considered. Specifically, microbial diversity indexes can help explain soil decomposition/mineralisation, while the PE seemed to be associated much more with microbial composition. Thus, we suggest two complementary approaches for future research.

First, continue ongoing investigation of relationships between microbial composition, functional traits and soil C dynamics. More detailed knowledge should help to better represent microbial diversity and its role in mechanistic SOM dynamics models, especially those including the PE.

Second, we argue that including microbial diversity indexes in mechanistic models could be as pertinent as including microbial composition. Quantitative relationships provided by our approach can help discover ways to do the latter. We suggest further investigation of the relationship between microbial diversity indexes and C-assimilation yield. One possibility is that the C-assimilation yield parameter could be modulated as a function of microbial diversity indexes in models.

3.2 Contribution temporelle des indices de diversité pour expliquer la variabilité de dynamique du C entre des sols échantillonnés à l'échelle d'un paysage

3.2.1 Introduction

La diversité microbienne du sol est fortement influencée par les propriétés du sol, ainsi que par le mode d'usage de ces sols (Maron *et al.*, 2011). Il est donc possible de disposer de sols ayant des diversités microbiennes non manipulées différentes en échantillonnant des sols aux propriétés et modes d'usage différents. Dans la section 3.1, cela nous a permis d'étudier le lien entre diversité microbienne et dynamique du C en échantillonnant des sols sous cultures et sous prairies à une échelle nationale et de montrer la pertinence des indices de diversité pour expliquer cette dynamique.

D'autres études ont montré ce lien entre diversité microbienne non manipulée et dynamique du C en échantillonnant des sols à l'échelle continentale (Tardy *et al.*, 2015) et/ou sous des modes d'usage contrastés (Lienhard *et al.*, 2013; Tardy *et al.*, 2015). Bien que ces résultats confirment la pertinence d'intégrer des indices de diversité dans les modèles de dynamique du C à des échelles où les facteurs influençant la diversité microbienne sont assez contrastés, la question de cette pertinence pour des sols échantillonnés à des échelles plus locales (e.g. paysage), où leurs propriétés et les modes d'usage sont moins contrastés, conduisant vraisemblablement à des diversités moins contrastées également, doit encore être étudiée.

Dans la section 1.2, nous avons vu comment les modèles mathématiques de dynamique du C différenciaient les pools de MO en compartiments, selon leur composition chimique, conduisant à différentes cinétiques de transformation (Manzoni and Porporato, 2009). Les résultats de la section 3.1 ainsi que d'autres études (e.g. Tardy *et al.*, 2015) ont montré une différence de lien entre diversité microbienne et dynamique du C selon la source du C. Pour l'étude du lien entre diversité microbienne et dynamique du C, une distinction systématique entre MOF et MOS pour l'acquisition de données, e.g. via l'utilisation d'isotope, est nécessaire.

D'autres modèles vont plus loin en subdivisant les pools de MOF et MOS en sous-pools représentant des composés plus ou moins récalcitrants et donc ayant des cinétiques différentes (e.g. Garnier *et al.*, 2001; Moorhead and Sinsabaugh, 2006). Il est souvent souligné que la diversité microbienne est fonctionnellement redondante pour des composés ayant une cinétique de décom-

position rapide (e.g. sucres) mais que des composés spécifiques plus récalcitrants pourraient être décomposés par des espèces clés (Niemenmaa *et al.*, 2008; Baumann *et al.*, 2013). Sous cette hypothèse, le lien entre diversité microbienne et dynamique de décomposition devrait être plus important pour cette deuxième catégorie et donc devrait s'exprimer de manière différente dans le temps, lorsque ces composés ont été suffisamment décomposés.

L'objectif de cette étude est d'examiner la contribution de la diversité des bactéries et des champignons de sols échantillonnés dans un bassin versant de 5km² dans la minéralisation de différents pools de C caractérisée par la cinétique d'émission de CO₂ de la même manière que dans la section 3.1. Quarante-et-un sols ont été incubés, avec et sans paille de blé marquée au ¹³C. La méthode utilisée est la même que la section 3.1, excepté que les cinétiques d'émission de CO₂ ont été découpées en 3 périodes afin d'étudier la dépendance temporelle du lien entre la diversité microbienne et la dynamique de transformation des différents pools de C.

3.2.2 Matériel et méthodes

3.2.2.1 Échantillonnage et analyse des sols

Les sols ont été prélevés dans le bassin versant de Kervidy-Naizin dans le département du Morbihan (56) comme présenté dans la section 2.1. Les sols sélectionnés pour cette analyse sont tous des sols cultivés et soumis à des rotations de cultures variées, présentant ou pas des prairies de plus ou moins longue durée et une variabilité de gestion de la fertilisation. Par rapport au jeu de données de la section 3.1, les mêmes types de propriétés physiques et chimiques sont disponibles auxquelles viennent s'ajouter les teneurs en N minéral et en P assimilable (Olsen) des sols.

La caractérisation de la minéralisation du C dans les sols a été effectuée selon le même protocole que l'étude précédente. Les sols séchés ont été incubés avec (amendé) ou sans (control) paille de blé marquée au ¹³C, après ré-humectation à 60% de la capacité au champ. Les flux de CO₂ marqués et non marqués ont été mesurés à 3, 7, 14, 21, 28, 42, 63 et 80 jours pour déterminer la minéralisation du sol (R_s) et du résidu (R_r) indépendamment. Le PE a été calculé selon les mêmes équations que précédemment (équations 3.1-3.3).

Toujours selon la procédure précédente, la biomasse moléculaire microbienne et les indices de diversité pour les bactéries et les champignons ont été mesurés. Une différence de nombre de séquences utilisées pour regrouper les OTUs empêche cependant une comparaison de ces indices avec ceux de DIMIMOS (cf. section 2.1).

3.2.2.2 Analyses statistiques

La méthode utilisée pour analyser le lien entre les covariables et les variables de dynamique du C est la même que celle présentée dans la sous-section 3.1.2.2 (Figure 3.1). Deux différences majeures sont à noter :

1. Chaque taux de minéralisation ($\Delta R/\Delta t$) est subdivisé en 3 périodes : 3-7 jours, 7-28 jours et 28-80 jours (Figure 3.5). Les variables à analyser caractérisant la dynamique du carbone sont alors les aires sous la courbe des taux sur chaque période (\propto taux moyen sur la période). Pour chaque cinétique de minéralisation (R_r , $R_{s,amended}$, $R_{s,control}$ et PE) et chaque sol, on obtient 3 valeurs d'aire sous la courbe (une par période). Chacune de ces variables a été analysée indépendamment. Cette approche nous permet d'analyser si les relations qui existent entre les covariables et les cinétiques de minéralisation varient avec le temps. En conséquence, la variable temps n'est plus une variable obligatoire du modèle et le modèle le plus simple devient :

$$Y = \mu \tag{3.7}$$

où Y est la variable réponse (i.e. l'aire sous la courbe du taux d'une cinétique de minéralisation sur une période) et μ représente la moyenne de Y . Pour vérifier la pertinence des périodes choisies, une ANOVA est effectuée entre les aires sous la courbe (variable réponse) divisées par le nombre de jours de la période (= taux moyen) et un facteur aléatoire représentant la période.

2. Au vu des propriétés du sol disponibles dans ce jeu de données, les covariables testées sont : les teneurs en argile, en limon, en sable et en argile+limon fin pour la texture, les teneurs en C organique, en N minéral et en P assimilable pour les nutriments du sol ; le pH et la biomasse moléculaire ainsi que la richesse, l'équitabilité (J'), l'indice de Shannon (H') et l'inverse de l'indice de Simpson ($1/D$) à la fois des bactéries et des champignons pour les propriétés biologiques.

3.2.3 Résultats

3.2.3.1 Variabilité temporelle des cinétiques de minéralisation

Pour toutes les cinétiques de minéralisation, les taux moyens de respiration sont significativement différents entre les 3 périodes ($P < 2e-16$, Figure 3.6). Les taux vont dans l'ordre décroissant de la période 3-7 jours à la période 28-80 jours. Le taux de minéralisation du résidu subit les

baisses les plus importantes au cours des périodes (75% puis 90%), suivi de la minéralisation du sol en présence du résidu (40% puis 45%) et des deux autres (24%, 22% puis 27%, 25%, respectivement pour $R_{s,control}$ et PE).

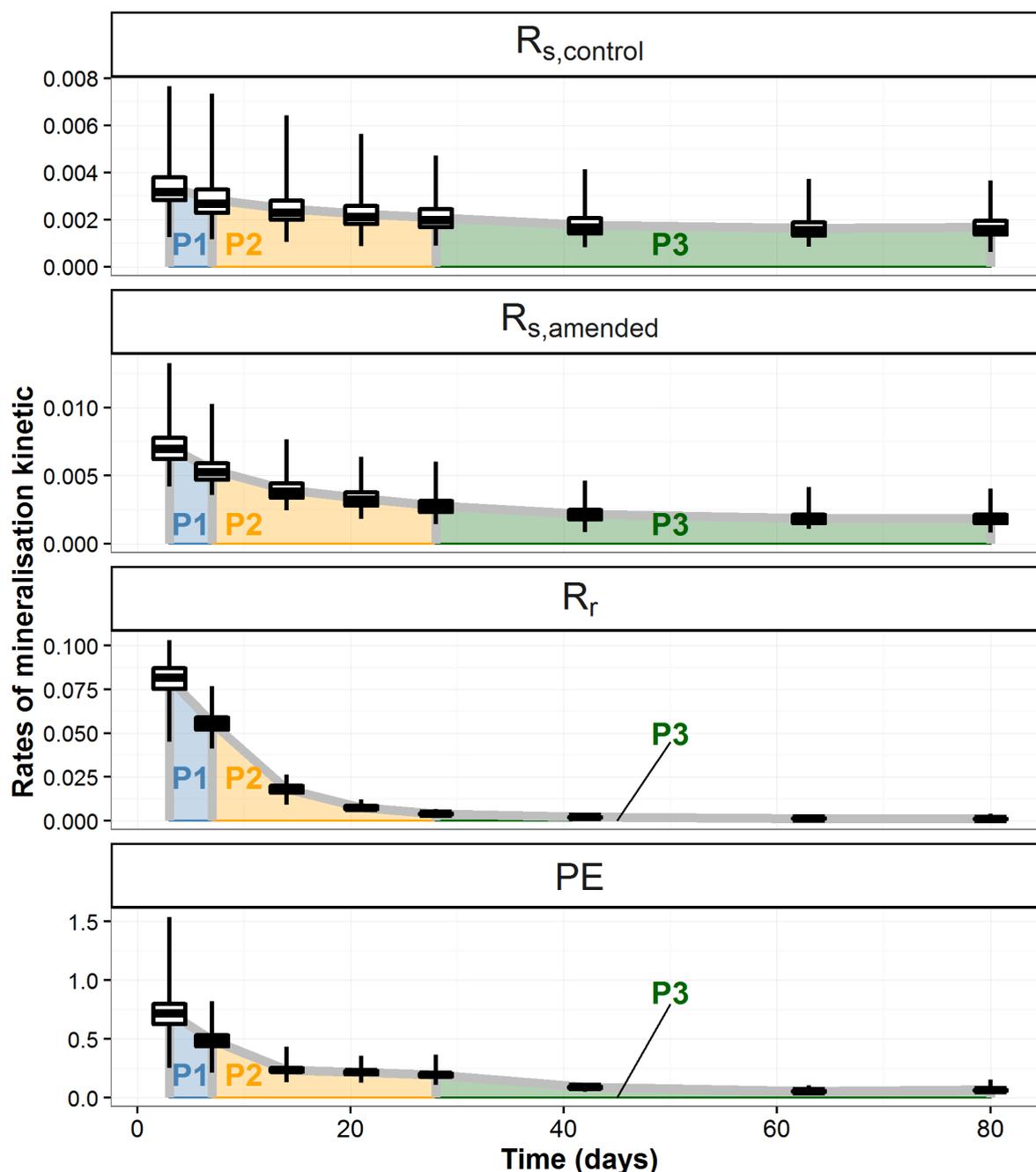


FIGURE 3.5 – Taux de minéralisation en fonction des périodes différenciées des cinétiques.

Les taux de minéralisation du sol, control et amendé ($R_{s,control}$ et $R_{s,amended}$ respectivement) et de minéralisation du résidu (R_r) sont exprimés en $\text{mg C-CO}_2 \text{ g}^{-1} \text{ sol jour}^{-1}$ et en jour^{-1} pour le PE. Les trois périodes sont identifiées par P1 (3-7 jours, bleu), P2 (7-28 jours, orange) et P3 (28-80 jours, vert). Les segments haut et bas des boîtes de dispersion représentent les valeurs maximales et minimales, respectivement.

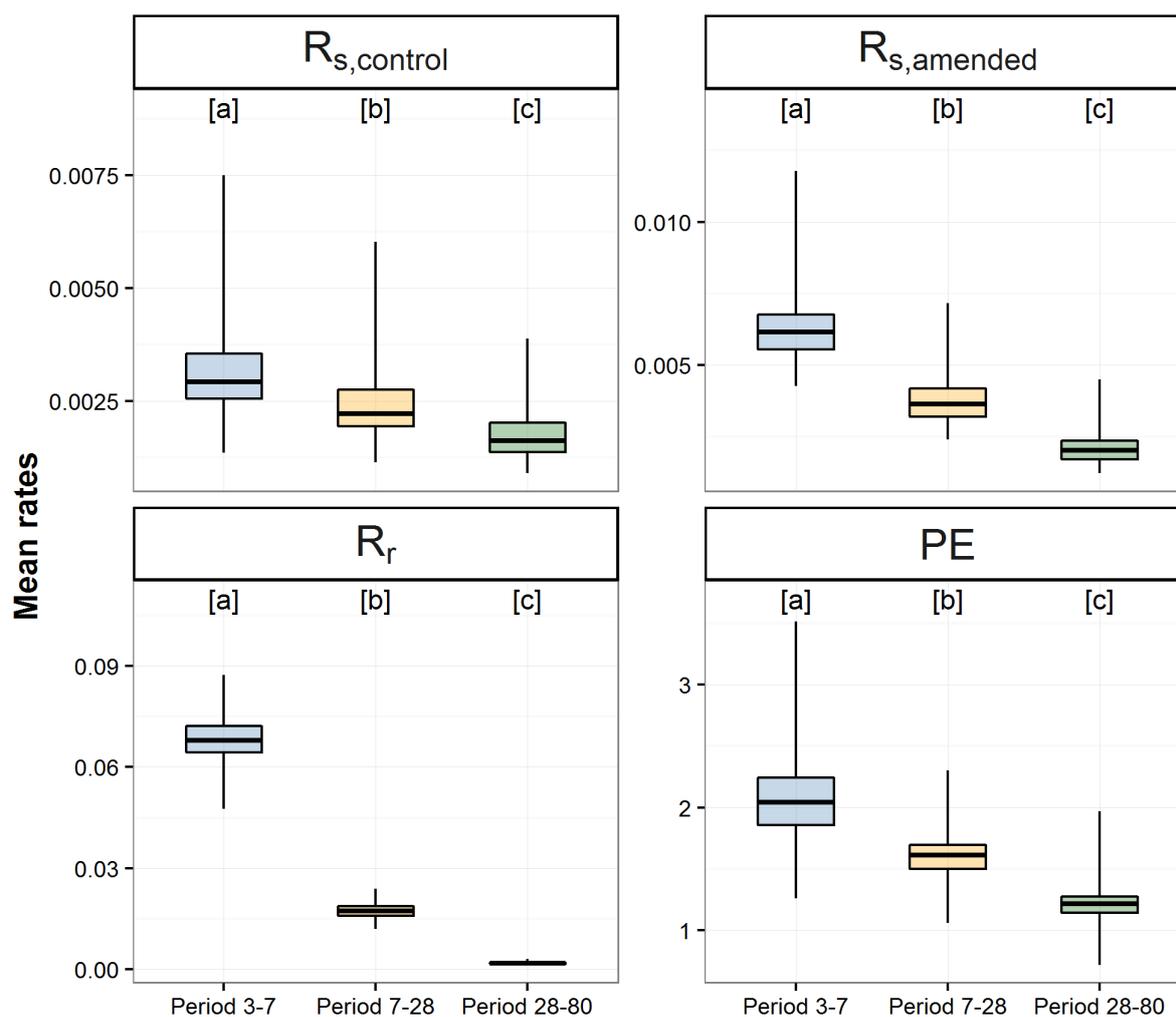


FIGURE 3.6 – Variabilité des taux moyens des cinétiques de minéralisation pour les périodes 3-7 jours (bleu), 7-28 jours (orange) et 28-80 (jours).

Les taux moyens des cinétiques de minéralisation du sol, control et amendé ($R_{s,control}$ et $R_{s,amended}$ respectivement) et de minéralisation du résidu (R_r) sont exprimés en $\text{mg C-CO}_2 \text{ g}^{-1} \text{ sol jour}^{-1}$ et en jour^{-1} pour le PE. Les segments haut et bas des boîtes de dispersion représentent les valeurs maximales et minimales, respectivement. Des lettres différentes au-dessus des boîtes de dispersion signifient une différence significative ($P < 0.05$) des taux d'une même cinétique entre les périodes.

3.2.3.2 Qualité des modèles et relations des cinétiques avec les covariables

Les qualités d'ajustement des modèles sélectionnés dépendent plus des cinétiques de minéralisation que de la période (Table 3.2). Ces qualités sont bonnes pour les cinétiques de minéralisation du sol (45% - 59% de variance expliquée) à moyennes voire basses pour le résidu et le PE (16% - 36% de variance expliquée). Les qualités prédictives sont quant à elles assez basses, les valeurs de RPIQ allant de 1 à 1,8 au maximum. Les variables sélectionnées sont toutes significatives (à

un seuil de significativité de 5% sauf pour 3 qui sont entre 5% et 10%) exceptée la teneur en argile pour le PE en période 28-80 jours.

La figure 3.7 résume l'importance des covariables sélectionnées dans les modèles ainsi que le type de relation obtenue avec les cinétiques de minéralisation. L'ensemble des relations est visible dans les annexes D.1 à D.4. Les covariables sélectionnées pour les cinétiques de minéralisation du sol ($R_{s,control}$ et $R_{s,amended}$) sont assez similaires. Les variables les plus importantes sont la teneur en carbone organique (2-12%) et la biomasse microbienne (7-20%) avec une relation positive.

Le ratio C/N est important pour les sols témoins uniquement (3-8%) avec une relation négative. Le P assimilable est très important seulement pour la période intermédiaire pour le sol témoin (30%), et pour toutes les périodes lorsque le sol est amendé (20-28%) avec une relation négative. Les relations avec la texture sont également similaires (1-5%) entre la minéralisation du C dans les sols témoins et amendés, et le type de relation dépend de la fraction granulométrique considérée. Globalement, la relation est négative avec les fractions les plus fines.

Les relations avec la diversité bactérienne sont négatives (2-6%) et sont valables pour toutes les périodes pour les sols témoins, et seulement pour les deux dernières pour les sols amendés. Pour la diversité en champignons, il existe une relation positive (2-5%) également pour toutes les périodes dans les sols témoins, et seulement pour les deux dernières pour les sols amendés.

La variable la plus importante pour le flux issu du résidu est le pH, avec une relation négative au début (7%) qui devient positive et de plus en plus importante au cours des deuxième et troisième périodes (2% puis 24%). Une autre variable importante est la diversité en champignons (13%) avec une relation positive mais seulement pour la période intermédiaire. On peut noter également une relation positive avec la covariable qui somme la teneur en argile et en limon fin pour les deux dernière périodes (4-9%).

Pour le PE, les variables les plus importantes se trouvent dans la catégorie des nutriments du sol, en particulier le P assimilable pour les deux premières périodes (11-12%) avec une relation positive et le ratio C/N à la dernière période (6%) avec une relation positive. De plus, seules des covariables concernant la diversité bactérienne ont été sélectionnées pour cette cinétique avec des relations positives à variables (1-7%).

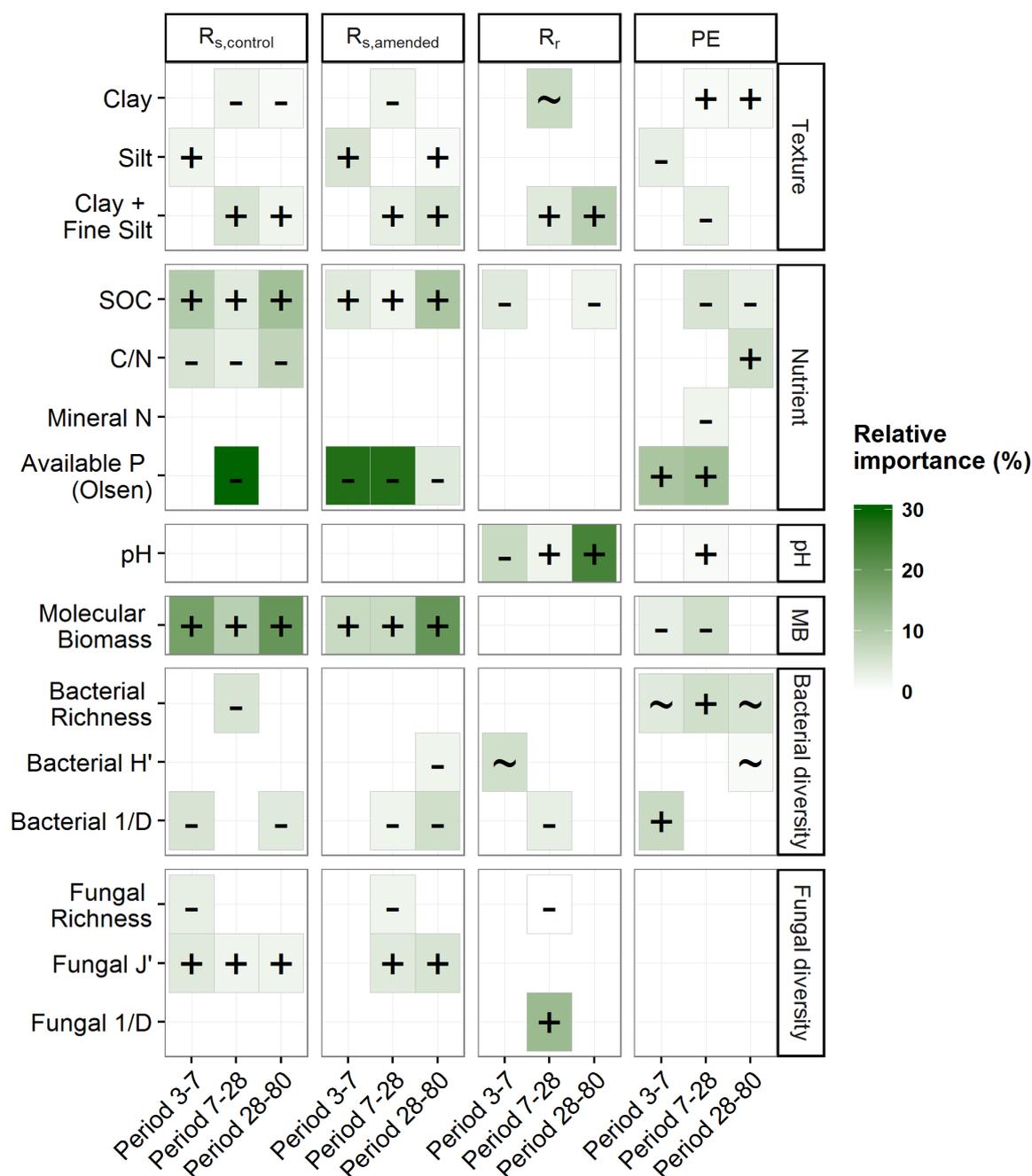


FIGURE 3.7 – Importance des covariables sélectionnées par période pour les cinétiques de minéralisation du sol témoins ($R_{s,control}$) et amendé ($R_{s,amended}$), de minéralisation du résidu (R_r) et du priming effect (PE).

La présence d'un carré vert et d'un signe signifie que la variable a été sélectionnée. Plus la couleur est sombre plus la part de variance expliquée par la variable est importante. Le signe correspond au type de relation sélectionnée entre la covariable et la cinétique de minéralisation : positive (+), négative (-) ou variable (~). Une relation variable correspond à une relation qui ne peut être qualifiée de positive ou négative car plus complexe (e.g. optimum, sinusoidale). **SOC**, teneur en carbone organique du sol; **C/N**, ratio entre le carbone organique et l'azote total du sol; **N**, azote; **P**, Phosphore; **H'**, indice de Shannon; **J'**, Équitabilité; **1/D**, inverse de l'indice de Simpson.

TABLE 3.2: Évaluation des modèles et des covariables sélectionnées pour chaque cinétique de minéralisation.

| Kinetic | Period | VarExp ^a | RPIQ | Covariates | RelImp ^b | Significance ^c |
|------------------------------|--------|---------------------|------|---------------------|---------------------|---------------------------|
| R_{s,control} | 0-7 | 47 | 1.6 | Silt | 2 | ** |
| | | | | SOC | 10 | *** |
| | | | | C :N ratio | 5 | *** |
| | | | | Molecular biomass | 18 | *** |
| | | | | Bacterial 1/D | 5 | ** |
| | | | | Fungal Richness | 3 | *** |
| | | | | Fungal J' | 4 | *** |
| | 7-28 | 59 | 1.8 | Clay | 2 | *** |
| | | | | Clay + Fine Silt | 5 | ** |
| | | | | SOC | 4 | *** |
| | | | | C :N ratio | 3 | *** |
| | | | | Available P (Olsen) | 30 | *** |
| | | | | Molecular biomass | 10 | *** |
| | | | | Bacterial Richness | 5 | ** |
| | 28-80 | 49 | 1.8 | Clay | 1 | ° |
| | | | | Clay + Fine Silt | 2 | ** |
| | | | | SOC | 12 | *** |
| | | | | C :N ratio | 8 | *** |
| | | | | Molecular biomass | 20 | *** |
| R_{s,amended} | 0-7 | 45 | 1.5 | Silt | 5 | ** |
| | | | | SOC | 4 | *** |
| | | | | Available P (Olsen) | 28 | *** |
| | | | | Molecular biomass | 7 | *** |
| | 7-28 | 49 | 1.7 | Clay | 2 | * |
| | | | | Clay + Fine Silt | 3 | ** |

Évaluation des modèles et des covariables (suite)

| Kinetic | Period | VarExp ^a | RPIQ | Covariates | RelImp ^b | Significance ^c |
|----------------------|--------|---------------------|------|---------------------|---------------------|---------------------------|
| | | | | SOC | 2 | ** |
| | | | | Available P (Olsen) | 28 | *** |
| | | | | Molecular biomass | 7 | *** |
| | | | | Bacterial 1/D | 2 | * |
| | | | | Fungal Richness | 2 | ** |
| | | | | Fungal J' | 4 | *** |
| | 28-80 | 53 | 1.7 | Silt | 1 | * |
| | | | | Clay + Fine Silt | 5 | * |
| | | | | SOC | 11 | *** |
| | | | | Available P (Olsen) | 4 | *** |
| | | | | Molecular biomass | 20 | *** |
| | | | | Bacterial H' | 2 | *** |
| | | | | Bacterial 1/D | 6 | ** |
| | | | | Fungal J' | 5 | ** |
| R_r | 0-7 | 17 | 1.4 | SOC | 4 | *** |
| | | | | pH | 7 | *** |
| | | | | Bacterial H' | 6 | * |
| | 7-28 | 28 | 1.6 | Clay | 7 | ** |
| | | | | Clay + Fine Silt | 4 | ** |
| | | | | pH | 2 | *** |
| | | | | Bacterial 1/D | 3 | * |
| | | | | Fungal Richness | <1 | *** |
| | | | | Fungal 1/D | 13 | *** |
| | 28-80 | 36 | 1.5 | Clay + Fine Silt | 9 | *** |
| | | | | SOC | 2 | *** |
| | | | | pH | 24 | *** |
| PE | 0-7 | 29 | 1.4 | Silt | 3 | *** |
| | | | | Available P (Olsen) | 11 | *** |
| | | | | Molecular biomass | 3 | *** |

Évaluation des modèles et des covariables (suite)

| Kinetic | Period | VarExp ^a | RPIQ | Covariates | RelImp ^b | Significance ^c |
|---------|--------|---------------------|------|---------------------|---------------------|---------------------------|
| | | | | Bacterial Richness | 4 | *** |
| | | | | Bacterial 1/D | 7 | ** |
| | 7-28 | 36 | 1.2 | Clay | <1 | ° |
| | | | | Clay + Fine Silt | 3 | ° |
| | | | | SOC | 5 | *** |
| | | | | Mineral N | 2 | * |
| | | | | pH | 1 | * |
| | | | | Available P (Olsen) | 12 | ** |
| | | | | Molecular biomass | 6 | *** |
| | | | | Bacterial Richness | 6 | *** |
| | 28-80 | 16 | 1 | Clay | <1 | ns |
| | | | | SOC | 3 | *** |
| | | | | C :N ratio | 6 | *** |
| | | | | Bacterial Richness | 5 | ** |
| | | | | Bacterial Shannon | 1 | ** |

RPIQ, Ratio entre l'espace inter-quartile et la racine carré de l'erreur quadratique moyenne de prédiction; **SOC**, Carbone Organique du Sol; **J'**, équitabilité; **H'**, indice de Shannon; **1/D**, inverse de l'indice de Simpson

^a Pourcentage de variance expliquée par le modèle

^b Significativité des covariables sélectionnées d'après l'ANOVA : *** $P < 0.001$; ** $0.001 < P < 0.01$; * $0.01 < P < 0.05$; ° $0.05 < P < 0.1$; ns (non significative) $P > 0.1$

^c Importance Relative (%) des covariables sélectionnées d'après la partition de variance

3.2.4 Discussion

3.2.4.1 Confirmation de la pertinence des indices de diversité microbienne pour expliquer la dynamique du C dans les sols

Des indices de diversité microbienne ont été sélectionnés pour toutes les cinétiques de minéralisation étudiées. Ces indices semblent importants dans la minéralisation du C du sol amendé et non amendé. Une relation négative a été trouvée avec la diversité bactérienne. Ce résultat

vient donc confirmer les résultats de DIMIMOS où une relation négative a été trouvée, mais qui souffrait d'une confusion entre les valeurs de la diversité bactérienne et le mode d'usage. Cela donne donc plus de poids à l'hypothèse soulevée d'une amélioration du rendement d'utilisation du C avec une augmentation de la diversité bactérienne (Six *et al.*, 2006).

Cette amélioration serait due soit à une augmentation de la présence d'espèces qui assimilent efficacement le C, i.e. dont la proportion de C assimilé réellement utilisée pour la croissance est plus importante (moins de minéralisation) (Schimel and Schaeffer, 2015), soit à une augmentation d'interactions positives qui augmentent la CUE à l'échelle de la communauté (Hooper *et al.*, 2005), soit aux deux. Les indices de diversité bactérienne sélectionnés sont l'inverse de l'indice de Simpson ($1/D$) pour les première et dernière périodes et l'indice de Shannon (H') pour la période intermédiaire. Ces deux indices sont très positivement corrélés (coefficient de corrélation de Spearman = 0,83, Figure 2.5) donc probablement interchangeables. Ils sont également très corrélés avec l'indice d'équitabilité (J') sélectionné pour le jeu de données DIMIMOS (coefficient = 0,94 et 0,87, respectivement, Figure 2.4). Le choix de l'indice a donc peu d'importance à partir du moment où il intègre de l'information sur l'abondance relative des OTU (i.e. équitabilité J' , indice de Shannon H' ou inverse de l'indice de Simpson $1/D$).

Une relation négative a également été trouvée entre les cinétiques de minéralisation du C du sol et la richesse en champignons en accord avec de précédentes études (Tardy *et al.*, 2015) et les résultats pour les sols amendés des données DIMIMOS. Cette sélection est cependant plus ponctuelle (période 3-7 jours pour les sols témoins et période 7-28 jours pour les sols amendés) et la même hypothèse peut être avancée, suggérant une contribution de la diversité en champignons dans l'amélioration de la CUE mais plus ponctuellement par rapport aux bactéries.

Cette amélioration pourrait venir d'une possible interaction entre bactéries et champignons : les champignons, plus efficaces pour utiliser des composés complexes, permettraient d'augmenter l'accessibilité de ces composés aux bactéries (de Boer *et al.*, 2005). Ceci expliquerait la sélection d'indices de diversité des champignons seulement à la première période (3-7 jours et 7-28 jours respectivement pour les sols témoins et amendés), et ensuite la sélection d'indices de diversité des bactéries qui utilisent des composés dont l'accessibilité a été facilitée par les champignons.

Une relation positive a également été trouvée entre l'équitabilité en champignons et les cinétiques de minéralisation du C du sol amendé ou non. Cette relation n'était pas présente dans DIMIMOS pour lequel les champignons étaient probablement affectés par les conditions de stockage (section 3.1). Une relation positive entre l'inverse de l'indice de Simpson des champignons

(intégrant richesse et l'abondance) et la cinétique de minéralisation du résidu pour la période 7-28 jours a également été trouvée. Une telle relation positive entre dynamique du C, qu'il vienne du sol ou du résidu, et diversité en champignons a déjà été observée (Baumann *et al.*, 2013; Tardy *et al.*, 2015) et est interprétée comme une augmentation de la décomposition avec la diversité en champignons, du fait de leur plus grande capacité à dégrader des substrats complexes comme la MOS ou la paille de blé (de Boer *et al.*, 2005). Au vu des indices sélectionnés, cette relation positive semble plus pertinente avec des indices intégrant de l'information sur l'abondance relative des OTU (J' , H' ou $1/D$).

Seuls des indices de diversité bactérienne ont été sélectionnés pour le PE avec des relations positives à variables (de type quadratique). Les relations positives sont avec l'inverse de l'indice de Simpson à la période 3-7 jours et avec la richesse à la période 7-28 jours. Pour ces indices et à ces périodes, des relations négatives ont été sélectionnées dans les sols témoins mais pas dans les sols amendés. L'interprétation de ces relations réside donc probablement dans le calcul même du PE qui est le ratio entre la minéralisation dans les sols amendés et les sols témoins. Les relations étant négatives dans les sols témoins et inexistantes dans les sols amendés, il est logique de trouver des relations positives dans le PE.

Les relations variables sont interprétées de la même manière et montrent la difficulté d'isoler les effets sur le PE des effets sur la minéralisation du sol. Cependant, l'absence d'observation de relations franches entre le PE et les indices de diversité microbienne est en accord avec l'hypothèse que le PE serait plus influencé par la composition des communautés microbiennes que par la richesse ou l'équitabilité (Fontaine *et al.*, 2011; Pascault *et al.*, 2013; Tardy *et al.*, 2015, section 3.1)

3.2.4.2 Contribution temporelle de la diversité microbienne

Les contributions négatives de l'inverse de l'indice de Simpson des bactéries et positive de l'équitabilité en champignons sur la minéralisation du C organique du sol se retrouvent à toutes les périodes dans les sols témoins, suggérant que cette contribution est constante dans le temps. Après ajout de la paille de blé, cette contribution n'apparaît cependant qu'à la deuxième période et est maintenue ensuite. Une baisse de la diversité a déjà été observée dans des sols juste après l'ajout de MOF (Bernard *et al.*, 2007). Cette diminution est due à la croissance rapide de microorganismes compétitifs pour la décomposition de la MOF (Fontaine *et al.*, 2003), devenant donc dominants (Bernard *et al.*, 2007), ce qui pourrait expliquer l'absence de contribution de la

diversité lors de la première période. À partir de la deuxième période, la croissance des organismes à croissance plus lente permet un retour de la diversité comparable à celle de départ, ce qui expliquerait le retour de sa contribution à la minéralisation du sol.

Cette diminution de la diversité pourrait expliquer aussi la temporalité observée pour la minéralisation du résidu. La contribution positive de l'inverse de l'indice de Simpson des champignons n'est observée qu'en période 7-28 jours. Une autre hypothèse est que la diversité en champignons ne serait importante que pour des composés complexes, se décomposant plus lentement et dont la contribution dans les émissions de CO₂ n'est pas visible en début d'incubation.

De plus en plus de modèles représentent la MOF en plusieurs pools ayant des vitesses de décomposition de plus en plus lentes : sucre, (hémi-)cellulose et lignine (Hansen *et al.*, 1991; Garnier *et al.*, 2001; Moorhead and Sinsabaugh, 2006). La diversité étant fonctionnellement redondante pour des composés simples comme les sucres, sa contribution au début d'incubation est nulle et ne s'exprime que plus tard, pour la décomposition des composés plus récalcitrants qui pourraient avoir besoin d'espèces clés (Niemenmaa *et al.*, 2008; Baumann *et al.*, 2013). L'absence de la diversité en troisième période pourrait dans ce cas être expliquée par un temps d'incubation trop court, où les composés vraiment plus récalcitrants n'ont pas été assez décomposés pour avoir une contribution significative dans la variabilité des émissions de CO₂.

3.2.4.3 Importance des propriétés physiques et chimiques

Malgré les limites de la méthode (section 3.1.4.4), les relations entre minéralisation et propriétés physiques et chimiques du sol sont cohérentes avec celles observées dans d'autres études. Tout comme pour le jeu de données DIMIMOS, l'importance positive de la teneur en C organique ainsi que la contribution négative de la qualité de la MOS (caractérisée par le ratio C/N) n'est plus à démontrer (Nicolardot *et al.*, 2001; Fang *et al.*, 2005). Le ratio C/N, de même qu'avec les données DIMIMOS, n'a une importance que pour les sols témoins. En présence du résidu, un C/N élevé ne semble donc plus être limitant pour la décomposition de MOS. Dans cette étude, on trouve également une grande importance positive de la biomasse microbienne suggérant une meilleure prise en compte nécessaire de ce paramètre dans les modèles (Wutzler and Reichstein, 2008; Manzoni and Porporato, 2009).

Il existe également une forte contribution négative entre le P assimilable et les émissions de CO₂ due à la minéralisation du sol en présence et absence de résidu. Cette relation a déjà été observée et pourrait être le résultat d'une diminution de l'activité microbienne avec l'augmentation

du P assimilable (Song *et al.*, 2011). Malgré une relation négative seulement à la période 7-28 jours dans les sols témoins et à toutes les périodes dans les sols amendés, il en résulte une relation positive pour le PE aux deux premières périodes. Il semble donc que malgré une diminution globale de l'activité de décomposition du carbone du sol avec le P assimilable, cette disponibilité du P permet une augmentation du PE. Cela montre bien la nécessité de comprendre le couplage entre les nutriments du sol pour réellement comprendre la dynamique d'un seul nutriment.

Dans ce jeu de données, les différences de texture entre les sols sont faibles et expliquent la faible importance de ce paramètre. Malgré tout, il existe une relation négative entre la fraction fine et les cinétiques liées à la minéralisation du sol. En particulier, la relation linéaire négative observée entre la teneur en argile et la décomposition du C du sol jusqu'à la teneur maximale de 250 mg/g sol est en accord avec celle utilisée pour ajuster le paramètre de décomposition dans certains modèles (e.g. Hansen *et al.*, 1991).

Enfin, pour la décomposition du résidu, le pH ressort comme un paramètre important. La relation est d'abord négative et devient positive aux deuxième et troisième périodes. Alors que la relation négative est surprenante, la relation positive est en accord avec les relations observées et utilisées dans les modèles et s'explique par une plus grande activité microbienne et enzymatique lorsque le pH augmente (Leifeld *et al.*, 2008).

3.2.5 Conclusion

L'utilisation d'un jeu de données utilisant des sols échantillonnés à l'échelle d'un paysage a permis de confirmer l'intérêt de la diversité microbienne pour expliquer la dynamique du carbone dans les sols. En particulier, cette diversité semble influencer le rendement d'assimilation du C du sol lors de la décomposition de MO du sol, mais également la décomposition de certains composés d'un résidu complexe comme le blé. Tout comme pour l'étude avec les données DIMIMOS, les relations trouvées ont besoin d'être confirmées par des expérimentations où les facteurs potentiellement influençant la dynamique sont plus contrôlés, pour éviter d'éventuelle confusion entre eux.

Sans oublier l'importance des covariables physiques et chimiques plus classiques, cette étude, avec d'autres de la littérature, vient confirmer la pertinence des indices de diversité microbienne pour expliquer la dynamique du C du sol et suggérer leur intégration dans des modèles mathématiques de dynamique du carbone.

3.3 Discussion de l'implication des résultats quant à l'intégration des indices de diversité dans un modèle de dynamique du C

3.3.1 Comparaison des résultats DIMIMOS et MOSAIC

Le tableau 3.3 résume et compare les résultats de sélection de covariables pour expliquer les cinétiques de minéralisation entre les 2 jeux de données. En observant les importances relatives calculées pour chacune des covariables, elles ont été classées en trois catégories d'importance globale décroissante. Ce classement est uniquement valable pour les données disponibles et dans les gammes de variabilité des covariables. Une généralisation de ces valeurs d'importance demanderait de comparer ces résultats à d'autres analyses.

Les variables ayant le plus d'importance pour la transformation du C de la MOS sont la teneur en C organique, la biomasse microbienne et le P assimilable (non testée dans DIMIMOS). Les deux premières sont des variables classiques des modèles de dynamique de C et, étant dynamiques, sont généralement directement prises en compte dans les équations comme présenté en section 1.2.1 (Wutzler and Reichstein, 2008; Manzoni and Porporato, 2009). La relation négative entre le P assimilable et les émissions issues de la minéralisation du sol a été interprétée comme une diminution de l'activité enzymatique microbienne avec l'augmentation du P assimilable (Song *et al.*, 2011). Le P est rarement pris en compte dans les modèles de dynamique de la MO (Manzoni and Porporato, 2009) et son importance ici plaide pour un couplage plus important de ce nutriment avec le C.

De manière moins importante, mais non négligeable, le ratio C/N et la diversité bactérienne sont ressortis dans les deux jeux de données. Cependant, l'influence du ratio C/N n'a été détecté qu'en l'absence du résidu de culture. La plupart du temps, le ratio C/N du sol est pris en compte dans les modèles lorsqu'ils modélisent la dynamique du C et N (Manzoni and Porporato, 2009). Les mécanismes en jeu pour expliquer le lien entre les dynamiques de ces deux nutriments sont variables mais correspondent à une prise en compte des relations stœchiométriques entre les besoins des micro-organismes pour ces deux nutriments et ce qui est disponible dans le sol.

Lorsque les besoins ne sont pas remplis, en défaveur de l'azote, soit il existe une limitation de la décomposition, soit les micro-organismes rejettent le surplus de C sous forme de CO₂ et de polysaccharides (« C-overflow ») (Manzoni and Porporato, 2009). Dans le deuxième cas, on observerait une augmentation de la respiration des micro-organismes et donc une augmentation des émissions de CO₂ avec l'augmentation du ratio C/N. La relation négative observée à la fois

pour DIMIMOS et pour MOSAIC suggère donc que l'on se place dans le premier cas et que l'augmentation du C/N inhibe la décomposition de la MOS.

Pour la diversité bactérienne, comme expliqué dans les deux sections précédentes, l'hypothèse soulevée est un meilleur rendement de l'assimilation du C avec l'augmentation de cette diversité. Cette relation n'est pas observée dans la première période d'incubation de MOSAIC, probablement en raison d'une baisse de la diversité durant cette période (Bernard *et al.*, 2007), mais du fait de son importance dans DIMIMOS où l'analyse a été effectuée sur la totalité de l'incubation, l'influence temporelle est considérée comme négligeable.

L'hypothèse d'un meilleur rendement d'assimilation du C avec une l'augmentation de la richesse en champignons convient également, mais cette importance est beaucoup plus faible. La diversité en champignons, caractérisée par l'équitabilité, a également une influence positive faible sur les émissions de CO₂, interprétée comme une augmentation de la décomposition avec la diversité fongique.

La dernière covariable sélectionnée est la texture. La relation positive avec la teneur en sable dans DIMIMOS, très fortement corrélée négativement avec la teneur en argile (coefficient se Spearman = -0,74) et la relation négative avec la teneur en argile dans DIMIMOS montre donc une relation négative de la décomposition avec la teneur en fraction fine. Cette variable est souvent prise en compte dans les modèles dans des fonctions qui viennent diminuer la décomposition avec l'augmentation en fraction fine (teneur en argile, e.g. Hansen *et al.*, 1991). L'importance faible de ce paramètre est principalement due aux données MOSAIC qui ne capturent qu'une gamme restreinte de valeurs de cette variable.

Beaucoup moins de covariables ont été sélectionnées pour la décomposition du C du résidu. Les principales sont le pH, dont l'augmentation permet une augmentation de l'activité enzymatique des micro-organismes et donc de la décomposition (Leifeld *et al.*, 2008). L'autre est la diversité en champignons prenant en compte l'équitabilité au même titre que pour la MOF, mais cette fois avec une importance assez élevée. Cependant, elle ne concerne pas les composés rapidement décomposables du résidu (e.g. sucre) car son influence n'est pas visible en début d'incubation.

Pour le PE, les variables les plus importantes sont celles qui concernent les nutriments du sol (C,N,P). Les données utilisées ne mesurant pas la dynamique de N et de P, il n'est pas possible d'aller plus loin dans l'explication des mécanismes du PE liés à ces nutriments. Ces résultats suggèrent cependant qu'une modélisation réussie du PE passera par un meilleur couplage des

dynamiques C, N et P dans le sol. Pour ce qui est de la diversité microbienne, les résultats obtenus ne permettent pas de dégager des interprétations claires de l'influence des indices de diversité sur le PE. Cela vient confirmer que ces indices, représentant la richesse et l'équitabilité, ne suffisent pas à expliquer le PE qui serait plus influencée par la composition des communautés microbiennes.

TABLE 3.3 – Comparaison et interprétation des liens entre covariables et dynamique du C entre DIMIMOS et MOSAIC.

| | Covariables | DIMIMOS | MOSAIC | Interprétation | Importance ^a |
|------------|-----------------------|---|--|---|-------------------------|
| MOS | Biomasse microbienne | ∅ | + | Pris en compte dans les équations de flux | Forte |
| | C organique | + (amendés) | + | Pris en compte dans les équations de flux | Forte |
| | Disponibilité P | NT | - | Baisse de la décomposition | Forte |
| | C/N | - (control) | - (control) | Baisse de la décomposition | Moyenne |
| | Diversité bactérienne | - | - | Augmentation du rendement d'assimilation | Moyenne |
| | Diversité fongique | ∅ | Équitabilité + | Augmentation de la décomposition | Faible |
| | | Richesse - | Richesse - | Augmentation du rendement d'assimilation | Faible |
| Texture | Sable + | Argile - | Baisse de la décomposition avec la fraction fine | Faible | |
| MOF | pH | ∅ | + | Augmentation de la décomposition | Forte |
| | Diversité fongique | ∅ | + (Période 7-28 jours) | Augmentation de la décomposition du pool de récalcitrance intermédiaire | Forte |
| PE | Nutriment | Forte importance des nutriments du sol C,N,P | | | Forte |
| | Diversité microbienne | Composition plutôt que richesse et équitabilité | | | ? |

MOS, Matière Organique du Sol; **MOF**, Matière Organique Fraiche; **PE**, Priming Effect; **NT**, Non Testé; +/-/∅, contribution positive/négative/nulle

^a Le niveau d'importance est une interprétation libre en fonction des importances relatives calculées dans les deux jeux de données

3.3.2 Discussion sur les conséquences pour l'étape de modélisation

L'interprétation des relations entre les covariables et les dynamiques du C dans le sol nous permet de dresser une stratégie d'intégration de la diversité microbienne dans un modèle de dynamique du C. Les résultats sur le PE n'ayant pas aboutis à des interprétations précises du

lien entre les covariables de diversité et ce mécanisme de la dynamique du C, il ne sera pas considéré dans la suite. Du fait qu'il existe tout de même une différence entre la minéralisation de la MOS en présence ou non de résidu, les dynamiques du C dans les sols témoins et les sols amendés seront analysées, mais indépendamment.

L'idée générale est de prendre en compte les covariables dans des fonctions qui viennent moduler la valeur des paramètres de décomposition/minéralisation comme présenté dans l'équation 1.4 de la section 1.2.3. Ces fonctions sont généralement comprises entre 0 et 1 et sont prises en compte de manière multiplicative, faisant donc l'hypothèse, en 1^{ère} approche, qu'il n'existe pas d'interactions entre l'effet des covariables. Prendre en compte des interactions compliqueraient fortement la construction de ces fonctions.

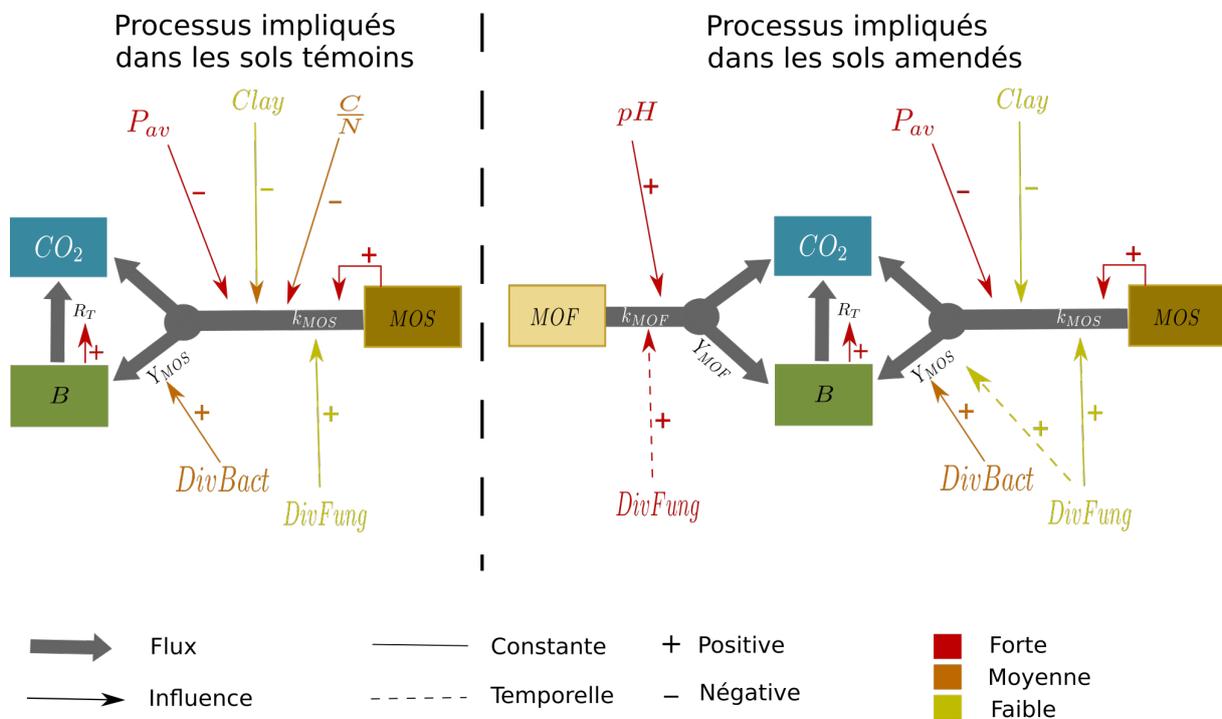


FIGURE 3.8 – Influence des différentes propriétés du sol dans un modèle simple de dynamique du C dans des sols témoins et des sols amendés.

MOS, Matière Organique du Sol; **MOF**, Matière Organique Fraîche; **B**, Biomasse microbienne; k_i , constante de décomposition du pool i ; Y_i , Rendement d'assimilation par les micro-organismes du C pris du pool i ; R_T , Respiration microbienne autre que la respiration de croissance (e.g. maintenance) P_{av} , Teneur en phosphore disponible; **Clay**, Teneur en argile; **C/N**, ratio C et N du sol; **DivBact**, Diversité bactérienne; **DivFung**, Diversité en champignons.

En l'absence de PE, les mécanismes considérés sont la décomposition de la MO, contrôlée par un paramètre de décomposition k ; l'assimilation par les micro-organismes du C décomposé, contrôlée par un paramètre de rendement d'assimilation Y ; et la respiration des micro-organismes autre que pour la croissance, contrôlée par un paramètre de respiration R_T . La figure 3.8 sché-

matise ces flux (décomposition, assimilation, minéralisation) et montre l'influence des différentes covariables les plus importantes sur les paramètres ad-hoc lorsque les sols sont amendés ou non amendés (témoins), par rapport aux interprétations faites dans la section précédente. L'influence de la diversité en champignons sur le paramètre Y_{MOS} n'est pas représentée car elle est supposée négligeable.

3.3.2.1 Intégration de la diversité microbienne

TABLE 3.4 – Fonctions de modulation possibles des paramètres d'un modèle de dynamique du C dans le sol.

| Covariable | Paramètres ajustés | Fonction possible | Référence |
|------------|------------------------|--|-----------------------------|
| Clay | k_{MOS} | $f_{clay} = \begin{cases} 0,5 & Clay > 250g \cdot kg^{-1} \\ 1 - 2.10^{-3} \times Clay & sinon \end{cases}$ | Hansen et al., 1991 |
| C/N | k_{MOS} | $f_{CN} = \begin{cases} 1 & \frac{C}{N} < \frac{b}{Y_{MOS}} \\ \frac{b}{Y_{MOS} \times C/N} & sinon \end{cases}$ | Manzoni and Porporato, 2009 |
| DivBact | Y_{MOS} | $f_{bact} = \frac{1}{1+ae^{-r \times DivBact}}$ | - |
| DivFung | k_{MOS} et k_{MOF} | $f_{fung} = \frac{1}{1+ae^{-r \times DivFung}}$ | - |
| P_{av} | k_{MOS} | $f_P = \frac{c+(2-c)e^{-d \times P_{av}}}{1+e^{-d \times P_{av}}}$ | - |
| pH | k_{MOF} | $f_{pH} = \frac{g(pH)}{g(pH=8)}$ $g(pH) = \frac{\alpha}{1+e^{-\frac{(pH-pH_0)}{\beta}}}$ | Leifeld et al., 2008 |

Il n'existe pas de fonction de modulation de la diversité microbienne actuellement. D'après nos hypothèses, toutes les relations liées à la diversité microbienne sont positives (une relation positive avec le paramètre Y conduit à une diminution de CO_2 d'où les relations négatives obtenues dans la sélection GAM). Les relations trouvées avec la méthode GAM sont des relations linéaires. Il serait donc possible de créer une fonction linéaire dont la borne supérieure est 1 comme il en existe déjà (e.g. pour l'argile, Hansen *et al.*, 1991). Cependant, il est vraisemblable que les données utilisées ne capturent pas toute la gamme de variabilité possible de ces indices de diversité et que les relations linéaires ne représentent qu'une partie d'une relation plus complexe. En particulier dans la figure 1.4 de la section 1.3.2.2, nous avons proposé un modèle conceptuel de lien entre

diversité microbienne et fonctionnement du sol. Ce modèle, représenté par une fonction logistique (Tableau 3.4), a l'avantage d'être compris entre 0 et 1 et constitue un bon candidat pour intégrer la diversité dans un modèle de dynamique du C :

$$\begin{aligned} f_{Div} &= \frac{1}{1+ae^{-r \times Div}} \\ &= \frac{1}{1+ae^{-\frac{\ln(a) \times Div}{J}}} \end{aligned} \quad (3.8)$$

où $J = \ln(a)/r$ représente l'abscisse du point d'inflexion de la fonction, c'est à dire la valeur pour laquelle la fonction est égale à 0,5 et la croissance de la fonction ralentit. Les paramètres a et r (ou J) diffèrent vraisemblablement entre la diversité bactérienne et la diversité en champignons.

3.3.2.2 Autres propriétés du sol

L'intégration de la diversité ne peut se faire sans la prise en compte de covariables classiques qui ont montré une influence dans notre étude. Ces covariables appartiennent à deux catégories : les nutriments du sol (P assimilable et C/N) et les autres propriétés (Texture et pH).

Habituellement, les covariables de la première catégorie ne sont pas intégrées dans des fonctions de modulation des paramètres du modèle *per se*. Elles sont prises en compte par un couplage des dynamiques des différents nutriments du sol dans le modèle (e.g. couplage C et N, Manzoni and Porporato, 2009). Ici, les données disponibles ne nous permettent pas d'effectuer ce couplage sans trop complexifier la démarche et s'éloigner de l'objectif premier d'évaluation de l'intégration de la diversité microbienne. Il est cependant possible d'imaginer des fonctions candidates sur la base des connaissances de la littérature et des observations de nos études. Pour le ratio C/N, Manzoni and Porporato (2009) rapportent l'utilisation d'un facteur d'inhibition de la décomposition avec l'augmentation du C/N qui joue un rôle similaire à une fonction de modulation de paramètres (facteur φ_{ON}). Ce facteur prend également en compte le ratio C/N_b de la biomasse microbienne, le rendement d'assimilation du C (Y_{MOS}) et celui de N (η). C/N_b et η sont des inconnus et Y_{MOS} est un paramètre que l'on estime. Les inconnus peuvent être représentées par un seul paramètre b et la fonction de modulation possible peut donc se résumer à (Tableau 3.4) :

$$f_{CN} = \begin{cases} 1 & \frac{C}{N} < \frac{b}{Y_{MOS}} \\ \frac{b}{Y_{MOS} \times C/N} & \text{sinon} \end{cases} \quad (3.9)$$

Le cycle du P a beaucoup moins été modélisé et il n'existe pas dans la littérature de fac-

teur/fonction qui pourrait être utilisé ici, à notre connaissance. Cette fonction doit être décroissante d'après la relation trouvée dans le modèle GAM (Annexes D.1 et D.2) et atteint une asymptote à partir d'une certaine teneur en P assimilable (autour de 0,1 mg.g⁻¹de sol). Une fonction possible pour représenter cette relation est (Tableau 3.4) :

$$f_P = \frac{c + (2 - c)e^{-d \times P_{av}}}{1 + e^{-d \times P_{av}}} \quad (3.10)$$

Le paramètre c représente la valeur de la fonction quand le plateau est atteint et le paramètre d contrôle la valeur de P assimilable à laquelle l'inflexion a lieu.

Pour les deux autres paramètres, les fonctions de modulation peuvent être trouvées dans la littérature. Pour le pH, Leifeld et al. (2008) propose la fonction :

$$f_{pH} = \frac{g(pH)}{g(pH=8)} \quad (3.11)$$

$$g(pH) = \frac{\alpha}{1 + e^{-\frac{(pH - pH_0)}{\beta}}}$$

Pour l'argile, moins importante dans notre cas, la fonction utilisée dans le modèle DAISY (Hansen *et al.*, 1991) correspond à nos observations :

$$f_{clay} = \begin{cases} 0,5 & Clay > 250g \cdot kg^{-1} \\ 1 - 2.10^{-3} \times Clay & \text{sinon} \end{cases} \quad (3.12)$$

À l'aide d'approche statistique ad hoc, l'étude du lien entre diversité microbienne et dynamique du C sur les données DIMIMOS et les données MOSAIC nous permet de répondre à plusieurs points soulevés dans la section 1.4 et posés dans la première question de recherche :

1. Les indices de diversité microbienne sont pertinents pour expliquer les variabilités des dynamiques de C observées dans des sols aux propriétés différentes qu'ils aient été échantillonnés à une échelle nationale ou une échelle du paysage.
2. En particulier, la pertinence de ces indices dépend du mécanisme considéré et de la source de C : la diversité bactérienne est liée négativement avec les émissions de CO₂ liées à la minéralisation de la MOS, la diversité en champignons est liée positivement avec la décomposition de la MOF, probablement avec les composés plus récalcitrants, et avec la

décomposition de la MOS ; les indices ne semblent pas pertinents pour expliquer le PE.

3. Les relations entre les indices de diversité et la dynamique du C ne semblent pas être dépendantes du mode d'usage des sols
4. Cependant, ces relations peuvent dépendre du temps. En particulier, la diversité en champignons semble être reliée positivement avec des pools de MOF qui se décomposent plus lentement car plus récalcitrants.

L'ensemble des observations nous a aussi permis de proposer des fonctions de modulation de certains paramètres de décomposition dans les modèles de dynamique de MO. Il faut désormais tester l'impact de l'intégration de ces fonctions, en plus de celles correspondant aux propriétés physiques et chimiques du sol, sur le comportement et la qualité des modèles de dynamique de MO.

Chapitre 4

Intégration des indices de diversité dans un modèle mécaniste de dynamique du C

4.1 Introduction

Le nombre important de modèles mécanistes de dynamique du C dans les sols montre toute la complexité des processus intervenant dans ces dynamiques (Manzoni and Porporato, 2009). Malgré près d'un siècle de développement, le nombre de ces modèles ne cesse d'augmenter ainsi que la variété des modèles existants. Cette augmentation s'explique par l'intérêt d'utiliser ces modèles pour synthétiser la connaissance et prédire les dynamiques de C (Kutsch *et al.*, 2009; Manzoni and Porporato, 2009).

Dans le premier cas, le modèle est utilisé comme moyen de tester des hypothèses en les traduisant mathématiquement et en comparant les sorties du modèle à des observations. Il constitue alors un outil de formalisation des connaissances. Ces modèles peuvent être utiles pour pallier les limitations techniques empêchant des expérimentations, pour corroborer des hypothèses ou pour éviter la mise en place d'expérimentation lourde et coûteuse.

Dans le deuxième cas, les modèles sont utilisés pour prédire des dynamiques du C selon différents scénarios tels que des scénarios de changement climatique ou d'évolution de changement d'usage des sols. Le choix du modèle dépendra alors de plusieurs paramètres tels que les facteurs et processus pris en compte, la précision recherchée, les échelles spatiales et temporelles des prédictions.

Le contexte dans lequel s'inscrit ce travail montre que i) les bactéries et les champignons du sol sont des acteurs majeurs de la dynamique de C, ii) que l'abondance et la diversité de ces micro-organismes sont très influencées par les facteurs pédo-climatiques et les pratiques agricoles et iii) que ces pratiques agricoles ont tendance à se diversifier pour répondre à des exigences agro-écologiques. Pour prédire l'évolution de la dynamique du C dans un tel contexte, il est nécessaire de prendre en compte le rôle des micro-organismes du sol dans la dynamique du C.

En particulier, alors que la prise en compte explicite de l'abondance des micro-organismes augmente dans les modèles, améliorant la précision des mécanismes impliqués dans la dynamique du C (Wutzler and Reichstein, 2008), la diversité des micro-organismes n'est que très peu intégrée malgré son influence (section 1.3). Les premiers modèles développés qui intègrent une notion de diversité des micro-organismes ont principalement un but de test d'hypothèse et ne sont pas adaptés pour une utilisation comme outil de prédiction. Ils prennent en compte la diversité microbienne par la multiplication de compartiments non mesurables par des analyses (section 1.3).

Nous proposons d'intégrer la diversité à l'aide d'indices de diversité qui ont montré leur pertinence pour expliquer la dynamique du C dans le chapitre 3. Ces indices sont pris en compte dans des fonctions de modulation des paramètres de décomposition des modèles de dynamique de C. Des hypothèses ont été faites sur la forme des fonctions et les paramètres du modèle à moduler. Il s'agit dans ce chapitre de tester le gain de qualité prédictive d'un modèle de dynamique du C suite à la prise en compte de la diversité microbienne à l'aide de ces fonctions de modulation.

Le modèle choisi est un modèle adapté aux données d'incubation utilisées ici. L'objectif n'est pas d'aboutir à un modèle de prédiction finalisé mais de tester si la méthode proposée d'intégration de la diversité à l'aide de fonctions de modulation de paramètres est pertinente. L'évaluation de la prise en compte de la diversité microbienne est réalisée en parallèle de la prise en compte d'autres propriétés du sol influentes sur la dynamique du C. Pour répondre à des problèmes d'identifiabilité des paramètres (Sierra *et al.*, 2015), une calibration bayésienne est effectuée.

4.2 Matériels et méthodes

4.2.1 Données utilisées

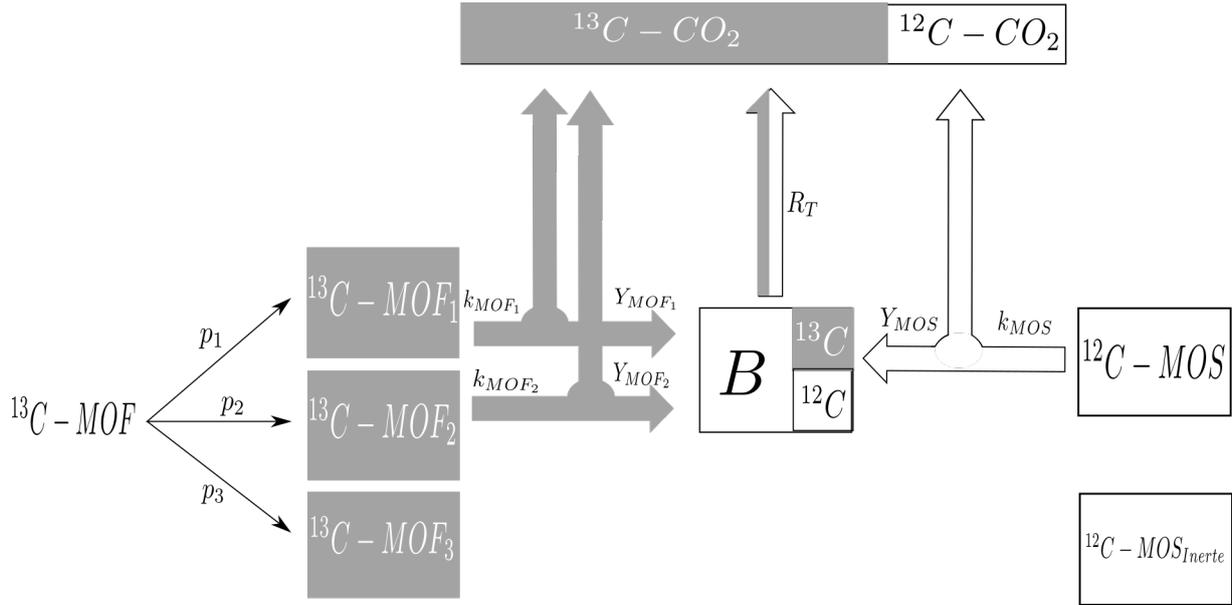
Ce travail repose principalement sur de la calibration et de l'évaluation de modèle. Comme les relations trouvées dans le chapitre précédent sont sensiblement similaires entre les jeux de données DIMIMOS et MOSAIC, et que le jeu de données MOSAIC comprend plus de données, seul ce dernier sera utilisé dans cette partie. Le nombre plus important de données permet une calibration plus robuste des modèles. Cependant, comme la texture du sol varie peu dans ce jeu de données, cette propriété du sol ne sera pas prise en compte afin de se concentrer sur les propriétés qui influencent le plus les dynamiques du C.

4.2.2 Choix du modèle de dynamique du C

4.2.2.1 Description du modèle

Le modèle de dynamique du C utilisé est fondé sur la décomposition de deux pools de MO, représentant le résidu ajouté (*MOF*) et la MO native du sol (*MOS*), alimentant un troisième pool représentant la biomasse microbienne (*B*) (Figure 2.4). Le pool de MOF est divisé en trois sous-pools avec des cinétiques de décomposition différentes. Ce formalisme nous permettra de tester l'effet de la diversité fongique sur les composés les plus récalcitrants uniquement (section 3.3). Une telle division est souvent adoptée dans la littérature et est basée sur le fractionnement Van Soest du résidu apporté (Van Soest and R.H., 1967), i.e. une fraction p_1 du C du résidu représente les composés solubles facilement décomposables (MOF_1), une fraction p_2 représente l'hémicellulose et la cellulose moyennement décomposable (MOF_2) et une fraction p_3 représente la lignine plus récalcitrante (MOF_3) (Paul and Clark, 1989; Whitmore *et al.*, 1997; Henriksen and Breland, 1999b; Garnier *et al.*, 2001).

Pour simplifier le modèle et ne pas trop augmenter le nombre de paramètres, on fait l'hypothèse ici que les flux de C issus du pool MOF_3 , dont la cinétique de décomposition est lente, sont négligeables sur la durée d'incubation de 80 jours. Ce pool est donc considéré non décomposé (donc constant). De même, une fraction f de la MOS est considérée stable à cette échelle de temps (MOS_{inerte}) (Petersen *et al.*, 2005a,b; Franko *et al.*, 2011).

FIGURE 4.1 – **Modèle de dynamique du carbone.**

Les pools et flux contenant du C marqué à l'isotope ¹³C sont représentés en gris. Les pools et flux contenant du C non marqué sont représentés en blanc. MOF_i , sous pool i de la matière organique fraîchement ajoutée; MOS , matière organique native du sol; B , biomasse microbienne; k_i , constante de décomposition du pool i ; Y_i , paramètre de rendement d'utilisation du C du pool i ; R_T , paramètre de respiration des micro-organismes autre que la respiration pour la croissance et p_i , fraction du pool MOF_i .

4.2.2.2 Cinétiques de décomposition/minéralisation du modèle

Équations différentielles

La décomposition d'un pool i de MO suit une cinétique du 1^{er} ordre : à chaque pas de temps, une proportion k_i (constante de décomposition) du C présent dans le pool i est décomposée. Ainsi, pour les pools MOF_1 , MOF_2 et MOS , les équations des cinétiques de décomposition s'écrivent :

$$\frac{dC_i}{dt} = -k_i C_i \quad (4.1)$$

où, C_i est la quantité de C dans le pool i à un instant t (jour) et k_i la constante de décomposition.

Une partie Y_i (rendement d'assimilation du C) est utilisée pour la croissance des micro-organismes (B) et l'autre partie $(1 - Y_i)$ est minéralisée sous forme de CO_2 . Les micro-organismes respirent également une proportion R_T de leur C sous forme de CO_2 à chaque pas de temps. Cette proportion correspond à l'ensemble des flux respirés par les micro-organismes (maintenance, production d'enzymes, cf. Figure 1.2 section 1.2) et permet dans le modèle de prendre en compte l'influence de la biomasse microbienne. Pour simplifier la calibration (cf section suivante), la

production de composés carbonés par les micro-organismes (e.g. enzymes, polysaccharides) est négligée dans ce modèle. Pour les pools microbiens et de CO_2 , le modèle distingue la partie marquée au ^{13}C (i.e. issue du résidu) de la partie non marquée (i.e. issue de la MOS). Cette distinction permet l'utilisation séparée des données marquées et non marquées de CO_2 pour la calibration et ainsi l'amélioration des résultats (Blagodatsky *et al.*, 2010). La croissance des micro-organismes, selon la source de C utilisée, est donc décrite par les équations suivantes :

$$\frac{dB13}{dt} = k_{MOF_1} \cdot Y_{MOF_1} \cdot C_{MOF_1} + k_{MOF_2} \cdot Y_{MOF_2} \cdot C_{MOF_2} - R_T \cdot B13 \quad (4.2)$$

$$\frac{dB12}{dt} = k_{MOS} \cdot Y_{MOS} \cdot C_{MOS} - R_T \cdot B12 \quad (4.3)$$

où $B13$ et $B12$ sont les quantités de C microbien marqué et non marqué respectivement.

Finalement les cinétiques concernant les flux de CO_2 s'écrivent :

$$\frac{d(^{13}C - CO_2)}{dt} = k_{MOF_1} \cdot (1 - Y_{MOF_1}) \cdot C_{MOF_1} + k_{MOF_2} \cdot (1 - Y_{MOF_2}) \cdot C_{MOF_2} + R_T \cdot B13 \quad (4.4)$$

$$\frac{d(^{12}C - CO_2)}{dt} = k_{MOS} \cdot (1 - Y_{MOS}) \cdot C_{MOS} + R_T \cdot B12 \quad (4.5)$$

En l'absence de résidu, les parties grisées de la figure 4.1 ainsi que les équations 4.1 (pour MOF_1 et MOF_2), 4.2 et 4.4 sont absentes du modèle.

Modulation des paramètres de décomposition par la diversité microbienne et les autres propriétés des sols

Comme décrit dans la section 1.2.3, les valeurs des paramètres du modèle peuvent être modulées à l'aide de fonctions dépendantes des propriétés des sols. Les conclusions du chapitre précédent suggèrent de moduler les paramètres k_{MOF_1} , k_{MOF_2} , k_{MOS} et Y_{MOS} comme suit :

$$k_{MOF_1} = kr_1 \cdot f_{pH} \quad (4.6)$$

$$k_{MOF_2} = kr_2 \cdot f_{pH} \cdot f_{fung} \quad (4.7)$$

$$k_{MOS} = k_s \cdot f_P \cdot f_{CN} \cdot f_{fung} \quad (4.8)$$

$$Y_{MOS} = Y_s \cdot f_{bact} \quad (4.9)$$

où kr_1 , kr_2 , k_s et Y_s sont les valeurs de base des paramètres k_{MOF_1} , k_{MOF_2} , k_{MOS} et Y_{MOS} , respectivement. Ces valeurs peuvent être modulées par des fonctions dépendantes du pH (f_{pH}), du ratio C/N du sol (f_{CN}), du phosphore assimilable (f_P), de la diversité bactérienne (f_{bact}) et/ou de la diversité fongique (f_{fung}). Pour la diversité bactérienne comme pour la diversité fongique l'inverse de l'indice de Simpson est considéré. Lorsque ces fonctions sont considérées constantes et égales à 1, la propriété correspondante n'a pas d'effet, lorsqu'elles sont égales à une fonction dépendante de la propriété correspondante, comme celles proposées en section 3.3 et rappelées dans le tableau 4.1, la propriété module le paramètre de décomposition.

TABLE 4.1 – Fonctions de modulation possibles des paramètres d'un modèle de dynamique du C dans le sol.

| Propriété | Paramètres ajustés | Fonction possible | Référence |
|-----------|------------------------|---|------------------------------|
| C/N | k_{MOS} | $f_{CN} = \begin{cases} 1 & \frac{C}{N} < b \\ \frac{b}{C/N} & \text{sinon} \end{cases}$ | Manzoni and Porporato, 2009 |
| DivBact | Y_{MOS} | $f_{bact} = \frac{1}{1 + ae^{-\frac{\ln(a) \times DivBact}{J}}}$ | - |
| DivFung | k_{MOS} et k_{MOF} | $f_{fung} = \frac{1}{1 + \gamma e^{-r \times DivFung}}$ | - |
| P_{av} | k_{MOS} | $f_P = \frac{c + (2-c)e^{-d \times P_{av}}}{1 + e^{-d \times P_{av}}}$ | - |
| pH | k_{MOF} | $f_{pH} = \frac{g(pH)}{g(pH=8)}$ $g(pH) = \frac{\alpha}{1 + e^{-\frac{(pH-pH_0)}{\beta}}}$ | Leifeld <i>et al.</i> , 2008 |

Il faut noter que la fonction associée au ratio C/N est légèrement différente de celle proposée dans le tableau 3.4. En effet, sachant que Y_{MOS} peut être ajusté par d'autres fonctions, sa prise en compte dans la fonction associée au C/N conduit à une interaction entre cette propriété et les propriétés des fonctions de modulation de Y_{MOS} . Ces interactions peuvent compliquer considérablement la calibration. De fait, Y_{MOS} a été « intégré » dans le paramètre b .

4.2.3 Stratégie de calibration

4.2.3.1 Complexité des modèles et identifiabilité des paramètres

Plus un modèle est détaillé (e.g. multiplication de compartiments, introduction de nombreuses fonctions de modulation), plus le nombre de paramètres à estimer augmente. L'identifiabilité des paramètres d'un modèle est définie comme l'obtention de valeurs de la variable simulée égales avec différentes valeurs des paramètres du modèle (Sierra *et al.*, 2015). Autrement dit, si les paramètres d'un modèle ne sont pas identifiables, le changement des valeurs d'un paramètre peut être compensé par de faibles changements de la valeur d'autres paramètres.

L'identifiabilité des paramètres est fonction du nombre de paramètres à estimer : un modèle non identifiable est donc dit sur-paramétré. De plus, elle ne peut être dissociée des données disponibles pour la calibration. En général, plus le nombre de variables intervenant dans le modèle et ayant été mesurées est grand, moins la calibration rencontre le problème d'identifiabilité des paramètres. Un modèle est non identifiable (ou sur-paramétré) étant donné les données disponibles (Sierra *et al.*, 2015).

Dans notre cas, seules les dynamiques du CO₂ sont disponibles. Il est possible de calculer un indice de colinéarité des paramètres à estimer qui renseigne sur les problèmes d'identifiabilité (Sierra *et al.*, 2015 ; Package FME, Soetaert *et al.*, 2010). La figure 4.2 représente la valeur de cet indice pour notre modèle en fonction du nombre de paramètres à estimer, sans inclure les paramètres des éventuelles fonctions de modulation, pour le modèle appliqué aux sols témoins (violet) et aux sols amendés (orange). Cette figure montre que plus on veut calibrer de paramètres, plus l'identifiabilité sera un problème (indice supérieur à 20, Sierra *et al.*, 2015), i.e. moins les valeurs calibrées des paramètres n'auront de sens.

En particulier, pour le modèle calibré sur les sols témoins, l'identifiabilité sera toujours un problème dès 2 paramètres à estimer et pour le modèle calibré sur les sols amendés, dès 6 paramètres à estimer. Ce problème sera plus important que l'on complexifie le modèle en ajoutant les fonctions de modulation des paramètres de décomposition, et donc en ajoutant des paramètres à estimer.

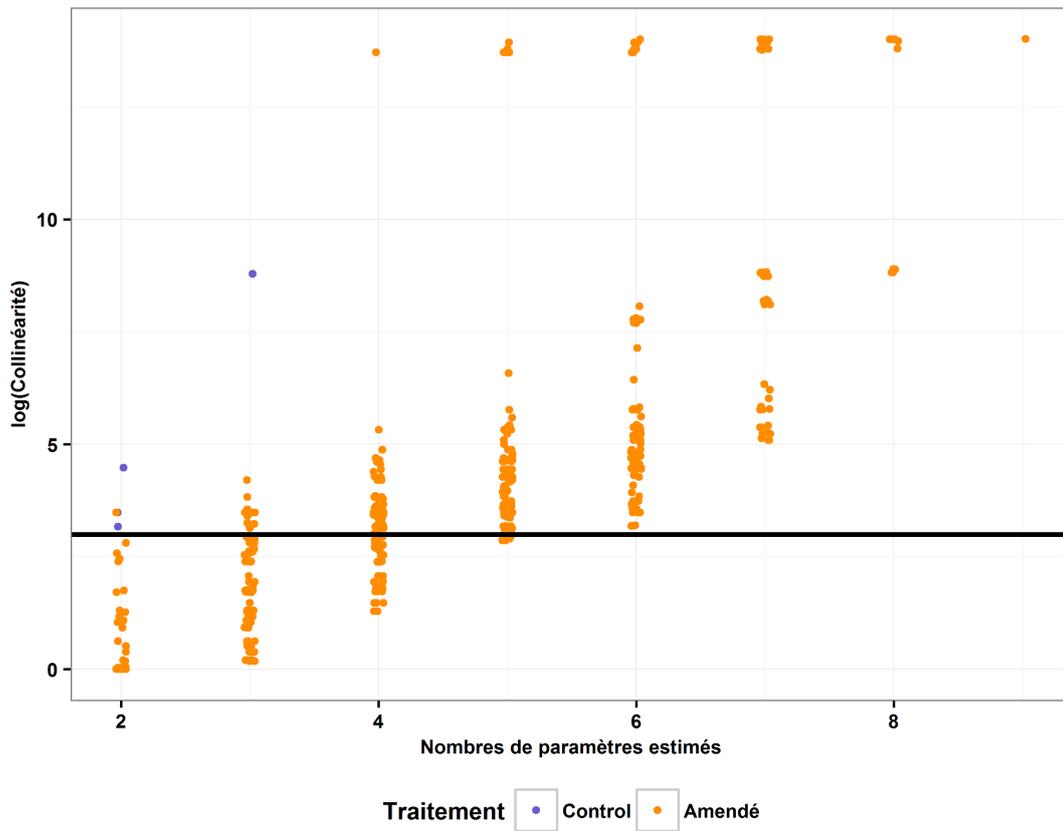


FIGURE 4.2 – Logarithme de l'indice de colinéarité du modèle en fonction du nombre de paramètres à estimer.

Chaque point représente le logarithme de l'indice de colinéarité pour une combinaison de n paramètres à estimer (axis des abscisses) parmi le nombre total (N) de paramètres dans le modèle sans modulation des paramètres de décomposition pour les sols control (violet, $N=3$) et les sols amendés (orange, $N=9$). La ligne horizontale noire représente le seuil de l'indice de colinéarité ($\log(20)$) pour lequel il y a un problème d'identifiabilité des paramètres s'il est dépassé (Sierra *et al.*, 2015).

4.2.3.2 Inférence bayésienne

Dans le cadre de l'inférence bayésienne, les paramètres du modèle (θ) sont considérés comme des variables aléatoires définies *a priori* (avant calibration) par une loi de probabilité *a priori* ($P(\theta)$) (Clark, 2007; Parent and Rivot, 2012). Cette connaissance *a priori* est mise à jour à l'aide de l'ensemble des données observées, $Y = y$, selon une équation basée sur le théorème de Bayes :

$$P(\theta, Z|Y = y) \propto P(\theta) \cdot P(Z|\theta) \cdot P(Y = y|\theta, Z) \quad (4.10)$$

où $P(\theta, Z|Y = y)$ est la distribution *a posteriori* (après calibration) des paramètres du modèle et des variables d'état Z sachant les données observées et $P(Y = y|\theta, Z)$ est une fonction de vraisemblance que l'on cherche à maximiser. Z représente ici l'ensemble des équations du

modèles et sont pris en compte de manière déterministe. Dans cette équation, on peut voir que la distribution *a posteriori* (posterior) des paramètres est facteur à la fois de la connaissance *a priori* apportée par le prior ($P(\theta)$) et des connaissances apportées par les données observées $Y = y$, en fonction des hypothèses du modèle. On parle aussi d'assimilation des données pour définir la calibration ou l'inférence bayésienne.

La possibilité de préciser une connaissance *a priori* sur la distribution des paramètres d'un modèle donne un avantage à la calibration bayésienne lorsque les modèles sont sur-paramétrés ce qui en fait une approche intéressante en cas de non-identifiabilité des paramètres (Reichert and Omlin, 1997; Huang *et al.*, 2006; Ahrens *et al.*, 2014). Très récemment, cette approche a été utilisée pour la calibration de modèles mécanistes de dynamique du C souvent sur-paramétrés (e.g. Hararuk and Luo, 2014; Hararuk *et al.*, 2015; Schädel *et al.*, 2013). C'est également celle choisie dans ce travail.

Pour l'inférence bayésienne, l'assimilation des données se fait en considérant les données observées, ici les émissions de CO_2 , comme des variables aléatoires qui suivent une certaine loi. Ici, on considère que les flux de CO_2 suivent des lois normales avec pour moyenne les solutions des équations différentielles et pour écart-type un paramètre à estimer :

$$P(^{12}C - CO_2 | \theta, Z) \sim \mathcal{N} \left(\text{solution} \left(\frac{d(^{12}C - CO_2)}{dt} \right), sd12 \right) \quad (4.11)$$

$$P(^{13}C - CO_2 | \theta, Z) \sim \mathcal{N} \left(\text{solution} \left(\frac{d(^{13}C - CO_2)}{dt} \right), sd13 \right) \quad (4.12)$$

où $\text{solution} \left(\frac{d(^{12}C - CO_2)}{dt} \right)$ et $\text{solution} \left(\frac{d(^{13}C - CO_2)}{dt} \right)$ sont les solutions des équations différentielles 4.4 et 4.5, respectivement, et $sd12$ et $sd13$ sont les écarts-types des flux de CO_2 .

En pratique, la calibration bayésienne (estimation des distributions *a posteriori*) repose sur des méthodes de simulations numériques comme les méthodes de Monte Carlo par Chaines de Markov (MCMC). Ces méthodes sont implémentées dans le logiciel OpenBUGS (Lunn *et al.*, 2009) qui sera utilisé à l'aide du package R2OpenBUGS (Sturtz *et al.*, 2010) du logiciel de statistique R (R Core Team, 2015).

4.2.3.3 Diagnostic de convergence

L'inférence bayésienne étant basée sur des simulations numériques (MCMC), il est important de vérifier la convergence de ces simulations. Dans cette étude, à chaque calibration, 3 chaînes MCMC sont effectuées en parallèle à l'aide de valeurs initiales différentes pour les paramètres et

tirées aléatoirement dans les distributions *a priori*. Le diagnostic de convergence peut se faire soit pour chaque chaîne, soit sur l'ensemble des chaînes. Plusieurs tests sont disponibles et utilisés pour vérifier la convergence de la calibration. Pour chaque calibration, les deux tests suivants sont utilisés :

Le test de Gelman et Rubin (Gelman and Rubin, 1992) vérifie que les chaînes ne sont plus influencées par leur point de départ (valeurs initiales des paramètres) et se fonde sur des comparaisons de variance intra- et inter- chaînes.

Le test de Heidelberger et Welch (Heidelberger and Welch, 1981) vérifie que chaque chaîne est issue d'un processus stationnaire.

4.2.4 Étapes de construction et d'évaluation des fonctions de modulation de la décomposition

Afin d'évaluer les gains d'ajustement et de prédiction associés à la prise en compte des propriétés du sol, et particulièrement de la diversité microbienne, le modèle choisi a été calibré selon différentes modalités consistant à intégrer au fur et à mesure les fonctions de modulation des paramètres de décomposition précédemment citées. Le modèle ne prend pas en compte l'interaction entre les décompositions de la MOF et de la MOS (i.e. le PE) et donc la calibration a été effectuée distinctement pour le modèle appliqué aux sols témoins (i.e. sans MOF, modèles « C ») et pour le modèle appliqué aux sols amendés (i.e. avec MOF, modèles « A »). L'ensemble des étapes permettant d'aboutir à l'évaluation des gains d'ajustement et de prédiction associés à la prise en compte des fonctions de modulation sont représentées dans le diagramme de la figure 4.3.

4.2.4.1 Initialisation et construction des distributions *a priori*

Initialisation des modèles

Comme pour tous les modèles dynamiques, les variables ont besoin d'être initialisées, i.e. d'avoir une valeur au temps $t = 0$. Ce temps $t = 0$ correspond au début des incubations dans cette étude. Le carbone du pool total de MOS correspond à la teneur en C organique mesurée pour chaque sol (SOC). Une fraction f de ce C est réellement disponible pour la décomposition tandis que le reste est considéré inerte. La valeur de cette fraction est prise dans la littérature et est fixée à 2% (Franko *et al.*, 2011).

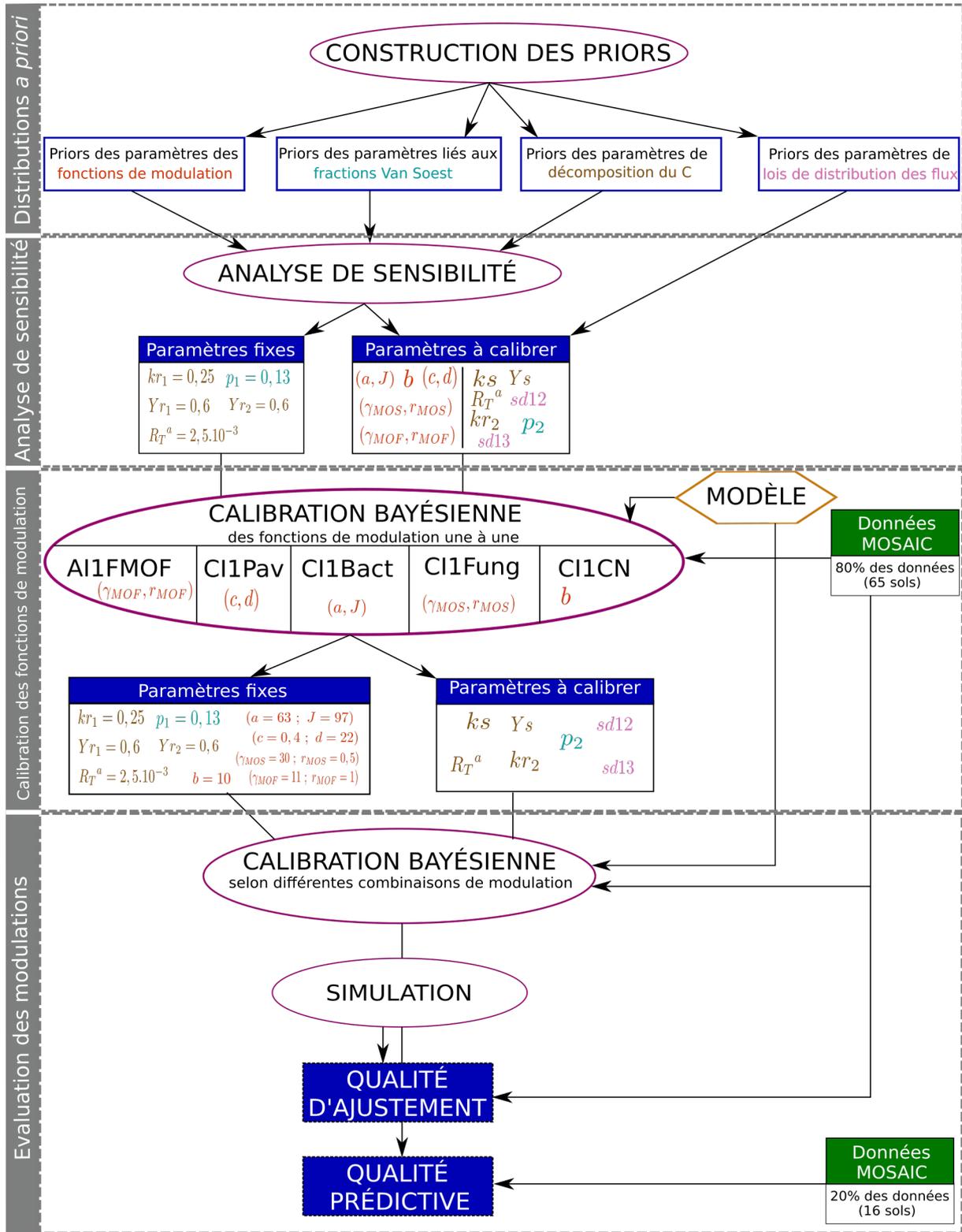


FIGURE 4.3 – Diagramme des étapes de construction et d'évaluation des fonctions de modulation.

Les données (en vert) sont analysées ou testées à l'aide de méthodes spécifiques (en violet) pour conduire à des résultats (en bleu).

La quantité R de C correspondant au pool MOF est égal à la quantité de C dans le résidu et a été mesurée ($R = 2,17$ g C/kg sol sec). La totalité de ce C est répartie dans les sous-pools MOF_1 , MOF_2 et MOF_3 selon les proportions p_1 , p_2 et p_3 (qui sont soit fixes, soit estimées), respectivement. Pour le modèle appliqué aux sols témoins, cette quantité est nulle car aucun amendement n'est effectué.

La quantité de C dans la biomasse microbienne n'a pas été mesurée dans cette étude. Suivant [Blagodatskaya et al. \(2014\)](#), la quantité de C microbien du pool $B12$ peut être estimée à partir de la biomasse moléculaire, i.e. à partir de la quantité d'ADN microbien mesurée dans le sol. En considérant que l'ADN correspond à 13% du poids des micro-organismes et que 45% du poids des micro-organismes est attribuable au C microbien, on peut estimer la biomasse microbienne selon l'équation :

$$B12(t = 0) = \frac{0,45 \cdot 10^{-3} \times ADN}{0,13} \quad (4.13)$$

où ADN est la quantité d'ADN microbien dans le sol ($\mu\text{g/g sol}$). Les valeurs de $B12(t = 0)$ varient alors de $5 \cdot 10^{-2}$ à $5 \cdot 10^{-1}$ g/kg sol, ce qui correspond à 0,2 à 2% du C organique total. Pour le pool de C microbien marqué au ^{13}C ($B13$), au début d'incubation, il est inexistant car aucune décomposition du résidu n'a encore eu lieu. Sa valeur est donc nulle. Il en est de même pour les valeurs initiales des pools de CO_2 .

Construction des distributions *a priori*

Pour la construction des distributions *a priori*, on distingue 4 groupes de paramètres :

1. Les paramètres des cinétiques du C (ks , kr_1 , kr_2 , Ys , Yr_1 , Yr_2 , R_T) dont les distributions *a priori* sont construites en fonction de valeurs de la littérature.
2. Les écarts-types des flux de CO_2 ($sd12$ et $sd13$).
3. Les paramètres p_1 , p_2 et p_3 dont les distributions *a priori* correspondent aux distributions des proportions de C correspondant aux fractions Van Soest de la paille de blé : fractions soluble, cellulose+hémicellulose et lignine respectivement.
4. Les paramètres des fonctions de modulation. Ces paramètres seront fixés si des valeurs existent dans la littérature. Sinon, les distributions *a priori* seront construites en fonction des relations du précédent chapitre ou des informations disponibles dans la littérature.

Les paramètres des cinétiques du C

Pour construire les distributions *a priori* des paramètres des cinétiques du C, des valeurs

de ces paramètres ont été sélectionnées dans la littérature pour des modèles similaires. Ces valeurs de paramètres sont considérées comme des variables aléatoires suivant une certaine loi de probabilité. Ces lois sont estimées par inférence bayésienne et constituent les distributions *a priori* des paramètres. En particulier, les constantes de décomposition étant des paramètres dont les valeurs sont strictement positives, des lois log-normales ont été ajustées sur ces valeurs. Pour les rendements d'utilisation du C et pour la proportion R_T de C respiré par les micro-organismes, tous compris entre 0 et 1, des lois de distribution bêta ont été ajustées. Les valeurs et références sélectionnées sont visibles en annexe E.1. Les résultats des constructions des distributions *a priori* sont visibles dans le tableau 4.3.

Les écarts-types des flux de CO₂

Il est plus difficile d'appliquer la méthode précédente sur les écarts-types des émissions de CO₂ car ces valeurs dépendent beaucoup de la situation, i.e. le protocole de mesure, les sols utilisés, etc. Les données DIMIMOS constituent une situation similaire. De fait, on utilise les écarts-types calculés sur les flux de CO₂ dans les données DIMIMOS pour un même sol, aux différents temps de mesure pour ajuster des lois gamma par inférence bayésienne. La loi gamma est souvent utilisée pour représenter les écarts-types d'une variable.

Les paramètres des fractions Van Soest

On cherche à connaître la distribution de la proportion de C (pC_{FRAC}) associée à chaque fraction Van Soest $FRAC$ d'une paille de blé. Ces proportions n'ont pas été mesurées pour la paille de blé utilisée dans cette étude et la littérature offre peu d'information. Cependant, elles peuvent être calculées comme suit :

$$pC_{FRAC} = \frac{x_{FRAC} \times p_{FRAC}}{pC} \quad (4.14)$$

où x_{FRAC} est la proportion du poids de la fraction $FRAC$ (i.e. fraction soluble, hémicellulose, cellulose ou lignine) attribuable au C (g C/g $FRAC$), p_{FRAC} est la proportion du poids de la paille de blé attribuable à la fraction $FRAC$ (g $FRAC$ /g paille de blé) et pC est la proportion du poids de la paille de blé attribuable à la totalité du C (g C/g paille de blé). Dans notre cas, seule pC est connue. x_{FRAC} et p_{FRAC} sont des données que l'on trouve dans la littérature. Pour p_{FRAC} , les données sélectionnées concernent seulement de la paille de blé. Pour x_{FRAC} , on fait l'hypothèse que la quantité de C dans chaque fraction ne dépend pas du résidu considéré, les

valeurs sélectionnées ne concernent donc pas seulement de la paille de blé. Toutes les valeurs et références sont visibles en annexe E.2.

À l'aide de ces valeurs, nous avons simulé des valeurs de pC_{FRAC} et ajusté, par calibration bayésienne, une loi de Dirichlet. Cette loi est une généralisation de la loi bêta à plus de deux proportions. Il est possible ensuite de déduire des distributions *a priori* de p_1 , p_2 et p_3 en déduisant les lois bêta des fractions solubles, d'hémicellulose+cellulose et de lignine, respectivement, à partir de la loi de Dirichlet ajustée.

Les paramètres des fonctions de modulation de la décomposition

Pour les fonctions de modulation, et toujours dans un souci de simplification de la calibration, il est décidé d'utiliser des valeurs de paramètres disponibles dans la littérature si tel est le cas. Sinon des déductions et hypothèses sont faites par expertise.

La seule fonction pour laquelle il existe des valeurs dans la littérature est celle concernant le pH (Leifeld *et al.*, 2008). Ces valeurs sont $\alpha = 0,1621$, $\beta = 0,8643$ et $pH_0 = 5,3453$. L'allure de cette fonction semble en accord avec les observations que l'on fait dans le chapitre précédent et nous avons donc choisi de l'utiliser et de ne pas calibrer de nouvelle fonction pour le pH, permettant ainsi de diminuer le nombre de paramètres à estimer.

Pour les autres fonctions, il n'existe pas de valeurs comme pour le pH, leurs paramètres sont donc à estimer et des distributions *a priori* doivent être proposées. Comme il existe peu d'information sur ces paramètres, des lois de distribution uniforme sont utilisées dont les paramètres sont à définir.

Pour f_{CN} : La fonction concernant le ratio C/N du sol ne comprend qu'un paramètre b (Tableau 4.1). Ce paramètre, selon Manzoni and Porporato (2009), est définie comme $b = \frac{\eta \times CN_B}{Y}$ avec η et Y les rendements d'utilisation du N et de C, respectivement, tous deux compris entre 0 et 1, et CN_B le ratio C/N de la biomasse microbienne dont les valeurs sont généralement comprises entre 5 et 15 (Manzoni and Porporato, 2009). De plus, la fonction est constante et égale à 1 tant que le C/N du sol est inférieure à b . Or, les données utilisées ici sont les données MOSAIC dont le C/N du sol varie entre 10 et 12. Si on veut que la fonction de modulation liée au C/N du sol varie dans la gamme des données MOSAIC, le paramètre b ne peut être supérieur à 10. Pour toutes ces raisons, la distribution *a priori* de ce paramètre est définie comme une loi uniforme comprise entre 2 et 10.

Pour f_{Pav} : La fonction associée au P assimilable n'existe pas dans la littérature. Elle dépend de deux paramètres c et d . Le paramètre c représente la valeur de l'asymptote vers laquelle la fonction tend pour des valeurs hautes du P assimilable. N'ayant pas d'indication particulière, une distribution uniforme entre 0 et 1 est choisie pour ce paramètre. Le paramètre d détermine le point d'inflexion de la courbe, i.e. la valeur du P assimilable à laquelle le plateau est atteint. Les résultats de la section 3.2, visible dans les annexes D.1 et D.1, suggèrent que ce plateau est atteint autour de 10 à 15 g P/kg sol. Par test graphique, une valeur comprise entre 20 et 40 pour le paramètre d semble correspondre à cette observation. Une loi uniforme comprise entre 20 et 40 est donc choisie.

Pour f_{bact} : La fonction associée à la diversité bactérienne (inverse de l'indice de Simpson), telle que présentée dans le tableau 4.1, dépend des paramètres a et J . Le paramètre J correspond au point d'inflexion de la fonction logistique, i.e. la valeur de diversité bactérienne pour laquelle la fonction f_{bact} est égale à 0,5 et après laquelle la croissance de la fonction décélère. Comme supposé dans la section 3.3, il est vraisemblable que les données MOSAIC ne capturent pas toute la gamme possible de diversité bactérienne. Il est difficile de trouver des valeurs comparables en termes de méthodologie. Cependant, une étude utilisant la même méthodologie dans une situation similaire (Constancias *et al.*, 2015) permet de faire l'hypothèse que les données de diversité bactérienne dans MOSAIC sont plutôt dans des valeurs élevées de la gamme possible. En conséquences, on peut faire l'hypothèse que le point J se situe au début de la gamme capturée par MOSAIC ou avant. On définit alors une loi uniforme comprise entre 50 et 150 pour le paramètre J . Pour le paramètre a , qui détermine la valeur de la fonction en 0, aucune information n'est disponible. On définit donc une loi uniforme comprise entre 0 et 100.

Pour $f_{fungMOS}$ et $f_{fungMOF}$: Pour les fonctions associées à la diversité fongique (inverse de l'indice de Simpson), le même raisonnement que pour la diversité bactérienne n'est pas possible car aucune information comparable pour le niveau de diversité n'a été trouvée dans la littérature. Pour déduire des distributions *a priori*, on se base sur l'allure de la courbe des résultats de la section 3.2 (annexe D.3) entre la diversité fongique et la dynamique du C non marqué et du C marqué à la période 2. Cette dernière atteint un plateau pour une valeur de diversité autour de 15, l'autre est une fonction linéaire. Par test graphique, on pose le paramètre γ comme suivant une loi *a priori* uniforme entre 0 et 50 et le paramètre r comme suivant une loi *a priori* uniforme entre 0 et 1.

4.2.4.2 Choix des paramètres à fixer : analyse de sensibilité multiple

Bien que l'inférence bayésienne permette de limiter le problème de sur-paramétrage du modèle, il est également possible de fixer la valeur de certains paramètres pour simplifier la calibration, par exemple en utilisant des valeurs préalablement estimées dans d'autres études. Le risque d'une telle approche est de fixer des valeurs non adaptées à l'étude en cours. Pour éviter ce problème, les paramètres à fixer sont choisis parmi ceux qui ont le moins d'importance dans le modèle, i.e. dont un changement de valeur aura peu d'impact sur les sorties du modèle. Cette importance est mesurable à l'aide d'une analyse de sensibilité globale.

Dans notre cas, certains paramètres ne peuvent être fixés même si leur indice de sensibilité est faible. C'est le cas des paramètres des fonctions de modulation, tant qu'elles ne sont pas calibrées, mais également des paramètres des cinétiques du C qui sont modulés par ces fonctions (i.e. ks , Ys et kr_2). Cependant, les indices de sensibilité de ces paramètres sont tout de même calculés afin d'avoir une vue d'ensemble de l'importance des différents paramètres.

Plusieurs méthodes existent pour mesurer des indices de sensibilité et celle adoptée dans ce travail est décrite dans Wang *et al.* (2013). Cette analyse de sensibilité est une analyse de sensibilité multiple i.e. prenant en compte tous les paramètres en même temps et leurs interactions, dont les étapes sont :

1. Générer aléatoirement plusieurs ensembles de valeurs des paramètres à analyser en accord avec leur distribution. Les distributions utilisées ici sont les distributions *a priori* des paramètres (cf. section précédente)
2. Faire tourner le modèle pour chaque ensemble de valeurs des paramètres et calculer la somme des erreurs quadratiques des simulations aux données observées. Les données observées sont définies comme les valeurs du modèle lorsque tous les paramètres sont fixés à leur valeur médiane. Les données observées sont différentes pour chaque paramètre en fonction des sorties du modèle sur lesquelles ces paramètres ont une influence. Par exemple, ks a une influence sur le CO₂ non marqué et sera analysé par rapport à cette sortie du modèle, kr_1 sur les flux de CO₂ marqué et sera analysé par rapport à cette sortie du modèle et R_T sur les deux et sera analysé par rapport à la somme des sorties du modèle correspondant au CO₂ marqué et non marqué. Cette information est visible dans le tableau 4.3.
3. Classer la première moitié des ensembles de paramètres qui ont des valeurs d'erreur qua-

dratique les plus basses comme « acceptable » et le reste comme « inacceptable ».

4. Normaliser les valeurs de chaque paramètre comme suit :

$$xn_{i,j} = \frac{x_{i,j} - x_{i,min}}{x_{i,max} - x_{i,min}} \quad (4.15)$$

où $xn_{i,j}$ et $x_{i,j}$ sont les valeurs j normalisées et non normalisées, respectivement, du paramètre i et $x_{i,min}$ et $x_{i,max}$ sont les valeurs minimales et maximales non normalisées du paramètre j . Cette normalisation permet d'avoir des valeurs comprises entre 0 et 1 pour tous les paramètres. Pour les paramètres suivant une distribution logarithmique (e.g. log-normal), il faut prendre le logarithme des valeurs non normalisées pour la normalisation.

5. Pour chaque paramètre, calculer la moyenne des différences entre les courbes de probabilité cumulée des distributions des sous-ensembles « acceptable » (xn^A) et « non acceptable » (xn^U) pour la mesure de l'indice de sensibilité :

$$SI_i = \frac{1}{n} \sum_{k=1}^n |xn_{i,j}^A - xn_{i,j}^U| \quad (4.16)$$

où SI_i est l'indice de sensibilité du paramètre i .

Les valeurs de l'indice de sensibilité peuvent être comparés entre les paramètres pour hiérarchiser les paramètres selon leur importance.

L'analyse de sensibilité présentée a été réalisée sur le modèle appliqué aux sols témoins et sur le modèle appliqué aux sols amendés, avec toutes les fonctions de modulations. La moyenne des indices de sensibilité calculés pour chaque paramètre pour ces 2 modalités est utilisée en tant qu'indice moyen de sensibilité globale. Les paramètres à fixer dans le modèle sont choisis en fonction des résultats de cette analyse de sensibilité.

4.2.4.3 Calibration des paramètres des fonctions de modulation de la décomposition

Prendre en compte toutes les fonctions de modulation dans le modèle pour calibrer à la fois leurs paramètres mais aussi les paramètres de décomposition conduirait à augmenter considérablement le nombre de paramètres à calibrer. Bien que l'inférence bayésienne et le fait de fixer certains paramètres permettent de lutter contre le problème de sur-paramétrage, le grand nombre de paramètres à calibrer suite à la prise en compte de toutes les fonctions de modulation, associé

à la colinéarité qui existe entre certains de ces paramètres, peut compliquer la convergence de la calibration bayésienne (Sierra *et al.*, 2015).

Afin de limiter le nombre de paramètres à calibrer, les fonctions de modulation ont été prises en compte une par une pour l'estimation de leurs paramètres propres. Ainsi, pour le modèle appliqué aux sols témoins, les modèles « CI1CN », « CI1Pav », « CI1Bact » ou « CI1Fung » sont respectivement les modèles prenant en compte le ratio C/N, le P assimilable, la diversité bactérienne ou la diversité fongique selon les équations 4.6 à 4.9. Par exemple, pour le modèle CI1CN, prenant en compte le ratio C/N, cela revient à définir f_{CN} par la fonction du tableau 4.1 et à fixer f_{Pav} , f_{Bact} et f_{Fung} à 1. Ces modèles ont pour objectif de calibrer les fonctions de modulation et ne seront pas utilisés pour une comparaison de qualité de modèles (Tableau 4.2).

Pour le modèle appliqué aux sols amendés, les fonctions de modulation communes entre le modèle appliqué à ces sols et le modèle appliqué aux sols témoins (i.e. toutes sauf f_{CN}) ont été réutilisées. La fonction concernant le pH provient de la littérature et n'a pas besoin d'être calibrée. Il reste donc la modulation du paramètre kr_2 par une fonction dépendante de la diversité en champignons. On suppose ici que cette fonction est différente de celle qui ajuste la constante de décomposition de la MOS et qu'il faut à nouveau la calibrer. On distinguera dans la suite les fonctions $f_{fungMOS}$ et $f_{fungMOF}$ les fonctions dépendantes de la diversité en champignons qui ajustent respectivement la constante de décomposition de la MOS (ks) et la constante de décomposition du sous-pool MOF₂ (kr_2). Ainsi, le modèle « A11-FMOF » est le modèle qui permet de calibrer les paramètres de $f_{fungMOF}$.

Une fois calibrés, les paramètres des fonctions de modulation seront fixés à des valeurs dépendantes de leur distribution *a posteriori*, lors de la prise en compte de différentes combinaisons des fonctions de modulation pour évaluer leur impact sur la qualité du modèle.

4.2.4.4 Évaluation de l'impact des fonctions de modulation de la décomposition sur la qualité du modèle

Combinaisons des fonctions de modulation prises en compte

Pour évaluer l'impact des fonctions de modulation ainsi calibrées sur la qualité du modèle, différentes combinaisons de prise en compte de ces fonctions, considérées comme valeurs fixes, sont intégrées dans le modèle. Après calibration des paramètres de cinétiques du C, les qualités d'ajustement et de prédiction des modèles sont calculées (cf. section suivante, Tableau 4.2).

TABLE 4.2 – Ensemble des modèles calibrés et modulations correspondants des paramètres.

| Combinaison Modulations correspondants ^a | | Utilisation | Calibration ^b |
|---|--|-----------------------------|--------------------------------------|
| Modèle pour les sols non amendés (témoin) | | | |
| CI1CN | $\forall i, k_{MOS,i} = ks \cdot f_{CN}(i), Y_{MOS,i} = Ys$ | Modulation de f_{CN} | $\theta_{dyn} + b$ |
| CI1Pav | $\forall i, k_{MOS,i} = ks \cdot f_P(i), Y_{MOS,i} = Ys$ | Modulation de f_P | $\theta_{dyn}Témoin + c$ et d |
| CI1Bact | $\forall i, k_{MOS,i} = ks, Y_{MOS,i} = Ys \cdot f_{bact}(i)$ | Modulation de f_{bact} | $\theta_{dyn}Témoin + a$ et J |
| CI1Fung | $\forall i, k_{MOS,i} = ks \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys$ | Modulation de $f_{fungMOS}$ | $\theta_{dyn}Témoin + \gamma$ et r |
| CBase | $\forall i, k_{MOS,i} = ks, Y_{MOS,i} = Ys$ | Comparaison qualité | $\theta_{dyn}Témoin$ |
| CI3-Fung | $\forall i, k_{MOS,i} = ks \cdot f_{CN}(i) \cdot f_P(i), Y_{MOS,i} = Ys \cdot f_{bact}(i)$ | Comparaison qualité | $\theta_{dyn}Témoin$ |
| CI3-Bact | $\forall i, k_{MOS,i} = ks \cdot f_{CN}(i) \cdot f_P(i) \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys$ | Comparaison qualité | $\theta_{dyn}Témoin$ |
| CI3-Pav | $\forall i, k_{MOS,i} = ks \cdot f_{CN}(i) \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys \cdot f_{bact}(i)$ | Comparaison qualité | $\theta_{dyn}Témoin$ |
| CI3-CN | $\forall i, k_{MOS,i} = ks \cdot f_P(i) \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys \cdot f_{bact}(i)$ | Comparaison qualité | $\theta_{dyn}Témoin$ |
| CI4 | $\forall i, k_{MOS,i} = ks \cdot f_{CN}(i) \cdot f_P(i) \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys \cdot f_{bact}(i)$ | Comparaison qualité | $\theta_{dyn}Témoin$ |
| Modèle pour les sols amendés | | | |
| AI1FMOF | $\forall i, k_{MOS,i} = ks \cdot f_{fungMOF}(i), Y_{MOS,i} = Ys, k_{MOF_1,i} = kr_1, k_{MOF_2,i} = kr_2 \cdot f_{fungMOF}(i)$ | Modulation de $f_{fungMOF}$ | $\theta_{dyn}Amendé + \gamma$ et r |
| ABase | $\forall i, k_{MOS,i} = ks, Y_{MOS,i} = Ys, k_{MOF_1,i} = kr_1, k_{MOF_2,i} = kr_2$ | Comparaison qualité | $\theta_{dyn}Amendé$ |
| AI3-F[MOF] | $\forall i, k_{MOS,i} = ks \cdot f_P(i) \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys \cdot f_{bact}(i), k_{MOF_1,i} = kr_1 \cdot f_{pH}(i), k_{MOF_2,i} = kr_2 \cdot f_{pH}(i)$ | Comparaison qualité | $\theta_{dyn}Amendé$ |
| AI3-F[MOS] | $\forall i, k_{MOS,i} = ks \cdot f_P(i), Y_{MOS,i} = Ys \cdot f_{bact}(i), k_{MOF_1,i} = kr_1 \cdot f_{pH}(i), k_{MOF_2,i} = kr_2 \cdot f_{pH}(i) \cdot f_{fungMOF}(i)$ | Comparaison qualité | $\theta_{dyn}Amendé$ |
| AI3-Bact | $\forall i, k_{MOS,i} = ks \cdot f_P(i) \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys, k_{MOF_1,i} = kr_1 \cdot f_{pH}(i), k_{MOF_2,i} = kr_2 \cdot f_{pH}(i) \cdot f_{fungMOF}(i)$ | Comparaison qualité | $\theta_{dyn}Amendé$ |
| AI3-Pav | $\forall i, k_{MOS,i} = ks \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys \cdot f_{bact}(i), k_{MOF_1,i} = kr_1 \cdot f_{pH}(i), k_{MOF_2,i} = kr_2 \cdot f_{pH}(i) \cdot f_{fungMOF}(i)$ | Comparaison qualité | $\theta_{dyn}Amendé$ |
| AI3-pH | $\forall i, k_{MOS,i} = ks \cdot f_P(i) \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys \cdot f_{bact}(i), k_{MOF_1,i} = kr_1, k_{MOF_2,i} = kr_2 \cdot f_{fungMOF}(i)$ | Comparaison qualité | $\theta_{dyn}Amendé$ |
| AI4 | $\forall i, k_{MOS,i} = ks \cdot f_P(i) \cdot f_{fungMOS}(i), Y_{MOS,i} = Ys \cdot f_{bact}(i), k_{MOF_1,i} = kr_1 \cdot f_{pH}(i), k_{MOF_2,i} = kr_2 \cdot f_{pH}(i) \cdot f_{fungMOF}(i)$ | Comparaison qualité | $\theta_{dyn}Amendé$ |

^aL'indice i signifie une distinction selon le sol considéré. Par exemple, l'équation $k_{MOS,i} = ks$ signifie que la constante de décomposition de la MOS est égale à ks quelque soit le sol.

^bIndique les paramètres à calibrer. $\theta_{dyn}Témoin$ et $\theta_{dyn}Amendé$ désignent l'ensemble des paramètres des cinétiques de dynamique du C pour le modèle appliqué à des sols témoins et des sols amendés respectivement (les différents k, Y, R_T, p), i.e tous les autres paramètres que ceux constituant les fonctions de modulation.

Pour le modèle appliqué aux sols témoins, la première combinaison correspond à un modèle de base (combinaison « CBase ») où aucune propriété du sol n'a été prise en compte, i.e. sans modulation des paramètres des cinétiques du C. En pratique, cela correspond à fixer les fonctions de modulation à une valeur de 1 dans les équations 4.6 à 4.9. Ces modèles permettent d'évaluer la qualité d'ajustement et de prédiction de base du modèle, i.e. lorsque l'on ne module pas les paramètres en fonction des propriétés des sols. Il est attendu que les qualités d'ajustement et de prédiction soient plus faibles qu'avec la prise en compte des modulations des paramètres.

Pour comparer la prise en compte des différentes modulations, une deuxième combinaison (« CI4 ») consiste à prendre en compte toutes les propriétés (i.e. toutes les fonctions de modulation) intervenant dans le modèle. Les autres combinaisons consistent à prendre en compte toutes les propriétés des sols sauf une pour évaluer la perte ou le gain de qualité suite au retrait de cette propriété (i.e. de sa fonction de modulation). Ainsi, on définit les combinaisons « CI3-Fung », « CI3-Bact », « CI3-Pav » et « CI3-CN » correspondant aux modèles prenant en compte toutes les fonctions de modulation moins celle concernant la diversité fongique, la diversité bactérienne, la disponibilité en P et le ratio C/N, respectivement.

Selon la même démarche, pour le modèle appliqué aux sols amendés, on définit la combinaison « ABase » comme le modèle sans prise en compte de fonctions de modulation, les combinaisons « AI3-F[MOF] », « AI3-F[MOS] », « AI3-Bact », « AI3-Pav » et « AI3-pH » prenant en compte les fonctions de modulation de toutes les propriétés considérées des sols moins $f_{fungMOF}$, $f_{fungMOS}$, f_{bact} , f_P et f_{pH} , respectivement et la combinaison « AI4 » prenant en compte toutes les modulations possibles.

L'ensemble des combinaisons, les modulations correspondantes, leur utilisation et les paramètres à calibrer correspondants sont résumés dans le tableau 4.2.

Critères de qualité d'ajustement et de qualité prédictive des modèles calibrés

Pour évaluer les modèles, et ainsi déterminer le gain associé à la prise en compte des différentes propriétés du sol, et en particulier des indices de diversité microbienne, le jeu de données MOSAIC est divisé en deux : 80% des sols (65 sols de calibration) sont utilisés pour la calibration des paramètres et 20% (tirage aléatoire) des sols (16 sols de validation) n'interviennent pas dans la calibration et constituent un jeu de données indépendant. De cette façon, en calculant des critères adéquats, on peut différencier la qualité d'ajustement des modèles, i.e. à quel point les modèles sont capables de simuler les dynamiques du C des sols qui ont servi à la calibration, et

la qualité prédictive, i.e. à quel point les modèles sont capables de prédire les dynamiques du C de nouveaux sols. Par définition, il est attendu que la qualité d'ajustement soit supérieure ou égale à la qualité de prédiction des modèles.

Les critères utilisés sont au nombre de deux et sont calculés pour chaque sol et chaque type de données (calibration vs validation). Ils sont calculés à chaque date de mesure t .

D'une part, le biais relatif est calculé selon l'équation :

$$pBias_{i,t} = \frac{1}{J} \sum_j \frac{sim_{i,t} - obs_{i,j,t}}{obs_{i,j,t}} \quad (4.17)$$

D'autre part, le coefficient d'efficience du modèle de Nash-Sutcliffe est calculé selon l'équation (Nash and Sutcliffe, 1970) :

$$E_{i,t} = 1 - \frac{\sum_j (sim_{i,t} - obs_{i,j,t})^2}{\sum_j (\overline{obs}_{\dots,t} - obs_{i,j,t})^2} \quad (4.18)$$

où $pBias_{i,t}$ est la valeur du biais relatif pour le sol i à l'instant t , $E_{i,t}$ est la valeur du coefficient d'efficience pour le sol i à l'instant t , $sim_{i,t}$ est la simulation de la variable de sortie issue du modèle pour le sol i à l'instant t , $obs_{i,j,t}$ est la valeur observée de la même variable pour le sol i à l'instant t et pour la répétition j et $\overline{obs}_{\dots,t}$ est la valeur moyenne sur l'ensemble des sols de la même variable à l'instant t . Les valeurs observées sont les flux cumulés de CO₂ non marqué pour les sols témoins. Pour le modèle appliqué aux sols amendés, les indices sont calculés séparément pour les flux marqués et non marqués. Si la valeur du biais relatif est négative, le modèle sous-estime les valeurs observées, si elle est positive, le modèle sur-estime les valeurs observées. De plus le biais relatif est exprimé en proportion de la valeur observée.

L'indice $E_{i,t}$ mesure où se situe le modèle en termes de qualité entre un modèle moyen qui correspond à une cinétique estimée par la moyenne de tous les sols à chaque instant et un modèle parfait qui simule sans aucune erreur les valeurs observées. Une valeur de 0 signifie que le modèle n'est pas meilleur que l'estimation par une moyenne générale. Une valeur de 1 signifie que le modèle est parfait. Si les valeurs sont négatives, le modèle estime/prédit moins bien qu'une moyenne générale.

Pour comparer de manière synthétique toutes les valeurs de $E_{i,t}$ entre les modèles, on trace les courbes de probabilité cumulée par modèle et par date de mesure, i.e. les courbes qui représentent la probabilité d'être inférieur à une valeur particulière de $E_{i,t}$, comme présenté en figure 4.4. Le pire modèle (courbe noire) est le modèle qui donne des valeurs de $E_{i,t}$ tout le temps égales à

$-\infty$. À l’opposé, le modèle parfait (la courbe rouge) est le modèle qui donne toutes les valeurs de E égales à 1 (estimation/prédiction parfaite).

Les modèles que l’on compare sont compris entre ces deux extrêmes (à l’instar des courbes orange et verte) et plus la courbe se rapproche de celle du modèle parfait (courbe verte), meilleur est le modèle. Pour comparer ces courbes, on calcule l’aire au-dessus de la courbe pour les valeurs de E entre 0 et 1 (A). Plus A est grand, plus la probabilité d’avoir des valeurs de E supérieures à 0 et proche de 1 est grande (i.e. meilleure que l’estimation par la moyenne).

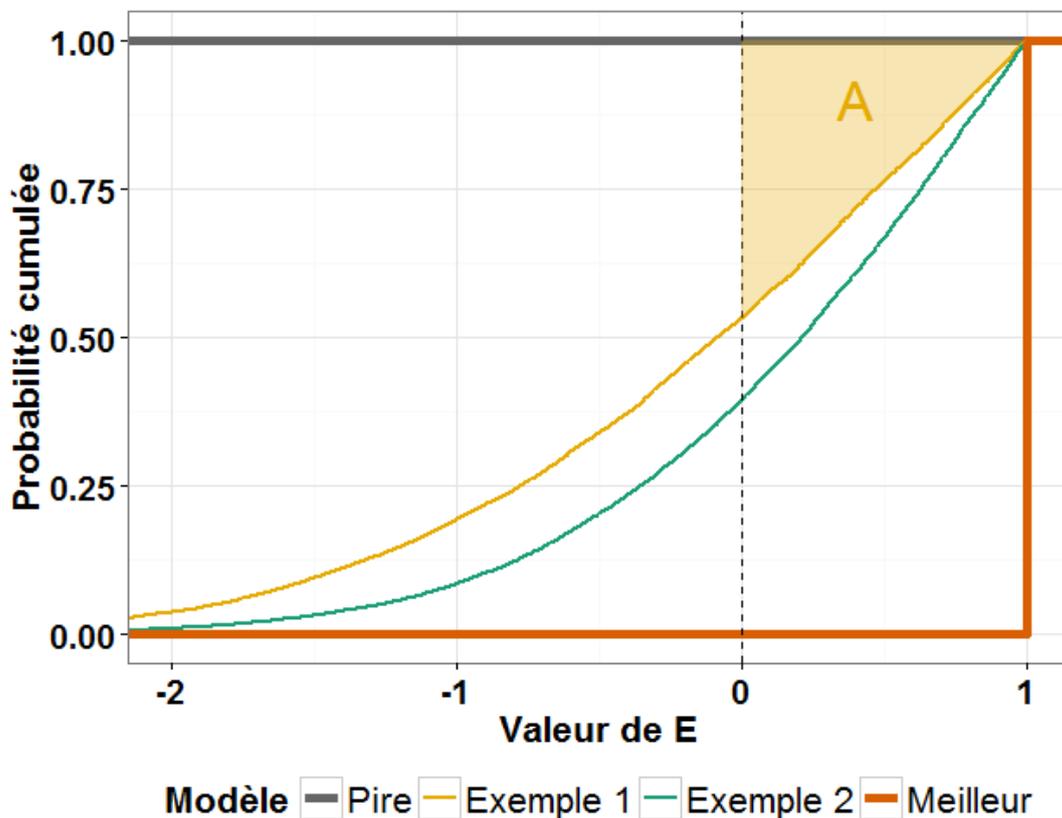


FIGURE 4.4 – Illustration de courbes de probabilité cumulée du coefficient E de Nash-Sutcliffe représentant la qualité des modèles.

Un point d’une courbe représente la probabilité (axe des ordonnées) d’obtenir des valeurs de E inférieures aux points correspondants sur l’axe des abscisses. Par exemple, la probabilité d’obtenir des valeurs de E en dessous de 0 avec le modèle « Exemple 1 » est de 0,53. La courbe noire représente le pire modèle, i.e. qui donne des valeurs de E tout le temps égales à $-\infty$. À l’opposé, la courbe rouge représente le modèle parfait où toutes les valeurs obtenues de E sont égales à 1. Les courbes oranges et vertes représentent des modèles intermédiaires où le modèle « Exemple 2 » (courbe verte) est meilleur que le modèle « Exemple 1 » (courbe orange). Pour comparer ces modèles, l’aire au-dessus de la courbe pour les valeurs de E entre 0 et 1 (A) est calculée. Plus cette aire est importante, plus la probabilité d’avoir des valeurs de E supérieures à 0 est grande.

4.3 Résultats

4.3.1 Distributions *a priori* et paramètres fixes

4.3.1.1 Distributions *a priori* des paramètres

Les distributions *a priori* des différents paramètres sont visibles dans le tableau 4.3. Le choix de la loi de distribution utilisée, ainsi qu'un coefficient de variation correspondant au ratio de l'espace inter-quartile sur la valeur de la médiane (CVQ), nous permettent de classer ces distributions selon le niveau d'information qu'elles apportent. Comme une loi uniforme est utilisée pour tous les paramètres des fonctions de modulation, les lois *a priori* de ces paramètres sont considérées comme peu informatives. D'autre part, en raison d'un coefficient de variation supérieur à 1, les lois des paramètres ks , R_T et $sd13$ sont également peu informatives.

Les coefficients des distributions des paramètres Ys , kr_2 et p_1 étant compris entre 0,2 et 1, ces lois sont considérées moyennement informatives. Pour les autres paramètres ($sd12$, p_2 et Yr_1) les lois sont considérées informatives. Une loi informative aura plus d'impact sur la distribution *a posteriori* qu'une loi non-informative.

4.3.1.2 Analyse de sensibilité globale *a priori* et choix des paramètres fixes

D'après l'analyse de sensibilité globale *a priori*, mis à part le paramètre r (fonction f_{MOF}), les deux paramètres les plus importants sont les constantes de décomposition de la MOS (ks) et de la MOF_2 (kr_2) (figure 4.5). La valeur plus faible pour le paramètre R_T vient du fait que celui-ci est testé pour l'ensemble des flux marqués et non marqués. Une décomposition entre flux montre que ce paramètre est très important pour le flux non marqué et très peu important pour le flux marqué (résultats non montrés). De plus, lorsqu'il est calibré, il a une valeur très proche de 0 (10^{-7}) dans le cas des sols amendés, ce qui est, d'un point de vue mécaniste, très peu probable. Ce paramètre sera donc estimé pour le modèle appliqué aux sols témoins et fixé pour le modèle appliqué aux sols amendés, à une valeur correspondant à la médiane de sa distribution *a posteriori* (0,0025) dans le modèle appliqué aux sols témoins sans modulation (CBase).

TABLE 4.3 – Description des paramètres potentiellement présents dans les modèles et distributions *a priori* associées.

| Paramètres | Description | Flux ^a | Priors | [Min ; Max] - Médiane ^b | CVQ ^c |
|---|--|----------------------|----------------------------------|---------------------------------------|------------------|
| Paramètres présents dans tous les modèles | | | | | |
| ks | Constante de décomposition du pool MOS | ^{12}C | $\log\mathcal{N}(-6, 9; 1, 8)$ | $[0; +\infty] - 1.10^{-3}$ | 3,1 |
| Ys | Rendement d'utilisation du C du pool MOS | ^{12}C | $Beta(1, 7; 3)$ | $[0; 0,65] - 0,31$ | 0,8 |
| R_T | Respiration microbienne autre que pour la croissance | ^{12}C et ^{13}C | $Beta(2, 1; 509)$ | $[0; +\infty] - 3,4.10^{-3}$ | 1 |
| $sd12$ | Écart-type des flux de CO_2 non marqué | - | $\Gamma(0, 44; 166)$ | $[0; +\infty] - 1.10^{-3}$ | 0,2 |
| Paramètres présents dans les modèles pour les sols amendés | | | | | |
| p_1 | Proportion de C qui forme le pool MOF_1 | ^{13}C | $Beta(13; 85)$ | $[0; +\infty] - 0,13$ | 0,4 |
| p_2 | Proportion de C qui forme le pool MOF_2 | ^{13}C | $Beta(75; 23)$ | $[0; +\infty] - 0,77$ | 0,1 |
| kr_1 | Constante de décomposition du pool MOF_1 | ^{13}C | $\log\mathcal{N}(-1, 4; 0, 6)^d$ | $[0; +\infty] - 0,25$ | 0,8 |
| Yr_1 | Rendement d'utilisation du C de pool MOF_1 | ^{13}C | $Beta(25; 25)$ | $[0; 0,65] - 0,5$ | 0,2 |
| kr_2 | Constante de décomposition du pool MOF_2 | ^{13}C | $\log\mathcal{N}(-3, 2; 0, 5)^e$ | $[0; +\infty] - 4,1.10^{-2}$ | 0,7 |
| Yr_2 | Rendement d'utilisation du C de pool MOF_2 | ^{13}C | $Beta(26; 42)$ | $[0; 0,65] - 0,38$ | 0,2 |
| $sd13$ | Écart-type des flux de CO_2 marqué | - | $\Gamma(1; 70)$ | $[0; +\infty] - 1.10^{-2}$ | 1,6 |
| Paramètres présents selon les fonctions de modulation à calibrer | | | | | |
| a | Paramètre de f_{bact} - Modèle CI1Bact | ^{12}C | $\mathcal{U}(0; 100)$ | $[0; 100] - 50$ | - |
| J | Paramètre de f_{bact} - Modèle CI1Bact | ^{12}C | $\mathcal{U}(50; 150)$ | $[50; 150] - 100$ | - |
| γ | Paramètre de $f_{fungMOS}$ et $f_{fungMOF}$ - Modèles CI1Fung et AI1FMOF | ^{12}C et ^{13}C | $\mathcal{U}(0; 50)$ | $[0; 50] - 25$ | - |
| r | Paramètre de $f_{fungMOS}$ et $f_{fungMOF}$ - Modèles CI1Fung et AI1FMOF | ^{12}C et ^{13}C | $\mathcal{U}(0; 1)$ | $[0; 1] - 0,5$ | - |
| b | Paramètre de f_{CN} - Modèle CI1CN | ^{12}C | $\mathcal{U}(2; 10)$ | $[2; 10] - 6$ | - |
| c | Paramètre de f_{Pav} - Modèle CI1Pav | ^{12}C | $\mathcal{U}(0, 1; 0, 9)$ | $[0,1; 0,9] - 0,45$ | - |
| d | Paramètre de f_{Pav} - Modèle CI1Pav | ^{12}C | $\mathcal{U}(20; 40)$ | $[20; 40] - 30$ | - |

$\log\mathcal{N}$, loi log-Normal ; $Beta$, loi bêta ; Γ , loi gamma ; \mathcal{U} , loi uniforme

^a Flux sur lesquels le paramètre a une influence : ^{12}C , flux non marqué ; ^{13}C , flux marqué

^b Les bornes supérieures de Ys , Yr_1 et Yr_2 ont été imposées

^c Coefficient de variation des quantiles, i.e. ratio entre l'espace inter-quartile et la médiane. Cet indice étant peu pertinent pour les lois uniformes, il n'est pas calculé.

^{d,e} Les écart-types ont été manuellement réduits car considérés trop grands.

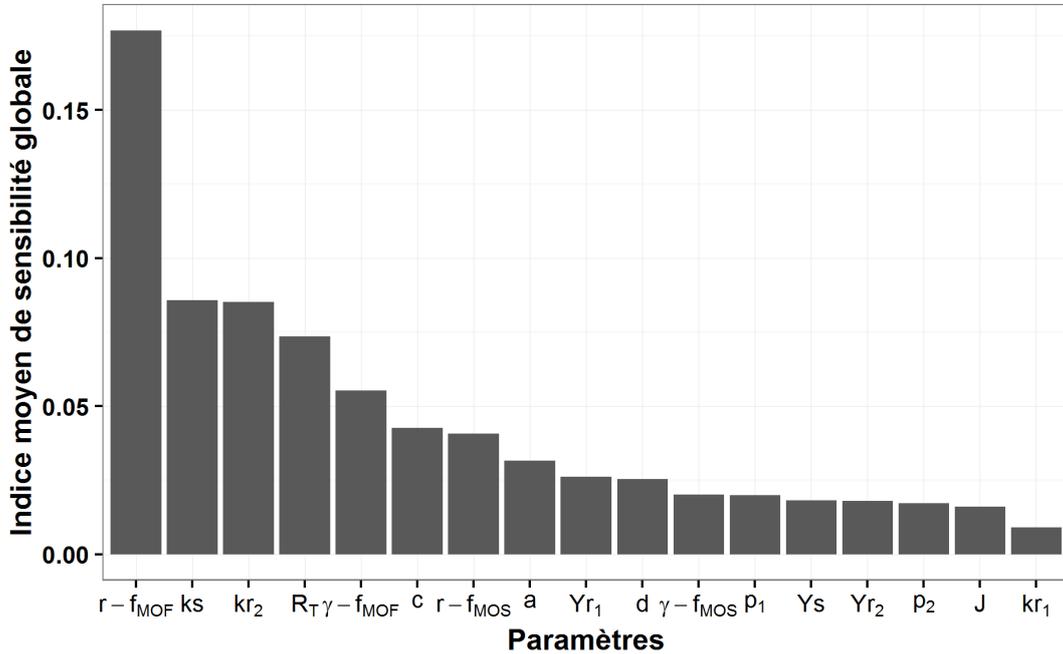


FIGURE 4.5 – Indices SI de sensibilité globale des différents paramètres du modèle de dynamique du C.

Les paramètres sont : ks , kr_1, kr_2 , les constantes de décomposition des pools MOS , MOF_1 , MOF_2 , respectivement ; Ys , Yr_1, Yr_2 , les rendements d'utilisation du C des pools MOS , MOF_1 , MOF_2 , respectivement ; p_1 et p_2 , les proportions de C de la MOF attribué à MOF_1 , MOF_2 , respectivement ; R_T , la respiration microbienne autre que pour la croissance ; a et J , les paramètres de la fonction f_{bact} ; b , le paramètre de la fonction f_{CN} ; c et d , les paramètres de la fonction f_P ; $\gamma - f_{MOS}$ et $r - f_{MOS}$, les paramètres de la fonction $f_{fungMOS}$; $\gamma - f_{MOF}$ et $r - f_{MOF}$, les paramètres de la fonction $f_{fungMOF}$.

Globalement, la faible sensibilité du modèle à l'ensemble des paramètres des fonctions de modulation conforte le choix de fixer ces paramètres après l'étape permettant leur calibration. Ce n'est pas le cas du paramètre r pour la fonction f_{MOF} pour lequel le modèle a la plus grande sensibilité. Une attention particulière devra être donnée à la distribution *a posteriori* pour ce paramètre. La valeur prise pour fixer les paramètres des fonctions de modulation dépend de leur distribution *a posteriori* dans les modèles où ils sont calibrés.

Tous les autres paramètres ont également des valeurs faibles d'indice de sensibilité et peuvent être fixés quelle que soit l'étape de calibration (i.e. calibration des fonctions de modulation et évaluation de la prise en compte des fonctions de modulation). Cependant, il apparaît difficile de bien simuler les flux marqués seulement avec le paramètre kr_2 (qui détermine la vitesse à laquelle le plateau du flux est atteint) sans un autre paramètre qui détermine la valeur du plateau atteint par le flux cumulé, i.e. Yr_2 ou p_2 . Ayant plus d'information sur le paramètre p_2 (en termes de nombre de données pour construire le prior et la valeur du CVQ), c'est ce paramètre que l'on

choisit de calibrer.

Les paramètres restants, i.e. p_1 , kr_1 , Yr_1 et Yr_2 , sont fixés lors de la calibration des fonctions de modulation ainsi que lors de la calibration des modèles avec les différentes combinaisons de prise en compte des fonctions de modulation pour évaluation sur la qualité du modèle. Les paramètres p_1 et kr_1 sont fixés à la médiane de leur distribution *a priori* (respectivement 0,13 et 0,25), et les paramètres Yr_1 et Yr_2 sont tous deux fixés à 0,6 ; valeur souvent utilisée dans la littérature (e.g. Garnier *et al.*, 2001).

4.3.2 Calibration des fonctions de modulation des paramètres des cinétiques de flux de C

4.3.2.1 Diagnostic de convergence et distributions *a posteriori* des paramètres des fonctions de modulation

TABLE 4.4 – Diagnostic de convergence des modèles utilisés pour calibrer les fonctions de modulation.

| Modèle | Gelman-Rubin | Heidelberger-Welsh | | |
|----------------|--------------|--------------------|----------|----------|
| | | Chaîne 1 | Chaîne 2 | Chaîne 3 |
| CI1Bact | ✓ | ✓ | ✓ | ✓ |
| CI1CN | ✓ | ✓ | ks^a | Ys^a |
| CI1Fung | ✓ | ✓ | ✓ | ✓ |
| CI1Pav | ✓ | ✓ | ✓ | ✓ |
| AI1Fung | ✓ | ✓ | ✓ | ✓ |

^a Ces paramètres ont montré un défaut de convergence ponctuel (i.e. pour une seule chaîne)

Le tableau 4.4 renseigne sur les paramètres qui n’ont pas convergé lors de l’inférence bayésienne pour la calibration des paramètres des fonctions de modulation (modèles « CI1 » et modèle « AI1 »). La plupart des chaînes ont convergé selon les deux tests utilisés exceptée la chaîne 2 pour le paramètres ks et la chaîne 3 pour le paramètre Ys dans le modèle CI1CN. Ce défaut de convergence n’a pas été amélioré après avoir triplé le nombre de simulations (jusqu’à 300000, pour environ 50 heures de temps de calcul). Ceci est probablement dû à la forte corrélation qu’il existe entre ks et b (corrélation = 0,9) imputable à la construction même de la modulation de k_{MOS} par une fonction dépendante du ratio C/N. Cette modulation étant égal à $k_{MOS} = ks \times \frac{b}{C/N}$, il est difficile de séparer les paramètres ks et b (multiplication directe). Cependant, le défaut de convergence pour les paramètres ne concernant qu’une chaîne à chaque fois, ces modèles ayant pour but de calibrer les paramètres des fonctions de modulation et considérant qu’il sera difficile d’atteindre une meilleure convergence, cette calibration est gardée.

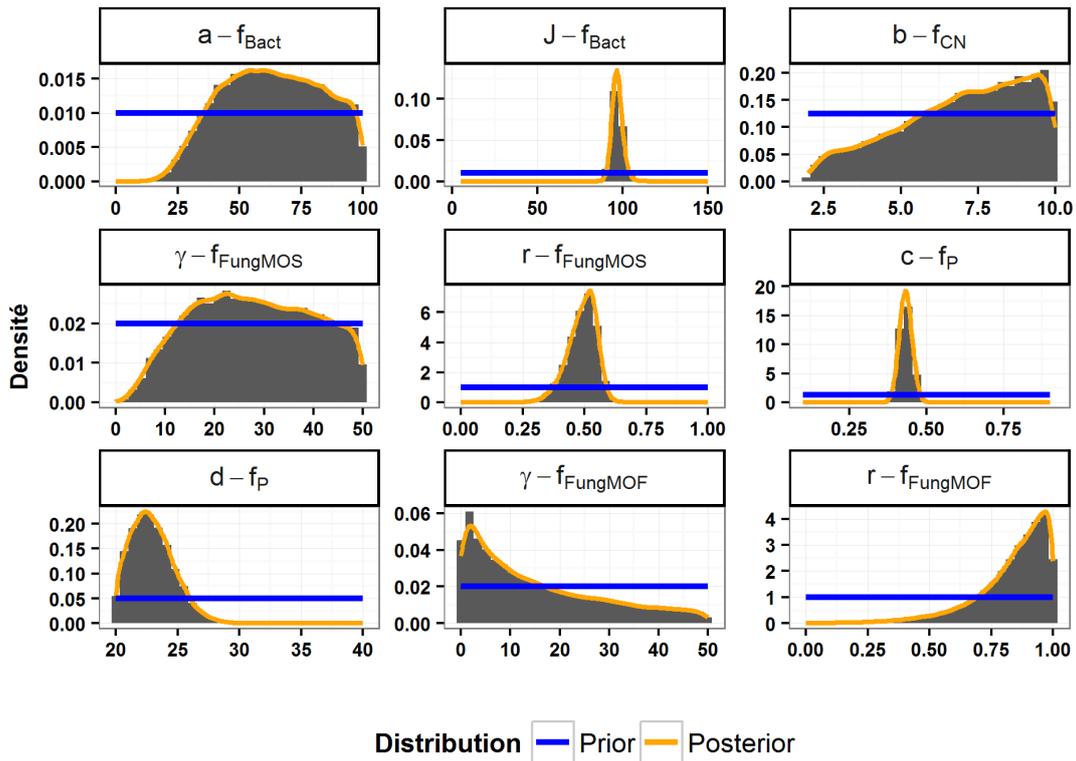


FIGURE 4.6 – Distributions *a priori* (lignes bleues) et *a posteriori* (histogrammes et lignes oranges) des paramètres des fonctions de modulation.

L'analyse des distributions *a posteriori* des paramètres des fonctions de modulation permet de classer ces paramètres en 3 groupes (Figure 4.6) :

- Les paramètres J (fonction f_{Bact}), r (fonction $f_{FungMOS}$), c et d (fonction f_P) ont des distributions *a posteriori* avec une dispersion assez faible donnant une calibration assez précise de leur valeur (CVQ = 0,04 ; 0,16 ; 0,06 et 0,11 respectivement).
- À l'inverse, l'inférence des paramètres a (fonction f_{Bact}) et γ (fonction $f_{FungMOS}$) semble n'avoir apporté que peu d'informations supplémentaires (CVQ = 0,51 et 0,71 respectivement).
- Le dernier groupe, constitué des paramètres b (fonction f_{CN}) et γ et r (fonction $f_{FungMOF}$), a des distributions *a posteriori* apportant peu d'information supplémentaire (CVQ = 0,44 ; 1,79 et 0,20 respectivement) et qui semblent avoir été contraintes par les bornes des distributions *a priori* suivant une loi uniforme ; en particulier pour les paramètres b , contraint à sa borne supérieure 10 et r , contraint à sa borne supérieure 1.

4.3.2.2 Formes estimées des fonctions de modulation

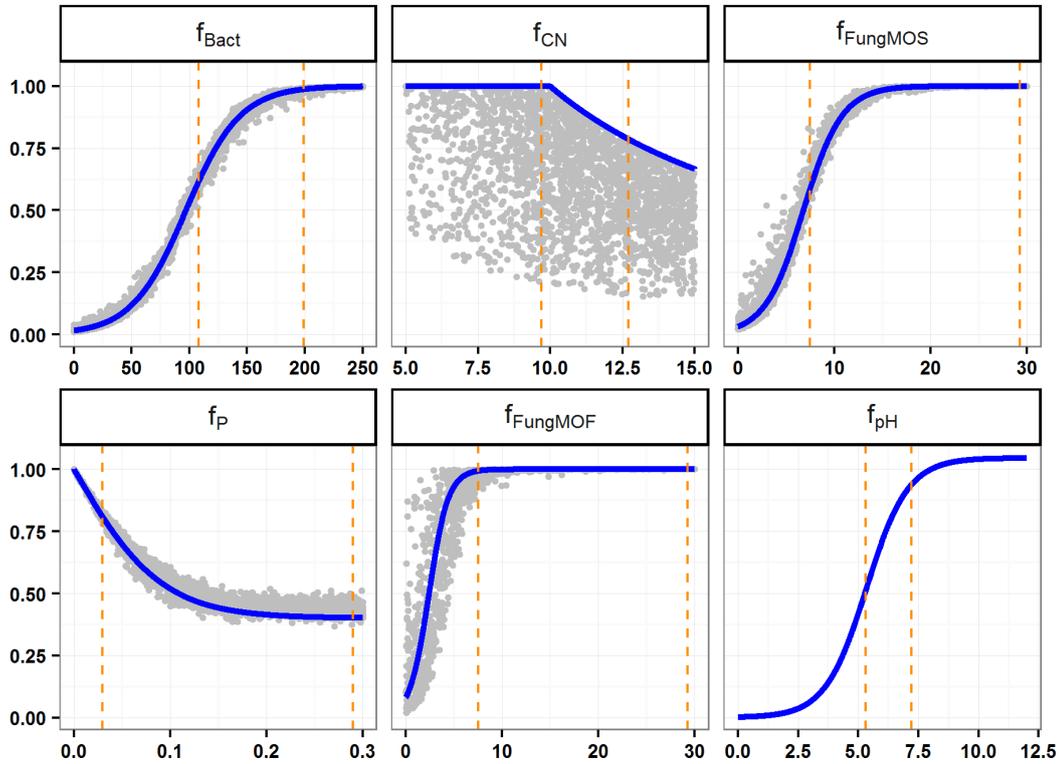


FIGURE 4.7 – Variabilité des fonctions de modulation estimées (points gris) et fonctions de modulation choisies (lignes bleues).

Les lignes verticales oranges représentent les valeurs minimales et maximales dans le jeu de données MOSAIC des propriétés du sol dont dépendent les fonctions de modulation. Pour f_{Bact} , la diversité est exprimée selon l'inverse de l'indice de Simpson bactérien et pour $f_{FungMOS}$ et $f_{FungMOF}$, la diversité est exprimée selon l'inverse de l'indice de Simpson fongique.

La figure 4.7 donne les valeurs des fonctions de modulation calibrées (points gris) ainsi que les formes des fonctions retenues pour la suite de l'analyse (lignes bleues). Les points gris ont été obtenus par tirage aléatoire de valeurs des paramètres dans les distributions *a posteriori*. Pour rappel, les indices de diversité qui interviennent dans les fonctions de modulation de la décomposition sont l'inverse de l'indice de Simpson ($1/D$) bactérien dans la fonction f_{Bact} et l'inverse de l'indice de Simpson ($1/D$) fongique dans les fonctions $f_{FungMOS}$ et $f_{FungMOF}$. La fonction liée au pH (non calibrée, définie d'après [Leifeld et al., 2008](#)) est également montrée.

Globalement, les fonctions n'ont pas une grande variabilité. En particulier, on peut noter que malgré la variabilité *a posteriori* des paramètres a (fonction f_{Bact}) et γ (fonction $f_{FungMOS}$), cette variabilité ne se retrouve pas dans les fonctions associées. Ceci s'explique par le fait que ces paramètres contrôlent les ordonnées à l'origine des fonctions f_{Bact} et $f_{FungMOS}$ dont les valeurs

sont égales à $1/(1+a)$ et $1/(1+\gamma)$, respectivement. Ainsi, les valeurs de ces ordonnées à l'origine atteignent rapidement une asymptote avec l'augmentation de la valeur de chacun des paramètres. À partir d'une certaine valeur du paramètre, la différence entre deux valeurs d'ordonnée à l'origine pour deux valeurs du paramètre est faible (e.g. pour des valeurs de a de 25 et 30, les valeurs d'ordonnée à l'origine de f_{Bact} sont respectivement de 0,038 et 0.032). Les bornes des valeurs des propriétés du sol dont dépendent ces deux fonctions (Figure 4.7, lignes oranges) montrent que les valeurs de l'inverse de l'indice de Simpson bactérien et fongique proches de l'ordonnée à l'origine ne sont pas disponibles dans le jeu de données MOSAIC, ce qui explique la difficulté d'estimer avec précision ces paramètres (Figure 4.6).

Il est intéressant de noter que l'estimation des paramètres de la fonction $f_{FungMOF}$ conduit à une fonction dont l'asymptote est déjà atteinte dans la gamme de $1/D$ fongique du jeu de données MOSAIC. Cette observation est vraie même en considérant la variabilité des distributions *a posteriori* de ses paramètres (i.e. points gris). Cette fonction n'aura donc vraisemblablement aucune influence sur les simulations de la dynamique du C lié à la MOF, contrairement à l'hypothèse faite. La sensibilité élevée du modèle au paramètre r de cette fonction n'a donc plus d'importance.

La seule fonction dont l'estimation des paramètres conduit à une fonction avec une grande variabilité et celle liée au C/N. Une valeur de $b=10$ est choisie, en tant que borne supérieure qui a contraint la distribution *a posteriori*.

4.3.3 Qualités des modèles et évaluation des combinaisons des fonctions de modulation

Dans cette section, les résultats concernent la comparaison des différentes combinaisons de prises en compte de fonction de modulation. Pour le modèle appliqué aux sols amendés, en plus du modèle de base, lors de la présentation des résultats concernant les flux non marqués (i.e. minéralisation de la MOS), seules les combinaisons de prise en compte des fonctions de modulation qui ont un impact sur ces flux sont présentées (i.e. AI3-Bact, AI3-Pav, AI3-F[MOS] et AI4). Lors de la présentation des résultats concernant les flux marqués (i.e. minéralisation de la MOF), seules les combinaisons de prise en compte des fonctions de modulation qui ont un impact sur ces flux sont présentées (i.e. AI3-pH, AI3-F[MOF] et AI4).

4.3.3.1 Diagnostic de convergence et distributions *a posteriori* des paramètres

Aucun problème de convergence n'a été détecté lors des deux tests pour tous les paramètres et tous les modèles appliqués aux sols témoins et aux sols amendés. La figure 4.8 montre les distributions *a posteriori* des paramètres des cinétiques du C dans le modèle avec les différentes combinaisons de modulation et appliqué aux sols témoins et amendés.

Paramètres liés à la minéralisation du C de la MOS

Les distributions *a posteriori* du paramètre ks ont toutes la majorité de leurs valeurs plus élevées que la distribution proposée *a priori*. Les valeurs de ce paramètre sont les plus distinctes entre les différentes combinaisons de modulation et augmentent selon cet ordre :

1. Le modèle avec toutes les modulations sauf celle liée au P assimilable (CI3-Pav ou AI3-Pav, médianes à 0,012 et 0,019 respectivement),
2. Le modèle de base (CBase ou ABase, médianes à 0,012 et 0,027 respectivement),
3. Les modèles avec toutes les modulations (CI4 ou AI4, médianes à 0,020 et 0,039 respectivement) et toutes les modulations sauf celle liée au ratio C/N et à la diversité en champignons (CI3-CN et CI3-Fung ou AI3-F_{MOF}, médianes à 0,018 et 0,020 ou 0,039 respectivement)
4. Le modèle avec toutes les modulations sauf celle liée à la diversité bactérienne (CI3-Bact ou AI3-Bact, médianes à 0,031 et 0,058 respectivement).

Les valeurs du paramètre ks doublent entre le modèle appliqué aux sols témoins et le modèle appliqué aux sols amendés.

Les valeurs de Y_s ont été contraintes par sa borne supérieure (0,65) sauf lorsque la modulation liée à la diversité bactérienne a été prise en compte dans le modèle appliqué aux sols témoins (i.e. CI3-Fung, CI3-CN, CI3-Pav, CI4; médiane entre 0,57 et 0,62). Pour le modèle appliqué aux sols amendés, la contrainte par la borne supérieure se retrouve dans toutes les fonctions de modulation sans exception.

Concernant l'écart-type des flux (sd_{12}), les valeurs des distributions *a posteriori* sont bien supérieures à celles de la distribution *a priori* pour tous les modèles, avec globalement une diminution de ces valeurs lorsque l'on prend en compte des modulations des paramètres. Les différences sont moins importantes pour le modèle appliqué aux sols amendés pour lesquels les valeurs de ce paramètre sont globalement plus élevées.

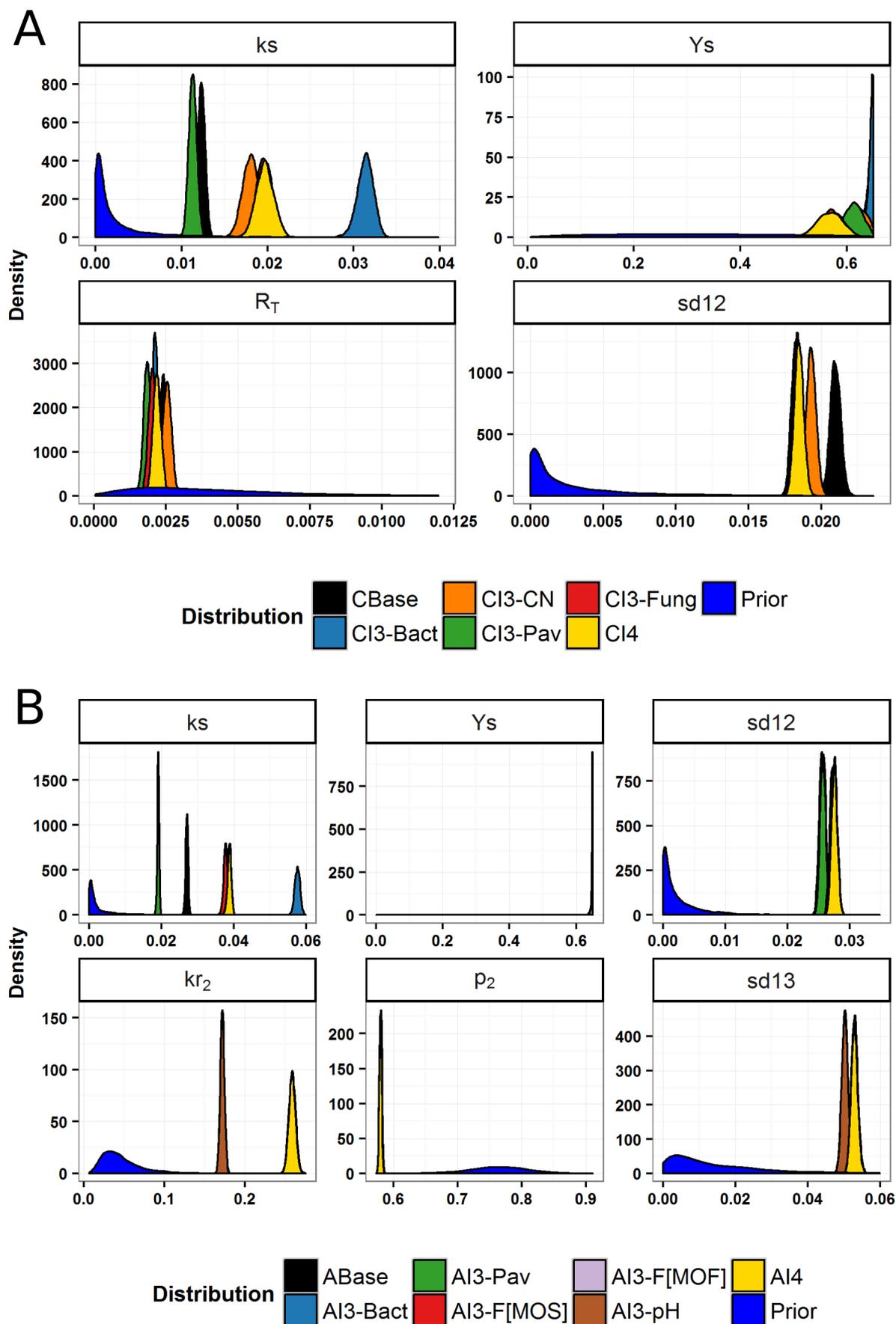


FIGURE 4.8 – Distributions *a priori* et *a posteriori* des paramètres des cinétiques de décomposition pour le modèle appliqué aux sols témoins (A) et aux sols amendés (B) en fonction des différentes modulations prises en compte.

Le paramètre R_T n'a été estimé que pour le modèle appliqué aux sols témoins. Les distributions *a posteriori* sont assez proches les unes des autres et de la distribution *a priori*.

Paramètres liés à la minéralisation du C de la MOF

De la même manière que la constante de décomposition du C de la MOS, les distributions *a posteriori* de celle du C du sous pool MOF₂ (kr_2) ont des valeurs plus élevées que la distribution *a priori*. Les modèles qui ne prennent pas en compte la modulation liée au pH (ABase et AI3-pH) ont des valeurs plus faibles (médiane à 0,17) que les autres. On retrouve cette même différence, beaucoup moins marquée, pour le paramètre $sd13$, écart-type des flux marqués, dont les valeurs sont également plus élevées que celles de la distribution *a priori*.

Quelle que soit la modulation prise en compte, les distributions *a posteriori* du paramètre p_2 sont confondues et leurs valeurs sont plus faibles d'environ 25% par rapport à celles de la distribution *a priori* (médiane à 0,58 vs 0,77).

4.3.3.2 Biais du modèle appliqué aux sols témoins et aux sols amendés quelles que soient les fonctions de modulation prises en compte

Biais des flux non marqués (MOS)

Qu'il soit appliqué aux sols témoins ou aux sols amendés, et quelles que soient les fonctions de modulation prises en compte, les valeurs du biais relatif de l'estimation des flux non marqués (i.e. minéralisation du C de la MOS) sont similaires dans le temps (Figure 4.9). Le biais est très négatif pour les estimations du début d'incubation, avec des valeurs entre -30% et -40% à 3 jours d'incubation, et croît progressivement pour atteindre des valeurs autour de 0% en fin d'incubation. Les valeurs sont globalement comparables si ce n'est que les biais négatifs du début d'incubation sont légèrement plus faibles (i.e. biais négatif plus important) lorsque des fonctions de modulation sont prises en compte. De plus, les valeurs sont légèrement plus proches de 0 pour le modèle appliqué aux sols amendés. Enfin, les valeurs entre les sols utilisés pour la calibration (ajustement) et ceux utilisés pour la validation (prédiction) sont sensiblement comparables.

Biais des flux marqués (MOF)

Pour les flux marqués, le biais est positif à la première date de mesure puis a des valeurs autour de 0 pour le reste de l'incubation, quelles que soient les fonctions de modulation prises en compte (Figure 4.9).

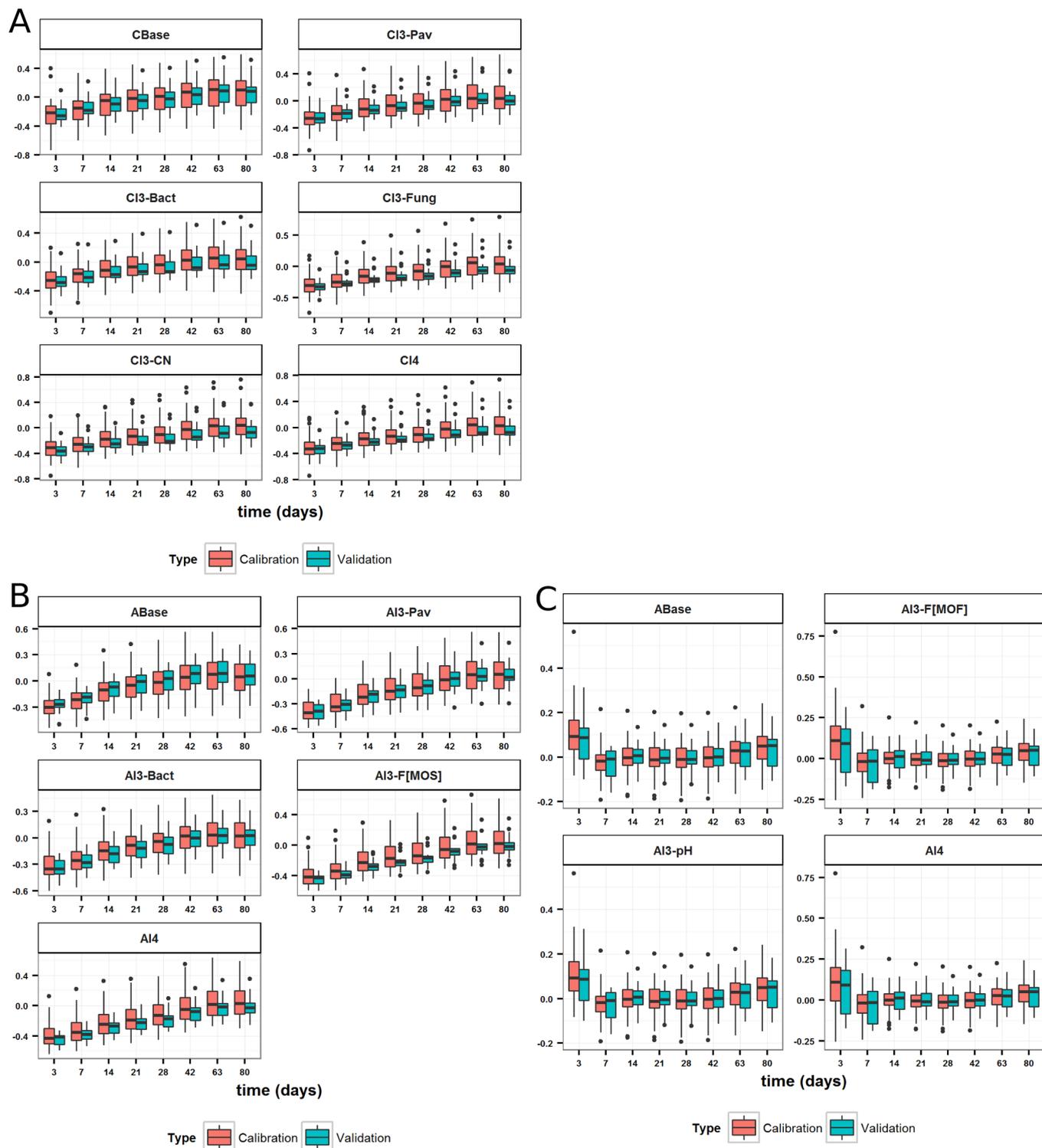


FIGURE 4.9 – Biais relatif d’ajustement (rouge) et de prédiction (bleu) des flux non marqués par le modèle appliqué aux sols témoins (A) et aux sols amendés (B) et des flux marqués (C).

4.3.3.3 Qualité du modèle de base

Flux non marqués (MOS) La qualité d'ajustement du modèle de base pour les flux non marqués, en sols témoins et en sols amendés, est meilleure en fin d'incubation qu'au début comme l'atteste les courbes de probabilités cumulées du coefficient E de Nash-Sutcliffe aux différentes dates de mesures (Figure 4.10). On retrouve cette observation pour la qualité de prédiction (validation).

Flux marqués (MOF) De la même manière, les qualités d'ajustement et de prédiction du modèle sont moins bonnes en tout début d'incubation (Figure 4.10).

4.3.3.4 Comparaison des qualités d'ajustement et de prédiction des modèles selon les combinaisons des fonctions de modulation

La Figure 4.11 montre les changements relatifs des aires (figure 4.4) au-dessus des courbes de probabilités cumulées du coefficient E de Nash-Sutcliffe pour les différentes situations (sols témoins vs amendés) et les différentes prises en compte des fonctions de modulation par rapport au modèle de base. Les modèles CI4 et AI4 nous renseignent sur la perte ou le gain de qualité lorsque toutes les fonctions de modulation sont prises en compte, en condition témoin et amendée respectivement. Les autres modèles représentent le modèle avec toutes les fonctions de modulation sauf une.

Ces modèles sont également à comparer au modèle où toutes les fonctions de modulation sont prises en compte pour voir l'évolution des qualités spécifique au retrait d'une fonction de modulation. La figure 4.12 montre des exemples de simulations suite à la calibration selon les différentes combinaisons de prise en compte des fonctions de modulation lorsque cette prise en compte améliore (A), n'a pas beaucoup d'impact (B) ou dégrade (C) l'ajustement ou la prédiction par rapport au modèle de base.

Flux associés à la minéralisation de la MOS (A et B, figure 4.11)

Lorsque le modèle est appliqué aux sols témoins (A, figure 4.11), la prise en compte des fonctions de modulation des paramètres dégrade globalement la qualité d'ajustement (calibration) des flux jusqu'à -34% du modèle de base et ceux jusqu'à 14 jours. À partir de cette date, la qualité d'ajustement est globalement améliorée jusqu'à 40% par rapport au modèle de base. Pour la qualité prédictive (validation), cette dégradation dure plus ou moins longtemps selon les

fonctions de modulation prises en compte et est plus élevée (jusqu'à -74%).

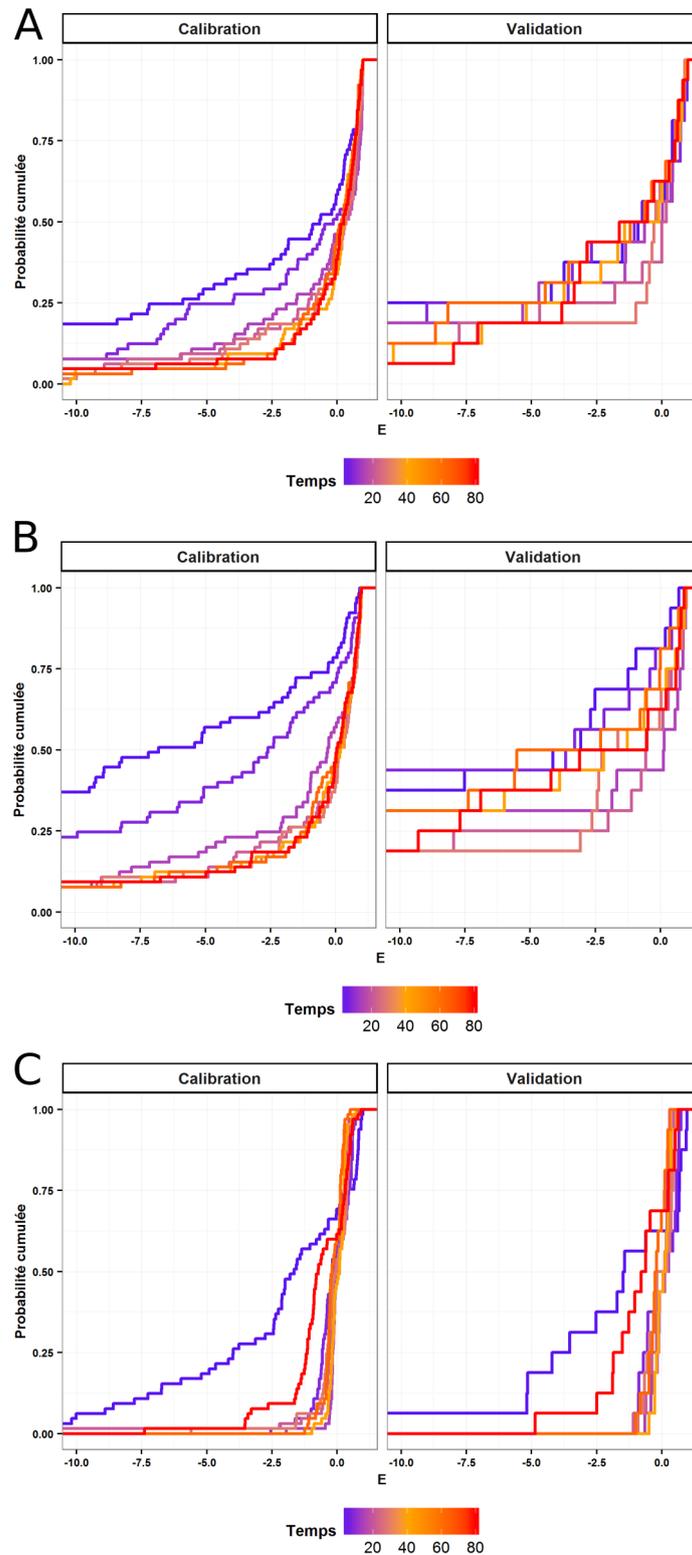


FIGURE 4.10 – Probabilités cumulées du coefficient E de Nash-Sutcliffe d’ajustement (calibration) et de prédiction (validation) du modèle sans modulation pour les flux non marqués du modèle appliqué aux sols témoins (A) et aux sols amendés (B) et pour les flux marqués (C) pour chaque date de mesure.

Cette dégradation est moins importante lorsque l'on ne prend pas en compte la fonction de modulation liée au P assimilable (CI3-Pav, vert ; jusqu'à -6% et -22% pour calibration et validation respectivement) ou celle liée à la diversité bactérienne (CI3-Bact, bleu ; jusqu'à -11% et -19%). Pour ces 2 prises en compte, dès 21 jours, l'observation n'est pas la même entre les qualités d'ajustement et de prédiction. La non prise en compte du P assimilable et, dans une moindre mesure, de la diversité bactérienne a dégradé la qualité d'ajustement comparé à leur prise en compte (i.e. modèle complet, CI4, jaune). La qualité prédictive est cependant améliorée par la non prise en compte du P assimilable (vert vs jaune, 67% vs 25% à 80 jours) alors qu'elle est systématiquement dégradée par la non prise en compte de la diversité bactérienne (bleu vs jaune, jusqu'à -28% vs 3% à 28 jours et 18% vs 25% à 80 jours).

Pour la fonction de modulation liée au ratio C/N, sa non prise en compte dégrade les qualités d'ajustement et de prédiction (saumon vs jaune). Pour la fonction de modulation liée à la diversité fongique, les valeurs des qualités d'ajustement et de prédiction lors de sa non prise en compte sont très proches du modèle complet (rouge vs jaune), exceptée pour la qualité prédictive aux deux dernières dates d'incubation qui double lorsque l'on ne prend pas en compte cette fonction de modulation (49% vs 21% à 63 jours et 53% vs 25% à 80 jours).

Dans le cas des sols amendés (B, figure 4.11), les observations sont sensiblement les mêmes. La dégradation au début d'incubation dure plus longtemps (jusqu'à 28 jours) et est plus élevée (jusqu'à -46% et -98% du modèle de base pour l'ajustement et la prédiction respectivement). L'amélioration en fin d'incubation est également plus élevée, en particulier pour la qualité prédictive (jusqu'à 26% et 196% du modèle de base pour l'ajustement et la prédiction respectivement).

Flux associés à la minéralisation de la MOF (C, figure 4.11)

Lorsque toutes les fonctions de modulation impactant la décomposition du résidu sont prises en compte (AI4), la qualité d'estimation est fortement dégradée jusqu'à 7 jours (jusqu'à -51% du modèle de base) et la qualité prédictive jusqu'à 21 jours (jusqu'à -66% du modèle de base). Après ces dates, les qualités oscillent entre faibles améliorations et dégradations (<10% en valeur absolue). La non prise en compte du pH (marron), et donc la seule prise en compte de la diversité fongique, donne des valeurs égales à celles du modèle de base. La non prise en compte de la diversité fongique ne change rien par rapport au modèle où elle est prise en compte. Le pH est responsable de la dégradation dans la première moitié d'incubation.

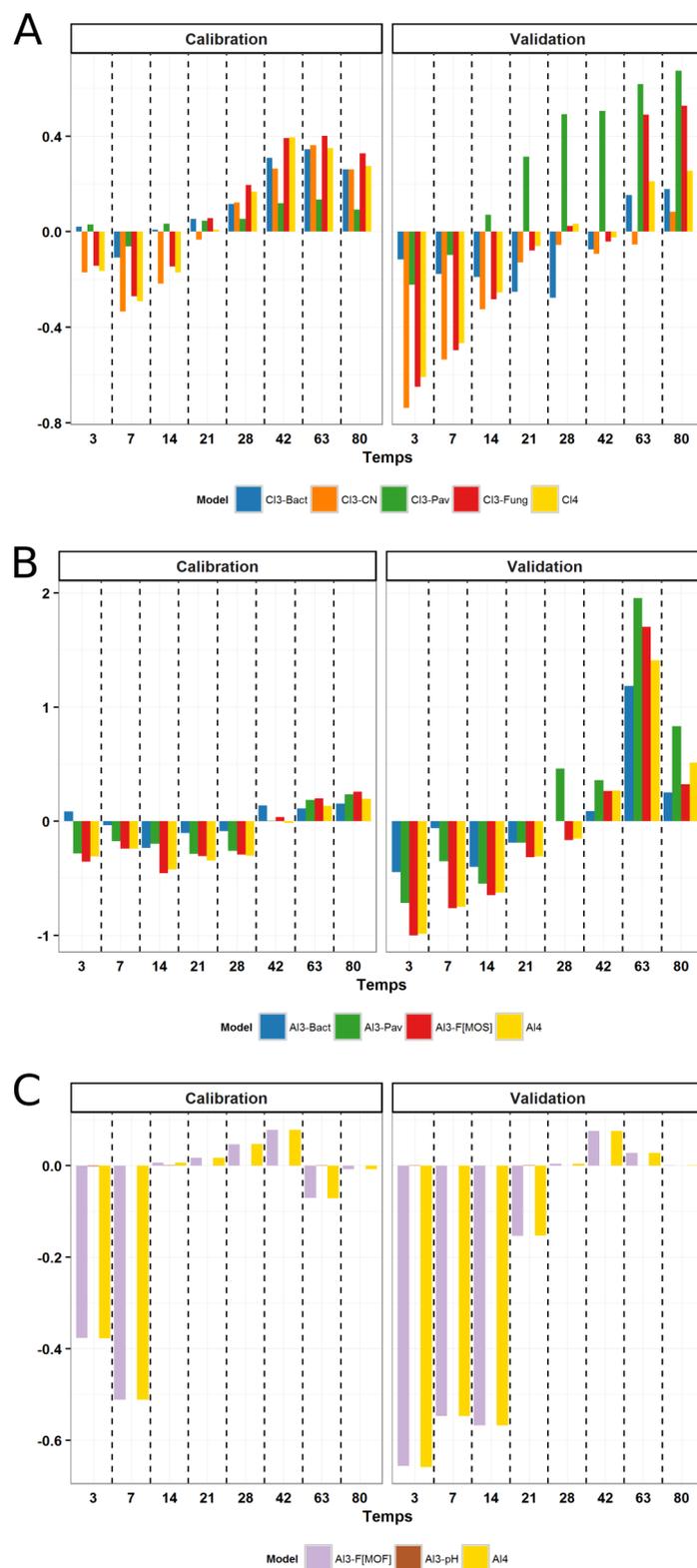


FIGURE 4.11 – Évolution relative par rapport au modèle de base de l'aire sous la courbe des courbes de probabilités cumulées du coefficient E de Nash-Sutcliffe pour la qualité d'ajustement (calibration) et la qualité de prédiction (validation) des modèles avec les différentes combinaisons de prise en compte des fonctions de modulation pour les flux non marqués du modèle appliqué aux sols témoins (A) et aux sols amendés (B) et pour les flux marqués (C) pour chaque date de mesure.

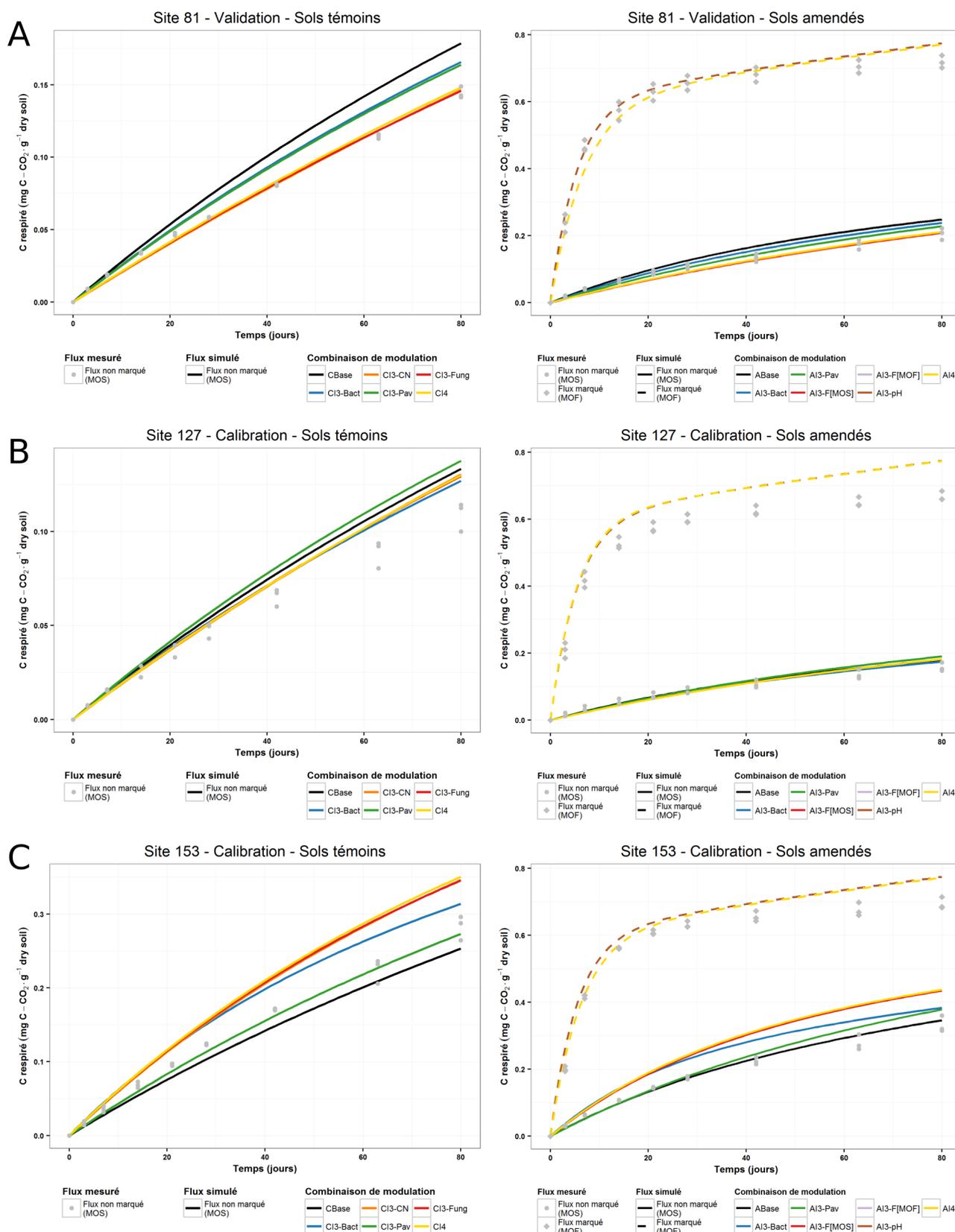


FIGURE 4.12 – Flux mesurés versus simulés pour le modèle appliqué aux sols témoins (gauche) ou aux sols amendés (droite) lorsque la prise en compte des fonctions de modulation améliore (A), n’a pas beaucoup d’impact (B) ou dégrade (C) la qualité de simulation par rapport au modèle de base (en noir).

4.4 Discussion

4.4.1 Comportement et qualité globale du modèle

Les qualités globales d'ajustement et de prédiction du modèle ont montré un défaut de simulation de la cinétique des flux non marqués (i.e. liés à la décomposition de la MOS) au début d'incubation et qui s'améliore avec le temps. Ce défaut de qualité se traduit par un biais négatif systématique. Le modèle choisi ne permet donc pas de simuler avec précision les dynamiques très court-terme de la décomposition de la MOS, que l'on soit en condition de sols témoins ou de sols amendés. Ce défaut de qualité étant systématique, il est attribué au modèle plutôt qu'aux différentes modulations.

En utilisant une méthode d'assimilation de données comme l'inférence bayésienne, Schädel et al. (2013) montrent que sur une période d'incubation comme celle utilisée ici (80 jours), la dynamique de décomposition de la MOS est mieux modélisée par la prise en compte de deux pools, un labile qui se décompose dans les 20 premiers jours et un autre plus récalcitrant avec une décomposition plus lente. La non prise en compte d'un pool labile peut expliquer la sous-estimation de notre modèle en début d'incubation. En effet, les valeurs estimées dans notre modèle de la constante de décomposition de la MOS (ks) sont du même ordre de grandeur que celles qu'ils estiment avec leur modèle ne prenant en compte qu'un seul pool ($1-3 \cdot 10^{-2}$). Ces valeurs ont des ordres de grandeur intermédiaires de celles qu'ils estiment pour le pool labile et le pool plus récalcitrant dans leur modèle prenant en compte deux pools de MOS.

La MOS est constitué généralement d'un sous-pool labile qui aurait pu être décomposé pendant la phase de pré-incubation (3 semaines). Cependant, ce pool labile a pu être alimenté pendant cette phase, en conséquence de la ré-humectation des sols séchés qui a pu activer des micro-organismes opportunistes (stratégie r) suite à l'amélioration des conditions d'humidité. Ces micro-organismes qui ont une durée de vie assez courte et leur croissance puis leur mort pendant la période de pré-incubation a pu conduire à l'augmentation de produits microbiens issus de la lyse cellulaire.

Ce défaut de qualité est plus important et dure plus longtemps dans le cas des sols amendés. L'impact de l'ajout de MOF sur la décomposition de la MOS (i.e. le priming effect) n'est pas du tout pris en compte dans ce modèle et peut expliquer cette observation. En effet, le priming effect est plus élevé en début d'incubation (figure 3.3) et l'absence de modélisation de cette interaction entre décomposition de MOF et MOS peut être responsable de la qualité plus mauvaise du

modèle dans le cas des sols amendés pour les flux non marqués.

La seule manière de prendre en compte cette interaction a été de séparer la calibration des paramètres entre le modèle appliqué aux sols témoins et le modèle appliqué aux sols amendés, ce qui a conduit à un doublement des valeurs du paramètres ks dans le cas des sols amendés. Cependant, une limitation de la décomposition de la MOS par la biomasse microbienne ou une co-limitation par la biomasse microbienne et la teneur en C, plutôt qu'une cinétique du premier ordre, semblerait plus appropriée (Wutzler and Reichstein, 2008). Cette manière de modéliser la décomposition de la MOS est moins courante dans la littérature et il serait plus difficile de renseigner des distributions *a priori*. Ce choix du modèle, participant aux valeurs élevées du paramètre ks , en comparaison de la distribution *a priori*, s'accompagne également des valeurs élevées du paramètre Ys , ces paramètres étant très corrélés (corrélation $>0,7$).

Pour les flux marqués (i.e. la décomposition du C de la MOF), le défaut de qualité générale n'est observable qu'à 3 jours d'incubation. La prise en compte de plusieurs pools a surement permis de mieux capturer la dynamique de cette décomposition, même court-terme. Le défaut de qualité à 3 jours provient surement du fait d'avoir fixé les paramètres de décomposition du pool le plus labile (MOF₁, paramètres kr_1 , Yr_1 et p_1) à des valeurs qui ne sont pas complètement adaptées à la dynamique observée de minéralisation du C de la MOF. Ces paramètres ont été fixés car leur impact sur les flux étaient faibles selon l'analyse de sensibilité globale. Cette analyse donne un indice qui concerne toute la période d'incubation ce qui explique la faible importance de ces paramètres qui concernent un pool qui se décompose très rapidement, et dont l'influence est plus grande seulement en début d'incubation, expliquant probablement le défaut de qualité à 3 jours.

Concernant la taille des sous-pools de la MOF, le paramètre p_2 , dont les valeurs sont les mêmes quelles que soient les modulations prises en compte, a une distribution *a posteriori* peu variable avec une valeur médiane de 0,58. Cette valeur est bien inférieure (-20%) à la valeur médiane de la quantité de C de cellulose et héli-cellulose dans un résidu de blé à laquelle se rapportait le pool MOF₂. Il n'est pas vraiment possible ici de savoir si cette valeur est liée au choix du modèle ou si le fractionnement Van Soest brut n'est pas le plus pertinent pour déterminer la taille des différents pools. Il existe dans la littérature des formules plus complexes, utilisant les fractions Van Soest, pour déterminer la taille de ces pools, qu'il serait intéressant de mettre en regard de ces valeurs (e.g. Whitmore *et al.*, 1997; Thuriès *et al.*, 2002). Il est à noter que les valeurs des distributions *a posteriori* du paramètres kr_2 sont similaires à celle fixée pour kr_1

(0,25), en particulier lorsque l'on prend en compte la modulation liée au pH. Cela suggère que la séparation des pools, en termes de cinétique, n'est pas aussi tranchée que celle utilisée ici car leur constante de décomposition ont des valeurs comparables.

Les distributions *a posteriori*, ayant des valeurs plus élevées des paramètres représentant les écarts-types des flux marqués et non marqués (*sd13* et *sd12*) que leur distribution *a priori*, s'explique par le fait que les distributions *a priori* ont été construites à partir des écarts-types des flux entre les répétitions d'incubation d'un même sol des données DIMIMOS. Autrement dit, ces écarts-types sont les écarts-types minimum que l'on peut observer compte tenu de la répétabilité des mesures de flux. Le modèle n'ajustant pas aussi bien les flux, l'estimation des écarts-types, qui inclut l'erreur liée au modèle, est forcément plus importante.

4.4.2 Impact global de la modulation de la valeur des paramètres

Globalement, la prise en compte de fonctions de modulation des paramètres de cinétique du C a dégradé les simulations des flux liés à la minéralisation de la MOS sur les 2 ou 3 premières semaine et a permis une nette amélioration de ces simulations pour le reste de l'incubation. La dégradation en début d'incubation est sûrement imputable pour partie à la mauvaise qualité du modèle pour simuler la dynamique sur cette période. Il est possible que la prise en compte d'une information supplémentaire, qui a sa propre variabilité, a amplifié cette mauvaise qualité.

Pour la période d'incubation pour laquelle le modèle de base a une meilleure qualité de simulation, la prise en compte d'une information supplémentaire via des fonctions de modulation a clairement permis d'augmenter la qualité par rapport au modèle de base, et ce pour le modèle appliqué aux sols témoins et aux sols amendés. Cela souligne la pertinence de cette approche de prise en compte des propriétés du sol et notamment de la diversité microbienne pour avoir des modèles relativement simples permettant de capturer au mieux la dynamique du C dans le sol. En effet, les fonctions de modulation utilisées, une fois validées, permettent de prendre en compte l'effet d'une propriété du sol en modulant les paramètres *ad hoc* d'un modèle sans avoir à affiner les processus pris en compte dans le modèle.

Les conséquences d'un modèle dont les processus pris en compte sont plus détaillés sont souvent une augmentation du nombre de paramètres dans le modèle entraînant une grande difficulté de calibration et conduisant à la nécessité d'émettre des hypothèses pour faire tourner le modèle qui ne sont pas toujours transférables d'une situation à une autre. C'est la limite des modèles prenant en compte la diversité microbienne par une multiplication de pools microbiens repré-

sentant différents groupes taxonomiques et/ou fonctionnels présentés au chapitre 1. À l'inverse, dans notre cas, le nombre de paramètres est assez réduit (6 au maximum) et leur modulation par les propriétés du sol permet de mieux simuler les dynamiques spécifiques à chaque site.

Cela pose tout de même deux questions. La première est la nécessité de bien identifier les paramètres qu'il faut moduler et de construire les fonctions de modulation adéquates. Historiquement, la constante de décomposition est le paramètre modulé principalement par la température, l'humidité et la texture (e.g. Hansen *et al.*, 1991). Nos résultats suggèrent que l'approche de modulation des valeurs de paramètre ne doit pas se réduire à ce seul paramètre. En particulier, la modulation du rendement d'utilisation du C (Y_s) par la diversité bactérienne a montré son intérêt. Cet intérêt est d'autant plus grand que ce paramètre est souvent un paramètre fixé, malgré la sensibilité des modèles à sa valeur, car il n'existe aucun consensus sur sa valeur et qu'il est parfois difficile à calibrer (car corrélé avec la constante de décomposition) (Manzoni and Porporato, 2009; Manzoni *et al.*, 2012).

Construire la bonne fonction a également son importance car une mauvaise fonction associée au mauvais paramètre peut dégrader la qualité du modèle. C'est le cas par exemple de la fonction liée au pH qui a globalement dégradé la qualité de simulation du flux lié à la minéralisation du résidu en modulant la constante de décomposition du sous-pool MOF_2 . De plus, aucune modulation du rendement d'utilisation des pools de MOF (e.g. Yr_2) n'a été appliquée dans le modèle ; ceci a conduit à des valeurs cumulées de flux marqués à 80 jours égales quel que soit le sol étudié et malgré une variabilité apparente (Figure 4.12). Identifier les propriétés permettant une modulation de ce paramètre permettrait d'améliorer les simulations.

La deuxième question concerne le choix du modèle. Il est important de bien identifier les processus à prendre en compte dans le modèle pour choisir le modèle le plus simple (moins de paramètres possibles) ainsi que les modulations à appliquer. Dans notre cas, le modèle est suffisant en termes de processus pour simuler la dynamique intra-annuelle du C suite à l'ajout d'un résidu de culture, alors qu'il est probable qu'une amélioration soit possible par la prise en compte d'autres modulations (e.g. pour Yr_2 , non identifié dans ce travail).

Cependant, pour une dynamique plus longue (inter-annuelle), d'autres processus sont à prendre en compte comme la stabilisation de la MOS ou le priming effect car il y aurait un apport régulier de résidu de culture. Il faudrait alors identifier les modulations de paramètres nécessaires dans ces situations. Cela renvoie à un avantage considérable de cette approche : les fonctions de modulation construites sont facilement transférables d'un modèle à un autre et leur

impact sur les simulations peut donc être testé facilement.

4.4.3 Prise en compte spécifique de la diversité microbienne dans le modèle au regard des autres propriétés du sol

La non prise en compte de la diversité bactérienne a tendance à dégrader les qualités d’ajustement et de prédiction du modèle. Cette observation est surtout vraie pour la qualité prédictive et à partir de 21 jours d’incubation. La fonction de modulation liée à la diversité bactérienne concerne le paramètre Y_s qui contrôle la quantité respirée du C décomposé, et dont l’influence est de plus en plus visible au cours du temps sur un flux cumulé. Cela peut aussi expliquer que la diversité bactérienne améliore le modèle surtout en milieu et fin d’incubation. L’amélioration, en particulier de la qualité prédictive, est en accord avec notre hypothèse que la prise en compte d’indice de diversité bactérienne peut aider à la prédiction de la dynamique du C et conforte notre hypothèse d’un meilleur rendement d’utilisation du C avec l’augmentation de la diversité bactérienne. Cette observation est cependant à confirmer dans d’autres études.

À l’inverse, l’hypothèse d’une modulation des valeurs de la constante de décomposition du C de la MOS par la diversité fongique semble ne pas être confirmée ici. Sa non prise en compte n’a pas eu d’effet sur les qualités du modèle, voire a permis une amélioration ponctuelle. La prise en compte de la diversité fongique pour moduler les valeurs le paramètre kr_2 n’a eu aucun impact comme il était attendu du fait de l’estimation de la fonction $f_{fungMOS}$ qui est toujours proche de 1 pour la gamme des valeurs de diversité fongique ($1/D$ fongique) du jeu de données MOSAIC.

Globalement, les hypothèses concernant la modulation par la diversité en champignons sont donc mises en défaut ici. Il n’est pas possible de savoir si la diversité en champignons n’a en fait pas d’effet significatif, si ce résultat est dû au choix de modélisation ou si l’hypothèse faite n’est pas la bonne et qu’il faudrait en tester d’autres. Par exemple, la physiologie des champignons en hyphes leur permet de s’infiltrer dans les agrégats ou dans le réseau matriciel cellulosique des plantes leur permettant une accessibilité accrue aux substrats par rapport aux bactéries (de Boer *et al.*, 2005). De plus, il est souvent avancé que les champignons produisent un panel d’enzymes permettant la dégradation de composés plus variés que les bactéries (de Boer *et al.*, 2005). Ces avantages pourraient être accrus par une augmentation de la diversité, ainsi la diversité en champignons serait responsable d’une accessibilité accrue au substrat à la fois pour eux mais aussi pour les bactéries par interactions positives (de Boer *et al.*, 2005; Hooper *et al.*, 2005). Sous cette hypothèse, la diversité fongique aurait plutôt un impact sur les paramètres p_2 et sur la

fraction (f) de MOS réellement accessible (fixée à 2% dans notre étude). Il serait intéressant de tester cette hypothèse.

Les modulations liées à la diversité microbienne ont été prises en compte avec des modulations liées aux propriétés physiques et chimiques du sol. Malgré une prise en compte plus simpliste qu'un couplage entre C et N, la non prise en compte du ratio C/N entraîne une diminution de la qualité du modèle confirmant l'importance de cette variable dans la dynamique du C. Au contraire, le P assimilable, bien que sa non prise en compte entraîne une dégradation de la qualité d'ajustement, elle entraîne aussi une diminution de la qualité prédictive. Cela suggère un sur-ajustement par la prise en compte de cette variable. Peu d'information a été trouvée dans la littérature concernant le lien entre P assimilable et dynamique du C dans les modèles et nos résultats suggèrent un plus grand examen de ce lien.

Enfin, pour le pH, sa prise en compte conduit à une diminution des qualités du modèle dans la première moitié d'incubation et semble ne pas avoir trop d'importance après. La dégradation en début d'incubation est en accord avec l'observation d'un lien négatif sur la période 0-7 jours dans la sélection GAM (section 3.2) alors que l'on prend en compte en lien positif (visible dans les 2 périodes suivantes). Comme déjà annoncé, le pH est lié à beaucoup d'autres variables et son effet n'est pas évident à expliquer. De plus, la fonction de modulation utilisée a été prise dans la littérature (Leifeld *et al.*, 2008) et n'est pas forcément la plus adaptée à notre étude.

4.4.4 Retour sur le choix de la méthode pour répondre à nos objectifs

Le choix d'utiliser la calibration bayésienne et de fixer certains paramètres nous a permis de calibrer des modèles de dynamique du C malgré les problèmes de sur-paramétrage au regard des données disponibles. Les distributions *a posteriori* estimées des paramètres ont des valeurs qui sont comparables à l'information proposée *a priori* et aux valeurs de la littérature vis à vis de la structure du modèle choisie.

Malgré de fortes corrélations entre certains paramètres comme ks et Ys , les distributions *a posteriori* sont globalement peu dispersées et l'information que l'on en retire permet de mieux comprendre le comportement du modèle. Cela nous a permis également de tester différentes hypothèses de modulation des paramètres. L'inférence bayésienne est encore peu utilisée dans le domaine de la dynamique du C mais a déjà prouvé son utilité (Schädel *et al.*, 2013; Ahrens *et al.*, 2014; Hararuk and Luo, 2014; Hararuk *et al.*, 2015), en particulier dans un domaine où les modèles sont souvent sur-paramétrés (Sierra *et al.*, 2015) mais où de l'information sur la valeur

possible des paramètres est largement disponible au vu du grand nombre de modèles existants (Manzoni and Porporato, 2009).

On peut s'interroger également sur l'impact de la calibration une à une des fonctions de modulation. En faisant cela, la variabilité liée à la non prise en compte des autres propriétés du sol peut rendre difficile la calibration de la fonction. Par exemple, la variabilité causée par une autre propriété peut être incorporée dans la fonction calibrée. Il aurait été intéressant de comparer la calibration de ces fonctions selon différentes combinaisons de prises en compte des modulations mais il existait une contrainte de temps des simulations numériques. Il serait sûrement plus judicieux de construire les fonctions de modulation en dehors de la calibration des autres paramètres du modèle. Cela nécessite i) soit d'avoir une expérimentation où seule la propriété varie, ii) soit de pouvoir calibrer les paramètres de cinétiques du C pour chaque sol avec suffisamment de confiance (i.e. peu de problème d'identifiabilité) pour pouvoir les mettre en regard des propriétés utilisées pour leur modulation. Ce point est détaillé dans la partie discussion générale.

4.5 Conclusion

En utilisant une stratégie de calibration pertinente, nous avons pu tester les hypothèses de prise en compte de fonctions de modulation de paramètres d'un modèle de dynamique du C, dépendantes de la diversité microbienne. Si les hypothèses concernant la diversité en champignons n'ont pas pu être confirmées, celle concernant la diversité bactérienne n'a pas été mise en défaut ce qui est encourageant quant à la prise en compte de la diversité dans les modèles de dynamiques du C dans le sol. Plus encore, l'amélioration du modèle à l'aide de fonctions de modulation, entre autres liées à la diversité microbienne, montre la pertinence d'une telle approche pour construire des modèles génériques simples.

Toutefois, des questions se posent quant au choix du modèle et ces résultats demandent à être confirmés dans d'autres études. L'encouragement qu'apporte cette étude doit nous inciter à continuer dans ce sens, en particulier dans le contexte actuel d'une meilleure prise en compte des processus biologiques auxquels la biodiversité microbienne participe.

Chapitre 5

Discussion générale

5.1 Apports et limites du cadre de modélisation de la dynamique du C proposé

5.1.1 Cadre de modélisation proposé

L'objectif général de ce travail avait pour ambition de mettre en œuvre une méthode pour prendre en compte la diversité microbienne dans un modèle de dynamique du C. L'objectif de modélisation se concentrait sur l'opérationnalité du modèle, i.e. i) un modèle dont le nombre de paramètres est limité pour faciliter la calibration, ii) dont la précision est satisfaisante et iii) dont les variables à renseigner dans le modèle sont facilement accessibles.

Une revue de la littérature (Figure 5.1, chapitre 1) nous a permis de constater que la diversité microbienne était relativement absente des modèles de dynamique du C malgré une démonstration de plus en plus forte de son rôle. De plus, les premiers modèles qui intègrent une notion de diversité microbienne le font en distinguant différents compartiments microbiens correspondant à différents groupes fonctionnels. L'intérêt de ces modèles est fort concernant la formalisation conceptuelle et le test d'hypothèses du lien entre diversité microbienne et dynamique du C. Cependant, l'ajout de groupes conceptuels de micro-organismes conduit à des modèles difficiles à renseigner dont l'augmentation du nombre de paramètres les rend peu opérationnels, s'éloignant de notre objectif.

Notre approche est basée sur la mise en œuvre d'un modèle mécaniste classique dont les flux de décomposition/minéralisation sont décrits par des cinétiques d'ordre 1 et dont les paramètres de ces équations sont modulés par des fonctions dépendantes de propriétés des sols, dont des indices

de diversité microbienne classiques en écologie (i.e. richesse, équitabilité, indice de Shannon et inverse de l'indice de Simpson).

La prise en compte de propriétés du sol dans des fonctions de modulation de paramètres de décomposition est classique en modélisation de dynamique du C (e.g. Hansen *et al.*, 1991; Garnier *et al.*, 2001). Cependant, cette modulation se fait systématiquement sur la constante de décomposition (paramètre k). Nous avons élargi ce concept en considérant que chaque paramètre du modèle pouvait être modulé, tel que le paramètre de rendement d'utilisation du C pour la croissance microbienne (paramètre Y).

La méthodologie générale mise en place suit les grandes étapes introduites dans le chapitre 1 (figure 1.6, flèches A, B, C) suite à la synthèse bibliographique et qui sont détaillées dans la figure 5.1.

La première grande étape concerne l'utilisation d'expérimentation pour trouver les descripteurs de la diversité microbienne pertinents pour expliquer les variabilités de dynamique du C à l'aide de méthodes statistiques. Deux jeux de données, provenant de programme ANR, ont été utilisés (Figure 5.1, chapitre 2). Dans ces jeux de données, les flux de CO₂ émis par différents sols incubés, avec ou sans paille de blé marqué au ¹³C (traitements témoin vs amendé), ont été mesurés pendant 80 jours pour caractériser la dynamique de C dans le sol. Les sols échantillonnés sont des sols agricoles et les deux bases de données diffèrent par l'échelle spatiale d'échantillonnage : nationale (DIMIMOS) vs paysage (MOSAIC).

Les conclusions de nos résultats concernent la dynamique du C à l'échelle intra-annuelle durant laquelle une grande partie du résidu (MOF) peut être décomposée alors que les variations de MO du sol (MOS) sont très faibles. Ainsi, la question du recyclage des différents pools de la MO dans le sol (MOF, MOS, biomasse microbienne) est évaluable à travers ces jeux de données mais la question du stockage/déstockage de MOS, intervenant sur des périodes plus longues, ne l'est pas. Néanmoins, cette notion pourra être discutée.

La pertinence des indices de diversité a été évaluée à l'aide de modèles statistiques additifs généralisés (GAM), avec lesquels une méthode de sélection des meilleurs prédicteurs des flux de CO₂ mesurés a été réalisée (Figure 5.1, chapitre 3). Les prédicteurs potentiels sont les différentes propriétés des sols et les indices de diversité pour les bactéries et les champignons indépendamment.

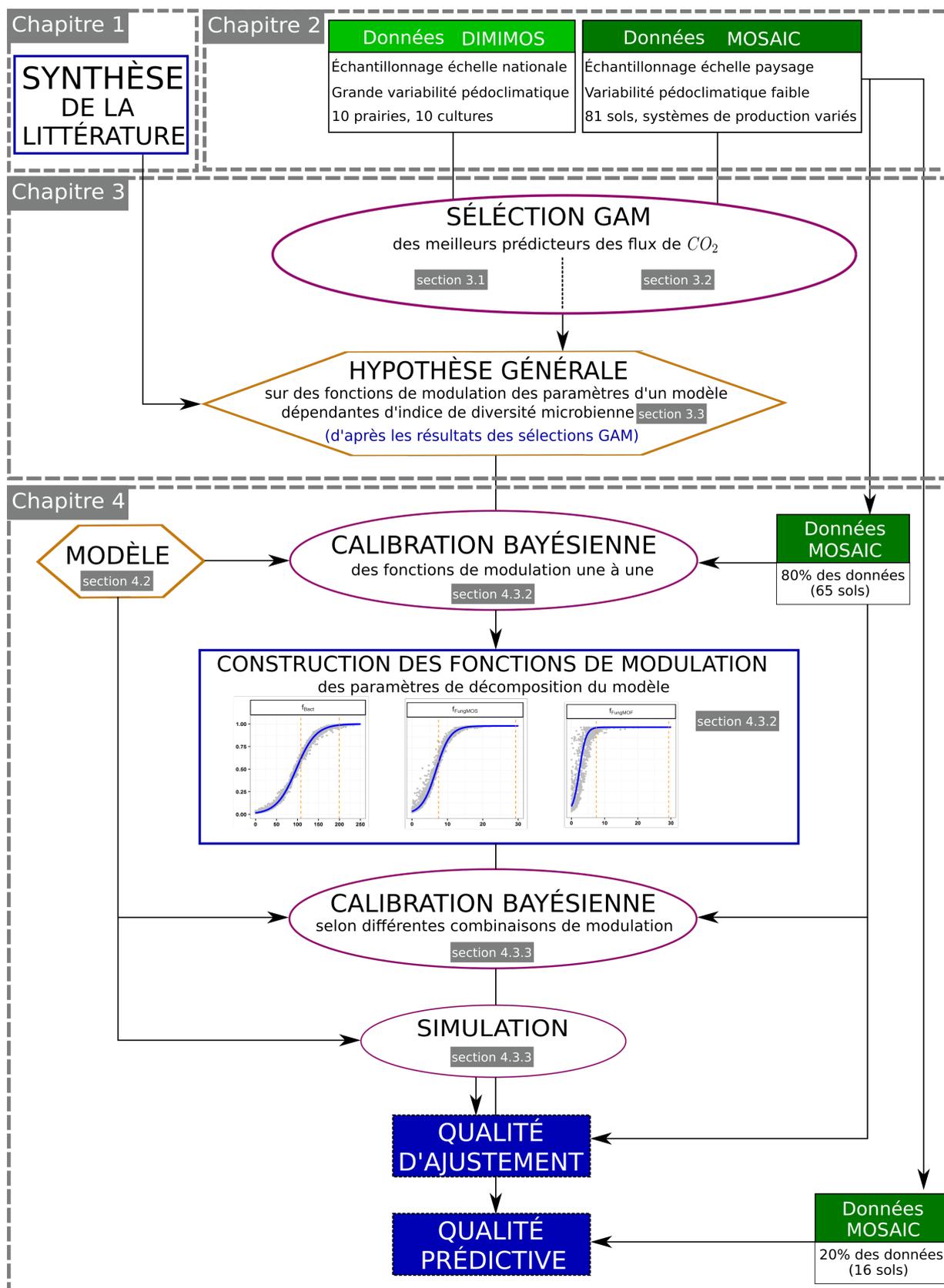


FIGURE 5.1 – Diagramme de la méthode générale appliquée.

Les données (en vert) ou hypothèse (en orange) sont analysées ou testées à l'aide de méthodes spécifiques (en violet) pour conduire à des résultats (en bleu) ou d'autres hypothèses.

La deuxième grande étape concerne la construction de fonctions de modulation des paramètres des cinétiques des modèles de dynamique du C, et dépendantes des indices de diversité identifiés à l'étape précédente. La troisième grande étape est l'intégration de ces fonctions et l'évaluation de leur impact sur la qualité prédictive des modèles. Ces deux étapes n'ont pas été séparées dans ce travail. Nous avons utilisé les résultats de la première étape, associée à la synthèse bibliographique, pour proposer des hypothèses sur la construction des fonctions de modulation. Ces fonctions ont été directement intégrées dans le modèle pour être calibrées.

Dans un premier temps, les paramètres des fonctions de modulation ont été calibrés dans un modèle mécaniste classique, à l'aide d'une calibration bayésienne sur 80% des données MOSAIC en intégrant ces fonctions une à une (Figure 5.1, chapitre 4). Ensuite, les paramètres intervenant dans les équations des cinétiques de dynamique du C du même modèle ont été calibrés selon la même procédure, en intégrant les différentes fonctions de modulation calibrées selon plusieurs combinaisons : aucune fonction de modulation ajoutée, toutes les fonctions de modulation ajoutées et toutes les fonctions de modulation ajoutées sauf une à chaque fois.

Les qualités d'ajustement et prédictive des différentes calibrations sont mesurées et comparées à l'aide des 80% de données MOSAIC utilisés pour la calibration et des 20% restants, respectivement. Cela nous permet d'évaluer le gain des qualités du modèle suite à la prise en compte ou non des fonctions de modulation et donc de la diversité microbienne dans la simulation de la dynamique du C.

5.1.2 Discussion sur les approches choisies pour la modélisation

Un modèle générique avec peu de paramètres

Le **modèle mécaniste** choisi avait pour ambition de pouvoir simuler les cinétiques de flux de CO₂ sur la période d'incubation tout en ayant un nombre réduit de paramètres et faisant intervenir des variables d'entrées facilement mesurables. Ce modèle est basé sur deux pools de MO, pour le résidu (MOF) et pour la MO du sol (MOS), qui sont décomposés suivant une cinétique du premier ordre contrôlé par deux paramètres : une constante de décomposition (k) et un rendement d'utilisation du C (Y). Cette représentation est la base de tout modèle mécaniste de décomposition et représente le minimum pour simuler les processus de décomposition et minéralisation du C (Manzoni and Porporato, 2009; Sierra *et al.*, 2015). Dans notre cas, c'est aussi une représentation obligatoire car des hypothèses de modulation sur les deux types de paramètres étaient à tester.

Une partie du C décomposé alimente un troisième pool, la biomasse microbienne. Un autre flux de C, contenant un paramètre (R_T), représente la respiration de la biomasse microbienne autre que pour la croissance. Ce flux permet de prendre en compte l'effet de la biomasse microbienne sur les flux de CO_2 , cette variable ayant montré son influence dans l'étape de sélection GAM.

Du fait d'une hypothèse sur la modulation de la constante de décomposition d'un sous-pool du résidu, ce dernier a été subdivisé en trois sous-pools, dont 2 dynamiques, avec chacun un lot de paramètres k et Y différents. Un pool est considéré labile, l'autre plus récalcitrant. Chacun des sous-pools se voit attribuer un paramètre contrôlant leur taille (proportion du carbone total du résidu).

Au total, dans le modèle de base, et dans les sols amendés, on dénombre 2 proportions + 3 x (2 paramètres de décomposition/minéralisation) + 1 paramètre de respiration = 9 paramètres. Ce nombre se situe dans la gamme basse des valeurs de nombre de paramètres dans les modèles de dynamique du C (Smith *et al.*, 1997). De plus, à l'aide de l'analyse de sensibilité pour fixer certains paramètres, jamais plus de 4 de ces paramètres n'ont été calibrés en même temps.

Le fait que le flux de CO_2 marqué (**minéralisation de la MOF**) atteigne une asymptote suggère que le résidu a fini de se décomposer ou que la partie restante a une dynamique de décomposition très lente et donc négligeable sur la période d'incubation. Le modèle, hormis un léger défaut de qualité jusqu'à 3 jours, capture très bien la dynamique du CO_2 marqué. Les valeurs des constantes de décomposition entre le pool le plus labile (valeur fixée à 0,25) et le pool moins labile (valeur estimée à 0,17 dans le modèle de base) montrent que la différence de vitesse de décomposition de ces deux pools n'est pas forte et qu'il serait possible de les réunir, simplifiant ainsi le modèle. Dans ce travail, cette réunification est d'autant plus justifiée que l'hypothèse qui a motivé cette division, i.e. la modulation de la valeur de la constante de décomposition du pool moins labile par une fonction dépendante de la diversité fongique, a été mise en défaut.

La simulation des flux de CO_2 non marqué (**minéralisation de la MOS**) par le modèle montre un défaut de simulation de la minéralisation de la MOS au début des incubations. Ce défaut de simulation peut être expliqué par la présence d'un sous-pool de MOS plus labile (Schädel *et al.*, 2013) qui n'est pas représenté dans le modèle. Dans les sols amendés, l'absence de modélisation du priming effect s'ajoute probablement pour amplifier ce défaut de simulation en début d'incubation.

La décomposition de la MOS étant un processus très lent, cette dynamique de début n'a que

peu d'influence et il est plus intéressant d'avoir une idée des sorties de C sur une durée plus importante. Les simulations du modèle étant non biaisées en moyenne sur la fin d'incubation, il permet de mieux simuler les flux à plus long terme (21-80 jours). Contrairement à la dynamique du C de la MOF, il existe déjà une différenciation des simulations par le modèle de base, entre les différents sols, car le C organique total et la biomasse microbienne de départ interviennent dans les équations du modèle.

Tout en ayant un nombre de paramètres limité, des variables à renseigner facilement mesurables (C de la MOS, C de la MOF et biomasse microbienne), le modèle choisi a la capacité de simuler les dynamiques de C suivant ces deux sources. Pour le C issu de la MOF, il s'agit d'une cinétique moyenne tandis que pour le C de la MOS, la cinétique est déjà différenciée entre les sols. Le cadre mécaniste choisi est très générique et peut s'appliquer pour différents sols et différentes situations (e.g. autres résidus, différents modes d'usage) (Tableau 5.1).

Une précision accrue possible par la modulation des paramètres

À partir du modèle de base, l'approche utilisée ici consiste à prendre en compte des facteurs influençant la dynamique du C, et en particulier la diversité, à l'aide de **fonctions de modulation de paramètres** de modèles mécanistes. Cette approche n'est pas nouvelle mais a été élargie ici à d'autres variables que celle classiquement prises en compte et surtout à d'autres paramètres que la constante de décomposition.

Bien que cette observation ne soit pas générale, la prise en compte des fonctions de modulation a permis d'**augmenter la précision des simulations** (Tableau 5.1). C'est le cas notamment pour les flux issus de la minéralisation de la MOS. Cela suggère donc qu'il est possible d'améliorer les prédictions d'un modèle à l'aide de cette approche et donc de mieux hiérarchiser les sols selon leurs flux. À titre d'illustration, lorsque l'on divise les flux non marqués cumulés, observés dans les sols témoins à 80 jours, en 4 groupes définis par leurs quartiles, et que l'on regarde si les simulations par les modèles classent les sols dans les bons groupes (i.e. les bons quartiles), on obtient un taux de bonne classification de 30% (24 sols) pour le modèle de base et de 44% (36 sols) pour le modèle avec toutes les fonctions de modulation prises en compte.

Si l'on se réfère à des relations bien connues, il est intéressant de noter que la fonction de modulation dépendante du ratio C/N a permis une amélioration importante des simulations. En général, la prise en compte de ce ratio se fait de manière plus complexe, i.e. en interactions stœchiométriques avec i) l'azote minérale, ii) le ratio C/N des micro-organismes et iii) les rendement

d'utilisation de C et N (Manzoni and Porporato, 2009). Dans notre cas, son effet n'est représenté que par une fonction de modulation générale. L'amélioration des simulations par sa prise en compte souligne l'intérêt de cette approche qui permet de représenter l'influence globale d'un processus plus ou moins complexe dans la dynamique du C sans avoir à détailler ce processus.

TABLE 5.1 – Avantages et limites ou contraintes du matériel et des méthodes utilisés.

| | AVANTAGES | LIMITES/CONTRAINTES |
|--|--|---|
| MATÉRIEL | | |
| Données disponibles | <ul style="list-style-type: none"> - Différences d'échelles d'échantillonnage - Cultures/Prairies pour DIMIMOS - Beaucoup de sols pour MOSAIC - Richesse d'information | <ul style="list-style-type: none"> - Comparaison difficile - Dynamique intra-annuelle - Suivi du CO₂ uniquement - Variabilité et confusion |
| APPROCHES CHOISIES | | |
| Indices de diversité microbienne | <ul style="list-style-type: none"> - Largement utilisés/connus - Simplicité de l'information - Facilement mesurables | <ul style="list-style-type: none"> - D'autres indices ? - Manque de référence |
| Intégration via des fonctions de modulation de paramètres | <ul style="list-style-type: none"> - Facilement transférables - Augmentation de la précision - Pattern général | <ul style="list-style-type: none"> - Représentation non explicite |
| Choix du modèle mécaniste | <ul style="list-style-type: none"> - Nombre de paramètres - Bonne qualité globale - Généricité | <ul style="list-style-type: none"> - Mauvaise qualité ponctuelle - Absence de mécanismes - Cinétique du 1^{er} ordre |
| Calibration des fonctions de modulation | | <ul style="list-style-type: none"> - Calibration une à une |
| Évaluation de la prise en compte des fonctions de modulation | | <ul style="list-style-type: none"> - Absence de certaines combinaisons de modulation |
| MÉTHODES UTILISÉES | | |
| Modèles additifs généralisés (GAM) | <ul style="list-style-type: none"> - Prise en compte de relations non linéaires | <ul style="list-style-type: none"> - Sur-ajustement, difficultés d'interprétation - Besoin d'un grand nombre de données |
| Inférence bayésienne | <ul style="list-style-type: none"> - Problèmes d'identifiabilité - Ajout de connaissances <i>a priori</i> - Diagnostic de calibration | <ul style="list-style-type: none"> - Temps de calibration - Impact du choix des distributions <i>a priori</i> non testé ici |

En termes de **diversité microbienne**, il est intéressant de noter que la mesure de l'importance relative de ces indices, lors de la sélection GAM, est comparable à celle d'autres propriétés des sols connues pour influencer la dynamique du C. Les résultats de l'étape de modélisation montrent également qu'il est possible d'améliorer la qualité d'un modèle en prenant en compte la diversité à l'aide de fonctions de modulation. Ce résultat est d'autant plus intéressant que le paramètre de rendement d'utilisation du C pour lequel la modulation a permis une amélioration de qualité est un paramètre présent dans la grande majorité des modèles de dynamique du C

(Manzoni and Porporato, 2009) et dont la valeur est souvent fixée sans qu'un consensus existe (Manzoni *et al.*, 2012).

De plus, les relations entre la diversité microbienne et la dynamique du C semblent être assez générales (e.g. pas de différence entre mode d'usage), ce qui n'altère pas la généralité du modèle de base. Cette observation est très intéressante car de telles fonctions sont facilement transférables d'un modèle à un autre et permettent d'affiner la précision des simulations sans complexifier le modèle.

Choix des indices de diversité : données d'entrées facilement accessibles

Le choix d'utiliser les **indices de diversité microbienne** que sont la richesse, l'équitabilité, l'indice de Shannon et l'inverse de l'indice de Simpson a été motivé par plusieurs facteurs. D'une part, ces indices constituent une information i) quantitative simple, ii) accessible à partir de méthodes de biologie moléculaire, iii) qui renseignent sur la diversité d'une communauté, iv) sensible aux conditions environnementales (e.g. Maron *et al.*, 2011) et v) concernant une propriété impliquée dans le fonctionnement des sols (e.g. Tardy *et al.*, 2015). Leur interprétation est aisée et leur utilisation très répandue en écologie.

Nos résultats sont encourageants quant à la pertinence de ces indices pour expliquer une partie de la dynamique du C. Cependant, ces indices ont été développés pour mesurer la diversité de macro-organismes dont la notion d'espèce est bien définie et où la rareté (en termes d'abondance) est moins marquée (Hughes and Bohannan, 2004). Bien qu'ils aient montré leur pertinence et robustesse pour décrire la diversité microbienne (Hughes *et al.*, 2002), en particulier pour les indices de Shannon et Simpson (Haegeman *et al.*, 2013), d'autres moins utilisés existent et pourraient s'avérer tout aussi pertinents (Hill *et al.*, 2003).

Il est vraisemblable que des indices spécifiques à la diversité microbienne du sol soient développés en même temps que la compréhension du lien entre diversité microbienne et fonctionnement du sol progresse. Par exemple, des classifications telles que la dichotomie copiotrophes-oligotrophes (Fierer *et al.*, 2007) pour les phyla bactériens permettent de faire le lien entre diversité taxonomique et traits fonctionnels et pourraient intervenir dans des indices de diversité. Il est donc prometteur d'utiliser l'approche proposée en relation avec des spécialistes de l'écologie microbienne afin d'en intégrer les travaux développés en termes d'indices descripteurs de la diversité microbienne.

Un réalisme modéré de la représentation de la diversité

La prise en compte de la diversité microbienne par des fonctions de modulation reste peu explicite. Hormis ce qui est représenté explicitement dans le modèle de base, les processus sous-jacents des fonctions de modulation ne sont pas représentés de manière réaliste. Il est alors légitime de questionner la limite de cette approche pour modéliser certains processus.

Par exemple, nos résultats montrent que ces indices de diversité ne sont pas suffisants pour expliquer le priming effect, qui serait plus influencé par la composition de la communauté (Fontaine *et al.*, 2011; Pascault *et al.*, 2013; Tardy *et al.*, 2015). En ce sens, si l'objectif du modèle est de représenter le priming effect, par exemple dans le but d'étudier l'effet de la gestion des résidus de culture sur ce mécanisme, le modèle de Perveen *et al.* (2014), distinguant un groupe microbien plutôt décomposeur de la MOF et un autre plutôt décomposeur de la MOS, paraît prometteur.

Toutefois, si notre étude n'a pas permis d'identifier de propriétés du sol intéressantes pour le priming effect, cela ne signifie pas qu'il n'est pas possible d'appliquer l'approche de modulation d'un paramètre dans la modélisation de ce processus. Cela signifie cependant qu'un effort supplémentaire doit être fait pour i) développer un modèle le plus simple possible qui prenne en compte ce processus et ii) identifier les modulations de paramètres nécessaires.

5.1.3 Avantages et limites des méthodes utilisées

Parmi les méthodes utilisées, les **modèles additifs généralisés** (GAM, Hastie and Tibshirani, 1990) ont été utilisés pour étudier le lien entre les indices de diversité et les dynamiques du C. La grande force de ces modèles réside dans leur capacité à prendre en compte des relations non linéaires lorsque la forme de ces relations n'est pas connue. Cela nous a permis par exemple d'obtenir la relation entre l'inverse de l'indice de Simpson fongique et la minéralisation du résidu à la période 7-28 jours (annexe D.3).

Toutefois, l'une des limites majeures, et déjà évoquée, est la tendance au sur-ajustement d'une telle approche. Bien que des techniques ont été mises en place pour éviter ce problème (cf. section 3.1.2.2), ce dernier a été rencontré avec les données DIMIMOS qui contenaient peu de sols par modalités de mode d'usage (10). De plus, certaines relations ont pu être sélectionnées sans qu'une raison biologique existe. Ceci constitue une limite supplémentaire au fait que les données proviennent d'échantillonnage *in situ* et n'ont pas été obtenues expérimentalement. Les relations obtenues ne peuvent être expliquées par un lien de causalité direct et seules des hypothèses

peuvent être soulevées. Cette notion est développée dans la section présentant les perspectives possibles.

En conclusion, nous recommandons vivement l'utilisation de cette approche pour déceler des relations non linéaires, mais nous attirons l'attention sur le besoin d'un nombre de données suffisant pour capturer ces relations et de garder un esprit critique vis à vis des relations obtenues.

Malgré le nombre limité de paramètres des cinétiques dans le modèle choisi, des problèmes d'identifiabilité des paramètres ont été soulevés dans ce modèle, principalement en raison d'un sur-paramétrage du modèle au regard des données disponibles, restreintes à la seule dynamique des flux de CO₂. Ce sur-paramétrage vient en particulier de la forte corrélation qui existe entre les constantes de décomposition et les rendements d'utilisation du C associés lors de la calibration. Le suivi d'autres variables au cours des incubations permettraient de contrer cette limite. Ce point est développé dans les perspectives. Pour pallier ce problème, le choix a été fait d'utiliser des méthodes d'**inférence bayésienne** permettant d'apporter une information *a priori*. Cela nous a permis d'obtenir des distributions *a posteriori* assez précises et de les comparer aux distributions *a priori*, issues des valeurs que l'on peut trouver dans la littérature.

Cette solution a un coût important en temps de calibration ce qui nous a restreint dans la **calibration des fonctions de modulation** qui ont du être effectuées une à une et n'ont pas été recalibrées à chaque combinaison de prise en compte de ces fonctions. De plus, certaines combinaisons de modulation de paramètres n'ont pas été testées. Il est probable que la calibration des fonctions de modulation toutes ensemble donnerait des résultats complémentaires sur la forme des fonctions ainsi que sur le gain prédictif de leur prise en compte. Elle constitue en cela une perspective de travail intéressante.

Il apparait cependant encore plus intéressant de pouvoir construire ces fonctions indépendamment de la calibration du modèle. En effet, l'ajout des fonctions de modulation à calibrer fait augmenter le nombre de paramètres, et donc le problème de sur-paramétrage du modèle, ce qui i) peut conduire à des problèmes de convergence comme cela a été le cas pour la fonction lié au ratio C/N, et qui ii) est contraire à l'objectif d'avoir un modèle limité en nombre de paramètres. Cela impliquerait d'avoir un jeu de données similaire mais où les confusions entre propriétés des sols échantillonnés sont moins prononcées. Cette approche serait un bon moyen de pouvoir faire à la fois un lien de causalité tout en construisant les fonctions de modulation indépendamment de l'étape de modélisation. Ce point est développé dans les perspectives.

5.1.4 Conclusion sur l'intérêt du cadre de modélisation proposé

Selon Levins (1966), la démarche de modélisation nécessite de réaliser un compromis entre i) généralité : le modèle doit capturer des patterns généraux, transférables à plusieurs systèmes différents, ii) réalisme : le modèle doit capturer fidèlement le fonctionnement d'un système particulier et iii) précision : les simulations du modèle doivent être les plus précises possibles. La démarche d'intégration de la diversité microbienne proposée ici peut être comparée à l'approche dominante de la littérature d'ajout de compartiments microbiens pour représenter différents groupes fonctionnels selon ces critères.

En choisissant un modèle mécaniste de base simple, fondée sur des représentations générales des processus de décomposition/minéralisation, nous avons vu que le modèle peut s'appliquer à plusieurs systèmes tout en capturant les patterns généraux des cinétiques d'émissions de CO₂. Le modèle de base utilisé est donc très générique. Cette **généricité** ne semble pas altérée par la prise en compte des fonctions de modulation, bien que ces dernières doivent être confirmées dans d'autres études.

La généralité est souvent altérée lorsque les mécanismes représentés deviennent plus précis. Par exemple, Perveen et al. (2014) utilisent deux groupes de micro-organismes pour mieux représenter le priming effect, Moorhead and Sinsabaugh (2006) utilisent trois groupes pour mieux représenter les interactions avec les différents sous-pools de MOF tandis que Miki et al. (2010) utilisent deux groupes pour représenter les relations entre diversité microbienne et diversité végétale. La représentation de la diversité dans ces modèles est pertinente voir performante pour l'objectif de modélisation fixé mais reste spécifique à cet objectif.

Les modèles représentant des mécanismes de manière détaillée augmentent leur réalisme vis à vis de l'objectif de modélisation. À l'inverse, la généralité de notre approche en fait un modèle capable de capturer une cinétique globale et de la hiérarchiser entre différents sols. La diversité n'est pas représentée de manière explicite, son influence sur la dynamique du C est seulement représentée à travers une fonction générale, sans représentation réaliste des mécanismes sous-jacents.

La précision des modèles est difficile à comparer car il faudrait évaluer les modèles sur les même jeux de données. Il n'est pas impossible que des modèles détaillés soient très précis pour l'objectif de simulation fixé. Toutefois, ces modèles sont souvent difficilement calibrable, rendant cette précision difficile à évaluer. La prise en compte de fonctions de modulation a amélioré la qualité prédictive de notre modèle de base. Cela met en évidence que cette approche a un impact

positif sur la précision du modèle.

En utilisant un modèle mécaniste simple, nous avons obtenu un modèle générique capable de capturer les cinétiques d'émissions du C. L'apport d'une approche plus empirique pour construire des fonctions de modulation des paramètres dépendantes de facteurs contrôlant la dynamique du C a permis d'augmenter la précision des simulations et surtout de mieux hiérarchiser les différences entre sols. Cette approche de modélisation se place alors dans une volonté d'optimiser généricité et précision, ce qui lui permet de tendre vers le développement de modèles opérationnels.

L'approche dominante de développement de modèle intégrant la diversité microbienne conduit à des modèles plus complexes, très réalistes et probablement précis vis à vis de l'objectif de modélisation. Ces modèles sont incontournables pour formaliser et tester des hypothèses concernant l'impact de la diversité microbienne sur la dynamique du C. Ils permettent une meilleure compréhension des processus sous-jacents et offre un cadre d'expérimentation théorique pour comprendre l'impact des processus représentés selon différents scénarios.

L'approche développée dans ce travail constitue alors un cadre méthodologique complémentaire pour une plus grande prise en compte de la diversité microbienne dans les modèles de dynamique du C. Elle n'est pas compétitive vis à vis de la représentation et compréhension des processus mais permet de tendre vers le développement de modèles opérationnels. Elle peut être utilisée en parallèle de l'approche dominante utilisée dans la littérature. Cette dernière, en précisant les processus permettrait d'alimenter le choix des modèles de base à utiliser selon un objectif donné, tandis que notre approche offre une alternative pour rendre les modèles complexes plus opérationnels.

En conclusion, le cadre d'intégration de la diversité dans les modèles de dynamique du C que nous proposons ici constitue une approche pertinente pour la représentation parcimonieuse de l'influence globale de la diversité sur la dynamique du C. Elle est facilement transférable car elle ne contraint pas l'utilisateur au modèle très simplifié utilisé dans ce travail, mais peut être aisément testée sur des modèles de complexités différentes, ainsi que pour d'autres propriétés des sols. Les méthodes GAM et d'inférence bayésienne utilisées sont des outils très intéressants dont nous recommandons l'utilisation pour accompagner l'approche de modélisation développée ici. Ces méthodes sont encore peu présentes dans le domaine du fonctionnement du sol et ont montré à travers ce travail qu'elles pouvaient répondre à certaines contraintes souvent rencontrées (e.g. relations non linéaires, sur-paramétrage de modèle) et aider dans cette démarche.

5.2 Implication des résultats sur l'influence de la diversité microbienne dans la modulation des cinétiques de décomposition et minéralisation du C

5.2.1 Cohérence du rôle de la diversité bactérienne dans la minéralisation de la MOS entre résultats, processus et modélisation

Parmi les fonctions de modulation prises en compte, celle concernant la diversité bactérienne a participé à l'amélioration des simulations du modèle en venant moduler les valeurs du rendement d'utilisation du C (Y_{MOS}). Ce choix vient de l'observation des résultats, lors de la sélection GAM, qui ont montré une relation négative entre les flux de CO₂ non marqué et la diversité bactérienne. À notre connaissance, cette relation négative n'a été trouvée que dans une très récente étude qui montre une diminution du taux de respiration du sol après diminution de la diversité par manipulation expérimentale (Zhang and Zhang, 2016). Le fait que cette observation soit rare dans la littérature peut s'expliquer par :

1. L'absence de distinction des flux issus de la minéralisation du C de la MOS et de la minéralisation de la MOF dans certaines études (e.g. Griffiths *et al.*, 2001). Dans ce cas, le flux total de CO₂ après ajout de MOF est largement influencé par la minéralisation du C de la MOF, qui représente environ 75% du flux total dans notre étude, ne permettant pas d'étudier le lien entre minéralisation de la MOS et diversité microbienne.
2. L'étude est réduite à l'analyse du lien entre dynamique du C et diversité microbienne globale, ne permettant pas de séparer l'influence des diversités bactérienne et fongique (e.g. Juarez *et al.*, 2013).

De fait, nos résultats suggèrent une utilisation systématique de MOF marquée et de l'étude séparée de l'influence des diversités bactérienne et fongique pour une meilleure compréhension du lien entre diversité microbienne et dynamique du C.

D'après la théorie sur l'efficacité d'utilisation du C (CUE, cf. section 1.2), qui décrit le devenir du C dans les sols, cette « diminution » de respiration peut provenir de la diminution des différents termes de respiration (maintenance R_M , croissance R_G , « carbon overflow » R_O ou production de composés carbonés R_E), et peut être expliquée par différentes hypothèses :

1. Une diminution de l'activité globale des micro-organismes par inhibition de la décomposition : Zhang and Zhang (2016) expliquent cette hypothèse par l'augmentation d'interactions négatives dans une communauté microbienne à diversité élevée. Ces interactions pourraient être la production de protéines augmentant la création de micro-agrégats et diminuant l'accessibilité au substrat (King, 2011), ou des interactions trophiques négatives (Allen *et al.*, 2010).
2. Une meilleure efficacité de la communauté microbienne qui peut avoir deux causes (Hooper *et al.*, 2005) :
 - (a) A l'échelle individuelle, l'augmentation de la diversité pourrait s'accompagner d'une augmentation du nombre d'espèces plus efficaces. Ce serait donc une augmentation de populations microbiennes avec une CUE_P (échelle individuelle) plus élevée conduisant à une CUE_C (échelle de la communauté) plus élevée. A cette échelle, il s'agirait alors plutôt du paramètre de maintenance (R_M) ou des paramètres d'anabolisme (R_G et R_E) (Geyer *et al.*, 2016).
 - (b) A l'échelle de la communauté, l'augmentation de la diversité pourrait s'accompagner d'une augmentation d'interactions positives comme : la complémentarité de niches écologiques à la fois entre les différentes bactéries (Allison, 2005; Hooper *et al.*, 2005) mais également entre bactéries et champignons (de Boer *et al.*, 2005; Johnston *et al.*, 2016), conduisant à des relations syntrophiques (Burns *et al.*, 2013) (paramètres R_G et R_E) ou permettant de mieux gérer les déséquilibres stœchiométriques (Mooshammer *et al.*, 2014, paramètre R_O).

Bien que nous ne pouvons pas définitivement exclure l'hypothèse de Zhang and Zhang (2016), celle-ci semble être moins probable dans notre étude. En effet, les différences de diversité qu'ils ont étudiés viennent d'une manipulation de la diversité d'une même communauté à l'aide de la méthode d'érosion (ou dilution) présentée en chapitre 1. Cette méthode a pu s'accompagner d'une diminution d'interactions négatives car la diminution de la diversité bactérienne a été accompagnée d'une diminution de la diversité d'autres groupes trophiques comme les champignons, responsables de la production des protéines données en exemple, ou les virus, responsables

des relations trophiques négatives données en exemple également. Dans notre cas, la diversité n'ayant pas été manipulée, celle concernant les bactéries n'est pas corrélée avec la diversité des champignons (Figure 2.5) et probablement pas avec les autres groupes trophiques.

Nous avons choisi d'admettre que la relation négative entre respiration et diversité bactérienne venait plutôt d'une relation positive entre CUE et diversité bactérienne, autrement dit que l'efficacité d'une communauté microbienne était liée, d'une manière ou d'une autre, à sa diversité bactérienne. Le suivi d'autres variables telles que la biomasse microbienne, ou encore des produits carbonés microbiens (Six *et al.*, 2006), permettrait d'être plus précis et de mieux cibler les paramètres du modèle à moduler.

En termes de modélisation, nous avons traduit cette hypothèse par une modulation du paramètre de rendement d'utilisation du C (Y_{MOS}) par une fonction dépendante de la diversité bactérienne. L'hypothèse sous-jacente est alors que la diversité bactérienne est liée à la respiration pour à la croissance microbienne (R_G). Cette modulation a permis d'augmenter la qualité prédictive du modèle de dynamique du C utilisé. Cependant, cela ne constitue pas une preuve de la véracité de cette hypothèse et il serait intéressant de tester la modulation d'autres paramètres de respirations (R_M , R_O , R_E). En particulier, Wang and Post (2012) ont souligné l'importance de séparer la respiration de croissance (R_G) de la respiration de maintenance (R_M) dans les modèles de dynamique de C ce qui pourrait aider à simuler l'effet de facteurs environnementaux sur le compromis entre maintenance et croissance microbienne (Lipson *et al.*, 2009). De plus, une meilleure prise en compte du « *carbon overflow* » (R_O) peut aider à modéliser efficacement le couplage de la dynamique du C avec la dynamique des autres nutriments du sol (N, P) (Manzoni and Porporato, 2009; Sinsabaugh *et al.*, 2013).

Comprendre les différents mécanismes qui régissent la CUE constitue un élément essentiel de l'évolution des modèles de dynamique du C. La suggestion de nos résultats sur un effet de la diversité bactérienne doit être étudiée plus profondément pour i) confirmer/infirmier notre hypothèse et ii) identifier dans quels mécanismes la diversité bactérienne a un impact non négligeable. Les analyses de sensibilité ont montré une importance forte de la CUE dans la prédiction des modèles (Six *et al.*, 2006; Lee and Schmidt, 2014). Bien modéliser ces mécanismes permettraient donc d'augmenter la prédiction des modèles, ce qui est confirmé par les conséquences fortes de la manière dont est modélisé ce paramètre sur les sorties des modèles (Allison *et al.*, 2010).

5.2.2 Un rôle ambivalent de la diversité fongique dans la dynamique du C

Malgré une relation positive observée entre la diversité fongique et la décomposition de la MOF, l'hypothèse émise d'une modulation positive de la constante de décomposition d'un sous-pool de la MOF n'a pas permis une amélioration du modèle. Cette hypothèse est en accord avec les connaissances sur l'influence des champignons sur la dynamique du C (de Boer *et al.*, 2005). Cependant, il a souvent été observé une redondance fonctionnelle rapidement atteinte avec l'augmentation de la diversité fongique, bien qu'il se pourrait qu'elle soit attribuable aux méthodes utilisées pour étudier la relation entre décomposition et dynamique du C (van der Wal *et al.*, 2013).

Cette redondance pourrait expliquer la forme de la relation observée (présence d'une asymptote) ainsi que la calibration de la fonction de modulation conduisant à une fonction constante sur la gamme de variabilité de la diversité en champignons. Nous n'avons pas réussi à comparer la gamme de variabilité de la diversité en champignons à d'autres valeurs du fait de l'hétérogénéité des protocoles. Il semble primordiale pour la compréhension du lien entre diversité fongique et dynamique de la MOF de tendre vers une meilleure harmonisation des protocoles.

Bien qu'il existe une variabilité de minéralisation de la paille de blé dans nos incubations, cette variabilité reste relativement faible compte tenu du coefficient de variation des flux cumulés qui ne dépasse pas 8% après 80 jours d'incubation. Cette variabilité restreinte peut être expliquée par l'utilisation d'un résidu de culture unique. En effet, le premier facteur de contrôle de la décomposition d'un résidu en condition contrôlée est sa qualité (Nicolardot *et al.*, 2001) devant même la diversité microbienne (Cleveland *et al.*, 2014). La diversité microbienne n'est alors peut être pas assez déterminante dans la décomposition du résidu, ce qui expliquerait sa faible contribution.

Cependant, d'autres études montrent l'importance de la diversité microbienne dans la décomposition des résidus de culture (Allison *et al.*, 2013; Miura *et al.*, 2015) en interaction avec la qualité du résidu (Moorhead and Sinsabaugh, 2006). L'étude a été restreinte à la seule décomposition de la paille de blé. Dans le contexte agro-écologique, une pratique possible est la multiplication des espèces cultivées et donc des résidus de culture à décomposer par les micro-organismes (Wezel *et al.*, 2014). Il serait alors intéressant d'étudier l'influence de la diversité

microbienne du sol sur la capacité des communautés microbiennes à dégrader différents résidus de culture.

En particulier, les champignons, en plus de leur contribution dans la décomposition des composés celluloseux, ont une importance très forte dans la décomposition de la lignine (de Boer *et al.*, 2005). Il est possible que dans notre étude, l'influence de la décomposition de la lignine sur le flux total n'ait pas été important en raison d'une période d'incubation trop courte, réduisant ainsi l'importance de la diversité fongique. Ainsi, dans leur modèle de décomposition de la litière, Moorhead and Sinsabaugh (2006) décrivent trois groupes de décomposeurs. L'un de ces groupes est performant pour décomposer la cellulose et la lignine et est composé majoritairement de champignons, qui ont la capacité de s'introduire dans le réseau cellulo-lignique. Analyser l'impact de la diversité fongique sur cette capacité pourrait aider à mieux appréhender l'influence de la diversité fongique sur la décomposition des résidus de culture et peut-être à trouver des fonctions de modulation et des paramètres à moduler adéquats pour réussir à simuler la variabilité de dynamique du C du résidu entre différents sols sans complexifier le modèle.

Une relation positive a également été trouvée dans MOSAIC, entre la diversité fongique et les flux de CO₂. Cette relation positive a été interprétée comme une augmentation de la décomposition avec la diversité fongique mais cette hypothèse a été mise en défaut dans l'étape de modélisation. Le lien entre diversité fongique et dynamique du C s'avère alors complexe à identifier. Une synthèse des différentes relations observées dans la littérature montre que des relations négatives, positives et neutres ont été trouvées pour les champignons (van der Wal *et al.*, 2013). Les auteurs ont conclu qu'aucune relation uniforme entre la décomposition de la MO et la diversité fongique n'a été démontrée pour le moment. Ils ajoutent que le type d'interactions prédominantes, la présence d'espèces capables d'une activité de décomposition extra-ordinaire, la composition du substrat ou encore l'échelle spatiale considérée peuvent être des déterminants de la nature de cette relation. La question de la diversité fongique demande donc une étude approfondie pour i) déterminer si son influence sur la dynamique du C est assez importante pour aider à hiérarchiser différents sols en fonction des sorties de C et ii) identifier pour quel(s) paramètre(s) il serait pertinent de moduler les valeurs par une fonction dépendante de la diversité fongique.

5.3 Limites et perspectives de recherche

5.3.1 Des données riches mais un besoin d'une harmonisation des mesures de diversité

L'une des plus grandes forces de ce travail réside dans les **données disponibles** pour répondre aux objectifs que l'on s'est fixés (Tableau 5.1). Ces données nous ont permis d'examiner le lien entre diversité microbienne et dynamique du C au travers de différentes questions : différence entre modes d'usage, dépendance temporelle, pertinence des indices de diversité pour expliquer la variabilité de dynamique du C dans un jeu de données où leur variabilité est moins importante (MOSAIC) que dans l'autre (DIMIMOS).

La valorisation de données existantes pour répondre à des objectifs qui ne sont pas ceux pour lesquels les données sont destinées à l'origine a tout de même permis de répondre aux objectifs. D'autres informations sont disponibles sur la structure et la composition des communautés microbiennes qui pourraient apporter d'autres résultats sur le lien entre dynamique du C et diversité microbienne.

Les limites de ces deux jeux de données résident dans l'impossibilité de comparer les valeurs de diversité en raison i) de protocoles de stockage des sols différents et ii) d'un manque d'harmonisation des protocoles de mesure de la diversité (Lemanceau *et al.*, 2015). Dans notre cas, ce manque d'harmonisation nous a limités pour la construction des distributions *a priori* des paramètres des fonctions de modulation car il était difficile de placer les valeurs de diversité des jeux de données dans la gamme possible de variation de ces valeurs. Dans l'optique d'utiliser ces indices dans les modèles de dynamique de C ou de simplement pouvoir comparer des valeurs obtenues dans un sol aux valeurs possibles dans une même situation, il est nécessaire aujourd'hui de renforcer l'harmonisation des protocoles et de construire des valeurs de référence.

5.3.2 Perspectives immédiates de modélisation

Le choix du modèle a été principalement motivé par une volonté de simplicité tout en représentant les mécanismes sur lesquels des hypothèses ont été faites. Malgré une qualité globale raisonnable, ce choix a conduit à un modèle avec une mauvaise qualité de simulation des dynamiques du C issu de la MOS en début d'incubation. L'utilisateur du modèle peut être amené à complexifier le modèle pour améliorer les simulations ou représenter certains processus nécessaires. Par exemple, les cinétiques utilisées du premier ordre ne sont peut être pas les plus

pertinentes et des équations plus complexes pourraient être envisagées (e.g. [Wutzler and Reichstein, 2008](#); [Wang and Post, 2013](#); [Wieder et al., 2013](#)). Tester différentes représentations pour approcher le meilleur modèle en termes de qualité prédictive constitue donc une première étape d'amélioration de l'étape de modélisation, afin d'évaluer au mieux l'apport de l'intégration de fonctions de modulation dépendantes de diversité microbienne.

Cette « complexification » possible des modèles s'accompagne d'un plus grand nombre de paramètres à calibrer, augmentant le sur-paramétrage et donc les problèmes de calibration déjà présents dans notre modèle « simple ». Bien que le suivi de plusieurs variables peut compenser le nombre de paramètres à calibrer (cf. section suivante), cet ajout de paramètres devra s'accompagner d'une construction de distributions *a priori* pour aider à la calibration bayésienne. Il semble alors nécessaire d'améliorer l'étape de construction des distributions *a priori* en effectuant une recherche détaillée de valeurs de paramètres possibles selon les hypothèses utilisées pour représenter les dynamiques de C .

Par exemple, lors de la construction des distributions *a priori* dans ce travail de thèse, nous avons vu que certains paramètres avaient des distributions peu précises. C'est le cas du paramètre ks dont le paramètre de dispersion a une valeur élevée, ce qui s'explique par un échantillonnage trop dispersé de valeurs de ce paramètre. Restreindre l'échantillonnage de ces valeurs à des situations qui correspondent parfaitement à celle utilisée aurait permis un choix de distributions *a priori* plus adéquates, ce qui peut influencer considérablement les distributions *a posteriori* obtenues ([Morita et al., 2010](#)). Bien choisir ces paramètres demandent un travail important de tri d'articles rapportant des valeurs de paramètres. Ce travail reviendrait à construire une base de données reprenant les différentes valeurs de paramètres selon les hypothèses utilisées dans le modèle et les situations dans lesquelles les modèles sont appliquées. Cette base de données auraient une valeur considérable quant à la construction de distributions *a priori* pertinentes.

En termes de modulation de paramètres, d'autres hypothèses ont été soulevées et pourraient être testées rapidement. C'est le cas des fonctions de modulation dépendantes de la diversité fongique, qui n'ont pas permis une amélioration de la qualité prédictive des modèles en modulant les valeurs des constantes de décomposition, mais qui pourraient également moduler les valeurs des paramètres définissant la taille des pools de MOF_2 et MOS dans le modèle (i.e. paramètre p_2 et fraction f , respectivement).

Enfin, dans le chapitre 1, le schéma conceptuel de lien entre fonctionnement du sol et diversité microbienne qui a été la base de la forme des fonctions de modulation dépendantes de la diversité

microbienne a été présentée avec une variabilité. Cette variabilité correspond à l'effet sous-jacent de la composition des communautés microbiennes : globalement le fonctionnement augmente avec la diversité microbienne mais cette augmentation est variable selon la composition de la communauté. C'est ce qu'on appelle l'idiosyncrasie (Nielsen *et al.*, 2011).

Cette variabilité pourrait être facilement estimée par inférence bayésienne en considérant les fonctions de modulation non pas comme des valeurs fixes dépendantes de la diversité microbienne mais comme des variables aléatoires suivant une loi de probabilité (e.g. loi normale) dont la moyenne serait égale à une fonction dépendante de la diversité et dont l'écart-type serait également estimé. Cet écart-type, fixe ou dépendant de la diversité également (i.e. hétéroscédasticité), correspond à cet effet d'idiosyncrasie. Il est important de noter que lors d'une calibration des fonctions de modulation une par une comme effectuée dans ce travail de thèse, l'estimation de l'écart-type serait sans doute surestimée car elle intégrerait la variabilité des autres facteurs non pris en compte. Malgré un temps de calcul plus long et des problèmes de convergence probables, une prise en compte de toutes les fonctions de modulation sembleraient donc plus appropriée.

5.3.3 Mise en place d'une expérimentation pour la confirmation des hypothèses et une meilleure calibration de modèle

5.3.3.1 S'affranchir des confusions entre diversité microbienne et autres propriétés du sol

TABLE 5.2 – Avantages et inconvénients des méthodes possibles.

| Méthodes | Avantages | Inconvénients |
|---|--|---|
| Échantillonnage pertinent dans une base de données | <ul style="list-style-type: none"> - Communautés réalistes - Différentes stratégies d'échantillonnages possibles (e.g. cLHS) | <ul style="list-style-type: none"> - Manque d'harmonisation des protocoles - Ré-échantillonnage de sols sur une échelle pouvant être large |
| Manipulation de la diversité par méthode de dilution | <ul style="list-style-type: none"> - Facilité de mise en place - Conditions très contrôlées | <ul style="list-style-type: none"> - « <i>Hidden effects</i> » (Huston, 1997) - Manipulation non distincte des diversités bactérienne et fongique |
| Inoculation d'un sol avec des communautés issus de différents sols | <ul style="list-style-type: none"> - Conditions très contrôlées - Communautés réalistes | <ul style="list-style-type: none"> - Ré-échantillonnage de sols sur une échelle pouvant être large - Effets sur la diversité? |

Le fait de ne pas avoir manipulé la diversité microbienne nous a empêché de pouvoir conclure à un lien direct entre cette diversité et la dynamique du C. Cette limite a été aggravée par la

multitude de corrélations, et donc de confusions, qu'il existe entre les différentes propriétés du sol. Pour tester les hypothèses et pouvoir faire un lien plus direct entre dynamique du C et diversité microbienne, une possibilité consisterait à éliminer les confusions avec les autres propriétés. Pour cela, différentes possibilités sont présentées dans le tableau et discutées ci-après.

Les trois propositions pour s'affranchir des confusions entre diversité et autres propriétés du sol sont, dans l'ordre subjectif de pertinence :

1. La confusion des effets dans les deux bases de données utilisées dans ce travail de thèse pourrait être éliminée en faisant un sous-échantillonnage pertinent des sols afin de sélectionner un ensemble de sols qui diffèrent par plusieurs de leurs propriétés, dont la diversité microbienne, mais dont les corrélations entre ces propriétés sont proches de zéro. De cette manière, on obtient des sols avec des communautés microbiennes réelles dans des sols dont les propriétés sont orthogonales entre elles, ce qui éliminerait le problème de confusion. Cet échantillonnage serait facilement réalisable à l'aide de méthodes comme les carrés latin conditionnés (cLHS, [Minasny and McBratney, 2006](#)) et pourrait être appliqué aux sols de MOSAIC par exemple. Cependant, nous avons vu qu'il serait plus pertinent d'estimer les fonctions de modulation indépendamment de la calibration des modèles. Dans cette optique, il serait plus intéressant d'avoir un ensemble de sols couvrant toute la gamme de diversité, ce qui n'est sûrement pas le cas dans MOSAIC.
2. Le moyen le plus sûr d'avoir un maximum de conditions contrôlées est d'utiliser un unique sol. Dans ce cas, pour manipuler la diversité, la méthode de dilution (appelée érosion de diversité) présentée dans le chapitre 1 serait appropriée. Cette méthode a l'avantage d'être relativement simple à mettre en place. Son principal inconvénient réside dans ce que [Huston \(1997\)](#) appelle les « effets cachés ». Ici cet effet est l'élimination préférentielle des espèces rares lors des dilutions successives pour diminuer la diversité ce qui peut biaiser les résultats. De plus, on ne peut pas séparer la manipulation de la diversité bactérienne et de la diversité fongique ce qui créerait une confusion entre ces deux propriétés.
3. Une autre méthode pourrait être un mélange des deux précédentes. Après avoir sélectionné des sols qui diffèrent par leur diversité, les communautés de ces sols pourraient être inoculées dans un unique sol. Cette méthode combine les avantages d'avoir des communautés très réalistes tout en étant certain de contrôler toutes les autres propriétés. Cependant, les autres propriétés du sol influençant la diversité, il est possible qu'une fois ré-inoculées, les diversités changent et convergent vers une même diversité. À notre

connaissance, cette méthode n'a jamais été appliquée mais il serait intéressant d'évaluer l'effet de la ré-inoculation sur la diversité.

À partir des ensembles de sols obtenus quelle que soit la méthode appliquée, la même approche générale que celle appliquée dans ce travail pourrait être mise en place. Un premier changement consisterait cependant à estimer les fonctions de modulation indépendamment de l'étape de calibration du modèle. Pour cela, il faudrait calibrer le modèle pour chaque sol (i.e. avoir un jeu de paramètres pour chaque sol étudié) et le mettre en relation avec les valeurs de diversité. Cela implique d'avoir une calibration précise de la valeur des paramètres qui pourrait être possible à l'aide d'une amélioration de l'expérimentation mise en place pour obtenir les données.

5.3.3.2 Mise en place d'une expérimentation pour une meilleure calibration du modèle

La principale amélioration apportée serait un suivi de la dynamique de plus de variables que les seuls flux de CO₂ au cours de l'incubation. Ces variables seraient *a minima* la biomasse microbienne mais pourrait être également le suivi des différents pools de MO, le suivi des produits microbiens (Six *et al.*, 2006) et également le suivi des éléments minéraux du sol (N et P), pour les raisons qui ont été exposées dans la deuxième section de ce chapitre. La contrainte apportée par le suivi de ces variables est que certaines nécessitent la destruction des échantillons. Suivre ces variables signifie alors de faire plus de répétitions au départ par modalités (e.g. par sols) car certaines de ces répétitions seront détruites au cours du suivi de la dynamique de ces variables. *In fine*, cela augmente considérablement le nombre d'incubations à effectuer, et donc impose un nombre limité de modalités de diversité à analyser. D'autre part, nous avons vu qu'il serait sans doute intéressant d'allonger la période d'incubation pour pouvoir analyser l'effet de la diversité sur le stockage de C sous forme de MOS.

À partir de ces remarques, nous proposons les modifications suivantes de l'expérimentation effectuée dans MOSAIC qui comprenait au départ 100 sols x 3 répétitions = 300 incubations par traitement (témoins vs amendés) :

- Réduire le nombre de sols/modalités de diversité à 25 et commencer avec 12 répétitions par sols/modalités de diversité. Les 25 sols pourraient être le résultat d'un échantillonnage en carré latin conditionné présenté dans la section précédente ou de différentes modalités de dilution pour manipuler la diversité.
- Selon les mêmes conditions, incuber les sols selon deux traitements : témoins vs amendés

avec un résidu marqué. Aux dates 1, 3, 7, 14, 21, 40, 80, 120 et 200 jours, mesurer le CO₂ marqué et non marqué accumulé. Ces dates sont données à titre indicatif, l'idée sous-jacente étant de faire des mesures plus rapprochées en début d'incubation.

- Aux dates 1, 7, 40 et 200 prélever 3 répétitions par sols/modalités de diversité pour effectuer les autres suivis de variables marquées et non marquées quand nécessaire (biomasse, produits microbiens, N, P). Les dates sont également données à titre indicatif mais correspondent globalement aux dates les plus pertinentes au vu de la dynamique possible de la biomasse microbienne sur 60 jours après ajout de glucose marqué (Blagodatsky *et al.*, 2010).

Une telle expérimentation permettrait de pallier la plupart des limites identifiées dans ce travail de thèse et serait donc intéressante. Il est à noter qu'il y aurait un coût supplémentaire du aux suivis des variables supplémentaires. Ce coût pourrait sans doute être amorti par la réduction du nombre de sols/modalités de diversité étudiés (e.g. 20 au lieu de 25). Une analyse plus fine des coûts est nécessaire pour bien définir cette expérimentation.

Une dernière modification peut être faite. Il s'agit de l'ajout de différents résidus à la place d'un unique. On pourrait par exemple étudier deux résidus avec des qualités contrastées (caractérisées par leur C/N) et réduire le nombre de sols/modalités de diversité pour avoir une expérimentation de la même envergure que celle dans MOSAIC.

5.3.4 Quels autres processus peut-on envisager de modéliser pour une meilleure gestion de la MOS ?

Certains processus sont absents du modèle : la dynamique de la biomasse microbienne et de ses produits (exsudats, enzymes, lyse cellulaire), la stabilisation/protection de la MOS et le couplage avec d'autres éléments. Ces processus pourraient être intéressants à modéliser selon les objectifs voulus du modèle. Suivant la même philosophie que le travail réalisé au cours de cette thèse, ces processus pourraient être pris en compte de la manière la plus simple possible tout en étant modulés par des fonctions dépendantes des propriétés des sols. Nous discutons ici de l'intérêt de modéliser ces processus.

Bien que sa prise en compte n'aurait eu que peu d'impact sur la quantité totale de MOS sur la période d'incubation étudiée, il est certains que le flux de C de la biomasse microbienne vers la MOS (i.e. *EX* et *BD*) n'est pas négligeable dans la période d'incubation et que sa prise en compte

permettrait de mieux simuler la dynamique de la biomasse microbienne. La complexification du modèle aurait demandé d'avoir mesuré la dynamique de la biomasse microbienne pour permettre une bonne calibration.

Le suivi de la biomasse aurait pu confirmer l'hypothèse d'une CUE plus élevée lorsque la diversité bactérienne est élevée. De plus elle permettrait d'en savoir plus sur le devenir de cette biomasse permettant une meilleure simulation de la CUE à différentes échelles de temps (Geyer *et al.*, 2016). Il est généralement admis qu'une CUE élevée est favorable à la dynamique du C dans le sol. En effet, une CUE élevée signifie plus de biomasse microbienne produite pour une même quantité de C décomposé.

Les conséquences de cette augmentation de biomasse microbienne sont variables et dépendent de la CUE. Sur quelques jours, une augmentation de la biomasse microbienne pourrait entraîner une augmentation de la production d'enzymes extra-cellulaires conduisant à une augmentation de la décomposition, et donc potentiellement augmentant le déstockage du C ainsi que la disponibilité en nutriments du sol pour les plantes (Manzoni *et al.*, 2012). Cependant, une CUE élevée signifie également une demande accrue pour les autres nutriments du sol pour répondre aux exigences stœchiométriques des micro-organismes, et donc une compétition accrue pour ces mêmes nutriments entre micro-organismes et plantes (Manzoni *et al.*, 2012). Cette remarque est d'autant plus importante si l'augmentation de biomasse concerne les bactéries qui ont un ratio C/N assez bas (~ 4 , Sylvia *et al.*, 2005).

Une augmentation de la biomasse signifie aussi à plus long terme (i.e. semaines, mois) une augmentation de composés microbiens et de produits issus de la mort des micro-organismes (Six *et al.*, 2006). Cotrufo *et al.* (2013) définissent le concept de « *Microbial Efficiency - Matrix Stabilization* » (MEMS) qui décrit la décomposition, transformation et stabilisation de la MO comme un continuum où la CUE microbienne et les mécanismes de stabilisation sont les facteurs clés du devenir du C dans le sol. Ce concept est appuyé par la proportion significative de biomasse microbienne morte qui constitue la MOS (Gougoulias *et al.*, 2014). Ainsi, bien qu'une augmentation de biomasse microbienne (due à une CUE élevée) ne peut être traduite directement par une séquestration de C dans le sol, car les facteurs de stabilisation sont prépondérants, elle constitue la première étape nécessaire.

La confirmation d'une augmentation de la CUE avec la diversité bactérienne constituerait un facteur important de la gestion de la MO dans le sol du fait des éventuelles conséquences sur la disponibilité des nutriments du sol pour les plantes et de la séquestration de ces mêmes nu-

triments. La diversité bactérienne étant un paramètre influencé par les pratiques agricoles, elle pourrait être un moyen de gérer correctement la MO du sol pour alterner entre fourniture des nutriments aux plantes lorsqu'elles en ont besoin et séquestration autrement. Cette notion de « banque de nutriments » a déjà été développée par Fontaine et al. (2011) soulignant l'importance des champignons dans la séquestration du C. L'importance des champignons pourraient alors marcher de concert avec la diversité bactérienne pour synchroniser la disponibilité des nutriments avec le besoin des plantes et contribuer à la séquestration à long terme. En ce sens, le schéma que propose Geyer et al. (2016) permettrait de savoir quelle CUE est affectée par la diversité bactérienne : la CUE_C qui est liée à la quantité d'enzymes produites (court-terme) ou la CUE_E plutôt liée à la mort des micro-organismes (long-terme). Bien évidemment, les CUE sont influencées par d'autres facteurs (humidité, ratio C/N,...) qu'il est important de considérer et dont l'importance relative par rapport à la diversité bactérienne est à déterminer (Manzoni et al., 2012).

Le rôle de la diversité fongique sur la dynamique du C issu de la MOS a été moins marqué dans notre étude. Une relation négative entre la richesse fongique et les flux de CO_2 a été trouvée dans les sols amendés pour DIMIMOS et dans les sols témoins et amendés aux première et deuxième périodes d'incubation respectivement pour MOSAIC. Cette relation peut être interprétée de la même manière que la relation négative entre dynamique du C de la MOS et diversité bactérienne discutée dans la section précédente. Une diversité fongique plus importante pourrait alors aussi être responsable d'une CUE plus grande, appuyant l'intérêt de prendre en compte à la fois diversité fongique et diversité bactérienne.

Cependant, en ce qui concerne les champignons, l'hypothèse de Zhang and Zhang (2016) d'une augmentation de l'activité de décomposition entraînée par une diminution de la production de protéines protégeant la MOS lorsque la diversité fongique diminue est également valable, ce qui ne facilite pas l'interprétation. Une diminution des flux de CO_2 en présence d'une diversité fongique élevée a également été observée pendant la décomposition de bois mort au Japon (Yang et al., 2016). Les auteurs ont également conclu à une diminution du taux de décomposition (par protection du substrat) lorsque la diversité fongique est élevée, ce qui est en accord avec d'autres études qu'ils citent, ainsi qu'avec l'importance reconnue aujourd'hui des champignons dans la séquestration du C (de Boer et al., 2005; Six et al., 2006).

Le caractère ponctuel de cette relation ainsi que les valeurs d'importance relative nous ont

amenés à la négliger dans l'étape de modélisation mais une étude approfondie semble nécessaire ici. En particulier, si l'importance de la diversité en champignons est en relation avec la séquestration du C, notre étude ne nous permet pas de l'évaluer en raison d'un temps d'incubation trop court. Pour déceler la relation négative des flux de CO₂ avec la diversité en champignons, Yang et al. (2016) ont mesuré ces flux sur 3 ans. Il semble donc qu'une étude plus approfondie concernant l'effet de la diversité en champignons sur la séquestration du C soit utile, en particulier dans une optique de modélisation de la dynamique du C sur le long terme.

De par leur physiologie et leur physiologie, les champignons sont souvent associés à la décomposition de molécules plus récalcitrantes que les bactéries (plus grande capacité enzymatique, C/N plus grand, turnover plus lent) et contribuent fortement à la séquestration du C (meilleure CUE, augmentation de la stabilité structurale) bien qu'ils peuvent être responsables d'un déstockage de C au vu de leur contribution dans le priming effect (de Boer *et al.*, 2005; Fontaine and Barot, 2005; Six *et al.*, 2006; Waring *et al.*, 2013).

Ainsi, la diversité en champignons pourrait avoir différents niveaux d'influence sur la dynamique du C de la MOS : un niveau continu sur la décomposition de la MOS dont la nature de la relation reste à déterminer (dualité entre décomposition accrue et protection de la MOS) et un niveau plus long-terme qui participe à la séquestration du C. L'intégration de la diversité fongique dans les modèles de dynamique de C pourrait se faire via la modulation de différents paramètres selon les objectifs du modèle :

1. Une modélisation de la dynamique intra-annuelle comme celle étudiée dans ce travail de thèse où la diversité aurait une influence sur les paramètres de décomposition ou sur la quantité de substrat disponible à la décomposition
2. Une modélisation plus long terme où il faudrait ajouter l'influence de la diversité fongique sur la séquestration du C

Compte tenu de l'importance possible de la diversité fongique et de son interaction avec la diversité bactérienne, il est nécessaire de conduire d'autres expérimentations pour y voir plus clair dans la relation entre diversité microbienne et dynamique du C de la MOS, et identifier leur place respective dans les modèles.

Le couplage avec d'autres nutriments n'était pas possible du fait d'un manque de mesure de la dynamique de ces nutriments. Au cours de ce travail, et sans surprise, des paramètres concernant les autres nutriments du sol (N, P) se sont avérés très importants dans la dynamique

du C renforçant l'idée qu'un couplage dans les modèles est nécessaire pour augmenter la qualité des prédictions de chacun des cycles. De plus, nous avons vu que par son impact sur la CUE, la diversité microbienne pouvait avoir un impact sur les autres nutriments (Manzoni *et al.*, 2012).

Le concept de CUE peut être élargi à l'azote (NUE) et au phosphore (PUE) (Mooshammer *et al.*, 2014). Notre étude n'a pas permis d'étudier la diversité microbienne sur les autres cycles mais au vu de l'importance du couplage, il serait intéressant de le faire. En particulier, si modifier la diversité microbienne selon le concept de banque de nutriments expliqué ci-dessus (Fontaine *et al.*, 2011) semble être une voie intéressante dans la gestion de la MO, connaître et prédire l'impact de la diversité microbienne sur les autres éléments du sol apparaît primordiale pour y arriver.

5.3.5 Perspectives d'application pour les transitions agro-écologiques

Le travail montre que la prise en compte des indices de diversité microbienne améliore la compréhension et la représentation de la dynamique du carbone, lorsqu'on est confronté à une diversité de systèmes de culture. Les indices qui sont particulièrement ressortis de ce travail sont la diversité bactérienne, à travers l'inverse de l'indice de Simpson, impliquée dans la dynamique du carbone du sol, et la diversité en champignons, à travers l'inverse de l'indice de Simpson et la richesse, impliquée dans la dynamique du carbone du sol et des résidus. Ces résultats restent à consolider par des travaux complémentaires.

Ces résultats concernent une diversité de systèmes agricoles :

- On montre que la diversité impacte la dynamique du carbone de la même manière sous culture que sous prairies de longue durée.
- On montre également que pour des gradients de systèmes de culture moins contrastés (rotations de cultures annuelles avec ou sans prairie temporaires, associées à des systèmes d'élevage), la diversité impacte aussi la dynamique du carbone.

En vue d'une application dans le cadre de transitions agro-écologiques, la démarche de modélisation initiée dans ce travail, fournit les bases pour une modélisation plus opérationnelle de la dynamique du carbone, intégrant le rôle de la diversité et s'élargissant à des échelles de temps plus longues.

Pour aller jusqu'à des recommandations de pratiques agricoles, il conviendrait aussi de faire le lien entre les gradients de diversité et de dynamique du C observés et les caractéristiques des systèmes de culture. Les jeux de données disponibles peuvent être mobilisés dans ce sens.

De plus, dans le cadre d'une transition agro-écologique, il conviendrait d'acquérir de nouvelles données dans des situations déjà conduites en agro-écologie et/ou en cours de transition.

Conclusion

Le travail effectué dans cette thèse s'intéresse à la prise en compte de la diversité microbienne dans la modélisation de la dynamique du C suite à l'ajout de paille de blé dans une expérimentation à court-terme (dynamique intra-annuelle). L'objectif était d'évaluer l'impact de l'intégration de la diversité microbienne sur la qualité prédictive des simulations de dynamique du carbone dans le sol avec la volonté de construire un modèle simple (i.e. peu de paramètres et données d'entrées accessibles). Pour cela, il a fallu i) proposer une stratégie d'intégration de la diversité microbienne dans un modèle de dynamique du carbone, ii) étudier les relations qui existent entre diversité microbienne et dynamique du carbone et iii) intégrer la diversité microbienne dans un modèle et évaluer l'impact sur la qualité prédictive de ce modèle.

L'approche novatrice utilisée ici était de considérer la diversité microbienne comme n'importe quelle autre propriété du sol et d'intégrer ce facteur de contrôle de la dynamique de la matière organique via des fonctions de modulation des paramètres du modèle de dynamique du carbone. Une première étape de sélection de prédicteurs des flux de C, à l'aide de modèles statistiques additifs généralisés, a montré que des indices classiquement utilisés en écologie sont pertinents pour faire le lien entre dynamique du C et diversité bactérienne et fongique. Ces indices, avec d'autres propriétés des sols, pris en compte dans des fonctions de modulation de paramètres d'un modèle de dynamique du C calibrées par inférence bayésienne, ont permis une amélioration de la qualité prédictive du modèle, tout en proposant un modèle générique simple.

L'ensemble de la méthode utilisée pendant ce travail de thèse constitue alors un cadre méthodologique facilement utilisable dans d'autres situations pour intégrer la diversité microbienne dans les modèles de dynamique de carbone au travers de fonctions de modulation des valeurs des paramètres de ces modèles. Les résultats trouvés ainsi que les interprétations apportées encouragent vivement à approfondir cette approche car elle constitue un bon compromis entre des modèles mécanistes détaillés, dont la complexité limite leur utilisation opérationnelle, et des modèles purement statistiques (i.e. empiriques), ne représentant pas les processus en jeu. À ce titre,

les limites de ce travail ont été identifiées et des propositions ont été faites pour les contrer dans des recherches futures.

En conclusion, le cadre d'intégration de la diversité dans les modèles de dynamique du C que nous proposons ici constitue une approche pertinente pour la représentation parcimonieuse de l'influence globale de la diversité sur la dynamique du C. Elle ne constitue pas en un développement d'un énième modèle de dynamique de carbone mais plutôt en un moyen transférable d'un modèle à un autre de prise en compte de la diversité microbienne. Compte tenu de l'importance de cette prise en compte dans la simulation de la dynamique de la matière organique, cette thèse apporte une réponse à une manière possible de le faire et qu'il faudra confirmer à l'avenir.

Annexes

Annexe A

Information sur les sols de DIMIMOS

TABLE A.1 – Information about soil sample locations in DIMIMOS database.

| Location | | Land-use | | | Climate | | | | |
|----------|---------------|------------------------|-----------------------|---------------|----------------|-----------------------------------|----------|----------|-------------|
| Site ID | Sampling year | Longitude ^a | Latitude ^a | Elevation (m) | Land-use class | Land-use description | AMT (°C) | AMP (mm) | AMETTP (mm) |
| 16 | 2003 | 60.4083 | 70.69637 | 134 | Grassland | Grassland < 10 years | 15.2 | 51.0 | 95.5 |
| 164 | 2003 | 71.4946 | 69.40768 | 45 | Grassland | Grassland < 10 years | 15.2 | 53.9 | 70 |
| 563 | 2005 | 53.7991 | 68.14326 | 176 | Cropland | Cropland + temporary grassland | 11.9 | 79.6 | 40.3 |
| 732 | 2006 | 45.7653 | 67.67023 | 55 | Grassland | Grassland < 10 years | 10.2 | 68.4 | 68.6 |
| 885 | 2003 | 36.1332 | 67.19848 | 47 | Cropland | Cropland + temporary grassland | 8.9 | 49.4 | 48.7 |
| 1184 | 2003 | 82.4147 | 66.20055 | 381 | Grassland | Grassland < 10 years | 7.5 | 48.1 | 58.5 |
| 1253 | 2004 | 61.6059 | 65.89839 | 397 | Cropland | Cropland + temporary grassland | 19.6 | 75.2 | 116.4 |
| 1376 | 2006 | 58.3687 | 65.42149 | 384 | Cropland | Cropland + temporary grassland | 4.7 | 136.8 | 14.7 |
| 1388 | 2006 | 77.5510 | 65.40535 | 439 | Grassland | Grassland < 10 years | 4.7 | 51.4 | 20.3 |
| 1409 | 2007 | 45.5674 | 65.27225 | | Cropland | Cropland | 17.2 | 70.6 | 85.9 |
| 1412 | 2006 | 50.3629 | 65.26829 | 137 | Cropland | Fallow | 16.2 | 83.4 | 77.4 |
| 1426 | 2004 | 72.7419 | 65.24956 | 320 | Cropland | Cropland | 14.7 | 78.4 | 107.5 |
| 1437 | 2007 | 90.3254 | 65.23458 | 390 | Cropland | Cropland + temporary grassland | 16.3 | 141.7 | 77.9 |
| 1589 | 2005 | 69.4910 | 64.61288 | 1140 | Grassland | Grassland < 10 years | 13.6 | 78.0 | 103.2 |
| 1714 | 2007 | 71.0490 | 64.13201 | 997 | Grassland | Grassland < 10 years | 10.7 | 78.6 | 86.3 |
| 1763 | 2008 | 80.6256 | 63.96405 | 445 | Grassland | Grassland < 10 years | 17.4 | 171.5 | 98.3 |
| 1829 | 2006 | 51.8288 | 63.66856 | 102 | Grassland | Grassland < 10 years | 7.0 | 76.3 | 16.7 |
| 1977 | 2007 | 70.9681 | 63.17302 | 850 | Cropland | Cropland + temporary grassland | 11.7 | 112.8 | 47.4 |
| 2114 | 2008 | 69.3295 | 62.69490 | 475 | Grassland | Grassland < 10 years | 19.5 | 39.5 | 161.7 |
| 2230 | 2008 | 64.4947 | 62.21948 | 244 | Cropland | Fallow | 7.2 | 68.2 | 37.6 |

AMT, Average Monthly Temperature; AMP, Average Monthly Precipitation; AMETTP, Average Monthly Evapotranspiration

^a Coordinates follow the Lambert-93 projection

Annexe B

Information sur les sols de MOSAIC

TABLE B.1: Information sur les sols de MOSAIC.

| Site ID | Longitude ^a | Latitude ^a | Culture présente à l'échantillonnage |
|---------|------------------------|-----------------------|--|
| 72 | 265008 | 6783282 | Pois |
| 73 | 265299 | 6783282 | Maïs grain |
| 78 | 264862 | 6783535 | Pomme de terre |
| 79 | 265153 | 6783535 | Pois |
| 80 | 265426 | 6783535 | Triticale |
| 81 | 265736 | 6783535 | Prairie moyenne durée |
| 85 | 264423 | 6783794 | Blé |
| 86 | 265008 | 6783787 | Maïs grain |
| 88 | 265627 | 6783788 | Blé |
| 89 | 265881 | 6783787 | Blé |
| 93 | 263988 | 6784039 | Blé |
| 94 | 264280 | 6784039 | Haricot |
| 96 | 264862 | 6784039 | Haricot |
| 97 | 265153 | 6784039 | Orge |
| 99 | 265736 | 6784039 | Blé |

^a Les coordonnées suivent la projection Lambert 93

Information sur les sols de MOSAIC (suite)

| Site ID | Longitude ^a | Latitude ^a | Culture présente à l'échantillonnage |
|---------|------------------------|-----------------------|--|
| 100 | 266027 | 6784039 | Orge |
| 101 | 266318 | 6784039 | Maïs grain |
| 102 | 266609 | 6784039 | Maïs grain |
| 103 | 264134 | 6784291 | Maïs grain |
| 104 | 264425 | 6784291 | Blé |
| 105 | 264716 | 6784291 | Colza |
| 106 | 265008 | 6784291 | Blé |
| 107 | 265299 | 6784291 | Maïs ensilage |
| 109 | 266172 | 6784291 | Blé |
| 110 | 266463 | 6784291 | Pomme de terre |
| 111 | 266755 | 6784291 | Maïs grain |
| 112 | 263988 | 6784543 | Prairie moyenne durée |
| 113 | 264280 | 6784543 | Maïs grain |
| 114 | 264571 | 6784543 | Blé |
| 115 | 264862 | 6784543 | Maïs grain |
| 116 | 265153 | 6784543 | Maïs grain |
| 119 | 266027 | 6784543 | Maïs grain |
| 120 | 266341 | 6784582 | Maïs grain |
| 121 | 266609 | 6784543 | Pomme de terre |
| 122 | 263843 | 6784795 | Prairie moyenne durée |
| 123 | 264134 | 6784795 | Maïs grain |
| 124 | 264425 | 6784795 | Maïs grain |
| 125 | 264716 | 6784795 | Prairie moyenne durée |

^a Les coordonnées suivent la projection Lambert 93

Information sur les sols de MOSAIC (suite)

| Site ID | Longitude ^a | Latitude ^a | Culture présente à l'échantillonnage |
|---------|------------------------|-----------------------|--|
| 126 | 265008 | 6784795 | Maïs ensilage |
| 127 | 265299 | 6784795 | Maïs grain |
| 130 | 266282 | 6784763 | Colza |
| 131 | 266463 | 6784795 | Triticale |
| 132 | 266755 | 6784795 | Blé |
| 133 | 264280 | 6785048 | Blé |
| 134 | 264571 | 6785048 | Maïs ensilage |
| 135 | 264862 | 6785048 | Prairie temporaire |
| 137 | 265736 | 6785053 | Maïs ensilage |
| 138 | 266024 | 6785046 | Maïs ensilage |
| 139 | 266425 | 6784977 | Maïs grain |
| 140 | 266609 | 6785048 | Haricot |
| 141 | 264705 | 6785300 | Blé |
| 142 | 265008 | 6785300 | Prairie moyenne durée |
| 143 | 265299 | 6785300 | Prairie moyenne durée |
| 144 | 265597 | 6785334 | Blé |
| 145 | 265883 | 6785278 | Blé |
| 147 | 265153 | 6785552 | Pomme de terre |
| 153 | 266333 | 6783021 | Prairie moyenne durée |
| 154 | 264725 | 6783803 | Maïs grain |
| 155 | 265893 | 6784303 | Maïs grain |
| 156 | 265145 | 6785040 | Prairie moyenne durée |

^a Les coordonnées suivent la projection Lambert 93

Information sur les sols de MOSAIC (suite)

| Site ID | Longitude^a | Latitude^a | Culture présente à l'échantillonnage |
|----------------|------------------------------|-----------------------------|---|
| 160 | 264585 | 6785051 | Maïs ensilage |
| 162 | 264464 | 6785199 | Blé |
| 163 | 265162 | 6785022 | Prairie moyenne durée |
| 167 | 265697 | 6785133 | Maïs ensilage |
| 168 | 265216 | 6784629 | Maïs grain |
| 170 | 265159 | 6785381 | Prairie moyenne durée |
| 171 | 264968 | 6783376 | Pois |
| 172 | 266642 | 6784013 | Blé |
| 174 | 266790 | 6783501 | Prairie moyenne durée |
| 176 | 263954 | 6784442 | Maïs grain |
| 177 | 266008 | 6784530 | Maïs grain |
| 180 | 265107 | 6784679 | Maïs grain |
| 183 | 264332 | 6784587 | Maïs grain |
| 184 | 264772 | 6785146 | Prairie moyenne durée |
| 186 | 264660 | 6784642 | Blé |
| 189 | 265078 | 6784561 | Maïs grain |
| 195 | 265229 | 6784658 | Maïs grain |
| 196 | 265481 | 6785169 | Maïs ensilage |
| 197 | 263998 | 6784439 | Maïs grain |
| 200 | 266172 | 6784601 | Non renseigné |
| 204 | 265761 | 6784306 | Maïs grain |

^a Les coordonnées suivent la projection Lambert 93

Annexe C

Résultats de la sélection GAM avec les phyla en guise de diversité pour les sols de DIMIMOS

TABLE C.1 – Comparison of quality between microbial diversity index-based models and phyla-based models.

| Microbial diversity index-based models ^a | | | | | | | |
|---|------|-------|------|------------------------|------|------|-------------------|
| | nVar | df | BIC | Variance explained (%) | nDiv | %Div | |
| Control soil mineralisation (Rs, control) | 4 | 18.98 | -235 | 93 | 1 | 44 | |
| Amended soil mineralisation (Rs, amended) | 5 | 13.98 | -5 | 90 | 2 | 4 | |
| Residue mineralisation (Rr) | 2 | 8.99 | 386 | 93 | 0 | 0 | |
| Priming effect (PE) | 7 | 19.98 | -462 | 74 | 1 | 9 | |
| Phyla abundance-based models ^b | | | | | | | |
| | nVar | df | BIC | Variance explained (%) | nDiv | %Div | ΔBIC ^c |
| Control soil mineralisation (Rs, control) | 4 | 18.38 | -221 | 93 | 2 | 21 | -14 |
| Amended soil mineralisation (Rs, amended) | 7 | 17.98 | -212 | 94 | 3 | 6 | 207 |
| Residue mineralisation (Rr) | 9 | 26.99 | 436 | 94 | 6 | <1 | -50 |
| Priming effect (PE) | 7 | 19.29 | -470 | 74 | 5 | 24 | 8 |

nVar, Total number of selected covariates ; df, degrees of freedom ; BIC, Bayesian Information Criterion ; nDiv, Number of selected microbial diversity covariates (indexes or phyla abundance) ; %Div, Total percentage of variance explained by all selected microbial diversity covariates

^a Models selected with microbial diversity indexes as potential microbial diversity covariates

^b Models selected with phyla abundance as potential microbial diversity covariates

^c Differences between BICs of phyla abundance-based models and microbial diversity index-based models. A positive value means a model based on phyla was better, while a negative value means a model based on diversity indexes was better.

Annexe D

Relations entre les covariables et les cinétiques de minéralisation pour MOSAIC

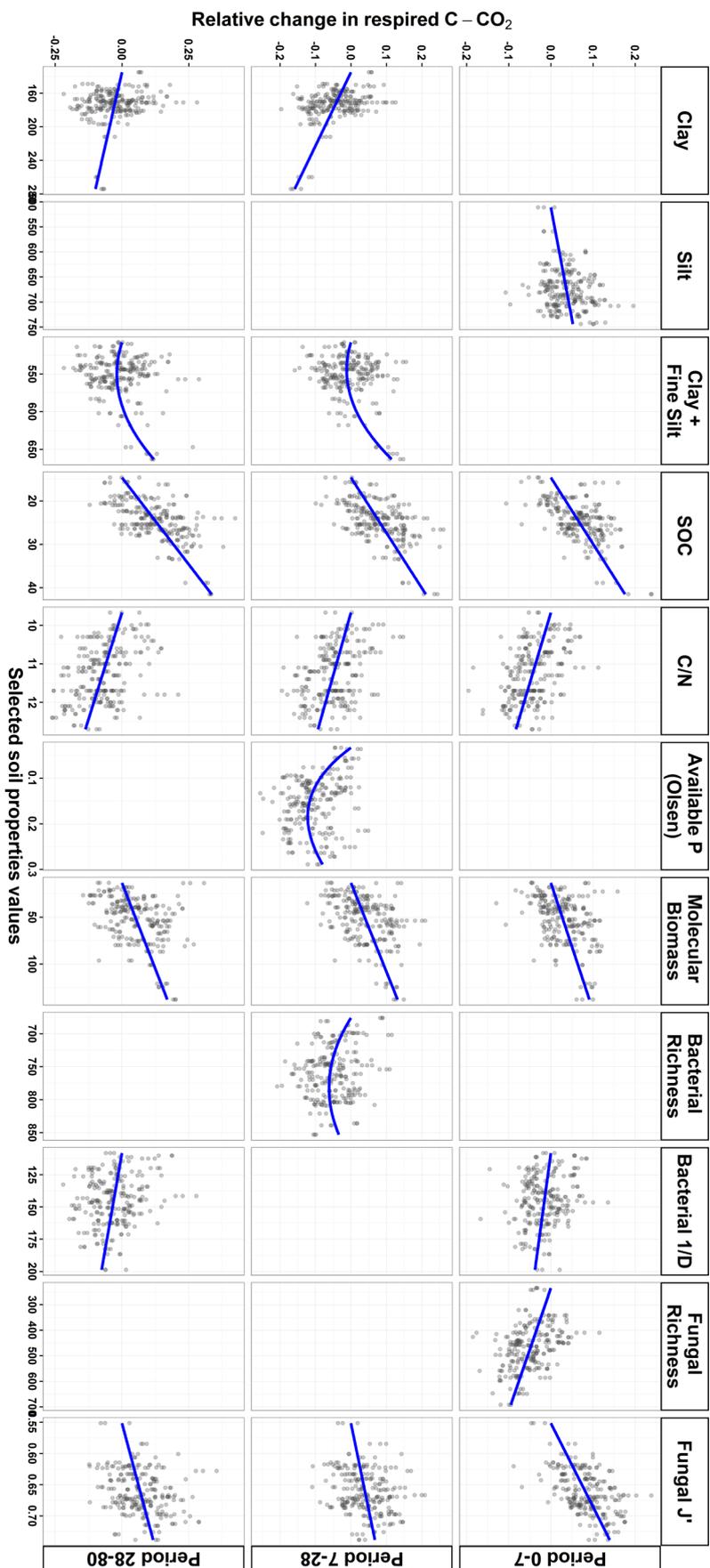


FIGURE D.1 – Propriétés du sol sélectionnées et relations estimées associées pour les flux non marqués des sols non amendés.

L'axe des abscisses représente les valeurs des propriétés des sols et l'axe des ordonnées (sans unité) représente le changement relatif de la cinétique de minéralisation lorsque les valeurs de la propriété du sol évoluent. Les propriétés de sols sont exprimées en $[g \cdot kg^{-1} \text{ sol}]$ pour la texture, la teneur en C organique du sol (SOC) et en phosphore assimilable (Available P), en $[µg \text{ DNA} \cdot g^{-1} \text{ sol}]$ pour la biomasse moléculaire, en [nombre d'unité taxonomique opérationnelle (OTU)] pour les richesses (richness) bactérienne et fongique et sans unité pour les autres propriétés. Pour faciliter la lecture, le changement relatif est égal à 0 lorsque la propriété du sol est égale à son minimum. Les lignes bleues représentent les relations estimées par la méthode GAM. Les points représentent les résidus partiels, i.e. les résidus du modèle lorsque toutes les propriétés sélectionnées sont dans le modèle sauf celle correspondant à la vignette considérée.

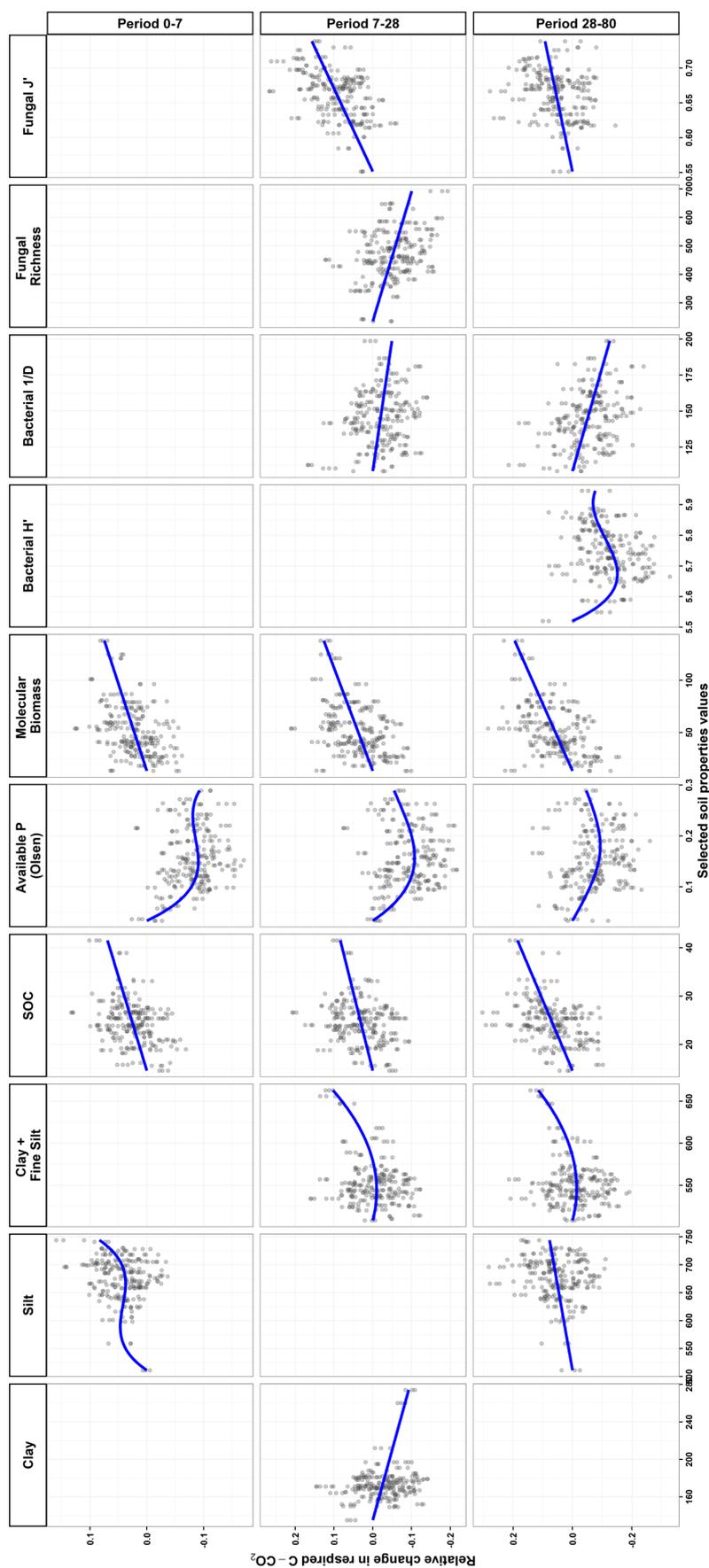


FIGURE D.2 – Propriétés du sol sélectionnées et relations estimées associées pour les flux non marqués des sols amendés.

L'axe des abscisses représente les valeurs des propriétés des sols et l'axe des ordonnées (sans unité) représente le changement relatif de la cinétique de minéralisation lorsque les valeurs de la propriété du sol évoluent. Les propriétés de sols sont exprimées en $[\text{g}\cdot\text{kg}^{-1}]$ sol pour la texture, la teneur en C organique du sol (SOC) et en phosphore assimilable (Available P), en $[\mu\text{g DNA}\cdot\text{g}^{-1}]$ sol pour la biomasse moléculaire, en [nombre d'unité taxonomique opérationnelle (OTU)] pour la richesse (richness) fongique et sans unité pour les autres propriétés. Pour faciliter la lecture, le changement relatif est égal à 0 lorsque la propriété du sol est égale à son minimum. Les lignes bleues représentent les relations estimées par la méthode GAM. Les points représentent les résidus partiels, i.e. les résidus du modèle lorsque toutes les propriétés sélectionnées sont dans le modèle sauf celle correspondant à la vignette considérée.

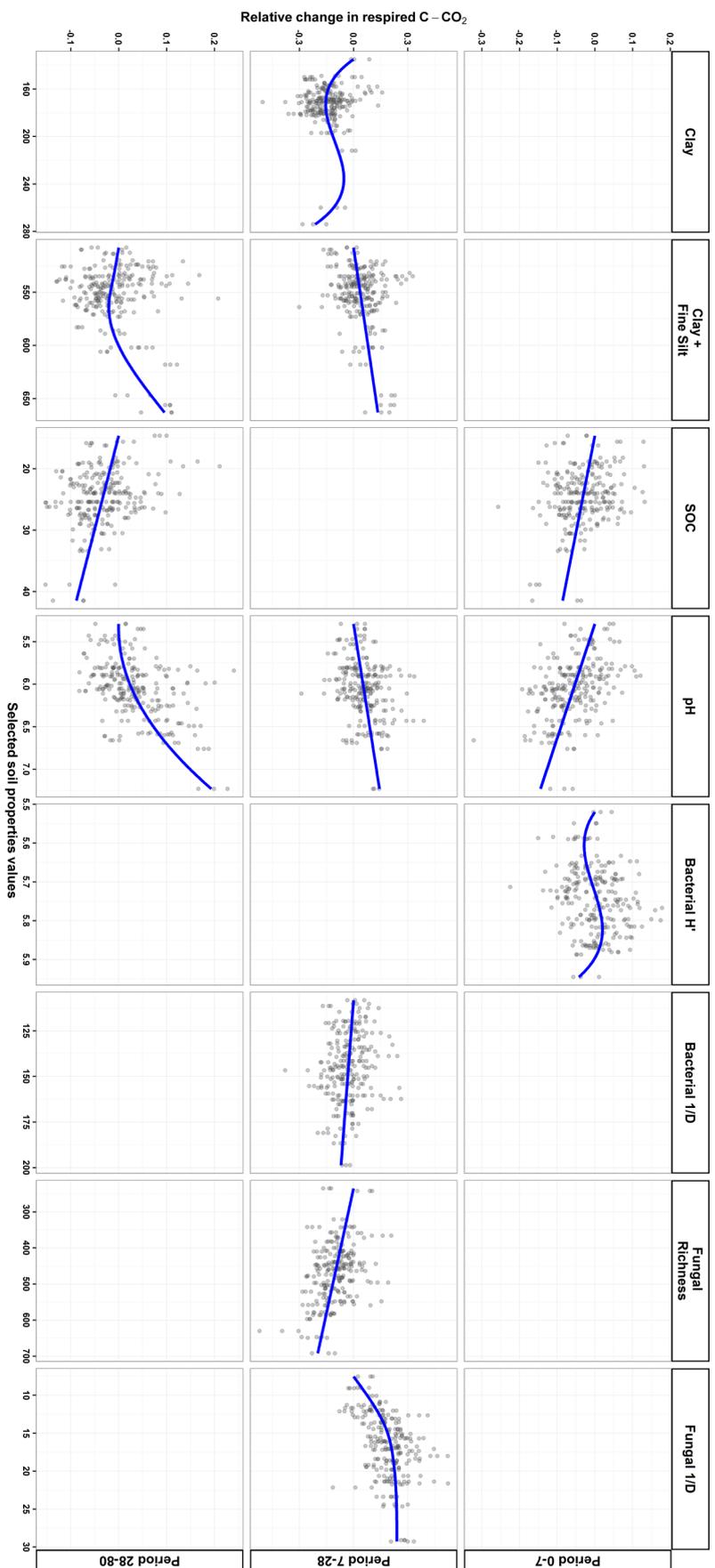


FIGURE D.3 – Propriétés du sol sélectionnées et relations estimées associées pour les flux marqués des sols amendés.

L'axe des abscisses représente les valeurs des propriétés des sols et l'axe des ordonnées (sans unité) représente le changement relatif de la cinétique de minéralisation lorsque les valeurs de la propriété du sol évoluent. Les propriétés de sols sont exprimées en $[\text{g} \cdot \text{kg}^{-1} \text{ sol}]$ pour la texture, la teneur en C organique du sol (SOC), en [nombre d'unité taxonomique opérationnelle (OTU)] pour la richesse (richness) fongique, en [unité pH] pour le pH et sans unité pour les autres propriétés. Pour faciliter la lecture, le changement relatif est égal à 0 lorsque la propriété du sol est égale à son minimum. Les lignes bleues représentent les relations estimées par la méthode GAM. Les points représentent les résidus partiels, i.e. les résidus du modèle lorsque toutes les propriétés sélectionnées sont dans le modèle sauf celle correspondant à la vignette considérée.

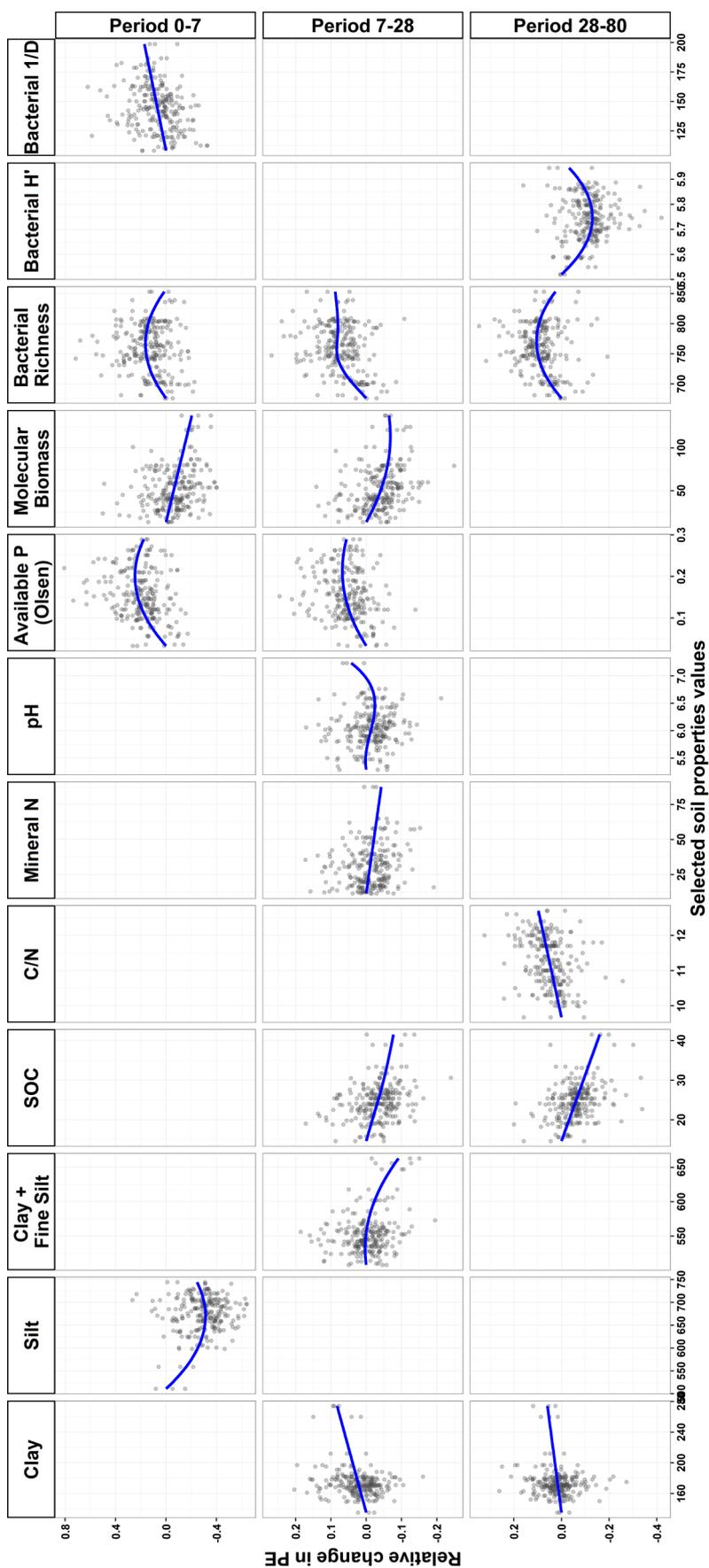


FIGURE D.4 – Propriétés du sol sélectionnées et relations estimées associées pour le priming effect.

L'axe des abscisses représente les valeurs des propriétés des sols et l'axe des ordonnées (sans unité) représente le changement relatif de la cinétique de minéralisation lorsque les valeurs de la propriété du sol évoluent. Les propriétés de sols sont exprimées en $[g \cdot kg^{-1} \text{ sol}]$ pour la texture, la teneur en C organique du sol (SOC), en phosphore assimilable (Available P) et en azote minéral (Mineral N), en $[\mu g \text{ DNA} \cdot g^{-1} \text{ sol}]$ pour la biomasse moléculaire, en [nombre d'unité taxonomique opérationnelle (OTU)] pour la richesse (richness) bactérienne, en [unité pH] pour le pH et sans unité pour les autres propriétés. Pour faciliter la lecture, le changement relatif est égal à 0 lorsque la propriété du sol est égale à son minimum. Les lignes bleues représentent les relations estimées par la méthode GAM. Les points représentent les résidus partiels, i.e. les résidus du modèle lorsque toutes les propriétés sélectionnées sont dans le modèle sauf celle correspondant à la vignette considérée.

Annexe E

Tableaux de références et valeurs
associées pour la construction des
distributions *a priori*

E.1 Paramètres des cinétiques du carbone (ks , Ys , R_T , kr_1 , Yr_1 , kr_2 , Yr_2)

| Références | ks | Ys | R_T | kr_1 | Yr_1 | kr_2 | Yr_2 |
|----------------------------------|-------------------|------|-----------------|--------|--------|--------|--------|
| Blagodatsky <i>et al.</i> , 2010 | 0,019 ; 0,023 | - | 0,0082 ; 0,0005 | - | - | - | - |
| Borgen <i>et al.</i> , 2011 | 0,000085 | 0,3 | - | - | - | - | - |
| Corbeels <i>et al.</i> , 2000 | - | - | - | 0,8 | 0,6 | 0,02 | 0,4 |
| Franko <i>et al.</i> , 2011 | 0,0009 | - | 0,0056 | - | - | - | - |
| Garnier <i>et al.</i> , 2003 | 0,00126 | 0,62 | - | 0,25 | - | 0,1 | - |
| Hadas <i>et al.</i> , 2004 | - | - | - | 1 | - | 0,051 | - |
| Henriksen and Breland, 1999b | - | 0,2 | - | 0,15 | 0,5 | 0,0134 | 0,5 |
| Kirschbaum and Mueller, 2001 | - | - | - | 0,1 | - | 0,041 | - |
| Kuka <i>et al.</i> , 2007 | 0,0012 | 0,65 | - | - | - | - | - |
| Moorhead <i>et al.</i> , 2014 | - | - | - | 0,1 | 0,4 | 0,047 | 0,3 |
| Mueller <i>et al.</i> , 1998a,b | 0,00014 | 0,45 | 0,0018 | - | - | - | - |
| Pansu <i>et al.</i> , 2004 | 0,004 ; 0,002 | - | - | - | - | - | - |
| Paul and Clark, 1989 | 0,0003 | 0,2 | - | 0,2 | 0,6 | 0,08 | 0,4 |
| Petersen <i>et al.</i> , 2005a,b | 0,000146 ; 0,0044 | - | 0,0052 | - | - | - | - |
| Wallman <i>et al.</i> , 2006 | - | - | - | 0,66 | - | 0,03 | - |
| Whitmore <i>et al.</i> , 1997 | 0,25 | - | - | 0,1 | 0,4 | 0,01 | 0,3 |

E.2 Paramètres liés au fractionnement Van Soest

E.2.1 Proportion des fractions Van Soest dans la paille de blé (p_{FRAC})

| Références | n ^a | p_{NDF} | p_{HEM} | p_{CEL} | p_{LIG} |
|---------------------------------------|----------------|-----------|-----------|-----------|-----------|
| <i>Abiven et al., 2007</i> | 1 | 0,17 | 0,33 | 0,44 | 0,06 |
| <i>Abiven and Recous, 2007</i> | 1 | 0,09 | 0,27 | 0,52 | 0,06 |
| <i>Gaillard et al., 2003</i> | 1 | 0,17 | 0,25 | 0,50 | 0,08 |
| <i>Gaillard et al., 2003</i> | 27 | 0,28 | 0,30 | 0,37 | 0,17 |
| <i>Henriksen and Breland, 1999b,c</i> | 1 | 0,13 | 0,41 | 0,42 | 0,05 |
| <i>Henriksen and Breland (1999a)</i> | 1 | 0,12 | 0,36 | 0,45 | 0,07 |
| <i>Motte et al., 2014</i> | 6 | 0,13 | 0,34 | 0,46 | 0,07 |
| <i>Nicolardot et al., 2007</i> | 1 | 0,23 | 0,29 | 0,43 | 0,05 |
| <i>Peltre et al., 2012</i> | 1 | 0,14 | 0,31 | 0,47 | 0,08 |
| <i>Peltre et al., 2012</i> | 3 | 0,12 | 0,31 | 0,50 | 0,08 |
| <i>Veeken et al., 2001</i> | 1 | 0,14 | 0,29 | 0,48 | 0,09 |

NDF, fraction soluble en détergent neutre; HEM, hémicellulose; CEL, cellulose; LIG, lignine

^a Lorsque plusieurs mesures ont été effectuées, seule la moyenne est prise en compte

E.2.2 Proportion de C dans chaque fraction Van Soest (x_{FRAC})

| Références | Résidu | x_{NDF} | x_{HEM} | x_{CEL} | x_{LIG} |
|---------------------------------------|--------|-----------|-----------|-----------|-----------|
| <i>Breland, 1997</i> | WC | - | 0,44 | 0,44 | 0,61 |
| <i>Coppens et al., 2007</i> | RA | 0,36 | 0,37 | 0,49 | 0,61 |
| <i>Coppens et al., 2007</i> | RY | 0,39 | 0,49 | 0,46 | 0,53 |
| <i>Henriksen and Breland, 1999c,b</i> | WS | 0,34 | 0,48 | 0,48 | 0,56 |
| <i>Henriksen and Breland, 1999c</i> | BS | 0,40 | 0,48 | 0,48 | 0,57 |
| <i>Vanlauwe et al., 1994</i> | ML | 0,38 | 0,40 | 0,40 | 0,58 |

NDF, fraction soluble en détergent neutre; HEM, hémicellulose; CEL, cellulose; LIG, lignine

WC, trèfle blanc; RA, colza; RY, seigle; WS, paille de blé; BS, paille d'orge; ML, feuilles de maïs

Bibliographie

- Abiven S, Menasseri S, Angers D, Leterme P (2007). “Dynamics of aggregate stability and biological binding agents during decomposition of organic materials.” *European Journal of Soil Science*, **58**(1), 239–247.
- Abiven S, Recous S (2007). “Mineralisation of crop residues on the soil surface or incorporated in the soil under controlled conditions.” *Biology and Fertility of Soils*, **43**(6), 849–852.
- Acosta-Martinez V, Dowd S, Sun Y, Allen V (2008). “Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use.” *Soil Biology and Biochemistry*, **40**, 2762–2770.
- Ahrens B, Reichstein M, Borken W, Muhr J, Trumbore SE, Wutzler T (2014). “Bayesian calibration of a soil organic carbon model using $\Delta^{14}\text{C}$ measurements of soil organic carbon and heterotrophic respiration as joint constraints.” *Biogeosciences*, **11**(8), 2147–2168.
- Allen B, Willner D, Oechel WC, Lipson D (2010). “Top-down control of microbial activity and biomass in an Arctic soil ecosystem.” *Environmental microbiology*, **12**(3), 642–648.
- Allison SD (2005). “Cheaters, diffusion and nutrients constrain decomposition by microbial enzymes in spatially structured environments.” *Ecology Letters*, **8**(6), 626–635. doi :10.1111/j.1461-0248.2005.00756.x.
- Allison SD (2012). “A trait-based approach for modelling microbial litter decomposition.” *Ecology Letters*, **15**(9), 1058–1070. doi :10.1111/j.1461-0248.2012.01807.x.
- Allison SD (2014). “Modeling adaptation of carbon use efficiency in microbial communities.” *Frontiers in microbiology*, **5**, 571.
- Allison SD, Lu Y, Weihe C, Goulden ML, Martiny AC, Treseder KK, Martiny JB (2013). “Mi-

- crobial abundance and composition influence litter decomposition response to environmental change." *Ecology*, **94**(3), 714–725.
- Allison SD, Wallenstein MD, Bradford MA (2010). "Soil-carbon response to warming dependent on microbial physiology." *Nature Geoscience*, **3**, 336. doi :doi:10.1038/ngeo846.
- Altieri MA, Nicholls CI, Henao A, Lana MA (2015). "Agroecology and the design of climate change-resilient farming systems." *Agronomy for Sustainable Development*, **35**(3), 869–890.
- Andr en O, Balandreau J (1999). "Biodiversity and soil functioning - from black box to can of worms?" *Applied Soil Ecology*, **13**(2), 105–108.
- Arrouays D, Jolivet C, Boulonne L, Bodineau G, Saby N, Grolleau E (2002). "A new projection in France : a multi-institutional soil quality monitoring network." *Comptes Rendus de l'Acad mie d'Agriculture de France (France)*.
- Bakker H (ed.) (1990). *The World Food Crisis : Food Security in Comparative Perspective*. Canadian Scholars Press.
- Banwart SA, Black H, Cai Z, Gicheru PT, Joosten H, Victoria RL, Milne E, Noellemeyer E, Pascual U (2015). *Soil carbon : science, management and policy for multiple benefits*, chapter The Global Challenge for Soil Carbon.
- Baumann K, Dignac MF, Rumpel C, Bardoux G, Sarr A, Steffens M, Maron PA (2013). "Soil microbial diversity affects soil organic matter decomposition in a silty grassland soil." *Biogeochemistry*, **114**(1-3), 201–212.
- Bell T, Newman JA, Silverman BW, Turner SL, Lilley AK (2005). "The contribution of species richness and composition to bacterial services." *Nature*, **436**, 1157–1160. doi :doi:10.1038/nature03891.
- Bellon-Maurel V, Fernandez-Ahumada E, Palagos B, Roger JM, McBratney A (2010). "Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy." *Trends in Analytical Chemistry*, **29**(9), 1073–1081.
- Bernard L, Moug el C, Maron PA, Nowak V, L ev eque J, Henault C, Haichar FeZ, Berge O, Marol C, Balesdent J, *et al.* (2007). "Dynamics and identification of soil microbial populations actively assimilating carbon from ¹³C-labelled wheat residue as estimated by DNA-and RNA-SIP techniques." *Environmental Microbiology*, **9**(3), 752–764.

- Berthrong ST, Buckley DH, Drinkwater LE (2013). “Agricultural Management and Labile Carbon Additions Affect Soil Microbial Community Structure and Interact with Carbon and Nitrogen Cycling.” *Microbial Ecology*, **66**, 158–170.
- Bing-Ru L, Guo-Mei J, Jian C, Gang W (2006). “A review of methods for studying microbial diversity in soils.” *Pedosphere*, **16**(1), 18–24.
- Blagodatskaya E, Blagodatsky S, Anderson TH, Kuzyakov Y (2014). “Microbial growth and carbon use efficiency in the rhizosphere and root-free soil.” *PloS one*, **9**(4).
- Blagodatskaya E, Kuzyakov Y (2008). “Mechanisms of real and apparent priming effects and their dependence on soil microbial biomass and community structure : critical review.” *Biology and Fertility of Soils*, **45**(2), 115–131.
- Blagodatsky S, Blagodatskaya E, Yuyukina T, Kuzyakov Y (2010). “Model of apparent and real priming effects : linking microbial activity with soil organic matter decomposition.” *Soil biology and biochemistry*, **42**(8), 1275–1283.
- Bölscher T, Wadsö L, Börjesson G, Herrmann AM (2016). “Differences in substrate use efficiency : impacts of microbial community composition, land use management, and substrate complexity.” *Biology and Fertility of Soils*, **52**(4), 547–559.
- Bonny S (2011). “L’agriculture écologiquement intensive : nature et défis.” *Cahiers Agricultures*, **20**(6), 451–462.
- Borgen SK, Molstad L, Bruun S, Breland TA, Bakken LR, Bleken MA (2011). “Estimation of plant litter pools and decomposition-related parameters in a mechanistic model.” *Plant and soil*, **338**(1-2), 205–222.
- Bosatta E, Agren GI (1994). “Theoretical analysis of microbial biomass dynamics in soils.” *Soil Biology and Biochemistry*, **26**(1), 143–148. doi :[http://dx.doi.org/10.1016/0038-0717\(94\)90206-2](http://dx.doi.org/10.1016/0038-0717(94)90206-2).
- Bosatta E, Agren GI (1995). “The power and reactive continuum models as particular cases of the q-theory of organic matter dynamics.” *Geochimica et Cosmochimica Acta*, **59**(18), 3833–3835. doi :[http://dx.doi.org/10.1016/0016-7037\(95\)00287-A](http://dx.doi.org/10.1016/0016-7037(95)00287-A).

- Bosatta E, Agren GI (1996). "Theoretical analyses of carbon and nutrient dynamics in soil profiles." *Soil Biology and Biochemistry*, **28**(10), 1523–1531. doi :[http://dx.doi.org/10.1016/S0038-0717\(96\)00167-8](http://dx.doi.org/10.1016/S0038-0717(96)00167-8).
- Breland T (1997). "Modelling mineralization of plant residues on soil : effect of physical protection." *Biology and fertility of soils*, **25**(3), 233–239.
- Buckley DH, Schmidt TM (2003). "Diversity and dynamics of microbial communities in soils from agro-ecosystems." *Environmental Microbiology*, **5**(6), 441–452.
- Burnham K, DR A (2003). *Model Selection and Multimodel Inference A Practical Information-Theoretic Approach Second Edition*. Springer.
- Burns RG, DeForest JL, Marxsen J, Sinsabaugh RL, Stromberger ME, Wallenstein MD, Weintraub MN, Zoppini A (2013). "Soil enzymes in a changing environment : Current knowledge and future directions." *Soil Biology and Biochemistry*, **58**, 216–234. doi :[doi:10.1016/j.soilbio.2012.11.009](https://doi.org/10.1016/j.soilbio.2012.11.009).
- Chambers A, Lal R, Paustian K (2016). "Soil carbon sequestration potential of US croplands and grasslands : Implementing the 4 per Thousand Initiative." *Journal of Soil and Water Conservation*, **71**(3), 68–74.
- Chena R, Senbayram M, Blagodatsky S, Myachina O, Dittert K, Lin X, Blagodatskaya E, Kuzyakov Y (2013). "Soil C and N availability determine the priming effect : microbial N mining and stoichiometric decomposition theories." *Global change biology*.
- Clark JS (2007). *Models for ecological data : an introduction*, volume 11. Princeton university press Princeton, New Jersey, USA.
- Cleveland CC, Reed SC, Keller AB, Nemergut DR, O'Neill SP, Ostertag R, Vitousek PM (2014). "Litter quality versus soil microbial community controls over decomposition : a quantitative analysis." *Oecologia*, **174**(1), 283–294.
- Coleman K, Jenkinson D (1996). "RothC-26.3-A Model for the turnover of carbon in soil." In *Evaluation of soil organic matter models*, pp. 237–246. Springer.
- Colloff M, Wakelin S, Gomez D, Rogers S (2008). "Detection of nitrogen cycle genes in soils for measuring the effects of changes in land use and management." *Soil Biology and Biochemistry*, **40**(7), 1637–1645. doi :<http://dx.doi.org/10.1016/j.soilbio.2008.01.019>.

- Constancias F, Terrat S, Saby N, Horrigue W, Villerd J, Guillemain JP, Biju-Duval L, Nowak V, Dequiedt S, Ranjard L, *et al.* (2015). "Mapping and determinism of soil microbial community distribution across an agricultural landscape." *MicrobiologyOpen*, **4**(3), 505–517.
- Coppens F, Garnier P, Findeling A, Merckx R, Recous S (2007). "Decomposition of mulched versus incorporated crop residues : modelling with PASTIS clarifies interactions between residue quality and location." *Soil Biology and biochemistry*, **39**(9), 2339–2350.
- Corbeels M, Hofman G, Van Cleemput O (2000). "Nitrogen cycling associated with the decomposition of sunflower stalks and wheat straw in a Vertisol." *Plant and soil*, **218**(1-2), 71–82.
- Cotrufo MF, Wallenstein MD, Boot CM, Deneff K, Paul E (2013). "The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization : do labile plant inputs form stable soil organic matter?" *Global Change Biology*, **19**(4), 988–995.
- Cox P, Wilkinson SP, Anderson JM (2001). "Effects of fungal inocula on the decomposition of lignin and structural polysaccharides in *Pinus sylvestris* litter." *Biology and Fertility of Soils*, **33**, 246–251.
- Creamer R, Hannula S, Van Leeuwen J, Stone D, Rutgers M, Schmelz R, de Ruiter P, Hendriksen NB, Bolger T, Bouffaud M, *et al.* (2016). "Ecological network analysis reveals the interconnection between soil biodiversity and ecosystem function as affected by land use across Europe." *Applied Soil Ecology*, **97**, 112–124.
- d'Annunzio R, Zeller B, Nicolas M, Dhôte JF, Saint-André L (2008). "Decomposition of European beech (*Fagus sylvatica*) litter : combining quality theory and ¹⁵N labeling experiments." *Soil Biology and Biochemistry*, **40**(2), 322–333.
- Davidson E, Belk E, Boone RD, *et al.* (1998). "Soil water content and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest." *Global change biology*, **4**(2), 217–227.
- de Boer W, Folman LB, Summerbell RC, Boddy L (2005). "Living in a fungal world : impact of fungi on soil bacterial niche development." *FEMS microbiology reviews*, **29**(4), 795–811.
- de Graaff M, Adkins J, Kardol P, Throop H (2015). "A meta-analysis of soil biodiversity impacts on the carbon cycle." *Soil*, **1**(1), 257.

- de Vries FT, Manning P, Tallwin JRB, Mortimer SR, Pilgrim ES, Harrison KA, Hobbs PJ, Quirk H, Shipley B, Comelissen JHC, Kattge J, Bardgett RD (2012). “Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities.” *Ecology Letters*, **15**, 1230–1239.
- Deacon JW (1985). “Decomposition of filter paper cellulose by thermophilic fungi acting singly, in combination, and in sequence.” *Transactions of the British Mycological Society*, **85**, 663–669.
- Decaens T (2010). “Macroecological patterns in soil communities.” *Global Ecology and Biogeography*, **19**(3), 287–302. doi :10.1111/j.1466-8238.2009.00517.x.
- Degens BP (1998). “Decreases in microbial functional diversity do not result in corresponding changes in decomposition under different moisture conditions.” *Soil Biology and Biochemistry*, **30**, 1989–2000.
- Dequiedt S, Saby N, Lelievre M, Jolivet C, Thioulouse J, Toutain B, Arrouays D, Bispo A, Lemanceau P, Ranjard L (2011). “Biogeographical patterns of soil molecular microbial biomass as influenced by soil characteristics and management.” *Global Ecology and Biogeography*, **20**(4), 641–652.
- Díaz S, Lavorel S, de Bello F, Quétier F, Grigulis K, Robson TM (2007). “Incorporating plant functional diversity effects in ecosystem service assessments.” *Proceedings of the National Academy of Sciences*, **104**(52), 20684–20689.
- Doran J, Elliott E, Paustian K (1998). “Soil microbial activity, nitrogen cycling, and long-term changes in organic carbon pools as related to fallow tillage management.” *Soil & Tillage Research*, **49**, 3–18.
- Doré T, Makowski D, Malézieux E, Munier-Jolain N, Tchamitchian M, Titttonell P (2011). “Facing up to the paradigm of ecological intensification in agronomy : revisiting methods, concepts and knowledge.” *European Journal of Agronomy*, **34**(4), 197–210.
- Fang C, Smith P, Smith J, Moncrieff J (2005). “Incorporating microorganisms as decomposers into models to simulate soil organic matter decomposition.” *Geoderma*, **129**(3), 139–146.
- Feng Y, Motta A, Reeves D, Beurmester C, van Santen E, Osborne J (2003). “Soil microbial communities under conventional-till and no-till continuous cotton systems.” *Soil Biology and Biochemistry*, **35**, 1693–1703.

- Fierer N, Bradford MA, Jackson RB (2007). "Toward an ecological classification of soil bacteria." *Ecology*, **88**, 1354–1364.
- Fierer N, Jackson RB (2006). "The diversity and biogeography of soil bacterial communities." *PNAS*, **103**, 626–631.
- Fierer N, Lauber CL, Ramirez KS, Zaneveld J, Bradford MA, Knight R (2012). "Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients." *The ISME journal*, **6**, 1007–1017.
- Foeroid B, Yearsley JM (2004). "Modelling the impact of microbial grazers on soluble rhizodeposit turnover." *Plant and soil*, **267**(1-2), 329–342.
- Fontaine S, Barot S (2005). "Size and functional diversity of microbe populations control plant persistence and long-term soil carbon accumulation." *Ecology Letters*, **8**(10), 1075–1087.
- Fontaine S, Henault C, Aamor A, Bdioui N, Bloor J, Maire V, Mary B, Revalliot S, Maron P (2011). "Fungi mediate long term sequestration of carbon and nitrogen in soil through their priming effect." *Soil Biology and Biochemistry*, **43**(1), 86–96.
- Fontaine S, Mariotti A, Abbadie L (2003). "The priming effect of organic matter : a question of microbial competition?" *Soil Biology and Biochemistry*, **35**(6), 837–843.
- Francis CA, Beman JM, Kuypers MM (2007). "New processes and players in the nitrogen cycle : the microbial ecology of anaerobic and archaeal ammonia oxidation." *The ISME journal*, **1**(1), 19–27.
- Franko U, Kolbe H, Thiel E, Ließ E (2011). "Multi-site validation of a soil organic matter model for arable fields based on generally available input data." *Geoderma*, **166**(1), 119–134.
- Gaillard V, Chenu C, Recous S (2003). "Carbon mineralisation in soil adjacent to plant residues of contrasting biochemical quality." *Soil biology and biochemistry*, **35**(1), 93–99.
- Garnier P, Neel C, Aita C, Recous S, Lafolie F, Mary B (2003). "Modelling carbon and nitrogen dynamics in a bare soil with and without straw incorporation." *European Journal of Soil Science*, **54**(3), 555–568.

- Garnier P, Néel C, Mary B, Lefol F (2001). "Evaluation of a nitrogen transport and transformation model in a bare soil." *European Journal of Soil Science*, **52**(2), 253–268. doi : 10.1046/j.1365-2389.2001.00374.x.
- Gelman A, Rubin DB (1992). "Inference from iterative simulation using multiple sequences." *Statistical science*, pp. 457–472.
- Gelsomino A, Azzellino A (2011). "Multivariate analysis of soils : microbial biomass, metabolic activity, and bacterial community structure and their relationships with soil depth and type." *Journal of Plant Nutrition and Soil Science*, **174**, 381–394.
- Geyer KM, Kyker-Snowman E, Grandy AS, Frey SD (2016). "Microbial carbon use efficiency : accounting for population, community, and ecosystem-scale controls over the fate of metabolized organic matter." *Biogeochemistry*, **127**(2-3), 173–188.
- Giacomini S, Recous S, Mary B, Aita C (2007). "Simulating the effects of N availability, straw particle size and location in soil on C and N mineralization." *Plant and Soil*, **301**(1-2), 289–301.
- Ginovart M, López D, Gras A (2005). "Individual-based modelling of microbial activity to study mineralization of C and N and nitrification process in soil." *Nonlinear analysis : real world applications*, **6**(4), 773–795.
- Gis Sol (2011). "L'état des sols de France." *Groupement d'intérêt scientifique sur les sols*, p. 188.
- Gömöryová E, Ujházy K, Martinák M, Gömöry D (2013). "Soil microbial community response to variation in vegetation and abiotic environment in a temperate old-growth forest." *Applied Soil Ecology*, **60**, 10–19. doi :10.1016/j.apsoil.2013.03.005.
- Gougoulias C, Clark JM, Shaw LJ (2014). "The role of soil microbes in the global carbon cycle : tracking the below-ground microbial processing of plant-derived carbon for manipulating carbon dynamics in agricultural systems." *Journal of the Science of Food and Agriculture*, **94**(12), 2362–2371.
- Graham EB, Wieder WR, Leff JW, Weintraub SR, Townsend AR, Cleveland CC, Philippot L, Nemergut DR (2014). "Do we need to understand microbial communities to predict ecosystem function? A comparison of statistical models of nitrogen cycling processes." *Soil Biology and Biochemistry*, **68**, 279–282.

- Grant R (2001). "A review of the Canadian ecosystem model ecosys."
- Grant R, Juma N, McGill W (1993). "Simulation of carbon and nitrogen transformations in soil : mineralization." *Soil Biology and Biochemistry*, **25**(10), 1317–1329.
- Griffith MB, Perry SA, Perry WB (1995). "Leaf litter processing and exoenzyme production on leaves in streams of different pH." *Oecologia*, **102**(4), 460–466.
- Griffiths B, Kuan H, Ritz K, Glover L, McCaig A, Fenwick C (2004). "The Relationship between Microbial Community Structure and Functional Stability, Tested Experimentally in an Upland Pasture Soil." *Microbial ecology*, **47**, 104–113.
- Griffiths B, Ritz K, Bardgett RD, Cook R, Christensen S, Ekelund F, Sørensen SJ, Bååth E, Bloem J, De Ruiter P, *et al.* (2000). "Ecosystem response of pasture soil communities to fumigation-induced microbial diversity reductions : an examination of the biodiversity–ecosystem function relationship." *Oikos*, **90**(2), 279–294.
- Griffiths B, Ritz K, Wheatley R, Kuan H, Boag B, Christensen S, Ekelund F, Sørensen S, Muller S, Bloem J (2001). "An examinataion f the biodiversity-ecosystem function relationship in arable soil microbial communaities." *Soil Biology and Biochemistry*, **33**, 1713–1722.
- Griffiths B, Robinson D (1992). "Root-induced nitrogen mineralisation : a nitrogen balance model." *Plant and Soil*, **139**(2), 253–263.
- Grime J (1998). "Benefits of plant diversity to ecosystems : immediate, filter and founder effects." *Journal of Ecology*, **86**(6), 902–910.
- Gubler S, Endrizzi S, Gruber S, Purves R (2013). "Sensitivities and uncertainties of modeled ground temperatures in mountain environments." *Geoscientific Model Development*, **6**(4), 1319–1336.
- Guisan A, Edwards TC, Hastie T (2002). "Generalized linear and generalized additive models in studies of species distributions : setting the scene." *Ecological modelling*, **157**(2), 89–100.
- Hadas A, Kautsky L, Goek M, Kara EE (2004). "Rates of decomposition of plant residues and available nitrogen in soil, related to residue composition through simulation of carbon and nitrogen turnover." *Soil Biology and Biochemistry*, **36**(2), 255–266.

- Hadas A, Molina J, Feigenbaum S, Clapp C (1987). "Simulation of nitrogen-15 immobilization by the model NCSOIL." *Soil Science Society of America Journal*, **51**(1), 102–106.
- Hadas A, Parkin T, Stahl P (1998). "Reduced CO₂ release from decomposing wheat straw under N-limiting conditions : simulation of carbon turnover." *European Journal of Soil Science*, **49**(3), 487–494.
- Haegeman B, Hamelin J, Moriarty J, Neal P, Dushoff J, Weitz JS (2013). "Robust estimation of microbial diversity in theory and in practice." *The ISME journal*, **7**(6), 1092–1101.
- Halfon E (1979). "Preview : theory in ecosystem analysis." *Theoretical systems ecology : advances and case studies*, **4**(6), 1.
- Hansen S, Jensen H, Nielsen N, Svendsen H (1991). "Simulation of nitrogen dynamics and biomass production in winter wheat using the Danish simulation model DAISY." *Fertilizer research*, **27**(2-3), 245–259.
- Hararuk O, Luo Y (2014). "Improvement of global litter turnover rate predictions using a Bayesian MCMC approach." *Ecosphere*, **5**(12), 1–13.
- Hararuk O, Smith MJ, Luo Y (2015). "Microbial models with data-driven parameters predict stronger soil carbon responses to climate change." *Global change biology*, **21**(6), 2439–2453.
- Hastie TJ, Tibshirani RJ (1990). *Generalized additive models*, volume 43. CRC Press.
- Hättenschwiler S, Fromin N, Barantal S (2011). "Functional diversity of terrestrial microbial decomposers and their substrates." *Comptes rendus biologiques*, **334**(5), 393–402.
- Hayatsu M, Tago K, Saito M (2008). "Various players in the nitrogen cycle : diversity and functions of the microorganisms involved in nitrification and denitrification." *Soil Science and Plant Nutrition*, **54**(1), 33–45.
- Heidelberger P, Welch PD (1981). "A spectral method for confidence interval generation and run length control in simulations." *Communications of the ACM*, **24**(4), 233–245.
- Henriksen T, Breland T (1999a). "Decomposition of crop residues in the field : evaluation of a simulation model developed from microcosm studies." *Soil Biology and Biochemistry*, **31**(10), 1423–1434.

- Henriksen T, Breland T (1999b). "Nitrogen availability effects on carbon mineralization, fungal and bacterial growth, and enzyme activities during decomposition of wheat straw in soil." *Soil Biology and Biochemistry*, **31**(8), 1121 – 1134. doi :[http://dx.doi.org/10.1016/S0038-0717\(99\)00030-9](http://dx.doi.org/10.1016/S0038-0717(99)00030-9).
- Henriksen TM, Breland TA (1999c). "Evaluation of criteria for describing crop residue degradability in a model of carbon and nitrogen turnover in soil." *Soil Biology and Biochemistry*, **31**(8), 1135–1149.
- Hill G, Mitkowski N, Aldrich-Wolfe L, Emele L, Jurkonie D, Ficke A, Maldonado-Ramirez S, Lynch S, Nelson E (2000). "Methods for assessing the composition and diversity of soil microbial communities." *Applied soil ecology*, **15**(1), 25–36.
- Hill TC, Walsh KA, Harris JA, Moffett BF (2003). "Using ecological diversity measures with bacterial communities." *FEMS Microbiology Ecology*, **43**(1), 1–11.
- Hooper DU, Chapin F, Ewel J, Hector A, Inchausti P, Lavorel S, Lawton J, Lodge D, Loreau M, Naeem S, *et al.* (2005). "Effects of biodiversity on ecosystem functioning : a consensus of current knowledge." *Ecological monographs*, **75**(1), 3–35.
- Huang Y, Liu D, Wu H (2006). "Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system." *Biometrics*, **62**(2), 413–423.
- Hughes JB, Bohannan BJM (2004). *Section 7 update : Application of ecological diversity statistics in microbial ecology*, pp. 3223–3246. Springer Netherlands, Dordrecht. ISBN 978-1-4020-2177-0. doi :[10.1007/978-1-4020-2177-0_701](https://doi.org/10.1007/978-1-4020-2177-0_701). URL http://dx.doi.org/10.1007/978-1-4020-2177-0_701.
- Hughes JB, Hellmann JJ, Ricketts TH, Bohannan BJ (2002). "Counting the uncountable : Statistical approaches to estimating microbial diversity." *Applied and environmental microbiology*, **68**(1), 448.
- Hunt H, Coleman D, Ingham E, Ingham R, Elliott E, Moore J, Rose S, Reid C, Morley C (1987). "The detrital food web in a shortgrass prairie." *Biology and Fertility of Soils*, **3**(1-2), 57–68.
- Hunt H, Stewart J, Cole C (1983). "A conceptual model for interactions among carbon, nitrogen, sulphur, and phosphorus in grasslands." *The Major Biogeochemical Cycles and Their Interactions*. Wiley, New York, New York, USA, pp. 303–325.

- Hunt H, Trlica M, Redente E, Moore J, Detling J, Kittel T, Walter D, Fowler M, Klein D, Elliott E (1991). "Simulation model for the effects of climate change on temperate grassland ecosystems." *Ecological Modelling*, **53**, 205–246.
- Huston MA (1997). "Hidden treatments in ecological experiments : re-evaluating the ecosystem function of biodiversity." *Oecologia*, **110**, 449–460.
- Ingwersen J, Poll C, Streck T, Kandeler E (2008). "Micro-scale modelling of carbon turnover driven by microbial succession at a biogeochemical interface." *Soil Biology and Biochemistry*, **40**(4), 864–878.
- IPCC (2014). *Climate change 2014 : synthesis Report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change*. IPCC.
- Jangid K, Williams MA, Franzluebbers AJ, Sanderlin JS, Reeves JH, Jenkins MB, Endale DM, Coleman DC, Whitman WB (2008). "Relative impacts of land-use, management intensity and fertilization upon soil microbial community structure in agricultural systems." *Soil Biology and Biochemistry*, **40**(11), 2843–2853.
- Jangid K, Williams MA, Franzluebbers AJ, Schmidt TM, Coleman DC, Whitman WB (2011). "Land-use history has a stronger impact on soil microbial community composition than aboveground vegetation and soil properties." *Soil Biology and Biochemistry*, **43**(10), 2184–2193.
- Jenkinson D, Powlson DS (1976). "The effects of biocidal treatments on metabolism in soil V : a method for measuring soil biomass." *Soil biology and biochemistry*, **8**(3), 209–213.
- Johnston SR, Boddy L, Weightman AJ (2016). "Bacteria in decomposing wood and their interactions with wood-decay fungi." *FEMS Microbiology Ecology*, **92**(11), fiv179.
- Jones PG, Thornton PK (2003). "The potential impacts of climate change on maize production in Africa and Latin America in 2055." *Global environmental change*, **13**(1), 51–59.
- Juarez S, Nunan N, Duday AC, Pouteau V, Chenu C (2013). "Soil carbon mineralisation responses to alterations of microbial diversity and soil structure." *Biology and fertility of soils*, **49**(7), 939–948.
- Kader M, Sleutel S, Begum SA, Moslehuddin AZM, *et al.* (2013). "Nitrogen mineralization in sub-tropical paddy soils in relation to soil mineralogy, management, pH, carbon, nitrogen and iron contents." *European Journal of Soil Science*, **64**(1), 47–57.

- Kaisermann A, Maron P, Beaumelle L, Lata J (2015). “Fungal communities are more sensitive indicators to non-extreme soil moisture variations than bacterial communities.” *Applied Soil Ecology*, **86**, 158–164.
- Kaisermann A, Roguet A, Nunan N, Maron PA, Ostle N, Lata JC (2013). “Agricultural management affects the response of soil bacterial community structure and respiration to water-stress.” *Soil Biology and Biochemistry*, **66**, 69–77. doi :doi:10.1016/j.soilbio.2013.07.001.
- Kätterer T, Reichstein M, Andrén O, Lomander A (1998). “Temperature dependence of organic matter decomposition : a critical review using literature data analyzed with different models.” *Biology and fertility of soils*, **27**(3), 258–262.
- Katul G, Porporato A, Oren R (2007). “Stochastic dynamics of plant-water interactions.” *Annual review of ecology, evolution, and systematics*, pp. 767–791.
- Kemmitt SJ, Wright D, Goulding KW, Jones DL (2006). “pH regulation of carbon and nitrogen dynamics in two agricultural soils.” *Soil Biology and Biochemistry*, **38**(5), 898–911.
- Kennedy A, Smith K (1995). “Soil microbial diversity and the sustainability of agricultural soils.” *Plant and soil*, **170**(1), 75–86.
- Kersebaum K, Richter O (1994). “A model approach to simulate C and N transformations through microbial biomass.” *European Journal of Agronomy*, **3**(4), 355–360.
- King GM (2011). “Enhancing soil carbon storage for carbon remediation : potential contributions and constraints by microbes.” *Trends in microbiology*, **19**(2), 75–84.
- Kirschbaum MU, Mueller R (2001). *Net Ecosystem Exchange : Workshop Proceedings*. Cooperative Research Centre for Greenhouse Accounting.
- Kivlin SN, Hawkes CV (2016). “Temporal and Spatial Variation of Soil Bacteria Richness, Composition, and Function in a Neotropical Rainforest.” *PloS one*, **11**(7), e0159131.
- Korsaeth A, Molstad L, Bakken L (2001). “Modelling the competition for nitrogen between plants and microflora as a function of soil heterogeneity.” *Soil Biology and Biochemistry*, **33**(2), 215–226.
- Kravchenko L, Shvytov I, *et al.* (2004). “Mathematical simulation of the dynamics of interacting populations of rhizosphere microorganisms.” *Microbiology*, **73**(2), 189–195.

- Kuijper LD, Berg MP, Morriën E, Kooi BW, Verhoef HA (2005). "Global change effects on a mechanistic decomposer food web model." *Global Change Biology*, **11**(2), 249–265.
- Kuka K, Franko U, Rühlmann J (2007). "Modelling the impact of pore space distribution on carbon turnover." *ecological modelling*, **208**(2), 295–306.
- Kutsch WL, Bahn M, Heinemeyer A (2009). *Soil carbon dynamics : an integrated methodology*. Cambridge University Press.
- Lal R (1999). "Soil management and restoration for C sequestration to mitigate the accelerated greenhouse effect." *Progress in Environmental Science*, **1**(4), 307–326.
- Lal R (2006). "Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands." *Land Degradation & Development*, **17**(2), 197–209.
- Lal R (2008). "Carbon sequestration." *Philosophical Transactions of the Royal Society of London B : Biological Sciences*, **363**(1492), 815–830. ISSN 0962-8436. doi :10.1098/rstb.2007.2185.
- Langer U, Böhme L, Böhme F (2004). "Classification of soil microorganisms based on growth properties : a critical view of some commonly used terms." *Journal of Plant Nutrition and Soil Science*, **167**, 267–269.
- Lauber CL, Ramirez KS, Aanderud Z, Lennon J, Fierer N (2013). "Temporal variability in soil microbial communities across land-use types." *The ISME journal*, **7**(8), 1641–1650.
- Lazcano C, Gómez-Brandón M, Revilla P, Domínguez J (2013). "Short-term effects of organic and inorganic fertilizers on soil microbial community structure and function." *Biology and Fertility of Soils*, **49**, 723–733.
- Lee ZM, Schmidt TM (2014). "Bacterial growth efficiency varies in soils under different land management practices." *Soil Biology and Biochemistry*, **69**, 282–290.
- Leffelaar P (1986). "Dynamics of partial anaerobiosis, denitrification, and water in a soil aggregate : Experimental." *Soil Science*, **142**(6), 352–366.
- Leffelaar P, Wessel W (1988). "Denitrification in a homogeneous, closed system : experiment and simulation." *Soil Science*, **146**(5), 335–349.
- Leifeld J, Zimmermann M, Fuhrer J (2008). "Simulating decomposition of labile soil organic carbon : Effects of pH." *Soil Biology and Biochemistry*, **40**(12), 2948–2951.

- Lemanceau P, Maron PA, Mazurier S, Mougél C, Pivato B, Plassart P, Ranjard L, Revellin C, Tardy V, Wipf D (2015). “Understanding and managing soil biodiversity : a major challenge in agroecology.” *Agronomy for Sustainable Development*, **35**(1), 67–81.
- Levins R (1966). “The strategy of model building in population biology.” *American scientist*, pp. 421–431.
- Li C, Aber J, Stange F, Butterbach-Bahl K, Papen H (2000). “A process-oriented model of N₂O and NO emissions from forest soils : 1. Model development.” *Journal of Geophysical Research : Atmospheres (1984–2012)*, **105**(D4), 4369–4384.
- Li C, Frolking S, Frolking TA (1992a). “A model of nitrous oxide evolution from soil driven by rainfall events : 1. Model structure and sensitivity.” *Journal of Geophysical Research : Atmospheres (1984–2012)*, **97**(D9), 9759–9776.
- Li C, Frolking S, Frolking TA (1992b). “A model of nitrous oxide evolution from soil driven by rainfall events 2. Model applications.” *Journal of Geophysical Research*, **97**, 9777–9783.
- Lienhard P, Tivet F, Chabanne A, Dequiedt S, Lelièvre M, Sayphoummie S, Leudphanane B, Chemidlin Prévost-Bouré N, Séguy L, Maron PA, Ranjard L (2013). “No-till and cover crops shift soil microbial abundance and diversity in Laos tropical grasslands.” *Agronomy for Sustainable Development*, **33**, 375–384.
- Lipson DA, Monson RK, Schmidt SK, Weintraub MN (2009). “The trade-off between growth rate and yield in microbial communities and the consequences for under-snow soil respiration in a high elevation coniferous forest.” *Biogeochemistry*, **95**(1), 23–35.
- Liu B, Gumpertz ML, Hu S, Ristaino JB (2007). “Long-term effects of organic and synthetic soil fertility amendments on soil microbial communities and the development of southern blight.” *Soil Biology and Biochemistry*, **39**, 2302–2316.
- Liu X, Herbert S, Hashemi A, Zhang X, Ding G, *et al.* (2006). “Effects of agricultural management on soil organic matter and carbon transformation-a review.” *Plant Soil and Environment*, **52**(12), 531.
- Lobell DB, Schlenker W, Costa-Roberts J (2011). “Climate trends and global crop production since 1980.” *Science*, **333**(6042), 616–620.

- Long T, Or D (2005). "Aquatic habitats and diffusion constraints affecting microbial coexistence in unsaturated porous media." *Water resources research*, **41**(8).
- Loreau M (2001). "Microbial diversity, producer–decomposer interactions and ecosystem processes : a theoretical model." *Proceedings of the Royal Society of London. Series B : Biological Sciences*, **268**(1464), 303–309.
- Loreau M, Naeem S, Inchausti P, Bengtsson J, Grime J, Hector A, Hooper D, Huston M, Raffaelli D, Schmid B, *et al.* (2001). "Biodiversity and ecosystem functioning : current knowledge and future challenges." *Science*, **294**(5543), 804–808.
- Lundquist E, Jackson L, Scow K, Hsu C (1999). "Changes in microbial biomass and community composition, and soil carbon and nitrogen pools after incorporation of rye into three California agricultural soils." *Soil Biology and Biochemistry*, **31**(2), 221–236.
- Lunn D, Spiegelhalter D, Thomas A, Best N (2009). "The BUGS project : Evolution, critique and future directions." *Statistics in medicine*, **28**(25), 3049–3067.
- Maggi F, Gu C, Riley W, Hornberger G, Venterea R, Xu T, Spycher N, Steefel C, Miller N, Oldenburg C (2008). "A mechanistic treatment of the dominant soil nitrogen cycling processes : Model development, testing, and application." *Journal of Geophysical Research : Biogeosciences (2005–2012)*, **113**(G2).
- Mansfield ER, Helms BP (1982). "Detecting multicollinearity." *The American Statistician*, **36**(3a), 158–160.
- Manzoni S, Porporato A (2009). "Soil carbon and nitrogen mineralization : Theory and models across scales." *Soil Biology and Biochemistry*, **41**(7), 1355 – 1379. doi :<http://dx.doi.org/10.1016/j.soilbio.2009.02.031>.
- Manzoni S, Taylor P, Richter A, Porporato A, Ågren GI (2012). "Environmental and stoichiometric controls on microbial carbon-use efficiency in soils." *New Phytologist*, **196**(1), 79–91.
- Maron PA, Mougél C, Ranjard L (2011). "Soil microbial diversity : Methodological strategy, spatial overview and functional interest." *Comptes Rendus Biologies*.
- Maron PA, Ranjard L, Mougél C, Lemanceau P (2007). "Metaproteomics : a new approach for studying functional microbial ecology." *Microbial Ecology*, **53**(3), 486–493.

- McGill W, Hunt H, Woodmansee R, Reuss J (1981). "Phoenix, a model of the dynamics of carbon and nitrogen in grassland soils." *Ecological Bulletins (Sweden)*.
- McGuire K, Treseder K (2010). "Microbial communities and their relevance for ecosystem models : Decomposition as a case study." *Soil Biology and Biochemistry*, **42**, 529–535.
- McLauchlan K (2006). "The nature and longevity of agricultural impacts on soil carbon and nutrients : a review." *Ecosystems*, **9**(8), 1364–1382.
- Meersmans J, Martin MP, De Ridder F, Laccarce E, Wetterlind J, De Baets S, Le Bas C, Louis BP, Orton TG, Bispo A, *et al.* (2012). "A novel soil organic C model using climate, soil type and management data at the national scale in France." *Agronomy for sustainable development*, **32**(4), 873–888.
- Miki T, Ushio M, Fukui S, Kondoh M (2010). "Functional diversity of microbial decomposers facilitates plant coexistence in a plant–microbe–soil feedback model." *Proceedings of the National Academy of Sciences*, **107**(32), 14251–14256.
- Minasny B, McBratney AB (2006). "A conditioned Latin hypercube method for sampling in the presence of ancillary information." *Computers & Geosciences*, **32**(9), 1378–1388.
- Miura T, Niswati A, Swibawa I, Haryani S, Gunito H, Shimano S, Fujie K, Kaneko N (2015). "Diversity of fungi on decomposing leaf litter in a sugarcane plantation and their response to tillage practice and bagasse mulching : implications for management effects on litter decomposition." *Microbial ecology*, **70**(3), 646–658.
- Molina J, Clapp C, Shaffer M, Chichester F, Larson W (1983). "NCSOIL, a model of nitrogen and carbon transformations in soil : description, calibration, and behavior." *Soil Science Society of America Journal*, **47**(1), 85–91.
- Moore JC, Berlow EL, Coleman DC, Ruitter PC, Dong Q, Hastings A, Johnson NC, McCann KS, Melville K, Morin PJ, *et al.* (2004). "Detritus, trophic dynamics and biodiversity." *Ecology letters*, **7**(7), 584–600.
- Moore JC, McCann K, de Ruitter PC (2005). "Modeling trophic pathways, nutrient cycling, and dynamic stability in soils." *Pedobiologia*, **49**(6), 499–510.
- Moorhead D, Lashermes G, Recous S, Bertrand I (2014). "Interacting microbe and litter quality controls on litter decomposition : a modeling analysis." *PloS one*, **9**(9), e108769.

- Moorhead D, Sinsabaugh R (2006). "A theoretical model of litter decay and microbial interactions." *Ecological Monographs*, **76**, 151–174.
- Mooshammer M, Wanek W, Zechmeister-Boltenstern S, Richter A (2014). "Stoichiometric imbalances between terrestrial decomposer communities and their resources : mechanisms and implications of microbial adaptations to their resources." *The Microbial Regulation of Global Biogeochemical Cycles*, p. 191.
- Morita S, Thall PF, Müller P (2010). "Evaluating the impact of prior assumptions in Bayesian biostatistics." *Statistics in biosciences*, **2**(1), 1–17.
- Motte JC, Escudié R, Beauvils N, Steyer JP, Bernet N, Delgenès JP, Dumas C (2014). "Morphological structures of wheat straw strongly impacts its anaerobic digestion." *Industrial Crops and Products*, **52**, 695–701.
- Mueller T, Jensen LS, Nielsen N, Magid J (1998a). "Turnover of carbon and nitrogen in a sandy loam soil following incorporation of chopped maize plants, barley straw and blue grass in the field." *Soil Biology and Biochemistry*, **30**(5), 561–571.
- Mueller T, Magid J, Jensen LS, Svendsen H, Nielsen N (1998b). "Soil C and N turnover after incorporation of chopped maize, barley straw and blue grass in the field : Evaluation of the DAISY soil–organic-matter submodel." *Ecological Modelling*, **111**(1), 1–15.
- Nannipieri P, Ascher J, Ceccherini M, Landi L, Pietramellara G, Renella G (2003). "Microbial diversity and soil functions." *European Journal of Soil Science*, **54**(4), 655–670.
- Nash JE, Sutcliffe JV (1970). "River flow forecasting through conceptual models part I A discussion of principles." *Journal of hydrology*, **10**(3), 282–290.
- Neill C, Gignoux J (2006). "Soil organic matter decomposition driven by microbial growth : a simple model for a complex network of interactions." *Soil Biology and Biochemistry*, **38**(4), 803–811.
- Nicolardot B, Bouziri L, Bastian F, Ranjard L (2007). "A microcosm experiment to evaluate the influence of location and quality of plant residues on residue decomposition and genetic structure of soil microbial communities." *Soil Biology and Biochemistry*, **39**(7), 1631–1644.

- Nicolardot B, Recous S, Mary B (2001). "Simulation of C and N mineralisation during crop residue decomposition : a simple dynamic model based on the C : N ratio of the residues." *Plant and Soil*, **228**(1), 83–103.
- Nielsen U, Ayres E, Wall D, Bardgett R (2011). "Soil biodiversity and carbon cycling : a review and synthesis of studies examining diversity–function relationships." *European Journal of Soil Science*, **62**(1), 105–116.
- Niemenmaa O, Uusi-Rauva A, Hatakka A (2008). "Demethoxylation of [O14CH3]-labelled lignin model compounds by the brown-rot fungi *Gloeophyllum trabeum* and *Poria* (*Postia*) *placenta*." *Biodegradation*, **19**(4), 555–565.
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2016). *vegan : Community Ecology Package*. R package version 2.4-0.
- Pansu M, Bottner P, Sarmiento L, Metselaar K (2004). "Comparison of five soil organic matter decomposition models using data from a 14C and 15N labeling field experiment." *Global Biogeochemical Cycles*, **18**(4).
- Parent E, Rivot E (2012). *Introduction to hierarchical Bayesian modeling for ecological data*. CRC Press.
- Parton W (1996). "The CENTURY model." In *Evaluation of soil organic matter models*, pp. 283–291. Springer.
- Parton WJ, Stewart JW, Cole CV (1988). "Dynamics of C, N, P and S in grassland soils : a model." *Biogeochemistry*, **5**(1), 109–131.
- Pascual N, Cécillon L, Mathieu O, Hénault C, Sarr A, Lévêque J, Farcy P, Ranjard L, Maron PA (2010). "In situ dynamics of microbial communities during decomposition of wheat, rape, and alfalfa residues." *Microbial ecology*, **60**(4), 816–828.
- Pascual N, Ranjard L, Kaisermann A, Bachar D, Christen R, Terrat S, Mathieu O, Lévêque J, Mougel C, Henault C, *et al.* (2013). "Stimulation of Different Functional Groups of Bacteria by Various Plant Residues as a Driver of Soil Priming Effect." *Ecosystems*, **16**(5), 810–822.
- Patten BC (1972). "A simulation of the shortgrass prairie ecosystem." *Simulation*, **19**(6), 177–186.

- Paul E, Clark F (1989). *Soil Microbiology and Biochemistry*. Academic Press Limited.
- Peltre C, Christensen BT, Dragon S, Icard C, Kätterer T, Houot S (2012). “RothC simulation of carbon accumulation in soil after repeated application of widely different organic amendments.” *Soil Biology and Biochemistry*, **52**, 49–60.
- Perveen N, Barot S, Alvarez G, Klumpp K, Martin R, Rapaport A, Herfurth D, Louault F, Fontaine S (2014). “Priming effect and microbial diversity in ecosystem functioning and response to global change : a modeling approach using the SYMPHONY model.” *Global change biology*, **20**, 1174–1190.
- Petersen BM, Berntsen J, Hansen S, Jensen LS (2005a). “CN-SIM : a model for the turnover of soil organic matter. I. Long-term carbon and radiocarbon development.” *Soil Biology and Biochemistry*, **37**(2), 359–374.
- Petersen BM, Jensen LS, Hansen S, Pedersen A, Henriksen TM, Sørensen P, Trinsoutrot-Gattin I, Berntsen J (2005b). “CN-SIM : a model for the turnover of soil organic matter. II. Short-term carbon and nitrogen development.” *Soil Biology and Biochemistry*, **37**(2), 375–393.
- Philippot L, Andersson S, Battin T, Prosser J, Schimel J, Whitman W, Maron PA (2010). “The ecological coherence of high bacterial taxonomic ranks.” *Nature Reviews Microbiology*, **8**, 523–529.
- Philippot L, Spor A, Hénault C, Bru D, Bizouard F, Jones CM, Sarr A, Maron PA (2013). “Loss in microbial diversity affects nitrogen cycling in soil.” *The ISME journal*, **7**(8), 1609–1619.
- Plassart P, Terrat S, Thomson B, Griffiths R, Dequiedt S, Lelievre M, Regnier T, Nowak V, Bailey M, Lemanceau P, Bispo A, Chabbi A, Maron PA, Mougél C, Ranjard L (2012). “Evaluation of the ISO Standard 11063 DNA Extraction Procedure for Assessing Soil Microbial Abundance and Community Structure.” *PLoS ONE*, **7**(9), e44279. doi :10.1371/journal.pone.0044279.
- R Core Team (2015). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rampelotto PH, ao de Siquiera FA, Barboza ADM, Roesch LFW (2013). “Changes in Diversity, Abundance, and Structure of Soil Bacterial Communities in Brazilian Savanna Under Different Land Use Systems.” *Microbial Ecology*, **66**, 593–607.

- Ranjard L, Poly F, Nazaret S (2000). "Monitoring complex bacterial communities using culture-independent molecular techniques : application to soil environment." *Research in Microbiology*, **151**(3), 167–177.
- Ratié C, Richer de Forges A, Berché P, Boulonne L, Toutain B, Saby N, Chenu JP, Laloua D, Ortolland B, Tientcheu Nguenkam ME, *et al.* (2010). "Le Conservatoire des Sols : la mémoire des sols de France."
- Raynaud X, Lata JC, Leadley PW (2006). "Soil microbial loop and nutrient uptake by plants : a test using a coupled C : N model of plant–microbial interactions." *Plant and Soil*, **287**(1-2), 95–116.
- Reed HE, Martiny JB (2007). "Testing the functional significance of microbial composition in natural communities." *FEMS microbiology ecology*, **62**(2), 161–170.
- Reeves D (1997). "The role of soil organic matter in maintaining soil quality in continuous cropping systems." *Soil and Tillage Research*, **43**(1), 131–167.
- Regan KM, Nunan N, Boeddinghaus RS, Baumgartner V, Berner D, Boch S, Oelmann Y, Overmann J, Prati D, Schloter M, Schmitt B, Sorkau E, Steffens M, Kandeler E, Marhan S (2014). "Seasonal controls on grassland microbial biogeography : Are they governed by plants, abiotic properties or both ?" *Soil Biology and Biochemistry*, **71**, 21–30.
- Reichert P, Omlin M (1997). "On the usefulness of overparameterized ecological models." *Ecological Modelling*, **95**(2), 289–299.
- Reidsma P, Tekelenburg T, Van den Berg M, Alkemade R (2006). "Impacts of land-use change on biodiversity : An assessment of agricultural biodiversity in the European Union." *Agriculture, Ecosystems & Environment*, **114**(1), 86–102.
- Robinson CH, Dighton J, Frankland JC, PA C (1993). "Nutrient and carbon dioxide release by interacting species of straw-decomposing fungi." *Plant and soil*, **151**, 139–142.
- Robinson D, Griffiths B, Ritz K, Wheatley R (1989). "Root-induced nitrogen mineralisation : a theoretical analysis." *Plant and Soil*, **117**(2), 185–193.
- Robson A (2012). *Soil acidity and plant growth*. Elsevier.

- Romaní AM, Fischer H, Mille-Lindblom C, Tranvik LJ (2006). “Interactions of bacteria and fungi on decomposing litter : differential extracellular enzyme activities.” *Ecology*, **87**(10), 2559–2569.
- Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, Knight R, Fierer N (2010). “Soil bacterial and fungal communities across a pH gradient in an arable soil.” *The ISME journal*, **4**(10), 1340–1351.
- Rousk J, Brookes PC, Glanville HC, Jones DL (2011). “Lack of correlation between turnover of low-molecular-weight dissolved organic carbon and differences in microbial community composition or growth across a soil pH gradient.” *Applied and environmental microbiology*, **77**(8), 2791–2795.
- Roy M, Mandal S, Ray S (2008). “Detrital ontogenic model including decomposer diversity.” *ecological modelling*, **215**(1), 200–206.
- Ruiter PCD, Moore JC, Zwart KB, Bouwman LA, Hassink J, Bloem J, Vos JAD, Marinissen JCY, Didden WAM, Lebrink G, Brussaard L (1993). “Simulation of Nitrogen Mineralization in the Below-Ground Food Webs of Two Winter Wheat Fields.” *Journal of Applied Ecology*, **30**(1), pp. 95–106.
- Rusco E, Jones R, Bidoglio G (2001). “Organic Matter in the Soils of Europe : Present Status and Future Trends.” *EUR 20556 EN, Office for Official Publications of the European Communities, Luxembourg*.
- Sagova-Mareckova M, Omelka M, Cermak L, Kamenik Z, Olsovska J, Hacki E, Kopecky J, Hadacek F (2011). “Microbial communities show parallels at sites with distinct litter and soil characteristics.” *Applied and Environmental Microbiology*, **77**, 7560–7567.
- Schädel C, Luo Y, Evans RD, Fei S, Schaeffer SM (2013). “Separating soil CO₂ efflux into C-pool-specific decay rates via inverse analysis of soil incubation data.” *Oecologia*, **171**(3), 721–732.
- Scharlemann JP, Tanner EV, Hiederer R, Kapos V (2014). “Global soil carbon : understanding and managing the largest terrestrial carbon pool.” *Carbon Management*, **5**(1), 81–91.
- Schimel J (2001). “1.13 - Biogeochemical Models : Implicit versus Explicit Microbiology.” In ED Schulze, M Heimann, S Harrison, E Holland, J Lloyd, IC Prentice, D Schimel (eds.),

- Global Biogeochemical Cycles in the Climate System*, pp. 177 – 183. Academic Press, San Diego. doi :<http://dx.doi.org/10.1016/B978-012631260-7/50015-7>.
- Schimel JP, Schaeffer SM (2015). “Microbial control over carbon cycling in soil.” *The causes and consequences of microbial community structure*, p. 155.
- Setälä H, McLean MA (2004). “Decomposition rate of organic substrates in relation to the species diversity of soil saprophytic fungi.” *Oecologia*, **139**, 98–107.
- Sierra CA, Malghani S, Müller M (2015). “Model structure and parameter identification of soil organic matter models.” *Soil Biology and Biochemistry*, **90**, 197–203.
- Sinsabaugh RL, Manzoni S, Moorhead DL, Richter A (2013). “Carbon use efficiency of microbial communities : stoichiometry, methodology and modelling.” *Ecology letters*, **16**(7), 930–939.
- Six J, Frey S, Thiet R, Batten K (2006). “Bacterial and fungal contributions to carbon sequestration in agroecosystems.” *Soil Science Society of America Journal*, **70**(2), 555–569.
- Smit E, Leeflang P, Gommans S, van den Broek J, van Mil S, Wernars K (2001). “Diversity and seasonal fluctuations of the dominant members of the bacterial soil community in a wheat field as determined by cultivation and molecular methods.” *Applied and Environmental Microbiology*, **67**(5), 2284–2291.
- Smith P, Smith J, Powlson D, McGill W, Arah J, Chertov O, Coleman K, Franko U, Frolking S, Jenkinson D, *et al.* (1997). “A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments.” *Geoderma*, **81**(1), 153–225.
- Soetaert K, Petzoldt T, *et al.* (2010). “Inverse modelling, sensitivity and monte carlo analysis in R using package FME.” *Journal of Statistical Software*, **33**(3), 1–28.
- Song C, Liu D, Song Y, Yang G, Wan Z, Li Y, Xu X (2011). “Effect of exogenous phosphorus addition on soil respiration in *Calamagrostis angustifolia* freshwater marshes of Northeast China.” *Atmospheric environment*, **45**(7), 1402–1406.
- Stange F, Butterbach-Bahl K, Papen H, Zechmeister-Boltenstern S, Li C, Aber J (2000). “A process-oriented model of N₂O and NO emissions from forest soils : 2. Sensitivity analysis and validation.” *Journal of Geophysical Research : Atmospheres (1984–2012)*, **105**(D4), 4385–4398.

- Stockmann U, Adams MA, Crawford JW, Field DJ, Henakaarchchi N, Jenkins M, Minasny B, McBratney AB, de Courcelles VdR, Singh K, *et al.* (2013). “The knowns, known unknowns and unknowns of sequestration of soil organic carbon.” *Agriculture, Ecosystems & Environment*, **164**, 80–99.
- Sturtz S, Ligges U, Gelman A (2010). “R2OpenBUGS : a package for running OpenBUGS from R.” URL <http://cran.rproject.org/web/packages/R2OpenBUGS/vignettes/R2OpenBUGS.pdf>.
- Sul WJ, Asuming-Brempong S, Wang Q, Turlousse DM, Penton CR, Deng Y, Rodrigues JL, Adiku SG, Jones JW, Zhou J, Cole JR, Tiedje JM (2013). “Tropical agricultural land management influences on soil microbial communities through its effect on soil organic carbon.” *Soil Biology and Biochemistry*, **65**, 33–38. doi :doi:10.1016/j.soilbio.2013.05.007.
- Sylvia D, Fuhrmann J, Hertel P, Zuberer D (2005). *Principles and Applications of Soil Microbiology, 2nd edition*. Pearson.
- Tardy V, Mathieu O, Lévêque J, Terrat S, Chabbi A, Lemanceau P, Ranjard L, Maron PA (2014). “Stability of soil microbial structure and activity depends on microbial diversity.” *Environmental Microbiology Reports*, **6**, 173–183.
- Tardy V, Spor A, Mathieu O, Lévêque J, Terrat S, Plassart P, Regnier T, Bardgett RD, Van Der Putten WH, Roggero PP, *et al.* (2015). “Shifts in microbial diversity through land use intensity as drivers of carbon mineralization in soil.” *Soil Biology and Biochemistry*, **90**, 204–213.
- Thuriès L, Pansu M, Larré-Larrouy MC, Feller C (2002). “Biochemical composition and mineralization kinetics of organic inputs in a sandy soil.” *Soil Biology and Biochemistry*, **34**(2), 239–250.
- Tilman D, Cassman KG, Matson PA, Naylor R, Polasky S (2002). “Agricultural sustainability and intensive production practices.” *Nature*, **418**(6898), 671–677.
- Tiunov AV, Scheu S (2005). “Facilitative interactions rather than resource partitioning drive diversity-functioning relationships in laboratory fungal communities.” *Ecology letters*, **8**, 618–625.

- Torsvik V, Øvreås L (2002). “Microbial diversity and function in soil : from genes to ecosystems.” *Current opinion in microbiology*, **5**(3), 240–245.
- Treseder KK, Balsler TC, Bradford MA, Brodie EL, Dubinsky EA, Eviner VT, Hofmockel KS, Lennon JT, Levine UY, MacGregor BJ, *et al.* (2012). “Integrating microbial ecology into ecosystem models : challenges and priorities.” *Biogeochemistry*, **109**(1-3), 7–18.
- Tukey JW (1949). “Comparing individual means in the analysis of variance.” *Biometrics*, pp. 99–114.
- United Nations, Department of Economic and Social Affairs, Population Division (2015). “World Population Prospects : The 2015 Revision, World Population 2015 Wallchart.” ST/ESA/SER.A/378., URL <https://esa.un.org/unpd/wpp/Publications/>.
- van der Heijden MG, Wagg C (2013). “Soil microbial diversity and agro-ecosystem functioning.” *Plant and soil*, **363**(1-2), 1–5.
- van der Wal A, Geydan TD, Kuyper TW, de Boer W (2013). “A thready affair : linking fungal diversity and community dynamics to terrestrial decomposition processes.” *FEMS Microbiology Reviews*, **37**(4), 477–494.
- Van Horn DJ, Van Horn ML, Barrett JE, Gooseff MN, Altrichter AE, Geyer KM, Zeglin LH, Takacs-Vesbach CD (2013). “Factors Controlling Soil Microbial Biomass and Bacterial Diversity and Community Composition in a Cold Desert Ecosystem : Role of Geographic Scale.” *Plos One*, **8**.
- Van Soest P, RH W (1967). “Use of detergents in the analysis of fibrous feeds. VI. Determination of plant cell constituents.” *Journal of Official Chemists*, **50**, 50–55.
- Vanlauwe B, Dendooven L, Merckx R (1994). “Residue fractionation and decomposition : the significance of the active fraction.” *Plant and Soil*, **158**(2), 263–274.
- Veeken A, Adani F, Nierop K, De Jager P, Hamelers H (2001). “Degradation of biomacromolecules during high-rate composting of wheat straw–amended feces.” *Journal of environmental quality*, **30**(5), 1675–1684.
- Waldrop MP, Firestone MK (2004). “Microbial community utilization of recalcitrant and simple carbon compounds : impact of oak-woodland plant communities.” *Oecologia*, **138**(2), 275–284.

- Wallman P, Belyazid S, Svensson MG, Sverdrup H (2006). “DECOMP—a semi-mechanistic model of litter decomposition.” *Environmental Modelling & Software*, **21**(1), 33–44.
- Walse C, Berg B, Sverdrup H (1998). “Review and synthesis of experimental data on organic matter decomposition with respect to the effect of temperature, moisture, and acidity.” *Environmental Reviews*, **6**(1), 25–40.
- Wang G, Post WM (2012). “A theoretical reassessment of microbial maintenance and implications for microbial ecology modeling.” *FEMS microbiology ecology*, **81**(3), 610–617.
- Wang G, Post WM (2013). “A note on the reverse Michaelis–Menten kinetics.” *Soil Biology and Biochemistry*, **57**, 946–949.
- Wang G, Post WM, Mayes MA (2013). “Development of microbial-enzyme-mediated decomposition model parameters through steady-state and dynamic analyses.” *Ecological Applications*, **23**(1), 255–272.
- Waring BG, Averill C, Hawkes CV (2013). “Differences in fungal and bacterial physiology alter soil carbon and nitrogen cycling : insights from meta-analysis and theoretical models.” *Ecology letters*, **16**(7), 887–894.
- Wertz S, Degrange V, Prosser JI, Poly F, Commeaux C, Freitag T, Guillaumaud N, Roux XL (2006). “Maintenance of soil functioning following erosion of microbial diversity.” *Environmental Microbiology*, **8**(12), 2162–2169.
- Wertz S, Degrange V, Prosser JI, Poly F, Commeaux C, Guillaumaud N, Le Roux X (2007). “Decline of soil microbial diversity does not influence the resistance and resilience of key soil microbial functional groups following a model disturbance.” *Environmental Microbiology*, **9**, 2211–2219.
- Wezel A, Bellon S, Dore T, Francis C, Vallod D, David C (2009). “Agroecology as a science, a movement and a practice. A review.” *Agronomy for Sustainable Development*, **29**(4), 503–515. doi :{10.1051/agro/2009004}.
- Wezel A, Casagrande M, Celette F, Vian JF, Ferrer A, Peigné J (2014). “Agroecological practices for sustainable agriculture. A review.” *Agronomy for sustainable development*, **34**(1), 1–20.

- Whitmore A, Klein-Gunnewiek H, Crocker G, Klir J, Körschens M, Poulton P (1997). “Simulating trends in soil organic carbon in long-term experiments using the Verbeme/MOTOR model.” *Geoderma*, **81**(1), 137–151.
- Wieder W, Grandy A, Kallenbach C, Bonan G (2014). “Integrating microbial physiology and physio-chemical principles in soils with the MIncrobial-MIneral Carbon Stabilization (MIMICS) model.” *Biogeosciences*, **11**(14), 3899–3917.
- Wieder WR, Bonan GB, Allison SD (2013). “Global soil carbon projections are improved by modelling microbial processes.” *Nature Climate Change*, **3**(10), 909–912.
- Wohl DL, Arora S, Gladstone JR (2004). “Functional Redundancy Supports Biodiversity and Ecosystem Function in a Closed and Constant Environment.” *Ecology*, **85**, 1534–1540.
- Wolf A, Blyth E, Harding R, Jacob D, Keup-Thiel E, Goettel H, Callaghan T (2008). “Sensitivity of an ecosystem model to hydrology and temperature.” *Climatic Change*, **87**(1-2), 75–89.
- Wood S (2006). *Generalized additive models : an introduction with R*. CRC press.
- Wood S (2011). “Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models.” *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, **73**(1), 3–36.
- Wutzler T, Reichstein M (2008). “Colimitation of decomposition by substrate and decomposers ? a comparison of model formulations.” *Biogeosciences*, **5**(3), 749–759.
- Yang Y, Schaefer DA, Liu W, Popescu VD, Yang C, Wang X, Wu C, Douglas WY (2016). “Higher fungal diversity in dead wood is correlated with lower CO₂ emissions in a natural forest.” *Scientific Reports*.
- Zak JC, Willig MR, Moorhead DL, Wildman HG (1994). “Functional diversity of microbial communities : a quantitative approach.” *Soil Biology and Biochemistry*, **26**(9), 1101–1108.
- Zehr JP, Jenkins BD, Short SM, Steward GF (2003). “Nitrogenase gene diversity and microbial community structure : a cross-system comparison.” *Environmental microbiology*, **5**(7), 539–554.
- Zelenev V, Van Bruggen A, Leffelaar P, Bloem J, Semenov A (2006). “Oscillating dynamics of bacterial populations and their predators in response to fresh organic matter added to soil : the simulation model BACWAVE-WEB.” *Soil Biology and Biochemistry*, **38**(7), 1690–1711.

- Zelles L (1999). "Fatty acid patterns of phospholipids and lipopolysaccharides in the characterisation of microbial communities in soil : a review." *Biology and fertility of soils*, **29**(2), 111–129.
- Zhang FG, Zhang QG (2016). "Microbial diversity limits soil heterotrophic respiration and mitigates the respiration response to moisture increase." *Soil Biology and Biochemistry*, **98**, 180–185.
- Zheng DW, Ågren GI, Bengtsson J (1999). "How do soil organisms affect total organic nitrogen storage and substrate nitrogen to carbon ratio in soils? A theoretical analysis." *Oikos*, pp. 430–442.
- Zheng DW, Bengtsson J, Agren GI (1997). "Soil food webs and ecosystem processes : decomposition in donor-control and Lotka-Volterra systems." *American Naturalist*, pp. 125–148.
- Zhong W, Gu T, Wang W Zhang B, Lin X, Huang Q, Shen W (2010). "The effects of mineral fertilizer and organic manure on soil microbial community and diversity." *Plant and Soil*, **326**, 511–522.

RÉSUMÉ

Prise en compte du rôle de la diversité microbienne dans la simulation de la dynamique de la matière organique du sol (MOS) dans un contexte de transition vers l'agro-écologie

La thèse propose et évalue une méthode de prise en compte de la diversité microbienne dans un modèle mécaniste opérationnel de dynamique du carbone (C). Partant d'un modèle classique dont les flux de minéralisation sont décrits par des cinétiques d'ordre 1, nous proposons des fonctions de modulation des paramètres des équations prenant en compte des indices de diversité microbienne. Elle s'appuie sur 2 jeux de données contenant des mesures de diversité microbienne et des cinétiques de minéralisation du C sur 80 jours, en conditions contrôlées, avec et sans apport de blé marqué au ¹³C, issus d'échantillons de sols agricoles diversifiés.

L'utilisation de modèles statistiques additifs généralisés (GAM) a permis d'identifier les indices de diversité pouvant expliquer la dynamique du C. Des fonctions de modulation ont été proposées selon les relations mises en évidence. L'ensemble a été calibré à l'aide de méthodes bayésiennes. L'ajout d'une fonction de modulation dépendante de la diversité bactérienne a permis d'améliorer la prédiction des flux issus de la minéralisation de la MOS. En revanche, l'intégration d'une fonction dépendante de la diversité fongique n'a pas permis l'amélioration des flux issus de la minéralisation du résidu ou de la MOS.

La méthode proposée d'intégration de la diversité microbienne dans les modèles génériques de dynamique du C constitue une approche prometteuse pour une représentation parcimonieuse de son influence. Il s'agit d'un cadre méthodologique complémentaire à la prise en compte détaillée de cette diversité dans les modèles de dynamique du C qui domine dans la littérature.

Mots-clefs : sol, matière organique, résidu de cultures, diversité microbienne, modélisation

ABSTRACT

Taking into account microbial diversity influence on soil organic matter (SOM) dynamic predictions in a context of agroecological transition

This thesis suggests and assesses a way to integrate microbial diversity in an operational mechanistic model of C dynamic. From a classical model where mineralisation fluxes are described by 1st order kinetics, we propose functions depending on microbial diversity indices to modulate the parameters of these kinetics. We used 2 datasets containing measures of microbial diversity and C mineralisation kinetics from controlled conditions incubations during 80 days, with and without ¹³C labelled wheat straw, from grassland and cropland soil samples.

In a first step, based on statistical generalizes additive models (GAM), we identified the microbial diversity indices relevant to explain C dynamic. Observed relations allowed us to submit modulation functions depending on microbial diversity to adjust model parameters. The whole has been calibrated through bayesian inference. Overall, the modulation function depending on bacterial diversity led to an improvement of SOM mineralisation fluxes predictions, while fungal based modulation functions led to no improvement.

To conclude, the proposed method to integrate microbial diversity was a relevant approach towards a parsimonious representation of microbial diversity influence on C dynamic in generic model. It forms a complementary methodological framework to the dominant explicit and detailed way of microbial diversity representation in C dynamic models.

Keywords: soil, organic matter, crop residue, bacterial and fungal diversity, modeling



AGROCAMPUS OUEST • Institut supérieur des sciences agronomiques, agroalimentaires, horticoles et du paysage
65 rue de Saint-Brieuc – CS84215 – F-35042 Rennes Cedex
Tél. : 02 23 48 50 00
www.agrocampus-ouest.fr

