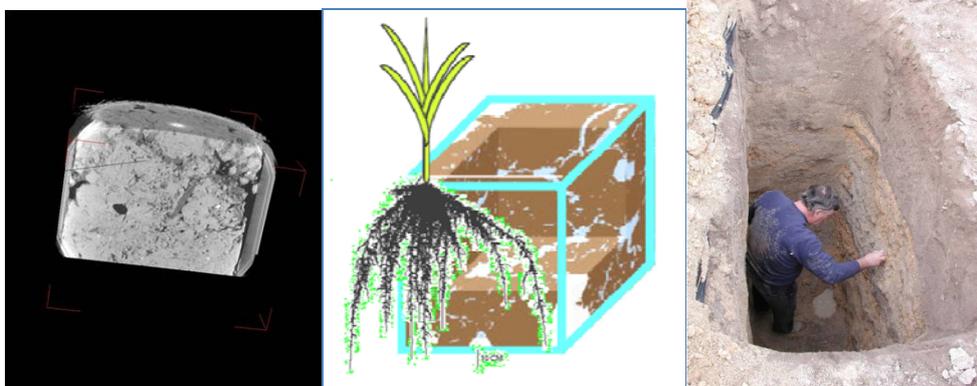


Plateforme SOL VIRTUEL

Spécifications fonctionnelles



Auteurs : Groupe Projet Sol Virtuel

Isabelle Cousin, Frédéric Gérard, François Lafolie, Nathalie Moitrier, Alain Mollier, Valérie Pot, Stéphane Ruy, Anatja Samouëlian

1. Introduction

Ce document est le cahier des charges fonctionnel du projet Sol Virtuel du Département Environnement et Agronomie de l'INRA. Il a pour objectifs de présenter succinctement les enjeux scientifiques du projet et les attentes des utilisateurs, sans préjuger des choix techniques ultérieurs. C'est un document de dialogue entre les scientifiques utilisateurs de la plateforme et les logisticiens chargés de la mettre en œuvre. En tant que tel, il n'est donc pas figé et sera amendé au cours du projet sous différentes versions. Il s'adresse donc à la direction du Département EA, aux directeurs et aux chercheurs des Unités du Département investis dans le projet, à l'ingénieur chef du projet Sol Virtuel et aux rédacteurs du cahier des charges non fonctionnel.

1.1 Contexte. Objectifs.

La construction de la plateforme Sol Virtuel vise à répondre à des objectifs scientifiques et organisationnels.

D'un point de vue scientifique, son objectif est de permettre la modélisation numérique de processus complexes se déroulant dans les sols et résultant de couplages entre champs disciplinaires différents, problèmes à grands nombres d'inconnues, résolutions numériques délicates, etc., afin de répondre à des problématiques émergentes concernant les fonctionnements biologiques, physiques et chimiques des sols naturels et cultivés. Il s'agit donc d'offrir une structure d'accueil pour la modélisation de ces phénomènes. Elle a aussi pour objectif connexe de permettre le couplage et la prise en compte des rétroactions entre ces processus. Elle doit permettre aux scientifiques de disposer d'une structure informatique d'accueil de leurs modules et elle doit faciliter le développement de nouveaux modèles.

D'un point de vue organisationnel, elle doit aider à fédérer les efforts de modélisation des scientifiques. Pour les modélisateurs, il faut qu'elle permette, d'accéder à un ensemble de modules du fonctionnement des sols, de partager les outils et de mutualiser les moyens pour le développement de modèle. Il ne faut pas que la plateforme se résume à un simple outil logiciel. Au-delà de cet aspect, elle doit être le moyen d'impliquer les ressources humaines (développeurs, numériciens) du département pour l'amélioration et le développement des modèles et des outils connexes. De façon plus large il faut qu'elle facilite pour tous l'utilisation des modèles. Elle a donc vocation à être un lieu d'accueil, d'assistance, de partage, de collaboration et à devenir un outil pédagogique.

1.2 Situation de ce document dans la démarche du projet

Pour amorcer le projet de la plateforme Sol Virtuel quatre axes thématiques ont permis de faire un état des lieux des modélisations existantes :

- extension des modèles 1D de représentation du sol à des thématiques émergentes,
- représentation architecturée de la dynamique des flux de nutriments, du pH et du pE dans l'ensemble système racinaire-sol,
- représentation des processus pédogénétiques et de l'évolution des propriétés et des constituants des sols,
- représentation explicite de la structure du sol, de sa dynamique et de la redistribution de matière solide.

Les réflexions menées dans le cadre de ces axes thématiques ont également mis en évidence les verrous des modèles actuels. Les principales faiblesses résident dans des processus et boucles de rétroactions encore peu ou pas modélisés à l'heure actuelle, comme par exemple : la dynamique de la structure et les interactions avec les fonctions du sol, la rhizosphère (mobilisation et prélèvement des éléments minéraux), la faune et l'écologie microbienne (écologie des populations).

Les échelles spatiales que nous avons considérées vont du minéral au profil de sol. L'échelle temporelle s'étend des temps brefs (min), par exemple pour simuler des écoulements préférentiels, à des temps longs (siècle) relatifs à l'évolution des sols.

Le présent document fait suite à ces réflexions initiales et a été réalisé par le groupe de travail missionné par le Département EA dans sa lettre du 18/02/2008. La démarche adoptée par le groupe de travail a été la suivante : le contenu de ce cahier des charges a été défini par les membres du groupe au cours du premier trimestre 2008. Ces premières réflexions ont été présentées et discutées dans chaque unité du Département investie dans le projet Sol Virtuel. La première version du cahier des charges fonctionnel est donc le fruit d'une réflexion collective.

1.3 Définitions.

Il est peut-être souhaitable, pour que les choses soient claires pour les informaticiens amenés à travailler sur la plateforme, de définir ce qu'est un sol et notre façon de voir les processus qui s'y déroulent.

Sol: *C'est un milieu poreux s'étendant de la surface jusqu'à une profondeur variable que l'on appellera « bas du profil ». Ce milieu est par essence tri-dimensionnel, mais, avec des hypothèses plus ou moins réalistes, on considère souvent que c'est un milieu monodimensionnel (dimension correspondant à la profondeur). Ce milieu est le lieu :*

- i/ de phénomènes de transferts de masse et d'énergie*
- ii / de processus chimiques*
- iii/ de processus biologiques*

Les divers processus sont en général représentés par des équations algébriques ou différentielles en temps et en espace. A noter que les constantes de temps des divers processus peuvent être très différentes. Certains des processus sont interdépendants. La

plateforme devra donc intégrer ces équations dans le temps et l'espace et les rétroactions entre processus.

Module : On entend par module un programme informatique pouvant être composé d'un ou plusieurs sous-programmes et réalisant une tâche « élémentaire ». Cette notion recouvre aussi bien des programmes que l'on pourrait qualifier d'utilitaires (solveurs par exemple) que des programmes correspondant à la simulation des mécanismes ou processus « élémentaires » devant être pris en compte. Par exemple, module du cycle Carbone – Azote, module de transfert hydrique, module de devenir d'un pesticide, module de géochimie, module de dynamique microbienne, etc. Ces modules sont destinés à être assemblés pour former des **modèles**. Par conséquent, la notion de tâche élémentaire est essentielle. Il est évident que des modules différents peuvent être disponibles pour représenter un processus donné. Il est aussi évident qu'un module requiert des entrées spécifiques qui ne peuvent être fournies que par certains modules en « amont ». Donc tout module ne pourra pas être connecté avec n'importe quel autre. Les modules doivent fournir certaines variables pour que les modules en aval ou couplés puissent fonctionner. Ils peuvent aussi fournir des variables annexes non utilisées par le reste de l'arborescence, mais disponibles pour les utilisateurs.

Modèle : Assemblage de **modules** permettant de simuler l'évolution du système bio-géophysico-chimique considéré. Cet assemblage ne peut pas être quelconque puisque les modules doivent être compatibles au sens où ils doivent pouvoir s'échanger les variables dont ils ont besoin. A partir d'un ensemble de modules on peut donc construire des modèles qui auront plus ou moins de modules en commun. Un modèle peut être constitué à la limite d'un seul module.

Squelette : Architecture prédéfinie au sein de la plateforme prenant en compte de la façon la plus exhaustive possible les couplages et les rétroactions entre processus. Un squelette s'appuie donc en premier lieu sur la connaissance scientifique que l'on peut avoir des processus et de leurs interactions dans le sol. Un squelette est donc une arborescence destinée à définir les liens entre processus. Le squelette ne fait pas référence à une représentation particulière d'un processus. Une réalisation particulière d'un squelette est un **modèle**. Un même squelette peut donc donner naissance à divers modèles.

2. Qui seront les utilisateurs de la plateforme ?

Avant de parler des « utilisateurs » sensu stricto, nous souhaitons préciser que deux types d'acteurs seront en interaction avec la plateforme :

-des acteurs dits « logisticiens » : ils ont vocation à développer et maintenir la plateforme. Ce sont donc des informaticiens mais également des scientifiques qui assurent la cohérence et la validation des nouvelles fonctionnalités à intégrer au sein de la plateforme.

-des acteurs dits « utilisateurs » : ils interagissent avec les logisticiens à des degrés divers que nous précisons ci-après. Selon les circonstances, un scientifique peut appartenir à l'une ou l'autre de ces catégories.

De nos visites dans les diverses unités, nous avons recensé un grand nombre de demandes (cf. synthèse en annexe) :

-des souhaits relativement simples comme par exemple « pouvoir accéder à des outils existants »,

-des souhaits plus ambitieux, par exemple : « pouvoir développer des modèles à partir de modules disponibles », « disposer d'un service pour la mise sous forme de programme informatique de mécanismes ou processus préalablement conceptualisés sous forme d'ensembles d'équations ».

Nous avons donc été confrontés à des demandes et des réactions extrêmement variées. Pour aboutir à un document utilisable par l'équipe en charge de la plateforme, nous avons essayé de les rassembler sous forme de « cas d'utilisation », la définition de ces cas permettant ensuite de définir concrètement les services que la plateforme doit offrir. C'est à partir de ces cas d'utilisation que pourront travailler les scientifiques et informaticiens en charge du développement. La définition des cas d'utilisation repose sur une classification des types d'utilisateurs. De nos contacts avec les unités, nous considérons que les utilisateurs potentiels de la plateforme peuvent être regroupés dans les quatre types suivants :

- a/ Utilisateur de modules ou de modèles « clé en main »
- b/ Utilisateur pour importer et coupler un module avec un modèle déjà disponible sur la plateforme
- c/ Utilisateur pour développer un module sur la plateforme
- d/ Utilisateur pour assembler des modules.

Nous détaillons ci-dessous les cas d'utilisation correspondant et ce qui est attendu de la plateforme dans chaque cas.

2.1 Cas d'utilisation numéro 1 : mise à disposition de modules et de modèles « clé en main »

Une demande récurrente est de pouvoir accéder à des modèles « clé en main ». Il est certain qu'un nombre important de modèles ont été développés (au niveau INRA, national et international) et sont utilisés au sein du département et de façon plus large dans la communauté scientifique nationale et internationale. Ces modèles traitent de thématiques diverses et offrent certainement des réponses toutes prêtes à des besoins d'utilisation de modèles.

La plateforme devra mettre à disposition des utilisateurs un ensemble le plus large possible des modèles disponibles. Pour ce faire elle devra proposer i) une aide à la sélection du modèle le plus approprié à la question posée, ii) une aide au choix des paramètres, iii) une aide au calcul, iv) des cas de références documentés permettant une prise en main efficace du modèle. Une liste des modèles qu'il est souhaitable d'avoir sur la plateforme est donnée en annexe.

Une variante de ce cas d'utilisation concerne des personnes qui souhaiteraient récupérer un module effectuant une tâche précise pour l'insérer dans un modèle plus global dont ils ont la maîtrise. La plateforme devra donc mettre à disposition de ces personnes une bibliothèque de modules. A noter que comme on le verra par la suite cette bibliothèque existera de-facto.

2.2 Cas d'utilisation numéro 2 : importation d'un module et couplage avec un modèle

Un deuxième cas d'utilisation correspond à la situation de chercheurs désirant coupler un module (en général développé par leurs soins) avec un modèle permettant une description plus globale du fonctionnement du profil de sol. Par exemple : module C/N, module devenir d'un pesticide, module géochimique, module de croissance de couvert, module de modification des propriétés hydrodynamiques, module de dynamique de populations microbiennes, module d'évolution des propriétés de surface, etc.

La plateforme devra donc permettre d'une part « l'importation » de ce module et d'autre part fournir les outils et/ou l'assistance pour que ce module puisse être utilisé en couplage au sein d'un modèle issu d'un squelette. L'opération d'importation devra pouvoir se faire de façon autonome ou avec assistance : un chercheur testant différentes versions d'un module devra pouvoir être autonome. Il est alors souhaitable qu'il n'ait pas à solliciter les compétences de l'équipe en charge de la plateforme. Par contre, le passage par l'équipe de la plateforme sera requis pour des raisons de sécurité et de mise en forme lorsqu'il s'agira de mettre un module à disposition de la communauté.

Dans certains cas on peut imaginer que le couplage avec un modèle de la plateforme sera aisé et pourra être réalisé au moyen de l'atelier logiciel (cf. §3.3), dans d'autres, on pourra aboutir à des problèmes numériques qui requerront une assistance : définition de processus itératifs, problèmes de convergence. Il faudra que l'équipe de la plateforme soit à même d'apporter cette assistance.

2.3 Cas d'utilisation numéro 3 : développement de modèles à partir de modules disponibles.

Un troisième cas d'utilisation est celui du développement de modèles à partir de modules disponibles. La plateforme devra donc rendre possible cette opération d'assemblage. L'objectif est très ambitieux. Il faudra en particulier régler le problème de « qui » peut-être couplé à « qui » pour éviter des assemblages conduisant à des aberrations. Il faudra aussi que la plateforme génère le code qui gère la boucle de temps et organise l'appel des divers modules. Si cette approche se révèle fonctionnelle elle pourrait alors en fait offrir une solution pour la plupart des besoins recensés. En effets les utilisateurs pourront, grâce à cette fonctionnalité, développer des modèles nouveaux mais aussi retrouver des modèles existants présentant des degrés de complexité variables.

2.4 Cas d'utilisation numéro 4 : développement de modules

Un des besoins exprimé est la mise sous forme informatique de processus dont on a une représentation plus ou moins formalisée. La plateforme doit permettre le développement de ces modules et ensuite leur intégration à la bibliothèque de modules utilisables. Cette spécificité peut concerner des aspects purement logiciels mais également nécessiter une assistance humaine comme par exemple pour le choix de méthodes de résolution des équations ou pour le traitement des problèmes numériques.

3. Que doit-on pouvoir trouver sur la plateforme ?

Au vu des besoins exprimés dans les cas d'utilisations, il nous a semblé que la plateforme devait offrir des services variés que nous avons classés en trois grands items.

3.1 Bibliothèque de modèles et de modules

Modèles prêts à l'emploi

En réponse au premier cas d'utilisation, la plateforme doit mettre à disposition des modèles prêts à l'utilisation et de mise en œuvre aisée. Cette mise à disposition devra comporter si nécessaire une aide au choix du modèle. Cette aide au choix devra prendre une forme conviviale qui « comme un expert » guide l'utilisateur vers le modèle correspondant le mieux à ses souhaits en tenant compte des contraintes résultant par exemple de la disponibilité des paramètres.

Chaque modèle sera accompagné d'une fiche synthétique, d'un manuel de l'utilisateur, de jeux de données test. Il devra aussi être accompagné de références et de documents décrivant le cadre théorique sur lequel il repose.

La plateforme fournira des modèles prêts sous forme d'exécutables. Il est hautement souhaitable que les modèles disposent d'une Interface Homme Machine (IHM), du moins pour les modèles faisant intervenir de nombreux phénomènes et donc requérant de nombreuses entrées.

De même, les IHM doivent offrir une visualisation des résultats.

Modules

La plateforme doit mettre des modules à disposition des utilisateurs (cf. glossaire). Ces modules doivent, comme les modèles, être accompagnés d'une fiche de synthèse, d'une description du cadre théorique (article par exemple) et d'une description précise des entrées et des sorties. Etant donné que ces modules sont destinés à être intégrés à des modèles internes (cf. la notion de squelette et de couplage de modules) ou externes à la plateforme, les sources informatiques devront être fournies.

3.2 Squelettes

Une réponse aux besoins exprimés aux cas d'utilisation 2.2, 2.3 et même 2.4 nous semble être le développement de « squelettes ». Comme défini au paragraphe 1.3, un squelette est une arborescence prédéfinie qui organise et fait communiquer les processus physiques, biologiques et chimiques au sein du domaine dans lequel ils se déroulent. Il nous semble que la plateforme devra héberger un nombre restreint de squelettes répondant à des besoins majeurs. Il faut donc penser les squelettes en termes de mécanismes, de couplages entre mécanismes et de rétroactions. Une étape importante sera de définir les squelettes c'est-à-dire de décrire les

mécanismes devant être considérés, les paramètres nécessaires à ces mécanismes ainsi que les liens entre ces mécanismes. Il faut que les squelettes soient « pensés » sans faire référence aux modèles existants, la réalisation informatique étant une étape ultérieure.

Il faut pouvoir changer les modules dans un squelette de façon à pouvoir tester de nouveaux modules. L'utilisation de différents modules pour décrire un processus amènera à générer des modèles différents au moyen d'un squelette. En pouvant utiliser différents modules pour un processus, un squelette permettra de générer des modèles qui diffèrent par la représentation de toute ou partie des processus considérés. Le squelette doit donc être capable de gérer la compatibilité des modules.

D'un point de vue informatique, le squelette pourrait correspondre à un programme qui génère un modèle en assemblant des modules. Si cela est faisable, le squelette pourrait donner naissance à une gamme élargie de modèles.

Il nous semble que d'un point de vue pratique on pourrait commencer à définir les squelettes en partant de la dimension d'espace. Ce mode de tri recouvre par ailleurs assez bien un classement des problématiques scientifiques importantes à traiter (cf. thèmes du programme). Une autre raison est de garder une liberté pour la définition de squelettes en 2 ou 3-D qui pourraient utiliser des concepts différents de ceux mis en œuvre pour le 1-D.

3.2.1 Squelette 1D

La plateforme devra mettre à disposition un squelette pour le cas 1D. Ce squelette devra permettre le couplage, c'est-à-dire tenir compte des échanges et des rétroactions entre les grands domaines suivants :

1. processus de transfert (eau, solutés, chaleur, gaz, colloïdes et particules)
2. processus impliquant des activités microbiennes (devenir de pesticides, cycle de la matière organique, dynamiques de populations microbiennes)
3. processus géochimiques (spéciation en phase liquide, précipitation, dissolution, échanges d'ions, adsorption, oxydo-réduction)
4. processus racinaires au sens large (absorption d'eau et d'éléments nutritifs, respiration, excrétion, etc.) et couplage avec le développement de la partie aérienne de la plante
5. modifications des propriétés physico chimiques induites par les processus biologiques, chimiques, mécaniques, et les interventions humaines au sens large.

Le squelette devra comporter un module de gestion du temps et des pas de temps. Ce module devra interagir avec les modules résolvant des équations différentielles. Il devra pouvoir donner naissance à des modèles applicables à différentes échelles de temps et à des modèles ayant des degrés de complexité différents dans la description des processus.

Les modèles générés par le squelette pourront être utilisés de façon déterministe ou stochastique, ce qui nécessitera leur couplage avec des outils de génération de valeurs de paramètres dans des lois de distribution statistiques diverses.

Les modèles générés pourront être utilisés avec des programmes d'estimation de paramètres (méthodes inverses).

Le squelette devra permettre le greffage de nouveaux modules ou la substitution de modules en fonction des besoins des utilisateurs : par exemple pouvoir utiliser divers modules du cycle C/N/P (cf. atelier logiciel).

Le squelette devra être évolutif. Il est essentiel que de nouvelles connaissances ou que les oublis des versions initiales puissent être pris en compte.

Pour les modèles 1D ayant besoin de prendre en compte une culture, la plateforme devra offrir l'accès à des modules de croissance du couvert végétal permettant de simuler la mise en place du couvert et des racines ainsi que les besoins en éléments nutritifs.

3.2.2 Squelettes 2D

La plateforme doit mettre à disposition un squelette 2D qui permette de simuler le fonctionnement d'un sol avec des concepts identiques à ceux utilisés en 1D. Par ailleurs, une spécificité du 2-D est de permettre une première prise en compte des hétérogénéités du milieu. Cette prise en compte peut-être explicite (définition de zones identifiées avec des propriétés particulières) ou passer par l'utilisation de distribution spatiale de propriétés, corrélées spatialement ou non. Il faut donc que le squelette offre les outils ad-hoc pour permettre cette prise en compte.

Une deuxième spécificité du 2-D est de permettre d'aborder des problèmes prenant en compte l'hétérogénéité spatiale de l'enracinement en tenant compte des phénomènes physico-chimiques au voisinage des racines (flux hydriques, prélèvements minéraux, excrétions de protons, de ligands organiques, etc.).

Ainsi, nous prévoyons que le squelette 2-D permette de coupler les flux d'eau, de solutés et de gaz aux processus biogéochimiques dans un domaine avec racines distribuées. Les associations symbiotiques (mycorhize, nodosités des légumineuses) devront également être considérées dans cette architecture.

3.2.3 Squelettes 3D

On peut penser ici à plusieurs types de squelettes en fonction de problèmes scientifiques précis.

La plateforme devra proposer une modélisation 3D qui permette de modéliser de façon explicite la structure du sol en tenant compte des rétroactions entre la structure et les fonctions du sol.

L'une des spécificités des squelettes 3D est d'offrir une modélisation qui tienne compte de l'organisation explicite du réseau de pores, dont la connectivité joue un rôle fondamental dans le transport de l'eau, des particules, des nutriments. C'est également l'organisation du réseau de pores qui détermine l'accessibilité des microorganismes responsables de la biodégradation des matières organiques du sol.

D'autres spécificités du 2D se retrouvent dans le 3D, à savoir : (1) pouvoir tenir compte d'hétérogénéités distribuées dans le domaine de simulation, (2) pouvoir prendre en compte de façon fine des fonctionnements hétérogènes et l'existence de gradients locaux pour des variables d'états, ces gradients apparaissant en raison de

la compétition entre des phénomènes de transfert et des termes puits/source (activités microbiennes, réactions).

Il faut donc que la plateforme propose un squelette qui permette une distribution hétérogène des paramètres ou des situations initiales selon des lois statistiques avec des corrélations spatiales à l'image de ce qui sera proposé pour les modèles 2D. Ce squelette doit permettre également de prendre en compte les boucles de rétroaction, notamment les modifications de l'organisation des pores.

Il y a donc plusieurs squelettes 3D envisageables en fonction des objectifs scientifiques. Il faudra donc préciser des priorités et définir précisément le contenu de chacun.

3.3 Atelier logiciel

L'atelier logiciel doit être une interface graphique qui permette de répondre aux objectifs suivants :

- être le lieu de développements de modules élémentaires en conformité avec les éventuelles règles de développement édictées dans le cadre de la plateforme,
- permettre à un utilisateur de tester un module en le greffant sur des modèles déjà développés ou présents dans la plateforme
- de construire des modèles à partir d'un squelette donné.

Cet atelier doit offrir une aide au choix des modules qui vont simuler les différents processus. L'interface doit avoir une fonction de contrôle des compatibilités entre les modules qui seront assemblés. Elle doit aboutir à un modèle exécutable au sein de la plateforme ou en local. Le modèle doit bien entendu être interfacé avec des outils de visualisation des résultats.

3.4 Outils périphériques

Nous présentons ici une liste non exhaustive d'outils périphériques qu'il nous semble intéressant voire indispensable de trouver sur la plateforme en plus des modèles/modules et squelettes. Nous les avons classés en plusieurs catégories :

- Des outils permettant l'exploitation des résultats et des données :
 - visualisation des résultats 1D, 2D, pseudo 3D et 3D produits par les modèles hébergés ou développés au sein de la plateforme
 - exportation des résultats vers des logiciels d'exploitation non hébergés par la plateforme
 - analyse permettant d'effectuer un traitement statistique et/ou géostatistique des résultats et/ou données. Il serait également utile de disposer d'outils de génération de distribution (méthode de Monte-Carlo par exemple)
 - analyse spatiale, géotraitement et gestion de données
- Des outils d'aide au paramétrage et à la validation des modèles :

- génération de distribution (éventuellement spatiale et/ou corrélée) des paramètres ou variables d'entrée,
- analyse de sensibilité,
- procédures d'inversion de modèles pour estimer des paramètres. Ces procédures devront être génériques, c'est à dire qu'elles pourront être utilisées pour n'importe quel modèle de la plateforme,
- assimilation de données.
- Des outils annexes tels que :
 - traitement d'images
 - calcul formel
 - solveurs d'équations aux dérivées partielles et d'équations différentielles ordinaires.

A noter que les outils de type « calcul formel » et « solveurs EDP » permettront d'aider au développement de modules/modèles sur la plateforme.

3.5 Accès à des bases de données

Afin de paramétrer les modèles ou modules mis en œuvre dans le cadre de la plateforme, il est nécessaire que cette dernière puisse être connectée à des banques de données sur les sols et leurs propriétés physiques, chimiques, microbiologiques, mais également sur le climat, la plante. Ces données pourront être organisées dans des bases de données (par exemple : Donesol, Physol, Unsoda, bases de données thermodynamiques des modèles de géochimie, etc...), notamment celles gérées par les unités de service du Département EA (Infosol, Agroclim), ou dans des fichiers plus simples (type excel ou html). Dans tous les cas, ces données devront être documentées.

Les données pourront être de deux types : i) des données issues d'observations ou d'expérimentation mises à disposition par certains utilisateurs. Dans ce cas, les données peuvent avoir des degrés de confiance variables ; ii) des jeux de données ou de paramètres test, parfaitement validés (voir par exemple les paramètres des propriétés hydriques par texture de sol, disponibles par défaut dans certains modèles de transfert d'eau).

4. Spécifications non fonctionnelles ou techniques

La plateforme offrira la possibilité de faire les calculs sur un serveur.

La plateforme permettra le téléchargement d'un modèle exécutable et des fichiers afférents pour une utilisation en mode local.

La plateforme permettra de télécharger des modèles exécutables dans les environnements habituels (Windows, Linux, Unix, Mac).

Au sein de la plateforme, l'accès à des modules en phase de développement sera restreint, d'une part pour éviter des problèmes de propriété intellectuelle, et plus généralement pour éviter l'utilisation d'outils non complètement validés, cette utilisation pouvant conduire à des problèmes divers.

La plateforme devra accepter des modules écrits dans les langages de programmation numérique les plus courant : FORTRAN, C, C++.

5. Priorités

Une des priorités sera le développement d'un squelette 1D offrant des possibilités de couplages variés et permettant d'aborder des problèmes avec des échelles de temps allant du phénomène bref à des temps très long. D'une part, ceci apportera une réponse à des préoccupations d'actualité (impact des activités anthropiques sur la qualité de l'eau, des sols et de leurs évolutions). D'autre part, le développement de ce squelette mettra à disposition de la communauté scientifique un outil permettant de répondre à des questions diverses. Enfin, de part sa conception ouverte, il pourra servir de structure d'accueil de modules pour une grande partie des thématiques abordées dans la communauté scientifique concernée.

Il existe par ailleurs des raisons fortes et des souhaits clairs pour développer des squelettes 2D ou 3D permettant :i) d'aborder le rôle des hétérogénéités de paramètres et/ou de loi de fonctionnement, ii) la simulation du fonctionnement en prenant en compte des descriptions explicites de la géométrie du réseau de pores. Pour le 3D, plusieurs squelettes sont envisageables comme indiqué (§3.2.3). Le cas des squelettes 3D est plus ambitieux et on peut penser qu'il conduit probablement à des représentations plus proches de la réalité que le cas 2D. Il a par contre la faiblesse de conduire à des tailles de problèmes et des temps de calcul qui peuvent devenir prohibitifs.

Quelques soient les priorités, une tâche incontournable et primordiale est de récupérer les modules disponibles qui correspondent à ce ou ces chantiers. Cette étape cruciale ne sera réalisée qu'avec la participation active des « modélisateurs ». La récupération de modules et de modèles permettra de proposer rapidement une bibliothèque de ces outils. Le développement conjoint du ou des squelette(s) pourra alors s'appuyer sur l'existence de cette bibliothèque.

6. Questions diverses que l'on se pose et que l'on nous pose. Sentiments. Réflexions.

Nous tenons à faire part de quelques réflexions et idées collectées lors de nos visites ou ayant surgi lors de nos discussions.

Gestion de la plateforme

Certaines décisions stratégiques concernant l'évolution de la plateforme à court ou moyen terme ne relèveront pas exclusivement de compétences informatiques. Il s'agit de questions scientifiques ou non, telles que, par exemple :i) la mise à disposition d'un module et de son environnement, ii) le choix entre différents outils périphériques (cf. §3.4), iii) la gestion des priorités en matière de développement, iv) l'arbitrage entre des demandes trop nombreuses d'assistance sur

la plateforme, etc. Il nous semble donc nécessaire de mettre en place un comité de gestion de la plateforme composé de scientifiques et d'informaticiens.

Plateforme : outil informatique ou plus ?

La plateforme doit-elle être simplement un outil informatique ou faut-il avoir une vision plus large ? Plus précisément, il est souhaitable que la plateforme ne se limite pas à une mise à disposition d'outils aussi sophistiqués soient-ils ; mais elle doit être une structure d'aide au développement, ce qui suppose la mise à disposition du temps « humains » sur des problèmes donnés. On peut même aller plus loin et considérer la « plateforme » comme un acteur de la recherche qui apparaisse explicitement dans des réponses à des appels d'offre (ANR, Europe, Projets INRA, Projets Innovant EA, etc.). Elle pourrait tirer bénéfice de cette position en se faisant financer sur ces projets du temps d'informaticien ou autre compétence.

Scientifiques et plateforme

Lors des réunions de préparation de ce document, il nous est apparu clairement que la plateforme aura du mal à voir le jour sans une adhésion totale des scientifiques, lesquels seront au cœur de la réflexion initiale sur le contenu des squelettes et sur la spécification des données échangées par les modules. Cette adhésion au projet signifie également, pour ces scientifiques, de devoir passer du temps pour mettre à disposition de la plateforme des modèles et des modules, mais aussi d'être disponible pour assurer une assistance scientifique qui ne saurait être demandée aux informaticiens. Il est essentiel que cette implication soit reconnue lors des évaluations. Nous nous faisons ici l'écho de remarques souvent entendues lors de nos visites dans les unités. Un autre point est de s'assurer d'une reconnaissance de la propriété intellectuelle. Enfin, l'absence d'alimentation de la plateforme par des modules nouveaux est un risque à ne pas négliger, si l'on en croit l'expérience d'autres plateformes. Si le flux est à sens unique, la plateforme deviendra vite obsolète et inutilisée.

Formation

Il faut prévoir de mettre en place dès que la plateforme sera opérationnelle des sessions de formation à son utilisation. Ces formations pourront prendre la forme d'écoles chercheurs sur des thématiques ciblées, ce qui constituerait un moyen d'agrèger les scientifiques autour de l'outil. Mais pour être efficace cette activité aura un coût en temps important qui encore une fois ne relèvera pas que des informaticiens. L'implication des scientifiques fournisseurs de modèles et modules, dans la réalisation de ces actions de formation et d'échanges, est incontournable.

Outils non INRA

Des modèles non développés à l'INRA (Hydrus, MIN3P par exemple) sont utilisés dans nombre d'unités. Ces outils sont quelquefois gratuits mais peuvent aussi être payants (Hydrus 2D/3D). De plus, ces outils peuvent être imposés dans le cadre de certains projets internationaux. Une question non résolue concerne donc la question de leur présence et/ou des conditions de leur présence sur la plateforme.

Outils payants

Quelle attitude avoir face à l'utilisation de logiciels payants qui sont souvent des outils de post traitement? Faut-il les mettre dans la plateforme parce que certains en ont besoin ? Faut-il se limiter à des outils gratuits ? Une sortie à des formats compatibles n'est-elle pas suffisante pour permettre un traitement hors plateforme?

7. Glossaire de modules

Ce glossaire présente une liste non exhaustive des modules que nous voudrions voir figurer dans une première version de la plateforme Sol Virtuel et qui devra être complétée au fur et à mesure des besoins des utilisateurs.

Modules de transfert hydrique (réservoirs, Richards, écoulements préférentiels, Richards + écoulements préférentiels)

Module de transfert de chaleur

Module permettant le couplage avec le bilan d'énergie et les transferts dans l'atmosphère

Modules de transfert de substances dissoutes (CDE, MIM, LEA, Cinétiques, Multi-sites, etc.)

Modules de transfert de gaz (advection, Fick, Gaz Poussière)

Modules de transfert de colloïdes biotiques et abiotiques

Modules du cycle de la matière organique (Century, Cantis, ROTHC, etc.)

Modules de dénitrification

Modules de volatilisation

Modules de décomposition de molécules xénobiotiques (plusieurs)

Modules de dynamiques microbiennes

Modules de géochimie (module géochimique de MIN3P, PHREEQ)

Modules de développement aérien et souterrain d'une plante

Modules pour le calcul des termes puits de la culture pour l'eau et les nutriments

Modules d'évolution des propriétés hydrodynamiques et de la structure.

8. Glossaire de modèles

Ce glossaire présente une liste non exhaustive des modèles existants INRA et hors-INRA qui pourraient figurer dans une première version de la plateforme Sol Virtuel :

Modèles existants INRA :

- PASTIS (Predicting Agricultural Solute Transport In Soil : Modèle 1D de transport eau- chaleur – solutés (MIM) – gaz avec couplage avec C/N, dénitrification, pesticides, plante croissance et absorption eau et azote, décomposition de résidus en surface ou non) (F. Lafolie, EMMAH)
- CANTIS (Carbon and Nitrogen Turnover In Soil : module de transformation biologique et dynamique du Carbone et de l'Azote – 0D) (B. Mary, AGRO-IMPACT ?)
- VOLT'AIR (Modèle de volatilisation de l'ammoniac et pesticides au champ – 1D) (S. Genermont, EGC)

- Modèle de prélèvement d'eau par un système racinaire architecturé 2D et 3D (C. Doussan)
- TRT LB (Modèle de transport eau et solutés avec prise en compte de la structure du sol – 3D) (V. Pot, EGC)
- KDW : Modèle écoulement préférentiels (1-D) (L. Di Pietro, EMMAH)
- GEOCHEM Modèle couplage Géochimie activités microbiennes (0-D) (P. Renault, EMMAH)
- KDWP et autres : Modèles colloïdes (1-D) (E. Michel, EMMAH)
- TEC Transfert couplé Eau-Chaleur Bilan d'énergie (1-D) (L. Bruckler et A. Chanzy, EMMAH)

Modèles existants issus de collaboration INRA avec autre partenaire :

- MOSAIC (Modelling of Soil Microbial Activity Investigated by Computer vision – 3D) (O. Monga de Geodes IRD, Patricia Garnier et V. Pot, EGC)
- FUSSIM (Transport et prélèvement d'eau et solutés, couplé à la croissance racinaire, 2D). Modalités de mise à disposition à voir.

Modèles existants hors INRA :

- MIN3P (géochimie et transport-réactif 1-D à 3-D, écoulements préférentiels 1-D, embryons de processus biologiques liés aux racines et aux microorganismes)
- PHREEQC (géochimie et couplage avec les transferts d'eau et de solutés en milieu saturé)
- HYDRUS -1,2,3D (transport eau – chaleur – solutés dans le sol avec modules de transfert préférentiel (1D), de transfert de particules (colloïdes, virus, bactéries) (1D), de production et transport de CO₂ (1D), de géochimie du Carbone (1D)) (Seule la version 1D est gratuite)
- HP1 (transport – réactif 1-D, couplage des modèles Hydrus 1D et Phreeqc)
- RETC (RETention Curve: modèle de description et prédiction des courbes de rétention et de conductivité hydraulique des sols)
- RothC (modèle de dynamique du C organique des sols – 0D)
- PEARL (transport eau-chaleur- solutés avec module de volatilisation de pesticides – 1D)
- PRZM (Pesticide Root Zone Model : transport eau – solutés (pesticides) avec module de croissance de plante (racines, parties aériennes), interception par le couvert végétal, de volatilisation de pesticides – 1D)
- MACRO (transport eau – solutés avec modules de transport préférentiel et interception par le couvert végétal– 1D)
- PELMO (transport eau- solutés avec module de volatilisation de pesticides, module ruissellement de surface et érosion– 1D)
- SWAP (transport eau – plante)
- MIOR (Modèle Individu centré de minéralisation des matières organiques par les microorganismes. 3D)
- SWORM (modèle individu-centré pour les effets des vers de terre sur les propriétés du sol)

9. Synthèse des visites des unités

Un groupe de 14 unités (UR et UMR) désignées dans la lettre de mission 4289-0208 ont été contactées par les membres du groupe de travail. Ces structures avaient manifesté leur intérêt pour le programme « sol virtuel ». Pour mémoire, il s'agissait des unités Science de Sol (Orléans), EMMAH (Avignon), BEF (EFPA, Nancy), EGC (Grignon), LSE (Nancy), BSR (Montpellier), TCEM (Bordeaux), INFOSOL (Orléans), BIOEMCO (Paris), PESSAC (Versailles), GSE (Aix en Provence), MSE (Dijon), SAS (Rennes), LISAH (Montpellier), et LBE (Narbonne).

Un exposé commun a été présenté devant toutes ces structures. L'objectif de cet exposé était de recueillir les souhaits et les préoccupations vis-à-vis de la mise en place d'une plateforme de modélisation du fonctionnement bio-physicochimique d'un sol. Son contenu est mis en annexe de ce document.

L'ensemble des unités a accueilli favorablement nos propositions. Leurs remarques ont servi à l'élaboration du présent document. Le tableau 1 synthétise les réactions recueillies dans les différentes unités. Les principaux points soulevés ont porté sur :

- l'organisation du pilotage de la plateforme (qui, où ?) ;
- le contenu opérationnel de la plateforme, aussi bien matériel et humain (localisation, moyens) que scientifique. A ce titre, devant la difficulté et l'ampleur de la tâche, des priorités devront être rapidement données et les premiers modèles/modules précisés ;
- l'ouverture de nos réflexions à l'international, et l'association d'autres instituts français (INRIA, BRGM, IRD, CIRAD) à celles-ci ;
- l'ouverture disciplinaire (écologie du sol en particulier) ;
- le bilan des plateformes et modèles existants. Anticiper les couplages entre plateformes.
- les bases de données (sol, géochimiques, hydriques etc.), les fonctions de pédo-transfert, les outils d'évaluation, d'inversion etc.
- le rapport coût/bénéfices pour le scientifique impliqué dans la plateforme ;
- le type de plateforme à développer dans un premier temps (type « supermarché » ?)

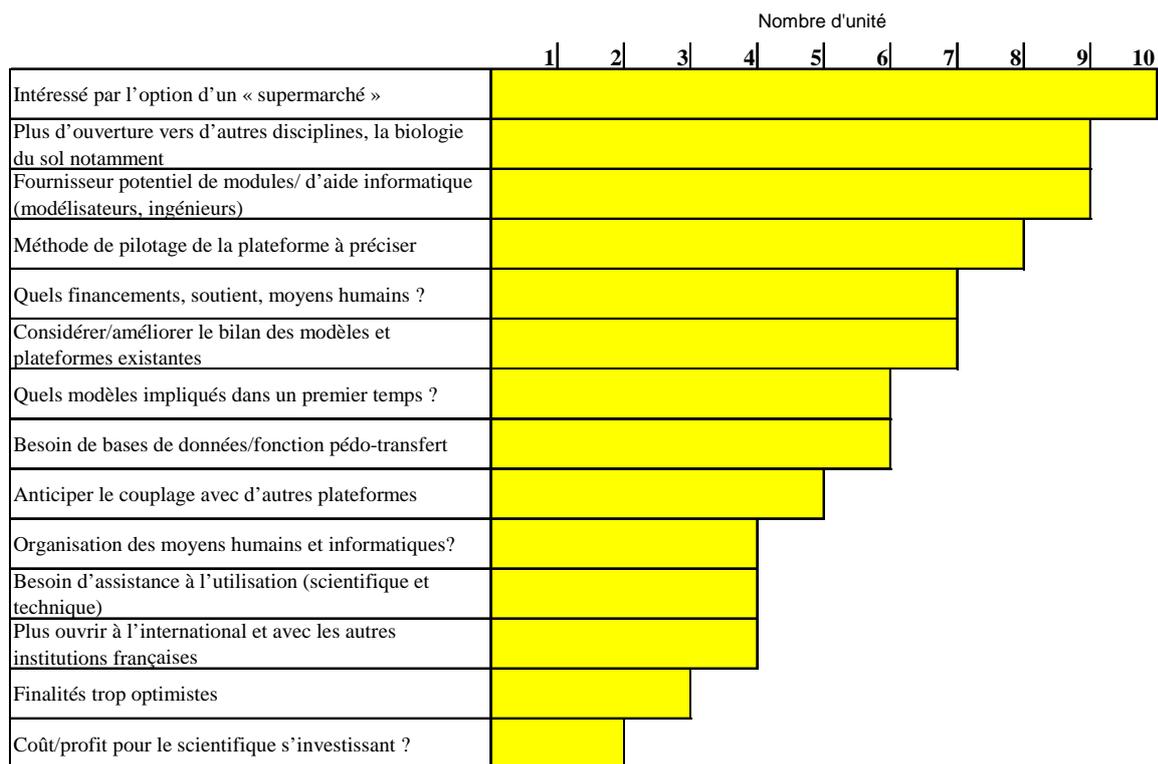


Tableau 1 : Récapitulatif des souhaits émis par les unités visités.