

ATELIERS

MOBILISER ET VALORISER LES DONNÉES SUR LES SOLS FORESTIERS

Compte rendu de l'atelier 4

THOMAS EGLIN – MANUEL MARTIN – DAMIEN MAURICE – MANUEL NICOLAS – CÉLINE PERRIER – MIRIAM BUITRAGO – GUY LANDMANN

Afin de lutter contre le changement climatique, la politique française de développement des énergies renouvelables⁽¹⁾ offre une place de premier plan à la biomasse solide et à la filière forêt-bois. La production et la récolte de bois devraient s'intensifier d'ici les prochaines années et les sols forestiers être davantage sollicités. Dans le cadre du Grenelle de l'Environnement, les partenaires forestiers majeurs se sont ainsi accordés sur l'objectif de « *produire plus de bois tout en préservant mieux la biodiversité* », notamment « *la biodiversité ordinaire qui accompagne et conforte les processus de production* » et en « *optimisant ressources et récoltes* » (FNE et al., 2007). Parallèlement, l'usage et la gestion de ces sols devront préserver les multiples services écologiques fournis par les écosystèmes forestiers : stockage et régulation des flux d'eau, stockage de carbone et régulation des émissions de gaz à effet de serre...

En raison de la durée des révolutions en forêt, la définition des objectifs de production et de récolte doit tenir compte des évolutions environnementales, notamment climatiques. Une meilleure connaissance des propriétés et de l'évolution des sols, en particulier de leur fertilité (nutriments, teneurs en matières organiques, réserve utile, sensibilité au tassement, biodiversité), en lien avec les pratiques de gestion, les pollutions atmosphériques et le climat, sera un élément important à considérer dans la conduite future des peuplements forestiers.

La pratique des analyses de sol n'est guère répandue parmi les gestionnaires forestiers, contrairement au milieu agricole. Les analyses effectuées en forêt relèvent principalement de projets d'étude plus ou moins ponctuels, de recherches (exemple : dispositifs expérimentaux) ou de suivi continu. Les résultats ne sont pas toujours répertoriés dans des bases de données. Certains restent encore au format papier ou dans des tableurs isolés. Valoriser ces données pour produire des références et des indicateurs utiles aux forestiers constitue donc un enjeu qui nécessite de s'intéresser à plusieurs questions :

- Quels sont les enjeux et les besoins ?

La définition des besoins est en effet déterminante pour définir les modalités d'acquisition et de mobilisation des données et leurs limites d'exploitation.

- Quelles sont les ressources existantes et pertinentes, comment y accéder et les exploiter ?

Il existe une ressource importante et sous-exploitée de données sur les sols forestiers, que ce soit au niveau des organismes de développement ou des laboratoires de recherche. Il s'agit d'identifier

(1) Plan d'action national en faveur des énergies renouvelables 2009 – 2020 : http://www.developpement-durable.gouv.fr/IMG/pdf/o825_plan_d_action_national_ENRversion_finale.pdf (consulté le 7/05/2014).

les démarches et modalités techniques existantes pouvant faciliter l'accès à l'information et la mise à disposition des données.

– Quelle coordination de la production et de la gestion des données ? Vers une centralisation systématique ?

L'intérêt de disposer de différentes sources de données est souvent accru lorsque que celles-ci peuvent être mises en commun pour répondre à une même question. La centralisation des données peut être, dans certains cas, une piste intéressante pour favoriser une amélioration de la qualité et de la comparabilité des données produites, et des économies d'échelle.

Le présent article propose une analyse de ces questions basées sur les discussions de l'atelier « Mobiliser et valoriser les données sur les sols forestiers » des journées REGEFOR.

BESOINS DE RÉFÉRENCES POUR MIEUX INTÉGRER LE SOL DANS LA GESTION ET LES POLITIQUES FORESTIÈRES

Les besoins en données et en références sur les sols discutés au cours de l'atelier sont liés à divers enjeux auxquels les forestiers sont confrontés :

- une meilleure **maîtrise de la gestion forestière** dans un contexte d'augmentation de la production et de la récolte forestières. Pour raisonner l'augmentation éventuelle des exportations, il faudra des indicateurs fiables permettant d'optimiser le retour des nutriments au sol (exemple : par apport de cendres issues de la combustion de biomasse sans adjuvant ou par prélèvement raisonné des rémanents) selon les types de sol et les peuplements (*cf.* Legout, Richter *et al.*, ce numéro). Outre des données sur la fertilité minérale, il faudra de meilleures connaissances sur la sensibilité au tassement, mais également sur la réserve utile (*cf.* ci-dessous) en lien avec les pratiques de gestion (*cf.* Ponette *et al.*, ce numéro).

- **l'adaptation des forêts à l'évolution du climat** qui devrait se manifester notamment par une augmentation de la température moyenne annuelle et par une modification du régime des précipitations (alternance de périodes sèches et de périodes avec de fortes précipitations) (*cf.* Ferry *et al.*, ce numéro). Cette évolution met en exergue la question de la disponibilité en eau pour les plantes, pour partie déterminée par les caractéristiques physicochimiques des sols et par la profondeur de sol accessible aux racines (*cf.* Biljou^{©(2)}, site dédié au modèle de bilan hydrique forestier). En croisant scénarios climatiques, caractéristiques des sols, écologie des essences et pratiques sylvicoles, les gestionnaires peuvent espérer mieux apprécier les contraintes hydriques futures.

- **l'atténuation du changement climatique** (*cf.* Ferry *et al.*, ce numéro). Les forêts contribuent à atténuer l'augmentation de la concentration en dioxyde de carbone dans l'atmosphère selon deux principaux mécanismes :

- la fixation et la séquestration de carbone dans le bois (sur pied, mort, racines), la litière et les matières organiques des sols,
- la substitution de ressources fossiles *via* l'utilisation du bois pour produire de l'énergie ou comme matériau.

C'est sur la base de ce principe que la décision européenne n° 529/2013/EU va obliger les États membres à considérer dans la comptabilisation de leurs émissions de gaz à effet de serre l'impact de la gestion forestière sur les stocks de carbone en forêt, dont la litière et les sols (UE, 2013). De nouvelles méthodes de suivi du carbone des sols basées sur des mesures sont recherchées. Cette

(2) <https://appgeodb.nancy.inra.fr/biljou/> (consulté le 16/12/2013).

comptabilisation impose des exigences particulières car elle doit pouvoir s'appuyer sur des données suffisamment transparentes et permettant de quantifier l'évolution des stocks dans le temps.

L'atelier n'a sans doute pas épuisé la question de l'adéquation entre les besoins en références sur les sols forestiers pour différents enjeux et différentes échelles spatiales (massif, sylvo-éco-région, territoire national) et l'offre potentielle de données. Il a cependant permis de dresser un premier panorama utile sur ces questions.

L'EXISTANT : LE CONNAÎTRE, Y ACCÉDER ET L'EXPLOITER

Connaître l'existant

Pour être utilisées, les données de sols forestiers existantes doivent avant tout être connues et facilement accessibles. Or, il n'existe à ce jour aucun recensement complet de ces données. Elles sont stockées par différents organismes forestiers, de recherche et de développement essentiellement. Leur mise en base de données n'est pas systématique. Quand elles sont intégrées dans des bases, ces dernières ne sont pas toujours compatibles entre elles, rendant difficile une comparaison ou une mise en commun des données. Par ailleurs, retrouver le protocole suivi pour la collecte n'est pas toujours évident. Le réseau mixte technologique AFORCE⁽³⁾ a récemment engagé une enquête auprès de ses partenaires afin d'initier un recensement des besoins en données dans le contexte particulier du changement climatique et pour les faire connaître.

La collecte des données

Un premier tour d'horizon des données existantes permet de distinguer trois types d'actions pour lesquelles il y a un recueil de données de sol (tableau I, p. 562) :

- les dispositifs de suivi et d'inventaire nationaux : protocoles relativement homogènes au sein d'un dispositif, données géo-référencées et stockées dans une base de données propre à chaque réseau ;
- les dispositifs expérimentaux pluriannuels : diversité des protocoles et des formes de stockage des données (parfois en base de données), géo-référencement non systématique ;
- les études et recherches ciblées, ponctuelles dans le temps (études stationnelles, projets de recherche et développement, etc.) : diversité des protocoles qui ne sont pas toujours référencés, données dispersées, stockage partiellement informatisé, opérateurs variés et parfois peu familiers des bases de données.

Les laboratoires d'analyses de sol collectent aussi un nombre important de données. L'importance de cette ressource est néanmoins inconnue et probablement partiellement redondante avec les autres sources listées dans le tableau I (p. 562). Il est à noter qu'un programme de mise en commun des données d'analyse de terre existe déjà pour les sols agricoles (base de données d'analyses de terre (BDAT) du GIS Sol) et pourrait être adapté aux sols forestiers.

Les bases de données

Il existe plusieurs bases de données (BD) à l'échelle nationale, parmi lesquelles :

- la BD du projet européen BIOSOIL⁽⁴⁾ – *Forest Soil and Biodiversity Monitoring in the EU*,

(3) <http://www.foretpriveefrancaise.com/aforce>

(4) Pour les rapports sur le suivi des sols forestiers en Europe, voir <http://www.flanders.be/en/search?keyword=biosoil> (consulté le 22/04/2015).

- la BD du réseau RENECOFOR⁽⁵⁾ (Réseau national de suivi à long terme des écosystèmes forestiers),
- la BD nationale des informations pédologiques DONESOL⁽⁶⁾, qui regroupe les données des programmes IGCS (inventaire, gestion et conservation des sols) et RMQS (réseau de mesures de la qualité des sols) gérées à InfoSol⁽⁷⁾ dans le cadre du GIS Sol,
- la BD phytoécologique EcoPlant gérée par l'équipe Écologie forestière – UMR LERFoB qui recense 5 838 descriptions dont 5 388 avec analyses de sol (Gégout *et al.*, 2005).

Ces bases de données ont été construites indépendamment les unes des autres. Elles n'ont pas de structure commune, et n'optimisent pas toutes le lien entre les descriptions de sol et les analyses physicochimiques correspondantes. Certaines présentent des informations partiellement redondantes entre elles, comme DONESOL et EcoPlant.

TABLEAU I Principales ressources de données identifiées

	Intitulé	Type de données	Nombre de points en forêt (évaluation 2014)	Forme de stockage
Dispositifs d'inventaire nationaux	Inventaire forestier national (IGN)	Description de sondages Données qualitatives	~ 160 000 relevés	Numérique
	Inventaire, Gestion et Conservation des Sols (IGCS)	Description de sondage et de profils	5 475 profils identifiés	Numérique
Dispositifs de suivi (cf. Nicolas <i>et al.</i> , ce volume)	RENECOFOR	Description de profils Analyses quantitatives sur plusieurs profondeurs	102 placettes permanentes	Numérique
	Réseau systématique de suivi des dommages forestiers, 16 x 16 km	Description de profils Analyses quantitatives sur plusieurs profondeurs	~ 550 placettes permanentes	Numérique
	RMQS 16 x 16 km (forêt)	Description de profils Analyses quantitatives sur plusieurs profondeurs	~ 600 points	Numérique
Dispositifs expérimentaux	Divers organismes (exemple : INRA, ONF, CNPF, FCBA, universités...)	Divers	Inconnu, estimé > 3 000	Stockage sous des formats très divers, de plus en plus en bases de données
Études et recherches ciblées (dont catalogues et guides de stations)	Divers organismes (CNPF, ONF, bureaux d'études, chambres d'agriculture...)	Divers	Données très nombreuses mais pour majeure partie difficilement accessibles	Local, surtout format papier

(5) <http://www.onf.fr/renecofor/> (consulté le 16/12/2013).

(6) <http://www.gissol.fr/outil/donesol/donesol.php> (consulté le 16/12/2013).

(7) <http://www.gissol.fr/programme/rmqs/rmqs.php> (consulté le 16/12/2013).

Les métadonnées

Les ressources concernant les sols forestiers sont donc multiples et assez mal connues de la communauté forestière. Pour porter à connaissance leur existence, le catalogage de métadonnées (ou description de données) se développe. Il permet également de répondre à certaines exigences réglementaires, telles que la directive européenne INSPIRE (*Infrastructure for Spatial Information in Europe*) qui concerne les données géographiques à caractère environnemental et impose, à des fins d'interopérabilité, une norme prédéfinie pour la diffusion des métadonnées (UE, 2007). Si trois niveaux de richesse d'information sont décrits classiquement dans les métadonnées – découverte, exploration et exploitation – l'effort de renseignement réalisé se limite le plus souvent au niveau découverte qui permet de diffuser une information minimale de manière homogène. Les principales informations attendues dans le cadre d'INSPIRE sont : présentation (titre, description, lien), couverture spatiale et temporelle, mots-clés (thésaurus ou libre), contacts, origine et modalités de création des données et contraintes d'accès et d'utilisation. Le respect des normes dans les catalogues de métadonnées permet de réaliser des recherches sur différents catalogues *via* le mécanisme de « moissonnage ». Les métadonnées cataloguées *via* des outils accessibles en ligne sont libres d'accès. Parmi les initiatives dans le milieu forestier, le GIP ECOFOR a créé le catalogue des sources d'informations sur la forêt – CaSIF⁽⁸⁾ – qui a pour vocation d'accueillir les descriptions de diverses ressources selon un grain assez large : bases et banques de données, systèmes d'information, portails web et réseaux de sites, de personnes. Il rassemble à ce jour plus d'une centaine de métadonnées renseignées par les producteurs, grâce à une interface simple de saisie en ligne (figure 1, ci-dessous).

FIGURE 1 COPIES D'ÉCRAN DES RÉSULTATS D'UNE INTERROGATION DU CATALOGUE CA-SIF ET APERÇU DU DÉTAIL D'UNE MÉTADONNÉE

The screenshot displays the CaSIF search interface. At the top, there is a navigation bar with 'Recherche' highlighted. Below it, the search results for 'sol' are shown, with 11 results found. A detailed view of a metadata record is open, showing the following information:

Détail du document	
Présentation	Projet Biosol
Intitulé de la donnée	Dans le cadre du programme européen «Forest Focus», une dizaine de projets français ont été retenus en 2004. Le projet européen BioSol, partie prenante du programme Forest Focus, a pour but de réaliser un inventaire des caractéristiques chimiques des sols et de la biodiversité forestière au niveau des parcelles du dispositif européen de suivi des dommages forestiers. Vingt deux pays ont ainsi soumis leurs données à la Commission Européenne sur environ 4300 sites. Ces données sont corrélatées lors des trois étapes précédant l'intégration dans la base de données Biosol.
Description résumée	
Type de donnée	Données géographiques
Liens	http://www.fr.fr/risq/
Langue de la donnée	fr
Langue de la métadonnée	fr
Représentation spatiale	textTable
Thématiques	
Liste des thèmes ISO	environnement
Liste des thèmes INSPIRE	Habitats and biotopes Environmental monitoring facilities Soil Land use
Mots clés	Sylviculture
Mots clés	forêt

(8) <https://www.casif.gip-ecofor.org> (consulté le 16/12/2013).

L'alimentation des catalogues de métadonnées nécessite un important travail d'animation afin d'inciter les différents producteurs de données à renseigner les métadonnées correspondantes pour améliorer la visibilité et la valorisation potentielle de ces ressources. Cette animation est d'autant plus importante que les catalogues se multiplient aujourd'hui, induisant un risque de dispersion, et qu'il est important de ne pas solliciter plusieurs fois un même producteur de données. Les échanges de l'atelier ont montré que l'intérêt des catalogues de métadonnées, comme première étape de la connaissance de l'existant était bien compris. Des interrogations se posent cependant sur l'intérêt pour l'usager de l'information (assez rudimentaire) du niveau découverte de ces catalogues.

ACCÉDER À L'EXISTANT ET L'EXPLOITER

Les modes d'accès aux ressources

Parfaitement applicables au domaine des sols forestiers, les métadonnées et le catalogage en ligne permettent de progresser dans le porter à connaissance des différentes ressources. L'effort de description fourni par le producteur ou fournisseur de données constitue une véritable valeur ajoutée en créant un point d'entrée pour accéder aux données elles-mêmes.

L'accès aux données peut s'établir par :

- un contact direct avec le producteur ou fournisseur de données (numériques ou non) à partir du champ « contact pour les données » dans les métadonnées (exemple : RMQS, RENECOFOR). Le dialogue entre utilisateur et fournisseur permet d'identifier précisément les données pertinentes et leurs limites pour chaque projet d'utilisation. Les données sont échangées gratuitement, éventuellement sur la base d'une licence de mise à disposition ;
- un accès direct en ligne à la visualisation ou au téléchargement du jeu de données décrit dans un catalogue de métadonnées. Geoportail permet ainsi de rechercher des métadonnées dans le Geocatalogue, de visualiser les données décrites et de les télécharger (si accès public) ;
- un accès direct à une interface web d'interrogation spécifique multicritères (thématique, spatial, temporel...) d'une (ou plusieurs) base(s) de données. L'exécution des requêtes ainsi définies par l'utilisateur permet de visualiser ou télécharger les fichiers de données construits à la volée. L'outil Indiquasol⁽⁹⁾ permet, par exemple, de visualiser sur une carte des interrogations de la BD Indicateurs de la Qualité des Sols.

L'usage des métadonnées peut aussi faciliter l'accès aux données comme le montre le développement actuel des infrastructures de données spatiales (IDS). Ces infrastructures exploitent l'interopérabilité des métadonnées grâce au respect des normes existantes et partagent des informations géographiques distribuées indépendamment des technologies utilisées pour la gestion des données grâce à des protocoles de communication et d'échange standardisés. Les fonctionnalités principales sont de rechercher des données grâce au catalogage des métadonnées, de visualiser les données, de les interroger en croisant critères spatiaux et non spatiaux, de les extraire ou encore de les éditer. Les plateformes PIGMA⁽¹⁰⁾ et Geobretagne sont deux exemples d'IDS basées sur la solution libre Geochestra. Les bénéfices du partage de données que proposent ces outils reposent également sur la qualité des données qui les alimentent, une responsabilité de leurs fournisseurs.

La mutualisation et la mise à disposition des données sont aujourd'hui largement encouragées (et parfois obligatoires). Des outils en ligne tels que ceux mentionnés ci-dessus se développent. Les extractions de données réalisables de manière automatique concernent majoritairement des données élaborées ou semi-élaborées, plus rarement des données brutes.

(9) <http://www.gissol.fr/programme/bdiqs/bdiqs.php>

(10) <https://www.pigma.org/>

Les droits d'accès et conditions d'utilisation

Si l'accès aux métadonnées ne rencontre que rarement des restrictions, il n'en va pas de même pour l'accès aux données. Les propriétaires et gestionnaires de données doivent mener le plus tôt possible une réflexion sur leur politique de diffusion et de réutilisation des données. Les outils informatiques créés doivent intégrer des fonctionnalités de gestion des droits sur les données pour permettre de respecter la diversité des cas rencontrés. Cela va de la simple authentification d'un utilisateur pour des données publiques à l'application de droits spécifiques temporaires appliqués à des critères spatiaux, temporels et thématiques sur les données. Les facilités d'accès varient selon les opérateurs, l'accès est généralement facile pour les données d'état initial des stations, plus restreint pour les données coûteuses, et soumis à un délai de valorisation pour les données de la recherche.

Les instituts publics s'inscrivent aujourd'hui dans la réflexion nationale et internationale sur l'ouverture des données (« open data ») en établissant des règles qui visent à favoriser une politique de diffusion et d'accès aux données plus ouverte, tout en clarifiant les droits des producteurs (Gaspin *et al.*, 2012).

Comment faciliter l'exploitation des données ?

La diffusion des données au moyen d'outils informatiques, tels les systèmes d'information (SI), amène l'utilisateur autorisé à accéder directement aux jeux de données. Les informations fournies par les métadonnées concernent surtout le niveau « découverte » et ne permettent pas l'exploration et l'exploitation des données. Il est donc nécessaire d'accompagner l'accès aux données par un ensemble d'informations complémentaires utiles voire requises pour l'utilisation des données.

Ces informations vont par exemple fournir à l'utilisateur :

- une description de la station, du dispositif ou encore de l'expérimentation associée aux données ;
- une définition des différentes variables et paramètres fournis et les unités dans lesquelles sont fournies les données ;
- les traitements effectués sur les données : s'agit-il de données brutes ou sont-elles le produit d'un traitement ? Quel est-il ? Des informations complémentaires peuvent être associées comme le nombre de répétitions ou l'écart type d'une valeur moyennée ;
- des éléments relatifs à la qualité des données (protocoles, instruments, indicateurs de la qualité des données...).

Il est également utile de pouvoir éclairer les potentialités et limites d'utilisation par la mise à disposition de la documentation issue de valorisations précédentes. Pour faciliter l'exploitation des données, les efforts de gestion et de mise à disposition des données devraient donc intégrer l'identification, la gestion et la diffusion de ces informations complémentaires. Essentielles pour pérenniser l'exploitation des données, celles-ci sont bien souvent moins structurées et plus dispersées que les données elles-mêmes et nécessitent un effort particulier, d'autant plus quand l'historique des données est important. C'est donc une bonne pratique à promouvoir en amont de la production de nouvelles données.

Par ailleurs, un utilisateur peut souhaiter comparer ou exploiter ensemble des données de nature similaire mais de sources différentes et gérées au sein d'une même base de données ou dans des bases de données distinctes. Il doit pour cela disposer de suffisamment d'informations pour déterminer si cela est possible et comment. L'utilisation de référentiels est encouragée pour éviter les incertitudes sémantiques sur la nature des données. Donner un sens précis et unique à chaque terme, par exemple en utilisant un thésaurus, permet de contrôler le vocabulaire et d'établir une

certaine interopérabilité sémantique entre des données produites ou gérées différemment. Pour aller plus loin en matière d'interopérabilité, les ontologies, qui permettent de définir des sémantiques bien plus complètes qu'un thésaurus, pourraient constituer une perspective nouvelle pour l'interopérabilité des différentes bases de données, comme cela a été discuté lors du *workshop* « Sémantique et Ontologies » organisé dans le cadre de l'infrastructure nationale AnaEE France⁽¹¹⁾.

Cela étant, les producteurs de données ne peuvent envisager toutes leurs utilisations potentielles. L'utilisation pertinente des données nécessite donc d'encourager les utilisateurs à échanger avec les producteurs.

QUELLE COORDINATION DE LA PRODUCTION ET DE LA GESTION DES DONNÉES ? VERS UNE CENTRALISATION SYSTÉMATIQUE ?

Il existe une multiplicité d'objectifs et de contextes dans lesquels les données sur les sols forestiers peuvent être acquises. Une coordination poussée de la production et de la gestion des données, voire une centralisation sous un même système d'information, devrait *a priori* permettre d'accroître fortement l'intérêt des données et l'importance de leur valorisation tout en permettant des économies d'échelle. L'exemple de la cartographie des sols dans le cadre du GIS Sol (encadré 1, p. 567) et celui du réseau national de suivi des écosystèmes forestiers RENECOFOR (encadré 2, p. 568) montrent tous les deux l'intérêt d'un lien étroit entre production et gestion des données. Ils montrent aussi que la centralisation des données est pertinente si le système d'information répond aux objectifs et au fonctionnement des divers programmes de collectes concernés. Des approches moins intégrées visant à rendre possible l'interopérabilité entre les bases de données permettent de tenir compte de cette limite.

Coordonner la production

On entend ici par coordination de la production, l'entente entre les différents partenaires, en amont, sur la stratégie d'échantillonnage (exemple : choix des sites, densité de sondages pour la description des types de sol dans un pédopaysage, nombre de prélèvements sur une surface élémentaire pour assurer un certain niveau de représentativité), les méthodes analytiques, les modalités de description des fosses (exemple : utilisation de fiches STIPA) ou encore sur un but commun d'acquisition des données (exemple : le programme IGCS au 1/250 000). Il semble peu évident que toute production puisse ou doive être systématiquement coordonnée, la diffusion de standards d'acquisition et la communication autour des applications dérivées qu'ils autorisent semble importante. Par exemple, les données de granulométrie sont nécessaires au développement de nombreuses applications, telles que des fonctions de pédotransfert permettant d'estimer la réserve utile ou des outils de modélisation du bilan hydrique (exemple : Biljou©) utilisables par toute la communauté. Étant donné le faible coût de cette analyse supplémentaire (entre 10 et 20 € en fonction du nombre de classes demandées), une coordination en amont pourrait inciter à la rendre systématique. Dans le cas de la cartographie des sols français, avec l'exemple des programmes coordonnés par le GIS Sol, la centralisation de la gestion des données a été un moteur de l'harmonisation des programmes de collecte de données et a permis de mieux assurer la cohérence et la qualité de la donnée.

Coordonner la gestion des données

La coordination de la gestion des données consiste d'abord à un partage des modèles conceptuels de données et l'assurance d'une interopérabilité des différentes bases de données.

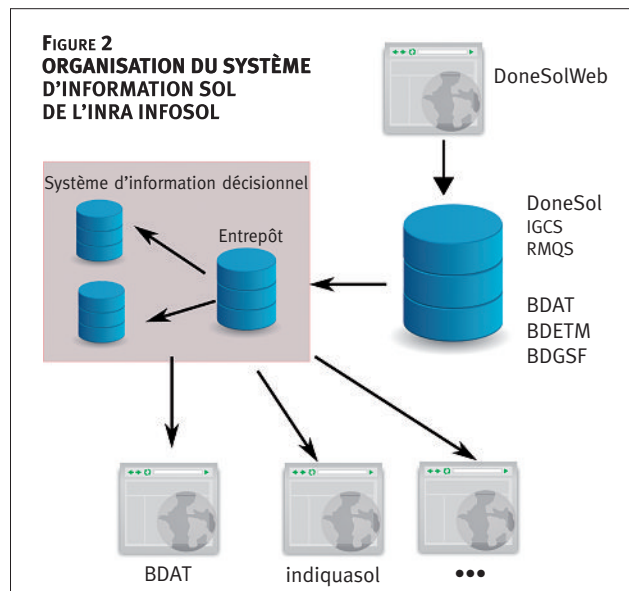
(11) <http://www.anaee-s.fr/spip.php?article56>

Encadré 1 – L'exemple du GIS Sol

L'objectif du GIS Sol est de constituer un système d'information sur les sols de France et sur l'évolution de leur qualité, sur la base de cinq grands programmes d'acquisition des données : le programme de la base de données géographiques des sols de France, le programme IGCS, le RMQS, la BD des éléments traces métalliques (BDETM) et la BDAT. Les données sont regroupées au sein du SI SOL géré par InfoSol.

La centralisation des programmes d'acquisition de données sol, et des données elles-mêmes, au sein d'une unique unité de service apporte une série de bénéfices : la possibilité d'en coordonner les acteurs ; le stockage des données dans un format national harmonisé ; la mise en place de contrôles qualité en continu au moyen d'une application SIVERCOH (SI pour la vérification de la cohérence des données) mobilisable sur la variété des jeux de données sol du SI et pouvant déboucher sur un processus d'attribution d'un niveau de qualité à certaines études (exemple : les études IGCS) ; la création, le maintien et l'évolution de la BD DoneSol et du dictionnaire des données associé ; l'organisation de séminaires et de formations à l'utilisation du SI et de ses outils (exemple : DoneSolWeb) ; le développement d'outils de consultation cartographique des données sur le web (bdat.gissol.fr, indiquasol.gissol.fr).

Des outils connexes ont été développés pour promouvoir et tracer l'utilisation des données sols. Citons les outils Repedo (répertoire national des organismes intervenant en pédologie, sciences du sol ou pollution des sols) et Refersols (répertoire national des études cartographiques de sol recensées par l'INRA dans sa base de données DoneSol). Le système d'information décisionnel, composante du SI Sol et mis en place récemment est d'ores et déjà utilisé comme support au traitement des données et au fonctionnement d'application web de consultation des données et devrait, à terme, faciliter la mise à disposition des données, le déploiement de services web et d'outils d'aide à la décision. L'ensemble des outils développés autour des bases de données liées aux différents programmes répondent à une diversité d'objectifs (depuis l'aide à la conduite de programmes tels que le RMQS, la mise en commun de données hétérogènes, la production d'outils de diffusion de la donnée et des métadonnées associées) tout en conservant les spécificités et en exploitant au mieux les caractéristiques propres de ces différentes bases.



Encadré 2 – L'exemple de RENECOFOR

RENECOFOR est un réseau national de placettes permanentes de suivi des écosystèmes forestiers, comprenant entre autres des données de description et de suivi des propriétés des sols (cf. Nicolas *et al.*, ce numéro). Il constitue la partie française du réseau de monitoring de niveau II du programme paneuropéen ICP Forests.

Face à la diversité des contextes écologiques et institutionnels des pays participants, la coordination du monitoring forestier paneuropéen se base sur un socle commun de contenu, de fréquence et de qualité des données (manuel ICP Forests et formats de données, exercices d'intercalibrations, circuits interlaboratoires...). Chaque pays participant organise lui-même sa collecte de données suivant ses contraintes pratiques et alimente annuellement la BD centrale ICP Forests *via* une plateforme web opérant des contrôles qualité sur les données entrantes. Les groupes d'experts internationaux mutualisent les compétences dans chaque domaine de suivi, animent la valorisation scientifique des données et discutent le cas échéant des modifications à apporter au manuel ICP Forests.

Au plan national, la BD du réseau RENECOFOR a été conçue en même temps que ses dispositifs de terrain et continue à être développée à mesure des besoins. À chaque protocole de collecte de données est associé un processus de saisie/transmission/intégration. Les formats de saisie sont adaptés à l'ergonomie de chaque type de mesure. L'intégration dans la base s'effectue au moyen de programmes permettant la réalisation de contrôles d'exhaustivité et de cohérence des données entrantes. La détection de lacunes ou d'anomalies conduit à la recherche d'informations supplémentaires ou de mesures correctives voire d'éventuelles contre-mesures dans un délai suffisamment court. La sécurité des données est assurée par des sauvegardes régulières à distance et l'ensemble des objets de la base est documenté (tables, champs, programmes). En outre, la BD contient les informations nécessaires à la gestion administrative du réseau, qui repose sur la coordination de centaines d'agents au sein de l'ONF et de nombreux partenaires externes. La BD RENECOFOR est donc non seulement un outil de sauvegarde mais aussi un maillon central de l'organisation du réseau. Sa gestion continue au plus près des besoins de fonctionnement est indispensable à l'efficacité et à la cohérence de la collecte de données.

Concernant la diffusion, les données sont mises à disposition gratuitement sur présentation d'un projet d'utilisation, aussi bien au plan national (BD RENECOFOR) qu'au plan paneuropéen (BD ICP Forests).

Poussée davantage, la coordination peut aboutir à une centralisation autour d'un même système d'information (SI), comme c'est le cas pour le GIS Sol. Cela favorise la mutualisation de différents développements nécessaires à la sauvegarde des données, au maintien de leur qualité et à la maintenance des fonctionnalités du SI (saisie, consultation et exploitation des données). Le fait de se limiter à un nombre restreint de SI permet des économies d'échelle sur certains composants clés, qu'il serait coûteux et redondant de redévelopper. Par exemple, les outils permettant d'assurer la qualité des données (automatisation des requêtes de test de la qualité des données lors de leur insertion dans le SI, requêtes de validation des données déjà insérées...) sont complexes à définir et à mettre en œuvre. Ils peuvent cependant présenter un degré de généricité les rendant applicables à différents jeux de données sol. Enfin, il est important de rappeler que la centralisation des données sous un même SI n'implique pas de perte de propriété et de gestion des accès.

Néanmoins, la centralisation est limitée par la capacité des SI à s'adapter aux objectifs des dispositifs au niveau desquels la collecte des données est réalisée en premier lieu. Ainsi, il semble difficile de prétendre intégrer et gérer tous types de données existantes et à venir dans un cadre conçu *a priori* sans risquer d'être réducteur. De plus, les données collectées sur le terrain doivent pouvoir être rapidement enregistrées, structurées et analysées pour éviter qu'elles ne soient perdues et pour être en mesure de contrôler la cohérence de la collecte voire de réviser les protocoles mis en œuvre. Un SI ne doit pas être seulement considéré comme un outil de sauvegarde et de partage de données mais aussi comme un outil de pilotage des dispositifs de collecte. Ainsi, la BD RENECOFOR alimente la BD ICP Forests tout en restant en gestion directe et rapidement adaptable aux évolutions des dispositifs de terrain.

CONCLUSIONS

L'atelier a permis de mettre en évidence l'importance de développer la connaissance des sols forestiers et d'améliorer les référentiels pour interpréter leur fonctionnement, notamment en termes de fertilité. La mobilisation et la valorisation des nombreuses ressources de données existantes sur les sols, qui restent encore pour la plupart sous-utilisées, est une démarche *a priori* intéressante sur le plan coût-efficacité.

Des outils informatiques sont développés pour faire connaître données et bases de données, et pour les rendre plus accessibles sur le réseau Internet : métadonnées et catalogues en ligne, outils de visualisation et de téléchargement de jeu de données, outils d'interrogation multicritères de bases de données. Face à la dispersion et à la faible harmonisation des jeux de données, un travail important reste cependant à faire : mise en base de données des données non numérisées et gestion active des bases produites, clarification des droits de propriétés et d'accès, production de métadonnées respectant les normes et suffisamment explicites pour les divers usagers, utilisation de référentiels sémantiques communs, procédures de contrôle-qualité, utilisation de protocoles d'échanges de données entre applications.

Les pistes d'amélioration concernent les données existantes, mais aussi celles qui seront produites demain. Il est donc important que les producteurs de données assurent la production des métadonnées et structurent l'information. Si les jeux de données ont leurs spécificités et sont souvent collectés pour répondre à des objectifs particuliers, des facilités de traitement et des économies d'échelle peuvent être obtenues par la coordination des méthodes de production et de gestion des données, voire par une centralisation sous un même système d'information. Ainsi, le GIS Sol, qui regroupe des données de qualité diverses, cherche à faciliter *ex-ante* la mutualisation des données et leur exploitation commune dans la cartographie des sols. Parmi les actions proposées au niveau de la collecte des données, citons des listes de paramètres « sols » standards à associer à chaque prélèvement, des guides partagés de protocoles expérimentaux et de méthodes d'échantillonnage ou des référentiels d'interprétation. La centralisation ne peut cependant pas être perçue comme une solution systématique car la gestion des données doit rester suffisamment souple pour s'adapter aux objectifs et aux fonctionnements des différents dispositifs de collecte.

Thomas EGLIN – Miriam BUITRAGO
ADEME (Agence de l'Environnement et de la Maîtrise
de l'Énergie)
Direction Productions et Énergies durables
Service Agriculture et Forêt
20 avenue du Grésillé
BP 90406
F-49004 ANGERS CEDEX 01
(thomas.eglin@ademe.fr) (miriam.buitrago@ademe.fr)

Manuel MARTIN
INRA Centre Val de Loire
US 1106, Infosol
2163 avenue de la Pomme de Pin
Ardon
F-45075 ORLÉANS CEDEX
(manuel.martin@orleans.inra.fr)

Damien MAURICE
 INRA-Université de Lorraine
 UMR 1137, Écologie et écophysio­logie forestières
 F-54280 CHAMPENOUX
 (damien.maurice@nancy.inra.fr)

Manuel NICOLAS
 OFFICE NATIONAL DES FORÊTS
 Département Recherche et Développement
 Boulevard de Constance
 F-77300 FONTAINEBLEAU
 (manuel.nicolas@onf.fr)

Céline PERRIER
 CNPF
 Institut pour le développement forestier
 47 rue de Chaillot
 F-75116 PARIS
 (celine.perrier@cnpf.fr)

Guy LANDMANN
 GIP ECOFOR
 42 rue Scheffer
 F-75116 PARIS
 (guy.landmann@gip-ecofor.org)

BIBLIOGRAPHIE

- FERRY (B.), EGLIN (T.), BISPO (A.), DAMBRINE (E.), CHENU (C.). — Matières organiques des sols forestiers et changements climatique et atmosphérique. — *Revue forestière française*, vol. LXVI, n° 4, 2014, pp. 425-436.
- FNE, ONF, FNCOFOR, PPF. — Forêt : Produire plus de bois tout en préservant mieux la biodiversité : Une démarche territoriale concertée dans le respect de la gestion multifonctionnelle des forêts. — 2007. Consulté le 16/12/2013 sur http://www.fne.asso.fr/_includes/mods/kb_upload/File/FORêt/Protocole%20accord%20Fne%20Forestiers%202007%20-%20version%20signee.pdf.
- GASPIN (C.), PONTIER (D.), COLINET (L.), DARDEL (F.), FRANC (A.), HOLOGNE (O.), LE GALL (O.), MAURIN (N.), PERRIÈRE (G.), PICHOT (C.), RODOLPHE (F.). — Rapport du groupe de travail sur la gestion et le partage des données. — 2012. [En ligne] : <http://prodirna.inra.fr/ft/E261B98C-86C6-42D8-AD89-622FAB942C55> (Rapport au conseil scientifique de l'INRA).
- GÉGOUT (J.-C.), COUDUN (C.), BAILLY (G.), JABIOL (B.). — EcoPlant: A forest site database linking floristic data with soil and climate variables. — *Journal of Vegetation Science*, vol. 16, issue 2, 2005, pp. 257-260.
- LEGOUT (A.), HANSSON (K.), VAN DER HEIJDEN (G.), LACLAU (J.-P.), AUGUSTO (L.), RANGER (J.). — Fertilité chimique des sols forestiers : concepts de base. — *Revue forestière française*, vol. LXVI, n° 4, 2014, pp. 413-424.
- NICOLAS (M.), JOLIVET (C.), JONARD (M.). — L'Apport des dispositifs de suivi vis-à-vis des enjeux de fonctionnement et de gestion des écosystèmes en relation avec les sols. — *Revue forestière française*, vol. LXVI, n° 4, 2014, pp. 491-500.
- PONETTE (Q.), BISPO (A.), BRÊTHES (A.), DUPOUEY (J.-L.), FEIX (I.), GÉGOUT (J.-C.), JABIOL (B.), POUSSE (N.), RANGER (J.). — Indicateurs de fertilité des sols forestiers tempérés : enjeux, approches et perspectives. — *Revue forestière française*, vol. LXVI, n° 4, 2014, pp. 501-515.
- UNION EUROPÉENNE. — Directive 2007/2/CE du parlement européen et du conseil du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne (INSPIRE) — 2007.
- UNION EUROPÉENNE. — Décision (UE) n° 529/2013 du 21/05/13 relative aux règles comptables concernant les émissions et les absorptions de gaz à effet de serre résultant des activités liées à l'utilisation des terres, au changement d'affectation des terres et à la foresterie et aux informations concernant les actions liées à ces activités. — JOUE L 165 du 18 juin 2013.

MOBILISER ET VALORISER LES DONNÉES SUR LES SOLS FORESTIERS. Compte rendu de l'atelier 4 (Résumé)

Une meilleure connaissance des sols forestiers et l'amélioration des référentiels pour interpréter leur fonctionnement est une attente des forestiers comme des acteurs des politiques publiques. Lors de REGEFOR 2013, l'atelier "Mobiliser et valoriser les données sur les sols forestiers" a identifié une ressource importante et sous-exploitée de données au niveau des organismes de développement et des laboratoires de recherche. Des outils existent pour les rendre plus visibles et accessibles sur le réseau internet. Face à la dispersion et à la faible harmonisation des jeux de données, des chantiers importants sont à mettre en œuvre pour favoriser leur exploitation (exemple : mise en base de données, production de métadonnées et de référentiels sémantiques communs, clarification des droits de propriété, protocoles d'échanges de données entre applications). Dans cette optique, l'atelier a aussi conduit à s'interroger sur les possibilités d'améliorer la coordination entre production et gestion de données, notamment en discutant les intérêts et limites d'une centralisation sous un même système d'information.

COLLECTING AND USING DATA ON FOREST SOILS. Report of the workshop 4 (Abstract)

Forest managers and decision-makers require more knowledge of forest soils and better references for understanding soil functioning. At REGEFOR 2013, the "Collecting and using data on forest soils" workshop identified a large under-exploited source of data in development organizations and research laboratories. Software was available for making these resources easier to discover and accessible *via* Internet. As the data is scattered and the data sets are not harmonized, much remains to be done to enable this data to be used (eg: incorporation into databases, production of metadata and common semantic standards, clarification of ownership rights, data transfer protocols between applications). The workshop also discussed ways of improving coordination between the production and management of data, in particular considering the advantages and limitations of centralization within a single information system.
