

Contexte et défis

La modélisation et la prédiction sont, de plus en plus, au cœur des recherches relatives aux systèmes biologiques et écologiques, que celles-ci soient orientées vers leur compréhension ou vers leur ingénierie et leur pilotage. Transversales aux cinq premières priorités scientifiques, les approches prédictives concernent un vaste spectre de domaines, tels que la biologie des systèmes et la biologie intégrative, les biotechnologies et la biologie synthétique, la conception et le contrôle de nouveaux procédés ou produits, le fonctionnement et le pilotage des écosystèmes, le contrôle des épidémies ou le suivi de la distribution spatiale d'espèces animales, végétales et microbiennes, la conception et l'évaluation de systèmes de production et de gestion de ressources renouvelables.

Simultanément porté par le formidable accroissement des capacités analytiques et des masses de données ainsi générées et par les opportunités offertes par les sciences et technologies de l'information et de la communication, l'essor des approches prédictives est motivé (i) par la généralisation de démarches cognitives couplant analyse, modélisation, observation et expérimentation, (ii) par la reconnaissance de la complexité des systèmes biologiques et écologiques, aussi bien que (iii) par le besoin de réduire les incertitudes liées au fonctionnement, au devenir, à la conception et au pilotage de ces systèmes.

Quelle que soit la question scientifique posée ou l'application visée, prédire suppose **un couplage plus étroit entre (i) modéliser** - de façon plus ou moins phénoménologique ou mécaniste, selon les systèmes et les niveaux d'organisation étudiés -, **(ii) gérer, intégrer et analyser des données, et (iii) expérimenter ou observer en produisant les données pertinentes**. Le premier défi concerne donc la maîtrise d'ensemble de cette chaîne dont chacune des composantes peut être critique pour la qualité des prédictions.

Le deuxième défi, à la fois scientifique et culturel, est corollaire du précédent. Il porte sur **l'interdisciplinarité**, et sur l'articulation de compétences (i) biologiques ou écologiques, relatives aux objets et phénomènes étudiés, et (ii) méthodologiques, relevant des mathématiques appliquées, de l'informatique ou des sciences de l'ingénieur. La mise en œuvre d'approches prédictives repose non seulement sur la mobilisation, interne ou en partenariat, de ces compétences, mais aussi sur diverses formes d'ingénierie (le développement et la diffusion de méthodes ; la maîtrise de technologies permettant de mesurer, modifier, concevoir ou contrôler les systèmes étudiés).

Questions de recherche et démarches

Au cours des dernières années, l'INRA a pris diverses initiatives pour accroître ses capacités de modélisation et de prédiction : lancement d'actions incitatives en biologie intégrative, comme pour la production de ressources biologiques ; création de plates-formes d'acquisition, de gestion et d'analyse de données, aussi bien que de modélisation ; renforcement des compétences en bioinformatique et des partenariats avec l'INRIA. Amplifier cette dynamique nécessite aujourd'hui de continuer à jouer sur différents registres qui relèvent de la gestion des compétences, de l'organisation du dispositif et du partenariat, et de la programmation scientifique :



- 1- **Poursuivre le développement et la mise en réseau d'infrastructures** pour accroître la fiabilité, la reproductibilité et le débit d'acquisition des données, et pour renforcer la capacité à les gérer, les intégrer, les partager et les analyser. L'état des lieux par rapport à cette question est contrasté selon les champs disciplinaires et les niveaux d'organisation considérés. Des verrous particuliers concernent : la maîtrise des (ou l'accès aux) technologies d'imagerie du vivant et d'étude à très haut débit de l'expression des génomes ; le phénotypage (diversité, exhaustivité et débit des mesures), notamment en milieu non contrôlé ; l'instrumentation des dispositifs d'observation de l'environnement ; la constitution de bases de données et de connaissances, ouvertes et structurées autour d'ontologies ; la création ou l'extension de plates-formes de modélisation et de simulation et des formalismes sous-jacents ; le dimensionnement des infrastructures et moyens de calcul, ou l'accès à des infrastructures externes *ad hoc* ;
- 2- **Soutenir des développements méthodologiques**, au sein de l'Institut ou en partenariat, notamment dans les domaines suivants : l'amélioration des méthodes de modélisation statistique qui offrent déjà un bon pouvoir prédictif dans différents domaines ; l'analyse, la visualisation, la modélisation et le contrôle de fonctions et phénomènes complexes mettant en jeu plusieurs niveaux d'organisation et de multiples échelles, spatiales et temporelles (par exemple, de réseaux d'interaction ou de systèmes dynamiques à structure dynamique rencontrés dans les phénomènes morphogénétiques) ; l'extraction automatique de connaissances et les méthodes de méta-analyse ;
- 3- **Favoriser des approches intégrées de modélisation** au travers de quelques programmes majeurs qui couvrent la diversité des systèmes et phénomènes étudiés à l'INRA (intégration de mécanismes cellulaires impliqués dans la réalisation d'une fonction biologique ; sélection génomique et prédiction des phénotypes ; métagénomique d'écosystèmes microbiens ; prédiction de l'évolution d'aires de répartition dans un contexte de changement climatique ; couplages entre processus écologiques et activités humaines dans les agro-écosystèmes et les paysages ; ...)
- 4- **Renforcer les compétences en bioinformatique et en modélisation** (par la formation des jeunes chercheurs, la formation continue et les recrutements) **et favoriser des dispositifs appropriés de mise en contact entre disciplines** (équipes, unités ou sites pluridisciplinaires ; partenariat avec d'autres organismes ou établissements, en France comme à l'étranger) ;
- 5- De manière transversale, l'accroissement des capacités de modélisation, de prédiction, voire de modification et de pilotage des systèmes biologiques et écologiques est porteur de mutations fortes dans les relations avec des partenaires techniques et socio-économiques comme dans la perception externe, par la société, des activités scientifiques. Conduire une approche réflexive sur ces changements est aussi, en soi, une question de recherche au cœur de laquelle les sciences sociales et économiques prendront toute leur place.

8- Cette note a été rédigée sur la base de la consultation interne organisée sur le thème « Vers une biologie et une écologie plus prédictives ».

9- A différents égards, les enjeux, défis et verrous évoqués ci-dessus concernent aussi les systèmes techniques et des systèmes socio-économiques étudiés par l'INRA.