

Agroécologie et intelligence artificielle

Christine Dillmann

Elodie Marchadier, Adrienne Ressayre, Renaud Rinent, Maela Semery



UMR Génétique Quantitative et Evolution - Le Moulon



université
PARIS-SACLAY

FACULTÉ
DES SCIENCES
D'ORSAY



AgroParisTech
Talents d'une planète soutenable

April 4, 2023

- 1 Agroécologie
- 2 Robotique
- 3 Analyse d'images
- 4 IA faible
- 5 Graphes de connaissance
- 6 Apprentissage supervisé
- 7 Un peu d'humanité

Changements sociétaux : de la révolution verte à l'agroécologie



1970's : Révolution verte. Sélection de variétés adaptées à des itinéraires techniques à forts intrants.



2010's : Services écosystémiques, agriculture écologiquement intensive.



2020's : Application des méthodes de l'écologie aux écosystèmes anthropisés gérés par l'agriculture et l'élevage.

- S'appuyer sur les régulations naturelles de l'agro-écosystème pour assurer la production agricole sans gaspiller les ressources.
- Modifier l'insertion de l'agriculture dans le système économique et social (services éco-systémiques, services culturels).

Des initiatives nationales

→ Stratégie Nationale Bas Carbone

Réduction de 40% des émissions GES d'ici 2030. Neutralité carbone en 2050.



→ Initiative 4 pour 1000

Stocker plus de carbone dans les sols agricoles



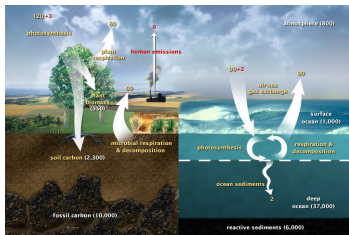
→ PEPR Agroécologie et Numérique

Données, agroéquipements et ressources génétiques au service de la transition agro-écologique et de l'adaptation aux aléas climatiques.



Usage des terres et émissions carbone

Flux de carbone sur la terre



Un bilan net de 4 gigatonnes de CO₂/an supplémentaires émis dans l'atmosphère du fait des activités humaines.

<https://earthobservatory.nasa.gov/features/CarbonCycle>

Coût de changement d'usage des terres



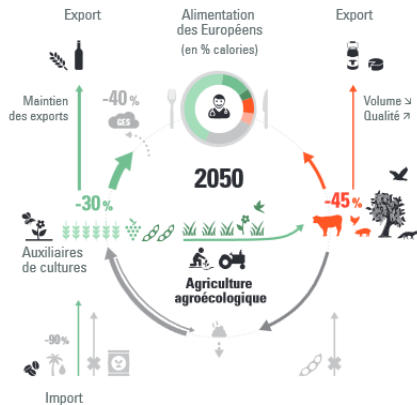
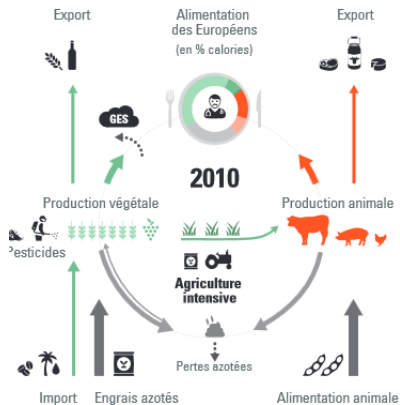
• Émissions de CO₂ par type d'espace et lors des changements d'affectation, en millions de tonnes équivalent CO₂ (valeurs 2013 - Source CITEPA 2015)¹⁶. (cf. tableau détaillé en page 8)

La transition agricole peut contribuer significativement aux réductions d'émissions de gaz à effet de serre.

<https://afterres2050.solagro.org/>

Il est possible de quantifier l'impact des activités humaines sur les émissions de gaz à effet de serre (GES) transcrites en unité de tonnes équivalent carbone.

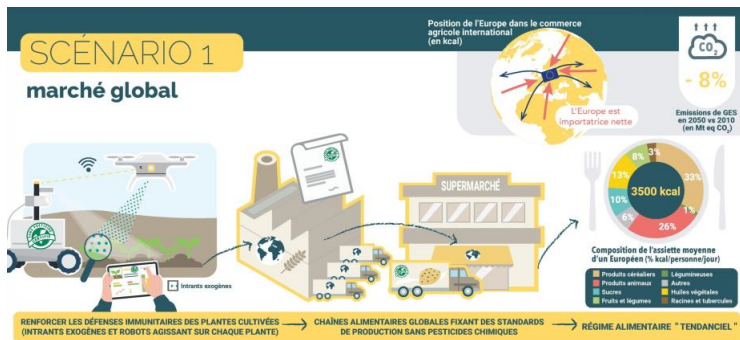
Scénarios pour la transition agro-écologique



<https://www.iddri.org/fr/projet/reussir-la-transition-agro-ecologique-en-europe>

Des scénarios existent qui impliquent des changements drastiques dans une fenêtre temporelle courte.

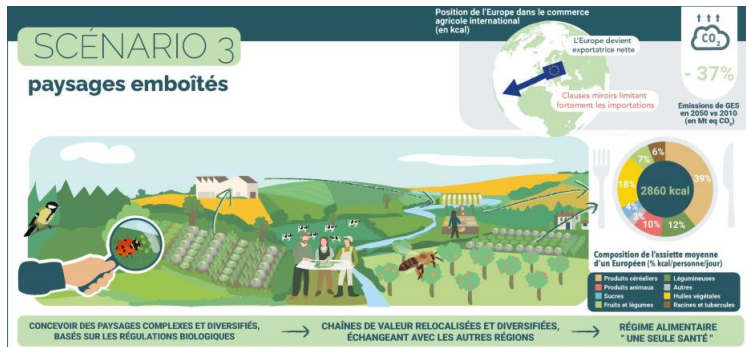
Trajectoire INRAE zéro-pesticides chimiques



Mora O. (coord), Berne J.A., Drouet J.L. et al. (2023). Agriculture européenne sans pesticides chimiques en 2050. Résumé de la prospective - INRAE, France

Tous les scénarios font appel à l'intelligence artificielle

Trajectoire INRAE zéro-pesticides chimiques



Mora O. (coord), Berne J.A., Drouet J.L. et al. (2023). Agriculture européenne sans pesticides chimiques en 2050. Résumé de la prospective - INRAE, France

Le paysage agricole se complexifie.

Les leviers de la transition pour l'agriculture

- **Boucler le cycle des nutriments**

- ▶ Réduire les pertes d'azote au cours du cycle et remplacer (*en partie*) l'azote de synthèse par l'azote d'origine symbiotique.
- ▶ Protéger les sols de l'érosion et recycler intégralement le phosphore.

- **Diversifier les cultures dans le temps (rotations, espèces)**

- ▶ Diminuer la pression des adventices et des parasites
- ▶ diminuer les besoins en engrais par le recyclage des éléments minéraux

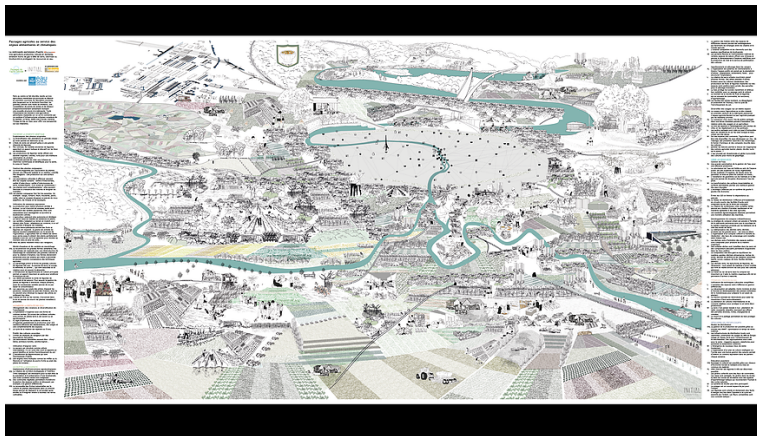
- **Associer cultures et variétés dans une parcelle (mélanges, ...)**

- ▶ Optimiser la photosynthèse
- ▶ Favoriser les mécanismes de régulation naturels

- **Mobiliser les facteurs de production vivants**

- ▶ Privilégier la lutte biologique via des Infrastructures AgroEcologiques (IAE, haies, bandes fleuries, ...)
- ▶ Cultiver des légumineuses et maintenir une couverture du sol et des systèmes racinaires profonds.

Vers de nouveaux paysages



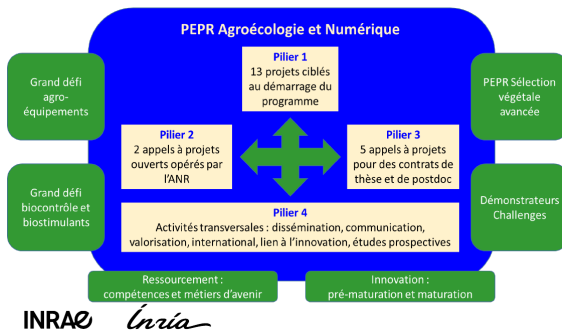
<http://www.paysages-apres-petrole.org/campagne-des-paysages-dafterres-2050/>

Un autre futur est possible !

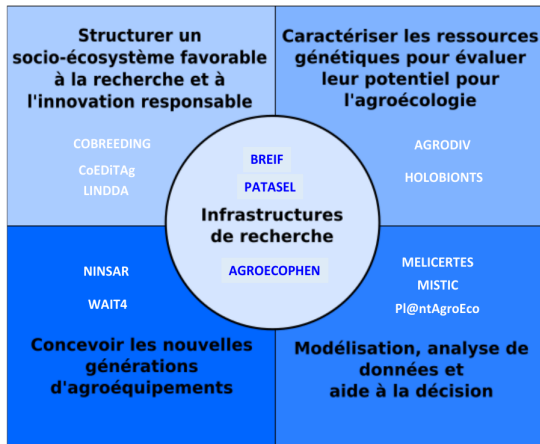
De nouvelles questions scientifiques impliquant des développements en IA

- **Suivi plus précis de cultures hétérogènes au cours du temps**
 - ▶ Evaluation des risques (ex. maladies) et décisions d'intervention
→ analyse et interprétation d'images
 - ▶ Interventions mécaniques complexes
→ robots multi-agent reconfigurables
- **Prédiction des performances en milieu hétérogène**
 - ▶ Evaluation multicritère
 - ▶ Prise en compte de la temporalité et du développement
 - ▶ Prédiction des interactions
- **Choix des meilleures associations**
 - ▶ Définition des critères
 - ▶ Elicitation de choix avec les utilisateurs
- **Biologie translationnelle**
 - ▶ Graphes de connaissance
 - ▶ Graphes génomiques

Systèmes agricoles durables et équipements agricoles contribuant à la transition écologique

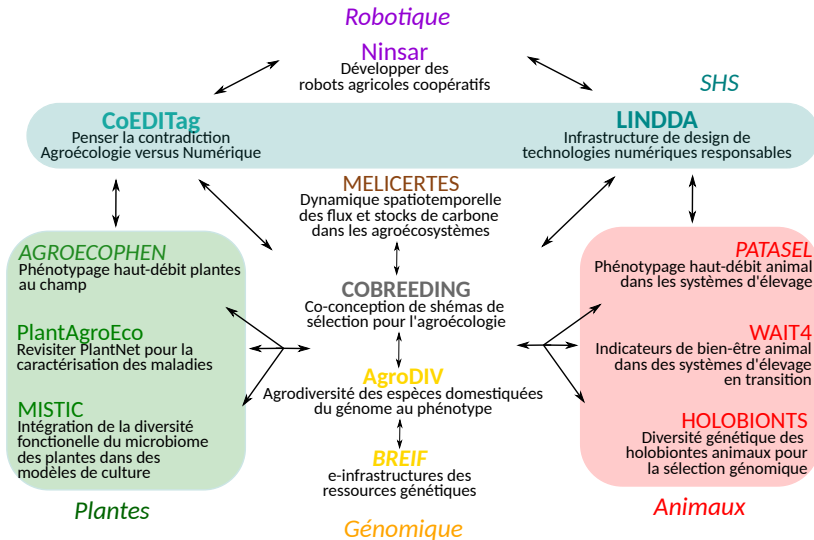


Des objectifs de résultat à court terme



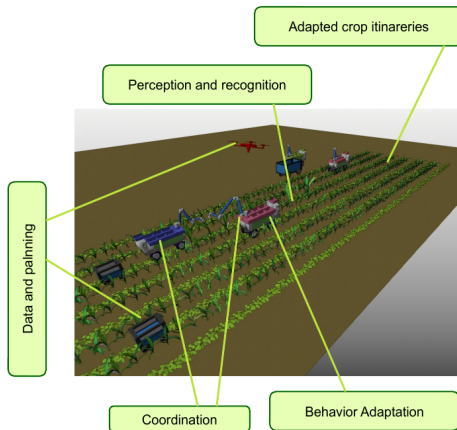
Principe de base : sobriété numérique

10 programmes cibles



- 1 Agroécologie
- 2 Robotique**
- 3 Analyse d'images
- 4 IA faible
- 5 Graphes de connaissance
- 6 Apprentissage supervisé
- 7 Un peu d'humanité

NINSAR



Des robots multi-agent coopératifs pour soulager le travail mécanique.

- 1 Agroécologie
- 2 Robotique
- 3 Analyse d'images**
- 4 IA faible
- 5 Graphes de connaissance
- 6 Apprentissage supervisé
- 7 Un peu d'humanité



Deux niveaux d'intervention

● High-Tech

- ▶ En station de *recherche*
- ▶ Produire des données quantitatives sur la croissance des plantes au cours du temps dans des couverts hétérogènes (mélange d'espèces).
- ▶ A partir de photos prises par drones équipés de caméras dans le visible ou de caméras hyperspectrales.
- ▶ Préalable = reconnaissance des espèces et traitement séparé par espèce.

● Low-Tech

- ▶ A la ferme
- ▶ Produire des données qualitatives sur l'environnement de la culture ou sur la reconnaissance des maladies.
- ▶ A partir de photos prises par des smartphone sur un mode participatif.
- ▶ **PlantAgroeco** Caractériser les maladies des plantes (en revisitant PlanNet)
- ▶ **Cobreeding** Caractériser l'environnement d'une exploitation agricole à partir de la flore.

Exemple : détection de dégâts sur des parcelles agricoles

A. Ressayre, coll. MIA PAris

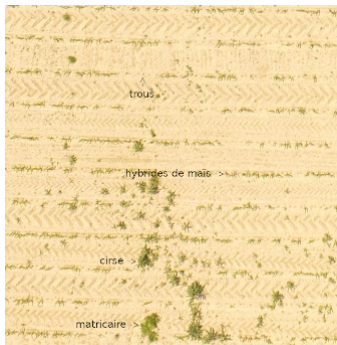


Avoir une estimation, sur une parcelle agricole, des dégâts de corvidés (plants arrachés) et de l'importance de la présence de chardons ou autre adventice (surface occupée).



Exemple : détection de dégâts sur des parcelles agricoles

A. Ressayre, coll. MIA Paris



Repérer les trous, les plants de maïs, les chardons, les matricaires, éviter les filets.... Quantifier.

- 1 Agroécologie
- 2 Robotique
- 3 Analyse d'images
- 4 IA faible**
- 5 Graphes de connaissance
- 6 Apprentissage supervisé
- 7 Un peu d'humanité

Classification non supervisée

- Faire émerger une structure dans les données
- **IA** → K-means, SOTA, forêts aléatoires,
- **STATS** → Modèles de mélange

Exemple : Tests de consommation

Sanane et al, 2021, J. Chemical Ecology

Ostrinia nubilalis

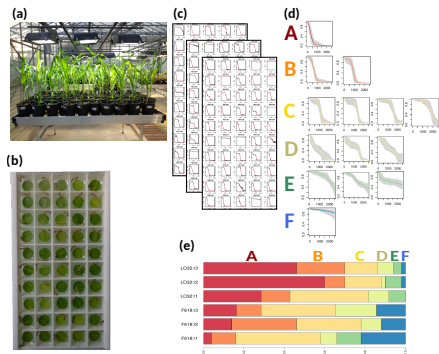


European stem borer.
1 or 2 generations/year
larvae feed on maize stems

Les pyrales ont-elles des préférences pour certaines variétés de maïs ?
Peut-on mesurer cette préférence à travers un test de consommation des larves sur des échantillons de feuilles ?

Classification automatique de courbes de consommation

Sanane et al, 2021, J. Chemical Ecology



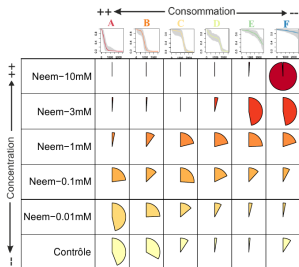
- Un réplicat = 50 disques foliaires d'un même traitement.
- Mesure de la consommation du disque foliaire par une larve de pyrale (analyse d'image).
- Algorithme SOTA pour faire émerger des types de comportement.
- Chaque échantillon est caractérisé par la proportion de chaque type de comportement.
- Classiquement, la variable mesurée était le temps moyen mis par une larve pour consommer 50% du disque foliaire.

La classification des courbes de consommation permet de prendre en compte la diversité des comportements individuels des larves.

Typologie versus quantification

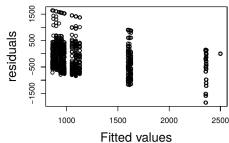
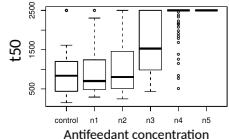
Exemple Mesure d'appétence pour des échantillons de feuilles traités à différentes concentrations d'une substance aux propriétés antiappétentes reconnues (Neemazal).

Analyse de la distribution des types de comportements



Possibilité de prendre en compte la variabilité des types de comportement et mettre en évidence des différences significatives entre traitements.

Analyse quantitative du t50



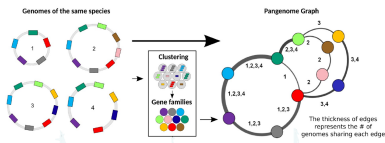
L'analyse statistique sur la variable quantitative t50 n'est pas possible car la variance résiduelle change selon le traitement.

Passer intelligemment du quantitatif au qualitatif peut augmenter le pouvoir de résolution des analyses.

Exemple : graphes pangénomiques

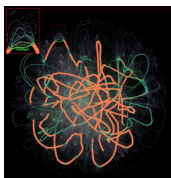
Gautreau et al, 2020, Plos Comp. Biol.

Les individus d'une même espèce biologique ne partagent pas les mêmes gènes



Le pangénome est l'addition des gènes présents au sein d'une espèce

Comment représenter graphiquement un pangénome ?



PPanGGOLiN utilise des informations génomiques (familles de gènes) et des informations positionnelles (ordre des gènes dans le génome).

Introduire des connaissances bien structurées facilite les découvertes.

- 1 Agroécologie
- 2 Robotique
- 3 Analyse d'images
- 4 IA faible
- 5 Graphes de connaissance**
- 6 Apprentissage supervisé
- 7 Un peu d'humanité

Enjeux agricoles de la Biologie translationnelle

- L'agroécologie mobilise de nouvelles espèces, parfois peu étudiées
- Les informations sur chaque espèce sont de différente nature
- Le plus souvent, des données de séquençage de l'ADN sont disponibles
- Est-il possible de propager les informations d'une espèce à l'autre en utilisant des graphes de connaissance ?

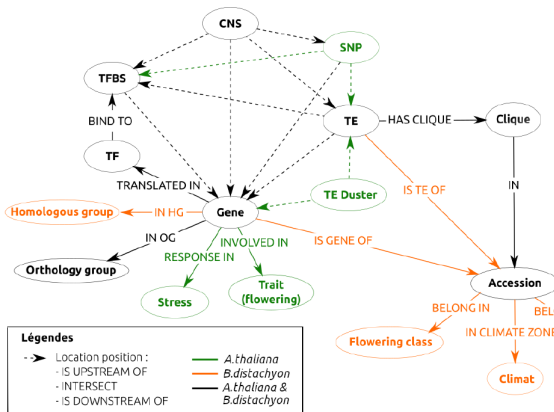
Programmes cibles **AgroDIV**, **BREIF**, du PEPR Agroécologie et numérique.

Graphe de connaissances génomiques

M. Sémary, J. Confais, N. Francillonne, URG1



Brachypodium distachyon est une espèce modèle des monocotylédones (blé, maïs) et peut-être comparée aux données disponibles sur *arabidopsis thaliana*, espèce modèle des dicotylédones (colza, chou). Les informations sur les deux espèces sont de nature et de qualité différente.



Graphe de connaissance. Les noeuds principaux sont des séquences d'ADN annotées en différentes classes (gènes, facteur de transcription, etc...)

La base de donnée orientée graphe permet, par exemple, de rechercher les séquences régulatrices ancestrales impliquées dans le déterminisme de la date de floraison.

- 1 Agroécologie
- 2 Robotique
- 3 Analyse d'images
- 4 IA faible
- 5 Graphes de connaissance
- 6 Apprentissage supervisé**
- 7 Un peu d'humanité

Enjeux agricoles de l'apprentissage supervisé

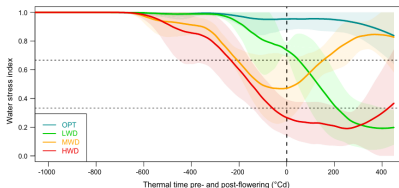
Dans des systèmes agricoles plus variés, mobilisant la diversité génétique à toutes les échelles, il s'agit de prédire des performances évaluées sur des critères multiples (à pondérer) de combinaisons (génomes, variétés, espèces) dans des environnements qui changent chaque année.

- Prédiction des interactions génotype x environnement
- Prédiction de la performance des associations
- Choix variétaux sur des critères multiples

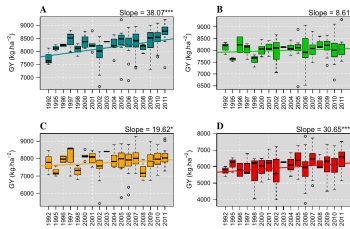
Interactions génotype x environnement

Touzy et al, 2019, Theoretical and Applied Genetics

210 variétés élite de blé mises sur le marché à différentes dates évaluées dans 35 environnements différents.



Utilisation d'un modèle de culture pour prédire la demande en eau selon les caractéristiques des environnements et classification des lieux en quatre types selon la nature du stress hydrique.



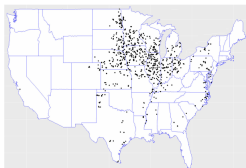
Le progrès génétique est important en situation optimale et en situation de fort stress hydrique. Il est pratiquement nul dans les situations de stress hydrique intervenant après la floraison.

Le classement des variétés de blé et les différences de rendement ne sont pas les mêmes selon les environnements de culture.

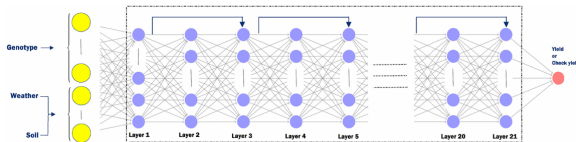
Syngenta Crop challenge

Khaki S, Wang L., 2019, *Frontiers in Plant Sciences*

Objectif : Confronter des méthodes d'apprentissage neuronal profond (DNN) à des méthodes statistiques pour prédire les performances de variétés d'hybrides de maïs dans différents lieux aux Etats-Unis, connaissant les performances des années précédentes.

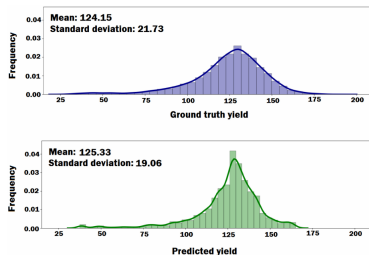


2267 hybrides de maïs évalués dans 2247 environnements entre 2001 et 2015. L'objectif est de prédire les rendements de l'année 2016 connaissant la séquence climatique, les caractéristiques du sol et la variété de maïs, identifiée par son génome.



Syngenta Crop challenge : victoire de l'IA

Khaki S, Wang L., 2019, *Frontiers in Plant Sciences*



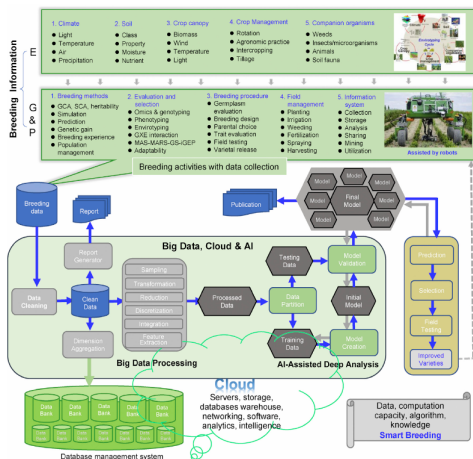
Distribution des rendements observée (top) et prédits par l'approche DNN (bottom)

Les approches d'apprentissage neuronal profond sont plus performantes que les approches statistiques classiques type *LASSO* ou *ElasticNet*.

Les méthodes d'apprentissages profond captent la non-linéarité des réponses.

Smart Breeding

Xu et al, 2022, Molecular Plant



Il y en a qui rêvent ...

Prédiction des performances hybrides (en mélange)

Cobreeding

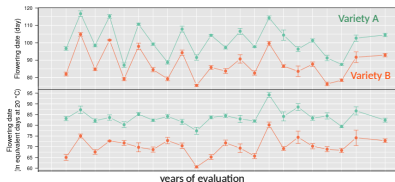
P1						
P2						
P3						
P4						
P5						
P6						

Peut-on prédire la performance d'hybrides connaissant les parents sur la base de données incomplètes de la valeur des hybrides ?

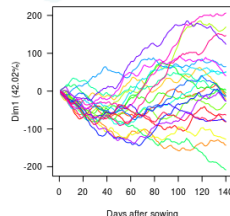
- 1 Agroécologie
- 2 Robotique
- 3 Analyse d'images
- 4 IA faible
- 5 Graphes de connaissance
- 6 Apprentissage supervisé
- 7 Un peu d'humanité**

Warmrules, DATAIA, F. Sahis, E. Marchadier

Quelles variables climatiques, à quel moment, déterminent la date de floraison du maïs?

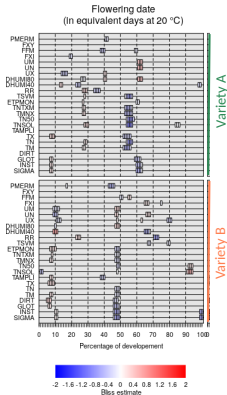


Variation des dates de floraison de deux variétés de maïs mesurée dans un même lieu durant 20 années. Des effets années mais aussi des interactions géotype x année.

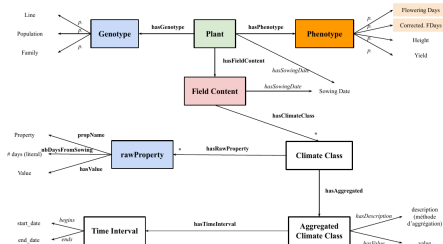


Chaque année est caractérisée par une séquence climatique particulière au cours de la saison.

Warmrules : statistiques versus IA



Approche Statistique Des
évènements climatiques particuliers
à des stades précis du
développement du maïs
déterminent la date de floraison



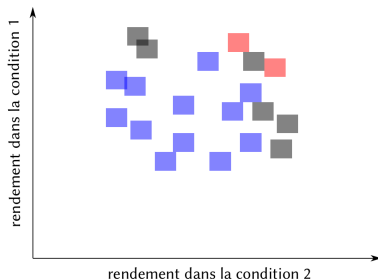
$(Pid1 \wedge Pid2 \wedge \dots) \wedge (Cvar1 \wedge Cvar2 \wedge Gvar3 \dots) \Rightarrow (Pvar1 \wedge Pvar2 \wedge \dots)$
 contextual identity cause phenotype
 climatic variables flowering time
 genotypic variables plant height

Approche IA Détermination de règles de logique formelle se propageant dans le graphe de connaissances.

Elicitation de préférences

E. Kuhn, J.B. Leger, R. Rincent, ANR Stat4Plants

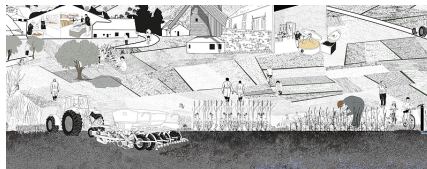
En sélection, choisir les meilleurs candidats résulte d'un compromis entre de nombreux critères, pas toujours objectifiables. Dans un espace multidimensionnel, la notion de front de Pareto dépend de relations de dominations. Les données sont connues avec incertitude.



Utiliser des méthodes d'élicitation de préférences pour choisir les candidats à conserver sur la base d'objectifs multicritères.

Conclusions

- Une partie des questions de recherche soulevées par la transition agro-écologique pourrait être résolue par l'intelligence artificielle.
- Les plus attendus concernent surtout l'apprentissage supervisé, dont l'analyse d'images.
- Pour l'apprentissage supervisé, l'enjeu est dans la collaboration avec les biologistes pour l'élaboration de modèles de connaissance efficaces.



La réforme des retraites met la terre en danger

1. Le GIEC (<https://www.ecologie.gouv.fr/publication-du-6e-rapport-synthese-du-giec>) alerte sur le fait que la fenêtre temporelle d'action pour limiter l'impact des changements climatiques se rétrécit chaque jour. Les rapports du groupe 3 du GIEC préconisent sans relâche la diminution des inégalités, la redistribution des richesses et la cohésion sociale comme des leviers pour mobiliser des changements rapides. Ces éléments d'information figurent dans les synthèses du GIEC pour les dirigeants. Pourtant, les syndicats ont démontré que cette réforme des retraites était génératrice d'inégalités sociales : inégalités hommes-femmes, inégalités riches-pauvres, etc...

=> Notre responsabilité, en tant que scientifiques, est de dénoncer l'incohérence de la politique du gouvernement vis à vis des enjeux environnementaux.

2. Les personnels de l'Enseignement supérieur et la Recherche sont également visés par la réforme : pour les ITA, des carrières trop longues, pour les chercheur-es et enseignants chercheur-es, des carrières courtes et morcelées. Pour de nombreux personnels de l'ESR, le départ en retraite sera à 67 ans avec une retraite incomplète [En 2021, l'espérance de vie sans incapacité était de 67,0 ans pour les femmes et 65,6 ans pour les hommes, <https://drees.solidarites-sante.gouv.fr/recherche?text=1258>].

=> La spécificité de nos métiers doit être reconnue.

3. Les projections du gouvernement qui justifient cette réforme se basent sur une volonté délibérée de diminuer la proportion des fonctionnaire parmi les actifs (<https://nosservicespublics.fr/>), soit un scénario inquiétant de stagnation des emplois de la fonction publique. Depuis de nombreuses années, tous les secteurs de la fonction publique ont vu leurs effectifs diminuer en deçà de seuils critiques (justice, police, hôpitaux, enseignement, recherche, ...). Pourtant, les services publics sont générateurs de cohésion sociale.

=> La Loi de Programmation de la Recherche ne prévoit pas d'augmentation du nombre d'emplois permanents. L'ESR subit de plein fouet la stagnation des emplois dans la fonction publique.

4. Face à l'urgence climatique, l'augmentation du niveau d'éducation des populations est un levier reconnu. La stagnation du nombre d'emplois publics, le vieillissement programmé de l'âge moyen des enseignants et des chercheurs et le nombre croissant d'étudiants sont des facteurs de risque, de même que la dégradation de l'enseignement public et sa privatisation [<https://www.cafepedagogique.net/2023/01/30/la-cour-des-comptes-relance-la-privatisation-de-lecole/>].

=> Les offres de formations privées vivent grâce à des subventions de l'Etat. On nous parle d'innovation plus que de recherche. Existe-t-il vraiment des exemples où des entreprises à but lucratif font augmenter le niveau général d'éducation d'une population et sont facteurs de réduction des inégalités et de promotion de la justice sociale ?

