Modélisation mathématique et simulation numérique de la croissance des plantes.

Véronique Letort, et al.

Digiplante (Ecole Centrale Paris) / Cybeletech

Forum Teratec - 2 juillet 2014

letort@mas.ecp.fr

Digiplante - CybeleTech

Digiplante : modélisation et méthodes mathématiques pour la croissance des plantes

- Equipe créée en 2002 au laboratoire de Mathématiques Appliquées aux Systèmes (ECP)
- Equipe pluridisciplinaire (maths appliquées, statistique, informatique, biologie)
- Importantes collaborations académiques (INRA, CIRAD, CAU, TU Munich, Univ. Wageningen ...)

CybeleTech : technologies numériques pour l'agriculture et la forêt

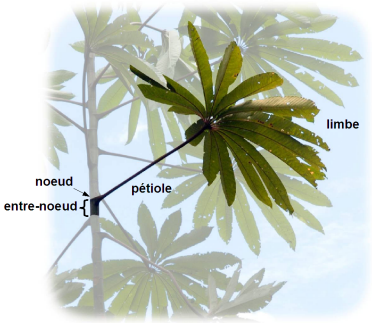
- Société crée en 2011, partenariat fort avec l'équipe de recherche Digiplante
- CybeleTech propose des services d'aide à la gestion de la ressource végétale par (i) développement de modèles et logiciels dédiés, (ii) réalisation d'études et activités de conseil

Plan

1. De la biologie végétale et la botanique vers des modèles mathématiques
   - Positionnement
   - Les briques de base : une histoire de factorisation botanique
2. Prévision du rendement
   - Quel modèle et comment le paramétriser ?
   - Assimilation de données
3. Lorsque les modèles aident à déchiffrer l'interaction G $\times$ E
4. De la plante à la population : croissance en compétition et simulations large échelle
   - Scènes virtuelles
   - Scènes calibrées sur des observations réelles

Positionnement de notre approche de modélisation

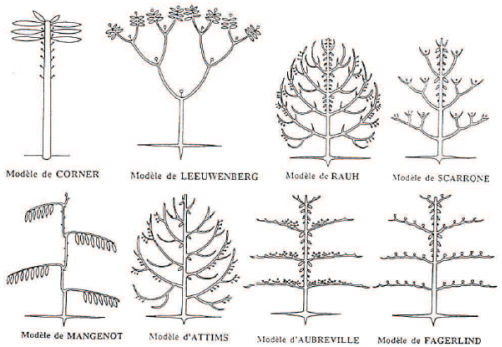
Modélisation intégrative de la croissance des plantes à l'échelle de l'individu

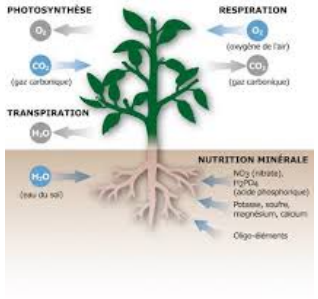






Modélisation Structure - Fonction





Généricité

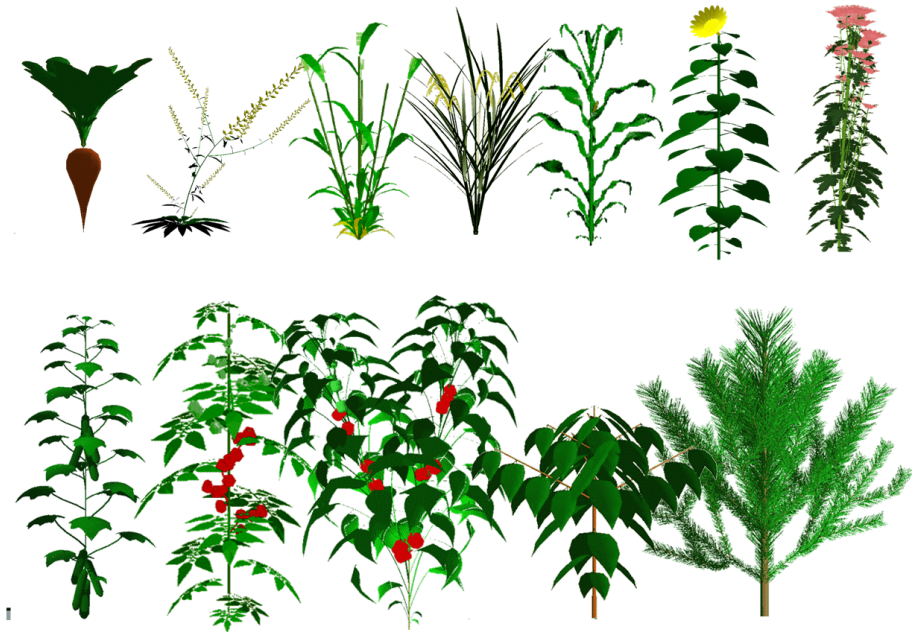




Conception de modèles

Des concepts biologiques aux équations mathématiques

- Des modèles originaux : GreenLab [Reffye et Hu, 2003], LNAS

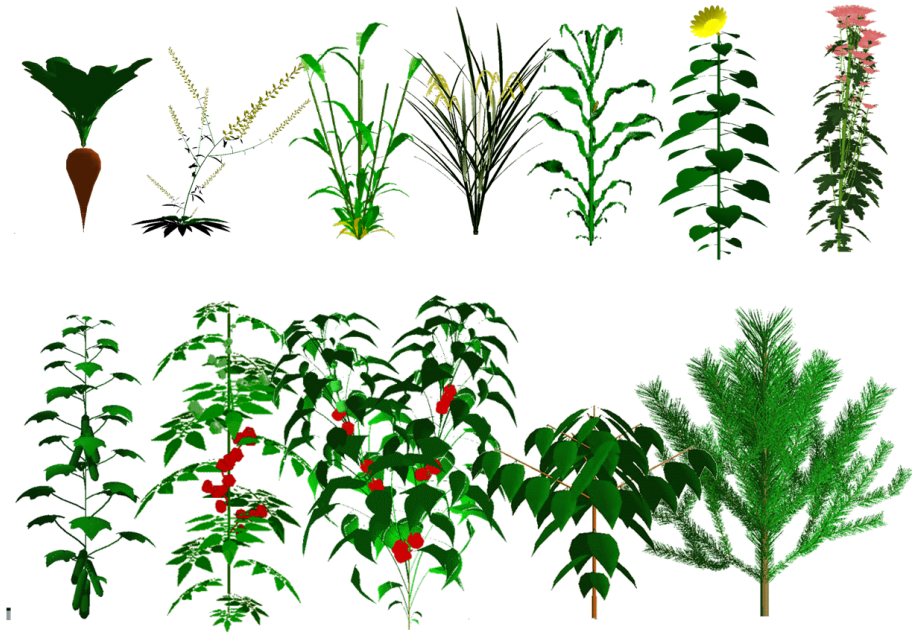


- Benchmarking, études sur des modèles classiques (STICS, CERES, PILOTE, SUNFLO...)
- Etude des interactions plante-environnement, en particulier interaction avec la ressource en eau et en azote.

Conception de modèles

Des concepts biologiques aux équations mathématiques

- Des modèles originaux : GreenLab [Reffye et Hu, 2003], LNAS

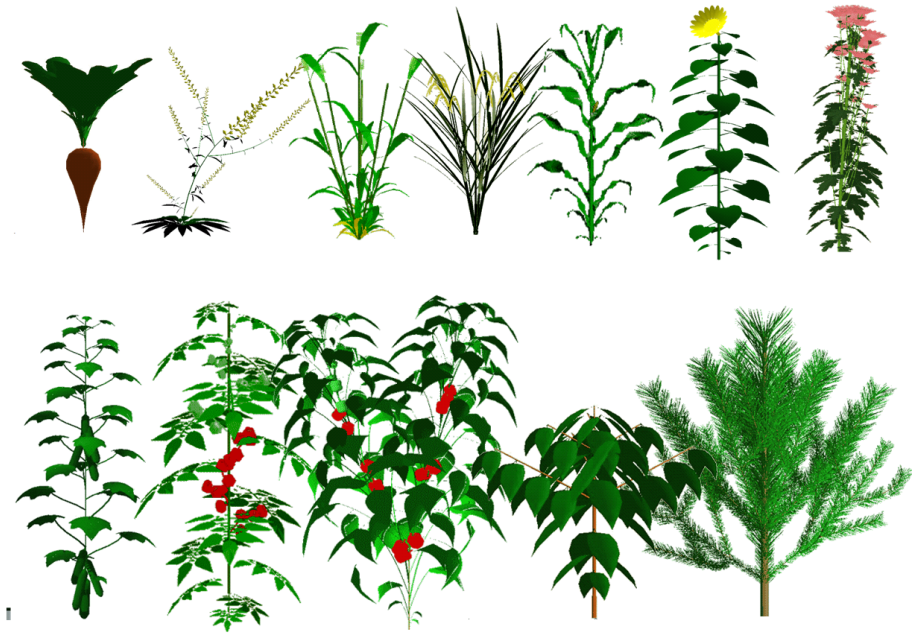


- Benchmarking, études sur des modèles classiques (STICS, CERES, PILOTE, SUNFLO...)
- Etude des interactions plante-environnement, en particulier interaction avec la ressource en eau et en azote.

Conception de modèles

Des concepts biologiques aux équations mathématiques

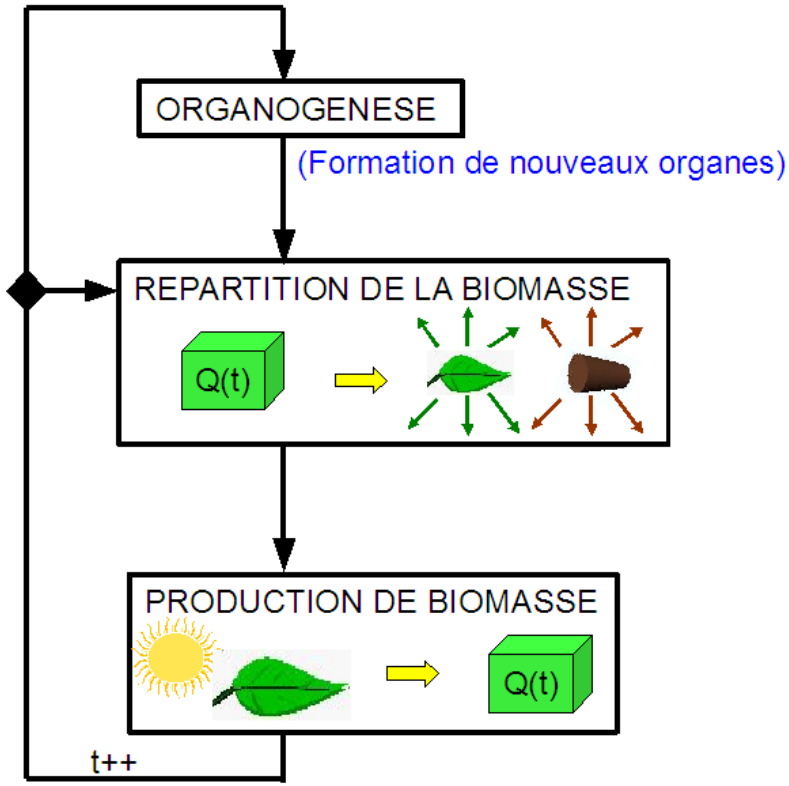
- Des modèles originaux : GreenLab [Reffye et Hu, 2003], LNAS



- Benchmarking, études sur des modèles classiques (STICS, CERES, PILOTE, SUNFLO...)
- Etude des interactions plante-environnement, en particulier interaction avec la ressource en eau et en azote.

Système Dynamique de la Croissance des Plantes

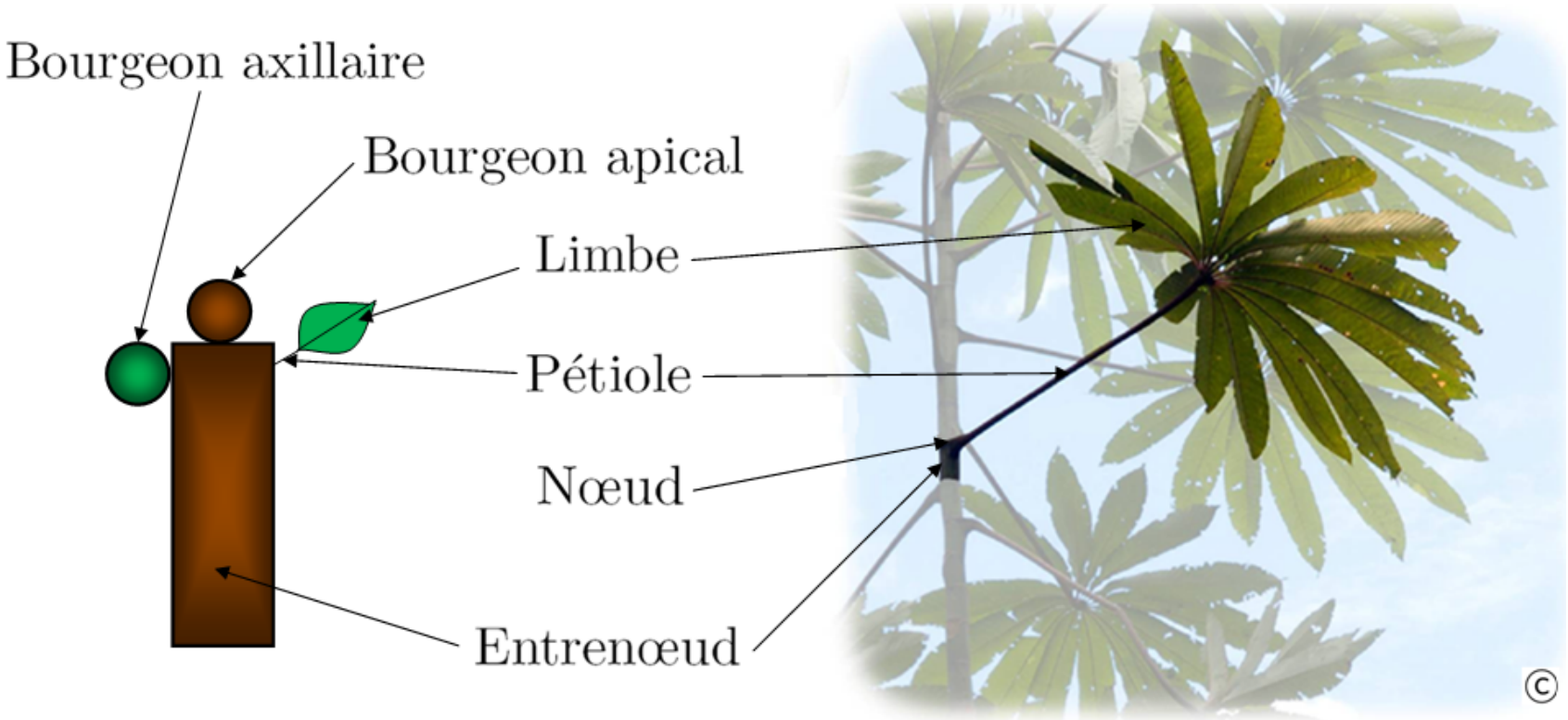
$$\frac{dX}{dt} = F(X(t), U(t), P, t)$$



- $X(t)$: variables d'état $\Longrightarrow$ masses des organes
- $F \Longrightarrow$ lois biophysiques
- P : paramètres $\Longrightarrow$ génétiques
- $U(t)$: variables de contrôle $\Longrightarrow$ conditions environnementales

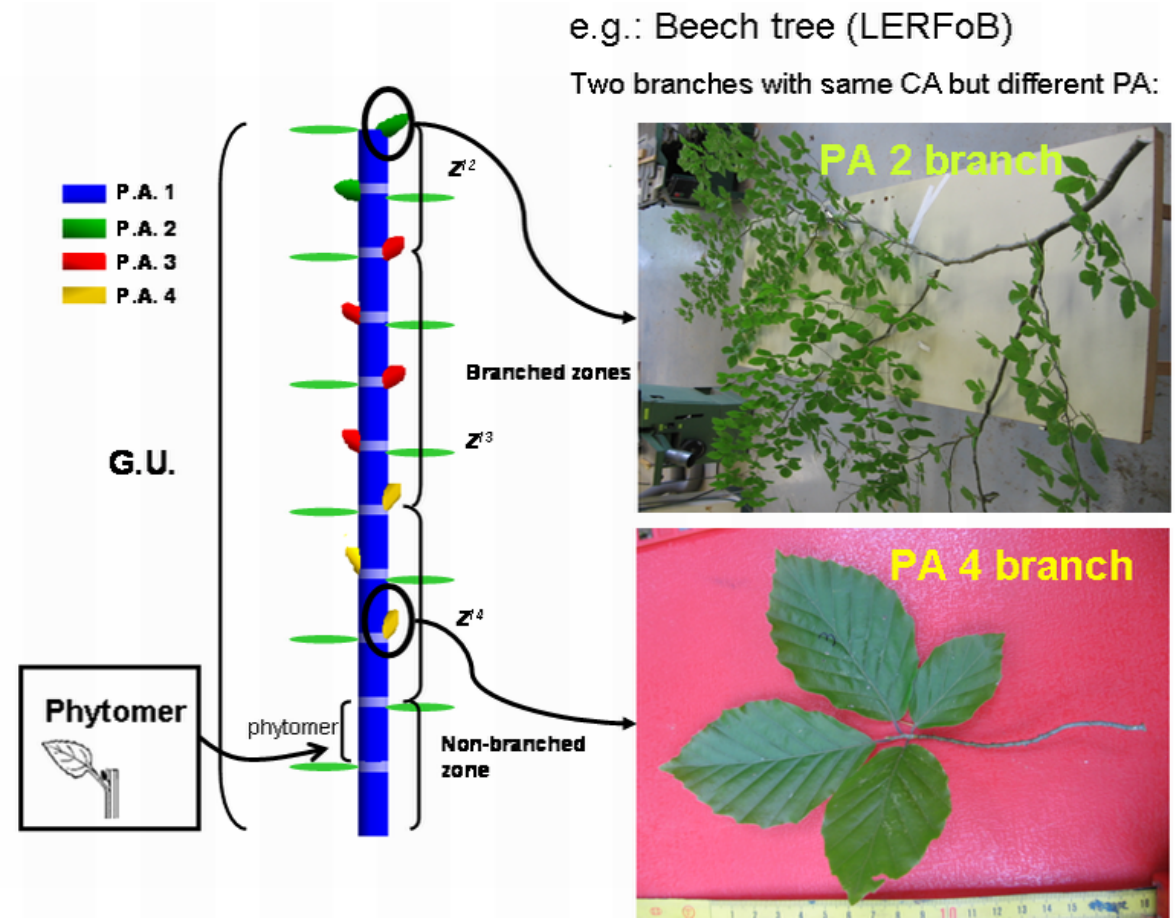
Une factorisation botanique

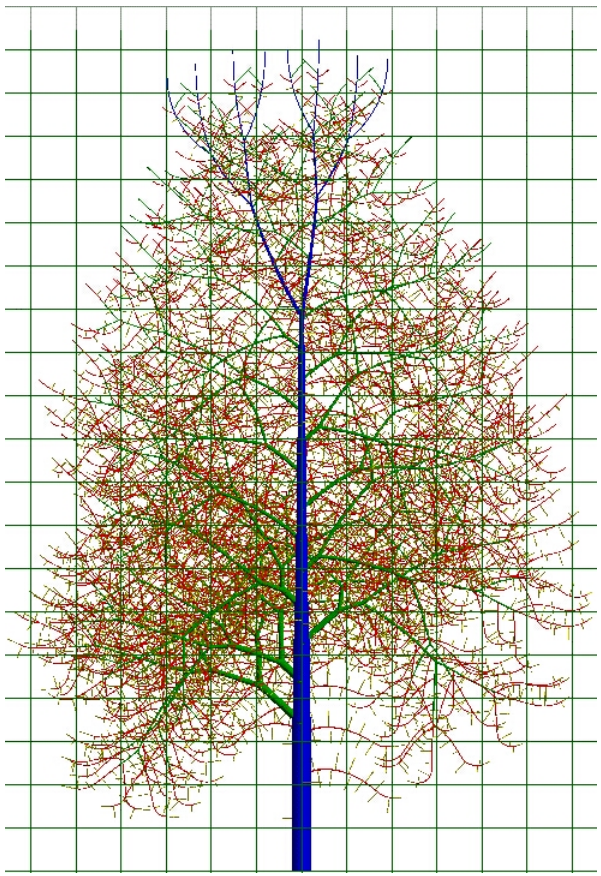
'Brique' de base : phytomère



Une factorisation botanique

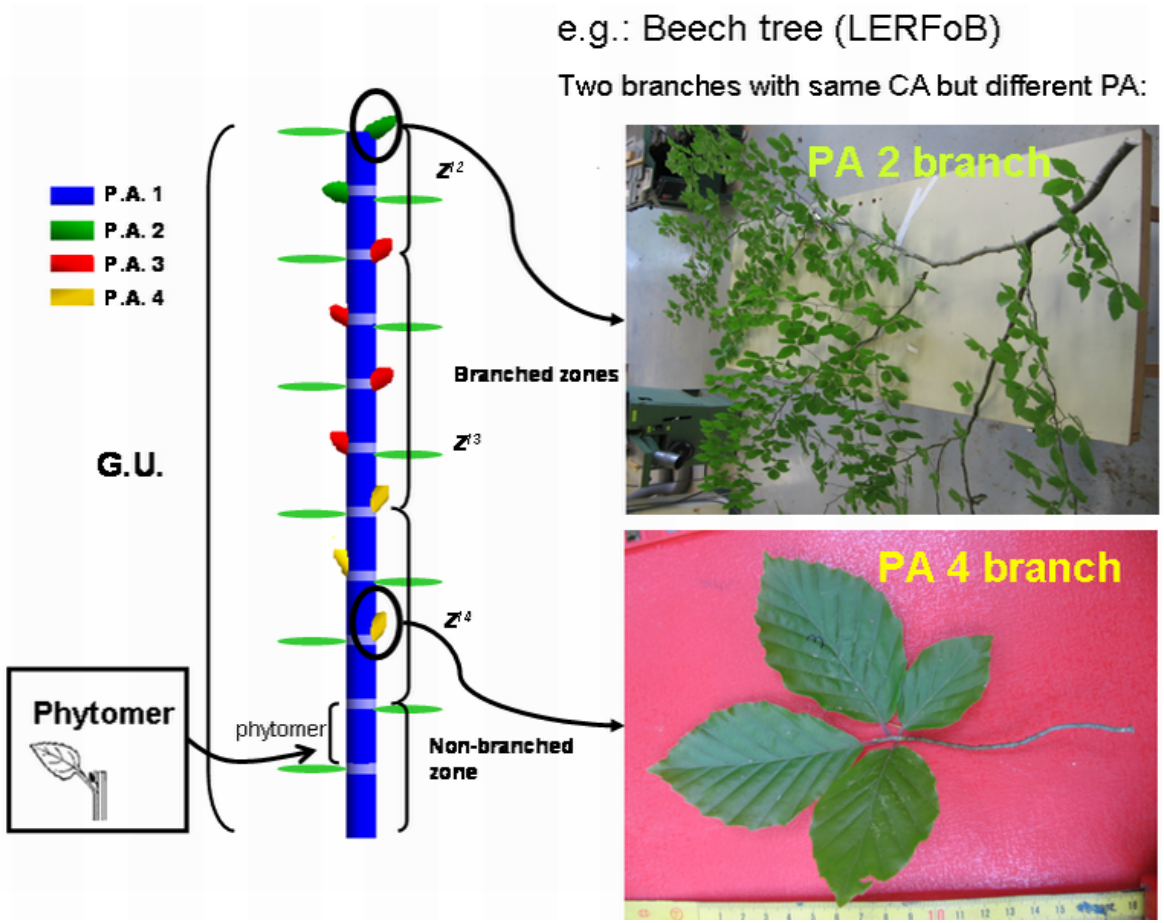
≪ Plants are modular organisms that develop by the repetition of elementary botanical entities(...) ≫
[Barthélémy and Caraglio, 2007]

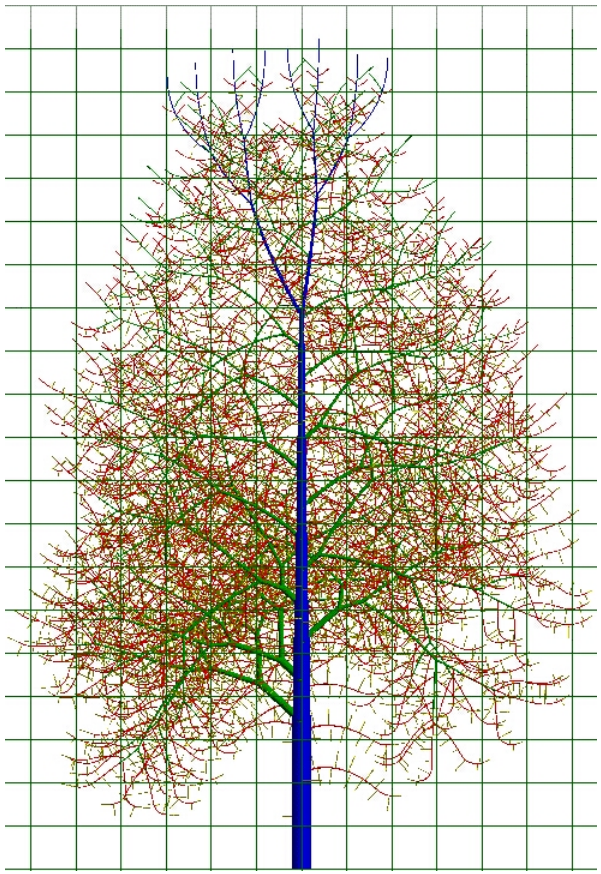




Une factorisation botanique

Définition de grammaires de croissance ou L-systemes s'appuyant sur l'âge physiologique du bourgeon

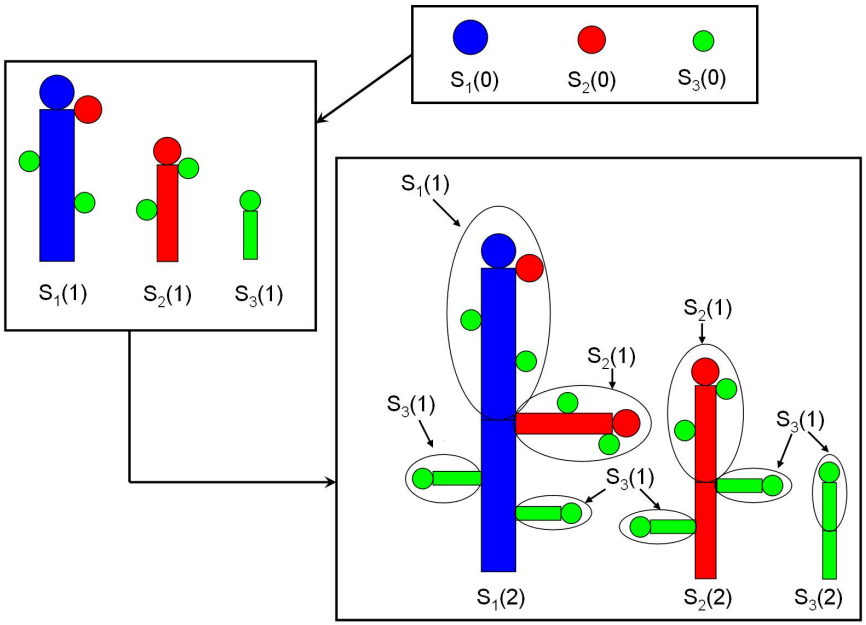




Une factorisation botanique

Factorisation des grammaires formelles de l'organogenèse [Cournède et al., 2006]

$$S_p^t(n) = \left[\prod_{p \leq q \leq P} \left(m_{pq}^t(n)\right)^{u_{pq}(t+1-n)} \left(S_q^t(n-1)\right)^{b_{pq}(t+1-n)} \right] S_p^t(n-1) \quad .$$



$\Longrightarrow$ les sous-structures de la plante sont construites récursivement, ce qui réduit le temps de calcul spectaculairement (d'exponentiel avec le temps à linéaire ou quadratique !)

Extension au cas stochastique [Loi et Cournède, 2008, 2010] $\Longrightarrow$ Liens théoriques des grammaires stochastiques au cadre Markovien

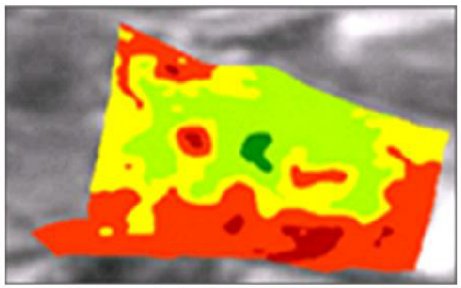
- Processus de Galton-Watson pour déterminer par récurrence les fonctions génératrices du nombre d'organes dans la plante
- Méthodes symboliques de la combinatoire pour les fonctions génératrices

Plan

1. De la biologie végétale et la botanique vers des modèles mathématiques
   - Positionnement
   - Les briques de base : une histoire de factorisation botanique
2. Prévision du rendement
   - Quel modèle et comment le paramétriser ?
   - Assimilation de données
3. Lorsque les modèles aident à déchiffrer l'interaction G $\times$ E
4. De la plante à la population : croissance en compétition et simulations large échelle
   - Scènes virtuelles
   - Scènes calibrées sur des observations réelles

Challenge global : Aide à la décision en agriculture

Productivité et Durabilité



Informations Parcelles + Sol



Informations Météo (T, L, P...)



Informations Culturales (Variété, Irrigation, traitements, ...)

Modèle Mathématique - Simulation Numérique

Prévisions de rendements;

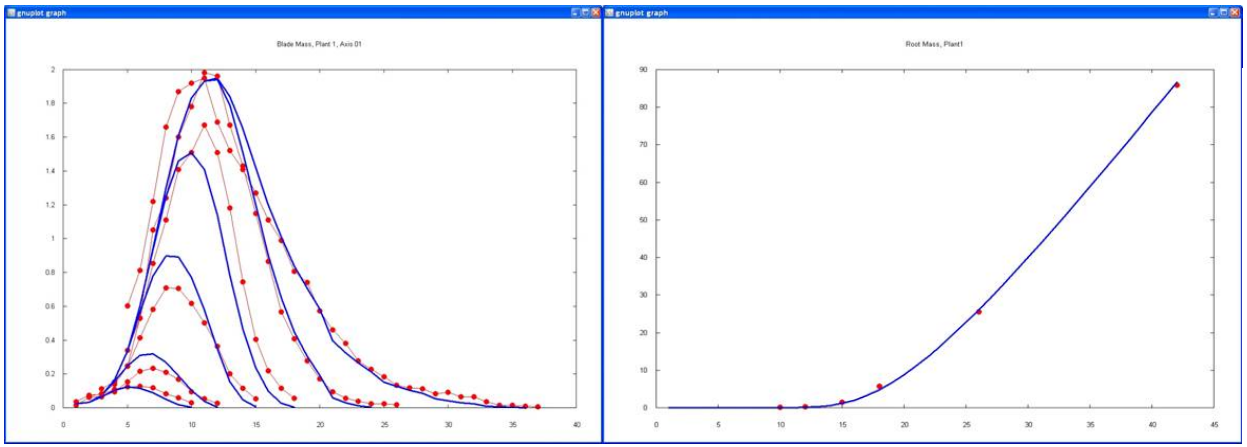
Etude de scénarios (variabilité...);

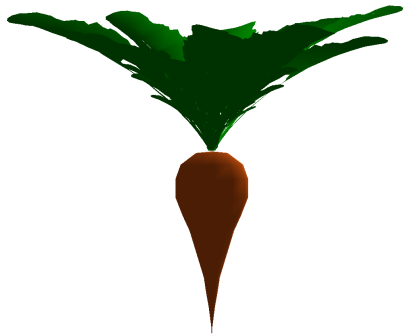
Optimisation et contrôle optimal;

Estimation Paramétrique des Modèles

$$\begin{cases} \text{State equ. : } X_{n+1} = F_n(X_n, U_n, P) \\ \\ \text{Observation func. : } Y = G(X_N, P), \text{ corresponding to exp. data } Y^{exp} \end{cases}$$

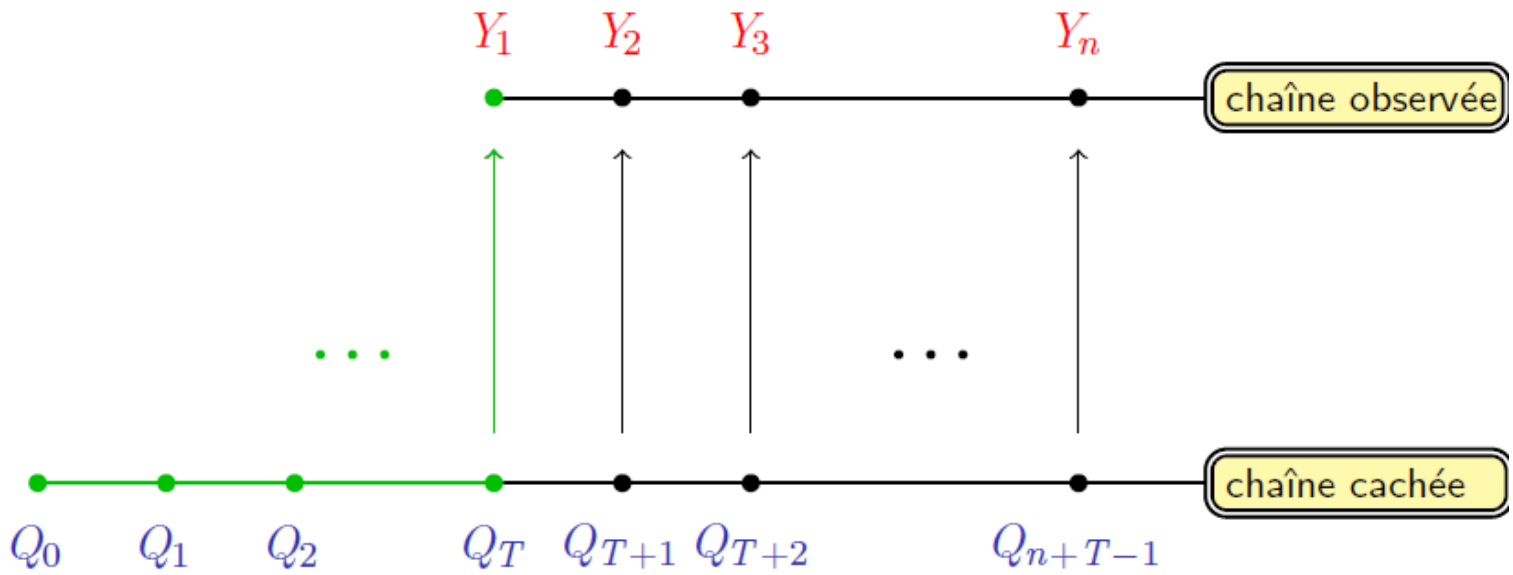
- Parameter estimation = estimate P^* to minimize $||Y^{exp} - Y(P)|| \Longrightarrow$ 2-Stage Aitken Estimator.
- Example, Sugar-Beet ITB, 6 dates. Multi-Fitting, Complete Description : $t_1, t_2, \ldots, t_6$, and the masses of all organs at all dates [Lemaire et al., 2008].





Identification Paramétrique

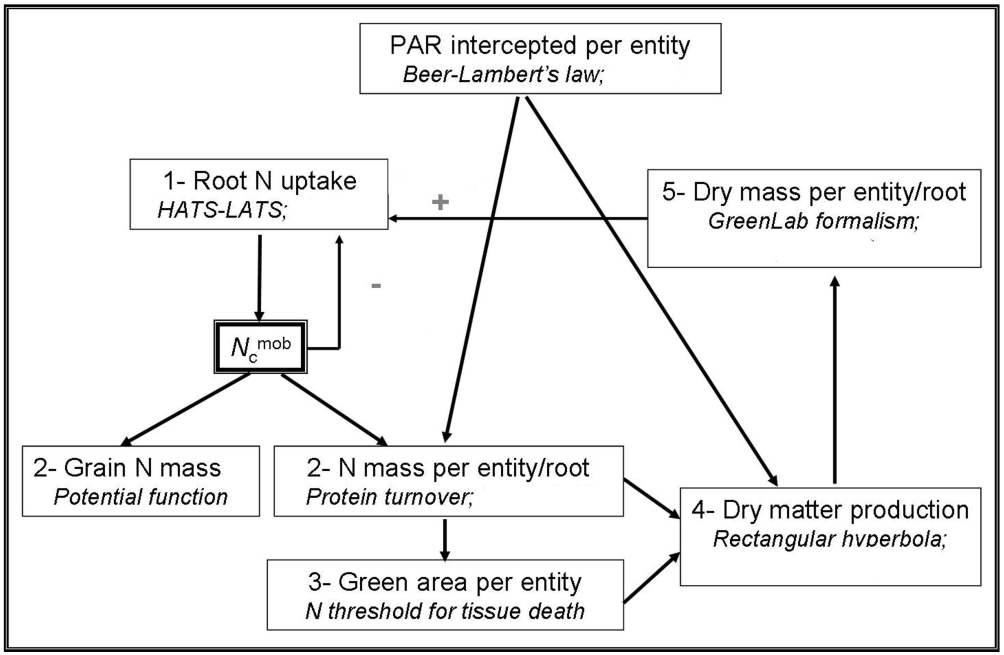
- Formalisation de GreenLab dans le cadre des modèles de Markov cachés en utilisant la suite architecturale (et non la suite temporelle) [Trevezas et Cournède, 2012]



- Stochastic EM avec Sequential MonteCarlo ou MCMC et Bootstrap pour la distribution des paramètres

Analyse de Sensibilité Globale, ex du Modèle NEMA

NEMA : modélise métabolismes azote et carbone pour le blé
[Bertheloot, et al. 2011]



5 modules biologiques :

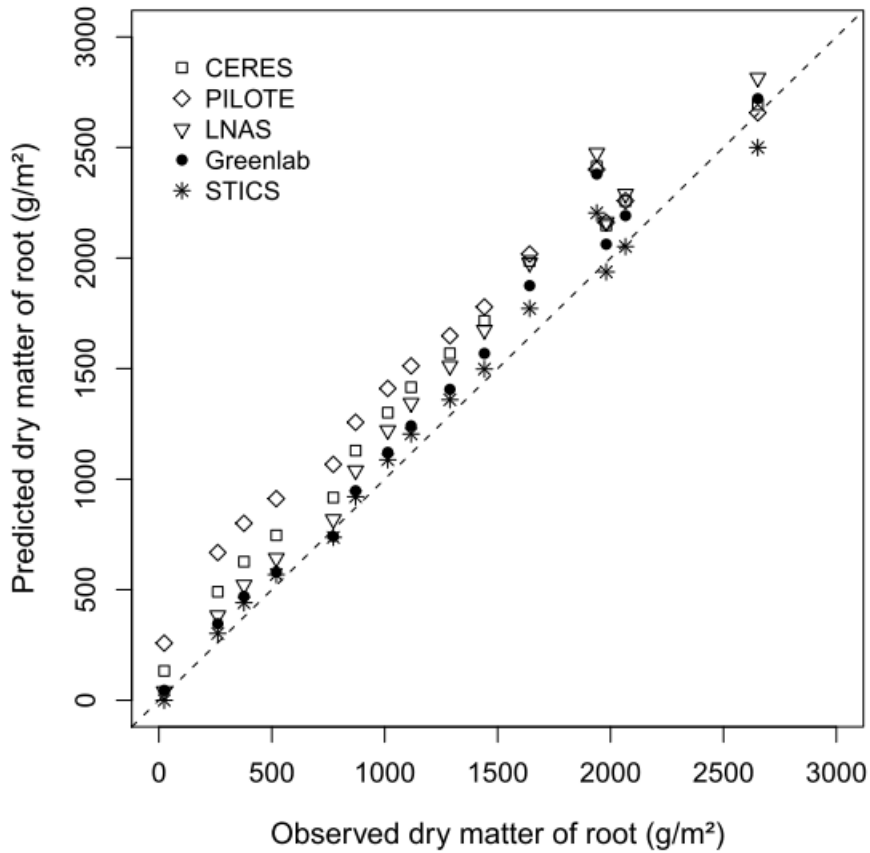
- N acquisition by roots (34 param.)
- N distribution (28 param.)
- C acquisition via photosynthesis (10 param.)
- C distribution (5 param.)
- Senescence (5 param.)

- Analyse de sensibilité globale : Indices de Sobol, basés sur décomposition de la variance : $S_i = \dfrac{V(E(Y|X_i))}{V(Y)}$
  $\Longrightarrow$ module par module, permet de sélectionner 17 paramètres influents (sur 82) : simplification du modèle drastique !

Benchmarking de modèles

Evaluation de la capacité prédictive de 5 modèles pour la betterave sucrière

- Data set 2010 for Calibration + Data Set 2011 to test the Predictive Capacity (With ITB)

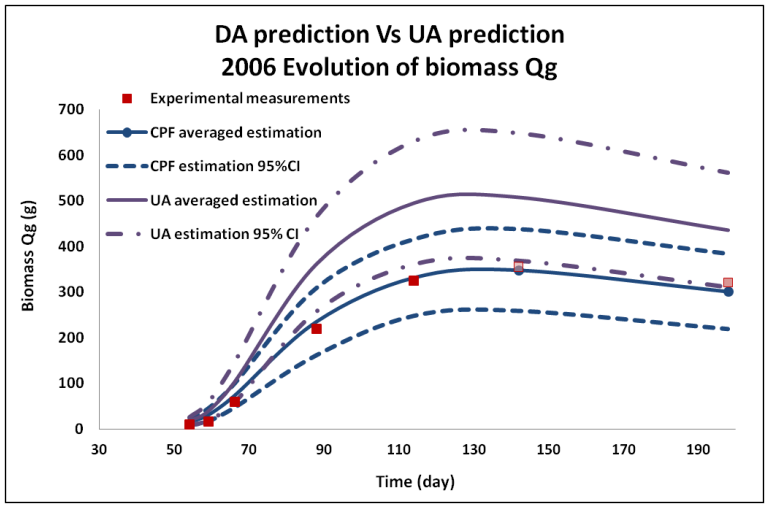


$\Longrightarrow$ Models with HI seem less predictive

Assimilation de Données pour la Prévision

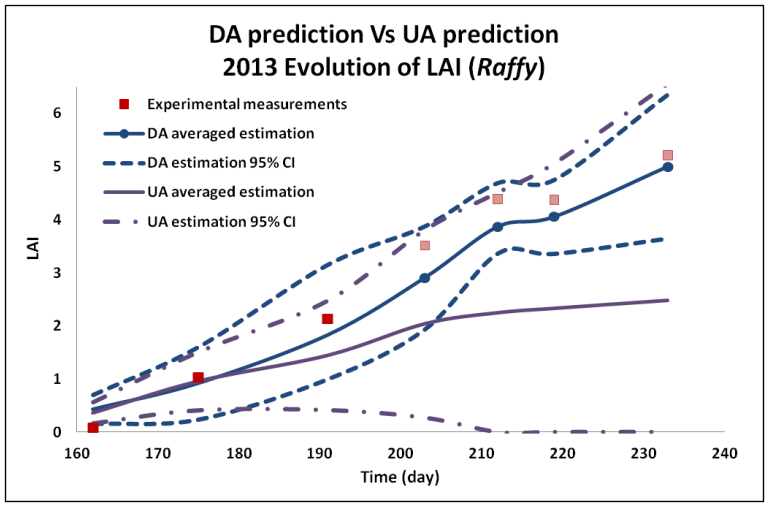
$\Longrightarrow$ Données en début de culture permettent d'affiner la prévision rendement.

Cas de la betterave (LNAS)



	Real Data 2006	DA estimates (relative error in %)	95% CI	UA estimates (relative error in %)	95% CI
$Q_b\,(t_{142})$	355.2	348.1 (2.0%)	[258.7; 437.4]	507.8 (43.0%)	[368.4; 647.3]
$Q_b\,(t_{198})$	320.6	301.3 (6.0%)	[219.0; 383.6]	435.7 (35.9%)	[384.3; 560.7] *
$Q_r\,(t_{142})$	1459.2	1716.2 (17.6%)	[1427.9; 2004.5]	1930.7 (32.3%)	[1603.0; 2258.4] *
$Q_r\,(t_{198})$	2400.0	2644.3 (10.2%)	[2209.4; 3079.2]	2942.9 (22.6%)	[2455.0; 3430.7] *

Cas du blé (STICS)



	Real Data 2013	DA estimates (relative error in %)	95% CI	UA estimates (relative error in %)	95% CI
$hur\,(t_{198})$	0.276	0.272 (1.7%)	[0.260; 0.283]	0.290 (5.0%)	[0.286; 0.295]
$hur\,(t_{210})$	0.274	0.303 (10.3%)	[0.284; 0.322]	0.308 (12.2%)	[0.303; 0.313]
$hur\,(t_{226})$	0.289	0.297 (2.6%)	[0.278; 0.316]	0.303 (4.7%)	[0.295; 0.310]
$hur\,(t_{255})$	0.239	0.262 (9.7%)	[0.240; 0.283]	0.272 (14.0%)	[0.253; 0.291]
$hur\,(t_{275})$	0.257	0.260 (1.2%)	[0.229; 0.291]	0.267 (4.0%)	[0.233; 0.302]
$LAI\,(t_{203})$	4.235	4.514 (6.6%)	[3.144; 5.883]	3.482 (17.8%)	[0.471; 6.494]
$LAI\,(t_{212})$	4.735	4.627 (2.3%)	[3.192; 6.062]	3.842 (18.9%)	[0.000; 7.693]
$LAI\,(t_{219})$	**4.225**	4.867 (15.2%)	[2.394; 7.340]	3.994 (5.5%)	[0.000; 8.638]
$LAI\,(t_{233})$	4.910	5.215 (6.2%)	[0.969; 9.462]	4.274 (13.0%)	[0.000; 11.214]
$magrain\,(t_{266})$	819.38	804.33 (1.8%)	[611.68; 996.98]	550.58 (32.8%)	[128.97; 972.20]

Plan

1. De la biologie végétale et la botanique vers des modèles mathématiques
   - Positionnement
   - Les briques de base : une histoire de factorisation botanique
2. Prévision du rendement
   - Quel modèle et comment le paramétriser ?
   - Assimilation de données
3. Lorsque les modèles aident à déchiffrer l'interaction G $\times$ E
4. De la plante à la population : croissance en compétition et simulations large échelle
   - Scènes virtuelles
   - Scènes calibrées sur des observations réelles

Problématique : l'Interaction Génotype × Environment





[Rey and Lecoeur, 2003]



[Allirand and Jullien, 2005]

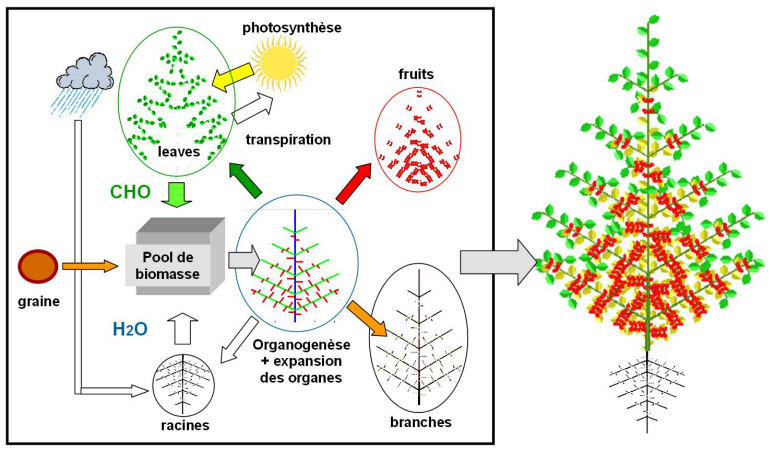
Problématique : Interaction Génotype $\times$ Environment

Caractérisation de la production végétale implique la compréhension fine de l'interaction génétique - environnement (phénotypage).
$\Longrightarrow$ Les modèles biophysiques doivent permettre de comprendre cette interaction : $\ll$ 1 genotype = 1 jeu de paramètres du modèle stables $\gg$

Objectifs

- Phénotype = f1(Paramètres, Environnement)
- Paramètres = f2(Génétique)

f1 : Système Dynamique de la Croissance des Plantes



$$\frac{dX}{dt} = F(X(t), U(t), P, t)$$

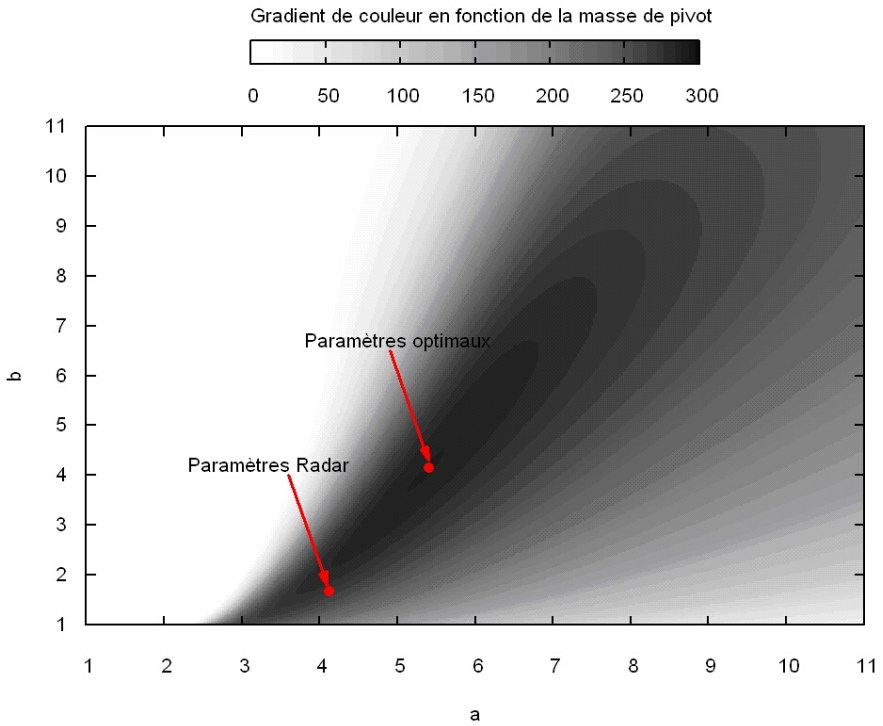
- $X(t)$: variables d'état $\Longrightarrow$ masses des organes
- $F \Longrightarrow$ lois biophysiques
- P : paramètres $\Longrightarrow$ génétiques
- $U(t)$: variables de contrôle $\Longrightarrow$ conditions environnementales

Amélioration variétale

Création de nouvelles variétés par les industries semencières

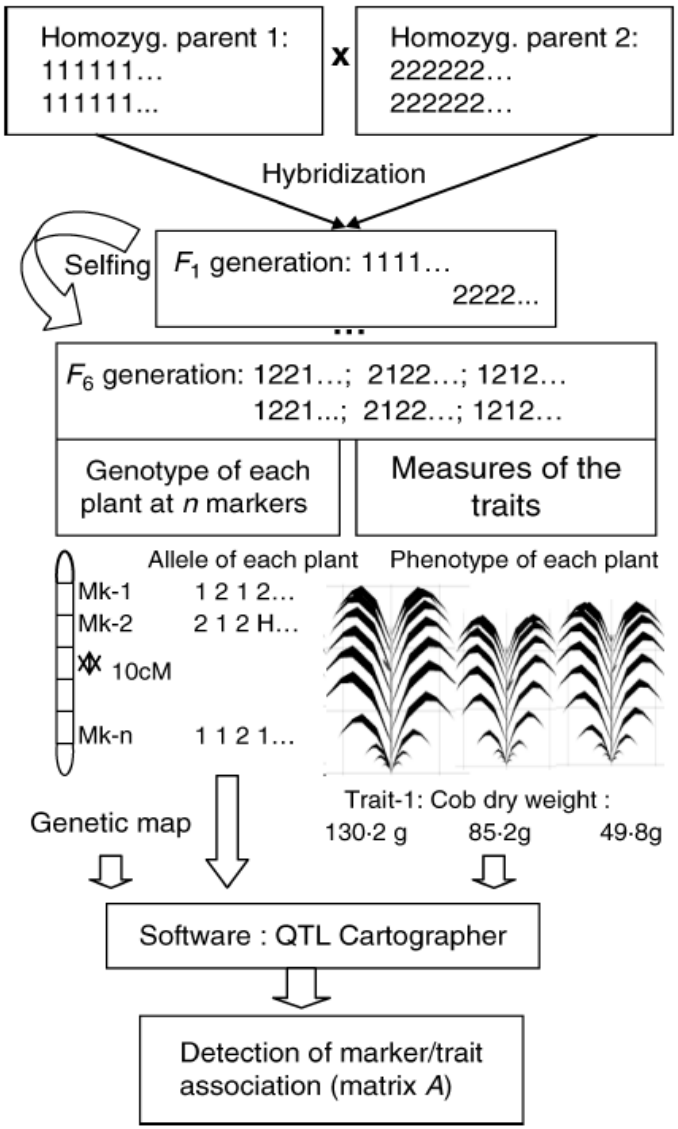
- Expérimentations virtuelles
- Optimisation multidim. dans l'espace des paramètres pour la détermination d'idéotypes

Ex. Betterave Sucrière, optimisation de la stratégie d'allocation à la racine

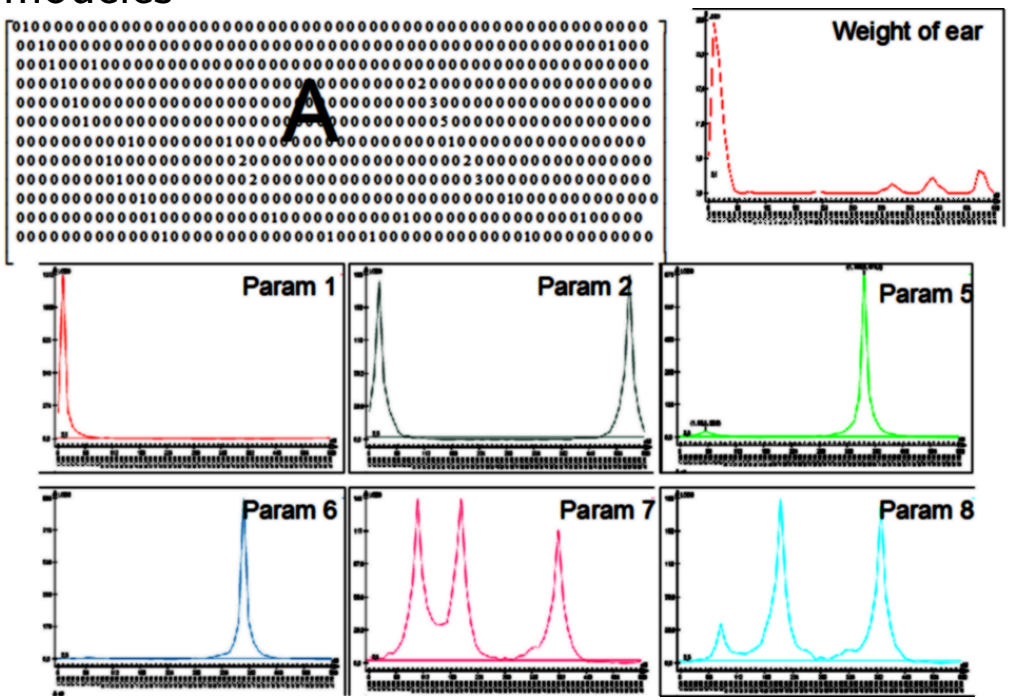


- Prise en compte de la dérive climatique dans cette optimisation, ≪ optimisation robuste ≫
- Paramètres=f2(Génétique) (génétique quantitative, réseaux de gènes ...)

Amélioration variétale



Détection des QTLs associés au paramètres des modèles



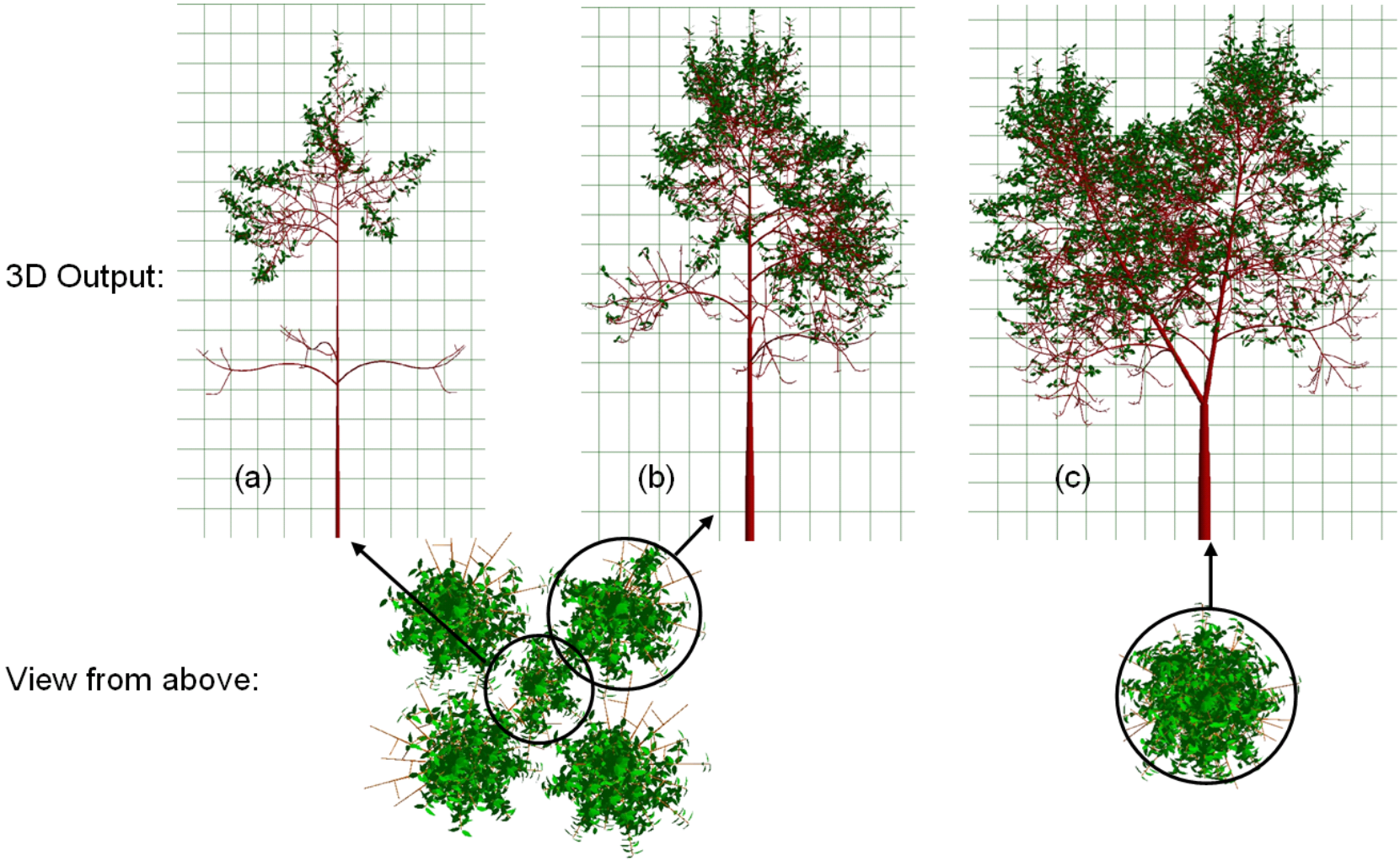
Sélection assistée par marqueurs ou (futur) sélection génomique + bioinfo

Plan

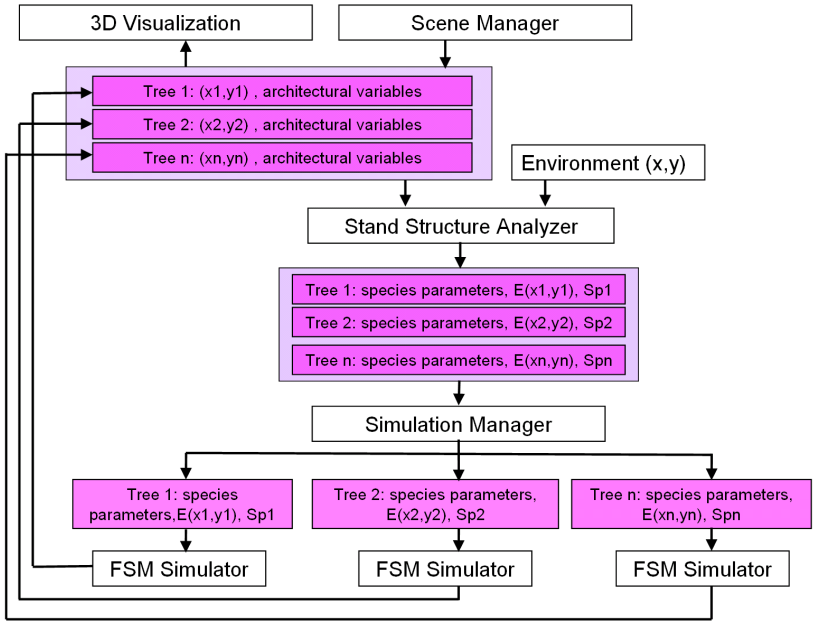
1. De la biologie végétale et la botanique vers des modèles mathématiques
   - Positionnement
   - Les briques de base : une histoire de factorisation botanique
2. Prévision du rendement
   - Quel modèle et comment le paramétriser ?
   - Assimilation de données
3. Lorsque les modèles aident à déchiffrer l'interaction G $\times$ E
4. De la plante à la population : croissance en compétition et simulations large échelle
   - Scènes virtuelles
   - Scènes calibrées sur des observations réelles

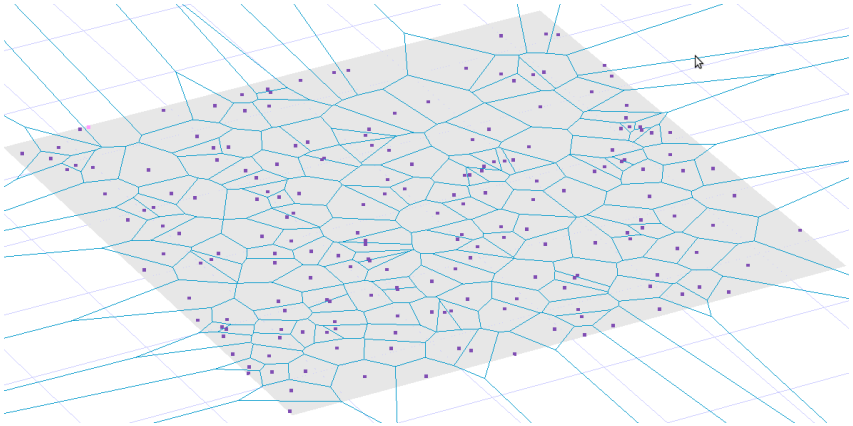
De la plante à la population

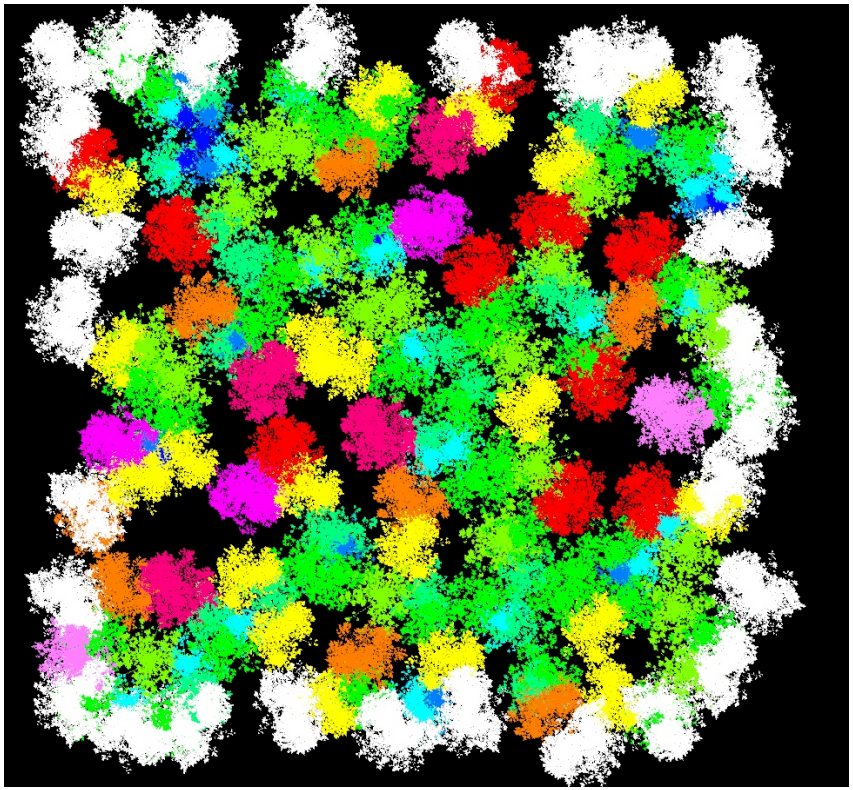
Modélisation de la rétroaction du fonctionnement sur la structure [Mathieu et al., 2009] et de la compétition pour la lumière [Cournède et al., 2008].

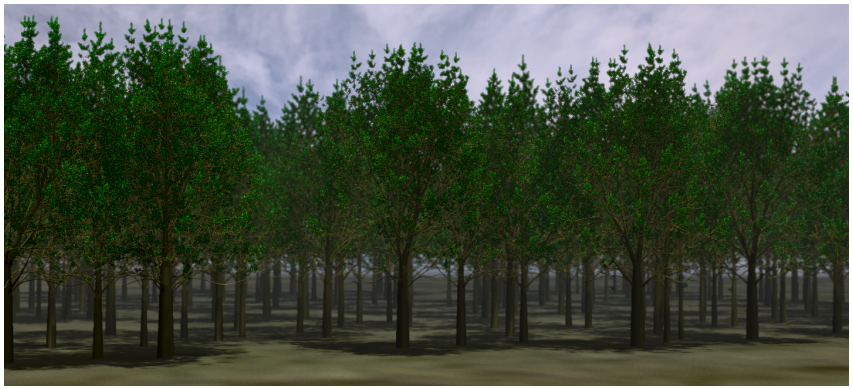


Simulations grande échelle



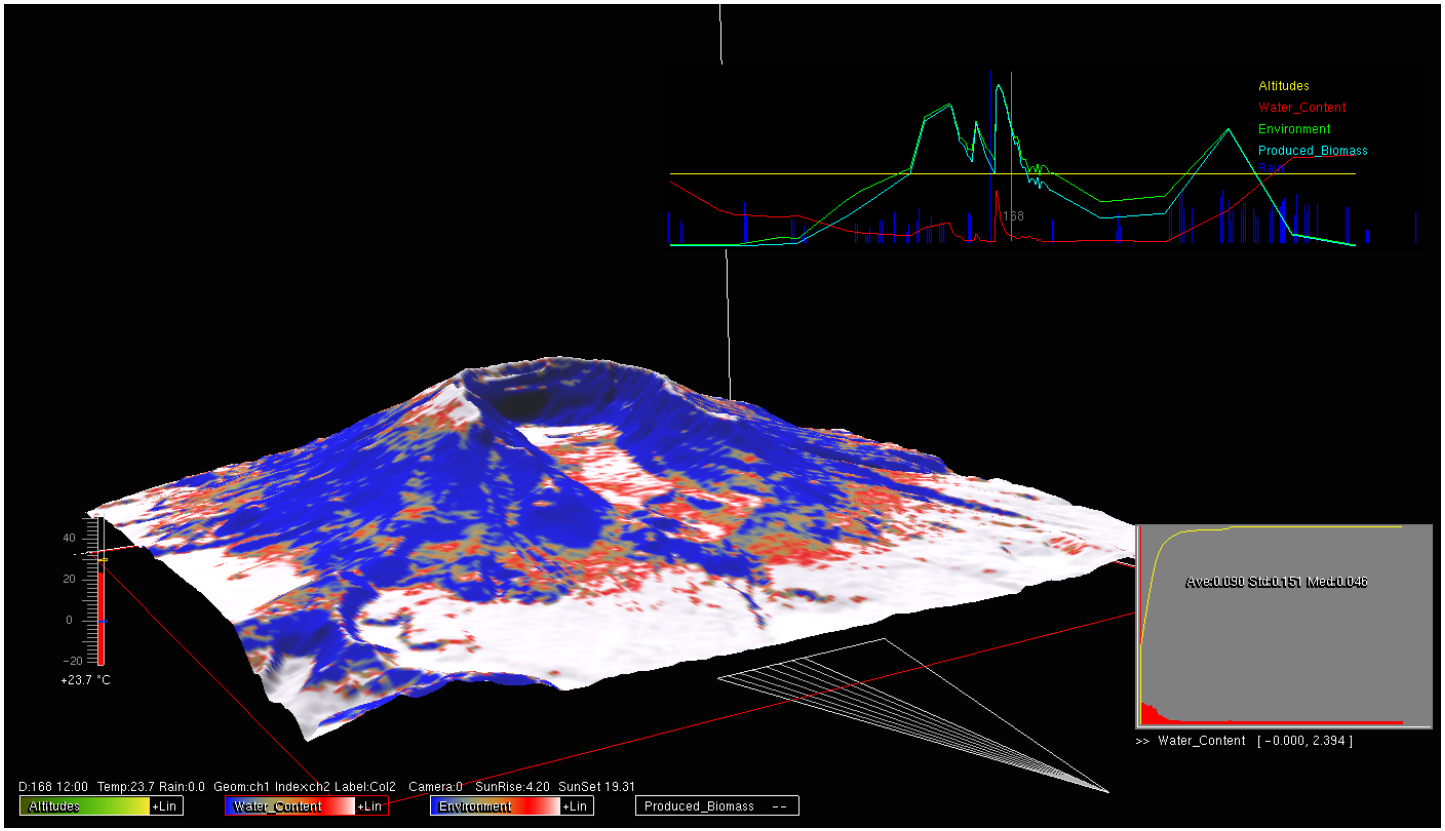






Simulations grande échelle

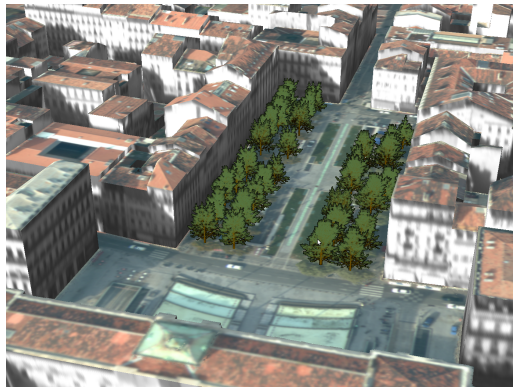
Prise en compte SIG, climat, phénomènes naturels [Le Chevalier et al., 2007]

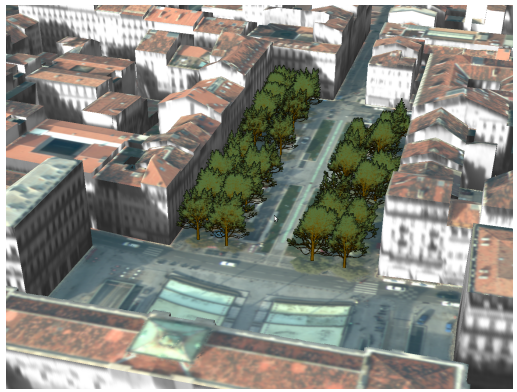


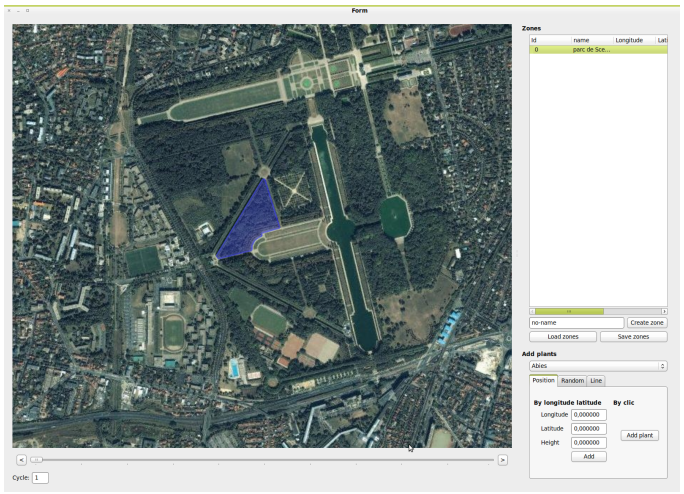
Assimilation de données d'imagerie pour reconstruction et prévision

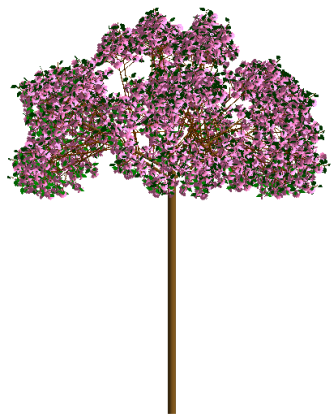
Dans le cadre d'un projet avec IGN

- Analyse d'image : individualisation d'arbres géoréférencés, extraction de caractéristiques biophysiques, reconnaissance d'espères
- Possibilités : assimilation de données, reconstruction, prévision.











MERCI

Avec les contributions de : Paul-Henry Cournède, Amélie Mathieu, Philippe de Reffye, Marc Jaeger, Patrick Heuret, Samis Trevezas, Corina Iovan, Benoit Bayol, Thomas Guyard, Charlotte Baey, Cédric Loi, Vincent Le Chevalier, Qiongli Wu, Jessica Bertheloot, Fenni Kang, Yuting Chen, Sébastien Lemaire et coll.