

**Projet CasDar :
Associer un niveau d'incertitude aux
prédictions des modèles**

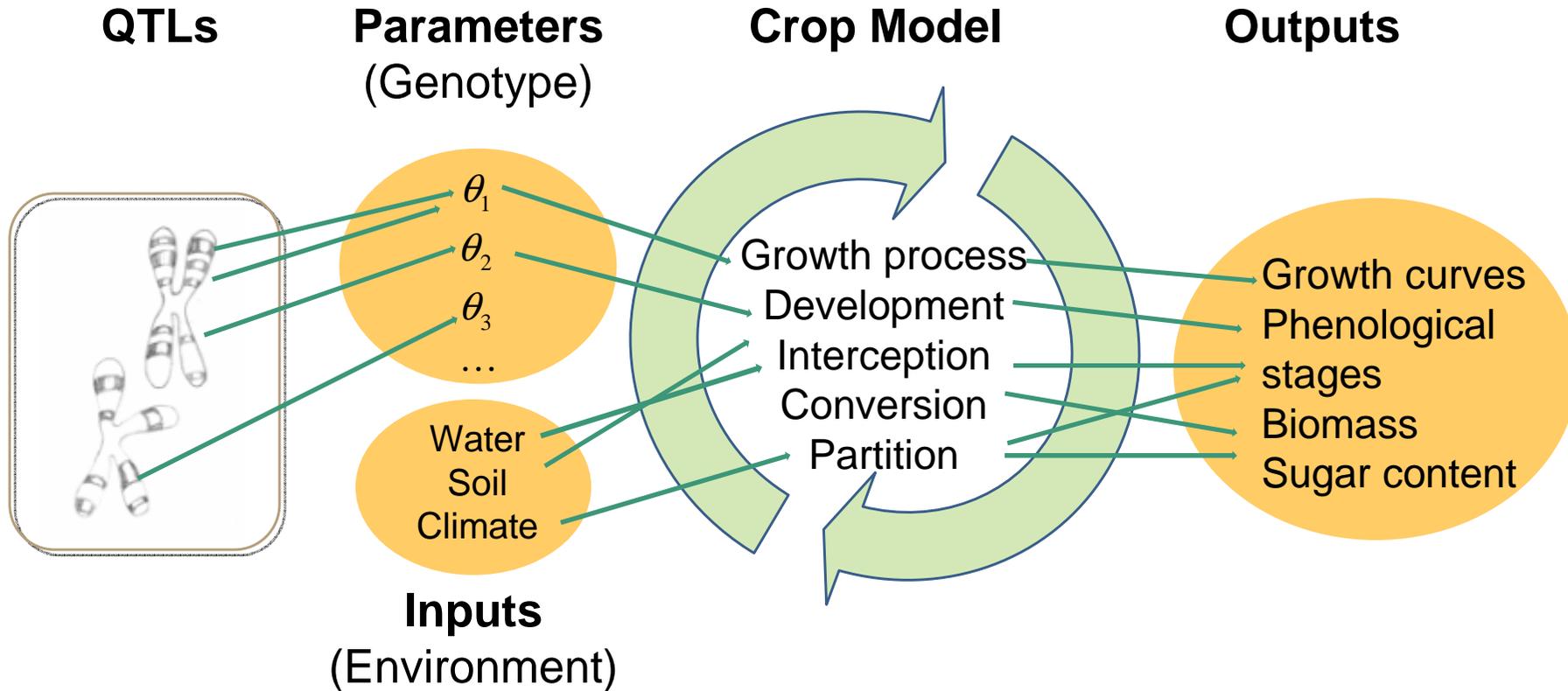
**Prédire la biomasse aérienne et le bilan
hydrique d'une culture de canne à
sucre**

Eric Gozé, Lauriane Rouan et Philippe Letourmy



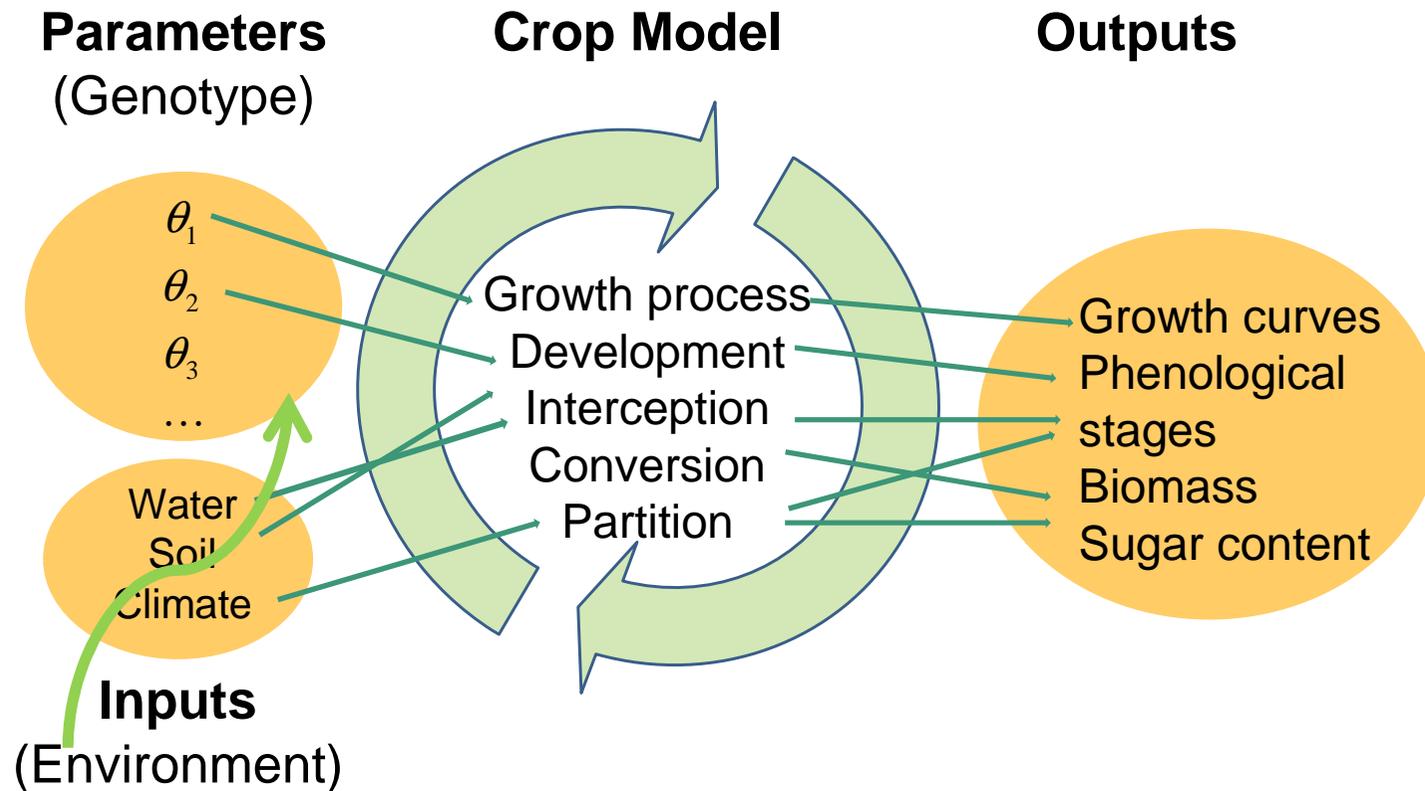
Paris 19/06/2013

Le projet DELICAS

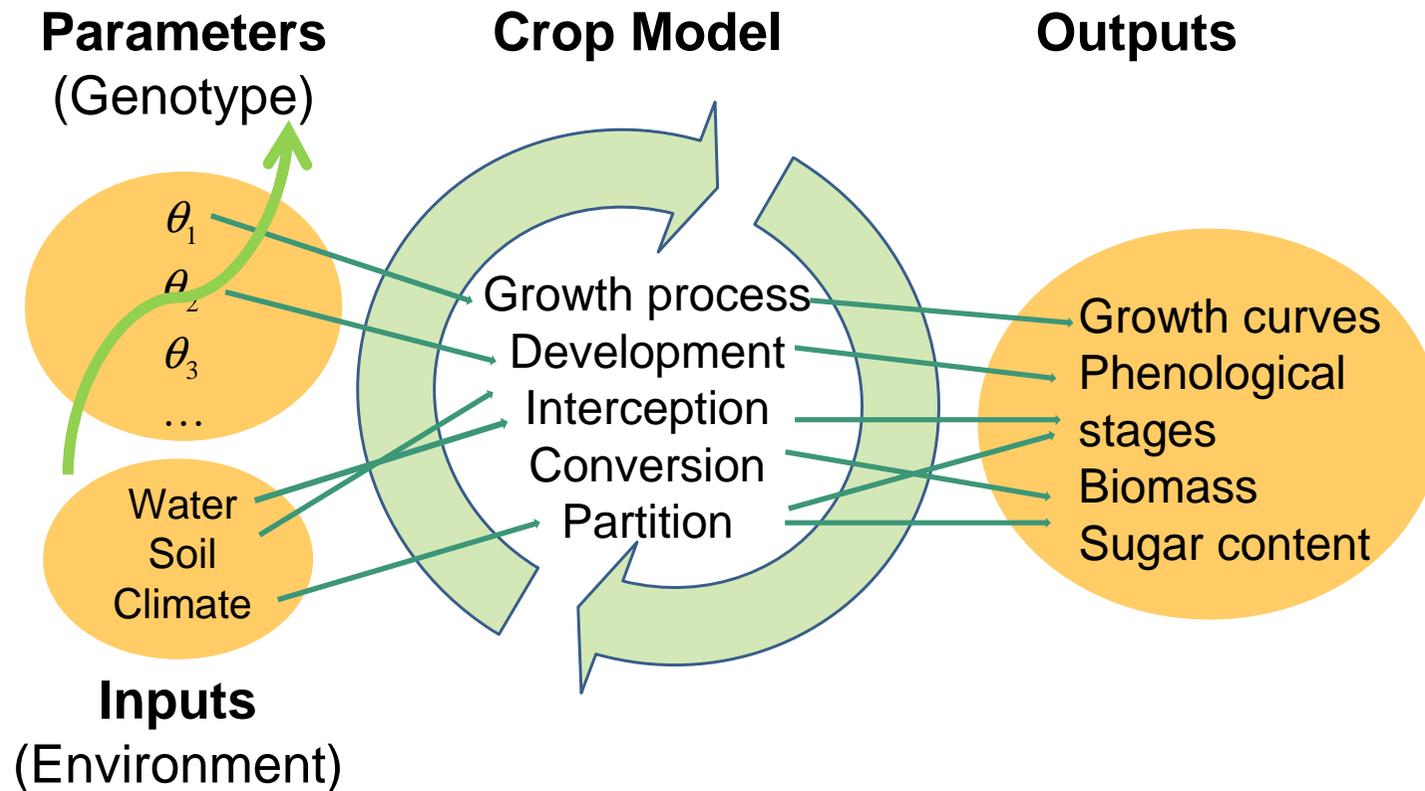


Association mapping et phénotypage par modèle pour l'identification de marqueurs moléculaires liés à l'élaboration et à la limitation du rendement chez la canne à sucre.

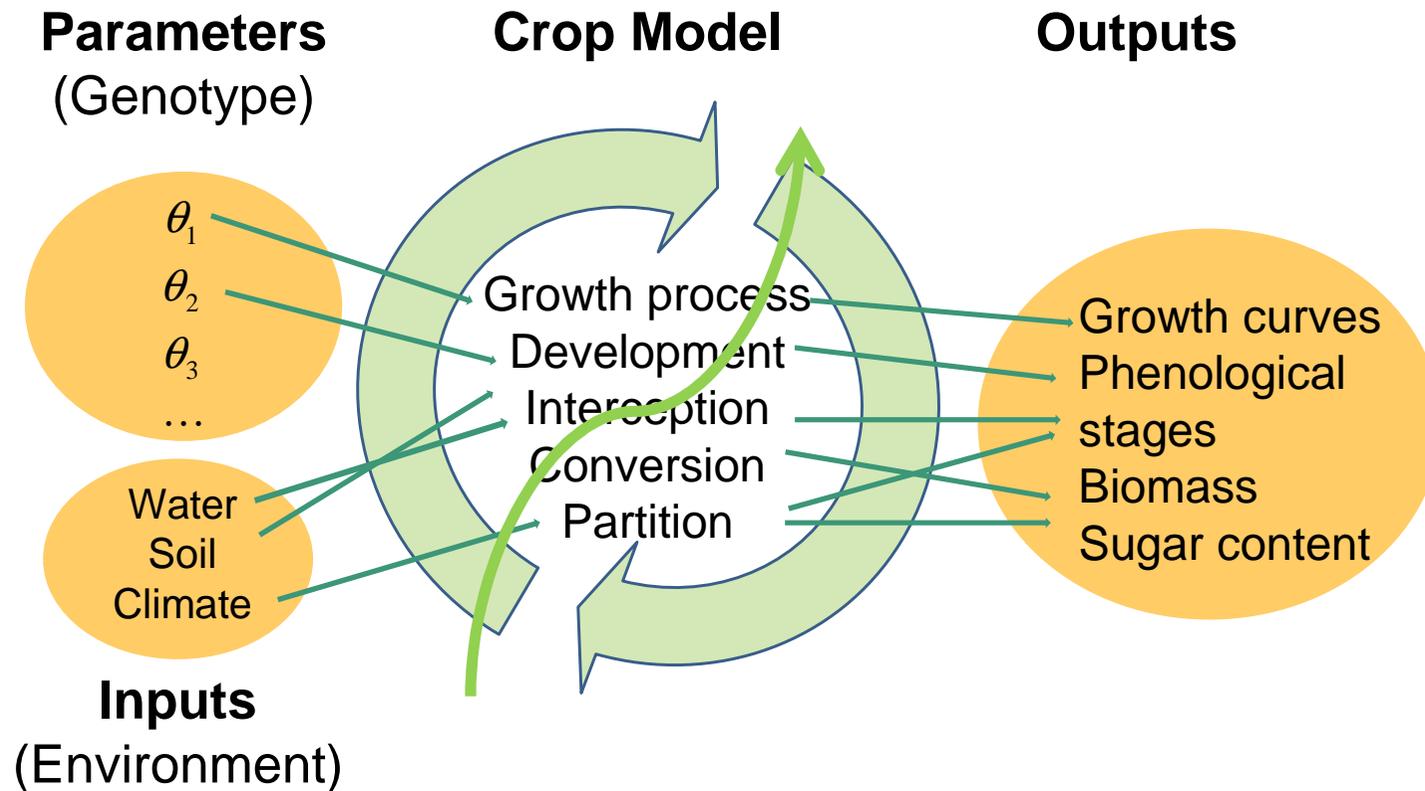
Le projet DELICAS



Le projet DELICAS



Le projet DELICAS



Le projet DELICAS



Vue Belle (altitude 650m)



La Mare (altitude 21 m)

- Phénotypage par modèle : expérimentation variétale en vue d'estimer des paramètres de modèle.
- Essai à 20 variétés x 3 blocs puis 200 variétés x 3 blocs dans deux lieux contrastés.



Propagation de l'incertitude sur les paramètres

DONNEES UTILISEES

- Chaque estimation utilise les données d'un seul bloc de chacun des deux sites
- Ajustement sur plusieurs observations ou séries d'observations :
 - lai* : surface foliaire /surface au sol: 1 mesure
 - ei* : interception de la lumière : 6 mesures
 - rdcan* : rendement en matière fraîche (g/m²)
 - msa* : matière sèche aérienne (g/m²) : 2 mesures

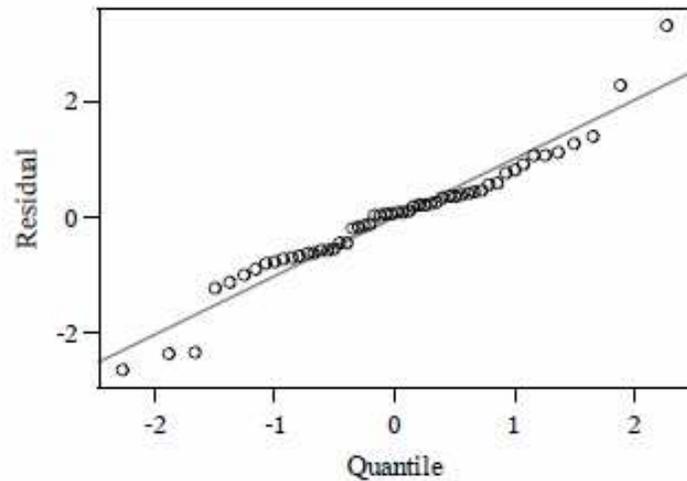
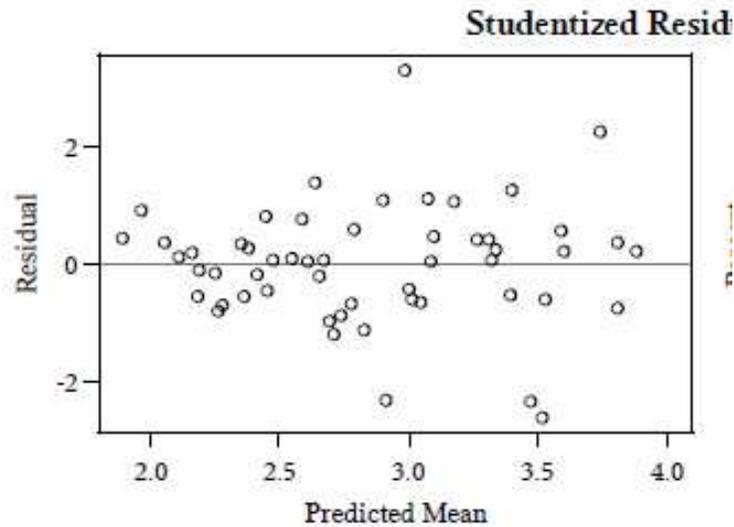
METHODE ET ALGORITHME

- Moindres carrés généralisés : l'on minimise $(\hat{Y} - Y)' \Sigma^{-1} (\hat{Y} - Y)$
- Σ est estimé par Manova

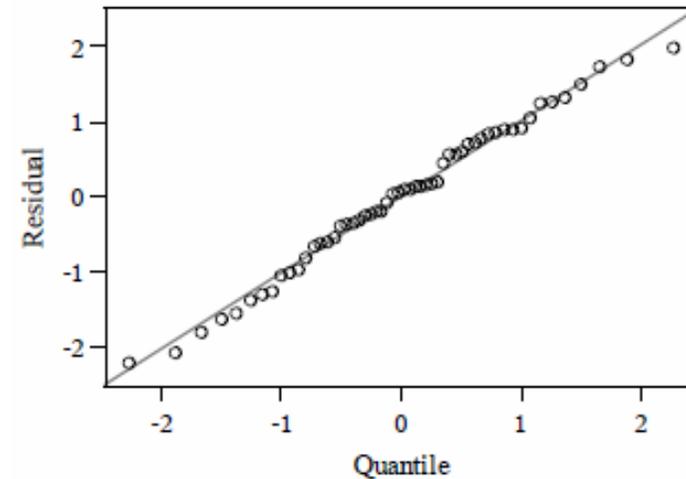
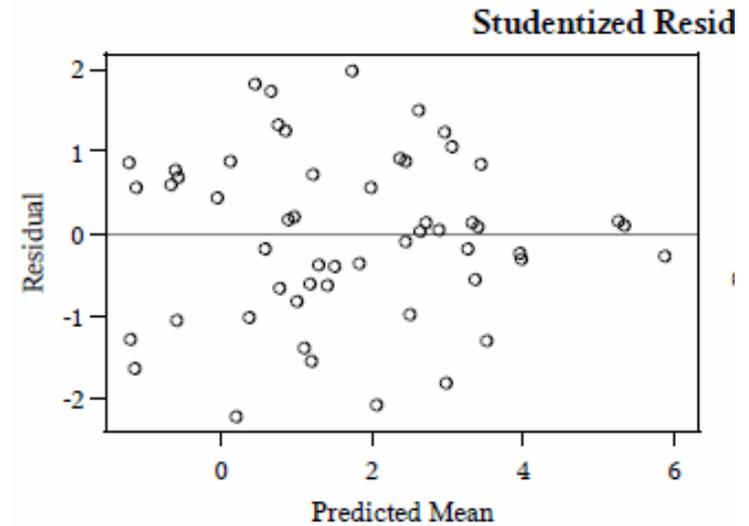


Distribution des erreurs sur les 6 paramètres estimés

ruemax



Log(laideb - 50)



Une Manova sur les paramètres permet d'estimer une matrice de variance-covariance des erreurs



Impact sur la précision de la sortie biomasse

	N	Biais moyen à LM	Ec-type propagé des paramètres à LM	Biais moyen à VB	Ec-type propagé des paramètres à VB
resCP76	1000	-1 250	520	-4290	1428
resF152	1000	152	534	-3224	591
resF160	1000	-2 251	498	-6255	612
resH564	1000	-4 528	549	-7739	1219
resLF66	1000	649	662		
resM376	1000	-5 243	834		
resMEX6	1000	-4 936	567	-8900	1293
resNCO3	1000	4	454	-993	555
resNIF5	1000	-686	571	-5656	1568
resQ110	1000	-923	466		
resR570	1000	-2 830	579	-6822	1729
resR573	1000	-1 613	588	-8797	1924
resR579	1000	-606	619	-6793	1138
resR832	1000	-942	533	-5700	1579
resR932	1000			-7463	1581
resSM81	1000	-258	569		
resSP73	1000			-7312	1601
resSP80	1000			-3530	1033
variance totale		3 580 994	332127	5 114 168	1807153
-1/1000 variance propagée		-332		-1 807	
-1/3 variance résiduelle		-343 791		-941 184	
variance du biais		3 236 872		4 171 177	
éc-type du biais		1 799		2 042	

Production typique : 100 T/ha = 10 000 g/m²



Conclusion d'étape et limites

L'incertitude sur la production de matière sèche engendrée par l'imprécision des paramètres dépend de l'environnement : l'écart-type propagé atteint 5% de la moyenne à La Mare, et entre 5% et 15% à Vue-Belle, suivant les variétés.

Le biais de prédiction est considérable, varie suivant la variété et le lieu.

Pour bien évaluer la distribution du biais de prédiction, il faudrait bien d'avantage de lieux. Pb les essais multilocaux en sont pas arrosés et mal renseignés du point de vue météo... autre source d'imprécision



Bilan du cas d'étude 2: équipe projet

- **Philippe Letourmy**, biostatisticien, Montpellier
- **Eric Gozé**, biostatisticien, Montpellier
- **Jean-François Martiné**, modélisateur (modèle Moticas: modèle dynamique pour la biomasse), La Réunion
- **Jean-Louis Chopart**, modélisateur (modèle Probe: modèle dynamique pour le bilan hydrique), Guadeloupe
- **Lauriane Rouan**, biostatisticienne, Montpellier

- 3 personnes ont œuvré de façon très utile: **Mamadou Diene Diop** en 2010, **Ibrahim Coulibaly** en 2011 et **Oumar Niang** en 2013 (jusqu'en septembre)



Objectifs généraux dans le projet

I – Prédiction de la biomasse aérienne
à 12 mois en tenant compte du facteur eau

II – Etude spécifique de la précision du bilan hydrique

Quelle précision avons-nous sur chacune de ces
prédictions ?



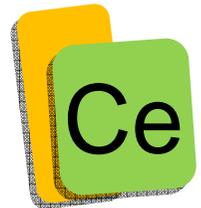
Modèles utilisés dans le projet

- 2 modèles dynamiques à l'échelle journalière:
 - **Mosicas** (objectif I, avec Jean-François Martiné)
 - **Probe** (objectif II, avec Jean-Louis Chopart)
- Les 2 modèles sont utilisés à l'échelle spatiale de la parcelle



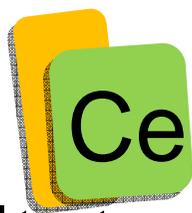
Modèles : Variables d'intérêt

- Pour le modèle **Mosicas**, la matière sèche de la biomasse aérienne à 12 mois (msa)
- Pour le modèle **Probe** les termes du bilan hydrique journalier :
 - Dr** la quantité d'eau en excès dans le sol cultivé (mm),
 - ETR** l'évapotranspiration réelle (mm),
 - ETMc** l'évapotranspiration maximale ajustée (mm),
 - ETR/ETMc** le taux de satisfaction des besoins en eau de la culture



Ce qui a été fait

- Etude de la sensibilité, aux données d'entrée et aux paramètres, de la MSA à 12 mois prédite par **Mosicas** et des 4 termes du bilan hydrique prédits par **Probe** (Mamadou Diène Diop 2010)
- Etude des incertitudes sur la MSA à 12 mois, prédite par **Mosicas**, liées aux données météo (Ibrahim Coulibaly 2011)
- Etude de la l'imprécision des paramètres variétaux de **Mosicas**, de son impact sur la prédiction de la MSA à 12 mois, et du biais de modèle (Oumar Niang 2013)



Ce qui a été fait (2)

Résultats d'Analyse de sensibilité

- Sur les paramètres utilisés pour étudier la sensibilité de la matière sèche aérienne avec **Mosicas**, les plus influents étaient des coefficients liés à la conversion du rayonnement, alors que les paramètres de LAI étaient peu sensibles, sauf un, et de sensibilité très variable selon les conditions de l'essai
- Pour le modèle **Probe**, les entrées ou les paramètres les plus influents étaient l'ETP et le coefficient cultural (K_c), et, dans une moindre mesure, la cote de drainage (CD) et la réserve maximale en eau utile (RU)



Ce qui a été fait (3)

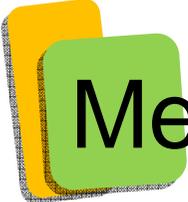
Résultats de l'étude d'incertitude

- Analyse de la distribution des données météo :
 - On a constaté que l'altitude était un facteur important à prendre en compte, apparemment plus que l'éloignement
 - De plus il était important de différencier les 2 bassins nord-est et sud-ouest de l'île
- Effet du remplacement d'une station par une autre :
 - Le biais sur la MSA à 12 mois (simulée avec **Mosicas**) était d'autant plus important que la distance et la différence d'altitude entre les stations étaient grandes
 - La côte est présentait une variabilité importante de la pluviométrie, même à faible distance, ce qui avait des conséquences importantes sur la prévision de la biomasse



Questions et Perspectives

- **Sur Mosaicas :**
 - Estimer l'incertitude globale, due aux entrées, aux paramètres et au modèle (fin du stage d'Oumar Niang)
 - Publication en vue sur la prise en compte de l'effet variétal: est-il utile d'estimer des paramètres variétaux pour faire une prédiction efficace ?
 - Etait-il préférable d'estimer les paramètres par les moindres carrés généralisés, plutôt que séquentiellement ?
 - Modèles alternatifs : EcoMeristem (modèle structure-fonction), Dssat-Canegro (autre big leaf) ?
- **Sur Probe :**
 - Etudier la possibilité d'améliorer la précision de la MSA à 12 mois avec le bilan hydrique de Probe



Merci de votre attention



Quand la fonction-modèle est de classe C^2

$$\begin{aligned}y &= f(x, \vartheta) + b(x, \vartheta) + \varepsilon' \\ &= f(x_{obs}, \hat{\vartheta}) + \left(f(x, \vartheta) - f(x_{obs}, \hat{\vartheta}) \right) + b(x, \vartheta) + \varepsilon'\end{aligned}$$

$$E(y - f(x_{obs}, \hat{\vartheta})) = \left(\frac{\partial y}{\partial \theta} \right)' E(\vartheta - \hat{\vartheta}) + b(x, \vartheta) + 0$$

$$\text{var}(y - f(x_{obs}, \hat{\vartheta})) = \left(\frac{\partial y}{\partial \theta} \right)' \Sigma_{\theta} \left(\frac{\partial y}{\partial \theta} \right) + \left(\frac{\partial y}{\partial x} \right)' \Sigma_x \left(\frac{\partial y}{\partial x} \right) + \text{var}(b) + \text{var}(\varepsilon')$$

L'on a négligé le biais et l'imprécision du modèle, que l'on ne peut estimer que par validation sur des essais multilocaux où les entrées du modèle sont mesurées.