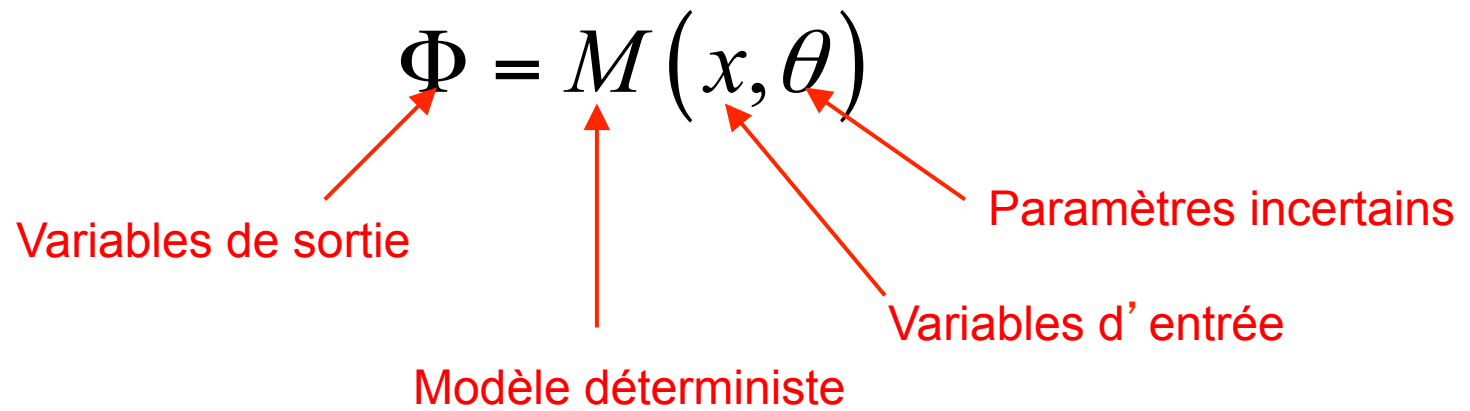


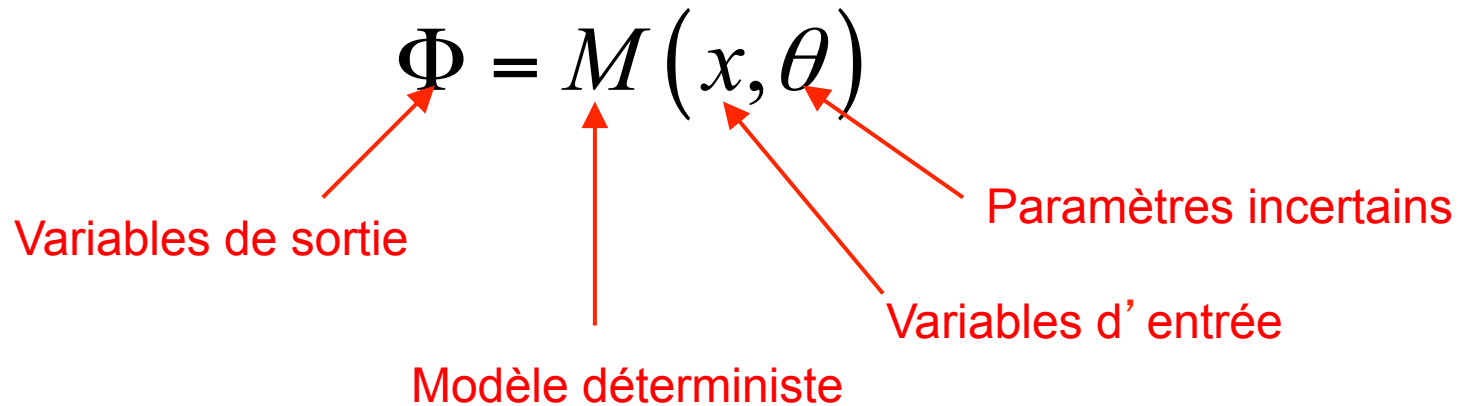
Combinaison de différentes sources d'information par Bayesian melding

Application à la modélisation du LAI

Contexte



Contexte



Exemples:

Variables de sortie: Rendement, LAI, etc.

Variables d'entrée: Température, rayonnement, stress hydrique etc.

Paramètres: LAI_{max}, ϵ_{imax} etc.

Contexte (suite)

Trois types d'information disponibles pour estimer θ :

- **Information a priori sur θ** (valeurs plausibles des paramètres)
ex: publications, valeurs estimées à partir de données passées etc.
- **Information a priori sur Φ** (valeurs plausibles des sorties)
ex: contraintes biologiques, gammes de valeurs observées sur le terrain etc.
- **Valeurs mesurées des sorties Φ**
ex: nouvelles expérimentations

Comment combiner ces trois types d'information pour estimer les paramètres θ ?

Méthode: Bayesian melding

Poole, Raftery. 2000. Inference for deterministic simulation models: The Bayesian melding approach. JASA 95, 1244-1255.

Représentation des connaissances sous forme de distributions de probabilité:

- Densité *a priori* des paramètres incertains θ : $q_1(\theta)$
- Densité *a priori* des sorties Φ : $q_2(\Phi)$
- Vraisemblance reliant les observations à Φ : $L(\Phi)$
- Estimation de la distribution *a posteriori* des paramètres θ

Il peut y avoir une vraisemblance reliant les observations directement à θ .

Méthode: Bayesian melding (suite)

L' *a priori* sur les paramètres incertains θ , $q_1(\theta)$, induit un *a priori* sur les sorties Φ . Sa densité est notée $q_1^*(\Phi)$.

Il existe donc deux *a priori* sur les sorties Φ :

- un *a priori* induit, de densité $q_1^*(\Phi)$
- un *a priori* direct, noté $q_2(\Phi)$.

Combinaison des deux *a priori* sur les sorties Φ :

$$q(\Phi) \propto q_1^*(\Phi)^\alpha q_2(\Phi)^{1-\alpha}$$

Combiné Induit Direct Pondération (e.g 0.5)

Algorithme

Poole, Raftery. 2000. Inference for deterministic simulation models: The Bayesian melding approach. JASA 95, 1244-1255.

Sampling-Importance-Resampling

1. Génération d'une série de valeurs de paramètres à partir de $q_1(\theta)$
2. Calcul des sorties correspondantes du modèle $M(\theta)$
3. Estimation non-paramétrique de $q_1^*(\Phi)$
4. Calcul des poids pour chaque valeur générée du vecteur de paramètres en utilisant la vraisemblance L

$$w_i = \left[\frac{q_2(M(\theta_i))}{q_1^*(M(\theta_i))} \right]^{1-\alpha} L(M(\theta_i))$$

5. Re-échantillonnage des valeurs des paramètres à partir des poids.

Illustration de la méthode

Modèle LAI (Leaf Area Index)

$$\text{LAI}(\theta) = L_{\max} \left\{ \frac{1}{1 + e^{-A[T-T_1]}} - e^{B[T-T_2]} \right\}.$$

$$T_2 = \frac{1}{B} \log[1 + \exp(A \times T_1)]$$

T : somme de température

L_{\max} , T_1 , A , B : quatre paramètres

Baret, 1986

Objectif : estimer les paramètres pour deux génotypes

Informations disponibles pour le génotype 1

Distribution a priori uniforme pour les quatre paramètres

Parameter	Unit	Uncertainty range	
L_{\max}	-	7	12
T_1	C	700	1100
A	-	0.003	0.01
B	-	0.0011	0.003

Distribution a priori uniforme pour le LAI à T =700 °C jour

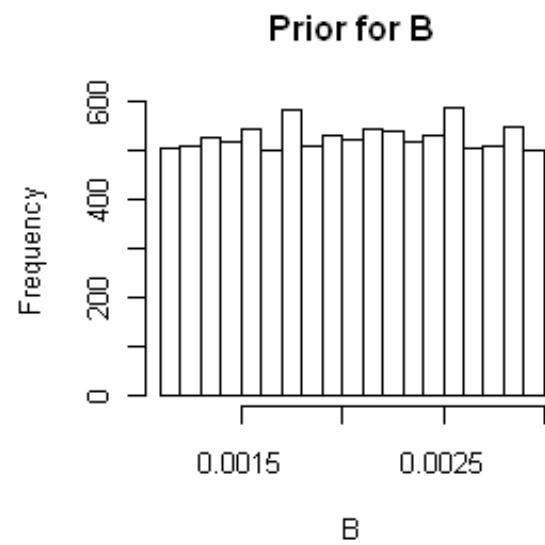
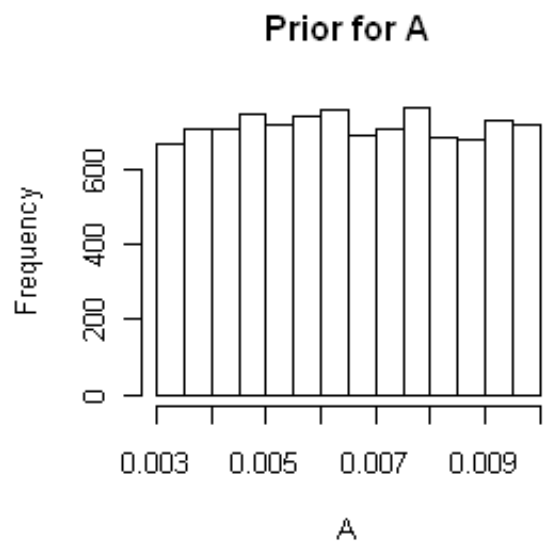
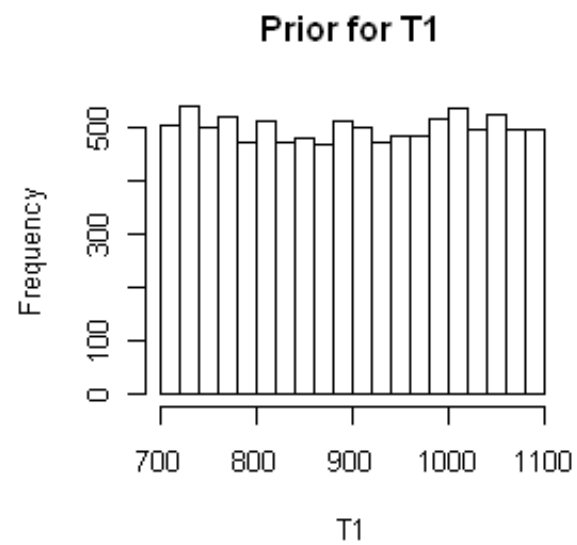
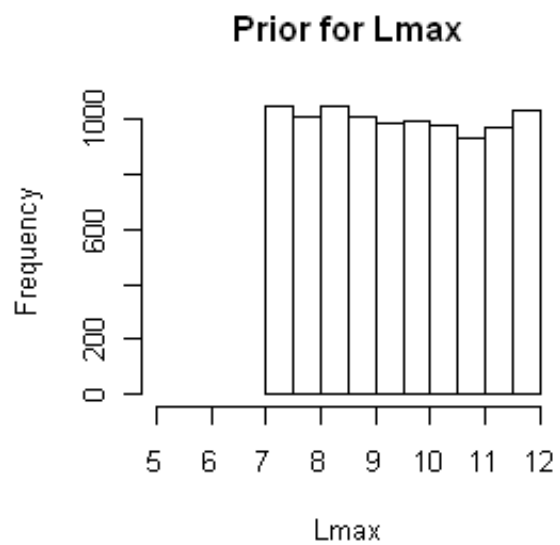
LAI ~ Unif (3, 6)

Trois observations

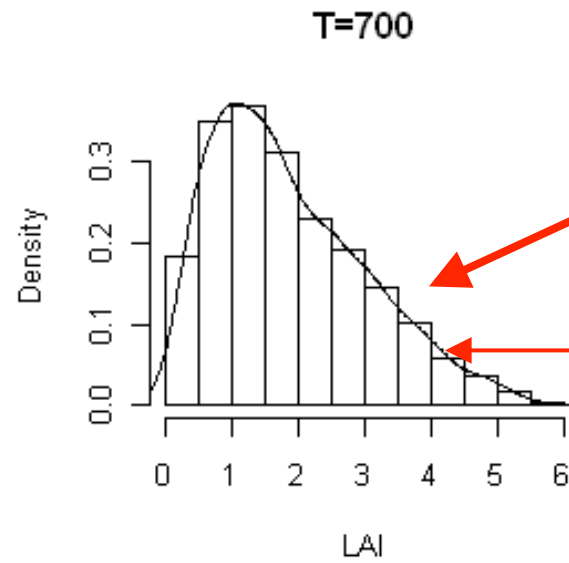
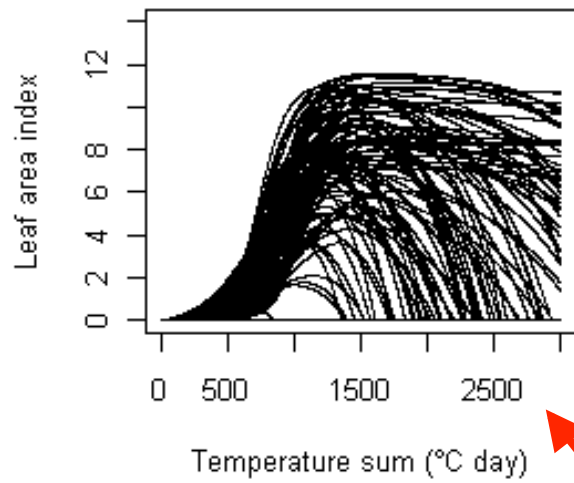
Y1 = 2.1 (+-0.53) à T=500 °C jour

Y2 = 4.5 (+- 0.6) à T=1000 °C jour

Y3 = 6.2 (+- 0.7) à T=1500 °C jour



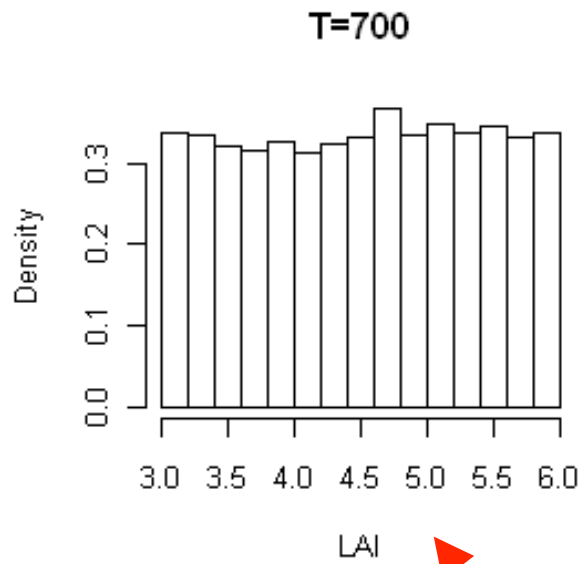
génotype 1



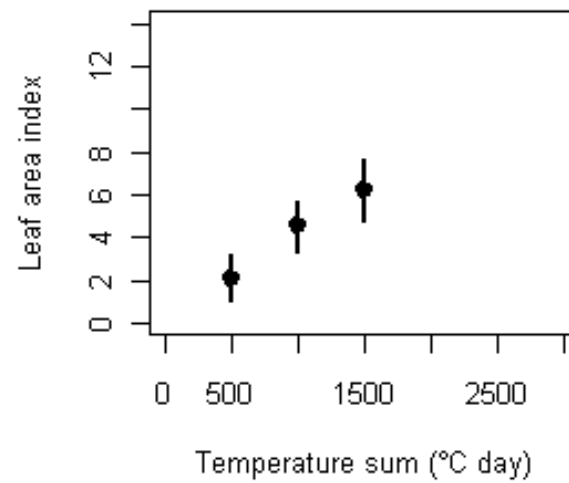
a priori induit par les paramètres sur le LAI

Estimation non-paramétrique de la densité

Réponses induites par l'a priori sur les paramètres

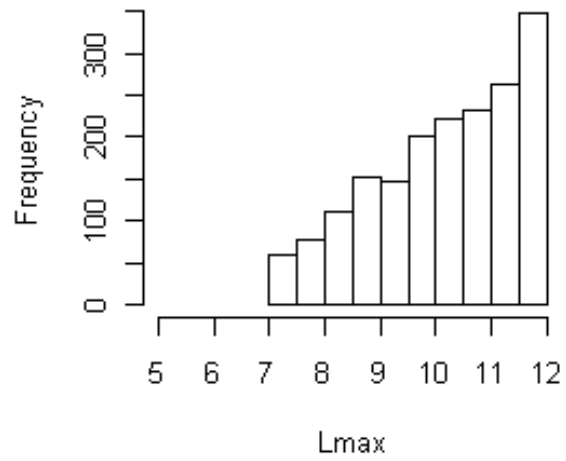


a priori direct sur le LAI

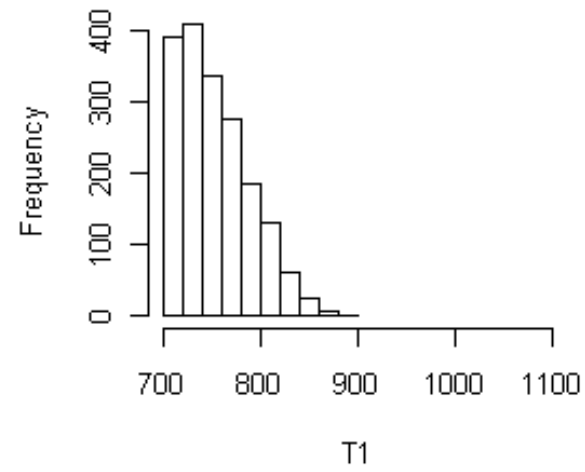


Données

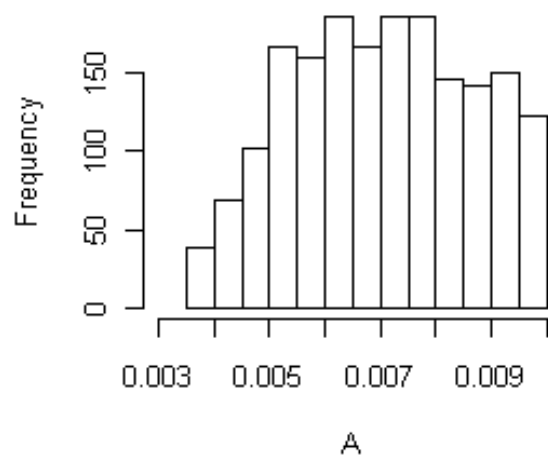
Posterior for Lmax



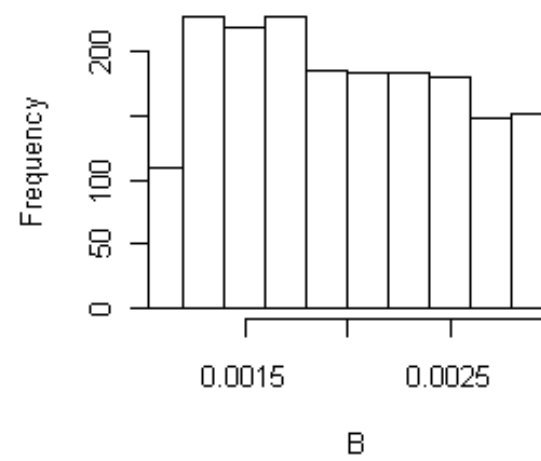
Posterior for T1



Posterior for A



Posterior for B



Informations disponibles pour le génotype 2

Distribution a priori uniforme pour les quatre paramètres

Parameter	Unit	Uncertainty range	
L_{\max}	-	5	10
T_1	C	700	1100
A	-	0.003	0.01
B	-	0.0011	0.003

Distribution a priori uniforme pour le LAI à T =700 °C jour

LAI ~ Unif (1, 4)

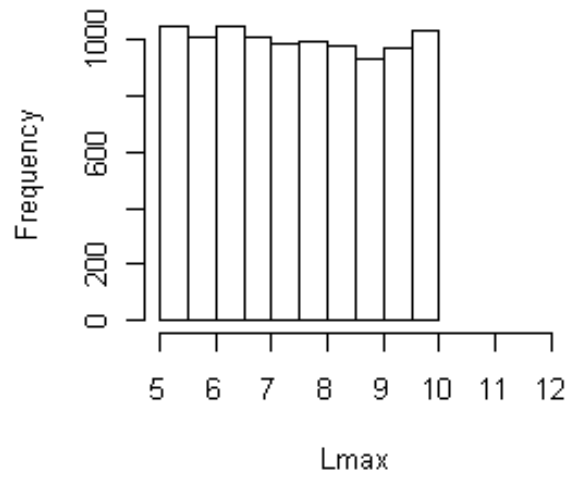
Trois observations

Y1 = 1.2 (+-0.4) à T=500 °C jour

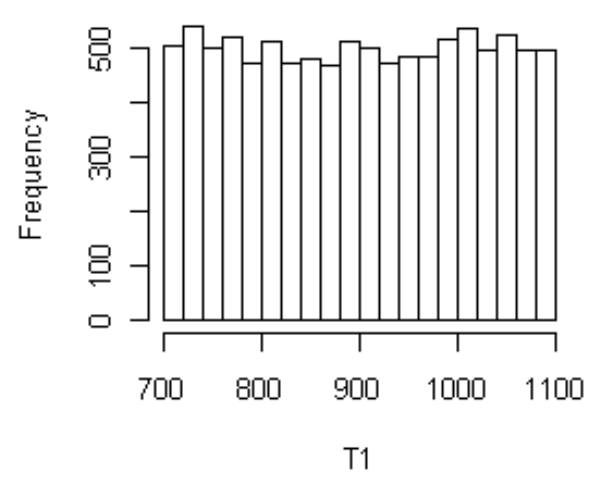
Y2 = 3.5 (+- 0.45) à T=1000 °C jour

Y3 = 5 (+- 0.5) à T=1500 °C jour

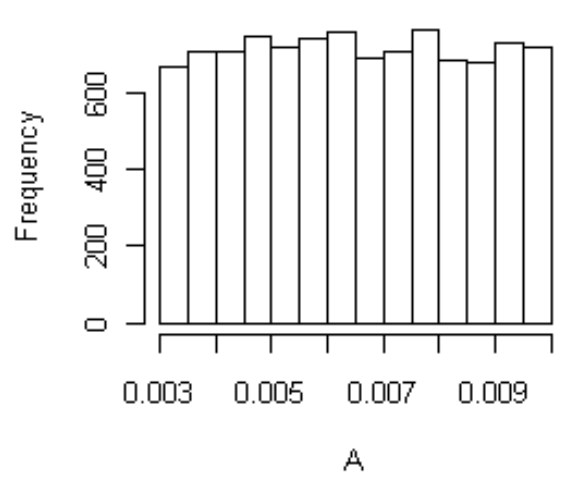
Prior for Lmax



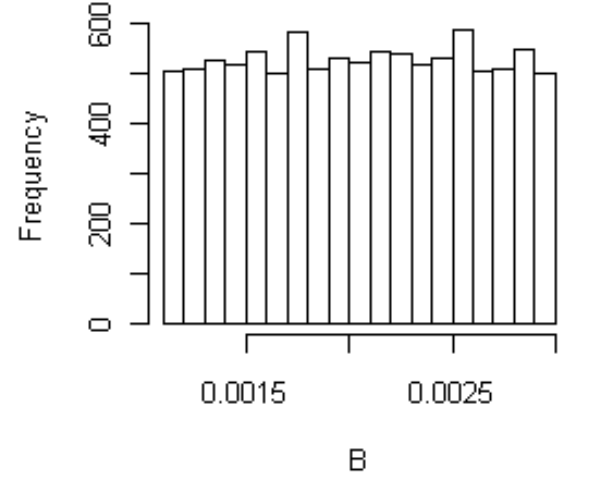
Prior for T1



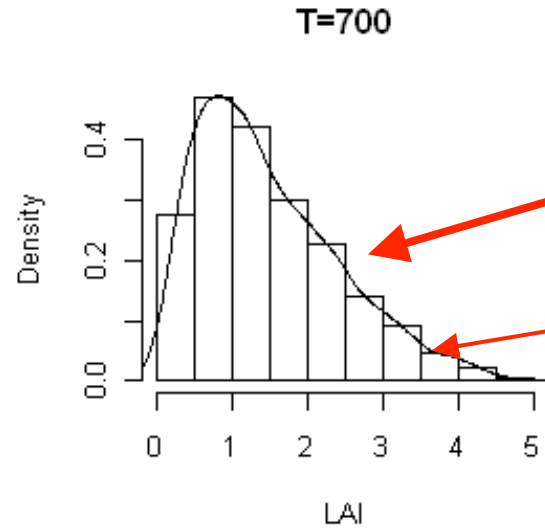
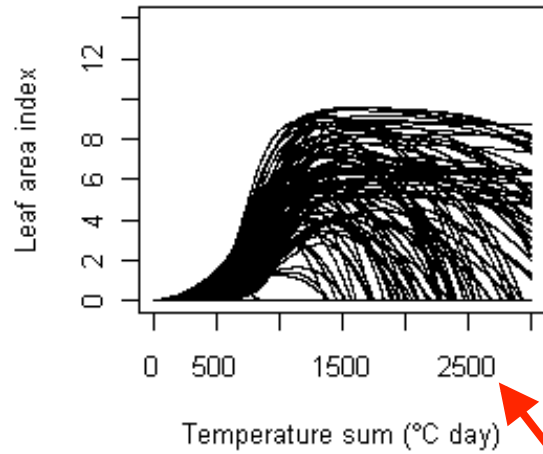
Prior for A



Prior for B



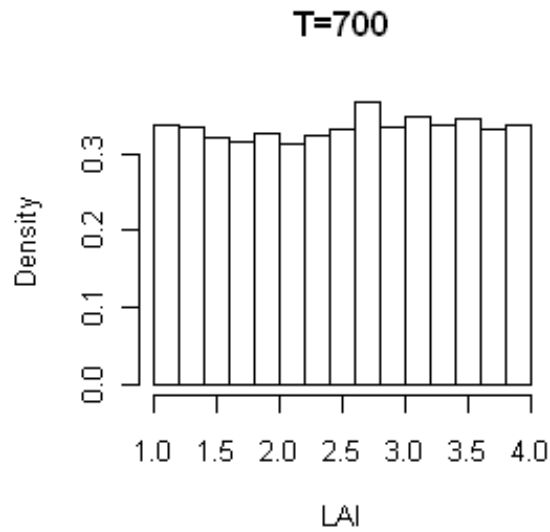
génotype 2



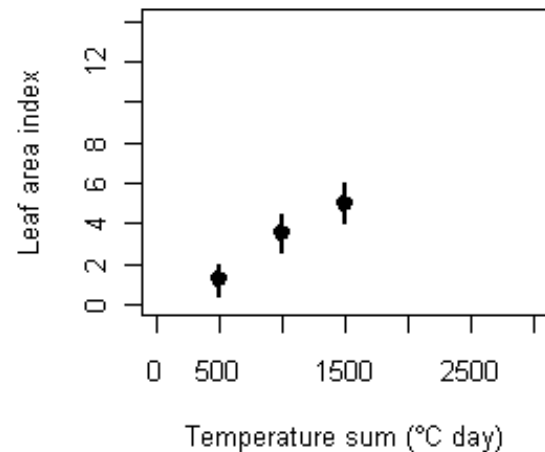
a priori induit par les paramètres sur le LAI

Estimation non-paramétrique de la densité

Réponse induite par l'a priori sur les paramètres

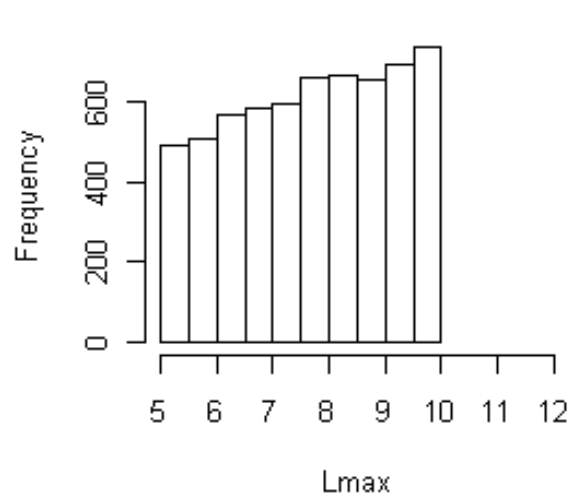


a priori direct sur le LAI

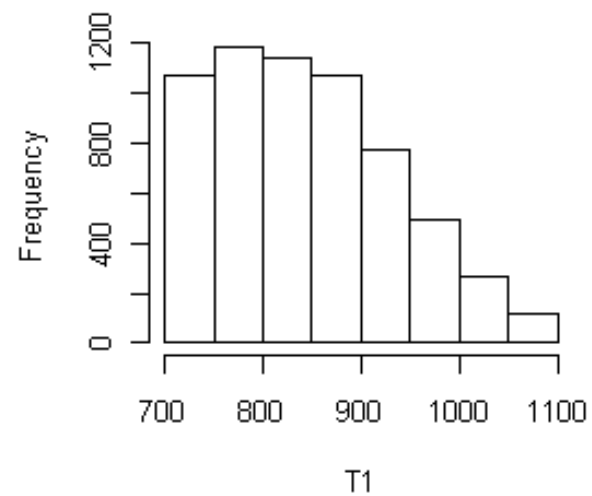


Données

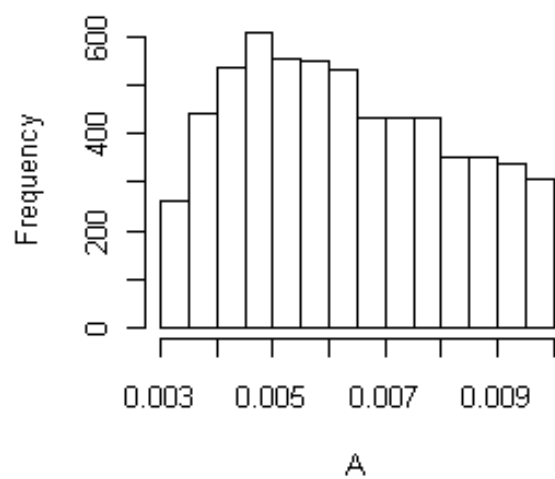
Posterior for Lmax



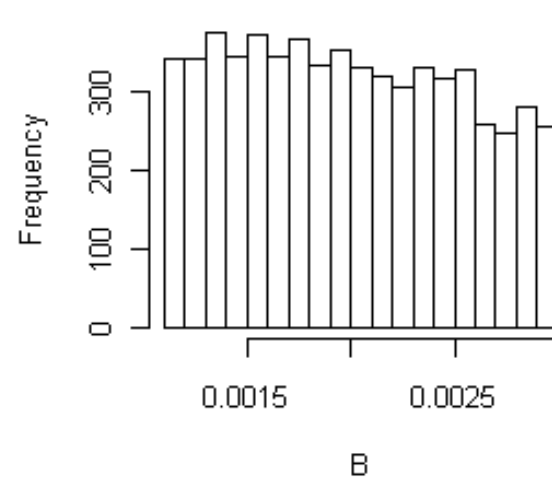
Posterior for T1



Posterior for A



Posterior for B



Moyennes et écart-types des distributions a posteriori des paramètres obtenus pour les deux génotypes

Parameter	Unit	Genotype 1		Genotype 2	
		Mean	SD	Mean	SD
L_{\max}	-	10.16	1.31	7.67	1.42
T_1	C	752.7	36.42	844.35	91.38
A	-	0.0071	0.0016	0.0063	0.0019
B	-	0.00198	0.00054	0.002	0.0005

Conclusion / Perspective

- Les algorithmes développés pour appliquer les méthodes statistiques bayésiennes ouvrent de nouvelles possibilités.
- Ces méthodes permettent de synthétiser différentes sources d'information et tenant compte de leurs niveaux de précision.
- Il est possibles de combiner des informations issues des théories biologiques, des réseaux expérimentaux, des publications scientifiques, des connaissances de terrain.
- Perspective:
 - application à divers modèles
 - évaluation de l'intérêt de la démarche