Ecole chercheur ASPEN

Avril 2013, Les Houches, France

Analyse d'incertitude, analyse de sensibilité. Objectifs et principales étapes

David Makowski

INRA

makowski@grignon.inra.fr

MEXICO

1. **Définitions et objectifs**
2. **Analyse d’incertitude**
3. **Analyse de sensibilité**
4. **Etude de cas**

1. Définitions et objectifs

Sources d'incertitude dans un modèle

$$G\left(x,\theta\right)$$

Equations | **Variables d'entrée** | **Paramètres**

Types d'incertitude

• ***Manque de connaissance***

Ex: Température optimale pour le développement d'un champignon pathogène

• ***Erreur de mesures / Echantillonnage***

Ex: Erreur de mesure de la densité de plantes dans une parcelle agricole

• ***Variabilité des caractéristiques du système***

Ex: Variabilité de la « température moyenne journalière » entre années

Notation

z = variables d'entrée et paramètres incertains

= facteurs incertains

$z = (z_1, z_2, \ldots, z_p)$

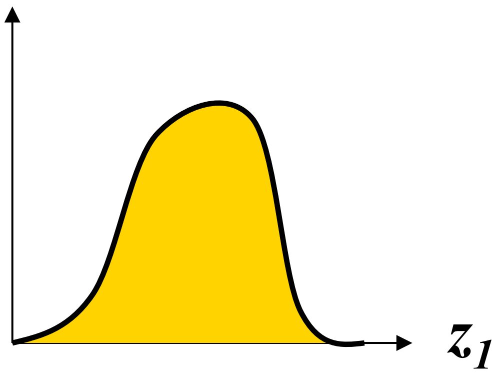
Sortie du modèle $\quad y\left(z_1, z_2, .., z_p\right) = y(z)$

Analyse d'incertitude

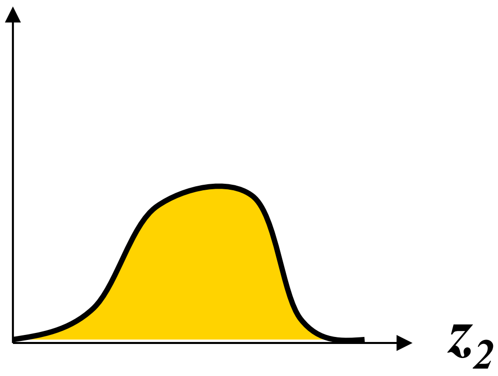
Permet de répondre à la question suivante:

« Quel est le niveau d'incertitude dans $y(z)$ qui résulte de l'incertitude dans z ? »

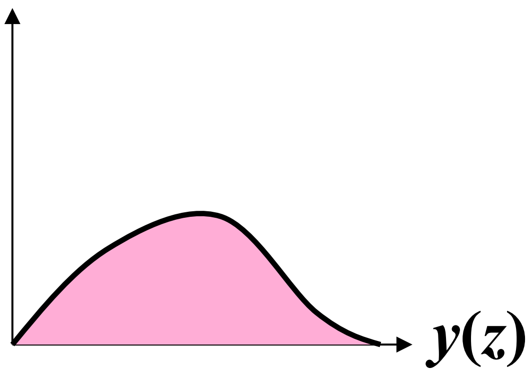
On a



et



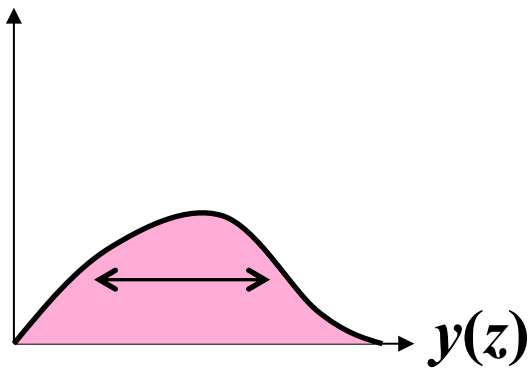
On veut déterminer



Analyse de sensibilité

Son objectif est de répondre à la question:

« Quelles sont les principales sources d'incertitude parmi z_1, z_2, …, z_p ? »



Variance de $y(z)$ = effet de z_1 + effet de z_2 + …

Intérêt pratique

de l'analyse d'incertitude

- donner des informations sur l'incertitude associée aux prédictions d'un modèle
- optimiser des variables décisionnelles

de l'analyse de sensibilité

- identifier les paramètres et les variables d'entrée qui ont une forte influence sur les sorties d'un modèle

→ ***Important de les connaître avec précision***

- identifier les paramètres et les variables d'entrée qui ont une influence moindre sur les sorties

→ ***Moins important de les connaître avec précision***

Exemples de questions pouvant être traitées par AI ou AS

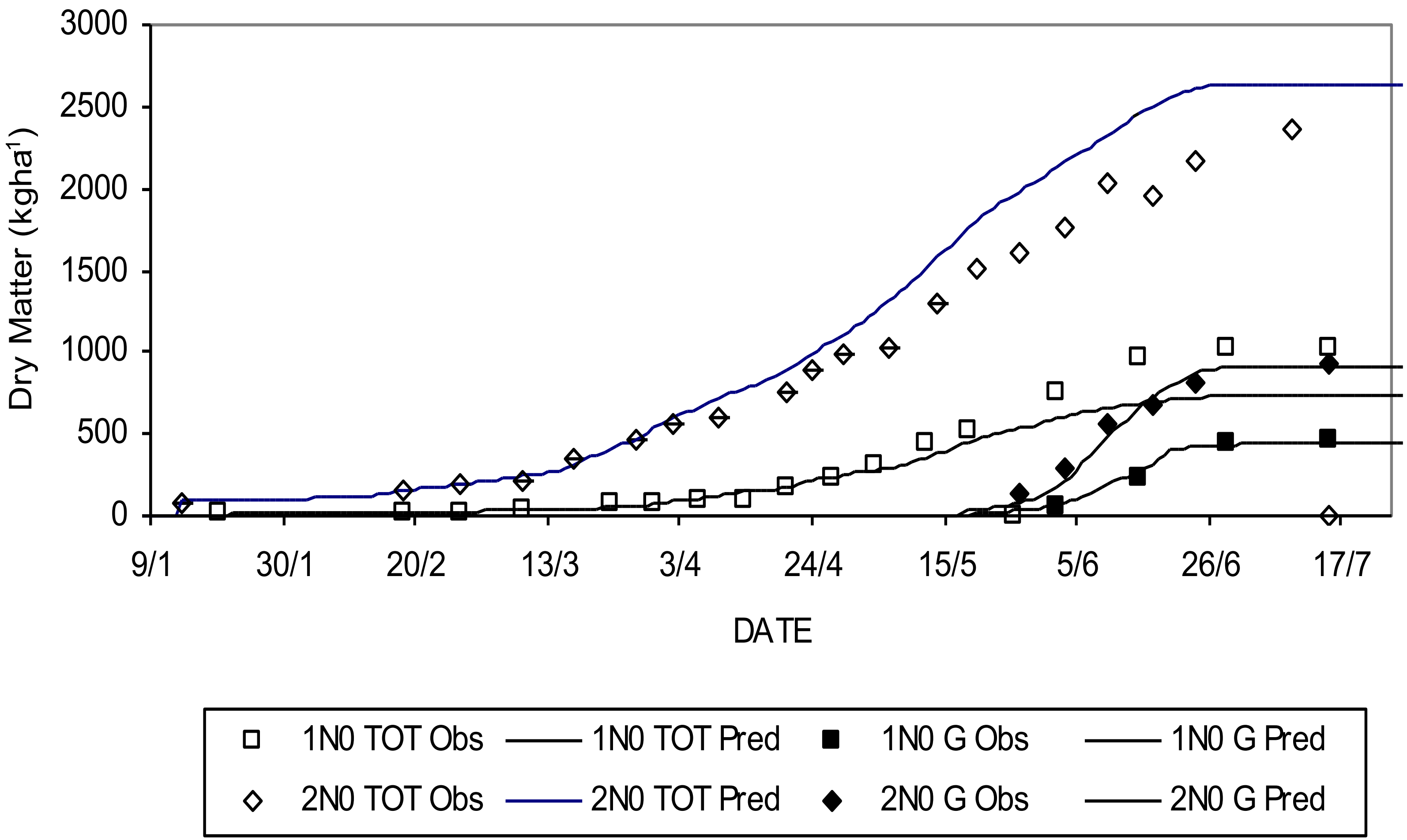
- Est-il important de mesurer précisément les caractéristiques du sol pour prédire le rendement d'une culture ?

- Probabilité qu'une nouvelle mesure de gestion du stock de langoustines soit plus efficace que la mesure actuelle ?

- Quelle est la probabilité de perdre plus de 0.2 t ha^{-1} si la dose d'engrais appliquée sur du blé est réduite de 20%?

- Quels sont les paramètres d'un modèle de culture à estimer en priorité génotype par génotype ?

Simulations de la biomasse du blé à l'aide du modèle dynamique AZODYN



Variables d’entrée

- caractéristiques du sol
- données climatiques
- pratiques agricoles

AZODYN

Paramètres

Biomasse

Rendement

Teneur en protéines des grains

N résiduel du sol…

Jeuffroy et Recous, 1999

Incertitude associée à 13 paramètres potentiellement génotypiques

Parameter	Definition	Range	Unit
RDTMAXVAR	Maximal yield	10.0 - 13.7	$t.ha^{-1}$
Ebmax	Radiation use efficiency	2.7-3.3	$g.MJ^{-1}$
D	Ratio of leaf area index to critical nitrogen	0.02-0.045	-
REM2	Fraction of remobilized nitrogen	0.5-0.9	-
K	Extinction coefficient	0.6-0.8	-
Eimax	Ratio of intercepted to incident radiation	0.9-0.99	
Tep.flo	Duration betw een earing and flow ering	100-200	°C.day
R	Ratio of total to above ground nitrogen	1.0-1.5	-
P1GMAXVAR	Maximal w eight of one grain	47-65	mg
Lambda	Parameter for calculating nitrogen use efficiency	25-45	-
Mu	Parameter for calculating nitrogen use efficiency	0.6-0.9	-
DJPF	Temperature threshold	150-250	°C.day
NGM2MAXVAR	Maximal grain number	107.95-146.05	-

Quels paramètres doit-on estimer ?

Exp. 1

Génotype
« Soissons »

Paramètres pour « Soissons »

Exp. 2

Génotype
« Recital »

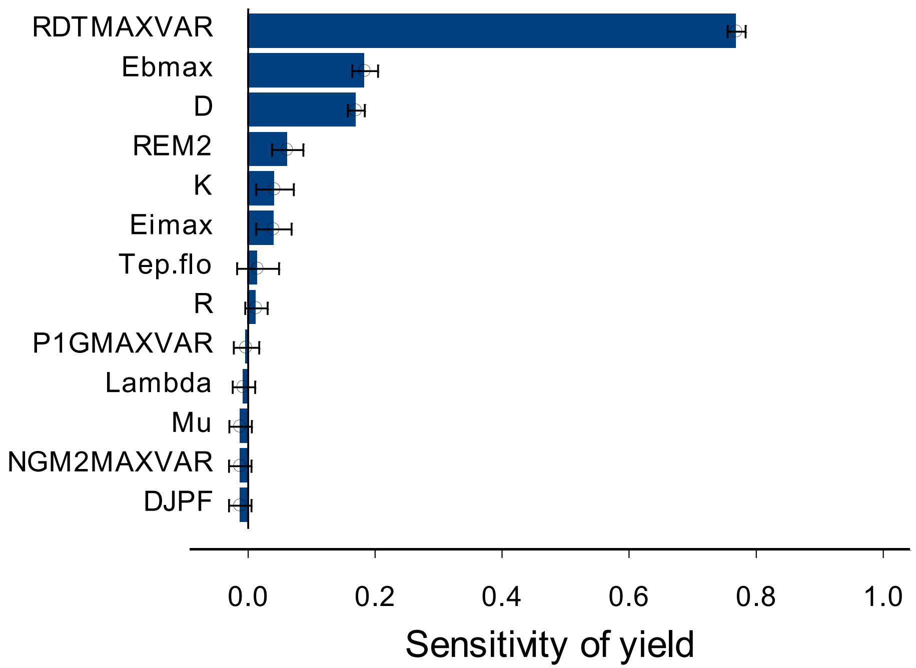
Paramètres pour « Récital »

…

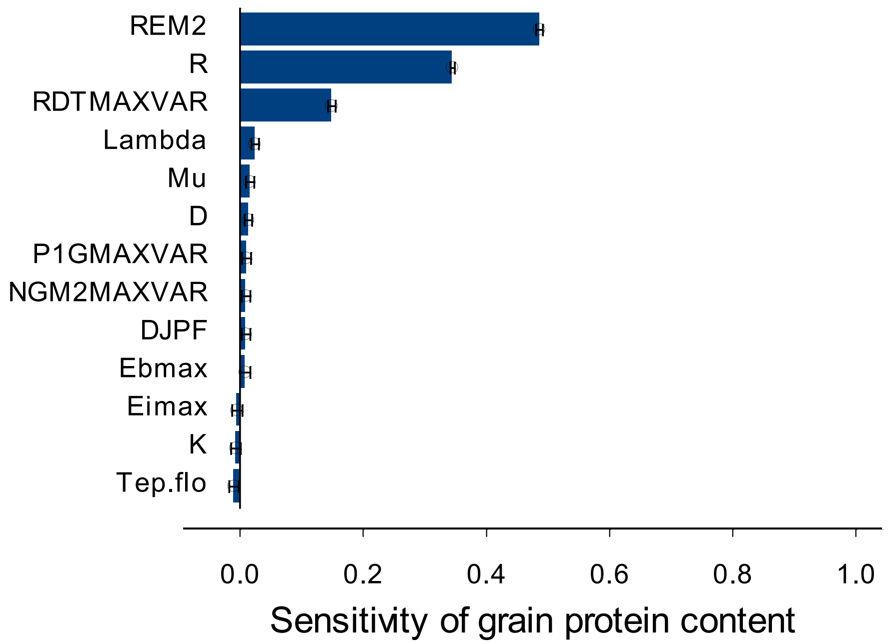
Coûteux !

Indices de sensibilité totale pour les simulations de rendement et de teneur en protéines

Rendement



Teneur en protéines



Makowski et al. 2005

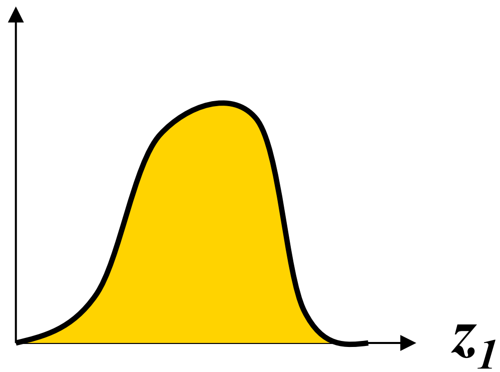
2. Analyse d'incertitude

Analyse d'incertitude

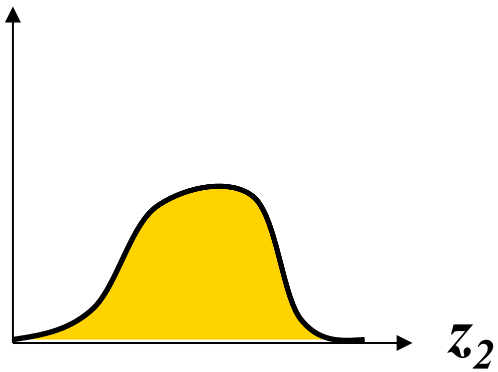
Permet de répondre à la question suivante:

« Quel est le niveau d'incertitude dans $y(z)$ qui résulte de l'incertitude dans z ? »

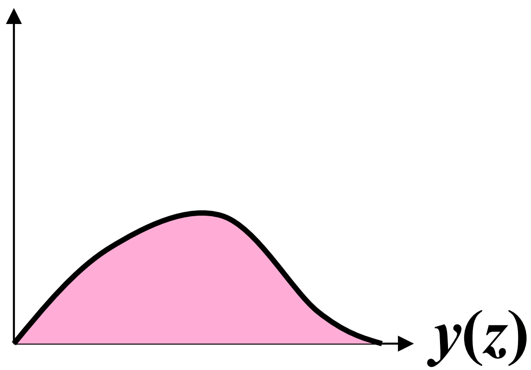
On a



et



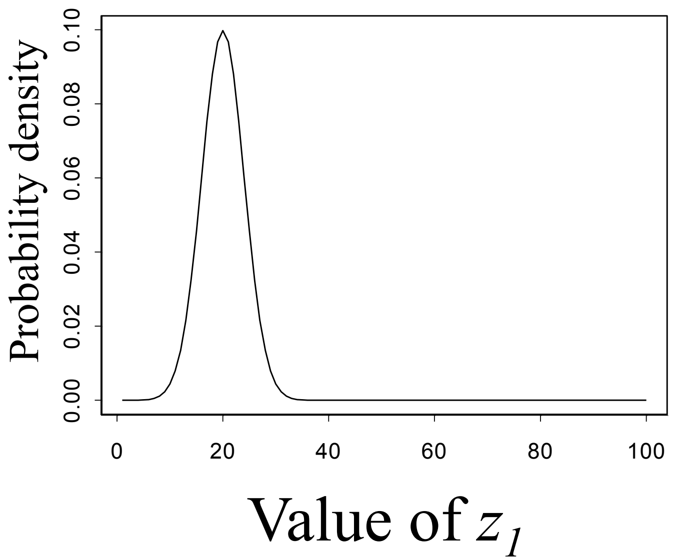
On veut déterminer

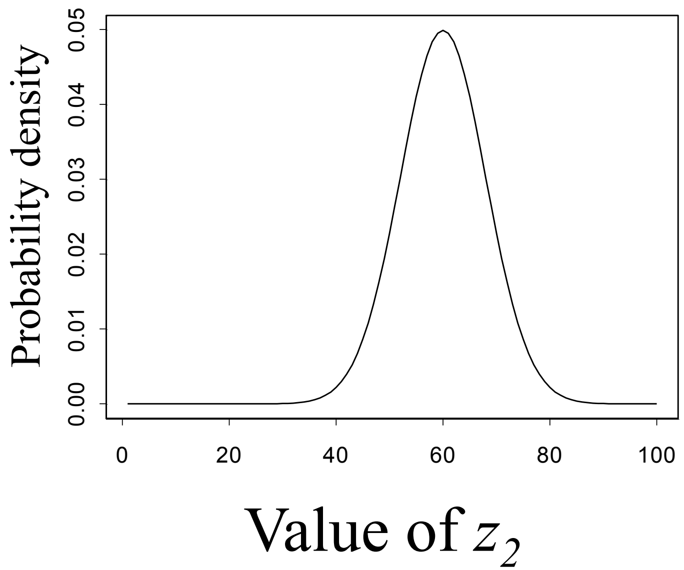


Application à un modèle très simple

Equation: $y(z_1, z_2) = z_1 + 2\, z_2$

Incertitude sur z_1 et z_2 : $z_1 \sim N(20, 16)$ et $z_2 \sim N(60, 64)$





Question: Réaliser une analyse d'incertitude

Application à un modèle très simple

« Vous devez déterminer la distribution de probabilité de $y(z_1, z_2)$ à partir des distributions de z_1 et z_2 » .

Propriétés:

Si z_1 et z_2 sont deux variables indépendantes de distribution Gaussienne alors

$A\,z_1 + B\,z_2$ suit une distribution Gaussienne

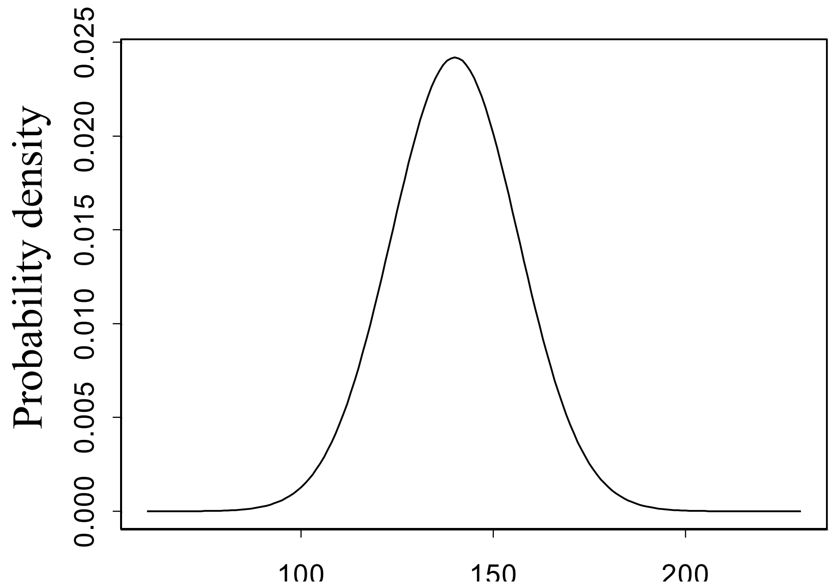
$$E(A\,z_1+B\,z_2)=A\,E(z_1)+B\,E(z_2)$$

$$\mathrm{var}(A\,z_1+B\,z_2)=A^2\,\mathrm{var}(z_1)+B^2\,\mathrm{var}(z_2)$$

Application à un modèle très simple

Pour ce modèle simple, on peut déterminer l'expression exacte de $y(z_1,z_2)$:

$$y(z_1,z_2) \sim N(140, 272)$$



Value of $y(z_1, z_2)$

En général, c'est plus dur !

• **Equations plus complexes, relation non linéaire entre** $y(z)$ **et** z

→ Pas possible de déterminer l'expression analytique de la distribution de $y(z)$

• **La distribution de** z **n'est pas toujours connue**

→ Choix subjectif

• **Temps de calcul parfois long avec certains modèles**

→ Le nombre de simulations est limité

Quatre étapes

1. **Définir les distributions de $z_1, \ldots, z_p$.**

2. **Générer des échantillons à partir des distributions définies à l'étape 1**

3. **Calculer $y(z)$ pour chaque série de $z_1, \ldots, z_p$ générée**

4. **Estimer la distribution de $y(z)$**

Étape 1. Définition des distributions

Les distributions de probabilité des facteurs incertains (paramètres ou variables d'entrée) peuvent être définies en utilisant :

- **La littérature scientifique et l'expertise**

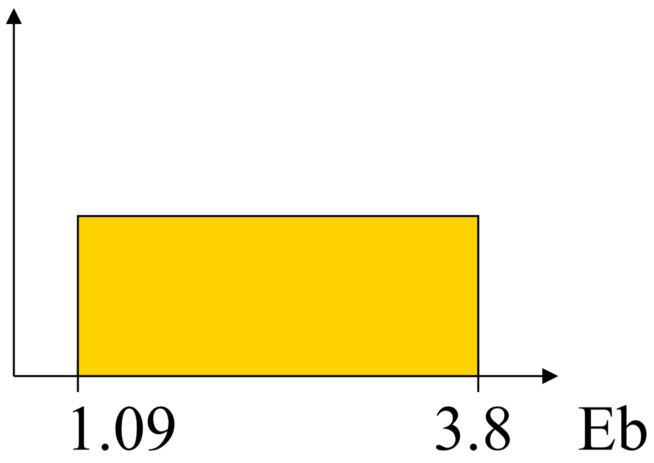
- **Des séries de mesures (série climatique…)**

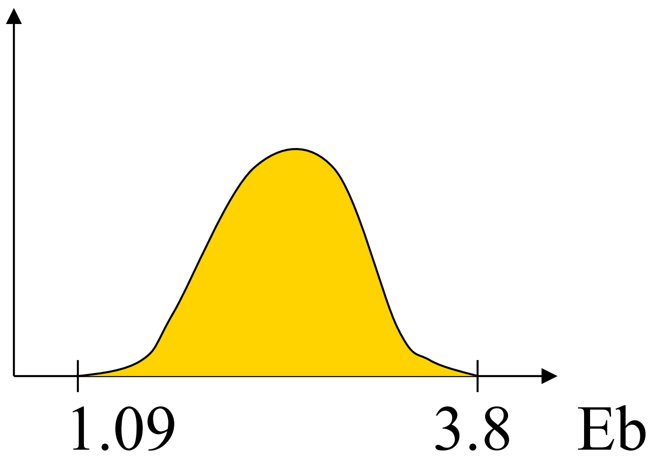
- **Les valeurs des paramètres estimées**

Étape 1. Définition des distributions

Exemple:

d'après un article publié par Jeuffroy et Recous en 1999 dans EJA, l'efficacité d'utilisation de rayonnement intercepté varie entre **1.09 et 3.8 g.MJ^{-1}** pour le blé





Parfois, plusieurs choix sont possibles

1. Définition des distributions de $z_1, \ldots, z_p$.

2. Génération d'échantillons à partir des distributions définies à l'étape 1.

Étape 2. Génération d'échantillons à partir des distributions de $z_1, \ldots, z_p$

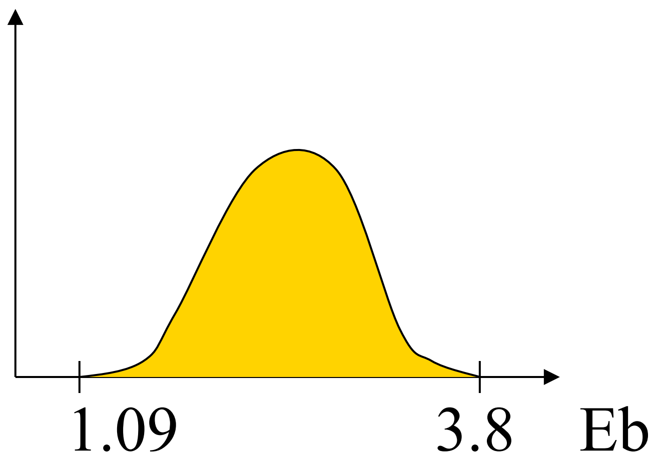
• Il faut générer suffisamment de valeurs de $z_1, z_2, \ldots, z_p$

• Différentes méthodes d'échantillonnage peuvent être utilisées:

- échantillonnage aléatoire
- échantillonnage en hypercube latin
- …

• En pratique, on utilise un logiciel pour générer N valeurs de $z_1, z_2, \ldots, z_p$ (ex: N=20000).

Étape 2. Génération d'échantillons à partir des distributions de $z_1, \ldots, z_p$



On génère un échantillon de valeurs de Eb issues de sa distribution :

1.2, 1.9, 2.1, 2.2, 2.3, 2.5, 2.7, 3.1, 3.7…

Étape 2. Génération d'échantillons à partir des distributions de $z_1, \ldots, z_p$

	z_1	z_2	…	z_p
Série 1	1.21	0.85	…	0.99
Série 2	1.97	0.72	…	0.92
…	…	…	…	…
Série N	3.70	0.75	…	0.91

1. **Définition des distributions de z_1, …, z_p.**

2. **Génération d'échantillons à partir des distributions définies à l'étape 1.**

3. **Calcul de $y(z)$ pour chaque série z_1, …, z_p générée.**

Étape 3. Calcul de $y(z)$ pour chaque série de $z_1, \ldots, z_p$ générée

• La difficulté de cette étape dépend du niveau de complexité du modèle.

• Le temps de calcul peut être long avec certains modèles particulièrement complexes.

Étape 3. Calcul de $y(z)$ pour chaque série $z_1, \ldots, z_p$ générée

	z_1	z_2	…	z_p	$y(z)$
Série 1	1.21	0.85	…	0.99	90.9
Série 2	1.97	0.72	…	0.92	95.2
…	…	…	…	…	…
Série N	3.70	0.75	…	0.91	81.5

1. Définition des distributions de $z_1, \ldots, z_p$.

2. Génération d'échantillons à partir des distributions définies à l'étape 1.

3. Calcul de $y(z)$ pour chaque série $z_1, \ldots, z_p$ générée.

4. Approximation de la distribution de $y(z)$.

Étape 4. Approximation de la distribution de $y(z)$

- **Décrire les N valeurs de $y(z)$ calculées à l'étape 3.**

- **Étape souvent assez facile.**

- **Différentes approches possibles**

  - calcul de la moyenne et de la variance,

  - calcul de quantiles (quartiles, déciles…),

  - histogramme,

  - fonction de distribution cumulée,

  - box plot …

Application au modèle simple

• Approche en 4 étapes pas nécessaire pour ce modèle car on peut calculer analytiquement la distribution de $y(z_1, z_2)$

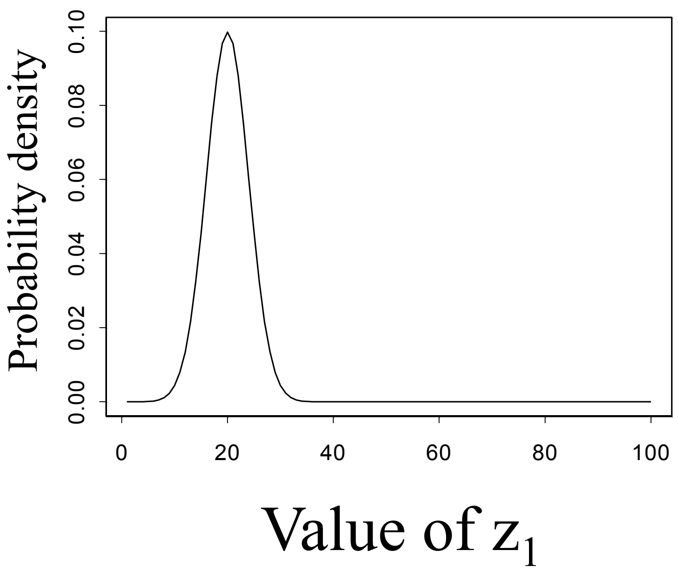
• On applique cette approche à ce modèle uniquement pour montrer qu'elle marche bien.

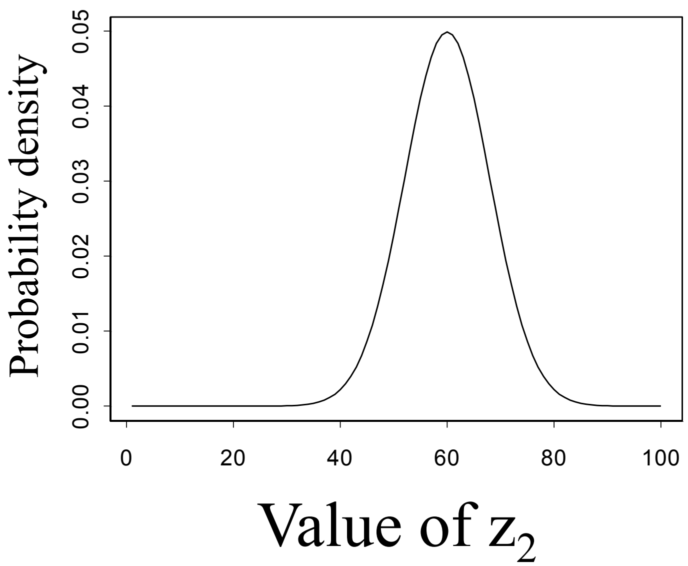
Application au modèle simple

Etape 1

Equation : $y(z_1, z_2) = z_1 + 2\, z_2$

Incertitude sur z_1 et z_2 : $z_1 \sim N(20, 16)$, $z_2 \sim N(60, 64)$





Application au modèle simple

Etape 2

- N valeurs de z_1 et z_2 sont générées

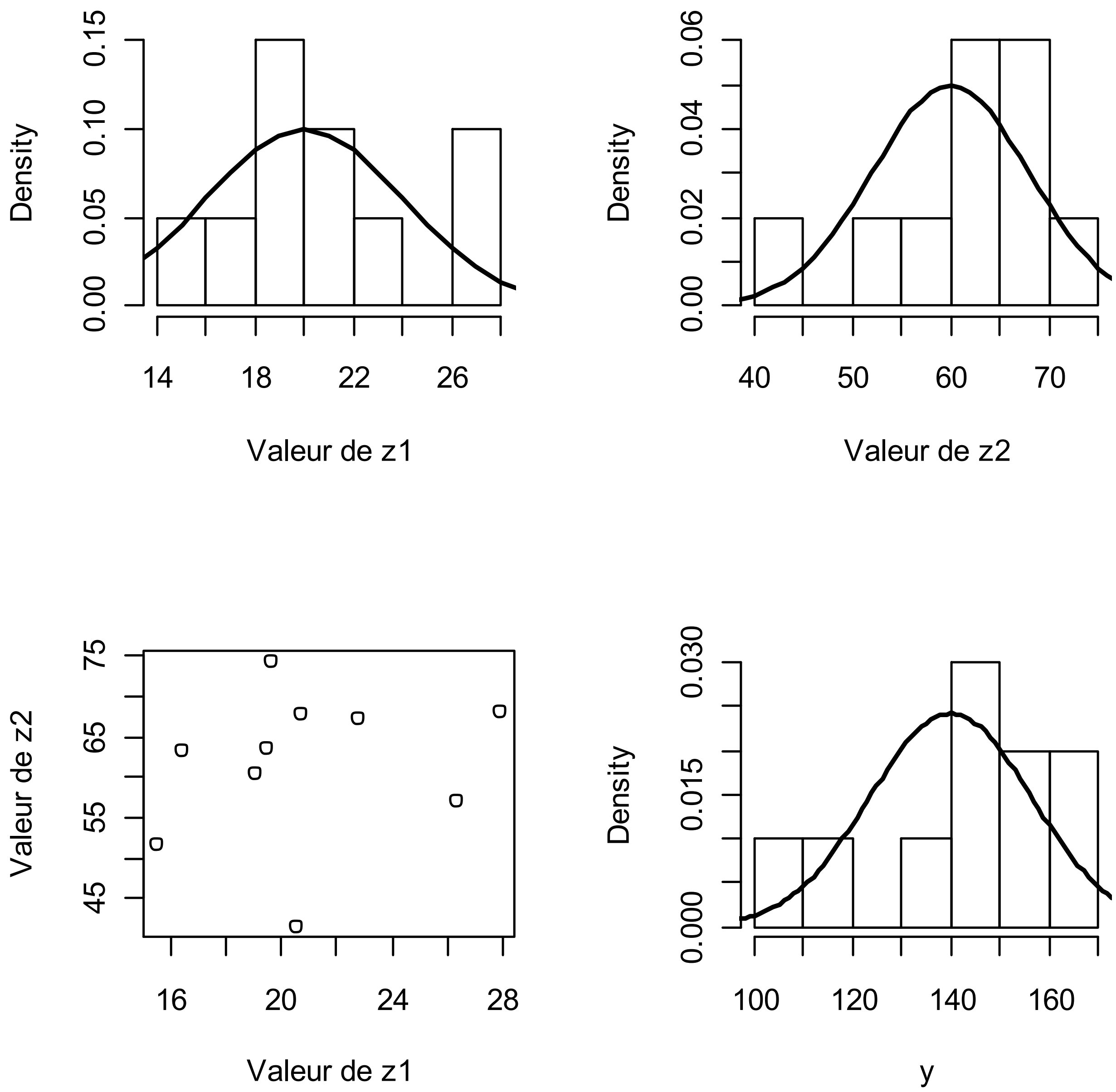
- Plusieurs valeurs de N sont considérées successivement

$$N = 10$$

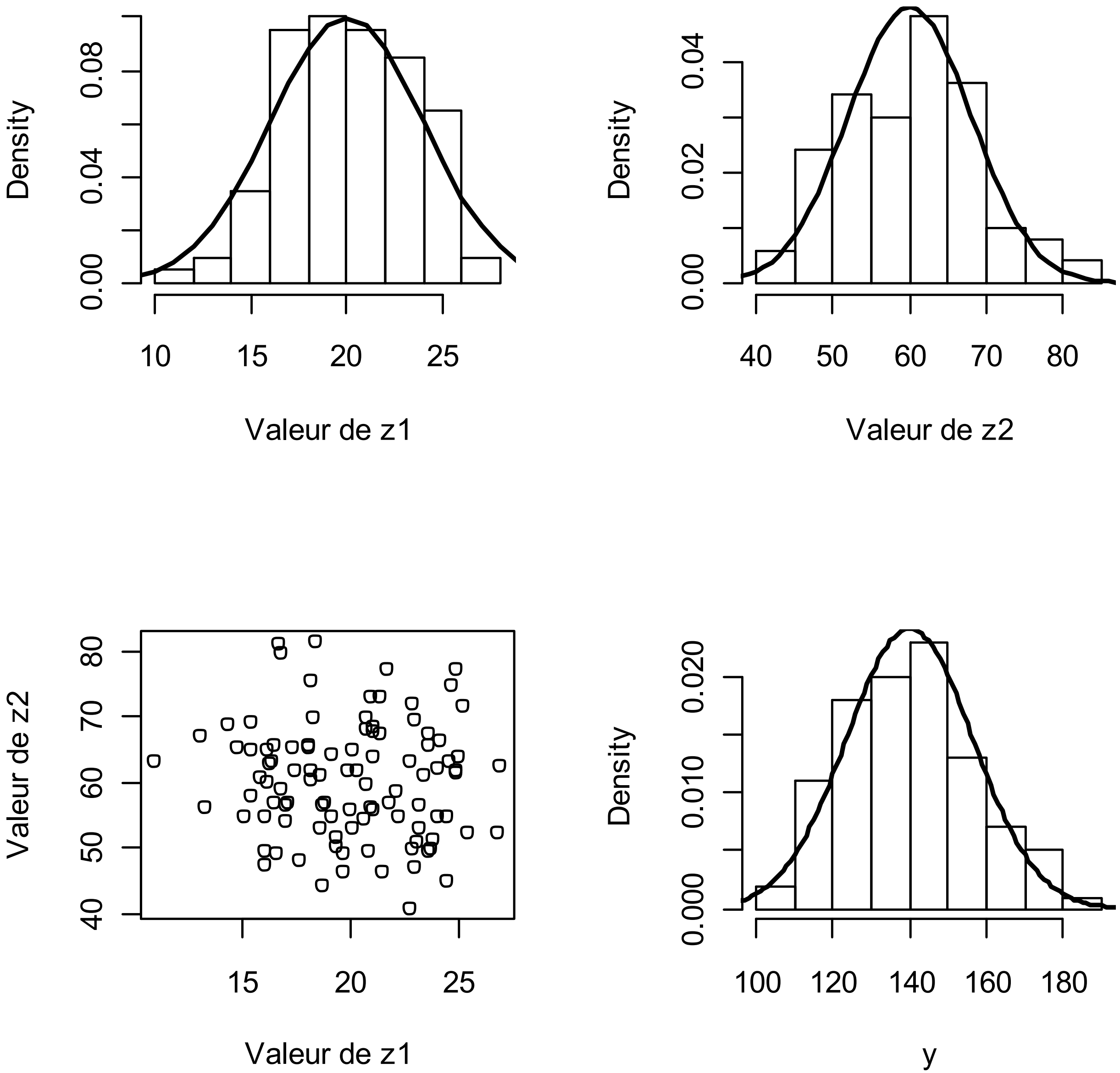
$$N = 100$$

$$N = 1000$$

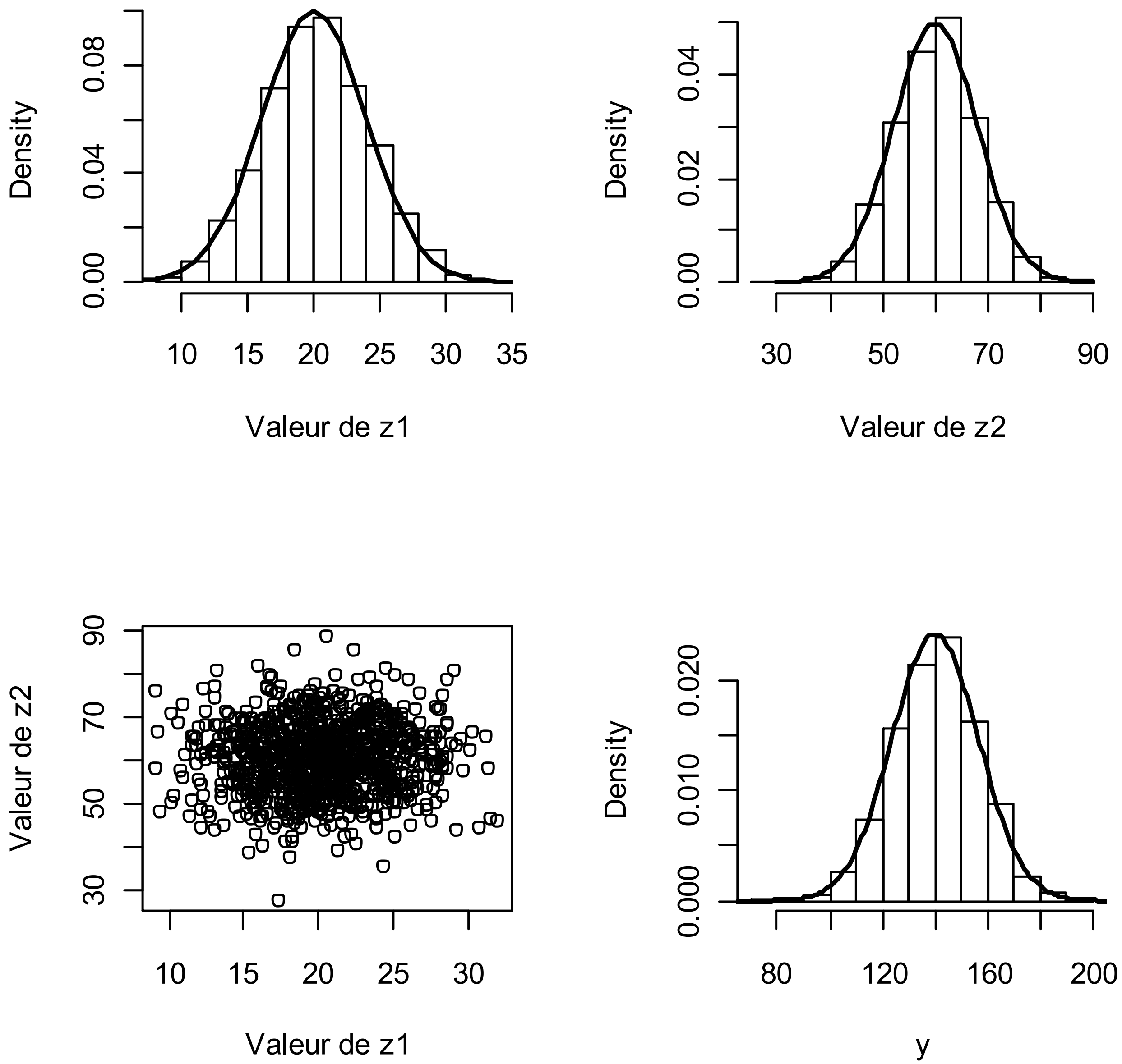
Application. Etape 2. N=10



Application. Etape 2. N=100



Application. Etape 2. N=1000



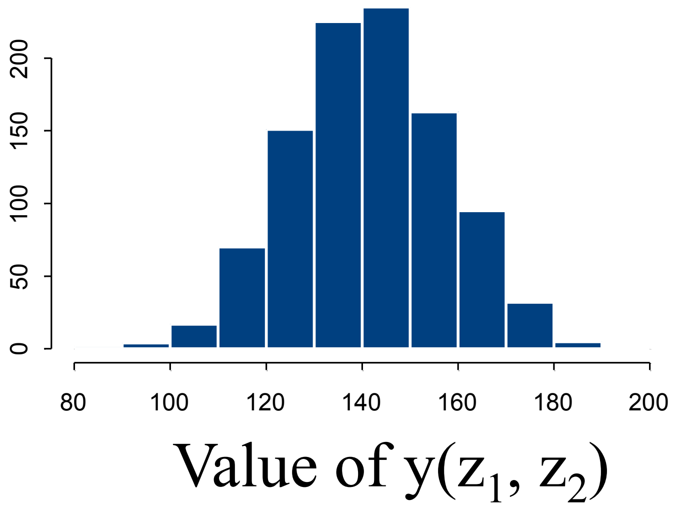
Application. Etape 3

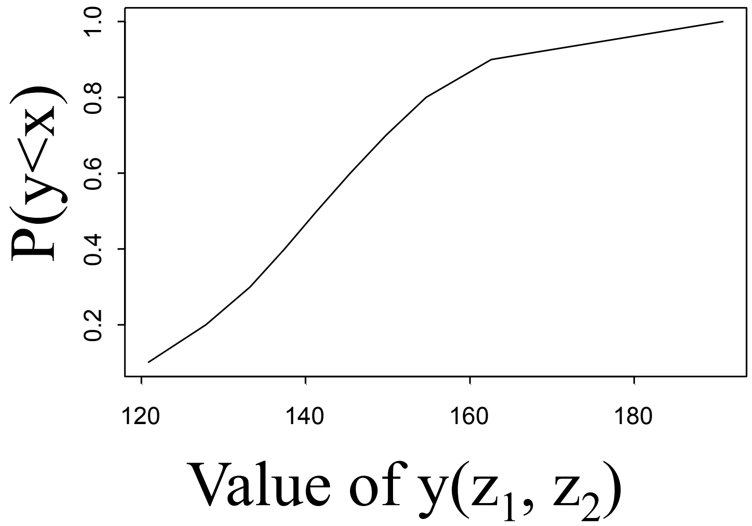
z_1	z_2	$y(z_1,z_2)$
16.83	59.30	
23.18	52.33	
16.43	57.85	
20.45	49.25	
25.48	66.11	
25.67	55.53	
24.67	61.55	
17.88	52.58	
23.69	58.54	
17.69	47.38	

Application. Etape 3

z_1	z_2	$y(z_1,z_2)$
16.83	59.30	135.43
23.18	52.33	127.84
16.43	57.85	132.13
20.45	49.25	118.95
25.48	66.11	157.71
25.67	55.53	136.73
24.67	61.55	147.77
17.88	52.58	123.04
23.69	58.54	140.78
17.69	47.38	112.45

Application. Etape 4. N=1000





Application. Etape 4

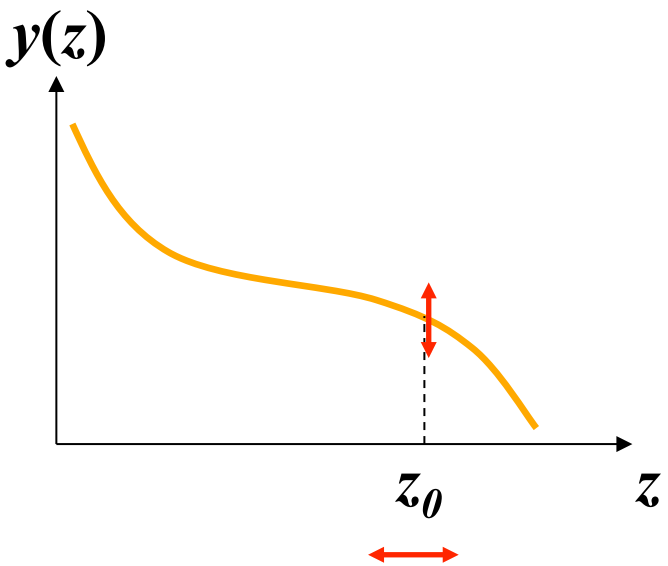
	Mean	Variance	Standard-deviation
N =10	133.28	183.85	13.56
N = 100	138.71	294.96	17.17
N = 1000	141.34	258.23	16.07
N = 5000	139.72	272.51	16.51
N = 7000	139.90	269.45	16.42
True values	140	272	16.49

3. Analyse de sensibilité

Analyse de sensibilité locale
ou
Analyse de sensibilité globale ?

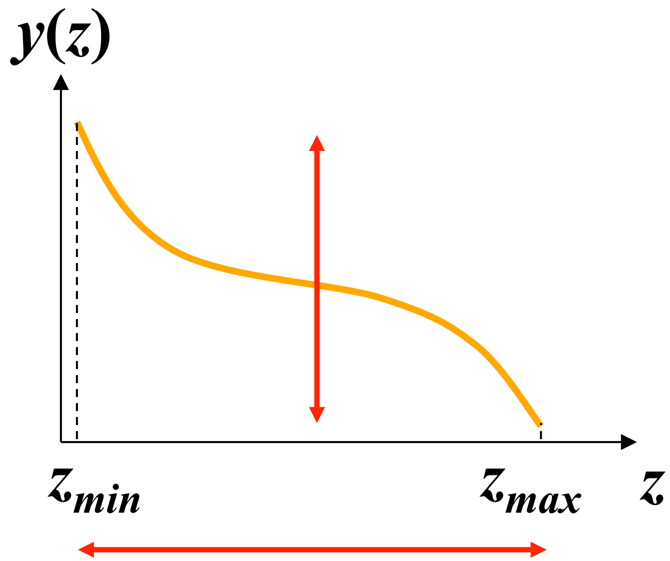
AS locale

Variation de $y(z)$ « autour » z_0



AS globale

Variation globale de $y(z)$ quand z varie dans son domaine d'incertitude



Intérêt pratique de l'analyse de sensibilité

i) Identifier les paramètres et les variables d'entrée qui influencent fortement les sorties du modèle
→ ***Important de les connaître précisément***

ii) Identifier les paramètres et les variables d'entrée qui n'ont pas une forte influence sur les sorties du modèle
→ ***Moins important de les connaître précisément***

iii) Analyser le comportement du modèle

Analyse de sensibilité locale

Basée sur le calcul de dérivé

Analyse de sensibilité globale

Elle consiste à

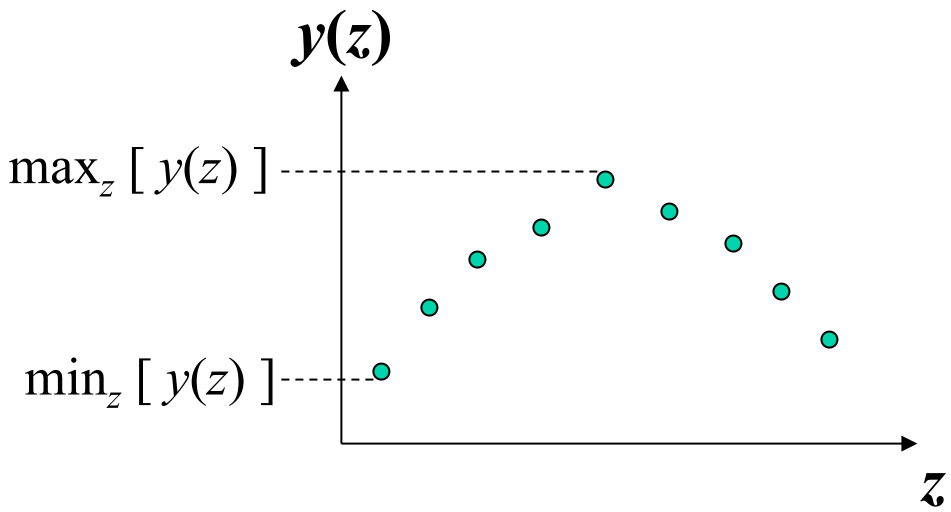
- Définir des indices de sensibilité
- Calculer ces indices en faisant varier les facteurs incertains $z_1, \ldots, z_p$ sur leurs domaines

Un indice de sensibilité simple

Bauer and Hamby (1991)

- On définit une **série de valeurs** pour chaque facteur.
- On fixe tous **les facteurs sauf** z_i à des **valeurs de référence**.
- On calcule pour **le facteur** z_i l'indice:

$$\mathbf{I}_{z_i} = \{ \max_{z_i} [\, y(z) \,] - \min_{z_i} [\, y(z) \,] \} / \max_{z_i} [\, y(z) \,]$$



Application

Equation: $y(z_1, z_2) = z_1 + 2\, z_2$

Définir cinq valeurs pour z_2: 40, 50, 60, 70, 80.

Fixer z_1 à 20.

Quelle est la valeur de l'indice de Bauer-Hamby index pour z_2?

Application

$\max_{z2}$ [$y(z_1{=}20, z_2)$] = 20 + 2*80 = 180

$\min_{z2}$ [$y(z_1{=}20, z_2)$] = 20 + 2*40 = 100

I_{z2} = (180 - 100) / 180 = 0.444

Limite de l'indice de Bauer-Hamby

- Chaque facteur est analysé séparément
- La valeur de l'indice peut dépendre des valeurs de référence

Exemple:

$y(z_1, z_2, z_3) = z_1 + 2*z_2*z_3.$

$I_{z2} = 0$ si $z_3=0.$

$I_{z2} \neq 0$ si $z_3 \neq 0.$

Interactions entre facteurs non prise en compte

AS globale = les <u>*trois premières étapes*</u> de l'AI
+ une <u>*quatrième étape*</u> spécifique

1. **Définition des distributions de $z_1, \ldots, z_p$.**

2. **Génération d'échantillons à partir des distributions définies à l'étape 1.**

3. **Calcul de $y(z)$ pour chaque série $z_1, \ldots, z_p$ générée.**

4. **Calcul d'indices de sensibilité.**

Il existe de nombreuses méthodes pour calculer les indices de sensibilité

ANOVA

Corrélation

Régression

Morris

Sobol

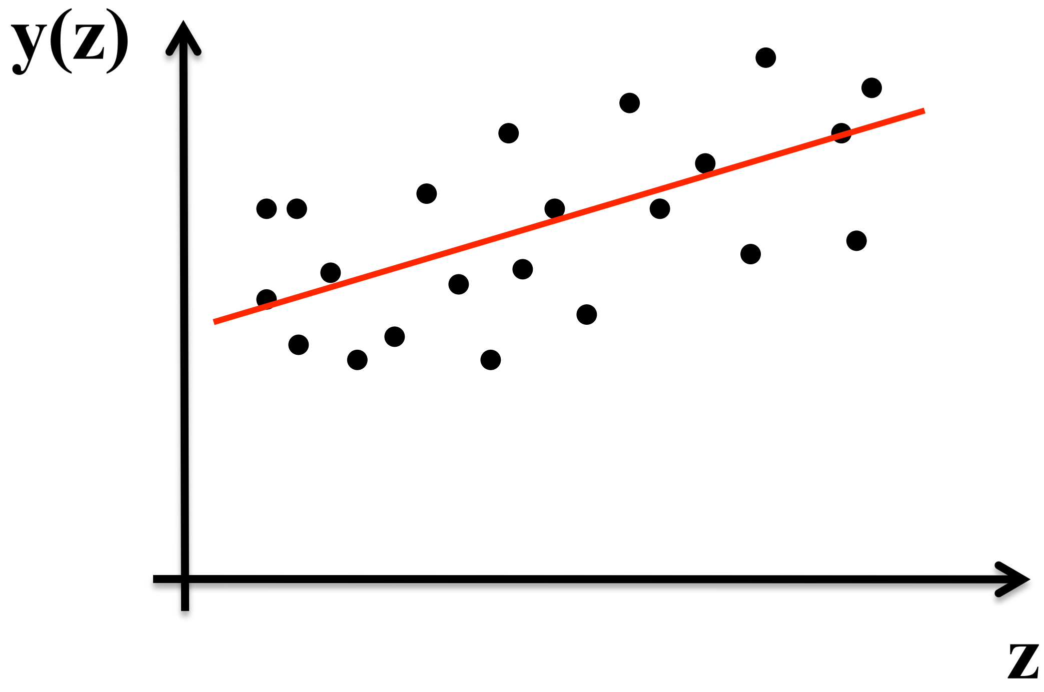
FAST/FAST étendu

etc.

Trois méthodes (relativement) simples

- Regression/Correlation
- ANOVA
- Morris

Regression/Correlation



Regression/Correlation

Différents coefficient de corrélation peuvent être utilisés

- Coefficient de corrélation linéaire

$$\frac{\operatorname{cov}\left(y(z),z\right)}{\sigma_z \sigma_{y(z)}}$$

- Corrélation des rangs (Spearman)

ANOVA

• Définir un plan d'expérience en combinant k valeurs de chacun des facteurs incertains

• Faire une ANOVA de y(z) en fonction de $z_1, \ldots, z_p$

$$y_{i_1, i_2, \ldots i_p} = \mu + \alpha_{i_1} + \beta_{i_2} + \ldots$$

Indices de sensibilité basés sur une décomposition de la variance

$$\underbrace{\mathrm{Var}[\,y(z)\,]}_{\text{Variance totale de la variable de sortie}} = \underbrace{V_{z1} + V_{z2} + V_{z3}}_{\text{Effets principaux des facteurs incertains}} + \ldots + \underbrace{V_{z1.z2} + V_{z1.z3}}_{\text{Termes d'interactions}} + \ldots$$

Indice de premier ordre de $z_1 = V_{z1}$ / Var[y(z)]

Indice de sensibilité total de $z_1 = (V_{z1} + V_{z1.z2} + V_{z1.z3} + \ldots)$ / Var[y(z)]

$$y(\mathbf{z}) = f(z_1, \ldots, z_s)$$

Indice de sensibilité de 1er ordre pour z_i

$$\frac{\mathrm{var}\left[\mathrm{E}[y(z) \mid z_i]\right]}{\mathrm{var}[y(z)]}$$

Indice de sensibilité total pour z_i

$$\frac{\mathrm{E}\left[\mathrm{var}[y(z) \mid z_j, j \neq i]\right]}{\mathrm{var}[y(z)]}$$

Signification de l'indice de sensibilité totale

- **Indice de sensibilité total de z_i (IT_i) = Fraction de la variance totale de y si seulement z_i est inconnu.**

- **IT_i est compris entre 0 et 1.**

$\boxed{IT_i \textbf{ proche de } 0}$ → Pas nécessaire d'estimer précisément z_i

$\boxed{IT_i \textbf{ proche de } 1}$ → Probablement important d'estimer précisément z_i

Méthode de Morris

- Define a design by combining k values of the p uncertain factors
- Add a « jump » Δ_{ij} to the i^{th} uncertain factors
- Compute an « elementary effect »

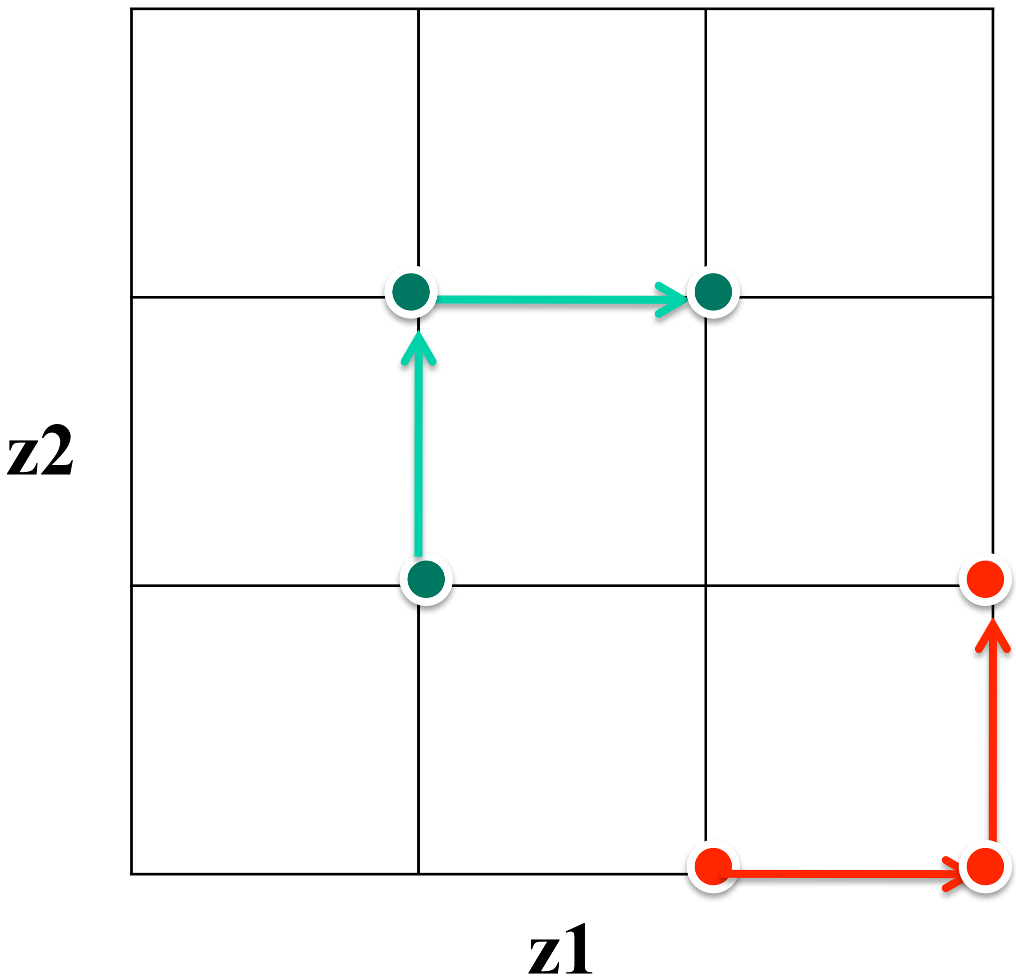
$$d_{ij} = \frac{\left[y\left(z_1,\ldots,z_{i-1},z_i+\Delta_{ij},\ldots,z_p\right) - y(z) \right]}{\Delta_{ij}}$$

- Repeat the procedure for all uncertain factors (i=1, …, p)
- Replicate r times (j=1, …, r)
- Compute mean and variance of elementary effects from r replicates

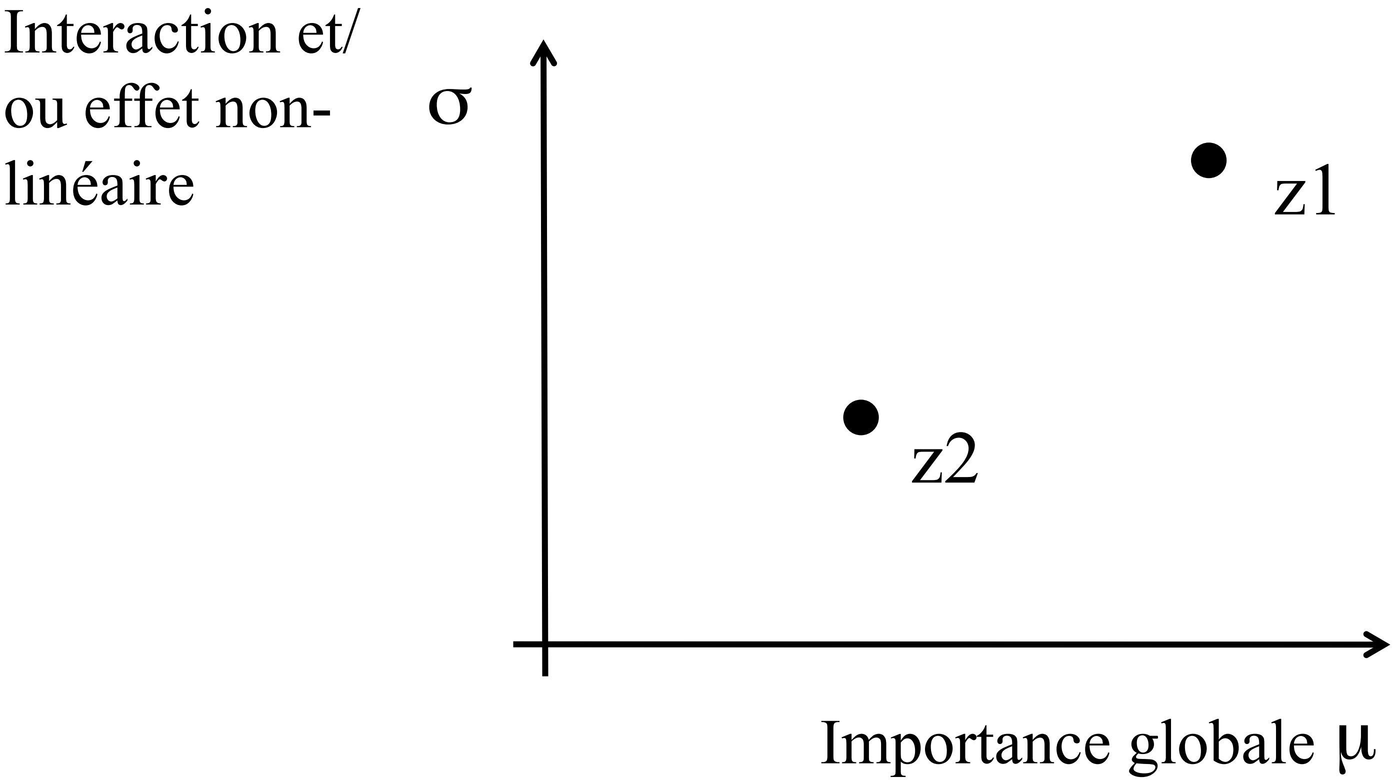
$$\mu_i = \frac{\sum_{j=1}^{r} d_{ij}}{r} \qquad \sigma_i = \sqrt{\sum_{j=1}^{r} \left(d_{ij} - \mu_i\right)^2 / r}$$

Morris

Deux exemples de trajectoires (*p*=2, *k*=3, *r*=2)



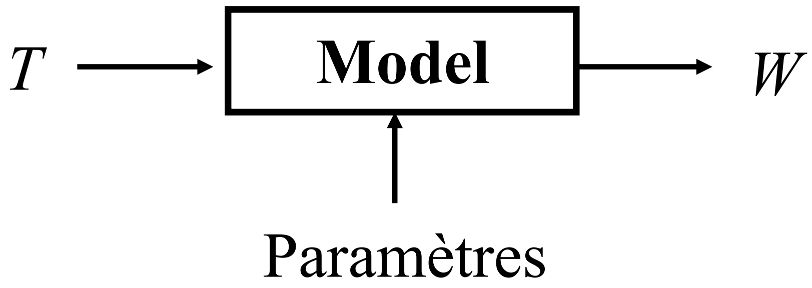
Morris
Présentation des résultats



Etude de cas

Un modèle générique pour calculer la durée (en heures) requise d'humidité pour qu'un champignon puisse infecter une plante

(Magarey *et al*., 2005)



W = durée d'humidité requise (h)
T = température moyenne (°C)

Un modèle générique pour calculer la durée (h) requise d'humidité pour qu'un champignon puisse infecter une plante

(Magarey *et al.*, 2005)

$W = W_{\min} / f(T)$, mais inférieure à $W_{\max}$

$$f(T) = \left(\frac{T_{\max} - T}{T_{\max} - T_{opt}}\right)\left(\frac{T - T_{\min}}{T_{opt} - T_{\min}}\right)^{(T_{opt} - T_{\min})/(T_{\max} - T_{opt})}$$

Cinq paramètres : T_{min}, T_{opt}, T_{max}, W_{min}, W_{max}

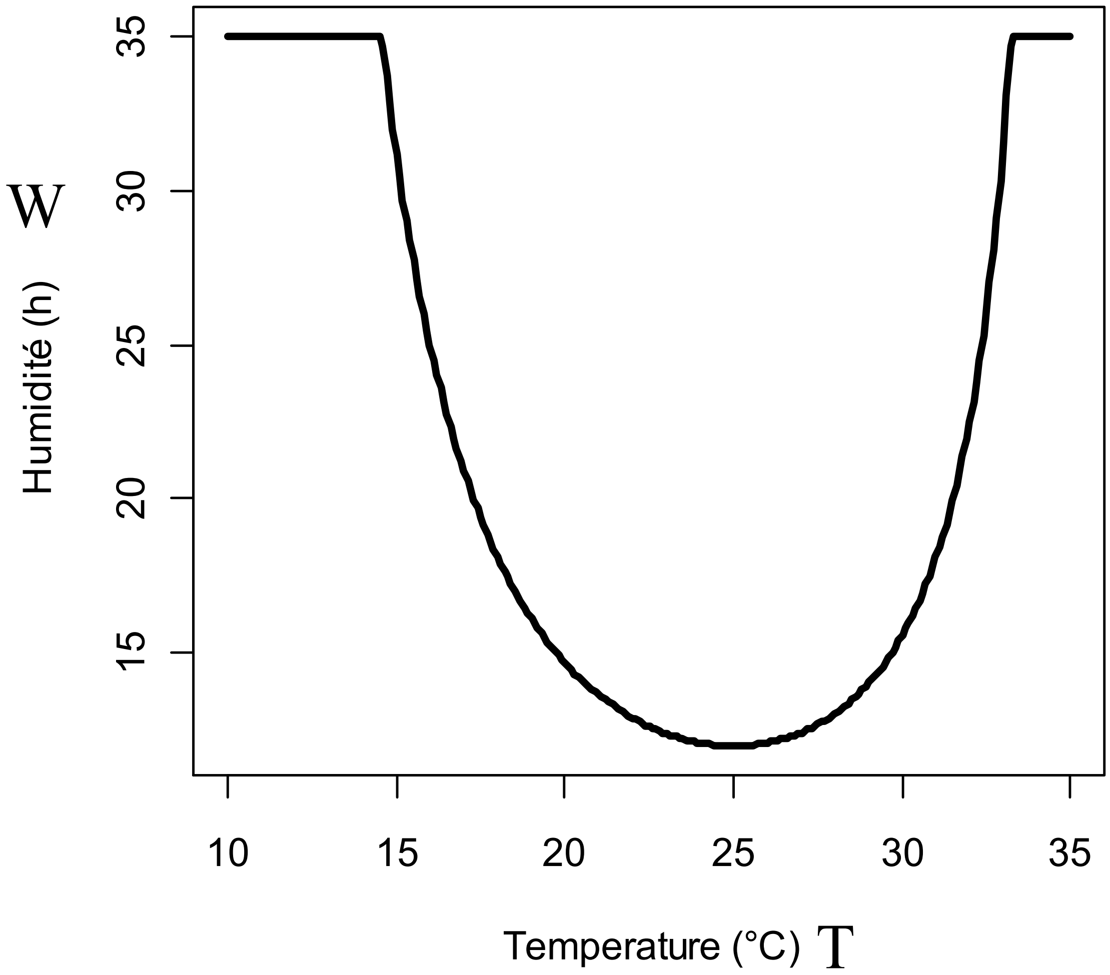
- **Les paramètres peuvent être estimés à partir de données et d'articles scientifiques pour différents champignons pathogènes**

- **Il reste des incertitudes sur ces paramètres**

- **Important**
  - **d'analyser l'incertitude induite par les paramètres sur W**
  - **d'identifier les paramètres les plus influents afin de réaliser des expérimentations spécifiques**

Exemple de valeurs estimées de paramètres pour les pycnidiospores de *Guignardia citricarpa* Kiely et valeurs simulées de W.

Tmin= 10 °C, Topt= 25 °C, Tmax=35 °C, Wmin=12 h, Wmax= 35 h



Incertitude sur les valeurs des paramètres
(pycnidiospores de *Guignardia citricarpa* Kiely)

		Min	Max
Tmin	(°C):	10	15
Tmax	(°C):	32	35
Topt	(°C):	25	30
Wmin	(h):	12	14
Wmax	(h):	35	48

Panel on Plant Health, EFSA (2008)

Questions

1. **Réaliser une analyse d’incertitude pour W**
2. **Réaliser une analyse de sensibilité sur W**

1. Analyse d’incertitude pour W

i. Définir les distributions des paramètres

ii. Générer *N* séries de valeurs de paramètres (*N*=500)

iii. Calculer *W* pour chaque série

iv. Décrire la distribution de *W*

Une fonction R pour calculer W

```
Wetness <- function(T, Tmin, Topt, Tmax, Wmin, Wmax) {

fT <- ((Tmax-T)/(Tmax-Topt))*(((T-Tmin)/(Topt-Tmin))
                              ^((Topt-Tmin)/(Tmax-Topt)))

W <- Wmin/fT
W[W>Wmax] <- Wmax
return(W)

}
```

T, Tmin, Topt, Tmax, Wmin, Wmax

Wetness

W

Génération des valeurs des paramètres

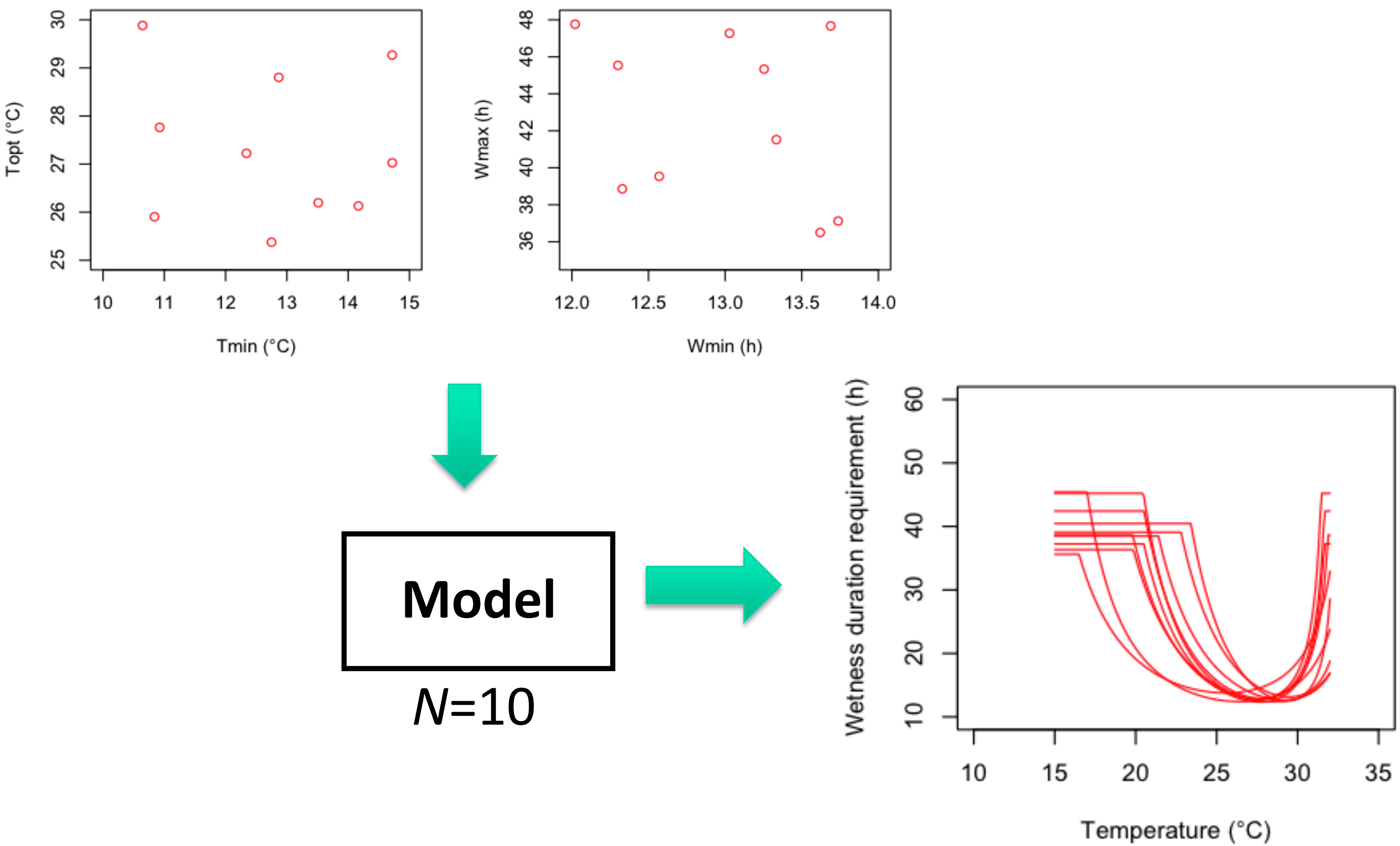
```
Num <- 500

Tmin_vec <- runif(Num, 10, 15)

Topt_vec <- runif(Num, 25, 30)

Tmax_vec <- runif(Num, 32, 35)

Wmin_vec <- runif(Num, 12, 14)

Wmax_vec <- runif(Num, 35, 48)
```

Simulation de W

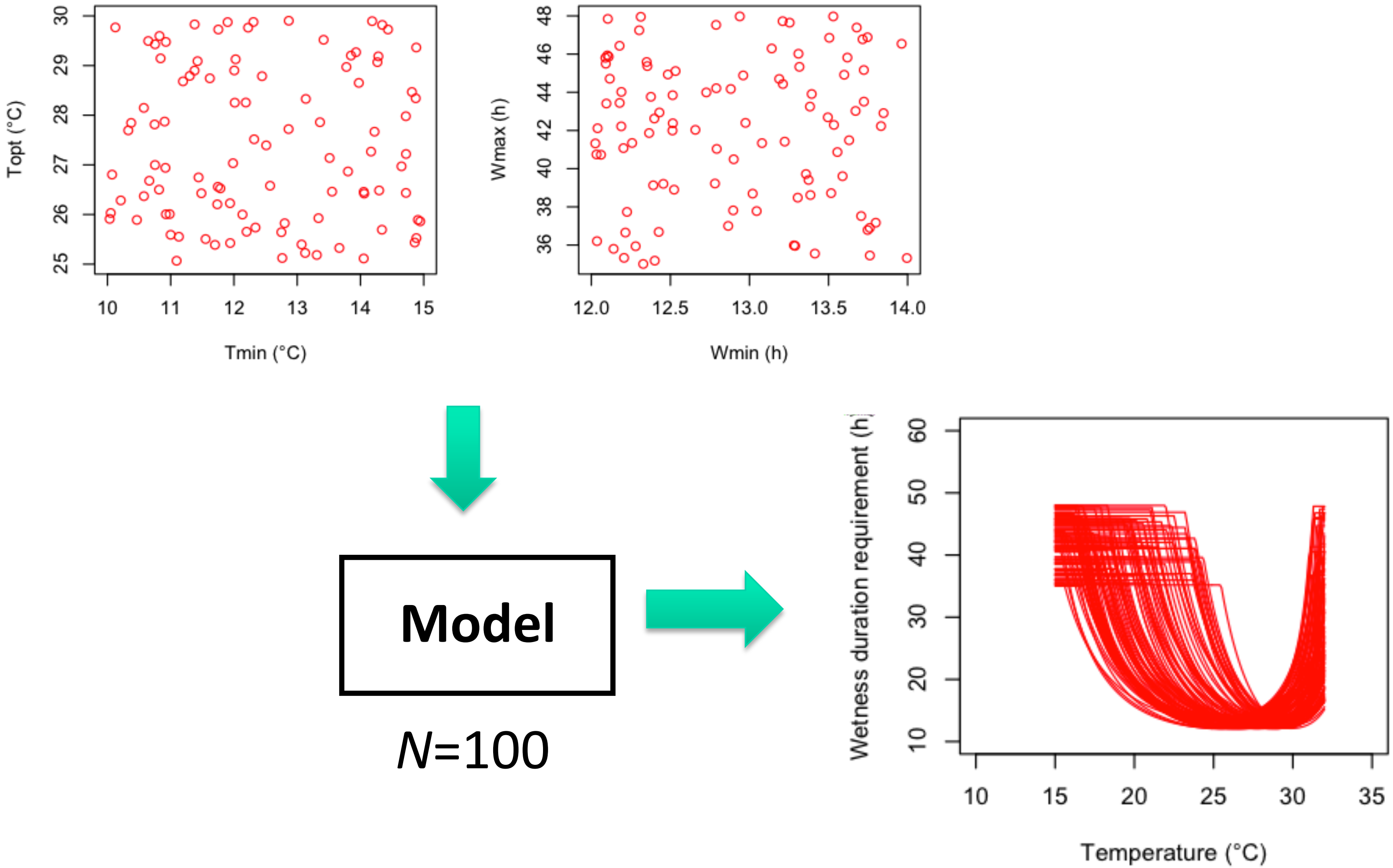
```
T_vec <- seq(from=15, to=32, by=0.1)

W_mat <- matrix(nrow=Num, ncol=length(T_vec))

for (i in 1:Num) {

W_mat[i,] <- Wetness(T_vec, Tmin_vec[i], Topt_vec[i],
                  Tmax_vec[i], Wmin_vec[i], Wmax_vec[i])

lines(T_vec, W_mat[i,])

}
```

Analyse des sorties

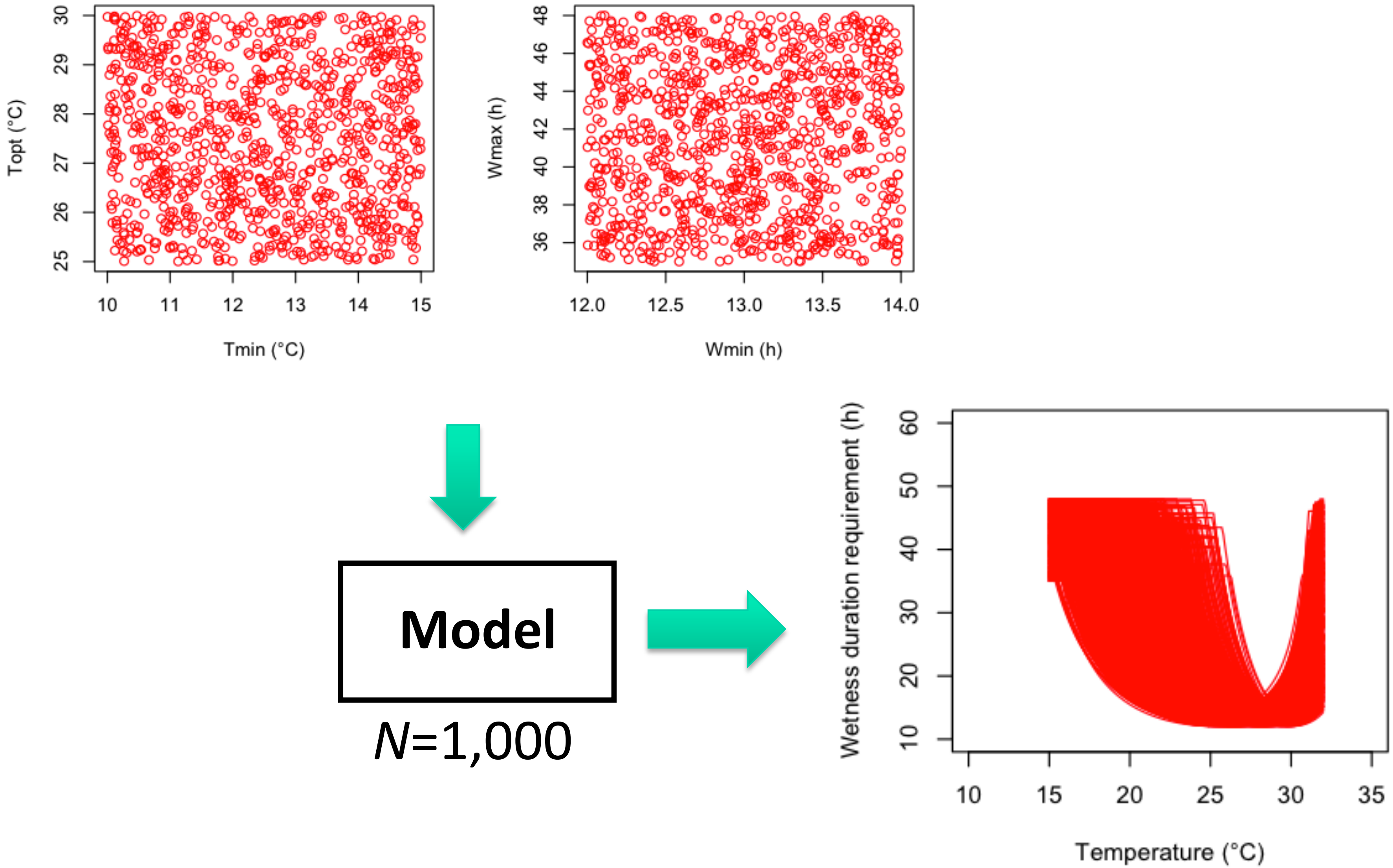
```
mean_vec <- apply(W_mat, 2, mean)
Q0.01_vec <- apply(W_mat, 2, quantile, 0.01)
Q0.1_vec <- apply(W_mat, 2, quantile, 0.1)
Q0.9_vec <- apply(W_mat, 2, quantile, 0.9)
Q0.99_vec <- apply(W_mat, 2, quantile, 0.99)

plot(c(0), c(0), pch=" ", xlab="Temperature (°C)",
ylab="Wetness duration requirement (h)", xlim=c(10, 35),
ylim=c(10, 60)
lines(T_vec, mean_vec, lwd=3)
lines(T_vec, Q0.9_vec, lty=2)
lines(T_vec, Q0.1_vec, lty=2)
lines(T_vec, Q0.99_vec, lty=9)
lines(T_vec, Q0.01_vec, lty=9)
```

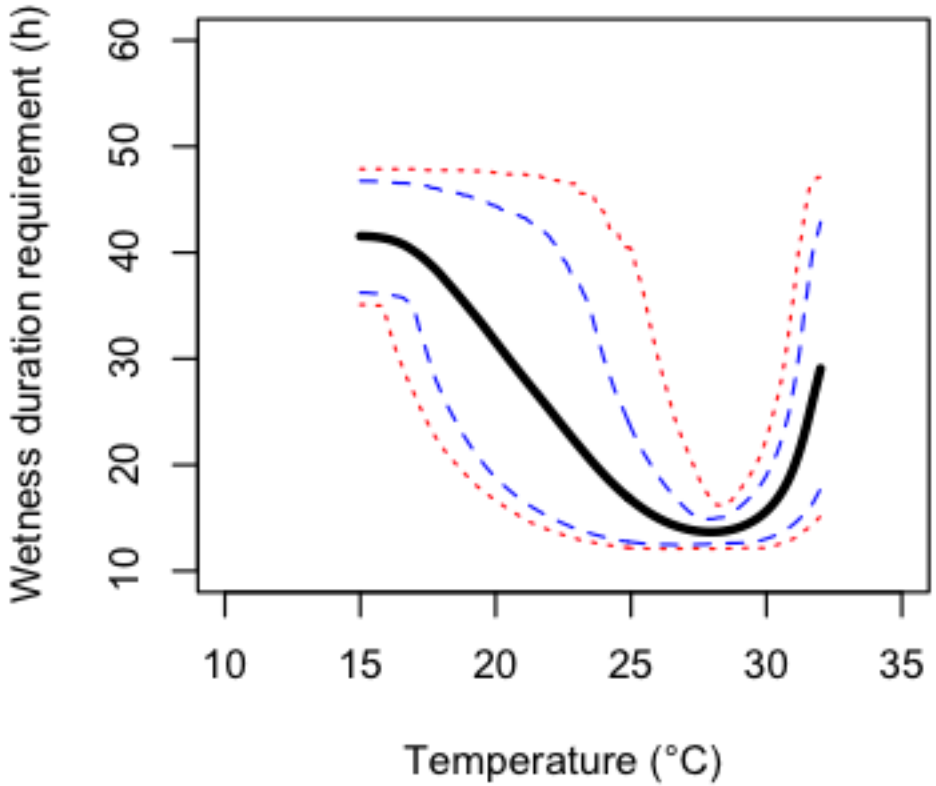


Topt (°C)
30
29
28
27
26
25
10
11
12
13
14
15
Tmin (°C)
Wmax (h)
48
46
44
42
40
38
36
12.0
12.5
13.0
13.5
14.0
Wmin (h)
Model
N=10
Wetness duration requirement (h)
60
50
40
30
20
10
10
15
20
25
30
35
Temperature (°C)



Topt (°C)
Tmin (°C)
Wmax (h)
Wmin (h)
Model
N=100
Wetness duration requirement (h)
Temperature (°C)


Topt (°C)
Tmin (°C)
Wmax (h)
Wmin (h)
Model
N=1,000
Wetness duration requirement (h)
Temperature (°C)



Q99

Q90

Mean

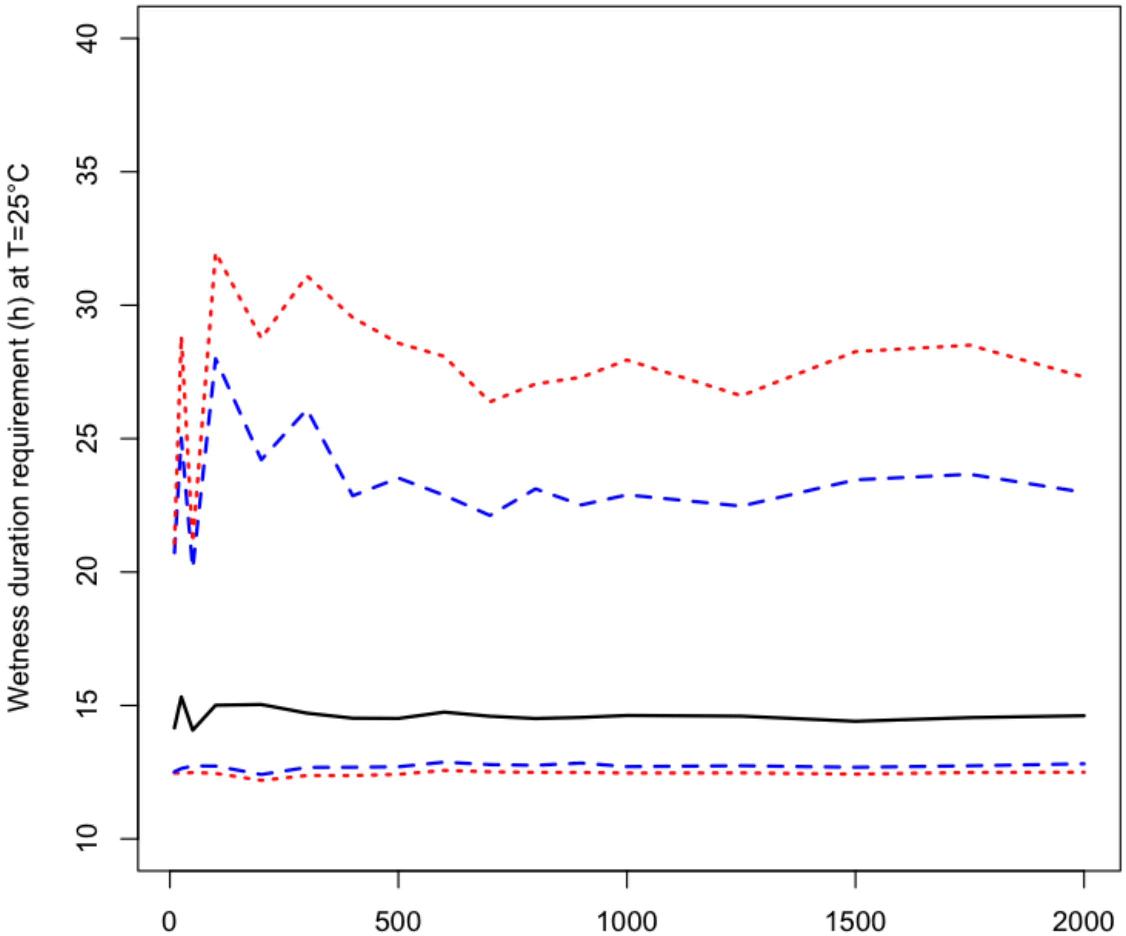
Q10

Q01

N=1,000

Estimated extreme quantiles may be unstable!

Wetness duration requirement (h) at T=25°C



Q99

Q90

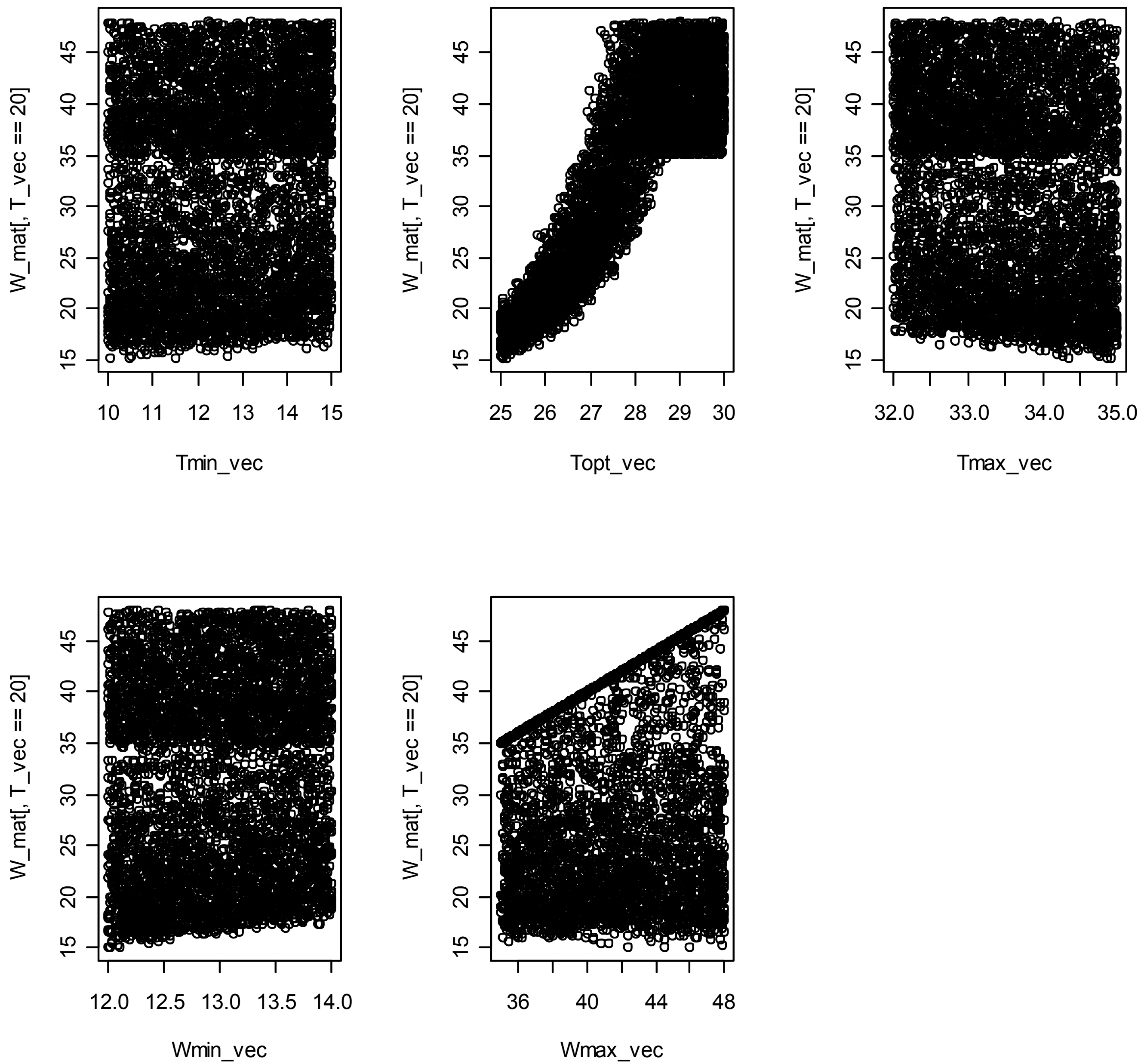
Q50

Q10

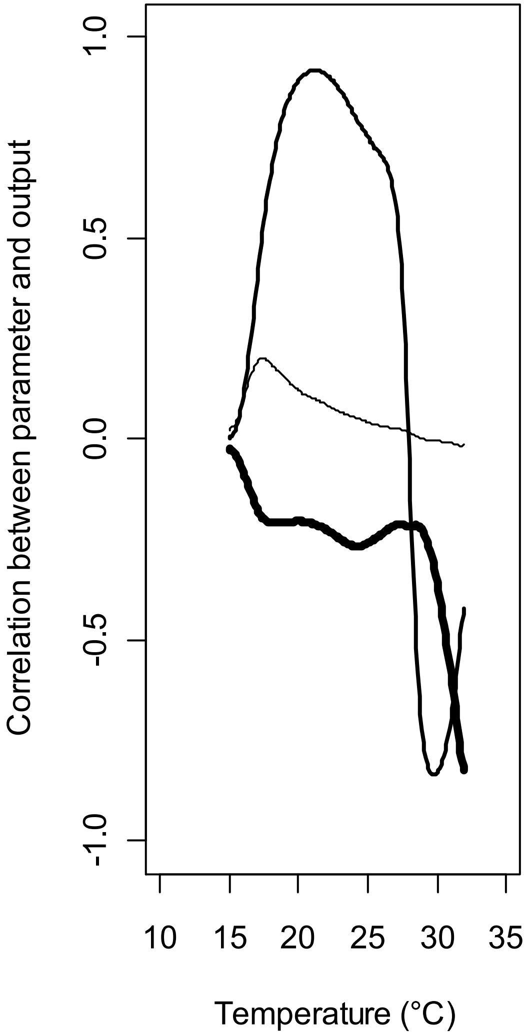
Q01

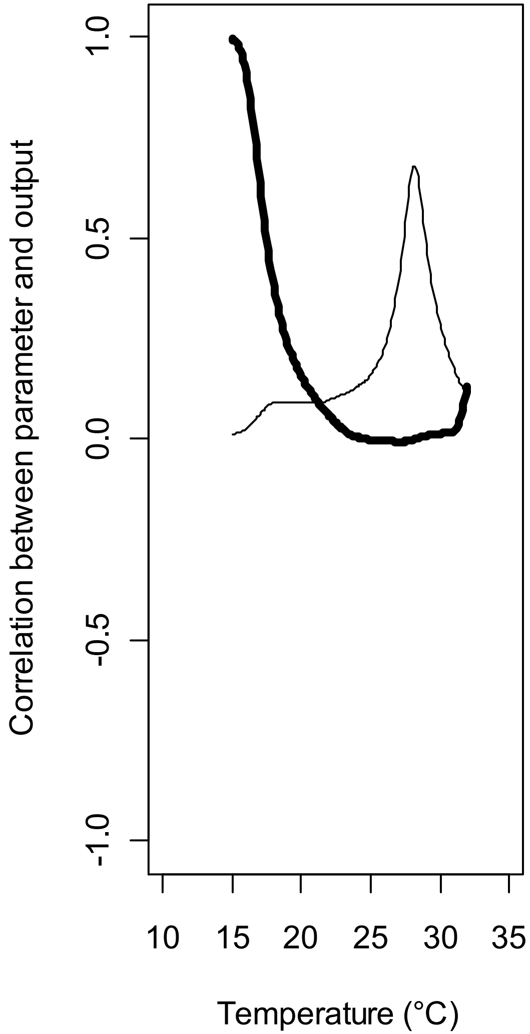
Nb. of simulations

Plots of model outputs obtained by Monte Carlo simulations



Correlation-based sensitivity analysis





2. Analyse de sensibilité pour *W* par ANOVA

i. Définir un plan d'expérience (plan fact. complet avec trois valeurs par paramètre)

ii. Générer toutes les combinaisons possibles

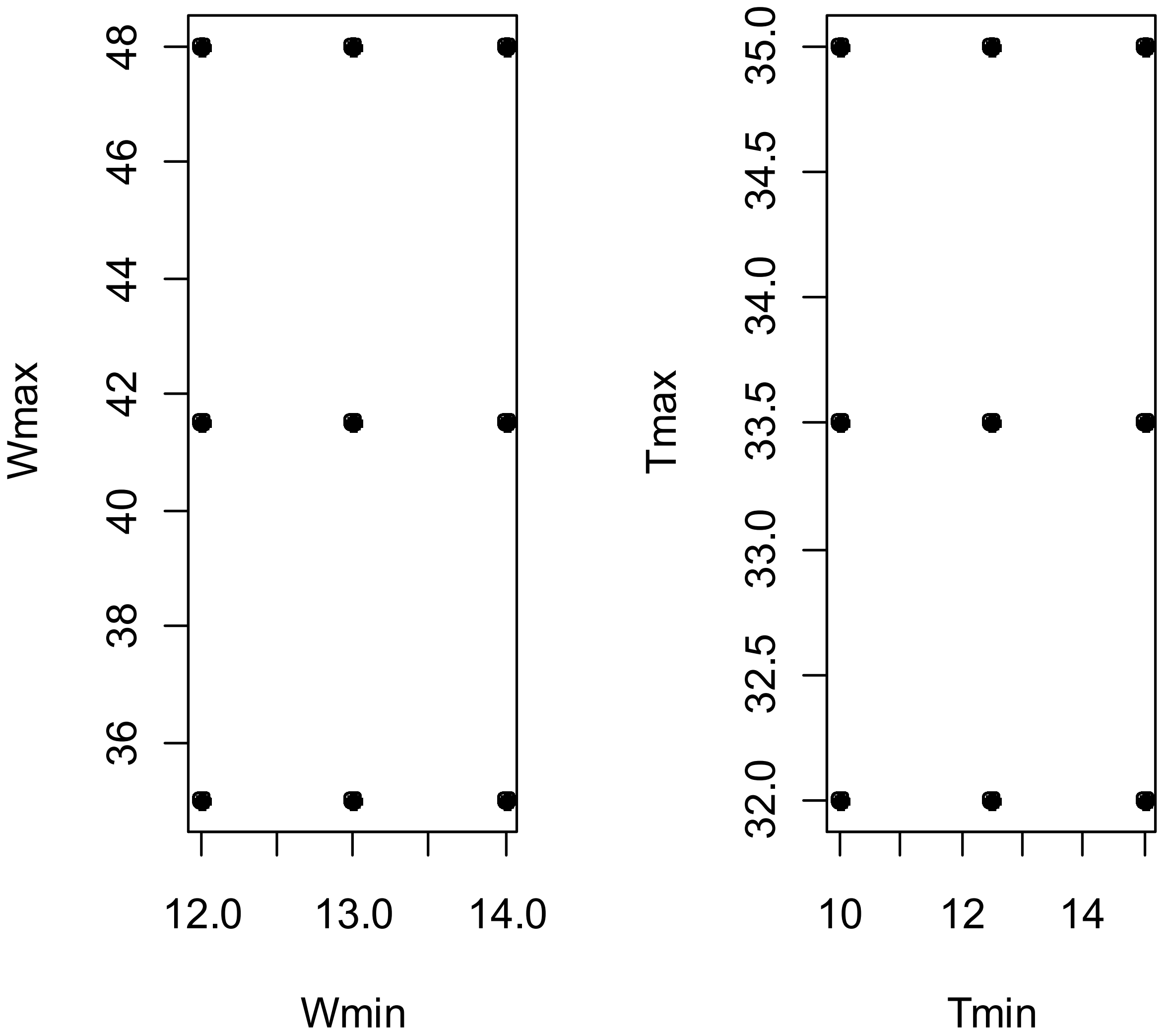
iii. Calculer *W* pour chaque combinaison

iv. Réaliser une ANOVA et calculer les indices de sensibilité

Plan d'expérience

```
# Tableau incluant 243 valeurs de paramètres

para.mat <- expand.grid(Tmin=c(10, 12.5, 15), Topt=c(25, 27.5,
30),Tmax=c(32, 33.5, 35), Wmin=c(12, 13, 14), Wmax=c(35, 41.5, 48))
print(para.mat)

plot(para.mat$Wmin, para.mat$Wmax, pch=19)
plot(para.mat$Tmin, para.mat$Tmax, pch=19)
```

	Tmin	Topt	Tmax	Wmin	Wmax
1	10.0	25.0	32.0	12	35.0
2	12.5	25.0	32.0	12	35.0
3	15.0	25.0	32.0	12	35.0
4	10.0	27.5	32.0	12	35.0
5	12.5	27.5	32.0	12	35.0
6	15.0	27.5	32.0	12	35.0
7	10.0	30.0	32.0	12	35.0
8	12.5	30.0	32.0	12	35.0
9	15.0	30.0	32.0	12	35.0
10	10.0	25.0	33.5	12	35.0
11	12.5	25.0	33.5	12	35.0
12	15.0	25.0	33.5	12	35.0
…..					
243					


Wmax
48
46
44
42
40
38
36
12.0
13.0
14.0
Wmin
Tmax
35.0
34.5
34.0
33.5
33.0
32.5
32.0
10
12
14
Tmin

Calcule de W pour chaque combinaison

```
# Temperature values

T.vec <- c(20, 25, 30)

# Create an empty matrix to store the simulated values

W.Mat <- matrix(nrow=243, ncol=3)

# Loop for simulating W

for (i in 1:243) {

W.mat[i,] <- Wetness(T.vec, para.mat$Tmin[i], para.mat$Topt[i],
para.mat$Tmax[i], para.mat$Wmin[i], para.mat$Wmax[i])

}
```

Indices de sensibilité

```
#Define the sets of parameter values as factors

Tmin <- as.factor(para.mat$Tmin)
Topt <- as.factor(para.mat$Topt)
Tmax <- as.factor(para.mat$Tmax)
Wmin <- as.factor(para.mat$Wmin)
Wmax <- as.factor(para.mat$Wmax)

#Select the simulations obtained for T=30
W <- W.mat[,3]

#Create a table
TAB <- data.frame(W, Tmin, Topt, Tmax, Wmin, Wmax)
```

```
#ANOVA (sum of squared associated with main effects and interactions)

Fit <- summary(aov(W~Tmin*Topt*Tmax*Wmin*Wmax, data=TAB))
print(Fit)

#Computation of sensitivity indices
SumSq <- Fit[[1]][,2]
Total <- 242*var(W)
Indices <- 100*SumSq/Total
print(Indices)
TabIndices <- cbind(Fit[[1]],Indices)
print(TabIndices)
TabIndices <- TabIndices[order(Indices, decreasing=T),]
print(TabIndices)
```

```
> print(TabIndices)
```

	Sum Sq	Mean Sq	Indices
Topt	2.315226e+03	1.157613e+03	6.362759e+01
Tmax	5.907681e+02	2.953841e+02	1.623563e+01
Topt:Tmax	4.555308e+02	1.138827e+02	1.251901e+01
Wmin	2.570847e+02	1.285423e+02	7.065261e+00
Topt:Wmin	9.133042e+00	2.283260e+00	2.509964e-01
Tmin:Topt	3.191415e+00	7.978539e-01	8.770723e-02
Tmin	3.029813e+00	1.514906e+00	8.326603e-02
Tmax:Wmin	2.330446e+00	5.826115e-01	6.404587e-02

Conclusions

• L'analyse de sensibilité et l'analyse d'incertitude ont différents objectifs

• Plusieurs méthodes existent : Appliquez les et comparez les résultats !

• Important d'être transparent sur les hypothèses (ex : distributions de probabilité)

• Le temps de calcul peut être un facteur limitant

• Logiciel:
  - Tableurs (@risk, crystalball)
  - R (package sensitivity et bientôt mtk)
  - C, Fortran routines…

Some references

EFSA. 2008. Scientific Opinion of the Panel on Plant Heath on a request from the European Commission on Guignardia citricarpa Kiely. *The EFSA Journal* 925, 1-108

Lacroix, A., N. Beaudoin, D. Makowski. 2005. Agricultural water nonpoint pollution control under uncertainty and climate variability. *Ecological Economics* 53:115-127

Lamboni M., Makowski D., Lehuger S., Gabrielle B., Monod H. 2009. Multivaiate global sensitivity analysis for dynamic crop models. *Field Crop Research* 113, 312-320

Magarey RD, Sutton TB, Thayer CL. 2005. A simple generic infection model for foliar fungal plant pathogens. *Phytopathology* 95, 92-100.

Makowski, D., C. Naud, M-H. Jeuffroy, A. Barbottin, H. Monod. 2006. Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model predictions. *Reliability Engineering and System Safety* 91:1142-1147.

Monod, H., C. Naud, D. Makowski. 2006. Uncertainty and sensitivity analysis for crop models. *In: Working with dynamic crop models.* D. Wallach, D. Makowski, J. Jones Eds, Elsevier. p. 55-100.

Saltelli, A., S. Tarantola, F. Campolongo, M. Ratto. 2004. « *Sensitivity analysis in practice, a guide to assessing scientific models* ». Wiley.