

***Comparaison d'approches fréquentiste
et bayésienne pour la modélisation du
risque septoriose du blé***

Vendredi 11 Décembre 2015 - Paris

Contexte

Méthodologie statistique

Finalité de notre cas d'étude

Comparaison des 2 approches

Contexte

Bulletin de santé du végétal

Céréales à paille du 19/05/2015

N° 21



Blé tendre

STADE

Contexte d'observations

67 parcelles du réseau ont fait l'objet d'une observation entre le 14 et le 19 mai (semaine 21). Les stades sont centrés sur fin épiaison (27%). 27% sont entre gonflement et mi épiaison, tandis que 46% des parcelles sont entre début et fin floraison.

ROUILLE JAUNE

Rappel des conditions favorables et des résistances variétales : cf. BSV Centre n°10s10 du 03/03/15.

Contexte d'observations

La rouille jaune progresse légèrement. Sur 55 observations, 9 parcelles dont 5 protégées récemment ont des symptômes : sur 10 à 40% des F3, 10 à 20% des F2, voire 10% des F1 du moment. Le seuil de nuisibilité est franchi pour toutes ces situations. En se référant aux notes de résistance actualisées en 2014, les variétés concernées sont sensibles à moyennement sensibles (Adhoc, Boregar) voire résistantes (Aligator, Diamento, Cellule).

Attention : les races de rouille jaune peuvent évoluer rapidement d'une année à l'autre. Il est donc important de surveiller régulièrement toutes les variétés, même celles présentant un haut niveau de résistance à cette maladie.

Seuil de nuisibilité

A partir du stade 1 nœud, le seuil de nuisibilité est atteint dès l'apparition des premières pustules dans la parcelle.

Prévision

Avec les alternances de temps ensoleillé et couvert et la baisse des températures, le risque rouille jaune augmente pour les jours à venir. Surveiller particulièrement les variétés les plus sensibles et les parcelles présentant déjà des symptômes.

OÏDIUM

Contexte d'observations

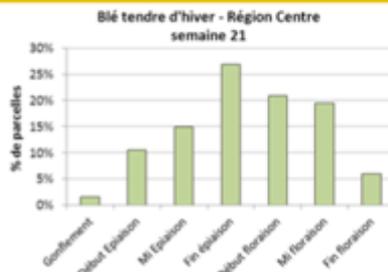
Sur 49 observations, 3 parcelles dont 2 protégées récemment présentent des symptômes sur 10 à 60% des F3, 40% des F2 voire 10% des F1 du moment. Les variétés concernées sont sensibles (Pakito) à peu sensible (Alixan, Rubisko). Une situation en Sologne est à nouveau au seuil de nuisibilité (variété peu sensible avec plus de 50% des F3, F2 ou F1 touchées). Des symptômes sur tiges et gaines sont notés à Amilly (45).

Seuil de nuisibilité

A partir du stade épi 1 cm, en fonction des sensibilités variétales, le seuil de nuisibilité est atteint lorsque :

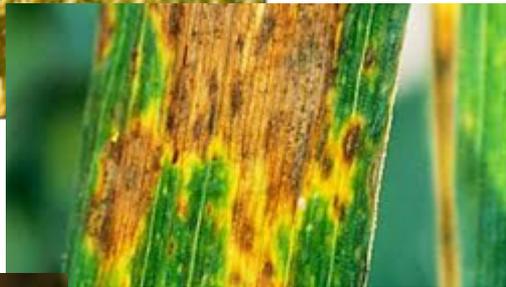
- pour les variétés sensibles : plus de 20% des 3^{èmes} ou 2^{èmes} ou 1^{ères} feuilles sont atteints,
- pour les autres variétés : plus de 50% des 3^{èmes} ou 2^{èmes} ou 1^{ères} feuilles sont atteints.

Echelle des sensibilités variétales à l'oïdium : cf. BSV n°13s13 édité le 24/03/2015.

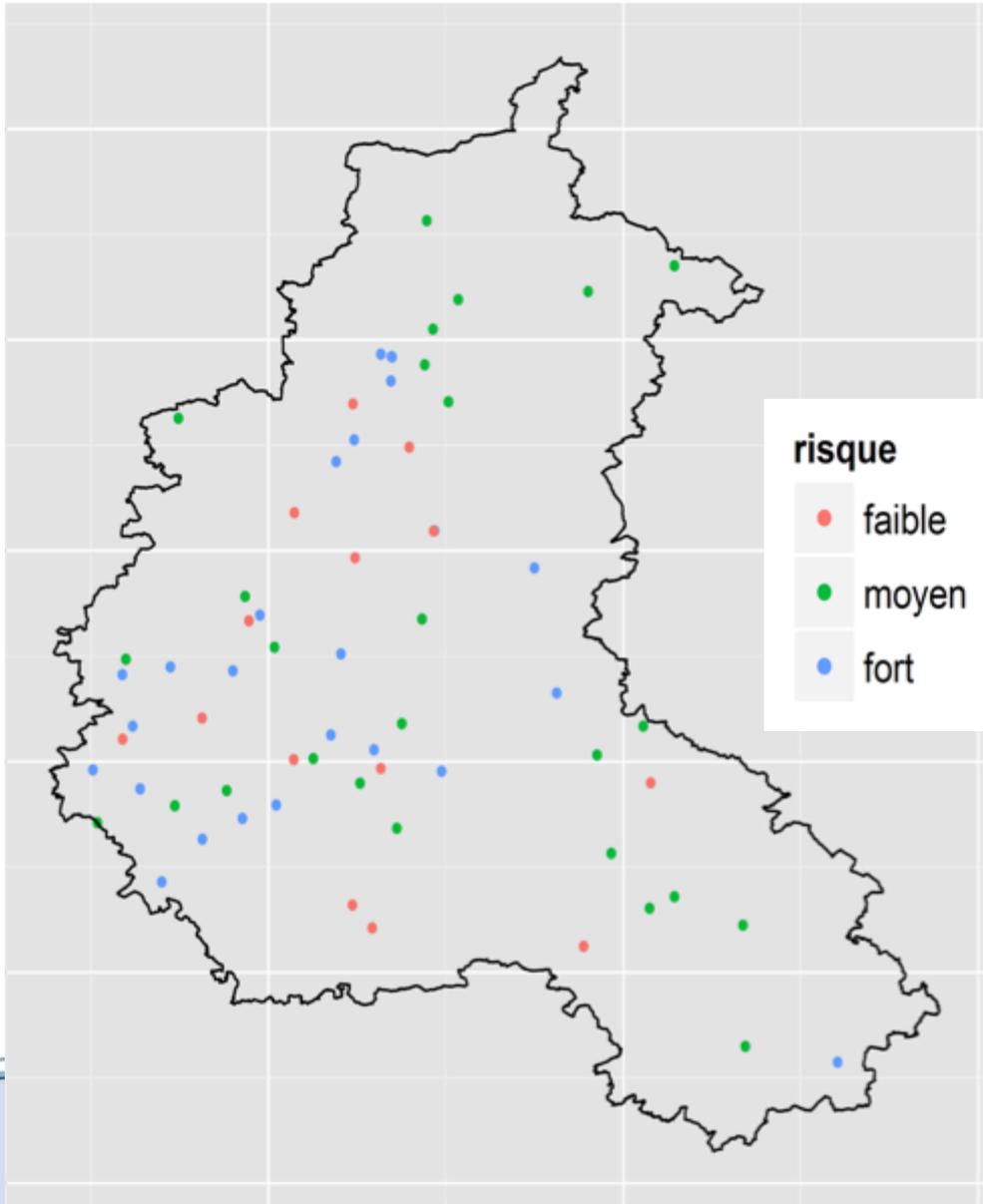


- document public et gratuit
- par région et par culture
- contenant notamment
 - état sanitaire au jour J
 - prédiction de son évolution
 - évaluation du risque (prise en compte nuisibilité)
- à partir
 - réseau régional d'observations
 - utilisation des modèles épidémiologiques
 - suivis biologiques en laboratoire
 - expertises

Septoriose du blé en Champagne-Ardenne

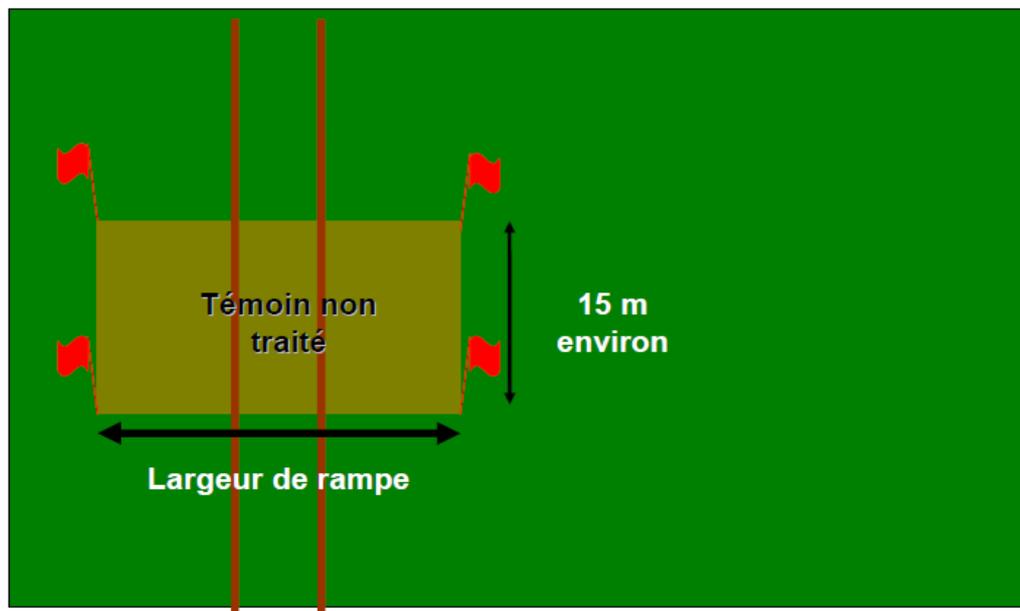


Blé Champagne-Ardenne



- **Période d'observation: épi 1cm (fin mars-début avril) à maturation (juin/juillet).**
- **Prélèvement de 20 plantes dans le témoin non traité puis observation des 3 dernières feuilles développées (F3 à F1).**
- **Comptage du nombre de feuilles touchées/20 pour F1, F2 et F3 (divisé par 2).**

témoin non traité (sans fongicides ni insecticides) pour observation

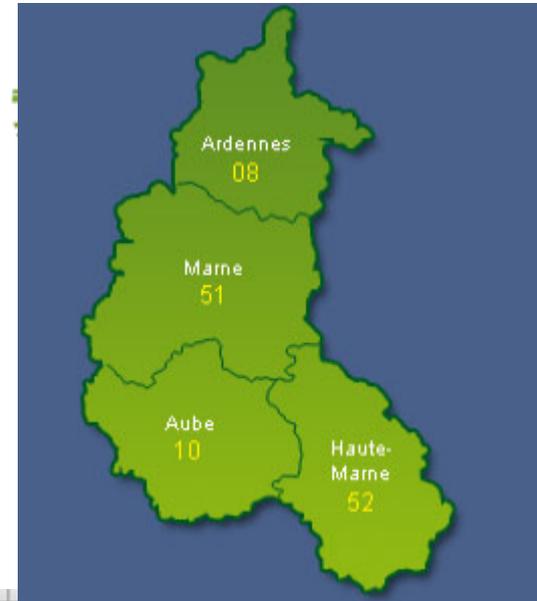


couper la rampe lors des traitements fongicides et insecticides

Méthodologie statistique

Observations

- Entre 62 et 91 sites observés par année au mieux 1 fois/semaine.
- 494 sites-années entre 2009 et 2015



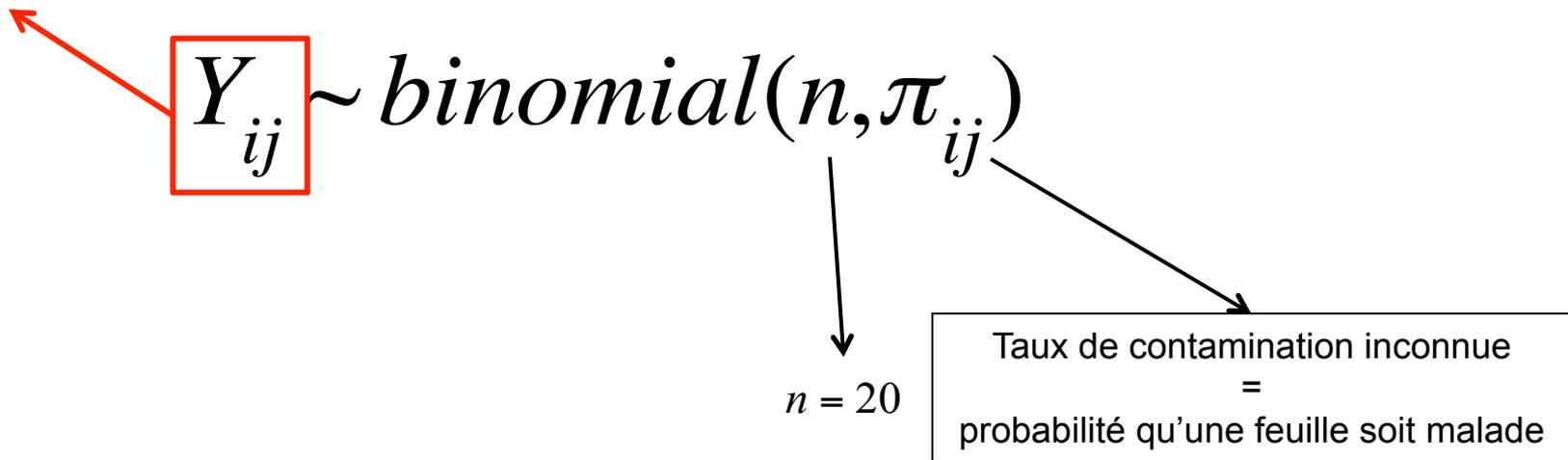
id_plot	variete	date_semis	annee	observation	obs_val_num	dateobs	sensibilite	risque	rm	rf	feuille.m	feuille.s	time
111872	CAPHORN	29/09/11	2012	F1	0	30/04/12	6	moyen	1	0	0	20	121
111872	CAPHORN	29/09/11	2012	F3	1	09/05/12	6	moyen	1	0	2	18	130
111872	CAPHORN	29/09/11	2012	F3	1	14/05/12	6	moyen	1	0	2	18	135
111872	CAPHORN	29/09/11	2012	F2	1	21/05/12	6	moyen	1	0	2	18	142
111872	CAPHORN	29/09/11	2012	F3	2	21/05/12	6	moyen	1	0	4	16	142
111872	CAPHORN	29/09/11	2012	F1	1	29/05/12	6	moyen	1	0	2	18	150
111872	CAPHORN	29/09/11	2012	F2	2	29/05/12	6	moyen	1	0	4	16	150
111872	CAPHORN	29/09/11	2012	F3	4	29/05/12	6	moyen	1	0	8	12	150
111872	CAPHORN	29/09/11	2012	F1	1	04/06/12	6	moyen	1	0	2	18	156
111872	CAPHORN	29/09/11	2012	F2	2	04/06/12	6	moyen	1	0	4	16	156
112131	SAMURAI	05/10/11	2012	F1	0	24/04/12	7	faible	0	0	0	20	115
112131	SAMURAI	05/10/11	2012	F1	0	01/05/12	7	faible	0	0	0	20	122
112131	SAMURAI	05/10/11	2012	F2	0	14/05/12	7	faible	0	0	0	20	135
112131	SAMURAI	05/10/11	2012	F3	0	14/05/12	7	faible	0	0	0	20	135
112131	SAMURAI	05/10/11	2012	F1	0	14/05/12	7	faible	0	0	0	20	135
112131	SAMURAI	05/10/11	2012	F2	0	21/05/12	7	faible	0	0	0	20	142
112131	SAMURAI	05/10/11	2012	F3	2	21/05/12	7	faible	0	0	4	16	142
112131	SAMURAI	05/10/11	2012	F1	0	21/05/12	7	faible	0	0	0	20	142
112131	SAMURAI	05/10/11	2012	F2	0	30/05/12	7	faible	0	0	0	20	151
112131	SAMURAI	05/10/11	2012	F3	2	30/05/12	7	faible	0	0	4	16	151
112411	APACHE	20/10/11	2012	F1	0	24/04/12	4.5	fort	0	1	0	20	115
112411	APACHE	20/10/11	2012	F1	0	30/04/12	4.5	fort	0	1	0	20	121
112411	APACHE	20/10/11	2012	F2	0	30/04/12	4.5	fort	0	1	0	20	121
112411	APACHE	20/10/11	2012	F1	0	07/05/12	4.5	fort	0	1	0	20	128
112411	APACHE	20/10/11	2012	F2	0	07/05/12	4.5	fort	0	1	0	20	128
112411	APACHE	20/10/11	2012	F3	1	07/05/12	4.5	fort	0	1	2	18	128
112411	APACHE	20/10/11	2012	F1	0	14/05/12	4.5	fort	0	1	0	20	135

- **Régression logistique: modèle statistique permettant d'étudier les relations entre**
 - **Un ensemble de variables X**
 - **Le temps**
 - **Une variable Y**
 - **Les données d'observations**

- **On veut prédire Y connaissant X = modélisation du nombre de feuille malade en fonction du temps.**

- **Modèle mixte qui prend en compte**
 - **un effet site-année aléatoire**
 - **Un effet groupe de risque fixe**

Nombre de feuilles
malades parmi n pour la $j^{\text{ème}}$ observation du site-année i .



➤ **Modèle à effet site-année aléatoire et effet groupe de risque fixe**

$$\text{logit}(\pi_{ij}) = \alpha_0 + b_{0i} + g_M \times Z_{Mi} + g_H \times Z_{Hi} + (\alpha_1 + b_{1i}) \times t_{ij}$$

α : paramètre médian

β : paramètre effet site-année

g : paramètre effet groupe de risque

Z_M & Z_H : variable binaire (0,1)

$Z_M = 1$ si le site-année i est caractérisé par un groupe de risque moyen et $Z_M = 0$ sinon

(idem pour Z_H)

i : Indice du site-année

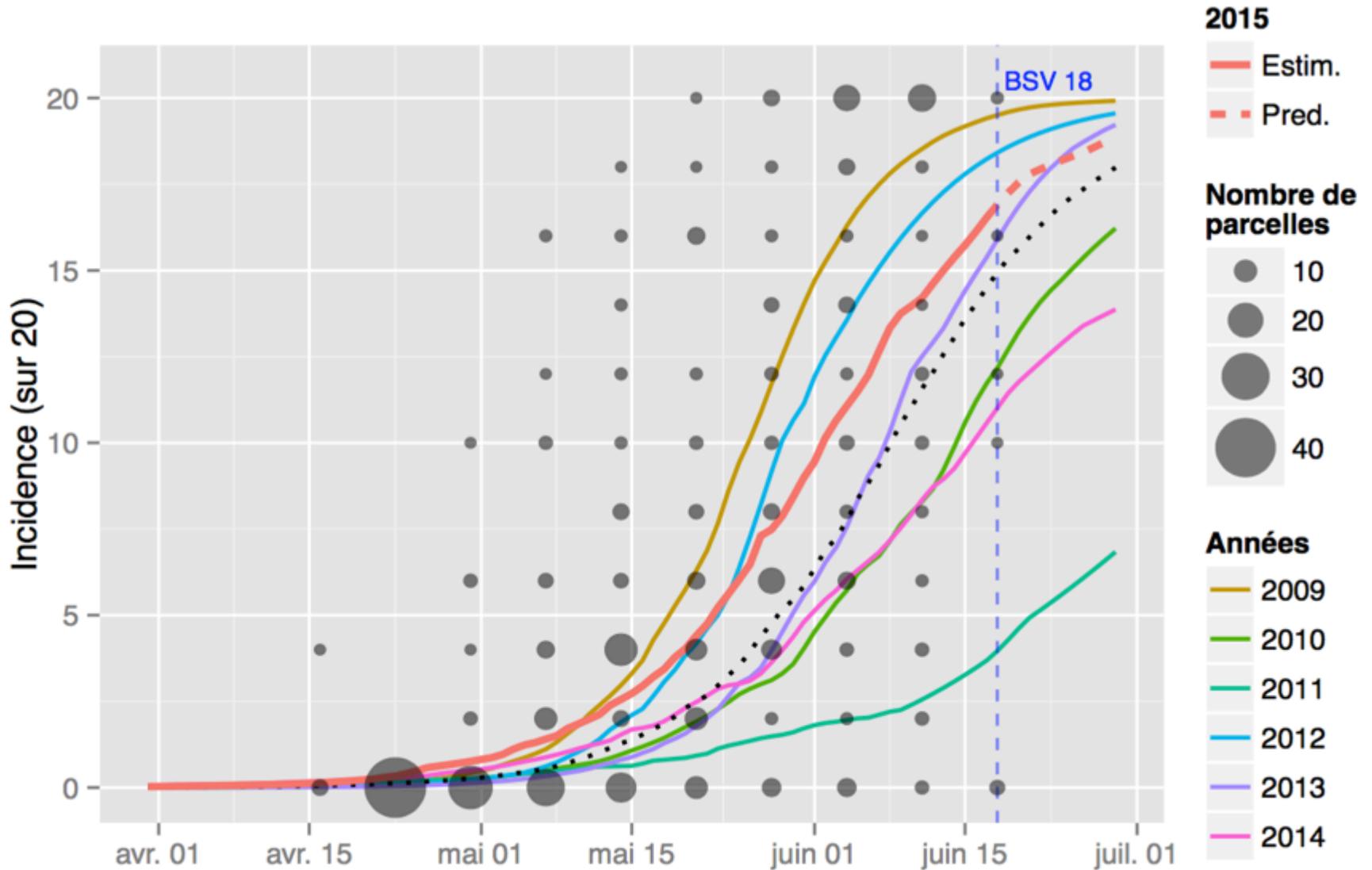
j : Indice de l'observation

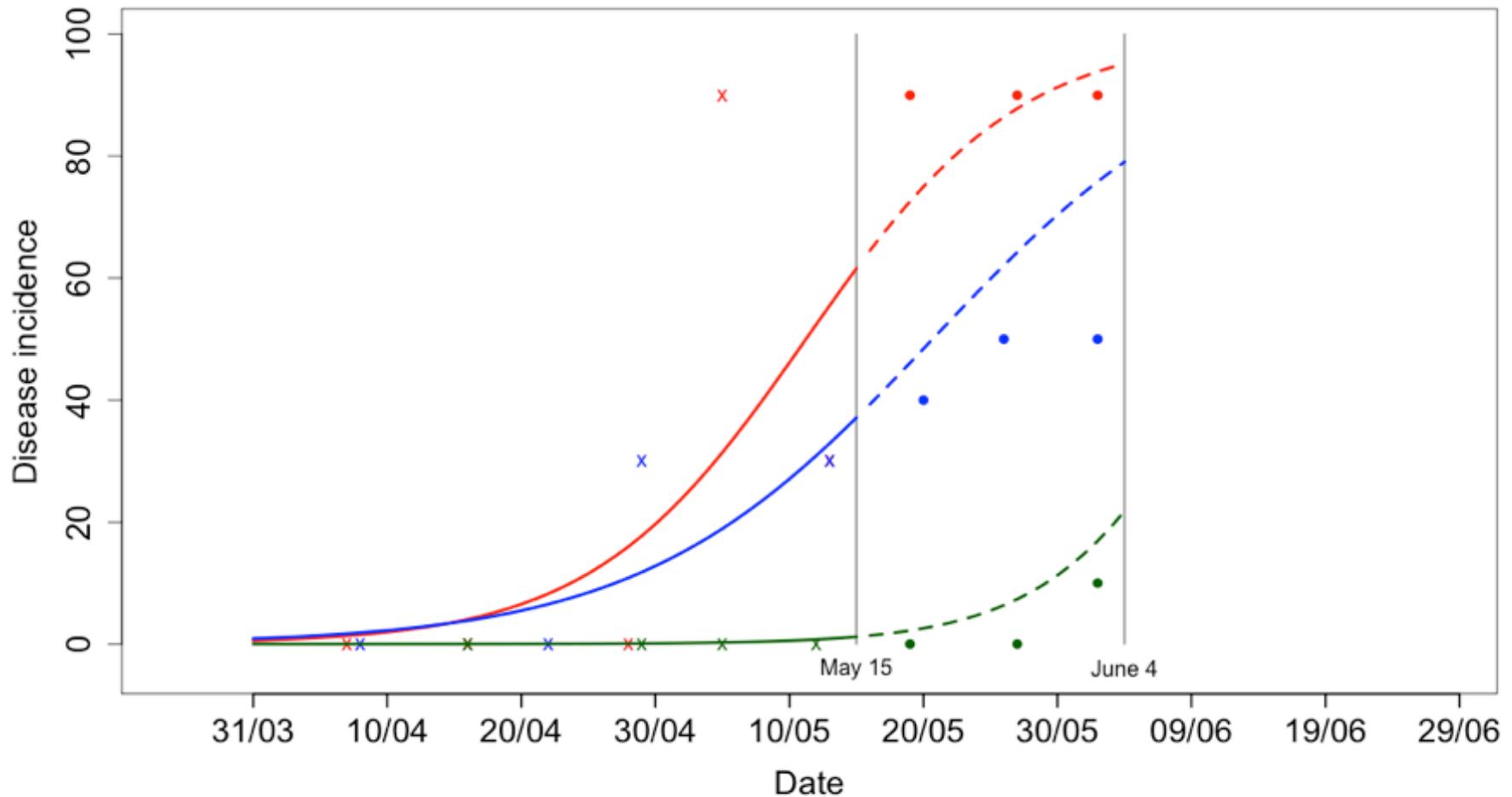
➤ **Modèle à effet site-année aléatoire et effet groupe de risque fixe**

$$\pi_{ij} = \frac{\exp(\alpha_0 + b_{0i} + g_M \times Z_M + g_H \times Z_H + (\alpha_1 + b_{1i}) \times t_{ij})}{1 + \exp(\alpha_0 + b_{0i} + g_M \times Z_M + g_H \times Z_H + (\alpha_1 + b_{1i}) \times t_{ij})}$$

Ce qui nous permet de faire les graphiques de prédictions

Finalité:
**outil d'analyse de la dynamique de l'incidence de la maladie à
l'échelle régionale et du groupe de risque septoriose.**





Comparaison: Fréquentiste VS Bayésien

- écriture du modèle**
- paramètres du modèles**
- RMSE**
- estimation**

Fréquentiste VS Bayésien

```
Model <- glmer(cbind(feuille.m, feuille.s) ~ time_b + risque
+ (1 + time_b | id_plot), family=binomial, data=TAB_glob)
```

Temps de calculs



```
model {
  for (i in 1:Ndata){

    y[i]~dbin(p[i],20)

    lambda[i] <- max(0,(alpha1+theta[siteyear[i],2]))

    logit(p[i]) <- alpha0+theta[siteyear[i],1]
    +lambda[i]*(time[i])+betarm*rm[i]+betarf*rf[i]
  }

  for (j in 1:Nsiteyear) {
    theta[j,1:2]~dmnorm(mu[1:2],tau.theta[1:2,1:2])
  }

  #prior
  alpha0~dnorm(0,1.E-6)
  alpha1~dnorm(0,1.E-6)
  tau.theta[1:2,1:2] ~ dwish(R[1:2,1:2],2)
  betarm~dnorm(0,1.E-6)
  betarf~dnorm(0,1.E-6)
}
```

Temps de calculs



Leaf 1

Parameters	Model 4F	Model 4B
α_0	-16.971 (-18.503 ; -15.438)	-17.230 (-18.57 ; -16.04)
α_1	0.238 (0.214 ; 0.263)	0.243 (0.224 ; 0.263)
g_H	2.550 (1.439 ; 3.661)	2.649 (1.624 ; 3.696)
g_M	0.942 (-0.096 ; 1.980)	1.005 (0.048 ; 2.014)

Leaf 2

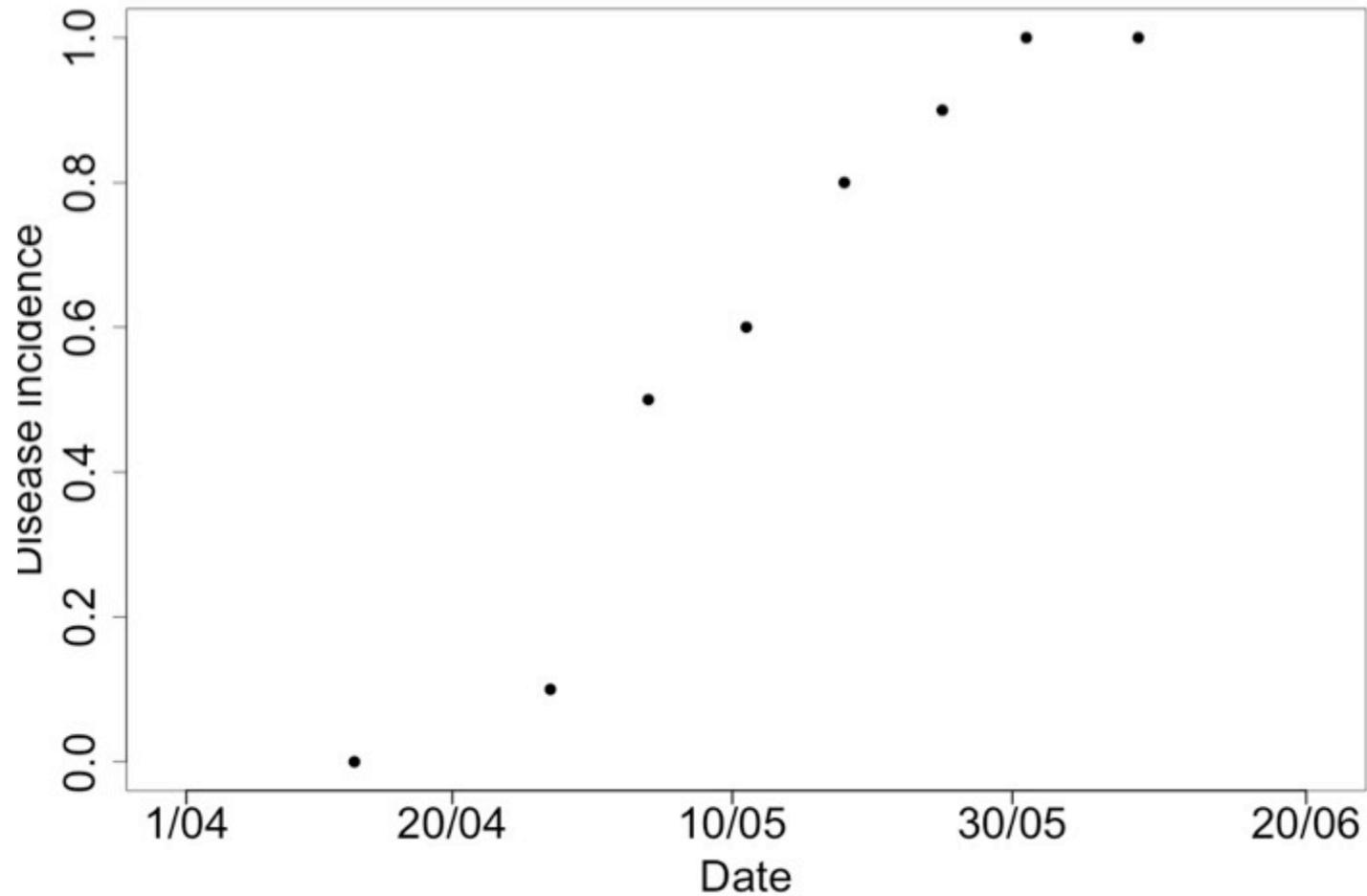
Parameters	Model 4F	Model 4B
α_0	-11.641 (-12.472 ; -10.811)	-11.717 (-12.5 ; -10.88)
α_1	0.168 (0.155 ; 0.181)	0.17 (0.156 ; 0.183)
g_H	1.766 (1.072 ; 2.459)	1.804 (1.135 ; 2.517)
g_M	0.250 (-0.394 ; 0.894)	0.27 (-0.375 ; 0.942)

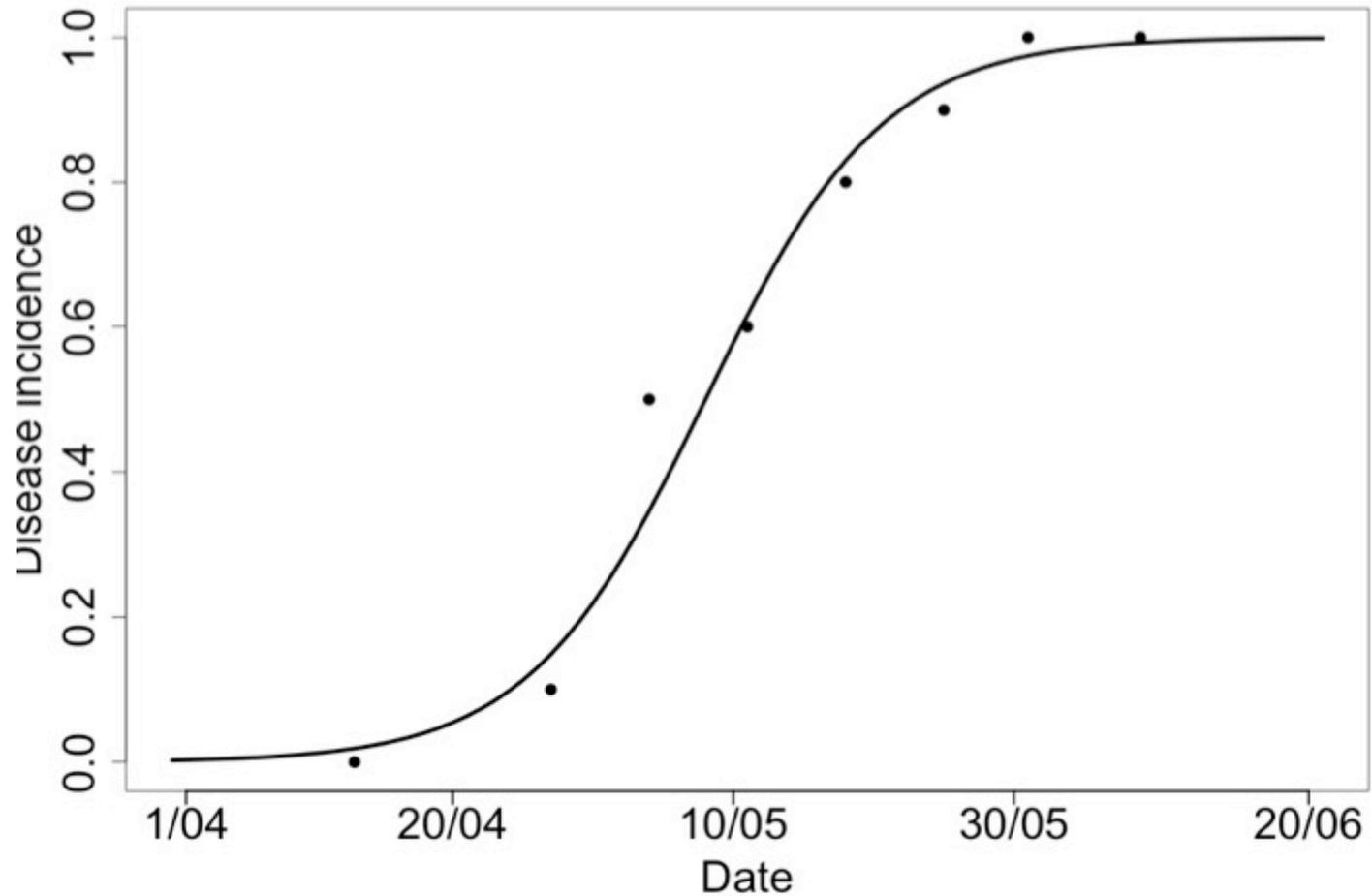
Leaf 3

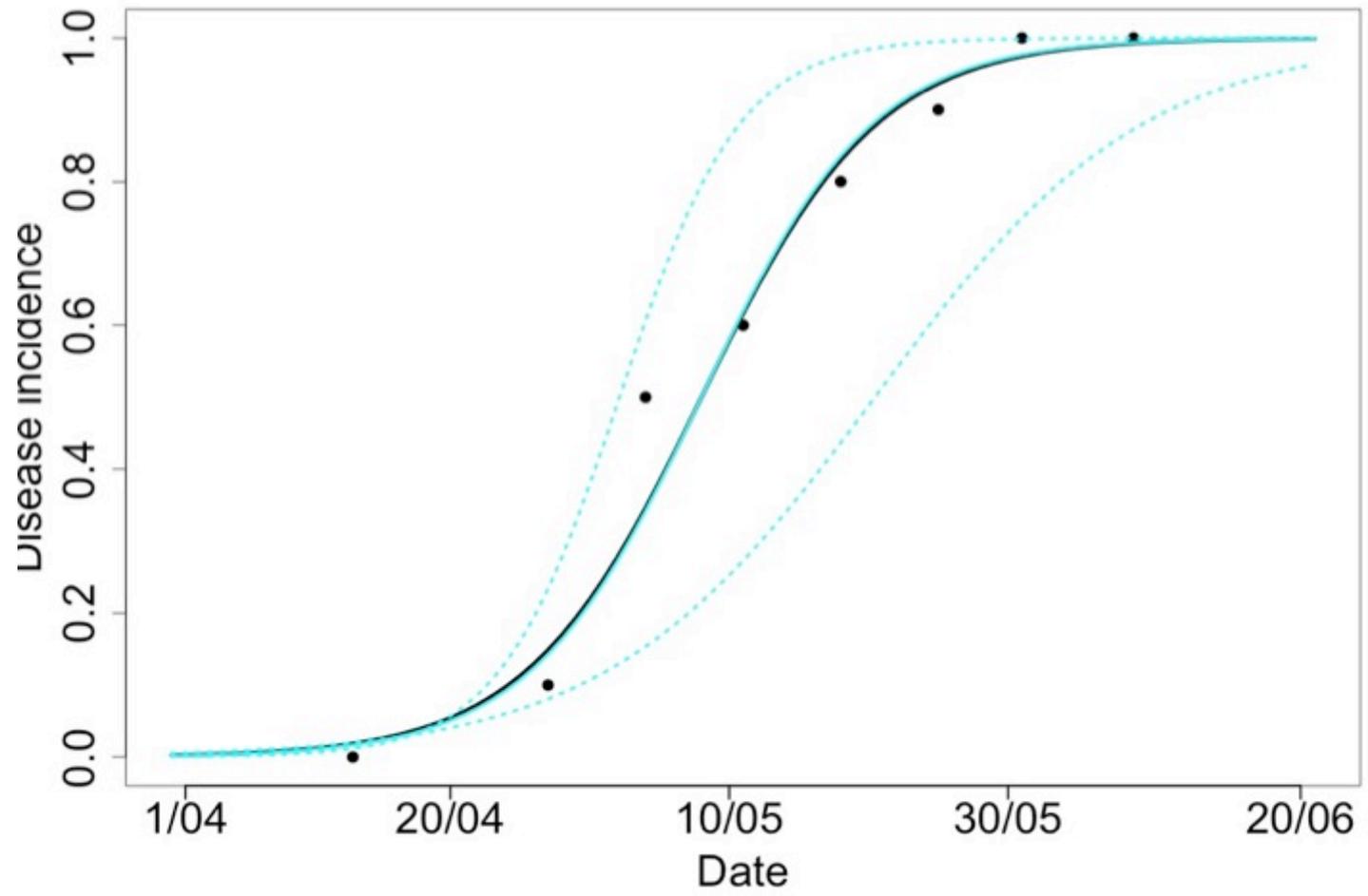
Parameters	Model 4F	Model 4B
α_0	-9.519 (-10.176 ; -8.862)	-9.389 (-10.15 ; -8.688)
α_1	0.133 (0.123 ; 0.144)	0.131 (0.118 ; 0.143)
g_H	1.195 (0.648 ; 1.743)	1.231 (0.753 ; 1.782)
g_M	0.509 (0.011 ; 1.007)	0.533 (0.03 ; 1.057)

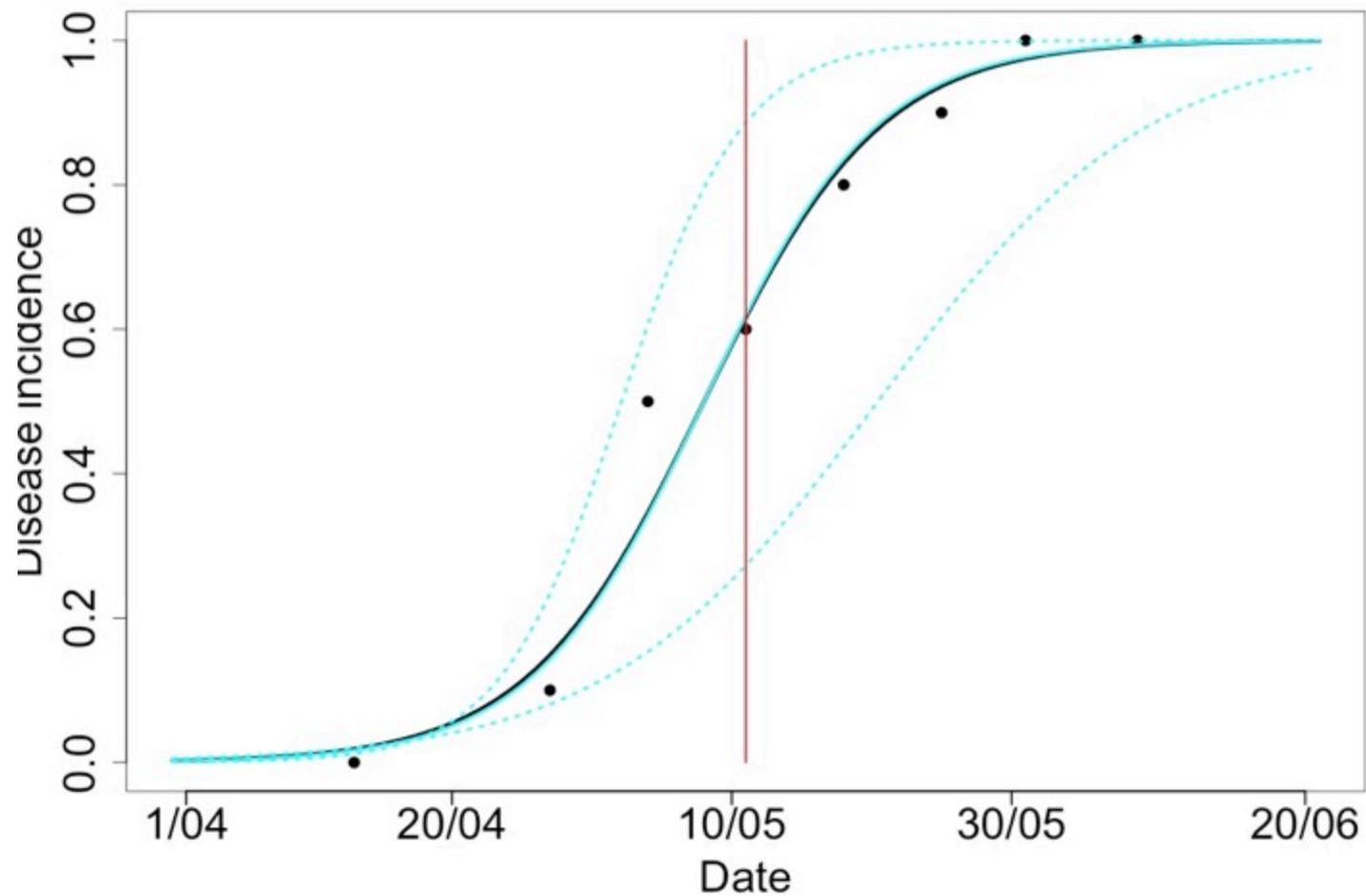
$$RMSE = \sqrt{\frac{1}{M} \sum_{i,j} (O_{ij} - P_{ij})^2}$$

Model	Statistical method	Risk level effect	Site-year random effect	Leaf 1	Leaf 2	Leaf 3
				RMSE	RMSE	RMSE
4F	Frequentist	Yes	Yes	24.33 (5.37)	26.13 (7.3)	28.23 (10.37)
4B	Bayesian	Yes	Yes	24.32 (5.44)	26.1 (7.4)	29.18 (10.52)









Conclusion dans notre cas d'étude

- **Paramètres du modèles, RMSE et RMSEP sont semblables.**
- **Le temps de calculs est plus long pour le Bayésien**
- **Intérêts du Bayésien: distribution de probabilité pour les outputs du modèles**