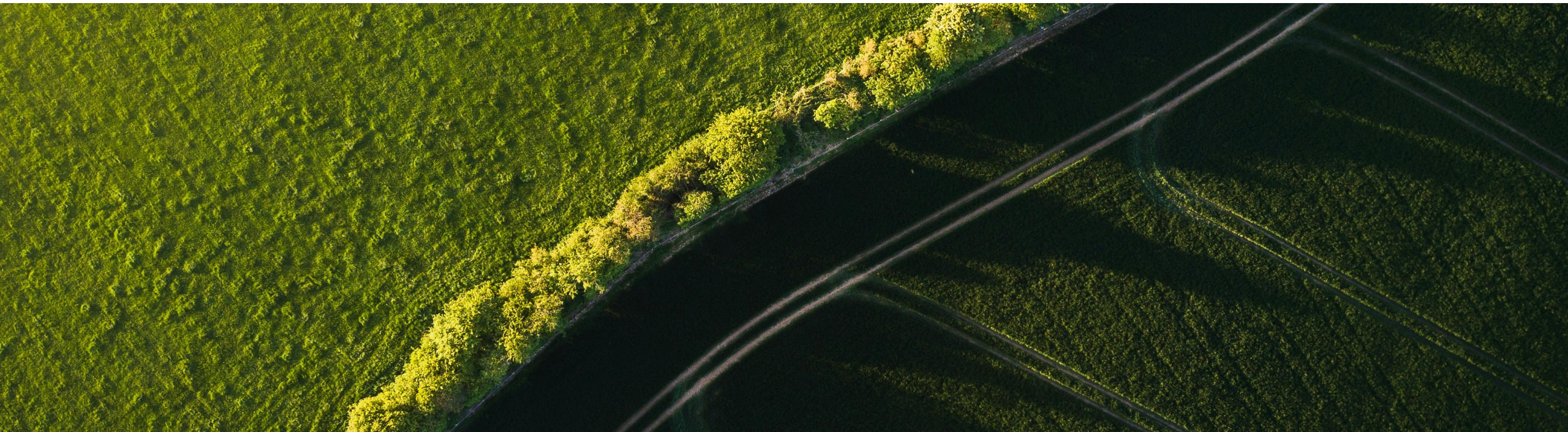


Utilisation de la méthode Doubly Robust pour évaluer l'effet de la présence de légumineuses à graines sur les performances des fermes du réseau DEPHY

Romain Nandillon - 13 février 2025



- En Europe, décroissance de la surface dédiée à la culture de légumineuses de 5.8M ha in 1961 à 1.8M ha en 2013¹.
- En France, seul 3% de la surface agricole dédiée à la culture de légumineuses en 2016².



¹(Zander et al., 2016)

²(Magrini et al., 2016)

- En Europe, décroissance de la surface dédiée à la culture de légumineuses de 5.8M ha in 1961 à 1.8M ha en 2013¹.
- En France, seul 3% de la surface agricole dédiée à la culture de légumineuses en 2016².
- Moins de progrès génétiques et agronomiques.



¹(Zander et al., 2016)

²(Magrini et al., 2016)



- En Europe, décroissance de la surface dédiée à la culture de légumineuses de 5.8M ha in 1961 à 1.8M ha en 2013¹.
- En France, seul 3% de la surface agricole dédiée à la culture de légumineuses en 2016².
- Moins de progrès génétiques et agronomiques.
- Moindre attractivité économique par rapport aux cultures dominantes.

¹(Zander et al., 2016)

²(Magrini et al., 2016)



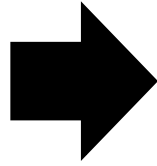
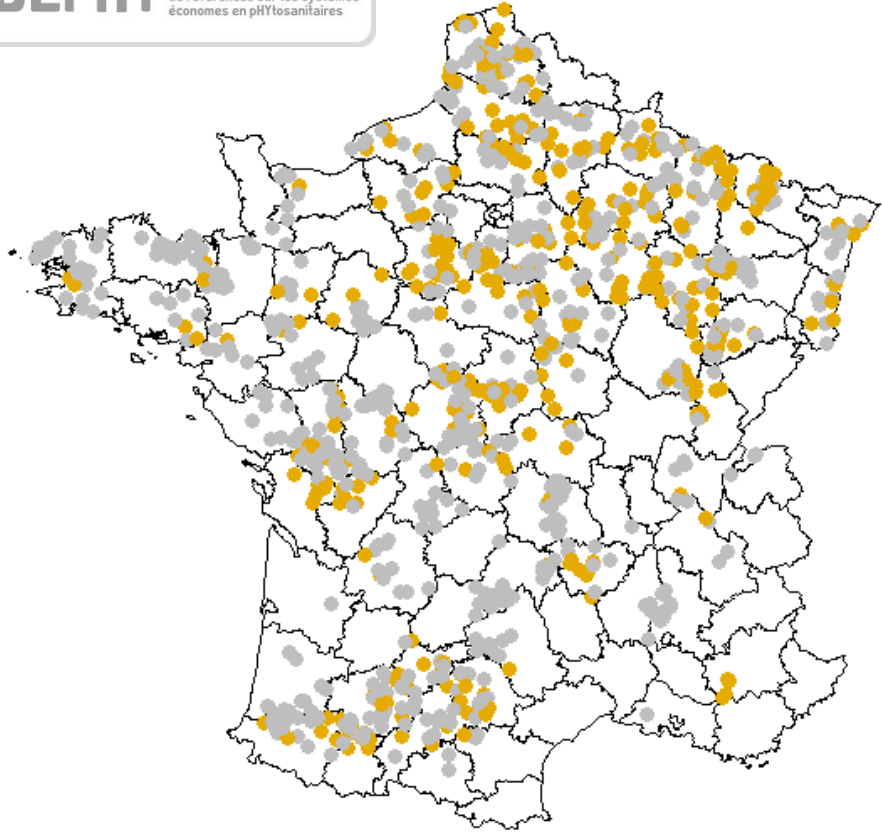
- En Europe, décroissance de la surface dédiée à la culture de légumineuses de 5.8 M ha in 1961 à 1.8M ha en 2013¹.
- En France, seul 3% de la surface agricole dédiée à la culture de légumineuses en 2016².
- Moins de progrès génétiques et agronomiques.
- Moindre atractivité économique par rapport aux cultures dominantes.

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

¹(Zander et al., 2016)

²(Magrini et al., 2016)

Fermes DEPHY en grandes cultures et en polyculture-élevage



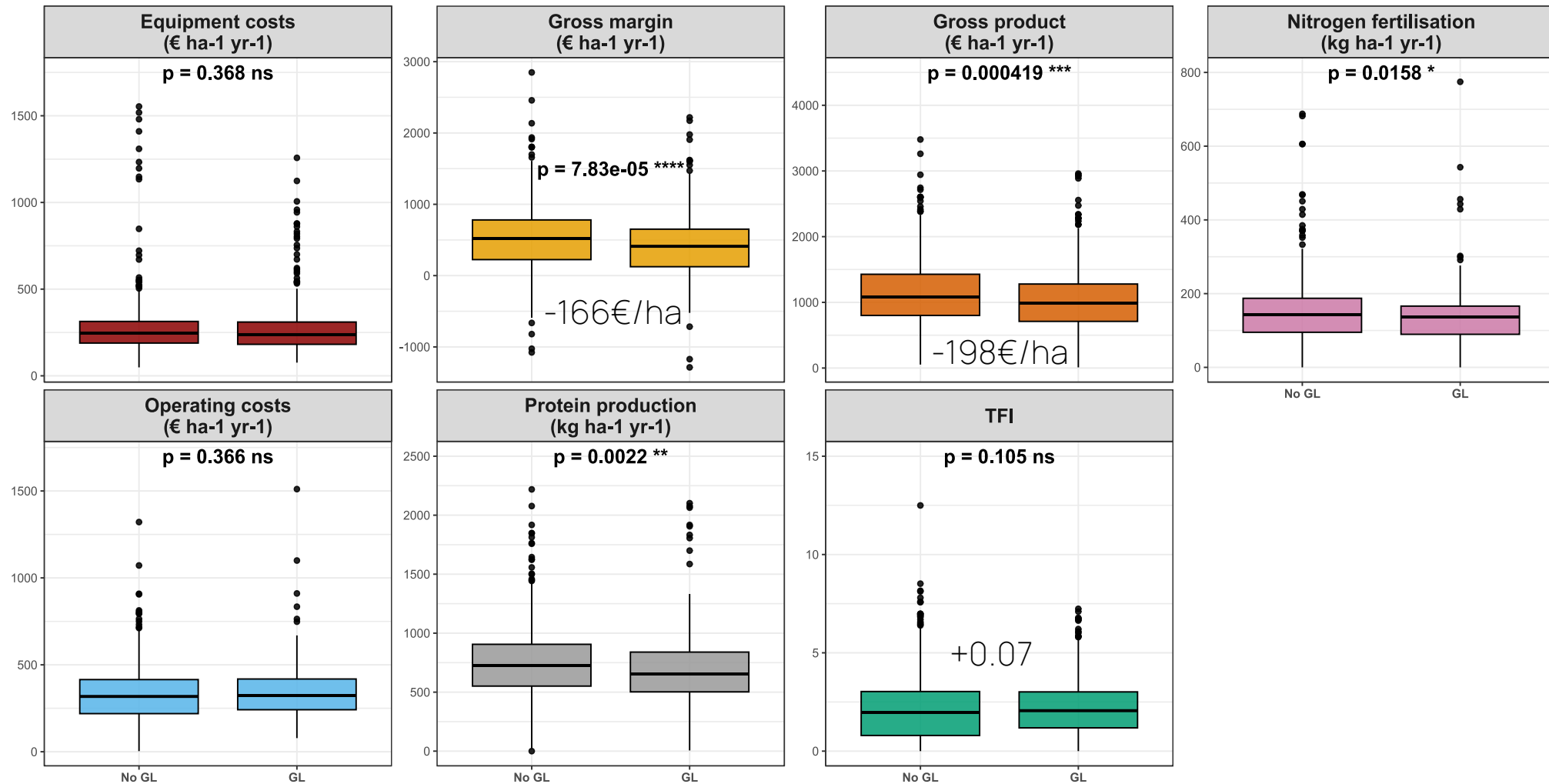
Données détaillées sur :

- les pratiques (e.g. rotations, intrants, travail du sol),
- les performances (e.g. rendements, coûts, marges brutes, temps de travail),
- les caractéristiques des systèmes.



Système de culture = 1 ferme DEPHY × 1 année

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?



Ces différences peuvent être liées au contexte

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

Méthode d'inférence causale pour estimer l'effet causal de contenir des LAG : « Doubly robust »

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

Méthode d'inférence causale pour estimer l'effet causal de contenir des LAG : « Doubly robust »

- 1108 fermes $\times$ 4.4 années = 4941 systèmes de culture en grandes cultures ou polyculture élevage

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

Méthode d'inférence causale pour estimer l'effet causal de contenir des LAG : « Doubly robust »

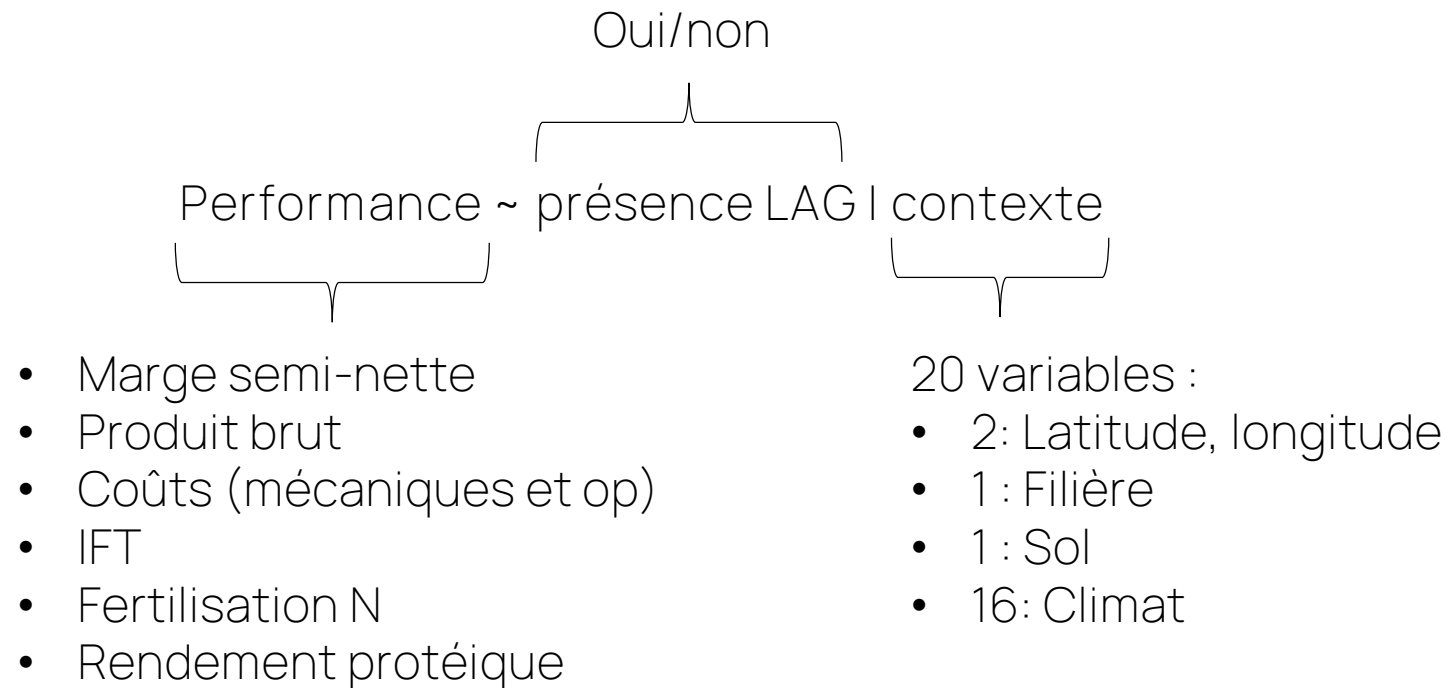
- 1108 fermes $\times$ 4.4 années = 4941 systèmes de culture en grandes cultures ou polyculture élevage

Performance $\sim$ présence LAG | contexte

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

Méthode d'inférence causale pour estimer l'effet causal de contenir des LAG : « Doubly robust »

- 1108 fermes $\times$ 4.4 années = 4941 systèmes de culture en grandes cultures ou polyculture élevage



Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

1. Modèle de propension : $TREAT \sim \text{contexte}$

2. Modèle d'outcome : $PERF \sim TREAT + \text{contexte}$



Effet moyen de contenir des LAG

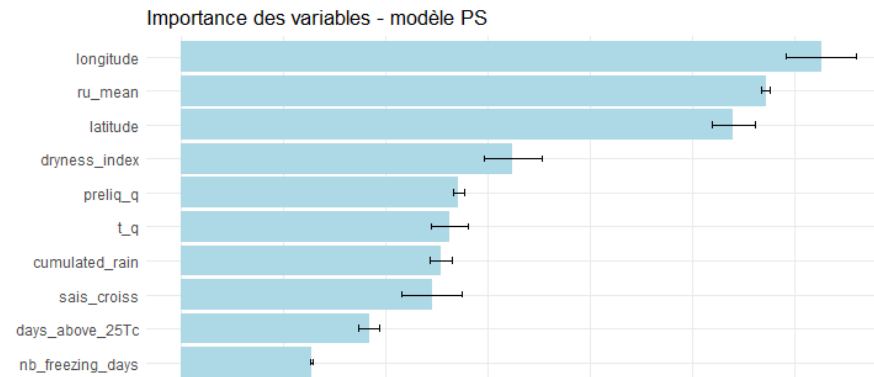
Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

1. Modèle de propension :

TREAT ~ filiere + ru_mean + t_q + preliq_q + t_base_32 +
nb_freezing_days + first_freeze + last_freeze + days_above_25Tc
+ cumulated_rain + dryness_index + sais_croiss + dt_drnj_tsup5 +
dt_1erj_tsup5 + co2 + lenght_heat_wave + nb_of_heat_waves +
nb_of_dry_periods + latitude + longitude

Sélection du meilleur GLM & GBM

Vaildation croisée sur l'année



Sélection de variable
> AUC

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

1. Modèle de propension :
$$\text{TREAT} \sim \text{filiere} + \text{ru_mean} + \text{t_q} + \text{preliq_q} + \text{t_base_32} + \text{nb_freezing_days} + \text{first_freeze} + \text{last_freeze} + \text{days_above_25Tc} + \text{cumulated_rain} + \text{dryness_index} + \text{sais_croiss} + \text{dt_drnj_tsup5} + \text{dt_1erj_tsup5} + \text{co2} + \text{lenght_heat_wave} + \text{nb_of_heat_waves} + \text{nb_of_dry_periods} + \text{latitude} + \text{longitude}$$
2. Modèle d'outcome :
$$\text{PERF} \sim \text{TREAT} + \text{filiere} + \text{ru_mean} + \text{t_q} + \text{preliq_q} + \text{t_base_32} + \text{nb_freezing_days} + \text{first_freeze} + \text{last_freeze} + \text{days_above_25Tc} + \text{cumulated_rain} + \text{dryness_index} + \text{sais_croiss} + \text{dt_drnj_tsup5} + \text{dt_1erj_tsup5} + \text{co2} + \text{lenght_heat_wave} + \text{nb_of_heat_waves} + \text{nb_of_dry_periods} + \text{latitude} + \text{longitude}$$

Même étape de sélection mais pour
chaque indicateur de performance
> R^2

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

1. Modèle de propension :

TREAT ~ filiere + ru_mean + t_q + preliq_q + t_base_32 +
nb_freezing_days + first_freeze + last_freeze + days_above_25Tc
+ cumulated_rain + dryness_index + sais_croiss + dt_drnj_tsup5 +
dt_1erj_tsup5 + co2 + lenght_heat_wave + nb_of_heat_waves +
nb_of_dry_periods + latitude + longitude

2. Modèle d'outcome :

PERF ~ **TREAT** + filiere + ru_mean + t_q + preliq_q + t_base_32 +
nb_freezing_days + first_freeze + last_freeze + days_above_25Tc
+ cumulated_rain + dryness_index + sais_croiss + dt_drnj_tsup5 +
dt_1erj_tsup5 + co2 + lenght_heat_wave + nb_of_heat_waves +
nb_of_dry_periods + latitude + longitude

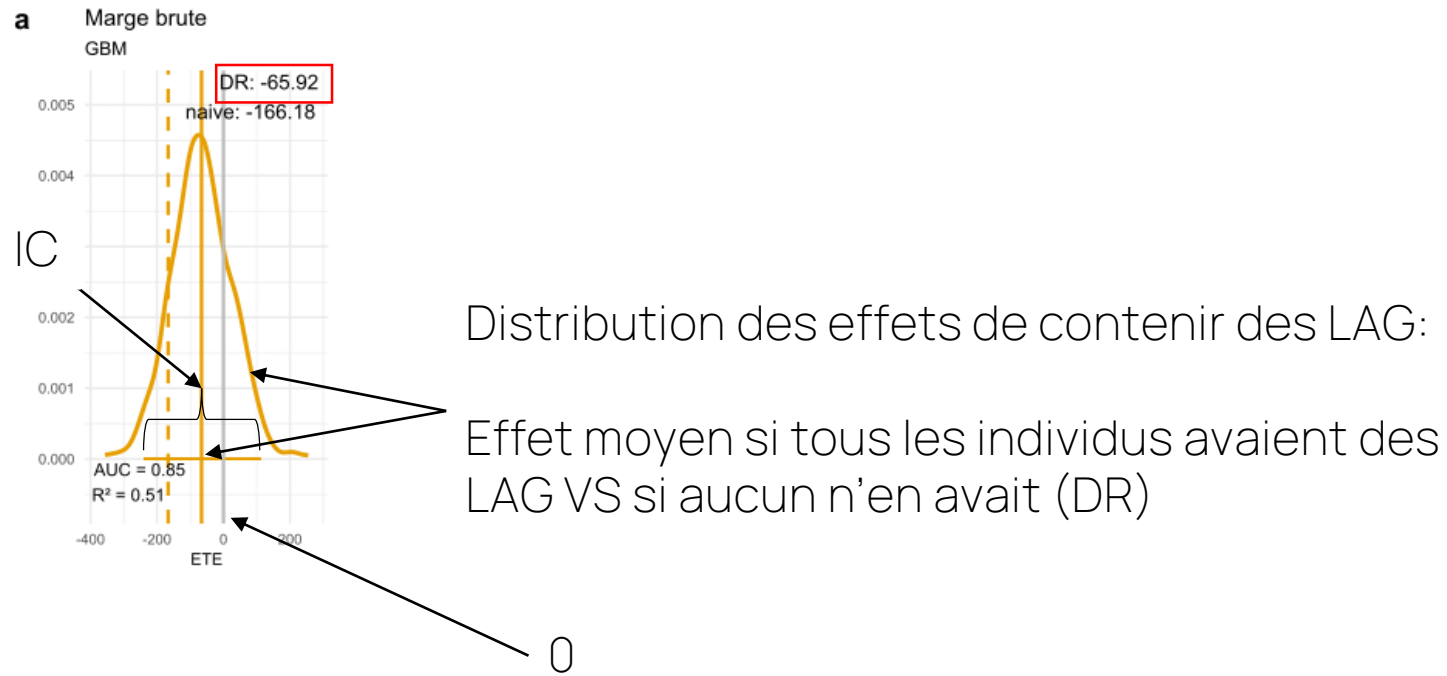


Implémentation DR avec 1000
bootstraps sur l'année

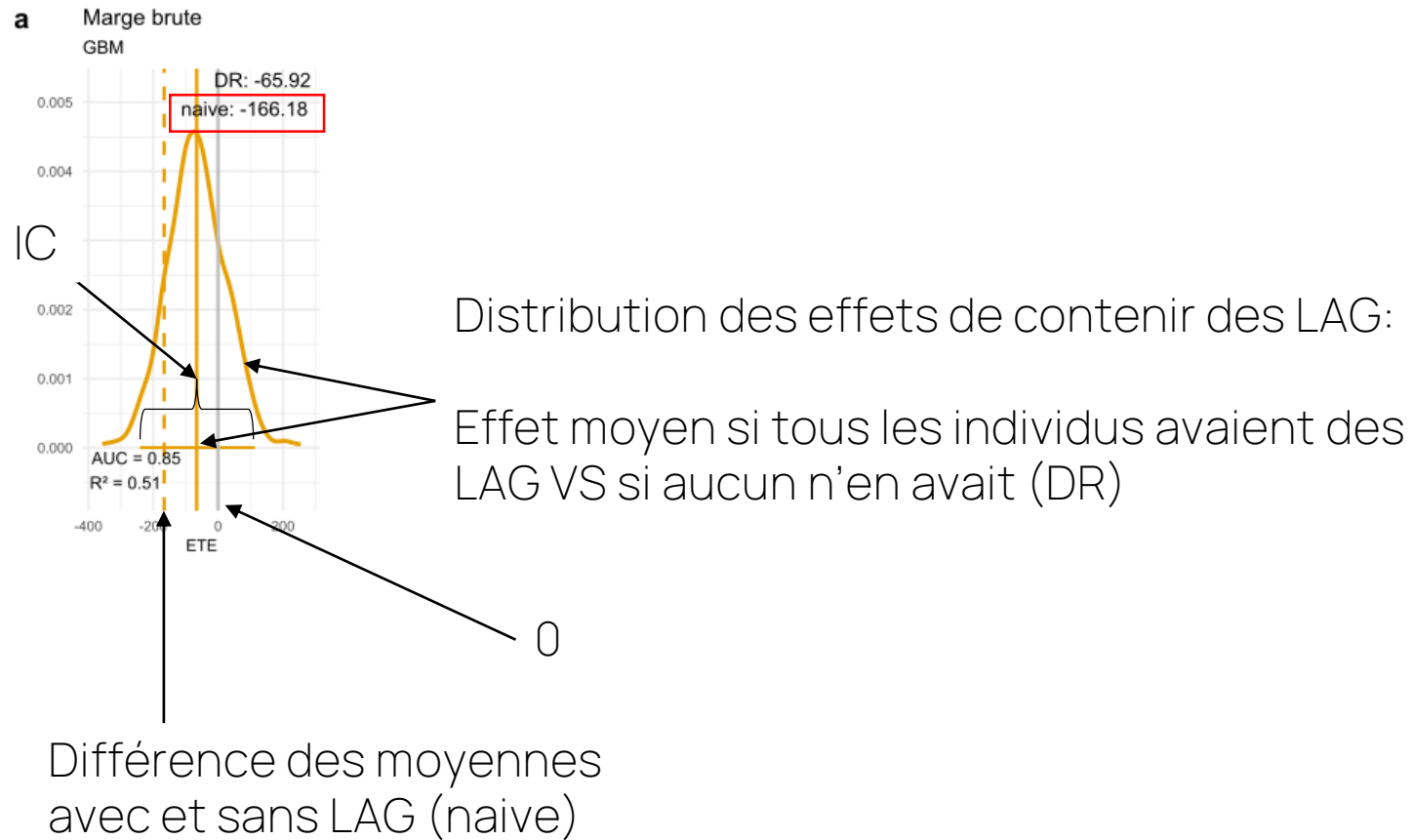


Effet moyen de contenir des LAG

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

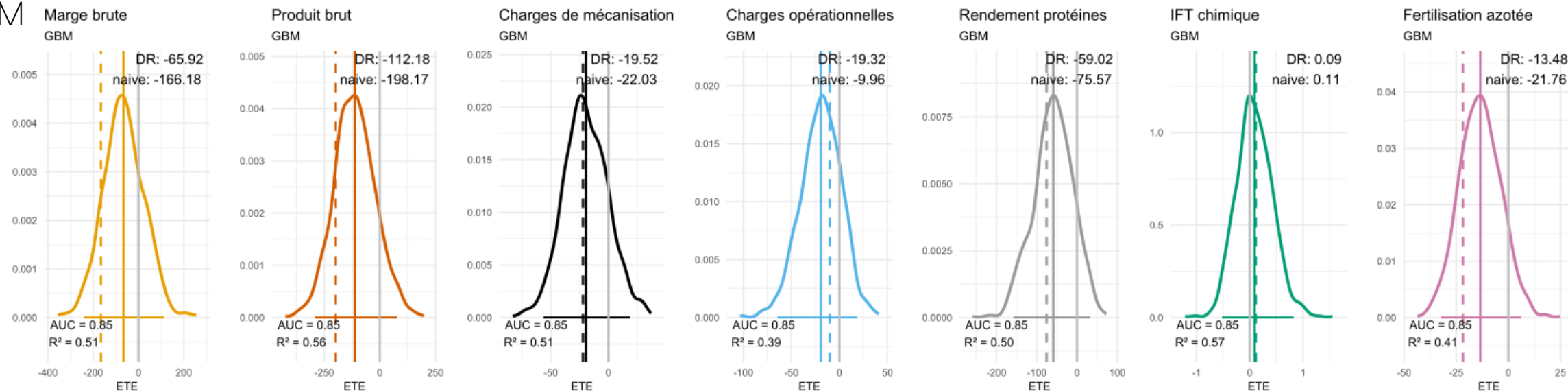


Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?



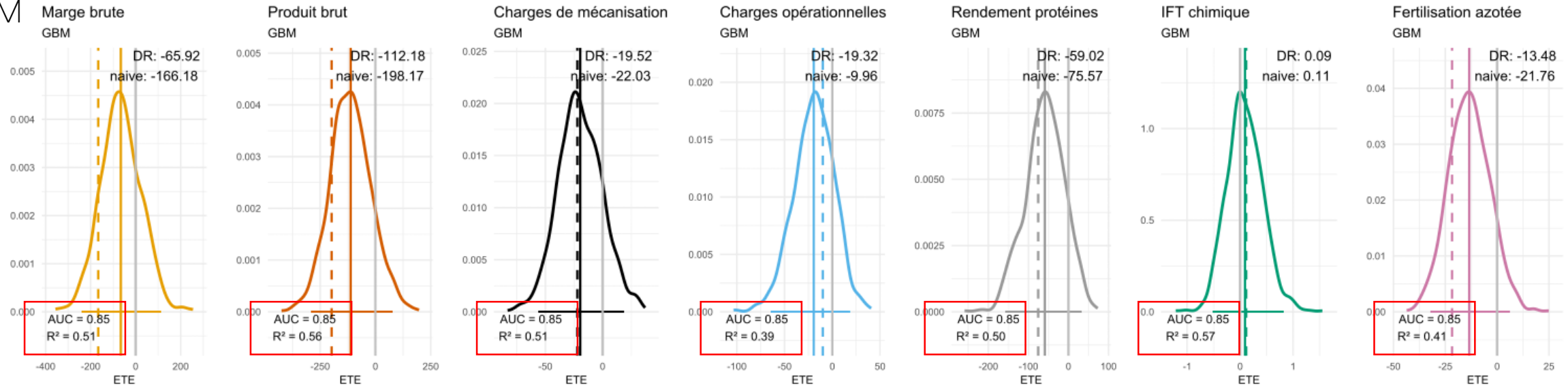
Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

GBM

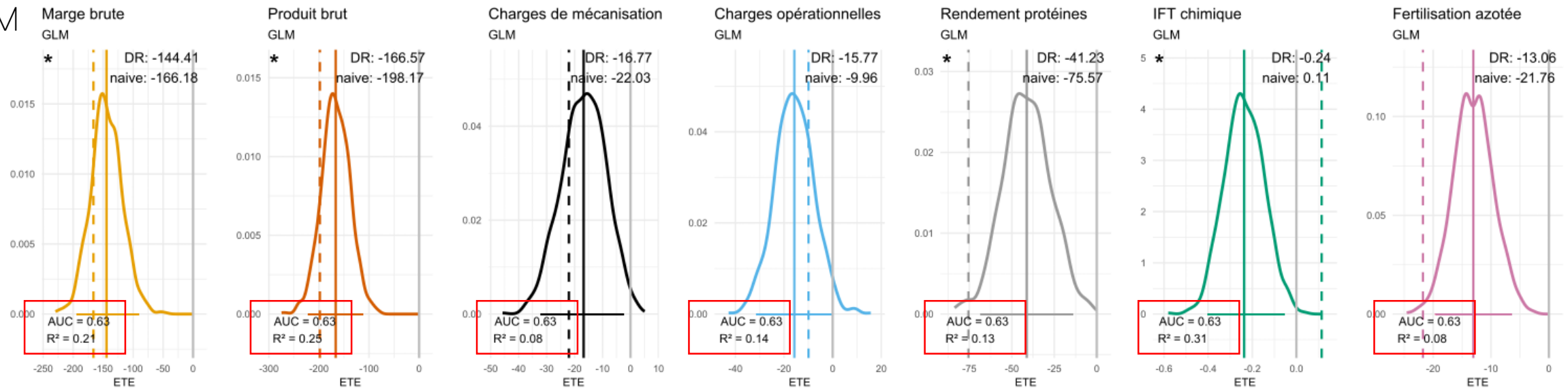


Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

GBM

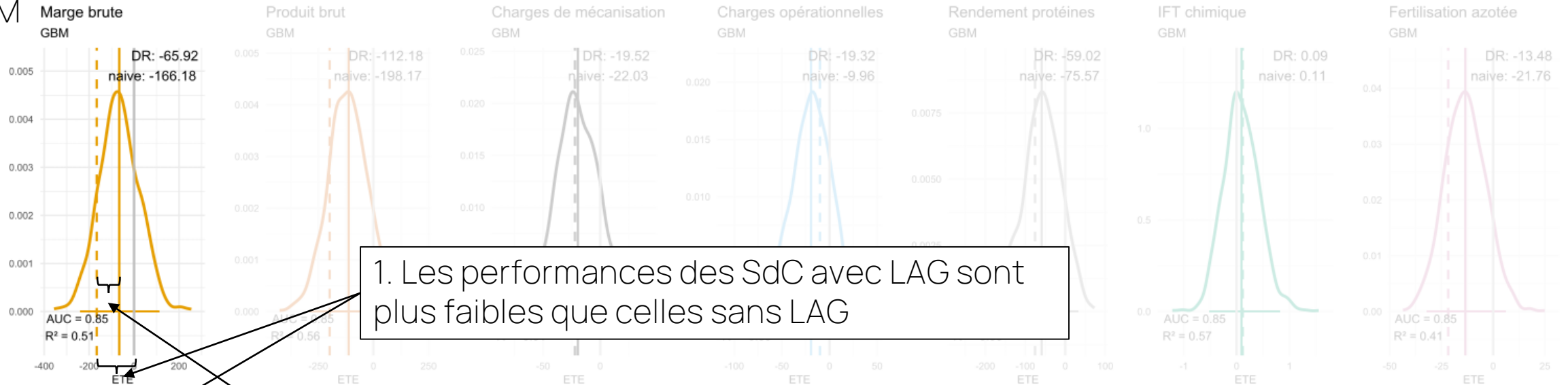


GLM



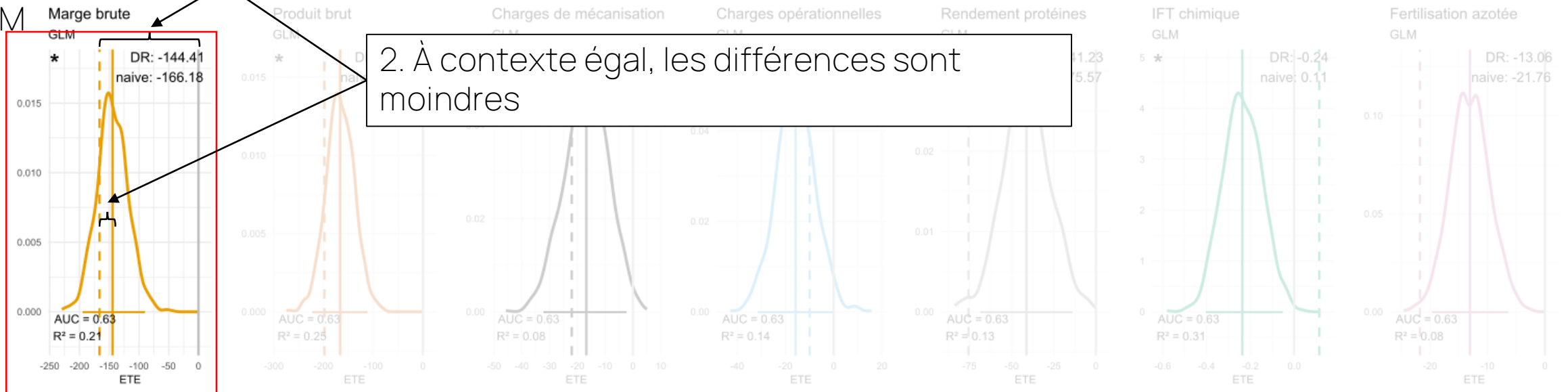
Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

GBM



1. Les performances des SdC avec LAG sont plus faibles que celles sans LAG

GLM



2. À contexte égal, les différences sont moindres

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

- Pas d'effet significatif selon les GBM
- Plusieurs effets significatifs selon les GLM
 - À contexte égal:
 - L'IFT est inférieur de 0.24 (9.6%) dans les systèmes avec LAG

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

- Pas d'effet significatif selon les GBM
- Plusieurs effets significatifs selon les GLM
 - À contexte égal:
 - L'IFT est inférieur de 0.24 (9.6%) dans les systèmes avec LAG
 - Sans prise en compte du contexte:
 - L'IFT est supérieur de 0.07 (3%) dans les systèmes avec LAG

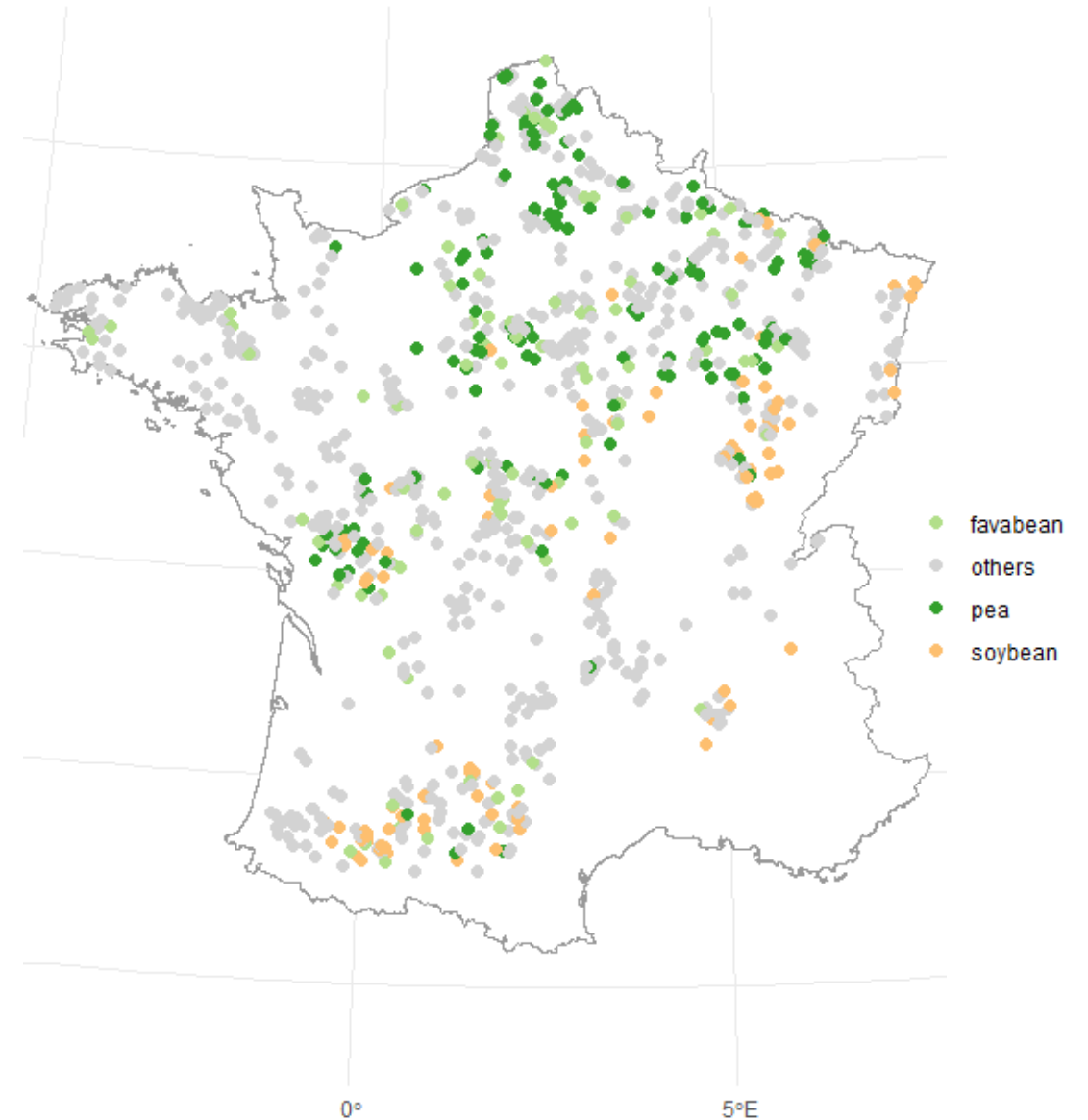
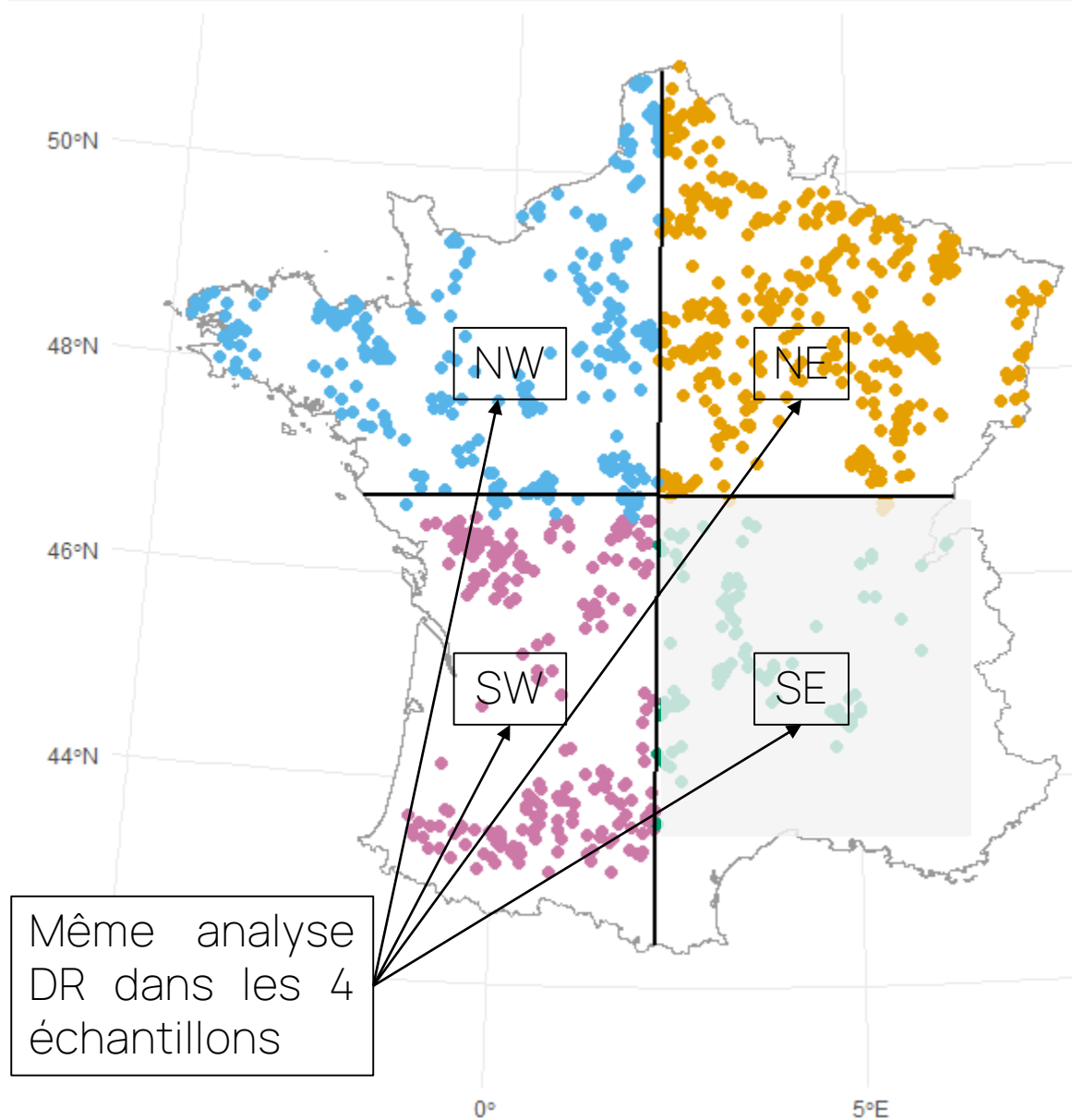
Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

- Pas d'effet significatif selon les GBM
- Plusieurs effets significatifs selon les GLM
 - À contexte égal:
 - L'IFT est inférieur de 0.24 (9.6%) dans les systèmes avec LAG
 - Le produit brut est inférieur de -167 €/ha (12%) dans les systèmes avec LAG
 - Sans prise en compte du contexte:
 - L'IFT est supérieur de 0.07 (3%) dans les systèmes avec LAG

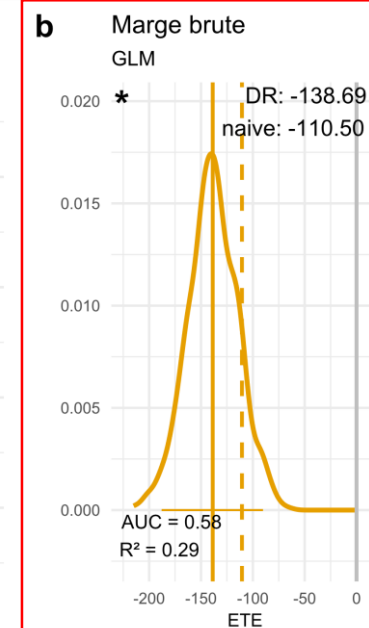
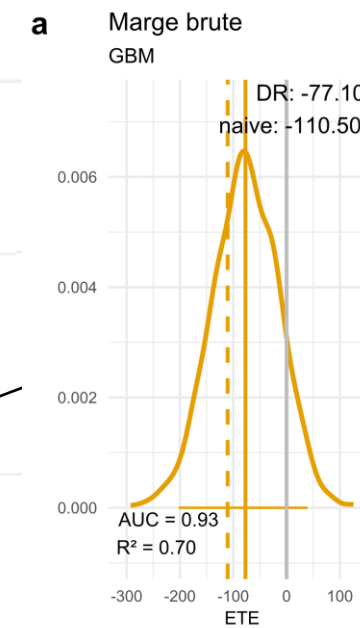
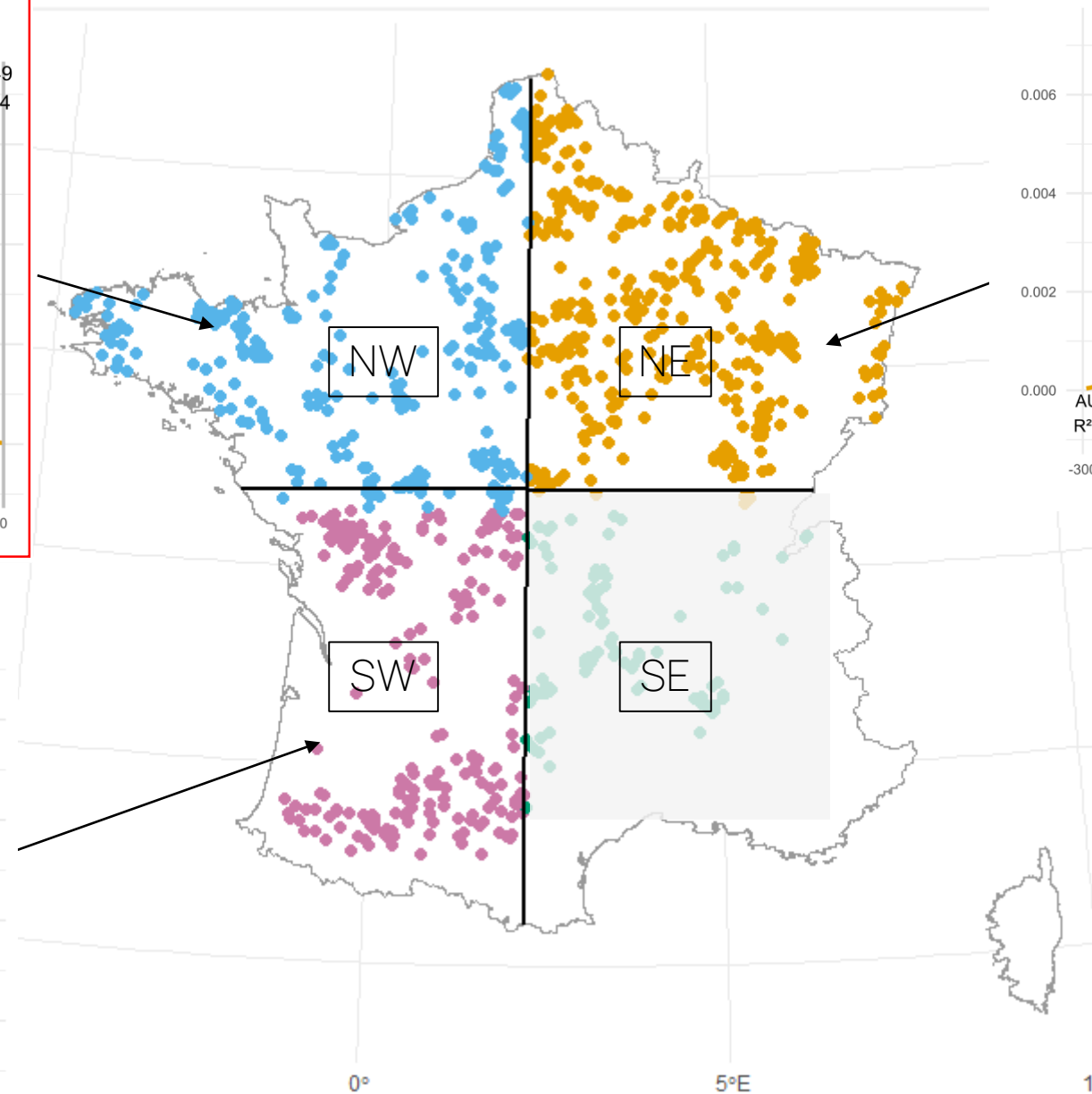
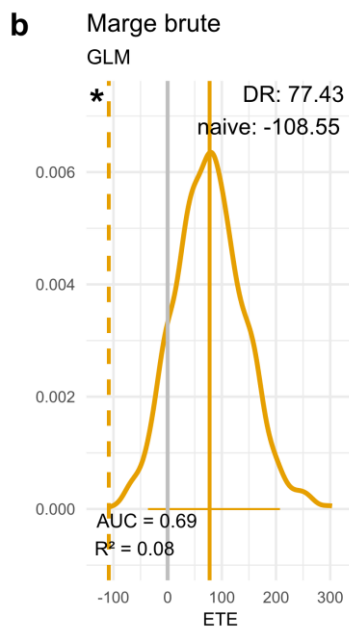
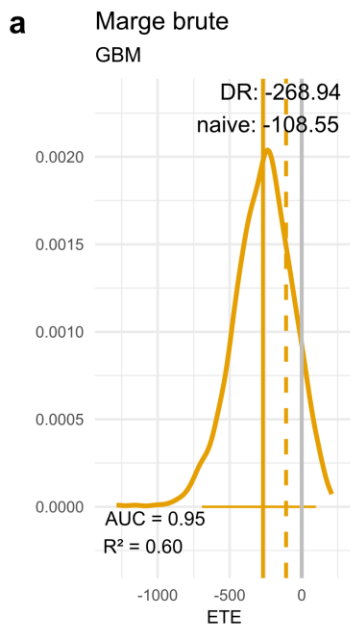
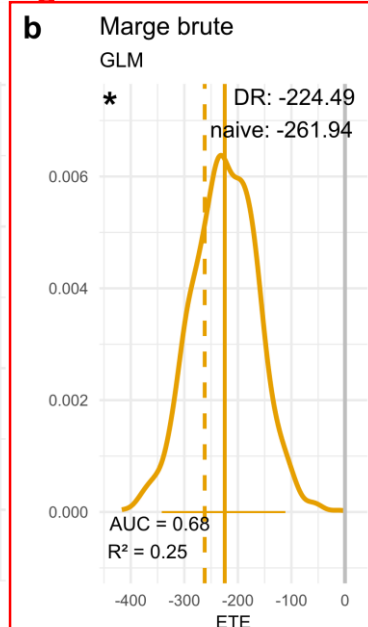
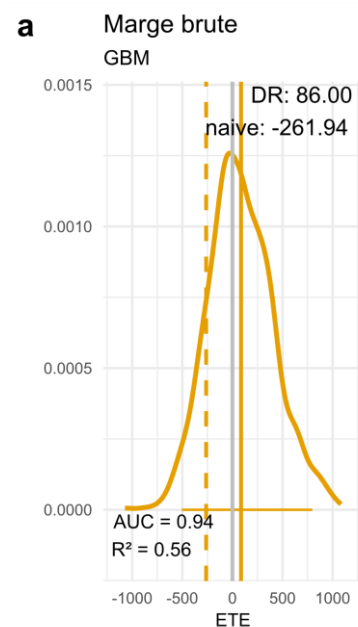
Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

- Pas d'effet significatif selon les GBM
- Plusieurs effets significatifs selon les GLM
 - À contexte égal:
 - L'IFT est inférieur de 0.24 (9.6%) dans les systèmes avec LAG
 - Le produit brut est inférieur de -167 €/ha (12%) dans les systèmes avec LAG
 - Sans prise en compte du contexte:
 - L'IFT est supérieur de 0.07 (3%) dans les systèmes avec LAG
 - Le produit brut est inférieur de -198 €/ha (15%) dans les systèmes avec LAG

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

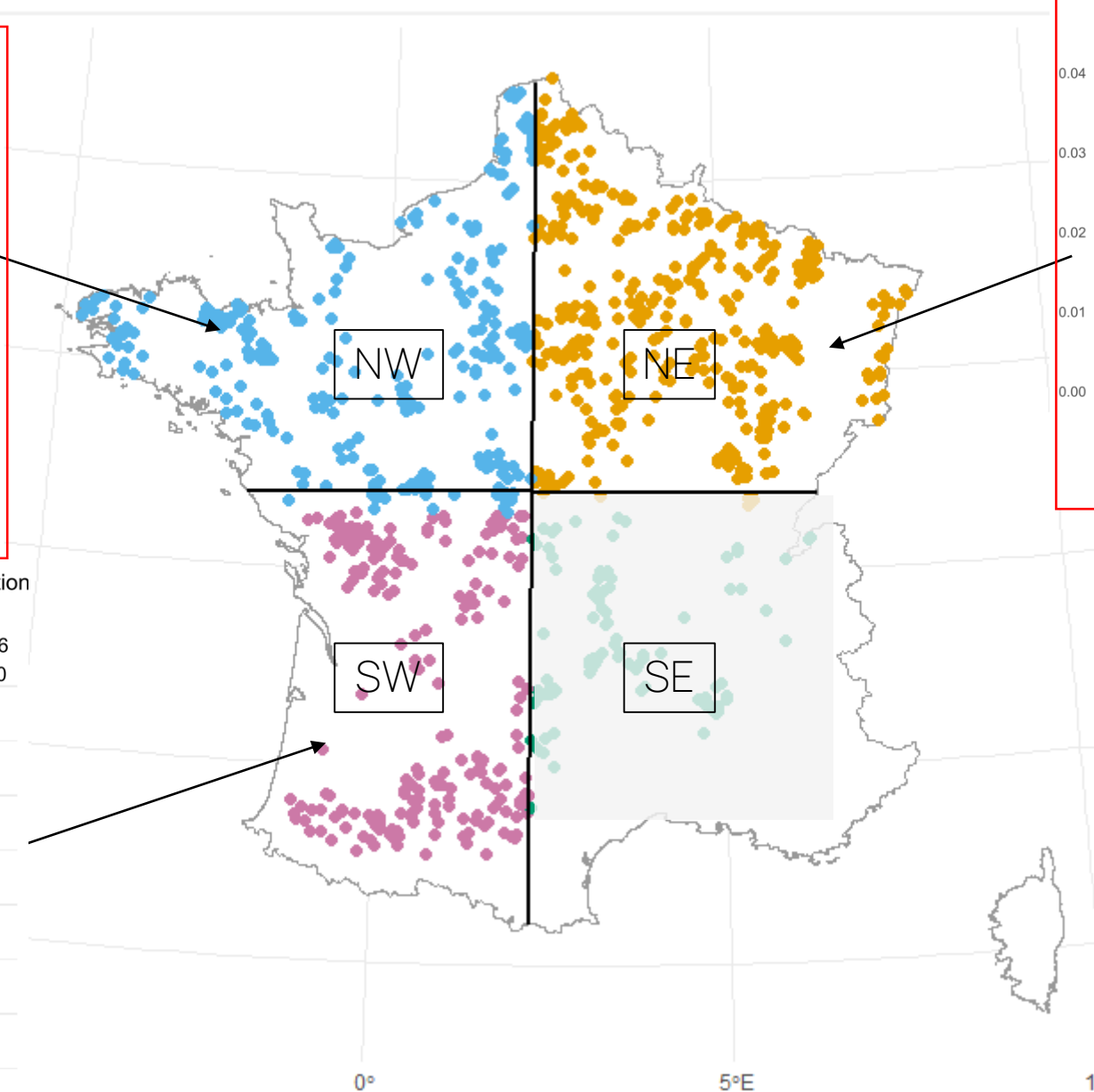
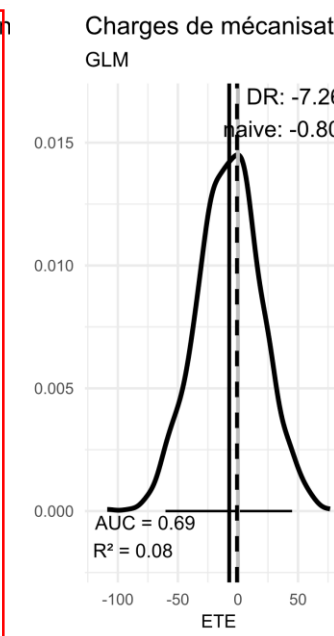
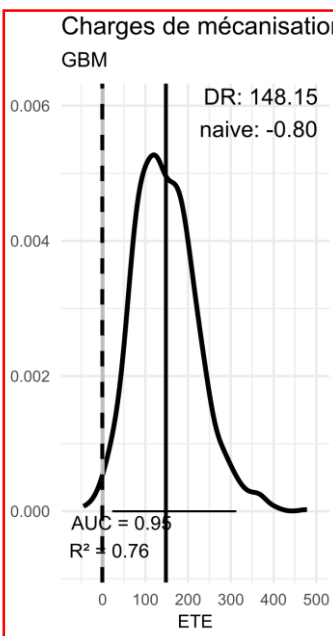
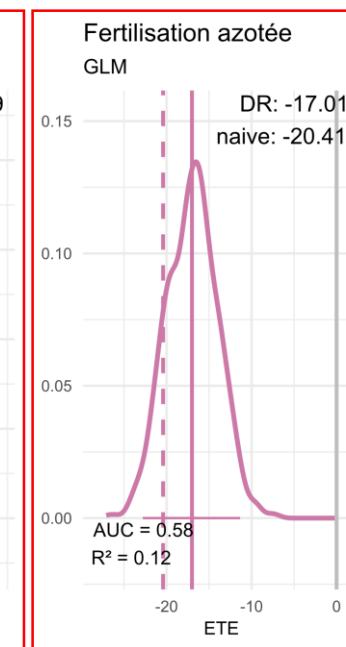
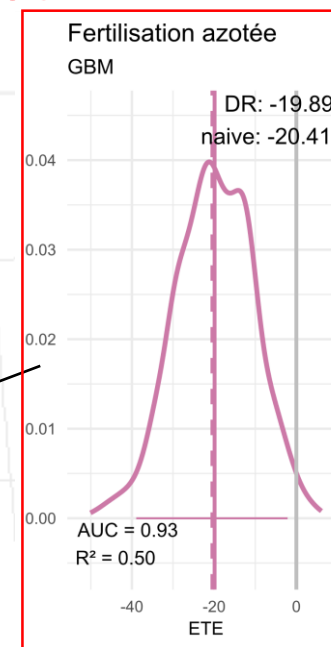
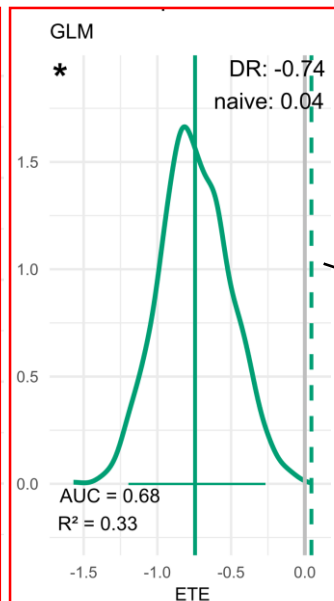
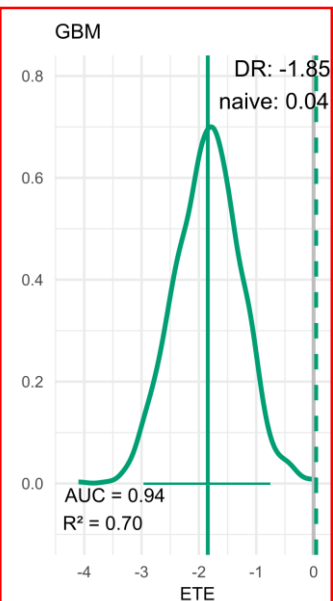


Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?



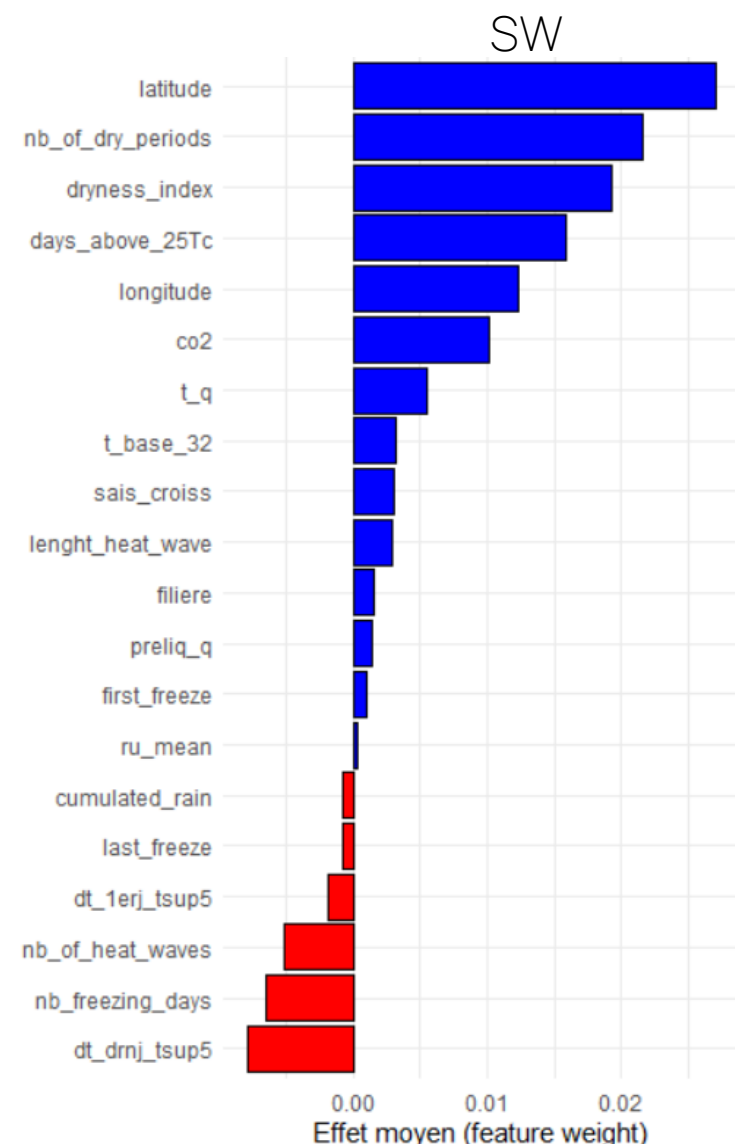
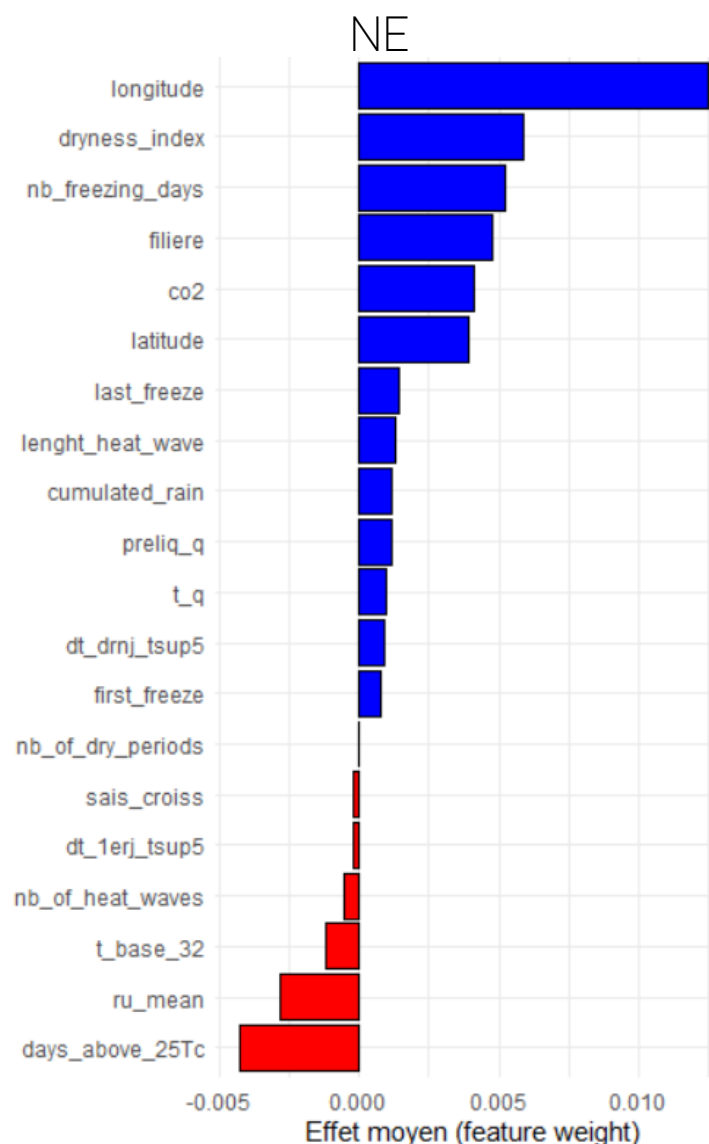
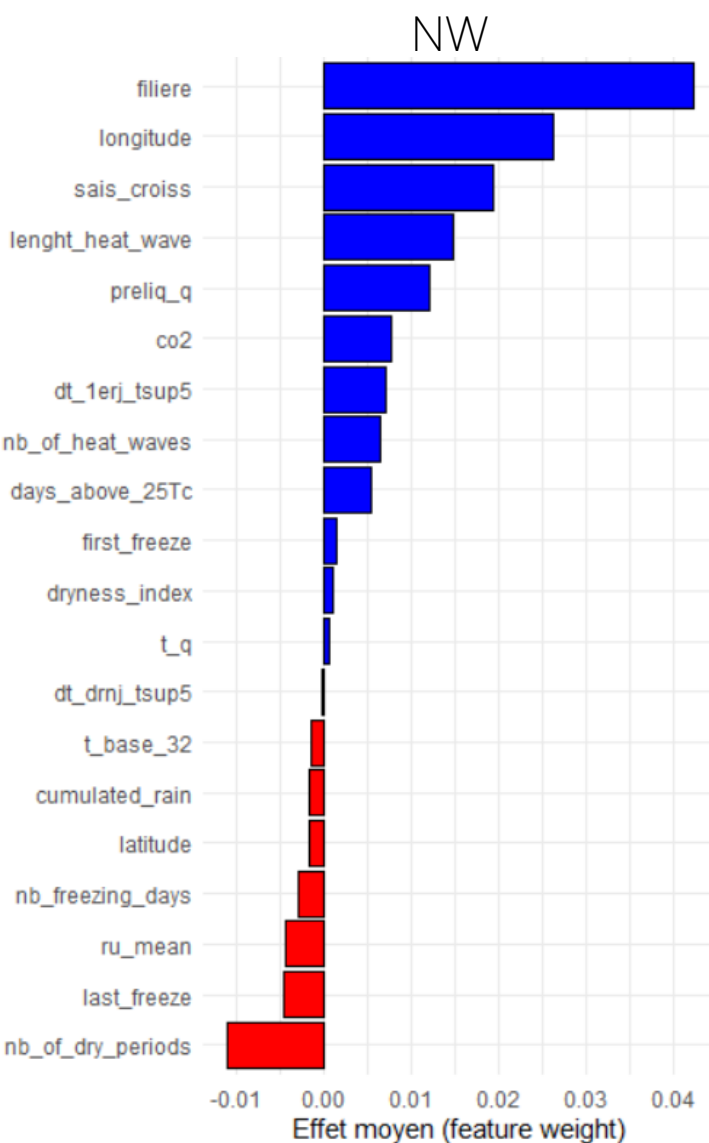
Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

IFT



Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

Quels éléments de contexte sont associés à la présence de LAG au sein des quarts de France?



Merci de votre attention

romain.escassut-nandillon@inrae.fr



© Diapositives : Slidego
© Images : Unsplash & Pixabay



Agroécologie
Dijon
Unité de Recherche



Insérez Les
Réussir avec les légumineuses



INRAE

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

3. Implémentation Doubly Robust

```
if (all(is.na(c(Pred_1, Pred_0)))) {  
  
  dr_gbm <- c(dr_gbm, NA)  
  message("DR: NA (Pred_1/Pred_0 NA)")  
  
} else {  
  
  DR1 <- mean(Pred_1 + (TAB$X / TAB$Score) * (TAB$Y - Pred_1), na.rm = TRUE)  
  
  DR0 <- mean(Pred_0 + ((1 - TAB$X) / (1 - TAB$Score)) * (TAB$Y - Pred_0), na.rm = TRUE)  
  
  dr_gbm <- c(dr_gbm, DR1 - DR0)  
  
}
```

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

3. Implémentation Doubly Robust

```
if (all(is.na(c(Pred_1, Pred_0)))) {  
  
  dr_gbm <- c(dr_gbm, NA)  
  message("DR: NA (Pred_1/Pred_0 NA)")  
  
} else {  
  
  DR1 <- mean(Pred_1 + (TAB$X / TAB$Score) * (TAB$Y - Pred_1), na.rm = TRUE) ← Estimation sous traitement  
  
  DR0 <- mean(Pred_0 + ((1 - TAB$X) / (1 - TAB$Score)) * (TAB$Y - Pred_0), na.rm = TRUE)  
  
  dr_gbm <- c(dr_gbm, DR1 - DR0)  
  
}
```

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

3. Implémentation Doubly Robust

```
if (all(is.na(c(Pred_1, Pred_0)))) {
```

```
  dr_gbm <- c(dr_gbm, NA)  
  message("DR: NA (Pred_1/Pred_0 NA)")
```

```
} else {
```

```
  DR1 <- mean(Pred_1 + (TAB$X / TAB$Score) * (TAB$Y - Pred_1), na.rm = TRUE)
```

Prédiction outcome

Pondération IPW

Résidu outcome

Estimation sous traitement

```
  DR0 <- mean(Pred_0 + ((1 - TAB$X) / (1 - TAB$Score)) * (TAB$Y - Pred_0), na.rm = TRUE)
```

```
  dr_gbm <- c(dr_gbm, DR1 - DR0)
```

```
}
```

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

3. Implémentation Doubly Robust

```
if (all(is.na(c(Pred_1, Pred_0)))) {  
  
  dr_gbm <- c(dr_gbm, NA)  
  message("DR: NA (Pred_1/Pred_0 NA)")  
  
} else {  
  
  DR1 <- mean(Pred_1 + (TAB$X / TAB$Score) * (TAB$Y - Pred_1), na.rm = TRUE)  
  
  DR0 <- mean(Pred_0 + ((1 - TAB$X) / (1 - TAB$Score)) * (TAB$Y - Pred_0), na.rm = TRUE)  
  
  dr_gbm <- c(dr_gbm, DR1 - DR0)  
  
}
```

Estimation sous non-traitement



Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

3. Implémentation Doubly Robust

```
if (all(is.na(c(Pred_1, Pred_0)))) {
```

```
  dr_gbm <- c(dr_gbm, NA)
  message("DR: NA (Pred_1/Pred_0 NA)")
```

```
} else {
```

```
  DR1 <- mean(Pred_1 + (TAB$X / TAB$Score) * (TAB$Y - Pred_1), na.rm = TRUE)
```

```
  DR0 <- mean(Pred_0 + ((1 - TAB$X) / (1 - TAB$Score)) * (TAB$Y - Pred_0), na.rm = TRUE)
```

```
  dr_gbm <- c(dr_gbm, DR1 - DR0)
```

```
}
```

Prédiction outcome

Pondération IPW

Résidu outcome

Estimation sous non-traitement

Y-a-t-il des différences de performances entre les systèmes avec ou sans légumineuses à graines (LAG) ?

3. Implémentation Doubly Robust

```
if (all(is.na(c(Pred_1, Pred_0)))) {  
  
  dr_gbm <- c(dr_gbm, NA)  
  message("DR: NA (Pred_1/Pred_0 NA)")  
  
} else {  
  
  DR1 <- mean(Pred_1 + (TAB$X / TAB$Score) * (TAB$Y - Pred_1), na.rm = TRUE)  
  
  DR0 <- mean(Pred_0 + ((1 - TAB$X) / (1 - TAB$Score)) * (TAB$Y - Pred_0), na.rm = TRUE)  
  
  dr_gbm <- c(dr_gbm, DR1 - DR0) ← Effet estimé du traitement (ETE)  
  
}
```