Using risk analysis to compare the accuracy of enteric methane emissions models of dairy cattle

P. Blondiaux[1], E.H. Cabezas-Garcia[2], T. Senga Kiessé[3], R. Muñoz-Tamayo[1], K.F. Reed[2] and M. Eugène[4]

[1] INRAE, Université Paris-Saclay, AgroParisTech, UMR Modélisation Systémique Appliquée aux Ruminants, 91120, Palaiseau, France

[2]Department of Animal Science, Cornell University, Ithaca, NY, United States

[3] INRAE, UMR SAS, Institut Agro, 35000 Rennes, France

[4] INRAE, UCA, VetAgro Sup, UMR Herbivores 1213, 63122 Saint-Genès-Champanelle, France

EAAP 2024, September 1st 2024, Firenze, Italy

Context

Enteric CH_4 emissions produced from the **rumen fermentation** contribute the most to the greenhouse gases (GHG) emitted from ruminants

➡ 66% of total GHG emissions from agricultural sector in 2019 (CITEPA, 2021)

Context

Enteric CH_4 emissions produced from the **rumen fermentation** contribute the most to the greenhouse gases (GHG) emitted from ruminants

- 66% of total GHG emissions from agricultural sector in 2019 (CITEPA, 2021)

Empirical models are used to:

- develop the national inventories of greenhouse gases

- assess the efficiency of dietary mitigation strategies

Context

Enteric CH_4 emissions produced from the **rumen fermentation** contribute the most to the greenhouse gases (GHG) emitted from ruminants

→ 66% of total GHG emissions from agricultural sector in 2019 (CITEPA, 2021)

Empirical models are used to:

→ develop the national inventories of greenhouse gases

→ assess the efficiency of dietary mitigation strategies

Empirical models are usually assessed and compared by computing several evaluation criteria quantifying the prediction error

Gaps → **The uncertainty is not considered when comparing models**

Aim and method

Aim: Quantifying and comparing the uncertainty of 7 empirical models of enteric CH_4 production

Assumption → **Integrating the uncertainty can provide a useful information to compare and recommend models**

Aim and method

Aim: Quantifying and comparing the uncertainty of 7 empirical models of enteric CH_4 production

Assumption → **Integrating the uncertainty can provide a useful information to compare and recommend models**

2 analyses were conducted

→ **Q1**: What is the probability of models to predict with an accuracy of x% around the CH_4 observed?

→ **Q2**: What is the probability of models to over or under predict a diet scenario?

Bayesian linear regression models

Bayes formula

$$P(\theta|X) = \frac{P(X|\theta).P(\theta)}{P(X)}$$

Posterior distribution
update the initial knowledge with the data

Prior distribution
initial knowledge on X

Bayesian linear regression models

Bayes formula

Posterior distribution
update the initial knowledge with the data

$$P(\theta|X) = \frac{P(X|\theta).P(\theta)}{P(X)}$$

Prior distribution
initial knowledge on X

Bayesian linear regression: Transition of the linear regression in the **probabilistic framework**

$$Y \sim N(\mu = \beta_0 + \beta_1.X, \sigma = \varepsilon)$$

Bayesian linear regression models

Bayes formula

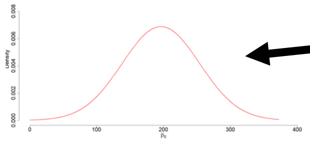
Posterior distribution
update the initial knowledge with the data

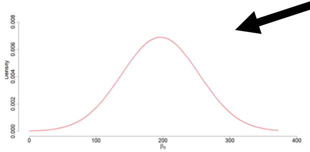
$$P(\theta|X) = \frac{P(X|\theta).P(\theta)}{P(X)}$$

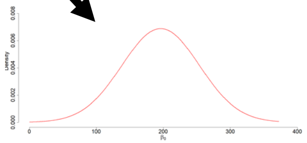
Prior distribution
initial knowledge on X

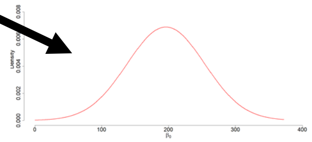
Bayesian linear regression: Transition of the linear regression in the **probabilistic framework**

$$Y \sim N(\mu = \beta_0 + \beta_1 . X, \sigma = \varepsilon)$$









Bayesian linear regression models

Bayes formula

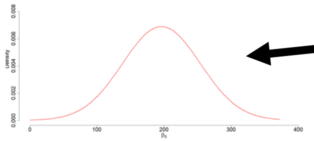
Posterior distribution
update the initial knowledge with the data

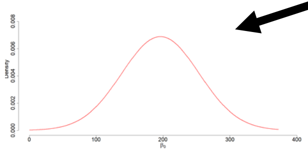
$$P(\theta|X) = \frac{P(X|\theta).P(\theta)}{P(X)}$$

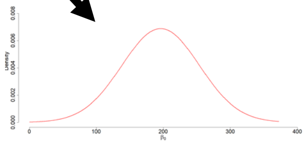
Prior distribution
initial knowledge on X

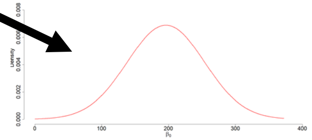
Bayesian linear regression: Transition of the linear regression in the **probabilistic framework**

$$Y \sim N(\mu = \beta_0 + \beta_1 . X, \sigma = \varepsilon)$$

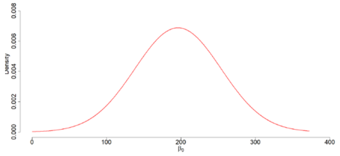








DMI = 10 kg/d → $\widehat{CH_4}$ ~



Uncertainty quantification

Bayesian linear regression models

Bayes formula

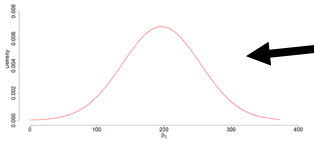
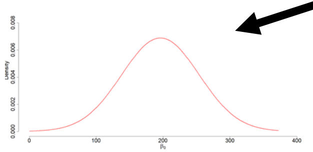
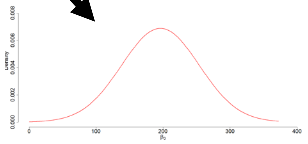
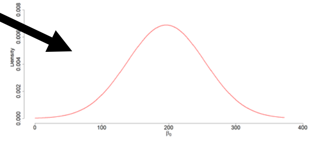
Posterior distribution
update the initial knowledge with the data

$$P(\theta|X) = \frac{P(X|\theta).P(\theta)}{P(X)}$$

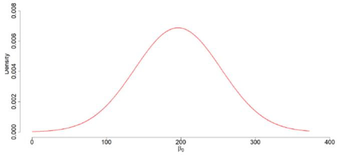
Prior distribution
initial knowledge on X

Bayesian linear regression: Transition of the linear regression in the **probabilistic framework**

$$Y \sim N(\mu = \beta_0 + \beta_1.X, \sigma = \varepsilon)$$

DMI = 10 kg/d ⟶ $\widehat{CH_4} \sim$



Uncertainty quantification

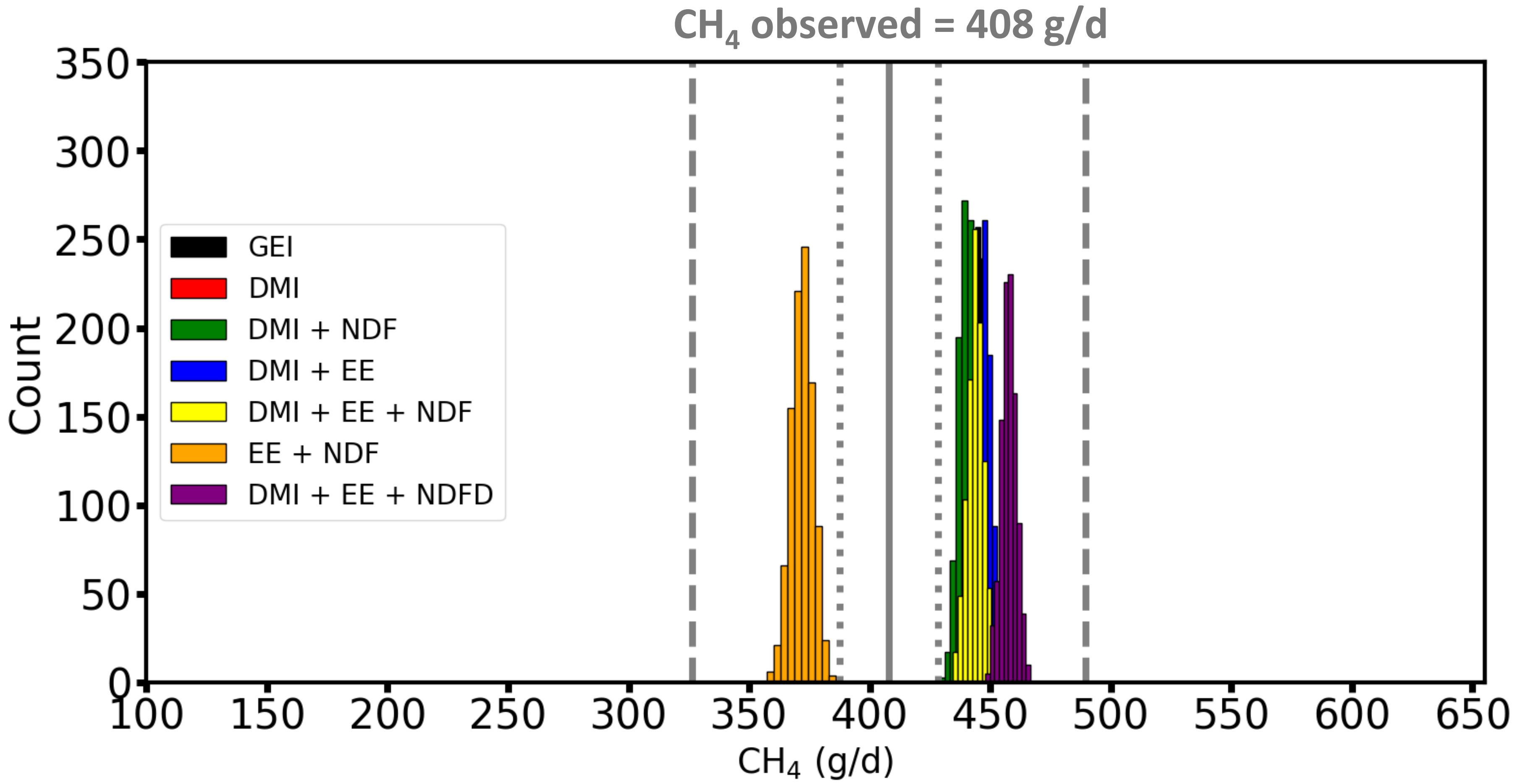
Database of enteric CH_4 emissions (g/d) of dairy cattle (de Ondarza et al., 2023)

Database of empirical models of enteric CH_4 production

Development of 7 Bayesian linear regression models

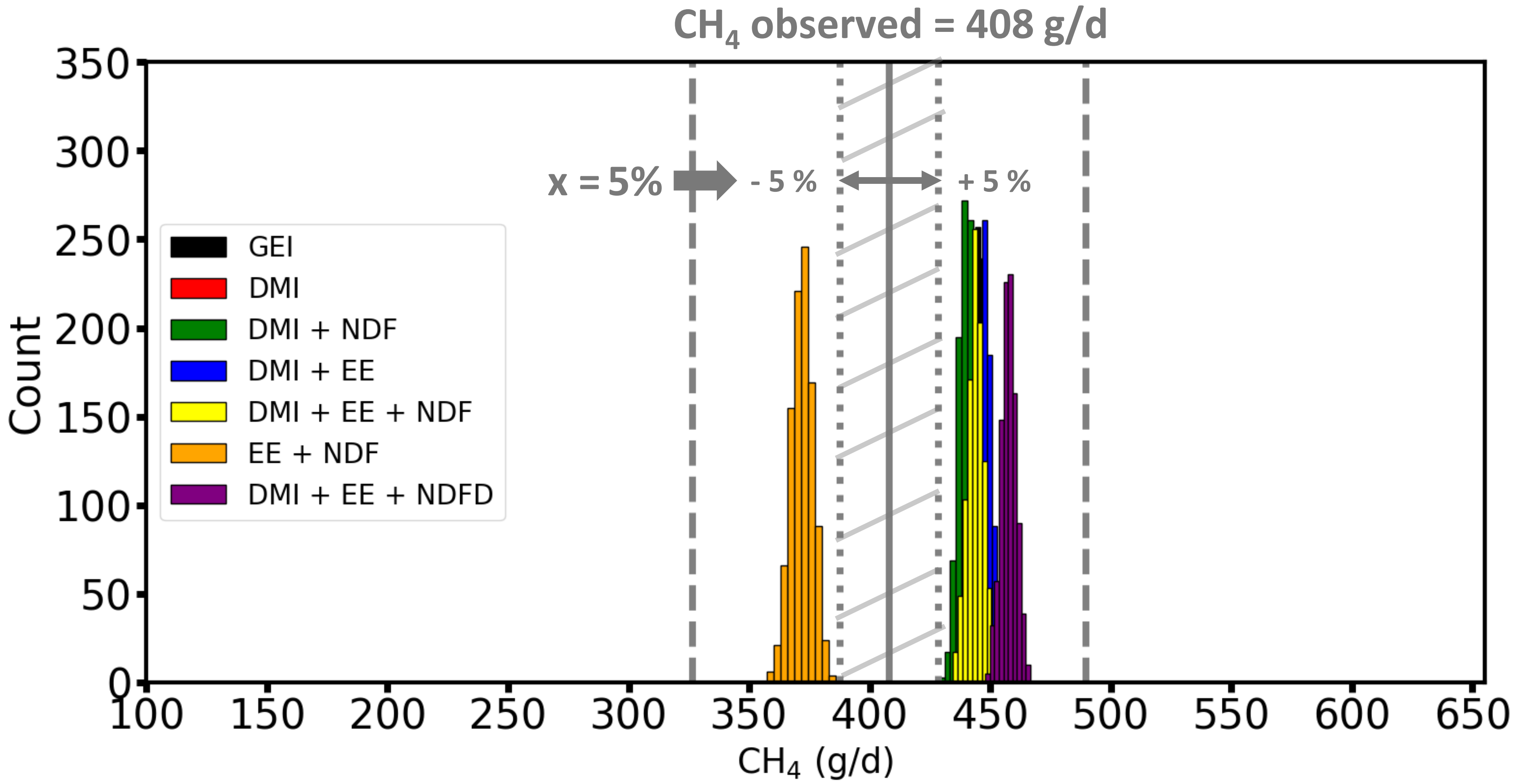
Q1: Assessment of model accuracy - Approach

What is the probability of models to predict with an accuracy of x% around the CH_4 observed? **P(CH_4 predicted ∈ [CH_4 observed ± x %])** with x ranging from 1 to 20%



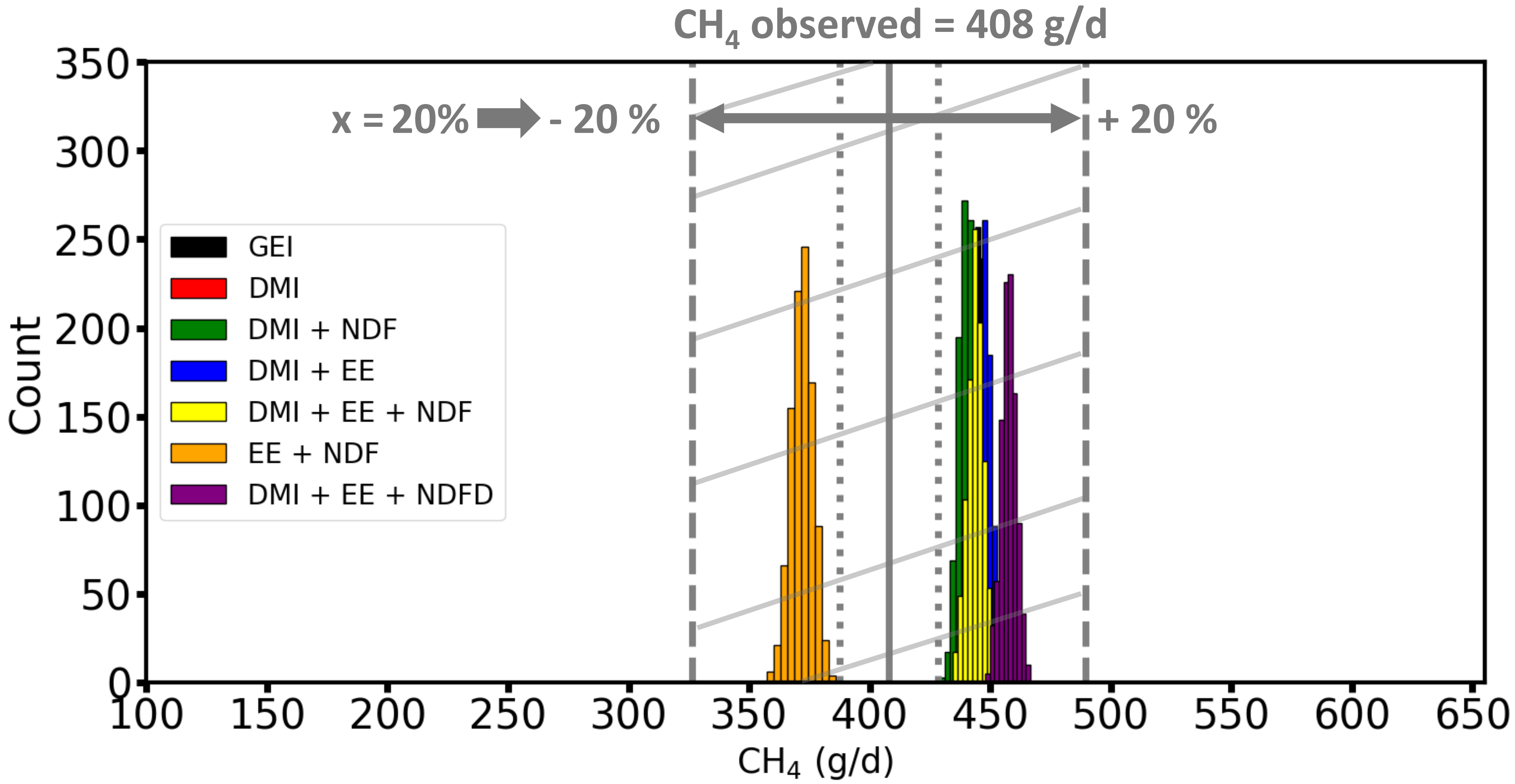
Q1: Assessment of model accuracy - Approach

What is the probability of models to predict with an accuracy of x% around the CH_4 observed? P(CH_4 predicted ∈ [CH_4 observed ± x %]) with x ranging from 1 to 20%



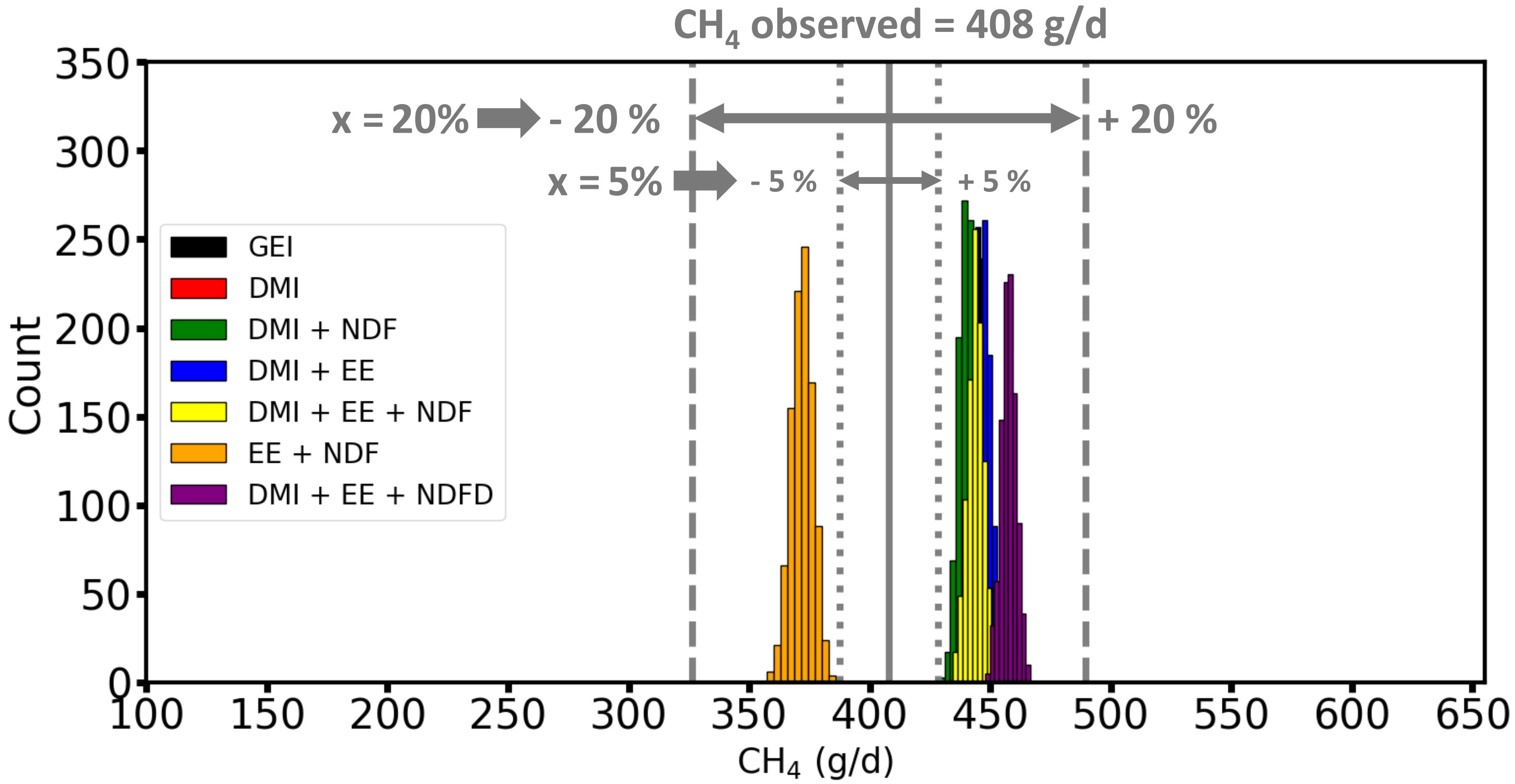
Q1: Assessment of model accuracy - Approach

What is the probability of models to predict with an accuracy of x% around the CH_4 observed? **P(CH_4 predicted ∈ [CH_4 observed ± x %])** with x ranging from 1 to 20%



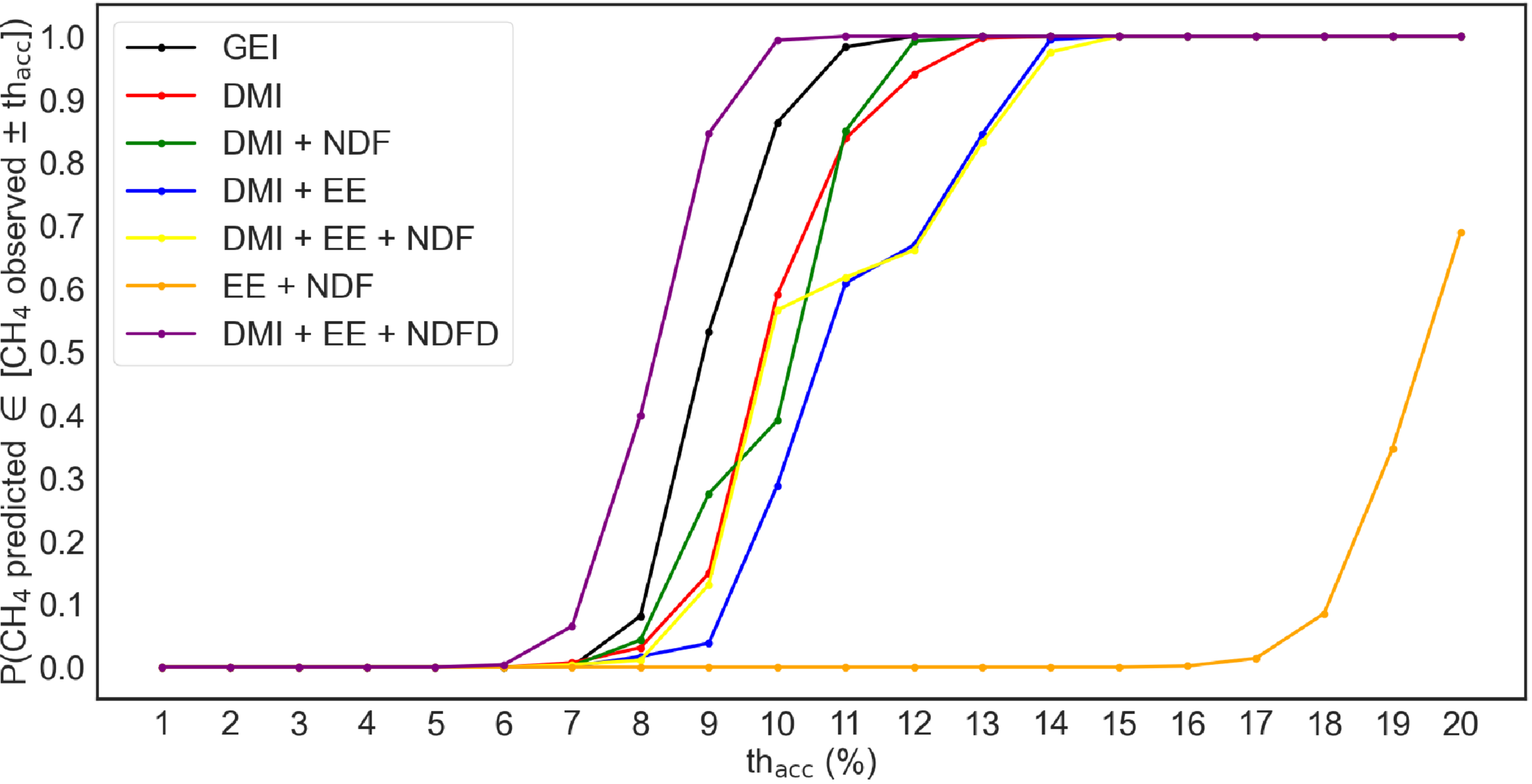
Q1: Assessment of model accuracy - Approach

What is the probability of models to predict with an accuracy of x% around the CH_4 observed? **P(CH_4 predicted ∈ [CH_4 observed ± x %])** with x ranging from 1 to 20%



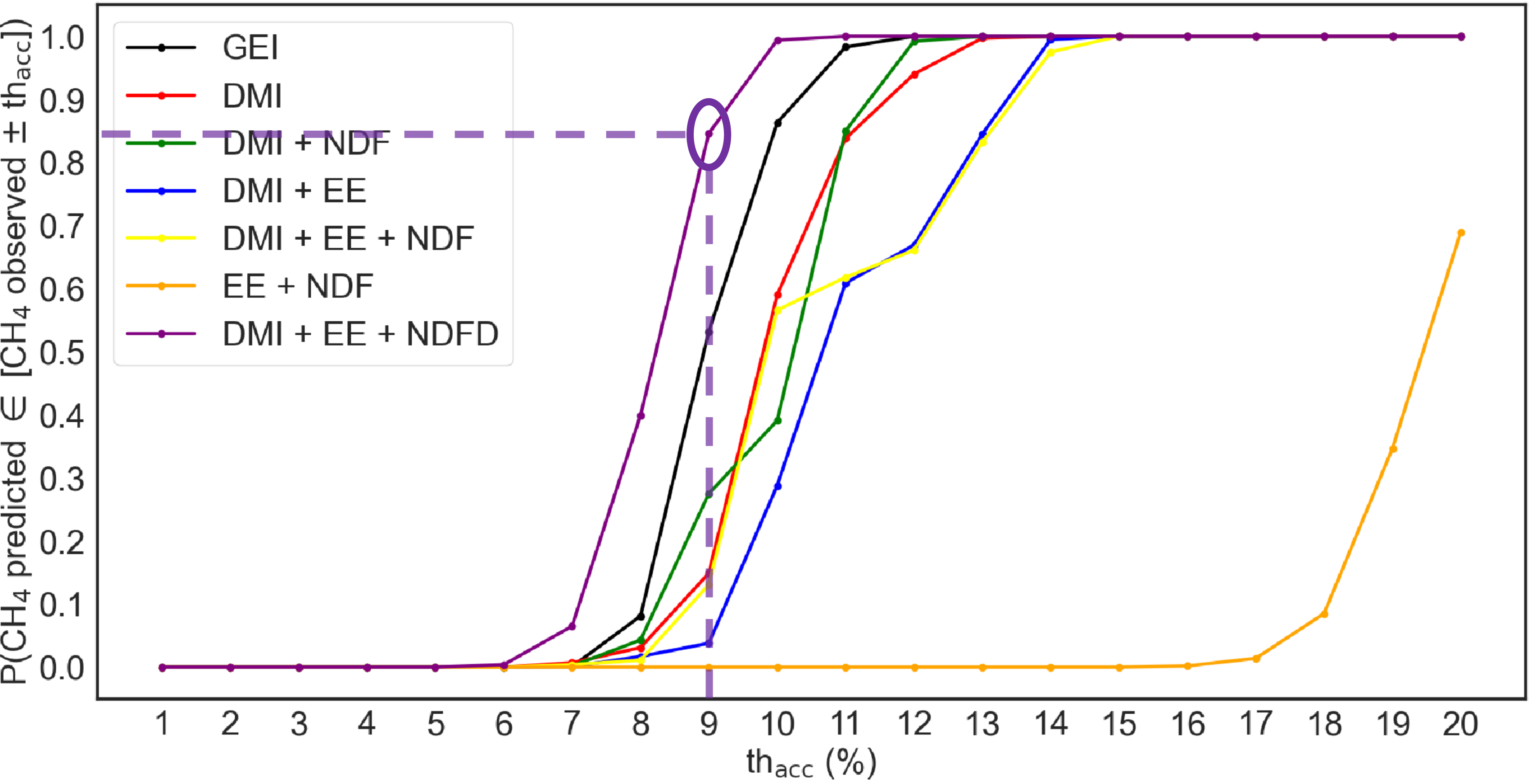
Q1: Assessment of model accuracy - Results

The probability was computed for the 43 test records and the median was displayed (n = 172 training records)



Q1: Assessment of model accuracy - Results

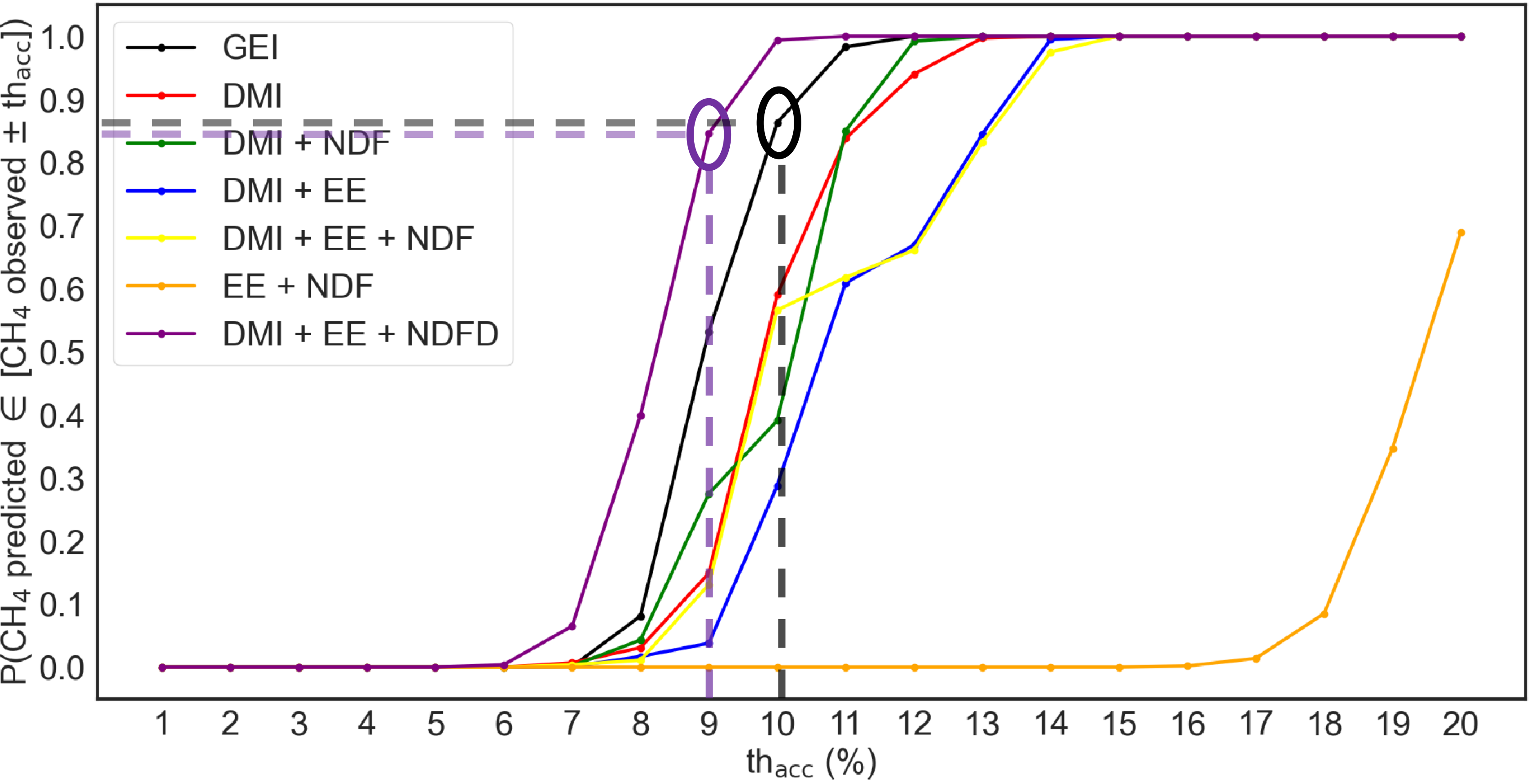
The probability was computed for the 43 test records and the median was displayed (n = 172 training records)



- Model **DMI + EE + NDFD** showed the best accuracy with a high probability to predict within a target area at 9% around the true value (P = 0.85)

Q1: Assessment of model accuracy - Results

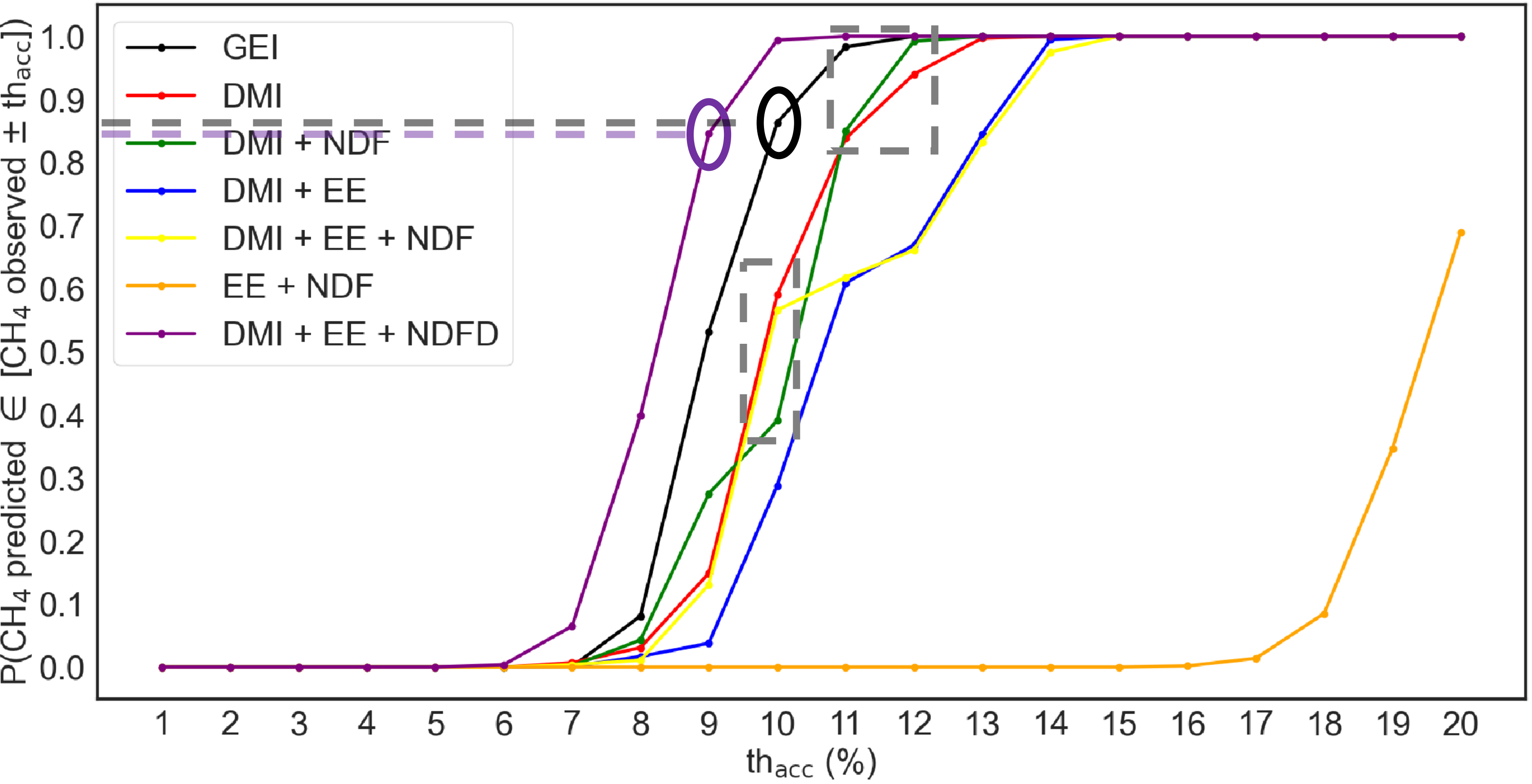
The probability was computed for the 43 test records and the median was displayed (n = 172 training records)



- Model **DMI + EE + NDFD** showed the best accuracy with a high probability to predict within a target area at 9% around the true value (P = 0.85)
- Model **GEI** showed the second best performance with a probability of 0.87 to predict within a target area at 10% around the true value

Q1: Assessment of model accuracy - Results

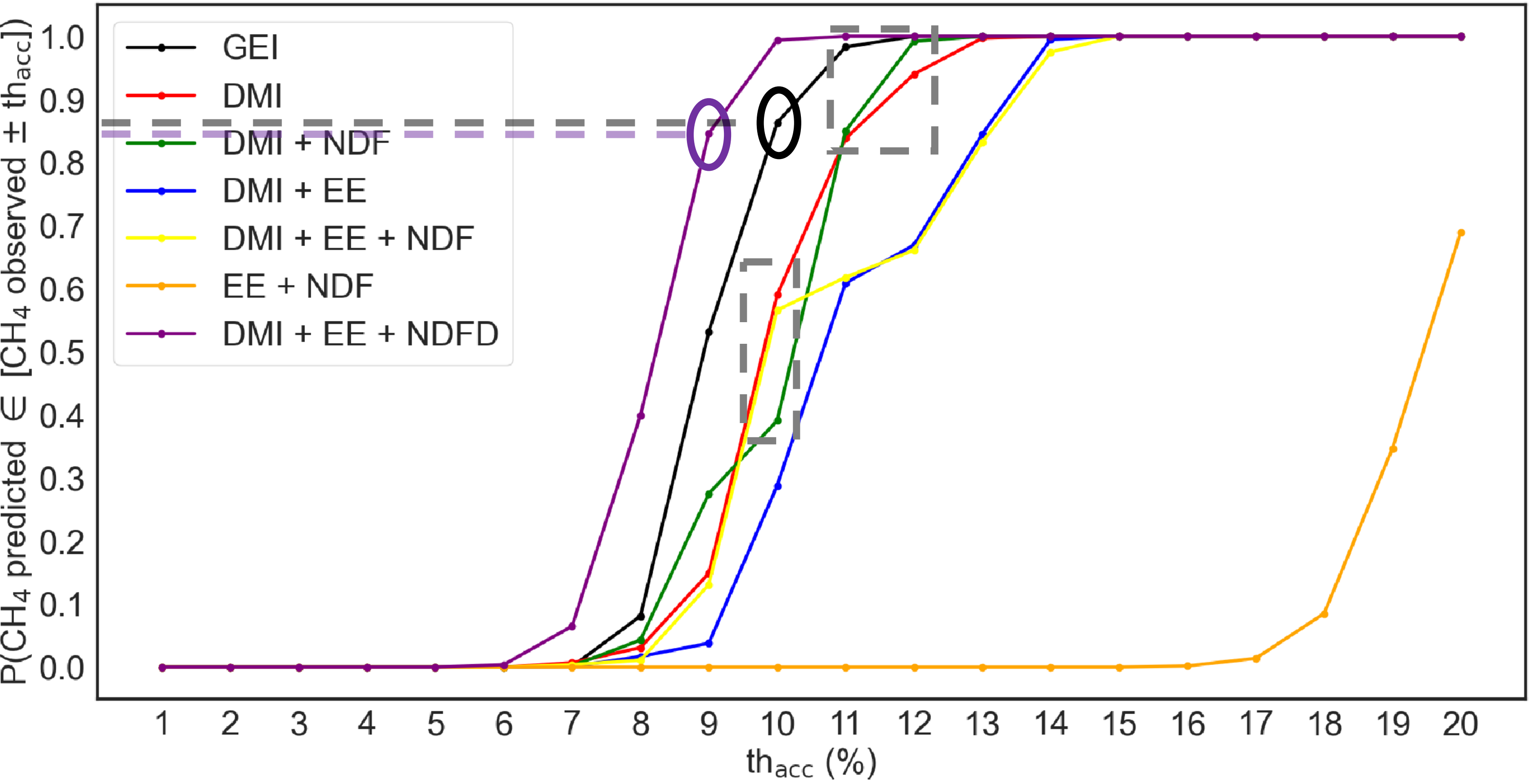
The probability was computed for the 43 test records and the median was displayed (n = 172 training records)



- Model **DMI + EE + NDFD** showed the best accuracy with a high probability to predict within a target area at 9% around the true value (P = 0.85)
- Model **GEI** showed the second best performance with a probability of 0.87 to predict within a target area at 10% around the true value
- Models using DMI alone or with diet composition variables showed a rather similar accuracy up to x = 10%, then models **DMI** and **DMI + NDF** were better

Q1: Assessment of model accuracy - Results

The probability was computed for the 43 test records and the median was displayed (n = 172 training records)

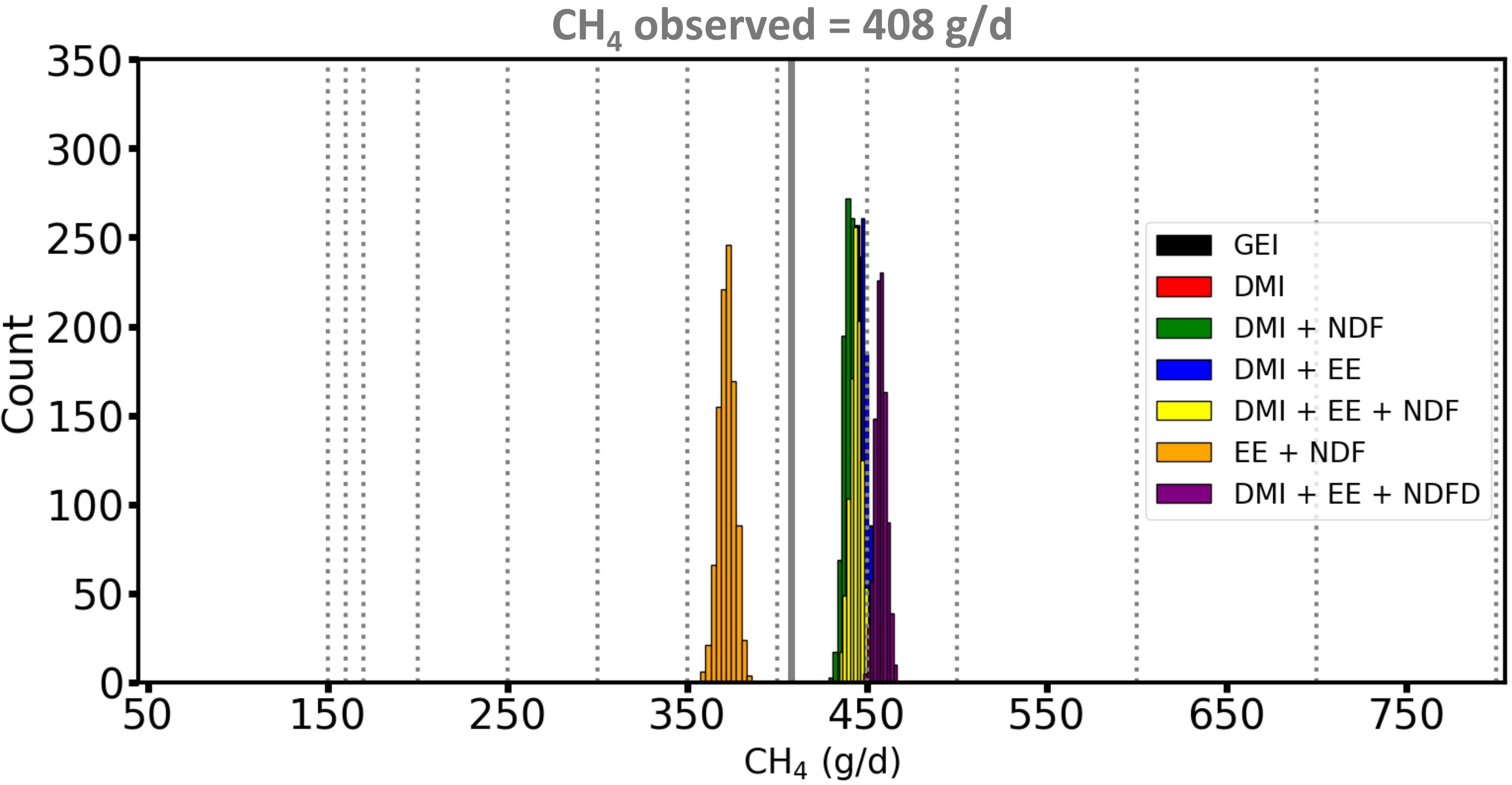


- Model **DMI + EE + NDFD** showed the best accuracy with a high probability to predict within a target area at 9% around the true value (P = 0.85)
- Model **GEI** showed the second best performance with a probability of 0.87 to predict within a target area at 10% around the true value
- Models using DMI alone or with diet composition variables showed a rather similar accuracy up to x = 10%, then models **DMI** and **DMI + NDF** were better
- Model **EE + NDF** provided largely the lowest accuracy

Q2: Assessment of a dietary scenario - Approach

What is the probability of models to predict below a CH_4 (g/d) threshold value?

$P(CH_4$ predicted $\leq CH_4$ threshold | dietary scenario$)$



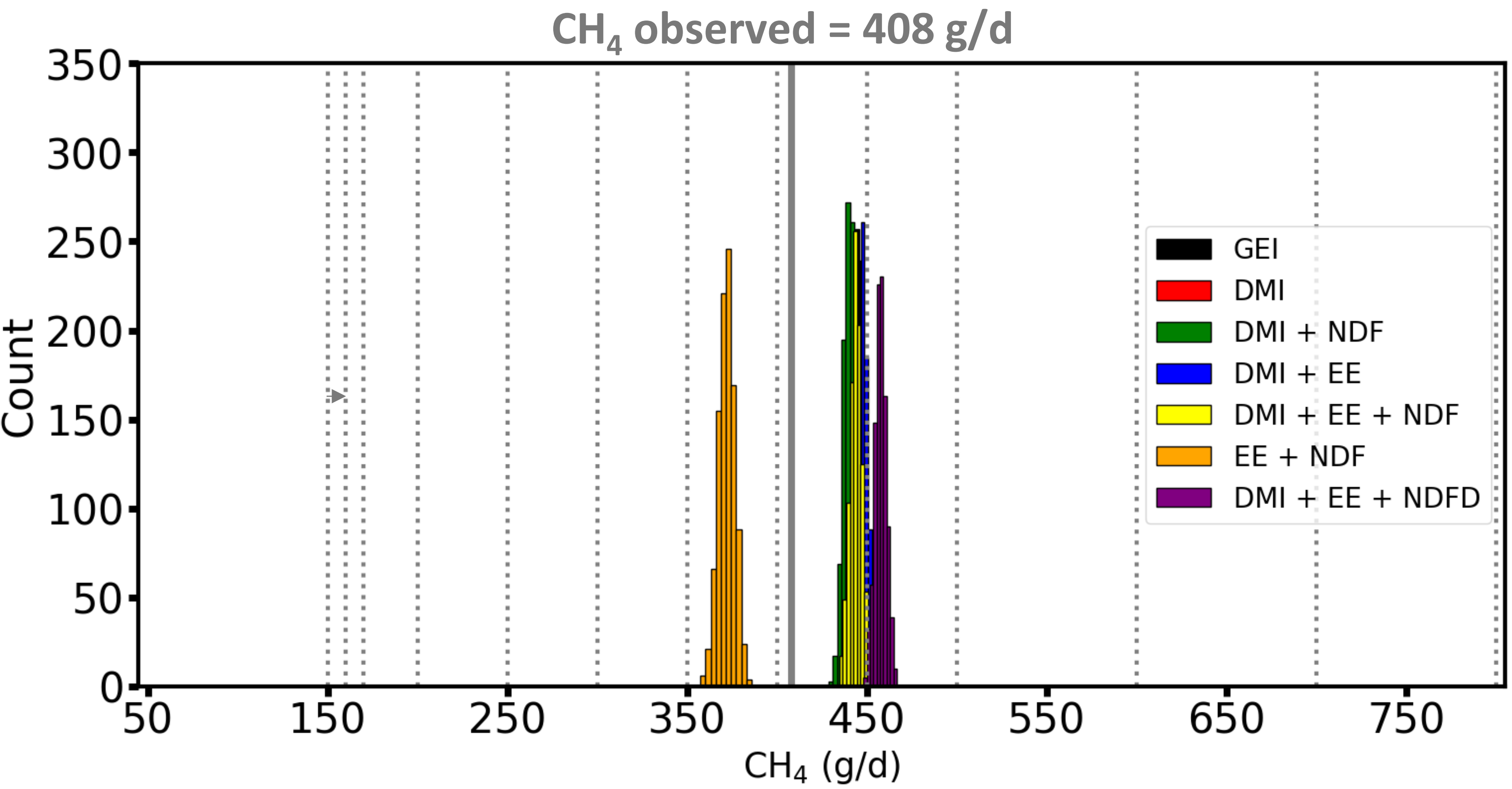
Dietary scenario based on the level of NDF (%DM)

	CH_4 (g/d)	DMI (kg/d)	NDF (%DM)
Low	290	16.7	25.6
Medium	326	16.8	33.7
High	77.6	4.33	73.1

Q2: Assessment of a dietary scenario - Approach

What is the probability of models to predict below a CH_4 (g/d) threshold value?

P(CH_4 predicted ≤ CH_4 threshold | dietary scenario)



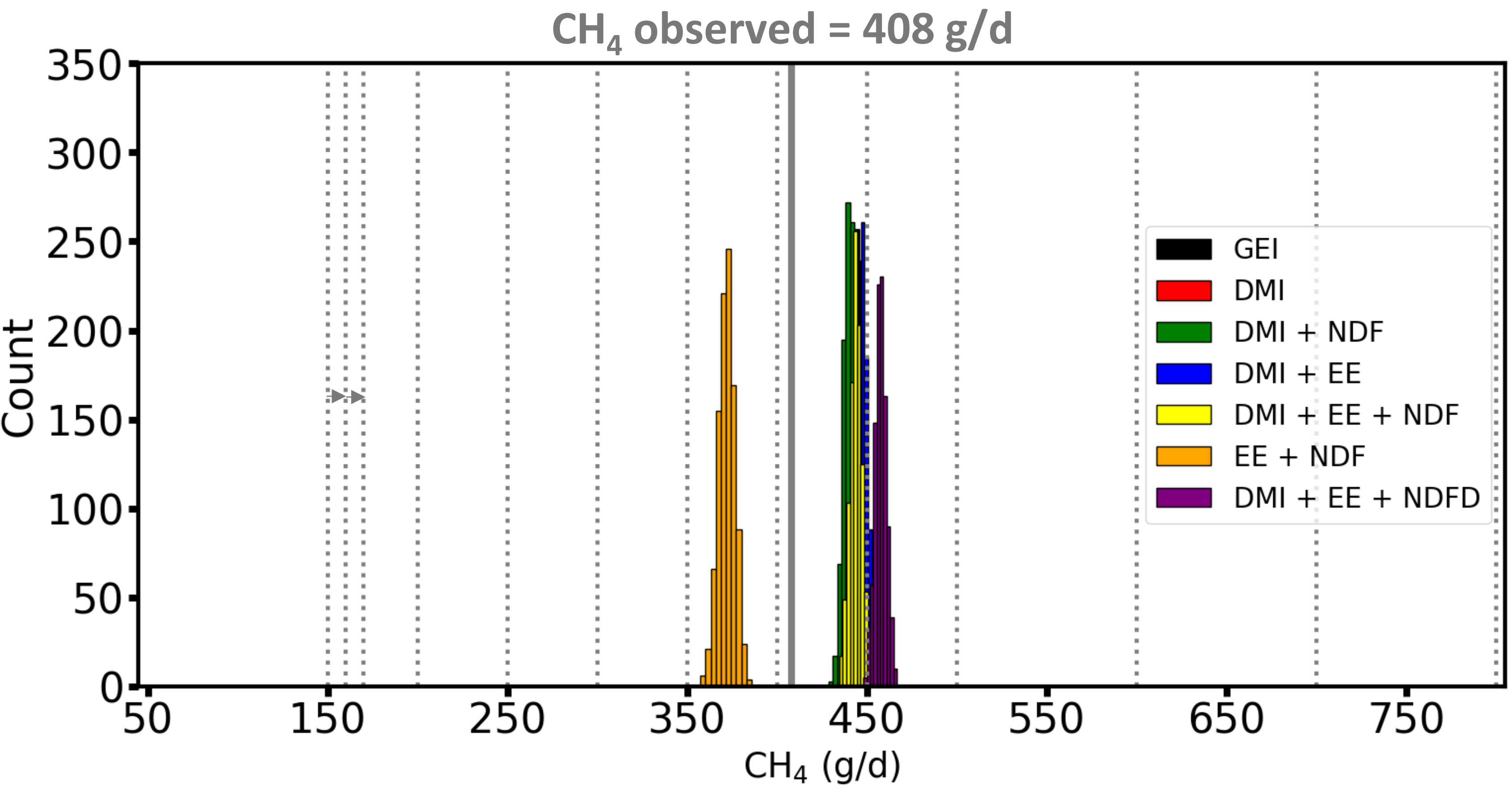
Dietary scenario based on the level of NDF (%DM)

	CH_4 (g/d)	DMI (kg/d)	NDF (%DM)
Low	290	16.7	25.6
Medium	326	16.8	33.7
High	77.6	4.33	73.1

Q2: Assessment of a dietary scenario - Approach

What is the probability of models to predict below a CH_4 (g/d) threshold value?

$P(CH_4 \text{ predicted} \leq CH_4 \text{ threshold} | \text{ dietary scenario})$



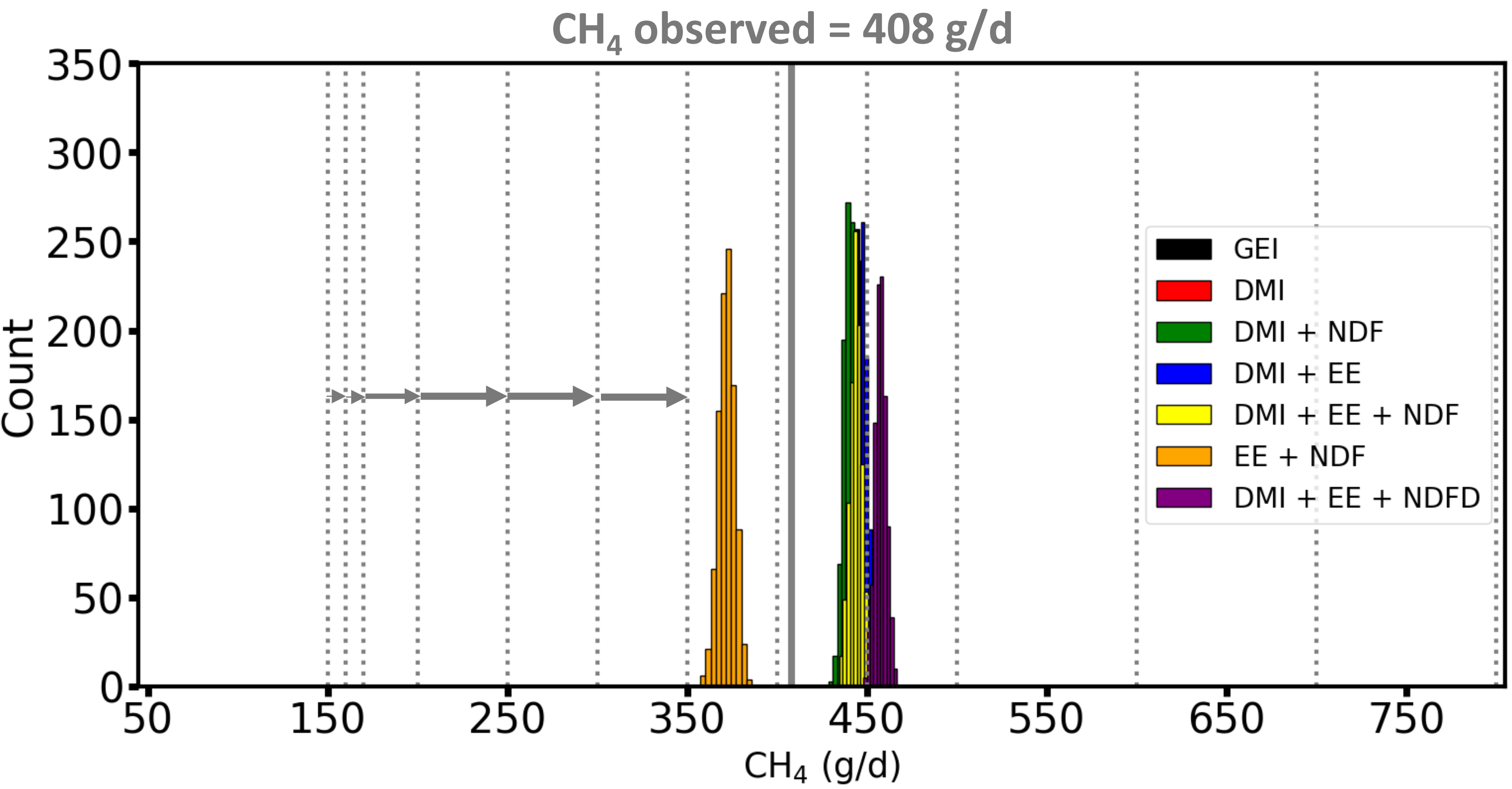
Dietary scenario based on the level of NDF (%DM)

	CH_4 (g/d)	DMI (kg/d)	NDF (%DM)
Low	290	16.7	25.6
Medium	326	16.8	33.7
High	77.6	4.33	73.1

Q2: Assessment of a dietary scenario - Approach

What is the probability of models to predict below a CH_4 (g/d) threshold value?

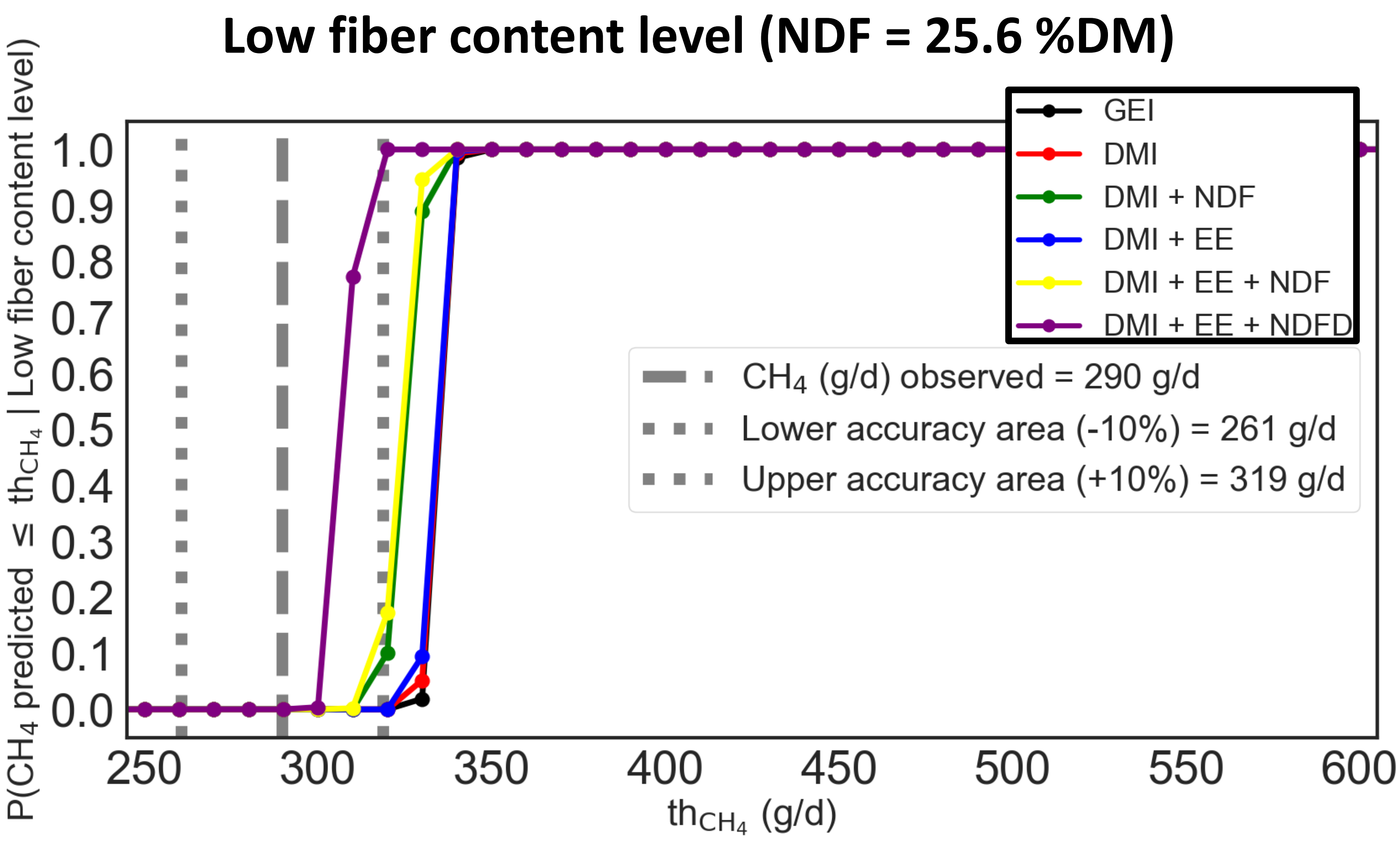
$P(CH_4 \text{ predicted} \leq CH_4 \text{ threshold} \mid \text{dietary scenario})$



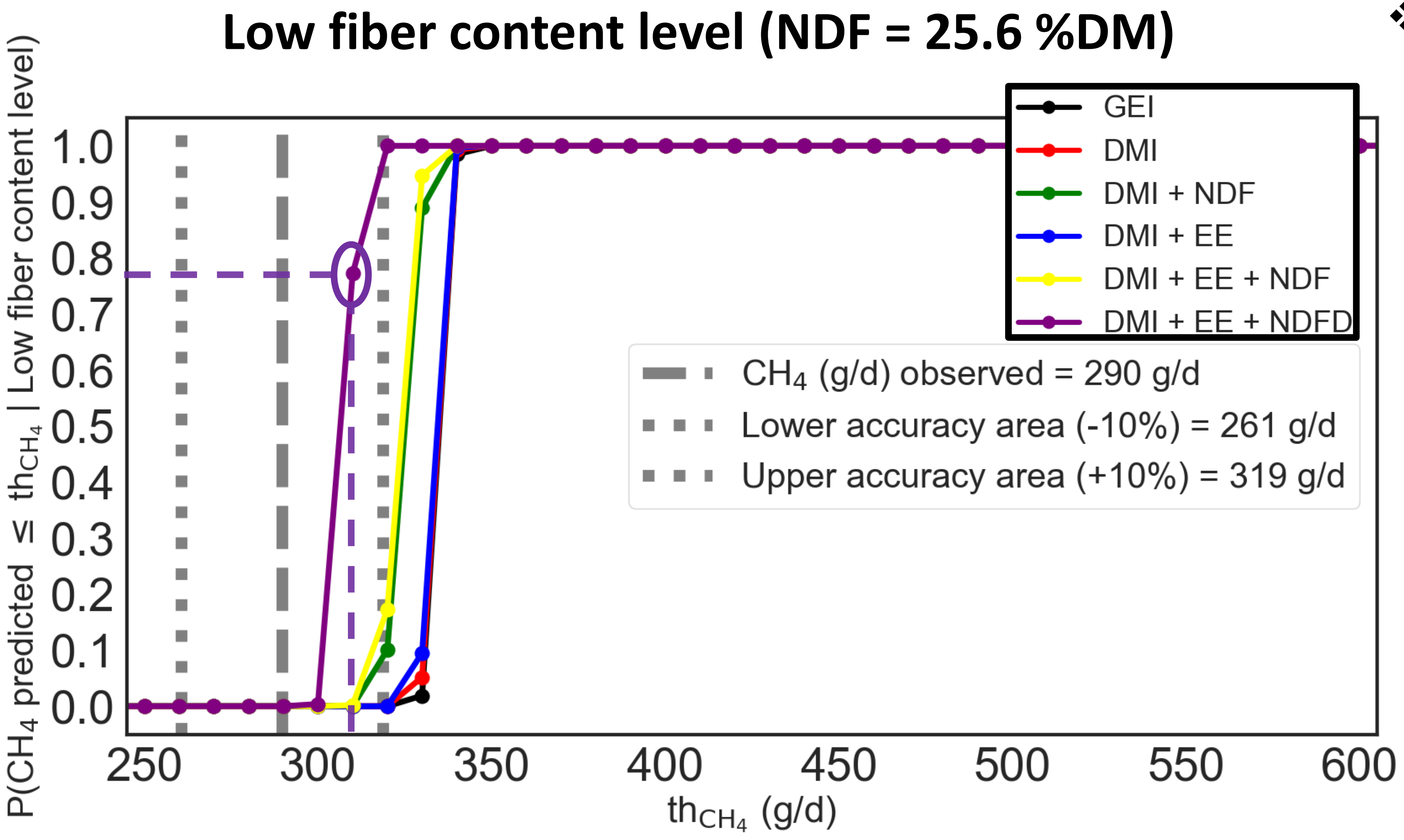
Dietary scenario based on the level of NDF (%DM)

	CH_4 (g/d)	DMI (kg/d)	NDF (%DM)
Low	290	16.7	25.6
Medium	326	16.8	33.7
High	77.6	4.33	73.1

Q2: Assessment of a dietary scenario - Results

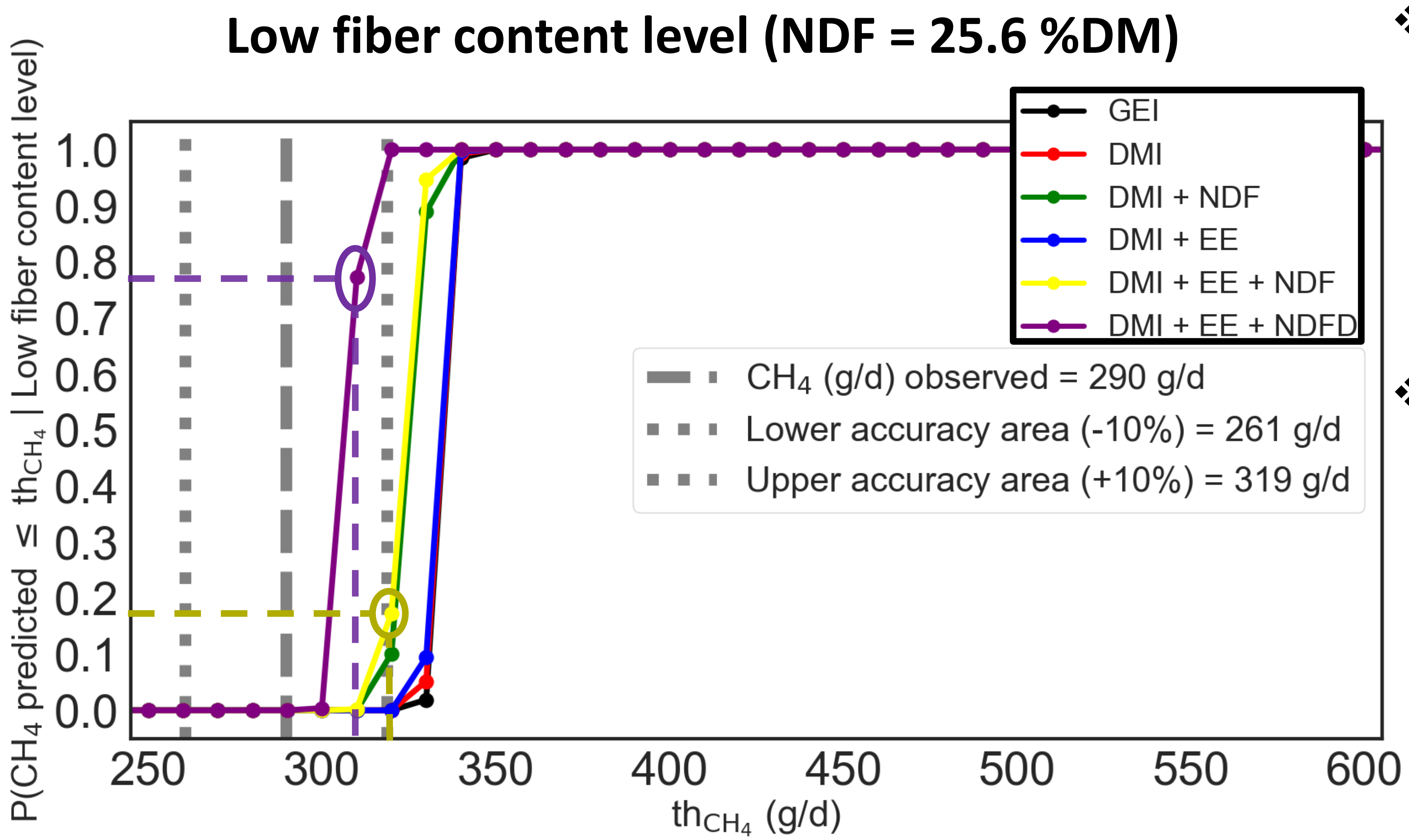


Q2: Assessment of a dietary scenario - Results



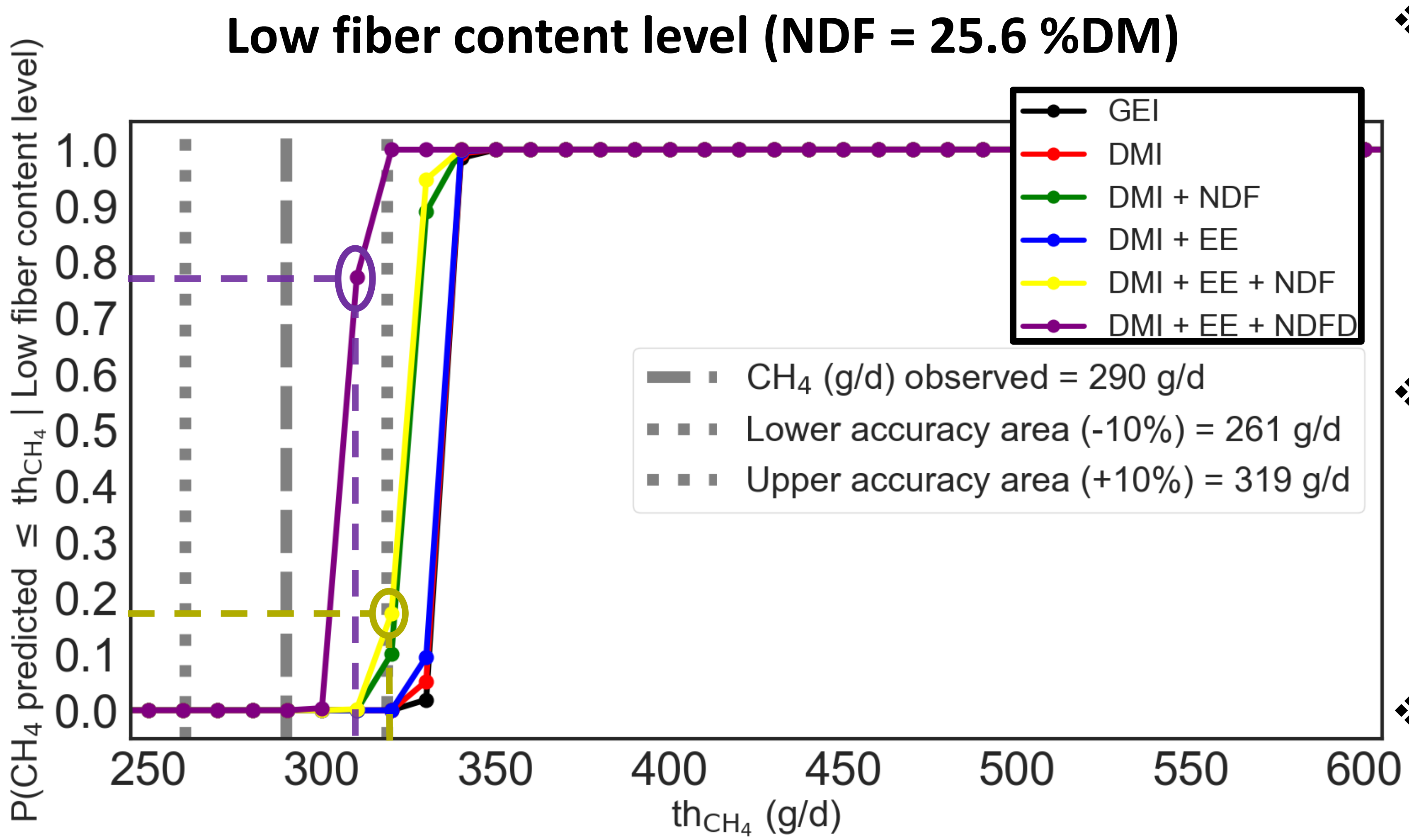
- Model **DMI + EE + NDFD** showed the best performance with a high probability (P = 0.79) to predict values below 310 g/d (within the target area)

Q2: Assessment of a dietary scenario - Results



- Model **DMI + EE + NDFD** showed the best performance with a high probability ($P = 0.79$) to predict values below 310 g/d (within the target area)
- Other models largely overestimated the true value with a low probability to predict within a target area at 10% around the true value ($P < 0.19$)

Q2: Assessment of a dietary scenario - Results



- Model **DMI + EE + NDFD** showed the best performance with a high probability (P = 0.79) to predict values below 310 g/d (within the target area)
- Other models largely overestimated the true value with a low probability to predict within a target area at 10% around the true value ($P < 0.19$)
- Models not using NDF-related variables (**DMI + EE**, **DMI** and **GEI**) were associated with the poorest performances (overestimation > 40 g/d compared with the true value)

Take home messages and perspectives

➡ Incorporating uncertainty into model comparison provides a useful information

Take home messages and perspectives

- Incorporating uncertainty into model comparison provides a useful information
- Model DMI + EE + NDFD was the most accurate (high probability of predicting within ±9% of observed CH_4); best ability to predict the low fiber content dietary scenario
- Model using GEI alone was associated with a better accuracy (high probability of predicting within ±10% of observed CH_4) than models with DMI and diet composition variables

Take home messages and perspectives

- Incorporating uncertainty into model comparison provides a useful information
- Model DMI + EE + NDFD was the most accurate (high probability of predicting within ±9% of observed CH_4); best ability to predict the low fiber content dietary scenario
- Model using GEI alone was associated with a better accuracy (high probability of predicting within ±10% of observed CH_4) than models with DMI and diet composition variables
- Models using NDF-related variables showed the best ability to predict the low fiber content dietary scenario

Take home messages and perspectives

- Incorporating uncertainty into model comparison provides a useful information
- Model DMI + EE + NDFD was the most accurate (high probability of predicting within ±9% of observed CH_4); best ability to predict the low fiber content dietary scenario
- Model using GEI alone was associated with a better accuracy (high probability of predicting within ±10% of observed CH_4) than models with DMI and diet composition variables
- Models using NDF-related variables showed the best ability to predict the low fiber content dietary scenario

Next steps

1. The framework developed in our study can be used to provide recommendations on which models to use in specific situations (region, CH_4 measurement method, treatment…)

2. Extend the approach of reaching CH_4 thresholds to reference dietary scenarios (INRAE, USDA, FAO)

Thank you for your attention!

paul.blondiaux@inrae.fr

This study was conducted as part of a scientific stay at the department of animal science of Cornell University. This stay was supported by INRAE, the doctoral school ABIES and the graduate school Biosphera of Paris-Saclay University