Predicting ammonia volatilization from field applied manure using Machine Learning

Armand Favrot - INRAE

David Makowski - INRAE

Sophie Génermont - INRAE

Céline Decuq - INRAE

Jean-Marc Gilliot - INRAE

March 26, 2024

Context

- Ammonia is an inorganic chemical compound of nitrogen and hydrogen : **NH_3**

Context

- Ammonia is an inorganic chemical compound of nitrogen and hydrogen : **NH3**
- It poses environmental issues
  - Euthropication, acidification
  - Atmospheric pollutant : precursor of fine particles (PM2.5, PM10)

Context

- Ammonia is an inorganic chemical compound of nitrogen and hydrogen : **NH3**
- It poses environmental issues
  - Euthropication, acidification
  - Atmospheric pollutant : precursor of fine particles (PM2.5, PM10)
- It originates from agriculture at 94%
  - field fertilization
  - animal husbandry

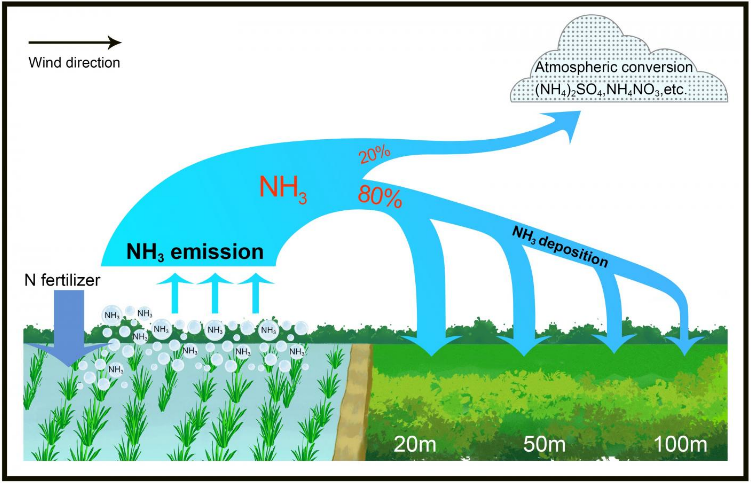
Context

- Ammonia is an inorganic chemical compound of nitrogen and hydrogen : **NH3**
- It poses environmental issues
  - Euthropication, acidification
  - Atmospheric pollutant : precursor of fine particles (PM2.5, PM10)
- It originates from agriculture at 94%
  - field fertilization
  - animal husbandry
- Ammonia emissions represent an economic loss for farmers

Context

- Ammonia is an inorganic chemical compound of nitrogen and hydrogen : **NH3**
- It poses environmental issues
  - Euthropication, acidification
  - Atmospheric pollutant : precursor of fine particles (PM2.5, PM10)
- It originates from agriculture at 94%
  - field fertilization
  - animal husbandry
- Ammonia emissions represent an economic loss for farmers
- **Objective** : build a model to predict ammonia emission from field-applied fertilizers

Descriptive scheme



Why machine learning?

- Very little used to predict ammonia volatilization

Why machine learning ?

- Very little used to predict ammonia volatilization
- Handles complex patterns, takes non-linearity into account

Why machine learning ?

- Very little used to predict ammonia volatilization
- Handles complex patterns, takes non-linearity into account
- Non parametric approach : no hypothesis on the form of the problem

Why machine learning ?

- Very little used to predict ammonia volatilization
- Handles complex patterns, takes non-linearity into account
- Non parametric approach : no hypothesis on the form of the problem
- High predictive accuracy

Why machine learning ?

- Very little used to predict ammonia volatilization
- Handles complex patterns, takes non-linearity into account
- Non parametric approach : no hypothesis on the form of the problem
- High predictive accuracy

Other approaches : Empirical / statistical models, mechanistic models, Semi-empirical models

Data description : ALFAM2

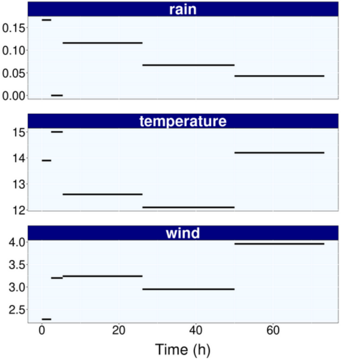
id	t.start	t.end	j.NH3	temp	wind	rain	man.source	app.mthd	dose	TAN	pH	man.dm	incorp	t.incorp
182	0.00	4.00	**1.79**	8.20	8.10	0.00	pig	bc	31.8	122.11	7.35	3.7	none	0
182	4.00	21.00	**0.07**	4.45	3.98	0.00	pig	bc	31.8	122.11	7.35	3.7	none	0
182	21.00	44.75	**0.15**	7.22	6.57	0.01	pig	bc	31.8	122.11	7.35	3.7	none	0
183	0.00	6.00	**0.58**	13.77	4.31	0.00	pig	bsth	21.6	58.32	7.71	2.8	none	0
183	6.00	20.50	**0.14**	9.88	3.38	0.00	pig	bsth	21.6	58.32	7.71	2.8	none	0
183	20.50	45.20	**0.00**	10.64	3.51	0.00	pig	bsth	21.6	58.32	7.71	2.8	none	0
184	0.00	5.15	**1.61**	11.15	8.27	0.00	pig	bc	24.9	78.68	7.61	3.4	none	0
184	5.15	20.15	**0.08**	8.43	6.23	0.00	pig	bc	24.9	78.68	7.61	3.4	none	0
184	20.15	45.15	**0.07**	10.00	7.30	0.00	pig	bc	24.9	78.68	7.61	3.4	none	0

- j.NH3, temp, wind, rain : **dynamic variables**
- man.source, app.mthd, dose, TAN, pH, man.dm, incorp, t.incorp : **fixed variables**

Data description : ALFAM2

Predictors

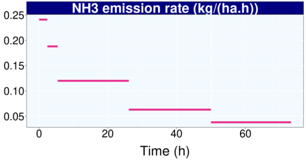
Dynamical variables



Fixed variables

- **application method**
  => broadcast
- **manure source**
  => pig
- **total ammonia nitrogen**
  => 72.7
- **application rate**
  => 40
- **incorporation**
  => shallow
- **dry matter**
  => 8.13
- **pH** = 6.75

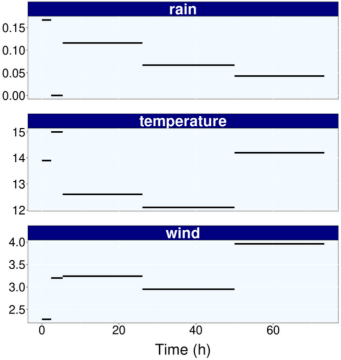
Response



Data description : ALFAM2

Predictors

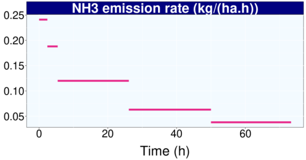
Dynamical variables



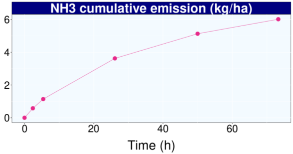
Fixed variables

- **application method**
  => broadcast
- **manure source**
  => pig
- **total ammonia nitrogen**
  => 72.7
- **application rate**
  => 40
- **incorporation**
  => shallow
- **dry matter**
  => 8.13
- **pH** = 6.75

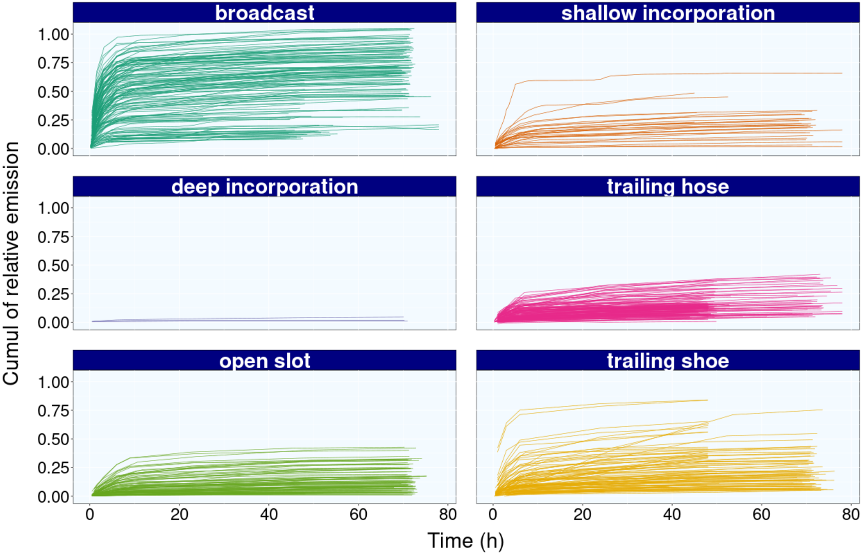
Response



<=>

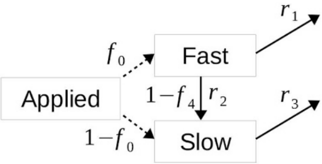


Data description : ALFAM2



A semi-empirical model : alfam2

Hafner et al - 2019 - *A flexible semi-empirical model for estimating ammonia volatilization from field-applied slurry*



$$F_0 = f_0 A$$
$$S_0 = (1 - f_0)A$$

$$F_i = F_{i-1} e^{-r_{1i}\Delta t_i}$$
$$S_i = S_{i-1} e^{-r_{3i}\Delta t_i}$$

f_0, r_1, r_2, r_3, f_4 : function of the data

$$E_i = A - F_i - S_i$$

$\Rightarrow$ Calibrated and evaluated using two different subsets of the **ALFAM2** database

A semi-empirical model

Table 3
Model performance statistics based on parameter values given in Table 2 applied to both data subsets.

Response	Pearson's r	ME[a]	MAE[b]	MBE[c]
Test dataset → Evaluation subset				
Flux (kg ha^{-1} h^{-1}) (n = 418 intervals)	0.832	0.663	0.433	−0.099
72 h cum. emission (kg ha^{-1}) (n = 48 plots)	0.861	0.620	5.63	−2.77
72 h relative cum. emission (frac. applied TAN) (n = 48 plots)	0.794	0.582	0.114	−0.0435
Training dataset → Calibration subset				
Flux (kg ha^{-1} h^{-1}) (n = 5371 intervals)	0.844	0.686	0.462	−0.194
72 h cum. emission (kg ha^{-1}) (n = 490 plots)	0.823	0.610	6.63	−2.94
72 h relative cum. emission (frac. applied TAN) (n = 490 plots)	0.802	0.530	0.118	−0.054

A semi-empirical model

Table 3
Model performance statistics based on parameter values given in Table 2 applied to both data subsets.

	Response	Pearson's r	ME[a]	MAE[b]	MBE[c]
Test dataset →	Evaluation subset				
	Flux (kg ha^{-1} h^{-1}) (n = 418 intervals)	0.832	0.663	0.433	−0.099
	72 h cum. emission (kg ha^{-1}) (n = 48 plots)	0.861	0.620	5.63	−2.77
	72 h relative cum. emission (frac. applied TAN) (n = 48 plots)	0.794	0.582	0.114	−0.0435
Training dataset →	Calibration subset				
	Flux (kg ha^{-1} h^{-1}) (n = 5371 intervals)	0.844	0.686	0.462	−0.194
	72 h cum. emission (kg ha^{-1}) (n = 490 plots)	0.823	0.610	6.63	−2.94
	72 h relative cum. emission (frac. applied TAN) (n = 490 plots)	0.802	0.530	0.118	−0.054

⇒ **Can we do better with machine learning ?**

Prediction tests with Machine Learning

- 3 methods tested : random forest, xgboost, lasso
- prediction of the cumul only (last observation)
- transformation of the dynamic variables (= climatic variables) into plot level variable

id	t.start	t.end	ct	j.NH3	e.cum	temp
182	0	4.00	4.00	1.79	7.15	8.20
182	4	21.00	21.00	0.07	8.29	4.45
182	21	44.75	44.75	0.15	11.83	7.22

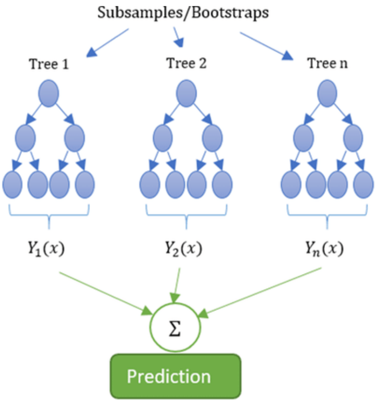
⬇

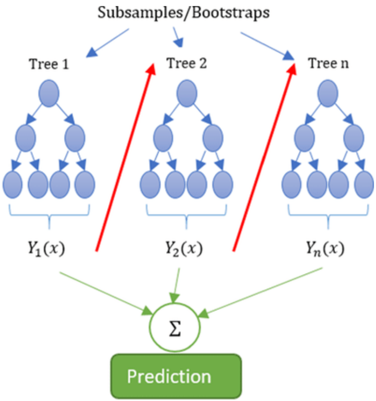
id	ct	e.cum	temp_1	temp_2	temp_3	temp_4	temp_5	temp_6	temp_7
182	44.8	11.831	8.2	4.45	4.45	4.45	4.45	6.5275	7.22

temp i = average temperature on the time interval [4(i-1) ; 4i] (i = 1, ..., 6)

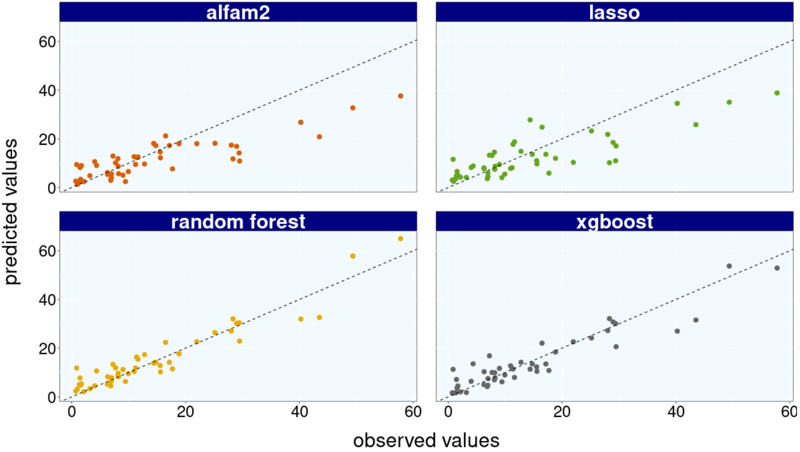
temp 7 = average temperature after 24h

Random forest and **Gradient boosting**

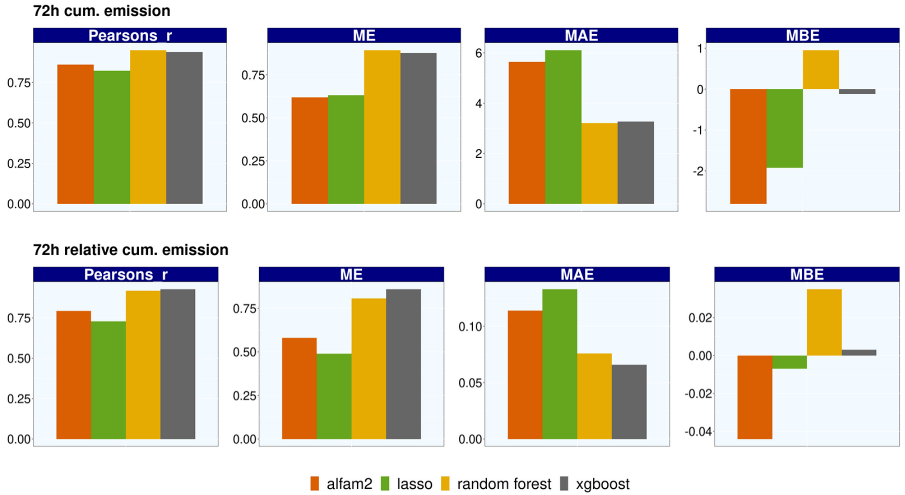




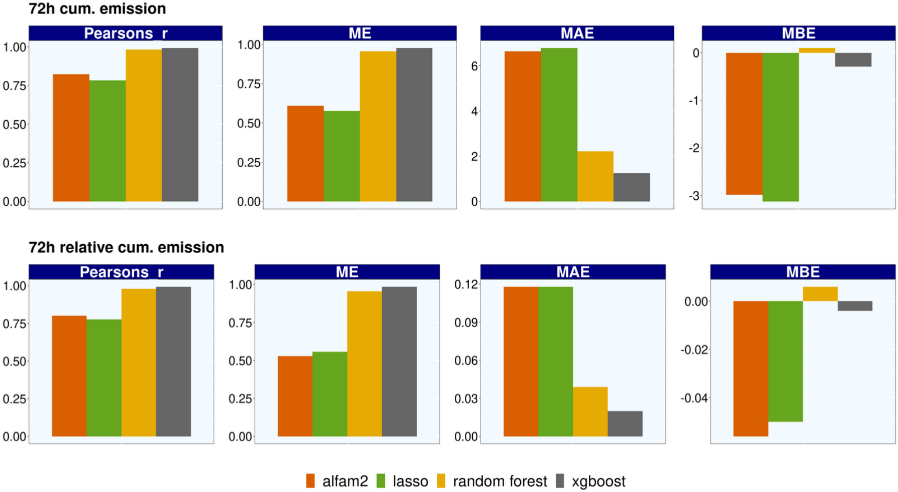
Results - Observed vs predicted values (test dataset)



Results - Test dataset



Results - Training dataset



Perspectives

Predict the dynamic of the volatilization

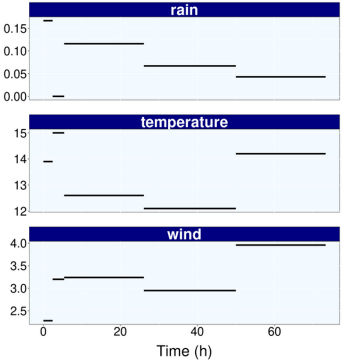
Perspectives

Predict the dynamic of the volatilization

Predictors

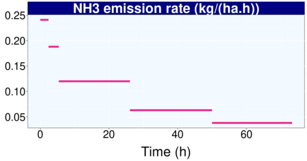
Response

Dynamical variables



Fixed variables

- **application method**
  => broadcast
- **manure source**
  => pig
- **total ammonia nitrogen**
  => 72.7
- **application rate**
  => 40
- **incorporation**
  => shallow
- **dry matter**
  => 8.13
- **pH** = 6.75



<=>