26th EAAP Congress, 28th August 2023

Sensitivity Analysis of INRAtion® V5 for Dairy Cows: Sobol' Indices

<u>Seoyoung Jeon[1]</u>, S. Lemosquet[2], T. Senga Kiessé[3], P. Nozière[1*]

[1]UMR Herbivores, INRAE-VetAgro Sup, France
[2]UMR PEGASE, INRAE-Institut Agro, France
[3]UMR SAS, INRAE-Institut Agro, France

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

INRAE

VetAgro Sup

L'INSTITUT agro

INRA 2018 feeding system for ruminants

Research Question

✓ ***How will the output variable be affected by this change?***

animal nutrient requirement and supply prediction model for cattle and small ruminants

- **INRAtion®V5**: Software (rationing tool)
  - **PrevAlim**: Generate feed table through given feedstuff information
  - **Diet optimizer:** Calculate diet composition suitable for rationing objective and/or predict animal's multiple responses

1. Direct measurement (in vivo, in sacco)

2. Predict with chemical components

3. Enzymatic method (in vitro)

INRAtion®V5

Feed information → **PrevAlim** → Feed table

« Feed characterization »
- Chemical component
- Digestibility
- Degradability

Animal information

Rationing objective

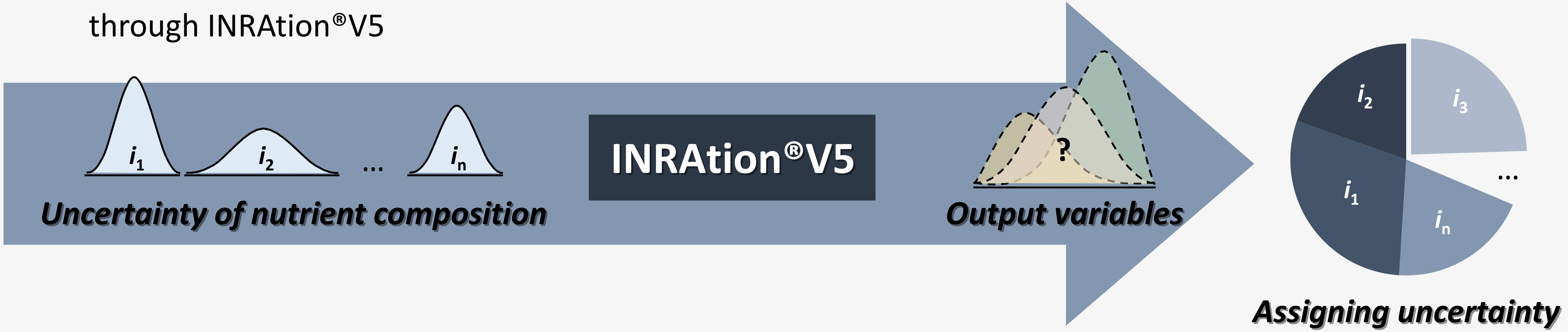
Diet optimizer

See Poster # 42391 (P. Nozière et al.) for more details

Sensitivity analysis

❑ Objective

- To evaluate the uncertainty of feedstuff's nutrient composition affect the output variables predicted through INRAtion®V5



❑ Sensitivity analysis

- How the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input (Saltelli, 2002)?
- **Global sensitivity analysis (GSA):** Change all variables at once
  - Interaction between input variables can be considered
  - Compared the relative influence of the input variables

Modified Groen et al. (2016) & Razavi et al. (2021)

Simulation conditions – Animal/Diet

❑ Animal information

- Multiparous (2nd parity, 43 kg/d of potential peak milk yield), calving at 50 months of age
- 608 kg of BW (BCS 2.29), wks 14 of lactation

❑ Diet

- **Rationing objective:**
  - Reach animal potential milk production
- **6 common dairy cow diet** formulated by INRAtion®V5
  - **RF**: Fresh perennial ryegrass based diet
  - **GH1,2,3**: Grass hay based diets
  - **GS**: Grass silage based diet
  - **CS**: Corn silage based diet

Table 1. Diet composition (%DM) for multiparous

	RF	GH1	GH2	GH3	GS	CS
Fresh perennial ryegrass	72.9	-	-	-	-	-
Grass hay (1st growth)	-	54.2	-	3[1]	-	-
Grass hay (2nd growth)	-	-	84.2	52.9	-	-
Grass silage	-	-	-	-	62.1	-
Corn silage	-	-	-	-	-	75.6
Soybean meal	3.6	-	-	-	4.2	13.6
Rapeseed meal	-	21.1	-	18.6	-	-
Barley	-	-	-	-	-	10.8
Corn	23.5	24.7	15.8	58.5	33.7	-

[1] 3kg with fixed amount

Simulation conditions

- **Simulation** → Selected as being the most influent on nutrient supply (Jeon et al., 2023)
  - **Input variables** (N = 5)
    - **CP** (crude protein), **GE** (gross energy), **OMd** (OM digestibility), **ED6_N** (effective N degradability), and **dr_N** (true intestinal N digestibility)
  - **Outputs**
    - **Protein related:** PDIA, PDIM, PDI, PDI/UFL, RPB, LysDI, MetDI, PDI efficiency, etc.
    - **Energy related:** OMd, UFL, UFL balance, etc.
    - **Intake related :** Predicted DMI, Forage intake, concentrate intake, SRg, etc.
    - **Animal performance related** : Milk yield, Milk yield / potential MY, Milk protein yield, Milk protein contents, N utilization efficiency, etc.
    - **Environment related**: Urinary N, Fecal N, Total excreted N, Urinary N / Total excreted N, ECH4, etc.

Simulation conditions

❑ Simulation

→ Selected as being the most influent on nutrient supply (Jeon et al., 2023)

- **Input variables** (N = 5)
  - **CP** (crude protein), **GE** (gross energy), **OMd** (OM digestibility), **ED6_N** (effective N degradability), and **dr_N** (true intestinal N digestibility)
- **Outputs**
  - **Protein related:** PDIA, PDIM, PDI, PDI/UFL, RPB, LysDI, MetDI, PDI efficiency, etc.
  - **Energy related:** OMd, UFL, UFL balance, etc.
  - **Intake related :** ***Predicted DMI***, Forage intake, concentrate intake, SRg, etc.
  - **Animal performance related** : Milk yield, Milk yield / potential MY, ***Milk protein yield***, Milk protein contents, ***N utilization efficiency***, etc.
  - **Environment related**: Urinary N, Fecal N, Total excreted N, ***Urinary N / Total excreted N***, ***ECH4***, etc.
- N = 10,000 quasi-random samples
  - Around pivot situation (±3.0 ~ ±5.5% CV)
  - Assume that 5 input variables have independent normal distributions

Sensitivity index – Sobol index

- **Sobol index** (Sobol, 1993)
  - Variance-based sensitivity analysis
  - **(1) First order index (S_i)**: Evaluate the importance of one input variable ***without considering interaction***
  - **(2) Total indices (ST_i)**: Evaluate the importance of one input variable ***considering interaction*** with other input variables
  - **(3) Interaction (ST_i – S_i)**: Level of interaction with other input variables

- **R program**: *sensobol* packages

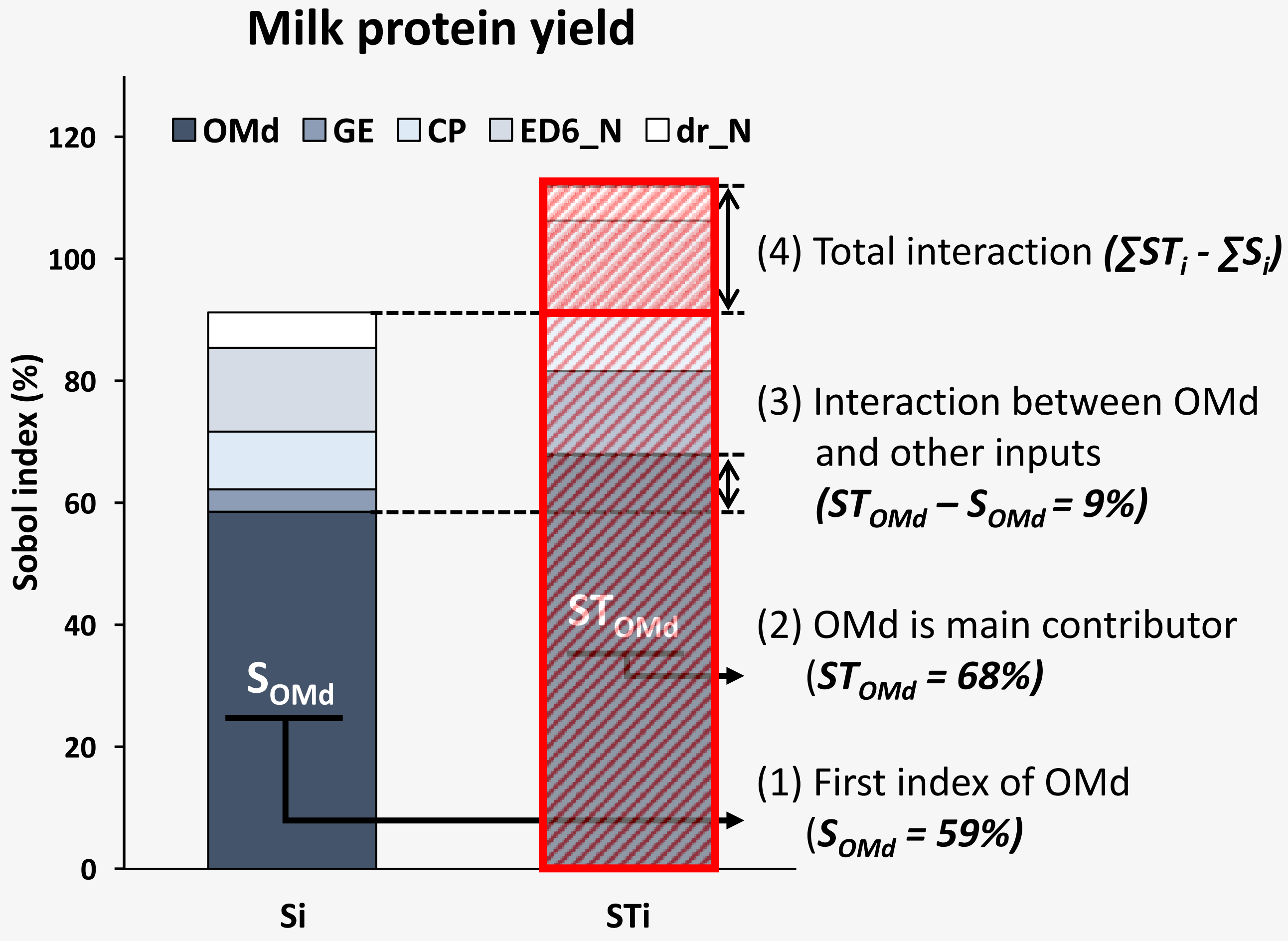
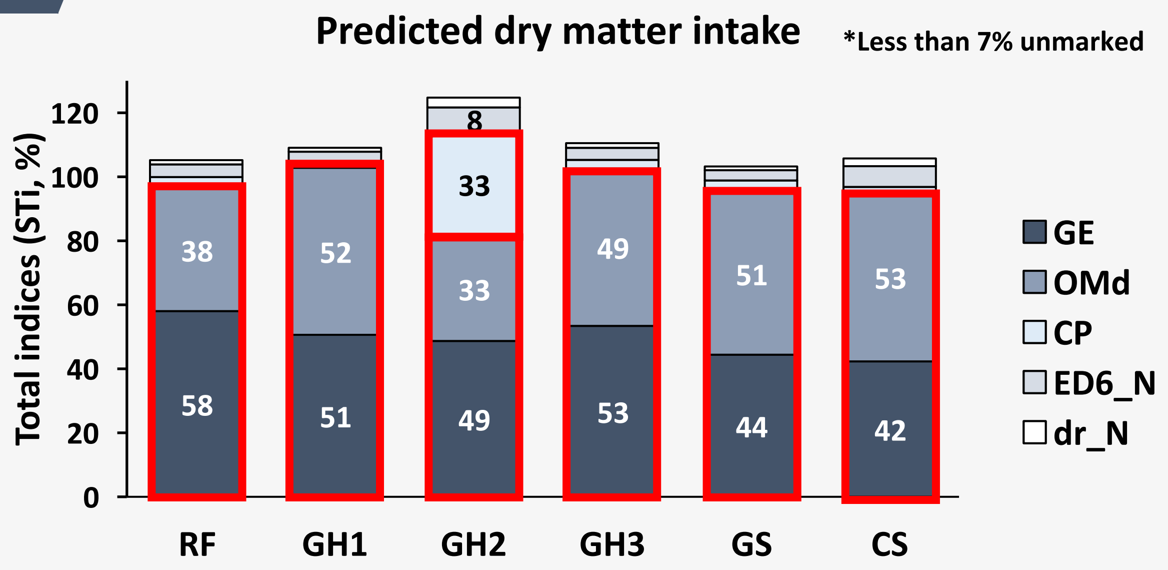


Figure 1. Response of milk protein yield to change in 5 input variables

RF: Perennial ryegrass fresh, **GH**: Grass hay, **GS**: Grass silage, **CS**: Corn silage

Predicted dry matter intake (DMI)



- DMI is sensitive mainly to Energy related-variables (GE and OMd)
  - CP was one of the main variable affecting DMI in GH2
    - High content (84% DM) of high CP forage (20% CP)

Table 2. Contribution of interactions (% of total interactions)

	RF	GH1	GH2	GH3	GS	CS
$\sum ST_i - \sum S_i$	8.8	18.0	46.3	21.1	4.7	9.5
% of total interaction						
GE	42.1	46.3	41.5	46.4	31.6	43.4
OMd	51.4	44.3	36.3	43.1	57.7	31.4
CP	1.4	-	1.1	-	4.3	-
ED6_N	5.1	7.7	14.5	8.5	6.5	23.3
dr_N	-	1.8	6.6	1.9	-	1.9

ED6_N: Effective N degradability, **dr_N**: true intestinal N digestibility

- Input variables highly interact in some diet
  - In all, there are GE and OMd were most interacting variables
    - GE and OMd are used for net energy calculation
  - Interaction from ED6_N and dr_N cannot be ignored (e.g. GH2).
    - It would be related to the CP level of the GH2

RF: Perennial ryegrass fresh, GH: Grass hay, GS: Grass silage, CS: Corn silage

Milk protein yield (MPY)

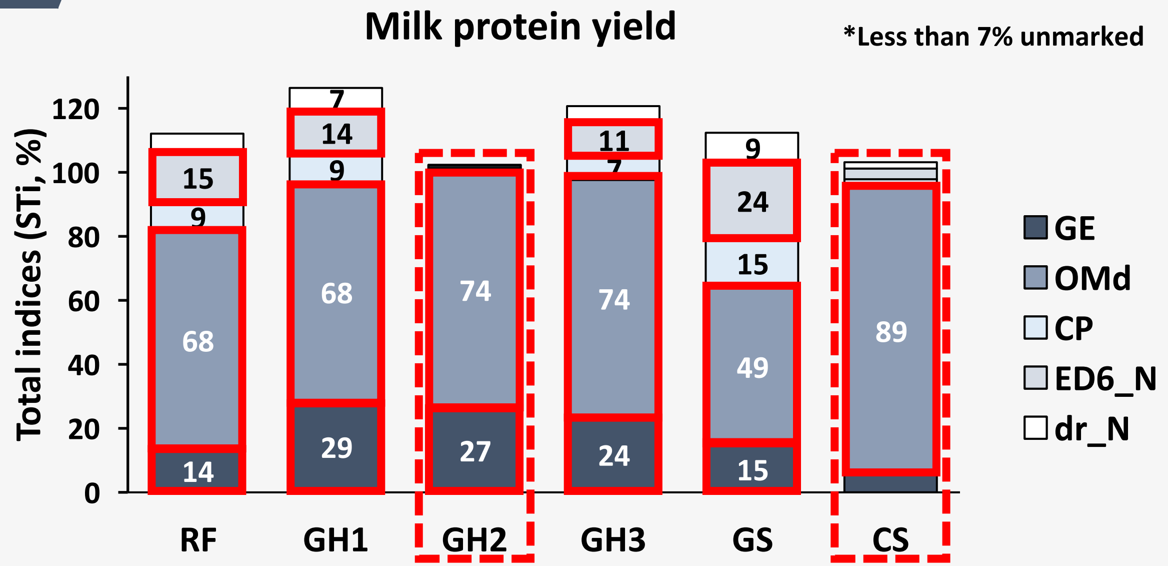


Table 3. Contribution of interactions (% of total interactions)

	RF	GH1	GH2	GH3	GS	CS
$\sum ST_i - \sum S_i$	21.1	50.1	4.1	38.9	20.9	5.2
% of total interaction						
GE	47.6	46.0	41.0	45.9	44.8	46.4
OMd	44.1	41.9	30.6	41.1	45.5	29.9
CP	-	-	5.5	-	-	-
ED6_N	8.2	9.2	15.8	9.9	9.6	20.2
dr_N	0.1	2.8	7.1	3.2	-	3.6

ED6_N: Effective N degradability, **dr_N**: true intestinal N digestibility

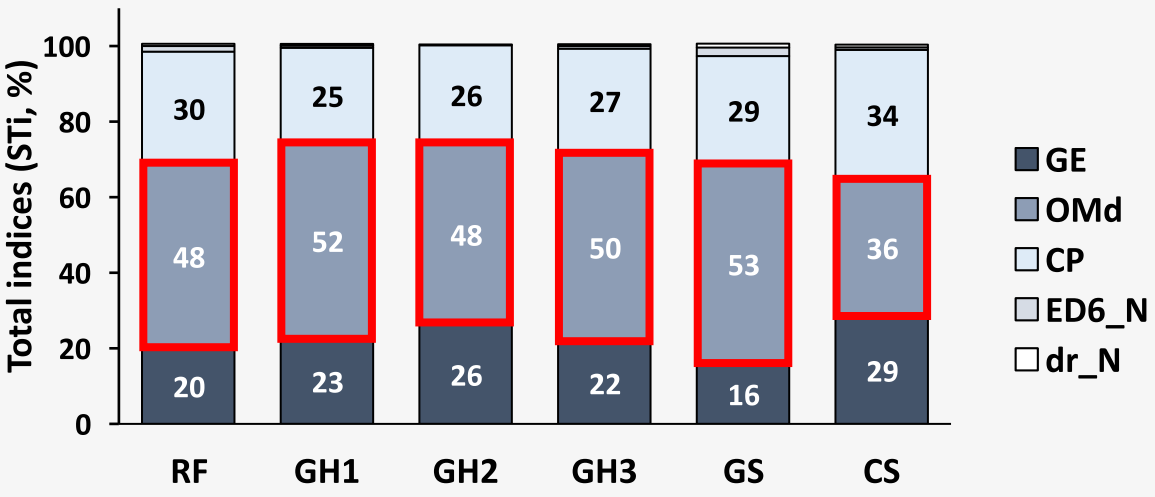
- MPY is sensitive mainly to OMd
  - GE and ED6_N also high contribute depending on the diet
  - In the GH2 and CS, there was almost no contribution from variables other than OMd and GE.

- Input variables highly interact in some diet
  - In all, there are GE and OMd are the most interacting variables.
  - The interaction of ED6_N cannot be ignored.

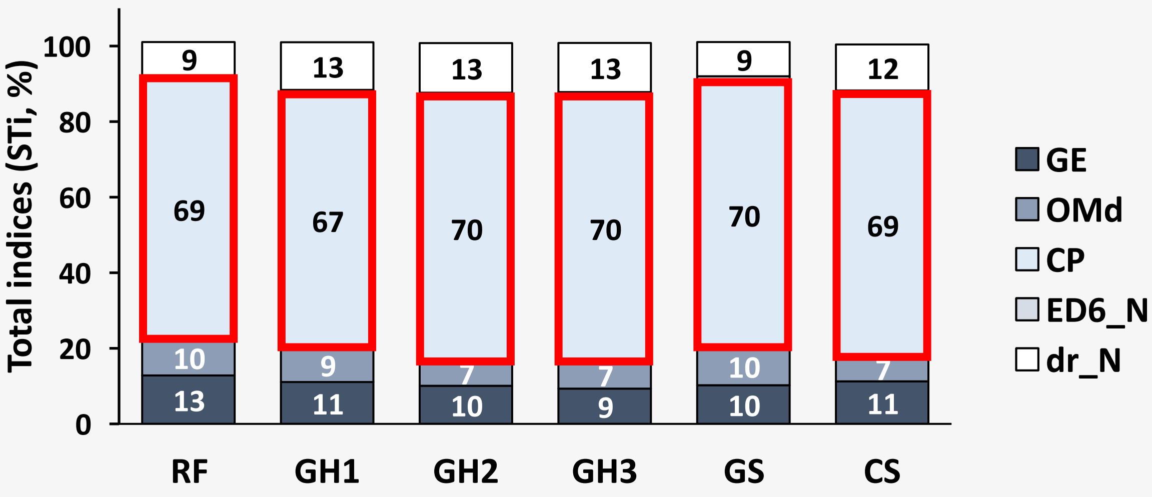
Environmental output variables

RF: Perennial ryegrass fresh, **GH**: Grass hay, **GS**: Grass silage, **CS**: Corn silage

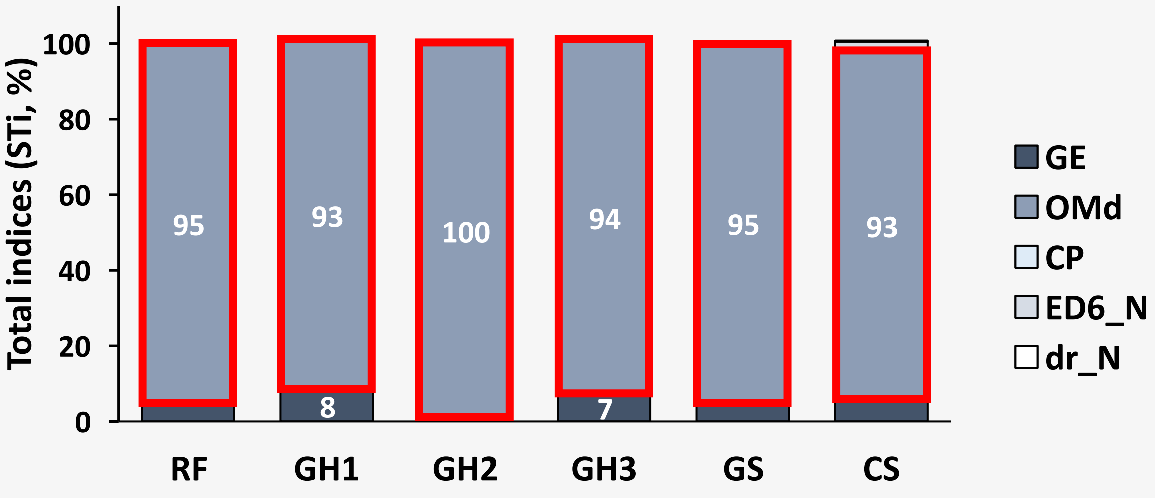
Nitrogen utilization efficiency *Less than 7% unmarked



Urinary N / Total N *Less than 7% unmarked



Energy in Methane *Less than 7% unmarked



- NUE is sensitive mainly to OMd, GE, and CP
  - NUE = N in milk / N intake
- UN/TN is sensitive mainly to CP
  - UN/TN = Urinary N / (Urinary N + Fecal N)
- ECH4 is sensitive mainly to OMd, and the others has a few impact on ECH4
- There is little interaction among input variables

Conclusion

Objective of present work

✓ ***To evaluate the uncertainty of feedstuff's nutrient composition affect the output variables predicted through INRAtion®V5***

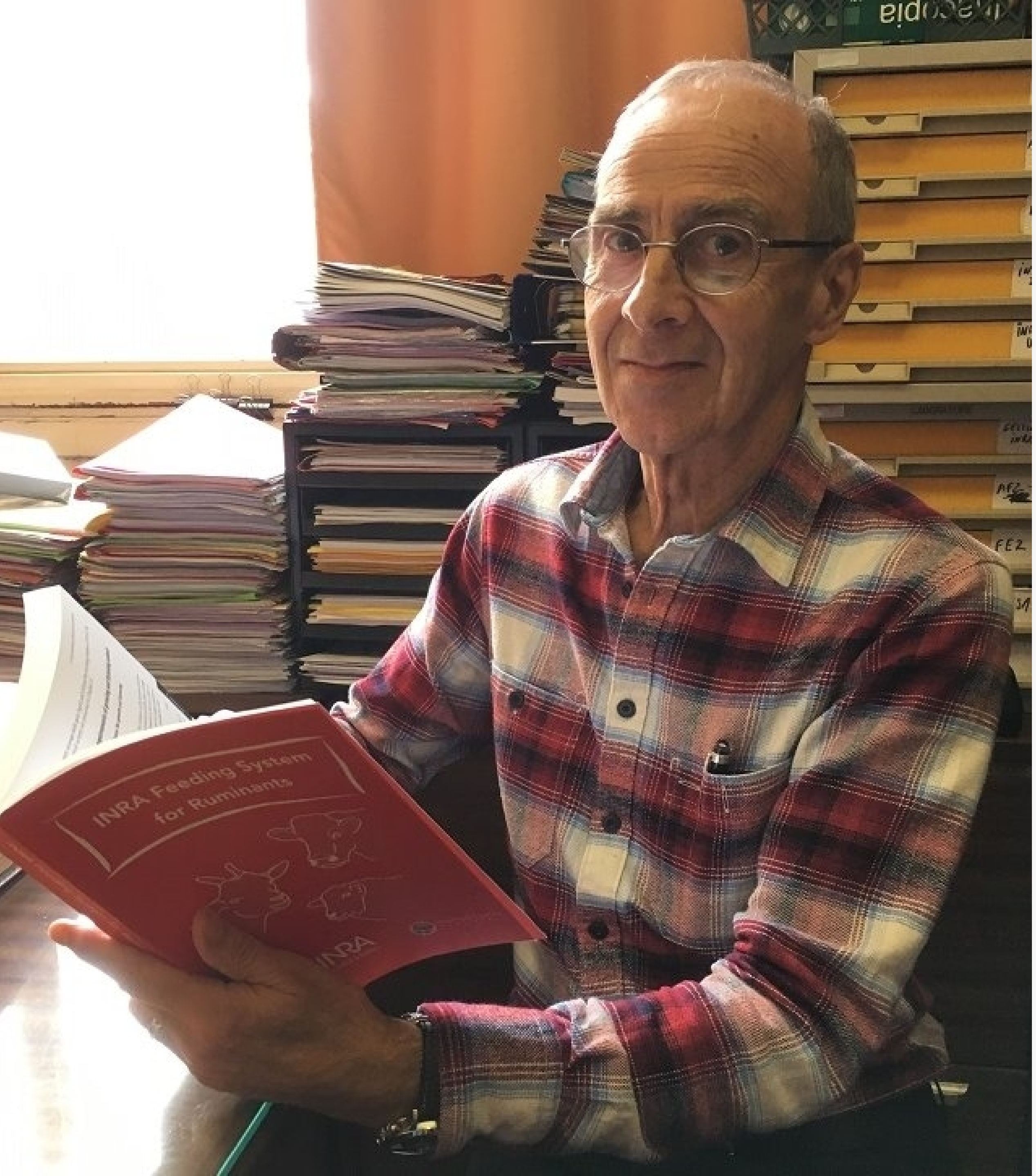
❑ The response of the output to the uncertainty of each input variable around animal potential

- **GE:** influential in all predicted animal responses except ECH4.
- **OMd:** have a great impact on all predicted animal responses except UN/TN.
- **CP:** cause variation in nitrogen excretion related variables.
  - Depending on the diet, intake and MPY are also affected.
- **ED6_N and dr_N:** have an impact on MPY and UN/TN, respectively.
- The responses of DMI and MPY were different depending on the diet.

❑ Perspective

Our next step

- Trend finding (GSA) → Quantification (Jeon, 2023; ADSA) → Generalization
- Continue to conduct sensitivity analysis under more diverse conditions (diets, animal characteristics, rationing objectives, and so on)



THANK YOU FOR YOUR ATTENTION

Contact

- seoyoung.jeon@inrae.fr / seoyoung203@gmail.com

Supervisors

- Pierre NOZIÈRE
- Sophie LEMOSQUET
- Tristan SENGA KIESSÉ

Committee

- Rémy DELAGARDE
- René BAUMONT
- Luc DELABY
- Yayu HUANG
- Patrick CHAPOUTOT
- Gaëlle MAXIN

Acknowledgement

- Anne-Cécile TOULEMONDE (Intern)

See Poster # 42391 (P. Nozière et al.)

Nutritional specifications of simulation diets

Table 4. Nutritive values (g/kg DM) of each diet for multiparous

	RF	GH1	GH2	GH3	GS	CS
OM	926	937	**922**	939	928	**955**
CP	143	150	**183**	150	**140**	141
NDF	414	443	**486**	446	389	**345**
ADF	199	239	**245**	237	207	**175**
EE	37	**25**	31	26	**34**	28
GE (kcal/kg DM)	**4339**	4448	4527	4456	**4481**	4495
UFL (/kg DM)	**0.98**	**0.81**	0.93	0.83	0.94	0.93
OMd (% DM)	**73.9**	**61.3**	70.2	62.6	69.8	68.3
ED6_N (% DM)	**66.3**	**58.2**	63.1	58.7	64.5	62.5
dr_N (% DM)	86.0	**79.3**	**86.9**	80.6	82.7	83.5
PDI	94.5	97.6	**109.2**	98.0	**89.3**	92.6
PDI/UFL (g/UFL DM)	96.4	**120.7**	117.5	118.8	**94.6**	99.2

Red: Highest value; Blue: lowest value

05 What if there is no interaction between input variables?

- **Sobol index** (Sobol, 1993)
  - Variance-based sensitivity analysis

  (1) First order index (S_i): Evaluate the importance of one input variable (*no interaction*)

  (2) Total indices (ST_i): Evaluate the importance of one input variable considering ***interaction*** with other input variables

  (3) Interaction (STi – Si): Level of interaction with other input variables

- **R program**: *sensobol* packages

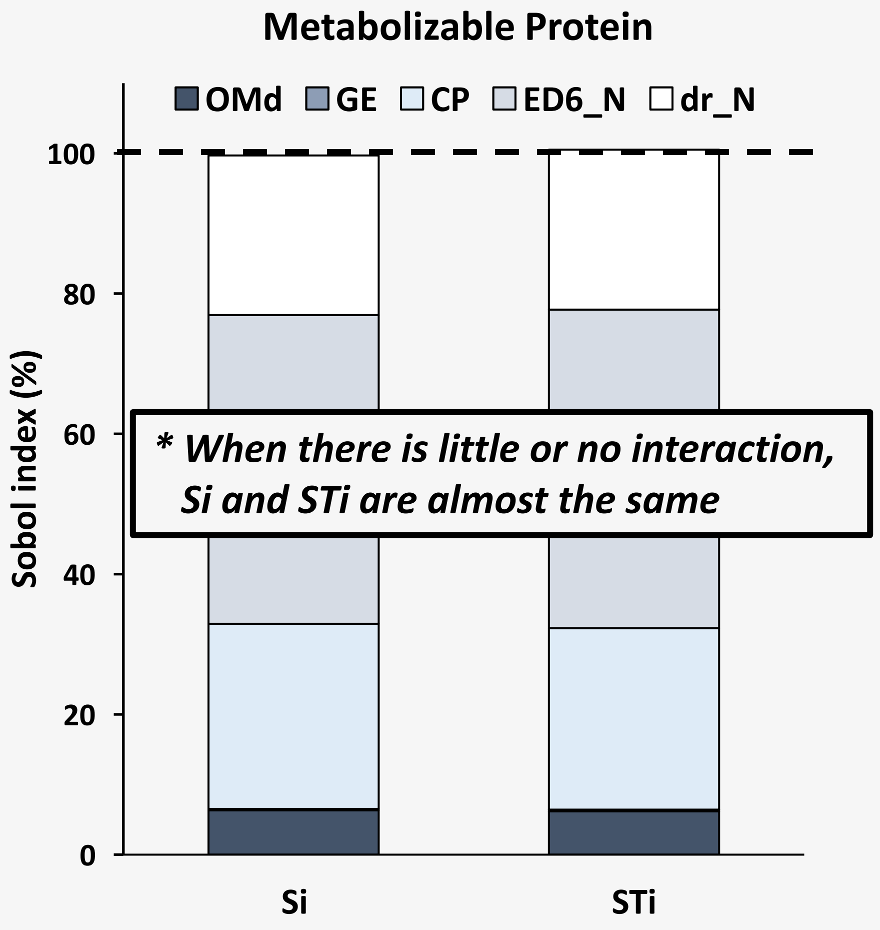


Figure 1-2. Response of metabolizable protein to change in 5 input variables

Why sobol?

- Big assumption of 'Pearson' and 'Spearman' correlation coefficient
  - 1) Linearity and/or 2) Monotonicity
  - However, most models of the INRA feeding system didn't satisfy this requirement
- What if we apply it to a model that is not linear or monotonicity?
  - Y1 = 10·X1 + X2 + X3 + X4 + X5
  - Y2 = 10·X1^2 + X2 + X3 + X4 + X5

* X1, X2, X3, X4, X5 ← Randomly select between -25 to 25 (n = 700; n = 100 for sobol)

Table 5. Sensitivity indices of input variables for **Y1**

	Pearson	Spearman	Sobol (ST_i)
X1	0.98	0.98	1.00
X2	0.08	0.08	0.01
X3	0.03	0.04	0.01
X4	0.02	0.05	0.01
X5	0.09	0.09	0.01

Table 6. Sensitivity indices of input variables for **Y2**

	Pearson	Spearman	Sobol (ST_i)
X1	0.06	<0.01	0.96
X2	0.06	0.02	<0.01
X3	0.07	0.06	<0.01
X4	<0.01	0.03	<0.01
X5	<0.01	<0.01	<0.01

05 Sobol analysis sequence using R

1. Create Sobol matrix

	OMd	ED6_N	CP	dr_N	GE
1	0.50	0.50	0.50	0.50	0.50
2	0.75	0.25	0.75	0.25	0.75
…	…	…	…	…	…
n	0.41	0.26	0.30	0.48	0.70

2. Assigning value to the Sobol matrix

Feed1

	OMd	ED6_N	CP	dr_N	GE
1	87.6	44.2	89.1	88.7	4497.5
2	90.9	42.6	92.4	85.5	4664.4
…	…	…	…	…	…
n	86.6	42.7	86.6	88.5	4623.9

Feed2

	OMd	ED6_N	CP	dr_N	GE
1	74.4	69.0	378.7	79.4	4604.7
2	77.2	66.4	392.8	76.4	4775.5
…	…	…	…	…	…
n	73.5	66.5	368.1	79.1	4734.1

Feed3

	OMd	ED6_N	CP	dr_N	GE
1	69.4	67.6	88.0	86.9	4503.2
2	71.9	65.1	91.3	83.6	4670.2
…	…	…	…	…	…
n	68.5	65.2	85.6	86.6	4629.7

3. Extract sobol matrices
- For PrevAlim running

4. **PrevAlim** running
→ Obtain full feed table of each feed

5. **Diet optimizer** running
- Obtain output matrix

6. Calculate Sobol index

Number of sample in sobol matrix

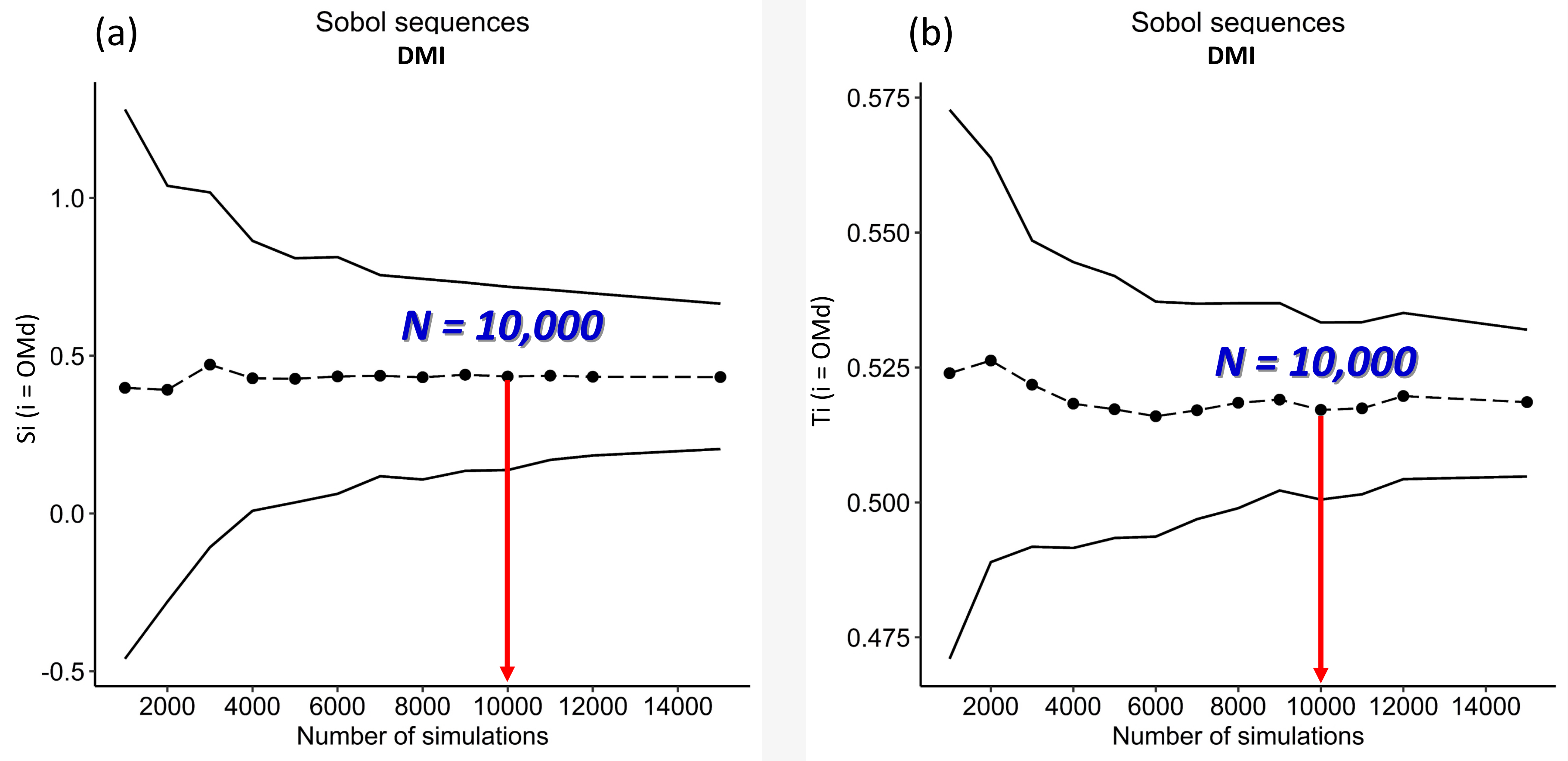


Figure 2. (a) First sobol index (Si, round) and confidence interval (solid line)
(b) Total sobol indices (Ti, round) and confidence interval (solid line)

05 Why does ∑STi > 100% ?

❑ Due to redundant counting of values for interaction when calculate $\sum ST_i$

For example, there are 3 input variables (A, B, C) for an output variable (i)

- $\mathbf{ST_A}$ = S_A + **Inter(A,B)** + **Inter(A,C)** + **Inter(A,B,C)**
- $\mathbf{ST_B}$ = S_B + **Inter(A,B)** + **Inter(B,C)** + **Inter(A,B,C)**
- $\mathbf{ST_C}$ = Si_C + **Inter(B,C)** + **Inter(A,C)** + **Inter(A,B,C)**

- $\sum ST_i$ = [S_A + Inter(A,B) + Inter(A,C) + Inter(A,B,C)] + [S_B + Inter(A,B) + Inter(B,C) + Inter(A,B,C)] + [S_C + Inter(B,C) + Inter(A,C) + Inter(A,B,C)]

  = S_A + S_B + S_C + **2 Inter(A,B)** + **2 Inter(A,C)** + **2 Inter(B,C)** + **3 Inter(A,B,C)**

Selection process of 5 input variables

Feedstuff approach

- Average: Red book table
- Variation: BIPEA report

Collect average & variation values

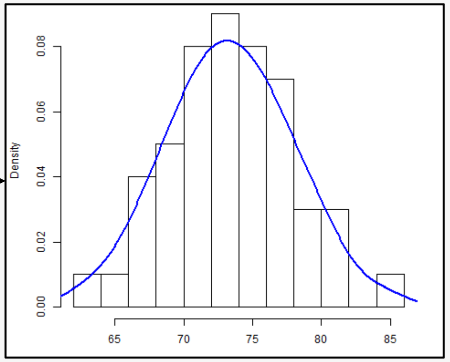
N < 9 → Normalization

N ≥ 9

Latin hypercube (LHS) random sampling (n = 1,000)

Normalized average & SD value

Random sampling (LHS, n = 50)



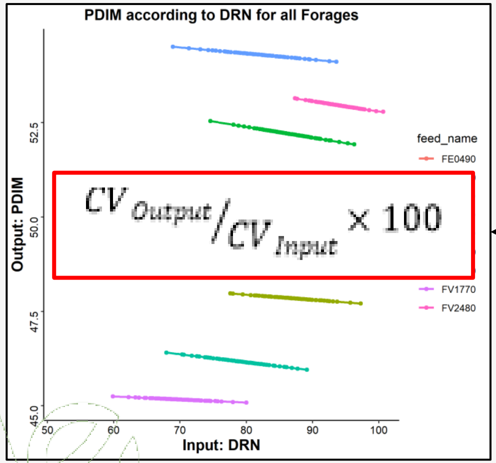
Change 1 input variable at a time — One-at-a-time

Global SA — Change all input variables simultaneously

PrevAlim running

Sensitivity analysis

GAMs fitting



Select important input variables

OMd, dr_N, ED_6_N, GE, CP

Diet approach

Forage

- **Feed list (9)**
  - **Fresh**: grass, ryegrass, whole crop corn, white clover
  - **Hay**: grass, alfalfa
  - **Silage:** grass, corn
- **Inputs (9)**: ADF, FA, OMd, dr_N, ED6_N, GE, GP, Ash, NDF
- **Outputs (17)**: PDIA, PDIM, PDI, UFL, UFV, UEM, UEB, UEL, RPB, LysDI, MetDI, ED6_N, ED_N, dr_N, OMd, GE, FL_ref

Concentrate

- **Feed list (13)**
  - Barley, Corn, Mixed concentrate, Faba baen, Soybean, Beet pulp, Wheat bran, Corn gluten feed, Corn gluten meal, Rapeseed meal, Soybean meal
- **Inputs (12)**: ADF, ADL, FA, Starch, OMd, dr_N, ED6_sta, ED6_N, GE, CP, Ash, NDF
- **Outputs (17)**: PDIA, PDIM, PDI, UFL, UFV, FVc, RPB, LysDI, MetDI, ED6_N, ED_N, dr_N, ED6_sta, ED_sta, OMd, GE

Quantification in One-at-a-time approach

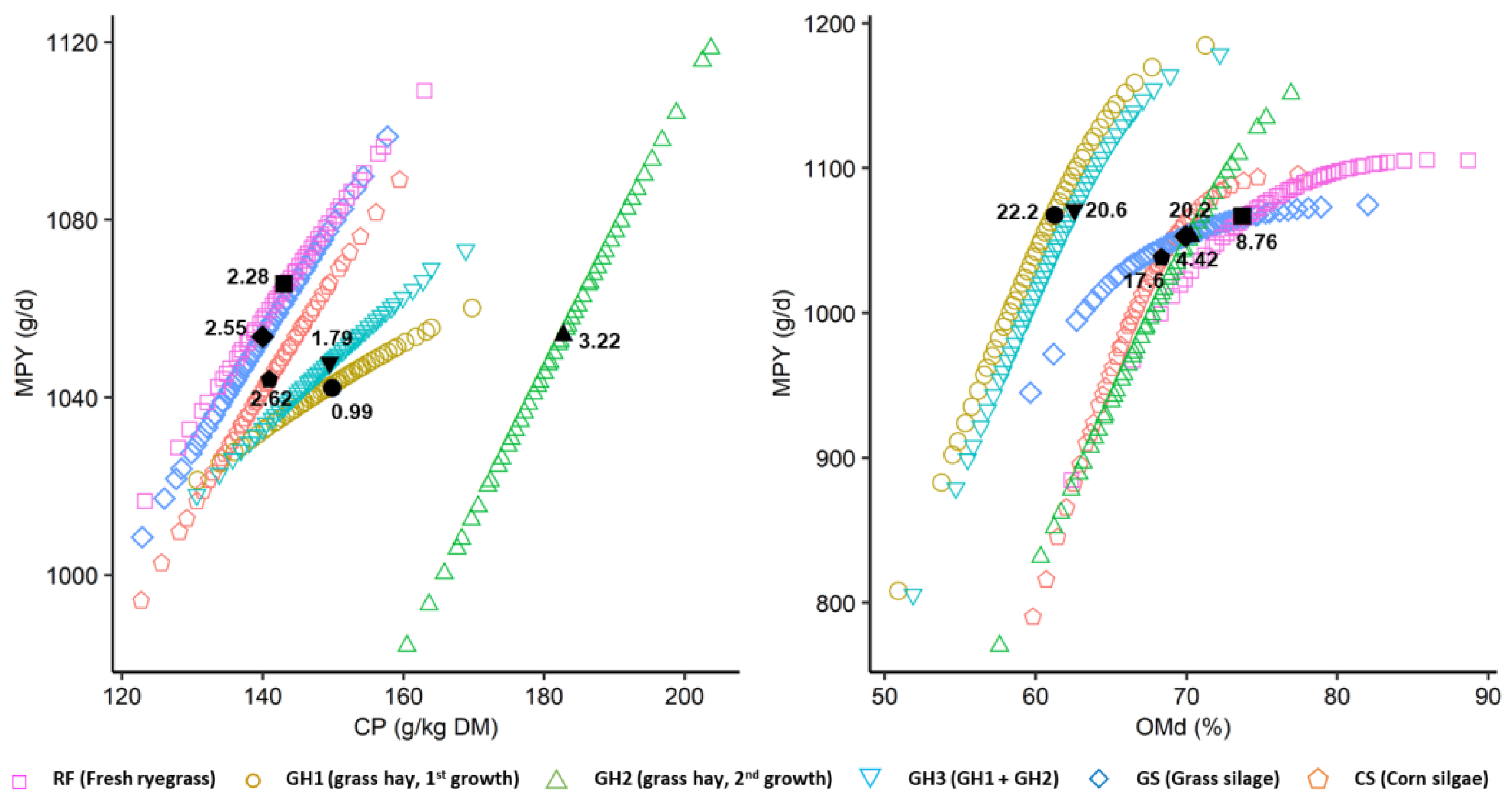


Figure 3. Response of milk protein yield when CP (g/kg DM) and OMd (%) are varied. Black filled mark are the pivot point for each diet, and the digits around it are the tangent values at each point (output unit / input unit)

Table 7. Normalized sensitivity coefficient (NSC[1], %)

Diet	CP	OMd
RF	30.6	60.5
GH1	14.3	127
GH2	55.7	135
GH3	25.6	121
GS	33.9	29.4
CS	35.4	116

$$^{1}\mathrm{NSC}_{ij}(\%) = \frac{\Delta y_j}{\Delta x_i} \times \frac{Avg(x_{1i},x_{2i})}{Avg(y_{1j},y_{2j})} \times 100 = \frac{y_{2j} - y_{1j}}{x_{2i} - x_{1i}} \times \frac{Avg(x_{1i},x_{2i})}{Avg(y_{1j},y_{2j})} \times 100$$

Why is the DMI of GH2 greatly affected by CP?

❑ Factors determining predicted DMI in INRA 2018 feeding system

- It was OMd, CP, and SRg of forage that affected DMI. SRg was mainly affected by GE and OMd.
- In the case of GH2 diet, the forage content was highest (84%) as well as forage CP was high (20%), so, it seems that the influence of CP was higher than the others.

IC → ~~BW, Wk of lactation, Wk of gestation, Age~~, **PDI/UFL**

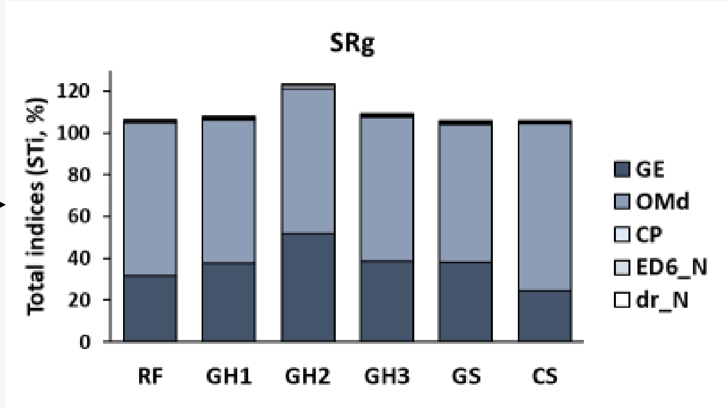
$$\frac{IC}{FVf \times (1 - PCO) + FVc \times PCO}$$

PCO → ~~Forage: Concentrate ratio~~

FVf → UEL/BW$^{0.75}$ kg → 140/VDMI$_L$

FVc → FVf x SRg

- RF = 66.3 + 0.6550·OMd + 0.098·CP + 0.626·DM – 3.7
- GH = 82.4 + 0.491·OMd + 0.114·CP +4.6
- GS = 99.3 + 0.167·OMd + 0.128·CP -1.4
- CS = -76.4 + 2.39·OMd + 1.44·DM



Results of other output variables