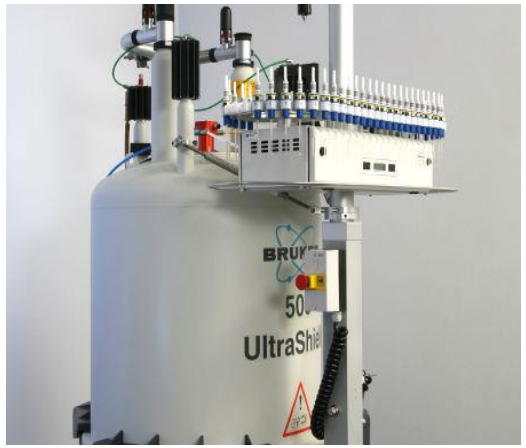
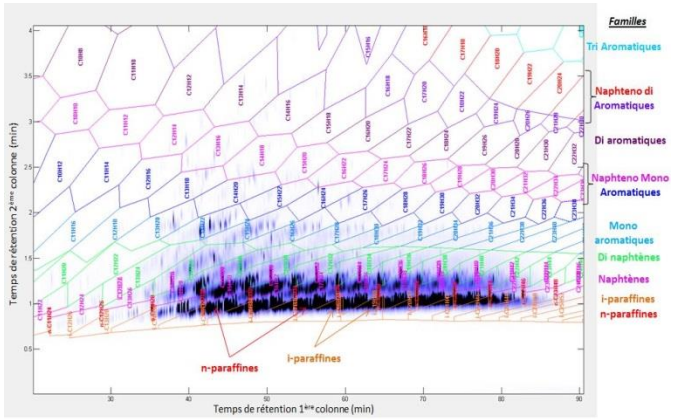
ON THE USE OF SPARSE-PLS TO EXPLAIN PETROLEUM PROPERTIES AT THE MOLECULAR LEVEL

J-J. Da Costa Soares[1], F. Chainet[1], B. Celse[1], M. Lacoue-Nègre[1], C. Ruckebusch[2], N. Caillol[1], D. Espinat[1]

[1]IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP3, 69360, Solaize, France

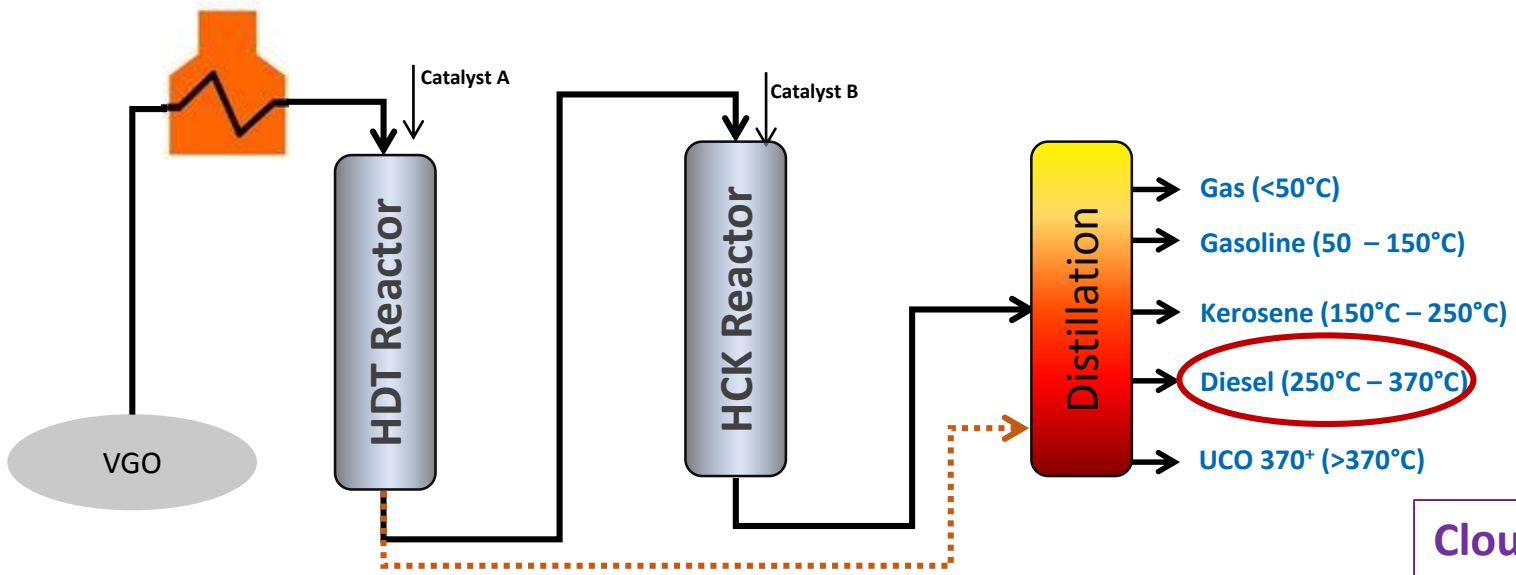
[2]LASIR, Université de Lille, Sciences et Technologies, 59655 Villeneuve d'Ascq Cedex, France





CONTEXT

RESPONSIBLE
OIL AND GAS



Standard measurement methods (ASTM and EN methods)

- Process optimization: time consuming and costly experiments → process modelling (feedstock properties, catalyst…)
- But some properties are poorly predicted

→ **Need to understand and explain the property at the molecular level**

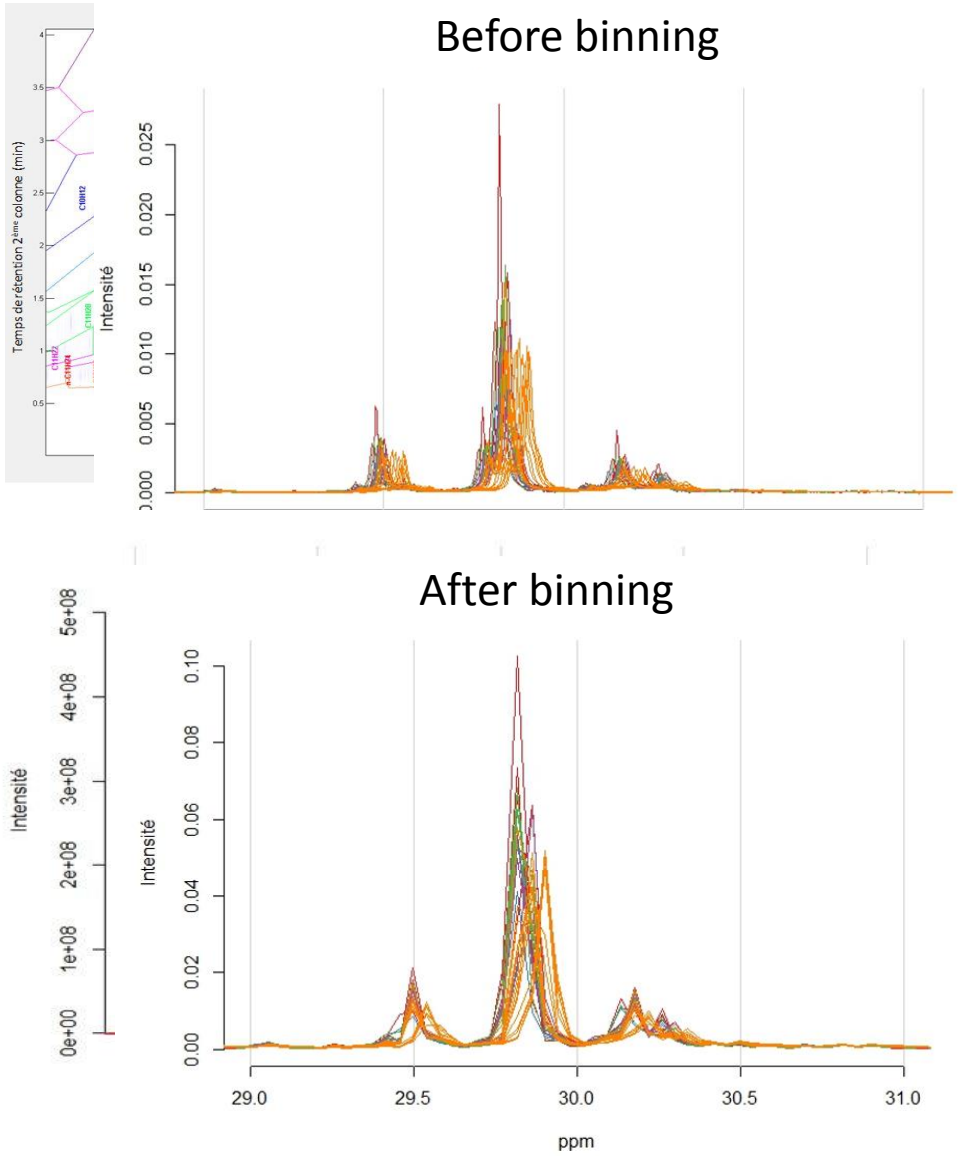
Cloud Point (CP)

- ASTM D2500 or NF ISO 23015
- The temperature at which wax crystals begin to form in a petroleum product as it is cooled

Analytical techniques: ^{13}C NMR and GCxGC

ANALYTICAL METHODS

- GCxGC → molecular families
  - 152 variables
  - Preprocessing : mean centering and autoscaling
- ^{13}C NMR → carbons molecular environment
  - Up to 65k variables → variable selection: 0 to 60 ppm (13k variables)
  - Preprocessing : binning , normalisation and mean centering



CHEMOMETRICS

RESPONSIBLE
OIL AND GAS

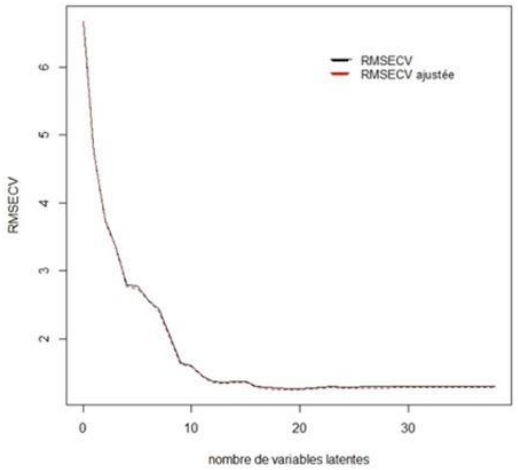
- Methodology :
  - Development of a model to predict the cloud point from NMR and GCxGC data
  - Interpretation of the loadings to extract the relevant information related to the property

- 1/ PLS
  - Estimation of the number of LVs
  - Interpretation of the loadings
- 2/ Sparse-PLS
  - Sparsity when constructing the direction vectors
  - Variable selection penalty → suppression of the lowest coefficients

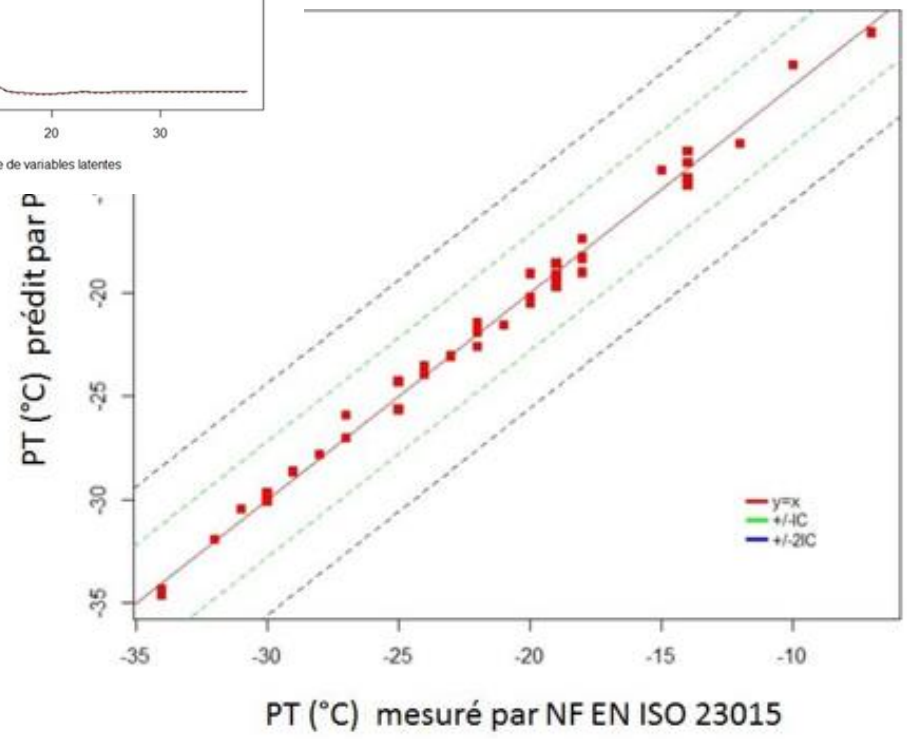
RESULTS – GCXGC DATA (1/3)

RESPONSIBLE OIL AND GAS

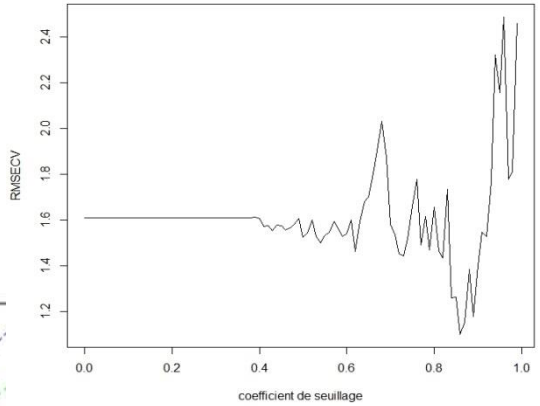
Model	LV	RMSECV (°C)	RMSEC (°C)	$\tau_{\pm IC}$ (%)
PLS	10	1,6	0,56	100
sparse PLS	10	1,2	0,48	100



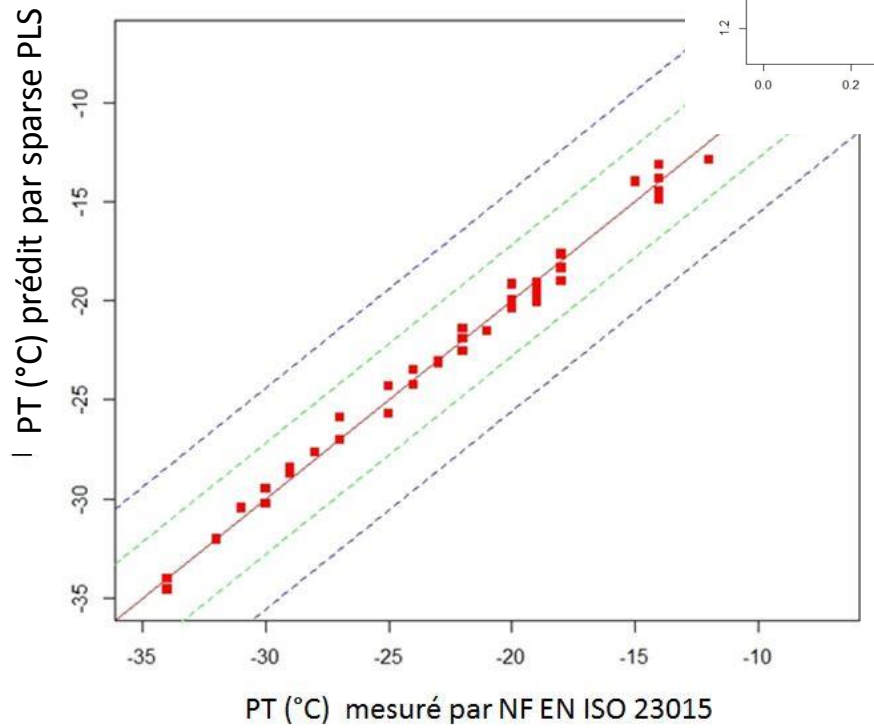
PLS model



Sparse PLS model



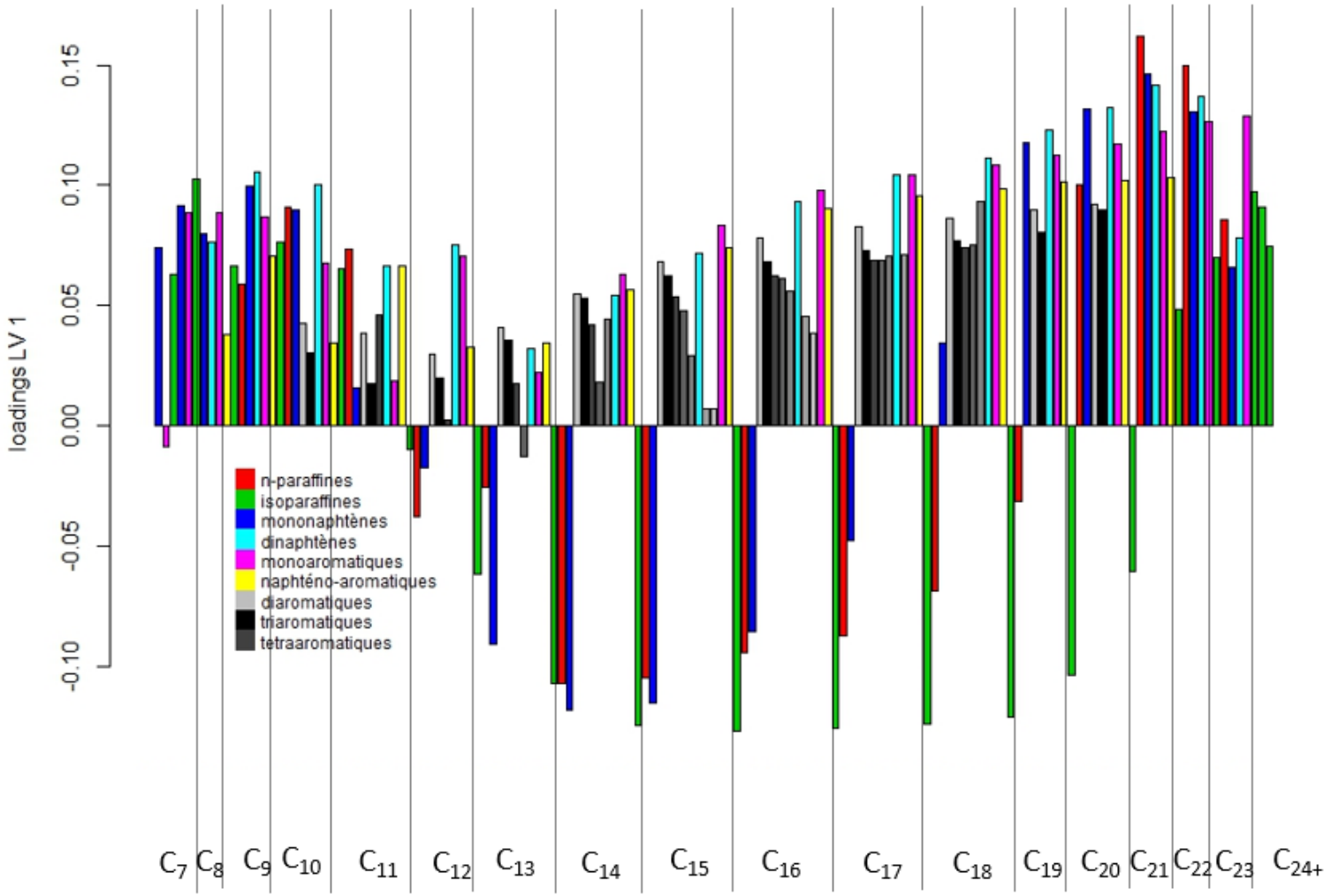
$\eta = 0{,}86$



RESULTS – GCXGC DATA (2/3)

RESPONSIBLE
OIL AND GAS

Example: Loading PLS LV1

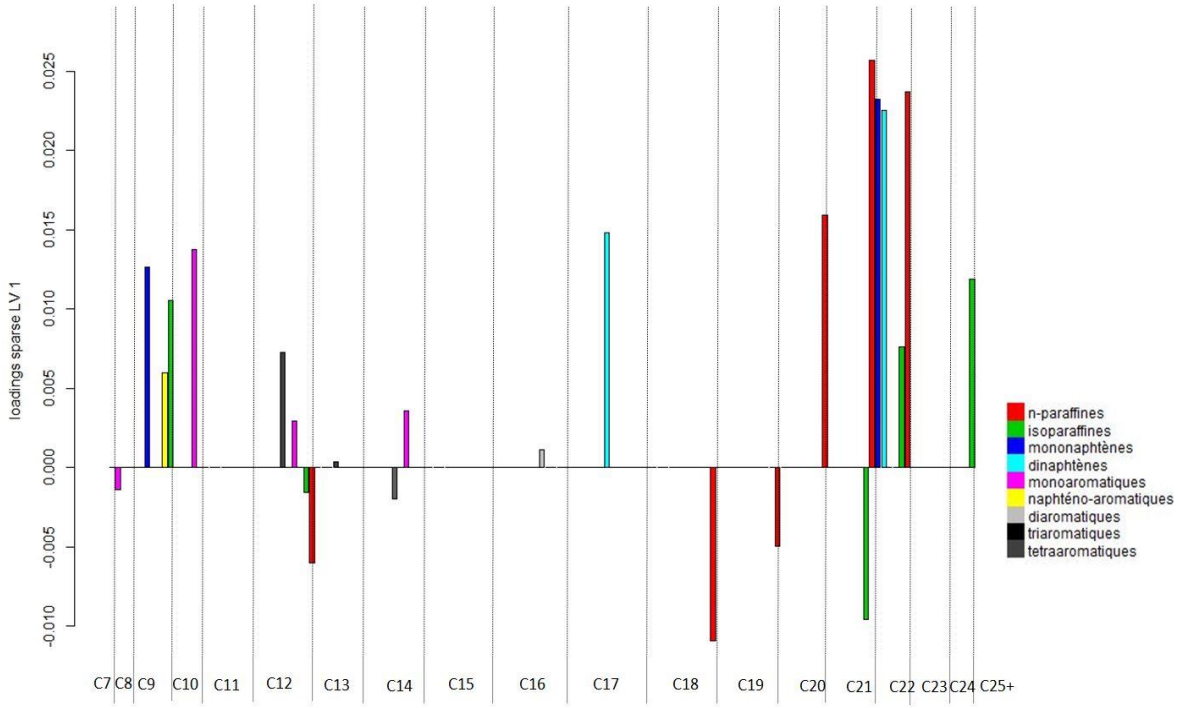


Where is the relevant information to explain the cloud point

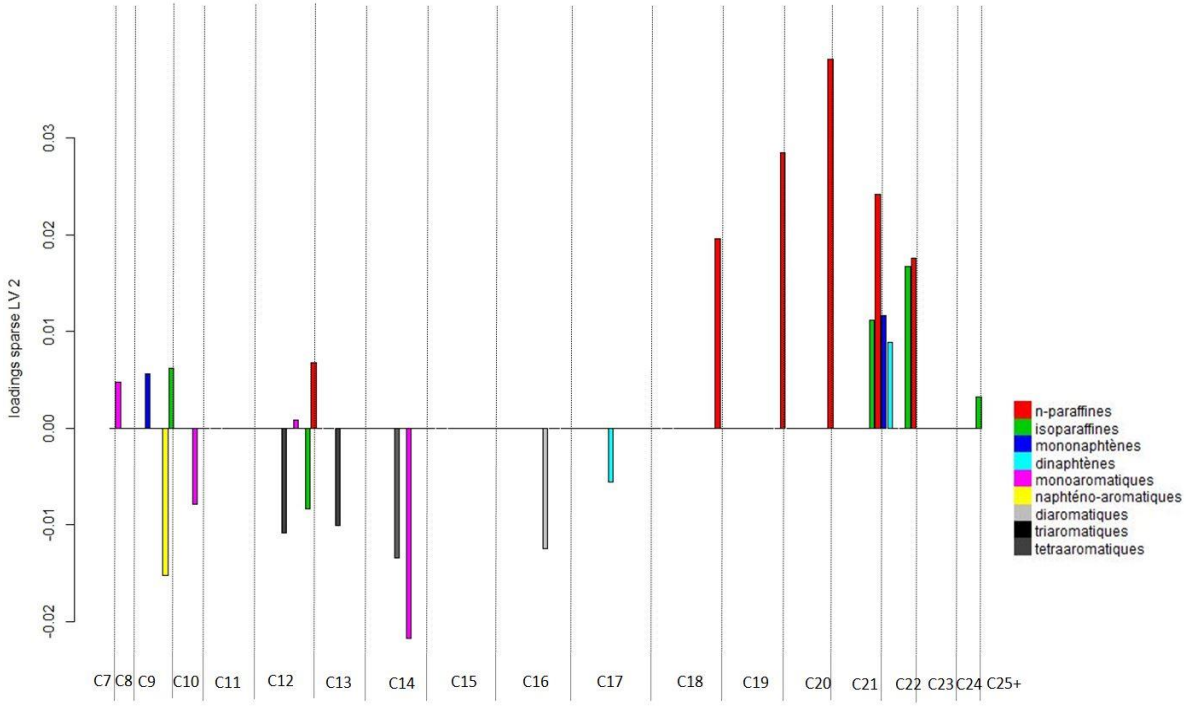
RESULTS – GCXGC DATA (3/3)

RESPONSIBLE
OIL AND GAS

Loading sparse-PLS LV1



Loading sparse-PLS LV2

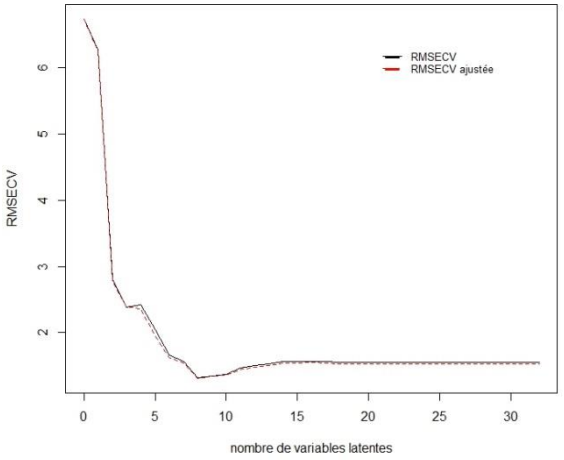


152 → 24 selected variables

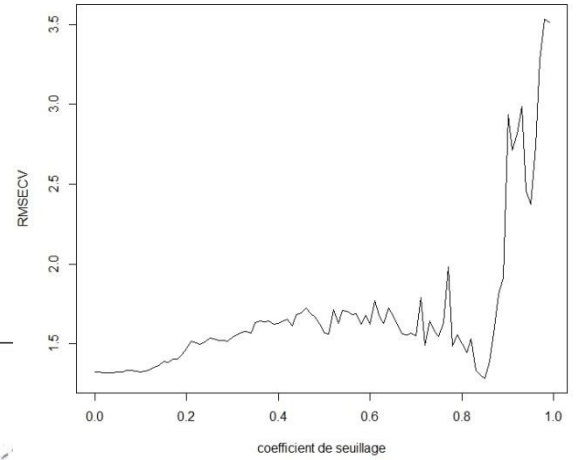
- N-paraffin from C18 to C21 ↗ **CP**
- Isoparaffin C21 and C22 ↘ **CP**

RESULTS – ^{13}C NMR DATA (1/2)

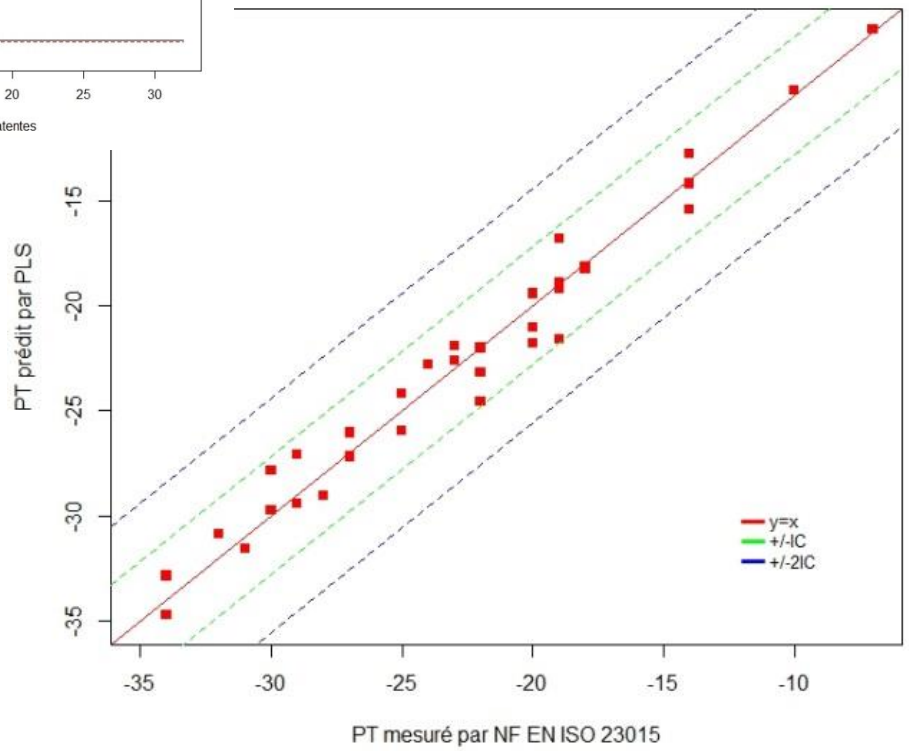
RESPONSIBLE
OIL AND GAS



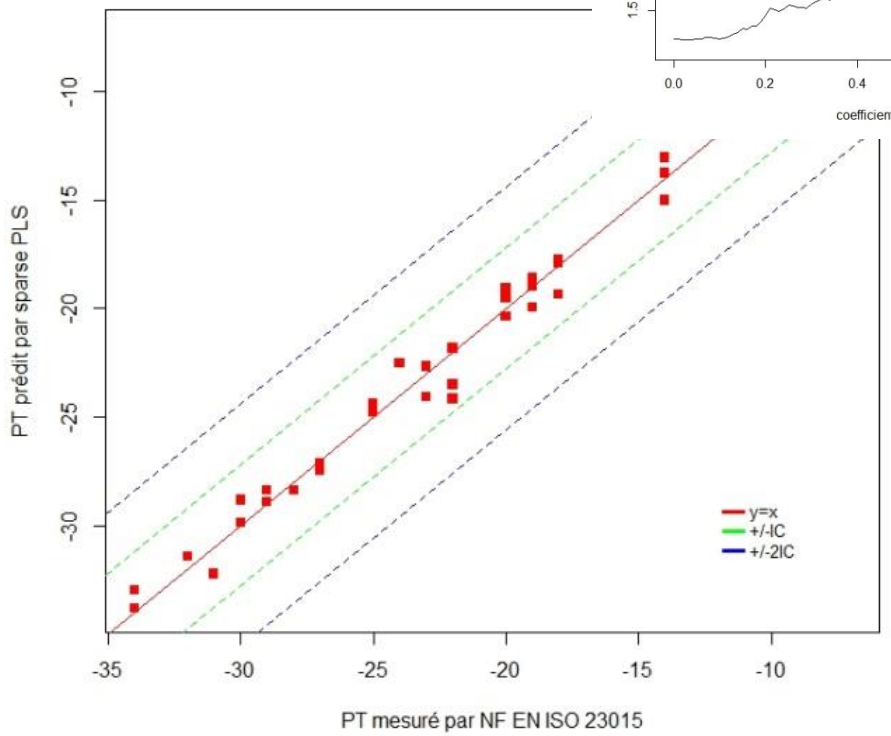
Model	LV	RMSECV (°C)	RMSEC (°C)	$\tau_{\pm IC}$ (%)
PLS	8	1,9	0,9	100
sparse PLS	8	1,3	0,8	100



PLS model



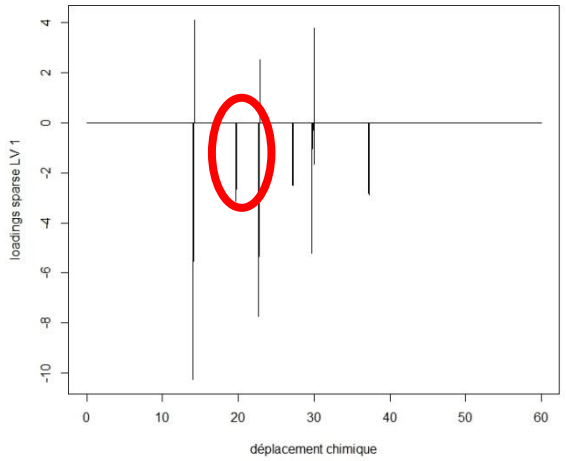
Sparse PLS model

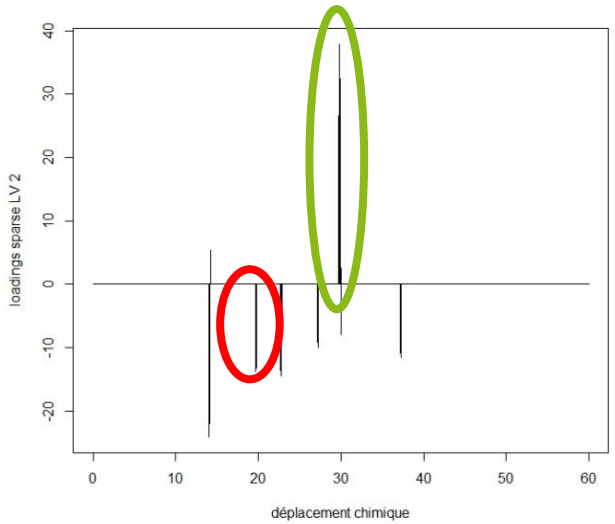


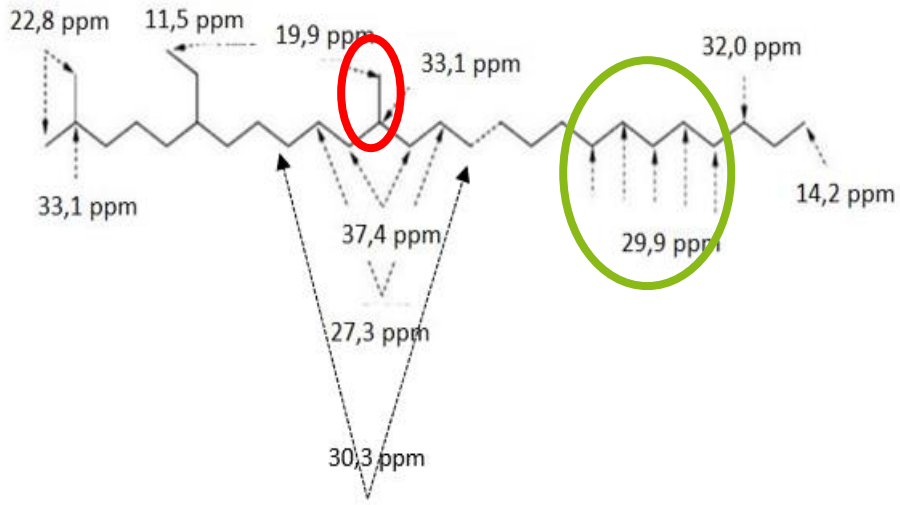
$\eta = 0{,}85$

RESULTS – ^{13}C NMR DATA (2/2)

RESPONSIBLE
OIL AND GAS







29.9 ppm ↔ N-paraffin: ↗ CP

19.9 ppm ↔ isoparaffin ↘ CP

Consistent with GCxGC

… More to investigate …

CONCLUSIONS

RESPONSIBLE
OIL AND GAS

- PLS model
  - Quite easy to develop
  - Loadings interpretation could be very long and complex
- Sparse-PLS
  - Quite easy to develop
  - Loadings easily interpretable
  - → Identification of the relevant variables to explain the cloud point

Thank you for your attention

Innovating for energy

Find us on:

www.ifpenergiesnouvelles.com

@IFPENinnovation