Application of kriging for modelling physicochemical properties in petroleum products

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

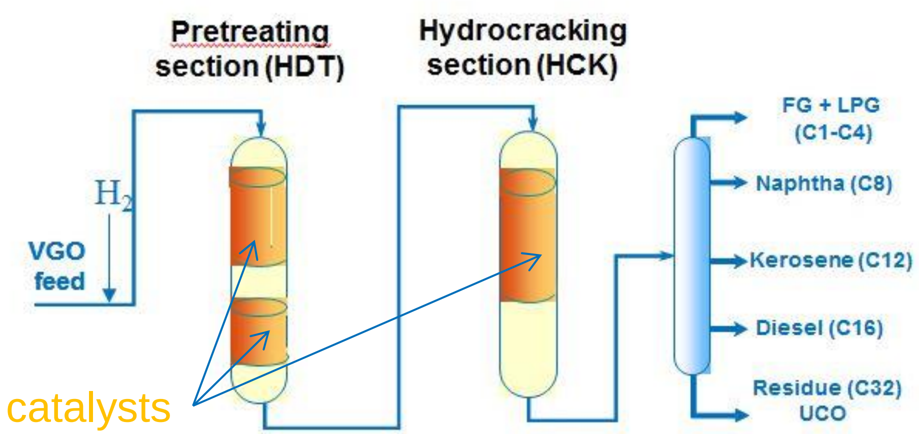
[1] IFP Energies Nouvelles Solaize
[2] LASIR CNRS Université Lille – Nord de France

CHIMIOMÉTRIE 2016

IFP Energies nouvelles

Hydrocraking in petroleum industry

- Global framework
  - Growth of demand for transport fuels
  - Insufficient proportions of light and middle distillates in crude oils
  - Need to develop refining processes as Hydrocracking



Feed characteristics
- Boiling range (350-550°C)
- Sulfur 0.3-3 % wt
- Nitrogen 600-4000 ppm wt

Experimental conditions
- Temperature (350-430°C)
- Total pressure 100-200 bar
- Contact time 0.2-3 hours
- Use of catalysts

Two targets:
- Maximize yield of middle distillates
- Upgrade UCO to high-quality lubricants

Background

- Hydrocracking process requirements
  - Set up pilot tests
  - Various physicochemical analysis
- Consequences
  - Time consuming
  - Costly
- Interests of modelling
  - Estimate products properties from petroleum analysis
  - Costs reduction
  - Short implementation time

➔

Current works

Objectives ➔ Improve prediction of petroleum products properties

Problems to solve
- Find relevant descriptors
- Define an adapted model frame

Present work ➔ Use of kriging model

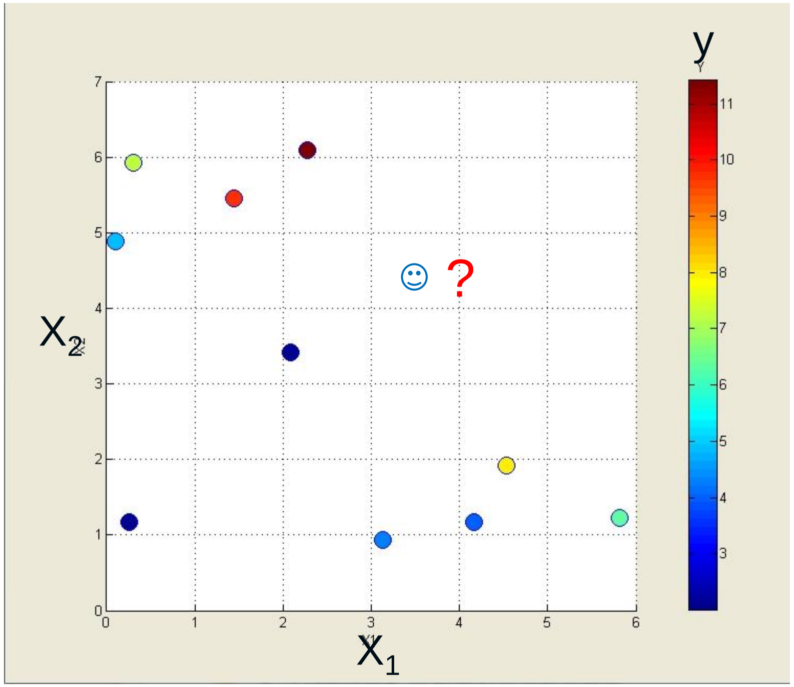
Outlines

- Theoretical aspect of kriging
  - Definition
  - Kriging predictor
  - Correlation model
- Chemometrics application
- Conclusion

Theoretical aspect

Definition

- Method of spatial interpolation that provides the best linear prediction of the intermediate values under suitable assumptions



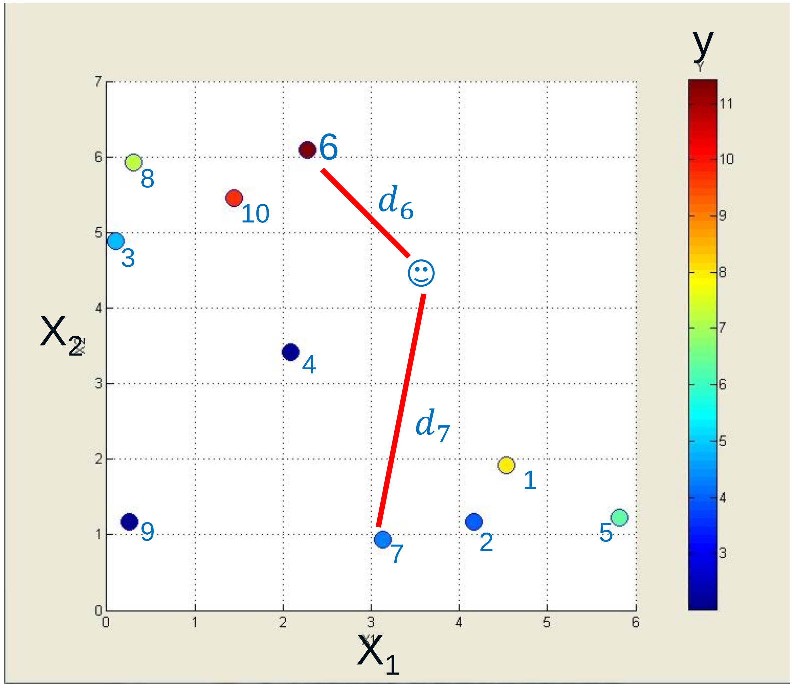
Interpolation problem

- X_1 and $X_2 \leftrightarrow$ explanatory variables (experimental conditions, feeds properties)
- $Y \leftrightarrow$ variable to predict (product property)
- Example of bivariate field ➔ $Y = f(X_1, X_2)$
- Colored points $\leftrightarrow$ observations sites
- Color $\leftrightarrow$ scale of Y values

How to estimate Y at the new location ?

Theoretical aspect

- **Kriging predictor**
  - Linear estimator ➔ $\hat{y}_{new} = p_1 y_1 + \cdots + p_{10} y_{10}$



- $p_i \leftrightarrow$ weight given to the i^{th} observation
- p_i depends on distance from i^{th} observation site to new location
- p_i globally decreases with the related distance
  $d_6 < d_7$ ➔ $p_7 < p_6$

Notes:

- Distance does not necessary refer to an euclidian distance
- Generally $d = f(|\Delta X_1|, |\Delta X_2|)$ where $\begin{cases} \Delta X_1 = X_{1,new} - X_{1,i} \\ \Delta X_2 = X_{2,new} - X_{2,i} \end{cases}$

How the weights are spatially distributed?

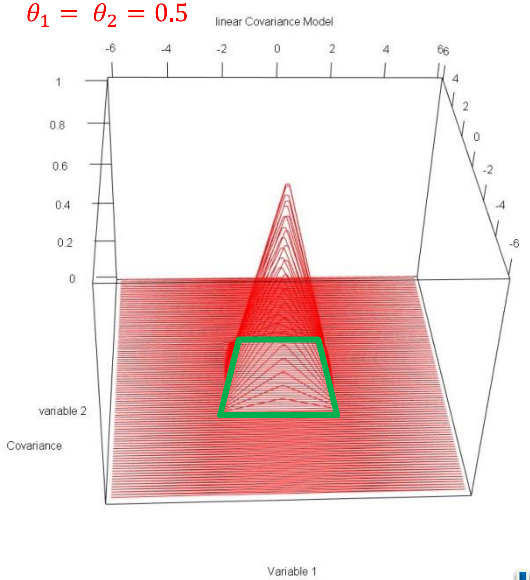
Theoretical aspect

Correlation model

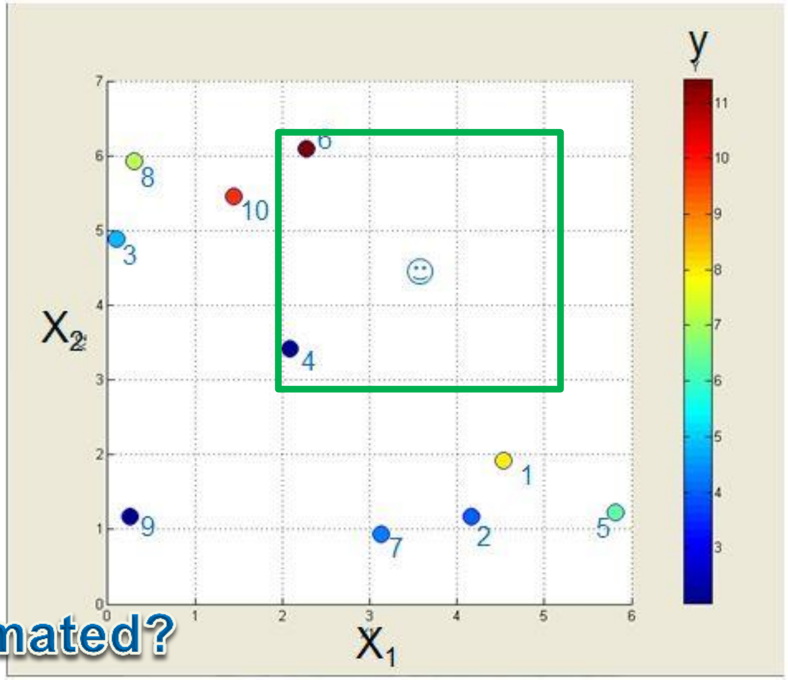
- Weights distribution is governed by a predefined correlation model
  - Linear:

$$\rho(y_i, y_j; \theta_1, \theta_2) = \begin{cases} (1-\theta_1|\Delta X_1|)(1-\theta_2|\Delta X_2|) \ if \ |\Delta X_1| < \frac{1}{\theta_1} \ and \ |\Delta X_2| < \frac{1}{\theta_2} \\ 0 \ else \end{cases} \ où \ \begin{cases} \Delta X_1 = X_{1,i} - X_{1,j} \\ \Delta X_2 = X_{2,i} - X_{2,j} \end{cases}$$

$\theta_1 = \theta_2 = 0.5$



Rectangular neighborhood



How p_i are estimated?

- Linear decrease of weights

$p_6, p_4 \neq 0$

$p_1, \ldots, p_3, p_5, p_7, \ldots, p_{10} = 0$

Theoretical aspect

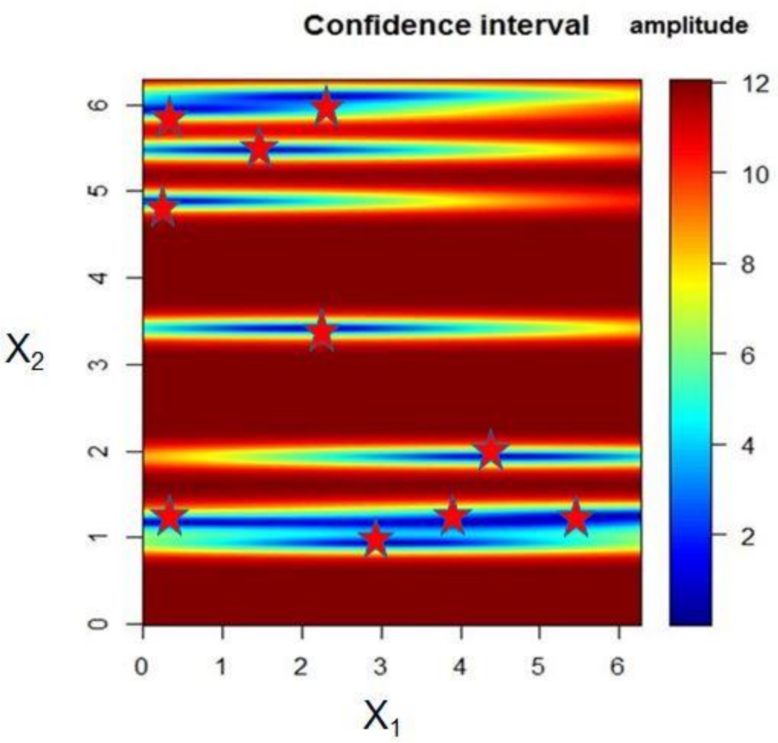
- Estimation of weights
  - Kriging tries to choose the optimal weights (p_i) that produce the minimum estimation error
  - Linearity
    - $\hat{y}_{new} = p_1 y_1 + \cdots + p_{10} y_{10}$
  - Optimal weights are those that produce unbiased estimates and give a minimum variance of Y value at new location

$$\Rightarrow (S)\begin{cases} p_1\rho(y_1, y_1) + \cdots + p_{10}\rho(y_1, y_{10}) = \rho(y_{new}, y_1) \\ \vdots \\ p_1\rho(y_{10}, y_1) + \cdots + p_{10}\rho(y_{10}, y_{10}) = \rho(y_{new}, y_{10}) \end{cases}$$

Estimate p_i → Solve linear System (S)

Theoretical aspect

- Limits of interpolation methods
  - Both following characteristics are requi
    - Uniform distribution throughout the stuc
    - Fair density of data
- Advantages of kriging
  - Kriging provides measures of predictio
  - Help to compensate for effects of clust
    - Assign individual points within a cluster less weight than an isolated data point
    - Have good performance in case of relevant irregularity
  - Incorporate prior information related to analytical uncertainties
    - Nugget effect



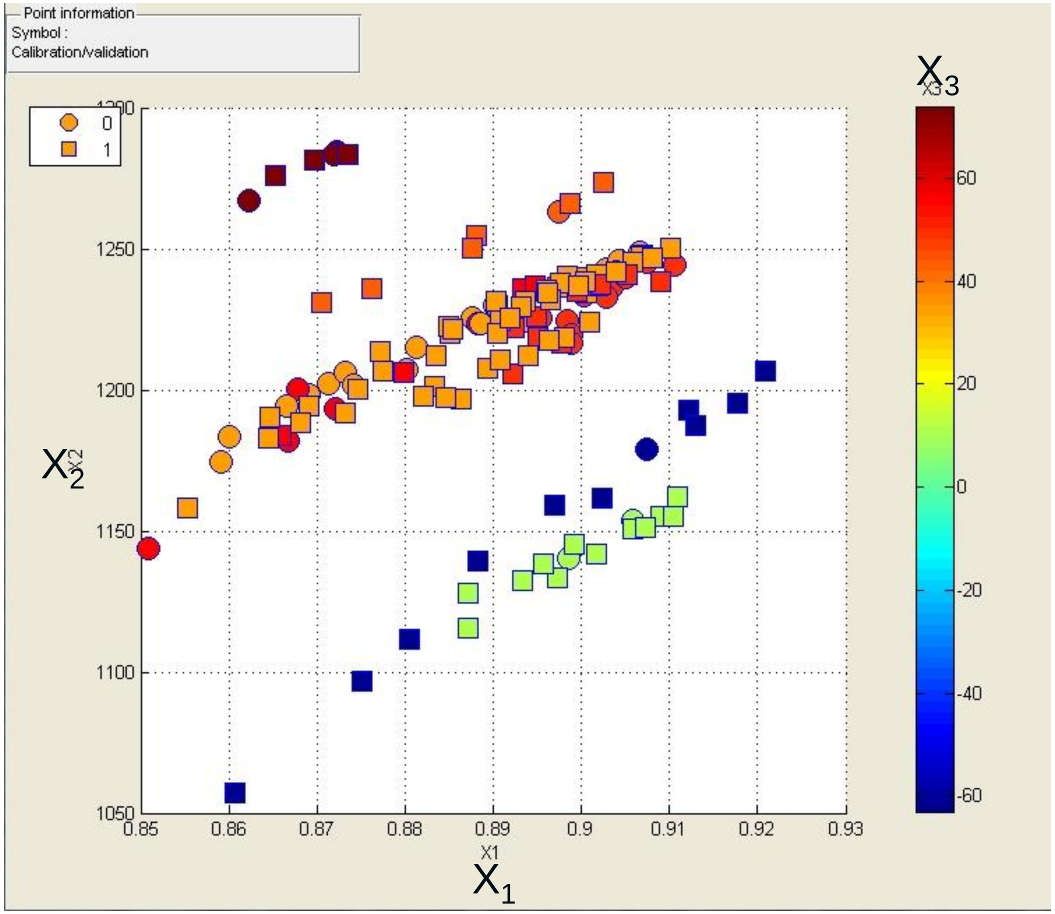
- Bibliography

  - P. Goovaerts, *Geostatistics for natural ressources*. New York: Oxford University Press, 1997
  - E. H. Isaaks and R. M. Srivastava, *Applied geostatistics*. New York: Oxford University Press, 1989

Outlines

- Theoretical aspect of kriging
- Chemometrics applications
  - Database description
  - Results and discussion
- Conclusion

First application



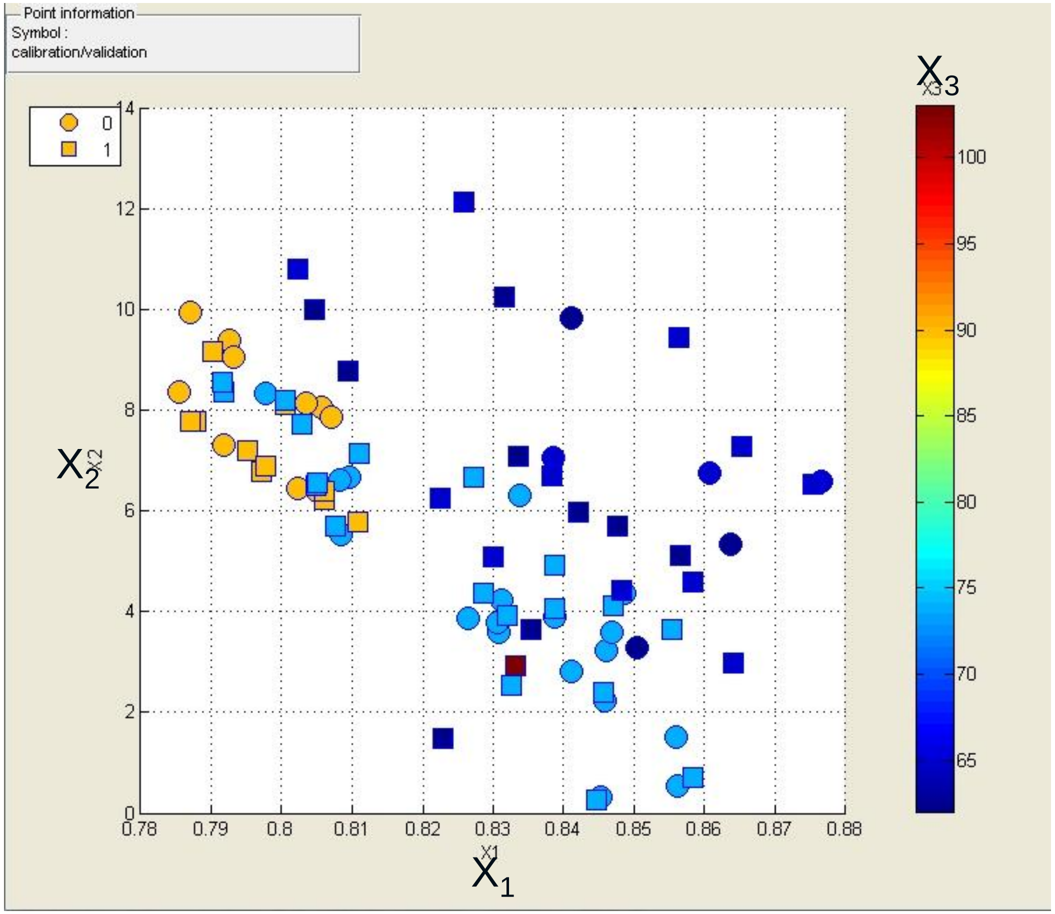
<u>Database used</u>

- Calibration set : 95 samples
- Validation set : 40 samples
- Input variables from physicochemical petroleum analysis : 6
- P : property to predict

<u>Data repartition</u>

- Squares: calibration points
- Circles: validation points
- Calibration points were selected using space filling algorithm

Second application



<u>Database used</u>

- Calibration set : 50 samples
- Validation set : 32 samples
- Input variables from physicochemical petroleum analysis : 7
- P : property to predict

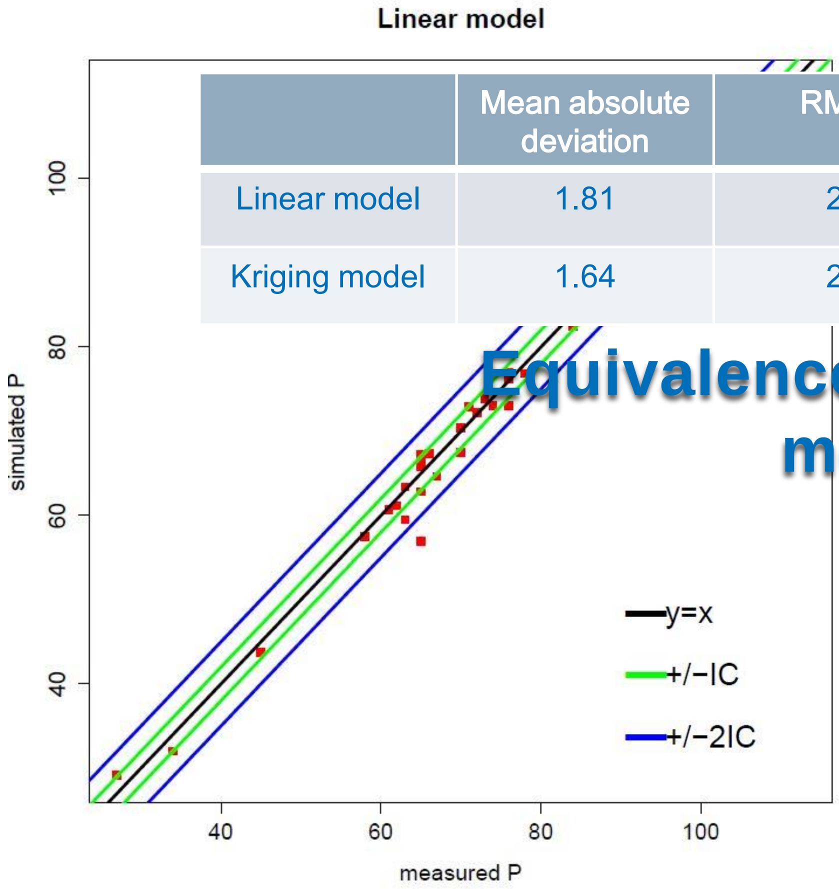
<u>Data repartition</u>

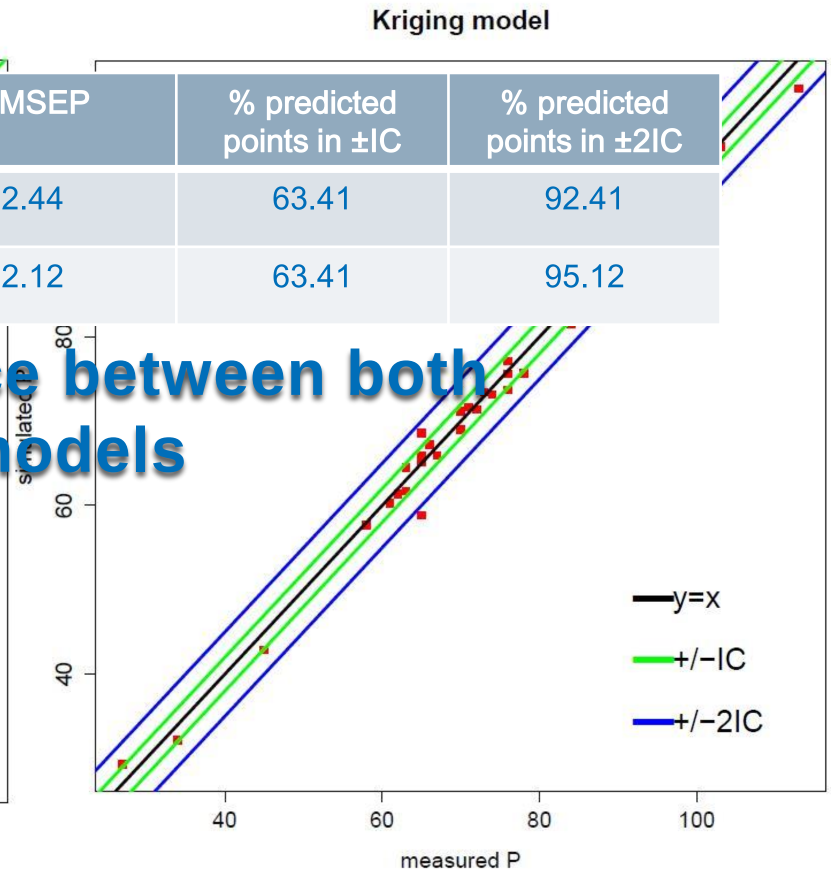
- Squares: calibration points
- Circles: validation points
- Calibration points were selected using space filling algorithm

Chemometrics application

- **First application results**

	Mean absolute deviation	RMSEP	% predicted points in ±IC	% predicted points in ±2IC
Linear model	1.81	2.44	63.41	92.41
Kriging model	1.64	2.12	63.41	95.12



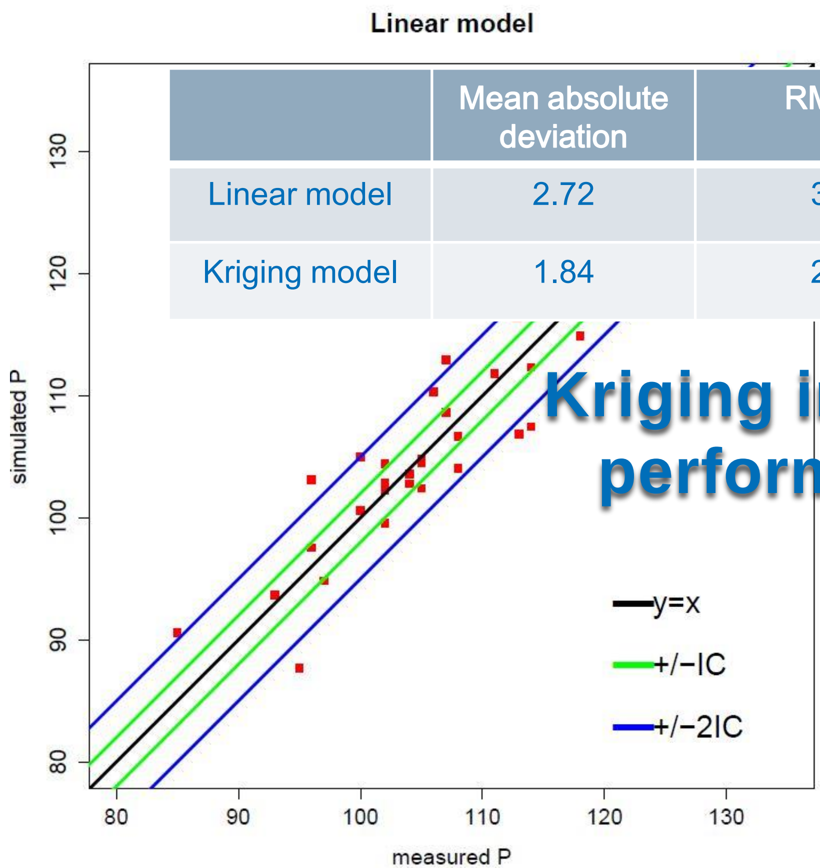


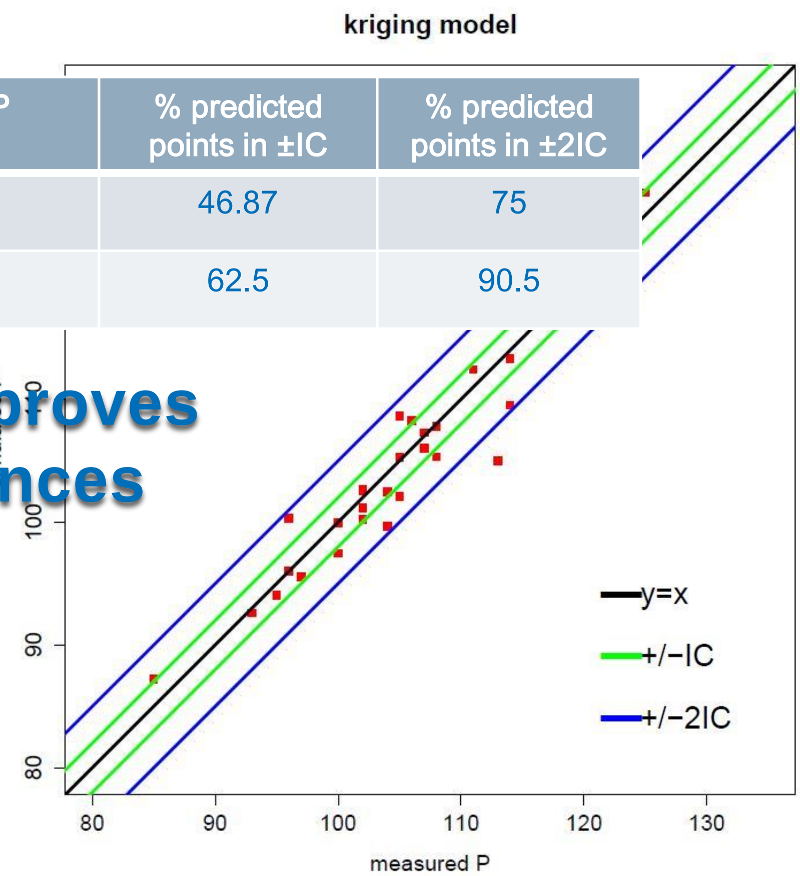
Equivalence between both models

Chemometrics application

- Second application results

	Mean absolute deviation	RMSEP	% predicted points in ±IC	% predicted points in ±2IC
Linear model	2.72	3.52	46.87	75
Kriging model	1.84	2.57	62.5	90.5





Kriging improves performances

Conclusions and perspectives

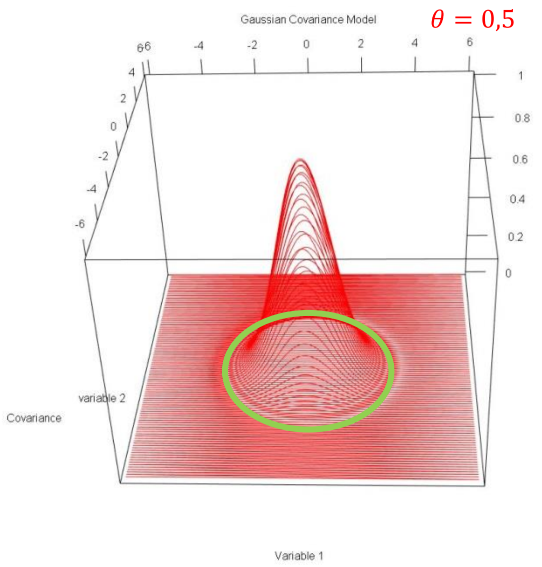
- Interest of kriging
  - Adaptability to linear and nonlinear situation
    - Good performance was obtained in modelling nonlinear property
    - No predefined model structure
  - Possibility to introduce more prior informations about data
    - Analytical uncertainties inclusion
    - Choose appropriate correlation model among various types
- Implementation constraints
  - Have data locations uniformly distributed around study area
  - More important optimization time than classical chemometrics methods
- Available R package « DiceKriging » April 24, 2015
  - Free
  - Many correlation models are available (gaussian, exponential,…)
- Future works
  - Use of Kriging in high multivariate regression

QUESTIONS

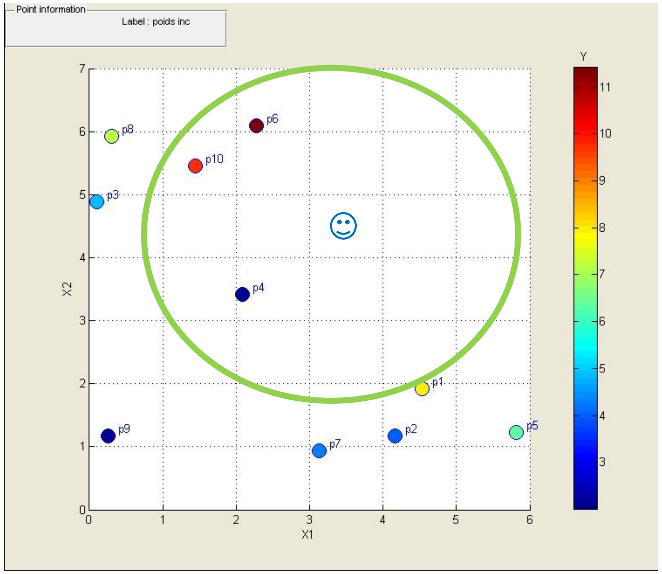
Contact: jean-jerome.da-costa-soares@ifp.fr

Annexe: Gaussian correlation model

- **Gaussian** ➔ $\rho(y_i, y_j, \theta) = e^{-\theta[(\Delta X_1)^2 + (\Delta X_2)^2]} = e^{-\theta d^2(i,j)}$
  - $d(i,j) \leftrightarrow$ euclidian distance



Ellipsoïdal neighbourhood



How p_i are numerically estimated?

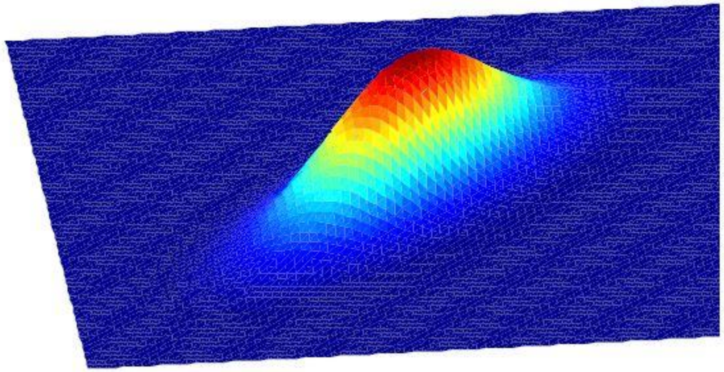
- Exponential decrease of weight
- euclidian distance

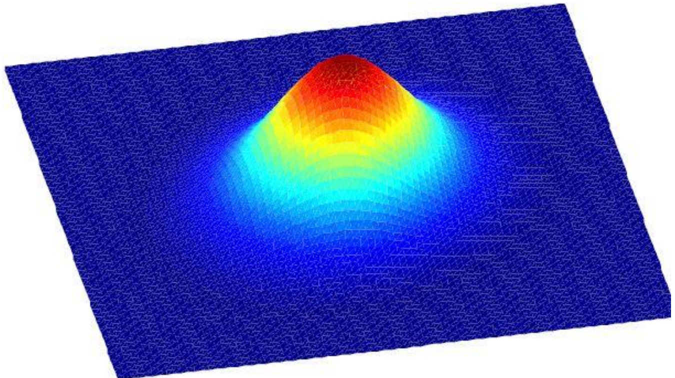
$p_i \neq 0 \; \forall \, i$

$|p_4| \gg |p_3|$

Annexe: Gaussian stochastic process

- **Hypothesis about errors 1**
  - $(\varepsilon_1, \ldots, \varepsilon_n) \sim$ Gaussian stochastic process
  - Let $\mathbf{X} = (X_1, \ldots, X_n)$ a random vector
    - X is said k-variate normally distributed if every linear combination of its k components has a univariate normal distribution.
    - $f_X(x) = \frac{1}{(2\pi)^{n/2}\det(\Sigma_X)^{1/2}} \exp(-\frac{1}{2}(x - m_X)^T {\Sigma_X}^{-1}(x - m_X))$
  - Example of distribution in two-dimensional case for different coefficient of correlation
    - **Left** : $\rho = 0.8$; **Right** $\rho = 0.2$.

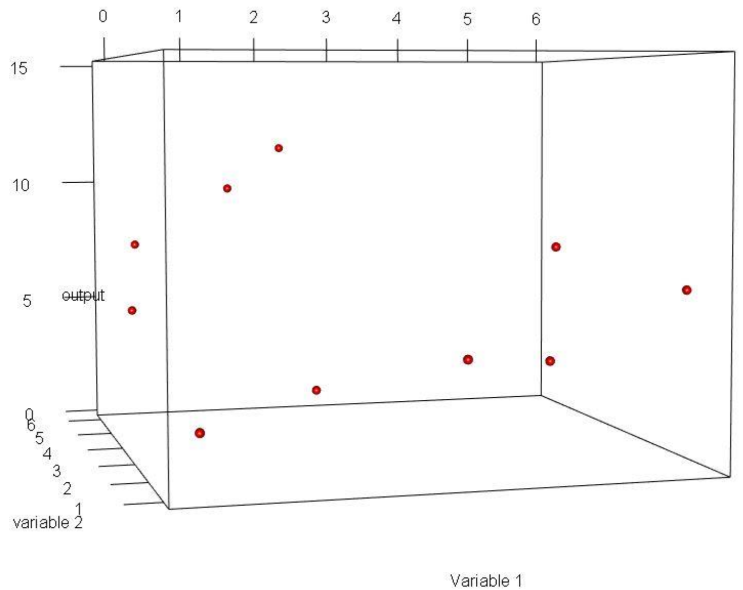


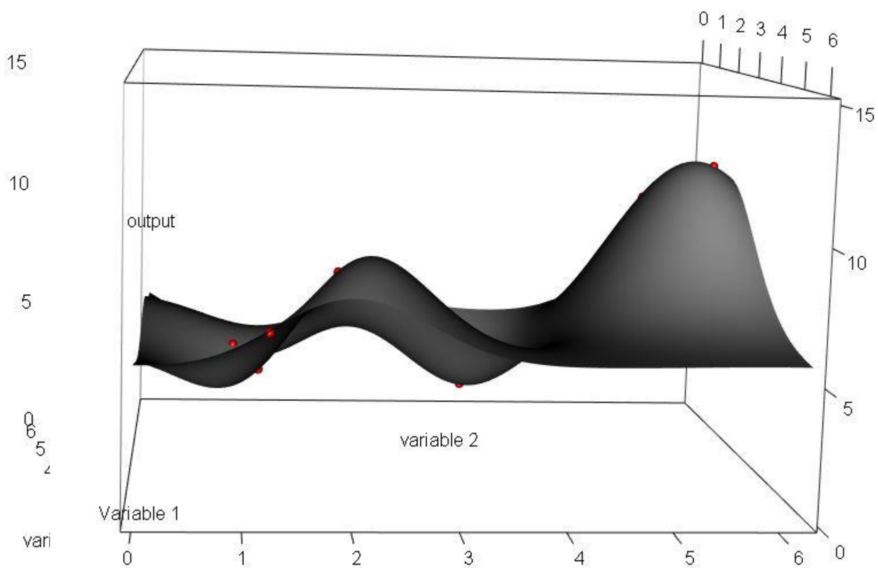


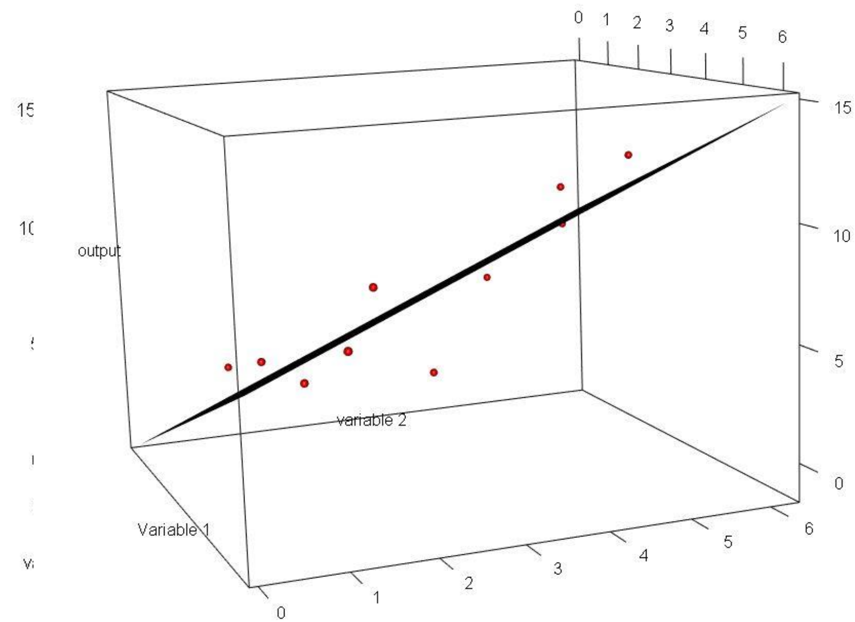
SK vs MLR

Kriging Model

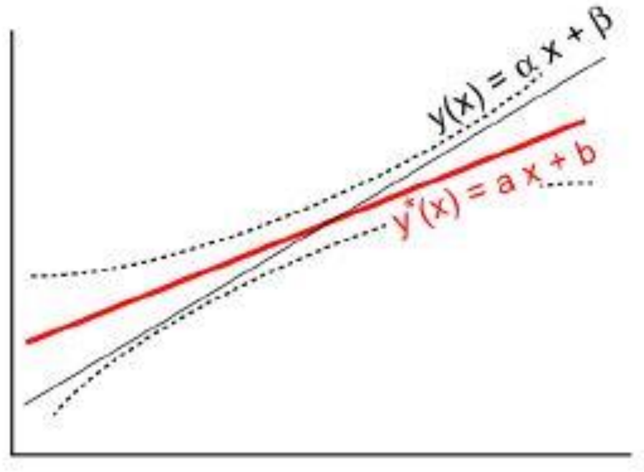
- statistical modelli … odelling
  - $y_i = m + \varepsilon_i, i =$ … $+ \varepsilon_i, i = 1, \dots, n$
    - $(\varepsilon_1, \dots, \varepsilon_n) \sim$ Gau … $\sigma^2), i.i.d.$







Annexe: confidence interval of linear model



Variation hyperbolique de l'amplitude de l'intervalle de confiance

$$y(x) \in ax + b \pm t_{\frac{\alpha}{2}} \times \hat{\sigma} \sqrt{\frac{(x-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \frac{1}{n}}$$

Annexe: confidence interval amplitude