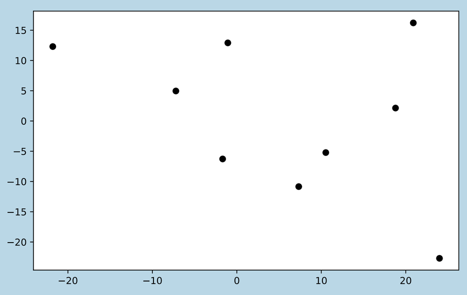
Gaussian Processes Indexed by Clouds of Points: a study Babacar SOW (EMSE, LIMOS), Rodolphe LE RICHE (CNRS, LIMOS)

Babacar Sow, Rodolphe Le Riche, Julien Pelamatti, Sanaa Zannane, Merlin Keller

Gaussian Processes Indexed by Clouds of Points: a study

Babacar SOW (EMSE, LIMOS), Rodolphe LE RICHE (CNRS, LIMOS)
Julien PELAMATTI(EDF R&D), Sanaa ZANNANE (EDF R&D), Merlin KELLER (EDF R&D)
Universités: Ecole Nationale Supérieure des Mines de Saint-Etienne(EMSE), Université Clermont Auvergne(UCA)
Laboratoire: LIMOS **Email of contact:** babacar.sow@emse.fr

Context And Problematic



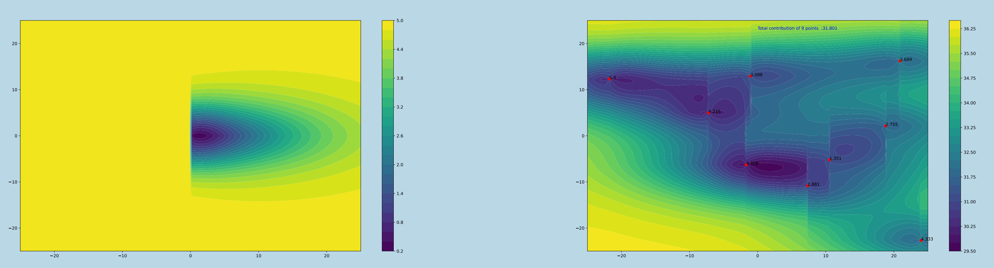
- Metamodel a function over clouds of points using Gaussian process.
- A cloud is a set of points invariant under permutation $\{x_1, ..., x_n\}$ with $x_i \in \mathbf{R}^{dim}$

Test Function

The following test function mimicks a wind-farm production:

$$F(\{x_1, ..., x_n\}) = \sum_{i=1}^{n} \sum_{\substack{j \\ x_{j,1} \leq x_{i,1}}} f_p(x_j, x_i) f_0(x_i)$$

where $f_p(x_j, x_i)$ expresses the energy loss over x_i that is caused by x_j and f_0 is a constant.



Kernels

Substitution kernel with MMD

- We want to construct a kernel between two clouds of the form $K(X,Y) = \sigma^2 \exp(-\frac{d^2(X,Y)}{2\theta^2})$ where d is an Hilbertian [2] distance.
- For two clouds $X = \{x_1, ..x_n\}$ and $Y = \{y_1, ..., y_m\}$, $P_X = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}$ and $P_Y = \frac{1}{m}\sum_{j=1}^{m} \delta_{y_j}$ are the respective associated empirical uniform distributions.
- There exists a Reproducing Kernel Hilbert Space, $\mathcal{H}$ with a characteristic kernel such as $k_{\mathcal{H}}(x, .) = exp(-\frac{||x-.||^2}{2\theta^2})$.
- The characteristic nature guarantees the injectivity of the embedding map [1]: $P_X \longmapsto \mu_X = \int P_X(x) k_{\mathcal{H}}(x, .)\mathrm{d}x$.
- $MMD^2(P_X, P_Y) = ||\mu_X - \mu_Y||^2_{\mathcal{H}}$
- For any kernel $k_{\mathcal{H}}$ of the RKHS, the uniform empirical laws gives $MMD^2(P_X, P_Y) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} k_{\mathcal{H}}(x_i, x_j) + \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m} k_{\mathcal{H}}(y_i, y_j) - 2\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} k_{\mathcal{H}}(x_i, y_j)$
- The correlation kernel $K_{sub\_mmd}(X, Y) = \sigma exp(-\frac{||\mu_X - \mu_Y||^2_{\mathcal{H}}}{2\theta^2})$ is symmetric and definite positive.

References

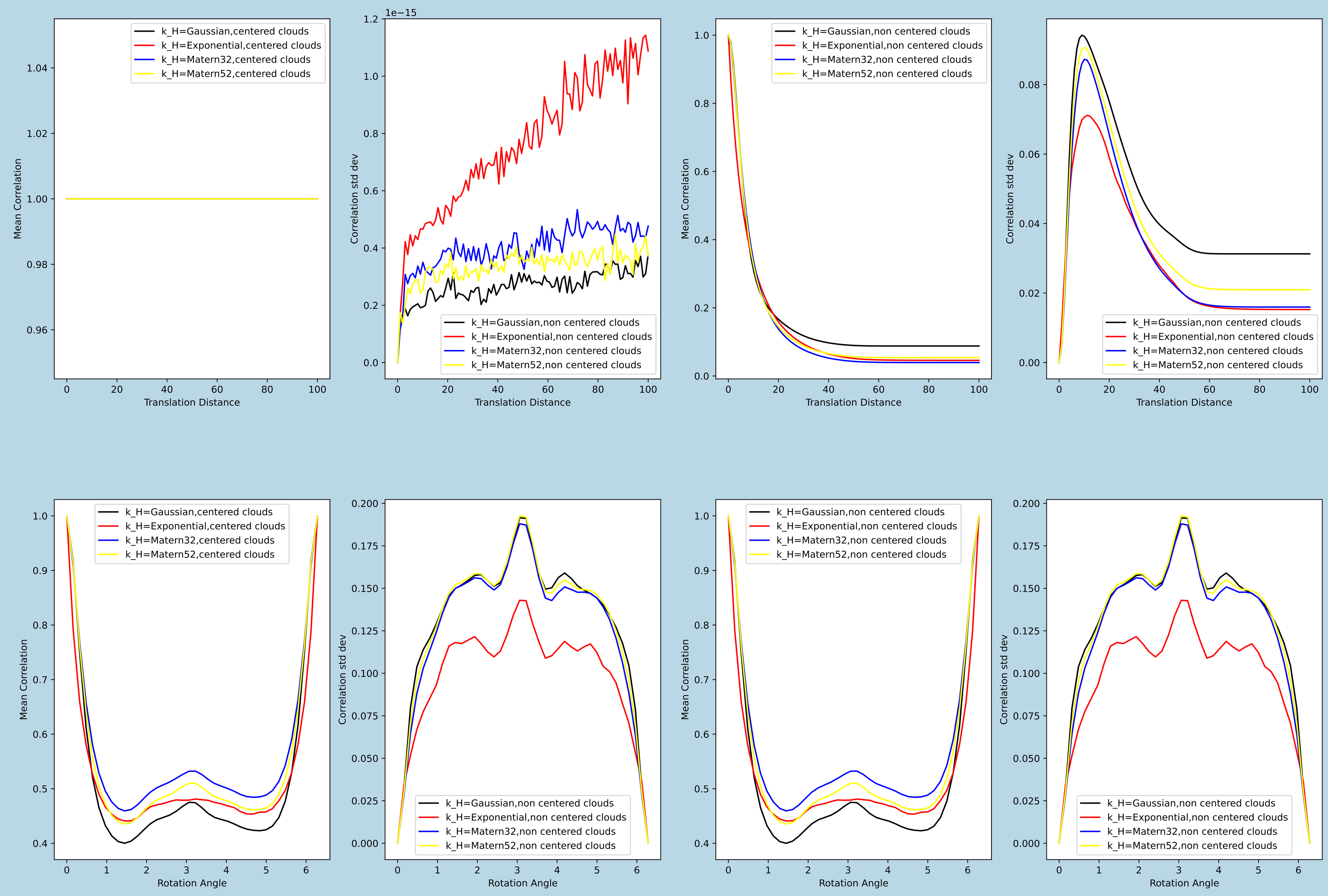
[1] Muandet, Krikamol and Fukumizu, Kenji and Sriperumbudur, Bharath and Schölkopf, Bernhard and others: *Kernel mean embedding of distributions: A review and beyond*, Foundations and Trends® in Machine Learning

[2] Haasdonk, Bernard and Bahlmann, Claus: *Learning with distance substitution kernels*, Springer

Geometrical Properties of the kernels

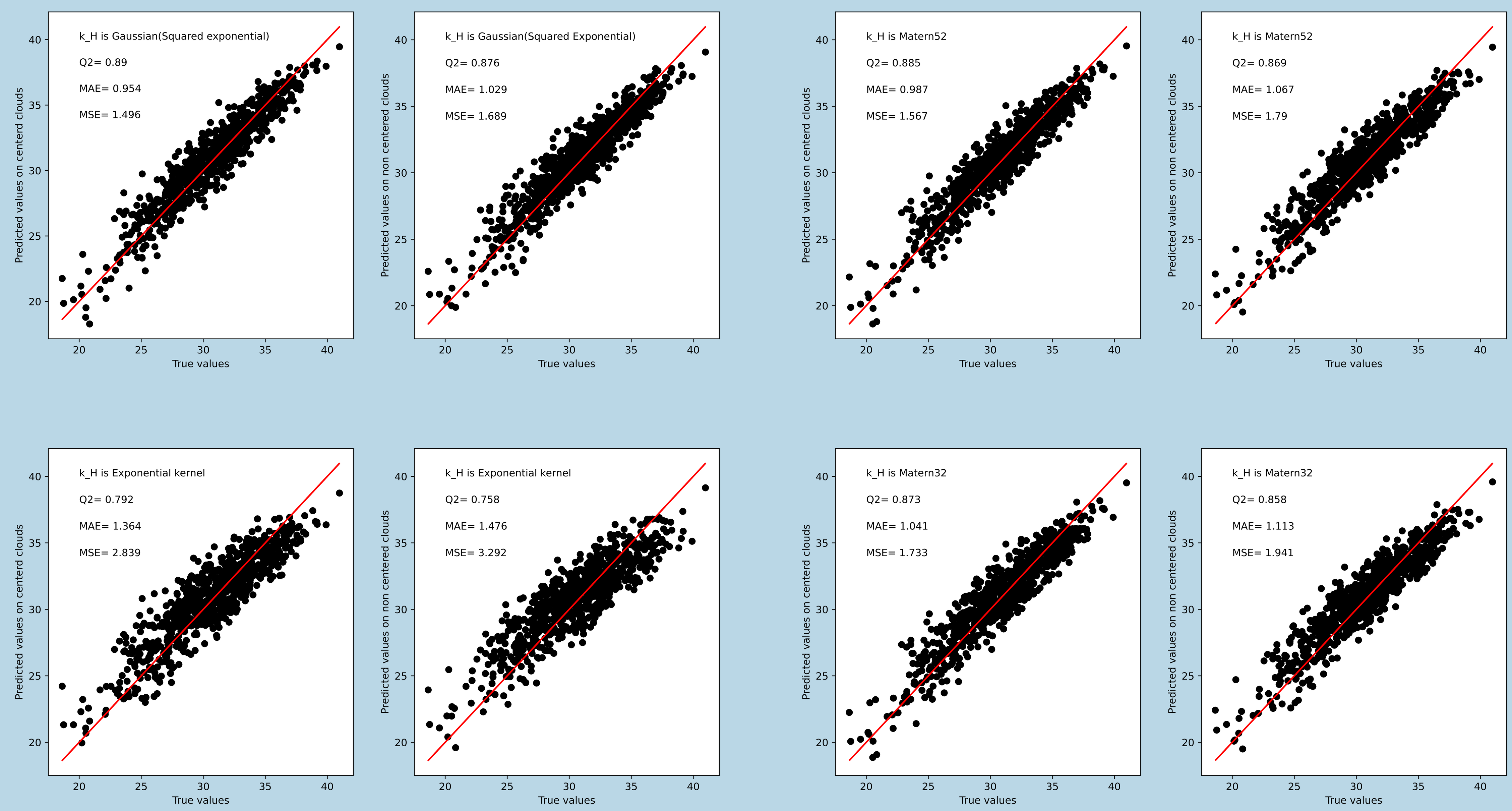
- Below is represented the correlation between a cloud and its image by a geometric transformation. Considered transformations are rotations and translations.
- We compare two scenarios: centered clouds and non-centered ones.
- The different kernels of the Hilbertian Space are the Exponential, the Gaussian(Squared Exponential), the Matern32 and the Matern52.

centered cloud **non-centered cloud**



Prediction Results on the analytical Function F

- We metamodel the wind-farm proxy function F with a Gaussian process of kernel K_{sub_mmd}
- We consider a set of 1000 clouds of 10 points each.
- Each point of a cloud is drawn uniformly in a square.
- The kernel parameters are learned using 200 clouds by maximizing log-likelihood with BFGS.
- On each plot, we represent predicted values vs. true ones on the remaining clouds , obtained with the different kernels.
- The corresponding Q2, MAE and MSE are also displayed.



- The best results are obtained when k_H is the Squared Exponential and clouds are centered. This can be explained by the fact that the quadratic modified distance with two length scales of this k_H complies better with the geometric properties of F.