



HAL
open science

Bayesian multi-objective optimization with noisy evaluations

Bruno Barracosa, Julien Bect, H elo ise Dutrieux Baraffe, Juliette Morin, Josselin Fournel, Emmanuel Vazquez

► To cite this version:

Bruno Barracosa, Julien Bect, H elo ise Dutrieux Baraffe, Juliette Morin, Josselin Fournel, et al.. Bayesian multi-objective optimization with noisy evaluations. MASCOT PhD student 2020 Meeting, Sep 2020, Grenoble, France. . <hal-03022267>

HAL Id: hal-03022267

<https://centralesupelec.hal.science/hal-03022267v1>

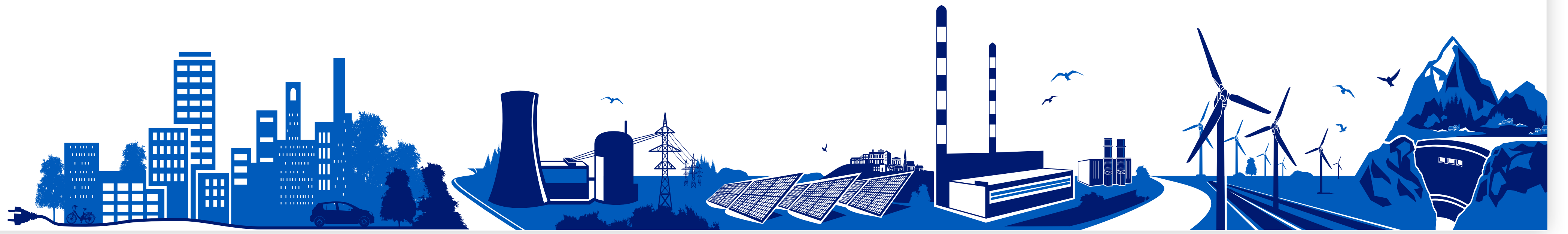
Submitted on 24 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destin e au d p t et   la diffusion de documents scientifiques de niveau recherche, publi s ou non,  manant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv s.



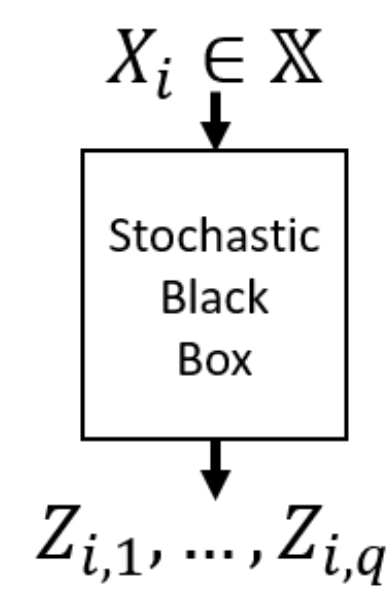
Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License



BAYESIAN MULTI-OBJECTIVE OPTIMIZATION WITH NOISY EVALUATIONS

1. CONTEXT

- Multi-objective optimization of the parameters of a planning strategy for the multi-year planning of the electricity distribution grid [DUT15]
- A stochastic black box provides **noisy evaluation results** of the objective functions f_1, \dots, f_q defined on a discrete search domain $\mathbb{X} \subset \mathbb{R}^d$
- Previous n evaluations at $X = (X_1, \dots, X_n)$ assumed:
 - $Z_{i,1} = f_1(X_i) + \varepsilon_{i,1}, \dots, Z_{i,q} = f_q(X_i) + \varepsilon_{i,q}$ where $\varepsilon_{i,j}$ s are zero-mean random variables



2. OBJECTIVE

Estimate the **Pareto-optimal solutions** (or Pareto set Γ) of the problem:

$$x^* = \operatorname{argmin}_{x \in \mathbb{X}} f_1(x), \dots, f_q(x)$$

Defined as:

$$\Gamma = \{x \in \mathbb{X} : \nexists x' \in \mathbb{X} \text{ such that } f(x') < f(x)\}$$

where $<$ stands for the Pareto domination rule:

$$y = (y_1, \dots, y_q) < y' = (y'_1, \dots, y'_q) \Leftrightarrow \begin{cases} \forall i \leq q, y_i \leq y'_i \\ \exists j \leq q, y_j < y'_j \end{cases}$$

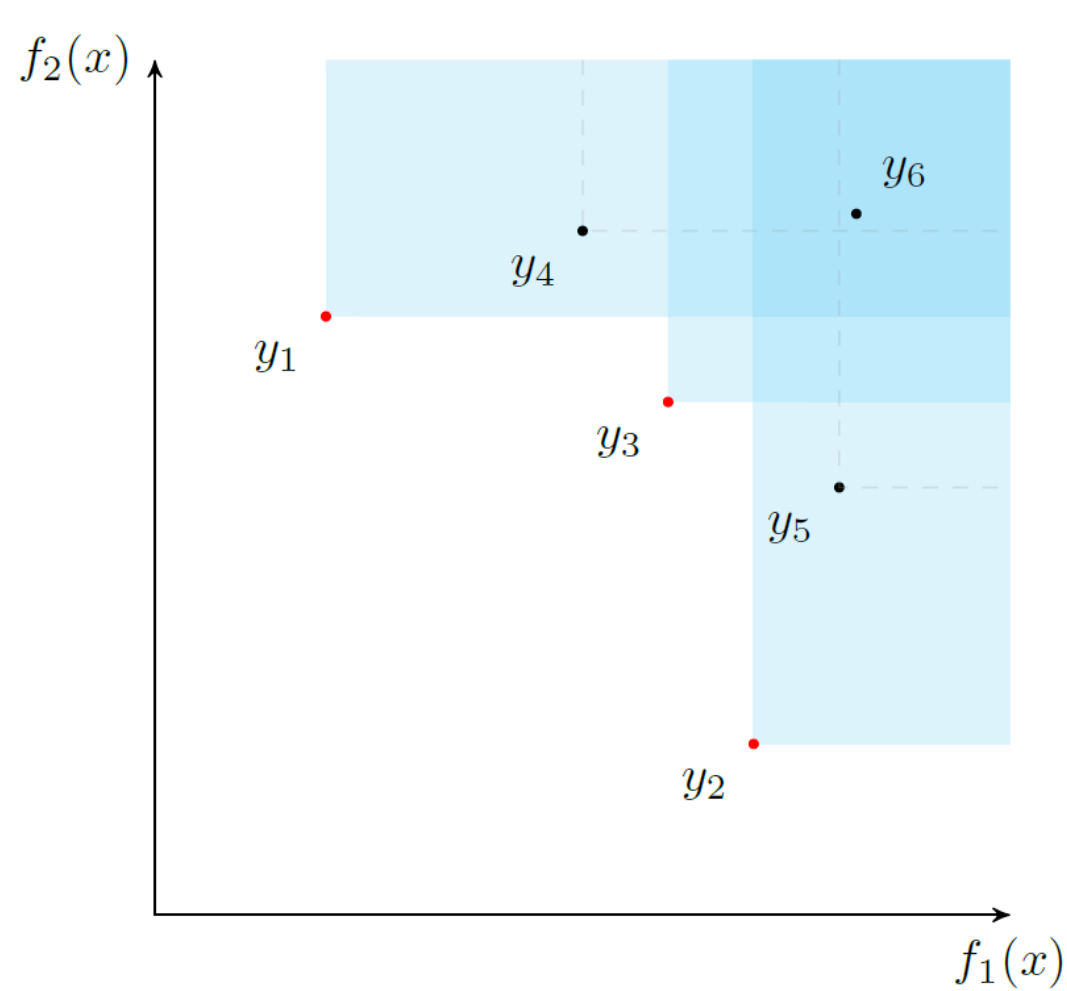
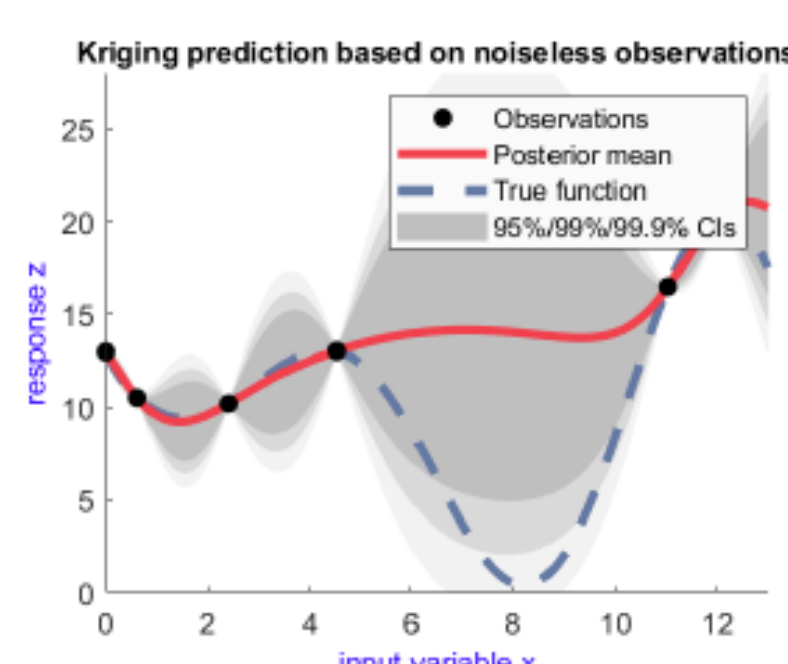


Illustration of the Pareto domination rule

- y_1, y_2, y_3 are non-dominated points
- y_4 is only dominated by y_1
- y_5 is only dominated by y_2
- y_6 is dominated by all other points

3. BAYESIAN OPTIMIZATION

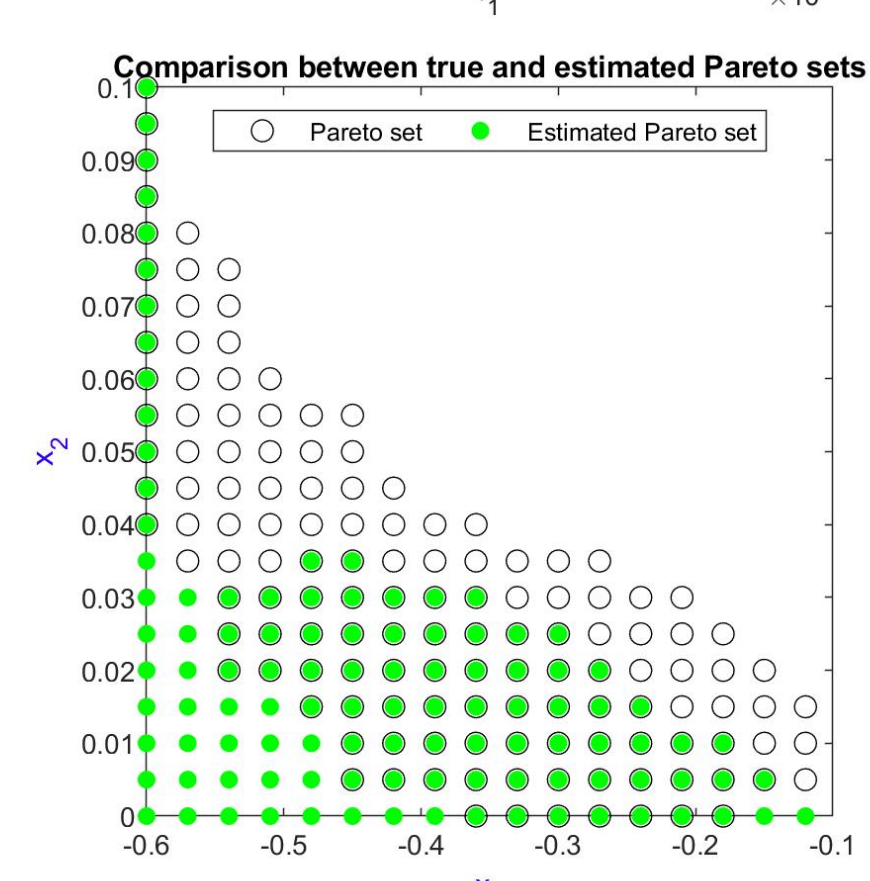
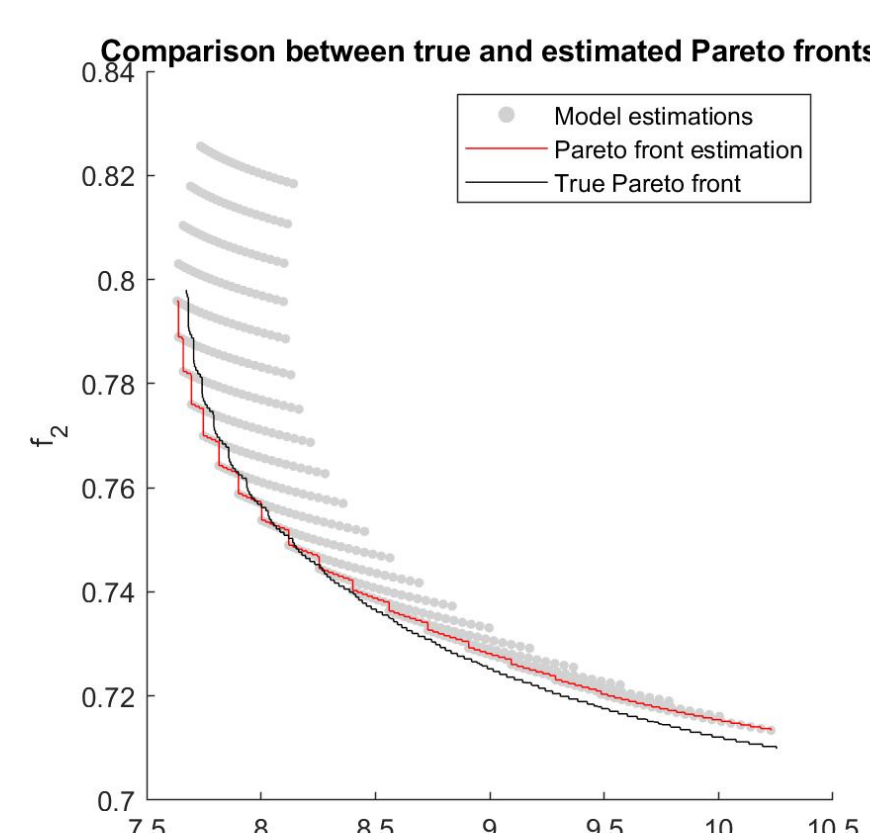
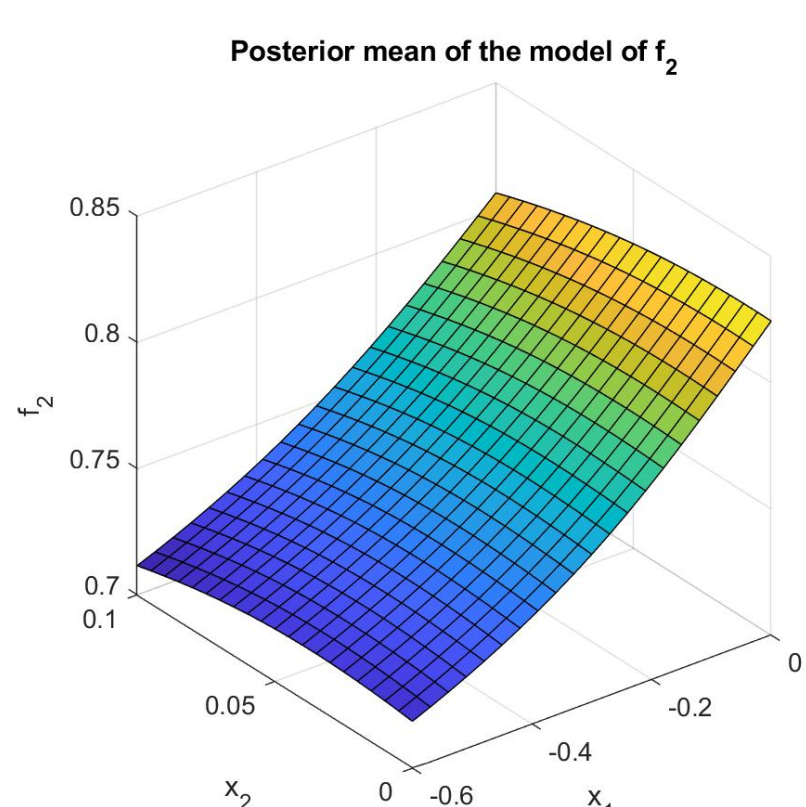
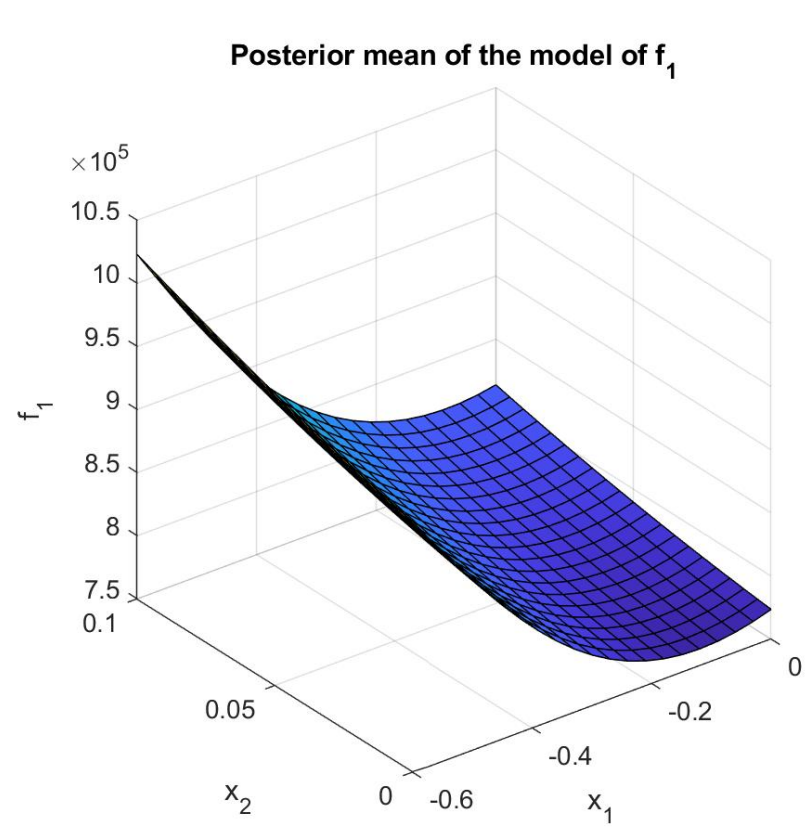
- Define a **probabilistic model** for each f conditional on previous observations
- Use a **sampling criterion** to select new evaluation points



Provides estimate of f and a measure of uncertainty of the estimation

PARETO-OPTIMAL ESTIMATES

Built from the estimates of f



PROPOSED APPROACH

- Replace the multi-objective problem by the minimization of a single **augmented Tchebycheff function** [KNO06]:

$$\tilde{f}(x) = \max_j [\omega_j f_j(x)] + \rho \sum_j \omega_j f_j(x), \quad \sum_j \omega_j = 1, \rho > 0$$

- At each iteration, generate random weights ω_j and apply this function to the n previous observations:

$$\tilde{Z}_i = \max_j [\omega_j Z_{i,j}] + \rho \sum_j \omega_j Z_{i,j}, \quad i = 1, \dots, n$$

- Assume a homoscedastic Gaussian noise model and fit to $\tilde{Z} = (\tilde{Z}_1, \dots, \tilde{Z}_n)$ a Gaussian Process model $\tilde{\xi}_n$ with parameters estimated by maximum likelihood
- We use the **Knowledge Gradient (KG)** criterion [FRA09] to select new point X_{n+1} based on previous observations. The idea is to identify a point that is expected to reduce the minimum of the posterior mean of $\tilde{\xi}$:

$$X_{n+1} = \operatorname{argmax}_{x \in \mathbb{X}} KG(x)$$

with:

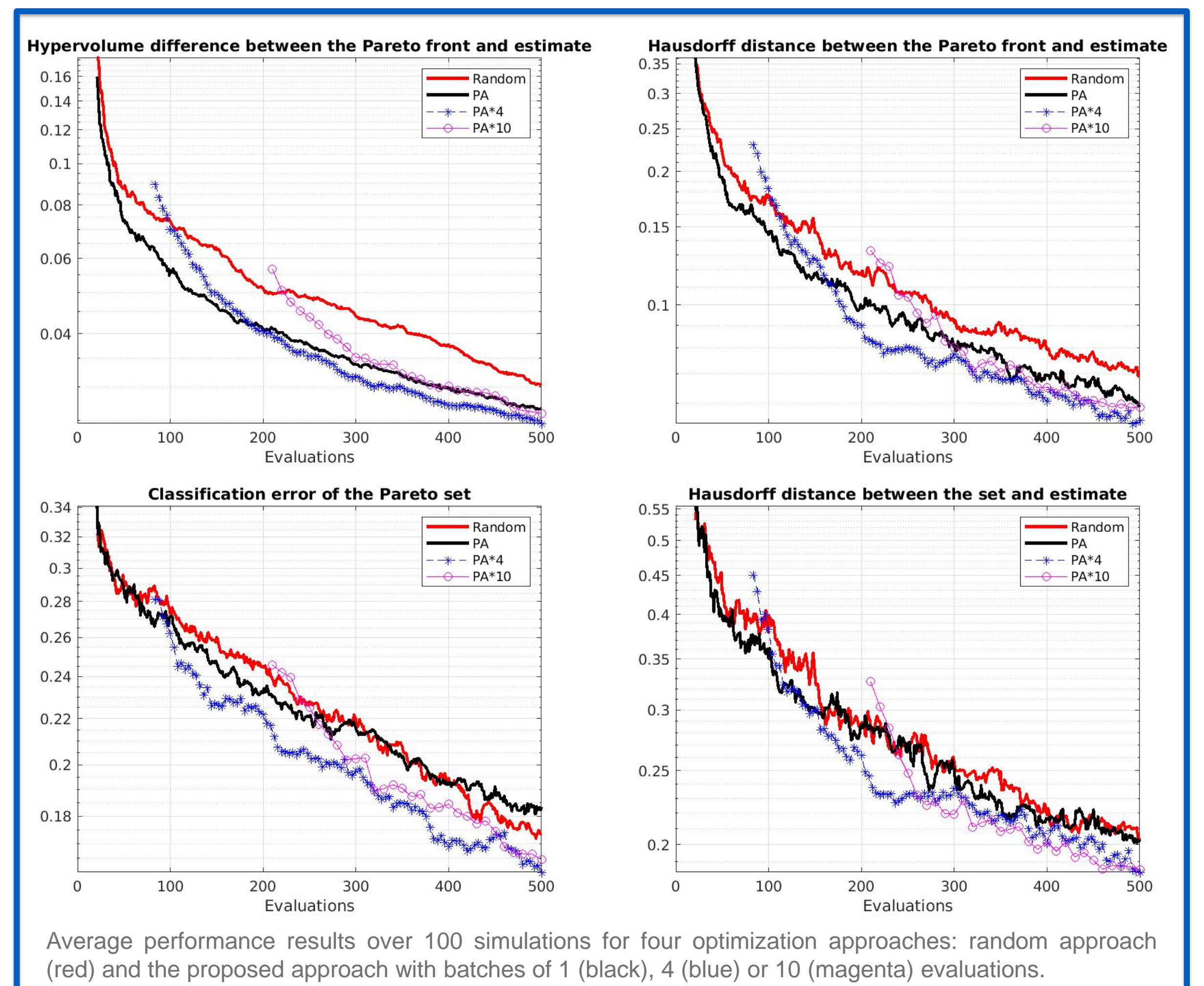
$$KG(x) = \min_{x' \in \mathbb{X}} \mathbb{E}[\tilde{\xi}_n(x')] - \mathbb{E} \left[\min_{x' \in \mathbb{X}} \mathbb{E}[\tilde{\xi}_n(x') | \tilde{Z}_{n,x}] \right]$$

where $\tilde{Z}_{n,x}$ denotes a new observation of $\tilde{\xi}_n$ at x

- Update the model with new observation and iterate until stopping criterion is met

4. NUMERICAL EXPERIMENTS

- Compare the proposed approach (PA) with random selection of points in a bi-dimensional bi-objective problem
- Compare use of batches of 1, 4 or 10 evaluations



5. OPEN QUESTIONS

- 'Ideal' batch of evaluations?
- Performance comparison of the proposed approach to other methods in the literature?

REFERENCES

- [DUT15] Dutrieux, H. (2015). "Méthodes pour la planification pluriannuelle des réseaux de distribution. Application à l'analyse technico-économique des solutions d'intégration des énergies renouvelables intermittentes". Doctoral Thesis, Ecole Centrale de Lille.
- [FRA09] Frazier, P., Powell, W., and Dayanik, S. (2009). "The knowledge-gradient policy for correlated normal beliefs". *INFORMS Journal on Computing*, 21(4):599–613.
- [KNO06] Knowles, J. (2006). "ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems". *IEEE Transactions on Evolutionary Computation*, 10(1):50–66.

Contacts:

Bruno BARRACOSA* (EDF R&D, L2S, CentraleSupélec), Julien BECT (L2S, CentraleSupélec), Héroïse DUTRIEUX BARAFFE (EDF R&D), Juliette MORIN (EDF R&D), Josselin FOURNEL (EDF R&D), Emmanuel VAZQUEZ (L2S, CentraleSupélec)
*bruno.tebbal-barracosa@edf.fr