RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

ONERA
THE FRENCH AEROSPACE LAB

Bayesian optimization to solve black box problems with hidden constraints

Nathalie Bartoli (ONERA - DTIS/M2CI)

Thierry Lefebvre, Rémi Lafage, Paul Saves

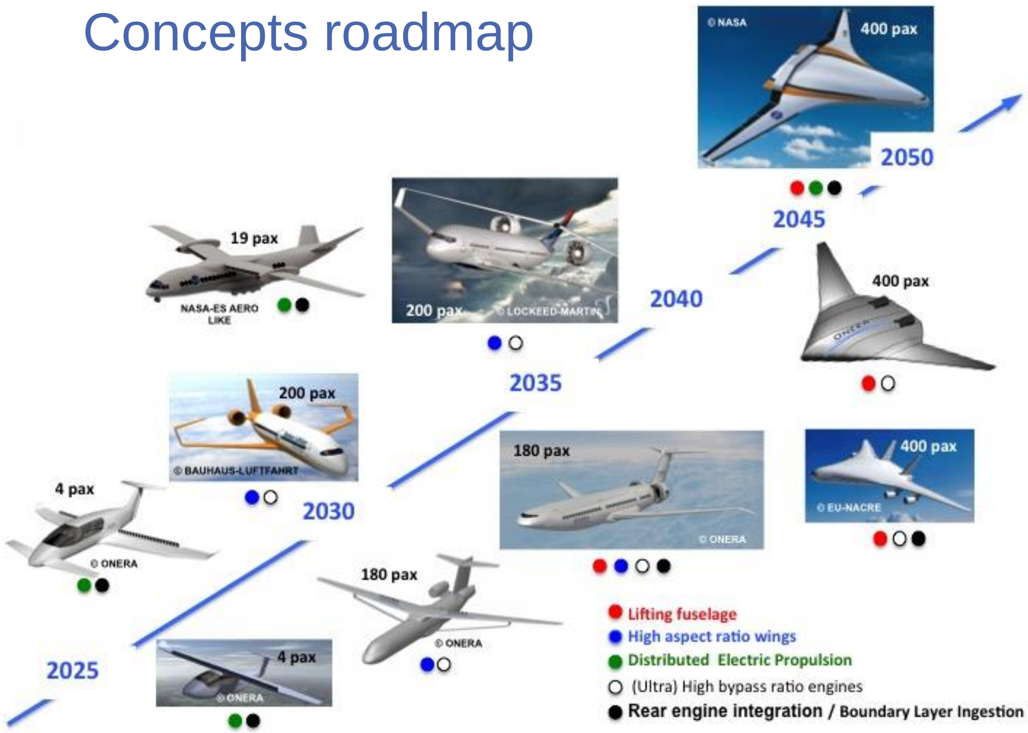
POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

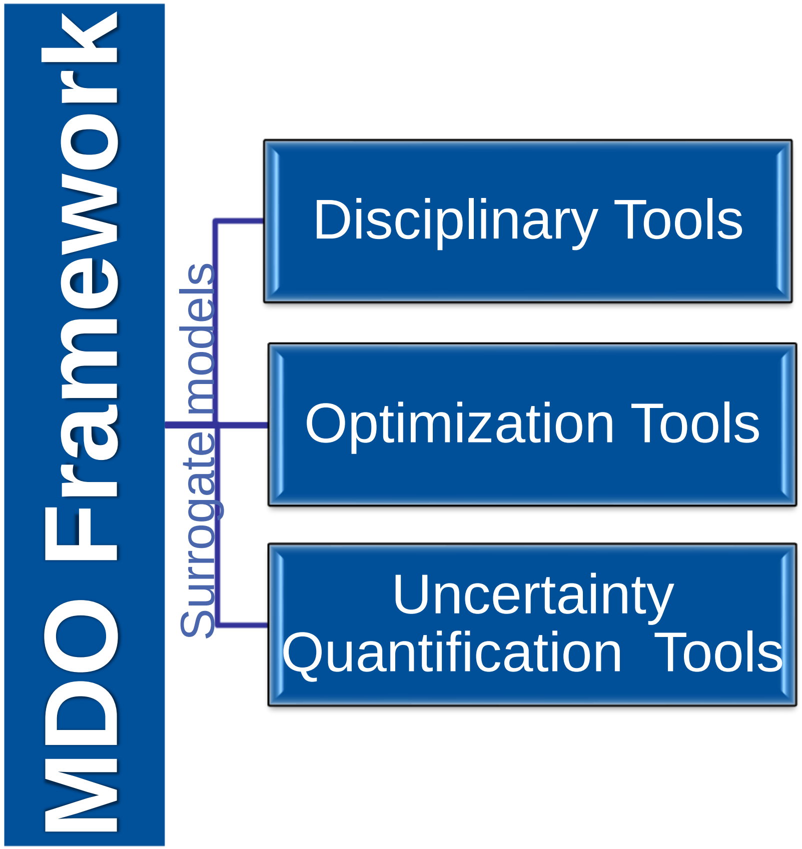
Youssef Diouane
Joseph Morlier

ISAE SUPAERO
Institut Supérieur de l'Aéronautique et de l'Espace
ICA

Tools to design the new aircraft configurations

Concepts roadmap



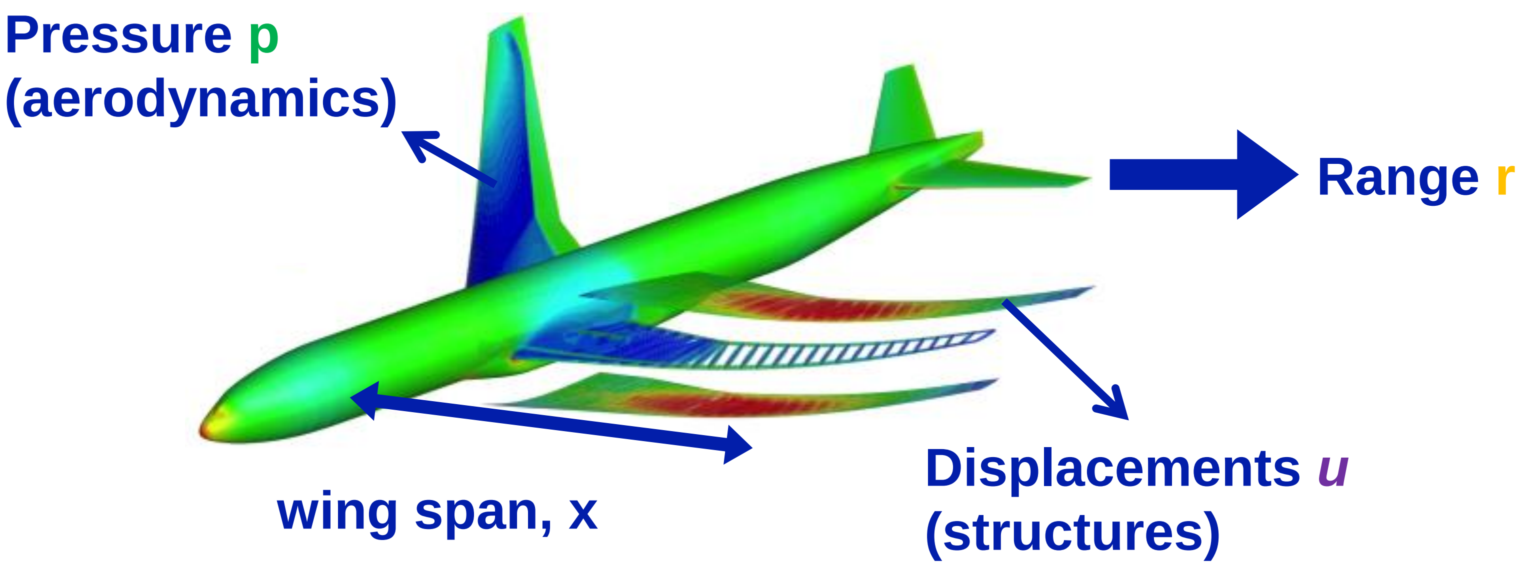


➔ Take into account more information earlier in the design process

➔ Keep a design space as large as possible

State of the art: Multidisciplinary Design Analysis and Optimization

Goal: Aircraft/drone design optimization



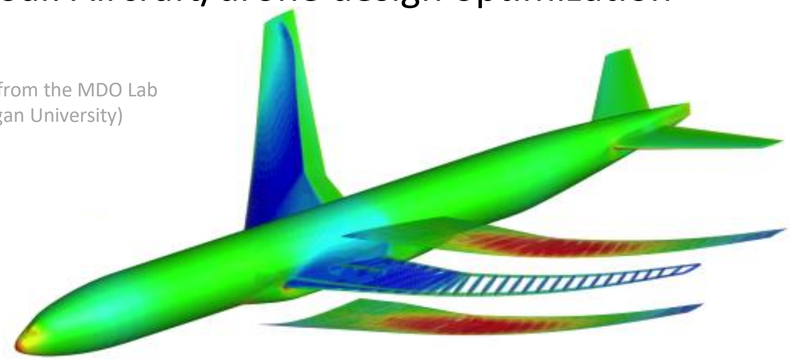
Exemple:
max Range
with respect to wing span
such that the aircraft is balanced (fixed point solved between aerodynamic and structure disciplines)

Martins J. R. R. A. and Ning A. Engineering Design Optimization. Cambridge University Press, 2020.

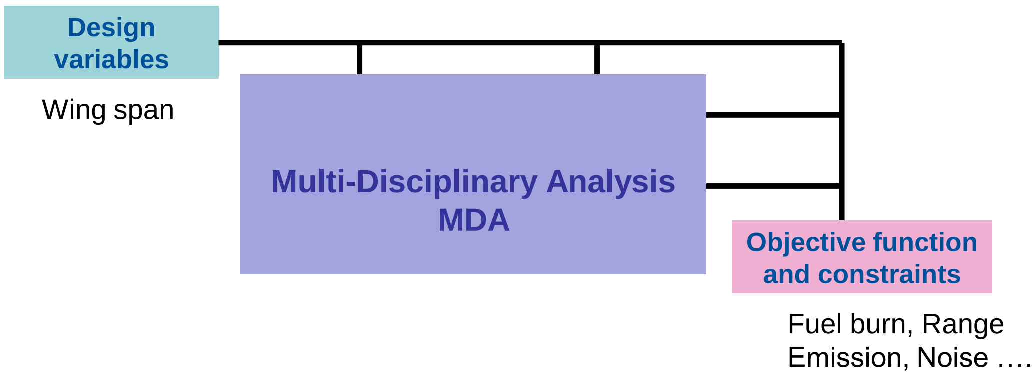
State of the art: Multidisciplinary Design Analysis and Optimization

Goal: Aircraft/drone design optimization

Image from the MDO Lab
(Michigan University)



MDO needs:
- A non-linear solver (Gauss-Seidel, Newton,…) to solve MDAs
- An optimizer (gradient based, gradient free, …) to solve the optimization problem

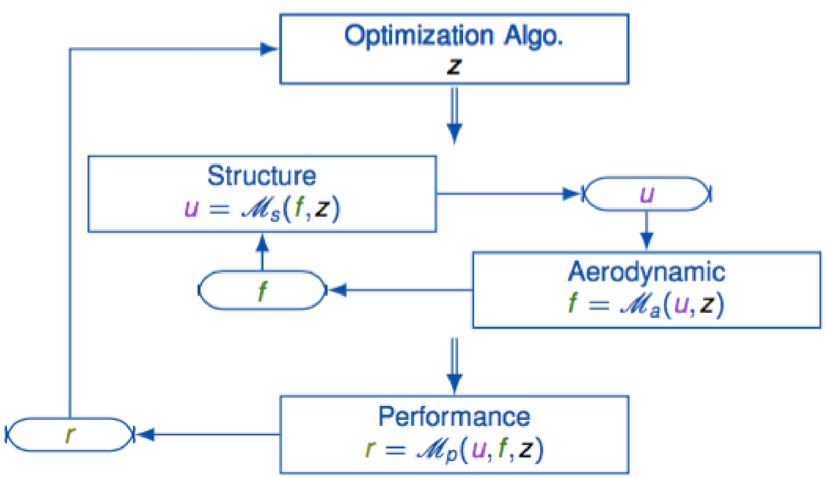


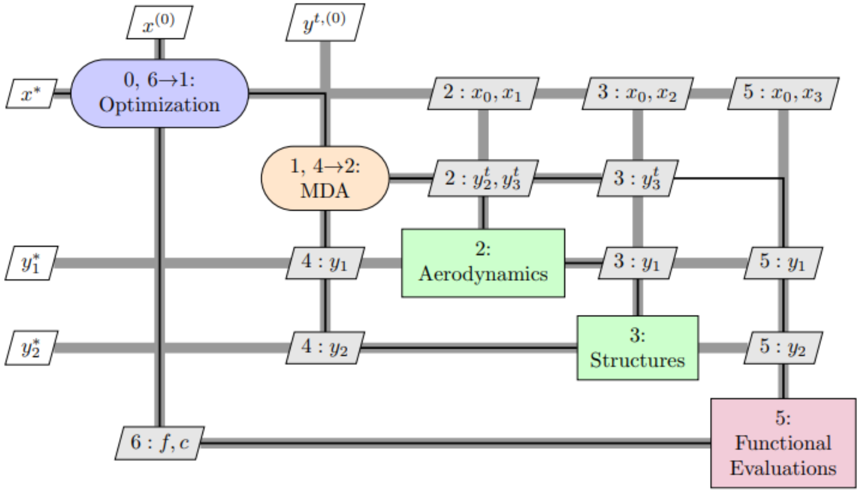
<u>Expensive black box problems</u>

→
- Gradient based optimization
- Evolutionary algorithms
- Bayesian optimization

Martins J. R. R. A. and Ning A. Engineering Design Optimization. Cambridge University Press, 2020.

Different strategies





Lambe, A. B., & Martins, J. R.R.A. (2012). Extensions to the design structure matrix for the description of multidisciplinary design, analysis, and optimization processes. Structural and Multidisciplinary Optimization, 46(2), 273-284.
Jasa, J. P., Hwang, J. T., & Martins, J. R. (2018). Open-source coupled aerostructural optimization using Python. Structural and Multidisciplinary Optimization, 57(4), 1815-1827.
Hwang, J. T., & Martins, J. R. (2018). A computational architecture for coupling heterogeneous numerical models and computing coupled derivatives. ACM Transactions on Mathematical Software (TOMS), 44(4), 37.
Gill, P. E., Murray, W., & Saunders, M. A. (2005). SNOPT: An SQP algorithm for large-scale constrained optimization. SIAM review, 47(1), 99-131.

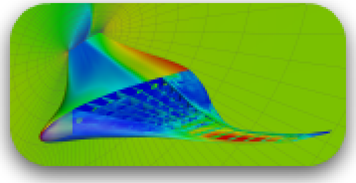
Different strategies

Gradient-based algorithm

Derivatives of objective and constraints

Methods for computing derivatives (coupled adjoint)

Source codes available

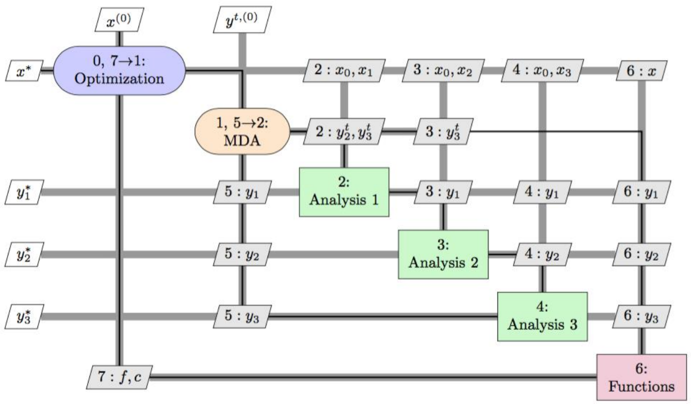


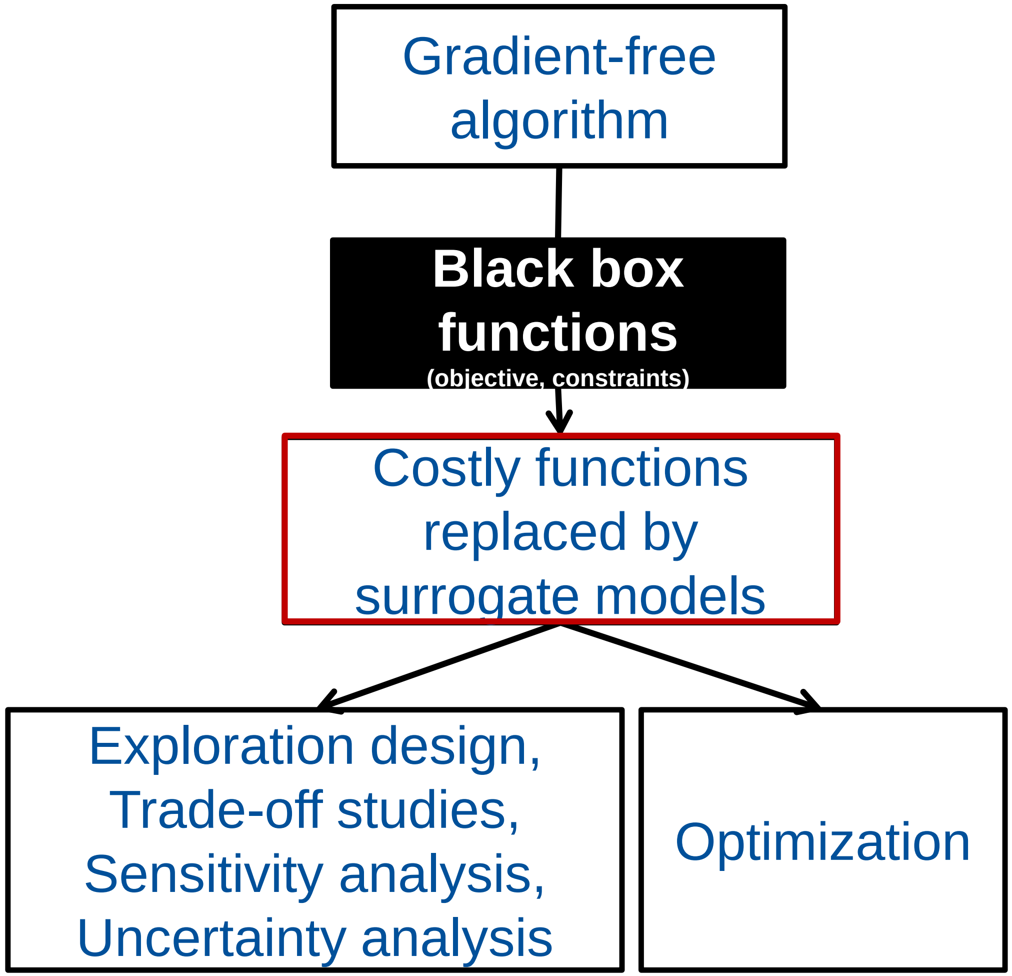
MDOlab strategy within OpenMDAO

Python user script Setup up the problem: objective function, constraints, design variables, optimizer and solver options					
Optimizer interface *pyOptSparse* Common interface to various optimization software		Aerostructural solver *AeroStruct* Coupled solution methods and coupled derivative evaluation		Geometry modeler *DVGeometry/GeoMACH* Defines and manipulates geometry, evaluates derivatives	
SNOPT	Other optimizers	Flow solver *ADflow* Governing and adjoint equations	Structural solver *TACS* Governing and adjoint equations		

Lambe, A. B., & Martins, J. R.R.A. (2012). Extensions to the design structure matrix for the description of multidisciplinary design, analysis, and optimization processes. Structural and Multidisciplinary Optimization, 46(2), 273-284.
Jasa, J. P., Hwang, J. T., & Martins, J. R. (2018). Open-source coupled aerostructural optimization using Python. Structural and Multidisciplinary Optimization, 57(4), 1815-1827.
Hwang, J. T., & Martins, J. R. (2018). A computational architecture for coupling heterogeneous numerical models and computing coupled derivatives. ACM Transactions on Mathematical Software (TOMS), 44(4), 37.
Gill, P. E., Murray, W., & Saunders, M. A. (2005). SNOPT: An SQP algorithm for large-scale constrained optimization. SIAM review, 47(1), 99-131.

Different strategies

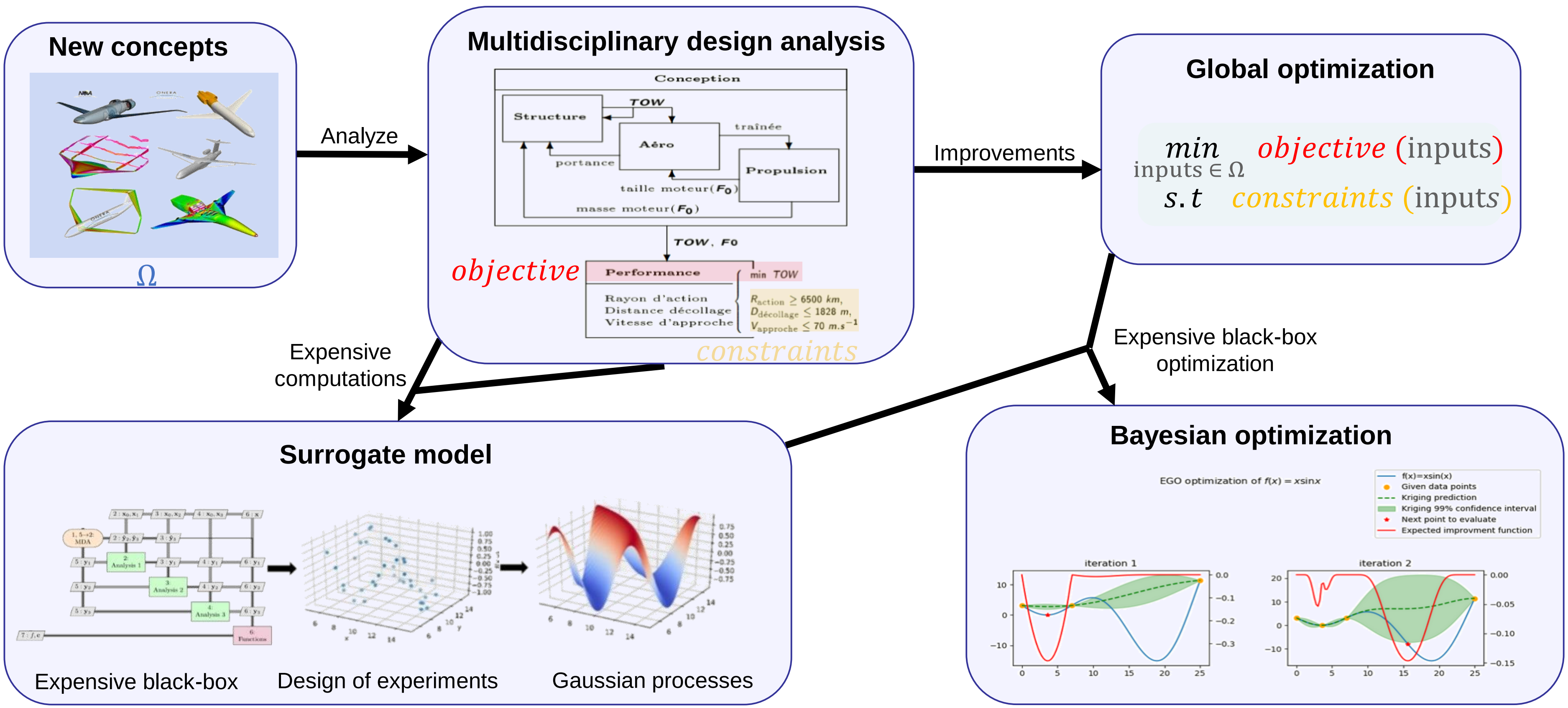




Generic approach
Extension to multiple disciplines

Audet, C., & Hare, W. (2017). Derivative-free and blackbox optimization. Berlin: Springer International Publishing
Powell, M. J. (1994). A direct search optimization method that models the objective and constraint functions by linear interpolation. In Advances in optimization and numerical analysis (pp. 51-67). Springer, Dordrecht.
Forrester, A., Sobester, A., & Keane, A. (2008). Engineering design via surrogate modelling: a practical guide. John Wiley & Sons

Overview



New concepts
Ω
Analyze
Multidisciplinary design analysis
Conception
Structure
TOW
Aéro
traînée
portance
Propulsion
taille moteur(F_0)
masse moteur(F_0)
TOW, F0
objective
Performance
min TOW
Rayon d'action
Distance décollage
Vitesse d'approche
$R_{action} \geq 6500\ km$,
$D_{décollage} \leq 1828\ m$,
$V_{approche} \leq 70\ m.s^{-1}$
constraints
Improvements
Global optimization
$\min_{inputs \in \Omega}$ objective (inputs)
s.t constraints (inputs)
Expensive computations
Expensive black-box optimization
Surrogate model
Expensive black-box
Design of experiments
Gaussian processes
Bayesian optimization
EGO optimization of $f(x) = x\sin x$
f(x)=xsin(x)
Given data points
Kriging prediction
Kriging 99% confidence interval
Next point to evaluate
Expected improvment function
iteration 1
iteration 2

Outline

- Kriging based surrogate models
- Bayesian optimization
  – mono & multiobjective
  – hidden constraints
- Applications
  – DRAGON: ONERA hybrid electric aircraft
  – Jet engine architecture
  – BRAC: BOMBARDIER conventional aircraft

Methodology developments: Surrogate Models

Definition of a metamodel library dedicated to Aircraft design

- Models to handle a large number of design variables
  - ➡ New Kriging models: KPLS & KPLS-K
- Models to handle heterogeneous functions
  - ➡ Mixture of experts (MOE)
- Models to handle heterogeneous variables
  - ➡ Kriging based on continuous relaxation
- Models to handle multifidelity data
  - ➡ Co-Kriging (MFKPLS, MFKPLS-K)
  - ➡ Co-Kriging with heteroscedastic Noise

open source python toolbox for surrogate models

SMT: Surrogate Modeling Toolbox

github.com/SMTorg/smt

➔ **Surrogate models with some focus on derivatives**

➔ **Included some Jupyter notebooks**

SMT 2.8 features (November 2024):

- Models to handle a large number of design variables (KPLS – KPLSK – MGP)
- **Mixture of experts** to handle heterogeneous functions (MOE)
- Different covariance kernels added
- **Multi-fidelity models** (MFK – MFKPLS – MFKPLSK)
- Noisy kriging to handle uncertainties on data
- Kriging models for **mixed variables** (continuous, discrete, categorical) & associated kernels
- Kriging models for hierarchical variables (meta, neutral, decreed) & associated kernels
- Sparse GP models to handle large database
- **Bayesian optimization** (EGO without constraint) for continuous and mixed variables

Bouhlel, M. A., Hwang, J. T., Bartoli, N., Lafage, R., Morlier, J., & Martins, J. R. (2019). A Python surrogate modeling framework with derivatives. Advances in Engineering Software, 135, 102662.
Saves, P., Lafage, R., Bartoli, N., Diouane, Y., Bussemaker, J., Lefebvre, T., Hwang J. & Martins, J. R. (2024). SMT 2.0: A Surrogate Modeling Toolbox with a focus on hierarchical and mixed variables Gaussian processes. Advances in Engineering Software, 188, 103571.

SMT: Focus on derivatives

$$y = f(x, xt, yt)$$

(xt, yt) Training data
(x, y) Prediction data

Surrogate modeling methods provided by SMT.

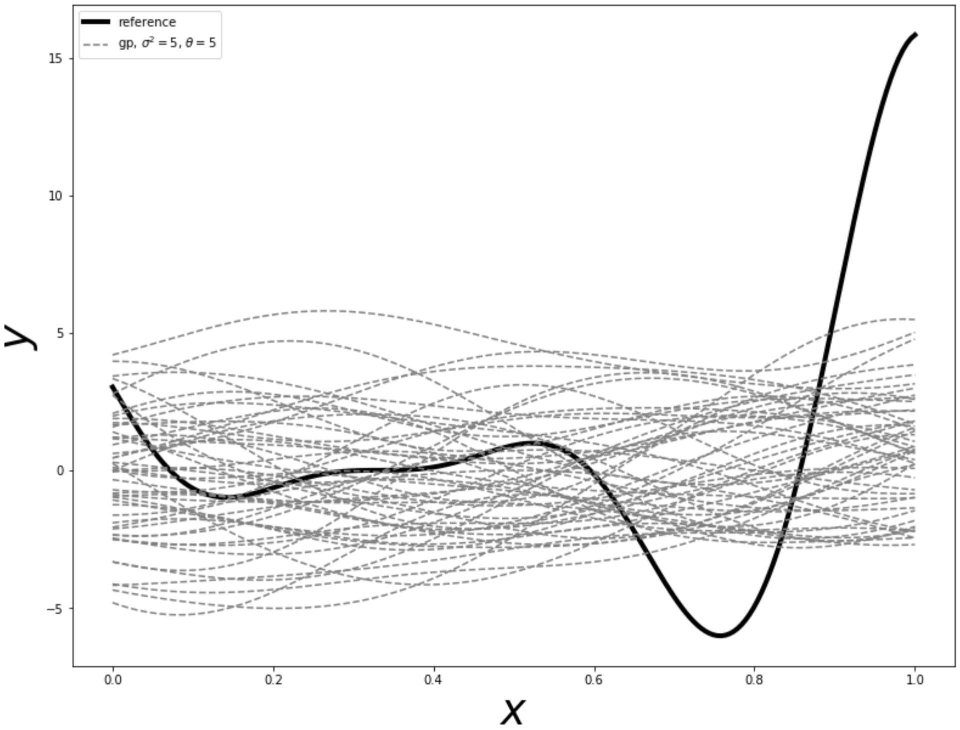
Method	Advantages (+) and disadvantages (−)	Derivatives		
		Train.	Pred.	Out.
Kriging	+ Prediction variance, flexible − Costly if number of inputs or training points is large − Numerical issues when points are too close to each other	No	Yes	No
KPLS	+ Prediction variance, fast construction + Suitable for high-dimensional problems − Numerical issues when points are too close to each other	No	Yes	No
KPLSK	+ Prediction variance, fast construction + Suitable for high-dimensional problems − Numerical issues when points are too close to each other	No	Yes	No
GE-KPLS	+ Prediction variance, fast construction + Suitable for high-dimensional problems + Control of the correlation matrix size − Numerical issues when points are too close to each other − Choice of step parameter is not intuitive	Yes	Yes	No
RMTS	+ Fast prediction + Training scales well up to 10^5 training points + No issues with points that are too close to each other − Poor scaling with number of inputs above 4 − Slow training overall	Yes	Yes	Yes
RBF	+ Simple, only a single tuning parameter + Fast training for small number of training points − Susceptible to oscillations − Numerical issues when points are too close to each other	No	Yes	Yes
IDW	+ Simple, no training required − Derivatives are zero at training points − Poor overall accuracy	No	Yes	Yes
LS	+ Simple, fast construction − Accurate only for linear problems	No	Yes	No
QP	+ Simple, fast construction − Large number of points required for large number of inputs	No	Yes	No

- (dyt/dxt): training derivatives used for gradient-enhanced modeling
- (dy/dx): prediction derivatives
- (dy/dyt): derivatives with respect to the training data

Bouhlel, M. A., Hwang, J. T., Bartoli, N., Lafage, R., Morlier, J., & Martins, J. R. (2019). A Python surrogate modeling framework with derivatives. Advances in Engineering Software, 135, 102662.
Saves, P., Lafage, R., Bartoli, N., Diouane, Y., Bussemaker, J., Lefebvre, T., Hwang J. & Martins, J. R. (2024). SMT 2.0: A Surrogate Modeling Toolbox with a focus on hierarchical and mixed variables Gaussian processes. Advances in Engineering Software, 188, 103571.

Gaussian process or Kriging model

$x \in \mathbb{R}^d$
$y(x) \in \mathbb{R}$



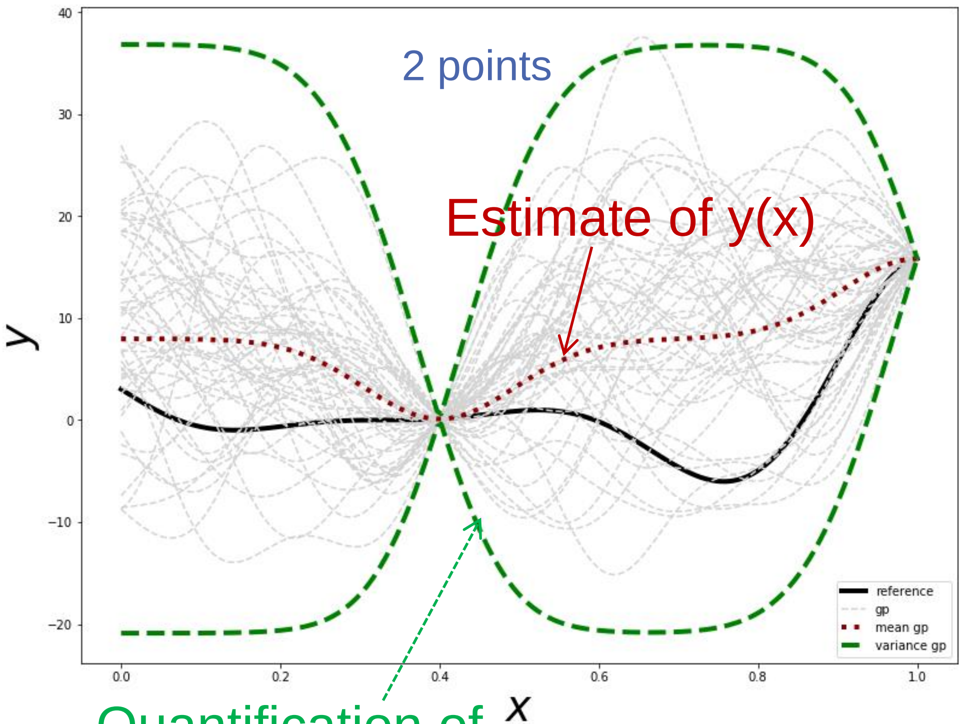
Gaussian process characterized by:

- its mean (or trend)
  $\mu(x) \in \mathbb{R}$
- its covariance Kernel
  $\mathrm{k}(x, x') \in \mathbb{R}$

D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy, 52(6):119–139, 1951
C. E. Rasmussen and C. K. Williams. Gaussian processes for machine learning, volume 1. MIT press Cambridge, 2006.

Gaussian process or Kriging model

$x \in \mathbb{R}^d$
$y(x) \in \mathbb{R}$



Quantification of the uncertainties in these estimates

Gaussian process characterized by:

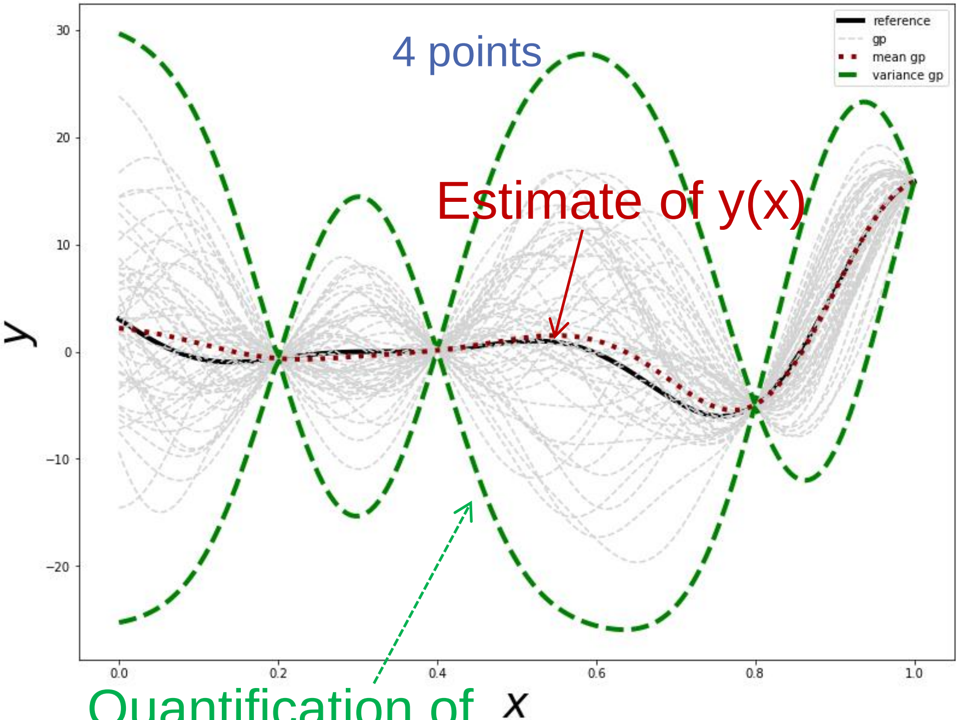
- its mean (or trend)
  $\mu(x) \in \mathbb{R}$
- its covariance Kernel
  $\mathrm{k}(x, x') \in \mathbb{R}$

+ information provided by data

D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy, 52(6):119–139, 1951
C. E. Rasmussen and C. K. Williams. Gaussian processes for machine learning. volume 1. MIT press Cambridge. 2006.

Gaussian process or Kriging model

$x \in \mathbb{R}^d$
$y(x) \in \mathbb{R}$



Quantification of the uncertainties in these estimates

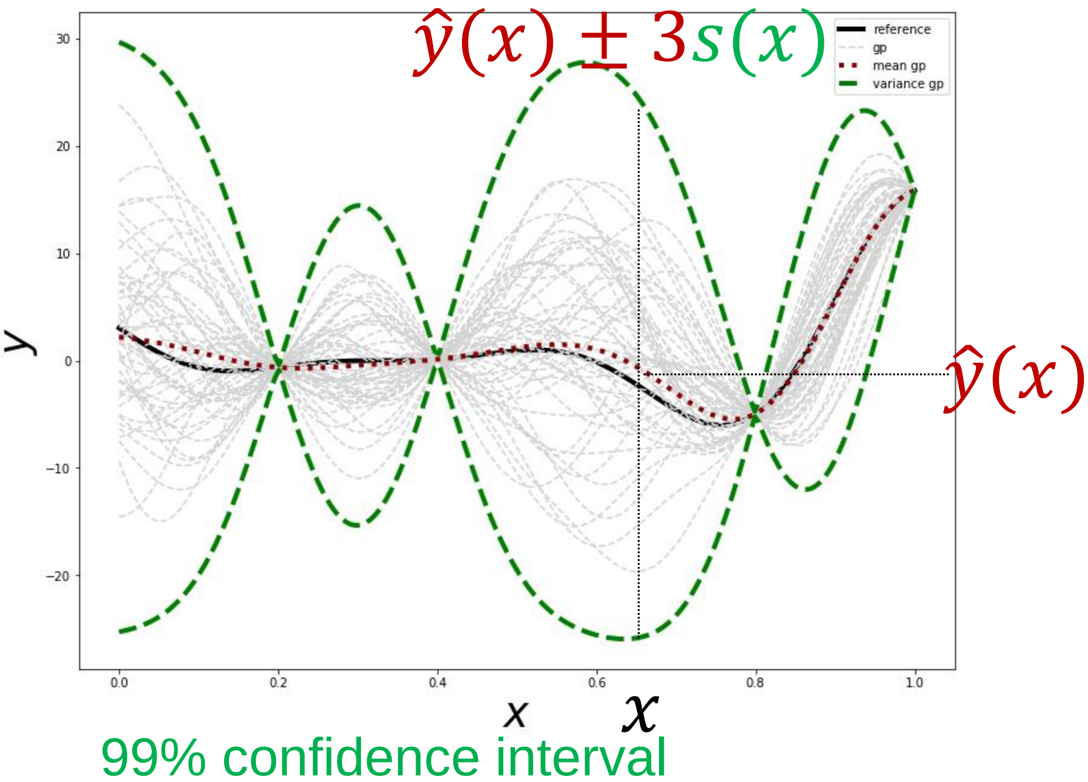
Gaussian process characterized by:

- its mean (or trend)
  $\mu(x) \in \mathbb{R}$
- its covariance Kernel
  $\mathrm{k}(x, x') \in \mathbb{R}$

+ information provided by data

D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy, 52(6):119–139, 1951
C. E. Rasmussen and C. K. Williams. Gaussian processes for machine learning, volume 1. MIT press Cambridge, 2006.

Gaussian process or Kriging model



$$f(x) \Rightarrow Y(x) = N(\hat{y}(\boldsymbol{x}), s^2(\boldsymbol{x}))$$

$x \in \mathbb{R}^d$

$y(x) \in \mathbb{R}$

DOE (LHS)

↓

Estimation of hyperparameters θ_i ($i=1,\dots,d$) by MLE

↓

Covariance matrix

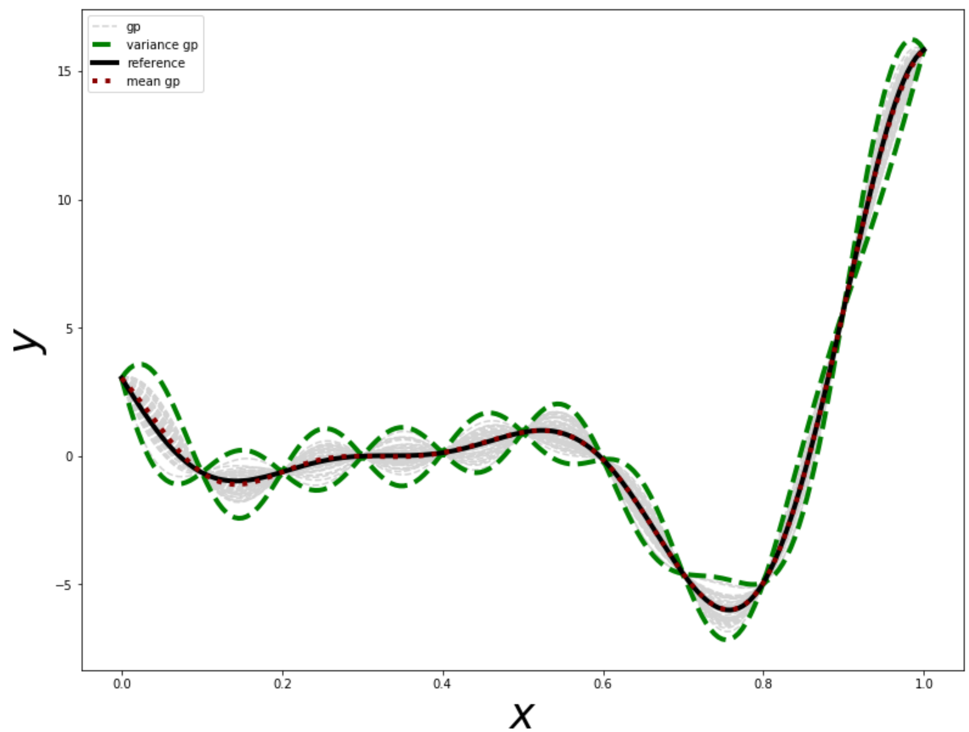
↓

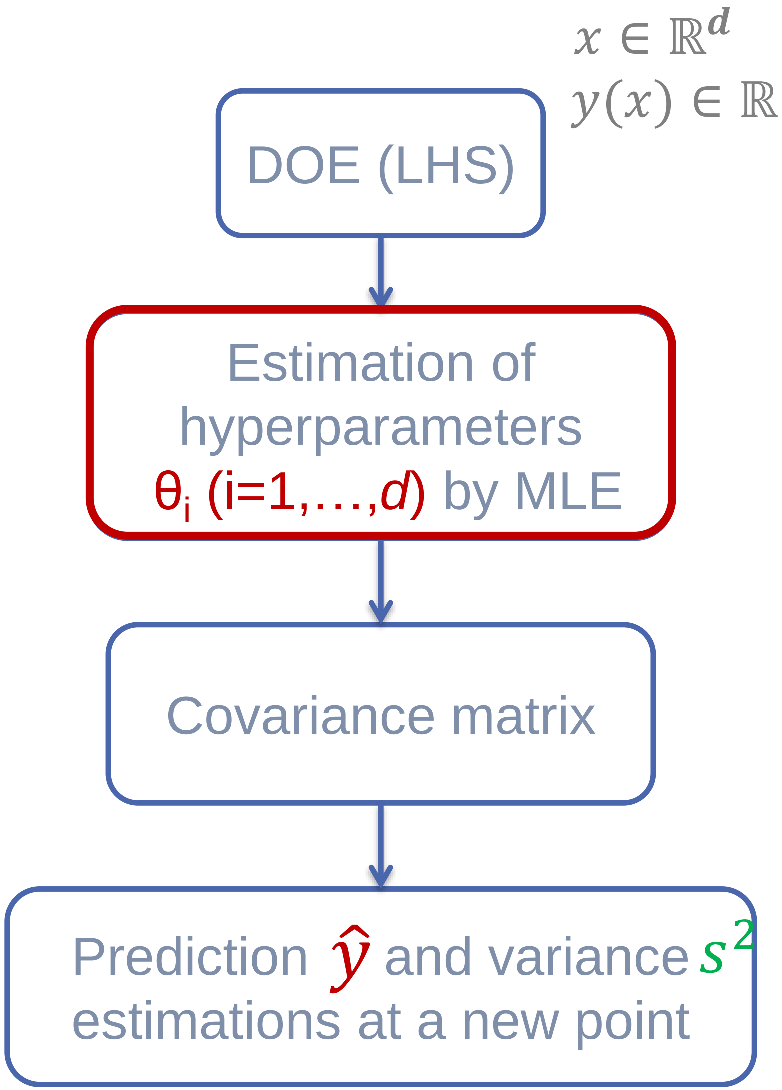
Prediction $\hat{y}$ and variance s^2 estimations at a new point

D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy, 52(6):119–139, 1951

C. E. Rasmussen and C. K. Williams. Gaussian processes for machine learning, volume 1. MIT press Cambridge, 2006.

Gaussian process or Kriging model





- ➔ Hyperparameters tuning
- ➔ Number of hyperparameters increases with the dimension *d* (number of design variables)
- ➔ Curse of dimensionality

D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy, 52(6):119–139, 1951
C. E. Rasmussen and C. K. Williams. Gaussian processes for machine learning, volume 1. MIT press Cambridge, 2006.

Models to handle a large number of design variables

Kriging models: KPLS & KPLS-K

➔ Exploitation of information provided by PLS (Partial Least Squares) in the construction of the Kriging model to reduce the dimension: KPLS and KPLS-K models

Ordinary Kriging

$$k(x,x') = \sigma^2 \exp\left(-\sum_{i=1}^{d} \theta_i |x_i - x_i'|^{p_i}\right)$$

with

d parameters θ_i to evaluate

Covariance kernel

KPLS

$$k_{PLS}(x,x') = \sigma^2 \exp\left(-\sum_{i=1}^{d} \eta_i |x_i - x_i'|^{p_i}\right)$$

with

$$\eta_i = \sum_{j=1}^{h} \theta_j \left|w_{i,j}\right|^{p_i}$$

h parameters θ_j to evaluate

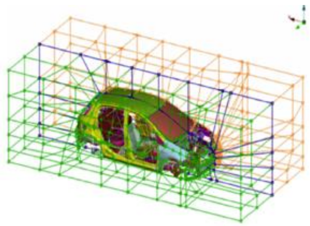
- $\left|w_{i,j}\right|_{\,i=1,\cdots,d}$ describes how sensitive the *j*-th principal component is to each design variable *i* ➔ **PLS**
- θ_j describes how sensitive the function is to each principal component (max *h*≈ 4) ➔ **MLE**
- If $h = d$ ➔ classical Kriging (exponential kernels)

Wold H (1966) Estimation of Principal Components and Related Models by Iterative Least squares, Academic Press, New York, pp 391–420
Bouhlel, M. A., Bartoli, N., Otsmane, A., and Morlier, J., “Improving kriging surrogates of high-dimensional design models by Partial Least Squares dimension reduction,” Structural and Multidisciplinary Optimization, Vol. 53, No. 5, 2016, pp. 935–952.
Bouhlel, M. A., Bartoli, N., Otsmane, A., and Morlier, J., “An Improved Approach for Estimating the hyperparameters of the Kriging Model for high-dimensional problems through The Partial Least Squares Method”, Mathematical Problems in Engineering, Vol. 2016(4), May 2016

Models to handle a large number of design variables

Kriging models: KPLS & KPLS-K

MOPTA test case function from automotive industry



d=124 inputs 1 output
training: 500 points LHS,
validation: 100 points

$$RE = \frac{\|y-\hat{y}\|_2}{\|y\|_2}100$$

Surrogate	RE (%)	CPU time
Ordinary Kriging (Scikit-Learn)	3.28e-7	17 min 23 s
KPLS *h*=4	4.52e-7	37 s

Time / 28

Intel(R) Core(TM) i7-4500U CPU@1.80GHz, 6.00 Go RAM

SNECMA test case (turbomachinery)

d=98 inputs 1 output
training: 340 points LHS,
validation: 24 points LHS

Surrogate	RE (%)	CPU time
Ordinary Kriging (Snecma ref)	2.24	1min 33s
KPLS *h*=1	1.62	0.90 s
KPLS *h*=2	1.62	1.56 s

Time / 60

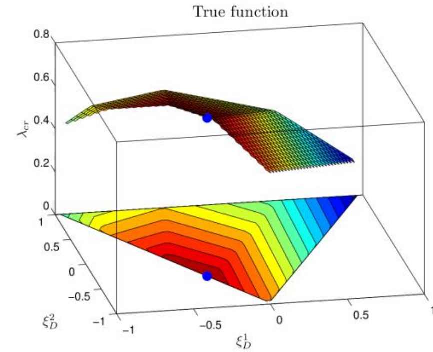
Intel(R) Xeron(R) CPU W3565@3.20GHz, 7.98 Go RAM Quad core

➔ CPU time drastically reduced: interest for adaptive enrichment optimization method

➔ Automatic choice for the number of PLS components

Jones, D., "Large-scale multi-disciplinary mass optimization in the auto industry," MOPTA 2008 Conference (20 August 2008)
Bouhlel, M.-A., Ph.D. thesis, ISAE-SUPAERO, 2016, https://hal.archives-ouvertes.fr/tel-01293319

Models to handle heterogeneous functions



Mixture of Experts (MOE)

→ **Mixture of experts technique**

- Divide the database into K clusters (Expectation-Maximization)
- Build a local surrogate model on each cluster (RBF, Polynomial functions, Kriging,…)
- Recombine the K local models into a global model

$$\hat{f}(x) = \sum_{i=1}^{K} P(k = i/X = x)\hat{f}_i$$

K number of clusters (Gaussian components)

$P(k = i/X = x)$ probability to be in the cluster *i*

(posterior probability given by the Expectation-Maximization algorithm)

$\hat{f}_i$ local expert build using the points in cluster *i* (RBF, Polynomial functions, Kriging,…)

Jordan, M. I., Jacobs, R. A, "Hierarchical mixtures of experts and the EM algorithm", Neural Comput. 6 (1994) 181–214.

Bettebghor, D., Bartoli, N., Grihon, S., Morlier, J., and Samuelides, M., "Surrogate modeling approximation using a mixture of experts based on EM joint estimation," Structural and Multidisciplinary Optimization, Vol. 43, No. 2, 2011, pp. 243–259

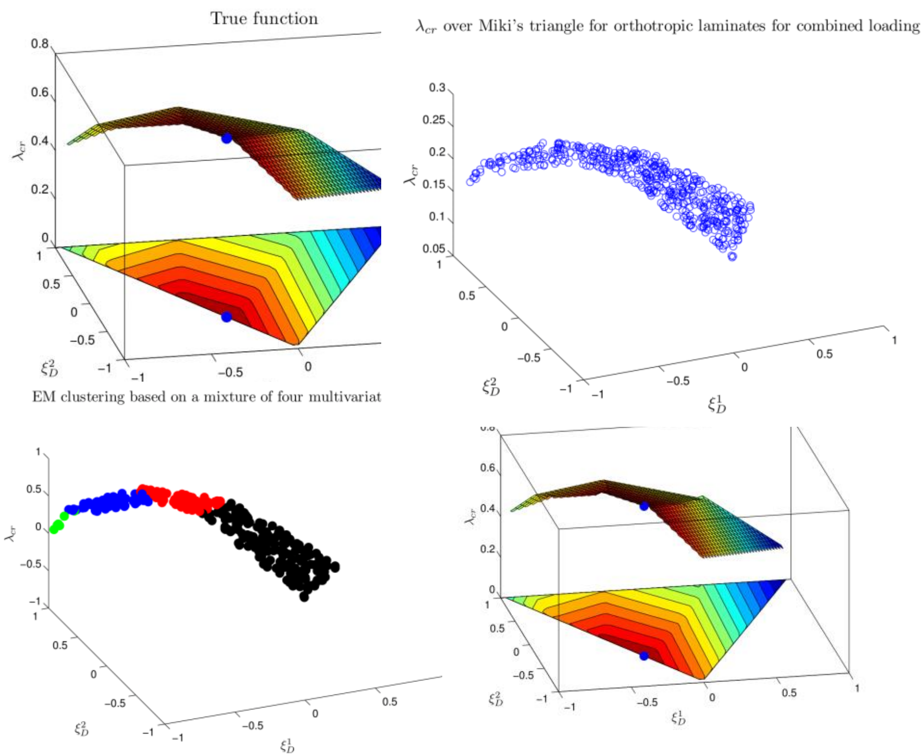
Liem, R. P., Mader, C. A., and Martins, J. R. R. A., "Surrogate Models and Mixtures of Experts in Aerodynamic Performance Prediction for Mission Analysis," Aerospace Science and Technology, Vol. 43, 2015, pp. 126–151

Models to handle heterogeneous functions

Mixture of Experts (MOE)

→ Mixture of experts technique

- Divide the database into K clusters (Expectation-Maximization)
- Build a local surrogate model on each cluster (RBF, Polynomial functions, Kriging,…)
- Recombine the K local models into a global model



Comparison on Buckling critical loads PhD D. Bettebghor 2011

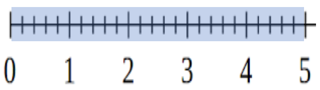
Jordan, M. I., Jacobs, R. A, "Hierarchical mixtures of experts and the EM algorithm", Neural Comput. 6 (1994) 181–214.
Bettebghor, D., Bartoli, N., Grihon, S., Morlier, J., and Samuelides, M., "Surrogate modeling approximation using a mixture of experts based on EM joint estimation," Structural and Multidisciplinary Optimization, Vol. 43, No. 2, 2011, pp. 243–259
Liem, R. P., Mader, C. A., and Martins, J. R. R. A., "Surrogate Models and Mixtures of Experts in Aerodynamic Performance Prediction for Mission Analysis," Aerospace Science and Technology, Vol. 43, 2015, pp. 126–151

Models to handle mixed variables (continuous, discrete, categorical)

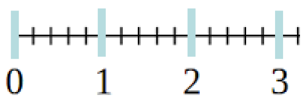
Hybrid variables

Variables types:

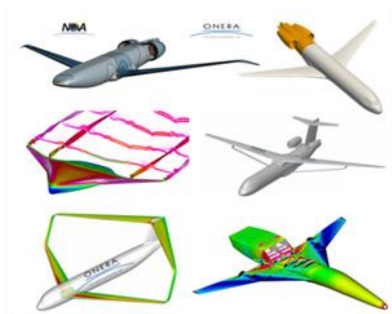
- **Continuous (x) Ex:** wing length



- **Integer (z) Ex:** winglet number



- **Categorical (u) Ex:** Plane shape / material properties



Categorical variables: n variables,
n=2
u1= shape
u2= color

Levels: L_i levels for I in 1,..n,
L_1=3, L_2=2
Levels(u1)= square, circle, rhombus
Levels(u2)= blue, red

Categories: $\prod_{i=1}^{n} L_i$, **2*3=6**
- **Blue square**
- **Blue circle**
- **Blue rhombus**
- **Red square**
- **Red circle**
- **Red rhombus**

6 possibilities

State of the Art approach: Continuous Relaxation

Ex: Garrido-Merchán and Hernández-Lobato model

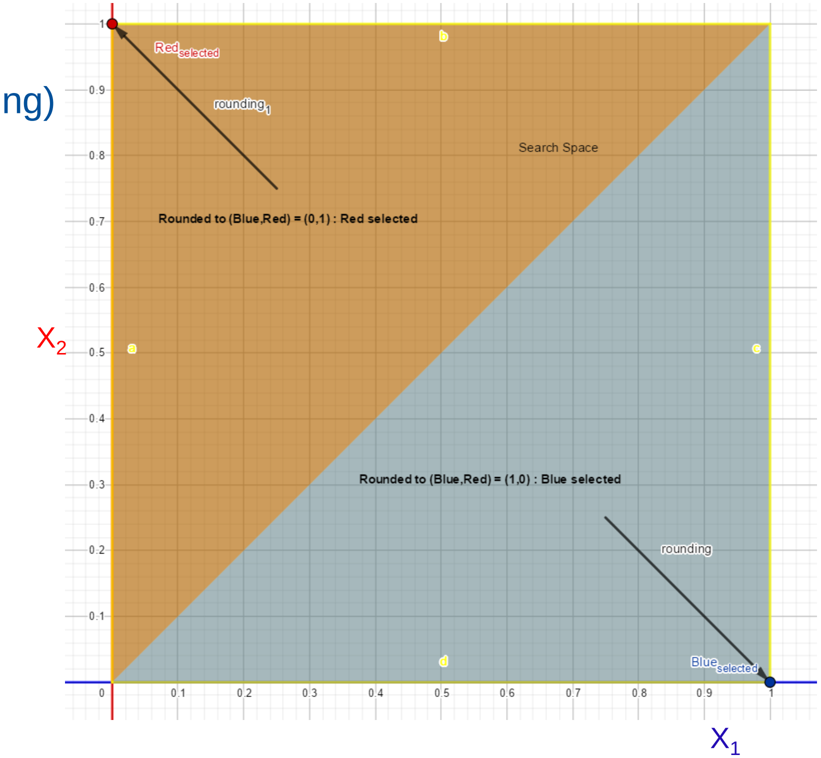
➔ Model as a **Continuous Relaxation** (one-hot-encoding)

Example with 1 categorical variable X and two levels

- Red color
- Blue color

➔1 Categorical variable replaced by 2 continuous variables denoted by X_1 and $X_2 \in [0,1]$

- If $X_1 > X_2$ => (1., 0.) => Blue color
- If $X_1 < X_2$ => (0.,1.) => Red color



A continuous kernel
(in the relaxed dimension)

Garrido-Merchán E. C., Hernández-Lobato D. "Dealing with categorical and integer-valued variables in Bayesian Optimization with Gaussian processes". Neurocomputing, vol. 380 (2020), pages 20-35
Saves, P., Bartoli, N., Diouane, Y., Morlier, J., & Lefebvre, T. (2021, March). Enhanced Kriging models within a Bayesian optimization framework to handle both continuous and categorical inputs. In SIAM CSE21

State of the Art approach: Continuous Relaxation

Ex: Garrido-Merchán and Hernández-Lobato model

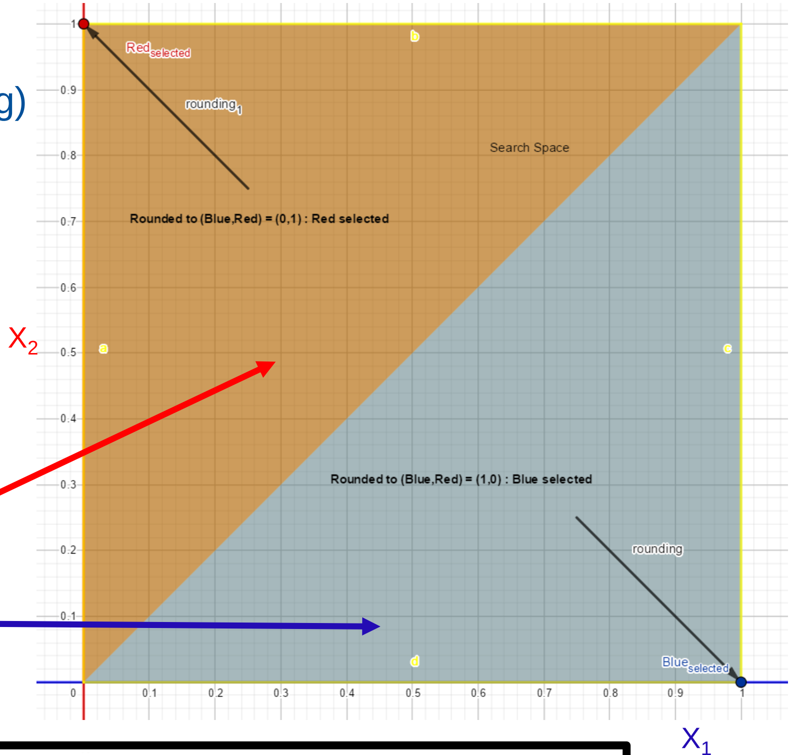
➔ Model as a **Continuous Relaxation** (one-hot-encoding)

Example with 1 categorical variable X and two levels

- Red color
- Blue color

➔1 Categorical variable replaced by 2 continuous variables denoted by X_1 and $X_2 \in [0,1]$

- If $X_1 > X_2$ => (1., 0.) => Blue color
- If $X_1 < X_2$ => (0.,1.) => Red color



A continuous kernel (in the relaxed dimension)

$$k(x^r, x^s, \theta^{cont}) = \prod_{j=1}^{d'} \exp\left(-(x_j^r - x_j^s)\theta_j^{cont}(x_j^r - x_j^s)\right)$$

Garrido-Merchán E. C., Hernández-Lobato D. "Dealing with categorical and integer-valued variables in Bayesian Optimization with Gaussian processes". Neurocomputing, vol. 380 (2020), pages 20-35
Saves, P., Bartoli, N., Diouane, Y., Morlier, J., & Lefebvre, T. (2021, March). Enhanced Kriging models within a Bayesian optimization framework to handle both continuous and categorical inputs. In SIAM CSE21

State of the Art approach: Continuous Relaxation

Ex: Garrido-Merchán and Hernández-Lobato model

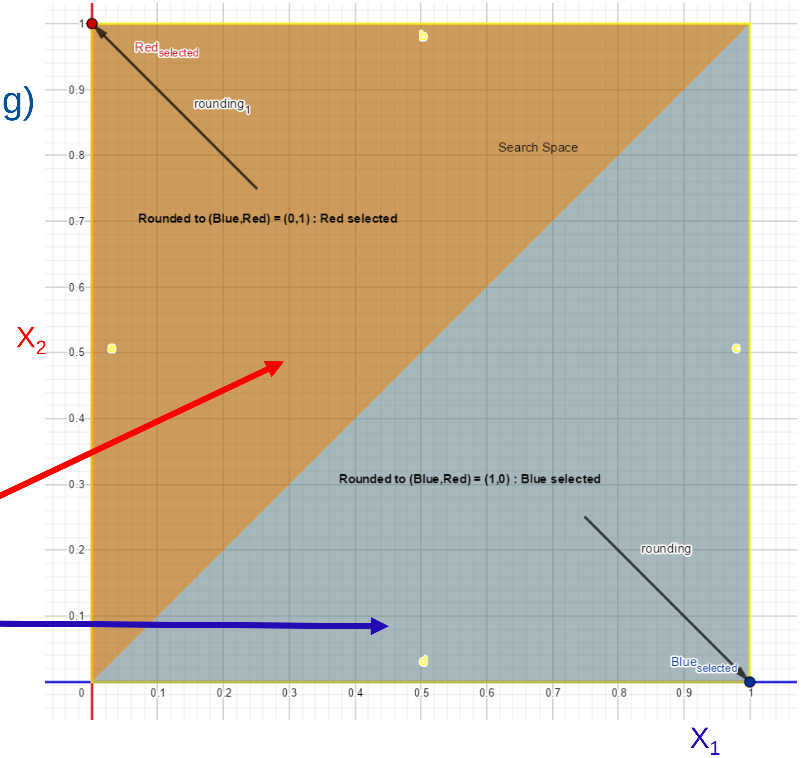
➔ Model as a **Continuous Relaxation** (one-hot-encoding)

Example with 1 categorical variable X and two levels

- Red color
- Blue color

➔1 Categorical variable replaced by 2 continuous variables denoted by X_1 and $X_2 \in [0,1]$

- If $X_1>X_2$ => (1., 0.) => Blue color
- If $X_1<X_2$ => (0.,1.) => Red color



A continuous kernel (in the relaxed dimension)

➔ **Increase the dimension**

Garrido-Merchán E. C., Hernández-Lobato D. "Dealing with categorical and integer-valued variables in Bayesian Optimization with Gaussian processes". Neurocomputing, vol. 380 (2020), pages 20-35
Saves, P., Bartoli, N., Diouane, Y., Morlier, J., & Lefebvre, T. (2021, March). Enhanced Kriging models within a Bayesian optimization framework to handle both continuous and categorical inputs. In SIAM CSE21

State of the Art approach: Continuous Relaxation

Ex: Garrido-Merchán and Hernández-Lobato model

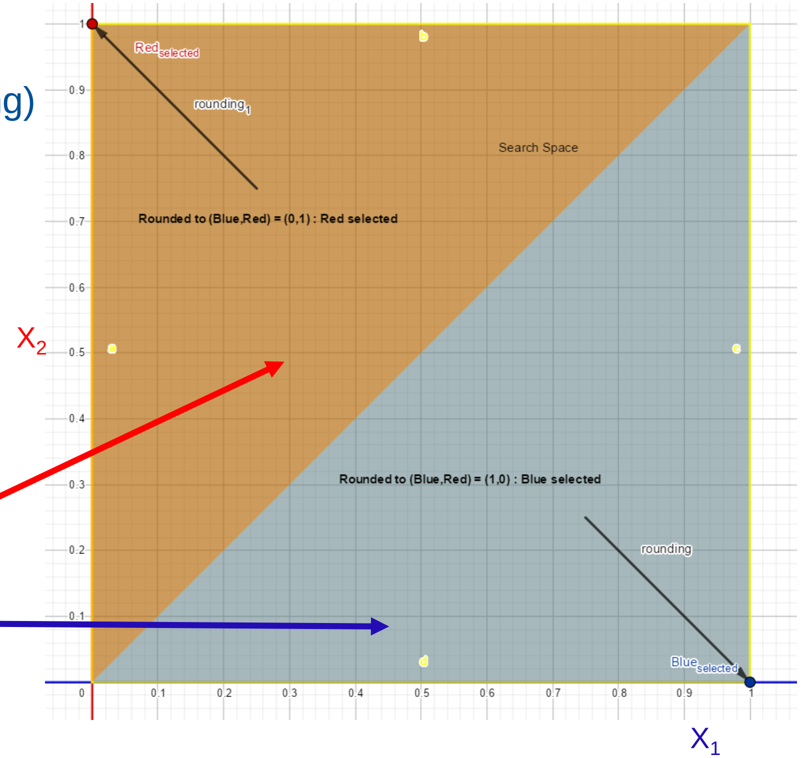
➔ Model as a **Continuous Relaxation** (one-hot-encoding)

Example with 1 categorical variable X and two levels

- Red color
- Blue color

➔ 1 Categorical variable replaced by 2 continuous variables denoted by X_1 and $X_2 \in [0,1]$

- If $X_1 > X_2$ => (1., 0.) => Blue color
- If $X_1 < X_2$ => (0.,1.) => Red color



➔ Use of KPLS models to decrease the dimension

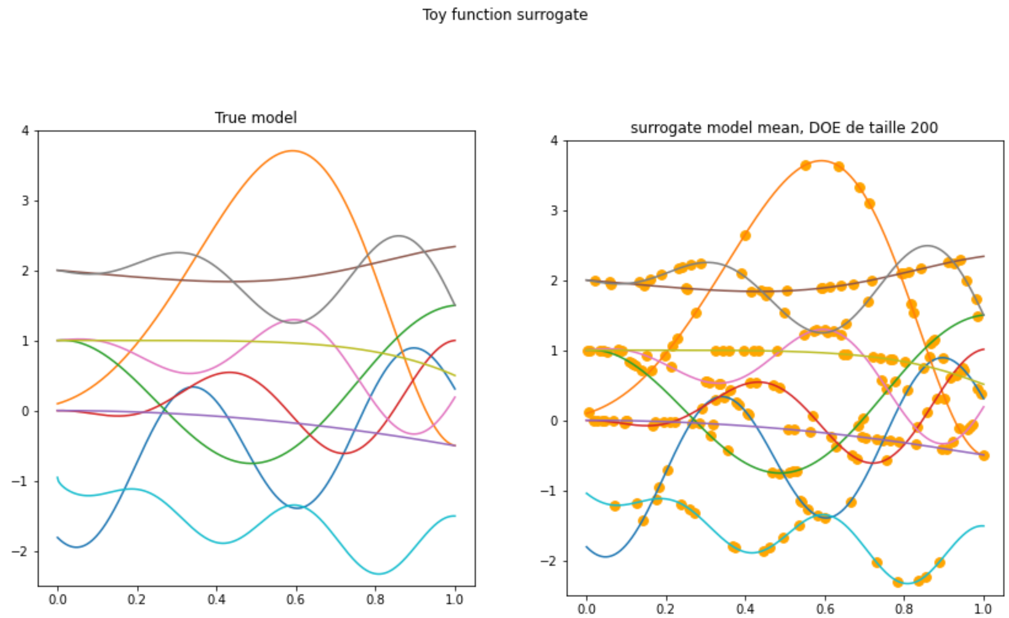
Garrido-Merchán E. C., Hernández-Lobato D. “Dealing with categorical and integer-valued variables in Bayesian Optimization with Gaussian processes”. Neurocomputing, vol. 380 (2020), pages 20-35

Saves, P., Bartoli, N., Diouane, Y., Morlier, J., & Lefebvre, T. (2021, March). Enhanced Kriging models within a Bayesian optimization framework to handle both continuous and categorical inputs. In SIAM CSE21

Models to handle mixed variables (continuous, discrete, categorical)

Toy function with a categorical variable (10 levels)

$$
f(x,z) = \begin{cases}
\cos(3.6\pi(x-2)) + x - 1 & \text{if} \quad z = 1, \\
2\cos(1.1\pi \exp(x)) - \frac{x}{2} + 2 & \text{if} \quad z = 2, \\
\cos(2\pi x) + \frac{1}{2}x & \text{if} \quad z = 3, \\
x\left(\cos(3.4\pi(x-1)) - \frac{x-1}{2}\right) & \text{if} \quad z = 4, \\
-\frac{x^2}{2} & \text{if} \quad z = 5, \\
2\cos(\frac{\pi}{4}\exp(-x^4))^2 - \frac{x}{2} + 1 & \text{if} \quad z = 6, \\
x\cos(3.4\pi x) - \frac{x}{2} + 1 & \text{if} \quad z = 7, \\
x(-\cos\left(7\frac{\pi}{2}x\right) - \frac{x}{2}) + 2 & \text{if} \quad z = 8, \\
-\frac{x^5}{2} + 1 & \text{if} \quad z = 9, \\
-\cos(5\frac{\pi}{2}x)^2\sqrt{x} - \frac{\ln(x+0.5)}{2} - 1.3 & \text{if} \quad z = 10.
\end{cases}
$$

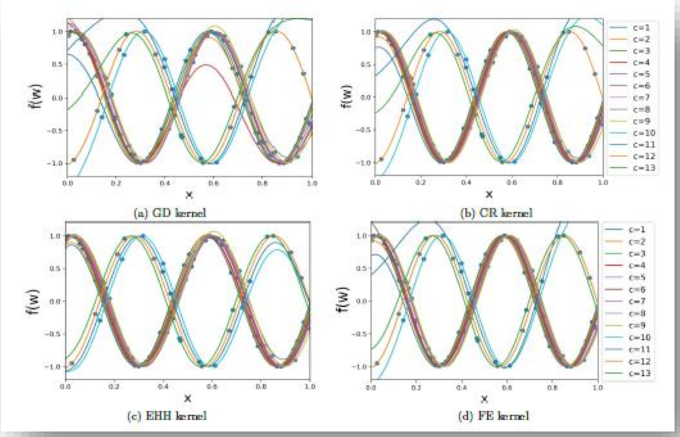


1 continuous + 1 categorical variable (10 levels) ➔ 11 continuous variables

Recent activities on mixed kernels

- **Mixed** kernels integration (Phd P. Saves)

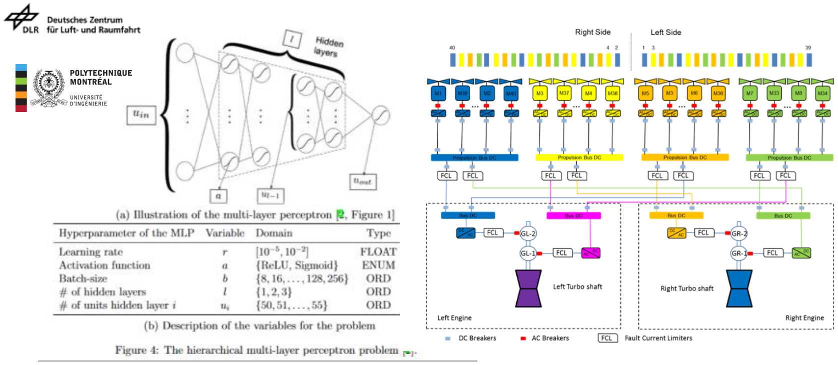
  Available kernels
  Continuous Relaxation,
  Gower distance,
  Homoscedastic hypersphere,
  Exponential Homoscedastic hypersphere
  + KPLS for dimension reduction with automatic choice for number of PLS components



(a) GD kernel (b) CR kernel (c) EHH kernel (d) FE kernel

- Extension to **hierarchical** variables (variable-size problems)

  Consider conditionally active distances
  New proposed kernels (Phd P. Saves & Collab. E. Hallé-Hannan Polytechnique Montréal)
  New application cases (Collab. J. H Bussemaker DLR)



(a) Illustration of the multi-layer perceptron [?, Figure 1]

Hyperparameter of the MLP	Variable	Domain	Type
Learning rate	r	$[10^{-5}, 10^{-2}]$	FLOAT
Activation function	a	{ReLU, Sigmoid}	ENUM
Batch-size	b	$\{8, 16, \ldots, 128, 256\}$	ORD
# of hidden layers	l	$\{1, 2, 3\}$	ORD
# of units hidden layer i	u_i	$\{50, 51, \ldots, 55\}$	ORD

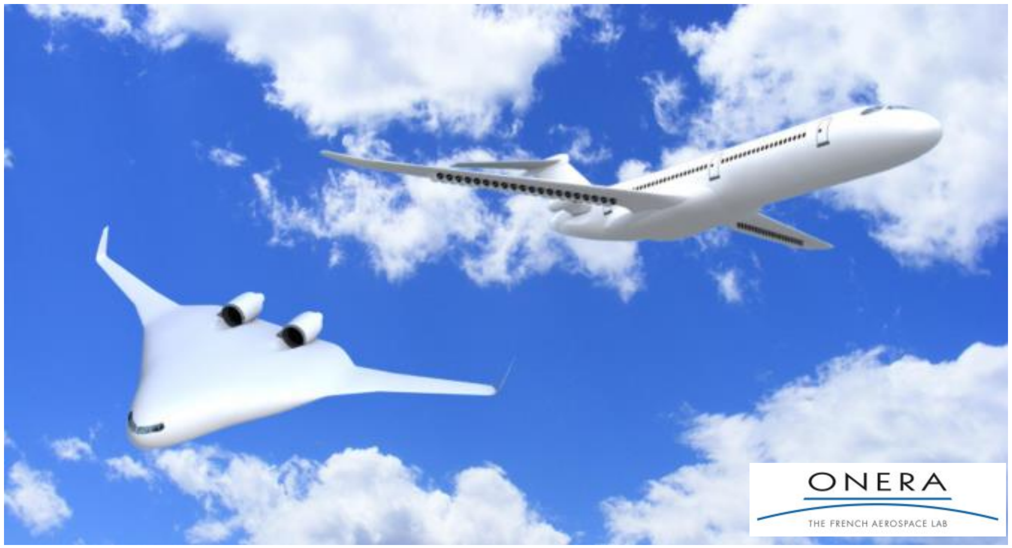
(b) Description of the variables for the problem

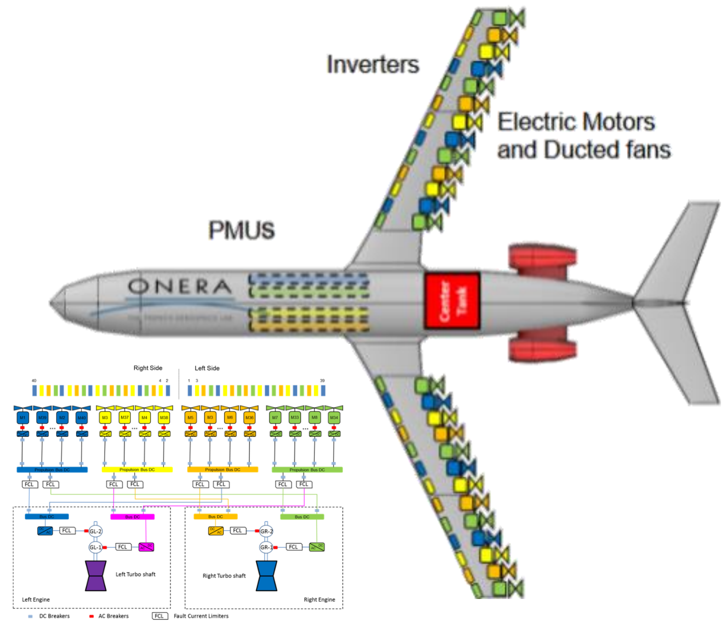
Figure 4: The hierarchical multi-layer perceptron problem [?].

Saves, P., Diouane, Y., Bartoli, N., Lefebvre, T., & Morlier, J. (2023). A mixed-categorical correlation kernel for Gaussian process. Neurocomputing, 550, 126472.
Bussemaker, J. H., Bartoli, N., Lefebvre, T., Ciampa, P. D., & Nagel, B. (2021). Effectiveness of Surrogate-Based Optimization Algorithms for System Architecture Optimization. In AIAA AVIATION 2021 FORUM (p. 3095).
Audet, C., Hallé-Hannan, E., & Le Digabel, S. (2023, February). A general mathematical framework for constrained mixed-variable blackbox optimization problems with meta and categorical variables. In Operations Research Forum (Vol. 4, No. 1, p. 12). Cham: Springer International Publishing.
Hallé-Hannan E, Audet A, Diouane Y, Le Digabel S., Saves P., A graph-structured distance for heterogeneous datasets with meta variables, 2024, Neurocomputing, Under review.

Multidisciplinary Design Analysis and Optimization for new configurations

Goal: Aircraft/drone design optimization





min objective function f(x,y(x))

with respect to design variables x (continuous, discrete, categorical, hierarchical)

subject to constraints g(x,y(x))

➔ y(x) are solution of a non linear system (MDA)

➔ Objective and constraint functions could be costly

Optimization problem in the field of aircraft design

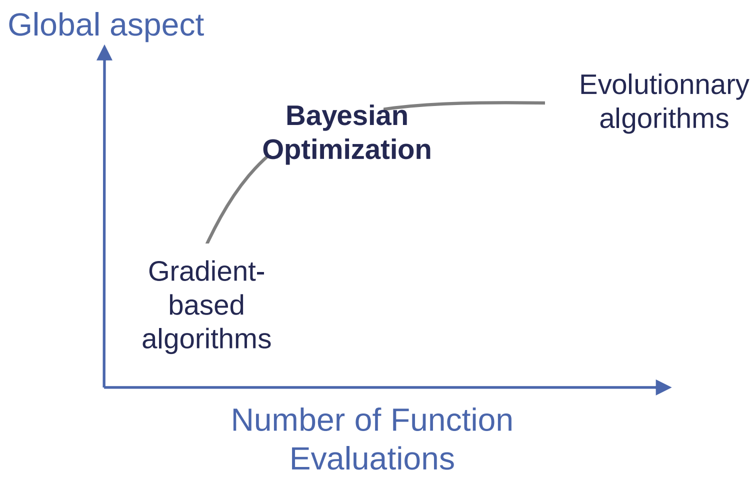
$$\begin{cases} \min\limits_{\boldsymbol{x}\in\mathbb{R}^d} \quad \boldsymbol{f}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_n(\boldsymbol{x})] \\ \text{s.t.} \\ c_1(\boldsymbol{x}) \leq 0 \ldots c_j(\boldsymbol{x}) = 0 \ldots c_m(\boldsymbol{x}) \leq 0 \end{cases}$$

1 to n objectives
d design variables
m mixed constraints (eq. & ineq)

- Main characteristics for aircraft design problem
  - **Mono & Multi** objective, multi-constraints (1 ~ 100 constraints)
  - Intermediate dimension problem (1 ~ 100 variables), mixed variables
  - Costly evaluation (CFD, FEM, objective and/or constraints)
  - Handling non linear constraints (black box, no derivative available)
  - Handling **hidden constraints**
- Applications
  - Disciplinary solvers (aerodynamic, structure, propulsion, …)
  - Overall aircraft design process (MDA)

How to build an efficient iterative process?

- Find the global minimum with a limited budget of function evaluations
- Use Bayesian information to detect interesting and promising areas (exploitation/exploration trade-off)

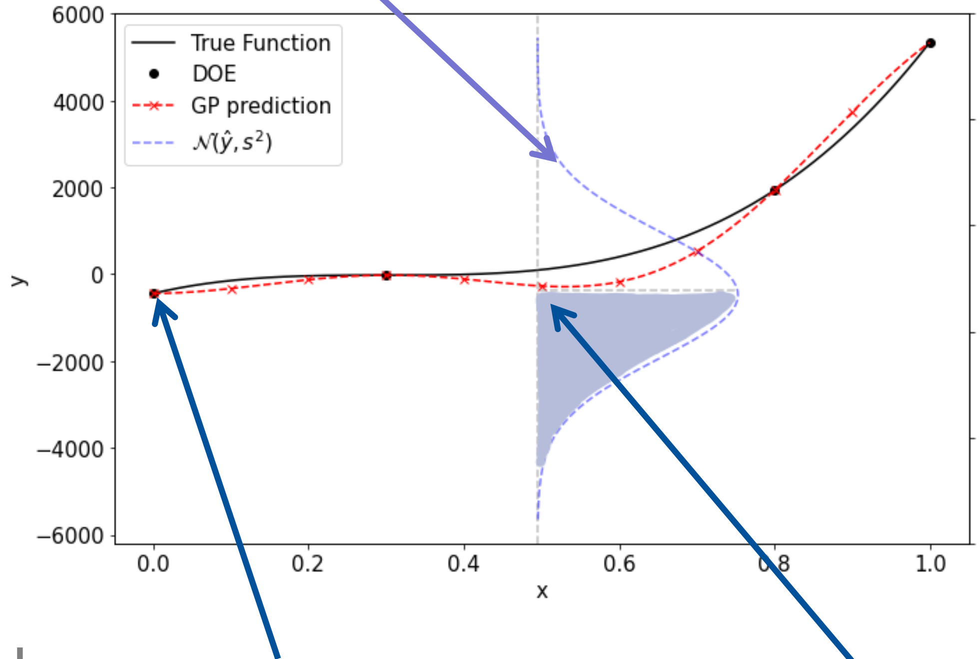


Bayesian optimization

1. **Probabilistic model** (surrogate model) uses data and Bayes theorem to compute posterior distribution

2. Optimization done via an **acquisition function** uses the posterior distribution to decide which data to obtain

$$Y(x) = N(\hat{y}(\boldsymbol{x}), s^2(\boldsymbol{x}))$$



Current Minimum y_{min}

Potential minimum

Example: BO to tune NN hyperparameters within AlphaGo

Chen, Y., Huang, A., Wang, Z., Antonoglou, I., Schrittwieser, J., Silver, D., & de Freitas, N. (2018). Bayesian optimization in alphago. *arXiv preprint arXiv:1812.06855*.

Enrichment infill sampling criterion

$f(x) \Rightarrow Y(x) = N(\hat{y}(\boldsymbol{x}), s^2(\boldsymbol{x}))$ **Expected Improvement criterion (EI)**

$$\mathrm{EI}(\mathbf{x}) = \mathbb{E}[\max(y_{\min} - Y(\mathbf{x}), 0)]$$

Kriging or Gaussian process of the objective function

Φ cumulative distribution function
ϕ probability density function of $\mathcal{N}(0,1)$

$$\mathrm{EI}(x) = (y_{min} - \hat{y}(\boldsymbol{x}))\Phi\left(\frac{y_{min} - \hat{y}(\boldsymbol{x})}{s(\boldsymbol{x})}\right) + s(\boldsymbol{x})\phi\left(\frac{y_{min} - \hat{y}(\boldsymbol{x})}{s(\boldsymbol{x})}\right)$$

Exploitation **Exploration**

Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient global optimization of expensive black-box functions," Journal of Global optimization, Vol. 13, No. 4, 1998, pp. 455–492.
Bartoli, N., Lefebvre, T., Dubreuil, S., Olivanti, R., Priem, R., Bons, N., Martins, J. R. A. A., Morlier, J. Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design. Aerospace Science and Technology, 90:85–102, 2019.

Enrichment infill sampling criterion

$f(x) \Rightarrow Y(x) = N(\hat{y}(x), s^2(x))$

Kriging or Gaussian process of the objective function

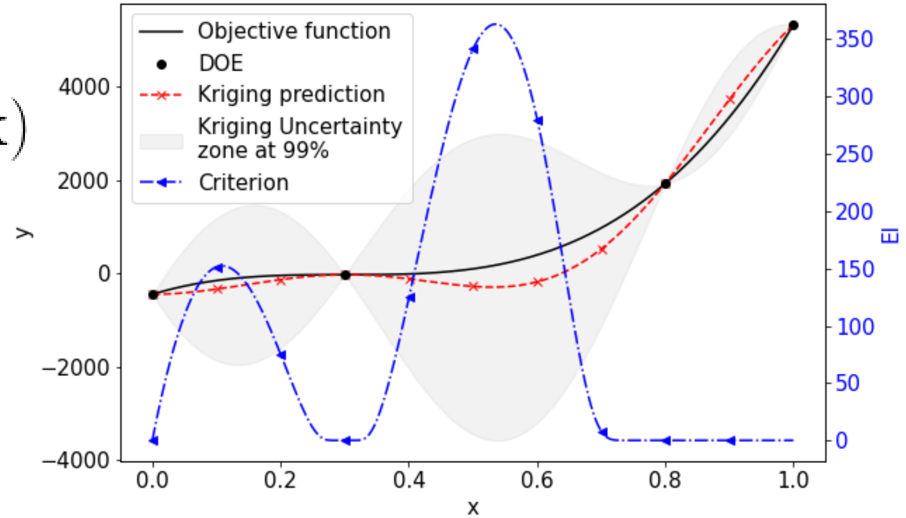
Expected Improvement criterion (EI)

$$\mathrm{EI}(\mathbf{x}) = \mathbb{E}[\max(y_{\min} - Y(\mathbf{x}), 0)]$$

$$\begin{cases} \min_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x}) \\ \text{s.t.} \\ c_1(\mathbf{x}) \leq 0 \\ \vdots \\ c_m(\mathbf{x}) \leq 0 \end{cases}$$

Surrogate models
(objective & constraints)

$$\begin{cases} \max_{\mathbf{x}\in\mathbb{R}^d} \mathrm{EI}(\mathbf{x}) \\ \text{s.t.} \\ \hat{c}_1(\mathbf{x}) \leq 0 \\ \vdots \\ \hat{c}_m(\mathbf{x}) \leq 0 \end{cases}$$



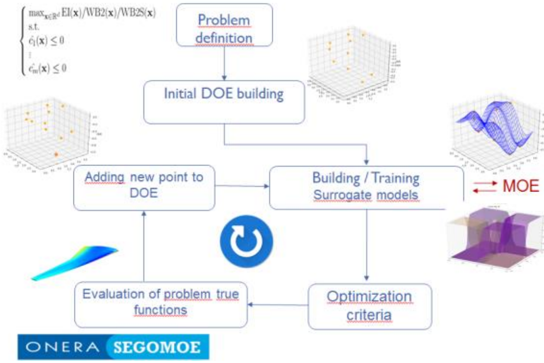
➔ **Different criteria available for the acquisition function (EI, WB2, WB2S)**

Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient global optimization of expensive black-box functions," Journal of Global optimization, Vol. 13, No. 4, 1998, pp. 455–492.
Bartoli, N., Lefebvre, T., Dubreuil, S., Olivanti, R., Priem, R., Bons, N., Martins, J. R. A. A., Morlier, J. Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design. Aerospace Science and Technology, 90:85–102, 2019.

SEGOMOE main characteristics

$$\begin{cases} \min\limits_{x \in \mathbb{R}^d} \quad f(x) = [f_1(x), f_2(x), \cdots, f_n(x)] \\ \text{s.t.} \\ c_1(x) \leq 0 \ldots c_j(x) = 0 \ldots c_m(x) \leq 0 \end{cases}$$

1 to n objectives
d design variables
m mixed constraints



- **Mono** & **multi objective** Bayesian optimizer
- **Mono** & **Multi fidelity** sources
- Equality & inequality constraints (1 ~ 100 constraints)
- Intermediate dimension problem (1 ~ 100 variables)
- **Heterogenous variables** (continuous, discrete, categorical, hierarchical)
- Costly evaluation (CFD, FEM, objective and/or constraints)
- Handling non linear objectives & constraints (black box, no derivative available) and **hidden constraints**
- Based on **SMT** toolbox for surrogate models

- Remote access via a web interface ONERA WhatsOpt
- Opensource version in EgoBox (Mono Obj & Mono Fidelity)

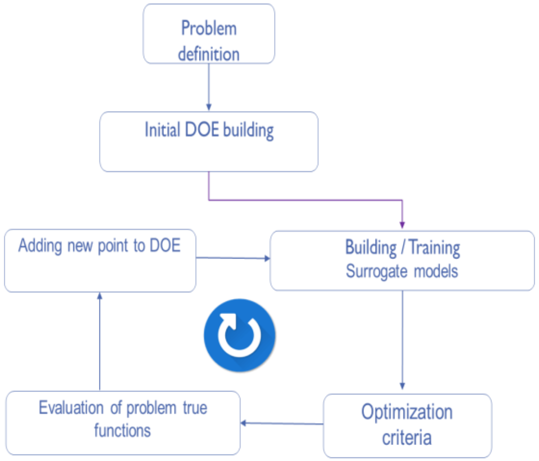
The Journal of Open Source Software
egobox, a Rust toolbox for efficient global optimization

Bartoli, N., Lefebvre, T., Dubreuil, S., Olivanti, R., Priem, R., Bons, N., Martins, J.R.R.A., Bouhlel, M.-A. Bouhlel & Morlier, J. (2019). Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design. Aerospace Science and technology, 90, 85-102.
Lafage, R. (2022). egobox, a Rust toolbox for efficient global optimization. Journal of Open Source Software, 7(78), 4737

SEGOMOE

SEGO

Super Efficient Global Optimization

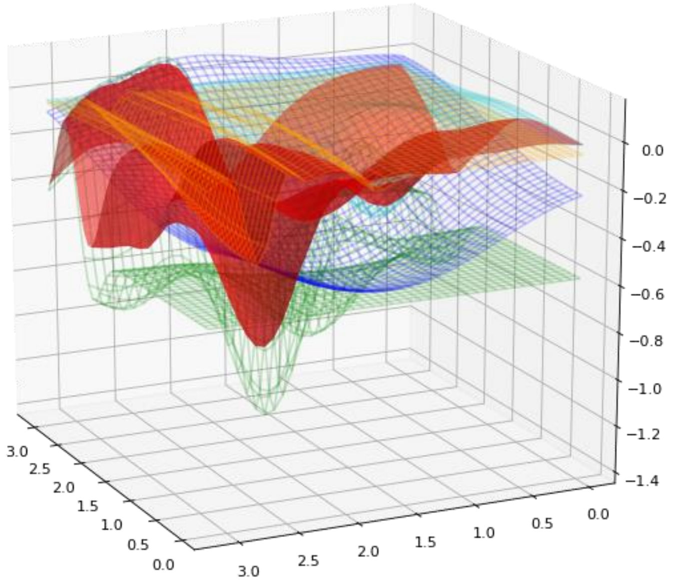


Global optimization with limited number of function evaluations

Jones, D. R., Schonlau, M., and Welch, W. J., “Efficient global optimization of expensive black-box functions,” Journal of Global optimization, Vol. 13, No. 4, 1998, pp. 455–492.
Sasena, M., Flexibility and efficiency enhancements for constrained global design optimization with Kriging approximations, Ph.D. thesis, niversity of Michigan, 2002

MOE

Mixture Of Experts

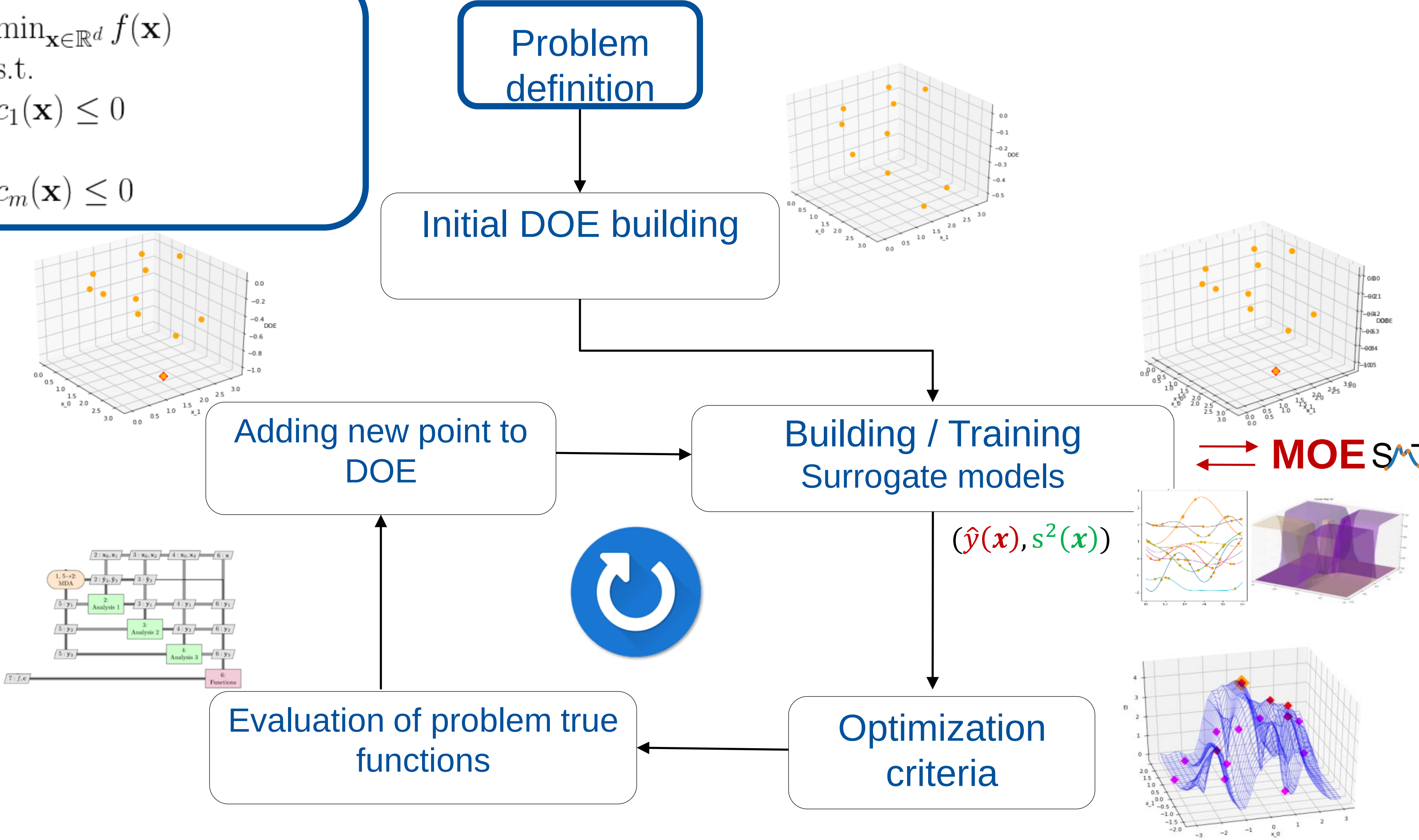


Combination of surrogate models

Jordan, M. I., Jacobs, R. A, “Hierarchical mixtures of experts and the EM algorithm“, Neural Comput. 6 (1994) 181–214.
Bettebghor, D., Bartoli, N., Grihon, S., Morlier, J., and Samuelides, M., “Surrogate modeling approximation using a mixture of experts based on EM joint estimation,” Structural and Multidisciplinary Optimization, Vol. 43, No. 2, 2011, pp. 243–259
Liem, R. P., Mader, C. A., and Martins, J. R. R. A., “Surrogate Models and Mixtures of Experts in Aerodynamic Performance Prediction for Mission Analysis,” Aerospace Science and Technology, Vol. 43, 2015, pp. 126–151

SEGOMOE algorithm – Mono objective

$$\begin{cases} \min_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x}) \\ \text{s.t.} \\ c_1(\mathbf{x}) \leq 0 \\ \vdots \\ c_m(\mathbf{x}) \leq 0 \end{cases}$$



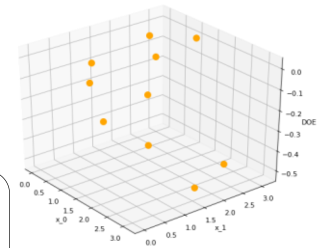
Bartoli, N., Lefebvre, T., Dubreuil, S., Olivanti, R., Priem, R., Bons, N., Martins, J.R.R.A., Bouhlel, M.-A. Bouhlel & Morlier, J. (2019). Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design. Aerospace Science and technology, 90, 85-102.

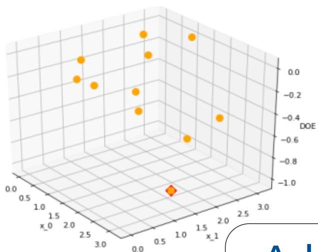
SEGOMOE algorithm – Mono objective

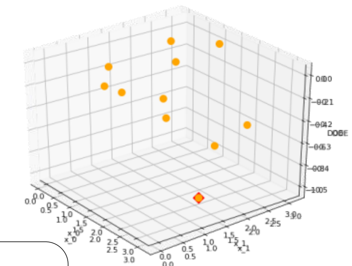
$$\begin{cases} \max_{\mathbf{x}\in\mathbb{R}^d} \mathrm{EI}(\mathbf{x})/\mathrm{WB2}(\mathbf{x})/\mathrm{WB2S}(\mathbf{x}) \\ \text{s.t.} \\ \hat{c_1}(\mathbf{x}) \leq 0 \\ \vdots \\ \hat{c_m}(\mathbf{x}) \leq 0 \end{cases}$$

Problem definition

Initial DOE building



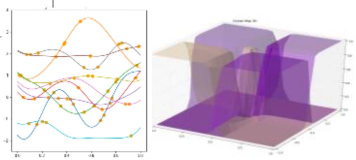




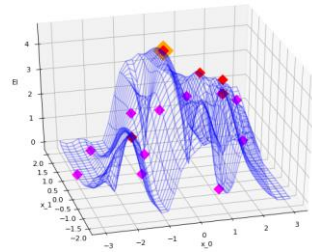
Building / Training Surrogate models

MOE SMT

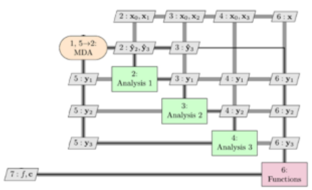
$(\hat{y}(x), s^2(x))$



Optimization criteria



Evaluation of problem true functions



Adding new point to DOE

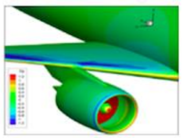
Bartoli, N., Lefebvre, T., Dubreuil, S., Olivanti, R., Priem, R., Bons, N., Martins, J.R.R.A., Bouhlel, M.-A. Bouhlel & Morlier, J. (2019). Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design. Aerospace Science and technology, 90, 85-102.

Some SEGOMOE application examples

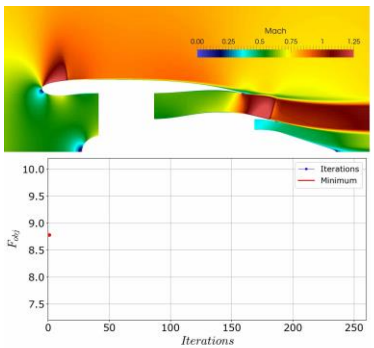
Phd M-A Bouhlel 2016, R. Priem 2020, R. Charayron 2023, P. Saves 2024

Nacelle aerodynamic

Mono, Continuous
d = 18, c = 2

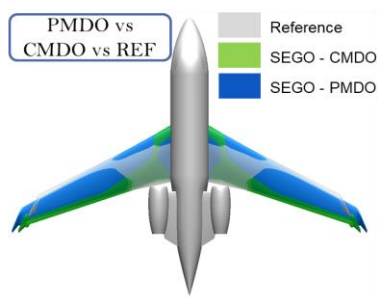


Mach number distribution



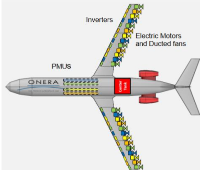
Bombardier Research Aircraft Configuration

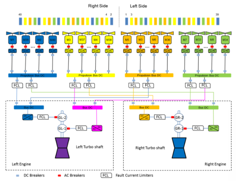
Mono, Continuous
d = 18, c = 8

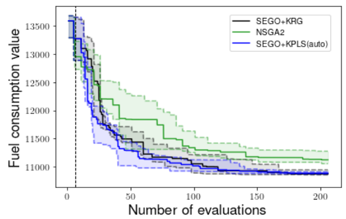


Dragon hybrid electric aircraft

Mono, Mixed
d = 29, c = 5

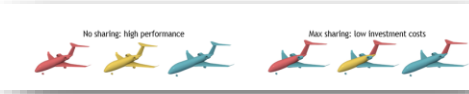


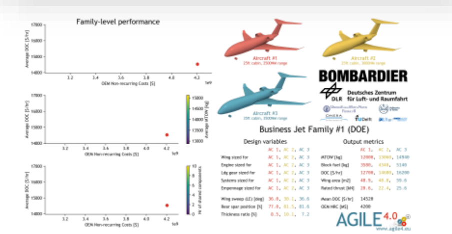




Business jet family

Multi, Mixed
d = 19, c = 2





Bartoli N, Lefebvre T, Dubreuil S, Panzeri M, d'Ippolito R, Anisimov K, Savelyev A. *Robust Nacelle Optimization design investigated in the AGILE European Project,* 19th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, AIAA AVIATION Forum, (AIAA 2018-3250)

Priem, R., Gagnon, H., Chittick, I., Dufresne, S., Diouane, Y., & Bartoli, N. (2020). An efficient application of Bayesian optimization to an industrial MDO framework for aircraft design. In AIAA aviation 2020 forum (p. 3152).

Saves, P., Bartoli, N., Diouane, Y., Lefebvre, T., Morlier, J., David, C., ... & Defoort, S. (2022). Bayesian optimization for mixed variables using an adaptive dimension reduction process: applications to aircraft design. In AIAA SCITECH 2022 Forum (p. 0082).

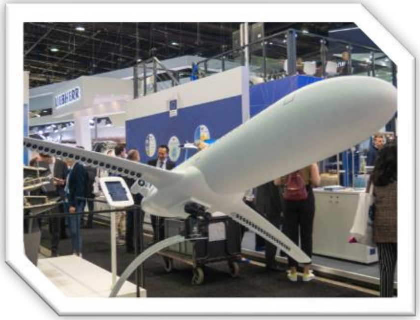
Bussemaker, J. H., Ciampa, P. D., Singh, J., Fioriti, M., Cabaleiro De La Hoz, C., Wang, Z., ... & Mandorino, M. (2022). Collaborative Design of a Business Jet Family Using the AGILE 4.0 MBSE Environment. In AIAA Aviation 2022 Forum (p. 3934).

Recent methodological developments

SEGOMOE capabilities

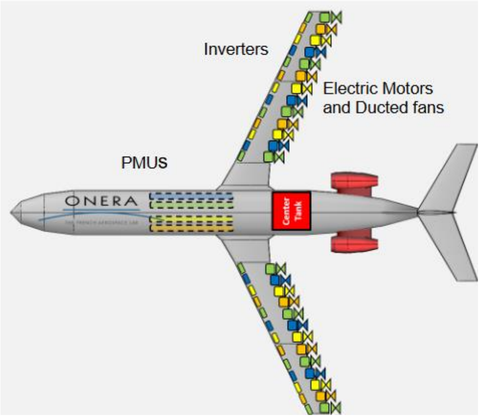
- To handle a large number of design variables
  → KPLS based models
- To handle heterogeneous functions
  → Mixture of experts models
- To handle highly non-convex constraints
  → Adapted acquisition function
- To handle mixed integer variables
  → Continuous relaxation & KPLS models
- To handle multifidelity models
  → 2-step approach based on multifidelity Kriging
- To handle multiple objectives
  → Predicted Pareto Front approach
- To handle hidden constraints
  → Comparisons of different strategies

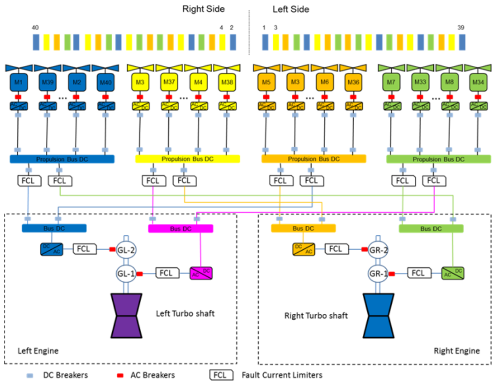
Application to mixed categorical optimization problem



DRAGON green aircraft concept

- ✓ **30% reduction of CO2 emissions by 2035**
- ✓ **Distributed electric propulsion aircraft: propulsive efficiency**
- ✓ **150 passengers over 2750nm**
- ✓ **Transonic cruise speed (M0.78)**



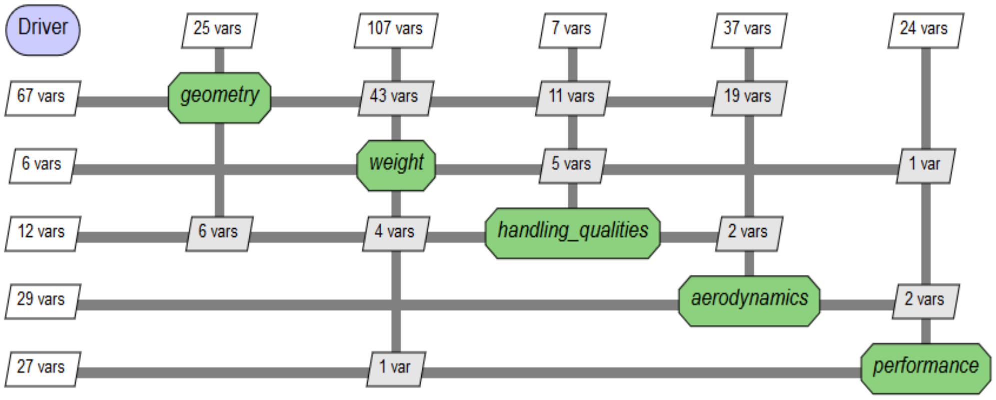


Phd P. Saves in collaboration with E. Nguyen Van, C. David, S. Defoort

P. Schmollgruber, C. Doll, J. Hermetz, R. Liaboeuf, M. Ridel, I. Cafarelli, O. Atinault, C. Francois, and B. Paluch. “Multidisciplinary Exploration of DRAGON: an ONERA Hybrid Electric Distributed Propulsion Concept”. In: AIAA Scitech 2019, 2019

FAST-OAD: an OpenMDAO based aircraft sizing tool

Code overview



https://github.com/fast-aircraft-design/FAST-OAD

FAST OAD

Future Aircraft Sizing Tool - Overall Aircraft Design

FastOAD* conceptual design framework:

- OpenSource framework developed by ONERA/ISAE-SUPAERO
- Based on OpenMDAO
- Automates MDA/MDO for simple and rapid OAD studies, concept evaluation and optimization
- Includes Level 0 disciplinary models for transport aircraft (geometry, weight, HQ, aerodynamics, mission/performance…)
- Modularity of each discipline model to include higher fidelity modelling

David C., Delbecq S. Defoort S., Schmollgruber P., Benard E., Pommier-Budinger V., From FAST to FAST-OAD: An open source framework for rapid Overall Aircraft Design, 2021 IOP Conf. Ser.: Mater. Sci. Eng.1024 012062.

DRAGON green aircraft concept

	Function/variable	Nature	Quantity	Range
Minimize	Fuel mass	cont	1	
	Total objectives		**1**	
with respect to	Fan operating pressure ratio	cont	1	$[1.05, 1.3]$
	Wing aspect ratio	cont	1	$[8, 12]$
	Angle for swept wing	cont	1	$[15, 40]$ $(°)$
	Wing taper ratio	cont	1	$[0.2, 0.5]$
	HT aspect ratio	cont	1	$[3, 6]$
	Angle for swept HT	cont	1	$[20, 40]$ $(°)$
	HT taper ratio	cont	1	$[0.3, 0.5]$
	TOFL for sizing	cont	1	$[1800., 2500.]$ (m)
	Top of climb vertical speed for sizing	cont	1	$[300., 800.](ft/min)$
	Start of climb slope angle	cont	1	$[0.075., 0.15.](rad)$
	Total continuous variables		10	
	Turboshaft layout	cat	2 levels	{1,2}
2 possibilities	Architecture_cat	cat	17 levels	{1,2,3, ...,15,16,17}
	Number of cores	int	1	{2,4,6}
	Number of motors*	int	1	{8,12,16,20,...,40}
	*graph-structure dependence to the core value			
subject to	Wing span < 36 (m)	cont	1	
	TOFL < 2200 (m)	cont	1	
	Wing trailing edge occupied by fans < 14.4 (m)	cont	1	
	Climb duration < 1740 (s)	cont	1	
	Top of climb slope > 0.0108 (rad)	cont	1	
	Total constraints		**5**	

Categorical
or
Hierarchical

- ***10 continuous design variables***
- ***2 categorical design variables***
  - ***Electric propulsion Architecture: 17 choices***
  - ***Turboshaft layout: 2 choices***

→ 29 variables in relaxed dimension
→ 14 variables in relaxed dimension

- **5 inequality constraints (MC)**
- **Fuel mass to minimize**

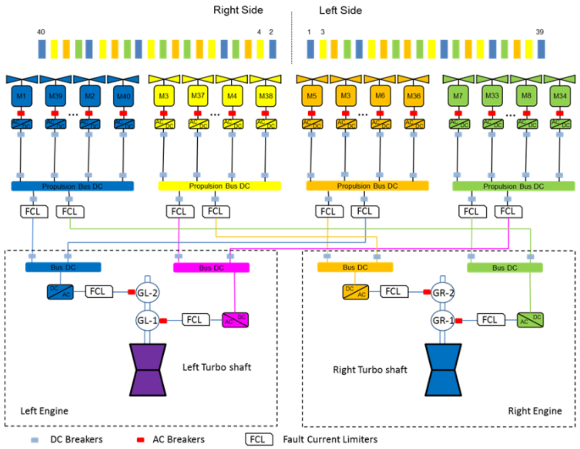
Saves, P., Bartoli, N., Diouane, Y., Lefebvre, T., Morlier, J., David, C., ... & Defoort, S. (2022). Bayesian optimization for mixed variables using an adaptive dimension reduction process: applications to aircraft design. In *AIAA SCITECH 2022 Forum* (p. 0082).

DRAGON optimization test case

Architecture	cat	17 levels	{1,2,3, . . . ,15,16,17}
Turboshaft layout	cat	2 levels	{1,2}
Total categorical variables		2	
Total relaxed variables		**29**	

Architecture number	number of generators	number of motors
1	2	8
2	2	12
3	2	16
4	2	20
5	2	24
6	2	28
7	2	32
8	2	36
9	2	40
10	4	8
11	4	16
12	4	24
13	4	32
14	4	40
15	6	12
16	6	24
17	6	36

Neutral



layout	position	y ratio	tail	VT aspect ratio	VT taper ratio
1	under wing	0.25	without T-tail	1.8	0.3
2	behind	0.34	with T-tail	1.2	0.85

2 possibilities

➔Categorical choice:
29 variables in relaxed dimension (10+17+2)

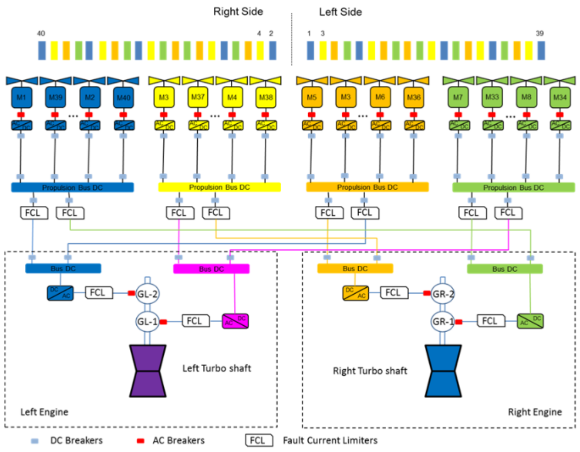
Saves, P., Bartoli, N., Diouane, Y., Lefebvre, T., Morlier, J., David, C., ... & Defoort, S. (2022). Bayesian optimization for mixed variables using an adaptive dimension reduction process: applications to aircraft design. In *AIAA SCITECH 2022 Forum* (p. 0082).

DRAGON optimization test case

Architecture	cat	17 levels	{1,2,3, . . . ,15,16,17}
Turboshaft layout	cat	2 levels	{1,2}
Total categorical variables		2	
Total relaxed variables		**29**	

Architecture number	number of generators	number of motors
1		8
2		12
3		16
4		20
5	2	24
6		28
7		32
8		36
9		40
10		8
11		16
12	4	24
13		32
14		40
15		12
16	6	24
17		36

Neutral Meta Decreed



layout	position	y ratio	tail	VT aspect ratio	VT taper ratio
1	under wing	0.25	without T-tail	1.8	0.3
2	behind	0.34	with T-tail	1.2	0.85

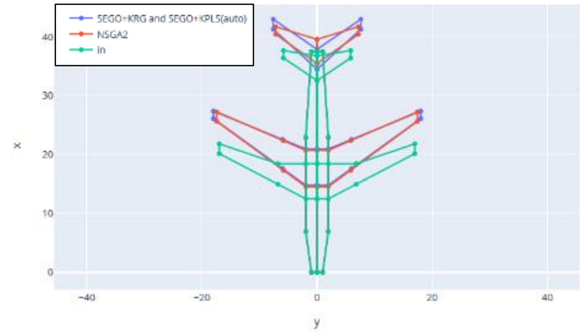
2 possibilities

➔**Categorical choice:**
29 variables in relaxed dimension (10+17+2)

➔**Hierarchical choice:**
14 variables in relaxed dimension (10+2+2)

Saves, P., Bartoli, N., Diouane, Y., Lefebvre, T., Morlier, J., David, C., ... & Defoort, S. (2022). Bayesian optimization for mixed variables using an adaptive dimension reduction process: applications to aircraft design. In *AIAA SCITECH 2022 Forum* (p. 0082).

DRAGON optimization results



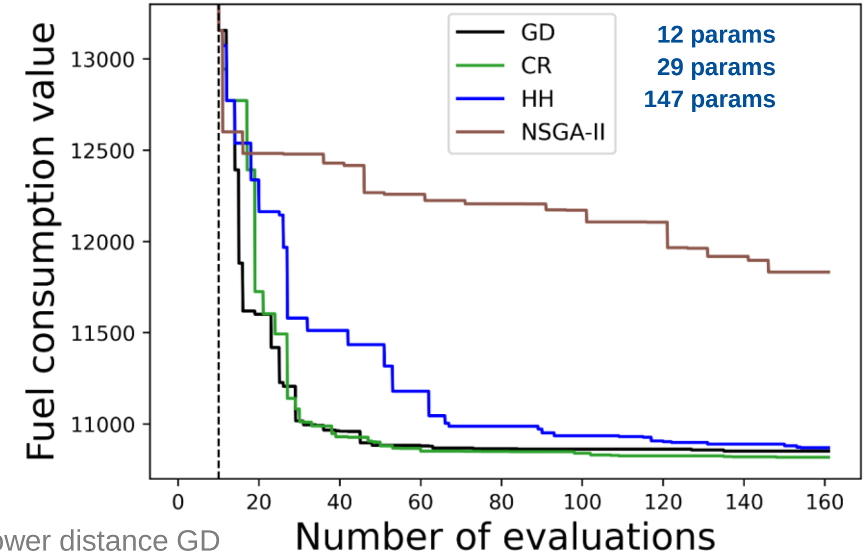
Comparison of BO with different kernels & NSGAII

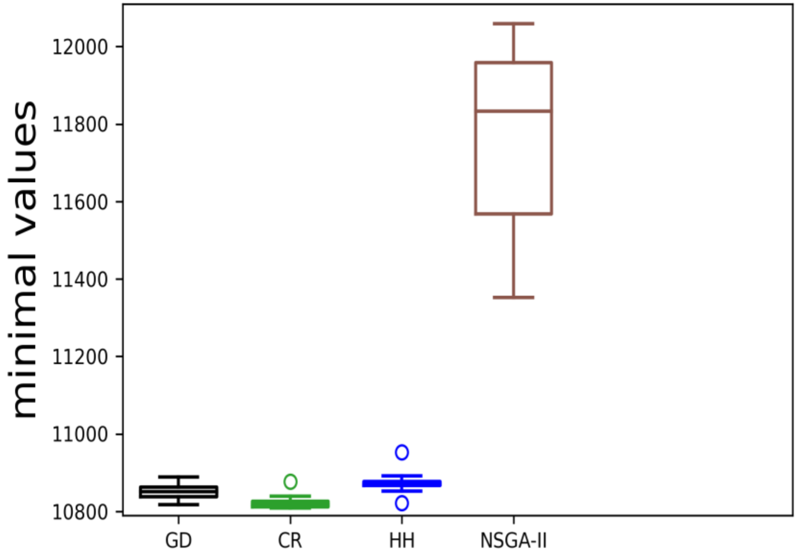
10 runs of (10 + 150) iterations

Convergence plots

Categorical variants without PLS

Boxplots after 160 evaluations

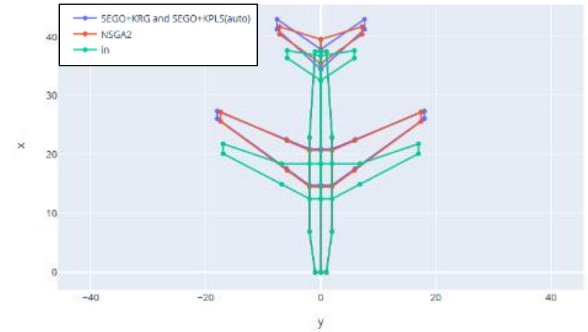




Gower distance GD
Continuous Relaxation CR
Homoscedastic hypersphere HH

Saves, P., Bartoli, N., Diouane, Y., Lefebvre, T., Morlier, J., David, C., ... & Defoort, S. (2022). Bayesian optimization for mixed variables using an adaptive dimension reduction process: applications to aircraft design. In *AIAA SCITECH 2022 Forum* (p. 0082)
Blank J. and Deb K., pymoo: Multi-Objective Optimization in Python, in IEEE Access, vol. 8, pp. 89497-89509, 2020, doi: 10.1109/ACCESS.2020.2990567

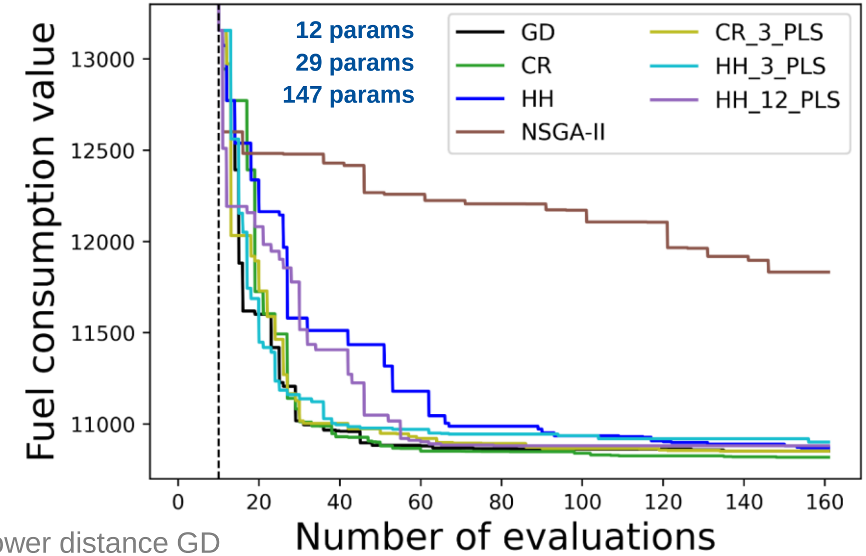
DRAGON optimization results



Comparison of BO with different kernels & NSGAII

10 runs of (10 + 150) iterations

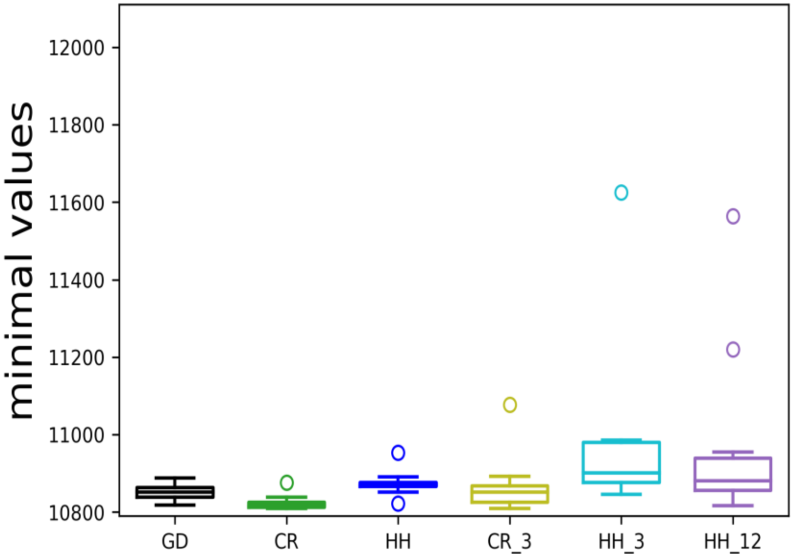
Convergence plots



Categorical variants with PLS

3 params
3 params
12 params

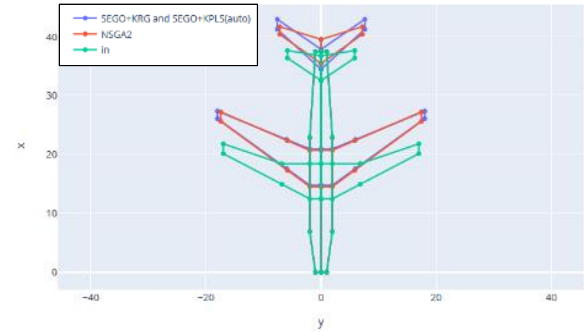
Boxplots after 160 evaluations



Gower distance GD
Continuous Relaxation CR
Homoscedastic hypersphere HH

Saves, P., Bartoli, N., Diouane, Y., Lefebvre, T., Morlier, J., David, C., ... & Defoort, S. (2022). Bayesian optimization for mixed variables using an adaptive dimension reduction process: applications to aircraft design. In *AIAA SCITECH 2022 Forum* (p. 0082)
Blank J. and Deb K., pymoo: Multi-Objective Optimization in Python, in IEEE Access, vol. 8, pp. 89497-89509, 2020, doi: 10.1109/ACCESS.2020.2990567

DRAGON optimization results



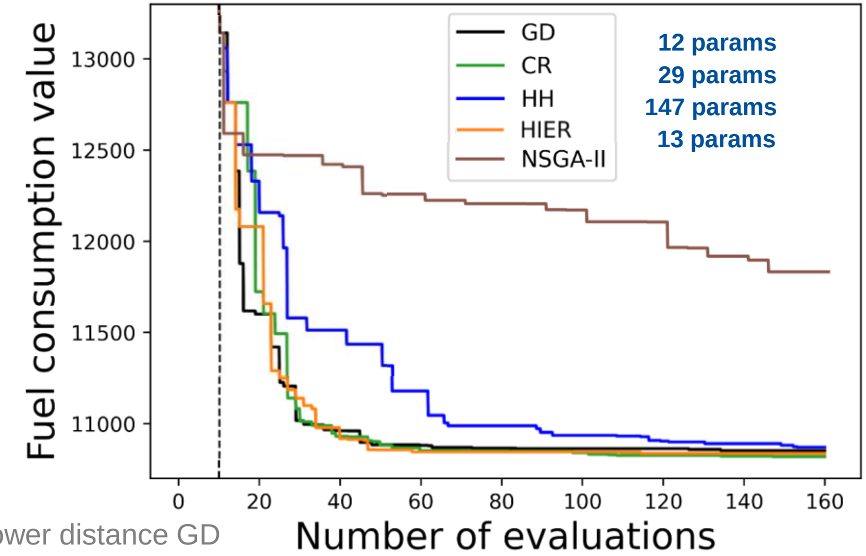
Comparison of BO with different kernels & NSGAII

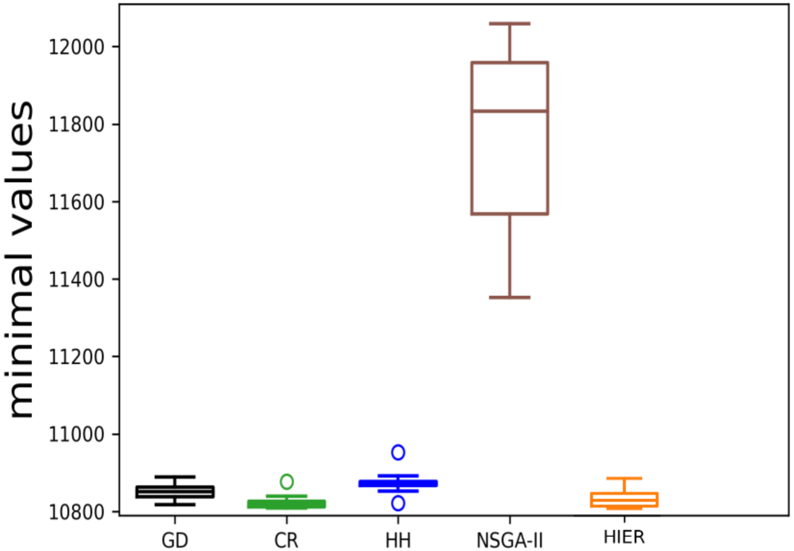
10 runs of (10 + 150) iterations

Convergence plots

Categorical or Hierarchical variants

Boxplots
after 160 evaluations

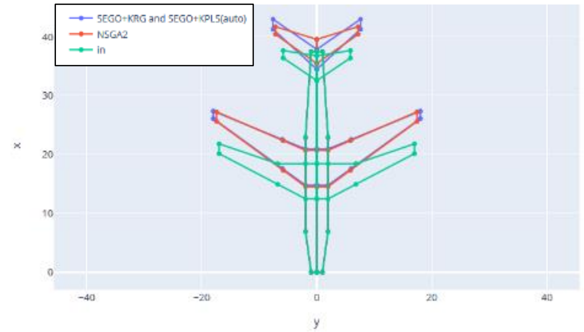




Gower distance GD
Continuous Relaxation CR
Homoscedastic hypersphere HH
Hierarchical HIER

Saves, P., Bartoli, N., Diouane, Y., Lefebvre, T., Morlier, J., David, C., ... & Defoort, S. (2022). Bayesian optimization for mixed variables using an adaptive dimension reduction process: applications to aircraft design. In *AIAA SCITECH 2022 Forum* (p. 0082)
Blank J. and Deb K., pymoo: Multi-Objective Optimization in Python, in IEEE Access, vol. 8, pp. 89497-89509, 2020, doi: 10.1109/ACCESS.2020.2990567

DRAGON optimization results

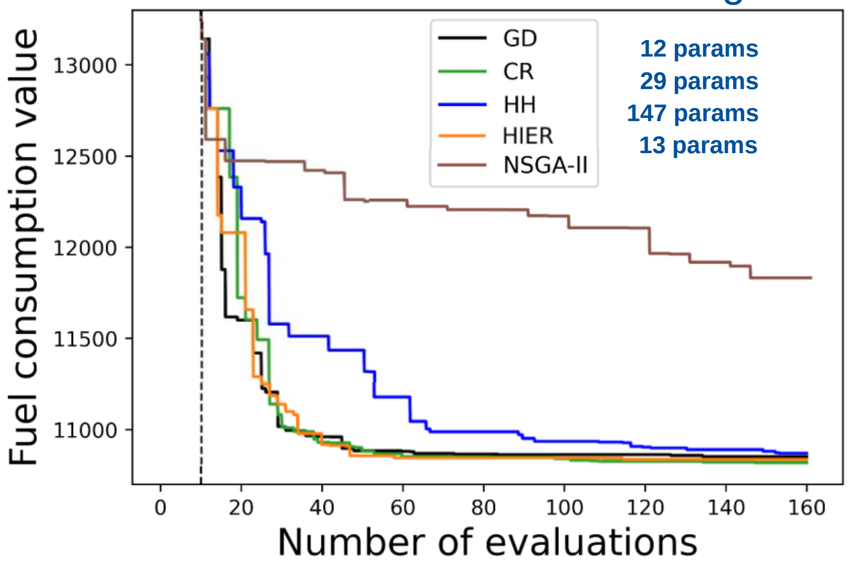


Comparison of BO with different kernels & NSGAII

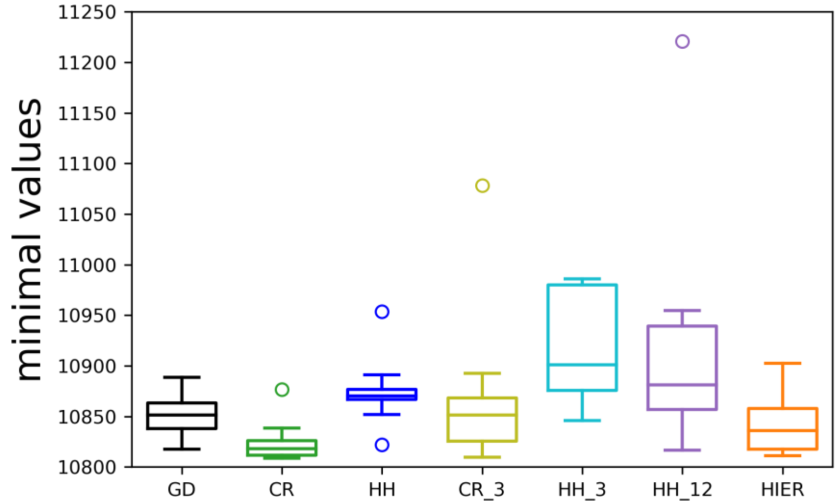
10 runs of (10 + 150) iterations

Convergence plots

Categorical or Hierarchical variants



Boxplots
after 160 evaluations



➔Hierarchical choice: best trade off convergence & CPU time

Saves, P., Bartoli, N., Diouane, Y., Lefebvre, T., Morlier, J., David, C., ... & Defoort, S. (2022). Bayesian optimization for mixed variables using an adaptive dimension reduction process: applications to aircraft design. In *AIAA SCITECH 2022 Forum* (p. 0082).

How to handle hidden constraints?

A failed simulation in an optimization:

a simulation that terminates unexpectedly resulting in an error (NaN) in the outcomes (objectives or constraints)

- $f_m(x) \rightarrow NaN$ and/or $g_k(x) \rightarrow NaN$
- Deterministic and requires an evaluation
- Non-quantifiable and unrelaxable
- Hidden (hence “hidden constraint”)

➔called hidden or unknown constraints

➔Bayesian Optimization (BO) to adapt

Hidden constraints also known as:

Unknown, unspecified, forgotten, virtual, and crash constraints

Le Digabel, S., & Wild, S. M. (2024). A taxonomy of constraints in black-box simulation-based optimization. Optimization and Engineering, 25(2), 1125-1143.

How to handle hidden constraints?

- Different strategies:
  – Remove the failed points from the DOE and add a constraint to avoid neighborhoods around these points
  – Use some techniques to avoid the expensive computation

➔ Collaboration with J. H. Bussemaker (DLR)

➔ Collaboration with A. Tfaily (Bombardier)

Tfaily, A., Diouane, Y., Bartoli, N., & Kokkolaras, M. (2024). Bayesian optimization with hidden constraints for aircraft design. Structural and Multidisciplinary Optimization, 67(7), 123.
Bussemaker, J. H., Saves, P., Bartoli, N., Lefebvre, T., & Nagel, B. (2024). Surrogate-Based Optimization of System Architectures Subject to Hidden Constraints. In AIAA AVIATION FORUM AND ASCEND 2024

Strategies for satisfying hidden constraints

BOMBARDIER DLR

- Cannot train a surrogate model on NaN
- Naive approach: **reject** failed points and train only using viable points
- **Replace** failed points
  - Neighborhood values
  - Predicted values (different values α)

$$y_{\text{replace}}\left(\mathbf{x}_{\text{failed}}\right) = \hat{y}\left(\mathbf{x}_{\text{failed}}\right) + \alpha \cdot \hat{s}\left(\mathbf{x}_{\text{failed}}\right)$$

- **Predict** location of failed region
  - Train model to predict the Probability of Viability (*PoV*)
    - Binary labels for each x: 0 = failed, 1 = viable
    - Classification model → $PoV(x')$ = probability that x' belongs to class "1"
    - Regression model → $PoV(x') = \hat{y}(x')$
  - Apply as f-penalty (modify acquisition function)
  - or as PoV_{min} constraint in infill optimization

$$f_{\text{m,infill,mod}}(\mathbf{x}) = 1 - \left(\left(1 - f_{\text{m,infill}}(\mathbf{x})\right) \cdot \text{PoV}(\mathbf{x})\right)$$

$$g_{\text{PoV}}(\mathbf{x}) = \text{PoV}_{\text{min}} - \text{PoV}(\mathbf{x}) \leq 0$$

Bussemaker, J. H., Saves, P., Bartoli, N., Lefebvre, T., & Nagel, B. (2024). Surrogate-Based Optimization of System Architectures Subject to Hidden Constraints. In AIAA AVIATION FORUM AND ASCEND 2024

Forrester, A. I., et al., "Optimization with missing data," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2006.

Lee, H., et al., "Optimization Subject to Hidden Constraints via Statistical Emulation," UC Santa Cruz, Apr. 2010.

Huyer, W., and Neumaier, A., "SNOBFIT – Stable Noisy Optimization by Branch and Fit," ACM Transactions on Mathematical Software, Vol. 35, No. 2, 2008, pp. 1–25.

Alimo, S. R., Beyhaghi, P., and Bewley, T. R., "Delaunay-Based Global Optimization in Nonconvex Domains Defined by Hidden Constraints," Computational Methods in Applied Sciences, Springer International Publishing, 2018, pp. 261–271.

Strategies for satisfying hidden constraints

BOMBARDIER DLR

- Cannot train a surrogate model on NaN
- Naive approach: **reject** failed points and train only using viable points
- **Replace** failed points
  - Neighborhood values
  - Predicted values (different values α)
- **Predict** location of failed region
  - Train model to predict the Probability of Viability (*PoV*)
    - Binary labels for each x: 0 = failed, 1 = viable
    - Classification model → $PoV(x')$ = probability that x' belongs to class "1"
    - Regression model → $PoV(x') = \hat{y}(x')$
  - Apply as f-penalty (modify acquisition function)
  - or as PoV_{min} constraint in infill optimization

Collaboration J. H. Bussemaker
Comparisons between SOTA different strategies

Collaboration A. Tfaily
Modification of the acquisition function

Comparison of hidden constraint strategies

DLR

Collaboration with J. H. Bussemaker (DLR)

- 18 test problems (mono and multi-obj)
  - 2 – 9 continuous x; 0 – 6 discrete x
  - 1 – 3 objectives; 1 – 9 constraints
  - With and without hierarchy
  - 0% – 83% failure rate
- BO settings
  - $n_{doe} = (2 \cdot n_x)/(1 - 60\%)$
  - $n_{infill} = 50$
  - 16 repetitions
  - SBArchOpt* implementation
- Performance comparison using HyperVolume regret ΔHV
  - Integral over $\Delta HV_i = (HV_{known} - HV_i)/HV_{known}$
  - Ranking per test problem: rank 1 has the best (lowest) ΔHV regret
  - Best strategy achieves rank 1 and 2 most often

Strategy	Sub-strategy	Configuration
Rejection		
Replacement	Neighborhood	Global, max
		Local
		5-nearest, max
		5-nearest, mean
	Predicted worst	$\alpha = 1$
		$\alpha = 2$
Prediction	Random Forest Classifier	$PoV_{min} = 50\%$
	K-Nearest Neighbors	$PoV_{min} = 50\%$
	Radial Basis Function	$PoV_{min} = 50\%$
	GP Classifier	$PoV_{min} = 50\%$
	Variational GP	$PoV_{min} = 50\%$
	Mixed-discrete GP	$PoV_{min} = 50\%$

All test problems, experiments, and algorithms are available open source!

SBArchOpt

https://sbarchopt.readthedocs.io/

Strategy comparison results

Collaboration with J. H. Bussemaker (DLR)

- Best strategies:
  1. Prediction using a Random Forest Classifier
  2. Prediction using a Mixed-discrete GP
     - GP with categorical kernels
  3. Replacement (predicted $\alpha = 1$)
- Training + infill times are increased by 70% – 90% (compared to rejection)

ΔHV regret

		Rank 1	Rank ≤ 2
Rejection		11%	33%
Replacement	Global max	11%	44%
Replacement	Local	6%	39%
Replacement	5-nearest, max	22%	67%
Replacement	5-nearest, mean	39%	83%
Replacement	**Predicted worst (α=1)**	56%	83%
Replacement	**Predicted worst (α=2)**	39%	72%
Prediction	RFC	78%	94%
Prediction	KNN	17%	67%
Prediction	RBF	50%	83%
Prediction	GP Classifier	61%	78%
Prediction	Variational GP	56%	78%
Prediction	MD GP	72%	94%

RFC: Random Forest Classifier (RFC)
kNN: k Nearest Neigbors
RBF: Radial Basis Function
MD GP: Mixed Discrete GP

Breiman, L. (2001). “Random Forests.” Machine Learning, 45, 5–32.
Saves, P., et al. "A general square exponential kernel to handle mixed-categorical variables for Gaussian process," AIAA AVIATION 2022 Forum, Chicago, USA, 2022.

Parameter studies

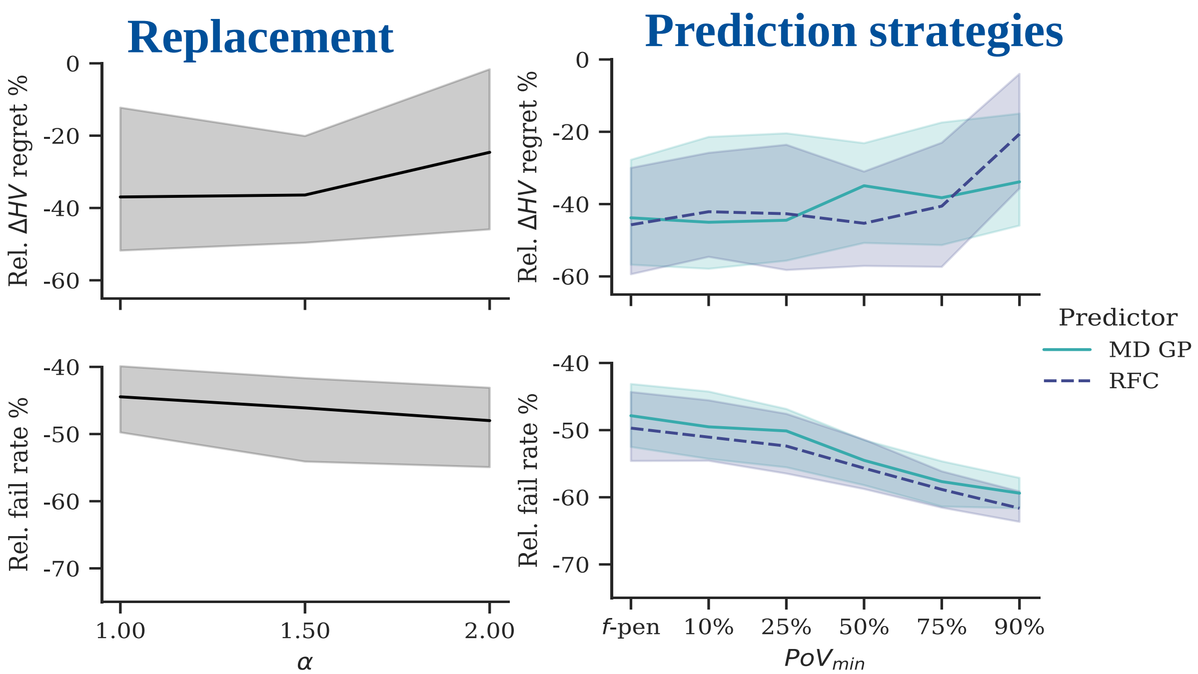
Collaboration with J. H. Bussemaker (DLR)

- Predicted worst replacement
  - Increasing α → more conservative, less exploration, less performance

$$y_{\text{replace}}\left(\mathbf{x}_{\text{failed}}\right) = \hat{y}\left(\mathbf{x}_{\text{failed}}\right) + \alpha \cdot \hat{s}\left(\mathbf{x}_{\text{failed}}\right)$$

- Prediction
  - Increasing PoV_{min} → same trends
  - Best at low PoV_{min} or as f-penalty

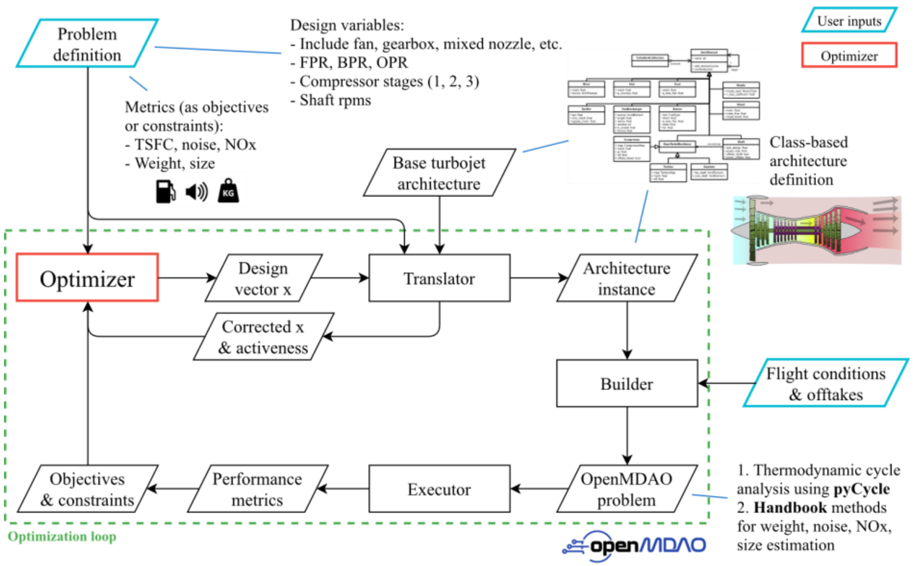
- Recommendation
  - Prediction with Mixed Discrete GP (MDGP) or Random Forest Classifier (RFC)
  - $PoV_{min} = 25\%$



Breiman, L. (2001). "Random Forests." Machine Learning, 45, 5–32.
Saves, P., et al. "A general square exponential kernel to handle mixed-categorical variables for Gaussian process," AIAA AVIATION 2022 Forum, Chicago, USA, 2022.

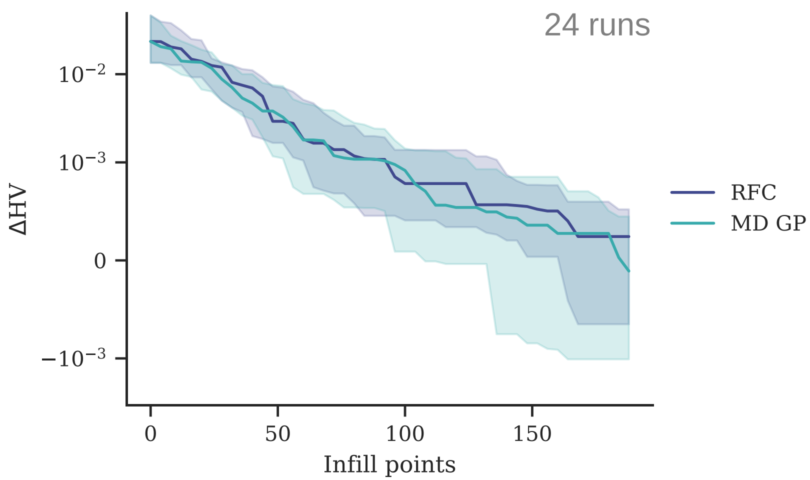
Application Case: Jet engine architecture optimization

Collaboration with J. H. Bussemaker (DLR)



- Evaluation budget: 300 (DOE 113 points + 187 infill points)
- $PoV_{min} = 25\%$
- Prediction strategy with RFC and MD GP perform similarly
- Able to find the optimum in 300 evaluations

  vs 3250 evaluations using an evolutionary algorithm (92% reduction)



- System Architecture Optimization test problem framework
- Minimize Thrust-Specific Fuel Consumption (TSFC)
- 3 discrete, 3 categorical and 9 continuous design variables
- 11 hierarchical variables
- 1 – 5 minutes per evaluation
- 50% failure rate

Bussemaker, J.H., et al., "System Architecture Optimization: An Open Source Multidisciplinary Aircraft Jet Engine Architecting Problem," AIAA AVIATION 2021 FORUM, Virtual Event, 2021

Modification of the acquisition function

BOMBARDIER
McGill UNIVERSITY

Collaboration with A. Tfaily (BOMBARDIER)

- Using

$$\mathrm{EI}(x) = \begin{cases} (y_{\min} - \hat{y}(x))\,\Phi\left(\frac{y_{\min} - \hat{y}(x)}{\hat{s}(x)}\right) + \hat{s}(x)\,\phi\left(\frac{y_{\min} - \hat{y}(x)}{\hat{s}(x)}\right), \\ 0, \quad \text{if } \hat{s} = 0, \end{cases}$$

- a new feasibility enhanced expected feasible improvement $\mathrm{EFI}_{\mathrm{FE}}$ function is defined

$$\mathrm{EFI}_{\mathrm{FE}}(x) = \begin{cases} \boxed{p_{\mathrm{nf}}(x)}(y_{\min} - \hat{y}(x))\,\Phi\left(\frac{y_{\min} - \hat{y}(x)}{\hat{s}(x)}\right) + \boxed{p_{\mathrm{nf}}(x)}^{\boxed{\alpha(x)}}\hat{s}(x)\phi\left(\frac{y_{\min} - \hat{y}(x)}{\hat{s}(x)}\right) \\ 0, \quad \text{if } \hat{s} = 0, \end{cases}$$

where the factor α enables the reduction of the impact of the classifier on the exploration part of EI

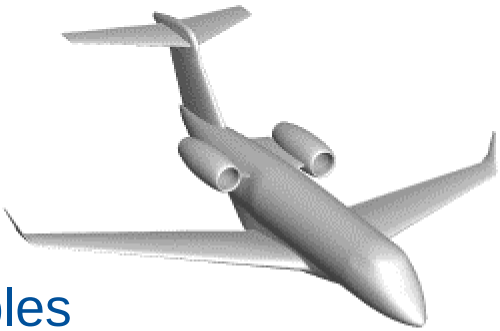
- p_{nf} is the probability of non failure based on a classifier (kNN, GP, RF, SVM, …)
- exploration factor: $\alpha(x) \in [0,1]$ with a dynamic calculation based on the variance

➔ allows the acquisition function to explore closer to a failure region even if p_{nf} is low

Bachoc, F., Helbert, C., and Picheny, V., "Gaussian process optimization with failures: classification and convergence proof," Journal of Global Optimization, Vol. 78, No. 3, 2020,
Tfaily, A., Diouane, Y., Bartoli, N., & Kokkolaras, M. (2024). Bayesian optimization with hidden constraints for aircraft design. Structural and Multidisciplinary Optimization, 67(7), 123.

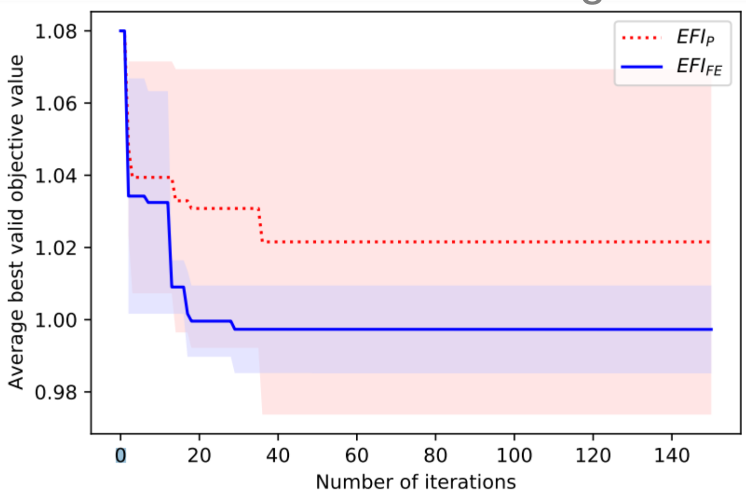
Modification of the acquisition function

Collaboration with A. Tfaily (BOMBARDIER)



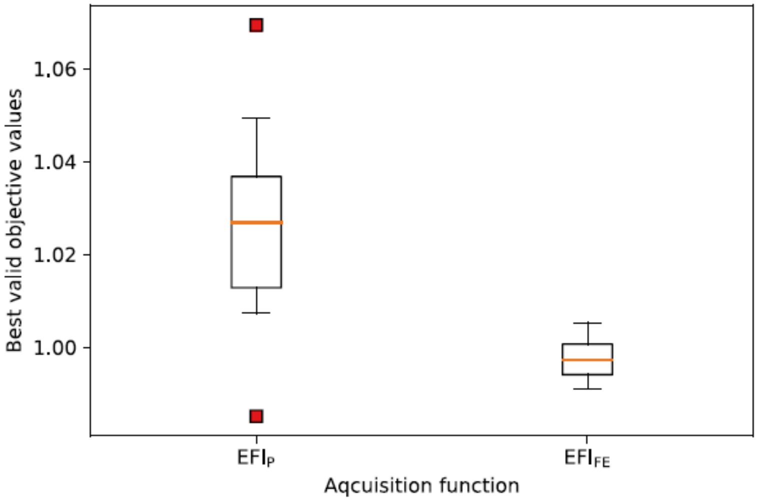
- Aircraft conceptual design problem:
  Bombardier Research Aircraft (BRAC)
- minimization of aircraft weight using 12 design variables
  and subject to 8 inequality constraints
- A landing gear design code fails in certain wing/fuselage configurations

Average of 20 optimization runs - kNN classifier with k=3



α=1 for EFI_P

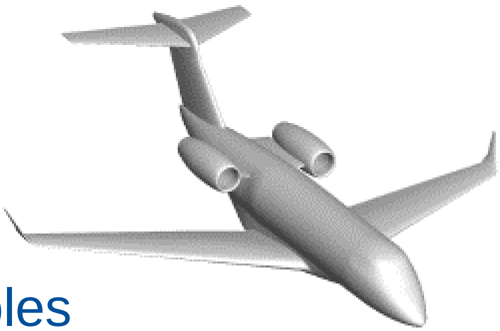
$\alpha(x)$ for EFI_{FE}



Tfaily, A., Diouane, Y., Bartoli, N., & Kokkolaras, M. (2024). Bayesian optimization with hidden constraints for aircraft design. Structural and Multidisciplinary Optimization, 67(7), 123.

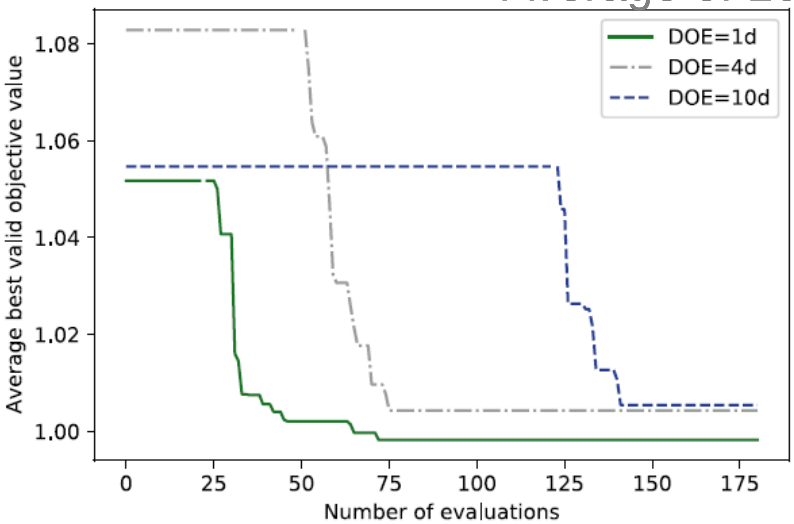
Modification of the acquisition function

Collaboration with A. Tfaily (BOMBARDIER)

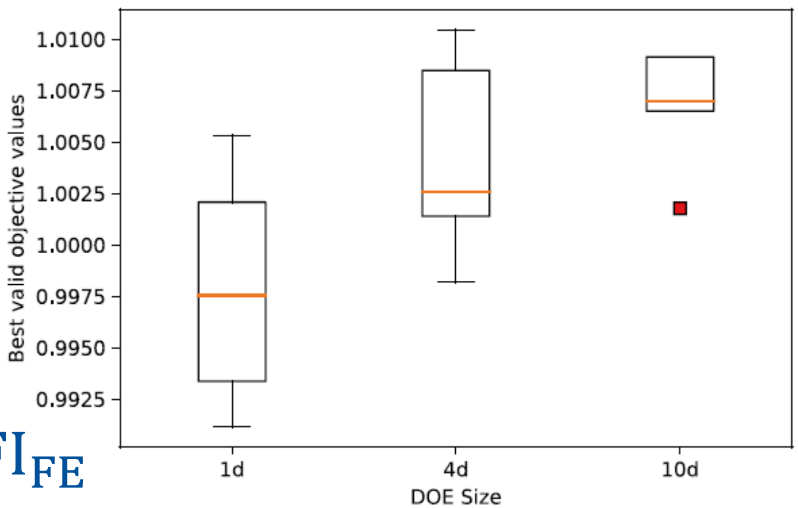


- Aircraft conceptual design problem:
  Bombardier Research Aircraft (BRAC)
- minimization of aircraft weight using 12 design variables
  and subject to 8 inequality constraints
- A landing gear design code fails in certain wing/fuselage configurations

Average of 10 optimization runs - kNN classifier with k=3



EFI_{FE}



Tfaily, A., Diouane, Y., Bartoli, N., & Kokkolaras, M. (2024). Bayesian optimization with hidden constraints for aircraft design. Structural and Multidisciplinary Optimization, 67(7), 123.

Conclusions and perspectives

- Bayesian Optimization for MDO
  - Mixed-discrete, hierarchical, mono and multi-objective, constrained
  - Subject to hidden constraints
- Hidden constraint strategies
  - Neighborhood constraint
  - Failed area prediction through Probability of Viability (*PoV*)
    - PoV_{min} =25% with a RFC or MD GP model works best
  - Modify acquisition function (EI, WB2, WB2S)
  - ➔ Integrated within SEGOMOE framework
- Application to more complex problems

COLOSSUS
COLLABORATIVE SYSTEM OF SYSTEMS

wildfire fighting case

Urban air mobility

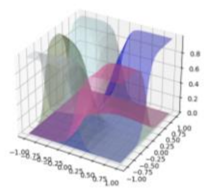
Two frameworks

SMT

github.com/SMTorg/smt

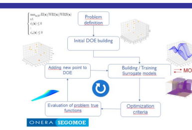
SMT: Surrogate Modeling Toolbox

- Open source python toolbox: surrogate modeling methods, sampling techniques, and benchmarking functions
- Focus on derivatives (training derivatives used for gradient-enhanced modeling, prediction derivatives)
- New Kriging based surrogate models for higher dimension (KPLS and KPLS-K)
- Noisy Kriging to handle uncertainties on data
- Multifidelity Kriging with or without n MFKPLS)
- Mixture of experts technique for heterogeneous functions

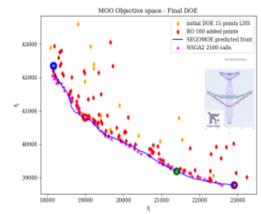


- Mixed integer Kriging to handle **discrete and categorical variables**

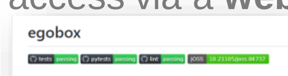
SEGOMOE



- **Mono & multi objective Bayesian optimizer**
- **Mono & Multi fidelity sources**
- Handling non linear objectives & constraints (black box, no derivative available)
- Equality & inequality constraints (1 ~ 100 constraints)
- Intermediate dimension problem (1 ~ 100 variables)



- Heterogenous variables (continuous, discrete, categorical)
- Costly evaluation (CFD, FEM, objective and/or constraints)
- Hidden constraints
- Based on SMT toolbox for surrogate models
- Remote access via a **web interface**



ONERA WhatsOpt