Machine Learning for Industrial Engineering and Applications

E. Burnaev

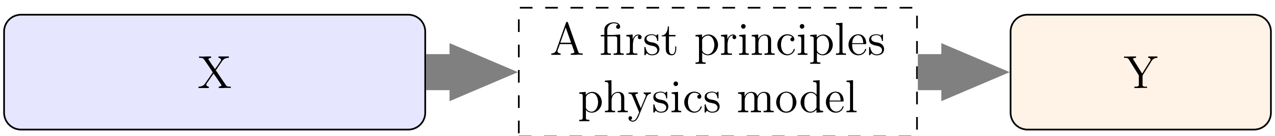
Skoltech, IITP, HSE

Outline

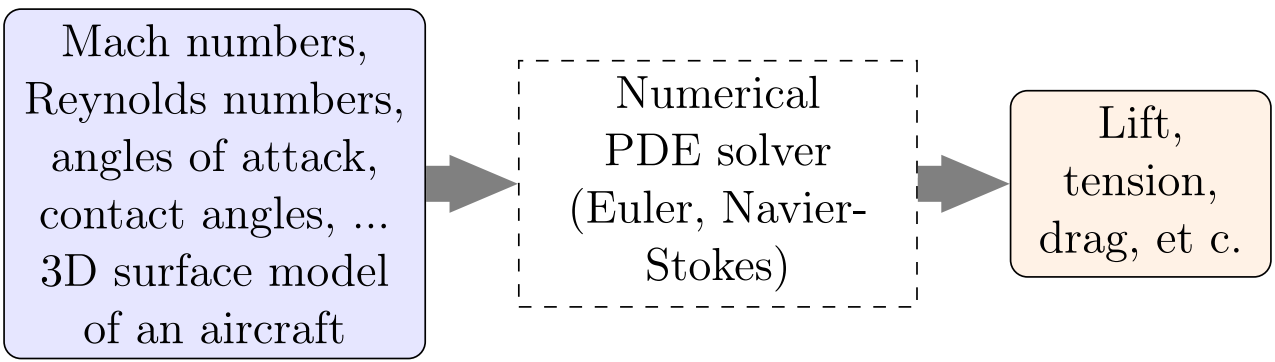
1. Challenges
   - Predictive Modelling
   - Potential
   - Customer Expectations
   - Surrogate Modelling in Engineering
2. Surrogate modelling
   - Constructing Surrogate Models
   - F1 car optimization
3. Special features of engineering problems
   - Data peculiarities
   - Model capabilities
4. Examples of ML solutions
   - Gaussian Processes on multidimensional grid
5. Other Problem Statements in Industrial Engineering

Predictive Modelling

Traditional approach based on the first principles

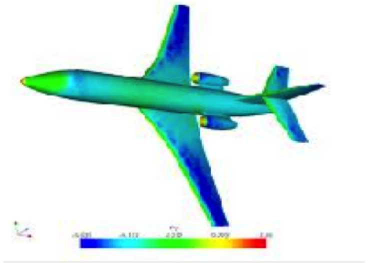


Example: Aircraft Aerodynamics Prediction model



PM: Potential

$$\frac{\partial v}{\partial t} + (v \cdot \nabla) v = -\frac{1}{\rho} \nabla p + \nu \nabla^2 v + f \, ,$$
$$\nabla \cdot v = 0 \, .$$



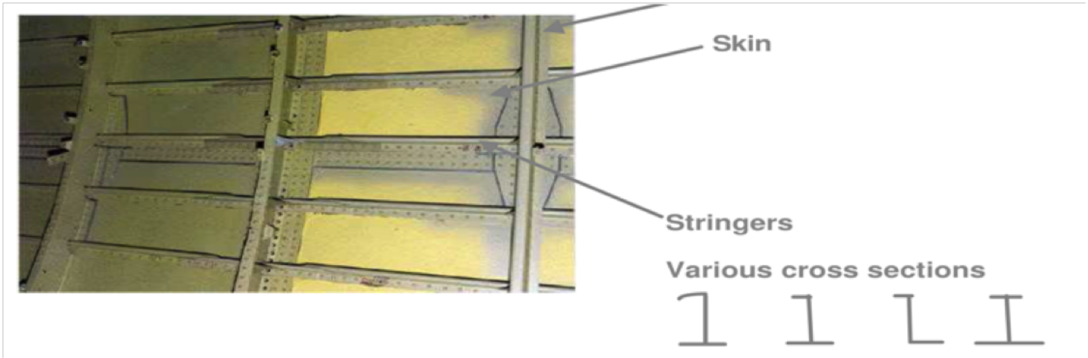
Predictive modelling in engineering:

- **1990-s** Typical volume of experiments (around 10-100) is enough to compare solutions, but not enough to carry out fully fledged optimization;
- **2000-s** Advances in High Performance Computing make engineering optimization economically feasible

Fact: The demands of the industry grow much faster than the computational capacity

PM: Customer Expectations

Airbus: Software for computation of Reserve Factors of aircraft structural elements (stringers) for given geometry, material parameters, loads and applied forces.



Expectations (Airbus): A 100-fold drop in the running time of this software shortens the full cycle of structure optimization from several days to several hours

SM in Engineering

PM is used in:

- "What-If" and Sensitivity Analysis, Design Space Exploration;
- Design Optimization with respect to specified efficiency criteria

Prohibitive volumes and/or run-time costs:

- from thousands to millions of computational experiments;
- running time of an experiment ranges form seconds to days;

Surrogate models: fast approximations, which substitute the original models without a significant loss in accuracy

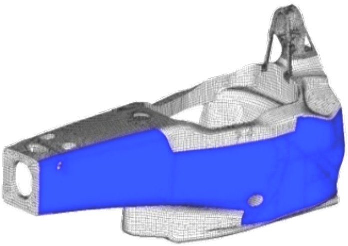
Constructing Surrogate Models

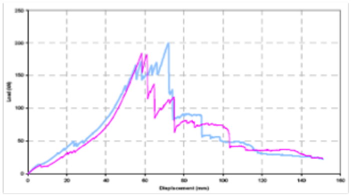
The original model $f_0 : \mathcal{X} \mapsto \mathcal{Y}$ is based on the "first principles".
Major stages of modelling:

1. **Initialization:** carry out experiments to get the initial sample $S_0 = \bigl(x_i, f_0(x_i)\bigr)_{i=1}^{m}$;
2. **Construction:** learn a fast computable approximation $\widehat{f} \approx f_0$ over some domain $U \subset \mathcal{X}$;
3. **Assessment:** measure the accuracy of $\widehat{f}$;
4. **Exploration:** pick an $x \in U$ at which to evaluate f_0, and update S_0 with $(x, f_0(x))$;
5. Repeat steps 2-4 until satisfactory *accuracy* is achieved, or the *computational budget* is exhausted.

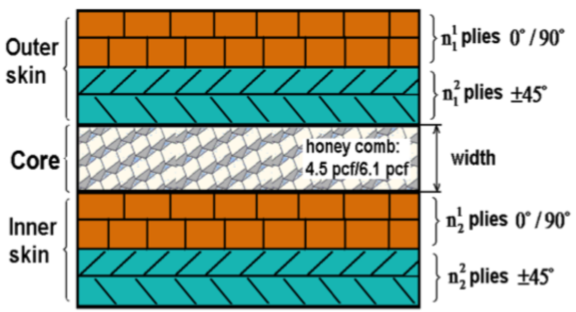
Design of the side panel of Formula 1 car

- **Objective**: Minimize mass of a side panel of Formula 1 car (number of layers, ply structure, thickness), subject to strength constraints;
- **Challenges**: Finite element analysis of the composite structures is slow and offers limited accuracy; experimental and numerical results are scant;
- **Solution**: Construct a surrogate model using experimental and numerical data (data fusion), and use it in optimization;
- **Result**: 10% mass reduction with fewer simulations or full order experiments.





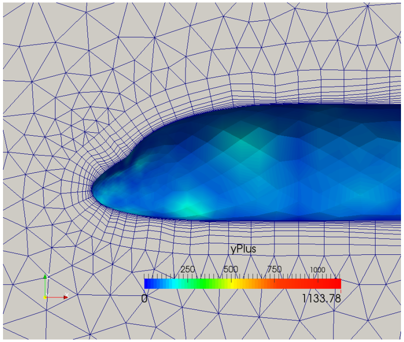
Load and Energy Plots



Several data sources

- Values of physical quantities can be obtained
  - by conducting an experiment;
  - by performing computer simulations;
  - by solving a system of equations;
  - ...
- Different data sources have varying fidelity;

Wind tunnel experiment



CFD simulation

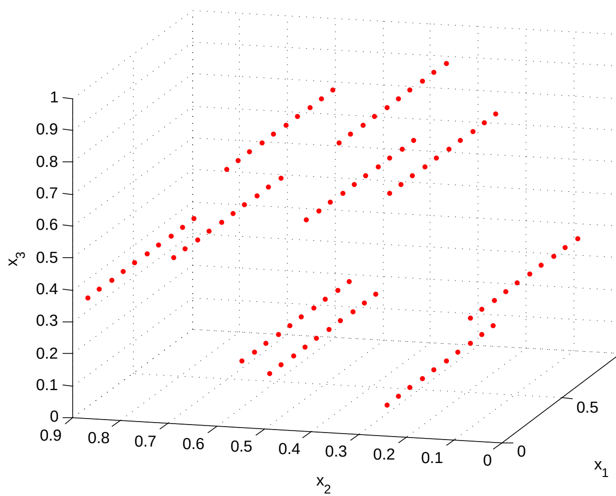
Fidelity of data sources **must** be considered during model construction.

Data anisotropy

Training data can be very **anisotropic** with respect to different groups of variables.

Example:

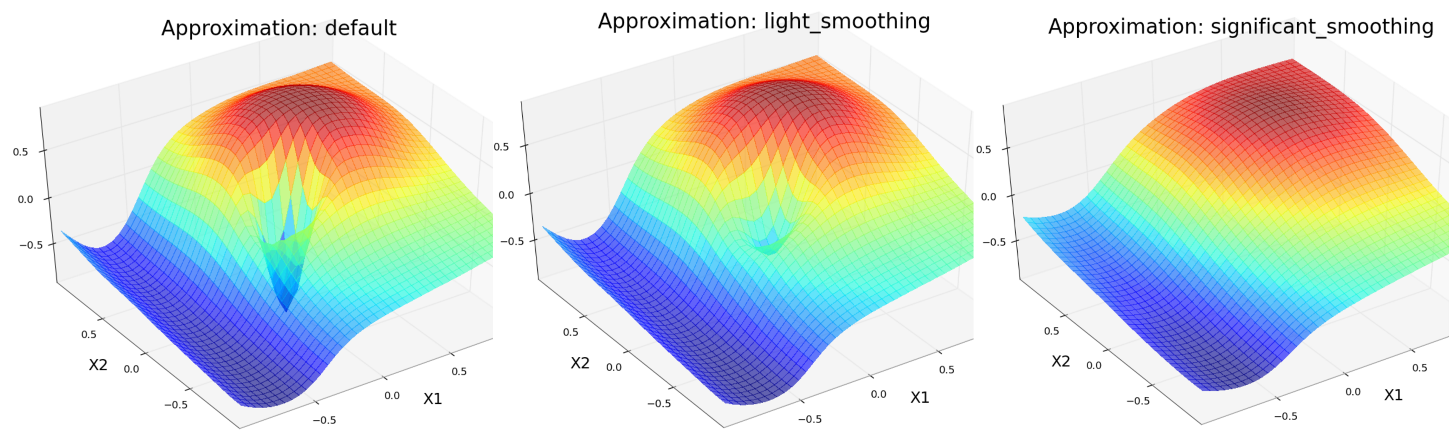
- Air pressure measured under different settings of Mach number and angle of attack;
- Air pressure detectors are abundant and regularly distributed;
- Number of different settings of Mach and angle of attack is scarce, parameters are distributed irregularly.



Model smoothness and availability of gradients

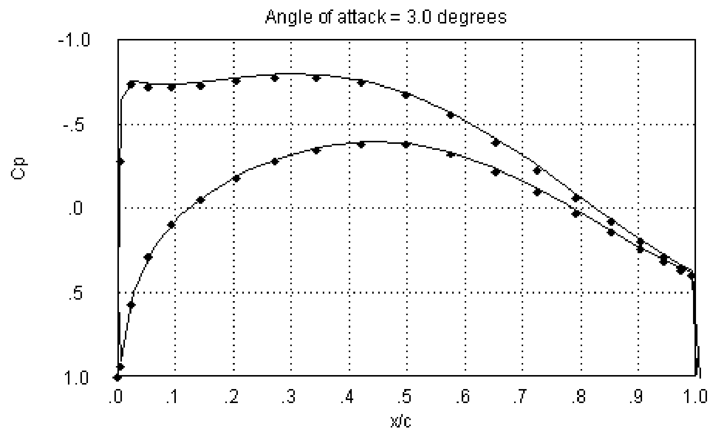
In engineering surrogate models are often used for optimization

- Analytical gradients are useful for gradient-based optimization;
- Model is expected to depend smoothly on input variables;
- Smoothness-accuracy trade-off $\Rightarrow$ it is helpful to control amount of smoothness of the model.



Multidimensional output

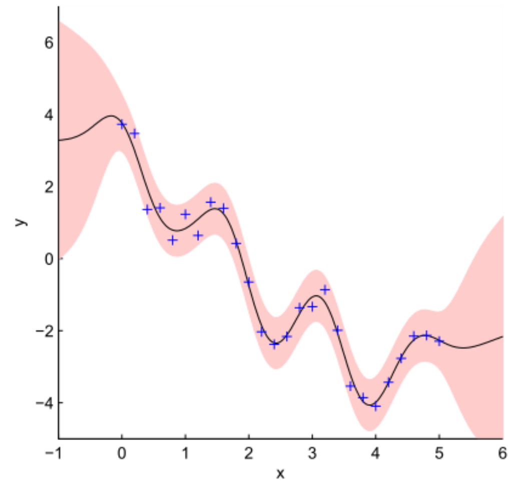
- Usually a surrogate model approximates function with scalar output;
- In some cases modeling of several physical quantities is needed, and such quantities can be highly correlated:
- **Example**: Model of the pressure distribution along an airfoil (as a function of its shape) is expected to have correlated outputs.



Local accuracy estimates

Many engineering problems require an approximation and an **accuracy estimate** of this approximation

- Accuracy estimate is rarely provided by existing software;
- The most popular algorithm that provides accuracy estimate is based on **Gaussian Processes**.



1. Challenges
   - Predictive Modelling
   - Potential
   - Customer Expectations
   - Surrogate Modelling in Engineering
2. Surrogate modelling
   - Constructing Surrogate Models
   - F1 car optimization
3. Special features of engineering problems
   - Data peculiarities
   - Model capabilities
4. Examples of ML solutions
   - Gaussian Processes on multidimensional grid
5. Other Problem Statements in Industrial Engineering

Gaussian Processes Regression

Let

- $S = (D, \mathbf{u}) = \{\mathbf{x}_i, u_i\}_{i=1}^{N}$ be a training set;
- $\mathbf{x}_i \in \mathbb{R}^d, u_i \in \mathbb{R}$.

Suppose

- $u_n = f(\mathbf{x}_n) + \varepsilon_n, \quad \varepsilon \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I});$
- $f(\mathbf{x})$ — is a Gaussian Process (GP)

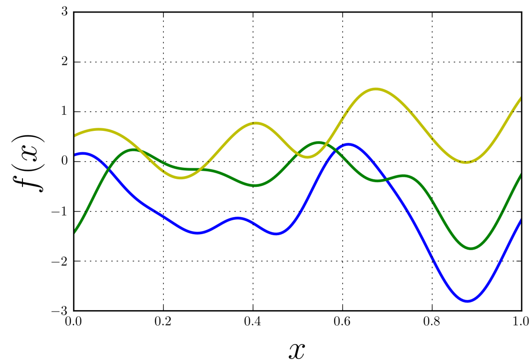
Gaussian Process Regression

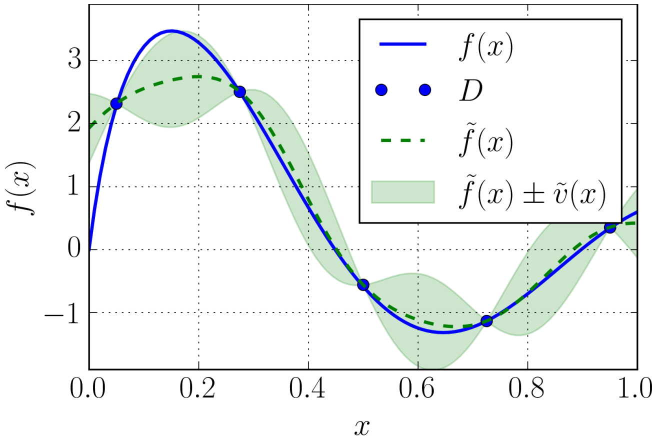
- $f(\mathbf{x})$ is a realization of a Gaussian Process;
- the Gaussian Process is stationary: its covariance function

$$\mathrm{cov}(f(\mathbf{x}), f(\mathbf{x}')) = R(\mathbf{x} - \mathbf{x}')\,;$$

- the spectral density

$$F(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} e^{2\pi \mathrm{i} \boldsymbol{\omega}^{\mathrm{T}} \mathbf{x}} R(\mathbf{x}) d\mathbf{x}\,.$$





To make predictions we use posterior mean

$$u_*|\mathbf{u} \sim \mathcal{N}(\mu(\mathbf{x}_*), \sigma^2(\mathbf{x}_*)),$$

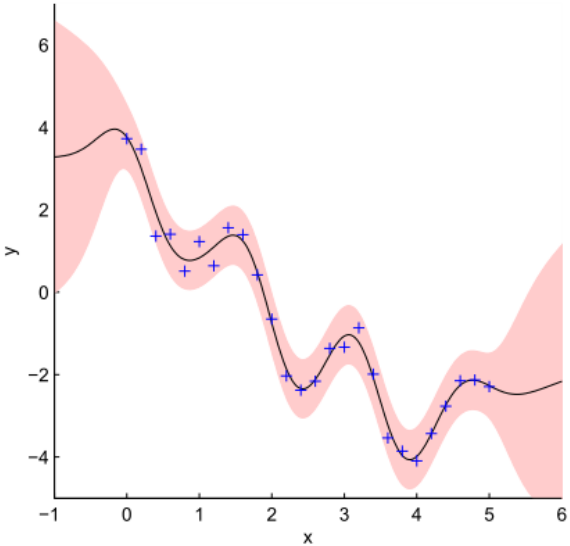
$$\mu(\mathbf{x}_*) = \mathbf{k}^T \mathbf{K}^{-1} \mathbf{u},$$

$$\sigma^2(\mathbf{x}_*) = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}^T \mathbf{K}^{-1} \mathbf{k},$$

where

$\mathbf{k} = (R(\mathbf{x}_* - \mathbf{x}_1), \ldots, R(\mathbf{x}_* - \mathbf{x}_N))^T$,

$\mathbf{K} = \{R(\mathbf{x}_i - \mathbf{x}_j)\}_{i,j=1}^N$



$\sigma^2(\mathbf{x}_*)$ is used as **Accuracy Estimate**

GP on a multidimensional grid

- In many engineering problems inputs of training data lie on a multidimensional grid;
- The size of the training data (N) grows exponentially with the number of input variables;
- Complexity of training a GP model is high — $\mathcal{O}(N^3)$;
- However, dataset has special structure which implies that

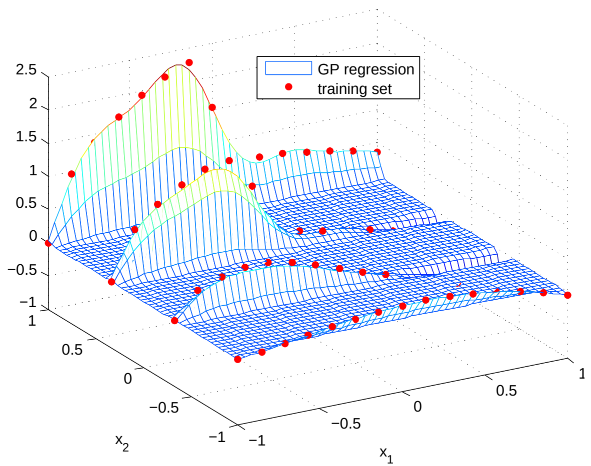
$$\mathbf{K} = \bigotimes_{i=1}^{d} \mathbf{K}_i \, ,$$

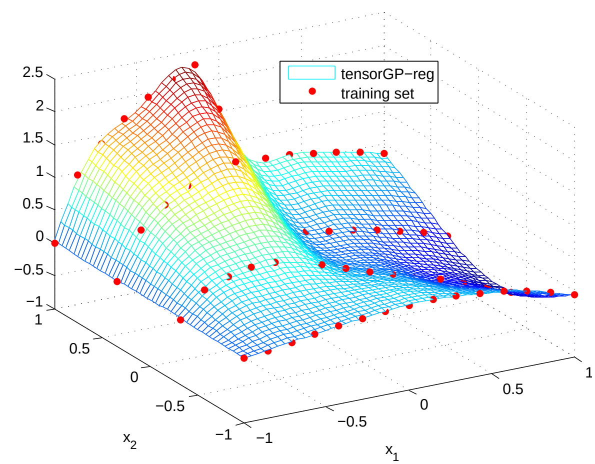
  where $\mathbf{K}_i$ — is a covariance matrix of i-th factor and $\bigotimes$ is the Kronecker product;
- Tensor algebra $\mathbf{K}^{-1}\mathbf{u}$ can be efficiently computed in this case.

GP on multidimensional grid

Special dataset structure can be exploited

- to reduce the complexity to $\mathcal{O}\left(N^{1+1/K}\right)$, where K is the number of input factors;
- to introduce special **regularization** to avoid degeneration of GP model;





Other types of problems are important for Industrial Engineering

1. Anomaly Detection;
2. Failure Prediction and Early warning systems.

Using Surrogate Modelling:

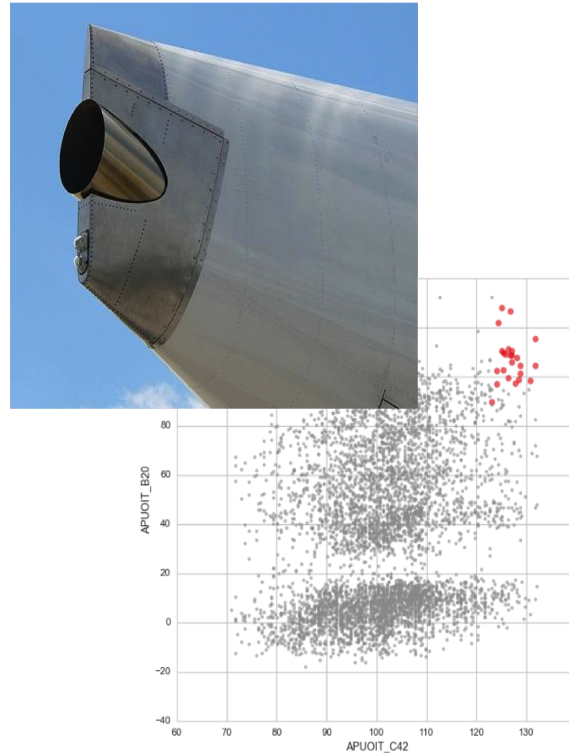
— we construct a model for a system in a normal state;

— detect anomalous behaviour against predictions of the model.

Other/newer methods are needed (imbalanced classification, semi-supervised anomaly detection, manifold learning, et c.)

Auxiliary Power Unit Failures Prediction

- **Objective**: Predict Failures of APU
- **Data**:
  - 30 aircrafts, 200+ parameters for each aircraft
  - Learning: 3 years (400 flights per year)
  - Testing: 1/2 year
- **Conclusion/Benefits**:
  - Early warnings about some types of failures
  - **Coverage** (detected failures) $\sim 34\%$
  - **Accuracy** of Failures Detection is $\sim 90\%$ (for 9 correctly predicted failures on average we get only 1 false alarm)



Conclusions

- We develop new machine learning approaches to efficient predictive modelling and engineering optimization;
- We collaborate with industrial companies (Airbus, ALTRAN, IPSEN, Louis Vuitton, etc.) and produce solutions, approved by the industry;
- We constantly seek new advanced ML problems, their industrial applications and new collaboration