POLYTECHNIQUE MONTRÉAL

BOMBARDIER

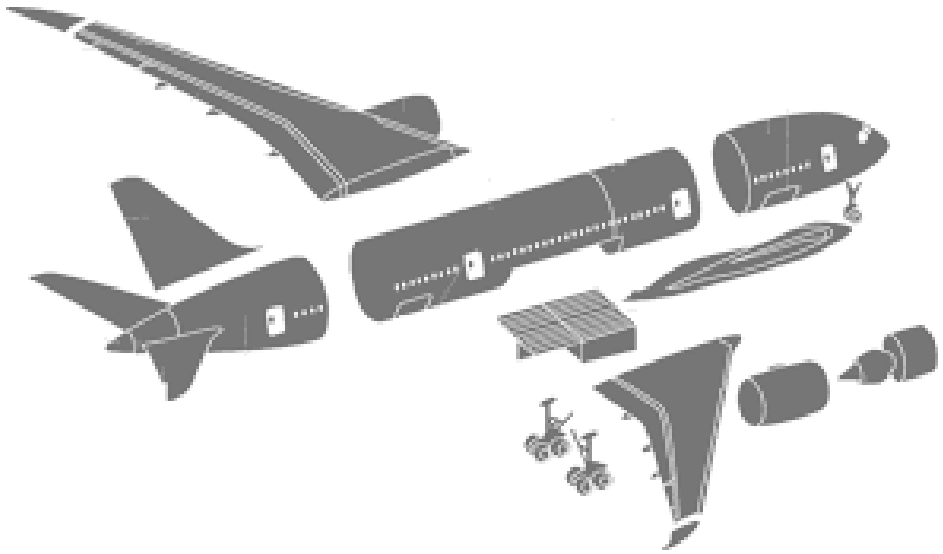
SmOOD: Smoothness-based Out-of-Distribution Detection Approach for Surrogate Neural Networks in Aircraft Design

Houssem Ben Braiek, Ali Tfaily, Foutse Khomh, Thomas Reid, and Ciro Guida

In The 37th IEEE/ACM International Conference on Automated Software Engineering



Aircraft Design



Aircraft design is a highly iterative process to balance competing criteria involving certification requirements and public demands.

Safety

Strength

Fuel Economy

CO_2 Emissions

Noise Pollution

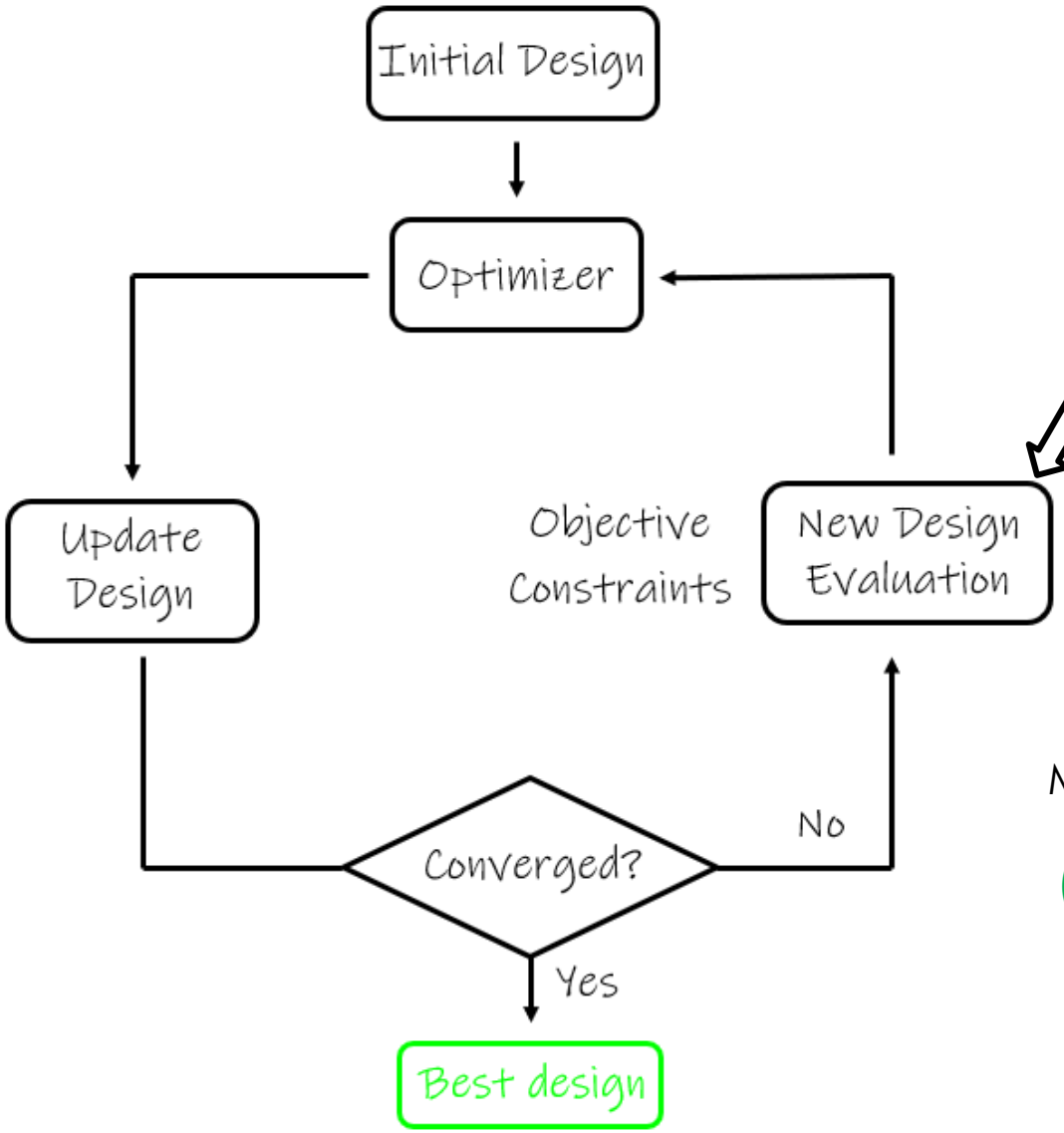
...

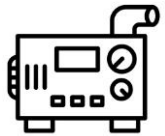
while maintaining aircraft development

Lead times

Costs

Surrogate Aircraft Design Optimization



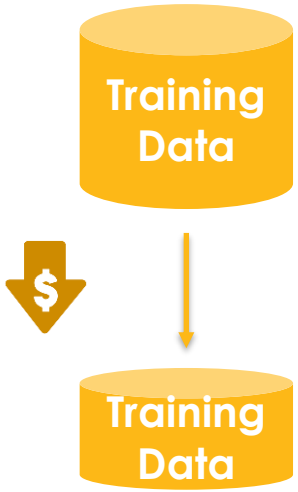


Physics-Based Model
(Expensive Simulations)



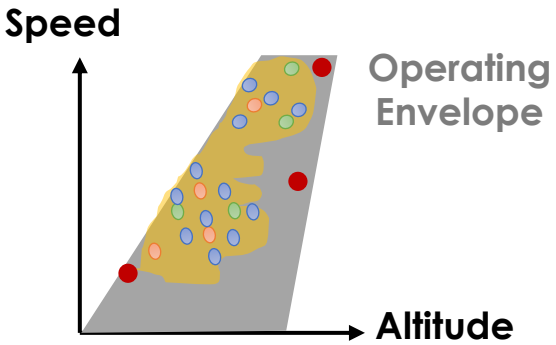
Machine Learning-Based Surrogate Model
(Cheap to run Statistical Inference)

Out-of-Distribution Entries in Surrogate Models and Hybrid Settings

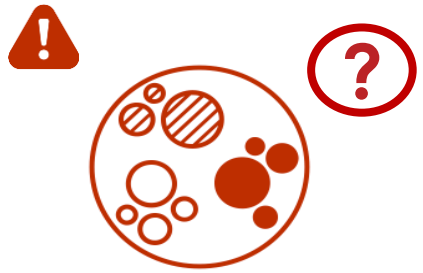


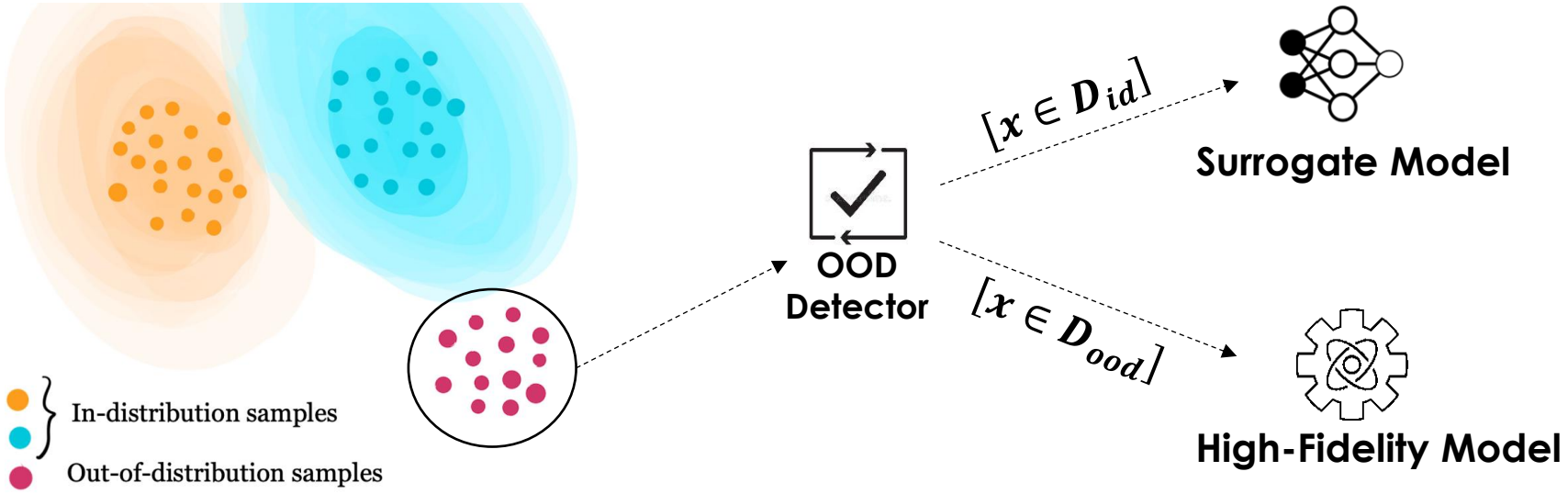


In-Distribution Data Fit



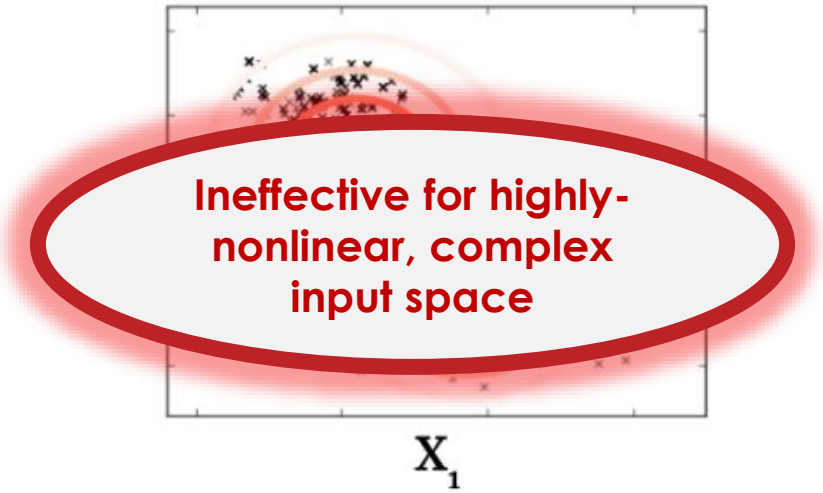
Out-of-distribution (OOD) entries are almost inevitable.





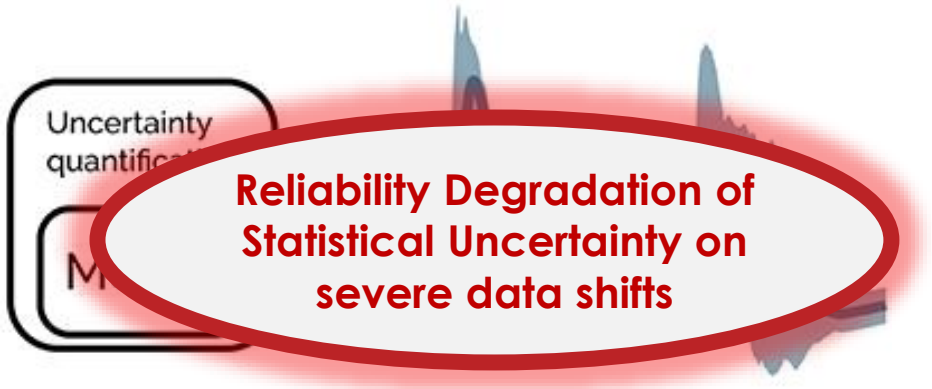
Out-of-distribution Detection Challenges

$P(X)$



If $P(x) < \tau$,
Then x is an OOD entry.

$y \sim \mathcal{N}(\mu, \sigma)$



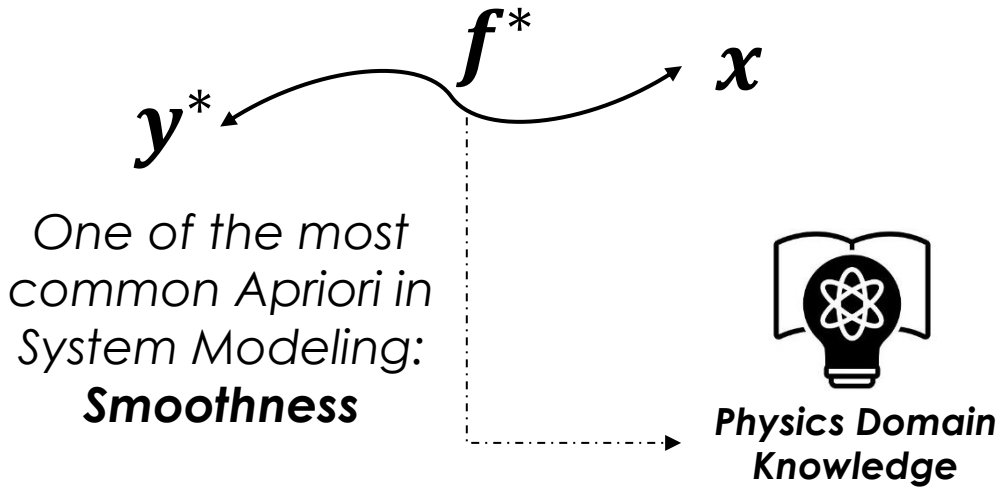
If $\sigma(x) < \tau$,
Then x is an OOD entry.

Apriori Smoothness Property of Neural Network Mapping Function

The sensitivity of the **learned mapping function**, f^*, of the model.

$$\forall\, x' \colon \|x - x'\|_p \leq \delta$$
$$\Rightarrow f(x) - f(x') \leq \epsilon$$

δ, ϵ : *depending on the modeled system*



- **Smooth dynamical systems.**
- **Piecewise-smooth dynamical systems.**

Definition of "Out of Domain"

Input Generation X using Latin Hypercube Sampling (LHS)

Physics-Based Model → (X, y) Ground truth data

Surrogate Model → $(X, \hat{y})$ Predictions data

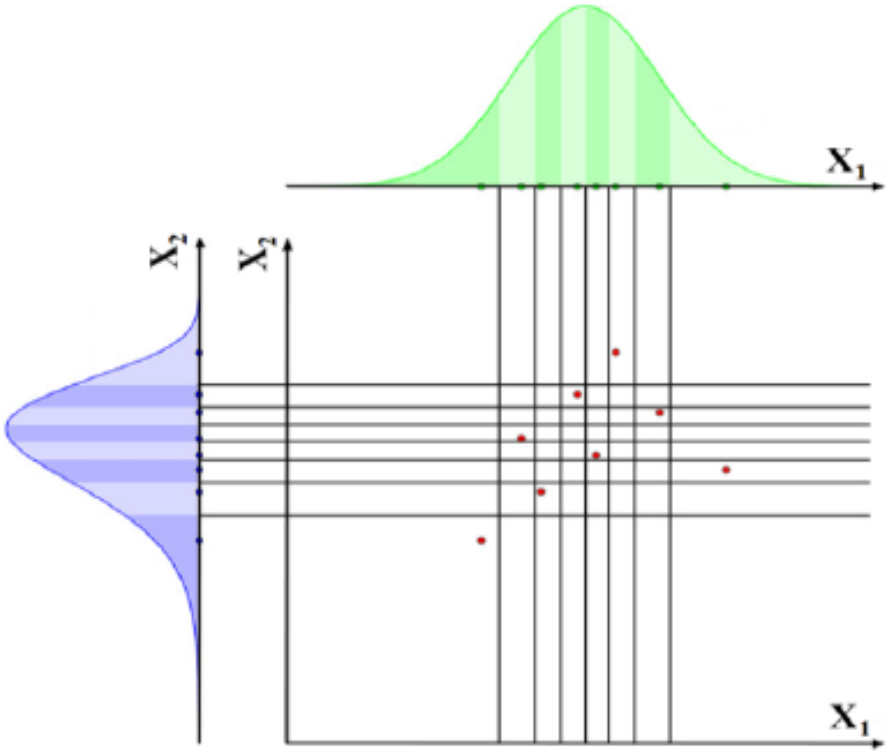


Illustration of LHS technique with two input variables

Observation 1: On training, the model tends to minimize the in-distribution error and eventually sacrifice some data points.

Observation 2: On validation, some data points would keep yielding high error scores despite hyperparameters tuning.

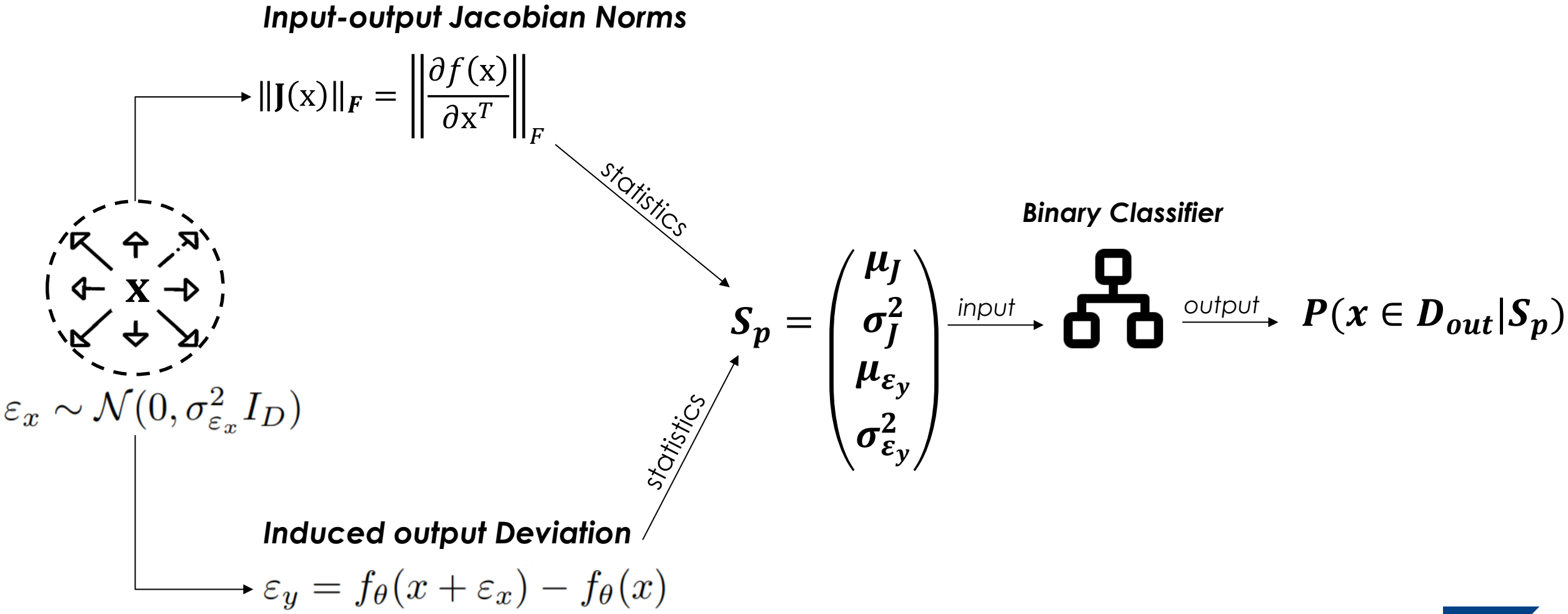
➔ These samples that could not be included within the model distribution, can be a proxy to recognize statistically the out-of-distribution regions.

Rule-based OOD Labeling

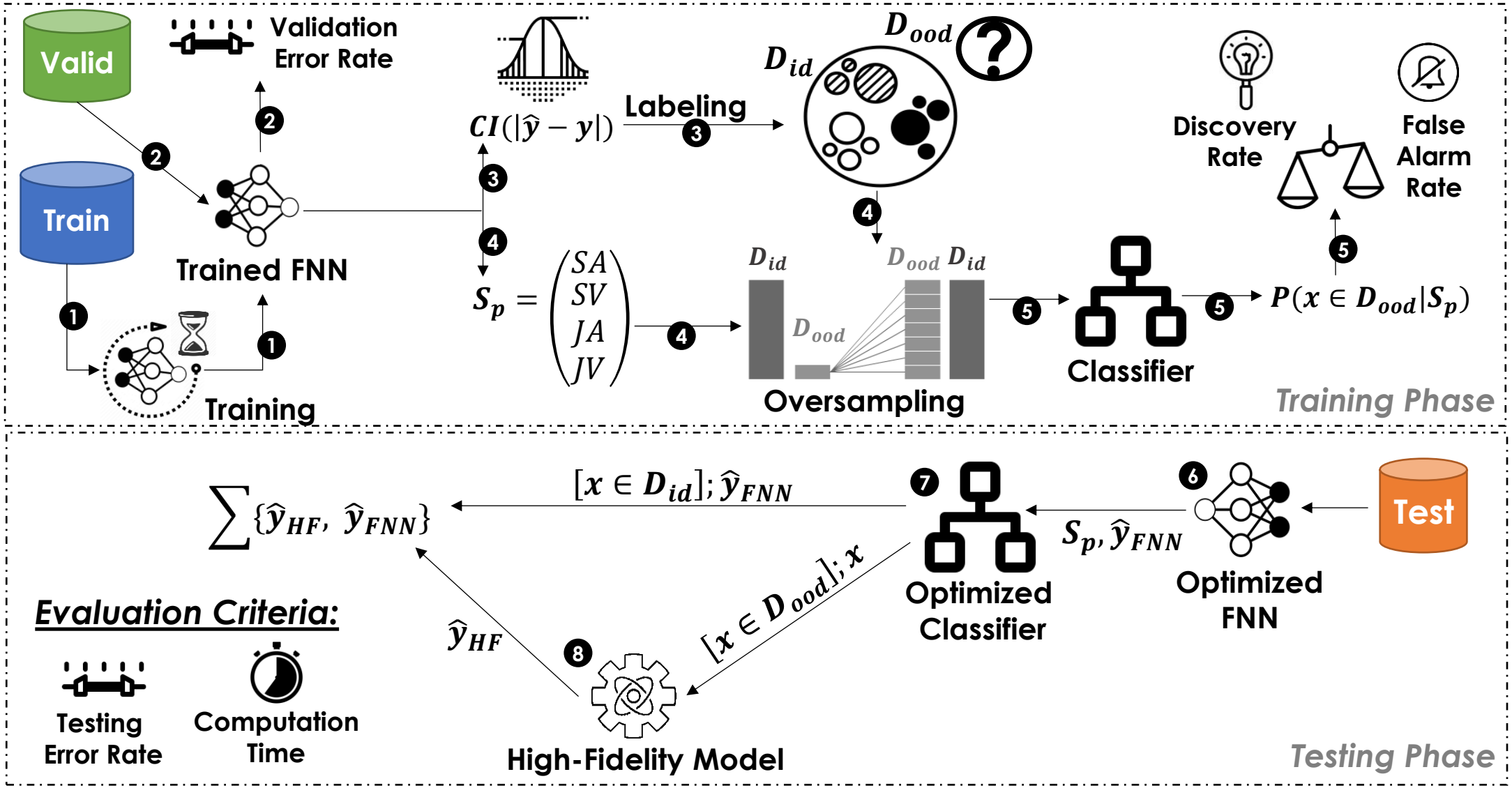
$$|y - \hat{y}| \leq \mu_{err} + coef \times \sigma_{err}$$

Depends on the Confidence Level (CL)
95%, **99%**, ...

Pointwise Sensitivity Profiling & Classification-based OOD Detection

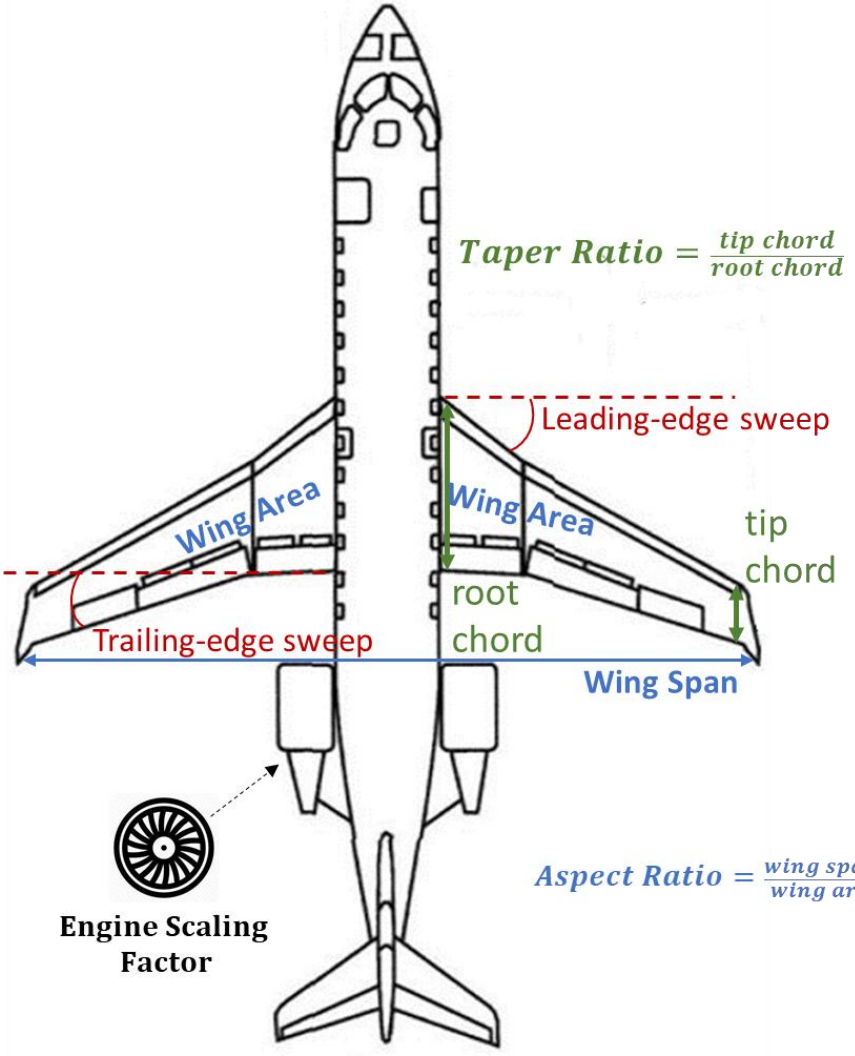


End-to-End SmOOD Workflow: Training and Testing Phases



Aircraft Performance Case Studies

Aircraft Design Variables



MTOW

TTC

BFL

Aircraft Performance Factors

Max (Takeoff Weight) = Basic Empty Weight + Payload + Fuel

Aircraft, crew & catering

Passengers, baggage & cargo

Time to Climb



RQ1. FNN vs GP in Surrogate Modeling Aircraft Performance Quantities

QoI	Model	Valid NRMSE	Test NRMSE
MTOW	GP	0.041	0.036
	FNN	0.035	0.032
TTC	GP	0.17	0.169
	FNN	0.076	0.08
BFL	GP	0.045	0.044
	FNN	0.043	0.036

$$\mathrm{NRMSE} = \frac{\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2/n}}{y_{max} - y_{min}}$$

Max Take Off Weight

FNN: 1,5% OOD Vs *GP:* 5% OOD

Time To Climb

FNN: 5% OOD Vs *GP:* 40% OOD

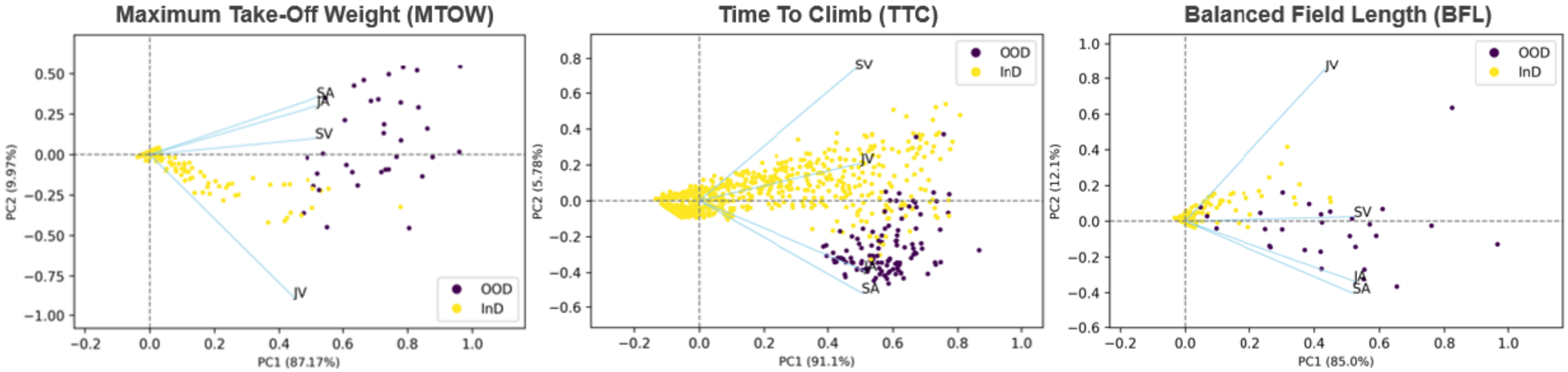
Balanced Field Length

FNN: 1,5% OOD Vs *GP:* 6% OOD

%: Ratio FNN: Feedforward Neural Network GP: Gaussian Process OOD: Out-Of-Distribution Inputs NRMSE: Normalized Root Mean Squared Error

RQ2/3. Effectiveness of Sensitivity Profiles in OOD Detection

Graph of Pointwise Sensitivity Profiles for InD and OOD w.r.t the derived 2 Principal Components



Max Take Off Weight

SmOOD: 86% Precision, 95% Recall

Vs

Baseline: 53% Precision, 53% Recall

Time To Climb

SmOOD: 85% Precision, 95% Recall

Vs

Baseline: 36% Precision, 47% Recall

Balanced Field Length

SmOOD: 88% Precision, 86% Recall

Vs

Baseline: 54% Precision, 53% Recall

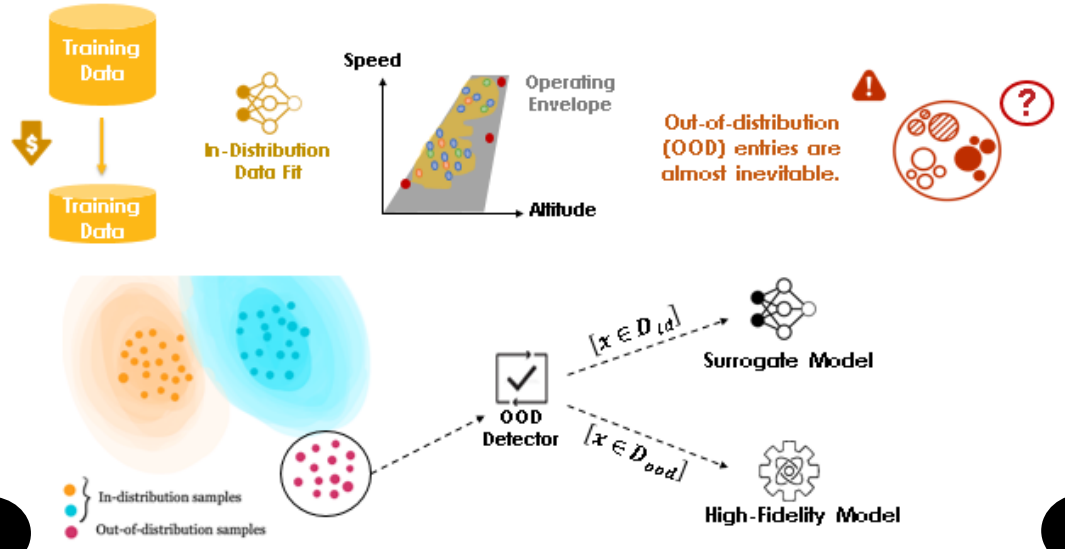
RQ4. Benefits of SmOOD in hybrid surrogate optimization settings

$$\text{SpeedUp} = \frac{\text{T}_{initial}}{\text{T}_{improved}}$$

QoI	Model	%NRMSE_Decr	Runtime	Speedup
MTOW	HF	-	02h:57m:58s	-
	FNN	-	0.81ms	3.57×10^4
	HybridFNN	47%	02m:38s	6.76×10
TTC	HF	-	02h:57m:58s	-
	FNN	-	15.51ms	6.88×10^5
	HybridFNN	45%	12m:31s	1.42×10
BFL	HF	-	02h:57m:58s	–
	FNN	-	4.58ms	2.33×10^6
	HybridFNN	12.25%	01m:54s	1.42×10

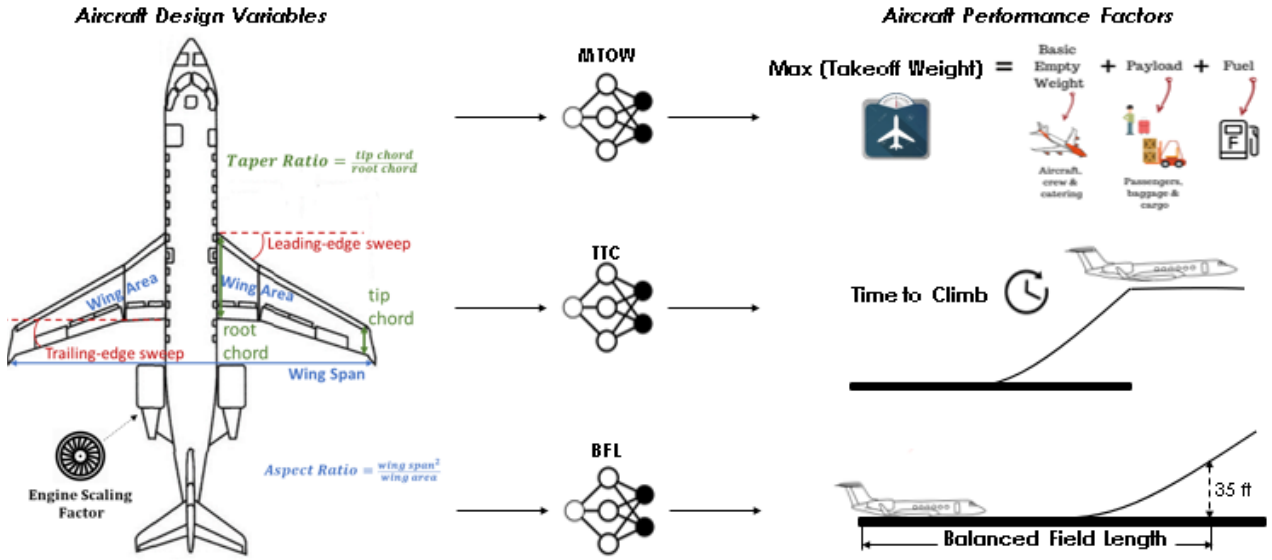
$$\%\text{NRMSE_Decr} = \frac{\text{preNRMSE} - \text{postNRMSE}}{\text{preNRMSE}} \times 100$$

Out-of-Distribution Entries in Surrogate Models and Hybrid Settings



1

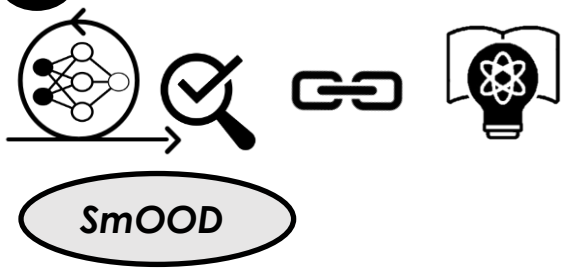
Aircraft Performance Case Studies



3

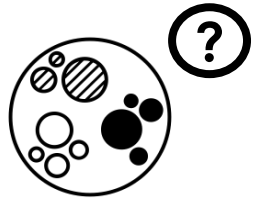
Your Questions ?

2

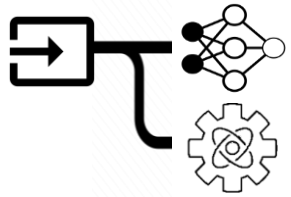




Apriori property of Smoothness



Out-of-distribution Detection



Trustworthy Selective Predictions

4



Active Learning

Beyond Physics and System Modeling:



Many Computer-vision models require usually smooth mappings.