Data Centric Operational Design Domain Characterization for Machine Learning Based Aeronautical Products

Fateh Kaakai

Thales, France

Ganesh Pai

KBR / NASA Ames Research Center, USA

Shreeder Adibhatla

Rockdale Systems, USA

Emmanuelle Escorihuela

Airbus Operations, France

SAFECOMP 2023

September 22, 2023. Toulouse, France

Outline

- Background
- Concepts
- Data centric characterization
- System layer analysis
- Conclusions and future work

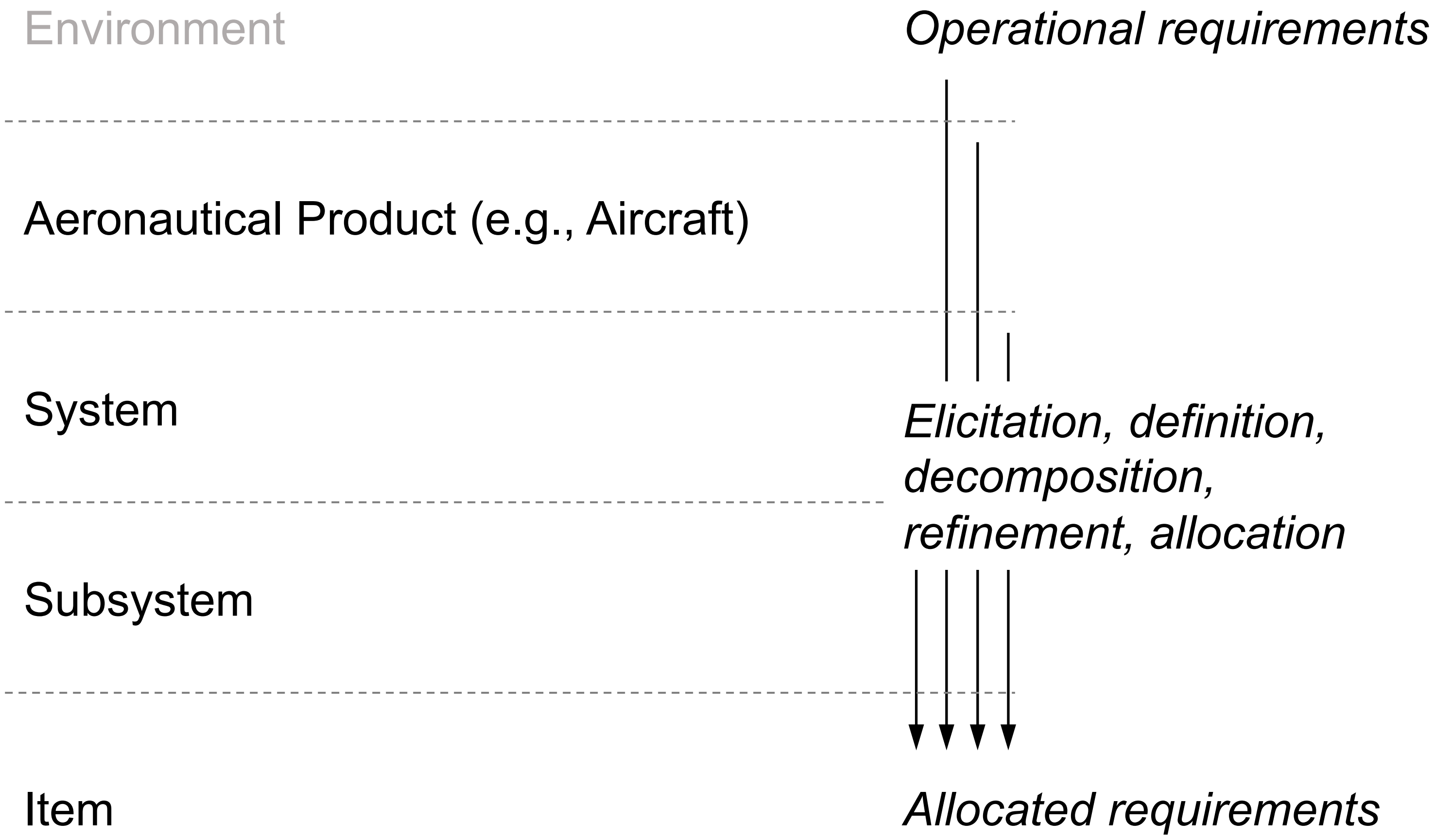
Background

- SAE G-34 and EUROCAE WG-114 jointly developing ARP 6983 (Corresponding ED to be defined), Process Guidance for Development and Certification/Approval of Aeronautical Safety-related Products Implementing AI
  - Focus on supervised, offline ML
  - For applications with up ML contribution to MAJOR severity of safety effect leading to Design Assurance Level (DAL) C or D according to the system architecture
  - (ODD) Working Group aiming to answer how to define, analyze, manage, allocate the operational conditions where ML will be used

Objectives

- To adjust conventional processes for determining operational conditions for ML functionality from system-layer operational requirements
- To take the specification of operational conditions into account for aeronautical system development, with safety constraints

Background



Environment

Operational requirements

Aeronautical Product (e.g., Aircraft)

System

Elicitation, definition, decomposition, refinement, allocation

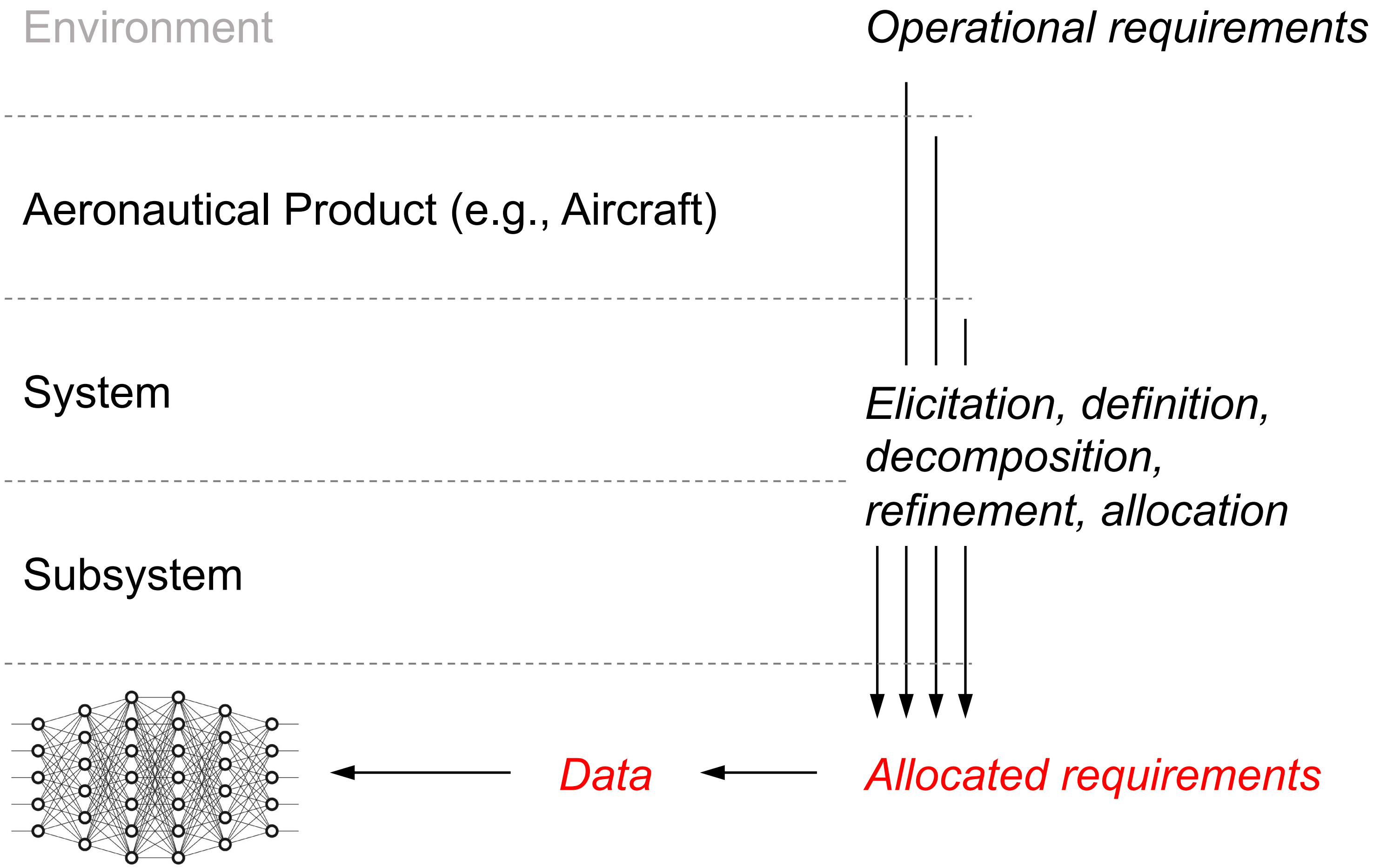
Subsystem

Item

Allocated requirements

Well understood for conventional systems

Background



Environment

Operational requirements

Aeronautical Product (e.g., Aircraft)

System

Elicitation, definition, decomposition, refinement, allocation

Subsystem

Data

Allocated requirements

Not as well understood for systems including Machine Learning

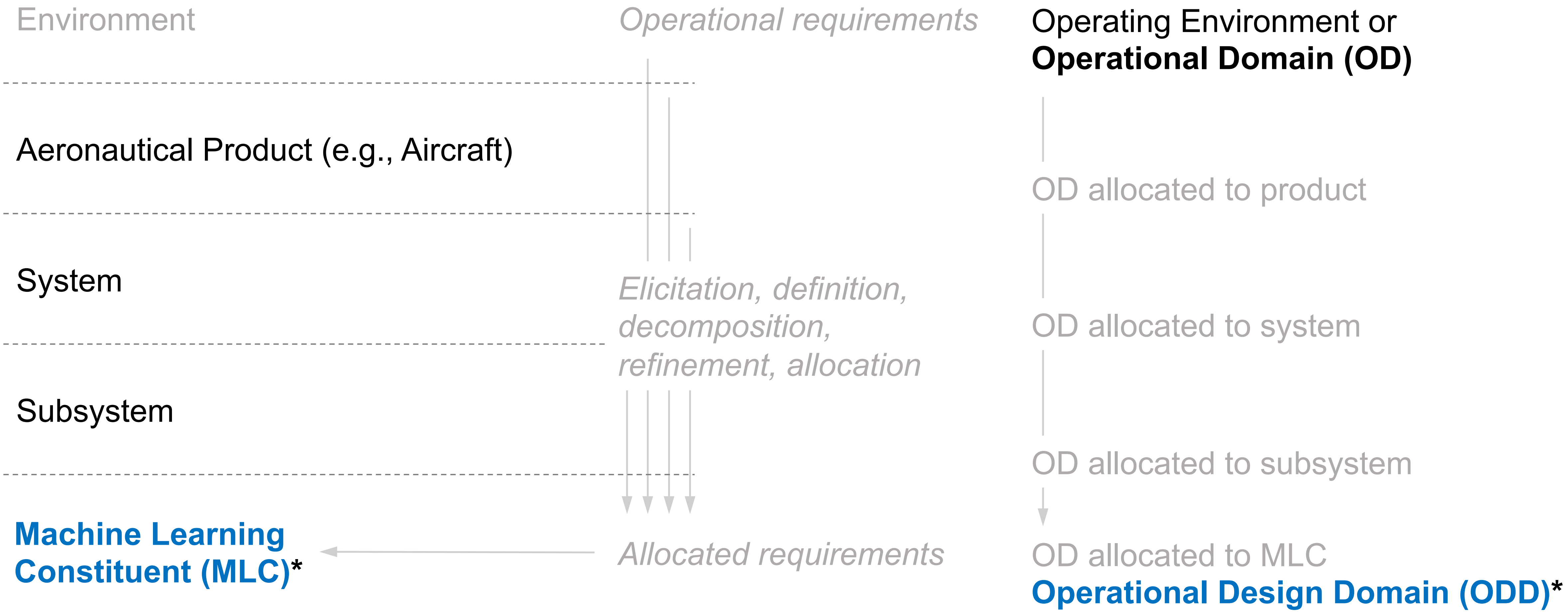
Reconciling data and functional intent

Outline

- Background
- Concepts
- Data centric characterization
- System layer analysis
- Conclusions and future work

Concepts

*New Concept



Environment

Aeronautical Product (e.g., Aircraft)

System

Subsystem

Machine Learning Constituent (MLC)*

Operational requirements

Elicitation, definition, decomposition, refinement, allocation

Allocated requirements

Operating Environment or
Operational Domain (OD)

OD allocated to product

OD allocated to system

OD allocated to subsystem

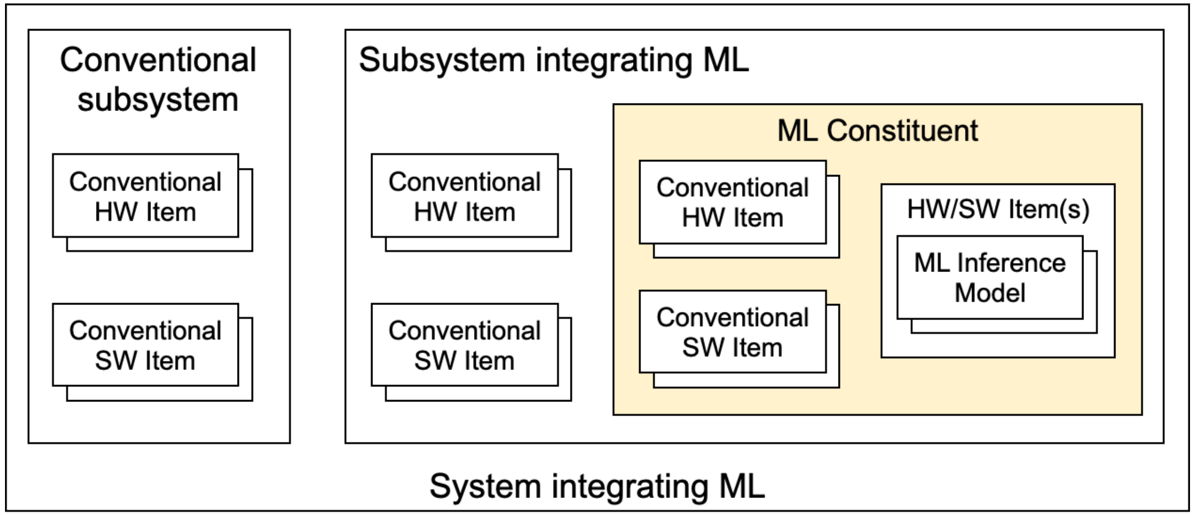
OD allocated to MLC
Operational Design Domain (ODD)*

Machine Learning Constituent (MLC)

- A (logical) grouping of hardware and/or software items implementing one or more ML Models (MLMs) and their associated data pre-processing and post-processing items

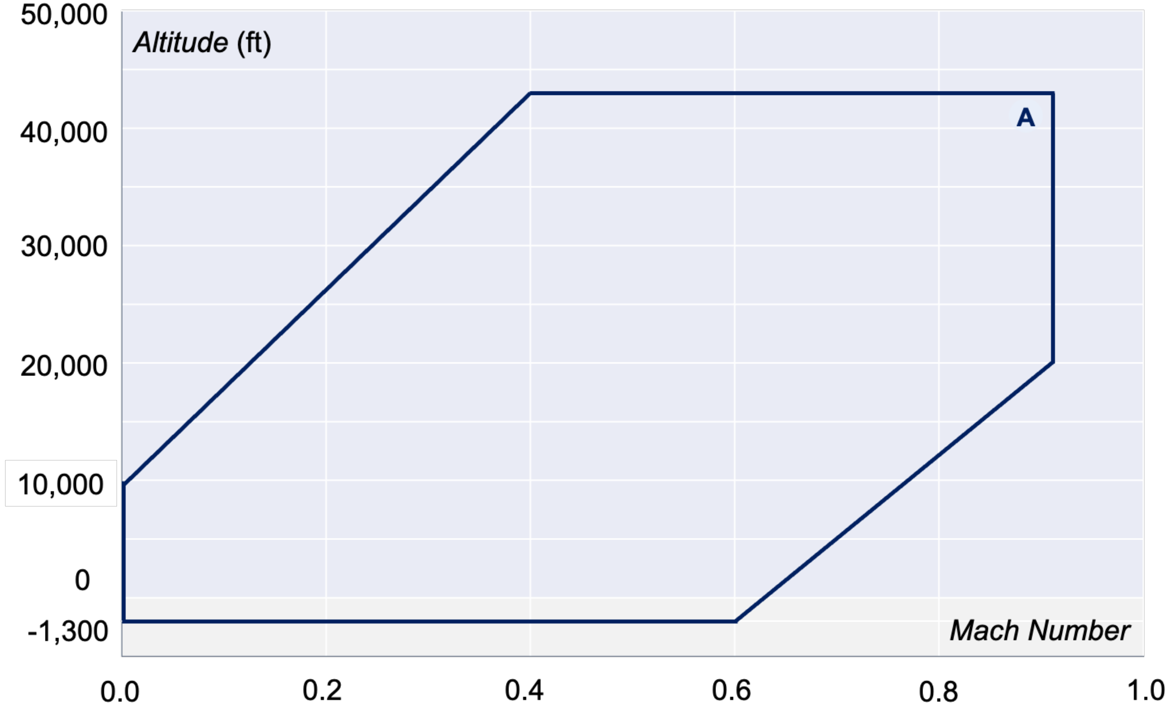
- Lowest layer of functional decomposition supporting a subsystem function

- Transition point from conventional system development and safety processes to ML Lifecycle Process



Operational Domain (OD)

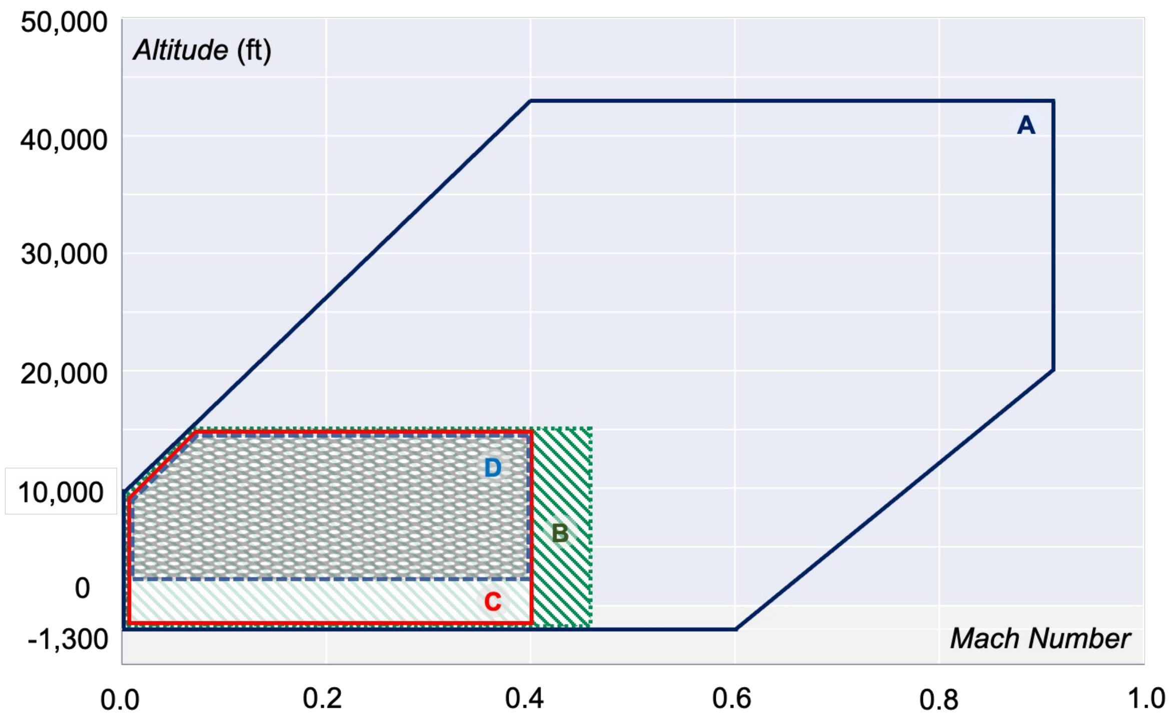
- Specification of all foreseeable conditions under which an end-product is expected to fulfill its missions
- Embodied in operational requirements
- Parameters: Environmental, Operational, System health; Values; Distributions



Region A: Full flight envelope / OD

Operational Design Domain (ODD)

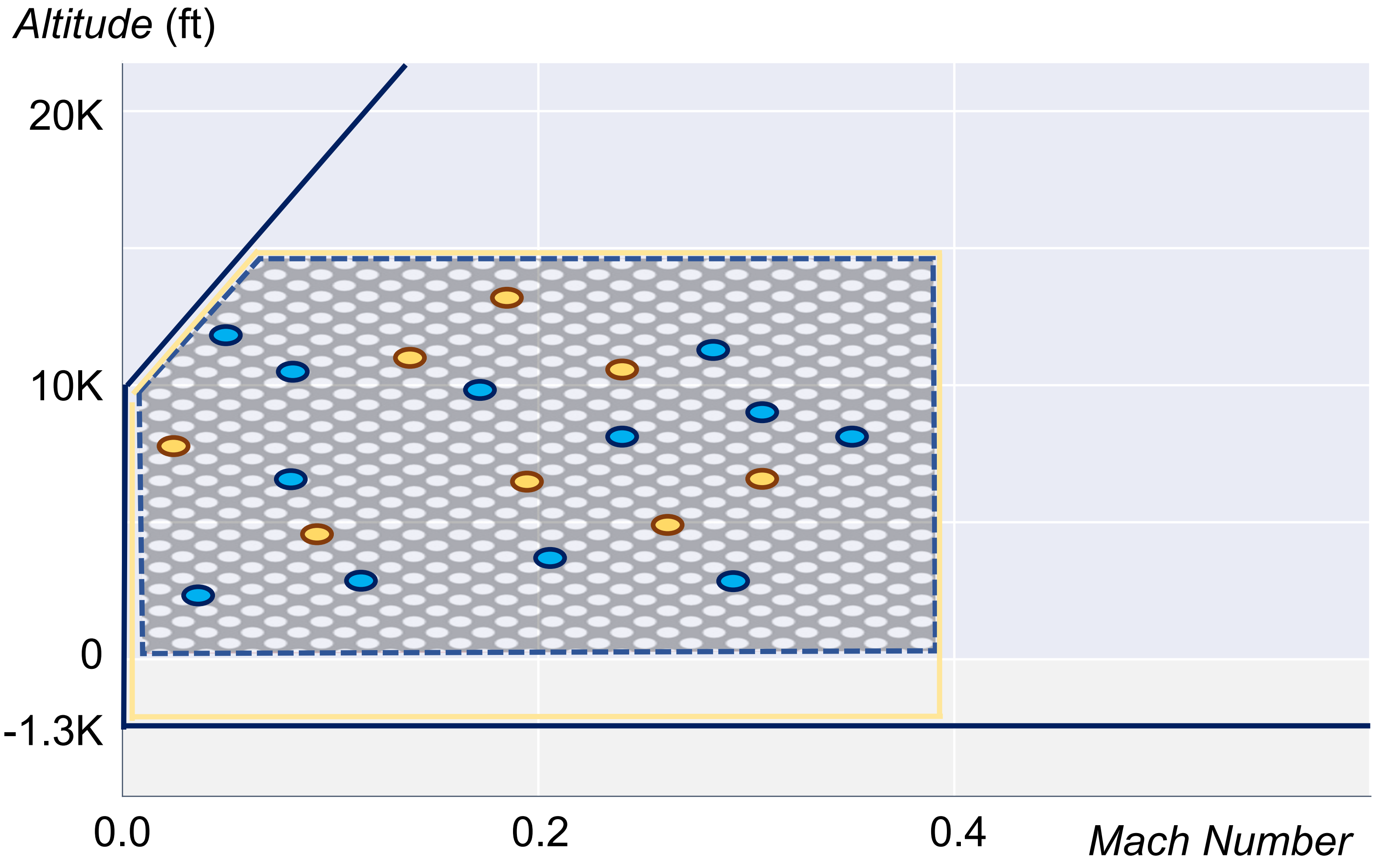
- An allocation of an OD (to a design layer of the product)
- Takeoff envelope allocated to MLC: As-specified MLC ODD (Region C)
- Takeoff envelope above sea level altitude allocated to MLM: As-specified MLM ODD (Region D)
- Takeoff envelope in operation: As-operated MLC ODD (Region B)

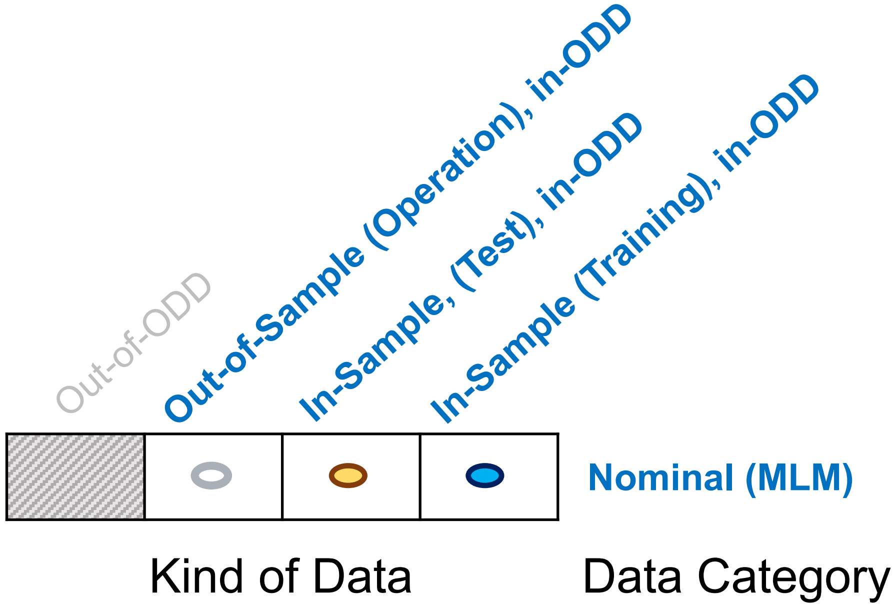


Outline

- Background
- Concepts
- Data centric characterization
- System layer analysis
- Conclusions and future work

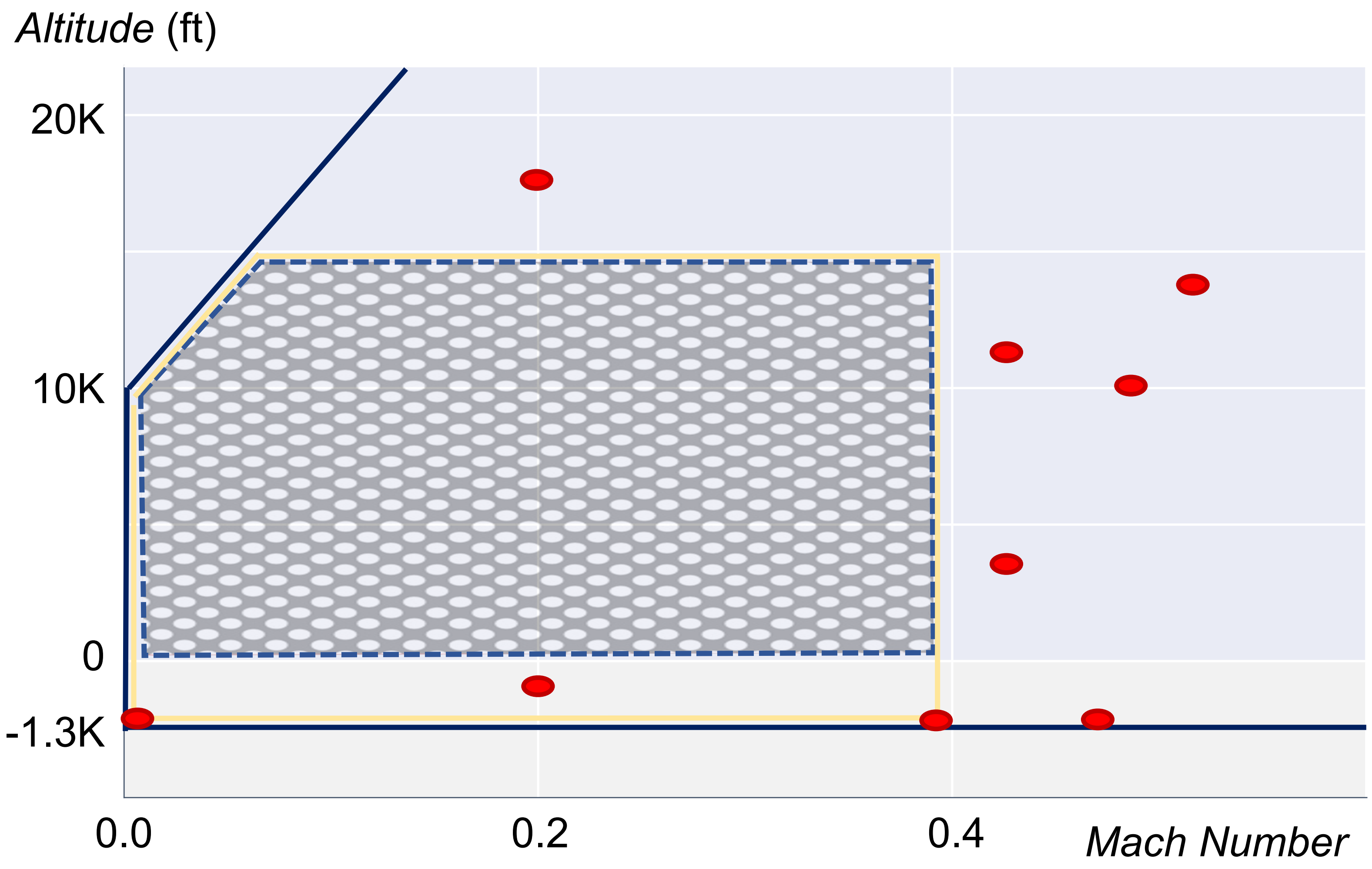
Data Centric Characterization – Nominal




Out-of-ODD
Out-of-Sample (Operation), in-ODD
In-Sample, (Test), in-ODD
In-Sample (Training), in-ODD
Nominal (MLM)
Kind of Data
Data Category

- **Generalization:** Produce required responses to in-sample, in-ODD, test dataset, and out-of-sample, in-ODD, operational data, after learning on in-sample in-ODD training data

Data Centric Characterization – Outlier



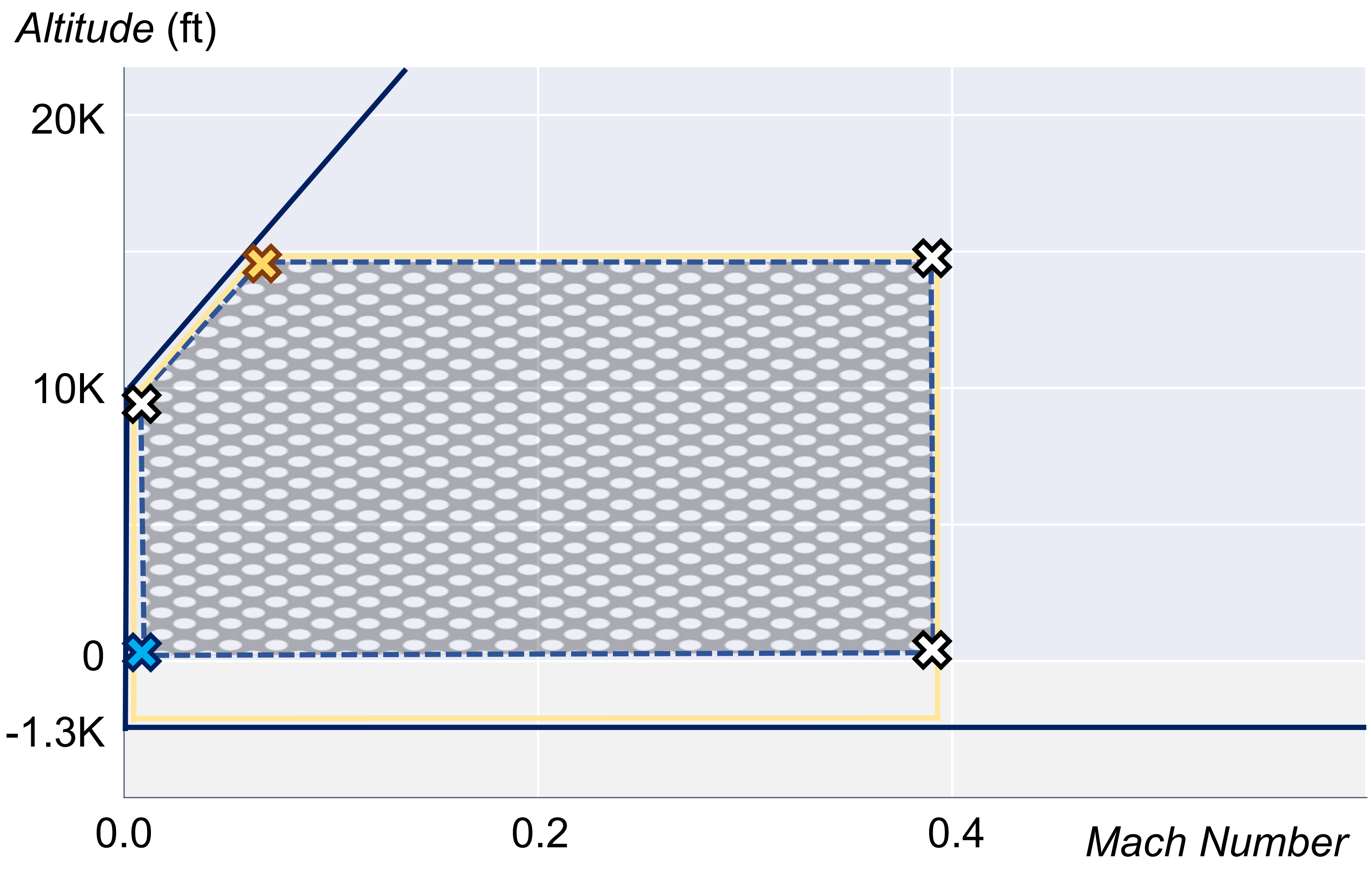
Data Category	Out-of-ODD	Out-of-Sample (Operation), in-ODD	In-Sample, (Test), in-ODD	In-Sample (Training), in-ODD
Nominal (MLM)	▨	○		
Outlier (MLM)	●	▨	▨	▨

Kind of Data

Data Category

- Producing a defined response to outliers changes functional intent and data category
- Therefore filter outliers from training data

Data Centric Characterization – Corner Case

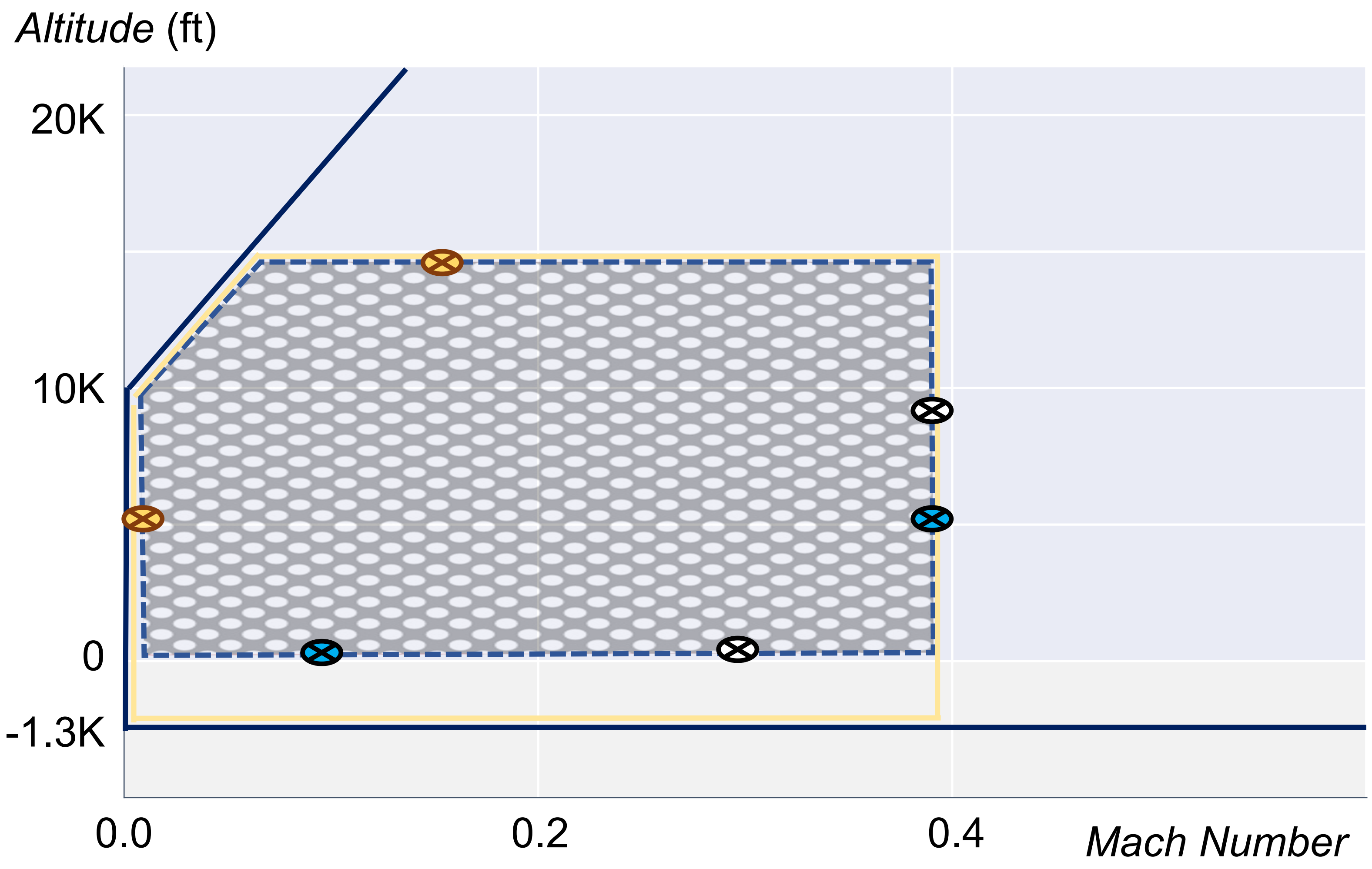


Out-of-ODD	Out-of-Sample (Operation), in-ODD	In-Sample, (Test), in-ODD	In-Sample (Training), in-ODD	
▨	○			Nominal (MLM)
	▨	▨	▨	Outlier
▨	✖	✖	✖	Corner Case (MLM)

Kind of Data

Data Category

Data Centric Characterization – Edge Case

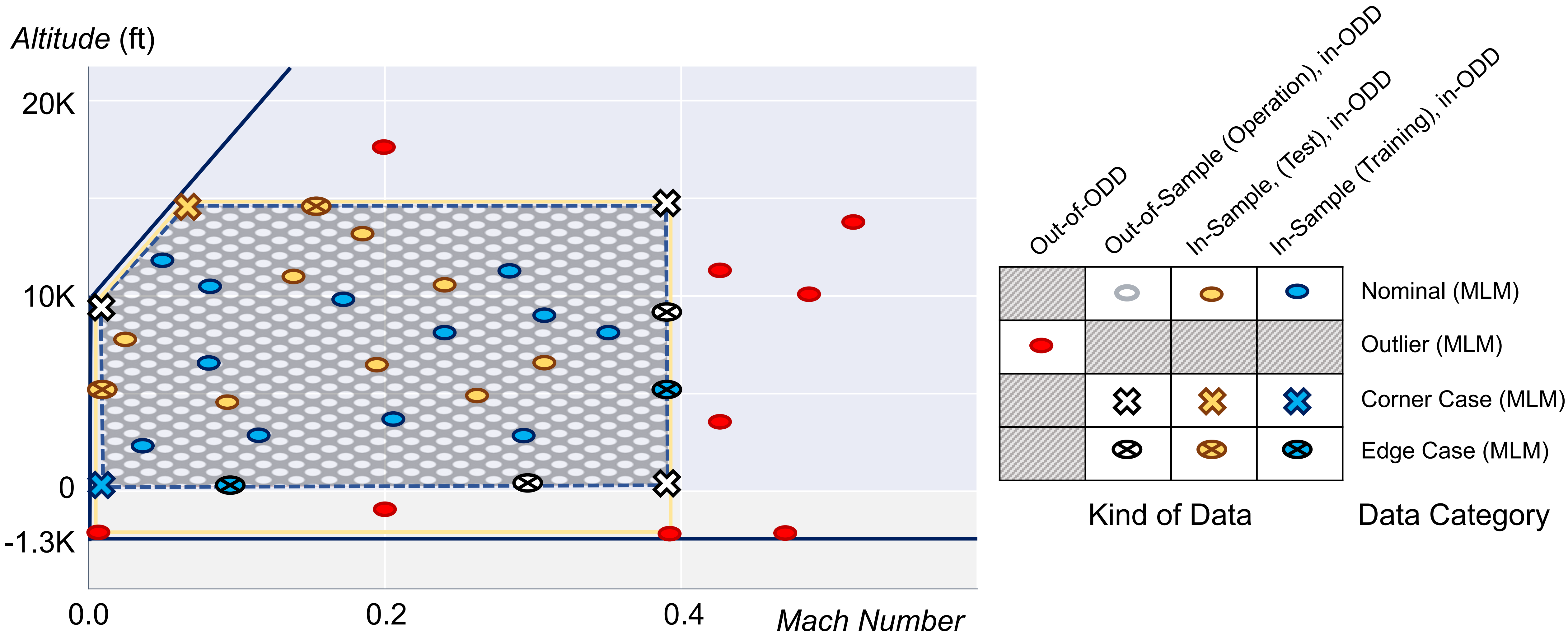


	Out-of-ODD	Out-of-Sample (Operation), in-ODD	In-Sample, (Test), in-ODD	In-Sample (Training), in-ODD
Nominal (MLM)	▨	○		
Outlier		▨	▨	▨
Corner Case	▨			
Edge Case (MLM)	▨	⊗	⊗	⊗

Kind of Data

Data Category

Data Centric Characterization – for MLM

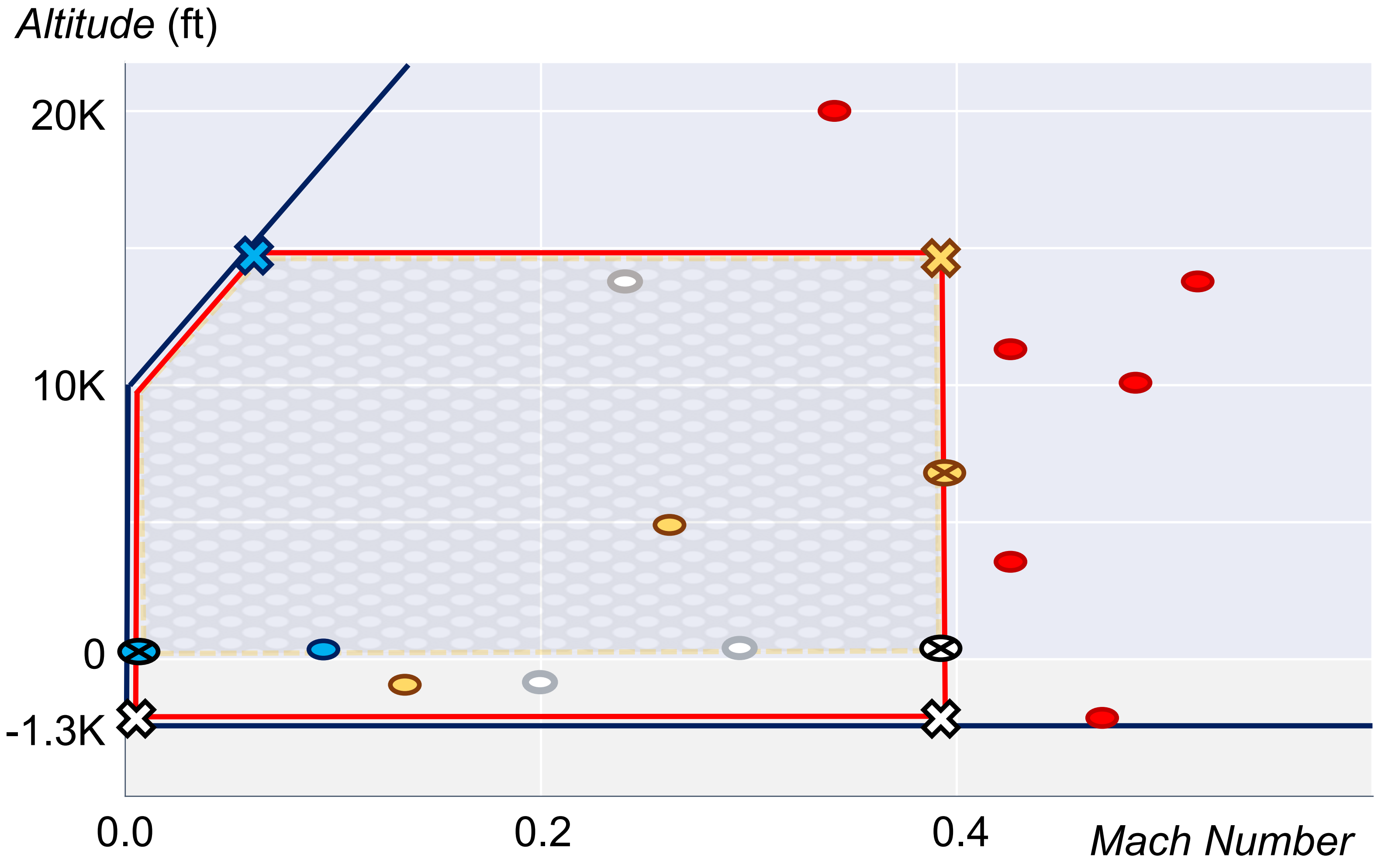


	Out-of-ODD	Out-of-Sample (Operation), in-ODD	In-Sample, (Test), in-ODD	In-Sample (Training), in-ODD
Nominal (MLM)	▨	○	●	●
Outlier (MLM)	●	▨	▨	▨
Corner Case (MLM)	▨	✖	✖	✖
Edge Case (MLM)	▨	⊗	⊗	⊗

Kind of Data

Data Category

Data Centric Characterization – for MLC



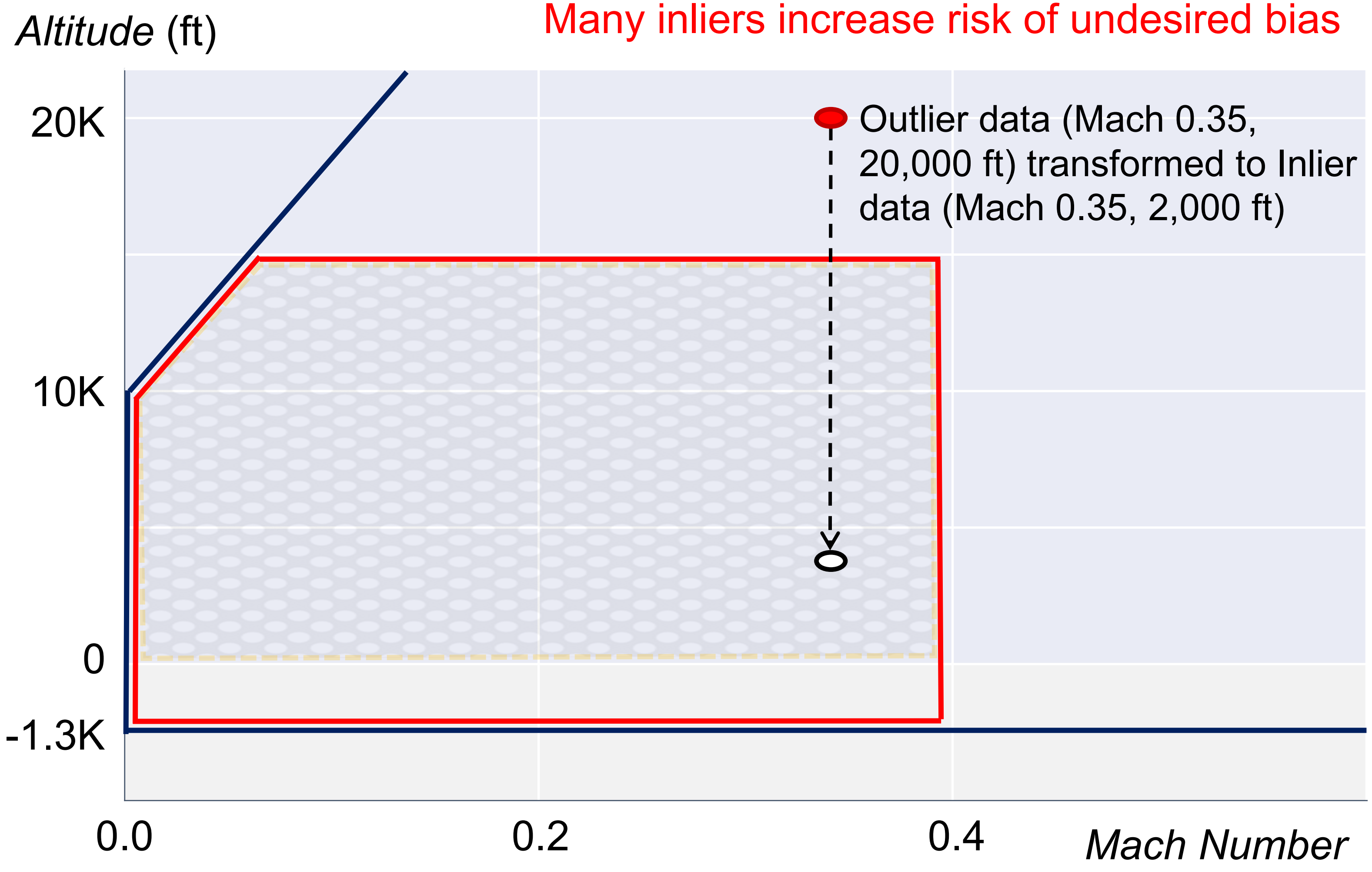
Out-of-ODD	Out-of-Sample (Operation), in-ODD	In-Sample, (Test), in-ODD	In-Sample (Training), in-ODD	
▨	○	●	●	Nominal (**MLC**)
●	▨	▨	▨	Outlier (**MLC**)
▨	✖	✖	✖	Corner Case (**MLC**)
▨	⊗	⊗	⊗	Edge Case (**MLC**)

Kind of Data

Data Category

Data Centric Characterization – Inlier

Many inliers increase risk of undesired bias



Outlier data (Mach 0.35, 20,000 ft) transformed to Inlier data (Mach 0.35, 2,000 ft)

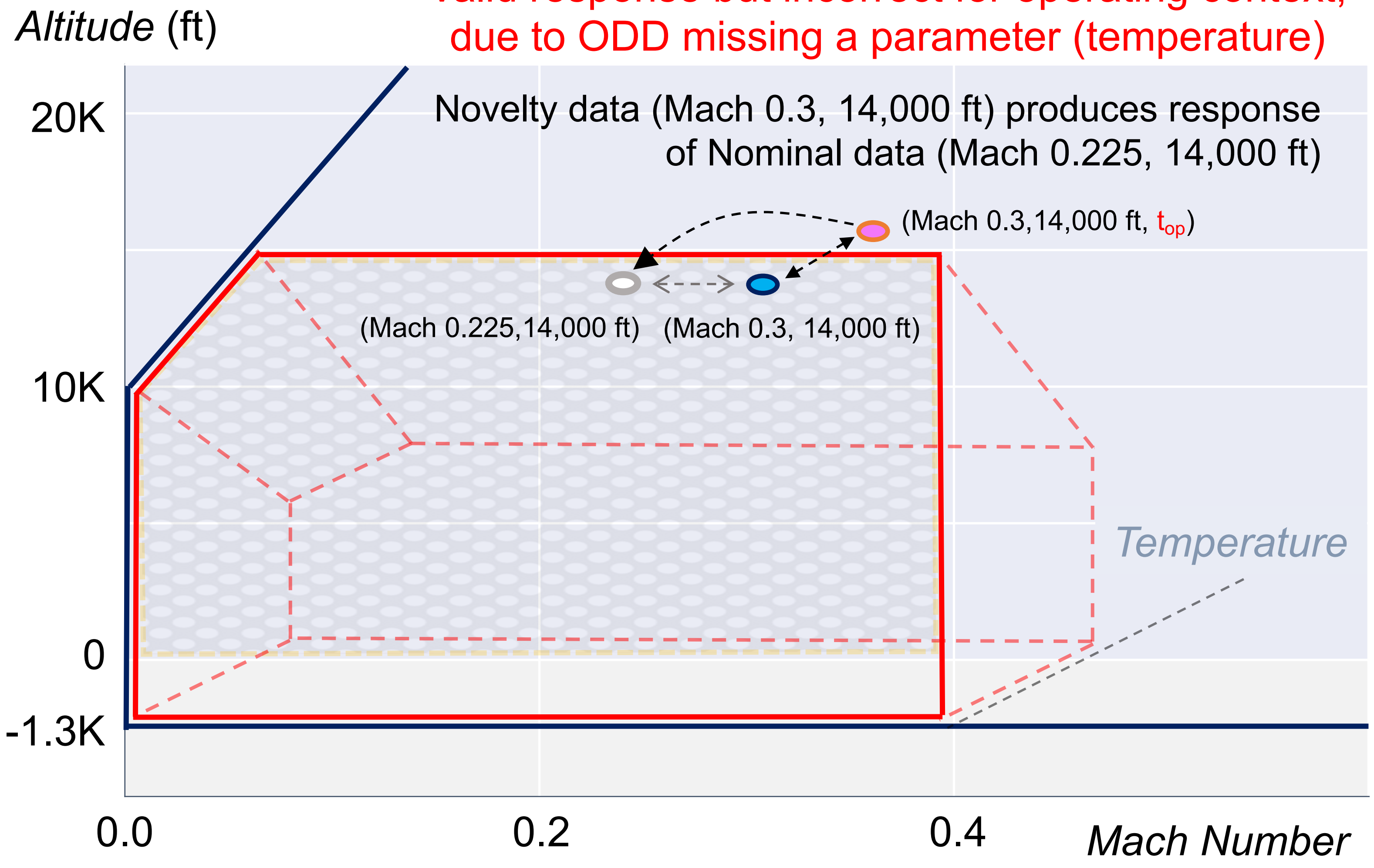
Data Category	Out-of-ODD	**Out-of-Sample (Operation), in-ODD**	In-Sample, (Test), in-ODD	In-Sample (Training), in-ODD
Nominal	▨			
Outlier		▨	▨	▨
Corner Case	▨			
Edge Case	▨			
Inlier (MLM, MLC)	▨	○	▨	▨

Kind of Data

Data Category

Data Centric Characterization – Novelty

Valid response but incorrect for operating context, due to ODD missing a parameter (temperature)



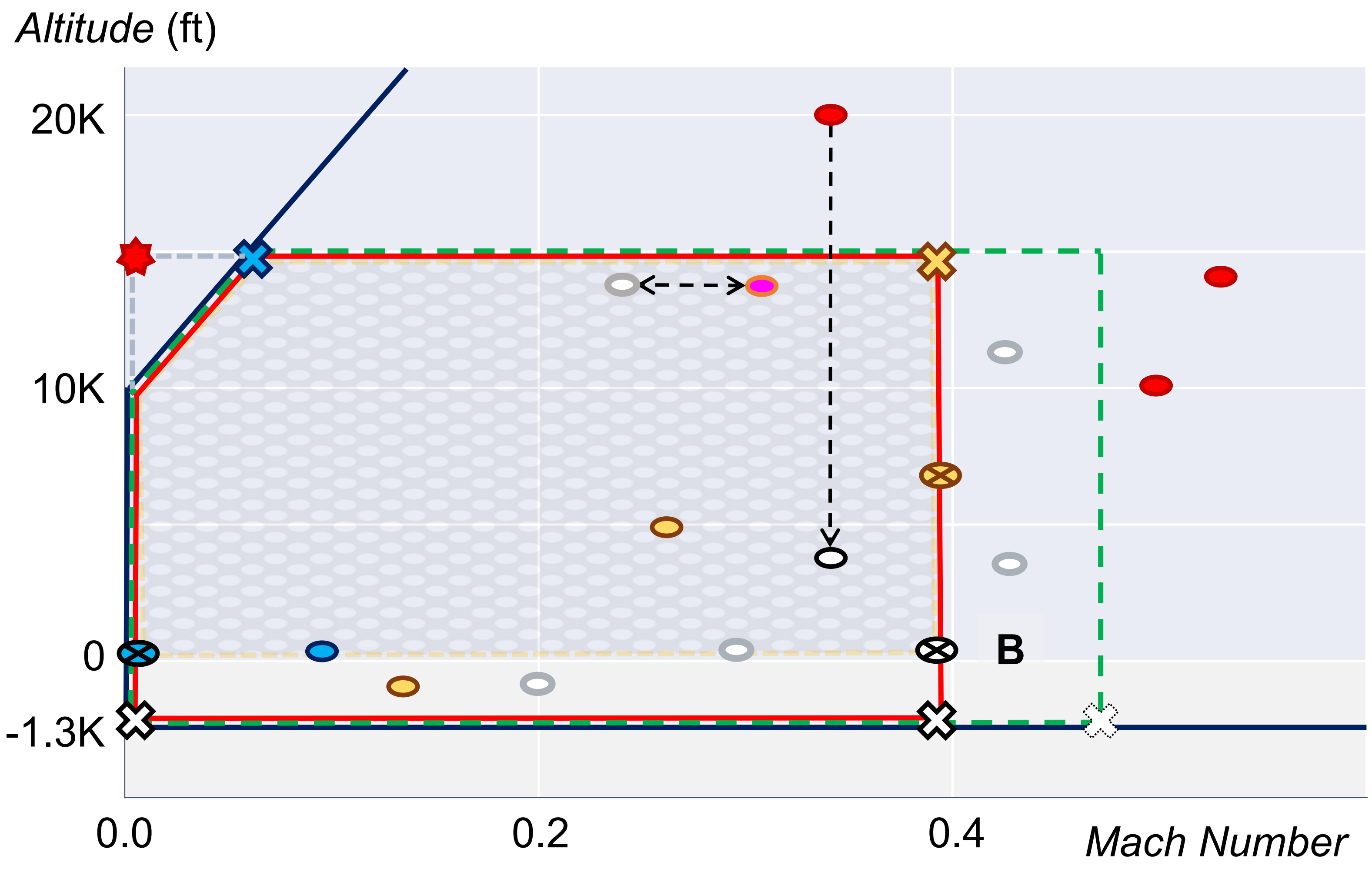
Novelty data (Mach 0.3, 14,000 ft) produces response of Nominal data (Mach 0.225, 14,000 ft)

Data Category	Out-of-ODD	Out-of-Sample (Operation), in-ODD	In-Sample, (Test), in-ODD	In-Sample (Training), in-ODD
Nominal	▨			
Outlier		▨	▨	▨
Corner Case	▨			
Edge Case	▨			
Inlier	▨		▨	▨
Novelty (MLM, MLC)	▨	●	▨	▨

Kind of Data

Data Category

Data Centric Characterization – Summary



Data Category	Out-of-ODD	Out-of-Sample (Operation), in-ODD	In-Sample, (Test), in-ODD	In-Sample (Training), in-ODD
Nominal (MLC)	▨	○	●	●
Outlier (MLC)	● ✶	▨	▨	▨
Corner Case (MLC)	▨	✖	✖	✖
Edge Case (MLC)	▨	⊗	⊗	⊗
Inlier (MLM, MLC)	▨	○	▨	▨
Novelty (MLM, MLC)	▨	●	▨	▨

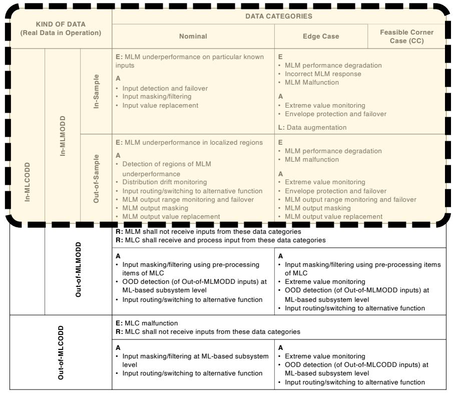
Kind of Data

Data Category

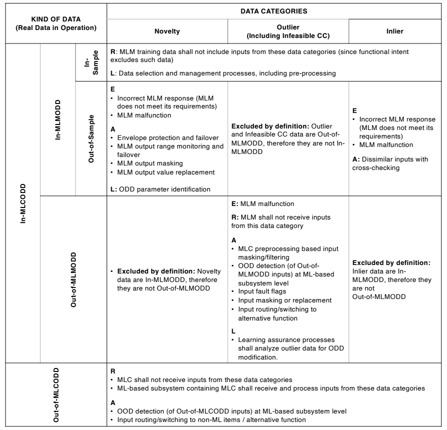
Outline

- Background
- Concepts
- Data centric characterization
- System layer analysis
- Conclusions and future work

Supporting System Layer Analyses



KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Nominal	Edge Case	Feasible Corner Case (CC)
In-MLCODD	In-MLMODD	In-Sample	**E:** MLM underperformance on particular known inputs **A** • Input detection and failover • Input masking/filtering • Input value replacement	**E** • MLM performance degradation • Incorrect MLM response • MLM Malfunction **A** • Extreme value monitoring • Envelope protection and failover **L:** Data augmentation	
		Out-of-Sample	**E:** MLM underperformance in localized regions **A** • Detection of regions of MLM underperformance • Distribution drift monitoring • Input routing/switching to alternative function • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	**E** • MLM performance degradation • MLM malfunction **A** • Extreme value monitoring • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	
	Out-of-MLMODD		**R:** MLM shall not receive inputs from these data categories **R:** MLC shall receive and process input from these data categories		
			A • Input masking/filtering using pre-processing items of MLC • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	**A** • Input masking/filtering using pre-processing items of MLC • Extreme value monitoring • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	
Out-of-MLCODD			**E:** MLC malfunction **R:** MLC shall not receive inputs from these data categories		
			A • Input masking/filtering at ML-based subsystem level • Input routing/switching to alternative function	**A** • Extreme value monitoring • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	



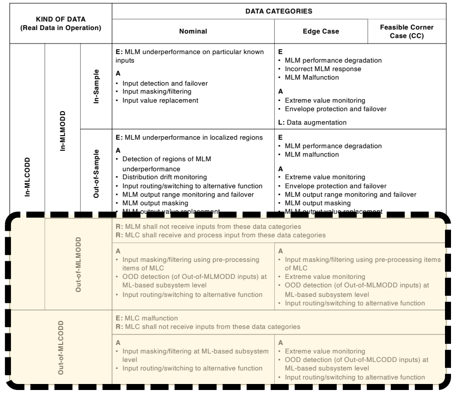
KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Novelty	Outlier (Including Infeasible CC)	Inlier
In-MLCODD	In-MLMODD	In-Sample	**R**: MLM training data shall not include inputs from these data categories (since functional intent excludes such data) **L**: Data selection and management processes, including pre-processing		
		Out-of-Sample	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A** • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement **L:** ODD parameter identification	**Excluded by definition:** Outlier and Infeasible CC data are Out-of-MLMODD, therefore they are not In-MLMODD	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A:** Dissimilar inputs with cross-checking
	Out-of-MLMODD		• **Excluded by definition:** Novelty data are In-MLMODD, therefore they are not Out-of-MLMODD	**E:** MLM malfunction **R:** MLM shall not receive inputs from this data category **A** • MLC preprocessing based input masking/filtering • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input fault flags • Input masking or replacement • Input routing/switching to alternative function **L** • Learning assurance processes shall analyze outlier data for ODD modification.	**Excluded by definition:** Inlier data are In-MLMODD, therefore they are not Out-of-MLMODD
Out-of-MLCODD			**R** • MLC shall not receive inputs from these data categories • ML-based subsystem containing MLC shall receive and process inputs from these data categories **A** • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to non-ML items / alternative function		

KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Nominal	Edge Case	Feasible Corner Case (CC)
In-MLCODD	In-MLMODD	In-Sample	**E:** MLM underperformance on particular known inputs **A** • Input detection and failover • Input masking/filtering • Input value replacement	**E** • MLM performance degradation • Incorrect MLM response • MLM Malfunction **A** • Extreme value monitoring • Envelope protection and failover **L:** Data augmentation	
		Out-of-Sample	**E:** MLM underperformance in localized regions **A** • Detection of regions of MLM underperformance • Distribution drift monitoring • Input routing/switching to alternative function • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	**E** • MLM performance degradation • MLM malfunction **A** • Extreme value monitoring • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	

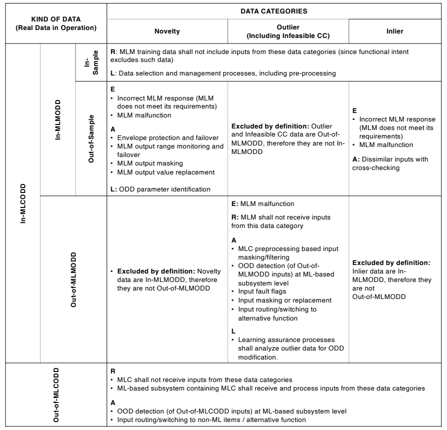
Potential effects of data

Partition of ODD as characterized by Data Kind x Category, results of system layer analyses

Supporting System Layer Analyses



KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Nominal	Edge Case	Feasible Corner Case (CC)
In-MLCODD	In-MLMODD	In-Sample	**E:** MLM underperformance on particular known inputs **A** • Input detection and failover • Input masking/filtering • Input value replacement	**E** • MLM performance degradation • Incorrect MLM response • MLM Malfunction **A** • Extreme value monitoring • Envelope protection and failover **L:** Data augmentation	
		Out-of-Sample	**E:** MLM underperformance in localized regions **A** • Detection of regions of MLM underperformance • Distribution drift monitoring • Input routing/switching to alternative function • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	**E** • MLM performance degradation • MLM malfunction **A** • Extreme value monitoring • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	
	Out-of-MLMODD		**R:** MLM shall not receive inputs from these data categories **R:** MLC shall receive and process input from these data categories		
			A • Input masking/filtering using pre-processing items of MLC • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	**A** • Input masking/filtering using pre-processing items of MLC • Extreme value monitoring • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	
Out-of-MLCODD			**E:** MLC malfunction **R:** MLC shall not receive inputs from these data categories		
			A • Input masking/filtering at ML-based subsystem level • Input routing/switching to alternative function	**A** • Extreme value monitoring • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	

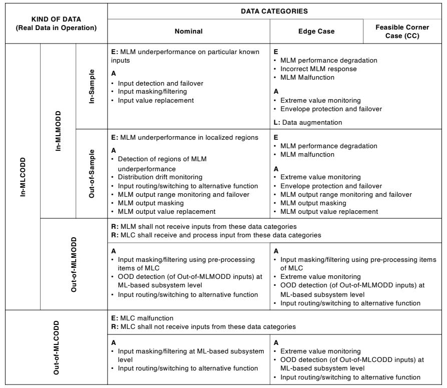


KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Novelty	Outlier (Including Infeasible CC)	Inlier
In-MLCODD	In-MLMODD	In-Sample	**R**: MLM training data shall not include inputs from these data categories (since functional intent excludes such data) **L**: Data selection and management processes, including pre-processing		
		Out-of-Sample	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A** • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement **L:** ODD parameter identification	**Excluded by definition:** Outlier and Infeasible CC data are Out-of-MLMODD, therefore they are not In-MLMODD	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A:** Dissimilar inputs with cross-checking
	Out-of-MLMODD		• **Excluded by definition:** Novelty data are In-MLMODD, therefore they are not Out-of-MLMODD	**E:** MLM malfunction **R:** MLM shall not receive inputs from this data category **A** • MLC preprocessing based input masking/filtering • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input fault flags • Input masking or replacement • Input routing/switching to alternative function **L** • Learning assurance processes shall analyze outlier data for ODD modification.	**Excluded by definition:** Inlier data are In-MLMODD, therefore they are not Out-of-MLMODD
Out-of-MLCODD			**R** • MLC shall not receive inputs from these data categories • ML-based subsystem containing MLC shall receive and process inputs from these data categories **A** • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to non-ML items / alternative function		

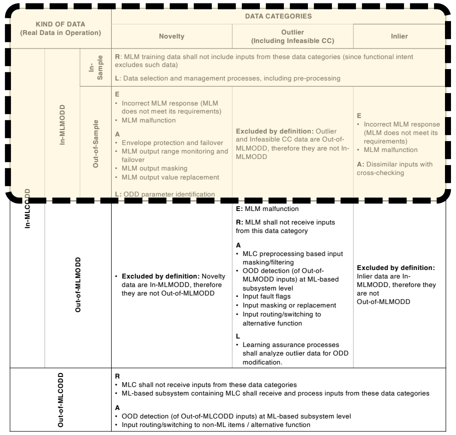
KIND OF DATA (Real Data in Operation)		DATA CATEGORIES		
		Nominal	Edge Case	Feasible Corner Case (CC)
In-MLCODD	Out-of-MLMODD	**R:** MLM shall not receive inputs from these data categories **R:** MLC shall receive and process input from these data categories		
		A • Input masking/filtering using pre-processing items of MLC • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	**A** • Input masking/filtering using pre-processing items of MLC • Extreme value monitoring • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	
Out-of-MLCODD		**E:** MLC malfunction **R:** MLC shall not receive inputs from these data categories		
		A • Input masking/filtering at ML-based subsystem level • Input routing/switching to alternative function	**A** • Extreme value monitoring • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	

Architecture modifications

Supporting System Layer Analyses



KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Nominal	Edge Case	Feasible Corner Case (CC)
In-MLCODD	In-MLMODD	In-Sample	**E:** MLM underperformance on particular known inputs **A** • Input detection and failover • Input masking/filtering • Input value replacement	**E** • MLM performance degradation • Incorrect MLM response • MLM Malfunction **A** • Extreme value monitoring • Envelope protection and failover **L:** Data augmentation	
		Out-of-Sample	**E:** MLM underperformance in localized regions **A** • Detection of regions of MLM underperformance • Distribution drift monitoring • Input routing/switching to alternative function • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	**E** • MLM performance degradation • MLM malfunction **A** • Extreme value monitoring • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	
	Out-of-MLMODD		**R:** MLM shall not receive inputs from these data categories **R:** MLC shall receive and process input from these data categories		
			A • Input masking/filtering using pre-processing items of MLC • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	**A** • Input masking/filtering using pre-processing items of MLC • Extreme value monitoring • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	
Out-of-MLCODD			**E:** MLC malfunction **R:** MLC shall not receive inputs from these data categories		
			A • Input masking/filtering at ML-based subsystem level • Input routing/switching to alternative function	**A** • Extreme value monitoring • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	

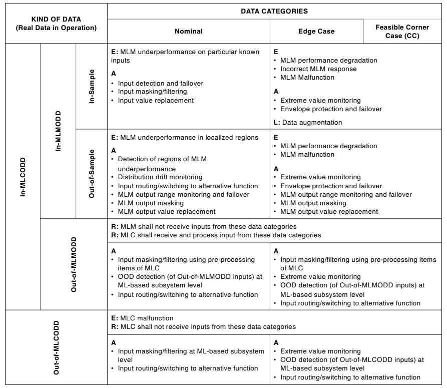


KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Novelty	Outlier (Including Infeasible CC)	Inlier
In-MLCODD	In-MLMODD	In-Sample	**R**: MLM training data shall not include inputs from these data categories (since functional intent excludes such data) **L**: Data selection and management processes, including pre-processing		
		Out-of-Sample	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A** • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement **L:** ODD parameter identification	**Excluded by definition:** Outlier and Infeasible CC data are Out-of-MLMODD, therefore they are not In-MLMODD	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A:** Dissimilar inputs with cross-checking
	Out-of-MLMODD		• **Excluded by definition:** Novelty data are In-MLMODD, therefore they are not Out-of-MLMODD	**E:** MLM malfunction **R:** MLM shall not receive inputs from this data category **A** • MLC preprocessing based input masking/filtering • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input fault flags • Input masking or replacement • Input routing/switching to alternative function **L** • Learning assurance processes shall analyze outlier data for ODD modification.	**Excluded by definition:** Inlier data are In-MLMODD, therefore they are not Out-of-MLMODD
Out-of-MLCODD			**R** • MLC shall not receive inputs from these data categories • ML-based subsystem containing MLC shall receive and process inputs from these data categories **A** • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to non-ML items / alternative function		

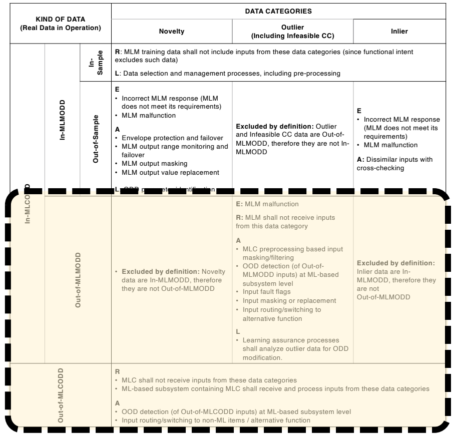
KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Novelty	Outlier (Including Infeasible CC)	Inlier
In-MLCODD	In-MLMODD	In-Sample	**R**: MLM training data shall not include inputs from these data categories (since functional intent excludes such data) **L**: Data selection and management processes, including pre-processing		
		Out-of-Sample	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A** • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement **L:** ODD parameter identification	**Excluded by definition:** Outlier and Infeasible CC data are Out-of-MLMODD, therefore they are not In-MLMODD	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A:** Dissimilar inputs with cross-checking

Learning Assurance Steps

Supporting System Layer Analyses



KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Nominal	Edge Case	Feasible Corner Case (CC)
In-MLCODD	In-MLMODD	In-Sample	**E:** MLM underperformance on particular known inputs **A** • Input detection and failover • Input masking/filtering • Input value replacement	**E** • MLM performance degradation • Incorrect MLM response • MLM Malfunction **A** • Extreme value monitoring • Envelope protection and failover **L:** Data augmentation	
		Out-of-Sample	**E:** MLM underperformance in localized regions **A** • Detection of regions of MLM underperformance • Distribution drift monitoring • Input routing/switching to alternative function • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	**E** • MLM performance degradation • MLM malfunction **A** • Extreme value monitoring • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement	
	Out-of-MLMODD		**R:** MLM shall not receive inputs from these data categories **R:** MLC shall receive and process input from these data categories		
			A • Input masking/filtering using pre-processing items of MLC • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	**A** • Input masking/filtering using pre-processing items of MLC • Extreme value monitoring • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	
Out-of-MLCODD			**E:** MLC malfunction **R:** MLC shall not receive inputs from these data categories		
			A • Input masking/filtering at ML-based subsystem level • Input routing/switching to alternative function	**A** • Extreme value monitoring • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to alternative function	



KIND OF DATA (Real Data in Operation)			DATA CATEGORIES		
			Novelty	Outlier (Including Infeasible CC)	Inlier
In-MLCODD	In-MLMODD	In-Sample	**R**: MLM training data shall not include inputs from these data categories (since functional intent excludes such data) **L**: Data selection and management processes, including pre-processing		
		Out-of-Sample	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A** • Envelope protection and failover • MLM output range monitoring and failover • MLM output masking • MLM output value replacement **L**: ODD parameter identification	**Excluded by definition:** Outlier and Infeasible CC data are Out-of-MLMODD, therefore they are not In-MLMODD	**E** • Incorrect MLM response (MLM does not meet its requirements) • MLM malfunction **A:** Dissimilar inputs with cross-checking
	Out-of-MLMODD		• **Excluded by definition:** Novelty data are In-MLMODD, therefore they are not Out-of-MLMODD	**E:** MLM malfunction **R:** MLM shall not receive inputs from this data category **A** • MLC preprocessing based input masking/filtering • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input fault flags • Input masking or replacement • Input routing/switching to alternative function **L** • Learning assurance processes shall analyze outlier data for ODD modification.	**Excluded by definition:** Inlier data are In-MLMODD, therefore they are not Out-of-MLMODD
	Out-of-MLCODD		**R** • MLC shall not receive inputs from these data categories • ML-based subsystem containing MLC shall receive and process inputs from these data categories **A** • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to non-ML items / alternative function		

KIND OF DATA (Real Data in Operation)		DATA CATEGORIES		
		Novelty	Outlier (Including Infeasible CC)	Inlier
In-MLCODD	Out-of-MLMODD	• **Excluded by definition:** Novelty data are In-MLMODD, therefore they are not Out-of-MLMODD	**R:** MLM shall not receive inputs from this data category **A** • MLC preprocessing based input masking/filtering • OOD detection (of Out-of-MLMODD inputs) at ML-based subsystem level • Input fault flags • Input masking or replacement • Input routing/switching to alternative function **L** • Learning assurance processes shall analyze outlier data for ODD modification.	**Excluded by definition:** Inlier data are In-MLMODD, therefore they are not Out-of-MLMODD
	Out-of-MLCODD	Requirements **R** • MLC shall not receive inputs from these data categories • ML-based subsystem containing MLC shall receive and process inputs from these data categories **A** • OOD detection (of Out-of-MLCODD inputs) at ML-based subsystem level • Input routing/switching to non-ML items / alternative function		

Outline

- Background
- Concepts
- Data centric characterization
- System layer analysis
- Conclusions and future work

Conclusions

- Rigorous data centric characterization of ODD concept using categories and kinds of data to partition and analyze
- Complementary to scenario-based approaches developed in the automotive domain
- Consensus position of aviation industry, anchoring concept in forthcoming process assurance guidance ARP 6983
- Could be applicable in other domains
- Real world validation ongoing (safe flight termination, airborne collision avoidance, time-based separation of transport aircraft in terminal environments)

Future Work

- Ongoing work to formalize ODD concept and data category definitions using topology theory
- Formalization of desirable properties: coverage of ODD, internal completeness
- Multiplicity of MLM / MLC and corresponding ODDs including overlaps and transitions
- Definition of underlying process for MLCODD characterization
- Relationship to equivalence classes and other ways of partitioning ODDs

Content of a forthcoming Journal paper

Acknowledgements

Members of the ODD working group in SAE G-34 and EUROCAE WG-114 contributed their time and expertise in the discussions leading to this paper