Challenges in Validating Complex Large-Scale Simulations for Operational Test

Karl Kuschner, Ph.D. (IDA)
Jay Smith, Ph.D. (JOTT)
Ben Lambert, Ph.D. (JOTT)
Jeremy Smith (JPO/NAVAIR)

NAVAIR Public Release 2017-795
Distribution Statement A – "Approved for public release"

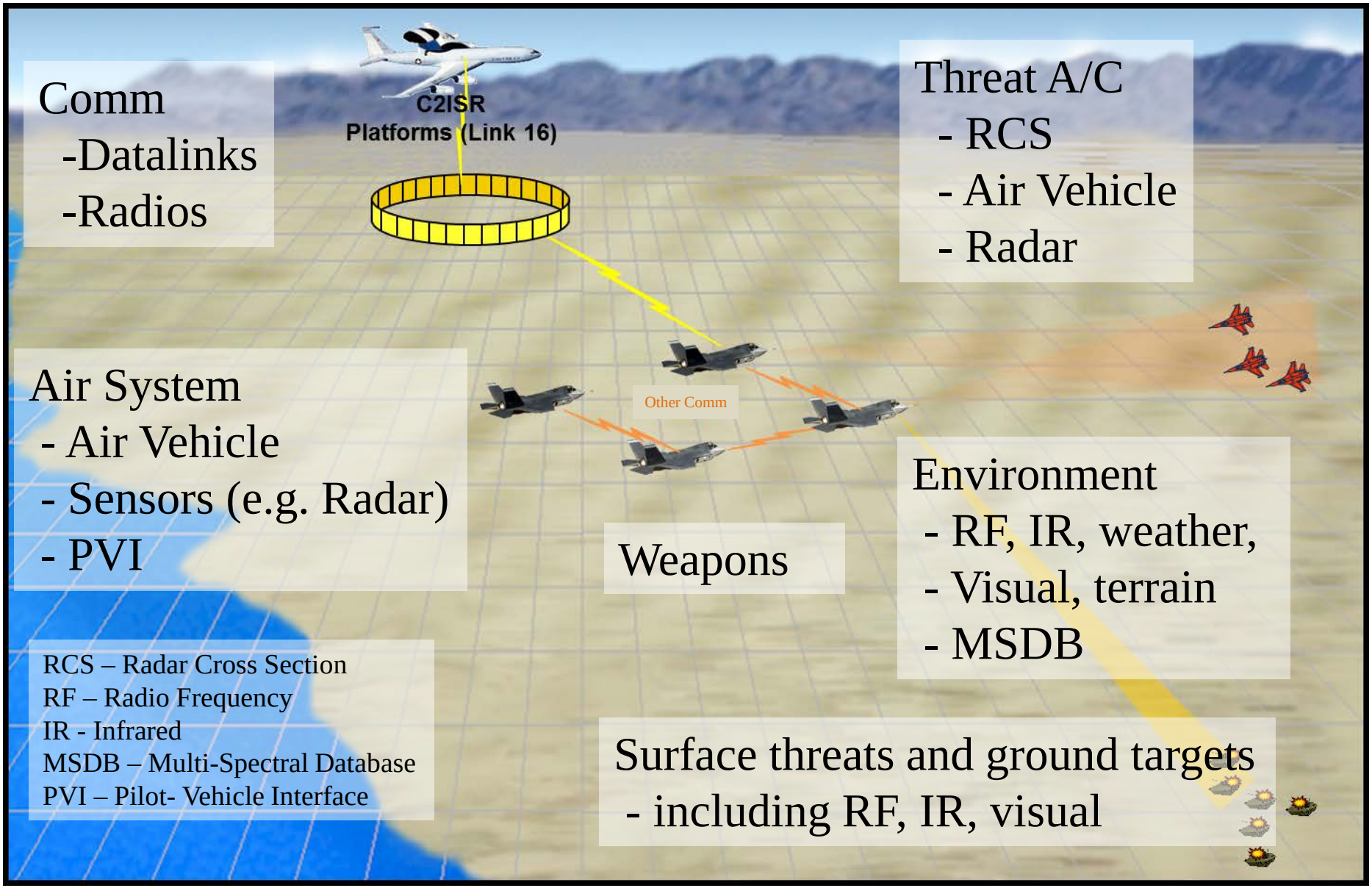
JPO Public Release Number JSF17-906
DISTRIBUTION A. Approved for public release: distribution unlimited.

CLEARED
For Open Publication
Sep 19, 2017
Department of Defense
OFFICE OF PREPUBLICATION AND SECURITY REVIEW

Background

- Need for these sims
  - Cost and limitations of live testing
- Intended use
  - Missions
  - Data collection
  - Number of trials
- Major requirements
  - Operational Flight Program (OFP)-in-the-loop
  - Man-in-the-loop threats

Example Synthetic Environment



Comm
- Datalinks
- Radios

Threat A/C
- RCS
- Air Vehicle
- Radar

Air System
- Air Vehicle
- Sensors (e.g. Radar)
- PVI

Weapons

Environment
- RF, IR, weather,
- Visual, terrain
- MSDB

RCS – Radar Cross Section
RF – Radio Frequency
IR - Infrared
MSDB – Multi-Spectral Database
PVI – Pilot- Vehicle Interface

Surface threats and ground targets
- including RF, IR, visual

Federations of Models

- Various environments
- Tying together with open interfaces

Early Validation Issues

- Typical issues encountered
  - User expectations
  - Sim instrumentation
  - Flight test data issues
  - Data tools
  - Model tuning

Accreditation Approach

- Accreditation challenges
  - Sim to be used to answer mission-level questions
    - Was the aircraft able to meet the Commander's intent?
    - What was the blue:red kill ratio?
  - Sim to be used in mission threat scenarios outside of that available on open-air ranges
    - Available mission-level V&V reference data limited
    - Different factors important to executing low-end (open-air) scenarios versus high-end (simulator) scenarios

Accreditation Approach

- How will Sim be accredited for the intended use?
  - Require detailed V&V to be conducted at the ‘component’ and ‘system’ level to assess risk of use at the mission level.
  - Accreditation risk would be a roll-up of risks at the component and system levels.

Mission Decomposition

Mission Level	**Mission**	**Defensive Counter Air (DCA)**
	Metric	**Proportion of Red Strikers that do not reach their Weapons Release Point**
System Level	**Interaction**	**Air-Air Kill Chain**
	Link	**Track**
	Metric	**System Track Accuracy**
Component Level	**Component**	**Radar**
	Mode	**Multi-Target Track (MTT)**
	Metric	**Track Accuracy**

Accreditation Methodology

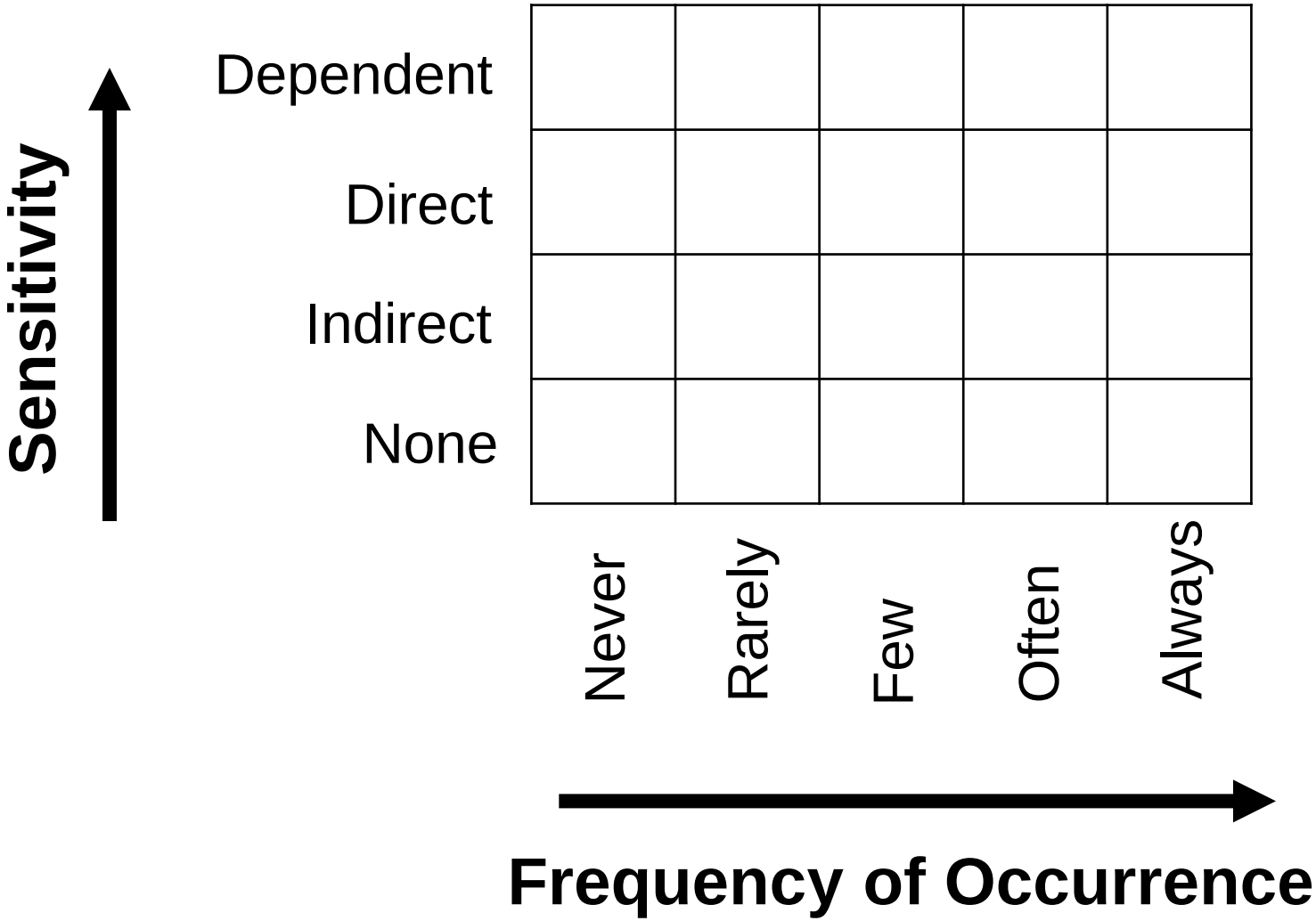
- Accreditation split into 3 parts:

Assess risk based on sensitivity and occurrence	• How important is the validation metric to mission-level evaluations?
Assess risk based on the size of the performance difference	• Is there a tactical impact?
Assess risk based on ability to mitigate	• Can model be fixed, tweaked, or re-tuned? • Can the risk be reduced by limiting the scope of the intended use?

Mission Interactions

Level	Task	Metric	Questions
Mission-Level	DCA	Proportion of Red Strikers that do not reach their Weapons Release Point	• How sensitive is this mission metric to a performance delta in this system metric? • How often will this mission metric include this system metric?
System-Level	A/A Kill Chain, System Track	System Track Accuracy	• How sensitive is this system metric to a performance difference in this component metric? • How often will this system metric include this component metric?
Component-Level	Radar, Multi-Target Track	Track Accuracy	

Interaction Importance



Statistical Approach

- Primarily at Component Level
- Perform a 2-sample non-parametric statistical test to identify if real-world (reference) and sim data could be from the same statistical distribution; minimum data size required
- Null Hypothesis: H_0 = reference data and sim data come from same distribution
- 'Pass' = failed to reject H_0, and minimum data requirements met; <u>no additional action required</u> (unless users demand further data)
- 'Fail' = rejected H_0
  - There is a performance difference between reference and sim performance, or failed to meet minimum data requirements
  - Assess potential tactical impact, and mitigation options

Value of Validation Test Design

- Structured engineering approach to VV&A
  - Accreditation Plan and V&V Plan interlinked
- Individual component, system, and mission test designs
  - Measures (detection range, false alarm rate, …)
  - Conditions (altitude, RCS, …)
  - Levels (high/low, large/small, …)
- VTDs provided reference matrix for ‘data miners’ to fill
  - Levels purposefully vague to maximize chances of finding suitable reference data
  - Gaps in reference data availability easily identified

Detailed V&V Plans

- The V&V plan focuses on:
  - How to determine what V&V needs to be done?
  - What V&V needs to be done?
  - What is needed in order to do the V&V?
  - What might affect V&V team's ability to do the V&V?
- Detailed V&V plans are critical as they enable assessment of the likelihood of successful accreditation
  - Aids user test planning, and adequate resource allocation
    - When is it likely to be able to be used?
    - For what uses will it likely be capable of being accredited?
  - Enables progress monitoring
    - Will it be ready on time?
  - Demonstrates competency of V&V agent

User Perspective

- Combined V&V and accreditation approach
  - Helps manage expectations
  - Reduces risk of misunderstanding
  - Accreditation often closely aligned with user, who is most knowledgeable on intended use
- Significantly more accreditation/user involvement than planned
- Efficient VV&A process dependent on externals:
  - Networks
  - Personnel

Value of a Government-led V&V

- Developers not validating their own product
  - Ensures separation even for Government-developed models
- Better understanding of intended simulation employment/mission use
- The Government generally has fewer barriers to information access
  - Security
  - Proprietary safeguarding
- Lessons learned can directly benefit future V&V efforts
  - Tools and processes
  - Robust team of subject matter experts
  - Complex organizational connections

Actions to ensure successful V&V

- Identify, pull, and analyze flight test data/reports to characterize real-world performance
- Perform continuous gap analysis – check validation data needs for real-world coverage and attempt to fill gaps
- Develop simulation scenarios
  - Balance efficiency with rigorous replication of flight test as executed
- Analyze simulation data
  - Generate the measures from the test design
- Perform comparative analysis
- Report findings
- Perform accreditation assessment

V&V Issues

- Scarcity of available expertise
  - V&V skills with SME knowledge is rare combination
  - Many projects require similar expertise
- Network connectivity and robustness
- Data
  - Discovery, validity for OT, and usability
- Contracting language and specifications
- Syncing with program schedules
- V&V timeline compression

Path to Overcoming Challenges

- Build and acquire expertise
- Ensure tuning is part of model development
- Plan rigorous V&V effort in detail
- Use enterprise to support V&V
- Stay tightly coupled with stakeholders