Cognitive Support for Airline Operations in Complex Environments

Dr. Robert Mauro,

- Decision Research &
- University of Oregon

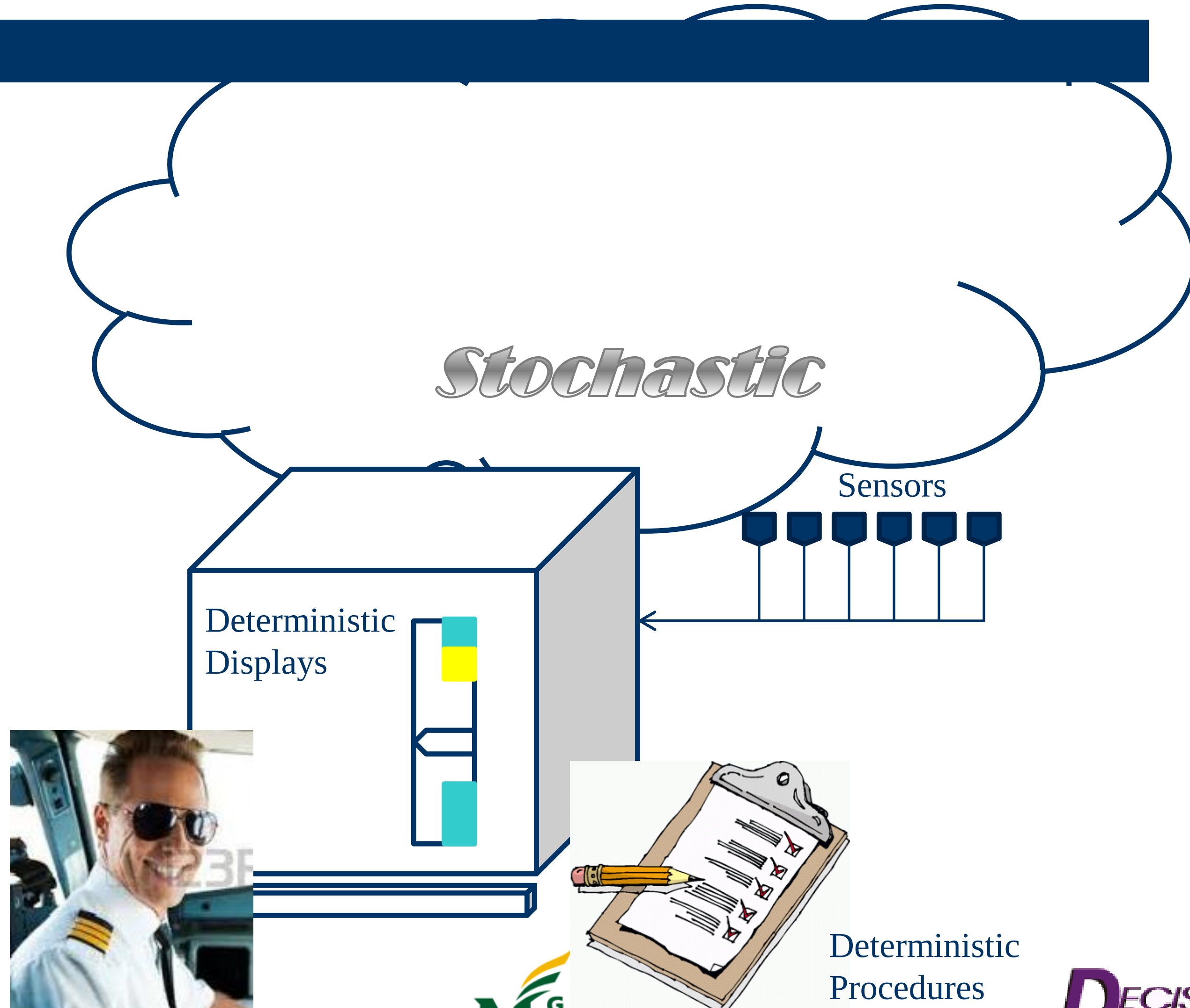
Dr. Lance Sherry,

- George Mason University Center for Air Transportation Systems Research

Outline

- Nowcasting Uncertain Events
  - Using Big Data and Machine Learning
- Risk Assessment
- Procedure Design

World View



Stochastic
Sensors
Deterministic Displays
Deterministic Procedures

Command & Control

1. Outside world is stochastic; captured by analog sensors
2. Sensor data is converted to deterministic representations
3. Operators match deterministic procedures with deterministic displays to execute the mission

Very Effective Approach …

- Stochastic world view → Deterministic procedure view is very, very successful

- In past 30 years of digital automation, the process has improved ***even*** as layers of complex procedures/functions have been added.
  - Simple → Complicated

…Except for Rare Events

- … erratic airspeed sensors (AF 447)
  - Automation to Operator: "your airplane, sir"
- … frozen A-o-A sensors (XL Germany)
  - Automation to Operator: "everything is fine"
- … failed Radio Altimeter coupled to Master Autothrottle (TK 1951)
  - Automation to Operator: "there are discrepancies, but we are going to ignore them."

Issues:

- Sensors and Sensor Checking logic cannot make sense of the stochastic data
  - Tolerances/Thresholds
  - Erratic data
  - Noise

NASA

GEORGE MASON UNIVERSITY

DECISION RESEARCH

… complex Dynamically Emerging Procedures

- Stable Approach:
  - On glide-path
  - On runway center-line
  - +/-10 knots of desired speed
  - Not excessive Rate-of-Descent (ROD)

Issues:
- Set-up is critical to downpath events/state
- What is tolerance?
- Abort is a complex decision (with significant workload consequences)

Simple/Complicated → Complex

- Snowden (2005)
- Simple/Obvious
  - Tightly constrained/No degrees of freedom (Linear System)
  - Procedures: Sense→ *Categorize*→ Respond
    - Decisions based on
  - Best Practice Procedure (i.e. optimized)
- Complicated
  - Governed by constraints/Tightly Coupled (Linear System with high combinatorics)
  - Procedures: Sense→ *Analyze*→ Respond
    - Decisions based on ruled
  - Good practice (i.e. not always best, but sufficient)

- Complex
  - Not governed by constraints/Tightly Coupled (non-linear, some uncertainty)
  - Procedures: *Probe*→ *Sense*→ *Analyze*→ Respond
    - Trial-and-error
  - Emergent/Novel practice

Transition
- **Instantaneous**
- **Silent**
- **Significant Consequences**

Challenge

- Provide operators solutions to execute missions in environment in which Stochastic world instantaneously is no longer compatible with deterministic procedures:
  - Rare events
  - Complex dynamic emergent procedures

Challenge

- How to deal with the residual unpredictability
- Our work in flightdeck automation, big data analysis, risk assessment, and procedure development is aimed at providing workable solutions

Nowcasting Uncertian Events Using Big Data and Machine Learning

Dr. Lance Sherry,

- George Mason University Center for Air Transportation Systems Research

Dr. Robert Mauro,

- Decision Research &
- University of Oregon

Good News

- Airline operations are increasingly captured by massive amounts of data:
  - Trajectory
  - Atmospheric
  - Vehicle system states
  - ATC/ATM/AOC system states
  - Operator Performance

Bad News

- Massive amounts of data
- Data sets not connected
- Data is noisy/variation
- Data from one operation (e.g. ILS approach) is not applicable same operations at another location
  - ILS approach at ABC is not the same as the ILS approach at XYZ

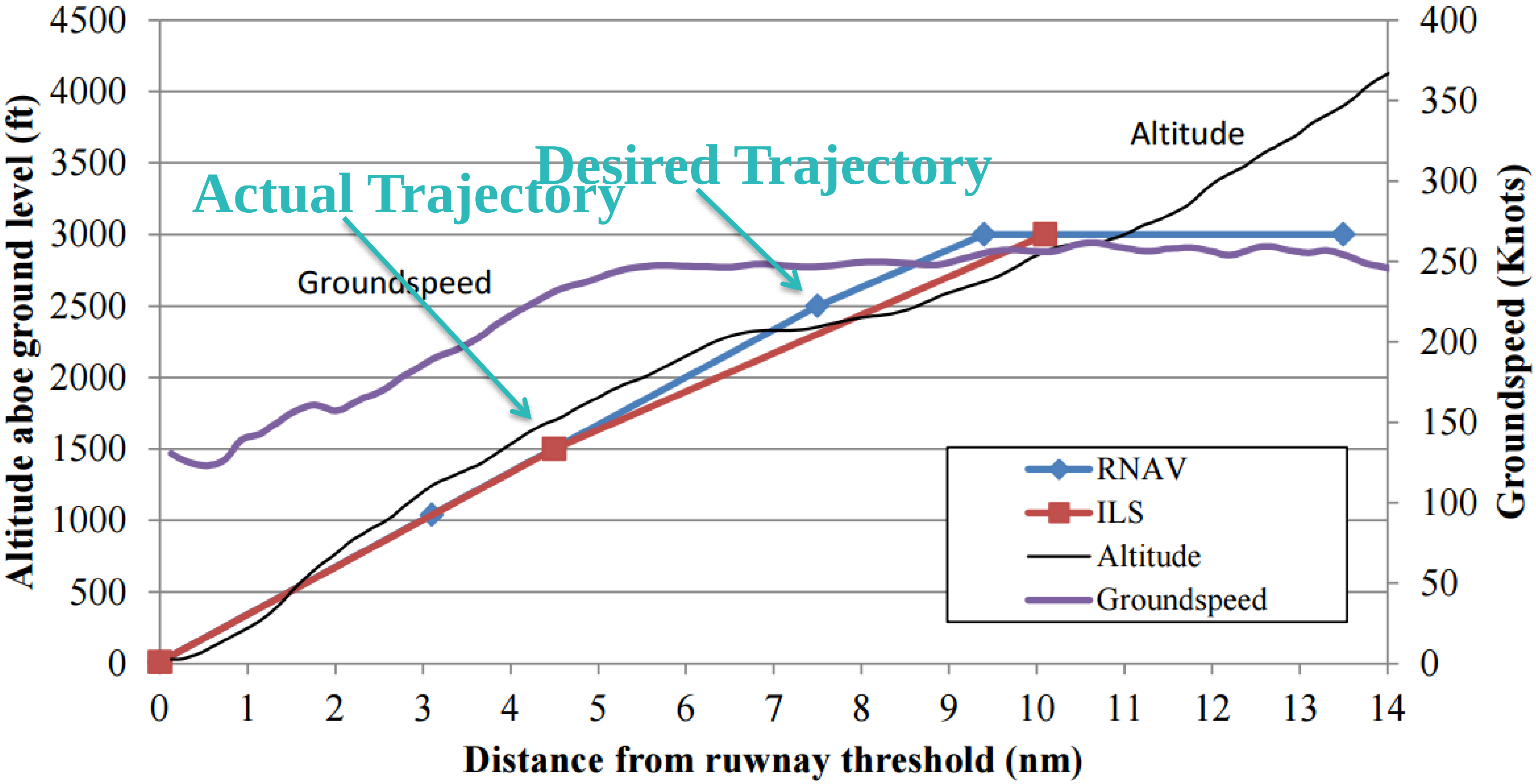
Magic Genie

- Machine Learning & Data Storage
  - Process massive amounts of data (from same location)
  - Tease out correlations and patterns

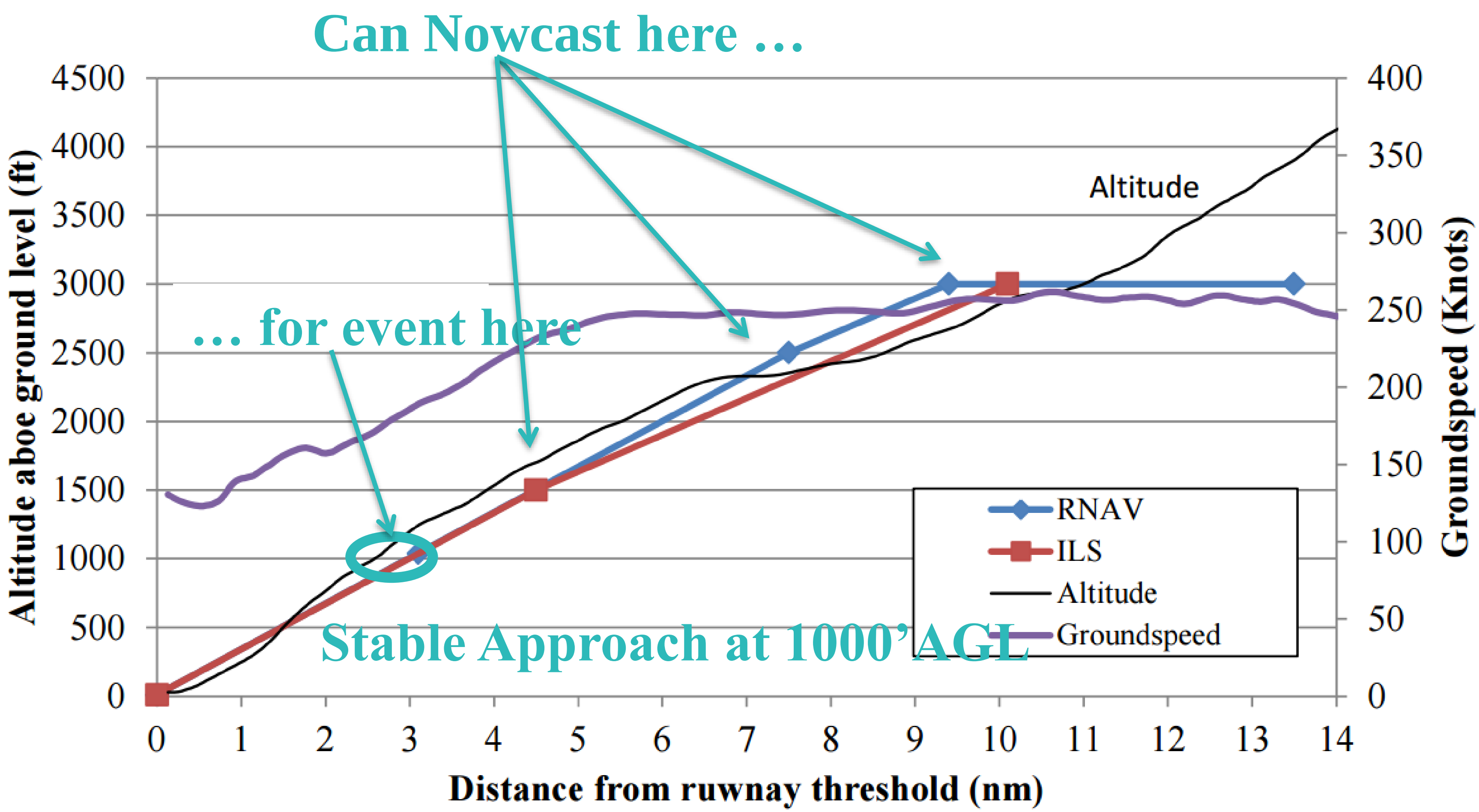
Nowcasting

- Use data from a specific procedure to *nowcast* down path performance
- *Nowcast* – forecast based on real-time data for events in the near future (minutes)

Example – Stable Approach



Example – Stable Approach

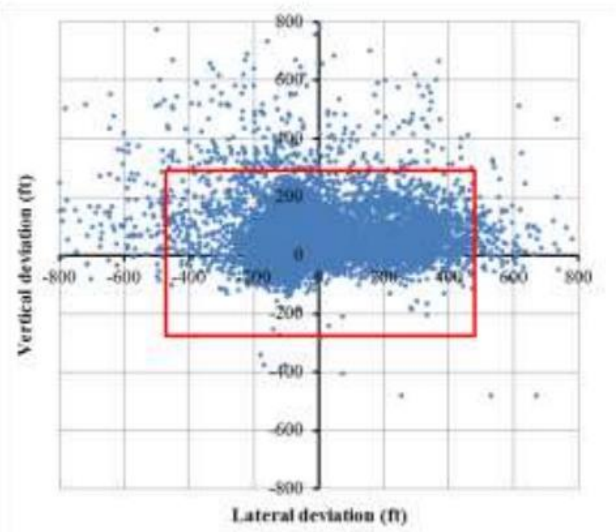


Data

- 28 days of “radar” surveillance track data
- 8237 flights
- Runway 22R EWR
  - 4-5 second update rate
  - track index
  - aircraft type
  - destination airport
  - seconds past midnight
  - latitude/longitude
  - Altitude

  \+ weather data

  \+ nav procedure data



Lateral/Vertical Deviation at 6 nm to Runway Threshold

- Results improved by use of FOQA/FDR data
  - 1 sec update rate
  - Aircraft configuration (slats, flaps)
  - Automation targets and modes
  - Aircraft maintenance log

NASA

GEORGE MASON UNIVERSITY

DECISION RESEARCH

Logistic Regression Model

Logistic Regression:

$$h_\theta(x) = \frac{1}{1+e^{-\theta x'}}$$

x = column vector containing all the feature values
θ(x) = row vector containing all the regression coefficients
hθ(x) = predicted probability that a flight with feature-vector x experiences an unstable approach after reaching 1000' AGL.

Cost Function:

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log h_\theta\left(x^{(i)}\right) + \left(1-y^{(i)}\right)\log\left(1-h_\theta\left(x^{(i)}\right)\right)\right]$$

$$\theta_j := \theta_j - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)}$$

Nowcast at 6nm/2000' AGL for Nowcast 1000' AGL

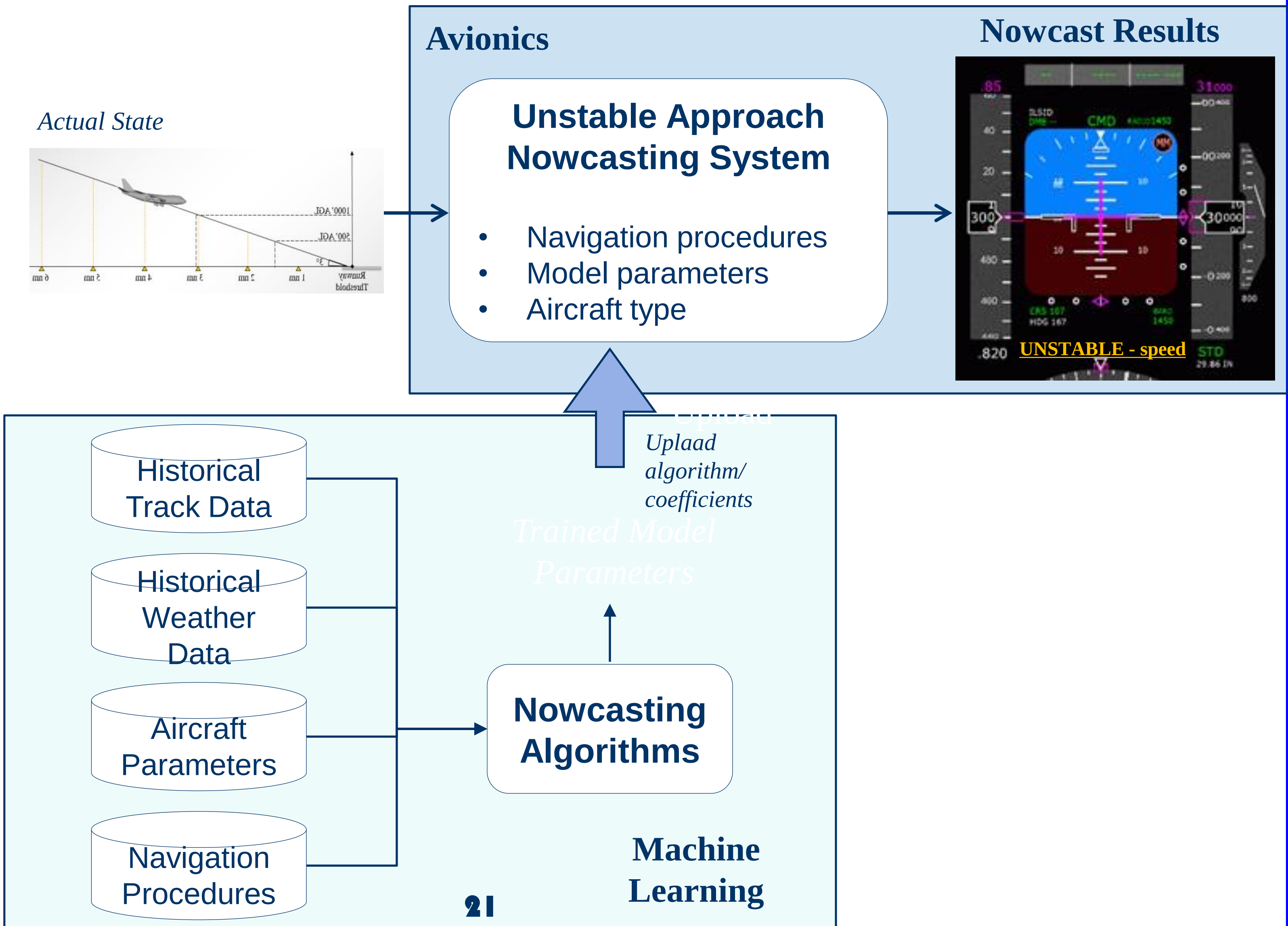
Recall 65%
% Predicated Unstable of Total Actually Unstable

Precision 77%
% Actually Unstable from Total Predicted Unstable

	Actual Unstable	Actual Stable
Predicted Unstable	962	281
Predicted Stable	518	1397

Accuracy 74%

Concept of Operations: Wise Associate



Nowcast Interpretation

- "Hey we got a strong crosswind today.

  When ATC vectors to localizer intercept late and there is this crosswind (tailwind on the base leg), flights tend to overshoot the runway centerline and have to fight to (over) correct"
  - So, anticipate to account for the tailwind in the intercept course
- "in peak arrival push, Small category aircraft have to keep their speed up in the initial approach and have to bleed-off speed in short distance while descending"
  - So, be prepared to add drag rapidly to avoid overspeed

Proposal: Wise Associate

- Automation that acts like a “back seat driver”
  - Always warning about potentially dangerous events
    - Events may not occur
    - Warnings may not always be accurate (< 25%)

- Benefits:
  - Allows operator to provide some attention to factors/events that have historically
- Applications:
  - Flight deck operations
  - Dispatch
  - Traffic Flow Management
  - Air Traffic Control

Research Questions

- Will human operators accept automation warnings?
- How accurate does it need to be to avoid "nuisance" alerts and be turned-off?
- What kind of display/aural alert would work best?

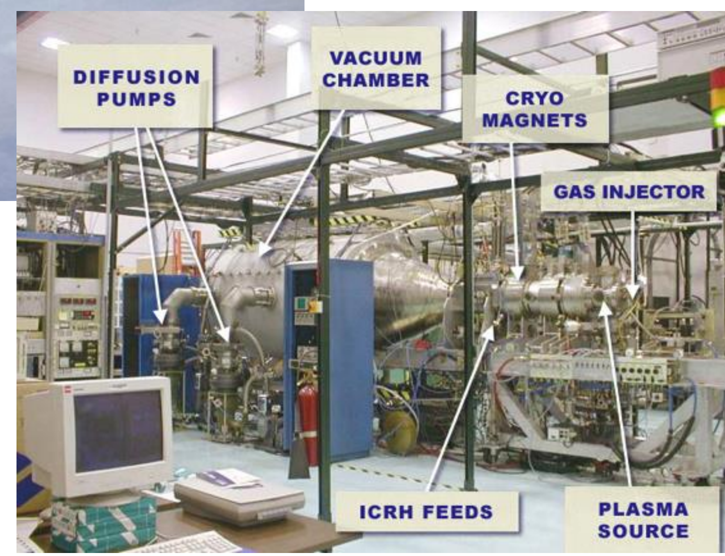
End Part 1

Risk Assessment

Dr. Robert Mauro
Decision Research &
Institute of Cognitive & Decision Sciences
University of Oregon

Dr. Immanuel Barshi
NASA Ames Research Center

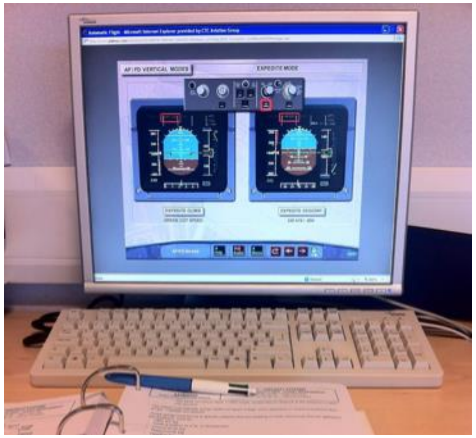
NASA Engineering Safety Center



GEORGE MASON UNIVERSITY

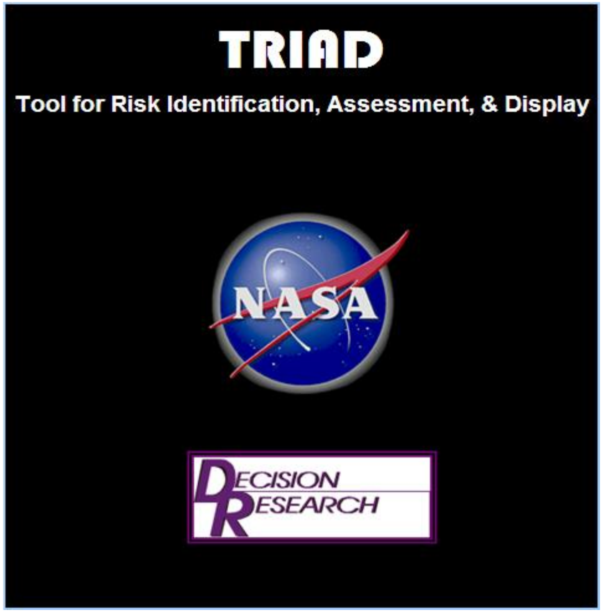
DECISION RESEARCH

Aviation Operations



Outline

- Topics
  - Thinking about Risk
  - Problems and Issues in Risk Assessment
  - Tool for Risk Identification Assessment & Display (TRIAD)



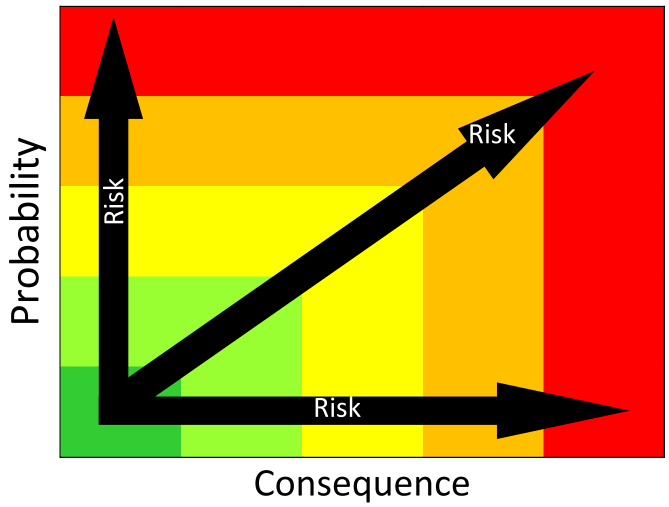
What is risk?

- Which poses the greatest risk?
  - Meteor strike on the Ops Center
  - W&B program error causes 2% decrease in fuel efficiency
  - All aircraft in a fleet grounded
  - Loss of an aircraft and crew

Dayton, Ohio

What is risk?

- Combination of
  - probability (likelihood)
  - consequences (threat value)
- Risk=f(p,c)

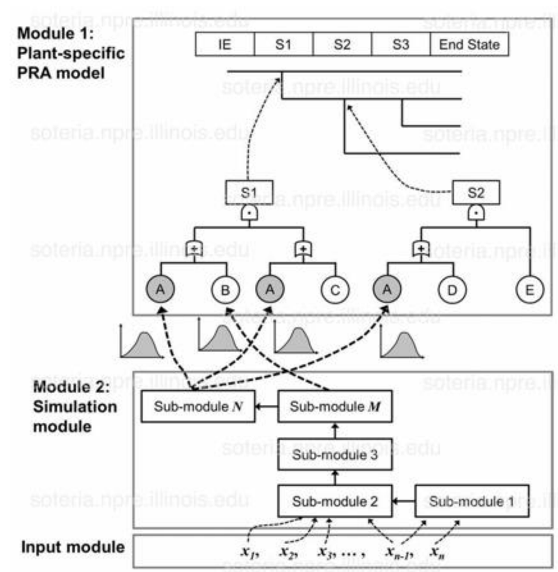


Common Methods of Risk Assessment

- Informal “seat-of-the-pants” approach
- Probabilistic Risk Analysis (PRA)

Probabilistic Risk Analysis

- Typical steps
  - Decomposition of paths leading to undesired state (e.g. FMEA)
  - Evaluation of probability functions associated with each path
  - Determination of costs associated with each possible undesired state
  - Integrated risk function produces risk value
- Problems
  - Complex process
    - Difficult to understand and communicate, so people avoid it
    - Conducted by decision analysts using input from domain experts
  - Costly and time consuming
    - So reserved for occasional "big" issues
  - Accuracy depends on values that are difficult to estimate and identification of paths that are difficult to explicate
    - Calculations may give misleading impression of precision
  - Difficult to obtain consistent cost metrics
    - Different types of consequences are hard to equate
  - Simplified versions don't fit all situations



Informal risk assessment

- Process
  - Managers consider problem and produce global judgment based on past experience and logical analysis
- Problems
  - Subject to numerous biases
    - Availability – whatever comes to mind
    - Representativeness – whatever fits expectations
  - Process is obscure
    - Leads to argument not discussion
    - Difficult to document
    - Difficult to improve

GEORGE MASON UNIVERSITY

DECISION RESEARCH

The Risk Matrix

- Semantics matter
  - Why use particular labels?
  - People often disagree on what the labels mean.
- Categories imply precision
  - E.g., what if an issue can range from “unlikely” to “possible”?
- Colors imply decisions
  - Who made the decision rules?
- Why 5 X 5 matrix?
  - Categories imply that one should treat everything in a “box” the same.

Likelihood	Consequences				
	Insignificant	Minor	Moderate	Major	Severe
Almost certain	M	H	H	E	E
Likely	M	M	H	H	E
Possible	L	M	M	H	E
Unlikely	L	M	M	M	H
Rare	L	L	M	M	H

Probability	Negligible	Marginal	Significant	Critical	Catastrophic	
5	R	U	I	I	I	Frequent
4	R	R	U	I	I	Probable
3	A	R	U	U	I	Occasional
2	A	A	R	U	U	Remote
1	A	A	A	R	R	Improbable
Severity	1	2	3	4	5	

The Risk Matrix

- Risk scores (numbers in the category boxes) imply “real” values (and ratio level measurement)
- This implies symmetry
  - E.g., all boxes with the same value are the same.
- But the scales are (at best) ordinal.
- This can result in reversals in which riskier combinations receive lower scores.

RISK ASSESSMENT SCORING MATRIX

LIKELIHOOD										
Certain	10	20	30	40	50	60	70	80	90	100
Almost certain	9	18	27	36	45	54	63	72	81	90
Very likely	8	16	24	32	40	48	56	64	72	80
Probable	7	14	21	28	35	42	49	56	63	70
Likely	6	12	18	24	30	36	42	48	54	60
Likely	5	10	15	20	25	30	35	40	45	50
May happen	4	8	12	16	20	24	28	32	36	40
Improbable	3	6	9	12	15	18	21	24	27	30
Unlikely	2	4	6	8	10	12	14	16	18	20
Very unlikely	1	2	3	4	5	6	7	8	9	10
	Insignificant injury	Minor injury	Minor injury	Illness - Injury	Illness - Injury	Major Injury	Major Injury	Single fatality	Fatality	Multiple Fatalities
KEY										SEVERITY

Not Significant		0 to 3	May be ignored, No further action Required
Very Low		4 to 12	
Low		13 to 25	Ensure safe working
Moderate		26 to 42	Refer to Risk Assessment, Safe Working Procedures
High		43 to 67	Monitor Control Measures
Very High		68 to 100	Avoid if Possible, Full Method Statement if Not

Risk Assessment Wish List

- Structure discussion
- Document the process followed
- Easy path to improve assessment when desired
- Relatively easy to understand and communicate
  - So it will be used
- Relatively inexpensive and time effective
  - So it will be used whenever needed
- Do not want to rely on misleading calculations
- Reveal (not hide) uncertainty
- Obtain consistent consequence metrics
- Valid in all situations

Middle way

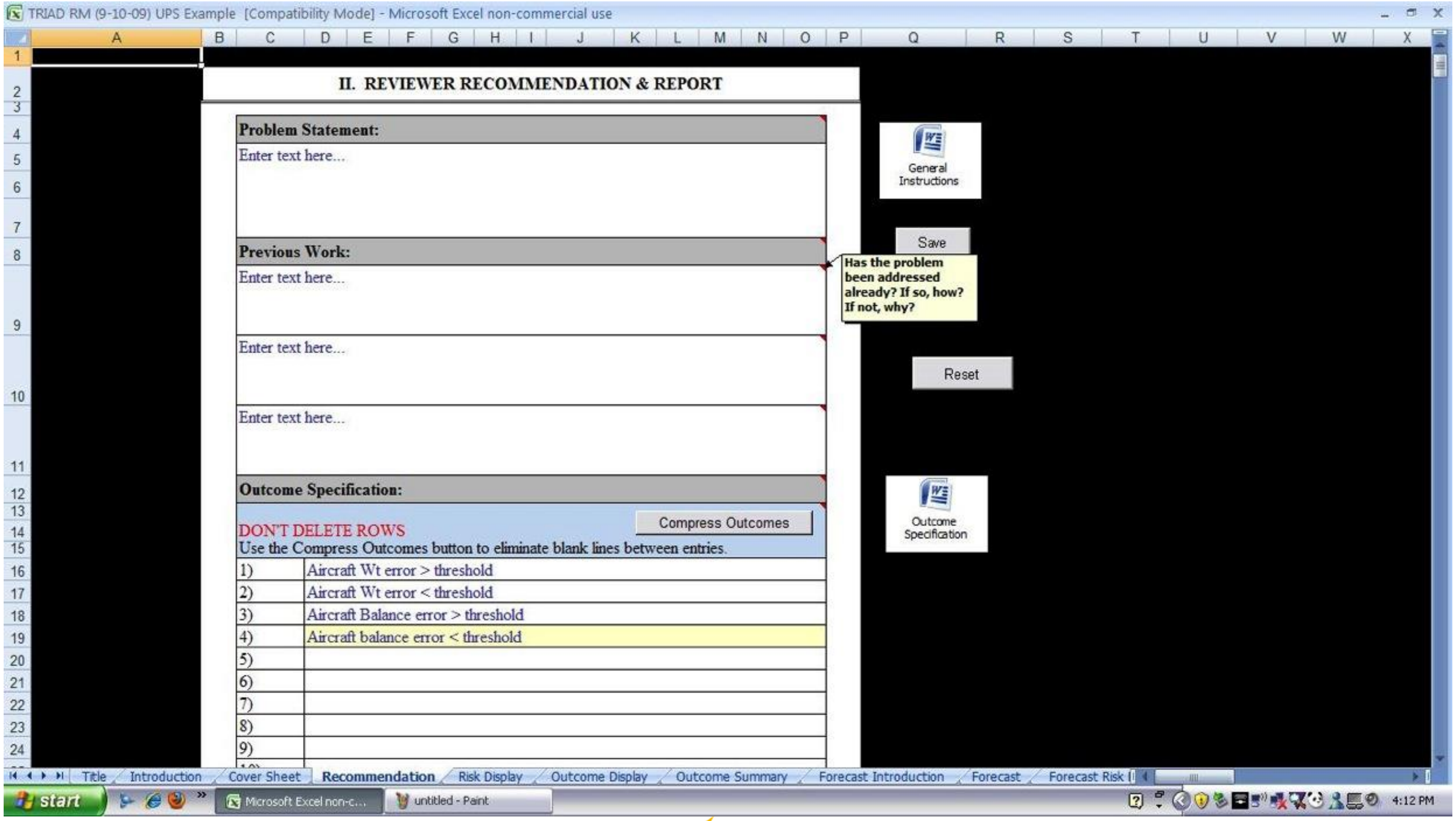
- Goals
  - Provide decision maker with information needed to make sensible risk assessments and decisions
  - Make bases of assessments and decisions explicit
  - Allow assessments to be made with different levels of precision and effort
  - Provide for different levels of analysis
- Systematic approach
  - Requires well-defined problems
  - Components of risk clearly specified
  - Different types of consequences handled separately
  - Uncertainties displayed
  - Risk functions not imposed

General Approach to Risk Assessment

- Current Risk Assessment
  - Define problem
  - Specify possible outcomes
  - Assess probabilities
  - Assess consequences
  - Display risk summaries
- Forecast Risk Assessment
  - Specify possible interventions
  - Assess expected effects of interventions
  - Display risk reduction summaries

Developing Solutions - TRIAD

1. Specify Problem & Outcomes



2. Likelihood Estimation

LIKELIHOOD ESTIMATION

Instructions: For each outcome, enter your best estimate of the likelihood. Specify minimum and maximum values. Values may be entered directly in the columns labelled "Min", "Best" and "Max". Likelihood categories (right) may be used to simplify the process: click on a cateogry then hit a button (Min, Max, Best, or All) to enter the values. The values must lie between 0 and 1 and "Min" ≤ "Best Estimate" ≤ "Max". Hit "Return" to go back to the Recommendations page.

Very Unlikely
Unlikely
Possible
Likely
Very Likely

Clear All

Set All

Set Max

Set Best

Set Min

To translate frequency/operation to events/day, enter operations/day here:

400

Return

Outcome		Min	Best	Max	Events/Day Translation: Min	Events/Day Translation: Best	Events/Day Translation: Max
1)	Aircraft Wt error > threshold	1/1,000,000	1/100,000	1/50,000	once every 2500 days	once every 250 days	once every 125 days
2)	Aircraft Wt error < threshold	1/4000	1/2000	1/400	once every 10 days	once every 5 days	1 times/day
3)	Aircraft Balance error > threshold	0	5/10,000,000	1/1,000,000	once every 0 days	once every 5000 days	once every 2500 days
4)	Aircraft balance error < threshold	1/1,000,000	5/10,000	1/1,000	once every 2500 days	once every 5 days	once every 3 days
5)							
6)							
7)							
8)							
9)							

3. Consequence Evaluation

LIFE & HEALTH EVALUATION

This scale measures the threat to the lives and health of humans that may occur as a direct result of the problem. The effects may be immediate or delayed.

Instructions: Select an outcome and then rate it on the evaluation scale. Choose the consequence level that best reflects the harm that would occur in the worst reasonable case of that outcome. Once you have rated an outcome, the rating will appear in the column on the right. Continue to select outcomes and rate accordingly until all listed outcomes have been rated.

Life & Health - Consequence Evaluation Scale

None	Very Low	Low	Moderate	High	Very High
None	Minimal Effects	Minor Injury	Major Injury	Single Death	Multiple Deaths
○	◉	○	○	○	○

1)	Aircraft Wt error > threshold	**Very Low**
2)	Aircraft Wt error < threshold	**None**
3)	Aircraft Balance error > threshold	**Very Low**
4)	Aircraft balance error < threshold	**None**
5)		

3. Consequence Evaluation

PROPERTY DAMAGE EVALUATION

This scale measures the financial costs that would be associated with each outcome. This includes the costs of repairing or rebuilding company equipment and vehicles, the costs of repairing or rebuilding structures belonging to other entities, and the cost of recovering from environmental damage, compensating individuals for related losses, and relocating people as a result of that damage.

Instructions: Select an outcome and then rate it on the evaluation scale. Choose the financial cost value that would be associated with the worst reasonable consequence of the outcome. Once you have selected a category, the rating will appear to the right of the outcome. Continue to select outcomes and rate accordingly until all listed outcomes have been rated.

Property Damage - Consequence Evaluation Scale

None	Very Low	Low	Moderate	High	Very High
None	\$0 to \$1 Million	\$1 Million to \$10 Million	\$10 Million to \$100 Million	\$100 Million to \$250 Million	More than \$250 Million

1)	Aircraft Wt error > threshold	Low
2)	Aircraft Wt error < threshold	Very Low
3)	Aircraft Balance error > threshold	Moderate
4)	Aircraft balance error < threshold	Very Low
5)		

3. Consequence Evaluation

MISSION SUCCESS EVALUATION

This scale measures the threat to the compny's mission posed by the problem under consideration. Direct financial costs for repairing or replacing equipment are included in the Property Damage scale. This scale seeks to measure the harm to the company that would result from mission objectives not being met. This includes direct costs of interruptions in service (e.g., lost baggage, delays, cancellations) caused by a problem and indirect costs caused by regulatory actions.

Instructions: Select an outcome and then rate it on the evaluation scale. Select the mission success consequence that would be associated with the worst reasonable case associated with the outcome. Once you have selected a category, the rating will appear to the right of the outcome. Continue to select outcomes and rate accordingly until all listed outcomes have been rated.

Mission Success - Consequence Evaluation Scale

None	Very Low	Low	Moderate	High	Very High
None	Disruption Single/Aircraft Flight	Disruption Multiple Aircraft/Flights	Disruption Station/Many Operations	Disruption Multiple Stations	Company Threatened
○	○	○	○	○	○

1)	Aircraft Wt error > threshold	**None**
2)	Aircraft Wt error < threshold	**Very Low**
3)	Aircraft Balance error > threshold	
4)	Aircraft balance error < threshold	

3. Consequence Evaluation

SOCIAL AMPLIFICATION EVALUATION

The damage caused by people's reaction to a failure can far outweigh the damage caused by the failure itself. This "social amplification" effect is a complex function of the physical consequences of the event, public perceptions, media effects, and political activity. The scale below is designed to estimate the size of these effects.

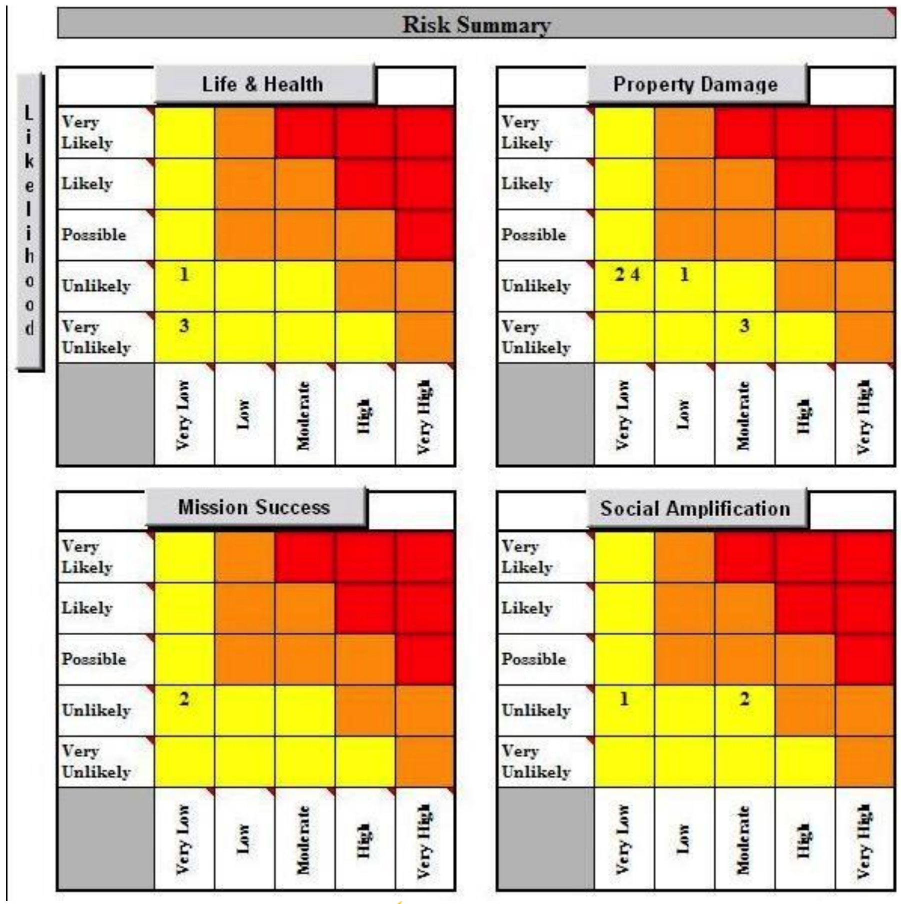
Instructions: Select an outcome and then rate it by checking all of the boxes on the evaluation scale that apply. Be sure that the outcome is highlighted. The rating will appear to the right of the outcome. Continue to select outcomes and rate accordingly until all listed outcomes have been rated.

None	Widespread Exposure?	Non-company Injuries?	Single Death?	Multiple Deaths?	Non-Company Damage?	Unavoidable Risk?	Unobservable Hazard?	Effects Delayed?	Company Failure?	Unfamiliar Technology?	Prior Media Interest?
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

How much of a social amplification effect would be expected if:

1)	Aircraft Wt error > threshold	**Very Low**
2)	Aircraft Wt error < threshold	**Moderate**
3)	Aircraft Balance error > threshold	

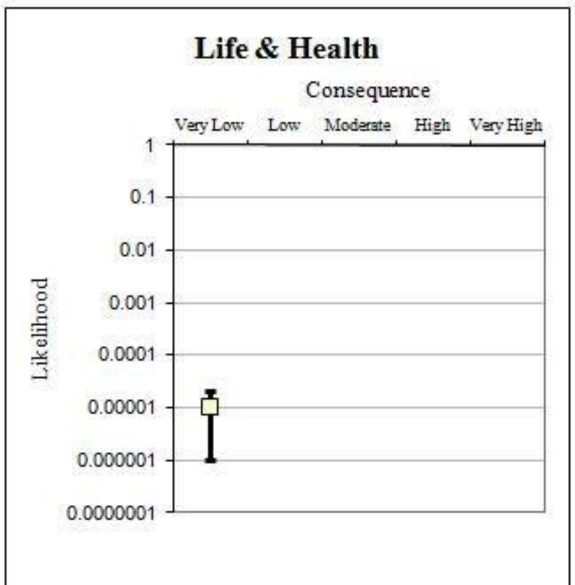
4. Risk Summary

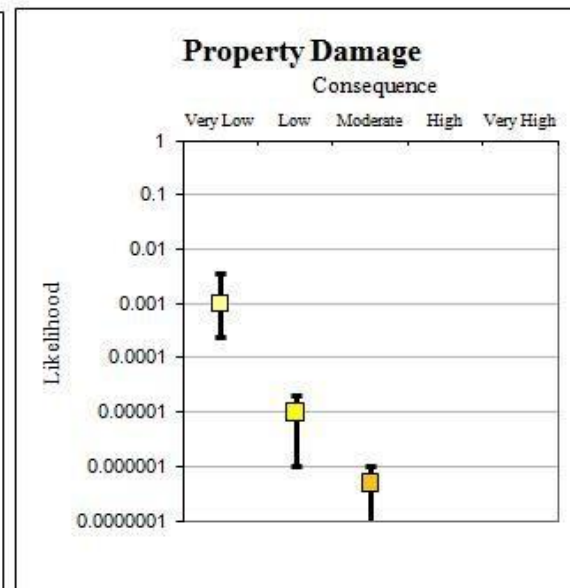


Logarithmic Risk Display

III. LOGARITHMIC RISK ASSESSMENT DISPLAY

These graphs present a visual summary of the risk profile associated with this issue. The values displayed are collapsed across outcomes but displayed separately for each level of each type of consequence. The error bars display the minimum and maximum probabilities.





NASA

GEORGE MASON UNIVERSITY

DECISION RESEARCH

Outcome Summary Table

V. OUTCOME SUMMARY TABLE

This table provides you with the probability and consequence scale ratings by outcome. The best estimate of the likelihood of each outcome is presented in digital form and coded by increasingly dense hatchings from white to black. The information in this table should not be changed or manipulated. To sort the outcomes in order of increasing probability, click on "Sort by Increasing Probability". To sort the outcomes in order of decreasing probability, click on "Sort by Decreasing Probability". To sort the outcomes alphabetically by label, click on "Sort by Outcome Label".

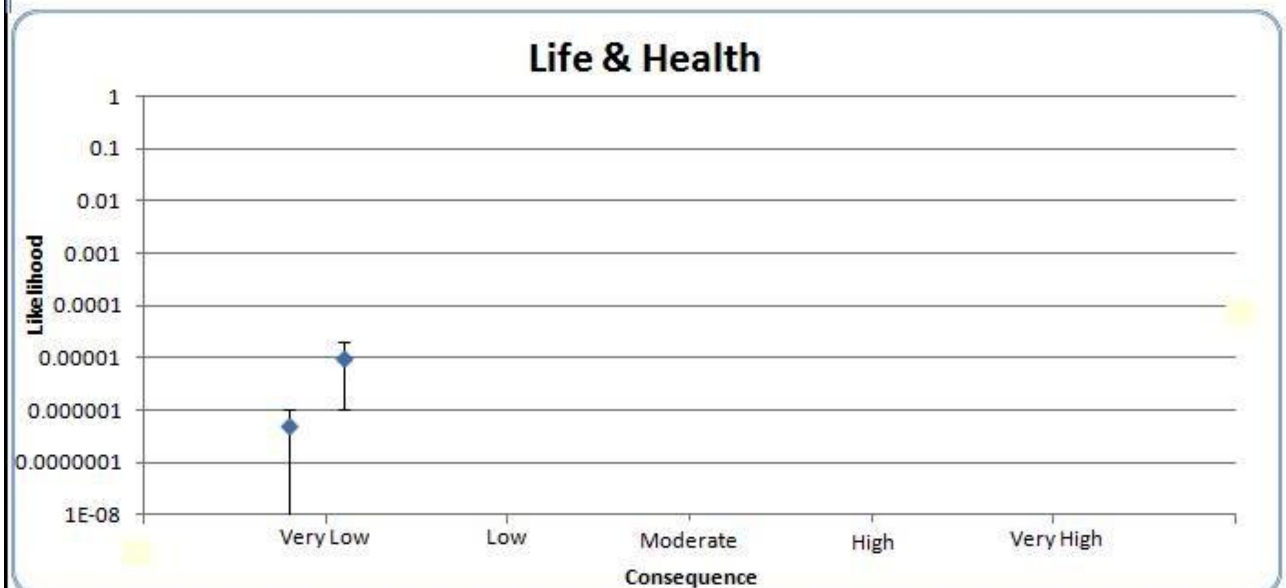
Sort by Outcome Label | Sort by Increasing Probability | Sort by Decreasing Probability

	Outcome	Probability		Life & Health	Property	Operation	Amplification
1)	Aircraft Wt error > threshold	0.00001		Very Low	Low	None	Very Low
2)	Aircraft Wt error < threshold	0.0005		None	Very Low	Very Low	Moderate
3)	Aircraft Balance error > threshold	0.0000005		Very Low	Moderate		
4)	Aircraft balance error < threshold	0.0005		None	Very Low		
5)							
6)							
7)							
8)							

Logarithmic Outcome Display

IV. LOGARITHMIC OUTCOME DISPLAY

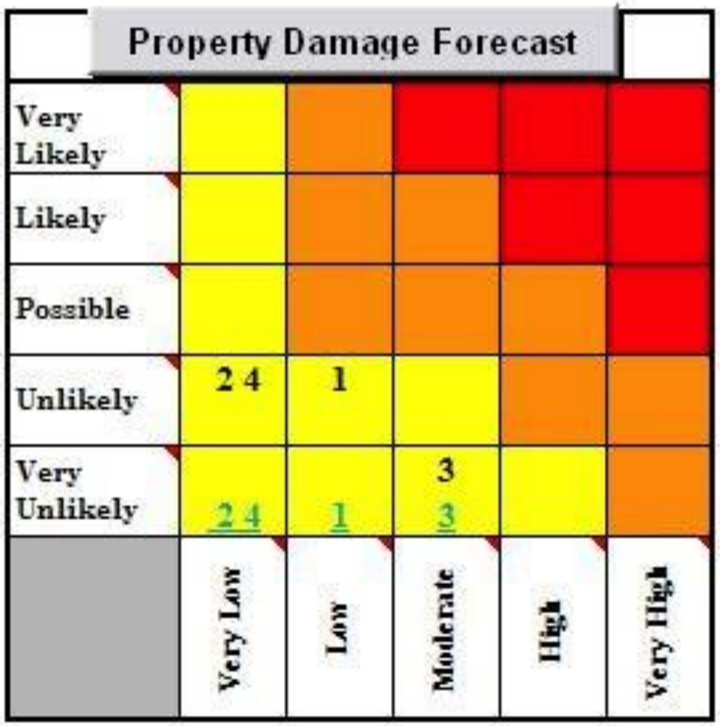
These graphs present a visual summary of the risk profile associated with this issue. The values for each outcome are displayed separately by type of consequence. The error bars display the minimum and maximum probabilities.

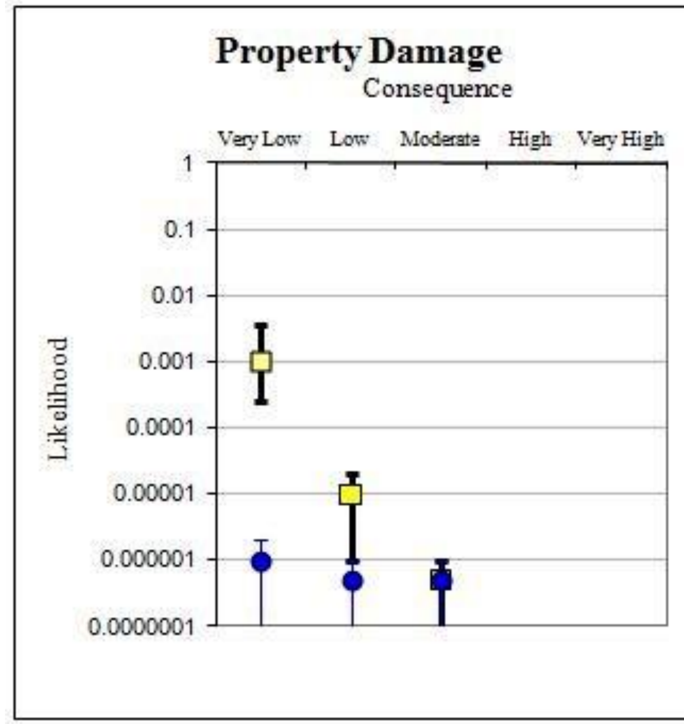


5. Forecast Risk

- Devise possible intervention(s)
- Consider previously determined outcomes
- Rate new likelihood given intervention
  - Particular intervention may/may not affect likelihood of particular outcomes
- Rate new consequence given intervention
  - Particular intervention may/may not affect consequences of particular outcomes
- Display Results

Forecast Risk Displays





Conclusion

- TRIAD
  - Easy to use tool for risk assessment under uncertainty
  - Provides information needed for decision without misleading oversimplification
  - Displays help guide group discussion and decision-making, not replace it

Conclusions

- TRIAD
  - Avoids using risk functions that do not match reality
  - Avoids use of arbitrary values to equate different types of consequences
  - Avoids logical / mathematical errors in combining assessments of different outcomes

End

End Part 2

Dr. Immanuel Barshi
NASA Ames Research Center
Immanuel.Barshi@nasa.gov

Dr. Robert Mauro
Decision Research
mauro@decisionresearch.org

Procedures that Work

Designing Excellent Procedures

Immanuel Barshi

NASA Ames Research Center

Asaf Degani

General Motors Advanced Technology Center

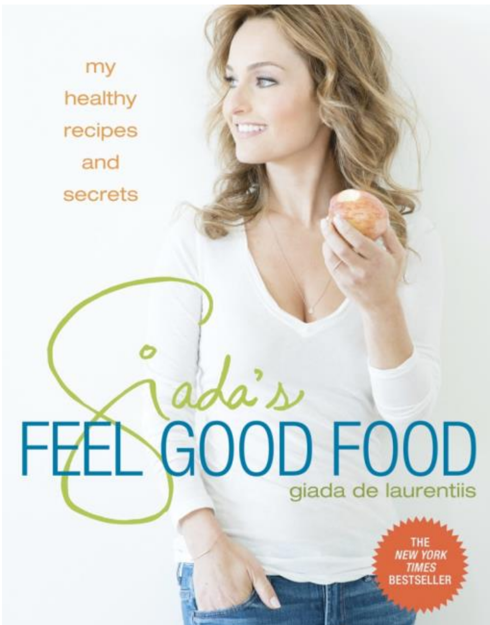
Loukia Loukopoulou

Swiss International Airlines

Robert Mauro

Decision Research & University of Oregon

Where



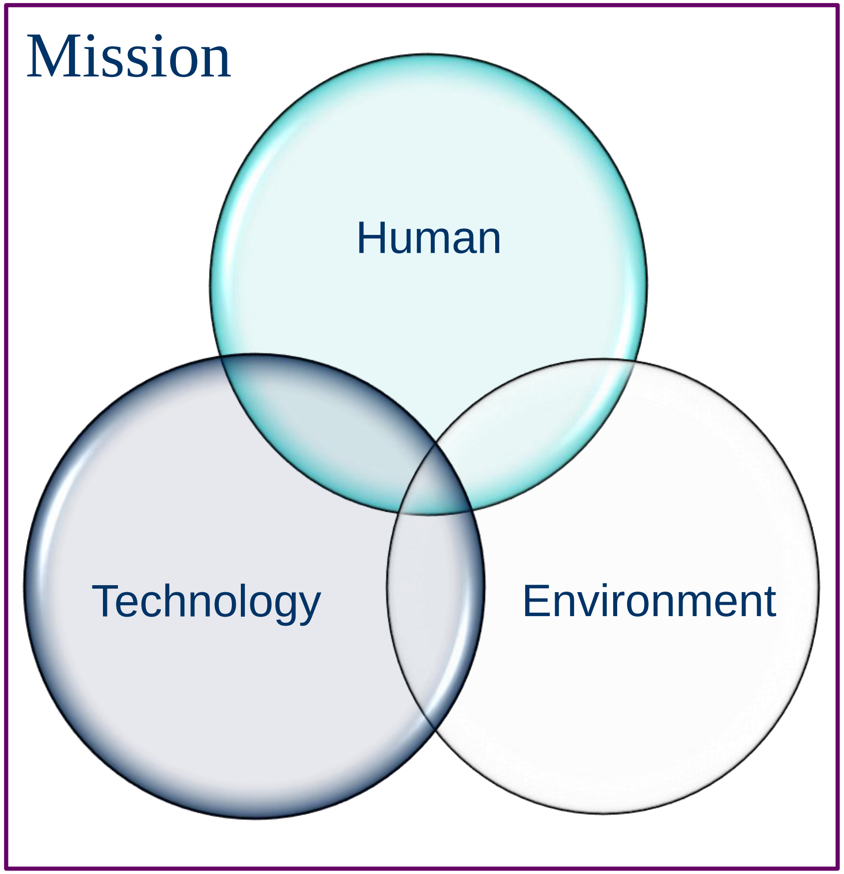
NASA GEORGE MASON UNIVERSITY DECISION RESEARCH

Why

- Operate
- Reduce error
- Substitutability
- Operational efficiency
- Organizational efficiency
- Managerial control
- Risk management

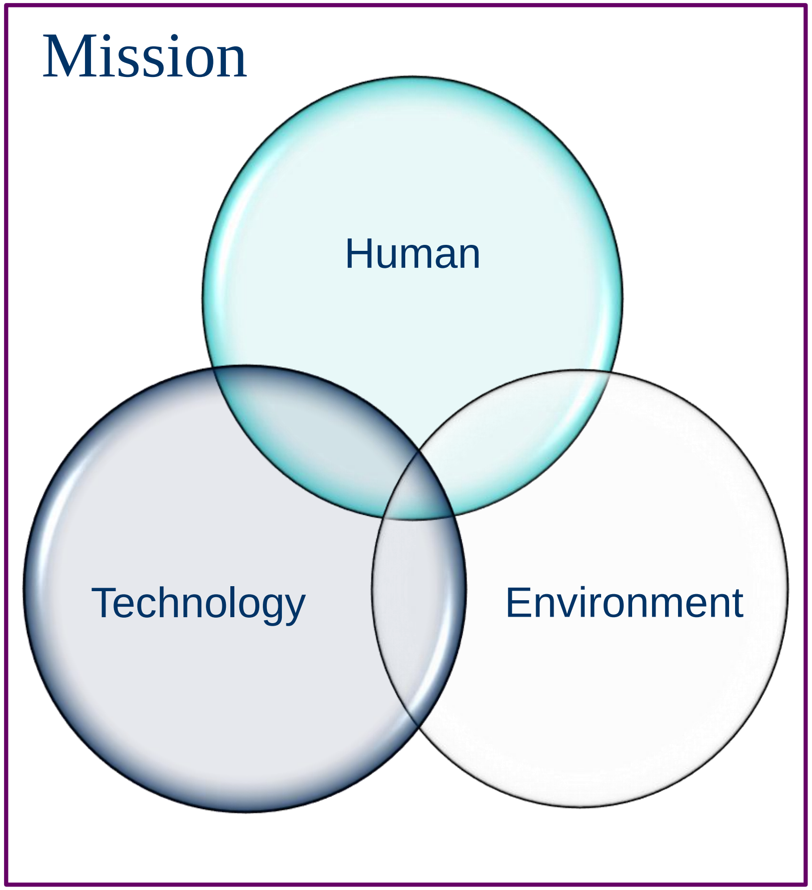
When

- Something is new
- Something has changed
- Something isn't working



Why Procedures Fail

- Bad procedures
  - Fail to take into account an important component of the task
  - Too narrow
  - Too broad
- Good procedures that went bad
  - Can't adapt (and aren't changed)
- Aren't followed



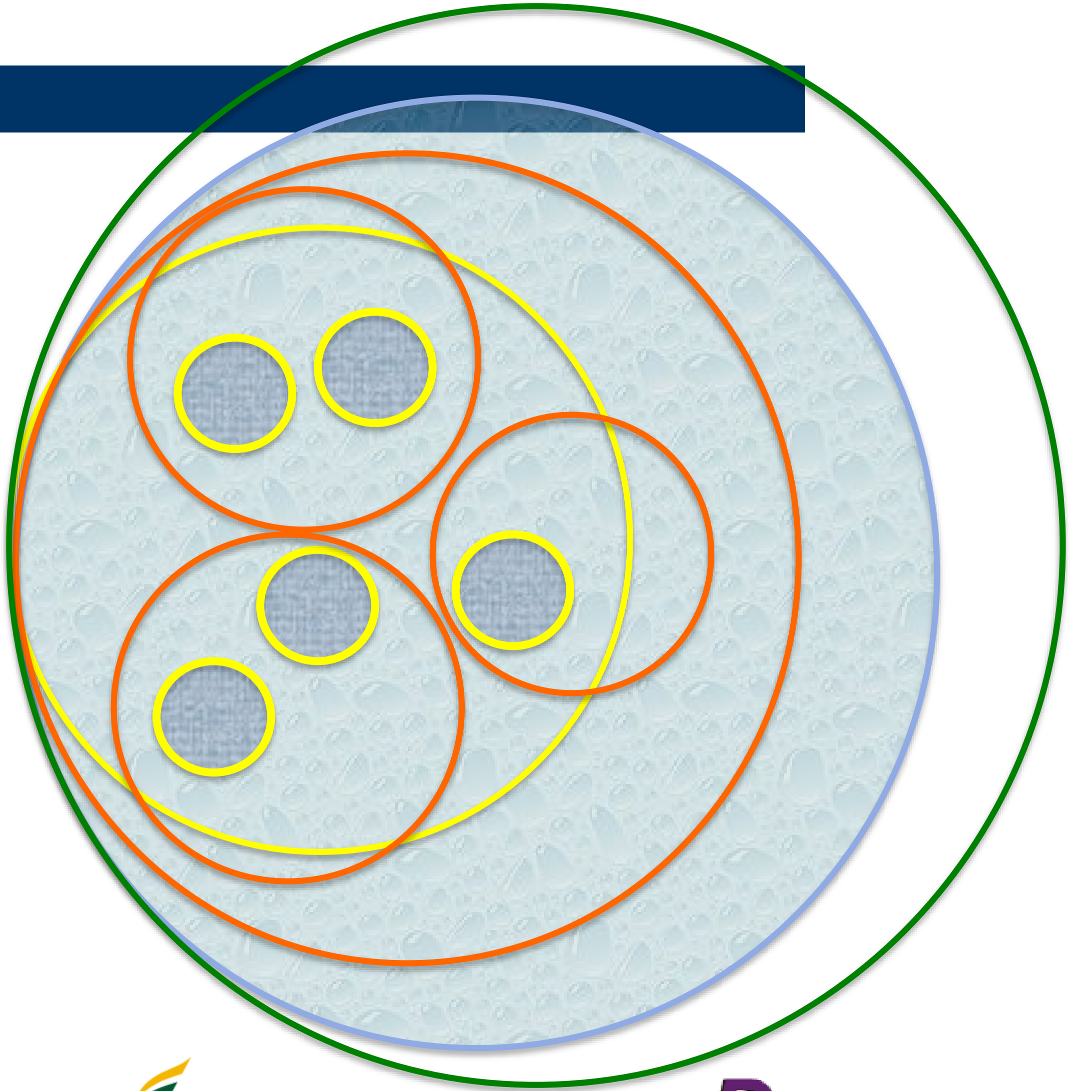
4P's

Practices

Procedures

Policies

Philosophy

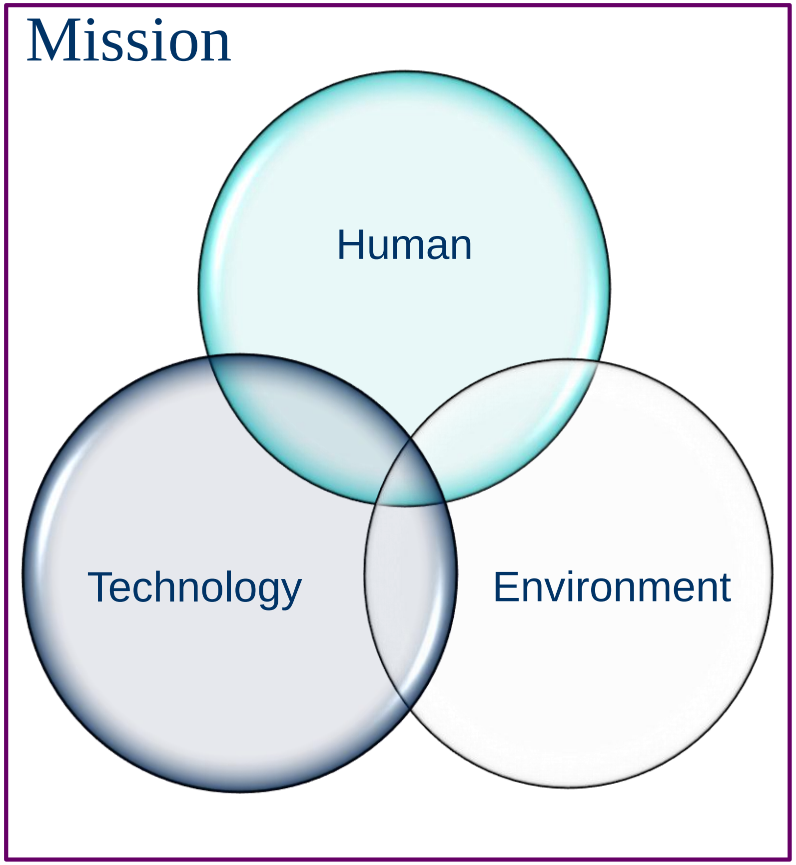


How

- Understand
  - Goals & Requirements
  - Task Analysis
- Design
- Implement
- Evaluate

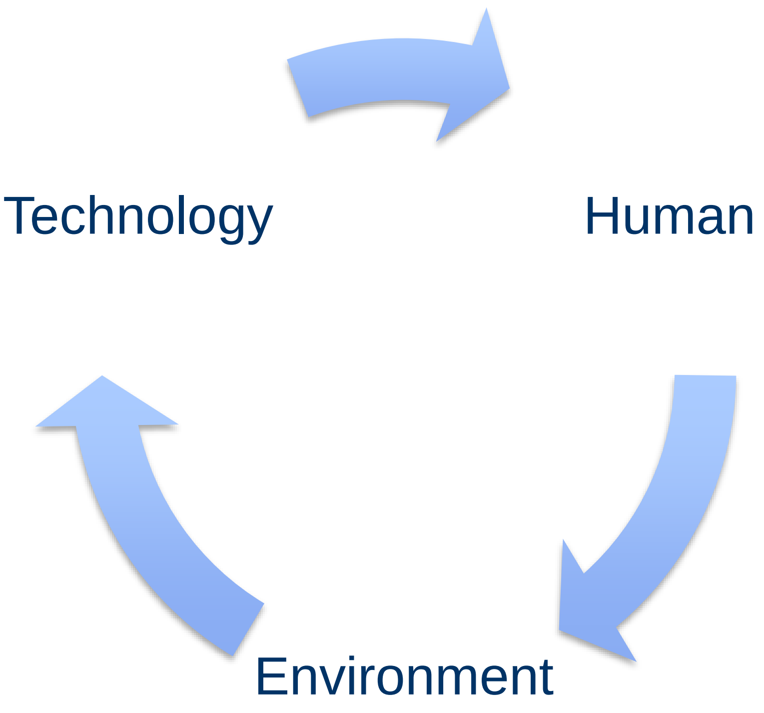
How: Understand

- Goals for procedures
- Goals of procedures
  - Correct
  - Reliable
  - Robust
  - Resilient
  - Efficient
- Task Analysis



How: Design

- Task Procedure
  - Domain
    - Technology
    - Human
    - Environment
  - Requirements
  - Conflicts
  - Prohibitions
    - Margins, Barriers, Buffers
  - Sequences
  - Timing
  - Formative Evaluations
- Phase Procedure
  - Sequential Task Procedures
  - Interwoven Task Procedures



How: Implement

- Training
  - How
  - Why
  - Why not
- Change Management

How: Evaluate

- Observations
  - General
  - Targeted
- Surveys
- Automated data collection (FOQA)
- Voluntary reports (ASAP, ASRS)

Results

Example

- Design Overview
  - Phase I: Old procedures
    - line pilots observed at random
  - Phase II: New procedure tests
    - a small group of test crews observed flying under old procedures and proposed procedures for a whole month
  - Phase III: New procedures after adoption
    - line pilots observed at random 4-5 months after adoption of new procedures
  - Phase IV: New procedures after stabilization
    - line pilots observed at random 8-10 months after adoption of new procedures (+ winter ops)

Example

- Measurement overview
  - Pilots observed using structured observation log
  - Pilots surveyed using open/closed ended questions returned directly to NASA
  - Observers' questionnaire
  - "Event" details
  - Observer training

NASA GEORGE MASON UNIVERSITY DECISION RESEARCH

Figure 1. Mean Number of Problems on Target Items per Flight

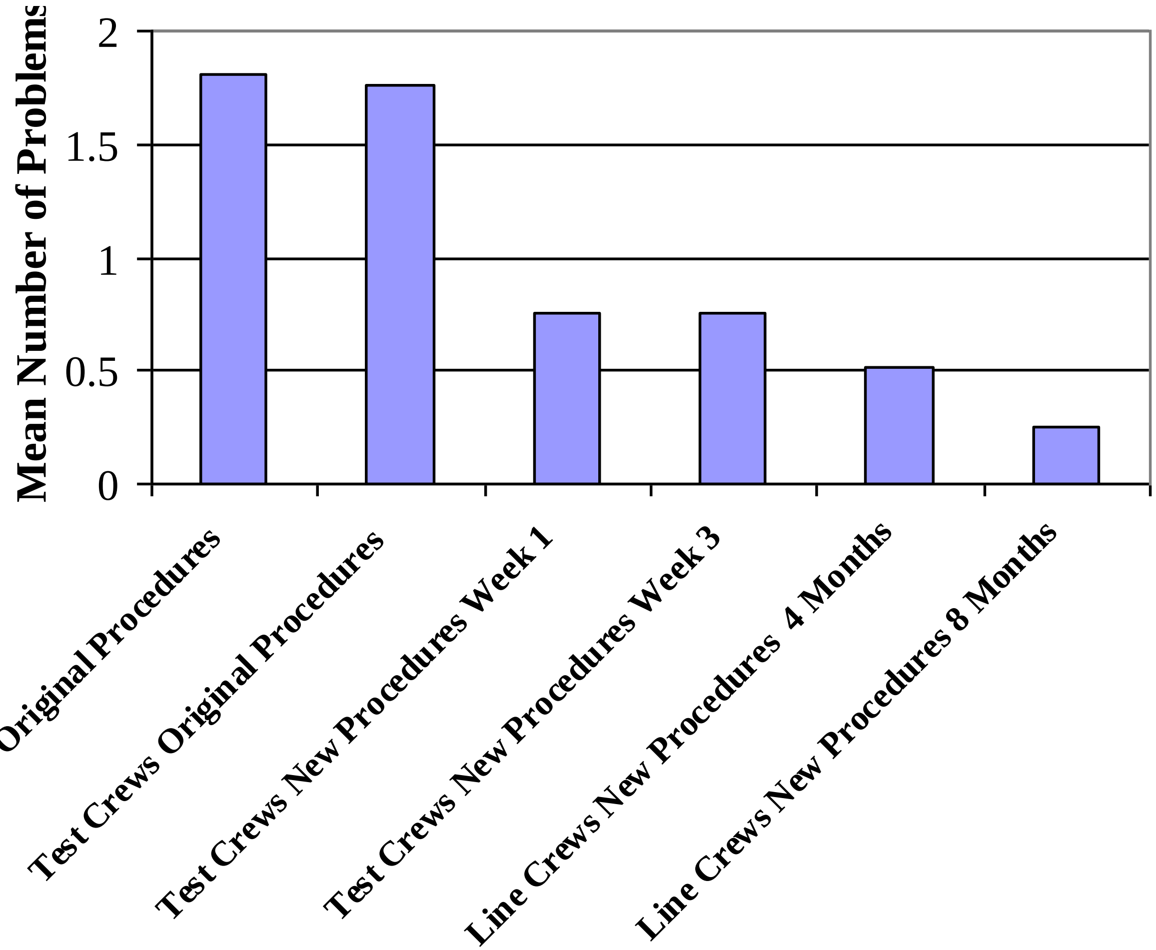
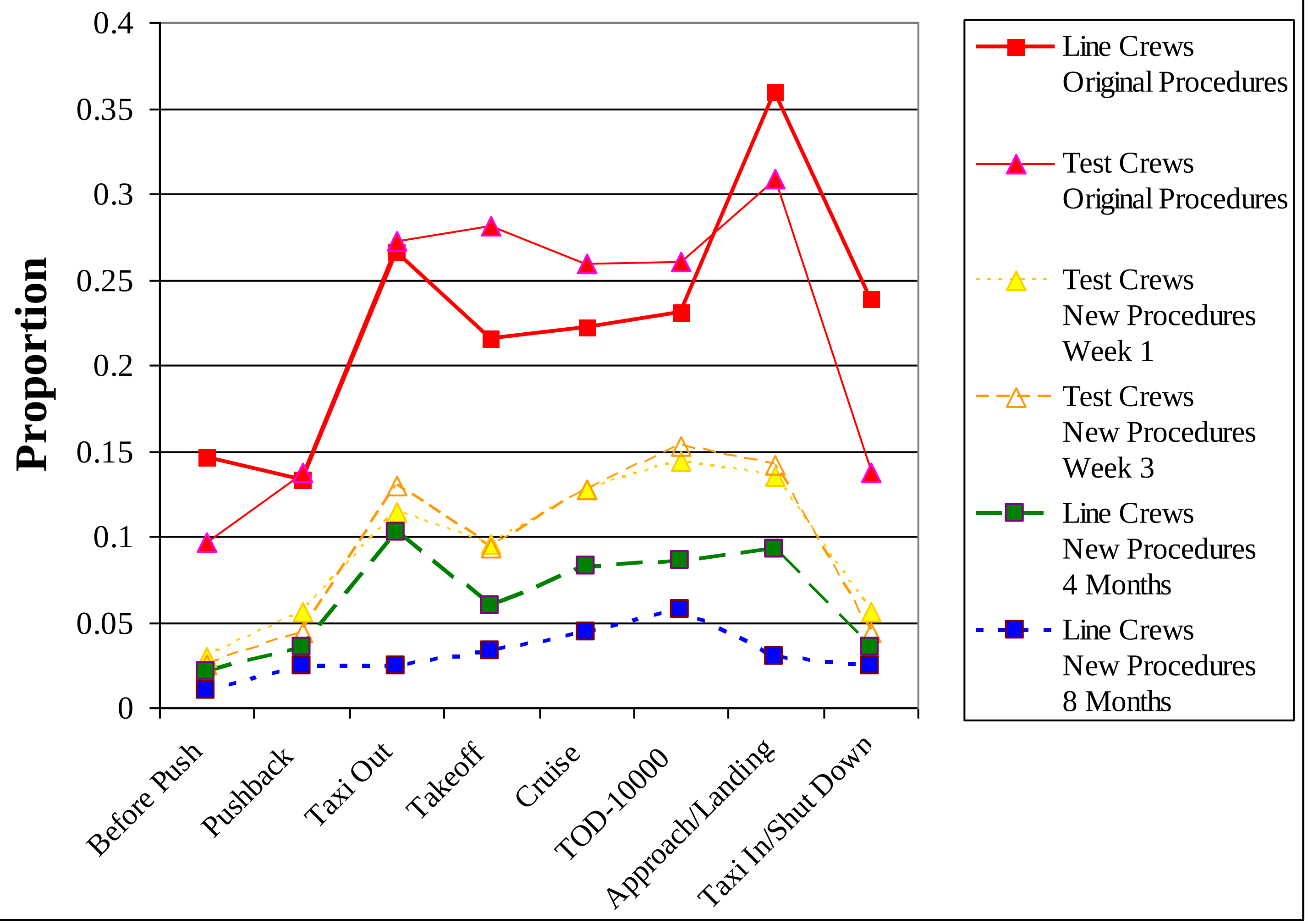


Figure 3. Problems on Target Items by Phase of Flight



Thank You!