INSTITUTE FOR DEFENSE ANALYSES

A Reliability Assurance Test Planning and Analysis Tool

John T Haman, Project Leader

Rebecca M Medlin
Emma P Mitchell
Keyla Pagán-Rivera
Dhruv K Patel

OED

September 2024

Distribution Statement A. Approved for public release: distribution is unlimited.

IDA Product ID 3003359

INSTITUTE FOR DEFENSE ANALYSES
730 East Glebe Road
Alexandria, Virginia 22305

About This Publication
This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-24-D-0020, Task no. BD-9-2299(90), “Methods Development,” and C9082, “CRP Statistics WorkGroup,” for the Office of the Director, Operational Test and Evaluation.

Acknowledgments
The IDA Technical Review Committee was chaired by Dr. Heather M. Wojton and consisted of Mr. Addison D. Adams, Dr. Curtis G. Miller, Dr. Jason P. Sheldon, Dr. John T. Haman, and Dr. Miriam E. Armstrong from the Operational Evaluation Division.

For more information:
Dr. John T Haman, Project Leader
jhaman@ida.org • 703-845-2132

Dr. Heather M. Wojton, Director, Operational Evaluation Division
hwojton@ida.org • (703) 845-6811

INSTITUTE FOR DEFENSE ANALYSES

IDA Product ID 3003359

A Reliability Assurance Test Planning and Analysis Tool

John T Haman, Project Leader

Rebecca M Medlin
Emma P Mitchell
Keyla Pagán-Rivera
Dhruv K Patel

Executive Summary

The Department of Defense develops and acquires some of the world's most advanced, sophisticated, and expensive systems. As new technologies emerge and are incorporated into systems, the test and evaluation community faces the challenge of ensuring that these systems undergo adequate and efficient testing prior to operational use. In this presentation, we highlight how the leveraging of Bayesian methods can help the test and evaluation community improved test efficiency in planning for the test and evaluation of a system's reliability.

When evaluating the suitability of a system, testers must gain an understanding of the system's reliability. However, the test and evaluation community must wrestle with the perennial challenge of scoping an operational test that will be long enough (whether in terms of miles, flight hours, cycles, etc.) to adequately evaluate system reliability. The length of such a test must be informed by the reliability requirement, the true reliability of the system, and the risk we are willing to accept from asserting an incorrect statement about the system's reliability.

Traditionally, we rely on reliability demonstration tests to determine a system's reliability. Such tests make use only of the data from the current test to assess whether the reliability-related quantity of interest meets or exceeds the requirement for the system. However, these tests often require lengthy testing periods, which means we cannot always accommodate them due to time and resource constraints.

In this presentation, we recommend the use of reliability assurance tests to address the dilemma posed by lengthy testing periods. These tests employ supplementary data and information sources to plan for reliability testing such as reliability models, prior test results, expert judgment, and knowledge of environmental conditions. Making use of such data and sources can often help reduce the amount of time required for testing. Bayesian test plans use supplementary data and information; hence, we refer to them as reliability assurance tests.

Unlike planning a reliability demonstration test, which a researcher can accomplish using tools such as Operating Characteristic Curves, planning a reliability assurance test adds a layer of complexity that requires the researcher to have knowledge of Bayesian methods and coding expertise.

The lack of readily available, user-friendly software that would allow researchers to apply Bayesian methods for reliability test planning is hindering the broader adoption and usage of reliability assurance tests. Currently, no such software tools are accessible to researchers who wish to use them to incorporate prior test data or supplementary information into their test planning process; nor is such software available to researchers undertaking the planned post-test analysis.

To address this need, IDA 2024 Summer Associate Emma Mitchel has developed an R Shiny application that can serve as a user-friendly software tool for the broader test and evaluation community to use Bayesian methods in reliability test planning and analysis. This presentation introduces the R Shiny application she developed and summarizes its features and intended uses. Her work builds on the research presented in a previous IDA briefing, "JSM 2023: Improving Test Efficiency: A Bayesian Assurance Case Study."[1]

The tool is meant to support researchers' efforts to incorporate test data and supplementary information throughout the planning and analysis phases. During the test planning phase, researchers can use the tool to use Bayesian methods to incorporate supplementary data for the purpose of determining appropriate test length.

The tool also enables researchers to incorporate assumptions in their planning about reliability degradation or reliability improvement from one phase of testing to the next. For example, when a researcher transitions from a developmental test environment to an operational test environment, the researcher often observes a 10 to 30 percent degradation in system reliability.

During the analysis phase, researchers can use the tool to apply Bayesian methods to combine information and more precisely quantify uncertainty using credible intervals.

For this presentation, we have opted to employ a case study approach that walks the audience or reader through the use of the reliability assurance test planning and analysis tool. It goes on to compare the test lengths generated using the tool to those generated using a traditional reliability demonstration test. The presentation demonstrates that using the tool to scope this test often results in a proposed test length that is shorter than that proposed using a traditional reliability demonstration planning approach.

Plans for future work by IDA on this project include releasing the application to the test community on the Test Science Website (testscience.org), authoring a paper for the

[1] Medlin, Rebecca. IDA Product 3000024: *JSM 2023: Improving Test Efficiency: A Bayesian Assurance Case Study*. August 1, 2023.

The ITEA Journal of Test and Evaluation to increase awareness of the method and application, and extending the methods discussed here to incorporate additional failure time distributions such as the Weibull.

This page intentionally left blank

This page intentionally left blank

A Reliability Assurance Test Planning and Analysis Tool

Emma Mitchell

Mentors: John Haman, Rebecca Medlin, Curtis Miller, Dhruv Patel, and Keyla Pagán-Rivera

Summer 2024

Institute for Defense Analyses
730 East Glebe Road ● Alexandria, Virginia 22305

Bottom Line Up Front

Motivation

- We would like to *size* a test for evaluating system reliability.
- We would like to *incorporate prior knowledge* of the system.

Problem

- It is often difficult to scope a test that is long enough to ensure adequate assessment of system reliability.
- No user-friendly software for IDA researchers to implement prior data into reliability assurance test.

Project Goal

Develop an R Shiny app for the Test Science Website

- Planning: Use Bayesian methods to leverage supplementary data when determining appropriate test length, which may reduce testing.
- Analysis: Use Bayesian methods to combine information and better quantify uncertainty through credible intervals.

Outcome

A Bayesian reliability tool for IDA Researchers to use in reliability planning and analysis.

Big Picture – *What is* Reliability and *why* is it important to DOT&E?

Definition

Importance

Reliability

Probability that a system performs its intended function under operating conditions for a specified period of time.

The reliability quantity of interest is typically the MTBF.

Example: Consider an elevator at IDA. We want this elevator to be reliable, meaning it should consistently operate without getting stuck during normal usage over time.

MTBF – Mean Time Between Failures

Big Picture – *What is* Reliability and *why* is it important to DOT&E?

Definition

Importance

Reliability

An unreliable system may **impact the mission**, resulting in **errors**, **delays**, or even **accidents**.

Unreliable systems also increase system costs over time.

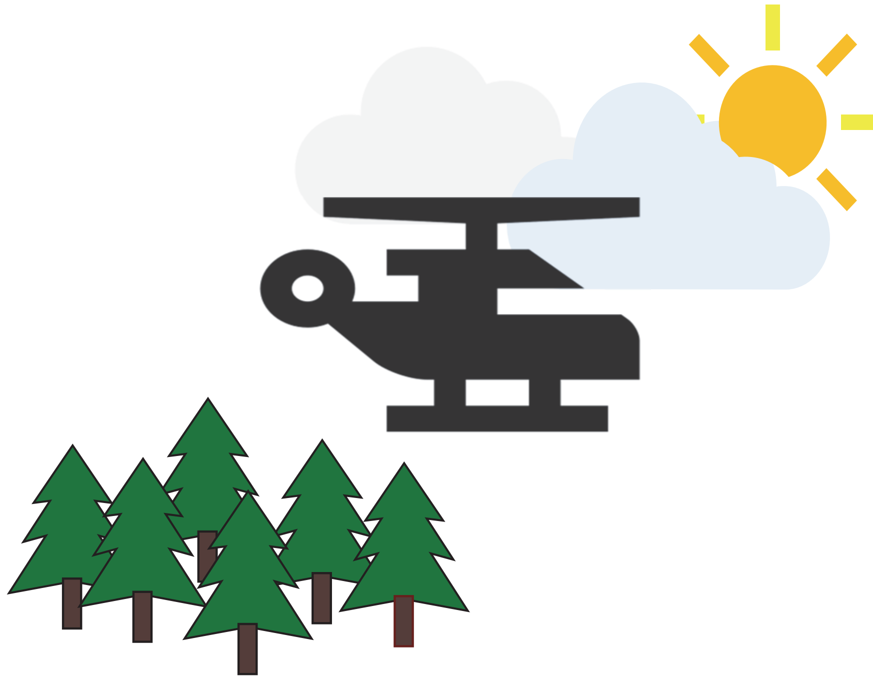
Planning a test with reliability in mind will inform us which systems are robust, cost-effective, and capable of meeting operational demands.

MTBF – Mean Time Between Failures

Case Study: Testing a New Variant of a Cargo Helicopter

Supports a variety of operations, such as:

- Air movement (e.g., carry cargo, personnel)
- Casualty evacuation
- Aerial recovery
- Area resupply



Case Study: Testing a New Variant of a Cargo Helicopter

- Reliability requirement of **35 flight hours** MTBMA.
- Reliability objective of **70 flight hours** MTBMA.
- Prior developmental testing included **329 flight hours** and **7 mission aborts.**
- **Some system changes** might imply an estimated **10 to 30 percent decrease** in reliability from DT to OT.

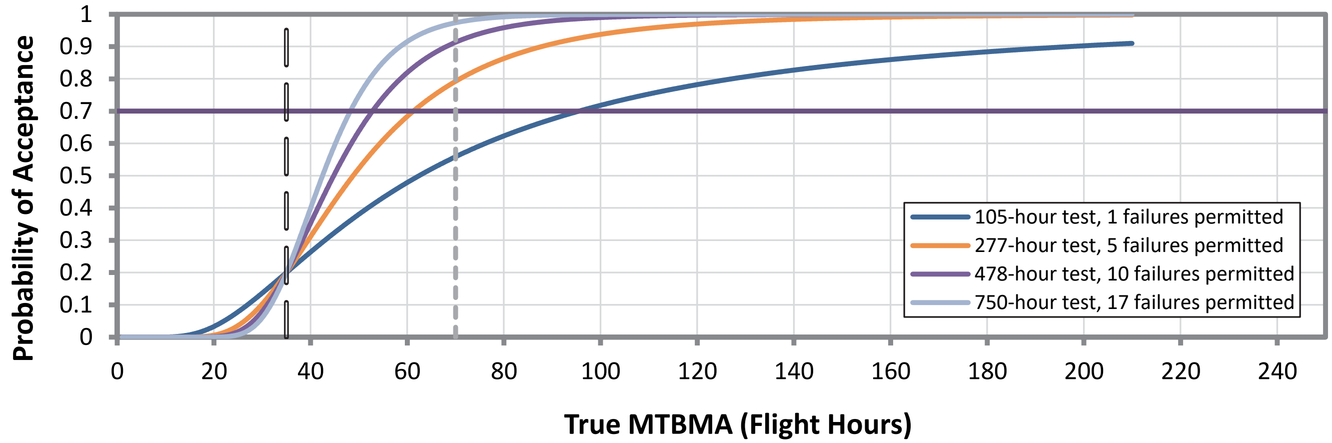
How long should we test the cargo helicopter, and what should the maximum number of failures allowed be when demonstrating the requirement with confidence?

MTBMA – Mean Time Between Mission Aborts

Reliability Demonstration Test – The Traditional Approach to Test Planning

A reliability demonstration test assumes we will **use only the data from the current test** to assess whether the reliability quantity of interest (e.g., MTBMA) meets or exceeds the requirement.

- An **Operating Characteristics Curve** is a useful tool for determining and comparing reliability plans.
  - The test length is informed by the reliability requirement, the true reliability of the system, and the risk we are willing to accept in making an incorrect statement about the system's reliability.



A Reliability Demonstration Test often requires lengthy testing periods or a very high true MTBF. (Unlikely!)

MTBF – Mean Time Between Failures; MTBMA – Mean Time Between Mission Aborts

Reliability Assurance Test – A Bayesian Approach to Test Planning

A Bayesian reliability assurance test **uses all contextual information** with the goal of reducing the required amount of testing.

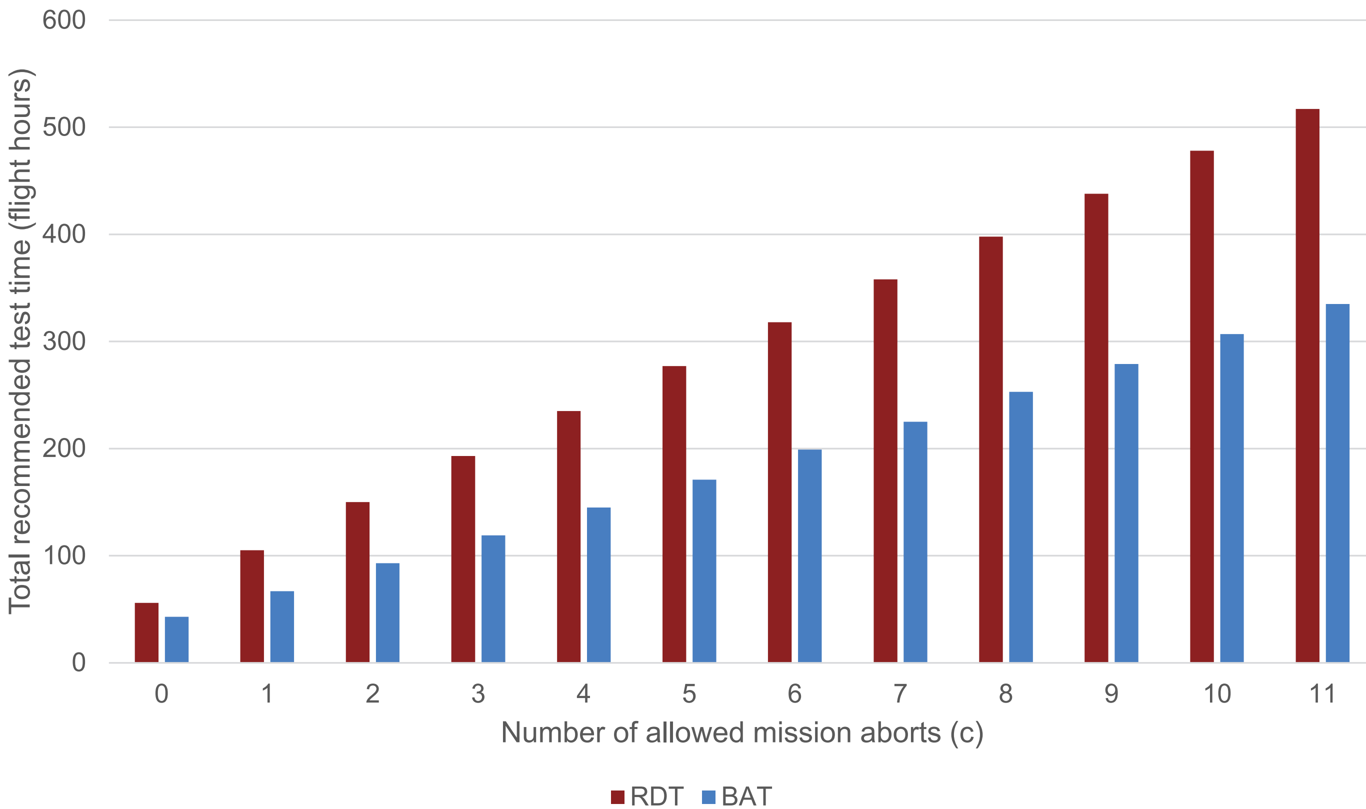
Contextual information includes:

- **Earlier test results on the same or similar systems**
- Expert judgment regarding performance
- Knowledge of the environmental conditions under which the systems are used
- Prior knowledge of possible failures

A Bayesian reliability assurance test often works best when…

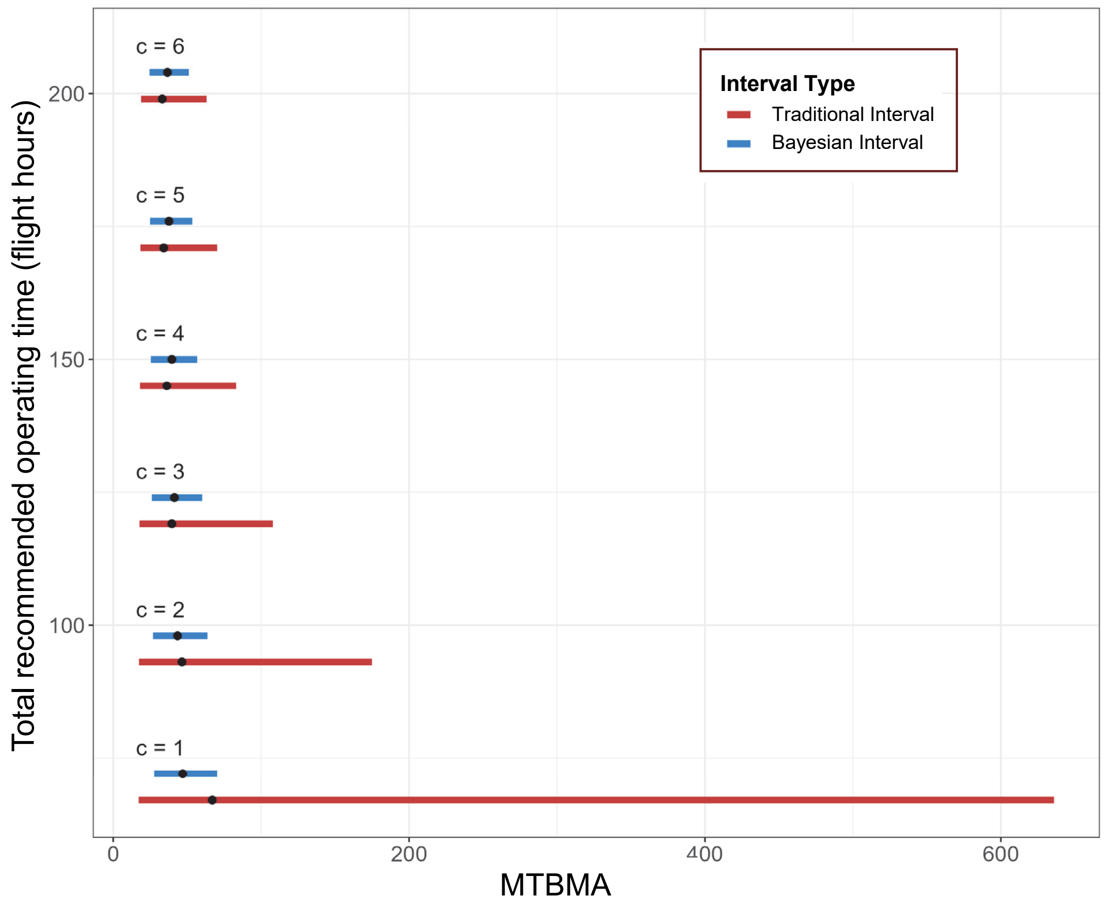
- **Minimal system changes** expected between testing
- **Longer previous testing periods**
- Proposed test length from a **reliability demonstration test is not feasible**

Bayesian Reliability Assurance test plans are <u>more efficient</u> than traditional Reliability Demonstration Plans in this case study.



BAT – Bayesian Assurance Test; RDT – Reliability Demonstration Test

Incorporating previous testing data reduces uncertainty around the MTBMA estimates.



c = maximum number of allowed mission aborts from the test plan.

MTBMA – Mean Time Between Mission Aborts

Bayesian Reliability Assurance Test – What are the steps?

Step 1: Define the prior, likelihood, posterior, threshold, and objective. Determine whether a change (degradation or improvement) from one testing phase to another is likely.

Step 2: Determine the maximum amount of consumer and producer risk you are willing to take.

Step 3: Solve for test plans.

Step 4: Compute the unconditional probability of passing for each test plan.

Reliability Assurance Test

Bayesian Reliability Assurance Test – What are the steps?

Step 1: Define the prior, likelihood, posterior, threshold, and objective. Determine whether a change (degradation or improvement) from one testing phase to another is likely.

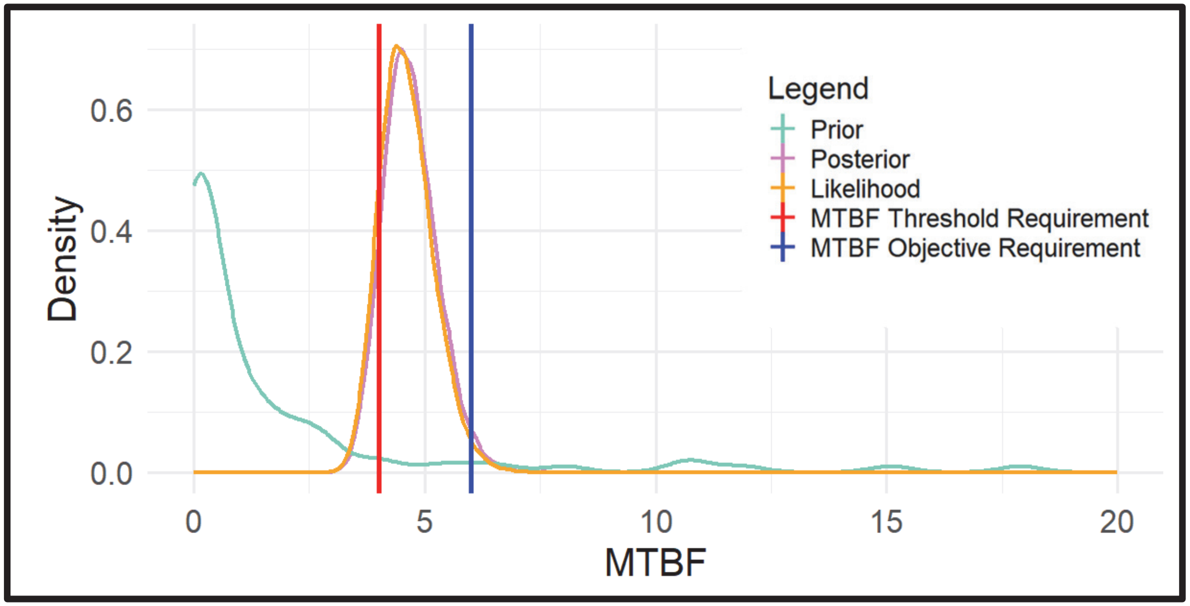
Step 2: Determine the maximum amount of consumer and producer risk you are willing to take.

Reliability Assurance Test

Step 3: Solve for test plans.

Step 4: Compute the unconditional probability of passing for each test plan.

First, define the Prior, Likelihood, and Posterior.



c = number of mission aborts from DT

T_i = time between each mission abort

$$T = \sum_{i=1}^{c} T_i = \text{total flight hours from DT}$$

λ = abort rate

$$MTBMA = \frac{1}{\lambda}$$

$$\lambda \sim gamma(a_1 = .001, b_1 = .001)$$

$$T_i \sim exponential(\lambda)$$

$$\lambda | T \sim gamma(a_1 + c, b_1 + T)$$

The posterior is the **updated distribution** of MTBF given our **supplemental data** and will serve as our prior belief about MTBF when we **solve for a test plan**.

DT – Developmental Testing; MTBF – Mean Time Between Failures; MTBMA – Mean Time Between Mission Aborts

Bayesian Reliability Assurance Test – What are the steps?

Step 1: Define the prior, likelihood, posterior, threshold, and objective. Determine whether a change (degradation or improvement) from one testing phase to another is likely.

Step 2: Determine the maximum amount of consumer and producer risk you are willing to take.

Step 3: Solve for test plans.

Step 4: Compute the unconditional probability of passing for each test plan.

Reliability Assurance Test

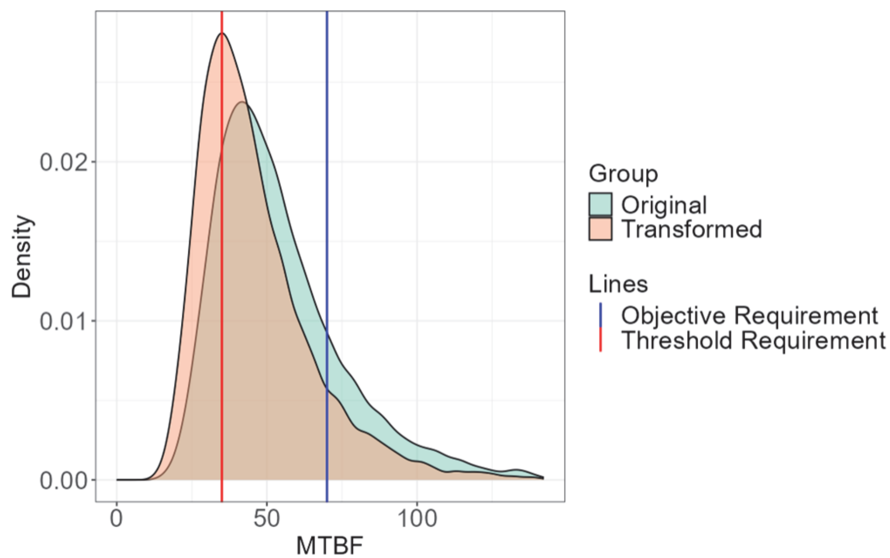
… And determine whether a change (i.e., a degradation or improvement) from one test phase to another is likely.

Example Phases: DT to OT

Developmental Testing (DT):
- Usually conducted in a more controlled environment.
- Tested by non-operationally representative users.
- Testing is often longer, with a focus on verifying system requirements.

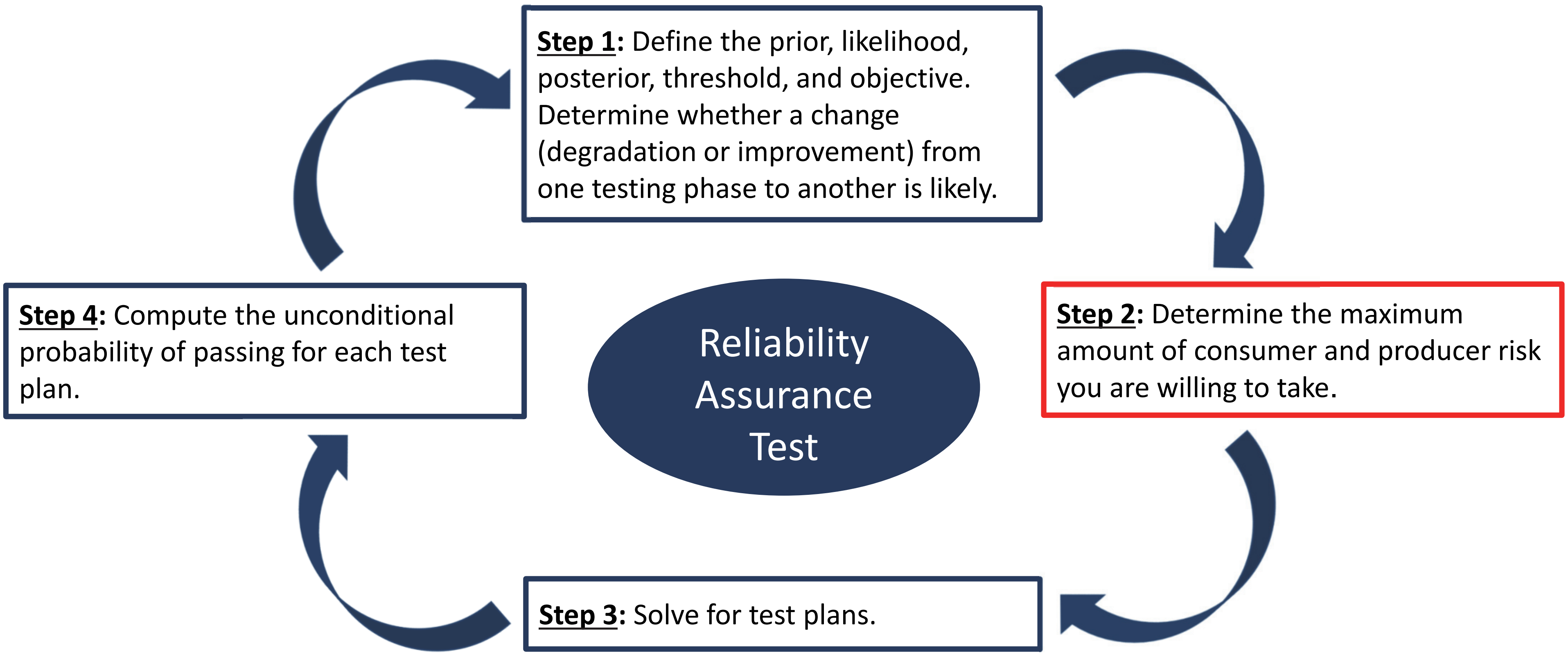
Operational Testing (OT):
- Production-representative systems.
- Conducted in operationally realistic environment.
- Tested by real users.
- Duration of testing is often limited.

- We may want to degrade or improve our prior information to be more representative of our current testing state.
- Failing to incorporate a degradation factor could lead to inadequate test plans.



Source: Gilman, J. F., Fronczyk, K. M., & Wilson, A. G. (2018). Bayesian modeling and test planning for multiphase reliability assessment. *Quality and Reliability Engineering International, 35*(3), 750-760. https://doi.org/10.1002/qre.2406

Bayesian Reliability Assurance Test – What are the steps?



Second, choose the maximum amount of risk you are willing to take.

Posterior Producer's Risk (α): The risk of having an MTBF above the objective, given the test has failed.

Context: If the test is failed, the producer requires there to be no more than a **30 percent chance** that the true MTBMA of the helicopter is **greater than 70 flight hours**.

Posterior Consumer's Risk (β): The risk of having an MTBF below the requirement, given the test has passed.

Context: If the test is passed, the consumer requires that there be no more than a **20 percent chance** that the true MTBMA of the helicopter is **less than 35 flight hours**.

A test is **passed** if we observe **c** or fewer failures in total operating time **T**.

MTBF – Mean Time Between Failures; MTBMA – Mean Time Between Mission Aborts

Bayesian Reliability Assurance Test – What are the steps?

Reliability Assurance Test

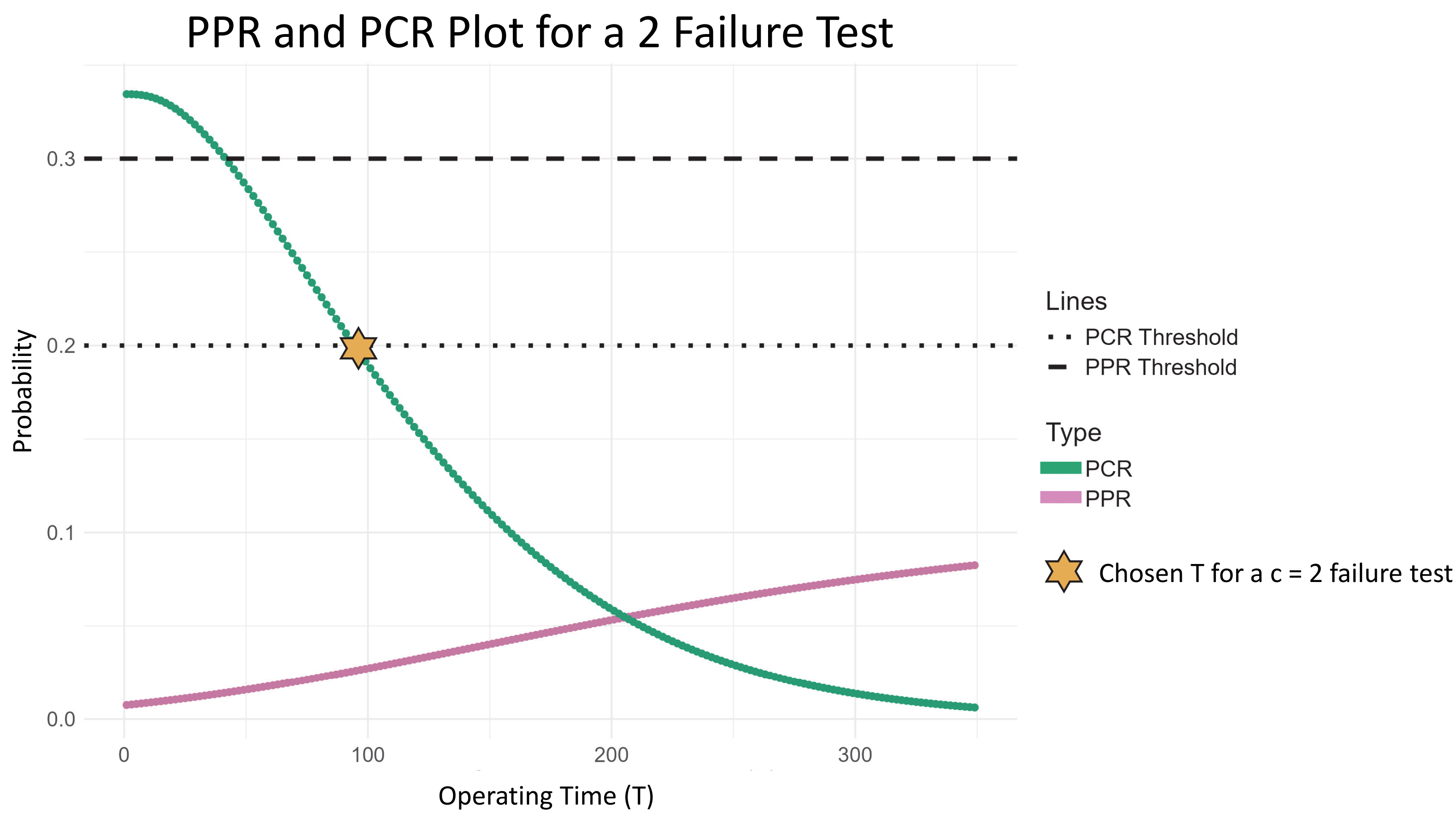
<u>Step 1</u>: Define the prior, likelihood, posterior, threshold, and objective. Determine whether a change (degradation or improvement) from one testing phase to another is likely.

<u>Step 2</u>: Determine the maximum amount of consumer and producer risk you are willing to take.

<u>Step 3</u>: Solve for test plans.

<u>Step 4</u>: Compute the unconditional probability of passing for each test plan.

For each c, the chosen test is the shortest test that satisfies both PPR and PCR requirements.



C – Maximum allowed number of failures; PCR – Posterior Consumer's Risk; PPR – Posterior Producer's Risk

Bayesian Reliability Assurance Test – What are the steps?

Reliability Assurance Test

<u>Step 1</u>: Define the prior, likelihood, posterior, threshold, and objective. Determine whether a change (degradation or improvement) from one testing phase to another is likely.

<u>Step 2</u>: Determine the maximum amount of consumer and producer risk you are willing to take.

<u>Step 3</u>: Solve for test plans.

<u>Step 4</u>: Compute the unconditional probability of passing for each test plan.

UPPT determines how likely your system is to pass.

UPPT can be used to **help the evaluator choose a test plan** in a situation where multiple test plans meet both PPR and PCR thresholds.

Low UPPT is bad: The system will likely **fail** future tests.

Historical data indicate low reliability.

Test plan is inadequate.

Requirements are stringent.

High UPPT is suspicious: The system will likely **pass** future tests.

- Historical data indicate high reliability.
- Test plan requires lots of testing.
- Requirements are lenient.

Low UPPT (<.5) Good UPPT (between .5 and .9) High UPPT (>.9)

OT – Operational Testing; DT – Developmental Testing; UPPT – Unconditional Probability of Passing Test

Low UPPT Example

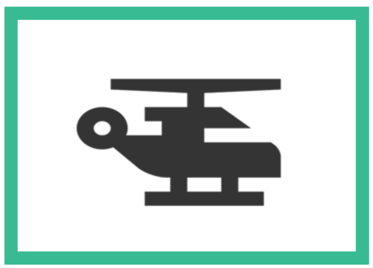
If …

Test Plan	+	DT Data	+	Requirement
0 allowed MA 43 hours		7 MA 329 hours		35 hours (MTBMA)

$\xrightarrow{\textit{implies}}$ $\mathrm{P}(\text{observe } \leq 0 \text{ fails in 43 hrs of OT given DT data})$
= **Small UPPT (<.5)**

Then regardless of …

OT Data



you will likely →

Decision

1 MA in 43 hours = 43 hours (MTBMA)

DT – Developmental Testing; MA – Mission Aborts; MTBMA – Mean Time Between Mission Aborts; OT – Operational Testing; UPPT – Unconditional Probability of Passing Test

Navigate with Ease

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Step 0: Input Parameters

Historical Data

DT Data

Historical Number of Failures:

7

Historical Operating Time:

329

The observed MTBF is 47

Mean Time Between Failure Thresholds:

MTBF Threshold Requirement:

35

MTBF Objective Requirement:

70

Risk Thresholds

Consumer Risk Threshold (β):

0 — 0.2 — 0.4

0 0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4

Producer Risk Threshold (α):

0 — 0.3 — 0.4

0 0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4

Test Plan Considerations:

How many failures would you like to consider in the test plan?

12

What is the longest operating time you would like to consider in the test plan?

350

Do you believe there to be a change in failure rate from previous testing?

Choose one:

- (•) Yes
- ( ) No

Please select one of the following:

- (•) Include degradation factor
- ( ) Include improvement factor

Lower Degradation Factor (%)

10

Upper Degradation Factor (%)

30

DT – Developmental Testing ; MTBF – Mean Time Between Failures

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Step 0: Input Parameters

Historical Data

DT Data

Historical Number of Failures:

7

Historical Operating Time:

329

The observed MTBF is 47

Mean Time Between Failure Thresholds:

Minimum acceptable reliability and desired reliability

MTBF Threshold Requirement:

35

MTBF Objective Requirement:

70

Risk Thresholds

Consumer Risk Threshold (β):

0 | 0.2 | 0.4

0 0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4

Producer Risk Threshold (α):

0 | 0.3 | 0.4

0 0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4

Test Plan Considerations:

How many failures would you like to consider in the test plan?

12

What is the longest operating time you would like to consider in the test plan?

350

Do you believe there to be a change in failure rate from previous testing?

Choose one:

- (•) Yes
- ( ) No

Please select one of the following:

- (•) Include degradation factor
- ( ) Include improvement factor

Lower Degradation Factor (%)

10

Upper Degradation Factor (%)

30

DT – Developmental Testing ; MTBF – Mean Time Between Failures

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Step 0: Input Parameters

Historical Data

DT Data

Historical Number of Failures:

7

Historical Operating Time:

329

The observed MTBF is 47

Mean Time Between Failure Thresholds:

Minimum acceptable reliability and desired reliability

MTBF Threshold Requirement:

35

MTBF Objective Requirement:

70

Risk Thresholds

Amount of risk we are willing to accept for making a wrong decision

Consumer Risk Threshold (β):

0 | 0.2 | 0.4

0 0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4

Producer Risk Threshold (α):

0 | 0.3 | 0.4

0 0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4

Test Plan Considerations:

How many failures would you like to consider in the test plan?

12

What is the longest operating time you would like to consider in the test plan?

350

Do you believe there to be a change in failure rate from previous testing?

Choose one:

- (•) Yes
- ( ) No

Please select one of the following:

- (•) Include degradation factor
- ( ) Include improvement factor

Lower Degradation Factor (%)

10

Upper Degradation Factor (%)

30

DT – Developmental Testing ; MTBF – Mean Time Between Failures

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Step 0: Input Parameters

Historical Data

DT Data

Historical Number of Failures:

7

Historical Operating Time:

329

The observed MTBF is 47

Mean Time Between Failure Thresholds:

Minimum acceptable reliability and desired reliability

MTBF Threshold Requirement:

35

MTBF Objective Requirement:

70

Risk Thresholds

Amount of risk we are willing to accept for making a wrong decision

Consumer Risk Threshold (β):

0 — 0.2 — 0.4

0 0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4

Producer Risk Threshold (α):

0 — 0.3 — 0.4

0 0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4

Test Plan Considerations:

How many failures would you like to consider in the test plan?

12

What is the longest operating time you would like to consider in the test plan?

350

What is the longest test time and largest number of failures you are willing to consider in the test plan?

Do you believe there to be a change in failure rate from previous testing?

Choose one:

- (●) Yes
- ( ) No

Please select one of the following:

- (●) Include degradation factor
- ( ) Include improvement factor

Lower Degradation Factor (%)

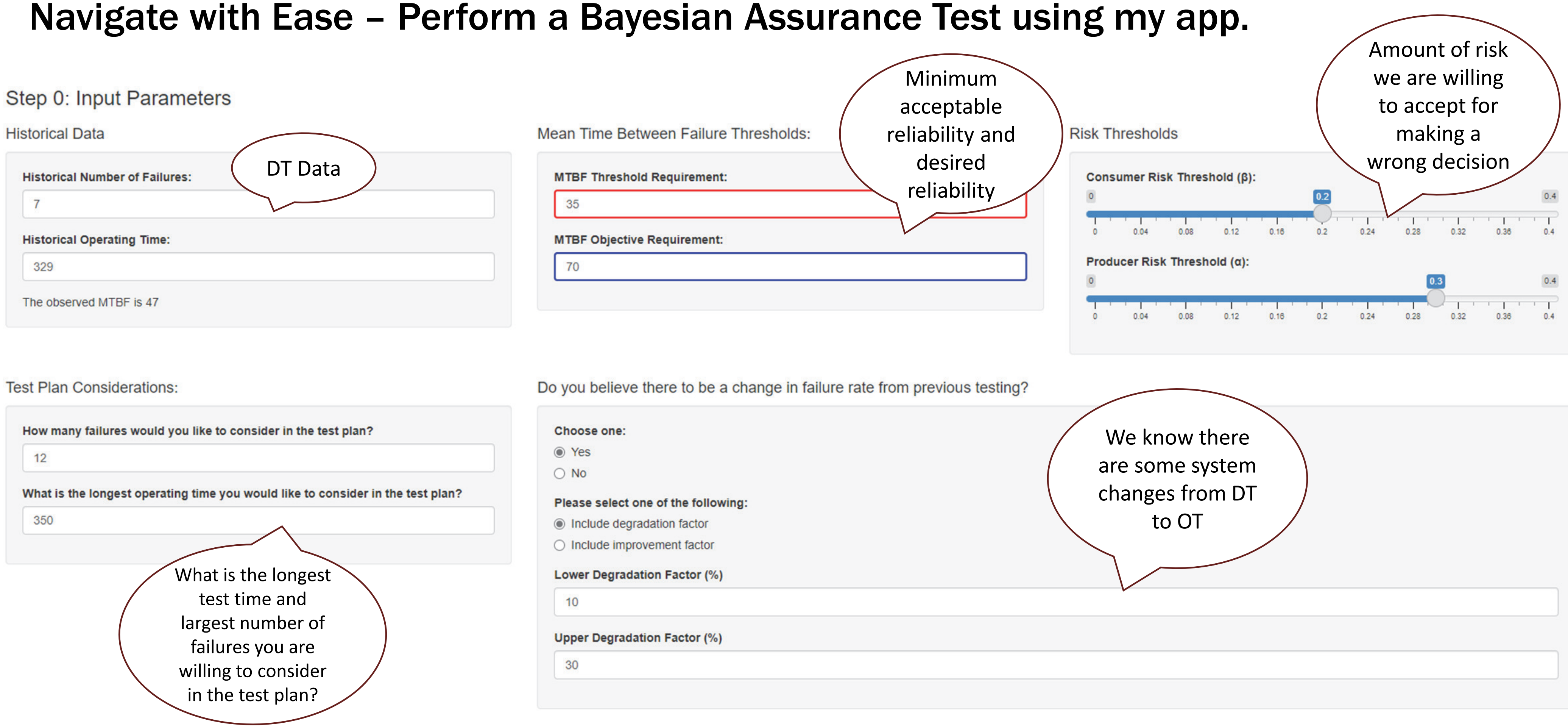
10

Upper Degradation Factor (%)

30

DT – Developmental Testing ; MTBF – Mean Time Between Failures

Navigate with Ease – Perform a Bayesian Assurance Test using my app.



DT – Developmental Testing ; MTBF – Mean Time Between Failures

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Step 2: Compute Test Plans

Unconditional Probability of Passing < 0.5 | Unconditional Probability of Passing between 0.5 and 0.9 | Unconditional Probability of Passing > 0.9

Show 25 entries

Search:

	Operating Time	Maximum Allowed Failures	PPR	PCR	Unconditional Probability of Passing
1	43	0	0.064	0.196	0.363
2	67	1	0.04	0.198	0.518
3	93	2	0.025	0.197	0.596
4	119	3	0.016	0.197	0.647
5	145	4	0.011	0.198	0.684
6	171	5	0.007	0.199	0.711
7	199	6	0.005	0.198	0.727
8	225	7	0.003	0.199	0.745
9	253	8	0.002	0.198	0.755
10	279	9	0.001	0.2	0.767
11	307	10	0.001	0.199	0.774
12	335	11	0.001	0.199	0.78

Showing 1 to 12 of 12 entries

Previous 1 Next

Download Table

All test plans have PPR less than .3

PPR – Posterior Producer's Risk ; PCR – Posterior Consumer's Risk; UPPT – Unconditional Probability of Passing Test

IDA

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Step 2: Compute Test Plans

Unconditional Probability of Passing < 0.5 | Unconditional Probability of Passing between 0.5 and 0.9 | Unconditional Probability of Passing > 0.9

Show 25 entries

Search:

	Operating Time	Maximum Allowed Failures	PPR	PCR	Unconditional Probability of Passing
1	43	0	0.064	0.196	0.363
2	67	1	0.04	0.198	0.518
3	93	2	0.025	0.197	0.596
4	119	3	0.016	0.197	0.647
5	145	4	0.011	0.198	0.684
6	171	5	0.007	0.199	0.711
7	199	6	0.005	0.198	0.727
8	225	7	0.003	0.199	0.745
9	253	8	0.002	0.198	0.755
10	279	9	0.001	0.2	0.767
11	307	10	0.001	0.199	0.774
12	335	11	0.001	0.199	0.78

Showing 1 to 12 of 12 entries

Previous 1 Next

Download Table

All test plans have PPR less than .3

...and PCR less than .2

PPR – Posterior Producer's Risk ; PCR – Posterior Consumer's Risk; UPPT – Unconditional Probability of Passing Test

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Step 2: Compute Test Plans

Unconditional Probability of Passing < 0.5 | Unconditional Probability of Passing between 0.5 and 0.9 | Unconditional Probability of Passing > 0.9

Show 25 entries

Search:

	Operating Time	Maximum Allowed Failures	PPR	PCR	Unconditional Probability of Passing
1	43	0	0.064	0.196	0.363
2	67	1	0.04	0.198	0.518
3	93	2	0.025	0.197	0.596
4	119	3	0.016	0.197	0.647
5	145	4	0.011	0.198	0.684
6	171	5	0.007	0.199	0.711
7	199	6	0.005	0.198	0.727
8	225	7	0.003	0.199	0.745
9	253	8	0.002	0.198	0.755
10	279	9	0.001	0.2	0.767
11	307	10	0.001	0.199	0.774
12	335	11	0.001	0.199	0.78

Showing 1 to 12 of 12 entries

Previous 1 Next

Download Table

Tradeoff between long test plans with little risk and shorter test plans with higher risk

All test plans have PPR less than .3

...and PCR less than .2

PPR – Posterior Producer's Risk ; PCR – Posterior Consumer's Risk; UPPT – Unconditional Probability of Passing Test

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Not all test plans have a good UPPT

Step 2: Compute Test Plans

Unconditional Probability of Passing < 0.5 | Unconditional Probability of Passing between 0.5 and 0.9 | Unconditional Probability of Passing > 0.9

Show 25 entries

earch:

	Operating Time	Maximum Allowed Failures	PPR	PCR	Unconditional Probability of Passing
1	43	0	0.064	0.196	0.363
2	67	1	0.04	0.198	0.518
3	93	2	0.025	0.197	0.596
4	119	3	0.016	0.197	0.647
5	145	4	0.011	0.198	0.684
6	171	5	0.007	0.199	0.711
7	199	6	0.005	0.198	0.727
8	225	7	0.003	0.199	0.745
9	253	8	0.002	0.198	0.755
10	279	9	0.001	0.2	0.767
11	307	10	0.001	0.199	0.774
12	335	11	0.001	0.199	0.78

Tradeoff between long test plans with little risk and shorter test plans with higher risk

Showing 1 to 12 of 12 entries

Previous 1 Next

Download Table

All test plans have PPR less than .3

...and PCR less than .2

PPR – Posterior Producer's Risk ; PCR – Posterior Consumer's Risk; UPPT – Unconditional Probability of Passing Test

Analyze with Ease

Navigate with Ease – Perform a Bayesian Assurance Test using my app.

Step 2: Compute Test Plans

Unconditional Probability of Passing < 0.5 | Unconditional Probability of Passing between 0.5 and 0.9 | Unconditional Probability of Passing > 0.9

Show 25 entries

Search:

	Operating Time	Maximum Allowed Failures	PPR	PCR	Unconditional Probability of Passing
1	43	0	0.064	0.196	0.363
2	67	1	0.04	0.198	0.518
3	93	2	0.025	0.197	0.596
4	119	3	0.016	0.197	0.647
5	145	4	0.011	0.198	0.684
6	171	5	0.007	0.199	0.711
7	199	6	0.005	0.198	0.727
8	225	7	0.003	0.199	0.745
9	253	8	0.002	0.198	0.755
10	279	9	0.001	0.2	0.767
11	307	10	0.001	0.199	0.774
12	335	11	0.001	0.199	0.78

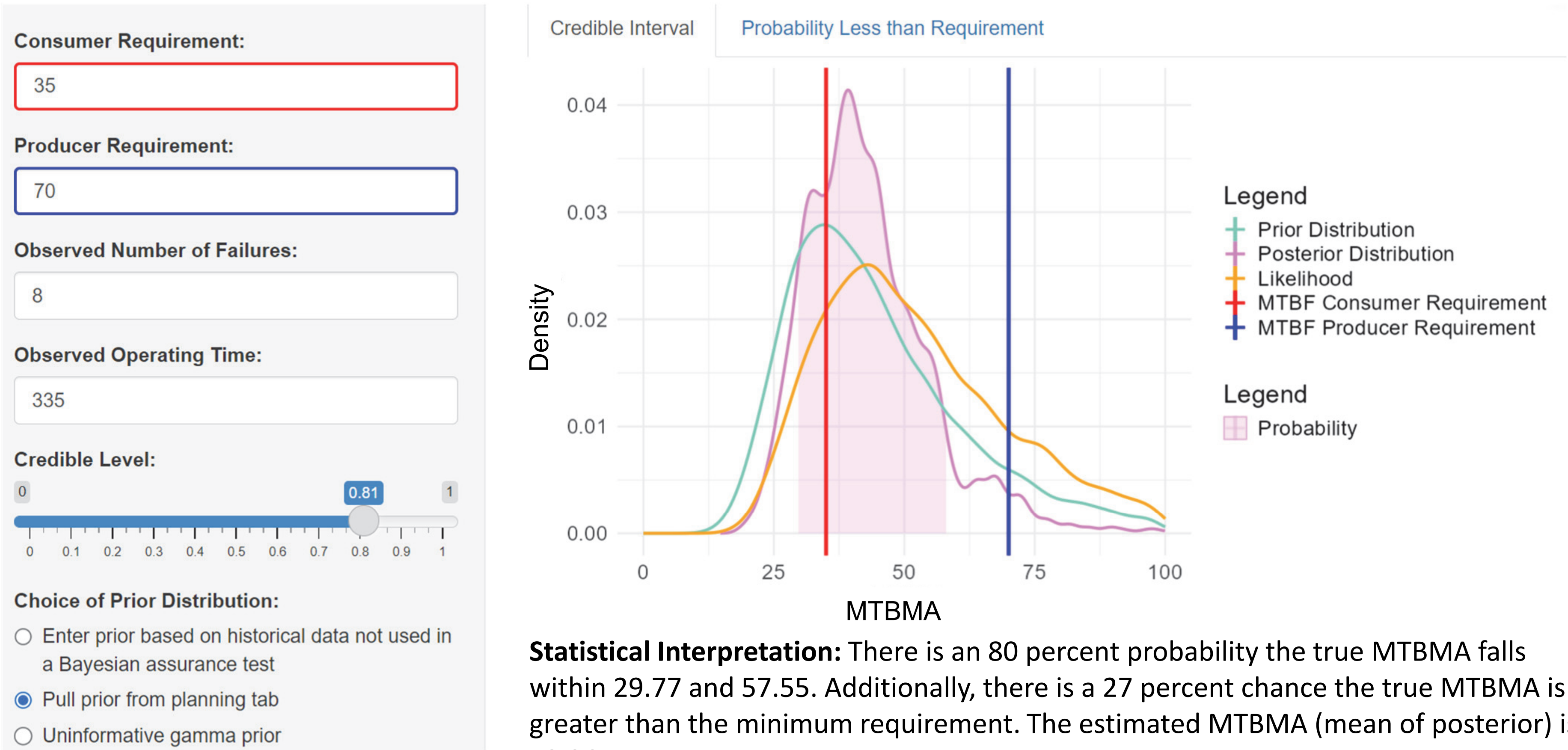
Showing 1 to 12 of 12 entries

Previous 1 Next

Download Table

PPR – Posterior Producer's Risk ; PCR – Posterior Consumer's Risk; UPPT – Unconditional Probability of Passing Test

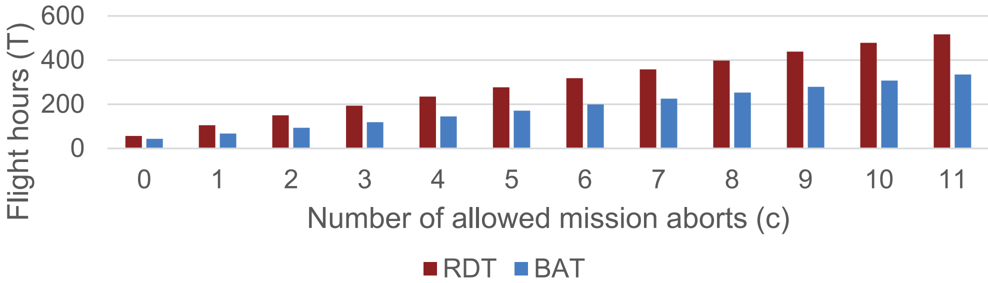
Navigate with Ease – Perform a Bayesian Assurance Test using my app.



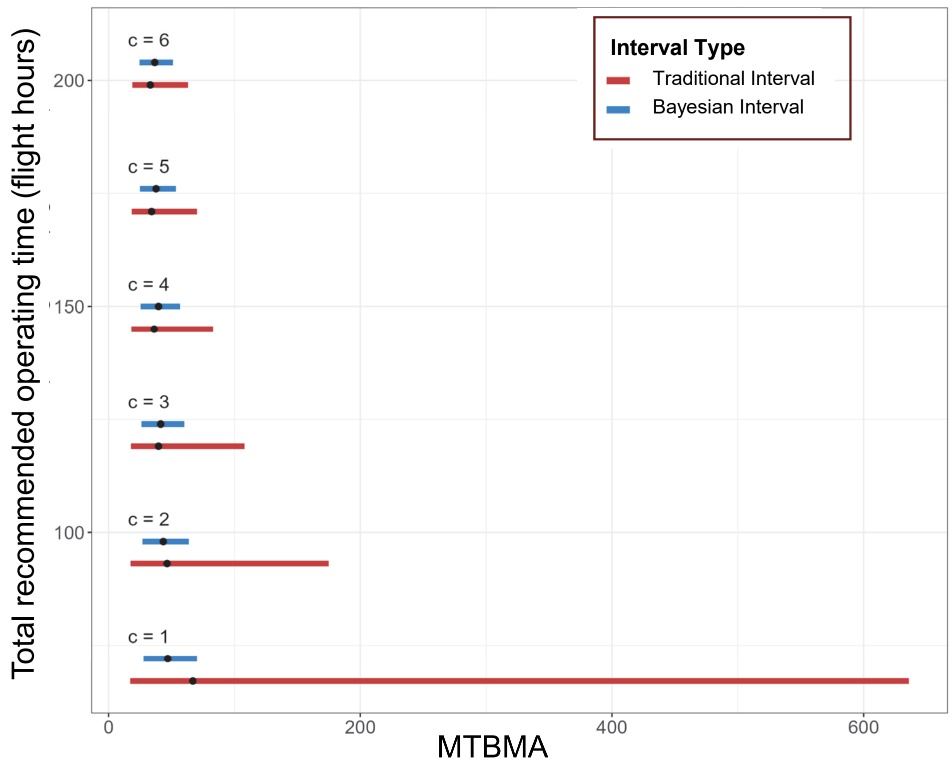
Statistical Interpretation: There is an 80 percent probability the true MTBMA falls within 29.77 and 57.55. Additionally, there is a 27 percent chance the true MTBMA is greater than the minimum requirement. The estimated MTBMA (mean of posterior) is 43.26.

MTBMA – Mean Time Between Mission Aborts

Closing Remarks



- Bayesian assurance test plans can be **more efficient** than traditional reliability demonstration test plans.
- Incorporating previous testing data can **reduce uncertainty** for the MTBF estimate.
- The app I developed provides OED researchers a **simple** and **convenient** way to implement Bayesian assurance testing.



RDT – Reliability Demonstration Test; BAT – Bayesian Assurance Test; MTBF – Mean Time Between Failures

Future Work

- Publish R Shiny Application on Test Science Website.
- Author a paper for *The ITEA Journal of Test and Evaluation* to increase awareness of method/application.
- Extend to additional failure time distributions (e.g., Weibull).

References

Gilman, J. F., Fronczyk, K. M., & Wilson, A. G. (2018). Bayesian modeling and test planning for multiphase reliability assessment. *Quality and Reliability Engineering International, 35*(3), 750-760. https://doi.org/10.1002/qre.2406

Hamada, M. S., Wilson, A. G., Reese, C. S., & Martz, H. F. (2008). *Bayesian reliability*. Springer. https://doi.org/10.1007/978-0-387-77950-8

Backup

So the second step is to decide on the maximum allowable producer and consumer risk you would like to take.

PCR

PPR

<u>Posterior Consumer's Risk</u>: The probability the system's reliability is worse than the MTBF requirement, given the test is passed and historical data on the system's reliability.

$$= P(\lambda \geq \lambda_1 | Test\ is\ Passed, \boldsymbol{x})$$

$$\approx \frac{\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\frac{\left(\lambda^{(j)}T\right)^{y}\exp\left(-\lambda^{(j)}T\right)}{y!}\right]I\left(\lambda^{(j)} \geq \lambda_1\right)}{\sum_{j=1}^{N}\left[\sum_{y=0}^{c}\frac{\left(\lambda^{(j)}T\right)^{y}\exp\left(-\lambda^{(j)}T\right)}{y!}\right]} \leq \beta$$

Where x is available historical data and $\lambda^{(j)}$ is the number of draws from the posterior given our historical data

Source: Michael S. Hamada et al., *Bayesian Reliability,* 2008, Chapter 10.

Note: The expression *Test is Failed* means that the number of observed failures is larger than the maximum number of allowed failures. That is, $y > c$. Similarly, *Test is Passed* means that $y \leq c$.

So the second step is to decide on the maximum allowable producer and consumer risk you would like to take.

PCR

PPR

Posterior Producer's Risk: The probability the system's reliability is better than the MTBF objective, given the test is failed and historical data on the system's reliability.

$$= P(\lambda \leq \lambda_0 | Test\ is\ Failed, \boldsymbol{x})$$

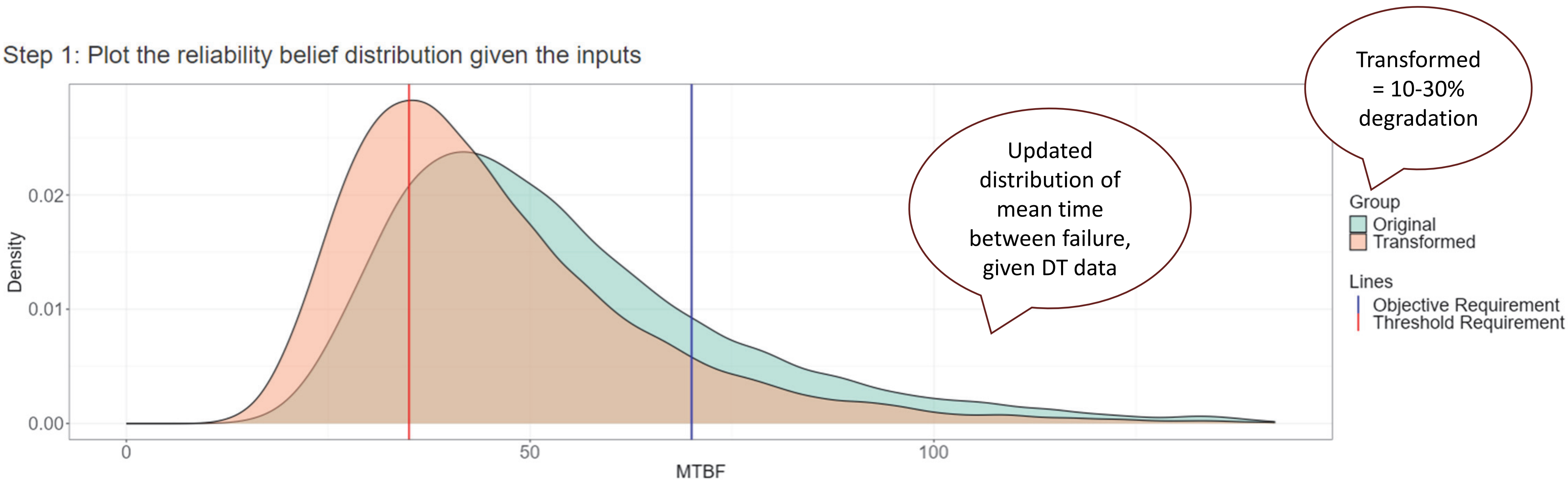
$$\approx \frac{\sum_{j=1}^{N} \left[1 - \sum_{y=0}^{c} \frac{\left(\lambda^{(j)}T\right)^{y} \exp\left(-\lambda^{(j)}T\right)}{y!}\right] I\left(\lambda^{(j)} \leq \lambda_0\right)}{\sum_{j=1}^{N} \left[1 - \sum_{y=0}^{c} \frac{(\lambda^{(j)}T)^{y} \exp(-\lambda^{(j)}T)}{y!}\right]} \leq \alpha$$

Where x is available historical data and $\lambda^{(j)}$ is the number of draws from the posterior, given the historical data.

Source: Michael S. Hamada et al., *Bayesian Reliability,* 2008, Chapter 10.

Note: The expression *Test is Failed* means that the number of observed failures is larger than the maximum number of allowed failures. That is, $y > c$. Similarly, *Test is Passed* means that $y \leq c$.

What does the Reliability Belief Distribution indicate?



Previous testing suggests reasonable reliability after degradation. A reliability assurance test may be useful.

What can we say about the true MTBF if we implement one of these test plans?

The analysis tab of the app allows the user to...

- **Quantify** uncertainty around the true MTBF with credible intervals
- **Determine** the probability of meeting the minimal acceptable reliability and desired reliability
- **Characterize** the distribution of the MTBF given our prior beliefs/data
- **Compare** the uncertainty around the estimated MTBF across multiple test plans and/or prior beliefs

MTBF – Mean Time Between Failures

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
XX-09-2024

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
A Reliability Assurance Test Planning and Analysis Tool

5a. CONTRACT NUMBER
HQ0034-19-D-0001

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Rebecca M Medlin
Emma P Mitchell
Keyla Pagán-Rivera
Dhruv K Patel

5d. PROJECT NUMBER
BD-9-2299(90) and C9082

5e. TASK NUMBER
229990 and C9082

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
730 East Glebe Road
Alexandria, Virginia 22305

8. PERFORMING ORGANIZATION REPORT NUMBER
3003359

3003359

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)
CRP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Statement A. Approved for public release: distribution is unlimited.

13. SUPPLEMENTARY NOTES
Project Leader: John Haman

14. ABSTRACT
This presentation documents the work of IDA 2024 Summer Associate Emma Mitchel. The work presented details an R Shiny application developed to provide a user-friendly software tool for researchers to use in planning for and analyzing system reliability. Specifically, the presentation details how one can plan for a reliability test using Bayesian Reliability Assurance test methods. Such tests utilize supplementary data and information, including reliability models, prior test results, expert judgment, and knowledge of environmental conditions, to plan for reliability testing, which in turn can often help in reducing the required amount of testing. In the planning phase, the application enables researchers to use Bayesian methods to incorporate supplementary data when determining appropriate test lengths. In the analysis phase, the tool allows researchers to combine information through Bayesian methods, resulting in better uncertainty quantification than traditional methods.

15. SUBJECT TERMS
Reliability, Assurance Testing, Bayesian, R Shiny, Tools

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified	Unclassified	Unclassified	SAR	51	John Haman

19b. TELEPHONE NUMBER *(Include area code)*
703-845-2132