INSTITUTE FOR DEFENSE ANALYSES

DATAWorks 2023: Back to the Future: Implementing a Time Machine to Improve and Validate Model Predictions

Benjamin A. Ashwell, Project Leader

Olivia Gozdz
Kyle E. Remley

OED Draft

April 2023

IDA Document NS D-33395

Log: H 2023-000045

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

The Institute for Defense Analyses is a nonprofit corporation that operates three Federally Funded Research and Development Centers. Its mission is to answer the most challenging U.S. security and science policy questions with objective analysis, leveraging extraordinary scientific, technical, and analytic expertise.

About This Publication
This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-19-D-0001, Task BA-9-4863, "Navy Sustainment Modeling and Analysis Support," for the Naval Supply Systems Command. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments
The IDA Technical Review Committee was chaired by Dr. V. Bram Lillard and consisted of Dr. Breeana G. Anderson, Ms. Rachel A. Haga, Dr. Thomas H. Johnson, Dr. Jason P. Sheldon, and Dr. James C. Thome Jr. from the Operational Evaluation Division, and Ms. Carrington A. Metts from the Strategy, Forces, and Resources Division.

For more information:
Dr. Benjamin A. Ashwell, Project Leader
bashwell@ida.org • (703) 845-2046

Dr. V. Bram Lillard, Director, Operational Evaluation Division
vlillard@ida.org • (703) 845-2230

Rigorous Analysis | Trusted Expertise | Service to the Nation

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-33395

DATAWorks 2023: Back to the Future: Implementing a Time Machine to Improve and Validate Model Predictions

Benjamin A. Ashwell, Project Leader

Olivia Gozdz
Kyle E. Remley

Executive Summary

At a time when supply chain problems are challenging even the most efficient and robust supply ecosystems, the Department of Defense (DOD) faces the additional hurdle of dealing primarily in low-volume orders of highly complex components that have multi-year procurement and repair lead times. When combined with the DOD's perennial budget shortfalls, these supply chain problems make it imperative for the DOD to spend money efficiently by ordering the "right components" at the "right time" to maximize readiness.

What constitutes the right components and the right time depends on model predictions that are based upon historical demand rates and order lead times. Given that the timescales between decisions and results are often years long, even small modeling errors can cause months-long supply delays or tens of millions of dollars in budget shortfalls. Therefore, we must quantify modeling errors by comparing predictions to actual observations; however, we cannot evaluate the accuracy and efficacy of today's decisions for some years to come.

To address the problem of long feedback times in forward-looking model comparisons, we have built "time machines" to pursue retrospective validation – for a given model, we rewind DOD data sources to some point in the past and compare model predictions, using only data available at the time, against known historical outcomes. This capability allows us to explore different decisions and the alternate realities that would manifest in light of those choices. In some cases, these studies are relatively straightforward, while in others they are complicated by problems familiar to any time-traveler: changing the past can change the future in unexpected ways.

In this presentation, we demonstrate our time machine capabilities for two use cases. The first case calculates the accuracy of historical procurement and repair lead time forecasts against the actual results over a multi-year span. This case is conceptually straightforward, though complex in its implementation. The second case examines the outcomes of different contracting policies – how do different sets of rules for prioritizing what to buy affect the DOD's financial and readiness risk? This example is challenging, because when we "replay" history using the different policies we have to account for how our decisions influence the future. Accomplishing this requires a tradeoff between realism and uniformity to ensure a fair comparison between contracting policies.

Contents

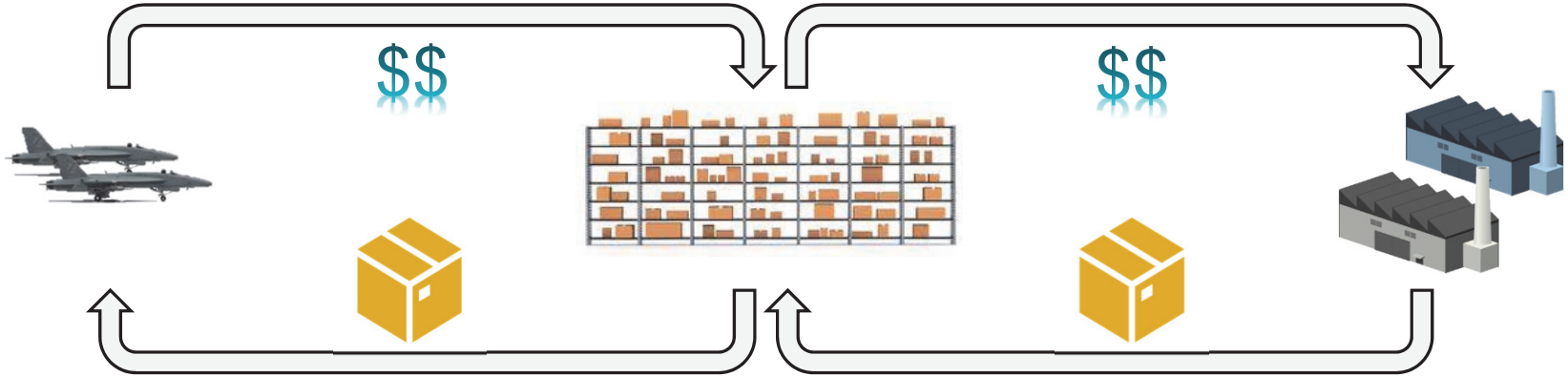
Back to the Future: Implementing a Time Machine to Improve and Validate Model Predictions

Kyle E. Remley
Olivia Gozdz

April 26, 2023

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

Sustainment is the science of maintaining readiness – can the DOD's assets perform their missions?



The role of wholesale is to ensure the right parts are on the shelf to equip the warfighter

Wholesale's budget is limited, and it relies on data-driven models to quantify the resultant risk to the sustainment system

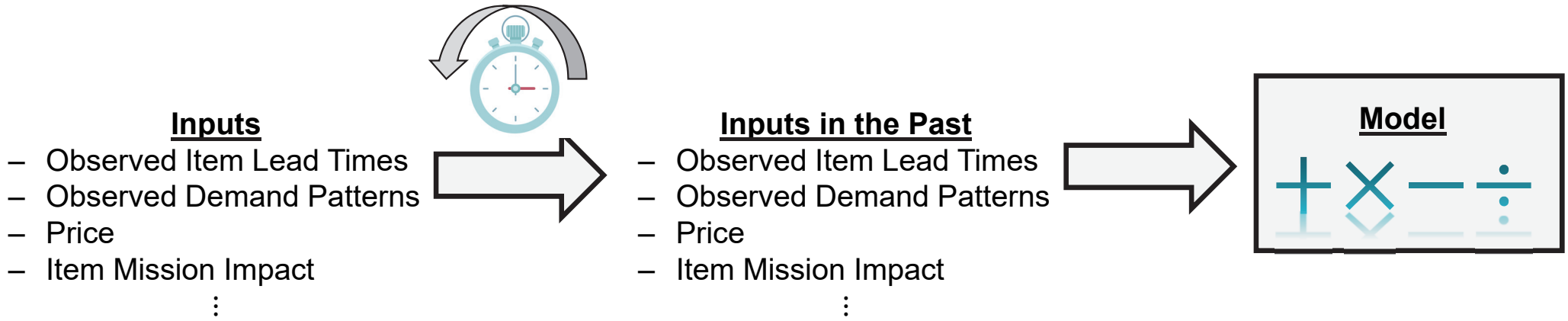
Accurate models are critical for wholesale to make the best decisions possible.

Validation is the key to ensuring models are useful, but there are challenges

Simplest approach: make a prediction and see if it comes true

But it takes a long time to get feedback

An alternative: use a "time machine" to validate model predictions against *historical* observations



Time Machine Overview

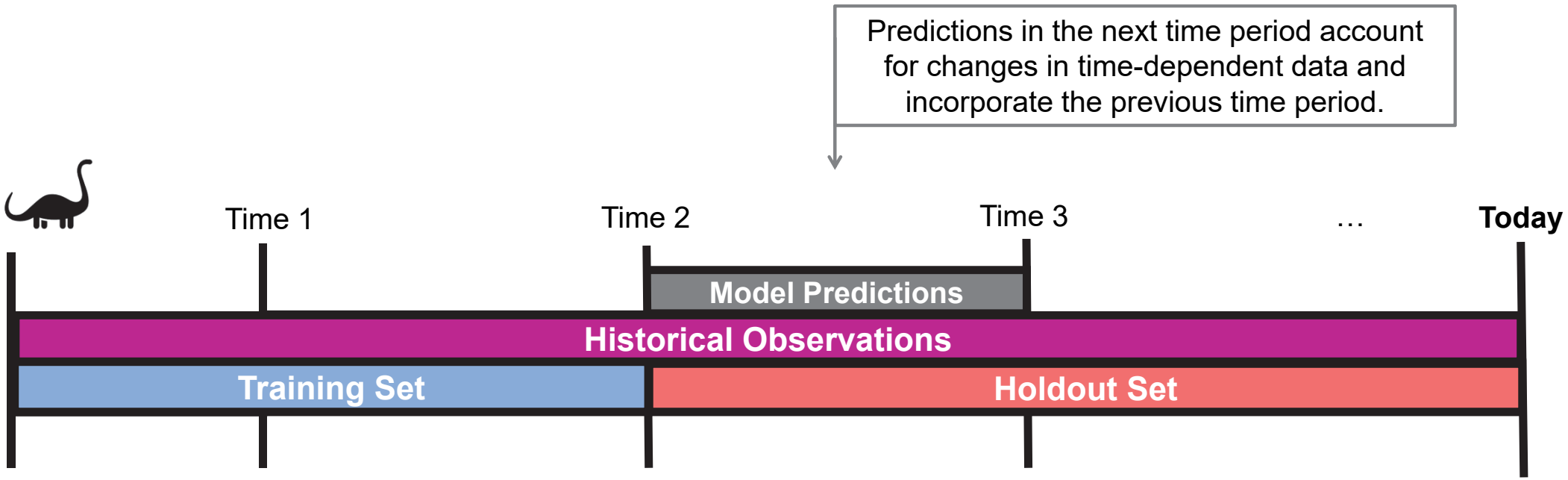
For the time machine to be effective, we have to accurately represent the past reality

One of the biggest challenges is ensuring there is no accidental cheating

Our models must use *only* data that would have been available at the point in the past that we're modeling

The time machine must account for models that "learn" with time

How do you build a time machine for a system that learns?



We iteratively produce model predictions for multiple time periods, comparing each set of predictions to known historical outcomes.

How do we ensure that our data are properly aligned at each time interval?

We don't necessarily have snapshots for every moment in time

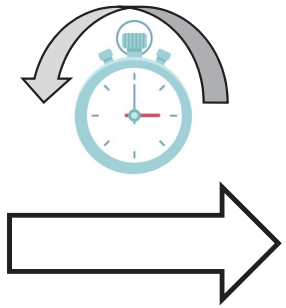
We must "rewind" current snapshot data to reflect the world as it was at any time interval

Today's contract delivery expectation:

Item: Flux Capacitor

Total Contracted: 10

Number Arrived: 8



Empirical delivery data

Yesterday's contract delivery expectation:

Item: Flux Capacitor

Total Contracted: 10

Number Arrived: 2

We need to rewind data to ensure model predictions don't cheat!

Time Machine Case 1: Contract Lead Time Validation

Background

The DOD manages stock levels by contracting with outside entities to produce or repair hundreds of thousands of different items

DOD item managers use estimated repair and procurement lead times to predict how long it takes to acquire new stock

These lead times are point estimates based on historical data and are calculated for each individual item

The ability of the DOD to maintain healthy stock levels heavily depends on having accurate lead time estimates.

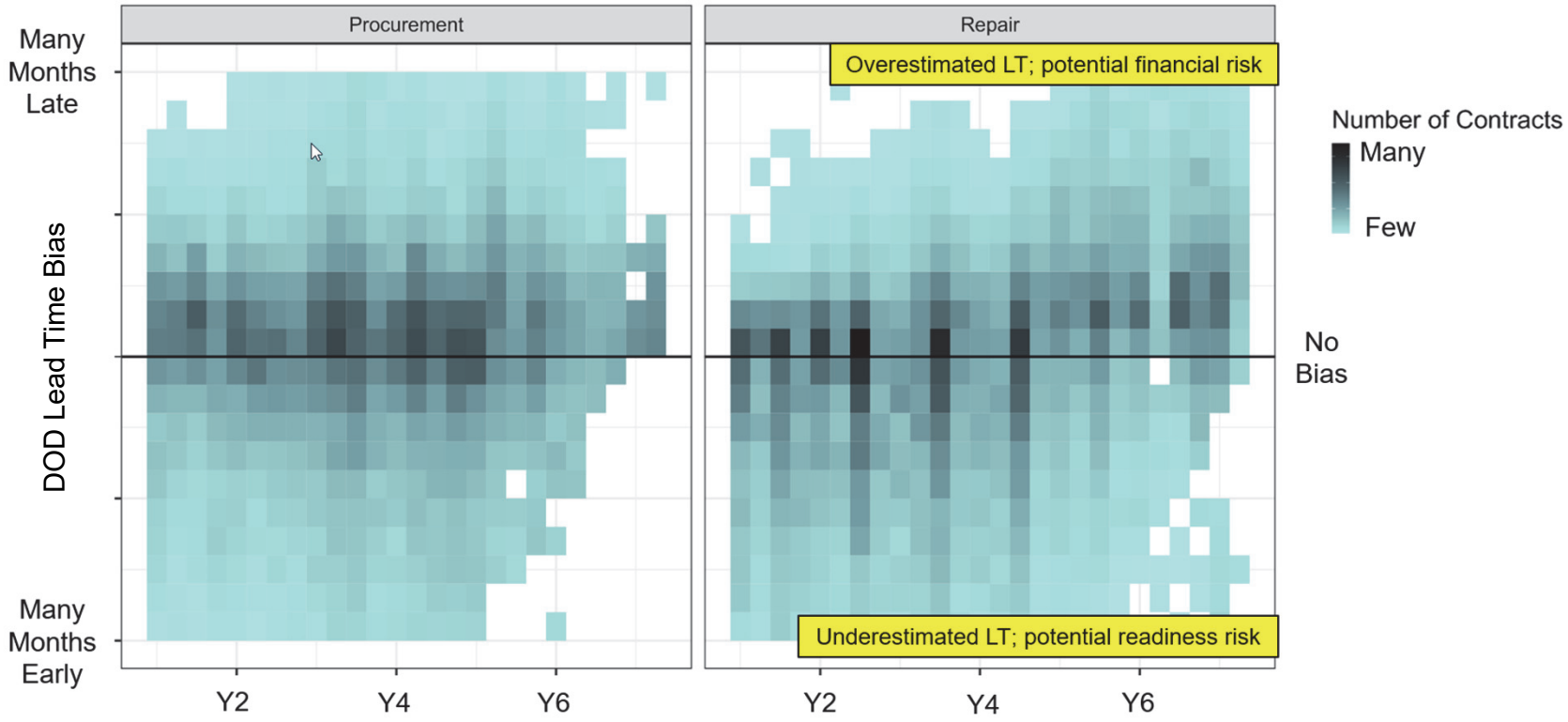
Decisions based on inaccurate lead times introduce financial and readiness risk

Underestimated lead times generally introduce **readiness risk** because you order too late

$$

Overestimated lead times generally introduce **financial risk** because you're paying too soon

The time machine can determine the accuracy and bias of lead time projections



- We can quantify the magnitude of the error of lead time projections.
- This error is significant in proportion to the magnitude of the lead times (not shown).
- We can identify bias in the error, which drives a different kind of risk.
- Our results give us ideas for how to improve these calculated lead times.

Time Machine Case 2: Contracting Policy Validation

The DOD proposed a new policy that focuses contracting efforts on the key items that drive most of the wholesale business

Wholesale manages many items and cannot afford to stock all of them; it must prioritize its contracting actions

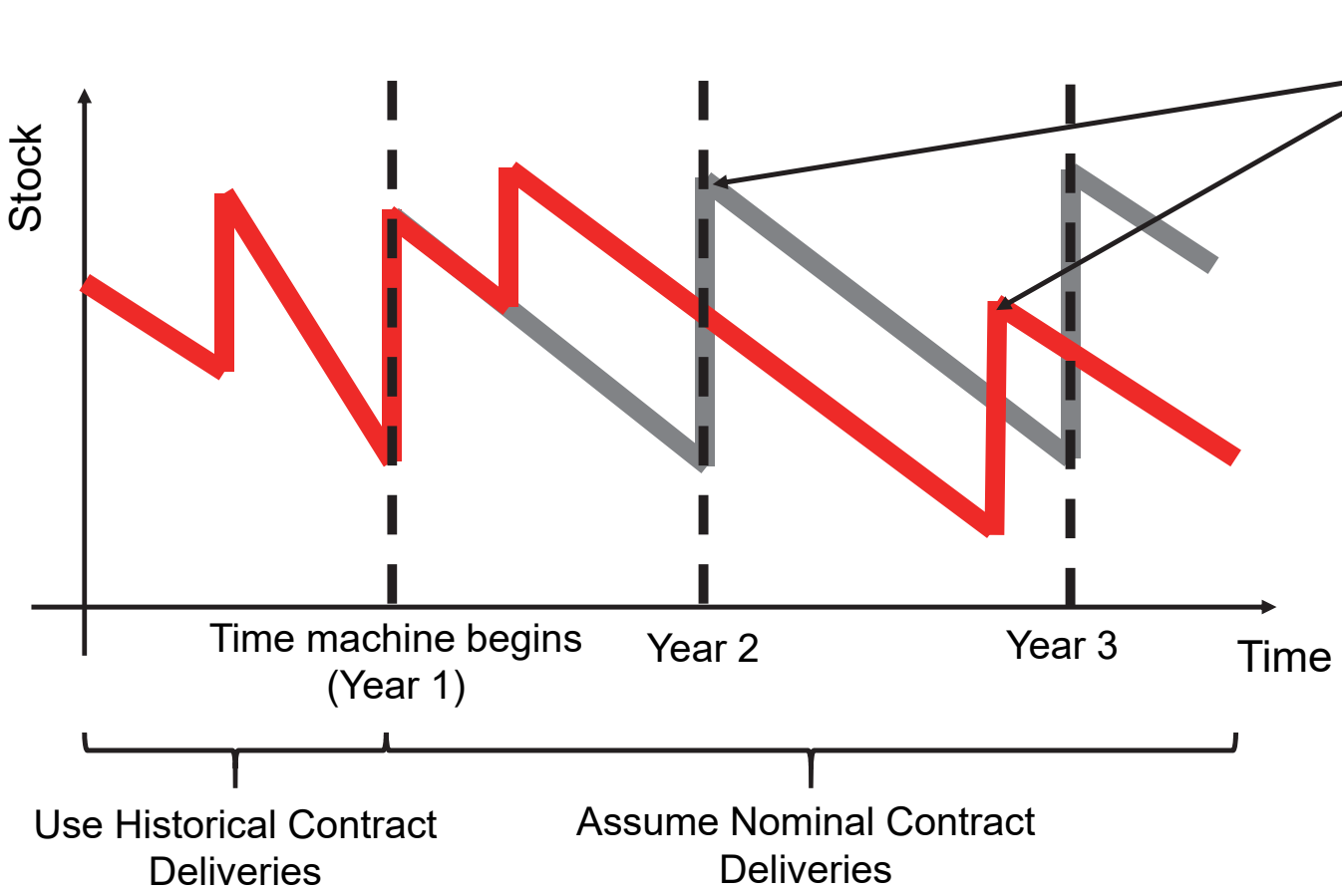
Item inventory can focus on breadth, depth, or some mix

A small number of items account for most of the demand

The DOD aims to limit its risk by increasing depth for this small number of items

This new policy makes new recommendations – “buy this, don’t buy that” – *what is the impact on wholesale*?

We used our time machine methodology to validate the new contracting policy



The new and historical policies are buying different items at different times

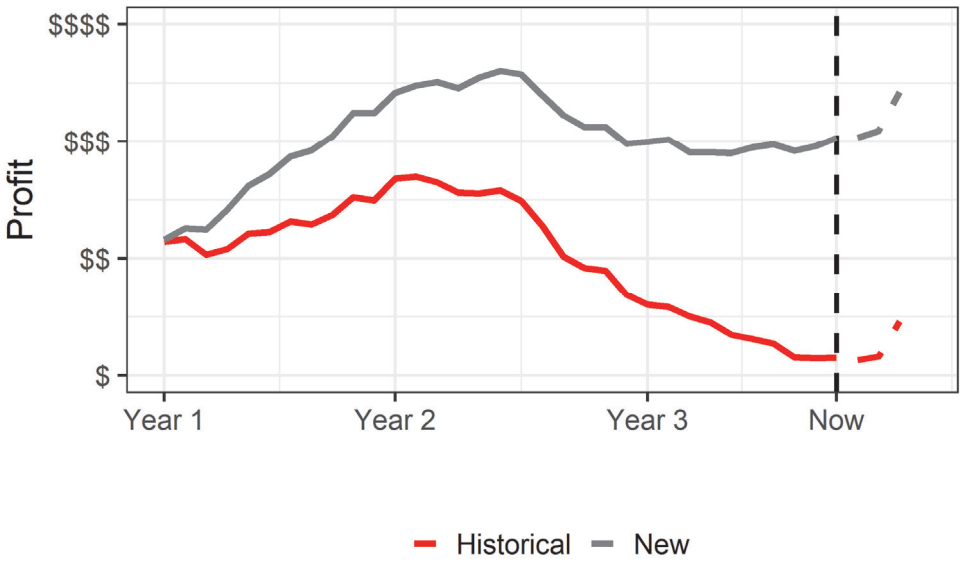
New

Historical

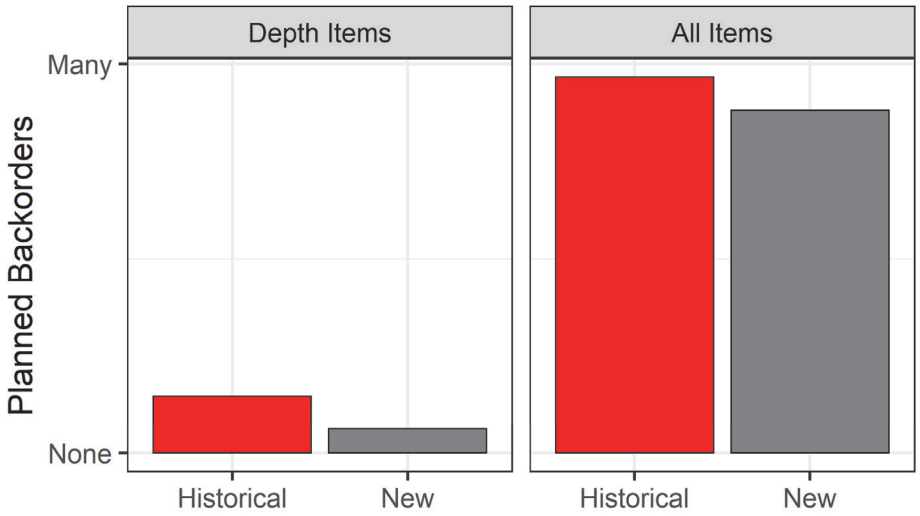
When changing the past, we need to simplify assumptions to generate an apples-to-apples comparisons.

The new policy is successful because it reduces financial risk without compromising readiness risk

Financial Risk

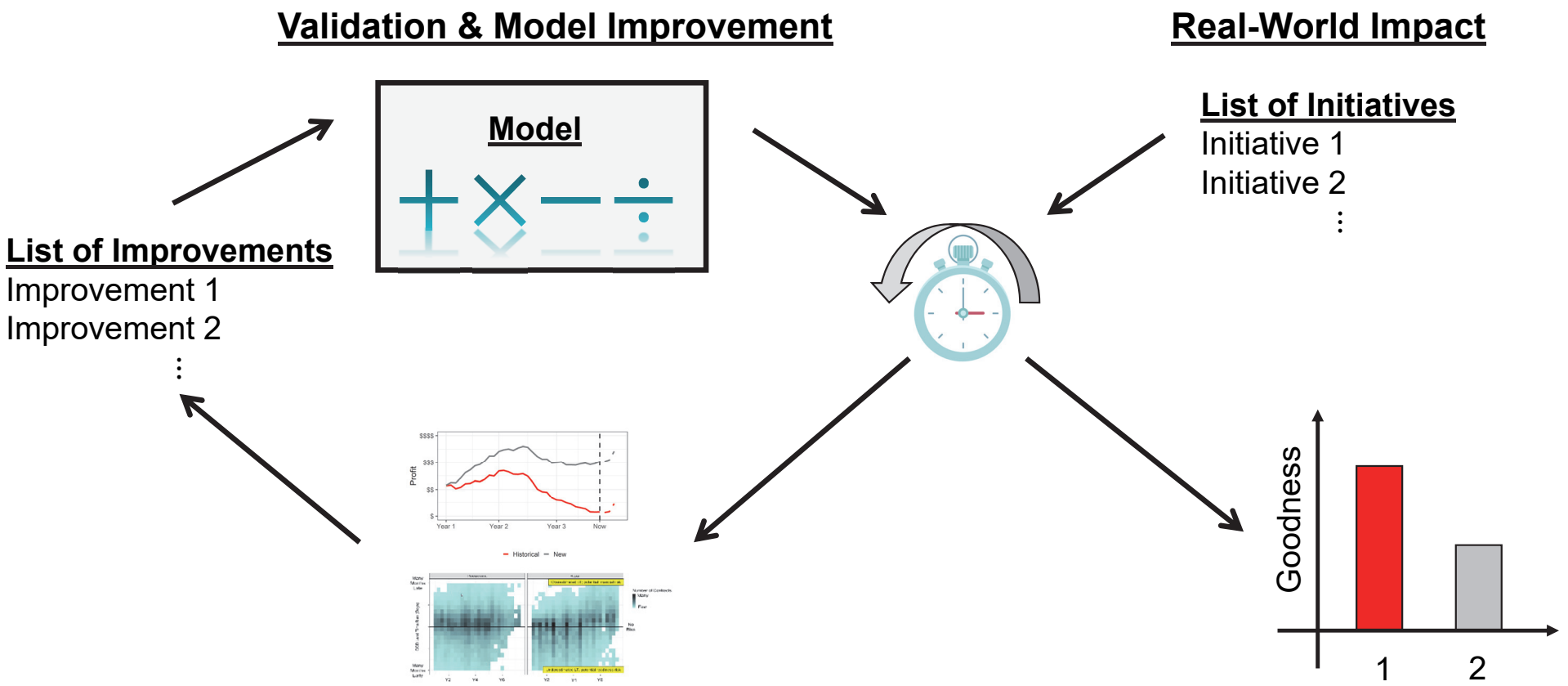


Readiness Risk



Using only the data available at the beginning of each year, the new policy can place orders that outperform the historical orders.

The time machine provides an important “yard stick” for the DOD’s modeling capabilities



We can use this capability to continuously help the DOD improve their models and risk assessments.

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
04-2023	Final	START DATE	END DATE Apr 2023

4. TITLE AND SUBTITLE
DATAWorks 2023 Back to the Future: Implementing a 'Time Machine' to Inform, Improve and Validate Model Predictions

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
HQ0034-19-D-0001		
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**
BA-9-4863	4863	

6. AUTHOR(S)
Gozdz, Oliva; Remley, Kyle, E.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Institute for Defense Analyses 730 East Glebe Road Alexandria, Virginia 22305	NS D-33395 H 2023-000045

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER
Supply Chain Management Policy & Performance Office Naval Supply Command, U.S. Navy 5450 Carlisle Pike Mechanicsburg, PA 17050-2411	NAVSUP	

12. DISTRIBUTION/AVAILABILITY STATEMENT
This OED Draft has not been approved by the sponsor for distribution and release. Reproduction or use of this material is not authorized without prior permission from the responsible IDA Division Director.

13. SUPPLEMENTARY NOTES
Project Leader: Ashwell, Benjamin A.

14. ABSTRACT
At a time when supply chain problems are challenging even the most efficient and robust supply ecosystems, the DOD faces the additional hurdles of primarily dealing in low volume orders of highly complex components with multi-year procurement and repair lead times. When combined with perennial budget shortfalls, it is imperative that the DOD spend money efficiently by ordering the "right" components at the "right time" to maximize readiness. What constitutes the "right" components at the "right time" depends on model predictions that are based upon historical demand rates and order lead times. Given that the time scales between decisions and results are often years long, even small modeling errors can lead to months-long supply delays or tens of millions of dollars in budget shortfalls. Additionally, we cannot evaluate the accuracy and efficacy of today's decisions for some years to come. To address this problem, as well as a wide range of similar problems across our Sustainment analysis, we have built "time machines" to pursue retrospective validation – for a given model, we rewind DOD data sources to some point in the past and compare model predictions, using only data available at the time, against known historical outcomes. This capability allows us to explore different decisions and the alternate realities that would manifest in light of those choices. In some cases, this is relatively straightforward, while in others it is made quite difficult by problems familiar to any time-traveler: changing the past can change the future in unexpected ways.

15. SUBJECT TERMS
Data Management; Risk Analysis; Data Analysis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	c. THIS PAGE	Unlimited	16
Unclassified	Unclassified	Unclassified		

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER
Kyle E.Remley	703-845-6898

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE.
Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE.
State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED.
Indicate the time during which the work was performed and the report was written.

4. TITLE.
Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER.
Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER.
Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER.
Enter all program element numbers as they appear in the report, e.g. 61101A

5d. PROJECT NUMBER.
Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER.
Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER.
Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S).
Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S).
Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S).
Enter report number as assigned by the sponsoring/monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT.
Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.