Innovative Techniques for Analyzing Incomplete Data to Improve Cost Estimates

George Bayer, MBA, PMP
Bryan Anderson
Cobec Consulting, Inc.

May 10, 2021
A g e n d a

• Introduction
• Government Capital Investments
• Failure Analysis
• Infrastructure Investments
• Data-Driven Cost Estimates & Business Cases
• Legacy System Cost Drivers & Failure Analyses
• Government Supply Chain
• Cost Estimating Process & Failure Analysis Methodology
  • Failure Analysis and Sustainment Methodology
  • Non-Standard Data
  • Trend Analysis, Challenges
  • Incomplete Data Sets
• Impacts of Incomplete Data on Cost & Decisions
• FAA Data Correlation Issues
• FAA Demand Data Analysis
• 2 Data Sources – Cost Accuracy & Impacts
  • Queries & Word Variations
  • Associated Words & Topic Analysis
• Conclusion
Introductions
Introduction

George Bayer

• Senior Director at Cobec Consulting
• Currently leads investment analysis consultant teams developing costs, benefits, and business cases for FAA acquisitions
• B.S. in Business Administration (Finance & English majors) from the University of Florida
• MBA in Corporate Finance from The University of Texas at Austin
• Project Management Institute (PMI) Project Management Professional (PMP)
• Over 20 years of Finance experience in capital investment valuation, forecasting & budgeting, cost estimation, benefits quantification, and business case development
• Developed discounted cash flow models in Investment Appraisal for major Power Generation capital investments at ConocoPhillips
• Evaluated major capital investments/acquisitions in the Business Case Group of Investment Planning & Analysis at the FAA
Introduction

Bryan Anderson

• Management Consultant and Programmer at Cobec Consulting
• B.S. in Economics and Mathematics from Augsburg College
• M.S. in Industrial & Systems Engineering from the College of Science and Engineering at the University of Minnesota – Twin Cities
• Over 8 years of experience in industrial engineering and systems engineering in the private and public sectors
• Active in ICEAA’s Machine Learning Working Group
• Leading database development efforts for Cobec Consulting’s Innovation Center
Government Capital Investments & Legacy Infrastructure
Federal Aviation Administration (FAA) and other civil agencies develop business cases to justify investment funding.

- Focus on Cost-Benefit Analysis to justify investments
- Brings private industry investment rigor to the agency for investment decisions
  - Identify the agency need
  - Why do we need to invest in this project?
  - What problem (shortfall) are we solving?
  - Who benefits and by how much? Benefits = FAA, airlines, public
- Two types of investments — (1) new investment, (2) infrastructure sustainment
- Quantify both costs and benefits for multiple alternative solutions
- Develop legacy reference case (compare to alternative cost models, basis of cost avoidance)
- Evaluate with Finance metrics — NPV, IRR, Payback, B/C Ration
Failure Analysis

• Government Agencies – Large Infrastructure
  • Aging infrastructure programs need operational sustainment
  • When should we replace aging infrastructure programs?
  • Establish criteria by which to justify infrastructure replacement.
    • Historical and projected parts failures
    • Difficulty to procure replacement parts
    • System obsolescence
    • Risk of loss of service

• Failure Analysis Answers:
  • What is vulnerable? Higher cost to sustain? Loss of Service?
  • Safety risk?
  • Sustainability – At what date can agency no longer sustain legacy system?
  • What parts are failing?
  • What is remaining inventory of parts, and how long will it last?
  • Are some old parts no longer manufactured and are difficult to procure? Consider lifetime buy.
  • How many parts must program procure to continue operation over lifecycle?
  • What is the total cost to sustain legacy system over lifecycle?
Infrastructure Investments

• Infrastructure Investment Decisions
  • Two Major Types of Government Investments – (1) infrastructure, (2) new capabilities
  • Approximately half of FAA business cases devoted to maintaining, improving, or replacing existing infrastructure.
  • To assess needs, timing of needs, & best spending of limited capital budgets, agency needs to:
    • Evaluate the condition of existing operations and services and their components,
    • Determine the logistics and spare parts sourcing requirements and availability,
    • Conduct an inventory analysis of spare parts’ sustainment,
    • Evaluate the supply chain of the infrastructure sustainment operations,
    • Compare infrastructure investments to determine which is more urgent,
    • Evaluate each investment and portfolio against a constrained capital budget and determine best time to invest in each project to minimize opportunity costs and maximize efficiency.

Why do we need to invest in this project?
Infrastructure Investments

• Four Cost Estimation Considerations:

  • **Cost to Sustain**  – What is the cost of sustaining operations with existing operational expenses versus replacing aging infrastructure in the National Airspace (NAS)?

  • **Ability to Sustain**  – At what point will continuing existing operations risk loss of service, or at what point will sustainment without significant investment no longer be feasible?

  • **Timing of Replacement**  – When is the best time to invest new capital to replace existing infrastructure? Consequences of not getting it right:
    
    • **Investing Too Early**
    • **Investing Too Late**

  • **Cost/Benefit Analysis**  – Justify capital investment – When do the costs of continued sustainment with increased parts failure or loss of service risk outweigh cost of replacement?
Data Driven Cost Est. & Business Cases

- **Data -Driven Estimates**
  - Cost estimates for government investments and acquisitions rely upon historical data, parametric data, and prior business cases to forecast cost.

- **Sustainment Business Cases & Legacy Case Cost Estimates**
  1) **Estimating End -of-Life**
    - How long can system be sustained without running out of parts?
    - Determine impact of EOL. Loss of Service? Impact to stakeholders?
  2) **Sustainment**
    - **Cost to Sustain** – Estimate cost to sustain legacy over program lifecycle.
    - **Further Investment** – If cannot be sustained over lifecycle, F&E investment required.
  3) **Cost Avoidance**
    - Cost of legacy case is baseline for comparison to replacement system/procurement.
    - Estimate of legacy case is estimate of costs that would be avoided if invest in replacement system. Usually, replacement system is cheaper to maintain.
Legacy System Sustainment Cost Drivers

• Legacy System Cost Driven by Parts/ System Failures
  • Agency tries to estimate cost of legacy system sustainment and timing of replacing legacy infrastructure.
  • Legacy infrastructure system costs driven by estimates of failures over lifecycle.

• Major Legacy System Sustainment Considerations:
  • Cost Per Maintenance Action
    • Logistics and maintenance costs for each outage
  • Labor
    • Repairs in the field, preparation of field spares, tracking outages
  • Analysis
    • Investment Planning & Analysis (IP&A) oversight group use legacy system sustainability as a driver for investment decision timing
    • Historical failure analysis helps demonstrate new investment urgency and helps agencies decide in which business cases to invest based on need and timing
Major Legacy System Sustainment Considerations (continued):

- **Parts Procurement, Storage, and Distribution**
  - **Procurement** – Critical Demand data drives procurement volume and operations costs to source replacement parts
  - **Storage** – Failures and parts demand drive inventory storage levels and cost
  - **Distribution** – Demand data used for “inventory readiness” and to repair right parts for distribution

- **Time to Repair**
  - Calculate Mean Time Between Failure (MTBF) to measure parts demand and inventory depletion

- **Ability to Source Parts & Manage Risk**
  - Some parts are customized; others no longer manufactured; some drive system end-of-life.
  - Agency manages parts sustainability risk by tracking parts procurement availability, cannibalizing parts, and making lifetime buys.

- **Condition of Data for Analysis/Assessments**
  - Is historical data accurate? What is the basis?
  - Analyze source data and assumptions for fidelity.
Government Supply Chain

Supply Chain Data Challenges for Failure Analysis

- Challenges to data collection and data integrity
- Failure analyses have challenges with single source of data
Cost Estimating & Failure Analysis Methodology
Failure Analysis & System Sustainment Methodology

- Aging: The cost team developed a standard methodology for estimating legacy system sustainment, including in-depth failure analyses by part and part type. This included examining procurement practices for each part at Logistics Center.

Failure Analysis Methodology

1. **Comprehensive Inventory of Systems and Parts**
   - Conduct audit of full list of system parts and the associated Lowest Replaceable Units (LRU) numbers.
   - Get initial inventory of parts in storage

2. **Parts Categorization**
   - Break parts supply into functional categories of parts (COTS, easy to procure, hard to procure, aging, custom for the FAA)
   - **Low Risk**: No procurement risk
   - **Medium Risk**: Supplier risk
   - **High Risk**: No supplier available

3. **Historical Demand**
   - Use Logistics demand data from parts returns to the Logistics Center to estimate historical demand.

4. **Trend Analysis**
   - Analyze historical demand data by part number (NSN) and check for failure trends.
Failure Analysis Methodology

5. **Failure Growth Rate** – Estimate failure growth rate for each NSN based on trend analyses
   - **Growth Rate Regression** – Develop growth curves to refine growth rates to realistic sustainable levels using regression analysis define three primary categories of growth for each NSN – zero growth, moderate growth, and high growth.
   - **Inventory Turnover** – Forecast growth rates for entire lifecycle with annual failure rate caps of full inventory turnover every three years for high failure rate parts.

6. Collect centralized inventory counts from the Logistics Center for Exchange & Repair (E&R) parts and a starting point for expendable parts

7. **Time Horizon** – Analyze by NSN the ability to sustain procurement of each E&R and Expendable part over the system lifecycle of 20 years.

8. **Forecast Procurement** – Forecast the annual inventory procurement of NSNs for each year using the demand forecast and appropriate failure growth rates.

9. **Set limits** on annual procurements, so growth rates cannot exceed a specific annual procurement level.
   - For this FAA project, we estimated that even with an aging infrastructure system, annual procurements and demand could not exceed total inventory turnover in a three-year period.
   - After total inventory turnover, failure growth rates would be reset to gradually escalate since old aging parts inventory were replaced.

Presented for the ICEAA 2021 Online Workshop - www.iceaaonline.com
Failure Analysis Methodology

10. Lifecycle Analysis  – Finalize inventory procurement and sustainability forecast for entire business case lifecycle.
   • Assess current inventory levels deployed and at the Logistics Center
   • Compare this inventory versus standard demand levels
   • Compare these levels to full inventory needed to get to system end of life or next Tech Refresh

11. Cost Estimation  –
   • **Procurement/Sustainment Cost**  – Multiply annual inventory procurement volume by the unit cost of each spare part = Price X Volume
   • Forecast annual cost requirements to procure spares in legacy cost estimate.
   • **Corrective Maintenance**  – Measure frequency of maintenance events by quantity of parts repaired by year and multiply by labor cost
Failure Analysis Methodology

1. Comprehensive Inventory of System and Parts
   - Collect list of parts per system, associated unique numbering system, and starting inventory.

2. Parts Categorization
   - Low Risk – No procurement risk
   - Medium Risk – Supplier risk
   - High Risk – No supplier available

3. Historic Demand
   - Currently based on parts orders rather than the actual failure rates.
   - Use historic demand/failures to forecast future failures, based on failure rates and failure growth.

5. Failure Growth Rate
   - Trends relating to:
     - Observed and Forecasted failure rate escalation year-over-year
     - Capped at 1/3 inventory turnover

10. Lifecycle Analysis
    - What are inventory needs for intended life of program?
    - Inventory procurement strategy to sustain legacy system until replaced
    - Critical parts determine sustainment needs and system end-of-life

11. Cost Estimation
    - Use Logistics Center current inventory as starting base of parts inventory for sustainment analysis

6. Centralized Inventory
    - Analyze by NSN the ability to sustain procurement of each E&I and Expendable part over the system lifecycle of 20 years.

7. Time Horizon

8. Forecast Procurement
   - Forecasting annual procurement and lifecycle procurement based on:
     - Failure & failure growth rates

9. Set Limits on Annual Procurements
   - Limit annual parts procurement to 1/3 of total inventory turnover:
     - Limit to failure growth rates
     - Once turnover inventory, new inventory lowers failure rate

ICEAA 2021 ONLINE WORKSHOP | COBEC CONSULTING, INC.
Impacts of Incomplete Data on Cost & Decisions
Impacts of Incomplete Data on Cost & Decisions

• **Legacy Investment Sustainment Analysis**  – FAA Air Traffic System
  • Estimated sustainment of FAA system using Logistics demand data
  • **Source of Demand Data:** Parts orders/returns to Logistics Center
    • Source of data incomplete and imperfect match
    • Demand data from spare parts returns from field and parts orders is a proxy for demand
    • Parts not consistently returned or ordered at the time of failure leads to:
      1. Inaccurate timing of failure,
      2. Volatile year-to-year failure data,
      3. Choppy forecasting of failure data
  • **Impact:** Failure and cost analyses from imperfect historical data could result in poor timing of investment decisions
  • **Analytical Solution:** Need another source of data to add fidelity and context to business case analysis and sustainment analysis
Impacts of Incomplete Data on Cost & Decisions

Incomplete data can adversely impact (1) cost estimates and (2) investment decisions.

Field Returns to Logistics Center – Proxy for Demand

- Parts not always returned.
- Parts sometimes returned later as bundle.
- Not when failures happen

Depot

Logistics Center

Demand Data

New Parts

Broken Parts Returned

Parts returned are proxy for demand or estimate of what failed, not an exact record of the time or system which failed.

Field

Program Management Organization (PMO)

Investment Decision

Procure Parts - to extend life of old system

Capital Investment Too Late - Potential Loss of Service

Capital Investment Too Early - Opportunity Cost of Capital

Tech Refresh - Timing of new investment to replace aging legacy system

Extend Sustainment for Too Long - Risk System Obsolescence

Retire Legacy System Too Soon - Unnecessary Capital Investment

If assumptions or timing is wrong, can either:
(1) Make Acquisition too soon – waste $$$
(2) Make Acquisition too late – System Fails & Loss of Service

ICEAA 2021 ONLINE WORKSHOP | COBEC CONSULTING, INC.
FAA Data Correlation Issues

2nd Source of Data Also Incomplete

• Adds context for each failure with notes from field at time of failure

• Free text fields – not standardized recording of systems or parts

• No means of searching across consistent parts numbers

• No correlation to “Lowest Repairable Unit” (lru) at Logistics Center

• Logistics Center uses “lru” – Lowest Repairable Unit
  • Repair E&R Spare Parts at FAA Logistics Center
  • Track using NSN and IFS unique numbering system

• Maintenance System uses “LRU” – Line Replaceable Unit
  • Track according to system field location and type
  • No correlation to NSN or IFS numbering system used at Logistics Center

• Only way to correlate failure with spare part number is for order requests in IFS and part returns

• Budget-Based Ordering – Little incentive to order parts at failure due to scarce “Store Credits”
Demand Data Analysis – Context from “Free Text”

• **Challenge #1** – Need system to interpret text to correlate maintenance logging source data in free text to demand projections by part from Logistics Center
• **Challenge #2** – Dependent on order data since no recording of parts at failure

• **“Topic Analysis”**
  • **Topic Analysis** – Machine data technique that analyzes text for word clustering and related words. Modern part example: correlated counts, added context, “reset” versus “replace”
  • **Key Word Search** – Addressed Challenge #1 by searching for contextual words that had meaning or potentially described a part
  • **Word Descriptions** – Word Choice. “Timecard” example, “Clock,” “CCA,” word context of actions
  • **Intent** – Tried to understand intent. How did part failures correlate with part numbers.
  • Easier with single NSN/IFS Number
  • Challenge to equate with parts orders due to “timing.”
Topic Analysis, Key Words, & Counts

Used Word Variations, Queries, Topic Analysis to Refine Data Sets

- Word queries based on multiple variations – trackball example – “ball,” “trackball,” “TRACK BALL,” “track,” “BEARINGS,” “WHEEL”
- Parsed out common words related to parts or actions
- Pivot table used to determine frequency counts and match popular words with words in component list
- Filtered for sentences that contain that word in free text
- Percentage of frequency captured and correlated with LRU and part numbers
Topic Analysis, Key Words, & Counts

Used Word Variations, Queries, Topic Analysis to Refine Data Sets

- Parsed out common words related to parts or actions
- Pivot table used to determine frequency counts and match popular words with words in component list
- Filtered for sentences that contain that word in free text
- Percentage of frequency captured and correlated with LRU and part numbers

- Chart shows frequency of part correlated with parsed words with most frequent as a larger block
# Topic Analysis, Key Words, & Counts

Used Word Variations, Queries, Topic Analysis to Refine Data Sets

<table>
<thead>
<tr>
<th>Part#</th>
<th>Parsed..</th>
<th>Fail Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>6674571-3</td>
<td>BALL</td>
<td>TO VENDOR FOR REPAIR. WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS. G...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BROKEN TOP LEFT CORNER NEAR THE MOUNTING HOLE. WORN FOAM PICK-OFF...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BROKEN TOP LEFT CORNER NEAR THE MOUNTING HOLE. WORN FOAM PICK-OFF...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EXCESSIVELY WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS ON ALL THE U...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INTERMITTENT OUTPUT FROM SWITCH CARD. WORN FOAM PICK-OFF WHEEL AN...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INTERMITTENT OUTPUT FROM THE SWITCH CARD. WORN FOAM PICK-OFF WHEE...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INTERMITTENT SWITCH OUTPUT. SWITCHES ARE STICKY. LIQUID SPILLED ON TH...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LEFT BUTTON DOESN'T FUNCTION. WORN FOAM PICK-OFF WHEEL AND BALL BEA...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LOWER RIGHT CORNER IS BENT. WORN FOAM PICK-OFF WHEEL AND BALL BEARI...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM EITHER SWITCH CARD OR MAIN PCB. BOTH CARDS ARE DEFEC...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM EITHER SWITCH CARD OR MAIN PCB. BOTH CARDS ARE DEFE...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM EITHER SWITCH CARD OR MAIN PCB. BOTH CARDS ARE D...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM SWITCH CARD. NOTICED CORROSION ON PC CARD. WORN FOA...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM THE ‘Y AXIS ENCODER. WORN FOAM PICK-OFF WHEEL AND B...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM THE UNIT. WORN FOAM PICK-OFF WHEEL AND BALL BEARIN...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM UNIT. WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS. GLO...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM UNIT. WORN FOAM PICK-OFF WHEEL AND BALL BEARINGS. GLO...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NO OUTPUT FROM EITHER SWITCH CARD OR MAIN PCB. BOTH CARDS ARE DEFE...</td>
</tr>
</tbody>
</table>

- Chart Parsed out common words related to parts or actions
- Counts the number of times the same sentence or word grouping is found within the data
- This allows matching of word description to part number
Two Data Sources – Cost & Accuracy Impact

Impact of Using Two Incomplete Data Sources

• **1st Data Source**
  • Delay between Logistics spare parts orders or returns and actual failure
  • Dependent on parts orders to estimate demand
  • Data source alone impacts (1) spare part procurement volume, (2) timing of future capital investments
  • Cost models are (1) heavily risk-adjusted for legacy sustainment and (2) F&E investments happen sooner to avoid loss of service risk

• **2nd Data Source**
  • Interpret free text to equal part LRU
  • No data standardization
  • For some spare parts, topic analysis provides some correlation
Two Data Sources – Cost & Accuracy Impact

Impact of Using Two Incomplete Data Sources

• Conclusion
  • Two data sets even though each incomplete and reported at different time intervals added context and fidelity to failure analyses and business case projections for legacy sustainment
  • Reducing sustainment risk and increase accuracy of end-of-life date forecast
  • Reduced part procurement waste

• Combined Data – 2 Sources – Cost Estimation Impact
  1) **Reduction in Procurement Cost** – Reduced risk in failure/demand forecast for legacy part procurement
  2) **Reduction in Procurement Waste** – More accurate demand forecast allows leaner inventory and fewer parts left over after decommission
  3) **Acquisition Timing** – Capital Investment Timing improved by more accurate estimate of legacy system end-of-life
    • **Too Early** – Spend limited capital dollars on system that doesn’t need it at expense of a system that does
    • **Too Late** – Risk loss of services with $ millions impact to NAS operations