Presented at the ICEAA 2024 Professional Development & Training Workshop - www.iceaaonline.com/min2024

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Advanced Earned Value Management Using Time Series Forecasts and Regression Models

2024 ICEAA Professional Development & Training Workshop

Anna B. Peters
Cost Analyst & Statistician
Anna.Taylor@jhuapl.edu

Mark Hodgins, PhD
Economist
Mark.Hodgins@jhuapl.edu

Adam Levin
Data Engineer
Adam.Levin@jhuapl.edu

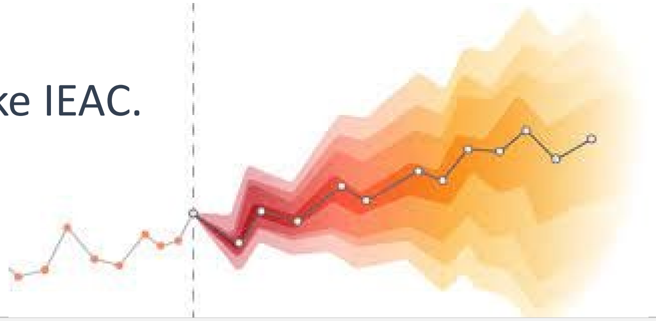
Agenda

1. Purpose
2. EVM Benefits & Limitations
3. Emergent Analytic Opportunities:
    1. Schedule estimating metrics (SEMs)
    2. Digitized acquisition data
    3. Data science tools and techniques
4. Time Series Analysis Overview
5. Autoregressive Integrated Moving Average (ARIMA) Forecasts
6. Time series Informed Independent Estimate at Completes (IEACs)
7. Macroeconomic Time Series Estimates
8. Path Forward & Conclusions

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Purpose

- The recent digitization of monthly contractor EVM data, specifically Integrated Program Management Data and Analysis Reports (IPMDARs), affords cost analysts a newfound ability to execute robust statistical and data science techniques that better predict total project cost and schedule realism.
- This presentation will discuss how several types of time series models and forecasts, prominent in private sector finance, have been employed on acquisition programs using Python based tools to raise program analysts' alertness to emergent acquisition risks and opportunities:
  - Auto regressive integrated moving average (ARIMA) models may capture the persistence and patterns in EVM variables of interest, like CPI.
  - Multi-variate time series models measure the relationship among contract performance, systems engineering data, and external macroeconomic considerations over time.
  - Both techniques, moreover, may forecast future changes in EVM variables interest, like IEAC.



Standard EVMS

Earned Value Management (EVM) is a project management technique that integrates scope, schedule, and resource measurements to assess project performance and progress.

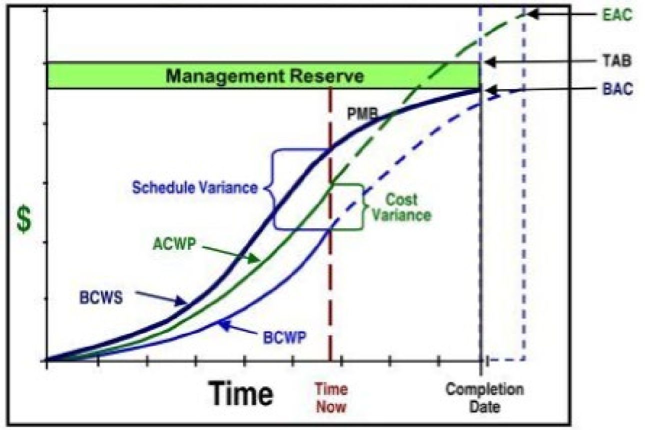
Benefits

1. **Performance Measurement:** EVM provides a comprehensive and objective way to measure project performance. It allows for a holistic view by integrating cost, schedule, and scope metrics.
2. **Early Warning System:** EVM can serve as an early warning system, identifying potential issues and variances from the project plan before they become critical problems.
3. **Objective Assessment:** EVM uses objective metrics to assess project performance, reducing the subjectivity associated with project status reporting.
4. **Cost and Schedule Integration:** EVM integrates cost and schedule information, providing a more complete understanding of the project's health.
5. **Benchmarking:** EVM allows for benchmarking and comparison across different projects, making it easier to evaluate the relative success of various endeavors.
6. **Management Visibility:** EVM provides clear visibility to project managers, stakeholders, and sponsors, enabling them to make informed decisions based on actual project performance.

Limitations

1. **Complexity:** EVM can be complex and requires a certain level of expertise to implement and interpret effectively. Small projects or those with limited resources may find EVM overly cumbersome.
2. **Data Accuracy:** EVM relies on accurate and timely data for effective implementation. Inaccurate data input can lead to misleading results and compromised decision-making.
3. **Resource-Intensive:** Implementing EVM requires additional effort and resources, including specialized training for project managers and team members.
4. **Subject to Manipulation:** There's a risk that project managers might manipulate data or make optimistic assumptions to make the project performance appear better than it actually is.
5. **Focus on Metrics:** EVM may lead to a focus on meeting metrics rather than achieving the project's ultimate objectives. This could result in a disconnect between project performance and the actual value delivered.
6. **Not Suitable for All Projects:** EVM may not be suitable for small, short-term projects or those where requirements are rapidly changing, as the overhead of EVM may outweigh its benefits in such cases.

EARNED VALUE MANAGEMENT 'GOLD CARD'



VARIANCES Favorable is Positive, Unfavorable is Negative

Cost Variance	CV	= BCWP – ACWP	CV % = (CV / BCWP) •100
Schedule Variance	SV	= BCWP – BCWS	SV % = (SV / BCWS) •100
Variance at Completion	VAC	= BAC – EAC	

OVERALL STATUS

% Schedule = $(BCWS_{CUM}$ / BAC) • 100
% Complete = $(BCWP_{CUM}$ / BAC) • 100
% Spent = $(ACWP_{CUM}$ / BAC) • 100

DoD TRIPWIRE METRICS

TW Cost Efficiency CPI = BCWP / ACWP Favorable is > 1.0, Unfavorable is < 1.0
TW Schedule Efficiency SPI = BCWP / BCWS Favorable is > 1.0, Unfavorable is < 1.0

TW **BASELINE EXECUTION INDEX (BEI) = A Schedule Metric**
BEI = Tasks with Actual Finish Date / (# of Baseline Tasks Scheduled to Finish Prior to Status Date + Tasks Missing Baseline Start or Finish Date)

TW **CRITICAL PATH LENGTH INDEX (CPLI) = A Schedule Metric**
CPLI = (CP Length $_{\text{(Time Now To Contract End)}}$) + Total Float $_{\text{(To Contract End Baseline Finish)}}$) / CP Length

Hit / Miss = Month's Tasks Completed ON or AHEAD / Month's Tasks Scheduled to Complete

\# **ESTIMATE AT COMPLETION** (EAC) = Actuals to Date + [(Remaining Work)/(Performance Factor)]

EAC_{CPI} = $ACWP_{CUM}$ + [(BAC – $BCWP_{CUM}$) / CPI_{CUM}]
= BAC / CPI_{CUM}
$EAC_{Composite}$ = $ACWP_{CUM}$ + [(BAC – $BCWP_{CUM}$) / (CPI_{CUM} • SPI_{CUM})]

\# § **TO COMPLETE PERFORMANCE INDEX** (TCPI)
$TCPI_{EAC}$ = Work Remaining / Cost Remaining = (BAC – $BCWP_{CUM}$) / (EAC – $ACWP_{CUM}$)

\# To Determine a Contract Level TCPI or EAC; You May Replace BAC with TAB
§ To Determine the TCPI $_{\text{BAC or LRE}}$ Replace EAC with BAC or LRE

Traditional EVM

- EVM is a mature standard and best practice to monitor and enforce adequate contract execution for the Department of Defense (DOD).
- EVM lays out the technical scope of a project, with the goal to minimize schedule delays and cost overruns, by creating an integrated baseline of program planning and performance. (Alexander, 2002)
- However, the limitations, specifically the manipulation and narrow metric focus, can manifest in significant issues
  - Schedule delays and cost growth are a normal occurrence for DOD acquisition programs
  - Government Accountability Office (GAO) reported that 98 major development acquisition programs (MDAPs) cumulatively overran their budget baseline by $402 billion and were an average of 22 months delayed in their schedules in 2010 (Hofbauer et al. 2011: 3).
  - Programs may incur a Nunn-McCurdy breach that necessitates a re-baseline or, worse yet, program cancellation

National Reconnaissance Office (NRO) Schedule Execution Metrics (SEMs)

- NRO created the SEMs to provide objective measures of schedule performance (NRO, 2022)
- The SEMs methods can be used to develop data driven methods, with data science and statistical techniques, to better predict program acquisition health (NRO, 2022)
- **Current Baseline Realism Index (BRI)**: Percentage of planned events that actually finished in the planning period. BRI indicates how well the contractor is following the plan within reporting period. BRI is applied using a six reporting period moving average. (NRO, 2021)
  - BRI ≥ 0.80 is favorable and can be considered “On Plan”
  - BRI ≤ 0.20 is unfavorable and can be considered “Way off plan”
- **Current Baseline Progress Index (BPI)**: Percentage of planned events that actually finished in or before the planning period. BPI indicates how many of the planned events within reporting period have been accomplished. (NRO, 2021)
  - BPI ≤ 0.35 is unfavorable and can be considered “Way off plan”
- **Current Baseline Execution Index (BEI)**: Percentage of total events finished in current reporting period. (NRO, 2021)
- APL and NRO are collaborating to further advance predictive analysis to improve acquisition program decision making and forecasting

Integrated Program Management Data and Analysis Report (IPMDAR)

- IPMDAR Data Item Description (DID) DI-MGMT-81861C contains data for measuring contractors' cost and schedule performance on Department of Defense (DoD) acquisition contracts > ~$20M.
- File format specification (FFS) provides precise digitization instructions for transparency, replicability, and auditability for the following:
  - Cost performance dataset
  - Schedule performance dataset
- Submittals are required on a monthly basis, which affords time series analysis on major development contracts
- Bottom line: Newly available IPMDAR data on current and emergent acquisition programs provides an opportunity to improve upon the limitations with standard EVMS

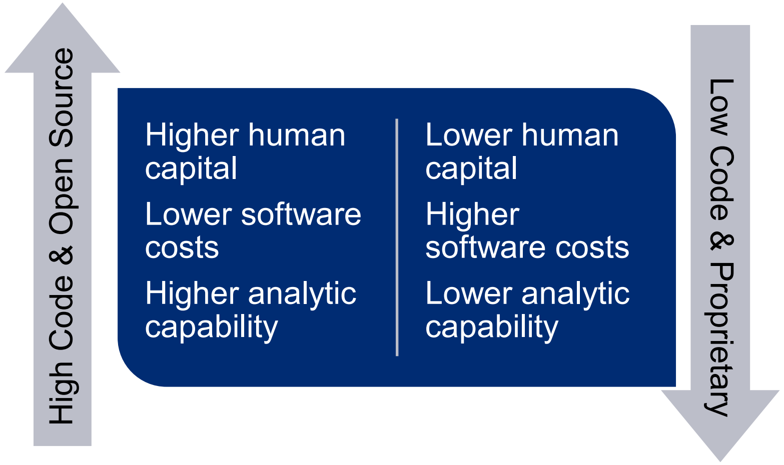
2.2.4 ContractData

Table	ContractData		
Entity	ContractData		
Fields	Name	Data Type	Nullable
	Quantity_Development	Decimal	Yes
	Quantity_LRIP	Decimal	Yes
	Quantity_Production	Decimal	Yes
	Quantity_Sustainment	Decimal	Yes
	NegotiatedContractCost	Decimal	Yes
	AuthorizedUnpricedWork	Decimal	Yes
	TargetFee	Decimal	Yes
	TargetPrice	Decimal	Yes
	EstimatedPrice	Decimal	Yes
	ContractCeiling	Decimal	Yes
	EstimatedContractCeiling	Decimal	Yes
	OriginalNegotiatedContractCost	Decimal	Yes
	ManagementEAC_BestCase	Decimal	Yes
	ManagementEAC_WorstCase	Decimal	Yes
	ManagementEAC_MostLikely	Decimal	Yes
	ContractBudgetBase	Decimal	Yes
	TotalAllocatedBudget	Decimal	Yes
	ContractStartDate	Date	Yes
	ContractDefinitizationDate	Date	Yes
	BaselineCompletionDate	Date	Yes
	ContractCompletionDate	Date	Yes
	ForecastCompletionDate	Date	Yes
	LastOTBDate	Date	Yes
Primary Key	[N/A]		
Foreign Keys	[N/A]		
Use Constraints	ContractData is a singleton.		

Analysis Execution with Python Native Environment

Overall Python Investment Rationale

- A fundamental tradeoff exists between **low-code / proprietary** and **high-code / open source** data science tools.
- Python is highly used and ranked data science programming language, with significant community support.
- Low-cost, open source language with highly readable and writable syntax/code, enabling programming ease.
- The complexity of emergent predictive analytics, evolution of software tools, sensitivity of most acquisition program data, and flexibility to interface with third party software **warrants native tool development, subject to a labor (e.g. talent and culture) investment, which aligns with the 2020 DoD Data Strategy.**



TIOBE Index* (Oct 2023)

Ranking	Language	Ratings
1	Python	14.82%
2	C	12.08%
3	C++	10.67%
4	Java	8.92%
5	C#	7.71%
6	JavaScript	2.91%

*The TIOBE Programming Community index is an indicator of the popularity of programming languages.

Analytic Environment – Native Python based DS

Raw Data Upload	PostgresSQL, Psycopg2	Together, these libraries allow for the robustness and flexibility of an SQL relational database management system
Data Cleaning & Transformations	pandas, NumPy	• Together, these libraries allow for data clean up and preprocessing necessary for model building • High Interoperability
Predictive Modeling	pmdarima, statsmodels	• These libraries are used for quantitative analysis and forecasting
Visualization & Exploratory Analysis	matplotlib, plotly, seaborn	• These libraries build static as well as dynamic and interactive visualizations
Reporting	jupyter, plotly | Dash	• Libraries allow for customized program reporting and dynamic visualizations • Low cost deployment

Time Series Analysis- ARIMA

- Time series analysis is used to forecast future trends in data with time as the independent variable. Historic data can be used over a variety of time horizons.
  - Technique has been applied with amazing success in the finance industry for many years
- Autoregressive integrated moving average (ARIMA) models combine two of the most common time series modeling formats, Auto Regressive (AR) and a Moving Average (MA) models.
  - The AR model is constructed so that Y depends only on its own lags.
  - The MA model is constructed so that Y depends on the lagged forecast errors.
- The theoretical ARIMA equation is below where Y_{t-p} and $\phi_q \epsilon_{t-q}$ are the AR and MA components of the model, respectively:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \cdots + \phi_q \epsilon_{t-q}$$

- Non-seasonal ARIMA(p, d, q) models are comprised of three terms: p, d, and q.
  - p is the order of the Auto Regressive term (the number of lags of Y that are used as predictors)
  - d is the minimum number of differencing needed to make the time series stationary
  - q is the order of the Moving Average Term (number of lagged forecast errors that should go into the ARIMA model)

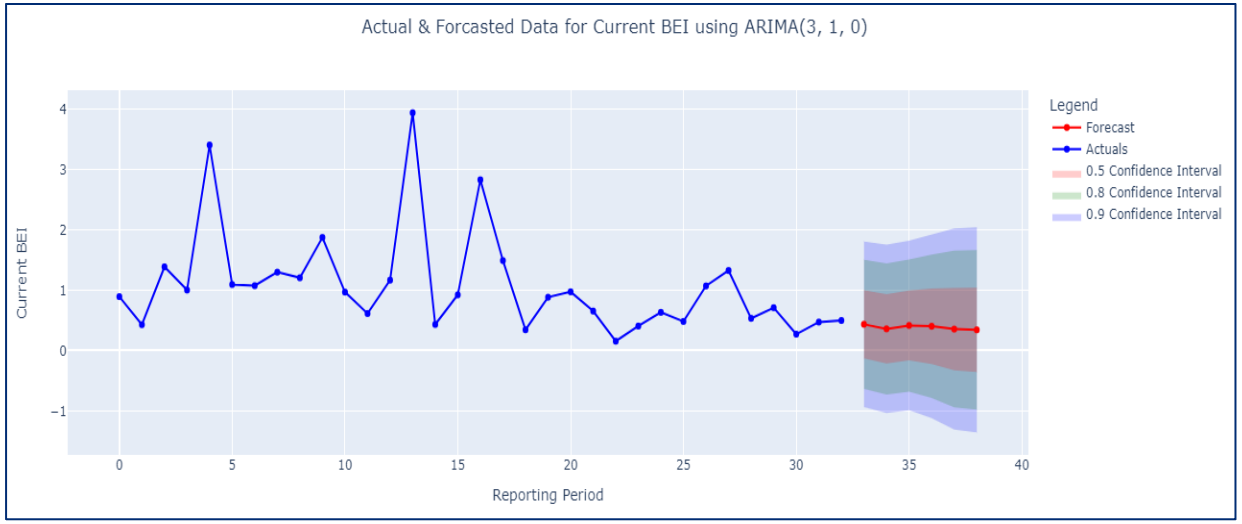
ARIMA Variables of Interest

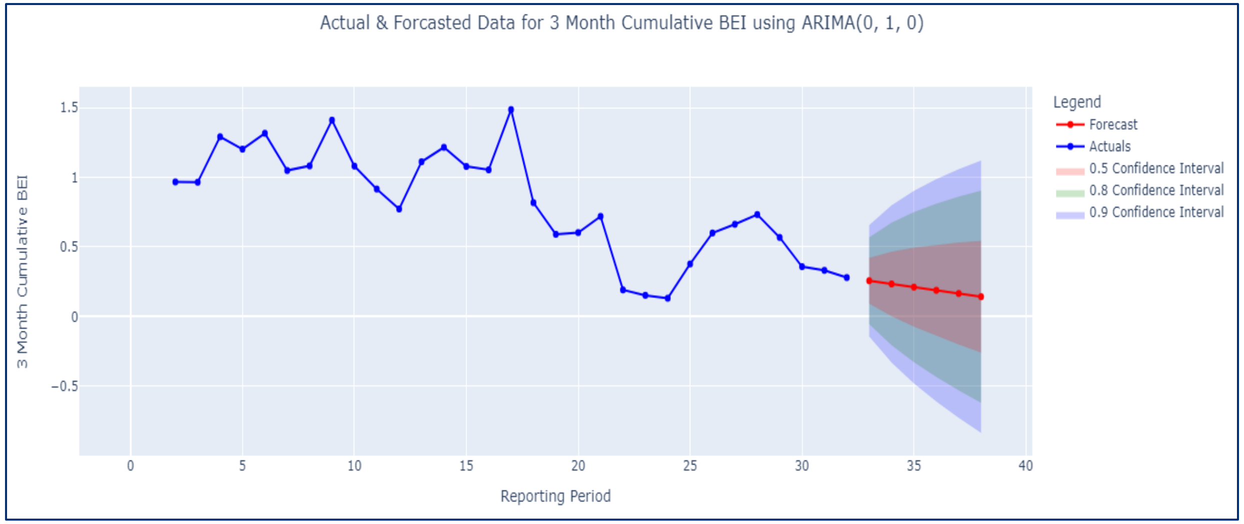
- Focused analysis on several standard and emergent acquisition metrics
  - **Cost Performance Index (CPI)**: method to estimate cost efficiency over time
  - **Schedule Performance Index (SPI)**: method to estimate schedule efficiency over time
  - **Current Baseline Realism Index (BRI)**: indicates how well the contractor is following the plan within reporting period.
  - **Current Baseline Progress Index (BPI)**: indicates how many of the planned events within reporting period have been accomplished.
  - **Current Baseline Execution Index (BEI)**: percentage of total events finished in current reporting period.
- ARIMA models were fit using Integrated Program Management Data and Analysis Report (IPMDAR) cost and schedule data from a major DoD acquisition program
- ARIMA models were fit using the metrics independently over various timeframes
  - SEM BPI and BEI: current/non-cumulative, 3 reporting period cumulative, and 6 reporting period cumulative
  - SEM BRI: 6 reporting period moving average
  - CPI and SPI: rolling cumulative
- The specific ARIMA(p, d, q) values for each model were chosen using Auto_ARIMA to optimize the best model. This is done using the Python package *pmdarima*.

Forecasting using ARIMA

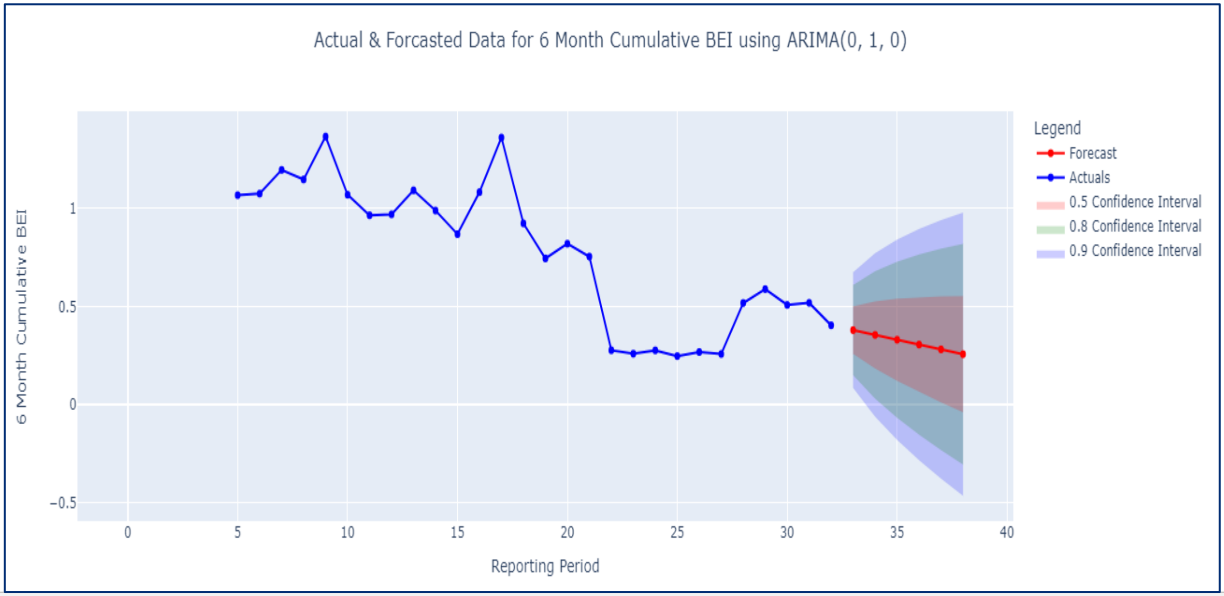
- Using the ARIMA models, metrics were forecasted for six future reporting periods that are based on the historical metric values.
- Followed the best practice of separating our datasets into test and validation subsets. The test data makes up 80% of the original dataset and is used to create models to explain the data available. The remaining 20% of the data is put into a validation subset, which is used to verify the models chosen using the test data set.
  - Time series analysis requires a continuous time series to use as the independent variable, so the data was separated based on the first 80% in time as the test dataset and the last 20% in time as the validation dataset.
  - After validating that the models adequately explain the data, they are applied to the entire dataset.
- Confidence intervals were added to each metric's forecast window. Confidence intervals are added to account for the level of uncertainty inherent in any modeling forecast. The confidence intervals display what range the actual results could have.
  - 90% confidence interval: most conservative and largest range
  - 80% confidence interval: mid point
  - 50% confidence interval: least conservative and smallest range

Schedule Performance Forecast – Current Baseline Execution Index (BEI) Example





- SEMs framework includes three models per metric
    - Current/non-cumulative
    - 3 reporting period cumulative
    - 6 reporting period cumulative

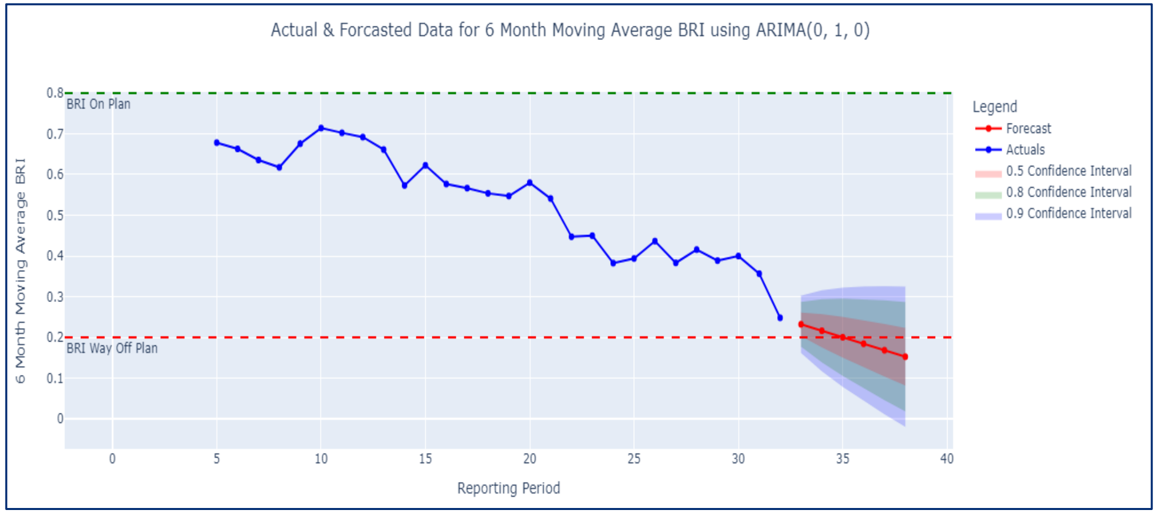


BEI is the percentage of total events finished in current reporting period.

Compare Schedule Execution Metrics: BRI and BPI

- **Current BRI** indicates how well the contractor is following the plan within a reporting period. BRI is applied using a 6 reporting period moving average. (NRO, 2021)
  - $Current\ BRI = \frac{\sum of\ tasks\ baselined\ to\ finish\ in\ RP\ \&\ completed\ in\ RP}{\sum tasks\ baselined\ to\ finish\ in\ RP}$
  - Each month the percentage of completed tasks that were baselined to execute is reduced; the schedule issues are forecasted to persist despite the recent improvement

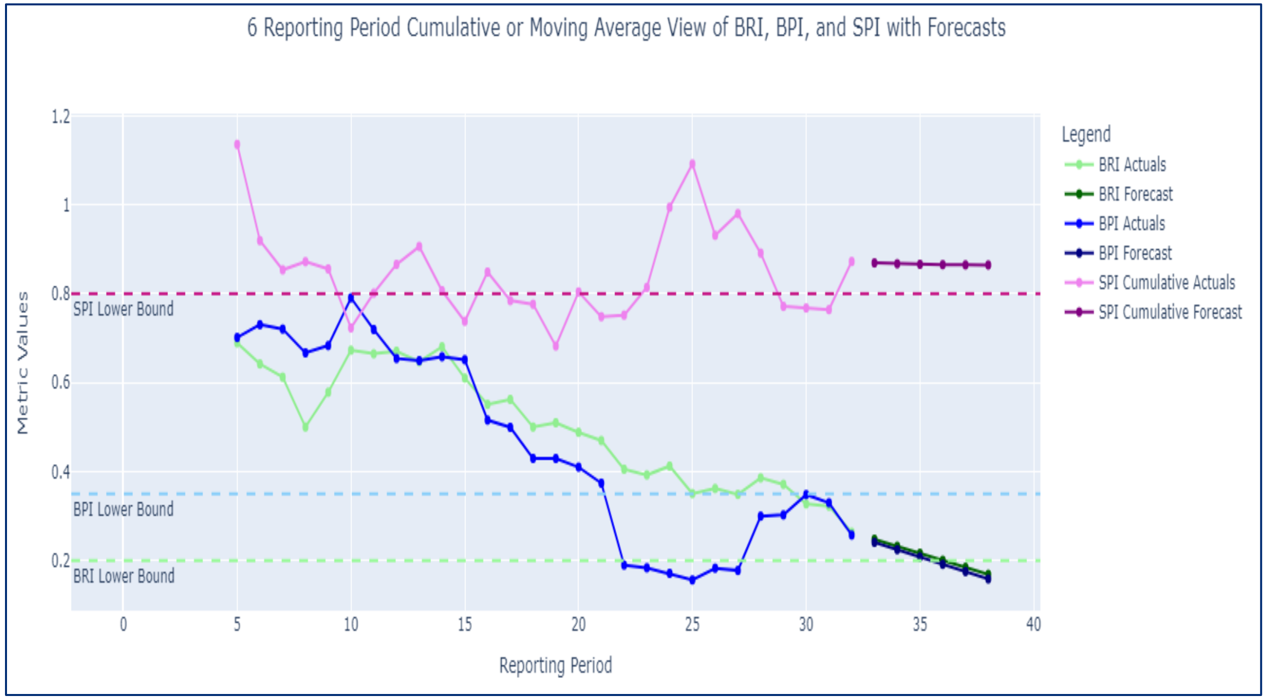
- **Current BPI** indicates how many of the planned events within a reporting period have been accomplished in or before the current period. (NRO, 2021)
  - $Current\ BPI = \frac{\sum of\ tasks\ baselined\ to\ finish\ in\ RP\ \&\ completed\ in\ or\ before\ RP}{\sum tasks\ baselined\ to\ finish\ in\ RP}$
  - Similar to BRI, the percentage of completed tasks that were baselined to be executed leading up to today reduces; the schedule issues are forecasted to persist.



BRI forecast high schedule risk as of reporting period 35, BPI reflects high schedule risk as of reporting period 22

How does SPI compare to BRI, and BPI?

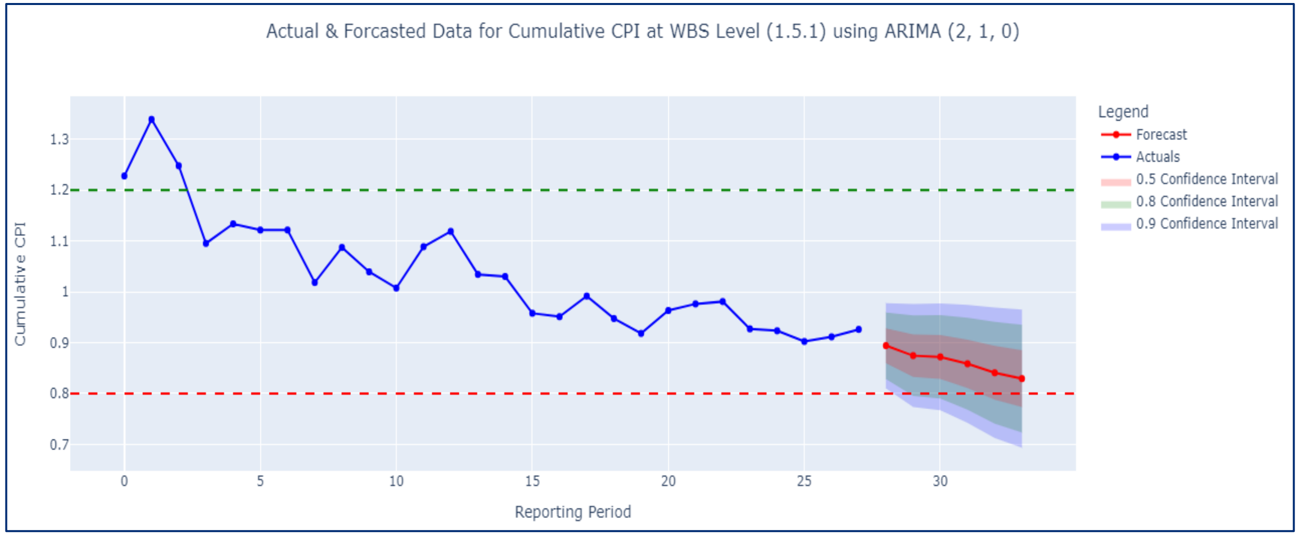
- SPI and BPI are shown using the 6 reporting period cumulative view. BRI is shown using the 6 reporting period moving average.
- BPI detects the program to be “Way Off Plan” starting in reporting period 22
- BRI forecasts the program to be ““Way Off Plan” starting in reporting period 35
- SPI occasionally drops below the lower threshold indicating there is some schedule risk
- In this example, the SEMs are more responsive and would allow for program personnel more time to address schedule issues before they became a critical issue

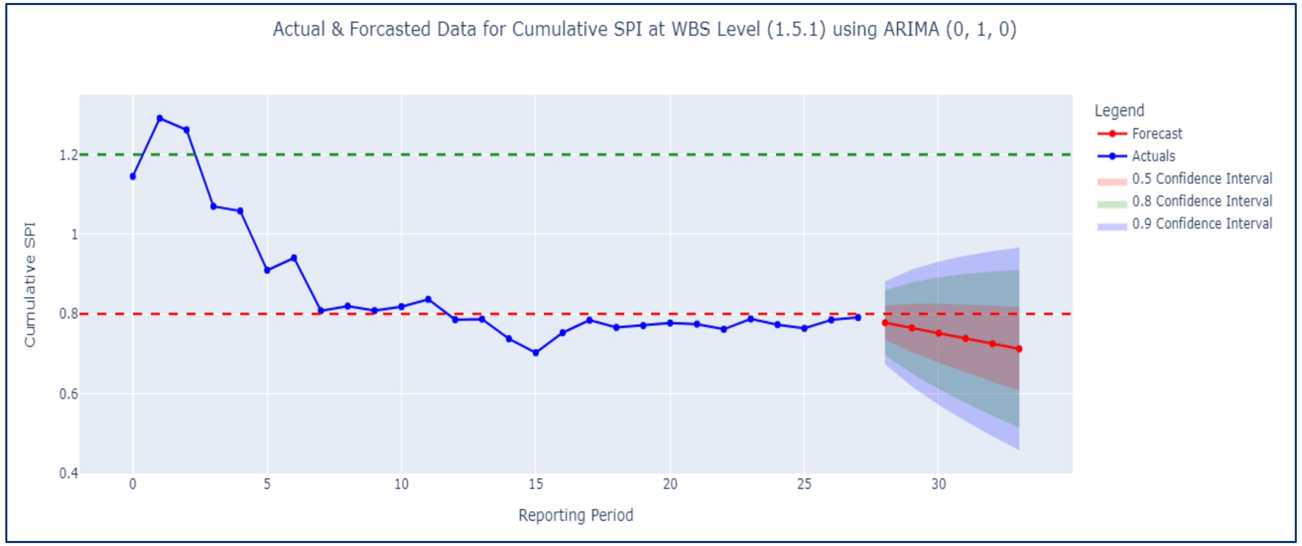


Implementing multiple metrics allows for more comprehensive program schedule analysis

Acquisition Performance Forecasting Fidelity

- ARIMA models are fit at varying work breakdown structure (WBS) levels
  - Total Contract
  - 2nd Level
  - 3rd Level
  - 4th Level
- Higher fidelity analysis limited due to noise (e.g. accounting revisions, missing data) that precludes meaningful results
- Drill down and Roll up features allow for root cause analysis by account managers.
- Models are refined and updated each reporting period.
- Example visualizations are using WBS Element 1.5.1: Developmental Test and Evaluation





Time series models enable insights into what portion of the system will drive cost & schedule

Independent Estimate at Complete Forecasting

- Forecasted metrics are also used to come up with independent estimates of a total contract cost (e.g. IEAC)
- Building on traditional descriptive IEAC metrics we have derived an IEAC metric that incorporates the results of the CPI and SPI ARIMA forecasts
- Forecasted IEAC

$$IEAC_t = ACWP_{cum} + \sum_{j=t+1}^{t+6} \left(\frac{\widehat{BCWS}_j * \widehat{SPI}_j}{\widehat{CPI}_j} \right) + \frac{BAC_t - BCWP_{cum} - \sum_{j=t+1}^{t+6} \widehat{BCWP}_j}{CPI_{cum} * SPI_{cum}}$$

$$where\ t = time, ACWP_{cum} = \left(\sum_{i=1}^{t-1} acwp_i \right), BCWP_{cum} = \left(\sum_{i=1}^{t-1} bcwp_i \right),$$

$$CPI_{cum} = \left(\sum_{i=1}^{t-1} \frac{bcwp_i}{acwp_i} \right), SPI_{cum} = \left(\sum_{i=1}^{t-1} \frac{bcwp_i}{bcws_i} \right)$$

$$\widehat{BCWS}_j = BCWS\ to\ complete\ within\ forecast\ window\ of\ ARIMA\ models,$$

$$\widehat{SPI}_j = forecasted\ value\ of\ SPI\ based\ on\ ARIMA\ models,$$
$$\widehat{CPI}_j = forecasted\ value\ of\ CPI\ based\ on\ ARIMA\ models$$

Third Level WBS Elements

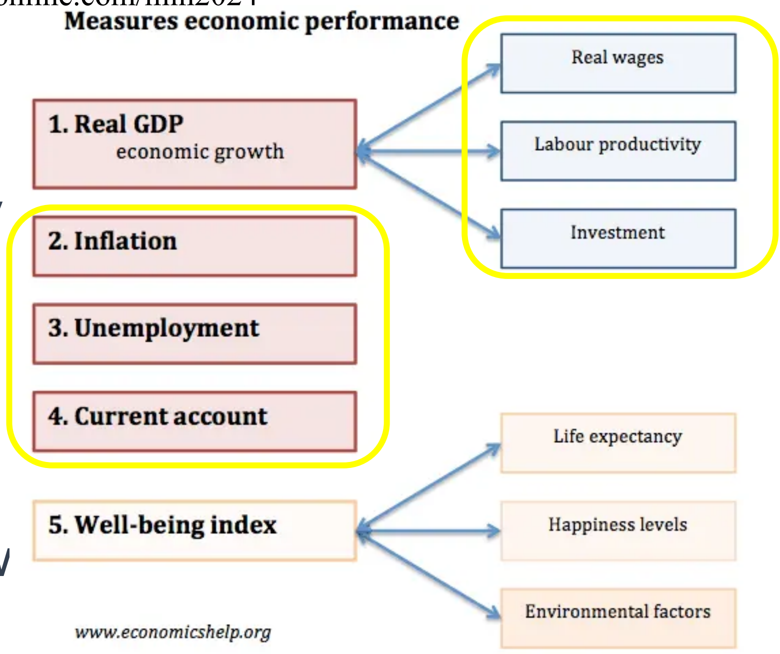
WBS Element	Name	IEAC	BAC	% Difference
1.5.1	Development Test and Evaluation	$ 170,182.95	$ 240,910.33	-29%
1.5.5	Test and Evaluation Support	$ 968,085.80	$ 497,392.23	95%
1.9.1	Test and Measurement Equipment	$ 170,801.37	$ 86,009.77	99%
1.9.2	Support and Handling Equipment	$ 106,128.77	$ 88,494.13	20%
1.10.2	Contractor Technical Support	$ 245,398.86	$ 230,503.80	6%

Fourth Level WBS Elements

WBS Element	Name	IEAC	BAC	% Difference
1.1.1.2	Hull/Frame/Body/Cab	$64,172.00	$27,093.56	137%
1.1.1.6	Vehicle Electronics	$227,918.56	$180,300.75	26%
1.1.1.13	Special Equipment	$513,538.24	$318,294.10	61%
1.8.1.3	Test and Measurement Equipment (Electronics/Avionics)	$200,636.94	$105,108.83	91%
1.9.2.1	Support and Handling Equipment (Airframe/Hull/Vehicle)	$92,069.64	$16,009.64	475%

Tracking the BAC and IEAC can help balance work priorities, re-plan remaining tasks, and adjust the technical approach to complete the project within the remaining resources.

Macroeconomic Regressions



- Purpose: Examine whether economic indicators that have a historically significant relationship with business cycle changes in gross output, possess a comparable relationship to DoD acquisition performance
- Example: Okun's Law measures negative linear association between cyclical unemployment and the output gap, implying that changes in unemployment will behave countercyclically with economic output grow (e.g. ↓ cyclical unemployment → ↑ output)

$$\Delta output\ gap = \alpha + \beta(\Delta[actual\ unemp\ rate\ - natural\ unemp\ rate]) + \in$$

Where alpha is a constant that represents the y-intercept, and epsilon is the error term

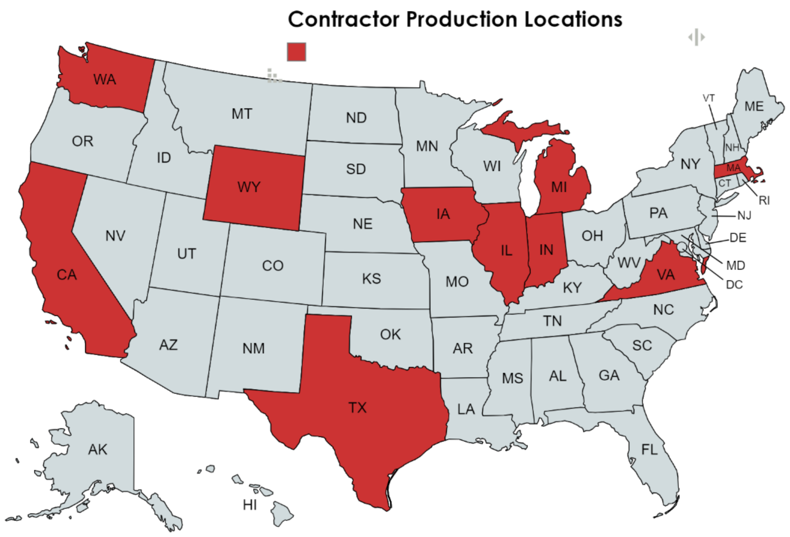
- Rationale:
  - External program factors, such as changes in broader economic conditions or a firm's supplier base, can impact program stability and lead to acquisition performance issues (Blickstein et al., 2012, p. 10; Arena, 2014, p. 28).
  - Mature, structured data sets for this information are available from the Bureau of Labor Statistics, Bureau of Economic Analysis, the Census Bureau, and Federal Reserve

BEA BUREAU OF ECONOMIC ANALYSIS U.S. DEPARTMENT OF COMMERCE
BLS BUREAU OF LABOR STATISTICS U.S. DEPARTMENT OF LABOR
United States Census Bureau
FRED ECONOMIC DATA | ST. LOUIS FED

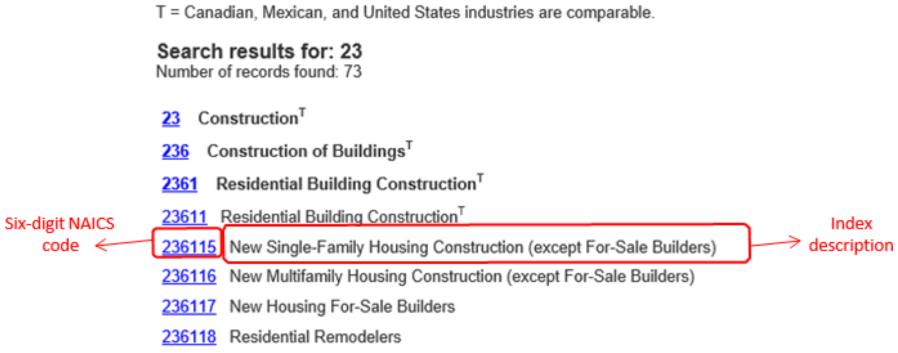
Macroeconomic Regressions

- Methodology: Execute linear regression with varying time lags to examine the relationship between economic indicators and acquisition performance
- Levels of Analysis
  - Geographic Location
    - National
    - State
    - County Level
  - Production Industry
    - 6 digit NAICS*
    - 4 digit NAICS
    - 2 digit NAICS

Variables of Interest	Independent Variables
CPI	Unemployment
SPI	Inflation
IEAC	Labor Productivity
	Real Wages



***North American Industry Classification System**


T = Canadian, Mexican, and United States industries are comparable.

Search results for: 23
Number of records found: 73

23 Construction[T]
236 Construction of Buildings[T]
2361 Residential Building Construction[T]
23611 Residential Building Construction[T]
236115 New Single-Family Housing Construction (except For-Sale Builders)
236116 New Multifamily Housing Construction (except For-Sale Builders)
236117 New Housing For-Sale Builders
236118 Residential Remodelers

Examples:

$$EVM\ metric_{cumulative} = \beta_{Intercept} + \beta_1 Unemployment\_rate_{lag=t} + \varepsilon,\ where\ i = county$$

$$EVM\ metric_{cumulative} = \beta_{Intercept} + \beta_1 PPI_{i,lag=t} + \varepsilon, where\ i = industry$$

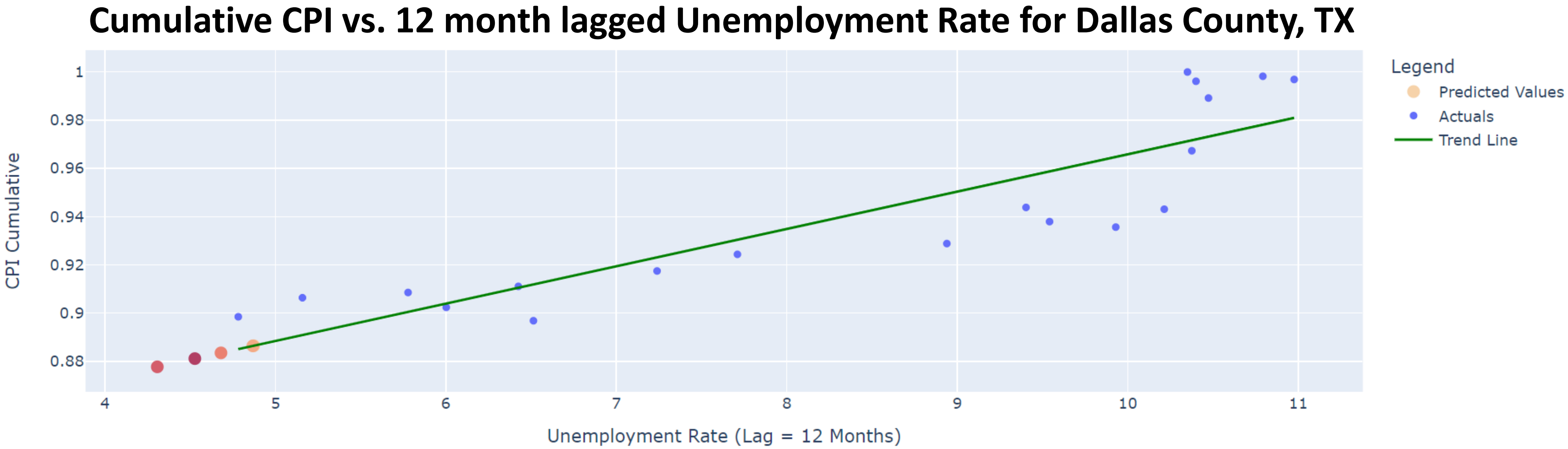
Macroeconomic Methodology

1. Collect monthly unemployment rates at county and national level from the BLS
2. Identify contractors and subcontractor demographic information by WBS in IPMDARs using Defense Logistics Agency (DLA) commercial and government entity (CAGE) codes and unique entity identifiers (UEIs)
3. Create new contractor table to include:
    1. location information down to zip code level, which is provided by CAGE codes and UEIs
    2. Supporting information, like the companies primary North American Industry Classification System (NAICS) codes
4. Map macroeconomic data to work locations and supporting information
5. Aggregate ACWP, BCWS, and BCWP at county level, then calculate CPI and SPI for each reporting period
6. Execute regressions and choose time series lag that provides best fitting linear regression model results at the analysis level of interest

Data Normalization Process allows for both cross sectional and longitudinal analysis

Macroeconomic Regressions

- Notional Results
  - Lagged unemployment is positively and significantly associated with cumulative CPI values for area of interest (Dallas Cty, TX)
  - Potential Interpretation: A “tight” labor market, exacerbated by limited labor supply with the requisite human capital to manufacture DOD systems, leads to worse acquisition performance.



- Key takeaway: IPMDAR affords the integration of geographic and industry specific macroeconomic information with program EVM data to examine how external program factors may impact the program both broadly and:
  - At key geographic areas of design and manufacturing
  - For specific industrial sectors
- The analytic value is likely substantial given current industrial base and supply chain considerations

Path Forward

- Further integration and exploration of how economic factors may influence defense acquisition performance
- Continue to expand portfolio of IPMDAR data through direct support of MDAPs and coordination with OSD
- Test and train the time series and macroeconomic models on the broader IPMDAR portfolio to:
  - Identify and verify best fitting analytics
  - Establish classifiers (i.e. objective and threshold values) to help inform decision makers when opportunities or risks are imminent so that preventive action can be taken
- Evaluate additional data science and statistical techniques
  - Panel regressions, which combine cross-sectional and time-series data, to control for unobserved dependencies and endogeneity. Potential research purposes:
    - Identify if work performance in a specific region of the U.S. is impacting one, or many, projects over time.
    - Compare acquisition performance by service, command, program size, prime system integrator, etc.
  - Clustering and classification techniques
    - K-nearest neighbors (K-nn): supervised classifier that minimizes distance of items from a centroid of measurements
      - Example: Classify the completion risk of activities over time, based on variables of interest (e.g. budgeted work remaining, active months)
    - Logistic regression: Probabilistic model of an event (e.g. acquisition issue or opportunity) taking place
      - Example: correlate cost and schedule variables of interest (e.g. average CPI or task mid-point SPI) with total cost outcomes (e.g. over budget or under budget) at varying WBS levels

Conclusions

- The study was constrained by several limitations, most notably data accessibility:
  - Current analysis is executed on an on-going program, rather than having completed projects, from which to assess the overall level of improvement provided by time series analysis. Nonetheless, forecasts have been timely in alerting the program of directional trends in cost and schedule health relative to standard EVM.
  - To conduct ARIMA analysis, multiple years of cost and schedule data (without re-baselines) is required to create a properly sized consecutive time series dataset
- Nevertheless, the methodological approach may offer several advantages over current EVM practices:
  - Time series forecasts on standard EVM data and the SEMs may be more responsive to trends in project performance relative to descriptive statistics, which may subsequently allow the analyst more time to identify and decompose potential cost and schedule risks, before they manifest into issues.
  - While traditional time series analysis has historically been conducted on a portfolio of completed programs or projects, this framework is set up to train, test, and execute analysis in real time on current projects.
    - Conducting time series analysis in real time, and at varying WBS levels allows for root cause analysis. Application creates an opportunity to prevent potential cost and schedule issues from emerging.
  - Finally, this analysis accounts for trends in external program factors, which can be a key reason for a cost or schedule breach that may impact program performance

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

References (1/2)

- Alexander, S. (2002). *Earned Value Management Systems (EVMS)*.
- Arena, Mark V., et al., *Management Perspectives Pertaining to Root Cause Analyses of Nunn-McCurdy Breaches: Contractor Motivations and Anticipating Breaches*, Vol. 6, Santa Monica, Calif.: RAND Corporation. MG-1171/6, 2014. As of January 27, 2019: https://www. rand. org/pubs/monographs/MG1171z6. html, 2014.
- Anton, Philip S., William Shelton, James Ryseff, Samantha Cohen, Grant Johnson, Stephen B. Joplin, David Kravitz, Megan McKernan, and Alejandro Vigo Camargo. "*Early Predictive Indicators of Contractor Performance: A Data-Analytic Approach.*" RAND Corporation, 15-42, 2022.
- Arena, Mark V., et al. *Management Perspectives Pertaining to Root Cause Analyses of Nunn-McCurdy Breaches: Contractor Motivations and Anticipating Breaches, Vol. 6, Santa Monica, Calif.: RAND Corporation*. MG-1171/6, 2014. As of January 27, 2019: https://www. rand. org/pubs/monographs/MG1171z6. html, 2014.
- Austin W Dowling, *Using Predictive Analytics to Detect Major Problems in Department of Defense Acquisition Programs*, Department of the Air Force Air University, Air Force Institute of Technology, March 2012.
- Blickstein, Irv, et al. "Methodologies in analyzing the root causes of Nunn-McCurdy breaches." (2012).
- Cryer, Jonathan D., and Kung-Sik Chan. Time series regression models. *Time series analysis: with applications in R*: 249-276, 2008.
- Data Item Description (DID) for Integrated Program Management Data and Analysis Report (IPMDAR), Number: DI-MGMT-81861C, Approval Date: 20210830.
- Department of Defense, Distribution Statements on DOD Technical Information., DOD Instruction 5230.24, Office of the Under Secretary of Defense for Research and Engineering, January 10, 2023.
- Department of Defense, Department of Defense Standard Practice Work Breakdown Structures for Defense Materiel Items, MIL-STD-881F, May 13, 2022.

References (2/2)

- “Earned Value Management System (EVMS) and Project Analysis Standard Operating Procedure (EPASOP)”, Office of Project Management, DOE-PM-SOP-05-2020, January 14 2020.
- GAO - United States Governance Accountability Office. GAO-16-89G Schedule Assessment Guide, 2015
- Hofbauer, Joachim, et al. Cost and time overruns for major defense acquisition programs. Acquisition Research Program, 2011. Need to find full reference information
- Johnson, J. C., “Cost Overruns, Schedule Slips: The Norm Not the Exception.” National Defense Magazine. May 1, 2018. https://www.nationaldefensemagazine.org/articles/2018/5/1/cost-overruns-schedule-slips-the-norm-not-the-exception
- Keaton, C. G., White, E. D., & Unger, E. J., “Using earned value data to detect potential problems in acquisition contracts”. *Journal of Cost Analysis and Parametrics, 4(2),* 148-159, 2011.
- Kwiatkowski, Denis, et al. “Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?." Journal of econometrics 54.1-3 (1992): 159-178, 1992.
- Montgomery, Douglas C., Cheryl L. Jennings, and Murat Kulahci. *Introduction to time series analysis and forecasting*. John Wiley & Sons, 2015.
- National Reconnaissance Office (NRO), “Advanced Schedule Analysis Techniques”, PowerPoint presentation, May 2021.
- National Reconnaissance Office (NRO), “Predictive Analytics: Schedule Execution Metrics”, PowerPoint presentation, September 2022. https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2022/ipmdFall/Bembers.pdf
- NDIA (National Defense Industry Association). "A guide to managing programs using predictive measures." (2014).
- Python API reference for plotly. https://plotly.com/python-api-reference/. Accessed May 2023.
- Schultheis, Candice, Ryan Carey, Robert Dunn. “An Analysis of Schedule Execution Metrics (SEMs) Developed by the National Reconnaissance Office (NRO)”, Naval Post Graduate School Thesis, March 2021.

Author Biographies

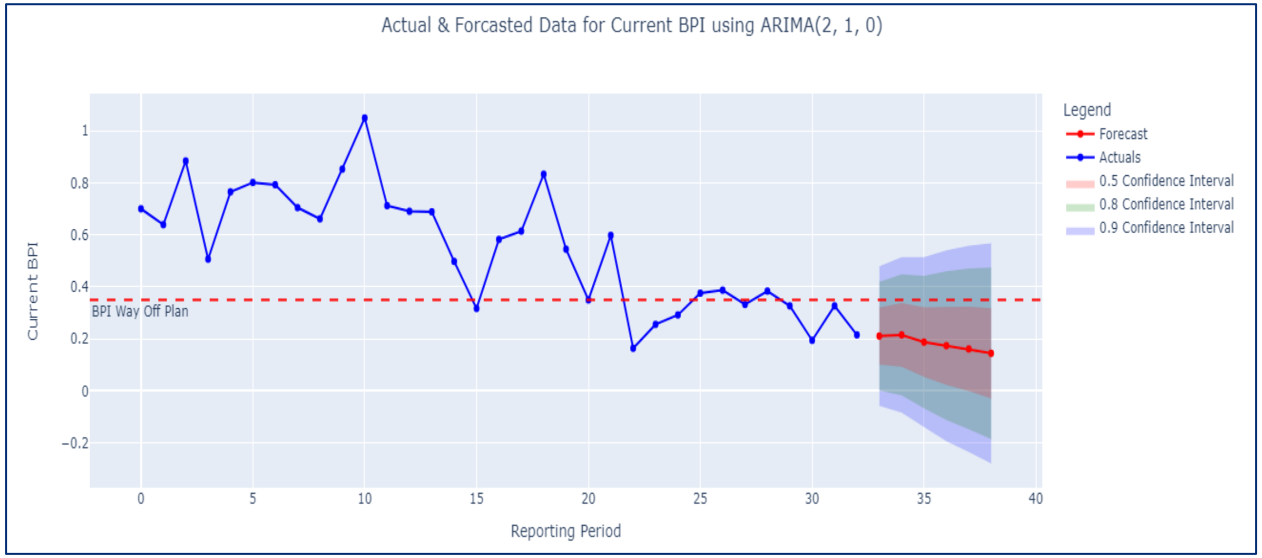
- Anna B. Peters is a cost analyst and data scientist at Johns Hopkins University Applied Physics Lab, contributing to a team that specializes in affordability analysis and assessments for a variety of Department of Defense sponsors. She holds a M.S. in Applied Statistics from Bowling Green State University and a B.S. in Industrial Math and Statistics from West Virginia University.

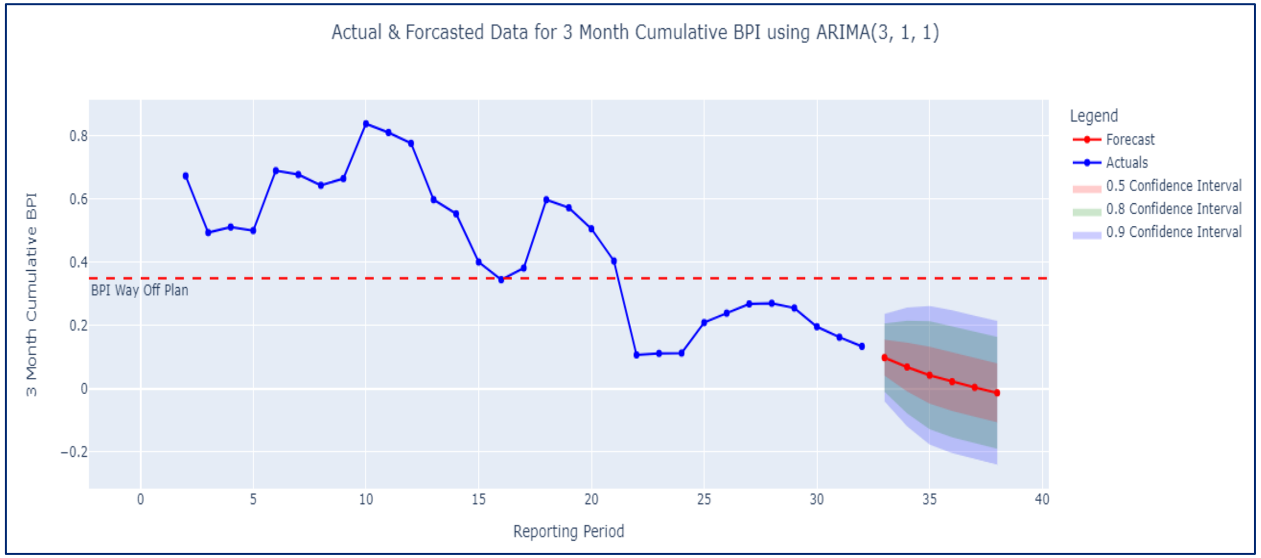
- Mark W. Hodgins is a member of the Principal Professional Staff at JHU APL, where he leads investment decision analyses for the DOD. Mark specializes in affordability assessments for strategic missile systems, mine countermeasures, and cybersecurity. He earned a Ph.D. in Economics from George Mason University, a graduate certificate in Business Analytics from Indiana University Bloomington, and a M.A./B.A. in Economics from Johns Hopkins University. He holds certifications as a CCEA from ICEAA and a PMP from PMI.

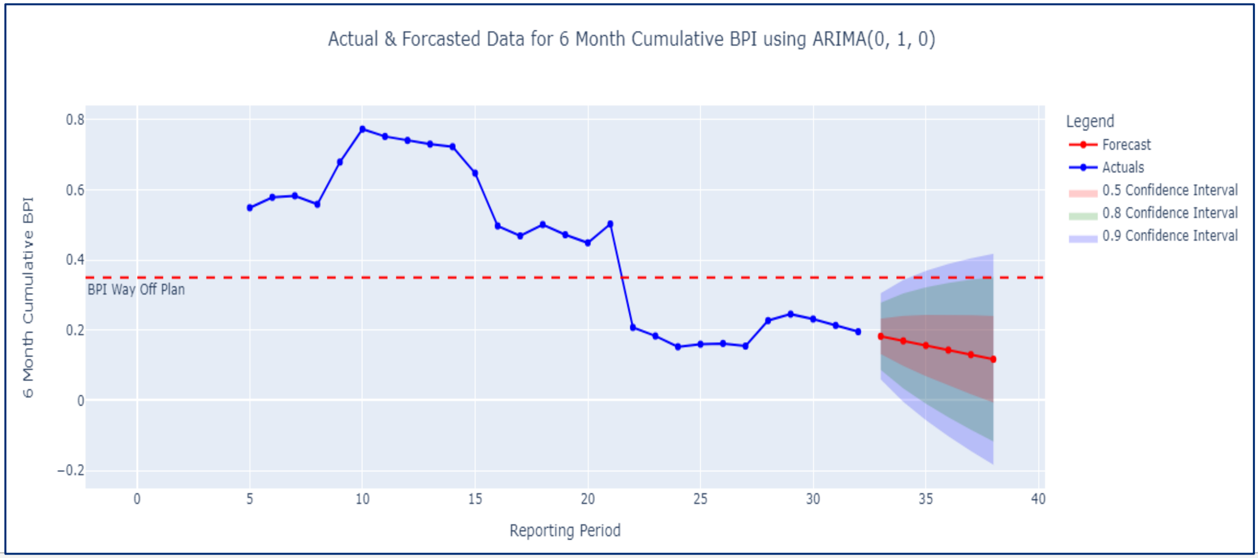
NRO Metric Equations

- $Current\ BRI = \frac{\sum of\ tasks\ baselined\ to\ finish\ in\ RP\ that\ were\ completed\ in\ RP}{\sum tasks\ baselined\ to\ finish\ in\ RP},$
  - Note: Current BRI is always applied as a six reporting period moving average
- $Current\ BPI = \frac{\sum of\ tasks\ baselined\ to\ finish\ in\ RP\ that\ were\ completed\ in\ or\ before\ RP}{\sum tasks\ baselined\ to\ finish\ in\ RP}$
- $Current\ BEI = \frac{\sum of\ tasks\ finished\ in\ RP\ even\ if\ not\ originally\ planned\ to\ finish\ in\ RP}{\sum tasks\ baselined\ to\ finish\ in\ RP}$
- $where\ RP = reporting\ period$

Schedule Performance Forecast - Baseline Progress Index (BPI) Example





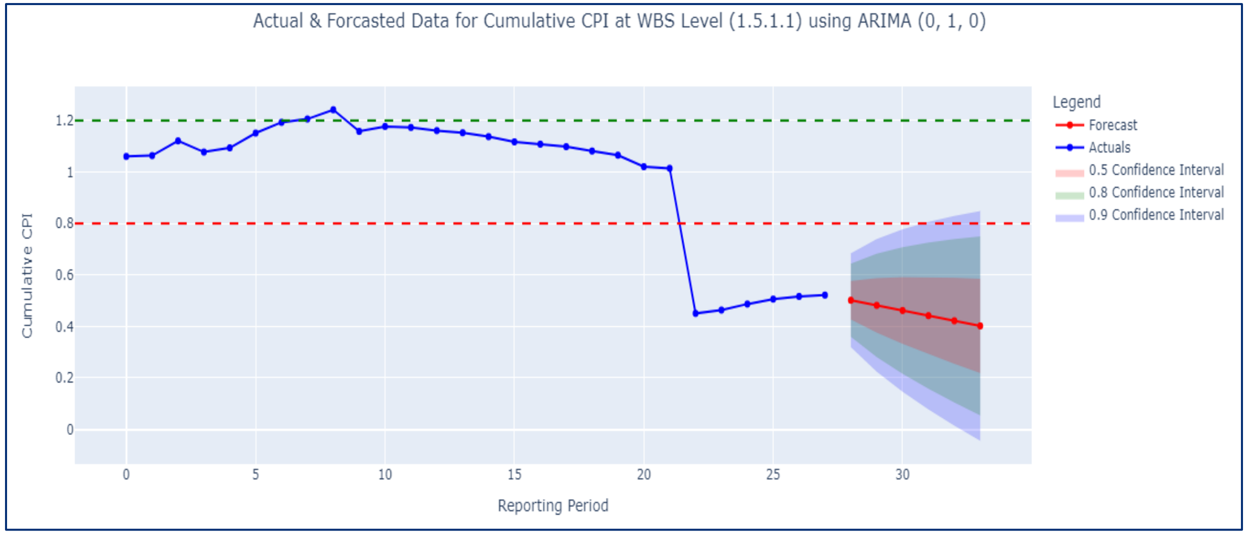


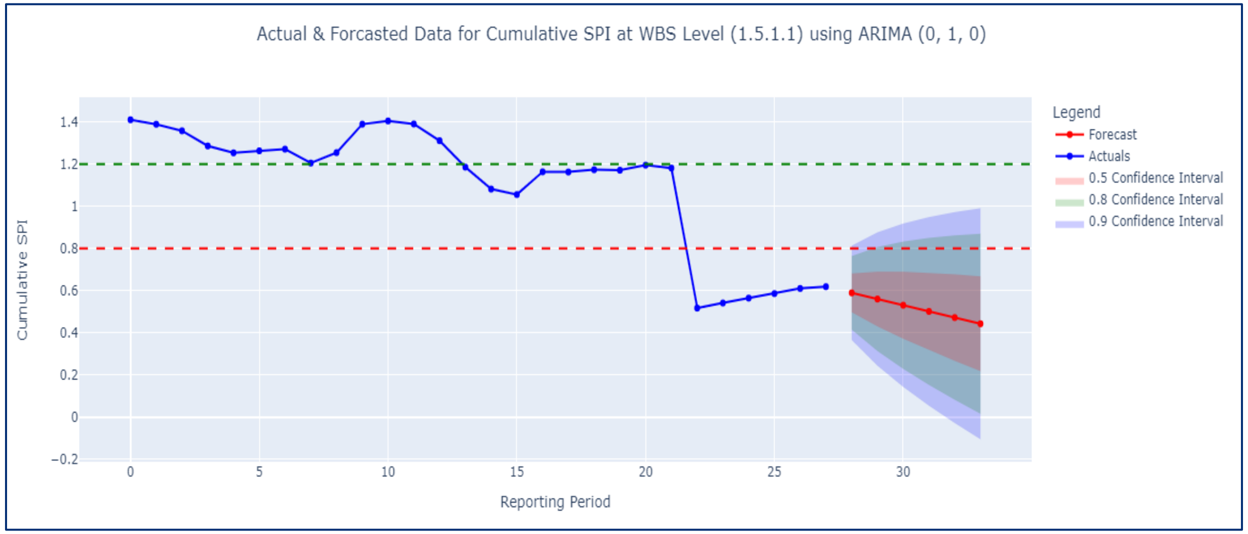
BPI indicates how many of the planned events within reporting period have been accomplished.

- ≤ 0.35 is unfavorable

Acquisition Performance Forecasting Fidelity: 4^{th} Level

- ARIMA models are fit at varying WBS levels
  - 4^{th} Level
- Example visualizations are using WBS Element 1.5.1.1: Cybersecurity Test and Evaluation





Python Packages

Python Library	Purpose
ARIMA, from statsmodels.tsa.arima.model	ARIMA time series modeling
auto_arima, from pmdarima.arima	Time series analysis, specifically auto_arima
dateutil, from relativedelta	Managing date-time formatted data
express, from plotly	Data visualization
graph_objects, from plotly	Data visualization
matplotlib.pyplot	Creating object-oriented plots
numpy	Scientific computing
pandas	Building and manipulating data structures
product, from itertools	Creating iterators for efficient looping
seaborn	Creating statistical data visualizations