MIT LGO

LEADERS FOR GLOBAL OPERATIONS

# LGO Class of 2025
# Knowledge Review Project Briefs

LEADERS FOR GLOBAL OPERATIONS

| First Name | Last name | Title | Host Partner Company |
|---|---|---|---|
| Shreeansh | Agrawal | Machine Learning Methods in Telecom | |
| Jack | Albright | Computer Vision for Cell Line Development | Amgen Inc. |
| Priya | Bhirgoo | Assessing the Feasibility of a Fully Electric Drug Substance Facility | Amgen Inc. |
| Linn | Bieske | Sensor simulation for autonomous vehicles: Diffusion based image and depth generation for driving scenes | Waymo, LLC |
| Conor | Briggi | The Value of Digitizing Manufacturing Environments | Stanley Black & Decker (SBD) |
| Alix | Carson | A Data-Driven Work Center Assignment and Pricing Strategy for a Job Shop | |
| Julie | Chong | Raw Material Inventory Planning | |
| Isabella | Didio | Impact Evaluation and Prioritization Framework for Manufacturing Inspection Technology Investment | Johnson & Johnson |
| Madeline | Dubelier | Systems Approach to Component Code Optimization for Wound Closure Portfolio | Johnson & Johnson |
| Andrew | Dugan | Fully Connected Digital Ecosystems within Hospitals - AI/ML Solutions for Improved Patient Care | Johnson & Johnson |
| Andrew | Epstein | Decarbonization of Gas Heating in Massachusetts: An Evaluation of Current Trends and Opportunities | National Grid |
| Andrew | Fenstermacher | Investigation Into Sources of Volatility in Sortation Center Processes to Improve Productivity and On-Time Delivery | Target |
| Gabriel | Gallardo | Transforming Unstructured Text into Actionable Insights: A Use Case of Generative AI in Operations Technology Problem Management | Amazon |
| Jeremy | Garber | Minimizing Cost of Intra-Yard Finished Vehicle Logistics Through Automation and Optimization | |
| Catalina | Garza Lozano | Predictive Model for Battery State of Health | NextEra |
| Adam | Gebner | Optimizing Raw Wire Inventory Management: A Data-Driven Approach to Demand Forecasting and Supply Chain Decision Support | Boeing |
| Jacob | Gerbino | Economies of Space: Developing a Lean Manufacturing Framework for Work Center Floorspace Reduction | Boeing |
| Viraat | Goel | Simulation Modeling of Drug Substance Tech Transfer Timelines at Amgen | Amgen Inc. |
| Yutao | Gong | Forecasting Automotive Production Using Theta Model | American Industrial Partners |
| Gretel | Gonzalez | Expanding Verizon's home broadband coverage through existing Low Earth Orbit mega-constellations | Verizon |
| Carlos | Gosen | Developing a Data-Driven Approach to Reducing Excess Inventory in a Multi-Echelon Supply Chain | Stryker |

| First Name | Last name | Title | Host Partner Company |
|---|---|---|---|
| Rachael | Harkavy | Forming the Future: A Digital Approach to Simulating Thermoplastic Manufacturing | Re:Build Manufacturing |
| Tony | Jiang | Research on Multi-level Framework for Video Understanding | Amazon |
| Chris | Johnson | Optimizing automotive production scheduling to reduce finished vehicle inventory | |
| Mohit | Kasliwal | An Integrated Optimization Model for Large-Scale EV Fleet Deployment: Balancing Emissions Reduction and Operational Costs | Verizon |
| Rachael | Knapp | Electric Vehicle Fleet Charging: A Simulation-Based Comparison of Charging Strategies and Cost Implications | |
| Carla | Lorente Anon | Multimodal Gen AI Chatbot for Root Cause Diagnosis in Predictive Maintenance | Amazon |
| Amna | Magzoub | Mapping Design Transfer to Accelerate New Product Development | Stryker |
| Bonny | Mahajan | Generative AI in Private Equity for Accumulative Advantage | LFM Capital |
| Maxwell | Malinowski | Data, Analytics, and Optimization for Production Planning | Northrop Grumman Corporation (NGC) |
| Will | McNulty | Standard Work for High-Mix Low-Volume Operations | Re:Build Manufacturing |
| Baraka | Minja | Design & Optimization of Shipping Container for Package-Less Units | Amazon |
| Sofie | Netteberg | Toward Smarter New Product Assignment: Optimizing Network Flexibility Using Demand Variability Insights | Nike |
| Michael | Norwalk | Decarbonized Cement Manufacturing via Advanced Production Technologies | NextEra |
| Olanrewaju | Oludipe | Optimizing Inventory Rebalancing: Strategies for Managing Excess Inventory in a Dynamic Supply Chain | Nike |
| Haoting | Pan | Analyzing Procurement Data for Cost Saving Application | Caterpillar |
| Karla | Perez Muñoz | An Optimization-Based Approach to Efficient Clearance Inventory Allocation | Zara (Inditex, S.A.) |
| Esteban | Ramirez Echavarria | Discrete Event Simulation as a Predictor for Factory Traffic Management | |
| Jennifer | Ray | Hydrogen Adoption Dynamics: A Flexible Modeling Framework for U.S. Industrial Applications | |
| Lizzy | Salata | Streamlining Diagnostics of Electrical-Connection-Related Defects in General Assembly Using Augmented Reality Wearables | Rivian |
| Kevin | Schurr | Towards Green Aluminum | |

| First Name | Last name | Title | Host Partner Company |
|---|---|---|---|
| Shweta | Sen | Multi-Objective Optimization of Container Load Plans for Modulating Inventory Flow | |
| M. Patrick | Serbent | Network Preparations for Networked Geothermal | National Grid |
| Sameed | Siddiqui | Advanced Architectures for Biological Sequence Modeling | Broad Institute |
| Julia | Sircar | Defect Management and Prevention Techniques to Improve Investment Casting Throughput | Blue Origin |
| Alex | Sirgo | A Techno-Economic Assessment of Hybrid Renewable Energy and Battery Storage Systems for Data Centers | NextEra |
| Steffan | Sowards | Data-Driven Key Performance Indicator Modeling for Robotic Mobile Fulfillment Systems | Amazon |
| Daniel | Tuana | A Technoeconomic Model for Maritime Applications of Green Power Technologies | Caterpillar |
| Arun | Varma | Diagnostics in Additive Manufacturing Using Image-Based Machine Learning | Blue Origin |
| Aaron | Wubshet | Closing the gap: Evaluation of On Body Injectors (OBIs) technologies through patient centric design | Amgen Inc. |
| Tiffany | Xi | Metal Additive Manufacturing Capabilities for Footwear Prototyping and Product Creation | |
| Mark | Zachary | Driving Manufacturing Best Practices Using Multimodal AI | LFM Capital |

# Machine Learning Methods for Churn Prediction and Infrastructure Resilience

## BUSINESS PROBLEM

Telecom companies seek to reduce customer churn through both short-term marketing strategies and long-term network reliability. The current forecasting approaches can be enhanced to better inform operational decisions around customer retention, upgrade promotions, and infrastructure resilience planning. The goal is to accurately forecast churn and proactively mitigate infrastructure-related churn risk.

## APPROACH

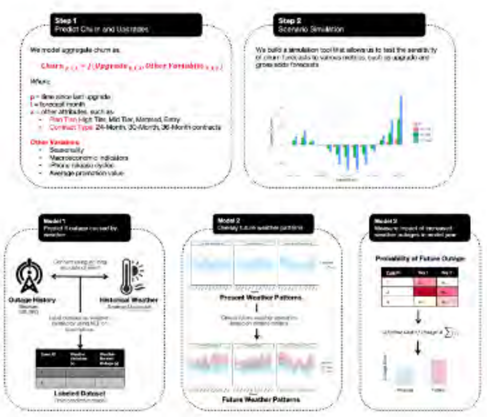
I created a cohort-based churn forecasting model integrating customer upgrade patterns, used synthetic data generation (MICE) for sparse cohorts, and applied interpretable linear regression with macroeconomic variables. For long-term resilience, I built predictive neural network models linking weather data to network outages and projected climate risk scenarios using clustering and simulations.

## DATA SOURCES

I used aggregated customer churn, upgrades, and contract tenure data provided by the company, alongside macroeconomic indicators from the Bureau of Labor Statistics (e.g., CPI). Infrastructure outage data came from the Company's internal systems, enriched with external weather data from Meteostat and future climate scenarios (CMIP5) via Openmeteo.

### Data Types and Format

Data included structured numerical customer records (monthly churn/upgrade rates, promo tenure), spatial-temporal outage incidents, qualitative incident descriptions, and hourly weather time series.




Author: Shreeansh Agrawal

## IMPACT

My churn model improves forecasting accuracy, enhancing operational efficiency and enabling targeted marketing, thereby optimizing marketing spend. Stakeholders gain clear, interpretable insights into churn drivers, fostering internal alignment between marketing and finance teams. Improved forecasting significantly impacts customer retention and financial performance. For infrastructure resilience, predictive modeling anticipates future climate-induced outages, enabling proactive planning and targeted hardening investments. By identifying and quantifying vulnerabilities through robust climate scenario modeling, the Company gains a competitive edge, mitigating long-term risk and ensuring reliable service delivery during extreme weather events. This proactive resilience strategy also enhances the Company's market reputation and customer trust, potentially improving market share in the long term. Together, these solutions position the Company as a leader in applied AI, integrating machine learning deeply into strategic decisions around customer retention and network resilience, which are critical for sustained market competitiveness and growth.

### DRIVERS

Increasing market commoditization and intense competition require the Company to prioritize customer retention via predictive insights. Concurrently, rising climate-related extreme weather events have exposed telecom infrastructure vulnerabilities. These industry pressures drive the strategic imperative to integrate AI-driven predictive analytics for immediate customer retention and long-term resilience planning.

### BARRIERS

Key barriers included data sparsity for newer contract types, requiring synthetic data techniques. Organizational dynamics between various teams posed challenges, demanding significant efforts in trust-building and stakeholder alignment. Additionally, limited integration of existing outage datasets with real-time network performance data restricted detailed outage attribution modeling.

### ENABLERS

Verizon's strategic commitment to becoming a leading AI-driven company provided strong organizational support. The Impact Analytics team's cross-functional position bridging AI and business operations was pivotal, facilitating stakeholder engagement and rapid operationalization of models. Additionally, Verizon's extensive historical customer data and structured outage records were foundational for modeling efforts.

### ACTIONS

I first cleaned and synthesized data using regression-based imputation and MICE, then developed an interpretable linear regression churn model, validated rigorously against recent forecasts. I also built neural network-based predictive models linking weather variables to outages, applied DBSCAN clustering for event attribution, and integrated Monte Carlo simulations with climate projections to predict future risks.

### INNOVATION

Innovations include the cohort-based churn model explicitly incorporating upgrade behaviors, and advanced imputation techniques for sparse datasets. Additionally, combining DBSCAN clustering with neural network modeling for weather-event attribution, integrated with CMIP5-driven climate scenario simulations, provides Verizon a cutting-edge framework for strategic infrastructure resilience planning.

### IMPROVEMENT

My churn model notably improves forecasting, directly enabling more effective marketing strategies and optimized customer retention efforts. Infrastructure modeling clearly quantifies long-term climate risk, allowing targeted investments in network resilience, thus significantly reducing vulnerability to future extreme weather disruptions and associated customer churn impacts.

### BEST PRACTICES

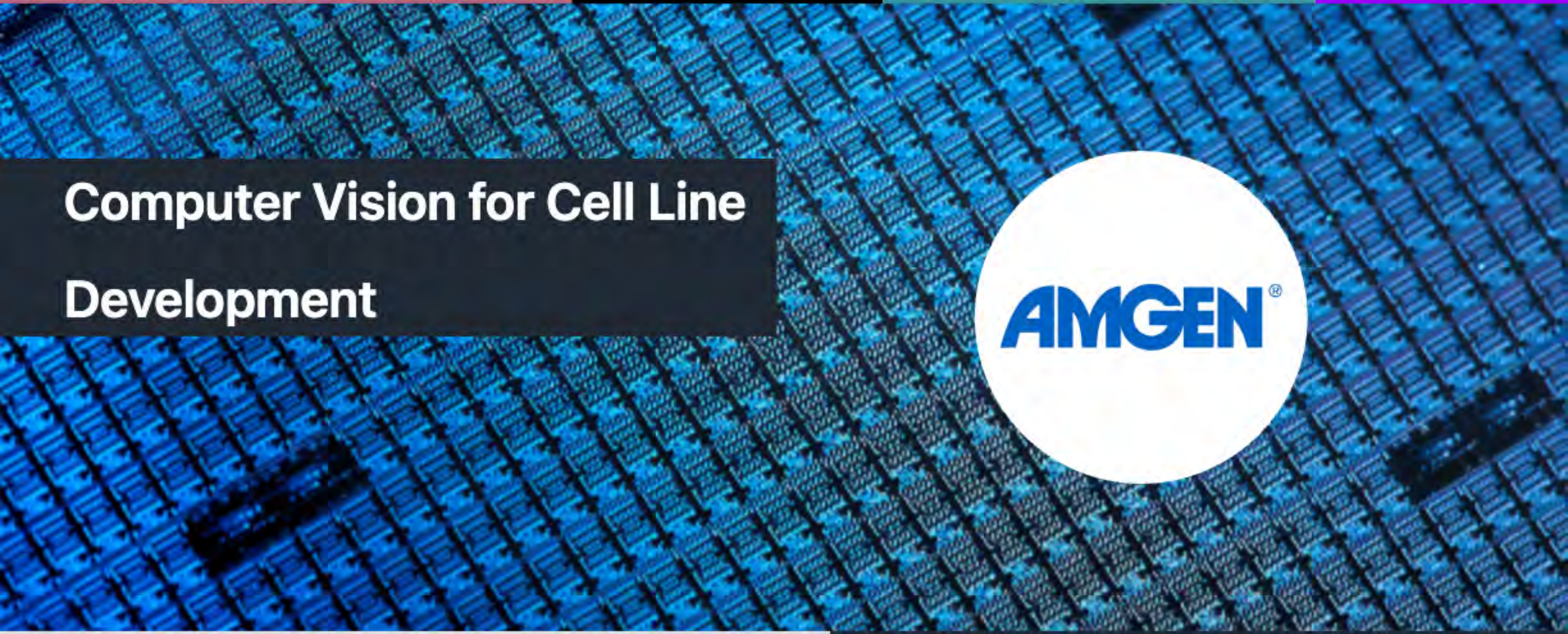
Prioritize stakeholder engagement and interpretability, particularly in enterprise contexts. Use cohort-based modeling to capture customer lifecycle dynamics, applying advanced imputation for data gaps. For resilience modeling, leverage robust spatio-temporal clustering methods combined with neural networks and climate scenario analysis, emphasizing transparency and explainability for practical business alignment.

### OTHER APPLICATIONS

This methodology can extend to other subscription-based industries (streaming, SaaS) for churn and upgrade analytics. Infrastructure resilience techniques can similarly apply to energy grids, water utilities, transportation networks, and critical supply chains facing climate risks. Predictive modeling frameworks can inform insurance underwriting, urban resilience planning, and governmental policy development for climate adaptation strategies.

Product design / development / engineering | Pharma / biopharma | Improved process performance | Auto-encoders

# Computer Vision for Cell Line Development



## BUSINESS PROBLEM

One of the most crucial interfaces in the drug development process is Cell Line Development (CLD). The general process consists of screening thousands of single cell clone candidates to determine a high quality Master Cell Bank (MCB). All future drug substance derive from this one step and the MCB. However, some candidates for the MCB are anomalies, requiring operators to spend hundreds of hours manually checking. Using the various forms of data generated in this process, Amgen looks to expedite this down-selection process and ensure anomalies are caught through the use of Machine Learning models.

## APPROACH

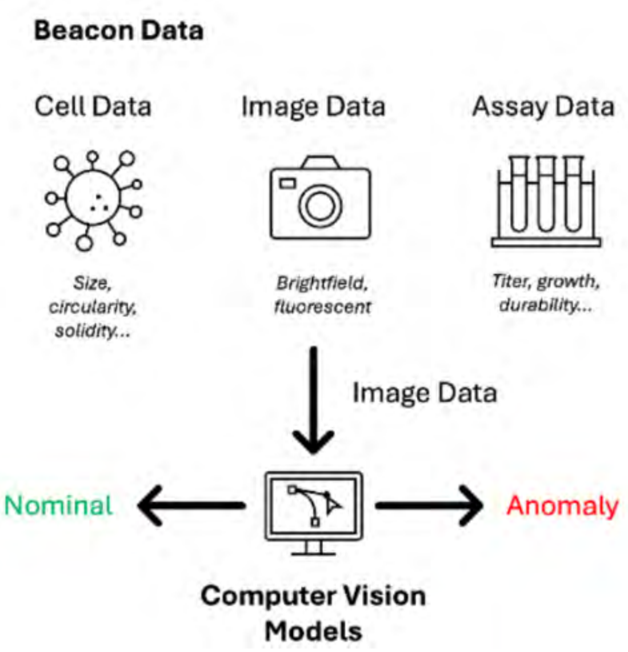
Three computer vision techniques were applied to identify anomalies without human intervention: a Fluorescent Image Autoencoder, an Edge Identification Convolutional Neural Network, and an RGB Channel Support Vector Machine (SVM). These techniques are evaluated for their effectiveness in anomaly detection within a historical dataset and the necessary resources.

## DATA SOURCES

The data primarily came from thee Beacon platform. Three types of data are produced inthe cell process: Cell Data, Assay Data, and Image Data. For a computer vision model, Image Data is the primary focus. Image Data consists of brightfield and fluorescent images, but models were only developed for fluorescent image anomaly detection.

### Data Types and Format

Image Data is generated as 22 brightfield and fluorescent png images. These images are spliced and split to produce 1,758 individual png images for model input and training.




Author: Jack Albright

## IMPACT

Implementing a computer vision model for anomaly detection has clear benefits for the CLD Process. To start, anomaly identification is a highly subjective process. Scientists in general have shared criteria that htey look for when making decisions on what is an anomaly versus a nominal image. However, scientists will focus on different small details that may lead them to separate conclusions. By introducing a model to supplement this process, the key factors that determine an anomalous image are incorporated to better standardize decision making. Reducing the subjective nature of the process means that the model reinforces selection of quality candidates for downstream development, while also providing an interface for scientists to exercise their expertise if they do not agree with the classification prediction. Concurrently, the amount of time scientists need to spend manually assessing whether cell clones are anomalies dramatically reduces. Originally, scientists could spend hundreds of hours going through this process, but by implementing a model to standardize this process, they now would only need to look over what is asserted as an anomaly. Further improving the model could result in scientists not needing to even dedicate time towards checking the validity of predicted anomalies, instead focusing on which clones are truly the top candidates for the MCB.

**DRIVERS**

Computer vision is widely used in biopharmaceuticals especially for detecting faults or quality issues. Leveraging the application of computer vision for challenges related to the biopharma manufacturing environment allowed focus on techniques most applicable for the CLD use case. Additionally, the data science experience within Amgen's TDC organization provided support.

**BARRIERS**

Fluorescent images coming from CLD have in-class variation that makes it difficult from a traditional autoencoder to learn what distinguishes a positive from negative class. Additionally, anomalies are outnumbered by nominal images. The class imbalance results in computer vision techniques readily skewing in favor of a nominal prediction as there is more data.

**ENABLERS**

The Transformative Digital Capabilities (TDC) organization at Amgen, the interface between data science experts and Process Development teams, was pivotal in the completion of this project. Working as part of TDC meant I had quick access to a range of expertise and support in establishing data pipelines to feed the models.

**ACTIONS**

I collaborated with the CLD team consistently to ensure the models developed would be useful for the scientists that would operate them. I also collaborated with another TDC team to ensure outputs from the computer vision models would serve as inputs to their more comprehensive models used for expedited clone selection.

**INNOVATION**

Per research on computer vision in biopharmaceuticals, there has not been an application of an autoencoder specifically used for fluorescent image anomaly detection. The innovative aspect comes from the ability of the autoencoder to be continuously trained on newer data after a batch run as been completed and labels for images are generated. The approach is self-reinforcing once implemented.

**IMPROVEMENT**

In its first iteration, the best performing autoencoder has an accuracy of 94% for anomalies and nominal images. Its precision for anomalies is also over 94%, while its recall remains suboptimal at 79%. However, its curve and composite metrics are also all significantly high, and the model correctly identifies 414 or 523 anomalies present in the evaluation set.

**BEST PRACTICES**

Establishing a strong data pipeline is critical so each model uses the same data as input. Additionally, it is important to establish a comparison baseline, especially in anomaly detection settings where class imbalances can impact accuracy and precision metrics.

**OTHER APPLICATIONS**

These solutions can be applied to other image types that may come from the rest of the process, such as brightfield images. The data pipeline for brightfield images has been established, so utilization of the models would merely require some architectural tweaking to determine whether this image type is suitable for making predictions.

# Evaluating The Feasibility of Electrified Process Heating for Drug Substance Manufacturing

## BUSINESS PROBLEM

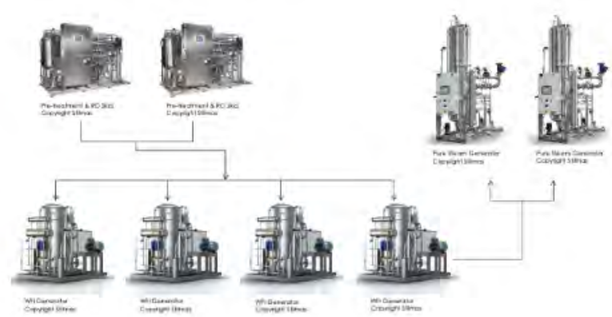
As Amgen strives for carbon neutrality by 2027, it faces the challenge of reducing Scope 1 emissions, which have risen by 12.5% since 2019 - despite progress in cutting Scope 2 emissions through renewable electricity. The primary source of Scope 1 emissions is the on-site combustion of fossil fuels to generate plant steam, used in high-temperature pharmaceutical manufacturing processes such as Water-for-Injection (WFI) generation, clean steam production, and equipment sterilization. The core business problem lies in decarbonizing these process-driven heat loads, which are currently dependent on fossil fuel-based plant steam systems.

## APPROACH

I conducted an analysis of Amgen's conventional plant steam system at a representative drug substance manufacturing site and explored electrification as a solution to eliminate usage of fossil fuels. Working with suppliers and stakeholders, I designed a fully electric system to meet the same high-temperature demands. A comparative analysis assessed cost, environmental impact, and feasibility.

## DATA SOURCES

The project used data from construction management platforms - Procore (equipment submittals, P&IDs) and Autodesk BIM360 (3D site building models). Insights into existing infrastructure came from site visits and engagements with site engineering and sustainability teams. Supplier websites, vendor calls, equipment quotes, and invoices provided data to evaluate electric equipment options.



### Data Types and Format

Primarily technical drawings, equipment specifications, and vendor documentation; supplemented by qualitative insights from engagements with key stakeholders at Amgen and with external suppliers.

Author: Priya Bhirgoo

## IMPACT

The proposed fully electric system has the potential to significantly advance Amgen's progress toward achieving carbon neutrality by 2027 by offering a viable alternative to fossil fuel-based process heating. The analysis indicates that capital investment requirements for the electric system are comparable to those of the conventional plant steam system, suggesting that upfront cost should not be a barrier to implementation. Additional construction savings are realized through the elimination of steam, condensate, and natural gas piping, as well as the boiler room. However, these may be offset by the need to upsize electrical infrastructure to meet increased power demand. The significantly higher cost of electricity relative to natural gas additionally emerges as a financial constraint. From a sustainability perspective, the electric system eliminates Scope 1 emissions associated with on-site combustion. While these are shifted to Scope 2 due to continued reliance on grid electricity, long-term environmental benefits will improve as the grid decarbonizes. Electrification also enhances operational flexibility by allowing decentralized equipment placement, modular facility design, and easier scalability. Overall, the findings support electrification as a viable strategy for reducing Scope 1 emissions and offer a scalable model that could be implemented across Amgen's global manufacturing network.

| | |
|---|---|
| **DRIVERS** | The primary driver was Amgen's corporate commitment to achieving carbon neutrality by 2027, with a focus on reducing Scope 1 emissions. The industry also faces growing pressure to decarbonize pharmaceutical manufacturing in response to evolving ESG expectations, regulatory trends, and investor priorities. Electrifying fossil fuel-based steam systems in manufacturing aligns with industry-wide decarbonization efforts. |
| **BARRIERS** | The most significant barrier to this project was addressing gaps in data. Cost data for certain existing equipment was unavailable due to contractual arrangements that did not include detailed cost breakdowns. Additionally, information on the natural gas rate and the local energy grid profile could not be provided by the corresponding suppliers. |
| **ENABLERS** | Several features of the company enabled the success of this project. I had a supportive and knowledgeable supervisor who was able to provide invaluable technical mentorship throughout my internship. Cross-functional collaboration is ingrained within the company culture and enabled access to key data and operational insights. The company's strong commitment to carbon neutrality also provided clear strategic alignment for the project's goals. |
| **ACTIONS** | To advance implementation, I presented my findings and recommendations to the Amgen team. I also ensured that my supervisor was connected with the vetted suppliers to support future electric equipment procurement and installation. My thesis will also serve as a pitch for the adoption of the fully electric system to leadership. |
| **INNOVATION** | The electrification of high-temperature heat demands represents a significant departure from established pharmaceutical industry practices regarding clean utility generation. It challenges the pharmaceutical industry's longstanding reliance on fossil fuels and offers a pathway towards carbon-neutral operations. |
| **IMPROVEMENT** | The solution identified electrification pathways that could eliminate up to 100% of Scope 1 emissions associated with clean utility generation at drug substance manufacturing sites. |
| **BEST PRACTICES** | Best practices for replicating this solution include conducting detailed site assessments to thoroughly evaluate existing process heating systems and enable effective collaboration with suppliers to design appropriate electrification-based solutions. Additionally, engaging key stakeholders such as engineering, utilities, and sustainability teams early is essential to ensure alignment between technical requirements and organizational objectives. |
| **OTHER APPLICATIONS** | This solution can be extended beyond pharmaceuticals to effectively decarbonize process heating across various industries, including the food and beverage sector, which similarly relies on fossil fuel-based steam boilers in its operations. |

# Enhancing Autonomous Vehicle Testing with Diffusion-Based Camera and LiDAR Sensor Simulation

## BUSINESS PROBLEM

Waymo invests significant resources in collecting test and validation miles to safely expand its autonomous ride-hailing service into new territories. This process is both costly and time-consuming. The business needs a more scalable, cost-efficient method to test critical driving scenarios. Generative AI, particularly diffusion models, provides an opportunity to reduce the need for physical testing by simulating realistic, multimodal sensor data—including RGB camera images, LiDAR depth maps, and 3D point clouds—thereby lowering operational expenses and expediting expansion efforts.

## APPROACH

This work leverages and extends stable diffusion models to simulate 360-degree driving scenes with synchronized camera RGB and LiDAR sensor data. The approach involves encoding data into the model's latent space, training a multimodal diffusion pipeline for generating consistent, multi-view data, and evaluating the results using realism, depth accuracy, and sensor alignment metrics.

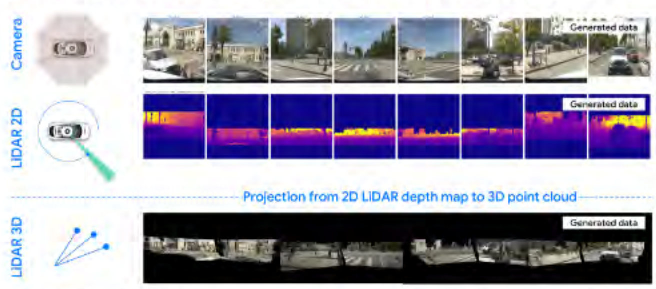
## DATA SOURCES

The project utilized multimodal sensor datasets captured from autonomous vehicle testing in urban environments. These datasets were internally sourced and reflect real-world driving scenarios, providing a rich foundation for training and evaluating generative models.

### Data Types and Format

The data included high-resolution RGB images from multiple vehicle-mounted cameras, 2D LiDAR depth maps, and 3D point clouds. Formats used were consistent with machine learning pipelines.

Methodology: A multimodal generative AI model simulates eight camera and 360' LiDAR sensor signals for new driving scenes




The trained latent stable diffusion model can enhance simulation realism for testing of autonomous vehicles. This could in-turn lead to a reduction of required on-road testing miles saving operational costs.

Author: Linn Bieske

## IMPACT

This work has the potential to enhance Waymo's simulation capabilities by enabling the generation of realistic, multimodal driving environments. By producing synchronized, high-fidelity RGB and LiDAR sensor data, the developed model could support more efficient iteration, targeted scenario exploration, and robustness evaluation of autonomous systems. The project may contribute to foundational generative AI research within the organization and facilitate internal knowledge sharing through well-documented pipelines and presentations. Looking ahead, the approach could inform future efforts aimed at real-time simulation and integration of additional sensor modalities. If further developed and validated, this direction might help reduce operational costs, improve scalability when entering new territories, and diversify training data sources for autonomous vehicle testing.

**DRIVERS**

The autonomous vehicle industry demands scalable, safe, and cost-effective methods for testing and validation. As companies aim to expand ride-hailing services into new territories, the need to reduce reliance on physical test miles—while maintaining scenario coverage and sensor fidelity—was a key catalyst for exploring generative simulation techniques using AI.

**BARRIERS**

Key challenges included aligning camera RGB and LiDAR data in the latent space, managing the computational complexity of training multimodal diffusion models, and ensuring the realism and consistency of generated sensor outputs. Limited off-the-shelf tools for such multimodal generative modeling required extensive custom development.

**ENABLERS**

The project was supported by a collaborative research environment, access to large-scale proprietary sensor data, and mentorship from experts in perception and simulation. Integration with ongoing internal research efforts also enabled feedback loops and alignment with strategic goals.

**ACTIONS**

To implement the solution, we curated training datasets, developed a multimodal latent diffusion model, and integrated a variational autoencoder (VAE) for encoding LiDAR data. We iteratively trained and evaluated the model on simulation quality metrics and documented the findings to support reproducibility and knowledge transfer.

**INNOVATION**

This work introduces a transformative simulation approach by unifying camera and LiDAR data in a shared latent space using a stable diffusion framework. It enables synchronized, high-fidelity 360° scene generation, bridging generative image synthesis and sensor emulation. This marks a leap beyond traditional tools and sets the stage for scalable, data-driven autonomy development.

**IMPROVEMENT**

The model demonstrated the ability to generate high-fidelity, spatially consistent sensor data across camera and LiDAR modalities. This approach has the potential to reduce reliance on costly physical test miles and significantly broaden the range and diversity of driving scenarios available for simulation and validation.

**BEST PRACTICES**

Success in similar projects depends on robust data preprocessing, thoughtful architectural integration of sensor modalities, clear performance metrics, and iterative feedback from domain experts. Documentation and close collaboration with internal teams significantly enhance adoption and scalability.

**OTHER APPLICATIONS**

Beyond autonomous vehicle testing, this generative simulation framework could be applied to synthetic data generation for robotics, AR/VR environments, smart city planning, or any domain requiring multimodal sensor fusion. It also lays the groundwork for real-time generative systems in hardware-in-the-loop testing environments.

Production | Industrial automation | Digital standard work / training | Linear / Non-Linear

# The Value of Digitizing Manufacturing Environments


## BUSINESS PROBLEM

The company is attempting to drive operational improvements to the flow and processes within sites by investing in digital and connected factory (Industrial Internet of Things) technology. The goal is to improve the fundamental operating systems and capabilities for lean and continuous improvement. The team requires a method to quantify the value generated from the implementation of digital tools in the manufacturing environment. The approach must be applicable to the large portfolio of digital tools employed and variety of production environments at SBD (100+ sites, 150k+ SKUs, 50k+ employees).

## DATA SOURCES

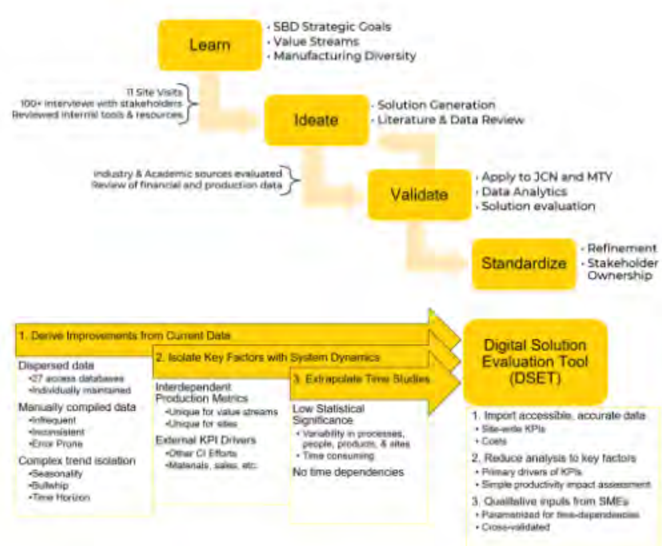
Data sources included legacy productivity databases and financial reporting spreadsheets derived from SAP. Qualitative inputs came from interviews with subject matter experts, cross-validated with time studies and alternate sources. Production data derived from the new digital tools became available later in the project.

### Data Types and Format

Data was tabulated in a combination of Excel tables, Access databases, Grafana dashboards, and Snowflake repositories. The productivity data was compiled into time series of KPIs.

## APPROACH

A system dynamics model analyzed the production environments and identified key factors that strongly correlated with the potency of the delivered digital solution. A simple, accessible, and repeatable Excel based Digital Solution Evaluation Tool (DSET) was built to qualitatively assess these factors through selective analysis of available data and numerous interviews with subject matter experts.





Author: Conor Briggi

## IMPACT

The digital transformation effort at SBD is capital-intensive and requires justification for continued investment while the manufacturing leadership needs to understand the risks and opportunities of the technology embedded into critical processes. Additionally, major problems occur during implementation of digitization efforts without monitoring systems to assess progress. By quantifying the value generation and costs from this digital push, SBD can optimize the delivered solution for each site and value stream. These solutions generate different ROI depending on the application, so mapping best practices and guiding the development and deployment strategy will dramatically improve the chance of success for this digital transformation at SBD. Unfortunately, there is currently no industry approved method for calculating the value created by digitizing manufacturing environments which prevents presenting a business case through a traditional approach. The wide spectrum of manufacturing environments, variety of digital tools, and dependency of results on time, culture, and people cause assessments to be ephemeral. Therefore, assessment workload will be inversely correlated with the applicability towards strategic insights. If it sacrifices efficiency for extremely precise results, those results will be incorrect by the time the assessment is complete and the assessment will be done infrequently, limiting the impact on digital strategy.

**DRIVERS**

Despite the long years of success as a company, SBD is in transition. Recent years of leadership challenges and supply chain bullwhip from the COVID-19 pandemic hit SBD particularly hard. To counteract those challenges, SBD has been focused heavily on supply chain optimization through operational excellence initiatives, site consolidation, and supply chain localization (partially due to changing tariff policies).

**BARRIERS**

Due to the size, diversity, and frequent acquisitions, there is significant variety in the people, products, and processes across the company. Additionally, the matrix structure encourages individual sites to tailor processes to the unique products they produce while generalized functional teams seek standardized solutions that apply across the company. A previous digital transformation attempt left sites opposed to new digitization efforts.

**ENABLERS**

Strong support from site digital champions was a major part of the project success. These individuals supported the effort and connected the team to pivotal manufacturing team members. Another factor was SBD leadership embracing rapid decision making and understanding that decisions often must be made from imperfect information. This enabled the team to act with speed to meet business needs.

**ACTIONS**

The productivity data was analyzed, and the production environments were mapped to a system dynamics model to identify causal relationships. Concurrently, the costs of implementing the digital solutions were tabulated for each of the stakeholders. A tool was built that compiled costs, immediate improvements, long-term productivity impacts, and positive externalities to assess financial performance of delivered solutions.

**INNOVATION**

The differentiating factor of the project was the adaptability to diverse digital solutions in various environments. The tool targets key factors identified from the system dynamics model for efficiency and flexibility. Time-dependent results are fit to generalized learning curve shapes and adjusted with risk factors. The tool also emphasizes transparency to achieve accessibility for all stakeholders, prioritizing straight-forward methodologies.

**IMPROVEMENT**

Direct impact from the tool usage has not been measured at this time. However, during development of the tool, various opportunities were identified and executed including 14% improvement to training efficacy and 7% improvement to labor efficiency for some value streams. Opportunities identified during the project and planned for execution in 2026 could achieve an additional $6.1M (2.9M risk adjusted) in savings per site deployment.

**BEST PRACTICES**

The digital transformation team developed an assessment procedure to evaluate value from different digital solutions at different sites. Taking an accurate baseline initially would enable more accurate accounting of solution performance. Implementation must be properly executed, focusing on areas with larger relative impacts such as inventory, attainment, and labor efficiency while strategically prioritizing the order of asset connection.

**OTHER APPLICATIONS**

The methodology and approach for this tool, while designed for digital tools in manufacturing environments, could be adapted to other strategic initiatives around productivity improvements. It could also be integrated into the initial assessment process for future digital tool deployments. Due to the inherent flexibility in the tool, applications could be found in other industries and manufacturing environments.

Operations | Aerospace and defense | Improved process performance | Constrained / Unconstrained

# A Data-Driven Work Center Assignment and Pricing Strategy for a Job Shop

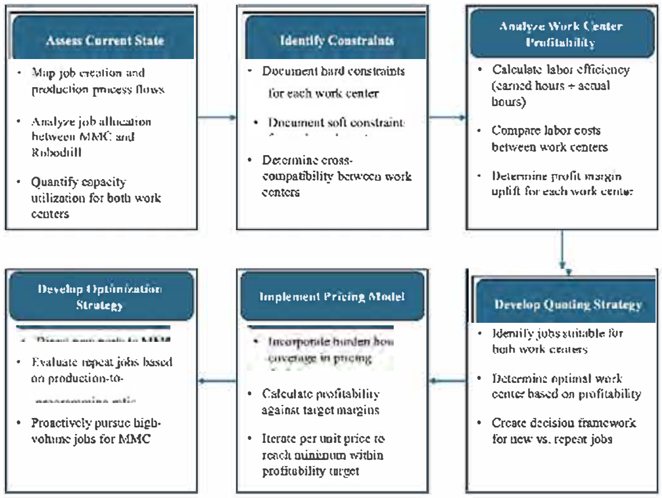
MITLGO

## BUSINESS PROBLEM

Machine Shop Baltimore (MSB) has a semi-autonomous CNC work center, the Makino Machining Center (MMC), but is struggling to win high-volume work best targeted for the machine. The shop lacks a structured quoting and pricing strategy making it difficult to win the target jobs. Although the shop tracks Industry 4.0 data, it does not utilize this information for work center assignments. A formal quoting and pricing framework that leverages shop data would enable MSB to optimize work center decisions, improve pricing alignment, and enhance competitiveness to target right-to-win high-volume jobs.

## APPROACH

The approach starts by analyzing historical data to identify gaps in work center utilization for the MMC and Robodrill. Interviews and observations map the quoting process and identify constraints for both work centers. A quoting and pricing framework is developed to compare job economics across both centers. MSB's cost structure is applied to find missed opportunities for future improvements.




## DATA SOURCES

The data includes both ERP records (Jan 2023-May 2024) tracking job metrics and machine monitoring software that tracks spindle time for MMC and Robodrill machines. Analysis compares job allocation to identify biases, calculate labor efficiency, and capacity utilization. These metrics were used to compare the profitability of each work center considering burden hours and labor efficiency.

### Data Types and Format

The data is time-series formatted per work center: ERP entries chronologically track operator tasks and machine monitoring software chronologically tracks machine spindle time.


Author: Alix Carson

## IMPACT

The two analytical frameworks transformed MSB's operations: a quoting model for work center allocation and a strategic pricing model. The quoting model addressed bias between Robodrill and MMC machines by applying data-driven analysis of labor efficiency, burden costs, and profitability. Analysis of September 2024 jobs revealed the MMC was consistently more profitable for new parts, while for repeat jobs, profitability depended on the ratio of production labor to programming/fixture labor. This created clear guidelines: new parts to MMC; Robodrill repeat jobs only switch when production hours significantly outweigh switching costs. The pricing model incorporated burden hour coverage into pricing decisions. Analysis of a high-volume aluminum job compared MSB-sourced materials versus customer-provided materials. The customer-provided scenario proved optimal, covering 65% of annual burden hours while using only 40% of machine capacity, yielding a 38% profit margin that exceeded the company's 35% target. It would also increase MMC utilization from 29% to 58%, exceeding the 2027 goal. These models enable MSB to: 1. Optimize work center allocation based on quantifiable metrics 2. Define "right to win" jobs based on costing, burden coverage, and profitability 3. Make pricing decisions that maintained profitability while meeting customer targets This data-driven approach allows MSB to confidently bid on favorable jobs rather than pursuing all available work.

**DRIVERS**

The job shop industry faces inefficiencies in quoting and pricing due to inaccurate labor tracking, misallocated resources, and rigid pricing structures. Despite having IoT capabilities, many shops fail to leverage real-time data for decision-making. This led to the need for a data-driven quoting and pricing model to optimize work center allocation, ensure burden cost recovery, and enhance competitiveness through strategic pricing adjustments.

**BARRIERS**

Barriers encountered include inaccurate data entries and disconnected systems. The ERP relies on operators properly clocking in/out, resulting in inaccurate data entries. MSB tracks burden hours differently between MMC and Robodrill, making Robodrill appear less burdened. Without integration between ERP and machine monitoring software, autonomous operations go unrecorded, distorting utilization and efficiency metrics across work centers.

**ENABLERS**

MSB's investment in Industry 4.0 technology provided real-time tracking of operator hours and machine utilization, enabling data-driven insights. The shop's semi-autonomous machining systems, MMC and Robodrill, offered flexibility in job allocation. Additionally, the team's openness to process improvements and willingness to explore data-driven decision-making created a foundation for optimizing quoting and pricing strategies.

**ACTIONS**

I applied the quoting and pricing models to MSB's sales pipeline. For September jobs, the quoting model showed a $13.5K projected loss due to unoptimized work center assignments. The pricing model analyzed two high-volume quotes, revealing that not all high-volume jobs are ideal. The optimal job had a projected profit of $750K, highlighting the importance of strategic job selection and pricing optimization.

**INNOVATION**

The solution integrates burden hour tracking into the quoting process. Tracking burden hours in the quoting process ensures that a machine shop accurately accounts for fixed costs, preventing underpricing and margin erosion. By understanding how many burden hours have been covered, shops can offer competitive rates when fixed costs are met. This leads to better resource allocation, optimized profitability, and data-driven pricing decisions.

**IMPROVEMENT**

MSB is positioned to increase its profits without additional resources. The pricing tool enabled MSB to bid for a high-volume job that would bring in $750k in profit and increase the MMC utilization by 58%, exceeding the shop's target of 39%. Additionally, the quoting model provides MSB has a framework for allocating future work between the MMC and Robodrill.

**BEST PRACTICES**

Ensure accurate IoT data collection for labor and machine utilization. Establish a clear costing structure that includes burden hours. Regularly assess burden coverage to adjust pricing dynamically. Prioritize high-margin jobs over volume alone.

Supply chain planning and inventory management | Med devices / MedTech | Inventory management / control | Constrained / Unconstrained

# Safety Stock Modeling for a Medical Devices Supply Chain

## BUSINESS PROBLEM

Historically, the med tech industry has been heavy in inventory. This mostly attributed to being very risk averse to stockouts and using excess inventory as a means to ensure against stockouts. This was exacerbated during the supply chain crisis following COVID-19. The business is trying to gain better visibility into their supply chain and strike the right balance between raw material and finished goods inventory through more robust inventory management strategies.

## APPROACH

I narrowed the focus of my study to safety stock targets. I started with a current state analysis, mapping out the process flow and supply chain and running diagnostic analyses. I then used single-echelon and multi-echelon modeling to calculate new safety stock targets. I compared the results of each model against the current values for two finished goods and their respective components.


## DATA SOURCES

Most of the data that I used came from the company's ERP systems that were centralized in a data lake. This included forecast data, actual inventory levels, safety stock targets, historical demand and consumption, etc. I also accessed bill of materials (BOM) data from the company's project management program.

**Data Types and Format**

CSV files, PowerBI dashboards and reports, interviews with key stakeholders

Author: Julie Chong

## IMPACT

My solution examines the business's inventory management practices and highlights areas of improvements. It identifies that finished goods inventories are excessively high due to conservative forecasting and arbitrary safety stock bounds, while raw material inventories are underestimated, leading to potential disruptions. The reliance on manual overrides and frequent interventions, such as expedited shipping, adds unnecessary costs and operational inefficiencies. The thesis recommends transitioning to data-driven safety stock calculations, adopting multi-echelon inventory modeling, and reevaluating service level targets to balance cost efficiency with customer satisfaction. It also emphasizes improving data visibility, tracking accurate lead times, and reforming the flow of information across teams. Implementing these strategies could significantly enhance the business's supply chain resilience and operational efficiency. For instance, MEIO could optimize inventory across all stages of the supply chain, while adjusting service levels slightly (e.g., from 99.8% to 98%) could save millions annually without compromising customer satisfaction. These changes would reduce costs, streamline processes, and improve overall supply chain performance.

### DRIVERS

There is a lot of potential for cost savings in the form of more efficient inventory management in the business and the medtech industry at large. Having the adequate tools to analyze tradeoffs will benefit the operations and performance of the business.

### BARRIERS

I had two primary barriers that impacted my project. The first was a changeover in one of the ERP systems that restricted access to data and tied up resources. The second was relatively recent organizational change that resulted in newer personnel in many of the functions that directly affected my work.

### ENABLERS

My team was really eager to help and very open to conversation and this was a huge positive in moving the project forward. Also, there was a fairly robust data architecture that made it easy access the data that was available.

### ACTIONS

1. Quantitatively and qualitatively assessed the current state and uncovered safety stock planning as the main bottleneck 2. Gathered/deduced "clean" data inputs 3. Produced single and multi-echelon inventory models 4. Calculated revised safety stock values for raw materials and finished goods 5. Performed cost-benefit analysis

### INNOVATION

Using a multi-echelon inventory model as opposed to the siloed single-echelon method that the business is currently using is one of the more innovative aspects of my solution. This approach leverages the relationship between stages, which are inherently interconnected. Another innovative aspect is reevaluating the stated service level such that a similar level of service can be provided to customers at dramatic cost savings.

### IMPROVEMENT

First, I discovered a discrepancy of up to 43% between the actual and theoretical (per the current safety stock levels) service level suggesting significant manual intervention. I also found a slight reduction in finished goods safety stock targets and an increase in raw material safety stock targets compared to current values with single and multi-echelon models, which translated to significant potential cost savings.

### BEST PRACTICES

The most important practice when attempting to replace my solution is using the right data. The business has a plethora of data sources and, at times, multiple truths for a desired data point, so discerning which to use is key.

### OTHER APPLICATIONS

My solution can be applied to other finished goods and raw material SKUs, including those belonging to other business units.

Engineering | Healthcare | Defect detection / product inspection | Clustering

# Impact Evaluation and Prioritization Framework for Manufacturing Inspection Technology Investment

## BUSINESS PROBLEM

The main business problem Johnson & Johnson Vision is trying to solve involves inefficiencies in its quality control processes for Acuvue contact lenses. These inefficiencies lead to unnecessary scrap, customer dissatisfaction, and lead time variability due to false rejections and acceptances, excessive manual inspection, and elevated customer complaints. A new inspection technology will address these challenges and needs to be implemented across their 100 manufacturing lines. Quantifying the impact and prioritization of the technology investment is critical to optimizing the rollout.

## APPROACH

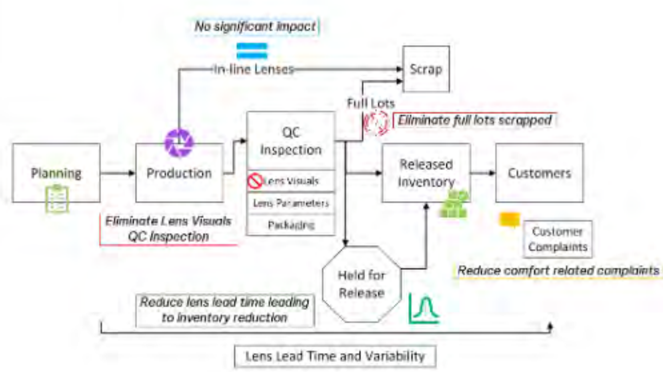
Data Analysis: Collecting historical data on yield, scrap rates, customer complaints, and inventory levels. Financial Modeling: Projecting future cash flow based on expected improvements. Engineering Risk Assessment: Evaluating technical compatibility and line longevity. Prioritization Framework: Combining cash flow impact and engineering risks to prioritize upgrades

## DATA SOURCES

Lot Description Data: Detailed information about each production lot. Customer Complaints: Historical data on customer complaints related to visual defects. Quality System Nonconformances: Records of quality issues detected during production. Scrap Data: Information on lenses scrapped due to defects. Demand and Lead Time Information: Data on production volumes and supply chain lead times.

**Data Types and Format**

Structured data formats, such as spreadsheets and databases, with some potential for unstructured data from customer feedback.




Author: Isabella DiDio

## IMPACT

The solution to upgrade Johnson & Johnson Vision's quality control processes for Acuvue contact lenses using advanced camera optics and machine learning has several key impacts on the business: Cost Savings: Reduces unnecessary scrap, manual inspection labor costs, and customer complaint costs. Scrap savings alone can be substantial, especially on high-volume lines. Operational Efficiency: Improves defect detection accuracy, minimizing false rejections and acceptances, which enhances overall product quality and reduces lead time variability. Inventory Management: While cycle stock remains unaffected, reductions in pipeline stock offer tangible cost savings due to decreased lead times. Customer Satisfaction: Decreases customer complaints by ensuring fewer defective lenses reach customers, enhancing brand reputation and loyalty. Strategic Positioning: Aligns with Johnson & Johnson's commitment to innovation and quality, setting new industry standards for medical device manufacturing

**DRIVERS**

Operational Inefficiencies: Current inspection methods led to unnecessary scrap and customer dissatisfaction. Technological Advancements: Machine learning and visual inspection technologies offered opportunities to enhance defect detection and reduce manual labor. Industry Competition: The competitive contact lens market necessitates high-quality products and efficient manufacturing processes. Regulatory Compliance: Strict regulatory standards

**BARRIERS**

Technical Complexity: Integrating advanced camera optics and machine learning across 100 manufacturing lines posed significant engineering challenges. Operational Disruption: Minimizing production downtime during upgrades was crucial to avoid impacting supply chains and customer delivery. Data Quality Issues: Ensuring accurate and reliable data for decision-making was a challenge due to the many different data sources

**ENABLERS**

Decentralized Organizational Structure: Johnson & Johnson's structure allowed for flexibility and autonomy, facilitating innovation and adaptation across different business units. Strong Corporate Culture: The company's commitment to quality and innovation, guided by "Our Credo," supported the integration of advanced technologies. Data-Driven Decision Making: Access to comprehensive historical data and advanced analytics tools

**ACTIONS**

Developed a Prioritization Framework: Combined historical data analysis, financial modeling, and engineering risk assessments to prioritize manufacturing line upgrades. Conducted Data-Driven Planning: Utilized dashboards to visualize projected business impacts and adjust implementation dates. Managed Stakeholder Engagement: Ensured collaboration between engineering, finance, and operations teams. Monitored Progress: Continuously reviewed data

**INNOVATION**

Advanced Camera Optics and Machine Learning: Integrating these technologies enhances defect detection accuracy, automates manual inspection processes, and reduces false rejections and acceptances. Data-Driven Prioritization Framework: This framework combines historical data analysis, financial modeling, and engineering risk assessments to strategically prioritize upgrades across manufacturing lines, maximizing return on investment.

**IMPROVEMENT**

Reduced Scrap: Elimination of full lot scrap due to enhanced in-line defect detection. Lower Customer Complaints: Decrease in customer complaints by up to 100% for visual defects. Labor Cost Savings: Significant reduction in manual inspection labor costs. Inventory Efficiency: Tangible cost savings from reduced pipeline stock due to decreased lead time variability.

**BEST PRACTICES**

Data-Driven Decision Making: Use historical data and analytics to inform prioritization and implementation strategies. Interdisciplinary Collaboration: Engage teams from engineering, finance, and operations to ensure alignment with business objectives. Phased Implementation: Roll out technology upgrades in phases, balancing engineering risks and cash flow impact. Continuous Monitoring: Utilize dashboards to track progress and adjust plans

**OTHER APPLICATIONS**

Pharmaceutical Manufacturing: Enhancing quality control for pharmaceutical products by detecting defects and contaminants. Medical Device Production: Improving inspection processes for other medical devices, such as surgical instruments or implants. Automotive and Aerospace: Applying similar technologies to inspect components in these industries, ensuring high-quality standards.

Supply chain planning and inventory management | Med devices / MedTech | Network optimization | Deterministic / Stochastic

# Systems Approach to Product Code Optimization in Wound Closure Supply Chain

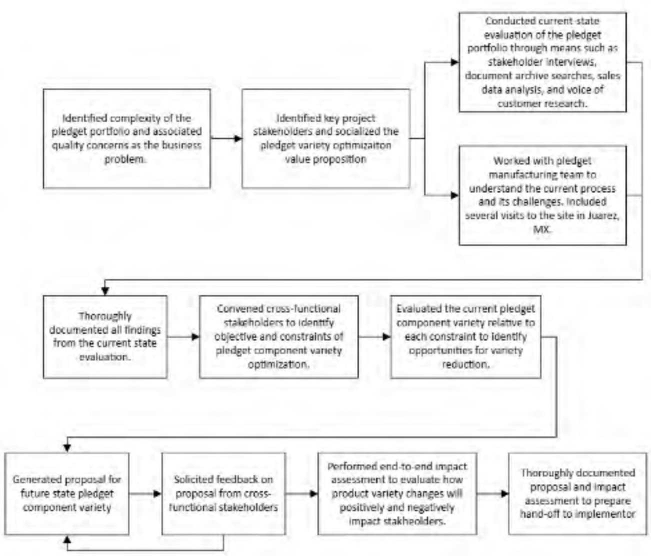
Johnson & Johnson

## BUSINESS PROBLEM

Pledgets are a critical component in non-absorbable suture systems used in cardiovascular and other surgical procedures. An increase in pledget quality complaints drew attention to these components that have undergone little lifecycle management in the last 40 years. In that time, the component portfolio had proliferated, introducing substantial product and process complexity. There is a need for a systematic, end-to-end supply chain management approach to decrease pledget component variety, reducing manufacturing complexity and enabling quality and process improvements.

## APPROACH

First, an end-to-end evaluation of the pledget variety captured the current-state of the portfolio. Second, cross-functional stakeholders defined the objective and constraints for pledget variety optimization. Third, the variety was evaluated with respect to each constraint. Fourth, a variety optimization strategy was proposed. Fifth, an end-to-end assessment evaluated impacts of variety changes.




## DATA SOURCES

Most data were pulled from the Tableau for the business unit that aggregates sales, production, and quality data. This includes global sales data, component production data, finished good production data, and bill of materials. Other internal data includes complaint and material waste data. Public product catalogs for J&J MedTech and its competitors were used to get data on product attributes.

### Data Types and Format

Product complaint images, product attributes (categorical), structured data for production and sales.

Author: Madeline Dubelier

## IMPACT

The proposed portfolio changes reduce the pledget component variety by 60%, exceeding the original target of 40% and meeting each of the three constraints. The constraints were informed by the cross-functional stakeholders' priorities for variety optimization: Continue to meet customer needs: this methodology reduced the component variety by 29% with no impact on the customer. The remaining reductions were informed by updated voice of customer research to ensure the future variety meets today's customers' needs. Protect J&J MedTech's competitive position in the market: the reduced component variety covers the same size range, ensuring that gaps in offerings cannot be leveraged by competitors. An evaluation of product code pricing also drew attention to opportunities for margin improvement. Reduce manufacturing complexity: The reduced component variety shifts volume to the more repeatable pledget punching manufacturing process. Reduced manufacturing complexity also brings clarity to the root cause of the product quality issues and which investments in capital equipment and process improvement will provide the most value to the business. This work presents a systematic, end-to-end approach to portfolio optimization to reduce manufacturing complexity and improve product quality and alignment between the portfolio and business strategy. This methodology offers a replicable model for similar legacy product portfolio optimization in the medical device industry.

**DRIVERS**

A recent increase in customer product complaints drew attention to the pledget portfolio. Many of these complaints originate from cardiovascular surgeons, as pledgets are commonly used in these surgeries. Given the increasing importance of the Cardiovascular portfolio to the Wound Closure & Healing business, there was concern that these product quality issues may lead to customer attrition.

**BARRIERS**

Data availability – there were gaps in process and supply chain data. For example, it was not possible to calculate the scrap rate for the different pledget processes or monitor impact of implemented solutions. Timeline – Implementation was not possible on the six-month timeline of this work given the administrative, quality system, and global regulatory changes that will be required.

**ENABLERS**

Stakeholder Buy-In – This project needed support and involvement from cross-functional stakeholders. With a compelling value proposition, stakeholders were willing to buy in despite sometimes conflicting priorities. Company Mission – J&J is a very mission-driven company, and most decisions come back to the responsibility to serve customers and patients, serving as a true guiding principle.

**ACTIONS**

This work culminated in a comprehensive review of the current state evaluation for the pledget portfolio and recommendations for component variety optimization. I presented the recommendations to multiple leadership teams, stakeholders, and those I was handing the project off to before I left. Solution implementation began after my project was complete.

**INNOVATION**

The literature on product variety management and optimization provides variety optimization methods that focus on a single stakeholder's perspective. Unilateral approaches could not be leveraged, as an effective, compliant solution will require input from cross-functional stakeholders. This project also proposes qualitative approaches to portfolio optimization given lack of data availability.

**IMPROVEMENT**

The five-stage of variety optimization proposal reduces component variety from 42 to 17 pledget codes, a 60% decrease compared to the 40% objective. The proposed variety spans the same size range as the original variety, preserving portfolio competitiveness. The proposal shifts 42% of total pledget volume to the more repeatable round corner punching process with a 32% raw material cost increase.

**BEST PRACTICES**

Executing a comprehensive baseline or current-state evaluation of the portfolio is critical, as justification for all recommendations will rely on this evidence and data. Identify and engage cross-functional stakeholders from the outset. Understanding their true priorities and individualized value proposition shaped the optimization constraints to inform the future-state portfolio recommendation.

**OTHER APPLICATIONS**

This method can be applied to other variety optimization cases with different motivations (inventory buildup, change in customer needs), objectives (increasing margins or capacity), and constraints (regulatory, change execution time). Ensure the scope of the variety optimization is not excessively broad, as a deep understanding of the individual component variety enabled the analysis.

R&D | Healthcare | Product design and testing | Ensemble-based (bagging, boosting, stacking)

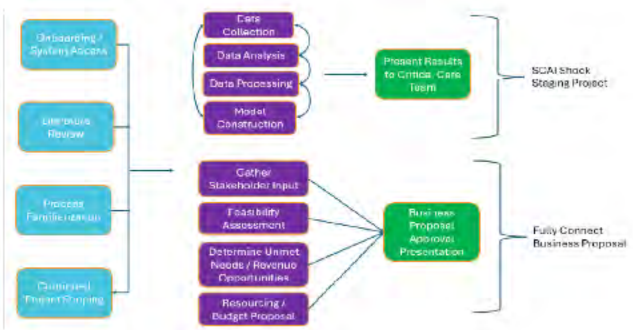
# Fully Connected Digital Ecosystems within Hospitals - AI/ML Solutions for Improved Patient Care

## BUSINESS PROBLEM

Abiomed is committed to its core value of "Patients First," aiming to save more lives through timely and effective interventions. I sought to determine if SCAI shock stages could better identify patients for Abiomed's Impella pumps and if EMR integration into its ecosystem was feasible. Siloed data and delayed cardiogenic shock interventions hinder timely care, despite Impella's real-time data, due to EMR integration challenges. My aim was to address these barriers to improve patient outcomes, aligning with Abiomed's patient-focused goals.

## APPROACH

I gathered data from prior studies and iteratively built an ML model to accurately predict SCAI shock stages, then assessed its usefulness and suitability as a tool. For EMR integration, I conducted multiple interviews with Abiomed stakeholders to gauge feasibility and strategic value. I synthesized these findings into a comprehensive assessment.




## DATA SOURCES

I utilized three data sources: two from studies on Impella pumps and a third from Abiomed's Impella Quality (IQ) Database. These datasets included diverse patient data (e.g., lactate levels, vasopressor use) alongside SCAI shock stage classifications, enabling analysis of shock severity and treatment outcomes.

### Data Types and Format

The data I used were tabular, consisting of structured clinical measurements from patients. These included static numerical values, categorical variables, and demographic details.


Author: Andrew Dugan

## IMPACT

The impact of my project lies in laying the initial research groundwork for assessing the feasibility of using SCAI Shock as a metric to determine if a patient needs an Abiomed product or care escalation. My ML models accurately predicted shock stages (up to 94% accuracy), but I found SCAI's reliance on treatment decisions limits its real-time impact. Instead, I propose nudging physicians to collect more data (e.g., lactate) or explore alternative metrics for Impella candidacy, potentially improving early interventions and outcomes. My EMR integration analysis showed it's feasible and realistic, achievable with a dedicated team, and could standardize CS care, boosting patient capture and Abiomed's leadership. This work sets the stage for future research into physiological markers and actionable data ecosystems.

**DRIVERS**

The catalyst for my solution was Abiomed's core value. "Patients First" I aimed to leverage Abiomed's unique Impella data position to enhance patient identification and outcomes, driving better CS management and hospital adoption.

**BARRIERS**

Major barriers included acquiring consistent data from diverse studies and grasping the complexity of hospital data ecosystems. Understanding how systems interconnect from nurses and field reps to physicians and patients-posed challenges in integrating impella data with EHRs and predicting SCAI shock stages effectively.

**ENABLERS**

The biggest enabler was the mentorship from Abiomed's team. Their expertise facilitated rapid onboarding into the problem space and provided access to vast knowledge about Impella products and the medical industry, enabling effective model development and EHR integration analysis with deep insights and guidance throughout the project.

**ACTIONS**

My key action was attending numerous meetings to absorb critical insights. I conducted multiple stakeholder interviews to assess EMR integration feasibility and worked diligently to understand the data sources, ensuring a solid foundation for building the SCAI shock prediction model and evaluating its alignment with Abiomed's goals.

**INNOVATION**

My key innovation was questioning. "Is SCAI shock the right metric for this goal?" and testing its validity. Adopting an outsider's perspective. I challenged assumptions about its utility for patient identification, driving a shift toward exploring alternative metrics and nudging physicians, enhancing the approach to improving CS care and impella use.

**IMPROVEMENT**

I developed a model with 94% accuracy in predicting SCAI shock stages, but its reliance on physician decisions revealed SCAI's limitations. This enabled me to demonstrate that SCAI shock isn't the ideal metric for patient identification, redirecting focus toward earlier data collection for better CS management and impella utilization, improving Abiomed's approach to saving lives.

**BEST PRACTICES**

To replicate my solution, prioritize deeply understanding the data-its sources, quality, and limitations-and how hospital systems (nurses, physicians, field reps, ENRs) interconnect. Master clinical variables like lactate and vasopressor use, map data flows across studies and devices, and engage stakeholders early to align technical models with real-world workflows, ensuring effective SCAI prediction and ENRR integration.

**OTHER APPLICATIONS**

My solution could extend beyond Abiomed to predict shock severity in other critical conditions using HL models with integrated EMR data. It could guide device use in heart failure or trauma, nudging clinicians for early intervention. EMR integration could also standardize care protocols across specialties, enhancing outcomes and device adoption in diverse hospital settings, leveraging real-time data ecosystems.

Engineering | Oil and gas | Network optimization | K-nearest neighbors

# Decarbonization of Gas Heating in Massachusetts: An Evaluation of Current Trends and Opportunities

## BUSINESS PROBLEM

The state of MA has ambitious carbon emissions reduction goals and the Department of Public Utilities has implemented regulations pushing the state's utilities to pursue decarbonization. National Grid operates a large gas distribution business and seeks to develop strategy for navigating a transition to a low-carbon future. In particular, the company seeks to understand and plan for electrification of gas heating which makes up a significant portion of gas demand in the state.

## APPROACH

This problem was addressed by evaluating: 1. Existing electrification of gas hearing and the state incentives driving that adoption 2. Which gas asset infrastructure projects could be foregone 3. The geographic correlation between where individual incentives for electrification are being used, opportunities for gas asset reduction, and communities facing housing burden.

## DATA SOURCES

Customer database (National Grid); History of heat pump adoption (National Grid); Market segments (National Grid)

**Data Types and Format**

Static databases, time series data


Author: Andrew Epstein

## IMPACT

This study helped provide clarity on existing adoption of heat pump technologies and identified future opportunities related to the electrification of gas heating in MA

| | |
|---|---|
| **DRIVERS** | The regulatory environment made this a priority |
| **BARRIERS** | Data in disparate datasets, uncertainty in future costs of electrification, uncertainty in the shape of future demand curves |
| **ENABLERS** | Welcoming colleagues at National Grid facilitated access to information |
| **ACTIONS** | A detailed modeling exercise of where existing adoption is occurring and where future benefits might be best located |
| **INNOVATION** | I built a cost/benefit model for evaluating traditional gas distribution infrastructure investments against investing in electrification of customers who would otherwise be served by those assets |
| **IMPROVEMENT** | No quantifiable improvement yet, but new approaches may be incorporated in typical processes in the near term |
| **BEST PRACTICES** | Ran sensitivity studies to understand impact of input assumptions across a range of reasonable values |
| **OTHER APPLICATIONS** | Other utilities, other environmental tradeoffs, etc. |

Supply chain planning and inventory management | Consumer packaged goods / Food and beverage | Warehousing automation / flow optimization | Discrete / Continuous

# Simulation of Sortation Centers To Improve Productivity and On-Time Delivery

## BUSINESS PROBLEM

Target's Sortation Centers face operational challenges driven by volatility in inbound package volume, reliance on manual labor, and inconsistent processes across sites. These factors often result in unproductive labor hours, delayed deliveries, and elevating operational costs. As demand continues to grow and expectations for fast delivery rise, the business aims to improve throughput and on-time delivery while reducing labor inefficiencies and minimizing the risk of volume carryover. These improvements must be achieved without significant capital investment, relying instead on smarter process design and better use of existing resources.

## APPROACH

A discrete event simulation was built and validated using SimPy to replicate Sortation Center operations. Leveraging operational data and site observations, process improvements were tested, like TLMD pre-sorting at store, to quantify their impact on productivity and on-time delivery performance.

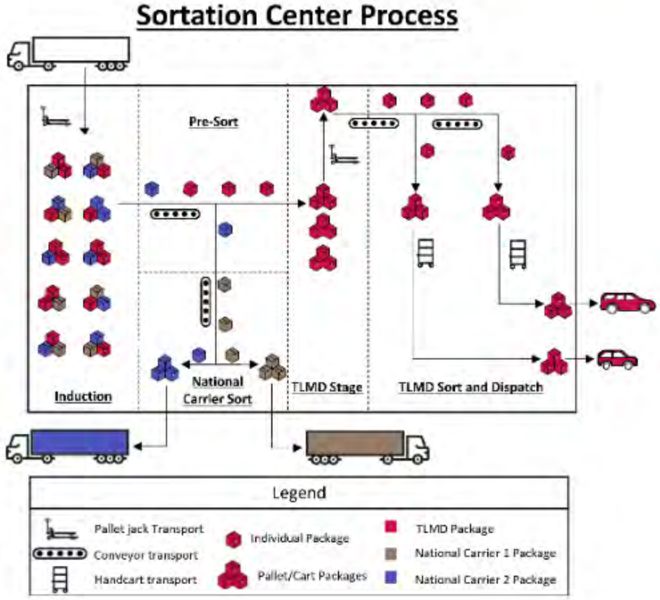
## DATA SOURCES

Operational data from Target's Sortation Centers, including package volume, labor hours, truck arrivals, and carrier distributions. Data sources included internal shift reports, timestamp logs, and in-person site observations.

**Data Types and Format**

Structured CSV data, shiftily operation logs, performance rates, staffing levels, and historical package-level tracking data from internal systems and observations.

**Sortation Center Process**




Author: Andrew Fenstermacher

## IMPACT

The simulation showed that implementing a TLMD pre-sort at the store level reduced overall processing time by 5.8% and cut carryover probability by up to 85% in unplanned high-volume excursion scenarios. The simulation also showed that likely due to the short batch style execution of the processes that the effects of process variance are low. These changes improve the system's resilience to demand volatility without requiring large capital investments or infrastructure overhauls. More generically, the change shifts the idea of process optimization from the sortation center alone to include the stores. The discrete event simulation enables Target to virtually trial process changes, offering a low-risk, data-driven way to improve Sortation Center performance. By adopting these recommendations, Target can standardize processes across its network and enhance service reliability, strengthening its position in the highly competitive last mile delivery space.

**DRIVERS**

The rapid rise of e-commerce and consumer expectations for fast delivery has shifted traditional retailers like Target to adopt last mile operations for quicker delivery times to better serve their guests. High variability, manual processes, and delivery delays in Sortation Centers created a need for data-driven process design and network-level optimization.

**BARRIERS**

Limited standardization across Sortation Centers, inconsistent data capture methods, and seasonal constraints on piloting new processes posed challenges. Non-standardized manual operations, time based variable shift headcounts, and unpredictable demand patterns added complexity to model validation and implementation.

**ENABLERS**

Target's Last Mile Delivery startup-like culture, openness to collaboration, and access to operational and field team members enabled site visits and access to all necessary data. Support from stakeholders across last mile delivery and supply chain teams fostered collaboration and insight.

**ACTIONS**

A SimPy-based discrete event simulation was built using operational data, validated it with historical performances, and ran experiments on process changes like TLMD pre-sorting and fluid loading. Results were shared with leaders to inform further investigation for future implementation.

**INNOVATION**

Using discrete event simulation to model high-variance manual processes at scale provided a virtual test bed for network-wide improvements. The approach quantified outcomes without disrupting operations and helped prioritize low-cost, high-impact changes. The model implements the current stochastic nature of inbound trucks and packages to determine a reasonable range of daily outcomes and timelines.

**IMPROVEMENT**

Having same day TLMD packages pre-sorted at stores could help to reduced carryover probability by up to 75% during significant volume surges and improve throughput by up to 5.8%. These results enhance delivery reliability and reduce operational costs with no major infrastructure changes, and minimal store based process changes.

**BEST PRACTICES**

Use detailed observations to inform model logic, validate simulation against real operations, and run matched experiments for each scenario. Engage cross-functional teams early and focus on scalable, low-friction process improvements. Simplify the model, but capture the necessary complexity to show potential improvements.

**OTHER APPLICATIONS**

This simulation framework can be applied to other warehouses, fulfillment operations, cross-docks, or manufacturing environments to test labor strategies, automation investments, or layout changes before real-world piloting or implementation.

OT | Retail | Improved process performance | Natural language processing

# A Use Case of Generative AI in Operational Technology Problem Management

**Amazon**

## BUSINESS PROBLEM

Amazon has deployed a highly technical and complex OT/IT infrastructure in almost every site in their global operations, heavily relying in its availability to drive real time optimization of their supply chain and give its customers fast and affordable deliveries. However, this complicates the root cause identification for each of their downtime events, usually leaving this task to specialized problem managers who incur in ad-hoc manual analysis which delays the potential identification and deployment of improvement projects to reduce the impact of this events.

## APPROACH

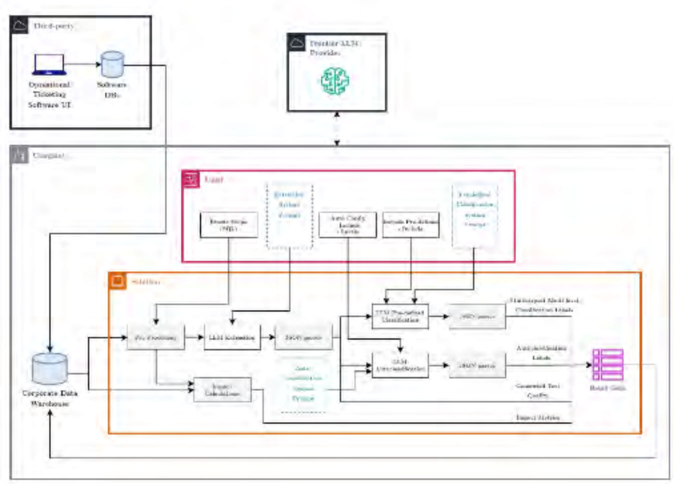
Implement an LLM enhanced software solution that automatically joints unstructured data from the company's data warehouse and process it through different LLM-based tasks (LLM-Extraction, LLM-Autoclassification, LLM-classification, and LLM-accuracy) to provide structured information that can be easily consumed and analyzed by problem managers.

## DATA SOURCES

Downtime event general data (from internal applications)
Incident management timeline (from internal applications) Post-mortem analysis report (from internal applications)

### Data Types and Format

Mostly unstructured text (from both unstructured sources and inside fields in structured tables)




Author: Gabriel Andres Gallardo

## IMPACT

Allowing problem managers to directly execute aggregated analysis on historical transformed data, accelerating the improvement opportunities identification and deployment, at 1% of the cost. It also allows us to granularly identify missing data in past events timelines and reports, making it possible to implement frequent feedback control over field and infrastructure engineers to gather more details about the recent events instead of letting it empty until manual analysis months or years later.

**DRIVERS**

Increasing availability of Generative AI end-user applications with decreasing usage cost

**BARRIERS**

Entry difficulty to internal software development and continuous deployment standards.

**ENABLERS**

Full access to any type of LLM through already prepared APIs

**ACTIONS**

Software implementation was needed to test the solution as a whole Cross validation sessions with subject matter experts to improve evaluation criteria and generate the ground truth Anonymization steps to be able to use external frontier LLMs

**INNOVATION**

The usage of the latest LLMs available in a traditional streamlined process

**IMPROVEMENT**

Reduction of 99% in the cost of processing downtime events unstructured data Reliability between 0.98 and 1.00 Accuracy between 0.89 and 0.96 Consistency between 0.79 and 0.99

**BEST PRACTICES**

Applying a complete evaluation framework measuring reliability, accuracy, and consistency. With a human always in the highest level. Identifying and quantify the failure modes for each of the LLM tasks in order to better focus improvement efforts.

**OTHER APPLICATIONS**

Expand the LLM-task for other characterization and classification tasks different from downtime events, for example high volume low impact events. Develop new LLM-tasks for additional downtime events characterization, like time-to metrics, impacted sites identification, troubleshooting process improvement opportunities, and logic assessments on the reported unstructured data.

Logistics Management | Automobile manufacturing | Warehousing automation / flow optimization | Deterministic / Stochastic

# Minimizing Cost of Intra-Yard Finished Vehicle Logistics Through Automation and Optimization

## BUSINESS PROBLEM

An OEM seeks to transform its finished vehicle logistics through automation, where newly manufactured vehicles drive themselves within the plant environment using Vehicle Plug-In (VPI) units. However, the company faces significant uncertainty in optimizing these operations. Key challenges include determining optimal VPI inventory levels (initially estimated at 300 units), identifying potential bottlenecks in the workflows, validating safety performance against human drivers, and evaluating whether mid-lifecycle LIDAR sensor upgrades would be cost-effective. These unknowns create financial planning challenges and operational risks.

## APPROACH

The research applies a three-layered systems engineering methodology: operational modeling using FlexSim simulation to optimize VPI inventory and replenishment cycles; safety analysis using Bayesian Networks to validate performance against human benchmarks; and technology evaluation to assess sensor architecture and potential upgrades throughout the system lifecycle.

## DATA SOURCES

Two-week yard movement data from the OEM's Yard Management System tracking vehicle movements; direct observations from the OEM's R&D team showing autonomous technology performance; commercial research on sensor technologies and tow truck specifications; and stakeholder interviews with FVL, Mobility and Innovation, and technical teams providing operational context and requirements.

**Data Types and Format**

Time-series vehicle movement data from the OEM's Yard Management System; autonomous system performance probability metrics; sensor and tow truck cost estimates; interview notes and flow diagrams.


Author: Jeremy Garber

## IMPACT

The three-layered systems engineering analysis validates the system's operational viability while revealing critical implementation parameters. Operational modeling demonstrates that 750 VPI units (significantly higher than initial 300-unit estimates), with 6-hour replenishment cycles, minimizes cost considering replenishment labor and depreciation factors. This modeling identifies a key bottleneck in stock lot VPI collection, where the current individual retrieval approach extends replenishment time to nearly two hours. Safety analysis using Bayesian Networks confirms system performance exceeds human driver benchmarks with 99.999% probability of accident-free operation, providing quantifiable risk metrics for decision-making. The technical evaluation establishes LIDAR-equipped RSUs as the optimal sensor architecture while determining that mid-lifecycle sensor upgrades yield a slightly negative NPV, indicating upgrades may become viable with further technology improvements. The implementation adds only 4.6% capacity impact to the existing End-of-Assembly Line (EOAL) workforce, minimizing operational disruption. Beyond immediate benefits, this framework positions Nissan for predictable intra-yard automation, and establishes a foundation for expanding autonomous systems throughout their manufacturing ecosystem. In effect, creating a bridge between controlled-environment autonomy and broader mobility applications.

**DRIVERS**

Industry trends toward manufacturing automation and the growing Automated Valet Parking market (projected to reach $42.38 billion by 2028) created competitive pressure and opportunity. The OEM's leadership mandated increased automation and asset optimization while creating higher-value jobs, providing executive sponsorship for implementation.

**BARRIERS**

Acquisition delays for an initial simulation tool required shifting to FlexSim after legal reviews. Computing power limitations prevented using actual OEM CAD files in simulations. Incomplete tracking data for accessory install operations required developing scaling factors for accurate processing time estimation. Geographic dispersion of teams created communication challenges affecting information gathering and stakeholder alignment.

**ENABLERS**

The controlled manufacturing environment avoided regulatory challenges of public-road implementation. The OEM's R&D team provided autonomous system expertise while yard operations staff offered process insights. MIT's operations management frameworks and FlexSim's simulation capabilities supported complex modeling, while leadership's automation vision provided organizational support.

**ACTIONS**

Developed a three-layered systems engineering approach to validate Nissan's automated yard operations. Conducted yard movement data analysis using Python, created FlexSim simulation models for VPI inventory optimization, implemented Bayesian Network safety analysis to quantify accident probabilities, and evaluated sensor technologies through technical assessment and NPV analysis.

**INNOVATION**

The integrated three-layered systems engineering approach connects operational modeling, safety analysis, and technology strategy to optimize multiple dimensions simultaneously. The modified inventory model extends Economic Order Quantity theory to account for circular VPI flow in autonomous logistics. The Bayesian Network safety validation provides quantifiable metrics directly integrating financial implications for NPV calculations.

**IMPROVEMENT**

Analysis established optimal parameters: 750 VPI units with 6-hour replenishment cycles maintaining 100 vehicles/hour throughput—far above initial 300-unit estimates. Safety validation confirmed 99.999% probability of accident-free operation, exceeding human benchmarks. Validated the implementation requires only 4.6% capacity impact to existing workforce at End-of-Assembly-Line (EOAL).

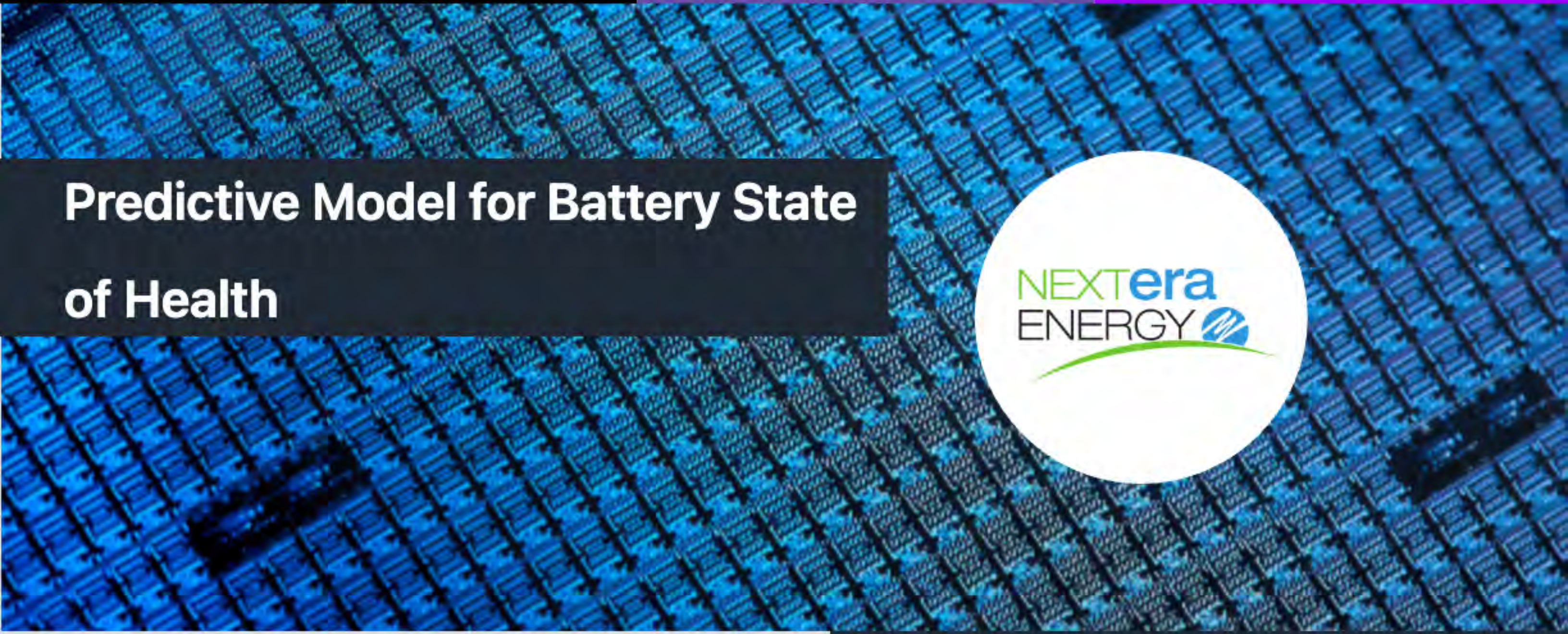
**BEST PRACTICES**

Begin with detailed process mapping before simulation development. Develop comprehensive stakeholder engagement spanning operational, technical, and management teams. Use multi-method validation combining simulation, probabilistic modeling, and financial analysis. Structure models with modular components for easier modification.

**OTHER APPLICATIONS**

The methodology can also apply to inbound logistics management, parts delivery, and carrier loading in automotive manufacturing. Similar approaches benefit port operations, distribution centers, and fulfillment operations. Beyond automotive, applications include pharmaceutical manufacturing logistics, airport baggage handling, and mining operations—all controlled environments suitable for autonomous vehicle technology.

Engineering | Energy | Predictive forecasting | Reinforcement learning



# Predictive Model for Battery State of Health

## BUSINESS PROBLEM

The main problem the business is trying to solve is the accurate estimation of the State of Health (SoH) of battery sites to inform and optimize their augmentation strategy. This is crucial for preventing both over- and under-augmentation of batteries, which can lead to unnecessary costs or insufficient capacity to meet contractual terms. With an accurate real-time SoH estimation, the business can plan timely augmentation and maintenance, minimize costs related to on-site capacity tests and associated downtime, and reduce expenses from suboptimal augmentation, ensuring efficient resource allocation and cost-effective operations.

## DATA SOURCES

The data came from NextEra's central server, sourced from Battery Management Systems and on-site sensors. Additional business inputs for the case study came from the Business Management team and the Power Generation Division.

### Data Types and Format

The data was time series format for current, voltage, temperature, State of Charge (SoC), and State of Health (SoH) from multiple battery sites.

## APPROACH

I utilized four models: a Kalman Filter, Long Short-Term Memory Recurrent Neural Network (RNN), Multitask RNN and Delayed Reinforcement Learning to predict the State of Health of battery sites using operational data such as current, voltage, temperature, and State of Charge. To evaluate the accuracy of these models, I validated predictions using internal capacity tests.


Author: Catalina Garza Lozano

## IMPACT

By providing accurate, real-time SoH assessments of utility-scale battery systems, the model enables more precise planning for augmentation, reducing both over- and under-augmentation. This leads to substantial capital expenditure savings by delaying augmentation until necessary and taking advantage of falling battery prices. It also minimizes operational inefficiencies and helps avoid penalties tied to underperformance in Power Purchase Agreements. The Delayed Reinforcement Learning model, in particular, achieved a 1.6-month prediction error—far outperforming the current Battery Management System (BMS), which had a 58-month error. Applied at scale, this level of accuracy could generate NPV savings in the millions across the fleet. Ultimately, the model supports smarter asset management, greater financial efficiency, and stronger compliance with contractual obligations, positioning NextEra to better manage the evolving demands of energy storage.

### DRIVERS

The rapid growth of utility-scale energy storage, falling battery costs, and the performance obligations in long-term PPAs created a need for better battery health forecasting. The industry is facing challenges in planning augmentations efficiently due to limited SoH visibility, risking unnecessary costs or compliance issues. This gap catalyzed the development of a predictive SoH model to enable smarter, data-driven decision-making.

### BARRIERS

Barriers included varying data quality across sites, limited capacity test data availability for some locations, and inconsistent data formats that required preprocessing. Additionally, training and tuning machine learning models demanded significant computational power, which limited the ability to rapidly test and iterate on different model architectures and hyperparameters.

### ENABLERS

The company's scale and access to extensive data and resources enabled the project, along with a strong culture of collaboration. My team sat between engineering and supply chain, giving it broad visibility. Many team members had both consulting and engineering backgrounds, which helped bridge technical and business perspectives effectively throughout the project.

### ACTIONS

To implement the solution, I learned how capacity tests work and how SoH is measured through them. I reviewed past approaches within the company to understand what had or hadn't worked. I researched SoH models in the literature and explored various machine learning algorithms, studying their implementation and identifying the best ways to run them efficiently in our context.

### INNOVATION

The solution's innovation lies in applying machine learning to predict battery SoH using the daily operational data from the sites. The Delayed Reinforcement Learning model yielded interesting results and shows strong potential to be further developed for SoH prediction.

### IMPROVEMENT

The final improvement was a significant increase in the accuracy of State of Health predictions. The best-performing model reduced prediction error from 58 months (using BMS data) to just 1.6 months. While there is still more work to be done, this improvement enables better augmentation planning, reduces unnecessary costs, and helps ensure compliance with long-term performance obligations.

### BEST PRACTICES

To replicate this solution, start by understanding how SoH is measured and ensure access to clean, consistent time series data. Review prior internal attempts and relevant literature to build on existing knowledge. Test model types on a small scale, validate against real capacity test results, and balance computational complexity with practical usage. Collaboration across technical and business teams is key to delivering actionable outcomes.

### OTHER APPLICATIONS

Beyond battery augmentation planning, this solution could be applied to warranty risk assessment, asset valuation, and predictive maintenance. It could also support portfolio-level planning, grid reliability analysis, and integration with energy trading strategies by providing more accurate, real-time insights into battery performance and degradation.

Supply chain planning and inventory management | Aerospace and defense | Inventory management / control | Classifiers

# Optimizing Raw Wire Inventory Management: A Data-Driven Approach to Demand Forecasting and Supply Chain

## BUSINESS PROBLEM

The main problems that the business aims to solve raw are wire inventory management issues to prevent shortages and achieve a 100% service rate. Key influences include COVID-19 supply chain disruptions, high production variability from customization, and expertise loss from retirements. Late bundle designs and long lead times (26-52+ weeks) complicate forecasting, causing shortages or overstocking. Traditional methods fail in this complex environment. The goal is to use data-driven tools to predict demand, optimize inventory, and ensure production efficiency, minimizing costs and delays. Improving communication is a parallel effort.

## DATA SOURCES

The thesis used live production & supply chain data: consumption from cutting machines (69% of 814 wire types), demand from production rates and work statements, inventory from off-site warehouse reports, and supply data from interviews. Spanning 2020-June 2024, it was collected from manufacturing systems, normalized for confidentiality, and analyzed to improve raw wire inventory management.

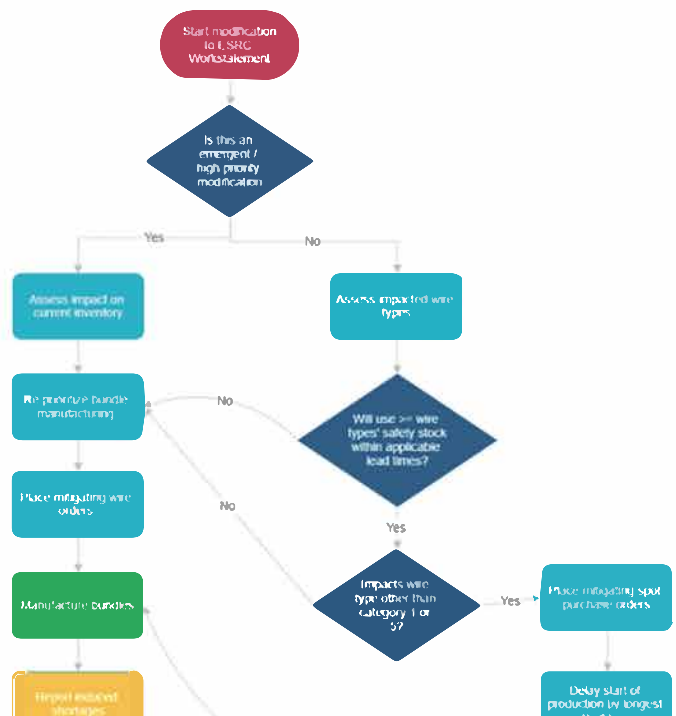
### Data Types and Format

Data included consumption (time series, numeric, database records), demand (time series, numeric, rates/statements), inventory (snapshot, numeric, reports), and supply (qualitative, text, interviews),

Author: Adam Gebner

## APPROACH

The thesis tackles Boeing's ESRC inventory issues with a data-driven approach: analyzing production data for demand patterns, using K-means clustering to categorize wire types, developing demand and inventory simulation models, and creating a decision support tool. It aims to improve forecasting accuracy and communication, despite data gaps, to optimize inventory and achieve a 100% service rate.



## IMPACT

The solutions proposed, though not fully predictive, impacts the business positively. The demand and inventory simulations underperformed, but the research offers insights and tools to boost supply chain resilience. The data analysis conducted clarifies demand patterns and waste (averaging 12%), and informs tailored inventory strategies for different wire types via K-means clustering. This could reduce shortages and optimize stock, cutting costs indirectly. The management decision support tool enhances communication, preempting inventory risks from work statement changes and reducing the bullwhip effect. This may prevent production delays, potentially saving thousands daily, though data gaps limit quantification and validation. While not directly yielding savings, the solution sets a foundation for improvement. Future steps like collecting better data, specifically regarding shortages, and advanced analytics could refine forecasting, aligning the supply chain with the build cycle employed to fulfill the customization needs. This work shifts the organization to proactive, data-driven management, with clearly identified channels of communication, promising efficiency and reliability gains long-term.

**DRIVERS**

The solution was driven by the facility facing supply chain disruptions from COVID-19, volatile production due to aircraft customization, and expertise loss from retirements. These factors exposed inventory vulnerabilities, with long lead times (26–52+ weeks) and late bundle designs clashing with the need for a 100% service rate, pushing for data-driven tools to predict demand and optimize raw wire inventory.

**BARRIERS**

Barriers impacting the project included incomplete bundle design data (retained only 6 months), missing consumption records for 25% of wire types, no verifiable shortage data, limited lead time info from interviews, and desynchronized production rates. These gaps hindered accurate demand forecasting and inventory simulation, limiting the ability to validate models and assess shortage risks effectively.

**ENABLERS**

Boeing's extensive production data (2020–2024) from the ESRC, including consumption and inventory records, enabled analysis. Team support from Boeing staff, access to software (Python, Excel), and the Leaders for Global Operations program facilitated tool development. Qualitative insights from interviews and a data-rich environment helped overcome gaps, supporting the creation of actionable inventory management solutions.

**ACTIONS**

I analyzed ESRC data using Python, developed a demand simulation model in excel, and built an inventory simulation tool. I applied K-means clustering to categorize wire types, created a decision support flowchart, and built a Tableau dashboard with the IEs for visibility. I conducted interviews, validated findings qualitatively, and recommended data improvements and pilot studies to refine forecasting and inventory management at Boeing.

**INNOVATION**

Innovative aspects include a data-driven demand simulation surpassing baseline methods, K-means clustering for wire categorization, and a decision support tool with a Tableau dashboard for real-time visibility. Integrating qualitative insights with quantitative models to address customization challenges and proposing future machine learning enhancements offer novel ways to tackle Boeing's complex inventory issues.

**IMPROVEMENT**

The solution didn't quantify savings due to data gaps, but it improved demand understanding and waste quantification (12% average). The decision support tool enhances communication, potentially reducing shortages. Simulations, though inaccurate, laid groundwork for future accuracy. Qualitatively, it promises fewer delays, saving thousands daily, pending better data and pilot studies for measurable impact.

**BEST PRACTICES**

Ensure comprehensive data collection (e.g., shortage records, full design specs), integrate qualitative insights with quantitative models, use clustering for demand patterns, and develop interpretable tools like dashboards. Validate with pilot studies, leverage team expertise, and iterate models with real-time data. Address data gaps early and align with business goals for effective inventory management replication.

**OTHER APPLICATIONS**

The solution could apply to other high-mix manufacturing (e.g., automotive, electronics) for inventory optimization, aerospace suppliers for demand forecasting, or any supply chain with customization challenges. The decision support tool and clustering approach could enhance logistics in healthcare (e.g., medical device

Production | Aerospace and defense | Improved process performance | Value Stream Mapping

# Economies of Space: Developing a Lean Manufacturing Framework for Work Center Floorspace Reduction

## BUSINESS PROBLEM

The main problem Boeing's Interiors Responsibility Center South Carolina (IRC SC) is trying to solve is the fully allocated production floorspace. This limitation prevents the addition of new product lines and hinders preparation for potential increases in production rates for the 787 Dreamliner Program. The goal is to eliminate wasted floorspace while increasing production capacity and efficiency without expanding the facility's physical footprint.

## APPROACH

We developed a lean manufacturing framework using the DMAIC methodology which employed lean tools like spaghetti diagramming and value stream mapping to identify and eliminate wasted space. The framework was tested on a single IRC SC work center and recommended changed were validated using discrete event simulation.

## DATA SOURCES

The data in this thesis includes detailed observations, time studies, and value stream maps from IRC SC. These data were collected through direct measurements, spaghetti diagramming, and discrete event simulation. The sources include Boeing's internal systems, process plans, and input from manufacturing technicians and industrial engineers.

### Data Types and Format

Data formats included Excel spreadsheets, manufacturing execution system production data, CAD models, Simio simulation models, and Visio documents.


Author: Jacob Gerbino

## IMPACT

The developed framework resulted in 25% reduction in floorspace and a 55% decrease in product throughput time in the studied work center. The lean manufacturing principles and scheduling optimizations discussed is expected to be applied to other work centers within IRC SC to further economize floorspace and streamline production processes in advance of potential rate increases. Additionally, discrete event simulation validated the proposed changes, ensuring they would achieve the desired improvements. The simulation demonstrated that the new system could perform satisfactorily even at higher production rates. This provides confidence that the framework can support future increases in production rates without compromising efficiency or quality. Overall, the impact addresses both immediate operational challenges and long-term strategic goals. By optimizing floorspace, improving efficiency, enhancing worker conditions, and supporting higher production rates, the framework positions IRC SC for continued success and growth in the competitive aerospace manufacturing industry.

**DRIVERS**

Boeing's IRC SC supports the 787 Dreamliner program and needs to be able to meet any potential new production or aftermarket demands. As such, the facility needs to perform advanced planning for such increases. Additionally, new product lines may be needed to meet supply chain priorities or regulatory requirements, and the facility needs to free up space on the production floor for these lines.

**BARRIERS**

Detailed data for high-confidence process time is not well captured in the manufacturing execution system. data collection required manual time-study data collection. Additionally, internal and regulatory requirements regarding product designs and materials constrained the possible changes to the work centers and processes.

**ENABLERS**

IRC SC provided broad access to personnel with in-dept knowledge of studied product lines as well as internal systems which contained information on processes, standards, and historical production data. IRC SC also ensured access to production floor and interaction with production teams.

**ACTIONS**

Actions taken to implement the solution included coordination with tooling department to support new equipment requirements, development of transition plans for floorspace and scheduling changes, as well as presentation of project methodology to the Lean Manufacturing and Industrial Engineering for further work in IRC SC and other Boeing facilities.

**INNOVATION**

This project developed innovative discrete scheduling methods to level-load daily production times. We also developed and constructed new solutions to integrate work-in-progress (WIP) storage that reduced the need for independent shelving and allowed segregation of production materials between shifts.

**IMPROVEMENT**

This project achieved major reductions in floorspace and product throughput times for the studied work center. Additionally the project achieved a lean framework for IRC SC to apply to other work centers to reclaim additional floor for future production needs.

**BEST PRACTICES**

Inclusion of production team in floorplan and process changes leads to improved and more sustainable outcomes. Consolidate of curing times when possible to reduce the amount of touch-days a product requires. Shift production scheduling from batch-and-queue to single part flow to reduce through-put time and WIP storage requirements.

**OTHER APPLICATIONS**

The framework outlined in this thesis should be applicable in most manufacturing contexts, especially in cases where production is carried out manually at movable workstations rather than on moving production lines.

Operations | Pharma / biopharma | Network optimization | Network analysis

# Simulation Modeling of Drug Substance Tech Transfer Timelines at Amgen

AMGEN®

## BUSINESS PROBLEM

Commercial tech transfers (TTs) are critical to scaling up and transferring the production of Amgen's drug portfolio as drugs are developed & acquired, new manufacturing sites are opened, and biomanufacturing innovations are made. Such TTs are highly complex with many interdependent teams, allowing bottlenecks to form in the system that slow transfers and suboptimally utilize resources. Historical data for executed timelines are also not centralized, creating challenges for quantitative analysis. Thus, Amgen is interested in modeling the TT process in silico to experiment, quantify process optimizations, and motivate changes to TT workflows.

## DATA SOURCES

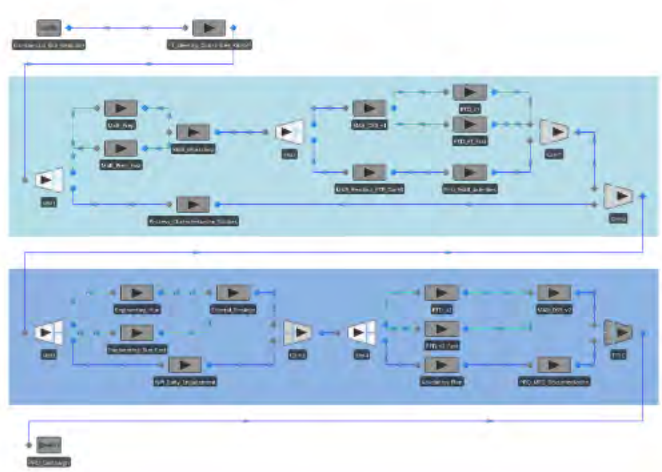
I had access to data on Amgen's previously completed tech transfers, which were available through the Medusa database centralizing transfer milestone data. I also could access process flow data from SOPs and previously constructed business process maps, which I was provided by SMEs across Process Engineering and Manufacturing as inputs to my model design.

**Data Types and Format**

Process flowcharts for TT progression, estimates for average task duration and stochasticity

## APPROACH

I used a business modeling software called Simio to reconstruct the network of tasks involved in high complexity commercial TTs and NPIs at Amgen, with each task being assigned an average time to completion and variability to model process stochasticity. I then run digital experiments in Simio of the base model and of altered network architectures to quantify model performance.



Author: Viraat Goel

## IMPACT

This project focused on the development of Simio models for high complexity commercial tech transfers in Amgen's drug substance division, as well as New Product Introduction models at Amgen's manufacturing sites. As a biotechnology giant with extensive in-house biomanufacturing expertise, Amgen is reliant upon successful TTs and NPIs to commercialize its suite of biologic products at scale, without error, and with regulatory compliance; as a result, these are vital business processes that merit analytical consideration for potential optimization. The models built in this thesis represent an analysis of system efficiencies in completing TTs and NPIs in a timely manner, with simulated experiments designed and run for each model testing how changes to the system affected the overall timeline. These models, developed iteratively through conversation with SMEs and trial-and-error, are intended to support decision-making at Amgen as the company seeks to optimize its business processes in an informed manner. The system perturbations tested in this project varied in their influence on overall timelines, highlighting various task optimizations and system designs that can accelerate TT completion. Thus, the models and analyses trialed in this thesis de-risk process optimization, and they can hopefully bridge process optimizations from early-stage hypothesis generation to an experimentally-informed process alteration worthy of testing in a real-world scenario.

### DRIVERS

Amgen needs to be able to commercialize and scale up its drug portfolio quickly if it wants to compete with a fast-paced biopharmaceutical industry. With major research and capacity-building investments underway and tech transfers on the horizon, Amgen needed to optimize its tech transfer business process to minimize inefficiencies and bring their new drug portfolio to market ASAP.

### BARRIERS

Amgen's manufacturing sites are still fairly decentralized and own most of their own processes; as a result, divergence often emerged between processes documented in SOPs and the actual business process conducted by one of Amgen's sites. Moreover, model-building required assumptions on the appropriate granularity with which to model a system, which created challenges for still modeling systems accurately.

### ENABLERS

There was plenty of inertia at Amgen surrounding tech transfer business process optimization, allowing me to benefit from materials recently generated by my team and easy generate buy-in for project questions & support from SMEs across Amgen Operations.

### ACTIONS

I iteratively built complexity into my models, designing them "from the ground up" to more closely represent the tech transfer workflow. During this process, I incorporated feedback from my supervisor, team, and adjacent SMEs regarding pain points in the tech transfer process and potential priorities for task-level optimization.

### INNOVATION

The application of Simio for timeline modeling of an abstract business process was new for Amgen. Generating a digital process model that could then be perturbed, simulated, and analyzed generated data useful for hypothesis generation about alterations to the real-world process, and this approach should be generalizable across a wide swathe of Amgen's Operations.

### IMPROVEMENT

My modeling identifies improvements on the order of weeks to months for various process optimizations, reducing commercial tech transfer completion times by 10-20%. Modeling comparisons also identify how closer communication between tech transfer and new product introduction workflows could reduce overall transfer time.

### BEST PRACTICES

Best practices to follow when conducting this kind of modeling are to consult thoroughly with SMEs familiar with the processes being modeled and to rigorously test assumptions by altering process architecture, stochasticity, and granularity. It is also prudent to design experiments with specific, actionable hypotheses for business process improvements in mind.

### OTHER APPLICATIONS

Other applications of my project include optimizations for any of Amgen's operating practices across its many divisions. Though applied specifically here for tech transfer optimization, creating a digital twin model for a business process forces the model designer to understand the base process, its interdependencies, and potential inefficiencies, and virtual experimentation allows costly process changes to be trialed quickly & inexpensively.

Sales | Automobile manufacturing | Predictive forecasting | Regression analysis

# Forecasting Automotive Production Volume Using Regression and Time Series Modelling

AMERICAN INDUSTRIAL PARTNERS

## BUSINESS PROBLEM

Automotive suppliers face significant challenges due to inaccurate vehicle production volume forecasts, often overly optimistic, either provided by Original Equipment Manufacturers (OEMs) or from industry benchmark forecasts. These inflated forecasts cause misallocation of production capacity, leading to both direct financial costs from underutilized resources and indirect opportunity costs by diverting capacity away from other market segments. The need for accurate forecasting is intensified by industry volatility and transformative trends like electrification and changing mobility patterns.

## DATA SOURCES

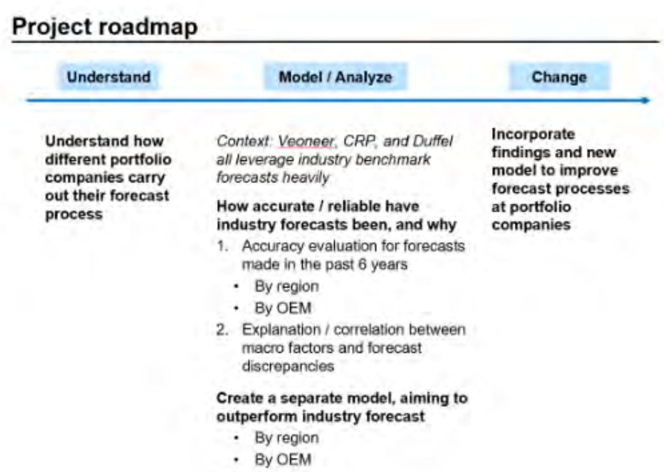
Primary data source was a third-party company who publishes monthly bottom-up automotive production forecasts with retrospectively adjusted actual production volumes dating back to 2000 and future forecasts extending seven years ahead. Macroeconomic indicator data was obtained from Trading Economics, including GDP growth, employment changes, inflation rates and other economic metrics.

### Data Types and Format

Data comprised numerical, monthly and quarterly time-series datasets reflecting historical automotive production volumes and macroeconomic indicators, structured in tabular format.

## APPROACH

Advanced statistical forecasting methods are evaluated, including ARIMA, linear regression, Lasso regression, Theta, and Boosted Theta models, using historical automotive production volumes and macroeconomic data. The Theta model was ultimately selected for implementation, balancing superior forecast accuracy with interpretability, implementation simplicity, and minimal data requirements.




Author: Yutao Gong

## IMPACT

The Theta model significantly improved forecasting accuracy compared to benchmark industry forecasts, reducing forecast error by 5-7 percentage points across multiple time horizons. This enhanced precision allows suppliers to make better-informed capacity allocation decisions, optimize customer mix across automotive and other segments, and improve capital allocation for facility expansions or equipment purchases. The solution was successfully implemented at Commonwealth Rolled Products where it directly influenced 2025 budget planning, with the sales team independently using the model to forecast key carlines rather than accepting potentially inflated customer projections. Financially, improved forecasting reduced potential losses from unused capacity, with millions of missed opportunity avoidance. Beyond immediate financial benefits, the implementation catalyzed a cultural shift toward more data-driven decision-making. A user-friendly Google Colab interface ensures sustainable adoption by non-technical users, while designated model owners at both portfolio companies and AIP maintain long-term viability. This practical application of advanced analytics provides automotive suppliers with a more reliable planning tool to navigate market volatility and evolving industry dynamics. It also offers valuable insights for future acquisitions and portfolio management at American Industrial Partners, through aligning business activities more closely with actual market demand.

**DRIVERS**

The major catalyst is the tendency of overestimation in OEM and industry forecasts, which created business planning challenges for suppliers who risked reserving manufacturing capacity for automotive demand that would never materialize. Increasing market volatility and structural shifts such as electrification amplified the need for more reliable forecasting solutions.

**BARRIERS**

One barrier is organizational inertia in shifting away from traditional forecasting methods, as they are industry norms and have been well-incorporated into internal processes of suppliers. Implementation barriers included limited technical expertise among industrial manufacturers, which not only increases the requirement for user-friendly interfaces, but also presents challenges in convincing users to adopt the new forecasting method.

**ENABLERS**

Active support from AIP Operations team helps to identify strong internal champions at portfolio companies, who have higher level of technical expertise and openness to adopting data-driven practices. Establishing a close, collaborative relationship with these champions facilitates model demo sesssions, builds trust among potential users of the model at portfolio companies, and eventually facilitates effective implementation.

**ACTIONS**

The implementation process began with developing a user-friendly Google Colab notebook that required no programming knowledge to operate, and provided interactive data visualizations for decision-making clarity. A champion was identified at Commonwealth Rolled Products to take ownership of the model. Comprehensive training was provided to this champion and subsequently extended to key account managers across the organization.

**INNOVATION**

The research balanced statistical rigor with implementation practicality, evaluating models not only on accuracy but also interpretability, complexity, and data dependencies. The Theta model's application to automotive forecasting represented an innovative crossover from general time series forecasting to a specialized industrial context. The implementation approach of creating an accessible interface for non-technical users was also innovative.

**IMPROVEMENT**

The Theta model reduced forecast error by 5-7 percentage points compared to industry benchmarks over recent 5-year periods. For 1-year forecasts in North America 2020-2024, Theta's average deviation is 10% versus Benchmark's 15%. These improvements directly influenced 2025 budget planning at Commonwealth Rolled Products, with forecasts for a key carline adjusted ~30% down.

**BEST PRACTICES**

1) Measure first - quantify historical performance of different forecast sources before exploring new approaches; 2) Start simple - prioritize interpretable models over complex ones; 3) Embed in process - identify internal champions and develop user-friendly interfaces aligned with existing planning cycles; 4) Keep iterating - compare multiple forecast sources rather than relying on a single approach.

**OTHER APPLICATIONS**

The model could be extended to forecast parts and components demand throughout the automotive supply chain beyond finished vehicle production. The approach could also be applied to other manufacturing industries with similar forecasting challenges, such as consumer electronics or industrial equipment. Beyond manufacturing, the balanced approach to model selection and implementation could serve as a template for analytics projects in any industry.

Business Analytics | Telecom | Network optimization | Linear / Non-Linear

# Expanding Home Broadband Coverage Through Existing Low Earth Orbit Megaconstellations

## BUSINESS PROBLEM

Internet Service Providers (ISPs) face high infrastructure costs when expanding broadband into rural and underserved areas, making traditional solutions financially unviable. Verizon seeks a scalable and cost-effective alternative to close the digital divide.

## APPROACH

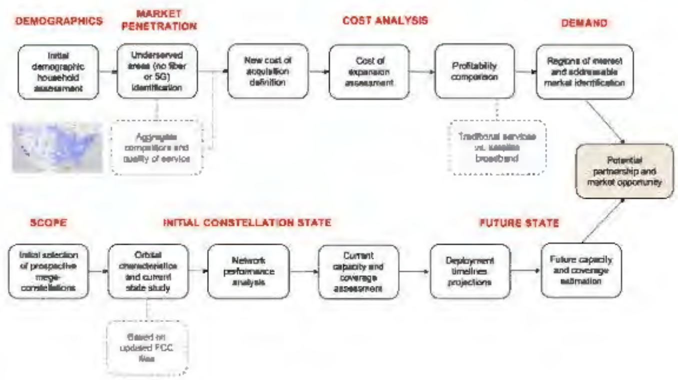
This project assessed the feasibility of leveraging Low Earth Orbit (LEO) satellite megaconstellations as an alternative broadband solution. It developed a customer demand model and a cost optimization framework to compare the profitability of traditional offerings vs. satellite solutions across diverse geographies, along with a technical capacity analysis of different satellite providers.

## DATA SOURCES

The analysis leveraged publicly available FCC broadband coverage datasets and U.S. Census demographic data at the census tract level. Internal Verizon cost and market data were also used, anonymized for confidentiality. Satellite provider capacity data and deployment plans were drawn from FCC filings and industry reports.

### Data Types and Format

Geospatial demographic census data, product tabular cost, time-series deployment projections, and FCC datasets.




Author: Gretel M. Gonzalez

## IMPACT

The analysis identified 17 million potential customers in underserved areas, with satellite broadband emerging as the most profitable solution for at least 1 million households. With minimal capital investment and favorable wholesale agreements, satellite technology enables new customer acquisition and rapid broadband expansion, positioning Verizon for strategic growth and improved digital equity.

**DRIVERS**

The main driver was Verizon's need to expand broadband to underserved areas where traditional infrastructure investments are not cost-effective. The rise of LEO satellite megaconstellations presented a timely opportunity to assess whether they could offer scalable, complementary solutions.

**BARRIERS**

Barriers included limited internal data on satellite broadband, uncertain satellite deployment timelines, and the challenge of comparing technologies with different cost structures and coverage models.

**ENABLERS**

Support from Verizon's Impact Analytics team, along with access to public regulatory data and internal broadband cost structures, enabled efficient development of demand models and profitability analyses.

**ACTIONS**

Built a customer demand model, cost optimization framework, and technical capacity assessment to evaluate LEO satellite broadband profitability relative to existing offerings across U.S. geographies.

**INNOVATION**

The integrated decision-making tool combined demographic demand modeling with broadband cost structures and satellite capacity analysis—bridging strategic, technical, and financial considerations in a single framework.

**IMPROVEMENT**

The analysis provided Verizon with previously unavailable data to support strategic decision-making and partner negotiations. It identified 17 million underserved households across the U.S. and pinpointed high-opportunity regions for satellite broadband deployment. Notably, the study revealed that satellite is the most profitable solution for at least 1 million of these households.

**BEST PRACTICES**

Start by integrating publicly available data sources to identify the market opportunity. Normalize cost structures across technologies to enable fair comparisons. Maintain flexibility in your modeling framework to adapt to new technical inputs or deployment updates. Finally, ensure cross-functional collaboration to incorporate both strategic and technical perspectives.

**OTHER APPLICATIONS**

This methodology could apply to satellite-to-device connectivity, international broadband expansion, or other telecom product rollout strategies.

Operations | Med devices / MedTech | Inventory management / control | Discrete / Continuous

# Developing a Data-Driven Approach to Reducing Excess Inventory in a Multi-Echelon Supply Chain

## BUSINESS PROBLEM

Stryker Spine has deployed hundreds of millions of dollars of inventory into the field to satisfy demand. However, if inventory deployment and allocation is not managed meticulously to ensure that it is in the right place at the right time then excess inventory arises. Spine has massive amounts of excess inventory in the field.

## APPROACH

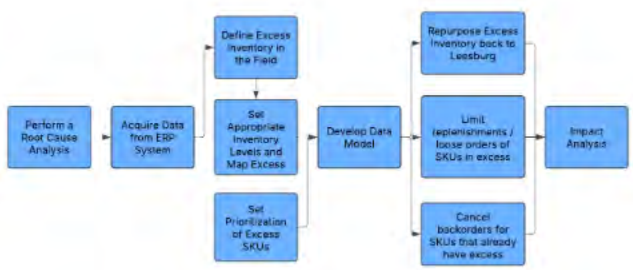
The objective of my thesis is to leverage a data-driven approach to define and reduce excess inventory at scale for the Stryker Spine business unit across the US. A robust data model / tool was developed to aid the internal Operations team to repurpose excess field inventory through the main distribution center and to automatically limit the deployment of inventory already in excess in the field.

## DATA SOURCES

The three main data sources for this research were Stryker Spine's Oracle ERP system called JD Edwards (JDE), internal Power BI dashboards, and internal Excel spreadsheets. Data was gathered at the SKU level and at the Set/Kit level related to historical sales/usage, surgery dates, orders, inventory on hand, demand, forecasts, location, $ cost, etc.

**Data Types and Format**

CSV files exported from JD Edwards, Excel and Power BI. Data types: Integer, String, Date/Time, Currency, Categorical




Author: Cayo Gosen

## IMPACT

By creating a scalable framework for defining, repurposing, and reducing implant excess inventory in the United States, Stryker Spine could unlock substantial amount of capital that is currently tied up, create flexibility in their supply chain to meet demand changes, and allow additional investment in innovation. Through the development of a data model and the implementation of three integrated workflows: repurposing excess to the distribution center in Leesburg, limiting replenishment of implants in overstocked locations, and canceling redundant backorders, this initiative targets multiple root causes of excess inventory. These workflows will be integrated into the current ERP system (JD Edwards), promoting automation, standardization, and long-term sustainability. The estimated financial impact includes the recovery of millions of dollars of excess inventory, reductions in unnecessary new purchases, and enhanced field inventory turnover. In addition to strengthening Stryker's competitive positioning in the spine implant sector, the project lays the foundation for a replicable model that can be extended to other business units across the enterprise.

**DRIVERS**

There is a large opportunity to leverage data to reduce the current amount of excess inventory in the field. A scalable and automated data-driven approach that could be easily integrated into Stryker's ERP system was necessary to alleviate the excess issue.

**BARRIERS**

Stryker recently organized their field inventory into cluster locations (groupings of hospitals) to improve the billing of implants; however, this removed the visibility of the exact location of each implant and instrument at the hospital level. Furthermore, for Legacy Stryker products it is unclear how much inventory is in sets vs loose parts as these products have never been tracked appropriately.

**ENABLERS**

Stryker is fully bought into solving the excess inventory problem. They are willing to invest in enhancing their ERP system and processes to reduce excess inventory at scale.

**ACTIONS**

Multiple methodologies to measure excess inventory in the field were examined and compared. One methodology was chosen, and it was incorporated into a data model. The locations of excess inventory were mapped across the US. The potential savings in $ value and units for excess inventory reduction were estimated.

**INNOVATION**

Establishing a scalable framework to measure excess field inventory, which can be integrated into the company's current ERP system to limit the deployment of unnecessary inventory, can generate significant savings.

**IMPROVEMENT**

Previously, Stryker did not leverage a data-driven approach to repurpose excess implant inventory from the field. Now, the internal Commercial Operations team can use the newly developed excess data model to potentially double the reutilization rate of implants once they are strategically sourced and repurposed. Additionally, Stryker can limit the deployment of unnecessary inventory saving millions of dollars.

**BEST PRACTICES**

Other departments in Stryker should have their inventory data available and well organized in their respective ERP system. Additionally, inventory management across all the field locations should be standardized before tackling excess inventory.

**OTHER APPLICATIONS**

The process that was followed (Root Cause Analysis, Acquire relevant data from ERP system, define/calculate excess field inventory, set optimal inventory levels and map excess, set prioritization of excess SKUs, develop data model, repurpose excess inventory through main distribution center, limit replenishments of SKUs that have excess and eliminate redundant backorders) can help other departments implement a similar solution.

Product design / development / engineering | Industrial automation | Improved process performance | K-nearest neighbors

# Forming the Future: A Digital Approach to Simulating Thermoplastic Manufacturing

## BUSINESS PROBLEM

Oribi faces long lead times for new product development due to the complexity of manufacturing thermoplastics. These delays make it challenging to compete with other non-thermoplastic composite manufacturers. The company lacks an efficient way to iterate on designs before physical production, leading to costly and time-intensive trial-and-error methods. By implementing a streamlined digital simulation process and creating a digital twin, Oribi aims to reduce development time, lower costs, and enhance its ability to compete in the broader manufacturing market while attracting new customers.

## APPROACH

To address these challenges, Oribi is developing a digital twin through a structured assessment of current lead times, product development workflows, and digital capabilities. The process involves integrating simulation software, leveraging cloud computing, and enhancing data accessibility. Validation is performed by comparing digital twin outputs to 3D scans and CAD files.

## DATA SOURCES

The thesis utilized data from three primary sources: CAD files provided by customers, SimSof simulations of thermoplastic composite forming processes, and ScanSof 3D scans of production parts. CAD files defined customer specifications, SimSof simulated forming processes, and ScanSof scans validated production accuracy,.

### Data Types and Format

The data types include 3D point clouds, STL files for geometric modeling, and simulation outputs such as deformation patterns, mesh sizes, and material properties.


Author: Rachael Harkavy

## IMPACT

The implementation of a digital twin will significantly reduce Oribi's current process of trial and error, leading to faster product development cycles and lower costs. Previously, multiple tooling iterations were required, each involving machinist work and production delays. With the digital twin, Oribi aims to cut these iterations down to a single production step, reducing both material waste and production costs. The ability to simulate and refine designs digitally will also enhance flexibility in responding to customer needs, making Oribi more competitive against traditional manufacturers. Beyond operational efficiencies, the project strengthens Oribi's position in the market by enabling quicker innovation cycles. The reduction in lead times will allow Oribi to take on more projects and expand its customer base beyond its current limited pool. Additionally, this initiative supports a broader industry shift toward digital manufacturing, positioning Oribi as a forward-thinking leader in advanced composites.

| | |
|---|---|
| DRIVERS | The catalyst was the inefficiency in onboarding new customers due to lengthy physical trial cycles in thermoplastic composite manufacturing. Oribi lacked standardized digital tools for validation, resulting in reliance on manual processes that delayed product development timelines. |
| BARRIERS | Barriers included limited access to accurate material cards for simulations, inability to model dynamic process parameters like pressure profiles, and inconsistencies in real-world data collection such as laminate temperatures and environmental conditions. |
| ENABLERS | Oribi's investment in advanced tools like SimSof and ScanSof enabled digital validation. Collaboration with subject matter experts ensured accurate experimental setups, while the company's focus on innovation supported the development of geometric similarity metrics. |
| ACTIONS | The project involved developing a geometric similarity measurement tool, conducting a Design of Experiment (DoE) with clustered simulations, validating simulation outputs against physical scans, refining material models, and standardizing data inputs for analysis. |
| INNOVATION | Innovations included the use of K-means clustering to reduce simulation complexity, development of a geometric similarity metric for validating digital twins, and leveraging point cloud data to compare simulated and real-world part geometries programmatically |
| IMPROVEMENT | The solution reduced time-to-first-article by enabling simulation-driven validation instead of costly physical trials. It also improved part quality by identifying optimal process parameters through digital modeling. |
| BEST PRACTICES | Best practices include ensuring accurate material characterization for simulations, using clustering techniques to minimize computational costs, integrating real-time sensor feedback into modeling processes, and automating scan-to-simulation alignment to streamline analysis |
| OTHER APPLICATIONS | While my application is meant as a piece in the pipeline of a digital twin, it can also be used as a validation tool to see if parts are within customer specifications |

Engineering | Technology | Product design and testing | Supervised Learning

# Domain Adaptation of VLM for Soccer Video Understanding

Amazon

## BUSINESS PROBLEM

Recent advances in vision language models (VLMs) have significantly improved their ability to process and understand visual inputs, including images and videos. However, while general-domain video VLMs have been extensively studied, there remains a gap in understanding how well these models can be adapted to domain-specific applications. This paper aims to address this gap by exploring an effective methodology and repeatable recipe for adapting a general-purpose video VLM to a specialized domain. As a case study, we investigate the adaptation of LLaVA-NeXT-Video, a state-of-the-art open-source video VLM, to the soccer domain in sports.

## DATA SOURCES

We source videos from two large-scale soccer datasets, SoccerNet-V2 and WyScout. WyScout is a subscription-based proprietary soccer dataset.

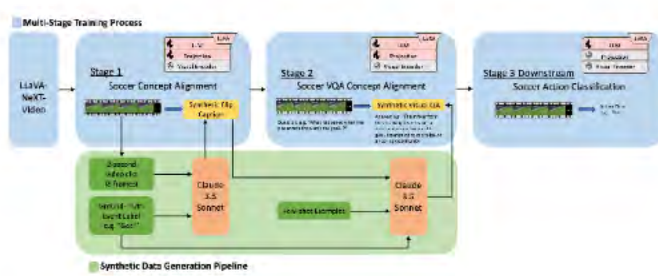
### Data Types and Format

Both datasets contain soccer match video footage and timestamped key soccer event labels.


Author: Tony Jiang


## APPROACH

We employ a three-stage finetuning process to adapt our base model to the soccer domain. 1. Concept Alignment: The model is first finetuned to develop an understanding of core soccer concepts. 2. Instruction Tuning: The model is then trained to follow diverse soccer-related instructions. 3. Downstream Task Finetuning: Finally, the model is further refined to perform specific soccer-related tasks



## IMPACT

This research, while not working with a specific AWS customer, aims to address this need by exploring best practices for adapting video VLMs to specialized domains. While the soccer domain serves as the initial case study for this research, the long-term business objective is to generalize our leanings so we can apply the same approach to other verticals or industries. The research also aims to develop a handbook or recipe of vision-language model finetuning broadly so it can help streamline domain adaption projects in the future for specific AWS GenAIIC customers. The ability of video VLMs to interpret and generate textual descriptions from videos opens up a wide range of commercial applications. The sports industry, in particular, stands to benefit greatly from video VLM advancements. Automated analysis of soccer games, for example, can provide insights into player performance, tactical formations, and key events like goals or fouls. Beyond sports, similar techniques can be applied in surveillance, autonomous driving, and entertainment, where understanding video content is essential.

| | |
|---|---|
| DRIVERS | The need for custom domain adapted video understanding tasks |
| BARRIERS | Technical challenges such as 1. VLM hallucination 2. Highly dynamic sports domain such as soccer 3. Lack of fully descriptive data labels |
| ENABLERS | 1. Team support and mentorship 2. Data sources 3. Computation resources |
| ACTIONS | 1. Data cleaning and curation 2. Full implementation of the model trained for the research. 3. Design and execution of the experiments and ablation studies |
| INNOVATION | 1. Multi-stage finetuning procedure 2. Synthetic data generation pipeline that uses LLM assistant 3. Data proprocessing |
| IMPROVEMENT | The final adapted model, trained using a curated dataset of 20k video clips, exhibits significant improvement in soccer-specific tasks compared to the base model, with a 37.5% relative improvement for the visual question-answering task and an accuracy improvement from 11.8% to 63.5% for the downstream soccer action classification task. |
| BEST PRACTICES | Please see section 6.3 in the thesis for best practices. |
| OTHER APPLICATIONS | 1. Sports match commentary 2. Refereeing assistance 3. Activity and event detection ... Adaptation to other custom domain and corresponding tasks. |

Logistics / Distribution | Automobile manufacturing | Transportation optimization | Single Objective

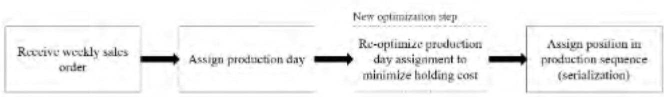
# Optimizing automotive production scheduling to reduce finished vehicle inventory

## BUSINESS PROBLEM

Automotive manufacturers face challenges managing finished vehicle inventory due to the misalignment between production scheduling and outbound logistics. Traditional production planning methods optimize for assembly efficiency, prioritizing smooth manufacturing workflows and component availability. However, these methods often fail to account for outbound transportation constraints, leading to inefficiencies in vehicle distribution. As a result, vehicles accumulate in storage yards while awaiting shipment, increasing inventory holding costs, reducing cash flows, and delaying dealer deliveries.

## APPROACH

This thesis proposes an optimization-based approach to align production scheduling with outbound logistics. The methodology reorders the daily production sequence to facilitate faster load formation while ensuring compliance with manufacturing constraints. By grouping vehicles into truckloads earlier in the production process, the model minimizes load assembly time, accelerating vehicle shipment.




## DATA SOURCES

Data came from three databases: the production, sales, and outbound logistics. The production schedule included vehicle ID, configuration, and assigned day. The sales database provided assigned dealerships, ZIP codes, and holding costs. The outbound system specified delivery mode (only truck was considered). These were merged to identify units that could be grouped into truckloads.

### Data Types and Format

Sales and production data was stored in SQL databases. Outbound logistics data was stored in spreadsheets.

Author: Chris Johnson

## IMPACT

The optimization model significantly reduced finished vehicle inventory. Back-testing showed a 63-65% average reduction in inventory load assembly time across two assembly lines. A real-world pilot confirmed the model's effectiveness, cutting load assembly time by 33% without disrupting manufacturing. This leads to faster cash flow, lower storage costs, and better delivery performance. OEMs also gain visibility into outbound inefficiencies, strengthening supply chain control. Despite scaling challenges when working with more than 600 vehicles, even partial solutions deliver substantial benefits. These results demonstrate that production and logistics integration offers meaningful operational and financial value.

**DRIVERS**

The automotive industry separates production and outbound logistics planning, which is inefficient. While a large body of literature discusses how production and outbound logistics planning can be combined in practice, discussion of real implementations is non-existent. Not only would solving this problem be academically interesting, but it would help improve cashflows, which was an important objective for the business.

**BARRIERS**

One major challenge was data fragmentation—logistics, production, and sales systems were separate and had not been integrated before. Computational scalability was another issue: large problem sizes solved more slowly, or not at all. Additionally, outbound carriers were independent and their pickup times couldn't be changed, limiting control to just load assembly.

**ENABLERS**

Leadership supported innovation and trying new ideas. There was also great support from the production planning team, which sat in a different leadership vertical than me. They were very willing to discuss the planning process and supported me in running a pilot where I adjusted the live production schedule.

**ACTIONS**

I developed the optimization in Python using real data, validated it through back-testing, and shared results with the planning team. We selected a pilot week and implemented an optimized schedule. After tracking outcomes, we confirmed a 33% inventory reduction during the pilot. I created documentation and training materials to support future use.

**INNOVATION**

The proposed approach is the first in the literature to discuss real-world testing of an integrated production and outbound distribution model in the automotive sector. The approach used a heuristic to maintain manufacturing feasibility while improving solvability. It also operates within existing systems, requiring no major workflow overhaul.

**IMPROVEMENT**

The solution reduced inventory of affected vehicles by an average of 63-65% in simulation and 33% in real-world tests. This translates to lower inventory costs, improved cash flow, faster delivery times.

**BEST PRACTICES**

Limit scope as much as possible to prevent problem from becoming too large to solve. Engage production planning team early to understand relevant constraints of inbound supply chain and manufacturing and work together to identify heuristics that will simplify optimization. Production planning is a challenging and demanding task; it is critical that any new addition to the planning process will create little to no disruption to existing workflows.

**OTHER APPLICATIONS**

This approach can apply to any manufacturing environment where outbound logistics depends on batch formation.

# Optimizing EV Fleet Charging Infrastructure Deployment for Maximum Emissions and Savings Impact

## BUSINESS PROBLEM

Verizon is planning on electrifying its fleet of 10,000+ light-duty vehicles spread across 1,000+ sites as part of a broader effort to meet its corporate sustainability goals. However, deploying EV charging infrastructure at such a dispersed fleet presents challenges: site-specific constraints, diverse vehicle usage patterns, location-based grid emissions, labor costs, utility rates and incentives. A scalable, data-driven approach is therefore required to support decision making about what order to electrify sites in and how many vehicles to convert on each site to minimize capital and operational costs and maximize emissions reduction.

## DATA SOURCES

Verizon Telematics Data (fleet usage, mileage, fuel purchase data), Verizon Real Estate Data (coordinates, type & ownership status), utility rate schedules, Bureau of Labor Statistics (zip based labor rates), government sources (State & Federal incentives, and state RPS targets), National Centers for Environmental Information (weather data), previous EV projects by Verizon (capital expense data).

**Data Types and Format**

Mainly time-series (telematics, weather) and location-based (labor rates, RPS targets) tabular data. Incentives data were unstructured text - converted to structured tables for ease of analysis.

Author: Mohit Kasliwal

## APPROACH

An integrated optimization model was developed to identify the optimal sequence and scale of charger deployment across sites. The model incorporates site-level vehicle usage, capital and operational expenses, incentive structures, and emissions profiles. Scenarios optimize for emissions reduction, OpEx savings, or a hybrid of both using an internal carbon cost (weighted objective function).

## IMPACT

The optimization model delivers a data-driven roadmap for EV charger deployment aligned with Verizon's sustainability goals. By incorporating a hybrid objective that balances emissions reductions with operational savings, it outperforms a baseline ranking by fleet size, achieving 72% higher $CO_2$ reductions and 87% greater cost savings under an ICC of $200 per metric ton $CO_2$ for a single-year planning horizon. The effective ICC, calculated post optimization, is $113, much lower than the input of $200/m.ton CO2 suggesting that the model chooses sites which have a better balance of operational savings and emissions reductions. Although the multi-year analysis results are only 2-3% better than baseline, it front loads high-impact sites, delivering early benefit as compared to baseline. The model incorporates conservative factors of safety for battery degradation (both long-term and weather based), ensuring only vehicles whose daily rides are within the effective range are electrified, avoiding the need for on-the-go charging. The optimization can re-evaluate site conditions if incentive programs or corporate goals change (change ICC), or if new EV technologies become available. Overall, the model ensures that each dollar spent on electrification is directed where it yields the highest net benefit, creating a replicable, robust blueprint for other large-scale fleet operators aiming to accelerate emissions reduction without compromising on financial viability.

**DRIVERS**

Stringent corporate sustainability targets at Verizon, including a 53% reduction in Scope 1 and 2 emissions by 2030 and net-zero operational emissions by 2035, demanded a scalable fleet-electrification plan. Evolving regulations like California's ACF rule, and consumer expectations around green operations spurred the need for a data-driven EV-charging infrastructure strategy.

**BARRIERS**

High initial capital expenditure, along with potential to disrupt operations while upgrading sites is the biggest barrier to rollout. EVs are also prone to disruptions due to power outages. This, along with other factors such as range anxiety, uncertainty of weather impacts on range make it difficult to get some site operators on-board.

**ENABLERS**

Strong executive sponsorship, robust internal data analytics capabilities, extensive fleet data, and cross-functional collaboration (Real-Estate, Fleet Operations) helped accelerate solution development.

**ACTIONS**

Collected and analyzed telematics data to identify usage trends. Engaged stakeholders to define scope. Gathered labor, incentive, grid-emissions, and utility-rate data. Built single- and multi-year models of capital/operational costs, emissions, and incentives. Incorporated load management to reduce demand charges. Applied a hybrid objective (emissions + OpEx) to prioritize high-impact sites.

**INNOVATION**

A hybrid optimization model simultaneously targeted emissions reduction and cost savings, using an internal cost of carbon. The model accounted for demand charges, weather-driven range loss, and location-based incentives. It also accounted to operational limitations while developing a easy to use ranking system for site deployment.

**IMPROVEMENT**

Compared to a baseline approach which electrifies sites in descending order of size, the single-year hybrid solution yields up to 72% additional emissions reduction and 87% more operational savings based on an ICC of $200/m.ton CO2. Over a long-term planning horizon, the model yields 2-3% improvement, but does so by front-loading high-impact sites and lowering capital expenditure by making effective use of incentives.

**BEST PRACTICES**

1. Tailor EV deployment to real site and vehicle usage profiles - avoid one-size-fits-all. 2. Incorporate weather impact and battery degradation assumptions. 3. Don't ignore smaller sites as they might be high impact due to high annual mileage. 4. Use load management to reduce demand charges. 5. Leverage an internal cost of carbon to balance profitability and sustainability goals.

**OTHER APPLICATIONS**

This integrated optimization framework can extend to commercial fleets in logistics, utilities, and public transit. Future expansions could include medium- and heavy-duty vehicles, backup-power integration, or vehicle-to-grid options.

Product design / development / engineering | anonymous | Improved process performance | Discrete / Continuous

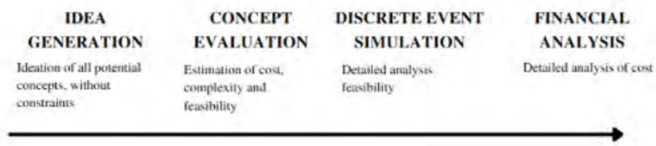
# EV Fleet Charging: A Simulation-Based Comparison of Charging Strategies and Cost Implications

## BUSINESS PROBLEM

Fleet charging strategies must balance cost efficiency, reliability, and operational constraints. A common approach is to deploy a dedicated 10kW charger for each EV. This ensures charger availability and simplifies operations, however, it leads to high infrastructure costs and results in underutilized chargers since not every vehicle will need a full charge each night. Conversely, shared charging systems improve charger utilization and may reduce upfront capital expenditures but also introduce logistical challenges and complexities.

## APPROACH

The project examines charging methods for EV fleets by ideating potential charging methods, down-selecting based on determined Company requirements, then evaluating the operational feasibility through a discrete event simulation and conducting a financial analysis.




## DATA SOURCES

Internal historical site data including EV arrival times, state of charge, fleet size, qualitative interviews and site visits to observe existing operations

**Data Types and Format**

Quantitative data from Company in Microsoft Excel, Qualitative data from site visits, interviews

Author: Rachael Knapp

## IMPACT

By using a discrete event simulation and financial analysis, fleet managers can identify charging methods that unlock significant cost savings while still fully charging the fleet. This discrete event simulation delivered insights into the effectiveness and cost of different charging methods. The baseline dedicated 10kW charger is simple in terms of operational complexity and results in all vehicles being charged before the target time, but it incurs a high 7-year cost per vehicle charged compared to other methods. Simultaneous multi-port charging offers lower cost per vehicle charged while completing charging before the target time. Additional analyses were conducted on opportunities to improve charging methods, such as through shifting associate hours, adding associates, or introducing a directed path algorithm to instruct which chargers to use for each EV.

| | |
|---|---|
| **DRIVERS** | As fleet electrification expands, fleet managers want to understand the most reliable yet cost-effective way to charge the vehicles. The study fills gaps in existing literature, which often overlook the interactions between operational feasibility and financial considerations. |
| **BARRIERS** | Integrating multiple groups in the company for holistic solutions, including charging product engineers, operators, site development, etc. |
| **ENABLERS** | Engrained in the culture, the company is willing to investigate unique approaches to EV charging. There is a substantial amount of historical data to reference for testing the model. There are numerous experts in the area of EV charging within the company. |
| **ACTIONS** | Delivered a discrete event simulation and financial model to compare charging methods. The outcome also included near-term recommendations to continue advancing the project. |
| **INNOVATION** | This research provides insights into the balance between charging infrastructure investment and operational efficiency. The comparison of using fewer DC-fast chargers integrated with manual swapping versus deploying more L2 chargers without manual intervention has not been thoroughly documented for fleet operations. This research bridged these gaps through a discrete event simulation and financial evaluation of charging strategies. |
| **IMPROVEMENT** | Initial analyses prove that multi-port simultaneous charging significantly reduces the cost per vehicle charged compared to dedicated 10kW chargers, while maintaining operational goals. |
| **BEST PRACTICES** | Visit sites and engage with stakeholders early to fully understand the complexities of the problem. Leave time to iterate after initial results. Many of the learnings in this project came from testing different variables beyond the initial concepts. |
| **OTHER APPLICATIONS** | Any EV fleet such as busses, robo-taxis, municipal vehicles and delivery vans. |

Maintenance | Industrial automation | Improved process performance | Natural language processing

# Multimodal Generative AI Chatbot for Root Cause Diagnosis in Predictive Maintenance

Amazon

## BUSINESS PROBLEM

Predictive maintenance systems are valuable for detecting early signs of equipment failure using sensor data such as vibration and temperature. When anomalies occur, alerts notify maintenance teams to take action before breakdowns happen. However, while these systems excel at signaling that a problem exists, they provide little support in diagnosing the underlying issue or guiding technicians through resolution. As a result, teams face delays, extended downtime, and inconsistent fixes. Technicians often rely on manual searches through complex documentation and individual expertise to identify the root cause.

## DATA SOURCES

The solution uses two primary data sources: time-series sensor data capturing vibration and temperature at one-second intervals, and PDF-based repair manuals containing images, tables, and procedural text. The sensor data enables anomaly detection, while the manuals provide diagnostic context for root cause analysis.

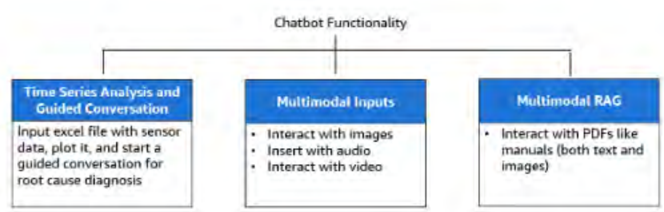
### Data Types and Format

Time-series sensor data (CSV), equipment images (JPEG/PNG), video clips (MP4), audio recordings (WAV), and PDF repair manuals with structured and unstructured text.

Author: Carla Lorente Anon

## APPROACH

I developed a generative AI chatbot using AWS tools to help technicians diagnose equipment issues. The assistant analyzes sensor trends, retrieves relevant repair steps, and supports image, video, and audio inputs. It uses multimodal retrieval-augmented generation (RAG) and LLMs to generate guided, real-time troubleshooting conversations tailored to the issue.



## IMPACT

The AI assistant helps technicians resolve issues faster by reducing time spent searching through manuals and guessing next steps. In pilot testing, diagnostic time dropped by over 30%, saving thousands per incident in avoided downtime. It also improves the use of existing knowledge by retrieving relevant repair history and technical content. Instead of relying on static, generic work orders, technicians receive case-specific, targeted guidance. By narrowing down possible causes, the tool increases first-time fix rates, reducing repeated maintenance and improving equipment uptime. Once issues are resolved, the assistant supports better reporting by guiding technicians to log accurate closure codes. This leads to cleaner maintenance records and improves model feedback for future predictive systems. The assistant also acts as a virtual mentor for new technicians, providing consistent, guided troubleshooting that accelerates onboarding and reduces dependence on tribal knowledge. Although built for manufacturing, the architecture is flexible and can be adapted to other industries like logistics, utilities, or oil and gas—anywhere sensor data and maintenance documentation are used together.

**DRIVERS**

Industrial operations face rising pressure to reduce downtime and improve maintenance efficiency. While sensors can detect issues early, diagnosing them remains slow and manual. The lack of tools that connect sensor data with actionable, technician-friendly guidance created a clear need for a scalable generative AI solution that supports faster, more accurate troubleshooting.

**BARRIERS**

A major barrier was the difficulty in measuring long-term impact before full implementation. Metrics like reduced downtime, improved fix rates, and model retraining benefits require extended usage to quantify. Early results were promising, but demonstrating sustained value depends on continued adoption and integration into daily technician workflows.

**ENABLERS**

Strong collaboration with maintenance and data engineering teams enabled access to high-quality sensor data and real-world feedback. The company's openness to innovation and support from AWS resources accelerated prototyping and deployment. Cross-functional input from technicians ensured the solution addressed practical needs on the ground.

**ACTIONS**

I visited the company's fulfillment center to observe operations firsthand and gather insights directly from technicians. This helped identify key pain points in troubleshooting. Based on these findings, I built a generative AI assistant that analyzes sensor data, interprets technician inputs, and retrieves relevant repair guidance to support real-time diagnostics.

**INNOVATION**

Traditional work orders were static and generic, offering limited value during troubleshooting. This solution introduced dynamic, AI-generated guidance tailored to real-time sensor data. While retrieval augmented generation (RAG) is widely used across other company projects, this implementation was the first to integrate multimodal RAG, retrieving both text and images. The solution's novelty led to a patent filing.

**IMPROVEMENT**

The solution reduced diagnostic time by over 30% in pilot tests, leading to faster issue resolution and decreased equipment downtime. It also increased the accuracy of closure codes, improving maintenance logs and enabling better predictive model retraining. These improvements directly contribute to higher technician productivity and operational efficiency.

**BEST PRACTICES**

Start by identifying critical assets whose failure would disrupt operations. Install sensors to log vibration and temperature data and send it to the cloud. Gather OEM repair manuals for those assets and, if possible, collect historical maintenance records. This foundation ensures the AI assistant can deliver accurate, relevant, and context-aware diagnostic guidance.

**OTHER APPLICATIONS**

The solution is adaptable across industries where minimizing equipment downtime is essential, such as manufacturing, logistics, oil and gas, and healthcare. It can support predictive maintenance for motors, fans, gearboxes, bearings, conveyors, and other components, as long as sensors capture vibration and temperature data.

# Design Transfer as a Lever for Accelerated Medical Device Innovation: A Case-Based Mapping Approach

## BUSINESS PROBLEM

Stryker aims to accelerate its New Product Development (NPD) to deliver new products to market faster and more efficiently. A subset of NPD, Design Transfer, describes the procedures governing the correct translation of a new design into a set of manufacturing instructions and procedures that could be transferred to manufacturing operations. Design Transfer, as it currently stands, is a one-size-fits-all procedure that is over-prescriptive. There have been attempts to parallelize some of the tasks and bucket deliverables for phase exits, but there has not been a study on the effect of this parallelization on the amount of rework.

## APPROACH

I analyzed the as-is process for four projects across two business units that vary in complexity, spend, and procedure. I visualized the process steps using a swimlane flowchart and Design Structure Matrix to highlight feedback loops and rework cycles. Subsequently, I tracked the effect of task parallelization on the amount of rework occurring in the development phase across cases.

## DATA SOURCES

Primary sources: data gathered via interviews with subject matter experts in Design Transfer from Advanced Operations Engineers, Advanced Operations Project Managers and Advanced Operations Functional Managers. Secondary sources: data gathered from historical project records, including project schedules, task lists, and other relevant documents.

**Data Types and Format**

Microsoft Project Plans, written interview notes, Excel sheets, Powerpoints


Author: Amna Magzoub

## IMPACT

The methodology used sheds light on the Design Transfer process that drives new product development. Mapping Design Transfer allows Stryker's decision-makers to make informed choices on where to optimize the process and minimize chances for rework, aligning with the company's overall goal to accelerate new product development. The case study analysis reveals key inefficiencies caused by poor cross-functional communication, late-stage validation testing, and inadequate supplier integration. Addressing these pain points through improved knowledge-sharing practices, early subassembly testing, and set-based concurrent engineering with suppliers helps reduce delays and costly rework. The approach enhances transparency across teams and builds a shared understanding of the Design Transfer process, enabling more predictable execution. Importantly, the methodology is scalable across business units, making it a repeatable blueprint for driving efficiency in complex, regulated development environments.

**DRIVERS**

The industry has room to adopt the best new product development practices, as it lags behind other industries due to regulation. Stryker, driven by growth, is very interested in adopting strategies to accelerate their operations.

**BARRIERS**

As Design Transfer is a business process, there was no Gemba to visit, and it was difficult to understand the steps to complete the process, especially since there was sparse documentation of the process as executed and issues that arose throughout. Furthermore, the execution of Design Transfer was slightly different across projects, and it required some extra work and best estimates to make the case studies comparable.

**ENABLERS**

The company is highly networked, which made it easy to reach out to any stakeholder I needed. My company supervisor was great at introducing me to people that could support my project.

**ACTIONS**

I engaged with stakeholders throughout and until the end of my project to ensure their buy-in. Stakeholders included Design Transfer Divisional Process Owners (DPOs) who are experts in the process and gave valuable feedback and enthusiasm towards the recommendations shared at the conclusion of my project.

**INNOVATION**

Although swimlane mapping is quite common, DSM mapping is not—especially the sequencing feature to minimize feedback loops, which revealed, for example, the need for later concept freezes.

**IMPROVEMENT**

The solution provides a framework that, if adopted, will lead to a shorter, streamlined design transfer process that reduces the overall new product development cycle and hastens time-to-market.

**BEST PRACTICES**

Engage with stakeholders early and often, especially during the process mapping phase. Be ready to create multiple iterations of the map, as it changes quickly with feedback, and try to seek consensus across stakeholders of each project. Explicitly ask for descriptions of breakpoints in the process and rework, as that is not always clearly documented in the development phase, and fact-check what is reported against what is documented.

**OTHER APPLICATIONS**

The solutions recommended are applicable to any design transfer project across the organization. The methodology used is beneficial in mapping out any complex process that is non-linear with branches and feedback loops to provide a better understanding of its strengths and pitfalls.

# Generative AI in Private Equity for Accumulative Advantage

LFMcapital

## BUSINESS PROBLEM

Generative AI is quickly gaining hype, but the challenge lies in effectively applying it within private equity, especially across diverse portfolio companies. In PE-owned firms, pressures to meet ambitious targets create tight constraints on time, capital, and resources. This research addresses how GenAI can be practically incorporated under these limitations, strategically aligning AI capabilities with organizational goals while simultaneously empowering employees and driving measurable improvements.

## APPROACH

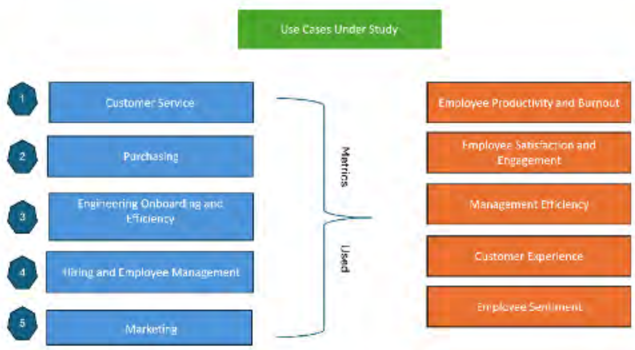
This research employed a case-study approach, systematically exploring targeted mini use cases across key departments—including customer service, purchasing, engineering, employee management, and marketing—within a private equity-owned firm. Each case involved deploying custom-built or publicly available GenAI tools to uncover practical strengths, limitations, and unique implementation challenges.

## DATA SOURCES

The data for this research was primarily qualitative, gathered from targeted mini case studies conducted within a private equity-owned, resource-constrained manufacturing firm. This included direct observations, employee interviews, departmental workflow assessments, and performance metrics collected before and after implementing custom-built or publicly available Generative AI tools.

**Data Types and Format**

Oral, Text and Quantitative




Author: Bonny Mahajan

## IMPACT

From the private equity firm's perspective—especially those managing diverse portfolios—this research demonstrates that strategic integration of Generative AI can significantly accelerate the scaling of portfolio companies. When companies are experiencing rapid growth, internal resources often become strained, creating resistance or difficulty in adopting new technologies. The practical insights provided by our case studies illustrate how GenAI solutions can be effectively implemented across various departments, clearly identifying areas where the technology enhances workflow efficiency and where caution is warranted. Crucially, by emphasizing the critical role of employees in the adoption process, our findings show that successful GenAI integration requires careful consideration of organizational culture, user interface design, and domain expertise. Consequently, the PE firm benefits from clearer decision-making criteria for investing in GenAI initiatives, ensuring alignment with both organizational goals and employee empowerment. Ultimately, the solution outlined here helps portfolio companies better leverage Generative AI to achieve sustainable productivity gains, foster a culture of continuous improvement, and support employee satisfaction, positioning them strategically for sustained growth.

**DRIVERS**

Historically, traditional industries like manufacturing have been underserved by technology providers, creating significant potential for innovative solutions. Rather than viewing Generative AI as an ultimate solution, this research was catalyzed by the recognition that GenAI can function effectively as a practical tool—especially as these technologies become increasingly accessible—helping employees achieve greater efficiency and satisfaction.

**BARRIERS**

Barriers impacting this project included employee unfamiliarity with Generative AI, creating concerns about potential disruptions to existing workflows. Additionally, it was crucial to carefully identify genuine opportunities for GenAI integration rather than assuming it could address all challenges. Thoroughly analyzing workflows to pinpoint improvement areas was essential, as was employing iterative experimentation and thoughtful UI/UX design.

**ENABLERS**

The project was enabled by a highly supportive senior leadership team that encouraged piloting diverse use cases to identify optimal Generative AI applications. Additionally, an open-minded employee base, enthusiastic about experimentation, facilitated adoption. Given the company's rapid growth phase, the team's readiness for workflow improvements helped mitigate employee some fatigue risks.

**ACTIONS**

Implementation began with a detailed workflow analysis across key departments to identify potential areas for improvement through Generative AI. This analysis informed decisions on whether custom-built or off-the-shelf tools would best meet departmental needs. Subsequently, multiple iterative feedback rounds with employee teams were conducted to evaluate, refine, and address limitations of the proposed GenAI solutions within the company context.

**INNOVATION**

Innovative aspects of this solution include the deliberate rejection of a "one-size-fits-all" approach, instead customizing Generative AI tools uniquely to departmental needs, such as distinct solutions for customer support versus HR or purchasing. Additionally, tools were thoughtfully integrated into existing workflows to minimize disruption, emphasizing effective change management among employees unfamiliar with GenAI technologies.

**IMPROVEMENT**

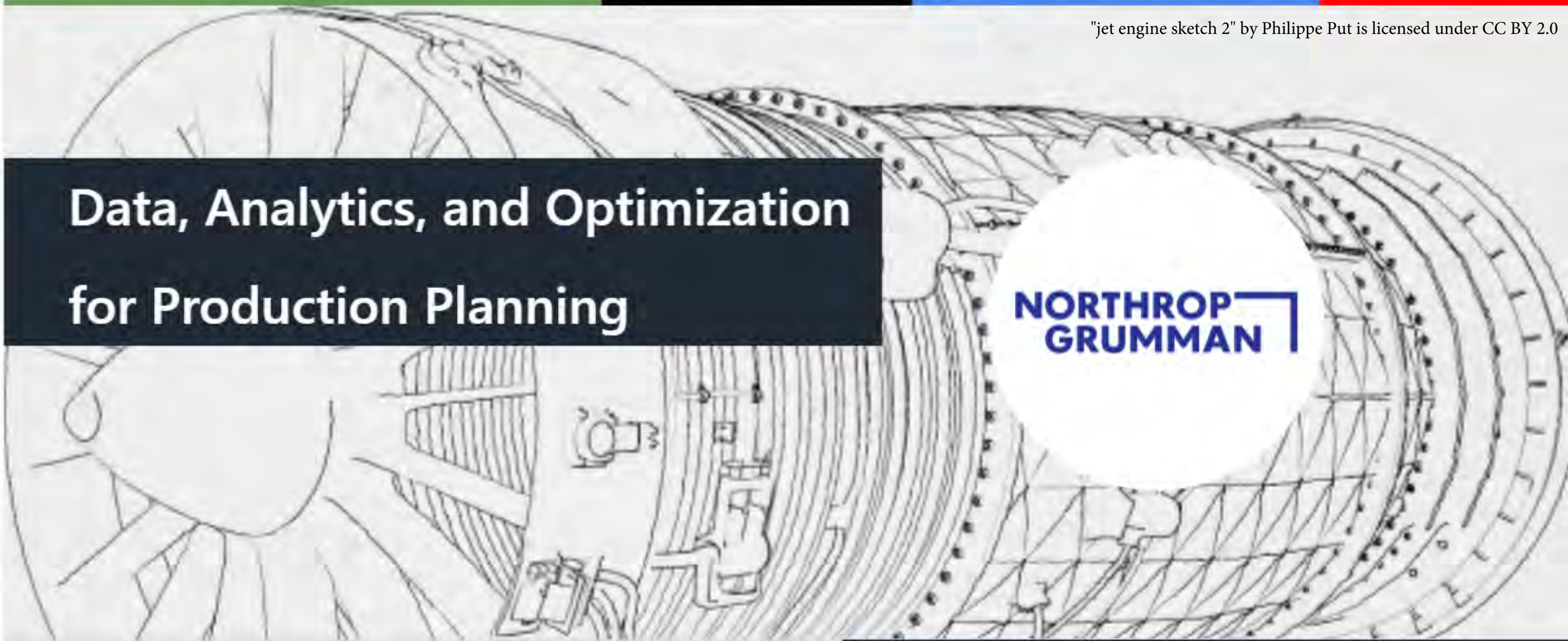
The implemented Generative AI solution produced measurable efficiency improvements across departments: customer support saw daily task reductions of 10-20 minutes per employee, purchasing achieved a 70-80% decrease in email response times, software engineering productivity improved by 5-10%, and marketing saved 1-2 hours per task. These quantitative gains were further enhanced by notable qualitative benefits in employee experience & satisfaction.

**BEST PRACTICES**

Best practices for replicating this solution involve carefully aligning Generative AI initiatives with the organization's core objectives. Specifically, it is essential to first identify and prioritize the top three to five business goals. Adopting this targeted approach ensures that GenAI integration is strategic, impactful, and sustainably aligned with the company's long-term priorities rather than merely following industry trends.

**OTHER APPLICATIONS**

Other potential applications of this solution include leveraging successful case studies to facilitate adoption in additional private equity-owned portfolio companies, particularly to secure senior leadership buy-in. Additionally, the findings offer valuable insights for firms developing and commercializing AI tools, Supply chain planning and inventory management | Aerospace and Defense | Inventory management / control | Linear / Non-Linear




# Data, Analytics, and Optimization for Production Planning

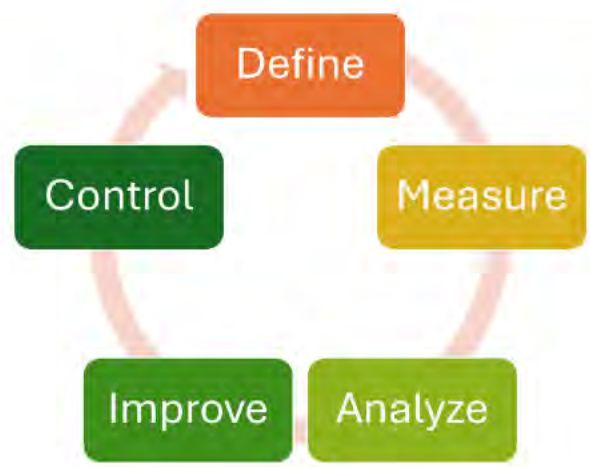
NORTHROP GRUMMAN

## BUSINESS PROBLEM

Northrop Grumman's Mission Systems business sector wanted to improve kitting throughput for its Surface Mount Technology (SMT) factory. The key challenge was that high volume SMT reels could not be shared across multiple orders that needed the same component. While total inventory of a component could satisfy forecasted demand, it might not have enough reels to adequately distribute that inventory. A framework for identifying these components and their impact to the business did not exist.

## APPROACH

The approach to this project was to compare inventory data to forecasted demand data for tape and reel components. Through data manipulation, the number of reels in inventory could be compared to the number of reels needed to support production. Components with insufficient reel quantities were identified and characterized based on cost and accounting structure.




## DATA SOURCES

Data sources were predominantly outputs from the MRP system which described current inventory state, upcoming demand, and BOMs for that demand.

**Data Types and Format**

Data was in .csv format and composed of strings and integers.



Author: Max Malinowski

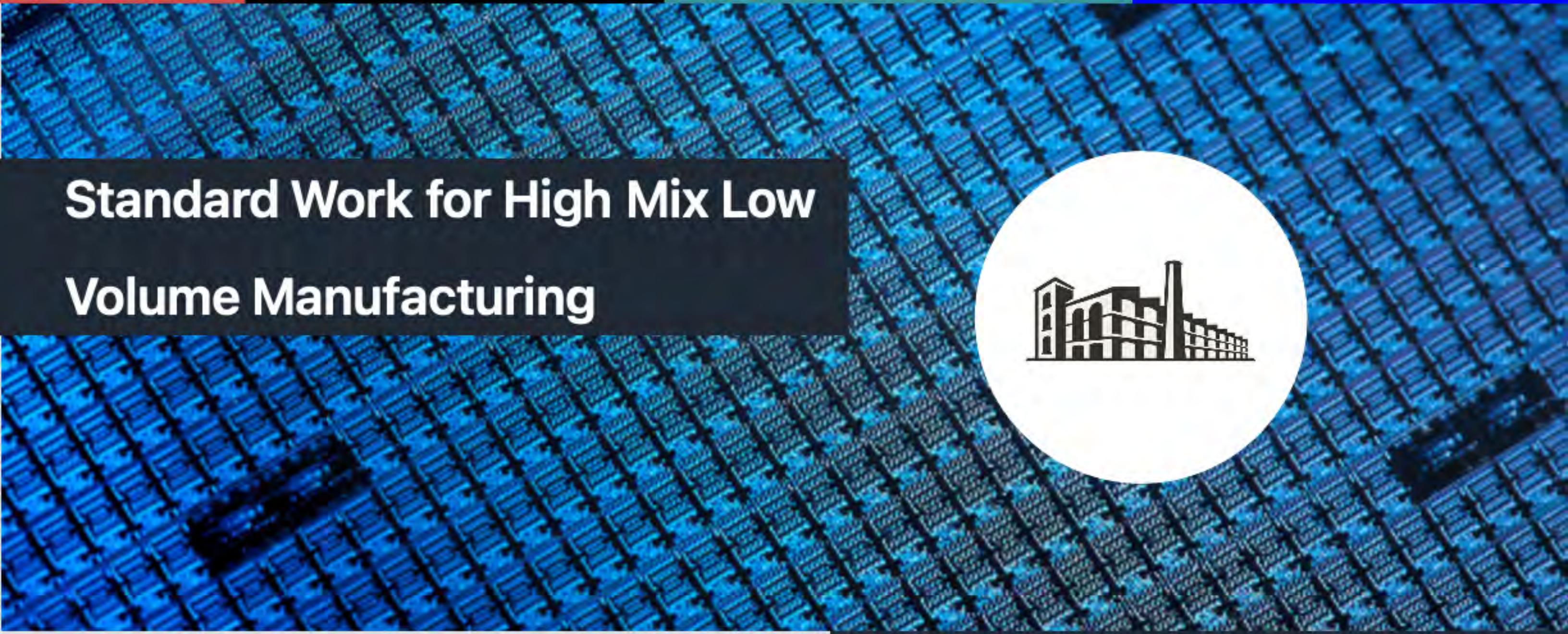
## IMPACT

This project identified that Mission Systems procured plant stock based on historical data rather than demand forecasts, and confirmed that these components with insufficient reels had substantial impact on kitting throughput for the SMT factory. Additionally, a framework and automated software tool were developed to identify these components based on existing company data sources. Ultimately, Mission Systems better understood a root cause of its kitting throughput challenges for SMT and was given tools to help alleviate this issue. Eliminating the reel shortages identified by this project will necessarily help Mission Systems improve SMT production throughput.

| | |
|---|---|
| DRIVERS | The high mix low volume nature of the SMT factory was a key contributor to the constraints that gave rise to the shared reel problem. |
| BARRIERS | The value stream for the SMT process, beginning with procurement and ending with a Circuit Card Assembly, was incredibly long and spanned multiple organizations. Coming up to speed on a complex process was quite challenging. |
| ENABLERS | The accessibility and availability of clean data relevant to the problem was tremendously helpful for producing data-driven insights. |
| ACTIONS | An optimization model to better allocate inventory was piloted, and an automated software tool to screen for components with insufficient reels was rolled out to the planning team. |
| INNOVATION | Engineering data features that better illustrated key characteristics of the system was a key innovation that enabled the success of this project. |
| IMPROVEMENT | By alleviating the shared reel issue, factory throughput is expected to increase by ~20%. |
| BEST PRACTICES | Document assumptions and vet them early during data analysis. Keep modeling simple to keep stakeholders informed. |
| OTHER APPLICATIONS | Optimization and exploratory data analysis are known to be widely generalizable to just about any problem. |

Production | Composites Manufacturing | Improved process performance | Graph-based (Markov model)

# Standard Work for High Mix Low Volume Manufacturing

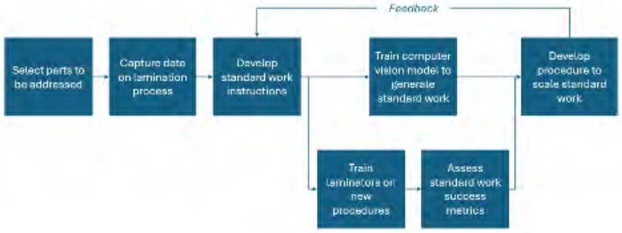


## BUSINESS PROBLEM

The thermoset composite industry currently faces rapid demand growth. Because the lamination process for TSC components in high-mix low-volume operations is highly manual, this demand surge will require a substantial increase in skilled labor forces which risks straining existing capacity and undermining quality; a challenge currently faced by Re:Build Composite Resources. Under these conditions, a major problem becomes how the business can systematically scale and up-skill its labor force to meet escalating demand while preserving product quality and improving overall efficiency.

## APPROACH

We investigated the efficacy of formalizing best practices through standard work at scale. We aimed to develop a prioritization process to identify addressable parts and a robust framework for efficiently deploying standard work methodologies that can facilitate more effective training, reduce variability, minimize scrap, and improve productivity for high-mix low-volume operations.




## DATA SOURCES

We had production data available from CR's Job Boss ERP system including material and labor costs, revenue, production volume, and time logs. We also had access to a spreadsheet the quality department used to track scrapped parts. In addition, we manually observed processes and captured videos that served as a baseline dataset for the cycle times and the current state.

### Data Types and Format

Production data spreadsheets, videos of processes, notes from operator interviews.

Author: Will McNulty

## IMPACT

Standard work at Composite Resources significantly improved quality and production consistency for a selected pilot part. For this part, work instructions for a new "best-known" process helped decrease the scrap rate by over 50% from previous months, while production volume simultaneously increased. For Part B, designing a one-piece flow schedule has prepared CR to increase throughput to meet growing customer demand in 2025. Beyond immediate gains, the thesis presents a framework for scaling standard work across the business' portfolio by prioritizing part groupings according to business impact—focusing on high-cost or high-volume parts. The proposed frameworks for standard work implementation and scaling position CR for demand expansion while maintaining or even improving quality and efficiency.

| | |
|---|---|
| **DRIVERS** | The business and thermoset composites market as a whole are growing. This has created the need to implement methods to help scale the labor force and production volume while still continuing to improve quality. |
| **BARRIERS** | Batched tracking of parts was a barrier to our project. Parts involved in our project were tracked, moved, and inspected in quantities of around 30 pieces. This practice impacted our ability to perform root cause analysis for defects at the operator and part level. Additionally, due to customer constraints, production for the second part addressed was paused until the start of 2025 which prevented implementation of the one-piece flow schedule. |
| **ENABLERS** | The culture at CR was a major enabler for our project; everyone from plant floor operators up to the GM was open to new ways of doing things and wanted to improve processes which allowed for support of my project. The lamination department leadership team was also key to implementing and sustaining proposed process changes. Finally, my managers and the IT team at CR made onboarding seamless from day one. |
| **ACTIONS** | Training and auditing were the most significant actions we took to implement our solution. We held a group training for all operators to introduce the new process and work instructions. We selected an experienced operator to lead the training and ensured leadership was present to communicate buy-in. We then set up process audits led by department leadership to ensure that process changes were sustained. |
| **INNOVATION** | We explored the use of computer vision to automate time-study analyses. This would serve as a tool to expedite the development of standard work instructions. However, our research found the current state and cost of computer vision software solutions were not applicable to CR's operation. |
| **IMPROVEMENT** | Standard work instructions helped reduce the monthly scrap rate from an average of greater than 20% to less than 5% for the first pilot part. The solution also contributed to an over 70% increase in monthly shipped volume to that same customer. We could not implement the one-piece flow schedule for the second pilot part, however simulations show it is likely to improve daily production volume and consistency. |
| **BEST PRACTICES** | Start with prioritization. Developing standard work consumes time from engineers so it is critical to make sure such efforts are directed to the parts that will have the greatest business impact. Further, when developing new processes, start by building a relationship with operators and understanding why they do things a certain way before suggesting any changes. Often, operators are the best source for process improvement opportunities. |
| **OTHER APPLICATIONS** | Our standard work frameworks could be applied in other high-mix low-volume operations by following the prioritization matrix presented in our research. |

Logistics / Distribution | Transportation and logistics | Product design and testing | Network analysis

# Design and Optimization of Shipping Container for Package-Less Units

**Amazon**

## BUSINESS PROBLEM

Amazon aims to ship units without any Amazon added packaging on certain transportation legs. This is in line with their sustainability goals and worldwide cost to serve reduction goals. It would involve process, and equipment changes across their end-to-end supply chain. One particular change is the introduction of a means to carry these package-less units across the supply chain, which is expected to be done through the use of a re-usable container. Therefore, this container needs to be optimally designed to integrate into the expected process stream.

## APPROACH

Hard and soft constraints were established based on Amazon's requirements. A container sizing simulation was then done to ensure the container is sized sufficiently based on the constraints. This was followed up by a process and reverse logistics analysis to understand what optimal container type should be used. Lastly, NPV calculation and recommendations were made based on the analysis.


## DATA SOURCES

Internal data from company tables was used. This data contained product information related to size and weight in combination with the expected unit entitlement for package-less units. For relocation planning, linehaul distance and cost data between facilities was used. In process cost analysis, process engineering data from process engineering team was used.

**Data Types and Format**

Historical data tables formed majority of the data used in the sizing simulation, and relocation optimization. Linehaul cost regression data and process calculation data tables were additional types.

Author: Baraka Minja

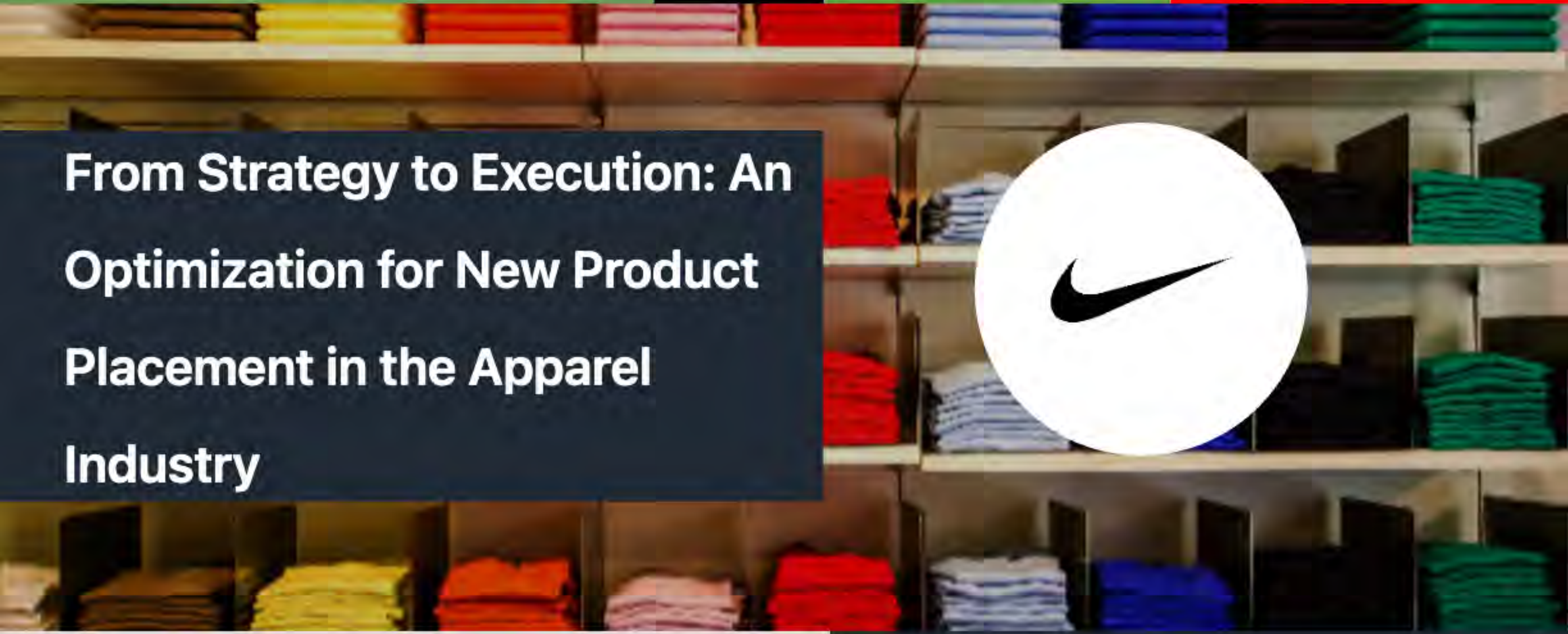
## IMPACT

The solution is expected to inform business decisions as they refine the next steps in developing the end-to-end process pipeline. The analysis shows that container fill rate is sensitive to weight, and this should be a key consideration in developing the upstream process to optimize for this. It also shows that reverse logistics flow will be sub-optimal and thus should be looked at from a different perspective to ensure cost per unit is minimized. Once this end-to-end process stream has been completed, the provided analysis can be revalidated with the new data and stated assumptions to ensure the recommendations align with this finalized process flow. Thereafter, container specifications can be provided to vendors for manufacturing and piloting. Doing this analysis parallel to process development ensures a lot of money and time is not wasted in a launch phase due to the development of a sub-optimal process and container match but rather ensures that a container is designed that fits to the process and a process designed that can accommodate the efficient use of such a container.

| | |
|---|---|
| **DRIVERS** | Amazon has a supply chain on which the solution had to integrate with. Therefore, the use of a container was already established, and the question was what features it should have to create the most upside. These features were designed with respect to Amazon's processes to ensure optimal performance in its end-to-end implementation. |
| **BARRIERS** | As this was a new initiative being developed, it was often open to changes upon periodic review. This sometimes was not ideal for the project timeframe as it often meant some of these changes could not be incorporated in the analysis. |
| **ENABLERS** | Amazon had an extreme pool of resources to get up to speed on their processes quickly. It also had expertise in each facet that helped inform and refine the analysis conducted. |
| **ACTIONS** ➡ | I went about extensive meetings with a variety of stakeholders that would be involved in the implementation or use of the potential end-to-end solution to understand their needs and potential pain points they may have in its implementation, as well as any hard and soft constraints they have. An implementation roadmap of analysis was then drafted to decide what the optimal features are and what features are critical for performance. |
| **INNOVATION** | The relocation analysis conducted looks into designing container features that optimize for relocation transport costs and sustainability impact. Doing this analysis also gives a holistic view of what the overall implementation costs will be for the different use cases and provides a means to simultaneously optimize for design and process. |
| **IMPROVEMENT** | As Amazon continues to develop this process, this solution gives them insight into the critical container features required for package-less shipping. It shows what an ideal size would be to maximize fill rate. Additionally, it shows that current relocation means are sub-optimal and can be re-designed to significantly lower relocation costs per unit. |
| **BEST PRACTICES** | Understanding the constraints is critical to inform the analysis that follows and create an acceptable and applicable solution. Testing and validating any required assumptions should also be considered to ensure a realistic solution is feasible. Initially, design should be conducted towards creating a minimum viable product to test customer experience and thereafter scale form there. |
| **OTHER APPLICATIONS** | Solution does not need to be isolated to package-free shipping but can also be used as a package aggregator to reduce touches per package across the supply chain, and in doing so reduce process costs. |

Supply chain planning and inventory management | Retail | Network optimization | Deterministic / Stochastic

# From Strategy to Execution: An Optimization for New Product Placement in the Apparel Industry



## BUSINESS PROBLEM

This work is motivated by effectively integrating the supply chain group's long-term strategic production objectives into the seasonal planning process for new product introductions. While the business has developed an advanced digital twin to identify long-term cost, emissions, and risk-reducing network changes, the current methods for selecting factories for new products rely on institutional knowledge, potentially creating a disconnect between strategic aims and execution in the medium term. The business seeks a proactive, data-driven capability to ensure that each product-factory placement aligns with future strategy.

## APPROACH

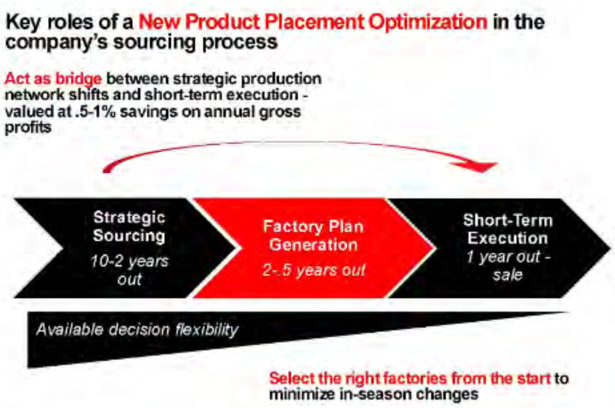
A deterministic optimization model was developed with company sourcing and analytics experts. The model focuses on maximizing network-wide profits while constrained by adherence to the long-term production strategy. The model uses early product demand forecasts and considers real-world business constraints like factory capacity utilization and capabilities.

## DATA SOURCES

Available data included: Historical geography-level demand predictions from Iota's forecast teams; Production and transportation cost coefficients, as well as price coefficients used in the optimization, from financial and operational databases; Factory capacities from supply chain management systems; Long-term production network strategy targets from the latest long-term strategy run.

### Data Types and Format

Included data was exclusively numerical data of demand, costs, and capacities. It was available for at least the past five years of production planning, which was the full timeframe relevant.




Author: Sofie Netteberg

## IMPACT

This new optimization model leads to a more continuous digital twin ecosystem at the company through connecting strategy to execution in network decision-making, as well as by generally expanding data-driven decision-making inside the organization. In the proof of concept, the adherence to long-term strategy improved by 11% from the baseline plan generated by the existing planning process. The model also shows a measurable increase in network profits. By making real-time production decisions closely follow the company's strategic direction, the model helps achieve the company goals of low cost, short lead time, low emissions, and low risk. Further work building on the foundation of this thesis is being carried out at the company that can enhance network stability and potentially reduce costly late-stage adjustments.

| | |
|---|---|
| DRIVERS | The catalyst was the complexity of such a large global supply chain and the need to make optimal decision-making less time-intensive. |
| BARRIERS | Barriers included gaining a sufficient understanding of the complex process and decision-making criteria for selecting a suitable factory for new products. An additional barrier was data availability and the time needed to process data sources. Restricting the model scope (product and factory set) was necessary due to runtime and data cleaning constraints. |
| ENABLERS | Collaboration with Iota's supply chain experts provided invaluable expertise and insights. Their deep knowledge of networks, strategy, and optimization was foundational. Iota's investment in a digital twin ecosystem provided dataflows, direction in selecting an optimization approach, and an integration pathway. |
| ACTIONS | Implementing a solution involved concurrently developing a deterministic MILP optimization model and consistently meeting with supply chain business counterparts to collect real-world constraints and sanity check initial data inputs and results. Model development included collecting and cleaning data, and defining decision variables, coefficients, an objective function, and constraints. The model was executed and outputs were discussed. |
| INNOVATION | The solution's innovation is the development of a method to ensure the consistency of production network changes across time horizons, and focusing that optimization method on the critical choice of where to place new products. Exploratory work on historical demand variability and the parameterization of demand volatility also opened the possibility of probabilistic forecasting and statistically determined optimal order quantities. |
| IMPROVEMENT | The proof of concept optimization with a single product category sold by the company demonstrated an 11% absolute improvement in adherence to the strategy success metric (from 86% to 97%). It also resulted in a measurable increase in network profits, from a standardized baseline of $1M to $1.14M. |
| BEST PRACTICES | Best practices for a new optimization model with novel constraints and success metrics include early collaboration with domain experts to understand the business context and data. Iterative model development and refinement based on initial results and feedback are crucial. Thorough data collection, cleaning, and validation are essential for model accuracy. Starting with a focused scope (e.g., one product category) can streamline development. |
| OTHER APPLICATIONS | The optimization methodology can be applied within any industry to companies with complex production networks run by multiple individuals, and who hope to strategically shift their high-level supply (whether by geography, supplier, product mix, and more). The approach of using optimization for new product placement decisions specifically can be adopted by other CPG companies. |

# Decarbonized Cement Manufacturing via Advanced Production Technologies

## BUSINESS PROBLEM

NextEra Energy, Inc. is one of North America's foremost electric power and energy infrastructure companies. The broad economy-wide transition from highly polluting, fossil fuel-based processes to renewable electricity-based processes creates business opportunities for NextEra Energy. Cement production contributes 7% of all CO2 emissions and cannot be decarbonized with existing commercial technologies. Accelerating the development of technologies that reduce the carbon intensity of cement production while increasing the usage of renewable electricity provides the opportunity for NextEra Energy to access a new market.

## APPROACH

The economic returns for three candidate decarbonization technologies were calculated for five different potential cement production locations. Models incorporated technology and location-specific production costs in addition to taxes, logistics costs, and available economic incentives. For market viability, the returns of a new production technology must exceed those of conventional technology.

## DATA SOURCES

Data for technology-specific costs were sourced from publicly released reports. Electricity prices were calculated from hourly locational marginal prices combined with the relevant capacity, transmission, and REC costs. Industrial natural gas prices were derived from the Energy Information Administration. Logistics costs were sourced from relevant industry publications and US government reports.

### Data Types and Format

Data types were generally .csv files for tabular data and in direct numerical format for the majority of capital and operating costs, sourced from the relevant data owner or report.

Calcination Reaction at Core of Cement Production:


Author: Michael Norwalk

## IMPACT

Cement production is considered a "hard-to-decarbonize" industry due to the high temperatures needed for production, necessitating the use of fossil fuels, as well as the inherent carbon dioxide-releasing chemistry of the production process. Its hard-to-decarbonize nature combined with its ubiquity and importance worldwide have spurred significant R&D efforts resulting in numerous proposed technology pathways for decarbonization. This analysis provides a framework for NextEra Energy to evaluate potential investments in the low carbon cement space aligned with the company's business objectives and to develop appropriate demonstration & implementation strategies. The study first assesses the technology landscape to determine three candidate technologies for further evaluation. The results of the analysis then inform which of the candidate technologies evaluated, if any, should be prioritized for investment and where further R&D effort should be focused in order to increase competitiveness with traditional production methods. The results also illustrate which markets are best positioned for cement production and consumption. Finally, the study proposes several policy recommendations that can accelerate the adoption of low carbon cement in the United States.

**DRIVERS**

Cement production does not currently require large amounts of electricity. As new technologies are developed that can reduce the carbon intensity of cement production, the penetration of electricity into the industry may increase, generating a market opportunity for electricity producers. In addition, the presence of economic incentives in the U.S. and E.U. can increase the attractiveness of otherwise uncompelling projects.

**BARRIERS**

Many of the technologies under assessment are immature and lack significant publicly available information resulting in greater uncertainty on costs. In addition, many of the inputs to the models can be highly volatile, making the selection of a baseline value more difficult. Many of these barriers can be overcome with sensitivity analysis and through further technology development efforts.

**ENABLERS**

The project was enabled by a highly supportive and experienced team at NextEra Energy. The team enabled conversations with relevant technology providers and facilitated cross-functional interactions within the company to develop the next steps and implementation plan based on the work in this study.

**ACTIONS**

Low carbon cement technology developers and experts were engaged to understand the competitive landscape and their view of industry dynamics and strategies. With the results of the study, the team's efforts could be focused on technologies and applications that are best positioned for success. Internal stakeholders were engaged to understand market opportunities and next steps.

**INNOVATION**

The study assesses a set of leading-edge technologies applicable to many industries. Their use in cement production would be innovative as they have not yet been validated or demonstrated at scale in that context. In addition, the notion of producing low carbon cement in the U.S. for export to the E.U. market to take advantage of U.S. tax incentives and the E.U. carbon pricing mechanism is a novel approach to entering the market.

**IMPROVEMENT**

The decarbonization pathway capable of producing the greatest economic returns was identified as point-source carbon capture. Producing low carbon cement with this technology for export to the E.U. generates economic returns larger than those achieved with conventional methods. For the domestic market, with an approximately 50% reduction in carbon capture CAPEX, economic returns can also become competitive with those of conventional methods.

**BEST PRACTICES**

A full understanding of the both the technology landscape as well as the strategic objectives of the interested company is required to ensure that the technologies pursued align with the business need. In addition, thorough interrogation of public cost data is required to justify appropriate back-to-back comparisons. Finally, the assessment of electricity prices on an hourly basis is critical for evaluating thermal storage technologies.

Logistics Management | Retail | Network optimization | Deterministic / Stochastic

# Optimizing Inventory Rebalancing: Strategies for Managing Excess Inventory in a Dynamic Supply Chain

## BUSINESS PROBLEM

In recent years, Nike has focused inventory into regionalized nodes closer to digital consumers to improve speed and service for its direct-to-consumer (DTC) channel. These smaller distribution centers must operate with agility to meet customer demand while maximizing capacity, especially by minimizing excess inventory. This study investigates whether changes to inventory policies and rebalancing strategies, such as enabling transfers between regional centers (inventory rebalancing) and streamlining excess inventory policies, can reduce excess inventory and improve agility, while carefully evaluating the impact on overall supply chain perform

## DATA SOURCES

The data used in this study came from the company's internal inventory database. This included SKU-level data on historical sales and MSRP. A second dataset provided associated cost information, such as fulfillment and transportation costs. All data were aggregated and anonymized to avoid disclosing any confidential company information

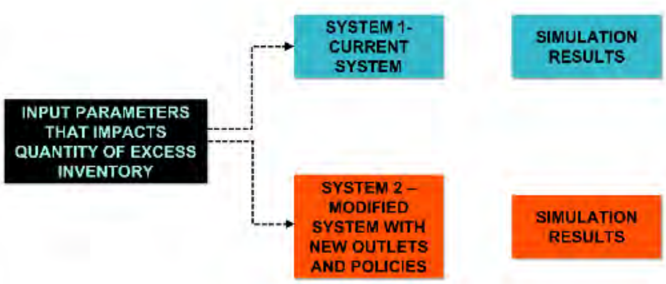
**Data Types and Format**

Microsoft Excel Spreadsheets (via Snowflake Database)

Author: Lanre Oludipe

## APPROACH

I developed a simulation model to compare Nike's current inventory system with a proposed system that featured inventory rebalancing and modified excess inventory policies. Using Monte Carlo simulations, I evaluated each system's performance across key metrics like capacity utilization, replenishment volume, and margin.



## IMPACT

The proposed solution uncovered insights that could have operational and strategic impact for Nike. Simulation results showed that implementing inventory rebalancing can reduce replenishment pressure on the main distribution center (MDC) by 75% by relying on transfers from other regional centers. This not only frees up inventory for other non-DTC channels but also provides additional outlets for regional centers to move excess inventory within the network. The availability of these additional channels reduces the pressure of accurate demand forecasting, as overstocking in one location can later be redistributed across the network through rebalancing.

| | |
|---|---|
| DRIVERS | The primary drivers of this project were the operational constraints of Nike's smaller regional centers. As demand for faster delivery increased, so did the pressure on these regional centers to operate with greater agility. High levels of excess inventory highlighted inefficiencies in inventory flow and policy limitations. The project aimed to identify scalable solutions to reduce inventory buildup |
| BARRIERS | Time constraints limited the ability to fully model the complexities of Nike's supply chain. As a result, simplifying operations into a simulation required assumptions that may not capture all real-world dynamics. These complexities were intentionally scoped out to focus on the most actionable levers within the six-month project timeline. |
| ENABLERS | Strong support from my manager and the Applied Analytics team enabled this project during both the scoping and execution phases. Nike's warm and open culture made it easy to reach out across the organization for brainstorming and idea testing. The company's innovative mindset encouraged curiosity and openness, which helped me gather feedback, incorporate new suggestions, and ultimately deliver a more polished result |
| ACTIONS | I developed a simulation model comparing the current system to a proposed one with inventory rebalancing and modified policies. I collaborated with team members to define key parameters, sourced data from internal systems, and coded the simulation in Python. I then ran Monte Carlo simulations to evaluate performance across margin, capacity utilization, and replenishment metrics |
| INNOVATION | One innovative aspect of the solution was modeling inventory rebalancing between regional centers, an approach not currently implemented at the company but supported in academic literature as an effective method for managing inventory across distributed networks |
| IMPROVEMENT | In 1000 simulations of both systems, the proposed solution improved capacity utilization at regional centers by up to 66% in some weeks and reduced replenishment needs from the main distribution center. These efficiency gain did not come at the expense of margin, in fact, margin increased by 2.2 percentage points due to reduced reliance on discounted factory store sales |
| BEST PRACTICES | Clearly define the scope and prioritize the most impactful levers, especially when working within time constraints, as modeling the entire supply chain may not be necessary. Design your simulation to be modular and testable for easy debugging. Use real historical data to set realistic parameters. Validate assumptions early with stakeholders, and apply Monte Carlo methods to capture demand variability |
| OTHER APPLICATIONS | This approach could be applied to other multi-node distribution networks facing storage constraints, including retail, pharmaceuticals, and e-commerce supply chains. |

# Analyzing Procurement Data for Cost Saving Application

CATERPILLAR

## BUSINESS PROBLEM

Caterpillar spends approximately $25 billion annually on parts procurement across a global network of over 160 facilities and 28,000 suppliers. Due to fragmented data systems and localized procurement practices, implementing centralized, data-driven cost reduction strategies can be a challenge. While the Lifecycle Cost Management (LCM) tool was developed to identify overpriced purchase orders using machine learning, its adoption was limited at first. The core business problem was twofold: improving the accuracy and usability of the LCM tool, and embedding it sustainably into the procurement workflow to drive measurable cost savings.

## DATA SOURCES

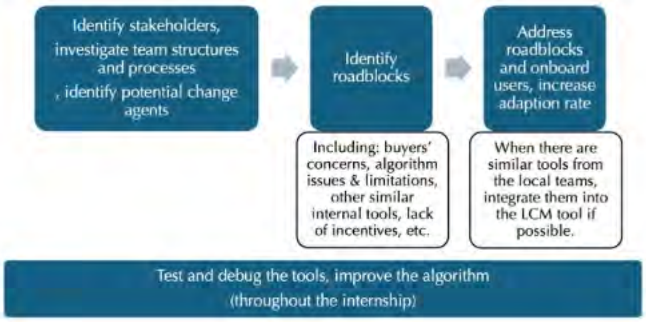
The CAT procurement databases were available to me via Snowflake. There are various ERP systems used at CAT, and the majority of them have been integrated into one central data lake. Other data are collected by interviews and research.

**Data Types and Format**

audio, csv, pdf.

## APPROACH

The project enhanced the LCM tool through: 1. Improving data sourcing and cleaning, exploring machine learning models, and validating cost outliers against other cost models. 2. Iterative UI/UX design and testing. 3 change management strategies including buyer interviews, stakeholder engagement, and pilot testing with international teams to drive adoption.



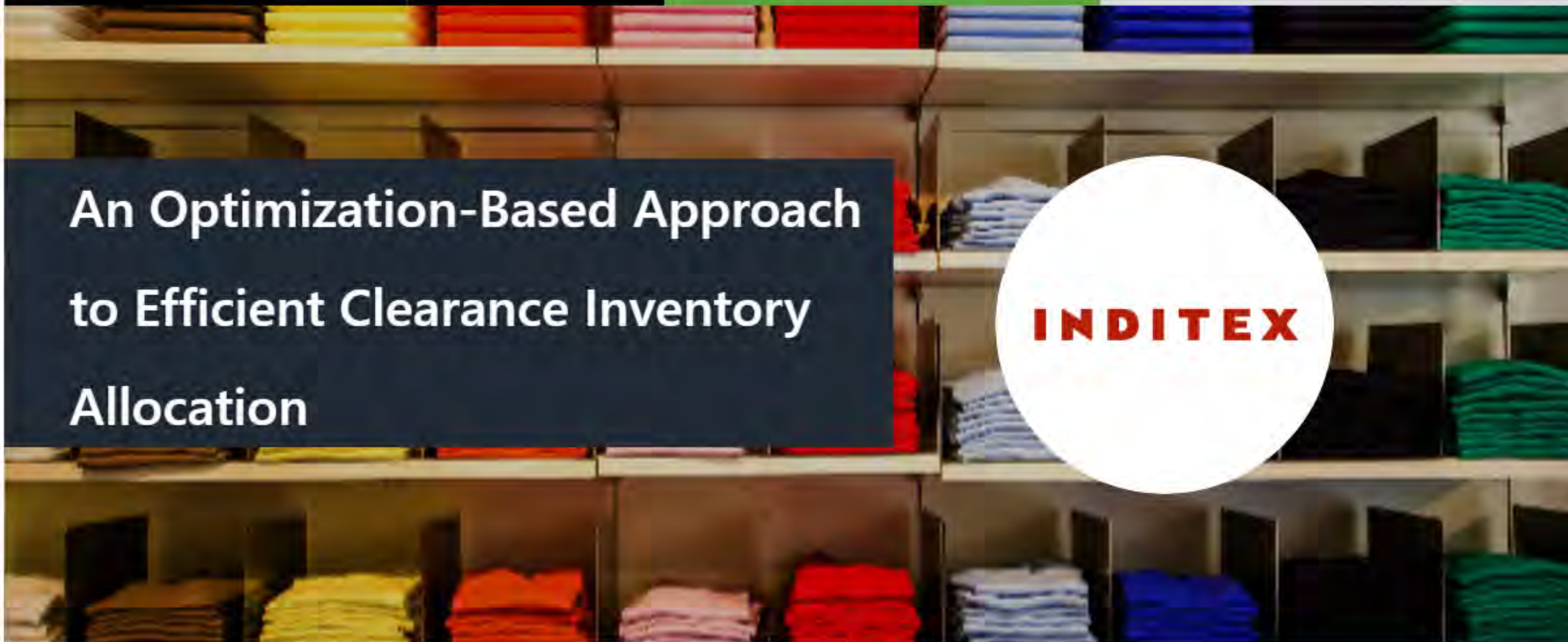

Author: Haoting Pan

## IMPACT

Review and execution teams were established within all four business units and sub-divisions. A clear process flow is established along with review team assignment logic built out. The UI has been redesigned to create a better user experience. A case study with the Construction Industries team in Brazil demonstrated the practical value of the tool: two machine-flagged purchase orders resulted in $160,000 in realized savings after switching to a supplier. Metrics such as validation rate, savings conversion rate, and user adoption were introduced to track performance. The tool now flags opportunities every quarter and features built-in visualization and justification tools to aid decision-making. Adoption has increased from under 10 users to 30+ active users, with ongoing training and stakeholder engagement expanding its reach. By embedding LCM into procurement workflows, Caterpillar is making measurable strides toward more consistent, transparent, and data-driven cost management across its global network.

| | |
|---|---|
| **DRIVERS** | Procurement data can be fragmented and scattered, especially for traditional manufacturing companies. CAT has a global footprint and 100 years of history, making data management challenging. Applying machine learning methods to find cost-saving opportunities and engaging procurement professionals (who are not data scientists) to test and iterate the digital tool were the drivers of this project. |
| **BARRIERS** | Part costs are dependent on both quantifiable variables (location, quantity, materials, etc.) and non-quantifiable variables (such as supplier relationship and geo-political factors), making prediction difficult. Buyer assignment is complex and dynamic at CAT, making assignment of cost-saving opportunities challenging. As a manufacturing company, CAT is not naturally set up to make digital products. The team was in lack of a UI designer. |
| **ENABLERS** | Many procurement professionals (buyers) at CAT have decades of experience and subject matter expertise in cost estimate, which was incredibly helpful for me to understand the business context. They are essential to the pilot program and providing feedback to the LCM tool. CAT has a global footprint and over 2M active parts, which ensures that there are plenty of data to analyze. |
| **ACTIONS** | Using the design thinking approach, I started by interviewing buyers from various divisions and locations to understand their pain points (empathize). Then I defined the problems to solve (buyer assignment, visualization and better backup data). Then, I ideated and took the iterative approach to design the user experience including process and user interface. |
| **INNOVATION** | Instead of establishing KPI through dollar values (i.e. savings realized), I prioritized a faster turnaround in executing cost-saving opportunities. |
| **IMPROVEMENT** | More active users |
| **BEST PRACTICES** | understanding underlying data sources and user pain points, as well as business context. |
| **OTHER APPLICATIONS** | When the cost prediction models generated by LCM reach a higher confidence level and become a routine practice of the procurement teams, the predictions can also be used on new parts procurement. Currently, buyers rely on experience and manually built Should-Cost Models to estimate the cost of a new part. With machine learning algorithms, buyers can have another reference point and potentially streamline the cost estimation process. |

# An Optimization-Based Approach to Efficient Clearance Inventory Allocation




## BUSINESS PROBLEM

Zara, a leading global fashion retailer, faces a recurring challenge during the clearance season: how to optimally distribute remaining inventory across stores to maximize revenue while minimizing overstocking and missed sales opportunities. During this period, product margins can become negative, yet it is still preferable to generate some revenue rather than none, given that production costs are already sunk. The business problem is to determine which articles and how many units to ship from secondary distribution centers to stores during the clearance period, ensuring decisions are made within a tractable timeframe.

## DATA SOURCES

We queried and transformed internal tables to obtain inventory levels, store assignments, demand forecasts, shipment targets, and clearance group mappings. Additional data included dynamic price and margin info, e-commerce demand estimation, and historical sales to refine forecasts and ensure accurate, margin-aware distribution planning.

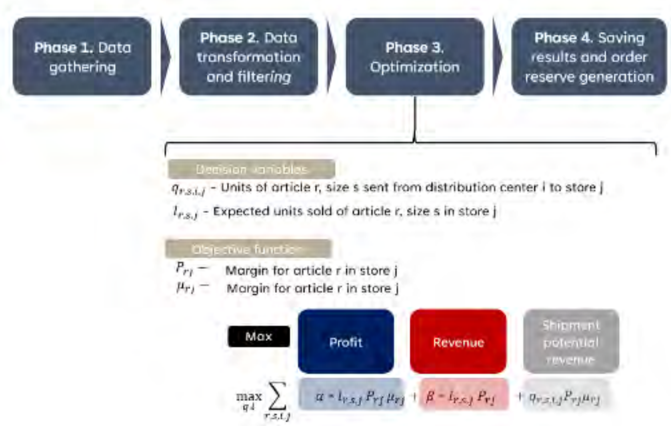
### Data Types and Format

The data was primarily structured in tabular format, and all transformations were handled in Python using the Databricks environment.

Author: Karla Perez Muñoz

## APPROACH

The challenge of optimizing the distribution of clearance inventory at Zara was solved by developing a data-driven optimization model. We formulated a mixed-integer linear programming model that incorporates business constraints such as product availability and expected demand. The model was implemented in Python and solved with Gurobi, using an optimized approach to ensure computional efficiency.



## IMPACT

The implementation of the proposed optimization model to the secondary distribution centers demonstrated a clear improvement over the baseline distribution strategy. By leveraging store-level demand signals and incorporating margins, the model achieved a more balanced and targeted allocation of clearance inventory across the store network maximizing the potential profit. This led to better product availability in high-demand locations and a reduction in excess inventory in lower-performing stores. Quantitative results revealed a 4.5 percentage point absolute increase in margin for the optimized shipment compared to the single-item approach currently in place. This uplift highlights the financial potential of the model during the clearance season—a period typically associated with lower profitability. Beyond the numerical gains, the project also demonstrated the feasibility of building a tractable, scalable, and data-driven model that can support real-time decision-making during a commercially sensitive period. We successfully reduced computation times, enabling the model to be run multiple times per day using the most up-to-date information.

**DRIVERS**

In the retail-fashion industry, where product life cycles are short and new inventory constantly replaces old, clearing excess stock efficiently is crucial. This highlighted the need for a more data-driven, scalable, and responsive distribution strategy. The nature of the fashion industry was the catalyst for our solution, driven by the need to make allocation decisions rapidly—often multiple times a day—to respond to dynamic demand.

**BARRIERS**

The development of the optimization model faced several challenges, including fragmented data from multiple internal systems that required extensive cleaning and alignment. Weekly price fluctuations added complexity to margin calculations, requiring frequent input updates. Additionally, the model had to remain computationally efficient to support multiple daily runs with real-time data.

**ENABLERS**

Despite these challenges, several factors enabled the project's success. Centralized data architecture made it easier to access key information, and support from the Business Analytics team helped guide and validate the approach. Tools like Python and Gurobi allowed for fast, flexible implementation, while standardized exchange rates and definitions simplified data integration.

**ACTIONS**

Relevant data were cleaned and transformed to ensure accurate pricing, inventory, and demand inputs. A mixed-integer linear programming model was developed to optimize clearance product allocation from distribution centers to stores, incorporating constraints on stock, demand, and capacity. Implemented in Python and solved with Gurobi, the model enabled efficient daily runs and was benchmarked against existing heuristics.

**INNOVATION**

A key innovation was the integration of price and margin dynamics directly into the optimization model, allowing it to prioritize more profitable allocations. Another novel aspect was the refined treatment of e-commerce data, which was previously aggregated at the distribution center level. Finally, the code was optimized to efficiently handle large datasets and ensure fast execution.

**IMPROVEMENT**

The optimized solution yielded a substantial improvement in allocation performance. Specifically, the model demonstrated a 4.5 percentage point absolute increase in shipment margin compared to current heuristic approaches. Additionally, the model improved operational efficiency by enabling faster, more consistent decisions supported by real-time data inputs.

**BEST PRACTICES**

First, ensuring consistent and accurate data preprocessing is critical—especially when dealing with dynamic pricing and regional data variations. Second, building a scalable and tractable model that can run quickly is essential for integration into real-world decision-making cycles. Incorporating margin calculations and disaggregating e-commerce trends also added substantial value.

**OTHER APPLICATIONS**

While the model was designed for clearance inventory distribution, its framework is flexible and can be extended to other use cases. For example, it could support initial stock allocation for new product launches, replenishment planning, or real-time inventory balancing between e-commerce and brick-and-mortar stores. It can also be extended to other industries where there's a need to take decisions regarding shipments in short time.

Production | Aerospace and defense | Transportation optimization | Network analysis

# Discrete Event Simulation as a Predictor for Factory Traffic Management

## BUSINESS PROBLEM

Currently, AGV movements rely on operator remote-control input, which enables material flow in the factory but limits the vehicles' autonomous potential and can still create traffic congestion. Increasing automation and incorporating real-time material tracking would smooth the flow and equip stakeholders with the information they need for proactive decision-making.

## APPROACH

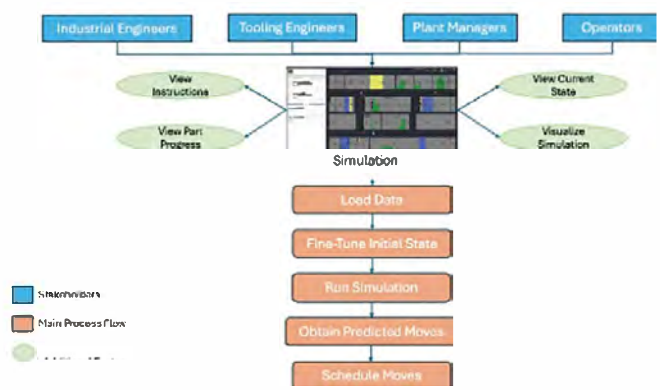
The problem was approached by developing a discrete event simulation using SimPy with a React.js frontend to create a digital twin of the factory. This solution models material flow and predicts AGV/crane/cart movements with 89.6% accuracy. I incorporated path planning algorithms, custom business rules, and visualization tools to provide factory visibility and optimize scheduling.

## DATA SOURCES

Primary data is captured through operator-recorded "stamps" in Boeing's database that log part progress at each milestone. Supplementary inputs include Excel-based priority lists, Gamma part configurations, and time-study results.

### Data Types and Format

Qualitative interview notes, Excel spreadsheets, time studies, process information from SQL database.




Author: Esteban Ramirez Echavarria

## IMPACT

The discrete event simulation solution delivers significant business impact at Boeing's Composite Wing Center through improved operational visibility and predictive scheduling. With 89.6% accuracy, it enables proactive AGV movement coordination, reducing congestion and allowing better maintenance planning. Management can evaluate potential modifications through scenario testing before implementation, transforming capital decisions into data-driven processes. The system establishes the foundation for incremental automation while preparing for production increases by identifying bottlenecks before they become critical. Its modular architecture enables integration with other factory systems, promoting a data-driven decision-making culture that improves resource utilization and positions the facility for future production demands.

| | |
|---|---|
| **DRIVERS** | Traffic congestion can be mitigated by predictively evaluating material flow using simulation software. Industry-wide advances in automation (Industry 4.0) and the push for data-driven decision-making in manufacturing further guided solution development, particularly given the intricate logistics of moving large composite wing components. |
| **BARRIERS** | Key challenges centered on harmonizing multiple data sources, building resilient validation protocols, and developing a detailed model flexible enough to capture the factory's wide array of workflows. Addressing these hurdles required close collaboration between engineering teams and targeted upskilling to streamline data integration and maintain analytic rigor throughout the project. |
| **ENABLERS** | Theprojectwas enabledby open accesst ofactoryoperations,supportiveand collaborativeteams,andleadershipthatvaluedexperimentationwithdigitaltools. Primary enablers were the AGV Management team and the Industrial Engineering team, who possess crucial information for the development of the simulation. |
| **ACTIONS** | Conducted interviews with stakeholders to understand needs. Evaluated approaches including Machine Learning and Simulation. Scoped the project based on feasibility. Developed a discrete event simulation in Python, validated it with historical data, and used it to explore traffic flow scenarios and optimizations. |
| **INNOVATION** | The innovative aspects include using discrete event simulation as a digital twin to predict factory traffic management, integrating real-time data for validation, and leveraging Python's object-oriented design for modularity and scalability. The approach combines traditional simulation with data-driven insights to optimize AGV movement and resource allocation in a dynamic environment. |
| **IMPROVEMENT** | The solution enables prediction of AGV traffic patterns, identification of bottlenecks, and evaluation of "what-if" scenarios. It provides a solution for visibility, and enables data-driven scheduling. |
| **BEST PRACTICES** | Deeply understanding the factory layout and constraints through stakeholder interviews. Use modular, object-oriented code for flexibility. Validate the simulation with real data early and often. Keep the scope focused, iterate quickly, and align outputs with operational KPIs to ensure relevance and adoption. |
| **OTHER APPLICATIONS** | The simulation can be applied to factory enhancement analysis, bottleneck reduction, and AGV policy optimization. It enables management to make data-driven investment decisions by quantifying the impact of layout or scheduling changes on production rates. |

Operations | Private Equity | Predictive forecasting | Linear / Non-Linear

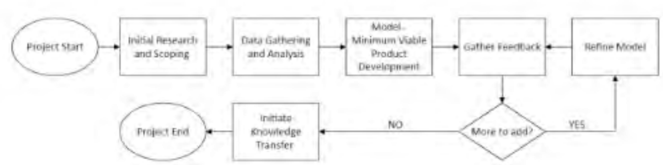
# Hydrogen Adoption Dynamics: A Flexible Modeling Framework for U.S. Industrial Applications

## BUSINESS PROBLEM

Company X aims to map the ecosystem and value-chain of emerging decarbonization technologies, with a specific focus on hydrogen. The company seeks to develop a better understanding of the hydrogen economy to generate a diversified investment thesis for new Platform or Add-On investments. This understanding will enable Company X to identify opportunities within or adjacent to the hydrogen economy, focusing on areas most relevant to their portfolio companies. By thoroughly analyzing the hydrogen technology landscape, Company X can position itself strategically in a rapidly evolving sector, supporting both its investment strategy and ESG pillar.

## APPROACH

The researcher developed a flexible modeling framework for hydrogen adoption across multiple industries. The approach included data collection from industry sources, creating demand models for six industrial sectors, developing a price framework with policy and technology scenarios, implementing scenario analysis, and conducting sensitivity analysis to identify key market drivers.




## DATA SOURCES

The data came primarily from established industry and government sources. The research drew extensively from the Argonne National Laboratory report "Assessment of Potential Future Demands for Hydrogen in the United States" for most industrial applications, and the Department of Energy report "DOE National Clean Hydrogen Strategy and Roadmap" for energy storage projections.

**Data Types and Format**

numerical

Author: Jennifer Ray

## IMPACT

The solution provides Company X with a strategic framework that identifies potential investment opportunities across the hydrogen value chain by revealing which industries are likely to adopt hydrogen first and which will follow later based on economic viability. It provides insight into market timing, showing the expected sequence of industry adoption as hydrogen costs decline. The framework quantifies the potential market size under various scenarios, demonstrating how government support and technological advancement could dramatically expand the hydrogen economy. It highlights key market drivers through sensitivity analysis, identifying which parameters most significantly affect overall hydrogen demand. This flexible decision-making tool allows Company X to navigate market uncertainties as the decarbonization landscape evolves, rather than providing a single-point forecast that would quickly become outdated. The tool's value lies in providing a framework that helps stakeholders understand how different variables could affect the hydrogen market, giving Company X a competitive advantage in identifying promising investment opportunities in the hydrogen economy.

**DRIVERS**

The catalyst for the solution was Company X's need to understand the emerging hydrogen economy for strategic investment decisions. As decarbonization efforts accelerate, hydrogen represents a significant opportunity, but the market timing, adoption sequence, and economic viability across industries remained unclear. Company X needed a flexible framework to identify promising investment opportunities within this rapidly evolving sector.

**BARRIERS**

The research faced several challenges, including limited public data on hydrogen adoption rates, rapidly evolving technology assumptions, and uncertainty around future government policy support. The interdependencies between industry sectors further complicated modeling efforts, while geographic variability in production costs created additional complexity in developing accurate market projections.

**ENABLERS**

Company X's investment expertise and portfolio company relationships provided valuable industry insights and context. The team's commitment to understanding decarbonization opportunities created strong support for my work. The company's emphasis on flexible, adaptable tools aligned perfectly with my modeling approach.

**ACTIONS**

The researcher developed a comprehensive hydrogen adoption model in Microsoft Excel, prioritizing usability, transparency, and adaptability. The model analyzes six key industries with flexible inputs and clear documentation to ensure it could be easily updated as market conditions evolve. This approach allowed stakeholders to test various scenarios without requiring specialized technical knowledge.

**INNOVATION**

The most innovative aspects of the solution include the multi-scenario framework that combines both policy support and technological advancement variables, the flexible input system that allows for parameter adjustment as market conditions evolve, and the integration of threshold price activation mechanisms that model real-world economic decision points for industry adoption. This approach avoids the "black box problem."

**IMPROVEMENT**

The solution provides a flexible modeling framework that helps Company X navigate investment decisions in the evolving hydrogen market. It reveals adoption sequences across industries, identifies key market drivers, and quantifies potential market size under various scenarios. This adaptable tool allows stakeholders to test assumptions and update inputs as market conditions change.

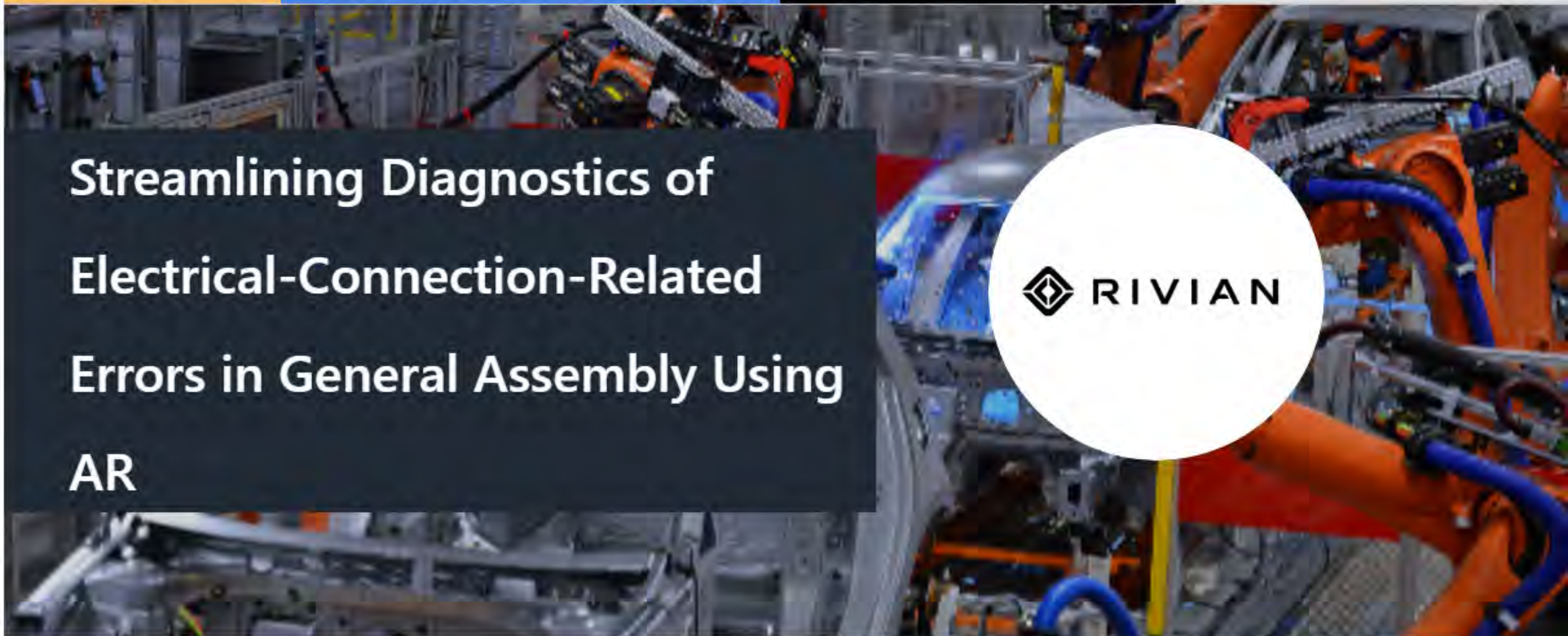
**BEST PRACTICES**

To model hydrogen adoption, one should create a flexible tool that tracks multiple industries, compares prices against thresholds, tests different scenarios, identifies key growth drivers, and stays adaptable to market changes. The goal isn't a perfect prediction, but understanding possible futures by exploring how technology, policy, and economics interact.

**OTHER APPLICATIONS**

The modeling framework provides practical tools for analyzing industrial hydrogen transitions, though its accuracy depends on continuously updated data. It helps stakeholders test policy scenarios and investment timing, but should be viewed as one of several decision-support tools rather than a precise predictive model. Its greatest value comes from exploring interactions between technological advancement, policy support, and economic thresholds.

Quality | Automobile manufacturing | Error Diagnostics | Statistical Analysis



# Streamlining Diagnostics of Electrical-Connection-Related Errors in General Assembly Using AR

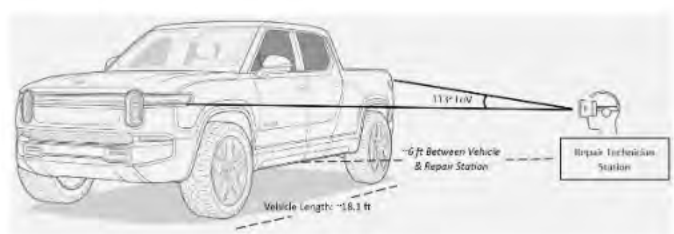
RIVIAN

## BUSINESS PROBLEM

An electrical-connection-related error (ECRE) is an issue that arises that could impair vehicle functionality. To minimize rework costs, it is critical to diagnose and repair these errors during the initial Trim line stations in General Assembly (GA) when wiring harness are still exposed and easily accessible. Currently, the time required to locate the connector containing the error often exceeds the station cycle time. As a result, repairs for ECREs are delayed until after the entire assembly process is complete, leading to significantly higher repair times and increased costs in both labor and replacement parts.

## APPROACH

The approach began with identifying key stakeholders with subject matter expertise in ECREs or those who would use the new diagnostic process. Collaborating with these stakeholders, the most common ECREs were selected for prototyping. Extensive research highlighted AR capabilities as a promising solution, leading to hardware selection and prototype development of the AR Overlay Process (AROP).




## DATA SOURCES

The Low-Voltage Systems engineers track ECREs in a database. The information stored in the database provides insight into 1) where in the electrical system harness the error is located, 2) the category of ECRE, such as a missed or incomplete connection, and 3) where in GA the error is diagnosed and repaired, such as Trim lines or rework bays.

### Data Types and Format

The data type is categorical and stored as text, with values limited to predefined categories.

Author: Lizzy Salata

## IMPACT

The majority of error diagnostic standard operating procedure time is spent locating the connector containing the error, specifically, the time spent reviewing lineside instruction screens and walking around the vehicle to match repair plan images with the actual vehicle. This lengthy diagnostic process has downstream consequences: only 42% of all ECREs are diagnosed and repaired within the takt time on Trim lines. The remaining 58% are repaired in the rework bays after GA, where the average repair time can take 13.5 times longer on average than on Trim lines. Furthermore, 70% of all ECREs are due to missed or incomplete connections. Once these errors are located, the repair itself takes less than 5 seconds, requires no tools, and demands minimal subject matter expertise. Assuming the diagnostic and repair location distributions for missed and incomplete connections align with those of all ECREs, 40.6% of all ECREs fall into this category and are not repaired until they reach the post-GA rework bays. The experiment conducted for this master's thesis demonstrates that the AROP process takes and average of 11 seconds to identify the connector and serves as an equally effective and more accessible diagnostic method. It reduces connector location times by 75%, improves consistency, and reduces overall cognitive load. Additionally, the AROP has the potential to prevent 40.6% of all ECREs from requiring lengthy, costly, and inefficient repairs.

**DRIVERS**

Rivian prioritizes innovation and novel problem-solving, with error reduction being crucial for profitability. AR has proven effective in improving task performance, reducing cognitive load, and enhancing training and diagnostics. Experimental studies analyzing qualitative and quantitative data demonstrate AR's impact on manufacturing efficiency. Leading automotive companies have successfully leveraged AR to address challenges in other domains.

**BARRIERS**

Limited computation resources constrained the AR device's functionality, affecting processing speed and overall performance. Additionally, the availability of participants was limited, making it challenging to fully validate the proof-of-concept prototype. Environmental constraints in the experimental setup further restricted which connectors could be tested, limiting the scope of the evaluation.

**ENABLERS**

Rivian's culture of innovation and novel thinking fostered an environment that supported this project. Key stakeholders had a deep understanding of ECREs and their impact on manufacturing, which led to an initial level of buy-in. Additionally, Rivian's design studio had already been using AR for vehicle design iteration, providing a valuable resource base for developing and refining the prototype.

**ACTIONS**

Key relationships were built with four stakeholder groups. Low-Voltage Systems engineers provided insights into the scope of the problem, while repair technicians helped map the current standard operating procedure. Manufacturing line operators shared their needs and potential hesitations as future users of the tool. Additionally, Visualization Design Studio designers contributed by sharing expertise on AR capabilities.

**INNOVATION**

The AR solution streamlines the diagnostic process for ECREs, reducing cognitive load and improving task performance by providing visual and spatial guidance for error identification. Notably, the solution is accessible for users with no prior experience in repairs. The solution also leverages existing AR technology from Rivian's design studio, applying it in a new context for vehicle assembly, demonstrating an adaptive use of resources.

**IMPROVEMENT**

The AROP reduced connector identification time by 75%. With this significant decrease, it has the potential to prevent 40.6% of all ECREs from requiring lengthy, expensive, and inefficient repair processes.

**BEST PRACTICES**

Prioritize a human-centered design approach, especially when working with emerging technologies. Since AR hardware is not yet fully optimized for automotive manufacturing environments, it is crucial to involve key stakeholders early and maintain constant communication with them throughout the process. This ensures that new procedures are developed with user needs in mind, increasing the likelihood of successful adoption.

**OTHER APPLICATIONS**

AR can enhance visual inspections, such as measuring body flushness and gaps, ensuring specified quality tolerances are met. Its image processing can verify the accurate printing of regulation decals, critical for compliance. For training, AR offers a cost-effective alternative to using actual parts, enabling new hires to learn without disrupting production. Additionally, AR can verify future product and manufacturing plant ergonomic designs.

Operations | Metals and mining | Improved process performance | Ranking Systems

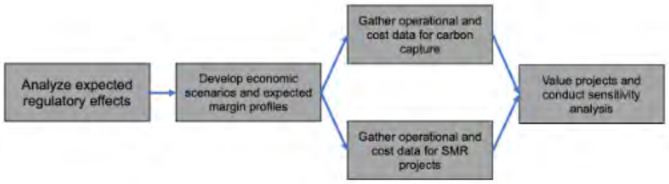
# Towards Green Aluminum

## BUSINESS PROBLEM

Recent changes in the regulations governing industrial carbon dioxide emissions in Europe are set to increase production costs for aluminum smelters. Aluminum smelting is a carbon intensive process, and the reducing emissions necessarily demands technological changes at existing smelters. Therefore, smelters such as the anonymized Aluminum of Europe must develop means to project the margin impacts of changing regulations and the value of the emissions reduction technologies available to them to make appropriate investment decisions and operational plans. This thesis aimed to build the methodology to inform operational decarbonization plans.

## APPROACH

I used global aluminum production cost data and emissions factors to model the effect of carbon cost implementation on the European supply curve. I projected the margin effects of the new regulations. I used these values and the technical specifications of a set of emissions reduction technologies to conduct a discounted cash flow (DCF) analysis and produce project valuations.




## DATA SOURCES

The data used for this project came from a combination of information collected onsite and recorded by Aluminum Dunkerque and a set of data gathered from industry and academic publications.

**Data Types and Format**

Tabular and timeseries data recorded historical production values at Aluminum Dunkerque. Industry and academic sources provided costs and operational details for the technologies under evaluation.

Author: Kevin Schurr

## IMPACT

The results of this work are two-fold. The economic analysis revealed that due to its position as a low-cost and low-carbon aluminum smelter, Aluminum of Europe is well-positioned to maintain its margin profile even as its carbon costs increase. The technological evaluation suggested that either carbon capture or integration of an on-site small modular nuclear reactor (SMR) could result in meaningful emissions reductions. However, only carbon capture is economically feasible in the short term and would require either very high carbon quota prices or significant subsidies. SMRs would need to achieve major cost reductions to make them competitive with grid electricity.

| | |
|---|---|
| **DRIVERS** | Aluminum is an important material to aid in the energy transition. As one of the world's lowest carbon smelters, Aluminum Dunkerque contributes to the emissions reductions of the economy already. Further, to remain profitable, Aluminum Dunkerque must avoid implementation of any systems that would depress its output. |
| **BARRIERS** | The largest barrier to the implementation of emissions reduction technology at Aluminum of Europe is the immaturity of the technology that could do it. Carbon capture systems that could curtail process emissions are not well-suited to low carbon dioxide concentration gas streams. SMRs that could curtail emissions from electricity are in their infancy. There is currently little data to support investment in these technologies. |
| **ENABLERS** | Aluminum of Europe has a team focused on engaging with technology providers, building operational plans, and testing carbon capture systems. Additionally, AOE's owner has a dedicated ESG pillar that provides support and guidance to companies undertaking sustainability-focused initiatives. These factors helped accelerate the work on this techno-economic analysis. |
| **ACTIONS** | I engaged with several companies that offer emerging technological carbon capture or SMR technologies and assessed their applicability to Aluminum of Europe's use case. I developed valuation models to inform the ongoing decarbonization and technology integration at the smelter. |
| **INNOVATION** | Neither carbon capture nor SMR power have been implemented for emissions reduction purposes at any aluminum smelter globally. This project clarifies the economic and technological conditions under which these types of projects will become feasible for industrial businesses. |
| **IMPROVEMENT** | Aluminum of Europe's emissions reduction program will continue through 2050 as it aims for a 70% reduction in its emissions. The technical and operational evaluation of decarbonization solutions informed a discounted cash flow analysis. This valuation suggests that under the correct conditions, this carbon reduction goal is achievable and can generate an additional €83 million in free cash flow for AOE. |
| **BEST PRACTICES** | This project necessarily involved engaging with technology providers who operate outside of the core functionality of an aluminum smelter. Prior to these engagements and decisions to pursue partnership, it is vital to conduct thorough investigations of academic and industry literature. This allows for more fruitful and candid discussions and helps the smelter more quickly down select to the most feasible, affordable providers. |
| **OTHER APPLICATIONS** | As other industries and regions embark on emissions reduction efforts, the use of carbon capture and low-carbon on-site power generation will be increasingly important. Carbon capture system integration will likely become more common outside of its current main use cases in the oil and gas industry. Large technology companies are already investing heavily in SMR technology to power large-scale data centers. |

# Multi-Objective Optimization of Container Load Plans for Modulating Inventory Flow

## BUSINESS PROBLEM

Global retailers face growing challenges in managing the flow of goods from manufacturing sites to distribution centers. Conventional container load planning (CLP) approaches focus solely on maximizing container utilization, which overlook broader supply chain considerations and lead to mistimed delivery, inventory pileups, higher storage and transportation costs, and operational inefficiencies at distribution centers. The core business problem is to rethink how CLP can be optimized to balance cost efficiency, space utilization, and delivery timing to lower costs, improve service levels, and ensure smoother inventory flow.

## APPROACH

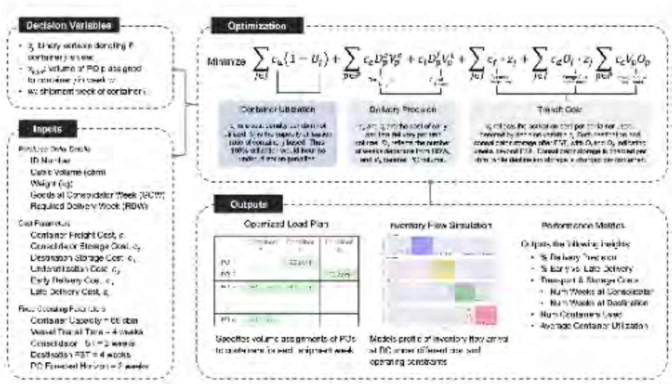
A multi-objective MILP-based emulator using a branch-and-bound structure was developed using a dataset of 200K+ purchase orders. The model evaluates trade-offs between transportation and storage costs, container utilization, and delivery precision and generates an optimized load plan that specifies container selection, shipment week, and allocation of purchase order volumes per container.

## DATA SOURCES

The model uses 200K+ purchase order records from a high-volume shipping lane. Data was sourced from ERP systems (order volumes, product details), sales order systems (required delivery dates), and transportation management systems (shipment timestamps), providing full visibility into baseline container usage, shipment timing, and purchase order volume allocations.

### Data Types and Format

Structured tabular data, including time series (PO volumes by shipment week), categorical variables (product type), and numerical fields (purchase order volume and quantities).




Author: Shweta Sen

## IMPACT

The solution enables a more strategic and prescriptive approach to container load planning (CLP). Rather than relying on FIFO or manual consolidator decisions, businesses can assess the impact of different loading configurations on key metrics such as delivery precision (early vs. late delivery rate), container utilization, transportation costs, and storage fees. The model allows logistics practitioners to evaluate metrics trade-offs, simulate different planning scenarios in real time, and visualize inventory flow across downstream warehouses and distribution centers. Simulation results highlight that rigidly maximizing container utilization can lead to high early delivery rates, inflating inventory holding costs and causing warehouse congestion. By balancing multiple operational objectives, the model reduces deviations in inventory arrival time and decreases pre-season inventory surges. The model also enables targeted policy analysis—for example, testing the business impact of relaxing co-loading restrictions or shipment volume thresholds—to understand which constraints can be modified to improve performance. Overall, the solution provides logistics practitioners with a decision-support tool that customizes container load planning with broader operational goals.

**DRIVERS**

Key drivers faced by the global retail industry include rising freight transportation and inventory storage costs, volatile freight rates, and increasing complexity in inventory flow management due to seasonal demand and omnichannel fulfillment.

**BARRIERS**

Key barriers included limited visibility into third-party consolidator decision-making, lack of standardized digital load plans, and fragmented data across systems. Modeling complexity was also high due to multiple, often conflicting objectives and operational constraints (e.g., co-loading rules, shipment thresholds).

**ENABLERS**

Key enablers included strong cross-functional collaboration across supply chain, logistics, and data teams, as well as access to high-quality, granular PO-level data across systems. Leadership support for innovation in transportation planning and openness to testing advanced analytics also created the runway for model development and scenario simulation.

**ACTIONS**

A site visit to a major shipping port provided firsthand insight into container operations. Interviews with subject matter experts allowed us to map the end-to-end load planning process to identify decision points and visibility gaps. These insights informed the development of the optimization model, integration of cross-system data, and simulation test cases for scenario analyses.

**INNOVATION**

Innovative aspects include a multi-objective MILP-based emulator that balances container utilization, delivery precision, and transportation and storage costs—moving beyond traditional volume-maximization models. The model supports real-time scenario simulation, policy testing (e.g., load rule relaxation), and visualizes downstream inventory flow, enabling more adaptive and data-driven load planning decisions.

**IMPROVEMENT**

The container load plans generated through the optimization model reduced early deliveries by up to 35%, increasing delivery precision (the percent of purchase orders arriving exactly on their target delivery week) by 73.9% from the baseline. The model also demonstrated a roughly 9% reduction in the number of containers used and fewer pre-season inventory peaks.

**BEST PRACTICES**

Best practices for replicating the approach include: 1) full process mapping to identify where optimization will have the most impact, 2) start with a focused lane or use case to pilot the model before scaling, 3) incorporate business constraints (like co-loading rules) to ensure practical applicability, and 4) enable scenario testing to build stakeholder buy-in and support decision-making.

**OTHER APPLICATIONS**

Potential applications include truck load optimization for middle- and last-mile operations. There is also opportunity in supporting strategic network design by simulating flow patterns under different flow models or demand scenarios. Additionally, the modeling framework can be extended to multi-echelon supply chains, enabling synchronized planning across factories, consolidators, and distribution centers.

Product design / development / engineering | Oil and gas | Life Cycle Costs | Recommender Systems

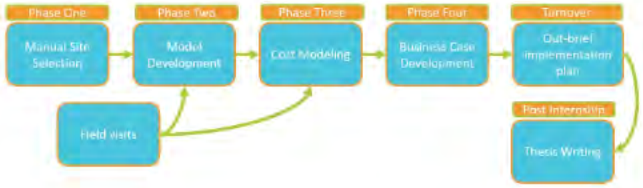
# Gas Network Preparations for Networked Geothermal


## BUSINESS PROBLEM

National Grid aims to lead in clean energy by exploring network geothermal systems, which offer efficient heat transfer for space heating compared to air-sourced heat pumps. Using High Density Polyethylene (HDPE) piping in leak-prone pipe replacement programs could enable future geothermal networks, repurposing assets and retaining skilled workers. The company seeks to identify areas for HDPE upgrades and develop criteria to justify the costs of installation, service, and repair for geothermal reuse.

## APPROACH

National Grid's project to select sites for HDPE piping in geothermal systems has four phases: Phase 1 identifies candidate sites using current processes. Phase 2 gathers design needs from existing projects and research. Phase 3 calculates the extra cost of HDPE vs. MDPE replacement. Phase 4 uses Phases 1-3 data to model site suitability and build a business case for HDPE adoption.




## DATA SOURCES

Data came from various gas engineering teams, field visits, and leveraging the business connections of the gas engineering teams. Data including planning models, cost estimators, operator interviews, and material data sheets.

**Data Types and Format**

Data was in form of notes, model outputs, and GSEP government data.

Author: Patrick Serbent

## IMPACT

The transition from natural gas to networked geothermal energy presents a viable and strategic opportunity for reducing carbon emissions while modernizing existing gas infrastructure. This study has explored the feasibility of utilizing high-density polyethylene (HDPE) piping as a future-proofing measure within Massachusetts' Gas System Enhancement Plan (GSEP). By aligning infrastructure investments with long-term decarbonization goals, National Grid and other utility companies can play a pivotal role in advancing sustainable energy solutions. This project has determined that using HDPE is technically feasible, cost competitive, and likely to garner regulatory support, and highlights doing the right thing through big work.

**DRIVERS**

Massachusetts' ambitious climate goals, finite energy dollars, and a desire to minimize stranded assets drove this problem.

**BARRIERS**

A lack of defined regulations for future heating systems, and the subsequent lack of cohesive plan impacted the ability to adequately and accurately calculate the net present values of the two materials compared.

**ENABLERS**

Company buy-in, access to systems, and access to leaders enabled the rapid gathering of the available required data.

**ACTIONS**

A turnover of concrete actions to the company enables recommendation implementation.

**INNOVATION**

Looking at costs and benefits beyond the dotted lines of an individual project unlocked a large potential benefit to adopting the new material, which could have been overlooked if a more narrow lens was used.

**IMPROVEMENT**

Further work is required to quantify the exact value of improvements.

**BEST PRACTICES**

Start with a wide aperture, go to the gemba, ask the "dumb questions" to everybody, and make it known everybody is playing for the same team.

**OTHER APPLICATIONS**

Other utility infrastructure could have similar repurposing potential during install. Technology is rapidly evolving, so planning a system around upgradability may deliver cost benefits by reducing risk.

R&D | Pharma / biopharma | Predictive forecasting | Natural language processing

# Machine Learning and Biosecurity in the Age of Pandemics: Advancing Biological Research & Safeguards

## BUSINESS PROBLEM

The Broad Institute faces dual operational challenges in biological sequence analysis: computational constraints and implementation of new biosecurity protocols. Current state-of-the-art models for sequence analysis require massive infrastructure - often hundreds of GPUs running for weeks - making them inaccessible to many research environments. This computational burden creates operational bottlenecks and significant infrastructure costs. Additionally, new federal nucleic acid screening requirements threaten to introduce delays in critical research workflows, particularly during public health emergencies where rapid response is essential.

## DATA SOURCES

All datasets used in this work were downloaded from publicly available sources, with citations provided throughout the paper as applicable.

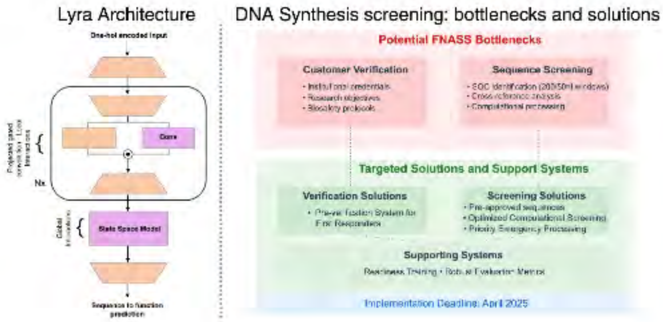
**Data Types and Format**

Genomics data: RNA, DNA, or protein sequences were used, alongside with functional properties associated with those sequences in tasks such as RNA splice site prediction or protein fluorescence.

Author: Sameed Siddiqui

## APPROACH

The project employs a comprehensive strategy centered on the development and deployment of Lyra, an efficient architecture for biological sequence analysis, alongside operational protocols for biosecurity implementation. The Lyra architecture introduces a novel approach to sequence modeling that achieves state-of-the-art performance, while in biosecurity we propose solutions to current hurdles.



## IMPACT

In this work, we introduce Lyra, a highly efficient machine learning architecture for biological sequence modeling; and analyze national biosecurity frameworks for synthetic biology, with notable implications for biotech and the Broad Institute. Lyra's Impact on Biotech: Lyra achieves state-of-the-art (SOTA) results in crucial tasks in biological sequence-to-function modelling; including in CRISPR guide prediction, protein activity prediction, and RNA structure prediction, using vastly fewer computational resources—sometimes over 100,000 times fewer parameters than recent foundation models in biology. This efficiency increases access to advanced modeling, potentially accelerating research and development in therapeutics, diagnostics, and basic biological understanding. Lyra is grounded in the biological principle of epistasis and offers a novel mathematical approach to sequence modeling that enables state-of-the-art results while democratizing access to researchers. Biosecurity and Research Agility: The thesis also addresses biosecurity, drawing on the author's and lab's first-hand experiences as infectious disease researchers responding to pandemics. It analyzes the Federal Nucleic Acid Synthesis Screening Framework, identifying potential bottlenecks. To balance rapid research with safeguards during outbreaks, it proposes solutions like pre-verifying researchers and emergency priority channels for vital synthetic DNA orders.

**DRIVERS**

Limitations in existing machine learning models for analyzing complex biological sequences, such as DNA, RNA, and proteins, spurred the development of the novel "Lyra" architecture. Furthermore, an urgent need to prevent the potential misuse of powerful nucleic acid synthesis technologies while maintaining agility in outbreak response prompted an analysis of current DNA synthesis screening frameworks.

**BARRIERS**

Existing machine learning models often present significant barriers due to their high computational costs and extensive data requirements, which limit their practical application, especially for analyzing long biological sequences. The project's original research scope, which included developing models for specific diagnostic assays, was altered because of challenges related to data availability and confidentiality.

**ENABLERS**

The collaborative and interdisciplinary research environment at the Broad Institute, coupled with the Sabeti Lab's specialized expertise in genomics, computational biology, and machine learning, were key enablers for this project. Mentorship from Pardis Sabeti, Jim Collins, and Roy Welsch, along with vital collaborations, particularly with Krithik Ramesh on Lyra and Jon Arizti-Sanz, Arya Rao, and Al Ozonoff on biosecurity research, was crucial.

**ACTIONS**

Key actions included the development of the Lyra machine learning architecture, which is biologically-grounded in the principle of epistasis and utilizes state space models (SSMs) with projected gated convolutions; this involved creating new mathematical theory and conducting extensive experimental validation. Another significant action was the detailed analysis of the federal nucleic acid synthesis screening framework to propose new solutions.

**INNOVATION**

The Lyra architecture represents a significant innovation in biological sequence modelling, which achieves state-of-the-art performance with substantially fewer parameters and faster inference times compared to traditional transformer models. These advancements make high-performance biological sequence modeling more accessible to biologists.

**IMPROVEMENT**

The Lyra architecture demonstrated significant quantifiable improvements by achieving state-of-the-art performance on various biological tasks while using up to 120,000-fold fewer parameters and running approximately 64 times faster in inference than comparable transformer-based models.

**BEST PRACTICES**

For developing Lyra-like models, best practices include grounding the architecture in established biological principles like epistasis and employing mathematically principled designs such as using SSMs for polynomial approximation. It's also critical to organize your data well, and to code cleanly and reproducibly :)

**OTHER APPLICATIONS**

The Lyra architecture has potential applications in therapeutic development, such as accelerating the design of cell-penetrating peptides and optimizing viral vectors for gene therapy. It can also be applied to pathogen surveillance and diagnostics, including rapid viral protein identification and the optimization of PCR and LAMP diagnostic assays.

Quality | Aerospace and defense | Improved process performance | Regression analysis

# Process Optimization and Proactive Quality Control to Increase Investment Casting Throughput

BLUE ORIGIN

## BUSINESS PROBLEM

Blue Castings is Blue Origin's investment casting plant located in the greater Seattle metroplex. The plant wants to increase throughput of the BE-4 engine components which it produces. Increasing throughput at Blue Castings is challenging as investment casting is a complex process that involves long rework times, high natural defect rates, and significant process variability. Inability to increase throughput can lead to schedule delays and inefficiencies that will impede Blue Origin's ability to meet its ambitious production goals to support growth in the commercial space industry.

## APPROACH

The researcher analyzed gaps between goal and actual throughput, identified root causes of the gaps, and investigated solutions. Solutions explored include process improvement to reduce unnecessary rework, cross-functional collaboration to increase design producibility, and use of machine learning to generate insights and aid in decision-making to better prevent and manage defects.


## DATA SOURCES

Data was collected via on-site qualitative interviews, sampling of defect data documented in weld maps, and querying of Blue Origin's manufacturing execution systems and quality management databases. Additionally, manufacturing process data relating to the slurry was collected from Blue Castings' production testing logs – this includes measurements such as viscosity, pH, and conductivity.

### Data Types and Format

Spreadsheets, PDF files of weld maps, Output files from manufacturing execution system, transcripts/notes from interviews

Author: Julia Sircar

## IMPACT

Reducing unnecessary rework on machined surfaces can increase the capacity of the rework work center by up to 70%. This improvement alone could potentially increase overall system throughput by up to 15%. The reduction in rework leads to direct and support labor cost savings. Additionally, improved producibility and defect prevention mechanisms will save money by reducing non-conformances. Additionally, digitizing defect tracking and integrating machine learning for predictive analytics improves data quality and enables better root cause analysis. If implemented, this will facilitate more accurate predictions of rework severity, allowing for better planning and resource allocation. Finally, establishing a producibility working group to formalize design best practices enables better communication between design, manufacturing, and quality teams. This leads to more "castable" designs, reduces defects, and improves overall production efficiency.

**DRIVERS**

The demand for cost-effective, high-throughput production of rockets, combined with pressure to reduce schedule delays and improve manufacturing efficiency, drives the need for innovative defect management and prevention. In response to this rapid growth in commercial space exploration and the resulting pressure on the aerospace industry, Blue Origin has set ambitious goals for BE-4 production.

**BARRIERS**

Limited historical data and inconsistent data capture made it challenging to perform data analysis and modeling. The lengthy casting process (4-6+ months/part) made it difficult to evaluate the impact of improvements within the project timeline. Equipment failures, urgent business needs, and personnel changes caused delays in piloting solutions. Finally, concurrent initiatives made it difficult to isolate the impact of specific interventions.

**ENABLERS**

High willingness to establish cross-functional collaboration between design, manufacturing, quality, and process engineers to improve design producibility. Commitment from Blue Origin and Blue Castings leadership to provide guidance and resources. Engineers with 20+ years' experience who provided valuable insights into investment casting processes. Blue Origin's leadership principles of bias for action, customer focus, and operational excellence.

**ACTIONS**

Created, piloted, and evaluated a process improvement to reduce rework. Created cross-functional group to analyze trends in casting quality issues and create chapter for Blue Origin's design handbook to improve producibility. Conducted cost-benefit analysis of defect tracking methods, piloted digital defect tracking approach, and evaluated potential return on investment of select methods and use of machine learning for predictive analytics.

**INNOVATION**

Use machine learning to predict rework severity at this facility, identify root causes of defects, and enhance Blue Castings' quality control. Evaluate importance of factory modernization to automate defect documentation, improve data accuracy, and enable trend analysis. Apply line-balancing strategies commonly used in the emergency medicine industry to optimize throughput and reduce bottlenecks.

**IMPROVEMENT**

Roadmap for Blue Castings to increase throughput via a mixed-method approach. Casting producibility best practice guide released company-wide. Business case and results to encourage Blue Castings to fully implement machine surface rework reduction to increase overall throughput by up to 15%. Business case and estimated timeline to onboard digital defect tracking software and build enough data to enable machine learning for predictive analytics.

**BEST PRACTICES**

Engage members from design, manufacturing, and quality early to discuss production challenges and possible solutions. Consider the impact and level of effort of more robust data collection than the status quo, and evaluate the return on investment of data improvements. Small-scale pilots and subsequent iteration can validate solutions and build consensus prior to implementation. Create mechanisms for regular review and refinement of processes.

**OTHER APPLICATIONS**

Predictive analytics and digital defect tracking can be applied across many of Blue Origin's commodities beyond investment casting, such as additive manufacturing or machining. Improved design producibility to reduce non-conformances can be implemented across manufacturing, sub-assembly, and inspection work centers. These principles can be useful across industries such as automotive manufacturing to optimize throughput and improve quality.

# A Techno Economic Assessment of Hybrid Renewable Energy and Battery Storage Systems for Data Centers

## BUSINESS PROBLEM

The rapid expansion of data centers is significantly increasing electricity demand, while companies are committing to 24/7 carbon-free energy use. However, current renewable energy procurement strategies—like PPAs and RECs—don't guarantee real-time carbon-free energy delivery. The challenge is to design an optimized energy system that minimizes costs and grid dependence while reliably providing renewable energy to meet the constant, high demand of data centers.

## APPROACH

I developed a techno-economic optimization model that simulates hourly energy generation and consumption from solar, wind, and battery storage systems co-located with a data center. The model incorporates system degradation, market pricing, and financial variables like CAPEX, OPEX, tax incentives, and storage repowering. I then ran simulations to identify optimal configurations that balance carbon

## DATA SOURCES

The project used real-world datasets including solar irradiance and wind speed from the National Renewable Energy Laboratory (NREL), energy price forecasts from the Cambium model, and infrastructure cost data from the NREL Annual Technology Baseline. A synthetic data center load profile was created to simulate energy usage.

### Data Types and Format

Data types included hourly time series data for solar, wind, load demand, and market prices—structured in CSV format. These were used to model the behavior of the energy system over a 20-year period.

1. **Define the Problem**
   Identify the challenge of powering data centers with 24/7 carbon-free electricity while maintaining cost-effectiveness and reliability.
2. **Select the Use Case**
   Focus on a hypothetical data center in Lubbock, Texas, due to its strong solar and wind resources and growing tech infrastructure.
3. **Collect Real-World Data**
   Gather hourly data for solar irradiance, wind speeds, energy prices, and a theoretical data center load profile.
4. **Build Simulation Models**
   Develop an energy simulation model to track hourly renewable generation, battery storage behavior, and grid interaction over 20 years.
5. **Integrate Financial Modeling**
   Layer in a financial model to calculate the Levelized Cost of Energy (LCOE) using CAPEX, OPEX, tax credits, and degradation rates.
6. **Run Optimization Scenarios**
   Simulate multiple system configurations (varying solar, wind, and battery sizes) to assess trade-offs between cost, carbon-free energy use, and grid reliance.
7. **Analyze Results**
   Identify optimal system designs that maximize carbon-free energy usage while remaining economically competitive.
8. **Draw Conclusions & Insights**
   Summarize key findings, including optimal configurations, cost thresholds, and recommendations for energy procurement strategies.

Author: Alex Sirgo

## IMPACT

The analysis demonstrates that a hybrid system of solar, wind, and battery storage can significantly reduce a data center's carbon footprint while remaining cost-competitive. It provides a decision-making framework for corporate sustainability and procurement teams, helping them optimize energy investments and move closer to true 24/7 carbon-free operation. These insights are also valuable to utilities and policymakers shaping future renewable infrastructure and regulation.

**DRIVERS**

Explosive data center growth, combined with aggressive corporate decarbonization targets (e.g. 24/7 carbon-free energy by 2030), created urgent demand for reliable, cost-effective renewable energy solutions tailored to large, constant-load facilities.

**BARRIERS**

Key barriers included the intermittency of renewables, battery degradation and cost, and the complexity of aligning real-time energy demand with renewable generation. Modeling long-term performance with limited historical data was also a challenge.

**ENABLERS**

Support from experts at NextEra, access to validated industry datasets, and strong cross-functional collaboration between business and engineering teams enabled the technical and financial depth of the project. MIT's LGO program also provided structure and guidance.

**ACTIONS**

I built a hybrid simulation and financial model that explored thousands of renewable energy system combinations. The model accounted for degradation, energy arbitrage, and hourly grid interactions to determine optimal configurations and costs.

**INNOVATION**

The integration of a co-located renewable system design—combining solar, wind, and battery storage with hourly simulation—enabled high-resolution analysis of 24/7 carbon-free strategies. The approach moves beyond traditional annual offsets or PPAs by modeling real-time matching.

**IMPROVEMENT**

The model identified configurations that could deliver up to 85-90% carbon-free energy usage with significantly reduced reliance on the grid, while remaining cost-competitive at under $45/MWh LCOE in high-resource regions like West Texas.

**BEST PRACTICES**

Use hourly time series modeling rather than annual averages to capture the real performance of intermittent resources. Include battery degradation, repowering, and market pricing in financial calculations to ensure realistic cost projections.

**OTHER APPLICATIONS**

This framework can be applied to any large, energy-intensive facility aiming for decarbonization—such as semiconductor fabs, hydrogen electrolysis plants, or AI training clusters—especially those considering co-located renewable infrastructure.

Strategy | Industrial automation | Performance management | Supervised Learning

# Data-Driven Key Performance Indicator Modeling for Robotic Mobile Fulfillment Systems

**Amazon**

## BUSINESS PROBLEM

E-commerce companies operating robotic mobile fulfillment systems (RMFS) face significant challenges in predicting and optimizing brownfield warehouse performance across complex networks. Traditional modeling methods excel in providing design stage insights but struggle to anticipate the dynamic nature of these systems, which constantly evolve due to rapid market changes, continuous experimentation, and shifting customer demands. Companies with large, digitized networks can benefit from data-driven approaches that leverage this operational data to provide accurate predictions and actionable insights for decision-making.

## APPROACH

Machine learning techniques were employed to develop predictive models for RMFS performance metrics. Gradient boosted tree ensembles were selected across model types. Template-specific models were created to address unique facility configurations. Interpretability techniques, including SHAP analysis and permutation feature importance, provided insights into system behavior and KPI drivers.


## DATA SOURCES

The research utilized extensive operational data from Amazon Robotics' Robotic Mobile Fulfillment Systems (RMFS) and was ultimately obfuscated in publishing. Original data included hourly time-series data, as well as site characteristics such as zone configurations and equipment details. This data spanned multiple years and sites across North America.

### Data Types and Format

Data utilized was tabular time-series sampled at various frequencies as well as unstructured semi-static site characteristics and classes.

Author: Steffan Sowards

## IMPACT

The developed predictive framework demonstrated the viability of this form of brownfield data-driven modeling in warehouse automation. Models achieved a mean $R^2$ value of 0.7838 across all templates and KPIs, with particularly strong performance in the most predictive metrics (mean $R^2$ of 0.9660). The framework was used to generate immediate practical artifacts such as lookup tables that visualize performance trade-offs between different parameters. These tables allow operators to quickly assess the impact of proposed changes and identify optimal parameter combinations that maintain desired performance levels. The model's scenario modeling capability also supported proactive planning by simulating "what-if" scenarios for inventory layouts, robot fleet management, and peak season preparation. This work also enhanced system understanding through interpretability techniques that identified key performance drivers, non-linear performance behavior, variable interactions, and system tipping points. This knowledge helps has been captured internally to help prioritize operational adjustments and system design modifications, focusing resources on factors with the greatest impact. By bridging the gap between basic analytics and complex simulation tools, this framework enables more efficient resource allocation, supports strategic planning, and facilitates rapid adaptation to changing market conditions.

| | |
|---|---|
| DRIVERS | The e-commerce industry's rapid growth and technological advancements enables this type of work. The catalysts for this solution were the need for rapid data-driven insights in RMFS operations, dynamic market conditions, complex system interactions, and the availability of rich operational data. |
| BARRIERS | Barriers included the scope and complexity of RMFS systems and the need to balance model interpretability with predictive accuracy. The rich availability of data required careful handling, engineering, and selection to maintain the utility and control of model inputs at the expense of perceived model performance. |
| ENABLERS | Amazon's massive active brownfield Robotic Mobile Fulfillment Systems network was central to this work. Key enablers were the collaboration with Amazon Robotics, access to extensive operational data, and the use of advanced machine learning techniques like CatBoost and SHAP analysis. |
| ACTIONS ➡ | Initial project actions included data collection and preprocessing, feature engineering, model development using AutoML and CatBoost, performance evaluation, and model interpretation using SHAP and permutation feature importance. Enduring impact was enabled through findings reports, look-up table generation, Sagemaker code documentation, and technical handoff to Amazon Robotics internal analytics and data-science personnel. |
| INNOVATION | Innovative aspects of this work included making use of large operational datasets in RMFS system modeling, the use of machine learning to replace traditional simulation methods, the development of a novel and generalizable feature engineering framework, the generation of practical operations artifacts, and the application of SHAP analysis for system performance behavior assessment. |
| IMPROVEMENT | The solution achieved a mean $R^2$ value of 0.7838 across all templates and KPIs, with particularly strong performance in the top metric (mean $R^2$ of 0.9660). This performance enables reliable forecasting and operational optimization in a format sensitive to non-linear system behavior without the time and cost associated with high fidelity simulation. |
| BEST PRACTICES | Best practices for replicating this solution include leveraging real-world operational data, assessing a broad range of advanced machine learning techniques, and focusing only on actionable inputs for model interpretability and utility. Ongoing use additionally required continued model performance monitoring, retraining, and validation against evolving system and market behavior. |
| OTHER APPLICATIONS | Potential applications extend beyond RMFS to other complex warehouse automation systems, such as Automated Storage and Retrieval Systems (AS/RS) and Autonomous Mobile Robots (AMRs). The approach could also be applied in other industries requiring predictive modeling of complex systems, like general manufacturing and logistics. |

Engineering | Industry components | Predictive forecasting | Regression analysis

# A Technoeconomic Model for Maritime Applications of Green Power Technologies

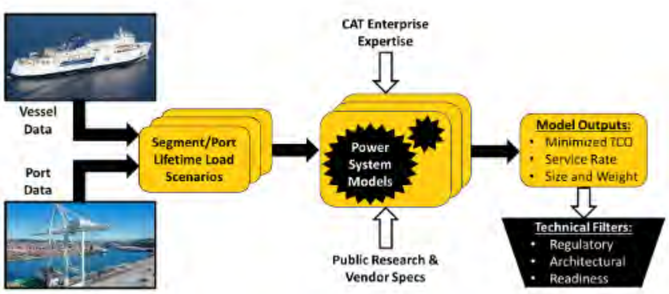
CATERPILLAR

## BUSINESS PROBLEM

With increasing demand for carbon-neutral and carbon-lite power generation both within the marine industry and Caterpillar Inc.'s (CAT) broader service areas, it is critical for CAT to understand how emerging technologies might affect their primary markets in the future. Given the variety of industries and diversity of customer needs that CAT serves, the development of tools that allow for a sector-by-sector approach to new technology investigation is critical to understand where and to what degree individual markets will be disrupted by the adoption of new technologies.

## APPROACH

Expand CAT's technoeconomic investigative capabilities into new power technologies by developing a flexible model to predict and compare the competitiveness (discounted total cost of ownership) of small and micro nuclear reactor (S/MMR), hydrogen fuel cell, and battery-electric power systems to incumbent fossil-fuel power sources for a variety of on-water and onshore marine applications.




## DATA SOURCES

The model was built from several sources of knowledge. For constructing load profiles and internal combustion engine (ICE) system models, the majority of the data came from CAT and other vendor documentation. For batteries and fuel cells, the data came from supplier data sheets, publications, and CAT experts. For nuclear systems, the data came from publications, textbooks, and academic experts.

### Data Types and Format

Data formats predominantly included time series for load profile construction, physical property data sets to build the technical components of the model, and cost data for the economic portion.


Author: Daniel Tuana

## IMPACT

The solution equips CAT with a configurable tool based on first principles, empirical correlations, and reasonable assumptions to assess and compare the performance of several alternative technologies of interest to incumbent fossil fuel power systems in the marine space. The tool predicts total cost of system ownership, lifetime carbon emissions, and component volume and mass based on segment-specific, characteristic input data. In scanning across a diverse set of use cases both within and adjacent to CAT's current marine market, the tool identified opportunities for the adoption of battery technologies in short-trip vessel segments where operations would not be disrupted by an increased need for charging. The tool also identified opportunities for the adoption of nuclear power technologies at large ports and on large cargo ships. However, these use cases are specific and fall outside of what can be considered the core components of CAT's marine business - the model predicts that it will be very difficult to economically decarbonize most of the vessel segments and port archetypes studied using solely the alternative technologies considered. Hybrid diesel-battery powertrains and drop in alternative fuels seem to show more potential in this vein. In addition to the high-level segment analysis conducted for the thesis, the tool may prove useful for individual vessel/port-specific modeling given the bespoke nature of marine power architecture, particularly for larger systems.

**DRIVERS**

CAT is concerned with informing their long-term technology strategy as markets shift towards incentivizing decarbonization due to numerous international and organizational carbon emission reduction mandates. CAT deals in markets that are reliant on fossil fuels and difficult to decarbonize (e.g., marine power). A structured yet flexible data-driven method for assessing the adoption of green power technologies could guide this strategy.

**BARRIERS**

Lack of extensive real-world vessel and port operating data in some of the segments that the analysis aimed to explore made it difficult to forecast the power demand requirements for these scenarios with a high level of certainty. Additionally, a lack of recent, public small-scale nuclear power projects made it difficult to forecast the cost of these systems precisely.

**ENABLERS**

The diversity of the team's knowledge base was a major factor in the success of the project. Specifically on the ICE and electrical system modeling front, having access to experts streamlined model development. The matrix structure of the company made it possible to reach across teams and leverage connections at partner companies and regulators to gain insights into the marine market which further informed the conclusions of the analysis

**ACTIONS**

Since the tool was designed to inform strategic decision-making, communicating both how the model works and the results of the analysis were crucial to adding value to the company. This communication included several presentations and a Q&A with CAT senior management and several hours of recorded training with CAT simulations team to ensure a complete knowledge transfer.

**INNOVATION**

Building on several theses that explore applications of nuclear power technologies for large cargo shipping, the solution developed takes a more ground-up approach to demand modeling, allowing for granular case-specific demand modeling rather than making high level assumptions about power usage. The model also includes an innovative simplified radiation shielding calculation method that is computationally inexpensive and scenario agnostic.

**IMPROVEMENT**

The solution provides CAT with the ability to leverage significantly more data to better inform their predictions on the economic competitiveness and technical fit of alternative power technologies in their primary markets. The configurability of the tool allows for the scanning of uses-cases across multiple segments of the marine industry to identify the aggregated addressable market and business opportunity for a given green power technology.

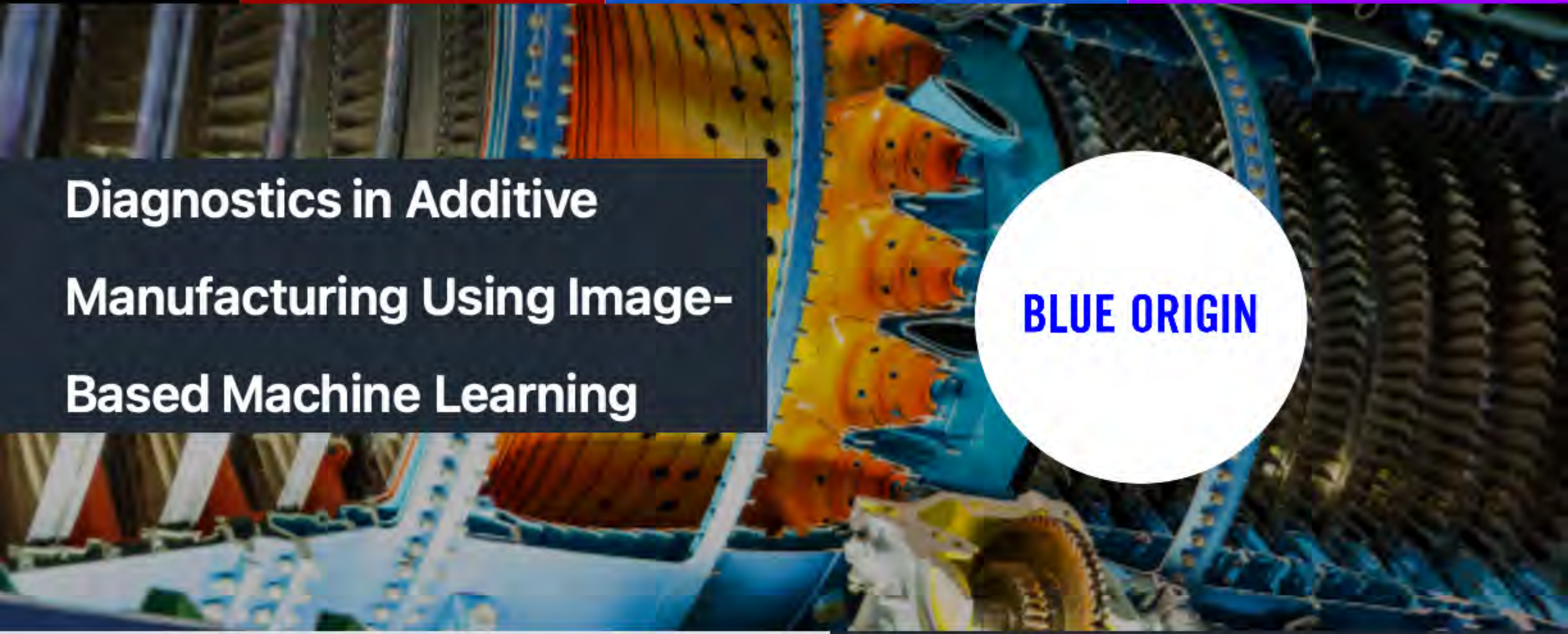
**BEST PRACTICES**

Documenting and substantiating the assumptions that contribute to the construction of the model is critically important to both increase confidence in the subsequent results and ensure that other members of the organization understand its function. In future, validating the model using available lifetime cost data for real-world power systems would develop further confidence in its forecasting abilities.

**OTHER APPLICATIONS**

The logic and architecture of the marine power system model could easily be adapted to compare the application of green power technologies in other areas of CATs business and expand to include a wider area of technologies.

Engineering | Aerospace and defense | Defect detection / product inspection | Convolutional neural networks



# Diagnostics in Additive Manufacturing Using Image-Based Machine Learning

BLUE ORIGIN

## BUSINESS PROBLEM

Blue Origin faces a critical bottleneck in Additive Manufacturing: non-destructive testing (NDT). Each 3D-printed engine part undergoes volumetric NDT inspection (CT or radiography), creating unsustainable constraints. Each scan can cost $5-10k, and, due to limited resources within the US, have a lead time of 30 days, creating production delays. More critically, there are not enough aerospace-grade CT machines worldwide to support projected throughput. Without process change, meeting future customer demand will become impossible. Blue Origin needs a solution that alleviates this bottleneck without sacrificing rigorous quality standards.

## APPROACH

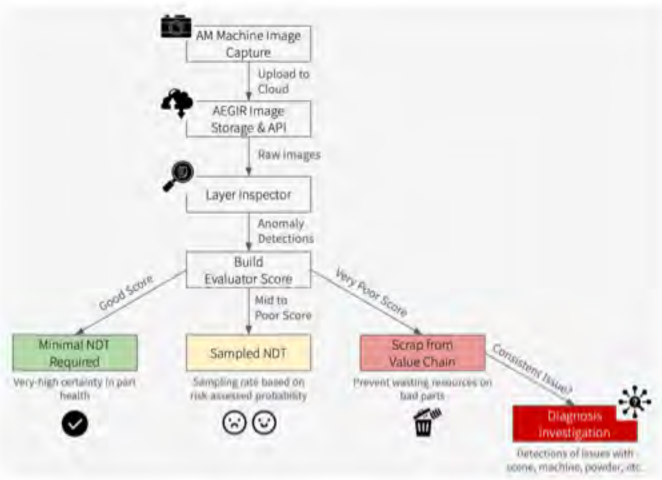
We developed a diagnostic system that uses machine learning to analyze layer-by-layer photographs taken during 3D printing. The system identifies anomalies in individual layers using U-Net neural networks, then evaluates overall build health through an ensemble of tree-based models. This generates a process-based evaluation that enables intelligent, risk-informed allocation of NDT resources.

## DATA SOURCES

Our primary data were layer-by-layer photographs of the 3D printing process. These came from AM machines that captured this "powder bed imagery." Our team also created the "Challenge Set," a dataset of 306 builds manually classified by subject matter experts. Each build contains between 1,000-20,000 layers of powder bed imagery, with classifications of "nominal" or "anomalous."

### Data Types and Format

The data consisted of grayscale image sequences (640x640 pixels) captured during printing – two images per layer: one after powder delivery ("Recoat") and another after laser welding ("Exposure").




Author: Arun Varma

## IMPACT

Our solution turns what began as an unsustainable bottleneck into an opportunity for enhanced quality control and long-term manufacturing strategy. First, it provides visibility into the build process. Unlike end-stage NDT inspections, our system assesses real-time process signatures, revealing when and how deviations occur. This enables better root cause analysis and process improvements across manufacturing operations. Second, it alleviates the NDT bottleneck. The system can confidently bypass NDT for low-risk builds while inspecting parts that most need it. Conservative implementations reduce NDT requirements by 37-54%, and far outperform alternative approaches in anomaly detection. These calculations represent a minimum expected benefit, as Blue Origin generates a much lower percentage of anomalous builds than in our test dataset. Third, it delivers financial benefits. While exact figures are proprietary, annual cost savings reach into the millions of dollars due to reduced NDT expenditures and prevention of wasted post-processing on defective parts. These reductions also translate to accelerated production cycle times and manufacturing agility. Beyond these immediate benefits, the system creates a foundation for future manufacturing intelligence – notably, early detection of process drift, and data-driven insights that transform routine quality control into a strategic advantage.

**DRIVERS**

Blue Origin's (and the aerospace industry's) growing demand for rocket engines created a need to dramatically scale AM production. The global limitation of high-energy CT machines created a barrier to this scaling by constraining Blue Origin's ability to run volumetric NDT.

**BARRIERS**

Aerospace regulations and standards (e.g. ITAR, AS9100) limit the possible solution space. Large volumes of image data create computational challenges for iterative development. Low-information photographs hinder classification. Ambiguity exists in visual anomaly detection. The consequences of aerospace component failure mean we must take extreme caution in balancing efficiency with quality.

**ENABLERS**

Our "NDT Peacemakers" tiger team combined talent from Manufacturing and Engines organizations, bringing together expertise in AM machines/data and in engine design. The prior development of the AEGIR tool provided critical infrastructure for processing high-quality powder bed imagery. Blue Origin's culture of balancing innovation with necessary precaution created space for exploring creative approaches.

**ACTIONS**

We developed a 2-step ML system: a U-Net to identify anomalies in individual layers, and a tree-based ensemble to assess build health. We engineered statistical features to capture physical anomaly patterns. We implemented a probability sampling approach that can scale NDT allocation according to risk level. We created a rigorous testing methodology using k-fold cross-validation to ensure results maintain detection sensitivity.

**INNOVATION**

Our solution transforms quality control from reactive inspection to proactive process monitoring. The dual-model approach (separate for Recoat and Exposure images) captures distinct failure signatures that might be missed by a single model. Our feature engineering process extracts physical meaning from visual data, creating explainability. The logistic sampling probability approach provides flexibility in balancing detection and efficiency.

**IMPROVEMENT**

Our most conservative implementation achieves 100% anomaly detection while reducing NDT requirements by 37%. More balanced parameters reduce NDT by up to 54% while maintaining 80% anomaly capture – far outperforming random sampling. With NDT cycle times spanning a full month, these reductions translate to millions of dollars in cost savings and accelerated production timelines.

**BEST PRACTICES**

Prioritize explainability in critical applications like aerospace, and design for asymmetric risk profiles where false negatives are far more consequential than false positives. Choose designs that physically align with expert knowledge. Involve subject matter experts throughout development. Mitigate confirmation bias when reclassifying data based on model outputs.

Engineering | Pharma / biopharma | Customer service / connectivity | Regression analysis

# Evaluating EM Drug Delivery Devices via OBI Market Landscape and Autoinjector Temperature Prediction

AMGEN®

## BUSINESS PROBLEM

The main problem the business is trying to solve is improving the patient experience with on-body injectors (OBIs) and autoinjectors by addressing the limitations of current drug delivery devices. This includes enhancing comfort, usability, and safety for patients, particularly those with chronic illnesses requiring regular medication. The focus is on developing innovative features like temperature sensing to ensure optimal drug efficacy and patient comfort, while also aligning product strategies with market trends and patient needs to maintain a competitive edge in the biopharmaceutical industry.

## APPROACH

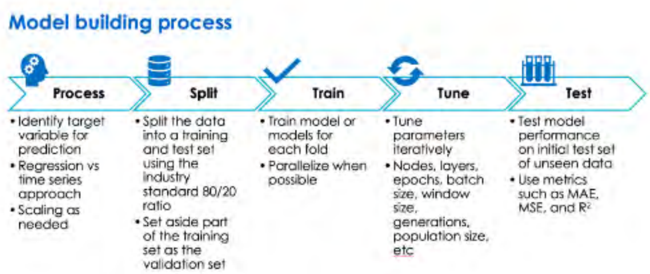
I investigated the market dynamics of OBIs and autoinjectors, identifying key patient needs, and developing innovative features like temperature sensing. This involved modifying existing devices, gathering and analyzing temperature data, and implementing predictive algorithms to enhance drug delivery accuracy and patient comfort. Strategic initiatives and partnerships were also explored.

## DATA SOURCES

External market reports and conversations with internal and external SMEs made up the bulk of the market landscape data. For the temperature prediction portion, I mainly collected my own data.

### Data Types and Format

PDF market reports as well as timestamped temperature probe data




Author: Aaron Wubshet

## IMPACT

The solution provides a proof of concept for a tangible way to significantly enahce patient satisfaction and adherence through the introduction temperature sensing in drug delivery devices. This improvement directly addresses the discomfort and usability issues patients face, particularly those with chronic illnesses requiring regular medication. By ensuring optimal drug efficacy and patient comfort, the solution fosters better health outcomes and increases patient trust in Amgen's products. Strategically, the solution aligns Amgen's product offerings with evolving market trends and patient needs, positioning the company as a leader in the biopharmaceutical industry. The development of temperature sensing capabilities not only improves the existing AutoTouch platform but also sets the stage for future advancements in on-body injectors (OBIs) and other drug delivery devices. This innovation enhances Amgen's competitive edge, driving market growth and expanding the company's footprint in the drug delivery market. Overall, the solution boosts Amgen's revenue and market share by delivering superior patient-centric products, fostering strategic partnerships, and maintaining a competitive edge in the biopharmaceutical industry. This comprehensive approach ensures that Amgen remains at the forefront of innovation, meeting the needs of patients and healthcare providers while driving business growth and success.

| | |
|---|---|
| **DRIVERS** | The catalyst for the solution was the increasing prevalence of chronic illnesses and the growing demand for self-administered treatments. Patients need more comfortable, user-friendly, and effective drug delivery devices. The biopharmaceutical industry's focus on patient-centric design and technological innovation, along with Amgen's commitment to improving patient outcomes and maintaining a competitive edge, drove the development |
| **BARRIERS** | Some barriers that impacted the project included the computational limitations of the existing device platform, such as limited memory, processing power, and battery life. Additionally, gathering sufficient and accurate temperature data for model training was challenging. Ensuring the predictive models were both accurate and efficient enough to run on the device without significant modifications also posed a significant challenge. |
| **ENABLERS** | Several features of the company and team helped enable the project. Strong R&D Focus: Amgen's commitment to research and development provided essential resources. Cross-Functional Collaboration: Effective teamwork across departments ensured a comprehensive approach. Patient-Centric Design: Focus on improving patient experience drove innovation. |
| **ACTIONS** | I gathered stakeholders from human factors, engineering, and legal to begin developing patent protections, testing the UI/UX with patients, and narrowing down on the appropriate notification to provide to patients when the drug had reached temperature. |
| **INNOVATION** | The use of temperature sensors that are designed to measure if the device has overheated instead as sensors for the actual drug temperature is the most novel aspect of the solution. |
| **IMPROVEMENT** | Improved patient experience in the comfort they experience during the injection. |
| **BEST PRACTICES** | Gathering as much data at the highest possible granularity as early as possible. Ultimately, solutions like this that rely on the data gathered can often be limited by the quality and quantity of data gathered. |
| **OTHER APPLICATIONS** | The same solution could be ported to on body injectors not just auto injectors. |

# Metal Additive Manufacturing Capabilities for Footwear Prototyping and Product Creation

## BUSINESS PROBLEM

The business seeks to enhance the efficiency of its footwear product creation timeline, as accelerating the product creation process improves the alignment of product supply with forecasted consumer demand. Specifically, the company aims to integrate Metal Additive Manufacturing (MAM) into its onsite prototyping facility to expedite the creation of physical footwear samples from a proposed product brief. By incorporating MAM, the business can streamline the prototyping process, enabling more effective design iteration and bolstering its market responsiveness and competitive edge.

## APPROACH

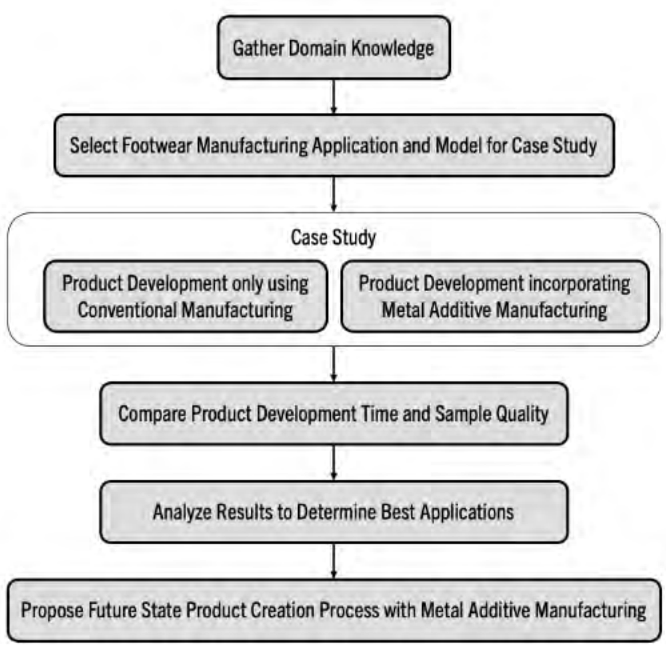
A case study approach on rubber outsole development was used to address the business problem. The study examined the impact of MAM on the speed and quality of footwear sample development. By comparing conventional manufacturing and MAM processes, the study evaluated time reduction and sample quality, focusing on rubber outsole mold fabrication and incorporating textures directly into the molds.

## DATA SOURCES

Information about MAM was obtained from literature review, exploring existing internal knowledge capture, and educational materials from the vendor. Information about footwear specific applications of MAM also came from the company's manufacturing partners. Three-dimensional part files, engineering drawings, and visual inspection standards were obtained from company and manufacturing partner datab

**Data Types and Format**

Process standard operating procedures and parameters, machine manuals, three-dimensional files, two-dimensional drawings, photographs, traction testing results, development time study, photographs




Author: Tiffany Xi

## IMPACT

Incorporating MAM into the product creation process was found to have an impact on the footwear product creation timeline. By reducing the time required for mold design and fabrication by 32%, the solution accelerates the overall product development timeline, allowing the company to bring innovative footwear designs to market more rapidly. This speed advantage is particularly valuable for projects involving textured outsoles, as MAM enables the direct fabrication of detailed textures, eliminating the need for outsourced etching processes. Although the visual quality of samples produced using MAM is not production-equivalent, it is adequate for product development purposes. The traction properties of the MAM-produced samples are comparable to those of the manufacturing partner's samples, indicating that the functional quality is sufficient for wear-testing. The integration of MAM capabilities enhances the business's prototyping facility's ability to serve as an accelerator for high-complexity or novel products, enabling quicker design revision cycles and more efficient product development. This improvement in speed and efficiency helps the company better align supply with consumer demand, reducing the risk of mismatched supply and demand due to changing consumer preferences over time. The solution contributes to the company's competitive edge in the fast-paced footwear industry by enabling faster market responsiveness and innovation.

### DRIVERS

The catalyst for the solution was the need to accelerate the product creation process in the fast-paced athletic footwear industry. The company faces challenges with long timelines for developing novel products, especially with overseas manufacturing partners. Incorporating MAM aimed to reduce these timelines, improve efficiency, and better align supply with consumer demand.

### BARRIERS

Communication challenges arose from navigating a matrix organizational structure alongside organizational changes and requesting information from overseas manufacturing partners. Technical issues, such as corrupted texture files and difficulties with mold alignment and part removal, also posed challenges. Additionally, resource limitations, including the availability of raw materials and suitable software, affected the project's progress.

### ENABLERS

Close relationships with manufacturing partners facilitated knowledge sharing and collaboration. I had total support of the department director and all the engineering managers, and the collaborative nature of the daily work helped remove barriers and prioritize the project. A departmental and company-wide culture in investment of new technologies and innovations, such as the MAM machine, boosted support for the project.

### ACTIONS

I explored the possible applications and context around the project. I conducted a case study that involved developing and making physical parts to obtain learnings and propose a solution. The current and proposed manufacturing processes were mapped and designed. The points of assessment were chosen for the case study and the results communicated to the organization to guide implementation of the technology.

### INNOVATION

Some innovative aspects of the solution include the use of MAM to directly fabricate detailed textures into molds, reducing prototyping development time and eliminating the need for etching, and the hybrid and modular design approach of combining subtractive and additive manufacturing techniques for footwear prototyping molds.

### IMPROVEMENT

The final improvement provided by the solution is a potential 32% reduction in the time required for mold design and fabrication by incorporating MAM. This speed advantage is primarily due to the ability to directly fabricate detailed textures into the mold, eliminating the need for outsourced etching and streamlining the overall product creation process.

### BEST PRACTICES

To replicate the solution, ensure close collaboration with manufacturing partners as well as other internal departments for knowledge sharing. Spend time understanding any evolutions in the needs and goals of the organization to deliver value-adding solutions. Ensure proper understanding of the work of each individual contributor to get a comprehensive understanding of potential improvement points.

### OTHER APPLICATIONS

Other potential applications of the solution include creating textured molds for pressurized bladder cushioning units, foam midsole molds, and plastic injection molds for cleated footwear. The approach used could also be used by organizations in other industries that aim to understand how MAM can affect their prototyping and manufacturing processes.

# Driving Manufacturing Best Practices with Multimodal AI

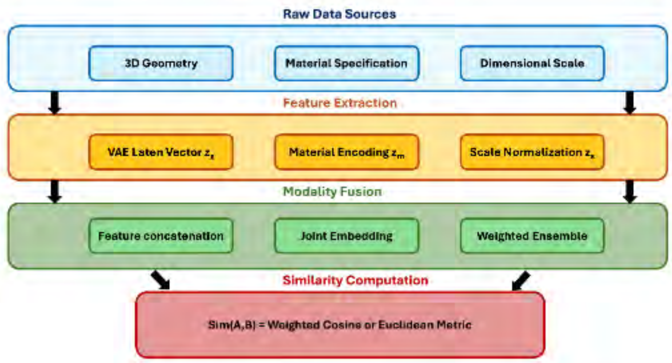
LFMcapital

## BUSINESS PROBLEM

The main problem is inconsistent quoting and engineering processes in contract manufacturing. Quoting relies heavily on individual estimator experience, causing significant variations that affect profitability and customer satisfaction. Engineers plan manufacturing processes from scratch for new projects despite having produced similar components previously, creating bottlenecks and preventing standardization. This is especially challenging in aerospace manufacturing where material and scale significantly impact production approaches beyond geometric similarity.

## APPROACH

A multimodal AI system that analyzes multiple aspects of manufacturing parts was developed. The system examines 3D models alongside material and scale information to find historically similar components. This approach helps estimators make more consistent quotes based on past data and allows engineers to reuse existing manufacturing strategies instead of starting from scratch with each new part.




## DATA SOURCES

The solution leverages three primary data sources from the company's systems: 3D CAD models stored on company drives, material specifications from the enterprise resource planning (ERP) database, and dimensional information extracted from technical drawings. Historical cost and production time data was also utilized from completed work orders.

### Data Types and Format

3D geometric data (CAD models), categorical data (material specifications), numerical data (dimensional measurements), and structured tabular data (historical cost & production records)

Author: Mark Zachary

## IMPACT

The part similarity tool delivers significant operational and financial benefits across multiple facets of the contract manufacturing business. In the quoting process, the solution reduces subjectivity and variability by enabling estimators to base new quotes on historically similar parts with known costs and production times. This data-driven approach improves quote accuracy, preventing both underpricing that erodes margins and overpricing that loses business opportunities. The consistency in quoting also enhances customer confidence and improves win rates on new bids. For engineering operations, process planning team can now quickly identify similar parts with proven manufacturing approaches, allowing them to adapt existing methodologies rather than creating them from scratch. This drastically reduces manufacturing process time for new components and standardizes manufacturing methods across similar part families, leading to shorter lead times and more predictable production schedules. The system also supports knowledge retention and transfer within the organization. As experienced personnel retire or leave, their tribal knowledge about manufacturing approaches becomes embedded in the system through historical part relationships, making this expertise accessible to newer staff members. This institutional memory helps maintain consistent manufacturing quality and efficiency despite workforce changes.

| | |
|---|---|
| **DRIVERS** | Industry inefficiencies in aerospace and defense contract manufacturing was the catalyst behind the part similarity tool enabled by multimodal AI approaches. Fragmented quoting, inconsistent engineering processes, and siloed expertise led to high costs and delays, highlighting the need for a data-driven, AI approach to standardize workflows and improve overall operational efficiency. |
| **BARRIERS** | Barriers included fragmented data systems, inconsistent documentation, and limited data quality. Legacy infrastructure and high process variability compounded these issues, while heavy reliance on tacit expertise further hindered data standardization and integration, challenging the deployment of a robust multimodal AI solution. |
| **ENABLERS** | Supportive leadership, strong cross-functional collaboration, and a culture of innovation were key enablers. The team's deep technical expertise and commitment to continuous digital improvement provided the foundation to integrate multimodal AI into manufacturing processes, driving operational excellence and digital transformation. |
| **ACTIONS** | We built a part similarity tool by training a VAE to encode 3D geometries and integrating material and scale data. We then implemented an ANN indexing system for rapid similarity retrieval, piloted the solution at Company X, and refined it iteratively using feedback from quoting and engineering teams. |
| **INNOVATION** | Innovative aspects include multimodal integration of 3D geometry, material properties, and scale data using VAEs to create a unified latent space. This approach enables rapid similarity retrieval through ANN indexing, streamlines quoting and engineering workflows, and leverages human feedback for continuous improvement. |
| **IMPROVEMENT** | Our solution improves quoting accuracy, reduces manufacturing cycle times, and shortens lead times. By standardizing workflows and leveraging multimodal AI, it delivers quantifiable cost savings and boosts competitive positioning through data-driven, consistent decision-making. |
| **BEST PRACTICES** | Standardize and pre-process all data sources to ensure high-quality input. Use robust multimodal models (e.g., VAEs) with efficient ANN indexing for similarity search. Engage cross-functional teams early, pilot in real-world settings, iteratively refine based on feedback, and document each step for replicability. |
| **OTHER APPLICATIONS** | Our approach extends to quality control and predictive maintenance, streamlining defect detection and process optimization. It also supports design reuse in rapid prototyping, digital twin creation, and cost estimation across industries like automotive, medical devices, and consumer electronics, where integrated multimodal insights drive efficiency. |