

ABSTRACT

In the window manufacturing value stream, glass is shipped from different suppliers, and various processes occur at the manufacturing facility to ship the windows to distributors. Each manufacturing process parameter might lead to the potential risk for window breakage which leads to increased costs. Our objective in this study is to **optimize process settings that decrease window breakage**. During the manufacturing process, we collect data from each step of the glass making process and explore features that might impact glass quality. Using linear regression analysis, logistic regression analysis, and optimization we provide an approach to predict the window breakage in stage I and then optimize the manufacturing process settings in stage II to improve the process and save the firm money with decreased breakage costs.

INTRODUCTION

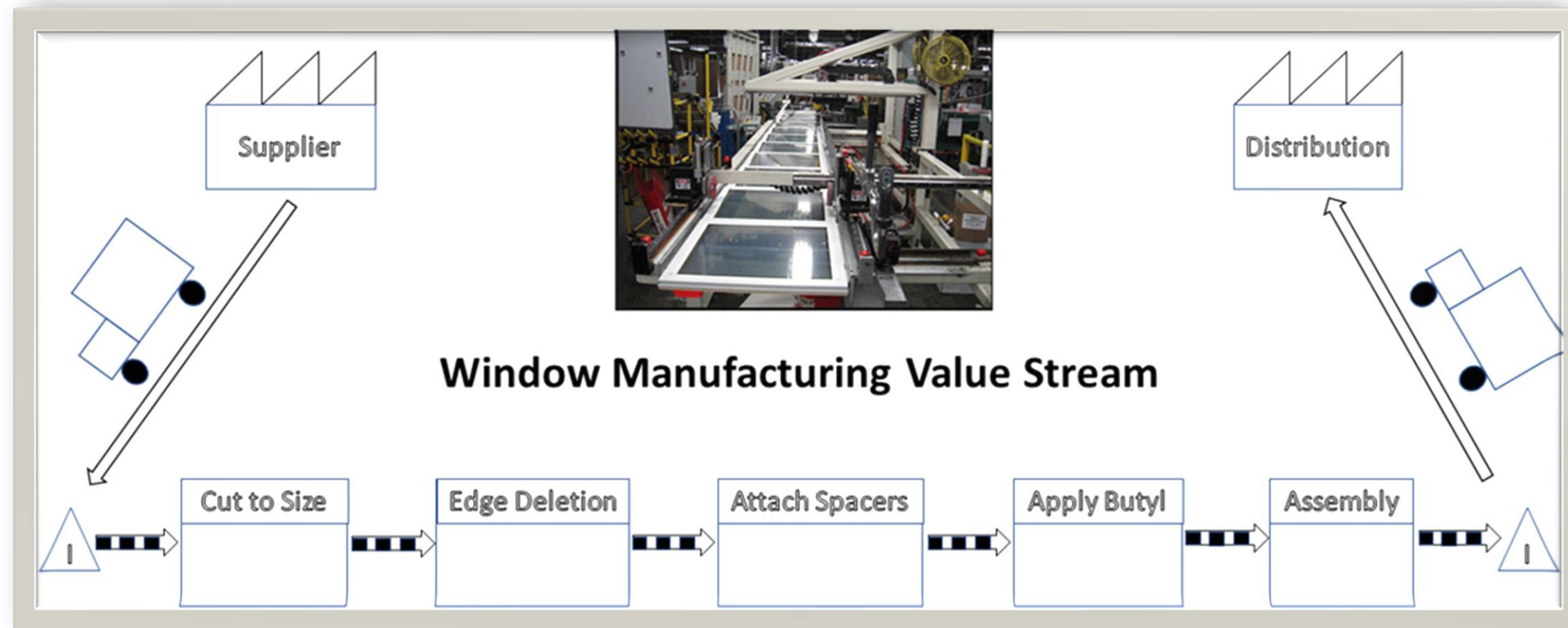


Fig 1. Process Variables in Value Stream (source: SAS JMP Pro)

A glass manufacturing company has an average opportunity cost of \$100 for each broken window order with 1m² of size. If the glass breaks easily, we will have to pay out more refunds. **Currently, 1 out of 16 windows break in the manufacturing process.** If we have 100 refunds per month, this will decrease our profit by \$10,000! Thus, to maximize our profit we are challenged to develop an analytics solution to set process settings for the manufacturer to decrease glass breakage.

Input Feature	Manufacture information
Glass Thickness	0.005+/-0.018 (supplier)
Glass Supplier	(supplier)
Window Color	65+/- 30 (supplier)
Cut Speed	(process)
Edge Deletion Rate	(process)
Spacer Distance	(process)
Window Size	(customer spec)
Window Type	(customer spec)

Table 1. Manufacturing Parameters

RESEARCH OBJECTIVE

What analytical approach would allow a manufacturer to optimally set window process settings to decrease breakage?

LITERATURE REVIEW

Year	Author	Study
2021	Adam N. Elmachtoub	Smart "Predict, then Optimize"
2020	Dimitris Bertsimas, Nathan Kallus	Management Science From Predictive to Prescriptive Analytics
2019	Bryan Wilder, Bistra Dikina, Milind Tambe	Melding the Data-Decisions Pipeline: Decision-Focused Learning for Combinatorial Optimization

Table 2. Summary of Academic Literature

METHODOLOGY

We identified a design that allowed us to estimate the relationship of process settings on breakage rate using parametric models (linear & logistic regression). Then we formulated an optimization model where the predictive model served as the objective function we were trying to minimize as shown in Fig 2.

Stage I: Estimate/Predict

Approach 1: Regression Formulation

Exploratory Data Analysis

- Data Cleaning & Reduction
- Standardization
- Box-cox Transformation

Linear Regression Model Building

- Weighted Least Squares Regression
- 5-Fold Cross Validation

Mathematical Theory

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 + \dots + \hat{\beta}_k$$

Approach 2: Classification Formulation

Exploratory Data Analysis

- Data Cleaning & Reduction
- Categorical Feature Transformation

Linear Regression Model Building

- Backward Selection (AIC)
- Down-sampling & Up-sampling
- SMOTE & ROSE

Mathematical Theory

$$\Pr(\hat{y}) = \frac{e^{\beta_0 + \beta_1 + \dots + \beta_k}}{1 + e^{\beta_0 + \beta_1 + \dots + \beta_k}}$$

Stage II: Optimize

Linear Regression Optimization Model

Objective

- Minimize the risk of window breakage rate with certain specification

Model Without Constraints

- Have negative input features which is not realistic

Summary Statistic Constraints

- Using standardized 1st and 3rd quartile as constrain to avoid outliers

Constraints with Interaction terms

- Achieved minimum window breakage rate

Mathematical Theory

$$\min(\hat{y}) = \hat{\beta}_0 + \hat{\beta}_1 + \dots + \hat{\beta}_k$$

Logistic Regression Optimization Model

Objective

- Maximize the non-breakage rate of window with certain specification

Model Without Constraints

- Have negative input features which is not realistic

Summary Statistic Constraints

- Using standardized 1st and 3rd quartile as constraints to avoid outliers

Constraints with Interaction terms

- Achieved maximum window non-breakage rate

Mathematical Theory

$$\max(\hat{y}) = \hat{\beta}_0 + \hat{\beta}_1 + \dots + \hat{\beta}_k$$

Fig 2. Methodological Design

STATISTICAL RESULTS

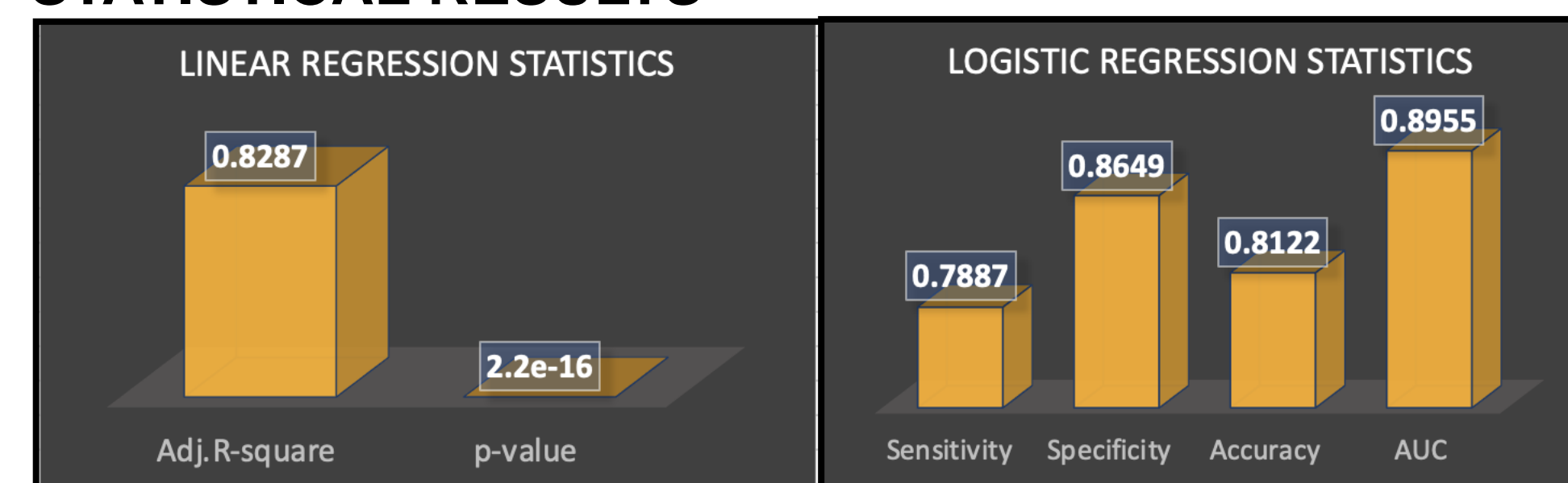


Fig 3. Model Performance Statistic For Linear and Logistic Regression

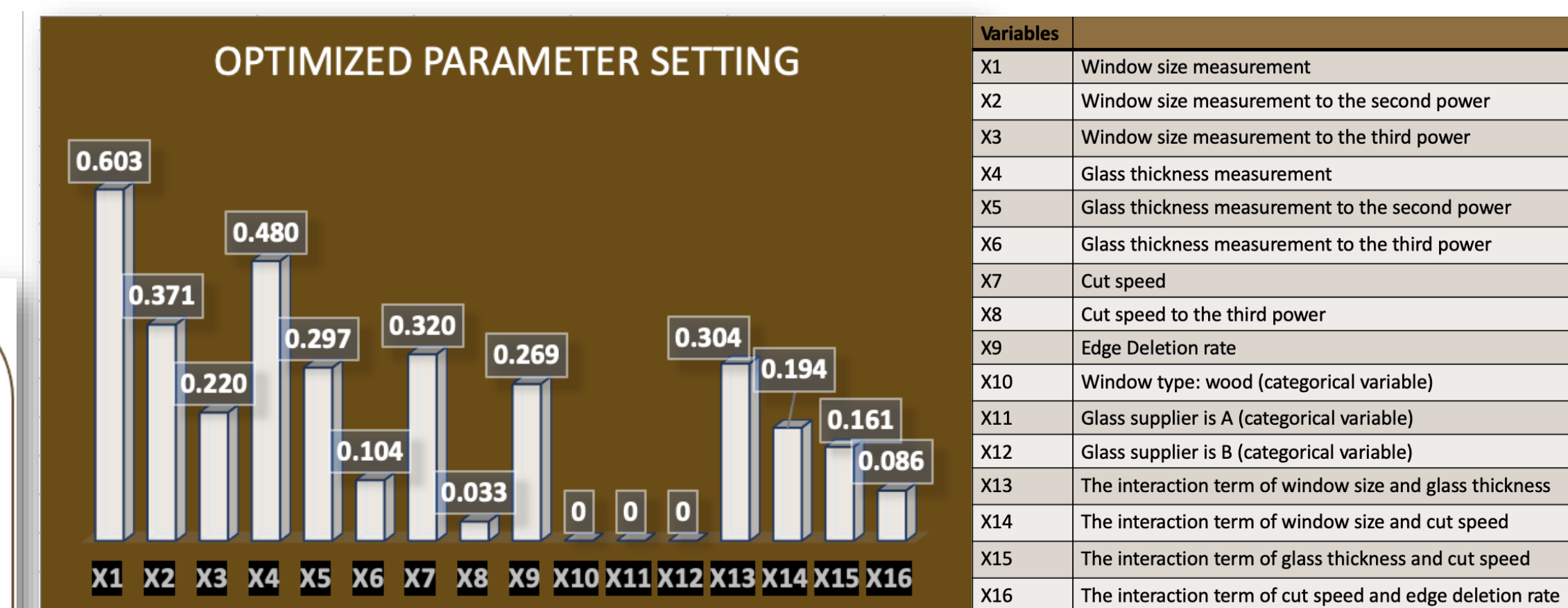


Fig 4. Optimized Parameter Settings

Expected Business Implications

If the technician uses the linear regression optimization formulation, the breakage rate will get as close to zero as possible (less than 0.1%). This means now only 1 out of 1,000 windows we make would be expected to break in the manufacturing process. Thus, we could save the cost of 63 broken windows in a batch of 1,000, which leads to a **+6% improvement or \$6,300 cost savings**.

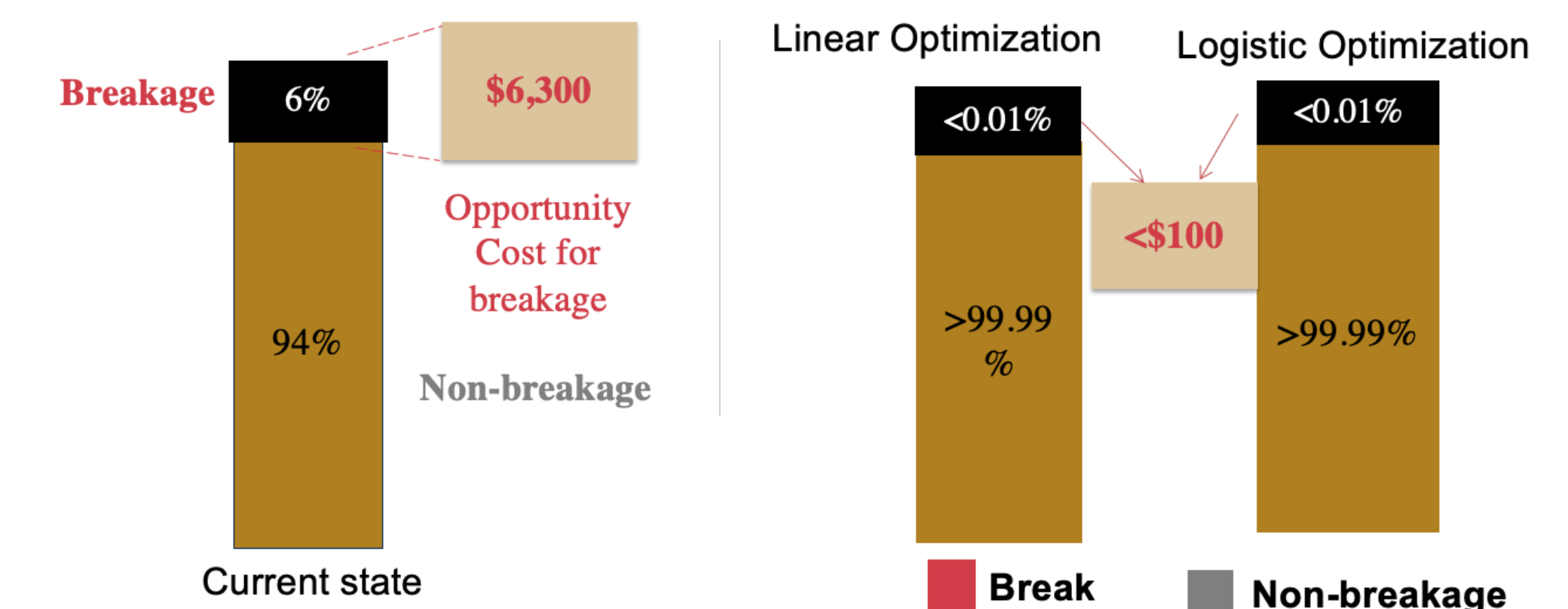


Fig 5. Process performance improvement

CONCLUSIONS

By interfacing predictive and prescriptive analytics in a two-stage fashion, we realized our lowest window breakage rate was close to 0. We concluded from the latest published research in *Management Science* and supported by our experiments that framing this particular problem as a linear regression model (e.g. response variable is a window breakage rate per batch rather than binary failed/good batch) first, then optimizing, was the ideal approach to support the technician and led to better process setting decision-making. The reason is the parameters would make more sense to the technician compared to odds-ratios in the classification case and would also be easier to update in Microsoft Excel and integrate into a solver.

FUNDING & ACKNOWLEDGEMENTS

We would like to thank the Summer Stay Scholarship program (<https://www.purdue.edu/summerstay/>) for funding this research, the Krenicki Center for Business Analytics & Machine Learning (<https://krannert.purdue.edu/centers/krenicki-center/>) for providing the research scope, and Frontline Solvers (www.solver.com) for providing a license to their user-friendly software that easily allows integration of prediction and prescriptive analytics.