

Factors predictive of pH in ABC Beverage products

[Mauricio Claudio](#), Chief Data Scientist

21 April 2023

This report presents the findings of an analysis to determine the factors predictive of pH in ABC Beverage products. The analysis applies five predictive Machine Learning models to beverage manufacture process data for each of the four ABC Beverage products – Brand A, B, C and D – and selects one model based on predictive performance on resampled data. Twenty models were tested in total, and their error performance measured with the Root Mean Square Error (RMSE) metric. The results of the analysis are summarized below:

- The data has a substantial number of missing entries, pointing to possible gaps in data collection and management, and/or sensor malfunction and need for maintenance.
- Only **two (2)** models -- the Random Forest and a Gradient Boosted Machines model for brand B -- achieved the minimum company R-squared threshold of 0.70 for selection.

The R-squared and resampled RMSE values for the two models is as follows:

- **Random Forest, brand B**
 - R-squared: 0.75, RMSE: 0.085
- **Gradient Boosted Machines, brand B**
 - R-squared: 0.71, RMSE: 0.092
- Manufacturing process factors most predictive of pH for the Random Forest model are, in descending order of importance, are: *Mnf.Flow, Bowl.Setpoint, Air.Pressurer*

Full factor importance for the Random Forest model is provided in the following page.

A limitation of the analysis is its estimates are based on resampled, cross-validated data. A next step is to test the two selected models for brand B with unseen test data. Another task is to seek different models that could meet the minimum R-squared threshold for brands A, C and D.

Predictions for brand B are available for inspection and download. No predictions for brands A, C and D were performed. Full details on the methodology of the analysis are found in the Technical Report.

Manufacturing process factor importance on product pH

Random Forest, brand B: predictor importance

