How to Use the Golden Batch Model, XAI, Causal Analysis for Root Cause Analysis in Industrial Batch Performance Analytics

09 • 25 • 2023

How to Use RCA - HMI

- **Step 1**: Search for knowledge-based questions to explore data insights that can potentially uncover the root causes.
  - Sample Questions:
    - Provide A scatter plot showing the relationship between the dissolved oxygen concentration and the aeration rate
    - Can you confirm the aeration in the sparge is in range? For all batches
    - Please provide A scatter plot showing the relationship between the dissolved oxygen concentration and the aeration rate
    - What was the min and max penicillin concentration value for batch 3
    - Give the O2 uptake rate? For all batches?
- **Step 2:** Find business insight from the Generative AI leveraged charts and answers
- **Step 3:** Store the top 5 historical question – response pairs for future retrieval

Summary:

The Batch Data Analyzer is designed for a rapid assessment of anomalies and the extraction of insights within batch data, offering a significant benefit by harnessing domain knowledge to uncover deep data insights. This is achieved through the use of Generative AI, which facilitates the identification and understanding of complex data patterns, aiding stakeholders in making informed decisions.

Rockwell Automation **Industry Solutions**
Batch Performance Analytics

- Batch Data Analyzer
- Event Analysis
- Causal Discovery

Batch Data Analyzer

Give the O2 uptake rate? For all batches?

Best Answer

The Oxygen Uptake Rate (OUR) for all batches has been visualized in a line plot. The plot shows the changes in OUR over time for each batch. This can help in understanding the variations in OUR across different batches and at different times.

Best Answer

The Oxygen Uptake Rate (OUR) for all batches has been visualized in a line plot. The plot shows the changes in OUR over time for each batch. This can help in understanding the variations in OUR across different batches and at different times.

Oxygen Uptake Rate over Time for All Batches

How to Use RCA Golden Batch Model for Root Cause Analysis

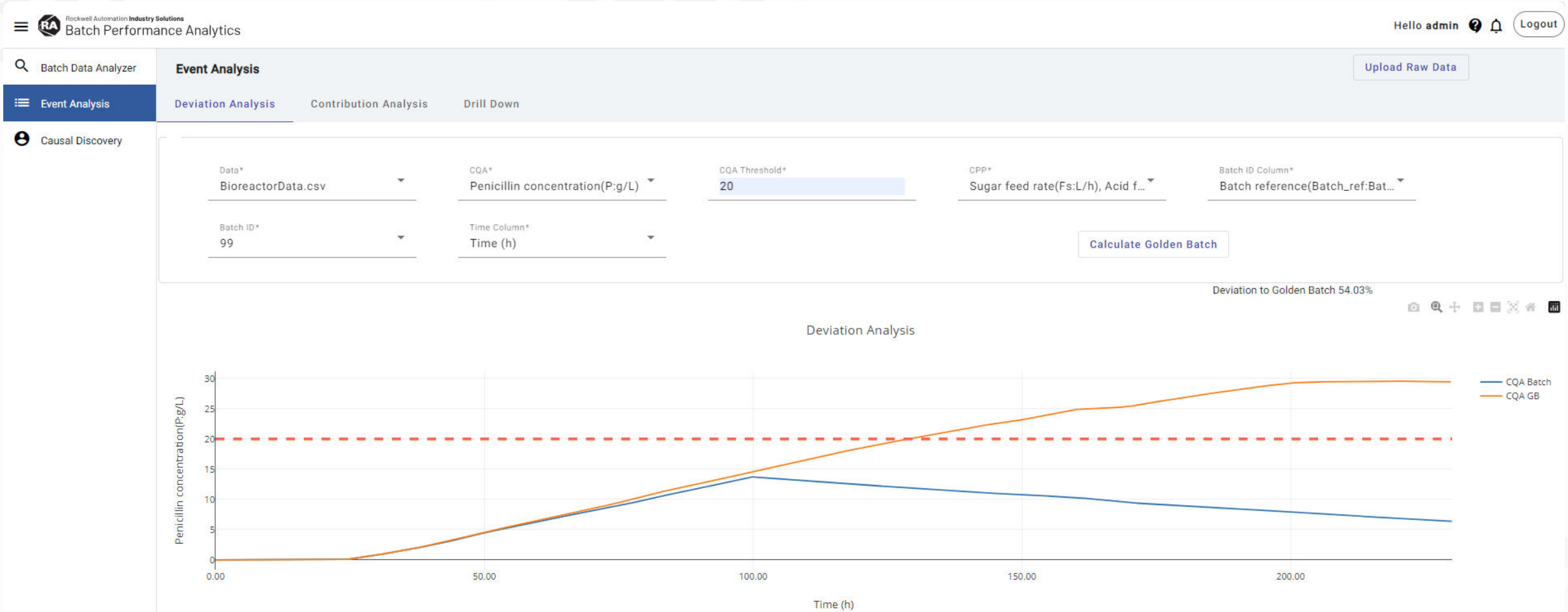
- **Step 1**: Go to the **Event Analysis tab** and select the Deviation Analysis view. Then, choose the data and select the CQA (critical quality attribute). Next, choose the CQA threshold which defines a batch as either good (CQA >= CQA threshold) or bad (CQA < CQA threshold). Select CPP (see CPP selection in slide 4). Pick the batch ID column, choose the batch ID (for example, 99), and choose the time column. Finally, click "**Calculate golden batch**" (Slide 3).
  - This will compute the golden batch, lower bound, and upper bound based on all the good batches identified by the CQA threshold.
  - Examine the chart. You will observe that **at time = 100 hours, batch 99 exhibits a deviation from the golden batch**, and the CQA value is less than 20 (the threshold). Therefore, **batch 99 is deemed a bad batch**.

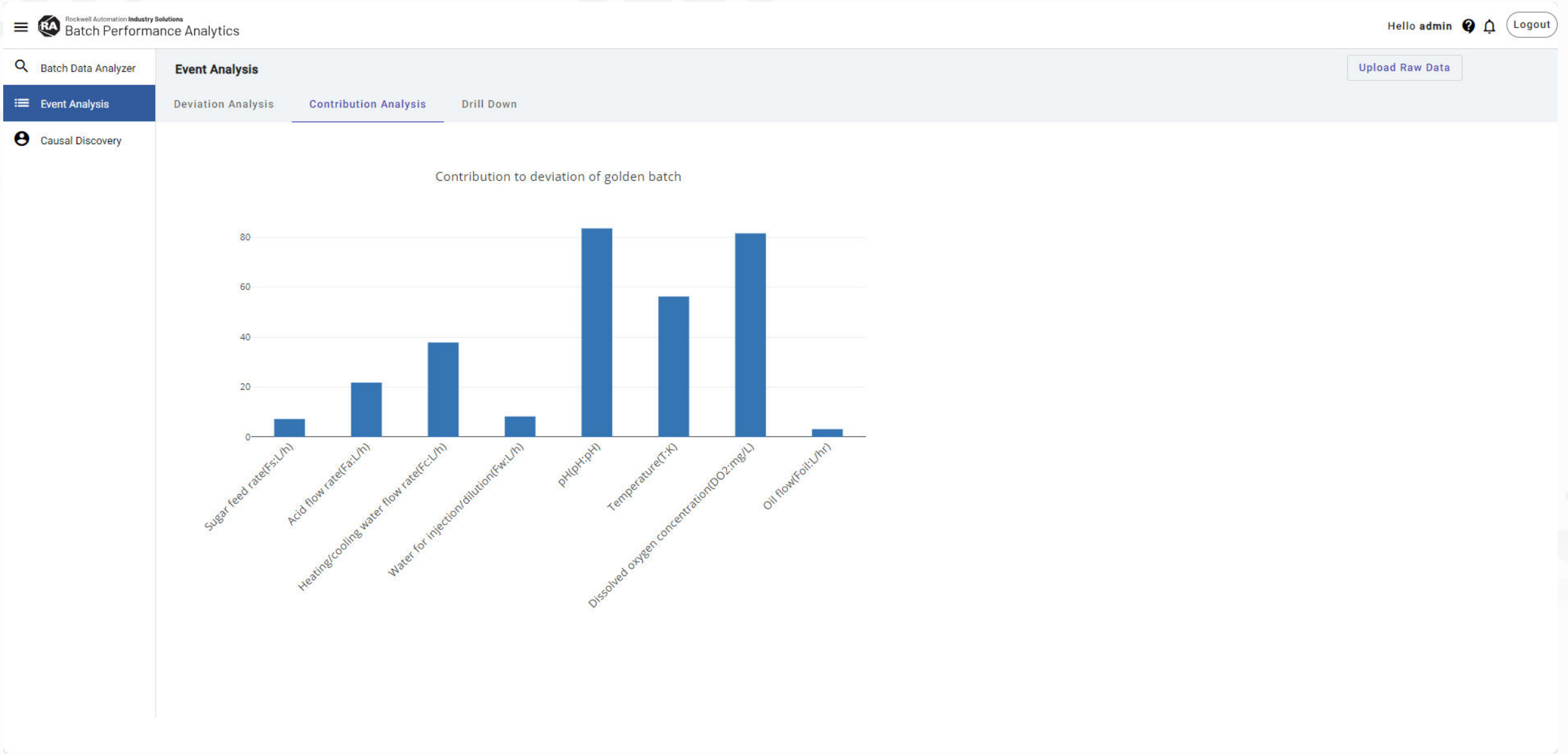
- **Step 2:** Navigate to the **Contribution Analysis** section (slide 4). Here, you will discover that **pH**, **Temperature**, and **DO2** are the top three CPPs with deviations from their golden batches. Click on pH, and it will take you to the Drill Down view.

- **Step 3:** In the **Drill Down** view:
  - The **pH** chart (Slide 5) **displays a spike at around 100 hours**. This is one of the causes of the CQA being out of specification (OOS).
  - Select **temperature** in the CPP file (Slide 6). You will notice a significant variation above the upper bound for an extended period. This is the reason why the CQA value is much lower than the CQA threshold of 20 at the end of the batch.
  - Choose **DO2** in the CPP field (Slide 7). From the chart, you can see that the DO2 batch exhibits increased variation after time = 100 hours. Therefore, DO2 is also a contributing factor to the OOS problem.

- **In summary**, by employing the golden batch model, you can determine if a batch is good or bad based on the CQA threshold, and compare it with the golden batch. If a bad batch (OOS) is identified, you can utilize contribution analysis to identify the top contributing CPPs. Following this, you can conduct a detailed investigation to ascertain **when, where, and how the root cause occurred**, and how these CPP factors influenced the CQA (OOS).

CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.


Rockwell Automation Industry Solutions
Batch Performance Analytics
Hello admin
Logout
Batch Data Analyzer
Event Analysis
Causal Discovery
Event Analysis
Upload Raw Data
Deviation Analysis
Contribution Analysis
Drill Down
Data*
BioreactorData.csv
CQA*
Penicillin concentration(P:g/L)
CQA Threshold*
20
CPP*
Sugar feed rate(Fs:L/h), Acid f...
Batch ID Column*
Batch reference(Batch_ref:Bat...
Batch ID*
99
Time Column*
Time (h)
Calculate Golden Batch
Deviation to Golden Batch 54.03%
Deviation Analysis
Penicillin concentration(P:g/L)
30
25
20
15
10
5
0
0.00
50.00
100.00
150.00
200.00
Time (h)
CQA Batch
CQA GB



CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.

Rockwell Automation Industry Solutions
Batch Performance Analytics

Hello admin | Logout

- Batch Data Analyzer
- Event Analysis
- Causal Discovery

Event Analysis

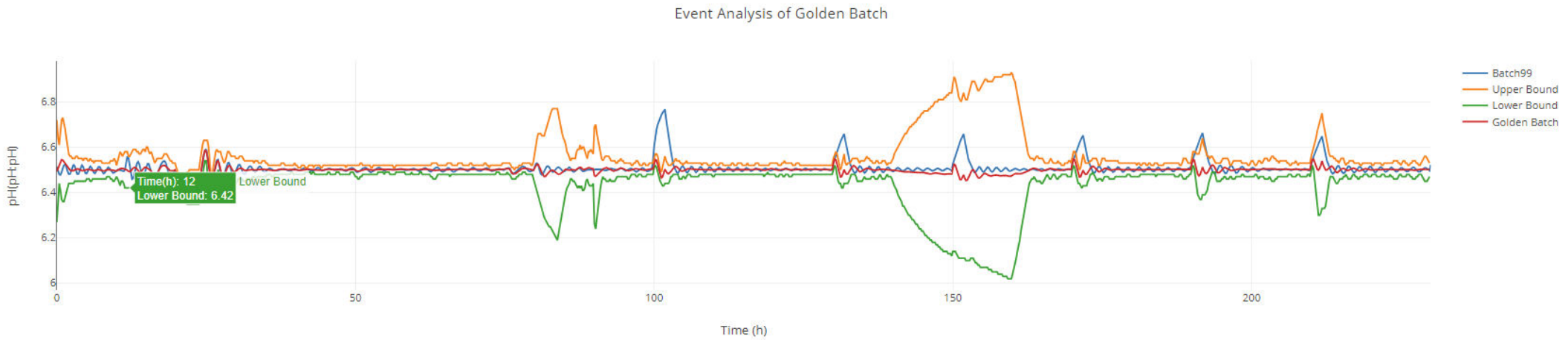
Upload Raw Data

Deviation Analysis | Contribution Analysis | Drill Down

Data Source: BioreactorData.csv

Batch ID Number: 99

CPP: pH(pH:pH)



Rockwell Automation Industry Solutions
Batch Performance Analytics

Hello admin | Logout

- Batch Data Analyzer
- Event Analysis
- Causal Discovery

Event Analysis

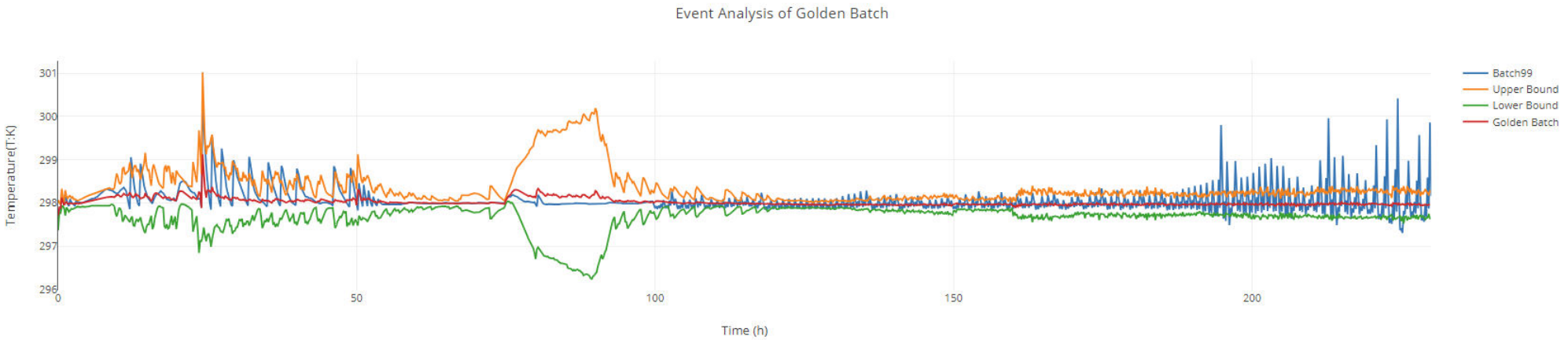
Upload Raw Data

Deviation Analysis | Contribution Analysis | Drill Down

Data Source: BioreactorData.csv

Batch ID Number: 99

CPP: Temperature(T:K)



CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.

Rockwell Automation Industry Solutions
Batch Performance Analytics

Hello admin | Logout

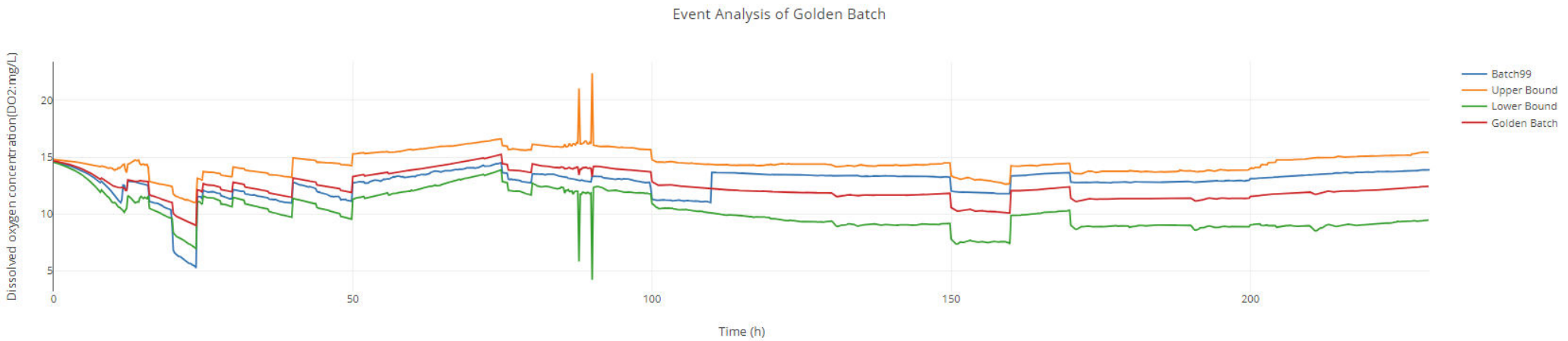
- Batch Data Analyzer
- Event Analysis
- Causal Discovery

Event Analysis

Upload Raw Data

Deviation Analysis | Contribution Analysis | Drill Down

Data Source: BioreactorData.csv
Batch ID Number: 99
CPP: Dissolved oxygen concentr...



CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.

Causal Discovery using XAI model

- **The causal discovery XAI page enables users to drill down to specific record level and find out the features that are affecting the target variable for a specific record.**
- **Step 1**: Pick the Record ID from Event Analysis Golden batch deviation page.
  - Select dataset, batch ID and Record ID for the record you want to analyze, GB deviation analysis page can be referred for finding these details.
  - Select target and input variables.
  - Select required number of top features and click "**Calculate XAI**" button.
  - There are 3 visualizations generated through this.
  - First, Actual vs Predicted plot showing actual target value against model predicted value to verify the accuracy of the model.
  - Second, Feature impact plot , showing top features as per the user input, their range and their impact score on the target variable.
  - Lastly, the data table showing top features, its actual value and its impact score in a tabular format.

- **Step 2:** The sliders of top features would be auto generated along with their existing values in the What-if prediction section.
- Using these sliders, user can adjust values of these top features and look at the impact that change would have on the value of target variable.
- All the variables can be adjusted using the sliders, the new value would be showed in the textboxes next to the sliders for each feature.
- Predicted value would show the latest value of target after adjustments in the input variables and help the end user determine which variables to adjust up to what extent to achieve a desired value for target.

CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.

Batch Data Analyzer
Event Analysis
Causal Discovery

Data*
BioreactorData.csv

CQA*
Penicillin concentr…

CQA Threshold*
20

CPP*
Aeration rate(Fg:L…

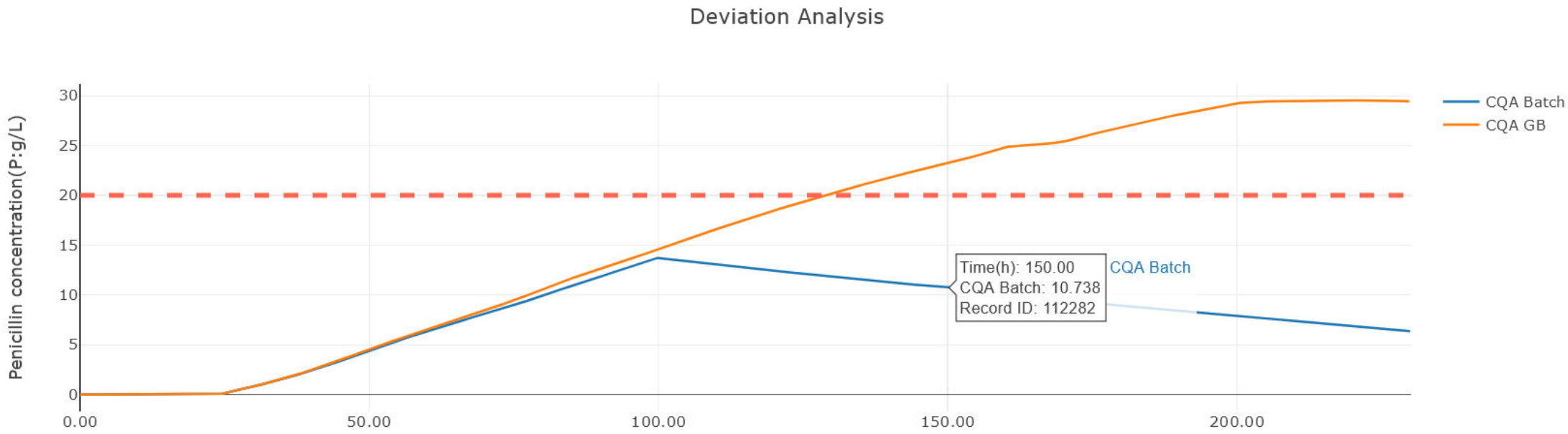
Batch ID Column*
Batch reference(B…

Batch ID*
99

Time Column*
Time (h)

Calculate Golden Batch

Deviation to Golden Batch 54.03%



CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.

Batch Data Analyzer
Event Analysis
Causal Discovery

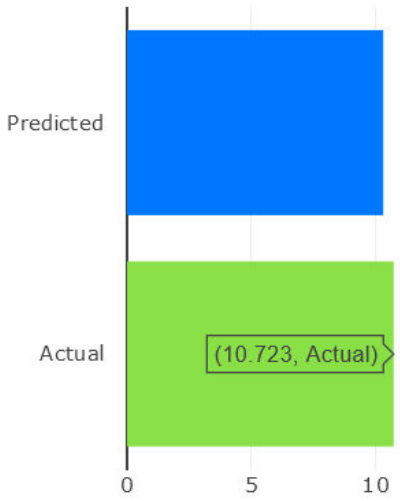
Data*: BioreactorData.csv | Model Type*: Regression | Batch ID: 99 | Record ID: 112282 | Select Target*: Penicillin concentrati...

Select Input Variables*: Aeration rate(Fg:L/h)... | Top Features*: 6 | Target: 10.738

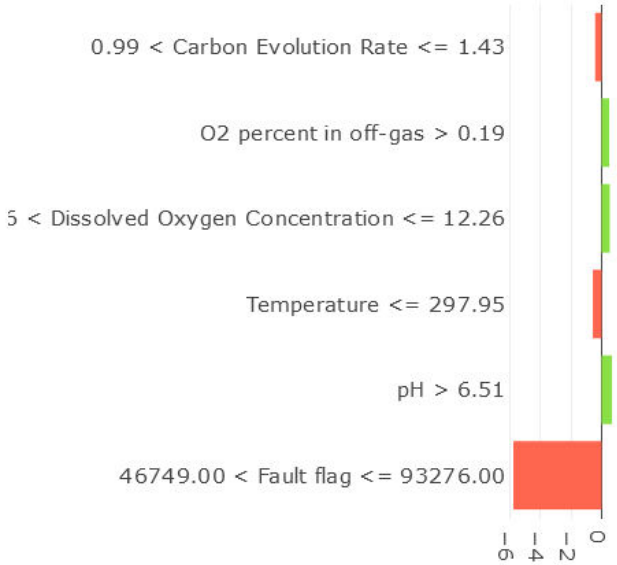
Calculate XAI

Explainable AI

Actual vs Prediction Comparison



Feature Impact Plot



Sample Data Table

Feature	Value	Impact
pH	6.5608	0.68
Temperature	297.86	-0.57
Dissolved Oxygen Concentration	11.996	0.54
O2 percent in off-gas	0.19608	0.5
Carbon Evolution Rate	1.3132	-0.42
Fault flag	70681	-5.78

CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.

Batch Data Analyzer
Event Analysis
Causal Discovery

What-if Prediction

Aeration rate

Sugar Rate

Base Rate

Heat/Cool Rate

Temperature

Dissolved Oxygen Concentration

Aeration rate
75

Sugar Rate
90

Base Rate
77

Heat/Cool Rate
1092.0001

Temperature
297.86

Dissolved Oxygen Concentration
16

Predicted Penicillin Conce

Calculate Prediction

CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.

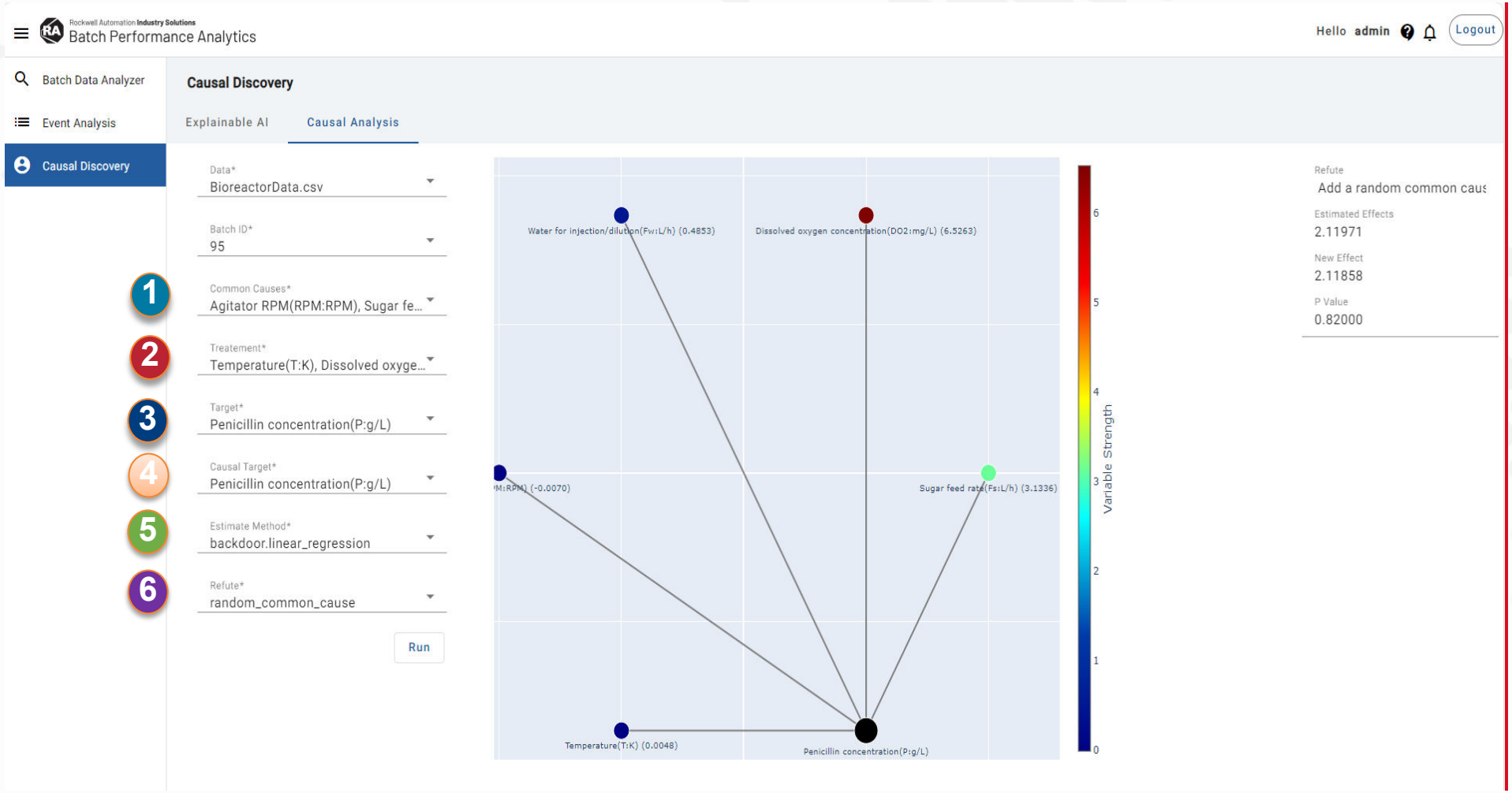
Causal Discovery using Causal Machine Learning

- **The Causal Discovery 'Causal Analysis' page enables the user to answer the question "How strong is the causal influence from a cause to its direct effect?".**
- **Step 1**: : First, create the causal graph for the causal machine learning estimation by specifying the main variables for the analysis, including the common causes, treatment, and target/CQAs.
  - Common causes- Those variables that potentially influence both the treatment and target variables.
  - Treatment: We are primarily interested in understanding the effect or impact of the treatment variables on the target
  - Target: The variable(s) for which we seek to understand how treatment variables affect it.
- **Step 2:** Specify the causal target
- This is the variable of interest in the causal strength graph for which the user wants to quantify the strength or causal influence from common causes and/or treatment.
  - Options will comprise users' combined selections from Treatment and Target.
  - If the selected variable is from Treatment, the user aims to understand the causal strength between the Common Causes and the selected treatment. If the selected variable is equal to the Target, the user aims to understand the causal strength between Common Causes and Treatment, on the target.
- **Step 3:** Specify the learning method
  - used to estimate the relationship between the variables selected in common causes, treatment, and target.
  - Default: backdoor.linear_regression
  - Options backdoor.distance_matching, backdoor.propensity_score_stratification, and backdoor.propensity_score_weighting can ONLY handle a single treatment variable and the selected treatment must be binary.
- **Step 4:** Specify the refute method
  - Evaluate or test causal assumptions based on the user input causal graph and assess the validity of the causal relationships inferred from the data

CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.

Illustrative Example with selection

1. Common causes: 'Agitator RPM', 'Sugar feed rate', and 'Water for injection/dilution' **2.Treatment:** Temperature, Dissolved oxygen concentration
3. Target: Penicillin concentration **4. Causal Target**: Penicillin concentration **5.Estimate Method**: backdoor.linear_regression **6. Refute**: Add a random common cause



- p-value measures whether the new effect is significantly different from the Estimated effect. In our example, there is no significant difference as denoted by the p-value, which implies that the estimator does not fail this test.
- A p-value > 0.05 implies that the estimator is robust.

For example, we observe that the direct influence from Dissolved Oxygen concentration (DO2) to Causal Target (Penicillin Concentration) (~6.3932) is stronger (by ~2 times) than the direct influence from Sugar Feed rate to the causal target (Penicillin concentration) (~3.3932.)

CONFIDENTIAL | Copyright ©2019 Rockwell Automation, Inc.