Advanced Analytics for Optimization of stage-wise ISO Brightness gain in Kraft Pulp Bleaching
Overview of Kraft Pulping Processes

**Major Processes**

- **Wood Chipping**: Reduction of wood logs into smaller size chips
- **Digesters**: Cooking of wood chips with white liquor and steam to remove lignin.
- **Screening and Washing**: Cleaning of pulp from impurities & uncooked wood. Washing of pulp to remove the soluble lingo salts and chemicals.
- **Bleaching**: Removal of residual lignin and to make the pulp whiter and brighter
Major variable cost for Kraft bleached pulp production are:
1. Wood cost
2. Chemicals.
Opportunities for reduction on Variable Cost of Pulp

**Wood Cost**
- Very limited control due to Scarce availability.
- Huge variability due to location of procurement.
- Non availability of measurable Data.

**Chemical cost**
- Depending on the process adopted.
- Depends on the various combination of process sequences.
- More measurable data is available.

**R & M Cost**
- Varies depending on the specialty of equipment.
- Can be optimized by adopting various best practices like TPM etc.,
- Not much of variation over years except the price variation.

**Fuel Cost**
- Limited control over cost & Quality as driven by external forces.
- Best efficiencies are being achieved by adopting newer technologies.
Evolution of Process Controls & Optimization

- Manual control with limited instrumentation
  - Single loop controls
    - Pneumatic instruments
    - PLC based controls
- Distributed control system
  - Electronic instruments
  - Computers.
- Advanced process control
  - Over a localized process and limited data set.
- Advanced analytics
  - Over wide range of processes and larger data set.
Industry 4.0 – An Introduction

Industry 4.0 is the philosophy of applying Internet of Things (IoT) at industry.

IoT is a network of interconnected devices, machines, sensors, data, etc. accessible to humans.

Consequence of Global markets’ search to beat the competition by finding new ways to improve productivity.

Direct result of an exercise to leverage huge volumes of data generated from rapidly evolving electronic & embedded systems and using the same to perform and automate complex tasks with relative ease.

In general, implementation of Advanced Data Analytics is one of the first steps taken at organizations to leverage Industry 4.0 initiatives and envisage the impact, allowing them to integrate Industry 4.0 in their future vision.
Advanced Analytics - Methodology

Similar to traditional six-sigma DMAIC methodology, but tweaked to suit the data size.
Define Problem

- To optimize stage-wise ISO Brightness gain in Kraft Pulp Bleaching
- To reduce the bleaching chemicals consumption

Scope
- Kraft Pulp Bleaching

Process Boundaries
- Brown Stock Washing to Final Bleaching Stage

Main Target Function
- Overall Chemical Consumption in Kg/T of Bleached Pulp
- Deconstructed to identify stage-wise target functions

<table>
<thead>
<tr>
<th>Bleaching Chemical</th>
<th>Baseline (Kg/T of Bleached Pulp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>X1</td>
</tr>
<tr>
<td>C2</td>
<td>X2</td>
</tr>
<tr>
<td>C3</td>
<td>X3</td>
</tr>
<tr>
<td>C4</td>
<td>X4</td>
</tr>
</tbody>
</table>

Baseline - Measured value of target functions before improvement.
Capturing & Structuring Data

**DCS Data Logger**
- 4500 Data Tags
- 6 months data
- 1 hour level

**SAP**
- 450 Data Tags
- 6 months data
- 1 day/shift level

**DCS Lab Logbook**
- 700 Data Tags
- 6 months data
- 2 hour level

+ 75 synthetic tags

**Stitching**

**1 Flat File**

**Most Important Step**

**Things to Ensure!**
- All possible process variations need to be encompassed.
- Necessary to avoid future ineffectiveness due to previously uncaptured variations.

- During data stitching, stage-wise time lags were considered. They are equivalent to average retention time in the stage.
- The variables were moved ahead/moved back based on time lags for synchronizing a data row of all variables to be roughly corresponding to the same pulp, using R or Python Programming Languages.
Cleaning Data

Outlier Treatment
Removal of practically impossible values appearing due to disturbances in plant and/or zero errors.

Multicollinearity
Removal of one or more of highly correlating input variables.
Cleaning Data

Outlier Treatment. Values marked in red were practically infeasible and all values below/above them were filtered out.

Multicollinearity matrix. Darker the cell colour, higher the correlation.
Data Visualization & Initial Analytics

Visualization of relationship between target functions and theoretically important control parameters.

Output pulp ISO brightness vs. Input ISO brightness in a bleaching stage. The dot size corresponds to incoming pulp Kappa number.

Bleaching Chemical Dosage in Kg/T vs. Input ISO brightness. **Strong inverse relationship is observed.**

Pulp ISO Brightness Gain vs. Bleaching Chemical Dosage in Kg/T. It is observed that **ISO Brightness gain peaks at a certain dosage.**

Bleaching chemical efficiency (ISO Brightness gain per Kg/T) vs. Bleaching Chemical Dosage in Kg/T. Bleaching chemical efficiency is defined as ISO Brightness gain per unit dosage of that Bleaching chemical.

Most of the observations reinforced the chemistry behind Bleaching reaction.
Data Visualization & Initial Analytics

Visualization of relationship between target functions and theoretically important control parameters.

Pulp ISO Brightness Gain vs. Chemical Dosage in Kg/T for consecutive bleaching stages. It was observed that dosage beyond a certain value does little to improve the brightness gain. The dots’ colour represents the input pulp ISO Brightness range. The chance of overdosing in B3 is higher because of relative mix-up across input Brightness ranges.

Strengthened the notion of uniqueness corresponding to behaviour variation for different chemicals.
Process Modelling

Identification of other significant non-intuitive control parameters using different modelling methods.

Random Forest: An ensemble of decision trees

Graphical Representation of decision tree

Graphical Representation of Random Forest
Process Modelling

Identification of other significant non-intuitive control parameters using different modelling methods.

**Random Forest**: An ensemble of decision trees

**Caution**: Number of decision trees in Random forest needs to be carefully chosen to avoid any of the undesirable cases of over-fitting or under-fitting of dataset.

<table>
<thead>
<tr>
<th>Underfitting</th>
<th>Desired</th>
<th>Overfitting</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Graphical Representation of model fit
Identification of other significant non-intuitive control parameters using different modelling methods.

**Random Forest**

An ensemble of decision trees

Results of Random Forest for all bleaching stages, showing the top 20 important control parameters. The parameters with highest relative importance are mostly the ones which were theoretically known to have significant relationship with the target function, pulp ISO brightness gain in this case.
Process Modelling

Identification of other significant non-intuitive control parameters using different modelling methods.

Cluster Analysis: Method to identify optimization opportunity within the process, as a function of control parameter variation

K Means Clustering - Classification of Dataset into K Clusters with highest possible degree of differentiation across the said K clusters.

Optimal value of K is determined by calculating average within-cluster sum of squares for each K and finding out the point of inflection for the same.

Graphical Representation of K- Means Clustering

Point of inflection for average within-cluster sum of squares determines K
Process Modelling

Identification of other significant non-intuitive control parameters using different modelling methods.

Cluster Analysis: Method to identify optimization opportunity within the process, as a function of control parameter variation

**K-means Clustering** revealed 4 clusters out of which one cluster was over-dosing due to slow responsiveness to the changes in incoming pulp properties. This cluster can be eliminated by automating the dosage based on multivariate regression of Clusters 1, 2 & 3.
Defining Optimization Logic

### Initial Analytics
- Dosage beyond a certain value resulted in little to no Brightness gain

### Random Forest Algorithm
- The variation of most of control parameters, other than variables identified by process experts for Initial Analytics, does not have a significant effect on bleaching reaction at each stage

### K-Means Clustering
- By modelling the data from Clusters1,2&3 and automating the dosage, the overdosing in Cluster4 can be eliminated in most cases.

#### Feature Engineering is considering a function of a control parameter, like a square root, exponential, logarithm, etc. instead of the control parameter itself, for a better correlation.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Chemical</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1 C1</td>
<td>C1= CON1*(K1)+CON2</td>
<td></td>
</tr>
<tr>
<td>B1 C2</td>
<td>C2=CON3<em>LOG(CON4</em>B1 pH)+CON5</td>
<td></td>
</tr>
<tr>
<td>B2 C3</td>
<td>C3=CON6+(CON7<em>K2)-(CON8</em>B1 Brightness)</td>
<td></td>
</tr>
<tr>
<td>B3 C4</td>
<td>C4=CON9<em>LOG(CON10</em>B3_pH)-CON11</td>
<td></td>
</tr>
<tr>
<td>B3 C5</td>
<td>C5=CON12+CON13(CON14-B22)-(CON15*B2 Brightness)</td>
<td></td>
</tr>
<tr>
<td>B4 C6</td>
<td>C6=CON16-{CON17*B3 Brightness}</td>
<td></td>
</tr>
</tbody>
</table>
Piloting, Validation & Operationalization

Implementation of Auto-Dosage Logic

Formulation of Exceptions and Modifications

<table>
<thead>
<tr>
<th>NFL#1</th>
<th>NFL#2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dosing Switch</td>
<td>Compliance %</td>
</tr>
<tr>
<td>C1</td>
<td>100</td>
</tr>
<tr>
<td>C2</td>
<td>82</td>
</tr>
<tr>
<td>C3</td>
<td>87</td>
</tr>
<tr>
<td>C4</td>
<td>88</td>
</tr>
<tr>
<td>C5</td>
<td>97</td>
</tr>
<tr>
<td>C6</td>
<td>89</td>
</tr>
<tr>
<td>C7</td>
<td>81</td>
</tr>
</tbody>
</table>

Month-wise tracking of adherence% to auto-dosage.

<table>
<thead>
<tr>
<th>Savings Over Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
</tr>
</tbody>
</table>

Month-wise tracking of Chemical Consumption reduction against Baseline.
Results & Conclusion

Benefits of I4.0 Platform

- Reduction in Bleaching Chemicals by 4%
- Reduction in Overdosing cases by 75%
- Reduction in Standard Deviation by 8%
- Potential to increase production by 1.3%
- Reduction in Final brightness above USL by 23%
- Reduction in Final brightness below LSL by 39%
End of Presentation