Real time Prediction of Gas Suspension Calciner Performance Using ML Based Soft

Sai Hitesh B[1], Kiran D Bhor[2], Chandrakala K[3], Vilas Tathavadkar[4], Pradeep Das[5], Ameet Chaure[6], Rajesh Khuntia[7], Kausikisaran Misra[8] and Uttam Giri[9]
1. Scientist*
2. Senior Scientist*
3. Lead Scientist (Domain Leader)
4. Joint president & CTO
Aditya Birla Science and Technology Company Pvt. Ltd, Navi Mumbai, Maharashtra, India
(* Ex-employee)
5. Deputy Manager- Technical
6. Asst. General Manager- Technical
7. Asst. General Manager- Process
8. Assistant Vice President- Process
9. Assistant Vice President- Process
Utkal Alumina International Limited, Rayagada, Odisha, India
Corresponding author: chandrakala.kari@adityabirla.com

Abstract

In alumina refineries, the major energy intensive unit operation is the calcination process, consuming almost 1/3rd of total energy required for the production of Alumina from Bauxite. Technological advancements with new generation gas suspension calciners have offered gradual reduction in energy consumption. The calcination process involves both removal of moisture from hydrate, and phase transformations to produce desirable smelter grade alumina. However, fraction of particles having spent lower residence time leads to a high Loss on Ignition LOI whereas particles exposed to higher temperatures with higher residence time leads to conversion from γ-alumina to α-alumina. For better control on product quality, operator needs to take decision of the set point based on the product quality (Alpha content and LOI) obtained through laboratory analysis. This analysis takes considerable amount of time and the lagged information is used for decision making by operators retrospectively. Therefore, extensive work was carried out to generate real time prediction of these quality parameters which will assist the operator to take proactive process control decisions without waiting for the lab data to achieve operational excellence along with consistent product quality.

Through this work, a Machine Learning (ML) based predictive models were developed and validated based on plant data historian of calcination process. Entire data was split into 70 % for training the model while remainder has been used for testing the model accuracy. Advanced algorithms viz., random forest, Extreme Gradient Boosting algorithm XG -boost was used over the pre-processed data with optimized hyper parameters to achieve higher accuracy for a predictor variable. The model accuracy for the prediction of alpha alumina content was found to be in the acceptable range of 85 % to 90 %.

The predictions were validated using real-time Distributed Control System (DCS) plant data on a minute-wise basis. Through a web-based Graphical User Interface, GUI platform, real time predictions from the model and actual plant measurements were pulled, displayed and compared to showcase the critical parameters for the prediction accuracy.

Keywords: Alpha alumina, Gas suspension calciner, Machine learning model, Soft sensor.

1. Introduction

Thrive of Alumina refineries hinges on energy as well as alumina quality. The most critical process operation contributing these factors is the calcination process. The calcination process is the most energy intensive process consuming ~30 % of total energy of refineries [1]. The calcination of aluminium tri-hydrate to alumina is represented by Equation (1). Although its theoretical energy requirement is in the order of 2393 kJ per kg of alumina, calciners are operated with 12-22 % higher energy in practical [1, 3]. Minimizing these variations not only saves energy but also produces consistent product as calcination process is sensible to operation variations.

$$2Al(OH)_3 \xrightarrow{800-1000\,^{\circ}C} Al_2O_3 + 3H_2O \quad (1)$$

Calcination (Equation (1)) seems to be a simple one step reaction happening at 800 – 1000 °C to produce desirable alumina phases with an LOI less than 1 %. Alumina calcination process involves various stages viz., unbound moisture removal, phase and structural transformations in which hydroxides are converted to oxides. Residual hydroxide content starts to decrease with temperature and time in a calcination process. Most of the reduction in free moisture present in the feed hydrate is observed in temperatures lower than 250 °C after which at a reducing rate of LOI, phase and structural transformations takes place.

Phase transformations play a crucial role in an alumina calcination process for the production of desired alumina phases. During the calcination process, alumina has a significance of existing in various metastable forms including γ, χ, κ, δ, η, θ and in the final stable α-alumina phase [1]. The reaction pathway to several alumina phases are majorly characterized by particle sizes and heating rates. On-set of phase transformation to alumina mono hydroxide (Boehmite) from gibbsite starts at temperatures lower than 250 °C after which several phase transformations based on heating rates and particle sizes to χ, and desirable γ-alumina takes place at temperatures between 300 to 800 °C. Stable alpha alumina formation occurs at temperatures greater than 1000 °C.

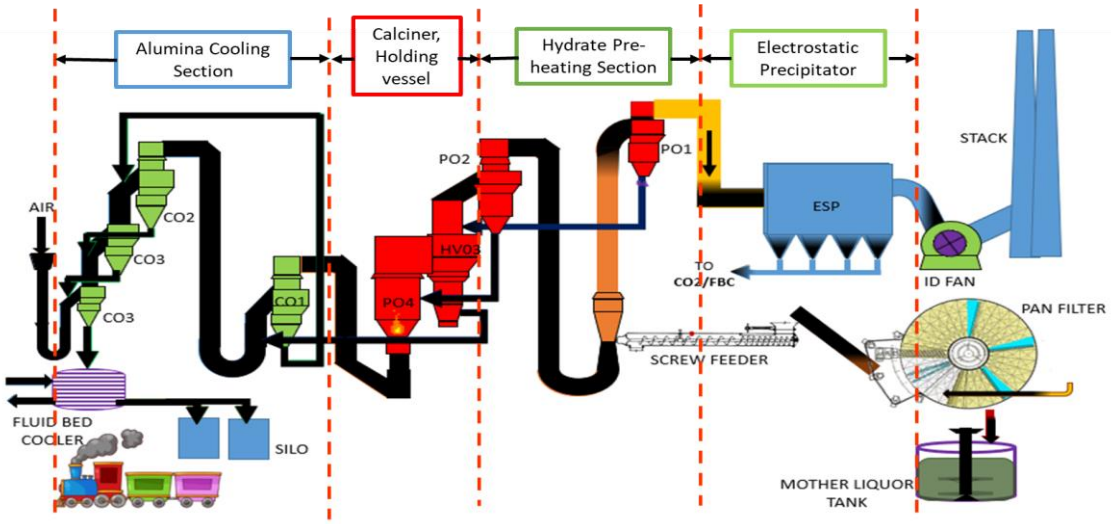

Figure 1. Process flow diagram for alumina calcination.

In modern day Gas Suspension Calciners (GSC) as in Figure 1, gibbsite particles are fed into a series of preheaters maintained at around 300-500 °C for removing moisture and initiating the conversion of aluminium tri-hydroxide to mono-hydrated alumina (boehmite) and transition to alumina. In the calciner furnace and holding vessel sections, production to a desirable smelter grade alumina is carried out. Post calcination, heat associated with alumina is recovered using a

series of cooling cyclones where ambient air gets pre-heated using heat from calcined alumina [2]. The quality of the alumina is critical for smelters if alpha alumina is crossing value of 6 % or higher then energy consumption increases proportionally in smelting pot which is not desirable [3]. Current practice for quality control is based on the feedback loop where quality is analysed for a shift (every 8 h) then corrective action is taken for the next shift operation. This practice of operation causes deviation in alumina quality. Hence there is a need of predictive models for alumina quality which can predict quality in real time. This helps to take pro-active corrective action for quality control. In this paper, predictive model was developed using time stamped data and advanced machine learning algorithms for predicting the alpha phase in the calcined alumina.

2. Technical Approach

The strategic path followed for the development of predictive model utilized as a soft sensor for real time predictions of alpha phase in alumina is depicted in Figure 2. Currently, the alpha value in the plant is found to vary between 3-6 % in the smelter grade alumina. The alpha phase formation is dependent on temperature, particle size and time spent by particles at temperatures greater than 950 °C.

A higher value of alpha alumina (> 5 %) would not be beneficial in an aluminium smelter, the predictive model would thus aid as a real time effective decision-making tool for plant operation personnel. The current operational best practices can control variation up to certain extent. Further an in-depth study and advanced tools are required for the precise quality control. Hence problem has been formulated as controlling variation in alpha alumina by data analysis in combination with advanced machine learning tools.

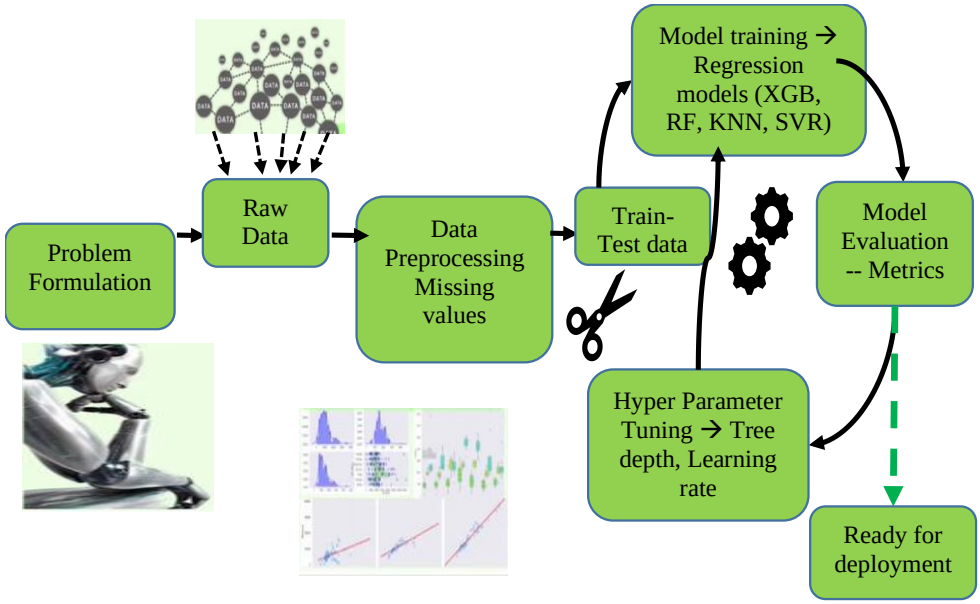

Figure 2. Path followed for the development of the predictive model.

The important step after formulating the problem is to gather the historical process data with minute-wise frequency for the entire calcination circuit. Upon performing exploratory univariate and multi-variate data analysis, several process variables are shortlisted, while the data tags which are irrelevant to the process are discarded. The practical process data has lot of missing data/zero values due to malfunction of the sensors, issues in networking, sudden power cut or scheduled shutdown. Thus, these missing values or arbitrary data should be removed from the data set in order to have healthy database for model development. Once sufficient database is available, data pre-processing was initiated for the database using outlier treatment with three sigma filtering.

The outlier treatment was carried out in Python version 3.8.1. The sample of alpha phase in alumina was measured in analytical laboratory with a frequency of once in 8 h (shift-wise). Hence, data with a frequency of 3 data points per day was available. Other real time DCS data was obtained every minute from plant historian. Thus, it was important to map the time stamp with the measured values. Time stamping was carried out by averaging all DCS variables from data set for the time 10 min prior of sample collection and 10 min post the sample collection. The data sets have more than 100 variables and each of these variables are affecting the quality parameter. In order to ease the model development, most affecting parameters were selected and the interdependent variables along with least affecting variables were discarded. The exploratory data analytics technique was applied by plotting correlation matrix using heat maps generated using seaborn library. This helps in the identification of correlation of all the independent variables with respect to the dependent variable (i.e., alpha alumina) and classify the interdependent and least important variables from complete dataset.

Further the entire data set was split into two parts whereas 70 % data used for developing the predictive models while remaining 30 % data used for testing the model capabilities in terms of alpha predictions. The step of data training/multivariable regression involves use of advanced algorithms like K-Nearest Neighbour (KNN), Support Vector Regression (SVR), Random Forest and Extreme Gradient Boosting (XG-Boost). Out of these, XG-Boost was used with minimum error. The error value was calculated by predicting the predictor values and validating with the actual plant sample analysis, here evaluation matrix was alpha alumina content from laboratory analysis. The hyperparameters were optimized based on the evaluation metrices and final model was ready for the offline deployment. This model was tested for more than 7 months dataset and found to be robust enough to be used for online deployment. These models are machine learning models hence after every 30 days of operation model database is upgraded with additional data and retraining the model with latest data splits optimized hyper parameters. The Graphical User Interface (GUI) is a simplified version of accessing the results of predictive model this can be easily utilized by the operators for taking corrective action based on the prediction of content of alpha alumina. The GUI was developed using web-based platform where Online data from DCS of the process plant was accessed and processed as an input to python model. On the other hand, data from laboratory analysis was compared with the predictions of the alpha alumina content. The details of the python model and GUI would be discussed in the results and discussion section.

3. Results and Discussion

3.1 Exploratory Data Analysis

Univariate, multi variate and time series data analysis was carried out for the process parameters using Tableau and Excel. The in-depth data analysis provided several interesting insights on product quality (alpha content) with respect to seasonal effect, effect of process parameters such as temperature, flow rate of fuel as Heavy Furnace Oil (HFO) and air etc. The data is segregated for different seasons and was depicted in Figure 3. For the detailed data analysis both for specific energy consumption and alpha content, two multivariate analysis has been illustrated i.e., hydrate feed rate as independent variable while dependent variable is specific energy consumption. While another plot shows effect of calciner temperature on alpha phase content in alumina. The energy consumption during winter session is always above 2800 kJ/kg of Al_2O_3, while during summer energy consumption has been on lower than winter session. When alpha alumina is compared for the seasonal effects, it can be clearly observed that during summer season the alpha content is below 4 %, and during winter there is high deviation ranging from 3.5 % to 6 %. This deviation could be a result of higher temperature in calciner i.e. majority of data point are above 1100 °C. Higher temperature is one of the reasons for high alpha alumina as per process understanding for calcination. Therefore, effects were made to co-relate the process data to understand effect of process variables on the calciner temperature and was shown on Figure 4.

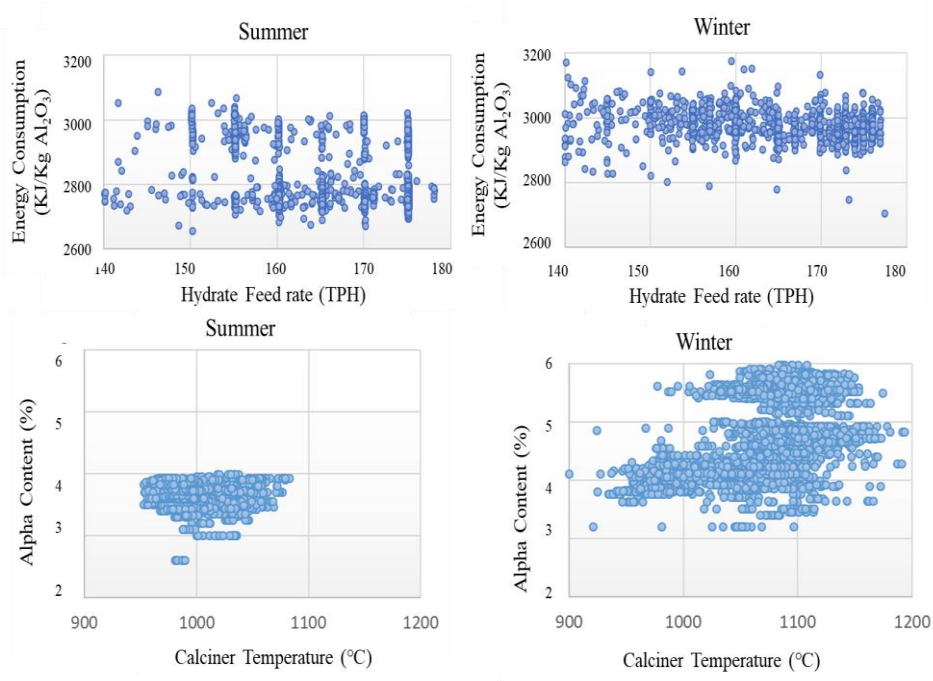

Figure 3. Seasonal effect on energy consumption and alpha alumina.

For simplicity of understanding the data has been segregated based on the calciner temperatures. The process parameters such as fuel (HFO) and air has been normalized to the input hydrate feed rate and is shown in Figure 4. Although, most of operating data for HFO flow rate per ton of hydrate is in the range of 40-45, lower alpha content can be achieved by maintaining lower values 40-43. Comparatively, the effect of air to fuel ratio is more evident which elucidates that lower the air/ fuel ratio, higher will be the alpha content. This is mainly due to the lower energy utilized for the controlled air provided to calciner. This segregation of data in-turn helped to identify the best process parameters to have lower temperature as well as lower alpha content from the segregated using best fuel to hydrate (40-43) as well as air to fuel ratios (22-24).

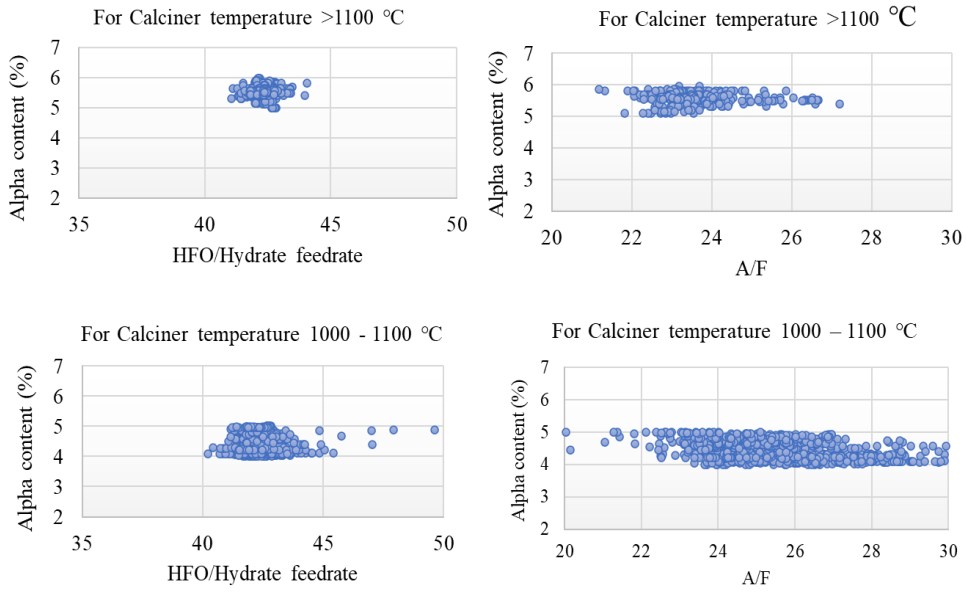

Figure 4. Effect of critical operating parameters on alpha alumina content.

3.2 Predictive Model Development

Python 3.8.1 was used involving various steps as discussed in technical approach in order to develop the machine learning model. The plant time stamped data (10 min average around the sampling time) along with the measurements of the quality variables for 2 years has been utilized. All the variables have been scaled to a range of '0' to '1' based on their minimum and maximum value according to the Equation 2. Scaling was performed in order to bring all the variables into one uniformity. This aids in removing the unnecessary biasness. In python, Standard Scaler is used on the data set.

$$X_{Scaled} = \frac{X - X_{min}}{X_{Max} - X_{min}} \quad (2)$$

To quantify the effect of each predictor variable on the quality variables, the Principal Component Analysis (PCA) and Projection on Latent Structures (PLS) technique have been utilized. However, these techniques could capture only 24 % of variance in the data.

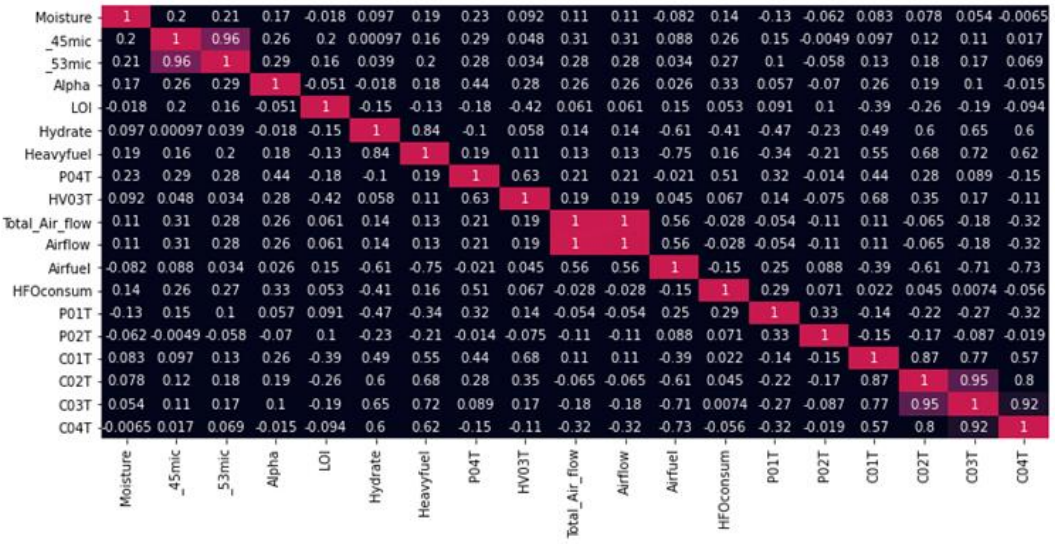

Figure 5. Corelation heatmap for the variables in calciner.

Corelation matrix using heat map generated from seaborn library of python for all the variables critical in the quality parameters has been plotted in Figure 5. A correlation heatmap shows 2D correlation matrix between two discrete dimensions, using coloured cells to represent data from usually a monochromatic scale. Correlation matrix explain the degree to which independent variables have an effect on the dependent variable (Alpha content). A positive value indicates positive relationship i.e. with increase in value proportionally increases and vice versa. Each square box in heatmap shows the corelation between the variables on each axis. Corelation ranges from -1 to +1 value. The values closer to zero implies that there is no linear trend between the two variables. The value close to 1 indicate that the correlation is more positive. A correlation closer to -1 is similar but instead of both increasing one will increase while other will decrease. In addition to the correlation matrix, process knowledge was applied while finalizing the variables for predictive model development of alpha alumina. It was observed that the particle size of -45 µm has strong relationship with the alpha alumina but this variable could not be considered since online measurements are not possible with current technology available in the market.

The entire data was processed by eliminating outliers and scaling to the desired range for each variable was performed as per Table 1. These values for maximum and minimum were carefully

selected from the process expertise and historical data analysis performed on two-year data. Due to the high non-linearity in the data, a robust machine learning algorithm XGBoost has been considered for model development. After pre-processing the raw data by eliminating outliers and scaling to the desired range, 2022 data points have been obtained out of which 1 617 have been utilized for training and 405 for model validation purposes.

Table 1. List of Predictor variable and their operating range.

Sr. No.	Predictor Variable	Unit	Minimum	Maximum
1	Hydrate feed-rate	TPH	115	200
2	Heavy fuel oil flowrate	kg/h	4 600	8 500
3	Calciner Temperature	˚C	850	1 200
4	Holding Vessel Temperature	˚C	800	1 150
5	Air Flowrate	kNm^3/h	105	150
6	Preheater Temperature P01	˚C	140	205
7	Preheater Temperature P02	˚C	250	340
8	Cooling Cyclone Temperature C01	˚C	590	800
9	Cooling Cyclone Temperature C02	˚C	360	580
10	Cooling Cyclone Temperature C03	˚C	198	410
11	Cooling Cyclone Temperature C04	˚C	100	208
12	Ambient temperature	˚C	8	45

The hyper parameters have been tuned using randomized grid search algorithm with 5-fold cross validation and the optimum values of the tuned parameters have been mentioned in Table 2. The hyper parameter tuning was a crucial step for a good accuracy of the model. Thus, a randomized grid search runs iteratively in such a way that the accuracy was intact. Using the finalized hyper parameters, the model was trained and tested using the data after train-test split as discussed in technical approach. Once, the training and testing were completed, it was important to rescale the data to the original in order to compare and analyse the degree of model accuracy. For the rescaling inverse laplace transform was applied in python. Further, several metrics i.e., mean absolute error, RMSE (Root mean square deviation) etc., available in python were used. Further, a new data which was not used in training and testing is utilized for validation of the predictive model.

Table 2. List of hyperparameter and their value.

Sr. No.	Hyper Parameter	Value for alpha model
1	Subsample	0.6
2	n_estimators	1 000
3	Min_child weight	1
4	learning_rate	0.004
5	Colsample_bytree	0.9
6	Reg_lambda	0.5
7	Colsample_bylevel	0.4
8	Max_depth	9

The predictions from the validated data was extracted to excel to evaluate the predictions with respect to the actual alpha content. The predictions values of Alpha % have been presented in Figure 5.

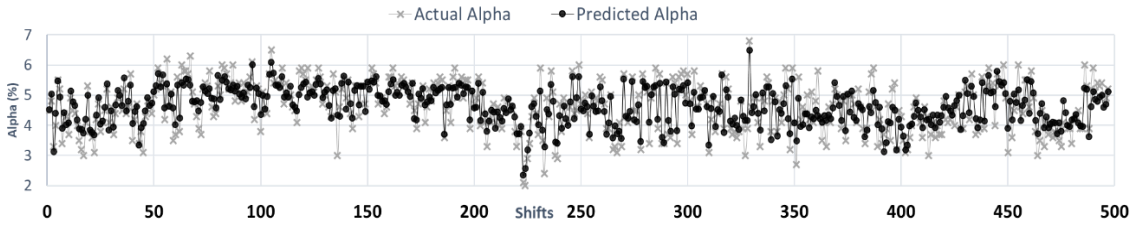

Figure 5. Predictions for the alpha alumina shift-wise.

In Figure 5, x-axis shows shift number (as a time) while y-axis shows predicted values alpha. There are 2 plots in Figure 5 where black circle marker shows predicted values and grey cross marker shows actual alpha measurement from the laboratory. Total time duration for the predictions was of 6 months. The plot clearly indicates the prediction capabilities of the model where it was able to predict the trend very well although magnitude of the value has accuracy from 80-85 % in most of the data. The accuracy was reduced in certain intervals during shifts of 320-325 this behaviour was because of instability in operation of calciner, similar incidents are observed at shift 450 and 220 -225.

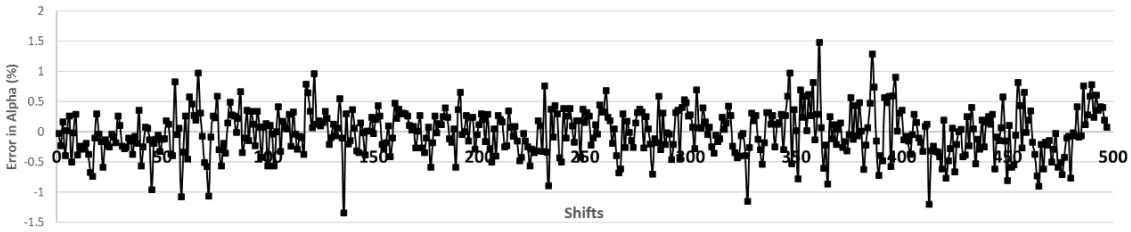

Figure 6. Absolute error % in alpha prediction value.

Table 3. showcases the percentage of data falling in different bands of error percentage showing that 85 % of the predictions are falling in the acceptable range.

Table 3. Model performance for Alpha predictions

Parameter	Region	Percentage points
Alpha prediction	Points lying within the deviation of alpha measurement $0 \leq \lvert e_{abs} \rvert \leq 0.4$	61.6 %
	$0.4 \leq \lvert e_{abs} \rvert \leq 1$ (Absolute error tolerance = 0.6)	23.6 %
	$\lvert e_{abs} \rvert > 1$	14.8 %

To understand the model accuracy in a better way, absolute error in value was calculated which was difference of predicted and actual value in Figure 6 for the same time duration. The values of alpha measured in laboratory using X-ray diffraction method, this method itself have an error of 0.5 %. From graph it can be observed that more than 85 % data is within 0.5 % error while remaining data shows error of 1 %.

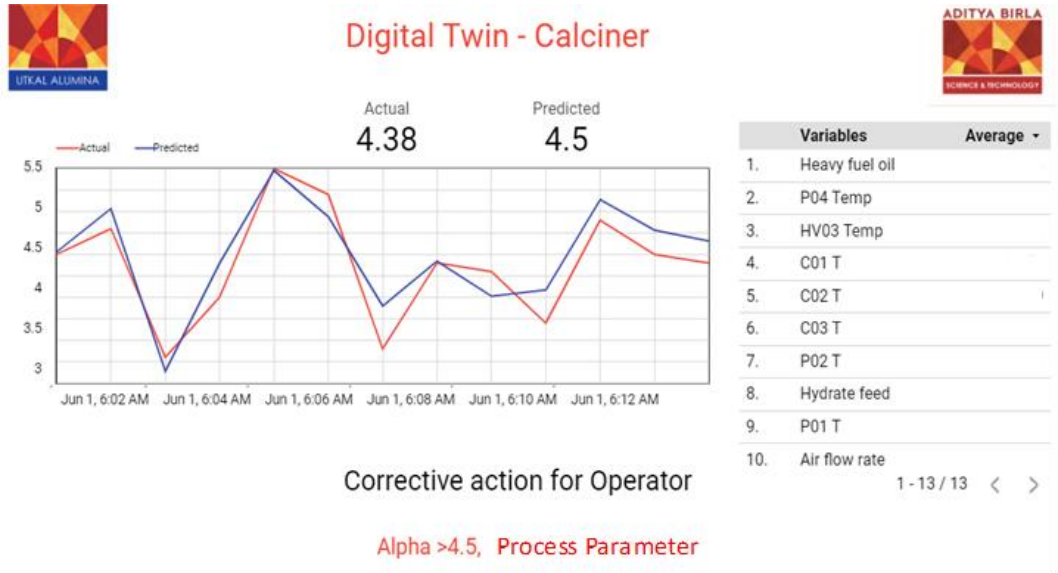

Figure 7. Sample dashboard for the real time predictions of alpha alumina.

The sample dashboard designed for GUI development; graphic user interface/dashboard was developed to display the results of the predicted alpha alumina content minute wise/hourly along with actual alpha content (Figure 7). In addition, it also involves the display of important variables /features which played a key role in the model prediction in the order of their respective ranking. On top of that, past trends of each variable, predictions and the reports can be generated for the benefit of the DCS operating personnel to understand the past trends and the respective decisions made during the period.

4. Conclusions

On the conclusive note, the current work focused on carrying the following technical methodologies viz., A) Data Analysis: Identification of critical variable affecting the process. B) Predictive model development based on data analytics techniques for product quality. A strong seasonal effect on the operating variables and the corresponding variations in heavy fuel oil consumption and product quality in terms of alpha content are observed. After testing various models XG- boost algorithm gave a better result in terms of predicting the Alpha alumina content. The developed Machine learning model is capable to predict the Alpha alumina with accuracy of 85-90 %. This model is further utilized in developing the web-based platform where predicted and actual alpha alumina will be displayed along with other critical parameters of process. This platform will be useful for operators to control process in a better way and take corrective action against the current practice of taking corrective action post laboratory analysis of the alumina. Further study focuses on developing the web-based platform which will accommodate the predicted values of alpha alumina from ML models as well as actual lab measurements of alumina.

5. References

1. B.E. Raahuge, *Smelter Grade Alumina Quality In 40+ Year Perspective: Where To From Here*, Light Metals, (2015), 74-79.
2. C. J. P. Steiner, D. P. H. Hasselman, R. M. Spriggs, *Kinetics of the Gamma-to-Alpha Alumina Phase Transformation*, Journal of American Ceramic Society, Vol. 54, No. 8, (1971), 412-413.
3. B.E. Raahuge, *Thermal Energy Consumption in Gas Suspension Calciners*, FLSmidth, ICSOBA, 2017

4. Aurelien Geron, *Hands-on Machine learning with Scikit-Learn, Keras and Tensor flow* , 2nd Edition 2019.