Enabling Value added Product (UTR) Rolling using Artificial Intelligence based Adaptive Model

Wasib Qadri 1[a]*, Ranjay Kumar Singh 2[b], Sanjeev Kumar 3[c], Alok Kumar4[d]

[a] *Sr. Technologist, Automation Dept, Tata Steel India Limited, Jamshedpur [1]*
[b] *Head Rolling Mill, LD3 TSCR Dept, Tata Steel India Limited, Jamshedpur [2]*
[c] *Head OA-Mills, Automation Dept, Tata Steel India Limited, Jamshedpur [3]*
[d] *Head Thin Slab Caster, LD3 TSCR Dept, Tata Steel India Limited, Jamshedpur [4]*

---***---

Abstract - *The width deviation is a critical statistic for evaluating the quality of a hot-rolled strip in steel production systems. This research proposes a machine-learning based Linear Regression Model to develop prediction model for predicting Final width in CSP Plant esp. for rolling special material with a tight width tolerance band than existing OEM Mill design. The overall objective of this effort is to create an online prediction model that can predict accurate width to be set at entry stage, thereby avoiding under gauge or over gauge issues at Mill , and can lead to rejection of the Coils.*

An original evolutionary algorithm and generalized linear regression was developed by choosing a representative variable from each process parameter impacting width and accordingly a prediction model was developed by using machine learning approach.

***Key Words*: Quality, Hot Rolled Coils, Width TOP OF COPPER, Machine Learning, Artificial Intelligence**

1. INTRODUCTION

Only at the mold stage in the caster can the width be adjusted, and only at the IMS Gauge following the F6 stand can the width be measured. To achieve a tighter tolerance limit for enabling rolling of special material than normally rolled material, having much higher acceptable width tolerance. As such Width Top of Copper in CSP Plant must be set accurately.

The major challenge was in creating an online model that would assist operations in reaching desired Width in Mill. As a result, when creating the model, consideration was given to the casting, furnace, and mill conditions by taking all the related data from all these three areas into account.

A Machine learning based model using Linear regression approach and a self-learning model was developed that considers the real time status of the Plant into consideration during prediction. The model's Computational results for varied mix of Grades, Width and Thickness sections were compared to the results of Live trials, and the model's predictions were verified in each of the five trials that were conducted for all sections. Finally, the model has been effectively applied online. This has made it possible to roll this unique grade of special material in mass quantities and an alternate route of production for this special grade has started which is a great achievement and first of its kind across all similar kind of Plants. Mass Rolling of this grade with special tolerance band has been going on now normally for the past six months without any quality-related concerns.

Due to the complexity of the task, the model must be self-learning and take into consideration the current mill condition for which an adaptation coefficient has been added. The model results need to be matched with ongoing trials with model results to be verified from actual trials for varied GWTL and sections.

Furthermore, the implementation of this model uncovers which process variables are correlated with the Final strip width and devises a path for the goal of setting Width Top of Copper for achieving desired strip width in Mill.

1.1 PROJECT STEPS

To obtain a predictive model using ML the following steps were followed:

1.1.1 Dataset Extraction: Develop a script to extract all possible information from several data sources like CSP areas.

1.1.2 Define the Goal: To select the best ML model and its parameters we define a metric by which the models will be compared and try using various machine learning techniques like Random Forest, regression.

1.1.3 Model / Feature Selection: Perform several iterations with different models and different sets of features and compare the selected metric for zeroing down on final model.

1.1.4 Try development final equation basis PDI Width range from narrow to wide section wherein UNTRIMMED ROLLING coils will be rolled.

1.1.5 Evaluation: The results are evaluated in terms of predictive Width Top of Copper and model results to be verified with various ongoing UNTRIMMED Rolling trails.

1.2 DEFINING THE DATASET

We define the *dataset* as the collection of labeled samples. The accuracy of any Machine Learning algorithm depends on the quality of the dataset, in our case this means that we need:

- Last one year data of Coil data.
- Last one year data of casting parameters.
- Last one year data of Target PDI from Level 3

1.3 Action Plan

- To do a basic exploratory data analysis of width raw data set to identify current Coil width Variations in CSP Plant.
- Identify pool of process parameters from Caster, Furnace and Mill having impact on final strip width deviation.
- Try to Identify a correlation among all process parameters on Final strip width.
- Develop a width Prediction Model based on all process parameters identified.

1.4 Process Background

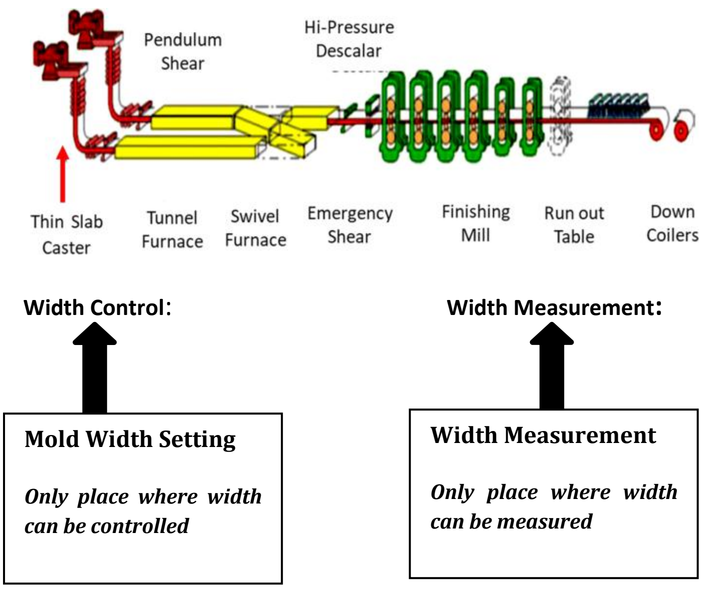


Figure 1- Process Background for Width setting and width control in CSP PLANT

In CSP Plant, the only place where width can be controlled is in Caster stage and the final output where width is measured is in Mill. Once the slab is generated from CSP Plant caster, operator can come to know final width only when Slab has reached Mill, which has a minimum time gap of 30 minutes. The model plans to predict Width to be set in Caster, to achieve desired width in CSP Plant Mill.

1.5 Data Landscape

Initially we tried to analyze the master Data set to determine the existing width deviation scenario in CSP Plant basis past historian data. We tried to check for any data skewedness and overall width distribution.

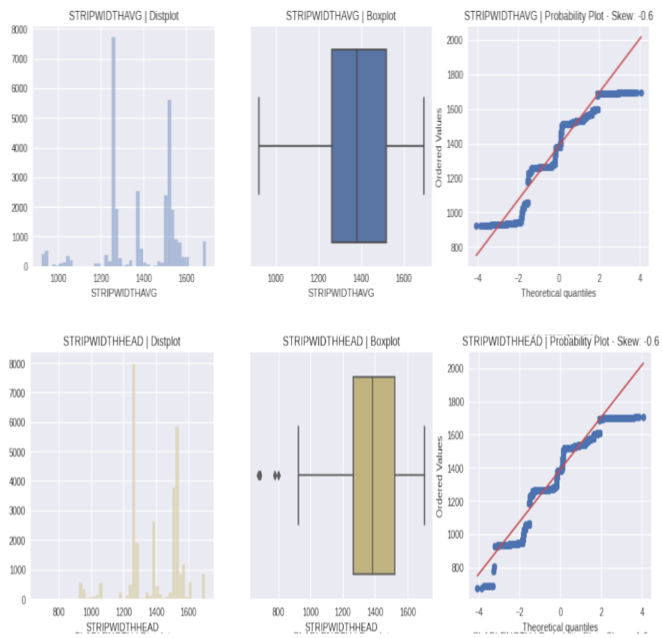


Figure 2- Overall scenario of Width distribution in head and overall

As analyzed from data, most of the coils are rolled around narrow Section of width as well as wider section width. The deviation in head and tail part of coil is more than the width deviation is rest of the coil as tension is yet to build and coil is moving freely after F4 stand. This indicates that width control needs to be more carefully considered in head and Tail part

The main points which can be inferred from above exploratory data analysis is as:

- Special Grade Rolling width acceptable Target is leaser than normal material.
- Special Grade Production will be done around two major width sections – around narrow & wider section
- Average Width Deviation is found to be more than the desired tolerance band for this type of rolling from past data.

- Average Head end Dev. is found to be way greater than desired tolerance band indicating head end control for enabling special Rolling from CSP Plant is more challenging than body

The next important step in the development of the width prediction model is to find correlation between various CSP Process parameters impacting final width.

Initially, a heat map algorithm wherein a matrix layout with color and shading to show the relationship between two categories of values. The values are presented along each axis and the corresponding cell is then color-coded to represent the relationship between the two categories. Here, we have used a heat map to identify correlation between various process parameters. It is a matrix which helps us to identify strong correlated variables.

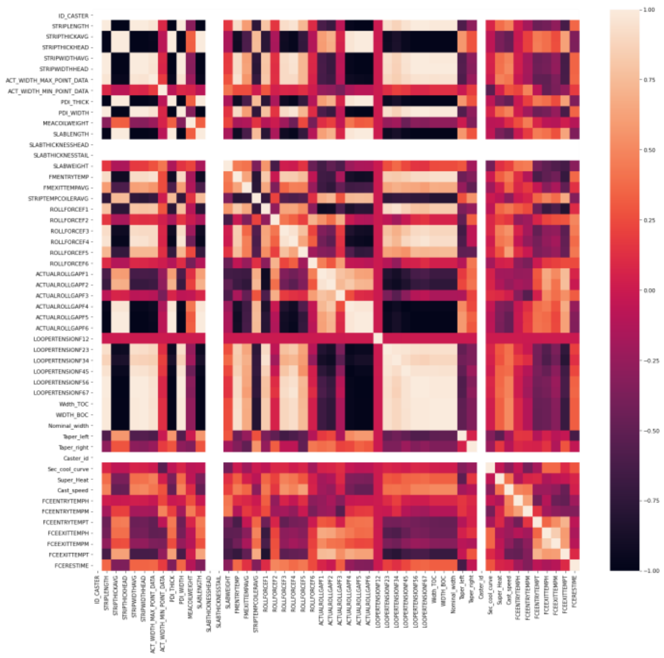


Fig 3 Showing Co-relation using co relation matrix algorithm

Parameters identified with strong correlation with Strip width:

- Width TOP OF COPPER
- Taper Right
- Taper Left
- Cast Speed
- Superheat
- Roll Force.
- Actual Roll Gap.
- Looper Tension
- FM Exit Temp.
- Slab Weight
- Coil Weight
- FM Entry Temp.
- Coiling Temp.

The co relation between various variables was further cross verified by using Pair plot co relation technique. The technique reiterated the same strong co relation between the variables already established by Co relation matrix, indicating our right approach.

Pair Plot is a method which allows to see the distribution of data and relationship of various process parameters with target variable

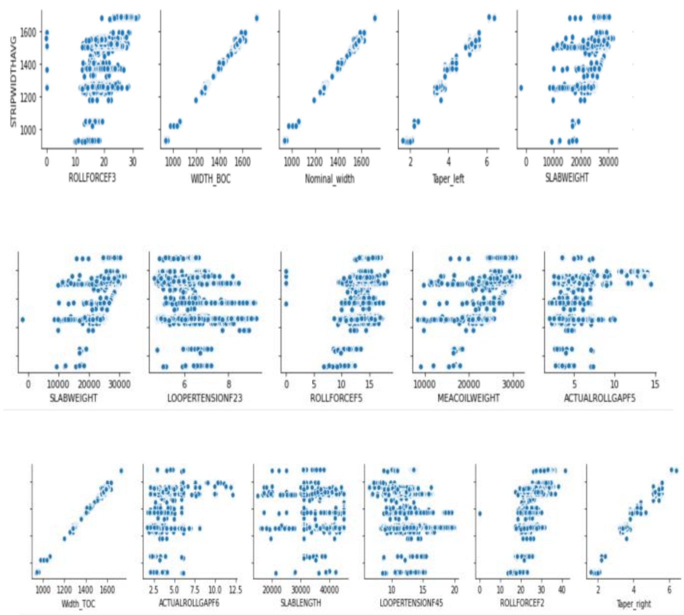


Fig 4 Showing Co-relation using pair plotting algorithm

From the Pair plotting, the process parameters selected from heat map have good strong correlation with Width.

Width TOP OF COPPER, Width Bottom of Copper, Taper Right, Taper left, Roll Force Stand 2, Slab weight are showing a good correlation with width.

2 Feature Engineering

For developing models, initially we have tried the Linear regression approach and the random forest approach.

Linear regression is a basic and commonly used type of predictive analysis. The overall idea of using a regression approach is to determine whether process variables selected for development of the model are good enough in predicting Mold width. These regression estimates are used to explain the relationship between one dependent variable and one or more independent variables.

Regression approach is helpful to identify the effect of independent variables on dependent variables. Also, it can be used to understand the impact of change in dependent variable with regards to independent variable which is helpful in our case.

Another approach we tried is Random Forest approach which is widely used for regression problems as well as identifying feature importance of process variables.

2.1 Random Forest Approach

For all the process parameters identified, we first used a random forest approach which helped us to identify the feature and process variables selected. The model results are as below:

```
Variable: Width_TOC              Importance: 0.17
Variable: WIDTH_BOC              Importance: 0.16
Variable: Nominal_width          Importance: 0.15
Variable: Taper_right            Importance: 0.13
Variable: Taper_left             Importance: 0.13
Variable: SLABWEIGHT             Importance: 0.06
Variable: MEACOILWEIGHT          Importance: 0.06
Variable: ROLLFORCEF1            Importance: 0.03
Variable: ROLLFORCEF4            Importance: 0.02
Variable: ROLLFORCEF2            Importance: 0.02
Variable: ROLLFORCEF3            Importance: 0.01
Variable: STRIPTHICKAVG          Importance: 0.01
Variable: LOOPERTENSIONF67       Importance: 0.01
Variable: ACTUALROLLGAPF6        Importance: 0.01
Variable: SLABLENGTH             Importance: 0.01
Variable: STRIPTEMPCOILERAVG     Importance: 0.0
Variable: ROLLFORCEF6            Importance: 0.0
Variable: LOOPERTENSIONF12       Importance: 0.0
Variable: STRIPTHICKHEAD         Importance: 0.0
Variable: LOOPERTENSIONF23       Importance: 0.0
```

Fig 5: Model results showing importance Parameters

In the next step, features with importance less than 0.004 were not considered and were removed from final model development.

To improve model prediction, we tried Removing few Features which were having importance less than 0.004. Accordingly, a final list of 27 parameters with feature importance were selected.

The final model results are as below:

Target Variable [Strip Width]	Training Set Mean Absolute error	Test Set Mean Absolute error	Inference
Linear Approach	3.21	3.3	Both these approaches showing almost same result. To cross check overfitting or any complexity
Lasso Approach	3.25	3.34	

Fig 6: Showing Test and train set results

2.1 Linear Regression Results

As the co relation among variables was high and good enough, so we used linear regression model to understand width deviations from target as well achieve final model equation to be deployed for getting final target width Top of Copper.

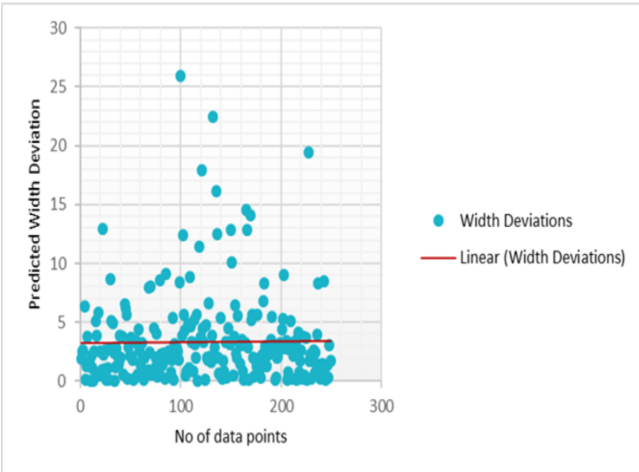


Fig 7: Model results showing No of width deviations

2.2. Mathematical equations:

As special grade Rolling will be done between two main sets of PDI Widths, so the dataset for training model was divided into main sets considering narrow and wider section widths into account. The purpose of dividing the dataset into two separate widths is to improve model accuracy and width prediction as set up calculations undergo major variation in narrow and wider sections of rolling.

PDI Width for Narrow section

PDI Width for wider section

Dividing data set into two parts narrow / wider and finally devising mathematical equation to predict final Width Top of Copper for operation visualization was finalized as below:

Equation Results for Narrow section

Total Data	Within 0-4 mm	Within 0-5 mm	Within 0-7 mm	Within 0- 9mm	Within 0-12 mm	Out of tolerance > 0 -12 mm
250	186/250	200/250	227/250	231/250	237/250	13/250
Prediction %	74.4	80	90.8	92.4	94.9	5.2

Equation Results for Wider section:

Total Data	Within 0-4 mm	Within 0-5 mm	Within 0-7 mm	Within 0- 9mm	Within 0-12 mm	Out of tolerance > 0 -12 mm
250	177/250	193/250	219/250	230/250	236/250	14/250
Prediction %	70.8	77.2	87.6	92.0	94.4	5.6

Result of this approach is as under:

Target Variable [Strip Width]	Training Set Mean Absolute error	Test Set Mean Absolute error	Inference
1100 to 1350 mm	2.64	2.71	This approach showed improvement from around 3.30 to 2.71 in Mean Absolute Error.
1350 to 1600 mm	3.43	3.55	However, this model approach showed slight increase from previous 3.30 to around 3.55

3. Live Trails basis model

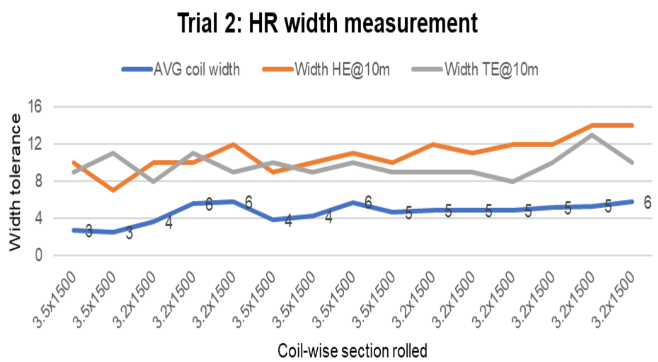
Several trials for special grade Rolling were taken in CSP PLANT to check the feasibility for enabling special grade Rolling . During first rolling, model development had not commenced, and it was not a success. During the third trial, model development had completed and the basis PDI Width desired, the final mold width in Caster was set as per model predicted Width TOP OF COPPER. The third trial was successful, and all coils passed quality clearance and were within accepted tolerance band

The comparative analysis was done to understand the difference between process parameters of both trails.

	trial_1_mean	trial_2_mean	trial_1_min	trial_2_min	trial_1_max	trial_2_max
CAST_SPEED	5.027333	4.073333	4.70000	3.50000	5.52000	5.10000
SUPER_HEAT	19.514267	27.300000	11.00000	15.30000	27.00000	37.30000
TAPER_RIGHT	5.399891	5.100379	5.39943	5.10026	5.40011	5.10049
TAPER_LEFT	5.400095	5.099968	5.39954	5.09992	5.40079	5.10004
WIDTH_TOC	1571.800000	1538.600000	1543.00000	1537.00000	1592.00000	1540.00000
WIDTH_BOC	1560.866667	1527.666667	1532.00000	1526.00000	1581.00000	1529.00000
SLABTHICKNESSHEAD	62.000000	62.000000	62.00000	62.00000	62.00000	62.00000
SLABTHICKNESSTAIL	62.000000	62.000000	62.00000	62.00000	62.00000	62.00000

- First Trail taken was not successful wherein avg. Width Deviation was more than the acceptable width range.
- 2nd Trail mean deviation was well within tolerance band and was a huge improvement and step forward towards enabling special Rolling in CSP PLANT.
- Comparative analysis of parameters identified clearly explains the difference between two different trail results.

The Trail 2 was satisfactory, and the model calculated width was set in Caster giving optimum results, thereby giving a huge leap towards enabling special grade Rolling in CSP PLANT. Below is the summarized view of trail 2 width measurements



4 CONCLUSIONS

The Model accuracy was high and matching with ongoing trails in CSP PLANT for predicting Width. A visualization has been developed for viewing model output and visualizing overall width profile of coil wherein its deviation in head, body and tail is generated dynamically. This helps operation team to analyses profile of last rolled coil and values of all process parameters having impact on final width.

Afterwards, third and fourth trail was also done with different other grades and sections. Model accuracy was accurate, and trails were successful. Now, Mass special grade Rolling has started from CSP PLANT Route in TATA Steel India LTD, Jamshedpur making an alternate route of production for special grade Rolling apart from conventional HSM Mill.

REFERENCES

[1] William L Roberts : "Hot Rolling of Steel", Volume 1 , Issue 2 , April 2008

[2] Arif S. Malik ; Ramana V. Grandhi: "A computational method to predict width profile in Rolling Mills", IEEE Transactions on Instrumentation and Measurement, Volume 62 Issue 3. March 2013.

[3] YingjunJi , Shixin Liu , Mengchu Zhou , Ziyan Zhaoa , XiwangGuo, Liang Qi: "A machine learning and genetic algorithm-based method for predicting width deviation of hot-rolled strip in steel production systems" .

[4] Nitesh V. Chawla , Kevin W. Bowyer, Lawrence O. Hall W. Philip Kegelmeyer : "SMOTE: Synthetic Minority Over-sampling Technique" Journal of Artificial Intelligence Research 16 (2002) 321–357. June 2002.

[5] Shai Shalev-Shwartz, Shai Ben-David. Understanding Machine Learning: From Theory to Algorithms, page 360. Cambridge University Press 2014.