# HIGH PRECISION PREDICTION MODEL OF MATERIAL PROPERTIES BY LOCALLY WEIGHTED REGRESSION METHOD WITH DIMENSION REDUCTION

Hiroyuki TAKAGI [1a], Shuji KUYAMA [1b], Hiroyasu SHIGEMORI [1c]

[1]JFE Steel Corporation, Kawasaki, Japan
[a]h-takagi@jfe-steel.co.jp, [b]s-kuyama@jfe-steel.co.jp, [c]h-shigemori@jfe-steel.co.jp

**Abstract**

To reduce production cost of steel products, it is necessary to control quality stabilization for preventing nonconformance of material properties. The quality stability is improved by constructing a highly accurate prediction model of material properties and incorporating it into the quality control system. In this paper, we aim to establish the highly accurate model of the full-length and full-width prediction of material properties. A locally weighted regression model with dimension reduction has been newly proposed and introduced into the conventional coupled model which consists of a just-in-time statistical model and a metallurgical microstructure model. In the proposed model, the data are projected from the input variable space to a dimension reduction space by a feature extraction method and a regression model with a local weighted method is created from the data near a query point in the projection space. It can properly fit complex process data which have non-linear and co-linear characteristics by mutual complement of these methods. As the result of introducing the proposed model, the standard deviation of an estimated error of tensile strength has decreased remarkably, therefore the target of the estimated accuracy for practical operation has been achieved. The new model has enabled to increase ratio of keeping the material properties within the allowable limits. Hereafter, practical operation by the developed technology has been promoted.

**Keywords:** Quality stabilization; material properties; full-length and full-width prediction; locally weighted regression model with dimension reduction.

## 1. INTRODUCTION

In order to reduce the manufacturing cost of steel products, quality control is important to prevent incompatibility of product specifications. However, there may be risks that quality incompatibility products leave a factory and enter the market because quality is not guaranteed for the full-length and the full-width of all products by the conventional sampling inspection with few samples.

Therefore, a high precision quality prediction technology in the full-length and the full-width of the steel products has been required. In this paper, we aim to establish the highly accurate model of material properties, and study results of accuracy verification of the prediction model are described.

## 2. TECHNICAL OVERVIEW OF PREDICTION OF MATERIAL PROPERTIES

For stable quality control of steel products, several quality prediction technologies of material properties [1,2] have been considered. Among them, as shown in **Figure 1**, the model which consists of three parts - the actual operation data collection part, the metallurgical parameters estimation part and the just-in-time statistical material properties prediction part - has been developed. Relatively high prediction accuracy has been expected by this model because of using both metallurgical knowledge and statistical knowledge. The prediction procedure (1) ~ (5) by this model is described below:

1) Actual operation data which include size, ingredients and temperature history of each product are sampled.
2) Metallurgical parameters are estimated with the sampled operation data as input.

3) The sampled operation data and the estimated metallurgical parameters are input to the model which statistically predicts material properties from the past operation data and the metallurgical parameters stored in a database, and a material properties prediction calculation is executed. Then prediction results of material properties are obtained as output.
4) The sampled operation data and the estimated metallurgical parameters are recorded in the database.
5) Actual material properties data are recorded in the database associated with the recorded operation data and the metallurgical parameters.

The actual operation data including the two-dimensional surface temperature of the products by the thermometer installed during the manufacturing process are collected and used for the estimation of the metallurgical parameters and the prediction of the material properties.

The metallurgical parameters estimation part estimates internal temperature of the product during manufacturing process and changes in the particle size, the phase fraction and the cumulative strain of the metal structure based on the actual operation data.

The just-in-time statistical prediction part predicts material properties based on the actual operation data and the metallurgical parameters of the prediction target from the large quantity of recorded operation data, the metallurgical parameters and the actual material properties data by the material tests.

In this study, sophistication of the statistical model applied to material properties prediction part has been examined. The target material property to be predicted is the tensile strength which is one of the capital quality indicators.

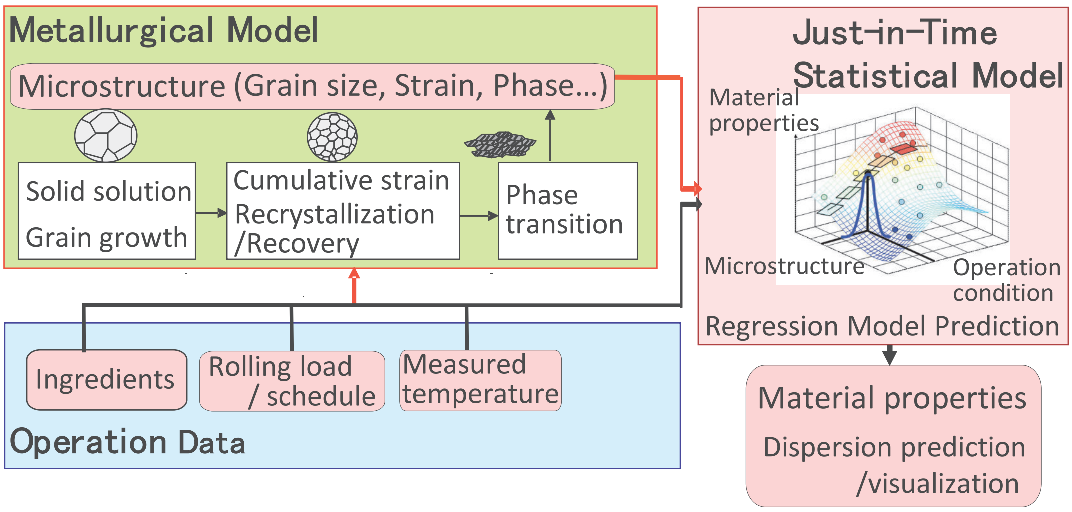


**Figure 1** Prediction model of material properties

## 3. APPLICATION OF LW-PCR MODEL TO MATERIAL PROPERTIES PREDICTION

In order to further reduce prediction errors of material properties, the more sophisticated data-science technology has been applied to the just-in-time statistical prediction part.

As the conventional prediction model, a Locally Weighted Regression (LWR) model [3] has been used in many cases. As shown in **Figure 2**, the LWR model derives a local regression formula by placing emphasis on the past actual data in the vicinity of the prediction target data. This model can deal with the nonlinearity of the entire data set as shown in **Figure 3**. On the other hand, when the diversity of data is insufficient, such as a part with small vicinity data, the prediction accuracy may decrease due to multicollinearity. It is a property that the reliability of regression decreases when there is a strong correlation between input variables.

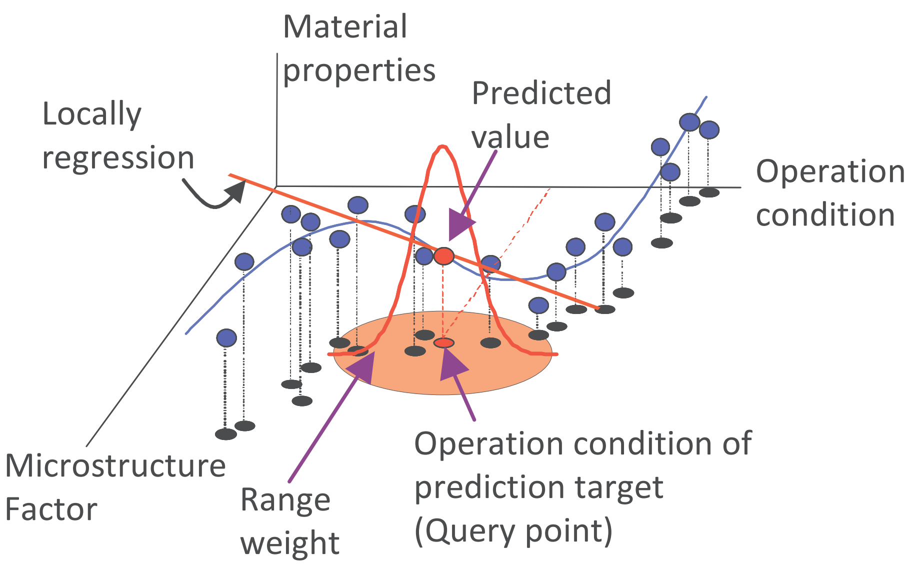


**Figure 2** Concept of LWR model

A well-known technique for avoiding the problem of multicollinearity is a Principal Component Regression (PCR) model [4]. As shown in **Figure 4**, the PCR model calculates the principal component which is a linear combination of input variables so that variance of data in the space becomes large, and calculates the regression formula from the data set on the projective space based on the principal component. On the other hand, the PCR model cannot cope with the nonlinearity of the entire data set.

Therefore, a Locally Weighted Principal Component Regression (LW-PCR) model [5] has been introduced. This model can be regarded as a model combining the LWR model and the PCR model. An outline of the LW-PCR model is shown in **Figure 5**. In the LW-PCR model, data are projected from the input variable space to the principal component space, and a local regression model is created placing emphasis on the data near the query point in the projective space. By using the LW-PCR model, multicollinearity can be avoided by the features of PCR while responding to nonlinearity by the features of LWR. They can complement each other by making use of mutual advantages, so that an improvement of prediction accuracy is expected.

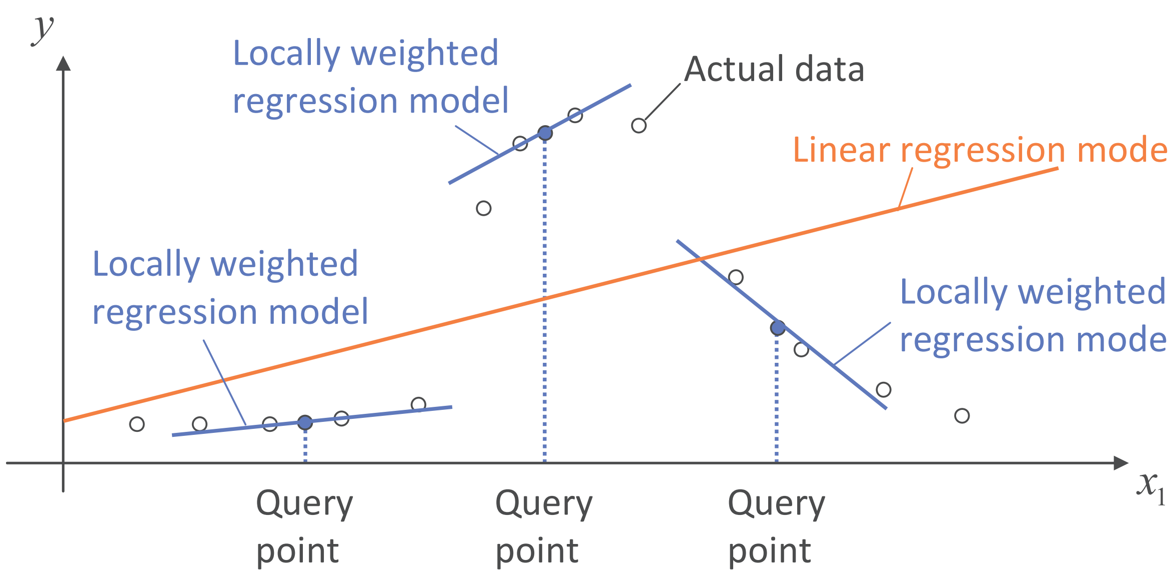


**Figure 3** Feature of LWR model

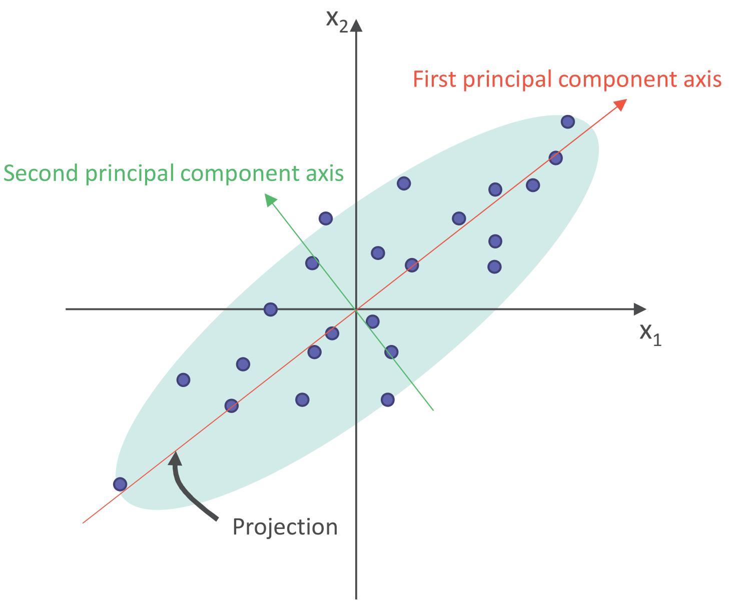


**Figure 4** Concept of PCR method

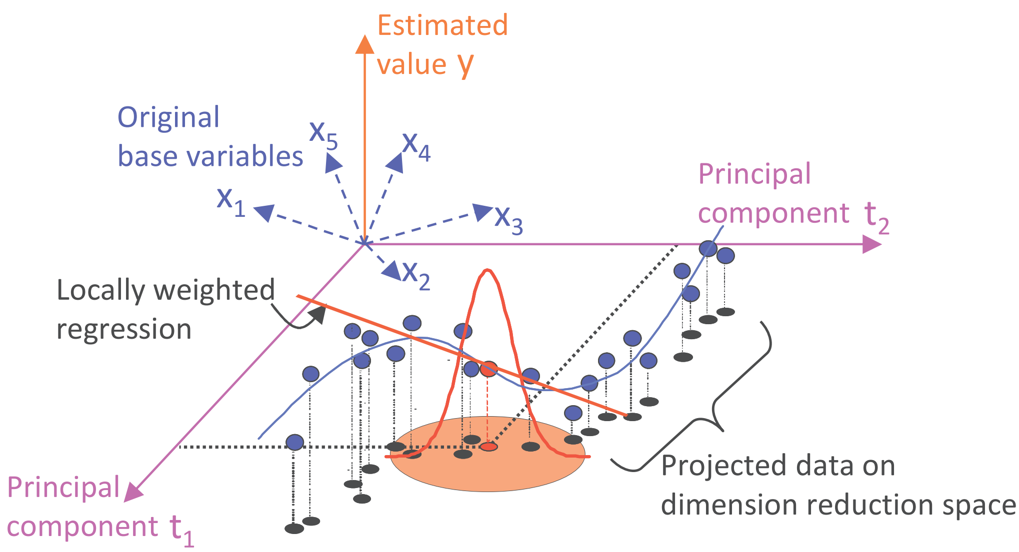


**Figure 5** Concept of LW-PCR model

## 4. MATERIAL PREDICTION ACCURACY VERIFICATION USING ACTUAL DATA

The prediction accuracy of the tensile strength using the actual data was estimated. The data was 473 samples for which the material test was conducted. Input data set for the accuracy verification include several items such as the size of the product, the amount of ingredients added in the steel, and the temperature history during production of the product as explanatory variables, and the tensile strength by actual test result of as true value. The accuracy verification was carried out with leave-one-out cross validation.

The comparison before and after the introduction of the LW-PCR model was made. As shown in **Figure 6**, the standard deviation of the prediction error with LW-PCR model was reduced compared to the conventional one. As a result, the tensile strength falls within the allowable range of the product specifications sufficiently, reaching the applicable level for the practical operation.

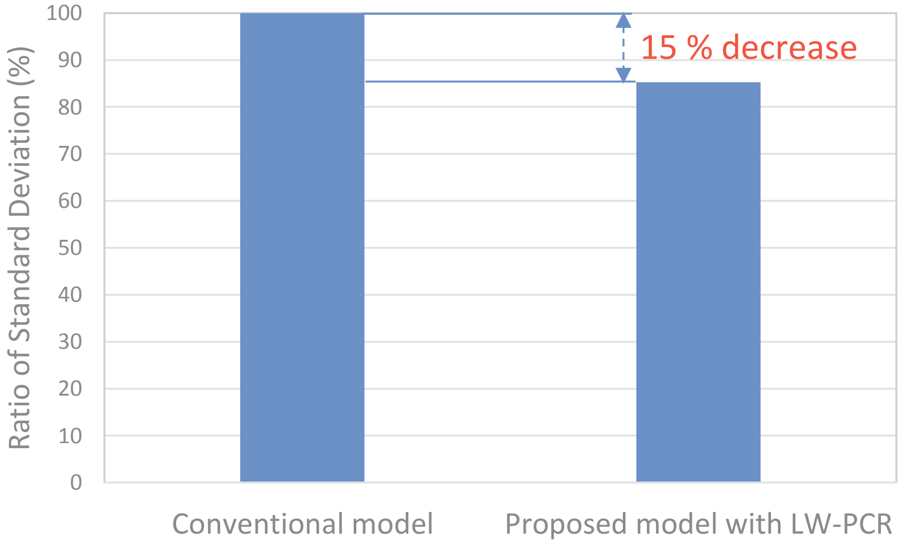


**Figure 6** Accuracy of estimated value of TS

## 5. CONCLUSION

For high quality control of steel products, high precision material properties prediction model has been developed. Using the actual operation data over the full-length and the full-width of the product, the LW-PCR has been adopted in the statistical prediction part in the material prediction model, and realized highly accurate material prediction.

## REFERENCES

[1] JUUTILAINEN, I., RONING, J., MYLLYKOSKI, L. Modelling the strength of steel plates using regression analysis and neural networks. *Proceedings of International Conference on Computational Intelligence for Modelling*, 2003, pp. 681-691.

[2] TOMOTA, Y., UMEMOTO, M., KOMATSUBARA, N., HIRAMATSU, A., NAKAJIMA, N., MORIYA, A., WATANABE, T., NANBA, S., ANAN, G., KUNISHIGE, K., HIGO, Y., MIYAHARA. M. Modelling the strength of steel plates using regression analysis and neural networks. *ISIJ International*, 1992, vol. 32, no. 3, pp. 343-349.

[3] SHIGEMORI, H., NAGAO, R., HIRATA, N., NAMBU, K., IKEDA, N., MIZUSHIMA, N., KANO, M., HASEBE, S. Quality Control for Steel Products through Locally-weighted Regression. *Journal of the Society of Instrument and Control Engineers*, 2008, vol. 44, no. 4, pp. 325-332.

[4] FUJIWARA, M., MINAMIDANI, T., NAGAI, I., WAKAKI, H. Principal Components Regression by Using Generalized Principal Components Analysis. *Journal of the Japan Statistical Society*, 2013, vol. 43, no. 1, pp. 57-78.

[5] CHARLTON, M., BRUNSDON, C., DEMSAR, U., HARRIS, P., and FOTHERINGHAM, S. Principal Components Analysis: from Global to Local. In *13th AGILE International Conference on Geographic Information Science*. Guimaraes: AGILE, 2010.