ISSN 1027-5495. Functional Materials, 29, No.1 (2022), p. 154-163.
doi:https://doi.org/10.15407/fm29.01.154 © 2022 — STC "Institute for Single Crystals"

Prediction of mechanical properties of hot rolled strips based on support vector quantile regression with adaptive QLASSO

Xiao-xia He[1,2], *Si-yu Huang*[1,2], *Xin Zhang*[1,2], *Wei-gang Li*[3]

[1]College of Science, Wuhan University of Science and Technology, 430065 Wuhan, China
[2]Hubei Province Key Laboratory of Systems Science in Metallurgical Process, Wuhan University of Science and Technology, 430065 Wuhan, China
[3]Engineering Research Center of Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, 40081 Wuhan, Hubei, China

Received January 4, 2022

The prediction of mechanical properties of hot rolled strips can be used for online dynamic control of product properties and optimal design of new steel grade. This paper proposes a prediction model of support vector quantile regression (SVQR) which combines support vector regression (SVR) with quantile regression (QR), to solve the nonlinear problems and data heterogeneity problems in modeling. First, based on the measured data collected from a hot rolling production process, the Quantile-LASSO (QLASSO) method is introduced to identify the factors affecting the tensile strength. The experimental results show that FT (reheating temperature), FET (finishing entry temperature), CT (coiling temperature), RT (roughing exit temperature), FDH (finishing delivery thickness), Si, Mn, V, Ti, NbC (niobium carbide), NbN (niobium nitride) and Cs (residual carbon) have a great influence on the tensile strength. A predictive model is then created using these influencing factors as independent variables. Compared with other models, the model proposed in this paper has better prediction accuracy, and the prediction errors of its training set and testing set are 2.42 % and 2.55 % respectively.

Keywords: prediction of mechanical properties, support vector machine, quantile regression, hot rolled strips.

Прогнозування механічних властивостей гарячекатаних смуг на основі квантильної регресії опорних векторів з адаптивним QLASSO. *Xiao-xia He, Si-yu Huang, Xin Zhang, Wei-gang Li*

Пропонується модель прогнозування квантильної регресії опорних векторів (SVQR), яка поєднує регресію опорних векторів (SVR) з квантильною регресією (QR) для вирішення нелінійних проблем та проблем неоднорідності даних при моделюванні. По-перше, на основі виміряних даних, зібраних у процесі виробництва гарячої прокатки, вводиться метод Quantile-LASSO (QLASSO) для виявлення факторів, що впливають на межу міцності при розтягуванні. Експериментальні результати показують, що FT (температура повторного нагрівання), FET (температура на вході чистової обробки), CT (температура намотування), RT (температура на виході чорнової обробки), FDH (товщина подачі чистової обробки), Si, Mn, V, Ti, NbC (карбід ніобію), NbN (нітрид ніобію) і Cs (залишковий вуглець) мають великий вплив на межу міцності при розтягуванні. Потім створюється модель прогнозування, що використовує ці фактори, що впливають, як незалежні змінні. Порівняно з іншими моделями модель, запропонована в даній статті, має кращу точність передбачення, а помилки передбачення її навчальної вибірки та перевірочної вибірки становлять 2,42 % та 2,55 % відповідно.

1. Introduction

Predicting the microstructure and mechanical properties of hot rolled steel, as a complex metallurgical frontier technology, has always been a challenge in the steel industry. Accurate prediction of the mechanical properties of hot rolled strip products is helpful in optimizing steel grade composition, reducing strip sampling, controlling mechanical properties of products and designing new products [1].

There have been many researches on predicting the mechanical properties of steel products (see Li et al. [2] and references therein). At present, the methods used for prediction modeling of the mechanical properties of hot rolled steel are divided into two types: mechanism modeling and statistical modeling. The mechanism modeling mainly uses a large number of data such as chemical composition of strips and rolling process parameters to carry out complicated mathematical calculations according to physical models. Zhang et al. [3] researched the effects of the mechanical properties of C–Mn steel weld metal with a micromechanics-based finite element method. Wang et al. [4] developed a prediction system for microstructure and properties of hot rolled strips based on the metallurgical physical model in hot rolling. At the same time, statistical modeling methods are divided into traditional statistical methods and artificial intelligence methods. Ding, Tian and Sutthaweekul [5] built a statistical linear MBN feature fusion model to predict hardness of various materials by applying multivariate linear regression (MLR). Khalaj et al. [6] used an artificial neural network method to predict the ultimate tensile strength of the API X70 steels after thermomechanical treatment. Sui and Lv [7] proposed a new extreme learning machine (ELM) model which combined ELM with the attribute reduction method to deal with the problem of high-latitude and strong coupling in the process quality parameters. Shen et al. [8] combined machine learning with artificial intelligence in physical metallurgy to obtain the best mechanical properties of ultra high strength stainless steel.

However, the above two modeling methods have several possible shortcomings. The statistical modeling method can hardly ensure the reliability and generality of the model; and the problem of the mechanism modeling method is that the mechanism of microstructure evolution is too complicated. Therefore, many scholars have combined the statistics and mechanism modeling methods to predict the mechanical properties of hot rolled strips in recent years. Li [9] proposed a global additive model for predicting mechanical properties of hot rolled strips by combining industrial data and metallurgical mechanisms, which divides the complex problem into several sub-problems, and then establishes several sub-models for each of influencing factors.

The above latest researches have improved the precision and adaptability of the model for predicting mechanical properties of hot rolled strips to certain degree, but could not consider the whole data distribution of the hot rolling process. To further improve the practicality and generalization of the built model, this paper proposed a new model combining the support vector machine (SVM) and quantile regression (QR). It takes into account the characteristics of the entire data distribution and is insensitive to outliers, which provides good reliability.

There are many factors that affect the mechanical properties of hot rolled strips, such as chemical composition of steel and process parameters. Therefore, dimensionality disasters often arise during modeling. The problem can be effectively solved by reducing the number of independent variables as much as possible while ensuring the accuracy of the model's prediction.

Wu et al. [10] used the mean impact value to select the effective process parameters which have a great influence on the mechanical properties, and then adopted a Bayesian neural network to construct a mechanical property prediction model. Yang et al. [11] selected influencing factors by the random forest algorithm and then built a mechanical property prediction model for hot rolled strips.

In order to improve the prediction accuracy of the model, this paper firstly selects the factors influencing on tensile strength using a variable selection method, and then constructs the prediction model of support vector quantile regression. The rest of the paper is organized as follows. Section 2 introduces the model of SVQR; section 3 details the variable selection method for SVQR; section 4 presents an algorithm for the parameters; section 5 describes the details and results of the experiments of the proposed model; and section 6 summarizes all the research.

2. Support vector quantile regression (SQVR)

As is known, SVM was proposed by Vapnik [12] based on the theory of statistical learning and the principle of minimizing structural risk. In SVM, instead of a nonlinear mapping to a multidimensional space, the kernel function of the inner product is used, which can well solve the non-linear problem in mathematical modeling applications. At the same time, the final SVM solution function is determined by only a few support vectors that can help understand key samples, thereby avoiding a dimensional disaster and obtaining excellent reliability. When SVM is applied to the regression problem, it is called support vector regression (SVR). The linear function of SVR in the feature space is

$$f(x) = m(x) = \omega \cdot \varphi(x) = b, \tag{1}$$

where ω denotes weights vector, b is the threshold and $\varphi(\cdot)$ is the nonlinear mapping function.

Quantile regression is proposed by Bassett and Koenker [13]. Compared with ordinary linear regression, it is insensitive to the data of outliers, has better robustness, and can handle data heterogeneity well. Meanwhile, QR takes into account the data distribution characteristics and thus can predict the mechanical properties of hot rolled strips in different quantiles.

Let $f_\tau(x) = m(x) = \omega_\tau\varphi(x) + b_\tau$ be the linear τ-th quantile in the feature space. Then from Takeuchi and Furuhashi [14], we know that the estimator for f_τ with a L_2 regularization term and Vapnik's ε-intensive tube is given by

$$\min_{\omega_\tau, b_\tau} \frac{1}{2}\|\omega_\tau\|^2 + C\sum_{i=1}^{n}\rho_\tau(y_i - f(x_i)), \tag{2}$$

where

$$\rho_\tau(u) = \begin{cases} \tau u, & u > \varepsilon, \\ 0, & |u| \le \varepsilon, \\ (\tau - 1)u, & u < -\varepsilon. \end{cases} \tag{3}$$

Setting $\varepsilon = 0.5$ leads to Vapnik's ε-intensive Support Vector Regression; and setting $\varepsilon \to 0$ and $C \to \infty$ leads to Quantile Regression for f_τ. In fact, introducing the ε-insensitive tube is somewhat unfavorable as a conditional quantile estimator, so the value of ε is usually fixed to be 0.

A combination of SVR and QR algorithms was proposed in some papers. Shim et al. [15] introduced a semiparametric method which combined SVR with QR to construct a semiparametric SVQR model. And He [16] proposed a short-term method for predicting the load power probability density using a kernel-based SVQR.

In this paper, QR and a variable selection method are introduced to SVR to predict the mechanical properties of hot rolled strips. First, as in [17], the input variables included in the parametric and nonparametric parts are assumed to be vectors z_i and x_i respectively. Furthermore, we assume that an input vector z_i has a linear effect on the response variable, while the other input vector x_i has a non-linear effect on the response variable. Then the quantile regression problem by the formulation for SVQR can be expressed as

$$\min_{\omega_\tau, b_\tau, \beta_\tau} \frac{1}{2}\|\omega_\tau\|^2 + C\sum_{i=1}^{n}\rho_\tau\left(y_i - b_\tau - \beta_\tau^T z_i - \omega_\tau^T\varphi(x_i)\right). \tag{4}$$

In order to solve the optimization problem (4), we introduce slack variables to construct Lagrange functions; then the above problems can be transformed into the following quadratic programming problems.

$$\min \frac{1}{2}\omega_\tau^T\omega_\tau + C\sum_{i=1}^{n}\left(\tau\xi_i + (1 - \tau)\xi_i^*\right)$$

subject to

$$\begin{cases} y_i - b_i - \beta^T z_i - \omega^T\varphi(x_i) \le \xi_i \\ -y_i + b_\tau + \beta^T z_i + \omega^T\varphi(x_i) \le \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases}$$

Then the estimators of the coefficient and the regression equations are obtained as follows

$$\begin{cases} \omega_\tau = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\varphi(x_i) \\ (b_\tau, \beta_\tau)^T = (U^TU)^{-1}U^T(y - K(\alpha - \alpha^*)) \\ Q_\tau(y_i|z_i, x_i) = b_\tau + \beta_\tau^T z_i + \sum_{j=1}^{n}K(x_j, x_i)(\alpha_j - \alpha_j^*) \end{cases}$$

where α_i, α_i^*, represent the optimal Lagrange multipliers, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $\boldsymbol{\alpha}^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_n^*)$.

According to the principle of support vector machine, the index set of the support vector in SVQR model is obtained by the Karush-Kuhn-Tucker condition.

$$I_{SV} = i = 1,2,3,\ldots,n|0<\alpha_i<\tau C, 0<\alpha_i^*<(1-\tau)C.$$

Here, $U = (1, z_i^T)$, and $\mathbf{y} = \{y_i|$ i ^{roman symbol л} I sub SV $\}$ and the matrix

$$\boldsymbol{K} = \begin{pmatrix} K(x_1,x_1) & K(x_1,x_2) & ,,, & K(x_1,x_n) \\ K(x_2,x_1) & K(x_2,x_2) & \ldots & K(x_2,x_n) \\ \ldots & \ldots & \ldots & \ldots \\ K(x_n,x_1) & K(x_n,x_2) & \ldots & K(x_n,x_n) \end{pmatrix}$$

where K is the kernel function of the input space, which is equal to the inner product of the sum in the feature space, and

$$K(x_s,x_t) = \varphi(x_s)^T\varphi(x_t),\ s,\ t \ \text{®}\ I_{SV}.$$

In practical application, the linear kernel function, the Gaussian kernel function, and the polynomial kernel function are often used. The linear kernel function and the Gaussian kernel function are respectively:

$$K(x,z) = x^T z,$$
$$K(x,z) = \exp\left(-\frac{(x-z)^2}{2\sigma^2}\right).$$

3. Variable selection for SVQR

The factors influencing on steel mechanical properties are numerous and have complex interactions. In this subsection we introduce a QLASSO method, which can select the influencing factors of tensile strength to improve the prediction accuracy.

3.1 Variable selection based on QLASSO

To identify significant variables in the linear part, we propose an adaptive QLASSO penalized method. Zhao et al. [18] proposed a variable selection procedure for models with quantile variable coefficients. We extend the procedure to the SVQR situations. The corresponding loss function is given by

$$L(b,\beta,\gamma) = \tag{5}$$
$$= \frac{1}{2}\gamma^T K\gamma + C\sum_{i=1}^{n}\rho_{\tau,\delta}(y_i - b - \beta^T z_i - K_t\gamma) + \sum_{j=1}^{p}\lambda_j\frac{\|\beta_j\|^2}{\|\hat{\beta}_{j,\lambda}^{(m)}\|},$$

where $\rho_{tau,\delta}$ will be defined in section 4. Then, the minimizer of Eq. 5 is given by β_λ^{m+1}, which can be given by

$$\hat{\beta}_\lambda^{(m+1)} = \tag{6}$$
$$= (Z^T WZ + D^m/C)^{-1}(Z^T Wy - Z^T W1b - Z^T WK\gamma),$$

where $\mathbf{D}^{(m)}$ is a $d \times d$ diagonal matrix with j-its diagonal component given by $\frac{\lambda_j}{\|\hat{\beta}_{j,\lambda}^{(m)}\|}$. Denote the limit values of $\hat{\beta}_\lambda^{(m+1)}$ by β_λ. Since the estimate $\hat{\beta}_\lambda$ is based on a combination of the regression quantile and the adaptive LASSO method, we refer to it as a Quantile-LASSO(QLASSO) estimate.

3.2 Tuning parameter

We follow the ideas of Zhao [18], Zou [19], Wang [20] and Wang [21] to simplify the tuning parameters as

$$\lambda_j = \frac{\lambda_0}{(n^{-1/2}\|\tilde{\beta}_j\|)^\gamma} \text{ with } \gamma = \frac{3}{2},$$

where $\tilde{\beta}_j$ is the j-th column of the unpenalized estimate $\tilde{\beta}$. Consequently, the original d-dimensional problem about $\lambda \in R^D$ becomes a univariate problem about $\lambda_0 \in R^1$, where λ_0 can be selected according to the following BIC-type criterion

$$BIC_\lambda = \frac{1}{n}\log(RSS_\lambda) + df_\lambda \times \frac{\log(n)}{n}, \tag{8}$$

where $0 \le df_\lambda \le d$ is simply the number of nonzero coefficients identified by β_λ; and RSS_λ is defined as

$$RSS_\lambda = \frac{1}{n}\sum_{i=1}^{n}\rho_{\tau,\delta}(y_i - b - \beta^T z_i - K_t\gamma_t).$$

Then the tuning parameter can be obtained as

$$\hat{\lambda} = arg\min_{\lambda} BIC_\lambda.$$

4. Solution algorithm for SVQR with QLASSO

In this subsection an iterative reweighted least squares (IRWLS) procedure proposed by Shim [15] is applied to estimate the parameters of the model SVQR, in which b_τ, β_τ and $\gamma_\tau = \alpha - \alpha^*$ are estimated simultaneously. To overcome the non-differentiability of ρ_τ at zero, we employ the modified check function $\rho_{\tau,\delta}$ (Nychka [22]) to get a good approximate solution to the original optimal problem. Explicitly, the function $\rho_{\tau,\delta}$ differs from ρ_τ only in the small enough interval $(-\delta, \delta)$ such as

$$\rho_{\tau,\delta}(r) = \tau\frac{r^2}{\delta}I(r \geq 0) + (1-\tau)\frac{r^2}{\delta}I(r<0). \quad (9)$$

Now Eq. 5 becomes the one of obtaining b_τ, β_τ, γ_τ to minimize

$$L(b,\beta,\gamma) = \quad (10)$$
$$= \frac{1}{2}\gamma^T K\gamma + C\sum_{i=1}^{n}\rho_{\tau,\delta}(y_i - b - \beta^T z_i - K_t\gamma) + ,$$
$$+ \sum_{j=1}^{p}\lambda_j\frac{\|\beta_j\|^2}{\|\hat{\beta}_{j,i}^{(m)}\|} \text{ for } \tau \to (0,1)$$

where $\mathbf{K}_i$ is the i-line of the matrix $\mathbf{K}$. Taking partial derivatives of Eq. 10 with regard to b_τ, β_τ, γ_τ respectively, leads to the optimal values to be the solution to

$$\begin{pmatrix} b_\tau^{(l+1)} \\ \beta_\tau^{(l+1)} \\ \gamma_\tau^{(l+1)} \end{pmatrix} = \begin{pmatrix} 1^T WI & 1^T WZ & 1^T WK \\ Z^T WI & Z^T WZ + D/C & Z^T WK \\ WI & WZ & WK + I/C \end{pmatrix}^{-1} \quad (11)$$

where $\mathbf{Z}$ is a $n \times p$ matrix consisting of z_i^T for $i = 1,2, \ldots,n$. $\mathbf{W}$ is a diagonal matrix with the diagonal element w_{ii} obtained from the derivative of the modified check function as

$$w_{ii} = \begin{cases} 2\tau/\delta, & 0<r_i<\delta \\ \tau/r_i, & r_i>\delta \\ 2(1-\tau)/\delta, & -\delta<r_i<0' \\ (\tau-1)/r_i, & r_i<-\delta \end{cases} \quad (12)$$

where $r_i = y_i - b_\tau - \beta_\tau^T z_i - K_i\gamma_\tau$.

The solution to Eq. 10 cannot be obtained in a single step, since the matrix $\mathbf{W}$ contains the parameters b_τ, β_τ, γ_τ therein. Thus, we apply the IRWLS procedure as following

0. Set the initial values $b_\tau^{(0)}$, $\beta_\tau^{(0)}$, $\gamma_\tau^{(0)}$.

1. Calculate $\mathbf{W}$ in Eq. 12 using δ, τ and $r_i = y_i - b_\tau^{(l)} - (\beta_\tau^{(l)})^T z_i - K_i\gamma_\tau^{(l)}$.

2. Obtain $b_\tau^{(l+1)}$, $b_\tau^{(l+1)}$, $\gamma_\tau^{(l+1)}$ from

$$\begin{pmatrix} b_\tau^{(l+1)} \\ \beta_\tau^{(l+1)} \\ \gamma_\tau^{(l+1)} \end{pmatrix} = \begin{pmatrix} 1^T WI & 1^T WZ & 1^T WK \\ Z^T WI & Z^T WZ & Z^T WK \\ W1 & WZ & WK + I/C \end{pmatrix}^{-1} \quad (13)$$

3. Loop iteration until convergence.

5 Simulation results and discussion

5.1 Data collection and evaluation criterion

Table 1. Composition of the data set

Variable	Minimum	Maximum	Average
FT	1127	1265	1211
FET	910	1103	1010
FDT	806	932	857
CT	402	636	568
RT	983	1124	1073
FDH	1.49	17.338	3.823
Si	0.005	0.199	0.03173
Mn	0.377	1.73	1.04
P	0.0054	0.09	0.0153
S	0.0003	0.0112	0.0046
Cu	0	0.5043	0.0131
Ni	0	0.4014	0.0064
Cr	0.0084	0.0413	0.01479
Mo	0	0.006	0.001224
V	0	0.0032	0.00023
Ti	0	0.0097	0.00125
B	0	0.0012	0.00003
NbC	0.011	0.095	0.033
NbN	0.0147	0.068	0.036
Cs	0.04048	0.15146	0.069
Ns	0.000032	0.00095	0.00038
Ts	342	770	532.92

The data applied in the current experiments are from a hot rolling production line of a large-scaled steel company in China. The data set is obtained after calculating the mass fraction of niobium carbide, niobium nitride, residual carbon and residual nitrogen in the rolling process. It contains 21 influencing factors and one response variable called tensile strength (Ts), which are listed in Table 1.

In Table 1, FT is the reheating temperature; FET is the finishing entry temperature; FDT is the finishing delivery temperature; CT is the coiling temperature; RT is the roughing exit temperature; and FDH is the finishing delivery thickness.

To evaluate the model properties, the evaluation indicators include mean absolute percentage error (MAPE) and root mean square error (RMSE), and the calculation formula is

$$MAPE = \frac{1}{q}\sum_{i=1}^{q}\frac{|f_i - m_i|}{m_i}, \quad (14)$$

$$RMSE = \sqrt{\frac{1}{q}\sum_{i=1}^{n}(f_i - m_i)^2},$$

where f_i is the predicted value of the model, m_i is the measured value of the sample and q is the sample size of the data.

5.2 Simulation results

The simulation experiments are carried out on a personal computer with a Windows 10 64 bit operating system, which is equipped with an Intel (R) Core (TM) i7-9700k CPU processor and two 16 GB memory modules (the total RAM is 32 GB); and a 2080Ti GPU are also used. R-3.6.1 and Python 3.5.3 are also utilized in the experiments.

5.2.1 Variable selection based on QLASSO

This study takes Ts (tensile strength) as the response variable, and all the predictors as the X-variables. In order to use the QLASSO method to analyze the data, all the predictors are linearly transformed to [0, 1] before applying the proposed model. In the empirical study, the penalty parameters were set as 200, δ as 10 and τ as 0.5, respectively, and the optimal shrinkage parameter selected by the BIC criterion is given by $\lambda_0 = 0.85$. The variable selection results are given in 0.

QLASSO compresses variables to reduce the coefficients of irrelevant variables or variables with small correlation to zero, and therefore the variables whose coefficient is not zero is considered to be the final results of the variable selection. According to the variable selection results in 0, FT, FET, CT, RT, FDH, Si, Mn, V, Ti, NbC, NbN and Cs are taken as the main influencing factors.

5.2.2 Prediction with different kernel functions

By taking the selected variables as the independent variables of the model and combining with the support vector quantile regression model, the mechanical properties of hot rolled strips can be predicted. To choose an appropriate kernel function for the SVQR, this study compares the accuracy of the linear and Gaussian kernel functions.

A support vector quantile regression model is established as follows

$$Q_\tau(y|z_i,x_i) = b_\tau + \beta_\tau^T z_i + \sum_{i=1}^{n} K(x_i,x)(\alpha_i - \alpha_i^*).$$

Table 2. Variable selection results for the strip rolling data

	coefficient
Intercept	215.6
FT	17.22
FET	−3.33
FDT	0
CT	−45.23
RT	14.05
FDH	−60.58
Si	20.93
Mn	84.75
P	45.96
S	0
Cu	0
Ni	0
Cr	0
Mo	0
V	−6.58
Ti	−26.73
B	0
NbC	69.46
NbN	52.22
Cs	82.02
Ns	0

The specific parameters of the above model are obtained by the IRWLS procedure. The estimation results are shown in 0.

According to the estimation results in 0, the coefficient estimation results with Gaussian kernel function show FT, RT, Si, V, NbC, NbN, Cs and Ts are positively correlated, and FET, CT, FDH, Mn, Ti and Ts are negatively correlated. Meanwhile, the coefficient estimation results with a linear kernel function show that FT, FET, V and Ts are positively correlated, and CT, RT, FDH, Si, Mn, Ti, NbC, NbN, Cs and Ts are negatively correlated.

The above two models were tested and their prediction accuracy results are shown in 0. It is obvious that the prediction accuracy with Gaussian kernel function is better than that with linear kernel function. Therefore, Gaussian kernel function is applied in the subsequent experiments.

5.2.3 Prediction with training and test sets

The experiment data were divided randomly into a training set and a testing set, the percentages of which are 70 % and

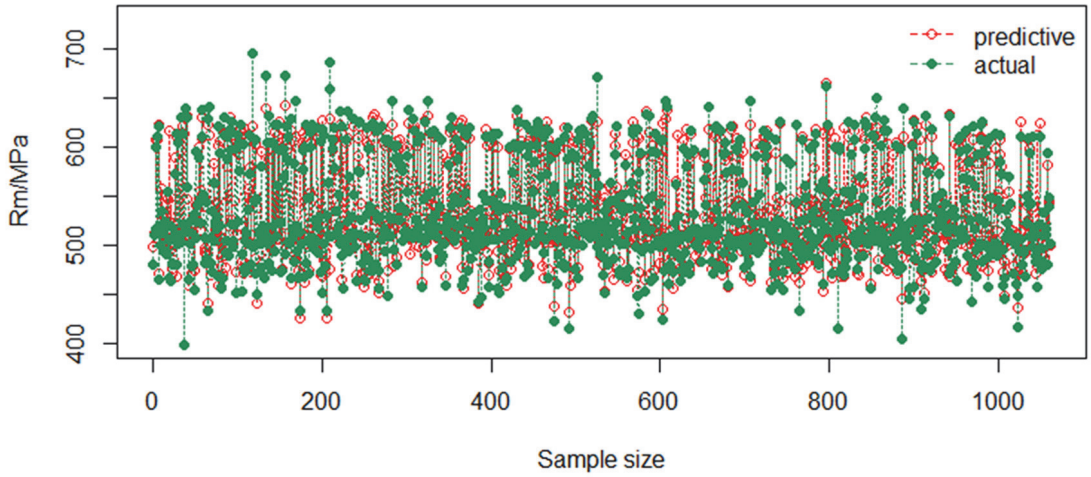

Fig. 1. Comparison of the predicted and measured tensile strengths.

30 % respectively. The model parameters were obtained from the training set, and then the built model was used to predict the mechanical property of hot rolled strips on the testing set. The results show that the prediction error of the training set is 2.42 %, and the prediction error of the test set is 2.55 %. The predicted and actual tensile strength values from the test set are compared in Fig. 1, which shows little difference between the prediction and the measured values.

The predicted and measured tensile strength results from the training set and the testing set are shown in Fig. 2, which shows that the predicted value is highly correlated with the measured value both on the training set and the testing set, and the deviations between the predicted and measured values of the tensile strength are in a reasonable range. In addition, the calculation process does not need manual intervention or correction, and the new model has a higher prediction accuracy.

5.3 Comparison of different models and variable selection methods

To further verify the validity of the model, the results of the proposed model were compared with some other methods: SR, RR and SVR. At the same time, three different variable selection methods including QLASSO, forward selection (FS) and random forest (RF) were tested in this work to make a comparison. Among them, RF is a supervised learning algorithm which core idea is to combine the prediction results of many decision trees and obtain the final prediction by the majority voting method (for the classification problem) or the result averaging method (for the regression problem). As a result, the advantage of RF is that it can be used for regression and classification tasks, and it is easy to see the relative importance of model input factors. The forward selection method is an independent variable qselection method for regression models. Based on this method, the variables with a large influence on output can be selected from the data sets of many variables, thereby reducing the couplings between the variables. The final input variable of the FS model is the important influencing factor. Based on the same data, the selected factors for QLASSO (QL), FS, RF are shown in 0; and twelve different results from different methods are shown in 0.

Table 3. Estimates of the coefficients with different kernel functions

Estimates	Gaussian kernel function	Linear kernel function
Intercept	468.53	583.16
FT	8.97	148.24
FET	−5.29	182.63
CT	−62.56	−1.51
RT	22.98	−206.21
FDH	−72.06	−65.59
Si	31.85	−7.35
V	80.06	319.29
Mn	−0.06	−138.31
Ti	−37.12	−178.85
NbC	85.79	−51.55
NbN	46.90	−74.86
Cs	90.68	−109.48

From the horizontal analysis of 0, the variables selected by the QLASSO method

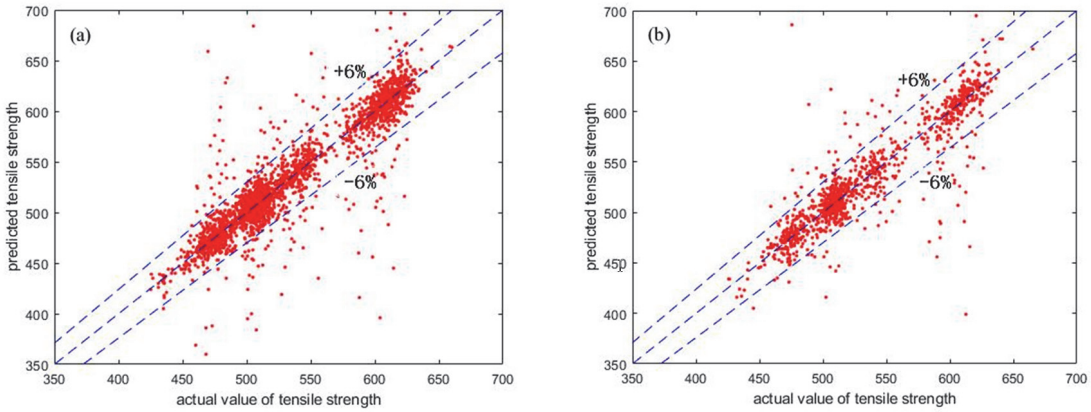

Fig.2. Predicted and measured tensile strengths. (a) Training set; (b) Test set.

Table 4. Prediction accuracy results of the two models

Predictive model	RMSE, MPa	MAPE, %
Gaussian kernel function	21.45	2.43
Linear kernel function	35.68	15.42

Table 5. Composition of influencing factors of different methods

	Influence factors
QL	FT, FET, CT, RT, FDH, Si, Mn, V, Ti, NbC, NbN, Cs
FS	FT, FET, CT, RT, FDH, Si, Mn, P, V, NbC, NbN, Cs
RF	FT, FET, RT, FDH, Si, Mn, P, S, NbC, NbN

are used to establish RR, SR, SVR, and SVQR models respectively. The model established by SVQR is the best, considering the training set or the test set. Similarly, using the variables selected by the RF and FS methods, the corresponding RR, SR, SVR, SVQR models are established, and the accuracy of the model established by SVQR is the best. From the above analysis, the variables selected from the three methods were found; the accuracy of the SVQR model was relatively high. Regarding the longitudinal analysis, in the SVQR model, the model selected by the QLASSO method is the most accurate. And in regard to the overall analysis, the accuracy of the SVQR model built based on the variables selected by the QLASSO method is the highest compared with the accuracy of the models established by other methods; this shows that the proposed model (SVQR) has the better prediction accuracy for the mechanical properties of hot-rolled strips compared with other methods.

SR (stepwise regression) and RR (ridge regression) are new linear regression models based on linear regression transformation, without considering the non-linear relationship of the model; while the SVQR model considers the linear relationship of the model as well as the non-linear relationship of the model. In practical application, there are both linear and non-linear relationships between variables. In such a context, compared with SR and RR, SVQR is more comprehensive.

In combining with the QR method, the prediction accuracy of SVQR is improved compared to SVR. This is mainly due to the fact that quantile regression is a very robust estimation method and the estimation results of the model will not be affected by abnormal data. In addition, QR does not require any assumptions of error distribution.

But, according to the results shown in 0, there is no significant difference in the accuracy between SVR and SVQR. The main reason is that compared with ordinary linear regression, quantile regression can well deal with data heterogeneity, is not sensitive to outliers, and is more "robust". And, the data in this experiment basically have no heterogeneous data and extreme values, so the results of the two models are not much different. At the same time, it is found that there is a strong linear relationship between the independent variables and the dependent variables. Although there is a non-linear part in the SVQR model, the linear part is still the main part of prediction.

QLASSO compresses variables to reduce the coefficients of irrelevant variables or variables with small correlation to zero so

Table 6. Prediction errors of different methods

		RR		SR		SVR		SVQR	
		M	R	M	R	M	R	M	R
		AP	M	AP	M	AP	M	AP	M
		E	SE	E	SE	E	SE	E	SE
QL	T	2.7	23	2.7	23	2.4	22	2.4	22
	ain	42	67	53	70	32	93	22	83
	T	2.6	23	2.6	23	2.5	23	2.5	23
	st	66	18	82	23	69	29	50	18
	Tr	2.6	23	2.6	23	2.4	22	2.4	22
	tin	91	56	99	57	48	94	40	83
FS	Te	2.6	23	2.6	23	2.5	23	2.5	23
	st	33	14	40	17	90	35	54	17
	Tr	2.6	23	2.6	23	2.4	23	2.4	23
	tin	94	67	99	67	72	20	67	20
RF	TE	2.6	23	2.6	23	2.5	23	2.5	23
	st	59	28	65	29	80	21	61	14

as to filter variables. FS and RF, by sorting the degree of influence of factors, finally determine to select the factors with a large degree of influence as variables. If the variables are not irrelevant, but are in the last position in the ranking of the degree of influence of factors, they will be ignored or even excluded. In the experiment in this study, Ti was deleted in the methods of FS and RF, which may be the main reason for the poor effect of these two methods compared with the QLASSO method.

6. Conclusions

In this paper we propose a SVQR approach based on adaptive QLASSO, a modeling method combining SVM and QR, for the prediction of mechanical properties of hot rolled strips. SVM uses an inner product kernel function instead of nonlinear mapping to high-dimensional space, which can well solve nonlinear problems in mathematical modeling applications. QR considers data distribution, is insensitive to data outliers and has good robustness, which can predict the mechanical properties of strips under different quantile points. So, the proposed SVQR can perform better than the ordinary mean regression models.

To improve the prediction accuracy, the QLASSO method is used to select influencing factors. And the BIC criterion is employed to select the appropriate adjustment parameters. The compression of the variable is used to reduce the coefficients of the irrelevant variables or the less relevant variables to zero, and directly select the correlation with the dependent variable. This study finally selects 12 variables including FT, FET, CT, RT, FDH, Si, Mn, V, Ti, NbC, NbN and Cs as the main influencing factors of the response variables. Then the selected variables are used to establish the SVQR model.

To evaluate the proposed model, four different models including SR, RR, SVR and SVQR, and three different variable selection methods including QLASSO, FS and RF are carried out to compare the prediction effect. The model is verified by field data, which shows the prediction errors of the proposed model are the smallest among these models. The prediction errors of its training set and testing set are 2.42 % and 2.55 % respectively. The results show that the proposed model has better prediction accuracy than the other models, and the model has a strong generalization. This study provides a new approach to predict the mechanical property of hot rolled strips.

Acknowledgements. The paper is supported by National Social Science Foundation of China (Grant No.11201356, 51774219), Educational Commission of Hubei Province of China (Grant No.B2020013), and the Hubei Province Key Laboratory of Systems Science in Metallurgical Process (Wuhan University of Science and Technology) (Grant No. 201715).

References

1. W.G.Li, W.Yang, Y.I.Zhao et al., *Journal of Iron and Steel Research International*, **26**, 230 (2019).

2. Wei-gang Li, Lu Xie, Yun-tao Zhao et al., *Journal of Iron and Steel Research International*, ???? (2020).
3. C.Z.Zhang, B.M.Gong, C.Y.Deng, D.P.Wang, *Materials Science & Engineering A*, **685**, 310 (2017).
4. L.Wang, D.Tang, Y.Song, *Iron and Steel*, **11**, 73 (2016).
5. S.Ding, G.Y.Tian, R.Sutthaweekul, *NDT & E International*, **107**, 102 (2019).
6. G.Khalaj, T.Azimzadegan, M.Khoeini, M.Etaat, *Neural Computing and Applications*, **23**, 2301 (2013).
7. X.Y.Sui, Z.M.Lv, *The International Journal of Advanced Manufacturing Technology*, **85**, 1395 (2016).
8. C.G.Shen, C.C.Wang, X.L.Wei et al., *Acta Materialia*, **179**, 201 (2019).
9. W.G.Li, W.Yang, Y.T.Zhao, H.F.Hu, *Journal of Iron and Steel Research*, **30**, 302 (2018).
10. S.W.Wu, Z.Y.Liu, X.G.Zhou, N.A.Shi, *Journal of Iron and Steel Research*, **28**, 1 (2016).
11. W.Yang, W.G.Li, Y.T.Zhao et al., *Iron and Steel*, **53**, 44 (2018).
12. V.N.Vapnik, The Nature of Statistical Learning Theory, Springer, New York (1999).
13. G.Bassett, R.Koenker, *Econometrica*, **46**, 33 (1978).
14. I.Takeuchi, Q.V.Le, T.D.Sears, A.J.Smola, *J. Mach Learn Res*, **7**, 1231 (2006).
15. J.Shim, Y.Kim, J.Lee, C.Hwang, *Comput. Stat.*, **27**, 685 (2012).
16. Y.Y.He, R.Liu, H.Y.Li et al., *Applied Energy*, **185**, 254 (2017).
17. Shim Jooyong, Kim Yongtae, Lee Jangtaek, Hwang Changha, *Comput. Stat.*, **27**, 685 (2012).
18. Zhao Weihua, Zhang Riquan, Lv Yazhao, Liu Jicai, *Journal of the Korean Statistical Society*, **42**, 343 (2013).
19. H.Zou, *Journal of the American Statistical Association*, **476**, 1418 (2006).
20. H.Wang, R.Li, C.Tsai, *Biometrika*, **3**, 553 (2007).
21. H.C.Wang, Y.C.Xia, *Journal of the American Statistical Association*, **486**, 747 (2009).
22. D.Nychka, G.Gray, P.Haaland et al., *J.Am. Stat.Assoc.*, **90**, 1171 (1995).