CHEMICAL ENGINEERING TRANSACTIONS

VOL. 51, 2016

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Tichun Wang, Hongyang Zhang, Lei Tian

**ISBN** 978-88-95608-43-3; **ISSN** 2283-9216

DOI: 10.3303/CET1651050

# Determination and Evaluation of Affecting Factors of Raw Material Cost Based on Rough Set Theory

Yajun Wang*, Zhangxuan Yin, Huijie Yan, Shuangjiao Fan, Junli Shi, Maojun Zhou

Department of Industrial Engineering, Dalian Polytechnic University, Dalian
wangyajun2004@hotmail.com

Effective cost control in electric arc furnace steelmaking relies on accurate analysis of factors that affect raw material cost and the causes of cost variation. This paper proposes a method to evaluate the importance of each factor affecting raw material cost. With the importance of each factor, the deterministic and nondeterministic factors can be identified by solving a minimal reduction of a decision table. Advice on further improvements for steelmaking production can then be provided. The simplicity and effectiveness of this method is illustrated by an example at the end of this paper. This method can be applied to quantitative analysis of factors that affect raw material cost in steelmaking.

## 1. Introduction

In steel production, raw material cost is defined as the amount of iron and ferroalloy required to produce one ton steel, which is the major technical economic index in steelmaking. Approximately 80% of the total cost in steelmaking is raw material cost. This index reflects the effectiveness of cost control in steelmaking. The evaluation method of analyzing the importance of each factor affecting raw material cost is the most critical part of cost control, because it will directly affect production management. Methods based on rough set theory as proposed in references. References (Cheng et al., 2008) are more scientific and objective than conventional methods and more feasible than neural network methods (Zhang and Liu, 2004). To meet the demands for cost control in steelmaking industry, the paper proposes an importance evaluation method from the perspective of knowledge discovery for affecting factors raw material cost in steelmaking based on real production and raw material cost data and it establishes a raw material cost analysis and decision system. By decertification and simplification of real production data, major affecting factors raw material cost and their importance can be determined, revealing the causes of cost variation and providing a basis for cost control decisions.

## 2. Rough set theory

Rough set theory was developed by Polish scientist Pawlak in the 1980s while studying the logical characters of information system, and it is becoming more and more popular in recent times (Pawlak,1991 and 1982). Rough set theory has been applied to areas such as machine learning, knowledge discovery, data mining, decision support and analysis, pattern reorganization, and intelligent control (Pawlak and Miao, 2008 and 2000). Related definitions in rough set theory are listed below:

Definition 1: Knowledge representation system and decision making system. In rough set theory, a knowledge representation system can be represented as, where $U$ is a non-empty, finite set of objects called the Universe, $A=C\cup D$ is a non-empty finite set of attributes which fulfills $C\cap D\neq\Phi$, Subsets $C$ and $D$ are sets of conditional attributes and decision attributes, respectively. $V=\cup_{a\in A} Va$ is the set of values of the attributes. $Va$ represents the possible values of $a\in A$ which is the range of $a$. $f: U\times A\rightarrow V$ is an information function that defines the attribute values of every object in $U$, i.e., for $\forall a\in A$, $x\in U$, and $f(x, a)\in Va$.

A knowledge representation system can be represented as *a* 2-dimensional table in which objects are presented in rows and attributes are presented in columns. The attribute value of each object is shown in the corresponding cell $f(x,a)$ to generate a decision table.

Please cite this article as: Wang Y.J., Yin Z.X., Yan H.J., Fan S.J., Shi J.L., Zhou M.J., 2016, Determination and evaluation of affecting factors of raw material cost based on rough set theory, Chemical Engineering Transactions, 51, 295-300 DOI:10.3303/CET1651050

Definition 2: Indiscernibility. In the Decision Table *T= (U, A)*, for a subset $P \subseteq A$, indiscernibility is defined as:

$$IND(P) = \{(x, y) \in U \times U \mid f(x, a) = f(y, a), \forall a \in P\} \tag{1}$$

If $(x,y) \in IND(P)$, $x$ and $y$ are indiscernible from each other by $P$. Indiscernibility is an equivalence relation in $U$. Thus owing to indiscernibility with respect to $P$, $U$ can be classified into $k$ equivalent classes, $x_1, x_2, \ldots, x_k$, which is noted as: $U/IND(P)=\{ x_1, x_2, \ldots, x_k \}$.
Definition 3: Set approximation. For any object set $X \subseteq U$ and attribute set $P \subseteq A$ the lower and upper approximations of $P$ are defined, respectively, as:

$$\underline{apr}_p(X) = Y\{Y \in U / IND(P); Y \in X\} \tag{2}$$

$$\overline{apr}_p(X) = Y\{Y \in U / IND(P); Y \cap X \neq \Phi\} \tag{3}$$

For attribute set $P$, equation (3) is the complete set of objects that can be unambiguously classified as belonging to $X$.
The equation (4) shows all objects that are possibly in $X$, which describes the minimum possible range.
Definition 4: Attribute dependency, attribute importance, and attribute redundancy. A rough set identifies the importance of an attribute by attribute dependency and filters those redundant attributes according to dependency. For attribute set $P$, $R \subseteq A$, the dependency of $P$ on $R$ is:

$$\gamma_R(P) = \frac{|POS_R(P)|}{|U|} \tag{4}$$

$$POS_R(P) = \bigcup_{X \in U / IND(P)} \underline{arp}_R(X) \tag{5}$$

$POS_R(P)$ is the positive region (lower approximation) of $R$ in $U/IND(P)$.
Definition 5: Core reduction. $B \subset C$ is a reduction of information system $S$, only if $POS_B(D) = POS_C(D)$, and every element in B is necessary for D. This reduction is noted as *RES (B, D)*.
If $\forall X \subseteq U/IND(D)$, $POS_C(X) = POS_B(X)$ and there is no B' such that $B' \subset B$ and $POS_C(X) = POS_{B'}(X)$, then $B$ is a minimal reduction of the information system.
The common set of all reductions of $C$ in decision table $D$ is called the core of $C$, which is denoted as

$$CORE(C, D) = \{a \in C : POS_C(D) = POS_{C-\{a\}}(D)\} \tag{6}$$

The core can be used as the basis data for finding the minimal reduction.

## 3. Affecting factors of steelmaking raw material consumption cost

### 3.1 Production process of steelmaking

UHP EAF is a high temperature, multiphase, fast metallurgical process (Zhang,2007), there are many the variables involved in the whole process, so a reasonable cost analysis model for steelmaking to increase productivity, reduce costs and improve product quality has important significance. In order to establish a reasonable cost analysis model, it must be first to analyze the system for the steelmaking process. For example, 3Cr13 stainless steel production process is divided into ingredients, smelting furnace, AOD(Argon-Oxygen Decarburization, AOD) refining, LF(Ladle Furnace, LF) refining, VD(Vacuum Desulphurization, VD) vacuum furnace, 150 square casting, slow cooling, transfer and other operations processes in figure 1.
After supplies of raw materials provided by the department entering the steel production process, the planning amount of scrap, pig iron, slag, raw slag, return steel and ferroalloy is initially identified by the ingredient process according to furnace capacity. Such as the EAF has nominal capacity of 30t, the actual amount of steel is 40t, Solid material loading is 45-46t, generally the total planned ingredients is 40t, 35t for chromium stainless steel (scrap 25t, slag 5t and raw slag 5t). The melting furnace smelting process mainly completes steel melting and oxidation stage of steelmaking, adjusts C, P levels and tests the Cu content. In an electric furnace it does not adjust the chemical composition of alloy element content, so it's kind of the steelmaking is essentially not be impacted, but it has some special requirements: in the end of life of the furnace, due to tap hole becoming large and with slag, the furnace cannot refine stainless steel; after exchanging furnace tap, the first smelting furnace is cleaning furnace steel and special requirements steel grade is not been refined in the first three furnaces; new furnace body and new ladle are not used for the same furnace number; smelting furnace with leaking cannot refine bearing steel, carbon steel can be refined; after high Cr steel , bearing steel

smelting furnace is equivalent to cleaning the furnace. AOD is referred to Argon oxygen decarburization, this method is major refining method for smelting low carbon steel, CO is diluted with Ar for decreasing pressure to achieve the effect of a vacuum, so that the carbon can be off to a very low level. AOD furnace tuyere is mounted on the side wall near the bottom of the furnace and the mixed gas of Ar and O2 is blown into furnace. The blowing process is divided into the oxidation period, the reduction of the refining period. AOD is the main production process of stainless steel. LF refining process mainly completes the reduction of steelmaking and adjustment of steel chemical composition. When the temperature appropriate, the laboratory analysis of the chemical composition is conducted, and as a basis to add ferroalloy, up to temperature before testing, substandard then add ferroalloy. And so forth until it reaches the composition and temperature requirements, then hanging steel bag. LF furnace time depends on the furnace conditions (chemical composition). VD requires vacuum tank furnace for mainly removing the gas (H, N, O, etc.), reducing inclusions and improving purity. After ingredients are qualified, the casting is conducted. In molding / casting process, sampling and laboratory analysis ingredients are carried out, and the results can be as the chemical composition of the final steel. The inspector compares the ingredients with the standard. If the ingredients are substandard then the steel is melted down, else turn to next process, fill out the "ingot / billet turn card" transferred together to the next branch.

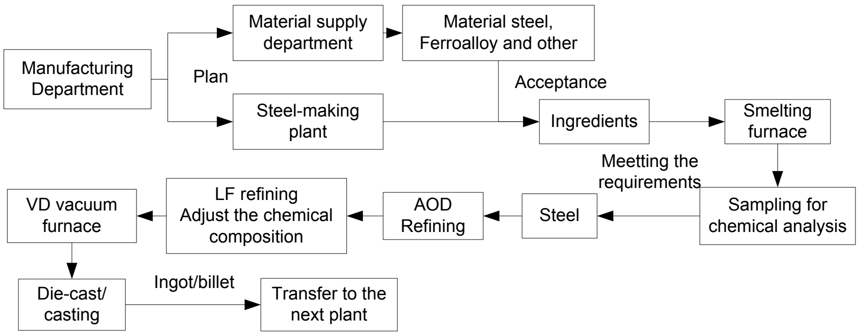


*Figure1: Production process for stainless steel smelting*

**3.2 Determination effecting factors of steelmaking raw material consumption cost**

From the steelmaking process, the raw material consumption costs of ingot or billet smelting production process on the one hand depend on the quality and quantity of different raw materials in different ratios, on the other hand depend on the yield of molten steel and molten steel into ingot (billet) rate, here ingot (billet) rate=(weight of qualified ingot) / (pouring molten steel weight), in which pouring molten steel weight = baked steel - steel recycled, the steel recycled is not double counted in the calculation of consumption of raw materials. The main material balance model of electric arc furnace is shown in figure 2.

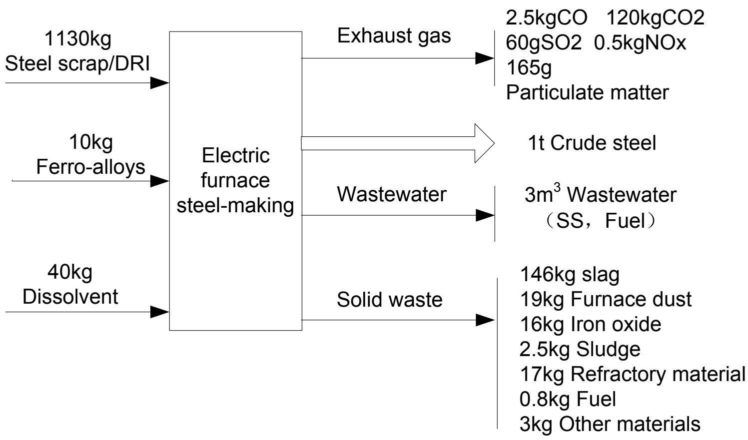


*Figure 2: Main material balance model of Electric Arc Furnace*

Building the steel metal balance equation

$$\sum_i X_i P_i a_{ij} = Qrb_j \tag{7}$$

Here, $Xi$ indicates raw material number $i$, $Pi$ means the grade of raw material number $i$, $a_{ij}$ represents the percentage of the $j$ metal in raw material number $i$, $Q$ represents amount of molten steel, $r$ represents the yield of molten steel, $b_j$ represents the percentage of $j$ metal in molten steel and is target of ingredients of production steel. By the material balance model and metal balance equation, when the amount of raw materials and quality indicators are certain, it is must to control waste and burning amount for making the ingot (billet) production as much as possible, and the quality and price of raw materials are comprehensively considered. So the consumption of raw materials can be controlled.

## 4. Analysis of affecting factors of steelmaking raw material consumption cost based on Rough Set

### 4.1 Calculation of the important degree of factors

The changes of one or several of the factors necessarily lead to changes of steelmaking raw material consumption cost. If one of the factors is ignored, that is the property representing the factor is deleted from the decision-making table, it will certainly undermine the indiscernible relationship between original objects of the decision-making table. Therefore, a condition attribute is removed from the decision table of raw material consumption cost analysis to examine the classification changes of decision table without this attribute. If the attribute will be removed the classification accordingly will changed, then the attribute is the effecting factor on changes of raw materials consumption cost. The larger the degree of change is, the greater intensity of the impact and higher importance this attribute has, and vice versa it illustrates the effect of the strength of the attribute is smaller, the lower importance that is.

It can be inferred, all attributes in minimum attribute simplicity of costs analysis decision table of steelmaking raw material consumption may be factors affecting the cost of raw materials consumed. The core attribute must be included in the minimum attribute simplicity, therefore the core attributes is necessarily affecting factors of the consumption cost and are certainty factors. While other attributes excepting the core in the minimum attribute simplicity may be affecting factors of the cost of raw materials consumed, are uncertainty factors.The $P_c(D)$ indicates the quantitative possibility of condition attribute c effecting on the decision-making attribute $D$. $T=(U,A)$ indicates the cost analysis decision table of raw material consumption, $P_c(D)$can be expressed as

$$P_c(D) = (\sum_{i=1}^{N} n_i) / N \tag{8}$$

$N$ is the number of all minimum attributes set of decision table, $n_i$ indicates whether the condition attribute $c$ is the element of minimum attribute simplicity of decision table. If it is, then $n_i=1$, otherwise $n_i=0$. Obvious $0 \le P_c(D) \le 1$, if $P_c(D)=0$, then attribute c is not an effecting factor of decision attribute $D$. If $0<P_c(D)<1$, then attribute $c$ may affect the decision attribute $D$. The greater the value is, the greater possibility is that the attribute affects the cost of consumption of raw materials. If $P_c(D)=1$, then attribute c is a core attribute, which is condition attribute definitely effecting the decision attribute $D$. Core attributes is a subset of the conditions attributes certainty affecting the decision attributes, but the importance of each attribute of core attributes to decision attributes are different. This importance is commonly assumed in advanced by the traditional method by auxiliary knowledge, and expressed with the "*weight*". In rough set theory, without the use of prior assumptions information, the effecting degree of an attribute to the steelmaking raw material consumption cost can be calculated by the existing attribute data in the decision table.

The dependence of decision attribute $D$ on attribute $R$ ($R \subseteq C$) is defined as

$$r_R(D) = Card(POS_R(D)) / Card(U) \tag{9}$$

Obvious for all $R \subseteq C$, $0 \le r_R(D) \le 1$, $r_R(D)$ reflects the importance of attributes $R$ for decision-making.

Define attributes $a \in C-R$, the degree of importance of decision attribute $D$ is expressed as:

$$r_a(D) = (Card(POS_C(D)) - Card(POS_{C-\{a\}}(D))) / Card(U) \tag{10}$$

Card(.) means number of elements in to a set, $r_R(D)$ measures the degree of change of decision classification after removing property a from the original decision table, larger value indicates that there are more changes in classification, the importance of property a is greater to decision attribute $D$.

### 4.2 Attribute reduction algorithm

Based on the above definition of attribute importance, attribute reduction algorithm is proposed. Input:

Step 1: A decision table $T=(U, C \cup D, V, F)$ , here $C$ and $D$ are the condition attribute and decision attributes.

Step 2: There is a core condition attribute relative to decision attribute $D$.

Output: A reduction of decision table.
Step 1: $R$=*Core*. $C$=$C$-$R$.
Step 2: In the condition attribute $C$, the attribute value a can be found to make $r_R(D)$ be the biggest value.
Step 3: If there are more than one attribute values to make $r_R(D)$ be the biggest value, an attribute is selected in which as attribute a that it forms a minimum number by combining with $R$.
Step 4: $R=R\cup\{a\}$, $C=C-\{a\}$.
Step 5: If $r_R(D)$ =1,terminate, otherwise, go to step 2.

### 4.3 Analysis steps for raw material consumption cost

Step 1: Establishment of initial information table. Based on domain knowledge and relevant prior knowledge the condition attributes and decision attribute set are determined. From historical production statistical data and cost data of raw materials of products the original database of factors - the cost of raw materials consumed is organized. The cost of raw materials consumed is a decision attribute, all the possible factors that affect the consumption of raw materials are condition attributes.
Step 2: Data pre-processing, Construction of decision table and discretization of continuous attributes. Before data mining, the value of the property is firstly mapped to a standard discrete symbol and each discrete symbol represents a data range.
Step 3: Attribute reduction. According to rough set attribute reduction algorithm, the redundant attributes are deleted from condition attributes, and ultimately get the smallest nuclear simplicity and property core.
Step 4: The importance and influence properties of condition attributes to decision attributes are calculated.

## 5. Application

In the case of the smelting cost analysis of raw material of 2Cr13 steel product in an enterprise, the data used is the performance statistical data of production and as the initial sample data shown in Table 1. According to the foregoing analysis, the yield of steel, a blank rate, the main raw material grade as the main effecting factors were analyzed.

*Table1: Original sample data*

| Furnace number | Steel yield rate % | Blank rate % | Grade of high-carbon ferrochrome % | Grade of high-carbon Ferromanganese % | Grade of ferrosilicon ％ | Raw materials cost (yuan/t) |
|---|---|---|---|---|---|---|
| 1 | 95.1 | 97.51 | 50 | 68 | 65 | 7635.33 |
| 2 | 93.5 | 97.21 | 58 | 65 | 72 | 8070.23 |
| 3 | 93.3 | 96.53 | 56 | 65 | 75 | 8079.56 |
| 4 | 98.11 | 97.01 | 52 | 75 | 70 | 7035.61 |
| 5 | 98.36 | 97.30 | 60 | 78 | 72 | 7040.77 |
| 6 | 96.53 | 95.87 | 50 | 70 | 68 | 7631.61 |

By using domain knowledge, the consecutive property values is conducted discrete processing：
(1) Steel yield rate represent the range: <95, [95, 97), ≥97
(2) Molten steel into ingot (billet) ratio indicates the range: <96.5, [96.5, 97.5), ≥97.5
(3) Grade of high carbon ferrochrome represents the range: <52, [52, 56), [56, 60), ≥60
(4) Grade of high-carbon ferromanganese represents the range: <66, [66, 69), [69, 72), ≥72
(5) Grade of ferrosilicon represents the range: <66, [66, 69), [69, 72), ≥72
(6) The unit cost of raw materials consumed represent the range: <7036, [7036,7384), [7384,7732), ≥7732
The sample data consists of consumption cost analysis decision table of steelmaking raw material of steel process. The condition attributes are steel yield, defined as a1, the molten steel into ingot (billet) rate, defined as a2,the grade of high carbon ferrochrome, defined as a3, the grade of high-carbon ferromanganese, defined as a4, and the grade of ferrosilicon, defined as a5. The decision attribute is the unit cost of raw materials consumed, defined as d. Condition attributes set is $C$={a1,a2,a3,a4,a5} and decision attribute set is $D$={d}. After discrete processing of sample, the data decision table is shown in Table 2.
According to Attribute reduction algorithm, attribute reduction calculation is conducted for the decision-making table to obtain the minimum simple set of attributes {a1, a2, a3, a4, a5}. The attribute core is {a1,a2,a3,a4,a5}, the possibilities of effect are $P_{a1}(d)=1$, $P_{a2}(d)=1$, $P_{a3}(d)=1$, $P_{a4}(d)=1$, $P_{a5}(d)=1$. It illustrates that steel yield rate, a blank rate, and grade of high-carbon ferrochrome, grade of high carbon ferromanganese and grade of

*Table2: Decision table*

| Domain | Condition attributes *C* | | | | | Decision attribute *D* |
|---|---|---|---|---|---|---|
| *U* | a1 | a2 | a3 | a4 | a5 | d |
| 1 | 2 | 3 | 1 | 2 | 1 | 3 |
| 2 | 1 | 2 | 3 | 1 | 4 | 4 |
| 3 | 1 | 1 | 3 | 1 | 4 | 4 |
| 4 | 3 | 2 | 2 | 4 | 3 | 1 |
| 5 | 3 | 2 | 4 | 4 | 4 | 2 |
| 6 | 2 | 1 | 1 | 3 | 2 | 3 |

Ferrosilicon is affecting factors on steelmaking raw material costs. Domain U={1,2,3,4,5,6}, a1,a2,a3,a4 and a5 have following equivalence classes:
$U$/a1={{1,6},{2,3},{4,5}}, $U$/a2={{1},{2,4,5},{3,6}}, $U$/a3={{1,6},{2,3},{4},{5}}, $U$/a4={{1},{2,3},{4,5},{6}}, $U$/a5={{1},{2,3,5},{4},{6}}.
By the decision table: $U$/Ind(d)={{1,6},{2,3},{4},{5}}, $U$/Ind($C$)={{1},{2},{3},{4},{5},{6}}.
According to the definition of a domain: $POS_C$ (d) = {{1}, {2}, {3}, {4}, {5}, {6}}.
Then, the dependence of the decision attribute $D$ on property $C$ is: $r_C$ (d) =$Card$ ($POS_C$ (d))/$Card$ ($U$) =1
The importance degrees of condition attributes to decision attributes were calculated. The set of attributes is C1={a2,a3,a4,a5}, C2={a1,a3,a4,a5}, C3={a1,a2,a4,a5}, C4={a1,a2,a3,a5}, C5={a1,a2,a3,a4}, then dependences of attributes were : $r_{C1}$(d)=$Card$($POS_{C1}$(d))/$Card$($U$)=4/6, $r_{C2}$(d)=$Card$($POS_{C2}$(d))/$Card$($U$)=2/6, $r_{C3}$(d)=$Card$($POS_{C3}$(d))/$Card$($U$)=4/6, $r_{C4}$(d)=$Card$($POS_{C4}$(d))/$Card$($U$)=4/6, $r_{C5}$(d)=$Card$($POS_{C5}$(d))/$Card$($U$)=4/6. So, the importance degrees of properties a1, a2, a3, a4, a5 were 2/6, 4/6, 2/6, 2/6, 2/6 for decision attributes. It illustrates that the rate of molten steel into billets has the biggest effect degree on the raw material consumption cost. Other attributes have the same effect degree on the steelmaking raw material consumption cost. Therefore, in the production of steel 2Cr13, when the ratio of raw materials is structurally similar and it takes the casting process, the yield of molten steel should be improved by avoiding quality accidents and increasing casting ratio. The continuous casting of some furnace can improve operating rate and the yield of the slab, and reduce production cost of materials.

## 6. Conclusion

According to Rough Set theory, the cost analysis and decision system for steelmaking raw materials is established. And through the real production data discretization and attribute reduction, definite and indefinite factors affecting the steelmaking raw material cost are determined and their influence extent is evaluated. Application of this method in the process of steel production cost analysis can extract the potential information hidden in the data, to ensure the objectivity of the cost analysis. An example shows that the method can be used to quantitatively analyze the influence of various factors on the cost of raw material consumption, and to solve the problem of estimating the influence degree of each factor on the cost.

### Reference

Chen W.D., 2004, Application of artificial neural networks in the rough tax warning system, Southeast University (Philosophy and Social Sciences), 6(1), 30-32.

Han J.W., 2000, Michline K, Data Mining Concepts and Techniques, Morgan Kaufmann Publishers.

Liu W., Zhang T., CHAI, 2004, Dynamic cost analysis based on neural networks, Information and Control, 33(2), 15-20.

Li Y., Tang J.F., Po Y., 2007, Quality Function Deployment final important method for determining the degree of customer needs, Computer Integrated Manufacturing Systems -CIMS, 13(4), 791-796.

Miao D.Q., Li D.G., 2008, Rough set theory, algorithms and applications, Beijing: Tsinghua University Press.

Pawlak Z., 1982, Rough Sets, International Journal of Information and Computer Science, (11): 341-356.

Pawlak Z, 1991, Rough set: Theoretical aspects of reasoning about data, Dordrecht: Kluwer Academic Publishers.

Pawlak Z, Skowron A, 2007, Rough sets and Boolean reasoning[J]. Information Sciences, 177:41-73.

Zhang H., 2007, Integrated steel manufacturing cost management and application, Dalian: Dalian University of Technology.

Zhang H., He B., Yang C., 2008, Site selection based on rough sets and multi-objective planning of Logistics Distribution Center, Industrial Engineering and Management, 2, 69-73.