




## Article

# Database Clustering after Automatic Feature Analysis of Nonmetallic Inclusions in Steel

Andrey Zhitenev \* , Maria Salynova , Alexey Shamshurin , Sergey Ryaboshuk and Vladislav Kolnysenko

Scientific and Technological Complex New Technologies and Materials,  
Peter the Great St. Petersburg Polytechnic University, 195251 St. Peterburg, Russia;  
m.salnova21@gmail.com (M.S.); sham\_a@mail.ru (A.S.); ryaboshuk.sv@gmail.com (S.R.); rjkysityrj@mail.ru (V.K.)  
\* Correspondence: zhitenev\_ai@spbstu.ru

**Abstract:** Non-metallic inclusions (NMIs) in steel have a negative impact on the properties of steel, so the problem of producing clean steels is actual. The existing metallographic methods for evaluating and analyzing nonmetallic inclusions make it possible to determine the composition and type of NMIs, but do not determine their real composition. The analysis of single NMIs using scanning electron microscope (SEM), fractional gas analysis (FGA), or electrolytic extraction (EE) of NMIs is too complicated. Therefore, in this work, a technique based on the automatic feature analysis (AFA) of a large number of particles by SEM was used. This method allows to obtain statistically reliable information about the amount, composition, and size of NMIs. To analyze the obtained databases of compositions and sizes of NMIs, clustering was carried out by the hierarchical method by constructing tree diagrams, as well as by the k-means method. This made it possible to identify the groups of NMIs of similar chemical composition (clusters) in the steel and to compare them with specific stages of the steelmaking process. Using this method, samples of steels produced at different steel plants and using different technologies were studied. The analysis of the features of melting of each steel is carried out and the features of the formation of NMIs in each considered case are revealed. It is shown that in all the studied samples of different steels, produced at different steel plants, similar clusters of NMIs were found. Due to this, the proposed method can become the basis for creating a modern universal classification of NMIs, which adequately describes the current state of steelmaking.

**Keywords:** nonmetallic inclusions; steel; automatic feature analysis; ca-treatment; cluster analysis; database of chemical compositions; classification



**Citation:** Zhitenev, A.; Salynova, M.; Shamshurin, A.; Ryaboshuk, S.; Kolnysenko, V. Database Clustering after Automatic Feature Analysis of Nonmetallic Inclusions in Steel. *Metals* **2021**, *11*, 1650. <https://doi.org/10.3390/met11101650>

Academic Editors: Eric D. van Hullebusch, João Pedro Oliveira, Belén Díaz Fernández, Leszek Adam Dobrzański, Jae-chun Lee and Alex Lanzutti

Received: 30 September 2021  
Accepted: 14 October 2021  
Published: 18 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Non-metallic inclusions (NMIs) have a negative effect on the casting of steel, leading to surface defects of continuously cast billets and sheets, and reduce the mechanical and corrosion properties of the final product [1–5]. Therefore, producing steels not contaminated with NMIs is one of the main tasks of metal products manufacturers.

The development of technological recommendations for obtaining clean steel is possible if information on the composition, size, and nature of inclusions is available; therefore, modern methods are required for their study and analysis. Existing metallographic methods for evaluating and analyzing make it possible to determine the composition and type of NMIs (ASTM E1245 and ASTM E45 [6,7]) but do not determine their real composition. Other methods, for example, analysis of single NMIs using a scanning electron microscope (SEM) [8], fractional gas analysis (FGA) [9], or electrolytic extraction (EE) [10] of NMIs are too complicated or unrepresentative for widespread use or do not provide statistically objective information.

For an objective statistical assessment of the amount, size, and composition of NMIs in industrial steels, it is possible to use databases obtained by the method of automatic particle analysis (AFA) [11–15] and containing information on a large amount of NMIs. For

this study, a metallographic sample is placed in a scanning electron microscope chamber equipped with an energy-dispersive spectrometer and a motorized stage. After that, in a given area of the sample, the composition, size, and other parameters of all non-metallic inclusions are determined. Thus, a representative database of compositions and sizes of NMI is created. The number of analyzed particles in AFA is comparable to the metallographic research methods using light microscopes. However, the capabilities of this method are not fully exploited. For example, in [16,17], this method is used only as an illustration of a statistically reliable determination of the average composition of NMIs. It is more productive to define groups of inclusions within the database. For this, systems of rules are often used, according to which NMIs, depending on the content of the components, are assigned to different classes. This method is proposed by the manufacturers of such systems [18–20], or independently developed by researchers [13]. However, this approach is not universal and contains uncertainty in the choice of rules in each specific case. Therefore, earlier the authors of this paper [14,15] used the k-means algorithm to identify the clusters of NMIs. However, the use of this method is not always convenient, since it requires a priori knowledge on the database structure and assumptions about the number of existing clusters. The analysis of descriptive statistics to identify the reliability of clustering [21,22] is difficult to formalize and can lead to incorrect results. Therefore, it is advisable to investigate other possible clustering methods. On the other hand, at present, a fairly large number of databases for different steels have been accumulated and a generalization of the results obtained is required.

Therefore, the purpose of this work is to test the AFA method for steels produced by different technologies and to compare the different methods for classifying NMIs.

## 2. Materials and Methods

In this work, samples of steels produced at different plants in electric furnaces and by the oxygen-converter method have been studied. Steels of ordinary quality, high-quality medium-alloyed, high-strength low-alloyed steels, and high-carbon steel for transport purposes have been investigated. Most metal products are used in a deformed state. Moreover, when an accident occurs due to NMIs, it is necessary to determine from the state of NMIs in the deformed metal what deviations from the optimal technology was made. Thus, samples of sheet and coils, and samples from finished railway wheels were studied.

Steel 1 (9MnSi5) of ordinary quality (Table 1) was obtained in an electric arc furnace (EAF). After EAF, the steel was processed in a ladle, the chemical composition was corrected, the melt was desulfurized with high basicity slag and deoxidized with an aluminum wire, and modified with SiCa. Then the steel was cast on a single-strand continuous casting machine and rolled on a reversing mill. The final sheet thickness was 15 mm.

**Table 1.** Chemical composition of the studied steels.

No.	Steel	Element. wt.%									
		C	Si	Mn	S	Cr	Al	Ti	Nb	V	Ca
		wt.%									
		ppm									
1	9MnSi5	0.07	0.5	1.3	0.010	0.1	0.019	-	0.01	-	16
2	B	0.7	0.4	0.9	0.019	-	0.004	0.01	-	0.04	5
3	X70	0.09	0.2	1.5	0.002	-	0.039	0.02	0.06	-	21
4	30CrMnSi	0.30	0.9	0.8	0.004	0.8	0.005	-	0.01	-	17

Steel 2 (grade B) was obtained in a basic-oxygen converter. The steel was processed in a ladle using non-aluminum technology [23]. It includes deoxidation with Si and Mn; further refining in ladle; modification with a minimum amount of SiCa at the end of treatment (consumption was 0.4–0.7 kg/ton of steel); processing in a RH-process and casting on a continuous casting machine into a round billet with a 430 mm in diameter, which were then rolled into wheels.

Steel 3 (grade X70) is high-strength low alloyed steel produced in a basic-oxygen converter. The steel was processed in a Ladle furnace, then treated in a chamber vacuum apparatus. At the end of the degassing, microalloying elements were added to the steel, finally deoxidized with an aluminum wire, and modified with Ca. The slabs were rolled on a hot strip mill to an 18 mm in thickness.

Steel 4 (30CrMnSi) was produced under the same conditions as Steel 1. However, during the ladle-treatment of this high-quality steel, a logistical error occurred, and the process was disrupted. The steel was desulfurized, adjusted to its chemical composition, and deoxidized with Al, and after that the ladle was sent to the waiting stand, where it waited for casting for 3 h. At the end of the wait, the aluminum concentration dropped significantly to 0.005% due to secondary oxidation. Before casting, the steel was modified with calcium. The thickness of the final rolled product was 25 mm.

To determine the chemical composition of the steels, the spark method was used, implemented in the Spectromax optical emission spectrometer.

For metallographic studies, samples were taken from coils, sheets, and wheels, which were then examined in the rolling direction. The samples were cut with abrasive discs on a cutting machine with a cooling emulsion supplied to the cutting zone, pressed into phenolic resin and grinded and polished to a state of “mirror shine”.

A technique based on the automatic analysis of a large number of particles on a scanning electron microscope (AFA) [13–20] was used to study the NMIs in this work. A Tescan electron microscope with an Inca Feature attachment was used for the study. The content of Al, Ca, Mg, Si, Mn, Si, Ti, Nb, V, N was analyzed. The content of light element O was not determined, but was calculated from the composition of oxides.

Database processing by different clustering methods was carried out using the Statistica software package.

### 3. Results and Discussion

#### 3.1. Comparison of Clustering Methods

After AFA, a database of compositions and sizes of NMIs was formed. However, in its original form, it is practically not suitable for analysis. First, the process of obtaining a database is associated with certain difficulties. The quality of preparation of specimens has a critical impact and there is a risk of obtaining a large number of analysis artifacts. For their automatic removal, a filter [14,15], which removes NMIs from the analysis, the total content of iron and oxygen in which is more than 95%, as well as particles with a high silicon content, introduced during grinding and polishing, has been developed. But even after filtering, the analysis of the database is difficult, since there is too much information. Therefore, it is necessary to separate the inclusions by composition into groups. In this work, cluster analysis was used. Cluster analysis is a group of techniques used to classify objects or events into relatively homogeneous groups called clusters [21,22]. Cluster analysis decomposes raw data into interpretable groups. Elements from one group should be as similar as possible, and elements from different groups should be as different as possible from each other.

Features of the cluster analysis methods [24]:

- (1) Absolutely all data cannot be used to select the groups (clusters). It is necessary to highlight only those signs that will help to effectively divide the objects into groups;
- (2) Different methods can show different results for the same data;
- (3) Cluster analysis can create artificial structures inside the database and highlight non-existent groups.

There are two well-known clustering methods. The first is hierarchical clustering [25], with the help of which it is proposed to carry out a preliminary analysis of databases in different areas of data processing. In this method, a tree of clusters is formed, and the presence of groups of inclusions of a similar chemical composition in a particular sample is preliminarily determined. During the analysis, all NMIs included in the database are compared with each other, the distance between them, expressed by the square of

concentrations (Euclidean metric), is estimated, and NMIs clusters are formed. Gradually increasing the linkage distance, one can see both the smallest groups of NMIs, which differ in a unique chemical composition, size and morphology, and large groups of inclusions, which provides complete information on the types of nonmetallic inclusions in the sample. The results of such clustering are presented in the form of dendrograms [25,26], on which the values of the overeating threshold are plotted along the Y axis, and the formed groups are marked along the X axis. Then, depending on the need, the clusters are selected at the required linkage distance. Further, these results were used as initial approximations for further processing by the second popular k-means method [14,15,21,22,24]. When using the k-means method, the most important uncertainty arises when choosing the initial number of clusters [14,15], and the use of the tree clustering method for preliminary determination of the number and type of clusters allows it to be eliminated.

Table 2 contains information about the received databases (Table 2). For each sample, information of 1500 non-metallic inclusions was collected. Depending on their size and nature of their mutual distribution, different sections of the section are scanned in automatic mode. In 9MnSi5 steel, the total contamination  $V_{tot}$  is high and amounts to 0.06%. The total area of all investigated NMIs ( $A_{NMIs}$ ) was 0.007 mm<sup>2</sup> with 11.16 mm<sup>2</sup> of the entire scanned area of the specimen ( $A$ ) and an average area of NMIs ( $A_{aver}$ ) of 4.8 μm<sup>2</sup>. In steel B, 0.0049 mm<sup>2</sup> of inclusions were viewed over an area of 8.28 mm<sup>2</sup>, with the same total contamination of 0.06%. The other two steels (X70 and 30CrMnSi) contain a significantly smaller number of inclusions, therefore, in order for the sample sizes to be the same, a much larger sectional area was examined in them. In steel X70, a total of 54.1 mm<sup>2</sup> was scanned, in steel 30CrMnSi, 62.35 mm<sup>2</sup>.

**Table 2.** Average chemical composition, the number of inclusions in the databases after AFA.

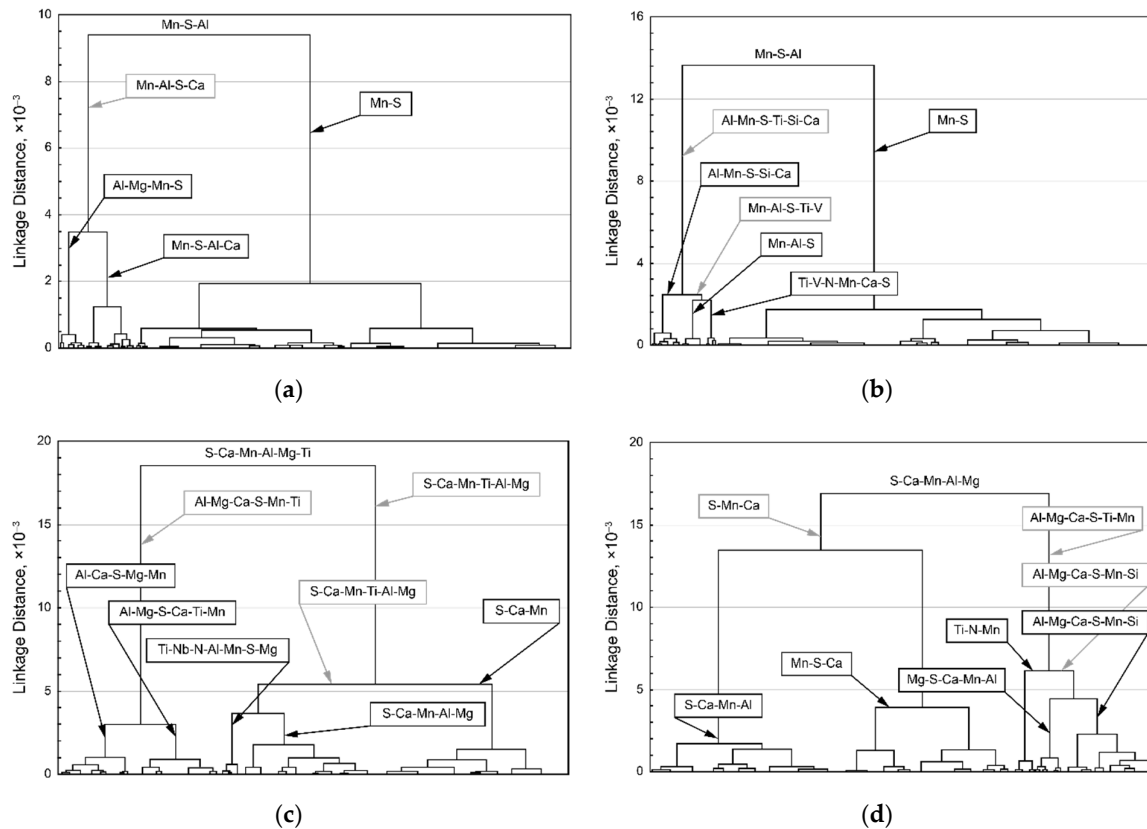
No.	Steel	Elements (C > 5%)	Concentration of Elements in NMI, without O (wt.%)	NMIs Number	$A_{aver}$	$A_{NMIs}$	$A$	$V_{tot}$
					μm <sup>2</sup>	mm <sup>2</sup>	%	
1	9MnSi5	Mn-S-Al	55-34-5	1466	4.8	0.0070	11.16	0.06
2	B	Mn-S-Al	58-32-5	1544	3.2	0.0049	8.28	0.06
3	X70	S-Ca-Al- Mn-Mg-Ti	28-25-16-11- 10-6	1607	6.8	0.0110	54.10	0.02
4	30CrMnSi	S-Ca-Mn- Al-Mg	31-22-20-11-9	1696	5.6	0.0095	62.35	0.015

For a preliminary analysis for each database its average chemical composition was calculated, and the name was taken out in descending order of elements with a concentration of more than 5%. The composition of the NMIs in steels modified by similar technologies or containing a lot of sulfur is practically the same. So, in steel 9MnSi5 with a sulfur content of 0.01% and in steel B with a sulfur content of 0.02%, the entire database consists of Mn, S, and Al with a close ratio of elements. There is much less sulfur in high-quality steels X70 and 30CrMnSi and the technology of calcium input is more optimal, therefore the effect of modification is higher and the inclusions consist of their similar elements, S-Ca-Al-Mn-Mg-Ti in steel X70 and S-Ca-Mn-Al-Mg in steel 30CrMnSi.

However, this analysis does not reveal the whole variety of inclusions that can be found in the studied steels. Therefore, it is impossible to study the influence of technological parameters on the change in the volume fraction of all inclusions or their types.

In order to identify the diversity of the types of NMIs, the clustering of the collected databases by different methods was performed. Using the database of compositions and sizes of NMIs in 9MnSi5 steel as an example, we investigated how both clustering methods work. Figure 1a shows a dendrogram of NMIs compositions in a sample of 9MnSi5 steel plate. Initially, the database is a cluster of Mn-S-Al (55-34-5) with a small concentration of other elements. With hierarchical clustering at a linkage distance of about 9300, two large branches are distinguished. The left branch in Figure 1a contains Mn-Al-S-Ca oxysulfides,

which make up 35% of the volume fraction of all inclusions. As the linkage distance decreases to 3500, two more clusters Al-Mg-Mn-S (11.45%) and Mn-S-Al-Ca (23.99%) are revealed. The right large branch contains the main mass of Mn-S inclusions, and its further division is not necessary, since clusters of the same chemical composition begin to appear.



**Figure 1.** Dendrograms of clusters of NMIs compositions in the studied steels. (a) 9MnSi5; (b) B; (c) X70; (d) 30CrMnSi.

Thus, in this database, it was possible to identify three NMIs clusters. Next, we performed k-means clustering for three clusters. The results of clusterization by the two methods are shown in Table 3. The chemical composition and names of the clusters are practically the same. There is a slight difference, for example, for the second cluster in the content of Mg and other elements, the concentrations of which are low. This is quite natural, since there are few inclusions containing these elements in the original databases; therefore, even small changes in the composition of a cluster lead to changes in its average composition for these elements.

Another difference was found in the mean areas of inclusions in each cluster. The average size of inclusions in the clusters identified by the hierarchical method is slightly higher than the clusters identified by the k-means method. However, an analysis of the databases of NMI compositions obtained for other samples showed that this difference is not of a systemic nature (Appendix A) and is associated, first of all, with the redistribution of inclusions of different sizes with similar chemical compositions between different clusters. The confidence band for determining the average area is on average 5–15%, so this difference between the two methods is small.



**Table 3.** Results of cluster analysis of the database in sample 1, steel 9MnSi5.

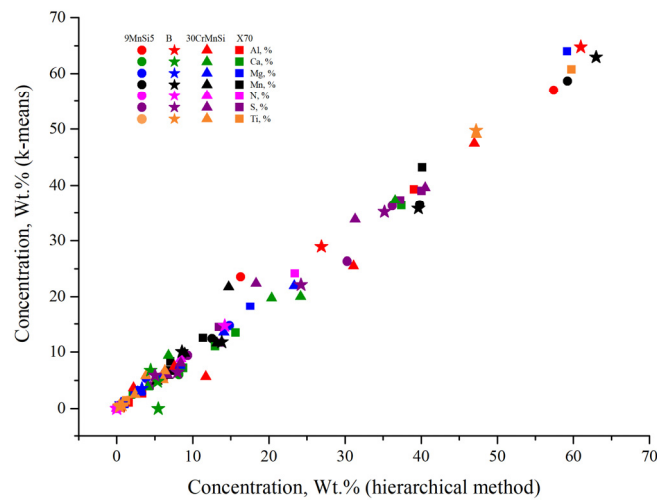
Cluster	Element, wt.%									N	Average Area, $\mu\text{m}^2$
	Al	Ca	Mg	Mn	N	Nb	S	Si	Ti		
hierarchical clustering											
Al-Mg-Mn-S	57.3	2	15	13	0.00	0.2	9	2	1.5	73	10.98
Mn-S-Al-Ca	16.2	8	4	40	0.00	0.2	30	1	0.4	149	11.27
Mn-S	0.3	2	0.3	59.5	0.00	0.3	36	1.4	0.2	1244	3.63
k-means											
Al-Mg-Mn-S	57.3	3	15	12	0.00	0.2	9	2	1.5	74	8.09
Mn-S-Al-Ca-Mg	24	6	5.2	37	0.00	0.2	26	1	0.6	93	7.84
Mn-S	0.4	2.5	0.4	59	0.00	0.3	36	1.2	0.2	1299	3.81

Dendrograms for the databases of samples of steels B, X70, and 30CrMnSi in Figure 1b–d are shown. In high-carbon steel B with a sulfur content of 0.019%, the dendrogram has an identical form with the dendrogram for 9MnSi5 steel (Figure 1b). Primarily, the database has a single Mn-S-Al cluster, which splits into two large branches at a linkage distance 13,500. The left branch in Figure 1b contains Al-Mn-S-Ti-Si-Ca inclusions, which make up 25.4% of the total amount of NMI. Further, this branch at the level of association of about 3000 is divided into clusters Al-Mn-S-Si-Ca, which is independent, and a cluster Mn-Al-S-Ti-V, which in turn further decomposes into clusters Mn-Al-S and Ti-V-N-Mn-Ca-S. The second large branch consists almost entirely of Mn-S inclusions with insignificant amounts of other elements.

High-quality steels X70 and 30CrMnSi contain much less sulfur, so they are better modified [27], since calcium is consumed in them mainly for the modification of oxides. This is reflected in the appearance of the dendrograms. NMIs in microalloyed steel X70 (Figure 1c) primarily form a cluster S-Ca-Mn-Al-Mg-Ti (28-25-16-11-10-6). As the database decomposes at the linkage distance of about 19,000, two large branches are spun out. The left branch in Figure 1b contains Al-Mg-Ca-S-Mn-Ti inclusions and this is about 47% of the volume fraction of all inclusions in this sample. With further decomposition, the left branch is divided into clusters Al-Ca-S-Mg-Mn and Al-Mg-S-Ca-Ti-Mn. The right branch contains S-Ca-Mn-Ti-Al-Mg inclusions and at the linkage distance in 5500 decomposes into a S-Ca-Mn sulfide cluster and an S-Ca-Mn-Ti-Al-Mg cluster, which subsequently decomposes into a cluster Ti-Nb-N-Al-Mn-S-Mg and S-Ca-Mn-Al-Mg.

The dendrogram for the 30CrMnSi steel specimen is outwardly similar to the dendrogram of the X70 specimen. Primarily, the database is a cluster S-Ca-Mn-Al-Mg (31-22-20-11-9), which, at the level of association of about 17,000, splits into two large clusters. The left cluster includes S-Mn-Ca sulfides, accounting for about 61% of all inclusions. Further decomposition makes it possible to reveal a more complex S-Ca-Mn-Al cluster at the linkage distance of about 13,000, and to separate pure sulfides into a Mn-S-Ca cluster. The second large Al-Mg-Ca-S-Ti-Mn branch at the linkage distance of about 6000 is divided into a Ti-N-Mn cluster and an Al-Mg-Ca-S-Mn-Si cluster, which decomposes into clusters Mg-S-Ca-Mn-Al and Al-Mg-Ca-S-Mn-Si.

Further analysis of these databases by the k-means method (Appendix A) showed that both methods systematically give a similar result, and the compositions of the identified clusters are close. For clarity, the coordinates of the centroids were compared for all clusters of the four studied databases (Figure 2). It can be seen that the centroids of the clusters determined by the two methods are very similar.



**Figure 2.** Comparison of hierarchical clustering and k-means.

Thus, the hierarchical method makes it possible to define the structure of the database and thus find natural clusters. The type of dendrograms for steels obtained using a similar technology is the same, and this can be used to compare them. Analysis by the k-means method requires additional procedures, for example, using the results of hierarchical clustering to determine the number of clusters. Therefore, further we used the method of hierarchical clustering, as completely suitable for independent use.

### 3.2. Study of NMIs Clusters in Different Steels

Let us further consider in detail the results of database clustering for the four studied steels. The results of clustering, as well as the results of recalculation of the compositions of cluster centroids into the phase composition by the method [8,14,15] are shown in Table 4. It was assumed that oxide-forming elements such as Al, Mg, Si form higher oxides in steel [28–30]. The calculation of sulfides was carried out taking into account the affinity of sulfide-forming elements for sulfur and oxygen [29]. So, if there is sulfur in the analysis, then the mass of calcium sulfide was determined, then the sulfur residue was converted into manganese sulfide. If there was an excess of calcium in the analysis, then it was converted into calcium oxide CaO. All nitride-forming elements, Ti, Nb, and V, were recalculated in TiN, NbN, and VN, respectively [31].

In ordinary steel 9MnSi5, the total volume fraction of non-metallic inclusions is 0.06%. The NMI themselves are represented by three types. The first type is inclusions with a high content of aluminum and magnesium oxides, containing a small amount of manganese and calcium sulfides, and their volume fraction is 0.0069%. The average area ( $A_{aver}$ ) of these NMIs is  $11 \mu\text{m}^2$ , and their nature is associated with incomplete modification of deoxidation products—magnesium spinel and corundum [16,32]. The second cluster Mn-S-Al-Ca contains significantly more inclusions—0.0144%. When the liquid metal is cooled and solidified, manganese and calcium sulfides are formed on the corundum inclusions that have appeared earlier [33], forming a second cluster consisting of 45% manganese sulfides, 10% calcium and corundum sulfides with a small amount of magnesium spinel. At the end of solidification, there is an intense formation of manganese sulfides due to the segregation of sulfur; therefore, a third Mn-S cluster was found, consisting almost of pure manganese sulfide. Tertiary inclusions cannot be removed from the solidifying metal; therefore, this cluster makes the largest contribution to the overall high contamination of steel ( $V = 0.0388\%$ ).

Table 4. Results of cluster analysis in the studied steels.

Cluster	Chemical Composition	$A_{aver}$ , $\mu\text{m}^2$	Mass Fraction of Oxide or Sulfide in NMIs, wt. %						$V_i$	$V_{tot}$
			$\text{Al}_2\text{O}_3$	MgO	$\text{SiO}_2$	MnS	CaS	Me* N		
9MnSi5										
Al-Mg-Mn-S	57-15-13-9	11	66	15	2.7	12.5	2.4	1.2	0.0069	0.06
Mn-S-Al-Ca	40-30-16-8	11.3	21.6	4.6	1.4	45.3	10.4	0.0144		
Mn-S	59-36	3.6	0.4	0.5		95	4	0.0388		
B										
Al-Mn-Si-S-Ca	61-14-8-8-5	7.9	73	3.5	11	7	6		0.0026	0.06
Mn-Al-S	40-27-24	4.9	39	1	6.6	43	6		0.0069	
Mn-S	63-35	2.8	0.8			98.5		0.5	0.0418	
Ti-V-N-Mn-Ca-S	47-18-14-9-6-5	5.7	1			3	8	89	0.0039	
X70										
Al-Mg-S-Ca-Mn	47-23-8-7-5	4.9	55	24.1	2.2	5.0	7.6	5.9	0.0026	0.02
Al-Ca-S-Mg-Mn	31-20-18-14-9	8.0	43.3	17.3	4	4	27	4.4	0.0069	
S-Ca-Mn-Al-Mg-Ti	31-24-15-12-8-6	5.1	17.6	11	2.5	25.9	34.6	8.4	0.0039	
S-Ca-Mn	41-37-13	5.1	3.9	4.2	2.67	10	75	4	0.0056	
Ti-Nb-N-Al-Mn-Mg-Ca	42-11-8-8-7-6-5	8	11.9	8.2	3	7.7	6.1	63	0.001	
30CrMnSi										
Al-Mg-Ca-S-Mn	39-18-16-14-7	5.6	50.1	19.9	8.1	1.8	19.1	1.1	0.0029	0.015
S-Ca-Mn-Al	40-37-11-5	4.5	8.6	3.5	4.8	24	58.8		0.0043	
Mn-S-Ca	40-37-13	5.3	5.7	4.8	2.9	65	20.6	1.1	0.0049	
Ti-N-Mn-S	60-23-7-5	8.2	3.2	1.9	1.6	11.1	1.1	81	0.0009	
Mg-S-Ca-Mn-Al	59-14-9-8-6	13.9	7.7	64.6	4.8	11.7	10.4	0.8	0.002	

\* Ti, Nb, V.

In wheel steel B, which is also slightly modified long before casting, the contamination is high and amounts to 0.06%. The first cluster is based on aluminum (73%), silicon (11%), magnesium (3.5%) oxides, and a small amount of calcium and manganese sulfides, and this cluster has the lowest volume fraction, 0.0069%. The second cluster contains less aluminum oxide (39%) and much more manganese sulfides (43%), because it includes NMI formed not only in the liquid metal, but also during solidification. The high content of silicon oxides in clusters containing oxides is associated with the low concentration of aluminum in this steel [23]. The third cluster consists of practically pure tertiary manganese sulfides, the volume fraction of which reaches 0.042%. The last cluster is formed by tertiary titanium and vanadium carbonitrides of a liquation nature [34]. This steel has a higher content of carbon and sulfur, which reduce the activity of magnesium [35], making the steel melt less aggressively to the refractory lining of the ladle. Therefore, the oxide cluster contains significantly less magnesium oxides than steel 9MnSi5, and the inclusions themselves are represented not by magnesium spinel, but by almost pure corundum.

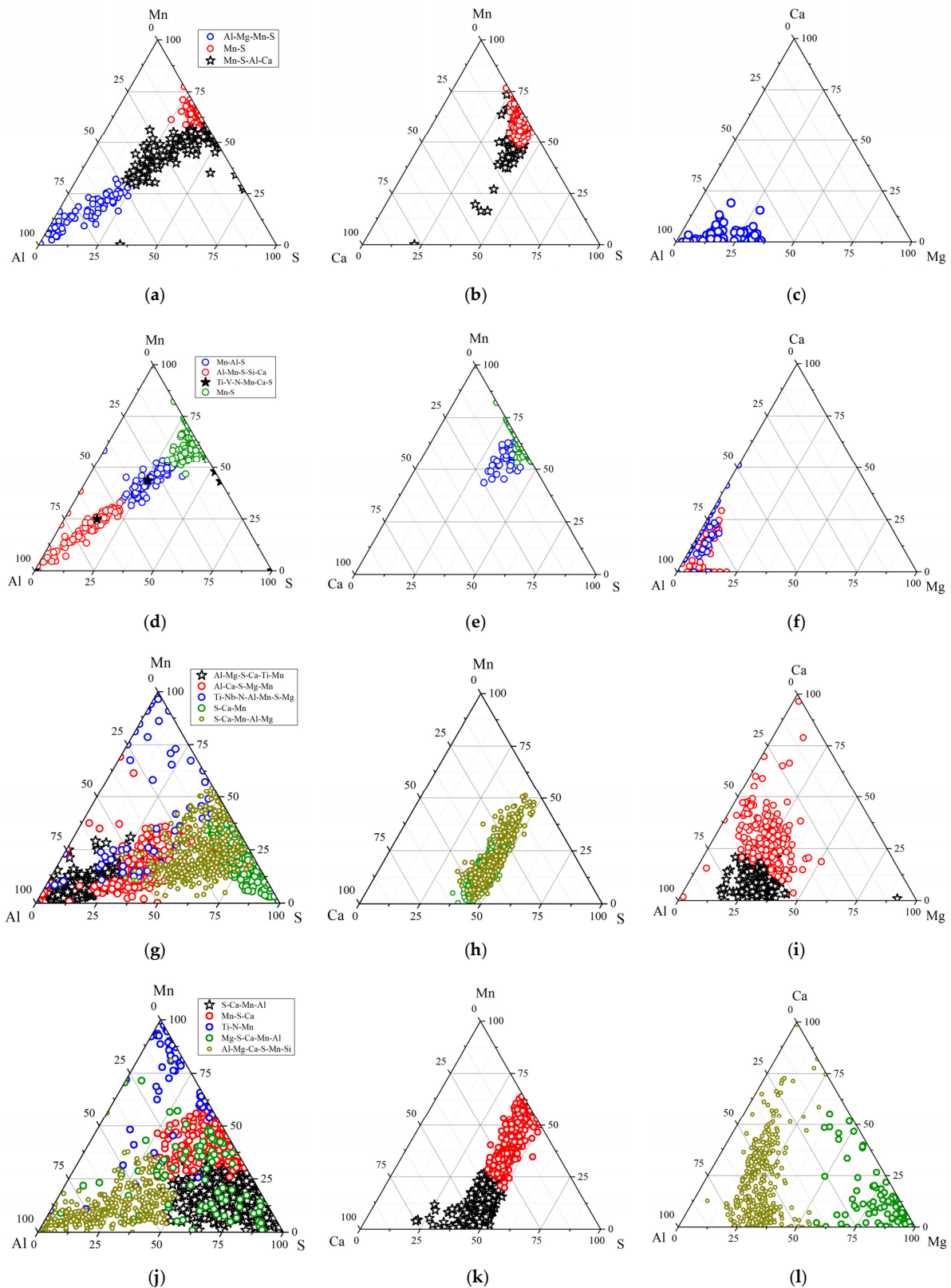
In optimally modified microalloyed steel X70, the total content of nonmetallic inclusions is much lower and amounts to 0.02%. This steel contains 4 oxysulfide clusters, Al-Mg-S-Ca-Mn (0.0026%), Al-Ca-S-Mg-Mn (0.0069%), S-Ca-Mn-Al-Mg-Ti (0.0039%), and S-Ca-Mn (0.0056%). Their nature is determined by the process of modification with calcium [12–17]. Initially, only corundum and magnesium spinel were present in liquid steel [12]. After the addition of calcium, modification began, and each cluster represents the products of this modification, and the higher the CaS content in the cluster, the more completely the inclusion is modified [12,34]. The last cluster of Ti-Nb-N-Al-Mn-Mg-Ca in this steel is dispersed titanium and niobium carbonitrides formed at the end of solidification. All clusters contain a certain amount of titanium and niobium nitrides, since this steel has high concentrations of carbonitride-forming elements.



In high-quality steel 30CrMnSi, the contamination with non-metallic inclusions is even lower, and amounts to 0.015%. The first cluster Al-Mg-Ca-S-Mn is a weakly modified magnesium spinel with a small amount of calcium and manganese sulfides. The volume fraction of this cluster is small and amounts to 0.0029%. Modified non-metallic inclusions with a large amount of calcium sulfide form an independent cluster S-Ca-Mn-Al, the volume fraction of which is higher and reaches 0.0043%. Tertiary inclusions of manganese sulfides, even at a low sulfur concentration, formed in solidifying steel due to liquation processes form an independent Mn-S-Ca cluster with a volume fraction of 0.0049%. Due to the titanium that is added to this steel, a Ti-N-Mn-S cluster was found containing titanium carbotitrides formed during solidification. This melt was carried out with a long downtime before casting. Because of this, there is practically no aluminum in this steel, which completely oxidized during the waiting time. Therefore, in this steel, the modification processes were most fully completed, since here as in wheel steel, an aluminum-free modification technology was implemented, which allows obtaining the largest amount of liquid non-metallic inclusions that are most efficiently removed from the ladle into the slag [23,27]. Therefore, this steel has the lowest NMI contamination. However, due to long-term interaction with the ladle lining, a big number of large single inclusions based on magnesium oxide were found in this steel, reaching an average area of  $13.9 \mu\text{m}^2$  and forming an independent cluster Mg-S-Ca-Mn-Al with a volume fraction of 0.002%.

For clarity, the figurative points of the compositions of nonmetallic inclusions that make up the clusters on ternary diagrams in the coordinates of the elements of which these clusters are composed was placed. We looked at several diagrams for each database. First, the Al-Mn-S diagram, which is the starting point of the non-metallic inclusions before modification. Then, the Ca-S-Mn diagram, on which we will place only clusters with a large quantity of sulfides in order to evaluate the degree of their modification and the Al-Ca-Mg diagram to evaluate the efficiency of oxide modification (Figure 3).

In steels 9MnSi5 and B, equally not optimally modified with calcium, in the Al-Mn-S coordinates (Figure 3a,d), all NMIs line up along a line from the Al angle to a point on the Mn-S axis corresponding to pure manganese sulfide. This is a natural trajectory of movement of the figurative composition of the non-metallic inclusions from pure  $\text{Al}_2\text{O}_3$ , which forms in liquid steel deoxidized with aluminum in the absence of other strong deoxidizers, to pure manganese sulfide, which is formed during solidification [30]. In unmodified or weakly modified steels, deoxidized with aluminum, the NMI compositions will always be located exactly along this line. In the Ca-Mn-S coordinates (Figure 3b,e), the inclusions are located near the point corresponding to pure manganese sulfide, since there are no other sulfide-forming elements. Since steel 2 (B) was modified too early, it contains more manganese sulfides than steel 1, which also contains some calcium sulfides. In the Al-Mg-Ca coordinates (Figure 3c,f) in 9MnSi5 steel, inclusions are grouped in the region around the Al angle corresponding to pure corundum and the point corresponding to magnesian spinel. In steel B, which is practically free of magnesium spinel, the oxides are grouped around the corner Al.



**Figure 3.** Chemical compositions of nonmetallic inclusions in the studied steels, presented on ternary diagrams in the coordinates of the elements most often found in NMI: (a–c) 9MnSi5, (d–f) B; (g–i) X70; (j–l) 30CrMnSi.

In optimally modified steels X70 and 30CrMnSi in the Al–Mn–S coordinates (Figure 3g,j), NMIs are located in a triangle between the aluminum corner, the sulfur corner, as well as

points on the Mn-S axis, corresponding to almost pure manganese sulfide. In steels treated with calcium with big consumption, in all clusters the number of Ca-S inclusions increases, and new clusters appear on their basis. Therefore, in the coordinates Al-Mn-S, a shift toward the angle S appears due to the appearance of other sulfide-forming elements except manganese. In the Ca-Mn-S coordinates (Figure 3h,k), for the same reason, inclusions are concentrated along the line from a point on the Mn-S axis, corresponding to pure manganese sulfide, to a point on the Ca-S axis, corresponding to calcium sulfide. It is along this line that the evolution of the composition of sulfides from MnS to CaS proceeds. The analysis of these diagrams shows that the modification of steel X70 (Figure 3h) was carried out more efficiently, since this steel does not contain pure manganese sulfides, in contrast to steel 30CrMnSi (Figure 3k). In the coordinates Al-Mg-Ca (Figure 3i,l), the oxide inclusions are located along the line from the calcium angle to the point corresponding to the spinel concentration. However, in steel 30CrMnSi, which had been waiting for casting for a long time, one of the clusters is displaced almost to the magnesium angle due to the large inclusion of refractory erosion products (Figure 3l).

#### 4. Conclusions

Various methods of cluster analysis of databases of compositions and sizes of NMIs after automatic SEM analysis of particles were investigated. It is shown that clustering by the hierarchical method and the k-means method gives a similar result. However, with hierarchical clustering, the structure of the database is more clearly presented and natural clusters are identified, therefore this method is more promising for further use. On the basis of this method, statistically reliable assessment of the compositions, sizes, and quantities of all NMIs in steel and their types was developed.

The developed method of cluster analysis makes it possible to assess the content of different types of NMIs in the finished metal and to determine how a particular technological operation influenced the contamination with different types of NMIs. This method allows for a short period of time using standard metallographic samples to obtain information about NMIs and can be used in the analysis of samples taken from the ladle, and in the routine or unscheduled analysis of finished products. This method is universal and can be easily applied to cast and deformed steels of different compositions.

Using the developed technique in different steels, produced according to different technologies, NMIs were studied. It is shown that in non-optimally modified steels, clusters with non-metallic inclusions based on  $\text{Al}_2\text{O}_3$  and  $\text{Al}_2\text{O}_3 \cdot \text{MgO}$  with a high content of manganese sulfides are presented. The total volume fraction of NMIs in these steels is at the level of 0.06%, and most of the inclusions are sulfides. In well-modified steels, the volume fraction of corundum and magnesium spinels is lower, while sulfides are mainly represented by CaS. In all steels, the volume fraction of calcium aluminates is low due to their removal from the ladle and transformation during solidification, therefore, one cannot find a separate cluster of calcium aluminates, because they are scattered over other clusters. Therefore, it is by the ratio of the amount of corundum and spinels and calcium sulfides that one can judge the degree of modification of the steel. In addition, the proposed method makes it possible to remove inclusions forming from the interaction of the refractory lining and the steel melt.

Each found cluster of NMIs is associated with a specific stage of ladle-treatment and casting. It is shown that they contain both common clusters associated with the similarity of deoxidation and modification schemes, and clusters that are unique for each sample, the type of which is determined by specific technological operations. The use of the developed method can help in the design of technology for obtaining clean steel with controlled modification of NMIs and obtain high-quality metal products and can become the basis for creating a modern classification of inclusions.

**Author Contributions:** Conceptualization, A.Z. and S.R.; methodology, A.Z., M.S. and A.S.; software, V.K.; review and editing, A.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research is partially funded by the Ministry of Science and Higher Education of the Russian Federation as part of World-class Research Center program: Advanced Digital Technologies (contract No. 075-15-2020-934 dated from 17 November 2020).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Results of clustering by different methods in steel B.

Cluster	Element, wt.%									NMIs Number	Average Area $\mu\text{m}^2$	
	Al	Ca	Mg	Mn	N	Nb	S	Si	Ti + V			
hierarchical clustering												
Al-Mn-Si-S-Ca	61	5.4	3.3	14			8	8			89	7.9
Mn-Al-S	27	5	1	40			24				77	4.9
Mn-S				63			35				1346	2.8
Ti-V-N-Mn-Ca-S		5		8.6	14		5		47 + 18		32	5.7
k-means												
Al-Mn-Si-S-Ca	65	5	3.5	12			6.5	8			74	7.8
Mn-Al-S-Ca-Si	29	7	1	36			22	5	0.5		92	5.0
Mn-S				63			35	1	0.4		1347	2.8
Ti-V-N-Mn-S				10	15		6		50 + 19		31	5.7

**Table A2.** Results of clustering by different methods in steel X70.

Cluster	Element, wt.%									NMIs Number	Average Area $\mu\text{m}^2$	
	Al	Ca	Mg	Mn	N	Nb	S	Si	Ti			
hierarchical clustering												
Al-Mg-S-Ca-Ti-Mn	47	7	2	5	0.0	1.3	8	2	6.1		283	4.9
Al-Ca-S-Mg-Mn	31	20	14	9	0.0	0.8	18	3	3.8		223	8.1
S-Ca-Mn-Al-Mg	12	24	8	15	0.0	1.9	31	1	6.3		423	5.1
Ti-Nb-N-Al-Mn-S-Mg	8	4	6	7	8.4	11.4	7	2	47.2		66	8.0
S-Ca-Mn	2	37	3	13	0.0	1	41	1	2.4		612	5.1
k-means												
Al-Mg-Ca-S-Ti	48	9	22	5	0.0	1.0	8	2	5.2		309	12.2
Al-S-Ca-Mg-Mn-Ti	26	20	14	10	0.0	1.4	22	2	5.8		314	5.9
S-Mn-Ca-Ti-Al	6	20	8	22	0.0	2.3	34	2	6.7		255	4.1
Ti-Nb-N-Al-S-Mn-Mg	6	4	6	6	8.8	11.4	6	2	49		62	8.5
S-Ca-Mn	4	37	3	12	0.0	0.9	40	2	2.5		667	5.6

**Table A3.** Results of clustering by different methods in steel 30CrMnSi.

Cluster	Element, wt.%									NMIs Number	Average Area $\mu\text{m}^2$	
	Al	Ca	Mg	Mn	N	Nb	S	Si	Ti			
hierarchical clustering												
Al-Mg-Ca-S-Mn-Si	39	16	18	7	0.0	0.2	13	5	1.3		334	5.6
Mg-S-Ca-Mn-Al	6	9	59	8	0.0	0.2	14	3	1		91	13.9
S-Ca-Mn-Al	5	37	2	11	0.0	0.5	40	3	0.3		615	4.5
Mn-S-Ca	3	13	3	40	0.0	0.3	37	2	1		589	5.3
Ti-N-Mn	2	1	1	7	23.4	0.3	4	1	60		67	8.2

Table A3. Cont.

Cluster	Element, wt. %									NMIs Number	Average Area $\mu\text{m}^2$
	k-means										
Al-Mg-S- Ca-Mn	39	13	18	8	0.0	0.2	14	5	1.4	337	4.8
Mg-S-Ca-Mn-Al	6	7	64	7	0.0	0.2	12	4	1	78	14.7
S-Ca-Mn-Al	5	36	3	13	0.0	0.5	39	3	0.3	721	4.9
Mn-S-Ca	3	11	4	43	0.0	0.3	37	1	1	495	5.4
Ti-N-Mn	1		1	7	24.1	0.3	4	1	61	65	8.3

## References

- Vinograd, M.I.; Gromova, G.P. *Inclusions in Alloyed steels and Alloys*; Metallurgiya: Moscow, Russia, 1972; p. 215.
- Kovačič, M.; Senčič, S. Critical inclusion size in spring steel and genetic programming RMZ. *Mater. Geoenviron.* **2010**, *57*, 17–23.
- Keskinkilic, E. Sulfide-type inclusion morphologies of a Ca-treated hot-rolled wheel steel. *J. S. Afr. Inst. Min. Metall.* **2011**, *111*, 417–422.
- Kazakov, A.; Kovalev, P.; Myasnikov, A.; Drobinin, A.; Zinchenko, S. Non-metallic inclusions and the nature of defects in cold-rolled sheet, Part 1. Defect of “Plena (Sliver)”. *Ferr. Met. Vol.* **2006**, *2*, 32–37.
- Hisser, H.; Tiekink, W.K. Solid inclusions in calcium treated aluminium-killed steel. *Clean Steel* **2007**, *7*, 45–51.
- ASTM E 1245-03. *Standard Practice for Determining the Inclusion or Second-Phase Constituent Content of Metals by Automatic Image Analysis*; ASTM: West Conshohocken, PA, USA, 2003; 8p.
- ASTM E 45-13. *Standard Test Methods for Determining the Inclusion Content of Steel*; ASTM: West Conshohocken, PA, USA, 2013; 19p.
- Kazakov, A.A.; Kovalev, P.V.; Ryaboshuk, S.V.; Mileikovskiy, A.B.; Malakhov, N.V. Investigation of the thermal-temporary nature of non-metallic inclusions in order to improve the metallurgical quality of high-strength pipe steels. *Ferr. Met.* **2009**, *12*, 5–11.
- Tabakov, Y.I.; Grigorovich, K.V.; Rummyantsev, B.A. Application of the method of high-temperature extraction in a carrier gas for the quantitative determination of TiN in rail steel. *Metally (Russ. J.)* **2016**, *6*, 28–35.
- Janis, D.; Inoue, R.; Karasev, A.; Jönsson, P. Application of Different Extraction Methods for Investigation of Nonmetallic Inclusions and Clusters in Steels and Alloys. *Adv. Mater. Sci. Eng.* **2014**, *2014*, 210486. [[CrossRef](#)]
- Singh, V.; Lekakh, S.; Peaslee, K. Using Automated Inclusion Analysis for Casting Process improvements. In Proceedings of the Steel Founders’ Society of America 62, Annual Technical and Operating Conference, Chicago, IL, USA, 1 December 2008; pp. 1–7.
- Pretorius, E.B.; Oltmann, H.G.; Cash, T. The effective modification of magnesium spinel inclusions by Ca treatment in LCAK steel. *Iron Steel Technol.* **2010**, *7*, 31–44.
- Ren, Y.; Wang, Y.; Li, S.; Zhang, L.; Zuo, X.; Lekakh, S.N.; Peaslee, K. Detection of Non-metallic Inclusions in Steel Continuous Casting Billets. *Metall. Mater. Trans. B* **2014**, *45B*, 1291–1303. [[CrossRef](#)]
- Kazakov, A.A.; Zhitenev, A.; Ryaboshuk, S. Interpretation and Classification of Nonmetallic Inclusions. *Mater. Perform. Charact.* **2016**, *5*, 535–543.
- Kazakov, A.A.; Zhitenev, A.I. Assessment and interpretation of nonmetallic inclusions in steel. *CIS Iron Steel Rev.* **2018**, *16*, 33–38. [[CrossRef](#)]
- Yang, S.; Li, J.; Zhang, L.; Peaslee, K.; Wang, Z. Evolution of MgO·Al<sub>2</sub>O<sub>3</sub> Based Inclusions in Alloy Steel During the Refining Process. *Metall. Min. Ind.* **2010**, *2*, 87–92.
- Zhu, S. *Low Cost Production of High Quality Clean Steel by the Implementation of Automated Inclusion Analysis Technology*; FEI Company: Hillsboro, OR, USA, 2014.
- Verma, N.; Pistorius, P.C.; Fruehan, R.J.; Potter, M.; Lind, M.; Story, S. Transient Inclusion Evolution During Modification of Alumina Inclusions by Calcium in Liquid Steel: Part II. Results and Discussion. *Metall. Mater. Trans. B* **2011**, *42*, 711–719. [[CrossRef](#)]
- Schamber, F.H.; Van Beek, C.G. *Understanding Particulate Contaminants via Automated Electron Beam Analysis*; Aspek Corporation: Delmont, PA, USA, 2003.
- Lamb, D. *Classification and Clustering*; Mir: Moscow, Russia, 1980; 390p.
- Borovkov, A.A. Summary of Reports Presented at Sessions of the Seminar on Probability Theory and Mathematical Statistics at the Institute of Mathematics of the Siberian Section of the USSR Academy of Sciences, September–December 1984. *Theor. Probab. Appl.* **1987**, *31*, 160–165.
- Grigorovich, K.V.; Shibaeva, T.V.; Arsenkin, A.M. Effect of a pipesteel killing technology on the composition and number of nonmetallic inclusions. *Russ. Metall. (Metally)* **2011**, *9*, 929–933.
- Jain, A.K.; Murty, M.N.; Flynn, P.J. Data Clustering: A Review. *ACM Comput. Surv.* **1999**, *31*, 264–323. [[CrossRef](#)]
- Ward, J.H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244. [[CrossRef](#)]
- Zhambyu, M. *Hierarchical Cluster Analysis and Correspondence*; Finance and Statistics: Moscow, Russia, 1988; 345p.
- Khoroshilov, A.D.; Grigorovich, K.V. Thermodynamic features of the modification of non-metallic inclusions with calcium in low-carbon steels deoxidized with aluminum. Proceedings of higher educational institutions. *Ferr. Metall.* **2019**, *62*, 860–869.
- Knyupel, G. *Deacidification and Vacuum Treatment of Steel. Per. with Him*; Metallurgy: Moscow, Russia, 1984; 414p.

28. Mchedlishvili, V.A. *Thermodynamics and Kinetics of Steel Deoxidation*; Metallurgy: Moscow, Russia, 1978; 287p.
29. Yavoiskii, V.I.; Rubenchik, Y.I.; Okenko, A.P. Nonmetallic inclusions and properties of steel. *Mosc. Metall.* **1980**, *173*, 169–172.
30. Inoue, K.; Ishikawa, N.; Ohnuma, I.; Ohtani, H.; Ishida, K. Calculation of Phase Equilibria between Austenite and (Nb,Ti,V)(C,N) in Microalloyed Steels. *ISIJ Int.* **2001**, *41*, 175–182. [[CrossRef](#)]
31. Kazakov, A.A.; Kovalev, P.V.; Ryaboshuk, S.V.; Zhironkin, M.V.; Krasnov, A.V. Management of the formation of non-metallic inclusions in the production of converter steel. *Ferr. Met.* **2014**, *C*, 91–96.
32. Imagumbai, M. Behaviors of Manganese-Sulfide in Aluminum-killed Steel Solidified Uni-directionally in Steady State Dendrite Structure and Inclusions. *ISIJ Int.* **1994**, *34*, 896–905. [[CrossRef](#)]
33. Abraham, S.; Klein, R.; Bodnar, R.; Dremailova, O. Formation of Coarse Particles in Steel as Related to Ferroalloy Dissolution Thermodynamics Part II: Crystallographic Study of Ferroalloys and Coarse Particles. *Mater. Sci. Technol.-Assoc. Iron Steel Technol.* **2006**, *2*, 109.
34. Buzek, Z. Zakladni termodynamické údaje o metalurgických reakciach. *Z. Buzek* **1979**, *20*, 5–111.
35. Guozhu, Y.E.; Jonnsson, P.; Lund, T. Thermodynamics and Kinetics of the Modification of Al<sub>2</sub>O<sub>3</sub>. *ISIJ Int.* **1996**, *36*, S105–S108.