

Automated Detection of Non-metallic Inclusion Clusters in Aluminum-deoxidized Steel



MOHAMMAD ABDULSALAM, MICHAEL JACOBS, and BRYAN A. WEBLER

This study presents a method to automatically identify inclusion clusters within a sample. Utilizing the output of scanning electron microscopy's automated feature analysis along with Density-based Spatial Clustering of Application with Noise, an unsupervised machine learning algorithm, inclusion clusters are identified based on their spatial position. The analysis was initially conducted on two samples known by manual analysis to be clustered and non-clustered to evaluate the applicability of this technique. A serial-sectioning analysis was performed to obtain a 3D representation of the inclusion distribution. The 2D and 3D results were consistent. To evaluate the area effected by clusters, the convex hull area was utilized rather than the total inclusion area in a cluster. The analysis was then applied to a series of samples from three aluminum-alloyed heats to investigate cluster evolution throughout the secondary steelmaking process. Several distinct types of clusters were identified. Agglomerated globular alumina inclusion clusters were observed after tapping, which then evolved to non-globular inclusion clusters. The same types of clusters were also observed for spinel inclusions, but they were not as pervasive as alumina inclusions. In addition, clustering of small micro-inclusions around a large macro-inclusion was occasionally observed.

<https://doi.org/10.1007/s11663-021-02312-5>

© The Minerals, Metals & Materials Society and ASM International 2021

I. INTRODUCTION

THE agglomeration of non-metallic inclusions during steel processing can lead to the formation of physical inclusion clusters. Although these inclusion clusters can promote flotation and removal of inclusions from liquid steel,^[1–3] they can have drastic effects on the properties of steel if they are present after solidification. One of the main concerns with inclusion clustering is their tendency to cause nozzle clogging,^[4–6] which can disrupt flow of liquid steel and sometimes completely block the nozzle.^[7] Large inclusion clusters can also cause stress concentrations in solidified steel, which can lead to cracks.^[8,9] Several studies have been conducted to understand the effect of inclusions clusters on steel properties, in addition to investigating their formation mechanisms and floatation behavior.^[1,10–13]

Zhang and Thomas^[14] studied the agglomeration of alumina inclusions, and concluded that in bulk, liquid steel agglomeration is controlled by Brownian, Stokes', and turbulent collisions. The collision regime is governed by inclusion size. After nucleation, diffusion and Brownian collision lead to the formation of fine spherical inclusions (around 1 to 2 μm radii), and turbulent collisions contribute to the growth of inclusions larger than 2 μm . Inclusion morphology is another important aspect that needs to be considered, and how the underlying type or shape of an inclusion defines the shape of an inclusion cluster. The shape of alumina clusters was examined in more detail by Tiekink,^[15] concluding that oxygen activity is the main factor that determines the shape of alumina inclusions and consequently their clusters.

Most of the previous research has focused on the formation and growth of alumina clusters specifically,^[1,14–16] since aluminum is a common deoxidizer used in the industry. Yin^[2,17] examined the formation of clusters using confocal laser scanning microscopy (CLSM), and concluded that there were strong long-range capillary attraction forces between pairs of alumina particles. This attractive force is weakest between a pair of liquid particles and strongest for a pair of solid particles; therefore, alumina inclusions (solid at steelmaking temperatures) have a higher tendency to agglomerate and form clusters.

MOHAMMAD ABDULSALAM and BRYAN A. WEBLER are with the Materials Science and Engineering Department, Center for Iron and Steelmaking Research, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, 15213. Contact e-mail: mabdulsa@andrew.cmu.edu MICHAEL JACOBS is with the Advanced Steel Processing and Products Research Center, Colorado School of Mines, 1301 19th St, Golden, CO, 80401.

Manuscript submitted May 5, 2021; accepted August 27, 2021.

Article published online September 21, 2021.

More recent work by Du^[18] conducted the same CLSM experiments with the addition of magnesium. The authors determined that the same attraction force existed between spinel particles; however, it was not as strong as the force between alumina particles. In addition, densification and deformation of spinel clusters were quicker compared to alumina clusters, and alumina clusters tend to form more dendritic structures which aid further agglomeration and growth. CLSM experiments by Kumar^[19] also concluded that alumina inclusions readily agglomerate forming a cluster of sintered inclusions. Kumar also noted the formation of spinel clusters can form either by magnesium pick up on alumina clusters or by collision of spinel inclusions; however, no sintering was observed with spinel clusters. Another CLSM study by Kimura *et al.*,^[20] on the behavior of magnesia and alumina-magnesia complex inclusions concluded that the tendency of coagulation for magnesium-containing inclusions is an order of magnitude less than that of alumina inclusions. The study attributed this effect due to the lower contact angle of magnesium-containing inclusions on the surface of molten steel, compared to that of alumina inclusions. On the other hand, CLSM experiments by Kang^[21] suggested that spinel inclusions do not agglomerate. A similar conclusion was reached by Dogan *et al.*^[22] While CLSM studies are very useful because they permit *in-situ* visualization of agglomeration, the processes are occurring at a liquid/gas interface and results may not be always directly applicable to behavior of inclusions surrounded by liquid steel.

One source of information about inclusion behavior in the bulk steel is automated scanning electron microscopy (SEM) coupled with Energy Dispersive X-ray Spectroscopy (EDS).^[23–28] Inclusion analysis is carried out using the automated feature analysis (AFA). AFA generates an abundance of data for each particle analyzed, though only one cross section is typically analyzed so all data are 2D. The most widely utilized information is the chemical composition produced by EDS analysis. Inclusion areas or equivalent circle diameters are also extracted to provide the inclusion size distribution. However, the analysis outputs more information that has not been as widely utilized, including shape metrics and spatial position of inclusions in the scanned area. This study presents an automated tool that can identify physical inclusion clusters from the spatial data produced from SEM analysis using machine learning methods.

Application of machine learning methods in the steel industry has gained a lot of interest over the past years. Recent work by Zhao *et al.*^[29] developed an automated tool for steel defect detection using deep learning. Cuartas *et al.*^[30] utilized machine learning to classify steel castings, and in another study^[31] to predict inclusion content of clean steels. All these studies employed “supervised” learning methods, where models are pre-trained on labeled data (inputs and outputs) to make predictions on unlabeled data (inputs only). The current work employs cluster analysis a form of “unsupervised” learning, where relationships in the data are inferred from inputs only.

“Cluster analysis” is a common unsupervised machine learning task in the field of data analytics. It is an automated method of arranging data into groups called clusters, where data points in one cluster are more closely related to each other than points outside the clusters. The physical agglomerations of inclusions are clusters in this general sense—groupings of AFA-identified inclusions in x and y-coordinate data space. The work presented here utilizes the inclusion spatial data, generated from inclusion SEM analysis, along with a clustering algorithm to identify physical clusters of inclusions. The spatial data present a 2D perspective of the inclusion distribution, and there has been concern whether areal scans are sufficient to describe volumetric inclusion distributions. Several studies have focused on investigating the effectiveness of analyzing a planar surface to describe a 3D distribution.^[32,33] This issue is also addressed in this study.

There are numerous clustering algorithms available, each having its own advantages and relevant applicability. The k-means^[34] is a popular clustering algorithm that has been widely used. It is an iterative centroid-based clustering technique. Some of the drawbacks of k-means are that the number of clusters has to be predefined and it assumes clusters have similar shapes and densities. Other examples of clustering algorithms include hierarchical clustering,^[35] DBSCAN,^[36] K-windows,^[37] or expectation-maximization algorithm.^[38] Each algorithm has its advantages and disadvantages. Therefore, a good understanding of the algorithm itself and the dataset being addressed is essential for selecting the appropriate clustering algorithm.

Work by Seleznev^[13] presented a similar approach of automated inclusion cluster detection. The authors used agglomerative hierarchical clustering, a classical bottom-up approach of cluster analysis, to estimate the effect of inclusion clusters on fatigue life. The work presented here employs another clustering algorithm, namely DBSCAN, an acronym for Density-based Spatial Clustering of Applications with Noise, developed by Ester *et al.*^[36] As specified by the name of the algorithm, DBSCAN is more suited for spatial pattern analysis compared to other clustering algorithms.

The aim of this work is to provide an automated method of inclusion cluster detection that can be generalized to any sample. Initially, the automated cluster detection technique was assessed on two samples to determine its performance. One sample is that clusters are expected to be present, and the other is free of clusters. To assess the area affected by a cluster of inclusions, the convex hull area was employed rather than the total inclusion area in a cluster. Since inclusion clusters can form concaved morphologies, the convex hull area is more indicative of the area affected by the cluster. In addition, the two samples were sectionally scanned over the same area, to provide a 3D perspective of the inclusion volumetric distribution. The cluster analysis was carried out on all 2D sections and the superimposed 3D distribution. The 2D and 3D analyses were compared and the inclusion clusters found on the 2D sections were consistent with those observed in the 3D volume.^[39] The work presented here utilized the tool

to examine clusters from a series of samples from three aluminum-alloyed heats, to investigate the evolution and types of clusters formed.

II. MATERIALS AND METHODS

This section is broken down into five sub-sections. The first section presents a summary of the steel samples from which the inclusion data were obtained. The following section illustrates the type of diagram utilized for inclusion compositions. The third and fourth sections are brief overviews of the DBSCAN algorithm and the computation of the convex hull area, respectively. The final section describes analysis of the serial sections.

A. Materials

Inclusion data were compiled from a sequence of samples from three heats. All heats were processed using basic oxygen furnace (BOF) steelmaking. The heats were aluminum deoxidized and alloyed; therefore, alumina inclusions were expected to dominate the inclusion population at the early stages, and they are known to agglomerate and form inclusion clusters.^[1,2] The heats included one sample taken after BOF tapping (L0), three or four samples from the ladle (L1 to L4), and a tundish sample (T). Heats 1 and 2 had four ladle samples, while heat 3 had three ladle samples. In total, 17 samples were compiled, six samples from Heats 1 and 2, and five samples from Heat 3.

Steel chemistries were measured by optical emission spectroscopy (OES), except for Mg, which was measured by inductively coupled plasma (ICP) OES. LECO analysis was used for Total O. A summary of the steel and slag chemistries, for the ladle samples only, is provided in Figures 1 and 2. The chemistries for the tundish samples were similar to the last ladle sample. The figures represent the change in steel and slag chemistry with respect to sample number, not time. The slag samples were taken 5 to 10 minutes before the steel samples on average. Slag FeO and MnO contents were less than 6 and 4 pct, respectively, for the L0 samples. They then decreased throughout ladle processing, and by the last ladle sample, they were both less than 0.5 pct.

B. Visualization of Inclusion Chemistry

Inclusion chemistry is commonly represented in the form of ternary plots, with compositions represented by anion or cation mass or mole fractions. In this study, the inclusion compositional distribution is displayed using ternary proportional symbol plots^[40] (e.g., Figure 5). The main advantage of the proportional symbol plots over conventional ternary plots is that they enable the visualization of the density of inclusions with a particular composition. The proportional symbol plot aggregates points with similar coordinates into triangles, where the size of each triangle is relative to the number of inclusions with that chemical composition. The inclusion phases are inferred from their compositions

by overlaying the ternary phase boundaries of the system at steelmaking temperatures (1600 °C), assuming minimal chemistry change during solidification.

The ternary proportional symbol plot can also be extended to visualize the area fraction of inclusions. By plotting the size of the triangles relative to the sum of inclusion areas rather than the number of inclusions. It is also common practice to plot the 50 pct liquid line, a boundary line, for certain inclusion systems (e.g., Mg–Al–Ca, S–Al–Ca), to show the region where more than 50 pct of the inclusions are in liquid phase.

C. DBSCAN Algorithm

The DBSCAN algorithm was developed in 1996,^[36] and since then, it has received wide attention in various fields. The main advantages of DBSCAN over other algorithms are as follows: (1) the number of clusters is determined based on the input data and algorithm's parameters, (2) it can identify arbitrarily shaped clusters, (3) it considers noise/outliers, and (4) it requires two parameters which can be easily determined based on the underlying data.^[41,42] Its disadvantages are that it is not entirely deterministic, meaning that reordering of points may lead to slightly different results. In addition, it is highly dependent on a distance measure. However, for the purpose of spatial clustering of inclusions, these issues are not of great concern. For spatial clustering on a polished cross section, the distance between points is the Euclidean distance, or length of the line that separates them (*i.e.*, a physical distance between inclusions). The variation in results, if the data are reordered, would also not be significant. If distinct inclusion clusters do exist in the data, they will always be identified by the algorithm. However, clusters may be labeled differently (e.g., cluster 1 might be labeled as cluster 3), depending on the initial point analyzed by the algorithm. There may also be minor differences in the number of individual inclusions assigned to a specific cluster; however, the objective is to identify clusters of inclusions and the assignment of a small number of inclusions to one cluster or another will not substantially affect the total number of clusters identified. This was tested on one sample; three trials were conducted by randomly reordering the data then performing the cluster analysis. The cluster assignment was the same for all three trials.

DBSCAN identifies clusters as dense regions in a data space, separated by areas of much lower density. In our case, isolated inclusions are considered as noise (low density regions) and physically clustered inclusions are identified as “clusters” (dense regions) based on their spatial position.

The DBSCAN algorithm is regulated by two parameters, epsilon (ϵ) and minimum points (MinPts). The ϵ parameter is analogous to the “neighborhood” around a data point, it defines the radius around a point (in n -dimensions, where n is the input variable dimensions). MinPts specifies the minimum number of points around a data point that are within ϵ , such that the main data point is in a cluster. Each data point is defined as one of the following:

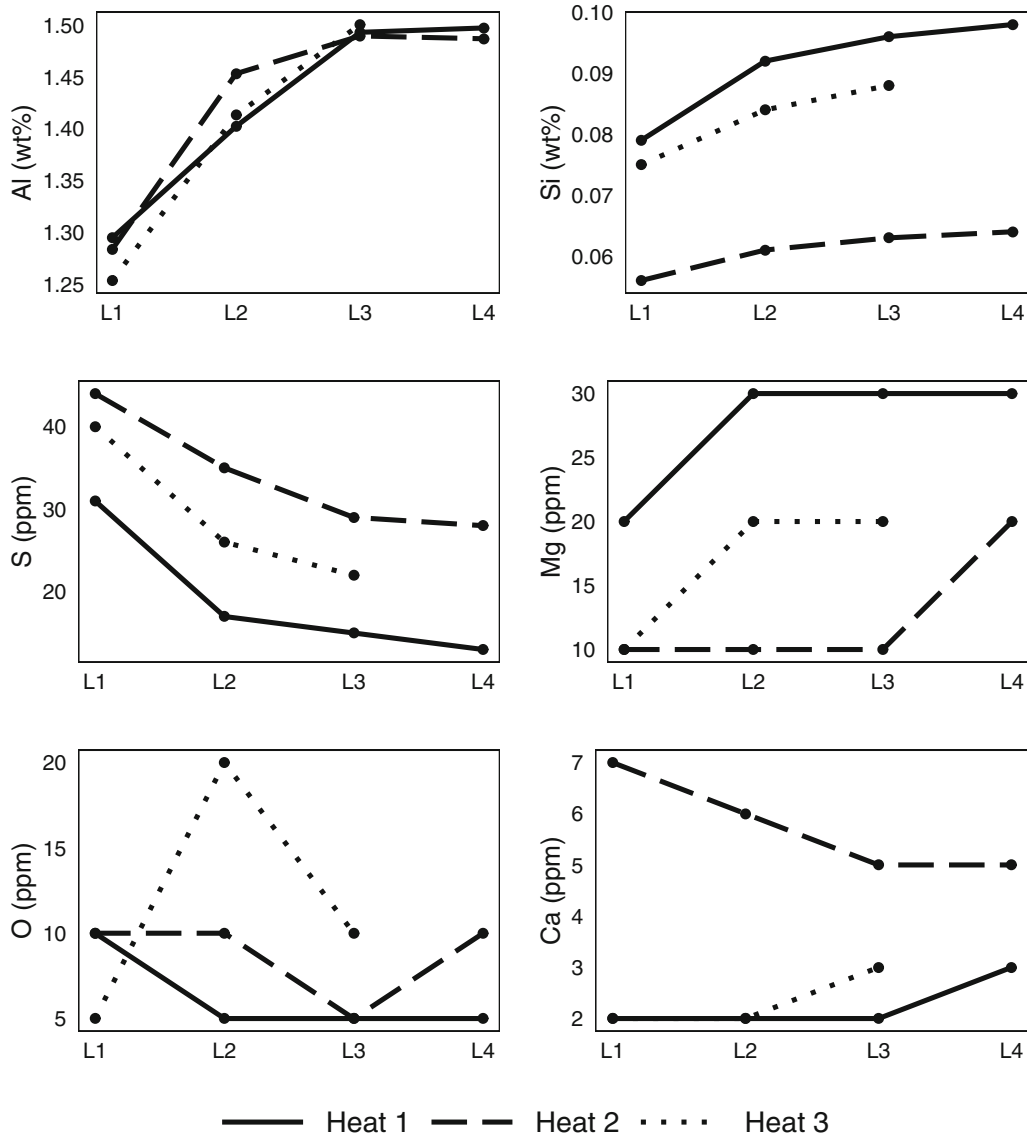


Fig. 1—Steel chemistry of ladle samples for all heats investigated in this work.

- Core point: if its neighborhood (ϵ) contains $\geq \text{MinPts}$,
- Border point: if its neighborhood (ϵ) contains at least one core point,
- Noise (or outlier): if it is not a core or border point.

Once all data points have been labeled, connected core and border points are assigned to a specific cluster, and all outliers are assigned as noise. Figure 3 presents an illustrative example of how points are labeled based on the parameters, ϵ and MinPts .

Although there are various clustering algorithms available, DBSCAN was applicable and relevant for our purposes. Since it focuses on spatial clustering and accounts for noise, and its main advantages over other algorithms are that the number of clusters is automatically generated, and any shape of cluster can be identified. The spatial coordinates of inclusions are the

inputs to the algorithm. Therefore, cluster detection is only dependent on the location of inclusions.

D. Convex Hull Area

To evaluate the area affected by a cluster within a sample, the convex hull area was employed rather than the total inclusion area in the cluster. The convex hull is basically the smallest convex polygon that encompasses all points. It can be visualized as a rubber band confining all points, as shown in Figure 4. Inclusion clusters can form in various shapes and sizes. In some instances, inclusion clusters can have concaved morphologies. Therefore, the convex hull is more indicative of the area effected by cluster than the total inclusion area. The work by Seleznev^[13] utilized the same metric in examining the effect of inclusion clusters on fatigue life.

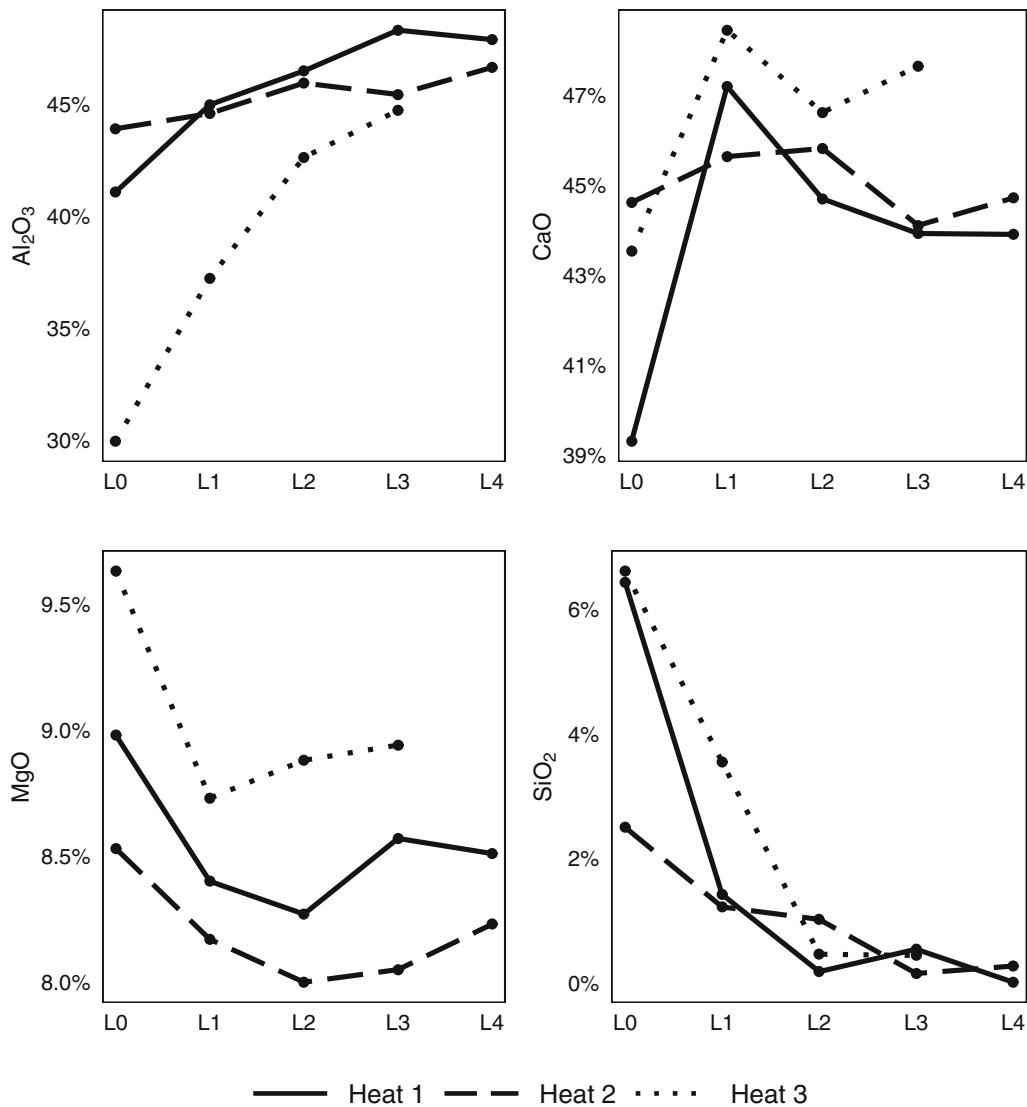


Fig. 2—Slag chemistry from all heats investigated in this work.

E. Serial Sectioning

To evaluate the 3D inclusion distribution, several serial sections were analyzed from samples L1 and T of heat 3. This was achieved by analyzing a specified area on the sample in an SEM, then polishing the sample to remove several micrometers off the surface and expose a new section for analysis. This process was reiterated until several serial sections were analyzed. To ensure the same area that is scanned at every serial section, the four corners of the area to be scanned were marked using a Vickers hardness indenter. A 5 kg load was used to produce indentations deep enough to allow for the analysis of multiple serial sections that are a few microns apart. Based on the length of the indent diagonals (d_1 and d_2), the height (h , effectively the depth of the indent) is calculated using the following equation:

$$h = \frac{d_1 + d_2}{4\sqrt{2} \tan \frac{\theta}{2}}, \quad [1]$$

where d_1 and d_2 are the lengths of the indent diagonals and $\theta = 136$ deg, a standard angle for the pyramidal indenter. Equation [1] is derived from the geometry of the Vickers hardness indenter.

After conducting the SEM analysis on a serial section, the sample was polished as follows:

- 9 μm polish for 5 minutes.
- 3 μm polish for 5 minutes.
- 1 μm polish for 2 minutes.

The same procedure was carried out following the SEM analysis of each serial section.

To allow the stacking of serial scans, one of the indentations on the first section analyzed was set as the reference point (*i.e.*, having coordinates [0, 0, 0]) from which the standardized coordinates of other points are calculated. It was assumed that the first section was level, such that all points on this section have a z -coordinate of 0, and consecutive sections have negative z -coordinates.

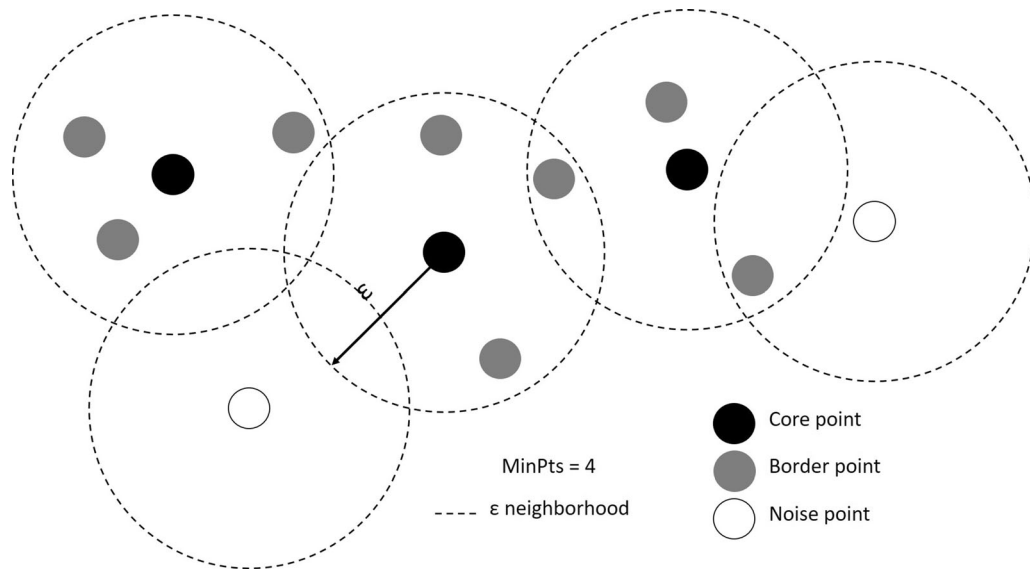


Fig. 3—Illustration of assignment of core, border, and noise (outlier) points in DBSCAN. Reprinted with permission from Ref. [39].

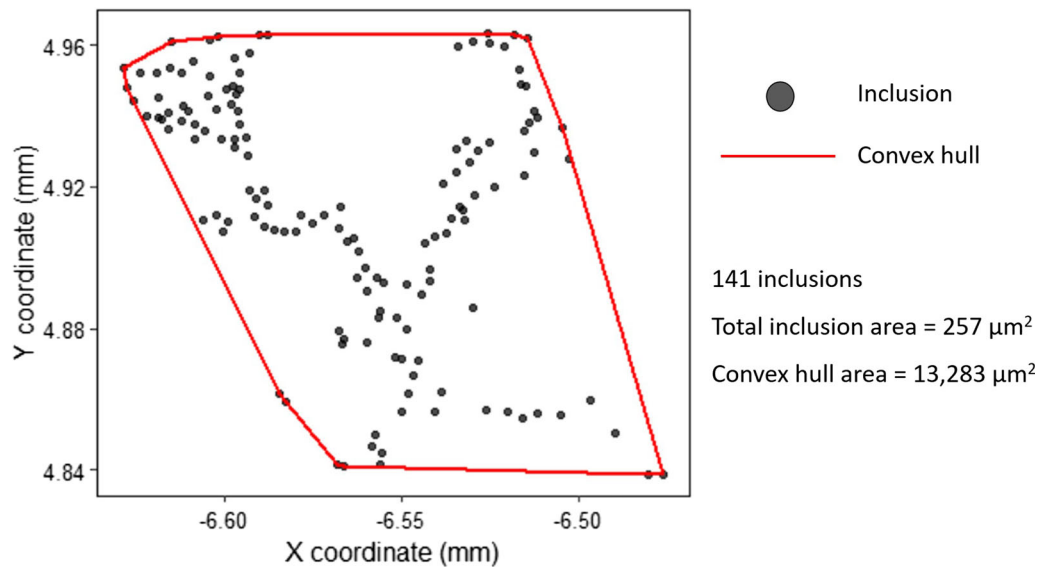


Fig. 4—Example of the convex hull area compared to the total inclusion area.

III. RESULTS

The results are summarized in three sub-sections. Initially, the SEM inclusion analysis results are presented for all samples in the three heats, displaying the overall inclusion composition distribution in each sample. The second section is focused on the cluster analysis of two samples, namely L1 and T from heat 3. This section provides a detailed examination of some of the inclusion clusters, to assess the performance of the automated cluster detection method. In addition, the results of the 2D and 3D analyses were contrasted. The final section presents the cluster analysis results for all samples to examine the clusters detected and their evolution through ladle processing.

A. Inclusion SEM Analysis

Inclusion analysis was carried out in FEI's ASPEX Explorer SEM, utilizing the AFA and EDS for compositional measurements. The accelerating voltage was set to 10 kV to minimize effects of the surrounding steel matrix.^[43] The different X-ray absorption effects will affect EDS composition measurements, not the spatial position of inclusions or cluster detection. SEM analysis was carried out on the inner cross section of the lollipop samples. Once the data were obtained, a manual filter was applied to the raw data to classify inclusions from non-inclusions, such as dirt, pores, or contaminants. The filter was based on the EDS measurements. Particles having Si counts greater than 100 or Fe content greater than 85 wt pct were identified as non-inclusions since the

only source of Si would be from contaminants and high Fe contents usually correspond to pores or erroneous (*e.g.*, false positives) readings.

Alumina inclusions are expected to dominate the inclusion population early in the ladle, since these heats were deoxidized and alloyed with Al. Thereafter, MgO reduction from the slag (with high basicity and low reducible oxide concentration) occurred in the ladle furnace leading to higher Mg contents in the liquid steel and subsequently the formation of magnesium aluminate (spinel) inclusions.

The measured composition distribution of inclusions is presented in Figure 5 in the form of ternary proportional symbol plots. In Figure 5, each triangle represents the area fraction of inclusions within the respective chemistry, displayed as cation mole fractions. The inclusion population was predominantly in the

Al₂O₃–MgO system with low CaO contents. As expected, the inclusion population was predominantly alumina after tapping from the BOF, and as the heats progressed through the ladle to the tundish the inclusion population evolves to spinel inclusions. This was also reflected by the increase in the steel's Al and Mg contents, and the decrease of MgO content in the slag, as shown in Figures 1 and 2.

B. Assessment of DBSCAN Algorithm on Inclusion Cluster Detection

To evaluate the performance of DBSCAN, initially two samples were analyzed—one sample where inclusion clusters were expected to occur and the other where no clusters were expected. The first sample selected for analysis was from the ladle metallurgy furnace, sample

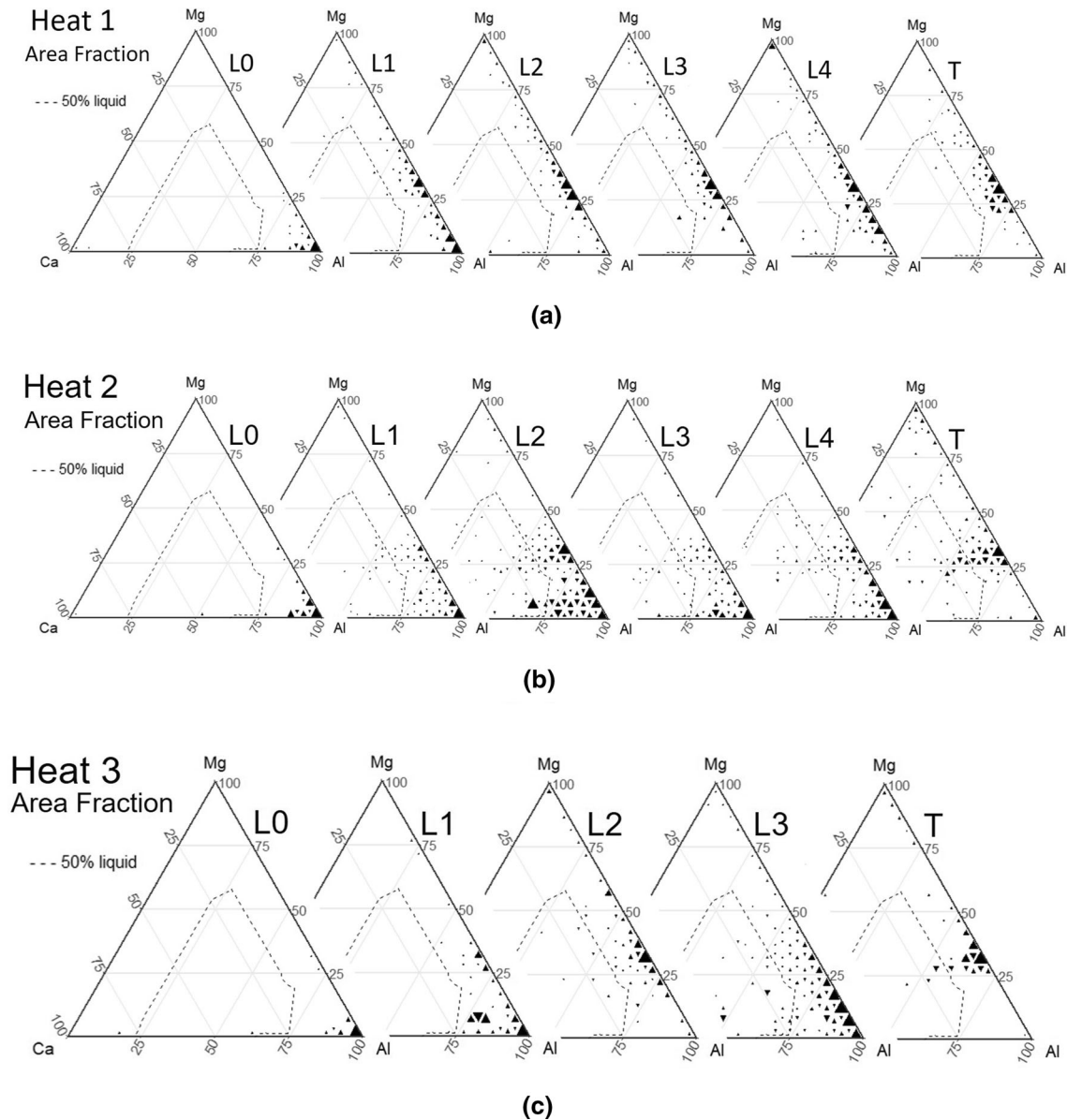


Fig. 5—Inclusion compositional distributions illustrated using ternary proportional symbol plots for (a) Heat 1, (b) Heat 2, and (c) Heat 3. The dotted line represents the 50 pct liquid region in the Mg–Al–Ca system at steel making temperatures, 1600 °C.

L1 from heat 3, taken approximately 40 minutes after BOF tapping. The second sample was a tundish sample (sample T from heat 3), taken approximately an hour and a half after BOF tapping at steady state casting conditions and where the tundish surface was well covered and protected from reoxidation. Alumina clusters were expected to occur in sample L1. While no clusters were expected to be detected in sample T, because all alloying additions have been made and enough time has passed to allow for inclusion floatation and removal. In addition, the sequential sectional scans were made on these two samples to compare the 2D and 3D inclusion distributions.

The cluster detection method was applied to the 2D serial sections separately. To select the algorithm's parameters, a sequence of values for both MinPts and epsilon were evaluated and plotted against the number of clusters identified in a sample. The ideal values would lie on the "plateau" region of the plots (*i.e.*, when the same clusters are identified for a wide range of ϵ and MinPts values). A more detailed account on the parameter selection methodology is provided in the authors' previous work.^[39] The selected values were 0.06 ϵ and 14 MinPts. It should be noted that the units for ϵ pertain to the units of the input variables (*i.e.*, mm in this case). The selected parameters were adopted to all serial sections of the L1 and T samples.

Five serial sections were analyzed in sample L1, and six in sample T. The material removal between serial sections varied slightly, on average, 7 μm of material was removed between sections. Sample L1 was composed of mostly of spinel and alumina inclusions, and sample T had mainly spinel inclusions (Figure 5(c)). In each serial section of sample L1, seven or eight clusters were detected, whereas no clusters were detected in any section of sample T. The inclusion summary for each section is given in Table I, along with a summary of the clusters in the last three columns. In terms of inclusion number and area density (area fraction), sample L1 displayed significantly higher densities and on average larger inclusions. The "Pct Inclusions Clustered" column is the number inclusions that were assigned to clusters as a percentage from the total inclusion count. The "Convex Hull Density" column is a summation of the convex hull areas of all clusters within the sample divided by the scan area, given in parts per million, analogous to the area density. As shown in the table, the convex hull density is roughly an order of magnitude larger than the inclusion area density. This implies that these clusters are not compactly agglomerated inclusions, and they essentially affect a larger area of the steel.

The spatial distribution of inclusions for the first section of both samples is presented in Figure 6. In the figure, inclusions are colored by cluster. Inclusions classified as noise are assigned to cluster 0, shown in gray. The coordinate axes are given in mm, while the inclusion size range is a few microns; therefore, size of points on this figure is not to scale. The point size is enlarged for visualization purposes only. A detailed

summary of each cluster from sample L1 is given in Table II, along with their respective average inclusion composition.

Figure 7 presents an example of the difference between the total inclusion area and the convex hull area in a cluster, for clusters 2 and 7 in sample L1. In this figure, the size of each point corresponds to the inclusion's equivalent diameter, and the shaded region to the cluster's convex hull area (points lying outside this region are "noise" inclusions).

To assess the extent to which the detected clusters on a 2D plane corresponded to 3D clusters, the inclusion distributions from each serial section were superimposed. Inclusion z -coordinates were calculated based on the indentation depths at each serial section. The same DBSCAN parameters were utilized to detect 3D inclusion clusters. In the L1 sample, 13 clusters were detected, most of which correspond to the same clusters identified in 2D. No clusters were identified in the 3D inclusion distribution of sample T. The clusters detected in both 2D and 3D, for sample L1, were compared to assess any discrepancies.

A total of 8245 inclusions were analyzed in all five sections of sample L1. The cluster assignment was consistent between the 2D and 3D analysis for 97 pct of the inclusions. In both analyses, 3202 (39 pct) were classified as noise inclusions, and 4754 (58 pct) inclusions were assigned to clusters. The remaining 289 inclusions (3 pct) were assigned to clusters in 3D only, while there were no inclusions assigned to clusters in 2D only. Figure 8 presents a visual comparison of these results, where gray points represent inclusions assigned to clusters in both 2D and 3D, and black points to inclusions assigned to cluster in 3D only. It should be noted the z -axis on this figure is not to scale and noise inclusions are not included in the figure.

C. Cluster Analysis of all Samples

The cluster detection method was applied to samples from all other heats. Again, the same parameters were utilized ($\epsilon = 0.06$ and MinPts = 14). Parameter plots were assessed for each sample, to verify that the selected parameters were within the suitable range. A summary of the inclusion analysis and the clustering results is presented in Table III. The inclusion density represents the number of inclusions per mm^2 , and the area density (area fraction) is the total inclusion area (μm^2) divided by the scanned area on the sample (mm^2), given in parts per million (ppm). The last three columns of Table III display the clustering results. A total of 116 clusters were identified in all samples. A better illustration of the evolution of clusters from tap to tundish is presented in Figure 9, displaying the variation of the convex hull density along the samples for each heat.

It should be noted that the tap sample (L0) in heat 1 had a conspicuously large number, area, and convex hull density. In contrast to the other tap samples, in heats 2 and 3, they had relatively low inclusion densities. This might be due to sampling time and/or location. All tap samples had predominantly alumina inclusions.

Table I. Summary of Inclusion and Cluster Analysis, for all Serial Sections of Samples L1 and T from Heat 3

Sample	Section	Total Inclusions	Inclusion Density (/mm ²)	Area Density (ppm)	Avg. Inclusion Area (μm ²)	Cluster Summary		
						No. of Clusters	Pct Inclusions Clustered	Convex Hull Density (ppm)
L1 (Heat 3)	1	994	57	339	6	7	60	6176
	2	1857	99	624	6.3	7	50	7937
	3	1701	92	454	5	7	51	7370
	4	1717	96	505	5.3	8	59	10,733
	5	1976	114	505	4.4	8	68	12,129
T (Heat 3)	1	202	12	27	2.3	0	0	—
	2	699	39	48	1.2	0	0	—
	3	339	19	36	1.9	0	0	—
	4	867	48	109	2.3	0	0	—
	5	448	25	37	1.5	0	0	—
	6	710	36	51	1.4	0	0	—

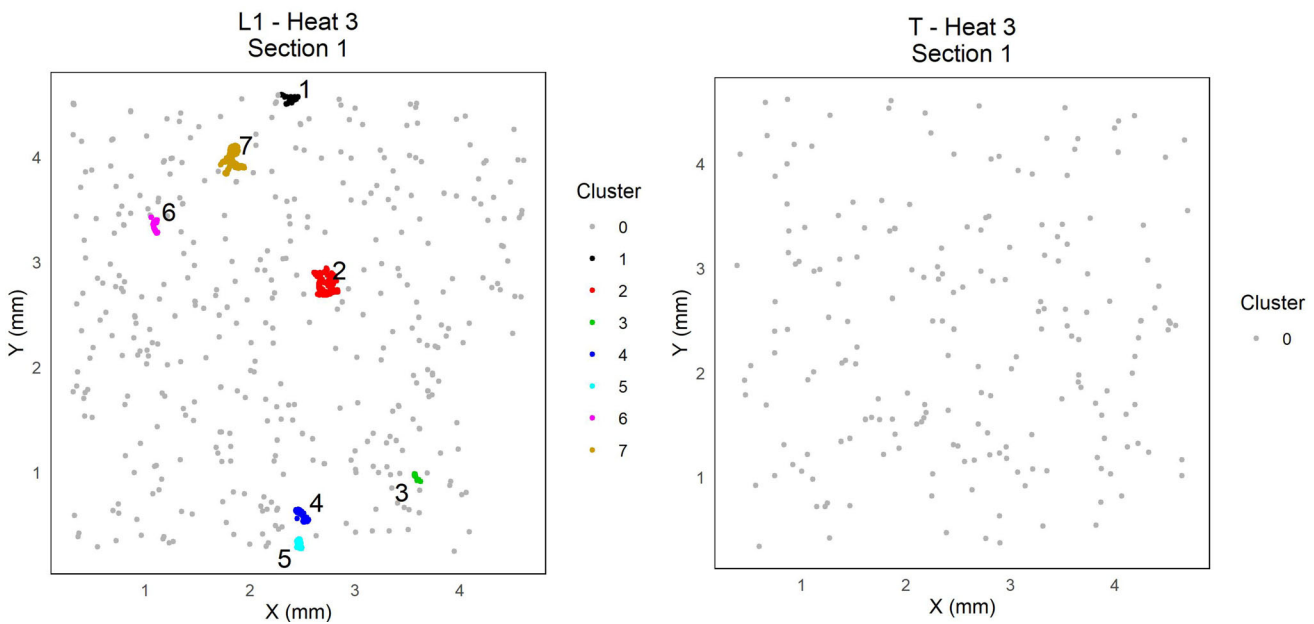


Fig. 6—Clustering (2D) results for the first section of samples L1 (left) and T2 (right) from heat 3. Reprinted with permission from Ref. [39].

Table II. Inclusion Cluster Summary, for Section 1 of Sample L1—Heat 3

Cluster	Total Inclusions	Avg. Inclusion Area (μm ²)	Total Inclusion Area (μm ²)	Convex Hull Area (μm ²)	Al (Pct)	Mg (Pct)	Ca (Pct)	S (Pct)	Mn (Pct)
0	400	3.4	1369	—	62	23	3	6	7
1	43	2.6	112	8816	84	14	0	1	1
2	235	12.8	3017	43,576	95	3	2	0	0
3	17	2.6	45	1335	88	12	0	0	0
4	76	4.5	345	8269	93	4	0	0	2
5	43	2.9	123	3501	93	7	0	0	1
6	34	2.1	70	5013	90	7	0	0	3
7	146	5.7	838	37,178	94	3	1	1	2

L1 - Heat 3
Section 1

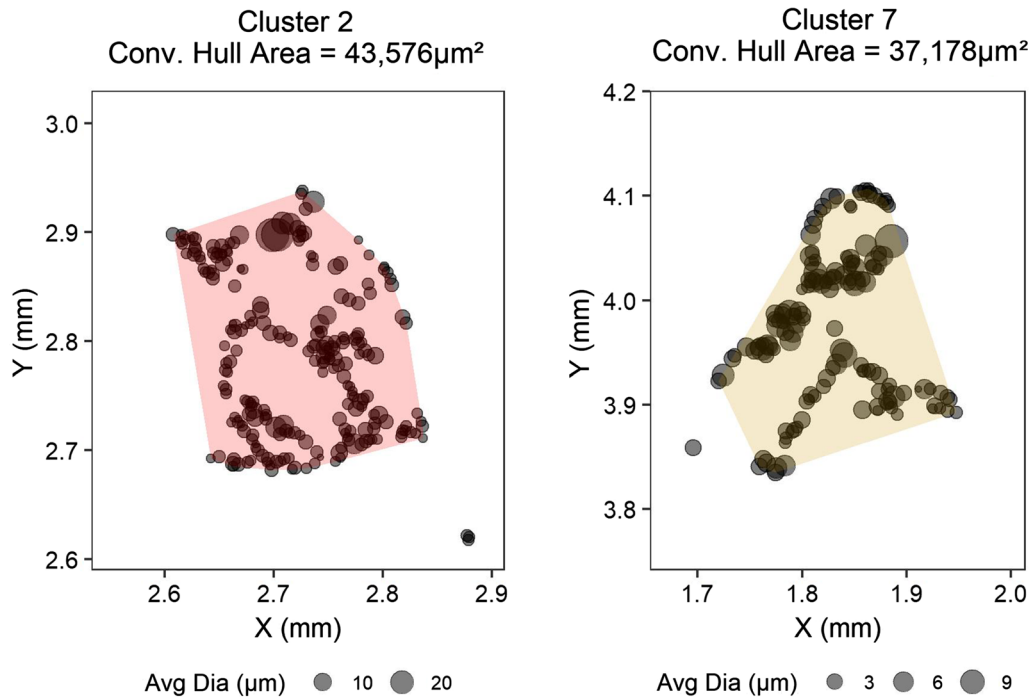


Fig. 7—Zoomed in spatial positioning plots for clusters 2 (left) and 7 (right), for section 1 of sample L1—heat 3. Reprinted with permission from Ref. [39].

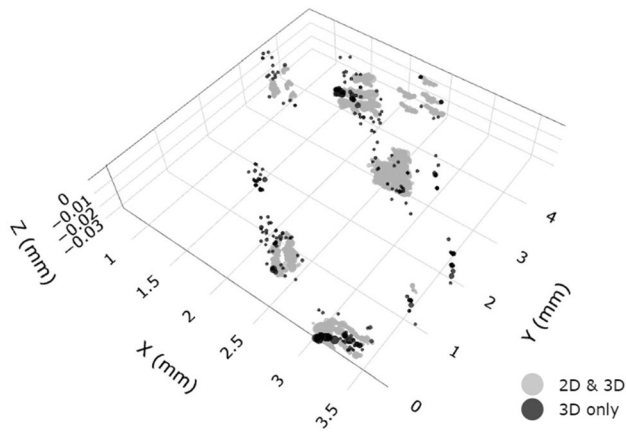


Fig. 8—2D vs 3D clustering, sample L1—Heat 3. Reprinted with permission from Ref. [39].

To illustrate the difference in compositions of clustered and noise inclusions in each sample, Figure 10 displays the average composition of clustered inclusions (top row) and noise inclusions (bottom row), for each sample. The columns of plots correspond to the different heats. Samples are represented by circles colored with different shades of gray, starting with white for L0 and gradually darker shades of gray for consecutive samples. It should be noted that the ternary plots are magnified on the Al corner for all plots in Figure 10. Both spinel and alumina clusters were observed. Alumina clusters

were more prominent earlier in the ladle (sample L0). No spinel clusters were observed in heat 2. The average magnesium content was higher for noise inclusions.

IV. DISCUSSION

Identification of inclusion clusters is critical for the assessment of steel cleanliness. Although relatively large clusters can be detected visually, it can be difficult to identify smaller ones. Such clusters can be detected by examining the spatial distribution of inclusions, generated by SEM AFA. However, manual inspection of steel samples and their corresponding spatial distribution can be arduous, time consuming, and prone to human error. The method presented here provides an automated method of cluster detection.

Initially the analysis was carried out on two samples, L1 and T from Heat 3, to assess the performance of the automated cluster detection technique. In addition, several cross sections were analyzed from both samples, to obtain the volumetric inclusion distribution. Several clusters were detected in the L1 sample, while no clusters were detected in sample T. In sample L1, the same clusters were identified in several serial sections. The largest two clusters in this sample are displayed in Figure 11, clusters 2 and 7 as per Figure 6 and Table II. The images are stitched BSE images taken during the AFA at low magnification.

Table III. Summary of Inclusion Analysis and Clustering Results

Heat	Sample	Inclusions	Inclusion Density (/mm ²)	Area Density (ppm)	Clusters	Pct Inclusions Clustered	Convex Hull Density (ppm)
1	L0	9931	2288	5335	16	100	323,368
	L1	2654	63	83	5	17	2145
	L2	3436	118	89	2	6	418
	L3	1159	37	81	4	8	303
	L4	1638	39	63	4	32	1735
	T	423	8	44	1	4	39
2	L0	321	10	24	4	70	764
	L1	3833	151	350	15	44	10,272
	L2	5892	87	221	16	19	2336
	L3	775	26	490	9	68	12,768
	L4	2757	68	178	7	49	5301
	T	3322	104	62	0	0	0
3	L0	298	7	10	8	62	353
	L1	975	53	306	7	61	5920
	L2	2151	67	59	2	3	271
	L3	2874	81	341	16	63	10,145
	T	202	12	27	0	0	0

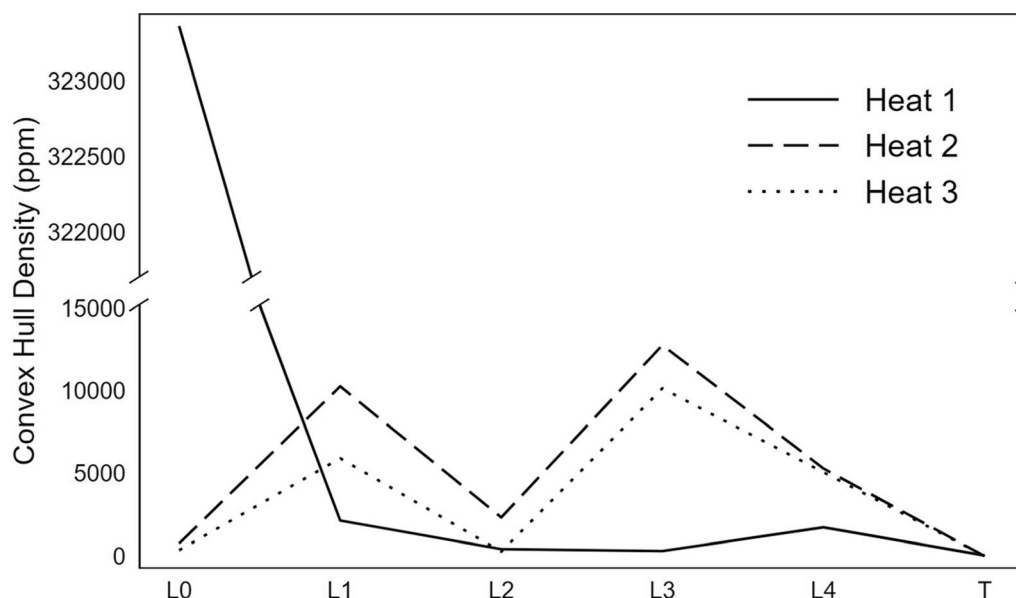


Fig. 9—The convex hull density of clusters in each sample, displaying the evolution of clusters in each heat from tap to tundish. There is no L4 sample for Heat 3.

The results from the 2D and 3D analyses were consistent, for both the clustered (L1) and un-clustered (T) samples. The clusters detected in 2D were compared with clusters detected from the 3D distribution. The same clusters were identified from both distributions. The only discrepancy was the additional small clusters detected in 3D and not in 2D, for sample L1—heat 3. Such variation is expected when a new dimension (z-coordinate) is introduced. Nevertheless, these inclusions are only 3 pct of the entire inclusion population. Therefore, it can be safely assumed that 2D areal scans provide a sufficient representation of the 3D volume.

The formation of large inclusion clusters can promote further agglomeration which can lead to nozzle constriction or blockage.^[4-6] As displayed by the results here, such large clusters envelope large areas compared to their total inclusion area. Therefore, the examination of the convex hull area of a cluster rather than its total inclusion area provides a better estimate of the area in the steel effected by the cluster. Another point worth noting is that large clusters (clusters with a convex hull area in tens of thousands μm^2) were predominantly Al rich. The non-clustered inclusions and some of the smaller clusters (*e.g.*, clusters 1 and 3 in Figure 6 and Table II) had on average higher Mg contents, closer to

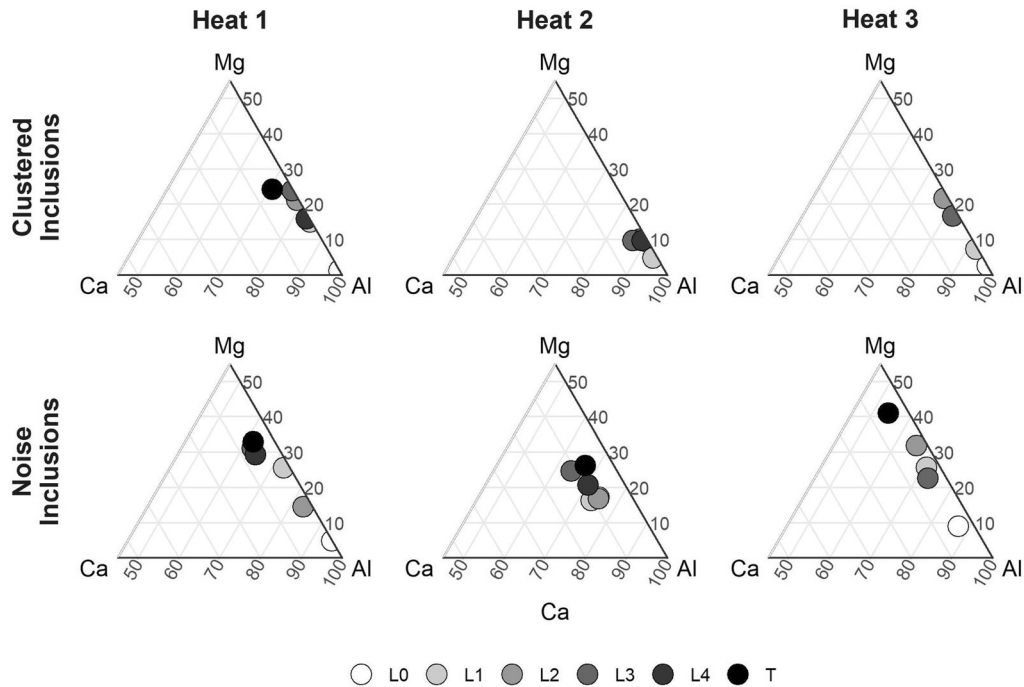


Fig. 10—Average chemical composition of clustered (top row) and noise (bottom row) inclusions. Each sample is represented by a circle with a different shade of gray, L0 displayed in white and consecutive samples in gradually darker shades of gray. Each heat is presented in a single column. Noise inclusions were generally richer in magnesium, compared to clustered inclusions.

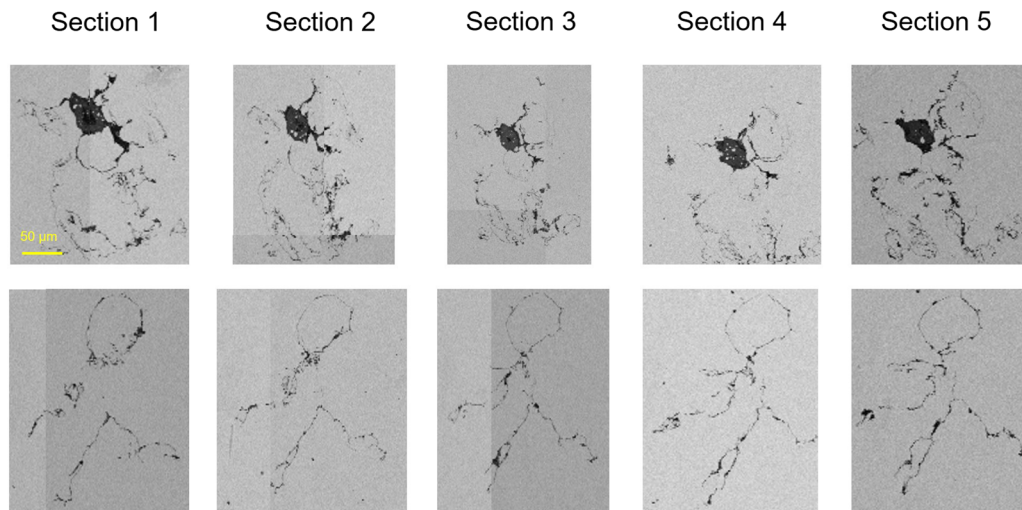


Fig. 11—Reconstructed field images of inclusion clusters in sample L1—Heat 3, from all serial sections. Top row of images is cluster 2 and bottom row is cluster 7, as per the designation given in Fig. 6 and Table II. Scale bar applies to all images (magnification 333x). Reprinted with permission from Ref. [39].

spinel inclusion chemistry. In these samples, alumina inclusions formed larger clusters, whereas spinel clusters were relatively smaller.

With respect to cluster evolution, heats 2 and 3 displayed a similar trend (Figure 9) in terms of cluster size (quantified by the convex hull densities). The density was low in the L0 sample, increased in L1, decreased in L2, then a greater increase in L3, until the cluster density reduced to nearly zero in the tundish sample (T). The fluctuations in cluster densities are attributed to

reoxidation events in the ladle, since a distinct peak in inclusion Al content was observed for the L3 samples of heats 2 and 3. However, the samples employed here were taken to track the general evolution of inclusions during ladle processing and were not matched to the timing of additions. With regards to heat 1, its L0 sample was densely clustered, much more so than the L0 samples of the other two heats, whereas other ladle samples from heat 1 (L1 to L4) displayed significantly lower cluster densities, compared to samples in heats 2 and 3,

although clusters were still observed in these samples. A possible explanation for this variation in cluster density between heats could be due to stirring practice. Heat 1 had bottom plug stirring, while heats 2 and 3 were stirred using a top lance. Tundish samples in all heats were free of clusters, except for heat 1, where only one small cluster was identified.

Two distinct cluster morphologies were identified. In all three heats, all clusters in the tap samples were composed of agglomerated and sintered globular alumina inclusions, as shown in Figure 12(a). Clusters of non-globular alumina inclusions were detected in all other ladle samples (L1 to L4), with an example is given in Figure 12(b). This type of cluster was the most prominent in all samples. Although the AFA classifies inclusions in this cluster as tens or hundreds of small inclusions, BSE images of these clusters indicate that they are composed of several elongated non-globular inclusions, that stretch over a large range but are relatively small in area.

The same types of clusters were also identified for spinel inclusions. However, spinel inclusion clusters were less frequent. In contrast to previous CLSM studies of spinel clustering,^[19,21,22] BSE images of spinel clusters indicated sintering of spinel inclusions. In addition, agglomeration of a spinel cluster and an alumina cluster was also found in some of the samples. Figure 13 presents an example of such a cluster. The non-globular inclusions in this cluster are alumina inclusions, while the globular inclusions are all spinels with some evidence of sintering between them. The algorithm identified these inclusions as a single cluster, it was from sample L1—heat 3.

Clusters of globular spinel inclusions were detected earlier in the ladle (L1 samples). Further ladle samples (L2 to L4) displayed clusters of non-globular spinel inclusions; an example is given in Figure 14. It should be noted that no Ca was detected in this cluster on this cross section.

The evolution of clusters from agglomerated globular to non-globular inclusion clusters was clearly observed for alumina inclusions, and to a lower extent with spinel inclusions. Previous studies by Braun^[1] and Tiekink^[15] on the morphology of alumina inclusions, suggest that the main factor controlling morphology is the oxygen activity. At relatively high oxygen activities (~ 100 of ppm), alumina inclusions tend to form globular inclusions, which are known to agglomerate and sinter. The work presented here encountered the same type of inclusions in the L0 samples. The oxygen contents at tap were 621, 717, and 492 ppm for heats 1, 2, and 3, respectively. Similar agglomeration of globular spinel inclusions was also observed; however, the oxygen concentration for these samples was relatively low (10 ppm or less). In addition, some of these globular spinel inclusions seemed to agglomerate and sinter (Figure 13—bottom right EDS map), countering some of the arguments made by earlier CLSM studies^[19,21,44] regarding the agglomeration and sintering of spinel inclusion.

Another type of cluster was also observed as shown in Figure 15. This cluster consisted of a central complex inclusion (calcium magnesium aluminat) surrounded by non-globular alumina. Other clusters like this were observed, but the large central feature was sometimes an alumina or spinel inclusion.

Such clusters, where micro-inclusions agglomerate around a single large inclusion, were common in the samples analyzed in this study. Figure 16 illustrates the extent of these clusters. The figure is a histogram of the largest inclusion size in each of the identified 116 clusters. Although most clusters were composed of relatively smaller inclusions (maximum inclusion area < 20 μm^2), 17 pct (20 clusters) had a maximum inclusion size greater than 100 μm^2 . There was no observed relationship between this type of cluster and sampling, since they were identified in all sample types except tundish samples, where only one small cluster was detected. This further corroborates the suggestion that macro-inclusions, in addition to being deleterious to the

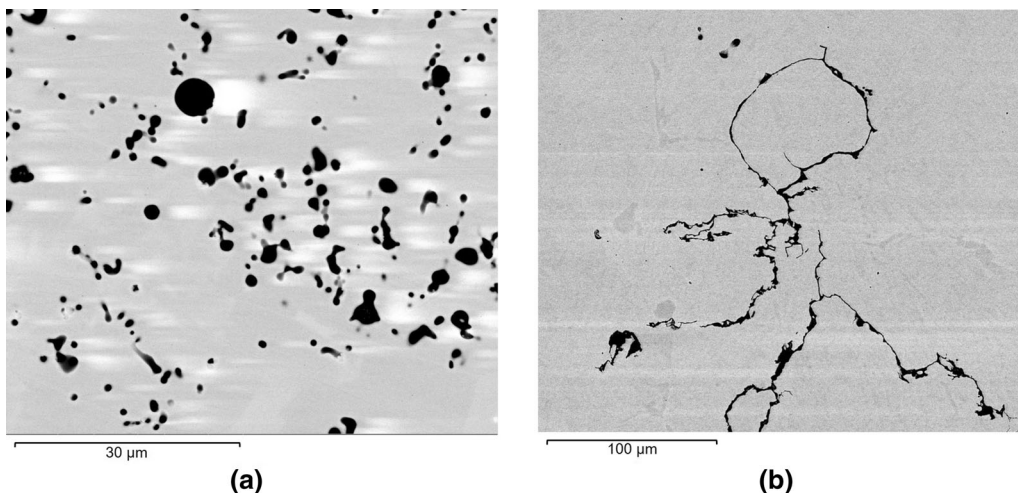


Fig. 12—BSE images of types of alumina clusters identified. (a) Agglomerated sintered globular alumina inclusions. (b) Non-globular alumina inclusions. Both images were taken at 20kV accelerating voltage.

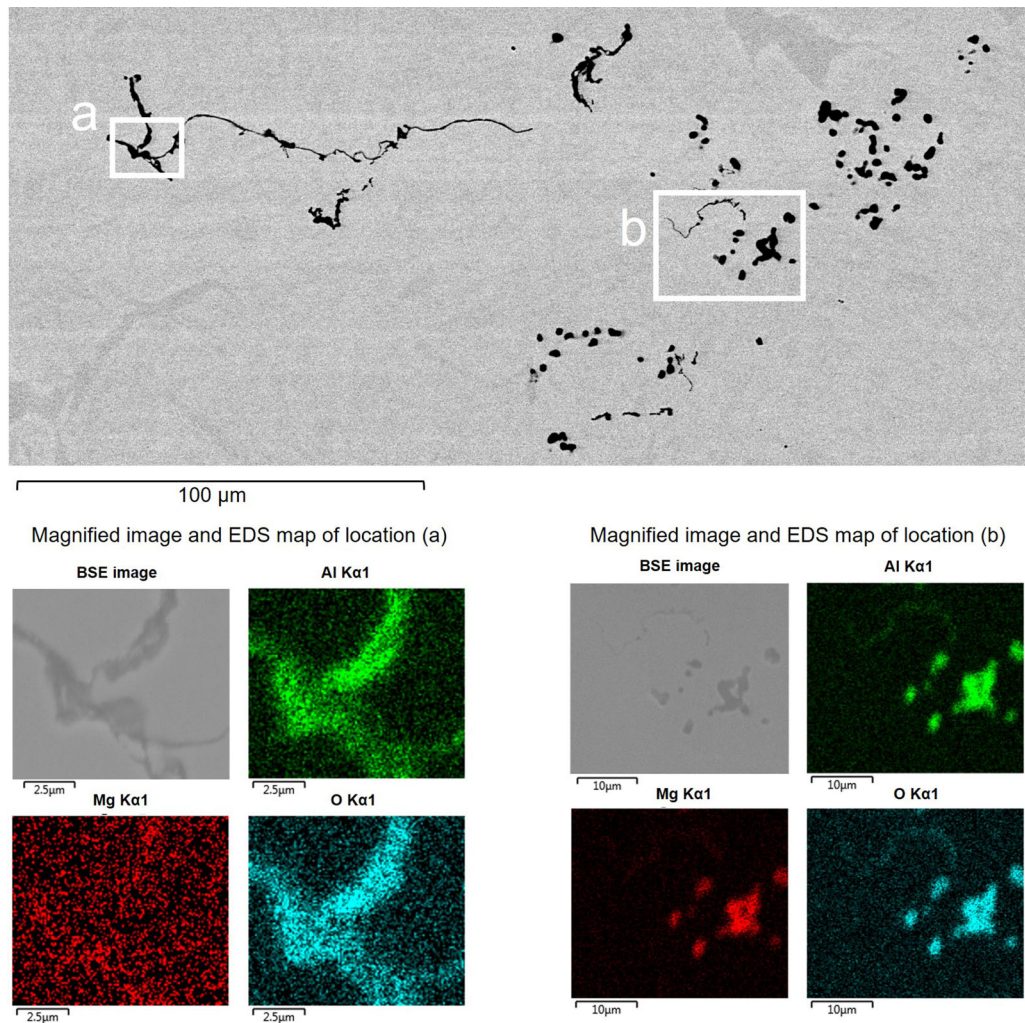


Fig. 13—BSE image of an inclusion cluster with clustered alumina (non-globular inclusions) and spinel (globular inclusions) inclusions (top). Magnified images and EDS maps of locations “a” (bottom left) and “b” (bottom right) of the BSE image on top.

mechanical properties of steel,^[24,45] can promote the agglomeration of additional smaller inclusions to form a larger inclusion cluster.^[5]

V. CONCLUSIONS

A density-based spatial clustering algorithm (DBSCAN) was applied to identify physical inclusion clusters from the output of inclusion SEM analysis. Initially, this approach was evaluated on two samples, a ladle and a tundish sample. Multiple serial sections were also analyzed from both samples to provide a 3D perspective on the distribution of inclusions. Once this tool was developed, it was applied to a series of samples from three aluminum-alloyed heats to investigate the evolution of clusters. The following observations were noted:

1. The work presented here can be readily applied to analyze the output generated from SEM AFA, providing an autonomous tool for inclusion cluster detection.
2. The stacking of serial sections enabled the analysis of clusters from a 3D perspective. Several clusters were identified in each 2D section of the ladle sample and none in the tundish sample. The results for the 3D analysis corroborated the findings in 2D with slight differences. The same clusters were identified in both 2D and 3D, with a few additional, but small, clusters identified in 3D only.
3. The cluster convex hull area was more indicative of the area in a sample affected by clustering, as opposed to the total inclusion area in a cluster.
4. Tap samples from all three heats displayed clustering of agglomerated globular alumina inclusions. Further in the ladle, alumina clusters were detected as elongated, non-globular inclusions. Spinel inclusions displayed a similar trend in cluster morphology, but fewer spinel clusters were observed. Agglomeration of globular spinel inclusions was observed in early ladle samples (L1), and BSE images suggested they sintered to form larger inclusions.

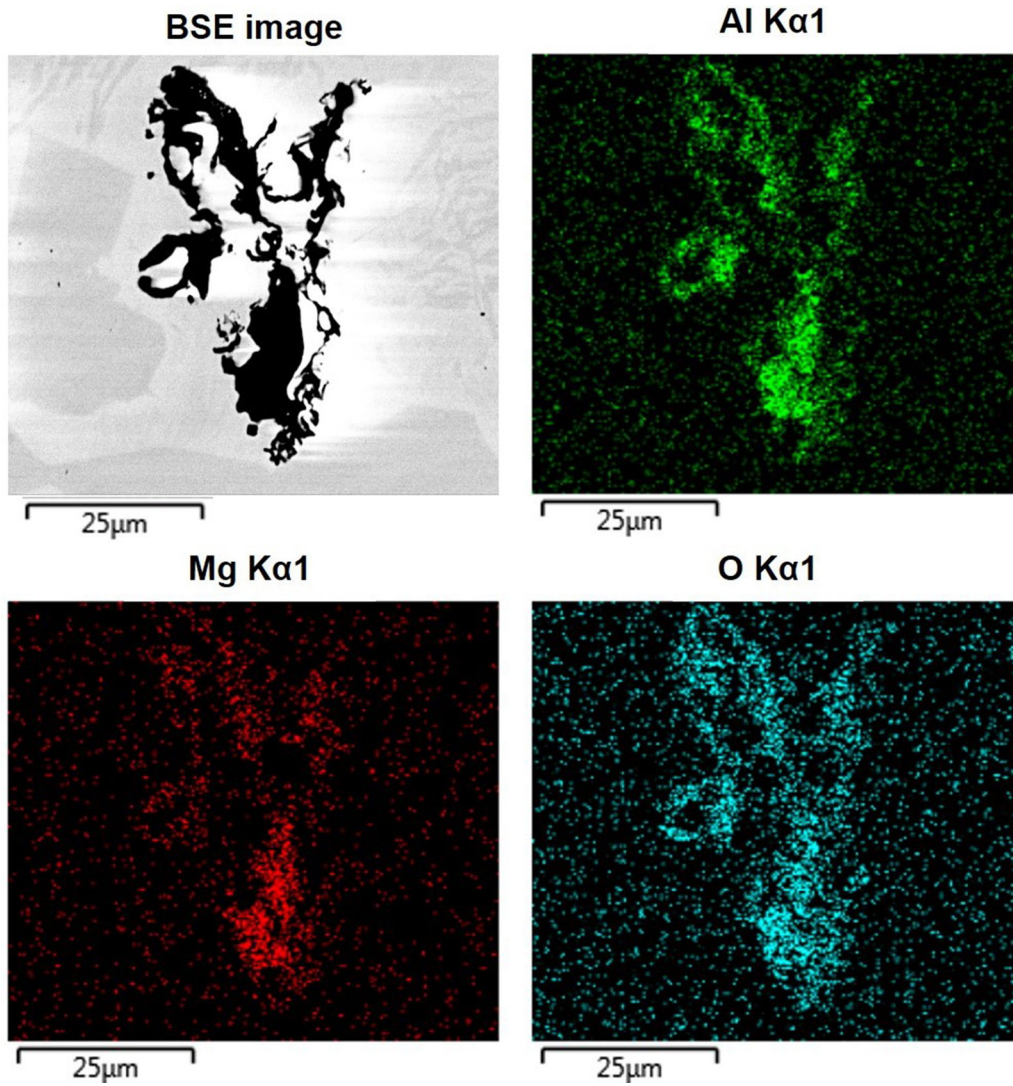


Fig. 14—BSE image and EDS map of a non-globular spinel inclusion cluster.

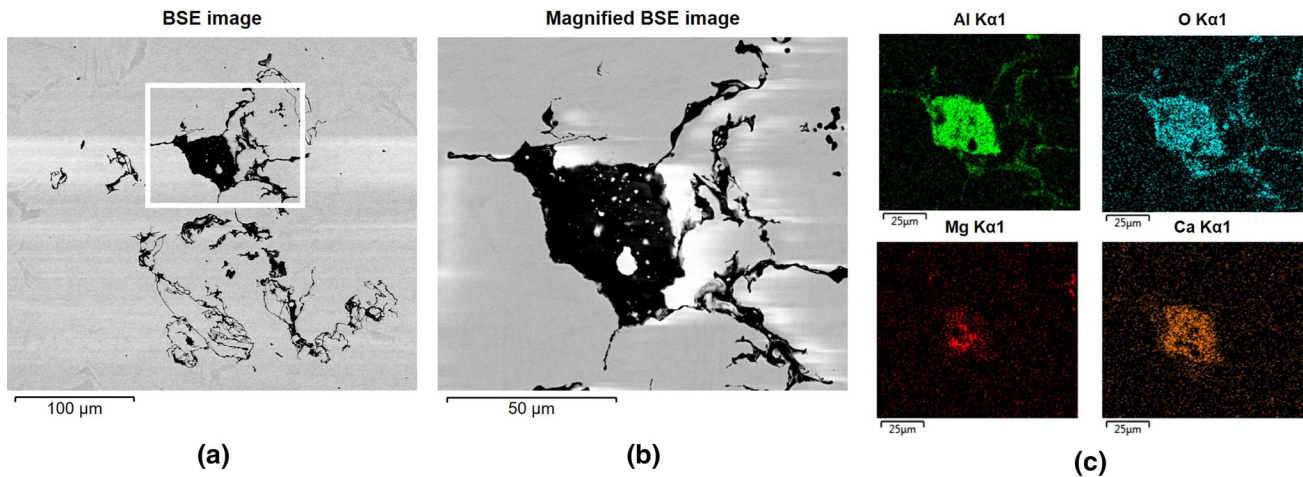


Fig. 15—(a) BSE image of agglomeration of small alumina inclusions around a large Mg-containing calcium alumina inclusion. (b) Magnified BSE image of the large inclusion in (a). (c) EDS mapping of BSE image in (b).

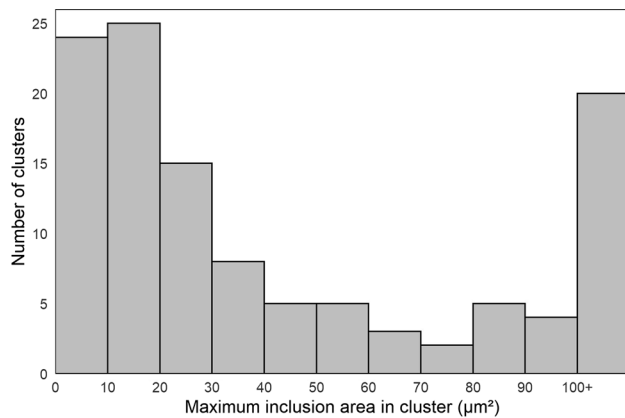


Fig. 16—Histogram of maximum inclusion area in each cluster.

- Approximately 20 pct of the clusters observed had the appearance of smaller micro-inclusions agglomerated around one large macro-inclusion.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support of the member companies of the Center for Iron and Steel-making Research and the use of the Materials Characterization Facility at Carnegie Mellon University, supported by grant MCF-677785.

REFERENCES

- T.B. Braun, J.F. Elliott, and M.C. Flemings: *Metall. Trans. B*, 1979, vol. 10, pp. 171–84.
- H. Yin, H. Shibata, T. Emi, and M. Suzuki: *ISIJ Int.*, 1997, vol. 37, pp. 946–55.
- A.L.V. Da Costa E Silva: *J. Mater. Res. Technol.*, 2018, vol. 7, pp. 283–99.
- S.N. Singh: *Met. Trans.*, 1974, vol. 5, pp. 2165–78.
- L. Zhang and B.G. Thomas: in *XXIV National Steelmaking Symposium2*, Morelia, Mich, Mexico, 2003, pp. 138–83.
- D. Janis, A. Karasev, R. Inoue, and P.G. Jönsson: *Steel Res. Int.*, 2015, vol. 86, pp. 1271–78.
- H. Bai and B.G. Thomas: *Metall. Mater. Trans. B*, 2001, vol. 32B, pp. 707–22.
- H.V. Atkinson and G. Shi: *Prog. Mater. Sci.*, 2003, vol. 48, pp. 457–520.
- L. Zhang: *Jom*, 2013, vol. 65, pp. 1138–44.
- R. Dekkers, B. Blanpain, P. Wollants, F. Haers, C. Vercruyssen, and B. Gommers: *Ironmak. Steelmak.*, 2002, vol. 29, pp. 437–44.
- T. Li, S.I. Shimasaki, S. Taniguchi, K. Uesugi, and S. Narita: *ISIJ Int.*, 2013, vol. 53, pp. 1943–52.
- T. Li, S.I. Shimasaki, S. Taniguchi, S. Narita, and K. Uesugi: *ISIJ Int.*, 2016, vol. 56, pp. 1989–95.
- M. Seleznev, K.Y. Wong, D. Stoyan, A. Weidner, and H. Biermann: *Steel Res. Int.*, 2018, vol. 89, pp. 1–9.
- L. Zhang and B.G. Thomas: in *7th Eur. Electr. Steelmak. Conf.*, 2002, vol. 2, pp. 77–86.

- W. Tiekink, R. Boom, A. Overbosch, R. Kooter, and S. Sridhar: *Ironmak. Steelmak.*, 2010, vol. 37, pp. 488–95.
- L. Zheng, A. Malfliet, P. Wollants, B. Blanpain, and M. Guo: in *Proc. 6th Int. Congr. Sci. Technol. Steelmak. ICS 2015*, 2015, vol. 55, pp. 731–36.
- H. Yin, H. Shibata, T. Emi, and M. Suzuki: *ISIJ Int.*, 1997, vol. 37, pp. 936–45.
- G. Du, J. Li, Z.B. Wang, and C. Bin Shi: *Steel Res. Int.*, 2017, vol. 88, pp. 1–9.
- D. Kumar: Carnegie Mellon University, 2018.
- S. Kimura, K. Nakajima, and S. Mizoguchi: *Metall. Mater. Trans. B*, 2001, vol. 32B, pp. 79–85.
- Y. Kang, B. Sahebkar, P.R. Scheller, K. Morita, and D. Sichen: *Metall. Mater. Trans. B*, 2011, vol. 42B, pp. 522–34.
- N. Dogan, R.J. Longbottom, M.H. Reid, M.W. Chapman, P. Wilson, L. Moore, and B.J. Monaghan: *Ironmak. Steelmak.*, 2015, vol. 42, pp. 185–93.
- P. Kaushik, J. Lehmann, and M. Nadif: *Metall. Mater. Trans. B*, 2012, vol. 43B, pp. 710–25.
- L. Zhang and B.G. Thomas: *ISIJ Int.*, 2003, vol. 43, pp. 271–91.
- E.B. Pretorius, H.G. Oltmann, and B.T. Scharf: in *AISTech*, Pittsburgh, 2013, pp. 333–45.
- M. Nuspl, W. Wegscheider, J. Angeli, W. Posch, and M. Mayr: *Anal. Bioanal. Chem.*, 2004, vol. 379, pp. 640–45.
- V. Singh, S. Lekakh, and K. Peaslee: in *SFSA Technical and Operating Conference, Steel Founders' Society of America (SFSA)*, 2008, pp. 1–18.
- S.R. Story, R.J. Fruehan, and M.S. Potter: *Iron Steel Technol.*, 2005, pp. 41–49.
- W. Zhao, F. Chen, H. Huang, D. Li, and W. Cheng: *Comput. Intell. Neurosci.*, 2021, <https://doi.org/10.1155/2021/5592878>.
- M. Cuartas, E. Ruiz, D. Ferreño, J. Setién, V. Arroyo, and F. Gutiérrez-Solana: *J. Intell. Manuf.*, 2020, vol. 32, pp. 1739–51.
- E. Ruiz, D. Ferreño, M. Cuartas, L. Lloret, P.M. Ruiz Del Arbol, A. López, F. Esteve, and F. Gutiérrez-Solana: *Metals*, 2021, vol. 11, p. 914. <https://doi.org/10.3390/met11060914>.
- M.A. Van Ende, M. Guo, E. Zinggrebe, B. Blanpain, and I.H. Jung: *ISIJ Int.*, 2013, vol. 53, pp. 1974–82.
- M.A. Van Ende, M.X. Guo, E. Zinggrebe, R. Dekkers, J. Proost, B. Blanpain, and P. Wollants: *Ironmak. Steelmak.*, 2009, vol. 36, pp. 201–08.
- J. MacQueen: in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, 1967, pp. 281–97.
- F. Nielsen: in *Introduction to HPC with MPI for Data Science*. Springer, 2016, pp. 195–211.
- M. Ester, H.-P. Kriegel, J. Sander, and X. Xu: *Kdd*, 1996, vol. 96, pp. 226–31.
- M.N. Vrahatis, B. Boutsinas, P. Alevizos, and G. Pavlides: *J. Complex.*, 2002, vol. 18, pp. 375–91.
- A.P. Dempster, N.M. Laird, and D.B. Rubin: *J. R. Stat. Soc. Ser. B*, 1977, vol. 39, pp. 1–22.
- M. Abdulsalam, M. Jacobs, and B.A. Webler: in *AISTech2020 Proceedings of the Iron and Steel Technology Conference*, AIST, 2020, pp. 1513–25.
- N. Verma, P.C. Pistorius, R.J. Fruehan, M.S. Potter, M. Lind, and S.R. Story: *Metall. Mater. Trans. B*, 2011, vol. 42B, pp. 711–19.
- E. Schubert, J. Sander, M. Ester, H.P. Kriegel, and X. Xu: *ACM Trans. Database Syst.*, 2017, vol. 42, p. 19.
- M. Hahsler, M. Piekenbrock, and D. Doran: *J. Stat. Softw.*, 2019, vol. 91, pp. 1–30.
- D. Tang, M.E. Ferreira, and P.C. Pistorius: *Microsc. Microanal.*, 2017, vol. 23, pp. 1082–90.
- N. Dogan, R. Longbottom, M.H. Reid, M.W. Chapman, and P. Wilson: *Chemeca: Challenging Tomorrow*, 2013, pp. 147–53.
- H.V. Atkinson, G. Shi, C.M. Sellars, and C.W. Anderson: *Mater. Sci. Technol.*, 2000, vol. 16, pp. 1175–80.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.