

# Data Mining Toward Increased Use of Aluminum Dross

Jiyoun Chang · Elsa Olivetti · Snorre Kjørstad Fjeldbo · Randolph Kirchain

Published online: 10 January 2015  
© The Minerals, Metals & Materials Society (TMS) 2015

**Abstract** Recycling provides a key strategy to move toward a more sustainable society by partially mitigating the impact of fast-growing material consumption. One main barrier to increased recycling arises from the fact that in many real world contexts, the quality of secondary (or scrap) material is unknown and highly variable. Even if scrap material is of known quality, there may be finite space or limited operational flexibility to separate or sort these materials prior to use. These issues around *identification* and *grouping* given the operational constraints create limitations to simply developing an appropriate sorting strategy, let alone implementing one. This study suggests the use of data mining as a strategy to manage raw materials with uncertain quality using existing data from the recycling industry. A clustering analysis is used to recognize the pattern of raw materials across a broad compositional range in order to provide criteria for grouping (binning) raw materials. This strategy is applied to an industrial case of aluminum recycling to explore the benefits and limitations in terms of secondary material usage. In particular, the case investigated is around recycling industrial byproducts (termed *dross* for the case of the

aluminum industry). The binning strategy obtained by the clustering analysis can significantly reduce material cost by increasing the compositional homogeneity and distinctiveness of uncertain raw materials. This result suggests the potential opportunity to increase low-quality secondary raw material usage before investment in expensive sorting technology.

**Keywords** Dross reprocessing · Aluminum recycling · Cluster analysis

## Introduction

The economic and environmental benefits of materials recovery, in the form of remanufacturing and recycling, are well documented and drive a 142 billion dollar industry in the US alone [1–4]. For recycling in particular, as a result of increased global manufacturing activity, the scrap material market has become more competitive, and recyclers are actively exploring alternative (generally lower quality) secondary materials. One key challenge with increased use of secondary materials drawn from a broader set of sources arises from the resulting increased variation in the quality of these streams.

Uncertain raw material quality has been one of the major barriers for efforts in recycling as well as remanufacturing [5–8]. Despite the fact that remanufacturing focuses on component recovery while recycling recovers materials, these two activities are similar in that the goal is to redirect or repurpose wastes as input resources. Although recycling and remanufacturing have different operational constraints originating from the nature of the different processes, blending and assembly, respectively, they share common difficulties in practice. Both recycling

---

J. Chang · E. Olivetti (✉)  
Department of Materials Science and Engineering,  
Massachusetts Institute of Technology, 77 Massachusetts Ave,  
Cambridge, MA 02139, USA  
e-mail: elsao@mit.edu

S. K. Fjeldbo  
Hydro Aluminium, Research & Technology Development  
(RTD), 6601 Sunndalsøra, Norway

R. Kirchain  
Engineering Systems Division, Massachusetts Institute of  
Technology, 77 Massachusetts Ave, Cambridge, MA 02139,  
USA

and remanufacturing processes are inherently subject to the uncertainty in input quality because secondary materials and components come from varied sources under varied conditions. Guide addressed the need for research in production planning and control of remanufacturing, given the uncertainties in recovered materials [9]. Thierry et al. pointed out that the benefits of remanufacturing and recycling processes can be maximized when companies are able to manage the quality of returned materials [8].

Many studies take the uncertainty of raw material quality into consideration in recycling and remanufacturing firms' decision-making processes. Previous work focused on improving the quality of secondary streams by increasing local homogeneity. Some studies explore different management strategies regarding material acquisition to manage the quality of returned products in remanufacturing [6]. Thierry et al. reported the case of a copier manufacturer with the strategy of reducing the types of materials in products in order to achieve a simplified and cost-effective recycling operation [8]. One of the most popular approaches to increase homogeneity of raw materials is sorting. Lund et al. perform an analysis of a centralized material-recovery facility to sort municipal solid waste using linear programming [10]. Research by Stuart and Lu demonstrates a model for reprocessing options for electronic scrap [11]. These papers explore the interaction between operational decision making and sorting in various contexts.

Sorting requires knowledge of the stream quality (identification) as well as a strategy to separate (group). In some cases, however, the composition of materials (i.e., the target objects of sorting) is unknown or not a constant. Galbreth and Blackburn consider the variability of used product conditions in remanufacturing and perform an analysis of optimal acquisition and sorting policies. The authors also point out the common assumption (often unsubstantiated) regarding homogeneous quality of returned products made in many other papers [12]. In most of these papers, there is only one decision around sorting: sort or not? There is no further discussion on the necessary degree of homogeneity of materials streams or required levels of sorting to achieve profitability. A recent study by Li et al. investigates the economic feasibility of separating various scraps into two categories, cast and wrought, and identifies the context that maximizes the benefit of sorting [13]. This study evaluates the impact of different recovery rates for cast and wrought scrap. However, the discussion around how to determine criteria to categorize raw materials and the effectiveness of these grouping methods has not been sufficiently addressed.

In this study, we suggest a way to improve the homogeneity of raw materials, using existing data from a recycler before investments are made into sorting technology. We propose a clustering analysis strategy to segment or

categorize raw materials, specifically dross from aluminum remelting, across a broad compositional space into a more homogeneous stream. This approach is shown for a case where the *identification* has been made but the method to *group* the raw materials is not clear.

### Reuse and Reprocessing of Dross

As mentioned above, recycler competition has led to interest in secondary materials from lower quality. Aluminum dross, which is a byproduct formed on the top of molten aluminum, has valuable entrapped metals, and could therefore be a good candidate for scrap that could see more intensive use. Environmental concern regarding dross disposal gives additional motivation for using it as a resource. The landfill disposal of dross is prohibited in European countries [14, 15] because dross can react with water and release explosive and noxious gases [16, 17]. Therefore, beneficial reuse or recycling of dross, instead of using expensive treatment methods to discard it, can offer both economic and environmental gains. Currently, dross management has involved metal recovery, either mechanically or chemically [18] as well as other repurposing for refractory materials, composites, and slag, among others [19, 20]. These latter methods have shown promise and lead to lower waste by volume due to salt management [21], but in some cases, metal recovery may be most beneficial either economically or environmentally [22]. Often metal recovery is done off site (termed tolling) in which an outside contractor processes dross and returns it to the remelter for a fee [23]. From an operational perspective, in order to use metal extracted from dross as a feed material to produce finished alloys, it must be first pre-processed in a rotary furnace. Recovering the metal from in-house dross (in other words, within the same facility where it was generated) may be especially beneficial since it has a composition similar to alloy products being made in that facility. This benefit can be maximized when the dross from a given alloy is used again to produce the same alloy. This study investigates in-house metal recovery from dross.

Achieving this goal, however, is challenging in practice. In a casthouse that produces multiple finished alloys, it is difficult to track the product from which each dross originates. Collecting dross separately by alloy product may provide a solution, but this requires as many dedicated lots for each type of dross as the number of products. Separate storage for dross is even more constrained for aluminum producers in Europe and Japan where storing dross outdoors is restricted due to the potential reaction with water. Therefore, this strategy is not practical in many cases and dross materials are combined before preprocessing. The significant loss in its economic value due to commingled

dross may be an issue not only for in-house processing, but also for off-site processing [22].

Rotary furnace operators may also add different scraps or dross from other sources to in-house dross in order to leverage the energy efficiency gains of operating a furnace at full capacity. Moreover, it is difficult to estimate the composition of dross from external sources until it is processed in the rotary furnace. This practice results in a situation in which the composition of output from the rotary furnace is different in every batch and potentially quite variable. If the measured composition of output after operating the rotary furnace happens to be similar enough to a product for the next batch in a melting furnace, it can be immediately used as hot liquid metal. Otherwise, it must be cast as an output (or sow) from the rotary furnace. Typically, the sows are aggregated and stored together within a facility.

Because of the myriad challenges described above resulting in increased compositional uncertainty, the use of the rotary furnace output may be limited to low-quality alloys (in other words, those with wider compositional specification). Considering the fact that some outputs could be valuable within a higher quality alloy if the composition fits well, aggregating is not the most efficient recycling strategy for in-house metal recovery from dross. However, for many plants, it is practically impossible to separate and store each output from every batch of the rotary furnace in an individual bin. A trade-off exists between having one aggregated bin and having several individual bins for each output from one batch of the rotary furnace, as described in Fig. 1a, b, respectively.

The former provides logistical simplicity but loses the collected compositional information of each sow by aggregating them and increases the uncertainty of raw materials for alloy production. The latter provides perfect information about raw materials for batch planning; however, separating is expensive and requires many lots or bins to store each material stream. It raises a question of how much information is enough or how much binning is enough for effective usage of cast sows.

The compositional characteristic of recovered dross is the key information for aluminum manufacturers who use it as a feed material because the composition of raw materials is directly related to profitability in alloy production. Previous research characterized the chemical and physical properties of aluminum dross and provides general estimates for those compositions [24, 25]. However, as several authors have pointed out, many factors influence the composition of dross. These factors include the skimming method, composition of the molten alloy, added salt flux composition, and dross-cooling process. The results from these studies suggest that the uncertainty from dross composition is unavoidable even if dross is separated by the original melt [17, 25].

In addition, the actual practices of dross processing, blending dross with scrap to maximize energy efficiency and using dross from external sources as explained above, increase the difficulty of consistently predicting accurate composition of outputs from the rotary furnace. Although many researchers have extensively tried to find optimal operational conditions to maximize recovery rates in the rotary furnace, which are important for profitable recycling operation [26–28], the question of how to increase the usage of recovered dross and scrap for alloy production in the actual operation setup has been unexplored.

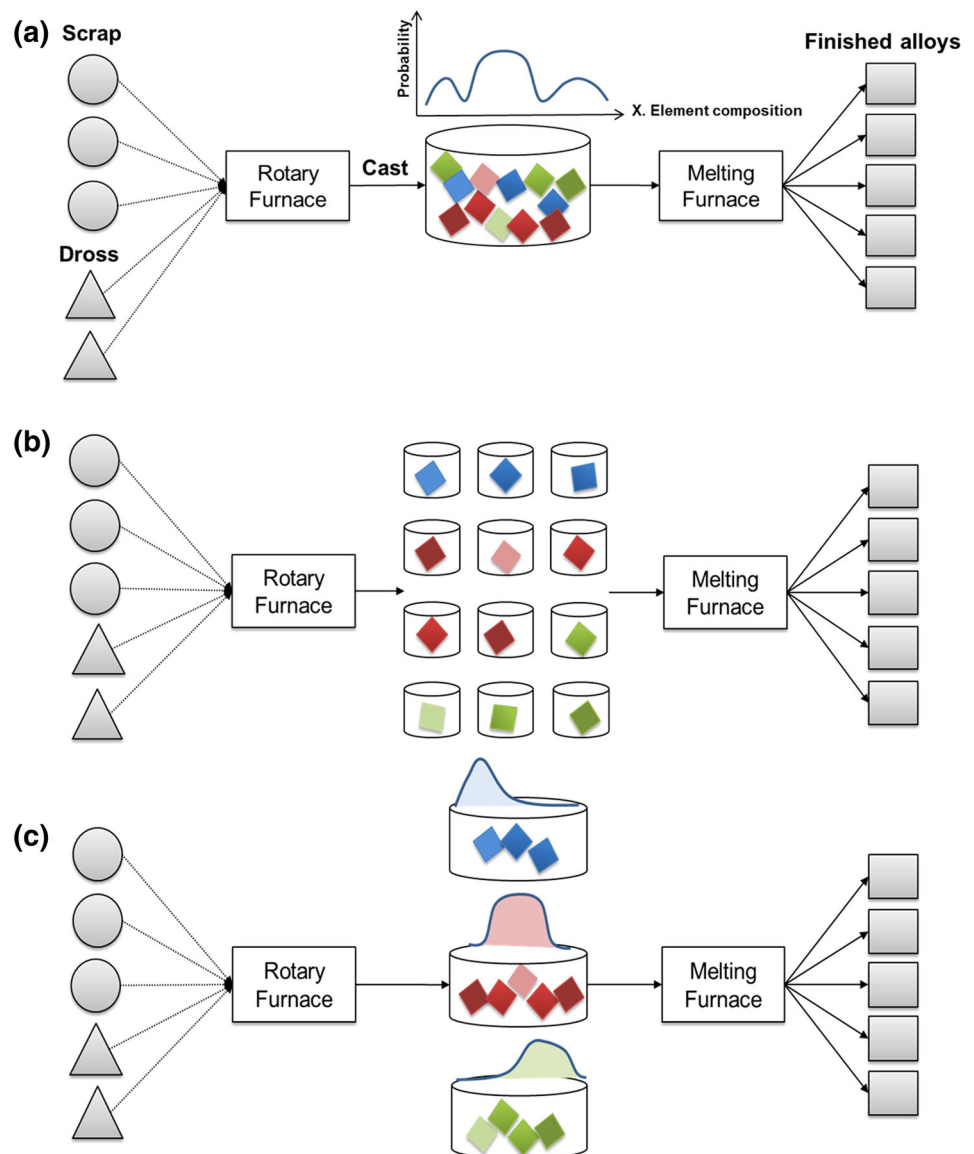
Figure 1 describes the general idea of our approach. In the suggested recycling operation, each output from the rotary furnace is assigned to a different bin based on its measured composition. Binning outputs from the rotary furnace as shown in Fig. 1c allows each bin to have relatively more similar raw materials compared to common recycling operations where all outputs from the rotary furnaces are mixed regardless of their composition as shown in Fig. 1a. Binning enables melting furnace operators to distinguish the specification of raw materials in different bins and use this information to model batches for the melting furnace. The clustering analysis in this study provides a way to define these bins.

Clustering analysis is one of several data mining methods that can be used to find patterns without any prior knowledge of what pattern exists. This method segments larger datasets into subsets, each of which are more homogeneous clusters of observations than the aggregate set as a whole. Therefore, clustering analysis can be used to recognize the patterns of raw materials with varied composition and group them into several categories.

Recent improvements in information gathering techniques in manufacturing allow firms to collect and store many types of data. Consequently, data mining has attracted attention as a tool for extracting information from these accumulated data pools [29]. However, applying data mining methods is less frequently done in manufacturing environments than in other areas such as finance or business. Several authors also point out that the use of accumulated data in manufacturing firms has been very limited, although the collected data embody valuable insights and knowledge [29, 30]. Many studies in recycling and remanufacturing that have employed historical data mostly focus on forecasting the expected outcome using statistical analyses [31, 32].

The approach of this study is motivated by actual practices in the recycling and remanufacturing industries. Most of these firms are likely able to acquire data about outputs from the first process, such as preprocessing at the rotary furnace or disassembly stage. Given the current common practice of measuring the composition of the outputs from the rotary furnace, data mining methods can

**Fig. 1** The diagram of aluminum dross recycling operation **a** current operation setup in which all cast sows are aggregated in the single bin **b** ideal operation setup where each output from the rotary furnace is individually binned **c** proposed operation setup where sows with relatively similar composition are binned together



provide valuable insights to improve current recycling operations.

In this article, we explore data mining of firms' operational data. There are two major questions this study addresses particular to the 'in-house' two-stage operation of dross reprocessing: (1) Does binning sows promote in-house recycling secondary materials? and (2) Is clustering analysis an effective method to categorize raw materials? These questions will be answered with an industrial case study.

## Methodology

### Cluster Analysis on Compositions of Sows

The cluster analysis method forms groups or clusters of similar records based on several measurements made on

these records. Among clustering methods, hierarchical algorithms are characterized as sequential clustering procedures, meaning each subsequent cluster cascades from the previous grouping. They can be categorized into two types of methods: agglomerative methods and divisive methods. Agglomerative methods start with a single point in each cluster and choose the pair of clusters to merge at each step based on the optimal value of an objective function until only one cluster is left. Divisive clustering methods are the reverse of the agglomerative methods. This category of methods begins with all data in one cluster and splits a cluster at each stage until each cluster has only single entity [33, 34].

Compared to partitioning algorithms that require the number of clusters a priori [33], hierarchical algorithms do not require any knowledge of the number of clusters. As a result, this category of algorithms produces the map of

hierarchy that represents the procedure by which clusters are merged or separated at every step, often described as a dendrogram or binary tree. The researcher can either use the entire hierarchy or select a level representing the specific number of clusters as needed [34].

The proposed binning strategy is analogous to the process of divisive hierarchical clustering methods. However, divisive hierarchical algorithms are not commonly used due to their computational complexity [33, 34]. We choose Ward’s minimum variance method in this study. Starting with many different clusters having only one object, this method finds the pair of clusters that leads to minimum increases in the total within-cluster variance at each step [35]. Since the goal of clustering analysis in this study is to reduce the uncertainty of raw materials, this method meets this goal. The distance between the two clusters A and B in Ward’s method is calculated as shown in Eq. (1)

$$D_{AB} = \frac{\|\bar{x}_A - \bar{x}_B\|^2}{\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}, \tag{1}$$

where  $\bar{x}_j$  is the center of cluster  $j$  and  $n_j$  is the number of points in it.

The historical composition data of outputs from the rotary furnace in a recycling facility that produces multiple alloy products for a 6-month period are used as clustering objects in this study. The commercial statistical software JMP is used to perform clustering analysis. Six elements of composition are chosen to calculate the distances because these elements are key components of alloy products in this facility. The six key elements are Si, Fe, Cu, Mn, Mg, and Zn. Also, other compositional elements vary relatively less. Although including other elements to calculate the distance between objects is possible, it decreases the contribution of these six elements to the overall distance. Therefore, using major alloy elements to calculate the distance leads to clearer distinctions between clusters for these elements.

### Chance-Constrained Batch Planning

In order to answer our research questions, it is essential to evaluate the impact of binned sows on their usage in a batch plan for finished alloy production. The goal of batch planning is to combine a variety of feeds such that the composition of their blend falls below maximum and above minimum targets. Therefore, this allowable range of final blends is often interpreted as a compositional window. We use the chance-constrained (CC) method to model batches for finished alloy production. The CC method, first introduced by Charnes and Cooper [36], allows users to explicitly specify the confidence level of each batch to meet the specifications of final products [37–39]. This

method provides an optimal batch plan based on the statistical parameters of the input materials. Due to its capability to control the batch error rate, it has been recently applied to the recycling area to model the blending operation for scrap with uncertain quality [37, 38].

The mathematical model of the blending problem with chance constraints can be written as below. Table 1 describes the nomenclature used throughout this article.

$$[\text{Obj}] \quad \min \sum_i \sum_m c_i x_{im} + \sum_l \sum_m c_l y_{lm} \tag{2}$$

Subject to

$$\sum_m x_{im} \leq A_i \quad \forall i \tag{3}$$

$$\sum_m y_{lm} \leq A_l \quad \forall l \tag{4}$$

**Table 1** Chance-constrained batch planning problem nomenclature

Indices	$i, j \in I$	Scrap and group of sows
	$l \in L$	Primary materials and alloying elements
	$m \in M$	Finished alloy products
	$k \in K$	Compositional elements
Parameters	$A_i$	Availability of scrap or group of sows $i$
	$A_l$	Availability of primary or alloying element $l$
	$c_i$	Cost of scrap or group of sows $i$
	$c_l$	Cost of a primary material or alloying element $l$
	$D_m$	The amount of demand for a finished alloy $m$
	$\varepsilon_{ik}$	The composition of an element $k$ in scrap or group of sows $i$
	$\bar{\varepsilon}_{ik}$	Mean of composition of element $k$ of scrap or group of sows $i$
	$\sigma_{(\varepsilon)ik}$	Standard deviation of composition of an element $k$ of scrap or group of sows $i$
	$\rho_{(\varepsilon)ijk}$	Correlation of composition of element $k$ between scrap or group of sows $i$ and $j$ $\rho_{(\varepsilon)ijk} = 1$ when $i = j$
	$\varepsilon_{lk}$	Composition of an element $k$ of primary or alloying element $l$
	$e_{mk}^{\max}$	The maximum specification of a compositional element $k$ in a finished alloy $m$
	$e_{mk}^{\min}$	The minimum specification of a compositional element $k$ in a finished alloy $m$
	$\alpha$	Confidence level to satisfy the maximum specification of finished alloy products
	$\beta$	Confidence level to satisfy the minimum specification of finished alloy products
	$X(\cdot)$	The inverse normalized cumulative Gaussian distribution function
	Variables	$x_{im}$
$y_{lm}$		The amount of primary or alloying element $l$ used in a finished alloy $m$



$$\sum_i x_{im} + \sum_l y_{lm} \geq D_m \quad \forall m \quad (5)$$

$$\Pr\left\{\sum_i \varepsilon_{ik} x_{im} + \sum_l \varepsilon_{lk} y_{lm} \leq \varepsilon_{mk}^{\max} D_m\right\} \geq \alpha \quad \forall m, k \quad (6)$$

$$\Pr\left\{\sum_i \varepsilon_{ik} x_{im} + \sum_l \varepsilon_{lk} y_{lm} \geq \varepsilon_{mk}^{\min} D_m\right\} \geq \beta \quad \forall m, k \quad (7)$$

$$x_{im} \geq 0 \quad \forall i, m \quad (8)$$

$$y_{lm} \geq 0 \quad \forall l, m \quad (9)$$

The objective function (2) is to minimize the sum of all raw material costs used in alloy production. Constraint (3) ensures that each raw material  $i$  with quality uncertainty, such as each scrap or sow group, is used in alloy products less than its availability,  $A_i$ . Similarly, constraint (4) limits the total amount of each primary material or alloying element used in alloy production to not more than its availability,  $A_i$ . Constraint (5) ensures that production volume of each alloy product satisfies demand. Constraints (6) and (7) enforce the maximum and minimum quality requirement for each final alloy product. Instead of two linear inequality constraints for quality requirement, the CC method requires those two inequality constraints to be satisfied with a given probability level,  $\alpha$  and  $\beta$ , where  $0 \leq \alpha, \beta \leq 1$ . Therefore, parameters  $\alpha$  and  $\beta$  represent likelihoods that the actual composition of blends will fall within the upper and lower limits of an alloy specification, respectively. Constraints (8) and (9) represent the non-negativity of decision variables.

Assuming that the compositions of raw materials with compositional uncertainty, which are indexed by  $i$ , follow a normal distribution, the two probabilistic constraints (6) and (7) can be transformed into their deterministic equivalents:

$$\begin{aligned} & \sum_i \bar{\varepsilon}_{ik} x_{im} + \sum_l \varepsilon_{lk} y_{lm} \\ & + X(\alpha) \left( \sum_i \sum_j \rho_{(e)ijk} \sigma_{(e)ik} \sigma_{(e)jk} x_{im} x_{jm} \right)^{\frac{1}{2}} \\ & \leq \varepsilon_{mk}^{\max} D_m \quad \forall m, k \end{aligned} \quad (10)$$

$$\begin{aligned} & \sum_i \bar{\varepsilon}_{ik} x_{im} + \sum_l \varepsilon_{lk} y_{lm} + X(1 - \beta) \\ & \left( \sum_i \sum_j \rho_{(e)ijk} \sigma_{(e)ik} \sigma_{(e)jk} x_{im} x_{jm} \right)^{\frac{1}{2}} \\ & \geq \varepsilon_{mk}^{\min} D_m \quad \forall m, k \end{aligned} \quad (11)$$

In this study, we assume that the compositions of raw materials follow a normal distribution. The compositional distribution of each bin of sows obtained from the cluster analysis varies with the element and the total number of bins, so this assumption may be limiting in some cases. The same six elements of composition used in the clustering analysis are tracked, and 99 % is used as a confidence level for the compositional constraint for each element.

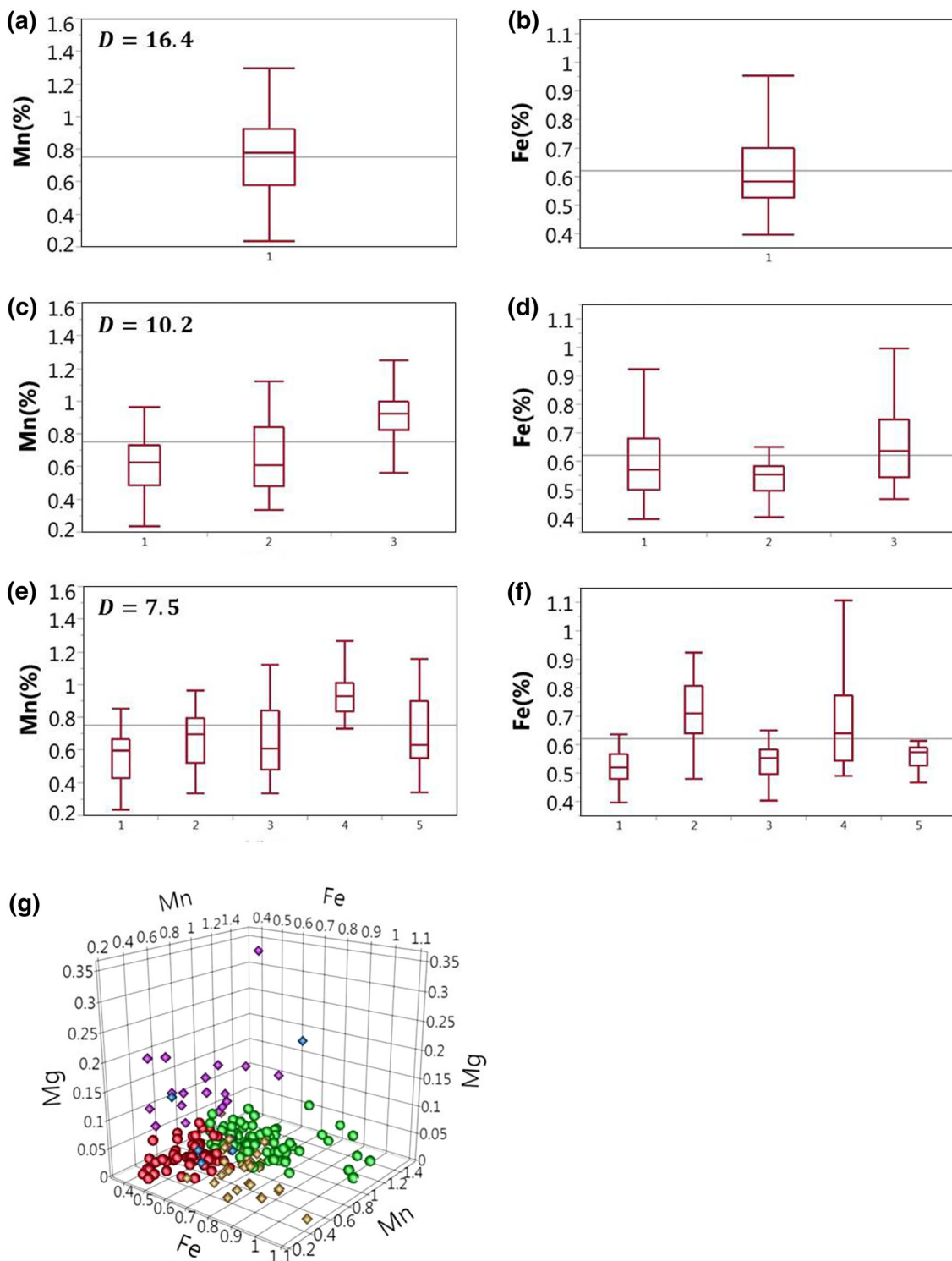
## Case Study

The case study in this article uses the operational setup of a casthouse located in Europe. This plant is equipped with rotary furnaces and melting furnaces. Eighteen different alloy products are produced in this facility. This means that there are eighteen different sources of dross. In addition to dross generated from alloy production, this casthouse uses primary metal, alloying elements, and eleven different scrap materials. After each batch of rotary furnace operation, the composition of recovered dross and scrap is measured before “being” cast as sow. We use the compositional measurement of 204 rotary furnace batches as the objects of clustering analysis in this study. The alloy products produced in each day vary dependent on demand and schedules. The performance of bins in blending operation is evaluated for various scenarios of demand for eighteen alloy products. The result presented in this paper is the one with the average performance.

## Result of the Clustering Analysis for Cast Sows

The clustering results can be obtained by cutting the dendrogram at different levels which represent the number of clusters. The result from each selected level of the dendrogram contains information about which sow belongs to which group. Each group can be interpreted as one separated bin for sows in the context of a production environment. Various levels are selected since there is no prior knowledge of which level will be most effective to indicate sows for use in alloy production. The compositional specification of each bin can be described by the statistical parameters, including the mean and standard deviation, of sows assigned in that bin.

Figure 2 shows the statistical characteristics of two of the compositional elements, Mn and Fe, with the selected numbers of bins as examples. Figure 2a, b represent the case when all raw materials are aggregated into one bin, which corresponds to the current operation at the case facility. The beginning of the clustering process, starting with only one object in each cluster (not included in Fig. 2), represents the opposite situation, where all outputs from batches of the rotary furnace are completely separated. In that case, the number of bins is equal to the number of batches in the rotary furnace. The compositional range of a bin is more distinctive as the number of clusters increases. For example, the one bin having compositional characteristics as shown in Fig. 2a, b is separated into three bins which are relatively characterized as low Mn and medium Fe, medium Mn and low Fe, high Mn and medium Fe, as shown in Fig. 2c, d. The ranges of different bins are not completely distinguishable because of the multi-



**Fig. 2** The statistical characteristics of each bin of sows with the selected number of bins, the case of one bin for **a** Mn and **b** Fe, three bins for **c** Mn and **d** Fe, and five bins for **e** Mn and **f** Fe. The number in the left top corner in **a**, **c**, and **e** represents the distance of bins, defined as total within-cluster variance, when the number of bins is

one, three, and five, respectively. This metric considers all six elements. **g** The scatter plot of sow compositions with limited elements, Fe, Mn, and Mg in the five-bin cases. Each *color* represents one bin, and each *dot* represents one output from the rotary furnace batch (Color figure online)

dimensionality of composition, which consists of six elements. Figure 2g represents how sows are assigned to different bins based on their composition in the case of five bins for three elements, Fe, Mn, and Mg. Each dot represents one batch output from the rotary furnace. Each color represents one bin. Dots with the same colors clearly congregate. This representation indicates that sows with relatively similar compositions are binned together. Although the graph is plotted with three elements, the distance between clusters is calculated based on all six elements.

### Performance of Bins in the Blending Operation

We evaluate the effects of varying the number of bins in daily batch planning of finished alloy production. Two performance metrics are used in this study: the percentage of the amount of sows used in production to the total available amount, and the ratio of material production cost to that of the current recycling operation where there is only one aggregated bin.

Figure 3 represents the result of a selected day's batch planning as an example. As the number of bins increases, more sows are incorporated to produce final alloys, replacing expensive primary metals and alloying elements. In the case of one bin, only 22 % of total available cast sows are used in alloy production, while all available cast sows are completely used in the case of ten bins. The production cost ratio of the 10-bin case to the single-bin case is 0.93.

As a benchmark, we also run the batch optimization model for the 204-bin case where individual sows are separately binned as Fig. 1b. This case, therefore, represents the situation in which there is no uncertainty associated with compositions of sows. In this case, 100 % of available cast sows are used in alloy production and its production cost ratio is 0.924. Compared to the 204-bin case, binning sows into ten bins by their compositions allows using the same amount of cast sows at a significantly lower number of bins and similar material cost. This result suggests that clustering by the compositions of sows is an effective binning strategy to increase usage of low-quality raw materials such as scrap and dross while reducing that of primary and alloying elements. This benefit is an incentive for material recyclers to maintain some compositional information from cast sows by grouping them into several categories rather than aggregating all of them.

Two different mechanisms explain the increase in performance with a higher number of carefully designed bins. The first mechanism is the reduced uncertainty of raw materials in each bin produced by binning sows. As

observed in Fig. 2, the composition of a bin for the case of a higher number of bins has a narrower distribution than in the case of a single bin. This reduced uncertainty of the sows allows use of more secondary raw materials, instead of using expensive primary metal or alloying elements. However, more use of cast sows, rather than other scrap, is attributed to their lower price. Second, a remelting furnace operator can take advantage of the more distinctive composition with the higher number of bins. In other words, the compositional distribution of each bin covers a relatively more distinct range and becomes more directly customized with particular products as the number of bins increases. For example, when the alloy specification is characterized as having high manganese content, one can reduce use of sows from the bin 1 and increase use of those from the bin 4 in the five-bin case if only considering the element manganese.

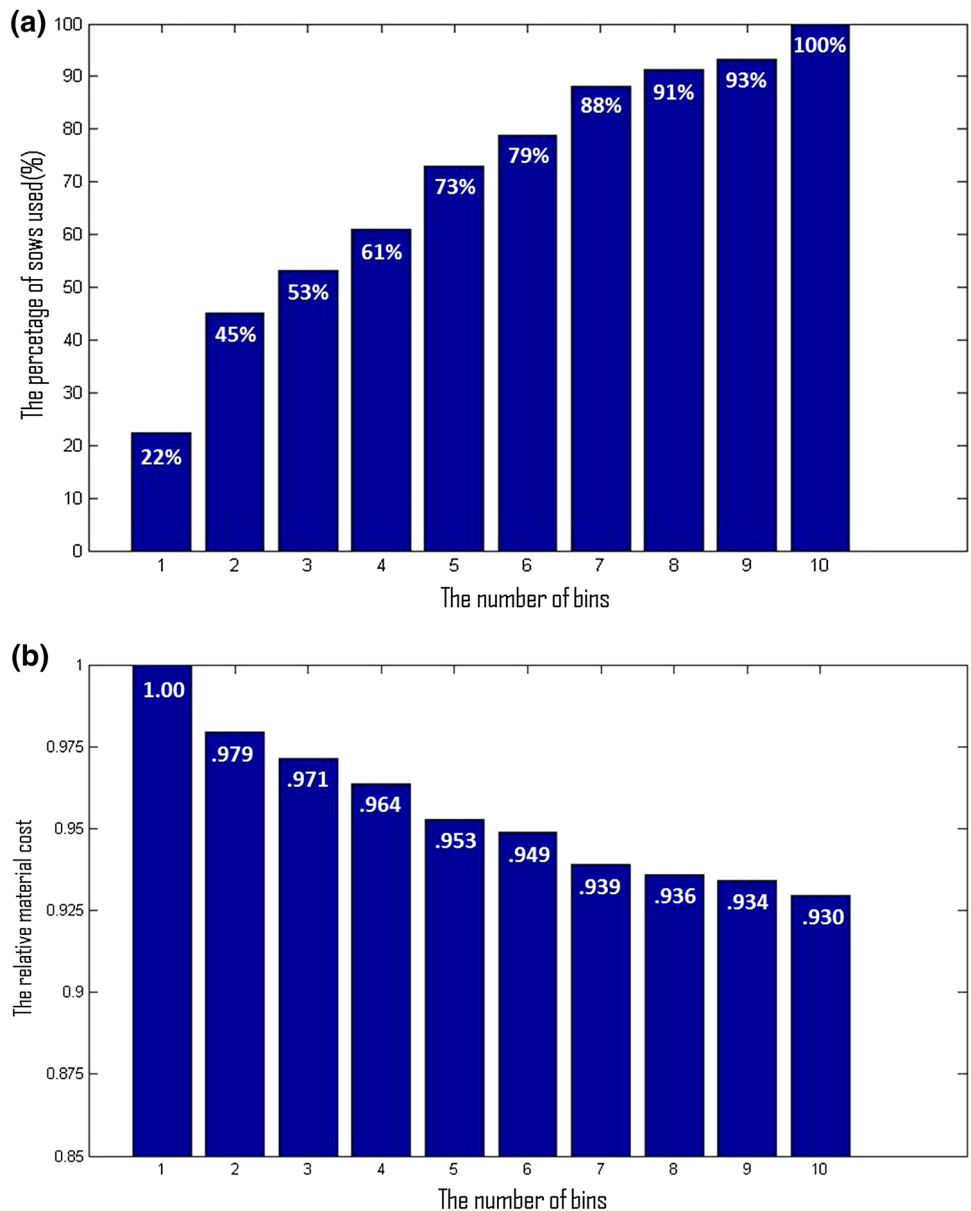
Understanding these mechanisms is easier if we look at constraints for the maximum and minimum specification requirements in the CC batch optimization model. Mathematically, the first benefit from the reduced uncertainty is related to the second term in Eqs. (10) and (11). It should be noticed that  $X(\alpha)$  is a positive number, whereas  $X(1 - \beta)$  is negative. These second terms play a role in narrowing the window of alloy specification depending on the compositional uncertainty of raw materials. The second term in Eq. (10),  $X(\alpha) \left( \sum_i \sum_j \rho_{(\varepsilon)ijk} \sigma_{(\varepsilon)ik} \sigma_{(\varepsilon)jk} x_{im} x_{jm} \right)^{\frac{1}{2}}$ , lowers the maximum limit of specification according to statistical parameters of uncertain raw materials and their usage. The terms in Eq. (11),  $X(1 - \beta) \left( \sum_i \sum_j \rho_{(\varepsilon)ijk} \sigma_{(\varepsilon)ik} \sigma_{(\varepsilon)jk} x_{im} x_{jm} \right)^{\frac{1}{2}}$ , elevate the minimum of specification. When the aggregated bin splits into two separate bins, the standard deviations of these bins become smaller. The smaller standard deviation of each newly formed bin,  $\sigma_{(\varepsilon)ik}$  or  $\sigma_{(\varepsilon)jk}$ , results in broadening the width of the given windows when the same amounts of raw materials are used. This broadening allows incorporating more raw materials with uncertainty into a batch plan if other conditions are unchanged.

The second mechanism of improved performance is relevant to the first term in Eqs. (10) and (11). As a consequence of clustering, the values of  $\bar{\varepsilon}_{ik}$  of bins are adjusted depending on the compositions of assigned raw materials. A remelting furnace operator can accordingly differentiate the composition of sows in different bins so that sows in each bin can be more customized to alloy products with similar compositions.

As the number of bins increases, the marginal increase in sow usage generally decreases, but not necessarily monotonically. The fact that the benefits of binning sows originate from two different mechanisms explains this



**Fig. 3 a** The percentage of the amount of sows used in alloy production to total available amount **b** the ratio of material production cost to that of the single bin case with the different number of bins



behavior. The benefit from the first mechanism, the reduced compositional uncertainty, becomes less significant as the number of bins increases. This can be explained relative to the clustering procedure. An agglomerative hierarchical clustering process starts with merging two most similar objects into one cluster that results in the least increase in the total within-cluster variation. In general, merging the final two clusters into one leads to the greatest increase in total within-cluster variation because these last two clusters are most unlike. Therefore, we see a large benefit going from one bin to two bins and so on. In other words, the first split results in the greatest decreases in compositional variation within each bin and further binning has a diminishing decrease in compositional uncertainty.

Once the compositional distribution of sows in each bin becomes smaller than the final alloy specification window, the sows in that bin can be fully utilized. Eventually, at a certain stage in the binning process, all available sows can be completely used in alloy production. However, the benefit from the second mechanisms, a more distinctive composition, is not necessarily related to the number of bins. For example, because the final alloy specification is not used in the clustering, there is no guarantee that the mean composition of sows in a bin for a two-bin case is more customized to the specification of alloy products than that of a bin in ten-bin case. Therefore, whether the benefit from the second mechanism is significant or not depends on the final alloy specification. Overall, the optimal number

of bins that allows complete usage is determined by the relationship between the statistical characteristics of bins at each level of clustering and the final alloy specification.

Although we observe improved performance from binning sows by employing the CC method to determine batch recipes, the mechanisms described above imply that similar results could be observed in other batch modeling approaches that consider uncertainty.

### Economic Analysis

Binning sows by clustering based on their compositions increases the homogeneity of raw materials available for production. Since sorting materials can be defined as an activity to separate the mixture of different materials into more homogeneous sets, this strategy can be considered as a different way to sort materials.

In that context, binning is an effective method to boost usage of cast sows because it improves the uniformity of raw materials. Since purchasing raw materials is one of the major cost factors for material manufacturers, the substitution of this new secondary material for expensive primary material can bring significant economic benefits. However, obtaining more compositional information on sows requires firms to purchase additional property (e.g., land) to accommodate sorting and storage in order to separate existing raw materials. Therefore, it is certainly meaningful for recycling firms to weigh the capital cost of bin setup versus the benefit of sorting. We perform a simple analysis to evaluate the expected economic benefits of the binning strategy based on our results of the daily batch planning above. Several assumptions are made for the purpose of this analysis. It is assumed that the rate of material substitution of cast sows for primary materials will be similar to our results above throughout the payback period. We also assume that there is no additional cost, such as a maintenance cost, other than the fixed cost to purchase a lot for storing raw materials. Other parameters used in the analysis are summarized in Table 2.

Table 3 represents the present value of total material cost savings over 3 years if bins are added to separate raw materials, compared to the current operation, which is equivalent to the single-bin case. For example, with an addition of two bins in which cast sows are separated into three groups, the expected cost saving from material substitution is US\$7.6 million or US\$3.8 million per bin. This value suggests an upper limit at which firms can invest to set up additional bins. Therefore, the benefit of adding a bin becomes less significant as the number of bins grows because the average cost saving per bin decreases with the increase in the number of added bins as shown in Table 3.

**Table 2** Parameters used to calculate the expected cost saving of binning strategy

Operation days per year	240 days/year
Discount rate	10 %
Payback period	3 years

**Table 3** Total expected material cost saving and average expected material cost saving per bin with the different number of bins

The number of added bins (total number of bins)	Total material costs savings	Average cost saving per bin
+1 (2)	(5.5)	(5.5)
+2 (3)	(7.6)	(3.8)
+3 (4)	(9.6)	(3.2)
+4 (5)	(12.5)	(3.1)
+5 (6)	(13.5)	(2.7)
+6 (7)	(16.2)	(2.7)
+7 (8)	(17.0)	(2.4)
+8 (9)	(17.5)	(2.2)
+9 (10)	(18.7)	(2.1)

Units: million dollar

In reality, expanding storage places for raw materials is often complicated and contextual. The cost of expansion varies from firm to firm and determining the size of lots for raw materials depends on many different factors. The optimal number of bins to bring firms the largest economic benefits must be chosen after careful consideration of the expected cost saving from material substitution as well as the capital costs needed to expand inventory spaces. However, the simple economic analysis in this study suggests that binning raw materials by their composition allows firms not only to increase the usage of low-quality raw materials in their production but also to realize an economic benefit.

Although we use one of the hierarchical algorithms in this study due to limited knowledge about the number of bins, different methods can be adapted depending on production environment. For instance, if a recycling firm knows the maximum amount of resources it has available to devote to expand raw material inventory, it can start with that number using popular partitioning algorithms such as k-means.

However, the goal of this study is to address a new opportunity presented by clustering analysis to increase the usage of low-quality raw materials while accounting for uncertainty. Also, clustering methods are heuristic in nature. A suitable clustering method can be changed based on characteristics of data or the goal of a study. Therefore, the effect of choosing different cluster algorithms is not the scope of this study. The relationship between clustering

methods and data are well reviewed in [34]. The researcher can choose appropriate clustering methods depending on the structure of targeting data and the context of manufacturing environment.

## Conclusion

Clustering analysis is used to recognize patterns in raw materials with compositional uncertainty. Such patterns provide criteria for separating raw materials to increase in their homogeneity. Binning cast sows by clustering analysis allows for full utilization of sows in alloy production without the need to maintain compositional information of all individual outputs from the rotary furnace. Therefore, clustering analysis is an effective method to separate raw materials. The results in this study suggest a new opportunity for material recyclers to maximize the usage of low-quality raw materials in alloy production using existing data. This new approach can be used not as an alternative but as a complement to the existing modeling tools and recycling technology.

**Acknowledgments** This work was supported by Norsk Hydro. The authors wish to acknowledge the support of the plant managers and operators for their input in the study, Robert Grønli, Johan Portin, and Jørgen Roar Jenssen.

## References

- Das SK (2011) Aluminum recycling in a carbon constrained world: observations and opportunities. *JOM* 63(8):137–140
- European Aluminium Association and Organisation of European Aluminium Refiners and Remelters (2006) Aluminium recycling in Europe: the road to high quality products
- Harris Z (2014) Scrap value: increasing industrial production and downstream demand will support revenue growth. *Recyclable Material Wholesaling in the US*, IBISWorld Industry Report. 42393
- United States International Trade Commission (2012) Remanufactured goods: an overview of the U.S. and global industries, markets, and trade. USITC Publication 4356, Investigation No. 332-525
- Ferrer G, Whybark DC (2001) Material planning for a remanufacturing facility. *Prod Oper Manag* 10(2):112–124
- Guide VDR, Van Wassenhove LN (2001) Managing product returns for remanufacturing. *Prod Oper Manag* 10(2):142–155
- Peterson RD (1999) Scrap variability and its effects on producing alloys to specification. *Light Met* 1999:855–859
- Thierry M, Salomon M et al (1995) Strategic issues in product recovery management. *Calif Manag Rev* 37(2):114–135
- Guide VDR Jr (2000) Production planning and control for remanufacturing: industry practice and research needs. *J Oper Manag* 18(4):467–483
- Lund JR, Tchobanoglous G et al (1994) Linear-programming for analysis of material recovery facilities. *J Environ Eng* 120(5):1082–1094
- Stuart JA, Lu Q (2000) A refine-or-sell decision model for a station with continuous reprocessing options in an electronics recycling center. *IEEE Trans Electron Packag Manuf* 23(4):321–327
- Galbreth MR, Blackburn JD (2006) Optimal acquisition and sorting policies for remanufacturing. *Prod Oper Manag* 15(3):384–392
- Li P, Dahmus J et al (2011) How much sorting is enough identifying economic and scrap-reuse benefits of sorting technologies. *J Ind Ecol* 15(5):743–759
- Council (2000) Commission Decision 2000/532/EC of 3 May 2000. *Official Journal of the European Communities* L226. E. Union: 3–24
- Council (2000) Council Directive 91/689/EEC of 12 December 1991 on hazardous waste *Official Journal of the European Communities* L377 E. Union: 20–27
- Prillhofer R, Prillhofer B et al (2009) Treatment of residues during aluminum recycling. TMS Annual Meeting, San Francisco, CA
- Xiao YP, Reuter MA et al (2005) Aluminium recycling and environmental issues of salt slag treatment. *J Environ Sci Health A Tox Hazard Subst Environ Eng* 40(10):1861–1875
- David E, Kopac J (2013) Aluminum recovery as a product with high added value using aluminum hazardous waste. *J Hazard Mater* 261:316–324
- Bajare D, Korjakins A et al (2012) Pore structure of lightweight clay aggregate incorporate with non-metallic products coming from aluminium scrap recycling industry. *J Eur Ceram Soc* 32(1):141–148
- Dai C, Apelian D (2012) Reuse of aluminum dross as an engineered product. In: *Proceedings of the 13th international conference on aluminum alloys (Icaa13)*, pp 1315–1320
- David E, Kopac J (2012) Hydrolysis of aluminum dross material to achieve zero hazardous waste. *J Hazard Mater* 209:501–509
- Nakajima K, Osuga H et al (2007) Material flow analysis of aluminum dross and environmental assessment for its recycling process. *Mater Trans* 48(8):2219–2224
- Peterson RD (2013) Metal contamination associated with dross processing. In: *Light Metals 2013*. John Wiley & Sons, Inc., pp 941–946
- Kevorkijan V (2002) Evaluating the aluminum content of pressed dross. *JOM* 54(2):34–36
- Manfredi O, Wuth W et al (1997) Characterizing the physical and chemical properties of aluminum dross. *JOM* 49(11):48–51
- Tzonev T, Lucheva B (2007) Recovering aluminum from aluminum dross in a DC electric-arc rotary furnace. *JOM* 59(11):64–68
- Unlu N, Drouet MG (2002) Comparison of salt-free aluminum dross treatment processes. *Resour Conserv Recycl* 36(1):61–72
- Zhou B, Yang Y et al (2006) Modelling of aluminium scrap melting in a rotary furnace. *Miner Eng* 19(3):299–308
- Wang K (2007) Applying data mining to manufacturing: the nature and implications. *J Intell Manuf* 18(4):487–495
- Choudhary AK, Harding JA et al (2009) Data mining in manufacturing: a review based on the kind of knowledge. *J Intell Manuf* 20(5):501–521
- Clotey T, Benton WC et al (2012) Forecasting product returns for remanufacturing operations. *Decis Sci* 43(4):589–614
- Goh TN, Varaprasad N (1986) A statistical methodology for the analysis of the life-cycle of reusable containers. *IIE Trans* 18(1):42–47
- Milligan GW, Cooper MC (1987) Methodology review—clustering methods. *Appl Psychol Meas* 11(4):329–354
- Xu R, Wunsch D (2005) Survey of clustering algorithms. *IEEE Trans Neural Networks* 16(3):645–678
- Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58(301):236
- Charnes A, Cooper WW (1959) Chance-constrained programming. *Manage Sci* 6(1):73–79

37. Gaustad G, Li P et al (2007) Modeling methods for managing raw material compositional uncertainty in alloy production. *Resour Conserv Recycl* 52(2):180–207
38. Olivetti EA, Gaustad GG et al (2011) Increasing secondary and renewable material use: a chance constrained modeling approach to manage feedstock quality variation. *Environ Sci Technol* 45(9):4118–4126
39. Shih JS, Frey HC (1995) Coal blending optimization under uncertainty. *Eur J Oper Res* 83(3):452–465