9th Global Innovations Symposium: Trends in Integrated Computational Materials Engineering for Materials Processing and Manufacturing <*Edited by…*>
(TMS) The Minerals, Metals & Materials Society, 2008

Modeling Methods to Guide Recycling Friendly Alloy Design: The Impact of Compositional Data Structure

Gabrielle Gaustad[1], Ray Peterson[2], Randolph Kirchain[1,3]
Materials Systems Lab, [1]Department of Materials Science and Engineering and [3]Engineering Systems Division, MIT, Massachusetts Institute of Technology
Room E40-442, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[2]Aleris International, Inc., 397 Black Hollow Rd., Rockwood, TN 37854, USA

Keywords: chance constrained optimization, aluminum remelting, uncertainty, batch optimization

Abstract

Recyclability is a context dependent materials property that has proven challenging to evaluate. However, previous studies by the authors have shown that certain modeling techniques, specifically chance constrained optimization, can provide a quantitative evaluation of recyclability. This information can be used to accelerate the evaluation and design of alloys (as well as products and operations) for improved recycling. One assumption in this work, however, is that the compositional variability of materials would be normally distributed. Previous work would indicate this may not be the case for all recycled materials. Performance, in terms of recyclability, may then be heavily influenced by the structure of the compositional data used. To test this, specific case studies involving recycled material streams were investigated. Results indicate that the underlying structure/distribution of the compositional data can have a profound effect on alloy design decisions to encourage usage of recycled materials.

Introduction

Modeling tools are available and broadly applied to support the decisions of secondary batch planners. Many producers make use of linear optimization (LO) techniques[1]. These LO models can improve decisions not only about raw materials mixing, but also purchasing, upgrading, and sorting of secondary materials[2-5]. The authors have previously proposed modifications[6] to traditional LO models used for batch planning that allow decisions to be made using more of the statistical information available. Specifically, these modifications are referred to as a chance-constrained (CC) formulation. The CC formulation allows a model to be constructed that embeds not only mean-based information, as with traditional LO models, but also some measure of dispersion (i.e., sample variance). This method has been shown to provide a quantitative evaluation of performance in regards to efficient use of recycled materials. However, to date, this modified formulation and the benefit it provides have only been explored in cases where the compositional variability of materials was assumed to be normally distributed. Previous work characterizing aluminum scraps[7] would indicate this may not be the case for all recycled materials. Performance, in terms of recyclability, may then be heavily influenced by the structure of the compositional data used.

This paper will statistically characterize the compositional distributions of five representative aluminum scraps in order to determine the accuracy of a normal distribution assumption. The consequences of these assumptions on production errors and finished alloy compositions will be examined using the CC optimization framework and Monte Carlo compositional simulations.

Methods

Curve Fitting Statistics

A chi-squared goodness of fit statistical test was performed on compositional data for five scrap categories collected over a ten month period. The chi-squared test determines how closely a real set of data follows a theoretical distribution with the metric χ^2 defined in Eq. 1, where O is an observed frequency and E is an expected frequency for the distribution of interest. The number of samples or observations is n.

$$\chi_o^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

Chance Constrained Optimization Model

Stochastic programming methods including chance constrained variants were first formulated by Charnes and Cooper[8] as mechanisms to embed a more rich set of statistical information into optimization based decision models. This method relates the desired level of confidence to the underlying standard deviations observed in the sampled raw materials. With the understanding that the compositional constraints will not be satisfied always due to inherent uncertainty, they can be rewritten as probabilistic expressions and transformed into their deterministic equivalents. For mathematical details, please refer to[6].

Monte Carlo Simulation

Monte Carlo simulations were executed to test the compositional acceptability of proposed batch mixing solutions provided by the model against scraps of varying composition. These simulations were carried out using Crystal Ball, an Excel based program. The Monte Carlo method uses pseudo-random numbers (i.e. not truly random in the sense that they are generated by numerical algorithms) to statistically simulate random variables. For this study, the top four best fitting distributions around the compositional mean of six of the scraps elemental considerations (Si, Mg, Fe, Mn, Cu, Zn) was assumed and the optimal solution was tested 25,000 times. The number of batches that had any errors (i.e. final composition of finished alloys fell out of specification) was recorded.

Case Study

Compositional data on five aluminum scraps was collected over an extended period of time in order to characterize their compositional distribution. These data sets tracked 19 elements with sample sizes ranging from 98 values to 225 values. A hypothetical aluminum secondary production case study was devised in order to test the consequences of distribution assumptions on production batch planning. Six representative alloys were chosen for production; their maximum and minimum compositional constraints(Table I) are based on guidelines set by the Aluminum Association and do not reflect production targets of any specific firm. Scrap types were based on those common to the scrap market and compositions were taken from EU standards(Table II). Prices for primary aluminum and alloying elements were taken from USGS 2005 averages[9] while scrap prices were estimated from [10].

Table I. Maximum and Minimum Compositional Specifications for Finished Alloys in weight fraction[11]

	Si Max	Si Min	Mg Max	Mg Min	Fe Max	Fe Min	Cu Max	Cu Min	Mn Max	Mn Min	Zn Max	Zn Min
2219	0.002	0	0.0002	0	0.003	0	0.068	0.058	0.004	0.002	0.001	0
3004	0.003	0	0.013	0.008	0.007	0	0.0025	0	0.015	0.01	0.0025	0
4032	0.135	0.11	0.013	0.008	0.01	0	0.013	0.005	0.005	0	0.0025	0
5052	0.0025	0	0.028	0.022	0.004	0	0.001	0	0.001	0	0.001	0
6061	0.008	0.004	0.012	0.008	0.007	0	0.004	0.0015	0.0015	0	0.0025	0
7075	0.004	0	0.029	0.021	0.005	0	0.02	0.012	0.003	0	0.061	0.051

Table II. Average compositions for scraps in weight fraction[12]

Scraps	Si	Mg	Fe	Cu	Mn	Zn
UBC	0.00225	0.00975	0.00375	0.0015	0.00825	0.000375
Mixed auto castings	0.10125	0.00225	0.00825	0.02625	0.00375	0.009
Cu- alum radiator	0	0	0.00525	0.3	0	0
Wire and cable scrap	0.001875	0.0045	0.003	0.000375	0.000375	0.000525
Mixed turnings	0.0675	0.00225	0.0075	0.02625	0.00375	0.01125
Litho sheets	0.006	0	0.006375	0.00125	0.006375	0.00075

Results and Discussion

Characterizing Compositional Distributions

Qualitative examination of the compositional distributions of the five aluminum scraps showed that there was a wide range in types of probability distributions. Some normal distributions existed; however, most were not normally distributed as hypothesized. To illustrate these differences, two example frequency histograms for compositions in a scrap are shown in **Figure 1**. One, **Figure 1**(A) where the compositions are clustered around the mean, suggest a normal distribution and another, **Figure 1**(B), where the values are clustered around the minimum values, suggest a skewed probability distribution.

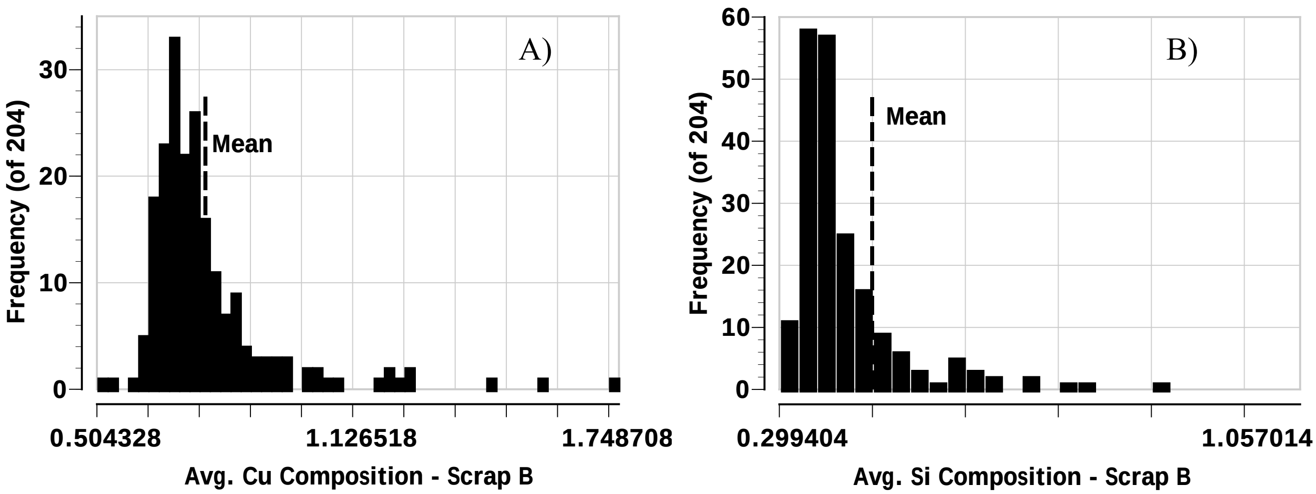

Figure 1. Frequency of average Cu and Si compositions for Scrap B. A) Compositions are clustered around the mean, B) compositions are clustered around the minimum.

To determine the appropriateness of other probability distributions, a chi-squared goodness-of-fit test was performed for the following distributions: normal, logistic, lognormal, maximum extreme, minimum extreme, gamma, beta, Weibull, exponential, triangular, uniform, and Student's t. For each of the five scraps, nineteen alloying elements were tested.

Table III summarizes the top eight results for two major alloying elements (iron and magnesium) and one trace element (boron) for Scrap A and Scrap B. A distinction is made between elements that are present in weight fractions greater than or equal to 0.10 (major alloying elements) and those that are lower (trace); major alloying elements include Si, Fe, Cu, Mn, Mg, and Zn while trace elements include Cr, Ni, Ti, Pb, Ca, Sn, Na, B, Ga, V, Bi, Cd, Zr. The lowest chi-squared values indicate the best fitting probability distribution. For some of the cases, boron in Scrap B for example, only one distribution provides a good fit (Student's t) while for others, such as magnesium in Scrap A, several different choices could be used to characterize the compositional distribution.

Table III. Chi-squared values for curve fits to various distributions for iron, magnesium, and boron.

	Fe		Mg		B	
Scrap A	Max Extreme	28.5	Gamma	5.3	Max Extreme	29.2
	Student's t	53.8	Max Extreme	9.7	Gamma	36.3
	Gamma	57.5	Lognormal	12.6	Weibull	56.4
	Logistic	76.6	Weibull	15.4	Lognormal	62.3
	Lognormal	76.1	Logistic	37.8	Logistic	67.3
	Weibull	112.5	Beta	42.5	Student's t	69.1
	Normal	138.9	Normal	42.5	Beta	73.5
	Min Extreme	380.0	Student's t	43.2	Normal	109.5
Scrap B	Lognormal	12.0	Max Extreme	9.0	Student's t	16.3
	Max Extreme	12.8	Weibull	10.2	Lognormal	51.7
	Gamma	14.6	Gamma	11.5	Max Extreme	57.1
	Beta	16.0	Beta	14.6	Gamma	105.7
	Weibull	25.7	Lognormal	27.3	Logistic	118.4
	Logistic	27.2	Logistic	31.6	Pareto	367.3
	Student's t	32.1	Exponential	34.7	Weibull	479.8
	Normal	39.7	Normal	35.9	Normal	650.1

Figure 2 summarizes the results for all five scraps and all nineteen elements tested. One can see that no single distribution provides the best match for a majority of the test cases; however, four top distributions make up 79 of the 95 (83%) total possible best fits. These are lognormal, Student's t, maximum extreme, and the gamma distribution. Some of the characteristics of these tests are described to provide insight as to how they might relate to the compositional distributions of scraps. Figure 3 shows the general shape of the distribution curves.

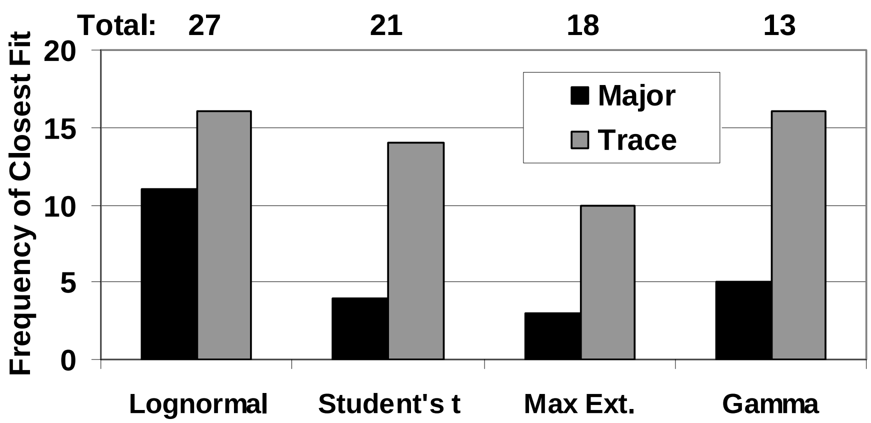

Figure 2. Ranking of the top four probability distributions for number of times it had the lowest chi-squared value of all probability distributions tested (ie. was the closest fit).

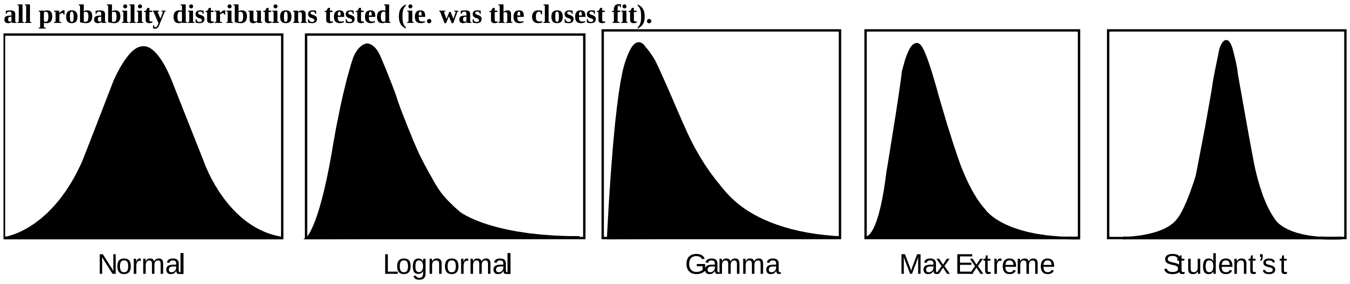

Figure 3. Abstract curve shape for the top four distributions compared to a normal distribution curve.

The lognormal distribution is widely used in cases where most of the distribution clusters around the minimum value, for example, in real estate valuation. Two characteristics of the lognormal distribution make it a good candidate for compositional data: 1) values are positively skewed and 2) values can increase without bound but cannot be negative. The gamma distribution is similar to the lognormal distribution; it describes independent occurrences that do not correlate to each other. The parameters for the gamma distribution are location, scale, and shape. The maximum extreme distribution, also known as the Gumbel distribution, is typically used to describe the largest value over a period of time. Its parameters are likeliest and scale and is a positively skewed distribution with a long tail extending towards the highest values. The student's t distribution is a normal distribution with higher kurtosis and thicker tail region (more outliers). It has more parameters than a normal distribution (midpoint, scale, and degree of freedom) which gives it additional flexibility. It is not entirely surprising that three of these top four distributions (excluding student's t) are positively skewed. The composition of recycled materials would have the most near the mean but accumulation of unwanted elements would result in a longer tail toward the high end. Because compositions of most of the trace elements are very near to zero but cannot be negative, a positively skewed curve would make sense.

Sensitivity Analysis: Effect of Assumptions on Error Rate

The chance constrained optimization model was used to determine an optimal batch plan of primary aluminum, alloying elements, and scrap materials to produce the six alloys outlined in the case set-up. The optimal batch plan is made assuming a normal distribution of scrap compositions. To test the effect a different distribution might have on the error rate for being within specification, Monte Carlo simulations were run on the optimal batch plan with varying distributions on the scrap compositions. The number of times the finished alloy had a composition outside the target specifications was recorded as the error rate, shown in Figure 3. One can see that most of the distributions result in a large increase in the error rate with the largest being for the maximum extreme and the gamma distributions. Interestingly, the lognormal distribution only resulted in a slight increase in error rate. By examining the simulated final compositions of one of the alloys (silicon composition in alloy 3004 in this case), one can see the mechanism behind this increase in error rate. Figure 5 shows that both the maximum extreme and gamma distributions have much larger tails into the higher composition regions than the normal distribution. This is of note because the consequences of compositional errors are inherently asymmetrical. Corrective measures rely upon either additives or dilutants with sometimes significantly different costs. Therefore, errors on the high side are especially important to avoid.

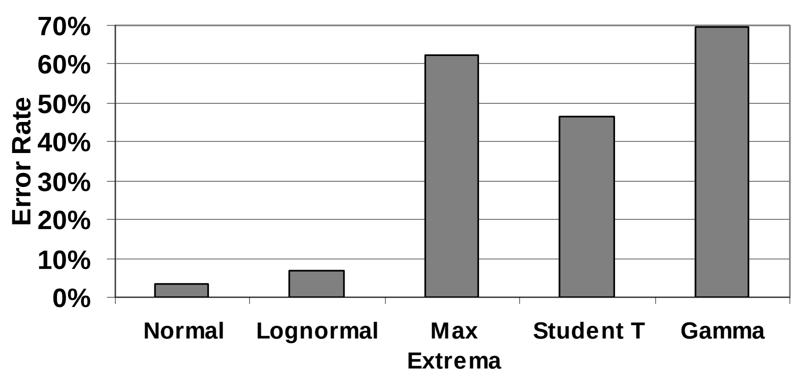

Figure 4. Error rate results for Monte Carlo simulations testing against optimal solution assuming normal distribution

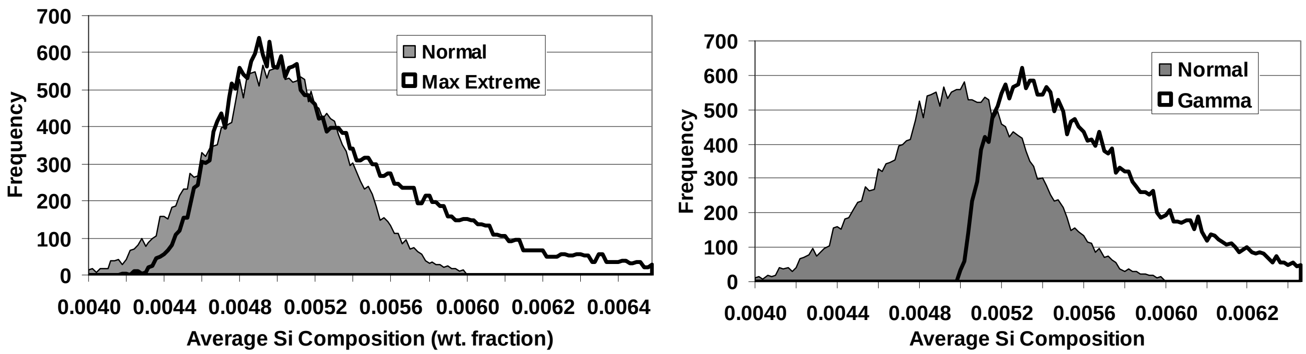

Figure 5. Monte Carlo simulation showing effect of different distributions on the final composition of finished alloys.

Conclusions

From the aluminum cases presented here, it is clear that a normal distribution may not be the best approximation for compositional data. Chi-squared goodness of fit tests showed that more likely a positively skewed distribution would be best, specifically lognormal, maximum extreme, gamma, or student's t distributions. Using Monte Carlo simulations, one can see that the consequence for making an incorrect assumption about the compositional distribution can have a profound effect on production decisions. Specifically, distributions with long tails cause significant increases in error rate because final compositions will be pushed closer to the maximum specifications than predicted by the optimal batch plan. Future work should focus on how to accurately accommodate these trends in production strategies that both increase the opportunity for scrap use and, therefore, lower production cost.

References

1. Lund, J.R., et al., *Linear Programming for Analysis of Material Recovery Facilities.* Journal of Environmental Engineering, 1994. **120**(5): p. 1082-1094.
2. Stuart, J.A. and Q. Lu, *A Model for Discrete Processing Decisions for Bulk Recycling of Electronics Equipment.* IEEE Transactions on Electronics Packaging Manufacturing, 2000. **23**(4): p. 314-320.
3. Cosquer, A. and R. Kirchain, *Optimizing The Reuse of Light Metals from End-of-Life Vehicles: Lessons from Closed Loop Simulations.* TMS: The Minerals, Metals, and Materials Society, 2003. **AA3**.
4. Shih, J.S. and H.C. Frey, *Coal blending optimization under uncertainty.* European Journal of Operational Research, 1995. **83**: p. 452-465.
5. Gesing, A., et al. *Separation of Wrought Fraction of Aluminum Recovered from Automobile Shredder Scrap*. in *The Minerals, Metals & Materials Society 2001*. 2001.
6. Gaustad, G., P. Li, and R. Kirchain. *Modeling Methods for Managing Raw Material Compositional Uncertainty in Alloy Production*. in *TMS: The Minerals, Metals, and Materials Society*. 2006. San Antonio, Texas.
7. Peterson, R.D. *Scrap Variability and its Effects on Producing Alloys to Specification*. in *TMS: The Metals, Minerals, and Materials Society*. 1999.
8. Charnes, A. and W.W. Cooper, *Chance constrained programming.* Management Science, 1959. **6**: p. 73-89.
9. Kelly, T., et al., *Historical Statistics for Mineral and Material Commodities in the United States*. 2004, USGS.
10. *Recycling Xchange*. 2006.
11. Gesing, A., *Recycling light metals from end-of-life vehicles.* Journal of Materials, 2001. **53**(11).
12. *EN 13920: Aluminium and Aluminium Alloys-Scrap*, B.B. Standards, Editor. 2003, European Committee for Standardization.