Manufacturing Processing Improvements Using Business Intelligence

Farid Bourennani[1], Jamal Alsadi[1], Ghaus M. Rizvi[1], Daniel Ross[2]
[1]Faculty of Engineering and Applied Science
University of Ontario Institute of Technology
Oshawa, Ontario
Canada
[2]Sabic Innovative Plastic Inc
Cobourg, Ontario
Canada
farid.bourennani, jamal.alsadi, ghaus.rizvi {@uoit.ca}, daniel.ross2@sabic-ip.com

ABSTRACT: *Plastics production is an important industrial sector in the world. However, producing the right plastic color with minimal reject is a challenge for plastic manufacturer. In this paper, Business intelligence is used for production data analysis in order to isolate the parameters the most susceptible of causing color mismatch. Consequently, the number of parameters to be further examined is significantly reduced. And, the processing parameters causing rejects can be quickly identified and solved.*

Keywords: Business intelligence, Intelligent manufacturing, Businsess data mining

Received: 26 April 2011, Revised 31 May 2011, Accepted 6 June 2011

© 2011 DLINE. All rights reserved

1. Introduction

Plastics production is an important industrial sector. It has grown substantially over the past few decades because of several reasons. Plastics are easier to manufacture in comparison with other materials. In addition, complicated parts can be modeled in fewer steps at lower production costs. Also, plastics are recyclable and their processing requires less energy than other conventional materials such as metals. Furthermore, one major benefit of using plastics is the vast array of colors that can be used for making products more attractive to consumers. However, producing the right color with minimal reject is a challenge for plastic producers (compounders). Especially, it is true for short lead time suppliers of small lot sizes which predominantly work on prototype development. Thus, quick realization of prototypes is a critical quality to be competitive in the global market.

Almost all the plastics used in manufactured products are colored. For the compounding industry, the main cause of material rejects is due to color mismatch. Many parameters might cause color mismatch problems such as formulations errors, poor pigment dispersion, pigment or resin degradation, interactions between materials, and processing conditions such as high operating temperatures or high processing sheer rates. The color mismatch problems have not been thoroughly studied by the plastic compounding industry as by the paint industry. In order to reduce rejects due pass color opportunities, the parameter(s) causing them needs to be identified.

SABIC, a recognized global industry leader, is among plastic manufacturers facing such problems. One of its manufacturing plant in Cobourg, Ontario, has in place very large databases that record the different formulations and polymer processing parameters. However, the amount of data is so large that it requires automated tools for data analysis.

In this paper, the use of statistical-based methods is proposed for determination of most susceptible parameters to cause color mismatch problems in the compounded polymers. It is shown that complex data manipulations, such as complex data queries often used in On-Line Analytical Processing (OLAP) can quickly isolate such group of parameters to be experimentally studied.

The rest of the paper is organized as follows: Section 2 and 3 discuss respectively the related work, and the background. Section 4 details the data to be processed. Section 5 presents the OLAP method and the results obtained from it. Section 6 describes the used data mining approach and the outcome of this method. Finally, Section 7 concludes the present paper.

2. Related Work

Business Intelligence (BI) is used for efficient business decision making. It encompasses several methods such as data warehousing, DM, OLAP, corporate performance management, analytics, benchmarking, text mining, and predictive analytics. In this paper, BI is used for detecting possible causes of polymer color mismatch. Especially, the DM answers this knowledge discovery concern as it is used to extract valuable knowledge from databases [1].

The combination of OLAP and DM is very strategical. It permits extracting valuable *knowledge,* discovered through DM, from the *necessary* data obtained through OLAP. The combination of DM and OLAP intests many industrial sctors, such as business management [2], clinical [3], agriculture [4], and others. In this paper, BI is used for detecting parameters susceptible to cause color mismatch. For example, pigment dispersion is one of the most likely causes of problems in color matching.

To the best of our knowledge, our literature review shows there is no other paper related to plastic color match problems using DM or OLAP. However, more general literature reviews about DM and OLAP applications in the manufacturing sector are available in ref [1] and [5]. And, Chaudhuri et al. [6] propose a litterature review about OLAP-based aplications. Shumueli et al. [7] present different DM algorithms applied to business problems in a user friendly envirnment using MS. Excel® and Xlminer®.

Usually, the plastic color match problems are solved ad-hoc by either experienced operators who have a feel for which pigment to add (due to experience) or based upon spectral analysis which indicates frequencies are lacking in the reflectance spectra and the pigments are adjusted accordingly in the new formulation [8]. Some other researchers tried to predict the output colors based on historical data by Neural-Networks [9].

3. Background

SABIC produces colored plastic on a small and medium scale for plastic processing companies that use it mostly for prototyping. Therefore, SABIC receives orders that need to be produced rapidly within a couple of days. They have several production lines for this purpose. However, before sending the ingredients to the appropriate line for mixing, they need to weigh each one of them. There are several hundreds of ingredients, which can be broadly categorized as: pigments, resins, and additives. The ingredients and additives combination form a plastic material of a certain grade. The pigments provide the color of the plastic.

There can be numerous causes, which can lead to color mismatch, some of which could be:

a) Slight variations in the amounts of pigments or dyes; some colors may have higher sensitivity to these variations.

b) Dispersion characteristics (mixing process, variations in pigment particle sizes due to processing, or resin and additives properties)

c) Interactions between pigments, resins and additives

d) Variations in the properties of the incoming raw material.

4. The data

The selected data covers the three operational years (2009-2011). Among this data, we selected two main tables, and extracted only the columns we need (Views). The first one (Table 1) is chosen because it contains the order-related information such as the order number, the grade of the plastic, the color of the plastic, and the line on which the order was processed.

order-id	grade	Color	line
2345	G1234	C1234	25
…	…	…	…

Table 1. Order related table

The second table (Table 2) contains essentially the batch-related information such as the order number, the batch number, the ingredient code, and the quantity of the ingredients used for preparing the batch. It is important to mention that an order may have several *adjusted* batches. In other words, if the order result is not satisfactory, another batch for the same order is processed again and again until the required result is achieved.

order-id	batch-id	ingredient	Quantity
2345	0	X145	630
2345	0	Z721	425
…	…	…	…

Table 2. Batch-related table

The extracted data will be processed using OLAP and DM methods.

5. On-Line Analytical Processing

(OLAP) The OLAP cubes can be presented as an extension to the two-dimensional spreadsheet array. For example, a user might wish to analyze data by more than two parameters such as by product, by time period, and by city. These additional parameters for analyzing the data are known as dimensions. Usually, the number of dimensions is three; that is why the method is called OLAP *cubes.*

The OLAP operations include rollup (increasing the level of aggregation) and drill-down (decreasing the level of aggregation or increasing detail) along one or more dimension hierarchies, slice_and_dice (selection and projection), and pivot (re-orienting the multidimensional view of data) [10]. All of these operations are used in this paper to answer the queries leading to the detection of the causes of rejects in plastic coloration by using *MS Access* [11].

5.1 Slice and Dice

Slice refers to the *selection* of a certain number of tuples respecting certain criteria. For example, in Table 1 we have selected only tuples having an order for the years 2009 and up.

The dice refers to the *projection* of a specific number of columns from a table. For example, Table 1 is a projection of four columns of interest. For example, the columns related to the dates are not part of the table as shown in Table 2 after dice operations.

In brief, the slice is a vertical way of selecting the data, while the dice is a horizontal way of selecting the data.

5.2 Drill Down/Up

Drilling down or up is an analytical technique where the user can navigate at different data levels ranging from a summarized (up) level to a more detailed (down) level. For example, Table 1 is too general to detect color mismatch causes because it does not contain ingredient-related information including the pigments. On the other hand, Table 2 contains this information. Consequently, merging these two tables – drilling down – will provide us the required level of detail to analyze further the ingredients of the orders that may cause the plastic color mismatch. The Table 3 is an illustration of drill-down operation's resulting table.

In order to continue data analysis, the data is divided into two groups: the successful orders versus the adjusted orders. The successful orders are the one that resulted in the correct plastic color from the first processing. The adjusted orders are the ones that required at least two processing *batches* before achieving the right plastic coloration. In other words, for every adjusted batch, the quantities of the ingredients – mainly the pigments – are re-adjusted each time to meet the required plastic coloration.

Order-id	Grade	Color	Line	Batch-ID	Ingre-dient
2345	G1234	C1234	25	0	X145
2345	G1234	C1234	25	0	Z721
…	…	…	…	…	

Table 3. A Consolidated table

5.3 Pivot

Pivot is a relational operation that allows data in rows to be exchanged for columns [12]. The result of a pivot operation is the summarization of a table's data, which becomes easier for the user to visualize. These visualization properties simplify greatly the data analysis and the pattern determination; you can imagine how complicated it is to analyze the data when you have hundreds of ingredients, several lines, and thousands of tuples describing the plastic orders.

In this paper, the *pivot tables* served to answer two questions. How many parameters have been involved in successful orders? And, how many parameters have been involved in adjusted (unsuccessful) orders? Therefore, we pivoted the table to make the parameters as a row, and the count of the parameters as a column. Table 4 is an example of pivot table that was processed from Table 3 from which only pigments were selected; the resins and additives were ignored. Also, from the adjusted orders, only the last successful batch of every order was selected; the intermediate batches which did not produce a successful color were eliminated.

Row Labels	Count of Ingredients
P145	4
P334	45
P347	211
...	...

Table 4. Pivot table - Pigments

6. Data Mining

The DM concept refers to the process of discovering new, valuable knwoledge from a large collection of raw data [13]. Other terms are also used for DM, such as data pattern processing, knowledge extraction, information discovery, etc.

In this work, a statistical manipulation of the data in was used. The aim of the statistics is to determine parameters susceptible to cause color mismatch such as resins, lines, grades and others.

6.1 Data cleaning

There are numerous pre-processing methods used in data mining, the most common ones are data cleaning, data integration, data transformation, and data reduction [14]. Usually when there is schema matching or data integration involved, the data cleaning is a necessary part of the work.

Data with discrepancies have been deleted. In addition, only one batch is kept among batches re-adjusted several times.

6.2 Pigment adjustment ratio

The parameters having high adjustment ratio were extracted using the following rule:

$$r_p = \frac{|S_p|}{|T_p|}$$

where, r_p is the ratio of adjusted orders, $|S_p|$ is the number of adjusted orders having the parameter *p*, while $|T_p|$ is the total number of parameters (successful and adjusted) having the parameter *p*.

Then, the factors or parameters that might cause high adjustments are being reported to Sabic processing group for further investigation. The next section describes the results.

7. Results and Discussion

Now the group of adjusted batches has been determines. Let's try to determine the parameters that may cause plastic color mismatches.

7.1 Pigment Color

The color is designated by a given hue or pigment. Possibly, some colors can be more sensitive to certain pigments or processing conditions. For example, some colors, as shown on Figure 2 have adjustments rates much higher than the average. These colors should be further investigated in the labs by processing engineers.

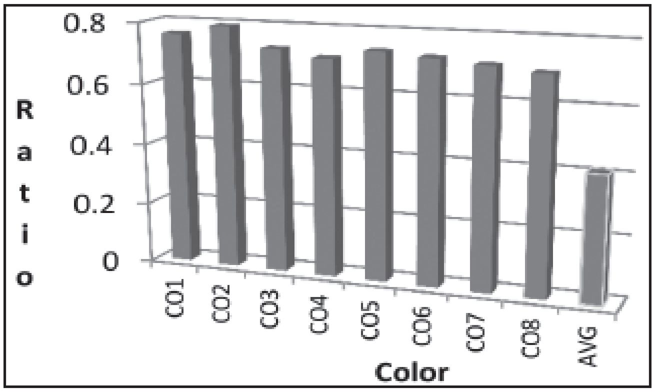

Figure 1. Color Adjustment Rates

7.2 Ingredients

The ingredients such as resins, additives, or other pigments can also cause problems. An ingredient is one of the components of a formula. Figure 2 shows some of the ingredients having the highest adjustment ratios in comparison with the average adjustment ratio. It can be seen that the ingredient N1 appears to be the worst ingredient. However, this does not necessarily mean that the ingredient is bad; but, this ingredient which is an additive should be further studied to see whether a combination with other ingredients is generating higher adjustments. Other possible hypothesis can be related to processing conditions.

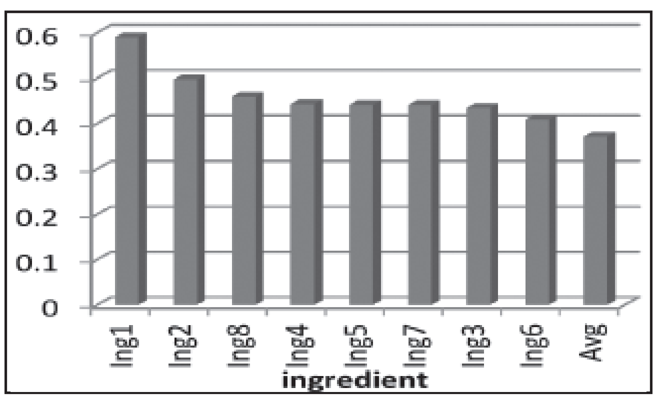

Figure 2. Ingredients Adjustments Rates

7.3 Grades

Grade is the internal designation for materials with a similar composition of resin and additives. As shown in Figure 4, several grades have adjustments rates higher than even the double of the averaged grades' adjustment rates. These grades are prospective sources of problems that should be examined by the mechanical engineering team.

7.4 Lines

SABIC has several production lines. By comparing production lines performances, the lines causing more problems with specific pigments can be further analyzed. On the other hand, the production lines offering good performances can be also further studied to maintain them with specific pigments if the ratio is low. As shown in figure 4, only one line seems involved in higher number of adjusted batches.

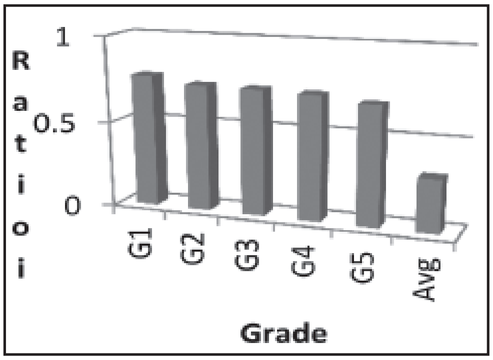

Figure 3. Grades Adjustment Rates

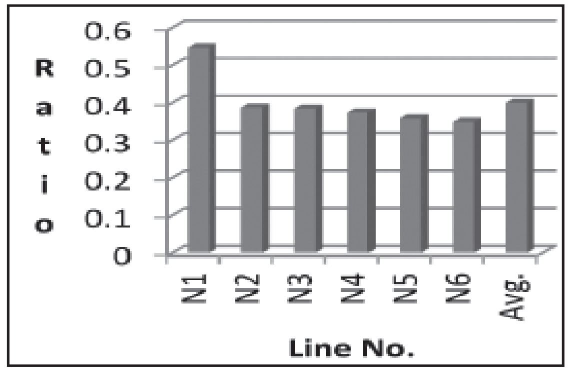

Figure 4. Lines Adjustment Rates

7.5 Products

In SABIC, given materials used in batches are from different resins. A product refers to the type of resin. The distinctive chemical makeup of each product is different. For example, a resin can be of polycarbonate product, PBT product, or others. The two products causing high adjustment rates, as shown in Figure 6 can be further analyzed because they have higher adjustment rates than the average.

7.6 Types

Polymer type refers to the resin type; e.g. translucent, opaque, or transparent. As shown in Figure 7, some polymer types such T1, T2, T3, and T4 have higher adjustment rates than the average. It would be interesting to examine the correlation between the high adjusted pigments and the polymer type. Such relationships have been discovered in a previous data mining project with smaller data [15].

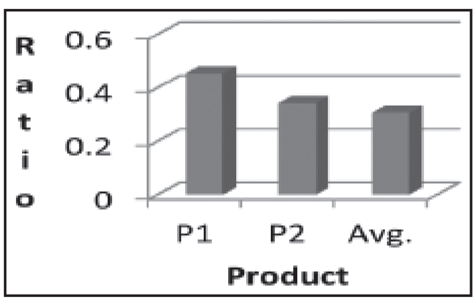

Figure 5. Product Adjustment Rates

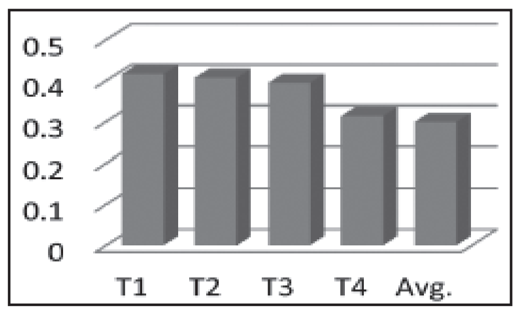

Figure 6. Polymer Type Adjustment Rates

8. Conclusions

In this paper, combinations of OLAP and data mining have been used for the determination of scientific reasons behind plastic color mismatching and adjustments.

The results permitted the extraction of a very limited number of parameters to be analyzed in the labs to determine the cause of plastic mismatch. Especially, six parameters were studied which are color, grades, ingredients, lines, products, and resin types. The parameters involved in high adjustments ratio were extracted. In general, color, grades, ingredients, and resin of plastics seemed to be the main parameters to cause color mismatch. For example, certain grades can double adjustment rates. The relation-ship between these factors should be further studied using DM methods. Moreover, lines and products seemed to not have any impact on adjustments with two or three exceptions.

In the future, we would like to build new algorithms such as fuzzy association rules, statistical models, or association rules for a deeper relationship analysis among parameters. Also, other processing parameters can be incorporated in the study such as temperature, RPM, feed rate and others in order to analyze the problem from a mechanical point of view.

References

[1] Wang, K. (2007). Applying data mining to manufacturing: the nature and implications, *Journal of Intelligent Manufacturing*, 18 (4) 487–495.

[2] Pintér, G. (2008). Integra*tion of OLAP and data mining for analysis of results from dependability evaluation experiments*, International Journal Knowledge Management Studies, 2 (4) 480-498.

[3] Palaniappan, S., Ling, C, S. (2008). Clinical Decision Support Using OLAP With Data Mining, IJCSNS International Journal of Computer Science and Network Security, 8 (9) 290-296.

[4] Wang, C., Wang, L., Dong, P., Qiao, X. (2008). Research of Decision Support System (DSS) for Greenhouse Based on Data Mining, IFIP International Federation for Information Processing: Springer Boston, 259, p. 1071-1076.

[5] Harding, J, A., Shahbaz, M., Srinivas, S., Kusiak, A. (2006). Data Mining in Manufacturing: A Review, *Journal of Manufacturing Science and Engineering,* 128 (4) 969-977.

[6] Chaudhuri, S., Dayal, U. (1997). An Overview of DataWarehousing and OLAP Technology. ACM SIGMOD, 26 (1) 65-74.

[7] Shmueli, G., Patel, NR., Bruce PC (2007). Data mining for business intelligence: Concepts, techniques, and Applications in Microsoft Office Excel with XLMiner, Wiley, Canada.

[8] Karbasi, A., Moradian, S., Asiaban, S. (2008). Computer Color Matching Procedures for Mass Colored Polypropylene, Polymer-Plastics Technology and Engineering, 47 (10) 1024–1031.

[9] Kuo CF, J., Huang, YJ., Su TL., Shih, CY. (2008). Computerized Color Distinguishing System for Color Printed Fabric by Using the Approach of Probabilistic Neural Network, Polymer-Plastics Technology and Engineering, 47 (3) 264–272.

[10] Wyss, C, M., Robertson, E, L. (2005). A formal characterization of PIVOT/UNPIVOT*., In:* Proceedings of the 14th ACM international conference on Information and knowledge management, p. 602 - 608.

[11] Access: office.microsoft.com/access [accessed: 01-06-2011]

[12] Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. Association for Computing Machinery, *Communications of the ACM,* 39 (11) 27-34.

[13] Fong, A, C, M., Hui, S,C., Jha, G. (2002). Data Mining for Decision Support, IEEE IT Professional, 4 (2) 9-17.

[14] Han, J., Kamber, M. (2006). Data Mining, Second Edition: Concepts and Techniques. San Francisco: Morgan Kaufmann, p. 72-97.

[15] Bourennani, F., Rizvi, G, M., Ross, D. (2010) Plastic color mismatch causes identification using OLAP and data mining. ICDIM 2010, Thunday-Bay, Canada, p. 69-74.