A comparison and application of pattern recognition techniques in materials selection using dimensionless numbers

Joseph Werner Schmidt-Castañeda, Francisco Javier Sandoval and Dante Jorge Dorantes-Gonzalez
Instituto Tecnológico y de Estudios Superiores de Monterrey - Campus Estado de México
{schmidt, ddorante, fsandova}@itesm.mx

Abstract **-The objective of this paper is to present a tool and a comparison of techniques that can help in the material selection process based on pattern recognition and using dimensionless parameters from Π theorem. It is known that the main disadvantage in materials selection is that a huge database is needed to make the selection and also to make clusters of materials with some characteristics, so we need an easier and automatic way to cluster materials with respect to their physical and mechanical properties, and particularly for material machinability. This paper also shows a possible application related to material selection using this approach to complement the conceptual design of a part coding system using neural network techniques.**

Index Terms - Materials selection, Π theorem, dimensionless numbers, physical and mechanical properties, neural networks, pattern recognition.

I. INTRODUCTION

Group technology is a technique and a philosophy that improves the efficiency in production by means of grouping a great variety of parts using information of their shape, dimensions and working process routing [1, 2].

The main requirement of group technology is to have a coding system and a part classification that describes part characteristics, as well as their geometrical form, material and working process routing to produce these parts with a code number, gathering this way the parts with similar codes in a specific manufacturing cell or group of machines [3-5].

Parts grouped into families are widely used in manufacturing in order to get profit from their similarities [6]. Grouping parts into families can make easier manufacturing, process planning and production [7].

Π theorem or Buckingham theorem is a tool that helps us to take into account various properties which can be used to give material characteristics such as machinability. As the units are dimensionless, they can assist in the design of a Flexible Manufacturing System (FMS) [8]. It is known that artificial intelligence techniques [9] like pattern recognition can help a lot in problems where there is a missing mathematical model or information about the case of study, so in problems like machinability [10] and materials selection we can use these techniques working with dimensionless numbers to make clusters of materials with the same characteristics of machinability to have an enormous database containing materials and their properties [11]. This enables us to generate material working parameters form only a limited selection of physical and mechanical properties.

This paper was supported by The Research Program on Automated Manufacturing Systems from the Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Estado de Mexico.

2. METHODOLOGY

Pattern recognition was selected as an intelligent tool for classification and recognition of groups or clusters with similar characteristics of machinability. The algorithms studied for this job were K-means, Fisher´s, lambda classification, expectation maximization and Isodata algorithms. All of them use statistical tools to separate clusters of points from different groups in different ways, one of them will be selected for the present application [12-17].

Materials selection process was the issue which initiated this work [18]. There are many ways of selecting a material, but the most used is the Ashby map [8], where material clusters are made considering two properties of the materials. The information defining these clusters is held within a large database (image 1). The aim of this thesis is to process this clustering without the need of a database and to include more properties in one representation.

As we need to have a many properties in one graph and we want to have dimensionless points in the graph we need to use dimensionless numbers theory that helps when developing experimental techniques.

Fluid Mechanics uses dimensional analysis which does not give a complete solution. The success of this analysis depends on the skill to define parameters that would be applied. If one of the variables is omitted, the result will be incomplete and incorrect conclusions will be generated.

The Π Buckingham theorem can be extrapolated to machining and metal-cutting applications, which can be applied to Group Technology (GT) directly [19-21]. Machining relationships are a good example showing how easily Π Buckingham theorem can be applied.

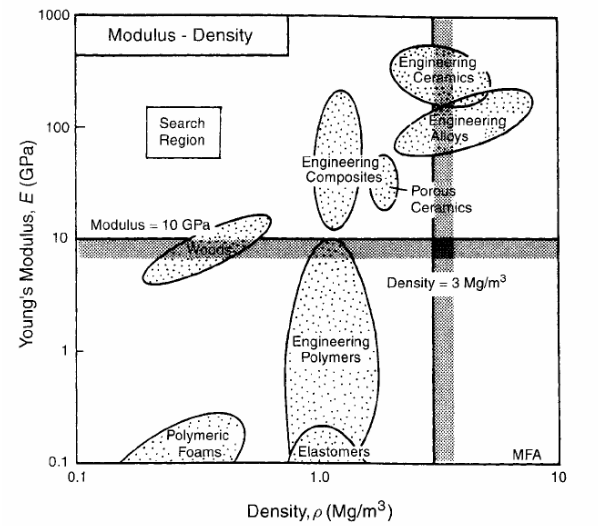


Image 1
Ashby map

The relationship argued by F.W. Taylor defined by equation (1), where cutting rate and tool life equals a constant makes it possible to establish a link with FLT system (force-length-time) Eq.(2):

$$VT^n = C \quad (1)$$
$$C = f(L,T) \quad (2)$$

A large version from the Taylor equation, defined by Eq. (3) exists, where variables such as: feed, cutting depth and the material hardness are involved:

$$C = V\,T^n\,d^x\,f^y \quad (3)$$
$$VT^n f^m d^p Hq = KT^n\,(ref) f^m(ref) d^p(ref) H^q(ref) \quad (4)$$

Where f is the feed rate, d is cutting depth and H is the hardness of the material as you can see in equation (4).

Exponents x, y, m, p and q can be determined experimentally. K is a constant value analogous to C.

We focused on the variables of the Taylor equation for making dimensionless numbers using the following mechanical and physical properties: Yield Strength, Vickers Hardness, Cutting Speed, Cutting depth, Tool life, Linear expansion, Thermal conductivity, and Density.

Now we used the Π theorem by means of making a dimensional analysis of each property and decompose them in their basic units that are length (*m*), weight (*kg*), time (*s*) and temperature (*k*) according to international units system.

2.1 The Buckingham's Π (Pi) theorem

The Buckingham's Pi theorem [22] establishes that with a physically meaningful equation which involves a certain number of physical variables (n) expressible in terms of k independent fundamental physical quantities. We can obtain an equation with a set of $p = n - k$ dimensionless variables. This is made up from the original variables and the original expression. For the purposes of this paper, we want a system of physical properties which share the same description in terms of these dimensionless numbers.
In mathematical terms, we are looking to have a physically meaningful equation such as:

$$f(q_1, q_2, \ldots, q_n) = 0$$

where the q_i are the n physical variables expressed in terms of k independent physical units. This equation can then be expressed by,

$$F(\pi_1, \pi_2, \ldots, \pi_n) = 0$$

where the π_i are dimensionless parameters constructed from the q_i with $p = n - k$ equations like,

$$\pi_i = q_1^{m_1} q_2^{m_2} \Lambda\ q_n^{m_n}$$

where the exponents m_i are constants.

The Buckingham π theorem provides a method for computing sets of dimensionless parameters from the given variables, even when the form of the equation is still known. However, the choice of dimensionless parameters is not unique, for purposes of this work we are looking after non dimensional numbers using a combination of physical and mechanical properties of materials; Buckingham's theorem only shows up a way of generating sets of dimensionless parameters. We are giving a meaning to each number generated related to materials selection in mechanical design and their easiness to their machining processes.

Two systems for which these parameters coincide are called similars; they are equivalent for the purposes of the equation, and we can choose the most convenient one for the purposes of this job.

From this theorem we obtain the following two dimensionless numbers:

$$\vartheta_1 = f\left(V_c, T_h, P_c, H, V, \sigma_Y\right)$$
$$\vartheta_2 = f\left(C_t, \rho, \alpha_t, V_c, P_c, \sigma_Y\right)$$

Where

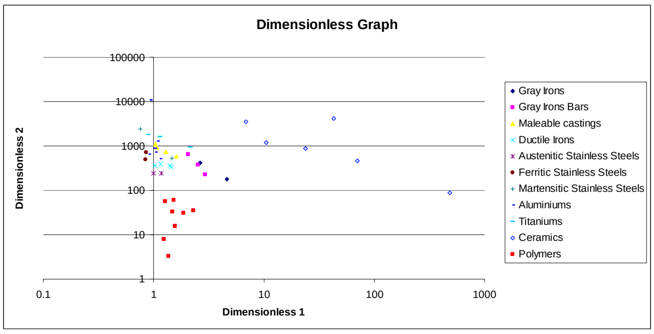
Yield Strength (σ_Y)
Vickers Hardness (HV)
Cutting Speed (V_C)
Cutting Depth (P)
Tool Life (T_h)
Linear expansion (α_L)
Thermal Conductivity (C_t)
Density (ρ)

Through some mathematical procedures we combine these properties [23] so we obtain the following expressions.

$$\vartheta_1 = C_1 \left(\frac{V_c T_h}{P_c} \right) \left(\frac{HV}{\sigma_Y T_h^{\ 2}} \right)$$

$$\vartheta_2 = C_2 \left(\frac{C_t}{\rho \alpha_t V_c^3 P_c} \right) \left(\frac{\sigma_Y}{\rho V_c^{\ 2}} \right)$$

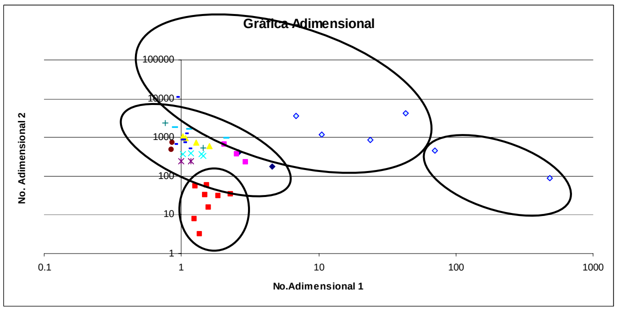
With this number we make a graph with 43 of the most common materials with different properties to view how the materials are spread over a two dimensional graph, as present below (graph 1).



Graph 1
Dimensionless Numbers Graph

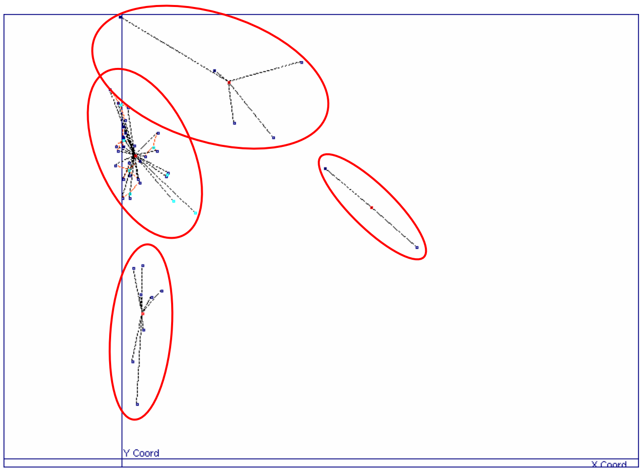
Over all the pattern recognition algorithms studied we compared K-means [13, 24, 25] and Isodata [26] algorithms due to a good separation over our groups of points and they are unsupervised algorithms that make clusters using Euclidean distances between each point.

Graph 2 shows how materials are grouped together according the material type, the group with more density corresponds to metals, the one with red points corresponds to polymers, and ceramics are represented in two groups with different properties indicated by blue points. These four clusters in this graph represent applications for different industries, not their machinability.



Graph 2
General clusters according to their machinability

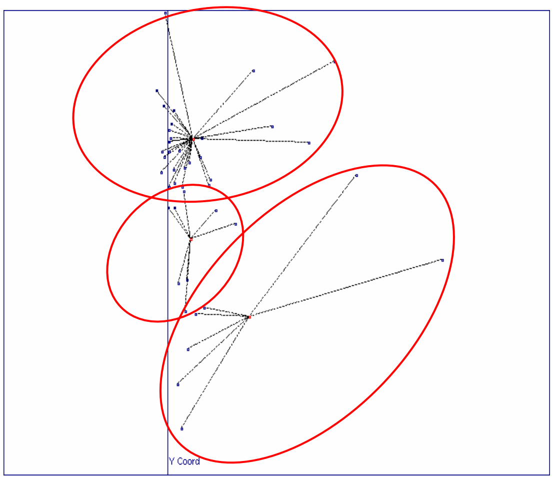
K-means, was the first method we tested, by definition is a nonhierarchical method that initially takes the number of components of the population equal to the final required number of clusters. In this step the final required number of clusters is chosen such that the points are mutually furthest apart. Next, it examines each component in the population and assigns it to one of the clusters depending on the minimum distance. The centroid's position is recalculated every time a component is added to the cluster and this continues until all the components are grouped into the final required number of clusters [13]. We have developed a tool with C# programming language to demonstrate graphically how these two methods are used for this specific application, and also to compare how good the results against their machining properties are.



Graph 3
Dimensionless numbers after k-means algorithm for four clusters

The results match expectations as we have 4 clusters in which there are two clusters of ceramics, one cluster of polymers and one cluster for all metallic materials that have similar characteristics of machinability.

For K-means method, the Graph above shows the C# output of four clusters generated with this dimensionless numbers (graph 3).



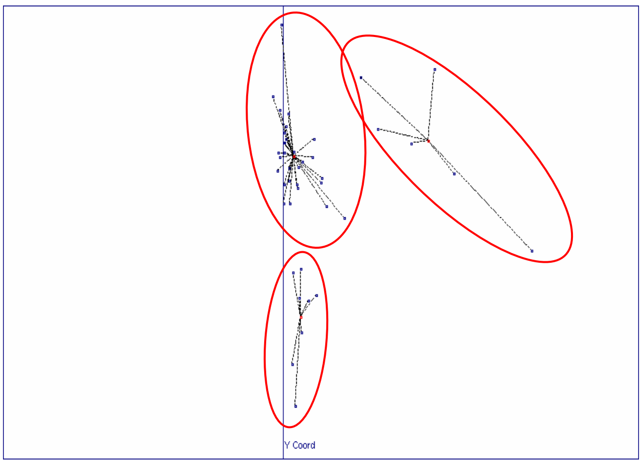
Graph 4
Dimensionless numbers after k-means algorithm for three clusters

In Graph 4, the three groups are not consistent with machinability expectations, so it is not able to model the representation that we are looking for. For accuracy using this method it is important to know how many clusters are required, however this information will not be known. Therefore this method does not provide acceptable results.

As we needed to compare these results with other methods, we tested the Isodata algorithm, which has some further refinements by splitting and merging of clusters. Clusters are merged solving if either the number of members in a cluster is less than a certain threshold or if the centers of two clusters are closer than a certain threshold. Clusters are spitted into two different clusters if the cluster standard deviation exceeds a predefined value and the number of members (points) is twice the threshold for the minimum number of members.

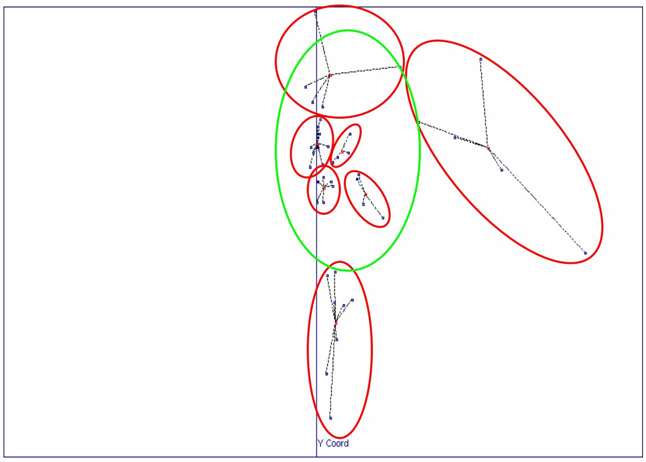
Isodata algorithm is similar to the K-means algorithm with the distinctive difference that Isodata makes it possible to work with different numbers of clusters while the K-means assumes that the number of clusters is known at priori. Also, Isodata algorithm can vary the number of clusters even when we put a desired certain number of clusters before the first iteration of the algorithm, thanks to the statistical tools used in this algorithm during its process [27, 28].

In graph 5, we tested the Isodata algorithm defining five clusters to be made, it consistently ended to three clusters. So we had different clusters of the same materials. In this same graph we can see that one ceramic material is merged to the metallic materials. As we know this material has similar machinability coefficients so we could say that the results of this method are better than the results with the K-means algorithm.



Graph 5
Dimensionless numbers after Isodata algorithm

As we have only 43 materials [29] in our sample, we do not want too many clusters. As we can see, in graph 6 we have obtained 7 clusters, which give a good representation of expected results. However, at this stage we are not looking for this level of detail. Further proving of this technique can be conducted at a later stage.



Graph 6
Dimensionless numbers after Isodata algorithm for twenty clusters

After this process, this program assigns a number for each cluster that has the input data to a Support Vector Machine (SVM) [30, 31] that is trained for entering new material data and assigning a group to the new point added to this system, all of this for purposes of integrate a pair of positions into a parts coding system for giving information about mechanical and physical properties, as well their machinability rate.

In the image 2 it is shown how the SVM gives the output to the new data that we entered to the developed software using the SVM^{multiclass} program that is simulator of multiclass SVM's of Cornell University [32].

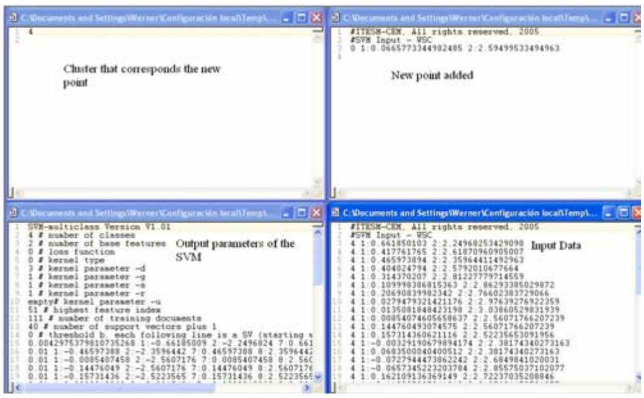


Image 2
SVM Outputs

3. Conclusions

In the present paper we can see how the union of techniques like Π theorem and dimensional analysis, in conjunction with artificial intelligence can be a powerful tool for automation in manufacturing processes. It is demonstrated that cluster separation via K-means algorithm helped spread out the main groups in this job, but the

Isodata algorithm gives more suitable results with less clusters in machinability manner.

The Π theorem can be extrapolated to machining and metal-cutting applications which can be applied to GT directly, where neural networks where implemented for information processing. Training the network with this input data was a straight forward task, so extracting information and testing new materials after using these pattern recognition techniques will also be a manageable task.

4. REFERENCES

[1] Greenwood N.1988.'Implementing Flexible Manufacturing Systems'. Macmillan Education. E.U.A.

[2] Hernández, Rafael. "Desarrollo de un algoritmo de asignación de operaciones tradicionales de corte con desprendimiento de viruta basado en el sistema de clasificación y codificación de piezas KK-3". Tesis. ITESM-CEM. 2001.

[3] Ferré R. 1988. 'La fábrica flexible'. Marcombo.

[4] J. J. Dhingra., K. L. Musser., & G. L Blankeship, Real time operations sheduling for flexible manufacturing systems, Winter simulation Conference, 1992.

[5] Drake R.G., Smith J.J., Peters B.A, 1995. 'Simulation as a planning and scheduling tool for flexible manufacturing systems. 'Winter simulation Conference 1995.

[6] Chen, S.J. and Cheng C.S. A neural network-based cell formation algorithm in cellular manufacturing. International Journal of Production Research, vol 33, No. 2, 1995, pp 293-318.

[7] Srinivas G., E. D. Fasse E.D., Marefat M.M. '1998 Retrieval of similarly shaped parts from CAD database'. IEEE International Conference on1998, Vol. 3, pp 2809-2814.

[8] ASHBY, MICHAEL F. "Materials Selection in Mechanical Design", Second Edition, Butterworth Heinemann, 1999.

[9] Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2nd Edition, Prentice Hall 2003.

[10] Bralla, James, Manual de diseño de producto para la manufactura. Guìa pràctica para producciòn a bajo costo, México, D. F., McGraw-Hill, 1993, 3 tomos.

[11] Nikolopoulos Chris, Shareef Iqbal, and Kalmes Donald; Neural Network Based Machinability Evaluation, R. Loganantharaj et al. (Eds.): IEA/AIE 2000, LNAI 1821, pp. 723-730, 2000.

[12] Tzeng Su C. 1994. 'A fuzzy approach for part family formation'. National Yunlin Institute of Technology, Taiwan, pp. 289-292.

[13] Duda, Richard and Hart, Peter, Pattern Classification, second edition, U. S. A., Wiley Interscience, 2001, 654 págs.

[14] Logendran, R. and Thomas, M.W. A comparison of methodologies for efficient part-machine cluster formation, Computers and Industrial Engineering, Vol. 21(1-4), May 1991, pp 285-2289.

[15] Rosen, Robert. Similitude, Similarity, and Scaling. SPB Academic Publishing. 1989, 207-216,.

[16] BISWAS, G., WEINBERG, J., AND LI, C. 1995. "A Conceptual Clustering Method for Knowledge Discovery in Databases". Editions Technip.

[17] Papamichail Georgios P. The K-Means Range Algorithm For Personalized Data Clustering In E-Commerce. 14th Mini Euro Conference, Human Centered Processes, Distributed Decision Making and Man-Machine Cooperation, 5-7 May 2003, Luxembourg.

[18] Ramos J. 1997. 'Algunas consideraciones sobre la competitividad internacional y su cuantificación'.ESEconomía. Instituto Politécnico Nacional. México, Vol.IV, No.17.

[19] J. L.Garibay J.L. 1999. 'Creación de una metodología para el diseño, evaluación y selección de sistemas flexibles de manufactura metalmecánica'. Tesis de Maestría ITESM.

[20] Groover, Mikell. "Automation, Production Systems and Computer Integrated Manufacturing." 2ª edición. Prentice Hall. E. U. A. 2001.

[21] Askin, R. and Standridge, C., Modeling and Analysis of Manufacturing Systems., New York, John Wiley & Sons, 1993, 480 pages.

[22] Hanche-Olsen, Herald, "Buckingham's pi-theorem", 09/2001.

[23] Kreyzig, Erwin, "Matemáticas Avanzadas para Ingeniería" Vol. I, Tercera Edición, Editorial Limnusa, 2003.

[24] Zhexue, Huang. Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values. Data Mining and Knowledge Discovery 2, 1998, pp 283–304

[25] K. Alsabti, S. Ranka, and V. Singh. An Efficient K-Means Clustering Algorithm. http://www.cise.ufl.edu/_ ranka/, 1997.

[26] Tou, Julius T. and Gonzalez, Rafael C, Pattern Recognition Principles., Pennsylvania, Addison-Wesley Pub. Co, 1974, 377 pages.

[27] Anderberg, M.R., Cluster Analysis for Applications, New York: Academic Press, 1973.

[28] BRAILOVSKY, V. L. 1991. A probabilistic approach to clustering. *Pattern Recogn. Lett. 12*, 4 (Apr. 1991), 193–198.

[29] Shakelford, James, Materials science and engineering handbook, third edition, New York, U. S. A., CRC Press, 2001, 1949 págs.

[30] Angulo Bahón, Cecilio, "Aprendizaje con Máquinas Núcleo en entornos de Multiclasificación", Tesis Doctoral, Universidad Politécnica de Cataluña, Especialidad en Matemáticas, Abril 2001.

[31] González Mendoza, Miguel, "Aprendizaje Estadístico, Redes Neuronales y Maquinas de Soporte Vectorial: Un enfoque global", ITESM CEM: Intelligent Transportation Systems Research Group, 02/2005.

[32] Joachims Thorsten; SVMmulticlassSoftware version: 1.01; Department of Computer Science, Cornell Univ.; Sept. 2004. Site link: http://www.cs.cornell.edu/People/tj/svm_light/svm_multiclass.html