Using Multi-Dimensional Scaling and Hierarchical Clustering to Improve the Process in a Hot-Dip Galvanizing Line

ANDRÉS SANZ GARCÍA, FRANCISCO JAVIER MARTÍNEZ DE PISÓN ASCACIBAR,
RUBÉN LOSTADO LORZA, ROBERTO FERNÁNDEZ MARTÍNEZ,
JULIO FERNÁNDEZ CENICEROS
Project Management Area. EDMANS research group (http://api.unirioja.es/edmans/)
Department of Mechanical Engineering. University of La Rioja
Luis de Ulloa 20. 26004. Logroño. La Rioja
SPAIN
andres.sanz@unirioja.es

Abstract: An application for the scheduling coils of a Hot-Dip Galvanizing Line is presented. In this case, a multi-dimensional scaling is used for the creation of a bidimensional coil map from the most significant parameters of the coils: chemical composition, dimensions, process's parameters, etc. This map shows the group degree and the existing distances between the coils so that the human expert can interact with the software to decide which are the more suitable groups of classification and to detect those coils that could be problematic within the industrial process. After that, a hierarchical clustering is utilized to find local clusters with similar dissimilarities to introduce them into the scheduling list. These combined methods can help in the generation of more effective sequences as well as in the prevention and reduction of the number of contingencies that habitually take place in the plant.

Key–Words: Scheduling, Hot-Dip Galvanizing Line, Multidimensional Scaling, Hierarchical Clustering

1 Introduction

The constant effort to increase product quality and reduce the expenses caused by failures in the manufacturing process is ongoing with Hot-Dip Galvanizing Lines, which are always trying to optimize their operative costs.

1.1 Hot Dip Galvanizing Line

Broadly speaking, the process within a Hot-Dip Galvanizing Lines (HDGL) (Figure 1) can be described in the following manner: The first stage in the line consists of the formation of a continuous strip using steel coils that come from a mill process. Next, the strip passes through a preliminary cleaning section at the entrance of the annealing furnace, where it receives a heat treatment; the stage prior to its immersion in the liquid zinc bath. This treatment is essential in order to improve the properties of the strip and its coating. After the bath, the strip exits vertically, passing through blade-like currents of air which regulate the thickness of the cover.

Finally, the steel is run through a molten-zinc coating bath, followed by an air stream wipe that controls the thickness of the zinc finish (see Fig. 1).

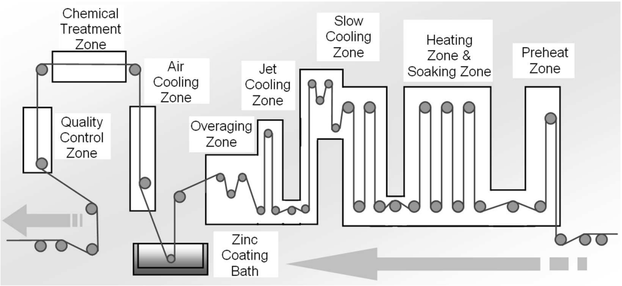

Fig. 1: Basic scheme of a HDGL

1.2 The Quality of the Galvanized Product

The quality of the galvanized product, according to several authors [1, 14], can be divided into two fundamental aspects: the anti-corrosive characteristics and the properties of the steel.

By examining previously exposed coils, one deduces that the control of the heat treatment in a HDGL is fundamental for the process of coating and for the improvement of the properties of the steel. One of the key aspects concerning temperature control consists in assuring that the suitable thermal cycle is reached for each coil while considering the limitations of the HDGL furnace inertia. This is fundamental for the process of coating and for the improvement of the properties of the steel and can only be accomplished when the flow of the steel mass for unit of time is relatively constant and the changes therein are carried out gradually. The need to limit changes in the mass flow rate from one coil to the next is a fundamental coil sequence constraint, but we must also consider other inconveniences: coils within each group have to be sequenced based largely on decreasing sort by width, coils have to be selected with similar steel properties and thermal requirements, etc.

Also, it must be borne in mind that defective coils can appear, or coils with characteristics that differ greatly from the others, and therefore, must be separated from the rest. For these reasons, it is crucial to order the coils that form the strip in such a way that the varying thickness, widths and types of steel do not differ too greatly [16, 22].

We see, therefore, that one of the most important tasks in the process in an HDGL is centered on obtaining sequences of coils that do not contain abrupt changes in the dimensions and types of steel in the consecutively treated coils.

1.3 Present Scheduling Techniques in the Steel Industry

There are few publications that deal with the problem of scheduling in an HDGL. The majority of the publications that address applications and techniques for scheduling in the steel industry are based on continuous casting machines. The systems presently employed in these processes are generally based on either heuristic scheduling strategies and/or are implemented using knowledge software technology [16, 18].

Tang et al. [22] offer a review of the more commonly used methods and scheduling systems in the steel industry. In this work, different methods of optimization are described according to many planning systems (actually integrated in several industrial plants in the steel industry).

As has been mentioned above, one of the primary aims of the HDGL is to determine the best sequence of coils, after which it conforms to the changes in the new dimensions and types of steel. This is essential in assuring that the quality of the coating of each coil and the production are optimal.

The present systems of planning in the HDGL divide the process into two general phases:

- The coils are grouped according to the type of steel (depending on its chemical composition) and its specifications (specifications of production, date of delivery of the product, specifications of quality, etc.).
- Each group of coils is ordered such that the changes in width and thickness are minimalized. Normally, the process of coil scheduling begins by selecting those coils of greater width, and gradually moving towards those that are more narrow. Also, the HDGL looks for variances which are minimal, so as to create a smooth transition.

For each kind of problem, multiple mathematical approaches and heuristic algorithms of optimization are selected [13, 20]. Other authors [5] focus the resolution towards problems of the type Traveling Salesman Problem (TSP) using the dimensions of each coil and other parameters. This methodology is based on the calculation of a triangular matrix of distances between all the coils to determine the best possible sequence. The most critical point of this consists in finding a trustworthy formulation that indicates the degree of existing similarity between two coils according to their physical and chemical properties, as well as restrictions in the process.

Other works, however, are based on the use of graphical methods for scheduling. For example, Tamura et al. [19] looks for optimal ways in distribution maps where each coil is represented according to two coordinates: thickness and width.

At the moment, these techniques are complemented with new techniques [2]: Prize Colleting VRP models, IA, multi-agents, methods of man-machine iteration, or combinations of all of the above. In the following sections, another experience is described, results obtained are discussed and final conclusions are drawn.

2 Proposed Technique

The present work makes use of the Sammon nonlinear projector [17] to visualize the equivalent Euclidean distances of the coils to be treated within a $\Re^2$ space. This algorithm employs non-linear transformations in order to map the original space, from the significant n weighted parameters so that they can influence the classification process, onto a low-dimensional visual space ($\Re^2$ or $\Re^3$), attempting to preserve the Euclidean distances between coordinates of patterns (logically, other linear or non-linear multidimensional scaling methods (MDS) can be used [4]. The first step consists in developing a data base with the n weighted parameters of each coil (considered important within the classification process). These parameters can be: the chemical composition of the steel, width and thickness of the coil, objective temperature, design quality parameters, etc.

Once defined, the coils like points within a space $\Re^n$ are projected by means of Sammon or another linear or non-linear projector in a bidimensional

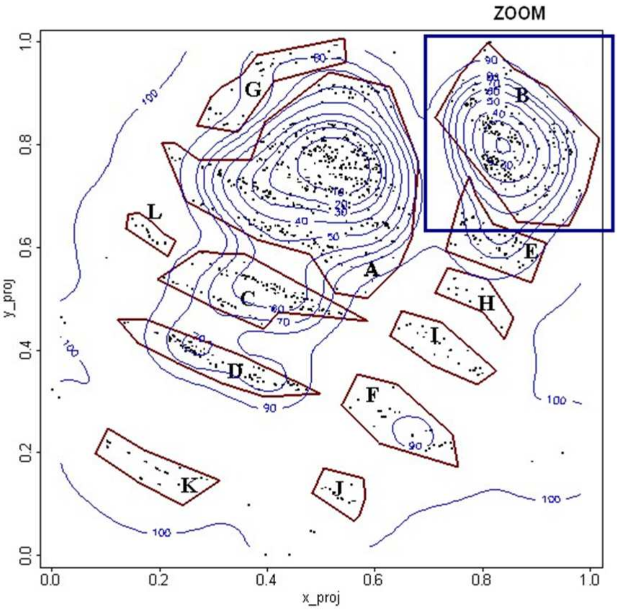

Fig. 2: Defined groups visually using a map

space. The bidimensional map generated represents the equivalent Euclidean distances between all the coils to organize. In this way, the degree of similarity between the coils and the obtaining of the standardized distance of its projections can be determined visually.

This map of points allows one to visualize in an objective manner the equivalent distances between all the coils, and to classify them according to all the considered parameters properly weighted. Logically, the coils with very close parameters will appear tightly-packed, whereas those coils whose values vary greatly amongst themselves will appear more loosely-grouped.

Once all the coils in a two-dimensional graph have been projected, the system allows for iteration with a human expert for the selection of main groups of coils with similar characteristics.

From the map of coils created, the groups of coils with similar characteristics are defined visually with a contour of lines that can be adapted by the user (see Fig. 3) and those different coils that do not correspond to any defined group are detected (Fig. 2).

When the main groups of coils are all selected, the coils of each group are locally grouped according to hierarchical clustering (see Fig. 4). This algorithm [9, 3] uses a set of dissimilarities for the n objects being clustered. Initially, each object is assigned to its own cluster and then the algorithm proceeds iteratively, at each stage joining the two most similar clusters, continuing until there is just one single cluster.

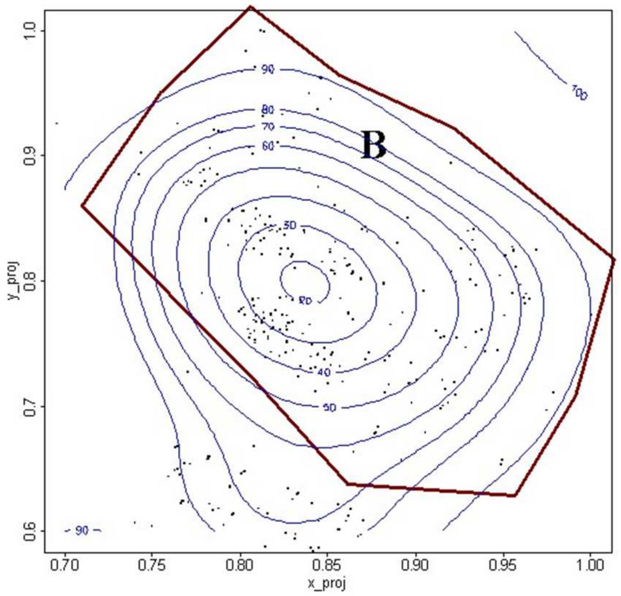

Fig. 3: Contour that defines the group B coils

Figure 4 shows a classification tree (dendrogram) where each node beneath the tree represents a coil from group B and each straight line represents the grouping of two clusters positioned by a Euclidean distance equal to the "height" of that straight line. The main advantage in using this type of graph is that the search for subgroups can be automated such that each individual coil does not exceed a determined distance.

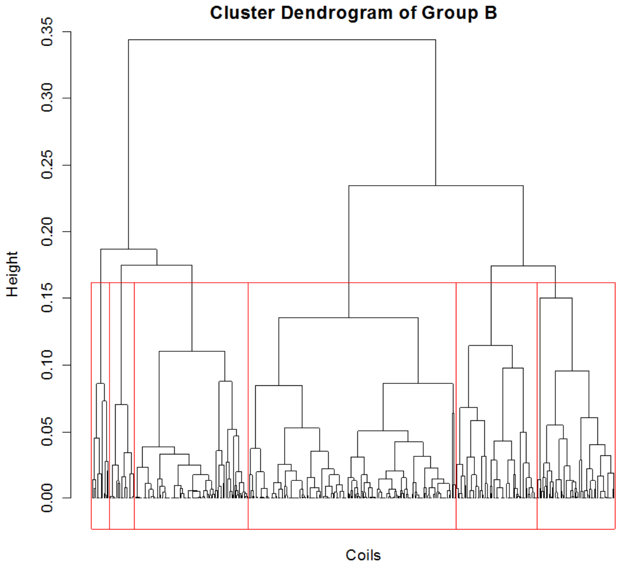

Fig. 4: Dendrogram of Group B formed from hierarchical clustering with six clusters

In Figure 5 are shown 6 subgroups whose Euclidean distance between coils does not exceed 0.15 and are

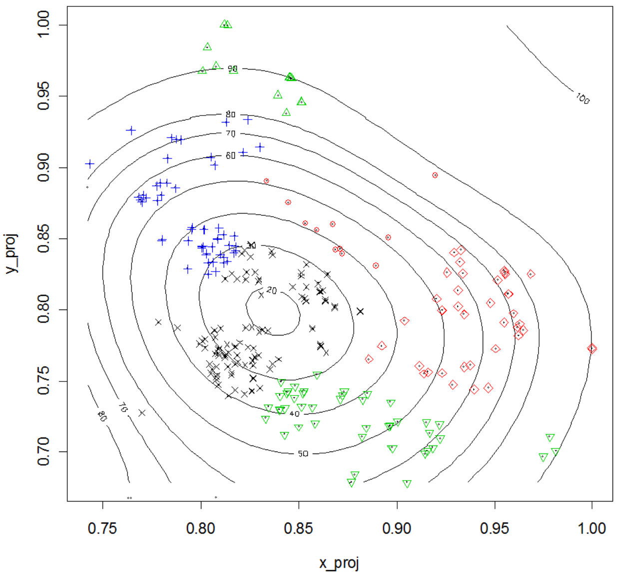

Fig. 5: Different subclusters into group B

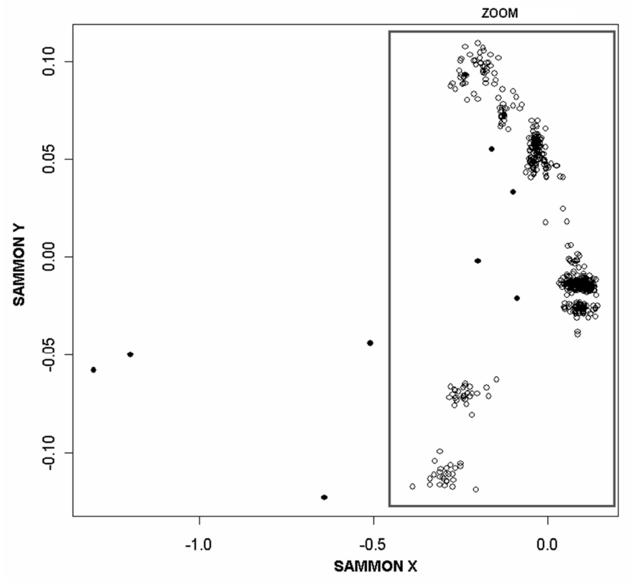

Fig. 6: Sammon coils projection according to chemical composition

distributed within a bidimensional space. The final scheduling is made automatically for each one of the defined subcluters according to one or several of the parameters employed: width, thickness, chemical components of the steel, target temperature, etc.

Once the local scheduling is made, it is necessary to verify the number of times that it surpasses a defined distance (number of gaps), as well as the maximum distances obtained between ordered coils.

If the number of gaps is big or the distances between consecutive coils great, the human expert can take different courses of action until the results of each local scheduling are optimal: i.e. modify the contours that define the groups, divide or fuse groups, choose another distance of clustering, etc.

Finally, once all of the groups are ordered, those coils that could not be classified in any group can be eliminated, ordered separately, or included in some of the already existing groups.

3 Results and Discussion

3.1 Analysis of a previous historical data using non-linear Multi-Dimensional Scaling Methods

The first interesting results arose when analyzing the projection of the 14 chemical components of the coils (C, Mn, Si, S, P, Al, Cu, Ni, Cr, Nb, V, Ti, B, N) for each one of the coils from previous scheduling and upon comparing them with the results obtained after the galvanizing process.

The use of the created map allowed, with a single look and an objective form, the groups of existing coils to be determined as well as to detect those coils whose chemical compositions were completely different. Figure 6 presents the Sammon projection and outliers coils that clearly appeared projected separately from the others. In addition, by extending the right part of the map, two main groups (cluster A and B) and three secondary ones (clusters C, D and E) are defined (see Fig. 7). This map can easily be used to classify new coils in one group or another.

However, the main advantage when using the MDS is observed when we compare the distances between consecutive coils within the scheduling, since it can be used to help in the prediction of potential problems in the process.

Some of the possible causes that can be attributed to problems in an HDGL are produced when one of the welded coils that forms a strip consists of a steel with mechanical characteristics different from the others, has a cross-section that differs substantially from the coils that precede and follow it, or has a defective weld. In Figure **??** is shown a case of potential, problematic events due to the inclusion of a coil with a steel different from the others within scheduling.

3.2 Results of the Experience

As has been mentioned, this methodology takes into account not only the chemical parameters of the steel coils, but also the dimensional parameters (thickness and width), target temperature, and other parameters of the process.

In the tests, the following parameters of each coil

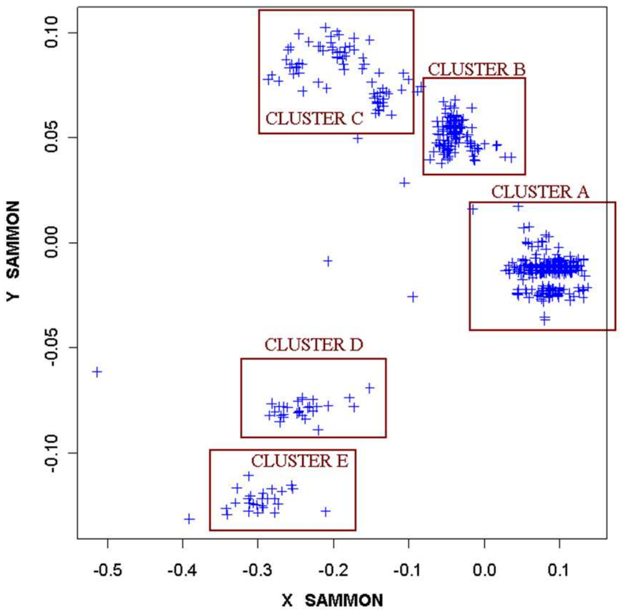

Fig. 7: Groups of coils according to the chemical composition of the steel

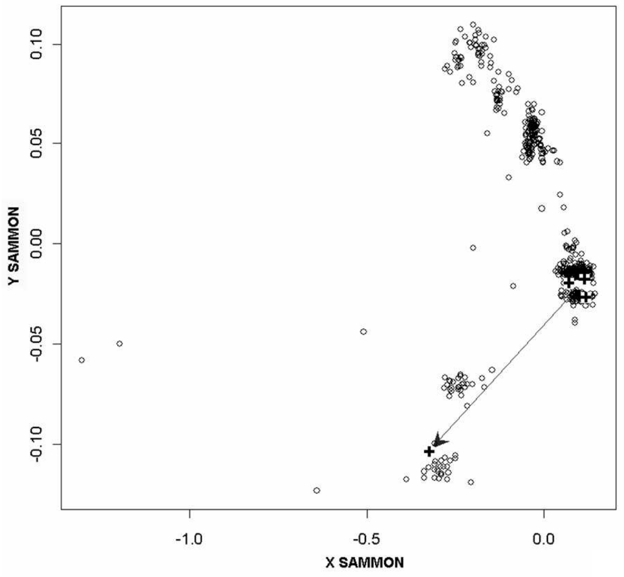

Fig. 8: A dangerous hypothetical event due to the presence of a coil from another steel family in scheduling

were included: width and thickness, chemical composition of the steel and desired final temperature; all of them properly weighted.

Introducing the database composed of 2,436 different coils, a nonlinear Sammon projection was made by obtaining a map similar to the one of Figure 2.

The handling of the map of projected coils was very user-friendly due to the following reasons:

- ○ The selection of the different groups from coils was extremely simple.
- ○ It allowed the expert to detect those coils with parameters which differed greatly from the others.
- ○ It helped to classify new coils easily within the process.

Once the groups of coils were selected, the scheduling for each one was generated semi-automatically according to hierarchical clustering. The number of gaps between coils was considerably reduced, as well as maximum distances between them. Even in the presence of a great number of gaps or distance between coils, this methodology proved highly effective in generating smaller groups to accommodate the gaps, and in adapting to the contours of greater distances.

4 Conclusions

This article presents a successful experience in the use of an MDS to generate bidimensional maps from the multitude of parameters that factor into the process of the creation of scheduling within an HDGL. The use of these bidimensional maps and the dendrograms allows one to graphically group the coils with the least distance between one another, thus ensuring that their scheduling presents no significant problems. An iterative process consisting of the following steps was proposed: projecting the coils according to the parameters selected in a bidimensional map, defining the groups of coils with the closest distances between them, selecting subclusters using hierarchical clustering and separately ordering them . In this way, lists of safer coil combinations are obtained, abrupt jumps between consecutive coils are reduced and potential problems spotted before occurring.

Acknowledgements: The authors thank the 'Dirección General de Investigación' of the Spanish Ministry of Education and Science for the financial support of project no. DPI2007-61090 and the European Union for project no. RFS-PR-06035. The authors also thank the Autonomous Government of La Rioja for its support through the no. 3 Plan Riojano de I+D+i.

References:

[1] V. L. Baum, J.M Stoll. LTV stell hot dip galvanizing line upgrade for exposed automotive products.*Iron and Steel Engineer 71*, 1994, pp. 32–37.

[2] P. Cowling, M. Johansson. Optimization in Steel Hot Rolling. Optimization in Industry. *European Journal of Operational Research*, 1391, 2002, pp. 230–244.

[3] R. D'andrade. U-Statistic Hierarchical Clustering. *Psychometrika* 4, 1978, pp. 58–67.

[4] M. Daszykowski, B. Walczak, D.L. Massart. Projection methods in chemistry. *Chemometrics and intelligent laboratory systems* 1345, 2003, pp. 1–16.

[5] J. Dorn, M. Girsch, G. Skele, W. Slany. Comparison of iterative improvement techniques for schedule optimization. *European Journal of Operational Research* 25, 1996, pp. 349–361.

[6] D. Frey, J. Nimis, H. Wörn, P. Lockemann. Benchmarking and robust multi-agent-based production planning and control. *Engineering Applications of Artificial Intelligence* 16, 2003, pp. 307–320

[7] B. Grabot. Artificial intelligence and soft computing for planning and scheduling: how to efficiently solve more realistic problems. *Engineering Applications of Artificial Intelligence* 14, 2001, pp. 265–267.

[8] I. Harjunkoski, I.E. Grossmann. A Decomposition Approach for the Scheduling of a Steel Plant Production, *Computer & Chemical Engineering* 25, 2001, pp. 647–1660.

[9] S. Johnson. Hierarchical clustering schemes. *Psychometrika* 2, 1967, pp. 241–254

[10] A. Karageorgos, N. Mehandjiev, G. Weichhart, A. Hämmerle. Agent-based optimisation of logistics and production planning. *Engineering Applications of Artificial Intelligence* 16, 2003, pp. 335–348

[11] B. Lally, L. Biegler, H. Henein. A model for sequencing a continuous casting operation to minimize costs. *Iron and Steelmaker* 14, 1987, pp. 63-70.

[12] L. Lopez, M. W. Carter, M. Gendreau. The hot strip mill production scheduling problem: A tabu search approach. *European Journal of Operational Research* 106, 1998, pp. 317–335.

[13] S. Moon, A. N. Hrymak. Scheduling of the batch annealing process, deterministic case. *Computer & Chemical Engineering* 23, 1999, pp. 1193–1208.

[14] J.B. Ordieres Meré, A. González Marcos, J. A. González, V. Lobato Rubio. Estimation of mechanical properties of steel strip in hot dip galvanising lines. *Ironmaking and Steelmaking* 1, 2004, pp. 43–50.

[15] H. Park, Y. Hong, S. Y. Chang. An efficient scheduling algorithm for the hot coil making in the steel mini-mill. *Production planning & control* 13, 2002, pp. 298–306.

[16] D. J. Renn, J. L. Stott, J. F. Vasko. Penalty-based sequencing strategy implemented within a knowledge-based system. Journal of the Operational Society. *Journal of the Operational Society* 50, 1999, pp. 205–210

[17] J. W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers* 18, 1969, pp. 401–409.

[18] M. Soo Suh, Y. Jae Lee, Y. Kwan. A two-level hierarchical approach for raw material scheduling in steelworks. *Engineering Applications of Artificial Intelligence* 10, 1997, pp. 503–515.

[19] R. Tamura, M. Nagai, Y. Nakagawa, T. Tanizaki, H. Nakajina. Synchronized scheduling method in manufacturing steel sheets. *Intelligent Transactions Operational Research* 3, 1998, 189–199.

[20] L. X. Tang, J. Y. Liu, A. Y. Rong, Z. Yang. A mathematical programming model for scheduling steelmaking-continuous casting production. *European Journal of Operational Research* 120, 2000, pp. 423–435.

[21] L. X. Tang, J. Y. Liu, A. Y. Rong, Z. Yang. A multiple traveling salesman problem for the hot rolling scheduling in Baoshan Iron and Steel Complex. *European Journal of Operational Research* 124, 2000, pp. 267–282.

[22] L. Tang, J. Liu, A. Rong, Z. Yang. A review of planning and scheduling systems and methods for integrated steel production. *European Journal of Operational Research* 133, 2001, pp. 1–20.

[23] H. Wai Chun. Scheduling as a multi-dimensional placement problem. *Engineering Applications of Artificial Intelligence* 9, 1996, pp. 261-273.