Chapter 14
Selecting Relevant Clustering Variables in Mass Customization Scenarios Characterized by Workers' Learning

Michel J. Anzanello[1]

Michel J. Anzanello
Federal University of Rio Grande do Sul, Brazil
Av. Osvaldo Aranha, 99 – 5° andar,
Porto Alegre – RS – Brasil - CEP 90.035-190
anzanello@producao.ufrgs.br

Abstract Clustering is an important technique in highly customized production environments, where a large variety of product models is typical. It allows product models with similar processing needs to be aggregated into families, increasing the efficiency of production programming and resources allocation. The quality of the clustering results, however, relies on using a set of relevant clustering variables. Our method selects the best clustering variables aimed at grouping customized product models in families. There are two groups of clustering variables: those generated by expert assessment on the features of products and those predicting the workers' learning rate, obtained by means of learning curve modeling. The method integrates an elimination procedure with a *k*-means clustering technique. The method is illustrated on a shoe manufacturing process.

Abbreviations

LC	Learning curve
MV	Model variables
SI	Silhouette index
SV	Specialists' variables

[1] Michel J. Anzanello is Professor at the Department of Industrial Engineering at Federal University of Rio Grande do Sul, Brazil. He received his Master's degree in industrial engineering from the Federal University of Rio Grande do Sul and his PhD in industrial and systems engineering from Rutgers, The State University of New Jersey. His research interests include multivariate process control, variable selection, and learning curves modeling. His research has been published in *Chemometrics* and *Intelligent Laboratory Systems*, and the *International Journal of Production Research*, among others.

14.1 Introduction and Background

Mass customization environments assume the manufacturing of a large variety of customer guided product models with reduced lot size. Although products may differ in terms of complexity and specific features, they usually require similar machinery and manual processing (Da Silveira *et al.* 2001). In that context, the clustering of models in families with analogous characteristics may enable a more efficient production programming and resource allocation of mass customized production systems.

Clustering tools have been widely used to assign observations (*i.e.*, product models) with similar characteristics to groups (see Jobson 1992, Kaufman and Rousseeuw 2005). Observations allocated in a group are similar to others also in the group and different from those allocated in other groups, without loss of information about the groups (Hair *et al.* 1995). In customized environments, product characteristics (*e.g.*, product complexity, number of operations and parts) have been traditionally used as clustering variables (see Anzanello and Fogliatto 2007). That generalizes information from existing product models to new ones.

In manual-based production environments, the use of clustering variables exclusively related to product characteristics may lead to unsatisfactory assignment of products to families. The way workers adapt themselves to the requirements of a new model should also be included in the clustering procedure. More specifically, the rate at which workers learn the required procedures could provide valuable information about the model's complexity (Uzumeri and Nembhard 1998, Nembhard and Uzumeri 2000), enabling a better assignment of that model to a family. Workers' learning rate can be efficiently estimated by means of learning curve (LC) modeling and then incorporated into the clustering procedure as variables.

The use of many clustering variables, however, may undermine the grouping procedure. As suggested by authors such as Milligan (1989) and Brusco and Cradit (2001), only a limited subset of variables is effectively relevant to establish the cluster structure. The use of irrelevant variables reduces the precision of clustering algorithms, due to the assignment of observations to improper clusters. In that context, selecting the most relevant clustering variables becomes a mandatory step to ensure the formation of consistent families of products.

The sections that follow present an iterative method to select the best clustering variables aimed at assigning customized product models to families with similar characteristics. Clustering variables are chosen from a combination of two groups of variables: (1) those generated by expert assessment on the complexity and features of existing products, and (2) those predicting the workers' learning rate when executing tasks on a new product, obtained by means of LC modeling of data collected from assembly procedures on similar products. The most relevant variables are identified by combining a "leave one variable out at a time" procedure with a *k*-means clustering technique. The clustering performance is evaluated by means of a silhouette index (SI), which indicates the variable to be removed. This iterative process is repeated until a lower bound of remaining variables is achieved, and a graph relating SI and number of remaining variables is generated. The maximum value of SI in that graph identifies the clustering variables to be used in future clustering procedures.

We also address a major pitfall of cluster analyses, namely: how many clusters should be formed? For that matter, the iterative process described above is replicated for a reasonable range of numbers of clusters. The maximum SI for that range identifies the ideal number of clusters.

We illustrate the proposed method in a shoe manufacturing application. We demonstrate that a reduced set of variables, consisting of both experts opinions on product features and LC parameters, leads to the best grouping performance. We also demonstrate that the clustering quality achieved by the selected variables is significantly higher than that obtained by using expert assessed variables alone.

We now provide a brief review of selected LC models, the *k*-means clustering technique, and the fundamentals of SI.

14.1.1 Learning Curves

LCs are mathematical representations of a worker's performance when submitted to a manual task repeatedly. Workers require less time to perform a task as repetitions take place, either due to familiarity with the task and tools required to perform it or because shortcuts to task completion are discovered (Teplitz 1991). There are several LC models proposed in the literature; most notably (1) power models, such as Wright's, (2) hyperbolic models, and (3) exponential models.

Wright's model is the best known LC function in the literature, mostly due to its simplicity and efficiency in describing empirical data. The curve is represented by

$$t = C_1 z^b, \tag{14.1}$$

where z represents the number of units produced, t denotes the average accumulated time (or cost) to produce z units, C_1 is the time (or cost) to produce the first unit, and b is the slope of the curve, such that $-1 \le b \le 0$ (Wright 1936). The parameter b can be assumed as the learning rate parameter, measuring how fast a worker becomes familiar with a new task or product model. For further discussion on b, refer to Jaber (2006) and Jaber and Guiffrida (2007).

Hyperbolic and exponential LC models enable a more precise description of the learning process if compared to Wright's model. The three-parameter hyperbolic model, reported by Mazur and Hastie (1978), is given by

$$y = \frac{m(x+p)}{(x+p+r)} \tag{14.2}$$

with $p + r > 0$. In (14.2), y describes worker's performance in terms of units produced after x time units of operation ($y \ge 0$ and $x \ge 0$), m gives the upper limit of

y ($m \geq 0$), p denotes previous experience in the task, given in time units ($p \geq 0$), and r is the learning rate parameter measured in time units demanded to reach $m/2$ (*i.e.*, half the maximum performance).

Uzumeri and Nembhard (1998) and Nembhard and Uzumeri (2000) modeled performance data from a population of workers exposed to new tasks using the hyperbolic model. The parameters in (14.2) were analyzed to determine workers' learning profiles. Results indicated that fast learners (workers whose LCs had low values of r) presented performance limits (m values) lower than those presented by slow learners (workers with high values of r). The authors recommended the assignment of fast learners to tasks with shorter production cycles, and *vice versa*. In customized environments, which are characterized by short production runs, workers (or teams of workers) associated to low values of r should be prioritized. The parameter m, which describes workers' final performance, is not important in mass customization settings since the number of repetitions in a production run is seldom enough to achieve that level.

One of the most important exponential LC models is the three-parameter model, which is presented in (14.3). Parameters of this model have the same meaning as those of the hyperbolic model,

$$y = m(1 - \exp^{-(x+p)/r}) \tag{14.3}$$

Knecht's model, which is represented in (14.4), is recommended for long production runs, where the learning parameter can present modifications as repetitions take place (Knecht 1974, Nembhard and Uzumeri 2000). The parameters are also as described before.

$$y = \frac{C_1 x^{b+1}}{(1+b)} \tag{14.4}$$

Although learning parameters in (14.1)–(14.4) assume different notations (*i.e.*, b and r) and magnitudes, they are equivalent in representing workers' learning rate and will be addressed as identical through our method.

14.1.2 Clustering Analysis and the Silhouette Index

Data clustering is a widely known multivariate analysis technique that inserts observations (objects) of a population into clusters (groups), such that observations within the same cluster have a high degree of similarity, while observations inserted in different clusters have a high degree of dissimilarity (Jobson 1992, Hair *et al.* 1995, Kaufman and Rousseeuw 2005). Clustering methods have been applied in many areas such as pattern recognition, decision making, and reliability analysis, among others (Taboada and Coit 2007).

There are two main branches of clustering algorithms: non-hierarchical and hierarchical methods. The most popular non-hierarchical clustering method is the *k*-means clustering algorithm, which is widely recognized for its efficiency in grouping observations from datasets (Jain and Dubes 1988).

The *k*-means algorithm inserts each observation into the cluster with the closest centroid. The centroid for each cluster may be calculated or randomly defined by the *k*-means algorithm. The objective function *f* to be optimized by the *k*-means algorithm is (Taboada and Coit 2008):

$$f = \sum_{j=1}^{n} \min_{i \in \{1,\ldots,k\}} \| \mathbf{v}_j - \mathbf{c}_i \|^2 \tag{14.5}$$

where $\mathbf{v}_j$ is the *j*th data vector, $\mathbf{c}_i$ is the *i*th cluster centroid, *k* is the number of clusters to be formed, *n* is the total number of vectors of observations, and ||•|| is the norm operator. The number of clusters *k* is defined by the analyst.

A graphical display, named silhouette graph, evaluates the performance of the clustering procedure by measuring how similar an observation is to observations in its own cluster compared to observations in other clusters (Kaufman and Rousseeuw 2005). An SI that ranges from +1 to –1 is associated to each observation *j*. A value close to +1 identifies observations that are distant from neighboring clusters (*i.e.*, were properly assigned to a cluster); SI_j close to 0 denotes observations that do not clearly belong to one cluster or another; and SI_j close to –1 indicates observations that were probably allocated in the wrong cluster. SI_j is estimated as in (14.6).

$$SI_j = \frac{b(j) - a(j)}{\max\{b(j), a(j)\}} \tag{14.6}$$

where *a*(*j*) is defined as the average distance from the *j*th observation to all the other observations belonging to *j*'s cluster, and *b*(*j*) is the average distance from the *j*th observation to all the observations assigned to the nearest neighbor cluster. Euclidean or Manhattan distances are normally used to calculated distance between observations.

The global quality of a clustering procedure can be assessed by averaging SI over the *n* clustered observations. It is important to mention that SI is independent of the clustering technique. Moreover, Rousseeuw (1987) and Rousseeuw *et al.* (1989) suggest that the SI can be used to determine the best value of *k* (*i.e.*, the number of clusters).

Finally, a major problem in cluster analysis is the selection of variables that truly define clusters with distinct characteristics. Studies have suggested that only a limited subset of variables is effectively important in defining the cluster structure (Fowlkes *et al.* 1988, Milligan 1989, Gnanadesikan *et al.* 1995, Brusco and Cradit 2001), and several approaches have been proposed to select the most relevant variables. The incorporation of irrelevant clustering variables may lead to inaccurate assignments of observations to clusters, in both hierarchical and non-hierarchical cluster analyses (Milligan 1980, 1989).

14.2 Method

The method to select the best variables for clustering purposes relies on two operational steps. In the first step we define the two groups of clustering variables to be used. The first group is subjectively defined based on production staff's expertise, and describe assembly complexity and product parts. The second group of clustering variables is represented by the parameters obtained *via* LC modeling on data collected from the assembly process. Several LC models are considered for that purpose, but only the parameters describing the learning rate are incorporated in the clustering procedure.

In the second step, the groups of variables from Step 1 are evaluated in terms of their efficiency in terms of clustering. We aim at defining which clustering variables are to be used, and the best number of clusters to be considered. For that matter, a "leave one variable out at a time" procedure is used, and the performance of the clustering procedure is evaluated by means of SI. This iterative process is replicated for a range of reasonable number of clusters. We now describe these two operational steps in detail.

14.2.1 Step 1

In Step 1 we define the two groups of clustering variables, which will enable an optimized grouping of product models. We initially select the products to be analyzed. Products with a large number of models (or variations) are preferred, since they potentially allow an *ad-hoc* clustering of models, which leads to an optimized data collection. In addition, market considerations play an important role in product selection: products chosen must be relevant to the company and must present a clear demand for customization.

The first group of clustering variables is obtained through expert analysis and is referred to as specialists' variables (SV). Product models are described in terms of their relevant characteristics, including physical aspects, number of parts, and complexity of its manufacturing operations. Such characteristics may be objectively or subjectively assessed, and either continuous or discrete scales can be used to describe product characteristics.

The second group of clustering variables comes from LC modeling and is referred to as model variables (MV). To obtain those MV readings we must select teams of workers, from which performance data will be collected. Teams must be comprised of workers familiar with the operations to be analyzed. We recommend collecting data from teams with low turnover in that the estimated LC parameters would be able to characterize teams across the time.

LC data is collected from teams performing bottleneck manufacturing operations in each product model. Bottleneck operations are seen as complex manual operations that demand more from workers in terms of learning time and dexterity.

The assignment of product model to teams may be performed as in Anzanello and Fogliatto (2007), or following the company's production plan. Performance data must be collected from the beginning of the operation and should last until no major modifications are noted on the data being collected. This data collection is performed by counting the number of units processed in each time interval.

Performance data collected from the process are analyzed using the LC models in (14.1)–(14.4). These models were chosen based on their performance when modeling learning data (see Nembhard and Uzumeri 2000, Anzanello and Fogliatto 2007). We use the outputs provided by the four LC models to ensure that variations on workers' learning rates are captured.

Estimates of the learning rate parameters may be obtained through nonlinear regression routines available in most statistical packages. The learning rate provided by each LC model will lead to a clustering variable, in Step 2. Note that we use only the learning rate parameter from the LC models. This is justified since production runs in customized environments are too short and do not enable final performance to be evaluated.

14.2.2 Step 2

In Step 2 the objectives are (1) to select the best clustering variables leading to an optimized product grouping procedure, and (2) to identify the ideal value for k (the number of clusters). Clustering variables from both groups (*i.e.*, SV and MV) should be evaluated since a combination of such variables may lead to the best clustering results. In addition, we recommend scaling both SV and MV variables before conducting the clustering process, since the variables may differ in units and magnitude.

We initially define a suitable interval of clusters $[k_{\mathrm{lb}}, K]$ to be evaluated in the iterative process, where k_{lb} is the lower bound on the number of clusters and K is the upper bound. We recommend a lower bound of two clusters ($k_{\mathrm{lb}} = 2$), while K is defined by the analyst. A k-means nonhierarchical clustering procedure using the specified k is run using all clustering variables (SV + MV), and SI is evaluated for that initial scenario. The value of SI obtained for that case is just a reference value, and may be used to assess the performance of the proposed clustering variable selection method.

Next, one variable at a time is left out of the clustering procedure, and an average SI value is computed for each instance. Note that a SI_j value is calculated for each observation j (product model) assigned to a family, and then an average SI is estimated. Once all clustering variables have been tested (*i.e.*, omitted once), the variable responsible for the maximum average SI is eliminated as the one that contributes the least in separating the products in families. The iterative procedure is then repeated for the SV + MV−1 remaining variables, and the average SI is again evaluated after each variable is omitted. We repeat this procedure until a lower bound of remaining variables is reached. A graph relating the average SI and the number of retained variables may be generated to identify the ideal number of variables to be used in clustering applications. A hypothetical example of the average SI profile generated by variable elimination is illustrated in Figure 14.1 for $k=3$. Note that the average SI increases when fewer variables are retained. In this case, the maximum average SI is obtained when 2 out of 10 variables are retained.

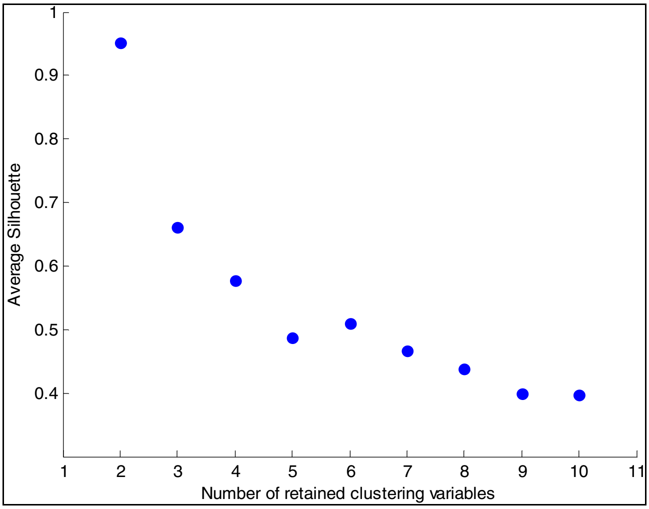

Figure 14.1 Hypothetical average SI profile with clustering variable elimination

In order to define the best number of clusters to be used, we then set $k=k+1$, and restart the iterative elimination procedure with the SV + MV clustering variables. The variable elimination is repeated as described above and the maximum average SI is stored for that k. The iterative process stops when $k=K$.

The maximum SI value for each k, as well as the variables leading to that value, may be represented in a table. The overall maximum SI indicates the best number of clusters k as well as the clustering variables to be used.

14.3 Numerical Case

The proposed method was applied in a shoe manufacturing plant. Shoe producers have been challenged by decreasing lot sizes in the past decade, forcing their mass production configuration to adapt to an increasingly customized market. In terms of production planning, it is mandatory to cluster such large variety of models to make resource allocation more efficient. The proposed method for selecting the best clustering variables was tested in the sewing stage of the shoe manufacturing company. The sewing is the bottleneck production stage, from which data for the LC modeling were collected.

20 shoe models were considered in the study. SV were defined with respect to manufacturing complexity of the upper part of the shoes: parts complexity (deployed into four categories), number of parts in the model, and type of shoe. The first five variables were subjectively assessed by company experts (assembly line supervisors and operators, and sales department personnel) using a three-point scale, where 3 denotes the highest complexity or number of parts. The variable type of shoe has two levels: one for shoes and sandals, and two for boots, which tend to be more complex in terms of assembly. Table 14.1 displays the 6 SV and the respective ID.

Performance data were collected from three teams of workers. Shoe models were directed to teams according to the company’s production planning. Performance data collected were registered as number of pairs produced in 10 min intervals, and were adjusted to the models in (14.1)–(14.4). The resulting learning parameters from the LC modeling were scaled in the interval 0–3 to ensure consistency with the SV and referred to as model variables (MV), as presented in Table 14.2.

The proposed method was run for *k* in the interval [2, 7]. Table 14.3 displays the average SI profile with the clustering variable elimination for each *k*. The bold value indicates the maximum average SI for each case, while the ID of the selected variables is presented at the bottom of the same table. A reduced number of variables is preferred in all cases, as implied by the increasing SI profile. That

Table 14.1 Specialists’ clustering variables (variable ID presented in parenthesis)

Shoe ID	Specialists’ clustering variables (ID in parenthesis)					
	Sewing complexity (1)	Adornments complexity (2)	Lining complexity (3)	Material complexity (4)	Number of parts (5)	Type of shoes (6)
Shoe1	1	1	1	1	2	2
Shoe2	1	1	2	2	1	1
Shoe3	1	1	2	1	2	1
Shoe4	2	1	1	1	2	1
Shoe5	1	2	2	1	2	1
Shoe6	1	1	1	1	1	1
Shoe7	1	1	1	2	1	1
Shoe8	2	3	1	1	1	1
Shoe9	2	2	1	2	2	1
Shoe10	2	2	1	2	1	1
Shoe11	1	3	1	1	2	1
Shoe12	2	2	1	2	2	1
Shoe13	1	2	1	2	2	1
Shoe14	1	2	1	2	2	1
Shoe15	2	3	2	3	2	1
Shoe16	1	3	2	3	2	1
Shoe17	2	2	2	2	3	1
Shoe18	2	3	2	3	2	1
Shoe19	2	3	2	2	2	2
Shoe20	2	3	2	2	3	1

Table 14.2 Model clustering variables (variable ID presented in parenthesis)

Shoe ID	Model clustering variables (ID in parenthesis)			
	Hyperbolic (7)	Exponential (8)	Wright (9)	Knecht (10)
Shoe1	2.00	3.00	0.73	2.48
Shoe2	1.97	1.54	1.26	2.61
Shoe3	2.27	3.00	0.86	2.53
Shoe4	2.12	2.00	1.01	2.58
Shoe5	0.60	0.94	0.51	2.46
Shoe6	0.14	0.24	1.82	2.72
Shoe7	1.67	1.52	1.07	2.60
Shoe8	0.97	0.65	1.02	2.56
Shoe9	0.58	0.44	1.53	2.70
Shoe10	0.36	0.54	1.17	2.64
Shoe11	0.81	1.44	2.41	2.90
Shoe12	1.09	0.89	2.69	2.92
Shoe13	0.42	0.57	3.00	3.00
Shoe14	0.61	0.65	2.59	2.89
Shoe15	0.93	1.92	0.80	2.52
Shoe16	0.88	1.30	1.29	2.65
Shoe17	0.50	0.59	1.80	2.74
Shoe18	3.00	1.69	2.97	2.94
Shoe19	0.26	0.49	0.63	2.49
Shoe20	0.71	1.12	1.56	2.69

indicates that the use of all clustering variables incorporates noise to the clustering procedure and decreases the grouping performance. In addition, we note that the best reduced sets for all values of k evaluated are composed of a combination of variables belonging to SV and MV. That demonstrates that both the specialists' assessment, represented by SV, as well as the workers' learning process, represented by MV, play an important role in the clustering procedure.

According to Table 14.3, $k=2$ is the best number of clusters to be considered when using a k-means procedure, and variables 6 and 10 (type of shoe and Knecht's learning rate parameter, respectively) should be used. An analysis based in four clusters (*i.e.*, $k=4$) may also lead to satisfactory results when variables 3 and 10 (lining complexity and Knecht's learning rate parameter, respectively) are considered.

It is important to mention that a random value of the order 10^{-4} was added to the SV variables due to the reduced number of points on the scale describing those variables. This enables the k-means algorithm to define clusters even when a reduced number of variables are considered, especially during the elimination steps for upper values of k. That modification does not significantly affect the precision of the clustering procedure, according to our experiments. The addition of a random value may be avoided if products are described by scales consisting of larger number of points (*e.g.*, a 1ten-point scale) or if a continuous scale is adopted.

Figure 14.2 brings the silhouette graph for $k=2$ when only the SVs are considered. This leads to an average SI of 0.4720. Each horizontal line represents the adherence of observation j (*i.e.*, a shoe model) to the cluster it was assigned to. In

Table 14.3 Average SI and selected clustering variables

Number of retained clustering variables	Number of clusters (k)					
	2	3	4	5	6	7
2	0,95	0,71	0,83	0,68	0,71	0,53
3	0,80	0,63	0,74	0,71	0,71	0,60
4	0,73	0,59	0,63	0,63	0,73	0,62
5	0,67	0,55	0,57	0,50	0,64	0,58
6	0,62	0,46	0,51	0,44	0,53	0,51
7	0,60	0,44	0,52	0,45	0,45	0,51
8	0,56	0,42	0,44	0,42	0,44	0,45
9	0,47	0,42	0,40	0,40	0,44	0,42
10	0,43	0,37	0,37	0,37	0,37	0,41
Retained clustering variable ID	6	2	3	4	3	2
	10	8	10	8	7	7
				10	9	9
					10	10

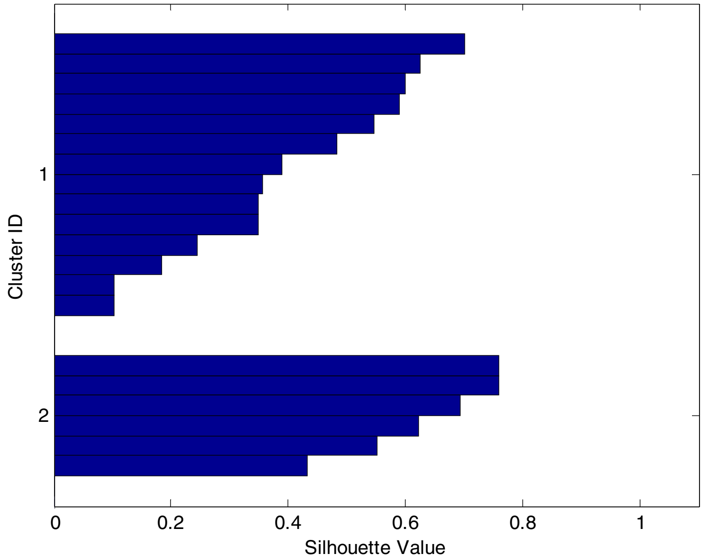

Figure 14.2 Silhouette graph using only SV

Figure 14.2, 14 shoes were assigned to cluster 1 and 6 to cluster 2. Some observations included in cluster 1 assume very low SI values, denoting an improper cluster assignment.

Figure 14.3 illustrates the silhouette graph when using the selected variables and *k* = *2*. There is a remarkable improvement in the adherence of the observations to the clusters. The same graph demonstrates that most product models actually belong to cluster 2, and not to cluster 1 as previously indicated by the SV alone. The average SI for this case is 0.9588.

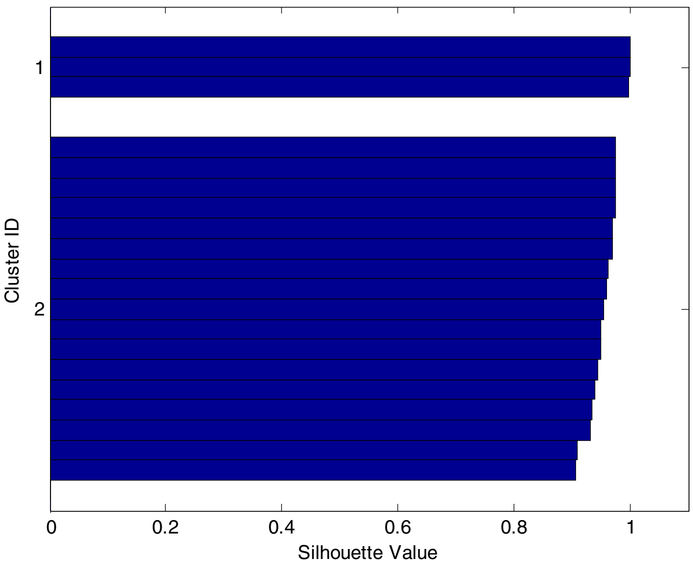

Figure 14.3 Silhouette graph using the selected variables

It is important to emphasize that small variations on the average SI may occur due to two factors: (1) the random value added to the SV variables, although small, may result in slightly different allocations of observations to clusters, and (2) the *k*-means algorithm used in this case study randomly defines its clusters centroids (also referred as "seeds"). This may lead to different allocations of observations to clusters even when using the same data and consequently affect the average SI.

14.4 Conclusion

Clustering is an important technique in highly customized production environments, where a large variety of product models and reduced lot sizes are typical. It allows product models with similar characteristics and processing needs to be aggregated into families, increasing the efficiency of production programming and resources allocation. The quality of the clustering results, however, relies on using a limited set of relevant clustering variables.

This chapter presented an iterative procedure aimed at selecting the most relevant clustering variables in processes where workers' learning takes place. Workers' learning rates were addressed by means of LC modeling, and the estimated LC parameters were incorporated in the grouping procedure as clustering variables. The best variables were identified by combining a "leave one variable out at a time" procedure with a k-means clustering technique. The less relevant variables were identified by means of the SI, which also defined the ideal number of clusters.

When applied to a shoe manufacturing case study, the method led to significant reduction of clustering variables needed for grouping, while increasing the clustering quality compared to using only the variables describing product's characteristics. We also demonstrated that a combination of variables assessed by production experts and variables generated by the LC modeling leads to the best set of clustering variables for a considerably wide range of numbers of clusters.

References

Anzanello M, Fogliatto F (2007) Learning curve modelling of work assignment in mass customized assembly lines. International J Product Research 45:2919–2938

Brusco M, Cradit J (2001) A variable-selection heuristic for k-means clustering. Psychometrika 66:249–270

Da Silveira G, Borestein D, Fogliatto F (2001) Mass customization: literature review and research directions. International J Production Economics 72:1–13

Fowlkes E, Gnanadesikan R, Kettenring J (1988) Variable selection in clustering. J Classification 5:205–228

Gnanadesikan R, Kettenring J, Tsao S (1995) Weighting and selection of variables for cluster analysis. J Classification 12:113–136

Hair J, Anderson R, Tatham R, Black W (1995) Multivariate Data Analysis with Readings. Prentice-Hall, Englewood Cliff, NJ

Jaber M (2006) Learning and forgetting models and their applications. In: Badiru AB (ed) Handbook of Industrial and Systems Engineering. CRC Press-Taylor and Francis Group, Baca Raton, FL

Jaber M, Guiffrida A (2007) Observations on the economic order (manufacture) quantity model with learning and forgetting. International Transactions in Operational Research 14:91–104

Jain A, Dubes R (1988) Algorithms for clustering data. Prentice Hall, Englewood Cliffs, NJ

Jobson J (1992) Applied Multivariate Data Analysis, Volume II: Categorical and Multivariate Methods. Springer, New York

Kaufman L, Rousseeuw P (2005) Finding Groups in Data: An Introduction to Cluster Analysis. Wiley Interscience, New York

Knecht G (1974) Costing, technological growth and generalized learning curves. Operational Research Q 25:487–491

Mazur J, Hastie R (1978) Learning as Accumulation: A Reexamination of the Learning Curve. Psychological Bulletin, 85:1256–1274

Milligan G (1980) An examination of six types of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika 45:325–342

Milligan G (1989) A validation study of a variable-weighting algorithm for cluster analysis. J Classification 6:53–71

Nembhard D, Uzumeri M (2000) An Individual-based description of learning within an organization. IEEE Transactions Engineering Management 47:370–378

Rousseeuw P (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J of Computational and Applied Mathematics 20:53–65

Rousseeuw P, Trauwaert E, Kaufman L (1989) Some silhouette-based graphics for clustering interpretation. Belgian J of Operations Research, Statistics and Computer Science 29

Taboada H, Coit D (2007) Data clustering of solutions for multiple objective system reliability optimization problems. Quality Technology and Quantitative Management J 4:35–54

Taboada H, Coit D (2008) Multi-objective scheduling problems: determination of pruned Pareto sets. IIE Transactions 40:552–564

Teplitz C (1991) The Learning Curve Deskbook: A Reference Guide to Theory, Calculations and Applications. Quorum Books, New York

Uzumeri M, Nembhard D (1998) A Population of learners: a new way to measure organizational learning. J Operation Management 16:515–528

Wright T (1936) Factors affecting the cost of airplanes. J Aeronautical Science 3:122–128