Mathematical formulation for a mixed-model assembly line balancing problem with stochastic processing times

Carlos D. Paternina-Arboleda, PhD
Professor, Departamento de Ingeniería Industrial, Universidad del Norte, Barranquilla, Colombia
Email: cpaterni@uninorte.edu.co

Jairo R. Montoya-Torres, PhD
Professor, Departamento de Procesos Productivos, Pontificia Universidad Javeriana, Bogotá, Colombia
Email: jairo.montoya@javeriana.edu.co

Edward Moscoso, Ing.
Professor, Departamento de Ingeniería Industrial, Universidad del Norte, Barranquilla, Colombia

Abstract
This paper focuses on the problem of line balancing in an assembly line. Although there are numerous studies published on the various aspects of the problem, most of the work focuses on the single-model problem, while the number of studies on the mixed-model assembly line balancing problem assume that processing times of tasks are deterministic. Task time variation however has a great impact with short cycle times. So, because there is always some variation when performing tasks, our research addresses stochastic processing duration of tasks. We propose an analytical analysis of the problem by using a multi-objective mathematical formulation that aims to minimize the total number of stations, to maximize system's productivity and to find a production sequence that smoothes system's operation.

Keywords
Mixed-model assembly line, stochastic processing times, mathematical formulation.

1. Introduction

In today's intensively competitive and customer centric market, manufacturers try their best to attract every single customer with more varieties of their products, by delivering product at shorter cycle times. A typical case of this situation is the automotive industry: any model has some options and customers can choose a model based on their preference and financial capability (i.e. with or without airbags, with or without air-conditioner, and so on) (Jin and Wu, 2002). As a consequence, different options mean that different parts should be put on the basic model. Due to the high cost of building and maintaining an assembly line, the manufacturers produce one model with different options or several models on a single assembly line. Under these circumstances, the mixed-model assembly line balancing problem arises to smooth the production and decrease the cost.

In the literature, the assembly line balancing problem has received considerable attention from researches in the last 40 years. The problem is *NP*-complete, since with a single model and tasks with no precedence relations, it is easy to reduce the problem to the *NP*-complete bin-packing problem. Hence, the combinatorial nature of the mixed-model line balancing problem makes difficult to obtain optimal solution, though the mixed-model line is the most frequently encountered type in industry due to the pressure of producing several models to attain high customers satisfaction (Gokcen and Erel, 1998). The mixed-model assembly line balancing problem can be stated as follows: Given *N* models, the set of tasks and a cycle time associated with each model, the performance times of the tasks and the set of precedence relations which specify the permissible ordering of the tasks for each model, the problem is to assign the tasks to an ordered sequence of stations such that the precedence relations are satisfied and some performance measure is optimized.

Although there are numerous studies published on the various aspects of the problem, most of the work focuses on the single-model assembly line balancing problem by proposing optimum-seeking procedures or heuristic algorithms. Interested readers may see (Baybars, 1986, Ghosh and Gagnon, 1989). The number of studies on the mixed-model assembly line balancing problem, however, is relatively small. Thomopoulos (1967) relates costs to different types of inefficiency, for example, idle time and utility time, and formulates the model to minimize the total penalty cost with a heuristic method for solving it. Roberts and Villa (1970) constructed a binary integer-programming model in order to propose an optimum-seeking procedure. Nevertheless, the extensive number of variables and constraints prohibits the applicability of the model for even small size instances. Gokcen and Erel (1998) proposed a binary integer formulation for the mixed-model assembly line balancing problem and developed some properties that prevent the fat increase in the number of variables. These authors, later, use a shortest-route formulation to minimize the task time for different models with precedence constraints (Erel and Gokcen, 1999).

These papers, however, assume that processing times of tasks are deterministic. Since there is always some variation when performing tasks, and since task time variation has a greater impact with shorter cycle times, this research addresses stochastic processing duration of tasks. To the best of our knowledge, only the work of McMullen and Frazier (1997) has addressed this stochastic version of the problem. These authors proposed a heuristic procedure by modifying the heuristic algorithm proposed in (Gaither, 1996) in order to incorporate stochastic task durations. In this paper, we propose an analytical analysis of the problem by using a multi-objective mathematical formulation that aims to minimize the total number of stations, to maximize system's productivity and to find a production sequence that smoothes system's operation.

2. Mathematical model

The general solution procedure consists on three steps:

1. From the requirements for each product, demands, processing times of tasks and precedence relations, an equivalent (or composite) single-product model is defined,
2. Cycle time and the minimum number of stations are defined, and
3. Productivity maximization, or minimization of total line idle time.

We use a hierarchical multi-objective model in which the line is firstly balanced for a equivalent single-product model in order to minimize the total number of stations in the line and to minimize the idle time. The concept of bottleneck station is finally used to determine the sequencing of products that maximize the objective.

For a given precedence relation matrix of a set of *N* models (products), the proposed model is based on the following assumptions:

- Tasks processing times associated with each model are normally distributed. Common tasks among the models do not need to have the same processing time.
- Processing times are independent and do not depend on workers' training.

- Precedence relations between tasks of each model are known.
- No work-in-process inventory buffer is allowed between stations.
- Demand of each model is deterministic and known beforehand.

2.1 Definition of the equivalent (composite) product

The equivalent product is defined as one resulting from the reduction of the stochastic task durations of the mixed-model production demand into composite stochastic task durations. The reduction procedure employed in this paper is the same proposed in (McMullen and Frazier, 1997). The variable d_n is the demand for product *n*, and the total demand, D_T, for all of the *N* different products demanded is given by $D_T = \sum_{n=1}^{N} d_n$. The weight w_n of product *n* is computed as $w_n = d_n / D_T$. With weights determined, proportions of product-mix for each product are known. This information is then used to compute weighted average task durations for each individual task. The mean processing time of task *i* of product *n* is denoted as $(t_i)_n$. The variable t_i represents thus the weighted mean (equivalent) processing time for task $_i$, and is estimated as $t_i = \sum_{n=1}^{N} w_n (t_i)_n$.

The estimates of the equivalent processing times for the equivalent (or composite) product have now been attained. These estimates are used in conjuction with a specified coefficient of variation (*CV*) to obtain standard deviations. Hence, the estimated standard deviation of task *i* for product *n*, denoted as σ_{in}, is computed as $\sigma_{in}=t_{in}\times CV_{in}$. The estimated of the standard deviation for the equivalent product task *i*, denoted as σ_i, is calculated as follows:

$$\hat{\sigma}_i = \sqrt{\left(\frac{1}{D_T - 1}\right)\left[\sum_{n=1}^{N} d_n\left((t_i)_n - t_i\right) + \sum_{n=1}^{N}(d_n - 1)\sigma_{in}^2\right]} .$$

The mixed-model has been reduced to a single-product model with processing times of tasks following a normal distribution with parameters N(t_i, $\hat{\sigma}_i$).

Cycle time, *C*, is determined as *C*=*pm*/*hp*, where *pm* represents the productive minutes per hour and *hp* is the desired hourly output. Cycle time is measured in minutes per unite. The theoretical minimum number of stations, $K°$, is computed as $K° = \left\lceil \sum_{i=1}^{A} t_i / C \right\rceil$, where *A* is the total number of tasks, and the symbol $\lceil\ \rceil$ is the ceiling function.

At this point, the actual line balancing procedure commences.

2.2 Line balancing

The actual line balancing procedure proposed in this paper is based on a mathematical formulation that considers the reduced single-product model studied previously. Binary decision variables are defined as:

$$x_{ik} = \begin{cases} 1 & \text{if task } i \text{ is assigned to the station } k \\ 0 & \text{othewise} \end{cases}$$

Objective function

The objective is to minimize the number of stations utilized:

$$Min\sum_{i=1}^{A}\sum_{k=1}^{K}C_{ik}x_{ik}$$

where $C_{ik}=(K^{\circ})^{k-1}$ is a logical cost depending on $K°$.

Constraints

<u>Capacity constraints</u>: Let S_k be the set of tasks assigned to the k-th station, then

$$E(S_k)=\sum_{i\in S_k}t_i x_{ik}$$

and

$$V(S_k)=\sum_{i\in S_k}\hat{\sigma}_i^2 x_{ik}$$

where $E(S_k)$ and $V(S_k)$ are, respectively, the expected value and the variance of the processing time at the k-th station. The capacity constraint for a 0.99 probability of completion of all the tasks at this station is

$$E(S_k)+2.33\sqrt{V(S_k)}\leq C,\ k=1,2,\ldots,K.$$

This constraint has a non-linear factor that increases the complexity of the model. Hence, in order to solve this, it is possible to use one of the following procedures.

a) Linearity: this is the suitable choice but the square-root contains integer decision variables that produces a discontinuous function that does not allow us to use Taylor's series.
b) Approximation: the left side of the equation is based on statistical laws that can be used to approximate the function. In fact, the non-linearity and non-continuous complexity of the function are due to the stochastic feature of the constraint. Let consider that the standard deviation of the capacity results from the sum of the standard deviation of processing times at that station. The restriction may be expressed as $\sum_{i=1}^{A}t_i x_{ik}+2.33\sum_{i=1}^{A}\hat{\sigma}_i x_{ik}\leq C$, $k=1,2,\ldots,K$. It is to note that this approximation will assure that the line is always balanced with a capacity utilization near to the optimum.

<u>Assignment constraints</u>: this set of constraints assures that tasks of each model are assigned to at most one station and can be written as

$$\sum_{k=1}^{K}x_{ik}=1,\ i=1,\ldots,A.$$

<u>Precedence constraints</u>: for the precedence diagram of the mixed-model problem, we define $CP=\{(u,v)\}$, where task v is an immediate follower of task u. The constraint is thus expressed as

$$x_{vb}\leq\sum_{j=1}^{b}x_{uj},\ b=1,\ldots,K \text{ and } (u,v)\in CP.$$

<u>Station constraints</u>: we define $CM=\{(u,v)\}$ as the set of task pairs that have to be performed on the same station, and $CD=\{(u,v)\}$ as the set of task pairs that can not be performed on the same station. This can be accomplished by introducing the following constraints:

$$\sum_{k=1}^{K}x_{uk}x_{vk}=1,\ (u,v)\in CM$$

$$x_{ub}+x_{vb}\leq 1,\ b=1,\ldots,K \text{ and } (u,v)\in CD$$

Bottleneck station: taking into account the previous considerations, the relative workload at a station is defined by the following approximation:

$$C_k = \sum_{i \in S_k} t_i + 2.33 \sum_{i \in S_k} \hat{\sigma}_i \text{ , } k=1,\ldots,K.$$

The bottleneck station is the station having the maximum workload, $k^b = \arg\max_k (C_k)$. Once the bottleneck has been identified, the percentage of utilization for each model (product) can be computed as

$$C_{bn} = \sum_{i \in S_k} (t_i)_n + 2.33 \sum_{i \in S_k} \sigma_{in} \text{ , } n=1,\ldots,N.$$

Let $CB=\{C_{b1},C_{b2},\ldots,C_{bN})$ be a $(1\times N)$-vector defined by those percentages.

Once the minimum number of stations has been determined, it is possible to minimize the total idle time as follows:

$$Min\left\{KC - \sum_{k=1}^{K} C_k\right\}.$$

At this point, it is possible to determine a sequencing strategy for the products.

2.3 Sequencing

For each model (or product) n, we want to produce a total of d_n units during the production planning horizon. Let $p=\max\{d_n\}$, the biggest demand among all products. The quantities of each model that most be produced during the planning horizon are defined as $U_n = d_n / p$, $n=1,\ldots,N$. The cycle must be repeated p times to satisfy demand of each model. Hence, the total number of unites produced per cycle are $U_T = \sum_{n=1}^{N} U_n$.

Once the workload at the bottleneck station is calculated, as well as the quantities for each product during the cycle, it is possible to define the sequencing strategy in order to minimize the maximum utilization of the bottleneck station for each product of the mixed-model with respect to the average utilization at this station.

Let define a binary variable y_{rn} as follows:

$$y_{rn} = \begin{cases} 1 & \text{if a unit of model } n \text{ is scheduled at the } r\text{-th position} \\ 0 & \text{othewise} \end{cases}$$

Before implementing the model, it is necessary to define the following notation.

W_n is the advance of the n-th model in the sequence, $W_n = \begin{cases} w_n & r=1 \\ rw_n - \sum_{q=2}^{r} y_{(q-1)} & \forall r \neq 1 \end{cases}$,

DC is standard deviation of the relative workload of n-th model against C_b during the sequencing of the bottleneck station, $DC = \begin{cases} \left|C_{bn} - C_b\right| & r=1 \\ \left|C_{bn} + \sum_{q=2}^{r}\sum_{n=1}^{N} C_{bn} y_{(q-1)} - C_b\right| & \forall r \neq 1 \end{cases}$.

In addition, it is mandatory to compute $\sum_{n=1}^{N} C_{bn} y_{1n}$ for r=1. This represents that the first product has been released.

Objective functions
The formulation searches for the optimization of the following objective functions at each step of the sequencing procedure:

$Min \sum_{n=1}^{N} DC_{bn} y_{rn}$, r=1,…,U_T, and

$Max \sum_{n=1}^{N} W_n y_{rn}$, r=1,…,U_T.

where *r n* states which product type will be assigned to the *r*-th position in the schedule.

Constraints
While running the sequencing procedure, it is mandatory to be sure that at least one unit of any product type is released to the program to be scheduled. This can be expressed by the following set of constraints:

$\sum_{n=1}^{N} y_{rn} = 1$, r=1,…,U_T.

3. Illustrative example

In this section, we apply the proposed model to a mixed-model assembly line balancing problem with stochastic processing times of tasks. Precedence relations, processing times, demands and coefficients of variation for each type of product are shown in table 1. Figure 1 is the precedence graph for this example.

Table 1. Performance data for the illustrative example

Task	Immediate predecessor	Model (or product)		
		Processing time (sec.)		
		P_1	P_2	P_3
A	-	30	25	35
B	A	15	10	10
C	A	5	7	5
D	A	40	40	45
E	D	50	45	45
F	D	30	40	40
G	B, C	15	15	20
H	E, G	65	60	65
I	H	25	25	0
J	F	50	50	60
K	I, J	45	40	40
L	K	5	5	5
Demand per week		500	250	750
Coefficient of variation (*CV*)		0.1	0.2	0.1

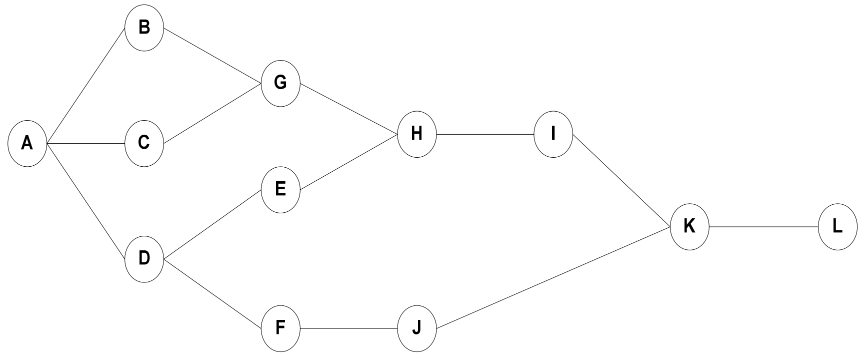

Figure 1. Precedence graph for the example.

The mathematical program was implemented on a PC AMD ATLON XP (1500MHz) using the student version of Premium Solver release 3.5. Under these conditions, the mathematical program was solved in 8 seconds. Results obtained for the minimization of the number of stations and the maximization of productivity, by using the proposed hierarchical solution procedure, are shown, respectively, in table 2 and 3.

Table 2. Results for minimum number of stations

Station (k)	S_k	C_k (sec)	Idle time (sec)
1	{A,D}	94.552	1.448
2	{B,C,F,G}	91.725	4.275
3	{E}	59.826	36.174
4	{H}	82.078	13.922
5	{I,J}	88.852	7.418
6	{K,L}	59.827	36.173
		Total	99.409

Table 3. Results for maximal productivity

Station (k)	S_k	C_k (sec)	Idle time (sec)
1	{A,D}	94.552	1.448
2	{C,E}	66.861	29.139
3	{B,F,G}	84.690	1.31
4	{J}	70.265	25.735
5	{H}	82.078	13.922
6	{I,K,L}	78.144	17.856
		Total	99.409

Looking at tables 2 and 3, we can see that the maximization of productivity allows us to minimize the number of stations (column Ck). The maximum idle time observed in table 2 corresponds to station 2 with 29.14 seconds, while in table 1 the maximum idle time is for stations 3 and 6. Another important factor is that tasks are distributed in a different manner when solving the second objective. This is a factor that allows to a simultaneous optimization of both objectives. Finally, it is to notice that the total idle time of the system is independent of the distribution of tasks on stations, but depending on the number of stations for a given instance.

This illustrative example was also solved using the approximation for the capacity constraint explained in section 2.3. Results are shown in table 4. The number of stations remains the same, as well as the total idle time of the system. The bottleneck station is still the station number 1, but the mathematical program sets the shortest tasks within the same station, for instance station number 2, while this heuristic does not by needing the same number of stations. This justifies the appropriateness of the proposed approximation for the capacity constraint in the mathematical program presented in section 2.3.

Table 4. Results using approximation procedure for capacity

Station (k)	S_k	C_k (sec)	Idle time (sec)
1	{A,D}	94.552	1.448
2	{E,B,C}	81.751	14.249
3	{F,G}	69.8	26.2
4	{H}	82.078	13.922
5	{I,J}	88.582	7.418
6	{K,L}	59.827	36.173
		Total	99.409

5. Conclusion

This paper presented a mathematical formulation for the mixed-model assembly line balancing problem with stochastic processing times of tasks. The formulation is based on the definition an equivalent multi-objective single-model. Experimental results at the implementation stage of the model illustrate the robustness of the proposed model, which suggests the practical applications of the model since simplifications allow for a short computation time while the solution value is near to the optimum. The proposed model serves a starting point for researchers in the field, and may be used as a validation tool for heuristic procedures solving multi-objective real-life line balancing problems.

References

Baybars I. (1986) A survey of exact algorithms for the simple assembly line balancing problem. *Management Science*, 32, 909-932.

Erel E., Gokcen H. (1999) Shortest-route formulation of a mixed-model assembly line balancing problem. *European Journal of Operational Research*, 116, 194-204.

Gaither N. (1996) *Production and Operations Management*, 7th Edition, Duxbury Press, Boston, USA.

Ghosh S., Gagnon J. (1989) A comprehensive literature review and analysis of the design, balancing and scheduling of assembly systems. *International Journal of Production Research*, 27, 637-670.

Gokcen H., Erel E. (1998) Binary integer formulation for mixed-model assembly line balancing problem. *Computers and Industrial Engineering*, 34, 451-461.

Jin M., Wu S.D. (2002) A new heuristic method for mixed-model assembly line balancing problem. *Computers and Industrial Engineering*, 44, 159-169.

McMullen P.R., Frazier G.V. (1997) A heuristic for solving mixed-model line balancing problems with stochastic task durations and parallel stations. *International Journal of Production Economics*, 51, 177-190.

Roberts S.D., Villa C.D. (1970) On a multiproduct assembly line balancing problem. *AIIE Transactions*, 2, 361-364.

Thomopoulos N.T. (1967) Line balancing sequencing for mixed-model assembly. *Management Science*, 14, 59-75.