Medium-Term Production and Staff Planning in the Automotive Industry

Claas Hemig, Jürgen Zimmermann
Clausthal University of Technology, Departement for Business Administration, Julius-Albert-Str. 2, D-38678 Clausthal-Zellerfeld, Germany, {claas.hemig, juergen.zimmermann}@tu-clausthal.de

In this paper we present a Dynamic Programming approach for an integrated production and staff planning problem in the automotive industry. The planning horizon has a length of three to six years. We focus on one shop with parallel production lines and search for a least cost solution subject to the given flexibility of the production resources. The flexibility is induced by the adjustment screws like production speed and time, permanent and temporary workers, and the distribution of a prescribed production program between the lines. In addition, we consider labor legislation, technical restrictions and in-plant-agreements. To obtain a tractable problem size, i.e. to guarantee the operational usability of the software tool to be developed, we modify the basic approach and introduce a heuristic to solve the embedded staff planning problem. A prototypical implementation shows that the approach works well.

Keywords: Production Scheduling, Real World Scheduling

1 Introduction

Since the automotive market becomes tighter and more dynamic the mid- to long-term planning of a flexible automotive production system is a task of increasing importance. Due to the various influences on its production systems the automotive industry has recognized that the mid-term planning of its production capacities and staff size obtains better solutions incorporating OR-methods. The highly complex problem arising from the concurrence of technical restrictions, labor legislation and in-company-agreements requires efficient solution techniques. In cooperation with one of our industrial partners we develop a software tool to support the production planners in several plants. After a detailed problem formulation in Chapter 2, we present a basic solution approach in Chapter 3, focussing on two subproblems to be solved. The paper is closed by a conclusion and an outlook on further research.

2 Problem Formulation

A typical automotive plant mainly consists of a body shop, a paint shop, and a final assembly. In each shop there may exist several parallel production lines that are organized as a continuous flow production. On the lines M types of products, called models, are manufactured. Each line can either be a solitary line (i.e. only one model can be produced) or a mix line (i.e. several different models can be produced). Buffers of limited capacities between the shops can store intermediate products to decouple subsequent shops. In this paper we focus on the final assembly with L parallel production lines where some models can be produced by more than one line. For simplicity we assume that at most two products can be manufactured per line, but we can easily extend the concept to more than two products per line. The following sections describe the decisions that have to be made in each period (usually a week or a month) of the planning horizon T using the flexibility instruments of a shop. Furthermore, we describe the most important restrictions as well as the objective function of our mathematical model.

2.1 Shift Model and Cycle Time

The fundamental decisions we make for every line l and period t are the shift model $sm_{lt} \in \{1,\ldots,|\mathcal{SM}|\}$ and the cycle time $ct_{lt} \in \{1,\ldots,|\mathcal{CT}_l|\}$. Every shift model induces the number of shifts SQD taking place on one day (early, late and night shift), where $SQD = 1$ indicates only early shifts, $SQD = 2$ indicates early and late shifts etc. In addition, every shift model induces the operating time OT along the way products are manufactured as well as the working time WT as the sum of OT and the paid breaks. Both times are measured in minutes. Third, if a shift model provides work breaks shorter than labor law allows, an extra percentage of employees SM_X is needed so that the workers can alternate while taking their breaks according to law. The cycle time ct_{lt} is measured in products per minute and determines the basic demand for employees BMR during one shift.

Let us assume a shift model with an operating time of 6220 minutes in three shifts and an extra amount of employees of 5% as well as a cycle time with a value of 0.4 and a BMR of 600 employees. Then the overall demand for employees results in $BMR \cdot SQD \cdot (1 + SM_X) =$ 1,890 employees. Independent of the shift model and cycle time we need extra employees taking into account that a certain percentage of the workers are absent due to illness or holidays. The production capacity of the considered production line is given by $OT \cdot ct =$ 2,488 products.

The selection of the shift model and the cycle time on each line is independent from each other. As a matter of principle we can choose them arbitrarily every period, but a change of the shift model as well as the cycle time evokes significant organizational changes. Consequently, a change is allowed to occur only once in a given number of subsequent periods $SM^{min}_{sm,l}$ or $CT^{min}_{ct,l}$, e.g.:

$$ct_{lt} \neq ct_{l,t+1} \Rightarrow ct_{l,t+i} = ct_{l,t+i+1} \quad \forall\, i = 1,\ldots,\min\{CT^{min}_{ct_{l,t+1},l}, T-t-1\} \tag{1}$$

A change of the shift model e.g. from a 2-shift- to a 3-shift-model resizes the production capacity significantly. Therefore, with respect to the demand for products it is possible to adjust the capacity by cancelling single shifts, e.g. in a 3-shift-model the night shift on Friday. This is not considered as a change of the shift model.

After each change of the cycle time the stations on a line are consolidated and/or new stations are built. In this case, the workers often are assigned to different tasks and require a learning phase ([6]). They are less productive and have to be supported by an additional fraction of workers ct^X_{lt} which is modeled by a piecewise constant function as shown in Figure 1. We divide the learning phase into three phases: In the first phase after a change of the cycle time in period t^* with a length of $CT^{LT1}_{ct,l}$ periods the value ct^X_{lt} is set to $CT^{X1}_{ct,l}$ and is greater than $CT^{X2}_{ct,l}$ in the second phase with length $CT^{LT2}_{ct,l}$. In the third phase no extra employees are needed any longer until the next change of the cycle time occurs.

2.2 Staff

The demand for employees at line l in period t, which is solely calculated from the selected shift model and cycle time of the line, can be met by permanent and temporary staff ps_{lt} and ts_{lt}, respectively. Typically, permanent employees can only be dismissed at a few points of time (e.g. at the end of a quarter) while contracts with temporary ones can expire at the end of any period. Permanent employees are trained better than their temporary colleagues and are evidently necessary

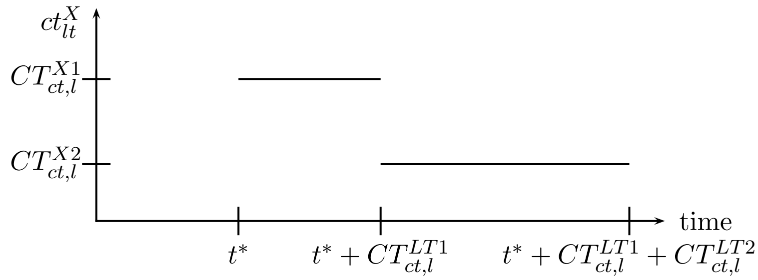

Figure 1: Extra amount of employees after changing the cycle time

for a continuously high quality. Thus, the fraction of temporary employees in the overall staff os_{lt} must not exceed a certain level TS^{max}:

$$ts_{lt} \quad \leq \quad os_{lt} \cdot TS^{max} \quad \forall\ l = 0, \ldots, L-1 \quad \forall\ t = 0, \ldots, T-1 \tag{2}$$

Especially regarding the staff the production lines cannot be decoupled from each other. For instance, the sum of all employees hired in one period t must not exceed the training capacities of the shop or due to their broad training permanent workers can be displaced them from one line to another (ds_{lkt}) to meet staff demand peaks.

2.3 Working Hours Summary

The working hours summary of a line where the workers collect their over- and undertime of each period is an important instrument to introduce flexible working hours (e.g. [3]). The over- and undertime is the difference between the attendance time WT of the selected shift model and the contracted working hours of the labor agreement. Due to the mid- to long-term planning horizon, we do not hold the summary for every single worker but for every single line as the average of all workers currently employed there, where the value whs_{lt} is bounded by in-company-agreements. Moreover, the summary has to take the value of 0 during a certain space of time (e.g. one year) or at a certain point of time (e.g. every last period of a year).

2.4 Production Program

Every period t the final assembly has to produce a certain amount D_{mt} of model m allowing a given relative deviation D^V, where the decision how many products of model m are produced on line l in period t is denoted by the variable p_{mlt}:

$$D_{mt} \cdot (1 - D^V) \leq \sum_{l=0}^{L-1} p_{mlt} \leq D_{mt} \cdot (1 + D^V) \quad \forall\ m = 0, \ldots, M-1 \quad \forall\ t = 0, \ldots, T-1 \tag{3}$$

During a certain range of time (e.g. one year) the overall demand of every model must be satisfied.

Since the production cost for a model m on a line l only depends on the selection of the cycle time and not on the load of a production line, some simplifying implications can be inferred. First, the lines should use the available capacity as long as there is enough demand for products. Second, if the capacity exceeds the demand one or more shifts have to be canceled so that the resulting capacity can be used to the maximum. Furthermore, if two products are produced on one line, for every cycle time—with respect to in-company-agreements—a preferred mix ratio of the products is given so that all stations of the line are nearly equally loaded (e.g. [4]). From this ratio the distributed production program can deviate only in a very tight range so we can determine the main part of the program's distribution.

2.5 Objective Function

The objective is to minimize a cost function which is given as the discounted sum of the staff, production, and changing costs over all periods. The staff costs C_t^S are composed of the basic wages, shift premiums, and costs for organizational modifications deriving from their hiring, dismissal and displacement to other lines. Dismissed employees get a compensation and possibly a payoff for their overtime hours on their current working hours summary. The production costs C_t^P are the sum of the variable costs for producing model m on line l using cycle time ct. The changing costs are invoked by changing the shift model (C_{lt}^{sm}) or the cycle time (C_{lt}^{ct}) on a line.

3 A Dynamic Programming Solution Approach

3.1 Basic Approach

We can model the considered problem as a MIP with a quadratic objective function and some restrictions neither linear nor convex. Particularly, due to the calculation of the staff demand (see Section 2.2) and the working hours summary, we are confronted with a non-convex MINLP. For this reason we selected a Dynamic Programming approach (see e.g. [1], [2]) to solve the problem. The main elements of any Dynamic Programming approach are states, decisions, cost functions to evaluate a state, and transformation functions. These elements are introduced in the following.

A state z_t associated with period t is characterized by the over- or underproduction Δp_{mt} of the different models m accumulated over all previous periods as well as a number of line-specific state variables. These are the actual shift model sm_{lt} and cycle time ct_{lt}, the number of periods gone by since their last change (s_{lt}^{sm} and s_{lt}^{ct}, respectively), the number of permanent and temporary employees ps_{lt} and ts_{lt}, as well as their working hours summary whs_{lt}.

A decision x_t contains the choice of the shift model $\widehat{sm}_{lt}$ and the cycle time $\widehat{ct}_{lt}$, the number of hirings, dismissals and displacements of permanent and temporary employees Δps_{lt}, Δts_{lt}, and ds_{lkt}, and the decision about the number of products manufactured on a certain line p_{mlt}.

Given some state z_t and an associated decision $x_t(z_t)$, the transformation function $t_t(x_t(z_t), z_t)$ generates a new state z_{t+1} in the next period. The chosen shift model and cycle time is adopted by $sm_{l,t+1} = \widehat{sm}_{lt}$ and $ct_{l,t+1} = \widehat{ct}_{lt}$, respectively. The number of employees is modified as

$$ps_{l,t+1} = ps_{lt} + \Delta ps_{lt} + \sum_{k=0}^{L-1} (ds_{klt} - ds_{lkt}) \quad \forall\, l = 0, \dots, L-1 \quad \forall\, t = 0, \dots, T-1, \tag{4}$$

$$ts_{l,t+1} = ts_{lt} + \Delta ts_{lt} \quad \forall\, l = 0, \dots, L-1 \quad \forall\, t = 0, \dots, T-1, \tag{5}$$

and the accumulated over- or underproduction is modified according to the production and the demand for each product

$$\Delta p_{m,t+1} = \Delta p_{mt} + \sum_{l=0}^{L-1} p_{mlt} - D_{mt} \quad \forall\, m = 0, \dots, M-1 \quad \forall\, t = 0, \dots, T-1. \tag{6}$$

Finally, the working hours summary is calculated and (in analogy for $s_{l,t+1}^{ct}$) $s_{l,t+1}^{sm}$ is set to $s_{lt}^{sm} + 1$ if no change of the shift model took place, or set to 0 otherwise.

The costs for changing the cycle time C_{lt}^{ct} is a linear function in ps_{lt} and ts_{lt} for all $l = 0, \dots, L-1$. The binary variables sm_{lt}^1 and ct_{lt}^1 indicate whether the correspondent changing costs have to be considered and the Bellman equation for the forward recursion can be written as

$$F_{t+1}^*(z_{t+1}) \quad = \quad \min \left(\left(C_t^P + C_t^S + \sum_{l=0}^{L-1} \left(ct_{lt}^1 \cdot C_{lt}^{ct} + sm_{lt}^1 \cdot C_{lt}^{sm} \right) \right) \cdot e^{-\alpha t} + F_t^*(z_t) \right). \tag{7}$$

Decisions as well as states can be infeasible and will then be eliminated. A decision is infeasible if e.g. the sum of hirings on all lines exceeds the training capacities of the shop. A state is infeasible if for instance the working hours summary is outside its admissible bounds. Infeasible decisions and those which lead to infeasible states are not generated to save computing time and memory.

The cost function serves to evaluate a state. If—disregarding costs—two states are identical then the more expensive one and the decision leading to it are eliminated. This way every state has a unique predecessor. Nevertheless, every state can have several successors that can be reached by different decisions. Starting at the least cost state z_T at the end of our planning horizon, we determine the optimal policy by backtracking through the predecessor states.

Due to the mid- to long-term planning horizon and the huge size of the feasible decision space it is reasonable and necessary to merge similar decisions to one. To give an example, we consider only decisions that—ceteris paribus—differ in the number of hired flexible workers in discrete steps of 20. Using this idea we can assure the operative usability of the developed software tool for the case of one single production line. In the case of parallel lines this simplification does not suffice to control the increasing complexity. For this reason, we adopt an idea by Schneeweiß ([5]) and separate the approach into two levels.

3.2 Separation into Top and Bottom Level

Both the distribution of the production program between the lines as well as the staff demand on the lines are dependent on the chosen shift models and cycle times. Both are independent from each other, so we can separate our approach from Section 3.1 into two levels. On the top level we generate, emanating from a feasible state in some period t, all feasible combinations of shift models and cycle times over all lines. We duplicate each of these combinations into as many decisions as there are optimal solutions with respect to the distribution of the production program between the lines (see Section 3.3). We complete each of these decisions by a greedy heuristic sizing the hiring, dismissal and displacement of staff (see Section 3.4). The quality of the heuristic for the decision concerning the staff is of utmost importance and has to take future periods into account. At this point, the decision on the top level is complete and the Dynamic Programming approach can be executed for the entire planning horizon. By this approach the number of decisions (and therefore the number of resulting states in the next period) is significantly reduced and our prototypical implementation shows the usability even for the case of parallel lines. We obtain an optimal policy over all periods keeping in mind that we use a heuristic for the staff decision.

In the second step, on the bottom level, for every period the decisions are partly preset, namely by adopting the values of $\widehat{sm}_{lt}$, $\widehat{ct}_{lt}$ and p_{mlt} from the optimal policy of the top level. We duplicate these decisions into as many as we can generate by enumeration of all meaningful decisions concerning the staff. There we rule out that e.g. permanent staff is displaced from line 1 to line 2 and simultaneously permanent staff is displaced from line 2 to line 3. Excluding such a decision is justified by the assumption that the costs for displacing staff only occurs on the line receiving the employees so that the triangle inequality always holds. Analogously, dismissing of permanent employees and simultaneous reception of permanent ones by displacements or hiring is an infeasible decision. Altogether, we complete the decisions adopted from the top level by those on the bottom level, start over the Dynamic Programming algorithm and determine an optimal policy.

3.3 Distribution of the Workload between the Production Lines

The distribution of the production program is restricted by two essentials: Firstly, by the choice of the shift model and the cycle time on every line the capacity is given, and secondly, the demand of each product and each period is given with respect to a certain relative deviation. This means that there is a minimal demand that has to be met and an additional, optional demand that may be met up to the maximum demand.

Before starting with the distribution of the production program, we test whether the maximal capacities for every single product exceed the maximal demands and whether the total capacity exceeds the total maximal demand. In this case, we reduce the capacity by canceling single shifts. Hence, the capacities never exceed the demands. If the capacities do not suffice for the minimal demand the decision is infeasible. Now the capacity of every line producing only one model is assigned completely to the production program of the related model. For every other line we assign, as indicated at the end of Section 2.4, the main part of the capacity to the production program of the two related models. The distribution of the production program still pending is now obtained by modeling the problem as a classical transportation problem where each line asks for and each model supplies production program. The transportation unit costs are derived from the variable production costs. In particular, we separate each model, e.g. denoted with A, into two suppliers, namely A_{min} offering the remaining minimal demand to be distributed after the considerations made above, and A_{add} offering the additional demand of A. Since the additional demand has not to be produced entirely, we introduce a dummy line to balance the transportation problem. Similarly to the splitting of the products, we separate every production line into as many destinations as products can be produced on that line. In doing so, we model the preferred mix ratio of the selected cycle time on this line and its allowed deviation as mentioned before.

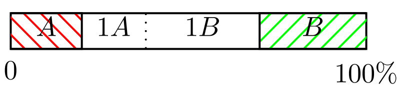

Figure 2: Distribution of the capacity between model A and B

Figure 2 illustrates the capacity of line 1 on which the models A and B can be produced. The dotted line denotes the preferred mix ratio, the left and right hatched rectangles represent the assignment of A's and B's production program to 1's capacity. A solution of the transportation problem provides the information whether the remaining, unmarked area (the yet unused capacity) should be filled with production workload of model A or B. The labels (see Figure 2) indicate the preferred model to be produced there to achieve the preferred mix ratio. This preferences are modeled by the transportation unit costs, i.e. the higher the cost the less is the preference to produce the correspondent model on the respective lines. The unit cost for the transportation of a model to the destinations built out of the line is set to the related variable production cost. The maximum of all these values is added to the unit costs for the transportation to the less preferred destinations, e.g. model A to destination $1B$. If a line cannot produce a model the transportation unit cost is set to a sufficiently large value $Big - M$. Setting the transportation unit costs, we do not differentiate between the two suppliers A_{min} and A_{add}. The costs for transporting from A_{min} to the dummy line are set to $Big - M$, from A_{min} it is set to a greater value than all transportation unit cost to any real line, but smaller than $Big - M$. In doing so, we interdict the production of A_{min}'s supply on the dummy line. Additionally, we can introduce a preference order between all products regarding the production of the additional demand by differentiating between the respective transportation unit costs to the dummy line. The preference can be e.g. induced by the accumulated over- and underproduction Δp_{mt}.

3.4 Heuristic Staff Planning

The heuristic solution concerning the staff planning on the top level is essentially an anticipation of the decision to be made on the bottom level. This anticipation includes a forecast of the future demand as well as the restrictions concerning the working hours summary. At first we introduce a preprocessing step determining the production capacity and the minimum number of required employees for every combination of shift model and cycle time for all lines. Incorporating the maximum number of employees to hire in one period and the current staff size, we can determine the number of periods needed to hire enough staff to meet the staff demand in a future period. We observe whether we have to hire additional employees already in the current period which can be seen schematically in Figure 3. The number of periods we look ahead is a parameter with a vital influence on the quality of our anticipation. Next, we have to test whether the working hours

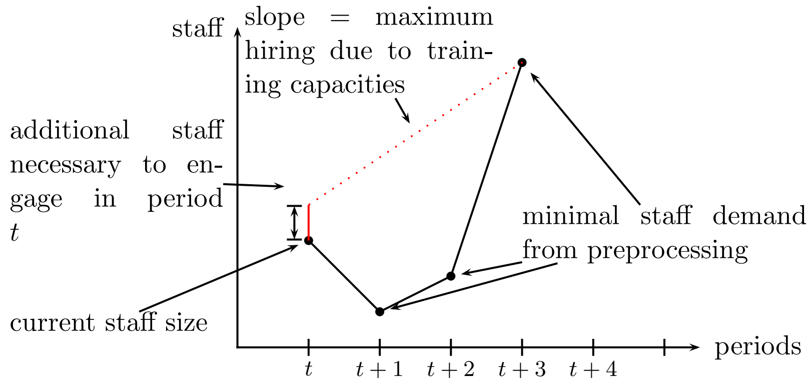

Figure 3: Anticipation of hiring staff for future demand

summary is feasible or not and whether actions are to be taken to improve a critical summary. The summary of a line is the average over all workers employed there. New employees added to the line have a summary of 0. So the average summary is moved towards 0 by new employees, independent of a positive or negative actual summary of the line.

The infeasibility of a working hours summary can be invoked by too many hours of over- or undertime. In the first case additional employees can be hired to bring back the summary into the feasible region. In the second case it is also possible to achieve feasibility by hiring employees in the short term. However, these additional workers cause an oversupply of labor in the following periods so that the summary will become infeasible again which will consequently lead to a vicious circle. For this reason, a state with an infeasible working hours summary caused by undertime is treated as infeasible.

The summary has to reach a total of 0 regularly, so in case of a feasible summary greater than 0 we have to hire additional employees in the following periods. The earlier an employee is hired, the more he can help to reduce the average summary. Hence, we observe that we should generally hire more employees in the current period. Thus, we avoid hiring noticeable more employees in the last periods before the summary has to be 0. If a feasible summary has a negative value, we encounter an analogous problem as described above. Again in a short-term view we can move the summary towards 0 by hiring additional employees but for the long-run the only possibility to reach the value of 0 is to choose a shift model that implies a greater attendance time than the contracted hours. This is outside the scope of this anticipation, so we punish a negative summary with artificial costs in our objective function, e.g. with the value of the accumulated undertime. At this point we know how many permanent and temporary employees we have to dismiss or hire on every line for the production in the current period as well as the surcharge for the future production program and a feasible working hours summary.

Finally, we have to examine whether displacements of permanent staff can avoid hirings. Therefore, we adopt the savings method from the vehicle routing problem and build up three savings lists. In the first list we calculate the savings for displacement instead of dismissing permanent workers on one line and hiring on another. In the second list the second line has a demand for flexible workers and in the third list the second line does not have a demand for employees but we could dismiss flexible workers on this line and receive permanent ones by displacements to this line. The second and third list are sorted in analogy to the first one. Beginning with the first list and according to the sorting sequence, the displacements are determined where negative savings are also taken into account because it is generally a reasonable assumption to prefer a small staff size.

4 Conclusion

In this paper we presented a Dynamic Programming approach to solve the problem of integrated mid-term production and staff planning in a shop of an automotive plant. We focused on achieving tractability even for the case of parallel production lines that are not independent of each other and disaggregated the solution approach into two levels. On the top level we modeled the problem of distributing of the production workload between the production lines as a transportation problem where the planning of the staff is solved approximately. On the bottom level we fixed all decisions beside the staff planning and determined the optimal staff planning. Areas of future research and development are: First, to integrate the techniques described in this paper into our partner's software tool which is already able to manage subsequent shops each with one production line and intermediate buffers in between them. Second, to improve i.e. to speed up the procedure of the bottom level, e.g. by using a local search approach instead of the Dynamic Programming.

References

[1] R.E. Bellman (1957), *Dynamic Programming*, Princeton University Press, New Jersey.

[2] D.P. Bertsekas (2000), *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont.

[3] L. Delsen, D. Bosworth, H. Groß, and R.M. de Bustillo y Llorente (eds.) (2007), Operating Hours and Working Times, Springer, Berlin.

[4] N. Leopold (1997), *Ein Planungsverfahren zur Kapazitätsbestimmung für Modell-Mix-Montagelinien am Beispiel einer Automobil-Endmontage*, Springer, Berlin.

[5] C. Schneeweiß (1994), Elemente einer Theorie hierarchischer Planung, *OR Spektrum* **16**, 161 – 168.

[6] T.L. Smunt (1986), Incorporate Learning Curve Analysis into medium-term Capacity Planning Prodedures: A Simulation Experiment, *Management Science* **32**(9), 1164 – 1176.