A review of methods to estimate haul fleet production

Douglas D. Gransberg, Texas Tech University
A Review of Methods to Estimate Haul Fleet Production

Dr. Douglas D. Gransberg, PE CCE

The past 3 decades have been an era of accelerating technology. Many advancements have been made in the development of larger, faster, and more productive construction machinery. Increased machine productivity has resulted in an increase in overall project size. These two factors have combined to produce a capital-intensive and risky environment in which construction contractors must operate. As a result, members of the construction industry have been forced to search for methods to reduce the high level of risk [12]. Historically, the lowest cost method for reducing risk has been to provide detailed estimating and planning before submitting a bid, and solid management throughout the course of a project [8]. Estimating and planning involve the judicious selection of equipment, the careful scheduling of time and resources, and the accurate determination of system production and costs. Management involves putting the plan into action. The key management ingredient is having predetermined standards to measure actual system outputs, upon which future decisions can be based.

Even a seemingly straightforward operation such as earthmoving is a highly dynamic system [2]. A hauling operation contains several components that interact in a very complex manner. Analytical methods, based on engineering fundamentals, have been developed to solve the problem of bringing these components together in a logical manner. These methods model hauling systems mathematically. Their solutions are numerical results that may be used in the decision-making process of estimating, planning, and managing an earthmoving project [9].

Early methods made the somewhat naive assumption that optimizing production based on the physical constraints of the environment would in turn minimize the overall production cost. Therefore, no effort was made to include cost or profit variables in those mathematical models. The models developed by Gates and Scarpa [3] were the first to recognize the importance of the cost function in overall system optimization. Many methods currently in use do not adequately model physical and economic conditions. They rely on the judgment and experience of the user, which may be very good or very bad, with corresponding outputs from the models [11].

The purpose of this article is to review five methods used to estimate haul fleet production based on optimal analysis and then compare the results predicted by those methods to actual field observations. The results of the comparison are then used to recommend various methods for estimating haul fleet production. With the exception of the Phelps method [10], the details of each are contained in the literature, so no detailed mathematical explanation of them is presented here. The reader is invited to look to the cited references if more amplification is required.

METHODS OF ESTIMATING PRODUCTION

Five analytical methods for estimating the production of hauling operations are found in the literature and are used to varying degrees in the construction industry. Each method brings its author's particular technical expertise to the same problem. The authors of the methods reviewed in this paper are:

- R.L. Peurifoy, as interpreted by Atcheson;
- F.H. Griffis;
- M. Gates and A. Scarpa;
- R.E. Phelps; and

The first four methods are mathematical models describing the hauling process. Inputs vary between models, but the outputs concern system production and the optimum size or number of haul units. These methods establish an analytical system to optimize haul fleet production by optimizing the composition of the haul fleet itself. The last method uses statistics derived from time and motion studies to determine the optimum productivity of an existing fleet. Each of the methods is reviewed below.

Peurifoy/Atcheson

The first haul fleet optimization method to gain wide acceptance was developed by R.L. Peurifoy in his seminal work, Construction Planning, Equipment, and Methods [7]. D.B. Atcheson took Peurifoy's work and built an algorithm that can be directly applied to most earthmoving systems to estimate system production [2]. Basically, Peurifoy's method involves determining all of the physical constraints involved in a hauling system and evaluating them to determine the system's ultimate performance. The primary input parameters are haul road rolling resistance and grades, haul unit (dump truck, scraper, etc.) loaded and empty weights, horsepower, and transmission characteristics, and loading, travel, and delay times for a typical haul unit.

Peurifoy's techniques allow an engineer to relate the hauling process to engineering fundamentals and make estimates of system productivity based on those fundamentals. The estimates are based on instantaneous production (the maximum theoretical production achievable at any given instant) rather than sustained production (the average realistic production achievable throughout the course of the project). This is the model's primary weakness. Additionally, it does not include cost factors that must be considered to fully optimize any production system.
Griffis

F.H. Griffis expands on Peurifoy's method by applying queuing theory for a finite population to determine optimum fleet size [5]. This method strives to determine the optimum tradeoff between idle time at the loading facility and the idle time of the hauling units. Queuing theory was first applied to this problem in 1964 [6] and involved extensive, complex calculations. Griffis was able to replace much of the calculation with a ratio of probabilities.

The Griffis model uses the same set of variables as Peurifoy to determine haul unit cycle time and loader productivity. Additionally, it uses the haul unit arrival rate, loader service rate, and a Poisson distribution to determine haul unit waiting time and loader idle time. It also uses the hourly cost of both the haul units and the loader with operators. The weakness of the Griffis model is that it does not include all of the costs involved in an earthmoving operation. No costs of supervision, overhead, mobilization, or demobilization are included. However, it is set up so these costs could be included for a truer picture of the relationship between different fleet sizes. Many of these costs are proportional, to some extent, to the number of units in the fleet. Additionally, costs such as supervision may take the form of a step function (1 to 4 haul units may require 1 supervisor while 5 to 8 units may require 2 supervisors). This could significantly change the outcome of the optimum fleet size.

Gates and Scarpa

This model concerns the determination of the optimum size of the hauling unit [3]. It was the first one to deal with the relationship between the size of the haul unit and the production characteristics of the loading facility. Additionally, this model recognizes the contribution of the relative cost function in the determination of an optimal condition and also relies on Peurifoy's fundamentals to obtain the optimum number of units and loader productivity. As a result, this model uses the same variables as Peurifoy, plus a few to model the cost relationships. These variables are cycle time components, loader productivity, the total hourly cost for a driver, and the ratio between the hourly ownership and operation cost of the haul unit and its capacity in cubic yards. This model is founded in theory and is sound to the point where instantaneous productivity is used to determine the optimum size and number of haul units in the system. It is obvious that earthmoving systems cannot always achieve the minimum instantaneous productions determined by physical constraints. None of the models reviewed thus far have included a factor to model sustained rates of production. Drivers and operators are human and subject to fatigue, and need to leave their machines for various reasons throughout the course of a normal work day. Because a model's purpose is to represent a real situation mathematically, the human factor must be included to allow an engineer to make realistic estimates and decisions.

The Gates and Scarpa model can be used to evaluate actual rates of production against a calculated optimum standard. Sustained production and cycle time values, rather than instantaneous values, can be used without any modification to the basic equation. This enhances the value of the model as a decision-making tool and furthers understanding of the dynamic interactions within the subject system. Because the model was designed for use as an equipment selection method, it is not evaluated here as an estimating tool. However, it was used to evaluate the expected production rate based on actual data for each of the case study projects.

Phelps

The model proposed by Phelps [10] is the first approach to use sustained cycle times for estimating productivity and determining the proper mix of equipment to obtain optimal conditions. As with the others, this model is based on the fundamentals established by Peurifoy, and therefore uses the same basic set of variables and virtually the same set of equations, in modified form. The striking difference between the Phelps model and other methods is that it does not fix physical constraints at a specific value but rather evaluates them over their natural range to determine the best possible situation. A good example of this approach is the treatment of rolling resistance. If a particular haul road is made of compacted earth and has a rolling resistance of 160 lb per ton (8 percent), the maximum velocity achievable on that road may be 20 mph (32 km per hour). By paving the road, the rolling resistance may be lowered to 50 lb per ton (3 percent), with a resultant possible increase in velocity to 60 mph (96 km per hour). The increase in productivity may justify the investment in temporary pavement. Additionally, Phelps bases his estimates on sustained productions by including estimates of delays due to human failurals.

This model requires a manager to examine each project in depth. As a direct result, the outputs are relatively consistent with the actual performance of the system. The basic drawback is the amount of detail required to achieve consistent results. The calculations tend to become very extensive due to the modeling of the dynamic interactions between components in the system. Additionally, the model does not seek to establish the relationship between the capacities of the loading facility and the haul units for top-loaded earthmoving equipment. To overcome this, one must carefully select haul units of a capacity that is a multiple of the capacity of the loading facility.

For example, if the loading equipment is a bucket loader with a heaped capacity of 1.5 yd³ (1.14 m³), dump trucks with a heaped cubic-yard capacity of 3, 4.5, 6, 7.5, etc., would permit the loader to load a complete bucket every loader cycle. Thus, as the haul units can only haul as much earth as the loader can load, making an equipment selection that permits the loading facility to operate at its maximum production rate permits the entire system to operate at its maximum sustained production without a constraint due to mismatched capacities. A good method for analyzing this relationship for more complex loading facilities is called the load growth curve method [4]. If this initial bit of equipment selection analysis is performed, the Phelps method can be used with good results.

It should be noted that if the project is constrained to a fixed set of equipment and no latitude is available to correct capacity mismatches, the estimator should calculate the sustained production rate of the loading facility as a function of the actual capacity of the haul units, or vice versa.
versa, depending on which capacity controls the actual job site production. For instance, if the 1.5-yd³ loader must load 6.5-yd³ trucks, it will actually load four buckets containing a total of 6 yd³ per truck. Thus, the estimator should use 6 yd³ per cycle to estimate system production. If the same loader is loading 7-yd³ trucks, the loader operator will probably attempt to load a partial fifth bucket to fill each truck. Therefore, an estimator should figure the loading facility sustained production rate by dividing 7 yd³ by the time it takes to load five buckets.

Adrian and Boyer

This model is called the method productivity delay model [1]. It makes no effort to optimize any part of a construction process. Its use is also not constrained to a particular method, so it can be used to estimate the productivity of almost any definable cyclic construction process. It is included here as an alternative to allow an engineer to evaluate productivity and make decisions. Additionally, it provides a good cross-check for comparison of the predicted results of the four previous models and the data obtained from field observations.

The model is based on input data derived from time and motion studies undertaken after a particular process has started. The basic variables that are timed are the overall cycle time and the durations of the various delays that occur during the cycle. The results of the time study are then analyzed statistically to determine the relative probabilities of the various types of delay. These are then applied to the ideal productivity to calculate or estimate the overall productivity of the system. Decisions then can be made to eliminate controllable delays and improve system productivity.

Obviously, the application of this model to estimating future occurrences depends on having data from previous projects. This begs the assumption that past projects are exactly the same as upcoming projects. The probability of two projects being exactly the same is very low, and this approach often leads to gross misjudgments. The value of the model is its ability to evaluate a project after it starts and "fine-tune" the equipment mix during the project. Additionally, it allows an engineer to directly measure the delays that are likely to occur during the operation. This increases understanding of the process and enhances a project engineer's ability to more accurately estimate the magnitude of the human factor on future projects.

FIELD OBSERVATIONS

The purpose of taking field observations was twofold. First, the described estimating methods needed to be evaluated to provide a comparison based on common data. Second, the strengths and weaknesses of each model became clear when predicted results were compared with actual field data. The hauling operations of three different construction projects located in Billings, MT, were observed. The projects were selected based on their fundamental differences in haul length, haul route complexity, haul unit size and type, dumping method, and number of units in operation. Unfortunately, none of the projects involved a totally simple haul route (a route without stops). However, a method that can closely model a complex situation can certainly model a simple situation. All of the projects involved the use of top-loaded hauling units, and in two cases, the dump end (the ability of production machinery to handle material brought to the project) of the haul cycle controlled the efficiency of the operation. Each project used the same size and model of haul unit within each project's haul fleet; the performance characteristics of the units within the fleet were assumed to be the same [10].

The following system variables were measured or obtained:

- haul length, rolling resistance (measured using the Phelps [10] stopwatch technique);
- grade;
- haul unit weights;
- haul unit performance characteristics; and
- cycle time components.

The hauling operations were observed from a location where the observer was both out of the way and theoretically unobtrusive so as not to prejudice the samples. However, the workers did find out that they were being observed. The productions observed during the sampling period were checked with previously recorded production rates. It appeared that the presence of an observer tended to raise system productivity slightly.

Descriptions of the Projects

Project A was the most complex. It involved hauling material from a stockpile 4.5 miles (7.2 km) to a subdivision street project. Five on-highway end dumps were used, and they were top-loaded by a front-end loader. Twelve turning movements accompanied by 12 stops were encountered along the haul route. Additionally, speed was restricted by a speed limit. The ability of a motor grader on the street project to spread the dumped material controlled the number of hauling units that could be used.

Project B consisted of hauling material from a borrow pit to a drive-over feeder. The feeder was connected to an aggregate processing plant by a belt conveyor. The haul was 3.5 miles (5.6 km) long. Tractor-trailer bottom dumps were used, and they were top-loaded by a front loader. In this case, the ability of the drive-over feeder to empty on to the conveyor controlled the maximum number of hauling units that could be used on the project. There were two stops and four turns each way.

Project C consisted of hauling material .75 miles (1.2 km) to a parking lot. This project used the same haul units as project A. Two turning movements and two stops were encountered on the haul. Heavy city traffic influenced the efficiency of the haul.

General Comments on Each Project

After the first day of observation on each project, the hauling system was evaluated to determine whether a mismatch between equipment capacities and numbers was constraining it from operating at maximum productivity. Projects A and C both were found to be operating at less than maximum. In both cases, it was recommended to the project manager that more haul units be added (one additional for A and two for C). On project A, the suggestion was implemented, and the total number of haul units was raised.
from five to six. It was then observed that the mean cycle time changed very little, and the daily production went up by the appropriate amount. Additionally, because this project's productivity was constrained by the motor grader's ability to distribute the material once it had been spotted and dumped, and the fact that only one grader was available, the amount of grader idle time was observed to drop to nearly zero. Therefore, it can be concluded that the system was operating at maximum productivity.

In the case of project C, there were no additional haul units available. As a result, the analysis described above could not be made. Project B was operating just below maximum productivity. The addition of another hauling unit of the same size would have exceeded the trap's ability to empty itself. Therefore, the Gates and Scarpa method [3] for determining the optimum haul unit size was used to evaluate the equipment size needed to maximize productivity. It was found that four 30-ton (27000 kg) haul units would have optimized the system with respect to haul unit size. However, legal haul restrictions prevented using this size of equipment. It is concluded that this system was operating at the maximum possible productivity within legal haul weight constraints.

Cycle time and daily production were of interest. Observed cycle times were compiled and analyzed statistically to determine the mean observed cycle time and the standard deviation of observed cycle times for each project. Since there were only three observations of daily production, the arithmetic mean was used to calculate a number for comparison. Table 1 shows the results.

Table 2 is a comparison of the values estimated by current methods and the observed values of cycle time and productivity. The difference from the mean observed cycle time and the mean daily production for each method is shown in table 3. Due to the small sample size of only three projects, a definitive analysis of these results is difficult. However, a certain amount of inference can be made about each method's performance in regard to the other methods. First, the Atcheson and Phelps methods seem to contain the greatest amount of error because the values they predict fall outside of the standard deviation expressed as a percentage of the mean on two out of three projects, for both cycle time and daily production. Adrian's method uses actual project information and therefore was very consistent when predicting cycle times, but it tended to estimate production on the high side. This is probably due to the inherent assumption that all trucks will haul at all times. When one breaks down or when a driver takes an extended break, an entire cycle or more of production can be lost. Since this method does not contain a parameter to account for lost cycles, it would be expected to overestimate the daily production. The Phelps method is the only one that predicted values for both cycle time and production that fell within the standard deviation for each project.
A number of conclusions can be drawn from these analyses. First, Atcheson's method is generally conservative, which is not necessarily desirable in this application. A conservative estimate of cycle time causes the optimum number of haul units to go up. If the actual cycle time is less than the estimate, more haul units than necessary will be assigned to the project. This causes a waste of time while waiting in the loading queue and increases costs. Also, Atcheson's estimates of production are inconsistent. This is due to the lack of a human factor in the model. Next, Griffis' method is generally low on determining the number of haul units required for optimum production. This is because of the nature of queuing theory itself. The Griffis model strives to strike a balance between the cost of trucks waiting to be serviced and the cost of the service facility waiting for trucks.

The Phelps method seems to be the most consistent. It is within less than one standard deviation high in all cases for cycle time. The productivity estimated is also very close to observed, particularly in the two projects (B and C) that did not have maintenance problems (this caused the loss of full production cycles). The disparity between the optimum number of haul units estimated for project C and the actual number used is due to the fact that this project was not operating at maximum productivity. Adrian's method also appears to be consistent. However, it uses information collected from the project itself, so it should be very accurate. As stated before, Adrian's method is best used as a means to evaluate a project once it is under way.

It should be noted that these conclusions cannot be considered statistically significant because of the small sample size. However, that notwithstanding, the study does provide an interesting comparison of a variety of methods that all seek to solve the same problem. Estimators and project managers alike should interpret the results as a message that calculating haul fleet production is a complex, highly dynamic operation and should receive due care and consideration if these estimates are to accurately model job site productivity.

ACKNOWLEDGMENTS

I would like to thank R.E. "Bob" Phelps and H.D. "Hal" Pritchett of Oregon State University for their support and mentorship. Their insights and experience helped bridge the gap between theory and practice. I also would like to thank Les Keebler, PE, of Midland Materials Company, who made it possible to take the field observations and adjust the equipment mixes.

REFERENCES