Lean Railroading for Improving Railroad Classification Terminal Performance

Bottleneck Management Methods

Jeremiah R. Dirnberger and Christopher P. L. Barkan

Although much attention has been focused on the growth of intermodal traffic over the past decade, manifest freight (or carload) traffic is a major revenue generator for railroads. The high potential profitability of carload traffic suggests that railroads should try to grow this segment of traffic further, especially in an era of limited railway capacity. To do this, they must meet the increasing logistical needs of their customers by providing more reliable service. The classification terminal is a key determinant in service reliability of manifest freight. Terminal performance also affects network efficiency. Regression analysis showed that, as average dwell time increased, average manifest train speed decreased. Inadequate terminal capacity is viewed by many as a barrier to improved service reliability and network efficiency. Because terminals can be considered production systems, insight is gained by adapting tools that have led to significant performance improvement in manufacturing. A new approach is introduced: lean railroading. The most important manufacturing process analog to improving terminal capacity is the bottleneck. The train assembly (pull-down) process has been identified as the bottleneck in a majority of classification yards. A sensitivity analysis conducted on three bottleneck management alternatives suggests that pull-down capacity can be increased by as much as 26%, compared with the baseline case without large labor or capital expenses, through better management of the process and its interactions with the system. To maximize efficient use of rail yard infrastructure and resources, more emphasis should be placed on the quality of the classification process, rather than on quantity.

Manifest or carload traffic has traditionally been one of the major sources of revenue for Class I railroads. Although the industry has placed substantial focus on intermodal traffic because of its substantial growth (about 40%) in the 10-year interval of 1996 to 2005 (*1*), growth in carload traffic is also important, particularly in light of its potential profitability. Norfolk Southern (NS) recently reported that 43% of its rail revenue is derived from carload shipments, and 18 of its top 50 customers are carload shippers (T. Bragman, Norfolk Southern Perspective on the Future of Single-Car Railroad Shipments). Union Pacific (UP) reports that manifest business already accounts for 41% to 45% of revenue and is its fastest growing business line (*2*).

J. R. Dirnberger, Canadian Pacific Railway, Gulf Canada Square, 401 9th Avenue, Southwest, Calgary, Alberta T2P 4Z4, Canada. C. P. L. Barkan, Railroad Engineering Program, Department of Civil and Environmental Engineering, University of Illinois at Urbana–Champaign, 1201 Newmark Civil Engineering Laboratory, 205 N. Mathews Avenue, Urbana, IL 61801. Corresponding author: J. R. Dirnberger, jeremiah_dirnberger@cpr.ca.

Transportation Research Record: Journal of the Transportation Research Board, No. 1995, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 52–61.
DOI: 10.3141/1995-07

A key operational distinction between traffic that runs in unit or intermodal trains versus traffic in manifest trains is that the latter frequently visits several terminals between origin and destination. Service quality is thus strongly affected by terminal performance as well as over-the-road operations. To meet the logistical requirements of shippers, railroads need to ensure reliable train connections and adequate terminal capacity (*3*).

Martland et al. names the classification terminal as a key determinant in the service reliability of manifest freight (*3*). In addition, Murray states that "Cars spend most of their time in terminals, and that's where the service battle is won or lost for carload business" (*4*). Two major North American railroads have reported that 59% (*5*) and 64% (Canadian Pacific Railway, "Rendering Value from IOP Compliance, a How-To Guide: Part I—the Hidden Truths") of railcar transit time is spent in yards, and these figures are probably typical of all the Class I railroads today. "This suggests that the reliability of car movements can be improved by reducing the time spent in those activities or by making them more reliable" (*6*). The transition to scheduled operations by all of North America's Class I railroads has increased the interaction between yard performance and service reliability (*7, 8*) because "efficient high-throughput classification yards are vital to scheduled railroading" (*9*).

IMPACT OF CLASSIFICATION TERMINAL OPERATIONS ON NETWORK EFFICIENCY

In addition to improving service reliability, better performing terminals result in more efficient railroad networks. A common measurement of network efficiency is average train speed. Train speed measures line haul movement between terminals. The average is calculated by dividing train miles by total hours operated, excluding yard and local trains, passenger trains, maintenance of way trains, and terminal time (*10*). A higher systemwide average train speed indicates a more fluid network because trains take less time to travel the same distance. Therefore, fewer cars and locomotives are required to move traffic because the equipment is cycled faster. Network efficiency can be thought of as a cycle (Figure 1) with terminal dwell linked directly to car velocity and indirectly to average train speed through a series of events.

One estimate of the impact of improved terminal performance on carload velocity is provided by Logan (*5*). Every 15% reduction in systemwide average terminal dwell time results in an approximate increase in carload velocity of 2 mph. The exact relationship between terminal dwell and average train speed varies on different railroads because of complex interactions between the factors shown in Figure 1. However, an idea of the relationship can be obtained by analyzing performance measures for each railroad.

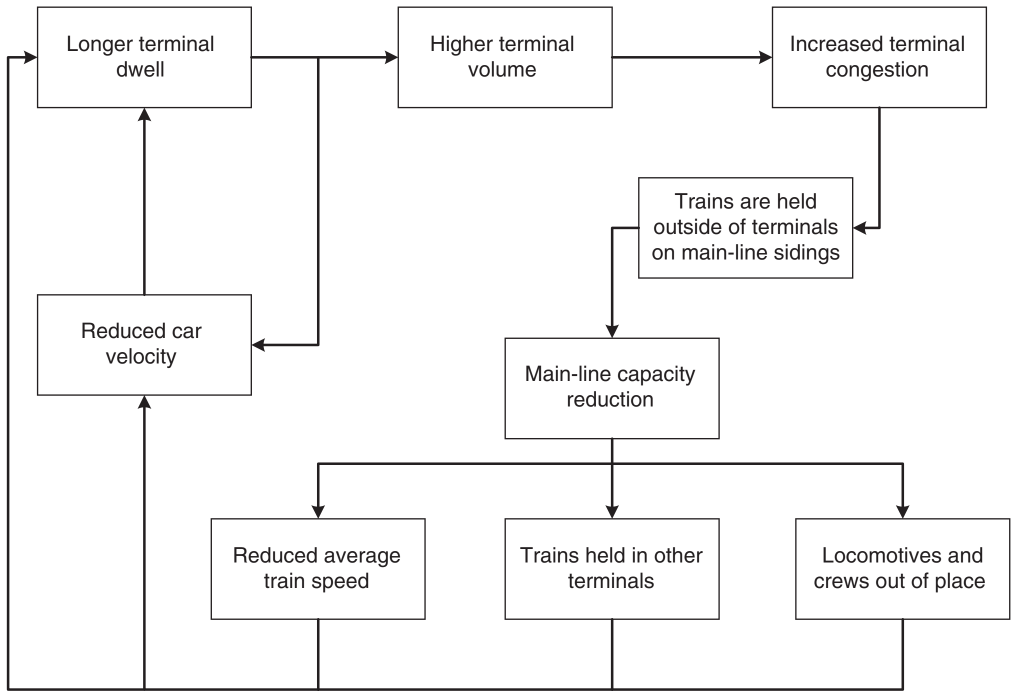

FIGURE 1 Railroad network efficiency cycle.

To understand better the relationship between average terminal dwell and average train speed, performance measurement data were analyzed for the six largest North American Class I railroads. Data for BNSF Railway (formerly the Burlington Northern Santa Fe Railway), CSX Corporation, Inc., NS, and UP were obtained from the American Association of Railroad's railroad performance measures website (*10*). Data for Canadian Pacific (CP) and Canadian National (CN) were obtained from their corporate websites (*11, 12*). For all of the railroads except CP, average weekly manifest (carload) train speed and corresponding average weekly terminal dwell for the entire railroad were obtained for the weeks ending May 27, 2005, through May, 26, 2006 (53 weeks). CP's website contained weekly averages for the weeks ending October 1, 2004, through May 26, 2006 (87 weeks). Data for Christmas week were excluded because the atypical operations characteristic of that week are distinctly different from the rest of the year. A simple linear regression model was applied to each railroad's data, and hypothesis tests at a 95% level of significance were performed. A scatter plot for the railroad with the most data, CP, is shown in Figure 2, and the scatter plots for the remaining railroads are similar (*13*).

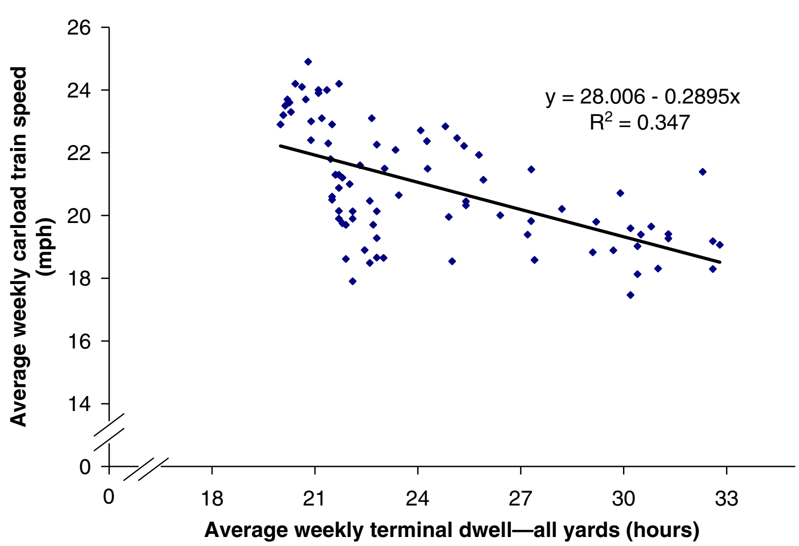

FIGURE 2 Scatter plot and regression results: CP, weeks ending October 1, 2004, to May 26, 2006.

TABLE 1 Regression and Hypothesis Test Results for Relationship Between Average Weekly Terminal Dwell and Average Weekly Carload Train Speed

Railroad	Regression Analysis Results	Railroad	Regression Analysis Results
BNSF	TS = 32.395 – 0.4943TD	CSX	TS = 24.613 – 0.2064TD
t-stat	11.56 –4.35	*t*-stat	16.96 –3.91
p	<0.001 <0.001	*p*	<0.001 <0.001
R^2	0.275	R^2	0.234
CN	TS = 28.290 – 0.285TD	NS	TS = 31.111 – 0.4579TD
t-stat	28.57 –3.61	*t*-stat	15.40 –5.18
p	<0.001 <0.001	*p*	<0.001 <0.001
R^2	0.207	R^2	0.349
CP	TS = 28.006 – 0.2895TD	UP	TS = 24.829 – 0.1857TD
t-stat	26.04 –6.64	*t*-stat	15.32 –4.61
p	<0.001 <0.001	*p*	<0.001 <0.001
R^2	0.347	R^2	0.2468

TD = terminal dwell; TS = terminal speed.

The potential network efficiency gain is seen in the regression results (Table 1). For each railroad, there is a statistically significant inverse relationship between carload train speed and terminal dwell ($p < 0.001$ for all six railroads). The R^2 values range from 20.7% to 34.7%, which is an indicator of the amount of the variability explained by the model. The remaining variation in average train speed is presumably explained by other factors such as locomotive availability, main-line train speeds, meets and passes, signal systems, weather, crew availability, line congestion, and so on. Considering all of these other factors affecting system performance, the percentage of variability accounted for by the single variable, dwell time, might seem surprising. However, the high percentage of time that freight cars spend in terminals and the 65% of root causes of delays related to terminal management point to terminal dwell time as a critical factor affecting railroad network efficiency (*14*).

INCREASING TERMINAL CAPACITY

Demand for rail transportation, measured in revenue ton miles, has continued the growth that began in the early 1990s (*15*). If demand continues to increase, as indicators suggest (*16*), average train speeds will trend downward unless more productivity can be extracted from network resources through infrastructure expansion or more efficient use, or both. As mentioned, one network resource having an increasingly important impact on network efficiency and service reliability is the classification terminal.

Within a classification terminal, connections are made by classifying cars from inbound trains into blocks that will be assembled into outbound trains. The objective is to sort cars and reliably connect them to the earliest possible candidate outbound train while minimizing cost (M. Barker, Network Solutions, CN, personal communication, Feb. 22, 2005). Kraft has extensively studied the connection reliability problem as it relates to dynamic car scheduling (*17*) and has developed a hump sequencing algorithm (*18*), a priority-based classification system (*7*), and a dynamic block to track assignment scheme with the goal of ensuring connections (*8*). Kraft raises the issue of inadequate terminal capacity as a barrier to improved service reliability (*18*). However, availability of capital and physical capability to expand some yards are constrained. Therefore, in addition to considering infrastructure expansion, railroads must also determine how to harness as much capacity from extant infrastructure as possible. This creates the need for new management and operational methods that will increase the capacity of existing facilities.

Investigating means of increasing the capacity of manufacturing facilities has been extensively studied, the management techniques that have been developed there can be adapted to analyze and improve railroad productivity as well. Lean railroading (*13*), an approach that adapts proven production management techniques to the railroad environment, can be used to guide improvement initiatives. In this paper, emphasis is placed on the bottleneck management component. The pull-down process is identified as the most common bottleneck in hump yards. The macroscopic evaluation method of Wong et al. (*19*) has been expanded by developing two additional equations to assess yard performance. These are then used to conduct sensitivity analyses on improvement alternatives using data collected at Bensenville Yard (CP) near Chicago.

LEAN RAILROADING

Because classification terminals can be considered production systems (*13*), their performance can be improved by adapting an integrated approach consisting of three proven production management techniques: lean, theory of constraints (TOC), and statistical process control (SPC or six sigma). Through the concept known as lean railroading (*13*), several railroads and railroad suppliers—including CP; UP; BNSF; NS; Belt Railway of Chicago, Illinois; and GE Yard Solutions—are actively applying all or parts of this approach to improving terminal performance. In addition, many of the precision railroading principles that CN has used to improve their operating performance can be considered lean.

The first step in any lean program is to define value for the ultimate customer and then work to increase value by eliminating waste in the system. Waste is defined as any step or process in a production system that, from the customer's standpoint, does not add value to the product (*20*). Waste can be classified into two types: direct waste and variability (*21*). Direct waste is most easily described as poor railroading practices such as unnecessary moves, mistakes that require an operation to be repeated, inadequate track maintenance, and unsafe operations, to name a few. Focusing on these practices is important, but the goal of eliminating direct waste is as old as the railroad itself (*22*).

Variability is a fundamentally different source of waste. Hopp and Spearman state, as a law of manufacturing, that, "Increasing variability always degrades the performance of a production system" (*23*). Railroad yards are no different: They are subject to both internal (outages, rework, sorting, etc.) and external (arrival times, weather, traffic volume, etc.) sources of variability. Another law of manufacturing from factory physics is that "variability in a production system will be buffered by some combination of inventory, capacity and time" (*23*). In a terminal, a capacity buffer takes the form of a process throughput greater than the process demand. A time buffer is the extra time built into each car's trip plan in order to ensure that the connection will be made and is seen in the terminal dwell.

Spearman states, "In many ways, the 'waste' discussed in Lean is the 'buffer' of Factory Physics. However, this is not always the case. If external variability creates the need for a buffer, is it waste?" (*21*). Providing different service levels increases variability, but would the railroad be better off if it were to only offer one service level? "The point is that while not all variability is waste, all variability will lead to a buffer which indicates that logistical (but not necessarily financial) performance has suffered" (*21*). Therefore, it becomes the task of yard management to reduce internal variability and the task of network management to manage external variability so that the bad sources (like arrival variability) are reduced and the good sources (like service level differentiation) increase profit.

Implementing Lean Railroading

With the advent of scheduled railroading, railroads have already taken an important first step in creating an environment that lean railroading can succeed in by reducing external variability for the yard. Implementation steps are as follows:

0. Eliminate direct waste. Take a fresh look at the yard as a system by drawing a value stream map (VSM) (*13, 20*) and try to eliminate obvious sources of waste. Step 0 is used to emphasize that railroads should already be working to eliminate direct waste in their operations.
1. Swap buffers. Decrease the time buffer (dwell time) by reducing the idle time between processes. This is synonymous with enabling continuous flow. Increase the capacity buffer by focusing on improving the performance of the bottleneck.
2. Reduce variability.
   a. Address problems in sorting, rework, car damage, down time, and setups (apply SPC or six sigma).
   b. Implement standardized work plans.
   c. Work with network management to increase on-time arrival of inbound trains.
   d. Level the production schedule in the yard and set the network operating plan.
3. Use continuous improvement.

> Once variability is significantly reduced, we can reduce the capacity buffer while continuing to identify and eliminate variability. Only at this point do we begin to make real gains in productivity. If we do not reduce variability, we will not be able to reduce the capacity buffer without hurting customer responsiveness. The result is a system that continues to improve over time (*21,* p. 5).

Theoretical Importance of Bottleneck

To decrease the time buffer, without a detrimental impact on connection performance, the capacity buffer must be increased. Capacity is defined as the upper limit on the throughput of a production process (*23*). The bottleneck process limits the throughput of a production system. As such, the processing rate (throughput) of the bottleneck process establishes the capacity of the entire system over the long term. Equation 1, Little's Law, can be used to estimate the benefits of improving the bottleneck rate (*23*):

$$\begin{aligned}\text{bottleneck rate} &= \text{yard throughput}\left(\text{cars per day}\right)\\ &= \frac{\text{volume}\left(\text{car count}\right)}{\text{dwell time}\left(\text{days}\right)}\end{aligned} \tag{1}$$

Increasing the bottleneck rate will reduce dwell time for any given volume level in the yard. Therefore, the avenue for the greatest capacity buffer increase lies with improving performance of the bottleneck.

Theory of Constraints

TOC provides a structured approach to improving production system performance by focusing on the system's bottleneck. Goldratt has established the general process in the TOC approach (*24*). For any production system, the TOC approach is as follows.

1. Identify the system's constraint.
2. Decide how to exploit that constraint.
3. Subordinate the remaining resources to the decision in the previous step.
4. Elevate the system's constraint.
5. If in the previous steps the constraint has been overcome, go back to Step 1.

Step 1 means identifying the actual constraints and focusing improvement efforts on the one that impacts the objective (or the goal, in TOC parlance) the most. From the factory physics standpoint, the most important constraint is the bottleneck. Exploiting the bottleneck (Step 2) means managing it in a way that maximizes its throughput. This goes hand in hand with Step 3 since the remaining resources (the nonconstraints) should be managed so that they provide the bottleneck exactly what it needs and nothing more. Efforts should continually be made to elevate the bottleneck (Step 4) until it is broken and a new constraint becomes the most limiting to the system (Step 5). At this point, the process begins again at Step 1 as the new system constraint is identified.

Identifying a Terminal's Bottleneck

The bottleneck can be identified by analyzing where cars spend time as they flow through the yard. A time-and-motion study conducted by the GE Yard Solutions group for one classification yard found that cars were idle for 71% of the 28.2-h average dwell time in the yard (*5*) (Figure 3). The largest portion of this time (14.6 h) was spent in the classification yard (or bowl). A disproportionately long wait time immediately upstream from a production process is a good indicator that process is the bottleneck. This indicates that the pull-down process is the bottleneck. This is consistent with previously published work (*7, 25, 26*) and railroad management experience at CP, CN, and UP (*27*).

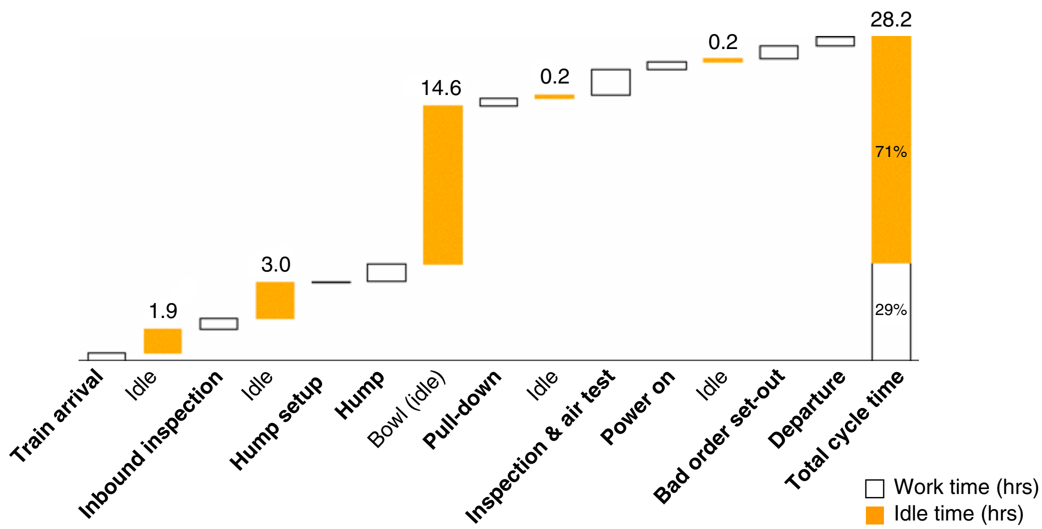

FIGURE 3 Average terminal process cycle time for major Class 1 railroad yard (*5*).

DETERMINING PULL-DOWN CAPACITY

The pull-down process (also called "trimming" or train assembly) consists of blocks of cars pulled from the classification tracks (bowl) and placed together to form outbound trains in the departure tracks. Despite the theoretical importance of the bottleneck in production systems, more work is needed to document and understand details of the pull-down process. A macroscopic evaluation method is presented in Wong et al. (*19*) for use designing new yards or redesigning old yards. The method also serves as a good starting point for evaluating the potential impact of different improvement strategies for existing yards. Equation 2, from Wong et al., estimates the capacity of the pull-down end for a parallel departure yard design (*19*):

$$C_P = \frac{T_M \quad N_E \quad N_C}{(T_H + T_L + N_D \quad T_D)(1.0 + C_F) + T_C} \tag{2}$$

where

C_P = capacity of pull-down end (cars/day),
T_M = productive crew time (min),
N_E = number of pull-down engines,
N_C = average number of cars per block pulled,
T_H = average travel time from classification yard to departure yard (min),
T_L = average travel time from departure yard to classification yard (min),
N_D = average number of doubling maneuvers to be made per pull,
T_D = time required to complete doubling maneuver (min),
C_F = conflict coefficient, and
T_C = average coupling time to couple average size block (min).

In this paper, this equation is refined with additional detail to increase its robustness and then conduct a sensitivity analysis to assess the effectiveness of several bottleneck management improvement options.

Operational Methods

Major activities performed by pull-down crews are coupling cars on the classification tracks and then pulling them to the departure yard (*19*). Pull-down operational methods are closely related to design of the pull-down end of the yard and orientation of the departure yard to the classification yard. Parallel departure yard and in-line departure yard orientations are two of the most common designs. In parallel departure yard designs, the method of making up trains can vary. The first method involves an engine pulling the cars on one track directly to the departure yard, referred to as single pull. In the second method, engines pull cars from several tracks and then move them as a group to the departure yard, referred to as multiple pull. In in-line departure yard designs, trains are usually built with the multiple pull method (*19*).

A detailed analysis of the pull-down process was conducted at the Bensenville Yard (CP) in Illinois, near Chicago. Bensenville has a parallel departure yard design, and both operational methods are used. However, because single pull was the predominant method used, N_D (the average number of doubling maneuvers to be made per pull) can now be used to reflect a similar activity, rework.

"Clean" and "Dirty" Tracks

The term "rework" is used because the pull-down process must correct the sorting of the hump process. Rework occurs on the pull-down end when tracks are "dirty." Kraft defines a dirty track as one that has more separations than blocks, with a separation defined as a group of cars in standing order having the same block (*18*). The number of separations is determined by looking at standing order of cars on the track and counting the number of times the block changes.

A slight modification of Kraft's definition is used here because separation might also be interpreted to mean a gap between cars on the classification track, so the term "separation" is replaced by the term "cut" as defined by Daganzo et al. (*28*). Cut is standard railroad terminology that can be used to describe any situation where a set of cars that shares a common destination track are sequenced together. A cut is a set of cars in standing order all having the same block. If there are more cuts than blocks, then at least one car must be out of place on that track (Figure 4). A switching move is required between every cut (*28*). Because the pull-down process is assembling blocks into trains, anytime there are more

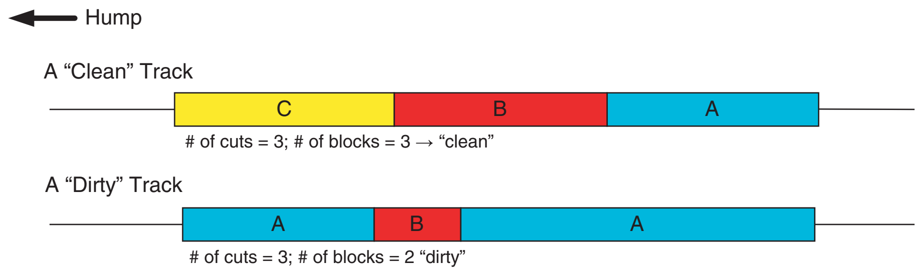

FIGURE 4 Modified definitions of clean and dirty tracks.

cuts than blocks, extra switching moves will be required to accomplish the task of building a particular train. This is the same as doubling a track, which is why N_D and T_D can be used to reflect the impact of rework.

Cycle Time Components

All time parameters in Equation 2 (T_M, T_H, T_L, T_D, and T_C) need to be determined by conducting time studies. The first, the productive crew time (T_M), is the time that the crew is doing productive work. The maximum possible productive crew time is 1,440 min per day, minus the total minutes for meals and breaks (*19*). However, this value should be further refined to reflect real work conditions. Crews will not maintain a maximum pace every minute of the work day because of interruptions, fatigue, and unavoidable delay (*29*). Also, crews will exert different effort levels depending on a variety of factors, such as skill level, motivation, and age. The result is a reduction in the productive crew time and can be accounted for with Equation 3.

$$T_M = \left(1{,}440 - M_B\right) \; P_F \tag{3}$$

where

T_M = productive crew time (min),
M_B = total meal and break time (min), and
P_F = performance factor.

The remaining time parameters can be added together to calculate the cycle time of the pull-down process (Equation 4). The cycle time is the time it takes to complete one cycle of the process.

$$C_T = T_L + T_C + B_R \;\; T_D + T_H + D_R \;\; T_D \tag{4}$$

where

C_T = average pull-down process cycle time (min),
T_L = average travel time from departure yard to classification yard (min),
T_C = average coupling time to couple average size block (min),
T_D = time required to complete doubling maneuver (min),
T_H = average travel time from classification yard to departure yard (min),
B_R = bowl rework occurrence integer (0 or 1),
D_R = departure yard rework occurrence integer (0 or 1), and
$B_R + D_R \leq 1$.

For the pull-down process, cycle time begins when the crew receives the switch list from the yardmaster. It ends when the crew uncouples from the cut of cars after placing them on the required track in the departure yard. The high-level process flow diagram in Figure 5 illustrates this procedure and breaks it into five cycle time components: setup (T_L), coupling (T_C), bowl rework ($B \cdot T_D$), transport (T_H), and departure yard rework ($D \cdot T_D$). It is assumed for this model that rework will occur at most only one time per pull: either in the bowl or in the departure yard. More detail for each cycle time component is found in Dirnberger (*13*).

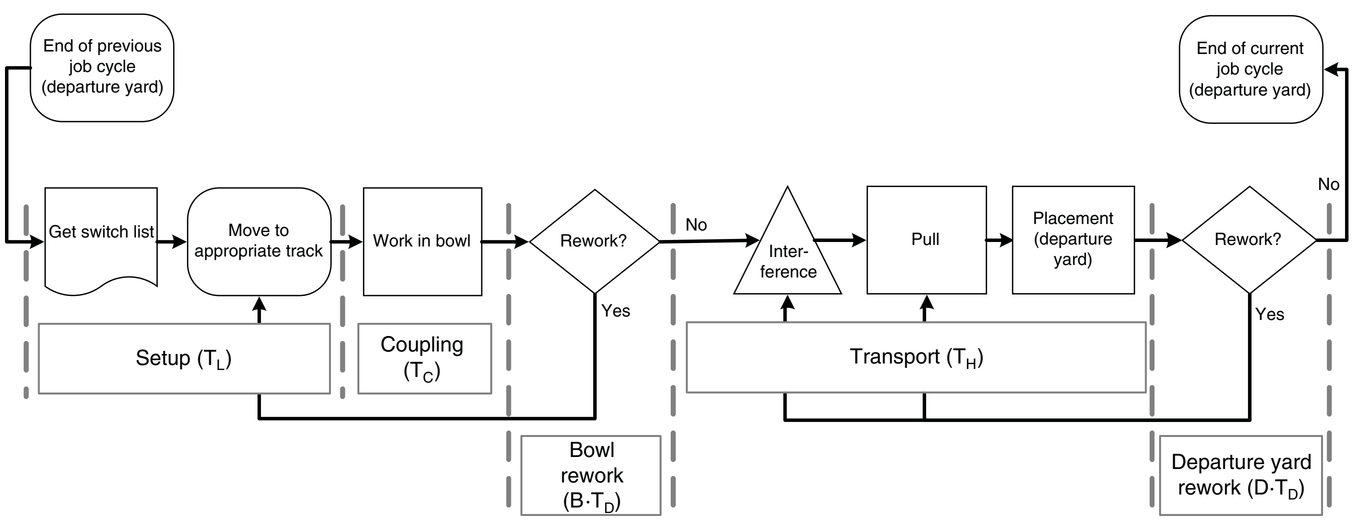

FIGURE 5 Pull-down process cycle time components.

TABLE 2 Calculated Parameters for Baseline Case, Bensenville Yard (CP)

Parameter	Value	Notes
M_B = total meal and break time (min)	135	30-min lunch, 15-min breaks, 3 shifts
P_F = performance factor	0.85	85% is the standard
T_M = productive crew time (min)	1,109	$T_M = (1{,}440 - M_B) \cdot P_F$
C_T = average pull-down process cycle time (min) (no rework)	79	Net travel time, no conflicts
C_T = average pull-down process cycle time (min) (with rework)	98	Net travel time, no conflicts
T_L = average travel time from the departure yard to the classification yard (min)	10	Net travel time, no conflicts
T_C = average coupling time to couple an average size block (min)	48	Calculated using bowl track authority logs, Dec. 26 to Jan. 2, for 22 cars
T_D = time required to complete a doubling maneuver (min) = rework	19	Net travel time, no conflicts
T_H = average travel time from the classification yard to the departure yard (min)	21	Net travel time, no conflicts
N_E = number of pull-down engines	3	Three crews per shift on average
N_C = average number of cars in a cut of a block (cars)	22	Dec. 24, 2005, to Jan. 10, 2006, average
N_D = average number of doubling maneuvers to be made per pull	0.17	Over 4-day period, average of 17% dirty tracks pulled per day
C_F = conflict coefficient	1.55	From Wong et al. (*19*), p. 149, configuration 1 with 3 engines

BOTTLENECK MANAGEMENT IMPROVEMENT ALTERNATIVES

Pull-down time studies were conducted at Bensenville over a period of 4 days during March 2006. The time of day that the observations were gathered was different each day. A total of 15 complete cycles were observed during the available time period. Data from those studies were used, along with other yard measurement data normally tracked by CP, to calculate the parameters for Equations 3 and 4 (Table 2). This allowed Equation 2 to be used to estimate the capacity of the pull-down process for a baseline case.

The baseline case has an estimated capacity of Bensenville is 541 cars per day. To check the accuracy of the estimate, the average daily process car count for 2004 was calculated. CP defines process cars as cars that go through all yard processes: arrival, classification, pull-down, and departure. Average throughput was 521 cars per day. Average throughput should be less than theoretical capacity (*23*); therefore, the estimate is acceptable.

Starting with the baseline, sensitivity analyses were conducted on the following parameters:

1. N_E, increase the number of pull-down engines.
2. N_D, decrease the average number of doubling maneuvers per pull that occur because of rework or "cherry-picking."
3. The primary cycle time components, T_L, T_C, and T_H, reduce the average time to complete each component.

The following sections describe each parameter analysis in detail.

Add Pull-Down Engines

One option to increase capacity at the pull-down end is to use additional engines. Adding one engine resulted in a capacity of 576 cars per day, a 6.5% increase compared with the baseline. Each additional engine increases capacity, but at a diminishing rate of return (Figure 6). The limiting factor when adding another engine is the

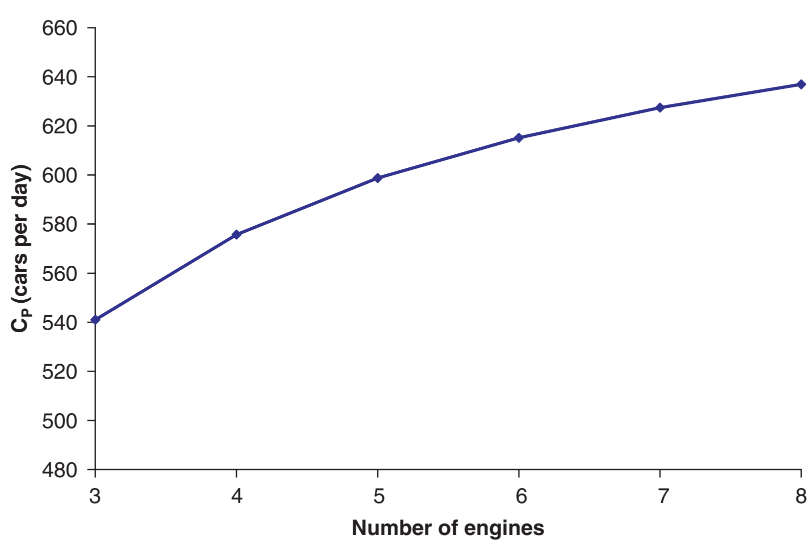

FIGURE 6 Pull-down capacity versus number of engines.

increase in the conflict coefficient. Every additional engine beyond two increases C_F by 1.0 (*19*), although other yard designs will have different conflict coefficients affecting the specific functional relationship. While the option of adding engines results in a relatively large increase in capacity, it is also the most expensive because of the additional engine and labor cost. As a result, adding pull-down engines is often not a cost-effective solution to increasing pull-down capacity.

Pull from Hump End When Idle

At Agincourt Yard (CP) in Toronto, Ontario, Canada, an option has been implemented that increases capacity without increasing interference or engine and labor costs. The hump engine is used to build trains when the hump is idle. This is done by placing the hump in trim mode (disabling the retarders), allowing the engine to enter the bowl and pull blocks from the hump end. This is possible because Agincourt, like Bensenville, has parallel receiving and departure yards. That solution would not be practical for yards with in-line designs.

This solution is consistent with the TOC approach; having identified the system's constraint, yard management was able to exploit the pull-down process by subordinating one of the other resources in the yard (the hump) to it. Hump productivity is generally governed by the number of cars available for sorting so productivity is usually less than the capacity of the hump (*30*). At Agincourt, for the first 6 months of 2005, the highest monthly average hump use was approximately 56% (Toronto report with shifts, unpublished data from CP Railway). Because of this low utilization rate, the hump could be used in trim mode part of the time and still be able to sort all of the required cars. If Bensenville, with a daily average hump utilization of 49% implemented a similar solution, capacity would increase to 586 cars per day, an 8% increase (Bensenville hump statistics, unpublished data from CP Railway). It is assumed that using the hump engine would result in an N_E value of 3.25: three engines from the baseline case plus the equivalent time of 0.25 of an engine from the hump end.

Reduce Number of Doubling Moves

As explained earlier, pull-down crews must perform doubling maneuvers whenever a track requires rework. At Bensenville, an average of 17% of the tracks pulled per day was dirty and required rework. Therefore, for the baseline case, it is assumed that the average number of doubling moves made per pull (N_D) is 0.17. If tracks are kept clean, less rework will have to be performed. This means crews will not have to dig cars out of tracks when they are assembling trains, and capacity would increase by six cars per day for every 3% reduction in N_D. If rework were eliminated, capacity would increase to 576 cars per day, the equivalent of adding an engine on the pull-down end. Keeping the tracks clean requires analyzing the interaction between the hump and the pull-down processes.

Address Cherry-Picking

Doubling maneuvers are also required when the crews must cherry-pick high-priority cars from a track. Cherry-picking is the most commonly accepted method of protecting connections in danger of being missed when outbound train capacity is exceeded. It "exacerbates the capacity bottleneck which already exists there, and reduces the throughput of the whole facility" (*7*). Kraft (*7, 8, 17, 18*) has extensively studied the issue of connection priority and developed many useful solutions but was not able to develop quantitative metrics on the performance of algorithms (*7*).

The additional switching required for rework and cherry-picking is similar. Therefore, modifying N_D can quantify some of the parameters needed to assess the impact of Kraft's proposed solutions on pull-down capacity. If it is assumed that cherry-picking occurs in 10% of the pulls, then N_D is 0.27 for the baseline case (0.17 for rework, 0.10 for cherry-picking) and capacity is 522 cars per day. Eliminating cherry-picking would increase capacity 3.6% to 541 cars per day. In addition to the capacity increase, it would improve connection performance.

Reduce Component Cycle Times

Faster cycle times result in increased process throughput. Setup (T_L) and transport (T_H) cycle times can be reduced by eliminating unnecessary moves, throwing fewer switches, increasing engine speed, preventing engine breakdown, using experienced crews, and so on. Coupling (T_C) cycle time can be reduced by eliminating gaps between cars on the classification tracks, better track inventory control, and quicker correction of out-of-alignment drawbars. Potential capacity increases are shown in Figure 7. The dashed line represents the resultant capacity increase for setup and transport time reduction because they have the same impact on capacity (10 to 12 cars per day per 1-min reduction). The solid line represents the impact of reducing coupling time, which has a smaller impact (4 cars per day per 1-min reduction).

SUMMARY OF RESULTS

Each option presented can increase the capacity of the pull-down process, but with differing incremental rates. Adding one engine at the pull-down end increases capacity by 6.5%, but the next engine only increases capacity by 4.0%. Other alternatives will most likely be more cost-effective because of the labor, equipment, and maintenance costs that result from adding pull-down engines. Eliminating rework results in the same 6.5% increase, and using the hump engine when it is idle provides an increase of 8.3%. Reducing cycle times also increase capacity but at different rates depending on which component of the pull-down process is improved.

Simply telling the crews to work faster is unlikely to result in significant cycle time reduction. Better management of the process and its interactions with other processes is required to achieve a meaningful and sustained reduction. That will result in capacity gains without large labor or capital expenses. This also means that multiple improvement options can be accomplished together. If the hump engine is used, rework eliminated and the average component cycle times are reduced (setup by 1 min, coupling by 3 min, and transport by 2 min) capacity would increase by approximately 26% to 681 cars per day.

CONCLUSIONS

Because a classification yard is a system, managing the interactions between the processes is just as important as managing the individual processes. Those options involving the interaction between the

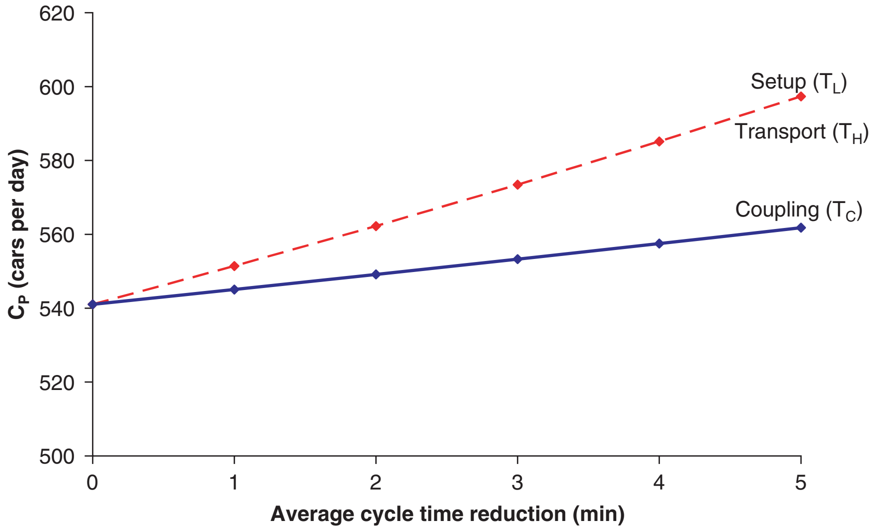

FIGURE 7 Pull-down capacity as function of cycle time reduction.

pull-down and the hump result in greater capacity increases. This is consistent with TOC. Eliminating rework by improving the sorting process results in the same capacity increase as adding a pull-down engine does. This alternative merits further analysis because of its potential to increase capacity without major capital, equipment, or labor expense.

One of the principal findings of this work is that the humping process should be subordinate to the pull-down process. Because the pull-down is the bottleneck, the hump should be managed and operated so that it provides the pull-down exactly what it needs when it needs it. The practice of measuring hump performance merely on number of cars processed can, and often does, contribute to poor pull-down performance because it encourages quantity rather than quality in the car classification operation. To manage the interaction between the hump and the pull-down processes better, a measurement of quality of sorting is needed. Dirnberger describes the development of a quality of sort metric to reduce occurrence of dirty tracks and measure adherence to a static track allocation plan if one is in place (*13*).

This paper presented the lean railroading approach and discussed the bottleneck management component. Increasing pull-down capacity will help enable railroads to swap the time buffer for a capacity buffer. This will reduce dwell time, leading to improved service reliability and network efficiency. By combining scheduled railroading with a version of lean in their yards, CP reported average terminal dwell fell from 30.4 h in March 2005 to 21.7 h in March 2006 (*11*). During the same period, average train speed increased by 3.6 mph.

ACKNOWLEDGMENTS

The first author, while at the University of Illinois, was supported by a CN railroad engineering research fellowship. Both authors are grateful for the assistance provided by CP Railway, CN Railway, and Prescott Logan of GE Yard Solutions during the course of this research. They also thank Darwin Schafer of the University of Illinois at Urbana–Champaign for his assistance in gathering and analyzing data.

REFERENCES

1. Association of American Railroads. Rail Intermodal Transportation. Background Papers www.aar.org/GetFile.asp?File_ID=143. Accessed July 17, 2006.
2. Gray, J. T. Union Pacific Perspective. Presented at 85th Annual Meeting of the Transportation Research Board, Washington, D.C., 2006.
3. Martland, C. D., P. Little, O. K. Kwon, and R. Dontula. *Background on Railroad Reliability.* AAR Report no. R-803. Association of American Railroads, Washington, D.C., 1992.
4. Murray, T. Wrong Train Running. *Trains,* July 2002, pp. 30–35.
5. Logan, P. Role of Yard or Terminal in Operating Performance and Capacity. Presented at 85th Annual Meeting of the Transportation Research Board, Washington, D.C., 2006.
6. Kwon, O. K., C. D. Martland, J. M. Sussman, and P. Little. Origin-to-Destination Trip Times and Reliability of Rail Freight Services in North American Railroads. In *Transportation Research Record 1489,* TRB, National Research Council, Washington, D.C., 1995, pp. 1–8.
7. Kraft, E. R. Priority-Based Classification for Improving Connection Reliability in Railroad Yards Part I of II: Integration with Car Scheduling. *Journal of the Transportation Research Forum,* Vol. 56, No. 1, 2002, pp. 93–105.
8. Kraft, E. R. Priority-Based Classification for Improving Connection Reliability in Railroad Yards Part II of II: Dynamic Block to Track Assignment. *Journal of the Transportation Research Forum,* Vol. 56, No. 1, 2002, pp. 107–119.
9. Ytuarte, C. Getting Over the Hump. *Railway Age,* Oct. 2001, pp. 29–30.
10. Association of American Railroads. Weekly Railroad Performance Measures. www.railroadpm.org. Accessed May 31, 2006.
11. Canadian Pacific Railway. Key Metrics. www8.cpr.ca/cms/English/Investors/Key+Metrics/default.htm. Accessed May 31, 2006.
12. CN Railway. Key Performance Measures. www.cn.ca/about/investors/perfmeasures/en_InvPerfMeasures.shtml. Accessed May 31, 2006.
13. Dirnberger, J. R. *Development and Application of Lean Railroading to Improve Classification Terminal Performance.* Master's thesis. University of Illinois, Urbana–Champaign, 2006.
14. Martland, C. D., P. Little, and J. M. Sussman. Service Management in the Railroad Industry. In *Conference Proceedings 2: Railroad Freight Transportation Research Needs,* TRB, National Research Council, Washington, D.C., 1994, pp. 89–103.
15. Association of American Railroads. Analysis of Class I Railroads, 2002–2005. www.aar.org. Accessed May 24, 2007.
16. AASHTO. Freight-Rail Bottom Line Report. *Transportation—Invest in America* series, Washington, D.C., 2002. transportation.org/bottomline/bottomline2002.pdf. Accessed May 24, 2007.

17. Kraft, E. R. Implementation Strategies for Railroad Dynamic Freight Car Scheduling. *Journal of the Transportation Research Forum,* Vol. 39, 2000, No. 3, pp. 119–137.
18. Kraft, E. R. A Hump Sequencing Algorithm for Real-Time Management of Train Connection Reliability. *Journal of the Transportation Research Forum,* Vol. 39, No. 4, 2000, pp. 95–115.
19. Wong, P. J., M. Sakasita, W. A. Stock, C. V. Elliott, and M. A. Hackworth. Railroad Classification Yard Technology Manual—Vol. I: Yard Design Methods. FRA/ORD-81/20.I. Final report. National Technical Information Service, Springfield, Va., 1981.
20. Rother, M., and J. Shook. *Learning to See: Value Stream Mapping to Create Value and Eliminate Muda,* Version 1.2. Lean Enterprise Institute, Brookline, Mass., 1999.
21. Spearman, M. L. To Pull or Not to Pull: What Is the Question? Part II: Making Lean Work in Your Plant. *White Paper Series,* Factory Physics, Inc., College Station, Tex. www.factoryphysics.net. Accessed Aug. 20, 2004.
22. Wellington, A. M. *The Economic Theory of the Location of Railways,* 6th ed. John Wiley and Sons, New York, 1900.
23. Hopp, W. J., and M. L. Spearman. *Factory Physics: Foundations of Manufacturing Management,* 2nd ed. McGraw-Hill, Boston, Mass., 2001.
24. Goldratt, E. M. *Theory of Constraints*. North River Press, Great Barrington, Mass., 1990.
25. Petersen, E. R. Railyard Modeling: Part I. Prediction of Put-Through Time. *Transportation Science,* Vol. 11, No. 1, 1977, pp. 37–49.
26. Petersen, E. R. Railyard Modeling: Part II. The Effect of Yard Facilities on Congestion. *Transportation Science,* Vol. 11, No. 1, 1977, pp. 50–59.
27. McClish, R. Analyst Meeting Presentation. Union Pacific Railroad. www.up.com/investors/attachments/presentations/2005/analysts_conf/mcclish.pdf. Accessed July 12, 2005.
28. Daganzo, C. F., R. G. Dowling, and R. W. Hall. Railroad Classification Yard Throughput: The Case of Multistage Triangular Sorting. *Transportation Research A,* Vol. 17, No. 2, 1983, pp. 95–106.
29. Niebel, B. W., and A. Freivalds. *Methods, Standards, and Work Design,* 10th ed. WCB/McGraw-Hill, Boston, Mass., 1999.
30. Wetzel, J. A. Optimal Use of Classification Yards. In *Transportation Research Record 1029,* TRB, National Research Council, Washington, D.C., 1985, pp. 36–41.

The Railroad Operating Technologies Committee sponsored publication of this paper.