Research Contribution 23 *December 1998*

STATISTICAL COMPARISON OF METHODS USED IN HARVESTING WORK STUDIES

by

Eldon D. Olsen
Mohammad M. Hossain
Mark E. Miller

College of Forestry

Forest Research Laboratory
Oregon State University

The Forest Research Laboratory of Oregon State University was established by the Oregon Legislature to conduct research leading to expanded forest yields, increased use of forest products, and accelerated economic development of the State. Its scientists conduct this research in laboratories and forests administered by the University and cooperating agencies and industries throughout Oregon. Research results are made available to potential users through the University's educational programs and through Laboratory publications such as this, which are directed as appropriate to forest landowners and managers, manufacturers and users of forest products, leaders of government and industry, the scientific community, and the general public.

The Authors

Eldon D. Olsen and Mark E. Miller are associate professor and research assistant, respectively, in the Department of Forest Engineering, College of Forestry, Oregon State University, Corvallis, Oregon. Mohammad M. Hossain is faculty member (teaching and research), Institute of Forestry, University of Chittagong, Bangladesh.

Disclaimer

To Order Copies

Copies of this and other Forest Research Laboratory publications are available from:

Forestry Publications Office
Oregon State University
227 Forest Research Laboratory
Corvallis, Oregon 97331-7401

Phone: (541) 737-4271
FAX: (541) 737-3385
email: forspub@frl.orst.edu
http://webdata.for.orst.edu/forestry/pubs/frl/

Please indicate author(s), title, and publication number if known.

Recycled
Paper

Statistical Comparison of Methods Used in Harvesting Work Studies

by

Eldon D. Olsen
Mohammad M. Hossain
Mark E. Miller

Abstract

Olsen, ED, MM Hossain, and ME Miller. 1998. STATISTICAL COMPARISON OF METHODS USED IN HARVESTING WORK STUDIES. Forest Research Laboratory, Oregon State University. Research Contribution 23. 41 p.

Work-study guidelines were developed using field data from thinning sites in the Oregon Cascade Range. Regression of detailed time study and shift-level data predicted harvesting production rates. Statistical analysis showed the relative difference in the discriminating power between shift-level studies versus detailed time studies. Indicator variables tested if there were significant differences between harvesting treatments. Confidence intervals demonstrated the improved effect of longer study lengths for both shift-level and detailed time studies.

Table of Contents

List of Tables

List of Figures

Introduction

In harvesting research, we often want to determine whether two alternative harvesting systems have different production rates. These alternatives may involve using different techniques, using different equipment, or working under different conditions. In this monograph, the examples for harvesting alternatives reflect actual field data for thinning under different residual tree densities.

The purpose of this monograph is to suggest ways in which production of harvesting alternatives can be compared more effectively. This improvement can be achieved by controlling or isolating unwanted variation in the data being collected.

The monograph will discuss how both detailed time studies and shift-level studies can be used to properly collect data. Improvements in the techniques traditionally used in analysis for these two study methods will enhance our ability to compare the production of harvesting alternatives. The main improvements are 1) using *error terms* to statistically compare production estimates, 2) using *indicator variables* in regression analysis, and 3) using regression analysis in shift-level studies. Appendix A gives simple definitions of some of the statistical terms used in this monograph; the terms defined in Appendix A appear in italics at their first use in the text.

In this monograph, traditional studies are described first, followed by the factors that contribute to variability. *Multiple least-squared linear regression* analysis of a typical detailed time study will be used to show the magnitude of variation that can be expected from each source. This analysis will demonstrate why the variation associated with a shift-level study is so much larger than that for a detailed time study. Next, indicator variables will be used for both detailed time study and shift-level analyses to attempt to find differences between thinning alternatives. The indicator variable method makes it possible to test more than two thinning alternatives simultaneously.

The monograph also demonstrates how study length affects the size of the standard error. This information will help researchers determine the length of study needed to detect significant differences between harvesting alternatives. A comparison of the relative size of the standard errors that can be achieved by detailed time and shift-level studies will help researchers determine when to use these two study methods.

Traditional Approaches

Studies of forestry operations are challenging because of the variability in the data. A researcher usually first attempts to study a sample of field operations in which all of the conditions except the one being investigated are as similar as possible. For those conditions that cannot be controlled adequately, the researcher attempts to measure the variations so that their possible influence can be isolated.

Data Collection

The most common methods for researchers to use in collecting production data are detailed time studies and shift-level studies. A detailed time study collects data on each production cycle. For a skyline system, this would be one round trip of the carriage between the landing and the stump. Such cycles typically last just a few minutes. The conditions of each cycle (such as yarding distance) are recorded. Any delay during the cycle is also recorded. With the shift-level method, the daily average production is calculated from the total output and the total hours for each day. Some data on the average conditions for the day are recorded, as are delays lasting 10 min or longer.

Analysis

Ideally, statistical tests are used to determine whether the production rates observed during the studies are significantly different. From the sample, an average production rate is estimated, usually with aid of regression analysis. The regression analysis is able not only to predict average production, but also to quantify the unwanted variability that accompanies the estimate. This unwanted variability in the sample is called the *standard error (S.E.) of the estimate*. To be declared significant, the difference between two harvesting techniques must be greater than the standard errors.

The traditional analysis of detailed time study data is by regression. The data for each alternative are used to create separate regression equations. Then production rates are estimated with a common set of conditions as input to the regression equation for each harvesting alternative. These production rates are then contrasted, usually without any statistical basis, to see whether a difference appears to exist.

With shift-level data, analysis is even simpler. The daily averages are plotted over some period, and the plots of alternative harvesting techniques are contrasted to see if any difference in the daily production can be inferred.

Improvements to the Traditional Approaches

Estimating the Standard Error

For both detailed time study and shift-level data, comparison of harvesting alternatives is aided by knowing the amount of variation, as measured by the standard error, in each of the production estimates. The smaller the standard error, the better any production difference between harvesting alternatives can be detected.

Using Indicator Variables

For analysis of detailed time study data, rather than each harvesting alternative being regressed separately, the data sets can be combined. Data from each harvesting alternative are denoted with a unique indicator variable. Any statistically detectable difference between harvesting alternatives is then automatically apparent during the regression analysis. If the indicator variable is statistically significant, it means the harvesting alternatives have different production rates.

The same indicator variable approach can be used in shift-level analysis, by combining the data sets for different harvesting alternatives. Again, if the indicator variable is significant, it indicates a production rate difference between the harvesting alternatives.

Regressing Shift-Level Data

To analyze shift-level data, rather than using just the daily average production, a researcher can conduct regression analysis. Such analysis requires that the data collection also record information on logging conditions.

Data Used in Examples

The field data used in this monograph were collected as part of a multi-year study, the Willamette Young Stand Project. Data for both felling and extraction activities were collected using detailed time study and shift-level methods concurrently. The overall project investigated differences in production between three thinning alternatives. The "heavy-thin" alternative left 53 residual trees/ac. The "light-thin" alternative left 115 trees/ac. The "light-with-openings" alternative had 0.5-ac clearcut openings on 20% of the area and light thinning elsewhere, which resulted in an average of 92 trees/ac. Other details are given in Appendix B.

Estimating the Standard Error in Production Studies

Just as land surveyors must understand the sources of error in each of their instrument readings, researchers conducting a production study must understand the source and magnitude of each error in their data, as well as how these error sources combine to give a total expected error. Production studies estimate the number of cycles per hour that can be expected while harvesting under various logging conditions. The amount of variation in the data affects the accuracy of the production average being estimated. Mean production rates of two different logging alternatives can be compared statistically if the amount of variation associated with each mean is known.

This section investigates the sources of variation in the production estimate. The methods of collecting and analyzing the data influence how much variation occurs. For instance, shift-level data have delay variation and unmeasured conditions embedded in the production estimate. Knowing the amount of potential variation will help determine which method of data collection should be used for the researcher to achieve a desired level of accuracy in the estimate.

Data sets from detailed time and shift-level studies each have their own relative amount of variation. Also, each method involves a trade-off between the cost of the study and the degree of accuracy provided. For instance, a detailed time study has high accuracy, can quantify the effect of such variables as extraction distance, and can isolate small delays; however, these studies are costly to conduct. Shift-level studies, on the other hand, are less expensive, but have high variability. Shift-level studies cannot quantify the effect of most variables or isolate small delays. In this section, we demonstrate the variability in detailed time studies and use this to explain the greater variability in shift-level estimates.

Sources of Variation

The objective of production studies is to quantify average cycle times and measured conditions. The sources of variation associated with randomly measured variables, unmeasured conditions, and delays are undesirable, and minimizing these sources will help achieve the objectives. However, estimated averages may vary in two studies of alternatives on the same site for at least five reasons:

1. The measurable harvesting conditions may be different. For instance, one study may have a longer average yarding distance.
2. Unmeasured conditions will also be a source of variation. For example, terrain or stand conditions may influence yarding difficulty.
3. The random occurrence of delays may cause a sampling error. Even though the two studies would have the same underlying delay distribution, the estimated percentage of delay may differ.
4. Randomly measured variables may cause a sampling error. For example, two samples may reflect a different number of pieces per cycle even though they came from the same source.
5. Logging methods may yield truly different average cycle times in the two studies. For instance, prebunching speeds up the cycle.

If the variation from these sources can be isolated, researchers can discriminate the true differences that are occurring. (This assumes that there is no measurement error in the variables that are recorded.)

Uncontrolled Variation

Sources of unwanted variation include the following:

- Level of training or skill: instruction, learning, aptitude
- Level of motivation
- Work pace
- Ergonomic considerations: conditioning, rest periods, human-machine compatibility
- Team coordination
- Incentive pay systems
- Proper engineering planning and layout
- Proper match of equipment to the task
- Weather conditions.

Ideally, all of these factors are standardized. Standardization can be accomplished by using the same equipment and crew to compare two harvesting systems. Sufficient time and training on both systems is needed for workers to reach equivalent levels of competency on each. Observations of the harvesting alternatives should also be randomized. Thus, any unknown or uncontrolled effects will be dispersed equally among the observed alternatives.

Even with standardization, productivity of the individual alternatives will vary, as a result of natural variations in stand and terrain conditions, machine performance fluctuations, and daily differences in crew performance. The S.E. of the estimate from the regression is a measurement of this internal variation. Appendix C discusses how to eliminate as much variation as possible during regression analysis. This enhancement produces production estimates with the least possible amount of variation.

When the goal is to compare the relative productivity of several harvesting systems, three main things can be done to minimize the residual, unwanted variability: 1) measure all significant variables, such as extraction distance; 2) measure all delays; and 3) conduct the observations over a long enough period.

Examples of Variation

To explore the sources of variation in production studies, we analyzed yarding production of a skyline thinning operation. A detailed time study of 3 days of operation provided the field data on production times, logging conditions, and delays. Distribution information was summarized for the variables and delays. Regression analysis provided the basic relationships and effect of each variable. Standard errors were then used to calculate the amount of variation that would occur in the estimation of the average cycle time. Finally, a simulation model used the regression equations, variables, and delays to generate sample cycle times. The variation in these simulated cycle times mimics what is occurring on a logging operation.

When comparing sample data from two harvesting alternatives, it is rare to have identical values for the variables. For instance, it is unlikely that the average extraction distance for both samples would be 600 ft. When samples do have different values, the cycle time is seriously biased, unless regression is used to isolate the effect. Through regression, the same distance can be used for both harvesting alternatives. Other variables that make a major contribution to the cycle time for skyline systems are the number of pieces and carriage height.

Table 1. Results of regression to estimate cycle time.

Variable	Mean	Standard deviation	Coefficient	Cycle time (min)	Relative effect (%)
Constant	1 min	0.416	1.312	1.312	35.3
Slope distance	600 ft	Controlled	0.00249	1.494	40.2
Lateral distance	21 ft	15	0.00932	0.196	5.2
Number of pieces	4.275	0.878	0.138	0.589	15.9
Carriage height	44.77 ft	13	0.0149	0.667	18.0
Slope	16.3%	7.51	0.00861	0.140	3.8
Multispan	1		–0.466	–0.466	–12.5
Preset logs	1		–0.220	–0.220	–5.9
Total time				3.714	

For this example, a sample of 200 cycle times was used to establish the regression equation. This equation is typical of the type of regression results obtained from many such analyses, with the *coefficient of determination* (R^2) in the range of 50%. Cycle time is the dependent variable. The independent variables used in the multiple regression equation are identified in Table 1.

Table 1 shows the amount of variation, measured by the *standard deviation*, inherent in each of the measured variables. The constant of 1.312 min can be thought of as the fixed portion of the cycle, which is not affected by the measured variables. The indicator variables, multispan and preset logs, are either "on" or "off". In this example we wish to show their effect, so our estimate of total cycle time assumes both variables were "on". Multiplying the regression coefficients by the average value of the remaining measured variables gives the number of minutes that each component contributes to the total cycle time, as shown in Table 1. The magnitude of each variable's effect on the delay-free cycle time is shown as a percentage. As a reference, the small delays, large delays, and landing changes each contribute an additional 10%–15%. The S.E. of the Y estimate [the standard deviation of the constant, which is 0.416 min (Table 1)] is the variation not explained by the independent, measured variables.

Small delays occurred about once per cycle, and their duration was exponentially distributed. The standard deviation of an exponential distribution equals the mean, which in this case was 0.65 min.

Simulation

A discrete-event simulator, ProModel PC, was used for the simulation. This standard industrial production simulator can generate variates, mimic activities, and summarize results.

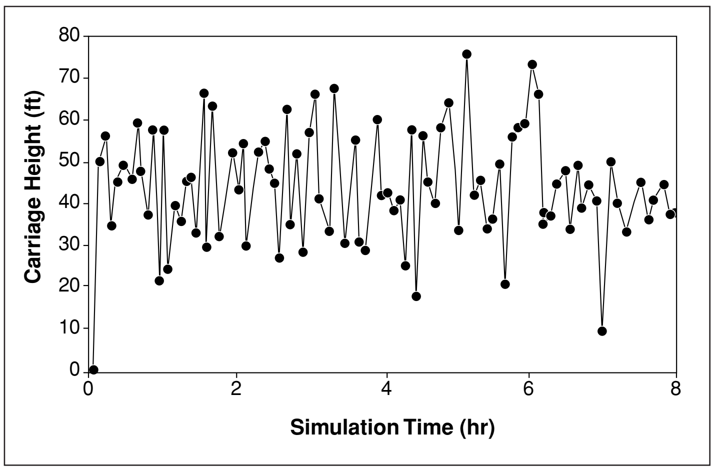

Figure 1. Simulated field condition of carriage height showing variability.

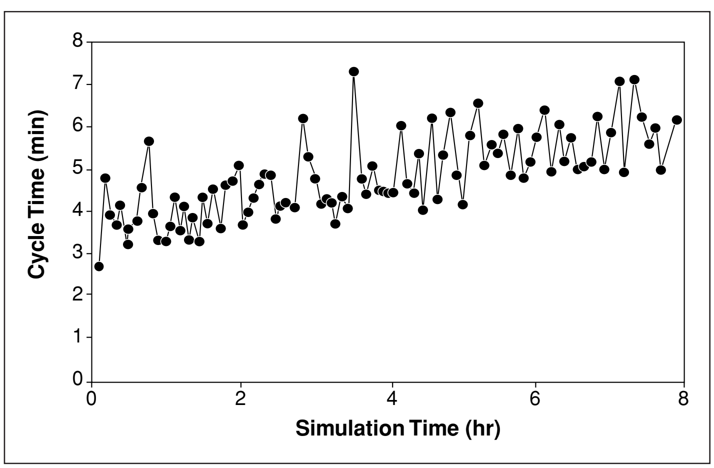

Figure 2. Simulated cycle time variability.

The regression equation obtained from a detailed time study was used as the basis for the cycle time. In each simulated cycle, a random value based on the means and standard deviations shown in Table 1 was generated for each variable. For example, Figure 1 shows carriage heights on each of the 100 consecutive cycles in one 8-hr day. Each randomly assigned value was multiplied by the regression coefficient to find the increase in cycle time caused by that variable. This process was repeated for each variable; combining values for all variables in the regression equation simulated a delay-free cycle time. The mean and S.E. for the day were calculated after 100 cycles were run. Figure 2 shows a representative plot of simulated cycle times for one day; the slight upward trend occurs because the slope yarding distance increased as the day progressed.

The regression and delay analyses of the detailed time study allowed us to separate and quantify the sources of variation in the production estimates. Simulation allows us to selectively recombine these sources of data to demonstrate why production estimates from shift-level studies vary more than those from detailed time studies. It also demonstrates the magnitude of unwanted variation that is removed during regression and delay analyses. Because all other conditions are held identical during a simulation, researchers can observe the individual sources of variation.

For the detailed time study, the S.E. of the Y estimate would be 0.416 min. The simulation was run again with variation in only the measured variables to find an additional source of error of 0.408 min. This result indicates that if regression is not used, this much additional variation can be expected. If small delays were not measured, an additional 0.650 min worth of error per cycle could be expected. Errors from all three sources are of similar magnitude. The first source is inherent variation, which is almost impossible to control. The other two, however, are eliminated in detailed time studies. Traditional shift-level analysis, unfortunately, retains variation from all three of these sources.

Using Indicator Variables to Denote Harvesting Conditions

A common purpose of a research study is to determine whether average cycle time differs between harvesting alternatives. Two approaches in multiple linear regression can be used, and this section demonstrates the use of both approaches on the same data set.

The first approach produces a separate regression equation for each alternative. A standard set of representative independent variables is then inserted into each regression equation. For instance, an extraction distance of 600 ft, which is near the average extraction distance for all of the data, might be used in each of the equations. Similar substitutions are made for all of the variables. We then obtain an average cycle time for each of the alternatives. The S.E. values from the regressions can be used in a *Student's t-test* to determine if the means differ significantly.

The second approach combines all of the alternatives into one large data set. The alternatives are identified with indicator variables, with one alternative designated as the reference category. The number of indicator variables needed is one less than the number of alternatives being tested. Thus, if there are four alternatives, three indicator variables are needed (Table 2). These variables are not measured in the way that quantitative variables are; rather, they are coded with 1 or 0 to indicate yes/no or on/off of a particular condition. For example, an indicator variable can be used to document whether preset chokers were in place during a given cycle: 0 = not preset; 1 = preset. The indicator variables then permit comparison of each alternative against the reference.

Table 2. Values of indicator variables to define four treatments.

Treatment	Indicator Two	Indicator Three	Indicator Four
One (reference)	0	0	0
Two	1	0	0
Three	0	1	0
Four	0	0	1

Regression of the combined data set gives a single equation. The statistical significance of the alternatives is automatically reported by the t-value for each indicator variable. If an indicator variable is significant, its coefficient represents a statistically significant difference between the average cycle time for that alternative and the average cycle time for the reference.

To demonstrate the advantages and disadvantages of these two approaches, we evaluated a skyline thinning op-

Table 3. Comparison of two approaches to regression: individual regressions approach (separate equation for each alternative) versus indicator variable approach (alternative data combined into one equation). Zero values indicate variable was not significant.

Regression results	Individual regressions approach: Light	Light between openings	Heavy	Opening	Indicator variable approach
Overall statistics					
N	285	255	301	96	855
S.E. (min)	0.8077	0.6373	0.6693	0.7385	0.7413
R^2 (%)	54.4	36.2	43.2	36.6	52.7
Coefficients (min)					
Constant	2.668	2.422	2.604	2.178	2.627
Slope	0	0	0	0.0252	0
Span	1.136	0	–0.282	0	0.136
Slope distance	0.002933	0.002165	0.002486	0	0.002650
Lateral distance	0.02218	0.01010	0.01288	–0.01070	0.01216
Logs	0.2287	0.2674	0.1098	0.4638	0.2391
Tops	0.1579	0.2146	0.0973	0.3328	0.1960
Fiber	0.1526	0.1146	0	0.1957	0.1180
Preset logs	0	0	–0.4740	0	–0.4684
Carriage height	0	0	0	0.01040	0.00265

eration involving four alternatives. Table 3 contains the regression results. The light-thin alternative is the reference in this example. However, the light-with-openings alternative was partitioned further into the clearcut openings themselves and the lightly thinned areas between the openings. We would expect results for the lightly thinned areas between the openings to be similar to those for the light-thin alternative. Results for the openings themselves might be similar to those for the heavy-thin alternative.

The sample size, N, for the indicator variable approach is the sum of the cycles for all individual alternatives, minus about 10% of the cycles that were reserved for validation.

The S.E. is the variation expected when observing individual cycles. It varied between 0.6373 and 0.8077 min. This value represents the inherent variation that the regression was not able to explain. The S.E. value for the indicator variable approach is similar to those in the individual regressions. However, the coefficient of determination, R^2, was lower on most of the individual regression equations than on the indicator variable regression.

The coefficients in the regression equations were fairly consistent. The constants were all around 2.6 min. The coefficients for slope distance, lateral distance, and piece counts (logs, tops, fiber) were similarly consistent. These variables have major impact on the cycle time. On the other hand, coefficients for the remaining variables, which include slope, span, and carriage height, all of which have less impact, were not consistent. The inconsistencies were most noticeable for the opening alternative. This alternative is a clearcut and so might be expected to have different operational characteristics from the thinnings. However, the sample size for this alternative was smaller and the range of slope distances was narrow, which is why slope distance was not a significant variable. The indicator variable for Preset logs applied mainly during the heavy-thin operation. The coefficient of –0.47 min indicates that a cycle took about 0.5 min less when the chokers were preset.

The regression equation developed using the indicator variable approach had coefficients for each type of thinning. These coefficients were –0.470 min for the light-between-openings alternative, –0.872 min for the heavy-thin alternative, and –0.729 min for openings. The coefficients represent how much less time a cycle took for these alternatives than for the light-thin alternative, which was the reference.

The next step is to insert representative values into the regression equations to predict the cycle times. For each variable, the same value was inserted into each equation, for instance 600 ft as the slope distance. Using these rep-

Table 4. Comparison of equations from individual regression and regression using indicator variables. The light-with-openings treatment was segregated into harvesting that was done in the openings and harvesting that was done in the lightly thinned areas between openings.

	Individual Regression (min)			Indicator Variable (min)		
Alternative	Mean	S.E.	N	Mean	S.E.	N
Light	5.665	0.048	285	5.434	0.025[1]	855[1]
Light between openings	4.851	0.040	255	4.964	—	—
Heavy	4.654	0.039	301	4.562	—	—
Openings	4.488	0.075	96	4.705	—	—

[1]Combined for all data.

resentative values eliminates any bias that might result from alternatives having different field conditions, such as shorter skyline corridors. The resulting cycle times for our example are shown in Table 4. When the alternatives were ranked by cycle time, the results for the two approaches were consistent. The light-thin alternative had the longest cycle time. The light-between-openings was next longest, as would be expected. The other two alternatives (heavy-thin and openings) had the smallest cycle times. Again this was expected.

A t-test was performed to determine whether the cycle times predicted based on the individual regressions were significantly different. The S.E. for these predicted cycle times, also called the *standard error of the mean*, is found by dividing the S.E. from the regression by the square root of the sample size. For instance, for the light-thin alternative, $0.8077/\sqrt{285} = 0.048$ min. The resulting S.E. values are shown in Table 4. All of the cycle times for alternatives were significantly different at the 95% confidence level.

Because the indicator variable approach leads to only one equation, it has only one S.E.: $0.7413/\sqrt{855} = 0.025$ min. The S.E. is about half the size of those for individual regressions because the sample size was much larger (the total, combined number of cycles). The software used for the indicator variable regression analysis automatically did t-tests on the indicator variable coefficients as well. All of the indicator variables were significant at the 95% confidence level.

Reducing the size of the S.E. value in the t-tests gives researchers an important advantage. With the smaller S.E. values, they can detect smaller differences among means. Conversely, for a given cycle time difference, they are able to assign a higher level of confidence to the results of the t-test.

It thus appears that in testing for significant cycle time differences among harvesting alternatives, the indicator variable approach has distinct advantages over the individual regression approach. First, the S.E. value is smaller and thus the test is more discriminating. Second, the significance of the alternative is tested automatically by testing the coefficient of the indicator variable. Third, regression equation coefficients do not vary among the alternatives, as they may in the individual regressions approach (Table 3).

Regressing Shift-Level Data

Shift-level data are usually collected over long periods of time. With several months of consecutive data, the long-term productivity of a harvesting system can be calculated. Shift-level data are summaries of the daily activity as recorded by one of the logging crew. On a skyline operation, the yarder operator is usually in the best position to record the data. An example of a shift-level form for a yarding operation is shown in Figure 3. Since large delays (≥10 min) are recorded, they can be subtracted from the total work time for the day. The remaining time represents the productive hours for the day and also includes the small delays (<10 min). A total piece count and the number of cycles for the day are also recorded. The productivity can then be calculated in cycles (or pieces) per hour.

Regression has been used infrequently with shift-level production data, but it may be useful for establishing relationships in the data and testing for long-term trends. The use of regression removes variability caused by the measurable variables. Unfortunately, the variables recorded during shift-level data collection are usually limited to a general description of the operation that day, as can be seen in Figure 3. The units of production also differ from the cycle time used in detailed time studies. Because of the way the data are collected in shift-level studies, production is reported as cycles per hour for skidding and trees per hour for felling. In future studies, it would be advisable to record such daily variables as the average extraction distance and the average piece size.

Indicator variables can be used in regression to identify treatment effects and operator influence. The use of regression will help minimize the S.E. and make treatment comparisons more discriminating. Three examples of regression analysis of shift-level data from skidding and felling operations are discussed below.

The same equipment and crew were used on each skyline operation. With both felling and the ground-skidding system, the production of individual operators could be observed. This also affected the production that was regressed. Although yarding distance was not recorded with skyline systems, whether it was short, medium, or long could be inferred from how long the equipment stayed on the same corridor. The longer the equipment stayed on the corridor, the greater the yarding distance must be.

The first example is from a skidding operation with four machine operators (Jeff, Tracy, Ron, and Mark) and two thinning alternatives (light-thin and heavy-thin). Combining the data from the four operators gave 136 days of data (Figure 4). In the combined data set, the indicator variables show which operator was involved. Each individual operator's indicator variable value is "1" on his cycles, while all of the other operators' indicator variables are "0". The following regression equation includes the variables that were significant at the 95% confidence level:

OREGON STATE UNIVERSITY
FOREST ENGINEERING DEPARTMENT
WILLAMETTE YOUNG STAND STUDY

YARDING PRODUCTION

NOTE: PLEASE START NEW FORM WHEN UNIT OR TREATMENT CHANGE OCCURS.

SALE Walk Thin
UNIT # 88
LANDING # 8
ROAD #(S) 11

HVY THIN x
LHT THIN ____
GAP CUT ____

DATE 3/4/95
START TIME 7:30 am
END TIME 4:00 pm
BREAK TIME —

****

Yarder Model Koller 501

Turns 76
Logs 146
Tops 189

Manpower (hours)

Yarder Engineer 8 1/2
Chaser #1 8 1/2
Chaser #2 ____
Hooktender 8 1/2
Rigging Slinger 8 1/2

Choker Setter #1 ____
Choker Setter #2 ____
Choker Setter #3 ____

****

Yarding DELAYS (greater than 10 minutes)
45 Minutes Problem Splice line
____ Minutes Problem ____
____ Minutes Problem ____

ROAD AND/OR LANDING CHANGES (nearest 10 minutes)
110 Minutes Problem Move yarder change roads 10 to 11
____ Minutes Problem ____
____ Minutes Problem ____

MECHANICAL DELAYS (greater than 10 minutes)
20 Minutes Problem Adjust clutch
____ Minutes Problem ____
____ Minutes Problem ____

OTHER DELAYS (greater than 10 minutes)
____ Minutes Problem ____
____ Minutes Problem ____
____ Minutes Problem ____

****

COMMENTS: Rigged to Tail Tree

Figure 3. Sample data form for a shift-level study.

$$\text{Cycles / hr} = 5.49 + (0.75 \times \text{treatment}) - (0.025 \times \text{days}) - (0.79 \times \text{Jeff}) - (0.72 \times \text{Tracy}) + (1.02 \times \text{Ron})$$

$R^2 = 29.7\%$ S.E. = 1.15

where

treatment	= 1 for heavy-thin
	= 0 for light-thin (reference condition)
days	= number of cumulative days, indicating learning curve effect, 1 to 136
Jeff	= 1 if Jeff was the machine operator
	= 0 otherwise
Tracy	= 1 if Tracy was the machine operator
	= 0 otherwise
Ron	= 1 if Ron was the machine operator
	= 0 otherwise.

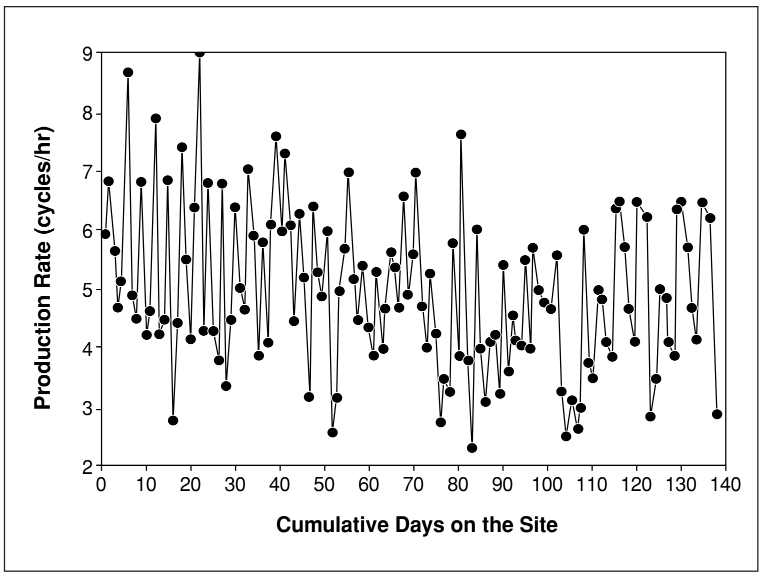

Figure 4. Production rate in a skidding operation based on a shift-level study.

In the reference condition (operator = Mark, treatment = light-thin, and days = 1), the production rate was 5.465 cycles/hr. The regression coefficient for the heavy-thin alternative indicates that the production rate was higher (0.75 cycles/hr more) for the heavy-thin than for the light-thin alternative. As an operator, Ron produced about 1.021 more cycles per hour than Mark. The other two operators, Jeff and Tracy, were even slower than Mark. For the period studied, the production data had a long-term downward trend (Figure 4), as indicated by the negative coefficient for the number of days.

Despite the large amount of variability in the data (S.E. = 1.15), the regression was able to detect operator differences, treatment differences, and a long-term trend. The long-term trend could be due to a combination of site conditions, weather, and worker attitude.

The second example is from a felling operation in which a pair of fellers, Paul and Arnold, worked under similar conditions each day. The production rates (trees per hour) of the two fellers rose and fell essentially in unison (Figure 5), which suggests that site conditions were responsible for much of the variation in production. The production rate had a long-term positive trend over the 51 days studied. The regression indicated that Paul consistently cut 5.2 more trees per hour than Arnold did. There was a standard deviation between workers of 2.4 trees/hr. The S.E. of the estimate (unexplained variation from day to day) was 5.0 trees/hr. The learning effect was an increase of 0.18 trees/hr for each consecutive day worked. The coefficient of determination (R^2) was 22.3%.

The third example is from a felling operation with three thinning alternatives (heavy-thin, light-with-openings, light-thin) and five fellers (Rod, Clint,

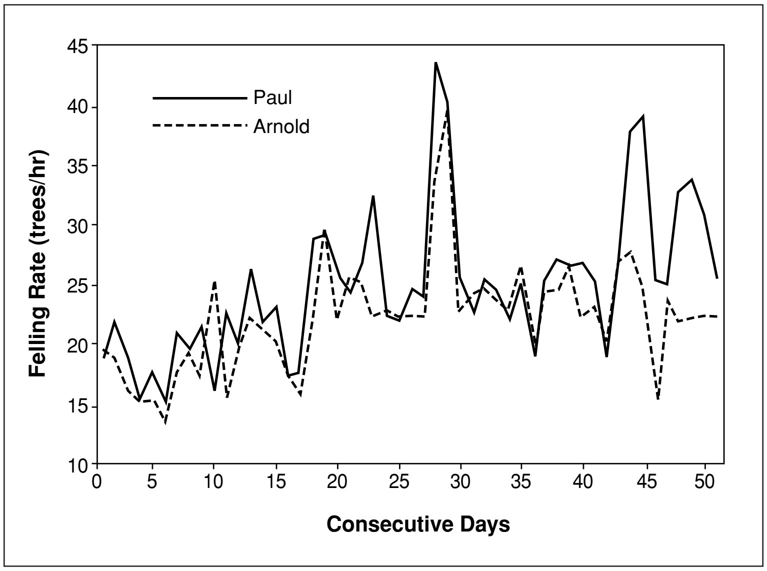

Figure 5. Variation and trend in felling rate based on a shift-level study.

Frank, Gene, and Wes). A total of 291 worker days were recorded. The regression results in Table 5 include variables significant at the 95% level (R^2 = 64.5%; S.E. = 5.0). The reference equation is for Wes as the feller and light-thin as the alternative. A variable indicating the length of the workday showed that more work hours per day had a negative impact on the hourly production rate.

By inserting the conditions on each day, values predicted from this regression equation were plotted (Figure 6). The first 35 days were light-with-openings alternative with Frank and Gene as fellers. Days 35 through 85 were heavy-thin alternative, also with Frank and Gene. Gene was about 3 trees/hr faster than Frank. A slight drop of 0.8 trees/hr occurred at day 86, when light-with-openings alternative resumed for a few days. Wes started at day 90, and he had the best production. Clint worked with Wes, but his production rate was 3.5 trees/hr slower. At day 110, work on the light-thin alternative began, and it was 2.7 trees/hr slower. The remaining variation in the graph is due to the length of the days worked, with short days getting 1.3 trees/hr better production.

These examples show the necessity of including all of the relevant variables in the regression. The goal of the regression is usually to find if a treatment effect exists. By including the operator and learning effects in the regression, it is easier to discriminate any treatment effect.

Table 5. Regression results of shift-level felling data.

Variable	Coefficient
Constant	31.06
Consecutive days	0.0864
Hours per day	–1.332
Heavy treatment[1]	3.435
Light with opening[1]	2.697
Rod[1]	–3.22
Clint[1]	–3.50
Frank[1]	–12.75
Gene[1]	–9.65

[1]Indicator variable (0 or 1 values).

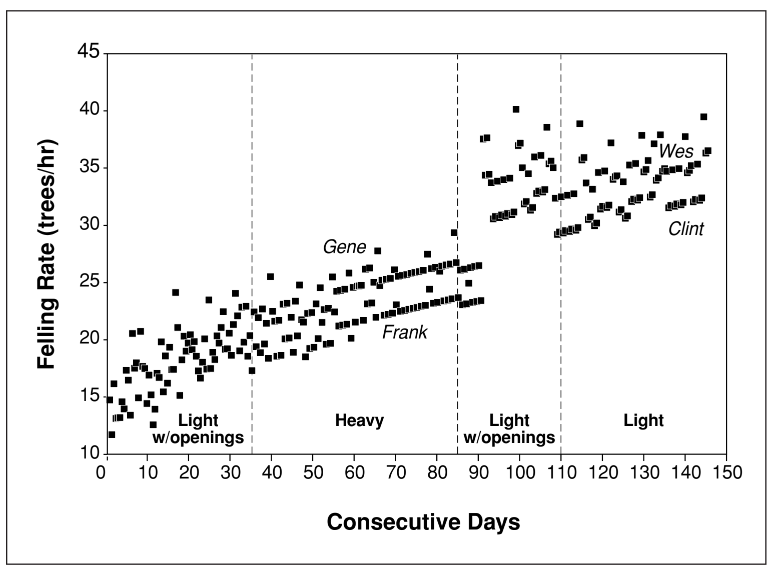

Figure 6. Shift-level regression results for felling.

Comparison of Detailed Time and Shift-Level Studies

Both detailed time and shift-level studies may be useful techniques in a production study, but these two techniques differ in labor costs, accuracy of results, and the type of information gathered. Since time and money are usually limited, researchers must decide how much effort to allocate between these methods. Neither alone can provide a complete picture. This section attempts to quantify the relative value of each technique.

Relative Advantages and Disadvantages of Each Technique

Detailed time study does an excellent job of comparing the delay-free production between alternatives. The frequent small delays (<10 min) are also documented well. Because of its limited duration, this type of study is of limited value in accurately estimating long-term trends. Other major drawbacks of detailed time studies, in addition to the cost, are associated with the limited sample size. Large delays, which occur on average only once per day, are not adequately sampled. The range of logging conditions is also limited when the study is done for only a few days.

The shift-level study, on the other hand, occurs over a longer period of time and is more likely to represent the actual range of conditions. The value of the shift-level data depends on a conscientious recorder. The recorder documents the piece count, number of cycles, large delays (≥10 min), crew size, and hours worked. Unfortunately, important logging conditions, such as yarding distance and piece size, are not usually documented. The two major defects of this technique are that it cannot account for other important variables and that small delays are confounded with the productive time.

In a detailed time study, trained research technicians collect the data, which have high precision (within 1 sec in most cases). In a shift-level study, a crew member records the data at the end of each day. The precision of these data (0.25 hr at best) affects the accuracy of the average production rate.

The ideal combination is to use the shift-level data to record large delays and time-related changes, such as long-term production improvements. The detailed time data can then be used to show the difference in delay-free production under standardized conditions.

Predictive Capabilities of Each Method

A detailed time study and a shift-level study were conducted in a yarding operation in order to contrast the predictive ability of the two methods. Data were collected simultaneously over parts of a 9-month period (March to November 1995). The harvesting method was skyline thinning of Douglas-fir, and there were three logging alternatives: light-thin, light-with-openings, and heavy-thin. Data for the shift-level study were collected for 78 days, and data for the detailed time study were for 704 cycles over parts of 15 different days.

Regression analysis was conducted for both studies; data sets for all treatments were combined, and alternatives were represented by indicator variables. Final regression equations for predicting production rates (cycles per hour) are reported below. In the detailed time study, the percentage of small delays was estimated so that results can be contrasted fairly with those from the shift-level study. Daily estimates from the 15 days in common to both studies were subjected to a paired t-test to determine if the studies estimated significantly different means. S.E. values are reported to demonstrate the variability in the estimates.

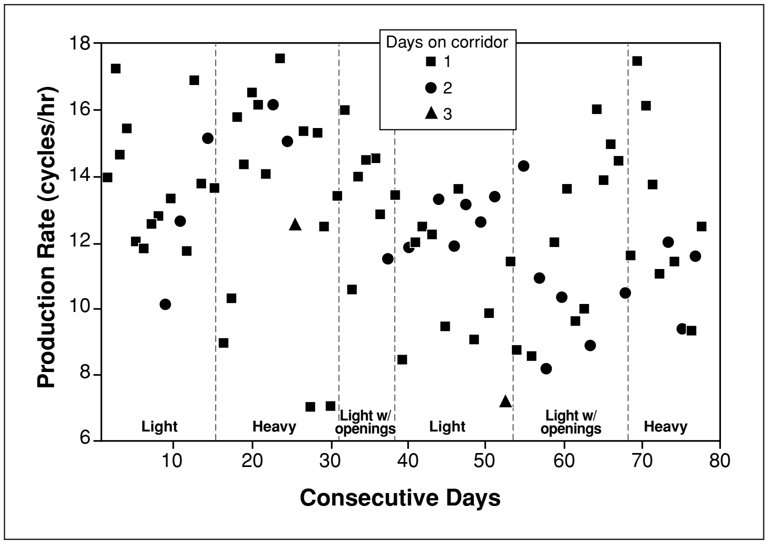

Figure 7. Observed production rate from a shift-level study of three alternatives (light, heavy, light with openings).

Results of Shift-Level Study

The daily average production rate (cycles per hour) for the study period appears to have a random pattern (Figure 7).

As shown in Table 6, average production during two replications of each treatment was calculated. The standard error of the average daily production was calculated by dividing the standard deviation by the square root of the sample size. The first three average production rates are not significantly different. Then a major downward shift of about 2 cycles/hr occurs. The remaining three averages are also not significantly different. Regression analysis can be used to see if any finer resolution is possible.

The final regression equation for the shift-level data was

$$Y = 14.32 - (0.0337 \times learn\_days) - (0.703 \times yrd\_day) + (1.021 \times extra\_crew)$$

$$N = 78 \text{ days} \qquad R^2 = 16.3\% \qquad \text{S.E. of Y estimate} = 2.48$$

where

Y = production rate (cycles/hr)

learn_days = 1 to 78, representing each day of the study

yrd_day = the number of consecutive work days on the corridor

extra_crew = 0 (no extra crew) or 1 (one additional crew member).

The yrd_day variable indicates that for each additional day on the corridor, production dropped 0.703 cycles/hr. This was the anticipated effect of longer yarding distances on corridors that took 2 or more days to complete. None of the indicator variables were significant, meaning no difference in production among the alternatives was found. The standard error of the mean was $2.48/\sqrt{78} = 0.2808$ cycles/hr, so any difference between alternatives in that range should have been detectable.

Table 6. Observed daily production rates from a shift-level study of a yarding operation.

Dates (1995)	Treatment	N (days)	Average daily production[1] (cycles/hr)	Standard deviation	Standard error
3/20–4/6	Light	15	13.62 A	1.99	0.51
8/2–8/25	Heavy	16	13.45 A	3.37	0.84
8/26–9/1	Light with openings	7	13.44 A	1.90	0.72
9/7–9/25	Light	15	11.37 B	2.03	0.52
10/6–10/26	Light with openings	16	11.56 B	2.62	0.66
10/27–11/9	Heavy	9	12.46 B	2.83	0.94

[1]Values with the same letter are not statistically significantly different.

Large delays and road changing times have been removed from the data set. Alternative (light-thin, light-with-openings, and heavy-thin), average pieces per cycle, and length of workday were not significant variables.

Figure 8 shows how the regression equation clarifies production once it has removed the random variation shown in Figure 7. Figure 8 is a plot of daily production based on the regression equation. The downward trend over the course of the study could be weather related or could reflect a change in the crew members or some other time-related effect. The top line of data begins when another crewmember was added. The lower lines of data show the decrease in production that occurred with increased yarding distance.

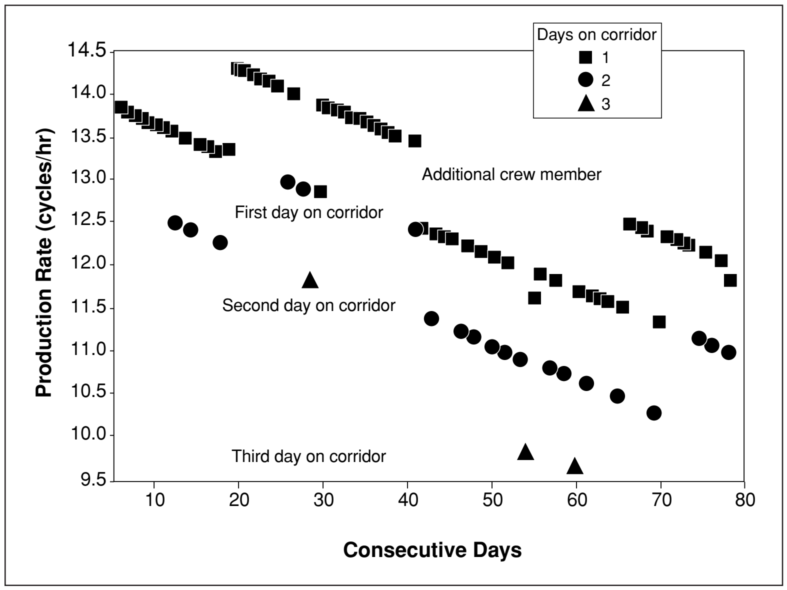

Figure 8. Shift-level regression results showing indicator variable effect.

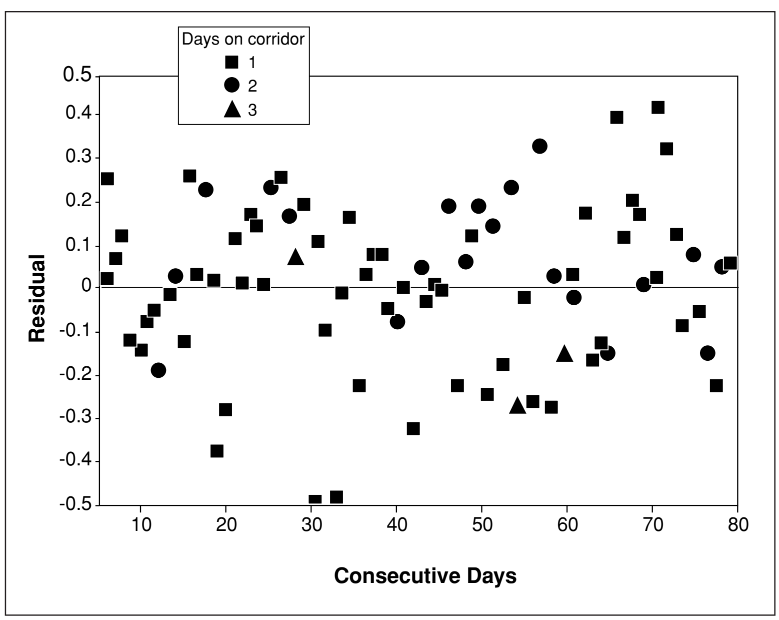

Figure 9. Residual plot of shift-level regression results.

Figure 9 shows the differences (residuals) in daily production between what was observed and what was predicted by the regression. There is no noticeable pattern, such as several consecutive points on the same side of the zero line, that would indicate that an additional variable is needed in the regression.

Results of Detailed Time Study

Data from the 704 cycles (portions of 15 days) were combined and regressed on a delay-free basis. The following regression equation includes the significant variables (t > 1.96) for this data set:

$$\begin{aligned} Y = {} & 25.79 + (1.165 \times \text{opening}) - (0.031 \times \text{slope}) - (0.011 \times \text{yarding\_distance}) \\ & - (0.065 \times \text{carr\_height}) - (0.0401 \times \text{pieces}) - (0.028 \times \text{lateral\_distance}) \\ & + (1.581 \times \text{preset}) + (0.572 \times \text{extra\_crew}) \end{aligned}$$

$$N = 704 \text{ cycles} \qquad R^2 = 54.5\% \qquad \text{S.E. of Y estimate} = 1.900$$

where

opening	=	0 (thinning) or 1 (clearcut opening)
slope	=	% of ground slope at hook site
yarding_distance	=	slope distance (ft) traveled by carriage
carr_height	=	height (ft) of carriage above ground at hook site
pieces	=	number of pieces yarded each cycle
lateral_distance	=	distance (ft) perpendicular from the skyline to the choked turn
preset	=	0 (no preset) or 1 (preset chokers).

Intermediate supports and indicator variables for alternatives were not significant variables. Small delays represented 16% of total cycle time. The data were collected over a relatively short period, so no time-related trend was observed.

Comparison of Techniques

Table 7 contains the observed and predicted average daily production rates for each of the 15 days when both detailed time and shift-level data were collected. The average observed conditions for the day were used in the regression equation to predict average daily production. The average shift-level production was 76% of the average detailed study production, which means that small delays occurred 24% of the time; the detailed time study estimated that small delays represented 16% of total cycle time. This discrepancy is rather serious, but the cause is not clear. The sampling period may have been too short to accurately estimate the percentage of small delays, or the shift-level data may contain an error, such as an incorrect daily tally of pieces. For this example, we assumed that 24% is correct, and we adjusted the daily values from the detailed study so that the overall average was the same for both studies.

After making the adjustment for small delays, there are still several reasons to accept the detailed time study results instead of the shift-level results. When the daily predicted value of the detailed time study data is subtracted from the observed value, this daily difference has a standard deviation of 0.38

Table 7. Observed and predicted (via regression) daily average production rates (cycles/hr) from detailed time and shift-level studies.

Date	Treatment	Detailed time[1]		Shift-level	
		Observed	Predicted	Observed	Predicted
9/18	Light	12.66	12.89	13.14	11.06
9/19	Light	10.45	10.19	9.09	12.08
9/20	Light	11.94	12.08	12.61	10.99
9/21	Light	12.09	11.99	9.87	12.01
Average	Light	11.78	11.79	11.18	11.53
10/16	Light with openings	10.57	10.35	13.65	11.64
10/17	Light with openings	14.17	14.01	9.63	11.61
10/19	Light with openings	10.78	10.47	8.86	10.48
10/23	Light with openings	13.25	13.29	13.92	12.48
10/24	Light with openings	11.88	11.70	14.92	12.45
10/25	Light with openings	12.05	11.82	14.50	12.41
10/26	Light with openings	11.71	11.77	10.45	10.29
Average	Light with openings	12.06	11.91	12.28	11.62
11/2	Heavy	10.60	11.52	12.00	11.13
11/3	Heavy	13.14	12.28	11.43	12.16
11/7	Heavy	7.99	8.08	9.32	12.08
11/8	Heavy	11.29	11.38	11.59	10.99
Average	Heavy	10.76	10.81	11.08	11.59
Grand average	All	11.64	11.59	11.66	11.59

[1]Adjusted to reflect the 24% difference in cycles/hr attributed to small delays.

cycles/hr. With the shift-level data, the daily difference between predicted and observed is 1.94 cycles/hr. This shows the higher variability in the shift-level data.

On a daily basis, the difference between detailed time and shift-level studies had a standard deviation of 2.16 cycles/hr, which seems large given that the average production was 11.64 cycles/hr.

The detailed time regression equation included eight significant variables whereas the shift-level equation included only three. The detailed time regression explained 54.5% of the variation, compared to 16.3% explained by the shift-level regression.

The plot of residuals comparing the daily average production (cycles per hour) of the observed detailed time data with the estimated average from the detailed time regression analysis appeared normally distributed, as well as free from unwanted time series.

Overall, the shift-level estimates of production are less reliable than those from the detailed time study. The detailed time study was able to detect a significant production increase in the clearcut openings. The shift-level data could not be partitioned to test the opening effect independently. Both methods agreed that there were no other production differences among treatments.

Confidence Intervals

Another way to visually demonstrate the superior resolution of one method over another is to compare *confidence intervals* (C.I.'s). C.I.'s are wide when there is a large amount of unwanted variability. After only a few days the limits for the detailed time study converge to a narrow band, whereas it takes weeks before the limits of the shift-level study converge significantly. C.I.'s around the average number of cycles per hour for the detailed time and shift-level data can be compared to show the relative accuracy of each. The daily standard deviation (S.D.) for the detailed time data is derived from the S.D. per cycle, which is called the S.E. of the Y estimate in the regression output.

If we assume 100 cycles in a day and an S.D. per cycle of 1.9, then

$$\text{S.D. daily} = (\text{S.D. per cycle}) / \sqrt{\text{cycles in one day}}$$

$$1.9 / \sqrt{100} = 0.19$$

The S.D. for the shift-level regression (2.54 cycles/hr) was given by the output. The 95% confidence limits are given by

$$\text{Upper / Lower Limit} = \text{mean} \pm 1.96 \times (\text{S.D. daily}) / \sqrt{\text{days in study}}$$

The C.I.'s (Figure 10) show how long each study would have to be conducted to reduce the C.I. to an acceptable bandwidth. For example, to attain a width of ±0.2 cycles/hr, a detailed time study would be run for 4 days (assuming 100 cycles/day); a shift-level study could achieve this level only after 620 working days. Figure 11 is a plot of the shift-level cumulative average production for the last half of the study, when the learning curve was stable. This figure demonstrates how the cumulative average for yarding production approaches its long-run mean slowly over time, just as the C.I. estimations predicted.

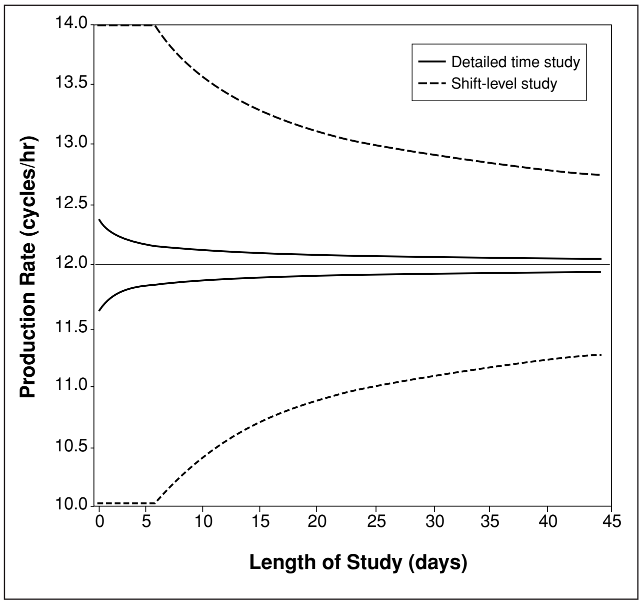

Figure 10. Confidence intervals for production rates from a detailed time study and a shift-level study.

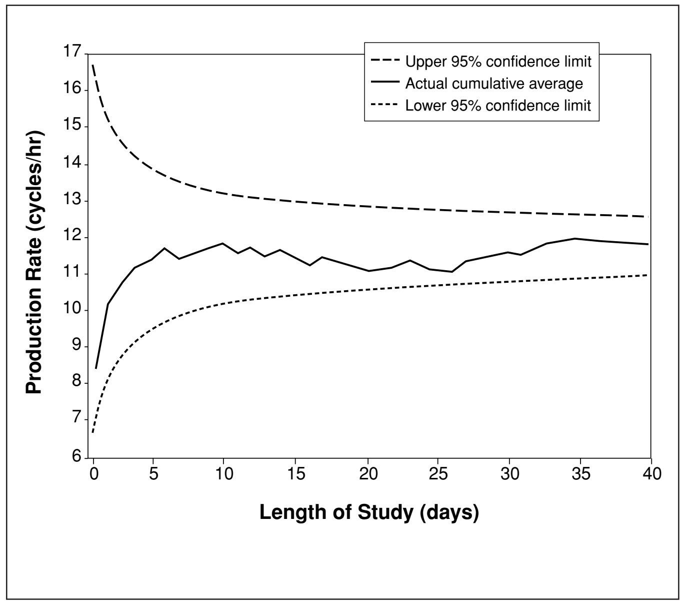

Figure 11. Cumulative average and 95% confidence interval for production rate based on a shift-level study.

Conclusions

Neither detailed time nor shift-level studies indicated that thinning alternatives had any effect, the determination of which was a main objective of the investigation. The two approaches also agreed on the effects of an extra crew member and increasing yarding distance. The shift-level study detected a learning curve effect over time that the detailed time study did not detect. The detailed time regression quantified the impacts of a greater number of significant variables than was possible in the shift-level analysis.

The C.I.'s on the shift-level data are an order of magnitude (10 times) wider than the detailed time study C.I.'s. Thus if alternative effects had been present, the detailed time study would have been more likely to be able to identify them. It appears that there is a distinct need for both methods, because each can quantify effects that the other cannot.

Determining Appropriate Length for Field Studies of Harvesting Operations

A common question among researchers conducting field studies is "How many days should the study last?" In the past, only crude approximations or rules of thumb were used to address this question. However, with the statistical analyses described in the prior sections, researchers can more rigorously determine study length. Indicator variables can be used to represent alternatives, and then the relationship can be established between the study length and the size of detectable treatment difference.

This section develops a procedure to determine the appropriate length for both shift-level and detailed time studies of harvesting operations. Statistically, an adequate sample size is influenced by the variation in the data, the confidence level desired, and the allowable error. The assumed objective of the field studies is to determine whether the average daily production rates of two harvesting alternatives are statistically significantly different.

We measured average daily production rates in the same units (cycles per hour) for the shift-level study and the detailed time study so that their relative accuracies could be compared easily. For the shift-level data, each day was a sample point, with one estimate of production for that day. For the detailed time study, the daily production rate was calculated from the approximately 100 cycles that could be measured each day.

We used regression to process the data. An indicator variable identified which alternative was present on each day: 0 = the reference (light-thin); 1 = the other alternative (heavy-thin). The regression coefficient for the indicator variable is the difference (cycles per hour) between the two alternatives. The longer the study, the more discriminating the tests are at detecting differences.

How long a study is needed is influenced by the amount of unexplained variation that is occurring in the production. This variation can be estimated a priori by finding the S.E. of a small sample.

With a detailed time study, a large percentage of the variation (usually 50% or more) can be explained with a regression equation. The coefficient of determination, R^2, is a measure of the percentage of variation that has been explained by the regression equation. The remaining variation ($1 - R^2$) is unexplained (random, unmeasured, or uncontrolled). This unexplained variation is also reported as the S.E. of the Y estimate in most statistical programs, and it can be used in calculating the appropriate study length.

For a shift-level study, little or none of the variation (25% or less) is explained via regression. In addition, small delays are mixed in with the daily production. These two factors can make the daily variation in the shift-level data an order of magnitude larger than in detailed time studies.

To develop a procedure for determining the sample size, we will first present the relevant statistical theory. Next the application of the theory to production studies will be demonstrated. Then actual field study data will be used as an example.

Theory

The regression output includes the indicator variable coefficient, B1, which represents the difference between the two alternatives. A t-test can be applied to see if this difference is statistically significant:

$$t = B1 / SB1$$

where SB1 is the standard deviation of B1. The t-value must be greater than 1.96 to be significant at the 95% confidence level. Therefore the minimum detectable difference between the treatment means will be

$$B1_{min} = t \times SB1 = 1.96 \times SB1$$

The S.D. for an indicator variable is

$$SB1 = \sqrt{MSE / \sum\left(X_i - \overline{X}\right)^2} = \sqrt{MSE} / \sqrt{\sum\left(X_i - \overline{X}\right)^2}$$

where

$$\begin{aligned} MSE &= \text{mean square error} \\ \sqrt{MSE} &= \text{S.E. of the Y estimate} \\ X_i &= 1 \text{ or } 0 \\ \overline{X} &= \text{average} \end{aligned}$$

S.E. will be used to denote $\sqrt{MSE}$:

$$SB1 = S.E./\sqrt{\sum\left(X_i - \overline{X}\right)^2}$$

For an indicator variable that has half the sample in each of the alternatives, the value of $\overline{X}$ is 0.5. The indicator variable X_i is either 1 or 0. The number of samples, N, can then be introduced to the equation:

$$\sum\left(X_i - 0.5\right)^2 = N \times (0.5)^2 = N / 4$$

Thus, the equation for SB1 can be rewritten as

$$SB1 = S.E./\sqrt{N/4} = S.E. \times 2 / \sqrt{N}$$

The minimum difference between treatments can be calculated as a function of N:

$$B1_{min} = 1.96 \times S.E. \times 2 / \sqrt{N}$$

Solving for N shows how the result depends on the minimum allowable error $B1_{min}$ (expressed in cycles per hour) and the unexplained variation S.E. (also expressed in cycles per hour).

$$N = \left(1.96 \times S.E. \times 2 / B1_{min}\right)^2$$

Since the detailed time regression is on the basis of cycles, the daily S.E. is calculated based on the number of cycles per day. The daily S.E. for the detailed regression is

$$\text{daily S.E.} = (\text{cycle S.E.}) / \sqrt{\text{cycles per day}}$$

For 100 cycles/day,

$$\text{daily S.E.} = (\text{cycle S.E.}) / \sqrt{100} = 0.1 \times (\text{cycle S.E.})$$

Application

For the detailed time example reported in the previous section, the yarding applications that we have studied had a cycle S.E. of about 2.0 with 100 cycles/day, which translates into a daily S.E. of 0.2 cycles/hr. A 4-day detailed time study is then able to detect differences of

$$B1_{min} = 1.96 \times 0.2 \times 2 / \sqrt{4} = 0.392 \text{ cycles / hr}$$

Thus, if the difference in production rates of two harvesting alternatives is greater than 0.392 cycles/hr, the 4-day study will be able to detect it. During the 4-day study, half of the 400 cycles observed should be from each alternative. If the difference between the alternatives is less than 0.392 cycles/hr, it will take a study longer than 4 days to detect it.

The measured S.E. for shift-level data is about 2.5 cycles/hr. To compare the two study types, the length of a shift-level study needed to detect the same minimum difference can be calculated:

$$B1_{min} = 1.96 \times 2.5 \times 2 / \sqrt{N}$$

Inserting 0.392 cycles/hr as the minimum and solving for N yields the following:

$$N = (1.96 \times 2.5 \times 2 / 0.392)^2 = 625 \text{ days}$$

Thus a shift-level study would have to be about 150 times longer than a detailed time study to detect the same difference in alternatives. Since most shift-level studies are only a few months long at most, the minimum difference that can be detected will be larger.

A typical shift-level study might last for 60 days (equally divided between the two alternatives), and the minimum difference that could be detected would be

$$B1_{min} = 1.96 \times 2.5 \times 2 / \sqrt{60} = 1.265 \text{ cycles / hr}$$

This analysis was for an operation with about 12.5 cycles/hr and indicates that a 60-day shift-level study could detect 10% (or greater) differences between the treatments.

The detailed time study compares the treatments as delay-free operations. The shift-level data, on the other hand, contain the small delays, which is one of the reasons the S.E. is larger.

Comparison of Shift-Level and Detailed Time Study

If both the detailed time study and the shift-level study predict cycles per hour, the relative sample sizes needed by each can be compared. To demonstrate the differences, data from a 28-day shift-level study will be used. From regression, we found the coefficient of the indicator variable (B1) is 0.4721 cycles/hr, which is the difference between the production of the light-thin alternative and that of the light-with-openings alternative. The regression output gives S.E.= 2.2302. A t-test can determine whether B1 is statistically different from 0:

$$SB1 = S.E. \times 2 / \sqrt{N}$$

$$SB1 = 2.2302 \times 2 / \sqrt{28} = 0.8451$$

and

$$t = B1 / SB1$$

$$t = 0.4721 / 0.8451 = 0.558$$

The calculated t-value of 0.558 is far less than the required value of 1.96. Therefore, the coefficient of the indicator variable (0.4721 cycles/hr) is not statistically different from 0, and no difference between the alternatives has been established. The difference between alternatives would have to be almost 4 times as large as this before a 28-day shift-level study could detect it.

Calculating Study Length

The previous discussion centered on making calculations after the study was completed. The same theory can be used to find an appropriate study length in advance. To do so, the S.E. must be estimated from field data that are collected either in a small sample or during previous similar field studies and then subjected to regression analysis as shown above. In lieu of regression, the standard deviation of the daily production can be used to estimate S.E., although the value will be overly large.

The choice of the minimum detectable difference, B1, is dictated by the end use of the analysis. A common value might be 10% of the average hourly production. Thus, if the average production was 15 cycles/hr, it would be reasonable for us to expect to detect differences between treatments that exceeded 1.5 cycles/hr.

This process can be done mathematically:

$$N = \left(1.96 \times S.E. \times 2 / B1_{min} \right)^2$$

As an example, let S.E. = 2.5 for a shift-level study. If the minimum difference ($B1_{min}$) = 1.5 cycles/hr, then N = 43 days. Figure 12 shows the relationship of S.E., B1, and N graphically.

Detecting Differences between Treatment Means

The purpose of many studies is to determine if the production of one harvesting treatment is larger than another treatment. If the average production of one treatment is much larger than the average production of another, then

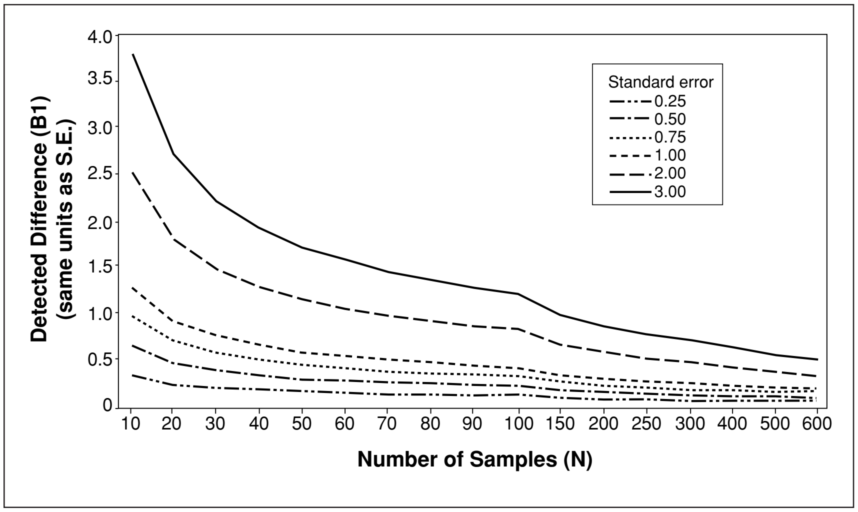

Figure 12. Required cycles for a time study.

a small sample will be able to detect that a difference exists. However, when only a small difference exists then the sample size must be large in order to detect statistically significant differences between the treatments. Detailed time study and shift-level data from a thinning study were regressed to determine if there was a difference in the average number of cycles per hour between the heavy-thin and light-thin alternatives. A summary of the results is shown in Table 8.

Because of the large number of observations, small differences in the alternative means could be detected. The shift-level and detailed time study regressions reported almost identical results: the light-thin alternative produced 2.1 fewer cycles per hour than the heavy-thin alternative. Either of the study methods is capable of concluding that alternative differences of this magnitude are statistically significant. Indeed, a difference of this magnitude could be detected with much smaller sample sizes.

Table 8. Detecting treatment differences with detailed time and shift-level studies.

Statistic	Detailed time study	Shift-level study
R^2 (%)	52.0	30.3
S.E. of Y estimate (cycles/hr)	2.00	2.02
No. of observations	960[1]	118[2]
Min. detectable difference (cycles/hr)	0.426	0.721
Difference in average (cycles/hr)	2.08	2.14

[1]Cycles.
[2]Days.

Delay Analysis Requirements

Use of a proper study length is also important for delay estimates. To get adequate estimates of the percentage of time spent in small delays, a detailed time study must be run for several days (Hossain 1998). Fortunately, these study lengths closely coincide with the sample size needed for the delay-free cycle time.

The percentages of time spent in large delays are best measured with the shift-level data. Even with long studies, the C.I.'s on the large-delay percentage are quite wide (Hossain 1998). From a practical point of view, the shift-level study should be run for as long as the unit is being harvested.

Conclusion

The length of study can be calculated a priori by using an estimate of the S.E. If regression with indicator variables is being used, the formula for the sample size reduces to the following (assuming a 95% confidence level):

$$N = (1.96 \times S.E. \times 2/B1_{min})^2$$

where

N = number of days

S.E. = daily standard error of Y estimate

$B1_{min}$ = minimum detectable difference (cycles/hr) between the treatments.

Because of the relatively large S.E. of shift-level data, a shift-level study must be run 150 times as long as a detailed time study to detect the same difference.

In choosing the length of a study, the cost of additional days of study must be weighed against the statistical improvement. If the error is allowed to be twice as large, the sample size would be reduced by a factor of 4. Conversely, if half as much error is allowed, data for 4 times as many days would be needed.

Summary

The main findings of the report are as follows:

- There is an equal amount of variation coming from the measurable variables, from the delays, and from the residual variation in the Y estimate. To the extent these sources can be isolated and measured, the production estimates will be statistically improved.
- Regression with indicator variables is equal or superior to regular regression analysis. The indicator variable approach makes comparisons relatively simple because the computer output can be used directly to test whether treatment means are different.
- Regression of shift-level data can show long-term trends. Indicator variables can be used effectively to identify treatments and binary variables such as equipment operators.
- In general, the detailed time study is more discriminating than the shift-level study. It can detect a smaller difference between means for alternatives than a shift-level study can detect.
- Comparison of confidence intervals is an effective way to demonstrate the effect of larger sample sizes. Even with studies lasting months, the shift-level confidence interval does not converge to the smaller magnitude obtained from several days of detailed time study data.
- Both detailed time and shift-level regression analyses are necessary to obtain an accurate pictures of production. Detailed time studies give accurate delay-free production estimates with narrow confidence intervals. Shift-level data can show long-term trends and more accurately reflect the infrequent large delays.

Bibliography

Appelroth, S-E. 1988. Classic and Nordic forest work-studies. IUFRO International Symposium on Developments on Work-Studies in Forestry, IUFRO WP 3.04.02. 11 p.

Bjorhedeh, R. 1988. New work-study methods help to decide processing technique in logging. Scandanavian Journal of Forest Research 3:569–574.

Cottell, PL, and HI Winer. 1969. Alternative methods for evaluating the productivity of logging operations: implications of a study of wheeled skidding. Pulp and Paper Research Institute of Canada, Vancouver, B.C. Technical Report TR-8. 48 p.

Evanson, AW, and M Kimberley. 1992. An analysis of shift-level data from six cable logging operations. New Zealand Forest Research Institute Ltd., Rotorua. Forest Research Institute Bulletin 174. 24 p.

Folkema, MP, P Giguere, and E Heidersdorf. 1981. Shift level availability and productivity: revised manual for collecting and reporting field data. Forest Engineering Research Institute of Canada, Pointe Claire, Quebec. 13 p.

Freese, F. 1967. Elementary Statistical Methods for Foresters. Agriculture Handbook 317. US Department of Agriculture. 87 p.

Hossain, MM. 1998. Young stand thinning in western Oregon: Cost comparison of harvesting alternatives and comparison of time study techniques. Ph.D. Dissertation, Forest Engineering Department, Oregon State University, Corvallis.

Howard, AF. 1992. Validating forest harvesting production equations. Transactions of the ASAE 35:1683–1687.

Matzka, PJ. 1997. Harvest system selection and design for damage reduction in noble fir stands (a case study on the Warm Springs Indian Reservation). M.S. Thesis. Forest Engineering Department, Oregon State University, Corvallis. 129 p.

Murphy, GE, and AW Evanson. 1992. Development of machine and system-based production models for steep terrain. P. 174–185 *in* IUFRO Proceedings of Computer Supported Planning of Roads and Harvesting Workshop, Feldafing, Germany.

Neter, J, W Wasserman, and MH Kutner. 1996. Applied Linear Statistical Models. 4th ed. Richard D. Irwin Co., Chicago. 1408 p.

Olsen, ED. 1988. Logging incentives systems. Forest Research Laboratory, Oregon State University, Corvallis. Research Bulletin 62. 19 p.

Olsen, ED, and LD Kellogg. 1983. Comparison of time-study techniques for evaluating logging production. Transactions of the ASAE 26:1665–1669.

Ortiz, D, E Olsen, and L Kellogg. 1997. Simulating a harvester-forwarder softwood thinning: a software evaluation. Forest Products Journal 47(5):36–41.

Rickards, J, R Skarr, S Haberle, K Apel, R Bjorheden, and MA Thompson. 1995. Forest work-study nomenclature. Swedish University of Agricultural Science, Garpenberg, Sweden. 16 p.

Riggs, J, D Bedworth, and S Randhawa. 1996. Engineering Economics. 4th ed. McGraw Hill, New York. 539 p.

Samset, I. 1990. Some observations on time and performance studies in forestry. Erfaringer angaende tids- og prestasjonsstudier I skogbruket. Moddelelser fra Norsk Institutt for Skogforskning 43(5):1–80.

Sondell, J. 1987. Better conditions for machine operators. Skogsarbeten No. 2. 4 p.

Thompson, MA. 1988. An analysis of methods used to report machine performance. American Society of Agricultural Engineers, St. Joseph, Michigan. ASAE Paper 88-7547. 56 p.

Winer, HI. 1980. Organizational factors affecting trials of new logging machines. Forest Engineering Research Institute of Canada, Pointe Claire, Quebec. FERIC Special Report SR-10. 18 p.

Appendix A: Simplified Explanation of Statistical Terms

The following definitions were adapted from Freese (1967).

Coefficient of determination: As a measure of how well a regression fits the sample data, we can compute the proportion of the total variation in Y that is associated with the regression. It is usually called the R^2 value.

Confidence interval (C.I.): The statistical way of indicating the reliability of an estimate is to establish confidence limits. For estimates from samples over 30 ($N > 30$) taken from normally distributed populations, the 95% confidence limits are given by

$$(\text{Estimate}) \pm 1.96 \times (\text{Standard Error})$$

This means that 95% of the time, limits set in this manner will contain the true population value.

Error term: This is also called the residual. It is the difference between the actual data value collected and the value predicted by the regression equation.

F-test: The F-test is a more sophisticated form of hypothesis testing than a t-test. An F-value of 4.0 is roughly equivalent to a t-value of 1.96.

Indicator variable: This is a qualitative variable that has a value of either 0 or 1. It is used to designate whether a given condition exists. The treatment effects can be treated as indicator variables. These are also called dummy variables or binary variables.

Multiple least-squared linear regression: A regression equation can be thought of as a moving average. It gives an average value of Y associated with a particular set of multiple X values. A least-squared regression minimizes the sum of the squared residuals. In linear regression, the relationship between the X values and the Y value is assumed to be linear. Inclusion of the term "multiple" means that more than one independent variable was used.

Standard deviation: The standard deviation characterizes dispersion of individuals about the mean. It gives us some idea whether most of the individuals in a population are close to the mean or spread out.

Standard error (S.E.) of the estimate: This is the standard deviation of the residuals. It is the amount of unexplained variation that remains after the regression equation has tried to fit the data. It is the square root of the mean squared error (MSE).

Standard error of the mean: This is the variation among means. It can be thought of as a standard deviation among sample means. It is a measure of the variation among sample means just as the standard deviation is a measure of variation within individual samples. The standard error is reduced roughly in half each time the sample size quadruples.

Student's t-test: Hypothesis testing checks whether a sample statistic such as a mean is statistically different from a given value. The t-test can also be used to test if two sample means are significantly different. For large samples the 95% confidence value of t is 1.96.

Appendix B: Description of Logging Sites

Thinning treatment	Logging system	Area (ac)	Volume (MBF)
Heavy	Tractor	29	192
Heavy	Skyline	19	128
Light	Tractor	33	320
Light	Skyline	60	580
Light with openings	Tractor	4	30
Light with openings	Skyline	32	270

Appendix C: Conducting Regression Analysis of Harvesting Production

This appendix describes how regression analysis is done on a logging operation. These procedures represent the typical steps taken during research on logging system productivity. Regression model building, indicator variables, and hypothesis testing have been discussed by Neter et al. (1996). The issue of candidates of independent variables in building a regression model and selecting the form of the equation was discussed by Samset (1990). The validation of regression models was discussed by Howard (1992).

In a research project, the production rate (cycles per hour) is one component of the cost estimate of a logging operation. This section outlines how production is estimated, assuming that a detailed time study will be the main source of information.

The major cost centers should be identified first. These are usually 1) planning and unit layout; 2) felling and bucking; 3) extraction, stump to landing; 4) loading and hauling; and 5) site preparation and reforestation. This discussion uses extraction as an example; the other cost centers would be similar.

The two main components of production are

- Delay-free cycle time
- Delay percentages.

Figure D-1 in Appendix D shows how these components are combined to calculate logging production. Each component significantly affects the final result, so equal attention should be paid to each in determining appropriate values. The resulting production will be valid only for a specific set of harvesting conditions.

Documentation of Conditions

Standard forms, such as those published by Thompson (1988), should be used to document each of the following field conditions:

- Date, weather condition, skyline corridor I.D.
- Logging procedures (e.g., thinning, presetting chokers, and intermediate supports)
- Equipment and crew
- Stand conditions (species, age, and distribution)
- Terrain (e.g., slope, brushiness, and wet areas)
- Unit layout
- Scheduled machine hour (SMH) per day.

Delay-Free Cycle Time

Multiple regression can be used to predict the cycle time (minutes per cycle). The independent variables observed during the study must be summarized. In a skyline yarding operation, for example, these variables would include

- Slope extraction distance
- Lateral extraction distance
- Pieces per cycle
- Classification of pieces (size, end use)
- Slope, intermediate supports, carriage height
- Number of chokers, preset chokers
- Number of crewmembers
- Operating in special zone (e.g., riparian area).

These variables should be recorded on each cycle. The average, standard deviation, maximum and minimum, and frequency of occurrence of each should be calculated. If the data are to be used for simulation, then their distribution must also be determined. For instance, piece size usually fits a log-normal distribution.

Regression analyses of detailed time studies are able to quantify the impact of each of these variables. In this manner their influence can be removed to detect production differences between the thinning alternatives. Since shift-level data traditionally do not or cannot measure these variables, their effect is confounded with the production estimate in shift-level studies. This creates a large S.E. that hinders the discrimination of any difference between thinning alternatives.

Sample Observations

The sample should represent as wide a range of conditions as possible. The choice of the logging unit and the number of days available will help determine how much can be accomplished. The other consideration in deciding on sample size involves estimating an average value for minutes per cycle, which has a small C.I. This is accomplished by reducing unwanted variation in the cycle time and by taking a large sample, usually more than 100 observations. The unwanted variation in the cycle time is caused mostly by random delays. These are identified, measured, and removed from the production data set before regression.

Elements of the Yarding Cycle

A yarding cycle can be desegregated into repetitive elements, usually the following:

- Outhaul
- Drop of chokers from carriage
- Lateral outhaul
- Hook
- Lateral inhaul
- Inhaul
- Unhook.

In screening and debugging the data, scatter diagrams are made of each of these elements. In the regression equations, the dependent variable, total cycle time, is the sum of these elements.

Indicator Variables

Indicator variables, also called binary qualitative variables (Samset 1990), can be used in multiple regression to identify qualitative conditions. Such variables can be used to distinguish between two conditions by using 0 to indicate the reference condition and 1 to indicate the alternative. For example, an indicator variable could be used to distinguish which of two loggers was the yarder operator: 0 = Jim (the reference logger in this case); 1 = Joe. In like manner, harvesting alternatives can be distinguished.

Data Collection

The following information is recorded and stored on each cycle:

- Length of time in centiminutes of each element of the cycle
- Length of each delay along with an identification code
- Value of each quantitative independent variable (e.g., number of pieces)
- Value (0 or 1) of each indicator variable for each qualitative condition.

Validating Regression Equations

The goal of regression is to obtain an equation that represents the general production of the system being observed. Several forms of regression equation can be fit with the same data set: linear, interaction, power, logarithmic, or a combination. Overfitting of data can occur if a researcher experiments with many different regression models with all possible combinations of terms and picks the one with the best fit.

The dependent variable in detailed time study regression is usually cycle time in minutes. In shift-level data, because of the way the data are collected, the dependent variable is usually cycles per hour. When detailed time studies are compared with shift-level studies, it is convenient to use cycles per hour as the basis of comparison. Once a regression equation has been chosen, an independent set of data should be used to validate the predicted results of the equations. Two approaches can be taken.

The first approach is to extract and reserve a portion of the original data set for validation purposes (Howard 1992). The sample must be large enough, at least 30 cycles, to allow a paired t-test to be performed. The regression is then done without these data, and the resulting regression equation is used to predict the cycle time on each of the reserved cycles, given the conditions on the cycles, such as yarding distance. The predicted cycle time is then paired with the actual observed time for a paired t-test, which will indicate if the equation is a good predictor of the observed times.

The second approach follows a similar procedure, but applies the regression equation to a new logging corridor. The researcher must be extremely careful that conditions on the new corridor are similar to those used in the regression formation.

Steps in Analysis of Field Data

The field data should be downloaded regularly (daily if possible) into a spreadsheet. At least weekly, the spreadsheet data should be checked to see if values appear reasonable. The following steps are then used to perform the regression analysis.

Screening for Outliers

An outlier is a data point that is more than three standard deviations from the mean in the Y direction. This deviation is usually on the positive side and is caused by a delay that was not separated properly from the data. The presence of even a single outlier value can seriously distort the regression results. Therefore, the data should be screened for outliers. This step isolates any entry errors and any delays that were inadvertently missed. Any outliers not located at this step in the analysis will be found on the final residual analysis step, and the whole regression will have to be redone after the outliers are removed. The following steps are used to screen data:

1. Print out the spreadsheet and scan the rows for 1) missing element times, 2) missing variable values, and 3) values that are unusually small or large compared with the other data. The missing values can be replaced with the average value of that variable, or the entire cycle can be deleted.
2. Create a scatter diagram that plots each of the time elements against the most important independent variable. Table C-1 shows the potential relationships. A scatter diagram can answer four questions: 1) Is the relationship linear? 2) Does there seem to be a significant trend? 3) Are there any outliers? 4) Is there a good range of data in the X direction? A software package such as Statgraphics can perform these functions statistically. Discard or replace any bad data.

Figure C-1 is an example of a scatter diagram in which a linear relationship exists; as carriage distance

Table C-1. Potential relationships between elements of the yarding cycle and independent variables.

Dependent element (time)	Independent variable
Outhaul and inhaul	Slope distance (ft)
Lateral in and lateral out	Lateral distance (ft)
Hook and unhook	Number of pieces
Drop of chokers from carriage	Carriage height (ft)

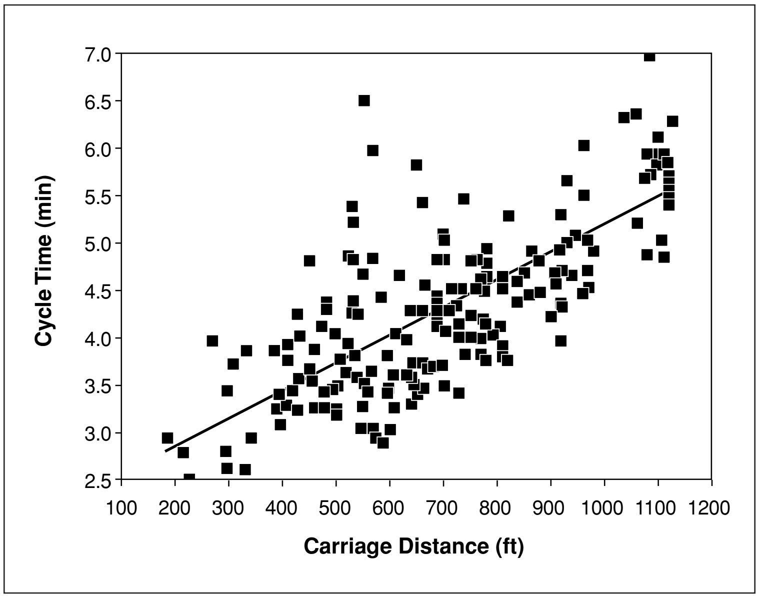

Figure C-1. Scatter diagram of data from detailed time study, showing a linear relationship.

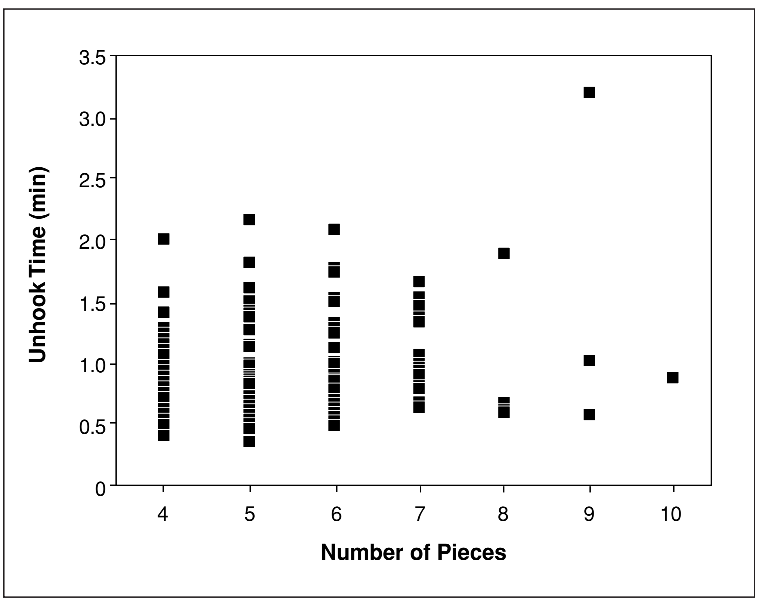

Figure C-2. Scatter diagram of data from detailed time study, showing no relationship.

increases, more time is required to retrieve the carriage to the landing. Outliers have been removed. Figure C-2 is an example of a scatter diagram in which no relationship exists between the time element (unhook time) and the independent variable (number of pieces in the cycle).

3. Replace outliers with the mean time and calculate the delay time. The data that were removed can be coded as a delay for the cycle and added to the total delay time. If the data appear to fit a nonlinear curve, consider using transformed variables in the regression; for instance, the lateral distance could be squared to reflect the increasing difficulty of manually pulling line to the side.

Regression Analysis

Candidates for Independent Variables

The scatter diagram will suggest candidates for the significant independent variables, but researchers need to be aware of any relationships between independent variables. Correlating the independent variables against each other can check this. A high correlation coefficient (r) suggests that only one of the two variables should be used; otherwise a spurious coefficient may result. For example, the slope extraction distance and number of pieces are sometimes correlated, meaning the number of pieces in a cycle increases with increasing yarding distance. In this situation, one of the variables should be dropped.

Stepwise Regression

Stepwise regression can be used to choose which variables to retain or add. Stepwise regression adds the most influential variable to the model first and reports the output. Then the next most influential variable is also added. This continues until the remaining variables produce an insignificant result. To obtain 95% confidence that the coefficient for the variable is statistically different from 0, researchers set the criteria at either t = 1.96 for t-tests or F = 4 for *F-tests*.

Researchers should confirm that the sign and magnitude of each coefficient is appropriate and should try to interpret the meaning of the indicator variables. For example, the coefficient of the distance variable is minutes per foot. This is the inverse of the line speed, feet per minute. This speed should be roughly equal to the line speed during inhaul and outhaul. This step in the process is an art form. Researchers may need to experiment with several combinations of variables to find one that is acceptable.

Residual Analysis

After a satisfactory regression equation has been found, residual analysis can be used to locate any remaining unwanted outliers. For each cycle, the independent variable values are inserted into the regression equation to predict cycle time. Subtracting the predicted from the observed cycle time gives the residual value. The residuals (Y variable) are plotted against the total cycle time (X variable). If any cycle has a residual greater than three standard deviations, the regression may need to be redone after that cycle is eliminated. The whole residual analysis can be conducted with a software package such as Statgraphics 2.0 for Windows.

If the residuals appear to have a pattern, then a nonlinear regression should be considered. One or more of the independent variables may have to be transformed, and the regression begun again. The residual plot shown in Figure C-3 is an example of constant variation and no trend.

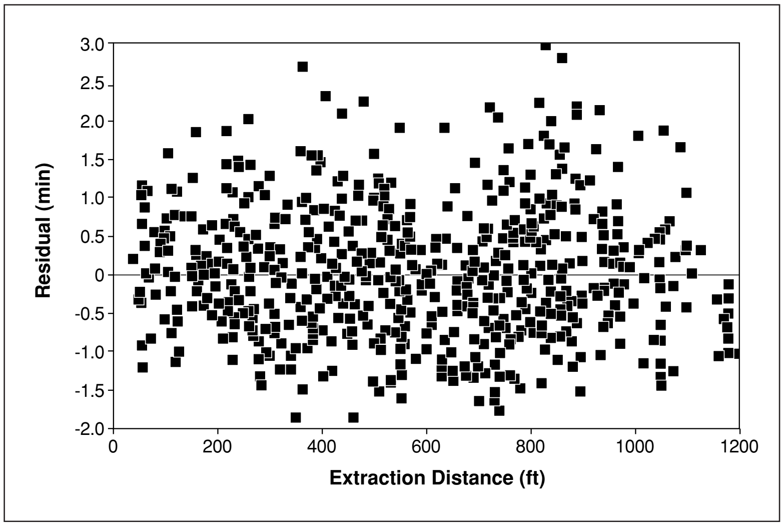

Figure C-3. Residual plot of detailed time study regression results.

Departures from Random Distribution of Residuals

If the regression equation is an apt model of the data, the residuals should be normally distributed. A chart of the residuals should show a random distribution about zero. Visual inspection can identify whether this is violated by one of the following conditions: 1) outlying residuals are much larger than the rest, 2) the residuals cluster in groups, or 3) the sizes of the residuals change as the Y estimate increases.

Large Sample Formulae

The formulae used in this paper are for 1) large samples, 2) sampling with replacement, and 3) no interaction among variables. The statistical formulae for small sample sizes (N < 30) require correction terms. With large samples of several hundred cycles, these corrections become inconsequential. In addition, sampling of harvesting production can be considered to be sampling with replacement, which also simplifies the formula being used. We further assume independence between the variables, which allows us to disregard covariance terms in the formula.

Appendix D: Logging Production Flow Chart

Figure D-1 is a flow chart of how pieces per hour can be calculated for a harvesting operation.

- Effective hour is the number of productive minutes in a scheduled work hour. It is calculated by first estimating the percentage of the workday that is lost in delays. For example, if delays represent 20% of the workday, then the effective hour would be

  60 min/hr × (100%–20%) = 48 min/hr

- Delay-free cycle time is the number of minutes in an average cycle with no delays. The value can be obtained from a regression equation, which takes into account the important variables, such as extraction distance. For example, 6 min/cycle might be typical.
- The number of cycles per scheduled hour is calculated by dividing the effective hour by cycle time. In our example, (48 min/hr)/(6 min/cycle) = 8 cycles per scheduled or total hour.
- The number of pieces per cycle is then multiplied by the number of cycles per hour to obtain the number of pieces per scheduled machine hour (SMH). In our example, (4 pieces/cycle) x (8 cycles/SMH) = 32 pieces/SMH.
- By going one step further, the volume per hour could be calculated. This was not done in this paper. The average volume per piece is multiplied by the pieces per SMH to obtain the volume per SMH.

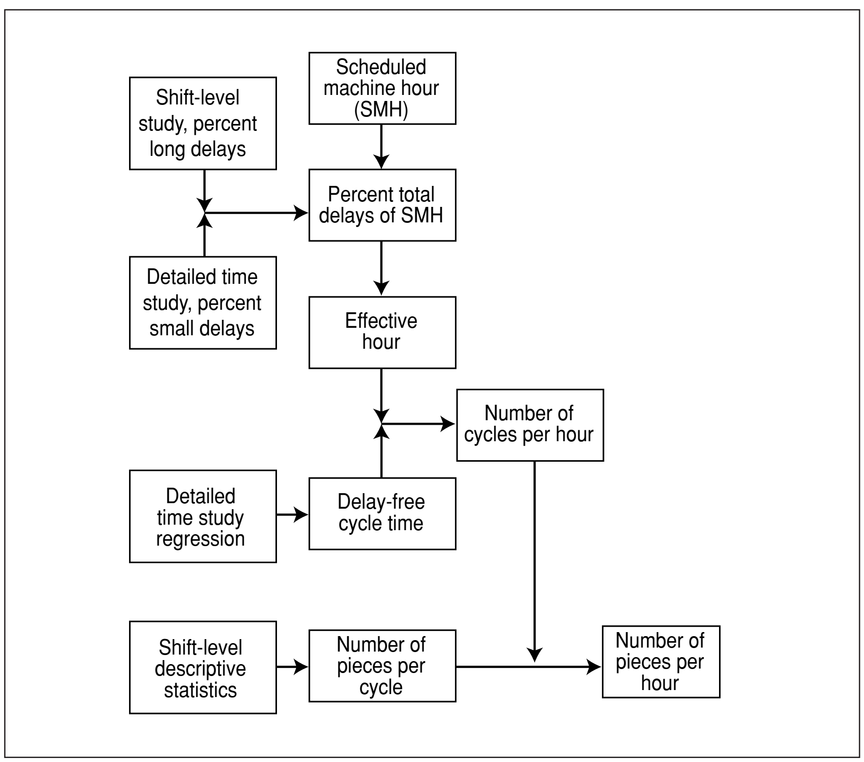

Figure D-1. Flow chart of information for production calculation.

Forestry Publications Office
Oregon State University
227 Forest Research Laboratory
Corvallis, OR 97331-7401

Address Service Requested

Non-Profit Org.
U.S. Postage
PAID
Corvallis, OR
Permit No. 200