HARVESTING TECHNICAL NOTE

Vol:2 Number: 8
2009

Comparison of Methods for Estimating Harvesting Payload

Summary

Different methods were investigated for estimating harvesting payload as a means of daily production monitoring or for production studies. A simulation model was used to estimate the prediction error from using different forms of a double sampling methodology (piece counting or measuring large end diameter or large end diameter and length). Beyond simple piece counting, the form of the regression model had very little effect on the prediction errors for daily production. The results showed that daily production can be predicted within 2% if double sampling with a sample size of at least 50 stems is used. The simple piece count method is probably acceptable for routine production monitoring within the harvesting crew, however it is not suitable for detailed production studies used for rate setting or research purposes.

Hamish Marshall, Interpine Forestry Ltd.

Introduction

Collecting actual production information on a harvesting crew is important for a number of performance monitoring and research purposes. This information is most commonly collected in work studies which involve researchers timing the various components that make up the harvesting cycle (Olsen *et al*., 1998). As the goal of harvesting production monitoring is to determine productivity in terms of volume of wood per productive hour, it is important to obtain an accurate measure of the volume of wood extracted from the harvest unit.

In shift-level studies and monthly production monitoring, volume is usually determined from truck load-out dockets (weighbridge dockets); however production studies and daily or hourly production monitoring require collection of more detailed volume measurements. This means that the volume of each extraction cycle (haul) needs to be either measured or estimated. Obtaining measurements of individual stem volume on a haul-by-haul basis often puts the worker collecting these data in a hazardous position, working in the area where stems are landed in close proximity to logging machinery and moving ropes.

This report investigates different methods of calculating individual haul size (payload) at a precision suitable for production monitoring.

In harvesting production studies, the most common way to determine the payload per production cycle is to use double sampling. This method requires detailed measurements of volume on a sub-sample of stems, plus other measurements that are easily taken on all stems during the production operation. For example, a sub-sample of data is used to create a relationship between large end diameter (LED) and stem length of each tree to stem volume.

Then during the normal production operation, the only data that have to be collected for all stems are of the easily measured variable of LED, and the volume is predicted from the relationships of LED to length and stem volume. In this way the data collection worker spends minimum time in the hazardous log landing zone. This double sampling method is quicker, cheaper and less hazardous than recording detailed measurements for all stems extracted to the landing.

There are a number of ways of applying this methodology in terms of the variables to collect and the size of the sub-sample. There are clearly trade-offs between the accuracy of the volume or haul size estimate and the time and cost to collect the data.

Methods

The accuracy of different double sampling methods is best demonstrated using a simulation model. The model simulated 20 days

Vol:2 Number: 8
2009

production of a cable yarding operation with an average target production of 250 m^3/day. The model randomly selected sets of 2-4 stems to simulate each haul cycle. The actual production was then calculated by adding up the volume from all the hauls.

The simulation model selected stems from a dataset of over 450 scaled stems of approximately 18 metres in length. These stems were all sectional measured at 0.3 metre intervals using the Timbertech optimisation callipers. These measurements were then used to calculate the sectional volume of each stem. The model was used to investigate the effect of the following:

- The type of predictive variable (i.e., those used to predict stem volume). In this study these were stem count, LED, stem length, the square of LED, and the square of the length;
- The size of the sub-sample used to calculate the predictive model. Sample sizes of 25, 50, 75 and 100 stems were trialled;
- The relationship type (i.e., linear or quadratic form used to relate the predictive variables to haul size (volume).

In the stem counting method, the number of stems was counted in each haul. This count was then multiplied by an estimate of the average stem volume. This was estimated by simply calculating the mean volume of the sub-sample of stems. In the other methods the regression model was generated using the sub-sample of stems for which volume was known. This regression model was then used to calculate the volume for each stem using LED, length or a combination of both. This was achieved by simply inserting the required variables (LED, length or both) into the regression equation, which produced a prediction of stem volume for each haul. The simulation model was run 1000 times to reduce the effects of random variability. The results were averaged for those 1000 simulations.

Results

The following results demonstrate the trade-offs in using the different methods for estimating the volume payload. The fit statistics, mean squared error (MSE), coefficient of determination (R^2), and the probability of statistical significance, (p-value) for all the relationships are given in Table 1.

Table 1. The fit statistics for all the regression-based methods

Volume = f (LED)			
Sub-sample	MSE	R^2	p-value
25	0.3015	0.8658	<0.001
50	0.3061	0.8672	<0.001
75	0.3071	0.8663	<0.001
100	0.3097	0.8647	<0.001
Volume = f (LED^2)			
Sub-sample	MSE	R^2	p-value
25	0.3056	0.8601	<0.001
50	0.3130	0.8595	<0.001
75	0.3132	0.8597	<0.001
100	0.3179	0.8557	<0.001
Volume = f (LED, Length)			
Sub-sample	MSE	R^2	p-value
25	0.2521	0.9106	<0.001
50	0.2591	0.9074	<0.001
75	0.2606	0.9054	<0.001
100	0.2623	0.9042	<0.001
Volume = f (LED^2, Length)			
Sub-sample	MSE	R^2	p-value
25	0.2818	0.9078	<0.001
50	0.2519	0.9109	<0.001
75	0.2526	0.9099	<0.001
100	0.2558	0.9075	<0.001

FUTURE FORESTS RESEARCH

HARVESTING TECHNICAL NOTE

Vol:2 Number: 8
2009

An MSE of zero means the model predicts observations with perfect accuracy. The R squared (R^2) is the proportion of variability in a data set that is accounted for by the statistical model. The R^2 values vary between 0 and 1 and the higher the value the better the outcomes are likely to be predicted by the model.

In these results the MSE and R^2 improved with the increasing complexity of regression (i.e., adding the length variable, then squaring the LED), however both measures deteriorated slightly with increasing sample size beyond 50 samples. In all cases the relationships were strong.

The average daily production from the simulations was 241.3 m^3. Table 2 shows the average absolute error in the prediction of volume production per day (m^3) from the 1000 simulations. There was a large decrease in error between the piece count and other methodologies. The difference in error between the regression-based methods was much smaller. The level of error also reduced as the sample size increased.

Table 2. Average absolute error (actual - predicted) in volume (m3 per day)

Sub-sample	Piece Count	LED	LED^2	LED and Length	LED^2 and Length
25	15.53	5.77	6.08	4.926	4.91
50	11.84	4.42	4.63	3.759	3.74
75	10.66	3.93	4.09	3.371	3.32
100	9.46	3.60	3.76	3.079	3.05
All Stems	6.88	2.56	2.69	2.167	2.15

Table 3. Average absolute error as a percentage of average daily production

Sub-sample	Piece Count	LED	LED^2	LED and Length	LED^2 and Length
25	6.4%	2.4%	2.5%	2.0%	2.0%
50	4.9%	1.8%	1.9%	1.6%	1.5%
75	4.4%	1.6%	1.7%	1.4%	1.4%
100	3.9%	1.5%	1.6%	1.3%	1.3%
All Stems	2.8%	1.1%	1.1%	0.9%	0.9%

Table 3 shows these absolute errors as a percentage of the average daily production. Using a piece count to estimate haul size resulted in errors, irrespective of sample size, of between 4% and 6%. This may be satisfactory for daily or hourly production monitoring purposes.

Using a regression of volume on LED based on a sub-sample of 25 stems gave an error of 2.4%, which reduced to less than 2% if based on a larger sub-sample.

Vol:2 Number: 8
2009

Table 4 shows the range of average error in the volume prediction over a 20-day production period averaged for the 1000 simulations.

The results for the piece count method based on average volume from a sample of 25 stems show the prediction in daily production could be within ±66m^3 of the actual productivity. This reduces to ±11m^3 by using a double sampling methodology using LED and a sub sample of 100 stems.

Table 4. Range of average error (m3) in the volume prediction for a 20-day period

Sub-sample	Piece Count	LED	LED^2	LED and Length	LED^2 and Length
25	66.04	19.46	20.68	18.21	17.20
50	46.77	17.42	20.96	11.64	14.82
75	29.26	12.30	13.02	10.91	10.79
100	32.52	10.96	11.59	9.22	8.60
All Stems	6.96	2.22	2.31	1.86	1.88

Conclusion

From the results of this simulation, it appears that the piece count method is acceptable only for routine production monitoring by the harvesting crew. It is probably not accurate enough for detailed production studies for rate setting or research purposes. In methods improvement studies, the errors in the volume estimates from using the piece count estimation method would likely be greater than any calculated change in productivity caused by varying the harvesting techniques.

There did not seem to be a major difference between the other methods of predicting volume and hence payload. In most cases the percentage error in estimation was less than 2% of daily production.

There seems to be little advantage in spending the extra time and effort in measuring stem length in this case, especially as measuring tree length stems can be both time-consuming and hazardous in a production logging operation.

In this simulation, the sample size had more influence on the level of error in the prediction than the predictive variables used. This study has shown that using any of the double sampling methods with a sample size of greater than 50 stems produced average prediction of haul volume acceptable for both production monitoring purposes and productivity improvement studies. Where these improvement studies are likely to result in only small increases in productivity, measuring the diameter and length of every stem may be advisable.

References

Olsen, E.D., Hossain, M.M., and Miller M.E. (1998): Statistical comparison of methods used in harvesting work studies. Research Contribution 23, Forest Research Laboratory, Oregon State University, Corvallis Oregon.