Determining Observer and Method Effects on the Accuracy of Elemental Time Studies in Forest Operations

RAFFAELE SPINELLI[1*], RUBEN LAINA-RELAŃO[2], NATASCIA MAGAGNOTTI[1] AND EDUARDO TOLOSANA[2]

[1] *CNR IVALSA, Via Madonna del Piano 10, I-50019 Sesto Fiorentino – Italy;*

[2] *E.T.S.I. Montes – Univ. Politecnica de Madrid, Ciudad Universitaria s/n, E-28040 Madrid – Spain*

Spinelli, R., Laina-Relańo, R., Magagnotti, N. and Tolosana, E. 2013. Determining Observer and Method Effects on the Accuracy of Elemental Time Studies in Forest Operations. *Baltic Forestry* 19(2): 301–306.

Abstract

The authors conducted a comparative study of continuous timing and work sampling, as applied by different observers. Twelve researchers were split into two groups and carried out the same study, switching method every half an hour. The study lasted 3 hours, corresponding to 6 iterations (3 iterations per method). Statistical analysis of data confirmed that both methods may determine elemental breakdown with the same accuracy. Observer-induced variability had a minor effect, and only on those time elements that were short and occurred less frequently. Error increased as the study progressed, as the likely result of observer fatigue. Again, this trend was significant for short elements and not significant for longer ones. This study may offer a warning against over-detailing elemental breakdown. Work cycles should be split in as many elements as strictly necessary for achieving the specific purpose of the study. Splitting them in too many elements just for the sake of description is counterproductive, because it places an excessive strain on the observer and it may increase the risk of errors. A simple study design will facilitate replication by a number of different researchers, regardless of method.

Key words: time studies, methodology, accuracy, work sampling, operations

Introduction

Originating over 100 years ago in the works published by Taylor (1895), work studies have become an integral part of forest research (Sundberg 1988, Samset 1990). They are used for many practical purposes, such as: setting work rates, scheduling harvesting activities, and comparing technologies or work methods (Björheden 1991).

The core of most work studies is a time study, whose purpose is to determine the time needed to carry out a specific task, the associated output per unit time and the effect of independent variables on time consumption (Anson 1953). Time studies offer different resolution depending on the level of detail in which they describe the studied process (Harstela 1988).

Elemental time studies are the most detailed and consist of splitting the work cycle into functional steps (elements) and then recording time consumption separately for each of them (Kanawaty 1992). The benefits of elemental measurement are: 1) indicating which specific process steps take more time, so that specific improvement measures will primarily target these steps; 2) separating effective work time from delay time (Björheden et al. 1995), since these two categories have different internal variability and could be modelled in different ways; 3) separating functional elements that react to different work characteristics, so that accurate sub-models can be developed (Bergstrand 1991).

Elemental breakdown can be obtained by timing individually all work steps as they progress, or by observing the work process at fixed or random intervals, and noting in which of the previously-defined functional steps the work team is engaged in that specific moment (Harstela 1991). The latter method is called ‘work sampling’, and offers the main advantage of allowing one researcher to follow more teams at a time, by organizing a sequence of observation intervals for the different teams (Olsen et al. 1998). While both methods should theoretically return the same results if applied to the same conditions, the question arises about whether they actually do, especially if one considers the effect of observer proficiency and understanding of the method. It may occur that different people may respond differently to the two methods, and that they may work better with one or the other (Motowildo et al. 1997). Despite the steady progress of automated time studies (McDonald and Fulton 2005), much of the work is still done manually by field researchers, with all its advantages and limits (Peltola 2003). For this reason, it is important to gauge the variability introduced by the observer, as well as the potential interaction between observer and method. Contrary to computers, human observers will tire over the work day, which may affect their performance and introduce further variability.

Therefore, the goals of this study were: 1) to determine if the two study methods return the same elemental breakdown when applied under the same conditions; 2) to gauge the variability introduced by the observer; 3) to check if record quality deteriorates over time, as a consequence of observer tiredness.

Materials and Methods

The study was carried out during the COST Training School organized in December 2011 (www.forestenergy.org). Forest engineering researchers from all over the world gathered in Florence for the event, offering an ideal pool for the experiment. Twelve young researchers were selected for the test coming from seven different European Countries. Before the experiment, all participants received specific lessons about time study techniques. This was done withn the purpose for harmonizing methods rather than teaching time study principles, since almost all participants were already familiar with conducting time studies, which they occasionally performed at home.

The researchers were then taken to the forest for studying a tree processing operation in a pine plantation. The forest was located in the Apennine Mountain near Acquerino, not far from Pistoia and about 50 km from Florence. In the operation, a small processor (Arbro 400S) was used mounted on a light excavator (9 tonne Caterpillar). With the machine whole trees were processed into 3 to 5 m sawlogs, separately piling logs and slash, for later recovery and chipping. Trees were skidded to the processor with a forestry fitted farm tractor in a hot-deck chain, which forced the processor to recurrent waiting when the tractor was late to deliver a new load. The operation was chosen because it was developed according to a steady routine, made of short and repetitive cycles. Once on site, the researchers were divided into two groups of six people. Each researcher was given the same notepad, pencil and manual stopwatch. They were all informed about the purpose of the study, which was to determine the elemental breakdown of processor work and especially the incidence of waiting time on the total time consumption. Before starting, time elements were jointly defined, and so were the break points between sequential work steps (i.e. the moment in which a work step ends and the following step begins). Time elements and breakpoints are described in Table 1. The process was divided into few meaningful steps, which lasted long enough for easy recording with a manual stopwatch. The field day was organized in six iterations of half an hour each, corresponding to about 5 tractor trips or 8 processor cycles per iteration. During each iteration, one group conducted a classic elemental time study based on continuous timing, while the other group performed a work sampling study. After half an hour, the groups switched time study technique. Researchers who had previously conducted the continuous time study switched to work sampling, while those engaged in work sampling switched to continuous time study. At the end, all researchers had worked half of the time with continuous time study and half of the time with work sampling. Work sampling was conducted at random intervals, whose duration has averaged 30 seconds (Miyata et al. 1981). All researchers were provided with random number tables and were instructed to sample according to the interval sequence contained in the tables. This was done to avoid the risk of accidental synchronization between observation interval and cyclic elements. Tables were different for each iteration and researcher.

Table 1. Description of the time elements

Manual processing: the operator delimbs and/or crosscuts with a chainsaw those branches and stem portions that are too big for the small processor.
Mechanical processing: picking up, handling, delimbing, crosscutting and stacking logs with the processor.
Waiting for wood: processor and operator are idle because no wood is available at the landing.
Other delays: any other interruptions of the work process, except for those caused by the study itself (these were excluded from recording).

The researchers also noted the time at the beginning and at the end of each iteration, as well as the diameter at breast height (DBH) of all trees processed during that period. This information allowed calculating productivity, after DBH data had been transformed into volume over bark using appropriate volume tables.

The study proper lasted for three hours, with a lunch break in between. Before starting data collection proper, researchers did a warm-up exercise with one of the teachers indicating the breakpoints as they occurred. This was done in order to reach a common understanding of the work process and the study procedure. All researchers stood at the same safe distance from the processor, none having a better view of the operation than the others. The day was clear and cold, with temperatures just few degrees above zero. That was assumed to cause tiredness and possibly affect observer performance.

The following day, each researcher organized his/her data in an Excel sheet for delivery to the course organizers. The authors then compiled all worksheets in a single master database. This consisted of a balanced factorial experiment containing 72 samples, resulting from the combination of 12 researchers × 2 methods × 3 replications. A replication was considered as a single half-hour time study bout. Elemental breakdown was calculated as the percent incidence over total work time, including delays. Before analysis, percentage data were a subject to an arcsine transformation in order to normalize their distribution. An arcsine transformation is the most suited when data is expressed as percent figures, and it did prove effective in restoring normality. The effect of treatments was gauged with classic Analysis of Variance (AnoVA) techniques, returning the strength and the significance of all main effects, namely: method and researcher.

Results

Processor performance was estimated at 7.6 m^3 over bark per scheduled machine hour (SMH), inclusive of all delays (Table 2). Time element breakdown is reported in Figure 1, which highlights the substantial incidence of waiting time. Manual processing was an important component of the work cycle and accounted for 11% of total scheduled time.

Table 2. Main data about processor performance

	Unit	Mean	SD	Min	Max
Net time	s per iteration	1102	334	328	1551
Delays	s per iteration	778	338	186	1518
Total time	s per iteration	1891	256	1290	2570
Output	m^3 per iteration	3.9	0.9	1.1	5.8
Output	trees per iteration	7.9	2.5	2.0	15.0
Piece size	m^3 per tree	0.535	0.184	0.283	1.161
Net productivity	m^3 per PMH	13.7	4.7	4.7	34.4
Gross productivity	m^3 per SMH	7.6	1.8	2.3	11.9

Notes: m^3 = stem volume over bark; PMH = productive machine hours, excluding delays; SMH = scheduled machine hours, including delays; SD = standard deviation.

Table 3 shows the results obtained with the two methods for each time element and iteration. Out of 24 possible matches, only 3 returned a statistically significant difference between methods. The standard *t*-test could not determine if these differences were actually related to the method or to the observer, since different observers applied different study methods during each iteration. There was no difference between methods once the results of all iterations were cumulated in one single three-hour long study session.

A full factorial ANOVA confirmed that study method was not a significant source of variation for any of the time elements under study (Table 4). In contrast, the observer introduced a minor (27%) but significant source of variation when estimating the incidence of one of the time elements, namely: "waiting for wood". In all cases, between 60% and 80% of the total variation was due to random factors, including the variability of the process itself. The interaction between method and observer had no significant effect on variability, indicating that no observer was systematically more proficient with one method compared to the other.

The effect of observer tiredness was checked by plotting the standard deviation from the mean incidence of each time element against the number of it-

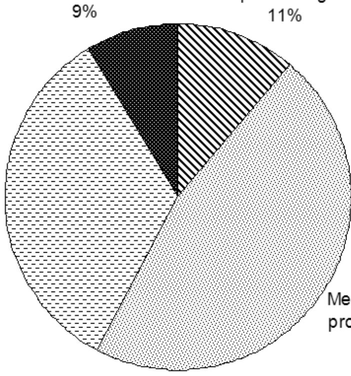

Figure 1. Average element breakdown

Table 3. Percent incidence of time elements resulting from the two methods, per iteration and total

Iteration	Manual processing			Mechanical processing			Waiting for wood			Other delays		
n°	C	WS	*P*-Value	C	WS	P-Value	C	WS	*P*-Value	C	WS	*P*-Value
1	16.3	17.2	0.6661	52.1	53.2	0.739	31.1	28.7	0.2988	0.1	0.2	0.8064
2	15.6	15.8	0.9091	56.5	57.3	0.7989	25.5	26.4	0.7155	1.1	0.0	0.1411
3	12.2	9.5	0.4637	40.8	36.1	0.3168	21.3	38.0	**0.0209**	23.0	16.1	0.0612
4	3.2	8.1	**0.0441**	12.6	27.5	0.1352	67.6	45.3	0.1403	11.3	12.8	0.8773
5	5.7	1.4	**0.0442**	64.2	62.5	0.7096	26.7	30.9	0.4341	2.4	3.2	0.6875
6	9.2	12.6	0.5289	48.9	47.0	0.7375	33.0	21.2	0.264	7.0	12.7	0.1939
All	9.8	9.9	0.9283	45.1	47.1	0.6414	33.8	31.5	0.5796	5.3	5.2	0.9555

Notes: C = continuous timing; WS = work sampling; *P*-Value = result of the *t*-test comparing C with WS; *P*-Values in bold highlight statistically significant differences at the 5% level between C and WS.

erations (Figure 2). As the study progressed, the standard deviation increased for all elements, indicating that some observers were becoming less accurate. Trends were modelled with the regression technique, assuming that the model had no intercept, since no deviation could exist before any measurement has been done (i.e. iteration 0). The ascending trend was significant only for those time elements that had the smallest incidence within the total cycle time, namely: "manual processing" and "other delays" (Table 5).

Discussion

The productivity levels recorded with the study were slightly higher than obtained from a previous study using the same machine and operation, possibly as a result of the different species and the larger tree size (Spinelli et al. 2009). In any case, these levels conformed to current Italian standards for job type and machine size (Spinelli et al. 2010).

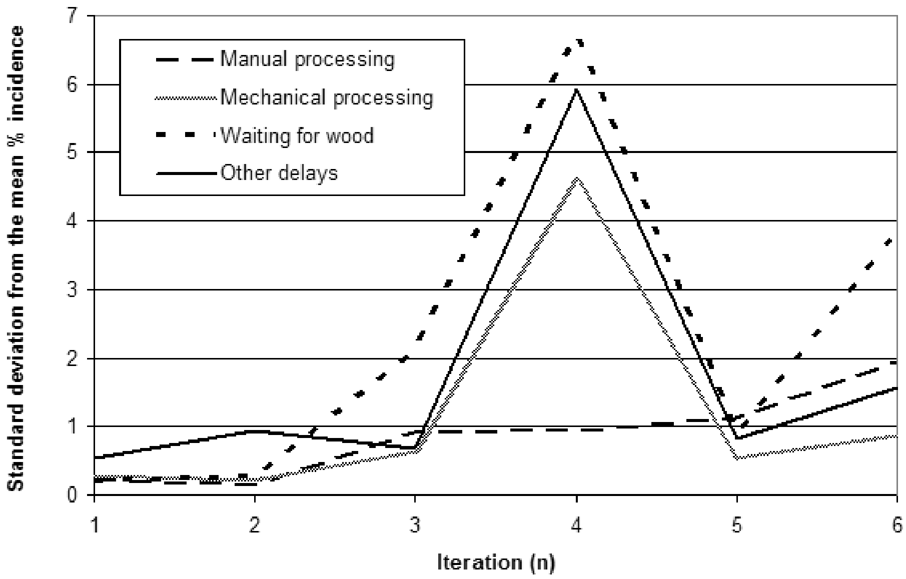

Figure 2. Standard deviation vs. study duration (in number of iterations)

Table 4. Result of the analysis of variance for the arcsine-normalized element incidence values

Element	Factor	SS	Variation	*F*-Value	*P*-Value	Power
Manual processing	Method	1.3E-04	0%	0.007	0.9328	0.051
	Observer	0.143	13%	0.711	0.7224	0.332
	Interaction	0.107	9%	0.531	0.8726	0.246
	Residual	0.881	78%			
Mechanical processing	Method	0.008	0%	0.183	0.6706	0.070
	Observer	0.148	6%	0.326	0.9761	0.157
	Interaction	0.289	12%	0.637	0.7879	0.296
	Residual	1.982	82%			
Waiting for wood	Method	0.011	0%	0.346	0.5591	0.087
	Observer	0.670	**27%**	1.961	**0.0544**	0.829
	Interaction	0.269	11%	0.788	0.6509	0.369
	Residual	1.491	61%			
Other delays	Method	6.6E-08	0%	4.1E-06	0.9984	0.050
	Observer	0.109	11%	0.608	0.8125	0.282
	Interaction	0.069	7%	0.382	0.9570	0.180
	Residual	0.785	82%			

Notes: *P*-Values in bold highlight statistically significant effects at 5% level.

Table 5. Regression equations describing the relationship between standard deviation and study duration

Model: Standard deviation = a Iteration^b				
Parameters	Manual processing	Mechanical processing	Waiting for wood	Other delays
a	0.297	0.303	2.093	0.334
b	1.330	0.944	1.747	0.662
r^2	0.881	0.318	0.332	0.607
F-Value	45.557	3.792	3.977	10.267
P-Value	**0.001**	0.109	0.103	**0.024**

Notes: *P*-Values in bold highlight statistically significant effects at 5% level.

Observer ability seemed only marginally impaired by fatigue, since variation did not increase sharply and significantly for the main time elements. Besides, these trends were estimated on relatively few data points, so that the result may be indicative, not conclusive.

The two timing methods on test returned the same time element breakdown for studies lasting longer than half an hour. Small but significant differences were occasionally found for individual iterations, lasting for up to half an hour. These differences concerned those time elements that had the lowest incidence within the total time consumption, such as short delays. Similar findings were also reported by Olsen and Kellogg (1983). Short elements like these are especially difficult to capture with a brief time study (Pehkonen 1978).

The observer-induced variability was minor, and it was seldom significant. This might be explained by two main factors. First, the relatively even level of experience of the participants, who were all researchers and had already conducted time studies, although with different frequency and intensity. Second, the simple elemental breakdown adopted for the study, which subdivided the process into few meaningful steps, lasting long enough for easy recognition and recording. It is likely that a more complex elemental breakdown and a faster work pace may have resulted in a higher variability of results (Nuutinen et al. 2008).

This is a warning against over-detailing elemental breakdown. Work cycles should be split in as many elements as strictly necessary for achieving the specific purpose of the study. Splitting them in too many elements just for the sake of description is counter-productive, because it places an excessive strain on the observer and it may increase the risk of errors (Magagnotti and Spinelli 2012). Designing a simple breakdown and extending study duration over some hours seem to be the best solutions for achieving accuracy and reliability.

On the other hand, a long study session may incur the negative effect of observer fatigue. Pehkonen (1973) stated that measuring accuracy declines after two hours of study, as the result of observer wear. Conducted under real field conditions, our study appears to confirm Pehkonen's statement. If fatigue is a problem, the elemental breakdown is articulate and the study cannot be automated, then video recording could be the answer (Björheden 1988). In this respect, there would be good (statistical) reasons to design a study with many iterations, on different days and times of the day.

Finally, it is worth noticing that similar results were obtained from a pool of observers, coming from different countries, where different time study practices are possibly used. Before conducting the experiment, all observers attended a two-day introduction course to time study methods, and received careful instructions on data collection. Therefore, our experiment may also point at the potential benefits obtained from common training actions in terms of methodology harmonization, which is a prerequisite for the efficient exchange of knowledge within Europe and beyond.

Conclusions

Continuous timing and work sampling may determine the elemental breakdown with the same accuracy, provided that the elemental breakdown is simple enough and that the study is extended over few hours. Under these conditions, observer error may have a minor effect, and only on those time elements that are short and/or occur infrequently. That accounts for researchers with some previous experience of time studies, not for novice investigators. After few hours of observation, observer fatigue may affect study accuracy, regardless of method. A simple study design will facilitate replication by a number of different researchers.

Acknowledgements

This study was made possible thanks to funding received from the Training School programme of Action COST FP902.

References

Anson, C.J. 1953. The development of time study. *Work Study.* 2 (12): 19-25.

Bergstrand, K.G. 1991. Planning and analysis of forestry operation studies. Skogsarbeten Bulletin n. 17. Skogsarbeten, Garpenberg, 63 pp.

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

Björheden, R. 1991. Basic time concepts for International comparisons of time study reports. *Journal of Forest Engineering* 2: 33–39.

Björheden R., Apel K., Shiba M. and Thompson, M.A. 1995. IUFRO Forest work study nomenclature. Swedish University of Agricultural Science, Dept. of Operational Efficiency, Garpenberg. 16 pp.

Harstela, P. 1988. Principle of comparative time studies in mechanized forest work. Scandinavian *Journal of Forest Research* 3: 253-257.

Harstela, P. 1991. Work studies in forestry. *Silva Carelica* 18: 41 pp.

Kanawaty, G. (Ed.) 1992. Introduction to work study. International Labour Office. Geneva, Switzerland. 523 pp.

Magagnotti, N. and Spinelli, R. (Ed.). 2012. COST Action FP0902 - Good practice guideline for biomass production studies. CNR IVALSA. Florence, Italy. 41 pp. ISBN 978-88-901660-4-4

McDonald ,T. and **Fulton, J.** 2005. Automated time study of skidders using global positioning system data. *Computers and Electronics in Agriculture* 48: 19-37.

Miyata, S., Steinhilb, H. and Winsauer, S. 1981. Using work sampling to analyze logging operations. USDA Forest Service Research Paper NC-213. North Central Experiment station, St. Paul MN, 56 pp.

Motowildo, S., Borman, W. and Schmit, M. 1997. A theory of individual differences in task and contextual performance. *Human performance* 10: 71-83.

Nuutinen, Y., Väätäinen, K., Heinonen, J., Asikainen, A. and Röser, D. 2008. The accuracy of manually recorded time study data for harvester operation shown via simulator screen. *Silva Fennica* 42: 63-72.

Olsen, E. and Kellogg, L. 1983. Comparison of time-study techniques for evaluating logging production. *Transactions of the ASABE* 26: 1665-1668.

Olsen, E., Hossain, M. and Miller, M. 1998. Statistical Comparison of Methods Used in Harvesting Work Studies. Contribution n° 23, Oregon State University, Forest Research Laboratory, Corvallis, OR, 31 pp.

Pehkonen, A. 1973. Aikatutkimustekniikoista ja nitide ajanmittausvaiheen tarkkuudesta. tutkimustiedote 5, Helsingin yliopiston maatalouden työtekniikan laitos, 159 pp. (In Finnish).

Pehkonen, A. 1978. Accuracy of timing in some time study methods. *Journal of the Scientific Agricultural Society of Finland* 50: 1-60.

Peltola, A. 2003. IT-time for mechanized forest work study. 2nd Forest Engineering Conference 12–15 May 2003, Växjö, Sweden. *Skogsforsk Arbetsrapport* 536: 107–112.

Samset, I. 1990 Some observations on time and performance studies in forestry. Communication 43.5 Norwegian Forest Research Institute, Ås, 80 pp.

Spinelli, R., Hartsough, B. and Magagnotti N. 2010. Productivity standards for harvesters and processors in Italy. *Forest Products Journal* 60: 226-235.

Spinelli, R., Magagnotti, N. and Nati, C. 2009. Options for the mechanized processing of hardwood trees in Mediterranean forests. *International Journal of Forest Engineering* 20: 30-35

Sundberg, U. 1988. The emergence and establishment of forest operations and techniques as a discipline in forest science. Communication 41.8 Norwegian Forest Research Institute, Ås, 37 pp.

Taylor, F.W. 1895. A piece-rate system being a step toward partial solution of the labor problem. *Transactions of the American Society of Mechanical Engineers* 16: 865-903.

Received 27 November 2012
Accepted 28 November 2013

ОПРЕДЕЛЕНИЕ ВОЗДЕЙСТВИЙ НАБЛЮДАТЕЛЯ И МЕТОДА НА ТОЧНОСТЬ ЭЛЕМЕНТНЫХ ИССЛЕДОВАНИЙ ВРЕМЕНИ В ЛЕСОХОЗЯЙСТВЕННЫХ ОПЕРАЦИЯХ

Р. Спинелли, Р. Лайна-Релако, Н. Магагнотти и Э. Толосана

Резюме

Авторы провели сравнительное исследование непрерывной синхронизации и выборки работ, использованые различными наблюдателями. Двенадцать исследователей были распределены на две группы и провели тот же эксперимент, меняя метод каждые полчаса. Исследование продолжалось 3 часа, что соответствует 6 повторениям (по 3 итерации на метод). Статистический анализ данных подтвердил, что оба метода могут определять элементное распределение с той же точностью. Индуцированная наблюдателем изменчивость имела незначительное влияние, причем только на те элементы времени, которые были короткими и проявлялись реже. Ошибка увеличивалась по мере продолжения исследования как вероятный результат, вызванный усталостью наблюдателя. Опять же, эта тенденция была значительна для коротких элементов и не имела существенного значения для более продолжительных. Данное исследование допускает возможность предотвращения чрезмерной детализации элементного распределения. Циклы работы должны быть разделены на многочисленные элементы, что строго необходимо для достижения конкретной цели исследования. Разделение их на слишком многочисленные элементы только желая указать контрпродуктивность описания, поскольку это накладывает чрезмерную нагрузку на наблюдателя, может увеличить риск возникновения ошибок. Несложный дизайн исследования будет способствовать репликации ряда различных исследователей, независимо от метода.

Ключевые слова: время исследования, методология, точность, выборки работы, операции