EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

COST Action FP-0902

WG 2 Operations research and measurement methodologies

GOOD PRACTICE GUIDELINES FOR BIOMASS PRODUCTION STUDIES

Editors: Magagnotti N., Spinelli R.

Contributors: Acuna M., Bigot M., Guerra S., Hartsough B., Kanzian C., Kärhä K., Lindroos O., Magagnotti N., Roux S., Spinelli R., Talbot B., Tolosana E., Zormaier F.

Reviewers: Björheden R., Kellogg L., LeBel L.

This publication is supported by COST

COST Vademecum (Part A) - Pay-as-you-go System

v05/05/2010 Page 44 / 67

Published by:

CNR IVALSA
Via Madonna del Piano, 10
I-50019 Sesto Fiorentino (FI)
ITALY
www.ivalsa.cnr.it
"Good practice guidelines for biomass production studies"
Year of publication: 2012
ISBN 978-88-901660-4-4

ESF provides the COST Office through a European Commission contract

COST is supported by the EU RTD Framework Programme

PEFC certified
This book is printed on paper from sustainably managed forests and controlled sources.
www.pefc.org
PEFC/18-31-124

Graphic lay out: Comunicambiente.net
Illustrations: Giovanni Tribbiani
Printed by: Litotipografia Alcione S.r.l.

COST- the acronym for European Cooperation in Science and Technology- is the oldest and widest European intergovernmental network for cooperation in research. Established by the Ministerial Conference in November 1971, COST is presently used by the scientific communities of 35 European countries to cooperate in common research projects supported by national funds.

The funds provided by COST - less than 1% of the total value of the projects - support the COST cooperation networks (COST Actions) through which, with EUR 30 million per year, more than 30 000 European scientists are involved in research having a total value which exceeds EUR 2 billion per year. This is the financial worth of the European added value which COST achieves. A “bottom up approach” (the initiative of launching a COST Action comes from the European scientists themselves), “à la carte participation” (only countries interested in the Action participate), “equality of access” (participation is open also to the scientific communities of countries not belonging to the European Union) and “flexible structure” (easy implementation and light management of the research initiatives) are the main characteristics of COST.

As precursor of advanced multidisciplinary research COST has a very important role for the realisation of the European Research Area (ERA) anticipating and complementing the activities of the Framework Programmes, constituting a “bridge” towards the scientific communities of emerging countries, increasing the mobility of researchers across Europe and fostering the establishment of “Networks of Excellence” in many key scientific domains such as: Biomedicine and Molecular Biosciences; Food and Agriculture; Forests, their Products and Services; Materials, Physical and Nanosciences; Chemistry and Molecular Sciences and Technologies; Earth System Science and Environmental Management; Information and Communication Technologies; Transport and Urban Development; Individuals, Societies, Cultures and Health. It covers basic and more applied research and also addresses issues of pre-normative nature or of societal importance.

Web: http://www.cost.eu

Table of contents

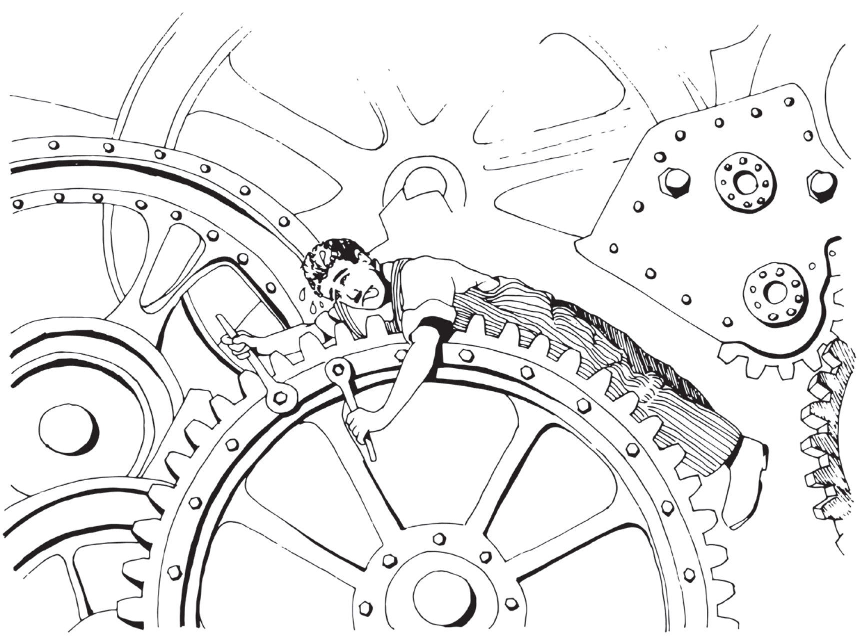
Good practice guidelines on biomass work studies

1. *Introduction*

Forest Work Science is an important branch of Forest Science, which has developed into an independent field since the late 1920's. Despite a strong international cooperation within the forest engineering community, the evolution of the discipline has inevitably generated local adaptations in response to different work environments and individual preferences. The mutual understanding once derived from common study methods has largely been lost. As a result, there is now much misunderstanding about time study methods, both at the theoretical and the practical level. Ambiguity arises especially regarding the terminology, the units of measure, the experimental design and the statistical treatment of data. Hence, there is the need for a good practice guideline (GPG), which must be simple and concise enough to encourage widespread adoption. The guide should ensure comparability of results and repeatability of experiments, both fundamental elements of the scientific method. In turn, this will facilitate international network building and research coordination, which are the ultimate goals of the EU COST Programme. The purpose of this guide is to answer this need. This is a simple and quick how-to guide that can help harmonize work study methods. It is designed for the field researcher who needs quick access to sound study practice, even when lacking strong theoretical skills in work science and/or statistics. Contrary to a scholarly book, this manual goes from practice to theory and not the reverse. In fact, this manual does not replace the many scholarly books dealing with operational studies and the related statistical methods. Readers are encouraged to consult them, if they wish to deepen their understanding of the subject. The reference section contains a partial list of authoritative sources, which readers can use to this very end.

2. *Background*

The origin of work studies is commonly credited to the paper "A piece-rate system being a step toward partial solution of the labor problem" published in 1895 by F.W. Taylor on the Transactions of the American Society of Mechanical Engineers. Taylor was convinced that for each task there was a quickest time in which it could be performed by a "first-class man", without depleting his work capacity. The "first-class man" was the man best fitted to perform that task, through natural and acquired capabilities, including proper equipment. The quickest time was referred to as the standard time, which could be determined through scientific investigation and used for work management. Determining a standard time was considered crucial to setting a fair piece rate and to find the "one best way" for performing a given task. Standard time was subdivided in three main categories: 1 - time actually spent working; 2 - time for overcoming fatigue (rest); 3 - time for overcoming delays. Complex



Scientific management has its roots in an exploitative era characterised by rapid industrialisation tasks were simply treated as a sum of elemental tasks, and their standard duration was considered as the sum of the standard duration of each elemental task. That is where elemental time studies comes from. Some of Taylor's original concepts were criticised early on. Already in 1930, the famous economist A.C. Pigou stated that everybody is continuously learning and that there is no one best way to perform any given task. Another common criticism is that the time to perform a complex task is not necessarily the sum of the times to perform its elemental sub-tasks, because there is often interaction between correlated sub-tasks, leading to time economies or diseconomies. Regardless of critics, Taylor's original philosophy has shaped work management as a discipline. His concepts are still echoed in modern work study techniques even 100 years after their original formulation.

3. Work measurement

Work science has multiple goals, achieved with different types of studies (Appendix 1). In this guide we are mostly interested with work measurement. The objective of work measurement is to describe the relationship between work inputs and work outputs, and the influence of process variables on that relationship. One may consider many types of inputs and outputs, depending on the goal of the study. In a simple work study, one may focus on mass output and time input. Energy is also a very good choice for both input and output, especially when dealing with energy biomass.

The direct relationship between product output and time input is called productivity. The inverse is called time consumption (per unit product). The variables affecting these relationships are many, and include such factors as technology, work technique, operator skill and environmental conditions. Some of these variables can be managed, while others are passively received.

Determining the effect of process variables on the input-output relationship has many practical uses, such as: setting work rates, scheduling harvesting activities, and comparing technologies or work methods.

3.1 Statistics in work measurement

Unfortunately, process variables come in almost endless combinations, which makes it difficult to determine the specific effect of the variable in which we are interested (target variable). That is where statistics come into play.

Good experimental design and statistical analysis of data allow contrasting the effect of target variables against the general effect of all the other variables combined (Figure 1). Target variables are often called "controllable factors", on the assumption that what is known and predictable can be managed, one way or the other. The other variables are called "nuisance variables", and their effect "background noise", or simply "nuisance". Experimental design offers several techniques to dull background noise, so that the target variable effects can emerge through proper statistical analysis.

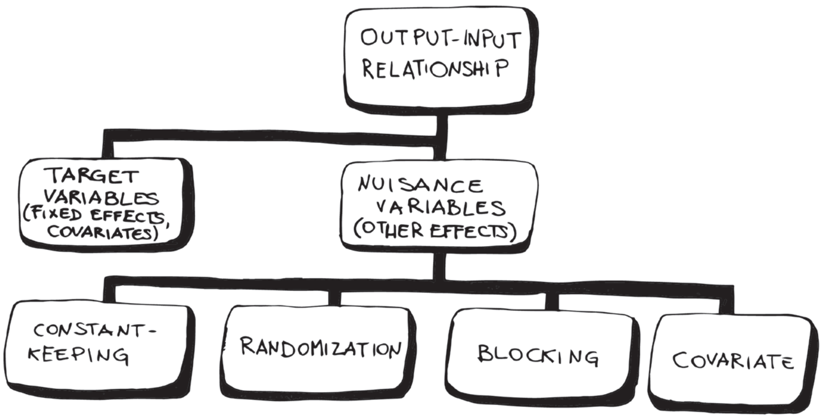


Figure 1. A generic model of a work study system, affected by target and nuisance variables. The lower tier shows the main strategies to dampen nuisance.

All variables can assume discrete nominal values (or levels - e.g. machine A, B and C) or continuously changing numerical values. Variables of the former type are commonly called "factors", whereas variables of the latter type are called "covariates".

3.2 Study types

Work measurement studies can be classified according to their scope, goals and characteristics, with different study types normally requiring different experimental designs and statistical techniques.

The scope of the study may be described after carefully defining the system boundaries. In general a study can concern a single worker, a single machine or a whole system.

As to goal, we can differentiate comparative studies from modelling studies. Comparative studies aim at determining if and how productivity or time consumption is affected by two or more operational alternatives (e.g. machine A vs. machine B). Basically, comparative studies try to disclose the effects of fixed "factors". In contrast, modelling studies aim at determining the effect of continuous variables, or "covariates" (e.g. tree size, extraction distance etc.). Normally, fixed effects are represented as nominal variables, whereas covariates are represented as numerical variables. Studies often involve a combination of comparative and modelling elements, with the specific goal of the study normally determining how the study is classified (comparative or modelling). For instance, modelling is often used in comparative studies of the highly variable forest environment in order to enable comparisons to be made under the same conditions (e.g. normalizing the comparison for the same mean tree size).

Based on its experimental characteristics a study can be defined as either "observational" or "experimental". In an observational study, influencing variables cannot be controlled. That may result in a rather weak study design, which will provide indicative rather than conclusive evidence about the effect of target variables. Many forest work studies are observational in character, yet they offer valuable insights into the studied processes and find their way to the scientific press. In contrast, experimental studies involve a stronger capacity to control process variables, with levels that can be suitably arranged in a strong experimental design. The insights obtained from these studies are stronger and much more reliable than those obtained from observational studies. Simulated environments offer an ideal opportunity to conduct the perfect experiment. Nevertheless, researchers playing with experiments must be careful not to build an experimental design that is too artificial to reflect real operational conditions.

Performing a work study is a complex job, involving several steps (Figure 2).

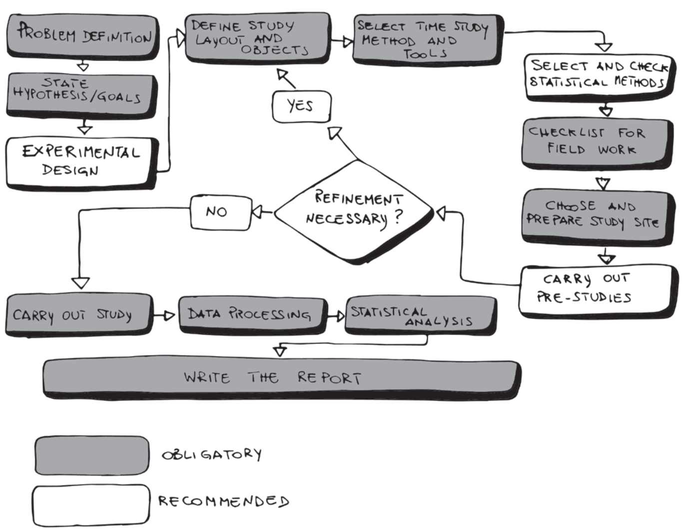


Figure 2 – Flowchart of the different steps required for a work study

4. Before you start

4.1 Study goal

A clear study goal will guide researchers through all steps of a good study. The exact definition of this goal is affected by: 1) specific problem to solve or knowledge to acquire, 2) foreseen use of study results and 3) available resources. For instance, the need for accuracy will differ between a low-budget study aimed at obtaining a rough and ready estimate of expected performance and a large-scale experimental study designed to produce reliable guidelines for official use. However, it is of the outmost importance to spend both time and consideration to formulate the goal, so that it meets both the expectation of the end user and the available budget. A clear goal statement is the foundation of a good study. A hypothesis statement is often part of the goal statement.

4.2 Experimental design

Experimental design is the process of planning a study to meet specified objectives. Planning an experiment properly is very important in order to ensure that the right type of data and a sufficient sample size are available to answer the research questions of interest as clearly and efficiently as possible.

Once the goal has been defined and the hypothesis formulated, the researcher is ready to draw his/her experimental design. The design must fulfill two conditions. First, the nuisance variables should be controlled efficiently and logically. Second, the design should lead to simple analysis, since it is the experimental design that decides how the collected data should be analyzed statistically.

In order to know what nuisance to control, an important step in the experimental design is to list all conceivable nuisance variables. Nuisance can be handled by the following methods: constant keeping, randomization and inclusion. By constant keeping the study is conducted under constant conditions - e.g. only chipping logs of a given size and species, which corresponds to one given nuisance level (where log size represents a nuisance variable in the study). Randomization consists in randomly allocating the treatments to the different levels of nuisance – e.g. randomly allocating the piles of similar, but not identical, material to be chipped. Inclusion is also called experimental control and consists of treating nuisance variables as additional target variables, measuring their levels and associating them to the corresponding levels recorded for productivity, time consumption or any other response variables.

These relationships are then used to correct the analysis. Depending on the characteristics of the nuisance variable included in the analysis, one will talk about “blocking” (for variables assuming fixed levels) or “introducing a co-variate” (for continuous variables).

Experimental design addresses the questions of how to combine the study treatments with the possible methods of nuisance control in such a manner that no treatment is systematically favored.

Comparative studies will normally use some kind of factorial design. The general similarities in factorial designs are that a certain number of repetitions are conducted for each combination of treatments and blocks (Figure 3). It is advisable to aim for equally many repetitions for all combinations, i.e. a balanced design. That facilitates simple and valid analysis. However, a balanced design is not compulsory in the analysis, and especially not so with large number of repetitions. The capability to accommodate for unbalance is very helpful, since unbalance often occurs due to unforeseen events during the experiment.

Box 1 – Example of goal statement and hypothesis statement
The goal of this study was to compare the technical and economic performance of terrain chipping and roadside chipping, applied to short-rotation biomass plantations. The null hypothesis was that there was no significant difference in the performance of the two work systems, when applied to short-rotation plantations.

A similar approach will be followed in modelling studies, with the difference that the focus is on how measurable and/or quantifiable nuisance factors of interest influence one or many controllable factors (treatments). Thus, instead of exerting passive statistical control on the nuisance variables, the ones of interest should be actively selected so they vary within a predefined range. To improve analysis, the number of observations should be balanced within the range of variation.

Each repetition of the same experiment is also called an observational unit, sample or replicate. The reason for observing several units that receive the same treatment is the expected variation in both response to treatments and in measurements. Thus, it is crucial to define the observational unit (i.e. what should be replicated) and the required number of repetitions. This can be calculated based on information about: expected mean value, sample variation and desired (statistical) accuracy.

The procedure for such calculations varies with the type of experimental design, and can be found in standard statistical textbooks. However, we include the example below as a reference. The equation is used as a basis for determining sample size in harvesting studies conducted at the cycle level (Murphy 2005):

number of replications = $t^2 * V/(E*Mean/100)^2$ [1]

where: t = Student's t-value (= 1.96 for a 95% confidence interval ⇨ t^2 = 3.842)

V = expected variance of work cycle time

E = level of precision required (e.g. 5%,)

Mean = expected mean of work cycle time

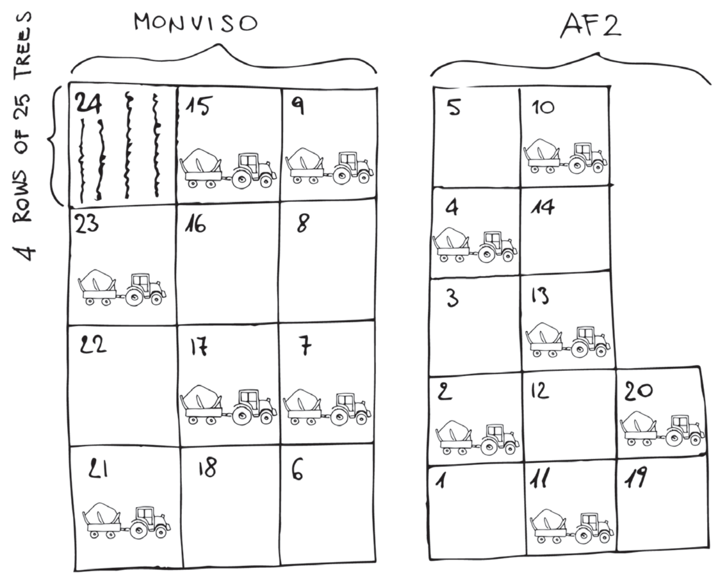


Figure 3 –Design of an experiment for comparing terrain chipping vs. roadside chipping in a short-rotation poplar plantation. Plots are allocated randomly to the two treatment levels (i.e terrain chipping = blank plots; roadside chipping = plots with a forwarder symbol). The experiment is blocked for two main clone types, i.e. Monviso and AF2. It is therefore a factorial 2 x 2 design, where each of the 4 treatments is repeated 6 times (i.e. total of 24 replications). Hence the design is balanced.

Although not formed by any kind of natural law, the precision level is generally set to 5%, which means that the researcher is willing to accept a 5% risk that the hypothesis is incorrectly evaluated.

The expected mean value and variation is less easy to quantify in ad-

Box 2 – Experimental design: example 1

We want to determine the difference between two chipper models (= one treatment with two levels) operated by three different operators (= one blocking factor with three levels) under identical (i.e. similar) conditions. Therefore, each operator will work with each chipper, so that we shall have 2 x 3 = 6 combinations. This is a factorial design. We decide to conduct 5 repetitions per combination, so that the total number of samples will be 6 x 5 = 30. The order in which we shall distribute samples should be random. So operators shall switch between machines randomly, until we have completed 5 repetition of each of the 6 operator x machine combinations. Under real conditions, a pure random sampling might be inconvenient, so that one may assign operator x machine combination randomly, but conduct the five repetitions sequentially for each combination. This procedure is formally incorrect, but it is often accepted if one can find other measures to mitigate the error deriving from sequential repetition, or if it can be rationally explained why it is expected that such violation of good practice does not represent a main source of error.

vance of the study. Ideally, pre-studies could be used to provide appropriate information for the calculation. Otherwise, this information can be obtained from previous similar studies. In fact, the decision about the number of repetitions is often based on educated guesses. However, this is related with the risk of having too few observations for detecting any differences (large variation compared to the size of the treatment effect) – or to spend excessive resources on too many observations (very limited variation and large treatment effect). Within a given experimental design, the sum of the samples (observational units) in all treatment combinations results in the total number of samples to be studied. This number is an important consideration in determining whether or not the stipulated experimental design will fit the study budget. If too resource demanding, alternative designs will have to be considered and/or the number of samples will have to be decreased. Thus, the actual experimental design will not only be the result of a preferred accuracy in statistical analysis, but also of the budget constraints.

4.3 Formulating a statistical model

Since it is the experimental design that decides how the collected data should be analyzed statistically, it is natural that the researcher should be aware of what statistical methods and models should be used for the chosen design. Thus, it is important to formulate the statistical model that will be used to analyze that data. If hypothesis and experimental design are well aligned and the experiment includes few treatments, blocks and co-variates, this is quite straight-forward. For instance, a design with one treatment under constant conditions would have the statistical model of:

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \qquad [2]$$

in which y is the response variable for observation i within treatment level j (e.g. time consumption for a given log chipped by chipper A), μ is the grand mean (total mean value), α is the main effect of the treatment j, and ε is the random error. This model would be evaluated in a one-way analysis of variance (ANOVA). If the study employs statistical control of, for instance, differences in log sizes, a co-variate is added to the model, according to

$$y_{ij} = \mu + \alpha_j + b \times x_{ij} + \varepsilon_{ij} \qquad [3]$$

where b is the slope of the co-variate x. This model would be evaluated in a one-way analysis of co-variance (ANCOVA).

4.4 Defining what to measure and how

Box 3 – Experimental design: example 2
When aiming to model the influence of log size on chipping productivity for a given machine, the design would be to define a range of log sizes to be studied and the total number of logs to be chipped. Ideally, the number of logs should be spread evenly along the log size range (instead of having 95% small logs, 2% medium sized logs and 3 % large logs).

The object of work studies is the relationship between work inputs and work outputs, and its reaction to the effects of process variables. A well planned and implemented work study will determine the inputs, the outputs and the process variables, trying to define their possible relationships with statistical methods. In particular, a work study will require that all of the following objects are measured:

4.4.1 Inputs

Being the characterizing element of work studies, time is obviously the very first object to be measured in a time study, e.g. the time input per cycle, per cycle element or per plot. Another crucial input to be measured is the energy used by the process under study. This is especially important when the object of the process is the manufacturing of an energy product – such as biomass fuel.

4.4.2 Outputs

As work is assumed to produce outputs, these outputs should be determined with sufficient accuracy for productivity studies, since productivity is defined as output divided by input. Outputs are both quantity and quality, equally essential to evaluating any work method and/or technology. In forest work, quality concerns two separate entities: product and environment.

Product quality is evaluated by comparing actual product characteristics with market specifications. In the case of fuel chips, for instance, these are: moisture content, particle size distribution, contamination level etc. The environmental quality of a given work technique is generally defined by its stand and soil impacts. Hence, the main outputs of a forest harvesting process are: product quantity, product quality, environmental quality.

4.4.3 Process variables

Process variables that may affect time consumption and/or productivity should be determined with accuracy, both in comparative and modelling studies. In the former, determining work conditions is crucial to ensure that all alternative treatments are applied under the same (controlled) conditions. In the latter, the validity of any model will depend on the accurate determination of the affecting independent variables.

Not all these objects must be measured in every study: object inclusion and measurement accuracy will be tailored to the goal of the study, and to the resources allocated to conduct it.

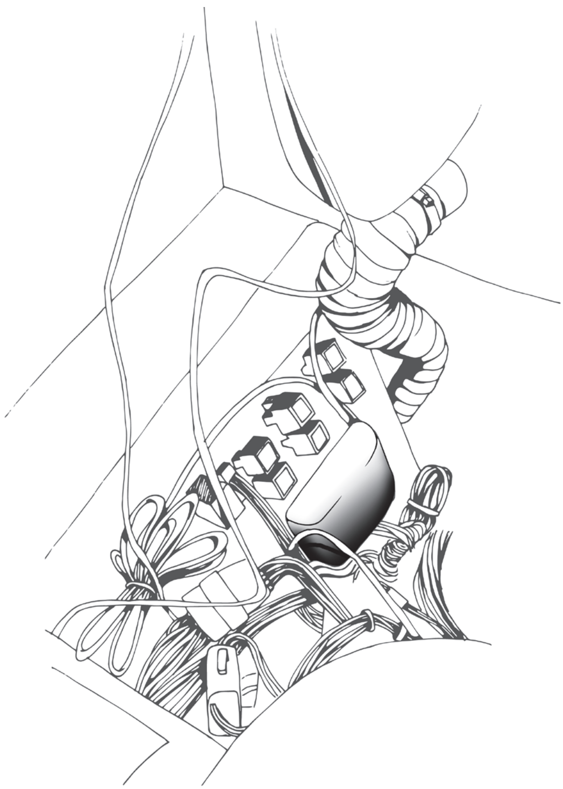
At the outset, the research team should define the goal of the particular study and determine the variables that must be recorded in order to meet these objectives.

If the study involves modelling, the researchers should generate a matrix of the relevant dependent variables (such as cycle time elements and production) and independent variables, anticipating which of the latter may influence the former. In Appendix 2 readers can find an example of the many variables that could be explored, when studying a range of different harvesting techniques.

4.5 Practical rules

4.5.1 Safety

Safety is the first and foremost requisite for all work activities, coming before productivity and environmental quality. Field researchers should always make sure that they are not exposing themselves and others to unnecessary risk. They have the legal and social obligation to comply with all safety requirements. If they are working in an environment still relatively casual about safety procedures, their obligation is also moral, because they can set an example that will help introduce a safety culture where this is badly needed. Regardless of how confident we are in our capacities, many operators look up to us because of our education and status and we should set a useful example that may save lives. Whenever entering an operation, researchers shall always:



Commercial GPS-tracking black-box unit

- wear high-visibility clothing (jacket or vest);
- wear a hard hat (with hearing protectors and visor when needed);
- ask the operator/s about the work routine, the safe zones and the risk zones, so that the researcher will always stay away from risk zones and inside the safe zones. Whenever losing sight of the researcher, the operator should stop work immediately;
- agree with the operator/s on a system of communication, so as to quickly and unambiguously transmit urgent information (e.g. radio phones);
- abstain from drinking alcohol and/or taking any drugs that may impair

one's alertness and judgment;

- and, more generally speaking, be compliant with any other requirements coming from regulations or internal rules adopted by the companies in charge of the forest operations.

Special attention should be paid when climbing into and out of containers to collect chip samples, as container edges are tall and slippery. When climbing is necessary, that should be done with caution, using the steps normally fitted on most containers.

Work at a landing can often be observed from a fixed station, including the researcher's own car, appropriately parked in a safe zone where it does not hinder the operation. This can offer much relief under rainy and/or cold weather conditions.

When studying felling, processing or harvesting machinery, a safe distance should be maintained in order to minimize the risk of injury in case of uncontrolled tree fall or saw chain breakage. In certain forest stands, safe distance may make it impossible for the researcher to observe the operation in such a detail as required for the study. The only safe place within the safe zone is in the machine cab, and it may happen that the researcher rides in the cab together with the operator. However, that it is not advisable unless the cab has been designed to take a passenger. Otherwise, the eventual passenger may not fit inside the internal survival volume remaining after a possible roll-over or impact. In such instances, researchers should consider remote data collection, as allowed by video-recorders, on-board computers and commercial GPS-tracking black-box units.

Field study researchers work outdoors and should take all precautions required by outdoor work, including: wearing appropriate clothing (comfortable, rainproof, warm, fresh - according to need); wear work boots or similar shoes; carry their own supply of water and food, as needed; use insect repellent or carry an appropriate weapon if harmful insects or animals may cause danger or discomfort; and get the appropriate vaccinations for any diseases that can be contracted in the specific work environment (tick-borne encephalitis, tetanus etc.).

4.5.2 Ethics

Work studies often represent an intrusion into the personal working space of individuals, crews, and enterprises. Studies can be initiated by companies that wish to know more about their own operations, by machine manufacturers that wish to test or enhance their design, or by researchers who wish to investigate one or more aspects of machine or system performance in an applied setting. Studies are almost never initiated by workers themselves.

Scientific management has its roots in an exploitative era characterised by rapid industrialisation and a need to quantify efficiencies and costs. While work studies are important, the integrity of the personal work space is protected by collective agreements, legislation and common respect.

Through work studies, the researcher unavoidably gains insight into the physical and intellectual capacities of the subjects involved. In motor-manual work this is explicit in the form of heart and lung performance measurement. In mechanised operations it might be the subjects' decision making or concentration ability.

This short passage cannot deal with the complexities of labour law and human rights in the 34 participating countries. It is a simple reminder that researchers should be familiar with the legal and ethical framework within which they operate. Work studies should be founded on dialogue, trust and confidentiality. Subjects should be made aware of the purpose, methodology and intended use of the results. Their consent should be obtained beforehand. The field of ethics also includes the relationship within your own study team and with other research teams. These are treated extensively in the USDA Code of Scientific Ethics, which provides an excellent example and is freely available on the Internet http://www.fs.fed.us/rm/analytics/ethics.htm

Similar attention must be paid to the relationship with the customer and all subjects involved in the study, and to the possible confidentiality obligations imposed on sensitive data.

5 Measurements in the field

5.1 Measuring time input

Time consumption measurements offer different resolution depending on whether they are conducted at the shift, cycle or elemental level.

5.1.1 Plot level

In plot level measurement the observation unit consists of a single plot, like those described in Figure 3. Hence, all the time input necessary for harvesting the plot is cumulated. Time input is measured directly by the researcher observing the operation, or automatically by appropriate sensors connected to a data logger. Data can also be recorded manually by a cooperative operator appropriately instructed.

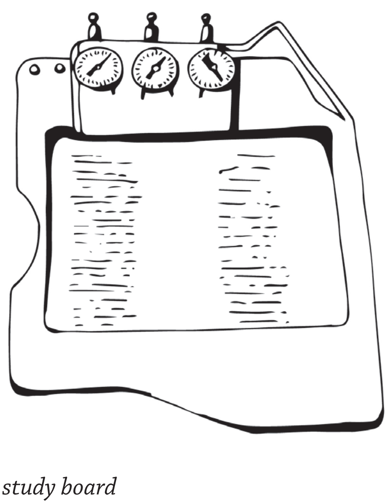
5.1.2 Shift level

Shift level measurement implies that the observation unit consists of a whole shift, whose duration and organization should always be indicated (e.g., 8 hours total time, including 6 hours of actual work and two hours of maintenance). Shift level measurements are conducted manually or automatically. In the former case, operators are given data collection forms and are instructed to note daily on the forms data such as: date, place, job type, starting and ending hour, estimated output (turns, trees, m^3 etc.), fuel consumption and any major delays.

The cause and estimated duration of all delays should also be noted. Much of the same information can be collected automatically through appropriate sensors, connected to a data logger. Most dedicated harvesters are already fitted with the necessary equipment to capture these data, for the purpose of operational optimization and cost control. Shift level measurement is generally the main technique used for long-term follow-up studies aimed at determining machine utilization, long-term productivity and incidence of delays.

5.1.3 Cycle level

In cycle level measurement, the observation unit is a single work cycle (e.g. the felling of a tree, the forwarding of a load etc.). Compared to shift-level measurement, cycle level measurement offers more detail and can help describe the work process with much more accuracy. It also helps identify the variability of a work process very quickly. Individual relationships can be isolated that could be difficult to pinpoint with shift level measurements. A number of different tools can be used for manual measurements of time consumption at cycle level, including: standard wristwatch, stopwatch, stopwatch board, or hand-held computer. All these instruments can determine the time elapsed



Time study board

between the start and the end of a previously defined work cycle, and this value is noted on paper or voice recorder by the researcher, or stored in the computer memory. Time consumption can also be captured automatically, if an action or a sequence of actions defining the start and the end of a work cycle can be identified by appropriate sensors. These can transfer captured data to the storage of an on-board computer (if fitted) or to an add-on external storage. The MultiDAT system developed by FP Innovations (formerly FERIC) and distributed by Castonguay Electronique Inc. in Canada is an example of a proven automated data collection system. Furthermore, modern fleet control and management systems offer similar capabilities and could be used for the purpose of collecting time and motion data.

5.1.4 Element level and work sampling

Element level measurement consists of splitting the work cycle into functional steps (elements) and then recording time consumption separately for each of them. This allows the work process to be described in more detail, which may contribute to a better understanding of process dynamics. In particular, 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, 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 more accurate sub-models can be developed.

Collecting time data

Elemental measurement is conducted with the same instruments listed in the previous paragraph. When very short elements must be captured, one may resort to video recording: elemental time consumption is then measured in the lab, using the slow/pause function and the time stamp of the video recorder. In all

Box 4 - Subdivision of cycle time into functional elements

As an example, one may consider subdividing a chipper cycle (defined as the process of filling up a container of known volume) into the following time elements:

- Moving the chipper along the wood pile or between adjacent wood piles. Starts when the outriggers are lifted off the ground and ends when they are firmly positioned into the ground at the next chipping station.
- Parking the container near to the chipper. Starts when the chipper is still, waiting for the container to be placed by its side and ends when the chipper begins chipping again.
- Chipping. Starts when the first wood load is moved to the chipper infeed and ends when no more wood is being fed to the chipper.
- Other work. Any other work process (e.g. piling, handling wood with the loader etc.)
- Delays. Any interruption of the work process (See next box).

When subdividing a work cycle into functional steps it is important to resist the temptation of producing too many time elements, since that may detract from recording accuracy, increase the possibility of errors and complicate experiment replication by others.

It is important to remember the purpose of the specific study and the main goal of elemental breakdown, which is the separation of process steps that are differently affected by different independent variables and/or require different improvement measures.

Separating more elements that are similarly affected by the same variables and/or improvement measure will produce no practical benefits. It is also helpful to look at other studies for ideas on elemental breakdown and on relevant variables to measure.

cases, it is crucial that the actions marking the beginning and the end of each functional step be clearly defined and described, so that the study can be interpreted and eventually replicated by fellow researchers. If more actions occur at the same time, they will overlap.

In this case, we shall have three options, depending on the goal of the study: 1) we may record their separate durations; 2) we may define a new combination element; 3) we may decide for a priority system that allocates the overlap time to one of the two separate activities. Let us consider the case of a feller-buncher handling cut trees while rolling on its tracks. If we need to separate the two activities, we shall ask a colleague for assistance so that two separate persons will time the two functions separately.

Box 5 - Delays

Delays are interruptions of the work process and are commonly subdivided into three main categories depending on their origin. Mechanical delays are caused by the need to service or repair the machine used for performing the work task. Personal delays are interruptions caused by the operator, and include rest breaks. Operational delays are related to organizational causes, such as a poor balance between the chipper and the supporting units (waiting) or an excessive concentration of machines on the same track (traffic), work planning and site reconnaissance.

A fourth delay type is represented by interruptions of the work cycle caused by the study itself (study delays). These are generally excluded from the analysis. The subdivision between evitable and inevitable delays requires a subjective judgment and should be discouraged. The main problem with delays is their large variability, due to erratic occurrence.

A reliable estimate of delay time (or overall time including delays) will require a very large number of replications and a comparably long observation time. Therefore, two main solutions have been devised to overcome this problem: 1) including into the study only those delay events that fall within a maximum duration limit (e.g. 10 or 15 minutes); 2) excluding delays from data recording and accounting for delay time through specific delay coefficients applied to productive time. The first strategy tends to underestimate the incidence of delays.

For example, long-term studies of chipping operations have shown that delay events shorter than 15 minutes represent over 80% of the occurrences, but only 32% of the total delay time. The second strategy is based on conducting a long-term follow-up study of the operation or on combining a large number of detailed time studies into a larger data pool, in order to extract the long-term incidence of delay time (delay coefficient). This should also be expressed as a percent of productive work time, and not of worksite time, since the latter form is flawed by inter-correlation.

The same result will be obtained by videotaping the operation and then playing the tape twice, so that one person can record the two separate times. Otherwise, we can define a special combination element (e.g. “handle and roll”) to contain overlap time.

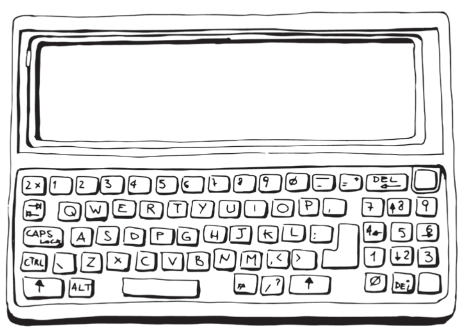
Finally, we can allocate overlap time to either the “handle” or the “roll” function. The function with priority will be the one clearly appearing in the study, whereas the other function will be "masked".

The description of time elements should clearly report which functions may overlap, and which will have priority when overlap occurs. Priority will be attributed on the basis of the study goal. If for instance our feller-buncher study aimed at predicting track wear by determining how much time a feller-buncher spent moving, then the "roll" function should receive priority and the "handle" function would be "masked" whenever the two occurred together.

Elemental time is recorded with two main techniques: continuous timing and snap-back timing. With continuous timing, the time of each element shift is noted, and the duration of each time element is calculated by sub-

Box 6 - Handhelds computers

Time study data can be collected with handheld computers running dedicated time study software. These computers are often ruggedized, so that they can withstand the outdoor forest environment. Different machines are used by different groups, but the most common are: the Husky Hunter (and subsequent models) running the dedicated Siwork3 time study software, which is still very popular in many English-speaking countries, as well as in Denmark, France and Italy; the Latschbacher family of field computers, widespread in Austria and Germany; the Rufco 900 in Finland. All these machines are relatively old and at times they present interface problems with modern laptops, as most of them use serial connection ports that have virtually disappeared from new personal computers. Modern potential replacements are new portable machines such as Allegro, Psion, Ranger and Toughbook. Interfacing potential, software availability, reliability and battery life are the main parameters to consider when choosing a handheld for time study purposes.



Husky Hunter hand-held computer

tracting two time marks (i.e. the time when the function was completed minus the time at which it was initiated). This technique requires back-calculation but is the only viable option when using a wristwatch for timing. Snap-back timing consists in restarting from zero at each element shift. That is done using the “lap” function available on most stopwatches. The advantage of this method is that one does not need any calculations to obtain the net elemental time.

Box 7 – Hawthorne effect
It is a well-known phenomenon, where workers modify their behavior just because they know that they are being studied. This may determine performance increases (or decreases) that are not caused by the technical changes introduced with the experiment. The Hawthorne effect may introduce a significant bias in short-term work measurements. For this reason, reliable productivity levels are best determined with long-term follow-up studies, or by analyzing long-term production statistics.

Work sampling (also known as frequency study) is another technique for measuring the elemental breakdown of time consumption. It consists of observing the 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. At the end of the study, the researcher will obtain a total time (duration of the study) and a relative frequency[1] of the different functional steps – which is one of the outputs expected of any elemental time study.

The advantage of work sampling is that it allows one researcher to follow more teams at a time, by organizing a sequence of observation intervals for the different teams. The disadvantage is that work sampling does not offer any information about cycle duration, since the observation interval cannot be synchronized with the variable duration of the work cycle. In fact, one should carefully avoid the synchronization of observation interval with cyclic work, which would return a biased representation of cycle time distribution.

Irregular sampling intervals are preferable to regular intervals, because they exclude the accidental synchronization with cyclic elements. Work sampling is often used for quantifying equipment and people interaction delays within a working team or work system.

1 *whence the alternative definition of “frequency study”*

5.1.5 Units

Time consumption is generally measured in hours, minutes and seconds, depending on the resolution of the study. Occasionally, very short process steps can be measured in smaller units, like the tenth of a second. Work studies are often conducted with clocks that measure minutes and centiminutes – i.e. hundredths of a minute – instead of minutes and seconds. That is a compromise aimed at transforming into a quasi-decimal system the traditional sexagesimal time measurement system. It elects the minute as the most representative unit and breaks it into hundredths, in order to simplify the eventual data processing (by allowing decimal calculation). It is a very effective measure, even in a time when computers can easily transform sexagesimal records into decimal records, and the reverse. However, the second is the only SI unit for time measurement, although the hour and the minute have been officially accepted for use with the International System[2]. Hence, time study data can be reported in scientific publications in any of these three units, whereas the use of centiminutes (cmin) could be rightly opposed by reviewers. In that case, an acceptable formulation could be min*10^{-2} or 1/100 min

5.1.6 Classification of time in forest studies

Time consumption can be subdivided and/or grouped according to the role of the specific work steps within the whole process. A number of classifications have been produced over time, generally deriving from the work of the Nordic Research Council and the American Pulp and Paper Association. Most previous classification efforts have already been consolidated and partly harmonized within IUFRO, and the best synthesis is still offered by the IUFRO Forest Work Study Nomenclature, published in 1995. In Section D, the IUFRO document presents a clear and comprehensive classification of time in forest work study.

That same classification is reported in Appendix 3 of the present manual, and adopted for the purposes of our good practice guideline without any further changes.

Some scholars question the subdivision into time elements, because of their possible inter-correlation. Correct statistical theory requires that variables are independent from each other. Hence the statistical treatment of separate time elements would be incorrect if these were found to be inter-correlated. In that case, it would only be correct to analyze

2 *Tab. 6, Page 105 of The International System of Units (S I) 1998, 7th edition 1998, Organisation Intergouvernementale de la Convention du Mètre.*

total time consumption as a whole. However, elemental time studies are still popular and useful.

5.2 Measuring energy input

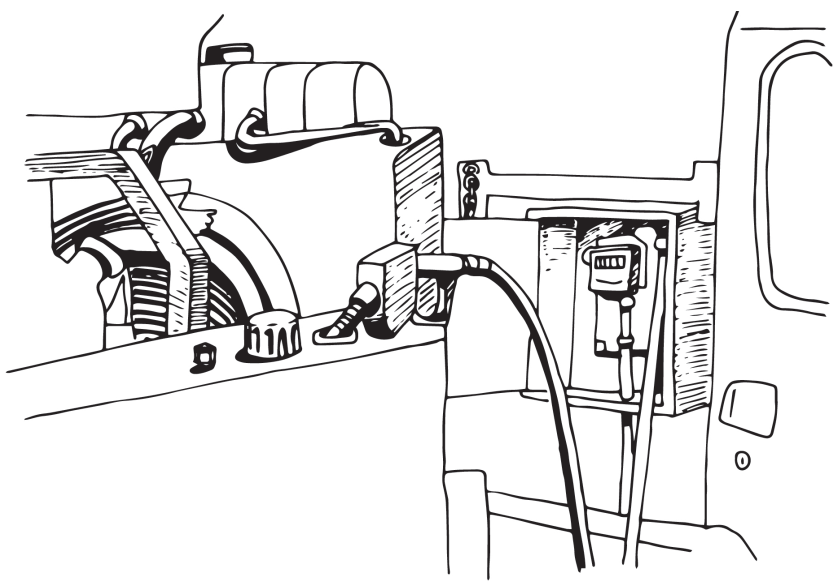
Reliable measurements of the energy used for the supply of energy biomass are crucial to the compilation of Life Cycle Analysis (LCA) studies, and ultimately to the formulation of policy suggestions. The work-study researcher is eminently well situated to provide accurate data on energy inputs. Direct energy consumption is normally measured by recording fuel consumption and then converting it to energy units through a constant that represents the energy content of the fuel.

Fuel consumption studies can be carried out at various levels of resolution, depending on the goal of the study. Table 1 lists the main techniques, with their pros and cons.

Any essential fuel consumption study should always provide at least the following data:

- Engine model, make, year of manufacture and displacement (cm^3);
- Litres or kilograms of fuel used for the duration of the study;
- Amount of biomass produced for the duration of the study.

Additional information could include: duration of the study, engine information relating to emissions (Euro standard or American tier system), etc.



Fuel is a main energy input

Measurement resolution	**Technology / Method**	**Pros & Cons**
Continuous	• Onboard flow meter (factory fitted) • Onboard flow meter (fitted for research)	☺ Coupled with electronic impulse from hydraulic valve bank, it allows consumption analysis on separate work elements, e.g boom movement, driving, chipping. ☹ Requires relatively sophisticated equipment
Operation or Shift level	• Onboard flow meter Standard flow meter on new machines provides accurate current and shift level data. • Pump-fitted flow meter Flow meter on an electric or manual pump is used to record fuel volume during refuelling. • Scale Scale can be used to weigh fuel before refuelling	☺ Shift or operation level data can provide more robust information, evening out erratic peaks. ☺ Requires less intensive observation / data management ☺ Short term – retain machine operator enthusiasm ☹ Lose individual work elements in the analysis ☹ Motivation - often requires that operator must record and log data alone
Daily or weekly	• Onboard flow meter • Pump fitted flow meter • Scale	☺ Similar to above, but less frequent measurement is required. ☹ Reduced accuracy in relation to outputs ☹ Not easy to keep operator motivated to fill forms - serious corruption of data if error or omission occurs
Monthly, quarterly or yearly	Typically based on data obtained from • Onboard flow meter • Fuel issued to machine (accounting system must identify machines) • Fuel purchase details from bulk supplier (for single machine) Total periodic fuel consumption averaged by machines	☺ Often reasonably accessible data (due to accounting laws) ☺ Robust data covering a wide range and depth of observations ☹ Cannot be used to develop specific models on operations ☹ Risk of data loss or corruption

Table 1 – Techniques to measure fuel consumption

5.3 Measuring product output

Productivity studies require that time consumption be associated with a product output, in order to determine the following relationships: Productivity = Product output/Time input; Specific time consumption = Time input/Product output. Specific energy use = Energy input/Product output. Product output can be measured using the different units listed below.

5.3.1 Count

Output can be measured just by counting the units produced, such as trees, logs, bundles, grapple loads or container loads. Unit count is a very approximate measure, which makes sense only when the units have a regular standard size, which can be easily quantified into mass or volume figures.

5.3.2 Solid Volume

Solid Volume is a very reliable entity for estimating biomass output. Once the measurement technique and the eventual inclusion/exclusion of bark is defined, the measurement is relatively robust. Solid volume can be measured with caliper and logger's tape, using many different techniques (Huber, Smalian etc.). Otherwise, it can be determined using the harvester measurement system, provided this is correctly calibrated. Another method to estimate solid volume consists of using volume tables, which return tree volume as a function of tree diameter and tree height. In this case, field measurements are limited to the diameter at breast height (DBH) of the trees to be processed, and to a certain number of tree heights, necessary to develop a diameter-height curve. However, all these methods will produce solid volume estimates for the stem and the main branches only, excluding the volume of smaller branches and twigs. That can be accounted for by using an empirical biomass expansion factor (BEF), which increases the stem volume estimate by a certain percentage, reflecting the contribution of smaller branches[3]. However, BEFs will only provide approximate estimates, partly defeating the benefit of using solid volume as the reference for estimating biomass output.

5.3.3 Bulk volume

Bulk Volume (or loose volume) is the volume physically occupied by a

3 See: Teobaldelli M., Somogyi Z., Migliavacca M., Usoltsev V. 2009 Generalized functions of biomass expansion factors for conifers and broadleaved by stand age, growing stock and site index. Forest Ecology and Management 257: 1004-1013.

certain quantity of biomass. This unit is often used with small logs, firewood and chips. Loose volume is very easy to determine, since it just takes a tape to measure the volume of the log stack or the internal volume of the chip container.

Loose volume can be converted into solid volume or weight by using appropriate coefficients, which should be estimated case by case with sampling. Although very easy to use, loose volume offers a somewhat approximate estimate, since the actual product mass will vary with the size and the form of the individual elements forming the stack or the pile. Furthermore, different chippers may “pack” chips with a different power, thus producing a more or less compact load, even for the same particle size distribution. Finally, loose volume can be determined with good accuracy only if the stack or the container has a regular shape, whereas determining the loose volume of chip piles may be difficult and will return approximate estimates.

5.3.4 Fresh weight

Fresh weight (or green weight) is considered the most direct measurement of actual mass output. However, its correct determination requires the use of accurate scales, often unavailable in the surroundings of the study site. In that case, loads can be scaled at delivery (they generally are) and their weight can be transmitted to the researcher, providing that each load is clearly and unambiguously identified. As an alternative, one can use portable scales of different types, applied to the loader

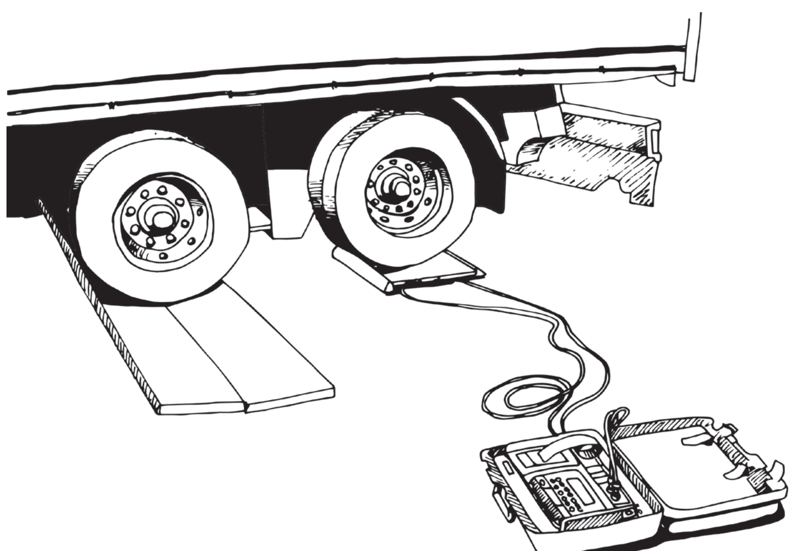


Plate scales used for axle weighing

boom or installed on large metal plates and used for axle weighing. Both methods can offer good results, providing that the plates are correctly calibrated, that they are placed on solid level ground and that all axles are at the same level when weighing.

5.3.5 Dry weight

Dry weight offers a better representation of true product value compared to fresh weight, because it excludes the inevitable contribution of water to mass output. Dry weight is an indirect measure obtained from fresh weight, after determining moisture content. Existing European standards define the methods for sampling (EN 14778), sample preparation (EN 14779) and moisture content determination (EN 14774). The accuracy of dry weight estimates will be affected by the errors accumulated during sampling.

Box 8 - Which units should one use to measure product output?
That is indeed a big question. All units have their pros and cons, and may be adopted depending on the goals and the circumstances of the study. Whenever indirect measures are provided (i.e. dry weight and energy content) it is essential that the researcher reports: the methods used for estimating them; the values actually measured for fresh weight and moisture content; the equations and parameter values used in the energy calculations. An indication of variability of the direct measurements would also be useful.
Please notice that the term "weight" applied here is formally inaccurate: what we are really measuring is a physical quantity defined as mass. However, understanding with managers and operators will be easier if we use the term weight, rather than mass. Hence, it is convenient to use "mass" in scientific papers and "weight" in everyday speech. After all, language is a convention.

5.4 Measuring energy output

Energy Content is another indirect measure of output value, and it has the merit of indicating the actual value for the end user, which simplifies communications with plant engineers (who will call it "lower heating value").

Energy content is obtained by multiplying dry weight for an appropriate energy density coefficient, then subtracting the energy absorbed by the water inside the product. A typical example for hardwoods would be the following one, which returns Giga Joule per metric tonne:

GJ/t = dry weight, t * 18.5 GJ/t – water weight, t * 2.5 GJ/t

This estimate is indirect and based on coefficients, and its accuracy is affected by the reliability of the coefficients and the eventual error with moisture content determination.

5.5 Measuring quality output

Job quality reflects on product quality and environment quality, the latter defined as the impact on the stand and the forest soil. In fact, environmental quality is a complex concept, going far beyond a simple determination of direct stand and soil impacts. However, determining the full extent of environmental impacts exceeds the scope of simple work studies.

5.5.1 Product quality

Product quality will be estimated in different ways, depending on target specifications. For the manufacturing of logs, measurement accuracy and superficial damage could be important quality indicators. The former will be checked with tape and caliper, the latter through visual inspection, or by capturing and processing digital pictures with image analysis software in order to estimate surface damage with more accuracy. In the case of chips, the actual work process can impact product quality especially for what concerns contamination and particle size distribution. Contamination with soil and stones can be estimated visually, or by separating wood and contaminants (manually or with the help of a solvent) and determining the weight ratio. Particle size distribution should be determined according to European Standard EN 15149.

5.5.2 Stand impacts

Stand impacts are generally determined by inspecting the residual stand after harvest, in order to detect and catalogue any eventual damage caused to residual trees and/or advanced regeneration. Inspection is generally conducted on sample plots of varying shape and size. The number and size of sample plots will be determined as a function of sample variability and desired accuracy. Tree damage is generally attributed a severity class, often related to wound size, type, position and depth. Superficial wounds smaller than 10 cm^2 are often neglected, since they do not seem to affect tree health, growth rate or wood quality[4].

4 *Whitney R. 1991. Quality of eastern white pine 10 years after damage by logging. For. Chronicle 67: 23–26.*

5.5.3 *Soil impacts*

Soil impacts are determined in a number of different ways, from very simple to very sophisticated. We may want to stick with the simple methods, assuming that the sophisticated methods will only be deployed for studies specifically devoted to the analysis of logging disturbance, and belong to soil science more than to work science. For our purposes, simple visual inspection could be enough, and it can be conducted scientifically with a standard method, such as that described by McMahon (1995), which is rapid and easy. According to this method, the harvest site is covered by a regular grid of inspection points with a mesh size sufficient to obtain the desired sampling intensity, on the basis of expected sample variability and desired accuracy. Then each point is visually inspected and attributed a predetermined disturbance class. As a result, one will obtain a reliable estimate for the frequency of different visible soil disturbance phenomena.

5.6 Measuring process variables

Process variables that may affect time consumption or productivity should be determined as accurately as possible, both in comparative and modelling studies. Such variables can be grouped in the following three large categories.

5.6.1 *Physical environment*

Terrain and forest characteristics have a major effect on work perfor-

Time study on a biomass operation

Box 9 - Measuring extraction distance

Extraction distance is a key independent variable in biomass extraction studies. Distance can be measured with several instruments, including: tape measure, hip chain, pacing, laser range-finder, machine odometer, GPS, map coordinates. It is always very important to indicate how distance was determined and if the distance reported is the map distance or the actual slope distance. It is also important to specify whether extraction occurred uphill or downhill.

mance, and can be described by a number of different indicators, generally pertaining to the fields of forest mensuration and topography. For the specific purpose of forest operations, terrain characteristics can also be described using the Swedish Terrain Classification System[5], which is widely adopted in Scandinavia, as well as in Ireland and in the United Kingdom[6] – with local variations. The system is simple, and the original manual offers reference pictures for the evaluator. It produces a single synthetic indicator capable of describing slope gradient, terrain roughness and ground bearing capacity.

When describing the physical environment it is important to distinguish between those variables that are essential for the study and those that are not, although potentially useful. The failure of many attempts at developing general data collection protocols is likely to rest in the overabundant requirements of such protocols, which put an unacceptable burden on the researcher, often constrained by budget or time limitations. Therefore, it may be better to restrain data collection to those variables that are most likely to affect the performance of the work process under examination. For instance, there is little need to define ground slope or residual stand density if the study concerns a chipper working at a landing.

In this case, the measurement may include average piece size (easily obtained by counting the number of pieces needed to fill a container of known volume or weight), landing surface, tree species and tree part (branches, logs, whole trees etc.), which have already been shown to have a significant effect on chipping performance. Other forest compartment data can help describe the general background of the experiment and are welcome if they come for free, but can also be omitted without much prejudice to the quality of the research.

5 *Berg S. 1992 Terrain classification system for forestry work. Skogsarbeten, Kista, Sweden. 28*

6 *UK Forestry Commission. 1995. Terrain classification. Technical Note 16/95. 5 p.*

In general, one should record with greatest accuracy all those characteristics of the physical environment that could be used as independent variables in the eventual data analysis, such as tree size in felling studies or extraction distance in forwarding studies.

5.6.2 Organization

Operation layout and work organization have a strong influence on productivity and time consumption. In general, it is enough to provide a simple description of how the whole operation is organized, how many units and crew members are involved, and what are their specific tasks. If the work organization generates specific risk for operator safety (e.g. interference between operators), it may be useful to identify this risks and suggest solutions.

Operator experience, skills and motivation have a major impact on productivity and time consumption. "Operator effect" has been shown to affect productivity for up to 40%, which accounts for the gap between the inexperienced and the very experienced operator. Ideally, work studies should be conducted with many different operators, in order to integrate operator variability into the study design. This often conflicts with the time and the financial constraints of most research projects. Operator rating would offer a practical way to deal with operator variability, and could be conducted with a number of different methods, often accurately codified. Unfortunately, no current method offers both easy application and objective evaluation, so that operator rating is either too complex or too subjective. For this reason, most researchers have discarded operator rating, and prefer to use general habilitation criteria, based on operator background. That leads to the exclusion of any operators considered inexperienced, unwilling, clumsy or slow. Evaluation is done by examining the operator work history, interviewing the operator and his/her colleagues and supervisors, and observing the operator at work. These precautions will not prevent operator effect from causing some variability in the results, but are most likely to contain the eventual error within acceptable limits.

The payment system can have a strong effect on operator motivation and help explain eventual inconsistencies between similar studies. Ideally, a study should provide suitable information on the compensation system (i.e. hourly rate, piece rate etc.) for comparison purposes.

5.6.3 Technology

All studies should integrate a full description of the technology being tested, such as machine make, model, type, power, year of manufacture. This information can be easily collected during work pauses, by reading the machine ID tag and interviewing the operator. Depending of the study goal other information such as vehicle weight, number of wheels, tire size and loader outreach would be useful.

6 Data analysis

Data analysis is closely related to experimental design, since a specific design will generally be geared to a specific data analysis technique. Hence, we are back to the two main study types: comparative and modelling. In fact, statistics offer many different techniques, most of which could be used for analyzing time study data. Here we shall consider only the most popular techniques, frequently the simplest to use. In general, data analysis will move through the following steps.

6.1 Descriptive statistics

The very first step in data analysis is describing the data. This is done through synthetic indicators that express the distribution of data and go under the general name of descriptive statistics. Essential descrip-

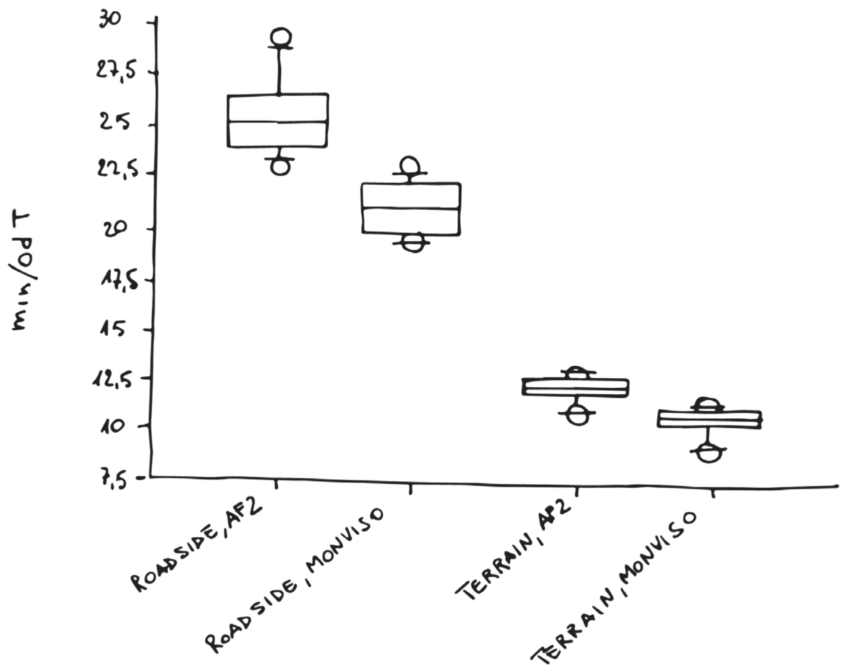


Figure 4 – Box Plot displaying the results of the comparison of terrain vs. roadside chipping in a short-rotation poplar plantation.

tive statistics are: mean, standard deviation or standard error, minimum and maximum. It is often useful to add the lower and upper quartiles, as well. Descriptive statistics will be updated at the end of the analysis, if the data have been purged of the eventual outliers. Data distribution can also be described graphically using Box Plots, which display the median (central line), the range and the 10^{th}, 25^{th}, 75^{th} and 90^{th} percentiles (Fig. 4). Box Plots are especially useful for displaying potential outliers (any dots much further away from the 10^{th} and 90^{th} percentile lines).

Box 10 - Statistical packages
A number of different statistical packages can be used for analyzing the data collected in forest work studies, and among them some of the most common are: SAS, SPSS, Minitab, R for Excel. They offer similar results but vary in cost and user-friendliness. R for Excel is a good tool, with a large capability and is freely available to all, although it takes some learning before it can be used correctly. The others are available at a price (from very moderate to high) and are generally quicker to master.

6.2 Checking for outliers

The data pool should be checked for outliers. The easiest way to do that is by extracting average and standard deviation for each data string, and checking how these values match expected figures for that given data type. Evident mismatches should arouse suspicion. For instance, if the average felling cycle has a duration of 40 seconds, it would be reasonable to suspect that a record indicating a duration of 4000 seconds is faulty. Then, this record should be extracted and examined for possible errors (e.g. erroneous inclusion of more cycles in the same record, transcription error, unwarranted inclusion of delay events etc.). A further method to detect possible outliers is to plot the data and just look for any points that seem unreasonably off the charts. Finally, there are formal outlier tests, essentially based on the criteria of "distance from the mean" and "distance from the nearest neighbor". Among these tests are the Grubbs' Test for the detection of single outliers and the Tietjen-Moore Test for the detection of multiple outliers. Standard or modified Z-scores can also be used for the detection of potential outliers. Suspected outliers should only be removed from the data pool if there is an objective reason (proof of error) to justify their exclusion. Otherwise we might be tampering with the data. When in doubt, a good strategy could be that of retaining the potential outliers and adopting a robust statistical technique that will not be unduly affected by outliers (outlier accommodation).

6.3 Checking for normality

Before analysis, data should be checked for normality by drawing a frequency distribution graph. If the data is normally distributed, then we can use parametric statistics; otherwise, we need to apply non-parametric statistics.

6.4 Data transformation

If data is not normally distributed, one could also try to alter its distribution through mathematical operations performed on each observation. This procedure is called data transformation and it has the purpose of bringing data distribution closer to normality. Common transformations are: square-root transformation for count data, logarithm transformation for size data and arcsine transformation for percent data. Transformed numbers are then used in the planned statistical tests. Test results must be reconverted to the original through back-transformation, by applying the inverse of the mathematical operation originally performed.

6.5 Making comparisons

The statistical significance of any difference of mean values returned by comparative trials can be checked with different statistical tests, depending on the number of treatments being compared, the relationships between repetitions in the treatments (paired or not) and the distribution of the data. If the data are normally distributed, the best way to analyze a typical factorial experiment will be through the technique called Analysis of Variance (ANOVA). The ANOVA table will provide information

	Effect	**DF**	**SS**	**%**	**F-Value**	**P-Value**	**Power**
min per odT	Treatment	1	840.157	89%	451.90	<0.0001	1.00
	Clone	1	52.628	6%	28.31	<0.0001	1.00
	Interaction	1	13.606	1%	7.32	0.0136	0.74
	Residual	20	37.184	4%			
MJ per odT	Treatment	1	9297.51	51%	48.25	<0.0001	1.00
	Clone	1	3789.88	21%	19.67	0.0003	0.99
	Interaction	1	1155.08	6%	5.99	0.0237	0.64
	Residual	20	3853.81	21%			
kg CO_2 per odT	Treatment	1	1.71	42%	32.46	<0.0001	1.00
	Clone	1	0.96	24%	18.15	0.0004	0.99
	Interaction	1	0.3	7%	5.59	0.0282	0.61
	Residual	20	1.057	26%			

Table 2 – ANOVA table calculated with the data obtained from the experiment represented in Figure 3

about the statistical significance and the strength of the effects derived from the treatments under analysis. Table 2 shows a typical ANOVA table calculated for the data obtained from the factorial experiment represented in Figure 3.

The ANOVA table taken as an example shows that treatment type (i.e. terrain chipping vs. roadside chipping) has a strong (89% of total SS) and significant effect ($p < 0.0001$) on time consumption expressed in minutes per oven-dry tonne. Clone type also has a significant ($p < 0.0001$) yet minor (6%) effect on specific time consumption. Please note that the percent value for each given effect is simply obtained by dividing the sum of squares (SS) for that effect by the total sum of squares. Also note that you can be interested in other types of input-output relationships than time consumption: in biomass energy studies it is also important to quantify energy use efficiency (MJ per tonne) and GHG emission levels (kg CO_2 per tonne).

In general, when we are checking the effect of one single variable (e.g. only treatment or only clone) we can conduct a standard t-test, if the variable being tested can assume only two levels (i.e. either roadside chipping or terrain chipping). If the variable can assume more than two levels, we can use one of the ANOVA post-hoc tests, such as Fisher PLSD, Scheffe, Bonferroni-Dunn, Tukey-Kramer etc. If data do not follow a normal distribution, the above tests will be replaced by their non-parametric equivalents, such as the Mann-Whitney test (two-way comparison, unpaired), the Wilcoxon Signed Rank test (two-way comparison, paired) or the Kruskal-Wallis test (more than two treatments). There are also other non-parametric tests that could be used to the same purposes, and the mention of specific tests made above is for reference only, without being exclusive. If the test includes both fixed effects and co-variates, then the ANOVA should be replaced by the similar technique called ANCOVA (Analysis of Co-Variance).

6.6 Modelling

The statistical significance of suspected relationships can be tested through regression analysis. The most commonly used regression type is ordinary least square regression. This technique is used to calculate an equation capable of representing the relationship between a dependent variable (typically time consumption) and one or more independent variables.

The predicting capacity of the equation is described by the coefficient of determination (R^2), which indicates the percent of the total variation

Box 11 - Caveats

Statistics are a specialist field and foresters are not always too versed or interested in mathematics (although some are really good at it). Hence there is a risk of making fundamental mistakes with data analysis. We list some of the most common mistakes encountered when examining forest harvesting studies, so that you may avoid them and get your manuscript through peer-reviewing with as little damage as possible.

- Productivity is a derived unit (output/time) and as such it is very unwieldy. Averaging the individual productivity values for a number of observations will return a value that will be different from the sum of output values divided by the sum of time consumption values, due to the skew in the distribution of single observations. Hence, it is always preferable to use time consumption in all calculations, since this figure is more stable and theoretically more appropriate. Productivity values are calculated in the very end of analysis, by inversing time consumption values. For instance, if time consumption is 0.2 hours/m^3, the productivity is 1/0.2=5 m^3/hour.
- In regression analysis, the predictors must be linearly independent, i.e. it must not be possible to express any predictor as a linear combination of the others. Basically, collinear variables contain information about the dependent variable and are redundant. Such redundancy will confound the individual effects of the variables, thus weakening their predicting capacity.
- The use of polynomial equations to describe machine performance or any work related phenomenon is considered illogical by most foresters, and such equations should not be used just because they provide better "fits" than other models. In fact, the form of the mathematical model should be based on what is known about the mechanics of the process. In many cases, linear models are also inappropriate, except as first-order approximations over limited ranges of the independent variables. For example, a linear model for forwarder travel time as a function of distance is consistent with the mechanics of forwarder travel. On the other hand, a linear model for number of trees accumulated by a feller buncher as a function of average tree size is guaranteed to fail as tree size increases.

explained by the numerical relationship just produced. A regression with R^2 = 0.8 will explain 80% of the total variation in the data pool, and will indicate a good predictor. Several indicators describe whether the effect of a given independent variable is statistically significant, and among them the p-value is most commonly used. This value is always individually associated to each independent variable, and it can be interpreted as the probability that the effect described by the equation can happen by chance. Hence, a very low p-value (< 0.05) is a good criterion for the inclusion of a variable in the regression equation. Beside R^2, the analysis of residuals can provide useful information on model quality.

When handling more than one independent variable, multiple linear regression is applied. By definition, this technique works best with independent variables that change linearly. Non-linear variables can be linearized by appropriate transformation, for instance by raising them to

Box 12 - Reporting

To ensure quality, clarity and repeatability, a report should include at least the following elements:

- introduction and background of the study, leading to problem statement
- clear and direct goal statement (e.g. The goal of this study was to…)
- description of the system under study, including a definition of system boundaries
- description of site conditions
- description of the experimental design, including number of replications and total study duration
- description of the techniques employed for statistical analysis
- definition of time concept used (what kind of delays were included etc.)
- definition of observational level used (shift level, cycle level etc.)
- definition of cycles and/or time elements (with break points and priority levels as needed)
- definition of output units and all required calculations for estimating indirect outputs
- description of measurement methods
- description of results, and comparison with the results from other similar studies
- inferences that can be made from the results of the study
- report of known constraints or limitations with the study and/or generalization of results

Use simple and clear language. No scientific report ever won the Pulitzer prize, so it is not worth trying now. If necessary, get the document revised by a professional language editor. Be concise, avoid redundant tables and figures.

the power of 2, if they show a quadratic behavior.

Regression equations can also be used to compare treatments. To this purpose, one of the treatments will be taken as the base case and the others will be configured just as other independent variables, reporting the value 0 or 1 respectively for the absence or the presence of the specific treatment.

For example, if we have a yarder working alternatively with standard and radio-controlled chokers, then we can add the variable “radio-controlled chokers”. This variable is set to 1 when the radio-controlled chokers are used, and to 0 when standard chokers are used. Since it indicates a treatment difference, this variable is defined as an “indicator” variable. In fact, the indicator variable is not a truly continuous variable (such as yarding distance), and for this reason it is also called a “Dummy” variable.

However dummy variables work well and their use in work studies is accepted, widespread and very effective [7]. Models should be verified and validated. A complete validation process normally includes several steps, but here we can recommend at least two of them: internal verification and independent validation. Internal verification consists of using the model to replicate some of the observations inside the data pool used for its construction.

The same predictor values will be input into the model and the predicted value will be compared with the actual one, using statistical analysis to detect if the eventual difference is statistically significant. Independent validation is a very similar process, whereas the observations being replicated come from outside the original data pool used to calculate the model. For this purpose, one may try and obtain data from other studies, or partition the study data pool into two subsets, one of which will be used for model construction and verification, and the other for model validation.

7. *Conclusive notes*

Much has been written about work studies, and this short guide can neither summarize all the knowledge on the subject, nor replace the many scholarly books that represent the foundations of work science. In fact, this guide only aims at providing a common platform for all people

7 *For a better explanation of the significance, the justification for use and the benefit obtained from dummy variables, see Olsen et al. 1998.*

approaching forest biomass work studies, so that misunderstandings can be avoided and communication improved. Ultimately, the success of this effort will depend on the contribution of all people involved in the COST Action, and on their adoption of this guide as the reference for their future work. We believe that this handbook is simple, clear and comprehensive enough for practical use in forest biomass work studies. What is more, this guide does not give prescriptions on how to do things, but rather offers insights on what could be done, leaving everyone free to develop their own specific approach to the work. Harmonization is not standardization. There is no need for everyone to do the same thing in the same way. That is contrary to academic freedom and progress. What we need to do is to understand what everyone has done, so that we can track back the process to the original elements and eventually "translate" the results. It is unrealistic to think that everyone should speak the same "scientific language". On the contrary, it is more practical to develop a "dictionary" that will allow effective communication regardless of language. That is the main purpose of this GPG, which offers advice, not directions, hoping that this advice can be useful and – only if useful – adopted by those who will read it.

8. *Relevant bibliography*

A number of different manuals and articles are available on the subject, so that it may be difficult and confusing to compile an exhaustive bibliography on work studies. However, the following texts may provide essential information on time studies, and their reading is strongly recommended:

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

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

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

Day R. 1975. How to write a scientific paper. ASM News, 41: 486-494.

Gullberg T. 1995. Evaluating operator-machine interactions in comparative time studies. International Journal of Forest Engineering, 7, 1, 51-61.

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

Howard A. 1989. A sequential approach to sampling design for time studies of cable yarding operations. Canadian Journal of Forest Research 19: 973–980.

ILO Guidelines for labour inspection in forestry Geneva, International Labour Office, 2006 ISBN 92-2-118081-6 (print) ISBN 92-2-118082-4 (web) Labour inspection, forestry, ILO Convention, comment, application. 04.03.5

Lindroos, O., 2010. Scrutinizing the theory of comparative time studies with operator as a block effect. International Journal of Forest Engineering 21: 20-30.

McMahon S. 1995. A survey method for assessing site disturbance. Project Report 54. Logging Industry Research Organisation, New Zealand, 16 pp.

Murphy G. 2005 Determining sample size for harvesting cost estimation. New Zealand Journal of Forestry Science 35: 166-169.

Nuutinen Y., Väätäinen K., Heinonen J., Asikainen A., 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., Hossain M., Miller M. 1998. Statistical Comparison of Methods Used in Harvesting Work Studies. Oregon State University, Forest Research Laboratory, Corvallis, OR. Research Contribution n° 23. 31 p.

Samset I. 1990. Some observations on time and performance studies in forestry. Communications of the Norwegian Forest Research Institute, n. 43.5, 80p.

Spinelli R., Visser R. 2009. Analyzing and estimating delays in wood chipping operations. Biomass and Bioenergy 33: 429-433.

Zerga, J. E. 1944. Motion and time study: a résumé and bibliography. Journal of Applied Psychology 28: 477-500.

Appendix 1 – Work science: definitions

Work science is the branch of knowledge associated with work and its measurement, including: the work itself, man at work, the machines, tools and other equipment employed in work and the organization and methods of work.

Work study is the systematic study of technical, psychological, physiological, social and organizational aspects of work. It provides for critical examination of existing and proposed ways of doing work. Work study is based on objective, unbiased observation and analysis. It is applied to establish or improve the efficiency of production.

Organization study is the systematic and critical analysis of organizational structures and relationships, in order to describe and improve the organization.

Method study is the systematic and critical analysis of ways of doing work, in order to make improvements.

Work measurement is the application of techniques designed to measure: 1) the input of resources into the productive process, 2) the methods and motions of work and 3) the output of production. For man at work the measurement may include: time consumption, movements and working motions, physical and mental workload etc. For machines and tools: time consumption, wear, movements and maneuvers, energy consumption etc. In addition to this, it is common to include descriptions of the work object (tree size etc), the work environment (terrain, weather etc) and the quantity and quality of production.

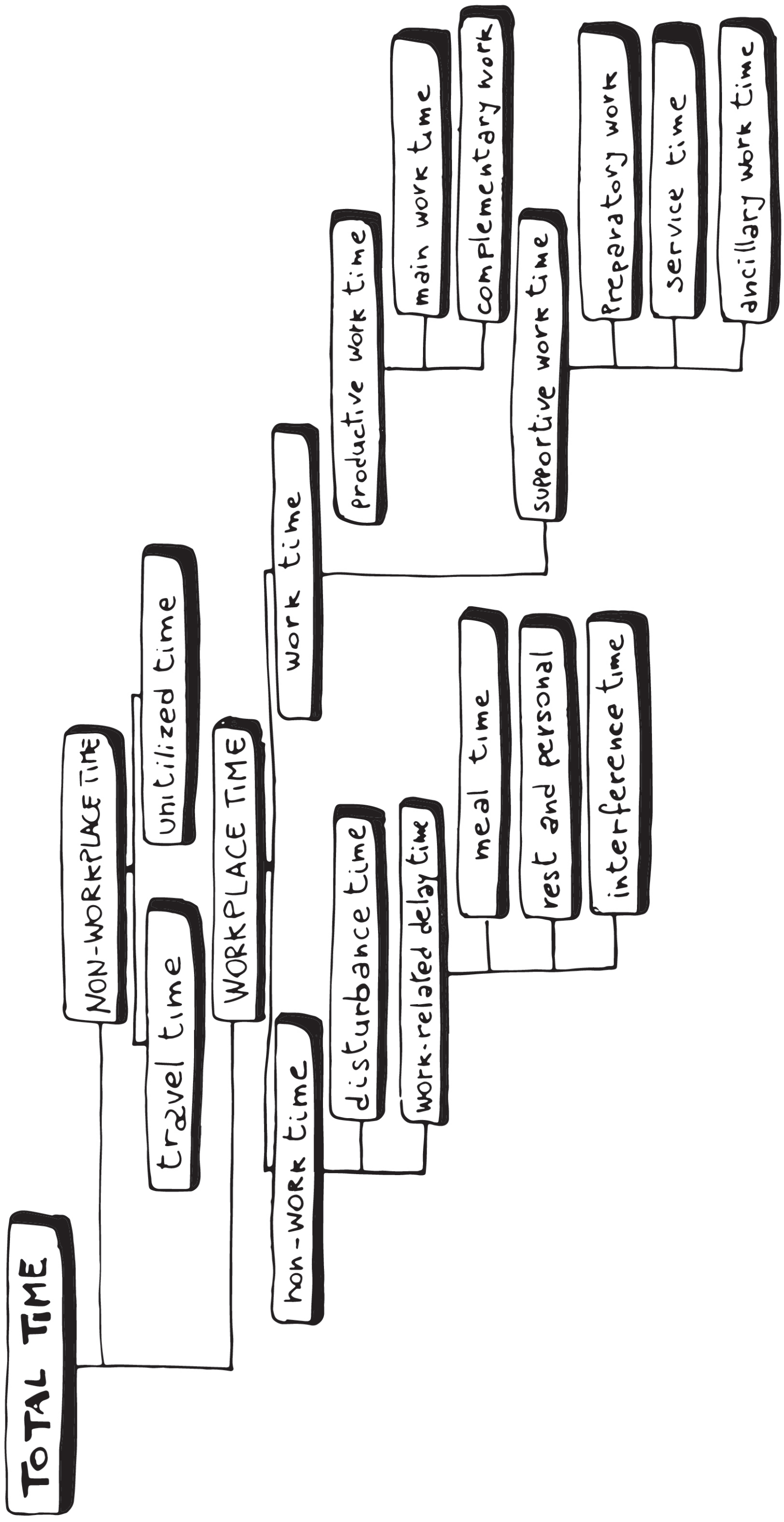
Time study is the measurement, classification and subsequent systematic and critical analysis of time consumption in work, with the purpose of eliminating useless time consumption.

Motion study is the systematic and critical analysis of working motions with the purpose of describing the motions, eliminating useless motions, and arranging the remaining motions in the best sequence for performing the operations.

Appendix 2 - Example of the main parameters most capable of affecting harvesting performance

Operation	Options	Cycle parameters	Units	Stand/operational parameters	Units
Felling and/ or processing	**Motor-manual**	→ Tree size:	m^3 – dry t (or green tons at a given moisture content) tree n°	→ Felling type (clear cut, systematic, selective thinning)	Felling type code
		- unit volume/weight		→ Felling intensity	Trees•ha^{-1}, m^3•ha^{-1}, t•ha^{-1}
	Mechanized	- dbh or other dimensions	cm	→ Stand type (Even-aged high forest, coppice, etc.)	Stand type code
		- number of trees per cycle (if more than one)	nr trees•$cycle^{-1}$	→ Weather conditions	Weather code
		→ Species	Species code	→ Terrain class /Aver. slope	Terrain class code, % slope
Extraction	**Forwarding**	→ Species / product sizes	Species code	→ Terrain class /Aver. slope	Terrain class code, % slope
		→ Load type (Whole trees, branches and tops, bundles, stump wood, roundwood)	Load code	→ Felling intensity	Trees•ha^{-1}, m^3•ha^{-1}, t•ha^{-1}
		→ Payload	m^3 – dry or humid tonnes/ cycle (trip) No. of bundles	→ Biomass and/or roundwood removals	Y/N
				→ If Y, bundle or pile size. Time in pile at harvesting before extraction	Trees•$pile^{-1}$, m^3•$pile^{-1}$, dry (or green) t•$pile^{-1}$
	Skidding	→ Extraction distances:	m		
		- Access/main road unloaded trip	m		
		- Strip road unloaded trip	m	→ Biomass average moisture, if significant	% m.c. (with indication of dry or wet basis)
		- Access/main road loaded trip	m		
		- Strip road loaded trip	m	→ Distance between strip roads	m
		- Loading distance		→ Access/main roads density	m•ha^{-1}
	Yarding	→ Slope, if significant	± % slope		
		- Access/main road unloaded trip	± % slope		
		- Strip road unloaded trip	± % slope		
		- Access/main road loaded trip	± % slope		
		- Strip road loaded trip			
Comminution	**Shredding**	→ Species / material sizes	Species code	→ Machine type	Machine type code
		→ Load type (Whole trees, branches and tops, bundles, stump wood, roundwood)	Load code	→ Engine power	HP – kW
				→ Infeeding system (Crane – type, capacity -, front loader,..)	r/min
				Communution mechanism	Feeding system code - data
	Chipping			→ Landing size	m^2
				→ Landing organisation	Code – draft scheme
				→ Transport (truck) type/s	Truck type code/s – payload/s.

Appendix 3 - Classification of time in forest work study (IUFRO 1995)



COST Action FP-0902 - “Good practice guidelines for biomass production studies”

Published by:

CNR IVALSA
Via Madonna del Piano, 10
I-50019 Sesto Fiorentino (FI)
ITALY
www.ivalsa.cnr.it
“Good practice guidelines for biomass production studies”
Year of publication: 2012
ISBN 978-88-901660-4-4