Chapter 16
Valuation of Airborne Laser Scanning Based Forest Information

Annika Kangas, Tron Eid, and Terje Gobakken

Abstract When an inventory is planned for, the decision makers typically strive for maximizing accuracy of the information with a given budget, or even maximize accuracy without considering any budget at all. Recent developments in airborne laser scanning and other remote sensing techniques facilitate the use of data obtained from such sources as an integrated part of the forest inventory process. However, a rational decision maker would not pay for information that is more expensive than the expected improvement in the value of the decisions based on the new information. The statistical accuracy that usually is provided in the remote sensing literature does not dictate the usefulness of the data in decision making. In this chapter we present methods to assess the value of information and go through the recent research related to this where we in different ways try to establish the links between the inventory effort level, decisions to be made and the value of information.

16.1 Introduction

Information has value to the decision maker if there is uncertainty concerning the correct decision. This value of information (VOI) can ex ante be calculated as the difference of expected value of a given decision with and without a source of new information (Lawrence 1999; Birchler and Bütler 2007). Some prior information is always assumed to be available, for instance in the form of old forest inventory data updated with growth models. Thus, VOI comes from the ability to make better

A. Kangas (✉)
Department of Forest Sciences, University of Helsinki, Helsinki, Finland
e-mail: annika.kangas@helsinki.fi

T. Eid • T. Gobakken
Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway

M. Maltamo et al. (eds.), *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*, Managing Forest Ecosystems 27, DOI 10.1007/978-94-017-8663-8_16,

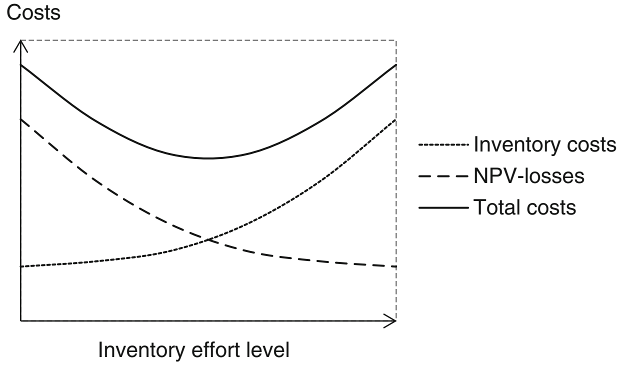

Fig. 16.1 Cost structure when considering the effort level of an inventory

decision with additional information, and a rational decision maker should not pay for information that is more expensive than the expected improvement in the value of the decision.

If the new information is without errors, i.e. perfect, then we can talk about the expected value of perfect information (EVPI). Usually the new information still contains some errors, however, and this means that we can talk about the expected value of imperfect information (EVII).

The problem can also be looked at from a different perspective by analyzing the losses due to sub-optimal decisions. Then, the value of information comes from the reduced expected losses when we have new information. This approach has been more commonly used in forestry, where different inventory methods have been compared using a so-called cost-plus-loss (CPL) analysis. In CPL, the total costs of an inventory consist of the inventory costs and the losses due to sub-optimal decisions, and the inventory method having the lowest total costs is defined as best (Hamilton 1978; Burkhart et al. 1978). Likewise, CPL could be used to define the optimal inventory intensity using a given method (Borders et al. 2008). The cost structure of a concept considering the effort level of an inventory is illustrated in Fig. 16.1. In general, the inventory costs increase and the losses in net present value (NPV) decrease when the effort level of the inventory increases. The ultimate goal in data acquisition planning should be to search for an effort level that minimises total costs.

In CPL, the benchmark is the perfect information, and the losses of a given inventory method are calculated related to that. In EVPI the benchmark is the given inventory, and the value of perfect information is calculated related to that. If we wish to consider the additional value of a new inventory method, it can be calculated if we know the expected losses (EL) or EVPI with prior information and with the new information as

$$EVII = EVPI_{prior} - EVPI_{new} = EL_{prior} - EL_{new}. \quad (16.1)$$

An objective function is used to calculate the VOI, and the VOI is expressed in the same units as the objective function. Thus, if the objective is to maximize NPV from the harvests, the VOI is also expressed as monetary units. If the decision maker was maximizing a multi-criteria utility function, then also the VOI is expressed as utils (Kim et al. 2003; Kangas et al. 2010a). However, it is possible to calculate the respective value in monetary terms, if the utility function involves at least one criterion that can be expressed in monetary terms.

The VOI depends on the application of the information. For example, the use of forest stand inventory data spans from forest property valuation to management planning. Since a number of different decisions can be based on the same information, the information may have different value depending on type of decision. For instance, Duvemo et al. (2012) calculated the losses in a hierarchical planning case including both strategic and tactical level planning. They observed that the losses were much higher, when traditional inventory data was used in both levels (about 166 €/ha) than in the case where perfect data was utilized in tactical level and the traditional data in the strategic level (about 15 €/ha).

The VOI depends also on the constraints involved in the problem formulation, especially if the constraints have no tolerance. For instance, requiring that there is an exactly even flow of timber from the area requires better information than just maximizing the NPV. On the other hand, if the constraints have some tolerance for violations, their effect on VOI is lower.

When a new inventory is planned for, the decision makers typically strive for maximal accuracy with the given budget. However, the statistical accuracy does not dictate the usefulness of the data in decision making (Ketzenberg et al. 2007). For instance, the better the prior information available, the less is the value of additional information (e.g. Kangas et al. 2010b). Moreover, the more sensitive the decisions are to the information, the higher the value of information (e.g. Eid 2000; Holmström et al. 2003; Duvemo et al. 2012). For instance when maximizing NPV in stands that are well past the final harvest age, additional information may have little value as the decision is always to harvest immediately. Therefore, modeling the VOI as a function of forest characteristics may not be straightforward (Islam et al. 2010). Finally, the VOI is the higher the larger the number of decisions that are based on the information. In forestry this directly means that data is the more valuable, the longer it can be used for decision making.

The role of VOI in data acquisition planning is to ensure an optimal level of efforts. For instance in planning an optimal sample in a given stand to be harvested, the optimal number of sample plots is achieved when the marginal costs are equal to marginal benefits (e.g. Birchler and Bütler 2007), i.e. when the cost of the last sample plot is exactly the same as the additional value of information obtainable from that additional plot.

Most of the applications in forestry have dealt with defining the expected losses (or EVPI) from a given data acquisition method, such as sampling based forest inventory, traditional visual forest inventory, photo interpretation based forest inventory or airborne laser scanning (ALS) based forest inventory (e.g. Holmström et al. 2003; Eid et al. 2004; Mäkinen et al. 2010; Borders et al. 2008). It is, however, possible to calculate the value for any piece of information. For instance, it is possible to calculate the value for one or more specific variables (Eid 2000; Kangas et al. 2010b), which helps in determining to which variables the data acquisition efforts should be concentrated. It is also possible to evaluate the value of a single sample tree measurement or single sample plot, for instance in determining whether it is profitable to measure the past growth from the plots or not (Mäkinen et al. 2012). A given prediction model or a given predictor in a model can also be evaluated, which can help in determining what are the most profitable input variables for a model (Eid 2003). The totality of a data acquisition strategy can be evaluated as well. For instance in ALS inventory, it would be possible to define the optimal ALS acquisition settings, optimal number of field plots and measurements, and optimal calculation system for the results. Finally, it would be possible to evaluate decision support systems (DSS), e.g. a system of growth and yield models (Rasinmäki et al. 2013). While it is quite difficult to account for non-sampling errors such as measurement errors or tally tree prediction errors when defining the accuracy of a given sampling method, for instance, all these aspects can be easily accounted for in VOI analysis.

16.2 Methods for Analyzing Value of Information

16.2.1 VOI Estimates Based on Empirical Data or Simulation

The simplest way to calculate the VOI or expected losses is through simulation. If we assume that the decision maker is maximizing the NPV in his/her forest estate without constraints, the loss in each stand i can be calculated as the difference (NPV_{loss}) between the true NPV from the optimal decision (NPV_{opt}) and the true NPV from the sub-optimal decision (NPV_{err})

$$NPV_{loss\ i} = NPV_{opt\ i} - NPV_{err\ i}. \tag{16.2}$$

In order to calculate the true NPV from the sub-optimal decision, the optimal decision based on the erroneous data is first calculated, and then the NPV is calculated by applying these decisions on the correct data (Fig. 16.2). The NPV obtained from the erroneous data cannot be used for calculating the VOI: the difference between true NPV and the erroneous one reflects erroneous expectations rather than actual losses. In fact, it may be either lower or higher than the true optimum.

It is useful to restrict the decisions included into the analysis to the time period that the collected data will actually be used. For instance, if the data is to be used for 10 years, then only the losses occurring during that time should be accounted for. This can be done so that after the period in question, the decisions are assumed to be correct, given the decisions already made (Holmström et al. 2003). Another

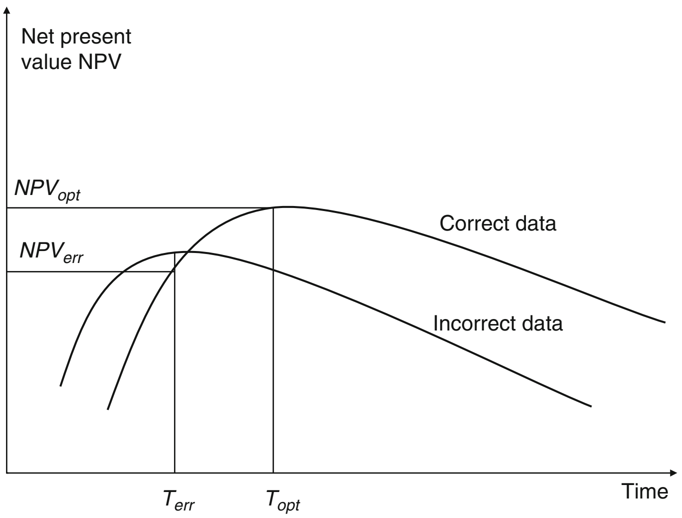

Fig. 16.2 The principle of cost-plus-loss analysis (see Ståhl 1994; Eid 2000). Time here means timing of a harvest, and the *curves* describe NPV obtainable with a given timing. *NPV* with the incorrect data is obtained by looking for the true *NPV* using erroneous timing

approach is to use an estimate of the value of the stand (i.e. net present value of expected future yields given the decisions so far) at the end of the period considered (Mäkinen et al. 2010; Pietilä et al. 2010).

In order to calculate the losses, the true and erroneous data both needs to be available for the areas in question. In some cases the test areas have been measured with several methods. Then, the most accurate method (typically a field sample) is considered as reference and ground truth, and the data obtained from e.g. remote sensing are assumed to be the erroneous data. The benefits of this approach are that the correlations between the different errors and their distribution, i.e. the error structure, will automatically be close to correct. The drawback of this approach is that perfect data for comparison can never be obtained. The field measurements also contain errors and the observed differences are a combination of errors from both methods. This means that if one of the methods is assumed to represent the truth, both sources of error are attributed to the other method. Another problem with this approach is that it is never possible to examine methods that do not yet exist, or even different variations of methods that do exist. It is impossible to determine, for instance, if it was profitable to measure some variables more accurately and some variables less accurately than with the currently used method.

The first study analyzing the VOI of forest inventory data obtained by means of ALS through CPL analysis was carried out by Eid et al. (2004). Based on actual data from practical management planning situations, they compared photo interpretation and ALS based inventory and used independent and intensive field sample plot inventories to represent the true values. The results were clear: while the inventory

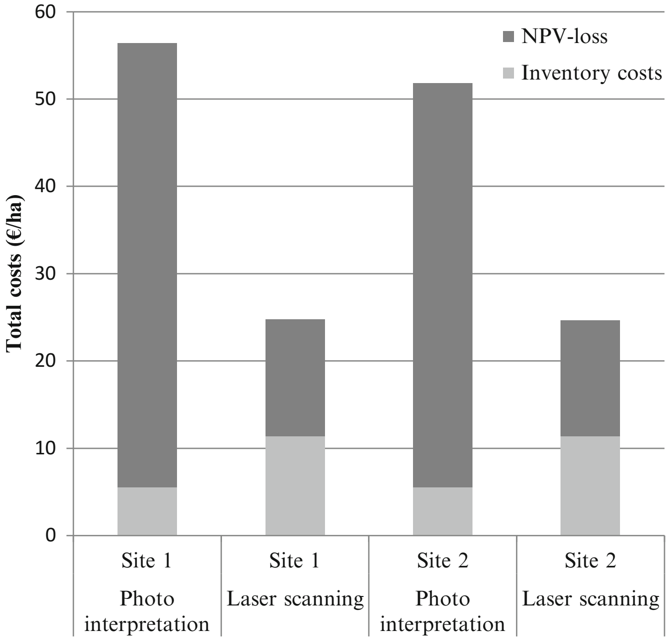

Fig. 16.3 The inventory costs and estimated NPV-loss of two inventory methods in two study areas (Eid et al. 2004)

costs of laser scanning were twice as expensive as that of the photo interpretation, the total costs were still less than those of photo interpretation (Fig. 16.3).

The above described example comparing photo interpretation and ALS is insufficient in the sense that it represents only one out many effort levels that could be applied for the methods. We do not know, for example, whether the effort levels of the ALS procedures carried for the two sites were near the optimum, or if they were too high or too low (see also Fig. 16.1). In general, the effort level and the VOI should be focused in any study related to innovative inventory methods. For ALS applications, both ALS acquisition settings (e.g. point density, footprint size, and scan angle) and efforts in field calibration (e.g. size of sample plots, number of sample plots, precision level required for georeferencing sample plots) could be considered in this context.

A more recent example of VOI analysis is the ALS study provided by Bergseng et al. (2013). The field data came from 23 large field plots in south-eastern Norway with georeferenced single trees collected. A single-tree growth and yield simulator was used in the analysis and NPV-losses when considering the timing of the final harvest were quantified. Losses were derived from four alternative inventory approaches, namely when using (1) mean values obtained by the area-based approach (ABA-MV), (2) diameter distributions obtained by the area-based approach (ABA-DD), (3) individual tree crown segmentation (ITC), and (4) an

Table 16.1 Maximum and mean loss in NPV (in percent) from the incorrect harvest timing decisions based on erroneous inventory data calculated by applying field reference data for the 23 plots

	NPV-loss (%)	
Inventory approach	Max	Mean
ABA-MV	23	5.4
ABA-DD	12	1.0
ITC	9	1.3
SITC	16	1.8

approach called semi-ITC, where field reference data within crown segments from the nearest neighboring segment are imputed. The results showed that using mean values from ABA (ABA-MV) may yield large NPV-losses (Table 16.1). When also inventory costs are taken into account, diameter distributions from ABA (ABA-DD) appear to provide a favorable combination of accuracy and costs.

Another option when applying CPL analysis would be to simulate the errors of different variables. Then, the field measurement is assumed to represent the true values, and errors are added to those values to obtain different erroneous datasets. Simulation allows for generating hundreds or even thousands of realizations or errors for each stand, and the estimate of stand-level loss is calculated as their mean. This enables a detailed analysis of the losses and their dependency on forest characteristics, and an analysis of the relative importance of different variables. It is thus possible to calculate the VOI for methods not yet available.

The first example on such a simulation approach was by Eid (2000). His results proved that for scheduling the final harvests for different forest stands, basal area and dominant height information had little value, while site index and age had much higher value. The low value of basal area, for instance, was due to the growth model used: if the growth models used do not utilize a variable as an input variable, then that variable does not have value. A similar approach was also used to evaluate and choose between different prediction models with different candidates regarding independent variables (Eid 2003). In addition to consider the usual fit statistics and independent data quantifying errors, the models were also evaluated by calculating NPV-losses due to errors related to field measurements of the candidate variables.

The simulation method has its pitfalls as well. In order to produce reliable results on the value of information, the structure and distribution of the errors need to be reliable. Assuming independent errors and normal distributions may produce misleading results. The shape of the marginal distributions and the dependencies between different errors can be accounted for using so-called copula approach (Mäkinen et al. 2010). The copula function directly describes the dependencies between the variables. In Gaussian copula, for instance, the dependencies are based on linear correlations as in the normal distribution, although the marginal distributions of each variable need not to be normal.

Mäkinen et al. (2010) tested the importance of different error structure approaches in a CPL analysis. They fitted a Gaussian copula with logit-logistic marginal distributions to observed errors of different forest variables based on ALS forest inventory, and used the estimated correlations and distributions to represent the true distributions in the simulation-based CPL analysis. They also

Table 16.2 The mean relative loss (%) estimated with different assumptions of the error distributions of ALS-based forest inventory for different forest operations in one test area inventory (Mäkinen et al. 2010)

	Assumption of the error structure			
Next operation	MLLTr	UN	ULLTr	MNTr
No operations	1.5	1.9	1.3	1.5
Thinning	6.4	5.5	6.6	6.7
Final harvest	9.8	5.0	9.7	9.6
Seed tree removal	34.2	26.4	33.2	34.8
All schedules	6.4	4.6	6.2	6.4

For the assumptions of error structure, see definitions in the text

accounted for the bias in the ALS assessments. The bias was observed to have a clear trend with large values generally underestimated and small values generally overestimated. They compared approaches where the errors were assumed to have a normal distribution, uncorrelated errors and no trend in the bias (UN), to have a multinormal distribution, correlated errors and trend in the bias (MNTr), to have logit-logistic distribution, uncorrelated errors and trend in the bias (ULLTr), or to have logit-logistic distribution, correlated errors and trend in bias (MLLTr).

It was clear that in the ALS data the estimated losses were markedly smaller when no trend in the bias was assumed (Table 16.2, column UN). The difference between multinormal or logit-logistic distribution assumptions was very small, indicating that the assumption of the shape of the distribution is not as important as the trend. Likewise, the difference between the correlated and uncorrelated logit-logistic distributions was small, so that the effect of correlations proved also quite small when compared to the effect of the trend. The effect of trend was so clear that for the case of final harvest the rank order between ALS and traditional visual forest inventory (results not shown here) changed when the trend was introduced (Mäkinen et al. 2010). It means that while with the UN assumption the best method was the ALS inventory, with the error structures actually estimated from observed errors (MLLTr), traditional forest inventory proved to be a better method due to the trend.

These results show that correct distributional assumptions are important, as those separate the different data acquisition methods from each other. If all the errors are assumed normally distributed, the only thing separating the different inventory methods would be the bias and standard error. In reality, the error structure of an ALS-based inventory is, however, clearly different from sample plot based field inventory, for example.

16.2.2 Bayesian VOI Analysis

In general, the ex-ante VOI can be calculated with (Lawrence 1999, p. 17, see also Eq. 9 in Kangas 2010)

$$VOI(S) = E_y \max_a E_{x|y}\, \pi\,(x, a) - \max_a E_x \pi\,(x, a) \tag{16.3}$$

where S denotes the information source (e.g. inventory method), x denotes the possible states of nature concerning the uncertain variable (the prior information), y denotes the possible messages from an information source (message denoting one piece of information), $\max_a E_x\pi(x,a)$ the expected payoff of the optimal decision a without the additional information, and $\max_a E_{x|y}\pi(x,a)$ is the expected payoff of the optimal decision (e.g. max NPV) with a given message y and state x, which is further averaged over all possible messages y. This is called preposterior analysis (Lawrence 1999 p. 65). In fact what is needed in such analysis is the prior distribution of the posterior mean of the uncertain x (Ades et al. 2004). If the information source S is perfect (i.e. message precisely describe the state of nature), the formula can be presented simply as

$$VOI(S) = E_x \max_a \pi(x,a) - \max_a E_x \pi(x,a). \tag{16.4}$$

Then, the posterior distribution of x given y is not needed, and Eq. 16.4 provides the EVPI, the upper bound for the acquisition cost of information y (e.g. Ades et al. 2004). This formula is closely related to calculating the mean loss from the simulated stand-level losses, only here the expectation is taken on the optimal decisions with true values and the decisions with incorrect or prior values rather than the differences of them, i.e. the losses. It is essentially the same as Eq. 16.1.

The Bayesian analysis method has not been used in forestry for many cases. The first application of Bayesian analysis in forest inventory was published by Ståhl et al. (1994). A more recent example is provided by Kangas et al. (2010b) where the value of wood quality information (information concerning the height of the living and dead crown) in a bidding situation was analyzed. The value proved to depend highly on the prior information available: if a model for the crown height was utilized as prior information, additional field information was only valuable in the potentially most valuable stands. However, when the prior information was poor, i.e. only the range of possible variation in wood quality was assumed known, the quality information obtained from crown height measurements could be valuable.

16.2.3 Value of Information Under Constrained Optimization

In most studies, VOI has been analyzed in a setting where the utility of decision maker only consists of maximizing NPV without any constraints. In practical decision-making, constraints are, however, common. Then, if there is uncertainty about the coefficients of the decision making problem, the feasibility of the optimal solution in all possible states of nature cannot be guaranteed. This complicates the VOI analysis: it is not enough to analyze the possible losses in the objective function value, but the costs of violating the constraints need to be accounted for as well. The infeasible solutions may well have a higher objective function value (i.e. higher NPV) than the true optimal decision. If the costs of violating of constraints are not accounted, VOI may then be negative.

In forestry decision-making situation with constraints, Islam et al. (2009) studied the losses, when the decision maker required a given level of growing stock at the end of the planning period and/or the incomes to be even between the planning periods. In order to solve the problem due to violated constraints, Islam et al. (2009) calculated the NPVs under uncertain information by using modified constraints. The constraints were modified as much as required to find a feasible solution as close as possible to the original problem setting. Even then, the costs of violating the original constraints were not included in the analysis.

The costs of violating the constraint depend on the nature of the constraint. The first aspect of importance is whether the constraint is a real constraint (external constraint) or part of the decision maker's utility (internal constraint). For instance, requiring even flow of incomes may be part of the forest owner's utility rather than a real constraint, i.e. the decision maker is giving up some of the objective function value to get an even flow of incomes. Then, the costs due to violating the constraint depend on the relative importance of the incomes from different periods. They can be defined from the shadow prices or the weights of the incomes from different periods in a utility function. An example of this approach is the study of Gilabert and McDill (2010). They maximized the NPV of the harvests plus the value of the forest at the end of the planning period, with the flow constraints for the harvests and minimum average ending age of the stands, and penalized the violations of constraints using shadow prices.

If the constraints are real, the costs of violating the constraints should also be real. For instance, in the case of budget constraint it might be possible to borrow more money, and thus the cost of violation of the constraint is the interest of that loan (e.g. Kangas et al. 2012). Then, optimal decision may be to violate the constraints to some extent and pay the penalty rather than requiring the constraints to be strictly fulfilled.

Whether the constraints are internal (set by the DM) or external (i.e. real), one possible avenue to solve these problems and to calculate the VOI is stochastic programming. In stochastic programming the decision is made to maximize the expected value of the decision over all possible scenarios of the state of nature, while requiring that the constraints will be met in every possible scenario. Thus, it means that if the constraints are strict, decisions are made in order to be prepared for the worst possible scenario. In such a case information can have very high value. If the violation of the constraints is possible but penalized, the problem could be solved by stochastic goal programming (or stochastic programming with simple recourse), i.e. by not requiring the constraints to be met in every possible state of nature. Then, very poor but highly improbable scenarios do not dictate the decision.

If the uncertainty can be resolved at some later point of time, the decision-making problem can be solved using, for instance, two-stage stochastic programming (e.g. Boychuck and Martell 1996). It means that the first stage decisions are carried out to maximize the expected value of the decision as above, but when data becomes available, the second stage decisions can be done to adjust to the observed scenario (see Eriksson 2006). Another possibility would be to use chance constrained programming (e.g. Bevers 2007) or robust programming (Palma and

Nelson 2009). In these approaches, the feasible solution is sought for by requiring a safety margin in the constraint. It means that the decision maker gives up some of the objective function value to be sure that the constraints are met under (nearly) all possible states of nature. In this case, the VOI stems from the possibility of reducing the safety margin or removing it altogether, which allows the decision maker to improve the objective function value.

An example of using stochastic programming to calculate the VOI is given in the following. The problem is to maximize the NPV from the first planning period, with the requirement that the NPV from the subsequent periods $k = 1, \ldots, K$ is at least as high as in the first. It is assumed that the constraints are not strict, but they form part of the decision maker's utility, and the deviations from the even flow are penalized using shadow prices of the constraints as penalties. This problem can be presented as a stochastic goal programming formulation. Assuming there are I different possible realizations of the uncertain coefficients, the stochastic goal programming formulation of the above problem is

$$
\begin{aligned}
&\max z = \sum_{j=1}^{n} E\left(c_{j1}\right)x_{j1} - \tfrac{1}{I}\sum_{i=1}^{I}\sum_{k=1}^{K-1}\left(p_k^+ d_{ik}^+ + p_k^- d_{ik}^-\right)\\
&\text{subject to}\\
&\sum_{j=1}^{n} c_{jik+1}x_{jk+1} - \sum_{j=1}^{n} c_{jik}x_{jk} + d_{ik}^+ - d_{ik}^- = 0, \forall i = 1,\ldots,I, k = 1,\ldots,K-1\\
&d_{ik}^-, d_{ik}^+ \geq 0 \forall\ i = 1,\ldots,I, k = 1,\ldots,K-1
\end{aligned}
\tag{16.5}
$$

where n is the number of stands, $E(c_j)$ is the expected net income (€/ha) from the stand j during the first planning period, c_{jik} is the net income in stand j realization i and period k and d_{ik}^+ is the negative deviation in realization i and period k and d_k^- is the respective positive deviation. The differences of this formulation and the ordinary goal programming formulation are that there are even flow constraints for all the I scenarios, meaning that there are I times as many constraints in the stochastic problem compared to the ordinary problem. Moreover, the penalties of not achieving the goals (p_k) are averaged over all the scenarios, i.e. the expected value of the penalties is minimized.

In case the true values of the uncertain coefficients are known, the optimal solution for the problem for realization i can be calculated as

$$
\begin{aligned}
&\max z = \sum_{j=1}^{n} c_{ji1}x_{ji1} - \sum_{k=1}^{K-1}\left(p_k^+ d_{ik}^+ + p_k^- d_{ik}^-\right)\\
&\text{subject to}\\
&\sum_{j=1}^{n} c_{jik+1}x_{jik+1} - \sum_{j=1}^{n} c_{jik}x_{jik} + d_{i1}^+ - d_{i1}^- = 0\\
&d_{ik}^-, d_{ik}^+ \geq 0 \forall k = 1,\ldots,K-1
\end{aligned}
\tag{16.6}
$$

For each possible realization i (i = 1, ..., I), the optimal solution for problem, z_i, is calculated where the optimal values of the decision variables are $x_{jil.}$ This is the optimal solution with full information. The value of stochastic solution can be calculated with the same objective function, but using the stochastic solution denoted with $\tilde{x}_{ji1}$, and the respective deviations $\tilde{d}_{ik}^-$ and $\tilde{d}_{ik}^+$ as

$$\tilde{z}_i = \sum_{j=1}^{n} c_{ij1}\tilde{x}_{ij1} - \sum_{k=1}^{K-1}\left(p_k^- \tilde{d}_{ik}^- + p_k^+ \tilde{d}_{ik}^+\right) \tag{16.7}$$

The loss due to incorrect information in realization i is then $loss_i = z_i - \tilde{z}_i$, and the value of perfect information EVPI is the mean loss over all the scenarios. Thus, essentially the VOI is calculated similarly to as in the unconstrained case, but in addition to the difference between the net incomes from the solutions with and without information, also the effect of penalties due to deviations from the constraints is accounted for. However, in stochastic programming the solution without information is made for all the scenarios at the same time, to account for the constraints.

The effect of constraints was calculated for a forest estate of 29 stands and 67.29 ha. It was assumed that stand volume has an uncertain initial value, and the development of a stand was predicted based on that information with a growth model for volume. In the case where the error estimate of the initial volume of the stands was 25 %, the optimal value of the objective function with true values was 80,608 €, and with the stochastic solution 79,926 €, meaning that the loss was 815 €, i.e. 12 € per hectare. If the solution were calculated using normal goal programming, the violations of the constraints would be clearly higher as the solution would not be selected so as to adjust to the different scenarios. In that case, the observed loss would have been 360 €/ha. In the case that the NPV of the stands was maximized without constraints, the respective value of information would have been 22 €/ha. Thus, the potential losses from the penalties may have a substantial effect on the value of information. In addition, just using stochastic programming might reduce the losses and thus reduce the value of information. It means that a stochastic solution may have value as such. In Kangas (2013) this case was extended to a case where the accuracy of basal area and height measurements were considered.

16.3 How Long Can the Data Be Used?

16.3.1 Direct Effect of Time

In forestry, time has many direct effects on the VOI, as the decisions considered in a plan typically cover a longer period, for instance 10 years or more. The first direct effect is that VOI is positively related to the number of decisions that can be made.

Table 16.3 Sub-optimality losses (€/ha) from four different inventory methods in two different test areas and with two different interest rates (2 and 4 %) (Holmström et al. 2003)

	Site 1		Site 2	
Method	2 %	4 %	2 %	4 %
k nearest neighbor imputation, stand record information	49	48	82	161
k nearest neighbor imputation, aerial photograph interpretation	44	39	63	138
Field inventory, 5 plots per stand	13	18	21	17
Field inventory, 10 plots per stand	8	11	14	9

It means that VOI will be higher if also thinning decisions are considered instead of just final harvest decisions. It also means that VOI will be higher for an area where there are many decisions, compared to an area with a lot of young stands where harvest decisions are not relevant. Finally, it means that the longer the same data can be used for decision making, the higher the VOI. This needs to be taken into account when comparing different studies on VOI. It could therefore be useful to calculate VOI per year in addition to per hectare, to emphasize the effect of time.

Another direct effect of time is the consequence of discounting: the further away the decisions are, and the higher the interest rate, the less are the costs of making sub-optimal decisions for the decision maker. However, the higher interest rate shortens the optimal rotation times, i.e. moves the final harvest decisions to earlier periods. The observed losses will be influenced by both these factors. Effects of interest rate were evaluated by Holmström et al. (2003) when they compared the sub-optimality losses from thinning and final harvest decisions (Table 16.3).

16.3.2 Growth Prediction Errors

Predictions concerning the future development of a stand are made with growth models. The errors of initial data and growth predictions propagate in time, meaning that the longer the prediction period, the lower the quality of the data at the end of the period (Gertner 1987; Mowrer 1991; Kangas 1997). As the errors propagate through the system, the expected losses will therefore increase and at some time the expected losses increase to a level where collecting new data is more profitable than using the old data. This defines the optimal inventory interval, i.e. the life span of the initial data (Pietilä et al. 2010; Mäkinen et al. 2012). There may also be interactions between the prediction errors and the initial accuracy, meaning that an initial data set with given accuracy is more valuable when the growth models are more accurate (c.f. Ståhl and Holm. 1994).

Pietilä et al. (2010) used a stand-level growth and yield simulator (SIMO) and error models that were based on observed errors in predictions. The development predicted with the stand-level models for basal area and dominant height were assumed true, and errors were simulated around that prediction to analyze the value of growth information. The erroneous growth predicted was generated by the error model as

$$\widehat{I}_t = I_t + \varepsilon_t, \tag{16.8}$$

where $\widehat{I}_t$ is predicted erroneous growth, I_t is true growth and ε_t is the growth prediction error. The growth prediction error was assumed to consist of stand effect (inter-stand variation) and period effect (intra-stand, or between growth period variation). Thus the growth prediction error in each stand was described as

$$\varepsilon_t = u_t + e_t, \tag{16.9}$$

where ε_t is total growth prediction error, u_t is stand effect and e_t is period effect within the stand. Intra-stand error, e_t, was assumed to follow the first-order autoregressive process (AR(1)). This model was used to predict the variation around the development of basal area and dominant height in each stand predicted with SIMO, resulting in a 6.5 % RMSE for basal area and 10 % RMSE for dominant height after 5 years. With an inventory interval varying from 5 to 60 years, the losses varied from 230 €/ha/60 year up to 860 €/ha/60 year, i.e. from 3.8 to 14.3 €/ha/year.

16.3.3 Life Span of Forest Information

Mäkinen et al. (2012) used a similar error model as Pietilä et al. (2010) except that the between-period variation was not assumed autocorrelated, and the autocorrelation was solely due to the stand effect (inter-stand variation). They introduced inventory errors between 0 and 25 % to the dominant height and basal area in addition to the errors due to growth model. The costs of the inventory were described with a hypothetical model to show the effect of the inventory cost structure on the life-span of inventory data with different parameters.

Quite obviously, the losses were smallest with shortest inventory interval (meaning smallest possible growth prediction errors) and perfectly accurate inventory with relative RMSE 0 % (Fig. 16.4). The difference between the 0 % RMSE and 5 % RMSE was, however, quite small. It is notable that when the losses with the 5-year interval increase from 188 to 591 €/ha (3.14-fold) as the inventory accuracy decreases, they increase only from 420 to 695 €/ha (1.65-fold) when the inventory interval is 30 years. In other words, with the least accurate inventories the losses do not much increase as inventory interval increases, but with the most accurate inventories the trend is very clear.

When the inventory costs were assumed to be 8 €/ha for the least accurate inventory and perfect inventory was assumed to cost 324 €/ha, the optimal life-span of the data proved to be 15 years with 10 % RMSE of basic forest information (Table 16.4). For a less accurate data the optimal life-span shortens, and for more

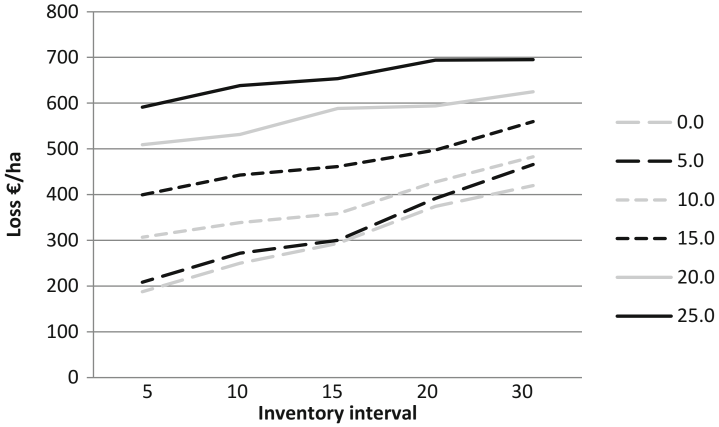

Fig. 16.4 The dependency of the losses on the inventory interval and the accuracy of the initial inventory data (Mäkinen et al. 2012)

Table 16.4 The average absolute total costs plus losses, in €/ha with different inventory interval (years) and accuracy levels (σ)

Accuracy level	Inventory interval (years)				
(σ)	5	10	15	20	30
0	1575.7	995.3	825.9	878.0	744.2
5	1017.7	706.3	610.4	685.4	655.1
10	718.6	560.1	**516.5**	577.2	579.1
15	570.9	534.7	527.3	559.6	599.7
20	567.6	563.0	611.0	615.4	638.9
25	625.5	656.9	666.8	706.7	703.3

The optimal accuracy and inventory period is given in bold (Mäkinen et al. 2012)

accurate data the life-span lengthens. The approach enables analysing the optimal life span with a fixed accuracy data or to optimize both the inventory interval and the data collection accuracy optimally at the same time.

16.4 Conclusions

In this chapter, we have dealt with different aspects regarding valuating forest information. Most of the theory presented applies to all kinds of data. However, actual VOI depends heavily on the data source used and the approach used to calculate the inventory results. ALS data may be offered as area based information and single tree information, or through diameter distributions as an approach somewhere in between. These approaches will differ in the error structure and therefore also in VOI. The differences between the approaches can be quite large, even if the data collected were the same. For each approach, the specifications used for collecting the ALS data, the field data used for interpreting the ALS data and all the models and methods used in different phases of the calculations have their effect on VOI. The importance of these specifications on the VOI has not yet been considered, but in the future the sensitivity of the VOI to them needs to be analyzed in order to maximize the benefits of the collected information.

The VOI also depends on its application, i.e. types of decisions to be made, and how long the data are used, i.e. we should search for a solution where the expected losses have increased to a level where collecting new data is more profitable than using the old data. Although quantifying VOI is not always is an easy task, the links between inventory effort level, decisions to be made and the VOI should be considered in any study related to innovative forest inventory methods including ALS applications. The accuracy measures usually presented in the remote sensing literature to describe a method are not providing a complete picture regarding the usefulness of the data.

References

Ades AE, Lu G, Claxton K (2004) Expected value of sample information calculations in medical decision modeling. Med Decis Mak 24:207–227

Bergseng E, Ørka HO, Næsset E, Gobakken T (2013) Valuation of information obtained from different forest inventory approaches and remote sensing data sources (Unpublished manuscript)

Bevers M (2007) A chance constrained estimation approach to optimizing resource management under uncertainty. Can J For Res 37:2270–2280

Birchler U, Bütler M (2007) Information economics. Routledge advanced texts in economics and finance. Routledge, London, 462 p

Borders BE, Harrison WM, Clutter ML, Shiver BD, Souter RA (2008) The value of timber inventory information for management planning. Can J For Res 38:2287–2294

Boychuck D, Martell DL (1996) A multistage stochastic programming for sustainable forest-level timber supply under risk of fire. For Sci 42:10–26

Burkhart HE, Stuck RD, Leuschner WA, Reynolds MA (1978) Allocating inventory resources for multiple-use planning. Can J For Res 8:100–110

Duvemo K, Lämås T, Eriksson L-O, Wikström P (2012) Introducing cost-plus-loss analysis into a hierarchical forestry planning environment. Ann Oper Res. doi:10.1007/s10479-012-1139-9

Eid T (2000) Use of uncertain inventory data in forestry scenario models and consequential incorrect harvest decisions. Silva Fenn 34:89–100

Eid T (2003) Model validation by means of cost-plus-loss analyses. In: Amaro A, Reed D, Soares P (eds) Modelling forest systems. CABI Publishing, Wallingford, UK, pp 295–305

Eid T, Gobakken T, Næsset E (2004) Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analyses. Scand J For Res 19:512–523

Eriksson L-O (2006) Planning under uncertainty at the forest level: a systems approach. Scand J For Res 21:111–117

Gertner G (1987) Approximating precision in simulation projections: an efficient alternative to Monte Carlo methods. For Sci 33:230–239

Gilabert H, McDill M (2010) Optimizing inventory and yield data collection for forest management planning. For Sci 56:578–591

Hamilton DA (1978) Specifying precision in natural resource inventories. In: Integrated inventories of renewable resources: proceedings of the workshop. USDA Forest Service, General technical report RM-55, pp 276–281

Holmström H, Kallur H, Ståhl G (2003) Cost-plus-loss analyses of forest inventory strategies based on kNN-assigned reference sample plot data. Silva Fenn 37:381–398

Islam N, Kurttila M, Mehtätalo L, Haara A (2009) Analyzing the effects of inventory errors on holding-level forest plans: the case of measurement error in the basal area of the dominated tree species. Silva Fenn 43:71–85

Islam N, Kurttila M, Mehtätalo L, Pukkala T (2010) Inoptimality losses in forest management decisions caused by errors in an inventory based on airborne laser scanning and aerial photographs. Can J For Res 40:2427–2438

Kangas A (1997) On the prediction bias and variance of long-term growth predictions. For Ecol Manag 96:207–216

Kangas A (2010) Value of forest information. Eur J For Res 129:863–874

Kangas A (2013) Effect of sustainability constraints on the value of forest information (in review)

Kangas A, Horne P, Leskinen P (2010a) Measuring the value of information in multi-criteria decision making. For Sci 56:558–566

Kangas A, Lyhykäinen H, Mäkinen H (2010b) Value of quality information in timber bidding. Can J For Res 40:1781–1790

Kangas A, Lyhykäinen H, Mäkinen H, Lappi J (2012) Value of quality information in selecting stands to be purchased. Can J For Res 42:1347–1358

Ketzenberg ME, Rosenzweig ED, Marucheck AE, Metters RD (2007) A framework for the value of information in inventory replenishment. Eur J Oper Res 182:1230–1250

Kim JB, Hobbs BF, Koonce JF (2003) Multicriteria Bayesian analysis of lower trophic level uncertainties and value of research in Lake Erie. Hum Ecol Risk Assess 9:023–1057

Lawrence DB (1999) The economic value of information. Springer, New York, 393 p

Mäkinen A, Kangas A, Mehtätalo L (2010) Correlations, distributions, and trends in forest inventory errors and their effects on forest planning. Can J For Res 40:1386–1396

Mäkinen A, Kangas A, Nurmi M (2012) Using cost-plus-loss analysis to define optimal forest inventory interval and forest inventory accuracy. Silva Fenn 46:211–226

Mowrer HT (1991) Estimating components of propagated variance in growth simulation model projections. Can J For Res 21:379–386

Palma CD, Nelson JD (2009) A robust optimization approach protected harvest scheduling decisions against uncertainty. Can J For Res 39:342–355

Pietilä I, Kangas A, Mäkinen A, Mehtätalo L (2010) Influence of growth prediction errors on the expected losses from forest decisions. Silva Fenn 44:829–843

Rasinmäki J, Kangas A, Mäkinen A, Kalliovirta J (2013) Value of information in DSS: models & data. In: Tuček J, Smreček R, Majlingová A, Garcia-Gonzalo J (eds) Implementation of DSS tools into the forestry practice, Reviewed conference proceedings. Published by Technical University in Zvolen, Slovakia, 167p. ISBN 978-80-228-2510-8

Ståhl G (1994) Optimizing the utility of forest inventory activities. Swedish University of Agricultural Sciences, Department of Biometry and Forest Management, Umeå. Report 27

Ståhl G, Holm S (1994) The combined effect of inventory errors and growth prediction errors on estimations of future forestry states. Manuscript. In: Ståhl G (1994) Optimizing the utility of forest inventory activities. Swedish University of Agricultural Sciences, Department of Biometry and Forest Management, Umeå. Report 27

Ståhl G, Carlson D, Bondesson L (1994) A method to determine optimal stand data acquisition policies. For Sci 40:630–649