500192538: Msc. Dissertation

Developing an inventory protocol of fruit trees for Brown bear (*Ursus arctos*) conservation in Pindos National Park, Greece

By

Kalyan Rai

A Dissertation submitted in partial fulfillment of the requirements

for the degree of

Master of Science (Msc) in Conservation and Land Management

Bangor University

School of Environment and Natural Resources

Bangor University, Bangor

Gwynedd LL57 2UW, UK

www.bangor.ac.uk

September, 2009

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Candidate …….

Date ……..

STATEMENT 1

This dissertation is being submitted in partial fulfillment of the requirements for the degree of Master of Science.

Candidate …….

Date …….

STATEMENT 2

This dissertation is the results of my own independent work/investigation except where otherwise stated.

Candidate ……

Date …….

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Candidate ……

Date ……

Acknowledgement

This project was made possible with the idea, thoughts, suggestions and guidance of my supervisor Dr. John Healey and co-supervisor Dr. Jenny Wong. I would like to thank both of them for providing me an opportunity to work in such a rich and splendid environment in Greece.

I am thankful to Erasmus Mundus commission for keeping faith in me and fulfilled my dreams to study in Europe.

I would also like to thank Dr. Yorgos Mertzanis and his team members of NGO CALLISTO and officials of Pindos National Park for getting me equipped with all the logistical requirements during the period of stay and field work in Greece.

I would like to extend my special thanks to the Mayor of Zagori for all his support throughout my stay in Zagori.

I am also grateful to my colleague Miss Tanwan Topoklang for assisting me in the field data collection.

I am also thankful to Dr. R.K Shetty for helping me to access the unavailable literatures during the project.

Last but not the least, I am grateful to all my friends Imrose, Nicola, Grace, Tim, Ravi and Nirav for being with me and encouraging throughout the project.

Dedication

To the Nature's Amazing creation

The Brown Bear

Dedicated to the welfare of declining population of

Brown bear in Pindos National Park

Abstract

Fruit trees in Pindos national park are relict orchard of the past and serve one of the key habitats of brown bear (*Ursus arctos)*. Inventorying such trees is important in formulating the management prescriptions and to address the brown bear habitat conservation. However, sampling fruit trees in Pindos National Park is challenging due to difficult terrain conditions, relative rarity of fruit trees and variability in spatial distribution (clustered distribution). Range of sampling designs and their combinations with plot designs were evaluated on the basis of statistical and field performance. A two-stage design without 'a priori' stratification with a combination of fixed-area plot and k-tree sample were found best among all the designs. The two-stage fixed-plot design performed better in terms of precision (SE %) over two-stage k-tree sampling. However in terms of capturing the variable of interest and time taken per unit area of sampling, k-tree was found superior. None of the stratification parameters used was found effective in reducing the variance of the two-stage design. An optimisation of number of primary sampling units and secondary sampling units was conducted for both two-stage fixed-area plot and two-stage k-tree plot by considering the precision and cost. A minimum of 27 primary sampling units were suggested based on the findings for both fixed-area plot and k-tree method in sampling fruit trees to achieve the desired level of precision. Finally, a recommendation of two-stage sampling design with a combination of either fixed-plot or k-tree is recommended with some form of stratification by land use.

Word Count

‘15440’ including tables, figure, footnotes and text boxes

‘15240’ without footnotes and text boxes

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 General Introduction

Sampling is an integral part of most forest inventory work due to the large extent of forest areas. There are range of sampling methods available, proper choice must be made as to which method suits best in given field conditions with an available budget (De Vries, 1986). The field conditions may vary between the variable of interest to be measured. For example the inventory conditions of oak trees are likely to be different than those of fruit trees in the same forest due to differences in the pattern of their spatial distribution. In the site of the present study, Pindos National Park[1] mainly consists of the vegetation dominated by oak and patches of beech. The presence of fruit trees within such vegetation type may be rare and the design suitable to inventory oak trees may not be applicable for inventorying fruit trees. Moreover in addition to its rarity, there is an expected variability in the spatial distribution of the fruit trees i.e. a clustered distribution. Fruit trees are one of the key habitats of brown bear[2] (*Ursus arctos*) (Kanellopoulos et. al., 2006; Paralikidis et. al., 2009) in Pindos National Park. Inventorying such trees will indirectly address the key issues of brown bear habitat conservation in Pindos National Park and help the conservation organisations to effectively manage such habitats.

The objective of the present thesis is to design an inventory protocol suitable for the field conditions[3] (difficult terrain conditions) of fruit trees (rarity and clustered distribution) in Pindos National Park based on a pilot survey. The inferences made are with respect to performance (both statistical and field) of a range of sampling methods. As far as possible an approximate costs and logistics were also taken into consideration for designing the optimal sampling strategy.

The present thesis is written in a different style than the recommended pattern and the whole document is divided into 10 chapters. The first two chapters deal with the background of the study and the literature review is split between these two chapters.

---

[1] Southernmost range of Alps located in North West Greece
[2] Largest mammal species present in Greece with high conservation status
[3] Difficult terrain conditions in Pindos National Park for sampling due to its location in the mountains

The chapters 3-8 deals with methodology, results and discussion while Chapter 9 deals with the optimization of the best sampling design based on the results obtained. Finally the last chapter concluded the outcomes of the study with future recommendations.

1.2 Brown bear conservation in Northern Pindos NP

Habitat conservation is one of the key issues facing environmental challenges today. Conservation of species cannot be done in isolation without conserving its habitat. Humans have long caused loss of forest habitat. Human activities such as agriculture, urban development, tourism and plantation forestry are strongly related to habitat alteration and loss of biodiversity (Glowka et al. 1994, IUCN). The conservation of such a landscape is the major focus of many of the environmental organizations. Much of the area of the Northern Pindos National Park was formerly intensive agricultural land which was abandoned some 60 years ago[4]. This abandonment resulted in modification of the nature of the landscape. Colonization by trees and shrubs means that much of the area is now forest and supports a population of brown bears. Migratory species such as the brown bear depends on more than one habitat and the loss of one of them seriously affects the survival of the species (Mertzanis, 2008). This area serves as one of the key habitat of brown bear in Northern Pindos National park (described in detail in chap. 3) and because of its continuous threat due to human caused disturbances and succession of forest trees over fruit trees, it requires urgent attention.

The current IUCN status of Brown bear species *Ursus arctos* is "Least Concern", but the locally small and isolated populations are at higher risk and becoming scarcer (McLellan, et. al., 2008). Hence according to IUCN,

"Least Concern does not always mean that species are not at risk. There are declining species that are evaluated as Least Concern".

The brown bear is a fully protected species in Greece under national and EU legislation. Since 1994 systematic monitoring and conservation programmes have been conducted by the NGOs CALLISTO[5] and ARCTUROS[6] as well as other competent authorities on a continuous basis. But despite all the efforts brown bear conservation status in Greece

---

[4] Local knowledge provided by NGO CALLISTO, Thessaloniki, Greece
[5] Conservation organization with head office in Thessaloniki, Greece, http://callisto.gr/
[6] Conservation organization with head office in Thessaloniki, Greece http://www.arcturos.gr/gr/

remains critical and faces major threats from human-caused mortality, habitat fragmentation, habitat loss and habitat degradation (Mertzanis, 1999a).

The brown bear distribution in Greece consists of two distinct nuclei located in the Pindos mountain range (NW Greece) and the Rodope mountain complex (NE Greece) (Mertzanis, et. al., 1994).The Pindos mountain range is the southernmost range of distribution of brown bear species in Europe reaching 39° parallel. There are about 115-145 individuals of brown bear reported in a continuous stretch of a 6200 km^2 area in the northern Pindos mountain range (Mertzanis, 1999b). The population of brown bear in recent years is expanding towards the eastern and southern parts of the species' former range (Mertzanis,et. al., 2008). This area was inhabited by humans 60 years ago and since then has been left abandoned. The knowledge of home range and habitat use of the brown bear is of much significance in formulating the management prescriptions for the conservation of this species (Mano, 1994; Wooding and Hardisky 1994).

The habitat of the brown bear is mostly remote and mountainous characterized by the presence of coniferous and hardwood forests along with patches of fruit trees, shrubs and grasses. The major vegetation zones of brown bear habitat in Greece comprises oak (*Quercus spp.*) forests (46%), the beech-fir (Fagus sp. - *Abies borissi regis*) forests (30%), and black pine (*Pinus nigra*) forests (19%). The remaining 6% is covered mainly by mountain coniferous forests composed of Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*).

1.3 Relationship between Brown bear conservation and fruit trees

Brown bears are omnivores and feed on a variety of plant and animal products. According to Vlachos et al. (2000) the global bear diet is composed 93% of food items belong to plant origin. Out of that, the major proportion is made up of fleshy and dry fruits. The rest (7%) is composed of insects, reptiles and livestock carcasses.

Brown bear habitat selection pattern is related to seasonal availability of food derived from groups of fruit trees present in the area (Kanellopoulos et. al., 2006). The bear is attracted towards such sites where high trophic value species occurs. Paralikidis et al. (2009) studied the dietary habits of brown bear and found *Pyrus* sp., *Morus* sp., *Prunus* sp. and *Rubus* sp. are important food source for brown bear in Pindos National Park. A list of fruit species present in Pindos mountains are wild apples, (*Malus* spp.), wild pears (*Pyrus amygdaliformis* and *Pyrus pyraster*), berries (*Rubus hirtus, Rubus*

canescens, Rubus idaeus), Sorbus *(Sorbus torminalis, S. domestica* and more rarely *S. aucuparia*), Rosa (*Rosa canina, R. agrestis, R. pulverulenta and R. heckeliana*), cherry (Prunus *domestica*), chestnuts (*Castanea sativa*), and hazelnuts (*Corylus avellana*) (refer Appendix 1). These species are associated with both forest and human settlements (Mertzanis, et. al., 2008). Hence the conservation of these trees and shrubs is extremely important from the point of view of bear conservation.

1.4 Problems of sampling clustered populations

1.4.1 Sampling

Fruit trees in Pindos National Park (PNP) (described in detail in chap. 3) are sparsely distributed. Moreover the spatial distribution of these trees occurs in clusters[7]. Such variability in population size and spatial distribution causes problems in monitoring for conservation purposes (Mcdonald 2004 cited in Thompson, 2004). Sampling such spatially variable sparse populations in the undulating terrain conditions of PNP is challenging and obtaining accurate estimates of population size can be extremely difficult, especially when the budget is limited (Morrison et al. 2008). Responding to this challenge using conventional systematic or random sampling approach would lead to a high number of sampling units, which may be highly inefficient with the information gathered not be justifying the sampling effort (Thompson, 2004).

Fruit trees in PNP are concentrated in relict orchards of the past human habitation[8]. Colonization of these orchards by trees and shrubs over the past 60 years or more has suppressed the regeneration of the associated orchard species. This led to succession towards a closed canopy forest dominated by taller tree species such as *Quercus* spp. leaving the fruit trees rare and clustered.

According to Thompson (2004)

"A rare population is one where it is difficult to find individuals, utilizing known sampling techniques, either because of small numbers, secretive and/or nocturnal behavior, or because of clumped distribution over large ranges, that is, a lot of zeros occur in the data".

---

[7] An assumption of clustered nature of fruit trees is made due to the abandonment of relict orchards since more than 60 years in Pindos National Park

[8] Assumption made by gathering information from the local authorities in Greece

It has been assumed that the status of fruit trees in PNP is rare due to geographic rarity. When a population is rare over a large geographic region, the estimate of abundance or distribution is difficult (Lancia et al. 1994). When the rare population is clustered it adds more problems to estimation. The effort required for an adequate sample using conventional approaches often cannot be justified for sampling such populations.

The problems associated with sampling rare/clustered plant populations can be summarized according to Christman (2000) and Thompson (2004)

- Rare and clustered nature of the species over a vast geographical area
- Unknown spatial distribution of the species
- Lack of appropriate knowledge on the ecology of the species that can be used to predict its habitat associations for pre-stratification of the sampling
- Lack of a well established and tested methodology for sampling such species
- Limitation of budget

Specific challenges that must be considered while selecting from the alternative sampling designs for rare and clustered elements are their accuracy, their risk of bias in the estimator, the shape of sampling distribution of their estimator and their ease of use (Christman, 2000). These issues are dealt with separately below.

1.4.2 Vast area to survey

The fruit trees in PNP are distributed over a large area (approx. 6200 km^2) which is very difficult to survey (Lancia et. al., 1994). Sampling in such a huge area is difficult as it needs high sampling effort and cost to generate a reliable estimate of the size of rare populations.

1.4.3 Accuracy of design

The most important factor which makes the estimate reliable is the variance given by the study design. The variance from sampling rare and clustered populations is high because a lot of zeros occur in the data (Thompson, 2004). Another important factor which is indirectly linked with that of variance is the size of the captured variable of interest (for example number of trees sampled) within a sampling unit. The greater the size of the captured variable the better will be the reliability of the sampling design.

High variance may sometimes be associated with small size of the captured variable. For example, a simple random sampling may have low variance and is superior compared with a cluster sampling. However if the size of the captured variable units is greater in cluster sampling, it will give more reliable estimate of variance. Hence an inference should be made by combining both variance and the number of units captured.

1.4.4 Bias of the estimator

When the survey of a rare species includes more than one observer, the relative precision of the estimate is reduced (Bodkin and Udevitz, 1999). This is because the ability of each observer to detect the rare species varies. Though the issue has its relevance in detecting animal abundance usually has less significance in plant density estimates (unless the plant species is very cryptic or hard to identify).

1.4.5 Ease of use

The decision regarding choice of sampling design should not be made solely on the basis of precision. However an overall score must be given to each alternative design considering its precision, cost and ease of sampling. For example fixed area plot based sampling could be easier to understand and lay out in the field conditions compared with strip lines (Wood 1990). On the other hand can be much easier and faster compared with both these designs (refer section 2.3.1 and 2.3.2 for description of the designs)

1.5 Research question

For this study I am assuming that fruit tree species in Pindos National Park are rare and have and clustered distributions. Sampling such populations in a difficult terrain condition is likely to need a sophisticated sampling design. To address this issue, the following research questions are addressed:

- Which is the best sampling method to quantify the density and spatial distribution of fruit trees in northern Pindos National Park?
- Why is this sampling method better than the alternatives

1.6 Research Approaches

The following research approaches have been selected to answer the above research questions:

- Comparison of sampling designs for clustered populations in a landscape-scale inventory
- Optimization of the best sampling design to achieve reliable precision with reduced cost

Chapter 2 Optimisation of sampling design for clustered populations

2.1 Sampling design

2.1.1 Choice of sampling design

By now we know that sampling rare and clustered population of fruit trees poses various problems in obtaining reliable density estimates especially if the survey area is large (section 1.4). Meeting this need by designing an effective sampling strategy can either be done by modifying existing ones or by creating a new one.

Familiar designs such as simple random sampling and its allied forms may not be the best approach, as the information gathered cannot justify the sampling effort required. Recently various sampling designs have been developed and more are under development for efficient density estimates (Thompson, 1992; Smith et. al., 1995; Christman, 1997; Salehi and Seber, 1997).

In most practical survey situations, it is necessary to choose a design which requires low effort, is cost effective and provides adequately accurate estimates; it may be unrealistic to achieve a high level of precision (e.g., Schreuder et. al., 1993). One of methods most often recommended for sampling clustered population is two-stage sampling (e.g., Cochran, 1977). Two-stage sampling is a special case of multi-stage sampling in which a population is broken down into a number of primary units. Each primary unit is further broken down into a number of secondary units which serve as a unit of measurement (Cochran, 1977; De Vries 1986). This design deviates from another form of two-stage design (Hankin, 1984; Skalski, 1994) where there is a random selection of sampling units as the first stage and a count or measurement as the second stage[1].

2.1.2 Two-stage cluster sampling

Two stage cluster sampling is a special case of two stage sampling (De Vries, 1986) in which clusters only a subset of secondary units within the primary units are sampled. It is often referred as two-stage sampling (e.g. See Cochran 1977, Shiver et al. 1996) or sub-sampling (De Vries 1986; Chacko 1965 p.55) in many textbooks. Two stage cluster samples reduce the number of measurements within a primary unit and thus allow time for more primary units to be sampled compared with single stage cluster sampling (another case of two stage sampling) in which all the secondary units within a cluster

are measured. The efficiency of single-stage cluster sampling was well quantified by Moisen et al. (1994), but for inventorying large areas two stage sampling is preferred (e.g. Bruce, 2005; Lillesand, 1994; Edwards et al., 1998; Nusser and Klass, 2003). Compared with simple random sampling the travelling time between the plots is also reduced with two-stage cluster sampling as the plots are clustered within a primary unit. The shape and size of primary and secondary units vary according to the local situation. For example the shape of primary plots varies from circular to square and irregular plots with variable size between the primaries (e.g. De Vries 1986, chap. 8). The selection of primary plots also varies from random to systematic (e.g. De vries, 1986, Chap.8).

Two-stage cluster sampling has been successfully applied in many forest inventories. The most common of them is inventorying timber in a forest composed of different stands (e.g. Shiver and Borders, 1996 chap 8). Other examples include estimating individual tree characteristics such as tree diameter, height and estimating biomass of tree components (e.g. Corona, 2006). A little modification from the existing two stage design in which the primary units are selected with probabilities proportional to size (PPS sampling) (see Cochran, 1977, chap. 11) of the primary units has been successfully applied in timber inventory of large acreages (Wickham et. al., 2004; Borders et. al., 2005). Another special sub-sampling design is called 3P (Probability proportional to prediction) sampling and was introduced by Grosenbaugh (1963) to estimate timber volumes in timber sales. It proved to be an effective method for sampling sparse species connected to certain types of substrates (Ringwall and Nicholas, 2005).The gain of standard error by 3P sampling is reported to be 30-55% over simple random sampling for estimation of red listed fungi and lichens (Ringwall and Nicholas, 2005).

General estimators in two-stage cluster sampling

Let the population consist of N primary units (PU) and M_i be the number of secondary units (SU) in the i-th PU. A simple random sample without replacement (SRSWR) of n PU´S selected from the population of N. From the i-th PU in this sample, a SRSWR of m_i is drawn, and the values y_{ij} are observed on the sample elements.

Then the unbiased estimator according to De Vries (1986) for the

Population total $$\hat{Y} = N.\frac{\sum_i^n M_i.\bar{\bar{y}}_i}{n}$$

Population mean per PU $\hat{\hat{Y}} = \frac{\sum_i^n M_i.\bar{\bar{y}}_i}{n}$

Population mean per SU is $\hat{\hat{\bar{Y}}} = N.\frac{\sum_i^n M_i.\bar{\bar{y}}_i}{n.M_O}$

Where $\bar{\bar{y}}_i = \frac{\sum_j^{m_i} y_{ij}}{m_i}$ is the sample mean per SU and $M_O = \sum_i^N M_i$

Whereas for variance

Total variance is $v\hat{a}r\ \hat{Y} = \frac{N}{n}.[\sum_i^n M_i^2 .\frac{M_{i-m_i}}{M_i}.\frac{S_i^2}{m_i} + N.\frac{N-n}{N}.\frac{\sum_i^n (\hat{Y}_i - \hat{\hat{Y}})^2}{n-1}]$ and

Variance per SU is $v\hat{a}r\ \hat{\hat{\bar{Y}}} = \left(\frac{1}{M_O^2}\right).v\hat{a}r\ \hat{Y}$

With $S_i^2 = \frac{\sum_j^{m_i} (y_{ij} - \bar{\bar{y}}_i)^2}{m_i - 1}$

The total variance component can be broken down into parts viz. between PU (S^2B) and within PU (S^2P)

Variance between PU $S^2B = \frac{N-n}{N}.\frac{\sum_i^n (\hat{Y}_i - \hat{\hat{Y}})^2}{n-1}$ and

Variance within PU $S^2P = \sum_i^n M_i^2 .\frac{M_{i-m_i}}{M_i}.\frac{S_i^2}{m_i}$

2.1.3 Stratification

A stratified two-stage cluster design is a method commonly used in forestry (Jessen, 1978). Stratification of two-stage sampling is similar to that of stratified simple random sampling and the variance estimators can be obtained by modifying the simple formulas (Cochran 1977, Shiver and Borders, 1996).

The main aim of stratification is to reduce the heterogeneity (Table 1) between the primary units within each stratum. Stratification of primary units can be done based on characteristics which group the units together. For example a forest composed of many stands can be broken into four strata based on stand origin (Shiver and Borders, 1996). Two-stage sampling is mostly applied in forest inventory with ‘a priori’ stratification (e.g. De Vries, 1986 chap 8). The gain in precision due to stratified random sampling over simple random sampling is usually evident, provided that the parameter chosen for stratification reduces the heterogeneity within the stratum (e.g. De vries 1986 Chap 2.6).

The same principle of gain in precision can be applied in stratified two-stage sampling (Table 1) compared with un-stratified two-stage sampling (e.g. Shiver and Borders, 1996 chap 8.1).

General estimators for two-stage stratified cluster sampling

The total and variance for each stratum can be estimated by using the equations as shown in section 2.1.2 for two-stage random sampling. The total across all the strata is obtained by summing individual strata, and variance is obtained by summing individual stratum variances (Shiver and Borders, 1996).

Comparison of stratified random sampling, two stage sampling and stratified two-stage sampling is done based on the discussion above and shown in Table 1 below.

Table 1 Comparison of stratified-random sampling, two-stage sampling and stratified two-stage sampling

Criterion	Stratified Random sampling (STRS)	Two-stage sampling (TSS)	Stratified two-stage sampling
Heterogeneity	< SRS (Simple Random Sampling)	< STRS	< TSS
Precision	> SRS	< STRS	> TSS
Cost of sampling	High	Low	Low
Number of measurements	High	Lower than STRS	Lower than TSS
Suitable for	abundant species and for a small geographical area	rare or clustered species for large area	rare or clustered species for large area
Employed when	sufficient budget is available together with requirement of high precision	Employed when scarcity of budget and reliable estimates are sufficient	Employed when scarcity of budget and reliable estimates are sufficient

2.2 Post stratification

Post stratification is broadly defined as "any method of data analysis which involves forming units into homogeneous groups after observation of the sample" (Smith, 1991). Post stratification is employed in the situations when variables that could have been used for pre-stratification are not present or when pre-stratification does not help in reducing the variance between the plots (Smith, 1991). Hence post stratification in one way is potentially more efficient in reducing the precision than pre-stratification (Holt and Smith, 1979). Post stratification for two-stage sampling receives little attention in most of the textbooks. Cochran (1977) and the more recent authors such as Shiver, et al. (1996) has only mentioned few examples on post stratification and only for stratified random sampling. Though the concept remains the same there is a difference in the variance estimation between pre- and post-stratification. Many types of estimators are used for variance estimation for post-stratified sampling by using a term for reducing the variance through ratio estimators. A list of them is Horvitz Thomson estimator, Rao, Hartley and Cochran method (Cochran, 1977 Chap 9A). Holt and smith (1979) and Smith (1991) went beyond the term (post-stratified estimator) narrowed it and considered separate bias and variance terms by following regression estimators (Bethelem and Keller 1987). Scott and Smith (1969), Royall (1976) and Valliant (1993) proposed a model based approach for reliable inferences of post-stratified estimators instead of following the random variable approach. A comparison of various estimators together with adjusted ones was shown by Valliant (1993) which followed the super-population approach.

2.3 Plot configuration

2.3.1 Fixed area plots

The plots are the ultimate sampling unit where the data is recorded. Hence the proper choice of sample plots is as important as that of choosing a sampling design.

Fixed area plot sampling is the oldest and still the most widely accepted form of sampling strategy used in many parts of the world (Wood, 1990). In this method a constant shape and size of fixed is chosen for the survey, and each of these plots is laid out in the field and the trees or objects tallied inside the plot are sampled for attributes.

Fixed-area plots are very popular and successfully applied in two-stage sampling. De Vries (1986. p. 161-167) lists various types of cluster design differing in their plot shape and arrangement. Also, Shiver and Borders (1996. Chap. 8) gives examples of two-stage designs with both circular- and square-shaped primary and secondary sampling units. But choosing to use fixed-area plots at the secondary stage is more common than choosing it at the primary stage. Apart from its wide application in inventory of temperate forests, the use of fixed-area plots in a two stage design is popular and applied in inventorying tropical forests also. For example Tokola and Shrestha (1999) used circular sub-plots within circular plots (nested plots) in a national forest inventory in southern Nepal.

For executing a survey using two-stage cluster samples an appropriate strategy must be made both at the first stage and the second stage. One of them would be to standardize the shape and size of the plots (Wong et. al., 2007). Fixed-area plots of any size can be used with any sampling design such as simple-random sampling, systematic sampling etc. Fixed-area circular, square or rectangular plots of various sizes have been recommended based on the size of the budget and the precision required (Kangas and Maltamo, 2006). Other factors considered for choosing the plot shape and size is the ease of laying them out in prevailing field conditions. The laying out of square plot is less vulnerable to field errors compared with rectangular plots (Loetsch et al., 1973; Shreuder et. al., 1993). On the other hand circular plots are easy to establish and can be handled by one person whereas rectangular and square plots need a minimum of two for efficient working (Avery, 1983). Circular plots are also vulnerable to error if the plot radius is very large (Shreuder et al. 1993; Loetsch et al. 1973) as the decision regarding whether trees are inside or outside the plot is difficult and prone to bias. In such a situation a combined circular plot[9] is recommended to increase the efficiency (Kangas and Maltamo, 2006).

Large numbers of small clustered plots are often preferred over small numbers of big plots (Kangas and Maltamo, 2006; Shreuder et al., 1993). The main reason is that the former will be better in capturing variability due to its wider spatial distribution. Another reason is that layout of small plots is easier than bigger plots. However laying

[9] Consists of several concentric circles where smaller circles are used for smaller trees and larger circles for larger trees

out more plots leads to high cost and a compromise must be sought between the cost and precision (Congalton, 1988a; Moisen et. al., 1994; Wickham et. al., 2004)). Another form of plot called strip plots are less favoured except for sampling rare populations compared with square, circular and rectangular ones (Kangas and Maltamo, 2006).

An assessment of the advantages and disadvantages of different plot shapes has been summarised by Wong et al. (2007) and is presented in Table 2 below.

Table 2 Factors considered in the choice of plot shape

Factor	Plot shape		
	Circle	Square	Rectangle(many times longer than wide = strip plots)
Suitable Plot size	Smaller (up to 25 m²)	Smaller (up to 5x5m) or larger if sub-divided into smaller squares)	Larger plots up to several ha
Tree density	Low, e.g. farmland	Any	High, selected plot width tends to be narrower in sites with a higher tree density
Edge effects	Minimised	Greater than for circle of same area	Maximised
Ease of checking edge trees	Generally good but can be fiddly passing rope back into centre to get around obstacles	Good, especially if sighting poles are set up at corners	Good, if plot is narrow – use rope from centre line

Source: Adapted from Wong et al. (2007)

Apart from all the new techniques that have been developed to estimate the population density or size parameter the traditional method of fixed-area plot sampling is still a popular method (Wood, 1990) (also refer 2.3.2 and 2.3.3). The reasons for this are that: fixed-area plots are easy to understand, supervise and execute in a wide range of field conditions. A survey conducted by Wood (1990) in 36 countries showed that the relative use of fixed-area plots is 46% compared with 36% for strip-lines and 19% for

variable radius plots (generally Bitterlich point samples) (see section 2.4). Scott and Algeria (1990) also concluded that fixed-area plots are very efficient and better than point samples in estimating tree density when the population includes many trees with small diameters classes.

2.3.2 k-tree plots

The terrain condition in Pindos National Park is undulating and difficult for field survey. Various techniques have been suggested and more are underway to combat the problem and perform approximate estimation (e.g. Sheil et al., 2003; Picard et al., 2005). There are many techniques becoming popular nowadays and plenty to choose from. One of them is distance sampling also known as plotless sampling or k-tree sampling (Kleinn and Vilcko 2006a).

In this method, from a selected sample point, distance and attributes of a fixed number of 'k' nearest trees are measured. The area of the circle is estimated by following different estimators, e.g. Kleinn (2006a), Eberhardt (1967). Subsequently the density is estimated as k trees in the estimated area of the circle, and per-plot values can be converted into per hectare (e.g Kleinn and Vilcko, 2006a).

The origin of distance sampling can be traced way back in 1949 where Cottam (1949) first applied the principles of tree to tree distance sampling in mixed hardwood stands of southern Wisconsin. But Kleinn and Vilcko (2006a) also discussed the possibility that in 1835 'Konig' (an unnamed American surveyor could have been the first person to use such a technique). The tree-to-tree distance technique was successfully applied in early 1950s by many foresters (Kleinn and Vilcko 2006 a). Later on the concept of point to tree sampling was introduced in 1950's and 1960's by German foresters (Kleinn and Vilcko 2006 a).

k-tree sampling is also sometimes termed as n-tree distance sampling and has been commonly applied in forestry. Lynch and Wittwer (2003) described the use of this method in sampling *Populus deltoids* in Cimmaron National grassland in Kansas, USA. Others, such as Thompson et al. (2006) have successfully employed this technique for estimating tree density in boreal forest stands. The method of k-tree sampling is widely applied and accepted as a fast way of sampling in temperate forest. However the technique received major attention in the last two decades when greater experience of its application has been gained and it has been particularly promoted for difficult terrain

conditions. One of the examples of such applications is distance-based tree density estimation in tropical dry savannahs in West Africa (Picard et. al., 2005). Other examples include that of Hall (1991) in Afromontane catchment forests and Lynch and Rusydi (1999) who compared the performance of distance sampling with other sampling methods for inventorying teak plantations in Indonesia.

Distance (k-tree) sampling is favored when fixed-area plot sampling is too difficult or costly (Sheil et al. 2003; Picard et al. 2005). The k-tree method is faster and suitable for inventorying irregular clustered population (Kleinn, 2006a). The major problem found with this distance sampling is the bias of estimates and this limits the wider applications of the method (White et. al., 2008). A large number of density estimators have been proposed to combat the problem of bias and a lack of biometrical robustness by adjusting the number of trees and radius of the imaginary circle demarcating the outer boundary of the sample (Byth, 1982; Patil et. al., 1982; Kleinn et. al., 2006a; White et. al., 2008). The density estimators of k-tree sampling are classified into two categories: first the empirical estimator and second the model-based approach (Kleinn and Vilcko 2006a). Attempts have been made at comparing these estimators between each other and also with that of general unbiased fixed-plot and point estimators. Kleinn and Vilcko (2006a) proposed a new empirical estimator and suggested calculations for determining effective radius by taking the mean distance between the kth and $(k+1)^{th}$ tree. They found that the new estimator is superior to estimator of both Eberhardt (1967) and Prodan (1968) for uniform map patterns, but the estimators of Eberhardt (1967) is the best for random and clustered map patterns. A comparison across other plot designs showed that k-tree estimators are less good than fixed-area plot estimators and relascope sampling estimators (Kleinn and Vilcko, 2006a). In another study, Takata and Kobayashi (1980) found the Prodan (1968) estimator to be less superior than their estimator.

The comparison of model-based approaches was made by Magnussen et al. (2007) for both distance estimators and plot and point estimators. The estimator of Kleinn (2006a,b) was reported to be superior among all the distance estimators by Magnussen et al., 2007 and ranked fifth in overall comparison with plot and point estimators. On the other hand the FIXED estimator was found to top the rank among all the estimators followed by the MORISITA, PERSSON, BYTH, KLEINN, ORBIT and GAMPOI estimators (Magnussen et. al., 2007).

From the above discussion it can be concluded that k-tree estimators are less good than other plot-based and point estimators. This is due to the bias associated with the estimators. Despite this bias, k-tree estimators are still very popular due to their wide applicability in dense forest with difficult terrain conditions (Lessard et. al., 2002; Sheil et. al., 2003; Picard et. al., 2005: Kleinn and Vilcko 2006a).

The k-tree sampling like that of fixed-area plots may be used in two-stage sampling designs. Little literature has emphasized this combination up till now.

2.3.3 Adaptive cluster sampling

Another very new method gaining popularity is adaptive cluster sampling (ACS) (Thompson, 1992) where 'clusters', or patches, of the target plant or animal form the sample unit (Noon et. al., 2006; Wong et. al., 2007). As the fruit tree patches occur singly as well as in clusters in Pindos National Park, this method could be very useful for achieving reliable results. The technique is efficient for populations which are clustered and found very rarely (Acharya, 2000). Another advantage of this sampling method is to generate a wealth of covariate data (Wong et. al., 2007) which is very useful to understand the ecology of the species. The technique has been successfully used for sampling clustered population of *Schima wallichi* in Nepal (Acharya, 2000). The technique is not suitable for species having a uniform distribution which may lead to the complete enumeration of the site.

In this technique the enlargement of the simple random plots is followed if the desired condition is met, for example target element (trees) is found in the plot (Thompson, 1992). Hence it is not considered as the plot size of its own but an adapted response design. A large number of population estimators were proposed for this design. However the most popular are Horwitz-Thompson (1952), Hansen-Hurwitz (1943) and Rao-Blackwell method (Cochran, 1977).

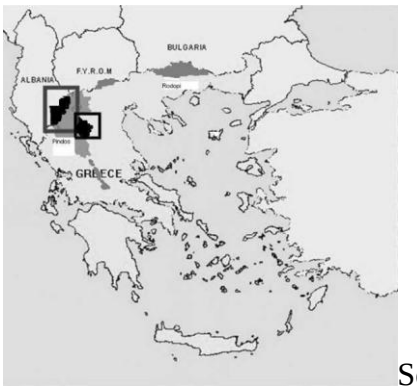
The precision and bias associated with ACS is discussed in detail by Thompson (1991, 1992), Thompson et al. (1992) and Thompson and Seber (1996). Traditional sampling designs are often associated with bias for estimating the population parameter especially when the population is rare and having varied spatial distribution (Thompson and Seber, 1996; Christman, 2000). The ACS method introduced by Thompson (1992) address these problems and proved efficient in reducing the bias and increases the precision of the population estimates (Thompson, 1992; Noon et. al., 2006). Christman (1997)

compared the ACS design for finite population and found efficient compared with conventional designs of simple random sampling and BSEC (Balanced sampling excluding contiguous units) (Hedayat et. al., 1988a).

ACS when combined with two stage sampling produce more efficient results as both the designs are well quantified and successfully applied for sampling rare species separately. Salehi and Smith (2005) has adopted this combination and successfully applied in sampling freshwater mussel population. In another study Christman (2002) also used such combination for comparing the efficiency of different empirical estimators of ACS.

Chapter 3 The Northern Pindos National Park

The Pindos National Park (PNP) is located in the western region of Greece sharing boundaries with Albania in the west (Fig. 1). The northern part of the PNP lies within two prefectures, Grevena on the northeastern side and Zagori on northwestern side (Fig. 2). Both the prefectures consist of a number of villages and monastery sites, both abandoned and active. Short descriptions of these sites are as follows.



Source : Mertzanis (2008)

Figure 1 Geographical location of Pindos National Park, Greece

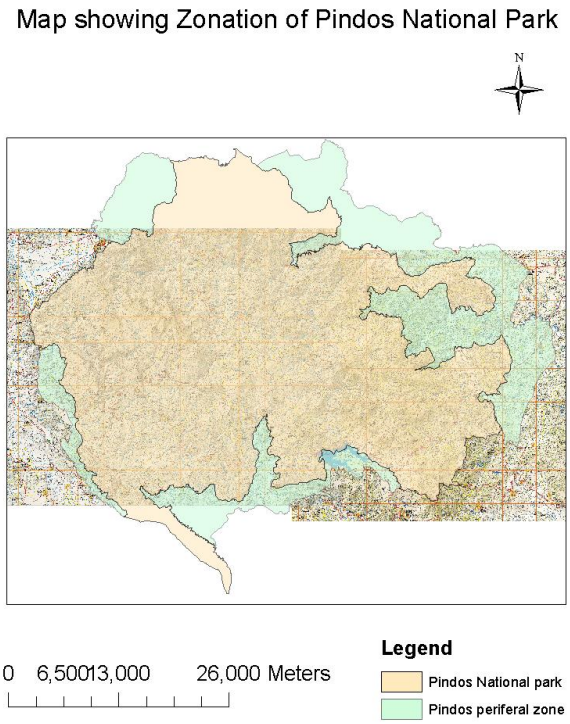


Figure 2 GIS map showing the boundaries of Pindos National Park

3.1Grevena

The first study site is the part of PNP lieing within the prefecture of Grevena in the northeastern part of the Pindos mountain range which extends over 800 km^2 of mixed forest and agricultural ecosystems (Mertzanis, 2008). The area is characterized by dense forest of oak (*Quercus* spp.), black pine (*Pinus nigra*) and beech (*Fagus* spp.) alongwith openings (bare land and grass) and small-scale cultivation (wheat and corn) (Kanellopoulos et. al., 2006). Trees such as wild cherry, wild apple, berries, prunus, chestnuts and other fruit-yielding species (for a complete list refer to Appendix 1) are widely cultivated by the local people in their homegardens as well as in orchards. There are also a number of abandoned villages and monastery sites present in Grevena where the probability of occurrence of these fruit trees is also high. Altitude ranges from 500 m to 2200 m asl. Mean annual precipitation (1990-1994) is 589 mm Mean monthly temperatures range from a minimum of -3.4° C to a maximum of 28.2 °C (Mertzanis, 1999). Almost all the native European mammal species including the brown bear (*Ursus arctos*) wolf (*Canis lupus*), roe deer (*Capreolus capreolus*), chamois (*Rupicapra rupicapra*), wild cat (*Felis sylvestris*), wild boar (*Sus scrofa*) and otter (*Lutra lutra*) are present here (Mertzanis, 2008). Grevena is well connected with other prefectures through a road network consisting of both highways and comparatively less developed forest roads. The newly built Ignatia highway connects Grevena with other countries neighbouring to the north and east.

3.2 Zagori

The second study site is the part of PNP within the prefecture of Zagori which is located in the northwestern part of the Pindos mountain range (Mertzanis, 2008). It extends over 850 km^2 with an altitude of 600 m to 2200 m asl. The prevailing vegetation types are similar to those in Grevena. The only difference is the presence of continuous forest cover with little opening. Mean monthly temperatures range from a minimum of -3.1° C to a maximum of 27.4° C. Mean annual precipitation is 814 mm. All the mammal species present in Grevena are also present here. Moreover the site is subject to human disturbance because of the construction of a road network (1.5 km/km^2) for timber transport and also due to hunting pressure (Mertzanis et. al., 2005).

3.3 Stratification and sampling design

3.3.1 Stratification and sampling design

At the beginning of the study the field data were collected by following a two-stage unstratified cluster sampling design separately for each study site Grevena and Zagori. Each of these study sites consists of a number of villages and monasteries (which were treated as the primary unit types in the survey). These divisions of the national park based on the two parameters (geographical boundaries and type of settlement) were utilized for 'a priori' stratification in the analysis of the data. The purpose of this stratification was to find out the effect of stratification on the improvement of sampling precision. This was done by comparing an unstratified design with stratified designs.

A two-way independent stratification methodology is followed to divide the population into strata (figure 3). The number of primary units sampled in each of the strata is shown in the table 3.

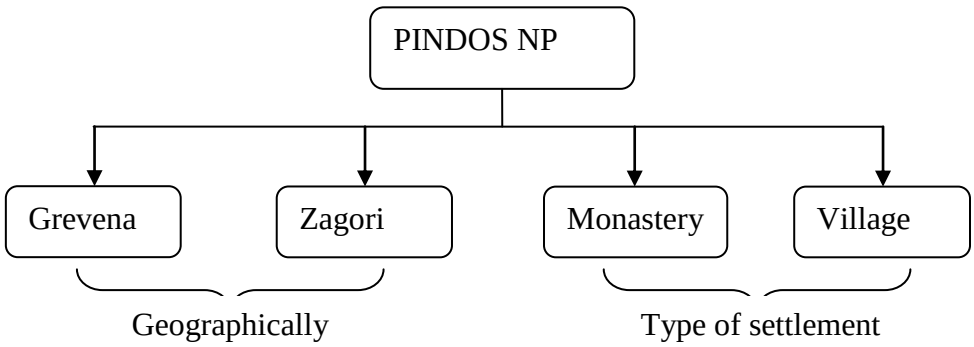


Figure 3 'A priori' stratification of Pindos National Park

Table 3 Two way stratification and allocation of primary sampling units in Pindos National Park

		Geographical boundaries		
		Grevena	Zagori	Total
Type of settlement	Village	5	5	10
	Monastery	3	1	4
	Total	8	6	14

Sampling Design

Three basic designs were adopted and used for comparison:

- Two stage unstratified cluster sampling
- Two stage stratified cluster sampling – Grevena and Zagori
- Two stage stratified cluster sampling – village and monastery

Plot Design

The data collection in the field was done by laying out plots using three different methodologies which are described in detail in chapters 4, 5 and 6.

- Fixed area plots
- k-tree plots and
- Adaptive cluster sampling

These three plot designs were initially used separately for collecting data for unstratified two-stage sampling. Due to the observed clustering of the fruit tree populations, all three plot designs were also combined in the stratified designs. This gives a total of nine unique designs (three sampling designs × three plot designs) for the comparison and optimization.

3.3.2 Allocation of first and second stage samples

The number of human settlements in the two study sites in PNP is 200 villages/monasteries (primary sample units - PSU). A first stage sample of n=14 was randomly selected[10] (without replacement) from this population across the two study sites.

For the stratified sampling the population of human settlements (PUs) was divided between the two design strata, Grevena containing N_{h1}=120 and Zagori containing N_{h2}=80. First-stage samples (n_{h1}) were selected from this population proportional to the size of the stratum (total number of villages in each stratum): n_{h1}= 8 PSUs in Grevena and n_{h2}= 6 PSUs in Zagori were randomly selected as a first stage sample.

[10] As far as possible the selection of primaries were kept random. But due to logistical problems, the selection of primary units is little biased towards selective sampling

A similar procedure was used for stratification by village and monastery sites; a random first-stage sample of n_{h1}= 10 from N_{h1}=140 for villages and n_{h2}= 6 from N_{h2}=60 for monastery sites was selected.

Within each PSU, a circular plot of 1 km was laid out on the map. The sizes of the PSUs were kept fixed at 3.14 km^2 (1 km radius). For the second-stage samples 10 random (without replacement) points were generated within each 3.14 km^2 PSU with the help of Arc GIS 9.2. Therefore a total of 140 random points were generated constituting 14 settlements (8 in Grevena and 6 in Zagori).The shape files for generating GPS-compatible maps were partly provided by the NGO Callisto and partly by the National Park authority (Fig. 4). Each of these 140 points served as the secondary sampling units (SSUs). The coordinates of each point were located in the field with a GPS (etrex Garmin).

Some of the random SSU points fell on very difficult terrain and were inaccessible. These points were excluded from the sample and replaced with newly generated random points.

From the figure 4, it is noted down that some of the PSU lies outside the National Park boundary. However the peripheral zone of the National park has been extending and basically the PSU shown outside the boundaries are within the peripheral zone.

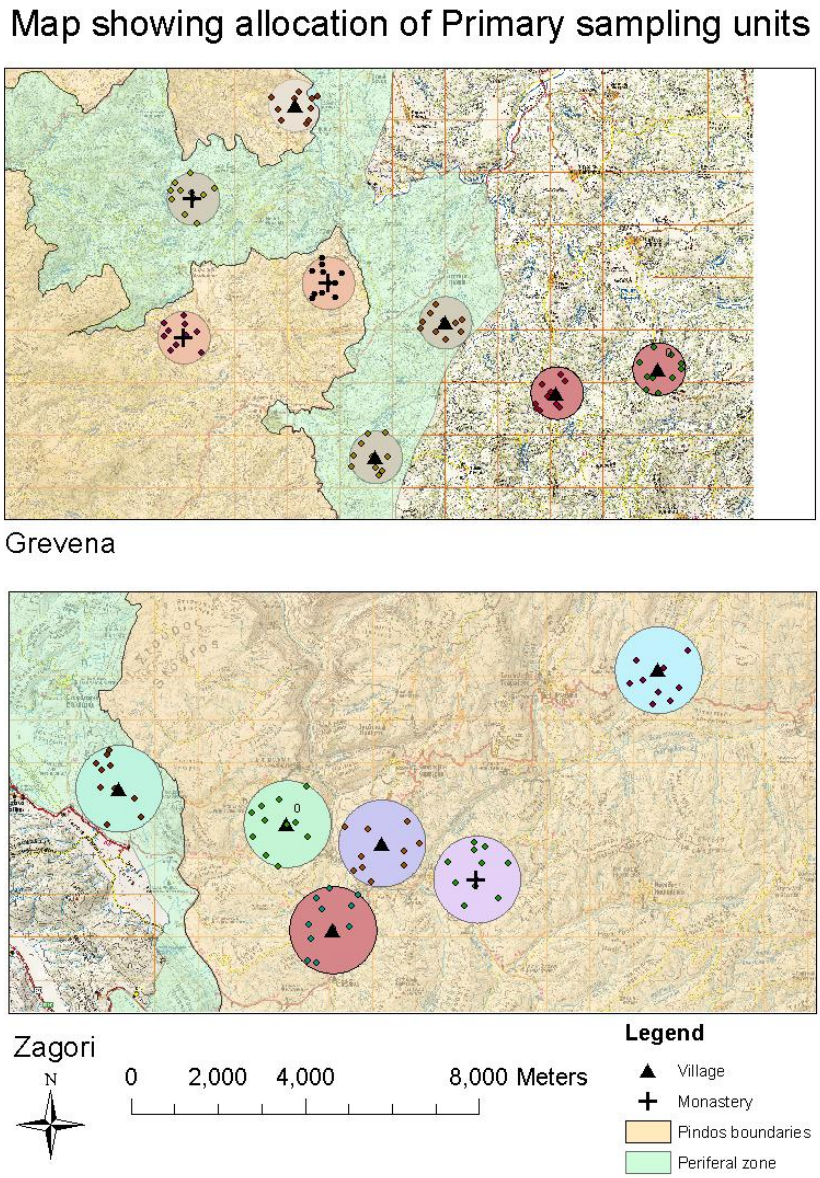


Figure 4 GIS map showing the location of the 14 primary sampling units (the dots within each primary units are the secondary sampling units) in the prefectures of Grevena and Zagori

3.3.3 Estimators for the two-stage sampling design for primary units of equal size

As the PSUs are of equal size (3.14 km^2), this is considered a special case of two-stage sampling for which estimators are given in the textbooks of De vries (1986) and Cochran (1977).

Let the population consists of N primary sample units (PSU) and M be the number of secondary sample units (SSU) in the i-th PSU. A simple random sample without replacement of n PSU´S selected from the population of N. From the i-th PSU in this

sample, a simple random sample of m SSU without replacement is drawn, and the values y_{ij} are observed on the sample elements. Then the estimators are:

Population total $\hat{Y} = NM.\frac{\sum_i^n \sum_j^m y_{ij}}{nm}$

Population mean per PSU $\hat{\bar{Y}} = M.\frac{\sum_i^n \sum_j^m y_{ij}}{nm}$

Population mean per SSU $\hat{\bar{\bar{Y}}} = \frac{\sum_i^n \sum_j^m y_{ij}}{nm}$

Total variance $v\hat{a}r\,\hat{Y} = (NM)^2.[\frac{1-\frac{n}{N}}{n}.\frac{S_b^2]}{m} + \frac{\frac{n}{N}\left(1-\frac{m}{M}\right)}{nm}.S_p^2$ and

Variance per SSU $v\hat{a}r\,\hat{\bar{\bar{Y}}} = \left(\frac{1}{(NM)^2}\right).v\hat{a}r\,\hat{Y}$

Variance per PSU $v\hat{a}r\,\hat{\bar{\bar{Y}}} = \left(\frac{1}{N^2}\right).v\hat{a}r\,\hat{Y}$

With $S_P^2 = \frac{\sum_i^n \sum_j^m (y_{ij} - \bar{\bar{y}}_i)^2}{n(m-1)}$ variance within village and

$S^2b = m.\frac{\sum_i^n (\bar{\bar{y}}_i - \bar{\bar{y}})^2}{n-1}$ Variance between villages

Where $\bar{\bar{y}}_i = \frac{\sum_j^{m_i} y_{ij}}{m_i}$ is the sample mean per SSU

3.3.4 Estimators for two-stage stratified sampling for primary sample units of equal size

The estimators for unstratified and stratified two-stage sampling are the same if the individual stratum is considered to be an unstratified two-stage sample. The total and variance for each stratum is estimated by using the equations shown in section 2.1.2 for two-stage unstratified sampling. The total across all the strata is obtained by summing individual strata, and variance is obtained by summing individual strata variances (Shiver and Borders, 1996).

The estimators given above were used for all the nine unique combinations of design (three sampling designs × three plot designs) together with the individual estimators of fixed-area plots, k-tree samples and adaptive cluster samples (figure 5).

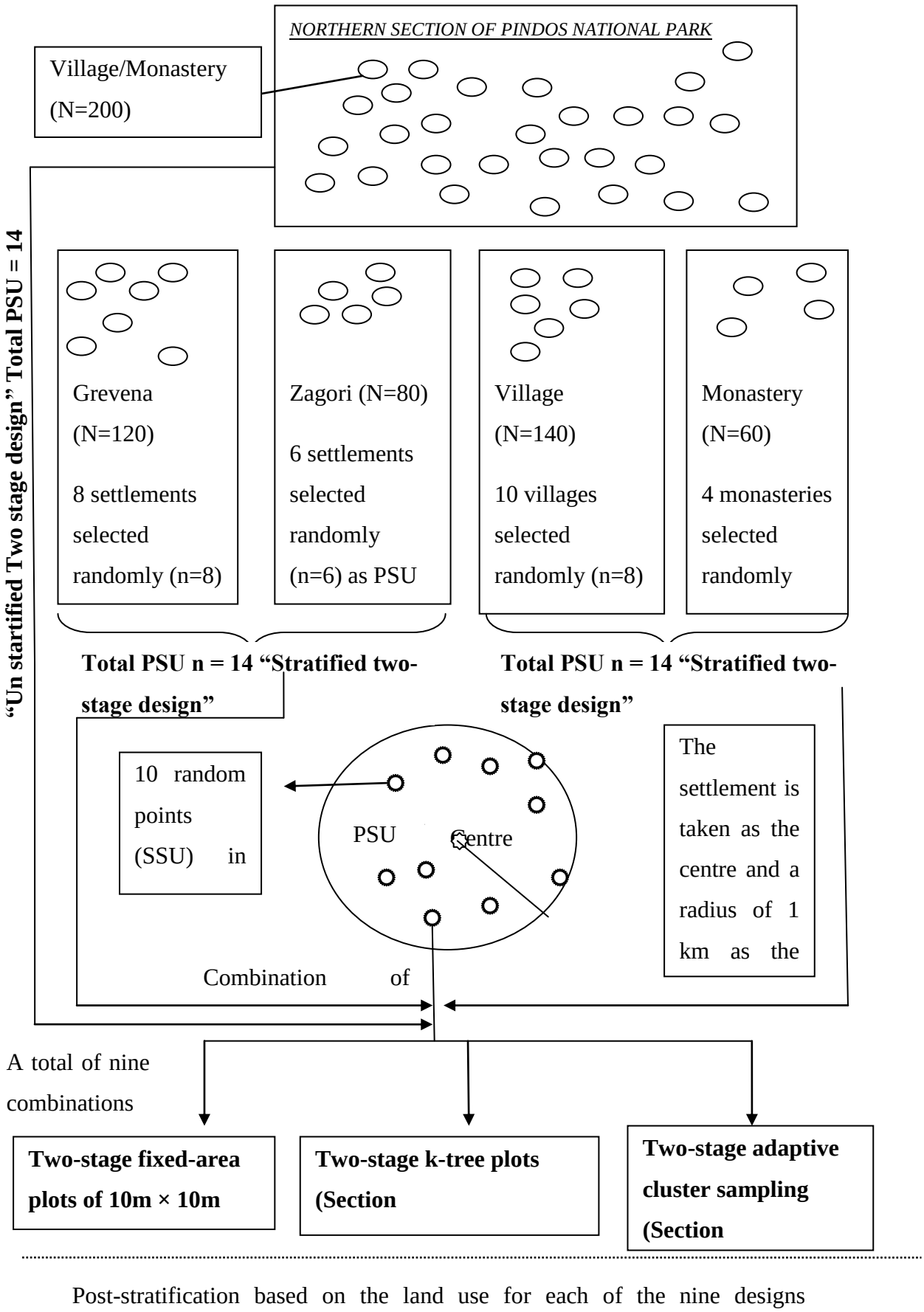


Figure 5 Flowchart showing the different sampling designs and their combinations together with allocation of PSU's and SSU's

3.4 Logistical considerations

Each of the first-stage units, i.e. village and monastery settlements were visited by vehicle. The start point was the centre of the settlement where its general description was noted. An appropriate strategy was made to visit all the SSUs prior to departure so as to reduce the travelling time between them. The travelling between the SSUs was mostly done by foot due to the lack of roads and difficult terrain. The travelling time between the PSU´s and SSU´s were not noted because of the non-uniformity in visiting the PSU's (by vehicle) and SSU's (by foot). However, an approximation of the ratio of these was used in investigating the optimal sampling strategy. The time taken in each of the SSUs for observation was recorded individually and an average calculated. The measure of time as a substitute for cost was used for developing an optimal sampling strategy (Wong et. al., 2007).

Chapter 4 Two-stage fixed-area plots

4.1 Field application

The sub-samples of 10 random points generated at the second stage (section 3.3.2) within each of the 14 settlements (PSUs) were located through GPS. At each of these points a square plot of size 10 m×10 m was laid out with a tape. A slope correction for each side of the square plot was made with the help of a clinometer and slope correction

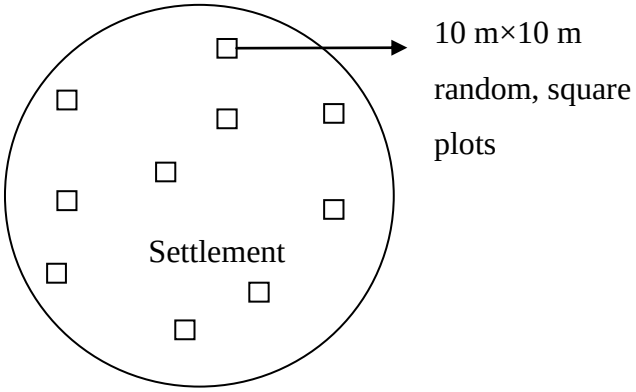


Figure 5 Layout of fixed-area plots at the second stage of two-stage cluster sampling

tables by following the methodology described by Kangas and Maltamo (2006) and Wong et al. (2007). In some of the places where the terrain is undulating, the slope correction was done more than once for each side of the plot. This slope correction was important for estimating accurate between plot variance (refer section 4.2) as the slope degree was highly variable between the plots.

Attributes measured

All the fruit trees present in each plot were measured for the following attributes:

- Species
- Number of separate stems at forking[11] point
- Diameter of each stem
- Height at which the diameter is recorded
- Crown position of each stem

[11] Forking is the splitting of a tree above the ground level (http://www.nativetreesociety.org/multi/low_branching_or_forking_trees.htm). The fruit trees in Pindos National Park generally found forked just above the ground level (low forking)

- Stem health

From these the following attributes were calculated:

- Number of trees per plot
- Basal area factor (BAF) of the stand

In addition, in each plot the following was recorded:

- Vegetation characteristics
- Time

As far as possible stem diameter was measured at 1.3 m height with the help of a diameter tape. However, due to the problem of the branching of fruit trees and very thin stems, the diameter of some trees was recorded at other heights. The height at which the diameter is measured was also recorded with a tape. The basal area of the stand was estimated by following the standard procedure of BAF calculation (Avery, 1975 p.166-167) where BAF is the ratio of DBH to distance from the point to the tree. The crown position and health of each tree, and the vegetation characteristics of the plot were recorded objectively on the basis of experience of the data recorder.

The time taken to sample each plot was recorded separately with respect to the sampling method (fixed-plot, k-tree or adaptive sample). The time is only recorded for the method which is followed first in the field.

Of all the attributes measured only the number of trees, species and time were used for the data analysis in this dissertation.

4.2 Analysis

The mean density of trees per plot and its variance was estimated by using the equations given in section 3.3.3 as those estimators pertain to two stage fixed area plots. There is no separate estimator for fixed-area plots (as was used for the k-tree and adaptive cluster sampling in chap 5 and 6 respectively). Similarly, the equations for stratified two-stage sampling were used to calculate the mean and variance of the individual strata and summed to estimate the total. A simple random sampling estimator was also calculated by pooling all the 140 plots (ignoring the clustering of plots).

The analysis was conducted in Microsoft Excel (the spreadsheets are shown in Appendix 2a) and an ANOVA table was generated by calculating the variance within-settlement (S_p^2) and between-settlement (S_b^2) together with pooled variance. A comparison between the designs was also made on the basis of mean and standard deviation. Summary statistics of density were calculated separately for each fruit tree species for each sampling method.

All the two-stage designs were compared on the basis of precision indicated by SE % (a statistic related to the coefficient of variation that is used by many authors (e.g. Chacko, 1965; Shiver and Borders (1996); Wong, 2007)).

4.3 Results and Discussion

The descriptive statistics for mean fruit tree density per settlement was generated along with respective standard deviation (Table 4). The result shows that in some of the settlements the density of fruit trees is quite high with a mean of 1.5 trees/100 sq.m (1.65)[12] and in others it is relatively very low with mean 0.1/100 sq.m (0.32) [13]. No trees were found in plots laid in 2 villages in Zagori). This shows that the density of fruit trees is variable amongst the 14 settlements. The high value of the between-SSU S.D. relative to the mean within each settlement shows that there is considerable within-settlement spatial variance in fruit tree density.

ANOVA was used to find the contribution of within-settlement (PSU) and between-settlement (PSU) variance (Table 5). The contribution of variance between settlements is 1.89 times that within settlements. However, the ratio was found to be not significant. To check this finding a similar analysis was carried out with the k-tree and adaptive-cluster sample data in the subsequent chapters.

[12] S.D at 95% CI
[13] S.D at 95% CI

Table 4 Descriptive statistics for the fruit tree sampling with a two-stage fixed-area plot design

Zagori	PSU type	Site code	Trees sampled/PSU	Mean trees/100m^2(SSU)	SD[a]
Grevena	Monastery	AG	11	1.1	1.45
Grevena	Village	DH	1	0.1	0.32
Grevena	Monastery	GH	6	0.6	1.07
Grevena	Village	GR	5	0.5	1.27
Grevena	Village	LA	4	0.4	0.71
Grevena	Monastery	MA	9	0.9	2.62
Grevena	Village	SP	4	0.4	0.97
Grevena	Village	VL	7	0.7	1.34
Zagori	Village	AN	0	0.0	0.00
Zagori	Village	DL	0	0.0	0.00
Zagori	Village	KL	1	0.1	0.32
Zagori	Monastery	PN	2	0.2	0.42
Zagori	Village	SK	15	1.5	1.65
Zagori	Village	VT	3	0.3	0.67
Total			68		
Average				**0.49**	**0.915**

[a] S.D at 95% C.I

Table 5 ANOVA test of the contribution of between-settlement and within-settlement variation in fruit tree density for two-stage fixed-area plot sampling

Source	SS	dof	SS/dof	Test
Between settlement	25.35	13	S^2b= 1.95	S^2b/S^2p = 1.89
Within settlement	129.78	126	S^2p= 1.03	
Total	155.13	139		

sig (0.05) F (13, 126)

Comparison between the sampling designs clearly shows the superiority of default simple random sampling (SRS) over the two-stage design. The term default has been used because the design of SRS is not the ideal one as the plots used for the analysis had been laid out according to a two-stage cluster sampling design in the field. Hence the spatial distribution of these plots depends on the location of the individual clusters (settlements). The results are in accordance with De Vries (1986) and Shiver and Bruce (1996) where they mentioned that two-stage designs are less precise than simple random designs. However one thing that should be kept in mind is that the sample size for SRS (14 settlements ×10 plots = 140) is increased over that for the two-stage designs

which has reduced the variance component drastically. Also the effort of laying out 140 plots randomly in the entire national park by following an ideal SRS methodology would be enormous, with a high cost. Hence the reduction in variance compared to the effort of sampling is marginal in SRS and therefore the comparison between the two is still valid. Under the practical constraints of resources using SRS it would probably not have been possible to lay out even 1/3rd of the number of plots that could have been laid by following a two-stage cluster design. Hence, even if the SRS performed better in statistical comparison with two-stage designs, it was not chosen for design optimization in the later chapters.

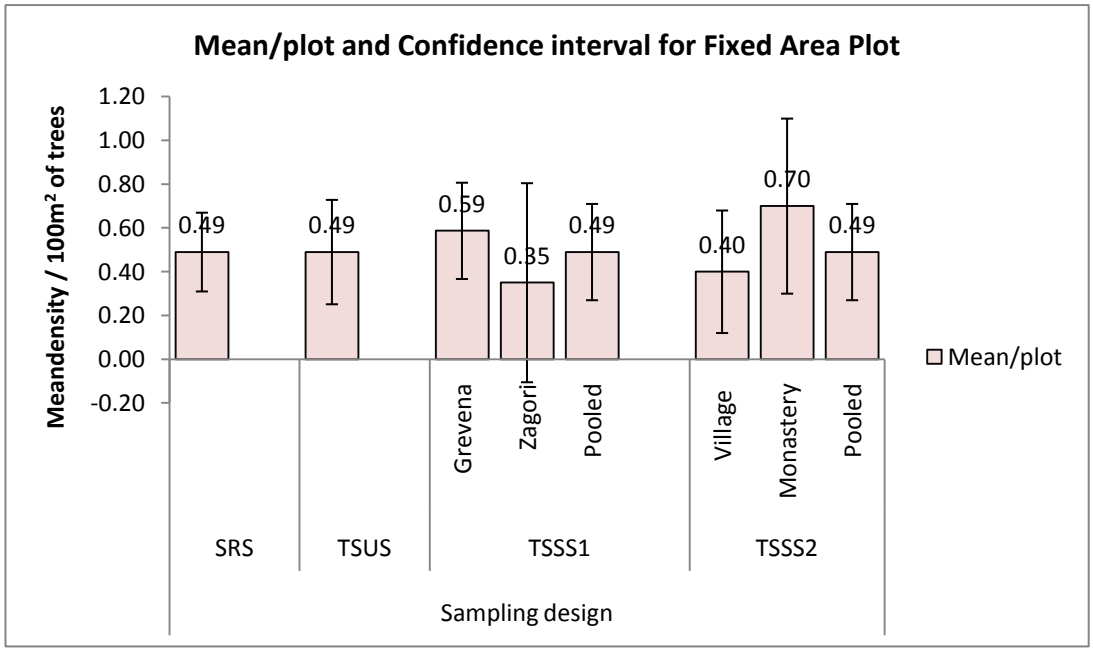
Comparison of the SE % between the sample designs shows that there is a benefit of a reduction of only < 1% of SE from unstratified to the stratified designs (Table 6). These results are a little ambiguous as the total variance does not reduce much after stratification (figure 6). This is due to the fact that the sample size in the individual strata is reduced considerably by half compared with unstratified design (Table 6). However, if the data had been collected over a large area with big sample sizes in individual strata, the story might have been different. At this stage it is not certain that the gain due to stratification is large enough to suggest that stratification by geographical boundaries and settlement type is the best approach. This means that the heterogeneity between the settlements remains the same even when 'a priori' stratification is carried out on the basis of geographical boundaries or settlement types (villages/monasteries). This suggests that the occurrence of fruit trees does not depend on region or type of settlement. Hence it can be concluded that geographical boundaries and PSU type (village/monastery) in PNP are not the best parameters to use for stratification in fruit tree survey and furthermore that pre-stratification as a whole may not be advantageous.

One important inference, however, shown in Figure 6 is that fruit tree density is far less variable in Grevena than in Zagori. In addition the density in villages is slightly more variable than in monasteries. This could be very useful information if a separate inventory plan had to be enacted in the future to know how much extra effort is needed to capture the variability in each of the regions of Zagori than Grevena or separately for villages and monasteries.

Table 6 Comparison of statistics between the sampling designs for fixed-area plot sampling of fruit trees

Sampling design for two-stage fixed-area plots								
	SRS[a]	TSUS[b]	TSSS1[c]			TSSS2[d]		
			Grevena	Zagori	Pooled	Village	Monastery	Pooled
Sample size	n=140	n=14	n=8	n=6	n=14	n=10	n=4	n=14
		m=10	m=10			m=10		
Mean/plot	0.49	0.49	0.59	0.35	0.49	0.40	0.70	0.49
SE	0.09	0.12	0.11	0.23	0.11	0.14	0.20	0.11
Mean/settlement	152.0	152.0	184.5	109.9	154.6	125.6	219.8	153.9
SE	28.86	36.4	35.2	71.4	35.5	43.5	61.3	35.6
Estimated Total	30503	30503	22137	8792	30929	17584	13188	30772
SE	5652	7287	4228	5713	7107	6094	3683	7120
SE%	18.51	23.89	19.09	64.98	22.98	34.66	27.93	23.14

[a] Simple random sampling
[b] two-stage unstratified sampling
[c] two-stage stratified sampling - Grevena and Zagori
[d] two-stage stratified sampling- village and monastery



SRS – Simple random sampling
TSUS – Two-stage unstratified sampling
TSSS1 – Two-stage stratified sampling - Grevena and Zagori
TSSS2 – Two-stage stratified sampling- village and monastery

Figure 6 Comparison of mean fruit tree density (with 95% confidence intervals) between the sampling designs for fixed-area plot sampling

The two-stage unstratified fixed-area plot design is judged to be superior among all the two-stage designs based on minimal SE% (Table 6). This result is due to the 'a priori stratification' factors that were chosen for the sampling proving to be inappropriate; they did not greatly reduce the heterogeneity in the individual strata. The potential for a priori stratification could be further researched by identification of more such factors and evaluating their use for impact on sample variance. One method that could be used for this is to test the effect of post stratifying the sample based on candidate factors such as land use, or to test the efficacy of alternative factors on the data for individual species (section 4.4). However, at this stage I have selected the two-stage unstratified fixed-area plot design as the best sampling design for further comparison with the best designs of k-tree and adaptive-cluster sampling.

4.4 Species wise comparison of descriptive

A species wise descriptive were generated without considering the design as a two-stage sample. This is done to find out the variability lies between the species which could help in estimating the effort needed to sample individual species.

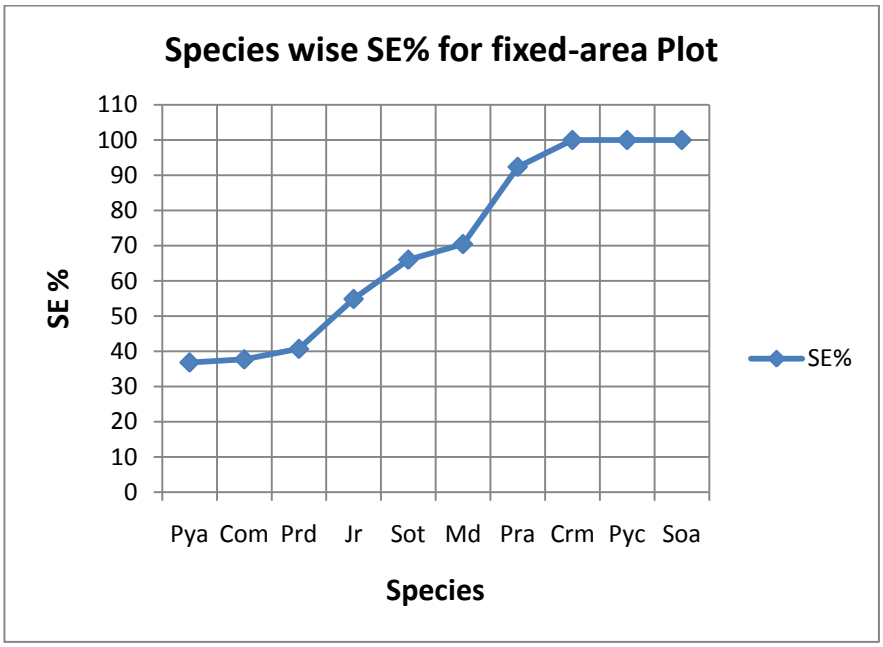


Figure 7 Relative rarity of the species in rank order shown in terms of their SE % for fixed-area plot sampling. Com, *Cornus mas*; Crm, *Cretaegus monogyna*; Jr, *Juglans regia*; Md, *Malus domestica*; Pra, *Prunus avium*; Prd, *Prunus domestica*; Pya, *Pyrus amigdaliformis*

The individual fruit tree species varied in density from 320 to 8074 trees per hectare (Table 7), which indicates the huge variation in the sampling effort that would be required to survey individual species with reliable precision across the whole national park. *Sorbus accuparia* (SE%-100), *Pyrus communis* (SE%-100), *Cretaegus monogyna* (SE%-100) and *Prunus avium* (SE%-92.39) would require extremely high sampling effort to estimate their density. Care would be needed in formulating inventory plans for these species and the costs involved would need to be justified by their importance.

Table 7 Descriptive statistics for species for fixed-area plot sampling. MD is mean density

Species	Trees sampled	M.D/ Plot	M.D/ village	Total density	MD/ ha	% density	SE%
Cornus mas	16	0.11	35.89	7177	5127	23.53	37.72
Cretaegus monogyna	3	0.02	6.73	1346	961	4.41	100
Juglans regia	7	0.05	15.70	3140	2243	10.29	54.88
Malus domestica	2	0.01	4.49	897	641	2.94	70.46
Prunus avium	2	0.01	4.49	897	641	2.94	92.39
Prunus domestica	13	0.09	29.16	5831	4165	19.12	40.70
Pyrus amigdaliformis	18	0.13	40.37	8074	5767	26.47	36.85
Pyrus communis	1	0.01	2.24	449	320	1.47	100
Sorbus accuparia	1	0.01	2.24	449	320	1.47	100
Sorbus torminalis	5	0.04	11.21	2243	1602	7.35	66.03
Total	**68**	**0.49**	**152.51**	**30503**	**21788**	**100**	**-**

Chapter 5 Two-stage k-tree sampling

5.1 Field application

The sampling strategy for the k-tree method differs from that of the plot-based method. Here, for the second-stage sample, instead of laying out square plots, the 'k' nearest trees to a single point was recorded (Kleinn 2006a). Hence the terminology "Two stage k tree sampling" is used throughout the text.

One hundred and forty k-tree points were sampled in the field, with 10 at each of the random points located by GPS within each of the 14 settlements. The points selected here are exactly the same points used for laying out fixed-area plots (refer figure 5 for clarification). Therefore some of the trees sampled by fixed-area plots were also sampled in k-tree[14]. From each of these points the distance to the k nearest trees was measured. The distance of the $(k+1)^{th}$ tree was also measured solely to estimate the effective radius (the mean of the distance to the kth and $(k+1)^{th}$ trees) and thus effective area of the plot. The slope correction for distance was calculated as recommended by Wong et al. (2007). The same attributes measured for the trees in the fixed-area plots were also measured for the selected k-trees.

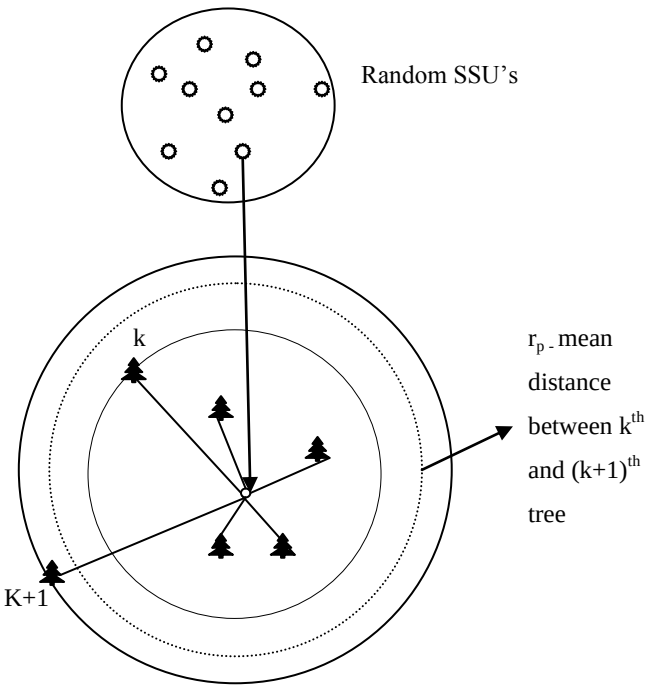


Figure 7 Lay out of the k-tree plots (SSU)

[14] Numbers were allotted to each tree by the first method applied to avoid repeated measurements

5.2 Analysis

5.2.1 Density estimation

For the fruit tree density calculation the Kleinn (2006a) emperical estimator was used:

$$r_{md} = \frac{1}{2}(d_{k+d_{k+1}}) \qquad \text{and} \qquad \frac{Density}{ha} = \frac{10{,}000}{\pi r_p^2}.k$$

Where d_k is the distance from the centre point to the kth tree and

$d_{k+1}$is the distance to the $(k+1)^{th}$ tree,

and where r_p is the effective radius of the plot p (calculated as the mean distance between k^{th} and $(k+1)^{th}$ trees)

The plot is termed the k_p tree plot.

Each of these k-tree plots was considered as a SSU for the two-stage sampling design with the PSUs remaining the same settlements of size 3.14 km^2 as the other sampling methods. All the per-plot value were converted into per 100 m^2 (instead of per ha to compare with fixed area plots of 10m×10m) by multiplying them by the respective expansion factor based on the effective area of each individual plot. This was done to enable summation of the plot data and compare the results with the other plot designs. The tree densities from the k-tree and fixed-area plots were standardised to 100 m^{-2}. Although the shape of former plots is circular and latter is square, these differences were not considered in the comparisons. The total mean and variance were calculated using the equation given in section 3.3.3 for two-stage k-tree plot sampling, with and without stratification. Estimators for both the unstratified and stratified two-stage k-tree plots was calculated by replacing the k_{ij} values with y_{ij} values in the equation for the two-stage cluster sampling design, where k_{ij} = density of trees 100 m^{-2}.

5.2.2 Bias

At the beginning of the study we kept the k value fixed as six. However, due to the unexpected scattered and clumped distribution of the fruit trees that we discovered, it was not possible to find six nearest trees at all 140 sample points. The possible approach could be to fix the lower 'k' value for all the plots and estimate the density following Kleinn (2006a) estimator. But following the same would have caused drastic reduction in the number of trees sampled and thus affected the robustness of the design. Therefore

I had to use a variable value of k in the survey. Because the estimator of per plot tree density of Kleinn (2006a) is valid only for fixed k values there is bias in the calculations of density made with these data. The bias we found here is expected to be negative compared with fixed 'k' values. However, I have had to use this estimator as there I have found no other density estimators valid for variable k values in the literature. This indicates both a disadvantage of the k-tree method for inventory of scattered and clumped trees and a research need for the development of a density estimator that would give unbiased results for variable k values.

5.3 Results and Discussion

The descriptive statistics obtained by the k-tree method show that the trend in the mean between the settlements is similar to that from the fixed-area plots (Table 8). The highest mean obtained with the k-tree method is 2.9 fruit trees per 100 m^2 (for village SK) and the minimum is 0 per 100 m^2 for only one settlement (village AN) (with a further three having 0.1 per 100 m^2). The village AN was found '0' value in both the k-tree method and the fixed-area plot method. But fixed-area plot method had an additional settlement having '0' value (Tables 4 and 9).

Between-settlement variance (6.84) was 98 times greater than within-settlement variance (0.07) with the k-tree sampling method and the ratio was significant (Table 9). There is therefore a real difference between the settlements (PSUs). This suggests that the fruit tree populations in the plots within each settlement are alike and differ a lot with those of plots in settlements. These ANOVA results are in contrast with those of the fixed-area plots (section 4.3) where no difference between the settlements was found. With the k-tree design the contribution of the between-settlement component is 99.85% of the total variation. These results are in accordance with Shiver and Borders (1996) who showed 99.8% of variation being between stands in a pine forest.

Table 8 Descriptive statistics for two-stage k-tree sampling

Block	PSU type	Site code	Trees sampled/PSU	Mean trees $100m^{-2}$(SSU)	S.D[a]
Grevena	Monastery	AG	28	1.63	0.37
Grevena	Village	DH	2	0.10	0.10
Grevena	Monastery	GH	27	0.30	0.01
Grevena	Village	GR	22	0.14	0.00
Grevena	Village	LA	3	0.01	0.00
Grevena	Monastery	MA	25	0.53	0.02
Grevena	Village	SP	15	0.28	0.04
Grevena	Village	VL	14	0.28	0.03
Zagori	Village	AN	0	0.00	0.00
Zagori	Village	DL	12	0.09	0.01
Zagori	Village	KL	10	0.15	0.01
Zagori	Monastery	PN	9	0.09	0.01
Zagori	Village	SK	30	2.85	0.41
Zagori	Village	VT	18	0.19	0.01
Total			215		
Average				**0.47**	**0.073**

[a] S.D at 95% C.I

Table 9 ANOVA of the variance components for two-stage k-tree sampling

Source	SS	dof	SS/dof	Test
Between settlement	88.92	13	S^2b= 6.84	S^2b/S^2p = 97.71*
Within settlement	8.82	126	S^2p= 0.07	
Total	97.74	139		

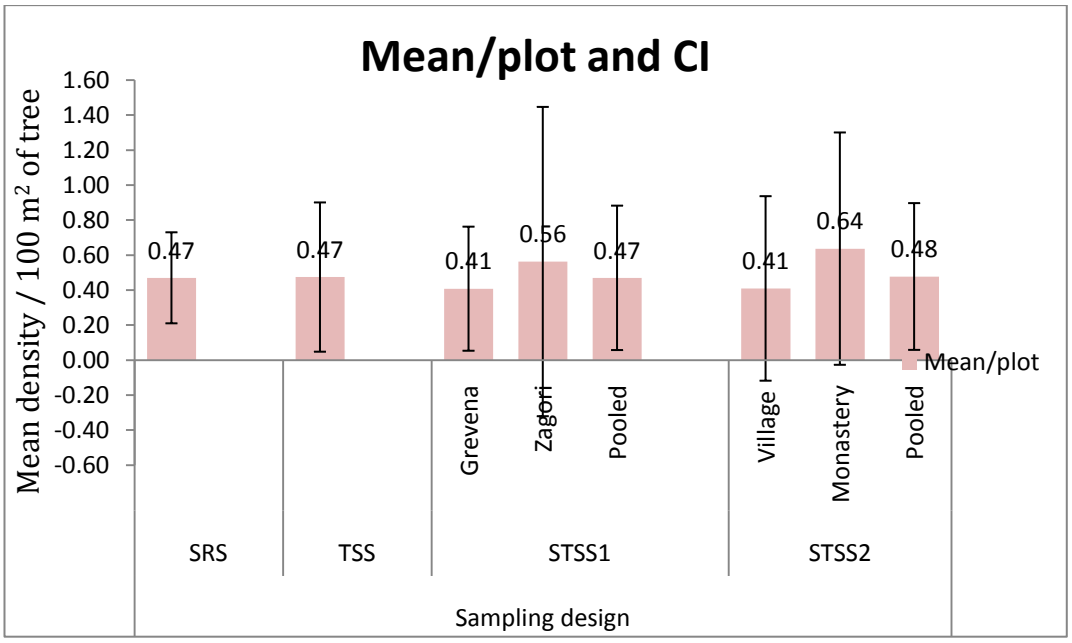
* sig (0.05) F (13, 126)

The comparison between the individual sampling designs yields similar results for k-tree sampling as for fixed-area plot sampling. The default simple random sampling (SRS) performed better than the two-stage designs (Figure 9, Table 10). The stratification factors used for the two-stage design did not reduce much of the pooled variance for the strata. There is a benefit of only a 1% reduction in SE for the two stratified designs over the two-stage unstratified design (Table 10). The reason for that is the same as for the fixed-area plot designs in which the sample size of plots within each individual stratum is too low to adequately show the actual reduction of variance that might occur with the use of pre-stratification. The result could be different if there had been bigger sample sizes of plots in the individual strata.

These results from both the k-tree method and fixed-area plot method show that the occurrence of trees does not depend on geographical location or on PSU (settlement) type. Therefore, post-stratification was again used to investigate whether other land use factors better explained the spatial variation in fruit tree density.

With the k-tree samples, as for the fixed-area plot samples, the confidence interval for Grevena was far less than for Zagori (Figure 9). Similarly, the confidence interval for monasteries as a percentage of the mean was again less than for villages (Table 10). Both sample methods show that the density of fruit tree populations in Grevena is less variable than in Zagori and in monastery settlements is less variable than in villages.

I conclude that the unstratified design is superior to the stratified designs for two-stage k-tree sampling. This is because the ‘a priori’ stratification factors chosen for the sampling proved to be inappropriate and did not help in reducing heterogeneity in the individual strata. Therefore the unstratified design is chosen for further comparison with the two-stage unstratified fixed-area plot design and the best design of adaptive cluster sampling.



SRS – Simple random sampling
TSUS – Two-stage unstratified sampling
TSSS1 – Two-stage stratified sampling - Grevena and Zagori
TSSS2 – Two-stage stratified sampling- village and monastery

Figure 8 Comparison of mean fruit tree density (with 95% confidence intervals) between the sampling designs for k-tree sampling

Table 10 Comparison of statistics (tree density) between the sampling designs for k-tree sampling

	Sampling Design							
	SRS	TSS	STSS1			STSS2		
			Grevena	Zagori	Pooled	Village	Monastry	Pooled
Sample size	n=140	n=14 m=10	n=8	n=6 m=10	n=14	n=10	n=4 m=10	n=14
Mean/plot	0.47	0.47	0.41	0.56	0.47	0.41	0.64	0.48
SE	0.13	0.21	0.18	0.44	0.21	0.26	0.33	0.21
Mean/settlement	148.9	148.9	128.0	176.7	147.5	128.5	199.9	149.9
SE	40.82	66.9	55.6	138.7	64.8	82.8	104.2	65.8
Estimated Total	29780	29780	15365	14137	29502	17990	11995	29984
SE	8164	13389	6677	11097	12951	11585	6252	13165
SE%	27.74	44.96	43.45	78.50	43.90	64.40	52.12	43.91

SRS – Simple random sampling
TSUS – Two-stage unstratified sampling
TSSS1 – Two-stage stratified sampling - Grevena and Zagori
TSSS2 – Two-stage stratified sampling- village and monastery

Chapter 6 Two-stage adaptive cluster sampling

6.1 Field Application

The adaptive cluster design (ACS) was executed together with the two-stage fixed-area plots. The same 10 random points in each settlement were also used with ACS to lay out plots of 10 m×10 m. Up to this stage there were no difference in the field methods used for the two types of sampling. The methodology followed for the second stage of the ACS sampling was as described by Thompson (1992), though he described its application for a primary stage of simple random sampling. If the initial sample plot had more than one fruit tree (i.e. satisfies the condition $y_{ij}>1$), then it was included in the sample. For each of these plots a second set of four 10 m x 10 m plots adjacent to each of its edges was laid out around it (figure 10). For each of these second-set plots,if they satisfied the same condition of $y_{ij}>1$, a further third set of 10 m x 10 m plots adjacent to each of its edges (where there was not already a sample plot) was laid out around it and the process continued until none of the newly laid out plots satisfied the condition. The process is shown diagrammatically in figure 10.

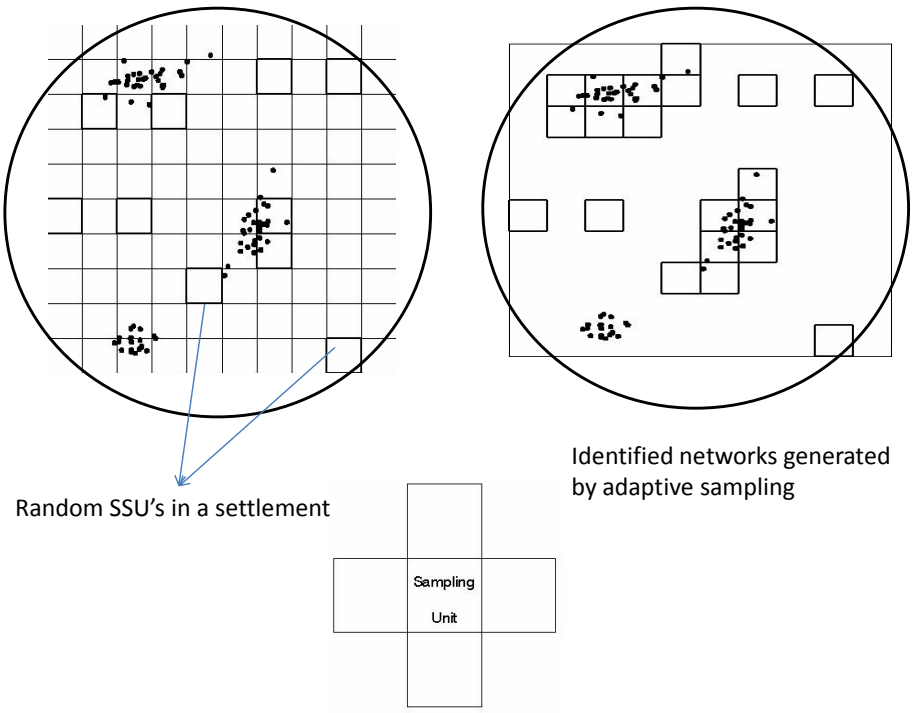


Figure 9 Layout of the two-stage adaptive cluster sampling adapted from Moser, 2004

6.2 Analysis

Adaptive cluster sampling needs high effort in the field (Acharya et. al., 2000). Considering the terrain conditions in Pindos national park and non availability of sufficient personnel in the field, the sample size obtained in this pilot survey was not enough to generate the reliable density estimates. However the data collected was collected in a team of two and the other colleague has used the data for generating a health model of the fruit trees in the orchards.

The second-stage samples (SSUs) were considered as simple random samples within each PSU (settlement) and the Hansen-Hurwitz estimator was used to calculate the mean and variance for each PSU (Hansen and Hurwitz, 1943). Subsequently, each of the PSUs were considered as the first stage of a two-stage sample, and between-settlement (S^2b) and within-settlement (S^2p) variance was calculated along with pooled variance by using the equations for two stage sampling (section 3.3.3).

Note that the difference between this sampling method and the two-stage fixed-area plot and k-tree methods is in the calculation of mean and variance per settlement. For ACS the Hansen-Hurwitz estimator is used whereas for the other two methods simple random estimators and Kleinn estimators were used, respectively. All the other estimators remain the same amongst the three methods.

6.3 Results and Discussion

The descriptive statistics for mean fruit tree density showed a completely different trend compared with fixed-plot sampling and k-tree sampling. The highest mean estimated here is 0.76 (0.39)[15] fruit trees per 100 m^2 (monastery AG) compared with 1.65 (table 4) and 2.9 (table 8) fruit trees for fixed-plot and k-tree respectively. The lowest value is obtained '0' which is same between the three designs. However a total of seven villages showed the value 0 compared with two in fixed-plot and one in k-tree sampling. The overall mean density of fruit trees per 100 m^2 was also found very low (0.16) compared with fixed-plot (0.49) and k-tree (0.47). The reason for large deviation of mean density of fruit trees here compared with other designs is the small sample size obtained in each village which has affected the efficiency of the design. If the sample size within each

[15] SD at 95% CI

village were considerable, the design could be comparable with other designs. Therefore the further comparison of this design is not discussed in the later chapters.

Between settlement variance (0.53) was 13.25 times more than the within village variance (0.04) and the ratio was found significant. The results are in accordance with k-tree sampling where ratio was also found significant. However the margin of difference between the two is quite high.

The reason for the large deviation in the descriptive statistics here compared with other designs is the small sample size obtained in each settlement which has affected the efficiency of the design. If the sample size within each settlement were considerable, the design could be comparable with other designs. Therefore the further comparison of this design is not discussed in the later chapters.

Table 11 Descriptive statistics for two-stage adaptive cluster sampling

Block	PSU type	Site code	Trees sampled/PSU	Mean trees/100m^2(SSU)	S.D[a]
Grevena	Monastery	AG	117	**0.76**	0.39
Grevena	Village	DH	0	**0.00**	0.00
Grevena	Monastery	GH	0	**0.00**	0.00
Grevena	Village	GR	47	**0.25**	0.24
Grevena	Village	LA	0	**0.00**	0.00
Grevena	Monastery	MA	46	**0.27**	0.27
Grevena	Village	SP	12	**0.40**	0.26
Grevena	Village	VL	11	**0.22**	0.22
Zagori	Village	AN	0	**0.00**	0.00
Zagori	Village	DL	0	**0.00**	0.00
Zagori	Village	KL	0	**0.00**	0.00
Zagori	Monastery	PN	0	**0.00**	0.00
Zagori	Village	SK	25	**0.36**	0.35
Zagori	Village	VT	0	**0.00**	0.00
Total			258	**0.16**	0.19
Average				**0.16**	**0.19**

[a] S.D at 95 C.I

Table 12 Table 12 ANOVA of the variance components for two-stage adaptive cluster sampling

Source	SS	dof	SS/dof	Test
Between settlement	6.89	13	S^2b= 0.53	S^2b/S^2p = 13.25*
Within settlement	0.352	126	S^2p= 0.04	
Total	7.243	139		

* sig (0.05) F (13, 126)

Chapter 7 Comparison of the methods: fixed-area plot versus k-tree sampling

7.1 Statistical approaches

The designs chosen for comparison were the best two-stage designs from the evaluation of the range of designs from each of fixed-area plot (chapter 4), k-tree plot (chapter 5) and adaptive cluster sampling (chapter 6). The two-stage adaptive cluster sampling was not included in the comparison as the data obtained were found to be not reliable due to limitations in the field (refer chap 6). For the other two methods the two-stage unstratified designs were selected for comparison. The data analysed for each design separately in chapters 4 and 5 were pooled for comparison (Table 13).

Table 13 Comparison of fruit tree density statistics between two-stage fixed-area plot and two-stage k-tree sampling

	Fixed area plots			**k-tree sampling**		
No. of trees sampled out of 140 plots	68			215		
No. of trees sampled/plot	**0.48**			**1.53**		
	Mean	SE	SE%	Mean	SE	SE%
Per plot	0.49	0.11	**23.89**	0.47	0.21	**44.95**
Per settlement	152.5	36.43		148.9	66.94	
Estimated total	30502	7287		29780	13389	
Sig. at F (13, 13) and 95% CI						

The designs were compared mainly on the basis of two parameters:

- the size of the captured variable of interest, i.e number of fruit trees sampled
- the precision achieved by each method, i.e SE %

The k-tree method performed much better than the fixed-area plot method in terms of the number of trees sampled (215 and 68 respectively, Table 13). This sample size of the variable of interest is a direct measure of the reliability of the method (Wong et. al., 2007). In terms of precision (SE %), the inference is opposite as the fixed-area plot method was more precise (23.89%) than the k-tree method (44.95%). The results are in accordance with the findings of Kleinn (2006a) and Magnussen et al. (2008) who also found that plot-based methods were more precise than the k-tree method.

7.2 Field work considerations

The experience of the observer in the difficult field conditions of the Pindos National Park (PNP) is also important in comparing the designs. The parameters chosen for this comparison are:

- ease and time taken to lay out the plots in the field
- ease and time taken to make the observations in the plots

The establishment of k-nearest tree samples was found to be easier than the fixed-area plots, as reported by Lynch and Rusydi (1999), Lessard et al. (2002), Picard et al. (2005) and Kleinn (2006a) because there was no need for initial demarcation of the boundaries of the plot which were, instead, estimated at the end. However, in PNP great difficulty was found in making the observations in many of the k-tree samples because at sample points only a few or no fruit trees were found within the vicinity of the point. As a result, when the distance from the point to the nearest 'k' trees was large this created a major problem in measuring the distances between the trees and the point. The net result of this extra effort was to greatly increase the area sampled over that with the fixed-area sample points, leading to the benefit of including many more trees in the sample, but at the cost of much greater time spent on the field work (paragraph below). An alternative strategy would be to keep the k value of the plots low (Kleinn and Vilcko 2006b). However, with a tree density of just 0.47 per 100 m^2 even a k value of 1, would entail an average of twice the sample area for each k-tree sample as for each fixed-area plot (Table 13).

The mean time taken to lay out and observe each fixed-area plot of 10 m × 10 m (100 m^2) was 13 minutes which is less than that of each k-tree sample (1668 m^2, 25 minutes). The time taken in the plots was less, this is because in many plots the trees were absent and the time taken was minimal. But those plots were also included for the calculation of time. The time taken to sample per hectare between the designs were also compared by multiplying the per plot time with an expansion factor. The mean time taken to lay out and observe k-tree plots was far less 150 minutes ha^{-1}compared with 1300 minutes ha^{-1} for fixed-plots. However, the time taken to observe each secondary plot was a small factor in the overall time taken to carry out the inventory, because the time taken to locate each primary sample area was much greater due to the difficult terrain and inaccessibility of the study area.

Chapter 8 Effect of Post stratification

8.1 Analysis

The pre-stratification based on geographical region and settlement type had very little benefit for the effectiveness of the sampling (chapter 4 and 5). The gain in SE% for both fixed-area plot and k-tree sampling was only approximately 1%. Therefore in this chapter I post-stratify the samples based on land use to investigate whether this would be a more effective factor to use in stratification. This analysis is conducted only for the fixed-area plots and k-tree sampling omitting the adaptive cluster sampling due to its small sample size.

No estimator for post-stratification of two-stage sampling is described in detail in any textbook and only a very few publications even mention such an analysis for two-stage sampling. There is therefore a knowledge gap in this analytical methodology. Therefore instead of using a specific two-stage post-stratified estimator, I had to use a normal stratified simple random sampling estimator. Shiver et al. (1996) reported that the gain in using a specific post-stratified estimator for random sampling as suggested by Cochran (1977) is very small.

The sizes of the strata that would be produced by the post-stratification were unknown before the field survey. They were estimated afterwards as proportional to the number of SSUs falling into each stratum. For example, if a total of 11 SSUs were in orchards out of the total of 140 SSUs sampled, then this ratio (11/140) was multiplied with that of the total number of possible plots at this sampling density that would fit in the whole national park i.e. 200×314 plots which gave the size of the orchard.

After estimating the stratum sizes, stratified random sampling estimators (Cochran 1977) were used to calculate the mean and variance for each stratum together with pooled mean and variance.

8.2 Results and Discussion

The results of the post-stratification were analysed as a comparison between the default simple random sampling and default post-stratified random sampling. The term default is used here as the sample plots were laid out in the field according to a two stage design which was then treated as simple random sampling for the analysis.

The descriptive statistic for both fixed-area plot and k-tree sampling (Table 14) shows similar results with the highest mean number of trees in the orchard land-use stratum (Figures 11, 12). This is not surprising as the fruit trees in the national park have a clustered distribution and the occurrence of these trees is high in and around orchards! The density was higher in meadows than in dense forest. This may be due to the high level of shade in the forest beneath the crowns of oak and beech trees. However, riverine forests showed higher fruit tree density than other forests. This may be due to the high seed dispersal rate near to the river.

There is little benefit of post-stratification in terms of smaller SE% which retains a very similar value when pooled across the land uses to that obtained with unstratified simple random sampling (Figures 11, 12, Table 14). This is because the sample size of plots in each of these six land use strata (farm, forest, meadow, orchard, other and riverine forest). This is a different situation from the pre-stratification which was done in different ways each of which divided the plots between only two strata.

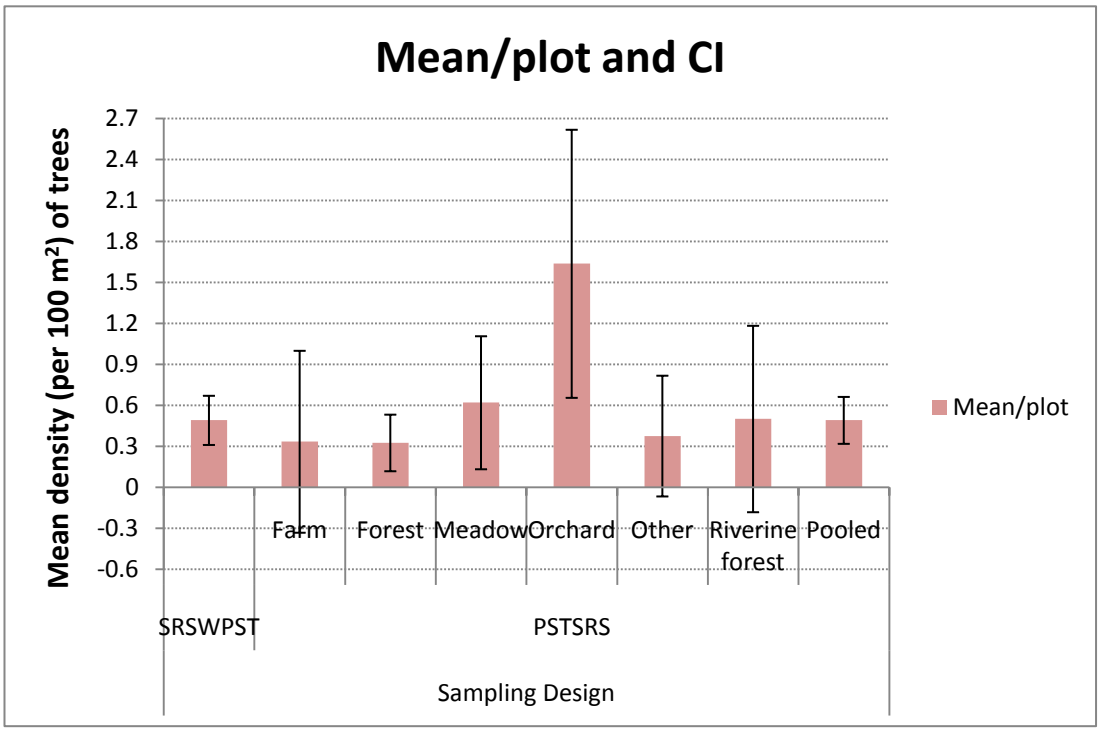
With such a small sample size the gain in SE% between the two designs were approximately same.

It is evident that in this pilot survey none of the three tested stratification parameters (two used for pre-stratification of the sampling, and one only in post-stratification analysis) were effective in reducing variance. However, in a large-scale inventory the situation might be different with the larger sample size in each stratum. In this case I recommend land use as the best parameter to be used in stratification.

Table 14 Comparison of fruit tree density statistics between sampling designs for fixed-area plot and k-tree sampling with post-stratification

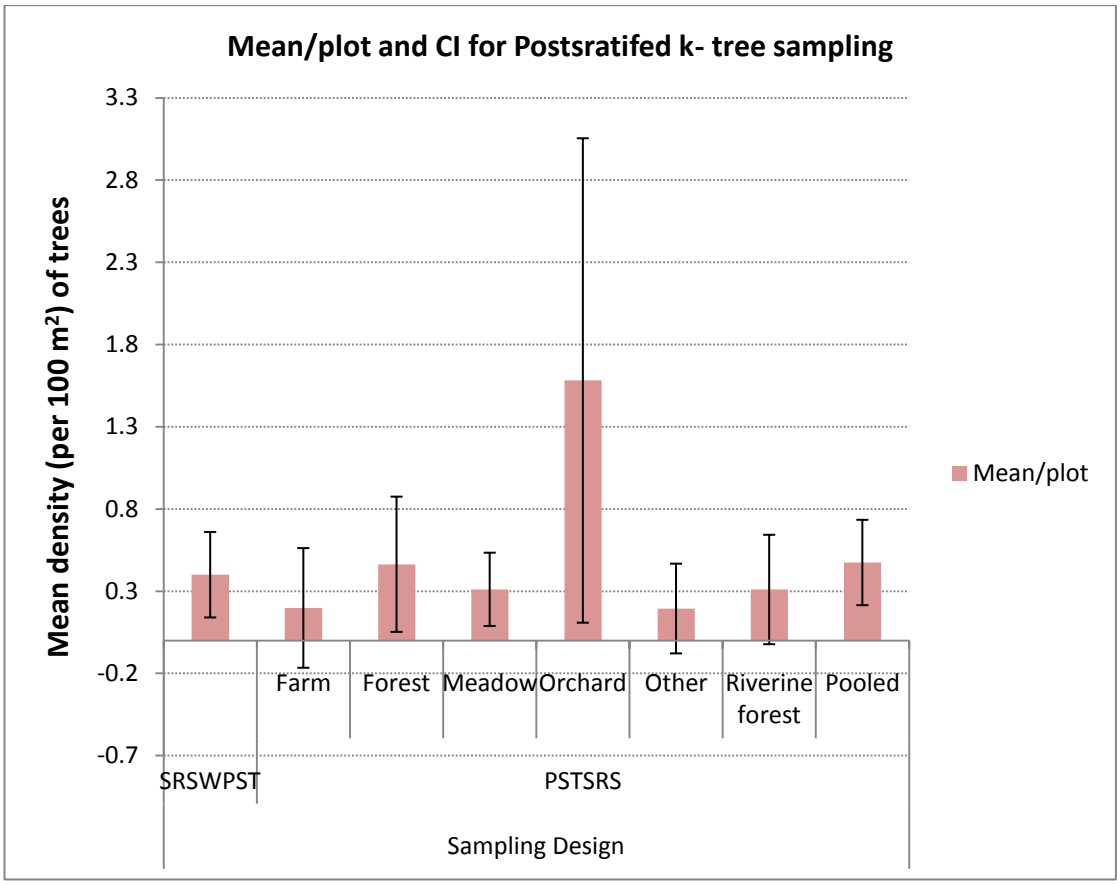
		Fixed area plots				k Tree sampling			
		SRSWPST		PSTSRS		SRSWPST		PSTSRS	
Land use	No. Of trees sampled	Mean/plot	SE %	Mean/plot	SE %	Mean/plot	SE %	Mean/plot	SE %
Farm	9			0.33	99.89			0.20	91.87
Forest	77			0.32	31.89			0.46	44.34
Meadow	21			0.62	39.33			0.31	35.80
Orchard	11			1.64	29.99			1.58	46.57
Other	16			0.38	58.94			0.19	70.17
Riverine forest	6			0.50	68.24			0.31	53.65
Pooled	140	0.49	18.4	0.49	17.68	0.47	27.74	0.47	27.37

SE at 95% CI with t =2
SRSWPST - Simple random sampling without post stratification
PSTSRS - Post stratified simple random sampling



SE at 95% CI with t =2
SRSWPST - Simple random sampling without post stratification
PSTSRS - Post stratified simple random sampling

Figure 10 Comparison of mean fruit tree density (with 95% confidence intervals) between the sampling designs for fixed-area plot sampling with stratification.



SE at 95% CI with t =2
SRSWPST - Simple random sampling without post stratification
PSTSRS - Post stratified simple random sampling

Figure 11 Comparison of mean fruit tree density (with 95% confidence intervals) between the sampling designs for k-tree sampling with stratification

Chapter 9 Design optimisation

After comparison of all the possible designs (chapters 4, 5, 6 and 7), it is evident that two-stage design without 'a priori' stratification is the best to use. Furthermore, post-stratification on the basis of land use had virtually no effect on variance (chapter 8).Therefore, this pilot survey provides no evidence for recommending that stratification would improve the efficiency of a full survey.

Whichever strategy is used to determine the primary distribution of sample units, a decision needs to be made on the form of sample units to use, and this pilot survey enables the comparison of fixed-area plots and k-tree sampling.

9.1 Optimisation of designs without consideration of cost

9.1.1 Analysis

Sample size optimisation was carried out by following the method given by Chacko (1964) and De vries (1986). The optimisation was based on the achievement of the highest precision with the best combination of primary and secondary sampling units. To do this the values of n (PSU) and m (SSU) which produce the minimum variance were determined (Chacko, 1964).

The method used is:

$$\hat{var}\,\hat{\bar{\bar{Y}}} = \frac{N}{n}.\left[\sum_{i}^{n} M_i^2 .\frac{M_{i-m_i}}{M_i}.\frac{S_i^2}{m_i} + N.\frac{N-n}{N}.\frac{\sum_i^n (\hat{Y}_i - \hat{\bar{Y}})^2}{n-1}\right]$$

When the primary units are of equal size, the formula reduces to:

$$\hat{var}\,\hat{\bar{\bar{Y}}} = \frac{1-\frac{n}{N}}{n}.\frac{S_b^2}{m} + \frac{n}{N}.\frac{1-\frac{m}{M}}{nm}.S_P^2$$

The desired level of precision was set as $\hat{var}\,\hat{\bar{\bar{Y}}}$. For example, if we set a margin of error of 20% (CI %), i.e a standard error of 10%, the $\hat{\bar{\bar{Y}}}$ (mean per SSU) should be multiplied by a factor of 10/100 and squared (Chacko, 1964). All the other variables remain the same. The values of n (PSU) and m (SSU) were estimated by inserting the values of all

the variables. The value of m is dependent on n. Thus the corresponding m values were obtained by varying the values of n.

Similarly SE% values of 10%, 15% and 20% were set as the basis for comparison, and the corresponding sample sizes (n and m) were obtained for each precision. The maximum number of PSUs (n) was set at 48.

9.1.2 Results and Discussion

Various combinations of primary (n) and secondary (m) sampling units were needed to achieve the desired level of precision (10%, 15% or 20%) separately for fixed-area plot (FAP) and k-tree sampling (figures 13, 14). To achieve a SE of 10%, a high sampling intensity is needed both at the first stage and at the second stage (> 27 PSUs (with 30 SSUs per PSU) for fixed-area plot and > 27 PSUs (with 90 SSUs per PSU) for k-tree sampling respectively). The sampling intensity to achieve SEs of 15% and 20% is lower (with as few as 7 PSUs (FAP) and 9 PSUs (KTS) respectively for SE of 20%). It has been recommended in many textbooks (De Vries 1986; Shiver, et al. 1996) to sample PSU's intensively rather than SSU's as most of the variation is likely to lie between the primaries. Moreover a minimum number of SSUs within a PSU must also be set as too few plots within a PSU destroy the efficiency of sampling (Shiver and Borders, 1996).

The results shown in figures 13 and 14 are in accordance with Shiver and Borders (1996) and considering the effort, my recommendation for large-scale inventory is to fix a minimum number of PSUs, and SSUs within each PSU, to achieve the required level of precision. Therefore, combining the views of Shiver and Borders (1996) and my results, I recommend a minimum number of PSUs between 30 and 32 and of SSUs of 10-11 for fixed-area plot sampling. Similarly for k-tree sampling I recommend a higher minimum number of PSUs of 45, with 20 SSUs within each PSU. Thus, without considering cost, the sample size must not fall below these levels.

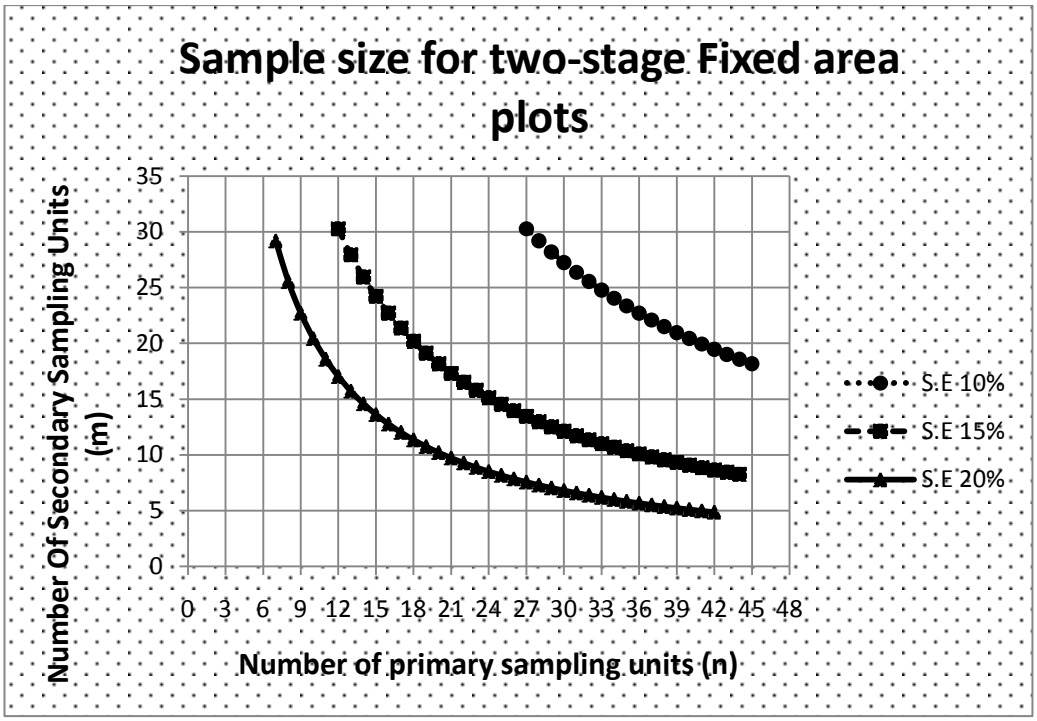


Figure 12 Modelled optimal sample size for two-stage fixed-area plots at three standard error levels (10%, 15% and 20%).

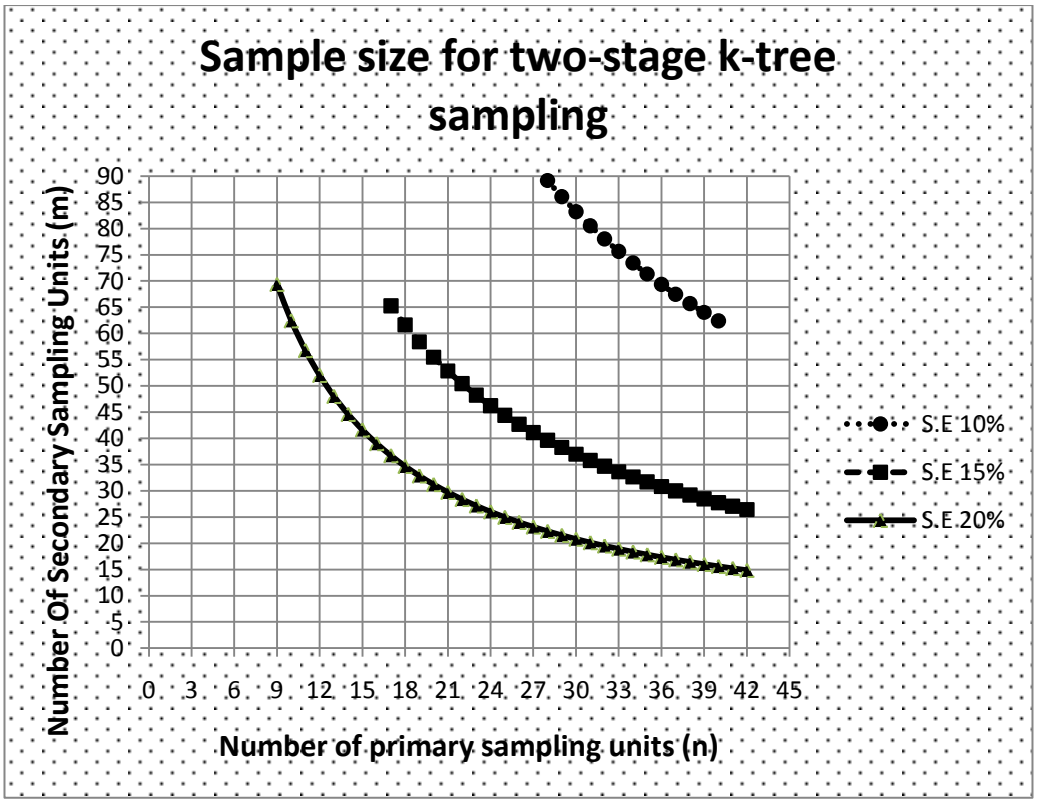


Figure 14 Modelled optimal sample size for two-stage k-tree sampling at three standard error levels (10%, 15% and 20%).

9.2 Optimisation of designs with consideration of costs

9.2.1 Analysis

The procedure for optimization with consideration of cost is similar to that without cost, except that, in addition to variance, the cost of sampling PSUs and SSUs is also considered for estimating the sample size. The ratio of cost of each PSU to that of each SSU is an approximation based on the experience in the field and kept fixed as 20:1. The overhead cost is not considered in the estimation. The analysis is designed to determine the values of n (PSU) and m (SSU) which produce the smallest variance at a fixed amount of budget (De Vries, 1986, chap. 8.3).

The method used is as follows. Let m_0 and n be the optimum sample size of SSU and PSU respectively. Then, according to Freese (1990 p. 60), the estimators for sample size optimization for two-stage cluster sampling are as follows:

Optimum size of SSU m (say m_0)

$$m_0 = \sqrt{\left(\frac{\sigma_{ii}^2}{\sigma_i^2}\right)\left(\frac{C_p}{C_s}\right)}$$

Where $\sigma_i^2 = S_p^2$, Variance within settlements and

$\sigma_{ii}^2 = \frac{S_b^2 - S_p^2}{m}$, where S_b^2 is the variance between village

C_p is the cost of locating each PSU

C_s is the cost of locating and observing each SSU

and the optimum size of PSU (n) is

$$n = \frac{\left(\sigma_i^2 + \frac{\sigma_{ii}^2}{m_0}\right)}{D^2 + \frac{1}{N}\left(\sigma_i^2 + \frac{\sigma_{ii}^2}{M}\right)}$$

where, D is the desired precision (standard error).

In this analysis, first the m_0 value was calculated at fixed cost which gives the minimum variance. Then corresponding n values were calculated for the desired precisions of 30%, 25%, 20%, 15% and 10%.

9.2.2 Results and Discussion

A calculation of minimum SSU's was conducted as mentioned in the analysis. It was found that a minimum of 15 SSU fixed-area plots are needed to minimize the within PSU (settlement) variance. Therefore, if this method is selected for a large-scale inventory a minimum of 15 SSU fixed-area plots per settlement is required. The corresponding number of PSU's was calculated for fixed 15 SSU's at 3 different levels of SE %. The results showed the corresponding number of PSU settlements to be included in the sample increases with the desired precision (figure 15). For example to achieve a SE% of 15% a total of 28 PSUs (settlements) with 15 SSUs (plots) within each PSU is required. The number of PSUs in this case is two times and number of SSUs in each and is 1.5 times that used in the pilot survey. This shows that it is feasible to achieve such precision if when using the fixed-area plot method.

Similar calculation of SSU's for the k-tree method give the result of only two SSU plots per PSU are required to capture the within PSU variation. This lower value is because the within-PSU variance is much less for the k-tree sampling method (0.07) than for the fixed-area plot method (1.03). However, two plots per PSU is a very low number and may yield a poor estimate of the true fruit tree density in the PSU settlement (Shiver and Bruce 1996). Therefore, in order to maintain the robustness of the design, 10 SSUs were used as the basis for calculating the number of PSUs required to achieve a given level of precision. To achieve a SE% of 15% with k-tree sampling a minimum of 75 PSUSs (settlements) need to be sampled (Figure 16), which is very high compared with fixed-area plot sampling (28). This shows clearly the superiority of the fixed-area plot sampling. However, it is still important to consider the ease of establishing k-tree samples, and the superior reliability of estimates of fruit tree density that were obtained with the k-tree samples than the fixed-area samples (Kleinn 2006a, Picard et.al 2005).

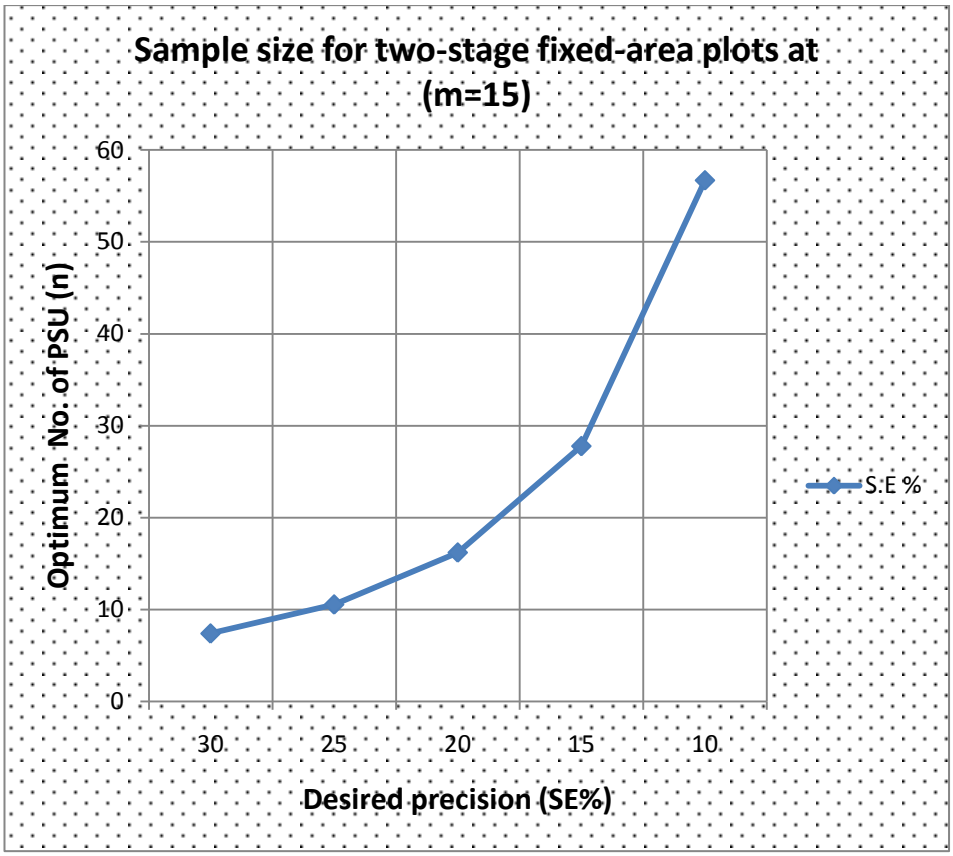


Figure 13 Modelled optimal sample size of primary sample units for two-stage fixed-area plots, with number of secondary sample units (m) fixed at 15.

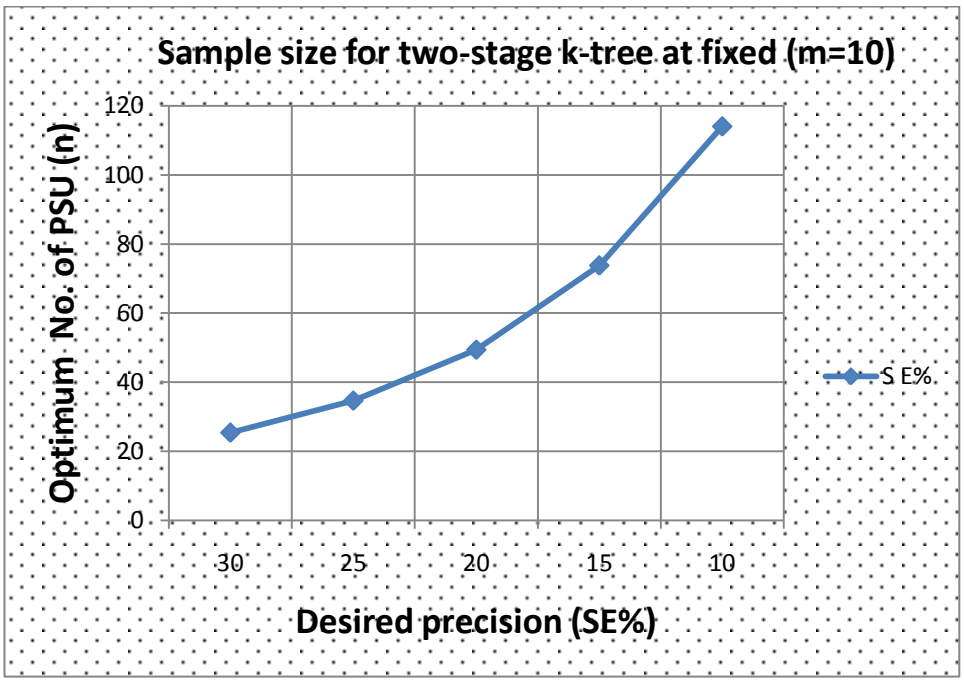


Figure 14 Modelled optimal sample size of primary sample units for two-stage k-tree sampling, with number of secondary sample units (m) fixed at 10

Chapter 10 Conclusion

10.1 Sampling fruit trees in Pindos National Park

Sampling fruit trees in Pindos National Park is challenging and poses a number of problems due to many factors identified in this study, such as difficult terrain conditions, relative rarity of fruit trees and variation in the spatial distribution of the trees with a strong tendency towards a clustered distribution. Common sampling designs such as simple random sampling may not be the best approach to quantify the density of fruit trees in such an area as the cost of sampling may not be justified by the information gained. A two-stage cluster sampling design with either fixed-area plots or k-tree sample plots proved to be an effective sampling strategy in these difficult field conditions. It may be possible to further improve either of these two designs if the sampling is pre-stratified with respect to land use such as between farms, forests, orchards etc. Therefore, the overall conclusion of this pilot study is the recommendation of a two-stage cluster sampling design for inventorying fruit trees in Pindos National Park, which should be preceded by further research to establish whether stratification based on land use is an effective pre-stratification parameter.

10.2 Optimal sampling design for clustered population

The optimal sampling design for two-stage cluster sampling depends on the available budget. If the budget is high, greater precision can be achieved. Better precision can be more easily obtained with fixed-area plots than with k-tree sample plots. However, the field-work effort required for k-tree plots is far less. Furthermore, the reliability of the estimates is also high with the k-tree sample method due to its ability to sample more trees in a given amount of time. Therefore, a trade-off between reliability and precision of the estimates must be taken into consideration in the choice of a design.

10.3 Recommendations

My personal recommendation for inventorying fruit trees in Pindos National Park is to follow the sampling strategy of two-stage design either with fixed-area plots or k-tree sampling with some form of stratification by land use. The terrain condition in Pindos National Park is difficult where capturing a variable of interest is extremely difficult. The k-tree method is effective in achieving that with a fair amount of precision. However, the method should not be applied unless a standard unbiased density estimator for variable k-values is obtained by further literature review or biometrical

research. It is hoped that this issue will receive general attention at an international level and a focus will be given to find such a technique. The other solution could be to fix a lower ‘k’ value for all the sampling points, but that needs high sampling effort. If fixed-area plot is employed care should be taken to standardize the size of the plots or kept high to capture more trees.

Limitations of the study

As far as possible the study tried to implement the designed methodology. However, due to many constraints, it was not always able to achieve this. The main limitations of the study are:

- Time constraints – Pindos National Park covers a large area and comprises many settlements, hence the field-work time allotted to achieve a representative sample size was not sufficient.
- Movement in the National Park – there was a small bias in the selection of the settlements (primary sample) against those that were very remote. There was insufficient time or resources to sample the settlements farthest from the camping place. This was despite the NGO, CALLISTO, helping me more than their capabilities could easily allow and assisting me in many ways to achieve the desired sample size.
- Lack of personnel – the field work was conducted by two persons. Although each fixed-area plot and a k-tree plot do not require much labour, adaptive cluster sampling proved to be extremely difficult to carry out by two persons in such a difficult terrain. This led to very few adaptive cluster plots being completed in the time available.

References

ACHARYA, B., BHATTARAI, G., DE GIER, A. & STEIN, A. (2000) Systematic adaptive cluster sampling for the assessment of rare tree species in Nepal. *Forest Ecology and Management,* 137**,** 65-73.

AVERY, T. G. (1983) *Forest measurements,* New York: London, McGraw-Hill.

BETHLEHEM, J. C. & KELLER, W. A. (1987) Linear weighting of sample survey data. *J. Offic. Statist,* 3**,** 141-153.

BODKIN, J. L. & UDEVITZ, M. S. (1999) An aerial survey method to estimate sea otter abundance. *Marine Mammal Survey and Assessment Methods***,** 13-26.

BORDERS, B. E., SHIVER, B. D. & CLUTTER, M. L. (2005) Timber inventory of large acreages using stratified two-stage list sampling. *Southern Journal of Applied Forestry,* 29**,** 152-157.

BYTH, K. (1982) On Robust Distance-Based Intensity Estimators. *Biometrics,* 38**,** 127-135.

CHACKO, V. J. (1965) *A manual on sampling techniques for forest surveys,* Delhi, Manager of Publications.

CHRISTMAN, M. C. (1997) Efficiency of some sampling designs for spatially clustered populations. *Environmetrics,* 8**,** 145-166.

CHRISTMAN, M. C. (2000) A review of quadrat-based sampling of rare, geographically clustered populations. *Journal of Agricultural Biological and Environmental Statistics,* 5**,** 168-201.

CHRISTMAN, M. C. (2003) Adaptive two-stage one-per-stratum sampling. *Environmental and Ecological Statistics,* 10**,** 43-60.

COCHRAN, W. G. (1977) *Sampling Techniques,* New York, Wiley.

CONGALTON, R. G. (1988) A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing,* 54**,** 593-600.

CORONA, P. & FATTORINI, L. (2006) The assessment of tree row attributes by stratified two-stage sampling. *European Journal of Forest Research,* 125**,** 57-66.

COTTAM, G. & CURTIS, J. T. (1949) A method for making rapid surveys of woodlands by means of pairs of randomly selected trees. *Ecology,* 30**,** 101-104.

DE VRIES, P. G. (1986) *Sampling theory for forest inventory : a teach-yourself course,* Berlin ; New York, Springer-Verlag.

EBERHARD.LL (1967) Some developments in distance sampling. *Biometrics,* 23**,** 207-&.

EDWARDS, T. C., MOISEN, G. G. & CUTLER, D. R. (1998) Assessing map accuracy in a remotely sensed, ecoregion-scale cover map. *Remote Sensing of Environment,* 63**,** 73-83.

FREESE, F. (1990) *Statistics for land managers : an introduction to sampling methods and statistical analysis for foresters, farmers and environmental biologists,* Jedburgh Paeony Press.

GLOWKA, L., BURHENNE-GUILMIN, F. & SYNGE, H (1994) A guide to the Convention on Biological Diversity. IUCN.

GROSENBAUGH, L. R. (1963) Some suggestions for better sample-tree measurement. *Soc. Am. Forest Proc.***,** 36-42.

HALL, J. B. (1991) Multiple-nearest-tree sampling in an ecological survey of afromontane catchment forest. *Forest Ecology and Management,* 42**,** 245-266.

HANKIN, D. G. (1984) Multistage sampling designs in fisheries research: Applications in small streams. *Canadian Journal of Fisheries and Aquatic Sciences,* 41**,** 1575-1591.

HANSEN, M. M. & HURWITZ, W. N. (1943) On the theory of populations from finite populations. *Annals of Mathematical Statistics,* 14**,** 333-362.

HEDAYAT, A. S., RAO, C. R. & STUFKEN, J. (1988a) Sampling designs excluding contiguous units. *Journal of Statistical Planning and Inference,* 19**,** 159-170.

HOLT, D. & SMITH, T. M. F. (1979) POST STRATIFICATION. *Journal of the Royal Statistical Society Series a-Statistics in Society,* 142**,** 33-46.

HORVITZ, D. G. & THOMPSON, D. J. (1952) A generalization of sampling without replacement from a finite universe. *Journal of the American statistical associations,* 81**,** 663-685.

JESSEN, R. J. (1978) *Statistical Sampling Techniques.,* New York, Wiley and Sons.

KANELLOPOULOS N, M. G., KORAKIS G, PANAGIOTOPOULOU M. (2006) Selective habitat use by brown bear (Ursus arctos L.) in northern Pindos, Greece. *Journal of Biological Research* 5**,** 23–33.

KANGAS, A. & MALTAMO, M. (2006) *Forest Inventory: Methodology and Applications,* Dordrecht, Netherlands, Springer.

KLEINN, C. & VILCKO, F. (2006) Design-unbiased estimation for point-to-tree distance sampling. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere,* 36**,** 1407-1414.

KLEINN, C. & VILCKO, F. (2006a) A new empirical approach for estimation in k-tree sampling. *Forest Ecology and Management,* 237**,** 522-533.

LANCIA, R. A., J.D. NICHOLS, AND K.H. POLLOCK. (1994) Estimating the number of animals in wildlife populations. IN BOOKHOUT, T. A. (Ed.) *Research and Management Techniques for Wildlife and Habitats 5th Edition.* Bethesda, MD., The Wildlife Society.

LESSARD, V., REED, D. D. & MONKEVICH, N. (1994) Comparing N-tree distance sampling with point and plot sampling in northern Michigan forest types. *Northern Journal of Applied Forestry,* 11**,** 12-16.

LILLESAND, T. M. (1994) Strategies for improving the accuracy and specificity of large-area, satellite-based land cover inventories. *Proceedings of the ISPRS Commision IV Symposium on Mapping and Geographic Information Systems.* Athens, GA.

LOETSCH, F. (1973) Forest inventory in federal republic of germany and other european states. *Mitteilungen Der Bundesforschungsanstalt Fur Forst-Und Holzwirtschaft***,** 73-89.

LYNCH, T. B. & RUSYDI, R. (1999) Distance sampling for forest inventory in Indonesian teak plantations. *Forest Ecology and Management,* 113**,** 215-221.

LYNCH, T. B. & WITTWER, R. F. (2003) n-tree distance sampling for per-tree estimates with application to unequal-sized cluster sampling of increment core data. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere,* 33**,** 1189-1195.

MAGNUSSEN, S., PICARD, N. & KLEINN, C. (2008) A gamma-Poisson distribution of point to k nearest event distance. *Forest Science,* 54**,** 429-441.

MANO, T. (1994) Home Range and Habitat Use of Brown Bears in the Southwestern Oshima Peninsula, Hokkaido. *Bears - Their Biology and Management***,** 319-325.

MCDONALD, T. L. (2004) Sampling rare populations IN W.L.THOMPSON (Ed.) *Sampling Rare or Elusive Species: Concepts, Designs, and Techniques for Estimating Population Parameters* Washington, Island Press.

MCLELLAN, B. N., SERVHEEN, C. & HUBER, D. (2008) (IUCN/ SSC Bear Specialist Group) 2008. Ursus arctos. The IUCN Red List of Threatened Species. IUCN.

MERTZANIS, G. (1999a) Supporting monitoring of key indices related to bear populations status and feedback : Monitoring of bear population levels and trends. NGO Arcturos ed., Thessaloniki Project LIFE96NAT/GR/003222, Final Report.& Final Report Annex 7.

MERTZANIS, G. (1999b) Status and management of the brown bear in Greece., Servheen C, Herrero S, Peyton M, editors. Status survey and conservation action plan. Greece: BEARS, IUCN/SSC Bear Specialist Group.

MERTZANIS, G., GRIVAS, C., IOANNIDIS, Y., BOUSBOURAS, D., GIANNATOS, G., PAPAIOANNOU, H., POIRAZIDIS, C. & GAETHLICH, M. (1994) The status of brown bear (Ursus arctos L.) in Greece. *Biologia Gallo-Hellenica,* 22**,** 225-232.

MERTZANIS, G., KALLIMANIS, A. S., KANELLOPOULOS, N., SGARDELIS, S. P., TRAGOS, A. & ARAVIDIS, I. (2008) Brown bear (Ursus arctos L.) habitat use patterns in two regions of northern Pindos, Greece - management implications. *Journal of Natural History,* 42**,** 301-315.

MERTZANIS, G. A. (1994) Brown Bear in Greece - Distribution, Present Status-Ecology of a Northern Pindus Subpopulation. *Bears - Their Biology and Management***,** 187-197.

MERTZANIS, Y., IOANNIS, I., MAVRIDIS, A., NIKOLAOU, O., RIEGLER, S., RIEGLER, A. & TRAGOS, A. (2005) Movements, activity patterns and home range of a female brown bear (Ursus arctos, L.) in the Rodopi Mountain Range, Greece. *Belgian Journal of Zoology,* 135**,** 217-221.

MOISEN, G. G., EDWARDS, J., T. C. & CUTLER, D. R. (1994) Spatial sampling to assess classification accuracy of remotely sensed data. IN MICHENER, W. K., BRUNT, J. W. & STAFFORD, S. G. (Eds.) *Environmental Information Management and Analysis: Ecosystem to Global Scales.* Philadelphia, Taylor & Francis.

MORRISON, L. W., SMITH, D. R., YOUNG, C. C. & NICHOLS, D. W. (2008) Evaluating sampling designs by computer simulation: a case study with the Missouri bladderpod. *Population Ecology,* 50**,** 417-425.

MOSER, E. B. (2004) *Fundamental Sampling Techniques.* *http://www.stat.lsu.edu/faculty/moser/exst7012/adaptive.pdf*

NOON, B. R., ISHWAR, N. M. & VASUDEVAN, K. (2006) Efficiency of adaptive cluster and random sampling in detecting terrestrial herpetofauna in a tropical rainforest. *Wildlife Society Bulletin,* 34**,** 59-68.

NUSSER, S. M. & KLAAS, E. E. (2003) Survey methods for assessing land cover map accuracy. *Environmental and Ecological Statistics,* 10**,** 309-331.

PARALIKIDIS, N. P., PAPAGEORGIOU, N. K., KONTSIOTIS, V. J. & TSIOMPANOUDIS, A. C. The dietary habits of the Brown bear (Ursus arctos) in western Greece. *Mammalian Biology - Zeitschrift fur Saugetierkunde,* In Press, Corrected Proof.

PATIL, S. A., KOVNER, J. L. & BURNHAM, K. P. (1982) Optimum Nonparametric-Estimation of Population-Density Based on Ordered Distances. *Biometrics,* 38**,** 243-248.

PICARD, N., KOUYATE, A. M. & DESSARD, H. (2005) Tree density estimations using a distance method in Mali savanna. *Forest Science,* 51**,** 7-18.

PRODAN, M. (1968) Punktstichprobe für die Forsteinrichtung. *Forst Holzwirt,* 23**,** 225-226.

RINGVALL, A. & KRUYS, N. (2005) Sampling of sparse species with probability proportional to prediction. *Environmental Monitoring and Assessment,* 104**,** 131-146.

ROYALL, R. M. (1976) Current advances in sampling theory - implications for human observational studies. *American Journal of Epidemiology,* 104**,** 463-474.

SALEHI M, M. & SMITH, D. R. (2005) Two-stage sequential sampling: A neighborhood-free adaptive sampling procedure. *Journal of Agricultural Biological and Environmental Statistics,* 10**,** 84-103.

SALEHI, M. & SEBER, G. A. F. (1997) Adaptive cluster sampling with networks selected without replacement. *Biometrika,* 84**,** 209-219.

SCHREUDER, H. T., WOOD, G. B., GREGOIRE, T. G. (1993) Sampling methods for multiresource forest inventory. *Wiley, New York.*

SCOTT, A. J. & SMITH, T. M. F. (1969) Estimation in multistage surveys. *J. Am. Statist. Assoc,* 64**,** 830-840.

SCOTT, C. T. & ALEGRIA, J. (1990) Fixed-versus variable-radius plots for change estimation. *State-of-the-Art Methodology of Forest Inventory : A Symposium Proceedings,* 263**,** 126-132.

SHEIL, D., DUCEY, M. J., SIDIYASA, K. & SAMSOEDIN, I. (2003) A new type of sample unit for the efficient assessment of diverse tree communities in complex forest landscapes. *Journal of Tropical Forest Science,* 15**,** 117-135.

SHIVER, B. D., BORDERS, B. E. (1996) *Sampling techniques for forest resource inventory,* New York, Wiley.

SKALSKI, J. R. (1994) Estimating wildlife populations based on incomplete area surveys. *Wildlife Society Bulletin,* 22**,** 192-203.

SMITH, D. R., CONROY, M. J. & BRAKHAGE, D. H. (1995) Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. *Biometrics,* 51**,** 777-788.

SMITH, D. R., VILLELLA, R. F. & LEMARIE, D. P. (2003) Application of adaptive cluster sampling to low-density populations of freshwater mussels. *Environmental and Ecological Statistics,* 10**,** 7-15.

SMITH, T. M. F. (1991) POST-STRATIFICATION. *Statistician,* 40**,** 315-323.

TAKATA, K. & KOBAYASHI, S. (1980) AN APPLICATION OF 6 TREE SAMPLING TO ARTIFICIAL AND NATURAL STANDS. *Journal of the Japanese Forestry Society,* 62**,** 419-424.

THOMPSON, I. D., ORTIZ, D. A., JASTREBSKI, C. & CORBETT, D. (2006) A comparison of prism plots and modified point-distance sampling to calculate tree stem density and basal area. *Northern Journal of Applied Forestry,* 23**,** 218-221.

THOMPSON, S. K. (1992) Sampling. *Wiley, New York.*

THOMPSON, S. K., RAMSEY, F. L. & SEBER, G. A. F. (1992) AN ADAPTIVE PROCEDURE FOR SAMPLING ANIMAL POPULATIONS. *Biometrics,* 48**,** 1195-1199.

THOMPSON, S. K. & SEBER, G. A. F. (1996) *Adaptive sampling,* New York, John Wiley and Sons.

THOMPSON, W. L. (2004) Sampling Rare or Elusive Species: Concepts, Designs, and Techniques for Estimating Population Parameters *Washington : Island Press,*.

TOKOLA, T. & SHRESTHA, S. M. (1999) Comparison of cluster-sampling techniques for forest inventory in southern Nepal. *Forest Ecology and Management,* 116**,** 219-231.

VALLIANT, R. (1993) Poststratification and Conditional Variance Estimation. *Journal of the American Statistical Association,* 88**,** 89-96.

VLACHOS, C. G., BAKALOUDIS, D. E., DIMITRIOU, M., KRITIKOU, K. & CHOUVARDAS, D. (2000) Seasonal food habits of the European brown bear (Ursus arctos) in the Pindos Mountains, Western Greece. *Folia Zoologica,* 49**,** 19-25.

WHITE, N. A., ENGEMAN, R. M., SUGIHARA, R. T. & KRUPA, H. W. (2008) A comparison of plotless density estimators using Monte Carlo simulation on totally enumerated field data sets. *BMC Ecol,* 8**,** 6.

WICKHAM, J. D., STEHMAN, S. V., SMITH, J. H., WADE, T. G. & YANG, L. (2004) A priori evaluation of two-stage cluster sampling for accuracy assessment of large-area land-cover maps. *International Journal of Remote Sensing,* 25**,** 1235-1252.

WONG, J., BAKER, N., BODEN, R. & KLEINN, C. (2007) Resource assessment of Non-Wood forest products: Reference Material: VOL. 2 *Non-Wood Forest Products Technical Papers* 13, F.A.O., Rome.

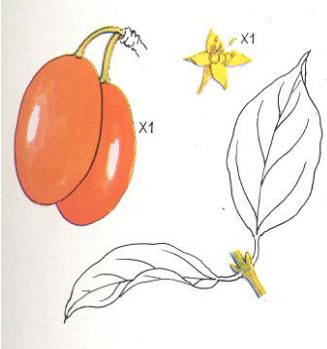
WOOD, G. B. (1990) Ground-sampling methods used to inventory tropical mixed moist forests. *Forest Ecology and Management,* 35**,** 199-206.

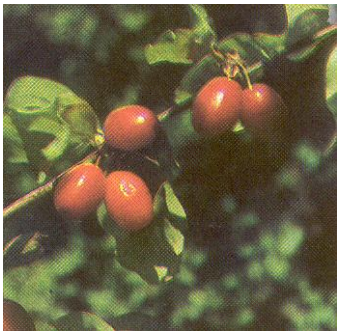
WOODING, J. B. & HARDISKY, T. S. (1994) Home Range, Habitat Use, and Mortality of Black Bears in North-Central Florida. *Bears - Their Biology and Management***,** 349-356.

Appendices

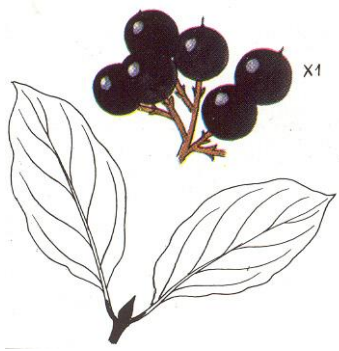
Appendix 1 List of common fruit trees in Pindos National park

(Source: NGO CALLISTO, Thessaloniki, Greece)

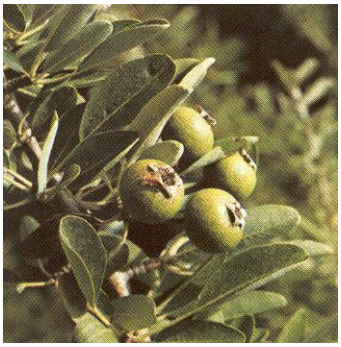
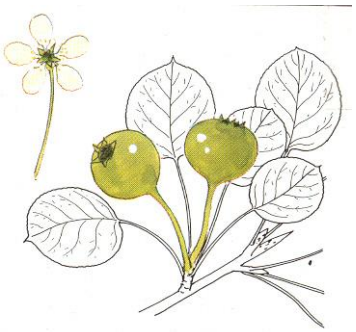




Cornus mas



Cornus sanguinea

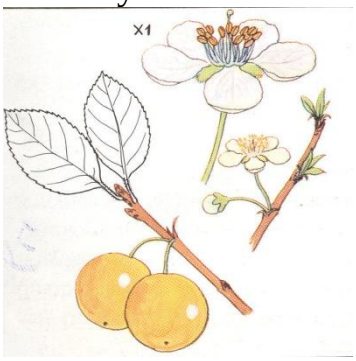



Pyrus amygdaliformis Pyrus communis

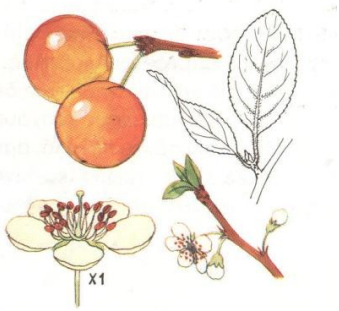
Malus sylvestris

Malus domestica

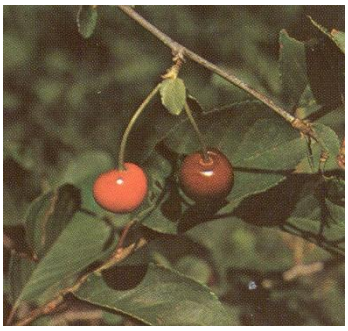
Malus dasyphylla



Prunus cerasifera

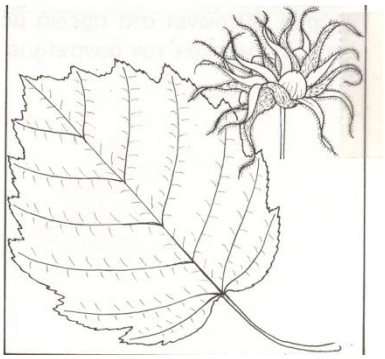


Prunus domestica



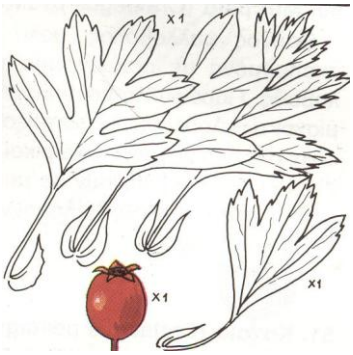
Prunus avium

Corylus avellana



Corylus colurna

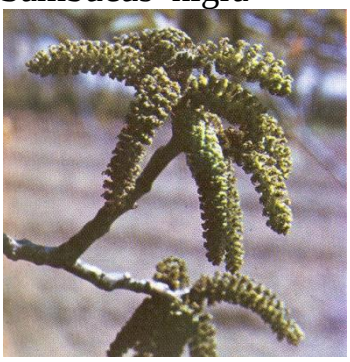
Castanea sativa



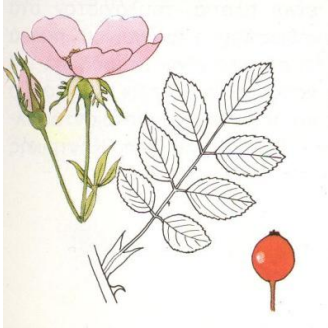
Crataegus monogyna

Sambucus nigra

Sambucus racemosa



Juglans regia

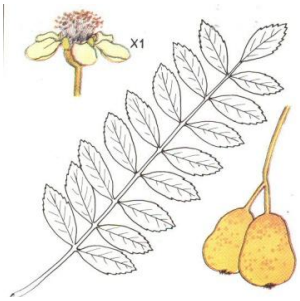


Rosa canina

Rosa arvensis

Rubus idaeus

Rubus canescens



Sorbus aucuparia

Sorbus domestica

Sorbus torminalis

Appendix 2 Calculation steps for Two-stage cluster sampling

Appendix 2a Two-stage fixed-area plot sampling

sampling design- Two stage cluster sampling (Fixed area plot method)					
Calculation of Mean, variance and standard deviation					
Number of Primary units in population (N) =			200		
Number of Primary units sampled (n) =			14		
Number of secondary units in i-th PU (M) =			314		
Number of secondary units sampled in i-th PU (m) =			10		
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village
Grevena	AG	11	1.100	345.4	2.1
Grevena	DH	1	0.100	31.4	0.1
Grevena	GH	6	0.600	188.4	1.155555556
Grevena	GR	5	0.500	157	1.611111
Grevena	LA	4	0.400	125.6	0.71
Grevena	MA	9	0.900	282.6	2.62
Grevena	SP	4	0.400	125.6	0.933333
Grevena	VL	7	0.700	219.8	1.788889
Zagori	AN	0	0.000	0	0
Zagori	DL	0	0.000	0	0
Zagori	KL	1	0.100	31.4	0.1
Zagori	PN	2	0.200	62.8	0.177777778
Zagori	SK	15	1.500	471	2.722222
Zagori	VT	3	0.300	94.2	0.455556
		68	**0.486**	**152.51429**	**1.033888881**
	Plot level	Estimated mean trees/plot	0.485714286		
		Variance/plot	0.013465002		
		SE/plot	0.116038797		
		C.I			
	Village level	Estimated mean total trees/village	152.5142857		
		Variance/village	1327.59538		
		SE/village	36.43618229		
		C.I			
		Variance within village (S2p)	1.033888881	1.0338889	population variance
		SE within village	1.016803266		
		Variance between village (S2b)	1.951648352	0.0917759	population variance
		SE between village	1.397014084		
		Ratio (S2b/S2p)	1.887677088	Steps overall variance	
	Population level	Estimated mean total trees in population	30502.85714	B/w village	within village
		Overall variance	53103815.2	0.93	0.067770701
		Overall SE	7287.236459	0.0066429	0.000484076
		C.I		51129997	1973817.94
		SE%	23.89034058	53103815	Overall variance

sampling design- Two stage cluster sampling (Fixed area plot method)						
Stratification by region: East of national park (Grevena): West of national park (Zagori)						
Number of Primary units in population (N) =			120		80	200
Number of Primary units sampled (n) =			8		6	14
Number of secondary units in i-th PU (M) =			314		314	314
Number of secondary units sampled in i-th PU (m) =			10		10	10
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village	
Grevena	AG	11	1.1	345.4	2.1	
Grevena	DH	1	0.1	31.4	0.1	
Grevena	GH	6	0.6	188.4	1.155555556	
Grevena	GR	5	0.5	157	1.611111	
Grevena	LA	4	0.4	125.6	0.71	
Grevena	MA	9	0.9	282.6	2.62	
Grevena	SP	4	0.4	125.6	0.933333	
Grevena	VL	7	0.7	219.8	1.788889	
			0.5875	1475.8	1.377361069	
Zagori	AN	0	0	0	0	
Zagori	DL	0	0	0	0	
Zagori	KL	1	0.1	31.4	0.1	
Zagori	PN	2	0.2	62.8	0.177777778	
Zagori	SK	15	1.5	471	2.722222	
Zagori	VT	3	0.3	94.2	0.455556	
Total trees observed		68	0.35	659.4	0.575925963	
			Grevena		**Zagori**	**Pooled**
	Plot level	Estimated mean trees/plot	0.5875		0.35	0.4925
		Variance/plot	0.012590413		0.051726147	0.012808732
		SE/plot	0.112207011		0.22743383	0.11317567
		C.I				
	Village level	Estimated mean total trees/village	184.475		109.9	154.645
		Variance/village	1241.364399		5099.991203	1262.889776
		SE/village	35.23300155		71.41422269	35.53716049
		C.I				
		Variance within village (S2p)	1.377361069		0.575925963	1.056787027
		SE within village	1.173610272		0.758897861	1.028001472
		Variance between village (S2b)	0.983928571		3.31	1.914357143
		SE between village	0.991931737		1.81934054	1.383602957
		Ratio (S2b/S2p)	0.714357762		5.747266511	1.811488118
	Population level	Estimated mean total trees in population	22137		8792	30929
		Overall variance	17875647.34		32639943.7	50515591.04
		Overall SE	4227.960187		5713.137815	7107.432099
		C.I				
		SE%	19.09906576		64.98109435	22.97983154
			Steps overall variance		Steps overall variance	
			B/w village	within villa	B/w village	within village
			0.933333333	0.0645435	0.925	0.072611465
			0.011666667	0.0008068	0.015416667	0.001210191
			16297918.8	1577728.5	32200138.99	439804.7098
			17875647.34	Overall var	32639943.7	Overall variance

sampling design- Stratified two stage cluster sampling (Fixed area plot method)						
Stratification by Village/Monastery						
Number of Primary units in population (N) =			140		60	200
Number of Primary units sampled (n) =			10		4	14
Number of secondary units in i-th PU (M) =			314		314	314
Number of secondary units sampled in i-th PU (m) =			10		10	10
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village	
Village	DH	1	0.1	31.4	0.1	
Village	GR	5	0.5	157	1.611111	
Village	LA	4	0.4	125.6	0.71	
Village	SP	4	0.4	125.6	0.933333	
Village	VL	7	0.7	219.8	1.788889	
Village	AN	0	0	0	0	
Village	DL	0	0	0	0	
Village	KL	1	0.1	31.4	0.1	
Village	SK	15	1.5	471		
Village	VT	3	0.3	94.2	0.455556	
			0.4		**0.633209889**	
Monastry	AG	11	1.1	345.4	2.1	
Monastry	GH	6	0.6	188.4	1.155555556	
Monastry	MA	9	0.9	282.6	2.62	
Monastry	PN	2	0.2	62.8	0.177777778	
			2.8	879.2	1.513333333	
		Total trees observed	68			
			Village		**Monastery**	**Overall**
	Plot level	Estimated mean trees/plot	0.4		0.7	0.49
		Variance/plot	0.019215666		0.038219674	0.012855
		SE/plot	0.138620584		0.195498528	0.113382
		C.I				
	Village level	Estimated mean total trees/village	125.6		219.8	153.86
		Variance/village	1894.587837		3768.307022	1267.496
		SE/village	43.5268634		61.38653779	35.60191
		C.I				
		Variance within village (S2p)	0.633209889		1.513333333	0.897247
		SE within village	0.795744864		1.230176139	0.947231
		Variance between village (S2b)	2.022222222		1.533333333	1.875556
		SE between village	1.422048601		1.238278375	1.369509
		Ratio (S2b/S2p)	4.439020059		1.013215859	2.090345
	Population level	Estimated mean total trees in population	17584		13188	30772
		Overall variance	37133921.61		13565905.28	50699827
		Overall SE	6093.760876		3683.192268	7120.381
		C.I				
		SE%	34.65514602		27.92836114	23.13916
			Steps overall variance		**Steps overall variance**	
			B/w village	**within villa**	**B/w village**	**within vill**
			0.928571429	**0.0691538**	**0.933333333**	**0.064544**
			0.009285714	**0.0006915**	**0.023333333**	**0.001614**
			36287710.04	**846211.56**	**12699164.8**	**866740.5**
			37133921.61	**Overall var**	**13565905.28**	**Overall va**

Appendix 2b Two-stage k-tree sampling

sampling design- Two stage cluster sampling (k tree method)							
Number of Primary units in population (N) =			200				
Number of Primary units sampled (n) =			14				
Number of secondary units in i-th PU (M) =			314				
Number of secondary units sampled in i-th PU (m) =			10				
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village		
Grevena	AG	28	1.630696411	512.03867	0.372916		
Grevena	DH	2	0.099196489	31.147698	0.098399		
Grevena	GH	27	0.302834795	95.090126	0.007414		
Grevena	GR	22	0.14058707	44.14434	0.002876		
Grevena	LA	3	0.00862293	2.7075999	0.000744		
Grevena	MA	25	0.525748649	165.08508	0.024431		
Grevena	SP	15	0.277239559	87.053222	0.038294		
Grevena	VL	14	0.277297635	87.071457	0.031894		
Zagori	AN	0	0	0	0		
Zagori	DL	12	0.091552733	28.747558	0.00568		
Zagori	KL	10	0.149346885	46.894922	0.005935		
Zagori	PN	9	0.087347257	27.427039	0.014987828		
Zagori	SK	30	2.854997028	896.46907	0.405812		
Zagori	VT	18	0.193432854	60.737916	0.009148		
Total trees observed		215	0.474207164	2084.6147	0.072752202		
	Plot level	Estimated mean trees/plot	0.474207164				
		Variance/plot	0.045452177				
		SE/plot	0.213195161	0.4263903			
		C.I					
	Village level	Estimated mean total trees/village	148.9010495		Steps overall variance		
		Variance/village	4481.402809		B/w village	within village	
		SE/village	66.94328054		0.93	0.0677707	
		C.I			0.006642857	0.0004841	
		Variance within village (S2p)	0.072752202	0.0727522	179117219.7	138892.68	
		SE within village	0.269726161		179256112.4	Overall variance	
		Variance between village (S2b)	6.836961573	0.6764209			
		SE between village	2.614758416				
		Ratio (S2b/S2p)	93.97600873				
	Population level	Estimated mean total trees in population	29780.2099				
		Overall variance	179256112.4				
		Overall SE	13388.65611	26777.312			
		C.I					
		SE%	44.95823283				

sampling design- Two stage cluster sampling (k-tree method)

Stratification by region: East of national park (Grevena): West of national park (Zagori)

Number of Primary units in population (N) =			120		80	200
Number of Primary units sampled (n) =			8		6	14
Number of secondary units in i-th PU (M) =			314		314	314
Number of secondary units sampled in i-th PU (m) =			10		10	10
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village	Column1
Grevena	AG	28	1.630696411	512.03867	0.372916	
Grevena	DH	2	0.099196489	31.147698	0.098399	
Grevena	GH	27	0.302834795	95.090126	0.007414	
Grevena	GR	22	0.14058707	44.14434	0.002876	
Grevena	LA	3	0.00862293	2.7075999	0.000744	
Grevena	MA	25	0.525748649	165.08508	0.024431	
Grevena	SP	15	0.277239559	87.053222	0.038294	
Grevena	VL	14	0.277297635	87.071457	0.031894	
			0.407777942	128.04227	0.072121	
Zagori	AN	0	0	0	0	
Zagori	DL	12	0.091552733	28.747558	0.00568	
Zagori	KL	10	0.149346885	46.894922	0.005935	
Zagori	PN	9	0.087347257	27.427039	0.014987828	
Zagori	SK	30	2.854997028	896.46907	0.405812	
Zagori	VT	18	0.193432854	60.737916	0.009148	
Total trees observed		215	0.516550165	1060.2765	0.073593805	
			Grevena		**Zagori**	**Pooled**
	Plot level	Estimated mean trees/plot	0.407777942		0.56277946	0.469779
		Variance/plot	0.031399242		0.195153553	0.042528
		SE/plot	0.177198313	0.3543966	0.441761874	0.206224
		C.I				
	Village level	Estimated mean total trees/village	128.0422739		176.7127504	147.5105
		Variance/village	3095.839666		19241.35974	4193.12
		SE/village	55.64027018		138.7132284	64.7543
		C.I				
		Variance within village (S2p)	0.072121		0.073593805	0.07271
		SE within village	0.268553533		0.271281781	0.269648
		Variance between village (S2b)	2.686376162		12.65283183	6.672958
		SE between village	1.639016828		3.557081927	2.583207
		Ratio (S2b/S2p)	37.24818239		171.927948	91.77482
	Population level	Estimated mean total trees in population	15365.07287		14137.02003	29502.09
		Overall variance	44580091.19		123144702.4	1.68E+08
		Overall SE	6676.832422		11097.05827	12950.86
		C.I				
		SE%	43.45460955		78.49644587	43.89811
			Steps overall variance		**Steps overall variance**	
			B/w village	**within villa**	**B/w village**	**within village**
			0.933333333	**0.0645435**	**0.925**	**0.072611**
			0.011666667	**0.0008068**	**0.015416667**	**0.00121**
			44497478.6	**82612.586**	**123088502.6**	**56199.76**
			44580091.19	**Overall var**	**123144702.4**	**Overall variance**

sampling design- Stratified two stage cluster sampling (k tree method)						
Stratification by Village/Monastery						
Number of Primary units in population (N) =			140		60	200
Number of Primary units sampled (n) =			10		4	14
Number of secondary units in i-th PU (M) =			314		314	314
Number of secondary units sampled in i-th PU (m) =			10		10	10
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village	
Village	DH	2	0.099196489	31.147698	0.098399	
Village	GR	22	0.14058707	44.14434	0.002876	
Village	LA	3	0.00862293	2.7075999	0.000744	
Village	SP	15	0.277239559	87.053222	0.038294	
Village	VL	14	0.277297635	87.071457	0.031894	
Village	AN	0	0	0	0	
Village	DL	12	0.091552733	28.747558	0.00568	
Village	KL	10	0.149346885	46.894922	0.005935	
Village	SK	30	2.854997028	896.46907	0.405812	
Village	VT	18	0.193432854	60.737916	0.009148	
Monastry	AG	28	1.630696411	512.03867	0.372916	
Monastry	GH	27	0.302834795	95.090126	0.007414	
Monastry	MA	25	0.525748649	165.08508	0.024431	
Monastry	PN	9	0.087347257	27.427039	0.014987828	
Total trees observed		215	2.546627113	799.64091	0.104937207	
			Village		**Monastery**	**Pooled**
	Plot level	Estimated mean trees/plot	0.409227318		0.636656778	0.477456
		Variance/plot	0.069455996		0.110115484	0.043944
		SE/plot	0.263545056		0.331836533	0.209628
		C.I				
	Village level	Estimated mean total trees/village	128.497378		199.9102283	149.9212
		Variance/village	6848.083421		10856.9463	4332.686
		SE/village	82.7531475		104.1966712	65.82314
		C.I				
		Variance within village (S2p)	0.0598782		0.104937207	0.073396
		SE within village	0.244700225		0.323940129	0.270917
		Variance between village (S2b)	7.475417208		4.71197824	6.646386
		SE between village	2.734120921		2.170709156	2.578058
		Ratio (S2b/S2p)	817.1641023		44.90283636	90.55527
	Population level	Estimated mean total trees in population	17989.63292		11994.6137	29984.25
		Overall variance	134222435.1		39085006.66	1.73E+08
		Overall SE	11585.44065		6251.800274	13164.63
		C.I				
		SE%	64.40065065		52.12173089	43.90515
			Steps overall variance		**Steps overall variance**	
			B/w village	**within villa**	**B/w village**	**within vill**
			0.928571429	**0.0691538**	**0.933333333**	**0.064544**
			0.009285714	**0.0006915**	**0.023333333**	**0.001614**
			134142414.8	**80020.268**	**39024905.35**	**60101.32**
			134222435.1	**Overall var**	**39085006.66**	**Overall va**

Appendix 2c Two-stage Adaptive cluster sampling

sampling design- Two stage cluster sampling (Adaptive cluster sampling method)						
Number of Primary units in population (N) =			200			
Number of Primary units sampled (n) =			14			
Number of secondary units in i-th PU (M) =			314			
Number of secondary units sampled in i-th PU (m) =			10			
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village	S.D
Grevena	AG	117	0.760599415	238.82822	0.15195669	0.3898162
Grevena	DH	0	0	0	0	0
Grevena	GH	0	0	0	0	0
Grevena	GR	47	0.247368421	77.673684	0.059242373	0.2433976
Grevena	LA	0	0	0	0	0
Grevena	MA	46	0.270588235	84.964706	0.07088621	0.2662446
Grevena	SP	12	0.4	125.6	0.068846426	0.262386
Grevena	VL	11	0.22	69.08	0.046858599	0.2164685
Zagori	AN	0	0	0	0	0
Zagori	DL	0	0	0	0	0
Zagori	KL	0	0	0	0	0
Zagori	PN	0	0	0	0	0
Zagori	SK	25	0.357142857	112.14286	0.123488886	0.3514099
Zagori	VT	0	0	0	0	0
Total trees observed		258	0.161121352	708.28946	0.037234227	0.1929617
	Plot level	Estimated mean trees/plot	0.161121352			
		Variance/plot	0.003520498			
		SE/plot	0.059333785			
					Steps overall variance	
	Village level	Estimated mean total trees/village	50.59210454		B/w village	within village
		Variance/village	347.1070229		0.93	0.0677707
		SE/village	18.63080843		0.006642857	0.0004841
		C.I			13813196.31	71084.608
		Variance within village (S2p)	0.037234227		13884280.92	Overall variance
		SE within village	0.192961725			
		Variance between village (S2b)	0.527254122			
		SE between village	0.726122663			
		Ratio (S2b/S2p)	14.16046897			
	Population level	Estimated mean total trees in population	10118.42091			
		Overall variance	13884280.92			
		Overall SE	3726.161687			
		C.I				
		SE%	36.82552565			

sampling design- Stratified two stage cluster sampling (Adaptive cluster sampling method)

Stratification by region: East of national park (Grevena): West of national park (Zagori)

Number of Primary units in population (N) =			120		80	200
Number of Primary units sampled (n) =			8		6	14
Number of secondary units in i-th PU (M) =			314		314	314
Number of secondary units sampled in i-th PU (m) =			10		10	10
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village	
Grevena	AG	117	0.760599415	238.82822	0.15195669	
Grevena	DH	0	0	0	0	
Grevena	GH	0	0	0	0	
Grevena	GR	47	0.247368421	77.673684	0.059242373	
Grevena	LA	0	0	0	0	
Grevena	MA	46	0.270588235	84.964706	0.07088621	
Grevena	SP	12	0.4	125.6	0.038726115	
Grevena	VL	11	0.22	69.08	0.046858599	
			0.237319509	596.14661	0.045958748	
Zagori	AN	0	0	0	0	
Zagori	DL	0	0	0	0	
Zagori	KL	0	0	0	0	
Zagori	PN	0	0	0	0	
Zagori	SK	25	0.357142857	112.14286	0.123488886	
Zagori	VT	0	0	0	0.455556	
Total trees observed		258	0.357142857	112.14286	0.096507481	
			Grevena		**Zagori**	**Pooled**
	Plot level	Estimated mean trees/plot	0.237319509		0.05952381	0.166201
		Variance/plot	0.007883016		0.003394145	0.003381
		SE/plot	0.08878635		0.058259292	0.058146
		C.I				
	Village level	Estimated mean total trees/village	74.51832581		18.69047619	52.18719
		Variance/village	777.2338444		334.6491303	333.348
		SE/village	27.87891397		18.29341768	18.25782
		C.I				
		Variance within village (S2p)	0.045958748		0.096507481	0.066178
		SE within village	0.214379916		0.310656532	0.257251
		Variance between village (S2b)	0.672508864		0.212585034	0.488539
		SE between village	0.820066378		0.461069446	0.698956
		Ratio (S2b/S2p)	14.63288032		2.202782953	7.382175
	Population level	Estimated mean total trees in population	8942.199097		1495.238095	10437.44
		Overall variance	11192167.36		2141754.434	13333922
		Overall SE	3345.469677		1463.473414	3651.564
		C.I				
		SE%	37.41215825		97.87561051	34.98526
			Steps overall variance		Steps overall variance	
			B/w village	within villa	B/w village	within village
			0.933333333	0.0645435	0.925	0.072611
			0.011666667	0.0008068	0.015416667	0.00121
			11139522.9	52644.459	2068056.689	73697.74
			11192167.36	Overall var	2141754.434	Overall variance

sampling design- Stratified two stage cluster sampling (Adaptive cluster sampling method)						
Stratification by Village/Monastery						
Number of Primary units in population (N) =			140		60	200
Number of Primary units sampled (n) =			10		4	14
Number of secondary units in i-th PU (M) =			314		314	314
Number of secondary units sampled in i-th PU (m) =			10		10	10
Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village	
Village	DH	0	0	0	0	
Village	GR	47	0.247368421	77.673684	0.059242373	
Village	LA	0	0	0	0	
Village	SP	12	0.4	125.6	0.068846426	
Village	VL	11	0.22	69.08	0.046858599	
Village	AN	0	0	0	0	
Village	DL	0	0	0	0	
Village	KL	0	0	0	0	
Village	SK	25	0.357142857	112.14286	0.123488886	
Village	VT	0	0	0	0	
Monastry	AG	117	0.760599415	238.82822	0.15195669	
Monastry	GH	0	0	0	0	
Monastry	MA	46	0.270588235	84.964706	0.07088621	
Monastry	PN	0	0	0	0	
Total trees observed		258	1.03118765	323.79292	0.055710725	
			Village		**Monastery**	**Pooled**
	Plot level	Estimated mean trees/plot	0.122451128		0.257796913	0.163055
		Variance/plot	0.002571108		0.030103753	0.003969
		Se/plot	0.050706091		0.173504331	0.063001
		C.I				
	Village level	Estimated mean total trees/village	38.44965414		80.94823056	51.19923
		Variance/village	253.5009325		2968.109632	391.3453
		SE/village	15.92171261		54.48036005	19.78245
		C.I				
		Variance within village (S2p)	0.029843628		0.055710725	0.037604
		SE within village	0.172753085		0.236031195	0.193917
		Variance between village (S2b)	0.274665966		1.286308236	0.578159
		SE between village	0.524085838		1.134155296	0.760367
		Ratio (S2b/S2p)	9.203504406		23.08905939	15.37502
	Population level	Estimated mean total trees in population	5382.951579		4856.893834	10239.85
		Overall variance	4968618.278		10685194.67	15653813
		Overall SE	2229.039766		3268.821603	3956.49
		C.I				
		SE%	41.40924794		67.30271888	38.63818
		Steps overall variance		**Steps overall variance**		
		B/w village	**within village**	**B/w village**	**within village**	
		0.928571429	**0.069153776**	**0.9333333**	**0.064543524**	
		0.009285714	**0.000691538**	**0.0233333**	**0.001613588**	
		4928735.73	**39882.54753**	**10653287**	**31907.53774**	
		4968618.278	**Overall variance**	**10685195**	**Overall variance**	

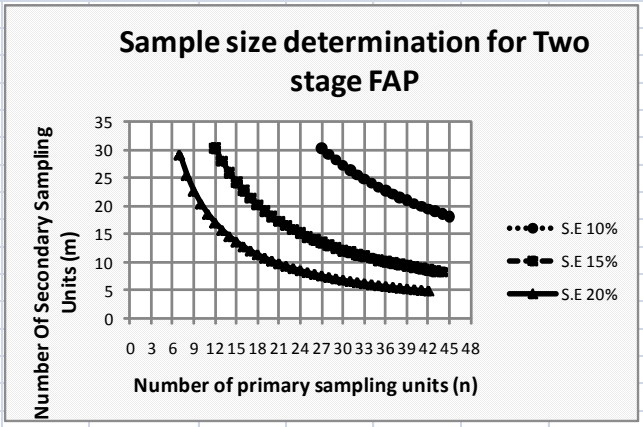
Appendix 3 Calculation sheet for Design optimisation

Appendix 3a Design optimisation without cost consideration

Design Optimisation for two stage fixed area plot (without cost consideration)

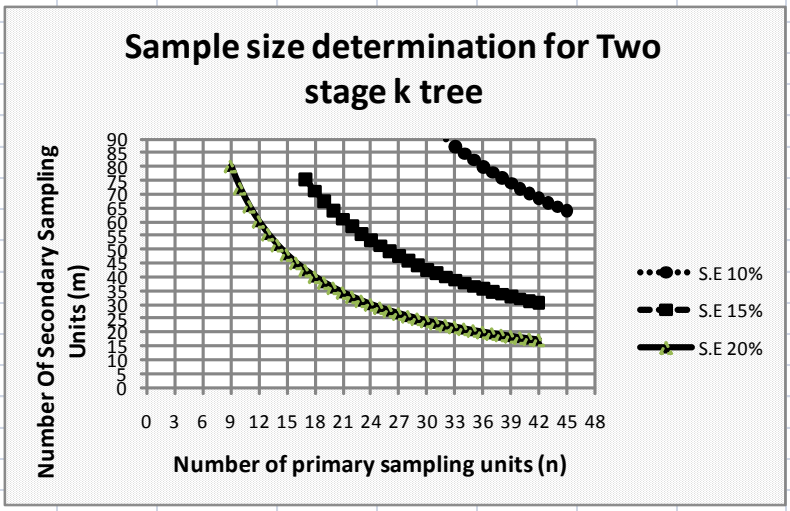
standard error 10%		standard error 15%		standard error 20%		standard error 30%			
n	m	n	m	n	m	n	m		
		52	7	42	5				
		51	7	41	5				
		50	7	40	5				
		49	7	39	5				
		48	8	38	5				
		47	8	37	6				
		46	8	36	6				
		45	8	35	6				
		44	8	34	6				
		43	8	33	6				
		42	9	32	6				
		41	9	31	7				
		40	9	30	7				
		39	9	29	7				
		38	10	28	7				
		37	10	27	8				
		36	10	26	8				
45	18	35	10	25	8				
44	19	34	11	24	9				
43	19	33	11	23	9				
42	19	32	11	22	9				
41	20	31	12	21	10				
40	20	30	12	20	10	14	6		
39	21	29	13	19	11	13	7		
38	22	28	13	18	11	12	8		
37	22	27	13	17	12	11	8		
36	23	26	14	16	13	10	9		
35	23	25	15	15	14	9	10		
34	24	24	15	14	15	8	11		
33	25	23	16	13	16	7	13		
32	26	22	17	12	17	6	15		
31	26	21	17	11	19	5	18		
30	27	20	18	10	20	4	23		
29	28	19	19	9	23	3	30		
28	29	18	20	8	26	2	45		
27	30	17	21	7	29	1	91	% Squared Mean	SE
		16	23					0.00002304	1.883304
		15	24						
		14	26						
		13	28						
		12	30						

560



Design Optimisation for two stage k-tree method (without cost consideration)

standard error 10%		standard error 15%		standard error 20%		standard error 30%			
n	m	n	m	n	m	n	m		
				42	17				
				41	18				
				40	18				
				39	18				
				38	19				
				37	19				
				36	20				
				35	21				
				34	21				
				33	22				
		42	30	32	23	26	12		
		41	31	31	23	25	13		
		40	32	30	24	24	13		
		39	33	29	25	23	14		
		38	34	28	26	22	15		
		37	35	27	27	21	15		
		36	36	26	28	20	16		
45	64	35	37	25	29	19	17		
44	65	34	38	24	30	18	18		
43	67	33	39	23	31	17	19		
42	69	32	40	22	33	16	20		
41	70	31	41	21	34	15	21		
40	72	30	43	20	36	14	23		
39	74	29	44	19	38	13	25		
38	76	28	46	18	40	12	27		
37	78	27	47	17	42	11	29		
36	80	26	49	16	45	10	32		
35	82	25	51	15	48	9	36		
34	85	24	53	14	51	8	40		
33	87	23	56	13	55	7	46		
32	90	22	58	12	60	6	53		
31	93	21	61	11	65	5	64		
30	96	20	64	10	72	4	80		
29	99	19	67	9	80	3	107		
28	103	18	71	8	90	2	160		
27	107	17	75	7	103	1	320	% squared mean	SE
								0.00002209	6.365944
								0.000016	1.936537



Appendix 3b Design optimisation with cost consideration

sampling design- Two stage cluster sampling (Fixed area plot method)

Calculation of Mean, variance and standard deviation

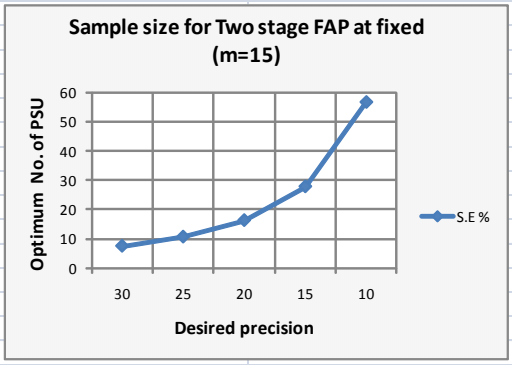
Number of Primary units in population (N) =	200
Number of Primary units sampled (n) =	14
Number of secondary units in i-th PU (M) =	314
Number of secondary units sampled in i-th PU (m) =	10

Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village
Grevena	AG	11	1.100	345.4	2.1
Grevena	DH	1	0.100	31.4	0.1
Grevena	GH	6	0.600	188.4	1.155555556
Grevena	GR	5	0.500	157	1.611111
Grevena	LA	4	0.400	125.6	0.71
Grevena	MA	9	0.900	282.6	2.62
Grevena	SP	4	0.400	125.6	0.933333
Grevena	VL	7	0.700	219.8	1.788889
Zagori	AN	0	0.000	0	0
Zagori	DL	0	0.000	0	0
Zagori	KL	1	0.100	31.4	0.1
Zagori	PN	2	0.200	62.8	0.177777778
Zagori	SK	15	1.500	471	2.722222
Zagori	VT	3	0.300	94.2	0.455556
		68	**0.486**	**152.51429**	**1.033888881**

	Plot level	Estimated mean trees/plot	0.485714286		
		Variance/plot	0.013465002		
		SE/plot	0.116038797		
		C.I			
	Village level	Estimated mean total trees/village	152.5142857		
		Variance/village	1327.59538		
		SE/village	36.43618229		
		C.I			
		Variance within village (S2p)	1.033888881	1.0338889	population variance
		SE within village	1.016803266		
		Variance between village (S2b)	1.951648352	0.0917759	population variance
		SE between village	1.397014084		
		Ratio (S2b/S2p)	1.887677088	Steps overall variance	
	Population level	Estimated mean total trees in population	30502.85714	B/w village	within village
		Overall variance	53103815.2	0.93	0.067770701
		Overall SE	7287.236459	0.0066429	0.000484076
		C.I		51129997	1973817.94
		SE%	23.89034058	53103815	Overall variance

Optimization of number of villages (n) and plots (m) within each village

Ratio of cost of locating primary(Cp) and cost of locating and observing each secondary(Cs)					
		Cp/Cs	20		
		Optimum number of plots to be sampled in each village (m_{opt})	15.01023504	15	
		Optimum number of villages to be sampled at 10% S.E (n_{opt})	56.69443075		
		Optimum number of villages to be sampled at 20% S.E (n_{opt})	23.3736519		
		15%	39.43272384		
SE%	30	25	20	15	10
n	7.402888433	10.5584316	16.21273344	27.786237	56.69443075



sampling design- Two stage cluster sampling (k tree method)

Number of Primary units in population (N) =	200
Number of Primary units sampled (n) =	14
Number of secondary units in i-th PU (M) =	314
Number of secondary units sampled in i-th PU (m) =	10

Block	Site code	Trees sampled/village	(Trees/village/plot) Y2sri	T2sri	Variance within village
Grevena	AG	28	1.630696411	512.03867	0.372916
Grevena	DH	2	0.099196489	31.147698	0.098399
Grevena	GH	27	0.302834795	95.090126	0.007414
Grevena	GR	22	0.14058707	44.14434	0.002876
Grevena	LA	3	0.00862293	2.7075999	0.000744
Grevena	MA	25	0.525748649	165.08508	0.024431
Grevena	SP	15	0.277239559	87.053222	0.038294
Grevena	VL	14	0.277297635	87.071457	0.031894
Zagori	AN	0	0	0	0
Zagori	DL	12	0.091552733	28.747558	0.00568
Zagori	KL	10	0.149346885	46.894922	0.005935
Zagori	PN	9	0.087347257	27.427039	0.014987828
Zagori	SK	30	2.854997028	896.46907	0.405812
Zagori	VT	18	0.193432854	60.737916	0.009148
Total trees observed		215	0.474207164	2084.6147	0.072752202

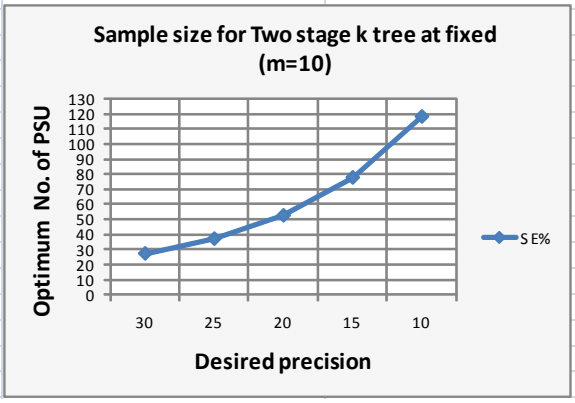
	Plot level	Estimated mean trees/plot	0.474207164		
		Variance/plot	0.042392386		
		S.D/plot	0.205894113		
		C.I	0.5±0.434		
	Village level	Estimated mean total trees/village	148.9010495		
		Variance/village	4179.719673		
		S.D/village	64.65075153		
		C.I	162±136.42		
		Variance within village (S2p)	0.072752202	0.0727522	
		S.D within village	0.269726161		
		Variance between village (S2b)	6.376347901	0.6303596	
		S.D between village	2.525143145		
		Ratio (S2b/S2p)	87.64474095		
	Population level	Estimated mean total trees in population	29780.2099	Steps overall variance	
		Overall variance	167188786.9	B/w village	within village
		Overall S.D	12930.15031	0.93	0.067770701
		C.I	32439±27284	0.0066429	0.000484076
		C.V%	43.41860029	167049894	138892.6839
				167188787	Overall variance

Optimization of number of villages (n) and plots (m) within each village

Ratio of cost of locating primary(Cp) and cost of locating and observing each secondary(Cs)

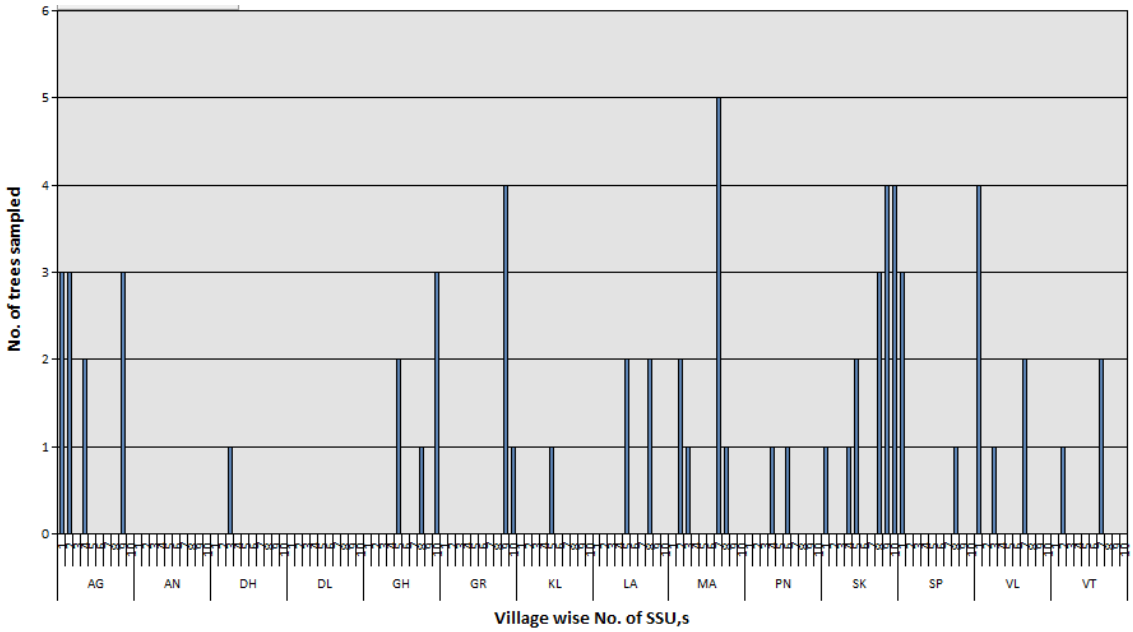
Cp/Cs	20	
Optimum number of plots to be sampled in each village (m_{opt})	1.519301246	10
Optimum number of villages to be sampled at 10% S.E (n_{opt})	123.4311526	
Optimum number of villages to be sampled at 20% S.E (n_{opt})	118.0437773	

SE%	30	25	20	15	10
n	27.25928051	37.05566772	52.48950191	77.64117	118.0437773

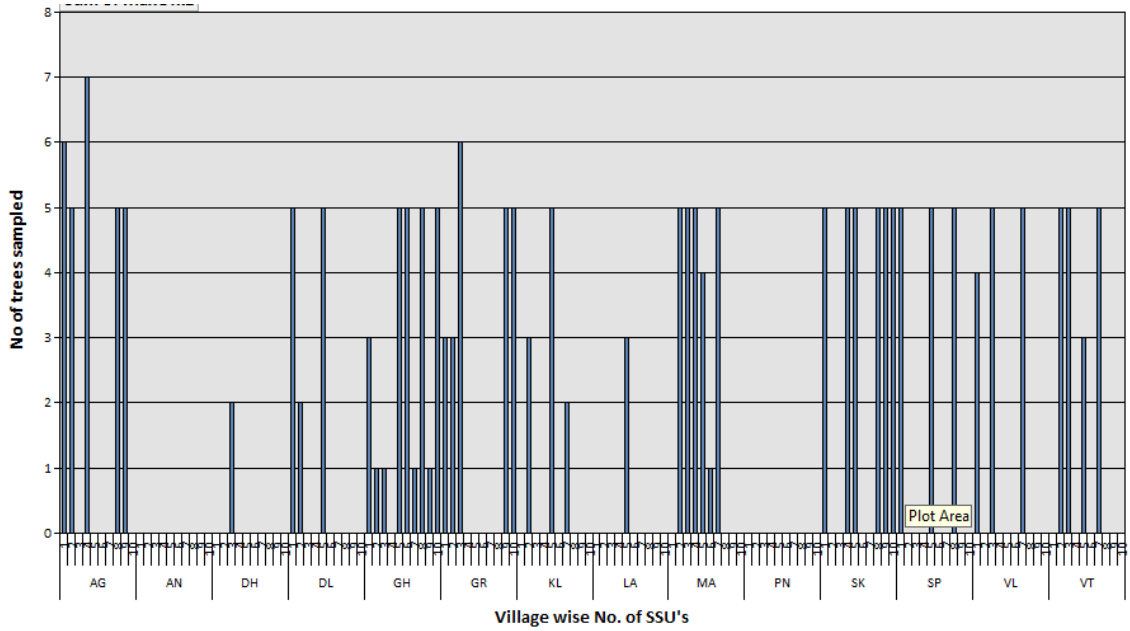


Appendix 4 Trees sampled per SSU

Number of trees sampled per SSU in fixed-area plot sampling



Number of trees sampled per SSU in k-tree sampling



Effective area of each of the k-tree plots (SSU)