DISSERTATION

SAMPLING STRATEGIES

FOR FOREST AERIAL DETECTION SURVEY IN COLORADO

Submitted by

Anh Quang Ha

Department of Forest and Rangeland Stewardship

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2015

Doctoral Committee:

Advisor: Robin M. Reich

William R. Jacobi
John E. Lundquist
Rajiv Khosla

ABSTRACT

SAMPLING STRATEGIES

FOR FOREST AERIAL DETECTION SURVEY IN COLORADO

Aerial detection survey (ADS) has been commonly employed in forest surveys in the United States for detecting forest damage and monitoring forest health. In Colorado, ADS by USDA Forest Service has conducted annual 100% census of government forested land for more than 20 years with the goal of achieving information about forest damage due to different causal agents and disorders. Sketchmapping has been commonly employed in ADS with the goal of detecting and documenting on maps mortality, defoliation and other visible forest change from aircraft. At medium and large scale, sketchmapping is a suitable technique for forest monitoring that provides valuable information in forest health. This dissertation deals with data of forest area damaged by five causal agents mountain pine beetle, spruce beetle, western spruce budworm, pin engraver, and Douglas fir beetle and two disorders subalpine fir mortality and sudden aspen decline. The combined areas damaged by all causes were also considered. Data were downloaded from ADS in Colorado from 1994 to 2013 as polygon shapefiles with associated information such as causal agents or disorders, area damaged, and type of forest. The goal of my dissertation was to identify an appropriate sampling strategies to archive good estimates of total area damaged, to decrease survey cost, and to increase safety by reducing the amount of flights. To approach this goal, four sample designs for estimating total area damaged caused by various causal agent were evaluated: simple random sampling, stratified random sampling, probability proportional to size, and non-alignment systematic sampling. A GIS layer of 150 transects

covering Colorado's forestlands was developed and represented the sample unit for my study. Each transect was 3.2 km wide and 625 km long and was numbered from 1 to 150 from south to north. Each sample design was evaluated using eight sample sizes (10, 15, 20, 25,30, 35, 50, and 70) and applied to the seven damages and the combined damaged area. The statistical properties were evaluated to determine the optimal sample design for estimating area damaged caused by different causal agents. The spatio-temporal characteristics of area damaged that influence precision and accuracy of estimate were considered. Most of the damaged forest areas by single causal agents and disorders showed aggregated spatial patterns; whereas the combined damaged areas were uniformly distributed across the landscape. A loss plus cost function was employed to determine the optimal sample size for each sample design and analyzed for the cost advantage of alternative sample designs. We found that stratified random sampling was the most optimal sample design by producing the highest percentage of unbiased estimates of total area damaged and the smallest variances. The next best sampling designs were simple random sampling and probability proportional to size. The non-alignment systematic sampling was the worst for estimating total area damaged both for individual causal agents and disorders and all causal agents combined. The optimal sample size varied by sample design and causal agents and disorders as well as the level of confidence. Optimal sample size increased with increasing variability in the population and as the desired level of confidence increased. Larger samples were required to simultaneously provide estimates for multiple causal agents and disorder with reasonable levels of precision when compared to a single causal agent. Stratified random sampling was the most cost effective when compared with other sample designs. For example, the cost advantage of stratified sampling over random sampling for estimating the damage from subalpine-fir mortality was $85,000 per year. In contrast, PPS sampling had a cost disadvantage

of -$13,000 per year when compared with simple random sampling and -$95,000 per year when compared with stratified sampling for estimating the total damage from all causal agents combined at the 0.95 level of confidence.

ACKNOWLEDGEMENTS

First of all, I would like to thank Vietnam Forestry University (VFU) for supporting me with a Vietnamese Government Scholarship. Thanks the Ministry of Education and Training of Vietnam for funding my Ph.D program. I also thank the Department of Forest and Rangeland Stewardship for giving me the opportunity to pursue the Ph.D degree at Colorado State University.

Thank you, all my teachers and friends here in Colorado State University who have been helping me, supporting me, and assisting me during the time staying here, such Sonya Lefebre, Christy Eylar, Dr. Yu Wei, Dung Nguyen – my first and last roommate during staying here. Your kindness will never be forgotten. Thank you, my committee members: Prof. Dr. William Jacobi, Dr. John Lundquist, and Dr. Rajiv Khosla for your kind lectures, for your valuable comments and suggestions. They have improved my knowledge as well as my life.

Special thanks goes to Prof. Dr. Robin Reich, my advisor. He is a wonderful teacher and advisor. He taught me a lot from the first day of my arrival at CSU, and gave me time to talk almost every day and therefore improving my English efficiency. He provided me basic knowledge about R and helped me apply R in my dissertation. I can't ever say thank you enough. Especially when he suddenly became ill, he still tried to work with my dissertation so I could graduate on time. Thank you very much and I highly appreciate your help, Robin.

To my family, I would like to thank to my wife, Phuong Luong, for her love, encouraging, and supporting me. She gave me two lovely kids, Minh Quang Ha (6 years old) and Bao Quang Ha (21 months old), who are my inspiration and motivation to overcome the program. You are always behind my any successes. I love all of you so much!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

AERIAL SURVEY APPLY IN NATURAL RESOURCES MANAGEMENT

Aerial survey has been well known as one common method of collecting data from using aerial equipment such aircraft, balloon, helicopter, unmanned aerial vehicles (UAVs), etc. It is helpful as an overview detection survey method that can provide information on many things not visible from the ground. Aerial survey has been widely applied in the mine industry, archaeology, land survey, forest health and wildlife monitoring, and many other fields.

Aerial survey in wildlife management

Aerial survey has been used in wildlife management for about 60 years (Caughley and Sinclair 1994). The most advantageous and impressive gain from aerial survey is providing an overview from a landscape perspective with 100% census. (Caughley 1977), Pollock and Kendall (1987), Walter and Hone (2003) suggested that aerial surveys are the only practical way for estimating wildlife animal population sizes and monitoring population growth rates over large areas (Pollock and Kendall 1987, Walter and Hone 2003). Sampling design plays an important role in aerial survey. Four types of sampling units widely used in aerial survey for wildlife animals are the line transect, line intercept, quadrat or block, and strip (Jolly 1981).

The *line transect* is widely used for detecting and estimating population size (Esseen et al. 2006), which is a plotless technique for sampling wildlife populations and basically described as transect lines randomly selected from main observer(s) travel line (Anderson et al. 1993). Group size of animals and their distance from a transect line are recorded. This technique was developed for surveying spread distributed populations (Buckland et al. 2001). A “good” line

transect survey, according to Thomas et al. (2007) is one that: randomly lays out transects, stratified if density is known to vary on a large scale, has an equal probability of being surveyed within a stratum in each location, produces an even distribution of transects throughout each stratum, and gives maximum efficiency per unit.

The *line intercept* is a standard sampling unit for estimating areas. This technique uses straight line, whose length intercepting selected categories of land use is recorded. Line intercept is suitable for assessing changes in plant species cover for most forest and rangeland communities (Caratti 2006), especially communities with shrub cover greater than 1m. Cover is recorded as the number of meter intercepted by each species along a transect, then percentage of cover is calculated by dividing the number of meter intercepted by each species by the total length of the transect (Caratti 2006). In aerial survey, during flight a fixed point is used as an imaginary line from which the length can be measured from distinguishable point on a map (Jolly 1981).

The *strip sample* is defined by its width and length. This sample unit is commonly employed for estimating size of bird and large animal population (University of Toronto 2002). Animals within the strip are recorded along with GIS and/or GPS. With aerial survey, strip width calibration is essential and strongly influences the sample errors. Many research papers report using strip sample surveys, however, mention is seldom made about how to determine the width of the strip (Jolly 1981).

The *quadrat or block* is a unit of probability sampling. According to Norton-Griffiths (1978), the distinction between quadrat and block is its shape, where a quadrat refers to a square or rectangle area and block refers to an irregularly shaped area. In most situations, the area of each quadrat must be known at least approximately (Jolly 1981). Blocks are also known as a

cluster in which the population being sampled is divided into groups (clusters) with as much variability as possible within each group (Scheaffer et al. 2006). In aerial survey, a block is commonly defined as a larger area than commonly used for ground survey for convenience conducting survey flight pathways. From cost effective point of view, cluster (block) sampling is more appropriate than simple random sampling or stratified random sampling in case a list of elements is not available or large dispersion of elements (Scheaffer et al. 2006).

Many factors such cost of survey, navigation, boundary effects, and sample errors influence the selection of an appropriate sample unit. Based on these factors, line transect sampling has usually been considered as the best choice, and quadrat sampling less favorable, but still preferable (Caughley 1977). In this regard, Chase (2007) used a Cessna single engine plane to estimate population size and distribution of large animals along Caprivi River Systems in Namibia. To provide 100% coverage a transect with a width of 500m was used. Flight transects were systematically flown along generally east/west axes at a speed of 100 knots. Wildlife only within the transect was counted and recorded using GPS and existing digital maps. Results showed that wildlife numbers were highest bordering conservation areas: in wetlands hippos occurred in the largest numbers (1,269); in woodlands buffalos occurred in the highest numbers (5,951); numbers of wildlife recorded from aerial survey in September 2007 was 79% higher than that in August 2004 (Chase 2007). Similarly, a census was published by Tanzania Wildlife Research Institute in 2013 reported using aerial survey to estimate population of large animals in Tanzania (Tanzania Wildlife Research Institute 2013). The line transect sample technique was employed by covering all critical interested areas with 5 and 10 km spacing between transects depending on the topography. Animals were counted within a strip width of 150m on either side of the aircraft (300m wide per transect) at a speed of 180 kilometer per hour.

In British Columbia, Zimmerman et al. (2002) used stratified random block sample design to estimate the population size of caribou and moose within four caribou herd areas and determine population compositions of the two species within these areas (Zimmerman et al. 2002). Previously (2000), transect sampling did not work well to count the total number of animals because animals were hiding in forested areas. With stratified random block sample design, each study area was divided into square grids of 25km^2 sample units using a weighted stratified random block sample strategy. The weights varied by species. For caribou (*Rangifer tarandus*) census, weights were based on percentage composition of Pine Lichen Winter Range (primarily low-elevation lodgepole pine (*Pinus* contorta) forests) and High Elevation Winter Range (primarily alpine and subalpine areas). For moose (*Alces alces*) census, weights were based on the percentage composition of Moose Winter Range preference classes (Zimmerman et al. 2002). Flight speed varied from 60 to 80 kilometer per hour and height-above-ground ranged from 50 to 200m. All animals observed were recorded with GPS and GIS to determine whether they were inside or outside the boundaries. The results suggested that the estimated precision from aerial survey was still influenced by sighting caribou when they are in forested habitats. This issue of hidden animals could be overcome by using marked animals survey method combined with the stratified random block sample design. The variance of estimate population size would be improved with larger sample sizes. Stratified random block sample design had an advantage to count caribou, which cluster together in groups. Another study in Alaska conducted by Evans et al. (1966) used aerial quadrat sampling method to estimate moose population size. The areas of interest were stratified into units of high, medium and low moose density based on current observations. Sampling units of 1 square mile in area were surveyed from the air. Results suggested that sampling protocol was useful to estimate moose population with reasonable cost

(Evans et al. 1966), especially in areas where access is limited. However, the total cost of the survey was not provided.

Helicopter aerial surveys with strip-count, mark-recapture and line transect techniques were conducted in the Australia Alps to estimate abundance and population density of wild horses (*Equus caballus*) (Walter and Hone 2003). Details of the mark-recapture aspect of the aerial survey were not clearly described. The horse population estimates using the strip-counts method were strongly dependent on strip width. Strips over 200m in width gave estimates that had a significant negative bias. The precision of strip estimate for individual animals was better than that for groups with lower coefficient of variation (CV). Estimated population from aerial survey using mark-recapture estimate technique gave 3 to 5 times greater precision than the 0-200m width strip estimates. By comparing these techniques, the results suggested that with surveys using strip, mark-recapture analyses or line transect, the effect of wide strip width should be examined for each specific object because it directly affects accuracy of the estimates. Aerial surveys using the line-transect system is what was recommended to monitor trends in wild horse abundance in the Australia Alps (Walter and Hone 2003).

In Selous-Niassa (Tanzania), a large wildlife census was conducted in the dry season of 2009 using aerial survey (Tanzania Wildlife Research Institute 2009). This survey applied systematic reconnaissance flights with 2.5 km width of transects oriented East to West. Global Positioning System (GPS) was used to record the starting and end points of each transect on the accompanied flight map. Each transect was divided into 30 sub-units for geo-referencing purposes within the survey area. Two observers on the plane, the front seat observer recorded the radar altimeter reading at the start of each sub-unit and GPS geo-referenced information on vegetation cover, fire and surface water presence for each sub-unit. The rear seat observer

recorded all wildlife observed within counting strips defined by streamers fitted to the wing struts on either side of the aircraft. Observation data were transcribed onto data sheets after each flight (Tanzania Wildlife Research Institute 2009). The results showed that diversity of wildlife species observed increased from 6 (in 2006) to 15 (in 2009). Elephants were the most numerous species (4,577± 1,126SE) followed by buffaloes (4,095±2.116SE) and others species such as sable antelope, roan antelope, hippo, zebra, baboon, etc. Report of this census also suggested that the survey should be conducted in the wet season as well to minimize missing of some species that are present only in the dry season.

Monitoring seabird population in marine systems with aerial strip-transects is an easily designed and low cost survey method, especially in a large scale survey (Certain and Bretagnolle 2008). A study on the bias of aerial strip-transects was conducted in the Bay of Biscay using 100.000km^2 of the French Atlantic coast, in the Northeast Atlantic Ocean (Certain and Bretagnolle 2008). Data from a survey of the entire seabird community between October 2001 and March 2002 were used to determine and compare whether the aerial strip-transect survey method produced temporal (i.e. variations of detection probability during the survey) or distance-based (i.e. variations of detection probability across the strip) bias. Results indicated that only temporal bias was attributable to sun glare while no distance bias was detected in a strip of 150 m width and the effect of distance up to 230m was very weak (Certain and Bretagnolle 2008). This suggested that there was no need to add distance bias analysis to survey protocol using line-transect method for monitoring seabird population. In another study, Ridgway (2010) employed line transect distance sampling to estimate density of double-crested cormorants (*Phalacrocorax auritus*) along the coastal areas of Georgian Bay and the North Channel on Lake Huron. This research found that even though the probability of detection of cormorants varied based on

different factors (e.g. season, group size, and location) and was often below 0.5 cormorants per km^2, distance sampling was useful for monitoring distribution and abundance the birds. Data from aerial survey in 2004 lead to estimate range from a peak of 2.3 cormorants per km^2 (95% CI = 1.72-3.03) in late July to 1.2 cormorants per km^2 (95% CI = 0.78-1.70) in late August (Ridgway 2010) .

In fishery science, aerial surveys have been widely applied to survey large marine species such whales, dolphins and manatees (Rowat et al. 2009, Alves et al. 2013). Alves et al. (2013) employed aerial survey to analyze the distribution of manatees, dolphins and sea turtles in the northeastern coast of Brazil. Using strip transects with a zigzag pattern covering 4026km, 36 sightings of manatees (41 individuals), 28 of dolphins (78 individuals) and 256 of sea turtles (286 individuals) were recorded. The study indicated a positive correlation between sea turtles and manatees. The researchers also found the number of manatees in marine protected area (MPAs) was significant higher than that in non-protected areas while dolphins and sea turtles were mostly observed in coral reefs areas (Alves et al. 2013).

In applying aerial survey for wildlife census, sample sizes have been considered in some studies. In his study on estimating elephant abundance, Walsh et al. (2001) found that increasing sample size was needed when aerial survey intensity increased (Walsh et al. 2001). With the monitoring of large wildlife mammals using aerial survey, sample size is important variable that directly affects the estimate of population size. Clumped distributions and low population densities may lead to outliers that reduce the precision of an estimate. In their study on elephant population in Africa, Ferreira and Aarde (2009) used an incremented sample size method that ranged from 5% to 100% coverage of the total study area by increasing the number of transects in a survey. The results showed that the accuracy and precision of estimates increased with

survey intensity and it depended on the population density and the distribution (clumped or random). To reduce the effects of clumped distribution on the precision of population estimate, survey intensities were suggested from 50 to 70% of a study area (Ferreira and Aarde 2009).

Aerial survey in forest health management

In the United States, aerial surveys have been used to monitor and report on the health of ecosystems by the nation Forest Health Monitoring (FHM) program since the late 1990s (Johnson and Ross 2008). As a part of remote sensing, aerial survey gives the connectivity between spatial patterns of landscape ecosystems and their ecological process by providing spatio-temporal information. Aerial photography and aerial sketch-mapping provide valuable information of forest health at medium to large-scale images (McConnell and Avila 2004).

McConnell et al. (2000) mentioned that flight path (pattern) is one of the most important factors that affect accuracy and cost of the survey. Flight patterns can be a contour (ridge contour, drainage contour) or parallel-line pattern (McConnell et al. 2000) (Figure 1.1). According to Klein et al. (1983) the parallel pattern is generally used in flat terrain such in the southeast United States or only in mountainous for cursory information surveys (McConnell et al. 2000).

Magnussen et al. (2012) recommended that aerial surveys provide valuable information on the scale and severity of defoliation and mortality caused by forest insects. The link of aerial survey data and severity of insect defoliation to tree ring series of radial growth was mentioned and employed to estimate losses of growth due to defoliation damage. In this research, the losses of radial growth from 1930 to 2005 caused by insect defoliation were estimated at 10% on average (Magnussen and Alfaro 2012). This approach was potentially useful for estimating the forest growth effects from their symptoms of damage by defoliating insects or diseases. Aerial surveys can be used as the first step of a multi-tiered process of detection, monitoring, and

evaluation, utilizing other remote sensing and ground sampling techniques to gather additional data on significant forest events or change (Heller and Wear 1969, Wulder et al. 2006, Johnson and Ross 2008). Measurement of error is a concern in all surveys (Lohr 2010). Minimizing errors that make biased estimates is a statistical objective and has been widely studied. The study on error of aerial survey in forest of Rocky Mountain in 2005 showed that when a spatial tolerance increased from 0m to 500m, the accuracy achieved was between 37% to 69%, respectively (Johnson and Ross 2008). This accuracy was better than what would be expected from the chance assignment of randomly drawn polygons to randomly selected pest categories. Aerial censuses have been faced with inaccurate problems caused by the observer missing a significant number of damage areas on the transects. Sight ability is affected by many factors that would influence bias such strip width, altitude, and speed of flying (Pollock and Kendall 1987, Marsh and D.F.Sinclair 1989). Inverse relationships occur between sight ability and speed of aircraft as well as width of the strip transect (Caughley 1974). Methods of eliminating errors from aerial survey by refining techniques have been discussed and rejected without technical support. An alternative strategy, known as linear regression method of correcting for bias, had been suggested for correcting the estimates based on measurement of bias. Seven steps of conducting aerial survey for collecting data of animal population have been suggested by Caughley (1974) following his study in Zambia, Kenya, Uganda to estimate wildlife animals (Caughley 1974): (1) divide the survey area into strata according to thickness of cover, (2) draw transects on a map and randomly assign to each treatments, (3) conduct survey and record information, (4) calculate the spatial regression of apparent density on strip width, speed, and altitude separately for each stratum, (5) estimate number of strata using the intercept constant, (6) sum the estimate of each stratum to get estimate of total, and (7) calculate a standard error for both each stratum and total.

The other popular technique that can use experts' skills to draw maps during flying is known as aerial sketch-mapping. It is also known as "real time photo interpretation" (Johnson and Ross 2008) based on its transfer time and technique. This technique has been commonly used in the United States, Zambia, Kenya and Uganda to estimate the extent and severity of forests damaged by insects and diseases (Caughley 1974, McConnell and Avila 2004). Aerial sketch-mapping is considered the most efficient and economical method of detecting and appraising recognizable pest damage over large remote forest areas (McConnell et al. 2000). Time of conducting survey is the main advantage of aerial sketch-mapping surveys because it provides a simple and quick method for detecting and recording forest damage on a map at large area in only a few days (McConnell and Avila 2004). By this method, the areal extent of damaged forests can be transferred to existing maps as polygons that are characterized by its size, shape, and location. These polygons might be coded with additional information such as type of forest, causal agent, and so on. From the scale of purposes of survey, aerial sketch-mapping would be divided into two types: the overview survey and the event-specific survey in which the overview (known as general) survey is the most common (McConnell et al. 2000). In forest inventory, the general survey is known as landscape-level survey in which large areas are covered. It is also a good method of surveying forest changes over time. The precision of overview survey varies depending on scale of map used for sketch-map. A larger scale map produces more precision survey outputs but increasing cost and time than smaller scale ones. In the United States, the standard scale for Forest Health Protection overview survey is 1:10,000 and 1:24,000 which is used by USGS (McConnell et al. 2000). The event-specific survey is used for specific interested events such a damage area by unique biological agent. This type of survey is commonly used following overview survey by employing larger scale maps (e.g. 1:24,000 scale) to assess abiotic damage such fires, hurricanes, floods, etc.

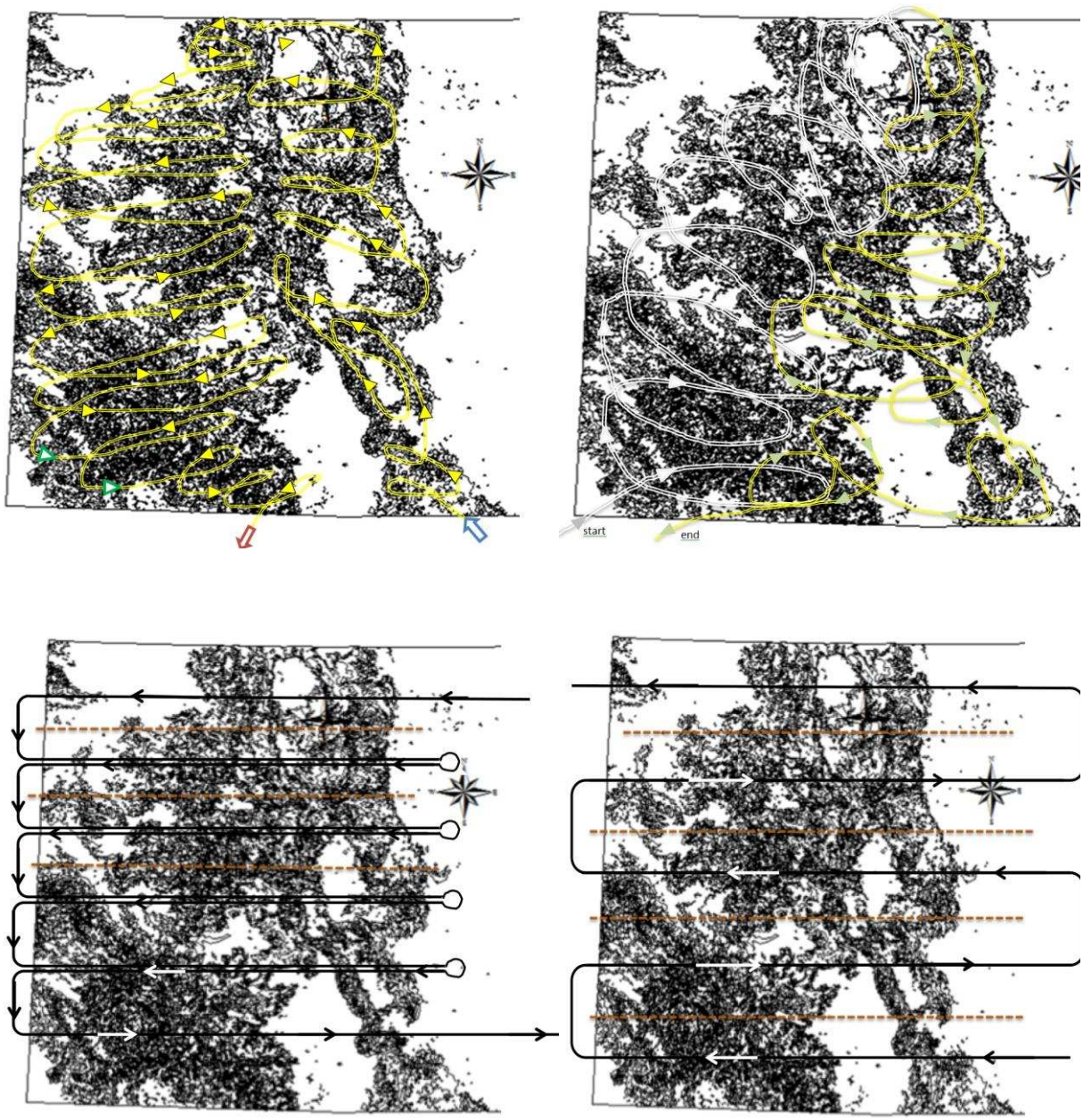


Figure 1.1. Ridge Contour flight pattern (top left) and Drainage Contour flight pattern (top right). Parallel flight path for single aerial (bottom left) observer and two aerial observers (bottom right). Dashed line indicates viewing distance for each observer (Modified from McConnell et al. (2000) and USDA Forest Service in applying for study site – forestland of Colorado). Flight path and number of observer in aerial overview survey are often corresponding and varies depending on terrain in which steep mountainous terrain only one observer, otherwise two commonly used.

Aerial sketch-map surveys have historically been undertaken annually in the western United States in order to locate areas of forest pest activity since the 1940s (Johnson and Ross 2008). The application of remotely sensed data (e.g. photographs, satellite images) for the detection and mapping of mountain pine beetle infestations has been ongoing since the early 1960s (Wulder et al. 2006). To identify insect and disease activity, the observer looks for characteristic signatures to distinguish the tree species and the type of damage that has occurred. Characteristics that observers use to determine the host tree species include the shape of the tree's crown, slope position, elevation and aspect. Variation in the color of the tree's foliage indicates the presence and type of insect or disease activity. These visual signatures can be implemented both manually (e.g. sketch-mapping) or computerized (e.g. analyzing spectrum of different reflection bands from remote sensing photographs). Woodall et al. (2009) compared aerial surveys with field inventories of forest health for oak forests in the northern United States and found that aerial damage surveys were weakly correlated with indicators of oak forest sustainability (e.g. oak seedling and saplings) but highly correlated with over-story attributes, such as tree mortality and standing dead. He also found that the highest correlations between the aerial damage surveys and oak mortality were found when the time between the aerial survey and subsequent forest inventory was 4-6 years later (Woodall et al. 2009).

The United States Department of Agriculture's Forest Service (USDA FS) has the Forest Health Monitoring Program (FHM) and is responsible for monitoring forest health of the US (Bechtold et al. 2007) in which aerial sketch-mapping technique has been used for monitoring of forest health (Steinman 2004). This technique relies on expert who can detect damage in the form of tree defoliation, mortality, and damage as associated with the occurrence of damaging insects, diseases, wind throw, and other biotic and abiotic forest disturbance (Conkling et al.

2005). The aerial observer is the key to a successful sketch map operation. Because it is created while flying with high speed in an aircraft and observing damage, sketch-mapping is an acquired and difficult skill (Johnson and Wittwer 2008). In the southeastern and northeastern United States, aerial surveys vary from 150m to 1525m above the sea level, while in the far West the flight altitudes vary from 1220m to 2745m or sometime higher. Flying at high speeds, often ranging from 115 kilometer per hour to 145 kilometer per hour, it is hard for observers to capture all area damaged. Johnson and Ross (2008) stated that the minimum damage area that can be drawn as a polygon by sketch-mapping technique is 0.4 ha while if very small infestation (e.g. less than 10 trees in a group) is designated as a dot (McConnell et al. 2000). Traditionally, sketch-maps of forest damage were drawn on paper-based maps. Recently, experts can use automated digital sketch-mapping system developed by the USDA Forest Service's Forest Health Technology Enterprise Team (FHTET) (Johnson and Wittwer 2008). Klein et al (1983) recommended that collecting data from aerial sketch-mapping should be regarded more as an art than an exact science (B.C. Ministry of Forests and Service 2000, Johnson and Wittwer 2008). Aerial sketch-mapping is still considered a low-cost remote sensing method that provides landscape level overview of forest conditions (e.g. USDA had spent $5million annual for doing forest aerial survey or roughly $0.25 per ha (USDA Forest Service 2004) or $3,000 per million acres (McConnell and Avila 2004). The accuracy of aerial sketch-mapping would be enhanced by combining with aerial photographs, especially in area of extensive pest damage (B.C. Ministry of Forests and Service 2000).

As previous mentioned, many remote sensing methods could be applied in forest inventories. Determining what alternative techniques should be used for forest inventories should consider its simplicity, cost and effectiveness (McConnell and Avila 2004) (Table 1.1).

Depending on specific survey purposes, each method has its advantages and disadvantages, a trade-off among these characteristics lead to a suitable method such a sketch-mapping, which has a bright future for conducting forest overview surveys with improving technology and combining good ground-checked information (McConnell 1999, McConnell et al. 2000). This technique has been an applicable method and widely used by the US Forest Service because it offers a simple, inexpensive and quick alternative to record forest pest infestation across a landscape.

Table 1.1 Cost, quality and scope features comparison of some remote sensing methods and their ability to detect, recognize, and identify forest pest damage (McConnell and Avila 2004).

Remote sensing method	Costs per million acres	Accuracy/Quality		Scope	Sensor capabilities to Detection, Recognition and Identification of forest pest
		Quality imagery pixel size (meter)	Damage area (acres)		
MODIS	$2,000	Low spatial 250/500	200	Large area analysis	Detect and recognize objects if they occur over very large area, no identification of trees
LANDSAT	$3,000	Low-med spatial 10/30	5	Large area analysis	Detect and recognize objects if they occur over very large area, no identification of trees
SPOT	$35,000	Medium spatial 2.5/10 or 20	2.5	Med-large area analysis	Detect and recognize objects if they occur over very large area, no identification of individual dead trees
Quick Bird	$292,000	High spatial 0.6/2.4	<1	Small-Med area analysis	Detect and recognize and identification individual dead trees, but low spectral resolution
Ikonos	$270,000	High spatial ¼	<1	Small-Med area analysis	Detect and recognize and identification individual dead trees (10 meters tree crown diameter)
Sketch-mapping	$3,000	Depend on map scale	Point	Small-Large area analysis	Detect and recognize and identification group or individual dead trees, but no spectral resolution

CHAPTER 2: FOREST MORTALITY AND DAMAGE BY INSECTS AND DISEASES IN COLORADO

Introduction

The 9,308,000 ha of forests in Colorado are dominated primarily by spruce-fir, ponderosa pine (*Pinus ponderosa* Laws.), lodgepole pine (*Pinus contorta* Dougl.), aspen (*Populus tremuloides* Michx) and pinion-juniper (Thompson et al. 2010). These forests provide wildlife habitat, improve water and air quality, mediate negative effects of climate change, and offer many other services and resources. Only about 18% (118,000,000 m^3) of annual growth on the stump is sold (USDA Forest Service 2011b). Most of the forests are used for conservation, water production, grazing, and recreation. In recent years, forests in Colorado have been subjected to a variety of agents that impact the health of the forests. These agents include human activity, fires, invasive exotic fungi and plant species, climate change, and insects and fungi that cause growth loss and mortality.

A number of damage causing insects and fungi occur in Colorado. Most insects and fungi are native and most of the time are found at endemic levels. However, five of the most important insects species that can develop into serious outbreaks include mountain pine beetle (*Dendroctonus ponderosae* Hopkins), spruce beetle (*Dendroctonus rufipennis* (Kirby)), piñon ips beetle (*Ips confusus* (Leconte)), western spruce budworm (*Choristoneura occidentalis* Freeman), and western balsam bark beetle (*Dryocoetes confusus* Swaine). Important forest diseases include dwarf mistletoe (*Arceuthobium spp.*) and armillaria root rot (*Armillaria sp*) (CSFS 2014a). A recent report on the condition of Colorado's state forest (2011) estimated that 11.33 billion m^3 of

growth loss per year was caused by dwarf mistletoe, of which 2.83 billion m^3 was associated with ponderosa pine (Wahl 2006). Many researchers report that these forest insects and diseases usually do not just individually impact trees but in many cases interact to cause damage (Hagle et al. 2003). For example, a tree will be weakened first by dwarf mistletoe, then various insects such as the mountain pine beetle can overcome the defenses of these weakened trees and kill them instead of attacking healthy trees. Similarly, about 62% to 75% of ponderosa pines on the Colorado's Front Range and the Black Hills of South Dakota were attacked by mountain pine beetles were also infected by root disease (USDA Forest Service 2011a). In this chapter, five agents and two disorders that damage forests are discussed, including the mountain pine beetle (*Dendroctonus ponderosae* Hopkins) (MPB), Douglas-fir beetle (*Dendroctonus pseudotsugae* Hopkins) (DFB), spruce beetle (*Dendroctonus rufipennis* Kirby) (SB), sudden aspen decline (*Populus tremuloides*) (SAD), subalpine fir (*Abies lasiocarpa var arizonica* (Merriam)) mortality (SUB), western spruce budworm (*Choristoneura occidentalis* Freeman) (WSB), and pine engraver (*Ips pini* (Say) (PE).

** Mountain pine beetle*

The mountain pine beetle is the most important forest insect in Colorado and can kill millions of trees during large outbreaks. Weather and host stand conditions are considered the two most dominant factors influencing the severity of MPB outbreaks (Safranyik et al. 2010). In northern Colorado and southern Wyoming, more than 607,000 ha of forests were affected by MPB during the drought period from the late 1990s to the early 2000s (USDA Forest Service 2015). Dense stands with large-diameter trees are more vulnerable to attack by the MPB (Romme et al. 1986, Gibson et al. 2009, Coops et al. 2012). In the Rocky Mountains, MPB infestation of lodgepole pine gradually increased from 1996 peaked in 2009 and collapsed by

2014. Lodgepole pine (*Pinus contorta*) forest mortality mainly occur in a cluster pattern (Chapman et al. 2012). Many authors report that an MPB infestation trend is hard to predict even though its host's spatial pattern is known (Safranyik et al. 2010, Chapman et al. 2012). The expansion of the beetle's population is affected by many unpredictable factors such as climate and weather.

** Engraver beetle*

Eleven species of Ips beetle (*Ips spp.*) also known as engraver beetle, occur in Colorado (Cranshaw et al. 2012). Of these *Ips pini* (Say) and *Ips emarginatus* are the most common. The pine engraver beetle's common hosts are ponderosa pine, lodgepole pine, and Jeffrey pines, while *Ips emarginatusis* is common in ponderosa pine, white, and Jeffrey pines. A few *Ips* beetle species attack spruce. *Ips confusus* is considered to be a single-host species because it mostly kills piñon trees and is rarely found on other pines (Cranshaw et al. 2012). There are many factors that influence pine engraver infestation such tree diameter, sunlight intensity, and precipitation (Hayes et al. 2008). In this regard, the pine engravers tend to attack trees with a large diameter. It was reported that even though infestation was found on trees with 5 cm dbh, fewer such trees were attacked when compared with trees with a diameter greater than 10 cm (Steed and Wagner 2004, DeGomez and Young 2015). Thatcher (1960) estimated that about 3.7 million m^3 of timber in the Southern United States was lost annually to a pine engravers infestation (Thatcher 1960) while Barker (1972) reported about 1.2 million m^3 was damaged annually in Florida alone (Baker 1972). *Ips* beetles are not considered as aggressive as other bark beetles (mountain pine beetle, spruce beetle, Douglas fir beetle) (Livingston 2010). The tops of large-diameter trees may be killed and entire trees are killed when beetle populations are high. On infested trees, red-orange boring dust is easily seen (Hagle et al. 2003). *Ips* beetles have

appeared commonly in Colorado's forests but reports on the location and amount of damage incurred are rare in the literature.

** Spruce beetle*

Spruce beetle can cause 1% to 99% mortality in stands of trees with a diameter larger than 25.4 cm (Schmid and Frye 1977). Thus, large diameter spruce trees have relatively high chance of being killed than smaller-diameter trees. (Baker and Kemperman 1974, Schmid and Frye 1977, Hart et al. 2015). The stand structure of spruce-fir forests can be significantly modified by spruce beetle by reducing the average tree diameter, height, and stand density (Hagle et al. 2003). Damage caused by the spruce beetle was noted during the mid-1890s, when 10% to 25% of the mature spruce forest on the White River National Forest, and 25% to 40% of the spruce on the Grand Mesa National Park were killed (Schmid and Frye 1977, Hart et al. 2015). Later, in the 1940s and 1950s, these regions again had extensive outbreaks that killed about 50% of the spruce trees, equivalent to 23 million m^3 (Massey and Wygant 1954). The estimate of annual mortality attributed to the spruce beetles through the 1960s was approximately 2 to 3 million m^3 (Schmid and Frye 1977, Hart et al. 2015). The most recent report of areas in Colorado infested by spruce beetles was 196,200 ha in 2014, an increase of 35,200 ha over that observed in 2013 (CSFS 2014a).

** Douglas fir beetle*

Douglas fir beetle kills Douglas fir trees, while inoculating affected trees with pathogenic blue stain fungi (Hagle et al. 2003, Furniss 2014b) which is similar to many other bark beetles. Trees that are injured by fire scorch, defoliation, wind, or root disease are more susceptible to attack (Withrow et al. 2013). Stand condition and weather, stand density, tree age, and precipitation are strong factors that influence Douglas-fir beetle activity (McMillin and Allen

2000). This beetle normally kills small groups of trees (Schmitz and Gibson 1996), and trees with diameter greater than 17.1 cm are relatively more susceptible to attack, while trees with < 13.5 cm dbh are resistant or otherwise survived outbreaks (Furniss 2014a). Devastating outbreaks occurred in western Oregon and Washington from 1950 through 1969. One outbreak, for example was 2 to 4 years long and 17,462,000 m^3 of timber were lost. Other notable outbreaks killed 1,888,000 m^3 of timber in California (1966) and 257,211 m^3 of Douglas fir in Idaho (1970-1973) (Schmitz and Gibson 1996). Significant mortality by Douglas fir beetle infestation were reported by Forest Health Management (1999) in Wyoming, where the basal area was reduced by 40% to 70%, tree diameter decreased by 8% to 40%, and the Douglas fir component of the overstory decreased by more than 15%. Additionally, conifer seedling regeneration increased about four-fold; 90% was Douglas-fir and there was a three-fold increase of understory vegetation, which suggest that even though the Douglas fir beetle causes a significant short-term impact, the long-term successional pattern have probably not changed (McMillin and Allen 2000). In Colorado, about 13,760 ha of tree mortality due to this beetle was reported in 2014 compared with 17,400 ha in 2013 (CSFS 2014a).

** Subalpine fir mortality*

Subalpine fir mortality is sometimes known as subalpine fir decline. Some research reports describe this conditions primarily caused by a combination of armillaria root disease and western balsam bark beetle (Williams et al. 1986). Other conditions/or disturbances associated with fir mortality include stem decay, drought, balsam woolly aphid, and many wood-rotting fungi that cause heart, trunk, butt, or root rots such as brown stringy rot, red heart rot, red ring rot, shoestring rot, brown cubical rot, white spongy root rot, and white pocket rot (Uchytil 1991). Subalpine fir is susceptible to decay and the amount of decay increases with a tree's age (Bier et

al. 1948). Subalpine fir decay was present in trees less than 80 years old but decay did not become extensive, while trees greater than 120 years old had a higher probability of having significant amount of decay (Hinds et al. 1960). *Stereum sanguinolentum* (Alb. & Schwein.) and *Echinodontium tinctorium* (Ellis & Everh.) were considered to be responsible for 87% of the total decay, of which *Stereum sanguinolentum* (Alb. & Schwein.) was the primary decay fungus and alone accounted for 47% of the total (Hinds et al. 1960). Other causal agents that can cause decay include heart-rot fungus (*Echinodontium tinctorium* (Ellis & Everh.)), *Coniophora carebella* (Pers.), *Polyporus tomentosus* Fr., and others. In the Rocky Mountains, Bigler et al. (2007) reported on the effects of drought on subalpine forest mortality concluding that large interspecific differences in drought-related mortality occurred with fir having the strongest effect, followed by spruce and pine. The effect of changing climate on tree mortality has the potential to bring about large-scale changes in the subalpine forests of the Rocky Mountains (Bigler et al. 2007).

** Sudden aspen decline*

Sudden aspen decline (SAD) refers to a complex disease that occur at landscape scales in aspen stands, which are predisposed by drought and damaged or killed by *Cytospora* canker, wood borers and other insects (Worrall et al. 2008, Marchetti et al. 2011). In 2008, aerial survey observers located over 220,000 ha (17% of the aspen cover type) with 45% of that area classified as "severely damage." From 2000 to 2010, about 535,000 ha of aspen forest were impacted by SAD in the Southern Rocky Mountains eco-region, of which 492,000 ha occurred in Colorado (Worrall et al. 2014). Several reports which testing the correlation between the severity of SAD and stand condition as well as site index concluded that the severity of SAD was weakly related to basal area (Fairweather et al. 2007), stem slenderness, and site index (Worrall et al. 2010).

Age of the overstory was unrelated to severity of crown loss, while slope position and elevation were strongly related (Fairweather et al. 2007) through their indirect effects. In a survey of areas outside of the SAD dominated areas, Dudley et al. (2015) found that damage and mortality were highest at low elevations and deep slope, where temperature is high and precipitation is low (Dudley et al. 2015). Drought and climate change, especially hot weather, have been considered to be important factors that accelerate dieback of aspen trees (Worrall et al. 2008, Rehfeldt et al. 2009, USDA Forest Service 2011c).

** Western spruce budworm*

Western spruce budworm is the most widely distributed and destructive defoliator of coniferous forests in Western North America (Fellin and Dewey 1992). The most common host species of WSB are Douglas fir (*Pseudotsuga menziesi* (Mirb.) Franco), grand fir (*Abies grandis* (Doug. ex D. Don) Lindl.), white fir (*Abies concolor* (Cord. and Glend.) Lindl. ex Hildebr.), subalpine fir (*Abies lasiocarpa* (Hook.) Nutt.), corkbark fir (*Abies lasiocarpa var. arizonica* (Merriam) Lemm.), blue spruce (*Picea pungens* Engelm.), Engelmann spruce (*Picea engelmannii* Parry ex Engelm.), white spruce (*Picea glauca* (Moench) Voss), and western larch (*Larix occidentalis* Nutt.). The stand structure and age of Douglas fir are considered accelerating factors of outbreaks (Hadley and Veblen 1993). Records suggest that there are no typical patterns or trends for WSB epidemics. However, Williams and Liebhold (2000), suggest that spruce budworm outbreaks produce cluster patterns that expand along an east-west direction (Williams and Liebhold 2000). Ryerson et al. (2003) used tree-ring analysis to identify at least 14 outbreaks of western spruce budworm during the past 350 years in the San Juan Mountains of Colorado with intervals of time between outbreaks of 25, 37, and 83 years (Ryerson et al. 2003). In Colorado, an epidemic of WSB was recorded starting in 1949 and persisting for more than 30

years, damaging 2,430,000 ha in the Northern Rocky Mountains (Fellin and Dewey 1992). In this region, damage ranged from 18,210 ha (1993) to about 1,983,000 ha (1958) (Blackford 2010). In Colorado, WSB defoliated approximately 72,000 ha in 2014 compared with 63,000 ha in 2013, 88,000 ha in 2012, and 62,000 ha in 2011 (CSFS 2014a). Most outbreaks in Colorado last for a few years and subsided naturally.

The main objective of this research project was to characterize the temporal and spatial distribution of area damaged by the five damage agents and two disorders based on aerial survey data from 1994 to 2013. Our specific questions addressed in this chapter were: (1) how did the amount of area damaged vary by forest type and (2) what was the pattern of damage and how did it change over temporally and spatially? These questions were addressed with individual agents and disorders and total overall damage. We expected that knowing the spatial and temporal characteristics of each causal agent would be helpful and contribute to understanding the relationship among causal agents and their hosts. We also expected that knowledge of the dynamics of damage dispersion might improve forest inventory and monitoring efforts.

Material and Methods

Study site

About 9,308,000 ha of the state of Colorado are covered by forest (Thompson et al. 2010), consisting of a variety of tree species that have been placed into nine different forest types: 1) aspen, 2) piñon-juniper, 3) spruce-fir, 4) mixed conifer, 5) oak shrublands, 6) ponderosa pine, 7) lodgepole pine, 8) montane riparian, and 9) plains (agroforestry) (CSFS 2014b). Our analysis was focused on the western part of Colorado in the Rocky Mountains (37-41^0 N, 102 –

109^0 W) where mostly are mountainous and covered by forestlands. Most of this area is administered by Region 2 of the United States Department of Agriculture (USDA) Forest Service.

GIS data

Digital maps and related data for area damaged by forest type and causal agent from 1994 to 2013 were obtained from the annual aerial survey of the USDA Forest Service, Forest Health Protection (FHP) by downloading from (http://www.fs.usda.gov/detail/r2/forest-grasslandhealth/?cid=fsbdev3_041629).

A GIS layer representing 155 parallel transects (3.2 km wide and 625 km long) was developed that covered the western part of the state. These transects were oriented from east to west and numbered from 1 to 155 from south to north. This GIS layer was overlaid on GIS layer of forest cover area of Colorado. Five transects did not contain any forestlands and were eliminated from further analysis. Transects were then overlaid on the layers of forest damaged and/or infested to obtain estimates of the amount of damaged on each transect. This was repeated for each of the 20 years in this study. Data were used for the analyses as describe below.

Spatial distribution tests

Spatial statistics tests (Reich and Davis 2011) were employed to test spatial autocorrelation (Moran's I test) and spatial distribution (runs test – test for non-randomness) of area damaged on individual transects. These tests were run in R version 2.5.12 using an RSpatial package for quantitative spatial statistics (Reich 2011).

Moran's I (Moran 1950) was used to test for spatial autocorrelation of the total area damaged by the various causal agents and disorders by year. Moran's I is defined as:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2} \quad [2.1]$$

where I is Moran's I; N is the number of transects indexed by i and j; X_i is the area of damage on the i^{th} transect; and w_{ij} is the element of a spatial weights matrix. If Moran's I is negative, this indicates the damage follows a cyclic pattern. If Moran's I is positive, this would indicate clustering of damage on adjacent transects, while a zero value would suggest damage is spatially independent among transects.

A runs test was used to test for non-randomness (Bradley 1968). A *run* is defined as a series of ones or a series of zeros. The number of runs is defined as the number of both zeros and ones. The number of sequential ones or zeros defines the length of the run (Penn State Science 2015). The run test is based on the assumption that the number of runs follows a binomial distribution. A *Z-statistic* was used to test the null hypothesis of spatial independence:

$$Z = \frac{R_{obs} - R_{exp}}{s_R} \quad [2.2]$$

where R_{obs} is the observed number of runs; R_{exp} is the expected number of runs; and S_R is the standard deviation of the number of runs. The values of R_{exp} and S_R were defined as:

$$R_{exp} = \frac{2n_1 n_2}{n_1 + n_2} + 1 \quad [2.3]$$

$$S_R = \sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)(n_1 + n_2 - 1)}} \quad [2.4]$$

where n_1 and n_2 denote the number of zeros and ones in the series, respectively.

The null hypothesis is rejected if $|Z| > Z_{1-\alpha/2} = 1.96$. Significant negative Z values indicate that damage on the transects resembles an aggregated (clustering) spatial pattern, while significant positive Z values indicate a uniform distribution of damage.

Severity classification

Four different levels of severity were defined according to percentage of area damaged: (1) light (1-10%); (2) moderate (11-30%); (3) severe (31-49%); and (4) very severe (≥50%). This classification was applied to each causal agent on the individual transects and years. The percentage of damage was calculated as the ratio of area damaged by each agent to the total amount of area damaged.

Results

Area damaged across forest types

Three forest types (aspen, mixed conifer, and western spruce-fir) (USDA Forest Service 2005), were impacted by the seven damage agents and disorders. The amount of area damaged per year of the three forest types generally increased from 1994 to 2007 (Table 2.1). The cumulative area damaged in mixed conifer was the largest (9,372,702 ha), followed by western spruce-fir forests (4,136,203 ha), and aspen forests (1,223,394 ha). Two different trends among forest types were observed in the period 1994 to 2013. A similar trend was found in the period from 1994 to 2007, when damage in all three forest types gradually increased. Another trend of area damaged in western spruce-fir forest was a fluctuation in amount which differed from the decreasing trend of the other two forest types from 2008 to 2013 (Figure 2.1). The total amount of area damaged in the mixed conifer forest peaked at 1,614,382 ha in 2007 and then dropped to 219,468 ha in 2013. The area damaged in aspen forests was mostly concentrated in the years 2005 to 2011. The trend in amount of damaged aspen was very similar to that of mixed conifer in the same period, which quickly increased from 17,707 ha (2005) to a peak at 349,746 ha in 2008, then gradually decreased to 19,411 ha in 2012 (Fig. 2.1). The trend of area damaged in western

spruce-fir increased from 1994 to 2013, with the highest amount occurring in 2013 (388,437 ha). The size of the area damaged in a given year ranged from less than 25,000 ha in aspen forests to less than 80,000 ha in western spruce fir forests to highly variable in the mixed confer forests (0 - 200,000 ha). Most areas of damage were concentrated at smaller size but varied by forest type.

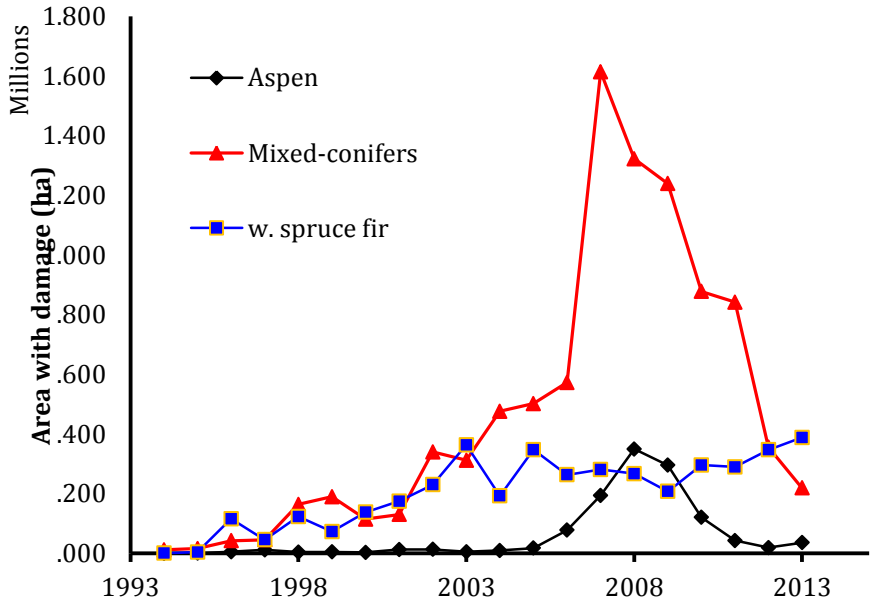


Figure 2.1. Temporal distribution of area damaged for three main forest types on Colorado forests from 1994 to 2013. Data based on USDA Forest Service annual aerial detection survey.

Table 2.1. Yearly combined area damaged in the three forest types mountain pine beetle, Douglas fir beetle, spruce beetle, sudden aspen decline, western spruce budworm, subalpine fir mortality, pine engraver, and others. Data were extracted from annual aerial detection survey from 1994 to 2013 by US Forest Service.

Year	Area damaged (ha)		
	Mixed conifer	Aspen	Western spruce-fir
1994	12021	-	650
1995	16126	877	4696
1996	42769	5492	115944
1997	26681	7907	29994
1998	164411	4445	123377
1999	190286	5066	73185
2000	114843	3410	138174
2001	130627	12788	174850
2002	340098	13291	230203
2003	312081	5275	363792
2004	476279	9986	192856
2005	502168	17707	347379
2006	572615	77951	263456
2007	1614382	194268	281392
2008	1322380	349747	267083
2009	1239250	295690	207688
2010	877999	121002	296332
2011	842041	42991	289704
2012	356175	19412	347011
2013	219468	36088	388437
Total	9,372,702	1,223,394	4,136,203

From a spatial or landscape scale, the area damaged associated with the three forest types had different starting points and expansion trends. The GIS layers of yearly area damaged showed that the area damaged in western spruce-fir forest type started expanding in 1996 followed by that in mixed conifer (1997) and aspen (2006), suggesting factors other than the hosts influenced the expansion of the infestations of which long-term drought during the period from 1990s to 2000s (USDA Forest Service 2015) would be played an important role. Spatial distribution of the cumulative amount of area damaged from 1994 to 2013 in the three forest types showed that mixed conifer damage was concentrated in the north and south of the state (Fig. 2.2). The very large area damaged were found in mixed conifer forests suddenly dropped in the far north of the state part would suggest the lack of abundant host trees after outbreak. Such an outbreak occurred in 2012 where beetles killed nearly all of the mature lodgepole pine trees in northern Colorado and southern Wyoming (USDA Forest Service 2015). Very few damaged areas were found in the central part of the state. The total amounts of area damaged were widely spread in the western part for aspen and throughout the entire state for western spruce-fir. Overall, northern part Colorado had the largest area damaged (8.1 million ha), followed by the southern (4 million ha) and central (2.6 million ha) parts of the state (Table 2.2). In most years, the areas damaged in the major forest types had a clustered distribution. Very few years were found where total area damaged was either random or of a uniform distribution.

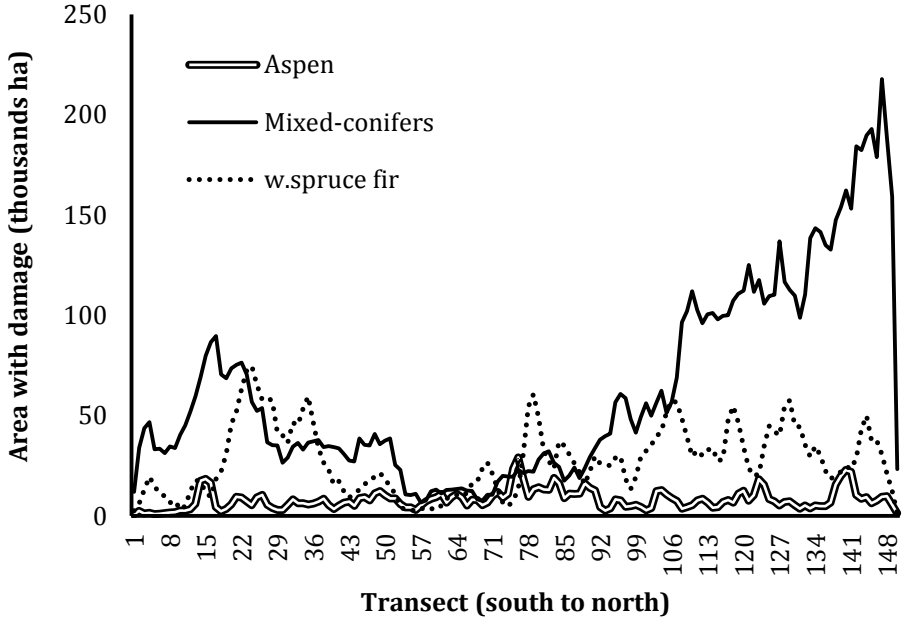


Figure 2.2. Spatial distribution of cumulative area damaged (1994-2013) of the three major forest types in Colorado based on aerial detection survey data. Sample transects were generated with 3.2 km wide x 625 km long covered the state's forestlands which oriented from east to west.

Table 2.2. Regional spatial distribution of the cumulative area damaged (ha) for the three major forest types (1994 to 2013). Damage was caused of all causal agents and disorders combined. Data were based on USDA Forest Service annual aerial detection survey.

Region	Aspen	Mixed-conifers	Western spruce-fir	Total
Southern	309,470	2,248,319	1,411,978	3,969,767
Central	489,138	1,194,213	971,239	2,654,590
Northern	424,786	5,930,170	1,752,987	8,107,943

Area damaged by causal agent

Spatio-temporal distribution of area damaged

Cumulative area damaged by region varied among causal agent (Table 2.3). The largest area damaged was found in the northern part of Colorado (160,320 ha), followed by the southern (70,309 ha) and central (48,763 ha) parts. WSB was the most destructive causal agent in the south, while SUB and MPB were the most destructive in the central and north parts of state, respectively. The contribution of all five causal agents and two disorders combined by region increased from a total area damaged of 62.0% in the south to 74.9% in the central region and 98.0% in the north.

The spatio-temporal distribution of area damaged caused by single causal agents and disorders had a strong relationship with the spatio-temporal distribution of relevant damaged forest types. In mixed conifer forests, the area damaged caused by MPB were found in almost every transect from south to north in the period 1996 to 2005. From 2004 to 2013, MPB distribution shifted to the north with no area damaged in the southern and central parts of the state. Area damaged caused by MPB gradually increased from south - central to northern part of the state (Fig. 2.3). A very similar trend of spatial distribution of cumulative area damaged was seen in mixed conifer forests (Fig. 2.2), indicating that MPB was the trend-leader for distribution of area damaged in those forests. Even though the area damaged by MPB had a decreasing trend from 2011 to 2013, its contribution (in percentage) to the total amount of damage caused by all agents combined still increased (Fig. 2.3), indicating that in the same period area damaged caused by the other causal agents combined decreased faster than those caused by MPB.

Table 2.3. Regional spatial distribution of the cumulative area damaged by causal agents and disorders over 20-year period from 1994 to 2013. Each region consisted of 50 transects. Area damaged (ha) was calculated as average area damaged within each region while average percentage was generated by dividing area damaged caused by each agent by the total area damaged caused by all agents combined. Data were based on USDA Forest Service annual aerial detection survey

Causal agents and disorders	South		Central		North	
	(ha)	(%)	(ha)	(%)	(ha)	(%)
Mountain pine beetle	2,950	2.7	14,674	22.2	116,405	69.9
Douglas –fir beetle	2,350	2.2	2,293	3.9	389	0.3
Spruce beetle	19,020	16.3	1,981	2.9	5,008	2.8
Pine engraver	1,284	1.3	542	0.9	410	0.3
Western spruce budworm	31,303	27.6	3,709	4.3	392	0.3
Sudden aspen decline	4,469	3.7	7,857	13.8	7,819	4.9
Subalpine fir mortality	8,933	8.1	17,708	26.9	29,896	19.6
Total	70,309	62.0	48,763	74.9	160,320	98.0

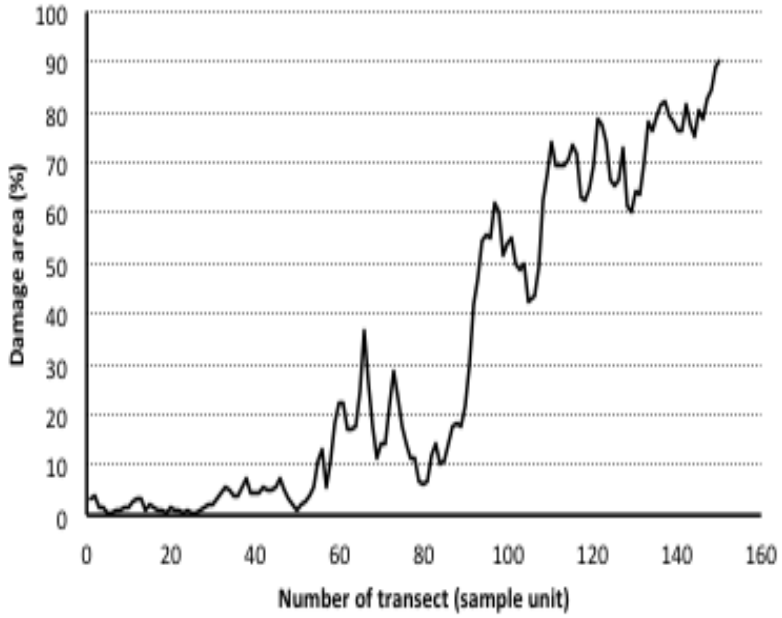


Figure 2.3. Spatial distribution of cumulative area with mountain pine beetle (MPB) damage in 20 years (1994 to 2013) compared with total amount of damaged by all causal agents and disorders combined. Data were based on USDA Forest Service annual aerial detection survey. Transects were numbered from south to north of the state with 3.2km wide x 625 km long.

The distribution of area damaged by DFB slowly shifted from the central region to the south, but was mostly concentrated in the central part of the state. The different trends of total damage in mixed conifer forests (Fig. 2.2) and total damage caused by DFB (Fig. 2.4) suggest that DFB did not have the strong impact on the mixed conifer forest type. The maximum contribution of area damaged by DFB to the total amount was 17.4%, located in the central part of the state (Fig. 2.4).

In contrast, WSB mostly infested mixed conifer forests (Table 2.6), while a small area damaged by WSB was found in western spruce-fir forests (Table 2.8). Area damaged by WSB had a clustered distribution and was concentrated in the southern part of the state (Fig. 2.4).

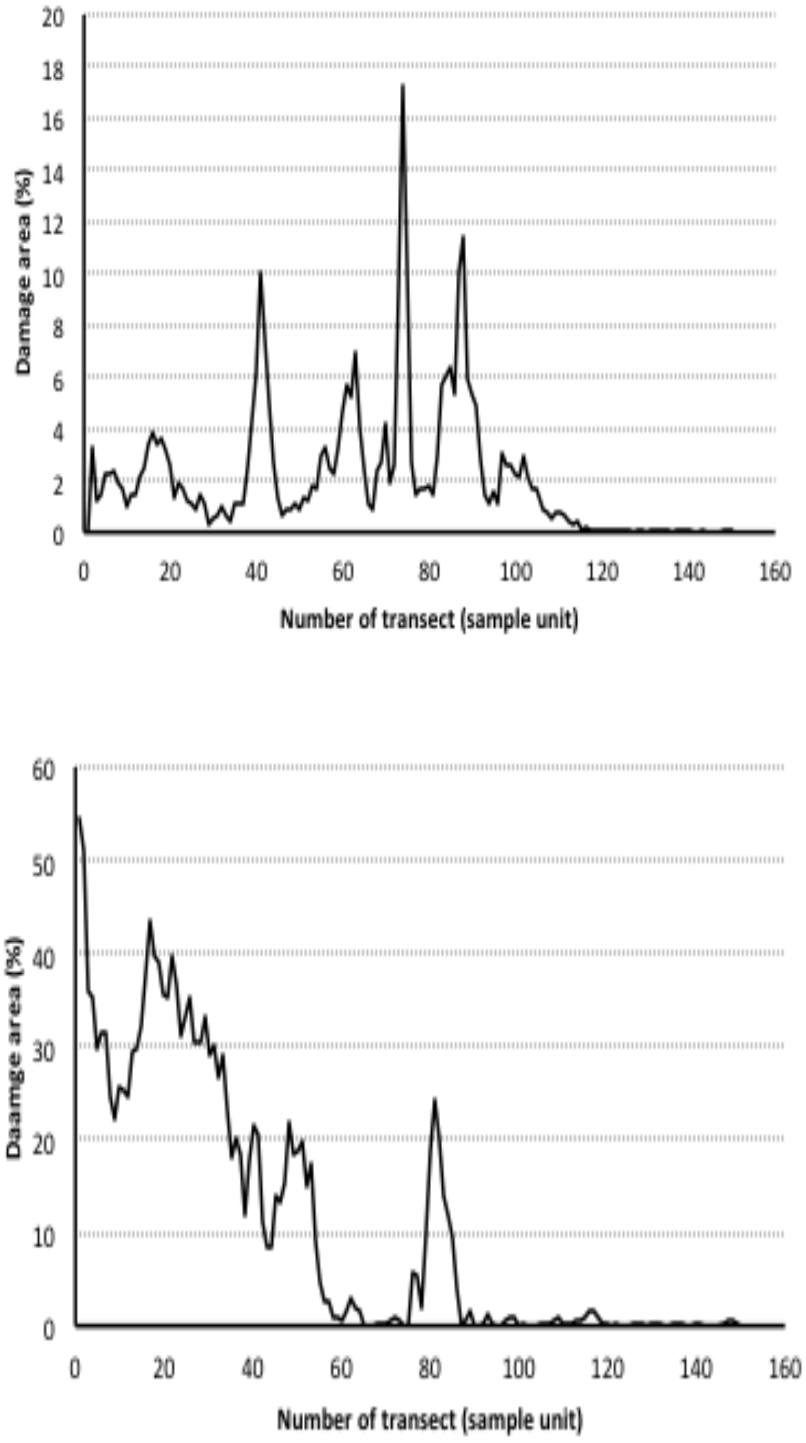


Figure 2.4. Spatial distribution of cumulative area with damage in 20 years (1994-2013) caused by Douglas fir beetle (above) and western spruce budworm (below) compared with total amount of damage by all causal agents and disorders combined. Data were based on USDA Forest Service annual aerial detection survey. Transects were numbered from south to north of the state with 3.2km wide x 625 km long.

SB and SUB were the main damage agents in the western spruce-fir forests (Table 2.8). The damage caused by SUB was concentrated in the central part of the state (Fig. 2.5), while area damaged by SB were distributed throughout entire western part of the state. SB with fewer area damaged tended to spread from north to south, especially from 2001 to 2013. SUB area damaged appeared in most transects, with the largest area damaged concentrated in the central part of the state (116,760 ha) during the period 2007 to 2013, followed by the northern part (105,776 ha) in the early period 1996 to 2006. SUB's trend of area damaged moved from south to north in the first period (1995-2004), then spread to the entire state. The largest contribution of damage caused by SB to the total amount of damage from all causal agents and disorders combined was 41.2%, in the southern part of the state (Fig. 2.5).

SAD affected only aspen forests (Table 2.7). It was found in most areas in the state, especially from 2005 to 2011, but the largest area damaged was concentrated in the central part of the state (Fig. 2.6). The yearly distribution of area damaged by SAD would suggest its damage cycle was about eight years; it started expanding from the central portion to the entire state from 2004 to 2011. By orientation from east to west, it was found that the cumulative area damaged in aspen forests concentrated in the western parts.

PE contributed a minor amount to the total amount of damage caused by all causal agents and disorders combined as well as the total amount of area damaged in mixed conifer forests (Table 2.6). The distribution of PE area damaged distribution had no obvious trend, but the largest cumulative of area damaged was generally concentrated in the southern part of the state and gradually decreased toward in the north part of the state (Fig. 2.6).

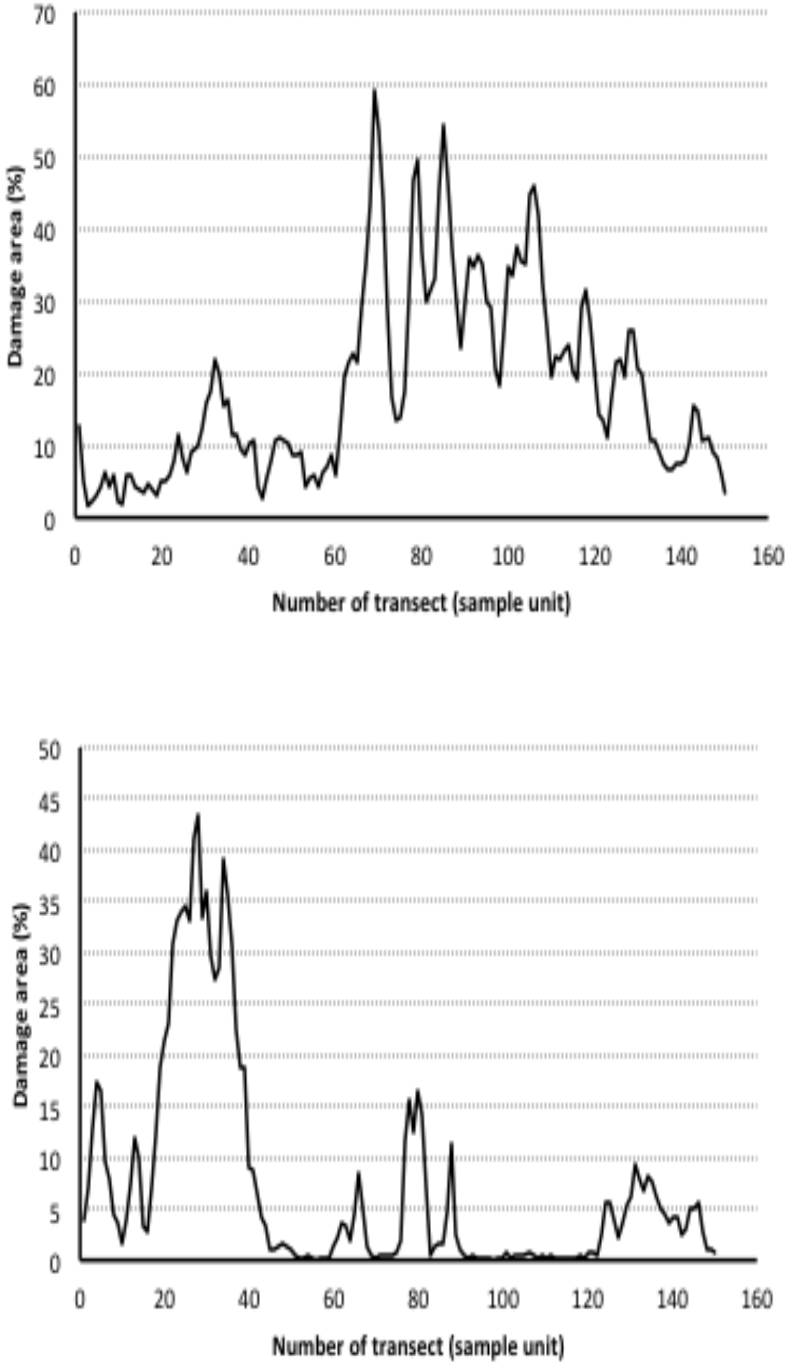


Figure 2.5. Spatial distribution of cumulative area with damage in 20 years (1994-2013) caused by subalpine fir mortality (above) and spruce beetle (below) in comparison with total area damaged caused by all causal agents and disorders combined. Data were based on USDA Forest Service annual aerial detection survey. Transects were numbered from south to north of the state with 3.2km wide x 625 km long.

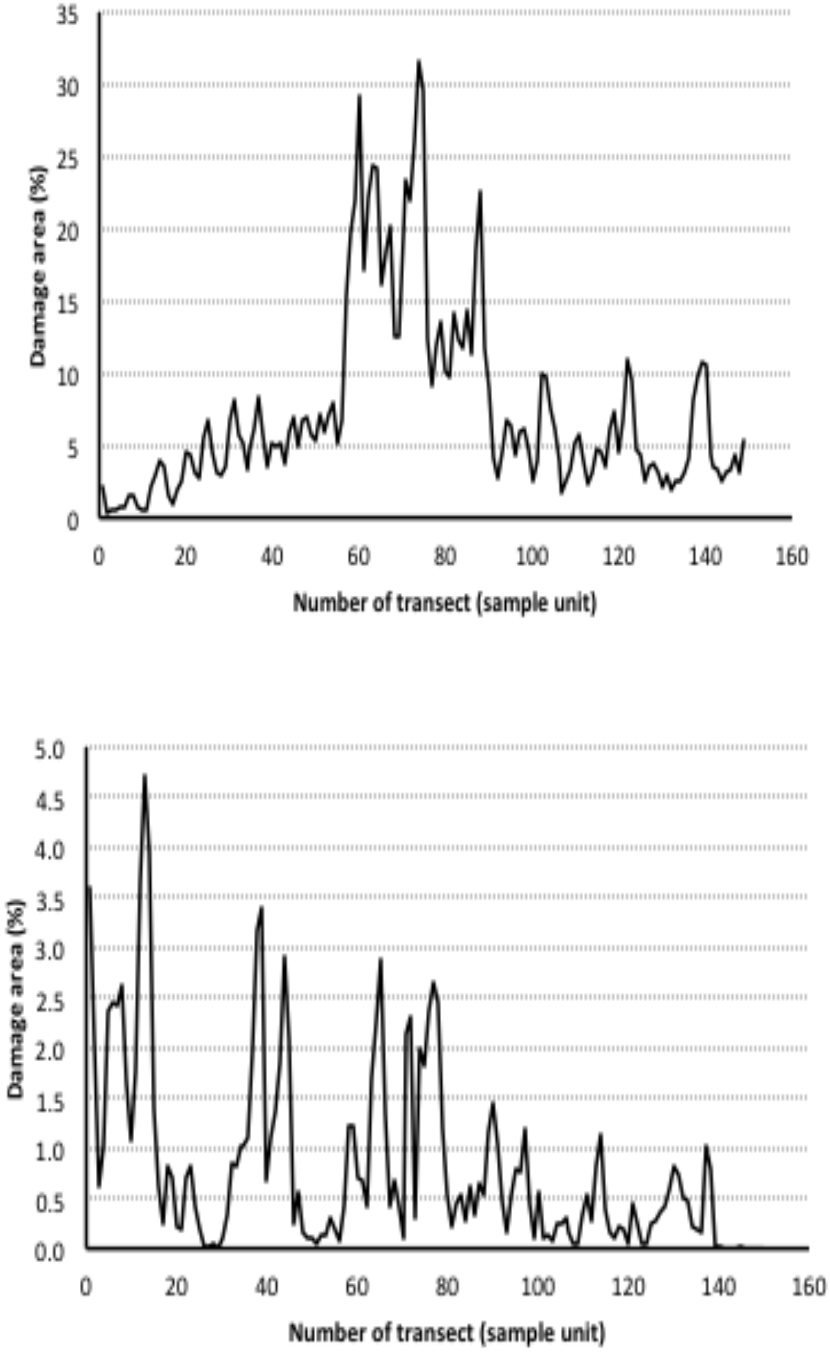


Figure 2.6. Spatial distribution of cumulative area with damage in 20 years (1994-2013) caused by sudden aspen decline (above) and pine engraver (below) in comparison with total amount of area damaged by all causal agents and disorders combined. Data were based on USDA Forest Service annual aerial detection survey. Transects were numbered from south to north of the state with 3.2km wide x 625 km long.

The total amount of area damaged by all causal agents and disorders gradually increased from 1994 to 2009, mostly concentrated from 2007 to 2013, reached a peak in 2010 with 1,297,797 ha, then gradually decreased to 660,000 ha in 2013 (Fig. 2.7). This trend was very similar to the trend of area damaged by MPB suggesting MPB was the main damage agent and had the strong impact on the total amount of damage in the state.

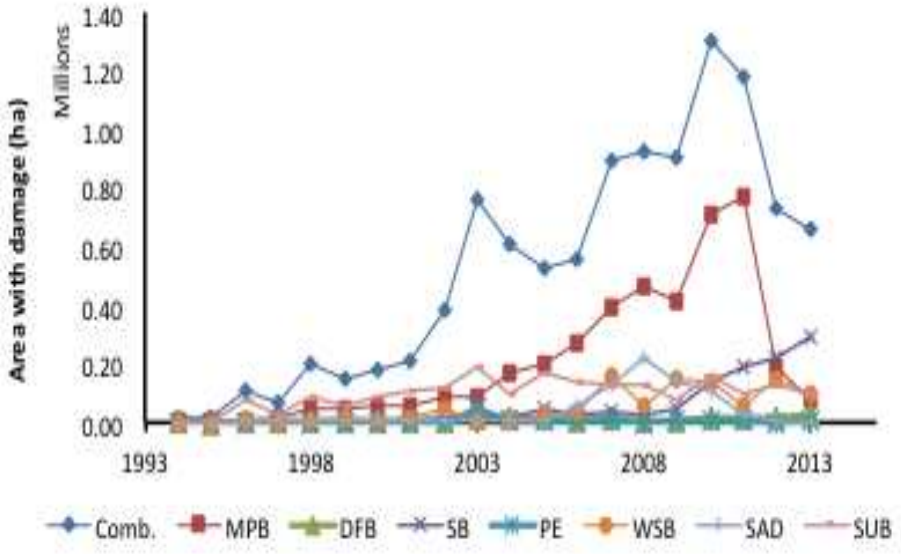


Figure 2.7. Temporal distribution of the total amount of area damaged caused by different eight causal agents and disorders from USDA Forest Service aerial survey data 1994 to 2013. Data came from the sum of area damaged in all transects for each year. Name of each causal agent as the following: Comb. – all causal agents and disorders combined; MPB- mountain pine beetle; DFB – Douglas fir beetle; SB – spruce beetle; PE – pine engraver; WSB – western spruce budworm; SUB – subalpine fir mortality; SAD – sudden aspen decline.

Spatial autocorrelation of area damaged

The amount of damage on individual transects showed significant positive spatial autocorrelation that varied by time and causal agent. Moran's I was used to measure the degree of spatial autocorrelation of the area damaged on adjacent transects for individual causal agents

and disorders over time. The degree of spatial autocorrelation varied from a low of -0.02 to a high of 0.95. For example, in 1997 both SB and PE showed no spatial autocorrelation in the amount of damage on an individual transect. As the infestation of SB increased over time, Moran's I increased to 0.93 in 2012, while the Moran's I for PE varied from 0.19 to 0.78. Other causal agents and disorders such as Comb., MPB, SUB and WSB showed an increasing trend in the degree of spatial autocorrelation over time, while the other causal agents had patterns similar to that observed for PE.

The frequency distribution of Moran's I of cumulative area damaged over a 20-year period is given Table 2.4. Most area damaged had more than 75% of Moran's I statistics greater than 0.85 except for the area damaged by DFB, PE, and SAD. The measure of spatial autocorrelations varied by causal agent and time. Area damaged caused by MPB, SB, and WSB had a similar overall trend of Moran's I that increased from 1994 to 2013, suggesting that these area damaged tended to cluster over time. Area damaged caused by DFB and SUB had positive Moran's I also, but were trendless. PE and SAD had weaker spatial autocorrelations with smaller Moran's I, suggesting a more scattered spatial distribution than the area damaged caused by MPB, SB, and WSB. Area damaged caused by SB, PE, and WSB each had one year with a Moran's I approaching zero (-.01, -.02, and -.01, respectively), indicating in those years the area damaged were independent among transects (Table 2.4), thus would be hard to determine severity using "hot spot" techniques and it would be difficult to suggest a sample size by which to achieve a desired boundary on the error of estimation for the individual causal agents and disorders.

According to Assuncao and Reis (1999) the spatial correlation of a significant Moran's I test implies that close areas tend to have similar risks that produce cluster of similar values

(Assuncao and Reis 1999). In this study, as mentioned above most the tests of spatial autocorrelation for area damaged varied by causal agents and disorders had significant positive Moran's I indicated strong spatial relationship between adjacent transects. Thus the higher probability of finding area damaged would be found next existing area damaged, therefore would be useful for predicting area damaged at specific small landscape scale. These properties also suggest the area damaged tend to be cluster distribution making a cyclic or periodical patterns on landscape scale. The Moran's I of area damaged caused by PE was quite consistent smallest over 20 years (0.42 on average) and fluctuation indicated the periodic outbreaks of PE over time. Based on PE's spatial autocorrelation properties accompany with its spatio-temporal distribution over time, one would believe that currently PE is not serious damage causal agent throughout the state. Other considered causal agents and disorders such Comb., MPB, SUB, and WSB with general increasing Moran's I over time would need to pay more attention.

Tests of non-randomness

The area damaged by individual causal agents and disorders had a cluster distribution in almost every year while area damaged by all causal agents and disorders combined (Comb.) had a more uniform patterns over the 20 year period (Appendix A2). These findings were very similar to the results of spatial autocorrelation tests suggesting that in each single year, one could find area damaged by a single agent or disorder in specific locations because of the clustering, while the area damaged by all causal agents and disorders combined could be found throughout the state. These findings also suggested that the precision of estimates would be strongly influenced when using different sample designs.

Table 2.4. Distribution of Moran's I for spatial autocorrelation of area damaged on individual transects over 20 year period (1994-2013). Moran's I ranges from -1 to +1 with negative value indicates the damage following a cyclic pattern. If Moran's I is positive spatial this would indicates clustering of damage on adjacent transects, while a value zero would suggest damage is spatially independent among transects.

Causal agents & disorders	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Mountain pine beetle	0.476	0.780	0.887	0.831	0.940	0.954
Douglas fir beetle	0.518	0.620	0.689	0.684	0.770	0.818
Spruce beetle	-0.010	0.619	0.878	0.717	0.913	0.930
Pine engraver	-0.020	0.286	0.409	0.424	0.520	0.783
Western spruce budworm	-0.008	0.772	0.847	0.753	0.892	0.950
Sudden aspen decline	0.195	0.502	0.602	0.560	0.659	0.728
Subalpine fir mortality	0.391	0.769	0.814	0.785	0.852	0.900
Combined	0.679	0.802	0.890	0.861	0.924	0.956

Severity of damage

Four causal agents, MPB, DFB, PE, and WSB infested mixed conifer forests (Table 2.6) with different levels of severity. MPB was the most destructive insect, causing an average of 38.4% of the total area damaged. The highest percentages of area damaged by MPB were found in 2010 (80%, equivalent 706,118 ha) and 2011 (91.0%, equivalent 766,625 ha). The other causal agents and disorders covered large percentages of area, with 42.3% on average (Table 2.6). The area damaged by DFB was the highest in 1994 at 65.2% and then sank to small portions in the following years. After reaching their peaks, the size of the area of damage caused by MPB and DFB both dropped. In contrast, WSB infested mixed conifer forests with small

percentages in the first 18 years but become more significant in the last two years. PE was consistently considered a minor infestation causal agent (Table 2.6).

Most of the years, SAD was the main agent that affected aspen forests. The area damaged by SAD increased from 1994 to 2008 then decreased, following the decreasing trend of total amount of aspen area damaged. The average area damaged caused by SAD was 55.7% of the total amount of area damaged, mostly damaged in the two periods from 1995 to 2001 and from 2005 to 2011 (Table 2.7). In 2012 and 2013, damaged aspen forest areas were not large since SAD represented a very small percentage of aspen forest damage (Table 2.7).

Three causal agents, WSB, SB, and SUB, infested western spruce-fir forests (Table 2.8). SB and SUB were the main causal agents and disorders of which SB was the most important in 1995 (83.5%). In the first period (1996-2008), SUB played an important role in damaging the western spruce-fir forest and accounted for 60.5% of the area damaged. There was an increasing role for SB in the following years. In 1997 and in the final four years (2010-2013), SB and SUB were the only causal agents and disorders noted in western spruce-fir damage (Table 2.8).

Three different levels of severity from “light” to “very severe” were classified for the seven causal agents and disorders based on what percent of the total cumulative amount of area damaged the agent caused (Table 2.5). Only MPB was a very severe damage agent (with 64.7% on average of total amount of damage). SUB, SB and WSB were moderate, while the three remaining causal agents and disorders (SAD, DFB, PE) accounted for light damage. These findings suggest that MPB was the most destructive agent, following by SUB, WSB and SB. However, for each single year the levels of severity varied from “light” to “very severe” depending causal agents and disorders.

Table 2.5. Severity class as defined by average percentage of cumulative area damaged in 20 years (1994-2013) by different causal agents and disorders. Data were calculated from average of total area damaged by each causal agent divided by total area damaged caused by all causal agents and disorders combined. Table also provides information on the average area damaged in each class in this period (1994-2013). Classification of severity based on: 1-10%: light; 11-30%: Moderate; 31-49%: Severe; ≥50%: Very severe.

Causal agent and disorder	Percentage	Area of damaged (ha)	Severity class
Mountain pine beetle	64.7	335,125	Very Severe
Subalpine fir mortality	27.3	141,344	Moderate
Western spruce budworm	17.1	88,509	Moderate
Spruce beetle	12.6	65,022	Moderate
Sudden aspen decline	9.7	50,386	Light
Douglas-fir beetle	2.4	12,579	Light
Pine engraver	1.1	5,590	Light

Table 2.6. Area damaged in mixed conifer forest by year based on aerial detection data, by main damage agents: mountain pine beetle (MPB), Douglas fir beetle (DFB), pine engraver (PE), western spruce budworm (WSB). The "others" damage agent including all causal agents and disorders which were out of seven considered causal agents and disorders.

Year	Area of Mixed Conifer Damage										
	Total (ha)	MPB (ha)	%	DFB (ha)	%	PE (ha)	%	WSB (ha)	%	Others (ha)	%
1994	12021	1542	12.8	7836	65.2	-	0.0	1	0.0	2563	21.3
1995	16128	2828	17.5	756	4.7	-	0.0	-	0.0	12522	77.6
1996	42769	5275	12.3	8119	19.0	26	0.0	10506	24.6	18812	44.0
1997	26681	12835	48.1	1432	5.4	-	0.0	193	0.7	12168	46.0
1998	164411	42363	25.8	5312	3.2	114	0.0	8910	5.4	107683	66.0
1999	190286	49423	26.0	2385	1.3	22	0.0	16616	8.7	121812	64.0
2000	114843	56389	49.1	4973	4.3	8	0.0	8230	7.2	45189	39.4
2001	130627	60948	46.7	3724	2.9	-	0.0	14347	11.0	51560	39.5
2002	340098	83982	24.7	5778	1.7	8	0.0	50461	15.0	199843	58.7
2003	312081	92073	29.5	18863	6.0	49487	15.9	8107	2.6	143499	46.0
2004	476279	173284	36.4	16460	3.5	1845	0.4	4249	0.9	280400	59.0
2005	502168	198487	39.5	13604	2.7	2590	0.5	26343	5.3	261101	52.0
2006	572615	267776	46.8	9409	1.6	5637	1.0	37349	6.5	252394	44.1
2007	1614382	395341	24.5	17999	1.1	1623	0.1	156857	9.7	1042538	64.6
2008	1322380	462100	34.9	11010	0.8	877	0.0	61897	4.7	786459	59.5
2009	1239250	416666	33.6	9099	0.7	2074	0.2	154363	12.5	657012	53.0
2010	877999	706118	80.4	19256	2.2	2942	0.3	142481	16.2	7119	0.8
2011	842041	766626	91.0	12711	1.5	2327	0.3	58172	6.9	2112	0.3
2012	356175	191614	53.8	13504	3.8	425	0.1	147709	41.5	2866	0.8
2013	219468	74289	33.9	22575	10.3	2425	1.1	95774	43.6	24359	11.1

Table 2.7. Area damaged in aspen forest by year in Colorado based on aerial detection survey 1994-2013 by main damage agent: sudden aspen decline (SAD). The "others" damage agent including all causal agents and disorders which were out of seven considered causal agents and disorders.

Year	Area Aspen Damage				
	Total (ha)	SAD (ha)	%	Others (ha)	%
1995	877	829	94.5	48	5.5
1996	5492	4101	74.7	1392	25.3
1997	7907	7907	100.0	0	0.0
1998	4445	3782	85.1	663	14.9
1999	5066	2790	55.1	2276	44.9
2000	3410	1614	47.3	1796	52.7
2001	12789	6299	49.3	6490	50.8
2002	13291	3170	23.9	10121	76.2
2003	5275	1685	31.9	3590	68.1
2004	9986	3269	32.7	6717	67.3
2005	17707	10864	61.4	6844	38.7
2006	77951	54993	70.6	22957	29.5
2007	194268	133163	68.6	61106	31.5
2008	349747	217879	62.3	131868	37.7
2009	295690	138278	46.8	157412	53.2
2010	121002	103931	85.9	17071	14.1
2011	42991	24938	58.0	18053	42.0
2012	19412	1720	8.9	17692	91.1
2013	36089	528	1.5	35559	98.5

Table 2.8. Area damaged in western spruce-fir forest by year in Colorado based on aerial detection survey 1994-2013, by main damage agent: western spruce budworm (WSB), spruce beetle (SB), and subalpine fir mortality (SUB). The "others" damage agent including all causal agents and disorders which were out of seven considered causal agents and disorders.

Year	Area Western Spruce-fir Damage								
	Total (ha)	WSB (ha)	%	SB (ha)	%	SUB (ha)	%	Others (ha)	%
1994	650							650	100.0
1995	4696			3919	83.5	133	2.8	560	11.9
1996	115944			83	0.1	75470	65.1	40391	34.8
1997	29994			27	0.1	29967	99.9	0	0.0
1998	123377			327	0.3	89550	72.6	33500	27.2
1999	73185			126	0.2	52882	72.3	20177	27.6
2000	138174	3701	2.7	928	0.7	83908	60.7	49634	35.9
2001	174850			3577	2.1	104714	59.9	66558	38.1
2002	230203	1132	0.5	21105	9.2	123719	53.7	84238	36.6
2003	363792	3187	0.9	28693	7.9	195034	53.6	136868	37.6
2004	192856	3897	2.0	20957	10.9	98318	51.0	69672	36.1
2005	347379			46402	13.4	166895	48.0	134069	38.6
2006	263456			26107	9.9	143081	54.3	94258	35.8
2007	281392			37501	13.3	125811	44.7	118067	42.0
2008	267083			24849	9.3	133674	50.1	108550	40.6
2009	207688			45210	21.8	73187	35.2	89269	43.0
2010	296332			141352	47.7	154980	52.3	0	0.0
2011	289668			188210	65.0	101458	35.0	0	0.0
2012	346969			219084	63.1	127885	36.9	0	0.0
2013	388437			288110	74.2	100327	25.8	0	0.0

Discussion

The amount of area damaged varied by forest type overall and by the type of damage. We found that the seven causal agents and disorders in this study mostly had a strong relationship with the location of their host plants, which was expressed by their contribution (in percentage) to the total amount of area damaged of each relevant forest type. Some damage agents that attacked trees in only one forest type, such sudden aspen decline were a dominant damage in that forest type. Other damage agents like western spruce budworm damaged trees in both mixed conifer forest and western spruce fir forest type so were not always the most dominate damage. Host availability is a major influence on distribution and severity, as pointed out in Lundquist and Reich's (2014) study of spatial patterns of dispersal behavior of forest insects across landscapes affected by host availability, population levels, mate-finding, and local density. Boyden et al. (2005) and Romme et al. (2006) also agreed that the high concentration of trees in each forest type should cause clump distributions of area damaged (Boyden et al. 2005, Romme et al. 2006, Paul Dunham 2008). In Colorado, a large variability of area damage was found in all three forest types. This suggests that the variability in area damaged is an important factor to consider when conducting sampling for forest inventory and monitoring since there is a need of a large sample size to get good estimates with low measures of error. These characteristics of distribution of area damaged should be incorporated in planning when doing inventory, especially in applying systematic sampling for estimating total amount of area damaged of interest. Without considering distribution of damage, one could under- or overestimate parameters.

Our findings that the patterns of damage were primarily clustered by individual causes but not so with total area of damage were expected based on host condition, distribution and pest biology. Previous reports on forest insects' infestation and infestation expansion characteristics

note that damage was clustered (Romme et al. 1986, Shore and Safranyik 1992, Hagle et al. 2003, Romme et al. 2006). In general, the direction and spread rate of a beetle infestation are hard to predict since there are many predisposing factors that we did not consider. However, this study confirms that forest insects and diseases usually spread to adjacent trees, resulting in cluster distributions of area damaged, which agrees with Williams and Liebhold (2000) (Williams and Liebhold 2000). Blackford (2010) stated that no typical patterns or trends were found in western spruce budworm epidemics (Blackford 2010). However, based on a landscape level and using long-term data, this study found that western spruce budworm was mostly concentrated in the southern part of the state, with severity decreasing from south to north.

The temporal change in the amount of damaged areas varied by forest type and agent involved and environmental factors involved. For example, the damage in aspen forests peaked in 2008 and was mostly concentrated from 2005 to 2011. These findings agreed with the previous reports by Chapman et al., (2012) and Safranyik et al., (2010) who mentioned the predisposing influence of long-term drought on forest insects and diseases infestation. Long-term drought affects trees both directly and indirectly. Directly, long-term drought slows trees growth and causes injury, dieback of the tree's crown or death. Indirectly, long-term drought weakens trees defense systems that make them more susceptible to insect pests and diseases. In Colorado, short-term droughts (3-month duration) covered as much as 80% of the state, while long-term droughts (2-4 years) have reached to about 70% of the state (McKee et al. 2000). These drought regime effects in aspen were also reported in strong relationship with elevation where lower elevations resulted in higher temperature and lower precipitation (Dudley et al. 2015). The decreasing trends of area damage in three forest types in the last few years of the study, also suggested the effects of stand condition such tree age, tree density, and tree diameter on insects

and disease outbreaks such it mentioned by Romme et al., (1986, 2006). This trend resulted from decreasing numbers of host trees available after and reduced tree density and large diameter after outbreaks.

Summary

In studying insects and diseases at the landscape level, spatial and temporal patterns are important indicators and spatial analysis is becoming a common tool for landscape ecology (Lundquist 2005). The largest area damaged were found in the north following by the south and central parts of the state. Moran's I and non-randomness tests indicated that most area damaged by single causal agents and disorders and by forest types had strong positive spatial autocorrelation and a clustered distribution in every year from 1994 to 2013, while area damaged caused by all causal agents and disorders combined were mostly distributed as uniform patterns.

The seven agents and disorders of forest damage appeared in Colorado forests in almost all years from 1994 to 2013. Mountain pine beetle, western spruce budworm, subalpine fir mortality, spruce beetle, sudden aspen decline, and Douglas fir beetle were considered as primary damage agents that contributed a noticable percentage to the total amount of area damaged in the state. Mountain pine beetle, subalpine fir mortality, spruce beetle and western spruce budworm were the most destructive agents and managers will need to be aware of and monitor them to prevent their development. Pine engraver was found in almost every year (exept 2001) and every part of the state, but its contribution to total amount of damage area was minor (1.1% on average of cumulative area damaged). Thus, pine engraver's threat to forests in Colorado may not be a serious problem if climate conditions do not change. This agrees with Livingston (2010), who considered pine engraver as an always present but occasionally damaging beetle (Livingston 2010).

CHAPTER 3: OPTIMAL SAMPLE DESIGN FOR ESTIMATING FOREST DAMAGE AND MORTALITY FROM AERIAL DETECTION SURVEY

Introduction

Aerial detection survey (ADS)

Aerial detection survey is a powerful tool for monitoring changing features in forested landscapes (Saeki 2005). Insect pests and tree pathogens are among the major drivers of these changes. The earliest flight to survey forest insect damage was made by J. M. Miller in 1925 over the Sierra National Forest in California (McConnell et al. 2000). One of the first aerial surveys to detect and map forest damage to mountain pine beetle was conducted in the 1960s (Wulder et al. 2006). More recently, the Forest Health Protection branch of the USDA Forest Service routinely assesses forest health using aerial surveys (Johnson and Wittwer 2008). Aerial detection survey in forests, also known as aerial sketch-mapping, involves using an aircraft with a trained observer(s) to systematically fly over a forested area to detect visible damage and/or mortality, and manually documenting observations to a map (McConnell et al. 2000, Johnson and Wittwer 2008). Initially, this technique involved paper maps, but with advances in computer and touch-screen display technology this information is now commonly recorded on digital maps (US Forest Service 2005). To do either paper or digital sketchmapping,1:100,000 maps with a standard projection and coordinate system are needed (US Forest Service 2005). These fundamental requirements are based on aerial sketch-mapping standards developed in 1997 (McConnell et al. 2000).

Forest health surveys use two types of aerial sketchmapping: 1) overview survey - to sketchmap all new forest change events during one flight, and 2) specific survey - to map

primarily just one forest change event usually at the peak of activity (McConnell et al. 2000). During flights at high speed, often ranging from 115 kilometer per hour to 145 kilometers per hour at varying altitudes from 150m to 1525m above sea level (Johnson and Wittwer 2008), observers are asked to draw on maps area damaged as polygons where pest impact is characterized by size, shape, and location. These polygons might be coded with additional information such as type of forest, causal agent. At the flight speed, observers have about 20 seconds to recognize, identify the cause, classify and record damage they see (Johnson and Wittwer 2008). Obviously, skills and visibility of trained observers affect the precision and accuracy of aerial detection survey data (Caughley 1974, McConnell 1999).

The advantage of aerial survey over ground-based survey is that it provides an overview from a landscape perspective with 100% census. In addition, aerial survey offers a simple, inexpensive and quick alternative to record forest pest infestations on a map over other remote sensing techniques. However, timing of the survey, sun angle, appropriate altitude above ground level, speed of flight, and maps quality, and other factors affect the quality of sketchmaps, and ADS works well only with visible damages that can be seen from aircraft. The cost of sending observers up in an aircraft is becoming increasingly expensive (McConnell 1999), and, of course, safety is an increasingly dominant priority with the USDA Forest Service. The desire to improve data quality, while maintaining safety standards, is a constant goal.

Sample designs in aerial survey for natural resources management

In statistics, sampling is concerned with the selection of a subset of individuals from a population to provide an estimate of the characteristics of the population as a whole. In most cases, it is the key to estimation of population parameters. Naturally, populations are often very

large and almost impossible to measure completely. Sampling in this case plays an important role. Getting good estimates of population parameters at minimum cost and time while maximizing the utility of data is one of the main objectives of survey sampling (Tokola and Shrestha 1999). Sample design is considered basic in sampling theory (Traat et al. 2004). Different sample designs have been employed depending on the objectives of the survey. The choice of a sample design also influences the size and shape of the sampling unit.

Even though different sampling techniques could be applied to natural resources inventories for monitoring, some sample designs have been widely used in these approaches, such as simple random sampling (SRS) (Nusser et al. 1998, Gregoire and Valentine 2007, Theobald et al. 2007, Ståhl et al. 2010), stratified random sampling (STRA) (Smith 1981, Gregoire and Valentine 2007, Ståhl et al. 2010), probability proportional to size (PPS) (McGinn 2004, Stevens and Olsen 2004, Gregoire and Valentine 2007), and so forth. In practice, each sampling method has some advantages and disadvantages depending on the population being sampled. With aerial survey for large animals, Caughley (1977) commented that systematic sampling could eliminate navigation problems associated with random sampling and would be the most efficient means of mapping the distribution of animals. But when money, manpower, or time is limited, stratified sampling is the most precise for estimating population sizes (Caughley 1977).

Legg and Nagy (2006), Field et al., (2007), and Lindenmayer and Likens (2009) reviewed some natural resource monitoring programs and found that many of them have been ineffective because of inadequate program support, poor planning, and lack of rigorous study design (Thompson et al. 2011). These studies also suggest that the challenges in developing a rigorous design include the need to maximize the spatial balance of a random sample of plots

while minimizing sampling effort; the need for sample sizes sufficient to detect change in heterogeneous or highly variable environments; and the long time periods required to detect change. A choice of inappropriate sample design would result in inadequate levels of accuracy (Taiti 1981). In his discussion about comparing several aerial sample designs, Smith (1981) concluded that stratified sample design would be best, while the choice between systematic and random sampling needs to be discussed. He also stated that there is no sample design that can claim to be superior in all, or even most, circumstances. The appropriate sample design(s) need to be suited to specific characteristics of the population surveyed and the objectives of the survey (Smith 1981). In British Columbia, a stratified random block sample design was used to estimate the size of caribou (*Rangifer tarandus*) and moose (*Alces alces*) population size within four caribou herd areas known as the Finlay, Pink Mountain, Chase, and Wolverine herds and determine population compositions for each of the two species within these areas (Zimmerman et al. 2002). With stratified random block sample design, each study area was divided into square grids of 25km^2 sample units and followed a weighted stratified random block sample strategy. The weights were varied based on species. For the caribou (*Rangifer tarandus*) census, weights were based on the percentage composition of Pine Lichen Winter Range (primarily low-elevation lodgepole pine (*Pinus contota*) forests) and High Elevation Winter Range (primarily alpine and subalpine areas). For the moose (*Alces alces*) census, weights were dependent on the percentage composition of Moose Winter Range preference classes (Zimmerman et al. 2002). Flight speeds varied from 60 to 80 kilometers per hour and height above ground ranged from 50 to 200m. All animals observed were recorded with help from GPS and GIS to determine whether they were inside or outside the boundaries. The results suggested that the estimated precision from aerial survey was still influenced by sighting caribou in forested habitats. This could be overcome by

using a marked animals survey method accompanied by a stratified random block sample design. The variance of estimated population size would be improved with larger sample sizes. Stratified random block sample design had an advantage in counting caribou, which cluster together in groups. Another study in Alaska conducted by Evans et al. (1966) used a quadrant sampling method in an aerial census to estimate moose population size. The areas of interest were stratified into units of high, medium, and low moose density based on current observations. Sampling units one square mile in area were surveyed from the air. Results suggested that sampling protocol was useful for estimating moose populations with a reasonable cost (Evans et al. 1966), especially in areas where access is limited. However, the total cost of the survey was not provided.

A complete census (100%) has been commonly conducted in aerial detection surveys for monitoring wildlife and forest insects and diseases, such as in Colorado, over regions of interest. Every year from 1994 until the present, the Forest Health Monitoring (FHM) program and its partners have been conducting aerial surveys of the entire state of Colorado and recording the changes in area damaged of all forestlands. Until recently, it has been rare to find papers describing the use of sampling designs in aerial surveys for monitoring forest damage caused by various insects or diseases. In this study we proposed and evaluated four sample designs applied in aerial detection surveying for estimating area damaged in the forests of Colorado with the question was "what sample design(s) would be appropriate for estimating total area damaged?" *The overall objective of this study was to determine the best sample design that can provide an estimation of total amount of area damaged that is similar to the current 100% coverage with fewer transects and therefore reducing cost, time, and the risks of flying*.

Material and Methods

Study site

The study was carried out in western Colorado which is dominated by forested lands covering about 9,308,000 ha (37 - 41^0N, 102 -109^0 W) (Thompson et al. 2010). This region has a wide range of topography, soils and environmental conditions that influence the diversity of forest types found in this area. The landscape ranges from plains to high plateaus to steep mountains with deep canyons and sloping foothills (Bailey 1980). Major forest types found in this area include 1) aspen, 2) piñon-juniper, (3) spruce-fir, 4) mixed-conifer, 5) oak shrubland, 6) ponderosa pine, 7) lodgepole pine, 8) riparian, and 9) plains (agroforestry).

GIS data

A GIS layer dividing the state into 155 parallel transects (3.2 km wide and 625 km long) was developed to cover the study area. All transects were oriented east to west and numbered from 1 to 155, south to north.

Two sources of GIS information were clipped with the state's forestland boundary and used to obtain the data used in this study. The first was a GIS layer of the major vegetation types of the state at a 30m spatial resolution. This information was used to create a binary surface indicating if a given raster cell was classified as being forested or non-forested. This layer was intersected with the GIS layer of transects to obtain estimates of the area of forested and non-forested on each transect. Five of the transects did not contain any forest lands and were deleted leaving 150 transects. The second were GIS layers of causal agents and disorders obtained from aerial surveys of the state carried out from 1994 to 2013. These layers were intersected with the GIS layer of transects to obtain estimates of the area of damage caused by eight agents: spruce beetle *(Dendroctonus rufipennis*) (SB), mountain pine beetle (*Dendroctonus ponderosae*)

(MPB), Douglas-fir beetle (*Dendroctonus pseudotsugae* Hopkins) (DFB), western spruce budworm (*Choristoneura occidentalis* (Freeman)) (WSB), sudden aspen decline (SAD), subalpine fir mortality (*Picea englmanii – Abies lasiocarpa*) (SUB), pine engraver (*Ips pini* (Say)) (PE), and all causal agents and disorders combined (Comb.).

Sample designs

The statistical properties of four sample designs were evaluated in estimating the total amount of damage caused by various causal agents and disorders agents that occurred in the study area. The sample designing included: (1) simple random sampling (SRS), (2) probability proportional to the area of forests on the transect (PPS), (3) non-alignment systematic sampling (NALIGN), and (4) stratified random sampling (STRA) and are brief described below.

Simple random sampling (SRS)

Simple random sampling is the most basic sample design in which a sample of size *n* is drawn from a population of size *N* in such a way that every possible sample of size *n* has the same chance (probability) of being selected (Schreuder et al. 2004, Scheaffer et al. 2006, Mandallaz 2007). SRS is the simplest probability sampling technique and is considered best suited for situations where not much information is available regarding the population of interest (e.g., spatial extent and severity of the damage). In this study, eight sample sizes of n = 10, 15, 20, 25, 30, 35, 50 and 70 transects were selected, without replacement. The total area damaged (τ) associated with the various casual and disorders agents was estimated using the following formula.

$$\hat{\tau} = N \frac{\sum_{i=1}^{n} y_i}{n} \qquad [3.1]$$

with estimated variance

$$\hat{V}(\hat{\tau}) = N^2 \left(\frac{N-n}{N}\right)\left(\frac{S^2}{n}\right) \qquad [3.2]$$

and 0.95 bound on the error of estimation (B)

$$B = 2\sqrt{\hat{V}(\hat{\tau})} \qquad [3.3]$$

where y_i is the area damaged on the i^{th} transect, s^2 is the sample variance, and N is the total number of transects in the population.

Stratified random sampling (STRA)

A stratified random sample is obtained by dividing populations into smaller groups called strata and then randomly selecting sample units from each stratum (Scheaffer et al. 2006). Stratified random sampling has three major advantages over simple random sampling. First, the variance associated with estimating the population total is usually more precise because the variability within strata is usually smaller than the overall population variance. Second, the cost of collecting and analyzing the data is often reduced. And third, separate estimates can be obtained for each individual stratum. In this study, the total number of transects covered by forestlands (150) was divided into five smaller, regularly spaced regions of equal size, and then a sample (n_i = 2, 3, 4, 5, 6, 7, 10 and 14) chosen randomly without replacement from each stratum. Estimates of the total area damaged was obtained using the following equation

$$\hat{\tau}_{stra} = N_i \sum_{i=1}^{L} \bar{y}_i \qquad [3.4]$$

with estimated variance

$$\widehat{V(\hat{\tau}_{stra})} = \sum_{i=1}^{L} N^2{}_i \left(\frac{N_i - n_i}{N_i}\right)\frac{{S_i}^2}{n_i} \qquad [3.5]$$

and bound on the error of estimation

$$B = 2\sqrt{\hat{V}(\hat{\tau}_{STRA})} \qquad [3.6]$$

where L = 5 is the number of strata; N_i = 30 is the number of transects per stratum; n_i is the sample size in stratum, i; $\bar{y}_i$ is the mean damage in stratum *i*; and s_i^2 is the sample variance of stratum *i*.

Probability proportional to size (PPS)

Probability proportional to size (PPS) is a sampling technique for use with surveys in which the probability of selecting a sampling unit is proportional to some characteristic that is correlated to the variable of interest (McGinn 2004). PPS sampling will be more precise than SRS if the selection probabilities (π_i) are correlated with the variable of interest (y_i). In this design, the selection probabilities were based on the proportion of the transect classified as being forested; thus, more heavily forested transects had a higher probability of being selected than lightly forested transects. The use of sampling with replacement with PPS sampling has been widely adopted in sample surveys (Sirken 2001) but in this study, the sampling procedure was changed so that the sample of transects was drawn without instead of with replacement. If the selection probabilities are known, an estimate of the population total is given by

$$\hat{\tau}_{PPS} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\pi_i} \qquad [3.7]$$

with estimated variance

$$\hat{V}(\hat{\tau}_{PPS}) = \frac{1}{n(n-1)}\left(\frac{N-n}{N}\right)\sum_{i=1}^{n}\left(\frac{y_i}{\pi_i} - \hat{\tau}_{PPS}\right)^2 \qquad [3.8]$$

where N is population size and n is sample size

The 0.95 bound on the error of estimation is given by

$$B = 2\sqrt{\hat{V}(\hat{\tau}_{PPS})} \qquad [3.9]$$

The same eight sample sizes used with SRS were evaluated using this design.

Non-aligned systematic sampling (NALIGN)

As an alternative to systematic sampling, a non-aligned systematic sampling was employed. In this method, the state was divided into zones of equal size where the number of zones was equal to the sample size. A transect was then randomly selected from each zone. This design basically combines simple random sampling and systematic sampling because the sample transects are spread uniformly and independently throughout the state. Since there is not a minimum distance between sample transects, this design should capture the spatial variability in most populations. As with systematic sampling, non-aligned systematic sampling will be more precise than simple random sampling if the population elements are ordered (e.g. gradient). If the population elements are randomly distributed, non-aligned sampling should be equivalent to simple random sampling and less precise if the population elements exhibit a cyclic or periodic distribution.

The estimation of the population total ($\hat{\tau}_{NALIGN}$) was calculated:

$$\hat{\tau}_{NALIGN} = \sum_{i=1}^{n} N_i \, y_i \qquad [3.10]$$

with estimated variance of the total population was given by:

$$\hat{V}(\hat{\tau}_{NALIGN}) = \left(\frac{N-n}{N}\right) s^2 \sum {N_i}^2 \qquad [3.11]$$

where N is the population size; n is the sample size; s^2 is the sample variance; y_i is the area damaged on the i^{th} transect; and N_i is the number of transects in i^{th} zone. If N_i is constant and divisible by 150, these equations are equivalent to the equations used for simple random sampling.

Evaluating the Statistical Properties of the Sample Designs

To evaluate the statistical properties of the four sample designs (D), each design associated with each year (20 years) and each causal agent was implemented M = 20,000 times for each of the eight sample sizes and the following statistics were computed:

The grand total:

$$\hat{\bar{\tau}}_D = \frac{1}{M}\sum_{i=1}^{M} \hat{\tau}_D \qquad [3.12]$$

The mean variance:

$$\bar{V}(\hat{\bar{\tau}}_D) = \frac{1}{M}\sum_{i=1}^{M} \hat{V}\,(\hat{\bar{\tau}}_{Di}) \qquad [3.13]$$

and the variance of the total:

$$\tilde{V}(\hat{\bar{\tau}}_D) = \frac{\sum_{i=1}^{M}(\hat{\tau}_{Di}-\hat{\bar{\tau}}_D)}{M(M-1)} \qquad [3.14]$$

If the sample design (D) provides an unbiased estimate of the population total, the grand total should equal the true population total (τ). Likewise, if the estimated variance is unbiased, the mean variance should equal the variance of the total, the latter of which is taken as the true variance. To evaluate each sample design associated with the different sample sizes (n=10, 15, 20, 25, 30, 35, 50 and 70), the estimators of interest (bias, mean variance, coverage rate, and ratio of variances) were compared by year and then ranked based on their values. The average

ranking for all 20 years determined the best sample design. The bias was calculated as the difference between the estimated total and the population total. The smaller the bias, the more accurate the sample designs would be. A *t-test* was used to test the hypothesis that the sample designs produce unbiased estimates of the population total at the 0.05 level of confidence.

To evaluate the variance estimates, the ratio of the mean variance to the variance of the total was calculated. According to Hevesi et al. (1992), the estimates of the variance were unbiased if the ratio of the variance fell within the interval $[1 \pm 2\sqrt{\frac{2}{M}}]$, where M is the number of simulations (Hevesi et al. 1992). With M=20,000, if this ratio falls within the range of 0.98 to 1.02, this would indicate the variance estimates are unbiased. If the ratio is greater than 1.02, this would indicate an overestimation of the variance, while a ratio less than 0.98 would indicate an underestimation of the variance. To compare estimated variances, Monte Carlo methods were employed using a Friedman test. This method was conducted using SPSS version 15 (SPSS Inc 2006) to produce a χ^2 value.

$$\chi^2 = \frac{12}{ba(a+1)} \sum_i {R_i}^2 - 3b(a+1) \qquad [3.15]$$

where: *a* is the number of sample designs (treatment); *b* is the sample size; and R_i is the sum of the rank for i^{th} sample design. Under the null hypothesis of equal variances $\chi^2 \sim \chi^2_{(n-1)}$. If the calculated χ^2 is greater than the critical value (χ^2_{crit}) for a given sample size then the null hypothesis was rejected.

The sample design efficiency coefficient (D_{eff}) was calculated as an alternative to comparing different sample designs with respect to their variances. Sample design effect has been popularized by Kish (1965) and was defined by Cornfiel (1951) as the ratio of the variance of a statistic under simple random sampling without replacement to the variance of the statistic

under complex design with the same sample size (Park and Lee 2001)

$$D_{eff} = \frac{\hat{V}_{\hat{\tau}}(c)}{\hat{V}_{\hat{\tau}}(SRS)} \quad [3.16]$$

where $\hat{V}_{\hat{\tau}}(c)$ is the variance of the statistic under complex design, and $\hat{V}_{\hat{\tau}}(SRS)$ is the variance of the statistic under simple random sampling. If D_{eff} equals 1, both sample designs are equivalent; if D_{eff} is greater than 1, the alternative sample design is less precise than SRS; otherwise, it is more precise than SRS. This coefficient (D_{eff}) was used to rank sample designs with respect to estimated variance.

The 95% confidence intervals were calculated assuming normality as follows.

$$CI = \hat{\tau} \pm t_{\frac{\alpha}{2},n-1}\sqrt{\hat{V}_{\hat{\tau}}} \quad [3.17]$$

where CI is confidence interval; $\hat{\tau}$ is estimated total; $\hat{V}_{\hat{\tau}}$ is estimated variance of the total that could be calculated by the previous equations. Coverage probabilities (coverage rates) were calculated as the proportion of confidence intervals that contained the true value. The actual coverage probability could either be less than or greater than the nominal coverage probability. The confidence interval is said to be "conservative" if the actual coverage probability is greater than the nominal coverage probability, otherwise, it is "anti-conservative" or "permissive" (Jiang et al. 2008).

Results

Normality

Normality is an important assumption attached to estimates of the population mean and total in survey sampling. It follows from the Central Limit Theorem that for any population with mean μ and variance σ^2, if the population is repeatedly sampled using the sample size, estimates of the population mean will be normally distributed with mean μ and variance σ^2/n.

To test this assumption the frequency distribution of the 20,000 estimates of the total damage associated with the individual causal and disorders agents for each of the 20 years and four sample designs were usually inspected.

Results of this process showed that the frequency distributions of estimates of the total damage were approximately normally distributed for the four sample designs. The frequency distribution approached normality with increasing sample size. Figure 3.1 provides an example of the frequency distribution for area damaged caused by sudden aspen decline. The frequency distributions for the other causal agents and disorders showed a very similar trend. Hansen (1953) mentioned in his book about the important role of testing normality before generating further statistical properties, of which in practical problems of sampling from finite population very often that the initial population from which the sample is drawn is far from normal, and thus the assumption of a normal distribution may lead to grossly wrong impressions as to the precision of variance estimates (Hansen et al. 1953). The ability to assume normality simplifies the interpretation of the statistical properties of the four sample designs.

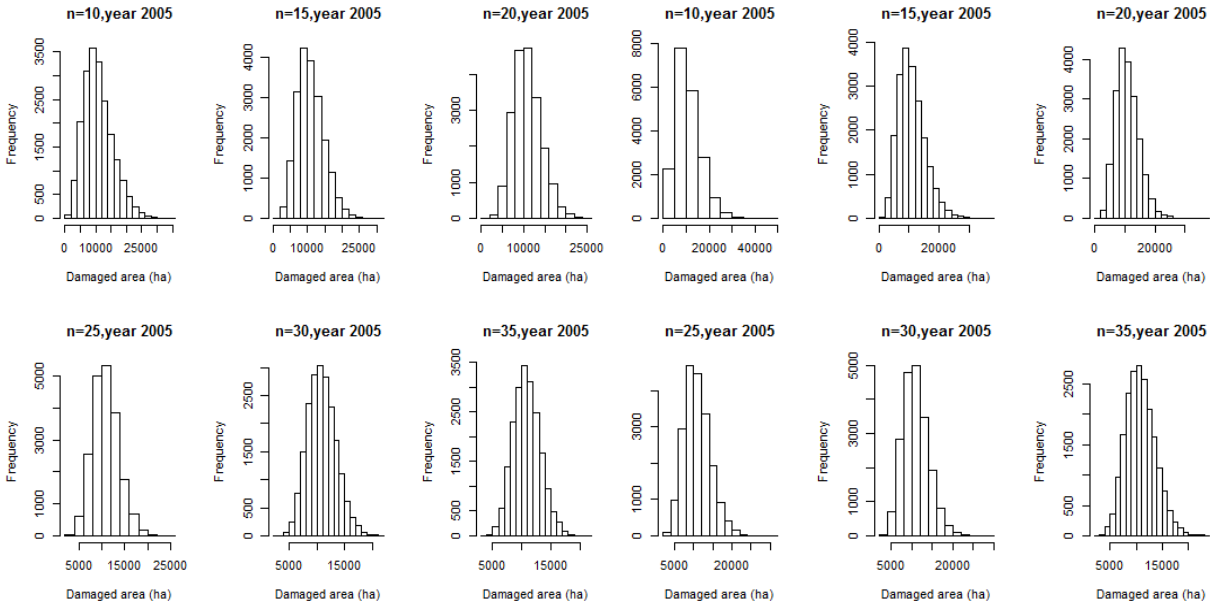
Bias

The majority of sample sizes evaluated for simple random sampling (SRS), non-alignmen systematic sampling (NALIGN) and stratified random sampling (STRA) produced unbiased estimates of the total area damaged, while probability proportional to size (PPS) estimates were biased. The percentage of unbiased estimates of the total area damaged produced by PPS decreased with increasing sample size irrespective of the causal agent (Fig. 3.2, Table 3.1). This result suggested that the selection probabilities (π_i) had a weak correlation to the variable of interest. These correlations were weaker with increasing variability in the population. The highest variability occurred when the area of damage was clustered throughout the landscape.

This clustering had a noticeable influence on the accuracy of the estimates. On average, PPS produced the smallest percentage of unbiased estimates of total area damaged when compared to other sample designs.

The PPS sample design had the smallest percentage of unbiased estimate at 33.8% for MPB and the highest value of 70.6% for DFB. The remaining sample designs (SRS, NALIGN, STRA) produced more accurate estimates, with at least 91.3% (Comb.) of the estimates being unbiased (Table 3.1). Overall, the percentage of unbiased estimates varied around 95% depending on sample designs and causal agents and disorders, of which STRA occupied the highest spot (95.2%), followed by NALIGN (95.0%) and then SRS (94.3%). Average ranking of bias produced by the four sample designs associated with the five causal agents and two disorders showed that on average STRA had the highest ranking with respect to producing unbiased estimates of the total area damaged, while PPS had the lowest ranking in its ability to produce unbiased estimates (Table 3.2)

SRS PPS



STRA NALIGN

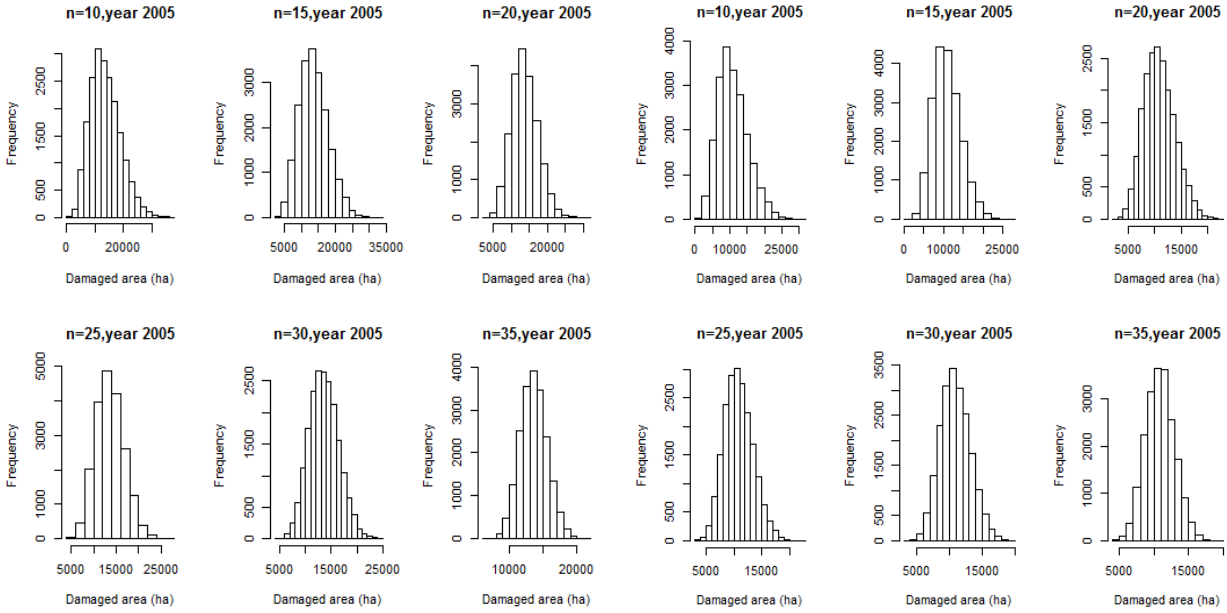


Figure 3.1. An example of the frequency distribution of 20,000 estimates of the total damage caused by sudden aspen decline (SAD) for the four sample designs and selected sample sizes. The x-axis is area damaged (ha), the y-axis is frequency.

Table 3.1. Percentage of unbiased estimate by sample design and causal agent. Percentages are averaged over eight sample sizes and 20-year time period from 1994 to 2013.

Causal and disorders	SRS[1]	PPS	NALIGN	STRA
Mountain pine beetle	95.6	33.8	96.9	96.9
Douglas-fir beetle	94.4	70.6	96.3	98.8
Spruce beetle	94.7	39.5	92.1	96.1
Pine engraver	95.6	53.7	94.9	92.7
Sudden aspen decline	95.4	54.0	96.1	98.0
Subalpine fir mortality	92.8	46.7	95.4	92.8
Western spruce budworm	94.7	38.8	96.1	92.8
Combined	91.3	36.3	92.5	93.8
Avg.	94.3	46.7	95.0	95.2

1. SRS-simple random sampling, PPS-probability proportional to size, NALIGN-nonalignment systematic sampling, STRA-stratified random sampling

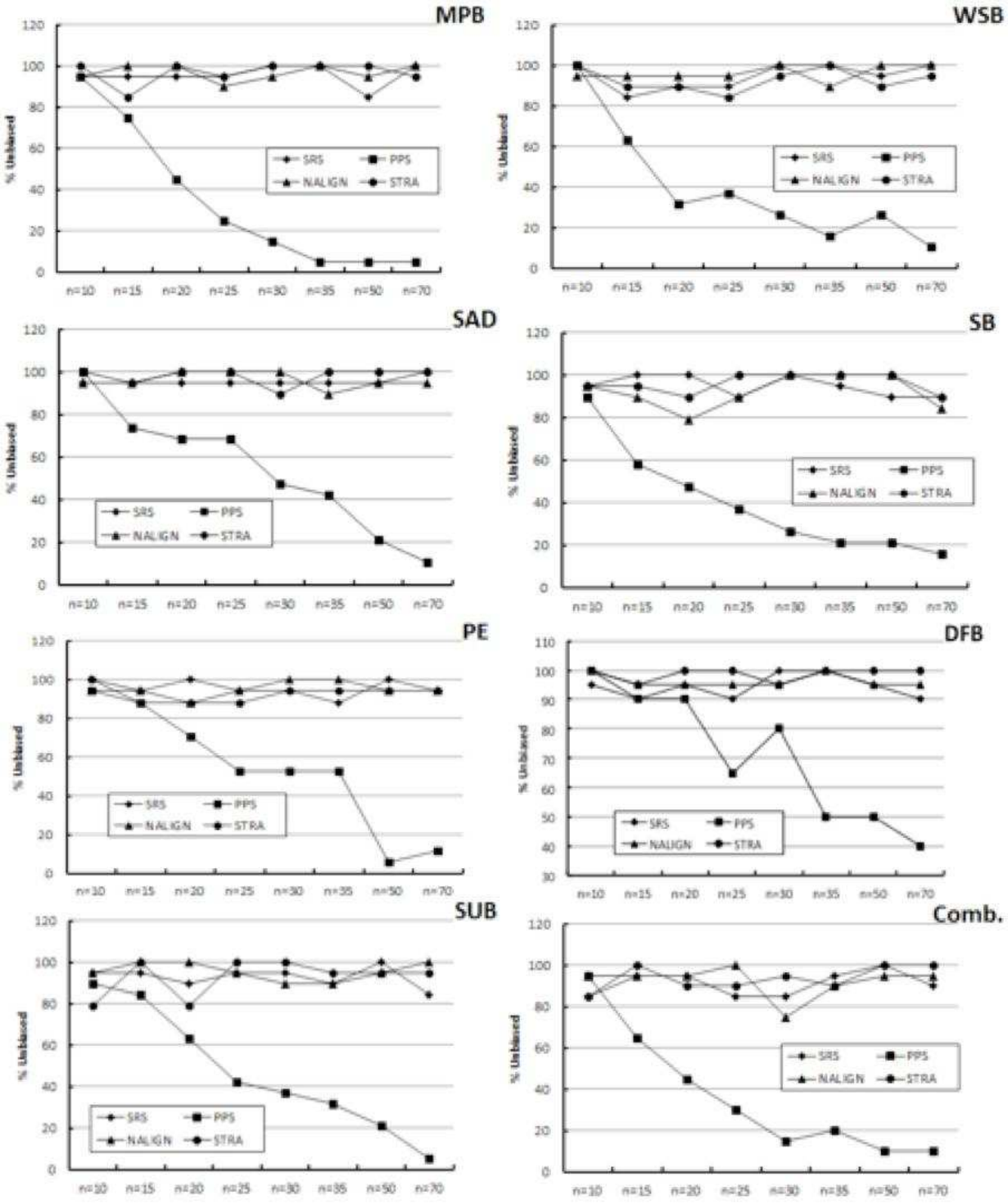


Figure 3.2. Percentage of unbiased estimates of the total damage areas for individual causal agents and disorders as a function of sample size (8) and sample design (4). Percentages are averaged over the 20-year period from 1994 to 2013. MPB-mountain pine beetle, DFB- Douglas fir beetle, SB-spruce beetle, PE-pine engraver, SAD-sudden aspen decline, SUB-subalpine fir mortality, WSB-western spruce budworm, Comb.-all causal agents and disorders combined.

Estimated mean variance

The estimated variance associated with estimating the total area damaged decreased with increasing sample size (Fig. 3.3). Similar trends were observed for all causal agents and disorders. The rate of decrease varied by sample design and causal agent. By ranking, STRA on the average always generated the smallest estimated variance for estimating the total area damaged by the seven different causal agents and disorders (except PE). This suggests that stratified random sampling provides the most precise estimates of the total area damaged by individual causal agents and disorders (except PE). On average, PPS had the smallest estimated variance of the total area damaged by PE. This suggests a strong correlation between the area damaged by PE and forest cover on the transects. The NALIGN sample design that combines systematic and simple random sampling produced the least precise estimates of the total area damaged for most causal agents and disorders (Table 3.3). The SRS sample design was ranked number two performing slightly better than PPS in estimating the variance of the total area damage.

Ratio of variance

The ratio of the variances associated with the four sample designs varied by causal agents and disorders (Table 3.4). Simple random sampling and stratified random sampling designs had ratios of the variances for all causal agents and disorders that fell within the interval 0.980 to 1.020, indicating unbiased estimates of the variance. The PPS sample design provided unbiased estimates of the variances for seven out of the eight causal agents and disorders. Average ratio of variance produced by PPS for area damaged caused by spruce beetle (0.974) fell below the lower critical value (0.980) indicating an under-estimation of the variance. NALIGN sampling, which

is basically a form of systematic sampling, produced a ratio of the variance greater than 1.02, indicating an over-estimation of the sample variance. This agreed with data in Table 3.3 where NALIGN had the worst ranking compared to other sample designs. These results suggested that SRS and STRA would be the most appropriate sampling designs for estimating the total area damaged. NALIGN was not considered a suitable sample design since it consistently over-estimated the variance.

Coverage rate

In general, all sample designs associated with the different sample sizes had coverage rates less than the nominal value of 0.95 for all causal agents and disorders evaluated. The coverage rates for all sample designs generally increased with increasing sample size independent of the causal agents and disorders. The rate of increase varied by year, causal agent, and sample design (Fig. 3.4). These results suggest that the four sample designs underestimated the level of confidence in the estimations despite having unbiased estimates of the sample variance.

Table 3.2. Average ranking of estimated bias for four sample designs associated with the eight causal agents and disorders. The smaller value, the higher rank was. The final rankings were based on averaging the individual ranking for each causal and disorders agents over the 20 years time period (1994-2013)

Causal and disorders	SRS[1]	PPS	STRA	NALIGN
Mountain pine beetle	3	4	1	2
Douglas-fir beetle	2	4	1	3
Spruce beetle	3	4	1	2
Pine engraver	3	1	4	2
Sudden aspen decline	3	4	2	1
Subalpine fir mortality	3	4	1	2
Western spruce budworm	3	4	1	2
Combined	3	4	1	2
Average	2.88	3.63	1.50	2.00

1. SRS-simple random sampling, PPS-probability proportional to size, NALIGN-nonalignment systematic sampling, STRA-stratified random sampling

Table 3.3. Average ranking of estimated mean variance for four sample designs associated with the eight causal agents and disorders. The smaller the value, the higher rank was. The final rankings were based on averaging the individual ranking for each causal and disorders agents over the 20 years time period (1994-2013)

Causal and disorders	SRS[1]	PPS	STRA	NALIGN
Mountain pine beetle	3	2	1	4
Douglas-fir beetle	2	3	1	4
Spruce beetle	3	1	2	4
Pine engraver	2	3	1	4
Sudden aspen decline	2	4	1	3
Subalpine fir mortality	3	2	1	4
Western spruce budworm	2	3	1	4
Combined	2	3	1	4
Average	2.38	2.63	1.13	3.88

1. SRS-simple random sampling, PPS-probability proportional to size, NALIGN-nonalignment systematic sampling, STRA-stratified random sampling

Table 3.4. Statistical distribution of ratio of the variance of four sample designs associated with the eight causal agents and disorders. If the ratio of the variance falls within the interval from 0.98 to 1.02, the estimated variance is unbiased. If the ratio of the variance is greater than 1.02, the estimated variance is overestimated, and underestimated if the ratio is less than 0.98.

Causal and disorders[1]	SRS[2]						PPS					
	Min	1st Quartile	Median	Mean	3rd Quartile	Max	Min	1st Quartile	Median	Mean	3rd Quartile	Max
MPB	0.962	0.995	1.002	1.001	1.007	1.027	0.854	0.967	0.997	0.988	1.013	1.072
DF	0.971	0.993	0.999	1.000	1.006	1.031	0.946	0.996	1.005	1.007	1.013	1.141
SB	0.978	0.993	0.999	0.999	1.004	1.021	0.868	0.956	0.977	*0.974*	0.992	1.057
PE	0.978	0.995	0.999	0.999	1.005	1.018	0.896	0.999	1.014	1.019	1.033	1.182
SAD	0.973	0.993	0.999	1.000	1.006	1.023	0.912	0.981	0.994	0.995	1.008	1.097
SUB	0.971	0.994	1.000	1.000	1.005	1.031	0.886	0.985	1.000	0.996	1.010	1.063
WSB	0.971	0.993	0.999	1.000	1.008	1.035	0.908	0.978	0.994	0.999	1.008	1.198
Comb.	0.972	0.996	1.001	1.001	1.007	1.030	0.891	0.976	0.993	0.991	1.010	1.096
	NALIGN						STRA					
MPB	1.074	2.013	4.078	6.040	9.147	25.014	0.968	0.995	1.002	1.001	1.008	1.038
DF	1.095	1.474	1.741	2.079	2.156	10.442	0.974	0.995	1.001	1.002	1.009	1.032
SB	0.441	1.546	2.760	3.461	4.469	13.867	0.976	0.993	1.000	1.000	1.006	1.026
PE	0.572	1.114	1.300	1.654	1.739	8.176	0.973	0.994	1.000	1.000	1.006	1.023
SAD	0.887	1.353	1.567	1.796	2.042	4.807	0.975	0.994	1.000	1.000	1.007	1.023
SUB	0.657	1.953	2.747	3.422	4.078	24.964	0.972	0.993	1.000	1.000	1.007	1.023
WSB	0.422	1.592	3.184	3.695	4.716	15.430	0.978	0.994	1.000	1.001	1.007	1.029
Comb.	1.377	2.471	4.160	5.054	6.840	20.622	0.969	0.992	1.000	0.999	1.006	1.023

1. MPB-mountain pine beetle, DFB- Douglas fir beetle, SB-spruce beetle, PE-pine engraver, SAD-sudden aspen decline, SUB-subalpine fir mortality, WSB-western spruce budworm, Comb.-all causal agents and disorders combined.
2. SRS-simple random sampling, PPS-probability proportional to size, NALIGN-nonalignment systematic sampling, STRA-stratified random sampling

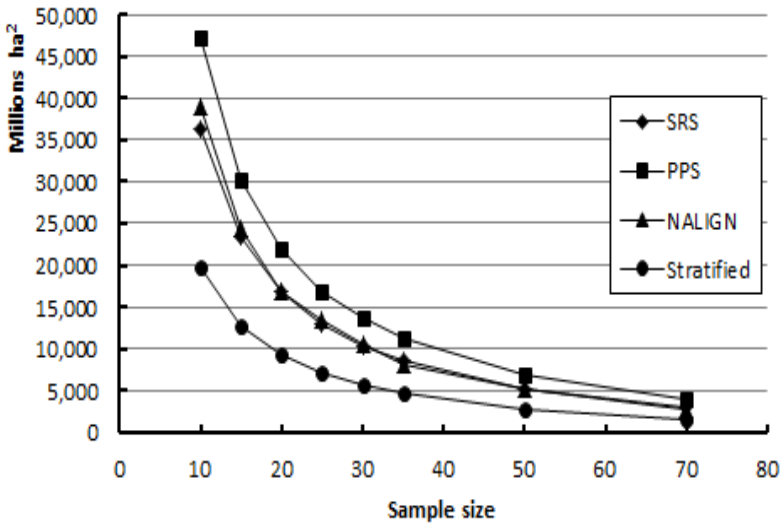


Figure 3.3. An example of estimated mean variance of area damaged caused by the combination of all causal agents and disorders associated with four different sample designs and eight sample sizes. Data presented for the year 2013. Where SRS-simple random sampling, PPS-probability proportional to size, NALIGN-nonalignment systematic sampling, Stratified-stratified random sampling.

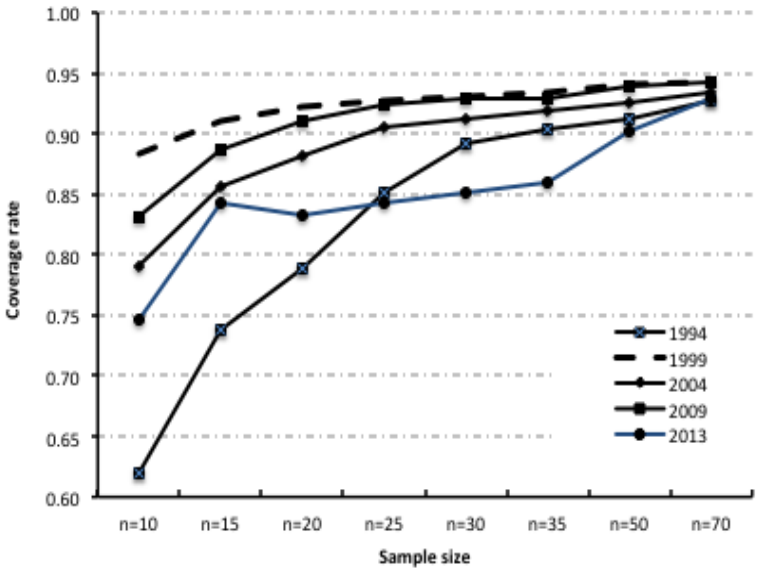


Figure 3. 4. Coverage rates for estimating the total area damaged by mountain pine beetle using stratified random sampling for five selected years. Similar trends were observed for the other causal and disorders agents and sample designs.

Discussion

As previous mention, the objective of this chapter was to determine the most appropriate sample design for estimating the total amount of area damaged caused by both single causal agents and disorders and multiple causal agents and disorders combined. Stratified random sampling was ranked number one in terms of producing unbiased and precise estimates of the total area damaged by all causal and disorders agents in each of the 20 years (1994-2013) evaluated in this study. This was followed by SRS, PPS and finally NALIGN. Since the NALIGN sample design consistently over-estimated the sample variance, it was not considered a viable sample design. While most sample designs produced unbiased estimates of the sample variance, coverage rates were less than the nominal rate of 0.95 especially when using smaller sample sizes. This suggests that the estimate of the total area damaged is more precise than they actually are.

One of the advantages of ADS is that it produces complete coverage of the damage caused by various causal and disorders agents. In contrast, sampling transects only produces a partial picture of damage occurring on the landscape. For survey sampling to become a practical alternative to the current ADS, it is important to be able to fill in the gaps between the transects selected for measurement. One approach would be to use satellite imagery to fill in these gaps. Statistical model could be developed relating what is observed on the transects selected for measurement and the spectral properties of the satellite imagery. The fitted models could then be used to interpolate between transects. The feasibility of this approach was demonstrated by Reich et al. (Reich et al. 2013). In their study Landsat imagery was used to estimate the area damage by aspen leaf miner in Alaska by using information contained on individual flight lines. For this approach to be feasible, it is important that the damage caused by various causal agents and

disorders be visible on the satellite imagery. If the damage cause by various causal agents and disorders are not visible on the satellite imagery it could not be possible to produce estimates of damage for all causal agents and disorders observed on the transects selected for measurement.

A second aspect of using survey sampling is how to sample overtime. With ADS it is simple process of subtracting one layer from another providing estimates of new damages continuing damage and areas were damage is no longer taking place. How this could be accomplished using survey sampling will depend on the objective of the survey. If the objective is to produce estimate in the current time period, the optimal approach is to select an independent sample in each time period. If the objective is to produce estimate of changes from one time period to the next, the optimal approach is to re-measure the same transects in each time period. If the objective is to produce estimate of change as well as current estimate, there is more flexibility in how the transects are selected for sampling. For example, one could use a form of double sampling. In this approach a subset of transects are selected for re-measurement and a new set of transects are selected for measurement, while keeping the sample size the same in each time period.

More research will be required to make survey sampling feasible on an operational basis that would meet of the current users of ADS data. This study demonstrated the feasibility of using survey sampling to estimate the damage caused by various causal agents and disorders. As the cost and risk of flying increases more focus will be placed on the use of survey sampling as a viable alternative to 100% surveys.

Summary

Overall, we found that three out of four sampling designs, simple random sampling, probability proportional to size, and stratified random sampling, were acceptable for estimating

damaged forest areas in Colorado using aerial survey parallel transects. All these sampling designs produced more than 90% unbiased estimate of total area damaged. These percentages of producing unbiased differed with different causal agents and disorders.

It is known that both efficient sample design of sample surveys and the evaluation of the precision of the estimates depend on a knowledge of the appropriate variances (and covariance or correlations) for the population from which the sample is drawn (Hansen et al. 1953). With respect to the smaller bias as well as smaller mean variance, the stratified random sampling design would be considered as the most appropriate for estimating total area damaged of forests in Colorado. In this study, most area damaged were clustered thus when using stratified random sampling we minimized variability within stratum while maximizing variability between each stratum resulting more precise estimates when comparing to other sampling designs. Because stratified random sampling provides more precise estimates than other sample designs, it suggested that using stratified random sampling would need fewer samples and therefore decreasing the cost and time. Following stratified random sampling, simple random sampling and probability proportional to size would be suggested. While stratified random sampling was the best sample design for estimating total area damaged of most causal agents and disorders, the convenience of employing simple random sampling and probability proportional to size to estimate total area damaged varied with different causal agents and disorders. These conclusions accompany the spatio-temporal distribution of area damaged caused by each individual causal agent, suggesting the strong effects of spatio-temporal distribution of area damaged on the potential of employing different sample designs.

CHAPTER 4: OPTIMAL SAMPLING STRATEGIES FOR AERIAL DETECTION SURVEYS USING ECONOMIC LOSS PLUS COST ANALYSIS

Introduction

Forests are damaged by biotic and abiotic disturbances of all kinds. Because forests grow over large heterogeneous landscapes, assessing the presence, severity and/or distribution of these disturbances and the damage they cause can be difficult. Furthermore, roadways and other pathways of access to forested areas are commonly sparse, and thorough pest assessments can be difficult to make. Aerial detection surveys enable assessments of inaccessible areas at a reasonably low cost and can cover vast areas in a relatively short time. According to a report in the 2009 Region 10 Forest Conditions Report (Lamb and Winton 2009), "no other method is currently available to detect subtle differences in vegetation damage signatures within a narrow time window at such low cost". As a consequence, aerial detection surveys (ADS) are used throughout the continental United States to provide essential information on the occurrence of insect pests, diseases and other forest disturbance agents (Johnson and Wittwer 2008). Information collected includes the spatial extent of the damage, the causal agent, and in some cases the level of mortality or the severity of the damage.

In Colorado, for example, ADS has been used to assess forest conditions since 1994 using a digital format. Colorado's forests have experienced significant changes over the past two decades (CSFS 2012), where "unprecedented mortality" in every major forest type "driven by poor resiliency to insects and diseases that have been exacerbated by warmer and drier weather conditions" (CSFS 2012). Major disturbance agents include:

- *Mountain pine beetle*: Until recently, the mountain pine beetle (*Dendroctonus ponderosae*) was the most widespread insect pest in the state. Approximately 1,380,000 ha of limber (*Pinus flexlis*), lodgepole (*Pinus contorta*) and ponderosa pine (*Pinus ponderosa*) forests have been impacted by this outbreak since it began in 1996. The epidemic reached a peak in 2009 affecting 423,300 ha and then began to decline due to the lack of available host trees. In 2013, only 39,700 ha of the mountain pine beetle mortality were mapped in Colorado's forests (CSFS 2013).

- *Spruce beetle*: The spruce beetle *(Dendroctonus rufipennis*) was the most widespread insect pest of Colorado's forests with increasing trends since it was first noticed in 1995. In 2012, active spruce beetle infestations were found on 132,000 ha and then increased to 161,000 ha in 2013 (CSFS 2013).

- *Subalpine fir mortality*: Subalpine fir decline occurs in the high-elevation spruce-fir (*Picea englmanii – Abies lasiocarpa*) forests across the state. Mortality is the result of attack by the western balsam bark beetle (*Dryocoetes confusus*). Attacks typically occur in trees infected and weakened by two species of fungi (*Armillaria* spp. and *Heterobasidion parviporum*) that attack the root systems of subalpine fir. A total of 72,000 ha of subalpine fir mortality were mapped in Colorado's forests in 2013 (CSFS 2013).

- *Sudden aspen decline*: Beginning in approximately 2004, many mature aspen (*Populus tremuloides*) forests in Colorado suddenly died off. In Colorado, approximately 220,000 ha of dead and dying aspen were mapped during aerial forest health surveys in 2008 (CSFS 2011). Since 2008, progressively smaller areas of SAD have been mapped each year: 138,500 ha in 2009; 77,000 ha in 2010; and only 18,600 ha in 2011. Several caterpillar species can defoliate aspen forests and can cause mortality with repeated episodes of defoliations. During outbreaks,

these caterpillars can cause complete defoliation, usually by mid to late June. In Colorado, two species of defoliating caterpillars can reach epidemic levels and cause widespread defoliation of aspen forests: western tent caterpillar (*Malacosoma californicum*) and large aspen tortix (*Choristoneura conflictana*).

Infested acre estimates for the 4 disturbance agents in Colorado listed above are based on complete surveys of, presumably, 100% of the landscape. Recently, the quality of aerial survey data have been questioned (Sapio 2012) as reports claim inflated infested acres, mistakenly low numbers of trees per acre (PA), uncertain spatial scales, and unclear damage trends due to 'less than optimum' planning and data inconsistencies. As a result, basic managerial questions cannot be answered, and survey results cannot be used to validate predictive pest and disease risk models (Sapio 2012). Furthermore, the Forest Service has recently intensified its effort for safety and "less time in the plane is equated with safety…" (Sapio 2012) As a consequence, the USDA Forest Service has focused a considerable energy on addressing these issues and improving aerial survey protocols.

In contrast to Colorado, some locations are much too big to be covered completely in a single season. Alaska is one such state. Alaska has around 51.4 million ha of forested area, of which approximately 13 million ha (or 25 %) are observed by aerial pest surveyors in any one year. Only infested areas actually seen by observers are reported, which leaves huge gaps of unsurveyed area between transects. In a recent study, predictive statewide spatial models of the distribution of aspen leaf miner (*Phyllocnistis populiella*) was over 20 times more abundant than what was observed by aerial surveyors and reported in the annual Forest Conditions Report (Lundquist pers. com.). Complete surveys like that conducted currently in Colorado may not be possible in the future. The question is how can ADS results be improved without increasing

flight time in an era of declining budgets and increasing sensitivity to safety? In this study, we examine sampling strategies to optimize aerial surveys using a method borrowed from economics called the loss plus cost analysis.

Loss plus cost analysis

Sample size calculations are rarely performed for aerial detection surveys. This may be due to uncertainty as to the type of causal agents and disorders one might encounter and the variability in size and spatial distribution of the damage caused by the various causal agents and disorders. In most forest inventories, sample sizes are selected that either minimize the variance of the sample estimates assuming constraints on the total cost of the survey or by minimizing the cost assuming constraints on the variance (Angelis and Stamatellos 2004). As an alternative, one can consider the solution to the sample size problem as a multiple objective optimization problem where the cost and the variance are simultaneously minimized (Angelis and Stamatellos 2004). This paper describes how economic loss functions plus the cost of the survey can be optimized to address these issues (Cochran 1977). The loss function places an economic value on the error associated with the survey. Increasing the sample size for a given sampling method will generally reduce the loss of information but at an increased cost. Selecting a sample size by managing the tradeoff between sampling cost and the cost of lost information can provide the minimum cost plus loss.

Loss functions can take on several forms such as a quadratic loss, absolute loss, step loss and the generalized loss of which the quadratic and absolute losses are special cases (Leung and Spiring 2004). Hamilton (1979) used an absolute loss function to determine the optimal sample size for individual timber sales. He assumed that the loss was proportional to the absolute error in the forest inventory. In the example cited, a loss of $100 was incurred for every thousand

board-feet the survey estimate of the total volume differed from the mill scale. Blithe (1945) used a similar approach to provide estimates of the number of logs to scale in a timber sale. In this study the loss was proportional to the sampling error. Such an approach can be used for any method of sampling and estimation in which the loss function is inversely proportional to the sample size and the cost a linear function of the sample size (Cochran 1977) (Figure 4.1).

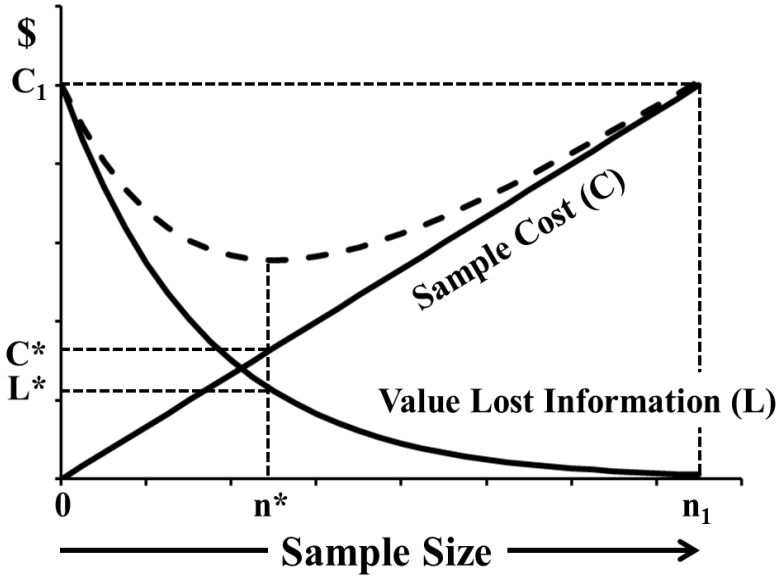


Figure 4.1. The value of lost information and costs for aerial surveys (Adopted from Freeman III et al. 1973).

In designing an aerial survey using this approach assume that there is only one agent causing damage on the landscape. Further assume that for a given sample design, the relationship between the sampling error and sample size is known and one can place dollar values on the loss of information based on the magnitude of the error associated with the survey. In Figure 4.1 this curve is labeled "Value Lost Information." These dollar values can be interpreted as the willingness of the users to minimize the errors associated with the aerial survey to avoid making incorrect decisions based on incomplete information on the damage being attributed to the causal agent. Figure 4.1 also depicts the sampling cost ($C), which is a function of the sample size

associated with the aerial survey. Moving to the right along the horizontal axis represents increasing sample sizes and increasing costs of the aerial survey. Conversely, moving to the left, sample sizes decrease and the value of the information lost ($L) increases due to increasing sampling errors associated with the aerial survey. Another way to look at this is to consider the gain in information associated with the survey. The gain (G) in information associated with the aerial survey can be defined as the total value a 100% aerial survey minus the loss ($G = C_1 - L$). If no survey is conducted ($n = 0$, the origin in Figure 4.1), the gain to the user will be zero, which is equivalent to a potential loss of C_1 dollars of information. If a 100% survey is conducted ($n = n_1$ in Figure 4.1), the loss will equal zero, and the user will gain by C_1 dollars of information.

The optimal sample size occurs where the total cost of the survey design is at a minimum, where the total cost of the survey design is defined as the vertical sum of the loss and sampling cost functions ($C + L$) (Freeman et al. 1973). The total cost of the survey design is depicted by the dashed line in Figure 4.1 and has its minimum at sample size n^*. Thus, n^* represent the optimal sample size that minimizes the value of lost information (L^*) due to incomplete information and the cost of conducting the survey (C^*). This point represents a cost savings of $C_1 - C^*$ by reducing the sample size from n_1 to n^*, and a gain in value of the information equal to $C_1 - L^*$.

To develop a framework to solve this problem there are several factors need to be taken into consideration. First is how much the government should invest in the aerial survey. Users of aerial detection survey information include land managers and forest health specialists from federal, state, and tribal agencies, private industry, and the public. Forest Health Protection, a part of the US Forest Service pays for the cost of the aerial surveys and does not charge the public or other state and federal agencies to access the data. Since the cost to the user is nothing,

the users of the aerial survey data will demand a 100% survey. On the other hand, the government should not be willing to spend more on the aerial survey than is necessary. Is a 100% survey necessary given the costs and high risks associated with conducting aerial surveys, especially in mountainous and remote regions? A reasonable attitude by the government would be to agree to spend as much on the aerial survey to produce estimates of damage for various causal agents and disorders with an appropriate and acceptable bound on the error of estimation.

Second, the number of causal agents and disorders and the spatial extent of the damage will also affect the solution. From the government's point of view large costs would be incurred if it was required to produce acceptable sampling errors for all causal agents and disorders. In this case the government might take the position that if the aerial survey produces an acceptable sampling error for estimating the total damage associated with all causal agents and disorders, the objective of the aerial survey has been met. However, this may not satisfy the users who might be interested in a particular causal agent. There may also be disagreement among users as to which causal agent should be used in designing the survey. A design that minimizes costs, for example, for one causal agent often is not the best design, or even adequate for other causal agents (Cochran 1977). Furthermore, since many management decisions are based on the outcome of the aerial surveys, reasonable accuracy for individual causal agents and disorders may be necessary. Therefore, in order to provide the most cost-efficient estimates it appears logical to assume that the problem must be solved for multiple causal agents and disorders.

Third, the method used to collect the aerial survey data should also be considered. There are numerous sample designs and methods from which to choose and each has its merits. Two aspects of sampling that have been evaluated for use with low-level aerial surveys is random sampling and stratification. The choice is influenced by many factors, such as the objective of

the survey, the spatial variability and extent of the damage associated with the population to be sampled, the number of causal agents and disorders to be measured and the availability of auxiliary information (Cochran 1977). Both the government and users of the information may not have a preference for a particular sample design, as long as it maximizes the value of the information collected at minimum cost. Thus, it is important to be able to evaluate the cost advantage of using one sample design over that of another design.

The purpose of this study is to 1) express the optimal sample size for aerial survey sampling in an economic framework, using the loss + cost approach, 2) calculate the cost savings of alternative sample designs (e.g., random sampling vs. stratification), and 3) demonstrate the use of an aggregated loss + cost function to select an optimal sample size that accounts for multiple objectives associated with multiple causal agents and disorders.

Material and Method

Sampling Frame

Before an area can be surveyed it must be divided into sample units that are exhaustive and non-overlapping (Caughley 1977). The four main types of sampling units used in aerial surveys are strip samples, quadrats or blocks, line intercepts, and line transects. Caughley (1977) points out that the choice of a sample unit for use in an aerial survey will involve trade-offs between "…maximizing safety and conditions of visibility while trying to minimize flying time, navigational problems and observer and pilot fatigue." There is general agreement in the literature that there are greater efficiency associated with using transects as sampling units when compared with quadrats (Caughley 1977).

In this study a strip sample or transect was chosen as the basic sample unit. The western half of Colorado was divided into 155 transects with a width and length of 3.2 km and 625 km,

respectively. Transects were oriented east-west to take into consideration solar light conditions while flying in mountainous terrain. Five of the transects did not contain any forestlands and were eliminated leaving 150 transects. The transects were numbered from 1 to 150 from south to north.

GIS Data

Two sources of GIS information were used to obtain the data used in this study. The first was a GIS layer of the major vegetation types of the state at a 30 m spatial resolution obtained from the Colorado Division of Wildlife as part of the Gap Analysis Program (http://ndis1.nrel.colostate.edu/cogap/cogaphome.html). The vegetation layer was used to create a binary raster layer indicating if a raster cell in the vegetation layer was classified as being forest or non-forest. This layer was intersected with the GIS layer of transects to obtain estimates of the area of forests and non-forests on each transect. Aerial Detection Survey (ADS) data were downloaded for the years 1994 to 2013 as GIS layers (http://www.fs.usda.gov/detail/r2/forest-grasslandhealth/?cid=fsbdev3_041629). These layers were intersected with the GIS layer of transects to obtain estimates of the total area impacted by all forest pests and area impacted by four causal agents and disorders of importance to the state: spruce beetle, mountain pine beetle, aspen mortality and subalpine-fir mortality.

Sample Designs

There are numerous sampling designs and methods from which to choose and each has its merits. The choice is influenced by many factors, such as the objectives of the survey, properties of the population to be sampled (e.g., spatial patterns and intensity of the damage), the number and type of causal agents and disorders to be measured, and the availability of auxiliary

information. In this study the discussion will be limited to three sample designs, two of which have particular interest in low-level aerial surveys, simple random sampling and stratified random sampling. The third design, which is a variation of random sampling was sampling with probability proportional to size.

Simple Random sampling.

Simple random sampling (SRS) is the most basic sample design in which a sample of size *n* is drawn from a population of size *N* in such a way that every possible sample of size *n* has the same chance (probability) of being selected. The selected sample units are surveyed and the area and cause of the damage recorded. For a given causal agent an estimate of the total damage in the state is given by

$$\hat{\tau} = N\bar{y} = N\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) \qquad [4.1]$$

with estimated variance

$$\hat{V}(\hat{\tau}) = N^2\left(\frac{N-n}{N}\right)\frac{s^2}{n} \qquad [4.2]$$

where $\bar{y}$ is the sample mean damage per transect and s^2 is the sample variance among transects.

Stratified random sampling (STRA)

As an alternative to random sampling, the state was divided into L = 5 regions or strata of equal size (N_i = 30 transects) and a simple random sample of n_i transects selected from each stratum. This form of stratification was selected to ensure sample units were distributed throughout the state and not necessarily to minimize the variability within the stratum. An estimate of the population total is given by:

$$\hat{\tau}_{STRA} = \sum_{i=1}^{L} N_i\bar{y}_i \qquad [4.3]$$

with estimated variance

$$\hat{V}(\hat{\tau}_{STRA}) = \sum_{i=1}^{n} N_i^2 \left(\frac{N_i - n_i}{N_i}\right) \frac{s_i^2}{n_i} \qquad [4.4]$$

Probability proportional to size (PPS)

Probability proportional to size (PPS) sampling includes a number of sample selection methods in which the probability of selecting a sample unit is directly proportional to a size measure, X which is known for all sample units and is approximately proportional to the value associated with the variable of interest (McGin, 2004). The increase in precision of PPS sampling over SRS will depend on the strength of the correlation between the selection probabilities (π_i) and the variable of interest (y_i). In this design, the selection probabilities were based on the proportion of the transect classified as being forested, thus more heavily forested transects had a higher probability of being selected than lightly forested transects. The use of sampling with replacement with PPS sampling has been widely adopted in sample surveys (Sirken 2001). In this study, the sampling procedure is changed so that the sample of transects was drawn without instead of with replacement (Sirken 2001). If the selection probabilities are known, an estimate of the population total is given by

$$\hat{\tau}_{pps} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{\pi_i} \qquad [4.5]$$

with estimated variance

$$\hat{V}(\hat{\tau}_{pps}) = \left(\frac{N-n}{N}\right) \frac{\sum_{i=1}^{n} \left(\frac{y_i}{\pi_i} - \hat{\tau}_{pps}\right)^2}{n(n-1)}. \qquad [4.6]$$

Exponential-with-nugget variance function

An exponential-with-nugget variance function was used to describe the influence of sample size on the estimated variance

$$\tilde{V}(\tau; n) = \begin{cases} \sigma^2 e^{-bn} & if \quad n > 0 \\ \delta^2 + \sigma^2 & otherwise \end{cases} \qquad [4.7]$$

where σ^2 is the partial sill, δ^2 is the nugget effect, $\delta^2 + \sigma^2$ is the sill and 1/b is the effective range (Irvine et al. 2007). The sill consists of the nugget effect, if present, and the partial sill represents the variability in damage from a causal agent on individual transects. The nugget represents the discontinuity at the origin and is typically attributed to micro-scale effects or measurement errors. In this application the nugget may be thought of as the efficiency of the sample design for sample sizes near zero. By definition the ratio, $\delta^2/\delta^2 + \sigma^2$ gives the relative magnitude of this effect, and is referred to in the literature as the relative nugget effect. For example, if the relative nugget effect for a given variance function is say 0.75, then this would indicate that the underlying sample design is relatively highly efficient. The partial sill may be thought of as the variance of the sample design near the origin. The ratio $\beta = \sigma^2/\delta^2 + \sigma^2$ can be defined as the relative partial sill of the sample design. The smaller this ratio the more efficient the sample design for sample sizes near zero. The range is the sample size at which the variance of the estimate levels off and begins to approaches zero. The effective range (1/b) is defined as the sample size beyond which the estimated variance is less than or equal to 5% of the sill: $0.05(\delta^2 + \sigma^2)$. For the exponential-with–nugget variance function the effective range (ξ) is defined as (Irvine et al. 2007):

$$\xi = \frac{-1}{b} ln\left(0.05 \frac{\delta^2+\sigma^2}{\sigma^2}\right) \qquad [4.8]$$

Variance estimation

To model the influence of sample size on the estimate of the variance, Monte Carlo procedures were used to simulate the three sample designs using eight samples of size: 10, 15, 20, 25, 30, 35, 50 and 70 transects. In the stratified design, an even number of transects were allocated to each of the five strata. Each sample size (n) – design (D) – causal agent (S) combination was simulated M = 20,000 times and the mean variance computed:

$$\tilde{V}(\hat{\tau}_{DS}; n) = \frac{1}{M}\sum_{i=1}^{M} \hat{V}(\hat{\tau}_{DSi}) \qquad [4.9]$$

Optimal sample size

Let C = c_1 * N equal the total amount of money the government is willing to spend on the aerial survey assuming a 100% survey, where c_1 is the cost of an individual transect and N is the total number of transects in the state. Further assume that the loss (L) of information, expressed in dollars is proportional to the relative standard error of the survey:

$$L = C * (s_{\hat{\tau}}/\hat{\tau}) \qquad [4.10]$$

where $s_{\hat{\tau}} = \sqrt{\hat{V}(\hat{\tau})}$, is the standard error of the total damage on the transects, and $\hat{\tau}$ is an estimate of the total damage in the state. As the sample size, *n*, increases the standard error of the total will decrease, but not necessarily in accordance to the Central Limit Theorem especially for sample designs other than simple random sampling. So as an alternative we propose to use the exponential-with-nugget model to describe this relationship.

$$\hat{V}(\hat{\tau}; n) = ae^{-bn}, \quad n > 0 \qquad [4.11]$$

where the parameters $a = \sigma^2$ is partial sill and b are as previously defined. Taking the square root of both sides and then dividing by the estimated total we get:

$$\sqrt{\hat{V}(\hat{\tau};n)}/\hat{\tau} = \frac{\sqrt{a}}{\hat{\tau}} e^{-bn/2} \qquad [4.12]$$

Simplifying, we see that the relative standard error (*rse*) has an exponential distribution.

$$s_{\hat{\tau}}/\hat{\tau} = a^* e^{-b^* n} \qquad [4.13]$$

where $a^* = \sqrt{a}/\hat{\tau}$ and b^* = b/2.

The loss function (L) then becomes

$$L(n) = \begin{cases} C, & n = 0 \\ Ca^* e^{-b^* n}, & 0 < n < N \\ 0, & n = N \end{cases} \qquad [4.14]$$

Now define the cost of the survey to be c_1 * n, where n is the sample size. The loss + cost function then becomes

$$LC(n) = Ca^* e^{-b^* n} + c_1 n \qquad [4.15]$$

To estimate the optimal sample size (n*) that will minimize the loss + cost function we take the first derivative with respect to the sample size. Setting this equal to 0 and solving for the sample size we get:

$$n^* = \frac{-1}{b^*} ln\left(\frac{c_1}{Ca^* b^*}\right) \qquad [4.16]$$

Noting that $C = c_1 * N$ we get

$$n^* = \frac{-1}{b^*} ln\left(\frac{1}{Na^* b^*}\right) \qquad [4.17]$$

Incorporating the Level of Confidence in the Loss Function

Any survey takes a sample from the population of interest and then generalizes the results to the whole population. This invariably leads to a possibility of an error because the entire population was not surveyed. This is captured in statistics as the bound on the error of estimation or the margin of error. The larger the bound on the error of estimation, the less likely the results of the survey reflect the true population value. In statistics the bound on the error of estimation (B) is equal to half the length of the confidence interval:

$$B = t_{n-1;1-\alpha/2} s_{\hat{\tau}}. \qquad [4.18]$$

where *t* is a student-*t* value with *n*-1degrees of freedom, with significance level α. This means that as the desired level of confidence increases (that is as α becomes smaller), the larger the bound on the error of estimation becomes for the same set of data. The bound on the error of estimation for an individual survey will also depend to a large extent on the sample size. Larger sample sizes tend to give more precise estimates than estimates based on smaller sample sizes. The bound on the error of estimation will also depend on the variability of damage on the individual transects.

The bound on the error of estimation is usually expressed as an absolute number, but in some cases, may also be expressed relative to the sample mean or total. To take into consideration the level of confidence associated with the estimates from the aerial survey, the relative bound on the error of estimation can be defined as

$$B_{rel} = t_{n-1;1-\alpha/2}(s_{\hat{\tau}}/\hat{\tau}) \qquad [4.19]$$

The loss + cost function becomes

$$Cta^{*}e^{-b^{*}n} + c_1 n \qquad [4.20]$$

with optimal sample size:

$$n^{*} = \frac{-1}{b^{*}} ln\left(\frac{1}{Nta^{*}b^{*}}\right) \qquad [4.21]$$

where t = $t_{n-1;1-\alpha/2}$.

Cost Advantage of Alternative Sample Designs

The cost advantage of one sample design over that of another sample design can be evaluated by subtracting the loss + cost (*LC*) functions associated with the different sample designs, *D*, evaluated at their respective optimal sample size, n^{*}:

$$LC\left(D_{1,}n_1^{*}\right) - LC\left(D_2, n_2^{*}\right) \qquad [4.22]$$

Positive values would indicate that sample design D_2 with sample size n_2^{*} has a cost advantage over using a sample of size n_1^{*} with sample design D_1. A negative value would indicate that sample design D_2 with sample size n_2^{*} has a cost disadvantage over sample design D_1 using a sample of size n_1^{*}.

Loss + Cost Function for Multiple Causal Agents

Consider that each causal agent represents an individual product with a unique set of users. In other words, the aerial surveys are designed for individual causal agents and disorders. If one can assume that the information provided by the aerial survey is a public good, in that

there is no limitation on who can access the information, and how they use the information. Under this scenario, the total loss function is obtained by the vertical sum of the individual loss functions for the various causal agents and disorders of interest (S).

$$LC(S,n) = \sum_{i=1}^{S} LC(S_i,n) + c_1 n \qquad [4.23]$$

Results and Discussion

Variation in transects

The variability of damage caused by the various causal agents and disorders in a given transect is the determining factor in the feasibility of using aerial survey sampling. Taking Colorado as an example the amount of damage on a given transect varied of 0 to 200,000 ha depending on the causal agent and year of sampling. The coefficient of variation associated with the damage on an individual transect varied considerably from one year to the next and from one causal agent to the next and ranged from 0.7 to 12.0. Thus, it would be very difficult to provide general guidelines outlining the number of transects that need to be flown in order to achieve a desired bound on the error of estimation for the individual causal agents and disorders in any given year.

Figure 4.2 displays the relationship between the coefficient of variation and Moran's I (Moran 1950), a measure of the spatial autocorrelation associated with the damage observed on the individual transects. The highest variability was observed when there was little to no spatial autocorrelation of damage among the transects. This variability decreased at an increasing rate as the spatial autocorrelation of damage on the transects increased. This suggests that causal agents and disorders that exhibit large contiguous patterns across the landscape are more efficiently estimated than causal agents and disorders that occur randomly on the landscape.

Fitting the Exponential-with-Nugget Variance Model

The exponential-with-nugget variance model provided an exceptional fit for describing the relationship between the relative standard error associated with estimating the total amount of damage and sample size for all causal agents and disorders and sample designs evaluated over the 20 year time period. The three sample designs had a median FIT statistic of 0.997, where FIT is defined as the correlation between the observed and fitted values squared (Table 4.1). PPS sampling had the most consistent fit (range in FIT = 0.0013), followed by simple random sampling (range in FIT = 0.0139) and then stratified random sampling (range in FIT = 0.0184). Figure 4.3 provides an example of the fitted exponential-with-nugget variance models for estimating the relative sampling error associated with estimating the total amount of damage due to all causal agents and disorders for each of the 20 years using simple random sampling.

The Effective Sample Size

Except for two observations, simple random sampling had an effective sample size of 115 transects (Figure 4.4, Table 4.2). This represents the sample size when the relative sampling error equals 5% of the coefficient of variation in the population. The relationship between the relative sampling errors and sample size closely followed what would be expected under the Central Limit Theorem which may explain the consistency in the estimates of the effective sample size. The relative sampling errors associated with stratified random sampling approached zero faster than simple random sampling with a median effective sample size of 106 transects (Figure 4.4, Table 4.2). Stratified random sampling had a minimum effective sample size of 16 transects and a maximum of 114 fight lines. The effective sample size for PPS sampling had characteristics similar to that of simple random sampling but exhibited more variability around the effective sample size of 115 transects (Figure 4.4, Table 4.2). This variability was due to the strength and

sign of the correlation between the amount of damage on the transects and the proportion of the transect that was forested. In general, the effective sample size was less than 115 transects when there was a strong positive correlation between the amount of damage and the proportion of forested lands on the transects; equaled 115 when there was no correlation, and greater than 115 light lines when there was a negative correlation.

Table 4.1. FIT statistics of the fitted exponential-with-nugget variance models used to describe the relationship between the relative sampling error and sample size for estimating the area damaged by various causal agents and disorders in an aerial survey for the years 1994 to 2013 for the three sample designs. Causal agents and disorders include mountain pine beetle, spruce beetle, aspen mortality, subalpine-fir mortality and all causal agents and disorders combined. The FIT statistic is defined as the correlation between the observed and predicted values squared.

Sample Design	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Simple Random	0.9833	0.9969	0.9969	0.9968	0.9970	0.9972
Stratified	0.9788	0.9964	0.9970	0.9953	0.9970	0.9972
PPS	0.9958	0.9965	0.9969	0.9968	0.9970	0.9971

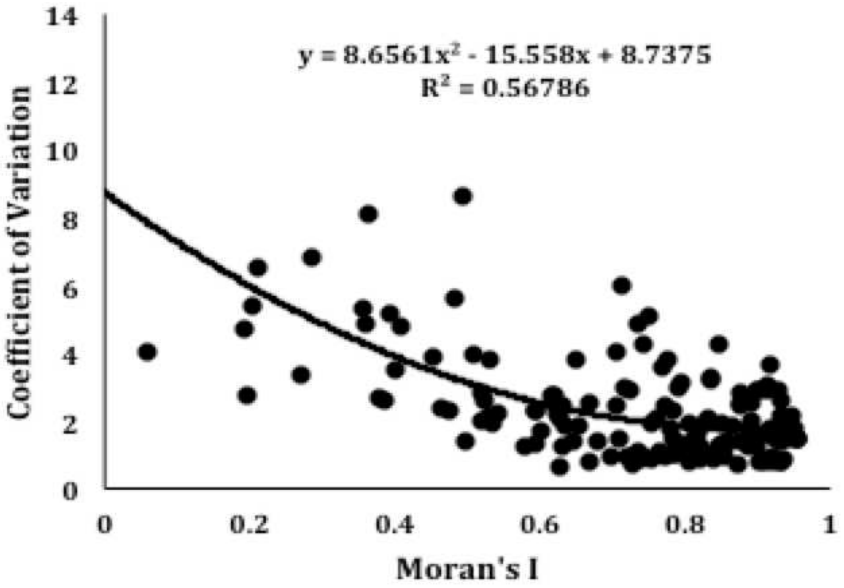


Figure 4.2. Relationship between the coefficient of variation and Moran's I associated with the area of damage on individual transects for the years 1994 to 2013. and disorders include mountain pine beetle, spruce beetle, aspen mortality, subalpine-fir mortality and all causal agents and disorders combined.

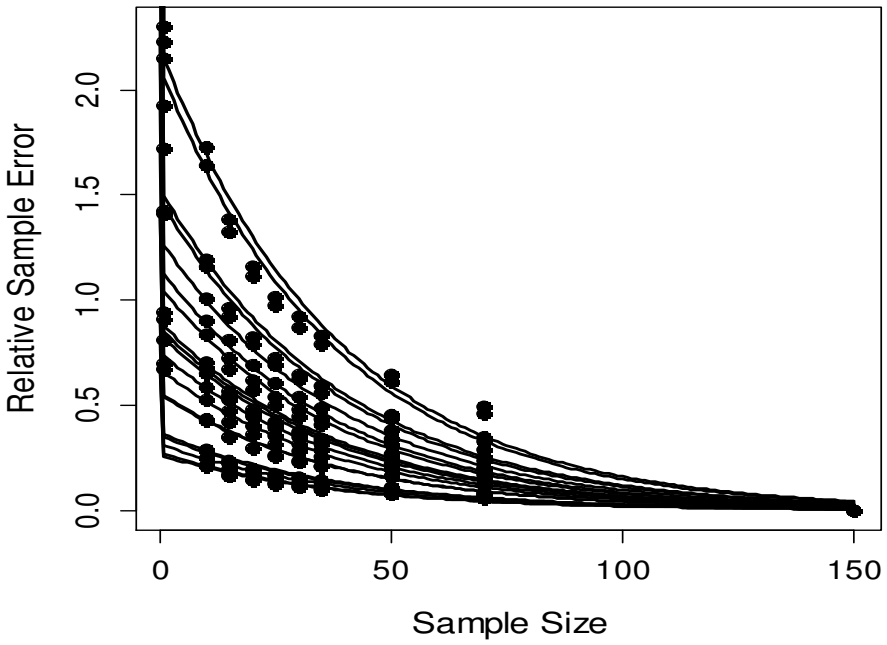


Figure 4.3. Example of the fitted variance-with-nugget model used to describe the relationship between the relative sampling error and sample size associated with estimating the total area of damage of all causal agents and disorders for the years 1994 to 2013. The points are the observed relative sampling errors for sample sizes, n=0, 10, 15, 20, 25, 30, 35, 50 and 70.

The Relative Partial Sill

The relative partial sill, β of the exponential-with-nugget variance model provides an estimate of the relative sampling error for a sample size near the origin. The smaller this parameter the greater the initial influence the sample design has on the precision of the estimates. In the case of simple random sampling the relative partial sill ranged between 0.38 and 0.39 with one extreme value of 0.46 (Figure 4.5, Table 4.3). The trend in the relative partial sill is consistent with the trend observed for the effective sample size. Stratified random sampling had relative partial sills less than or equal to the median relative partial sill for simple random sampling (Figure 4.4, Table 4.3). The relative partial sills ranged from 0.33 to 0.39 with a median value of 0.36. There was one extreme value with a relative partial sill of 1 (not shown in Figure 4). The relative partial sill for PPS sampling had characteristics similar to that of simple random sampling but exhibited more variability around the relative partial sill of 0.39 (Figure 4.5, Table 4.3). The relative partial sills ranged from 0.36 to 0.46 with a median value 0.40. The relationship between the relative partial sill and the correlation between the amount of damage on the transects and the proportion of forested lands on the transects was opposite of that observed for the effective sample size. PPS sampling was more effective in reducing the relative partial sill than simple random sampling when there was a positive correlation and less effective when there was a negative correlation.

Loss + Cost Functions

Optimal Sample Size For Different Levels of Confidence

The optimal sample size that minimizes the loss + cost of estimating the total area of damage of all causal agents and disorders for the years 1994 to 2013 for different levels of confidence is summarized in Table 4.4. Sample sizes are based on the assumption of a simple

random sample of transects. Increasing variability in the population as measured by the coefficient of variation required a larger sample size to minimize the loss + cost function. The optimal sample size also increased as the desired level of confidence (t-statistic) associated with the outcome of the aerial survey increased (Table 4.4). Similar trends were observed for stratified random sampling (Table 4.5) and PPS sampling (Table 4.6). At the lower levels of confidence, there was little difference between the optimal sample sizes for simple random and stratified random sampling. This difference increased as the desired level of confidence increased. At the 0.95 level of confidence stratified random sampling required 6 fewer transects than simple random sampling to minimize the loss + cost function, on the average. The optimal sample sizes for PPS sampling fluctuated around the optimal samples sizes for simple random sampling. Again the correlation between the amount of damage and the proportion of forested lands on the transects influenced whether the optimal sample size was larger or smaller than that for simple random sampling.

Optimal Sample Size for Multiple Causal Agents

Optimal sample size that minimizes the loss + cost of estimating the area of damage associated with four causal agents and disorders both individually and simultaneously for the years 1994 to 2013 at the 0.67 level of confidence (t=1) is presented in Table 4.7. Sample sizes are based on the assumption of a simple random sample of transects. Under this scenario aerial surveys are designed specifically for a given causal agent. If the user groups are willing to participate in a single aerial survey, it will require larger sample sizes at the 0.67 level of confidence than the individual causal agents and disorders at the 0.95 level of confidence to minimize the vertical sum of the individual loss functions plus the cost of the survey. Stratified random sampling (Table 4.8) required 11 fewer transects than simple random sampling while

PPS sampling (Table 4.9) required 5 additional transects on the average. The optimal sample size required to minimize the vertical sum of the individual loss functions + costs was 1.6 times larger than the largest sample size observed for an individual causal agent irrespective of the sample design.

Table 4.2. Distribution of the estimated effective sample size for three sample designs used to estimate the total area of damage using aerial survey for the years 1994 to 2013. Causal agents and disorders include mountain pine beetle, spruce beetle, aspen mortality, subalpine-fir mortality and all causal agents and disorders combined. The effective sample size is defined as the sample size that yields a relative sampling error equal to 0.05 * CV.

Sample Design	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Simple Random	87.7	115.1	115.3	115.4	115.6	152.9
Stratified	16.0	89.8	105.4	93.6	109.9	113.9
PPS	107.6	111.6	116.9	116.9	121.4	129.2

Table 4.3. Distribution of the estimated relative sampling error for sample sizes near zero for three sample designs used to estimate the total area of damage using aerial survey for the years 1994 to 2013. Causal agents and disorders include mountain pine beetle, spruce beetle, aspen mortality, subalpine-fir mortality and all causal agents and disorders combined. The relative sampling error for sample sizes near zero is defined as $(1- \delta^2/CV)$, where δ^2 is the nugget effect.

Sample Design	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Simple Random	0.3839	0.3905	0.3910	0.3917	0.3916	0.4610
Stratified	0.3326	0.3504	0.3631	0.4114	0.3782	1.000
PPS	0.3560	0.3765	0.3960	0.3998	0.4190	0.4613

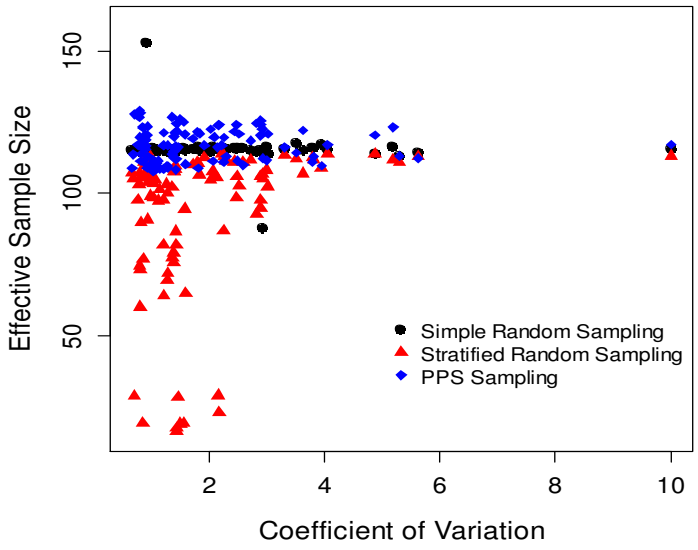


Figure 4.4. The estimated effective sample size plotted against the coefficient of variation for three sample designs used to estimate the total area of damage using aerial survey for the years 1994 to 2013. Causal agents and disorders include mountain pine beetle, spruce beetle, aspen mortality, subalpine-fir mortality and all causal agents and disorders combined. The effective sample size is defined as the sample size that yields a sampling error equal to 0.05 * CV.

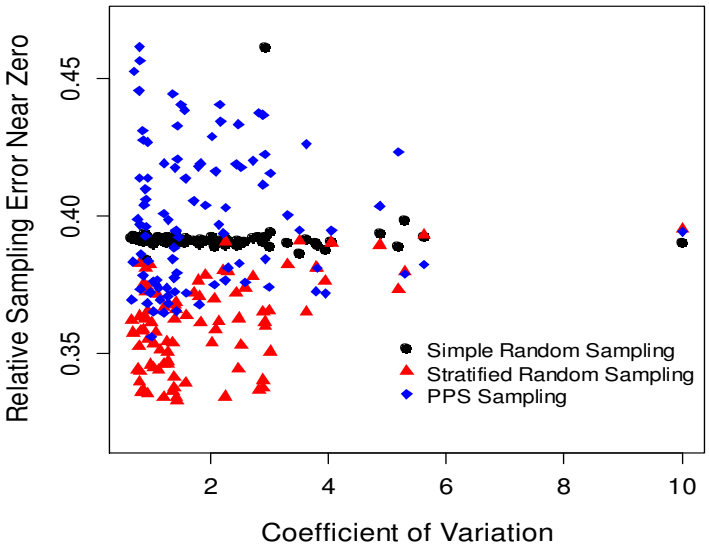


Figure 4.5. The estimated relative sampling error for sample sizes near zero plotted against the coefficient of variation for three sample designs used to estimate the total area of damage using aerial survey for the years 1994 to 2013. Causal agents and disorders include mountain pine beetle, spruce beetle, aspen mortality, subalpine-fir mortality and all causal agents and disorders combined. The relative sampling error for sample sizes near zero is defined as $(1- \delta^2/CV)$, where δ^2 is the nugget effect.

Table 4.4. Optimal sample size that minimizes the loss + cost of estimating the total area of damage of all causal agents and disorders for the years 1994 to 2013 for different levels of confidence. Sample sizes are based on the assumption of a simple random sample of transects (N=150 transects).

Year	CV[†]	Level of Confidence			
		t=1.0 (0.67)	t=1.3 (0.80)	t=1.6 (0.90)	t=2.0 (0.95)
1994	2.07	45	55	63	71
1995	3.00	59	69	77	86
1996	0.90	13	23	31	40
1997	1.40	30	40	48	56
1998	1.01	17	27	35	44
1999	0.94	14	24	32	41
2000	1.04	18	28	36	45
2001	0.82	9	19	27	36
2002	1.42	30	41	49	57
2003	0.86	11	21	29	38
2004	1.43	31	41	49	57
2005	0.78	7	17	25	34
2006	0.82	9	19	27	36
2007	0.81	9	19	27	36
2008	0.86	11	21	29	38
2009	0.71	4	14	22	30
2010	0.81	9	19	27	35
2011	1.37	29	39	47	56
2012	0.80	8	18	26	35
2013	0.94	15	25	33	41
Average		18	29	37	46

[†]CV – coefficient of variation

Table 4.5. Optimal sample size that minimizes the loss + cost of estimating the total area of damage of all causal agents and disorders for the years 1994 to 2013 for different levels of confidence. Sample sizes are based on the assumption of a stratified random sample with an equal number of transects in each of L = 5 strata (N=150 transects).

Year	CV[†]	Level of Confidence			
		t=1.0 (0.67)	t=1.3 (0.80)	t=1.6 (0.90)	t=2.0 (0.95)
1994	2.07	42	52	59	67
1995	3.00	55	65	72	80
1996	0.90	12	21	29	37
1997	1.40	27	37	44	52
1998	1.01	16	25	32	40
1999	0.94	14	23	31	39
2000	1.04	17	26	33	41
2001	0.82	8	18	25	33
2002	1.42	28	38	45	54
2003	0.86	14	21	26	32
2004	1.43	27	34	40	47
2005	0.78	7	16	22	30
2006	0.82	13	20	25	30
2007	0.81	17	22	27	31
2008	0.86	19	21	22	24
2009	0.71	20	23	25	27
2010	0.81	13	19	24	30
2011	1.37	26	32	38	43
2012	0.80	7	16	24	29
2013	0.94	14	22	28	33
Average		19	28	34	40

[†]CV – coefficient of variation

Table 4.6. Optimal sample size that minimizes the loss + cost of estimating the total area of damage of all causal agents and disorders for the years 1994 to 2013 for different levels of confidence. Sample sizes are based on the assumption of sampling with probability proportional the proportion of forested lands on each transect (N=150 transects).

Year	CV[†]	Level of Confidence			
		t=1.0 (0.67)	t=1.3 (0.80)	t=1.6 (0.90)	t=2.0 (0.95)
1994	2.07	43	53	60	69
1995	3.00	57	66	74	82
1996	0.90	13	23	31	40
1997	1.40	27	38	46	55
1998	1.01	15	24	32	40
1999	0.94	13	23	30	38
2000	1.04	17	27	35	43
2001	0.82	9	18	26	34
2002	1.42	31	41	49	58
2003	0.86	13	24	32	41
2004	1.43	29	38	45	54
2005	0.78	7	18	26	34
2006	0.82	9	19	27	36
2007	0.81	12	23	32	42
2008	0.86	13	24	32	42
2009	0.71	6	17	26	36
2010	0.81	12	23	32	42
2011	1.37	33	44	53	63
2012	0.80	10	22	30	40
2013	0.94	16	27	36	45
Average		20	30	38	47

[†]CV – coefficient of variation

Table 4.7. Optimal sample size that minimizes the loss + cost of estimating the area of damage associated with four causal agents and disorders both individually and simultaneously for the years 1994 to 2013 at the 0.67 level of confidence (t=1). Sample sizes are based on the assumption of a simple random sample of transects (N=150 transects).

Year	Mountain Pine Beetle		Spruce Beetle		Subalpine-fir Mortality		Aspen Mortality		All Four Agents
	CV[†]	n	CV	n	CV	n	CV	n	n
1994	2.52	52							52
1995	2.59	53	2.26	48	5.19	80	5.30	81	121
1996	2.26	48	4.06	71	1.29	27	1.93	42	103
1997	1.81	40	10.01	105	2.49	52	2.30	49	124
1998	1.37	29	3.52	65	1.29	27	1.72	38	96
1999	1.09	20	4.88	78	1.39	29	2.94	58	106
2000	1.16	22	3.81	68	1.29	27	5.62	83	111
2001	1.11	21	2.44	51	1.39	30	2.23	48	92
2002	0.99	16	2.90	58	1.27	26	3.80	68	101
2003	1.22	24	2.83	57	1.22	25	3.31	63	99
2004	1.60	35	2.48	52	1.21	24	2.15	46	93
2005	1.44	31	2.90	58	0.92	13	1.41	30	89
2006	1.47	32	2.10	45	0.83	10	0.94	15	81
2007	1.57	34	2.93	57	0.88	12	0.91	13	83
2008	1.45	31	3.02	59	0.95	15	0.70	3	86
2009	1.44	31	3.63	66	0.91	3	0.67	2	91
2010	1.51	33	2.03	44	0.86	13	0.81	9	80
2011	2.17	46	1.80	39	0.99	17	1.42	31	88
2012	2.18	47	1.84	40	1.05	18	3.95	70	101
2013	2.90	58	1.58	34	1.13	22	2.73	55	98
Average		36		58		25		42	95

[†]CV – coefficient of variation

Table 4.8. Optimal sample size that minimizes the loss + cost of estimating the area of damage associated with four causal agents and disorders both individually and simultaneously for the years 1994 to 2013 at the 0.67 level of confidence (t=1). Sample sizes are based on the assumption of a stratified random sample with an equal number of transects in each of L = 5 strata (N=150 transects).

Year	Mountain Pine Beetle		Spruce Beetle		Subalpine-fir Mortality		Aspen Mortality		All Four Agents
	CV[†]	n	CV	n	CV	n	CV	n	n
1994	2.52	47							47
1995	2.59	51	2.26	40	5.19	77	5.30	78	114
1996	2.26	48	4.06	70	1.29	24	1.93	41	100
1997	1.81	38	10.01	105	2.49	48	2.30	47	121
1998	1.37	26	3.52	64	1.29	25	1.72	36	88
1999	1.09	18	4.88	77	1.39	26	2.94	53	98
2000	1.16	20	3.81	66	1.29	25	5.62	82	105
2001	1.11	20	2.44	49	1.39	26	2.23	46	85
2002	0.99	16	2.90	49	1.27	24	3.80	66	93
2003	1.22	22	2.83	48	1.22	25	3.31	62	86
2004	1.60	30	2.48	45	1.21	22	2.15	43	77
2005	1.44	27	2.90	54	0.92	13	1.41	28	80
2006	1.47	27	2.10	41	0.83	10	0.94	12	62
2007	1.57	23	2.93	54	0.88	11	0.91	13	71
2008	1.45	21	3.02	53	0.95	13	0.70	3	68
2009	1.44	20	3.63	62	0.91	13	0.67	1	75
2010	1.51	23	2.03	40	0.86	10	0.81	8	61
2011	2.17	31	1.80	38	0.99	15	1.42	29	68
2012	2.18	28	1.84	37	1.05	17	3.95	66	84
2013	2.90	50	1.58	30	1.13	19	2.73	54	87
Average		30		54		24		41	84

[†]CV – coefficient of variation

Table 4.9. Optimal sample size that minimizes the loss + cost of estimating the area of damage associated with four causal agents and disorders both individually and simultaneously for the years 1994 to 2013 at the 0.67 level of confidence (t=1). Sample sizes are based on the assumption of sampling with probability proportional the proportion of forested lands on each transect (N=150 transects).

Year	Mountain Pine Beetle		Spruce Beetle		Subalpine-fir Mortality		Aspen Mortality		All Four Agents
	CV[†]	n	CV	n	CV	n	CV	n	n
1994	2.52	56							56
1995	2.59	51	2.26	46	5.19	86	5.30	79	122
1996	2.26	50	4.06	72	1.29	27	1.93	43	105
1997	1.81	37	10.01	107	2.49	51	2.30	48	124
1998	1.37	28	3.52	65	1.29	25	1.72	39	95
1999	1.09	19	4.88	81	1.39	29	2.94	57	107
2000	1.16	21	3.81	67	1.29	25	5.62	81	108
2001	1.11	20	2.44	55	1.39	28	2.23	48	93
2002	0.99	16	2.90	64	1.27	25	3.80	65	101
2003	1.22	23	2.83	63	1.22	27	3.31	64	103
2004	1.60	33	2.48	57	1.21	25	2.15	47	96
2005	1.44	30	2.90	61	0.92	14	1.41	32	92
2006	1.47	32	2.10	48	0.83	9	0.94	16	83
2007	1.57	39	2.93	63	0.88	13	0.91	14	93
2008	1.45	34	3.02	63	0.95	14	0.70	3	89
2009	1.44	35	3.63	72	0.91	14	0.67	1	94
2010	1.51	37	2.03	48	0.86	11	0.81	10	85
2011	2.17	52	1.80	42	0.99	16	1.42	31	92
2012	2.18	52	1.84	43	1.05	17	3.95	66	102
2013	2.90	64	1.58	37	1.13	20	2.73	59	103
Average		36		61		25		42	100

[†]CV – coefficient of variation

Cost Advantage of Sample Designs

Stratified random sampling had an $81,766 cost advantage over simple random sampling for estimating the total area of damage of all causal agents and disorders when averaged over the years 1994 to 2013 at the 0.95 level of confidence (Table 4.10). In general, the cost advantage of stratified random sampling over simple random sampling increased as the variability in the population increased. For example, the cost advantage of stratified sampling over random sampling for estimating the damage from subalpine-fir mortality was $85,318 (Table 4.11). The average coefficient of variation in the damage due to subalpine-fir mortality was 1.19 compared to an average coefficient of variation of 0.98 for the total damage due to all causal agents and disorders. In contrast, PPS sampling had a cost disadvantage of -$13,082 when compared with simple random sampling (Table 4.12) and -$94,848 when compared with stratified sampling (Table 4.13) for estimating the total damage from all causal agents and disorders at the 0.95 level of confidence.

Table 4.10. Cost advantage of stratified random sampling over simple random sampling for estimating the total area of damage of all causal agents and disorders for the years 1994 to 2013 at the 0.95 level of confidence (t=2).

Year	CV†	Simple Random	Stratified	Cost Advantage ($)
		n	n	
1994	2.07	71	67	45,352
1995	3.00	86	80	59,070
1996	0.90	40	37	35,124
1997	1.40	56	52	43,004
1998	1.01	44	40	42,834
1999	0.94	41	39	22,242
2000	1.04	45	41	50,990
2001	0.82	36	33	38,059
2002	1.42	57	54	33,729
2003	0.86	38	32	101,955
2004	1.43	57	47	123,272
2005	0.78	34	30	56,499
2006	0.82	36	30	102,589
2007	0.81	36	31	113,529
2008	0.86	38	24	244,648
2009	0.71	30	27	143,135
2010	0.81	35	30	96,818
2011	1.37	56	43	152,840
2012	0.80	35	31	46,809
2013	0.94	41	35	82,819
Average		45	41	81,766

†CV – coefficient of variation

Table 4.11. Cost advantage of stratified random sampling over simple random sampling for estimating the area of damage associated with subalpine-fir mortality for the years 1994 to 2013 at the 0.95 level of confidence (t=2).

Year	CV†	Simple Random	Stratified	Cost Advantage ($)
		n	n	
1994				
1995	5.19	107	103	40,990
1996	1.29	53	47	70,652
1997	2.49	79	72	66,773
1998	1.29	53	41	154,954
1999	1.39	56	44	145,427
2000	1.29	53	41	159,737
2001	1.39	56	44	152,341
2002	1.27	53	47	61,786
2003	1.22	51	39	165,985
2004	1.21	51	41	126,063
2005	0.92	40	38	23,489
2006	0.83	36	31	76,149
2007	0.88	39	37	20,308
2008	0.95	41	37	47,596
2009	0.91	38	39	47,667
2010	0.86	40	35	56,864
2011	0.99	43	38	63,415
2012	1.05	45	40	66,844
2013	1.13	48	42	73,999
Average		49	46	85,318

†CV – coefficient of variation

Table 4.12. Cost advantage of PPS sampling over simple random sampling for estimating the total area of damage associated with all causal agents and disorders for the years 1994 to 2013 at the 0.95 level of confidence (t=2).

Year	CV†	Simple Random	PPS	Cost Advantage ($)
		n	n	
1994	2.07	71	69	26,281
1995	3.00	86	82	37,134
1996	0.90	40	40	-1,102
1997	1.40	56	55	2,826
1998	1.01	44	40	40,491
1999	0.94	41	38	30,768
2000	1.04	45	43	20,581
2001	0.82	36	34	20,390
2002	1.42	57	58	-7,040
2003	0.86	38	41	-37,204
2004	1.43	57	54	37,091
2005	0.78	34	34	-5,257
2006	0.82	36	36	-4,332
2007	0.81	36	42	-67,555
2008	0.86	38	42	-43,442
2009	0.71	30	36	-63,103
2010	0.81	35	42	-74,362
2011	1.37	56	63	-73,511
2012	0.80	35	40	-56,849
2013	0.94	41	45	-43,440
Average		46	47	-13,082

†CV – coefficient of variation

Table 4.13. Cost advantage of PPS over stratified random sampling for estimating the total area of damage associated with all causal agents and disorders for the years 1994 to 2013 at the 0.95 level of confidence (t=2).

Year	CV†	Stratified	PPS	Cost Advantage ($)
		n	n	
1994	2.07	67	69	-19,072
1995	3.00	80	82	-21,936
1996	0.90	37	40	-36,226
1997	1.40	52	55	-40,178
1998	1.01	40	40	-2,343
1999	0.94	39	38	8,526
2000	1.04	41	43	-30,410
2001	0.82	33	34	-17,669
2002	1.42	54	58	-40,769
2003	0.86	32	41	-139,159
2004	1.43	47	54	-86,182
2005	0.78	30	34	-61,756
2006	0.82	30	36	-106,921
2007	0.81	31	42	-181,083
2008	0.86	24	42	-288,089
2009	0.71	27	36	-206,237
2010	0.81	30	42	-171,180
2011	1.37	43	63	-226,352
2012	0.80	31	40	-103,658
2013	0.94	35	45	-126,259
Average		41	47	-94,848

†CV – coefficient of variation

Summary

In this study, we present a cost + loss method for determining the optimal sample size for aerial detection surveys. The method can be applied either to the single objective optimization problem (when there is only one variance function associated with a single causal agent) or to the multiple-objective optimization problem (when several variance functions associated with multiple causal agents and disorders have to be minimized simultaneously). The approach is demonstrated using aerial survey data from Colorado collected over a 20-year period for multiple causal agents and disorders and three sample designs. The results were particularly encouraging, since the approach appeared to find the optimal solutions that made statistical and economic sense. For example, stratified random sampling was more cost-efficient than probabilities proportional to size sampling which resulted in a savings of almost $95,000 per year for estimating total area damage caused by all causal agents and disorders combined. The optimal sample size increased as expected with increasing variability in the population and as the desired level of confidence increased. Finally, larger samples were required to simultaneously provide estimates for multiple causal agents and disorders with reasonable levels of precision when compared to a single causal agent.

In the monitoring of forest health it is important to take into account that the amount of damage being attributed to a given causal agent and how this changes over time. Resource managers need reliable and timely information of these changes in order to make efficient decisions. In states like Alaska where 100% surveys are not economically feasible, sampling may be a cost-efficient alternative. However, a major drawback of aerial survey sampling is the lack of information between transects. As a consequence, the severity and distributions for many pests may be underestimated. Assessments would be more realistic if the areas omitted during

the survey flights could be taken into consideration. Recent work in modeling the spatial distribution of insect pests and diseases using field observations, satellite imagery, and spatial statistics offers some promise for providing more complete cost-effective coverage of pest distributions (Reich et al. 2013). In this case, the aerial survey data could be combined with satellite images and other geo-referenced data into a multi-source inventory system. Under such a system, the cost and the variance functions can be quite complex and the use of the loss + cost methods seem to be a promising approach for evaluating the cost effectiveness of such an inventory system. The important thing with using the loss + cost approach is the fact that it can be easily applied to a large number of optimization problems encountered in natural resource sampling.

REFERENCES

Alves, M. D. d. O., R. Schwamborn, J. C. G. Borges, M. Marmontel, A. F. Costa, C. A. F. Schettini, and M. E. d. Araújo. 2013. Aerial survey of manatees, dolphins and sea turtles off northeastern Brazil: Correlations with coastal features and human activities. Biological Conservation **161**:91-100.

Anderson, D., K. Burnham, and J. Laake. 1993. Distance sampling: estimating abundance of biological populations. Chapman &Hall, London, United Kingdom.

Angelis, L. and G. Stamatellos. 2004. Multiple objective optimization of sampling designs for forest inventories using random search algorithms. Computers and Electronics in Agriculture **42**:129-148.

Assuncao, R. M. and E. A. Reis. 1999. A new proposal to adjust Moran's I for population density. Statistics in Medicine. Statist. Med. 18, 2147-2162. John Wiley & Son, Ltd. 16pp.

B.C. Ministry of Forests and C. F. Service. 2000. Forest Health Aerial Overview Survey Standards for Bristish Columbia. Report 97-1. Version 2.0. Forest Practices Branch and Forest Health Network for the Resources Inventory Committee, Canadian Government Publications Center. 36 ISBN 0-7726-4312-1.

Bailey, R. G. 1980. Description of the ecoregions of the United States. US Department of Agriculture, Forest Service.

Baker, B. H. and J. A. Kemperman. 1974. Spruce Beetle Effects on a White Spruce Stand in Alaska. Journal of Forestry **72**:423-425.

Baker, W. L. 1972. Eastern Forest Insects. USDA Forest Service, Wahsington, DC 20402.

Bechtold, W. A., B. Tkacz, and K. Riitters. 2007. The historical background, framework, and application of forest health monitoring in the United States.*in* Korea forest conservation movement, 2007. Proceedings of the international symposium on forest health monitoring; 30-31 January: 2007; Seoul, Republic of Korea.

Bigler, C., D. G. Gavin, C. Gunning, and T. T. Veblen. 2007. Drought induces lagged tree mortality in a subalpine forest in the Rocky Mountains. Oikos **116**:1983-1994.

Blackford, D. 2010. Management Guide for Western Spruce Budworm. 6.1, USDA Forest Service.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to Distance Sampling. Oxford University Press.

Caratti, J. F. 2006. Line Intercept (LI) Sampling Method. USDA Forest Service Gen. Tech. Rep. RMRS-GTR-164-CD. 13p.

Caughley, G. 1974. Bias in Aerial Survey. The Journal of Wildlife Management **38**:921-933.

Caughley, G. 1977. Sampling in aerial survey. The Journal of Wildlife Management **41**:605-615.

Caughley, G. and A. R. E. Sinclair. 1994. Wildlife ecology and management. Blackwell Science.

Certain, G. and V. Bretagnolle. 2008. Monitoring seabirds population in marine ecosystem: The use of strip-transect aerial surveys. Remote Sensing of Environment **112**:3314-3322.

Chapman, T. B., T. T. Veblen, and T. Schoennagel. 2012. Spatiotemporal patterns of Mountain pine beetle activity in the southern Rocky Mountain. Ecological Society of America **93**:2175-2185.

Chase, M. 2007. Aerial Wildlife Census of the Caprivi River Systems - A survey of rivers, Wetland and Floodplains. Elephants Without Borders, Namibia.

Cochran, W. G. 1977. Sampling Techniques. John Wiley & Sons, New York. 428p. pp83-85.

Conkling, B. L., J. W.Coulston, and M. J. Ambrose. 2005. Forest health monitoring: 2001 national technical report. U.S Department of Agriculture, Forest Service. Southern Research Station. Tech. Rep. SRS - 81. 204 pp.

Coops, N. C., M. A. Wulder, and R. H. Waring. 2012. Modeling lodgepole and jack pine vulnerability to mountain pine beetle expansion into the western Canadian boreal forest. Forest Ecology and Management **274**:161-171.

Cranshaw, W., D. A. Leatherman, and B. Kondratieff. 2012. Insects that feed on Colorado trees and shrubs. Colorado State University Cooperative Extension, Colorado State University.

CSFS. 2011. Colorado Statewide Forest Resource Assessment - A foundation for Strategic Discussion and Implementation of Forest Management in Colorado. 40p.

CSFS. 2012. Colorado Statewide Forest Resource Assessment. Colorado State Forest Service; Fort Collins, CO, 34p.

CSFS. 2013. Colorado Statewide Forest Resource Assessment. Colorado State Forest Service; Fort Collins, CO, 22p.

CSFS. 2014a. 2014 Colorado Forest Insect and Disease Update - A supplement to the 2014 Report on the Health of Colorado's Forests. Colorado State University.

CSFS. 2014b. Colorado's Forest Types.

DeGomez, T. E. and D. Young. 2015. Pine Bark Beetles. Book University of Arizona Cooperative Extension Service and Agricultural Experiment Station Bulletin AZ1300-2015, College of Agriculture, University of Arizona, Tucson, AZ.

Dudley, M. M., K. S. Burns, and W. R. Jacobi. 2015. Aspen mortality in the Colorado and southern Wyoming Rocky Mountains: Extent, severity, and causal factors. Forest Ecology and Management **353**:240-259.

Esseen, P.-A., K. U. Jansson, and M. Nilsson. 2006. Forest edge quantification by line intersect sampling in aerial photographs. Forest Ecology and Management **230**:32-42.

Evans, C. D., W. A.Troyer, and C. J.Lensink. 1966. Aeral Census of Moose by Quadrat Sampling Units. The Journal of Wildlife Management **30**:767-776.

Fairweather, M. L., B. W. Geils, and M. Manthei. 2007. Aspen Decline on the Coconino National Forest. Proceedings of the 55th Western International Forest Disease Work COnference; 2007 October 15-19; Sedona, AZ. Salem, OR: Oregon Department of Forestry.

Fellin, D. G. and J. E. Dewey. 1992. Western Spruce Budworm. Forest Insect & Disease Leaflet 53. USDA Forest Service.

Ferreira, S. M. and R. J. v. Aarde. 2009. Aerial survey intensity as a determinant of estimates of African elephant population sizes and trends. South African Journal of Wildlife Research **39**:181-191.

Freeman, A. M. I., R. H. Haveman, and A. V. Kneese. 1973. Economics of environmental policy.

Furniss, M. M. 2014a. The Douglas-Fir Beetle in Western Forests - A historical perspective - Part 2. American Entomologist **60**:16.

Furniss, M. M. 2014b. The Douglas-Fir beetles in western forests: a historical perspective - Part 1. Entomological Society of America **84**.

Gibson, K., S. Kegley, and B. Bentz. 2009. Moutain Pine Beetle. USDA Forest Service, Portland, Oregon, USA.

Gregoire, T. G. and H. T. Valentine. 2007. Sampling Strategies for Natural Resources and the Environment. Taylor & Francis Group, Boca Raton, FL 33487-2742. International Standard Book Number: 13:978-0-203-49888-0.

Hadley, K. S. and T. T. Veblen. 1993. Stand response to western spruce budworm and Douglas-fir bark beetle outbreaks, Colorado Front Range. Canadian Journal of Forest Research **23**:479-491.

Hagle, S. K., K. E. Gibson, and S. Tnnock. 2003. A Field Guide to Disease and Insect Pests of Northern & Central Rocky Mountain Conifers. USDA Forest Service; State and Private Forestry; Northern Reigon, Missoula, Montana; Intermountain Region, Ogden, Utah. R1-03-08, Montana, USA. 197p.

Hansen, M. H., W. N. Hurwitz, and W. G. Madow. 1953. Sample Survey Methods and Theoy. Volume I. Methods and Application. Wiley Publications In Statistics. pp 425-435.

Hart, S. J., T. T. Veblen, N. Mietkiewicz, and D. Kulakowski. 2015. Negative Feedbacks on Bark Beetle Outbreaks: Widespread and Severe Spruce Beetle Infestation Restricts Subsequent Infestation. Plos One **10**.

Hayes, C. J., T. E. DeGomez, J. D. McMillin, J. A. Anhold, and R. W. Hofstetter. 2008. Factors influencing pine engraver (Ips pini Say) colonization of ponderosa pine (Pinus ponderosa Dougl. ex. Laws.) slash in Northern Arizona. Forest Ecology and Management **255**:3541-3548.

Heller, R. C. and J. Wear. 1969. Sampling forest insect epidemics with color films. Page 1157 *in* Remote Sensing of Environment, VI.

Hevesi, J. A., A. L. Flint, and J. D. Istok. 1992. Precipitation estimation in mountainous terrain using multivariate geostatistics. Part II: Isohyetal maps. Journal of Applied Meteorology **31**:677-688.

Hinds, T. E., F. G. Hawsworth, and R. W. Davidson. 1960. Decay of Subalpine Fir in Colorado. Station Paper No. 51, Rocky Mountain Forest and Range Experiment Station, Colorado, USA.

Irvine, K., A. Gitelman, and J. Hoeting. 2007. Spatial designs and properties of spatial correlation: Effects on covariance estimation. Journal of Agricultural, Biological, and Environmental Statistics **12**:450-469.

Jiang, W., S. Varma, and R. Simon. 2008. Calculating confidence intervals for prediction error in microarray classification using resampling. Statistical Applications in Genetics and Molecular Biology **7**.

Johnson, E. W. and J. Ross. 2008. Quantifying error in aerial survey data. Australian Forestry **71**:216-222.

Johnson, E. W. and D. Wittwer. 2008. Aerial detection surveys in the United States. Pages 212-215 *in* Australian Forestry, editor.

Jolly, G. M. 1981. A review of the sampling methods used in aerial survey. International Livestock Centre for Africa (ILCA), Addis Ababa, Ethiopia. Ser. No.3. Available on line at: http://books.google.com/books?id=ipNj97nz9tAC&pg=PA149.

Leung, B. P. and F. A. Spiring. 2004. Some properties of the family of inverted probability loss functions. Quality Technology and Quantitative Management **1**:125-147.

Livingston, L. 2010. Management Guide for Pine Engraver. Page 6 *in* F. H. P. a. S. F. Organizations, editor. USDA Forest Service, State of Idaho.

Lohr, S. L. 2010. Sampling: Design and Analysis. Richard Stratton, United State of America, Boston, MA 02210. ISBN 13-978-0-495-10527-5. Library of Congress Control Number: 2009938648.

Lundquist, J. E. 2005. Landscape pathology-forest pathology in the era of landscape ecology. Pages 155-165. American Phytopathological Society (APS Press), St. Paul.

Magnussen, S. and R. I. Alfaro. 2012. Linking aerial survey data of forest insect defoliation and tree ring data to estimate forest level growth losses. Dendrochronologia **30**:287-294.

Mandallaz, D. 2007. Sampling Techniques for Forest Inventories. Chapman and Hall/CRC.

Marchetti, S. B., J. J. Worrall, and T. Eager. 2011. Secondary insects and diseases contribute to sudden aspen decline in southwestern Colorado, USA. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere **41**:2315-2325.

Marsh, H. and D.F.Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Manage **53**:1017-1024.

Massey, C. L. and N. D. Wygant. 1954. Biology and control of the Engelmann Spruce beetle in Colorado. Circular. United States Department of Agriculture:35.

McConnell, T. and R. Avila. 2004. Aerial Detection Overview Surveys Futuring Committee Report. USDA Forest Service, Forest Health Technology Enterprise Team, Fort Collins, Colorado. FHTET - 04-07.

McConnell, T. J. 1999. Aerial Sketch Mapping Surveys the Past, Present and Future. USDA Forest Service Proceeding RMRS-P-12, Guradaljara, Mexico.

McConnell, T. J., E. W. Johnson, and B. Burns. 2000. A Guide to Conducting Aerial Sketchmapping Surveys. USDA Forest Service Forest Health Technology Enterprise Team. , Fort Collins, Colorado. FHTET 00-01.

McGinn, T. 2004. Instructions for Probability Proportional to Size Sampling Technique.*in* C. U. Heilbrunn Department of Population and Family Health Mailman School of Public Health, editor.

McMillin, J. D. and K. K. Allen. 2000. Impacts of Douglas-fir Beetle on Overstory and Understory condistions of Douglas-fir Stands Shoshone National Forest, Wyoming. Technical Report R2-64, USDA Forest Service, Rocky Mountain Region, Golden, CO 80401, Colorado, USA. R2-64.

Moran, P. A. P. 1950. Notes on Continuous Stochastic Phenomena. Biometrika **37**:17-23.

Nusser, S. M., F. J. Breidt, and W. A. Fuller. 1998. Design and Estimation for Investigating the Dynamics of Natural Resources. Ecological Applications **8**:234-245.

Park, I. and H. Lee. 2001. The Design Effect: Do we know all about it? , Maryland, USA.

Penn State Science. 2015. STAT 415: Intro Mathematical Statistics *in* O. l. 49, editor.

Pollock, K. H. and W. L. Kendall. 1987. Visibility Bias in Aerial Surveys: A Review of Estimation Procedures. The Journal of Wildlife Management **51**:502-510.

Rehfeldt, G. E., D. E. Ferguson, and N. L. Crookston. 2009. Aspen, climate, and sudden decline in western USA. Forest Ecology and Management **258**:2353-2364.

Reich, R. M. 2011. Spatial library for R package for Quantitative Spatial Statistics. Colorado State University handout package. URL: http://www.stat.colostate.edu/~rdavis/st523/.

Reich, R. M. and R. A. Davis. 2011. Quantitative Spatial Statistics: course handout for NR/ST 523. Colorado State University handout. Fort Collins, CO:481.

Reich, R. M., J. E. Lundquist, and V. A. Bravo. 2013. Characterizing Spatial Distributions of Insect Pests Across Alaskan Forested Landscape: A Case Study Using Aspen Leaf Miner (PbyUocnistis populieUa Chambers) Journal of Sustainable Forestry **32**:527-548.

Ridgway, M. S. 2010. Line transect distance sampling in aerial surveys for double-crested cormorants in coastal regions of Lake Huron. Journal of Great Lakes Reserch **36**:403-410.

Romme, W. H., J. Clement, J. Hicke, D. Kulakowski, L. H. MacDonald, T. L. Schoennagel, and T. T. Veblen. 2006. Recent Forest Insect Outbreak and Fire Risk in Colorado Forests: A Brief Synthesis of Relevant Research. Colorado Forest Restoration Institute.

Romme, W. H., D. H. Knight, and J. B. Yavitt. 1986. Mountain Pine Beetle Outbreaks in the Rocky Moutains: Regulators of Primary Productivity? The American Naturalist **127**:484-494.

Rowat, D., M. Gore, M. G. Meekan, I. R. Lawler, and C. J. A. Bradshaw. 2009. Aerial survey as a tool to estimate whale shark abundance trends. Journal of Experimental Marine Biology and Ecology **368**:1-8.

Ryerson, D. E., T. W. Swetnam, and A. M. Lynch. 2003. A tree-ring reconstruction of western spruce budworm outbreaks in the San Juan Mountains, Colorado, U.S.A. Canadian Journal of Forest Research **33**:1010-1028.

Saeki, I. 2005. Application of aerial survey for detecting a rare maple species and endangered wetland ecosystems. Forest Ecology and Management **216**:283-294.

Safranyik, L., A. L. Carroll, J. Régnière, D. W. Langor, W. G. Riel, T. L. Shore, B. Petter, B. J. Cooke, V. G. Nealis, and S. W. Taylor. 2010. Potential for range expansion of mountain pine beetle into the boreal forest of North America. Entomoligical Society of Canada **142**:415-442.

Scheaffer, R. L., W. M. III, and R. L. Ott. 2006. Elementary Survey Sampling - Sixth Edition. Duxbury, Belmont, California CA 94002-3098. ISBN 0-534-41805-8. Library of Congress Control Number: 2004112934.

Schmid, J. M. and R. H. Frye. 1977. Spruce Beetle in the Rockies. Rocky Mountain Forest and Range Experiment Station, USDA Forest Service.

Schmitz, R. and K. E. Gibson. 1996. Douglas-fir beetle. USDA Forest Service.

Schreuder, H. T., R. L. Ernst, and H. Ramirez-Maldonado. 2004. Statistical techniques for sampling and monitoring natural resources. USDA Forest Service. Rocky Mountain Research Station, Gen. Tech. Rep. RMRS-GTR-126. Fort Collins, CO. .

Shore, T. and L. Safranyik. 1992. Susceptibility and risk rating systems for the mountain pine beetle in lodgepole pine stands. Inf. Rep. BC-X-336. Forestry Canada, Pacific Forestry Centre, Victoria, BC.

Sirken, M. 2001. The Hansen-Hurwitz estimator revisited: PPS sampling without replacement. ASA Proceedings of the Survey Methods Section.

Smith, G. E. J. 1981. Some thoughs on sampling design. Low-level aerial survey techniques. Internaltional Livestock Center for Africa (ILCA). Report of an International Workshop, Nairobi, Kenya., Ethiopia.

SPSS Inc. 2006. SPSS 15.0 Brief Guide.*in* S. Inc., editor., Chicago, IL 60606-6412, USA.

Ståhl, G., S. Holm, T. G. Gregoire, T. Gobakken, E. Næsset, and R. Nelson. 2010. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, NorwayThis article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time. Canadian Journal of Forest Research **41**:96-107.

Steed, B. E. and M. R. Wagner. 2004. Importance of log size on host selection in reproductive success of Ips pini (Coleoptera: Scolytidae) in ponderosa pine slash of Northern Arizona and western Montana. Journal of Econ. Entomo. **97**:436-450.

Steinman, J. 2004. Forest Health Monitoring in the Northeastern United States. Disturbances and Condition during 1993-2002. USDA Forest Service, Northeastern Area, Newtown Square, PA. NA-TP-01-04. 7 pp.

Stevens, D. L. and A. R. Olsen. 2004. Spatially Balanced Sampling of Natural Resources. Journal of the American Statistical Association **99**:262-278.

Taiti, S. W. 1981. Aerial survey methods: the experience in Kenya. Low-level Aerial Survey Techniques. International Livestock Centre for Africa (ILCA), Addis Ababa, Ethiopia. Available online at: http://books.google.com/books?id=ipNj97nz9tAC&pg.

Tanzania Wildlife Research Institute. 2009. Aerial census in the Selous-Niassa wildlife corridor. TAWIRI Aerial Survey Report, Tanzania Wildlife Research Institute, Dar es Salam, Tanzania. Available online at: http://www.selous-niassa-corridor.org/uploads/tx_drblob/storage/Selous_Niassa_WC_2009_Census_Report_final.pdf. 35 pp.

Tanzania Wildlife Research Institute. 2013. Aerial Census of Large Animals in the Selous-Mikumi Ecosystem - Population Status of African Elephant. Tanzania Wildlife Research Institute and Wildlife Divison of Tanzania National Parks, Tanzania. Available online at: http://www.daressalam.diplo.de/contentblob/4102454/Daten/3821641/Download_Report_Selous_Elephants.pdf. 13 pp.

Thatcher, R. C. 1960. Bark beetles affecting southern pines: a review of current knowledge. Occasional Papers. Southern Forest Experiment Station **25**:180.

Theobald, D., D. Stevens, Jr., D. White, N. S. Urquhart, A. Olsen, and J. Norman. 2007. Using GIS to Generate Spatially Balanced Random Survey Designs for Natural Resource Applications. Environmental Management **40**:134-146.

Thomas, L., R. Williams, and D. Sandilands. 2007. Designing line transect surveys for complex survey regions. Journal of Cetacean Res. Manage **9**:1-13.

Thompson, M. T., J. A. Duda, L. T. Deblander, J. D. Shaw, C. Witt, T. A. Morgan, and M. C. Amacher. 2010. Colorado's Forest Resources, 2002-2006. Resource. Bull. RMRS-RB-11. USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Thompson, W. L., A. E. Miller, D. C. Mortenson, and A. Woodward. 2011. Developing effective sampling designs for monitoring natural resources in Alaskan national parks: An example using simulations and vegetation data. Elsevier **144**:1270-1277.

Tokola, T. and S. M. Shrestha. 1999. Comparison of cluster-sampling techniques for forest inventory in southern Nepal. Elsevier **116**:219-231.

Traat, I., L. Bondesson, and K. Meister. 2004. Sampling design and sample selection through distribution theory. Journal of Statistical Planning and Inference **123**:395-413.

Uchytil, R. J. 1991. Abies lasiocarpa. In: Fire Effects Information System. Rocky Mountain Research Station, Fire Science Laboratory, USDA Forest Service. Available online at: http://www.fs.fed.us/database/feis/.

University of Toronto. 2002. Censusing Populations in Algonquin Park - Appendix A: Counting Populations. University of Toronto, Canada. Available online at: http://bio150.chass.utoronto.ca/sampling/book/sampling.html.

US Forest Service. 2005. Aerial Survey Geographic Information System Handbook, Sketmaps to Digital Geographic Information. USFS. Available online at: http://www.fs.fed.us/foresthealth/aviation/resources/docs/GISHandbook_body_apndxA-C.pdf. 35pp.

USDA Forest Service. 2004. 2004 Insect and Disease Aerial Detection Surveys - Acres with Mortality. Forest Health Monitoring. Available online at: http://fhm.fs.fed.us/dm/maps/04/surveys_acres_mortality.pdf.

USDA Forest Service. 2005. Aerial Survey GIS Handbook - Appendix G: Forest Type Codes.*in* U. F. Service, editor.

USDA Forest Service. 2011a. Armillaria Root Disease - Most important root disease in the Rocky Mountain Region. USDA Forest Service, Forest Health Protection.

USDA Forest Service. 2011b. National Forest of Colorado. Pages Available on line at: http://www.foresthistory.org/ASPNET/Publications/region/2/colorado/sec1.htm. Last updated 12-Sep-2011. Electronic edition courtesy of the Forest History Society.

USDA Forest Service. 2011c. Sudden Aspen Decline in Colorado. Forest Health Protection, Rocky Mountain Region. pp.4.

USDA Forest Service. 2015. Mountain pine beetle epidemic. USDA Forest Service Medicin Bow-Routt National Forests, Thunder Basin National Grassland, Laramie, WY 82070.

Wahl, J. 2006. Dwarf mistletoe - The Quiet Kiss of Death. Page Available online at: http://www.landandwater.com/features/vol49no46/vol49no46_41.html. Wahl Marketing Communications.

Walsh, P. D., L. J. T. White, C. Mbina, D. Idiata, Y. Mihindou, F. Maisels, and M. Thibault. 2001. Estimates of forest elephant abundance: projecting the relationship between precision and effort. Journal of Applied Ecology **38**.

Walter, M. J. and J. Hone. 2003. A comparison of 3 aerial survey techniques to estimate wild horse abundance in the Australian Alps. Wildlife Society Bullentin **31**:1138-1149.

Williams, D. W. and A. M. Liebhold. 2000. Spatial synchrony of spruce budworm outbreaks in eastern North America. Ecology **81**:2753-2766.

Williams, R. E., I. C.G. Shaw, P. M. Wargo, and W. H. Sites. 1986. Armillaria Root Disease. USDA FS.

Withrow, J. R., J. E. Lundquist, and J. F. Negrón. 2013. Spatial Dispersal of Douglas-Fir Beetle Populations in Colorado and Wyoming. ISRN Forestry **2013**:10 pp.

Woodall, C. W., R. S. Morin, J. R. Steinman, and C. H. Perry. 2009. Comparing evaluations of forest health based on aerial surveys and field inventories: Oak forests in the Northern United States. Ecological Indicators **10**:713-718.

Worrall, J., S. Marchetti, L. Egeland, R. Mask, T. Eager, and B. Howell. 2010. Effects and etiology of sudden aspen decline in southwestern Colorado, USA. Forest Ecology and Management **260**:638-648.

Worrall, J., S. Marchetti, and R. G.E. 2014. Sudden Aspen Decline report for Spruce Beetle Epidemic and Aspen Decline Management Response EIS. Article, USFS, Forest Health Protection, Gunison Service Center.

Worrall, J. J., R. A. Mask, T. Eager, L. Egeland, and W. D. Shepperd. 2008. Sudden aspen decline in southwest Colorado. Phytopathology **98**:S173-S173.

Wulder, M. A., C. C. Dymond, J. C. White, D. G. Leckie, and A. L. Carroll. 2006. Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. Forest Ecology and Management **221**:27-41.

Zimmerman, K. L., R. S. McNay, L. Giguere, S. Walshe, G. A. Keddie, L. Wilson, K. Schmidt, P. E.Hengeveld, and A. M.Doucette. 2002. Aeral-Based Census Results for Caribou and Moose in the Mckenzie Timber Supply Area March 2002. Report No.044. Wildlife Infometrics Inc., Mackenzie, British Columbia. 33 pp.

APPENDICES

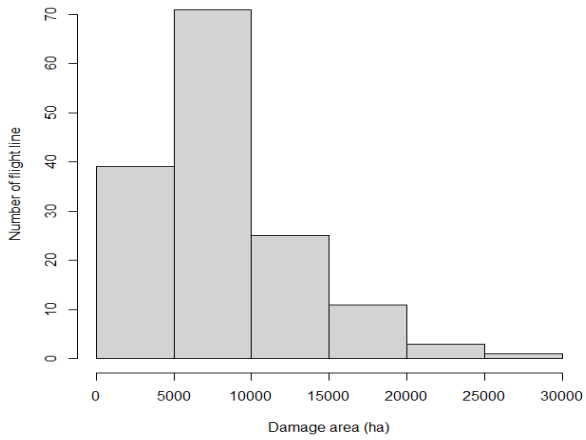


Figure A1a. Frequency distribution of accumulative area damaged of aspen forest in 20 years (1994-2013). The x-axis is damage area (ha), the y-axis is number of transect (transect).

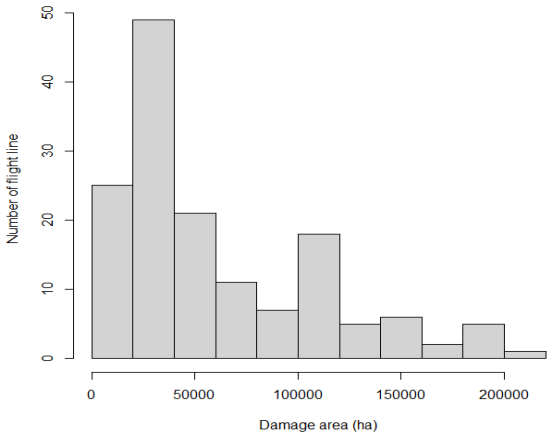


Figure A1b. Frequency distribution of accumulative area damaged of mixed conifer forest in 20 years (1994-2013).The x-axis is damage area (ha), the y-axis is number of transect (transect).

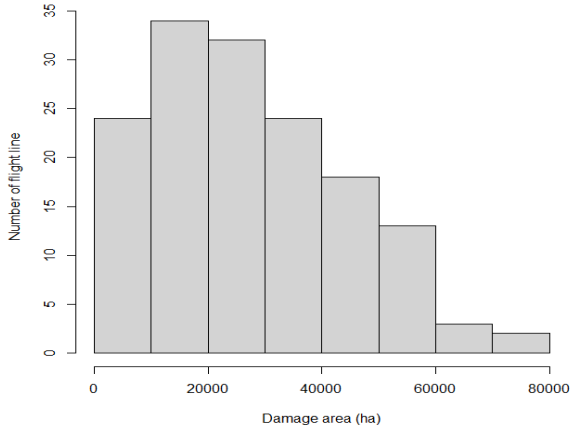


Figure A1c. Frequency distribution of accumulative area damaged of western spruce-fir forest in 20 years (1994-2013). The x-axis is damage area (ha), the y-axis is number of transect (transect)

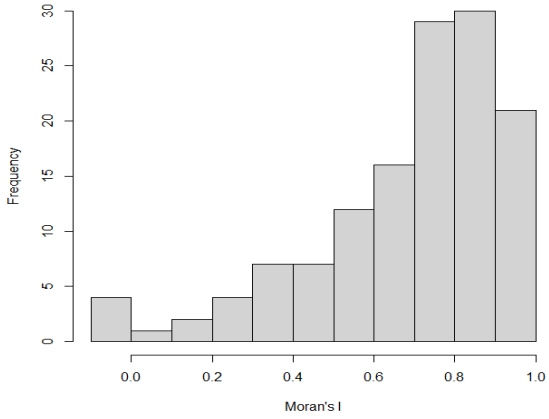


Figure A2a. Moran's I distribution of area damaged on adjacent transects caused by seven single causal agents and disorders combined in 20 years (1994-2013). The x-axis is Moran's I value, the y-axis is frequency.

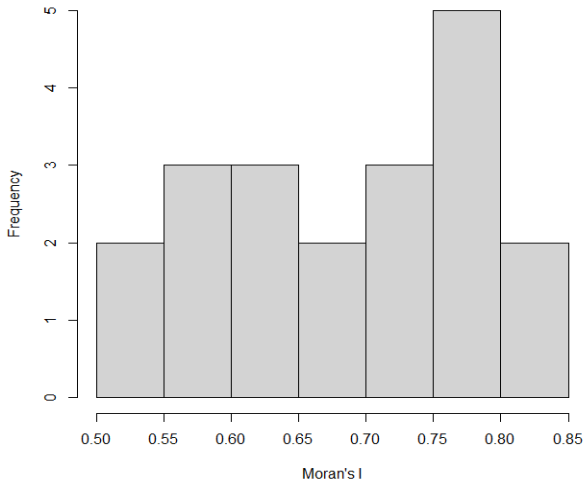


Figure A2b. Moran's I distribution of area damaged on adjacent transects caused by DFB in 20 years (1994-2013). The x-axis is Moran's I value, the y-axis is frequency.

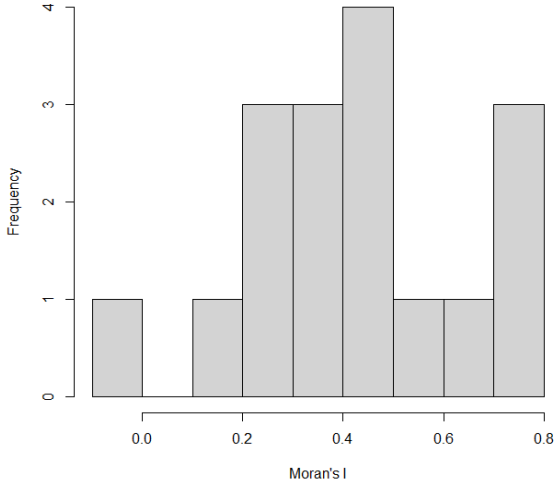


Figure A2d. Moran's I distribution of area damaged on adjacent transects caused by PE in 20 years (1994-2013). The x-axis is Moran's I value, the y-axis is frequency.

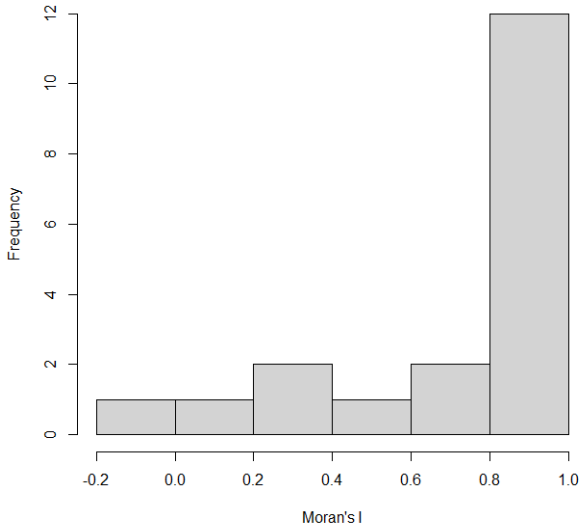


Figure A2c. Moran's I distribution of area damaged on adjacent transects caused by SB in 20 years (1994-2013). The x-axis is Moran's I value, the y-axis is frequency.

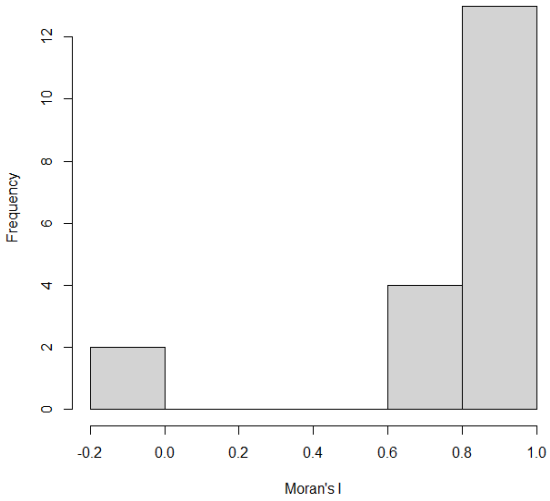


Figure A2e. Moran's I distribution of area damaged on adjacent transects caused by WSB in 20 years (1994-2013). The x-axis is Moran's I value, the y-axis is frequency.

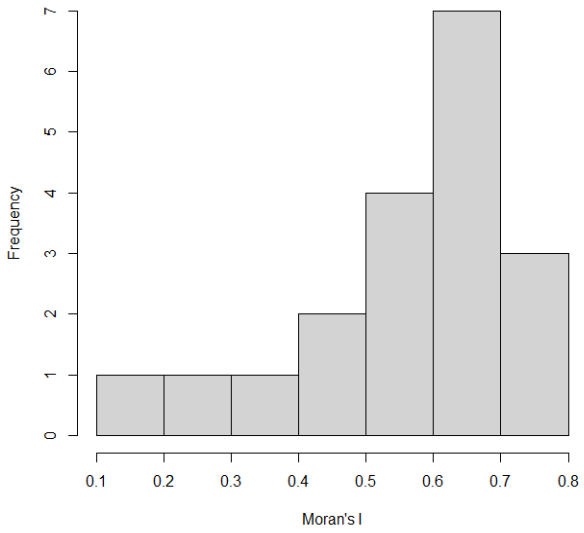


Figure A2f. Moran's I distribution of area damaged on adjacent transects caused by SAD in 20 years (1994-2013). The x-axis is Moran's I value, the y-axis is frequency.

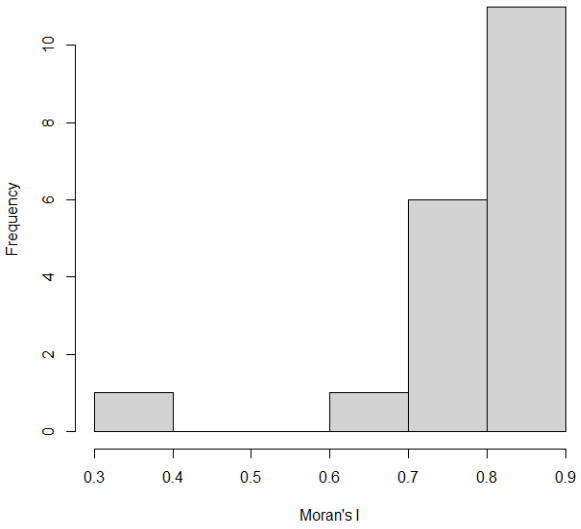


Figure A2g. Moran's I distribution of area damaged on adjacent transects caused by SUB in 20 years (1994-2013). The x-axis is Moran's I value, the y-axis is frequency.

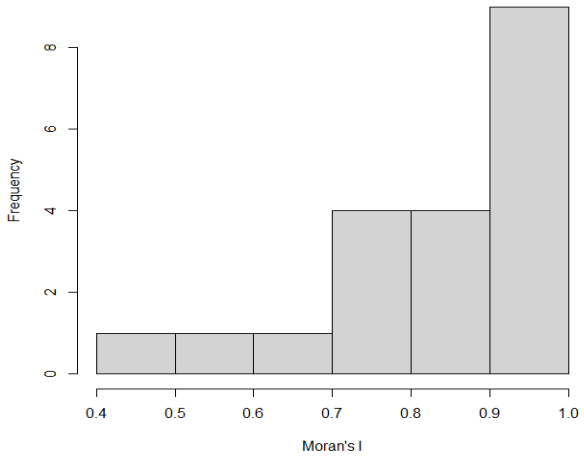


Figure A2h. Moran's I distribution of area damaged on adjacent transects caused by MPB in 20 years (1994-2013). The x-axis is Moran's I value, the y-axis is frequency

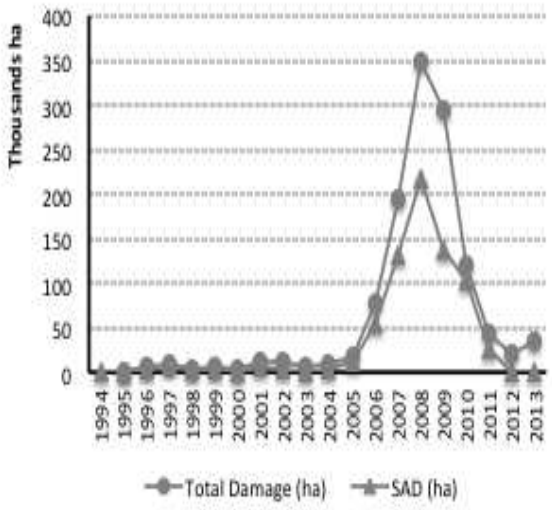


Figure A3. Accumulative total area damaged of Aspen forest and its area damaged caused by SAD from 1994 to 2013

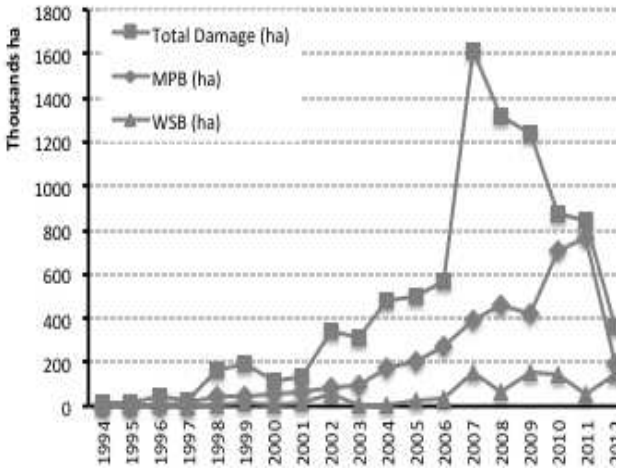


Figure A4. Accumulative total area damaged of mixed conifer forest and its area damaged caused by MPB and WSB from 1994 to 2013

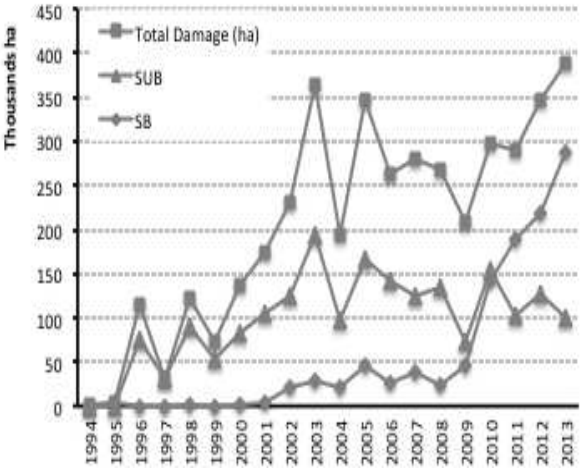


Figure A5. Accumulative total area damaged of western spruce-fir forest and its area damaged caused by SUB and SB from 1994 to 2013

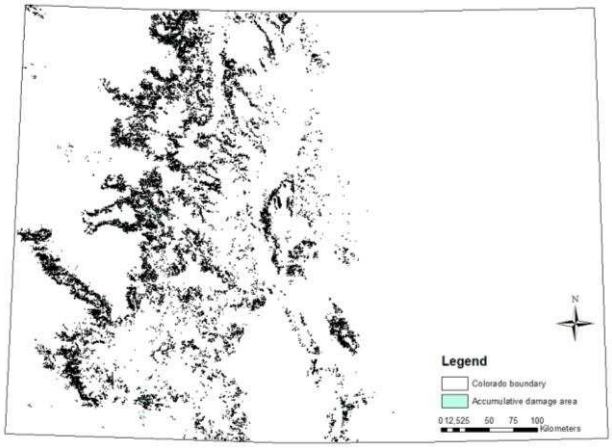


Figure A6a. Distribution of accumulative area damaged of aspen forest from 1994 to 2013

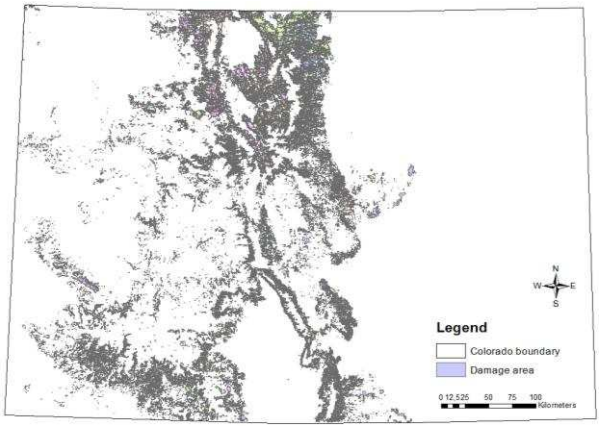


Figure A6b. Distribution of accumulative area damaged of mixed conifer forest from 1994 to 2013

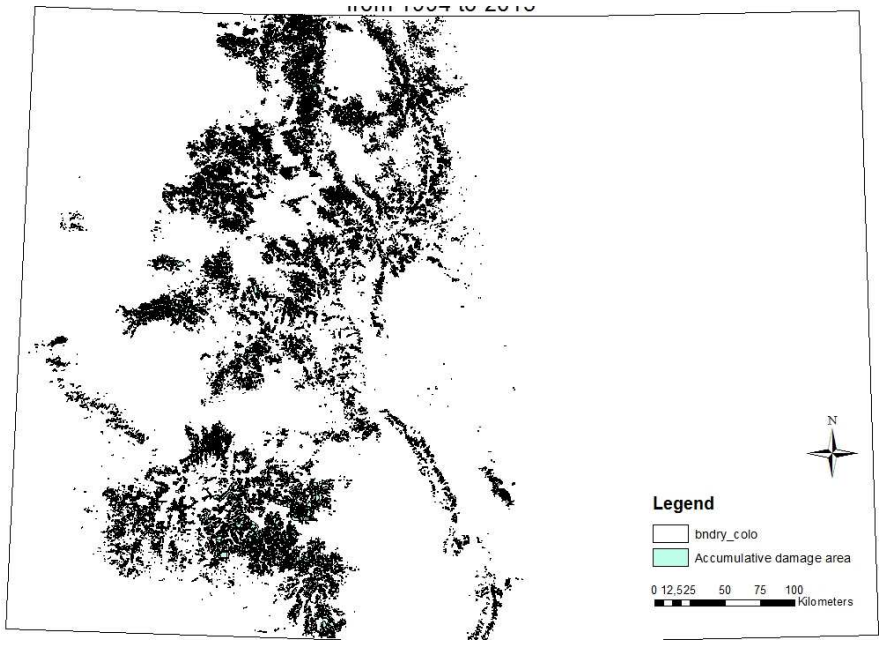
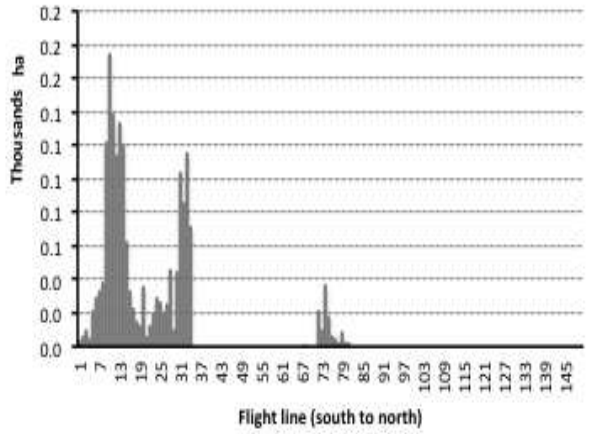
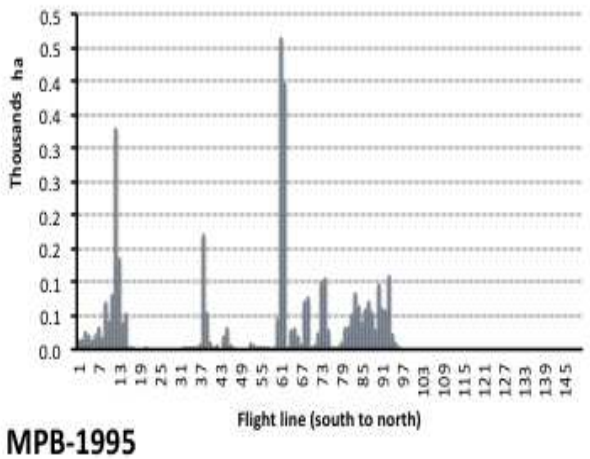


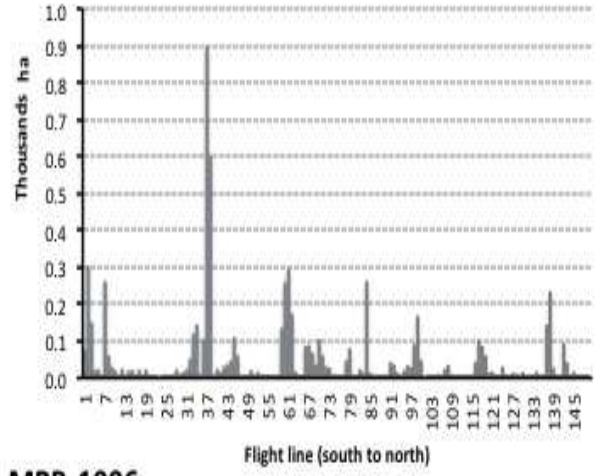
Figure A6c. Distribution of accumulative area damaged of western spruce-fir forest from 1994 to 2013



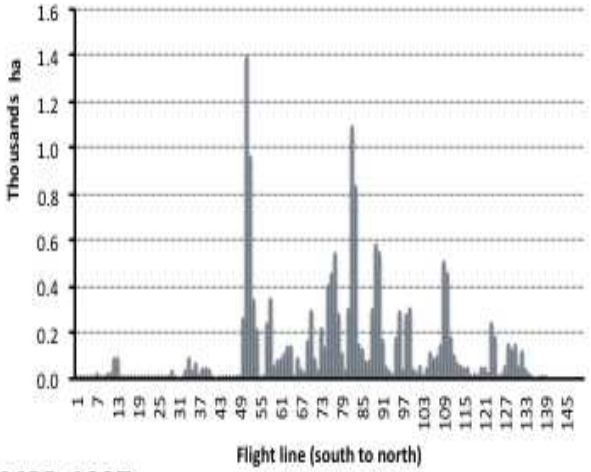
MPB-1994



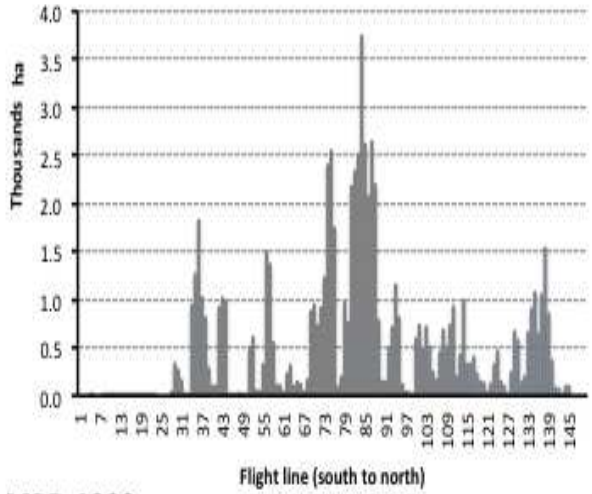
MPB-1995



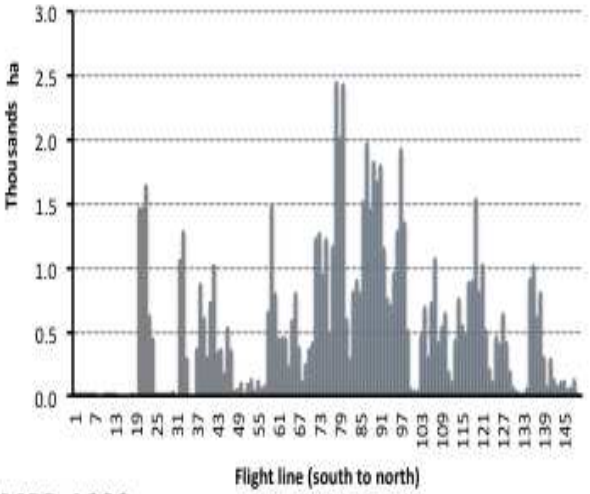
MPB-1996



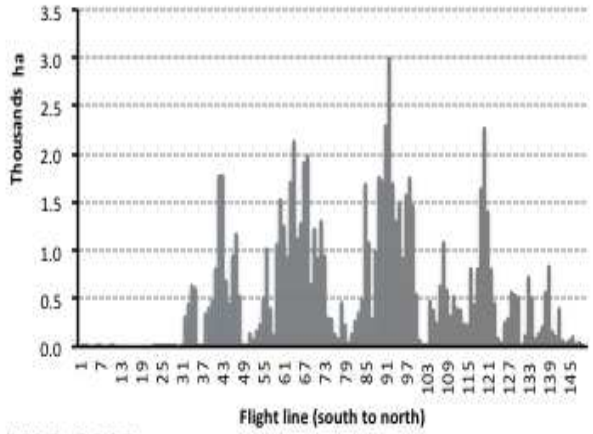
MPB-1997



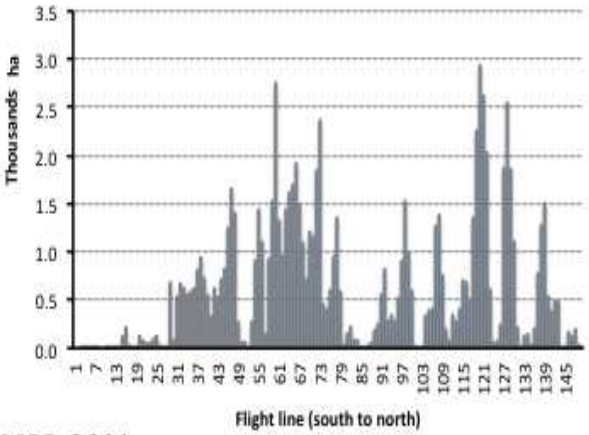
MPB-1998



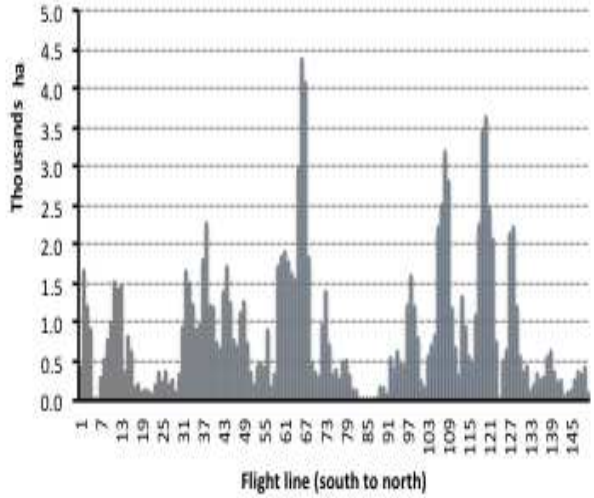
MPB-1999



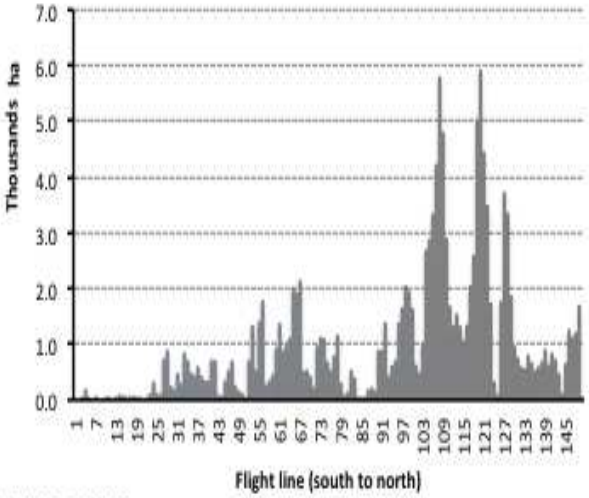
MPB-2000



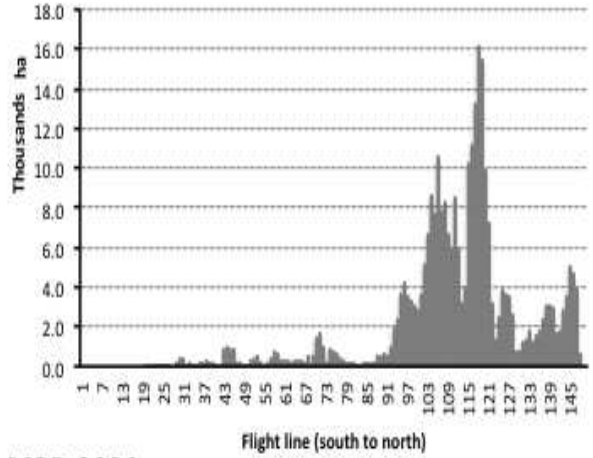
MPB-2001



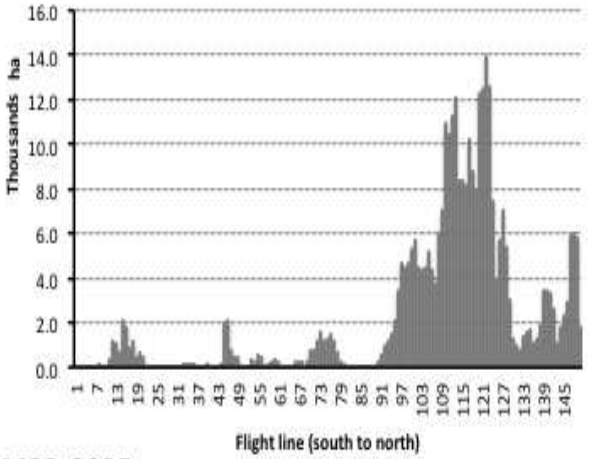
MPB-2002



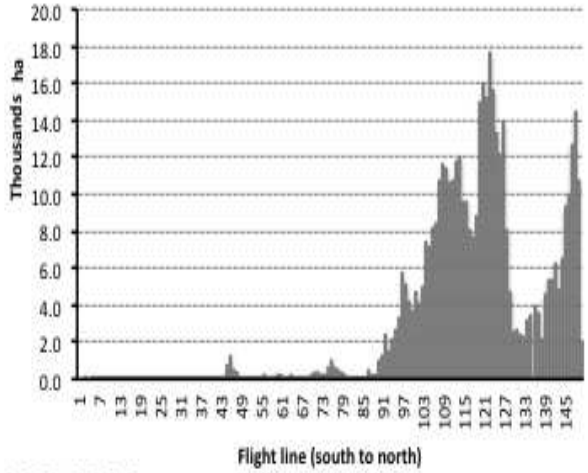
MPB-2003



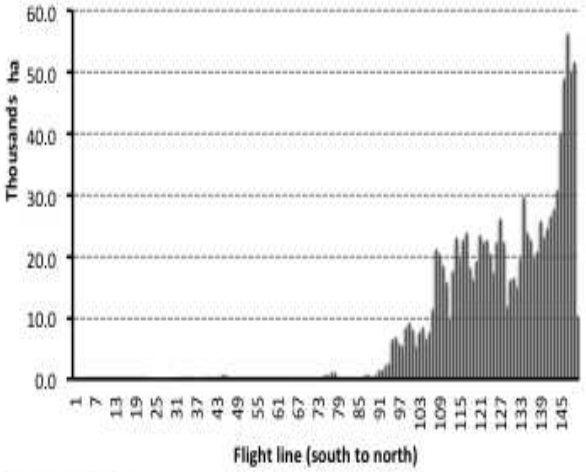
MPB-2004



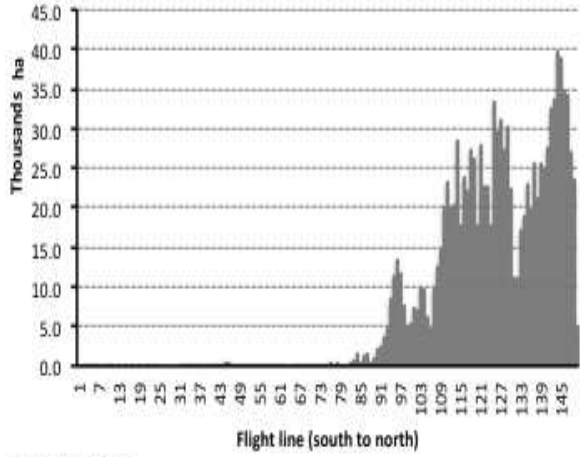
MPB-2005



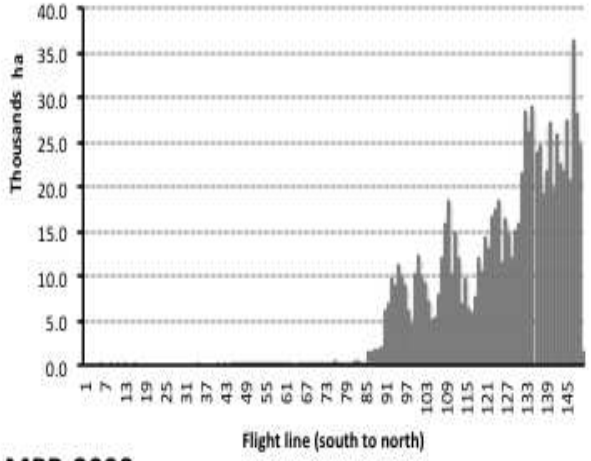
MPB-2006



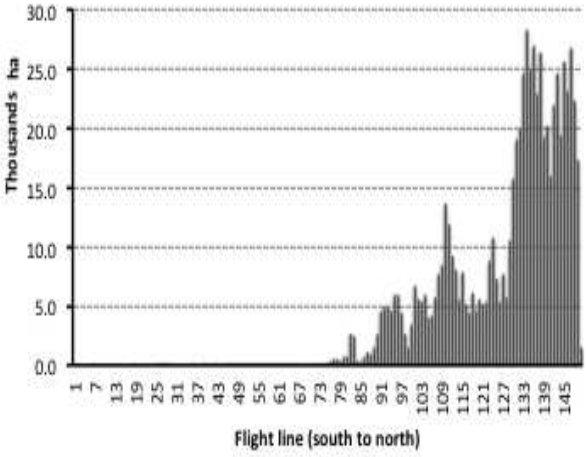
MPB-2007



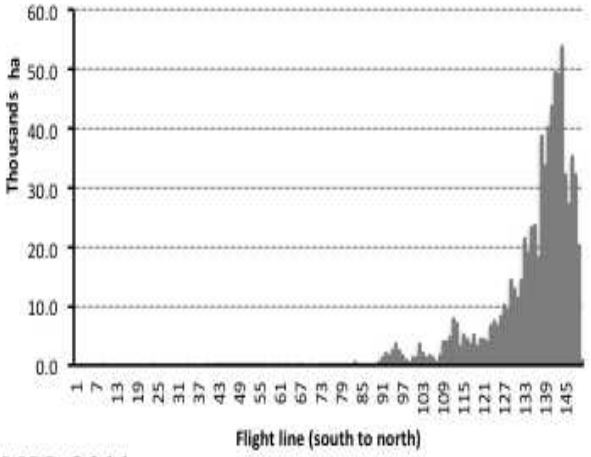
MPB-2008



MPB-2009



MPB-2010



MPB-2011

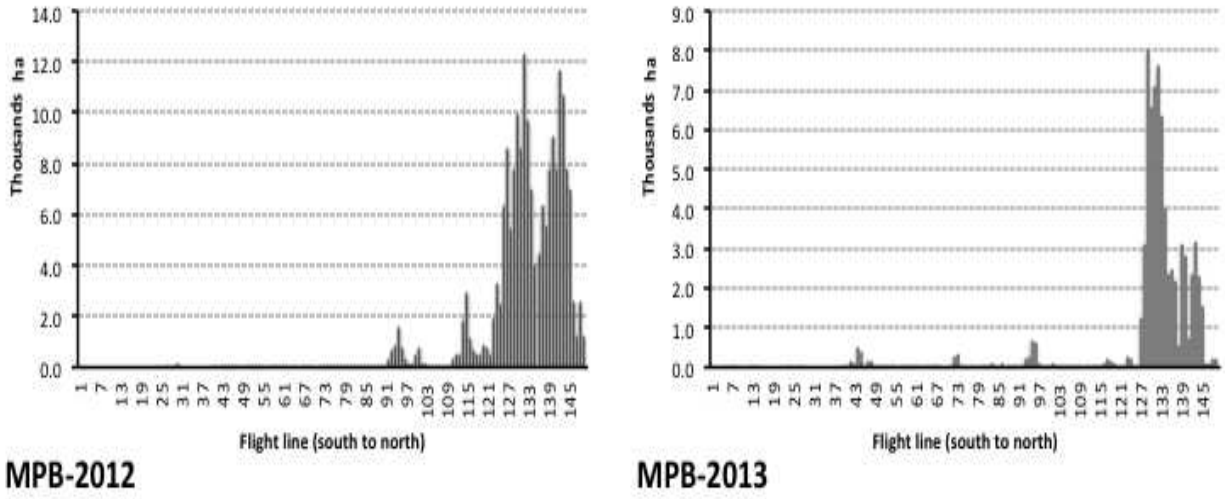
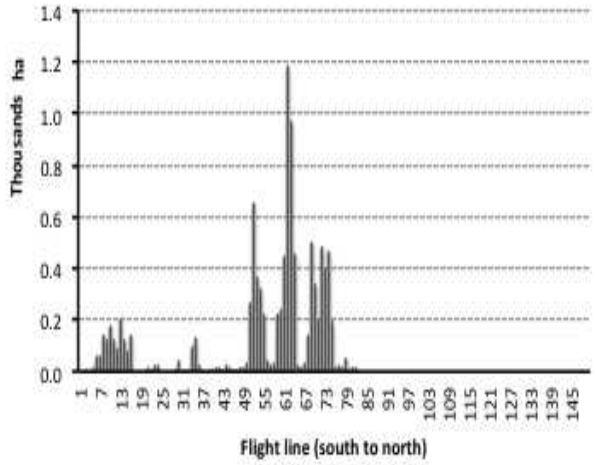
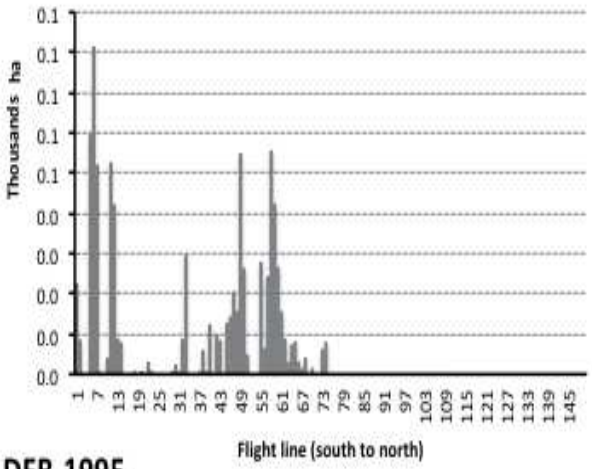


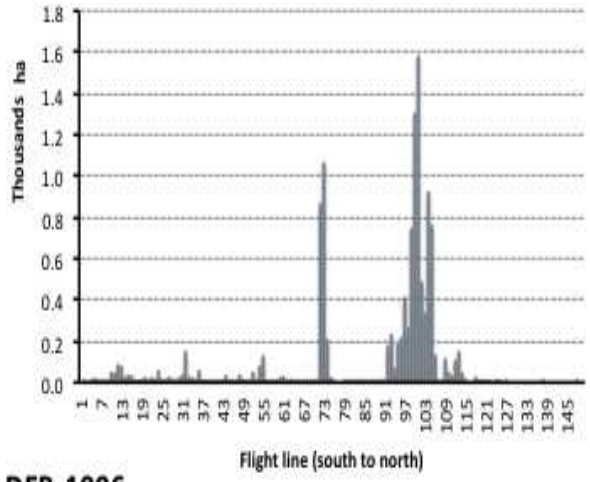
Figure A7. Yearly spatio-temporal distribution of area damaged caused by MPB. The x-axis is number of transect (flight line), the y-axis is damage area (thousand ha)



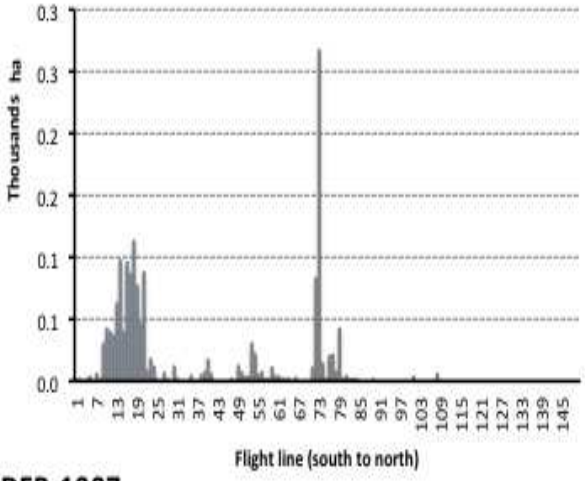
DFB-1994



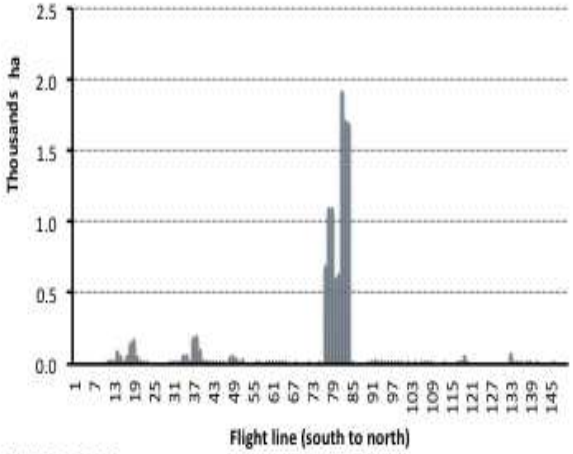
DFB-1995



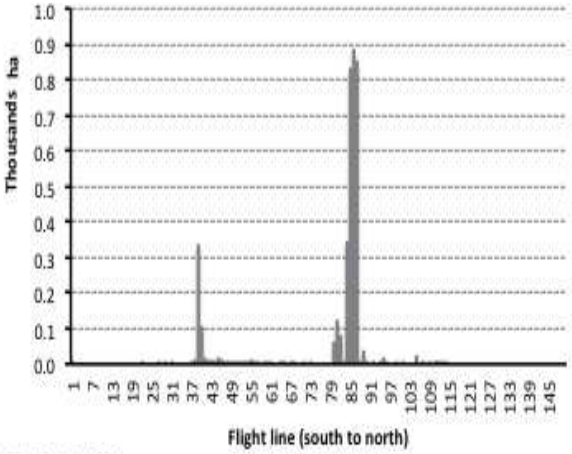
DFB-1996



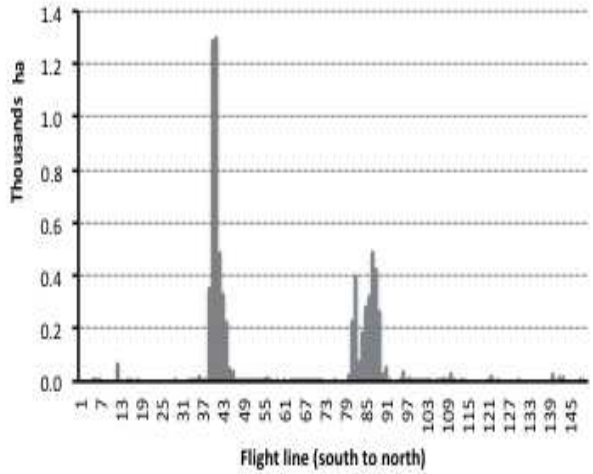
DFB-1997



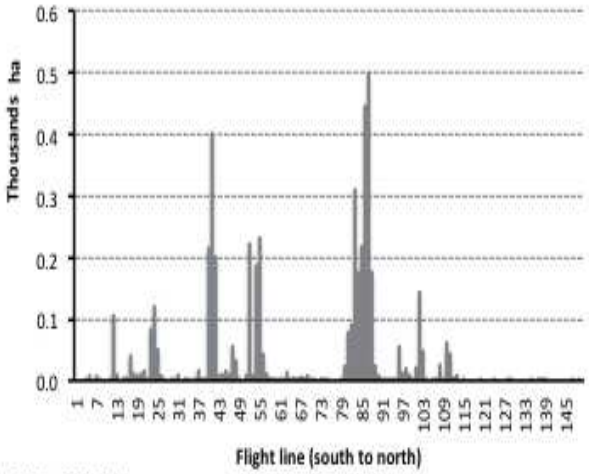
DFB-1998



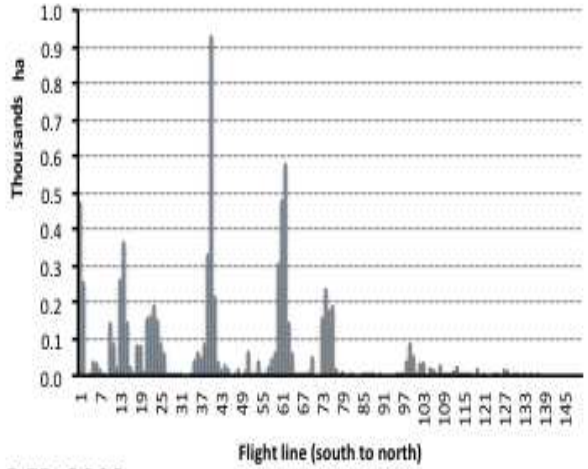
DFB-1999



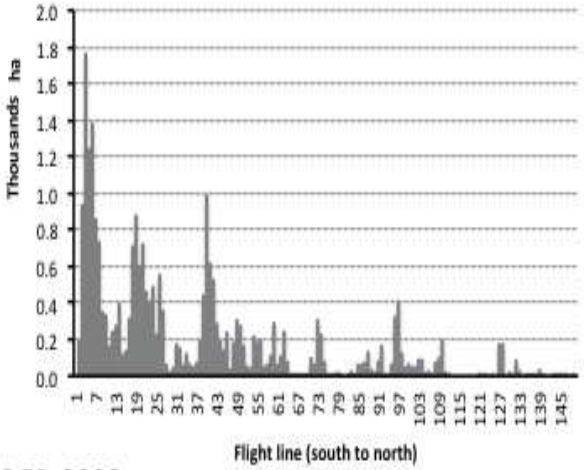
DFB-2000



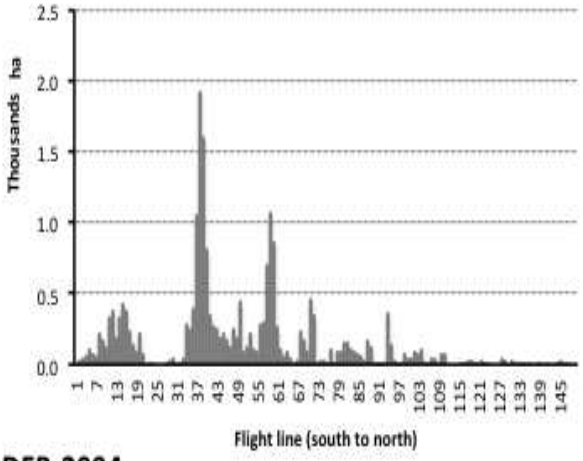
DFB-2001



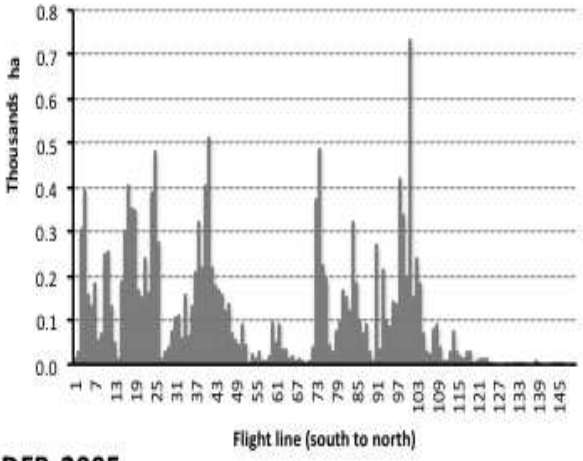
DFB-2002



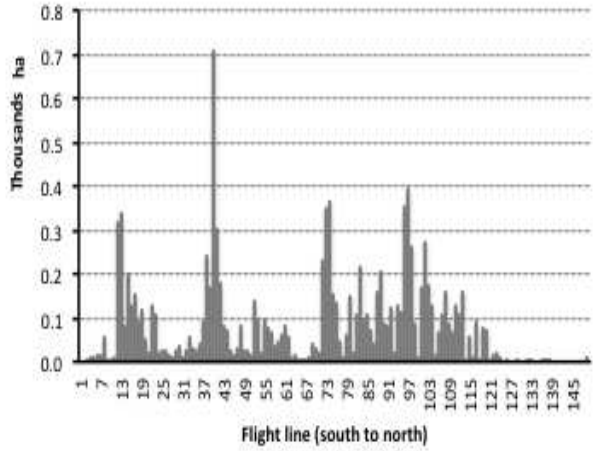
DFB-2003



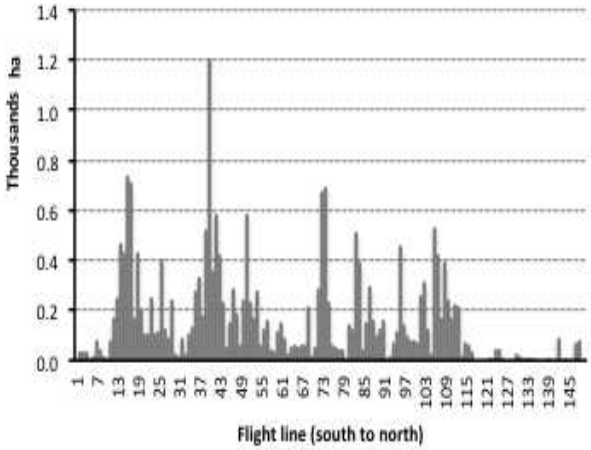
DFB-2004



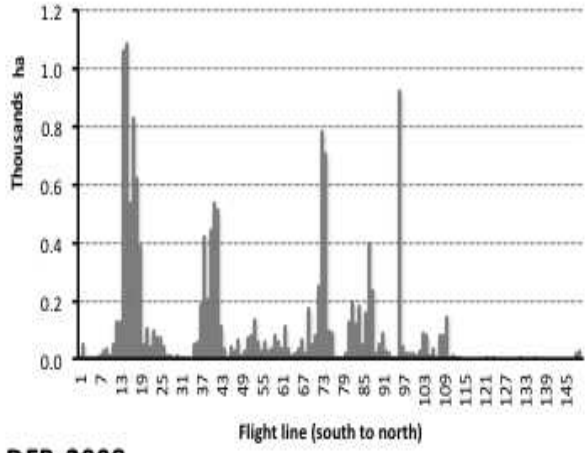
DFB-2005



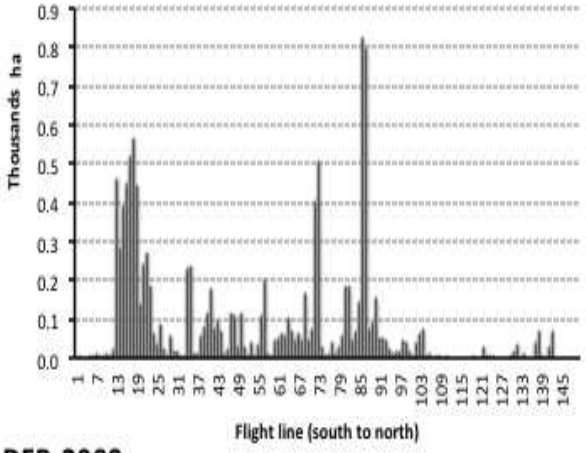
DFB-2006



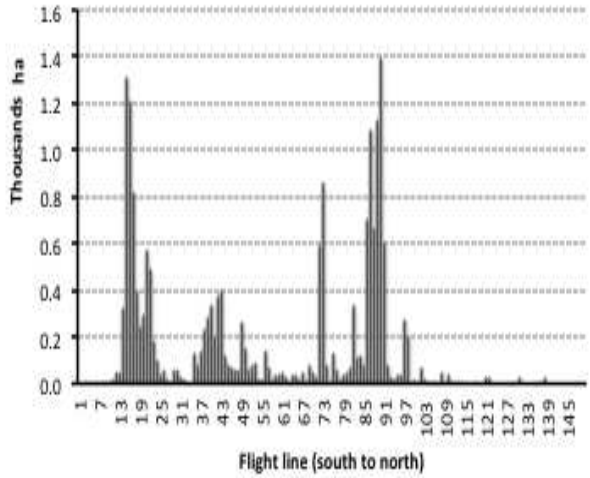
DFB-2007



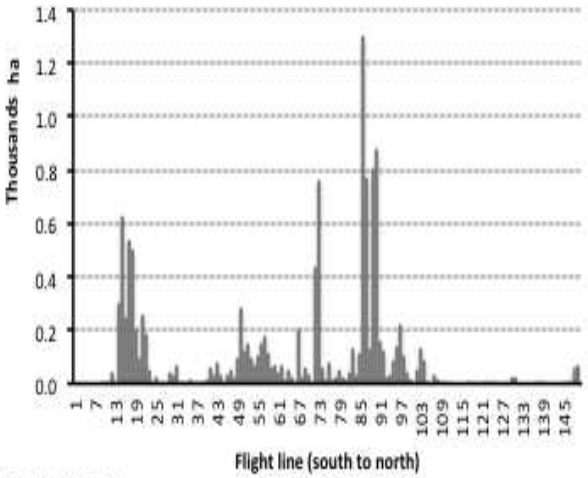
DFB-2008



DFB-2009



DFB-2010



DFB-2011

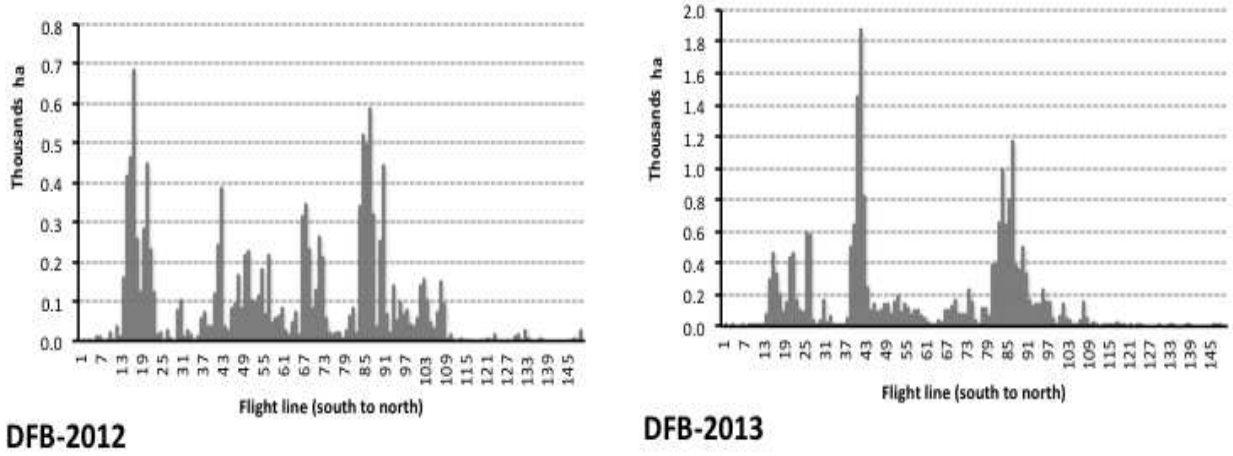
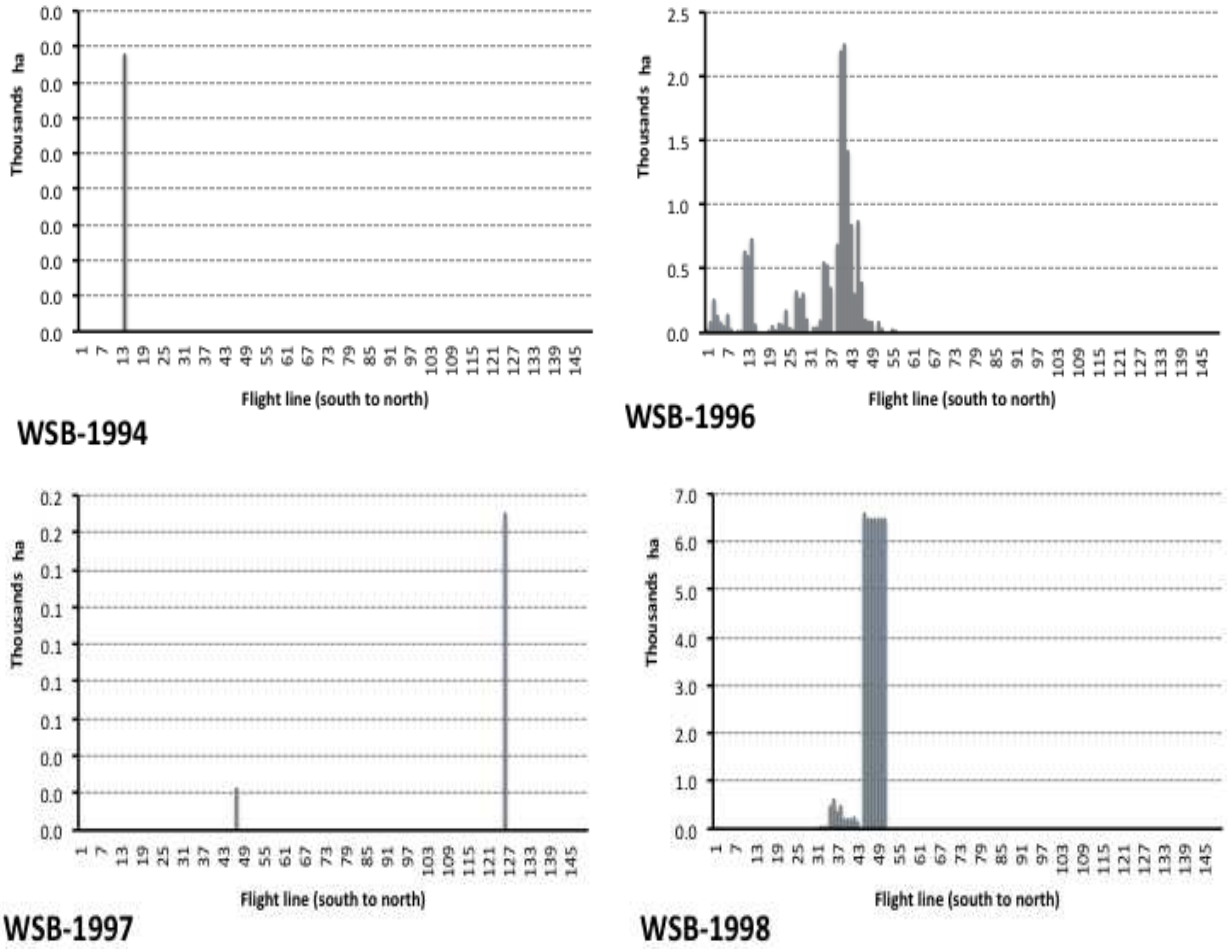
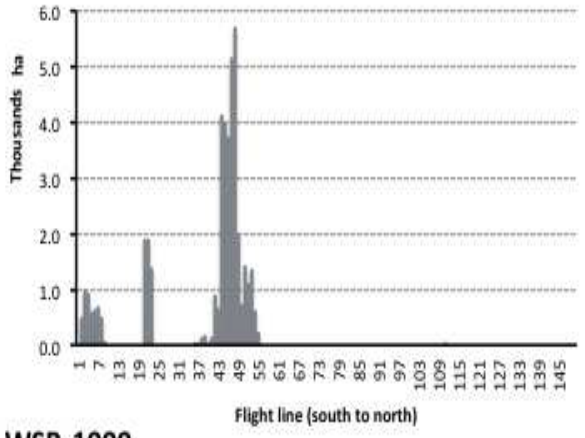
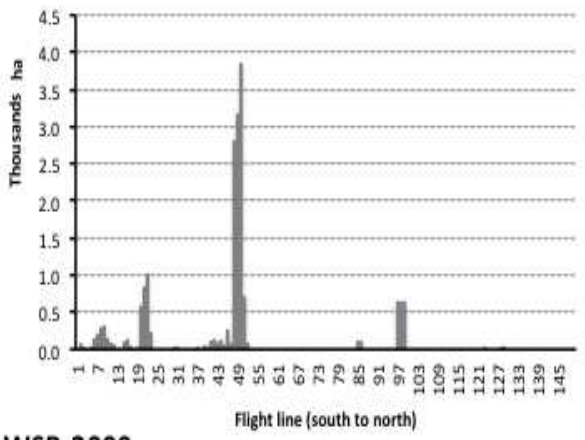


Figure A8. Yearly spatio-temporal distribution of area damaged caused by DFB. The x-axis is number of transect (flight line), the y-axis is damage area (thousand ha)

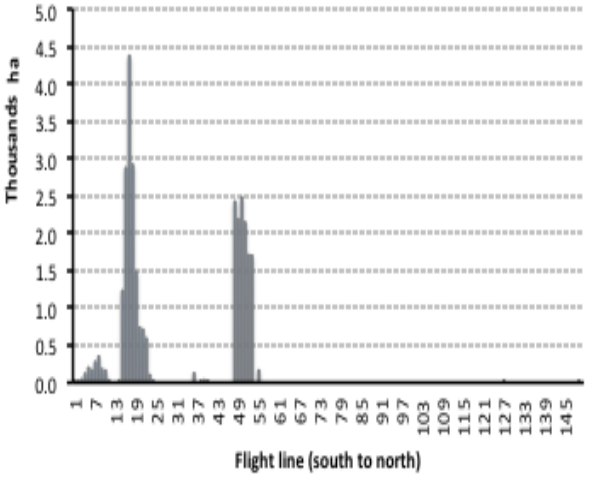




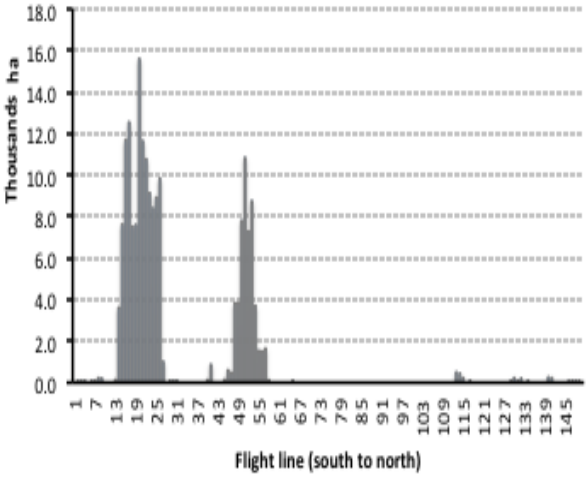
WSB-1999



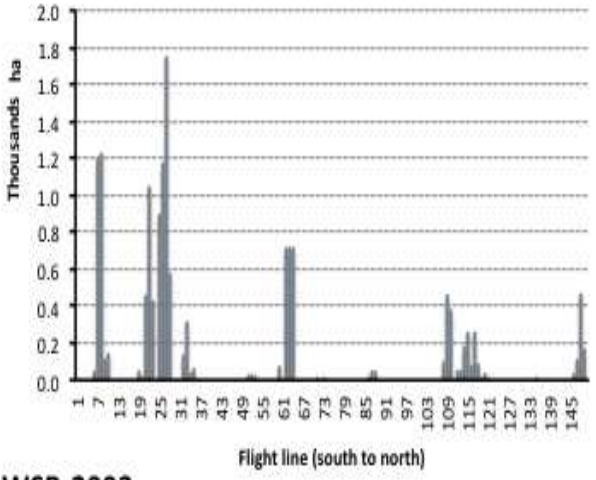
WSB-2000



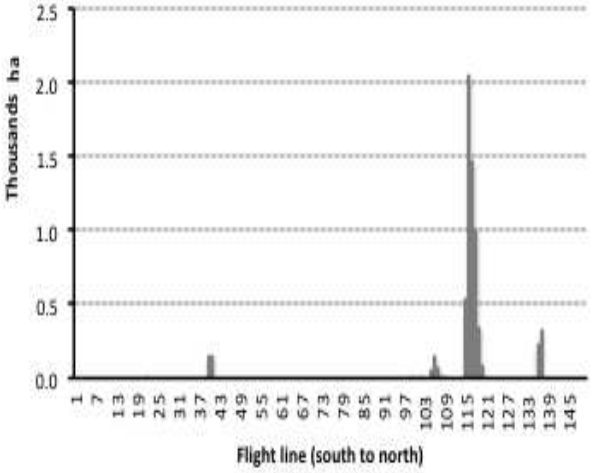
WSB-2001



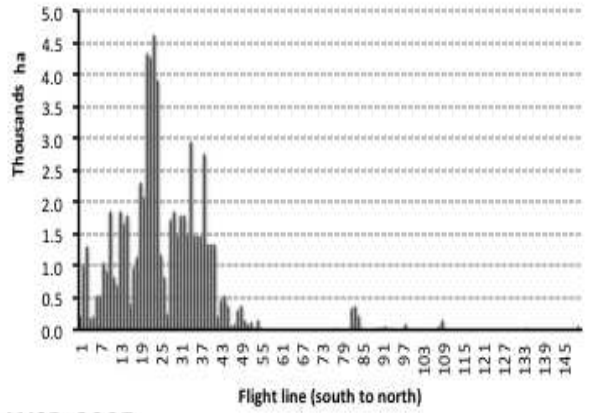
WSB-2002



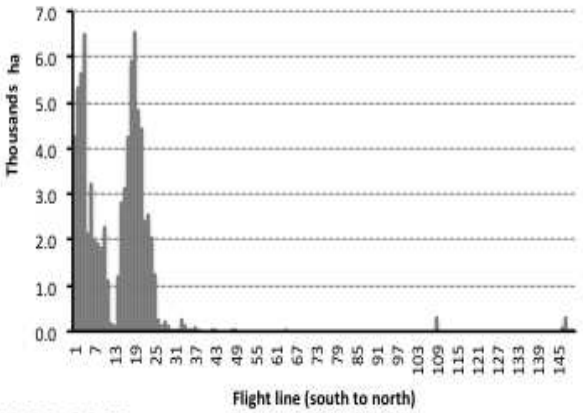
WSB-2003



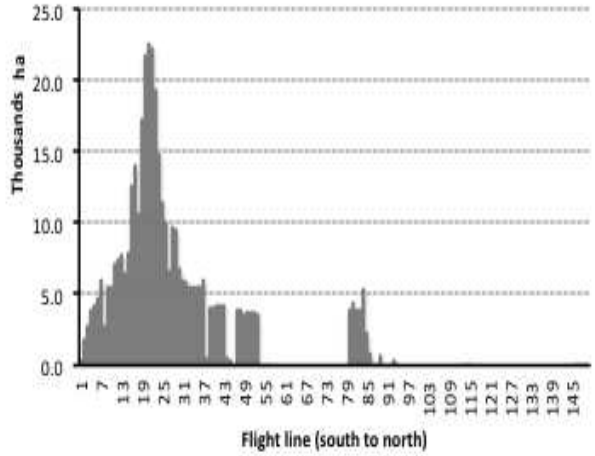
WSB-2004



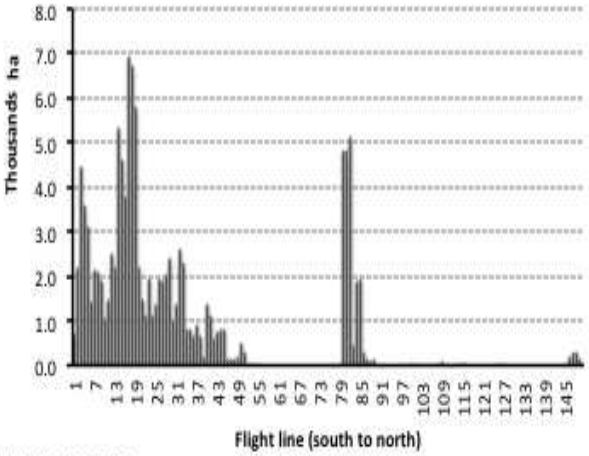
WSB-2005



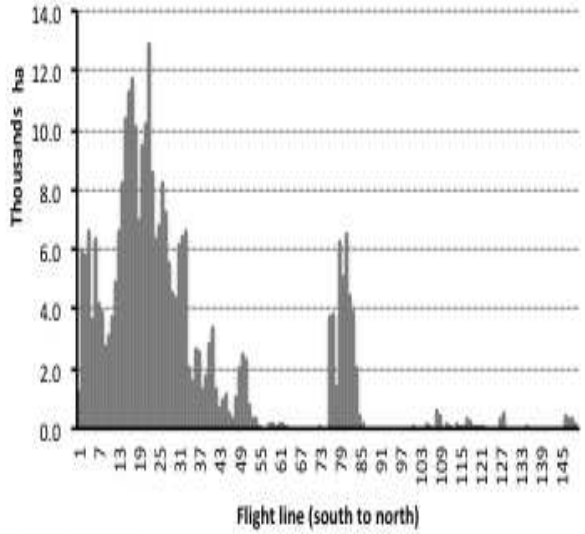
WSB-2006



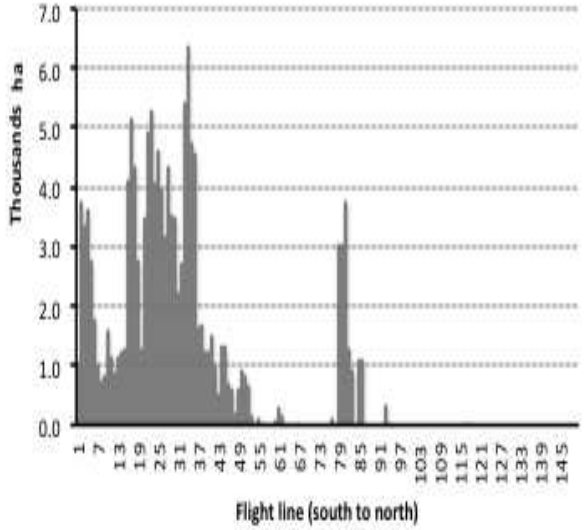
WSB-2007



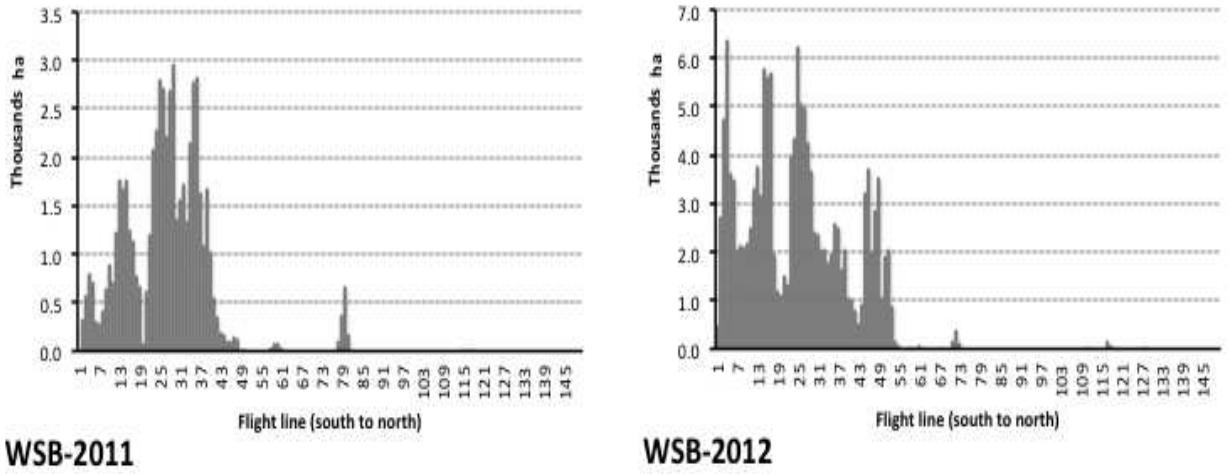
WSB-2008



WSB-2009



WSB-2010



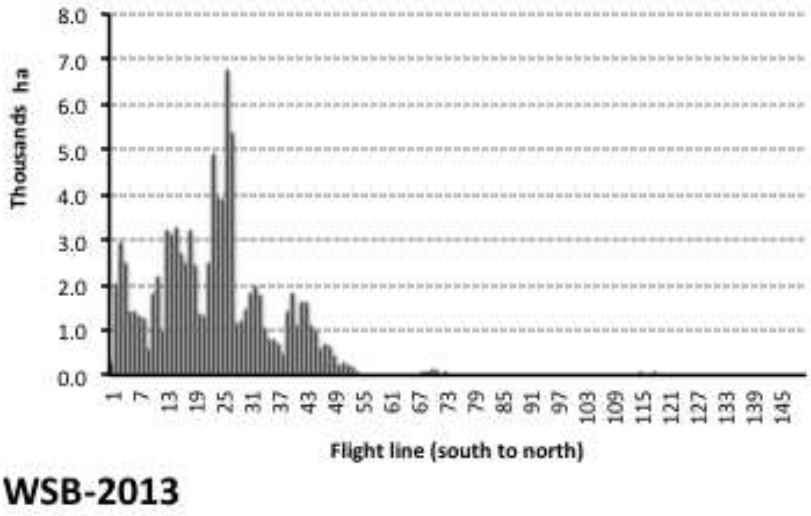
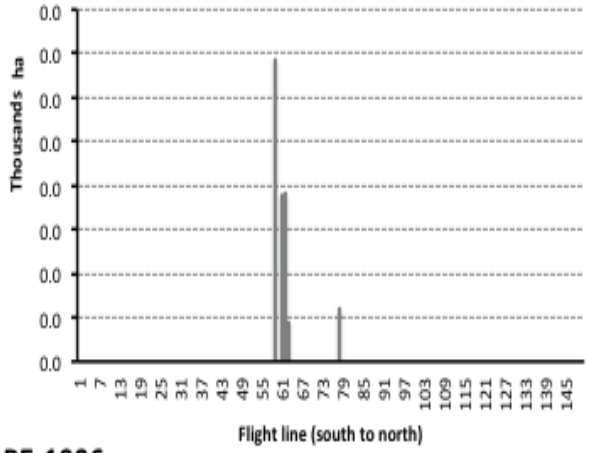
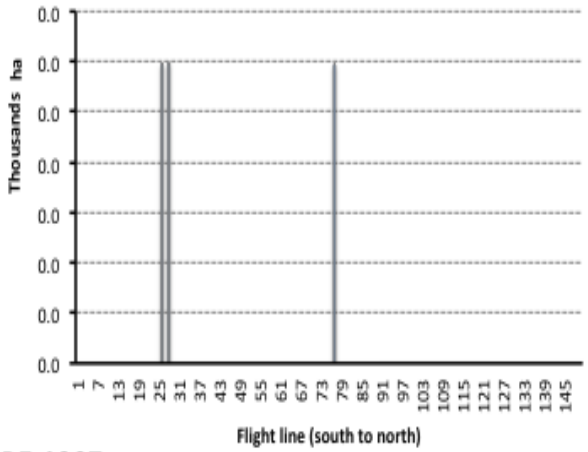


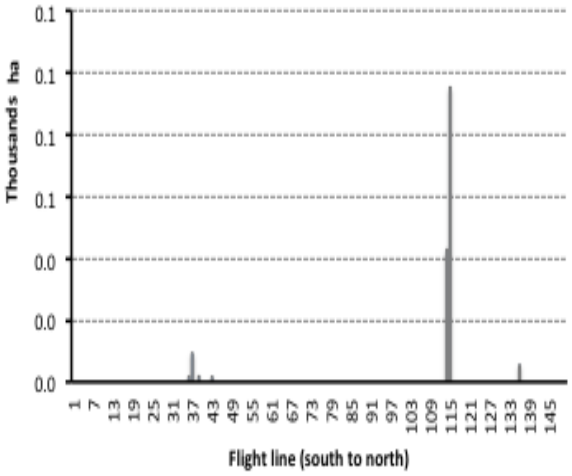
Figure A9. Yearly spatio-temporal distribution of area damaged caused by WSB. The x-axis is number of transect (flight line), the y-axis is damage area (thousand ha)



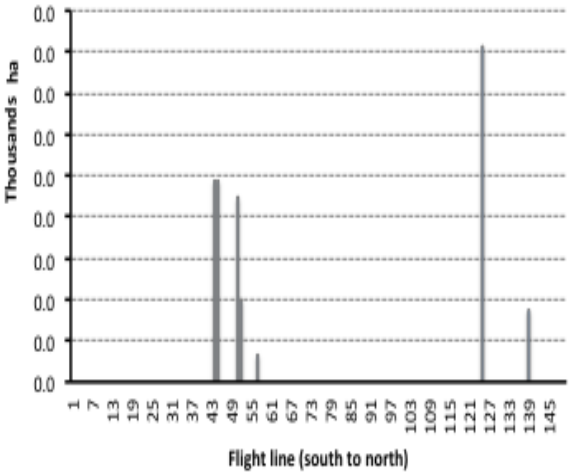
PE-1996



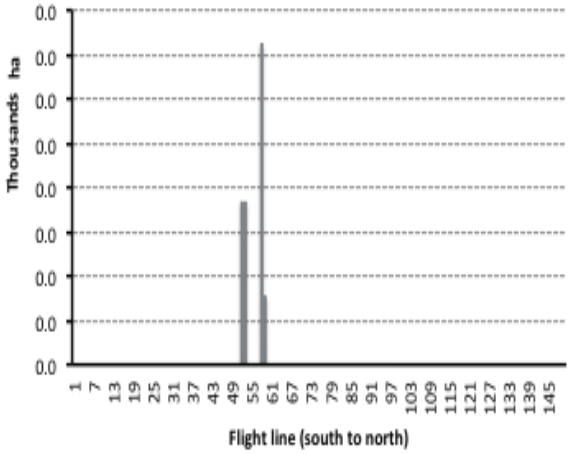
PE-1997



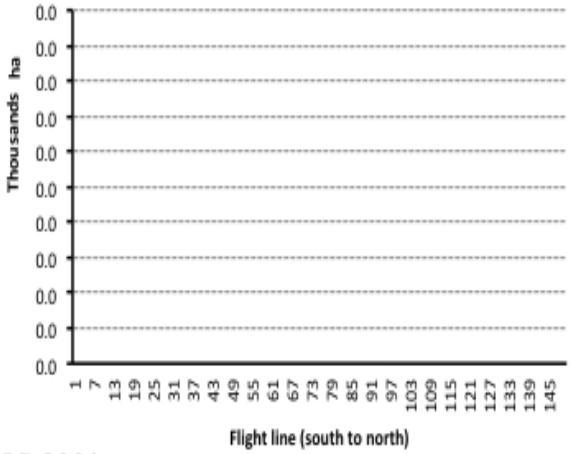
PE-1998



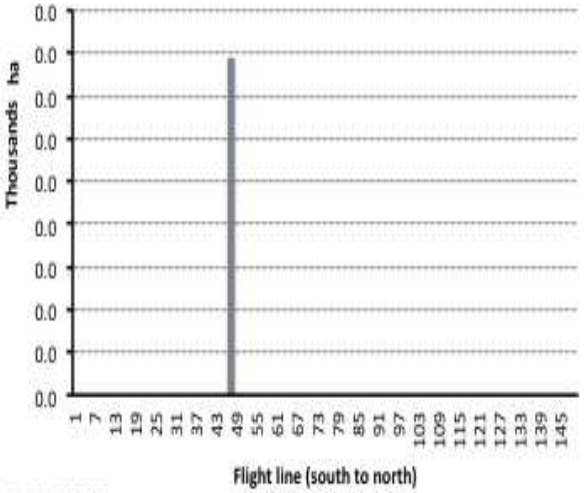
PE-1999



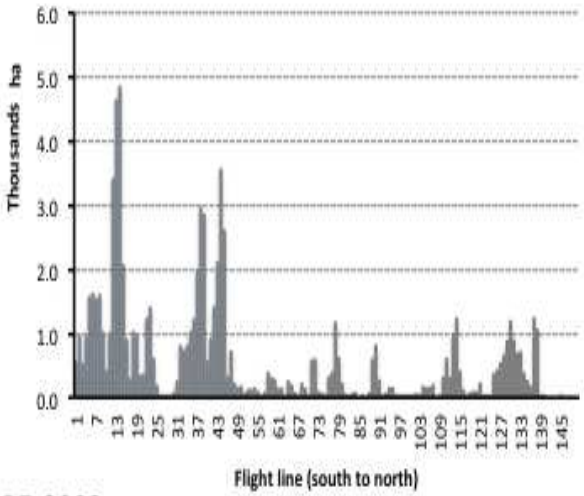
PE-2000



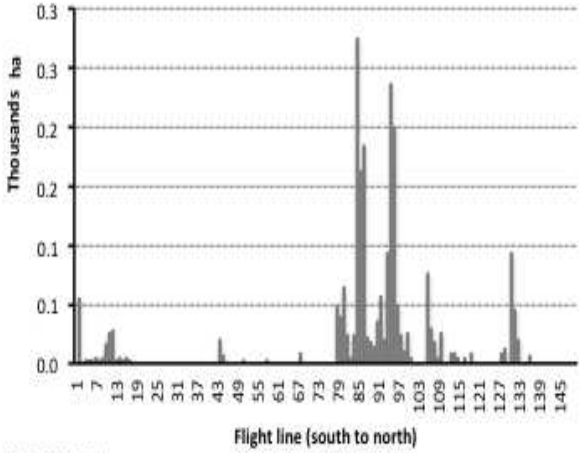
PE-2001



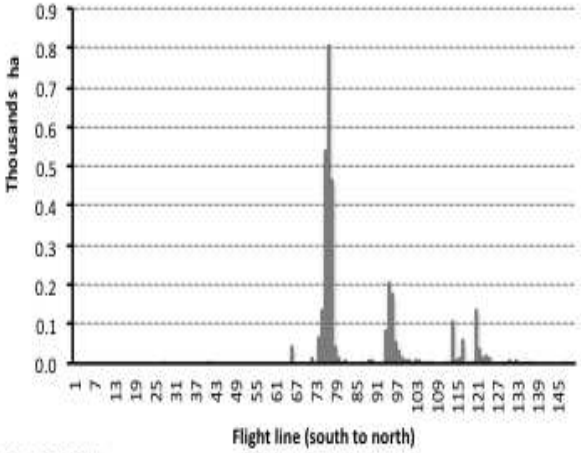
PE-2002



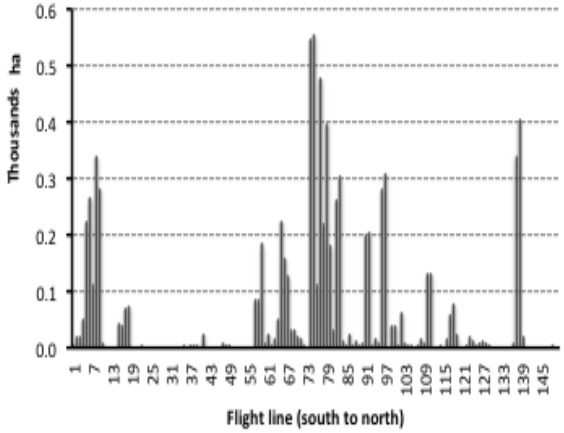
PE-2003



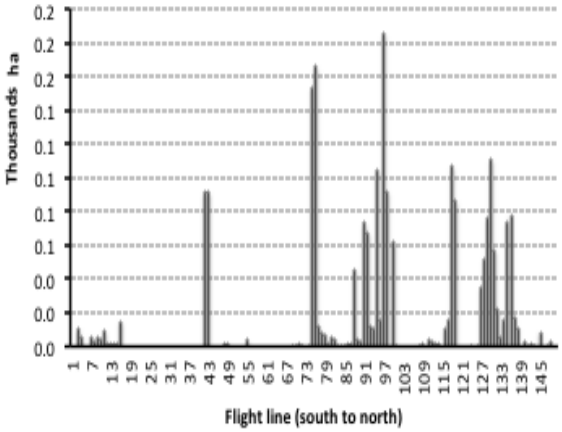
PE-2004



PE-2005



PE-2006



PE-2007

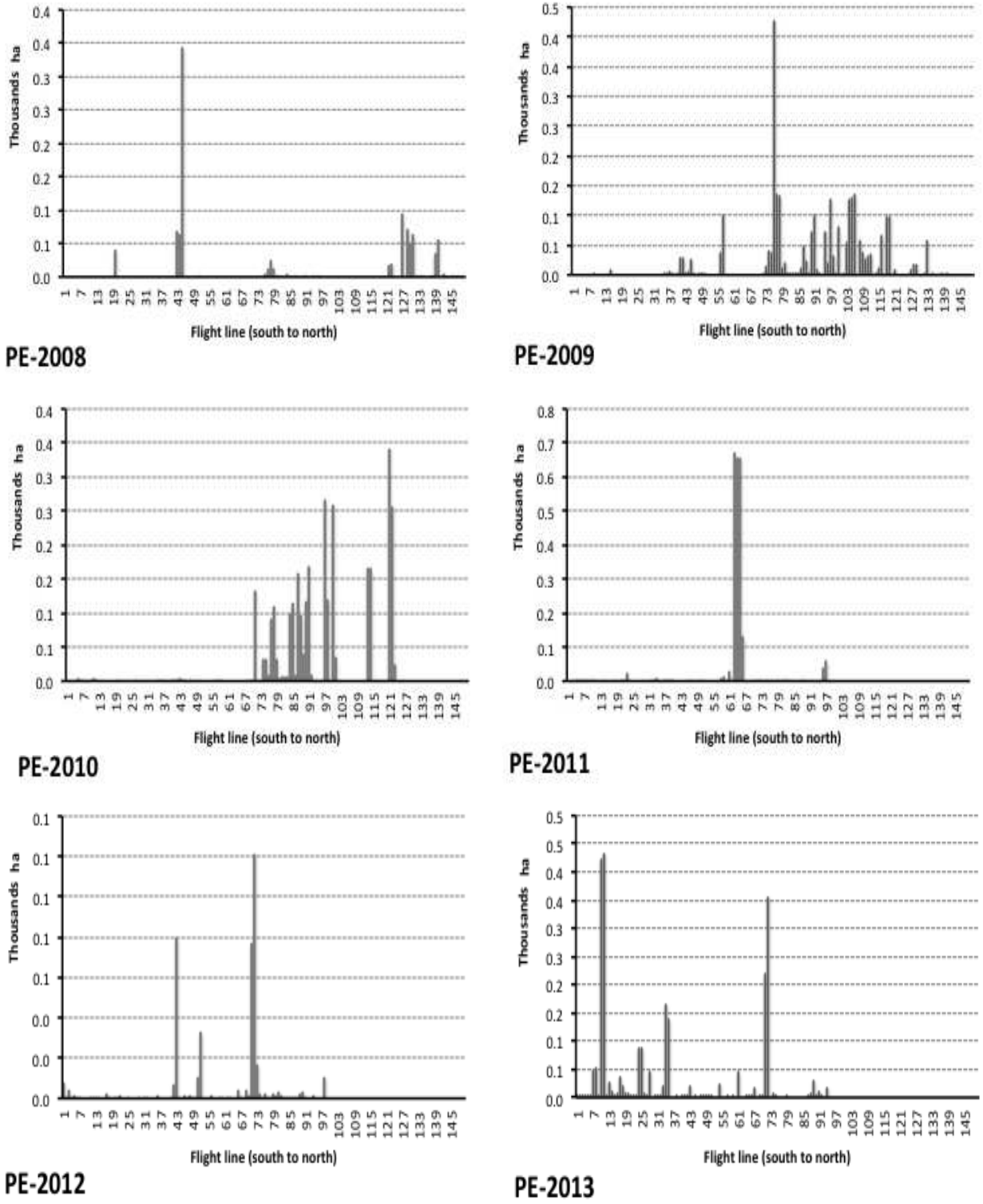
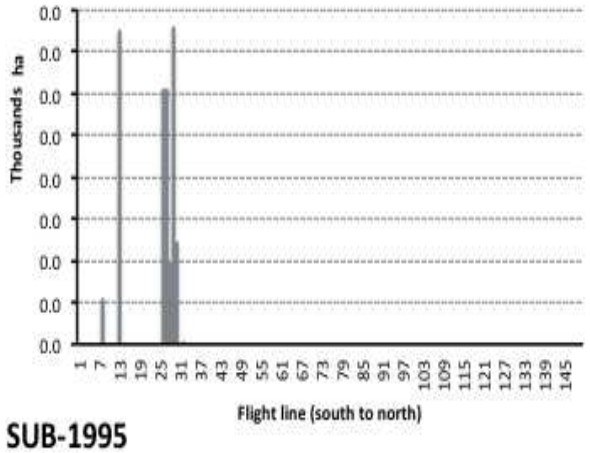
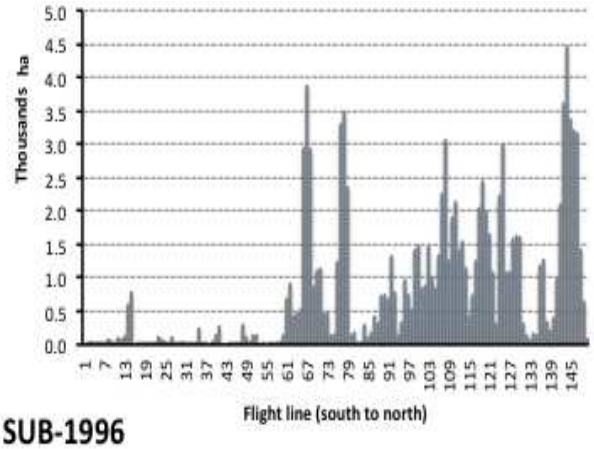


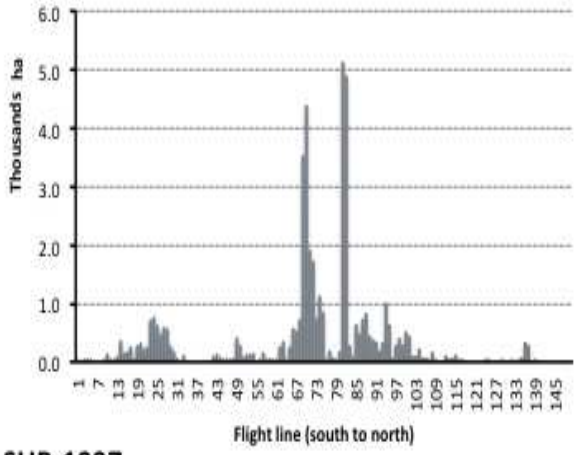
Figure A10. Yearly spatio-temporal distribution of area damaged caused by PE. The x-axis is number of transect (flight line), the y-axis is damage area (thousand ha)



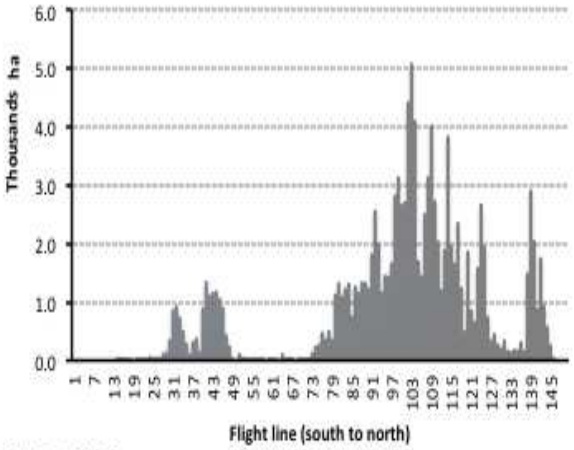
SUB-1995



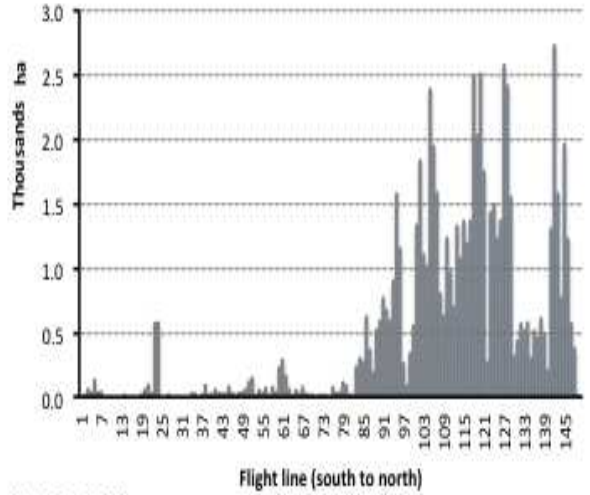
SUB-1996



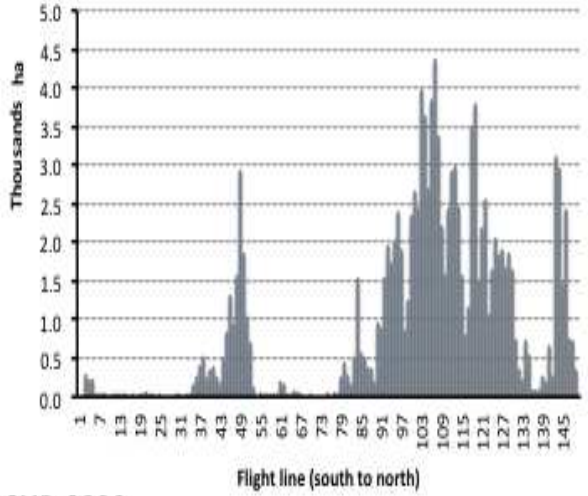
SUB-1997



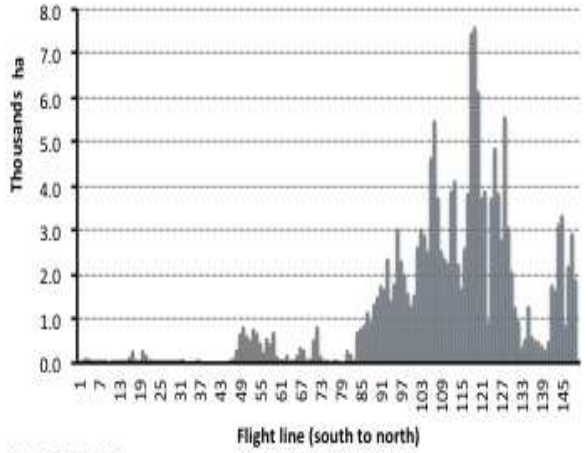
SUB-1998



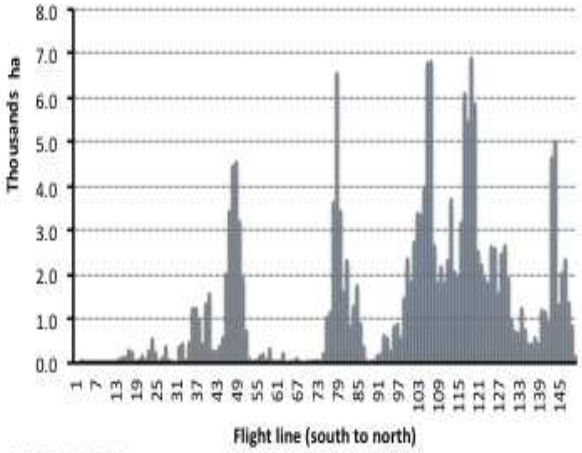
SUB-1999



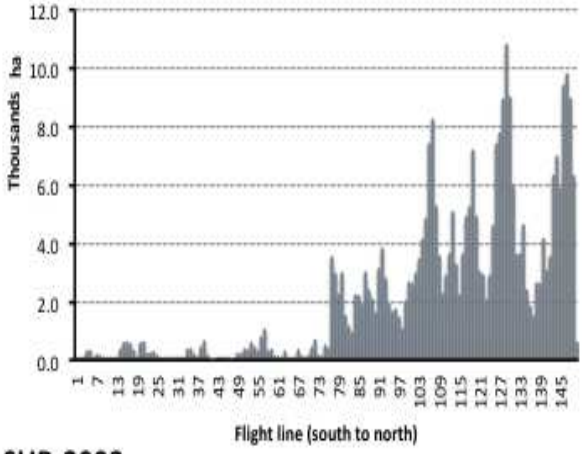
SUB-2000



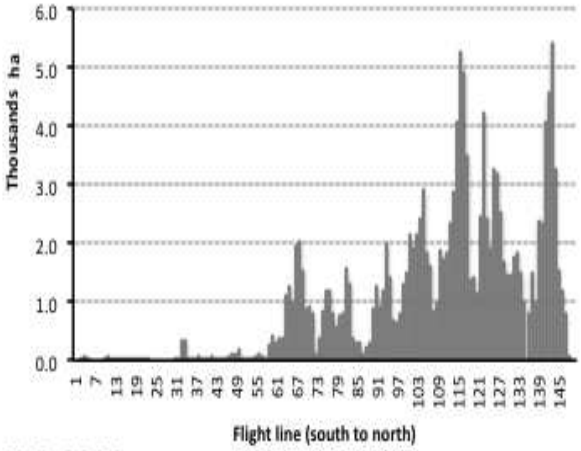
SUB-2001



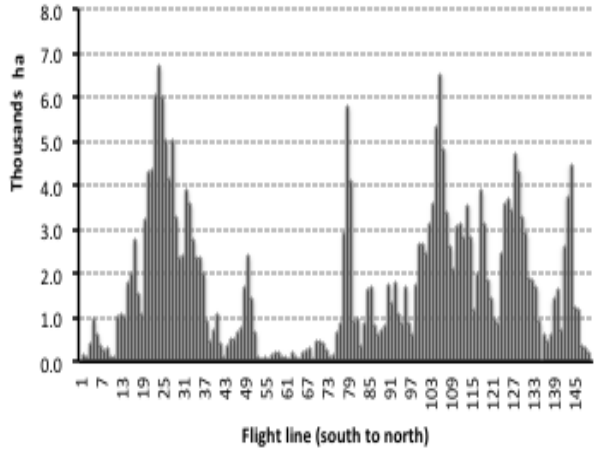
SUB-2002



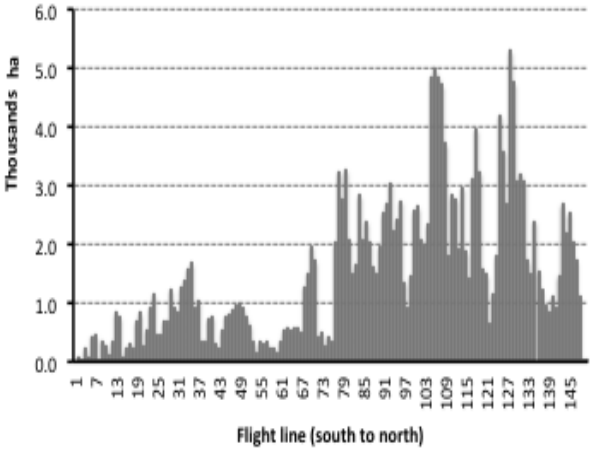
SUB-2003



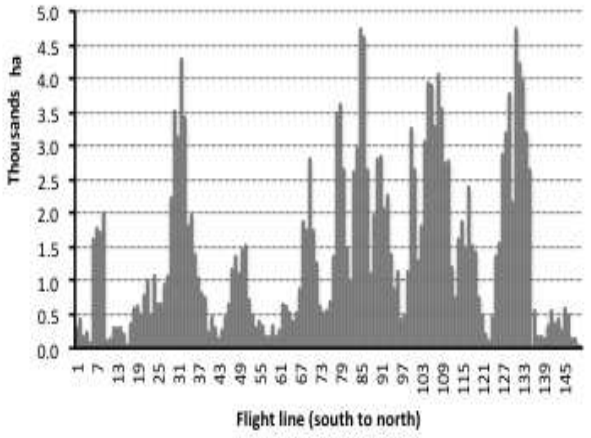
SUB-2004



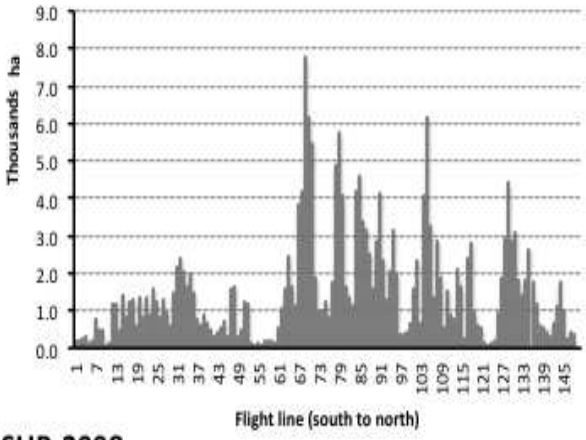
SUB-2005



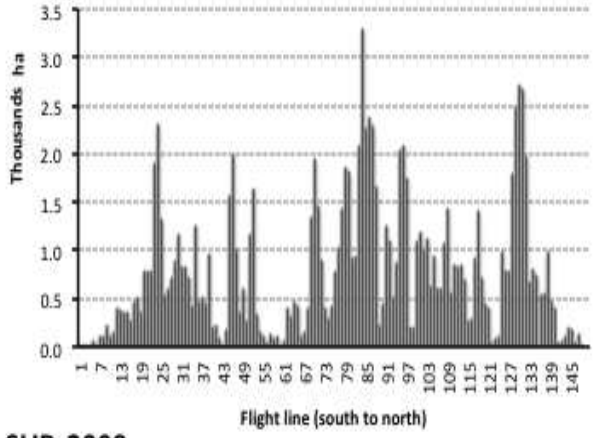
SUB-2006



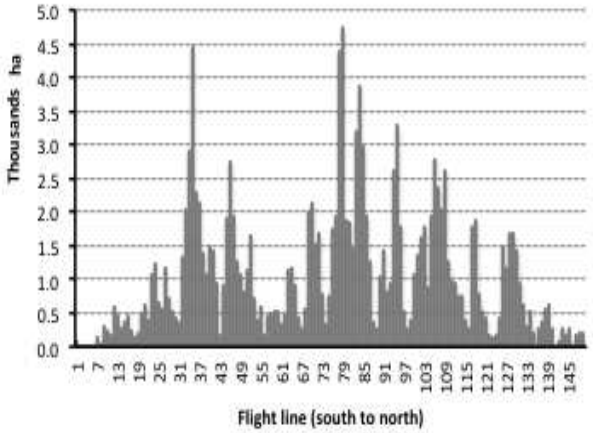
SUB-2007



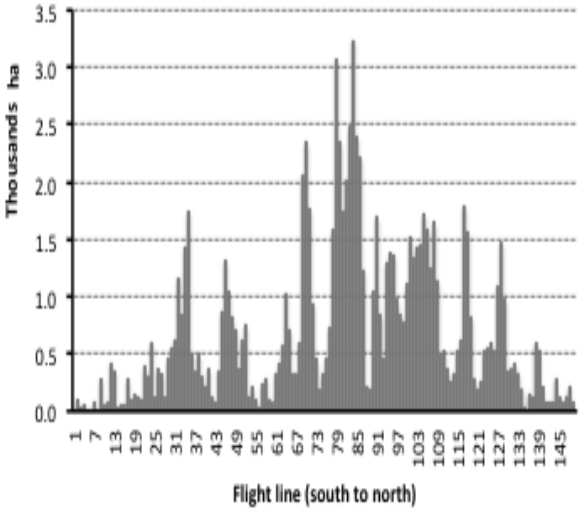
SUB-2008



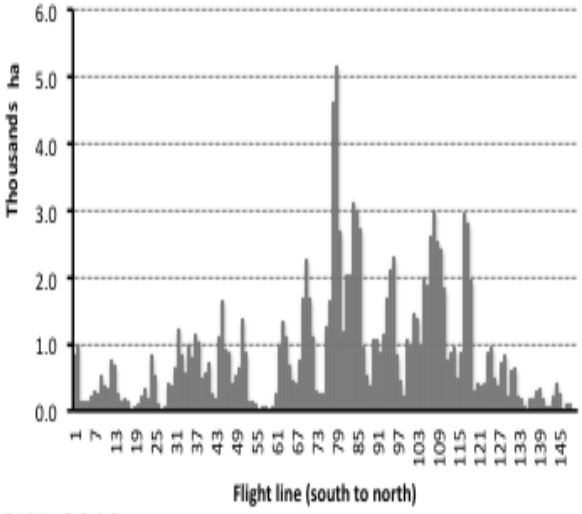
SUB-2009



SUB-2010



SUB-2011



SUB-2012

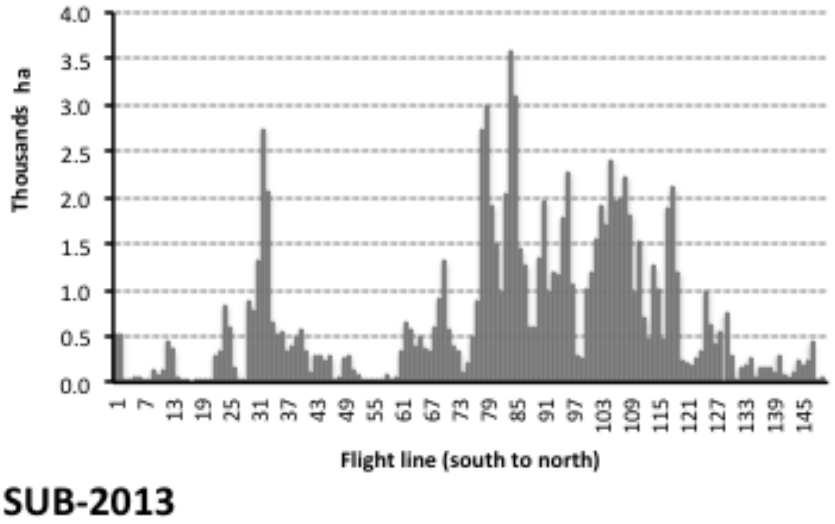
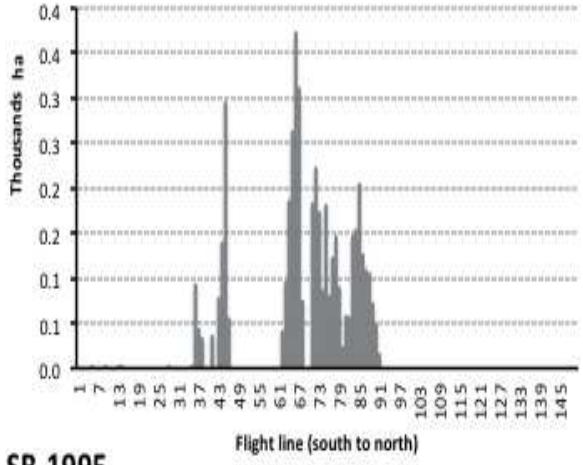
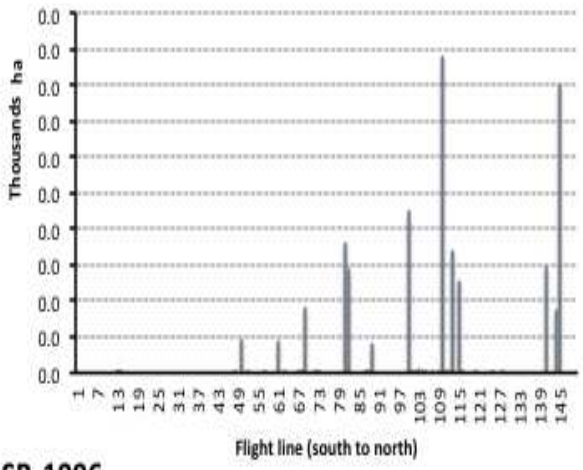


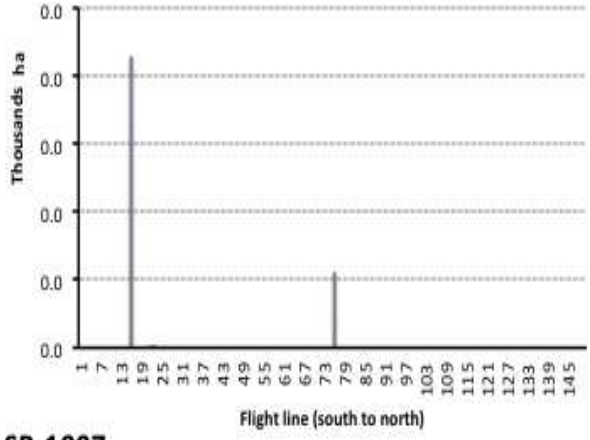
Figure A11. Yearly spatio-temporal distribution of area damaged caused by SUB. The x-axis is number of transect (flight line), the y-axis is damage area (thousand ha)



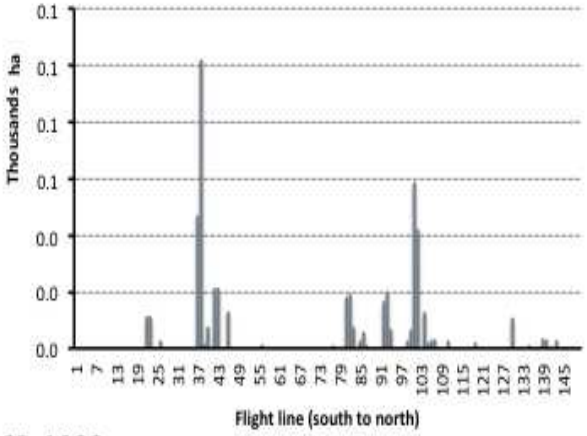
SB-1995



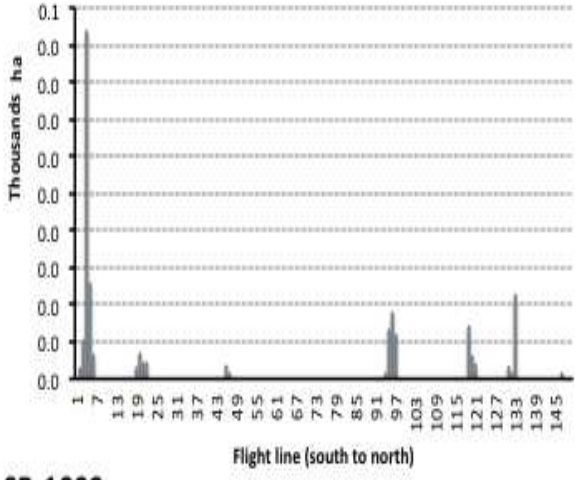
SB-1996



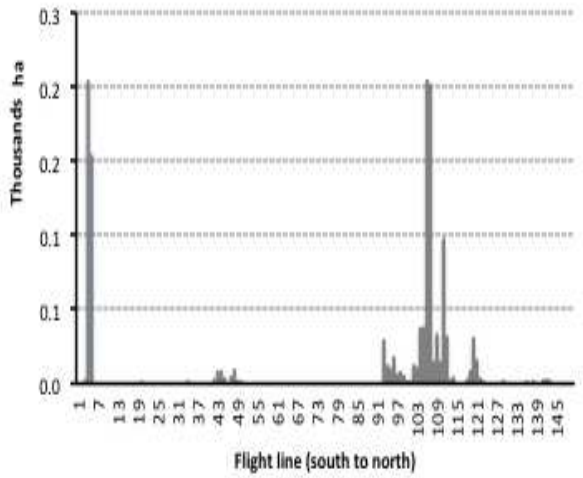
SB-1997



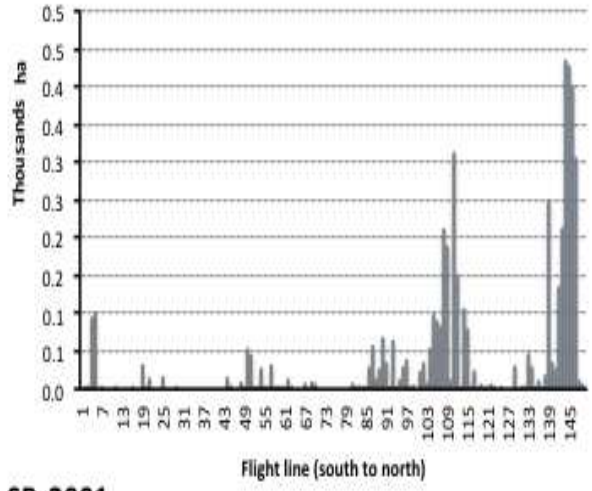
SB-1998



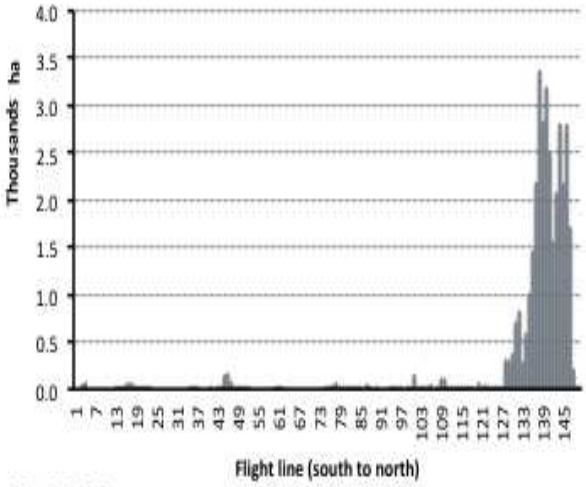
SB-1999



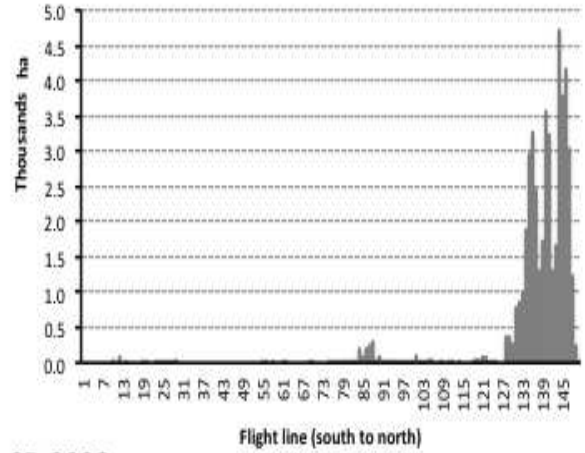
SB-2000



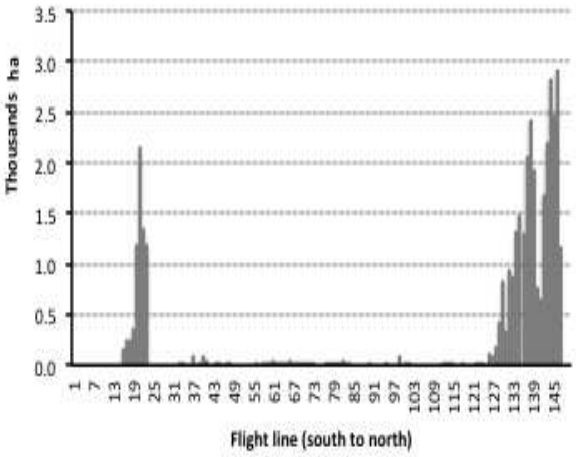
SB-2001



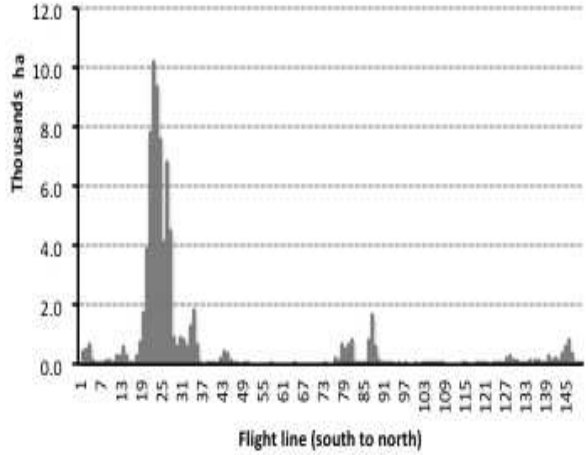
SB-2002



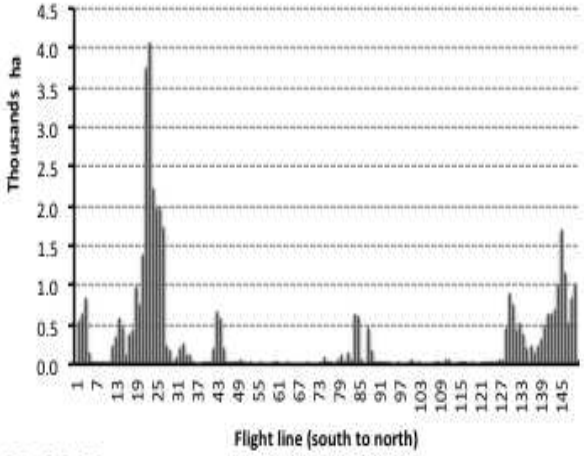
SB-2003



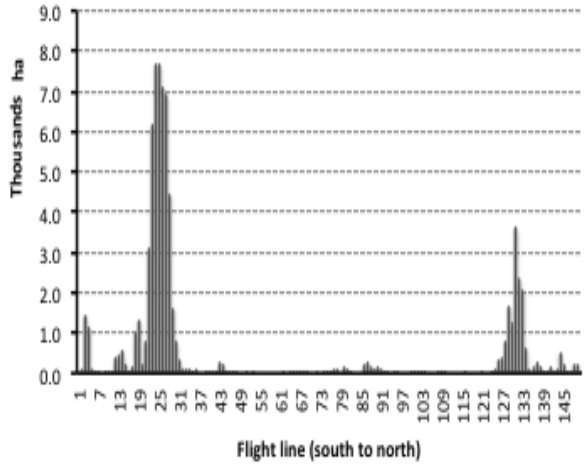
SB-2004



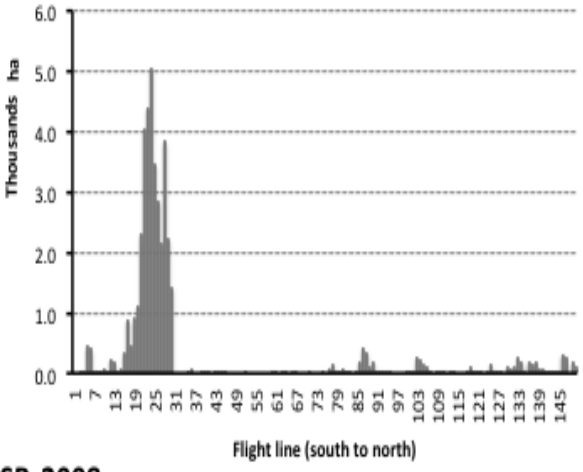
SB-2005



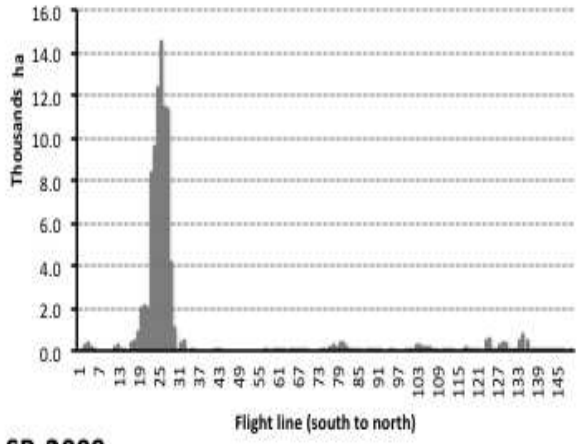
SB-2006



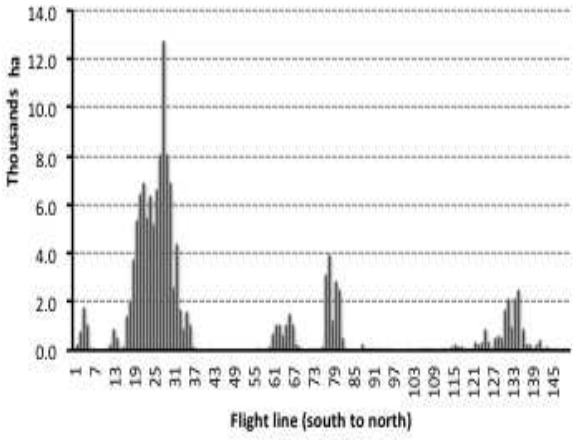
SB-2007



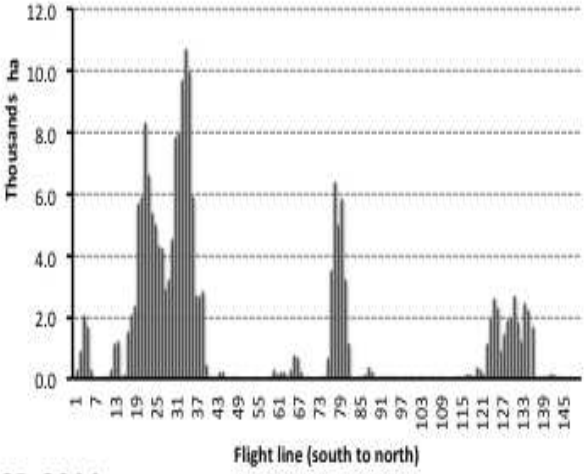
SB-2008



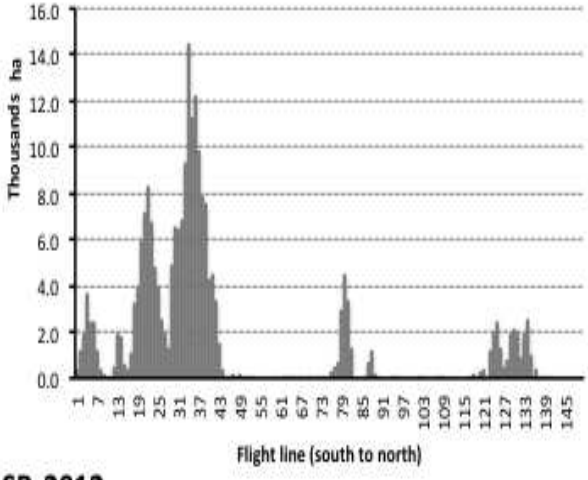
SB-2009



SB-2010



SB-2011



SB-2012

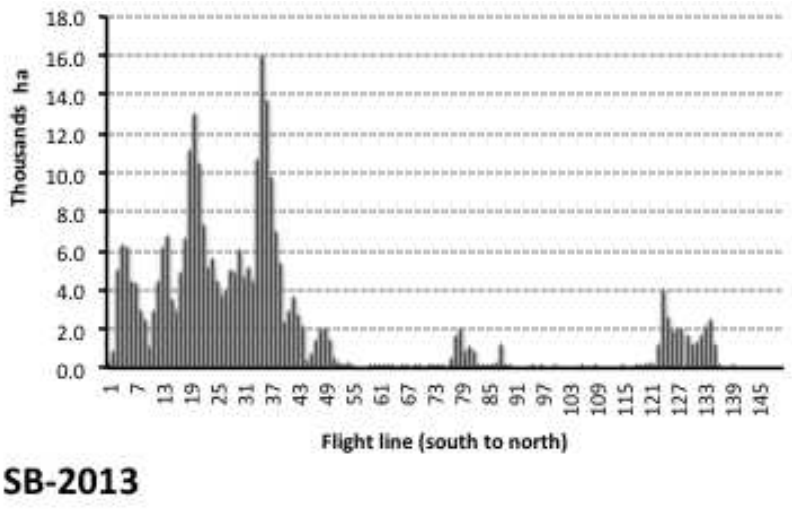
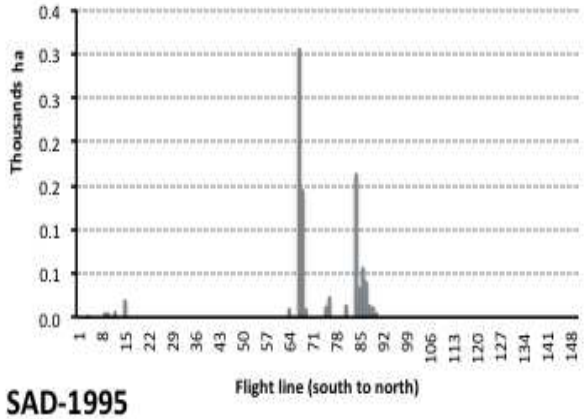
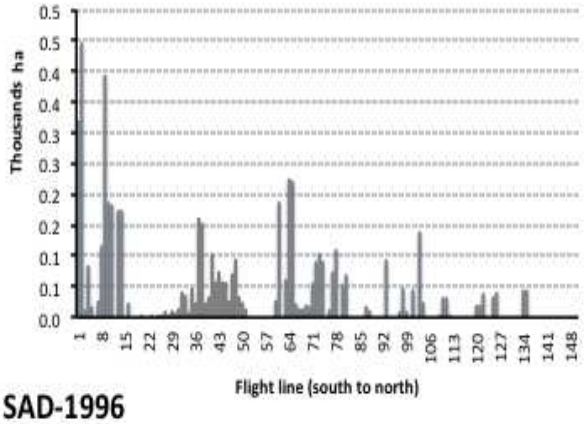


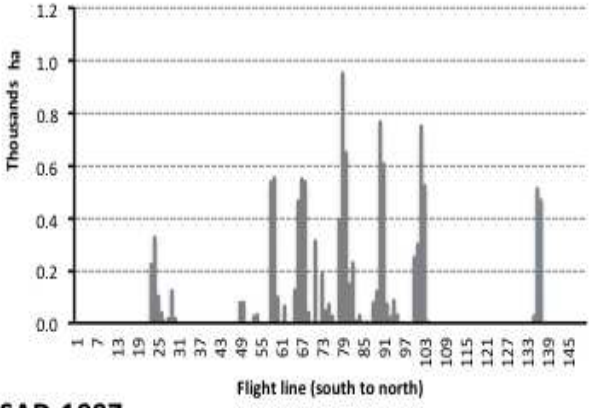
Figure A12. Yearly spatio-temporal distribution of area damaged caused by SB. The x-axis is number of transect (flight line), the y-axis is damage area (thousand ha)



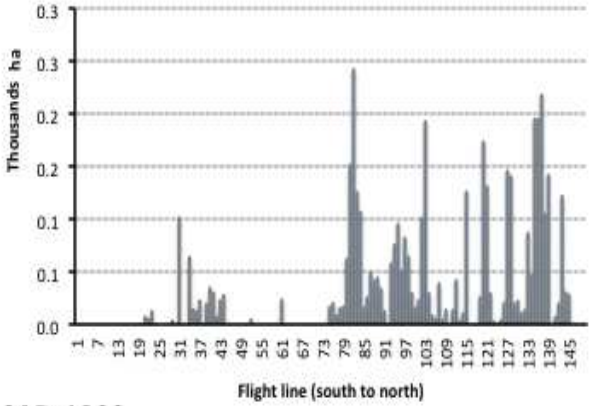
SAD-1995



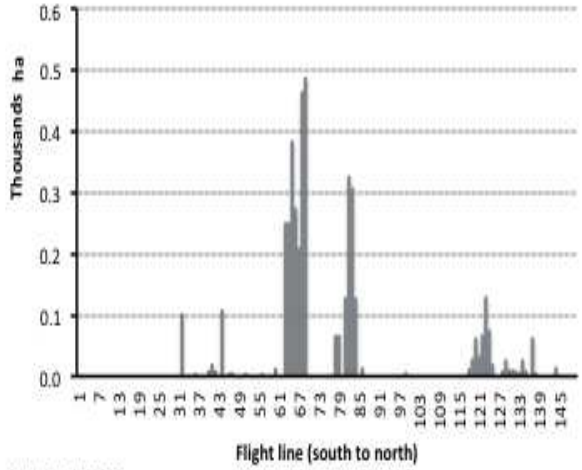
SAD-1996



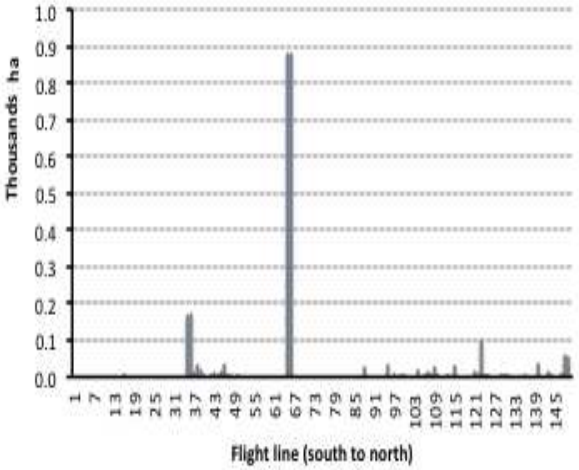
SAD-1997



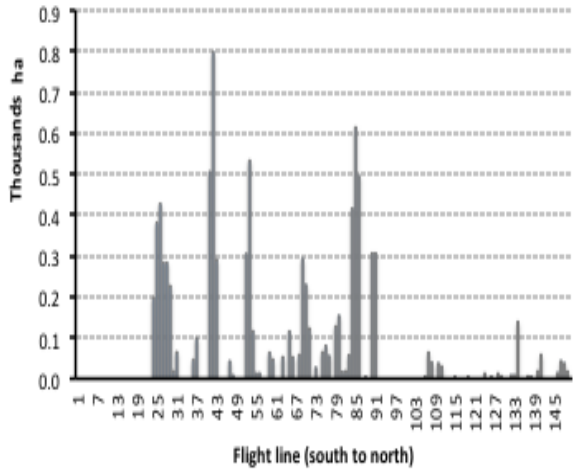
SAD-1998



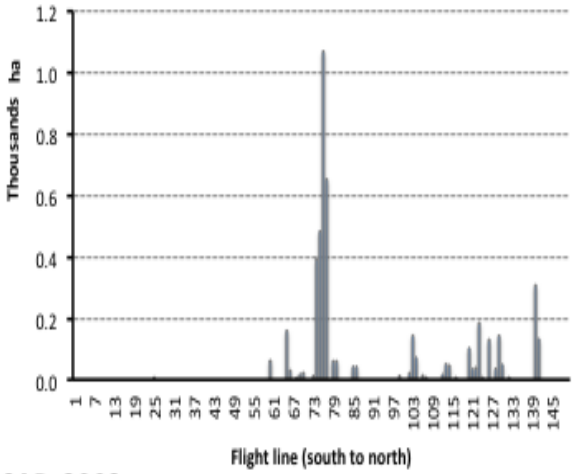
SAD-1999



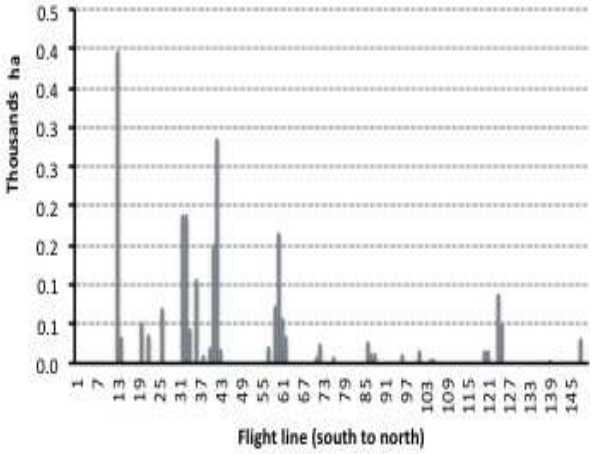
SAD-2000



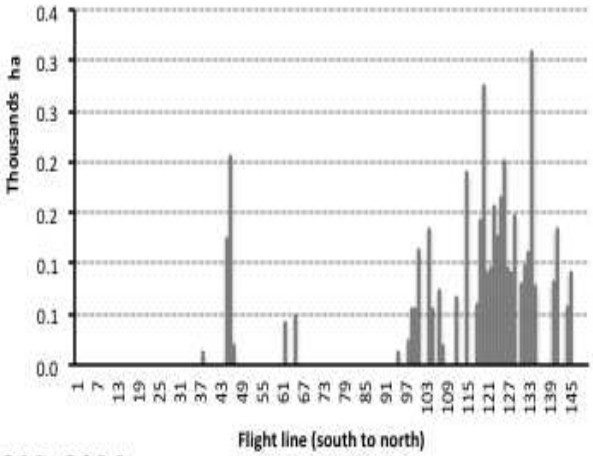
SAD-2001



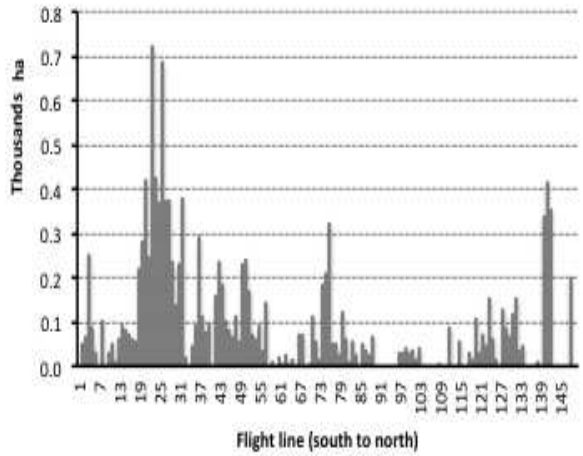
SAD-2002



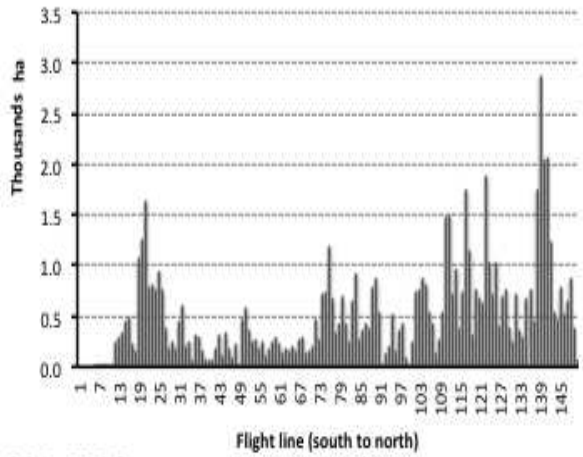
SAD-2003



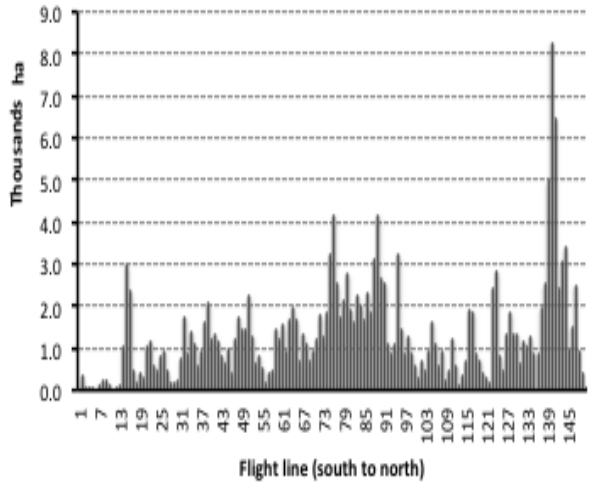
SAD-2004



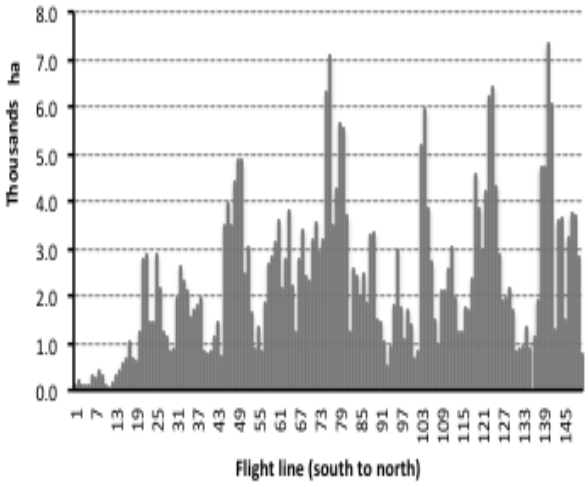
SAD-2005



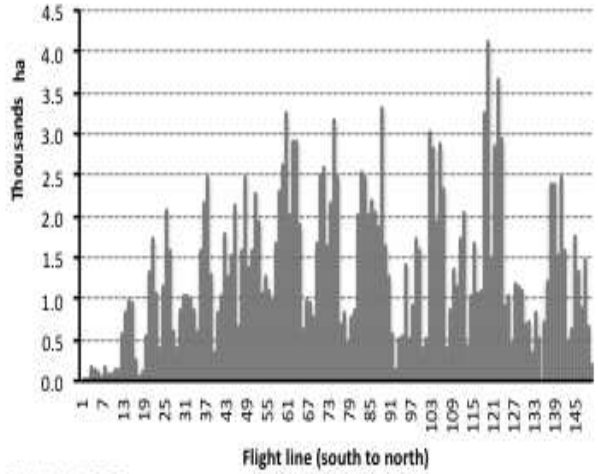
SAD-2006



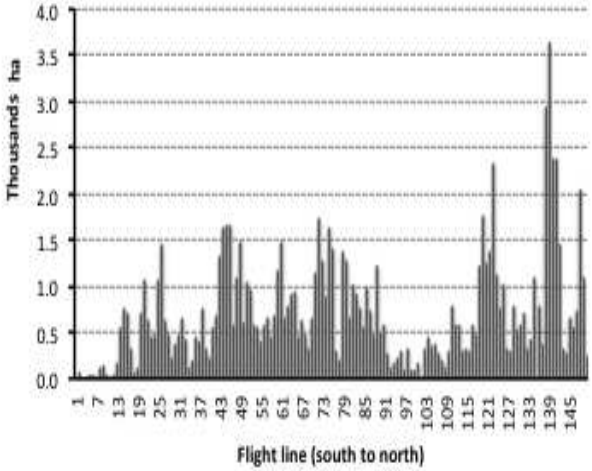
SAD-2007



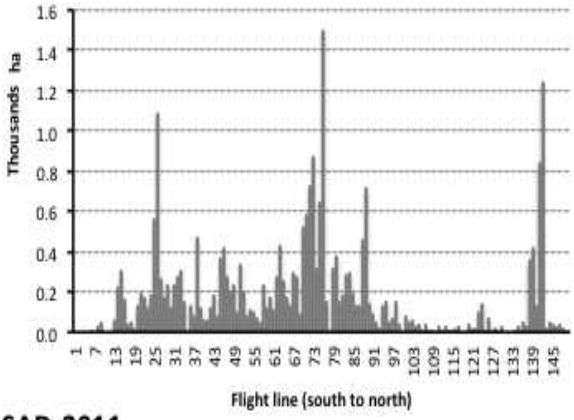
SAD-2008



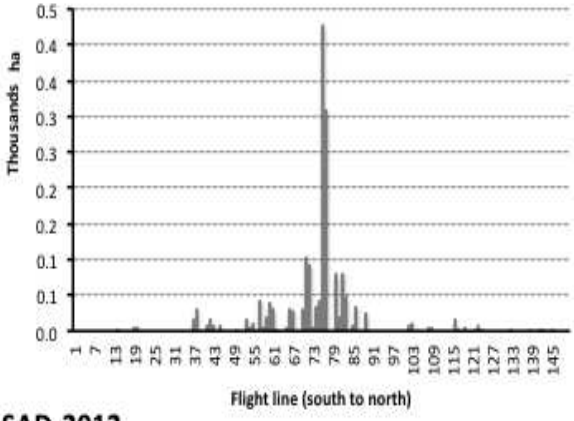
SAD-2009



SAD-2010



SAD-2011



SAD-2012

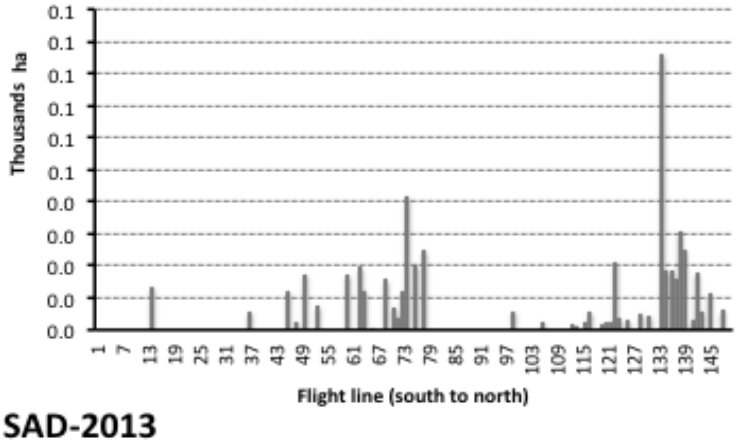


Figure A13. Yearly spatio-temporal distribution of area damaged caused by SAD. The x-axis is number of transect (flight line), the y-axis is damage area (thousand ha)

Table A1a. Test for non-randomness of area with damaged of Aspen forest type

Year	Z_{stat}	Spatial distribution	Year	Zstat	Spatial distribution
1994	-	-	2004	-8.1948	Clustering
1995	-5.672	Clustering	2005	-3.1458	Clustering
1996	-5.8577	Clustering	2006	-11.1202	Clustering
1997	-7.5025	Clustering	2007	-2.0053	Clustering
1998	-8.0016	Clustering	2008	NaN	Uniform
1999	-5.5957	Clustering	2009	-8.6023	Clustering
2000	-5.4995	Clustering	2010	-8.6023	Clustering
2001	-2.3913	Clustering	2011	-7.1699	Clustering
2002	-4.5298	Clustering	2012	-1.7901	Random
2003	-3.7443	Clustering	2013	0	Random

Table A1b. Test for non-randomness of area with damaged of mixed-conifers forest type

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	-12.1244	Clustering	2004	-4.284	Clustering
1995	-7.6273	Clustering	2005	-9.6748	Clustering
1996	-8.5324	Clustering	2006	-8.6023	Clustering
1997	-2.2314	Clustering	2007	-8.6023	Clustering
1998	-7.4851	Clustering	2008	NaN	uniform
1999	-8.4182	Clustering	2009	NaN	uniform
2000	-7.286	Clustering	2010	-8.6023	Clustering
2001	-4.5258	Clustering	2011	-8.6023	Clustering
2002	-7.7451	Clustering	2012	-8.6023	Clustering
2003	-3.5563	Clustering	2013	-10.5474	Clustering

Table A1c. Test for non-randomness of area with damaged of western spruce-fir forest type

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	-7.6273	Clustering	2004	-4.284	Clustering
1995	-8.5324	Clustering	2005	-9.6748	Clustering
1996	-8.5324	Clustering	2006	-8.6023	Clustering
1997	-2.2314	Clustering	2007	-8.6023	Clustering
1998	-7.4851	Clustering	2008	NaN	Uniform
1999	-8.4182	Clustering	2009	NaN	Uniform
2000	-7.286	Clustering	2010	-8.6023	Clustering
2001	-4.5258	Clustering	2011	-8.6023	Clustering
2002	-7.7451	Clustering	2012	-8.6023	Clustering
2003	-3.5563	Clustering	2013	-10.5474	Clustering

Table A2a. Test for non-randomness of area with damaged caused by Douglas-fir beetle (DFB)

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	-10.3103	Clustering	2004	-5.9551	Clustering
1995	-7.687	Clustering	2005	-7.7451	Clustering
1996	-8.1194	Clustering	2006	-8.0989	Clustering
1997	-7.8145	Clustering	2007	-6.9224	Clustering
1998	-5.9265	Clustering	2008	-9.0731	Clustering
1999	-5.5957	Clustering	2009	-7.1958	Clustering
2000	-5.5915	Clustering	2010	-7.8984	Clustering
2001	-4.5751	Clustering	2011	-8.6483	Clustering
2002	-3.9936	Clustering	2012	-5.9551	Clustering
2003	-5.3562	Clustering	2013	-5.8319	Clustering

Table A2b. Test for non-randomness of area with damaged caused by mountain pine beetle (MPB)

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	-11.328	Clustering	2004	-11.0543	Clustering
1995	-10.4852	Clustering	2005	NaN	Uniform
1996	-2.0053	Clustering	2006	-6.5628	Clustering
1997	-9.0202	Clustering	2007	-9.4474	Clustering
1998	-9.5108	Clustering	2008	-4.5515	Clustering
1999	-5.7634	Clustering	2009	-8.1194	Clustering
2000	-8.4885	Clustering	2010	-8.9011	Clustering
2001	-6.5628	Clustering	2011	-8.4955	Clustering
2002	NaN	Uniform	2012	-8.8315	Clustering
2003	-4.1332	Clustering	2013	-6.319	Clustering

Table A2c. Test for non-randomness of area with damaged caused by pine engraver (PE)

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	NaN	Uniform	2004	-7.8768	Clustering
1995	NaN	Uniform	2005	-7.3149	Clustering
1996	-4.8899	Clustering	2006	-6.9018	Clustering
1997	0.2734	Random	2007	-7.0307	Clustering
1998	-3.1823	Clustering	2008	-6.1124	Clustering
1999	-3.1823	Clustering	2009	-6.364	Clustering
2000	-6.3584	Clustering	2010	-7.0394	Clustering
2001	NaN	Uniform	2011	-5.4862	Clustering
2002	-6.9679	Clustering	2012	-6.7638	Clustering
2003	-7.4006	Clustering	2013	-7.1515	Clustering

Table A2d. Test for non-randomness of area with damaged caused by sudden aspen decline (SAD)

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	-	-	2004	-6.2388	Clustering
1995	-5.672	Clustering	2005	-4.3372	Clustering
1996	-5.8577	Clustering	2006	-11.5579	Clustering
1997	-7.5025	Clustering	2007	-2.0053	Clustering
1998	-8.0016	Clustering	2008	NaN	Uniform
1999	-5.664	Clustering	2009	-8.6023	Clustering
2000	-4.8828	Clustering	2010	-3.3885	Clustering
2001	-3.9769	Clustering	2011	-6.1393	Clustering
2002	-4.9366	Clustering	2012	-4.334	Clustering
2003	-3.1379	Clustering	2013	-2.1329	Clustering

Table A2e. Test for non-randomness of area with damaged caused by spruce beetle (SB)

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	-	-	2004	-6.709	Clustering
1995	-8.645	Clustering	2005	-6.456	Clustering
1996	-3.635	Clustering	2006	-4.440	Clustering
1997	-3.000	Clustering	2007	-5.493	Clustering
1998	-4.215	Clustering	2008	-4.884	Clustering
1999	-7.759	Clustering	2009	-5.796	Clustering
2000	-8.109	Clustering	2010	-6.352	Clustering
2001	-6.603	Clustering	2011	-4.484	Clustering
2002	-7.773	Clustering	2012	-7.221	Clustering
2003	-7.414	Clustering	2013	-6.732	Clustering

Table A2f. Test for non-randomness of area with damaged caused by Subalpine-fir mortality (SUB)

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	-	-	2004	-9.675	Clustering
1995	-5.960	Clustering	2005	-6.965	Clustering
1996	-2.231	Clustering	2006	-8.602	Clustering
1997	-7.485	Clustering	2007	NaN	Uniform
1998	-8.418	Clustering	2008	-10.547	Clustering
1999	-7.286	Clustering	2009	-10.547	Clustering
2000	-4.526	Clustering	2010	-8.602	Clustering
2001	-7.644	Clustering	2011	-8.602	Clustering
2002	-3.556	Clustering	2012	-10.547	Clustering
2003	-4.284	Clustering	2013	-3.389	Clustering

Table A2g. Test for non-randomness of area with damaged caused by (WSB)

Year	Z_{stat}	Spatial distribution	Year	Z_{stat}	Spatial distribution
1994	0.116	Random	2004	-7.166	Clustering
1995	-	-	2005	-9.491	Clustering
1996	-9.825	Clustering	2006	-10.228	Clustering
1997	0.191	Random	2007	-10.485	Clustering
1998	-10.901	Clustering	2008	-9.831	Clustering
1999	-10.371	Clustering	2009	-8.204	Clustering
2000	-7.894	Clustering	2010	-9.824	Clustering
2001	-8.878	Clustering	2011	-10.764	Clustering
2002	-7.772	Clustering	2012	-9.804	Clustering
2003	-7.025	Clustering	2013	-10.764	Clustering

Table A3. Accumulative area damaged caused by seven causal agents and disorders overtime. Unit: ha

Year	Causal agents and disorders							
	Comb.	MPB	DFB	SB	PE	WSB	SAD	SUB
1994	12269.17	1541.99	7836.58	0.00	0.00	1.56	0.00	0.00
1995	12804.53	2827.62	755.91	3919.11	0.00	0.00	828.82	133.49
1996	111426.29	5275.19	8118.81	83.20	25.73	10505.93	4100.57	75469.58
1997	65466.33	12834.52	1431.94	27.02	0.09	193.16	7907.22	29966.77
1998	199604.78	42363.27	5312.10	326.68	113.99	8910.46	3781.73	89549.53
1999	149824.47	49422.95	2385.33	126.27	21.95	16616.37	2790.16	52881.59
2000	174953.74	56389.13	4972.63	928.27	8.06	8230.79	1614.16	83907.66
2001	212615.01	60947.86	3723.89	3576.55	0.00	14346.78	6298.74	104713.69
2002	381727.10	83981.96	5777.49	21104.69	7.88	50461.29	3170.36	123719.18
2003	764205.57	92072.88	18862.77	28693.08	49487.40	8107.35	1684.82	195034.48
2004	605738.74	173284.23	16460.42	20956.58	1844.71	4249.15	3268.57	98317.91
2005	528388.78	198487.46	13604.43	46402.24	2589.97	26342.94	10863.50	166894.54
2006	560720.36	267776.50	9408.95	26107.18	5636.76	37348.96	54993.31	143080.76
2007	896583.09	395341.39	17998.27	37500.54	1622.70	156856.68	133162.75	125810.98
2008	928447.75	462100.36	11010.40	24848.92	877.44	61897.30	217878.71	133674.05
2009	900327.89	416665.72	9099.75	45210.17	2074.00	154363.25	138278.03	73186.83
2010	1297727.43	706118.32	19256.17	141352.00	2942.01	142480.88	103930.74	154979.59
2011	1175107.76	766625.69	12710.53	188209.94	2327.22	58172.20	24937.74	101458.45
2012	726564.14	191614.15	13503.54	219084.08	424.59	147709.47	1719.93	127884.66
2013	660085.06	74289.85	22575.06	288211.48	2425.06	95773.95	528.44	100327.29

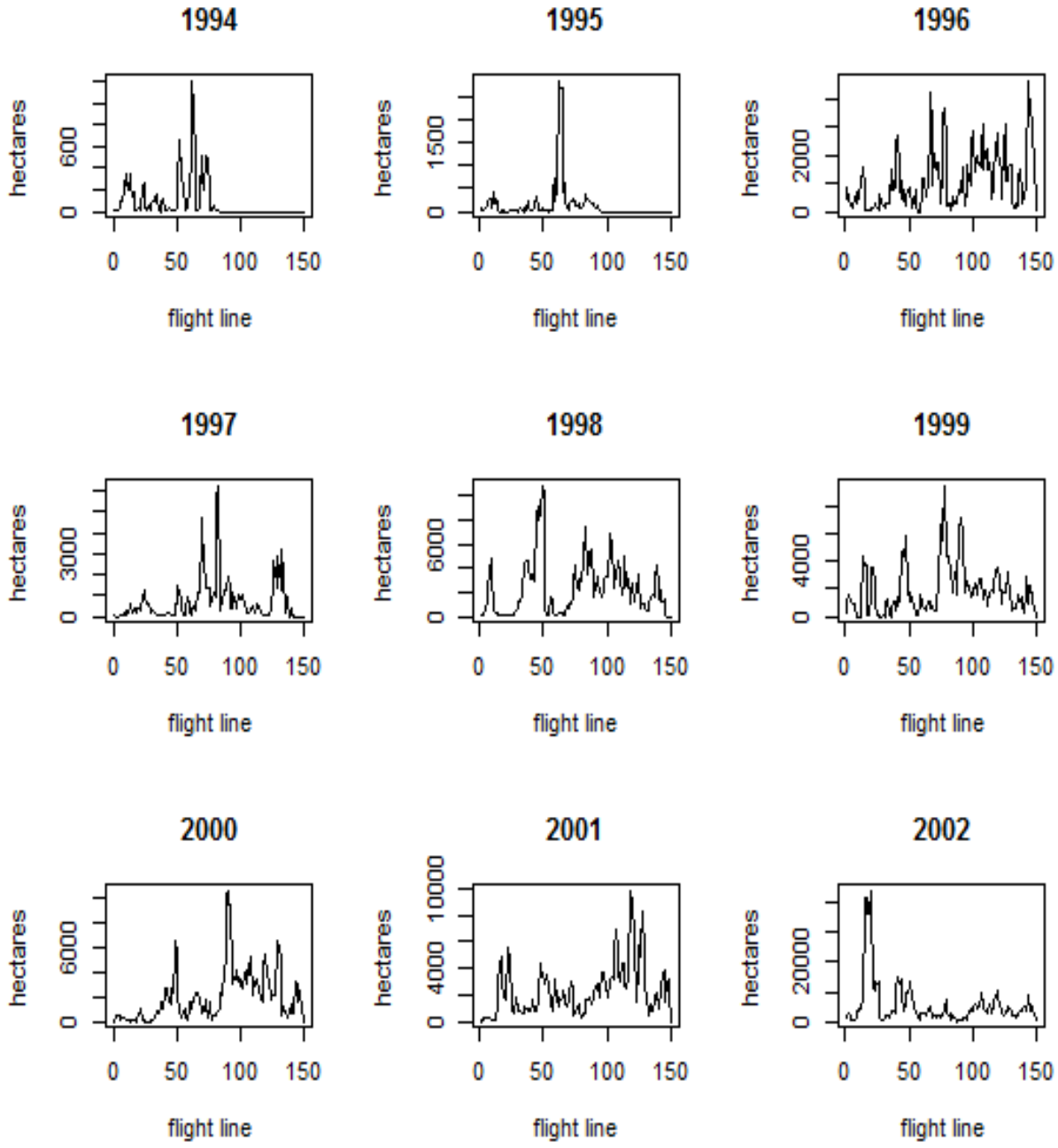
where Comb. – all causal agents and disorders combined, MPB – mountain pine beetle, DFB – Douglas fir beetle, SB – spruce beetle, PE- pine engraver, WSB – western spruce budworm, SAD – sudden aspen decline, and SUB – subalpine fir mortality.

Table A4. Spatial autocorrelation test using Moran's I for area damaged on adjacent transects for different damage agents over time 1994 to 2013. The Moran's I vary in range from -1 to +1. If Moran's I get is negative indicating that area damaged on adjacent transects were cyclic pattern. If Moran's I is positive spatial this would indicates clustering of damage on adjacent transects. A zero Moran's I would suggest damage is spatially independent among transects.

Year	Moran's I							
	Comb.	MPB	DFB	SB	PE	WSB	SAD	SUB
1994	0.76	0.88	0.77	-	-	*-0.01*	-	-
1995	0.79	0.52	0.63	0.78	-	-	0.36	0.39
1996	0.74	0.48	0.73	0.06	0.21	0.79	0.53	0.79
1997	0.68	0.63	0.52	*-0.01*	*-0.02*	*-0.01*	0.59	0.62
1998	0.84	0.81	0.78	0.40	0.36	0.85	0.60	0.86
1999	0.77	0.73	0.73	0.36	0.20	0.84	0.72	0.81
2000	0.84	0.79	0.77	0.53	0.29	0.71	0.48	0.85
2001	0.81	0.76	0.67	0.71	-	0.83	0.62	0.86
2002	0.88	0.79	0.59	0.92	0.49	0.89	0.65	0.82
2003	0.91	0.84	0.82	0.88	0.78	0.62	0.27	0.90
2004	0.96	0.93	0.81	0.88	0.62	0.75	0.53	0.89
2005	0.93	0.94	0.63	0.90	0.74	0.85	0.65	0.85
2006	0.93	0.95	0.59	0.83	0.52	0.93	0.70	0.84
2007	0.91	0.94	0.58	0.91	0.46	0.95	0.72	0.82
2008	0.93	0.95	0.63	0.91	0.19	0.82	0.73	0.73
2009	0.87	0.92	0.65	0.92	0.39	0.92	0.63	0.74
2010	0.92	0.95	0.75	0.89	0.38	0.89	0.67	0.76
2011	0.93	0.95	0.54	0.92	0.71	0.93	0.50	0.81
2012	0.90	0.93	0.71	0.93	0.41	0.89	0.51	0.78
2013	0.92	0.90	0.78	0.91	0.45	0.86	0.19	0.80
Average	*0.86*	*0.83*	*0.68*	*0.72*	*0.42*	*0.75*	*0.56*	*0.79*

where Comb. – all and disorders combined, MPB – mountain pine beetle, DFB – Douglas fir beetle, SB – spruce beetle, PE- pine engraver, WSB – western spruce budworm, SAD – sudden aspen decline, and SUB – subalpine fir mortality.

APPENDIX B - CHAPTER 3



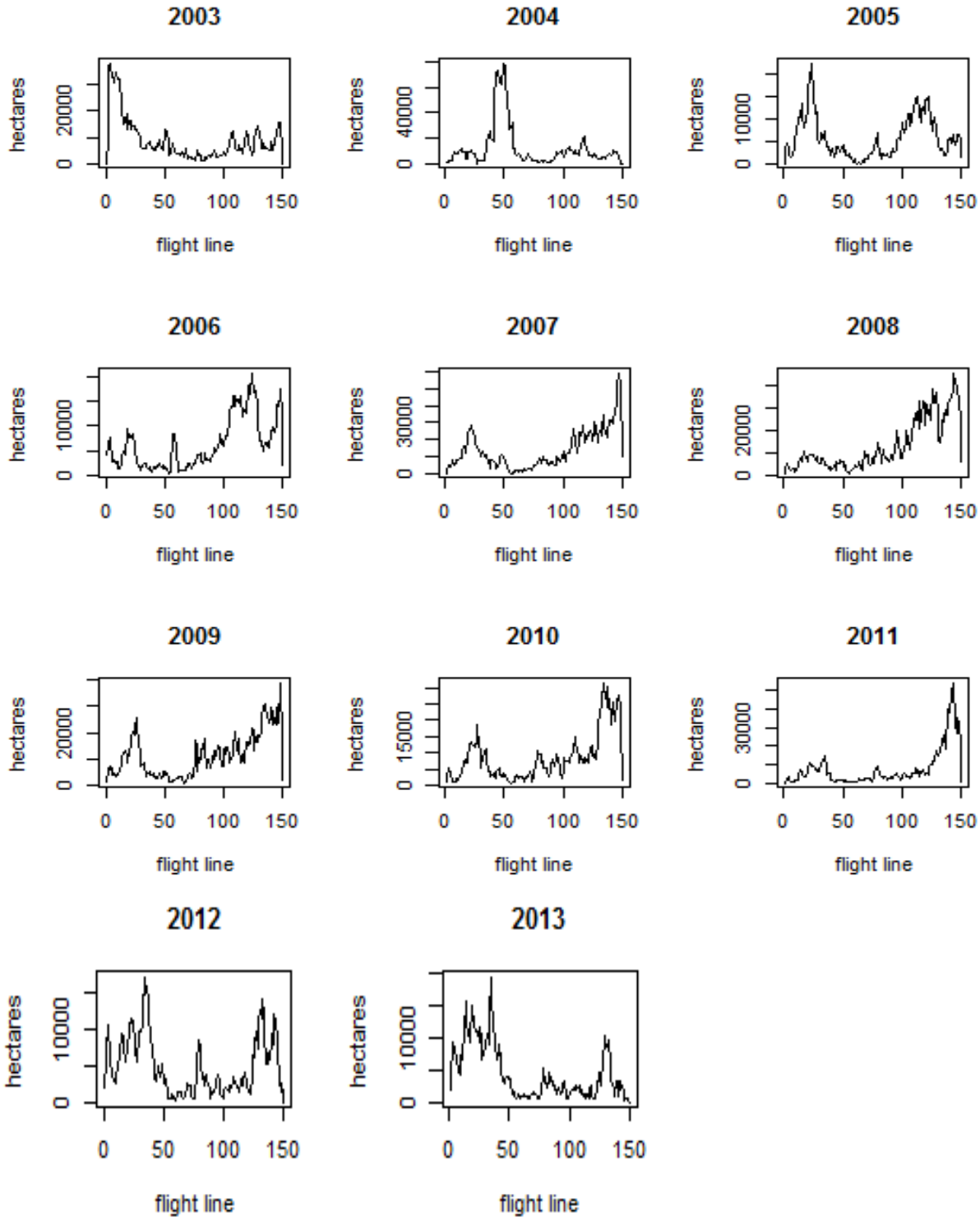
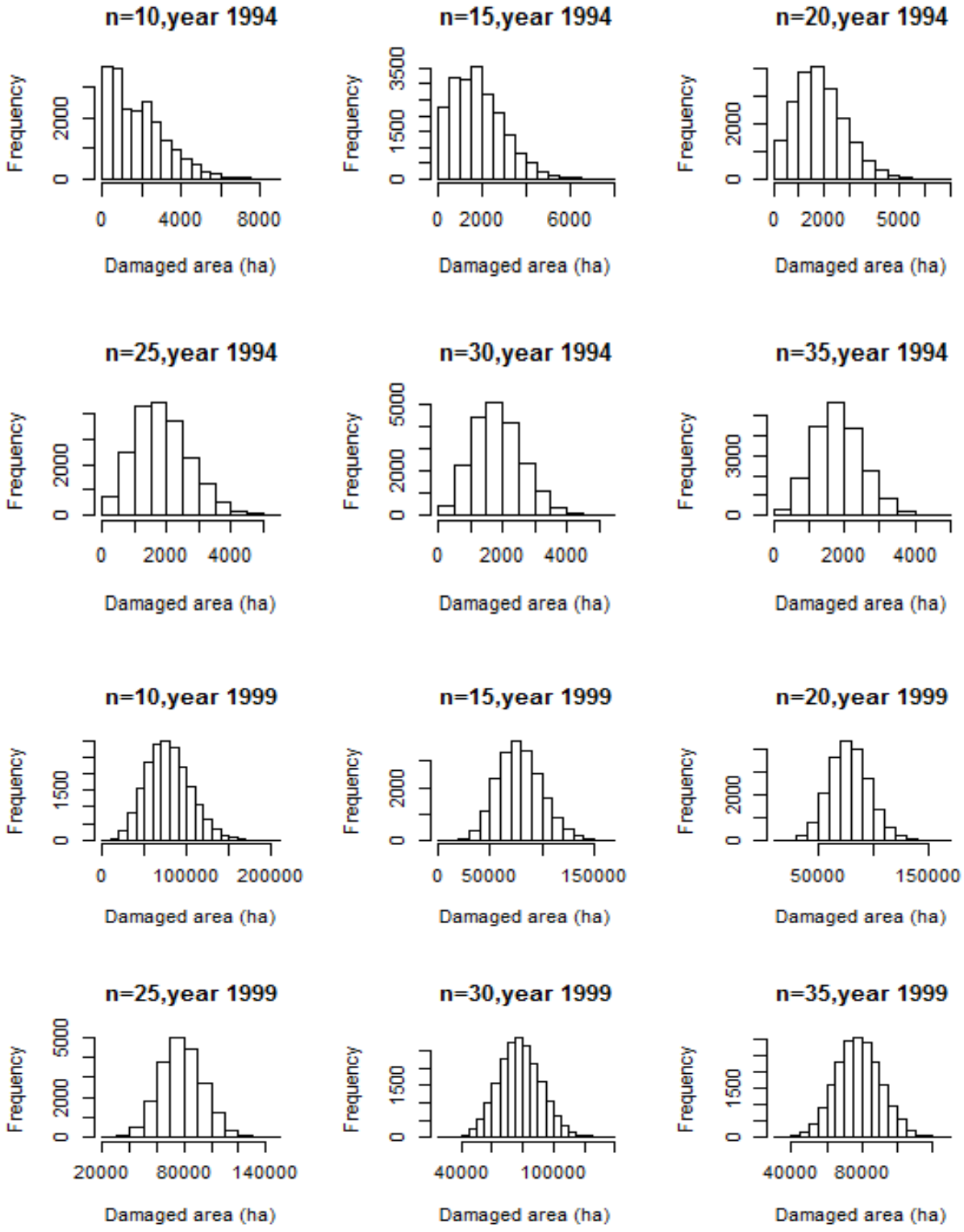


Figure B1. An example of a spatial cyclic patterns distribution of area with damage caused by all causal agents and disorders combined. These distribution patterns resulted very large and unpredictable estimated variances when increasing sample sizes.The x-axis is number of transect (transect), the y-axis is damage area (ha)

SRS - MPB


n=10,year 1994
n=15,year 1994
n=20,year 1994
n=25,year 1994
n=30,year 1994
n=35,year 1994
n=10,year 1999
n=15,year 1999
n=20,year 1999
n=25,year 1999
n=30,year 1999
n=35,year 1999
Frequency
Damaged area (ha)

PPS – MPB

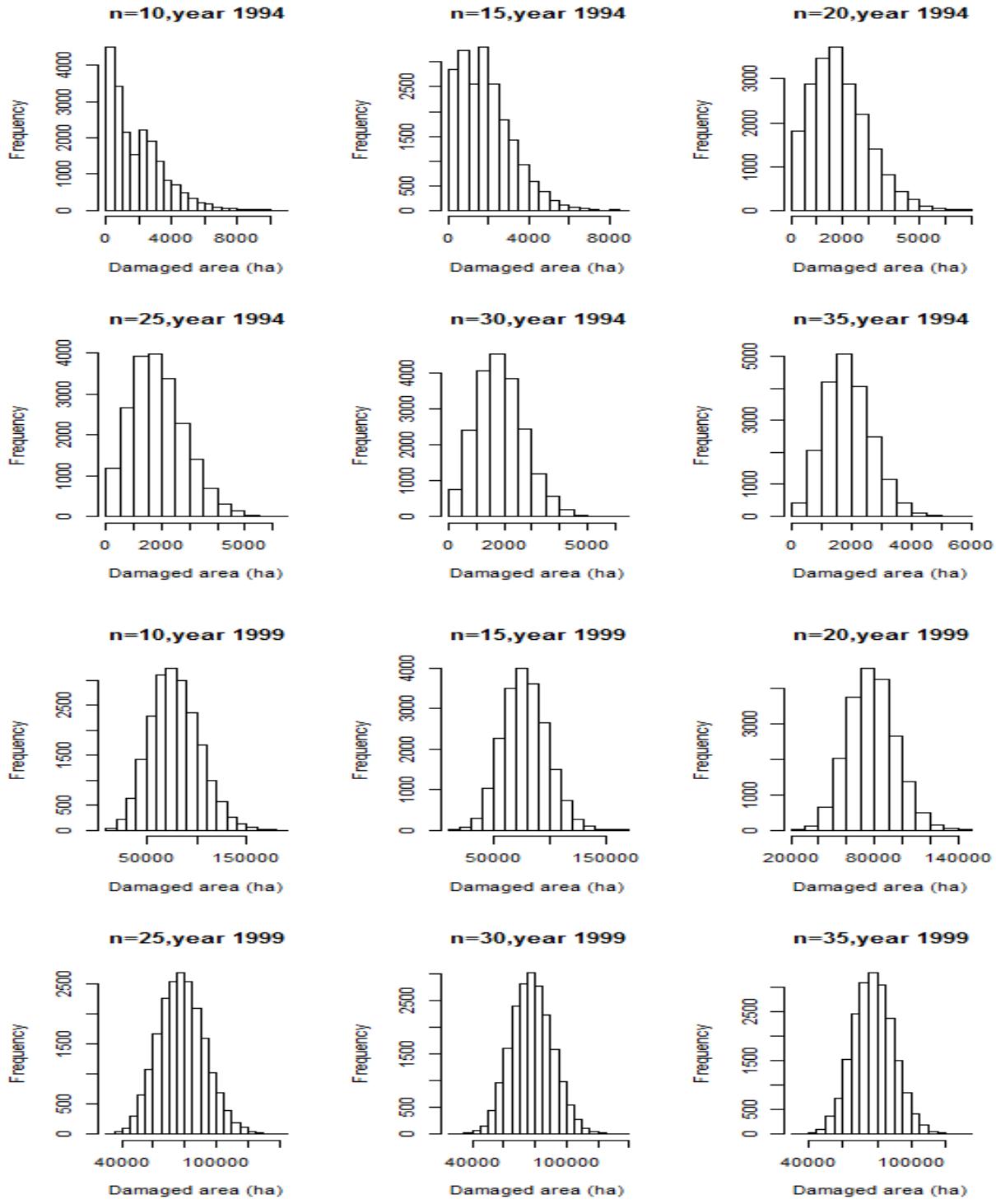


Figure B2. An example of area damaged caused by MPB approached normality distribution with increasing sample size

Table B1. Average ranking of estimated bias for four sample designs associated with eight causal agents and disorders.

Causal agent	SRS	PPS	STRA	NALIGN
DFB	3	4	1	2
MPB	2	4	1	3
PE	3	4	1	2
SAD	3	1	4	2
SB	3	4	2	1
SUB	3	4	1	2
WSB	3	4	1	2
Comb.	3	4	1	2
Average	2.88	3.63	1.50	2.00

Table B2. Largest estimated Bias of four sample designs associated with different causal agents and disorders and sample sizes in the year 1994 to 2013.

Unit: %

	MPB	DF	SB	PE	SAD	SUB	WSB	Comb.
SRS	1.13	1.17	1.29	3.23	2.49	7.47	4.60	0.55
PPS	*6.99*	*4.13*	*7.97*	*5.33*	*2.97*	*7.47*	*5.76*	*5.25*
NALIGN	0.59	0.85	2.04	2.49	1.38	0.84	2.86	0.66
STRA	1.16	1.29	2.43	2.10	2.14	2.23	3.56	0.79

Table B3. Example of testing homogeneity of estimated mean variance among four sample designs associated with various sample sizes using Monte Carlo method with Friedman test. Outputs are test for sample size equal 10 for combination of all causal agents and disorders. STRA is the best with smallest estimated mean variance.

Ranks

	Mean Rank
SRS10	2,95
PPS10	2,25
NALIGN10	3,75
STRA10	1,05

Test Statistics(a)

N			20
Chi-Square			47,160
df			3
Asymp. Sig.			,000
Monte Carlo Sig.	Sig.		,000
	95% Confidence Interval	Lower Bound	,000
		Upper Bound	,000

a Friedman Test

Table B4. Test of difference of estimated mean variance among four sample designs associated with different sample sizes in the year 1994 to 2013 using Friedman test index with significant level $\alpha = 0.05$. The null hypothesis was all estimated mean variances are equal. All results show that estimated mean variances of four sample designs were significant different associated with various sample sizes. Of which, STRA's estimated mean variance is the smallest, except PE's one causal agent where PPS is the smallest.

Causal agents and disorders		Sample size							
		10	15	20	25	30	35	50	70
Comb.	χ^2	47.16	49.32	38.76	47.46	47.46	37.98	42.54	37.20
	p-value	.000	.000	.000	.000	.000	.000	.000	.000
MPB	χ^2	34/98	36.60	30.30	36.30	34.20	33.18	34.20	34.98
	p-value	.000	.000	.000	.000	.000	.000	.000	.000
DF	χ^2	47.16	49.32	38.76	47.46	47.46	37.98	42.54	37.20
	p-value	.000	.000	.000	.000	.000	.000	.000	.000
SB	χ^2	38.56	39.76	38.43	43.86	46.89	38.81	47.02	31.99
	p-value	.000	.000	.000	.000	.000	.000	.000	.000
PE*	χ^2	21.35	26.01	14.51	23.82	27.85	17.40	24.74	23.89
	p-value	.000	.000	.000	.000	.000	.000	.000	.000
SAD	χ^2	34.26	26.05	21.32	29.34	28.71	27.57	28.45	27.95
	p-value	.000	.000	.000	.000	.000	.000	.000	.000
SUB	χ^2	41.08	41.08	39.95	41.08	41.08	39.00	38.94	35.97
	p-value	.000	.000	.000	.000	.000	.000	.000	.000
WSB	χ^2	27.06	32.87	26.81	32.94	30.92	28.71	36.73	24.92
	p-value	.000	.000	.000	.000	.000	.000	.000	.000

(*) Ranks of PPS were smallest

APPENDIX C – R CODES

Hear are R codes example for generating interested statistical properties

I. Simulation codes

1. SRS

```
`SRS.sim` <-
function (data,n,k=1,M=20000)
{
# Simple random sampling
#
# data - matrix of data
# n - sample size
# k - column of the variable to be sampled
# M - number of simulations
#
################################
#
data<-as.matrix(data)
N<-nrow(data)
stats<-matrix(0,nrow=M,ncol=2)
y<-data[,k]
mu<-mean(y)
tau<-sum(y)
```

```
cr<-0
tstat<-qt(0.975,n)
for(i in 1:M) {
yi<-sample(y,n)
stats[i,1]<-N*mean(yi)
stats[i,2]<-N*(N-n)*var(yi)/n
z<-(stats[i,1]-tau)/sqrt(stats[i,2])
if( abs(z) < tstat) {
  cr<-cr+1
  }
}
cr<-cr/M
cat("\n Simple Random Sampling \n")
cat("\n Sample Statistics \n")
cat("\n Number of Simulations = ", M)
cat("\n Sample Total = ", mean(stats[,1]))
cat("\n Mean Variance = ",mean(stats[,2]))
cat("\n Variance of the Mean = ",var(stats[,1]))
cat("\n Ratio of Variances = ",(mean(stats[,2])/var(stats[,1])))
cat("\n Coverage Rate = ",round(cr,4))
cat("\n")
cat("\n Population Statistics")
cat("\n Population Mean = ", mu)
```

```
cat("\n Population Total = ", tau)
cat("\n Population Variance = ",var(y))
cat("\n")
invisible(stats)
}
```

2. PPS

```
`PPS.sim` <-
function (data,n,kx,ky,M=20000)
{
# Sampling with probability proportional to size (PPS)
#
# data - matrix of data
# n - sample size
# kx - column of the auxiliary variable used to calculated
#      selection probabilities
# ky - column of the variable to be sampled
# M - number of simulations
#
################################
#
data<-as.matrix(data)
N<-nrow(data)
```

```
stats<-matrix(0,nrow=M,ncol=2)
y<-data[,ky]
x<-data[,kx]
prob<-x/sum(x)
indx<-seq(1,length(x))
mu<-mean(y)
tau<-sum(y)
cr<-0
tstat<-qt(0.975,n)
for(i in 1:M) {
yi<-sample(indx,n,replace=FALSE,prob=prob)
stats[i,1]<-mean(y[yi]/prob[yi])
stats[i,2]<-((N-n)/N)*var(y[yi]/prob[yi])/n
z<-(stats[i,1]-tau)/sqrt(stats[i,2])
if( abs(z) < tstat) {
  cr<-cr+1
  }
}
cr<-cr/M
cat("\n PPS Sampling \n")
cat("\n Sample Statistics \n")
cat("\n Number of Simulations = ", M)
cat("\n Sample Total = ", mean(stats[,1]))
```

```
cat("\n Mean Variance = ",mean(stats[,2]))
cat("\n Variance of the Mean = ",var(stats[,1]))
cat("\n Ratio of Variances = ",(mean(stats[,2])/var(stats[,1])))
cat("\n Coverage Rate = ",round(cr,4))
cat("\n")
cat("\n Population Statistics")
cat("\n Population Mean = ", mu)
cat("\n Population Total = ", tau)
cat("\n Population Variance = ",var(y))
cat("\n")
invisible(stats)
}
```

3. STRA

```
STRAT.sim <-
function (data,n,k=1,strata=n,M=20000,file="",write=F,append=F)
{
# Stratified Random Sampling
#
# data - matrix of data
# n - sample size
# k - column of the variable to be sampled
# M - number of simulations
```

```
# strata - number of strata
# file - fiepath and name to write sample statistics
# write - write sample statistics to the 'file'
# append - T = append statistics to an existing file; F = create a new file
#
################################
#
data<-as.matrix(data)
N<-nrow(data)
stats<-matrix(0,nrow=M,ncol=3)
y<-data[,k]
mu<-mean(y)
tau<-sum(y)
cr<-0
nobs<-floor(150/strata)
ul<-seq(nobs,150,nobs)
ni<-floor(n/strata)
for(i in 1:M) {
llim<-1
ulim<-ul[1]
m<-length(ul)
for(j in 1:(m)) {
smp<-seq(llim,ulim)
```

```
mi<-sample(smp,ni)
yi<-data[mi,k]
if(j == 1) {
yi.s2<-var(yi)
yi.mean<-mean(yi)
}
if(j > 1) {
yi.mean<-c(yi.mean,mean(yi))
yi.s2<-c(yi.s2,var(yi))
}
llim<-ulim+1
ulim<-ul[j+1]
}
if(sum(yi.s2) != 0) {
df<-(sum(yi.s2)^2/sum(yi.s2^2))*(ni-1)
}
else {
df<-ni-1
}
df<-df*5
tstat<-qt(0.975,df)
stats[i,1]<-nobs*sum(yi.mean)
stats[i,2]<-nobs^2*(nobs-ni)*sum(yi.s2/ni)/nobs
```

```
stats[i,3]<-ni*m
z<-(stats[i,1]-tau)/sqrt(stats[i,2])
if(abs(z) < tstat) {
  cr<-cr+1
  }
}
cr<-cr/M
output<-
cbind(M,mean(stats[,1]),mean(stats[,2]),var(stats[,1]),(mean(stats[,2])/var(stats[,1])),round(cr,4),
round(mean(stats[,3]),4),mu,tau,var(y))
cat("\n Stratified Random Sampling \n")
cat("\n Sample Statistics \n")
cat("\n Number of Simulations = ", M)
cat("\n Sample Total = ", mean(stats[,1]))
cat("\n Mean Variance = ",mean(stats[,2]))
cat("\n Variance of the Mean = ",var(stats[,1]))
cat("\n Ratio of Variances = ",(mean(stats[,2])/var(stats[,1])))
cat("\n Coverage Rate = ",round(cr,4))
cat("\n Average Sample Size = ",round(mean(stats[,3]),4))
cat("\n")
cat("\n Population Statistics")
cat("\n Population Mean = ", mu)
cat("\n Population Total = ", tau)
```

```
cat("\n Population Variance = ",var(y))
cat("\n")
if(write) {
write.table(output,file=file,append=append,quote=F,row.names=F,col.names=F)
}
invisible(stats)
}
```

4. NALIGN

```
NALIGN.sim <-
function (data,n,k=1,M=20000,file="",write=F,append=F)
{
# Non-aligned systematic sampling
#
# data - matrix of data
# n - sample size
# k - column of the variable to be sampled
# M - number of simulations
# strata - number of strata
# file - fiepath and name to write sample statistics
# write - write sample statistics to the 'file'
# append - T = append statistics to an existing file; F = create a new file
################################
```

```
strata<-n
data<-as.matrix(data)
N<-nrow(data)
Ni<-rep(0,n)mu=mean(aerial94[,6])
stats<-matrix(0,nrow=M,ncol=3)
y<-data[,k]
mu<-mean(y)
tau<-sum(y)
cr<-0
ns<-floor(N/strata)
ni<-1
ul<-seq(ns,150,ns)
L<-length(ul)
if(L != n) {
ul<-ul[1:n]
ul[n]<-N
}
for(i in 1:M) {
llim<-1
ulim<-ul[1]
for(j in 1:n) {
smp<-seq(llim,ulim)
Ni[j]<-length(smp)
```

```
mi<-sample(smp,ni)
if(j == 1) mj<-mi
if(j > 1) {
mj<-c(mj,mi)
}
llim<-ulim+1
ulim<-ul[j+1]
}
tstat<-qt(0.975,n)
yi<-y[mj]
stats[i,1]<-N*mean(yi)
#stats[i,1]<-sum(yi*Ni)
stats[i,2]<-N*(N-n)*var(yi)/n
#stats[i,2]<-sum(Ni*(Ni-1))*var(yi)
stats[i,3]<-n
z<-(stats[i,1]-tau)/sqrt(stats[i,2])
if( abs(z) < tstat) {
  cr<-cr+1
  }
}
cr<-cr/M
```

```
output<-
cbind(M,mean(stats[,1]),mean(stats[,2]),var(stats[,1]),(mean(stats[,2])/var(stats[,1])),round(cr,4),
round(mean(stats[,3]),4),mu,tau,var(y))
cat("\n Non-Aligned Systematic Sampling \n")
cat("\n Sample Statistics \n")
cat("\n Number of Simulations = ", M)
cat("\n Sample Total = ", mean(stats[,1]))
cat("\n Mean Variance = ",mean(stats[,2]))
cat("\n Variance of the Mean = ",var(stats[,1]))
cat("\n Ratio of Variances = ",(mean(stats[,2])/var(stats[,1])))
cat("\n Coverage Rate = ",round(cr,4))
cat("\n Average Sample Size = ",round(mean(stats[,3]),4))
cat("\n")
cat("\n Population Statistics")
cat("\n Population Mean = ", mu)
cat("\n Population Total = ", tau)
cat("\n Population Variance = ",var(y))
cat("\n")
if(write==T) {
write.table(output,file=file,append=append,quote=F,row.names=F,col.names=F)
}
invisible(stats)
}
```

II. Test for non-randomness (run tests)

```
runs.test <-
function (data,k=1)
{
y<-unlist(data[,k])
n<-length(y)
bin<-rep(0,n)
bin[y>0]<-1
n1<-sum(bin)
n2<-n-n1
r.exp<-(2*n1*n2)/(n1+n2)+1
s2<-2*n1*n2*(2*n1*n2-n1-n2)/(n1+n2)^2/(n1+n2-1)
#calculate number of runs
id<-bin[1]
r.obs<-1
for(i in 2:n) {
        if(id != bin[i]) {
        r.obs<-r.obs+1
        id<-bin[i]
        }
}
z<-(r.obs-r.exp)/sqrt(s2)
cat("\n")
```

```
cat("\n Sample size: ",n)
cat("\n Number of 0's: ",n2)
cat("\n Number of 1's: ",n1)
cat("\n Observed number of runs: ",r.obs)
cat("\n Expected number of runs: ",round(r.exp,4))
cat("\n Variance: ",round(s2,4))
cat("\n")
cat("\n Test statistic: z = ",round(z,4))
cat("\n Critical Region: Reject Ho if |z| > 1.96")
cat("\n Significant negative z values - nonrandom due to clustering")
cat("\n Significant positive z values - nonrandom due to regular distribution")
cat("\n NaN - nonrandom due to uniform distribution")
cat("\n")
}
```

III. Test for spatial autocorrelation (Moran's I test)

The spatial wts matrix for use with Moran's I

```
aerial.wt<-function(n) {
w<-matrix(0,ncol=n,nrow=n)
w[1,2]<-1
w[n,n-1]<-1
for(i in 2:(n-1)) {
w[i,i-1]<-1
```

```
w[i,i+1]<-1
}
invisible(w)
}
## Creat matrix 1's and 0's with 150 rows and 150 collumns)
matrix=aerial(150)
```

IV. Fitting variance model

An example using PPS for Douglas-fir beetle agent..

```
srs.dfb.var<-read.csv("C://projects//anh//srs-douglas-fir-beetle-var.csv",header=T)
srs.dfb.var1<-srs.dfb.var[2:10,]
srs.dfb.mean<-srs.dfb.var[1,2:21]
n<-seq(0,150,2)
yhat<-matrix(0,ncol=20,nrow=length(n))
n1<-srs.dfb.var1[2:9,1]
for(k in 2:21) {
i<-k-1
y<-log(srs.dfb.var1[2:9,k])
srs.dfb.n<-as.data.frame(cbind("n"=n))
srs.dfb.dat<-as.data.frame(cbind("n"=n1,"y"=y))
jnk<-lm(y~n,data=srs.dfb.dat)
#print(predict.lm(jnk,newdata=ndata))
yhat[,i]<-predict.lm(jnk,newdata=srs.dfb.n)
```

```
if(k ==2) {
r2<-summary.lm(jnk)$r.squared
a<-summary.lm(jnk)$coef[[1]]
b<-summary.lm(jnk)$coef[[2]]
}
else {
r2<-c(r2,summary.lm(jnk)$r.squared)
a<-c(a,summary.lm(jnk)$coef[[1]])
b<-c(b,summary.lm(jnk)$coef[[2]])
}
}
y.max<-max(sqrt(exp(yhat[1,])))
plot(seq(0,150,2),sqrt(exp(yhat[,1])),type="n",xlab="Sample Size",ylab="Standard Error of the
Mean",ylim=c(0,y.max))
for(i in 1:20) {
lines(seq(0,150,2),sqrt(exp(yhat[,i])),lwd=2,col=i)
points(srs.dfb.var1[2:7,1],sqrt(srs.dfb.var1[2:7,(i+1)]),col=i,pch=19)
}
r2.mean<-mean(r2)
a.mean<-mean(a)
b.mean<-mean(b)
hist(r2,density=30,xlab="R-square",ylab="frequency",main="")
abline(v=r2.mean,lwd=2)
```

```
hist(a,density=30,xlab="ln(s^2)",ylab="frequency",main="")

abline(v=a.mean,lwd=2)

hist(b,density=20,xlab="Slope",ylab="frequency",main="")

abline(v=b.mean,lwd=2)

for(i in 1:20) {

if(i == 1) {

se<-sqrt(exp(yhat[,i]))/srs.dfb.mean[,i]

}

else {

se<-cbind(se,sqrt(exp(yhat[,i]))/srs.dfb.mean[,i])

}
}
y.max<-max(se[1,])

plot(seq(0,150,2),se[,1],type="n",ylim=c(0,y.max),xlab="Sample Size",ylab="Relative Standard

Error of the Mean")

for(i in 1:20) {

lines(seq(0,150,2),se[,i],lwd=2)

}
truncated

for(i in 1:20) {

se1<-sqrt(exp(yhat[,i]))/srs.dfb.mean[,i]

se1[se1>1]<-1

if(i == 1) {

se<-se1

}
```

```
else {
se<-cbind(se,se1)

}
}
y.max<-max(se[1,])

plot(seq(0,150,2),se[,1],type="n",ylim=c(0,y.max),xlab="Sample Size",ylab="Relative Standard

Error of the Mean")

for(i in 1:20) {

lines(seq(0,150,2),se[,i],lwd=2)

}

fit relse models

n<-seq(0,150,2)

yhat<-matrix(0,ncol=20,nrow=length(n))

for(k in 1:20) {

y<-log(se[,k])

y.dat<-as.data.frame(cbind(y,n))

jnk<-lm(y~n,data=y.dat)

yhat[,k]<-predict.lm(jnk)

if(k ==1) {

r2<-summary.lm(jnk)$r.squared

a<-summary.lm(jnk)$coef[[1]]

b<-summary.lm(jnk)$coef[[2]]
}
else {

r2<-c(r2,summary.lm(jnk)$r.squared)

a<-c(a,summary.lm(jnk)$coef[[1]])
```

```
b<-c(b,summary.lm(jnk)$coef[[2]])

}
}
yhat.e<-exp(yhat)

a.mean<-mean(a)

b.mean<-mean(b)

hist(a,density=30,xlab="ln(s^2)",ylab="frequency",main="")

abline(v=a.mean,lwd=2)

hist(b,density=20,xlab="Slope",ylab="frequency",main="")

abline(v=b.mean,lwd=2)

y.max<-max(yhat.e[1,])

plot(seq(0,150,2),yhat.e[,1],type="n",ylim=c(0,y.max),xlab="Sample Size",ylab="Relative

Standard Error of the Mean")

for(i in 1:19) {

lines(seq(0,150,2),yhat.e[,i],lwd=2)
}
y.pred.1<-exp(ahat.1+bhat.1*n)

y.pred.2<-exp(ahat.2+bhat.2*n)

y.pred.3<-exp(ahat.3+bhat.3*n)

y.pred.4<-exp(ahat.4+bhat.4*n)

y.pred.5<-exp(ahat.5+bhat.5*n)


groupings

plot(seq(0,150,2),yhat.e[,1],type="n",ylim=c(0,y.max),xlab="Sample Size",ylab="Relative

Standard Error of the Mean",main="dfb decline")
```

LIST OF ABBREVIATIONS

SRS	Simple Random Sampling
PPS	Probabilities Proportional to Size
NALIGN	Non-Alignment Systematics Sampling
STRA	Stratified Random Sampling
MPB	Mountain Pine Beetle
DFB	Douglas-Fir Beetle
SB	Spruce Beetle
WSB	Western Spruce Budworm
SUB	Subalpine Fir Mortality
SAD	Sudden Aspen Decline
PE	Pine Engraver
Comb.	Combination of all causal agents and disorders including seven causal agents above and others