Collecting Information on Treatment Response

Prepared for
Jon Vivian
Resources Inventory Branch
B.C. Ministry of Forests

Project: MFI-401-064

March 31, 2001

J.S. Thrower & Associates Ltd. *Consulting Foresters*
Vancouver – Kamloops BC

ACKNOWLEDGEMENTS

We thank Pat Martin, *RPF* BC Ministry of Forests (MOF) Forest Practices Branch, Vera Sit *MSc* and Wendy Bergerud, *MSc,* BC MOF Research Branch, Val LeMay, *PhD RPF*, University of British Columbia and Bill Warren, *PhD* for their reviews of earlier versions of this document.

TABLE OF CONTENTS

List of Tables

List of Figures

1 INTRODUCTION

When asked what type of growth and yield (GY) information they need to improve their decision making, silviculturists respond with a multitude of answers. Most answers can be grouped into two categories:

1) Information to base treatment decisions on (primarily information on treatment response). For example, if *treatment A* is applied to *stand type X* on *site Y*, what happens to volume production, piece size distribution, wood quality, pest interactions, etc.? Will the results be more favourable if treatments "B" or "C" are applied?
2) Feedback on stand growth after treatment. They need to answer the questions "did we make the right decisions?", "were the silviculture dollars well spent?".

Meeting all silviculturists' information needs is a daunting task due to limited resources and the volume of information required to make informed decisions. British Columbia has very diverse forest types, site conditions, and possible management regimes. Furthermore, preferred management treatments and the managed portion of the forest landbase are continually changing.

Many different approaches have been used to obtain information to guide treatment application decisions and receive feedback on stand performance after treatment. Not all of these have been successful. Reasons for this failure include a lack of understanding of what growth after treatment and treatment response are, the suitability of various approaches used to obtain information, and the risks involved in extrapolating this information.

The goal of the report is to help silviculturalists understand the issues involved in quantifying the affect that silviculture treatments have on stand performance. The specific objectives are to:

1) Review the concepts of inference and statistical inference, and the need to understand and document the risks involved in extrapolating information (Section 2).
2) Define and distinguish between "growth after treatment" and "treatment response" (Section 3).
3) Review the basic principles of experimental design and sampling with reference to determining growth after treatment and treatment response (Section 4).
4) Discuss the strengths and weakness of different approaches for determining treatment response (Section 5).
5) Discuss the role of growth and yield monitoring in providing feedback on stand growth after treatment (Section 6).

2 INFERENCE AND STATISTICAL INFERENCE

Inference is the process of forming a conclusion about the state of something based on the state of something else. For instance, if you received one bad meal in a restaurant, should you infer that all meals served in the restaurant are not palatable? Statistical inference is a specific method of inference. Not only can we make inferences about a population based on a sample of the population, we can use the variability in the sample to develop a level of confidence in those inferences. Following the principles of statistical theory provides a scientifically defensible result.

The scope of the statistical inference refers to the population that the observed results are said to represent. Strictly speaking, the scope of inference is limited to the sampled population.[1] Other forms of inference are also commonly used to extrapolate the results further. This is often necessary due to limited time and resources. Personal experience or expert opinion can form the basis for extrapolating results. To be successful, the key is understanding and documenting the risks involved. This includes a full accounting of the assumptions and the effects of them being invalid. One must also realize this process is open to debate and is not as rigorously defensible as statistical inference.

Consider the following example. A fertilizer trial is replicated on good, medium, and poor sites that were purposively chosen with reference to their biogeoclimatic ecosystem classification (BEC) unit. To infer your results to other sites you must rely on: 1) the validity or credibility of the BEC system; and 2) some assurance that the sites chosen are typical of good, medium, and poor sites. Analyzing the trends in the data across sites or using the data to develop a model allows inference to sites not included in the trial. In this example, caution should be used when extrapolating the results to other sites because some sites may have additional nutrient deficiencies not addressed in the fertilizer prescription that limit the expected fertilizer response.

[1] The population from which the sample was drawn. Based on probability sampling theory all elements in the sampled population must have a known probability of being included in the sample.

3 GROWTH AFTER TREATMENT VERSUS TREATMENT RESPONSE

3.1 Growth After Treatment

Growth after treatment is simply the growth that occurs in a stand after it has been treated (Figure 1). It can be estimated by:

1) *Using re-measured permanent plots*
   - Permanent plots are established and measured in the stand prior to treatment to provide an estimate of pre-treatment stand yield.
   - At one or more times following the treatment, the plots are re-measured to provide an estimate of growth after treatment.

2) *Using independent temporary plots*
   - Temporary plots are located and measured in the stand prior to treatment to provide an estimate of pre-treatment stand yield.
   - At one or more times after treatment new, independent temporary plots are located and measured to provide an estimate of growth after treatment.

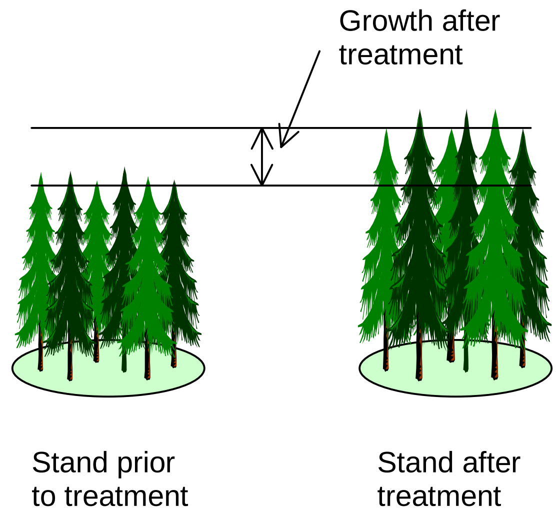


Figure 1. Growth after treatment. (For illustrative purposes, assume growth after treatment is measured in terms of height.)

Re-measured permanent plots are preferred because they should provide more precise estimates of growth after treatment. In addition, unlike temporary plots, re-measured permanent plots provide direct estimates of the components of net growth.[2] When measuring change, the use of independent temporary plots will generally result in increased sampling error and only provide estimates of net growth, not its components. However, the main advantage of temporary plots is the reduced cost.

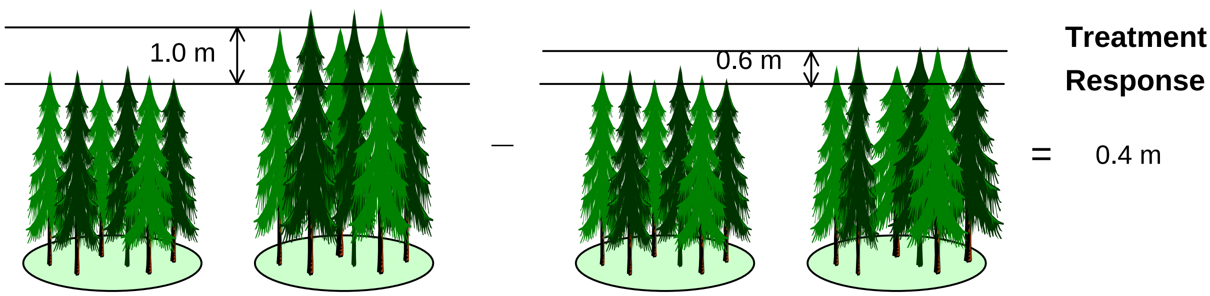
The phrase "growth after treatment" also describes the growth of a treated stand between any two points in time after treatment. For example, assume a stand was treated at age 10. The difference between measures of stand yield at ages 10 and 15 would provide the first five years growth after treatment. The difference between measures of stand yield at ages 15 and 20 would provide the second five years growth after treatment. If the objective is to estimate treatment response, growth after treatment is observed immediately following treatment (Section 3.2). When simply tracking growth after treatment as is done in growth and yield monitoring (Section 6.2), this may be done any time after treatment. Having a measure of pre-

[2] Net Growth = Survivor growth + Ingrowth - Mortality

treatment yield so that growth immediately following treatment can be estimated is not necessary. However, having the pre-treatment measurement will help explain why the stand is growing as it is after treatment.

3.2 Treatment Response

Treatment response is the growth after treatment of a treated stand *minus* the growth that would have occurred in that same stand if left untreated (Figure 2). It is what you are buying with your silviculture investment dollars. Treatment response is the incremental gain (or loss) due to the treatment.



Growth after treatment in the treated stand

Growth that would have occurred if the treated stand had not been treated

This can be estimated by sampling the treated stand.

This must be estimated by sampling an untreated stand (control) that is assumed to be similar to the treated stand.

Figure 2. What is treatment response? (For illustrative purposes, assume treatment response is being estimated in terms of height growth.)

A key factor in determining treatment response is estimating how the treated area would have grown without treatment. It is relatively straightforward to sample treated areas and estimate growth after treatment. To estimate the untreated growth requires selecting control (untreated) areas that are assumed to grow just as the treated areas would have if not treated. Too often the importance of this step is overlooked.

4 WHAT ARE THE BASIC DIFFERENCES BETWEEN SAMPLING AND EXPERIMENTAL DESIGN?

GY information is often collected by measuring tree characteristics within plots. There are many different types, shapes, and sizes of plots. Great care is often taken in designing a suitable plot type to provide the required information. Determining how and where the plots will be established is equally or more important as this affects your ability to drawn conclusions and make inferences. This is the role of sampling and experimental design.

Experimental design *is used to determine the effects of one or more treatments and their interactions on a population of interest*.

As discussed below (Sections 4.1 and 4.2), experimental design can be employed in research or operational trials.

Sampling *is used to describe and make inferences about a population of interest without having to measure all the elements within the population*.

Random sampling is used in numerous types of surveys to provide the basis for statistical inference. Schwarz[3] provides an excellent review of different surveys. Descriptive and analytical surveys (the two most common types used in forestry) are summarized from his paper (Sections 4.4 and 4.5).

It is possible to have a mixture of experimental design and sampling. Large experimental plots may be sampled (often referred to as sub-sampling) if it is too costly to measure the entire plots. The sites chosen for carrying out an experiment may, in turn, be a sample of all sites that make up a population of interest.

4.1 EXPERIMENTS (TRIALS)

Experiments provide a stronger basis for inferring treatment effects than sampling does. In experiments, the application of the treatments is controlled, as opposed to surveying treated and untreated areas where treatments have purposely been applied. Two common experimental designs used in forest research are the completely randomized design (Figure 3) and the randomized complete block design (Figure 4). In order to attribute observed differences to treatments, a good experiment should have the following characteristics:

[3] Schwarz, C.J. 1998. Studies of uncontrolled events. Pages 19 – 39 in Sit, V. and Taylor, B. (ed.s) Statistical methods for adaptive management studies. B.C. Min. For. Land Mgmt. Hndbk 42.

1) *A relatively homogeneous set of experimental units*[4] *(plots) that the treatments can be randomly applied to*. You are more likely to observe the treatment effects if the plots are similar. If necessary, plots should be grouped according to a characteristic that may effect treatment response (e.g., site productivity). This grouping, or blocking, ensures that treatment effects are not over-shadowed by additional effects (e.g., site effects).

2) *Random allocation of treatments to experimental units (plots).* Random allocation avoids the potential biases associated with the subjective allocation of treatments. From a statistical perspective, randomization ensures that the estimate of treatment effect will have an expected value equal to the true treatment effect[5] and allows for hypothesis testing.[6] Random, as opposed to subjective allocation, allows the highest degree of confidence in inferring that the observed differences are due to the treatments.

3) *All treatments are replicated*. Replication allows the calculation of experimental error (the variation not explained by treatments or blocks) which is then compared to the variation explained by the treatments (using an ANOVA F test).

4) *The chosen experimental units should be drawn from the population of interest.* (i.e., the population that the results of the experiment will be applied to). Experimental units can be chosen in different ways and this affects how the results can be applied. For example, if the experimental units are chosen randomly from a population, then the results of the experiment are applicable to the entire population. In contrast, if the experimental units are purposively located on specific sites, then the results of the experiment are applicable only to those sites. In the latter case, these trials are often referred to as case studies. Usually several case studies are required before it is appropriate to extrapolate the results to the larger population.

4.2 RESEARCH VERSUS OPERATIONAL EXPERIMENTS

The distinction between "research" and operational" experiments is not always clear; however, there are some commonly accepted differences (Table 1). There are often debates over which approach is more appropriate. Both meet specific needs and a combination of both is likely the best.

---

[4] Experimental units are the units that the treatments are applied to. They may be a single entity or a group. Sampling units are the measured portions of the experiment units. In many forestry applications, the experimental units are plots and the sampling units are trees. The experimental unit and the sampling unit may be one in the same. For example, a greenhouse experiment where the treatments are randomly assigned to individual trees and the individual tree growth is measured.

[5] Randomization ensures that across all possible randomizations, the average treatment response estimate (the expected value) will be equal to the true treatment response. It does not ensure that any individual treatment response estimate will be even close to the true treatment response.

[6] Many statistical procedures, including Analysis of Variance, assume observations (or errors) are independently distributed random variables. In most cases, randomization validates this assumption.

The trade-off between the two approaches is in the ability to detect treatment response and the applicability of the results to current operational treatments. The steps taken in a typical research trial to increase the chances of observing a treatment response often decrease the applicability of the results to current operational treatments. Conversely, the steps taken in an operational trial to increase the applicability of the results decrease the chance of observing a treatment response and the ability to make any statements about the treatment effect.

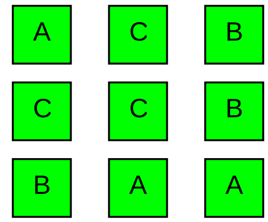


Figure 3. Example of a completely randomized design with treatments A, B, and C (control) laid out on a single site.

The strengths and weaknesses of each approach should be considered when basing decisions on trial results. In addition, when evaluating any trial, one should ensure that the trial (operational or research) possesses the elements of good design described in Section 4.1. If not, the results will have limited value and likely be misleading.

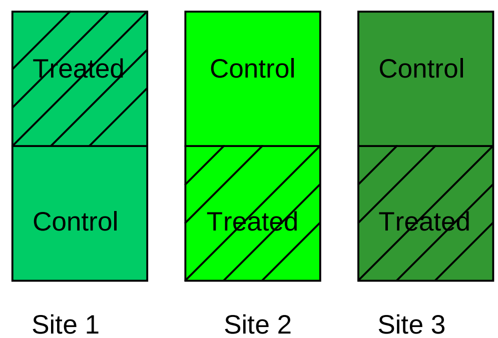


Figure 4. Example of a randomized block design with only one treated and control plot on each site.

Table 1. Summary of major characteristics of research and operational trials.

	Research Trials	Operational Trials
Treatment Application	Treatments are carefully applied to minimize application variability. For example, spacing treatments would be tightly controlled using detailed pre-treatment measurements, calculations, and choice of leave trees.	Treatments are applied following standard operational procedures. This results in greater application variability than research trials.
Plot Size	Typically have smaller plots as the objective is to make them as homogeneous as possible.	Typically have larger plots so treatments can be applied operationally.
Site Selection and Plot Location	Typically only a small number of sites, and often only one site, is chosen. Sites are purposively chosen to be as homogeneous as possible and represent sites of interest. Plots are purposively located to minimize between plot variability.	Typically the intent is to choose a large number of sites, either randomly from a defined population, or purposively to represent a range of conditions. Plots are then either purposively located to be as homogeneous as possible <u>or</u> are randomly located. Often the requirement for larger plots limits the number of plots to be established on a single site. Replicating across several sites compensates for this.
Applicability of Results to Current Operational Treatments	Theoretically, the results are only applicable to the plots in the trial. Each individual trial is a case study. Once there are several case studies across a range of conditions showing similar results, confidence in extrapolating the results increases. The tight control of treatment application means observed results might not be totally applicable to operationally applied treatments. Homogenous sites can be difficult to find. This requirement limits how well the chosen sites can represent the population to which the results are to be applied. Smaller plots may restrict observations on certain aspects of stand dynamics.	If plots are purposively located, then, as with research trials, theoretically the results are only applicable to the plots. If plots are randomly located on randomly chosen sites, then the results are applicable to the population from which the sites were chosen.[7] The operational application of treatments allows a higher degree of confidence in extrapolating the results to operationally applied treatments. The selection of more sites allows for better representation of the populations to which the results are to be applied. Larger plots allow a more comprehensive view of stand dynamics <u>if</u> all trees are re-measured over time.
Probability of Detecting Treatment Response	High probability of detecting a treatment response if one exists. Tight control over treatment application reduces differences due to inconsistencies in treatment application. Replication across homogeneous plots reduces confounding[8] with site and stand characteristics. The ability to detect treatment response is a function of the size of the response in relation to the natural variability. The smaller the response in relation to the natural variability, the harder it will be to detect. This necessitates, as in research trials, tight control over plot selection.	The probability of detecting a treatment response, if one exists, is much lower than research trials. Variability in the collected data may prohibit the detection of treatment effects unless treatment effects are large or a large number of plots are established. Variability is increased over that observed in research trials due to the operational application of the treatment, plots being less homogeneous, and the same nominal treatments applied on different sites being different in practice. For example, with brush removal, the treatment impact will depend on the initial amount, size, and distribution of the brush. This may vary considerably between sites.

[7] The population of managed stands is continually changing and defining a population to sample from is difficult. A population of managed stands sampled at one point in time will not be the same as the one to which you want to apply the results.
[8] Confounding results when the effects of two or more factors (e.g., treatment, site) cannot be isolated.

4.3 RECONNAISSANCE SURVEYS

Reconnaissance surveys are used to provide quick overview information about a population. Commonly, in a reconnaissance survey, a population (e.g., a silviculture opening) is subjectively sampled. This may consist of a walk-through with the surveyor taking notes, or plots may established at subjectively chosen points so the surveyor can calibrate their eye. Any conclusions drawn from the results of a reconnaissance should note the potential bias. Data from subjectively located plots cannot be used to calculate confidence intervals. Subjective sampling is often cost effective, but there must be an awareness of the potential bias and how it limits the strength of the inference. The degree of inference accepted will be based in part on the experience of the person conducting the subjective sample.

4.4 DESCRIPTIVE SURVEYS

Descriptive surveys are used to describe one or more characteristics about a population at a specific point in time (Figure 5). They could be used to check stand yield or other characteristics before or after treatment. A random or systematic sample with a random start provides the basis for statistical inference and allows confidence intervals to be calculated. The information collected is used to describe the state of the opening surveyed; no inferences can be made to other openings. Examples of descriptive surveys include regeneration and free-to-grow surveys.

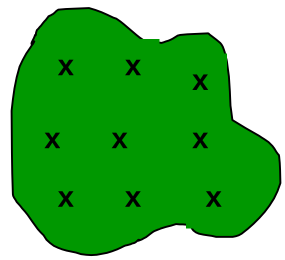


Figure 5. Example plot layout for a descriptive survey of a single opening.

4.5 ANALYTICAL SURVEYS

Analytical surveys use randomly selected sites from two populations to test the hypothesis that that there is a difference between the populations (Figure 6). For example, there appears to be a difference in plantation survival between two common site series within a particular subzone. Several openings located on each of the two site series are randomly chosen and surveyed with a random or systematic sample. Estimates of mean survival for each

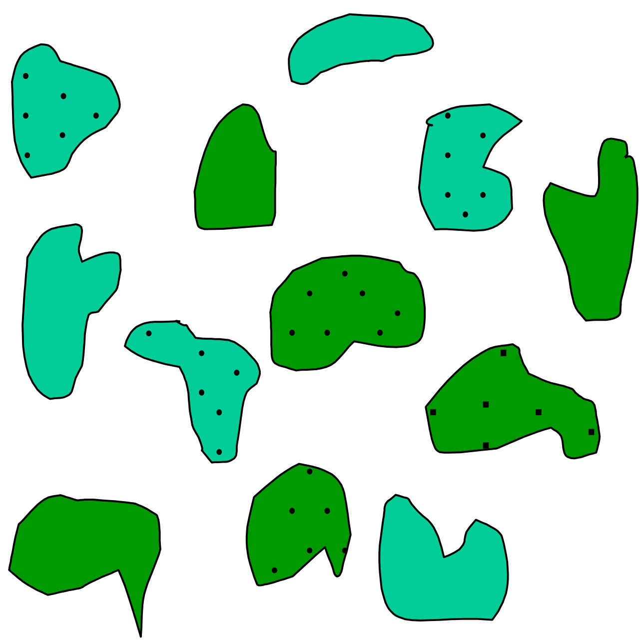


Figure 6. Example of an analytical survey. Three openings were randomly chosen from each of two site series. These openings were then surveyed.

site series are obtained along with measures of the variability within and between openings. With this information, the hypothesis that plantation survival differs between the two populations can be tested.

The difference between analytical surveys and experiments is that the explanatory variable (site series in the above example) is not controlled. In this example, if some other variable (such as stock type) caused survival to be different on the two site series, the wrong inference about the effect of site series on survival would be made.

5 EXPERIMENTAL DESIGN SHOULD BE USED TO DETERMINE TREATMENT RESPONSE

In an experiment, the explanatory variable (treatment) is randomly assigned to plots. This provides the highest level of confidence that observed differences between control and treated plots are due to the treatment.

In a managed forest, treatments are not randomly applied; they are subjectively applied to stands with specific characteristics. As a result, sampling treated and untreated stands and attempting to infer treatment effects can be very misleading (Section 5.3).

Sampling is valuable because in most cases the entire target population for which information is required can be assessed. This is much more difficult to do with experimentation. Properly done, sample surveys provide you with estimates of population parameters and the errors associated with those estimates. Systematic sampling can also provide a measure of the spatial variability in a population.[9]

5.1 RANDOMIZATION, REPLICATION, BLOCKING AND COVARIATES HELP TO ISOLATE THE TREATMENT EFFECT FROM NATURAL VARIATION

Stand growth is a function of site, age, stand structure, species composition, competition, genetic diversity, insect and disease impacts, and past treatment history. Therefore, there is often a great deal of natural variation within plots established to measure growth. This natural variation is often large relative to treatment response. This makes it difficult to estimate how the treated area would have grown without treatment. This in turn makes estimating treatment response difficult. As a result, researchers spend a lot of time trying to locate homogeneous areas to establish plots. They also insist on taking pre-treatment measurements so they can use techniques like regression and analysis of covariance to account for some of the natural variability in their attempts to isolate treatment response.

---

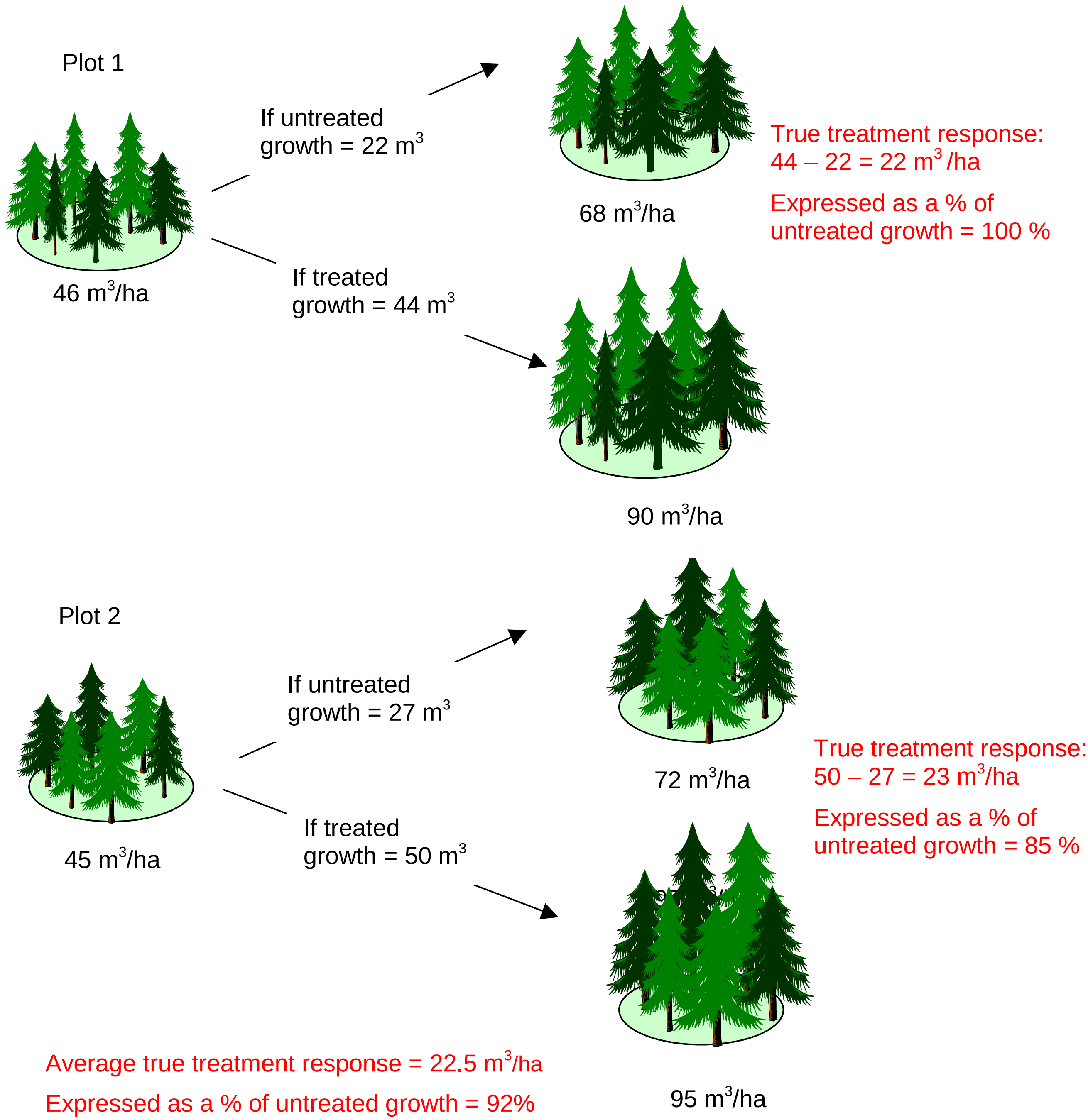
[9] This requires methods that are outside the scope of this report.

A simplified example of the problems introduced by natural variation is illustrated in Figure 7. Assume you establish two plots in a young spaced stand and measure the volume in these two plots. You are satisfied that the plots are well matched so you randomly choose one of them to fertilize. Several years after treatment, you measure the volume on both plots again. As shown in Figure 7, if plot 2 was treated and plot 1 was kept as the control, the estimate of relative response would be 127%. Conversely, if plot 1 was treated and plot 2 was the control, the estimate of relative response would be 63%.

So despite the homogeneous plots and a relatively large treatment response, the one estimate of relative response is double that of the other (and <u>this is one of the best case examples)</u>. It is quite feasible to have the response estimate range from positive to negative simply based on which plot is chosen as the control and which is chosen as the treated plot. This illustrates the need for replication (i.e., having several control and treated plots).

If the two plots in this example were dissimilar, the difference between the two possible response estimates would be greatly increased. This illustrates the need for homogeneous plots.

Randomization is used to avoid the potential biases associated with the subjective allocation of treatments. However, as this example shows, randomization does not ensure that any individual treatment response estimate will be close to the true treatment response. Rather, it ensures that across all possible randomizations the average treatment response estimate (the expected value) will be equal to the true treatment response. In this case, the average treatment response estimate is 22.5 m^3/ha (the average of 28 and 17 m^3/ha) which is equal to the true treatment response.



If plot 1 was chosen as the control, and plot 2 was treated, the response estimate would be: $50\ m^3 - 22\ m^3 = 28\ m^3/ha$ or 127%.

If plot 2 was chosen as the control, and plot 1 was treated, the response estimate would be: $44\ m^3 - 27\ m^3 = 17\ m^3/ha$ or 63%.

Figure 7. A simplified example of estimating treatment response with only 2 plots.

5.2 BEWARE THE PSEUDO-REPLICATION TRAP

It is very important to distinguish between the randomized replication of treatments in an experiment and random sub-sampling within a single plot. Sub-sampling is often referred to as "pseudo-replication". Consider the two scenarios illustrated in Figure 8 and Figure 9. In both cases measurements were taken on six control plots and six treated plots, but the inferences you can draw from these two scenarios are quite different. In the first scenario (Figure 8), the treatment and control were randomly assigned to two large plots and then six smaller plots were established within these large plots. In the second scenario (Figure 9), the treatment and control were randomly assigned to each of six plots. In the first scenario, there is only one true replicate of the control and treatment because there has been only one treatment application. In the second scenario, there are six replicates resulting from six treatment applications.

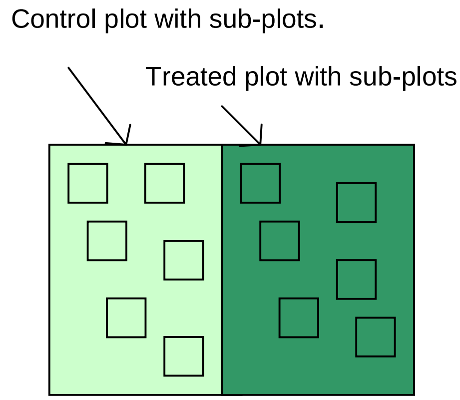


Figure 8. Example of sub-sampling (pseudo-replication).

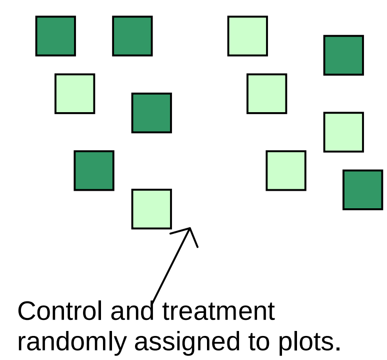


Figure 9. Example of true replication.

Following statistical theory, you cannot make any statements about the effect of the treatment in the first scenario. It is quite possible that differences in site, stand structure, or other factors between the two large plots could mask the treatment effect. In the second scenario, control and treatment applications have been randomly assigned across the site providing a much stronger basis for inference.

There may be instances where the random application of a treatment to several plots is not feasible and the only option is to sub-sample a large treated and control plot. In this instance, a mean and standard error can be calculated for both the control and treated sub-samples. The hypothesis that the two plot means are equal can then be tested. If this hypothesis is rejected and the two plot means are different, then the question becomes how much of this difference can be attributed to the treatment. If you are confident that the plots were similar prior to the experiment, then the level of confidence attributed to the observed difference to the treatment increases. At no point, however, can one state that the observed difference is a statistically significant treatment response (there is no basis for a statistical inference of treatment response).[10] There is also the risk that unobserved site or stand attributes are the cause of the observed difference between the two plots.

5.3 WHY CAN'T I SIMPLY COMPARE TREATED AND UNTREATED SITES TO GET AN IDEA OF TREATMENT RESPONSE?

The answer is that you can as long as you are aware of the large risks of drawing the wrong conclusion. You could conduct an analytical survey (see Section 4.5), randomly sampling both treated and untreated sites, and compare their yields (or growth) in an attempt to infer something about the treatment response. However, the following factors must be considered:

1) In a managed forest, treatments are not randomly applied. So in most cases the untreated and treated sites will differ in some way (i.e., sites chosen for treatment have certain characteristics that sites left untreated do not have).
2) The magnitude of the treatment effect is often small relative to the natural variability, making it difficult to isolate the treatment effect.
3) Pre-treatment measurements, which are often critical to understanding the effect of the treatment, are not obtained in most cases.

Consider the following example:

A silviculturist wants to assess the success of a brushing treatment on a particular site type. She locates six openings that were planted with the same stock seven years ago. Three of these openings were brushed three years ago, and three were not. She randomly chooses two brushed openings and two that were not brushed (Figure 10). She randomly establishes a large plot in each chosen opening and records the distribution, size and health of the seedlings. After summarizing the results, she is puzzled because the seedlings on the brushed sites are growing slower than those on the non-brushed sites.

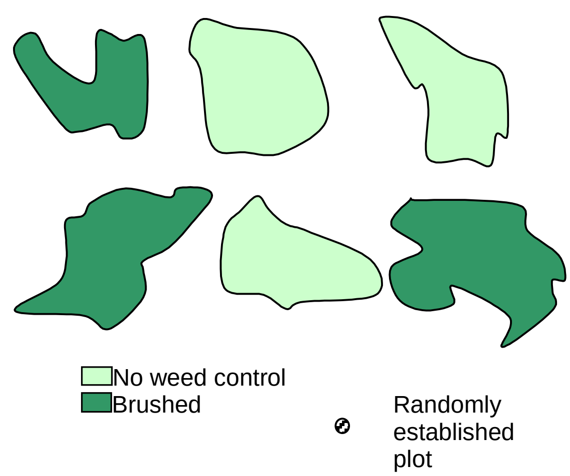


Figure 10. Sampling treated and untreated sites.

What is going on?

Has the brushing actually reduced the growth of the seedlings? Perhaps, but not likely. The more plausible explanation is that brushing was only done on sites where the brush was impacting seedling growth. Seedling growth was not being impacted on sites that were not

[10] This is true for individual sites in operational trial designs that have a single control and treated plot on each of several sites.

brushed. The treatments were not randomly applied. In this example, the silviculturist chose poor control sites for estimating what would have happened without the brushing treatment.

What could she have done?

Listed below is a range of some of her options. The likelihood of observing a treatment response increases from option 1 to 4. In contrast, the applicability of the results to current operational treatments decreases from option 1 to 4. The description for each option provides an overview of the steps needed. More detail on the key features of experiments is provided in Section 4.1.

1) *Randomly choose some openings to be left untreated.*

If a large treatment program is planned, randomly choose some of the openings scheduled for treatment and leave them untreated (this is sometimes referred to as establishing legacy areas) (Figure 11). These untreated and randomly chosen treated openings can be surveyed at a later date to provide an indication of the treatment response for the entire area treated. However, the variability between openings (differences in site quality, stand characteristics, etc) may make it difficult to determine the treatment response.

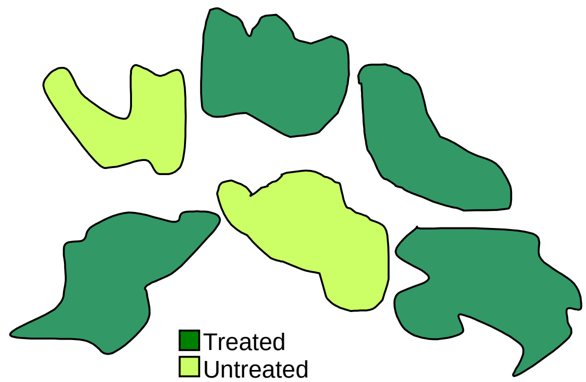


Figure 11. Randomly chosen openings from stands scheduled for treatment left untreated.

2) *Randomly choose portions of openings to be left untreated.*

On several openings scheduled for treatment, randomly choose a portion of the opening to set aside as control areas (Figure 12). At some point in the future, survey both treated and untreated portions, calculate a treatment response estimate on each opening and examine the trends across openings. This will provide a more precise estimate than option 1, as there will be less variability within openings than between openings.

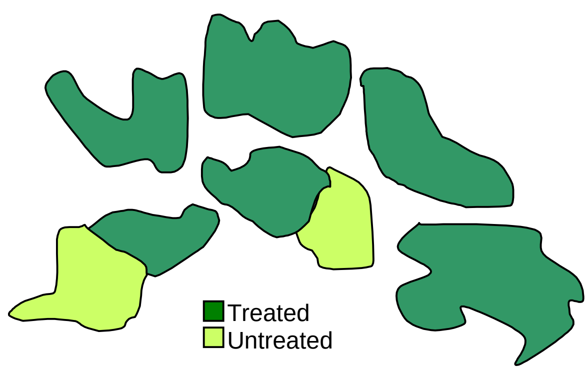


Figure 12. Randomly chosen portions of openings scheduled for treatment left untreated.

3) *Establish an operational trial.*

Choose several sites and establish a minimum of two large plots, which are as similar as possible (Figure 13). Record pre-treatment measurements and then randomly assign the treatment and control. Apply the operational treatment and then collect the post-treatment measurements to obtain estimates of treatment response.

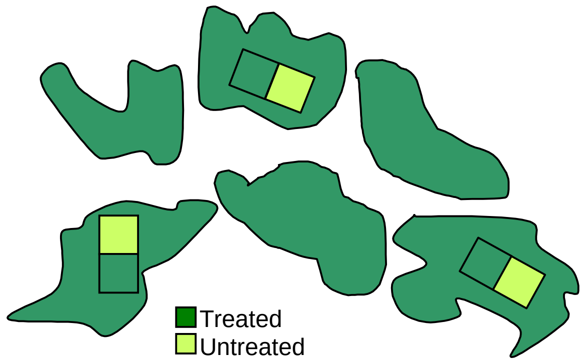


Figure 13. Example of an operational trial layout.

4) *Establish a research trial.*

Choose one or more sites and establish several relatively small plots, which are as similar as possible (Figure 14). Record pre-treatment measurements[11] and then randomly assign treatment and control to each of several plots on a single site. Carefully apply the treatment to minimize treatment application variability. Collect post-treatment measurements to obtain estimates of treatment response.

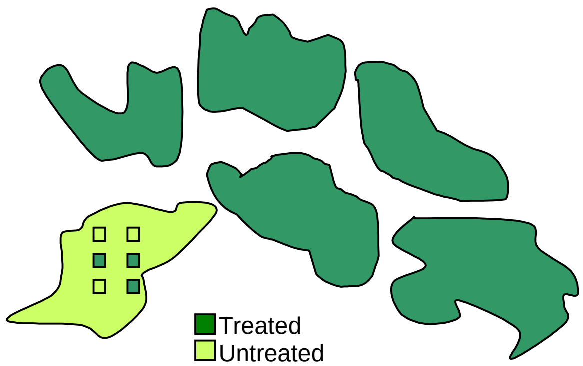


Figure 14. Example of a research trial layout.

[11] Pre-treatment data can be used to ensure plots are as similar as possible. It can also be used as covariates if there is difficulty in locating homogeneous plots.

6 SAMPLING SHOULD BE USED TO PROVIDE FEEDBACK ON STAND GROWTH AFTER TREATMENT

When silviculturists want feedback on treated stand growth they typically have a target population[12] in mind. This may be a single stand or it may be a group of similarly treated stands. In order to make statements about the entire target population, random or systematic sampling should be used because it provides a defensible statistical basis for inference to the entire sampled population.[13] If this is not done (for example plots are purposively established in fully stocked portions of the stand or simply located where there is easy access) the data collected will only represent the plot locations and there will be no basis for statistical inference to the target population. Growth and yield monitoring (Section 6.2) uses random or systematic sampling to provide feedback on a defined population of stands.

6.1 WHAT IS MONITORING?

"Monitoring" is a very widely used and ambiguous term. Generally, it is used to describe the process of checking or regulating some defined activity. For example:

- Remote sensing is used to monitor changes in forest cover.
- Insect and disease surveys are used to monitor the state of forest health.
- Projected timber yields are monitored for accuracy.
- Herbicide and pesticide applications are monitored for compliance with guidelines.
- The implementations of plans are monitored to ensure that objectives are being met.
- Water quality is monitored in streams adjacent to fertilizer operations to ensure they are not contaminated.
- Permanent sample plots are used to monitor the growth of the forest over time.
- The federal government has committed Canada to sustainable forest management and must monitor the overall management of the nation's forests.
- An RPF must monitor the work of his or her staff to ensure that it is carried out to specified standards.

Therefore, in any discussions using the term "monitoring", it is imperative that the monitoring objectives are well defined and clearly understood.

---

[12] The target population is the population about which information is wanted. Ideally the sampled population and the target population are one in the same. In some cases the sampled population is more restricted than the target population and in these cases you must remember that the inferences from your sample apply only to your sampled population. The validity of applying the inferences to the target population will depend on other sources of information and judgment.

[13] The sampled population is the population from which the sample is taken and to which inferences from the sample can be made.

6.2 WHAT IS GROWTH AND YIELD MONITORING?

In British Columbia ***GY monitoring*** has been defined as:

The process of comparing the actual GY of a forest or stand to the predicted or expected GY for that forest or stand.[14]

However, it may be more suitably called "*checking growth and yield estimates and predictions*". The essential feature of a GY monitoring program is that the sample must be representative of the population of interest. Two of the most widely accepted methods of obtaining a representative sample are a random or systematic[15] establishment of plots. If we wanted to check second growth projections used in timber supply analysis, GY monitoring plots would be established at systematically or randomly chosen locations throughout the population of second growth stands.

GY monitoring is a type of sampling. It does not include controls and therefore cannot be used to estimate or check treatment responses directly.

6.2.1 Why do we make the distinction between collecting new growth and yield information versus checking existing growth and yield information?

GY monitoring is not intended to provide new information, it is intended to check existing growth and yield information. This distinction is important because it can translate into different data collection strategies. Typically checking (monitoring) existing growth and yield information will require a smaller sample than the sample required for collecting new information. For example, to establish the amount and distribution of timber volume, we have developed detailed sampling procedures within the new Vegetation Resources Inventory. To check existing estimates of timber volumes, we utilize the inventory audit procedure. These are two different sampling strategies designed to meet two different objectives.

6.2.2 GY Monitoring tracks growth after treatment

GY monitoring plots can be established in a population of treated stands (preferably, but not necessarily, prior to treatment)[16] and re-measured over time to track the growth of the treated stands. This growth is then compared to the predicted growth of the treated stands (e.g., TIPSY output) used in both silviculture decision making and timber supply analysis. If large differences are observed between actual and predicted growth then action must be taken to determine the reason for the difference. It may be that incorrect inputs (e.g. stand density and site index) are being supplied to the models, or more seriously, that the models are flawed.

[14] BC Ministry of Forests Growth and Yield Monitoring Task Force.

[15] Typically done on a fixed grid with a random starting point.

[16] Establishing the plots prior to treatment may be logistically difficult and costly in some instances and simply not fit within the sampling plan in others. However, if feasible, obtaining pre-treatment measures provides a description of where the stand started from and helps to explain why it is growing as it is after treatment.

Having growth after treatment data also provides a historical record of how stands grew following treatments. Silviculturists can use this information to assess trends in stand performance.

6.2.3 Checking (monitoring) treatment responses

GY monitoring can be used to indirectly check predictions of treatment response. Data from plots established in treated and untreated stands can be compared to growth and yield projections for these stands. If both treated and untreated stands are being projected with reasonable accuracy, then one could assume that treatment responses predicted by the model are acceptable. However, the difference between the observed treated and untreated growth does not provide an estimate of treatment response that can be checked directly. In a managed forest, treatments are not randomly applied so in most cases the untreated and treated sites will differ in some way. Consider the following simplified example.

The local forester embarked on an aggressive fertilization program and as a result, all Douglas-fir stands suitable for fertilization have been fertilized. He wanted to check that he was obtaining the expected fertilizer response. As part of a GY monitoring program, plots were randomly located in fertilized and unfertilized Douglas-fir stands and re-measured on a 5-year cycle. All the fertilized plots are in site index 30 m stands, while the unfertilized plots are in site index 25 m and 38 m stands (Figure 15). The data from the fertilized stands were compared to predicted fertilized growth for site index 30 and the analysis indicated the predictions are, on average, acceptable. This is a check that growth after treatment is as predicted. The data from the unfertilized stands were compared to the growth projections for site index 25 and 38. Again the analysis indicated the predictions on average are acceptable. This process provides an indirect check that the predicted fertilizer response for site index 30 is also acceptable. The assumption is made that if the predictions for site index 25 and 38 unfertilized growth are acceptable, then the prediction for site index 30 unfertilized growth will be as well. A direct check is not possible, as there are no site index 30 unfertilized stands.

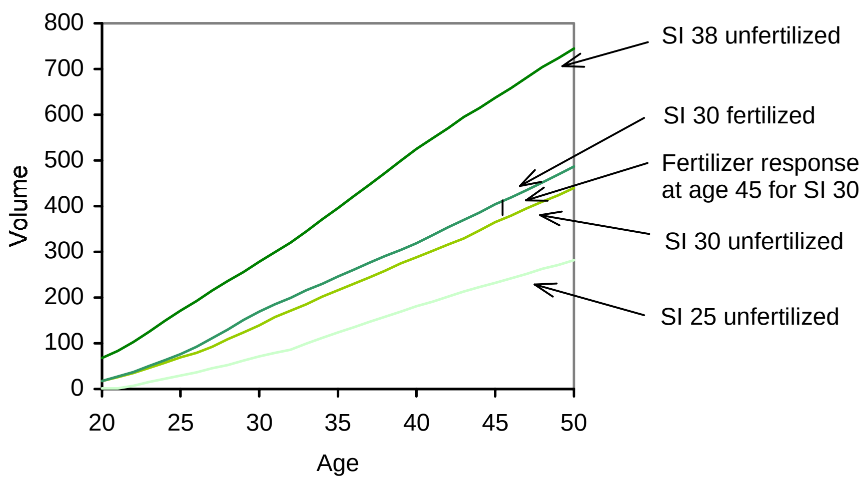


Figure 15. Volume versus age for fertilized and unfertilized Douglas-fir

7 SUMMARY

The GY information that most silviculturalists need can be put into two categories; 1) information to base treatment decisions on (treatment response information) and 2) feedback on stand growth after treatment. Treatment response information should be obtained from well-designed experiments and feedback on stand growth after treatment should be obtained by sampling target populations of treated stands.

Growth after treatment is simply the growth that occurs in a stand after it has been treated. By sampling a defined population (either a single treated stand or a group of similarly treated stands) and measuring growth after treatment, inferences can be made about the entire population that was sampled. Observed growth after treatment can provide feedback, indicate trends, and generate hypothesis to be tested in experiments. It can also be compared to predicted growth after treatment to check the predictions (known as GY monitoring).

Treatment response is the incremental gain (or loss) due to the treatment. It is what is purchased with silviculture expenditures. The key factor in determining treatment response is estimating how the treated stand would have grown without treatment. Developing this estimate requires controls. The more similar the controls are to the treated area, the better you can estimate how the treated areas would have grown without treatment. Poorly chosen controls (for example, the stand across the road from the treated stand or the back corner of the block that is too far for the spacing crew to walk to) result in poor and potentially very misleading treatment response estimates.

Experimental design, whether employed in research or operational trials, provides the strongest basis for inference of treatment effects because the application of the treatments is randomly applied across a set of experimental units. This is opposed to surveys of existing treated and untreated areas where treatments are likely to have been purposively applied to specific sites and stands (those in greater need of treatment).

Always consider the way treatment response information was obtained when using it to make decisions. Are the response estimates from survey or experimental data? Were plots randomly or subjectively located? How well does the data represent the population you are interested in? How similar to the treated areas are the controls? How were the treatments applied? How consistent was the treatment application? Answers to all these questions will influence the reliability of the information.

There is a great deal of natural variability in managed forests. This can make detection of treatment responses difficult. However, even relatively small responses can be economically important, making investments in sound methods to determine treatment responses justified.