



The Experimental Design of the Missouri Ozark Forest Ecosystem Project

Steven L. Sheriff and Zhuoqiong He¹

Abstract.—The Missouri Ozark Forest Ecosystem Project (MOFEP) is an experiment that examines the effects of three forest management practices on the forest community. MOFEP is designed as a randomized complete block design using nine sites divided into three blocks. Treatments of uneven-aged, even-aged, and no-harvest management were randomly assigned to sites within each block. Pre-treatment data have been collected to ensure that results can be adjusted in terms of pre-existing conditions. Interdisciplinary studies are conducted within this design to provide information about relationships of different forest components. MOFEP's design was selected to allow the most flexibility to forest managers during the implementation phase while accounting for among block variation in examining treatment effects.

Many studies of forest and wildlife resources have been conducted in the Missouri Ozarks. The objectives of these studies have covered the breadth of forest and wildlife management. Several have even examined forest-wildlife habitat relationships (e.g., Robinson *et al.* 1995, Thompson *et al.* 1992). Despite the number of studies conducted, controversy surrounding the impacts of forest management upon wildlife populations remains (Kurzejeski *et al.* 1993). The controversy is due to different factions basing their arguments on studies that were observational in nature and done under different conditions and at different times. Most wildlife studies are not designed to answer questions concerning management effects, but are designed to develop hypotheses about these possible effects (Romesburg 1981). To overcome these problems and issues, a study was planned that would examine how forest management affects the forest-wildlife community in the Missouri Ozark Plateau. In other words, a project was needed to test hypotheses that these other studies had established and to provide a reliable knowledge base for decision processes in forest management.

The Missouri Ozark Forest Ecosystem Project (MOFEP) is designed to collect data to estimate effects and test hypotheses. The design allows

the examination of cause-and-effect relationships within the forest ecosystem. MOFEP differs from earlier studies in many ways. First, MOFEP is a large-scale experiment conducted at the landscape scale used in forest planning and management in Missouri. Second, it examines management concerns that are not only pertinent today but will be of concern to future forest managers in Missouri's Ozark forests. Third, MOFEP is the first attempt to coordinate a multidiscipline approach for examining the effects of forest management practices on the forest ecosystem through an experimental approach.

In this paper, we describe (1) reasons for choosing an experimental approach for MOFEP, (2) components of the experiment, (3) experimental design selected, (4) overlap of complementary interdisciplinary studies, (5) limitations of the selected experimental design, and (6) MOFEP as an adaptive management approach.

WHY AN EXPERIMENT?

Forestry and wildlife studies can basically be divided into three conceptual designs: descriptive, correlational, and manipulative (White and Garrott 1990:14-16). These three conceptual designs are analogous to the respective three approaches that can be used in the scientific method—induction, retroduction, and hypothetico-deductive (Romesburg 1981).

¹ Wildlife Biometrics Supervisor and Wildlife Biometrician, respectively, Missouri Department of Conservation, 1110 S. College Ave., Columbia, MO 65201.

Studies that are designed using the descriptive approach observe and describe natural processes. These types of studies are useful in describing the natural history of a species or the structure of a forest. However, these studies do not test hypotheses. As the term indicates, the useful information derived from a descriptive study is a description of things measured. Often from these descriptive studies, hypotheses are formulated that can be tested under one of the other two approaches.

A more elucidating approach than the descriptive study is to formulate at least one hypothesis and design a correlational or retroductive study (Romesburg 1981). Using this approach, the researcher collects data on the subject over a broad range of environmental factors. For example, Thompson *et al.* (1992) conducted a correlational study of breeding birds. In their study, bird densities were examined on areas that had been either clearcut or on areas where no timber harvest had occurred in recent time. Because areas had been previously treated (clearcut or no harvest) with no randomization of treatments among areas, the analysis could indicate only if there were observed differences in the bird densities between the two types of areas. From this type of design, the forest management treatment cannot be inferred as the cause of differences in bird densities. The forest management treatments and the location of the treatments are mixed or confounded. The location of treatments may be tied, inadvertently or intentionally, to a process that would have shown treatment effects where none might have existed if a different assignment of treatments had been made to the locations. In other words, factors other than treatment may have been responsible for the observed responses due to the choices of areas studied.

Correlational or retroductive studies are very useful (Romesburg 1981). They can provide insights into hypotheses that should be further explored to determine cause-and-effect relationships. In other words, studies, like those reported by Thompson *et al.* (1992) and Robinson *et al.* (1995), should be used to formulate experimental approaches for determining treatment effects upon some set of response variables.

To infer cause-and-effect relationships, one must conduct a manipulative or hypothetico-deductive study (Eberhardt and Thomas 1991, Green 1979, James and McCulloch 1985,

Romesburg 1981, White and Garrott 1990). Under this approach, the system must be manipulated in a planned manner to determine if hypothesized cause-and-effect relationships exist. An experimental approach with properly defined treatments, randomization, and replication is used to determine cause and effect. Treatments may entail more than one type of manipulation that may be compared with each other or with a control treatment that remains untouched. Within the experiment, the results from areas treated the same are compared to results of areas treated differently. If proper experimental procedures are applied and data differ among treatments, cause and effect can be inferred. However, if all treatments show similar results, then one would conclude that the treatments had little effect upon the parameters being measured.

Given the public's desire to support a stronger conservation and stewardship ethic for forest management (Brookshire *et al.* 1997), MOFEP was designed as a manipulative or experimental approach. The goal was to determine the effect of forest management upon the forest and wildlife community of the Missouri Ozarks. When we use this scientific approach for determining cause and effect, any impacts or benefits that might be measured during this project may be attributed to forest management practices.

COMPONENTS OF AN EXPERIMENT

According to Hurlbert (1984), an experiment is composed of five components: (1) the hypothesis, (2) the experimental design, (3) the experiment execution, (4) the statistical analysis, and (5) the interpretation of results. Without the first component, the hypothesis, an experiment would be a failure. This would be true even if the other four components were carried out with great attention to detail and protocol. The hypothesis of any experiment is the key to the successful outcome of that experiment. For MOFEP, the hypothesis is that no differences among the selected forest management practices will be found when applied to the experimental units. This hypothesis is stated in terms of equivalence. Statistical procedures normally used in studies like MOFEP examine data under a null hypothesis that allows biologists to determine if equivalence can be supported by the experimental data. In other words, are treatment effects equal or do they exhibit differences? These questions must be

answered within the context of the scope and power of the experiment.

Experimental Design

An experimental design must provide observations that will support tests of hypotheses and estimation of parameters of interest. The description of an experimental design, according to Hurlbert (1984) and McAllister and Peterman (1992), includes (1) the nature of the experimental unit; (2) the number and kinds of treatments, including controls, to be tested in the experiment; (3) replication in time and space, which controls for stochastic factors among replicates that are inherent in the experimental units; (4) interspersing of differently treated units in space to control for properties of the experimental units; (5) randomization in allocating different treatments to experimental units so that biases and stochastic factors associated with the experimental unit do not become influential; and (6) statistically independent experimental units.

The experimental unit chosen for MOFEP was a site (Brookshire *et al.* 1997). Nine sites of 266 to 527 ha were found on Missouri Department of Conservation lands located in Shannon, Reynolds, and Carter Counties in Missouri. Brookshire and Hauser (1993) and Meinert *et al.* (1997) provide extensive descriptions of these nine sites. Three treatments—even-aged management, uneven-aged management, and no harvest management (the control)—are being applied to these sites (Brookshire *et al.* 1997). Visual observations were used to assign each site to one of three blocks based on their subjectively determined similarity. This blocking allows for replication of the three treatments in space, so that no treatment is assigned twice to the same block. These blocks are considered independent of one another. Due to the similarity of sites within each block, we expect that results will be more similar within blocks than among blocks if all sites were treated alike. Sites within each block also are assumed to be independent. We are assuming that the responses in one experimental unit are not related to responses in other units, except that they might share the same treatment.

The three treatments were randomly assigned to sites within a block. Sites from each block were ordered using a random numbers table. Each site within the random ordered list for a block was assigned a treatment number in its turn,

again, using a random numbers table. Then, an individual was asked to assign a treatment to each treatment number, without having any prior knowledge of the previous randomization results. Thus, a treatment was randomly assigned to a site within a block. The result of this randomization process is shown in figure 1 of Brookshire *et al.* (1997). Block 1 includes sites 1-3, block 2 includes sites 4-6, and block 3 includes sites 7-9. This design is commonly known as a randomized complete block design (Steel and Torrie 1980:196-197) or randomized blocks (Cochran and Cox 1957:106-107).

The 5 years before treatments were applied (i.e., before timber was harvested) were critical to the experimental design. During this period, data were collected about the characteristics of interest. This pre-treatment information will be critical in understanding if the impacts of treatment were due to treatment or were a continuation of the system as it existed before treatment. To illustrate the importance of the pre-treatment data, an example is shown in figure 1. Figure 1A shows a difference between two treatments, whereas figure 1B shows that there was no impact through time. Without pre-treatment information, we might conclude in both cases that a difference between treatments occurred. Pre-treatment data can be included in the statistical analysis model to increase precision for determining the treatment effects.

Experiment Execution

The execution of the experiment is the next crucial component in the experimental approach. Because MOFEP is a long-term study, it should extend through two or more full rotations of timber harvest, or about 200 years or longer. This length of time may be important in understanding the full and long-term impacts of each management strategy. However, results from shorter periods can be used by resource managers in the forests of Shannon, Reynolds, and Carter Counties. Information derived from MOFEP also can be used by managers in adjusting their approach to each of the harvest treatments. Also, variables measured do not have to be measured every year, but a systematic scheme, which ensures continuity of data collected on all nine sites through time, can be built to periodically remeasure certain variables during this project. The information from MOFEP will become more valuable as each year passes and subsequent

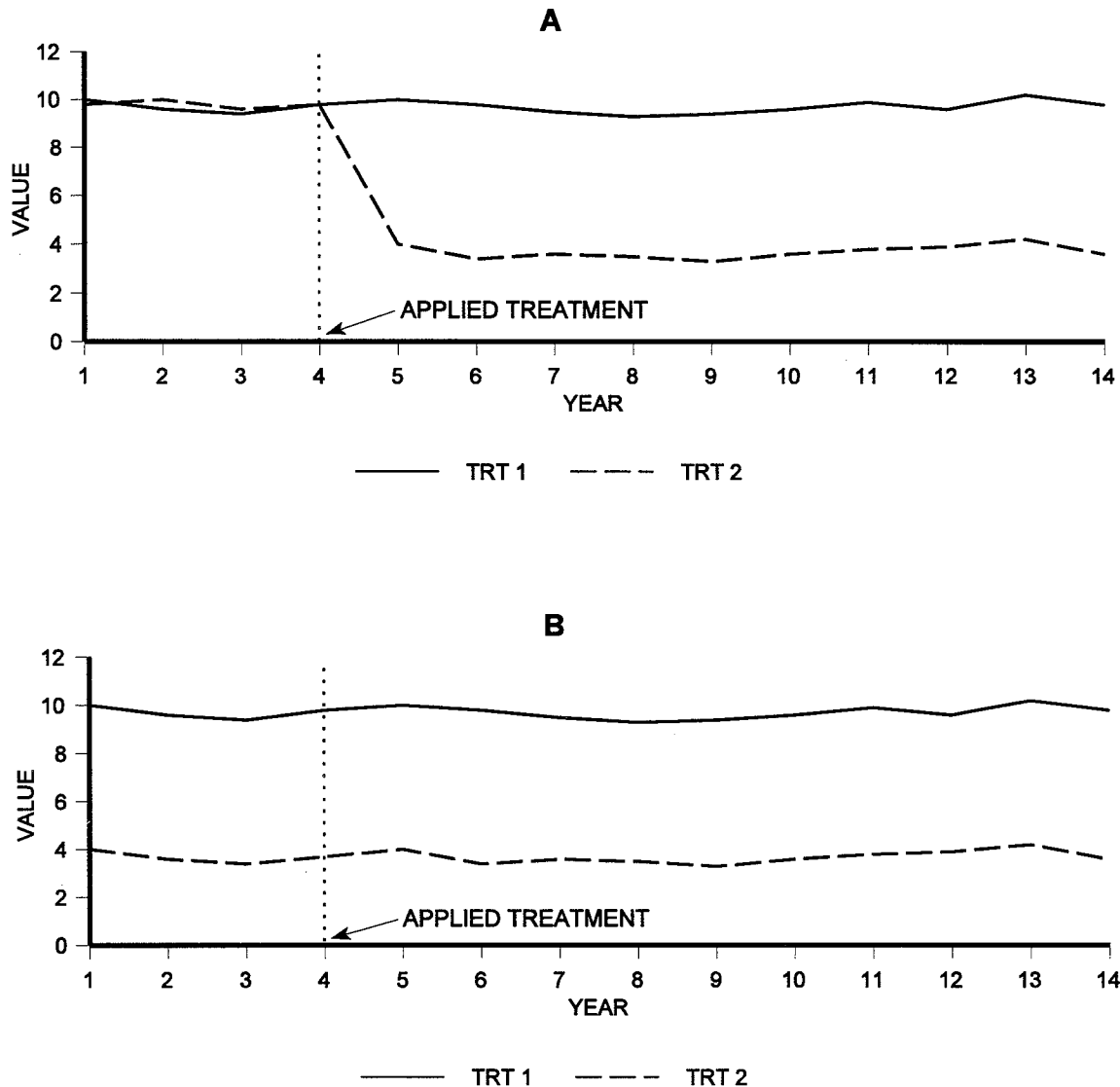


Figure 1.—Illustration of two experimental studies showing importance of pre-treatment data. Illustration A shows a treatment effect, whereas illustration B has no treatment effect.

data are added. Coordination of the data collection schedule in the long term will be critical for increasing ecological understanding.

One of the critical factors in executing MOFEP is the application of the two timber harvest treatments. The design requires that all six sites where timber is to be harvested receive their prescribed treatment within the same year. For example, the initial timber harvest had to be done within the 1996-97 cutting season (Brookshire *et al.* 1997). In subsequent re-entry periods, harvest should also be done within a cutting year unless the treatment prescriptions for uneven- and even-aged managed areas are

redefined due to some modification of standard forest management practices. If this does not occur, the experiment will not have adequate temporal replication because the applications of treatments will become staggered. If, at any point, timber harvesting is not completed on schedule, the entire experiment should be re-evaluated to determine the potential impacts on replication of the prescribed treatment.

Statistical Analysis

Many statistical models are available for analyzing data from MOFEP studies. The decision of which model to use must be based on the

nature of the response variables (continuous or discrete); assumptions (normality, independence, additive or multiplicative, structure of variance-covariance matrices, etc.); and the inference space, the extent that inferences can be applied in terms of landscape and temporal range restrictions. Also, the size of the MOFEP study (only nine sites are being used) must be taken into account. Because of the large number of possible statistical models to choose from, we will only discuss three basic models expressed in terms of analysis of variance to illustrate potential statistical approaches.

In our illustrations, we assume that the data adhere to the assumptions for the analysis of variance (Steel and Torrie 1980:167-170). The response variable for our illustration will be the differences between the pre-treatment and post-treatment means. In other words, we find the mean of the post-treatment data that were taken over a number of years and subtract it from the mean of the pre-treatment data taken over a similar number of years. This adjustment will take into account the problem illustrated earlier in figure 1.

The first and the simplest model (MODEL 1) can be used only to examine block and treatment effects (table 1). The error term used to test the hypothesis of no treatment effect is the interaction of the block and treatment effects. This understanding of the proper error term is important to remember when data used in the analysis are from a number of measurement plots that were randomly placed within each

site. These measurement plots are a subsample of the site and are used to estimate the response variable at the site level (Bergerud 1996). The among-measurement plot variation is not used in testing treatment effects. It is the variation at the site level that is important in this test. MODEL 1 can also be used to compare response variables measured at only one point in time (not repeatedly measured.)

Physiographic or some other characteristics of the sites can be influential in the ecosystem response to treatments. If sites are divided into physiographic characteristics, such as ecological land types (ELT), and studied separately to test the response of the variables to these characteristics, then a different statistical model is needed. The statistical model for this type of data is known as a split-plot analysis of variance (Steel and Torrie 1980:377-382). Table 2 illustrates this analysis of variance table for this design (MODEL 2). Within this model, two error terms exist. The first is the block-by-treatment interaction that is used to test the hypothesis of no treatment effects. This test is the same as in MODEL 1. The other error term, in this case, is used to test the hypothesis of ELT effects. As illustrated, this error term is the block-by-treatment-by-ELT interaction effect. This error term is used to test for ELT effect and ELT-by-treatment interaction effect. This error term can be pooled with the block-by-ELT interaction for testing purposes, but this must be done with caution (Hines 1996).

MODEL 1 and MODEL 2 do not use effectively

Table 1.—*Example of randomized complete block analysis of variance table with block and treatment effects (MODEL 1). This analysis would use site level data representing one point in time or a single measure from the sites.*

Source of variation	DF ¹	MS ²	F ³	P-value ⁴
Block	2			
Treatment	2			
Error ⁵	4			

¹Degrees of freedom.

²Mean square.

³Calculated F-statistic.

⁴Probability level of F-statistic.

⁵Error term for treatment effects. This error term is the interaction of block by treatment (Block*Treatment).

Table 2.—Example of split-plot design analysis of variance table with block and treatment main effects and each site split by Ecological Land Type (ELT) (MODEL 2). This analysis would use ELT information within each site representing one measure.

Source of variation	DF ¹	MS ²	F ³	P-value ⁴
Block	2			
Treatment	2			
Error a ⁵	4			
ELT ⁶	L-1			
Block*ELT	2*(L-1)			
Treatment*ELT	2*(L-1)			
Error b ⁷	2*2*(L-1)			

¹Degrees of freedom.

²Mean square.

³Calculated F-statistic.

⁴Probability level of F-statistic.

⁵Error term for treatment effects. This error term is the interaction of block by treatment (Block*Treatment).

⁶Ecological Land Type. L categories of ELT are used in this example.

⁷Error term for ELT effects and interaction of ELT by treatment. This error term is the three-way interaction of ELT, block, and treatment effects.

the repeated measures that occur through a sequence of years. In most of the studies that are conducted under MOFEP, the same plots are measured repeatedly through a sequence of years. MODEL 1 and MODEL 2 can only use data from one year at a time or by pooling data over the years, such as through a mean. Our final model for illustrative purposes uses repeated measures from plots across a sequence of years more efficiently than MODEL 1 or MODEL 2. Table 3 shows the split-plot design with repeated measures (MODEL 3). This profile analysis uses a multivariate analysis of variance approach (Littell *et al.* 1996, von Ende 1993). Data used in this approach will no longer be differences between pre-treatment and post-treatment means. The response variable for this analysis can take many forms. For example, the response variables can be the separate repeated measures through the pre-treatment and post-treatment periods. Polynomial growth curves are fit through time for each site. The polynomial coefficients are tested for differences among treatment and, in this case, ELT effects and their interactions. Another form the response variable might take is through differences between post-treatment measures for each year and an index of the pre-treatment measures. The index might be the

mean of the pre-treatment measures or even the measure that was taken during the last year of the pre-treatment period. This method will produce as many repeated measures as the number of post-treatment repeated measures used.

These models can also be used to analyze specific sets of data that do not overlap the time boundary between pre-treatment and post-treatment phases of MOFEP. For example, in this proceedings, most of the papers examine only pre-treatment information. During the pre-treatment phase, interest was focused on the block, year, and “pseudo-treatment” effects. We emphasize “pseudo-treatment,” because during this period harvesting of trees had not occurred on the sites assigned specific harvest treatments. For MODEL 1 and MODEL 2, data for these analyses would either be from a single year or a pooled measure across the study period (for example, the mean of a variable that was measured each year during the pre-treatment phase). For MODEL 3, data would be in the form of repeated measures (Littell *et al.* 1996) and would not have to be indexed.



Table 3.—*Example of split-plot design analysis of variance table with repeated measures (MODEL 3). Block and treatment are the main effects, each site is split by into ecological land type (ELT) effects, and year effects are a repeated measure of ELT within each site. Measurements would be made in each ELT category within each site over a number of years.*

Between site effects

Source of variation	DF ¹	MS ²	F ³	P-value ⁴
Block	2			
Treatment	2			
Error a ⁵	4			
ELT ⁶	T-1			
Block*ELT	2*(T-1)			
Treatment*ELT	2*(T-1)			
Error b ⁷	2*2*(T-1)			

Within site effects

Source of variation	Pillai's Trace ⁸	F ⁹	NumDF ¹⁰	DenDF ¹¹	P-value
Year					
Year*Block ¹²					
Year*Treatment ¹²					
Year*ELT					
Year*Block*ELT					
Year*Treatment*ELT					

¹Degrees of freedom.

²Mean square.

³Calculated F-statistic.

⁴Probability level of F-statistic.

⁵Error term for treatment effects. This error term is the interaction of block by treatment (Block*Treatment).

⁶Ecological Land Type. T categories of ELT are used in this example.

⁷Error term for ELT effects and interaction of ELT by treatment. This error term is the three-way interaction of ELT, block, and treatment effects.

⁸Pillai's Trace Statistic (Seber 1984:39-40).

⁹F-statistic derived from Pillai's Trace Statistic (Seber 1984:564).

¹⁰Calculated numerator degrees of freedom for the F-statistic.

¹¹Calculated denominator degrees of freedom for the F-statistic.

¹²Error matrix for testing these effects would be from the three-way interaction of year by block by treatment.

Interpretation of Results

The results of any experiment must be correctly interpreted within the constraints of the hypothesis, the experimental design, the execution of the experiment, and the statistical analysis to provide meaningful information about the impacts of treatments. Inferences made beyond the scope of these elements can be misleading and may cause harm to resources. Therefore, interpretation of results is a very important aspect of any experiment and is the ultimate purpose of the experiment. The responsibility of interpretation abides with both the researcher and the user of the information.

As much as the inference space relies upon the other components of the experiment, so do the experimental design, execution, and statistical analysis depend upon the desired inferences that researchers and managers wish from the project. MOFEP has been designed to allow some flexibility in the breadth of inferences that can be drawn from resulting data. This flexibility is granted through the assumptions that researchers and managers might wish to make when analyzing and interpreting results from different MOFEP studies.

For example, if a researcher is reporting to forest managers results that might be used in adapting a treatment during subsequent re-entry harvest periods, then the researcher might wish to assume that the blocks are fixed and represent themselves. In this manner, the blocks can be tested under an analysis of variance using blocks and treatments as fixed effects. However, if the researcher wishes to make statements about the impact of treatments on sites outside of the nine used in MOFEP, then the researcher would need to assume that the blocks represent random effects. In other words, the researcher wants to make statements about the potential treatment effects on a larger population of sites from which the nine sites used in MOFEP were randomly selected. Under the regime where blocks are assumed to be fixed effects, a fixed model analysis would be used in the analysis of variance, because both treatment and blocks are a "fixed" set of effects. Under the latter regime, where inferences would be drawn for sites beyond those used in MOFEP, a mixed model (Littell *et al.* 1996) would be appropriate. Blocks would represent the random effect and treatments would be fixed effects in this case.

Inferences may not be representative of the entire site due to availability of resources. Some researchers have had to confine their views to portions of each site. For example, the reptile and amphibian study examines only those reptile and amphibian populations within two major ELT classes (Renken 1997). Therefore, data from this study are not representative of the entire site, but are limited to the two ELT classes chosen. The experimental design will accommodate this restriction in study scope; however, the interpretation that might be made from these data must also be restricted.

WHY THE RANDOMIZED COMPLETE BLOCK DESIGN

The selection of a manipulative or experimental approach for MOFEP appears to be a logical choice, given the goal of showing cause-and-effect relationships among forest management practices. These impacts are believed to have an influence on biotic and abiotic components within the forest ecosystem. However, many other approaches could have been used in the design of MOFEP. A design could have been selected that would have used regression procedures as the basis for statistical analysis (Draper and Smith 1966). Or, we could have chosen a different experimental layout, such as completely randomized or an incomplete block design (Cochran and Cox 1957).

The regression procedure would have allowed for a wide variety of forest opening sizes to be tested at the site level. Under this design a site would have been randomly assigned a specified size of "clearcut" to be used for the duration of the study. These clearcut sizes could have ranged from zero acres for sites assigned as controls up to one-tenth of the size of a site given a 100-year rotation. The independent variable in the regression analysis would have been the sizes of the assigned "clearcuts" on which the dependent variables would have been regressed. This design would have restricted the options for forest managers in implementing the treatments and would not have allowed forest managers to use information in adapting forest management practices on the project.

Other experimental designs that use the pre- and post-treatment were considered. For example, an incomplete block design with four treatments replicated in three blocks each having three experimental units was considered. However, this design was discounted due



to the decreased power of the statistical analysis and inadequate replication of treatments.

A completely randomized design was also considered. Under this design each site would have been assigned a treatment at random so that each treatment would have appeared three times. This differs from the chosen design in that blocking of sites would not have occurred. This design forces the variation among sites in a completely randomized design into the experimental error and provides less accuracy than a design using blocking of the sites (Cochran and Cox 1957). On-site reconnaissance indicated that all the sites were not alike and that the three sites on Peck Ranch Conservation Area (PRCA) and their underlying soils were different from the other six sites.

Several other considerations that eliminated the use of the completely randomized design were discussed during the design phase of MOFEP. Discussions on the potential of adding other sites to MOFEP were an important factor in eliminating this design. Under a completely randomized design, adding other sites would not be possible without a re-randomization of treatments among sites. Also, the possibilities of site destruction due to some natural cause, such as tornado or fire, were discussed. If these problems impacted a block, then these impacts might also be studied and accounted for within a block design, but not under a completely randomized design. Under a completely random design, these problems could cause irreparable damage to MOFEP. To avoid these problems and to add flexibility, a completely randomized design was not chosen.

The reconnaissance of the MOFEP sites suggested that the nine sites could be divided into three blocks that were nearly homogeneous. That is to say, we would expect results from sites within each block to be more alike than results compared among blocks. Blocking would prevent the chance assignment of only one treatment type occurring on the three sites on PRCA. Therefore, the randomized complete block design was chosen. The randomized complete block design allows us to eliminate the variation due to differences among blocks (block effects) during data analysis. If variation among blocks is included as part of the experimental error, greater differences among treatments would be necessary before the impact of a

treatment might be found. This design also allows for flexibility in adding extra blocks at a later date and in different locations. Adding extra blocks makes the analysis more difficult and will require additional assumptions, such as impacts of temporal and spatial confounding.

We believe that the randomized complete block design is the best choice given the number of sites available. This design is simple and allows flexibility so that forest managers can adapt their practices to the state of the art at each re-entry period.

ECOSYSTEM PROJECTS WITHIN THE DESIGN

The beauty of MOFEP is the concept of allowing many different ecosystem components to be studied during the life of this project. The treatments will be replicated within the nine sites, and the response to the treatments will be documented through time. Because these sites will be consistently treated under a designed experiment, we have the opportunity to take measurements on a variety of environmental variables. MOFEP will be a valuable source of information for wildlife and forest researchers and managers.

As one might expect, not all types of variables are suitable for measurement within MOFEP. The restrictions on these variables are defined by statistical and practical considerations. During the process of including individual studies under the MOFEP umbrella, these restrictions were taken into account. In the future, as other individual studies are considered, we believe that these restrictions also will be applicable.

The first restriction concerns relationships among variables. Variables that are measured should not be the same ecologically. It does not make sense to measure essentially the same thing in several different ways. However, the selected variables may be related through their influence on each other in the ecological web of the forest ecosystem. MOFEP offers a design under which correlational responses of these interactions of ecological variables can be studied. Statistical modeling offers the opportunity of using data from several individual studies to explore and develop hypotheses about ecological connectivity among ecosystem components.

Methods for measuring each variable should not have an impact on other variables and community components. For example, if all the trees within each site needed to be cut to determine their weight, then tree mass probably should not be considered as a viable candidate to be measured. Therefore, methods and techniques for measuring the forest ecosystem need to be non-destructive in studies like MOFEP. If the process of measuring one variable causes an impact upon other components of the ecosystem, false inferences about treatment impacts could be the result. These false inferences could cause forest managers to make decisions that could damage the forest.

The size, shape, and juxtaposition of the sites needed to be considered in selecting proper variables to study. For example, it would be unwise to measure wild turkey densities on a site-size area. Because turkeys have such a large home range, the numbers of turkeys would vary greatly within any set of given days. This variability would most likely cause the measurement error of density within the sites to be greater than the amount of variability among treatments. The most logical conclusion from this highly variable measure of turkey density would be that treatment could not be shown to have an effect upon turkey density. Therefore, the area of influence that affects variables had to be taken into account, and some important forest ecosystem variables cannot be studied under the MOFEP design due to scale problems.

A restriction that occurs in every research project also affects MOFEP. This restriction is caused by a limited amount of resources—financial, space, and time. It must be cost-effective to collect the data. For example, ground litter invertebrates were found to be highly variable within a site (Weaver and Heyman 1997). To obtain a reliable and precise estimate of these invertebrates for a site would have required a large army of entomologists to collect and classify the samples. The expenditure would have been prohibitive for a ground litter invertebrate study that met the objective of determining the impact of forest management on these invertebrates. Therefore, the objective was changed for ground litter invertebrates to make it cost-effective and accomplishable within a reasonable time (Weaver and Heyman 1997).

Once a variable had been selected for study, proper statistical sampling procedures needed

to be identified to ensure that data were representative of the site or some smaller subdivision of the site (Cochran 1977, Thompson 1992). The sampling procedures had to include adequate sample size to obtain a reasonably precise estimate. If the estimates were not reasonably precise due to inadequate sample size or biased due to lack of randomization, then the data could lead to false inferences. Overlaying of individual studies on MOFEP's experimental design required that variables be ecologically dissimilar, the act of measuring them did not impact other ecological components, the precision of each variable was adequate so that the measurement error within sites did not exceed the variation among sites, and data could be collected in a timely and cost-effective manner. The selected variables had to be measured following proper sampling procedures to ensure that data would be representative of the population of interest.

LIMITATIONS OF THE MOFEP DESIGN

MOFEP has a solid experimental design. The randomized complete block design offers many opportunities to examine the impact of forest management on a broad array of ecosystem components; however, MOFEP does have limitations. MOFEP's biggest limitation is that statistical power may be low in most cases (Hurlbert 1984, McAllister and Peterman 1992, Peterman 1990, Steidl *et al.* 1997, Toft and Shea 1983). The statistical power will be low in detecting differences among treatments when the treatment effect is small relative to the experimental error. MOFEP has only three replicates for each treatment. The ability to detect a significant difference among treatments under a null hypothesis of equivalence is usually poor when so few replicates (i.e., small degrees of freedom) are used. The differences among treatments will have to be large in comparison to the experimental-wise error for a statistically significant difference to be detected. In all likelihood, researchers in the field probably will suspect biological differences before they are able to detect them through statistical analyses. We need to be cognizant that even though we might not reject a null hypothesis with the data, this does not mean that forest management practices are not impacting the system in some positive or negative manner.

The problem of not rejecting a null hypothesis when in fact a treatment effect exists (called Type II error) is a major issue concerning



MOFEP. The importance of knowing the probability of detecting a difference if it exists cannot be cast aside as irrelevant (Forbes 1990, Peterman 1990, Simberloff 1990, Steidl *et al.* 1997, Toft 1990, Toft and Shea 1983). So, the question is what can be done in light of low statistical power. A larger probability (the α -level) can be used for determining if the null hypothesis should be rejected. For example, instead of using the usual α -level of 0.05 for a statistical test to show a significant difference, a probability level of 0.10, or even 0.15, might be used. The α -level is inversely related to the probability of making a Type II error (Forbes 1990). Therefore, as the selected α -level becomes larger, the likelihood decreases that a false null hypothesis is accepted. It is important that the α -level be established before data collection and during the design phase of the experiment. Instead of setting large α levels, a better alternative might be the use of confidence intervals on the estimated differences between treatment means (Steidl *et al.* 1997, Gary White, personal communication). This method provides information about the range where differences between treatments are masked by the error.

Another limitation with the MOFEP design is the limited population of sites represented by the nine sites used in this project. In an attempt to find suitable sites that could be included in this long-term study, only the nine sites used in MOFEP met the criteria of age and homogeneity (Brookshire *et al.* 1997). These sites are relatively close in proximity (fig. 1 in Brookshire *et al.* 1997), and are all located on Missouri Department of Conservation lands. Because of their close proximity and land ownership, the "population" of sites represented by these nine sites, probably in strict terms and definitions, is these nine sites. Therefore, researchers and forest and wildlife managers will need to be very careful in making their inferences and extrapolating results beyond MOFEP project sites.

The small number of sites available also made it impossible to replicate the treatments temporally. Weather and possibly other abiotic and biotic components that vary annually impact results. The initial treatment (cutting of trees) was applied to all sites in the same year under one set of temporal impacts. A different set of results might be possible due to conditions in another year when treatments could have been applied. Not enough sites exist to apply timber

harvest to sets of sites over several years under this replicated design. For example, four additional sites per block would have been needed to replicate the two timber harvest practices over a 3-year period. This would have allowed us to determine if temporal effects were present during the 3 years when trees were cut, but we would have been unable to detect longer temporal trends. Simply put, the results from MOFEP will represent the "population" of sites that will be cut the same year as we applied the initial timber harvest in MOFEP.

A catastrophic event, such as wildfire or tornado, within one or more sites would cause a major problem for MOFEP because of the low statistical power of the design. If a single site were affected by a catastrophic event, then the design would be unbalanced (Littell *et al.* 1996). At worst, only the statistical power would be affected under this type of circumstance. If an entire block of sites were affected by the event, the design could accommodate this problem. If catastrophic events destroy more than one site in different blocks, judgments about merits of continuing MOFEP will have to be made. The design may be too heavily impacted by this problem to provide meaningful results for all treatments.

MOFEP AS ADAPTIVE MANAGEMENT

Walters and Hilborn (1976) presented the concept of using adaptive control processes in managing natural resources. From this basic concept, adaptive resource management has grown into a management concept of learning while managing (Walters 1986). MOFEP follows this concept of allowing forest managers to learn from the results and to adapt their practices to reach their management goals (Walters 1993).

The principal forest vegetation management practices used by the Missouri Department of Conservation are even-aged, uneven-aged, and no harvest. These practices are competing models of forest management. Each management practice has a different path in achieving the goal of maximum forest diversity over an infinite time horizon (Larsen 1997), but the impacts on specific forest ecosystem components are not known under each management model.

The experimental design of MOFEP allows forest managers to adapt their management style

within the "flexible" protocol established for each model. The most restrictive model is no-harvest management. This model does not generally allow the forest manager to manipulate any forest stands within these assigned sites. Under the two timber harvest models, forest managers actively manipulate and manage the forest for economic and biological gains (Brookshire *et al.* 1997).

This approach for MOFEP is very passive adaptive management, but it differs significantly from a pure experimental approach. Under a pure experimental approach, researchers would wait until the end of the experiment to analyze their data. Forest managers would be given a set of very restrictive prescriptions for each model, and they would not be allowed to deviate from these prescriptions throughout the life of the experiment. In other words, we could not learn from the results until after the experiment was completed (several hundred years from the start of MOFEP). Adaptive resources management, however, gives us the opportunity to learn while managing through a less restrictive experimental approach (Walters 1993). As Carl Walters says (personal communication), adaptive experiments are necessary to make learning ever happen in situations such as MOFEP.

As foresters adopt dynamic numeric models in their forest management planning, MOFEP will progress from very passive adaptive management to a more active adaptive management approach. Forest managers and researchers will be able to use the data that will be collected and analyzed to develop, evaluate, and change these dynamic models. Numeric procedures, such as stochastic dynamic programming (Lubow 1995, Lubow 1996, Puterman 1994), can be used to optimize timber harvest practices through adaptive resources management (Conroy and Crocker 1996). As data are collected on each site within MOFEP, forest managers can use this information to develop management plans that will establish a more rapid path for achieving optimal resource objectives (Walters 1986).

A cautionary note is important here, because no guarantee can be made that any of the three management practices under study in MOFEP is the "best" for achieving the goal of maximum diversity (Carl Walters, personal communication). Some other practice may actually be the "best." MOFEP can be used only to judge the

regime that is the "winner" among these three practices as forest managers adapt their management based on information that is obtained through this experimental approach.

SUMMARY

The experimental approach used in MOFEP will provide results demonstrating cause-and-effect relationships among forest management practices in the associated Ozark forest communities of Shannon, Carter, and Reynolds Counties. These results must be interpreted with the realization of the locational, scalar, and temporal limitations of MOFEP. What makes MOFEP such a unique project is the replication and randomization of treatments. Using these experimental procedures reduces the risk of biased or misleading results. Reliable knowledge about forest management and its impact on forest ecosystems can, and will be, gained under this experimental approach. However, because of the low statistical power of the MOFEP design, results that are not significant in a statistical sense will have to be scrutinized, through the use of confidence intervals of the differences, to determine if one of the treatments might have an impact in a biological sense (Steidl *et al.* 1977). Conversely, if results show statistical significance, we will be assured that differences were large among treatments.

Due to the limitations of the MOFEP design, no simple analytical model is the "best" procedure for determining treatment effects. We foresee that further research of better statistical analysis techniques will have to occur. To derive all the valuable insights possible, data collected under MOFEP will require sophisticated statistical methods that do not exist at present. Research into areas of variance-covariance structure modeling (Littell *et al.* 1996) is needed. As more data are collected, greater insights will be gained about the nature and structure of the information that MOFEP can supply.

The long-term nature of MOFEP is mind-boggling. Realizing that anyone born on the day that MOFEP started will not be alive to see the successful conclusion of this project, one becomes aware of the significance and magnitude of this research project. But throughout the life of MOFEP, managers will be able to use results obtained from individual studies within the project to establish better management practices. Researchers at the same time will be



able to develop new hypotheses to be tested and use new analytical tools to obtain more information from the data. Information from MOFEP will be invaluable to wildlife and forest managers for generations yet to come.

ACKNOWLEDGMENTS

The design and implementation of MOFEP cannot be credited to us or to any single person. MOFEP is the result of a large group of people who have cooperated in tackling a major question of concern to forest and wildlife managers. Cooperation has been the key to MOFEP and its success thus far. The spirit of cooperation emerged during a Wildlife Research Section staff meeting where Rick Clawson and John Faaborg presented their idea of studying birds in a forest that was being fragmented. This cooperative spirit has continued through the design phase and to the point we find ourselves now, at the end of the pre-treatment data collection period. For MOFEP to succeed, cooperation will have to be the standard until the last report is written in several hundred years.

MOFEP would not be possible without the invaluable commitment of the Missouri Department of Conservation. The Department offers cooperating scientists from the academic and natural resource management environments an opportunity to study and address issues in forest management through MOFEP. For the Department's commitment, we and all of the cooperating scientists are very thankful.

As for designing the experiment, many people helped in this regard, and all of these people deserve a great deal of credit for their foresight and insight. We wish personally to thank Larry Vangilder and Steve Westin for helping with the design and an earlier presentation and paper from which we have adopted this manuscript. Their professionalism and opinions have added greatly to MOFEP. We also wish to thank Gary Krause and Mark Eilersieck, University of Missouri, for their help in discussions about the merits of different statistical designs. Carl Mize, Iowa State University, has also provided many insightful comments on the limitations and statistical analysis approaches of MOFEP. We thank Gary White, Gary Krause, Carl Mize, Carl Walters, and an anonymous reviewer of our manuscript for their many suggestions and comments on our paper. Their comments definitely strengthened our message.

LITERATURE CITED

- Bergerud, Wendy A. 1996. Displaying factor relationships in experiments. *The American Statistician*. 50: 228-233.
- Brookshire, B.L.; Hauser, C.H. 1993. The Missouri Ozark Forest Ecosystem Project: the effects of forest management on the forest ecosystem. In: Gillespie, A.R.; Parker, G.R.; Pope, P.E.; Rink, G.R., eds. *Proceedings of the 9th central hardwood forest conference*; 1993 March 8-10; West Lafayette, IN. Gen. Tech. Rep. NC-161. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station: 289-307.
- Brookshire, B.L.; Jensen, R.; Dey, D.C. 1997. The Missouri Ozark Forest Ecosystem Project: past, present, and future. In: Brookshire, Brian L.; Shifley, Stephen R., eds. *Proceedings of the Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research*; 1997 June 3-5; St. Louis, MO. Gen. Tech. Rep. NC-193. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station: 1-25.
- Cochran, William G. 1977. *Sampling techniques*. 3d ed. New York, NY: John Wiley and Sons. 428 p.
- Cochran, William G.; Cox, G.M. 1957. *Experimental designs*. New York, NY: John Wiley and Sons. 611 p.
- Conroy, Michael J.; Crocker, J.E. 1996. Effect of uncertainty in mapped biodiversity data on optimal conservation decisions. In: Mowrer, H.T.; Czaplewski, R.L.; Hamre, R.H., eds. *Spatial accuracy assessment in natural resources and environmental science: 2d international symposium*; 1996 May 21-23; Fort Collins, CO. Gen. Tech. Rep. RM-277. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station: 685-692.
- Draper, Norman R.; Smith, H. 1966. *Applied regression analysis*. New York, NY: John Wiley and Sons. 407 p.

- Eberhardt, L.L.; Thomas, J.M. 1991. Designing environmental field studies. *Ecological Monographs*. 61: 53-73.
- Forbes, L. Scott. 1990. A note on statistical power. *The Auk*. 107: 438-439.
- Green, Roger H. 1979. Sampling design and statistical methods for environmental biologists. New York, NY: John Wiley and Sons. 257 p.
- Hines, W.G.S. 1996. Pragmatics of pooling in ANOVA tables. *The American Statistician*. 50: 127-139.
- Hurlbert, Stuart H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*. 54: 187-211.
- James, Frances C.; McCulloch, C.E. 1985. Data analysis and the design of experiments in ornithology. In: *Current ornithology*. Vol. 2. New York, NY: Plenum Press: 1-63.
- Kurzejeski, Eric W.; Clawson, R.L.; Renken, R.B.; Sheriff, S.L.; Vangilder, L.D.; Hauser, C.; Faaborg, J. 1993. Experimental evaluation of forest management: the Missouri Ozark forest ecosystem project. *Transactions of the North American Wildlife and Natural Resources Conference*. 58: 599-609.
- Larsen, D.R. 1997. Simulated long-term effects of the MOFEP cutting treatments. In: Brookshire, Brian L.; Shifley, Stephen R., eds. *Proceedings of the Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research; 1997 June 3-5; St. Louis, MO*. Gen. Tech. Rep. NC-193. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station: 347-355.
- Littell, Ramon C.; Milliken, G.A.; Stroup, W.W.; Wolfinger, R.D. 1996. SAS system for mixed models. Cary, NC: SAS Institute Inc. 633 p.
- Lubow, Bruce C. 1995. SDP: generalized software for solving stochastic dynamic optimization problems. *Wildlife Society Bulletin*. 23: 738-742.
- Lubow, Bruce C. 1996. Optimal translocation strategies for enhancing stochastic metapopulation viability. *Ecological Applications*. 6: 268-280.
- McAllister, Murdoch K.; Peterman, R.M. 1992. Experimental design in the management of fisheries: a review. *Journal of Fisheries Management*. 12: 1-18.
- Meinert, D.; Nigh, T.; Kabrick, J. 1997. Landforms, geology, and soils of the MOFEP study area. In: Brookshire, Brian L.; Shifley, Stephen R., eds. *Proceedings of the Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research; 1997 June 3-5; St. Louis, MO*. Gen. Tech. Rep. NC-193. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station: 56-68.
- Peterman, Randal M. 1990. Statistical power analysis can improve fisheries research and management. *Canadian Journal of Fisheries Aquatic Science*. 46: 1183-1187.
- Puterman, Martin L. 1994. Markov decision processes: discrete stochastic dynamic programming. New York, NY: John Wiley and Sons. 649 p.
- Renken, R.B. 1997. Pre-treatment conditions of herpetofaunal communities on Missouri Ozark Forest Ecosystem Project (MOFEP) sites, 1992-1995. In: Brookshire, Brian L.; Shifley, Stephen R., eds. *Proceedings of the Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research; 1997 June 3-5; St. Louis, MO*. Gen. Tech. Rep. NC-193. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station: 289-308.
- Robinson, Scott K.; Thompson, F.R., III; Donovan, T.M.; Whitehead, D.R.; Faaborg, J. 1995. Regional forest fragmentation and the nesting success of migratory birds. *Science*. 267: 1987-1990.
- Romesburg, H. Charles. 1981. Wildlife science: gaining reliable knowledge. *Journal of Wildlife Management*. 45: 293-313.
- Seber, G.A.F. 1984. Multivariate observations. New York, NY: John Wiley and Sons. 686 p.
- Simberloff, Daniel. 1990. Hypotheses, errors, and statistical assumptions. *Herpetologica*. 46: 351-357.



- Steel, Robert G.D.; Torrie, J.H. 1980. Principles and procedures of statistics: a biometrical approach. 2d ed. New York, NY: McGraw-Hill Book Co. 633 p.
- Steidl, Robert J.; Hayes, J.P.; Schaubert, E. 1997. Statistical power analysis in wildlife research. *Journal of Wildlife Management*. 61: 270-279.
- Thompson, Frank R., III; Dijak, W.D.; Kulowiec, T.G.; Hamilton, D.A. 1992. Breeding bird populations in Missouri Ozark forests with and without clearcutting. *Journal of Wildlife Management*. 56: 23-30.
- Thompson, Steven K. 1992. Sampling. New York, NY: John Wiley and Sons. 343 p.
- Toft, Catherine A. 1990. Reply to Seaman and Jaeger: an appeal to common sense. *Herpetologica*. 46: 357-361.
- Toft, Catherine A.; Shea, P.J. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. *American Naturalist*. 122: 618-625.
- von Ende, Carl N. 1993. Repeated-measures analysis: growth and other time-dependent measures. In: Design and analysis of ecological experiments. New York, NY: Chapman and Hall: 113-137.
- Walters, Carl J. 1986. Adaptive management of renewable resources. New York, NY: McGraw-Hill Book Co. 374 p.
- Walters, Carl J. 1993. Dynamic models and large scale field experiments in environmental impact assessment and management. *Australian Journal of Ecology*. 18: 53-61.
- Walters, Carl J.; Hilborn, R. 1976. Adaptive control of fishing systems. *Journal of Fisheries Research Board of Canada*. 33: 145-159.
- Weaver, J.; Heyman, S. 1997. The distribution and abundance of leaf litter arthropods in MOFEP sites 1, 2, and 3. In: Brookshire, Brian L.; Shifley, Stephen R., eds. Proceedings of the Missouri Ozark Forest Ecosystem Project symposium: an experimental approach to landscape research; 1997 June 3-5; St. Louis, MO. Gen. Tech. Rep. NC-193. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station: 318-331.
- White, Gary C.; Garrott, R.A. 1990. Analysis of wildlife radio-tracking data. New York, NY: Academic Press. 383 p.