PROJECT TITLE: Evaluating the use of the belt transect method in determining native plant composition changes in upland and wet meadow habitats on Lacreek National Wildlife Refuge: Project Update FY 2013

PROJECT DESCRIPTION: The 2006 Lacreek National Wildlife Refuge (NWR) Comprehensive Conservation Plan (CCP) established management goals of restoring and enhancing the upland and wet meadow plant communities to create a mosaic that reflects the habitat requirements of grassland birds of management concern. Specifically, the CCP states that these two habitat types will be managed to provide the plant community composition and structure necessary to support selected birds of concern that occur in this ecoregion. Birds of concern with similar habitat requirements were grouped based on patch size, vegetation height requirements, and tolerance to trees. The CCP objectives for the upland and wet meadow habitats is to increase floristic quality assessment (FQA) "C" scores (Stohlgren et al. 1995) with in tall, medium, and short patches that are a given distance from trees.

Following completion of the CCP, a monitoring program based on FQA was developed to assess progress in achieving the upland and wet meadow objectives. Initially, plots (20m x 50m) were established and measured twice in upland (n =5) and wet meadow (n = 5) communities to evaluate this methodology. This pilot effort identified several disadvantages of using this approach. Including the amount of time and personnel required to conduct this form of monitoring, and this method measures a "tolerance to disturbance." Therefore, a plot may have a low "C" score but be composed of native species that are tolerant to disturbances such as burning and/or grazing. Based on the pilot results, a more rapid method was necessary to determine changes in native plant composition and structure and to provide information for management direction. The purpose of this project was to assess the potential of using the belt transect method to determine changes in native species composition and structure on upland and wet meadow habitats on Lacreek NWR.

PRIMARY OBJECTIVE: Determine if the belt transect method can be used to detect a 10% increase in native plant composition on upland and wet meadow habitats on Lacreek NWR within a 10 year period with 80% confidence.

METHODS AND PROTOCOLS: The belt transect method (Grant et al. 2004) was used to measure vegetation in selected refuge units two or more times. These included units that received disturbance and two units that received no disturbance. This data was analyzed to provide an overall summary of composition for the unit, and will be used to determine a standard deviation of the difference of the paired measurements.

DATA ANALYSIS: (The following was taken directly from statistical report provide by Brian Cade-USGS). The framework for analyzing these data and determining sample size was a confidence interval on mean differences associated with a paired *t*-test on mean differences, where pairs are the individual transects measured at two different time points. A two-tailed 80% confidence interval was proposed as an acceptable degree of uncertainty to characterize estimates. I used both traditional power calculations for paired *t*-tests with power set to 0.80 to detect a mean difference of at least 10% and power for the two one-sided test approach for equivalence tests associated with an inequivalence null hypothesis. However, all power/sample size estimation can be a bit misleading as it doesn't directly indicate the magnitude of confidence intervals that will be estimable and interpreted with respect to quantities of interest (here the ≥10% equivalence region). Thus, following Bacchetti (2010), I provide a sensitivity analysis of sample size, estimated mean differences, and estimated standard deviations of differences on confidence interval half-widths (Millard and Neerchal 2001:471). For symmetrical confidence intervals around an estimated mean difference, plots of confidence interval half widths against estimated standard deviations by sample size provide a direct indication of whether for a given sample size you will have enough precision to interpret small estimated mean differences as conclusively <10% (Examples 1 and 3 in Figure 1) and whether you have enough precision and magnitude of estimated mean differences exceeding 10% to interpret them as conclusively ≥10% (Example 5 in Figure 1). The confidence interval half-widths serve both as a measure of precision and a measure of how much greater than 10% an estimate mean difference will have to be with that level of precision to be conclusively interpreted as exceeding 10%.

The power and confidence interval half-width approaches both require some reasonable range of standard deviations of the paired differences to be input into the statistical routines. I used the range of estimates for standard deviations of paired differences for cool season native grasses provided by preliminary data from Lacreek NWR. Unit 10SE-1 had $n = 11$ transects for 2008 and 2013 with an estimated standard deviation of differences (SD) = 39.2%. Unit SH-1 had $n = 5$ transects for 2009 and 2013 with an estimated SD = 31.9%. Unit 10NE-2 had $n = 11$ transects for 2007 and 2010 with an estimated SD = 58.2%. Unit 10SW-3 had $n = 5$ transects for 2010 and 2012 with an estimated SD = 21.4%. Therefore, I used a range of SD from 10% to 50% in the power analyses and confidence interval half-width analyses, providing sample size estimates for standard deviations estimated more precisely than those currently obtained to those estimated very imprecisely.

<u>Sample size estimates from power analyses</u>

The power analyses from a paired *t*-test with $\alpha = 0.20$, power = 0.80, two-tailed alternative to detect a difference of at least 10% yield estimated samples sizes of $n = 6$ to $n = 114$, with increasing size of SD (Table 1). The comparable power analyses for the two one-sided tests of null hypothesis of inequivalence for a paired design with $\alpha = 0.20$, power = 0.80, two-tailed alternative, with true mean differences of 15% and 20% for an equivalence region of ≥10% yielded estimate sample sizes that were much larger (Table 1). The required sample size for the same power is much greater for the equivalence hypothesis testing approach because it is trying to maximize the probability of keeping the 80% confidence intervals within the equivalence region of ≥10%. The more the true mean difference is assumed to exceed 10%, the smaller the required sample size for the same power for the null hypothesis of inequivalence in the two one-sided test.

Table 1. Estimated sample sizes to detect either a 10% difference from a paired *t*-test or to establish equivalence for a ≥10% equivalence region with assumed mean differences of 15% and 20% based on two one-sided paired *t*-tests for a range of standard deviations of the differences, α = 0.20, power = 0.80, two-tailed alternative.

Power analysis type	Standard deviation of mean differences				
	10%	20%	30%	40%	50%
Difference	6	19	41	73	114
Equivalence, 15%	24	92	205	364	568
Equivalence, 20%	7	24	52	92	143

<u>Sample size estimates from sensitivity analyses of confidence interval half-widths</u>

Plots of the half-widths of confidence intervals for a range of SD from 10% to 50% and by *n* = 10, 20, 30, 40, and 60 are shown in Figure 2. Obviously, the half-width of confidence intervals decreases with both increasing sample size and decreasing SD. A few simple rules can help interpret these relationships relative to detecting changes greater than 10%. Anytime the confidence interval half-width is less than 10%, an 80% confidence interval on an estimated mean difference = 0 would be completely less than the equivalence region of ≥10% providing conclusive evidence that the change has not exceeded 10% (as in Examples 1 and 3 in Figure 1). This is only true at for all SD from 10-50% for *n* = 60 but nearly true for *n* = 40. The confidence interval half-width also provides an estimate of how much greater than 10% an estimated mean difference would have to be given a sample size and SD for an 80% confidence interval to be completely within the equivalence region of ≥10% providing conclusive evidence that change has exceeded 10% (similar to Example 5 in Figure 1). For example, if we want an estimated 15% mean difference in cover to have 80% confidence intervals that are completely within the equivalence region then the half-width of the confidence interval must be <5%. This can be achieved with an *n* = 10 at SD = 10%, but will require *n* = 30 at SD = 20%, and *n* = 60 at SD = 30% (Figure 2). Similarly, an estimated 20% mean difference in cover will have 80% confidence intervals completely within the equivalence region if the half-width of the confidence interval is <10%. This can be achieved with *n* = 30 for SD = 10-40% but will require *n* = 60 for SD = 50% (Figure 2, Table 2). Confidence intervals for a range of estimated mean differences, standard deviations of differences, by sample size are provide in Table 2 to aid with this sensitivity analysis.

<u>Recommendations</u>

A default minimum sample size of *n* = 40 transects for each vegetation unit of interest would be a reasonable compromise if it is reasonable to expect the standard deviations of paired differences between 10 year time periods to be as high as 40%. The standard deviations that I used in my evaluations were based on paired differences between 3-5 years apart. With a sample size of *n* = 40, any mean differences near 0% would have 80% confidence intervals completely <10% providing conclusive evidence that there wasn't a 10% increase in mean differences, and if the estimated mean differences were 20% or greater would have 80% confidence intervals completely ≥10% providing conclusive evidence that the mean differences did exceed 10%. If standard deviations of the paired differences are 50% or greater, larger sample sizes would be required to make similar statements. Smaller sample sizes of $n = 20$ might be justified if there are reasons to think that the standard deviations of paired differences between 10 years are more likely to be $\leq$30%. If it is desirable to ensure that smaller mean differences such as 15% are estimated with sufficient precision to be declared conclusively $\geq$10%, then larger sample sizes will be required, e.g., $n = 60$ would be sufficient for standard deviations of paired differences $\leq$30%. Alternatively, reducing the standard deviations of paired differences to values $\leq$20% by either refining the measurement technique or by reducing heterogeneity through stratifying vegetation units into more homogeneous types would reduce the sample sizes required.

DATA MANAGEMENT: Data for the belt transects was summarized and is currently stored in Excel, Access, and R.

PARTNERS: Two people have been involved in the project: Shilo Comeau representing the US Fish and Wildlife Service (Lacreek NWR) and Brian Cade representing the US Geological Service (Fort Collins Science Center).

SOURCES OF SUPPORT: None to report.

CURRENT STATUS: Data has been summarized, analyzed and recommendations made as to how many transects are need per unit to detect a 10% change over a 10 year period.

CHALLENGES: Because of communication problems between the USFWS staff members the monies were obligated to USGS staff much later then planned so the end date was extended to the end of CY 2013, also the government shutdown has delayed the statistician from conducting analysis.

MORE INFORMATION:

Shilo Comeau
Lacreek NWR
(605)685-6508 (phone)
shilo_comeau@fws.gov

LITERATURE CITATION:

Bacchetti, P. 2010. Current sample size conventions, flaws, harms, and alternatives. BioMed Central Medicine 8:17.

Cade, B. S. 2011. Estimating equivalence with quantile regression. Ecological Applications 21: 281-289.

Cade, B. S., and P. R. Johnson. 2011. Quantile equivalence to evaluate compliance with habitat management objectives. Journal of Fish and Wildlife Management 2(2): 169-182.

Grant, T. A., E. M. Madden, R. K. Murphy, M. P. Nenneman, and K. A. Smith. 2004. Monitoring native prairie vegetation: the belt transect method. Ecological Restoration 22:106-111.

McBride, G. B. 1999. Equivalence tests can enhance environmental science and management. Australian and New Zealand Journal of Statistics 41: 19-29.

Millard, S. P., and N. K. Neerchal. 2001. Environmental statistics with S-plus. CRC Press, Boca Raton, FL.

Stohlgren, T.J., M.B. Falkner, and L.D. Schell. 1995. A modified-Whittaker nested vegetation sampling method. Vegetatio 117(2): 113-121.