Brood-count Power Estimates of Rio Grande Turkey Production in Texas

T. Wayne Schwertner, *Texas Parks and Wildlife Department, P.O. Box 1583, Mason, TX 76856*

Markus J. Peterson, *Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, TX 78744*

Nova J. Silvy, *Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, TX 78744*

Fred E. Smeins, *Department of Rangeland Ecology and Management, Texas A&M University, College Station, TX 78744*

Abstract: Brood counts are used frequently by state conservation agencies to estimate wild turkey *(Meleagris gallopavo)* recruitment. We performed power analyses for 25 years of Rio Grande Wild Turkey (*M. g. intermedia;* RGWT) brood-count data from five ecological regions of Texas in order to determine if these data had sufficient ($1–\beta \geq 0.80$) power to detect inter-annual and long-term changes in turkey production of 10%–20%, which we considered biologically meaningful. We then analyzed the data to determine trends in production. The analyses showed that a minimum annual sample of 200–500 turkey-group observations per region was required to detect an inter-annual change of 10%–20% in the proportion of poults in the hen:poult population. Historic annual sample size averaged 65–306. Existing data were not sufficiently powerful to detect long-term changes of 10%–20% in poult proportions. Brood counts, as currently conducted, appear to be ineffective at detecting biologically meaningful changes in RGWT recruitment in Texas. Further, non-random sampling methodology may render data unreliable. We recommend reevaluation of TPWD's efforts to monitor RGWT populations, as well as investigations into the sensitivity of RGWT populations to changes in recruitment. Our results show that power analysis offers a powerful tool for designing and evaluating population monitoring schemes.

Key words: Meleagris gallopavo intermedia, production, power, survey, Texas

Proc. Annu. Conf. Southeast. Assoc. Fish and Wildl. Agencies 57:213–221

Power analysis is a statistical technique whereby an investigator estimates the probability of committing a Type II statistical error, given the data examined. Whereas Type I error rate (α) is the probability of rejecting H_0 when H_1 is false, Type II error rate (β) is the probability of failing to reject H_0 when H_1 is true. Power of a statistical test ($1–\beta$), therefore, is the probability of rejecting H_0 when H_1 is true, and is a function of population standard deviation (σ), sample size (N), α, and the hypothesized (or actual) difference between population means or proportions ("effect size" or δ; Ott and Longnecker 2001).

Although statistical power is a fundamental statistical concept (Zar 1999), power analysis was rarely employed in the wildlife sciences prior to the mid-1990s (Steidl et al 1997). Since that time, however, it has enjoyed increasing prominence. The Wildlife Society (1995) suggested several ways in which power analysis could be used in wildlife research, including calculation of required sample sizes prior to performing wildlife studies and the *a posteriori* interpretation of study results (so-called "retrospective power analysis"). Although Gerard et al. (1998) questioned the validity of retrospective analysis on theoretical grounds, Steidle et al. (1997) observed that retrospective power analysis had utility if calculated using effect sizes other than the observed effect size.

Several investigators have used power analyses to design wildlife population monitoring efforts (Gibbs and Melvin 1997, Crouch and Paton 2002). Others have used power analysis to evaluate existing wildlife surveys. Lougheed et al. (1999) used retrospective power analysis to evaluate ongoing waterfowl surveys in Canada, finding that the surveys had sufficient power to detect a 5% trend had one existed, although power, and hence survey duration required to detect a trend, varied among species. Rice (2003) evaluated the power of ring-necked pheasant *(Phasianus colchicus)* call and brood counts in Washington, and determined that both methods had sufficient power to detect only very large (40%) year-to-year changes.

Recruitment may be the demographic parameter most important in determining wild turkey abundance trends (Roberts and Porter 1996). Hen:poult ratios, calculated from observations of turkeys during the brood-rearing season, are used as an index of recruitment by several states (Kurzejeski and Vangilder 1992). Observations of hens and poults are recorded by conservation personnel during the summer months either incidental to other duties (Schulz and McDowell 1957, Wunz and Shope 1980) or along predetermined routes (Shaw 1973, Menzel 1975, Bartush et al. 1985). Texas Parks and Wildlife Department (TPWD) has collected incidental RGWT brood observations across the range of the subspecies since 1976 (TPWD, unpublished data). Although usually referred to as a "survey," this technique is best classified as a "convenience" or "haphazard" sampling (Anderson 2001, Morrison et al. 2001). This is the only method by which RGWT populations are currently monitored; however, we found no published assessment of the power of brood counts to detect changes in turkey production. Therefore, the objective of this study was to evaluate the power of TPWD brood counts for detecting changes in RGWT production across broad spatial scales. Specifically, we calculated the power to detect differences among years and between two consecutive long-term data sets.

Methods

We evaluated RGWT production across the Edwards Plateau, Rolling Plains, Cross Timbers and Prairies, Post Oak Savannah, and South Texas Plains ecological regions. These regions encompassed the majority of RGWT range in Texas (Fig. 1).

Personnel from TPWD collected RGWT brood observations from 1 June through 15 August, 1976–2000. Observers recorded all RGWT hens and poults dur-

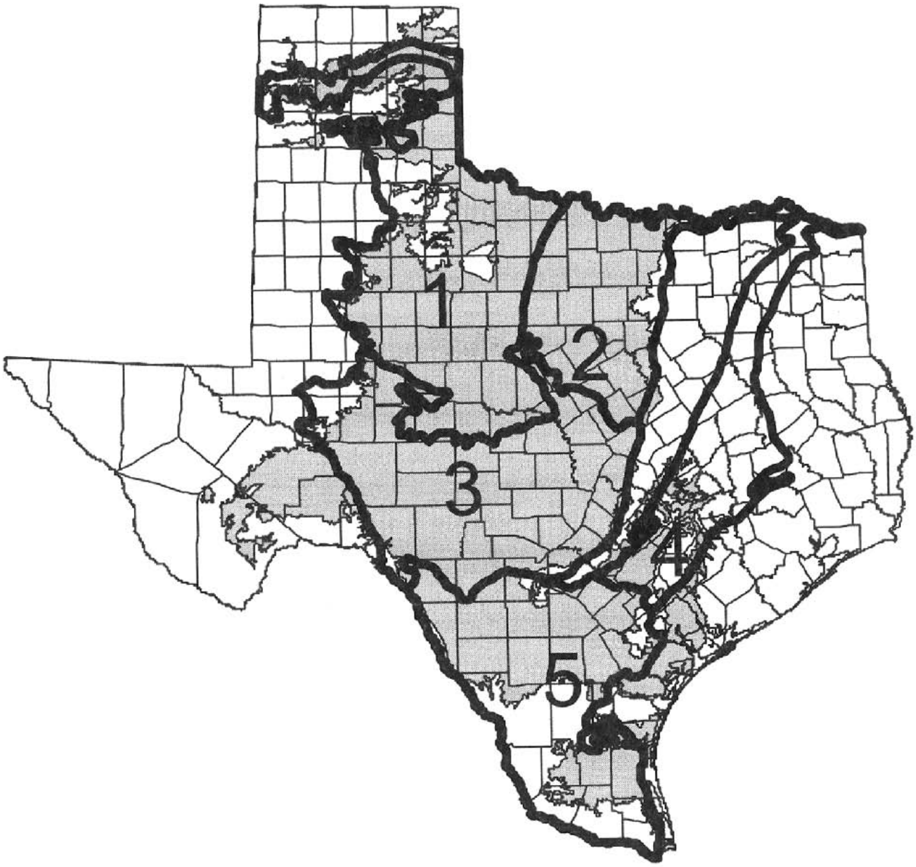

Figure 1. Ecological regions (Gould 1975) of Texas containing significant populations of Rio Grande wild turkey. Names of ecological regions are 1 = Rolling Plains, 2 = Cross Timbers and Prairies, 3 = Edwards Plateau, 4 = Post Oak Savannah, and 5 = South Texas Plains. Shaded area indicates approximate range of the Rio Grande wild turkey in Texas, adapted from Texas Parks and Wildlife Department (1997).

ing the course of routine daily activities. Counts were not conducted along standardized routes; rather, observers were encouraged to observe 10–25 hens per county during each 2-week period. Observations were recorded by county and latitude-longitude coordinates (Graham and George 2002).

Data Analysis

Brood-count Data.—We grouped each year's data according to ecological region prior to analysis. Data from the Edwards Plateau and Cross Timbers and Prairies were available for 1976–2000, data from the Rolling Plains and Post Oak Savannah were available for 1977–2000, and data from the South Texas Plains were available for 1977–1978 and 1980–2000.

We calculated total number of hens and poults observed per year in each ecological region. We then calculated RGWT poult production *(p)* per region as

$$p = \frac{N_p}{(N_p + N_h)}$$

where, N_p = number poults and N_p = number of hens. We also determined the total number of RGWT groups containing at least one poult or hen observed annually in each ecological region.

Power Analysis.—Steidle et al. (1997) advised that power analysis should be performed using biologically meaningful effect size. However, Gerard et al. (1998) noted that biologists often are reluctant to define what effect size is biologically meaningful, because it is a subjective decision, often with little data to support it. Published research addressing the sensitivity of turkey populations to changes in recruitment are sparse. Vangilder and Kurzejeski (1995) performed sensitivity analysis using a population model of eastern wild turkeys in northern Missouri to examine the effects of varying nest success and poult mortality, which are both important determinants of recruitment. They found that increasing annual nest success 10% and 20% increased the hypothetical population after 40 years by 937% and 12,696%, respectively; decreasing nest success 10% and 20% resulted in 13% and 88% declines in the population. Changes in poult mortality produced similar results. Increasing poult mortality 10% and 20% resulted in a population decrease of 68% and 98%, while decreasing poult mortality by 10% and 20% resulted in a population increase of 3,154% and 19,957%, respectively. These results suggested that changes in recruitment of 10%–20% where biologically meaningful; however, differences in climatic and habitat conditions between northern Missouri and Texas may lessen the applicability of the results to turkeys in Texas. Therefore, we chose to perform our analysis using a wide range of effect sizes. We estimated power of the brood counts to detect a change in poult production between consecutive years using the 1-proportion power calculation function in Minitab for Windows 12.2 (Minitab, Inc., State College, Pennsylvania). We calculated power to detect inter-annual differences in poult production (i.e., $p_1 – p_0 = \delta_p$), where $\delta_p = 0.05, 0.075, 0.10, 0.15, 0.20$ and $p_0 = 0.50$, for a range of sample sizes (25–500) representative of actual sampling effort. We set $p_0 = 0.50$ because power is lowest and required N is highest for this value, thus corresponding estimates are most conservative (Ott and Longnecker 2001:474). We set $\alpha = 0.05$ for all calculations.

We also estimated the power of the survey to detect long-term changes in poult proportion within each ecological region. We assumed that changes in production over time could be tested for by dividing the time series into two periods (labeled arbitrarily as periods No. 1 and 2) at the approximate mid-point of the time series and comparing the means of the period No. 1 and 2 using a Student's *t*-test. Sample size

Table 1. Pooled sample standard deviation and sample size (in years), by ecological region, used in power analysis of long-term recruitment trends.

Ecological region	S_{pooled}	n_1	n_2
South Texas Plains	0.2238	11	12
Low Rolling Plains	0.1848	12	12
Edwards Plateau	0.2192	13	12
Cross Timbers and Prairies	0.1287	13	12
Post Oak Savannah	0.1796	12	12

equaled length of each period in years. Hence, power of the test was determined using the two-sample *t*-test power analysis in Minitab. Because the results of Levene's test indicated sample standard deviation did not differ within ecological region between the two periods ($P = 0.492–0.910$), pooled sample standard deviation (S_{pooled}) was calculated per region (Table 1) and used as an estimate of population standard deviation in power calculations. We calculated power to detect difference in mean poult production between the two consecutive long-term data sets (i.e., $\mu_1 – \mu_2 = \delta_\mu$), where $\delta_\mu = 0.05–0.40$ in 0.05 increments.

We also determined the minimum number of years that the count would have to be conducted to detect a mean difference in poult production (δ_μ) between the periods No. 1 and 2 for each region, where $\delta_\mu = 0.05–0.40$ in 0.05 increments, $p_0 = 0.50$. and $1–\beta = 0.80$. For these analyses, we assumed that brood counts accurately estimated the mean poult proportion for each ecoregion.

Results

Power analysis indicated 50 turkey brood observations per year were required for ≥80% chance of detecting $\delta_p = 0.200$. For the same probability of detection, required group size increased to 100 for $\delta_p = 0.150$, 200 for $\delta_p = 0.100$, 350 for $\delta_p = 0.075$, and >500 for $\delta_p = 0.050$ (Fig. 2).

Power analysis indicated the current data set had power ≥0.80 to detect ≥0.30 difference in poult production between the two consecutive time series in all regions. Only the Cross Timbers and Prairies data had similar power to detect a difference of 0.20. No region's data had power ≥0.80 to detect a difference of ≤0.15 (Fig. 3)

Time-series data sets of 16–30 years had power ≥0.80 to detect long-term mean differences in poult production of 0.20 in the Rolling Plains, Cross Timbers and Prairies, and Post Oak Savannah regions. Counts of ≥40 years would be required for similar results in the Edwards Plateau and South Texas Plains (Table 2).

Discussion

Vangilder and Kurzejeski (1995) suggested that a 10%–20% change in turkey recruitment was biologically meaningful. This corresponds to $\delta_p = 0.050–0.10$ when $p_0 = 0.50$. Our results suggest that a sample size of $N = 200–>500$ turkey-group ob-

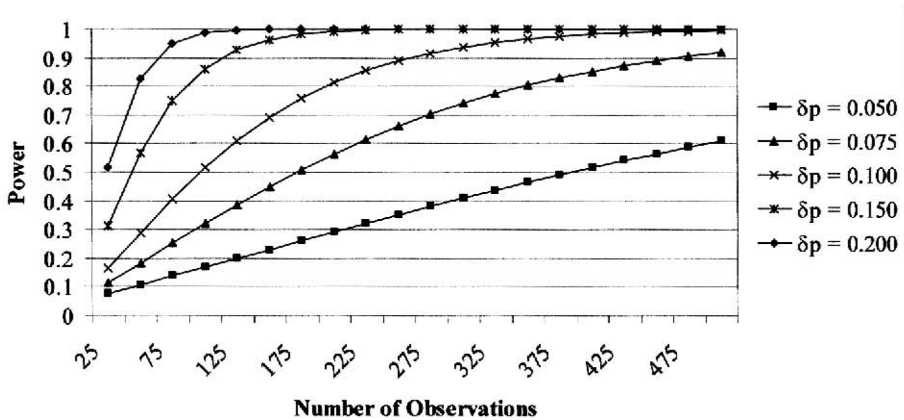

Figure 2. Power of TPWD brood surveys to detect inter-annual change of δ_p away from hypothetical proportion of 0.50 poults in the hen:poult population.

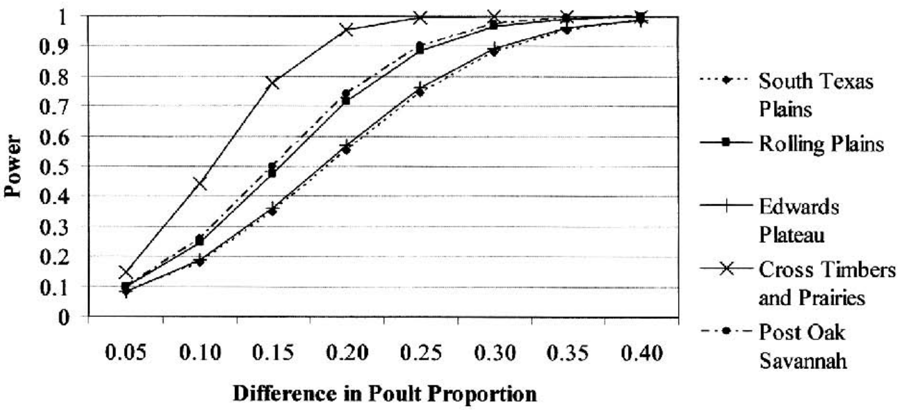

Figure 3. Power of current TPWD brood-count data sets to detect a given difference in mean poult production between two consecutive long-term data sets.

servations were needed to detect this level of inter-annual difference in poult production when power ≥0.80. Sample size averaged 65–306 for the five regions. Number of observations likely differed among regions due to sampling effort and turkey density.

Our results indicated that existing production data had very low power (<0.50) to detect a long-term change of <0.20%. Further, time series of 54–160 years would

Table 2. Minimum length (in years) of time series required to detect a long-term change (δ_μ) of 0.05–0.40 in poult proportion with power ≥0.80, for five ecological regions of Texas.

δ_p	South Texas Plains	Low Rolling Plains	Edwards Plateau	Cross Timbers and Prairies	Post Oak Savannah
0.05	632	432	606	210	408
0.10	160	110	154	54	104
0.15	72	50	70	26	48
0.20	42	30	40	16	28
0.25	28	20	28	12	20
0.30	20	16	20	10	14
0.35	16	12	16	8	12
0.40	14	10	12	6	10

be required to detect this effect size in all ecological regions. This low power resulted from the high degree of inter-annual variation in poult production.

A further complication is that collection of incidental brood count data is "haphazard" or "convenience" sampling, not a true survey. Samples are not random; therefore, samples may not be representative of the population. This may have biased estimates of turkey production.

Our evaluation of TPWD brood count data was based on the assumption that a 10%–20% change in recruitment is biologically meaningful to RGWT population dynamics in Texas, as it was for eastern wild turkey in Missouri (Vangilder and Kurzejeski 1995). There is some evidence to suggest that Texas populations may behave differently than those in Missouri. Annual turkey survival on four study sites in the Edwards Plateau was 0.566–0.737 (Beau Willesey, unpublished data), versus 0.445–0.693 used in Vangilder and Kurzejeski's model. Higher annual survival rates may lessen the sensitivity of turkey populations to changes in recruitment.

Rio Grande wild turkey brood counts, as currently conducted by the Texas Parks and Wildlife Department, have little value for detecting either inter-annual or long-term changes in turkey recruitment. Further, haphazard sampling may bias recruitment estimates. We suggest that TPWD reevaluate the way RGWT populations are monitored to address the issues we have raised. At a more fundamental level, we encourage research into RGWT population dynamics, in order to more adequately define the role of recruitment in regulating populations and determine the biologically meaningful effect size that surveys should be designed to detect.

Wild turkey management and the setting of harvest regulations require reliable information regarding turkey population dynamics, including recruitment. Power analysis is a powerful tool for designing and evaluating population monitoring efforts. Without a clear understanding of the statistical power, managers may falsely conclude that populations are stable when, in fact, changes are occurring. We encourage the use of power analysis in population monitoring efforts to strengthen the rigor and reliability of knowledge upon which management decisions are based.

Acknowledgments

We thank the numerous TPWD personnel who collected brood-count data during 1976–2000. We also thank M. Frisbie for providing access to TPWD brood-count database. We thank M. Gray, J. Taylor, and an anonymous reviewer for their helpful comments that greatly improved the quality of this manuscript. Data collection and analysis were supported by Federal Aid in Wildlife Restoration through TPWD.

Literature Cited

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29: 1294–1297.

Bartush, W. S., M. S. Sasser, and D. L. Francis. 1985. A standardized turkey brood survey method for northwest Florida. Proceedings of the National Wild Turkey Symposium 5: 173–181.

Crouch, W. B. III, and P. W. C. Paton. 2002. Assessing the use of call surveys to monitor breeding Anurans in Rhode Island. Journal of Herpetology 36:185–192.

Gerard, P. D., D. R. Smith, and G. Weerakkody. 1998. Limits of retrospective power analysis. Journal of Wildlife Management 62: 801–807.

Gibbs, J. P. and S. M. Melvin. 1997. Power to detect trends in waterbird abundance with call-response surveys. Journal of Wildlife Management 61:1262–1267.

Gould, F. W. 1975. Texas plants: a checklist and ecological summary. Texas A&M University, Agricultural Experiment Station, College Station.

Graham, G. L. and R. R. George. 2002. Operational Plan: Fiscal Year 2003. Texas Parks Wildlife Department, Austin.

Kurzejeski, E. W. and L. D. Vangilder. 1992. Population management. Pages 165–184 *in* J. D. Dickson, editor. The wild turkey: biology and management. Stackpole Books, Mechanicsburg, Pennsylvania.

Lougheed, L. W., A. Breault, and D. B. Lank. 1999. Estimating statistical power to evaluate ongoing waterfowl population monitoring. Journal of Wildlife Management 63: 1359–1369.

Menzel, K. E. 1975. Population and harvest data for Merriam's turkeys in Nebraska. Proceedings of the National Wild Turkey Symposium 3:184–188.

Morrison, M. L., W. M. Block, M. D. Strickland, and W. L. Kendall. 2001. Wildlife study design. Springer-Verlag, New York, New York.

Ott, R. L. and M. Longnecker. 2001. An introduction to statistical methods and data analysis, 5th edition. Duxbury, Pacific Grove, California.

Rice, C. G. 2003. Utility of pheasant call counts and brood counts for monitoring population density and predicting harvest. Western North American Naturalist 63:178–188.

Roberts, S. D. and W. F. Porter. 1996. Importance of demographic parameters to annual changes in wild turkey abundance. Proceedings of the National Wild Turkey Symposium 7:15–20.

Schultz, V. and R. D. McDowell. 1957. Some comments on a wild turkey brood study. Journal of Wildlife Management 21: 85–89.

Shaw, H. G. 1973. The roadside survey for Merriam's turkey in Arizona. Pages 285–293 *in* G. C. Sanderson and H. C. Schultz, editors. Wild turkey management: current problems and programs. University of Missouri, Columbia.

Steidl, R. J., J. P. Hayes, and E. Schauber. 1997. Statistical power analysis in wildlife research. Journal of Wildlife Management 61:270–279.

Texas Parks and Wildlife Department. 1997. Wild turkey distribution in Texas. Texas Parks and Wildlife Department Leaflet PWD LF W7000-038 (11/97), Austin.

Vangilder, L. D. and E. W. Kurzejeski. 1995. Population ecology of the eastern wild turkey in northern Missouri. Wildlife Monographs 130:1–50.

Wildlife Society, The. 1995. Journal news. Journal of Wildlife Management 59:196–198.

Wunz, G. A. and W. K. Shope. 1980. Turkey brood survey in Pennsylvania as it relates to harvest. Proceedings of the National Wild Turkey Symposium 4: 69–75.

Zar, J. H. 1999. Biostatistical Analysis, 4th edition. Prentice-Hall, Inc., Upper Saddle River, New Jersey.