Final Report for J8W07060004

Development of sampling designs for long-term monitoring protocols for the Klamath, Mojave, San Francisco Bay Area, Sierra Nevada, Upper Columbia Basin, Rocky Mountain, and Pacific Islands I & M Networks

Kirk Steinhorst
University of Idaho
May 16, 2011

This task agreement ended April 1, 2011. The purpose of the agreement was to provide statistical support for the development of long term monitoring protocols in the NPS I&M networks listed above. For each protocol there are statistical components to both status and trend analysis. During the five years that this TA was in effect, statistical support was supplied by Kirk Steinhorst, Leigh Ann Starcevich, Kathi Irvine, Steve Hayes, and at the beginning, Andrew Merton.

Our statisticians participated in protocol development at several levels.

A. In some cases we worked extensively as a part of the protocol development team helping with development of objectives, identification of population(s)/ sampling unit, development of sampling design, status and trend statistical analysis, evaluation of pilot data, and power analysis. Examples of such protocols include UCBN camas, SIEN lakes, KLMN landbirds, and UCBN sagebrush-steppe.
B. In other cases much of the groundwork had been laid and we spent most of our time on status/trend analysis and power analysis. Examples of these projects are SFAN coho, UCBN aspen, SFAN raptors, and UCBN water quality.
C. At the lower end, there were projects where the statistical input was less. For example, network staff may write a statistical section for something like osprey or pika and call one of us for confirmation. Leigh Ann had some smaller involvement in SIEN grazing and stock-use monitoring. The statistical analyses recommended for snowy plovers (SFAN) only took a few days work.

We also worked on “general” statistical tasks. We spent a lot of time developing statistical tools and techniques useful for long term monitoring. Statistical tools useful for assessing status of a vital sign fall under the survey sampling body of statistical literature. In some cases simple or stratified random sampling was recommended. Often though, we worked with network scientists to develop spatially balanced designs (GRTS). Sampling units were quadrats, transects, stream reaches, road and trail corridors, nests, etc. depending on the resource. Temporal design used the panel design notation of McDonald (2003).

For detecting trend we needed statistical tests that were appropriate and robust. We also needed tests where power analyses could help determine sample sizes that could detect biologically interesting trends over time. The most useful tool for this purpose was the mixed linear model (Piepho and Ogutu 2002),

$$y_{ijk} = \mu + \beta w_j + a_i + b_j + w_i t_j + c_{ij} + e_{ijk}$$

where $\mu + \beta w_j$ is the fixed linear trend over time, a_i is the random effect of the i-th site, b_j is the random effect of the j-th year, t_i is the random slope associated with the i-th site, c_{ij} is the random interaction effect between site and year, and e_{ijk} is random error. Errors were often assumed to be normal after appropriate transformation of the response variable.

We used two approaches for determining power for this mixed model—asymptotic methods (Urquhart et al., 1993) and Monte Carlo simulation (Sims et al., 2006). Our simulations have shown that large-sample theory overapproximates power as compared to the Monte Carlo simulation approach, but large-sample theory retains the appropriate relative power for comparing revisit designs (Starcevich, manuscript in preparation).

Not all protocols could use the mixed model approach. One of the first protocols we worked on was the camas protocol (UCBN) which uses permutation tests for detecting trend. Power analysis of this model used Monte Carlo simulation. In other protocols we applied occupancy models which were often couched in terms of zero-inflated count models (SFAN raptors, MEDN herps). The sagebrush-steppe protocol (UCBN) used proportional odds models of ordered categorical data. Both wetland data (ROMN) and vegetation data (PACN) used zero-inflated continuous density models for detecting trend. Monte Carlo simulation was used for power analysis. KLMN scientists asked us to develop multivariate methods for analyzing community level data (see Irvine, Dinger, and Sarr citation below).

The protocols that we were most heavily involved with are:

1. KLMN
    a. Landbird monitoring protocol
    b. Whitebark pine
    c. Water quality and aquatic communities-streams
    d. Caves
    e. Macroinvertebrates—wadeable streams
    f. Multivariate community analysis
    g. Invasive plants
2. MEDN
    a. Aquatic herps
    b. Vegetation
3. MOJN
    a. Water quality and macroinvertebrates
    b. Springs
    c. Integrated upland

    d. Streams and lakes
4. PACN
    a. Focal terrestrial plant communities
    b. Early detection of invasive plant species
    c. Marine fishes and benthos
    d. Anchialine pools
5. SFAN
    a. Coho
    b. Raptors and spotted owl
    c. Snowy plovers
6. SIEN
    a. Lake chemistry
    b. Rivers and Streams
    c. Wetlands
    d. Whitebark pine
7. UCBN
    a. Camas lily
    b. Sagebrush-steppe
    c. Water quality and benthos, integrated riparian
    d. Aspen
    e. Limber Pine

In addition this task agreement included work on wetlands for the Rocky Mountain network, on giant sequoia in the SIEN parks, and on the effects of visitor disturbances on the Merced River in Yosemite National Park.

We put on a 3 day statistics workshop for I&M scientists at the San Francisco Area Network in December of 2010.

This task agreement also produced refereed papers and talks at professional meetings:

1. Irvine, K.M. and T.J. Rodhouse (2009) Modeling Trends in Vegetation with Ordinal Cover Classes: Implications for Long-Term Monitoring Designs. TIES 2009. Corvallis, Oregon.
2. Starcevich, L.A.H., K.M. Irvine, and A.M. Heard (2009). Implications of Model Specification and Temporal Revisit Designs on Trend Detection. TIES 2009. Corvallis, Oregon.
3. Irvine, K.M., T.J. Rodhouse, and K.T. Vierling, Incorporating Spatial Correlation within a Bayesian Hierarchical Trend Model for Camas Lily Abundance, WNAR International Biometric Society 2010
4. Starcevich, L.A.H., K.M. Irvine, and A.M. Heard (2010). Implications of Model Specification and Temporal Revisit Designs on Trend Detection. Presentation at Oregon State University Statistics Department seminar. Corvallis, Oregon.

5. Irvine, K.M., E. Dinger, and D. Sarr (2010) The Complexities of a Multivariate Power Analysis: Possible Finite Solutions to an Infinite Problem. Joint Statistical Meetings 2010
6. Irvine, K.M. and T.J. Rodhouse (2010). Power Analysis for Trend in Ordinal Cover Classes: implications for long-term monitoring. Journal of Vegetation Science, 21, pp.1152-1161.
7. Starcevich, L.A.H Temporal Revisit Designs on Trend Detection. Manuscript submitted.

REFERENCES

McDonald, T.L. (2003). Review of environmental monitoring methods: survey designs. *Environmental Monitoring and Assessment*, 85, 277-292.

Piepho, H. and J. Ogutu (2002). A simple mixed model for trend analysis in wildlife populations. *Journal of Agricultural, Biological, and Environmental Statistic*s, 7, pp. 350-360.

Sims, M., S. Wanless, M.P. Harris, P.I. Mitchell, and D.A. Elston (2006). Evaluating the power of monitoring plot designs for detecting long-term trends in the numbers of common guillemots. *Journal of Applied Ecology* 43: 537-546.

Urquhart, N.S., Overton, W.S., & Birkes, D.S. (1993). Comparing sampling designs for monitoring ecological status and trends: impact of temporal patterns. *Statistics for the Environment*, V. Barnett and K.F. Turkman, eds. John Wiley and Sons Ltd.