THE WILDLIFE SOCIETY
5410 Grosvenor Lane • Bethesda, MD 20814-2144
Tel: (301) 897-9770 • Fax: (301) 530-2471
E-mail: tws@wildlife.org

Review of the Study on Golden-cheeked Warbler Population Distribution and Abundance and Associated Manuscripts

PEER REVIEW

18 July 2011

The Wildlife Society (TWS) appreciates the opportunity to facilitate the review of and comment on four (4) associated manuscripts (Collier et al. 2010, Laake et al. 2011, Collier et al. *in press*, and Mathewson et al. *in review*) and the overall findings of the document entitled *Golden-cheeked Warbler Population Distribution and Abundance* prepared by Texas Agrilife Research associated with Texas A&M University (TAMU) to provide a modeling tool and abundance and distribution estimates of the golden-cheeked warbler across the species' breeding range. The following comments are the product of a peer review team organized in response to TAMU's request for a scientific and technical review. The Wildlife Society (TWS) was founded in 1937 and is a non-profit scientific and educational association of over 10,000 professional wildlife biologists and managers, dedicated to excellence in wildlife stewardship through science and education.

The Wildlife Society assembled a team of reviewers to respond to TAMU's Statement of Work issued 1 April 2011. The team included experts in Golden-cheeked Warbler (GCWAs) ecology and biology, spatially-explicit habitat modeling as an indicator of abundance and distribution of songbirds, field testing and validation of predictive modeling, statistical analysis, GIS assessments, breeding survey techniques for neotropical migratory songbirds, and other appropriate biological training and experience.

Although The Wildlife Society facilitated the following review, the comments summarized herein are those of the independent reviewers; this is not an official statement of The Wildlife Society and does not represent the view of The Wildlife Society or its membership.

INTRODUCTION

At the request of TAMU and the U.S. Fish and Wildlife Service (the Service), TWS facilitated a two month, independent peer review of a package of materials relating to a recent modeling study for predicting occupancy and abundance within the GCWAs summer breeding range in Texas. In addition to the independent peer review, a brief question and answer/interview conference call occurred on 28 June 2011 (facilitated by TWS), enabling reviewers to seek clarification and a greater understanding of the research and interpretations associated with these studies. This review takes into account not only the documents themselves, but this question and answer session.

Peer Review Team

TWS' peer review team consisted of 5 independent and well qualified professionals with the collective expertise previously identified. The comments are those of the reviewers and do not necessarily represent the views of their employers or affiliated organizations. The review team consisted of the following individuals:

- *Jeffrey R. Dunk*, Dept. of Environmental Sciences & Mngmt, Humboldt State University
- *Paul B. Hamel*, U.S. Forest Service
- *Rebecca Peak*, U.S. Army Garrison – Fort Hood
- *Wayne Thogmartin*, U.S. Geological Survey
- *Frank R. Thompson III*, USDA Forest Service Northern Research Station

The Review Materials and Process

The following is a summary of general and specific review comments on all five of the requested manuscripts:

1. Collier, B. A., M. L. Morrison, S. L. Farrell, A. J. Campomizzi, J. A. Butcher, K. B. Hays, D. I. MacKenzie, and R. N. Wilkins. 2010. Monitoring endangered species occupying private lands: case study using the golden-cheeked warbler. Journal of Wildlife Management 74:140-147.

2. Laake, J. L., B. A. Collier, M. L. Morrison, and R. N. Wilkins. 2011. Point-based mark-recapture distance sampling. Journal of Agricultural, Biological, and Environmental Statistics. DOI:10.1007/s13253-011-0059-5.

3. Collier, B. A., K. E. Groce, M. L. Morrison, C. Newnam, A. J. Campomizzi, S. L. Farrell, H. A. Mathewson, R. T. Snelgrove, R. J. Carroll, and R. N. Wilkins. Predicting patch occupancy in fragmented landscapes at the rangewide scale for endangered species: an example of an American warbler. Diversity and Distributions, *In Press*.

4. Mathewson, H. A., J. E. Groce, T. M. McFarland, M. L. Morrison, J. C. Newnam, R. T. Snelgrove, B. A. Collier, and R. N. Wilkins. Estimating breeding season abundance of golden-cheeked warblers in Texas. Journal of Wildlife Management, *In Review*.

5. Morrison, M. L., R. N. Wilkins, B. A. Collier, J. E. Groce, H. A. Mathewson, T. M. McFarland, A. G. Snelgrove, R. T. Snelgrove, and K. L. Skow. 2010. Golden-cheeked warbler population distribution and abundance. Texas A&M Institute of Renewable Natural Resources, College Station, Texas, USA.

Due to the complexity of the statistical analyses and reviews performed by each reviewer, it is difficult to adequately and effectively summarize the degree of detail captured within each review. TWS recommends that in addition to using this summary as a basis for understanding,

each review be read and considered for its unique level of assessment regarding these manuscripts.

Based on the request from TAMU and the Service, there is significant emphasis on whether the best available science was used and interpreted in a reasonable way and whether or not the implications of this report will be useful and reliable for developing scientific approaches for management and recovery of GCWAs.

SUMMARY OF THE INDEPENDENT PEER REVIEWS

Responses to Review Questions

In conjunction with TAMU and the Service, TWS developed 9 questions for reviewers to address throughout the peer review process. The following is a summary of key points highlighted in those responses. Additional details are available within each independent review.

1. **Are the model (Morrison et al. 2010) and its implications supported by sound scientific data?**

   The overall study design, analyses, and inferences are supported by sound scientific data and analysis. However, the necessary information to judge this is really not provided in the documents – more information on training and quality control could be provided. In addition, when applied more locally the accuracy of the data and model is assumed to breakdown without more samples used to refine the values of the coefficients.

   Also, as observed by one reviewer, the conclusions of Morrison et al. 2010 are potentially harmed by several possible flaws. (1) No consideration was given to the underlying model uncertainty associated with how habitat is characterized – choosing a model that over-predicts suitable habitat resulting in extrapolated densities based upon an imprecise relationship between density and occupancy. (2) "The characterization and inferences from the model of occupancy are seemingly flawed to an unknown extent…Their model parameter estimates are not credibly different from zero, suggesting that the characterization of the spatial structure in occupancy is largely a result of their spline rather than effects of habitat. Thus they have no mechanism to restrict the extrapolation to forest patches because they don't have credible habitat characterizations associated with occupancy."

2. **Does the model(s) provide reliable conclusions and are the assumptions of the model appropriate given the scale and the scope of the project and its purpose?**

   Although the models were deemed appropriate given the scale and scope of the project, a fairly consistent response was generated by reviewers that, given the concerns voiced above, the reliability of the model(s) remains uncertain. Assumptions were identified and specified, and although unavoidable, were not tested sufficiently. Admitted over-estimation of habitat, imprecision and poor explanatory power in the extrapolation algorithm precludes a reliable inference. In addition, concern was raised regarding the modest sample size for the scale of the study and the acknowledged deviations from random sampling.

It is recommended that the authors better justify some assumptions and adjust the model to address these issues more transparently.

3. **Are the methods used to both develop and validate the model scientifically sound?**

   As stated by one reviewer, "this is exemplary work." The methods used for developing and validating the various models are scientifically sound, but a more rigorous cross-validation of the density model has been identified as a need. Concern is noted regarding the statistical inferences drawn from the models in question. As noted previously, there is a need to better address support for some assumption and potentially necessary to include additional factors in the models.

4. **Were the data collection and statistical procedures appropriate given the scale and scope of the project?**

   Yes – the methods for collecting and analyzing occupancy data were seemingly sound. However, it was noted that the results of those models were insufficient for delivering the inferences the authors desired. It was noted that the authors tried to balance the tradeoff between "spatial extent of an analysis and spatial resolution…by controlling for within patch heterogeneity by assuming the sampling design would address these factors; essentially fitting models to random samples they could make predictions for all patches." Due to the complexity of the problem/scale it is noted that it would be difficult to properly represent all of the factors discussed.

5. **Were the statistical procedures being used misinterpreted or do they have invalid assumptions?**

   The majority of reviewers felt that the statistical procedures were not misinterpreted. In addition, one reviewer noted that the "uncertainty around some estimates and the assessment of model fit and predictive ability need[ed] to be more thorough."

   However, one reviewer felt that, if not misinterpreted, the statistical procedures described by Collier et al. (*in press)* were over interpreted. Although the authors acknowledged that their methods were "amenable to information-theoretic approaches to model selection," they "failed to compare their completed model to less complex forms."

6. **Do the authors (Collier et al. *in review* and Mathewson et al. *in review*) appropriately interpret the results of the document (Morrison et al. 2010) in making their findings?**

   The general response from reviewers was that this question was not clear due to the fact that neither of the *n review/press* manuscripts spend much time interpreting results from Morrison et al. (2010). That said, it was noted that there is consistency in the results among these works and that these documents are simply "outgrowths" of the process that produced the Morrison et al. report. Although the findings are in line with Morrison et al. it is noted that this does not necessarily mean that they are valid for the previously stated reasons.

7. Are the implications of this report useful and reliable for developing scientific approaches for management and recovery of golden-cheeked warblers?

This question generated a mix response from the reviewers airing on the side of usefulness for developing scientific approaches for management and recovery. The primary implication noted is that there may be more GCWAs on the landscape than previously thought. The process implies how to improve the accuracy and applicability of the product, demonstrating the usefulness of the process.

It was noted by numerous reviewers that, even if these implications represent the best current approach for developing management and recovery actions, they do not represent the best possible information.

It is recommended that revisions to the approach and further work is pursued to "improve the reliability and demonstrate the degree of reliability of this (more GCWAs) conclusion."

8. Are the findings of this report reliable to the point the USFWS may effectively use them as best available science for recovery planning for the golden-cheeked warbler?

Consistently, reviewers felt that this could be used as "the best available science" for understanding this species. However, the question arose as to whether it is good enough science to drive recovery planning, particularly as smaller geographic scales than range wide. It was also noted that this should not be viewed as "the best available science"' at the exclusion of other available science.

If the revisions suggested in this peer review are pursued, one reviewer felt that this could result in the most rigorous range wide population estimate and the "best available science" for some aspects of recovery planning.

9. How could the model and subsequent documents be improved?

The following list is a comprehensive set of 13 recommendations/suggestions from all reviewers. In an effort to ensure clarity with these recommendations, all have been given verbatim from the individual reviews.

1) Because the modeling process appears to be rigorous and appropriately conducted and the coefficients of the various parameters in the model are in the main not different from zero, a process of continuing to incorporate more field observations into the validation of the model is warranted.

2) The delivery of the model is still a black box for any user interested in calculating land use or conservation implications of the apparent predictive ability of the model. Additional expression of this product in a form that is user friendly and tailored to the potential audience of land managers and local/county development and other government entities is imperative.

3) Activity that will be very useful is the explicit testing of the assumptions identified in the model building process, especially gathering sufficient data to incorporate patch habitat contents into conservation planning for Golden-cheeked Warbler.

4) The inclusion of reproduction and/or survival information (spatially-explicit) would be a large improvement.

5) Provide a series of "what if" scenarios of the use of their models. For example, they could ask "what if" all conservation efforts were restricted to patches >500 ha? What would the estimated abundance of GCWAs be? What percentage of the total estimated number of GCWAs would that represent? How many of the Recovery Regions would have more than some minimum threshold number of GCWAs in them (say 500)? Or, they could ask "What if the goal of the USFWS was to have a minimum of 10,000 singing male GCWAs in each of the modeling regions, what is the smallest estimated amount of area that could be conserved to reach this goal (by modeling region and overall)?"

6) Include an outline (perhaps a figure) in which an example of the use and refinement of the models developed is provided. For example, to what degree do the authors think that additional field sampling should be conducted on GCWA occupancy/abundance/density? What refinements to the model are likely to be made (mechanistic spatial covariates as opposed to knots; targeting sampling for those areas where estimates were highly variable, etc.)? How might the USFWS utilize the current models, make conservation planning decisions, gather new data, and incorporate new information into the model (refine it) over time?

7) Due to the residual uncertainty relating to how to best characterize GCWA habitat, interested parties should invest in a multi-model perspective, akin to what is done with, say, climate change scenarios.

8) To extrapolate density across occupied patches, it may be used to think in terms of quantile regression (see Cade and Noon 2003, Koenker 2005, Hao and Naiman 2007) to properly characterize the yawning patterns in the density-occupancy relationship. This yawning pattern in density-occupancy may either be a result of an inappropriate model form (supposedly addressed with a negative binomial regression) or lurking covariates.

9) Information-theoretic approaches seem warranted for evaluating parsimonious collections of occupancy models.

10) More rigorously address support or non-support for assumptions so we have more faith that an assumption is in fact supported, or if not that the results are simply conditional on it and describe possible effects on the results.

11) Reconsider the detectability issue in the abundance models.

12) Don't rely on the design to control some key factors affecting abundance (i.e. veg type, ownership, edge, year), address them in the abundance model.

13) More thoroughly address model fit and uncertainty in the population estimates.

General Comments Regarding the Review Package

Comprehensively all reviewers recognized and commended the authors for the quality and breadth of the research and the magnitude of the task of developing a model for occupancy that can be applied on a range-wide scale. The observation was made that tackling a range-wide abundance estimate that considers public and private lands, and that uses a statistical-based sampling design, was a large undertaking and the authors accomplished much in this regard.

Reviewers noted that this work both in its current form and with key revisions could be very applicable to the Service's efforts for GCWA recovery. One reviewer noted that the overall sampling design and analyses could likely serve as an excellent example for other large-scale single species conservation planning exercises and could eventually lead, based on the chronology of research, to the development of spatially-explicit estimates of reproduction and survival for GCWAs.

As noted by another reviewer, "given that the process involved in the papers reviewed herein produced an estimate of the species population 10 times greater than that accepted in the standard reference for North American songbird populations, the "Flight Plan" from Partners in Flight (Rich et al. 2004), this review is a useful one for those concerned with future management and listing or delisting decisions relative to GCWAs."

Although the overall effort of the authors and their teams was commended, key recommendations and concerns were identified by all reviewers for both the primary model as a whole and the subsequent interpretations and additional modeling inferred from the Morrison et al. study. Key concerns were raised in reference to patch level attributes. It was noted by numerous reviewers that the variation in terminology used to describe woodland type needs to be addressed. At various times throughout many of the papers the habitat classification selected for patches was classified as "woodland" or "oak-juniper woodland." The use of the different terminology was confusing to the reviewers and the use of "oak-juniper woodland" implies a much narrower habitat classification.

An additional concern noted regarding patch selection and survey was the relatively small number of patches surveyed (301) versus potential patches (63,616) identified through the habitat classification. Although 301 sample sites can be viewed as a sufficient number, in relation to the total potential patches, this number may not be representative due to sampling error. As noted by one reviewer "a sample of 301 patches by most standards would seem large, but as a representation of 63,616 patches it is modest. So without further evidence that the authors sample of patches is representative of all potential habitat, the reliability of the estimate is unknown. As a result another reviewer noted that "application of the model to more restricted areas [should only be done] with the utmost caution because no local region can be assumed to be a random representative of the entire range."

Along those lines, concern was raised regarding the inability for researches to access patches that were inaccessible through certain private landowners. An assumption was made by the authors that those patches where access was not granted were not significantly different in habitat and probability for occupancy than those sites in which access was granted. It has been recommended that some additional analyses should be conducted to support the authors'

assumption that there are not any differences between the areas they were able to access and the areas that they were not able to access that would bias occupancy or abundance estimates.

It has been emphasized throughout the comments that the real need within these studies is to "address additional patch level attributes in the models instead of relying on sampling design to address these" in order to ensure a representative sample of GCWA potential habitat. As one reviewer phrased it, "…the issue falls most squarely on how to properly characterize habitat rather than how birds are surveyed or populations are estimated." A recommendation has been made to "reanalyze the data with woodland type as a covariate instead of pooling all woodland types." If followed, this recommendation may lead to providing tangible information to natural resource professionals "regarding how much variation in the abundance estimate is explained by woodland type and thus, how to interpret these results."

Additional concerns were noted regarding the selection of Model III (Morrison et al.). Although this model seemed to do a good job describing available habitat, it suffers from "relatively high errors of commission" and over-predicts abundance of habitat.

Finally, due to the large discrepancy between the projected population numbers produced by the model and what is accepted in the standard reference, two key questions were raised: (1) How does the estimate of the population size fit with the authors' gut feelings on the front end of the study? and (2) Does a population of this size continue to warrant an ESA listing of Endangered?

Manuscript Specific Comments

Manuscript specific comments and reviews have a level of specificity, analysis and detail that is difficult to summarize across all reviewers. This summary reflects more of the high-level areas for support and/or concern. TWS recommends an extensive review of each reviewers comments to fully comprehend the detailed reasoning behind these higher level observations. Where applicable and effective, specific reviews have been noted for reference.

Collier et al. 2010

In general, reviewers felt that this article was well constructed and that the reasoning in the paper appeared to be sound. It was stated that the authors did demonstrate that occupancy modeling can be used to assess patch occupancy at broad scales for GCWAs and that the management implications stated by the authors is sound.

Reviewers noted concerns regarding the approach in which several paired surveys are considered to be independent due to the potential for surveys to be pseudoreplicates (p. 143, Fig. 3; Hamel). Concern was also expressed regarding the use of the centroid of Ft. Hood as the basis for determining distances between patches due to the potential for obscuring edge-to-edge distance relationships. One reviewer noted concerns regarding sampling covariates - although authors considered date, other covariates could be justified such as time of day or year. In addition, it was noted by Thompson that the most serious shortcoming of the study was to not consider other patch attributes.

In regard to management implications, one reviewer wondered why a recommendation was not made for large patches to be maintained to the degree possible given their limited number on the landscape.

A recommendation was made that future work should focus on intra-patch distribution, patch level abundance, and productivity and that a few additional patch attributes be incorporated into the model (Dunk).

Laake et al. 2011

Much like the previous article, this was viewed to be well written and the inferences drawn to be reasonable. The conclusion that MR can result in biased estimates and that MRDS methods deserve more consideration are sound and viable. As stated by one reviewer, this approach appears to be useful and that the acknowledgement of heterogeneity in detection is appropriate.

Two areas for concern were noted by reviewers:

1) Regarding the point count intervals (Peak): If the assumptions associated with point count sampling theory were not satisfied then the estimate within each patch is biased. Two assumptions for sampling theory should be considered (1) observers detected birds at the point with certainty, and (2) observers detected birds at their initial location. If the period is too short, all individuals may not sing and be detected; if too long, double counting may occur. The reviewer identified that studies have shown that conventional 5-minute count periods "produce substantial upward bias in abundance estimates for bird species with rapid movement rates."

2) Regarding ambiguity of "detection" in the sampling (Hamel): the paper does not specify what constitutes as a "detection." It does not seem to be consistent throughout the papers. In addition, if this approach was "such a good one to eliminate heterogeneity, why was it not used in the population estimation process in the Morrison et al. report?

Collier et al. *In Press*

The sampling design for this paper was viewed by reviewers to be strong. Strengths of this paper were identified as the range-wide sampling design and the model performance leading to appropriate steps in estimating range wide abundance. The following areas for improvement/concern, however, were noted:

(Hamel) It will be important to check the accuracy/sufficient precision of the 2007-08 LANDSAT derived habitat classifications for this project. Reviewers assume that the classification is sufficient for this scale, but there is concern regarding the precision of the classification for conservation planning and implementing conservation activities at the local scale based upon this larger effort

(Thogmartin) Regarding inferences, the authors suggest using the geoaddative approach because GLMs and GAMs often provide only a measure of habitat suitability rather than occupancy. The reviewer does not agree with the reasoning provided by the authors and, at the very least, felt that the reason was incomplete regarding this inference.

Additional concern was raised by the reviewer that the parameters of the model are not credibly different from zero – the model estimates represent variation in noise. Thogmartin provided an extensive statistical analysis of the model that should be referenced directly from the independent review for details. The resulting information generated from Thogmartin's analysis has left the reviewer believing that "the credibility of the model is simply the result of a highly parameterized spline rather than their model covariates, describing the spatial structure of their data but nor the environmental correlates causing that structure."

Further concern was expressed by reviewers regarding the woodland composition of patches. It was also noted that all the concerns about a single year wildlife study apply to this study. However, it was noted that do to the coarse resolution, large spatial scale, and focus on occupancy the results may not be greatly affected by these concerns, unlike the density estimations in the previous two papers.

Some reviewers felt that this study could have accomplished more. Covariates (juniper density, age of woodlands, etc) acknowledged but not effectively integrated in the study would have given valuable insight into warbler ecology than the approach used. Additional spatial variables were noted that should have been evaluated for some relationship to GCWA presence (temperature, precipitation, etc). Not evaluating these was noted to have consequences in the future use of the model (pg 9, lines 173-176; Dunk). For example, if the "spatial coordinates are really representing variation in development of land conversion pressure, and that presser either increases of decreases, your model's spatial effect will remain static. However, if you had included some measure of vegetation/community type, or development or land conversion as a covariate the model would be more flexible (and accurate) when used in the future." The current utility of the model may be accurate, but it's future use may be questionable.

One reviewer recommended additional model validation by including a Figure 5a and 5b. With 5a depicting the "distribution (percentage) of detections among various probability of occupancy bins for the training/developmental data and 5b depicting the test/independent data (exactly as Figure 5 is now)." The purpose is to look for "similarities within these two figures to evaluate whether the developmental/training data are a good representation in general" (Dunk).

An additional recommendation included an evaluation with a larger analysis window (radius around sample locations) to account for more landscape-scale fragmentation (pg 15, lines 302-304; Dunk).

Mathewson et al. *In Review*

Due to the statistical detail and nature of this paper and subsequent review, it is highly recommended to continually reference each independent review (primarily Thogmartin and Dunk) for the specifics relating to the model evaluation. Those details that can be easily and effectively summarized are captured here.

One reviewer highlighted that the work captured within this study addresses the objectives expressed within the paper and is indeed a valuable case study of local/regional conservation efforts. It was noted that the comparison in Table 3 between the method of estimation as well as the abundance corrected for variation in detecatability greatly relieved initial concerns regarding

the fact that simple counts were used to develop population estimates for this paper.

The primary comment associated with this paper revolved around the comparison and broadening of data from 2008 to 2009. Due to the differences in data collected, it is not statistically valid to infer results across years. By treating the data the same, fitting it to the model, and then applying the 2009 sampling time frame, the potential result could be a positive bias in the resulting population estimate. A recommendation for how to alleviate this bias would be "to fit the abundance model to both years but include a year effect, and when using the model to predict abundance for the 2009 sampling frame, set the value of year equal to 2009" (Thompson).

Concerns were also voiced regarding a potential bias from avoiding edges. "By deviating from random sampling the way [the authors] did, they have under represented woodland habitat near edges in their survey and then by applying models derived from these samples to all woodland, [the authors] create a positive bias in their estimates" (Thompson). A recommended change in approach is to "explicitly consider key covariates in the model (like edge density) rather than rely on the sample design to account for these."

Numerous reviewers expressed reservation regarding the statistical analysis relating to confidence intervals and R^2 values associated with the model. For additional information relating to the statistical validity, review manuscript specific comments in the Thogmartin review. The authors need to account for other uncertainty in addition to the probabilistic nature of site occupancy. In addition to Thogmartin's assessment, Dunk provided additional recommendations regarding the use of outliers – the analysis should be presented both with and without these outliers because exclusion of these observations because they are influential (in a negative way) is "troublesome."

Dunk provided additional recommendations for model validation by running a "rigorous cross-validation" of the model by "randomly sub-setting the 301 sample locations, removing 20-25% of them, refitting the model and then classifying with the with-held observations." For specific details relating to this recommendation, please see Dunk's independent review.

Additional recommendations include referring to the population abundance estimate as of males, rather than of warblers, given that the density estimate is an estimate of abundance of males only; and commenting on the discrepancy between the sizes of the largest patch – as both studies were attempts at range wide design, and must have been examining the same "largest" patch.

The final recommendation for this paper from one reviewer is for the authors to include a few sentences suggesting "next steps" in the process of understanding GCWA ecology at the scale of geographic range. "Specifically, studies evaluating survival and/or reproduction among the various Po bins. Each of the pieces of information… (sampling methodology, occupancy modeling, spatial occupancy modeling, etc) are valuable and very useful. However, this "next steps" is a logical follow through on this work and will provide [the Service] and others…nearly a full set of tools that can be used for scientifically-informed management."

Morrison et al. 2010

As expressed by one reviewer, this document is the clear basis or culmination of the effort involved in the 5 documents. Although this study was noted as a great accomplishment (general comments), numerous concerns were expressed by reviewers regarding the confidence related to the modeling and abundance projections. This summary will only capture those high level concerns, but substantive details are present in each independent review.

The voluminous and lengthy nature of this report seems to be directed at a more technically inclined audience, while the actual messaging and conservation value are relatively inaccessible by the population of users in conservation planning, implementation, mitigation, and economic development. It's not an easily absorbed tool and it would behoove the authors to describe on paper, as briefly as occurred on the conference call, the care and diligence that they took in surveying patches that had been randomly selected, and how inaccessible patches were replaced by other randomly selected patches.

Substantial inquiry was expressed by reviewers regarding Model selection. As indicated in the general comments, additional justification and reason is needed to validate why Model II truly outperforms the SWCA (2007) model. As noted by one reviewer, the desire for a model that over-predicts vs under-predicts "is a preference from a conservationist's perspective. From a private landowner's perspective, I'm sure they would much prefer omission errors than commission errors. Given the very large proportion of the GCWA's range that is privately owned, this seems like it might be an issue at least to mention." Another reviewer noted that the "consequence of over-predicting suitable habitat is an important one for population estimation – a model that over-predicts habitat would by extension over-predict abundance using the extrapolative methods employed here."

Additional concern along these same lines is the potential for the maps/habitat projections to prioritize a high GCWA abundance sink area at the expense of a lower abundance source area. As requested by the reviewer, these caveats should be made clear to those who are likely to use these products.

In regards to model comparison, as mentioned in the Mathewson et al. comments, the comparison of the two modeling approached was confusing to one reviewer in that the sizes of the "largest" patches do not seem to be the same between the two models – this discrepancy needs to be cleared up.

Another reviewer noted that more descriptive data on sources of heterogeneity that were not modeled and that the authors assumed were controlled by their design is needed. For example, "what was the land cover composition of all patches and of the patches you sampled? What area of the patched fell within a 100m buffer of the patch edges, and percentage of your point counts that were in the buffer or in the interior?"

In reference to the credibility of predicting occupancy, one reviewer stated, "As best as I can tell, Figures 1-13 and 1-15 derive from models incapable of credibly predicting occupancy based upon patch size and landscape composition, respectively, as the parameter estimates do not likely differ from 0. Therefore, it is unclear what role these two variables should play in subsequent

inferences." This important observation from the reviewer is not easily summarized. Additional information expanding upon this comment can be found in the Thogmartin review.

Finally, two key concerns from a reviewer were articulated that should be highlighted in this summary.

1) It was noted by the authors that "…because our model allows for accurate prediction, we can identify priority conservation areas and activities based on local and regional expected rates of warbler occurrence and associated vegetative characteristics." As stated by the reviewer, "to some degree this is true, and it greatly exceeds the quality of information that exists in many parts of the world (including the U.S.) for species/systems that have conservation plans developed, and apparently that existed in Texas on GCWA prior to this effort." However, the reviewer expressed the need to put caveats on the use of this information for identifying "priority conservation areas" as this model could be better refined. It can be used to inform the process but, as the reviewer noted, I should not be the beginning or the end of such a process.

2) The concern was raised that agencies might read the author's statement of "…could be used to answer spatially-explicit questions, evaluate the likely outcome of recovery plans, and compare policy scenarios." and believe that the model provides more insight into the future than what actually exists.

A final closing thought was provided by one reviewer that captures the majority of recommendations highlighted within this summary: The Precautionary Principle would favor under-estimating abundance rather than over-estimating it as is seemingly done here – it is far preferable to predict extinction when the species is still extant than to predict abundance when it has in fact been extirpated (Myers 1993).

CONCLUSION

The Wildlife Society would like to emphasize the importance of reviewing each independent review in addition to the prepared summary. As expressed throughout this document, the level of detail and diversity of information reviewed within this packaged produced an extraordinary level of detail and analysis in the reviewers comments. This diversity and detail cannot be captured within a single summary. In addition, it is also important to recognize that although this summary captures numerous concerns and recommended changes to the papers under review, all of the reviewers felt that this was a critically important task based on strong science that, once improved upon, could serve as an effective tool for predicating occupancy and/or abundance on a range-wide level.

It should be noted that one review expressed a desire to see more discussion between the discrepancy of previous population estimates and those produced and evaluated by these studies.

Thank you for considering the views of wildlife professionals during the evaluation of these critical conservation materials. If questions arise regarding this summary or the independent comments submitted by reviewers, please contact The Wildlife Society's Assistant Director of Government Affairs, Terra Rentz (terra@wildlife.org).

RECOMMENDED LITERATURE & LITERATURE CITED

Balph, D. F., and H. C. Romesburg. 1986. The possible impact of observer bias on some avian research. Auk 103:831–833.

Boyce, M.S., Vernier, P.R., Neilsen, S.E., Schmiegelow, F.K.A., 2002. Evaluating resource selection functions. Ecological Modelling 157, 281–300.

Breiman, L. 2001. Statistical Modeling: The Two Cultures. Statistical Science 16: 199-231.

Buckland, S. T. Point-transect surveys for songbirds: robust methodologies. Auk 123:245–357.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford, UK.

Cade, B. and B. Noon. 2003. A Gentle Introduction to Quantile Regression for Ecologists. Frontiers in Ecology and the Environment 1:412-420.

Carroll, C., D. S. Johnson, J. R. Dunk, and W. J. Zielinski. 2010. Hierarchical Bayesian Spatial Models for Multispecies Conservation Planning and Monitoring. Conservation Biology 24:1538-1548.

Carroll, C., Dunk, J.R., Moilanen, A.J. 2010. Optimizing resiliency of multi-species reserve networks to climate change in the Pacific Northwest, USA. Global Change Biology 16, 891–904.

Carroll, C., Johnson, D.S. 2008. The importance of being spatial (and reserved): assessing Northern Spotted Owl habitat relationships with hierarchical Bayesian models. Conservation Biology 22, 1026–1036.

Cimprich, D. A. 2009. Effect of count duration on abundance estimates of Black-capped Vireos. Journal of Field Ornithology 80:94–100.

Davis, F.W., Seo, C., Zielinski, W.J. 2007. Regional variation in home-rangescale habitat models for fisher (*Martes pennanti*) in California. Ecological Applications 17, 2195–2213.

Diefenbach, D. R., D. W. Brauning, and J. A. Mattice. 2003. Variability in grassland bird counts related to observer differences and species detection rates. Auk 120:1168–1179.

Elith, J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography **29:**129–151.

Granholm, S. L. 1983. Bias in density estimates due to movement of birds. Condor 85:243-248.

Hao, L., and D. Q. Naiman, 2007. Quantile Regression, Sage Publications, San Diego, California, USA.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modeling 199, 142–152.

Johnson, D. H. 2001. Validating and evaluating models. Pages 105–119 in T. M. Shenk and A. B. Franklin (editors), Modeling in natural resource management. Island Press, Covelo, California.

Koenker, R. 2005. Quantile Regression, Econometric Society Monograph Series, Cambridge University Press, Cambridge, UK.

Lee, D. C., and S. J. Marsden. 2008. Adjusting count period strategies to improve the accuracy of forest bird abundance estimates from point transect distance sampling surveys. Ibis 150:315-325.

Marques, T. A., L. Thomas, S. G. Fancy, and S. T. Buckland. 2007. Improving estimates of bird density using multiple-covariate distance sampling. Auk 124:1229–1243.

Meyer, J. S., L. L. Irwin, and M. S. Boyce. 1998. Influence of habitat abundance and fragmentation on Northern Spotted Owls in western Oregon. Wildlife Monographs 139:1–51.

Myers, N. 1993. Biodiversity and the precautionary principle. Ambio 22:74-79.

Norvell, R. E., F. P. Howe, and J. R. Parrish. 2003. A seven-year comparison of relative abundance and distance-sampling methods. Auk 120:1013–1028.

Peak, R. G. 2011. A field test of the distance sampling method using the Golden-cheeked Warbler. Journal of Field Ornithology 82:309–317.

Ralph, C. J., and J. M. Scott (EDS.). 1981. Estimating numbers of terrestrial birds. Studies in Avian Biology, No. 6.

Reynolds, R. T., J. M. Scott, and R. A. Nussbaum. 1980. A variable circular-plot method for estimating bird numbers. Condor 82:309–313.

Rich, T. D., C. J. Beardmore, H. Berlanga, P. J. Blancher, M. S. W. Bradstreet, G. S. Butcher, D. W. Demarest, E. H. Dunn, W. C. Hunter, E. E. Iñigo-Elias, J. A. Kennedy, A. M. Martell, A. O. Panjabi, D. N. Pashley, K. V. Rosenberg, C. M. Rustay, J. S. Wendt, T. C. Will. 2004. Partners in Flight North American Landbird Conservation Plan. Cornell Lab of Ornithology. Ithaca, NY. Partners in Flight website. http://www.partnersinflight.org/cont_plan/ (VERSION: March 2005).

SCWA Environmental Consultants. 2007. Preliminary deliverable golden-cheeked warbler 513 status review. San Antonio, Texas, USA.

Wahl, R., D. D. Diamond, and D. Shaw. 1990. The Golden-cheeked Warbler; a status review. Final report submitted to Ecological Services, U. S. Fish and Wildlife Service, Fort Worth, Texas.

Zabel, C.J., Dunk, J.R., Stauffer, H.B., Roberts, L.M., Mulder, B.S., Wright, A. 2003. Northern spotted owl habitat models for research and management application in California (USA). Ecological Applications 13, 1027–1040.