20 January 2025

RH: Diefenbach et al. • Turkey productivity analysis

**A framework for analyzing wild turkey summer sighting data**

Duane R. Diefenbach, U.S. Geological Survey, Pennsylvania Cooperative Fish and Wildlife Research Unit, Pennsylvania State University, University Park, PA 16802, USA

Frances E. Buderman, Department of Ecosystem Science and Management, Pennsylvania State University, University Park, PA 16802, USA

Mary Jo Casalena, Pennsylvania Game Commission, Harrisburg, PA 17110, USA

Michael Dye, Virginia Department of Wildlife Resources, Verona, VA 24482, USA

Robert Gates, School of Environment and Natural Resources, Ohio State University, Columbus, OH 43210, USA

Laura C. Gigliotti, U.S. Geological Survey, West Virginia Cooperative Fish and Wildlife Research Unit, Morgantown, WV 26506, USA

C. Robert Long, Maryland Department of Natural Resources, Wildlife and Heritage Service, Cambridge, MD 21613, USA

Katherine Martin, Virginia Department of Wildlife Resources, Forest, VA 24551, USA

Michael Muthersbaugh, New York State Department of Environmental Conservation, Division of Fish and Wildlife, Albany, NY 12233, USA

Michael L. Peters, West Virginia Division of Natural Resources, Farmington, WV 26571, USA

James Sloan, New Jersey Department of Environmental Protection, Fish and Wildlife, Trenton, NJ 08625, USA

Joshua Stiller, New York State Department of Environmental Conservation, Division of Fish and Wildlife, Albany, NY 12233, USA

Mark Wiley, Ohio Department of Natural Resources, Division of Wildlife, Columbus, OH 43229, USA

**Correspondence:** Duane R. Diefenbach, 466 Bigler Rd, Pennsylvania State University, University Park, PA 16802, USA. Email: [DRD11@psu.edu](mailto:DRD11@psu.edu)

**ABSTRACT** Wildlife agencies collect data on productivity (e.g., proportion of hens with poults and number of poults per hen) of wild turkey (*Meleagris gallopavo*) to monitor population status and trends. However, sampling protocols to collect productivity data rely on opportunistic observations reported by wildlife agency personnel and the public and have changed over time and differed among agencies. A protocol to standardize data collection was adopted by most state wildlife agencies in 2019, but long-term historical datasets exist that cannot be analyzed readily to make inferences about spatial and temporal patterns in wild turkey productivity. We developed statistical models to allow comparisons and model trends in productivity among and within states even though data collection protocols changed over time and differed among states. We found greater spatial variation in the proportion of hens with poults than number of poults per brood, which may reflect how environmental factors influence wild turkey productivity. Our models can also provide inferences about productivity when data are limited or temporally discontinuous for some spatial units. Additionally, we found that temporal and spatial variation in data collection even under the new protocol can affect inferences about trends in productivity. The statistical models we developed address the uncontrolled nature of when and where data are collected and offer the ability to investigate long-term patterns of productivity in relation to factors such as changing climate or habitat conditions.

**KEYWORDS** *Meleagris gallopavo*, poults per brood, poults per hen, productivity, recruitment, summer sighting survey, wild turkey

Many state wildlife agencies collect data on wild turkey (*Meleagris gallopavo*) productivity where sightings of hens and poults are used to calculate metrics of reproductive success (e.g., poults per hen) to monitor wild turkey population trends (Byrne et al. 2015, Wakeling et al. 2022). Survey data are opportunistic observations of turkeys that may be reported by both agency personnel and the public. Therefore, statistics derived from these data are an index to turkey productivity because the survey methodology does not provide the ability to account for factors that could influence the number of turkeys or poults observed (i.e., detection probability, Williams et al. 2001). Wakeling et al. (2022) noted that summer sighting surveys lacked statistical power to detect changes and comparisons among regions were difficult. Byrne et al. (2015) encountered difficulty in conducting a regional analysis of wild turkey productivity because of a lack of a standardized protocol and a mismatch in the timing of data collection among, and even within, states.

Recognizing potential benefits of a standardized, repeatable sighting survey, the National Wild Turkey Federation (NWTF) Technical Committee adopted a standardized sighting protocol (NWTF Technical Committee 2019). Thirty-three state wildlife agencies adopted the NWTF protocol with the objective to ensure consistent methods of data collection and analysis to improve insights into trends of productivity over larger spatial areas. However, there remain differences among state agencies in their data collection methods. For example, even after 2019, survey period length differed among states and location data for each sighting was recorded at different spatial resolutions, such as geographic coordinates, county, township, or management unit (MU).

The standardized protocol adopted by the NWTF Technical Committee (2019) may improve future monitoring of wild turkey productivity, but there also exist decades of survey data that could provide insights into turkey population dynamics. For example, West Virginia, USA, has collected turkey sightings since 1967; many states have time series of >20 years of sighting data. Survey protocols have changed over time for individual states, which has limited inference about temporal changes in productivity within regions (*cf*. Byrne et al. 2015). In addition, even though surveys now follow a standardized protocol, turkey sighting data are still collected according to an unstructured sampling process.

Our objective was to develop statistical models that could use historical data from summer wild turkey surveys collected by wildlife agencies in the mid-Atlantic region of the United States (Maryland, New Jersey, New York, Ohio, Pennsylvania, Virginia, and West Virginia) even when sampling protocols changed over time and differed among states. Observations include information on the proportion of hens with poults (HWP) and the number of poults per brood (PPB). We incorporated information on the date that observations were collected to account for temporal differences in data collection without having to limit data to common survey dates among differing protocols. We used mixed-effects models with spatial and temporal random effects to permit inferences about productivity even when sample sizes were small or data were missing. Also, we used mixed-effects models to demonstrate how to estimate temporal trends in productivity that are less likely to be confounded by factors related to the uncontrolled nature of how turkey sighting data are collected. Finally, we investigated whether the current NWTF sampling protocol, which has been adopted by many state agencies, remains confounded by temporal and spatial sampling problems that can be addressed with the models we developed.

**STUDY AREA**

The study area included the states of Maryland, New Jersey, New York, Ohio, Pennsylvania, Virginia, and West Virginia. The easternmost physiographic province was the Atlantic Coastal Plain and included Long Island, New York. To the west was the Piedmont that adjoined the Ridge and Valley provinces and Appalachian Mountains and Allegheny plateaus. Further west into Ohio, northwestern Pennsylvania, and western New York occurred the Great Lake plains (and till plains of Ohio). To the north in New York, the physiographic provinces included the Catskill and Adirondack mountains and Hudson, Champlain, and St. Laurence lowlands. The mid-Atlantic region contained a diverse range of natural and human-dominated landscapes where wild turkeys occurred, including cities with millions of inhabitants, suburban development, fragmented forest interspersed with human development, contiguous forests, and expansive areas of cropland.

Each state wildlife agency managed wild turkey harvest based on MU boundaries defined using political or some combination of physiographic and political boundaries. For most states, we consolidated MUs into larger units that likely had similar turkey population characteristics and ensured sufficient sample sizes by MU. We used the 5 MUs Maryland and Ohio delineated for wild turkey. New York has 23 MUs for wild turkey that we consolidated into 4 units based on physiographic regions and similar wild turkey population characteristics, except we excluded the western portion of Long Island that included New York City. For Pennsylvania, we consolidated 22 MUs into 10 units with similar turkey population densities. New Jersey has 18 MUs, which we consolidated into 3 units with similar landscape and turkey population characteristics. Both Virginia and West Virginia used counties as their MUs, so we consolidated data into 4 physiographic regions in West Virginia and 5 physiographic regions in Virginia.

**METHODS**

**Data collection**

State wildlife agencies conducted wild turkey production surveys during summer. Survey timing varied among states and years, such that data were collected sometime during June–September. Opportunistic sightings of a wild turkey, or group of turkeys, during the survey period were recorded with each sighting event treated as an observation. At a minimum, observers recorded date, county, number of hens observed, number of poults observed, number of males observed, number of turkeys that could not be identified to sex or age, and whether the observer believed they previously recorded the turkeys observed. Maryland, New York, and Virginia instructed observers not to record observations of turkeys that they believed were previously recorded. A definition of a repeat observation was not provided to observers and repeat observations were not linked to previous observations in the database. All states recorded the MU where the observation occurred. In addition, some states recorded physiographic province (West Virginia), county and township (Pennsylvania and Virginia), and latitude and longitude (New Jersey, New York, Ohio, and Pennsylvania). Observers were state wildlife agency personnel and the public, except in West Virginia and Virginia where only agency personnel and cooperators provided observations.

All states followed the NWTF Technical Committee (2019) protocol after 2019, except New York only collected observations during August. The NWTF protocol was similar to what was described heretofore but restricted observations to July and August and excluded observations when >25% of the turkeys could not be identified to age or sex, ≥8 hens were observed with no poults, poults were observed with no hens, the poult:hen ratio was >16, and observations were reported to be of turkeys that were previously recorded.

We defined HWP as the proportion of hens observed with poults. If multiple hens were observed, we treated each hen as an independent observation. For example, 3 hens observed with poults were treated as 3 observations of hens with poults and 2 hens observed with no poults were treated as 2 observations of hens without poults. When hens with poults were observed, we calculated PPB as the number of poults divided by the number of hens and treated each hen as an independent observation. The product of HWP and PPB yields the average number of poults per hen (PPH).

**Data filtering and survey timing variation**

We examined how data filtering (e.g., excluding repeat observations) reduced sample size by calculating the number of observations that would be lost due to the different NWTF recommended data filters. In addition, we examined how dates when observations were reported differed among states and over time within states, which may limit the ability to use data from multiple states to conduct regional analyses.

**Date of observation effect**

The first set of developed hierarchical models was to address the problem with data being collected over different dates among states and within states over time. We included calendar day (DOY) of the observation as a potential covariate, which would allow us to estimate HWP and PPB for 31 August, which was a common date of data collection for all states. We describe the models below assuming a quadratic effect of DOY. To illustrate more complex models that can be used to partition the spatio-temporal variation in HWP and PPB, we developed models that included random intercepts for MU or a random effect that allowed variation by both year and MU.

For the HWP model, we developed a hierarchical generalized linear mixed model with a binomial distribution and logit link function to estimate the probability a hen has poults (HWPi = 1 if poults present during observation *i*, 0 otherwise). The first model was a random-intercept model where the random effect of MU ( was constant among years:

(1)

where *i* indexes *n* observations, *j* indexes *J* years, *k* indexes *K* MUs, is the corresponding year intercept (i.e., fixed effect of year) for observation *i*, and are coefficients for the DOY quadratic fixed effect, DOY*i* is the calendar day of observation *i*, and .

We specified a second model that included a random effect that varied by MU and year

(2)

where such that the random effect varied by year and MU. We developed generalized linear mixed models for PPB that are similar in structure to equations 1 and 2, but using a gamma distribution with an inverse link function:

and (3)

. (4)

We note that other specifications of the random effects in equations 2 and 4 could be developed that address different temporal and spatial correlations over time and space that may be of interest; we used equations 2 and 4 simply to show that more complex models can be used to capture variation in the observations among MU and time.

We used Akaike’s Information Criterion (AIC) to identify the most parsimonious model for each of the HWP and PPB model sets. Models with ΔAIC >2 were not considered competitive (Burnham and Anderson 2002). The candidate model set included models without an effect of DOY, a linear effect of DOY, or a quadratic effect of DOY; each of these DOY models was fit in combination with the constant or MU×year random effect for MU. We used Program R (R Core Team 2024) and the package *glmmTMB* to fit all models using maximum likelihood estimation (REML = F) because we compared models with different fixed effects using AIC (Brooks et al. 2017).

Finally, we used the best model in the previous analysis to assess whether excluding repeat observations changed inferences about HWP and PPB using data from Ohio, Pennsylvania, and West Virginia. We fit these models using datasets that included and excluded repeat observations and examined how estimates of HWP and PPB changed.

**Multi-state analyses**

By including DOY in models of HWP or PPB we accounted for the date that observations were collected, such that we could conduct multi-state analyses when survey periods differed among states or changed over years. Furthermore, MUs may have variable sample sizes or missing data among years; modeling variation among MUs as a random effect still permits estimation of HWP or PPB when data at the MU scale are limited or lacking (Schaub and Kery 2012). We chose 3 states to illustrate the benefits of the modeling approach, but one could extend it to all states conducting sighting surveys. We modeled HWP for the 3 conterminous states of New Jersey, New York, and Pennsylvania, 2011–2023. Survey periods differed among states (New Jersey, Jun–Sep; New York, Aug; Pennsylvania, Jul–Aug); New Jersey lacked data for 2020–2021; and Pennsylvania lacked data for 2011–2016. The model was formulated as

(5)

in which the notation follows equation 1, with the addition of *s*[*i*], which indexed the state for observation *i*, , and . The and were random slope effects on the coefficients and for DOY and DOY2, respectively, which allowed us to estimate state-specific relationships with DOY.

**Detecting annual and spatial trends in productivity**

If timing of survey periods changes over time, then inferences about trends in productivity could be confounded with dates that observations were collected. For each state, we developed a generalized linear mixed-effect model that accounted for dates when observations were collected to estimate an annual trend over time. We illustrate this type of analysis using HWP where we propose a modification of equation 1 in which Year is treated as a continuous, as opposed to a categorical, variable:

(6)

where is the slope parameter of the change in HWP by year, ~ *N*(0,) is a year-specific random effect, and ~ *N*(0,) is a MU-specific random effect. Note that this model does not include state-level variation in the effect of DOY (equation 5) and is merely meant to illustrate how one could estimate an annual trend that accounts for different survey timing.

Additionally, we considered how productivity might be confounded with a landscape characteristic, because where turkey observations occur is uncontrolled. If HWP varied spatially, temporal trends in productivity could be confounded with temporal changes in the location where turkey observations occurred. We modified equation 6 by adding a parameter for the proportion of developed land (DEV) where the observation occurred:

(7)

We used data from Pennsylvania, 2017–2023, and Ohio, 2017–2022, because these states had a sufficient time series of observations with location data for each observation. For Pennsylvania data, the township in which the observation occurred was recorded, whereas for Ohio latitude-longitude coordinates were recorded. We used 30 m × 30 m 2019 National Land Cover Database data for Ohio and Pennsylvania (Bocinsky et al. 2024), where we reclassified open space and low, medium, and high intensity developed areas as developed; deciduous, evergreen, mixed forests, shrub/scrub, and woody wetlands as forest; and all else as non-forest. The proportion of developed land had the strongest relationship with HWP, which is why we used DEV in this example. For Pennsylvania data, we calculated the proportion of area in each township that was classified as developed and assigned that value to each observation based on the township in which the observation occurred. For Ohio data, we used latitude-longitude coordinates to calculate the proportion of developed land within 1 km of the observation location. We scaled proportions using a square root-arcsine transformation.

We compared the models of equations 6 and 7 to a similar model that did not incorporate DEV and DOY to assess how ignoring DEV and DOY affected inference about trends in productivity. This model was similar to equations 6 and 7, but excluded DEV and DOY parameters:

(8)

In all models (equations 1–8), we can estimate HWP or PPB for any given DOY; however, in models without DOY, the expectation of the response would be the same at any point during the survey period. In all results that include estimates of HWP and PPB, we estimated HWP and PPB for 31 August because all states in all years collected data during August and this date could be used to represent the product of HWP and PPB as a measure of recruitment.

**RESULTS**

Excluding observations when ≥25% of turkeys were of unknown sex-age, >8 hens without poults, and >16 poults per hen had little effect on sample size (Table 1). However, excluding repeated observations of turkeys reduced sample size by 2–18% in Ohio, Pennsylvania, and West Virginia. Dates when observations were collected varied greatly among states and even among years within a single state (Table 2). Following the NWTF Technical Committee (2019) protocol, which restricts observations to July and August, >30% of data from some states would be excluded. New York collected data only in August.

Mean dates of observations varied greatly among states and even within individual states (Figure 2). New York collected data during August and had the least variability but other states, such as Virginia, exhibited large variations in mean dates because of changes to the survey period. Even within a state, variation in mean date of observations over years and among WMUs was quite large: New Jersey, 98 days; New York, 12 days; Ohio, 22 days; Pennsylvania, 29 days; Virginia, 59 days; and West Virginia, 37 days. A trend towards earlier mean dates of observation was evident even in states with consistent survey periods, such as New Jersey and Pennsylvania (Table 2, Figure 2).

**Effect of observation date**

Calendar day (DOY) was an important variable in models of HWP and PPB and for all states except Maryland was best modeled as a quadratic relationship (Table 3). Models that did not include DOY and DOY2 were not competitive (ΔAIC > 44). For HWP, the proportion of hens observed with poults increased over calendar day (DOY) but appeared to reach an asymptote by 1 August (Figure 3). New York was an exception, however, where the proportion of hens with poults declined and then increased, but the effect of DOY was minor because their survey period was only the month of August. Brood size (PPB) declined quickly in June and tended to stabilize by mid-July for all states (Figure 4).

We found HWP models with a random effect that varied by year and MU (equations 2 and 4) were parsimonious (lower AIC values) compared to models that only included a MU random effect (Table 3). There are multiple ways to model random effects in these models that could be explored, although some require more parameters to be estimated and are more likely to have convergence problems. However, the model that included a random effect that varied by MU and year (equation 2) was able to converge for all datasets. Despite evidence HWP varied by year and MU, we included only MU-specific random effects in all other analyses to simplify analyses and interpretation of the other proposed statistical models (equations 5–8).

Models of PPB for all states indicated little evidence of spatial or temporal variation in PPB among MUs because the estimated variance for the random effect among all states and models was small (< 0.001). In most datasets, the random effects that differed by year and MU had the lowest AIC values but sometimes models with only a MU random effect had the lowest AIC value (Table 4). Also, not all models that were fit to the Virginia data converged.

Retaining observations reported to have been repeat observations of the same birds had little effect on estimates for both HWP and PPB (Figures 5, 6). Point estimates were slightly larger for Pennsylvania and West Virginia but the temporal pattern did not change for all 3 states. We included repeat observations in datasets in all subsequent analyses.

**Multi-state analyses**

The generalized linear mixed-effects model (equation 5) allowed us to combine data from New Jersey, New York, and Pennsylvania without excluding data that were not collected within the same survey period among states. We were able to address the effect of DOY on estimates despite the large differences among states in survey periods (compare Figure 7 [bottom] to Figure 3). In addition, we were able to estimate HWP for New Jersey during 2020–2021 and 2011–2016 for Pennsylvania when neither state had data available. Because the survey period in New York was August, excluding June, July, and September data collected by New Jersey and Pennsylvania would greatly reduce sample sizes. New Jersey would have to exclude 76% (*n* = 3,459) of their data and Pennsylvania would have to exclude 52% of their data (*n* = 14,144).

**Detecting trends in productivity**

The effect of ignoring calendar day when modeling trends in productivity influenced both the precision and point estimates of the slope parameter in the model, which represents the annual trend. Incorporating calendar day in the trend model for Pennsylvania indicated no change over time (ΔAIC = 0.0; 90% confidence interval encompassed 0; black error bars in Figure 8) whereas a decline would be inferred when excluding DOY in the model (ΔAIC = 22.8; blue error bars in Figure 8). In contrast, nearly opposite conclusions would be inferred for Ohio data because the trend was positive when DOY was not included in the model (ΔAIC = 103.9; blue error bars in Figure 8); no trend was detected when DOY was included (ΔAIC = 8.6; black error bars in Figure 8); and the trend was positive with the largest precision when DOY and habitat were included (ΔAIC = 0.0; orange error bars in Figure 8). Including DOY in the New York model had little effect on either the point estimate of the slope parameter because New York only collected data during August; however, there was little loss of precision for including DOY and DOY2 in the model as evident by nearly identical confidence intervals (Figure 8).

**DISCUSSION**

Recruitment in wild turkeys can be considered a sequential process whereby hens first need to complete egg-laying and incubation, then hatch eggs, and finally successfully rear poults. Consequently, we chose to quantify productivity in wild turkeys using the proportion of hens with poults (HWP, i.e., successfully hatched eggs), and then for those hens with poults to estimate brood size (PPB). The product of HWP and PPB can be used as an index to recruitment (poults per hen [PPH]). We believe separately estimating these 2 components of recruitment in wild turkey is preferable to estimating PPH because different factors are likely affecting nest success and poult survival. In fact, we found differences in variability between HWP and PPB, in which HWP exhibited greater spatial variation among MUs compared to PPB. More importantly, the timing of when observations of HWP and PPB are collected (DOY) affects these productivity parameters differently.

The patterns of change that we observed in HWP and PPH with DOY were expected from what is known about breeding biology of wild turkeys. The proportion of hens with poults increased asymptotically with date reflecting the protracted nesting and re-nesting chronology of wild turkeys and declining brood size reflected poult mortality. The patterns we found with DOY provide empirical evidence that summer observations are likely to provide useful and relevant data for management. Furthermore, we can estimate HWP and PPB for any given date because our models account for DOY. We used a DOY of 31 August, the latest date that all states collected summer sighting data, such that all nesting and most poult mortality had occurred so that HWP and PPB provided an index to recruitment (HWP × PPB) that was comparable among all states. Not only does including DOY in models account for variation in when observations occur but it allows direct comparison of estimates of HWP and PPB when sampling protocols differ spatially (i.e., among states) or over time.

A fundamental challenge with making inferences from the summer sighting survey remains the uncontrolled collection of observations over space and time. Despite the advantages of separately estimating HWP and PPB, numerous factors influence estimates obtained from opportunistic sightings of wild turkey during summer sighting surveys. Biologically, crèche behavior in wild turkeys means that multiple hens and poults may be sighted together, and one cannot assess whether all hens present are a parent of offspring observed. Counts of hens and poults are fraught with potential errors, such as undercounting poults because of vegetation obstructing the observer’s view and misidentification of poults and hens when poults approach adult size. Furthermore, both undercounting and misidentification could be confounded by phenological changes in vegetation during the survey period that may influence the ability of observers to accurately count and determine the age of turkeys. Consequently, the patterns we observed in DOY, although likely dominated by patterns in timing of nesting, nest survival, and poult mortality, also include factors that influence and introduce error in counts of turkeys.

Data collected to estimate productivity in wild turkey ideally would be completed during a short (nearly instantaneous) period and on the same dates each year; although this is logistically difficult. The protocol adopted by the NTWF Technical Committee (2019) attempts to address the temporal challenges associated with sighting surveys by restricting observations to July and August. However, given the extended period over which wild turkey nest and high mortality of poults during these 2 months, there remain important temporal patterns in the data. Changes to when data are collected can influence parameter estimates when the average observation date changes over time, even within the July-August timeframe. We have demonstrated that when and where observations are reported can change inferences from these data; for example, temporal trends in HWP will differ if calendar day or habitat conditions where observations occur are ignored. Consequently, the NWTF protocol does not eliminate the need for statistical models to account for DOY and the spatial distribution of observations.

Data filtering parameters of the NWTF Technical Committee (2019) protocol are designed to maximize comparability among datasets, such as removing 1) observations with ≥25% of birds unidentified to age and sex, 2) >8 hens observed without poults, 3) >16 poults per hen, and 4) if the group of birds were previously reported. Filtering data based on criteria 1–3 have little effect on sample size and unlikely to influence estimates of HWP or PPB. However, the decision to exclude observations of >8 hens without poults will increase estimates of HWP. Similarly, excluding sightings of hens with >16 poults will reduce estimates of PPB. The NWTF Technical Committee (2019) protocol was not designed to estimate a detection probability or misidentification rate, so filters 2 and 3 are simply ad hoc attempts to account for observations most likely to represent situations where not all birds were detected or some were misidentified.

The fourth data filter, excluding repeated sightings, is problematic for several reasons. First, repeat sightings can represent nearly 1 in 5 observations, which is an amount of potential data that could inform HWP and PPB. Second, the data collection protocol does not allow the observer to specify which observation is a repeat of previously reported birds. If repeated observations were linked, then statistical methods could readily account for these observations and provide insights into the recruitment process (e.g., poult survival). Third, it is impossible to determine when the same turkeys are reported by different observers. These types of repeat observations are unidentifiable and likely exist in the data.

There are no standards in the NWTF Technical Committee (2019) protocol that define a repeat observation. Observed turkeys are not marked, so there is no means to verify that the same turkeys are observed more than once; the observer must make the determination of a repeat observation. Quantitative methods in wildlife ecology have endeavored to address the uncertainty of counting animals and it is well established that detection probability for nearly all survey methods is <1.0 (Williams et al. 2001). Most estimators of population parameters rely on marked animals to identify repeat observations, with a relatively few estimators developed that rely on unmarked animals (e.g., Buckland et al. 2001). It may be useful to request observers to avoid specific types of repeat observations, such as when hens with poults repeatedly visit bird feeders, but our results suggest that repeat observations had little influence on HWP and PPB estimates.

Our analysis suggest that the spatial distribution of observations also could affect inference about trends in recruitment. If job duties of agency staff change, and if state agencies rely more upon observations provided by the public, then spatial and temporal patterns of observations may change over time. The difficulties of using observations from an unstructured sampling process, when uncorrected for sampling effort, to make inferences about wildlife populations is well known (e.g., eBird, <https://ebird.org/home>, accessed 29 June 2024; Fink et al. 2023). Using township landscape characteristics for Pennsylvania was a predictor of productivity and influenced estimates of the temporal trend in HWP. Similarly, Ohio collected latitude and longitude for most locations and incorporating habitat associated with observations improved the precision of the estimated temporal trend. We found that accounting for habitat type changed inferences about temporal trends in recruitment, which suggests it may be useful for summer sighting surveys to record coordinates of where observations occurred.

We demonstrated that accounting for when and where observations are obtained can greatly influence inferences from these data. Advantages of our statistical models are that analyses can encompass time periods over which the survey period changes within a state (e.g., Maryland and Virginia); analyses can use data from multiple states even when survey periods differ; and inferences can be made even when data for some MUs are limited or missing (e.g., New Jersey). Furthermore, our models can account for potential biases introduced because of the uncontrolled nature of how observations are collected temporally and spatially. Perhaps most importantly, our proposed approach offers the ability to investigate longer-term patterns of productivity in relation to factors such as changing climate or habitat conditions.

**ACKNOWLEDGMENTS**

We thank K. J. Smelter for creating Figure 1 and J. Trowbridge for creating the landcover covariates for Pennsylvania and Ohio data. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

**ETHICS STATEMENT**

The authors declare no conflict of interest.

**REFERENCES**

Bocinsky R. K. 2024. FedData: functions to automate downloading geospatial data available from several federated data sources. R package version 4.0.1, <https://CRAN.R-project.org/package=FedData> (accessed 5 August 2024).

Brooks M. E., K. Kristensen, K.J. van Benthem., A. Magnusson, C.W. Berg, A. Nielsen, H.J. Skaug, M. Maechler, and B.M. Bolker. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R Journal 9(2):378–400.

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, England, United Kingdom.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer, New York, New York, USA.

Byrne, M. E., M. J. Chamberlain, and B. A. Collier. 2015. Potential density dependence in wild turkey productivity in the southeastern United States. Proceedings of the National Wild Turkey Symposium 11:329–351.

Fink, D., A. Johnston, M. Strimas-Mackey, T. Auer, W. M. Hochachka, S. Ligocki, L. Oldham Jaromczyk, O. Robinson, C. Wood, S. Kelling, and A. D. Rodewald. 2023. A double machine learning trend model for citizen science data. Methods in Ecology and Evolution 14:2435–2448.

National Wild Turkey Federation Technical Committee. 2019. A standardized protocol for conducting wild turkey brood surveys. National Wild Turkey Federation, Edgefield, South Carolina, USA.

R Core Team. 2024. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Schaub, M., and M. Kéry. 2012. Combining information in hierarchical models improves inferences in population ecology and demographic population analyses. Animal Conservation 15:125-126.

Wakeling, B. F., J. M. Sandrini, S. P. Lerich, and C. Cardinal. 2022. A review of harvest and population monitoring methods for western wild turkeys. Wildlife Society Bulletin 46:e1281.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2001. Analysis and management of animal populations: modeling estimation, and decision making. Academic Press, London, England, United Kingdom.

Figure Captions

Figure 1. States and management unit boundaries (Maryland and Ohio), or boundaries of consolidated management units (New Jersey, New York, Pennsylvania, Virginia), or physiographic regions (West Virginia) with observations of wild turkey (*Meleagris gallopavo*) hens and poults, Maryland, New Jersey, New York, Ohio, Pennsylvania, Virginia, and West Virginia, USA.

Figure 2. Mean date observations were obtained of wild turkeys (*Meleagris gallopavo*) for summer sighting surveys by year for 2000–2023 in Maryland (MD), New Jersey (NJ), New York (NY), Ohio (OH), Pennsylvania (PA), Virginia (VA), and West Virginia (WV), USA.

Figure 3. Effect of day of year with 95% confidence intervals on the proportion of hens observed with poults during summer sighting surveys of wild turkey (*Meleagris gallopavo*), 2000–2023, in Maryland (MD), New Jersey (NJ), New York (NY), Ohio (OH), Pennsylvania (PA), Virginia (VA), and West Virginia (WV), USA.

Figure 4. Effect of day of year with 95% confidence intervals on the number of poults per brood during summer sighting surveys of wild turkey (*Meleagris gallopavo*), 2000–2023, in Maryland (MD), New Jersey (NJ), New York (NY), Ohio (OH), Pennsylvania (PA), Virginia (VA), and West Virginia (WV), USA.

Figure 5. Estimated proportion of hens with poults (HWP) during summer sighting surveys of wild turkey (*Meleagris gallopavo*) in Ohio, Pennsylvania, and West Virginia, USA, when repeat observations of the same birds, as reported by observers, are excluded (left column) and when all observations are included, 2016–2023. Error bars are 95% confidence intervals and horizontal lines are for reference to compare estimates.

Figure 6. Estimated poults per brood (PPB) during summer sighting surveys of wild turkey (*Meleagris gallopavo*) in Ohio, Pennsylvania, and West Virginia, USA, when repeat observations of the same birds, as reported by observers, are excluded (left column) and when all observations are included (right column), 2016–2023. Error bars are 95% confidence intervals and horizontal lines are for reference to compare estimates.

Figure 7. Multi-state wild turkey (*Meleagris gallopavo*) model of the proportion of hens with poults, by management unit, for New Jersey (NJ), New York (NY), and Pennsylvania (PA), USA (top) and the effect of calendar day on the estimator (bottom), 2011­–2023. The model demonstrates estimates can be obtained for years with missing data (New Jersey, 2020–2021; Pennsylvania, 2011–2017) and that different relationships for calendar day are retained (*cf*. Figure 3).

Figure 8. Annual change in proportion of hens with poults over time and 90% confidence interval observed during summer sighting surveys of wild turkeys (*Meleagris gallopavo*) when calendar day of observations is incorporated in the model (black), when it is not (blue), and when both calendar day and proportion of developed area associated with the observation are included (orange); Maryland (MD), 2018–2023; New Jersey (NJ), 2000–2023; New York (NY), 2008–2023; Ohio (OH), 2017–2022; Pennsylvania (PA), 2017–2023; Virginia (VA), 2008–2023; West Virginia (WV), USA, 2016–2023

Table 1. Percentage of data removed when excluding observations where ≥25% of wild turkeys (*Meleagris gallopavo*) were of unknown sex-age (Unknown), >8 hens sighted without poults present (Hens), >16 poults per hen (PPH), and whether the group had been previously reported by the observer (Seen before).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| State | Years | *n*a | Unknown (%) | >8 Hens (%) | >16 PPH (%) | Seen before (%) |
| Marylandb | 2003–2023 | 11,181 | 2.6 | 0.56 | 0.06 |  |
| New Jersey | 2000–2023c | 3,463 | 1.9 | <0.01 | <0.01 | 12.3c |
| New Yorkb | 2005–2023 | 13,490 | 2.9 | 0.73 | 0.07 |  |
| Ohio | 2009–2023 | 13,494 | 4.1 | 1.70 | 0.19 | 5.4 |
| Pennsylvania | 2017–2023 | 14,838 | 3.7 | 0.94 | 0.10 | 18.0 |
| Virginiab | 2008–2023 | 3,786 | 1.8 | 1.10 | 0.03 |  |
| West Virginia | 2016–2023 | 2,440 | 1.4 | 0.61 | 0.04 | 7.3 |

a Number of observations before filtering

b States that request observers not record repeat sightings.

c New Jersey did not collect data in 2000–2001 and only recorded whether an observation of turkeys was previously reported during 2022–2023 (*n* = 565).

Table 2. Number of observations of wild turkey (*Meleagris gallopavo*) by month and year prior to filtering data according to the National Wild Turkey Federation protocol. Maryland and Virginia changed the time period when data were collected.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | Number of observations by month | | | |
| State | Years | June | July | August | September |
| Maryland | 2003–2008 | 1,021 | 845 | 672 | 284 |
|  | 2009–2023 | 0 | 4,863 | 3,496 | 0 |
| New Jersey | 2000–2023 | 1,199 | 1,034 | 806 | 424 |
| New York | 2005–2023 | 0 | 0 | 13,490 | 0 |
| Ohio | 2009–2023 | 4,828 | 5,089 | 3,577 | 0 |
| Pennsylvania | 2017–2023 | 0 | 7,783 | 7,055 | 0 |
| Virginia | 2008–2009 | 157 | 147 | 158 | 0 |
|  | 2010–2017 | 0 | 0 | 1,467 | 0 |
|  | 2018–2023 | 0 | 943 | 914 | 0 |
| West Virginia | 2016–2023 | 877 | 851 | 712 | 0 |

Table 3. Model selection results for models of the proportion of hens with poults that did not include repeat observations of wild turkey (*Meleagris gallopavo*) during summer sighting surveys in mid-Atlantic states, 2000–2023.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | State | | | | | | | | | | | | | | | | | | | |
|  | Maryland | |  | New Jersey | |  | New York | |  | Ohio | |  | Pennsylvania | |  | Virginia | |  | West Virginia | |
| Modelc | Ka | ΔAICb |  | K | ΔAIC |  | K | ΔAIC |  | K | ΔAIC |  | K | ΔAIC |  | K | ΔAIC |  | K | ΔAIC |
| Year, DOY, DOY2 (MU) | 8 | 39.3 |  | 25 | 497.5 |  | 22 | 146.7 |  | 8 | 1.2 |  | 8 | 149.3 |  | 19 | 92.4 |  | 11 | 80.8 |
| Year, DOY (MU) | 7 | 38.1 |  | 24 | 531.2 |  | 21 | 189.7 |  | 7 | 76.7 |  | 7 | 162.6 |  | 18 | 132.8 |  | 10 | 151.8 |
| Year (MU) | 6 | 217.9 |  | 23 | 874.2 |  | 20 | 187.8 |  | 6 | 636.6 |  | 6 | 298.5 |  | 17 | 130.9 |  | 9 | 279.9 |
| Year, DOY, DOY2 (MU × Year) | 8 | 0.4 |  | 25 | 0.0 |  | 22 | 0.0 |  | 8 | 0.0 |  | 8 | 0.0 |  | 19 | 0.0 |  | 11 | 0.0 |
| Year, DOY (MU × Year) | 7 | 0.0 |  | 24 | 43.3 |  | 21 | 43.4 |  | 7 | 76.8 |  | 7 | 12.8 |  | 18 | 47.1 |  | 10 | 68.1 |
| Year (MU × Year) | 6 | 183.7 |  | 23 | 380.0 |  | 20 | 41.8 |  | 6 | 633.6 |  | 6 | 139.0 |  | 17 | 45.1 |  | 9 | 184.8 |

a Number of parameters in the model.

b Difference in Akaike’s information criterion (AIC) compared to the model with the lowest AIC value.

c Year was categorical, DOY was calendar day, and MU was management unit. Variables in parentheses describe random effect patterns (see Methods).

Table 4. Model selection results for models of the number of poults per brood that did not include repeat observations of wild turkey (*Meleagris gallopavo*) during summer sighting surveys in mid-Atlantic states, 2000–2023.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | State | | | | | | | | | | | | | | | | | | | |
|  | Maryland | |  | New Jersey | |  | New York | |  | Ohio | |  | Pennsylvania | |  | Virginiaa | |  | West Virginia | |
| Modeld | Kb | ΔAICc |  | K | ΔAIC |  | K | ΔAIC |  | K | ΔAIC |  | K | ΔAIC |  | K | ΔAIC |  | K | ΔAIC |
| Year, DOY, DOY2 (MU) | 9 | 5.2 |  | 26 | 0.0 |  | 23 | 3.7 |  | 9 | 17.6 |  | 9 | 0.0 |  | 20 | 4.1 |  | 12 | 6.6 |
| Year, DOY (MU) | 8 | 18.8 |  | 25 | 10.0 |  | 22 | 9.1 |  | 8 | 19.2 |  | 8 | 4.0 |  | 19 |  |  | 11 | 23.3 |
| Year (MU) | 7 | 3.2 |  | 24 | 74.5 |  | 21 | 32.4 |  | 7 | 128.9 |  | 7 | 124.3 |  | 18 | 25.6 |  | 10 | 44.8 |
| Year, DOY, DOY2 (MU × Year) | 9 | 2.0 |  | 26 | 7.7 |  | 23 | 0.0 |  | 9 | 0.0 |  | 9 | 10.4 |  | 20 | 0.0 |  | 12 | 0.0 |
| Year, DOY (MU × Year) | 8 | 0.0 |  | 25 | 18.3 |  | 22 | 4.9 |  | 8 | 1.9 |  | 8 | 14.2 |  | 19 |  |  | 11 | 16.1 |
| Year (MU × Year) | 7 | 15.5 |  | 24 | 79.6 |  | 21 | 28.8 |  | 7 | 111.5 |  | 7 | 134.8 |  | 18 | 20.6 |  | 10 | 38.3 |

a Not all models were estimable.

b Number of parameters in the model.

c Difference in Akaike’s information criterion (AIC) compared to the model with the lowest AIC value.

d Year was a categorical variable, DOY was calendar day, and MU was management unit. Variables in parentheses describe the random effect pattern (see Methods)