Sanchez *et al. Avian Res* (2021) 12:28
https://doi.org/10.1186/s40657-021-00258-5

Avian Research

METHODOLOGY

Open Access

A field guide for aging passerine nestlings using growth data and predictive modeling

Audrey A. Sanchez[1*], Andrew W. Bartlow[2], Allison M. Chan[3], Jeanne M. Fair[2], Aaron A. Skinner[1], Kelly Hutchins[1], Maria A. Musgrave[1], Emily M. Phillips[1], Brent E. Thompson[1] and Charles D. Hathcock[1]

Abstract

Background: Accurate nestling age is valuable for studies on nesting strategies, productivity, and impacts on reproductive success. Most aging guides consist of descriptions and photographs that are time consuming to read and subjective to interpret. The Western Bluebird (*Sialia mexicana*) is a secondary cavity-nesting passerine that nests in coniferous and open deciduous forests. Nest box programs for cavity-nesting species have provided suitable nesting locations and opportunities for data collection on nestling growth and development.

Methods: We developed models for predicting the age of Western Bluebird nestlings from morphometric measurements using model training and validation. These were developed for mass, tarsus, and two different culmen measurements.

Results: Our models were accurate to within less than a day, and each model worked best for a specific age range. The mass and tarsus models can be used to estimate the ages of Western Bluebird nestlings 0–10 days old and were accurate to within 0.5 days for mass and 0.7 days for tarsus. The culmen models can be used to estimate ages of nestlings 0–15 days old and were also accurate to within less than a day. The daily mean, minimum, and maximum values of each morphometric measurement are provided and can be used in the field for accurate nestling age estimations in real time.

Conclusions: The model training and validation procedures used here demonstrate that this method can create aging models that are highly accurate. The methods can be applied to any passerine species provided sufficient nestling morphometric data are available.

Keywords: Cavity-nesting, Nest boxes, Nestling development, Predictive models, Western Bluebird

Background

The nestling period is a sensitive time within the life cycle of altricial birds that strongly influences their survival and reproductive success (Langham 1972; Bryant 1978; Amiot et al. 2014). Quantitative measures of nestling growth and development are important for studying avian breeding biology and reproductive strategies (Amiot et al. 2014). Accurately aging nestlings at younger ages is challenging, but important for data quality throughout the nestling period. Often, nest box studies require collecting data on nestlings multiple times during the nestling period. If the initial age estimate is not accurate, there will be data quality issues throughout the study. Because young nestlings do not yet display some obvious physical characteristics (e.g., feather tract development), photographic guides and models based on morphometric measurements aid in accurately estimating age (Bortolotti 1984a; Wails et al. 2014; Costa et al. 2020).

The ability to accurately age nestlings is an important aspect of avian ecology that yields insight into the effects of different nesting strategies on nest success (O'Connor

*Correspondence: audrey_a@lanl.gov
[1] Environmental Stewardship, Los Alamos National Laboratory, Los Alamos, NM, USA
Full list of author information is available at the end of the article

© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

1978; Shaffer 2004) and helps identify impacts of environmental variables on fitness, feeding habits, growth, and reproductive success (Jongsomjit et al. 2007). Studies that quantify nest success often use models (e.g., Dinsmore et al. 2002; Shaffer 2004) that directly input nest age (determined by the nestlings within that nest) as a covariate to predict nest success, and thus accurate nestling age would also be critical for any study employing these models. Overall trends in population demographics and determining different stages of nestling development that are most impacted by adverse environmental conditions can benefit from accurately aged nestlings (Partridge and Harvey 1988; Brawn 1991). Without an accurate way to estimate age, productivity can easily be overestimated or underestimated depending on the methodology used (Wails et al. 2014). In addition, it is important for determining when nestlings can be banded (Murphy 1981; Costa et al. 2020). Missing these opportunities could result in incomplete and poor-quality data if birds fledge before they can be banded. Other examples of single opportunity sampling (i.e., only one chance to obtain data) are studies that rely on taking blood samples or sexing birds at certain ages.

The most accurate method of aging nestlings would result from daily nest checks until eggs hatch, which provides researchers with a specific hatch date. However, this is not always feasible; checking a nest daily is time consuming, expensive, causes stress on the birds, and can lead to increased predation (Wails et al. 2014). Some researchers have used feather tract development and physical characteristics in photographic guides for age determination for a variety of species (Murphy 1981; Podlesak and Blem 2002; Jongsomjit et al. 2007; Fernaz et al. 2012; Brown et al. 2013; Amiot et al. 2014; Wails et al. 2014; Costa et al. 2020). However, the use of photographs alone can be highly subjective due to quality, perspective, and image scale (Bechard et al. 1985; Brown et al. 2013). There are also guides that contain written physical descriptions of nestlings for each day (Pinkowski 1975; Amiot et al. 2014). Costa et al. (2020) developed a photographic guide for aging European Bee-eaters (*Merops apiaster*) that is accurate to within 3 days. This guide also contains written descriptions on eye development, bill size and color, feather color and stages of development, motor coordination, and overall size for 3-day intervals throughout the nesting period. For researchers that manage large nest box networks and have crew members with varying levels of experience, reading a written description for each nestling is time consuming and interpretation is subjective.

Morphometric data have been used to estimate nestling age of various species [see Wails et al. (2014) for a list of published studies]. Predicting age from morphometric measurements using linear regression, where the morphometric measurement is the independent variable, has been used so that researchers can age nestlings in the field (Bortolotti 1984b; Gilliland and Ankney 1992; Palacios and Anderson 2018). This method of nestling age estimation has been used in Bald Eagles (*Haliaeetus leucocephalus*) (Bortolotti 1984b) and California Brown Pelicans (*Pelecanus occidentalis californicus*) (Palacios and Anderson 2018). Validation to determine the accuracy of a given aging method is a critical step, so that researchers know the uncertainty around their estimation (Wails et al. 2014; Costa et al. 2020). Failing to know the accuracy of estimations could result in poor-quality data. One way to validate age estimates is to use predictive modeling and validation procedures. This is done by applying the model to a subset of known test data, which are withheld from the model-building dataset, to quantify (e.g., using root mean squared error) how close the model-predicted values match the known test data.

Another method of decreasing uncertainty in age estimates is to use multiple morphometric measurements. Brown et al. (2011) compared the accuracy between inexperienced and experienced researchers for two methods of age determination: feather tract development and morphometric data. They found age estimates were most accurate for both inexperienced and experienced researchers when more than one morphometric measurement was used. Additionally, using a combination of morphometric measurements and physical characteristics is recommended if time allows and birds are not stressed (Murphy 1981; Haggerty 1994; Podlesak and Blem 2002; Jongsomjit et al. 2007).

The Western Bluebird (*Sialia mexicana*) is a small territorial passerine that is widely distributed in western North America from southern British Columbia to southern Mexico (Dickinson and Leonard 1996; Keyser et al. 2004). They are a secondary cavity-nesting songbird that inhabits coniferous and open deciduous forests. Ponderosa pine (*Pinus ponderosa*) forests provide nest cavities and low perches for insect hunting and constitute one of the Western Bluebird's typical habitats (Kozma and Kroll 2010). Western Bluebirds also live and nest in pinyon-juniper woodlands comprised of pinyon pine (*Pinus edulis*) and juniper (*Juniperus* sp.) trees.

Bluebirds also use nest boxes for breeding when there is cavity competition (Brawn and Balda 1988; Brawn 1991) or when there is loss of breeding habitat (Keyser et al. 2004). Nest box programs not only provide suitable nesting locations, but they also provide a mechanism for data collection and a way to monitor and evaluate environmental and anthropogenic impacts to populations over long time periods (Musgrave et al. 2019; Wysner et al. 2019).

Despite various nest box studies on Western Bluebirds (Brawn 1991; Wang and Weathers 2009; Fair et al. 2010), no nestling aging guides exist. There is an image-based guide for aging Eastern Bluebird (*S. sialis*) nestlings (Pinkowski 1975), a non-peer reviewed guide describing general development of all bluebird species among different age ranges (Graham 2006), and an assessment testing the accuracy of three photographic guides for aging Eastern Bluebird nestlings (Wilkins and Brown 2015).

Here, we developed methods to age Western Bluebird nestlings using data collected from nest boxes over three breeding seasons in northern New Mexico. Our specific objectives were to create linear regression models built from morphometric data and present measurement ranges for tarsus, mass, and two different culmen measurements for each nestling age (0–21 days) to accurately age nestlings. Our regression approach included using multiple morphometric measurements to develop a series of validated predictive models that can be used to calculate age from measurements taken in the field. We provide information on how to use our data for researchers working with Western Bluebirds and how these methods can be applied to other passerine species using R code that is shared on a public data repository.

Methods

Study site

Our field work was conducted using an existing nest box network at sites on and around Los Alamos National Laboratory in Los Alamos County, New Mexico. Los Alamos County is located on the eastern flanks of the Jemez Mountains in north-central New Mexico, approximately 2200 m in elevation, on the Pajarito Plateau. The landscape consists of narrow mesas separated by steep-sided canyons. The dominant habitat type is dependent on elevation but there are four major habitat types in Los Alamos County: mixed conifer forest, ponderosa pine forest, pinyon-juniper woodland, and juniper grasslands (Fair and Myers 2002).

Field work

Nest boxes were placed on trees in locations dominated by ponderosa pine. A total of 14 nests over three breeding seasons (2015–2017) were monitored in this study. The number of nestlings studied varied each year with 10 in 2015, 26 in 2016, and 27 in 2017. We checked nest boxes every 2 weeks beginning in April until a clutch of eggs was discovered. Following egg discovery, we monitored those nest boxes daily to ensure the hatch day was recorded. Western Bluebirds generally lay one egg per day, averaging five eggs per clutch (Brawn 1991) and do not begin incubation until all eggs are laid (Guinan et al. 2020). They can display asynchronous hatching when looking at a much finer timescale (Ferree et al. 2010); however, in our study we did not have the ability to monitor the nest boxes hourly to determine the exact hatch order and we assumed all eggs hatched within a 24 h period.

Beginning with the hatch day (day 0) through fledging, we collected morphometric data daily for each nestling between 10:00 and 12:00 to ensure that a constant interval of development was recorded for each day. Two researchers were present for all measurements over all 3 years to limit observer bias. We weighed each nestling using a digital scale to the nearest tenth of a gram. Tarsus length and culmen length were also measured daily. We used digital calipers to measure the length of the tarsus and culmen to the nearest hundredth of a millimeter. Tarsus was measured from the tibiotarsal joint to the distal end of the last leg scale (Jongsomjit et al. 2007).

We used two different methods for measuring culmen. In 2015, we measured the total culmen length, which is a measurement from the tip of the bill to the edge of the skull (Baldwin et al. 1931; Winker 1998). In 2016 and 2017, we measured the culmen from the tip of the bill to the distal side of the nares (Baldwin et al. 1931; Borras et al. 2000). The different culmen measurements were a result of two different field crews interpreting methods differently.

Statistical analysis

All statistical analyses were performed using R version 3.5.1 (R Core Team 2018). For all three morphometric growth measurements (mass, tarsus, and culmen), we developed models for age prediction using linear mixed models (LMM), model training, and model validation. We used the 'lme4' package (Bates et al. 2015) for the LMMs. P-values presented for the mixed models were derived from Satterthwaite's degrees of freedom method using the 'lmerTest' package (Kuznetsova et al. 2017). The 95% confidence intervals are presented for each coefficient and were calculated using the confint function. We used the package 'caret' (Kuhn 2020) for splitting data that were used for model training and validation. Data were plotted using the 'ggplot2' package (Wickham 2009). The data analysis procedure is described below.

First, we plotted each measurement variable (mass, tarsus, and culmen) against nestling age to determine the age ranges for which predictions would be the most accurate and informative. We visually assessed the data in those plots to determine the age ranges selected for our analyses. Mass and tarsus were not as good of a predictor of age after 10 days as they were from 0 to 10 days. This was due to increased variation after day 10 (see Fig. 1a and Fig. 1b). Therefore, mass and tarsus were only used to predict nestling age up to day 10.

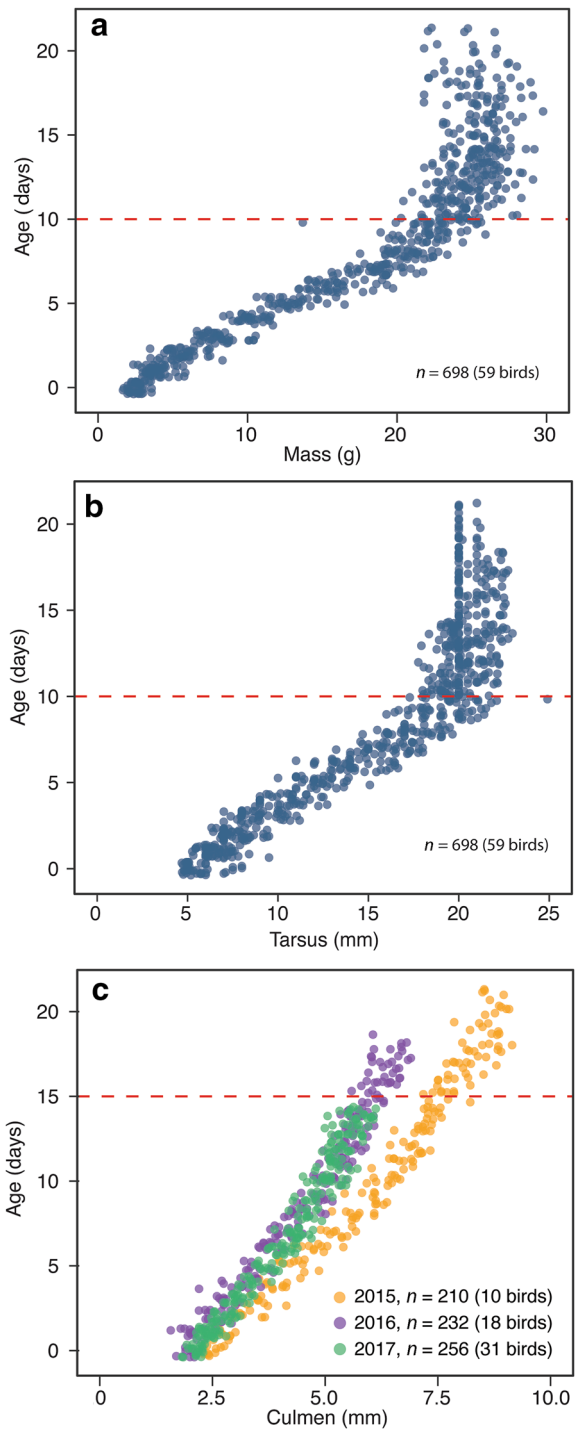

Fig. 1 Initial mass (**a**), tarsus (**b**), and culmen (**c**) measurements collected from nestlings for 0–21 days of age. For final models, mass (**a**) and tarsus (**b**) were analyzed from 0 to 10 days because variation increased after day 10 and predictive power of models was reduced. Culmen (**c**) was analyzed from 0 to 15 days because variation increased after day 15 for 2015 and data was not collected passed day 15 in 2017. Red dashed lines show the age below which data were used for model development. In all three panels, points are jittered to see the spread of the data

As expected, the two different culmen measurements were distinct from each other and could not be used in the same predictive model (Fig. 1c). Specifically, 2015 is different from both 2016 and 2017 because of the two different measurement methods, as described previously. Therefore, we completed the model training and validation separately for these two culmen measurements. We also determined that culmen was accurate through day 15. For 2015, age becomes slightly harder to predict after nestlings become 15 days old (greater amount of variation between culmen and age by visually assessing the data). Initial analyses also suggested that models were much less accurate if the full nestling period (0–21 days old) was included. Additionally, there are no data for 2017 after day 15 (Fig. 1c). Therefore, culmen was only used to predict nestling age up to day 15 for these 2 years as well.

Mass, tarsus, and culmen were highly correlated (Pearson; mass-tarsus: $r = 0.98$, $p < 0.001$; $n = 698$; mass-culmen: $r = 0.87$, $p < 0.001$; $n = 698$; tarsus-culmen: $r = 0.85$, $p < 0.001$; $n = 698$) due to all three measurements increasing as nestlings got older. Due to issues regarding multicollinearity in a multiple regression model, we created separate models for each of the measurement variables. We combined all 3 years for mass and tarsus in an effort to obtain models that were more general (i.e., did not over fit the data) for a given year. Again, culmen was separated into two data sets (2015 and 2016–2017), which gave us four measurement variables for the modeling procedure described next.

We created four LMMs with nestling ID as the random effect to account for repeated measures of the same nestling over the course of the nestling period. The fixed effect in each model was the measurement variable (mass, tarsus, culmen 2015, and culmen 2016–2017). Models were fit using restricted maximum likelihood. The following model training and validation was applied to each measurement variable. We randomly split the data into two separate datasets: a model training dataset containing 90% of the original data and a model test dataset containing 10% of the original data. Based on our samples sizes, we wanted to make sure that enough data were being used in model training to reduce the error surrounding the estimates as much as possible, but still have enough points to test the models.

The mixed model, with the structure described above, was created from the training dataset using the lmer function in the 'lme4' package. Since some relationships appeared curvilinear when graphed, we made two models for each variable: linear and quadratic (eight models total). For each model, the test dataset was used to validate the model derived from the training dataset by making predictions from the mixed models. We used

the 'predict' function in R to predict the age (in days) of nestlings using the individual nestling measurement. This resulted in a predicted age and an actual age (from the test data) for a given nestling. For each model, we calculated the root mean square error (RMSE). RMSE is reported in days and is a measure of how the predicted ages differed, on average, from the actual ages. We chose between the linear and quadratic models using RMSE, since this is a measure of how well the model predicts nestling age. If there was a substantial drop in RMSE from the linear to quadratic model (>10% decrease), then the quadratic model was used as the final model from which nestling age predictions were based. These final models were plotted with the best fit line, the actual ages (test data), and the predicted ages. Equations were obtained from the model coefficients and the *y*-intercept. The R^2 values were obtained using the r.squaredGLMM function in the 'MuMIn' package (Barton 2019).

The code and documentation for data processing, model training, and model validation can be found on GitHub (https://github.com/ChanLand/nestling_age_model). This code can be used to conduct these analyses for nestlings of any other bird species, provided there are sufficient data available on known ages and the corresponding measurement(s) of interest.

Results

The model selection process (here, RMSE) to determine whether the linear or quadratic model should be used for predicting nestling age are presented for each development variable (Table 1). For mass and tarsus, the quadratic model did not improve our predictive power based on RMSE values (Table 1). Therefore, we chose the linear mixed model as the final model used to make the predictions from mass and tarsus. Mass had a RMSE of 0.43, which means that mass is accurate, on average, to within a half day. Tarsus is accurate to within 0.68 days. For both culmen measurements, RMSE decreased from the linear to the quadratic model (Table 1). Because the predictive power got better, we chose the quadratic model as the final model. Culmen 2015 is accurate to within 0.59 days and culmen 2016–2017 is accurate to within 0.70 days.

Table 1 Root mean square error (RMSE) for the linear and quadratic mixed models

Variable	Power	RMSE
Mass	Linear[a]	0.43
	Quadratic	0.42
Tarsus	Linear[a]	0.68
	Quadratic	0.69
Total culmen	Linear	0.72
	Quadratic[a]	0.59
Culmen from tip of nares	Linear	0.86
	Quadratic[a]	0.70

RMSE is a measure of the difference between predicted age and actual age during model validation

[a] Denotes the final model that was chosen

The model validation is shown graphically in Fig. 2. The full dataset used is shown for mass (Fig. 2a), tarsus (Fig. 2b), culmen 2015 (Fig. 2c), and culmen 2016–2017 (Fig. 2d). In each panel, the full dataset is split into the training data (gray points) and the test data (purple points). The green points are the predicted nestling ages from the final mixed models (described above) using the nestling measurement variables from the test data. Each green point is connected to the corresponding purple point (actual nestling age) by a vertical black line. The average difference between the predicted age and the actual age is measured by the RMSE (Table 1).

The output for the final model (linear for mass and tarsus, quadratic for both culmen measurements) is presented along with an equation that was created from the model coefficients (Table 2). These equations can be used with measurement data for each variable to estimate the age of unknown nestlings. The Additional file 1: Table S1 contains the mean, minimum, and maximum values for the mass, tarsus, and both culmen measurements for each nestling age from 0 to 21 days old.

Discussion

The most accurate nestling age would result from checking a nest daily until eggs hatch to determine a hatch date, however that is not always feasible. Checking a nest daily is time consuming, expensive, can cause stress on the birds, and can lead to increased predation (Wails et al. 2014). In this study, we developed a set of field tools designed to improve age estimates of Western Bluebird nestlings for use by researchers of varying levels of experience without having to determine the exact hatch date. We transformed 3 years of morphometric field data into predictive models for estimating nestling age. All of the measurement variables we used to develop our models are accurate to within less than a day: mass is accurate to within 0.43 days, tarsus is accurate to within 0.68 days, culmen 2015 is accurate to within 0.59 days, and culmen 2016–2017 is accurate to within 0.70 days. Our high accuracy species-specific guide is a resource for ensuring quality data and research on the Western Bluebird. The novelty of our study is that the model training and validation methods used here can be applied to any other passerine species provided enough accurate daily growth data are collected to develop robust growth models. In addition, assigning process-based age estimates ensures all field crews in a given study and across different studies

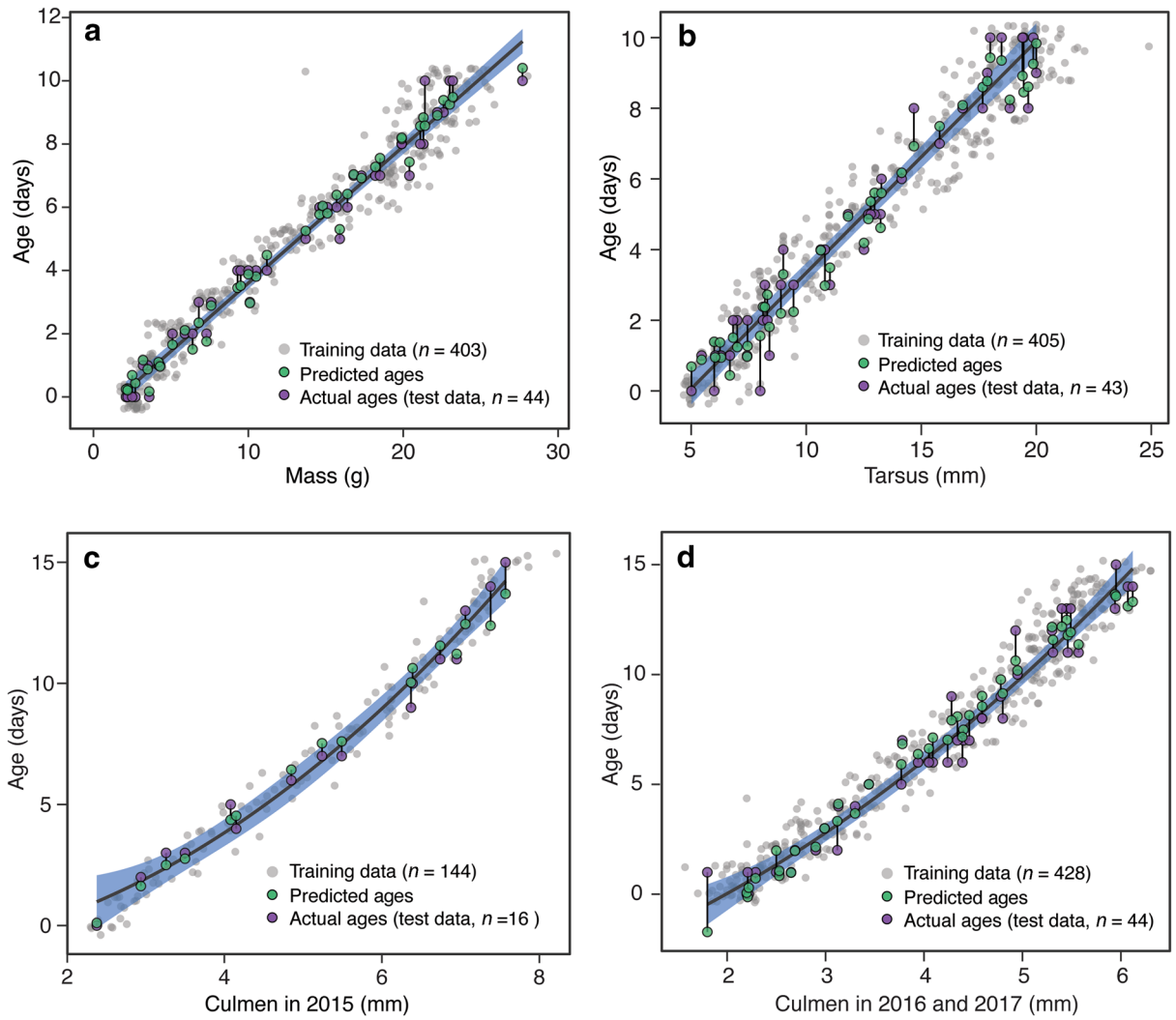

Fig. 2 Final model based on training data and model validation results for mass (**a**), tarsus (**b**), culmen 2015 (**c**), and culmen 2016–2017 (**d**). For each panel, gray points are the training data used for model training. Purple points are the test data (actual ages), while the green points are the ages predicted from the final model. The predicted age from the model is connected to the actual age by the vertical black lines. The black lines are the regression lines and are linear for mass and tarsus and quadratic for both culmen measurements. Shaded blue areas represent the 95% confidence intervals. In all four panels, points are jittered to see the spread of the data

are aging birds exactly the same way and it reduces the subjectivity of the estimates. This results in more robust comparisons between studies of a specific species.

Growth curves of most passerines follow a logistic pattern characterized by an inflection point that occurs when 50% of the asymptote is attained (Pinkowski 1975; Starck and Ricklefs 1998). Traditional growth models for avian species are presented with age as the independent variable (Holcomb and Twiest 1971; Lyons and Mosher 1983; Bortolotti 1984a; Bechard 1985; Carlsson and Hörnfeldt 1994; Rodway 1997) to demonstrate the logistical growth pattern. The predictive models we developed consider age the dependent variable determined by morphometric field measurements, as age is the variable we are predicting. By first graphing the data set, we were able to visually assess the accuracy range for each morphometric measurement. Model validation provided metrics of accuracy for each measurement variable.

Young nestlings do not display some obvious physical characteristics for aging (e.g., feather tract development). Therefore, growth models can aid in capturing early life stage data. Reliable methods for estimating the age of nestlings is important for investigating ecological and behavioral aspects of a species (Amiot et al. 2014).

Table 2 Linear mixed model results for each morphometric measurement that can be used to estimate age of Western Bluebird nestlings

Developmental measurement	Model output	Equation for estimating age	95% CI of coefficients	*t*-value (df)	Valid range (days)
Mass (g)	$R^2 = 0.96$, $P < 0.001$	$\text{Age} = 0.43x - 0.68$	0.424–0.439	114.3 (372.8)	0–10
Tarsus length (mm)	$R^2 = 0.95$, $P < 0.001$	$\text{Age} = 0.64x - 3.1$	0.625–0.648	112.7 (369.1)	0–10
Total culmen (mm)	$R^2 = 0.97$, $P < 0.001$	$\text{Age} = 0.14x^2 + 1.33x - 3.67$	x^2: 0.102–0.173	x^2: 7.58 (132.3)	0–15
			x: 0.962–1.69	x: 7.13 (132.3)	
Culmen from tip of nares (mm)	$R^2 = 0.96$, $P < 0.001$	$\text{Age} = 0.26x^2 + 1.47x - 3.82$	x^2: 0.205–0.318	x^2: 9.12 (402.8)	0–15
			x: 1.02–1.93	x: 6.39 (406.2)	

Final equations were obtained from the model coefficients. Mass and tarsus measurements should only be used to estimate ages 10 days old and younger. Both culmen measurement techniques can be used to estimate age to day 15. For the quadratic models, the 95% CIs, *t*-values, and degrees of freedom (df) are listed for each of the coefficients (x^2 and x) in the equation

CI confidence interval

The reliability of our research is dependent upon normal variation in individuals and may vary between growth parameters. In addition, factors such as asynchronous hatching on a finer timescale could add some variation between nestlings. Western Bluebirds are facultative cooperative breeders, with a small proportion of breeding pairs having helper males at the nest (Dickinson et al. 1996; Duckworth and Badyaev 2007; Potticary and Duckworth 2018). We did not have any way to identify that in this study, but it should be a consideration for future work, especially for other species using this tactic more often such as the Acorn Woodpecker (*Melanerpes formicivorus*) (Koenig et al. 2011). Single morphometric measurements can be used to predict nestling age with high accuracy. However, the combination of two or more measurements may increase the accuracy of age prediction.

The two different culmen measurements show the importance of identifying which technique should be used and training of field crews before collecting measurements to estimate age. Misinterpretation of measurement techniques by different field crews can result in inaccurate age predications. We provided both here so that both culmen measurement techniques can be used to predict nestling age. However, we emphasize that researchers should pick one technique for a given field season to avoid erroneous measurements and age estimations and to make sure the method used is documented.

If researchers plan to use our data for other Western Bluebird populations, we recommend that they test the models developed in this study with a few nestlings of known ages before relying exclusively on our models. It is known that passerine nestlings can exhibit variation in developmental patterns both within and among different geographic locations (Ardia 2006). The data ranges in the Additional file 1: Table S1 should be used to aid in age determinations in the field. If researchers plan to use these techniques to develop their own models for other species, we recommend gathering morphometric measurement data on individual nestlings for which the hatch date is known or can be inferred from frequent visits to the nest. Measurements should be taken on at least 30 nestlings. Ideally, this should be done over at least two field seasons to account for potential variation between years. Variation can occur from multiple factors including but not limited to, environmental influences, researcher bias, sibling competition, hatch order if eggs are hatching asynchronously, and whether or not there are helpers assisting with nest provisioning. When taking measurements, nestlings should be measured every day from day 0 to when they fledge. A greater number of nestlings in the dataset increases the accuracy of the models in predicting ages from the morphometric measurements taken. Once the morphometric data are collected, researchers should use the R code and documentation for data processing, model training, and model validation on GitHub referenced in the methods.

Conclusions

All measurement variables were accurate to within less than a day and two were accurate to within half a day. We discourage the use of the equations beyond the valid age ranges shown in Table 2. The models for calculating age from tarsus and mass are only valid for nestlings 10 days old and younger, as these metrics begin to plateau at this point, and thus observed changes are most likely due to observer bias (especially for tarsus), and variation in parental feeding success (mass). Similar patterns are seen in the culmen measurements, which make them less reliable at estimating nestling age greater than 15 days old. Once a given measurement is not reliable, the range of measurements for 0–21 days of age provided (Additional

file 1: Table S1) will aid in narrowing down the age estimation for older nestlings.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40657-021-00258-5.

Additional file 1: Table S1. The mean, minimum, and maximum values for the mass, tarsus, and both culmen measurements for each nestling age from 0 to 21 days old.

Acknowledgements
The authors would like to thank M.M. Salazar and J.T. Berryhill for their assistance in fieldwork and data management.

Authors' contributions
CDH conceived the study with design contributions from BET. AuAS led the manuscript writing and assisted with data collection and analysis. AWB and AMC led the data analysis and developed the model training and validation methods. JMF and CDH were the project managers and assisted with manuscript editing. AaAS, KH, MAM, EMP, and BET led fieldwork and data collection. All authors read and approved the final manuscript.

Funding
All authors and this work were supported by Los Alamos National Laboratory, managed by Triad National Security, LLC, which receives its core funding from the U.S. Department of Energy's National Nuclear Security Administration. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials
The datasets used in the present study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
This work was performed under an approved Institutional Animal Care and Use Committee protocol.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Environmental Stewardship, Los Alamos National Laboratory, Los Alamos, NM, USA. [2]Biosecurity and Public Health, Los Alamos National Laboratory, Los Alamos, NM, USA. [3]Water Program, N3B, Los Alamos, NM, USA.

Received: 27 October 2020 Accepted: 8 May 2021
Published online: 02 June 2021

References

Amiot C, Ji W, Hill SD. Using plumage and behavioral development to age New Zealand Fantail nestlings. N Z J Zool. 2014;42:35–43.
Ardia DR. Geographic variation in the trade-off between nestling growth rate and body condition in the Tree Swallow. Condor. 2006;108:601–11.
Baldwin SP, Oberholser HC, Worley L. Measurements of birds. Cleveland: Cleveland Museum of Natural History; 1931.
Barton K. MuMIn: multi-model inference. R package version 1.43.6; 2019. https://cran.r-project.org/web/packages/MuMIn/index.html. Accessed 17 Feb 2020.
Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015;67:1–48.
Bechard MJ, Zoellick BW, Nickerson M. Accuracy in determining the age of nestling red-tailed Hawks. J Wildl Manage. 1985;49:226–8.
Borras A, Pascual J, Senar JC. What do different bill measures measure and what is the best method to use in granivorous birds? J Field Ornithol. 2000;71:606–11.
Bortolotti GR. Physical development of nestling Bald Eagles with emphasis on the timing of growth events. Wilson Bull. 1984a;96:524–42.
Bortolotti GR. Criteria for determining age and sex of nestling Bald Eagles. J Field Ornithol. 1984b;55:467–81.
Brawn JD. Environmental effects on variation and covariation in reproductive traits of western Bluebirds. Oecologia. 1991;86:193–201.
Brawn JD, Balda RP. Population biology of cavity nesters in northern Arizona: do nest sites limit breeding densities? Condor. 1988;90:61–71.
Brown WP, Alexander AL, Alexander DA, Zuefle ME, Underwood TJ. Estimating ages of House Wren nestlings based on body mass, wing chord length, and feather tract development patterns. N Am Bird Bander. 2011;36:101–10.
Brown WP, Zuefle ME, Underwood TJ, Alexander AL, Alexander DA. House Wren nestling age can be determined accurately from a guide of digital images. N Am Bird Bander. 2013;38:150–9.
Bryant DM. Environmental influences on growth and survival of nestling House Martins (*Delichon urbica*). Ibis. 1978;120:271–83.
Carlsson BG, Hörnfeldt B. Determination of nestling age and laying date in Tengmalm's Owl: use of wing length and body mass. Condor. 1994;96:555–9.
Costa JS, Rocha AD, Correia R, Alves JA. Developing and validating a nestling photographic aging guide for cavity-nestling birds: an example with the European Bee-eater (*Merops apiaster*). Avian Res. 2020;11:2.
Dickinson JL, Leonard ML. Mate attendance and copulatory behaviour in western Bluebirds: evidence of mate guarding. Anim Behav. 1996;52:981–92.
Dickinson JL, Koenig WD, Pitelka FA. Fitness consequences of helping behavior in the western Bluebird. Behav Ecol. 1996;7:168–77.
Dinsmore SJ, White GC, Knopf FL. Advance techniques for modeling avian nest survival. Ecology. 2002;83:3476–88.
Duckworth RA, Badyaev AV. Coupling of dispersal and aggression facilitates the rapid range extension of a passerine bird. Proc Natl Acad Sci. 2007;104:15017–22.
Fair JM, Myers OB. Early reproductive success of western Bluebirds and Ash-throated Flycatchers: a landscape-contaminant perspective. Environ Pollut. 2002;118:321–30.
Fair JM, Hanelt B, Burnett K. Horsehair worms (*Gordius robustus*) in nests of the western Bluebird (*Sialia mexicana*): evidence for anti-predator avoidance? J Parasitol. 2010;96:429–30.
Fernaz JM, Schifferli L, Grüebler MU. Ageing nestling Barn Swallows *Hirundo rustica*: an illustrated guide and cautionary comments. Ringing Migr. 2012;27:65–75.
Ferree ED, Dickinson J, Rendell W, Stern C, Porter S. Hatching order explains an extrapair chick advantage in western Bluebirds. Behav Ecol. 2010;21:802–7.
Graham H. Monitoring guide monitoring your bluebird trail in California. Somerset: Hot Pepper Press; 2006.
Gilliland SG, Ankney CD. Estimating age of young birds with a multivariate measure of body size. Auk. 1992;109:444–50.
Guinan JA, Gowaty PA, Eltzroth EK. Western Bluebird (*Sialia mexicana*), version 1.0. In: Poole AF, editor. Birds of the world. Ithaca: Cornell Lab of Ornithology; 2020.
Haggerty TM. Nestling growth and development in Bachman's Sparrows. J Field Ornithol. 1994;65:224–31.
Holcomb LC, Twiest G. Growth and calculation of age for red-winged Blackbird nestlings. Bird-Banding. 1971;42:1–17.
Jongsomjit D, Jones SL, Gardali T, Geupel GR, Gouse PJ. A guide to nestling development and aging in altricial passerines. Biological Technical Publication BTP-R6008-2007. Washington, DC, US Department of Interior, Fish and Wildlife Service; 2007.
Keyser AJ, Keyser MT, Promislow DEL. Life-history variation and demography in western Bluebirds (*Sialia mexicana*) in Oregon. Auk. 2004;121:118–33.

Koenig WD, Walters EL, Haydock J. Variable helper effects, ecological conditions, and the evolution of cooperative breeding in the Acorn Woodpecker. Am Nat. 2011;178:145–58.
Kozma JM, Kroll AJ. Nest survival of western Bluebirds using tree cavities in managed ponderosa pine forests of central Washington. Condor. 2010;112:87–95.
Kuhn M. Caret: Classification and Regression Training. R package version 6.0–85. 2020. https://CRAN.R-project.org/package=caret. Accessed 17 Feb 2020.
Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. J Stat Softw. 2017;82:1–26.
Langham NPE. Chick survival in terns (*Sterna* spp.) with particular reference to the Common Tern. J Anim Ecol. 1972;41:385–95.
Lyons DM, Mosher JA. Age-estimation model for nestling broad-winged hawks. Wild Soc Bull. 1983;11:268–70.
Murphy MT. Growth and aging of nestling eastern Kingbirds and eastern Phoebes. J Field Ornithol. 1981;52:309–16.
Musgrave K, Bartlow AW, Fair JM. Long-term variation in environmental conditions influences host–parasite fitness. Ecol Evol. 2019;9:7688–703.
O'Connor RJ. Growth strategies in nestling passerines. Living Bird. 1978;16:209–38.
Palacios E, Anderson DW. Age determination in California brown pelican chicks (Pelecanus occidentalis californicus) chicks in the Gulf of California. Waterbirds. 2018;41:305–9.
Partridge L, Harvey PH. The ecological context of life history evolution. Science. 1988;241:1449–55.
Pinkowski BC. Growth and development of eastern Bluebirds. Bird-Banding. 1975;46:273–89.
Podlesak DW, Blem CR. Determination of age of nestling Prothonotary Warblers. J Field Ornithol. 2002;73:33–8.
Potticary AL, Duckworth RA. Environmental mismatch results in emergence of cooperative behavior in a passerine bird. Evol Ecol. 2018;32:215–29.
R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2018.
Rodway MS. Relationship between wing length and body mass in Atlantic Puffin chicks. J Field Ornithol. 1997;68:338–47.
Shaffer TL. A unified approach to analyzing nest success. Auk. 2004;121:526–40.
Starck JM, Ricklefs RE, editors. Avian growth and development: evolution within the altricial-precocial spectrum. Oxford: Oxford University Press; 1998.
Wails CN, Oswald SA, Arnold JM. Are morphometrics sufficient for estimating age of pre-fledging birds in the field? A test using Common Terns (*Sterna hirundo*). PLoS ONE. 2014;9:e111987.
Wang JM, Weathers WW. Egg laying, egg temperature, attentiveness, and incubation in the western Bluebird. Wilson J Ornithol. 2009;121:512–20.
Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer; 2009.
Wilkins N, Brown WP. The accuracy of eastern Bluebird (*Sialia sialis*) nestling age estimates produced from three different aging guides of digital images. N Am Bird Bander. 2015;40:1–11.
Winker K. Suggestions for measuring external characters of birds. Ornitol Neotrop. 1998;9:23–30.
Wysner TE, Bartlow AW, Hathcock CD, Fair JM. Long-term phenology of two North American secondary cavity-nesters in response to changing climate conditions. Sci Nat. 2019;106:54.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions