J Forensic Sci, July 2008, Vol. 53, No. 4
doi: 10.1111/j.1556-4029.2008.00744.x
Available online at: www.blackwell-synergy.com

Aaron M. Tarone,[1,†] *Ph.D. and David R. Foran,*[2] *Ph.D.*

Generalized Additive Models and *Lucilia sericata* Growth: Assessing Confidence Intervals and Error Rates in Forensic Entomology*

ABSTRACT: Forensic entomologists use blow fly development to estimate a postmortem interval. Although accurate, fly age estimates can be imprecise for older developmental stages and no standard means of assigning confidence intervals exists. Presented here is a method for modeling growth of the forensically important blow fly *Lucilia sericata*, using generalized additive models (GAMs). Eighteen GAMs were created to predict the extent of juvenile fly development, encompassing developmental stage, length, weight, strain, and temperature data, collected from 2559 individuals. All measures were informative, explaining up to 92.6% of the deviance in the data, though strain and temperature exerted negligible influences. Predictions made with an independent data set allowed for a subsequent examination of error. Estimates using length and developmental stage were within 5% of true development percent during the feeding portion of the larval life cycle, while predictions for postfeeding third instars were less precise, but within expected error.

KEYWORDS: forensic science, blow fly, development, Calliphoridae, confidence interval, postmortem interval, generalized additive model, forensic entomology

Daubert, et al. v. Merrell Dow Pharmaceuticals (509 U.S. 579 [1993]) was a pivotal ruling for forensic scientists, in which the U.S. Supreme Court declared that the Federal Rules of Evidence (particularly Rule 702), and not Frye (*Frye v. U.S.A.* (293 F. 1013, 1014, D.C. Cir. [1923]), were the standard for scientific evidence and expert testimony. In doing so, the High Court placed the burden of assessing the validity—and thus admissibility—of scientific evidence on the trial judge, based on five main criteria: Has the technique in question been tested; do standard operating procedures (SOPs) exist for the technique; has the technique been subjected to peer review and publication in the appropriate literature; is the technique widely accepted by the relevant scientific community; and finally, what is the known or potential error rate of the technique? DNA-based evidence has set the "gold standard" for meeting Daubert requirements, largely satisfying all of them. In contrast, many of the forensic sciences and resultant expert testimony are based on practitioners' training and experience, often with little consideration for SOPs, method testing, potential error rates, or publication, even when the technique is generally accepted. As an example, the National Institute of Justice recently posted a solicitation for the study of fingerprints/friction ridges, although certainly this method of identification is extremely well established. Other areas of forensic science fare far worse (1).

Forensic entomology falls between these extremes. The predictable growth of carrion-feeding flies has long been used to estimate the time a body has been exposed to insects, and thus to estimate a postmortem interval (PMI). Using larval size and developmental stage to approximate age is well supported by research and observations in developmental biology, and this forensic technique is widely described in the scientific literature (e.g., 2,3). Likewise, countless legal rulings have assured its admissibility, just as countless juries have been guided by entomological testimony. However, scientists have reported different growth rates for immature flies (2–6) and court qualified experts have come to incongruent conclusions about a PMI based on the same entomological evidence, depending on which growth data were utilized (e.g., *California v. Westerfield*, CD 165805 [2002]). This problem stems, at least in part, from a general failure to develop SOPs, and also from not fully considering the amount of variation present in larval growth (or more precisely, to account for error rates inherent in estimates of larval age), two of the major tenets of Daubert. The difficulty in estimating error is exacerbated by the fact that blow flies grow in a nonlinear fashion and have variable size distributions at different ages, unequally affecting age estimates of developmental stages (7).

The research presented here was designed to investigate the variability that occurs in larval and pupal growth of blow flies in order to discern which of a suite of variables have the largest influence on estimating age, and to explore the possibility of placing confidence intervals around juvenile age estimates. Using three regional strains of the blow fly *Lucilia sericata* (Diptera: Calliphoridae) (Meigen), collected in California (CA), Michigan (MI), and West Virginia (WV), a data set containing linear (developmental stage, strain, rearing temperature) and nonlinear (length and weight) measures was established. Generalized additive models (GAMs) were developed taking these variables into account, examining the level to which each influenced/predicted the percent of immature fly development (8,9). Similar GAMs have already been used to assess the effects of

[1]Department of Zoology, Michigan State University, East Lansing, MI 48824.
[2]Forensic Science Program, School of Criminal Justice and Department of Zoology, Michigan State University, East Lansing, MI 48824.
*A portion of this work was presented at the American Academy of Forensic Scientists Annual Meeting in Seattle, WA, February 2006.
†Present address: Molecular and Computational Biology Program, University of Southern California, Los Angeles, CA 90089.
Received 4 Nov. 2006; and in revised form 11 Jan. 2008; accepted 27 Jan. 2008.

cadmium on the growth of *L. sericata* cohorts (10), and were assessed here for their potential use as tools in predicting blow fly development percent. The utility of a model was then tested on an independent data set (larvae reared on rat carcasses), focusing on developmental stage and length. GAM predictions of larval development percent were plotted against true age to assess the error of the predictions and to define confidence intervals for these estimates.

Materials and Methods

Species Identification

Wild *L. sericata* were collected in CA, MI, and WV, from the UC Davis campus in June of 2005, the Michigan State University campus starting in May 2005 (which were provisioned with new flies occasionally throughout the summer), and from the West Virginia University campus in August of 2005. Adult individuals from each strain were identified by key (11,12), with independent confirmations, and through mitochondrial cytochrome oxidase 1 gene sequencing (13).

Growth Experiments

Cohorts of flies were raised in a round robin design, in which CA and MI were reared in one block, followed by CA and WV, and WV and MI, between September 1, 2005 and October 24, 2005. Flies ranged from two to five generations removed from their natural population. Cohorts were initiated by placing fresh liver into the cages of adult flies, which was checked regularly for eggs. When oviposition occurred, the time was recorded and meat and eggs were removed 1 h later. Cohorts were placed in either 20 ± 0.5°C or 33.5 ± 1.8°C incubators under a 12:12-h light cycle at 25 ± 5% relative humidity. Incubator temperature fluctuation was noted using a HOBO data logger (Onset Computer, Bourne, MA). Eggs were transferred to fresh liver, which was placed on a moist paper towel in 1-L jars, covered with a breathable fabric lid, based on rearing conditions previously found to best mimic those on carrion (13). Cohorts were given fresh liver daily until postfeeding third instars were observed, at which point 250 individuals (33.5°C treatments) and 375 individuals (20°C treatments) were transferred in batches of 125- to 1-L jars containing 500 mL of fresh sand as a pupation substrate.

Length and weight of 2559 larvae/pupae were recorded, starting *c*. 24 h after eggs were laid. Length was measured with a ruler based on the furthest extension of a larva to the nearest 0.5 mm. Wet weight of live individuals was measured on a Cahn 27 Automatic Electrobalance (Cahn Instruments, Cerritos, CA) to the closest 1/100 mg. Developmental stage was assessed by observing feeding larvae microscopically, by visible crop length and migrating behavior for postfeeding larvae, and puparium formation for pupae. Ten larvae were removed from a cohort and measured/weighed, twice daily (in the morning and late afternoon). Ten pupae were collected once daily and measured/weighed; five individuals were collected if less than 10 were available.

Earlier research showed that the destructive sampling of pupae delayed the appearance of adults (13). To account for this, pupal age was calibrated to the day of pupation. This means that pupal samples were assessed in groups that pupated within 24 h of each other (i.e., 0–1, 1–2, 2–3, etc. -day-old puparia) with the minimum development time for pupation being the minimum development time for any individual within a collective group of pupae.

Forensic entomologists generally assess fly growth progression using a measure of relative age, allowing them to take into account the substantial influence of temperature on development. Given that multiple variables had the potential to affect immature fly growth rates in the current research, including understood (e.g., temperature) and questioned (e.g., fly strain) factors, a method that would allow growth progression to be compared directly among all flies was required. Development percent, or the relative (developmental) age of an individual, was used to assess the extent to which a fly had progressed towards maturation (eclosion). This measure, often used for relative developmental comparisons (e.g., 6,14,15), permitted individuals at all points in development to be compared, which would be impossible if, for instance, temperature and fly strain varied in their influence on growth. Development percent was calculated by determining the age in hours of an individual, then dividing the age by the minimum total development time of that experimental replicate. As an example, if an individual was sampled 100 h after oviposition and the minimum development time for the replicate was 285 h, then the individual was considered 35% developed.

The laboratory growth of larvae on rats has been described previously (13) and differed from the measured cohorts primarily in food source and temperature (25°C). Three cohorts of Michigan *L. sericata* larvae were reared on rat carcasses, and the developmental stage and length of 12 individuals were recorded daily from each cohort through the first day that puparia were observed. These data were used to predict age. The ethical guidelines of the Michigan State University Laboratory Animal Resources unit were followed, adhering to IACUC requirements.

Statistical Analyses

GAMs were developed using the mgcv library in the R statistical package (16). The models use likelihood statistics to predict a value (e.g., age) based on various input data. GAMs relate nonlinear data such as fly length and weight to the predicted value (e.g., development percent) using smoothed, nonlinear mathematical functions (8–10). In this manner, the relationship of two nonlinear variables to each other can also be included in GAMs (9), so a length-by-weight term was also evaluated. Distributions must be applied to the functions used to make predictions in a GAM, which is performed through a link function. Based on the results of residual plots produced for the models, a gamma distribution (instead of a normal distribution) with a log-link function was most appropriate for the models evaluated. More detailed information on GAM can be found in Refs. (8–10,17).

Models generated several statistics. For linear models, the statistic used to explain how closely data match a model is R^2; as length and weight data are nonlinear, the apposite statistic for GAMs is the percent deviance explained (9). Degrees of freedom or estimated degrees of freedom (a nonlinear equivalent) were determined, as was a *p*-value, which was based on the likelihood of a variable being predictive of age. *p*-values in GAMs are considered estimates because likelihood statistics do not yield actual *p*-values, but do provide values that are similar and can be used to estimate the more familiar statistic. These estimates can vary by up to two times the actual *p*-value (9), thus terms were not considered significant unless *p*-values were <0.025. Additionally, multiple variables were included in some models, requiring a Bonferroni correction that resulted in a significance threshold of $p < 0.0042$. Given the inherent inaccuracy of estimated *p*-values, they were only used to identify informative terms or terms that were candidates for removal from a model owing to intermediate or nonsignificant *p*-values. The inclusion or removal of a term, however, was ultimately decided by the statistic used to compare models: the generalized cross-validation (GCV) score, which is an information criterion that is lower for better models (9).

Six terms were used to develop models: fly developmental stage, length, weight, length-by-weight, strain, and temperature. Stage, strain, and temperature were considered linear variables, and length, weight, and the two plotted against each other were nonlinear. This resulted in 63 possible models, hence only a subset is presented here. The first six models examined each variable by itself, while the remaining 12 combined variables to assess improvements gained (as measured by a decrease in GCV) from including specific terms. Developmental stage was considered the primary variable, as all forensic entomologists include this in PMI predictions. Body size is also often incorporated into PMI estimates, thus length and weight were added to several models, as well as being examined in combination. Next, the influences of strain and temperature were tested through inclusion with the more familiar variables (stage, length, weight). Similarly, as length-by-weight is a somewhat novel measure, it was evaluated in combination with the three standard variables, and then with all variables.

Each model also provided estimates of error based on plots of true (response) versus predicted (fitted) values for data used to construct the model. For simplicity's sake this will be referred to as the $Y = X$ line, or Y (predicted age) = X (true age). The most precise models have all predicted values clustered close to the $Y = X$ line, with no gaps in the line, given that flies were examined continuously throughout development. A gap in predictions results in inaccuracy because an individual of an age found in a gap will be identified as either older or younger than it actually is.

Finally, a GAM incorporating the standard variables used to age flies in forensic entomological enquiries, developmental stage and length (2–4,6), was tested against an independently derived data set. The model-based predictions of larval development percent for three previously collected fly cohorts raised on rats were plotted against their true development, comparing them to the predicted 95% confidence intervals for the model (precision) and the $Y = X$ line (accuracy). Confidence intervals were superimposed over the predictions made for rat cohorts (using the quantreg library in R) by plotting locally weighted sum of squares curves through the 97.5th and 2.5th percentiles.

Results

Species Identification

Flies collected from the three states were identified as *L. sericata* based on both visual verification, visual confirmation by an independent entomologist, and cytochrome oxidase 1 sequence data (accession numbers DQ868503, DQ868523, and DQ868524 for CA, MI, and WV, respectively). Sequences obtained from the CA, MI, and WV strains were 428 and 227 nonoverlapping base pairs, 774 continuous base pairs, and 776 continuous base pairs in length, respectively. BLAST results for the sequences showed the closest match for all was to *L. sericata*, with 100% similarity to at least one other *L. sericata* sequence. The next closest species match was *L. cuprina* with a 98–99% similarity (5–8 base pairs difference).

Immature Development

Figure 1 depicts a plot of fly length against percent juvenile development. The feeding portion of the life cycle makes up the initial 25%, and shows a linear increase in length. The postfeeding third instar, where body size decreases and variation in size increases, is found from *c*. 25–50%. The relatively unchanged second half of the plots is the pupal stage. Weight results displayed the same pattern (data not shown), and both demonstrated that the

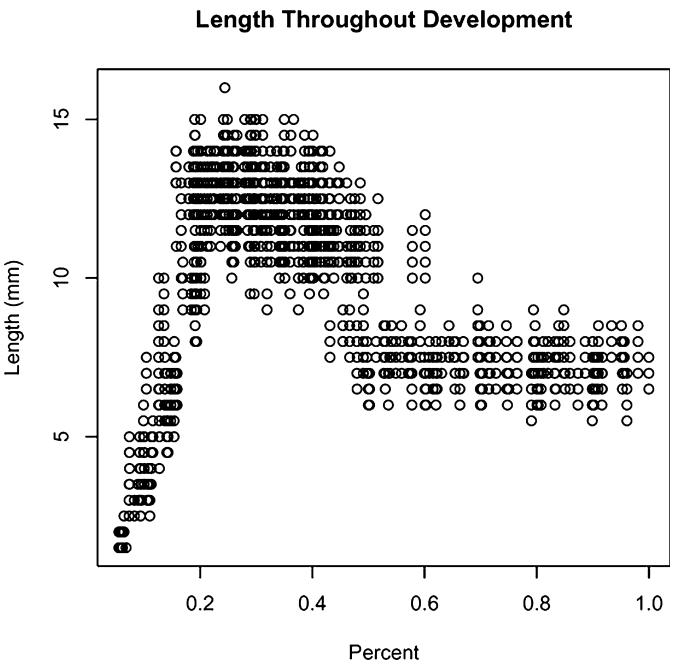

FIG. 1—*The lengths (mm) of 2559 immature* Lucilia sericata *throughout immature development (percent of development values are on a 0–1 scale). Note the tight distribution of sizes during the earlier, linear growth phase compared to the more variable postfeeding third instar and pupal stages.*

distribution of sizes in the feeding stages was much smaller than it was in postfeeding third instar larvae and pupae. Minimum and maximum development percents for each stage of development were: first instar = 5.5–11.0%; second instar = 7.4–15.4%; feeding third instar = 12.6–26.0%; postfeeding third instar = 19.1–60.1%; and pupa = 43.2–100% (Fig. 2).

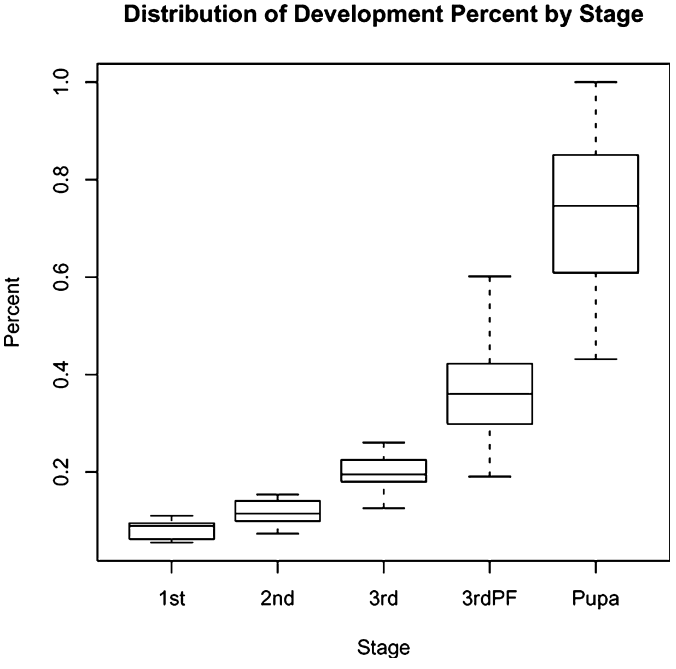

FIG. 2—*A plot of the distribution of development percents for individuals at each developmental stage. As development progressed, the proportion of the life cycle spent in a stage increased. 3rd indicates the feeding portion of the third instar; 3rdPF indicates the postfeeding stage of the third instar.*

Size was influenced slightly, but significantly, by temperature and strain. CA individuals tended to be larger than MI, which were larger than WV (Fig. 3). Differences in size among strains were not observed during feeding stages, but were observed once feeding ceased (Fig. 3) as each strain initiated the postfeeding third instar at different points in development, resulting in variation in average pupal sizes. Also, growth at 20°C yielded larger individuals on average than did growth at 33.5°C, presumably due to a change in the relative rate of development for feeding larvae (Fig. 3). Size differences caused by both strain and temperature were repeatable, although average differences were well within the variation observed for size traits (e.g., Fig. 1), resulting in an overlap of body sizes among all strains and both temperature treatments.

Assessing Statistical Models

A comparison of all models examined (Table 1) displayed the utility of GAMs to predict development percent when different variables were included. Stage was the singlemost informative variable (GCV = 0.045), while length and weight garnered less information (GCV = 0.126 and 0.144, respectively); all were statistically significant ($p < 0.0001$). The length-by-weight term (model 4) provided an intermediate level of information in assessing development (GCV = 0.059). Temperature and strain were not significant predictors of age by themselves (models 5 and 6; GCV = 0.358 for both) and only provided useful information ($p < 0.0001$ and a decreased GCV) when combined with other variables (e.g., model 18). Predictions with length and weight

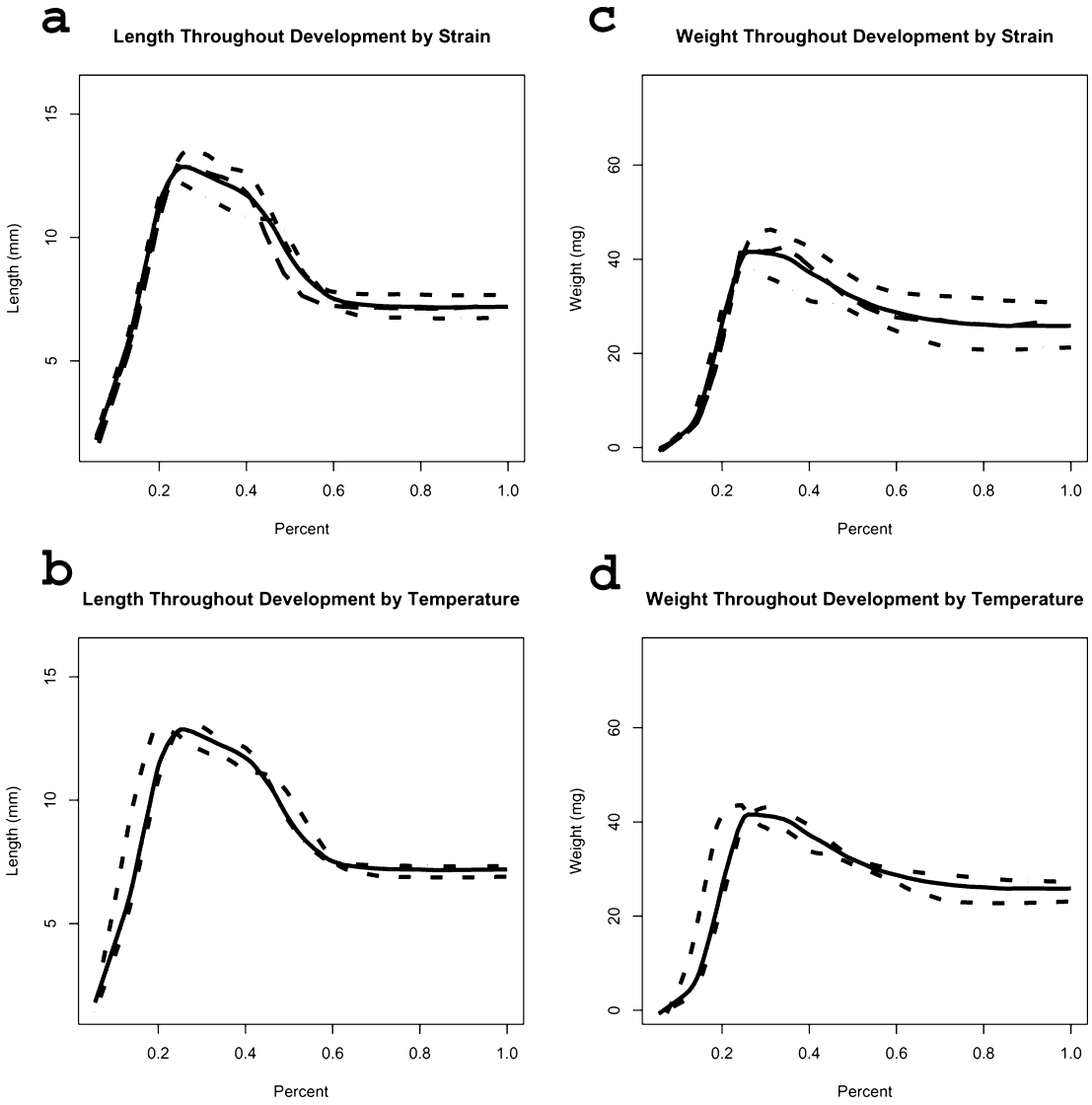

FIG. 3—*The lengths and weights of individuals throughout development from the six cohorts. Growth is compared by strain and by temperature. Solid lines represent the average for all strains or both temperatures. (a) Length (mm) plots for each strain. The largest strain, denoted by the line with short bars and spaces, was CA, and the smallest strain, designated by the line with short bars separated by dots, was WV. The MI strain was close to the average size and is represented by the spaced line with long bars and short spaces. Less size variation existed during the feeding portion of the life cycle (when size was increasing) than in the postfeeding and pupal stages. (b) Length plots comparing growth at 20°C versus 33.5°C. Growth at 20°C is represented by the spaced line with short bars separated by dots and 33.5°C is represented by the line with short bars and long spaces. The higher temperature resulted in a growth curve that had a steeper slope during the linear growth phase of development; individuals from these treatments peaked in body size proportionally faster than cooler treatments, which resulted in smaller body sizes as pupae. (c) Weight (mg) plots for each strain. Comparisons among strains were as in (a). (d) Weight plots for the two temperature treatments, with similar results as in (b).*

TABLE 1—*The 18 generalized additive models for predicting development percent assessed in this experiment.*

Model	Development Percent =	Percent	GCV
1	Stage	89.5	0.045
2	s(length)	63.3	0.126
3	s(weight)	65.7	0.144
4	s(length,weight)	86.8	0.059
5	Temperature	0.022	0.358
6	Strain	0.0041	0.358
7	Stage + strain	89.5	0.044
8	Stage + temperature	89.7	0.044
9	Stage + strain + temperature	89.7	0.044
10	Stage + s(length)	91.2	0.038
11	Stage + s(weight)	90.8	0.04
12	s(length) + s(weight)	85.9	0.064
13	Stage + s(length) + s(weight)	91.6	0.036
14	Stage + s(length) + s(weight) + s(length,weight)	92	0.035
15	Stage + s(length) + s(weight) + temperature	91.8	0.036
16	Stage + s(length) + s(weight) + strain	91.6	0.036
17	Stage + s(length) + s(weight) + temperature + strain	92	0.036
18	Stage + s(length) + s(weight) + s(length,weight) + temperature + strain	92.6	0.034

s(variable) indicates that a smoothed, nonlinear curve was used in the GAM for this variable. s(length,weight) indicates that a smoothed function of length plotted against weight was used in the GAM. Development Percent indicates the variables used in each model to predict development percent. Percent indicates the percent deviance explained. GCV, generalized cross-validation score; lower scores are better models for predicting development percent.

yielded similar results to model 4, approaching, but not improving upon, the explanatory power of stage alone (model 12; GCV = 0.064). Any model that included stage and at least one body size measure explained 90.8–92.6% of the deviance in the data and GCV scores of 0.04–0.034, with the model that included all variables garnering the highest percent deviance explained and the lowest GCV.

All models were limited in predicting the ages of postfeeding third instars and pupae, generating artificially narrow age ranges. Gaps between stages were most dramatic in model 1 (developmental stage alone), wherein individuals were simply predicted to be the average age of that stage, although true ages were continuous. Inclusion of body size improved predictive precision in the early stages, but not for postfeeding third instars and pupae. As an example, in model 10, which included developmental stage and length, postfeeding third instars that were 19.1–60.1% developed (Fig. 2) were given a restricted age range of 30.7–40.3% (95% confidence intervals in Fig. 4*a*). The gap between feeding and postfeeding third instars closed somewhat in more complex models; model 18, which utilized all available data, showed no gap between these stages (Fig. 4*b*), although the data still did not cluster along the $Y = X$ line at the level seen during feeding. The inaccuracy of predictions remained for pupae in all models, where true pupal ages were between 43.2% and 100% of immature development, but 95% of predictions for pupae using model 18 had fitted values between 61.9% and 81.2%. Interestingly, predicted ages throughout this range were made for pupae of any true age; that is, there was no slope to the pupal data as there was for the other stages.

GAM Validation with Independent Data

The utility of model 10 (developmental stage and length) was examined through analysis of the previously collected and independently produced rat carcass data set. Consistent with the above finding, error in larval age estimations increased with age (pupae were not considered here as length does not change during the stage). A plot of true versus predicted age (Fig. 4*c*) shows that age predictions generated for the rat data, when compared to known ages, spanned the $Y = X$ line and were generally consistent with (inside) the 95% confidence interval provided by the diagnostic plot for model 10. In the feeding stages (<26% of total development), the predictions were *c.* ±5% (or less) of the true age. However, in postfeeding third instars, ages were initially overestimated, then clustered close to the line, and eventually disbursed well below $Y = X$, consistent with the expectation that postfeeding individuals could not be precisely aged using length. The model also continued to predict a narrower range of ages for postfeeding larvae (32.5–40.1%) as compared to their true ages (22.9–50.2%).

Discussion

The requirements of Daubert necessitate standardized methodologies and knowledge of potential error, two criteria where several forensic sciences, including entomology, may be found lacking. Previously we examined how variation in published rearing protocols, which are not standardized among laboratories, affect growth rates of juvenile blow flies (13). In the current research, the ability to conduct statistical analysis of blow fly growth and aging, including confidence intervals, error rates, and model comparisons, was tested. From a practical standpoint, the methodology allows for direct estimates of error that should satisfy both scientific and Daubert considerations. For instance, if stage and length are used to estimate that a larva is 15% developed, model 10 generates a 95% confidence interval of *c.* 10–22% (Fig. 4*c*). Using a published minimum development time for *L. sericata* of 288 h at 26.7°C (4), an age estimate of 29–52 h is produced, with the requisite error described. For postfeeding third instars, an estimate of 40% developed (115 h) corresponds to a 95% confidence interval of *c.* 22–60%, or 63–173 h. Although the precision necessarily decreases in the latter stages, the window of time placed around that prediction is now mathematically defined. The methodology also has the flexibility to incorporate other data that may be collected.

Through the modeling used in this study, several key points became apparent. First, developmental stage was the singlemost predictive factor in the models assessed, explaining 89.5% of the deviance in the data. Logically this makes sense, as stage is a direct measurement of developmental progress. In contrast, the nonlinear measurements—weight and length—although still significant, proved far less effective in predicting development, while strain and temperature (genetic and environmental factors) were by

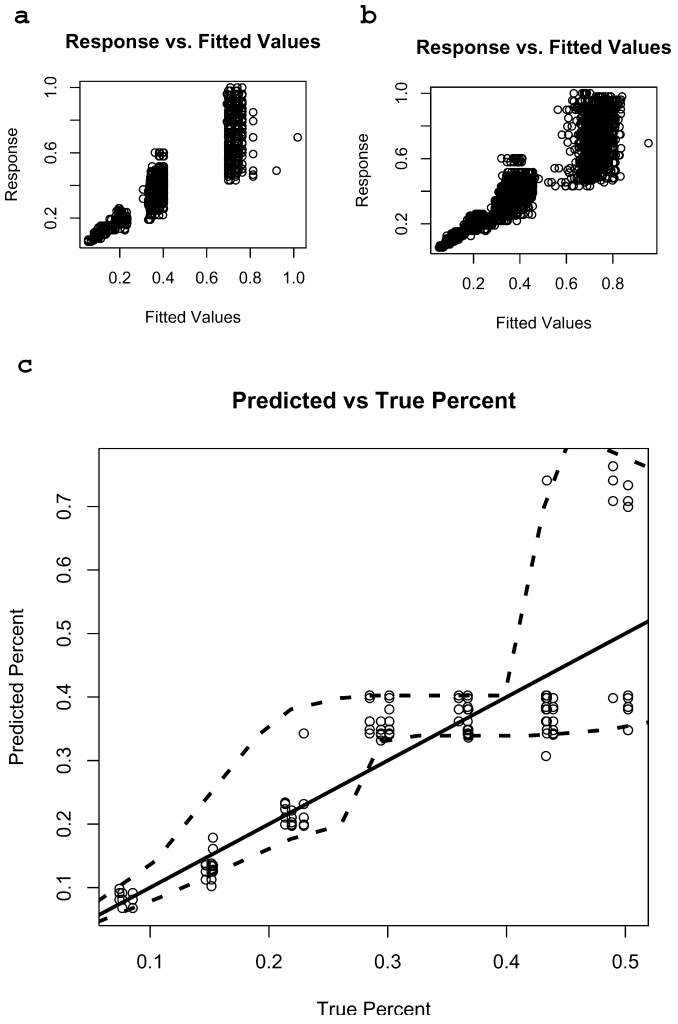

FIG. 4—*A comparison of diagnostic plots for model 10 (a), model 18 (b), and predictions made with model 10 using previously collected length and stage data (c). Model 10 predicted minimum development percent using length and stage, two criteria regularly used by entomologists for estimating a PMI. Model 18 used all available data to make predictions, explaining the most deviance in the data, and producing the lowest GCV. Panels (a) and (b) are plots of true (response) versus predicted (fitted) values for data used to construct the models. Fitted values accurately predicted true age through linear (feeding) development (or the first 25%), after which the first gap in predictions appeared. For model 10, this gap meant that larvae in that age category (c. 19–30% developed) had an overestimated developmental percent, while postfeeding third instars that were c. 40–60% developed were predicted to be younger than their true development percent. Model 18 did not produce the gap observed in model 10 between feeding and postfeeding third instar predictions, but still exhibited an increase in error during the stage. For pupae in both models, flies 43.2–100% developed corresponded to fitted values that represented a much smaller range of development (e.g., 61.9–81.2% for model 18), thus age was both over and underestimated. Panel (c): A plot of predicted versus true development percents of 252 larvae raised on rats as estimated with a generalized additive model for developmental stage and length (model 10). Development is displayed to 50% because length measurements ceased when pupation occurred. The solid line represents the* Y *(predicted age) =* X *(true age) line. Dashed lines represent 95% confidence intervals for the predictions based on the data in (a). Model 10 accurately (data span* Y = X*) and precisely (ranging c. 5% above and below* Y = X*) predicted age for the feeding portion of the life cycle (<26% of development), when body size increased in a linear fashion. As expected, precision decreased greatly once the postfeeding third instar was reached, although overall error for the model was within predicted levels.*

themselves not significant predictors. The ability of stage, length, and weight to estimate age was greatest during the earliest phases of development, but for different reasons. Egg, first instar, and second instar are by far the shortest developmental stages in flies (Fig. 2), thus identification of any of these simply described development more accurately than did the much longer third instar and pupation. Weight and length on the other hand are related to feeding, and changed in a relatively linear fashion during the early larval stages, including the first portion of the third instar (e.g., Fig. 1), but once feeding ceased their utility dropped dramatically due to the reversal in body size and larger overall variance in length and weight. Likewise, pupal size was of little utility as it is static throughout the stage. As a result, model 18, which used all available information, predicted a restricted pupal development of 61.9–81.2% (Fig. 4*b*, 95% confidence intervals) and showed no specificity (that is, the youngest or oldest individuals were equally likely to be placed anywhere within that range). This means a pupal development prediction with the best GAM was essentially the same as using stage alone. Given these effects on predictive ability, it is not surprising that adding weight and/or length to stage resulted in minimally improved models.

Second, error increased as development progressed for all models, indicated by the gaps in predictive ability and the widening confidence intervals for successive developmental stages (Figs. 4*a* and 4*b*), which were most pronounced in postfeeding third instars and pupae. The increasing inaccuracy of age approximation as fly development progresses has been noted in the literature (5,7), and forensic entomologists account for it in PMI estimates by giving large age ranges to postfeeding flies, although these rarely include an objective estimate of error. The latter study (7) used linear models to estimate blow fly age based on length data, and yielded an increase in error for older larvae. The similar findings indicate that there is a limit to the precision in blow fly age predictions that can be achieved when only developmental stage and body size are evaluated. Owing to this, alternative developmental data independent of basic growth are likely required to increase the accuracy of PMI predictions, and in the future, traits that change regularly during fly development, such as hormone titers or gene expression, may be useful in generating a more specific PMI.

Third, the limited influence of fly strain and rearing temperature on development is an important consideration as it indicates the models have value regardless of where flies are collected or at what temperature they develop. This is not to say that temperature is unimportant when making PMI estimates—it is critical, and is always considered when estimating PMI (usually as accumulated degree days). However, temperature did not alter developmental patterns to any large extent, although lower rearing temperatures did result in slightly larger individuals overall for all strains, a finding we continue to investigate. Similarly, the strains of flies examined had different average sizes during development (Fig. 3). For both criteria, the distributions of body size throughout development overlapped, so these data modified age predictions minimally. Also, there was little difference in size among strains during the feeding stages of the life cycle, where size best predicts age, thus size variability resulting from strain adds no confounding information during those stages.

Fourth, any given forensic case may present the entomologist with different data from which to estimate fly age. While developmental stage was the most useful datum for the development estimates in this study, other data, such as weight and length, increased their accuracy. Using a model that incorporates all available data can help ensure that investigators make the best possible prediction with the information they receive, while maintaining an

understanding of the limitations inherent in that model. An estimate of the relative reliability of a PMI prediction (based on the GCV and percent deviance explained) provides an understanding of its value in interpreting evidence. Overall, GAMs offer a useful means of incorporating information from multiple linear and nonlinear variables to predict blow fly age, variables that can be accommodated even if they change from one case to the next.

Finally and most importantly, a comparison of modeled development predictions to the independently derived rat data made it possible to assess error rates and produce confidence intervals in these estimates. Individuals <26% developed (feeding larvae) generated the most accurate predictions; when 12 individuals of the same age were sampled from a cohort, the predictions clustered around the known age in all instances (Fig. 4*c*). In contrast, postfeeding stages had a much larger error rate. It is worth noting that even when model 10 yielded its best estimates, there was still *c.* 10% total variance in predicted ages of larvae (note again, at a true age of 15%, the 95% confidence interval for rat data predictions was between *c.* 10% and 22%, thus stage and size were not "perfect" in estimating development even in the youngest individuals). The utility of the methodology presented here is that it establishes a defined way to produce confidence intervals around entomologically based PMI predictions, regardless of fly age. Until new and independent variables that change in a predictable manner during development are incorporated into age estimates, this error will necessarily exist, and increase with age; however through these models that error can objectively be determined. Equipped with such knowledge, the forensic entomologist can relay to the court the level of error found in a PMI prediction. Through this feat, one of the major requirements of Daubert is more fully addressed.

Acknowledgments

The authors would like to thank Kim Jennings for her help with gene sequencing, Jeff Tomberlin and Ryan Kimbirauskas who confirmed species identifications for the MI strain, Bob Kimsey who provided flies from CA, Rich Merritt for allowing the use of lab equipment and his insight into the research, and Dave Szymanski for helpful comments. Thanks to Jeff Wells for advice on modeling blow fly development and for providing WV flies. Randy Shoemaker and Constance Crew provided animal ethics consultation and euthanized rats. Programming in R and basic GAM theory originated from the skills of Brian McGill. This body of work was a portion of the PhD dissertation research of AT. This project was supported by Grant Number 2004-DN-BX-K005, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

References

1. Saks MJ, Koehler JJ. The coming paradigm shift in forensic identification science. Science 2005;309(5736):892–5.
2. Greenberg B. Flies as forensic indicators. J Med Entomol 1991;20(5):565–77.
3. Grassberger M, Reiter C. Effect of temperature on *Lucilia sericata* (Diptera: Calliphoridae) development with special reference to the isomegalen- and isomorphen-diagram. For Sci Int 2001;120:32–6.
4. Kamal AS. Comparative study of thirteen species of sarcosaprophagous Calliphoridae and Sarcophagidae (Diptera) I. Bionomics. Ann Entomol Soc Am 1958;51:261–70.
5. Wells JD, Kurahashi H. *Chrysomya megacephala* (Fabricius) (Diptera: Calliphoridae) development: rate, variation, and the implications for forensic entomology. Jpn J Sanit Zool 1994;45(4):303–9.
6. Anderson GS. Minimum and maximum development rates of some forensically important Calliphoridae (Diptera). J Forensic Sci 2000;45(4):824–32.
7. Wells JD, Lamotte LR. Estimating maggot age from weight using inverse prediction. J Forensic Sci 1995;40(4):585–90.
8. Hastie TJ, Tibshirani RJ. Generalized additive models. Boca Raton: Chapman and Hall/CRC, 1990.
9. Wood SN. Generalized additive models: an introduction in R. Boca Raton: Chapman and Hall/CRC Taylor and Francis Group, 2006.
10. Moe SJ, Kristoffersen AB, Smith RH, Stenseth NC. From patterns to processes and back: analyzing density-dependent responses to an abiotic stressor by statistical and mechanistic modeling. Proc R Soc Lond B Biol Sci 2005;272:2133–42.
11. Hall DG. The blow flies of North America. Baltimore, MD: Monumental Printing Co., 1948.
12. Gorham JR. Insect and mite pests in food: an illustrated key. Beltsville, MD: U.S. Department of Agriculture, Agriculture Handbook Number 655, 1987.
13. Tarone AM, Foran DR. Components of plasticity in the development of a Michigan population of *Lucilia sericata*. J Med Ent 2006;43(5):1023–33.
14. Rogina B, Helfand SL. Regulation of gene expression is linked to life span in adult *Drosophila*. Genetics 1995;141:1043–8.
15. Rogina B, Benzer S, Helfand SL. *Drosophila* drop-dead mutations accelerate the time course of age-related markers. PNAS 1997;94:6303–6.
16. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2004. ISBN 3-00051-00-3. Available at: http://www.R-project.org. Accessed on April 28, 2008.
17. Mansson RA, Frey JG, Essex JW, Welsh AH. Prediction of properties from simulations: a re-examination with modern statistical methods. J Chem Inf Model 2005;45:1791–803.

Additional information and reprint requests:
David R. Foran, Ph.D.
Forensic Science Program
School of Criminal Justice and Department of Zoology
560 Baker Hall
Michigan State University
East Lansing, MI 48824
E-mail: foran@msu.edu