Herpetological Review, 2003, 34(1), 35–39.

Testing for Repeatability in Measurements of Length and Mass in *Chthonerpeton indistinctum* (Amphibia: Gymnophiona), Including a Novel Method of Calculating Total Length of Live Caecilians

G. JOHN MEASEY [1,2,3]
JORGE BERNARDO SILVA [1]
and
MARCOS DI-BERNARDO [1]
[1] Laboratório de Herpetologia - Museu de Ciências e Tecnologia
Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Av. Ipiranga, 6681. CEP 90619-900, Porto Alegre, RS, Brazil

[2] Department of Zoology, The Natural History Museum
Cromwell Road, London, SW7 5BD, UK

[3] Present address and for correspondence:
Laboratoire d'Ecologie des Sols Tropicaux
Institut de Recherche pour le Développement
32 Avenue Henri Varagnat, 93143 Bondy Cedex, France
e-mail: john.measey@bondy.ird.fr

Individual size remains an important part of any study of reptiles and amphibians, be it systematic or ecological, with countless studies demonstrating the importance of measurements of body length in most taxa (snout to tail tip, snout–vent length, etc., e.g., Lowcock et al. 1992; Scolaro et al. 1985; Vrcibradic and Rocha 1996). However, very few studies have included measurements of live caecilians (Amphibia: Gymnophiona), relying instead on external measurements from museum specimens. For example, Gudynas et al. (1988:16) found that all body measurements were significantly correlated with total length in the semi-aquatic caecilian, *Chthonerpeton indistinctum*: "of 14 significant correlations, eight presented coefficients greater than 0.8." In ecology, size range can be useful in comparisons between populations, and the combination of length and mass is often used to infer the relative condition of individuals over time or between populations.

During ecological fieldwork there is a constant trade-off between collecting as much information as possible, and the reduction of handling time of each individual animal. Ultimately, it is the welfare of the animal that must remain paramount, especially when the normal behavior of the individual is a necessary requirement (such as in mark-recapture experiments). Thus, when an animal is alive, it is not possible to take the ideal full complement of measurements; this makes the precision of measurements that are taken all the more important. Of the numerous publications that include morphometric measurements of live animals, few test to see whether such measurements are repeatable. We believe that testing for repeatability of measurements from live animals should be an integral part of any ecological study that relies on morphometric data.

We consider three conditions important when undertaking measurements from a collection of live animals. Firstly, measurements should be taken by a single operator. Multiple operators tend to have undesired confounding effects on measurements, note taking, and behavior of animals. Secondly, anesthetic should not be used simply to measure an animal. We realize that some procedures require anesthetics (e.g., tagging) and it would seem prudent to take measurements from an animal at such a time. Lastly, we believe that the method chosen should be consistent for future comparison of data. A mixture of data taken with various methods and when animals are under different conditions (e.g., conscious versus unconscious) are likely to be less repeatable.

There are only two principal methods to determine the length of elongate amphibians and reptiles, although we are aware that there are variations on each theme.

1. A fixed rule with calibrated marks to which the animal is straightened, often having to measure the animal in parts as its movements allow. Variations on this theme range from a simple ruler to a rigid calibrated tube through which a live animal is persuaded to move.
2. The string method, in which string is placed along the length of the animal. The length of the string is then retrospectively determined.

The consistent difference between these methods, including all their variants, is that in the first, the measurement apparatus is fixed and, in the second, it is flexible. From a theoretical viewpoint the first method suffers from the problem that animals are not straight and forcibly straightening an animal is bound to lead to errors. Although a tube might seem an ideal way to measure elongate animals, animals tend not lie straight within tubes and in some burrowing taxa (including some caecilians) total length can vary during movement (see Gans et al. 1978; O'Reilly et al. 1997). The string method can also be problematic, especially when measurements need to be taken by a single operator, for example when venomous snakes or slippery caecilians are to be measured.

The magnitude of measurement error in morphometric characters is often surprisingly high (Merilä and Björklund 1995 and references therein). During our fieldwork with the caecilian, *Chthonerpeton indistinctum* (and previous fieldwork by GJM with caecilians, MDB with snakes, and JBS with amphisbaenids) we have experienced considerable difficulty using the aforementioned methods. In particular, we were concerned that measurements using our preferred method—a fixed ruler—were not repeatable. We decided to test this hypothesis by statistical analysis of repeated measures, using the caecilian *Chthonerpeton indistinctum* as a model. We also compare the fixed ruler to another method using the flexible string principle, involving images acquired with a digital camera. At the same time we tested the repeatability of field methods to measure mass, with spring balances and a digital balance.

TABLE 1. Informal terms to describe the measure of repeatability r_i, from Martin and Bateson (1986). Note that these can only be used if results are statistically significant.

R_i	Term
r_i *less than 0.2*	*Slight repeatability*
r_i *between 0.2 and 0.4*	*Low repeatability*
r_i *between 0.4 and 0.7*	*Moderate repeatability*
r_i *between 0.7 and 0.9*	*High repeatability*
r_i *greater that 0.9*	*Very high repeatability*

The semi-aquatic caecilian *Chthonerpeton indistinctum* was collected by excavation during ecological field studies in May 2001 from the Litoral Norte region of Rio Grande do Sul, Brazil. It was necessary to transport 19 specimens from two localities to the laboratory to assess the suitability of marking techniques (see Donnelly et al. 1994; Measey et al. 2001), and it was during this period that the following procedures were made.

To test for repeatability and measurement error in measurements of total length and mass, all 19 animals from both localities were used. Caecilians were temporarily placed within individual containers with a small amount of the substrate in which they were captured. Each container was labeled with a number (1 to 19) and a letter to denote the site of collection. Random numbers were generated using a hand-held calculator to determine the order in which each animal was measured. Animals were treated as if measurements were being taken in the field. Thus, individuals were carefully removed from their containers, cleaned of excess mud (if necessary), placed onto paper towel to remove excess moisture, weighed and then measured before being replaced. Note that were animals to be multiply measured and then replaced, this would constitute pseudoreplication. To avoid potential inter-operator bias, each procedure was carried out by a single operator (1, 2, and 3 by GJM, 4 by JBS):

1. Total live mass was measured using an ACCULAB® Pocket Pro® 250-B (Sartorius Group, Denver, USA), taped with a plastic dish on top of the pan. The animal was placed entirely within the dish and the mass of the stabilized reading recorded.
2. Total length was measured using a fixed plastic ruler with divisions in mm. The animal was placed venter first to the zero end of the ruler. If calmly handled the animal responded by relaxing the posterior body portion, allowing relatively swift measurement of at least two thirds of the body length. The length of this portion was noted on the ruler. The anterior third required some coercion of the animal to lie along the ruler, especially the head and collar region, although minimum manipulation was used to obtain the total length.
3. Two PESOLA® (Pesola, Switzerland) precision spring-balances (10g and 100g) were suspended from a fixed point and zeroed with a length of clear plastic folded to form a sling (ca. 150 x 150 mm, cut from a plastic bag) in which the animal was placed. The measurement of total live mass was then read from the scale.
4. An Olympus C-2100 digital camera was mounted on a tripod 0.5 m from a piece of laminated graph paper. Animals were placed in the center of the graph paper and an exposure made as soon as the caecilian lay in a single plane on the plastic. For most animals the exposure could be made within a couple of seconds as the animal was carefully placed. Occasionally the animal coiled around itself and it was necessary to reposition it before making an exposure. The image was saved as a JPEG file onto a personal computer and was later written to a CD. The image was loaded into Image Tool software (for Windows, Version 2.0). Image Tool is freeware, and the current version can be downloaded from the internet (Wilcox et al. 1997).

   Each image was magnified by a ratio of 1:2 for calibration on a 10 mm section of graph paper. This calibration was then checked by returning the image to 1:1 and measuring 100 mm of graph paper. A result within 1 mm was accepted, otherwise calibration was repeated. A continuous sequence of lines was then drawn over the image in the center of the dorsum of the animal from the snout to tail tip (total length); care was taken to keep the angle between preceding and following lines between 140° and 200° (Fig. 1). The software automatically calculated the total length of the line formed by the sequence of all lines. Image and software distortion is bound to have occurred, although we consider such problems to be insignificant given calibration accuracy.

This process was repeated on all 19 *C. indistinctum* four times, so that each animal was processed five times in total. Records were made without reference to previous measurements. Note that no anesthetic was used, nor were the animals cooled. The whole process took under three hours (excluding time required for analysis of images). The animals were returned to their collection location the following day.

Data were tested for normality using MINITAB. The repeatability of measurements was then tested with one-way analysis of variance (ANOVA, MS EXCEL). Two-tailed paired-sample *t* tests were conducted on averaged data of different methods (MS EXCEL).

The repeatability index or intra-class correlation coefficient (r_i equation 1) can be used to assess precision, i.e., whether an observer makes consistent measurements and whether a trait varies. It cannot be used to measure accuracy of measurements, or differences between observers. Table 1 shows the allocation of classes from the results of repeatability (equation 1). Measurement error (ME) is the percentage of total phenotypic variation in a trait attributable to errors during measurement, calculated on the proportion of within-individual variation to total variation (equations 2 and 3 from Bailey and Barnes 1990).

equation 1 (1 = perfect; 0 = none)

$$r_i = \frac{MS_{between} - MS_{within}}{\left(MS_{between} + (n-1)MS_{within}\right)}$$

equation 2

$$ME = \frac{MS_{within}}{\left(s^2 + MS_{within}\right)} \times 100$$

equation 3

$$s^2 = \frac{\left(MS_{between} - MS_{within}\right)}{n}$$

equation 4

$$AverageDeviation = \sum \frac{\left|\bar{x} - x\right|}{n}$$

Where n is the number of repeated measurements (here five), $MS_{between}$ is the Mean Squares between groups, MS_{within} is the Mean

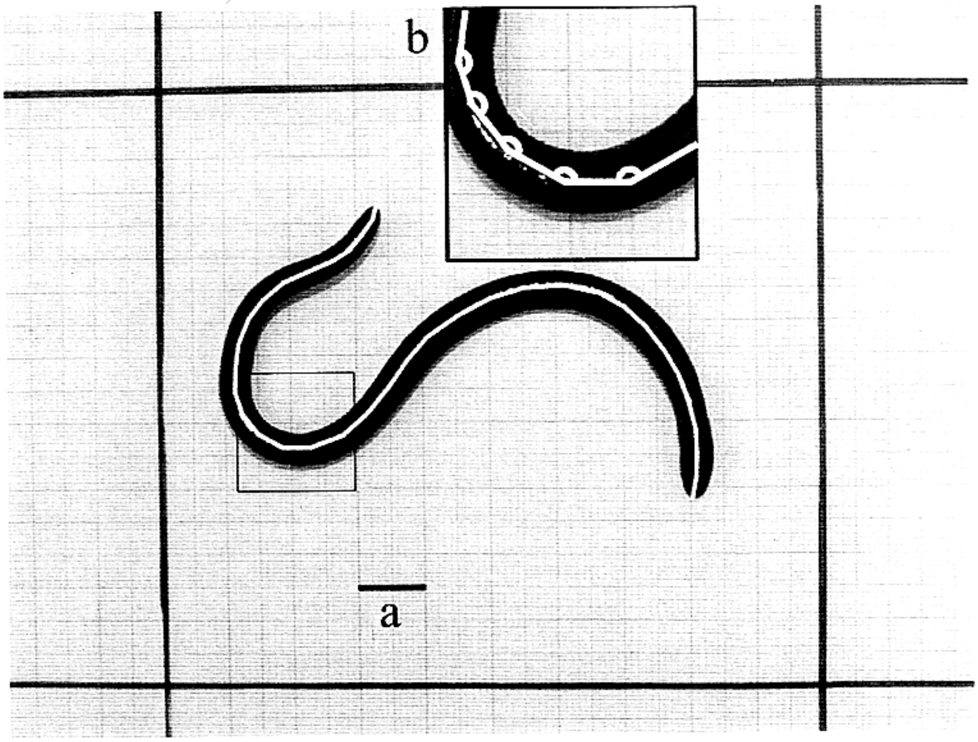

FIG. 1. A typical digital image of a *Chthonerpeton indistictum* juvenile with a sequence of lines drawn along the dorsum. a) 10 mm calibration line and b) inset of increased magnification showing a series of lines with angles (from left to right) of 150°, 163°, 154°, 154°, 166°.

Squares within groups and $\bar{x}$ is the mean of the samples (x).

Measurement of mass.—Both spring balances and pocket balance were easy to use and offered no special difficulties in operation or recording of mass. Operation time of the pocket balance was considerably faster than spring balances resulting in a shorter duration of handling. Fig. 2 shows the mean mass and variation in total length for each method.

Table 2 shows a summary of the data for repeated measures of mass with a very high repeatability for both the electronic and spring balances. Over the five measurements of mass, animals showed consistent loss between first and fifth measurements for the spring balance ($\bar{x}$ = 0.232 g) and for the electronic balance ($\bar{x}$= 0.159 g), although this loss was not significant (spring $F_{4,90}$ = 0.0021, P = 0.999; electronic $F_{4,90}$ = 0.0009, P = 0.999). Weight probably changed as a result of urination, which is common in caecilians when first handling them (GJM, pers. obs.). The t test showed significant difference between weights using the two methods ($t_{2,18}$ = –2.924, P = 0.009), although the mean difference was smaller than could be measured with these instruments (0.04 g or 0.65%).

Measurement of total length.—Manipulation of individuals for measurement of total length with the ruler was generally easier for larger animals which offered least resistance to being forcibly straightened. Smaller animals often coiled around themselves and were generally harder to work with. Similar behavior was observed when animals were placed on the graph paper under the camera, although the image was usually taken prior to this causing difficulties. Measurement using the ruler was far faster than using the digital camera, but the handling time and severity of handling was far greater with the ruler than when using the camera. Operation of the camera itself was swift, but more time was spent obtaining length from digital images, approximately 2 minutes for each image. This time involved transfer of data, processing of images with the Image Tool software, and backing up of image files.

The ruler was found to produce significantly longer measurements ($t_{2,18}$ = 6.235, P < 0.0001) with an average of 5.7 mm (3.3%). No consistent increase or decrease of length was noted with measurements over time, and differences between trials were not significant (ruler $F_{4,90}$ = 0.002411; P = 0.999; camera $F_{4,90}$ = 0.008118; P = 0.999).

For valid application of parametric analyses of variance and t tests, we must be able to assume that data from each group are obtained randomly from a normal distribution; also that the sampled populations have equal variances (homoscedastic) and have factors with additive levels (Zar 1999). Data in this study were found to deviate significantly from the normal distribution, however we consider that as the sampled populations are essentially the same (being the same animals) this would not invalidate the analysis of variance and t tests (see discussion in Zar 1999:185). Data were found to be log-normal, and application of statistical methods to log transformed data did not alter in significance within three decimal places.

Harper (1994) cautions against errors, when making conclusions from repeatability (r_i) calculations, when animals of greatly varying sizes or species are used. In this study two considerably larger adult animals were included, while all others were juveniles (see Fig. 2). It is possible that inclusion of these individuals would

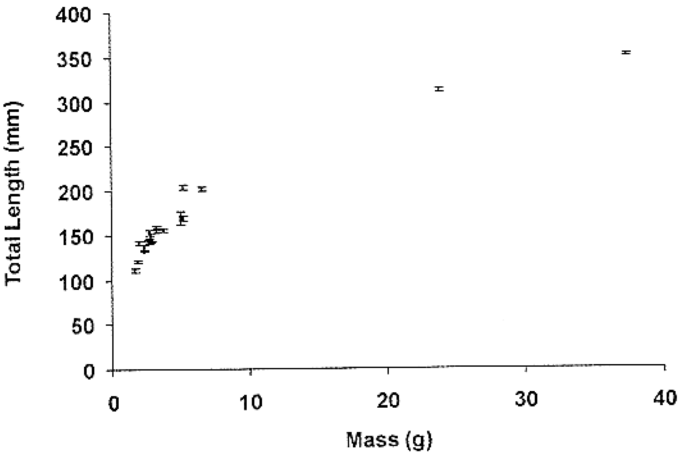

FIG. 2. The distribution of total length and mass of *Chthonerpeton indistinctum* from two localities in Litoral Norte, Rio Grande do Sul, Brazil. Points show the range of length data and averaged mass obtained using a fixed ruler and spring balance (the graph for digital camera and digital balance is visually identical).

produce a bias for the index of repeatability and so the calculations were repeated excluding them (see Table 2).

Precision refers to the closeness to each other of repeated measurements of the same quantity (Zar 1999). All methods presented are precise, showing 'very high repeatability' and an insignificant (i.e., less than 5%) amount of measurement error. This held true (with slightly reduced indices) when only animals of a similar size were measured (by removing the two large animals, Table 2). We do not consider the precision of many operators, as this is known to be considerable (up to 30% in measurements of skulls; Palmeirim 1998); for this reason we suggest that single operators are mandatory.

Accuracy is the nearness of a measurement to the actual value of the variable being measured (Zar 1999). Although it is possible to determine the accuracy of each balance, the accuracy of measurements of total length is somewhat more difficult. The fixed ruler method proved slightly more repeatable, but gave significantly longer measurements, than measurements made from digital images. We suggest that the choice of technique should therefore concentrate on the relative merits of each technique.

TABLE 2. Results of the multiple measurements on 19 *Chthonerpeton indistinctum* from Rio Grande do Sul, Brazil. Mean Squares, F and Critical value of F were derived from one-way ANOVA (MS Excel). Average deviations calculated from positivized data set. (Figures in parentheses are those from calculations excluding the two adult specimens, N = 17.) † indicates that the same measurement of an animal is producing the maximum or minimum measurement and percentage. * shows the same individual producing maximum and minimum percentages.

Measurement	Mass (electronic)	Mass (spring)	Total Length (ruler)	Total Length (camera)
Mean Squares between groups	399.8549 (9.643529)	404.3236 (9.656882)	20573.32 (3218.224)	18881.92 (2999.054)
Mean Squares within groups	0.006421 (0.005529)	0.031579 (0.005882)	9.289474 (9.911765)	19.69082 (13.91397)
F	62272.49 (1744.043)	12803.58 (1641.67)	2214.692 (324.6872)	958.9197 (215.5427)
Critical value of F	1.741189 (1.794557)	1.741189 (1.794557)	1.741189 (1.794557)	1.741189 (1.794557)
P	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Index of Repeatability (r_i)	0.999157 (0.99714)	0.9992 (0.996962)	0.996022 (0.984788)	0.992473 (0.977225)
Term	Very high repeatability	Very high repeatability	Very high repeatability	Very high repeatability
Measurement Error (ME)	0.008029% (0.286034%)	0.039039 % (0.303828%)	0.225358% (1.521203%)	0.519254% (2.277462%)
Max/Min deviation from mean	+0.20 g –0.20 g	+0.16 g –0.14 g	+12.8 mm † –6.2 mm †	+14.99 mm –15.17 mm †
Max/min as a % of measurement	+6.7% * –8.9% *	+4.3% –5.4%	+7.0%† –4.0%†	+8.9% –10.8%†
Average deviation ± Std error	0.052105 g ± 0.004965	0.093053 g ± 0.013291	1.991579 mm ± 0.191999	2.713053 mm ± 0.298792
% average deviation ± Std error	1.463956% ± 0.158015	1.744285% ± 0.111033	1.213712% ± 0.116015	1.597892% ± 0.166096

The fixed rule is a simple and inexpensive method for measuring the total length of live animals. Equipment is easily obtained, transported, replaced (if broken, stolen, etc.) and can be autoclaved to prevent transport of pathogens; this last point is especially important for herpetologists working with fossorial species (see Halliday 1998). However, we do acknowledge that operation of this equipment requires experience to produce repeatable measurements. Also, data taken with a fixed rule represent a "one off" measurement that cannot be checked by another authority.

The digital camera offers an archive image of each animal enabling measurements to be taken by various operators of the image analysis software. Images can also be used for taking various other measurements (ventral images can be recorded through glass), and other measurements can be obtained retrospectively, given that images are of sufficient quality and have been stored well. Further, images could also be used for identification of individuals (e.g., Donnelly et al. 1994), and charting the occurrence of scars and marks (known to occur in caecilians: e.g., *Scistometopum thomense,* Teodecki et al. 1998; *Gegeneophis ramaswamii,* Measey et al. 2001) on individually marked animals. However, this equipment is presently very expensive and comparatively complicated to use. Moreover, like many electronic products, it is prone to un-repairable error in the field. Time taken to analyze images with the software is significant.

We follow the philosophy of Fellers et al. (1994) who advocate the reduction in handling and handling time. In this study, a general reduction in handling would be made by using an electronic field balance and digital camera to determine the mass and total length of individual *Chthonerpeton indistinctum.* We caution against using the results from this study for other taxa, as behavior of each taxon during measurement might be different. However, we recommend that particularly in studies of growth where measurements are taken from live animals, a method of estimating measurement error is used.

Acknowledgments.—We thank CNPq for permission to conduct this research, the Royal Society for financial support, and Marcio Martins-Borges for discussion and fieldwork.

LITERATURE CITED

BAILEY R. C., AND J. BARNES. 1990. A new, old method of assessing measurement error in both univariate and multivariate morphometric studies. Syst. Zool. 39:124–130.

DONNELLY, M. A., C. GUYER, J. E. JUTERBOCK, AND R. A. ALFORD. 1994. Techniques for marking amphibians. *In* W. R. Heyer, M. A. Donnelly, R. W. McDiarmid, L.-A. C. Hayak, and M. S. Foster (eds.), Measuring and Monitoring Biological Diversity, pp. 277–284. Smithsonian Institutional Press, Washington, D.C.

FELLERS, G. M., C. A. DROST, AND W. R. HEYER. 1994. Handling live amphibians. *In* W. R. Heyer, M. A. Donnelly, R. W. McDiarmid, L.-A. C. Hayak, and M. S. Foster (eds.), Measuring and Monitoring Biological Diversity, pp. 275–276. Smithsonian Institutional Press, Washington, D.C.

GANS, G., H. C. DESSAUER, AND D. BAIC. 1978. Axial differences in the musculature of uropeltid snakes: the freight-train approach to burrowing. Science 199:189–192.

GUDYNAS, E., J. D. WILLIAMS, AND M. M. AZPELICUETA. 1988. Morphology, ecology and biogeography of the South American caecilian *Chthonerpeton indistinctum* (Amphibia, Gymnophiona, Typhlonectidae). Zoologische Mededelingen 62:5–28.

HALLIDAY, T. 1998. A declining amphibian conundrum. Nature 394:418.

HARPER, D. G. C. 1994. Some comments on the repeatability of measurements. Ringing and Migration 15:84–90.

LOWCOCK, L. A., H. GRIFFITH, AND R. W. MURPHY. 1992. Size relation to sex, hybridity, ploidy, and breeding dynamics in central Ontario populations of the *Ambystoma laterale-jeffersonianum* complex. J. Herpetol. 26:46–53.

MARTIN, P., AND P. BATESON. 1986. Measuring Behavior: An Introductory Guide. Cambridge University Press, Cambridge.

MEASEY, G. J., D. J. GOWER, O. V. OOMMEN; AND M. WILKINSON. 2001. Permanent marking of a fossorial caecilian, *Gegeneophis ramaswamii* (Amphibia: Gymnophiona: Caeciliidae). J. South Asian Nat. Hist. 5:141–147.

MERILÄ, J., AND M. BJÖRKLUND. 1995. Fluctuating asymmetry and measurement error. Syst. Biol. 44:97–101.

O'REILLY, J. C., D. A. RITTER, AND D. R. CARRIER. 1997. Hydrostatic locomotion in a limbless tetrapod. Nature 386:269–272.

PALMEIRIM, J. M. 1998. Analysis of skull measurements and measurers: Can we use data obtained by various observers? J. Mammal. 79:1021–1028.

SCOLARO, J. A., J. M. CEI, AND L. ARIAS-DE-REYNA. 1985. La identidad de las especies del grupo *Liolaemus fitzingeri-melanops* por medio del analisis discriminante de caracteres morfologicos (Sauria: Iguanidae). Historia Natural 5:13–22.

TEODECKI, E. E., E. D. BRODIE, JR., D. R. FORMANOWICZ, JR., AND R. A. NUSSBAUM. 1998. Head dimorphism and burrowing speed in the African caecilian *Schistometopum thomense* (Amphibia: Gymnophiona). Herpetologica 54:154–160.

VRCIBRADIC, D., AND C. F. D. ROCHA. 1996. Ecological differences in tropical sympatric skinks (*Mabuya macrorhyncha* and *Mabuya agilis*) in southeastern Brazil. J. Herpetol. 30:60–67.

WILCOX, C. D., S. B. DOVE, W. D. MCDAVID, AND D. B. GREER. 1997. UTHSCSA ImageTool Version 2.00 http://ddsdx.uthscsa.edu/dig/.

ZAR, J. H. 1999. Biostatistical Analysis. Prentice Hall International, London.

Herpetological Review, 2003, 34(1), 39–42.

Comparative Effectiveness of Two Trapping Techniques for Surveying the Abundance and Diversity of Reptiles and Amphibians Along Drift Fence Arrays

CHRISTOPHER L. JENKINS[1]
KEVIN McGARIGAL
and
LLOYD R. GAMBLE
Department of Natural Resources Conservation, University of Massachusetts Amherst, Massachusetts 01003, USA

[1] *Current Address: Herpetology Laboratory, Department of Biological Sciences Idaho State University, Pocatello, Idaho 83209, USA*
e-mail: christolj@hotmail.com

In the northeastern United States, there are growing concerns about the effects of habitat loss and degradation on vernal pool herpetofauna (Gibbs 1993; Kittredge 1996; Melvin and Roble 1990; Windmiller 1996). Conservation of these species requires effective and efficient methods for surveying their populations. Perhaps the most common method of surveying for adult amphibians involves drift fence arrays in combination with pitfall traps. This method has been used successfully to capture a variety of forest floor vertebrates (Bury and Corn 1987; Gibbons and Semlitsch 1981), including ambystomid salamanders (DeGraaf and Rudis 1990; Madison 1998; McWilliams and Bachmann 1988; Pechmann 1995; Stenhouse 1985; Whiteman et. al. 1994), other salamander species (Gill 1978a,b) and frogs (Guttman et. al. 1991; Yanosky et al. 1997).

Previous studies from around the world have compared the effectiveness of various terrestrial amphibian and reptile trapping techniques. These studies have found varying effectiveness of drift fence/pitfall trap arrays when compared to other methods such as cover boards, pipe traps, visual surveys, box traps and calling surveys (Christiansen and Vandewalle 2000; Crosswhite et al. 1999; Lohoefener and Wolfe 1984; Parris et al. 1999; Sutton et al 1999; Webb 1999). Another method, terrestrial funnel traps, was found to be successful for capturing amphibians along drift fence arrays in the southeastern United States (Enge 1997a). Comparisons of funnel traps to pitfall traps in the Pacific Northwest and the Southwest have shown that snakes and some lizards are more susceptible to capture in funnel traps (Bury and Corn 1987; Jorgensen et al. 1998).

To our knowledge, the only study in the northeastern United States comparing terrestrial trapping techniques for amphibians and reptiles found funnel traps in conjunction with drift fence arrays to be more effective than plastic cover sheets (Kjoss and Litvaitis 2001). A rigorous comparative evaluation of the effectiveness of funnel traps and pitfall traps has not been conducted in the Northeast. Because of rocky soils and wet conditions often found adjacent to amphibian breeding sites in the northeast comparing the effectiveness of terrestrial trapping techniques that can be used in these situations will be valuable.

The goal of this study was to evaluate the effectiveness of terrestrial funnel traps and pitfall traps for capturing amphibians and reptiles by placing the traps along drift fence arrays that encircle