Recommendations for Fat Scoring

Erica H. Dunn
Canadian Wildlife Service
National Wildlife Research Centre
Raven Road
Carleton University
Ottawa, ON KIA OH3

ABSTRACT

Fat scoring will be most valuable for pooled analyses if banders use the same scoring scale, but currently there are many systems in use. Moreover, most scoring systems have minimal descriptions of the score criteria, so there is a good deal of scope for individual interpretation. If fat scoring is to be done at all, I recommend adoption of the Kaiser scale that is used widely in Europe, which has been well tested and which minimizes individual variation in scoring because its descriptions are so detailed. Alternatively, other scales already in use could be improved along the Kaiser model by providing detailed descriptions and illustrations to delineate the lower and upper limits of each fat class.

INTRODUCTION

Fat scoring is a rapid and relatively simple way to estimate the fat reserves of birds, based on observing the quantity of fat visible under the skin. While the predictive power of fat scores for estimating true fat content of an individual bird is quite low (Scott et al. 1995), group mean fat scores are good indicators of group mean fat content and have proven valuable for a wide variety of studies (see examples in Krementz and Pendleton 1990). The technique has been recommended as the best non-invasive means for studying body condition in small passerines as long as sample sizes are large, observer variability is low, and exact prediction of the fat content for individual birds is not required (Conway et al. 1994).

In fact, it is known that fat scoring does vary among individual observers, and some authors caution against the pooling of fat data from different sources unless observer effects can be addressed in analysis (Krementz and Pendleton 1990, Rogers 1991). Many banders collect fat data with little or no intention of analyzing the data themselves, presumably in hopes that others may find their data of use. However, this will not even be a possibility if fat scores are too variable to permit pooling of data from different observers. If banders are going to collect fat data at all, they should strive towards a standard of measurement that will ensure consistency of scoring among observers.

Another obstacle to pooling fat scores from different sources is that different observers use different scoring systems. In compiling weight data for small passerines from 13 bird observatories across Canada (Dunn 2002), I discovered that eight different fat scoring systems were being used (Table 1). Most were either 4 class or 6 class scoring scales, but there were many different definitions for each class. Often data could be translated from one system to another, but in some cases they could not, precluding the possibility of even cursory comparisons of data across wide geographic areas. I suspected that the many variants were the result of each banding station developing its own interpretation and simplification of published fat scoring systems, in which case the published systems must not be very precise in their definitions.

This suspicion stimulated me to check the literature for published scoring scales and definitions of each fat class, with the aim of making recommendations on the best system to use. My literature search was far from exhaustive, and this review presents only enough examples to illustrate my main conclusions.

Table 1. Fat Scoring Scales in Use at 13 Bird Observatories in Canada.								
				Definition of Fat Class				
Number Sites Using This Scale	**0**	**T (trace)**	**1**	**2**	**3**	**4**	**5**	**6**
2	0		T-1/3	1/3-2/3	2/3-almost full	full	overflow	
1	0-T		<1/3	1/3-2/3	2/3-almost full	full	overflow	
2	0		T (<5%)	<1/3	1/3-2/3	2/3-full	overflow	greatly overflow
3	0	T	<1/3	1/3-2/3	2/3-overflow			
1	0	T	<1/4	1/4-full	overflow			
1	0	<1/3	1/3-2/3	2/3-full	overflow			
2	0	T	<1/2	1/2-full	overflow			
1	0	T	<1/3	1/3-full	overflow			

RESULTS

Scoring systems examined ranged from entirely qualitative, such as Mueller and Berger's (1966) 4 class scale, in which 0 = "no fat," 1 = "light fat," 2 = "moderate," and 3 = "heavy"; to a more quantitative 9 class scale with 31 sub-classes, proposed by Kaiser (1993).

Definitions for fat class in a selection of scoring systems are shown in Table 2. Most schemes assigned a single fat score for furcular and abdominal fat levels combined, or for fat in the furculum alone. (The latter system appears to have been adopted by many North American bird observatories; e.g., Table 1, Cherry 1982). Diamond (1974) followed a different approach, assigning three-class scores separately for fat stores in the furculum, axilla, abdomen, and rump. Scores were then added together for a total score, which could range from 0 -12.

Independently, Andreas Kaiser also conducted a comparison of fat scoring systems that were in use, including many European examples, and found a similar degree of variation among banding studies (A. Kaiser, unpubl. data).

DISCUSSION

Fat scoring will be most valuable if data can be compared directly among data sets. This will be possible only if everyone uses the same scoring system, or if scores can be converted to other scales with a high degree of confidence. In deciding on the best system of fat scoring to use, several criteria should be considered, each of which is discussed below.

Value to study of energy stores - For studies that pool data from many banders, bird weights alone provide a wealth of information on body condition, whether or not weights are adjusted for body size (e.g. Dunn 2002, Jones et al. 2002). Measurement of weight is far less subjective than is fat scoring, and usually can be done more quickly and with less handling. Nonetheless, weight cannot substitute for fat score for all research purposes (Rogers 2003), and under some field conditions it may be easier to score fat than to take weights. Kaiser (1993) reviewed other means of determining fat stores in birds and argued persuasively for the suitability of high quality fat scoring as the most useful field technique for wide-spread use.

Studies with specific goals of analyzing fat levels most often used the system of Helms and Drury (1960), or other 6-9 class systems that consider both furcular and abdominal fat (Nolan and Ketterson 1983, Krementz and Pendleton 1990, Kaiser 1993, Ralph et al. 1993). Assigning separate scores to each body region and then averaging the two scores provides better precision of fat score estimates (Meijer et al. 1994). Moreover, evidence is accumulating that species vary in the body regions in which fat first accumulates, and this can be studied only if scores are recorded separately for each bird. Systems with six or more fat classes allow better resolution of fat stores than do systems with four classes, and therefore are more suitable

Table 2. A Selection of Fat Scoring Schemes from the Literature. Some Definitions Have Been Shortened Slightly for Tabular Presentation.

Scheme	0	0.5	1	2	3	4	5	6	7	8
Helms and Drury (1960)	F:none or trace; A: none		F:present, but region deeply concave; clavicles visible; A:trace	F:filling, but region still concave; some covering of clavicles; A:not covering; some between intestinal folds and/or in small patches	F: filled, clavicles covered; some overflowing, but area still concave; A:covering pad, but not markedly mounded	F: filled, fat nearly level with pectoralis muscle, overflowing up interclavicles and over anterior pectoral muscles; A:mounded pad, becoming distended	F: convex pad, oveflowing length of the furculum, but never extending over ventral pectoral muscles; A: greatly distended mound			
Ralph et al. (1993)	F: no fat, region concave; A: no fat		F:trace, deeply concave, scattered patches, <5% filled; A:none, or trace	F:thin layer, <1/3 filled; A:trace or thin layer	F: 1/2 filled in small patches; A:small patches, not covering some areas	F: >2/3 filled, level with clavicles; A: covering pad, slightly mounded	F: slightly bulging; A: well mounded	F: bulging greatly; A:greatly distended mound	very large fat pads of furculum and abdomen meet	
Kaiser (1993)[1]	F: none to wedge; A: none to small trace, patchy		F:wide wedge to almost completely layered; A:trace to wide stripes (2 mm)	F:completely covered but deeply concave; A:strips of visceral fat completely filling area between intestinal loops, to small pad with at least 2-3 intestinal loops still visible	F:some fat covers ends of interclavicles, to almost filled but still concave; A:flat pad, one loop still visible, to bulge	F:filled to distal portion of interclavicles, to slightly bulging with central concavity; A:conspicuous bulge (2-4mm) to complete covering of abdominal structures	F:convex bulge to covering border of flight muscles a few mm; A:extreme convex bulge to covering border of flight muscles a few mm	F:fat covers flight muscles by several mm, to almost 3/4 A:covers abdominal part of flight muscles by several mm to almost 3/4	3/4 to nearly all of flight muscles covered	fat completely covers flight muscles; ventral side of bird completely covered

Table 2 (cont'd). A Selection of Fat Scoring Schemes from the Literature. Some Definitions Have Been Shortened Slightly for Tabular Presentation.

Scheme	0	0.5	1	2	3	4	5	6	7	8
Nolan & Ketterson (1983)	no visible fat in F or A		F:<33% full; A:<50%full	F:33-66% full; A:50-100% full but fat layer not even with pectoral region	F&A: filled and fat flush with pectoral musculature, but no bulging	as 3, but with F or A bulging	both F & A bulging			
Krementz & Pendleton (1990)	no visible fat in F or over A		trace in F or A but neither completely lined	F lined but not bulging; little fat on A	F full; some fat on A but not full	F and A full to bulging	F and A full with fat extending across apex of sternum			
Petterson & Hasselquist (1985)[1]	no visible fat in F or A		F:no fat; A:little fat between intestinal loops	F:small amount in bottom of pit; A: bands under and between intestines	F:bottom of pit filled; A:intestinal loops embedded in fat; some bluish loops and liver still visible	F: filled, but no fat outside pit; A:belly covered, liver partly visible	F:convex bulge, some fat outside pit; A:belly covered, often liver also	F:bulging fat covers pit and surrounding areas; trachea embedded in fat; A:whole belly covered with thick fat layer, noticeable edge against breast		
Cherry (1982)	F:none	F:trace	F:solid sheet of fat, concave	F:full of fat	F:fat bulging, but not meeting fat of abdomen	F:fat bulges and meets fat of abdomen				
Diamond (1974)[2]	none		little fat	substantial amounts fat, flush or nearly so with clavicles in furculum and forming a continuous sheet in other areas	fat mounded, stretching skin and bulging					

[1] Publication includes illustrations and description of colors to look for.
[2] Scoring done separately for furculum, axilla, abdomen, and rump, then summed for a possible total score of 0-12.
F = Furcular A = Abdominal

for research purposes if it can be shown that individual variability in scoring is low enough to justify having the extra classes. Kaiser (1993) found that average inter-observer variability in scores using his 9 class and 31 sub-class system was 1-2 sub-classes, indicating that high resolution can be obtained with good training.

Low observer variability - Given that there is wide individual variation in fat scoring (Krementz and Pendleton 1990), it is important to adopt a scoring system that leaves as little room as possible for interpretation. Unfortunately, most fat scoring scales define classes in qualitative terms, such as "furculum <1/3 filled." Such a description provides considerable latitude for individual judgment on how much constitutes "one third."

The greatest strength of the system developed by Kaiser (1993) is that there are very detailed, non-qualitative descriptions of each class (including discussion of tissue color, not included in Table 2). Kaiser's definitions for sub-classes within each fat class (two to four per class), provide guidance as to the low and high ends of each class. Even if scores are recorded only by class, having descriptions of sub-classes for reference increases the probability that observers will be consistent in assigning class scores. Kaiser also provided illustrations of the fat patches for each class that are more realistic and less ambiguous to interpret than the cross sections of the furculum that are often used by bird observatories (see illustration in North American Banding Council 2001). This, too, should contribute to consistency of scoring.

Ease of use - If a fat scoring system is too complex for banders to use easily, then it would not be used at all. This argues for ease of use as a criterion for selecting a system, and it is probably for reasons of user-friendliness that most fat scoring scales have short descriptions for each class that provide little detail or qualification (e.g. Table 2). On the other hand, there is no point in collecting fat score data at all if the results are not of good quality, so ease of use should not be the sole criterion for selecting a scoring system.

While Kaiser's (1993) 31-sub-class system is clearly the most complex of all those discussed here, it has been found that experienced banders learn to use it quickly, after handling only 10 to 20 birds (A. Kaiser, pers. comm.; although many more birds normally have to be handled before examples can be found of every fat class). Difficulty of use is reduced if birds are put into one of the nine main fat classes, rather than into a sub-class. Although more initial training may be required for Kaiser's system than for some others, that training will be less than is required for many other bander skills (such as extracting birds from nets or aging them by plumage), and the system is easy to use once learned.

CONCLUSIONS AND RECOMMENDATIONS

Fat scores have the potential to be a useful piece of information on the energetic condition of birds, but that potential can be realized on a large scale only if there is greatly improved consistency in the methods used to score fat. Kaiser's (1993) system is currently the best available in terms of encouraging consistency among observers, because of the detail it provides on end points of each fat class. Investigators of energy stores could use this system to gather detailed data (scoring furculum and abdomen separately and recording sub-classes), while other banders might prefer using it to record a combined furcular/abdominal score, by class only.

Although Kaiser's system is used widely in Europe, it is scarcely used in North America. It may, therefore, be preferable to improve one of the scoring scales most commonly in use here (possibly Ralph et al. 1993) by using the illustrations and descriptions provided by Kaiser to better define the fat classes. For example, if fat class 2 in the Ralph et al. (1993) system was deemed to be equivalent to Kaiser's sub-classes 1.75 to 2.25, then Kaiser's drawings and descriptors for those sub-classes could be adopted as the end-points of that fat class in the Ralph et al. system. Those drawings and descriptions should then be posted in banding rooms and printed in handbooks for easy reference. Even with these improvements, however, it will be important to check and compare observers regularly to ensure that scorers are not developing individual styles.

Finally, if a bander is reluctant to use an improved fat scoring system, consideration should be given to weighing birds but not scoring fat. In either case, the bander should record the time of day at which the bird was caught, as daily variation in weight and/or fat is often one of the variables of interest in studies of energetics.

ACKNOWLEDGMENTS

Andreas Kaiser and C. J. Ralph kindly reviewed the manuscript and gave feedback that greatly improved the final paper. Kaiser also provided unpublished information cited in the text and the Appendix, providing more detail on his system and on his experience with its use.

LITERATURE CITED

Cherry, J. D. 1982. Fat deposition and length of stopover of migrant White-crowned Sparrows. *Auk* 99:725-732.

Conway, C. J., W. R. Eddleman, and K. L. Simpson. 1994. Evaluation of lipid indices of the Wood Thrush. *Condor* 96:783-790.

Diamond, A. W. 1974. Annual cycles in Jamaican forest birds. *J. Zool.* 173:277-301.

Dunn, E. H. 2002. A cross-Canada comparison of mass change in birds during migration stopover. *Wilson Bull.* 114:368-379.

Helms, C. W. and W. H. Drury. 1960. Winter and migratory weight and fat field studies on some North American buntings. *Bird-Banding* 31:1-40.

Jones, J., C. M. Francis, M. Drew, S. Fuller, and M. W. S. Ng. 2002. Age-related differences in body mass and rates of mass gain of passerines during autumn migratory stopover. *Condor* 104: 49-58.

Kaiser, A., 1993. A new multi-category classification of subcutaneous fat deposits of songbirds. *J. Field Ornithol.* 64:246-255.

Krementz, D. G. and G. W. Pendleton. 1990. Fat scoring: sources of variability. *Condor* 92:500-507.

Meijer, T., F. J Möhring, and F. Trillmark. 1994. Annual and daily variation in body mass and fat of Starlings *Sturnus vulgaris*. *J. Avian Biol.* 25:98-104.

Mueller, H. C. and D. D. Berger. 1966. Analyses of weight and fat variations in transient Swainson's Thrushes. *Bird-Banding* 37:83-112.

Nolan, V. Jr. and E. D. Ketterson. 1983. An analysis of body mass, wing length, and visible fat deposits of Dark-eyed Juncos wintering at different latitudes. *Wilson Bull.* 95:603-620.

North American Banding Council. 2001. The North American banders' study guide. North American Banding Council, Point Reyes Station, CA (available through http://www.nabanding.net/nabanding/pubs.html)

Petterson, J. and D. Hasselquist. 1985. Fat deposition and migration capacity of Robins *Erithacus rubecula* and Goldcrest *Regulus regulus* at Ottenhy, Sweden. *Ringing and Migration* 6:66-76.

Ralph, C. J., G. R. Geupel, P. Pyle, T. E. Martin, and D. F. DeSante. 1993. Handbook of field methods for monitoring landbirds. USDA *Gen. Tech. Rep.* PSW-144. USDA Forest Service, Pacific Southwest Research Station, Albany, CA.

Rogers, C. M. 1991. An evaluation of the method of estimating body fat in birds by quantifying visible subcutaneous fat. *J. Field Ornithol.* 62:349-356.

Rogers, C. M. 2003. New and continuing issues with using visible fat classes to estimate fat scores. *J. Avian Biol.* 34:129-133.

Scott, I. A., P. I. Mitchell, and P. R. Evans. 1995. The reliability of fat scores as predictors of the mass of fat carried by individual birds. *Ardea* 83:359-363.

APPENDIX

Andreas Kaiser (pers. comm.) offered the following additional details of the scoring system as described in Kaiser (1993).

(1) To separate scores 3.75-4.25, check to see if the liver is visible. At score =3.75, the liver can always be seen; at score=4.0, it can sometimes be seen; and at score=4.25, the liver is obscured by abdominal fat.

(2) In Fig. 2 of Kaiser (1993), the broken line in the illustration for class 5 represents the limit of the area covered by fat, even though the area is not shaded in the diagram. Likewise, the drawing for class 6 does not show shading above the furculum, even though fat for birds in this class will be present over the entire area between the furculum and neck. (The drawing was done this way to emphasize the differences between classes in the amount of fat around the furculum.)