

## The use of regression equations for estimation of arthropod biomass in ecological studies

José A. Hódar

*Departamento de Biología Animal y Ecología, Facultad de Ciencias,  
Universidad de Granada, E-18071, Granada, Spain.  
Fax +3458 243238,  
e-mail jhodar@goliat.ugr.es*

Received: 31.5.96

Accepted: 30.9.96

### Abstract

The aim of this work is to analyse the use and applicability of different series of equations published for estimating biomass from length data in Arthropoda. A set of length-weight regression equations was developed for 53 groups of arthropods, as well as general equations for Arthropoda and Hexapoda, from a collection carried out in an arid zone of southeastern Spain. The significance of regression equations performed was generally very high, as well as  $R^2$  values, giving to these equations a high predictive capacity. From the same collection, a sample was selected to test the applicability of the different equation series published by other authors. Although there were no significant differences between the equations, results in crude estimation both of overall biomass and their distribution among the taxonomic groups present in the sample showed strong variations. These variations can strongly bias studies in which biomass is the main variable, such as those focused on availability, energetics and macroecology. Comparison among available equations suggests that these must be applied with caution and after testing the different applicabilities of available references. The use of specific equations is recommended over general equations, as well as the compiling of more new equation series.

**Keywords:** Arthropoda, biomass estimation, body size, body weight, ecological methodology, Hexapoda, regression equations.

### Résumé

Le but de ce travail est d'analyser l'emploi et l'applicabilité de différentes séries d'équations publiées pour estimer des biomasses à partir de longueurs en Arthropoda. Une série de régressions longueur-biomasse a été développée pour 53 groupes d'Arthropoda et Hexapoda, à partir d'une collection effectuée dans un site aride du sud-est de l'Espagne. La signification des équations de régression élaborées est généralement très élevée, ainsi que leurs valeurs de  $R^2$ , donnant aux équations une haute capacité prédictive. Un échantillon a été sélectionné dans la même collection pour y tester l'applicabilité de différentes séries d'équations publiées par d'autres auteurs. Bien qu'il n'y ait pas de différences significatives entre les équations, les résultats bruts d'estimation, tant en biomasse totale que dans la distribution de celle-ci parmi les différents groupes taxonomiques présents, ont montré des variations importantes. Celles-ci peuvent produire des erreurs notables dans les études qui prennent la biomasse comme variable essentielle, telles que celles qui s'intéressent à la disponibilité tant énergétique que macroécologique. La comparaison avec les équations disponibles suggère que

celles-ci doivent être appliquées avec précaution et après avoir testé le degré d'applicabilité de différentes références. Il est recommandé d'employer des équations spécifiques de préférence aux équations générales et d'élaborer de nouvelles séries d'équations.

## INTRODUCTION

Estimation of biomass in living organisms is a current necessity in many ecological studies, since it influences many physiological and behavioural processes in the individual (PETERS, 1983; SCHMIDT-NIELSEN, 1984; ROGERS & RANDOLPH, 1986; WILLIAMS, 1989), determines distribution and density patterns at the species level (PETERS & WASSENBERG, 1984; DAMUTH, 1991; GRIFFITHS, 1992; BLACKBURN *et al.*, 1993a; COTGREAVE, 1993), and is an appropriate estimate in studies on the economy of ecosystems, that is, matter and energy fluxes and resource distribution among species (HARVEY & GODFRAY, 1987; PAGEL *et al.*, 1991; BLACKBURN *et al.*, 1993b). The importance of this parameter has been stated in many recent works that have used it as a basis for macroecological approaches, in which energy and space partitioning between species are analysed (*e.g.*, MAURER & BROWN, 1988; BROWN & MAURER, 1989; PAGEL *et al.*, 1991; reviews in LAWTON, 1991; COTGREAVE, 1993). Nevertheless, most of these studies restrict their analysis to birds and mammals, because in these groups biomass is available in the literature (see *e.g.* table I in PAGEL *et al.*, 1991). Only some recent works (GASTON & LAWTON, 1988; STORK & BLACKBURN, 1993; BLACKBURN *et al.*, 1993b) analyse arthropod communities with the same approaches used in birds and mammals.

The diversity of species in arthropod communities implies that the species composition is different in almost every new research project, and consequently data on body size in most cases cannot be obtained from the literature: they have to be taken from the field. Biomass estimates have been usually made by weighing several individuals belonging to the same taxa, calculating a mean and then multiplying by the number of individuals collected for this taxa (ROGERS *et al.*, 1977; SAMPLE *et al.*, 1993). This procedure is tedious, and does not reflect the variance in weight for the different individuals. Individual weighing is not always a reliable option: the handling of a high number of individuals for measuring and/or weighing can be a time-consuming process, and sometimes their death is required, as in dry-weight measurements, which consequently prevents for studies with live specimens. But arthropods are ubiquitous, numerous, and serve as an important food resource for many types of organisms. Insects can ingest up to 80% of the primary production and represent 90% of secondary production in ecosystems (PRICE, 1984, and references therein). For these reasons, the ability to estimate arthropod biomass is a strong priority, and a valuable tool for many types of ecological studies.

Traditionally, biomass estimates have involved allometric indices or other measurements (HERRERA, 1978; SMILEY & WISDOM, 1982; CIBOROWSKY, 1983; GRIFFITHS, 1986). The most common procedures have been an estimation from the mean of each taxonomic group (a procedure that ignores the individual weight variance) or the application of regression equations from lineal measurements, either general for all arthropods (or all insects, ROGERS *et al.*, 1976) or specific for each taxonomic group (ROGERS *et al.*, 1977; SCHOENER, 1980; GOWING & RECHER, 1984; DIAZ & DIAZ, 1990; SAMPLE *et al.*, 1993). This last method seems to be the most accurate, quickest and easiest to apply. However, equations for all taxonomic

groups may be unavailable in references, and doubt concerning bias could arise when equations are made with specimens from different zones (SCHOENER, 1980; GOWING & RECHER, 1984). Despite these problems, many works have used the equations available in literature for several purposes, without appropriate caution on the effects of a careless application of these estimations.

The aim of this work is to examine these problems, specifically addressing the following questions: Are the available equation series similar and equally applicable to new data series? What tests can be made in order to confirm the applicability of available equation series to new series of data? When the equation series available are not applicable, is it advisable to formulate new equation series and, if so, what criteria should be followed? To analyse these problems, I formulated a new series of equations and obtained a reference sample from the same pool used to perform the equations. Then, I used the results of the application of different tests to compare equation series, as well as the crude results in the estimation of the sample biomass against the actual biomass.

## MATERIALS AND METHODS

### Sampling procedure

Adults, larvae and nymphs of many types of arthropods were caught during the period 1990-1992 in three different zones of the Guadix-Baza Basin (Granada province, southeastern Spain), including shrubsteppes, cereal crops, fallow lands and cleared oakwoods; a few specimens were caught in other zones within Granada province. For an entire year, the sampling procedure was carried out simultaneously in the three zones, including (expressed as samples/month/zone) pitfall traps (72, capture period 48 h), vegetation beating (100), baited traps with excrement (10) and carcasses (2), light traps (2-4), and direct capture (see for details SANCHEZ-PINERO, 1994). The representativity of sampling universe was ensured by combining several capture methods, covering different zones and microhabitats, throughout the year and at all times of day. In addition, the taxonomic complexity of the sample was also increased by the high diversity in each zone (SANCHEZ-PINERO, 1994), and only a few groups which are strictly aquatic during some phases of their life (*e.g.* Ephemeroptera, Trichoptera) were absent from the sample.

All the specimens were stored in Scheerpeltz preservative prior to measurements. This preservative can alter the external biometry of certain types of arthropods, especially soft-bodied groups like Araneae or Lepidoptera larvae. Nevertheless, I chose Scheerpeltz because this method of preservation, and others similar, such as 70% alcohol, are widely used by entomologists (ROGERS *et al.*, 1977; GOWING & RECHER, 1984; DIAZ & DIAZ, 1990), as opposed to live collecting or freezing (SAMPLE *et al.*, 1993) which more accurately preserves the live dimensions of the specimens.

### Laboratory procedure

I selected 48 groups, the numerically most important in the arthropod community sampled, using a taxonomic criteria (following BARRIENTOS, 1988 and RICHARDS & DAVIES, 1984), with the order level as a basis for establishing the groups but taking into account morphological variations within taxonomic groups, since pooling individuals with different body shapes can diminish the predictive power of equations (SCHOENER, 1980). For example, the order Heteroptera was divided in two sub-groups labelled *H* and *S*, the former for insects with heavy and wide bodies (*e.g.* Cydnidae, Pentatomidae, Scutelleridae) and the latter for those with slender and soft bodies (*e.g.* Lygaeidae, Miridae, Reduviidae). I used the term OTU (Operational Taxonomic Unit, sensu SNEATH & SOKAL, 1973) to define the 48 groups. In hemimetabolous insects, each OTU included several phases of the development –that is, nymphs and imago pooled (*e.g.* Heteroptera)– whereas in holometabolous (*e.g.* Lepidoptera, Coleoptera), larvae and imago were separated. I also calculated other equations for several OTUs together into the same Order: Orthoptera, Heteroptera, Diptera, Lepidoptera imago and Coleoptera.

The specimens included in the regression calculations were arbitrarily selected within the body-size gradient of the appropriate OTU, in an effort to cover the gradient as widely as possible. Each OTU included 6 to 26 (usually 10) specimens for the regression calculation. The body length of the individuals (hereafter  $L$ ) was measured (in mm) mainly with a binocular microscope 10-40 $\times$  equipped with an ocular micrometer, using a digital caliper only with large and heavy specimens, both instruments with 0.05 mm precision.  $L$  was always taken excluding appendages such as antennae, ovipositors or wings. In Scorpionida,  $L$  was measured as prosoma length. After the measurements, specimens were dried in a stove at 70°C for 24 hours and then dry weight (hereafter  $W$ ) was obtained by using an electronic balance (in mg, precision 0.01 mg). Extremely small specimens were grouped in packs of 5-10 of the same size and weighed together, in order to obtain a more accurate reading for the individual weight.

### Statistical analysis

Regressions were performed with data transformed to logarithms. This procedure usually reduces heteroscedasticity to the data (EDWARDS, 1985; ZAR, 1996), and in biomass estimations, as in this case, allows the conversion of a power equation  $W = aL^b$  in a linear regression (ROGERS *et al.*, 1976, 1977; SCHOENER, 1980; SAMPLE *et al.*, 1993). Equation series constructed to date fall into two basic groups: a single equation from a pool of diverse arthropods (hereafter general equations), or an equation for each separated taxonomic group of arthropods (hereafter specific equations). I have constructed both types of equations from my sample (see also GOWING & RECHER, 1984; SAMPLE *et al.*, 1993). Hereafter, I shall refer to all these as Guadix-Baza equations.

To examine the effect of sample size on the significance of equations performed, I calculated the sample size required for a significant correlation  $\rho \neq 0$  using as basis the  $R^2$  values obtained in the regressions (ZAR, 1996). To compare within the Guadix-Baza equations, I used ANCOVAs, in which the classification factor was the OTU, and body length the covariate. To compare with other equations in the literature, I used Z-tests (DOWDY & WEARDEN, 1991; see also SAMPLE *et al.*, 1993), comparing the slopes of the regressions and the intercept when the slopes were non-significant (ZAR, 1996). References used in this comparison were: ROGERS *et al.* (1976, 1977), SCHOENER (1980), GOWING & RECHER (1984), DIAZ & DIAZ (1990) and SAMPLE *et al.* (1993). I compared both general equations as well as specific equations when this was allowed by the taxonomic definition of the group.

To examine the reliability of crude estimates from the different series of equations, I selected a subsample of the arthropods caught in 864 pitfall traps set over an entire year in one of my sampling zones (the Barranco del Espartal, Baza; see SANCHEZ-PIÑERO, 1994). I randomly selected 24 pitfalls with more than six specimens greater than 1 mm  $L$ , in order to ensure a good sample size and to minimize the biases during measurement and weighing. None of the specimens belonging to this subsample was used in the formulation of the equations. I statistically compared these actual data of weight with the estimates obtained by using the equation series of the following references: ROGERS *et al.*, (1976, 1977), SCHOENER (1980, only the general equation), GOWING & RECHER (1984), SAMPLE *et al.* (1993), and the Guadix-Baza equations. In the comparison performed with data registered as individuals ( $n = 347$ ) I used paired  $t$ -test (ROGERS *et al.*, 1977; SAMPLE *et al.*, 1993) and Pearson  $r$  correlation coefficient. I also performed comparisons with data registered by taxonomic groups, by using paired  $t$ -test, Wilcoxon paired test and Pearson  $r$  correlation coefficient.

The significance of the tests was always corrected with post-hoc Bonferroni sequential adjustments (RICE, 1989).

## RESULTS

### General equations

A total of 526 specimens were used in the formulating of the equations. The application of the regression to all used specimens gave a general equation

for Arthropoda of  $W = 0.0397 L^{2.393}$  ( $R^2 = 0.83$ ,  $SEE = 0.972$ ,  $p < 0.0001$ ;  $SE_{lna} = -3.227$ ,  $SE_b = 0.049$ ). The  $b$  value was low in comparison with ROGERS *et al.* (1976) and SAMPLE *et al.* (1993) values, and higher than those of SCHOENER (1980). These references offer general equations only for Hexapoda. Therefore, I divided Arthropoda into non-Hexapoda and Hexapoda, and repeated the procedure; the new equations were  $W = 0.0935 L^{2.041}$  ( $R^2 = 0.82$ ,  $SEE = 1.051$ ,  $n = 78$ ,  $p < 0.0001$ ;  $SE_{lna} = -2.370$ ,  $SE_b = 0.110$ ) and  $W = 0.0315 L^{2.492}$  ( $R^2 = 0.83$ ,  $SEE = 0.923$ ,  $n = 448$ ,  $p < 0.0001$ ;  $SE_{lna} = -3.457$ ,  $SE_b = 0.053$ ), respectively. Although the two equations did not differ significantly ( $F = 0.63$ ,  $df = 1, 523$ ,  $p > 0.4$ , ANCOVA), the effect of non-Hexapoda diminished the  $b$ -value of the total equation. In Hexapoda,  $b$  remained almost identical to SAMPLE *et al.* (1993) and closer to ROGERS *et al.* (1976) and GOWING & RECHER (1984), and without significant differences (table I); however, it differed ( $Z = -3.36$ ,  $p < 0.01$ ) from that of the total general equation and from two out of the other three general equations (table I) offered by SCHOENER (1980).

### Specific equations

I constructed 53 specific equations estimating  $W$  from  $L$  (table II). The  $R^2$  values obtained were in general high, with only 17 below  $R^2 = 0.90$ , and only 3 below of  $R^2 = 0.81$  ( $r = 0.90$ ). This gives the equations a high predictive capacity. The significance of regressions is also very high, with only 1 equation up to  $\alpha = 0.05$  table-wide level (RICE, 1989), and 50 below  $\alpha = 0.01$ . That is, 98% of the Guadix-Baza equations are significant after a Bonferroni adjustment.

As a whole, the equations obtained showed significant differences ( $F = 27.84$ ,  $df = 47, 477$ ,  $p < 0.0001$ , ANCOVA), indicating that the length-weight relationships clearly differed between the OTUs considered (but see non-Hexapoda vs. Hexapoda, above). On the other hand, comparisons with the equations offered by other authors showed rare, in general nonsignificant, differences (table I), although some cases such as Araneae, Curculionidae or Neuroptera, differed strongly in parameters which were not supported by significant differences. Tenebrionidae using SAMPLE *et al.*'s (1993) equations, and Diptera, Lepidoptera and Coleoptera using SCHOENER's (1980) equations were the only comparisons that proved at all significant. The case of Hymenoptera non-Formicidae is noteworthy, with the  $b$  term significantly different only employing SAMPLE *et al.* (1993), and the  $a$  term significantly different using all the other equations except DIAZ & DIAZ (1990).

Calculations of the minimum significant sample size with a given  $r$  value (ZAR, 1996), taking the  $r$ -values obtained in the regressions as the basis, indicated that an  $n$  value relatively low can be good enough to perform reliable equations. If we wish to reject  $H_0 : \rho = 0$ , with a 99% probability of correctly rejecting  $H_0$ , and a  $r$  value  $> 0.90$  (94.3% of Guadix-Baza equations), we need an  $n = 12$  at  $\alpha = 0.05$ , and  $n = 15$  at  $\alpha = 0.01$ ; and with  $r > 0.95$  ( $r^2 \approx 0.90$ , 67.9% of Guadix-Baza equations),  $n$  are 9 and 10, respectively. That is, an  $n$  value of between 10 and 20 specimens appears to be adequate, and  $n = 10$  may be the minimum recommended.

TABLE 1. – Comparisons of the slope parameter and the intercept ( $\pm$ SE) obtained in the Guadix-Baza equations with that offered by ROGERS *et al.* (1976, 1977), SCHOENER (1980), GOWING & RECHER (1984), DIAZ & DIAZ (1990) and SAMPLE *et al.* (1993). In SCHOENER (1980) comparisons were performed with equations from the three different zones offered by the author (Mass.=Massachusetts, Guap.=Guapiles). Figures in italics denote significant differences ( $p < 0.05$ ) against Guadix-Baza estimates.

General	2.620 ± 0.030 – 3.490 ± 0.055 2.929 ± 0.294 – 3.106 ± 0.189	2.350 ± 0.061 – 3.730 ± 0.058	2.160 ± 0.046	2.110 ± 0.066	2.362 ± 0.094 – 3.466 ± 0.168 2.911 ± 0.056 – 3.311 ± 0.122 2.543 ± 0.172 – 4.591 ± 0.517 2.180 ± 0.121 – 4.049 ± 0.329 3.160 ± 0.343 – 4.582 ± 0.585	2.494 ± 0.032 – 3.628 ± 0.073 2.296 ± 0.202 3.909 ± 0.470 2.626 ± 0.271 2.844 ± 0.204 2.686 ± 0.187	
Araneae					2.669 ± 0.353 – 3.106 ± 0.741		
Diplopoda							
Chilopoda							
Isopoda							
Orthoptera	2.512 ± 0.105 – 3.020 ± 0.284 2.696 ± 0.046 – 3.308 ± 0.062 2.270 ± 0.081 – 2.998 ± 0.113 2.366 ± 0.078 – 3.293 ± 0.115 2.809 ± 0.146 – 5.137 ± 0.319	1.650 ± 0.227 – 1.427 ± 0.191 2.650 ± 0.136 – 3.772 ± 0.064 2.480 ± 0.279 – 3.352 ± 0.218 1.640 ± 0.121 – 3.817 ± 0.056	1.960 ± 0.408 – 2.477 ± 0.273 2.230 ± 0.180 – 3.507 ± 0.066 2.280 ± 0.354 – 3.612 ± 0.197 1.590 ± 0.116		2.714 ± 0.363 – 4.828 ± 1.070	2.225 ± 0.158 – 2.823 ± 0.273 3.075 ± 0.147 – 4.784 ± 0.313 2.213 ± 0.085 – 3.184 ± 0.184 2.959 ± 0.085 – 5.909 ± 0.257 3.122 ± 0.064 – 5.036 ± 0.157 1.200 ± 0.247	2.492 ± 0.081 – 3.247 ± 0.175 2.696 ± 0.083 1.900 ± 0.154
Homoptera		2.310 ± 0.226 – 3.730 ± 0.108 3.130 ± 0.335 – 5.116 ± 0.082 2.420 ± 0.096 – 3.817 ± 0.056				2.410 ± 0.412 – 2.919 ± 0.852 2.448 ± 0.448 – 3.650 ± 0.974 1.666 ± 0.543 – 2.513 ± 1.548	2.354 ± 0.247 2.668 ± 0.253 2.392 ± 0.158 2.571 ± 0.180 2.969 ± 0.170 2.669 ± 0.150
Heteroptera							
Diptera							
Lepidoptera larvae							
Lepidoptera imago		2.550 ± 0.528 – 4.269 ± 0.173	2.500 ± 0.146 – 3.650 ± 0.090	1.320 ± 0.121 – 2.551 ± 0.111			
Tenebrionidae	2.681 ± 0.080 – 3.050 ± 0.204						
Curculionidae	3.050 ± 0.164 – 3.915 ± 0.245						
Coleoptera tot.	2.790 ± 0.050 – 3.460 ± 0.105	1.990 ± 0.134	2.110 ± 0.086	1.910 ± 0.096	2.689 ± 0.148 – 3.306 ± 0.258	2.492 ± 0.147 – 3.297 ± 0.378	2.640 ± 0.080 1.900 ± 0.154
Hymenoptera	2.407 ± 0.097 – 3.817 ± 0.108	2.550 ± 0.106 – 4.135 ± 0.071	2.070 ± 0.091 – 3.147 ± 0.051	2.290 ± 0.136 – 3.817 ± 0.056	2.104 ± 0.078 – 3.336 ± 0.127		
Formicidae	2.572 ± 0.097 – 4.029 ± 0.171	2.190 ± 0.304 – 3.381 ± 0.140	2.720 ± 0.246 – 4.423 ± 0.108	2.310 ± 0.209 – 3.863 ± 0.119	2.489 ± 0.051 – 3.996 ± 0.101	2.208 ± 0.257 – 3.730 ± 0.405	2.919 ± 0.171 – 4.727 ± 0.350

TABLE II. – Parameter estimations for regression of weight (mg) on body length (mm),  $W = aBL^b$ .  
 $n$  = sample size,  $R^2$  = coefficient of regression,  $SEE$  = Standard error of the estimate,  $a$  = intercept,  
 $b$  = slope,  $SE$  = standard errors of  $a$  and  $b$ . Min. and Max. indicate the extreme values of body length  
considered in the regression for each OTU.

OTU	$n$	$R^2$	$SEE$	$a$	$SE_{Ina}$	$b$	$SE_b$	Min.	Max.
Scorpionida	7	0.942	0.313	0.0078	1.106	3.424	0.379	10.20	26.75
Soliphuga	6	0.957	0.133	0.0578	0.829	2.318	0.285	14.30	29.70
Araneae	18	0.890	0.717	0.1044	0.374	2.296	0.202	1.30	27.10
Opilionida	10	0.791	0.474	0.0405	0.757	2.916	0.531	2.60	6.40
Acarina	7	0.934	0.595	0.0530	0.256	2.494	0.297	0.40	3.90
Isopoda	10	0.960	0.374	0.0101	0.512	2.844	0.204	4.25	22.50
Diplopoda	10	0.896	0.643	0.00012	1.487	3.909	0.470	11.00	39.00
Chilopoda	10	0.922	0.626	0.0036	0.906	2.626	0.271	10.00	81.00
Collembola	8	0.918	0.182	0.0024	0.223	3.676	0.256	1.50	3.25
Thysanura	10	0.938	0.187	0.0747	0.292	1.601	0.146	3.00	10.00
Orthoptera	27	0.874	0.694	0.0255	0.561	2.637	0.200	4.00	63.00
Ensifera	10	0.909	0.709	0.0285	0.842	2.797	0.314	4.00	38.65
Caelifera	17	0.932	0.483	0.0163	0.538	2.686	0.187	4.70	63.00
Blattodea	10	0.897	0.327	0.0494	0.534	2.344	0.282	2.60	9.60
Mantodea	10	0.900	0.574	0.0017	1.225	2.953	0.347	16.00	66.55
Dermaptera	10	0.960	0.358	0.0015	0.659	3.497	0.251	5.50	25.20
Embioptera	10	0.989	0.148	0.0011	0.274	3.150	0.119	5.00	15.80
Psocoptera	6	0.852	0.249	0.0425	0.284	1.637	0.341	1.20	3.00
Thysanoptera	6	0.911	0.202	0.0071	0.195	2.537	0.397	1.15	2.00
Homoptera non-Aphid.	12	0.901	0.620	0.0548	0.449	2.354	0.247	2.80	25.40
Aphidae	6	0.588	0.490	0.0598	0.374	1.724	0.722	1.00	2.20
Heteroptera	21	0.855	0.510	0.0341	0.530	2.688	0.253	2.90	14.30
Heteroptera H	11	0.933	0.290	0.0399	0.534	2.802	0.251	4.90	14.00
Heteroptera S	10	0.980	0.197	0.0376	0.249	2.417	0.122	2.90	14.30
Diptera	36	0.872	0.702	0.0312	0.264	2.392	0.158	1.00	24.00
Nematocera	10	0.969	0.398	0.0210	0.233	2.081	0.132	1.00	24.00
Brachycera	26	0.954	0.387	0.0304	0.193	2.630	0.118	2.00	20.60
Lepidoptera larvae	16	0.936	0.370	0.0110	0.534	2.571	0.180	9.60	57.60
Lepidoptera imago	20	0.944	0.384	0.0095	0.455	2.969	0.170	5.00	35.00
Ropalocera	10	0.885	0.400	0.0139	0.965	2.778	0.355	6.50	22.85
Heterocera	10	0.972	0.354	0.0081	0.482	3.072	0.183	5.00	35.00
Neuroptera	10	0.702	0.716	0.0814	1.022	1.530	0.352	7.50	46.70
Carabidae larvae	10	0.898	0.488	0.0338	0.690	2.162	0.258	3.90	30.15
Scarabeidae larvae	10	0.982	0.367	0.0112	0.390	2.776	0.133	2.80	52.00
Tenebrionidae larvae	9	0.908	0.386	0.0031	1.012	2.955	0.355	8.00	26.00
Coleoptera total	156	0.878	0.673	0.0410	0.195	2.640	0.080	1.75	56.55
Carabidae	12	0.962	0.407	0.0080	0.500	3.214	0.203	4.30	26.00
Tenebrionidae	16	0.958	0.341	0.0513	0.407	2.669	0.150	4.00	38.60
Scarabeidae	10	0.986	0.192	0.0746	0.273	2.582	0.109	4.15	32.50
Chrysomelidae	10	0.918	0.511	0.0258	0.674	3.083	0.325	3.00	15.20
Curculionidae	12	0.881	0.667	0.1281	0.517	2.254	0.262	1.75	17.05
Aphodiidae	10	0.965	0.256	0.0246	0.373	2.824	0.191	3.25	13.20
Cetoniidae	10	0.923	0.321	0.0124	0.861	3.281	0.334	9.00	21.00
Dynastidae	10	0.928	0.490	0.0556	0.776	2.694	0.266	5.90	41.40
Histeridae	10	0.883	0.400	0.0568	0.704	2.706	0.348	4.20	13.60
Buprestidae	10	0.996	0.153	0.0105	0.211	3.173	0.091	4.70	27.80
Cerambycidae	10	0.997	0.293	0.0197	0.533	2.848	0.187	7.60	52.10
Meloidae	10	0.964	0.333	0.0309	0.526	2.483	0.170	8.10	56.55
Staphylinidae	10	0.918	0.820	0.0023	0.911	3.332	0.353	2.00	28.00
Coleoptera unid.	16	0.881	0.523	0.0664	0.430	2.192	0.216	2.30	14.50
Hymenopt. non Form.	24	0.874	0.644	0.1636	0.286	1.900	0.154	1.60	26.50
Formicidae workers	11	0.907	0.432	0.0104	0.524	2.763	0.294	2.70	12.30
Formicidae winged	10	0.862	0.455	0.0025	1.056	3.610	0.511	4.50	11.20

### Results of the application of the different estimations

The estimates of absolute biomass from the Baza subsample showed differences of more than fourfold from the lowest to the highest (fig. 1), depending on the series of equations used. General equations tended to underestimate the biomass with respect to specific equations of the same reference, except SAMPLE *et al.* (1993). The SCHOENER (1980) equation gave the lowest values, whereas ROGERS *et al.* (1977) specific equations gave the highest.

Comparisons with the actual values of biomass recorded in the selected sample show some contrasting results. Focusing on the value of the absolute estimate, the best results are given by the general of ROGERS *et al.* (1976), the specifics of GOWING & RECHER (1984), and the specifics of Guadix-Baza (table III, fig. 1). These results were corroborated by the *t*-tests performed on individuals, whereas comparisons by taxonomic groups gave always non-significant differences with the actual biomass (table III). The Wilcoxon test disagrees somewhat with the *t*-test, showing some significant differences where *t* gave non-significant ones. However, the *r* correlation coefficient gave a different result, showing that the best adjustment by groups between actual and estimated biomass was given by the specific Guadix-Baza equations. Figure 1 reinforces this observation, showing that the distribution of biomass among the different taxonomic groups of the sample is much more similar to the actual by using the Guadix-Baza equations than the others.

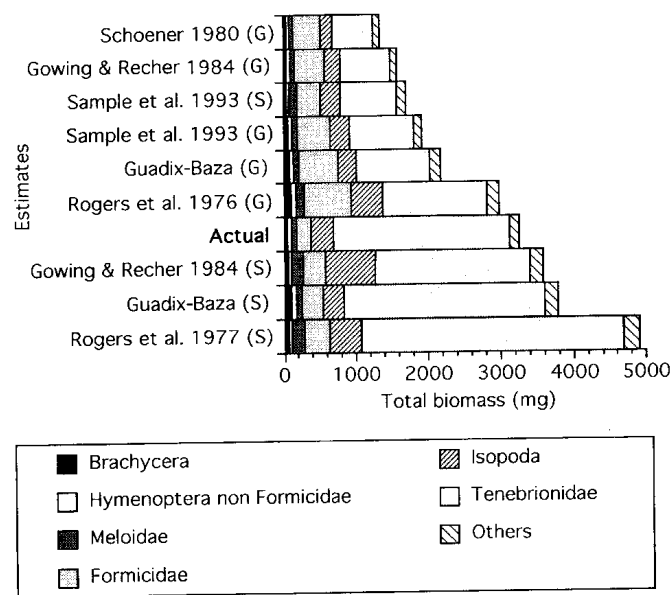


FIG. 1. – Distribution of biomass among the different taxonomic groups of the subsample, showing the actual values (=Actual) and the different estimates applied to the sample. G and S indicate general and specific equations, respectively.



TABLE III. – Comparisons between the actual biomass of a sample and the estimates obtained with the equation series of ROGERS *et al.* (1976, 1977), SCHOENER (1980), GOWING & RECHER (1984), SAMPLE *et al.* (1993) and Guadix-Baza (this study). In SCHOENER (1980) only the general equation was used. ns indicates  $p > 0.10$ .

	Rogers <i>et al.</i>		Schoener	Gowing & Recher		Sample <i>et al.</i>		Guadix- Baza	
	General Specifics		General	General Specifics		General Specifics		General Specifics	
<i>Individuals</i>									
Abs diff.	279.4	−1 635.5	1 929.2	1 681.2	−314.0	1 345.8	1 561.8	1 085.5	−519.7
Mean diff.	0.81	−4.71	5.56	4.85	−0.91	3.88	4.50	3.13	−1.50
SD	15.18	14.89	22.93	21.42	15.61	19.53	21.23	18.23	8.68
<i>t</i> -test	0.99	−5.90	4.52	4.21	−1.08	3.70	3.95	3.19	−3.21
<i>p</i>	ns	0.001	0.001	0.001	ns	0.002	0.001	0.005	0.006
Pearson <i>r</i>	0.933	0.968	0.933	0.933	0.970	0.933	0.942	0.946	0.967
<i>Groups</i>									
Mean diff.	11.66	−68.15	80.40	70.06	−13.09	56.07	65.07	45.22	−21.66
SD	93.34	61.03	141.77	133.33	70.33	122.53	126.03	113.64	27.64
<i>t</i> -test	0.25	−1.40	1.02	0.95	−0.59	0.85	0.94	0.74	−1.48
<i>p</i>	ns	ns	ns	ns	ns	ns	ns	ns	ns
Wilcoxon	ns	0.009	ns	ns	0.09	ns	ns	ns	0.038
Pearson <i>r</i>	0.940	0.934	0.937	0.939	0.951	0.940	0.944	0.949	0.977

## DISCUSSION

The results clearly show important variations between the different equation series for estimating arthropod biomass. This does not mean that the available equation series are not correct or precise, but simply that they are not equally applicable to any arthropod sample. The problem usually argued in this sense is the degree of accuracy with which equations developed in a zone can be applied to arthropods from other zones. Although significant differences among compared equations are rare (SAMPLE *et al.*, 1993; table I), some parameters in equations reveal important contrasts between zones (see SCHOENER, 1980 for a detailed reasoning of this phenomenon). However, it is noteworthy that in my case the best estimates came from a different continent but a similar habitat (ROGERS *et al.*, 1976), and from another different continent and habitat (GOWING & RECHER, 1984). Other factors apart from the precedence of equations seem to be important.

### Are the available equation series similar and equally applicable to new data series?

Caution appears to be seldom used when an estimate of biomass is needed. The most common practice is to use a general equation, regardless of the purpose of the research and without previous test, despite the application of specific rather than general equations is the usual recommendation in previous works (ROGERS *et al.*, 1976; SCHOENER, 1980). In fact, the predictive power of the equations always improves when the group is more specific (SCHOENER, 1980; SAMPLE *et al.*, 1993). The range of sizes for which the equations have been calculated is also ignored,

sometimes being too narrow (*in* ROGERS *et al.* [1976] 0.5-36.0 mm; SCHOENER [1980] does not offer the data), and the extrapolation outside these limits can be dangerous (SCHMIDT-NIELSEN, 1984; ZAR, 1996).

The consequence of the careless application of estimates seriously weakens the results of the research. For instance, in studies of abundance-body size relationships, different estimators can give different results (BLACKBURN *et al.*, 1993a, b). The same is true in studies on diet or composition of arthropod communities, in which the relative importance of different taxa can be widely biased (see fig. 1). An important question is that the best estimators for an absolute biomass from a sample could be a bad one when estimating the relative importance in biomass of the different groups, or individuals, of a sample (table III). That is, the optimal estimator may be one equation series or another, depending on the aim of the research.

#### **What tests can be made in order to confirm the applicability of available equation series to new series of data?**

The choice of the equations to apply must depend on the objectives of the researcher. When a large sample is available and the analysis requires only the estimation of the absolute biomass of a sample, a single general equation can be reliable, but the choice of the more accurate general equation from among the available must be made. When the sample size is smaller and the study carried out needs absolute estimates of biomass (*e.g.* food availability, diet selection, resource partitioning among the components of a community), it may be necessary to use specific equations, but once again the choice from among the published series must be made and its applicability tested. In my case, the ROGERS *et al.* (1976) equation offered an absolute estimate even better than the Guadix-Baza equations on the Baza subsample (table III), but Guadix-Baza equations gives a better picture of the relative importance in biomass of the different groups present in the sample (table III, fig. 1).

To objectively choose from among the available equations, it is important to test statistically their degree of fit with the sample to be estimated. In comparing actual weights with estimated weights, the tool used is paired *t*-test (ROGERS *et al.*, 1977; SAMPLE *et al.*, 1993; this study). However, its application can offer misleading results, because the significance of this test depends both on the mean difference between actual and estimated weights ( $H_0 = 0$ ), and the variance of this mean. The smaller the mean difference and the wider the variance, the higher is the probability to accept  $H_0$ . For instance, the Guadix-Baza specific equations gave a mean difference almost twice that of ROGERS *et al.* (1976) or GOWING & RECHER (1984), but an SD almost half that of these two works. On the contrary, the Guadix-Baza general equation gave *p*-values that were similar to the specific ones (0.0048 vs 0.0056 after Bonferroni corrections, see table III), despite that the mean difference with the actual values was more than twofold, because SD was also more than twofold greater. That is, a good estimate of mean but with a narrow SD can be significant, whereas a less accurate mean but with wider SD could be non-significant. The Wilcoxon paired test could be a reliable alternative; however, it works with ranks, and thus can misrepresent the true differences of estimates by reducing them to ranks. The *r*-correlation coefficient is a good indicator of the fit between the paired values, but is not indicative of the absolute estimation. Consequently, the aim of the estimate must be taken into account: when the interest

is an absolute estimate for a sample, a paired *t*-test may be a good tool for comparison, but for a good adjustment by groups or individuals, an *r* correlation coefficient appears to be more reliable.

**Is it advisable to formulate new equation series and, if so, what criteria are needed to follow?**

If there is no equation series that satisfies the degree of accuracy required, a solution may be to formulate a new equation series. In this case, it is important to take into account the number of individuals and taxonomic groups included in the regression, especially when formulating a general equation. The variations observed in the estimate between the different general equations arise, in large part, from the particularities of the pool of arthropods used in the regression, due to the number of individuals which each OTU accounts within the pool (see SCHOENER, 1980 for a detailed explanation). For example, the total *b*-value of the regression in the Guadix-Baza general equation was affected when non-Hexapoda were pooled together with Hexapoda, even without significant differences between equations. All references obtain significant specific equations and thus pooled the data: there are no clear criteria to determine the appropriate number for each OTU in the general regression, and this is a strong argument in favour of specific equations.

If specific equations, instead of general, are performed, the important questions are the criteria for selecting the groups and for selecting the individuals within a group. A combination of taxonomical and morphological bases is the best choice when selecting the groups to study (SCHOENER, 1980; SAMPLE *et al.*, 1993). On the other hand, when formulating specific equations, it is useful to take into account that a good regression equation does not need a high number of individuals to improve the fit (see Results). The key is to avoid outliers (SCHMIDT-NIELSEN, 1984) and to represent the whole range size of the group. An arbitrary but careful selection gives results as good as a large randomly selected sample.

## CONCLUSIONS

For the biomass estimations from a diversity of original data, accurate and precise tools are needed. However, the precision of the tool cannot correct the consequences of its incorrect use. All equation series may be equally accurate and precise, but they are not equally applicable to any sample, and researchers should take this into account. Consequently, previous testing is needed to select the series of equations to apply. The key is not to know which equation series is the best, but which series offers the best estimate for a particular sample and purpose. Because the performing of regression equations estimating biomass from length does not require the huge numbers that traditionally have been used, I strongly recommend preparing new equations in more zones and different taxonomic groups. This is especially needed for groups that, due to their number or biomass, are key in the new system to study, and in which the need for an accurate estimate is greater; also, this represents the only possibility when there are no available equations in the literature.

## ACKNOWLEDGEMENTS

Paco SANCHEZ-PIÑERO helped me in all the stages of this work, especially during the field work and selection and sorting of samples in laboratory. Mario DIAZ, José M. GOMEZ and Regino ZAMORA criticized several earlier versions of this manuscript, giving many valuable suggestions, David NESBITT improved the English, and Eva IÑESTA corrected the summary in French. Lodging for field work was provided by Conf. Hidrográfica del Guadalquivir (especially D. José HERNANDEZ) and the direction of Parque Natural Sierra de Baza. During this work, I was funded by a grant PFPI from Spanish Ministerio de Educación y Ciencia and a grant from DGICYT no. PB90-0852.

## REFERENCES

- BARRIENTOS J. A., (Coordinator) 1988. – *Bases para un curso práctico de Entomología*. Asociación Española de Entomología, Barcelona.
- BLACKBURN T. M., BROWN V. K., DOUBE B. M., GREENWOOD J. J. D., LAWTON J. H. & STORK N. E., 1993a. – The relationship between body size and abundance in natural animal assemblages. *J. Anim. Ecol.*, **62**, 519-528.
- BLACKBURN T. M., LAWTON J. H. & PIMM S. L., 1993b. – Non-metabolic explanations for the relationship between body size and animal abundance. *J. Anim. Ecol.*, **62**, 694-702.
- BROWN J. H. & MAURER B. A., 1989. – Macroecology: the division of food and space among species on continents. *Science*, **243**, 1145-1150.
- CIBOROWSKY J. J. H., 1983. – A simple volumetric instrument to estimate biomasses of fluid-preserved invertebrates. *Can. Entomol.*, **115**, 427-430.
- COTGREAVE P., 1993. – The relationship between body size and population abundance in animals. *Trends Ecol. Evol.*, **8**, 244-248.
- DAMUTH J., 1991. – Of size and abundance. *Nature*, **351**, 268-269.
- DÍAZ J. A. & DÍAZ M., 1990. – Estimaciones de tamaños y biomasa de artrópodos aplicables al estudio de la alimentación de vertebrados insectívoros. *Doñana Acta Vert.*, **17**, 67-74.
- DOWDY S. & WEARDEN S., 1991. – *Statistics for research*. John Wiley & Sons, New York.
- EDWARDS A. L., 1985. – *Multiple regression and the analysis of variance and covariance*. Freeman, New York.
- GASTON K. J. & LAWTON J. H., 1988. – Patterns in body size, population dynamics, and regional distribution of bracken herbivores. *Amer. Nat.*, **132**, 662-680.
- GOWING G. & RECHER H. F., 1984. – Length-weight relationships for invertebrates from forests in south-eastern New South Wales. *Austr. J. Ecol.*, **9**, 5-8.
- GRIFFITHS D., 1992. – Size, abundance, and energy use in communities. *J. Anim. Ecol.*, **61**, 307-315.
- GRIFFITHS R. A., 1986. – Feeding niche overlap and food selection in smooth and palmate newts, *Triturus vulgaris* and *T. helveticus*, at a pond in Mid-Wales. *J. Anim. Ecol.*, **55**, 201-214.
- HARVEY P. H. & GODFRAY H. C. J., 1987. – How species divide resources. *Amer. Nat.*, **129**, 318-320.
- HERRERA C. M., 1978. – Datos sobre la dieta invernal del colirrojo tizón (*Phoenicurus ochruros*) en encinares de Andalucía Occidental. *Doñana Acta Vert.*, **5**, 61-71.
- LAWTON J. H., 1991. – From physiology to population dynamics and communities. *Funct. Ecol.*, **5**, 155-161.
- MAURER B. A. & BROWN J. H., 1988. – Distribution of energy use and biomass among species of North American terrestrial birds. *Ecology*, **69**, 1923-1932.
- PAGEL M. D., HARVEY P. H. & GODFRAY H. C. J., 1991. – Species-abundance, biomass, and resource-use distributions. *Amer. Nat.*, **138**, 836-850.
- PETERS R. H., 1983. – *The ecological implications of body size*. Cambridge Univ. Press, Cambridge.
- PETERS R. H. & WASSERBERG K., 1984. – The effect of body size on animal abundance. *Æcologia*, **60**, 89-96.
- PRICE P. W., 1984. – *Insect ecology* (2nd ed.). John Wiley & Sons, New York.
- RICE W. R., 1989. – Analyzing tables of statistical tests. *Evolution*, **43**, 223-225.
- RICHARDS O. W. & DAVIES R. G., 1984. – *Tratado de Entomología Imms*. Omega, Barcelona. (Spanish translation of *Imms's General textbook of Entomology*, Chapman & Hall).

- ROGERS D. J. & RANDOLPH S. E., 1986. – Distribution and abundance of tsetse flies (*Glossina* spp.). *J. Anim. Ecol.*, **55**, 1007-1025.
- ROGERS L. E., HINDS W. T. & BUSCHOM R. L., 1976. – A general weight vs length relationship for insects. *Ann. Entomol. Soc. Amer.*, **69**, 387-389.
- ROGERS L. E., BUSCHOM R. L. & WATSON C. R., 1977. – Length-weight relationships of shrub-steppe invertebrates. *Ann. Entomol. Soc. Amer.*, **70**, 51-53.
- SAMPLE B. E., COOPER R. J., GREER R. & WHITMORE R. C., 1993. – Estimation of insect biomass by length and width. *Amer. Midl. Nat.*, **129**, 234-240.
- SANCHEZ-PIÑERO F., 1994. – Ecología de las comunidades de coleópteros de zonas áridas del sudeste peninsular. Ph. D. thesis, University of Granada, Granada, Spain.
- SCHMIDT-NIELSEN K., 1984. – *Scaling: Why is animal size so important?* Cambridge University Press, Cambridge.
- SCHOENER T. W., 1980. – Length-weight regressions in tropical and temperate forest-understory insects. *Ann. Entomol. Soc. Amer.*, **73**, 106-109.
- SMILEY J. T. & WISDOM C. S., 1982. – Photographic estimation of weight of insect larvae. *Ann. Entomol. Soc. Amer.*, **75**, 616-618.
- SNEATH P. H. A. & SOKAL R. R., 1973. – *Numerical taxonomy*. Freeman, San Francisco.
- STORK N. & BLACKBURN T. M., 1993. – Abundance, body size and biomass of arthropods in tropical forest. *Oikos*, **67**, 483-489.
- WILLIAMS P. H., 1989. – Why are there so many species of bumble bees at Dungeness? *Bot. J. Linn. Soc.*, **101**, 31-44.
- ZAR J. H., 1996. – *Biostatistical analysis* (3rd edition). Prentice Hall, New Jersey.