A multi-parameter artificial neural network model to estimate macrobenthic invertebrate productivity and production

Thomas Brey
Alfred Wegener Institute for Polar and Marine Research, P.O. Box 120161, D - 27515 Bremerhaven, Germany

Abstract

I developed a new model for estimating annual production-to-biomass ratio P/B and production P of macrobenthic populations in marine and freshwater habitats. Self-learning artificial neural networks (ANN) were used to model the relationships between P/B and twenty easy-to-measure abiotic and biotic parameters in 1252 data sets of population production. Based on log-transformed data, the final predictive model estimates log(P/B) with reasonable accuracy and precision (r^2 = 0.801; residual mean square RMS = 0.083). Body mass and water temperature contributed most to the explanatory power of the model. However, as with all least squares models using nonlinearly transformed data, back-transformation to natural scale introduces a bias in the model predictions, i.e., an underestimation of P/B (and P). When estimating production of assemblages of populations by adding up population estimates, accuracy decreases but precision increases with the number of populations in the assemblage.

Secondary production is the quantitative base of energy flow and of trophic interactions between heterotrophs and thus constitutes a key parameter of ecosystem functioning. The amount of secondary production and its distribution between populations are measures of trophic position and of linkage in aquatic food webs (Benke 2011) and may be indicative of ecosystem and population health (Buffagni and Comin 2000; Dolbeth et al. 2005). For a number of reasons, secondary production of the benthic compartment is of particular interest in aquatic ecology: (i) the sediment-water interface is, especially in shallow waters, a place of intense energy and matter turnover, relevant for both nutrient recycling and channeling of matter to higher trophic levels such as (commercially exploited) demersal fish; (ii) compared with pelagic organisms, benthic animals are less subject to random and/or short-term variability in space and time, because they are less mobile and, on average, live longer. Therefore, effects of external drivers such as increasing water temperature may be seen earlier and clearer in benthic populations and communities than in pelagic ones.

Demand for benthic secondary production studies is rising exponentially with the growing use of spatially explicit approaches (based on Geographical Information Systems, GIS) in aquatic ecosystem research (e.g., Ardron 2002; Dunton et al. 2005; Friedlander et al. 2007), which may require production estimates for many geographic reference points instead of the "average" population only. However, assessment of secondary production in aquatic systems is expensive and time-consuming. Multiple linear regression models (MLM), based on empirical data that predict population production P or production-to-biomass ratio P/B, are the favorite approach to shortcut this task (see Brey 2001, Cusson and Bourget 2005, and Dolbeth et al. 2005 for reviews and comparisons of published models).

Most of these predictive models, however, have their weak spot. Many models are restricted to taxonomic, functional, or habitat-defined subsets of macrobenthic populations (e.g., Plante and Downing 1989; Morin and Bourassa 1992; Tumbiolo and Downing 1994) or require independent variables such as life span, which are laborious to determine (e.g., Cusson and Bourget 2005). Furthermore, these published models lack a proper evaluation of their statistical properties, such as error distribution, accuracy, and precision. The model type used is a further concern. It is not yet clear whether or not there are universally valid scaling factors (e.g., Brown et al. 2004) between body mass or temperature and metabolic activity as well as its derivates such as production (Kozlowski and Konarzewski 2004; Glazier 2006; Seibel 2007; Brey 2010), i.e., MLM may not describe the existing relationships appropri-

ately. Moreover, MLM are quite sensitive to intercorrelation between independent parameters (Draper and Smith 1981). Artificial Neural Networks (ANN) provide an alternative approach: they can model complex, nonlinear, and noncontinuous relationships between independent and dependent variables by means of learning and generalizing from example data (Fausett 1994; Lek and Guégan 2000). Ecological applications of ANN range from taxonomic identification to diversity pattern prediction (e.g., Olden 2003, Willems et al. 2008). Brey et al. (1996) showed for the first time that ANN performed slightly but significantly better than MLM in predicting the P/B ratio of macrobenthic invertebrates. However, they applied specialized ANN software, which did not allow exporting the models in a generally applicable format.

This study has three objectives: (i) to build a general ANN model for the prediction of production-to-biomass ratio P/B and production P in aquatic macrobenthic invertebrate populations, (ii) to evaluate the statistical properties of this model, and (iii) to make this model available to the scientific community.

Materials and procedures

Data acquisition and conditioning

The initial database of this study consists of 1258 data sets collected from 522 publications (see spreadsheet "ProductivityANN01.xls" downloadable at http://www.thomas-brey/science/virtualhandbook). One data set represents one single measurement of population annual production P [mass m^{-2} y^{-1}], mean biomass B [mass m^{-2}], annual production-to-biomass ratio P/B [y^{-1}], and mean body mass M, accompanied by information on the animal taxonomy and lifestyle, on the environment, and on the methods applied. Relevant literature was tracked through the Aquatic Sciences and Fisheries Abstracts database (ASFA) of FAO (http://www.csa.com/factsheets/aquclust-set-c.php), through Web search engines, and through the catalogue and inventory of the Alfred Wegener Institute for Polar and Marine Research library. Each data source was carefully checked for comprehensible procedures, completeness of information, and methodical flaws.

Taxonomy follows a simplified seven-level taxonomic hierarchy, consisting of species, genus, family, superfamily/order, order/subclass, class, (sub-)phylum, following the Integrated Taxonomic Information System (ITIS) online database (http://www.itis.gov) and the World Register of Marine Species (http://www.marinespecies.org). Lifestyle is characterized by alimentation type (herbivorous, omnivorous, and carnivorous; species with symbionts were excluded a priori), by feeding type (grazer, suspension feeder, deposit feeder, predator/scavenger), and by mobility type (sessile, crawler, facultative swimmer). The environment is defined by the type of water body (marine, lake, river), by mean annual water temperature T, by water depth D, and by tidal elevation (intertidal, subtidal). A further parameter describes whether or not the population was subject to exploitation. Owing to the intrinsic

relationships between mortality rate and production-to-biomass ratio (Allen 1971; Brey 1999), exploitation per se does not put a population outside the empirical context considered here. The production calculation approach is recorded, too (size frequency method, increment summation method, mass specific growth rate method, mortality estimate; see Brey 2001 for detailed descriptions).

Data transformation and pre-analysis

I decided to predict production-to-biomass ratio P/B rather than production P directly, as the latter would require including biomass B as an independent variable, and the overwhelmingly strong effect of B on P may obscure other relationships. Original units of body mass were converted to Joule using factors provided by the conversion factor collection of Brey et al. (2010) and references therein.

Prior to model building, the variables P/B and the continuous independent variables body mass M, temperature T, and water depth D were transformed by approximating linear relationships with P/B according to theoretical considerations (e.g., Schmidt-Nielsen 1984; Brown et al. 2004) and to empirical evidence (e.g., Seibel and Drazen 2007), regarding the scaling of metabolic activity with bodymass, temperature, and water depth (see Brey 2010 for a full discussion). These transformations—log(P/B), log(M), 1/T, log(D)—aimed (i) at a more even distribution of data and of variance in the [M, T, D] space and (ii) at reducing non-linearity, which both facilitate ANN model fitting capabilities. All categorial variables were introduced in form of a binary (0,1) parameter for each category.

Pre-analysis by means of stepwise regression and ANN trial runs indicated that the best trade-off between number of input parameters, model predictive power, and model generality was achieved with a 20 parameter model using log(M), 1/T, log(D), five taxonomic categories, seven lifestyle categories, four environmental categories, and the exploitation marker (Fig. 1).

Multivariate outliers in the sample space [log(P/B), log(M), 1/T] were identified by Mahalanobis-Jackknife distances MJD (Barnett and Price 1995). We did not include depth in this analysis because the few deep sea values would always be identified as outliers. Data sets with MJD > mean MJD + 3 sigma were excluded from further analysis.
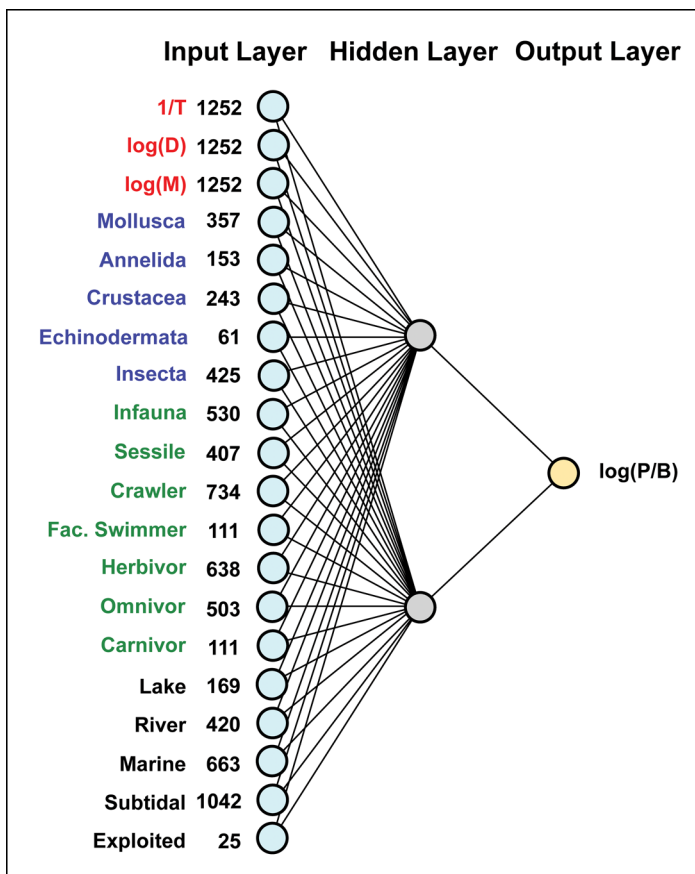
Predictive model construction

Artificial neural networks (ANN) of the backpropagation type (Hagan et al. 1996) consisted of 20 input nodes, two hidden nodes (H), and one output node, i.e., log(P/B), see Fig. 1. The network was parameterized as follows:

$$\log(P/B) = a_0 + a_1 \times H_1 + a_2 \times H_2$$

with

$$H_1 = \tanh(b_0 + b_1 \times \log(M) + b_2 \times 1/T$$
$$+ b_3 \times \log(D) + \ldots + b_{20} \times \text{Exploited})$$

Fig. 1. Scheme of the artificial neural network used to predict population production-to-biomass ratio (P/B) from three continuous parameters (temperature, water depth, body mass) and 17 categorical parameters (5 taxa, 7 lifestyle, 4 environment, and and state of exploitation). Number of positive cases ("yes") are provided for categorical variables.

$$H_2 = \tanh(c_0 + c_1 \times \log(M) + c_2 \times 1/T + c_3 \times \log(D)$$
$$+ \ldots + c_{20} \times \text{Exploited})$$

Note that internally the input data are normalized (mean = 0, SD = 1) and that the network parameter values are adjusted accordingly. To avoid overfitting of an ANN, the data were split randomly in 75% training data and 25% test data (Lek and Guégan 2000). The ANN "learned" the relationship between dependent and independent variables from the training data and was tested subsequently for its generalization capability by comparing prediction accuracy in training and test data as measured by the correlation between calculated P/B$_{calc}$ and predicted P/B$_{ann}$ values. Following a bootstrap approach (Efron and Tibshirani 1993), five ANN, each trained on a different subsample, were pooled into a composite prediction model, i.e., the predicted P/B$_{ann}$ value is the average of the predictions made by five ANN. This approach allows computing of confidence limits of the prediction, too. The ANN constructed in this study have been implemented in the Microsoft EXCEL spreadsheet "ProductivityANN01.xls" which

is available for download at http://www.thomas-brey/science/virtualhandbook (Brey 2001).

Assessment

Exploration of ANN model statistical properties

Accuracy (deviation from true value) and precision (variability of estimate) of model predictions of P/B and (including B) of P of the composite ANN are of particular concern, as the fitting of a least squares model to log-transformed data [here: log(P/B), log(M), log(D)] results in biased estimates when predicted values are transformed back to arithmetic units (Finney 1941; Smith 1993). I tested model performance in the estimation of population and macrobenthic assemblage production by means of a Monte Carlo approach. An assemblage was taken to consist of n (1, 2, 4,… 256) populations i randomly selected (without replacement) from the database ($n = 1252$). Assemblage production is the sum of P$_i$/B$_i$ × B$_i$ over these n populations. Model error E$_{ann}$ is computed by (P$_{calc}$ – P$_{ann}$)/P$_{calc}$. 1000 trial runs were made for each n except $n = 1$, where all 1252 data sets were used, and $n = 256$ with 500 runs. Additionally, I compared P$_{calc}$ and P$_{ann}$ in natural assemblages of ≥ 5 populations, i.e., as reported in studies of the production of several populations at one site that are part of the present database.

Mahalanobis-Jackknife distance analysis identified 6 outliers, thus reducing the data to 1252 data sets with 635 different species. There are 357 Mollusca data sets, 153 Annelida, 242 Crustacea, 61 Echinodermata, 425 Insecta, and 13 data sets referring to species from other taxa. The data cover a temperature range of almost 32° and a depth range of 2900 m, but center around shallow sites in temperate regions (Fig. 2). Mean body size covers nine orders of magnitude, but ranges between 1 and 1000 Joule in the majority of populations (Fig. 2).

ANN model statistics

Correlation between calculated log(P/B$_{calc}$) and log(P/B$_{ann}$) predicted by each of the five ANN models separately amounts on average to $r^2_{train} = 0.785$ (± 0.008 SD) for training data and to $r^2_{test} = 0.783$ (± 0.006 SD) for test data. r^2_{train} and r^2_{test} do not differ significantly (ANOVA, $P = 0.649$), i.e., the ANN models are not overfitted. Correlation between calculated P/B$_{calc}$ and the average P/B$_{ann}$ of the five ANN amounts to $r^2 = 0.801$ with a residual mean square RMS of 0.083 (Fig. 3). The residual plot indicates an overall precision range of about two orders of magnitude, but a very symmetric distribution of residuals (skewness = 0.205, kurtosis = 0.205) with 90% of all values within the interval –0.359 to 0.375 (Fig. 4). The composite ANN explains about four percent more variance in log(P/B$_{calc}$) than the best MLM obtained by means of stepwise multiple regression ($r^2 = 0.757$). This MLM indicates that body mass is by far the most significant factor (F value = 1036), followed by temperature (F value = 358) and taxonomic identity (F values between 24 and 107), whereas the remaining significant ($P < 0.05$) variables have F values < 15.

To evaluate the explanatory power provided by the 17 nominal variables additional to body mass, temperature, and
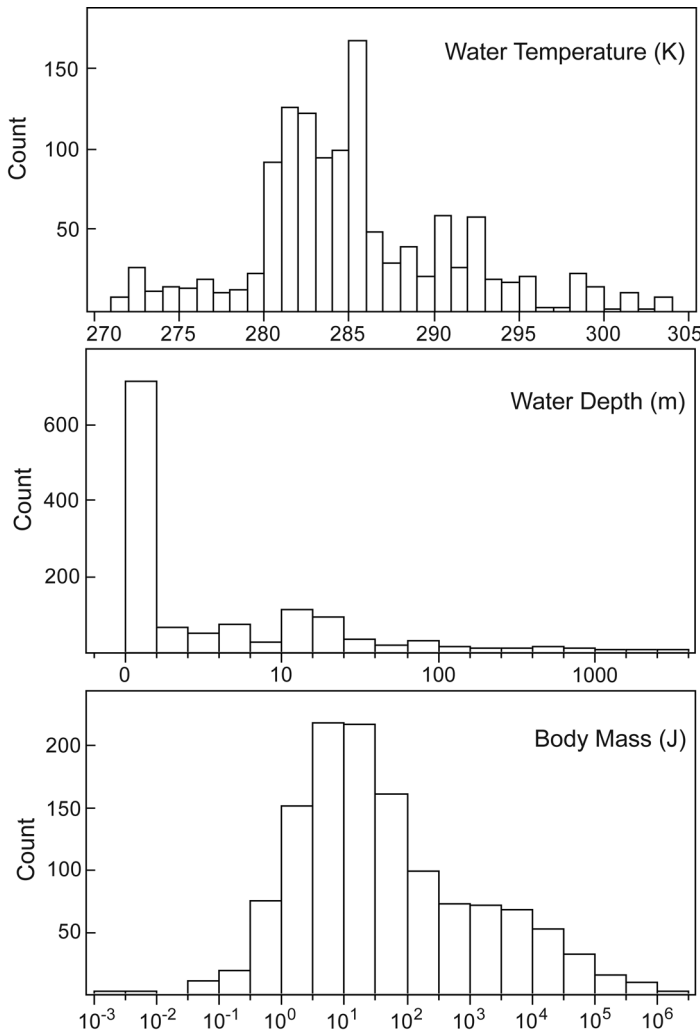
Fig. 2. Distribution of the 1252 data sets used for model building with respect to water temperature (Kelvin), water depth (meters), and mean body mass (Joule).

water depth, I constructed a further five ANN_3 models with the three input parameters $\log(M)$, $1/T$, and $\log(D)$ only. Correlation between $\log(P/B_{calc})$ and $\log(P/B_{ann})$ predicted by each of these ANN_3 separately amounts on average to $r^2_{train} = 0.709$ (± 0.004 SD) for training data and to $r^2_{test} = 0.712$ (± 0.010 SD) for test data. Both r^2_{train} and r^2_{test} are significantly ($P < 0.001$) smaller than the corresponding valued for the 20 parameter ANN. I.e., the 17 nominal variables explain about seven percent of the variance in $\log(P/B_{calc})$.

Accuracy and precision of production estimates

Back-transformation to arithmetic units introduces distinct asymmetry in the ANN model absolute prediction error (calculated as $P_{calc} - P_{ann}$) distribution. The residual plot (Fig. 5) indicates that the absolute prediction error increases with P/B, distinctly so beyond P/B > 5 y^{-1}, and preferably toward underestimation, thus causing an overall skewed error frequency distribution (skewness = 8.127, kurtosis = 89.330, not shown).
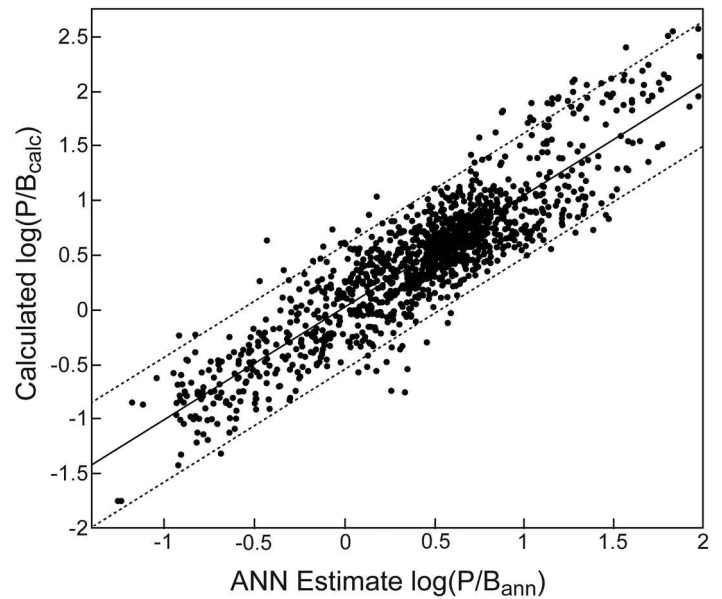

Fig. 3. Relationship between calculated P/B_{calc} and ANN predicted P/B_{ann} represented by a log-log linear regression model. Stippled lines indicate 95% confidence limits for individual estimates of P/B.
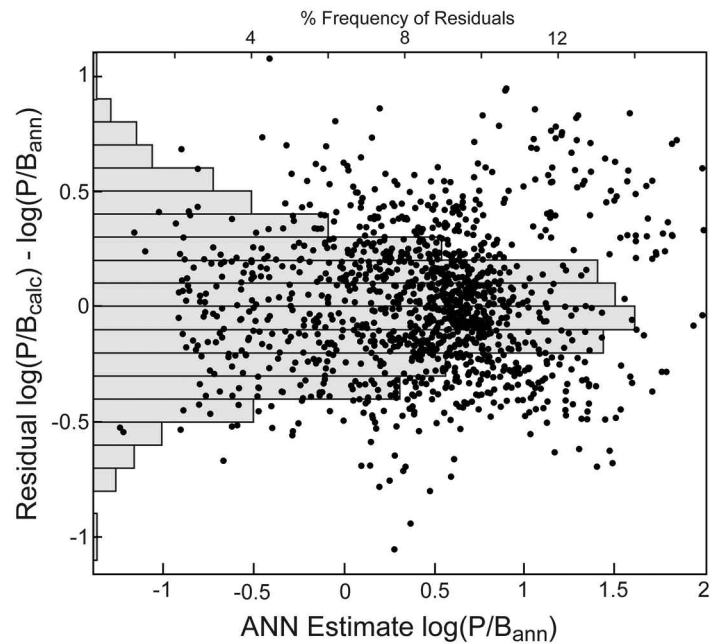
Fig. 4. Plot of residuals $\log(P/B_{calc}) - \log(P/B_{ann})$ versus P/B_{ann}. Superimposed frequency distribution of residuals (gray bars, upper scale).

This is manifest in estimates of single population production P, too, as indicated by the distribution of the error term $E_{ann} = (P_{calc} - P_{ann})/P_{calc}$ (Fig. 6, bottom graph, skewness = –2.829, kurtosis = 17.591). With increasing number of populations n (1, 2, 4,… 256) per assemblage, the precision of the production estimate increases distinctly. The 50% range (distance between 25th and 75th percentile) of E_{ann} decreases from
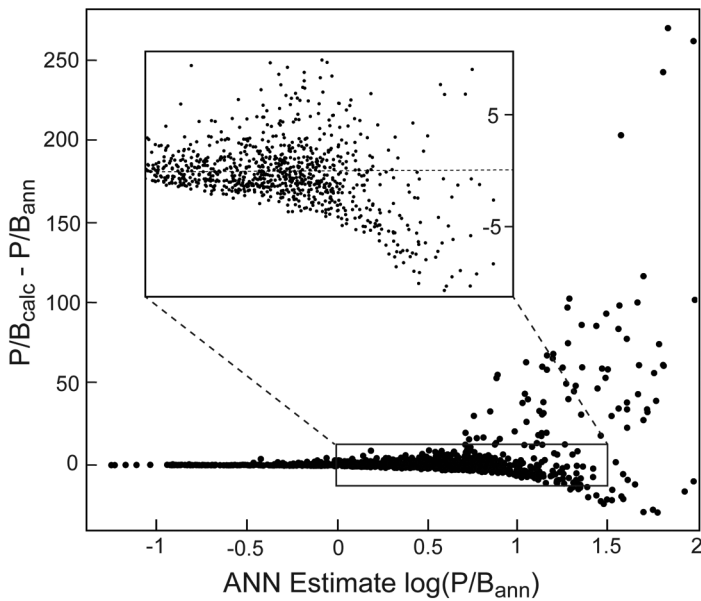
Fig. 5. Plot of model deviation $P/B_{calc} - P/B_{ann}$ versus $\log(P/B_{ann})$. Inlay shows a detailed view of error distribution within the range $0 \leq \log(P/B_{ann}) \leq 1.5$ (i.e., $1.0 \leq P/B_{ann} \leq 31.6$).

0.901 ($n = 1$) to 0.230 ($n = 256$), and the distribution of E_{ann} becomes more and more symmetric (Fig. 6; at $n = 256$ skewness = –0.215 and kurtosis = –0.315). At the same time, however, accuracy decreases from –0.013 (median E_{ann}) at $n = 1$ to significant ($P < 0.001$ for $n \geq 8$) underestimation of calculated assemblage production with an median E_{ann} of up to 0.218 at $n = 256$ (Fig. 5). Taking the whole database ($n = 1252$) as one assemblage results in an error of 0.253 (= underestimation of P by 25%). Based on all 8752 permutations of the Monte Carlo analysis, model error is predicted best by $E_{ann} = -0.164 + 0.159 \times \log(n)$ ($r^2 = 0.062$, $P < 0.001$).

Within the 21 natural assemblages of populations tested here, error E_{ann} of assemblage production ranged from –0.90 to + 0.80 with an average deviation of 0.10 and an average absolute deviation of 0.40 (Table 1). Furthermore, the model is significantly better ($P = 0.033$) in predicting production of taxonomically diverse assemblages ("Benthos," mean absolute deviation = 0.23) compared with more monophyletic assemblages (mean absolute deviation = 0.49).

Discussion

The composite ANN model predicts $\log(P/B)$ quite reasonably ($r^2 = 0.801$, RMS = 0.083, Fig. 3). However, back-transformation to P/B (and P, subsequently) produces the expected skewness in the error distribution, i.e., a bias is introduced when predicted values are transformed back to arithmetic units (Finney 1941; Smith 1993). Hence, applying an ANN model instead of an MLM does not circumnavigate this general disadvantage of all models built by least-square fitting of nonlinearly transformed data. There is little one can do against this shortcoming. Models fitted to untransformed data
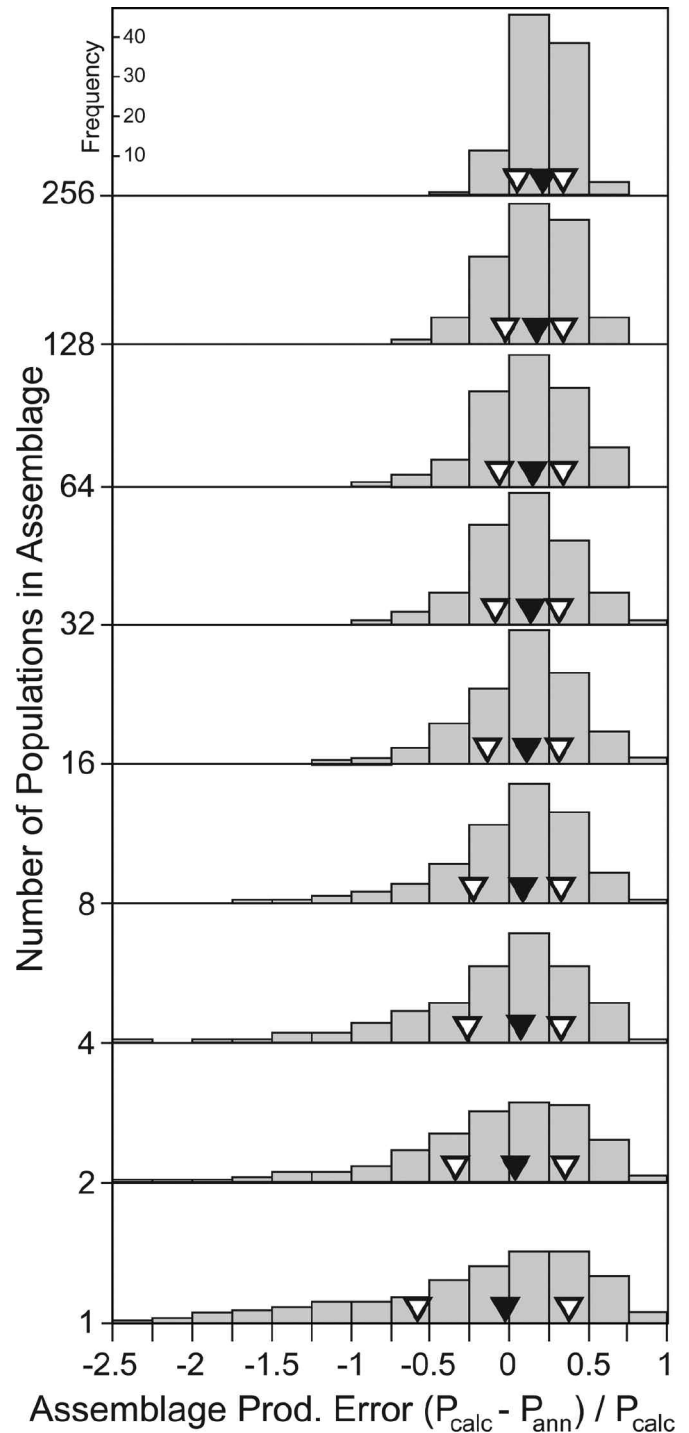
Fig. 6. Monte Carlo evaluation of accuracy and precision of assemblage production estimates by means of the ANN model. An assemblage consists of n (2, 4,...256) populations i randomly selected (without replacement) from the database ($n = 1252$). Assemblage production is the sum of $P_i/B_i \times B_i$ over these populations. Model error E_{ann} is computed by ($P_{calc} - P_{ann})/P_{calc}$. 1000 trial runs for each n except $n = 1$, where all 1252 data sets were used, and $n = 256$ with 500 runs. Gray bars: frequencies; black triangle: median of distribution; white triangles: 25% and 75% percentiles. Best fitting model: $E_{ann} = -0.164 + 0.169 \times \log(n)$, ($r^2 = 0.062$, $P < 0.001$, $n = 8752$).

Table 1. Calculated (P_{calc}) and estimated (P_{ann}) production of natural assemblages of macrobenthic populations. Deviation measured by ($P_{calc} - P_{ann}$)/P_{calc}.

Source	Nr Species	P_{calc} (kJ m^{-2} y^{-1})	P_{ann} (kJ m^{-2} y^{-1})	Deviation	P_{calc}/B (y^{-1})	Habitat, Assemblage
Benke and Wallace (1980)	6	28.5[‡]	19.8	0.30	5.3	River, Trichoptera
Benke et al. (2001)	10	918.3[‡]	984.5	–0.07	9.9	River, Insecta
Buchanan and Warwick (1974)	7	36.5[*]	51.3	–0.41	0.7	Marine, Benthos
Foberg (1976)	6	120.4[*]	63.9	0.47	2.7	Marine, Benthos
Gaines et al. (1992)	27	868.7[‡]	460.0	0.47	18.2	River, Insecta
Grzybkowska (1989)	7	300.4[‡]	99.6	0.67	33.5	River, Chironomidae
Grzybkowska (1989)	7	319.7[‡]	110.5	0.65	17.5	River, Chironomidae
Huryn and Wallace (1987)	21	46.1[‡]	76.9	–0.67	5.5	River, Insecta
Huryn and Wallace (1988)	6	47.6[‡]	48.7	–0.02	0.9	River, Insecta
Jackson and Fisher (1986)	5	1376.8[‡]	386.5	0.72	126.5	River, Insecta
Jónsson (1985)	10	292.5 [2]	239.7	0.18	8.1	Lake, Chironomidae
Lindegaard and Mortensen (1988) , Lindegaard et al. (1994)	7	279.0 [1]	470.5	–0.69	4.9	River, Chironomidae
Maitland and Hudspith (1973)	6	1258.3[*]	740.6	0.41	4.5	Lake, Chironomidae
Potter and Learner (1974)	14	585.9[†]	486.8	0.17	6.2	Lake, Benthos
Ramírez and Pringle (1998)	13	7.9[‡]	10.9	–0.38	19.4	River, Insecta
Sardá and San Martin (1993)	17	1279.1[*]	782.1	0.39	4.6	Marine, Benthos
Warwick and George (1980)	7	346.6[*†]	374.8	–0.08	1.3	Marine, Benthos
Warwick and Price (1975)	5	306.1[*†]	297.1	0.03	1.0	Marine, Benthos
Warwick et al. (1978)	10	661.0[*†]	626.3	0.05	0.6	Marine, Benthos
Welch (1976)	10	26.2[†]	49.6	–0.90	1.4	Lake, Chironomidae
Wilda (1984)	7	346.1[‡]	93.8	0.73	86.9	Lake, Chironomidae

[*]Increment summation method; [†]Mass specific growth rate method; [‡]Size frequency method.

perform poorly, because data variability is high and data distributions are skewed severely. Nevertheless, extent and direction of this bias should be reported with each model (see Fig. 5). Interestingly, the fact that the error distribution becomes more and more bell-shaped with growing assemblage size indicates that the estimate of assemblage production follows some central limit theorem (Rice 2007).

Estimates of natural assemblage production (Table 1) indicate that the ANN model underestimates assemblage production by 10% on average (assemblage size: 6 to 27 species), which is in good agreement with the findings of the Monte Carlo simulation. However, the model tends to underestimate calculated production much more in populations (see Fig. 5) and systems with very high P/B ratio, represented in Table 1 by assemblages of insect larvae from warm streams and lakes. A fully factorial ANOVA of the error E_{ann} for single population estimates versus log(B), log(P/B), and 1/T with all data as well as within the taxon Insecta confirms this view: underestimation of population P increases with log(P/B) ($P < 0.001$).

The poor model performance in high productivity systems may, on the one hand, be due to the small number of data the ANN can learn from (167 P/B values > 10 y^{-1} of which just 94 are > 20 y^{-1}). However, it might well be that the original production calculations are biased, too, as apparently the size-frequency method that is applied to such populations of fast-

growing and short-lived organisms is the least accurate of all standard methods of production calculation and is extremely sensitive to the estimate of maximum life span ("cohort production interval"; Cushman et al 1978; Benke 1984; Giberson and Galloway 1985; Morin et al. 1987; Cusson et al. 2006; Butkas et al. 2011). It is not surprising that predictions of production in taxonomically more diverse assemblages are more accurate; any empirical model should perform better the more the test assemblage reflects the diversity of the underlying data.

The current ANN displays about the same predictive power ($r^2 = 0.801$) as the ANN of Brey et al. (1996) ($r^2 = 0.799$), although the database has increased by some 40%. And, consistently, the ANN performs a few percent better than the best corresponding MLM. The tentative addition of a further five networks did not improve goodness of fit of the average prediction significantly. Apparently, the current set of "easy-to-measure" input parameters can explain some 80% of variance in log(P/B) of macrobenthic populations, but not more (see Brey 2001 and Cusson and Bourget 2005 for r^2 values of various published MLM). It seems unlikely that we can enhance prediction capability without using more elaborate predictors. For instance, life span has proven to be a powerful predictor of productivity (Robertson 1979, Cusson and Bourget 2005), but its determination is much too elaborate to make it a suitable parameter for the type of versatile predictive model we aim for.

Thus, single-populations estimates of P/B and P derived from the ANN model should be interpreted carefully, because precision is low; the 95% confidence range of the estimate is about one order of magnitude wide (Fig. 3). Precision increases distinctly when the model is used to estimate P/B and P of assemblages such as functional groups, feeding guilds, or whole communities. The fact that the model underestimates P/B and P of such assemblages by up to one quarter, i.e., estimate accuracy decreases with the number of populations pooled (see Fig. 6)—deems to be no fatal flaw. On the one hand, for benthic communities of "typical" species richness < 100 species (e.g., Bambach 1977; Gray 2000; Macpherson 2002; Reichert et al. 2010), this bias is small compared with sampling error in most benthic studies (e.g., Holmes and McIntyre 1984; Vezina 1988; Schacher and Wooldridge 1996). On the other hand, the bias is well known and described by a model (see legend to Fig. 6). Therefore, a correction for decreasing accuracy of assemblage production estimates is an option, but its benefits remain open for discussion.

Comments and recommendations

The ANN model presented here is a suitable tool for the estimation of production of macrobenthic assemblages (communities). When applied to single populations, a measure of error (standard deviation or confidence interval) should be provided together with the estimate of P/B or P. The user should keep in mind that accuracy and precision of model input data, particularly body mass, temperature, and water depth, intrinsically determine reliability of model output. It remains to be seen whether a broader database may reduce the actual model bias toward high productivity populations and assemblages.

References

Allen, K. R. 1971. Relation between production and biomass. Can. J. Fish. Aquat. Sci. 28:1537-1581.

Ardron, J. 2002. A GIS receipe for determining benthic complexity. *In* J. Breman (ed.), Marine geography: GIS for the oceans and seas. ESRI.

Bambach, R. K. 1977. Species richness in marine benthic habitats through the phanerozoic. Paleobiology 3:152-167.

Barnett, V., and T. Price. 1995. Outliers in statistical data, 3rd ed. Wiley.

Benke, A.C., and J.B. Wallace 1980. Trophic basis of production among net-spinning caddisflies in a southern Appalachian stream. Ecology 61:101-118.

———— 1984. Secondary production of aquatic insects, p. 289-322. *In* V. H. Resh and D. M. Rosenberg (eds.), Ecology of aquatic insects. Praeger.

————, J.B. Wallace, J.W. Harrison, and J.W. Koebel 2001. Food web quantification using secondary production analysis: predaceous invertebrates of the snag habitat in a subtropical river. Freshwater Biology 46:329-346 [doi:10.1046/j.1365-2427.2001.00680.x].

————. 2011. Secondary production, quantitative food webs, and trophic position. Nature Educ. Knowl. 2(2):2.

Brey, T. 1999. Growth performance and mortality in aquatic benthic invertebrates. Adv. Mar. Biol. 35:153-223 [doi:10.1016/S0065-2881(08)60005-X].

————. 2001. Population dynamics in benthic invertebrates. A virtual handbook. <http://www.thomas-brey.de/science/virtualhandbook/>

————. 2010. An empirical model for estimating aquatic invertebrate respiration. Methods Ecol. Evol. 1:92-101 [doi:10.1111/j.2041-210X.2009.00008.x].

————, A. Jarre-Teichmann, and O. Borlich. 1996. Artificial neural network versus multiple linear regression: Predicting P/B ratios from empirical data. Mar. Ecol. Progr. Ser. 140:251-256 [doi:10.3354/meps140251].

————, C. Müller-Wiegmann, Z. M. C. Zittier, and W. Hagen. 2010. Body composition in aquatic organisms—a global data bank of relationships between mass, element composition and energy content. J. Sea Res. 64:334-340 [doi:10.1016/j.seares.2010.05.002].

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Towards a metabolic theory of ecology. Ecology 85:1771-1789 [doi:10.1890/03-9000].

Buchanan, J.B., and R.M. Warwick 1974. An estimate of benthic macrofaunal production in the offshore mud of the Northumberland Coast. Journal of the Marine Biological Association U.K. 54: 197–222.

Buffagni, A., and E. Comin. 2000. Secondary production of benthic communities at the habitat scale as a tool to assess ecological integrity in mountain streams. Hydrobiologia 422:183-195 [doi:10.1023/A:1017015326808].

Butkas, K. J., Y. Vadeboncoeur, and M. J. Vander Zanden. 2011. Estimating benthic invertebrate production in lakes: a comparison of methods and scaling from individual taxa to the whole-lake level. Aquat. Sci. 73:153-169 [doi:10.1007/s00027-010-0168-1].

Cushman, R. M., H. H. Shugart, S. G. Hildebrandt, and J. W. Elwood. 1978. The effect of growth curve and sampling regime on instantaneous-growth, removal-summation, and Hynes/Hamilton estimates of aquatic insect production: a computer simulation. Limnol. Oceanogr. 23:184-189 [doi:10.4319/lo.1978.23.1.0184].

Cusson, M., and E. Bourget. 2005. Global patterns of macroinvertebrate production in marine benthic habitats. Mar. Ecol. Prog. Ser. 297:1-14 [doi:10.3354/meps297001].

————, J.-F. Plante, and C. Genest. 2006. The effect of different sampling designs and methods on the estimation of secondary production: A simulation. Limnol. Oceanogr. Methods 4:38-48 [doi:10.4319/lom.2006.4.38].

Dolbeth, M., A. I. Lillebø, P. G. Cardoso, S. M. Ferreira, and M. A. Pardal. 2005. Annual production of estuarine fauna in different environmental conditions: An evaluation of the estimation methods. J. Exp. Mar. Biol. Ecol. 326:115-127 [doi:10.1016/j.jembe.2005.05.010].

Dunton, K. H., J. L. Goodall, S. V. Schonberg, J. M. Grebmeier, and D. R. Maidment. 2005. Multi-decadal synthesis of benthic–pelagic coupling in the western arctic: Role of cross-shelf advective processes. Deep Sea Res. II 52:3462-3477 [doi:10.1016/j.dsr2.2005.09.007].

Draper, N. R., and H. Smith. 1981. Applied regression analysis, 2nd ed. Wiley.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall.

Fausett, L. V. 1994. Fundamentals of neural networks. Prentice Hall.

Finney, D. J. 1941. On the distribution of a variate whose logarithm is normally distributed. J. Royal Stat. Soc. Suppl. 7:155-161.

Foberg, M. 1976. Productionsberäkningar för nagra makroevertebrater i Lules skärgard sommaren 1976. Internal Report, Univ. Stockholm, Sweden.

Friedlander, M., E. K. Brown, and E. M. Monaco. 2007. Coupling ecology and GIS to evaluate efficacy of marine protected areas in Hawaii. Ecol. Appl. 17:715-730 [doi:10.1890/06-0536].

Gaines, W.L., C.E. Cushing, and S.D. Smith. 1992. Secondary production estimates of benthic insects in three cold desert streams. Great Basin Naturalist 52:11-24.

Giberson, D. J., and T. D. Galloway. 1985. Life history and production of Ephoron album (Say. (Ephemeroptera: Polymitarcidae. in the Valley River, Manitoba). Can. J. Zool. 63:1668-1674 [doi:10.1139/z85-248].

Glazier, D. S. 2006. The 3/4-power law is not universal: evolution of isometric, ontogenetic metabolic scaling in pelagic animals. BioScience 56:325-332 [doi:10.1641/0006-3568(2006)56[325:TPLINU]2.0.CO;2].

Gray, J. 2000. The measurement of marine species diversity, with an application to the benthic fauna of the Norwegian continental shelf. J. Exp. Mar. Biol. Ecol. 250:23-49 [doi:10.1016/S0022-0981(00)00178-7].

Grzybkowska, M. 1989. Production estimates of the dominant taxa of Chironomidae (Diptera) in the modified River Widawka and the natural River Grabia, Central Poland. Hydrobiologia 179:245- 259.

Hagan, M. T., H. B. Demuth, and M. H. Beale. 1996. Neural network design. PWS Publishing Company.

Holme, N. A., and A. D. McIntyre. 1984. Methods for the study of marine benthos, 2nd ed. Blackwell Scientific.

Huryn, A.D., and J.B. Wallace. 1987. The exopterygote insect community of a mountain stream in North Carolina, USA: life histories, production, and functional structure. Aquatic Insects 9:229-251.

———, and ———. 1988. Community structure of Trichoptera in a mountain stream: spatial patterns of production and functional organization. Freshwater Biology 20:141-155 [doi:10.1111/j.1365-2427.1988.tb00438.x].

Jackson J.K., and S.G. Fisher 1986. Secondary production, emergence, and export of aquatic insects of a Sonoran Desert Stream. Ecology 67:629–638.

Jónsson, E. 1985. Population dynamics and production of Chironomidae (Diptera) at 2 m depth in Lake Esrom, Denmark. Archiv für Hydrobiologie Supplement 70:239-279.

Kozlowski, J., and M. Konarzewski. 2004. Is West, Brown and Enquist's model of allometric scaling mathematically correct and biologically relevant? Funct. Ecol. 18:283-289 [doi:10.1111/j.0269-8463.2004.00830.x].

Lek, S., and J. F. Guégan. 2000. Artificial neural networks— Application to ecology and evolution. Springer.

Lindegaard, C., and E. Mortensen. 1988. Abundance, life history and production of Chironomidae (Diptera) in a Danish lowland stream. Archiv für Hydrobiologie Supplement 81:563-587.

———, K. Hamburger, and P.C. Dall. 1994. Population dynamics and energy budget of *Marionia southerni* (Cernosvitov) (Enchytraeidae, Oligochaeta) in the shallow littoral of Lake Esrom, Denmark. Hydrobiologia 278:291-301 [doi:10.1007/BF00142338].

Macpherson, E. 2002. Large-scale species-richness gradients in the Atlantic Ocean. Proc. Royal Soc. London B 269:1715-1720 [doi:10.1098/rspb.2002.2091].

Maitland, P.S., and P.M.G. Hudspith. 1973. Production studies of larval Chironomidae and other zoobenhtos in the sandy littoral of Loch Leven, Kinross. Proceedings of the Royal Society of Edinburgh 74:241-258.

Morin, A., and N. Bourassa. 1992. Modèles empiriques de la production anuelle et du rapport P/B d'invertébrés benthiques d'eau courante. Can. J. Fish. Aquat. Sci. 49:532-539 [doi:10.1139/f92-063].

———, T. A. Mousseau, and D. A. Roff. 1987. Accuracy and precision of secondary production estimates. Limnol. Oceanogr. 32:1342-1352 [doi:10.4319/lo.1987.32.6.1342].

Olden, J. D. 2003. A species-specific approach of modeling biological communities and its potential for conservation. Cons. Biol. 17:854-863 [doi:10.1046/j.1523-1739.2003.01280.x].

Plante, C., and J. A. Downing. 1989. Production of freshwater invertebrate populations in lakes. Can. J. Fish. Aquat. Sci. 46:1489-1498 [doi:10.1139/f89-191].

Potter, D.W.B., and M.A. Learner. 1974. A study of the benthic macro-invertebrastes of a shallow euthrophic reservoir in South Wales with emphasis on the Chironomidae (Diptera). Archiv für Hydrobiologie 74:189-226.

Ramírez, A., and C.M. Pringle. 1998. Structure and production of a benthic insect assemblage in a neotropical stream. Journal of the North American Benthological Society 17: 443-463.

Reichert, K., K. I. Ugland, I. Bartsch, J. Hortal, J. Bremner, and A. Kraberg. 2010. Species richness estimation: Estimator performance and the influence of rare species. Limnol. Oceanogr. Methods 8:294-303 [doi:10.4319/lom.2010.8.294].

Rice, J. 2007. Mathematical statistics and data analysis, 3rd ed. Duxbury Press.

Robertson, A. I. 1979. The relationship between annual production: biomass ratios and lifespans for marine macrobenthos. Oecologia 38:193-202 [doi:10.1007/BF00346563].

Sarda, R., and G. San Martin. 1993. *Streptosyllis verrilli* (Moore, 1907), new combination, life cycle, population dynamics and production from a salt marsh in southern New England. Bulletin of Marine Science 51:407-419.

Schacher, T. A., and T. H. Wooldridge. 1996. How sieve mesh size affects sample estimates of estuarine benthic macrofauna. J. Exp. Mar. Biol. Ecol. 201:159-171 [doi:10.1016/0022-0981(95)00198-0].

Schmidt-Nielsen, K. 1984. Scaling—Why is animal size so important? Cambridge Univ. Press [doi:10.1017/CBO9781139167826].

Seibel, B. A. 2007. On the depth and scale of metabolic rate variation: scaling of oxygen consumption rates and enzymatic activity in the Class Cephalopoda (Mollusca). J. Exp. Biol. 210:1-11 [doi:10.1242/jeb.02588].

———, and J. C. Drazen. 2007. The rate of metabolism in marine animals: environmental constraints, ecological demands and energetic opportunities. Phil. Trans. Royal Soc. B 362:2061-2078 [doi:10.1098/rstb.2007.2101].

Smith, R. J. 1993. Logarithmic transformation bias in allometry. Amer. J. Phys. Allom. 90:215-228 [doi:10.1002/ajpa.1330900208].

Tumbiolo, M. L., and J. A. Downing. 1994. An empirical model for the prediction of secondary production in marine benthic invertebrate populations. Mar. Ecol. Progr. Ser. 114:165-174 [doi:10.3354/meps114165].

Vezina, A. F. 1988. Sampling variance and the design of quantitative surveys of the marine benthos. Mar. Biol. 97:151-155 [doi:10.1007/BF00391255].

Warwick, R.M., and R. Price. 1975. Macrofauna production in an estuarine mud-flat. Journal of the Marine Biological Association U.K. 55:1–18.

———, C.L. George, and J.R. Davies. 1978. Annual macrofauna production in a *Venus* community. Estuarine, Coastal and Marine Science 7:215–241.

———, and ———. 1980. Annual macrofauna production in an *Abra* community. In: Collins, M.B., F.T. Banner, P.A. Tyler, S.J. Wakefield, and E. James (eds): Industrialized embayments and their environmental problems. Pergamon, Oxford, 517-538.

Welch, H.E. 1976 Ecology of Chironomidae (Diptera) in a polar lake. Canadian Journal of Fisheries and Aquatic Sciences 33:227-247.

Wilda, T.J. 1984. The production of five genera of Chironomidae (Diptera) in Lake Norman, a North Carolina reservoir. Hydrobiologia 108:145-152 [doi:10.1007/BF00014874].

Willems, W., and others. 2008. Where is the worm? Predictive modelling of the habitat preferences of the tube-building polychaete *Lanice conchilega*. Ecol. Model. 212:74-79 [doi:10.1016/j.ecolmodel.2007.10.017].