



**HAL**  
open science

## **Bias-corrected Pearson estimating functions for Taylor's power law applied to benthic macrofauna data**

Bent Jørgensen, Clarice G.B. Demétrio, Erik Kristensen, Gary T. Banta, Hans Christian Petersen, Matthieu Delefosse

### ► **To cite this version:**

Bent Jørgensen, Clarice G.B. Demétrio, Erik Kristensen, Gary T. Banta, Hans Christian Petersen, et al.. Bias-corrected Pearson estimating functions for Taylor's power law applied to benthic macrofauna data. *Statistics and Probability Letters*, 2011, 10.1016/j.spl.2011.01.005 . hal-00746104

**HAL Id: hal-00746104**

**<https://hal.science/hal-00746104v1>**

Submitted on 27 Oct 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

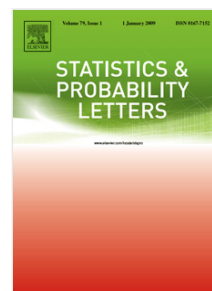
## Accepted Manuscript

Bias-corrected Pearson estimating functions for Taylor's power law applied to benthic macrofauna data

Bent Jørgensen, Clarice G.B. Demétrio, Erik Kristensen, Gary T. Banta,  
Hans Christian Petersen, Matthieu Delefosse

PII: S0167-7152(11)00012-5  
DOI: [10.1016/j.spl.2011.01.005](https://doi.org/10.1016/j.spl.2011.01.005)  
Reference: STAPRO 5880

To appear in: *Statistics and Probability Letters*



Please cite this article as: Jørgensen, B., Demétrio, C.G.B., Kristensen, E., Banta, G.T., Petersen, H.C., Delefosse, M., Bias-corrected Pearson estimating functions for Taylor's power law applied to benthic macrofauna data. *Statistics and Probability Letters* (2011), doi:10.1016/j.spl.2011.01.005

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Bias-corrected Pearson estimating functions for Taylor's power law applied to benthic macrofauna data

Bent Jørgensen<sup>a\*</sup>, Clarice G.B. Demétrio<sup>c</sup>, Erik Kristensen<sup>b</sup>, Gary T. Banta<sup>d</sup>,  
Hans Christian Petersen<sup>a</sup>, Matthieu Delefosse<sup>b</sup>

<sup>a</sup>Department of Mathematics and Computer Science, <sup>b</sup>Institute of Biology,  
University of Southern Denmark, Odense, Denmark

<sup>c</sup>Department of Exact Sciences, ESALQ, University of São Paulo, Piracicaba, Brazil

<sup>d</sup>Department of Environmental, Social and Spatial Change, Roskilde University, Denmark

January 5, 2011

## Abstract

Estimation of Taylor's power law for species abundance data may be performed by linear regression of the log empirical variances on the log means, but this method suffers from a problem of bias for sparse data. We show that the bias may be reduced by using a bias-corrected Pearson estimating function. Furthermore, we investigate a more general regression model allowing for site-specific covariates. This method may be efficiently implemented using a Newton scoring algorithm, with standard errors calculated from the inverse Godambe information matrix. The method is applied to a set of biomass data for benthic macrofauna from two Danish estuaries.

## 1 Introduction

A common feature of species abundance data is that sites with higher abundances tend to have higher variability. Let us consider independent abundance data  $Y_{ij}$ , where  $i = 1, \dots, k$  denotes site and  $j = 1, \dots, n_i$  denotes replicates within site, and let  $\mu_i = \mathbb{E}(Y_{ij})$  denote the mean abundance for site  $i$ . The most common mean-variance relationship found in practice is the power variance function with parameters  $a > 0$  and  $b$ ,

$$\text{Var}(Y_{ij}) = a\mu_i^b, \quad (1.1)$$

which is known as *Taylor's power law* (Taylor, 1961). For example Taylor *et al.* (1983) reported that the power law had been observed for 444 different species of birds, moths and aphids sampled over Great Britain. Kendal (2004) summarized a number of applications of the power law in areas like genomics and epidemiology, and proposed a theoretical explanation for the

---

\*Corresponding author.

*E-mail address:* bentj@stat.sdu.dk (B. Jørgensen)

power law based on the so-called Tweedie convergence theorem (Jørgensen *et al.*, 1994), see Tweedie (1984) and Jørgensen (1997, Ch. 4). Taylor's power law has also received attention in physics (Eisler *et al.*, 2008; Fronczak & Fronczak, 2010), where the phenomenon is known as *fluctuation scaling*.

There is particular interest in the power parameter  $b$  of (1.1), because Taylor (1961) suggested that  $b$  is a species-specific index of aggregation related to the social behaviour of the individuals of the species, an assertion that has generated much controversy (see e.g. Titmus, 1983; Downing, 1986; Kilpatrick & Ives, 2003). It is common in practice to observe values of  $b$  between 1 and 2, where the traditional interpretation of  $b$  is that the value 1 indicates a random dispersion pattern, whereas  $b > 2$  indicates an aggregated dispersion pattern (Pedigo & Buntin, 1994, p. 48). Values of  $b$  below 1 are almost never found in practice, whereas values of  $b$  above 2 are possible, but less frequent, although Taylor *et al.* (1983) reported a large proportion of such values in their study.

It is hence unfortunate that the traditional estimator for  $b$ , obtained by regressing the log empirical variances on the log empirical means (cf. Section 2), suffers from a problem of bias (Perry, 1981), especially for sparse data, i.e. when the number of replicates  $n_i$  are small. In response to this problem, Perry (1981) proposed two alternative estimators, one based on a gamma regression method for the empirical variances, and the other using weighted least squares.

Following Park & Cho (2004), we consider a more general model allowing for site-specific covariates, corresponding to for example a generalized linear model for the data  $Y_{ij}$  with two unknown variance parameters  $a$  and  $b$ , a model that may be fitted by combining a quasi-score method for the regression parameters with a pseudo-likelihood method for the variance parameters. This method depends on second-moment assumptions only, but like full maximum likelihood, the pseudo-likelihood method suffers from bias in the estimation of the variance parameters.

We shall address this problem in the context of unbiased estimating functions along the lines of Jørgensen & Knudsen (2004) and Holst & Jørgensen (2010), using a bias-corrected Pearson estimating function for the parameters  $a$  and  $b$  (Sections 3–4). In the simple case without covariates, it turns out that the new method produces the same estimates for  $a$  and  $b$  as Perry's gamma regression method, but with the additional advantage that standard errors for all parameters, including  $b$ , may be calculated from the inverse Godambe information matrix (Section 5). We apply the method to estimate Taylor's power law for a set of biomass data for benthic macrofauna from two Danish estuaries (Section 6).

## 2 Taylor's power law and the Tweedie distribution

We now review the basic setup for Taylor's power law and discuss its possible explanation in terms of the Tweedie distribution. Let us consider the simple case without site-specific covariates, which we shall call the *Taylor model*. In this case, estimation of the power law (1.1) is usually based on the site-specific summary statistics

$$\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij} \text{ and } S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2, \quad (2.1)$$

which are unbiased estimators of the mean  $\mu_i$  and variance  $a\mu_i^b$ , respectively, for each  $i = 1, \dots, k$ .

The classical *Taylor's regression estimator* is based on the intuitively appealing idea that a plot of  $(\log \bar{Y}_i, \log S_i^2)$  is approximately linear,

$$\log S_i^2 = \log a + b \log \mu_i + \epsilon_i, \quad (2.2)$$

where the error term  $\epsilon_i$  has, approximately, variance inversely proportional to the degrees of freedom  $n_i - 1$ . Taylor's regression estimator is obtained by weighted regression of  $\log S_i^2$  on  $\log \bar{Y}_i$ , with weights  $n_i - 1$ , providing an approximate solution to the linear regression model (2.2). However, as noted by Perry (1981), the substitution of  $\bar{Y}_i$  for  $\mu_i$  in (2.2) and the fact that the error term  $\epsilon_i$  has non-zero mean lead to underestimation of the power parameter  $b$  for sparse data, leading to a negative bias. By comparison, *Perry's gamma regression estimator* is based on a generalized linear model with a gamma distribution for  $S_i^2$ , again weighted by the degrees of freedom  $n_i - 1$ .

A general problem for the Taylor model is that empty sites and sites with zero variances have to be excluded from the regression, causing a selection bias. This problem is unavoidable when no site covariates are available, since there is no way of knowing if an empty site is due to chance, or to unsustainable living conditions for the species in question. The benthic biomass data of Section 6, containing no covariates, is typical in this respect. We shall hence consider a more comprehensive approach based on a regression model of the form  $\mu_i = f_i(\boldsymbol{\beta})$ , where  $f_i$  is a smooth function depending on an  $\ell$ -vector  $\boldsymbol{\beta}$  of parameters, where for simplicity we suppress the dependence on the covariates in our notation. The simplest special case is a generalized linear model involving site-specific covariates via a link function. Such regression models will be useful in future investigations, where the presence of covariates will allow models to be fitted to the complete set of data, including empty sites, thereby avoiding the above-mentioned selection bias of the Taylor model.

We now turn to the Tweedie distribution  $\text{Tw}_b(\mu_i, a)$  with mean  $\mu_i$ , dispersion parameter  $a$  and power parameter  $b$ , a model which is a natural exponential family with power variance function (1.1) (cf. Jørgensen, 1997, Ch. 4). This model appears as the limiting distribution in the Tweedie convergence theorem (Jørgensen *et al.*, 1994), which may explain the frequent occurrence of Taylor's power law for ecological data. Estimation of the parameters  $a$ ,  $b$  and  $\boldsymbol{\beta}$  under the Tweedie model  $Y_{ij} \sim \text{Tw}_b(\mu_i, a)$  may be based on maximum likelihood. This method has been implemented in the `tweedie` R function of Dunn (2009), which is, however, somewhat slow due to the use of a profile likelihood for  $b$ . The pseudo-likelihood method of Park & Cho (2004) is computationally lighter, but unfortunately it shares with full maximum likelihood the problem of bias in the estimators for the variance parameters  $a$  and  $b$ , when these are estimated together with  $\boldsymbol{\beta}$ . We show in Section 3 how this bias may be reduced by using a bias-corrected Pearson estimating function for  $a$  and  $b$ . The correction of bias is particularly important for the Taylor model, where  $a$  and  $b$  are estimated in the presence of the  $k$ -dimensional nuisance parameter  $\boldsymbol{\beta} = (\mu_1, \dots, \mu_k)^\top$ , see the discussion at the end of Section 3.

Let us interpret the Tweedie distribution in the most common case  $b \in (1, 2)$ , which is relevant for the biomass data analyzed in Section 6. In this case a Tweedie random variable  $Y$  follows a compound Poisson distribution (cf. Jørgensen, 1997, Ch. 4), which is given as the sum of  $N$  independent clusters,

$$Y = X_1 + \dots + X_N, \quad (2.3)$$

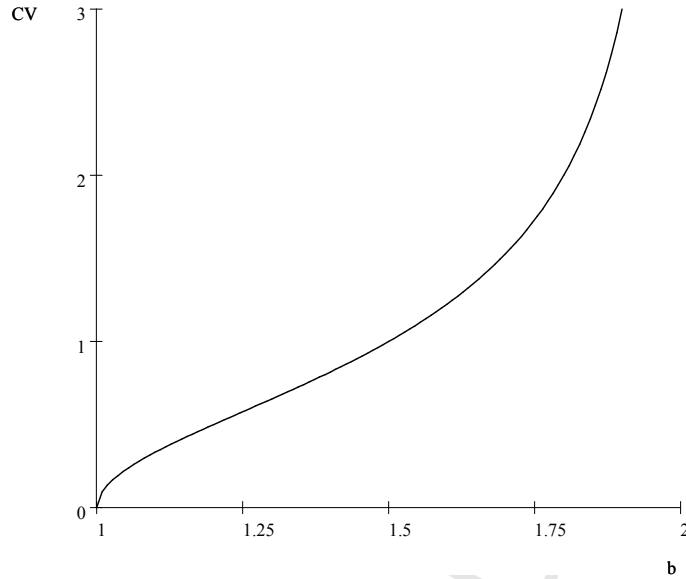


Figure 1: Plot of CV for cluster size as a function of  $b$ .

with the convention that  $Y = 0$  for  $N = 0$ . Here  $N$  is a Poisson variable with mean  $\mu^{2-b}/(2-b)$ , independent of the i.i.d. gamma random variables  $X_m$ , each with mean  $(2-b)\mu^{b-1}$  and coefficient of variation

$$CV(X_m) = \sqrt{\frac{b-1}{2-b}}.$$

We note that  $CV(X_m)$  is an increasing function of  $b$  (cf. Figure 1), and that both the Poisson and cluster means are increasing functions of the overall mean  $\mu = \mathbb{E}(Y)$ .

An important property of the compound Poisson distribution is that  $Y$  has a positive probability mass at zero, while being continuous for  $Y > 0$ . This agrees well with the behaviour of a variable such as biomass, which is zero when the species in question is absent, and continuous for positive values. Figure 2 shows plots of the Tweedie density function in the case  $b = 1.5$ , corresponding to an exponentially distributed cluster size with  $CV = 100\%$ . We note that the boundary cases  $b = 1$  and  $b = 2$  correspond to a scaled Poisson and a gamma distribution for  $Y$ , respectively.

The decomposition (2.3) allows us to interpret  $b$  as a measure of aggregation or clustering for the species in question. For  $b$  just above 1, the cluster CV is small (cf. Figure 1), corresponding to isolated and well-defined clusters. As  $b$  increases and the cluster CV becomes larger, the clusters become more dispersed, as illustrated by the case  $b = 1.5$  (cf. Figure 2). As  $b$  approaches the value 2, the cluster CV goes to infinity, corresponding to a more and more chaotic cluster structure. The value  $b = 2$  (the gamma distribution) is a critical point where no cluster structure remains. The case  $b > 2$  constitutes a different regime for the Tweedie distribution, corresponding to exponentially tilted positive stable laws (cf. Jørgensen, 1997, Ch. 4). We also note that values of  $b$  between 0 and 1 do not correspond to Tweedie distributions (cf. Jørgensen, 1997, Ch. 4). This

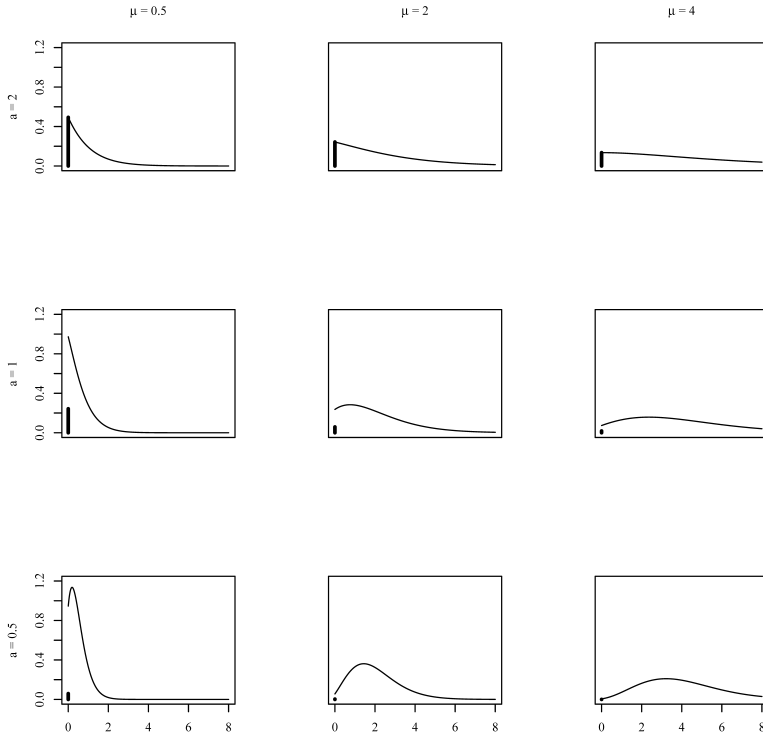


Figure 2: Plots of some Tweedie density functions for  $b = 1.5$ , for three different values of  $\mu$  (columns) and three different values of  $a$  (rows). The horizontal axis represents biomass. The probability mass at zero is represented by a vertical bar.

is consistent with the empirical fact that Taylor estimates of  $b$  below 1 are rare, and in fact such values may be spurious, in view of the above-mentioned negative bias of the Taylor estimator.

### 3 The Pearson estimating function

#### 3.1 General

We shall now introduce the Pearson estimating function and discuss the correction of bias, using terminology and results from Jørgensen & Knudsen (2004) and Holst & Jørgensen (2010). We consider data  $Y_{ij}$ , as defined in connection with (1.1), and the regression model  $\mu_i = f_i(\beta)$  introduced above. From now on we let the symbol  $\stackrel{\text{TM}}{=}$  denote results that hold under the Taylor model, inserting the estimator  $\bar{Y}_i$  for  $\mu_i$ .

Let us introduce the parameter  $\lambda = \log a$ , which will simplify many calculations in the following, and let  $\gamma = (\lambda, b)^\top$  denote the corresponding vector of variance parameters. The

quasi-score function for  $\boldsymbol{\beta}$  has the following form,

$$\boldsymbol{\psi}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{Y_{ij} - \mu_i}{e^{\lambda \mu_i^b}} \mathbf{x}_i,$$

where the  $\ell$ -vector  $\mathbf{x}_i$  is the  $\boldsymbol{\beta}$ -derivative of  $f_i$ . The Godambe information matrix for  $\boldsymbol{\beta}$  is the variance matrix for  $\boldsymbol{\psi}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma})$ ,

$$\mathbf{J}_\beta = \sum_{i=1}^k \frac{n_i}{e^{\lambda \mu_i^b}} \mathbf{x}_i \mathbf{x}_i^\top \stackrel{\text{TM}}{=} \text{diag} \left\{ \frac{n_i}{e^{\lambda \bar{Y}_i^b}} \right\}, \quad (3.1)$$

whose inverse yields the asymptotic variance of the quasi-likelihood estimator for  $\boldsymbol{\beta}$ .

Let  $C_i = e^{\lambda \mu_i^b}$  denote the variance of  $Y_{ij}$ , and define the weights  $W_{i\gamma}$  by

$$W_{i\gamma} = C_i^{-2} \frac{\partial C_i}{\partial \gamma} = C_i^{-1} \frac{\partial \log C_i}{\partial \gamma}, \quad (3.2)$$

where  $\gamma$  is a component of  $\boldsymbol{\gamma}$ , and where the required derivatives are

$$\frac{\partial \log C_i}{\partial \lambda} = 1 \text{ and } \frac{\partial \log C_i}{\partial b} = \log \mu_i.$$

Following Jørgensen & Knudsen (2004), the Pearson estimating function for the component  $\gamma$  has the following general form

$$\begin{aligned} \boldsymbol{\psi}_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma}) &= \sum_{i=1}^k \sum_{j=1}^{n_i} W_{i\gamma} \left[ (Y_{ij} - \mu_i)^2 - C_i \right] \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\partial \log C_i}{\partial \gamma} \left[ \frac{(Y_{ij} - \mu_i)^2}{C_i} - 1 \right]. \end{aligned} \quad (3.3)$$

which is an unbiased estimating function for  $\gamma$  based on the squared residuals  $(Y_{ij} - \mu_i)^2$  with mean  $C_i$ . This function is the derivative of the pseudo-likelihood function of Park & Cho (2004), and is equivalent to treating the squared residual  $(Y_{ij} - \mu_i)^2$  as a gamma variable, which is hence close in spirit to Perry's gamma regression method.

Let  $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}})$  denote the solution to the system of equations  $\boldsymbol{\psi}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{0}$  and  $\boldsymbol{\psi}_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{0}$ , where  $\boldsymbol{\psi}_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma})$  is obtained by stacking  $\boldsymbol{\psi}_\lambda(\boldsymbol{\beta}, \boldsymbol{\gamma})$  and  $\boldsymbol{\psi}_b(\boldsymbol{\beta}, \boldsymbol{\gamma})$ . We note that the quasi-score function  $\boldsymbol{\psi}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma})$  is  $\boldsymbol{\gamma}$ -insensitive, in the sense of Jørgensen & Knudsen (2004), meaning that  $\mathbb{E}_{\beta, \gamma} [\nabla_\gamma \boldsymbol{\psi}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma})] = \mathbf{0}$ , where  $\nabla_\gamma$  denotes the gradient operator. This implies that the quasi-likelihood estimator  $\hat{\boldsymbol{\beta}}$  depends relatively little on the choice of estimating function for  $\boldsymbol{\gamma}$  (cf. Jørgensen & Knudsen, 2004). In particular, under the Taylor model the quasi-likelihood estimator for  $\mu_i$  is  $\bar{Y}_i$ , independently of the choice of estimator for  $\boldsymbol{\gamma}$ .

By contrast, the pseudo-likelihood estimator  $\hat{\boldsymbol{\gamma}}$  may be heavily affected by the substitution of the fitted values  $\hat{\mu}_i = f_i(\hat{\boldsymbol{\beta}})$  into (3.3), inducing a nuisance parameter bias for  $\hat{\boldsymbol{\gamma}}$ . Although the estimating function  $\boldsymbol{\psi}_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma})$  is unbiased, this is not the case for the function  $\boldsymbol{\psi}_\gamma(\hat{\boldsymbol{\beta}}, \boldsymbol{\gamma})$ , but we shall now see how to correct this bias.



### 3.2 Bias correction

Following Jørgensen & Knudsen (2004), we now derive the bias-corrected estimating function  $\check{\psi}_\gamma(\boldsymbol{\beta}, \gamma)$  for  $\gamma$ , leading to the corrected estimator  $(\check{\boldsymbol{\beta}}, \check{\gamma})$ , which is the solution to the equations

$$\boldsymbol{\psi}_\beta(\boldsymbol{\beta}, \gamma) = \mathbf{0} \text{ and } \check{\psi}_\gamma(\boldsymbol{\beta}, \gamma) = \mathbf{0}. \quad (3.4)$$

The general form of the bias correction was derived by Holst & Jørgensen (2010). We now develop the bias-corrected estimating function in the special case of estimating  $\gamma$ .

For a given component of  $\gamma$ , the bias-corrected Pearson estimating function is defined by

$$\check{\psi}_\gamma(\boldsymbol{\beta}, \gamma) = \boldsymbol{\psi}_\gamma(\boldsymbol{\beta}, \gamma) - \text{tr} \left[ \mathbf{J}_\beta^{(\gamma)} \mathbf{J}_\beta^{-1} \right], \quad (3.5)$$

where  $\mathbf{J}_\beta^{(\gamma)}$  denotes the  $\gamma$ -derivative of  $\mathbf{J}_\beta$ . Since  $\mathbf{J}_\beta^{(\lambda)} = -\mathbf{J}_\beta$ , the bias-corrected Pearson estimating function for  $\lambda$  obtained from (3.3) and (3.5) is

$$\begin{aligned} \check{\psi}_\lambda(\boldsymbol{\beta}, \gamma) &= \sum_{i=1}^k \sum_{j=1}^{n_i} \left[ \frac{(Y_{ij} - \mu_i)^2}{e^{\lambda \mu_i^b}} - 1 \right] + \ell \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{(Y_{ij} - \mu_i)^2}{e^{\lambda \mu_i^b}} - (n - \ell) \end{aligned} \quad (3.6)$$

$$\stackrel{\text{TM}}{=} \sum_{i=1}^k \frac{(n_i - 1) S_i^2}{e^{\lambda \bar{Y}_i^b}} - (n - k), \quad (3.7)$$

where  $n = \sum n_i$  denotes the total sample size. The solution  $\check{\lambda}$  yields the conventional Pearson estimator for  $a$ , given by

$$\check{a} = e^{\check{\lambda}} = \frac{1}{n - \ell} \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{(Y_{ij} - \check{\mu}_i)^2}{\check{\mu}_i^b} \stackrel{\text{TM}}{=} \frac{1}{n - k} \sum_{i=1}^k \frac{(n_i - 1) S_i^2}{\bar{Y}_i^{\check{b}}}, \quad (3.8)$$

where  $\check{\mu}_i = f_i(\check{\boldsymbol{\beta}})$  denote the fitted values corresponding to the estimator  $\check{\boldsymbol{\beta}}$ .

Let us now turn to the estimation of  $b$ , with Pearson estimating function (3.3) given by

$$\boldsymbol{\psi}_b(\boldsymbol{\beta}, \gamma) = \sum_{i=1}^k \sum_{j=1}^{n_i} \left[ \frac{(Y_{ij} - \mu_i)^2}{e^{\lambda \mu_i^b}} - 1 \right] \log \mu_i. \quad (3.9)$$

To derive the bias-corrected estimating function for  $b$ , we again combine (3.3) and (3.5). To derive the correction term, we need the derivative  $\mathbf{J}_\beta^{(b)}$ , which is given by

$$\mathbf{J}_\beta^{(b)} = - \sum_{i=1}^k \frac{n_i \log \mu_i}{e^{\lambda \mu_i^b}} \mathbf{x}_i \mathbf{x}_i^\top \stackrel{\text{TM}}{=} - \text{diag} \left\{ \frac{n_i \log \bar{Y}_i}{e^{\lambda \bar{Y}_i^b}} \right\}. \quad (3.10)$$

Hence, the correction term becomes

$$\begin{aligned}
\text{tr} \left[ \mathbf{J}_\beta^{(b)} \mathbf{J}_\beta^{-1} \right] &= -\text{tr} \left( \sum_{i=1}^k \frac{n_i \log \mu_i}{e^{\lambda \mu_i^b}} \mathbf{x}_i \mathbf{x}_i^\top \mathbf{J}_\beta^{-1} \right) \\
&= -\sum_{i=1}^k \text{tr} \left( \frac{n_i \log \mu_i}{e^{\lambda \mu_i^b}} \mathbf{x}_i^\top \mathbf{J}_\beta^{-1} \mathbf{x}_i \right) \\
&= -\sum_{i=1}^k \frac{n_i \log \mu_i}{e^{\lambda \mu_i^b}} \mathbf{x}_i^\top \mathbf{J}_\beta^{-1} \mathbf{x}_i \\
&\stackrel{\text{TM}}{=} -\sum_{i=1}^k \log \bar{Y}_i.
\end{aligned} \tag{3.11}$$

The bias-corrected Pearson estimating function for  $b$  is hence

$$\begin{aligned}
\check{\psi}_b(\boldsymbol{\beta}, \gamma) &= \sum_{i=1}^k \sum_{j=1}^{n_i} \left[ \frac{(Y_{ij} - \mu_i)^2}{e^{\lambda \mu_i^b}} - 1 \right] \log \mu_i + \sum_{i=1}^k \frac{n_i \log \mu_i}{e^{\lambda \mu_i^b}} \mathbf{x}_i^\top \mathbf{J}_\beta^{-1} \mathbf{x}_i \\
&= \sum_{i=1}^k \sum_{j=1}^{n_i} \left[ \frac{(Y_{ij} - \mu_i)^2 + \mathbf{x}_i^\top \mathbf{J}_\beta^{-1} \mathbf{x}_i}{e^{\lambda \mu_i^b}} - 1 \right] \log \mu_i
\end{aligned} \tag{3.12}$$

$$\stackrel{\text{TM}}{=} \sum_{i=1}^k \left[ \frac{(n_i - 1) S_i^2}{e^{\lambda \bar{Y}_i^b}} - (n_i - 1) \right] \log \bar{Y}_i. \tag{3.13}$$

The correction term  $\mathbf{x}_i^\top \mathbf{J}_\beta^{-1} \mathbf{x}_i$  appearing in (3.12) is of order  $1/n_i$ , and positive, which will tend to make  $\check{b}$  bigger than  $\hat{b}$ , confirming the negative bias of  $\hat{b}$ .

Under the Taylor model, the estimating functions (3.7) and (3.13) may be interpreted as follows. The bias-corrected estimators  $\check{\lambda}$  and  $\check{b}$  may be obtained by solving the equations  $\bar{Y}_i = \mu_i$  for  $i = 1, \dots, k$  together with the two unbiased estimating equations

$$\sum_{i=1}^k (n_i - 1) \frac{\partial \log C_i}{\partial \gamma} \left[ \frac{S_i^2}{C_i} - 1 \right] = 0, \tag{3.14}$$

where  $\gamma$  is either  $\lambda$  or  $b$ . Comparing (3.14) with (3.3), the bias-corrected Pearson estimator is seen to be identical to Perry's gamma regression estimator under the Taylor model, based on treating  $S_i^2$  as gamma distributed with mean  $C_i$  and variance proportional to  $C_i^2/(n_i - 1)$ , but with the additional advantage that the inverse Godambe information matrix yields the correct standard errors for  $\check{b}$ , cf. Section 5. The unbiasedness of the equations (3.14) implies the consistency of the estimator  $\check{b}$  in the large  $k$  limit (even for sparse data), as well as for  $n_i \rightarrow \infty$  for all  $i$  (fixed  $k$ ). The large  $k$  limit, where the dimension of the nuisance parameter  $\boldsymbol{\beta} = (\mu_1, \dots, \mu_k)^\top$  goes to infinity under the Taylor model (cf. Section 2), is an example of a "Neyman-Scott" problem, see e.g. Li *et al.* (2003) for a recent discussion. Our solution to the Neyman-Scott problem involves correcting the bias of the estimating function rather than

that of the estimator, and hence provides a natural extension of the conventional solution. In fact, our bias-corrected estimate for  $a$  in (3.8) is identical to the conventional Pearson estimator, which involves a simple degrees-of-freedom correction.

## 4 The Newton scoring algorithm

We now consider the Newton scoring algorithm for the parameter vector  $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ , which is a version of the Newton method applied to the equations (3.4), based on replacing the derivatives of the estimating functions by their expectations. Jørgensen & Knudsen (2004) showed that by taking advantage of the  $\boldsymbol{\gamma}$ -insensitivity of the quasi-score function  $\boldsymbol{\psi}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma})$  (as noted above), a so-called *chaser algorithm* is obtained. The chaser algorithm oscillates between updating  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$  as follows:

$$\boldsymbol{\beta}^* = \boldsymbol{\beta} - \mathbf{J}_\beta^{-1} \boldsymbol{\psi}_\beta \quad (4.1)$$

$$\boldsymbol{\gamma}^* = \boldsymbol{\gamma} - \check{\mathbf{S}}_\gamma^{-1} \check{\boldsymbol{\psi}}_\gamma, \quad (4.2)$$

where  $\check{\mathbf{S}}_\gamma$  is the sensitivity matrix (the expected derivative) for  $\check{\boldsymbol{\psi}}_\gamma$  discussed below, see also Section 5. In each of the two update steps (4.1) and (4.2), the most recent values of  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$  are used as arguments in the functions on the right-hand side of the equations. The asymmetry of the  $\boldsymbol{\gamma}$ -insensitivity condition implies that convergence of the  $\boldsymbol{\gamma}$  part is conditional on the  $\boldsymbol{\beta}$  part having converged first, giving rise to the *chaser* qualification. A small simplification may be obtained if we insert the estimator  $\check{\lambda}$  from (3.8) into (4.2), which makes  $\check{\boldsymbol{\psi}}_\lambda(\boldsymbol{\beta}, \boldsymbol{\gamma}) = 0$ .

We shall now calculate the sensitivity matrix

$$\check{\mathbf{S}}_\gamma = \mathbb{E}_{\boldsymbol{\beta}, \boldsymbol{\gamma}} [\nabla_\gamma \check{\boldsymbol{\psi}}_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma})] = \mathbf{S}_\gamma - \mathbf{T}_\gamma, \quad (4.3)$$

say, where the last two terms correspond to the expected derivatives of the corresponding terms of (3.5). Thus,  $\mathbf{S}_\gamma$  denotes the sensitivity of  $\boldsymbol{\psi}_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma})$ , with entries of the form

$$S_{\gamma_1 \gamma_2} = - \sum_{i=1}^k n_i C_i^{-2} \frac{\partial C_i}{\partial \gamma_1} \frac{\partial C_i}{\partial \gamma_2} = - \sum_{i=1}^k n_i \frac{\partial \log C_i}{\partial \gamma_1} \frac{\partial \log C_i}{\partial \gamma_2},$$

where  $\gamma_1$  and  $\gamma_2$  denote either  $\lambda$  or  $b$ , giving

$$\mathbf{S}_\gamma = \begin{bmatrix} -n & -\sum_{i=1}^k n_i \log \mu_i \\ -\sum_{i=1}^k n_i \log \mu_i & -\sum_{i=1}^k n_i \log^2 \mu_i \end{bmatrix}.$$

The fact that  $\mathbf{S}_\gamma$  is constant as a function of the variance parameters  $\lambda$  and  $b$  shows that this particular parametrization has the advantage of approximately linearizing the estimating functions  $\psi_\lambda$  and  $\psi_b$ , which in turn helps speed up the algorithm. Using the Taylor estimates as starting values, the algorithm required between 5 and 18 iterations for the benthic macrofauna data of Section 6.

Three of the four entries of the correction term  $\mathbf{T}_\gamma$  turn out to be zero, as seen by direct inspection of the correction term  $n - \ell$  of (3.6). Only the  $(b, b)$  entry of  $\mathbf{T}_\gamma$  is non-zero, and has

the form  $\text{tr} \left[ \mathbf{J}_\beta^{(b)} \mathbf{J}_\beta^{-1} \mathbf{J}_\beta^{(b)} \mathbf{J}_\beta^{-1} - \mathbf{J}_\beta^{(b,b)} \mathbf{J}_\beta^{-1} \right]$  (cf. Holst & Jørgensen, 2010), where  $\mathbf{J}_\beta$  and  $\mathbf{J}_\beta^{(b)}$  are obtained from (3.1) and (3.10), respectively, and

$$\mathbf{J}_\beta^{(b,b)} = \sum_{i=1}^k \frac{n_i \log^2 \mu_i}{e^{\lambda \mu_i^b}} \mathbf{x}_i \mathbf{x}_i^\top$$

is the second derivative of  $\mathbf{J}_\beta$  with respect to  $b$ . In the case of the Taylor model, the fourth term is also zero, as seen directly from (3.11).

## 5 The Godambe information matrix

We shall now calculate the asymptotic variance of the estimators, as obtained from the inverse Godambe information matrix, whose general form is  $\mathbf{J}_\theta^{-1} = \mathbf{S}_\theta^{-1} \mathbf{V}_\theta \mathbf{S}_\theta^{-\top}$  for a vector parameter  $\boldsymbol{\theta}$ , where  $-\top$  denotes inverse transpose. Here  $\mathbf{S}_\theta = \mathbb{E}_\theta [\nabla_\theta \boldsymbol{\psi}_\theta(\boldsymbol{\theta})]$  denotes the sensitivity matrix for  $\boldsymbol{\theta}$ , as discussed above, whereas  $\mathbf{V}_\theta = \text{Var}_\theta [\boldsymbol{\psi}_\theta(\boldsymbol{\theta})]$  denotes the variability matrix. In the special case where  $\boldsymbol{\psi}_\theta$  is a quasi-score function, the Godambe information matrix reduces to  $\mathbf{J}_\theta = \mathbf{V}_\theta = -\mathbf{S}_\theta$ , as we have already seen in connection with (3.1).

In the present case of estimation of the vector parameter  $\boldsymbol{\theta}$  with components  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$ , Jørgensen & Knudsen (2004) showed that the  $\boldsymbol{\gamma}$ -insensitivity of  $\boldsymbol{\psi}_\beta$  implies that the asymptotic variance of  $\tilde{\boldsymbol{\beta}}$  is given by  $\mathbf{J}_\beta^{-1}$ , which is hence unaffected by the presence of the nuisance parameter  $\boldsymbol{\gamma}$ . We shall now derive the asymptotic variance for the estimator  $\tilde{\boldsymbol{\gamma}}$  using the general results of Holst & Jørgensen (2010), based on the full Godambe information matrix for  $\boldsymbol{\theta}$ .

In order to calculate the asymptotic variance for  $\tilde{\boldsymbol{\gamma}}$ , we need to calculate the variability  $\mathbf{V}_\theta$  (which is unaffected by the non-stochastic correction term of (3.5)) and the sensitivity  $\tilde{\mathbf{S}}_\theta = \mathbf{S}_\theta - \mathbf{T}_\theta$ , say, similar to (4.3). The variability for  $\boldsymbol{\theta}$  has the form

$$\mathbf{V}_\theta = \begin{bmatrix} \mathbf{V}_\beta & \mathbf{V}_{\beta\gamma} \\ \mathbf{V}_{\gamma\beta} & \mathbf{V}_\gamma \end{bmatrix},$$

where  $\mathbf{V}_\beta = \mathbf{J}_\beta$ , whereas  $\mathbf{V}_{\beta\gamma} = \mathbf{V}_{\gamma\beta}^\top$  and  $\mathbf{V}_\gamma$  depend on third and fourth moments of  $Y_{ij}$ , respectively. In order to avoid this dependence on high-order moments, we propose instead to use the empirical versions of  $\mathbf{V}_\gamma$  and  $\mathbf{V}_{\gamma\beta}$ , which are given by

$$\check{\mathbf{V}}_\gamma = \sum_{i=1}^k \sum_{j=1}^{n_i} \check{\boldsymbol{\psi}}_{\gamma ij}(\check{\boldsymbol{\beta}}, \check{\boldsymbol{\gamma}}) \check{\boldsymbol{\psi}}_{\gamma ij}^\top(\check{\boldsymbol{\beta}}, \check{\boldsymbol{\gamma}}) \quad \text{and} \quad \check{\mathbf{V}}_{\gamma\beta} = \sum_{i=1}^k \sum_{j=1}^{n_i} \check{\boldsymbol{\psi}}_{\gamma ij}(\check{\boldsymbol{\beta}}, \check{\boldsymbol{\gamma}}) \boldsymbol{\psi}_{\beta ij}^\top(\check{\boldsymbol{\beta}}, \check{\boldsymbol{\gamma}}).$$

Since we are mainly interested in the asymptotic variance for  $\tilde{\boldsymbol{\gamma}}$ , there is no need to explicitly calculate the full Godambe information matrix for  $\boldsymbol{\theta}$ . Using the results of Holst & Jørgensen (2010), we find that the  $\boldsymbol{\gamma}$  block of  $\mathbf{J}_\theta^{-1}$  is given by

$$\mathbf{J}_\theta^\gamma = \check{\mathbf{S}}_\gamma^{-1} \left( \mathbf{V}_\gamma + \check{\mathbf{S}}_{\gamma\beta} \mathbf{J}_\beta^{-1} \check{\mathbf{S}}_{\gamma\beta}^\top + \check{\mathbf{S}}_{\gamma\beta} \mathbf{J}_\beta^{-1} \mathbf{V}_{\beta\gamma} + \mathbf{V}_{\gamma\beta} \mathbf{J}_\beta^{-1} \check{\mathbf{S}}_{\gamma\beta}^\top \right) \check{\mathbf{S}}_\gamma^{-1}, \quad (5.1)$$

where  $\check{\mathbf{S}}_\gamma$  was derived above, and where we shall use the empirical version  $\check{\mathbf{J}}_\theta^\gamma$ , obtained by substituting the empirical variabilities  $\check{\mathbf{V}}_\gamma$  and  $\check{\mathbf{V}}_{\gamma\beta}$  into (5.1). One additional element of (5.1) required is the block  $\check{\mathbf{S}}_{\gamma\beta}$ , which we shall now calculate.

Consider the function  $\psi_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma})$  from (3.3), where  $\gamma$  is a component of  $\boldsymbol{\gamma}$ . Using the linearity of the expectation and derivative operators, we find that the matrix  $\mathbf{S}_{\gamma\beta}$  has entries defined by

$$\begin{aligned}\mathbb{E}_\theta [\nabla_\beta \psi_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma})] &= \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\partial \log C_i}{\partial \gamma} \frac{1}{C_i} \mathbb{E}_\theta \left\{ \nabla_\beta \left[ (Y_{ij} - \mu_i)^2 - C_i \right] \right\} \\ &= - \sum_{i=1}^k n_i \frac{\partial \log C_i}{\partial \gamma} \frac{1}{C_i} \nabla_\beta C_i \\ &= - \sum_{i=1}^k n_i \frac{\partial \log C_i}{\partial \gamma} \frac{b}{\mu_i} \mathbf{x}_i^\top.\end{aligned}$$

It follows that

$$\mathbf{S}_{\gamma\beta} = - \sum_{i=1}^k \frac{n_i b}{\mu_i} \begin{bmatrix} 1 \\ \log \mu_i \end{bmatrix} \mathbf{x}_i^\top.$$

The correction term  $\mathbf{T}_{\gamma\beta}$  has entries given by

$$\text{tr} \left[ \mathbf{J}_\beta^{(\gamma)} \mathbf{J}_\beta^{-1} \mathbf{J}_\beta^{(\beta_m)} \mathbf{J}_\beta^{-1} - \mathbf{J}_\beta^{(\gamma, \beta_m)} \mathbf{J}_\beta^{-1} \right],$$

where  $\beta_m$  denotes a component of  $\boldsymbol{\beta}$ . The entries corresponding to  $\gamma = \lambda$  are zero, since

$$\text{tr} \left[ \mathbf{J}_\beta^{(\lambda)} \mathbf{J}_\beta^{-1} \mathbf{J}_\beta^{(\beta_m)} \mathbf{J}_\beta^{-1} - \mathbf{J}_\beta^{(\lambda, \beta_m)} \mathbf{J}_\beta^{-1} \right] = \text{tr} \left[ -\mathbf{J}_\beta \mathbf{J}_\beta^{-1} \mathbf{J}_\beta^{(\beta_m)} \mathbf{J}_\beta^{-1} + \mathbf{J}_\beta^{(\beta_m)} \mathbf{J}_\beta^{-1} \right] = 0.$$

The entry for  $b$  is  $\text{tr} \left[ \mathbf{J}_\beta^{(b)} \mathbf{J}_\beta^{-1} \mathbf{J}_\beta^{(\beta_m)} \mathbf{J}_\beta^{-1} - \mathbf{J}_\beta^{(b, \beta_m)} \mathbf{J}_\beta^{-1} \right]$ , where

$$\mathbf{J}_\beta^{(\beta_m)} = - \sum_{i=1}^k \frac{n_i b x_{im}}{e^{\lambda \mu_i^{b+1}}} \mathbf{x}_i \mathbf{x}_i^\top + \sum_{i=1}^k \frac{n_i}{e^{\lambda \mu_i^b}} \left[ \mathbf{x}_i^{(\beta_m)} \mathbf{x}_i^\top + \mathbf{x}_i \mathbf{x}_i^{(\beta_m)\top} \right] \stackrel{\text{TM}}{=} - \frac{n_m b}{e^{\lambda \bar{Y}_m^{b+1}}} \mathbf{x}_m \mathbf{x}_m^\top,$$

and where under the Taylor model  $\mathbf{x}_m$  has 1 at position  $m$  and zero elsewhere. Here  $x_{im}$  is the  $m$ th component of  $\mathbf{x}_i$ , and  $\mathbf{x}_i^{(\beta_m)}$  is the  $\beta_m$ -derivative of  $\mathbf{x}_i$ . The mixed derivative of  $\mathbf{J}_\beta$  is

$$\begin{aligned}\mathbf{J}_\beta^{(b, \beta_m)} &= - \sum_{i=1}^k \frac{n_i x_{im} (1 - b \log \mu_i)}{e^{\lambda \mu_i^{b+1}}} \mathbf{x}_i \mathbf{x}_i^\top - \sum_{i=1}^k \frac{n_i \log \mu_i}{e^{\lambda \mu_i^b}} \left[ \mathbf{x}_i^{(\beta_m)} \mathbf{x}_i^\top + \mathbf{x}_i \mathbf{x}_i^{(\beta_m)\top} \right] \\ &\stackrel{\text{TM}}{=} - \frac{n_m (1 - b \log \bar{Y}_m)}{e^{\lambda \bar{Y}_m^{b+1}}} \mathbf{x}_m \mathbf{x}_m^\top.\end{aligned}$$

These results define the entries of the matrices  $\mathbf{S}_{\gamma\beta}$  and  $\mathbf{T}_{\gamma\beta}$ , from which we may calculate  $\check{\mathbf{S}}_{\gamma\beta} = \mathbf{S}_{\gamma\beta} - \mathbf{T}_{\gamma\beta}$ .

## 6 Analysis of the benthic macrofauna data

Benthic macrofauna was collected in 2008 from two Danish estuaries separated by about 100 km, Odense Fjord (OF) and Roskilde Fjord (RF), located on the islands of Fyn and Sjælland,

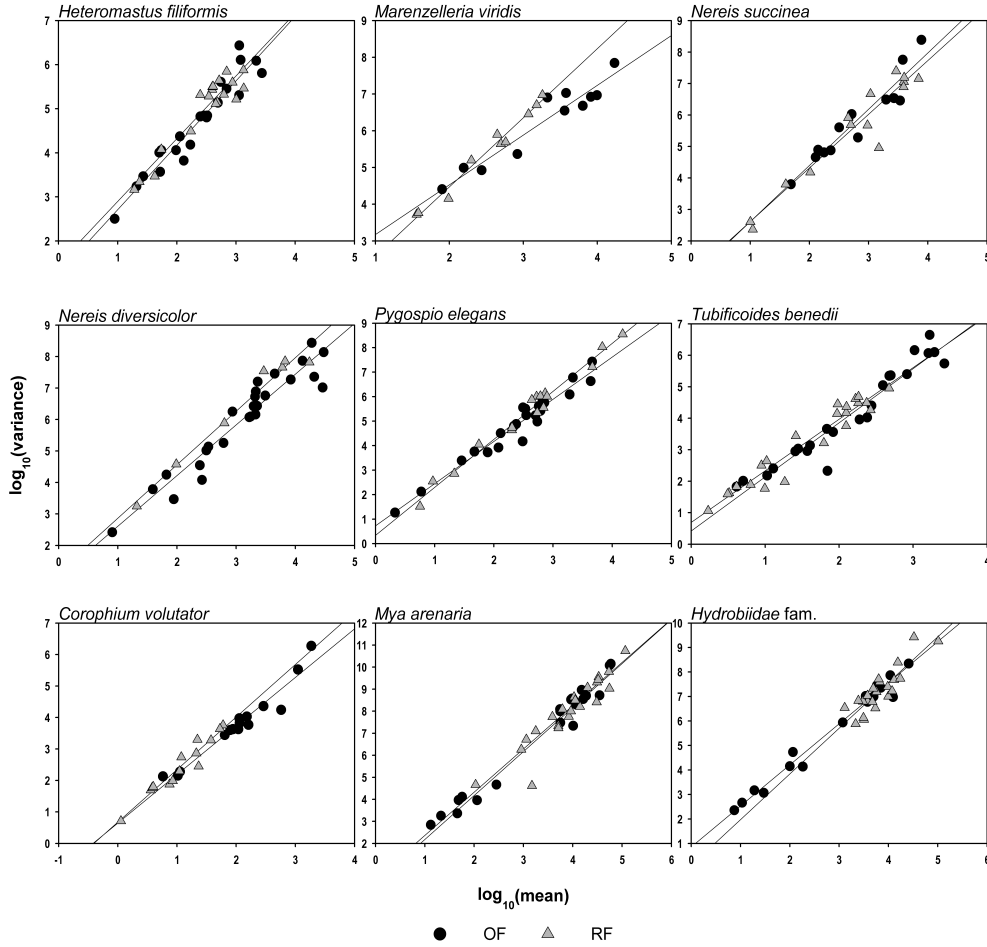


Figure 3: Plots of  $\log S^2$  versus  $\log \bar{Y}$  for the benthic biomass data.

respectively. Animals were collected from 28 and 24 sites, respectively, well spread out within each of the two estuaries, and representing the whole depth range of the estuaries (0 to 5 m). Four replicate sediment cores were taken at each site, separated by a few meters, using a 15 cm in diameter steel corer to a depth of 30 cm. Sediments were sieved at 1 mm and animals retained were identified to species, counted and weighed after drying at 60 °C overnight. Nine of the most common species present at both locations were selected for analysis: softshell clams (*Mya arenaria*), mud snails (family *Hydrobiidae*), mud shrimps (*Corophium volutator*), ragworms (*Nereis diversicolor*), and five other species of oligochaetes (*Tubificoides benedii*) and polychaetes (*Heteromastus filiformis*, *Marenzelleria viridis*, *Nereis succinea* and *Pygospio elegans*). The number of non-empty sites for each combination of species and location (see Table 1 below) varied between 29% and 93%.

Figure 3 shows plots of  $\log S^2$  versus  $\log \bar{Y}$  (base 10 logarithm) for the variable biomass, with straight lines fitted by Taylor's method. The plots show good agreement with the hypothesis

Species	Odense Fjord (OF)					Roskilde Fjord (RF)				
	$k$	MLE	Tayl	Pear	SE	$k$	MLE	Tayl	Pear	SE
<i>H. filiformis</i>	15	1.41	1.46	1.55	0.13	15	1.32	1.42	1.48	0.21
<i>M. viridis</i>	22	1.33	1.35	1.38	0.14	19	1.46	1.88	1.88	0.29
<i>N. succinea</i>	17	1.61	1.79	1.94	0.19	22	1.49	1.71	1.78	0.17
<i>N. diversicolor</i>	20	1.51	1.61	1.68	0.09	21	1.52	1.70	1.77	0.24
<i>P. elegans</i>	11	1.69	1.72	1.76	0.13	10	1.71	1.94	1.94	0.10
<i>T. benedii</i>	13	1.53	1.72	1.76	0.07	15	1.39	1.64	1.65	0.13
<i>C. volutator</i>	26	1.45	1.54	1.61	0.09	7	1.56	1.67	1.70	0.26
<i>M. arenaria</i>	23	1.76	1.96	1.98	0.09	18	1.68	1.98	1.93	0.14
<i>Hydrobiidae</i>	23	1.53	1.68	1.69	0.09	22	1.68	1.87	1.96	0.15

Table 1: Estimates of the power parameter  $b$  for Odense Fjord and Roskilde Fjord ( $k$  = number of non-empty sites, MLE = maximum likelihood estimate, Tayl = Taylor’s regression estimate, Pear = corrected Pearson estimate, SE = standard error for Pearson estimate).

of linearity in the double logarithmic scale, taking into account the small number of replicates ( $n_i = 4$ ). Table 1 compares estimates of  $b$  based on the maximum likelihood method, Taylor’s method and the corrected Pearson method. As expected (cf. the remark after Eq. (3.13)), the corrected Pearson estimates are, with one exception, bigger than the corresponding Taylor and maximum likelihood estimates, corresponding to a bias of as much as  $-0.15$  for the Taylor estimator and as much as  $-0.38$  for the maximum likelihood estimator. The estimates of  $a$  (Table 2) do not indicate any bias problems for the Taylor estimator, as compared with the corrected Pearson estimator, whereas the maximum likelihood estimates of  $a$  seem to be positively biased.

The corrected Pearson estimates of  $b$  in Table 1 range from 1.38 to 1.98, corresponding to cluster CVs in the range 0.8–7 under the Tweedie cluster model (2.3). The corresponding cluster structures range from quite variable to very chaotic, which suggests that the species studied here may vary considerably with respect to the social behaviour of the individuals of each species. A detailed interpretation of these results is outside the scope of the present paper, and will instead be published elsewhere.

The corrected Pearson estimates for  $b$  show good agreement between the estimates of  $b$  for OF and RF for each species. An overall test for species-wise equality of the values of  $b$  was performed by a  $\chi^2$  test, treating the estimators as normal with standard deviations known and equal to the SE values in the table ( $X^2 = 7.35$ ,  $df = 9$ ,  $p = 0.6$ ). This test does not indicate any significant geographical variation of the  $b$  values for the nine species analyzed here, in agreement with the hypothesis that  $b$  is a species-specific index of aggregation.

It is clear from these results that accurate estimation of the power parameter  $b$  is crucial for the correct interpretation of the results of fitting Taylor’s power law to ecological data. We conclude that the bias-corrected Pearson estimator for  $b$  introduced here is a good candidate for an accurate and easily calculated estimator.

Species	Odense Fjord (OF)				Roskilde Fjord (RF)			
	MLE	Tayl	Pear	SE	MLE	Tayl	Pear	SE
<i>H. filiformis</i>	27.82	17.63	13.65	0.75	52.19	29.74	26.42	1.29
<i>M. viridis</i>	90.26	65.84	67.04	1.17	80.60	5.06	5.57	1.74
<i>N. succinea</i>	38.31	6.69	3.59	1.28	56.25	8.08	7.53	1.23
<i>N. diversicolor</i>	29.94	10.00	9.24	0.75	76.11	14.66	10.91	1.90
<i>P. elegans</i>	8.07	5.29	5.27	0.79	13.67	2.51	2.84	0.62
<i>T. benedii</i>	9.06	2.64	2.96	0.38	13.55	4.94	5.94	0.59
<i>C. volutator</i>	7.51	4.45	3.78	0.46	8.84	4.76	4.94	0.72
<i>M. arenaria</i>	31.57	2.50	2.99	0.77	39.45	1.68	4.01	1.24
<i>Hydrobiidae</i>	26.27	7.17	7.87	0.72	8.58	1.27	0.91	1.35

Table 2: Estimates of the dispersion parameter  $a$  for Odense Fjord and Roskilde Fjord (MLE = maximum likelihood estimate, Tayl = Taylor’s regression estimate, Pear = bias-corrected Pearson estimate, SE = standard error for Pearson estimate). See Table 1 for the number of non-empty sites.

## Acknowledgements

We are grateful to the Editor and an anonymous referee for their useful comments on a earlier version of the paper. This research was supported by FAPESP and CNPq (Brazil), and the Danish Natural Science Research Council.

## References

- DOWNING, J. A. 1986. Spatial heterogeneity: evolved behaviour or mathematical artefact. *Nature*, **323**, 255–257.
- DUNN, P. K. 2009. Package tweedie: Tweedie exponential family models, Ver. 2.0.2. *R Contributed Packages*, <http://cran.r-project.org/web/packages/tweedie/tweedie.pdf>.
- EISLER, Z., BARTOS, I., & KERTÉSZ, J. 2008. Fluctuation scaling in complex systems: Taylor’s law and beyond. *Adv. Phys.*, **57**, 89–142.
- FRONCZAK, A., & FRONCZAK, P. 2010. Origins of Taylor’s power law for fluctuation scaling in complex systems. *Phys. Rev. E*, **81**, 066112.
- HOLST, R., & JØRGENSEN, B. 2010. Efficient and robust estimation for a class of generalized linear longitudinal mixed models. *arXiv:1008.2870v1*.
- JØRGENSEN, B. 1997. *The Theory of Dispersion Models*. London: Chapman & Hall.
- JØRGENSEN, B., & KNUDSEN, S. J. 2004. Parameter orthogonality and bias adjustment for estimating functions. *Scand. J. Statist.*, **31**, 93–114.
- JØRGENSEN, B., MARTÍNEZ, J. R., & TSAO, M. 1994. Asymptotic behaviour of the variance function. *Scand. J. Statist.*, **21**, 223–243.



- KENDAL, W. S. 2004. Taylor's ecological power law as a consequence of scale invariant exponential dispersion models. *Ecological Complexity*, **1**, 193–209.
- KILPATRICK, A. M., & IVES, A. R. 2003. Species interactions can explain Taylor's power law for ecological time series. *Nature*, **422**, 65–68.
- LI, H., LINDSAY, B. G., & WATERMAN, R. P. 2003. Efficiency of Projected Score Methods in Rectangular Array Asymptotics. *J. Roy. Statist. Soc. B*, **65**, 191–208.
- PARK, H., & CHO, K. 2004. Use of covariates in Taylor's power law for sequential sampling in pest management. *Journal of Agricultural, Biological, and Environmental Statistics*, **9**, 462–478.
- PEDIGO, L. P., & BUNTIN, G. D. 1994. *Handbook of Sampling Methods for Arthropods in Agriculture*. second edn. Florida: CRC Press.
- PERRY, J. N. 1981. Taylor's power law for dependence of variance on mean in animal populations. *Appl. Statist.*, **30**, 254–263.
- TAYLOR, L. R. 1961. Aggregation, variance and the mean. *Nature*, **189**, 732–735.
- TAYLOR, L. R., TAYLOR, R. A. J., WOIWOD, I. P., & PERRY, J. N. 1983. Behavioural dynamics. *Nature*, **303**, 801–804.
- TITMUS, G. 1983. Are animal populations really aggregated? *Oikos*, **40**, 64–68.
- TWEEDIE, M. C. K. 1984. An index which distinguishes between some important exponential families. *Pages 579–604 of: GHOSH, J. K., & ROY, J. (eds), Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference*. Calcutta: Indian Statistical Institute.