

# Application of non-linear quantile regression to macrozoobenthic species distribution modelling: comparing two contrasting basins

F. Cozzoli<sup>1,\*</sup>, T. J. Bouma<sup>1</sup>, T. Ysebaert<sup>1,2</sup>, P. M. J. Herman<sup>1</sup>

<sup>1</sup>Netherlands Institute of Sea Research (NIOZ), P.B. 140, 4400 AC Yerseke, The Netherlands

<sup>2</sup>Wageningen University, Institute for Marine Resources and Ecosystem Studies (IMARES), P.B. 77, 4400 AB Yerseke, The Netherlands

**ABSTRACT:** The occurrence and distribution of macrozoobenthos in estuaries are strongly related to sediment grain-size characteristics. However, statistical prediction of the distribution of benthic populations as a response to a single environmental gradient has proven to be difficult, because the focal variable may set upper limits to the abundance, but other (partly uncorrelated) variables may cause considerable deviation from the maximum. A multi-quantile regression approach is better suited to characterize biota–environment relationships than a single (average or boundary) estimation, because it shows the variation in responses and quantifies the relative importance of other unmeasured factors. Here, a univariate application of non-linear quantile regression is proposed to account for heteroskedasticity and non-linearity in the biological response to sediment grain size. The analysis was applied to a large macrozoobenthic dataset from the SW Delta area (The Netherlands) to compare the relationships between sediment granulometry and macrozoobenthos in 2 neighboring but differing temperate coastal ecosystems (Oosterschelde and Westerschelde). Preference of individual species for grain size was consistent between both systems, although in general, a slightly higher median grain size (ca. +60% in grain diameter) was preferred in the Oosterschelde than in the Westerschelde. The major difference in the community was, however, that mud-preferring species dominated the assemblage in the Westerschelde, and sand-preferring species dominated the Oosterschelde. Although the prevalence of muddy and sandy sediments in both systems is similar, in the Westerschelde, strong hydrodynamic stress is correlated with sandy habitats, causing impoverishment of assemblages at sandy sites. In the Oosterschelde, sandy sediments are usually associated with much more benign conditions and have the richest species assemblage.

**KEY WORDS:** Species distribution models · Macrozoobenthos · Sediment · Oosterschelde-Westerschelde · Heteroskedasticity · Quantile regression · Regression splines

— Resale or republication not permitted without written consent of the publisher —

## INTRODUCTION

Macrozoobenthic organisms play a fundamental role in estuarine ecosystem dynamics. They are central components of estuarine food webs (Herman et al. 1999) and they affect sediment biogeochemistry (Heip et al. 1995). The importance of estuaries for human activities (fisheries, aquaculture, sand extraction, transport of goods and people, recreation) leads to changes in habitats (Kraufvelin et al. 2001) and

directs research towards the study of the relationships between environmental conditions and distribution of macrozoobenthic organisms (e.g. Snelgrove & Butman 1994, Herman et al. 2001, Thrush et al. 2004, Degraer et al. 2008).

Among the environmental factors directly delimiting the potential niche of macrozoobenthic species, the sediment composition is of great relevance (Gray 1974). On one hand, sediment composition is a proxy for typically covarying stressors, such as the hydro-

\*Email: francesco.cozzoli@nioz.nl

dynamic regime (Allen 1985, Snelgrove & Butman 1994), depth (or emersion time) and resource distribution (Heip et al. 1995, Herman et al. 1999, van der Wal et al. 2008). On the other hand, it influences benthic biota directly as it constitutes the physical matrix in which benthic organisms live, gather food and construct burrows (Gray 1974). Sediment granulometry correlates with hydrodynamic stress (van Prooijen & Winterwerp 2010), water permeability (van Ledden et al. 2004), organic matter content (Mayer 1994, Curry et al. 2007) and microbial oxygen consumption (Sutherland et al. 1998). Degraer et al. (2008) found that sediment median granulometry and mud percentage are the most important environmental variables in determining macrozoobenthos spatial distribution in the Belgian North Sea. The influence of sediment composition on macrozoobenthic distributions is evident also in estuarine environments, where the salinity gradient can have major relevance (Ysebaert et al. 2002).

The variability in the morphological, hydrological and chemical characteristics (Day et al. 1989) of estuarine environments generates spatial and temporal patterns in species distribution (Wolff 1983, Ysebaert et al. 2003). The accurate statistical prediction of these patterns as a response to an environmental gradient has proven to be extremely difficult (Butman 1987, Snelgrove & Butman 1994, Chapman et al. 2010) as the observed distributions are the outcome of complex interactions between hydrodynamics, sediment dynamics and benthic biology (Herman et al. 2001). When unmeasured factors interact with the measured factor, they generate an error distribution with respect to the variables included in the model. Complicated forms of non-linear and heterogeneous response distributions can be expected in observational studies where many important processes interact (Cade et al. 2005). Stochastic factors also contribute to variability in the response of individual sample densities. Univariate species-environment relationships commonly show heteroskedastic error distributions (White 1980): variances are smaller in limiting conditions, but they increase at more suitable sites (Terrell et al. 1996, Thomson et al. 1996, Cade et al. 1999, Anderson 2008). The local spatial and temporal covariance structures between environmental factors may differ from one system to another, thereby decreasing the predictability of species' responses across different scales of observation (Thrush et al. 2005, Chapman et al. 2010, Kraufvelin et al. 2011).

Traditional regression methods like Ordinary Least Squares (OLS) focus exclusively on changes in the

means. They assume that the response variable has constant error variance, regardless of the value of the predictor variable. Thus, central estimators are not able to cope with data sets where the variance changes with the mean. In a regime of limitation by subsidiary factors (high prevalence of zero observations along the entire gradient), they are not representative of the higher densities and they may fail to distinguish real non-zero changes (Terrell et al. 1996, Cade & Noon 2003).

To overcome these confounding effects, recent studies on species distribution models focus on models that predict the likelihood of species presence (Franklin 2009). Presence/absence observations minimize variation by emphasizing the lower threshold (1 specimen per sample) in the observations. They give as output a measure of habitat suitability expressed as the frequency of positive observations (presence) given the environmental conditions. Habitat suitability models are often used to describe and compare relationships between macrozoobenthic species distribution and different environmental factors, including sediment texture (Ysebaert et al. 2002, Degraer et al. 2008, Merckx et al. 2011). A second approach to heteroskedasticity is to apply a regression analysis through the 'ceiling' of the distribution. Upper boundary regression emphasizes the upper limit (maximum attainable abundance), consistent with the ecological theory of the minimum (Cade et al. 1999, Cade & Noon 2003, Cade et al. 2005, Downes 2010). While all lower observations are constrained by subsidiary factors, the upper edge of the multivariate data cloud is mostly representative of the maximal potential response of the species to the measured variable. Examples of upper boundary regression applied to benthic species distribution modeling along the sediment granulometry gradient can be found in Thrush et al. (2003, 2005), Ellis et al. (2006), and Anderson (2008).

The methodologies mentioned above offer only an inaccurate (OLS) or partial (occurrence and maxima models) description of heteroskedastic distributions. While habitat suitability is more stable through time compared to abundances and is generally considered highly important for management strategies (Degraer et al. 2008), the magnitude of the effect organisms have on ecosystem processes is often influenced by the realized size and density rather than the presence or the potential abundance. The estimation of density-related effects of macrozoobenthos (e.g. the influence of macrozoobenthos on sediment dynamics or the availability of food for avifauna) is likely to require extra information about the

probability that a density is realized at given environmental conditions. The quantile regression model (Koenker & Basset 1978, Koenker & Machado 1999, Koenker & Hallock 2001, Koenker & Xiao 2006) can solve this problem. Quantile regression is emerging in ecology as a comprehensive approach to the statistical analysis of linear and nonlinear response models (Cade & Noon 2003, Austin 2007, Franklin 2009, Downes 2010). This method aims to fit any desired quantile of a response variable distribution to an independent variable. Regression quantile estimates can be used to construct predictions without assuming any parametric error distribution and without specifying how variance heterogeneity is linked to changes in means. The effect of the error distribution on parameters is indexed by a family of quantiles  $\tau(0 \leq \tau \leq 1)$ . By representing the regressions of different quantiles simultaneously, the quantile regression model expresses the expected variation in the form of a conditional probability density distribution. The multi-quantile regression models permit the estimation of the expected distribution of observations as (co-) determined by all other factors affecting occurrence and abundance (Cade et al. 1999, Cade & Noon 2003). Compared to presence/absence and maxima analysis, they offer not only a measure of habitat suitability (occurrence or potential maximum), but also a more complete view of biotic distributions along single gradients. From this point of view, a multi-quantile regression approach can be preferable to single rate models, not because it 'fits better', but because it provides better insights into observed distributions. The full method, including the simultaneous estimation of regression models for all quantiles, was suggested by Cade et al. (2005) to be the best solution when the interactions between measured and unmeasured environmental variables are unknown, and one has no *a priori* basis for the selection of one of the quantiles as the most representative.

In this article we propose the use of non-linear quantile regression to estimate the entire cumulative distribution of common macrozoobenthic species as a function of granulometry in temperate coastal basins. The primary aim of the analysis is to accurately describe how the biomass is distributed along a gradient of sediment composition. This description is fairly complete, as it yields a predicted probability distribution of the response variable for all values of the independent variable. To be properly validated, the predicted probability distribution should be compared with a complete population of values from an independent dataset. It was not possible to find,

within a single basin, 2 or more sufficiently large and independent data sets. We therefore approached the validation problem from a different angle, by comparing 2 neighboring coastal ecosystems: the Oosterschelde and the Westerschelde (Fig. 1). The sediment grain size composition per se is similar in both basins (see Fig. 2). Due to the different anthropogenic impacts, there is a different covariance structure between sediment composition, water salinity and local hydrodynamic forcing (Table 1). The basins share a common pool of macrozoobenthic species, but abundances and dominances often differ. The comparison of the Westerschelde and the Oosterschelde allows us to evaluate the consistency of the predicted responses despite different covariance structures with other variables. Apart from the statistical validation of the method, it also allows us to formulate hypotheses on causal factors (covariance structure within environmental variables) explaining the differences in the response of the macrozoobenthic community to a similarly distributed gradient in sediment composition.

## MATERIALS AND METHODS

### Study area: Oosterschelde-Westerschelde

The Schelde estuary, a macrotidal, turbid, nutrient-rich, coastal plain estuary, is situated near the border between The Netherlands and Belgium (Fig. 1). It was originally composed of 2 aligned and intercon-

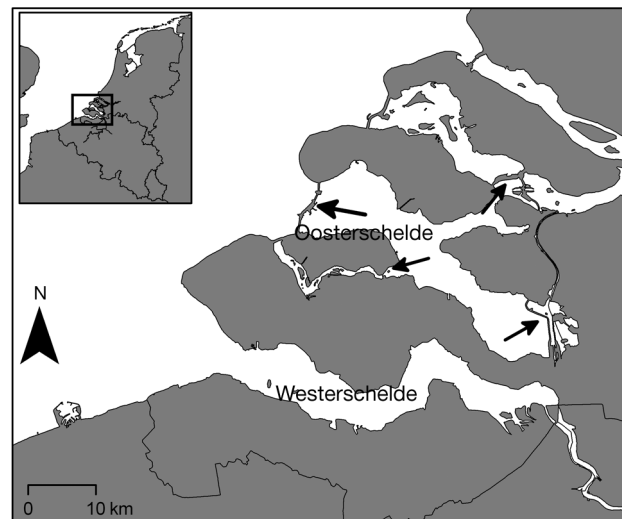


Fig 1. The study area. Note the back-barrier dams and the sea-side storm surge barrier in the Oosterschelde (black arrows)

Table 1. Some characteristics (yearly average) of the Westerschelde and the Oosterschelde. The Westerschelde values separated by a dash refer to the mouth (right) and to 80 km upstream (left)

Variable	Water surface (km <sup>2</sup> )	Depth (m)	Tidal range (m)	Tidal volume (m <sup>3</sup> × 10 <sup>9</sup> )	Salinity (PSU)	Freshwater load (m <sup>3</sup> s <sup>-1</sup> )	Suspended sediment (mg l <sup>-1</sup> )
Westerschelde	310	9.7–13.7	5.20–3.82	1	2.5–29	104	101
Oosterschelde	351	8	3.25	0.88	30	25	25

nected water bodies called Westerschelde and Oosterschelde. The present geomorphology and functioning of these basins is deeply affected by human interference (Louters et al. 1998).

Due to land reclamation, the Oosterschelde was progressively separated from the Westerschelde. The freshwater input from the Schelde river was definitively interrupted in 1903, transforming the estuary in an enclosed sea-arm with stable (marine) salinity (Table 1). The construction of 2 back-barrier dams in 1965 and 1969 had a significant impact on the Oosterschelde tidal hydrodynamics and sediment transport. The effects of these interventions were still ongoing when in 1986, the basin was partly closed off from the sea by a storm surge barrier (Delta Works) (Eelkema et al. 2012). After the Delta Works, the tidal prism of the Oosterschelde basin was reduced by approximately 30%. Current velocities have declined by 20–40% in the tidal channels and by over 40% around the tidal shoals and salt marshes (Louters et al. 1998). Simultaneously, the import of sediment from the coastal sea was cut off by the storm surge barrier. As a consequence, the channels tend to fill up using sediment eroded from the tidal flats. The availability of suspended mud for deposition on the flats has decreased considerably, with present suspended particulate matter concentrations only half those of the pre-barrier situation (Louters et al. 1998). The erosive trend of the Oosterschelde's tidal flats has increased to a warning level for coastal defense and habitat preservation, creating new challenges for the management of Oosterschelde (Bijker 2002).

The Westerschelde, due to its importance as a shipping route to the Antwerp port, kept its open connection to the sea. It is, however, subject to extensive dredging in order to enhance its function as a shipping route. The estuary is characterized by a fully developed salinity gradient and a complex network of channels surrounding several large intertidal flats and salt marshes. Current velocities are high, especially in the gullies (De Vriend et al. 2011). The suspended particulate matter concentrations are on average 4× higher in the Westerschelde compared to the Oosterschelde.

### Sampling and data processing

A total of 3926 macrozoobenthic samples, taken between 2000 and 2008, were used in the analysis. Samples were collected by Netherlands Institute of Sea Research (NIOZ)—Yerseke and Rijkwaterstaat. The use of such a large and multi-year dataset accounts for most of the spatiotemporal variability of environmental conditions and biotic responses. Data collection was carried out in spring and autumn (except for spring 2006, for which no data were available). While in the Westerschelde samples were randomly collected within 4 depth strata, in the Oosterschelde the sampling was repeated at the same sites across years in similar depth strata as used for the Westerschelde sampling. In both basins, 25% of the records come from intertidal sites, 50% of the records were collected above a depth of 5 m NAP (Normaal Amsterdams Peil, the Dutch height datum, 0 m NAP = mean sea level in Amsterdam) and 95% above 20 m NAP.

From each site, 3 cores of 8 cm diameter were taken and combined. Sediment was sieved through 1 mm mesh. Subtidal sites were sampled using a large (30 cm diameter cylinder) box-corer, from which the three 8 cm sub-cores were extracted. The collected animals were weighted and classified to the lowest practical taxonomical level. Abundances were expressed as number of individuals m<sup>-2</sup> and biomass as mg m<sup>-2</sup> of ash-free dry weight. Rare or occasional species (<6 observations) were removed from the dataset. After this skimming, the dataset comprised 194 taxa.

Sediment samples were analyzed for grain size distribution by laser diffraction performed with a Malvern Mastersizer 2000. For each sample, the median sediment grain size (d50), expressed in μm, and the percentage of different granulometric size-classes (mud < 63 μm < very fine sand < 125 μm < fine sand < 250 μm) were reported. In this analysis, d50 is predominantly used. Depth (m NAP) was measured on site.

Tidal-averaged bed shear stress (BSS, N m<sup>-2</sup>) and daily-averaged salinity (spring) were calculated

using the Delft3D FLOW model (version 3.55.05.00) with tidal forcing only (Lesser et al. 2004). The spatial resolution of the grid varies from more than 2000 m at the seaward boundary to around 100 m at the Oosterschelde inlet. For the Westerschelde, 25% of the records were collected at a salinity higher than 25, and 50% of the records were collected at a salinity between 25 and 18.

### Statistical analysis: non-linear quantile regression

Quantile regression was developed as an extension of the linear regression model (Koenker & Basset 1978). The  $\tau$ -th sample quantile of any random variable  $Y$ ,  $Q_Y(\tau)$ , is the value splitting the distribution in a portion  $Y \leq Q(\tau)$  containing a fraction  $\tau$  of the observations and a portion  $Y > Q(\tau)$  with a fraction  $(1 - \tau)$  of the observations. It can be calculated by minimizing the expectation of the loss function  $\rho_\tau(y_i - \xi)$ :

$$\operatorname{argmin}_{\xi \in \mathcal{R}} \sum_{i=1}^n (\rho_\tau(y_i - \xi)) = \operatorname{argmin}_{\xi \in \mathcal{R}} \left[ (\tau - 1) \sum_{y_i \leq \xi} (y_i - \xi) + \tau \sum_{y_i > \xi} (y_i - \xi) \right] \quad (1)$$

with respect to  $\xi(\tau)$  possible solutions, the smallest of which is  $Q_\tau(Y)$ . The simultaneous estimation of several quantiles permits to characterize the distribution of the variable  $Y$  by its quantiles  $Q_Y(\tau)$ . By extension, the linear conditional quantile  $Q_Y(\tau|X = x)$  linearly depending on an independent variable  $X$  can be estimated by replacing the term  $\xi$  in Eq. (1) with the term  $x_i\beta$ :

$$\hat{\beta}(\tau) = \operatorname{argmin}_{\beta \in \mathcal{R}^p} \sum \rho_\tau(y_i - x_i\beta) \quad (2)$$

where  $\hat{\beta}(\tau)$  is the unknown regression coefficient for the  $\tau$ -th quantile and  $\beta$  are the possible solutions with respect to  $x_i$ . With  $\tau = 1$ ,  $Q_Y(\tau|x_i)$  is the expected maximal response at  $x_i$ ; in case of biomass or abundance distributions, this value represents the expected response when (unmeasured) disturbance is at a minimum and (unmeasured) facilitation at a maximum. The succession of the underlying quantiles represents the expected distribution of responses given the actual distribution in the habitat of the unmeasured facilitating or inhibiting variables.

We extend the use of quantile regression to the non-linear case by the use of spline transformations (Koenker 1994, Thompson et al. 2010). The B-spline smoothing of the explanatory variable produces flexible curves without linking the response shape to a predetermined function. Regression splines produce a similar output to general additive models, but they

are computationally faster and better suited to the analysis of large spatio-temporal datasets (Elith & Leathwick 2007). Quantile regression can easily be coupled with a spline transformation of the independent variable. A B-spline function is a piecewise polynomial of a given order satisfying second-order continuity conditions between the pieces. Breakpoints are automatically selected by an iterative calculation (Barsky & Thomas 1981). Because of its property of continuity, the polynomial resulting from a B-spline expansion of the explanatory variable forms a smooth and flexible curve. B-spline quantile regression analysis is particularly suited to describing the heterogeneous responses of species abundances to environmental factors (Anderson 2008). The main variable controlling the flexibility of the spline shape is the degree of the polynomial, resulting in the number of parameters in the regression model. As suggested in previous studies (Cade et al. 2005, Anderson 2008), we selected spline degrees from 2 to 5 (see Tables S1 & S2 in Supplement 1 at [www.int-res.com/articles/suppl/m475p119\\_supp/](http://www.int-res.com/articles/suppl/m475p119_supp/) for a summary of the dataset). Measurements of goodness of fit like  $R^2$  and its equivalent for quantile regression (Koenker & Machado 1999) cannot decrease with increasing number of parameters, thus they are not useful for selecting the most appropriate spline degree. Conversely, the Akaike Information Criterion (Akaike 1981) offers a relative measurement of the goodness of fit based on the tradeoff between the accuracy and the complexity of the model; the lowest AIC value provides a summary measure of the best fit with the fewest number of parameters. Following Burnham & Anderson (2004), in this analysis we used a small sample size corrected version of the Akaike Information Criterion (AICc) developed by Hurvich & Tsai (1990).

Non-linear multi-quantile regression models are able not only to reproduce the asymmetry of the distributions along the  $x$ -axis, but also the asymmetry along the  $y$ -axis (i.e. the rate at which the abundances drop to zero beyond the upper boundary of the distribution). In this study, B-spline quantile regression was used to characterize relationships between the median sediment grain size (d50) ( $\mu\text{m}$ ) and the principal community descriptors total biomass ( $\text{mg m}^{-2}$ ), abundance ( $\text{m}^{-2}$ ), number of taxa ( $S$ ) and Shannon-Wiener diversity index

$$H' = - \sum_{i=1}^s p_i \log(p_i)$$

where  $p_i$  is the proportion of individuals belonging to the  $i$ th species (see Tables S3 & S4 in Supplement 1 for a summary of the dataset). The method was also

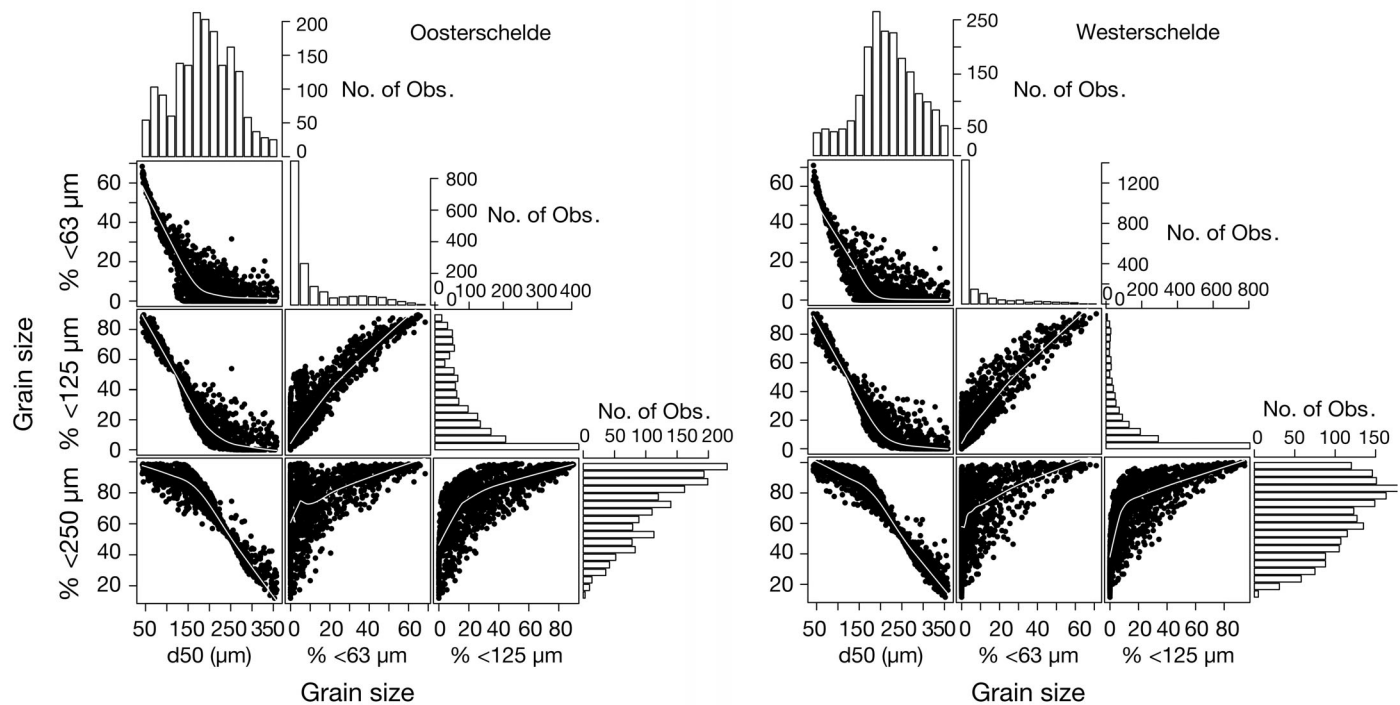


Fig. 2. Correlation of d50 (median sediment grain size) and different sediment fractions in the Oosterschelde and Westerschelde. Histograms show the number of observations. White lines show the correlations between  $x$  and  $y$  values as predicted by locally weighted regression

used to regress the biomass of the 25 most frequently observed (relative number of occupied samples) taxa on sediment d50 (see Tables S5 & S6 in Supplement 1). We analyzed the distribution of the most abundant taxa irrespective of the level of taxonomical accuracy at which they were identified. We made this choice to offer as complete and realistic an overview as possible of the overall benthic assemblages. Focusing the analysis only at species level would force us to reduce the numbers of groups accounted for in the analysis due to scarcity of data available for single species. Species optimal d50 were assessed as the average of the optimum predicted by each quantile of the biomass distributions predicting occurrence ( $x_i$  for which  $Q_Y(\tau|X = x_i) = \max(Q_Y(\tau|X))$  and  $\max(Q_Y(\tau|X = x_i)) > 0$ ).

Environmental variables were significantly affected by spatial autocorrelation (Moran's  $I$   $p$ -value  $< 10^{-10}$  in both basins). However, the spatial trend was not removed from the dataset, as Cade et al. (2005) showed that quantile regression models have high performance in explaining the observed variance also in the presence of spatial autocorrelation of environmental variables. Data scarcity at the extremes of the (long-tailed) d50 distributions can affect the behavior of the spline function, leading to overestimation of the expected densities. In order to avoid this effect, the

analysis was limited between the first (42  $\mu\text{m}$ ) and the 99th (358  $\mu\text{m}$ ) percentile of the Oosterschelde range of median grain sizes (see Tables S7 to S14 in Supplement 1 for a summary of the data excluded from analysis). To make the analysis still more resistant, the regression was 'shrunk' by computing the fit with a lasso algorithm (Tibshirani 1996) using as a threshold the parameter  $\lambda = 10$  for analysis of community indices and  $\lambda = 1$  for single taxa analysis. The full quantile range from  $\tau = 0.01$  to  $\tau = 0.99$  was fitted, with increments of 0.005.

All statistical analyses were implemented in R (R Development Core Team 2010) using the packages 'splines' and 'quantreg' (Koenker 2010). The R script of the analysis is available in Supplement 2 at [www.int-res.com/articles/supp/m475p119\\_supp/](http://www.int-res.com/articles/supp/m475p119_supp/).

## RESULTS

### Basic data analysis

#### Abiotic variables

The Oosterschelde and Westerschelde sampling sites showed similar sediment composition and distribution (Fig. 2). The ratios between the different

granulometric classes of sediment also appeared similar across basins. Measures of sediment composition related to the percentage of different granulometric size-classes (mud < 63  $\mu\text{m}$  < very fine sand < 125  $\mu\text{m}$  < fine sand < 250  $\mu\text{m}$ ) were truncated to 0 or 100% at

the extreme of the gradient, while sediment median grain size was more continuously distributed (Fig. 2). For this reason, the latter variable was preferred in this study. Some examples using different proxies for sediment composition can be found in Figs. S1 to S3 in Supplement 1.

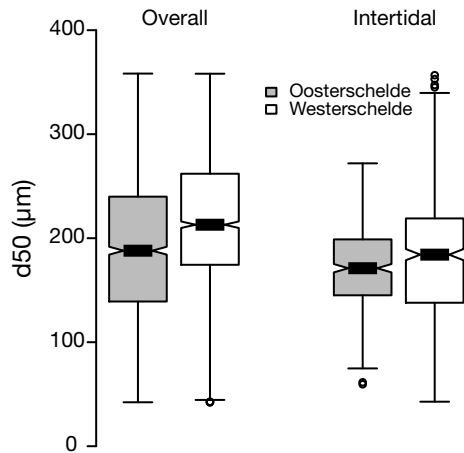


Fig. 3. Sediment d50 distribution for the overall dataset and for intertidal samples only. The bottom and the top of the boxes are the 25th and 75th quantiles of the distribution, the band near to the middle of the boxes is the 50th. The bottom and the top whiskers are the 2.5th and the 97.5th quantiles

On average, the sediment in the Oosterschelde was slightly but significantly ( $p < 10^{-10}$ ) finer than in the Westerschelde (Fig. 3). In the intertidal samples, the d50 distribution was wider in the Westerschelde than in the Oosterschelde. The relationships between sediment d50, bottom shear stress and depth differed across basins (Fig. 4). In the Westerschelde, fine sediments were rare at deeper sites and at higher hydrodynamic stress values. In the Oosterschelde, the relationships with depth and hydrodynamic stress were less significant than in the Westerschelde. In the case of bottom shear stress, the distribution of the explanatory variable was very different between the 2 systems: in the Oosterschelde, the maximal tide-averaged BSS was  $0.92 \text{ N m}^{-2}$ , while in the Westerschelde, BSS was  $> 3 \text{ N m}^{-2}$  (Lesser et al. 2004). Sediment composition in the Westerschelde was not correlated with the average salinity gradient ( $p = 0.1107$ ).

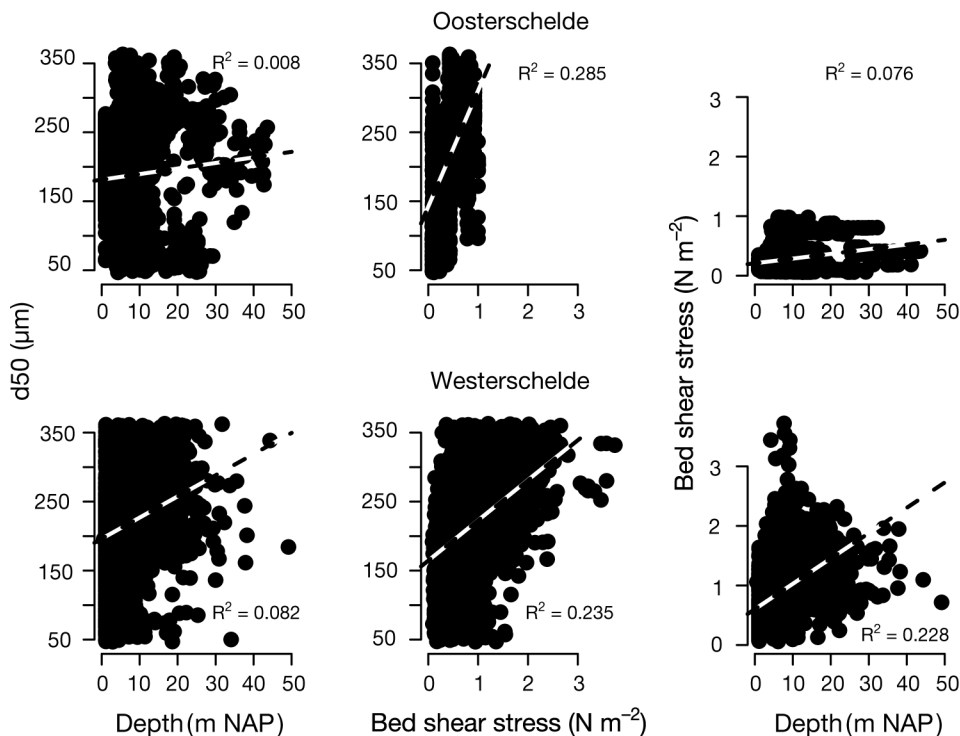


Fig. 4. Correlation of sediment d50 with measured depth and modelled bed shear stress. Dashed lines show the correlation between x and y values as predicted by linear regression. NAP = Normaal Amsterdams Peil, the Dutch height datum, 0 m NAP = mean sea level in Amsterdam

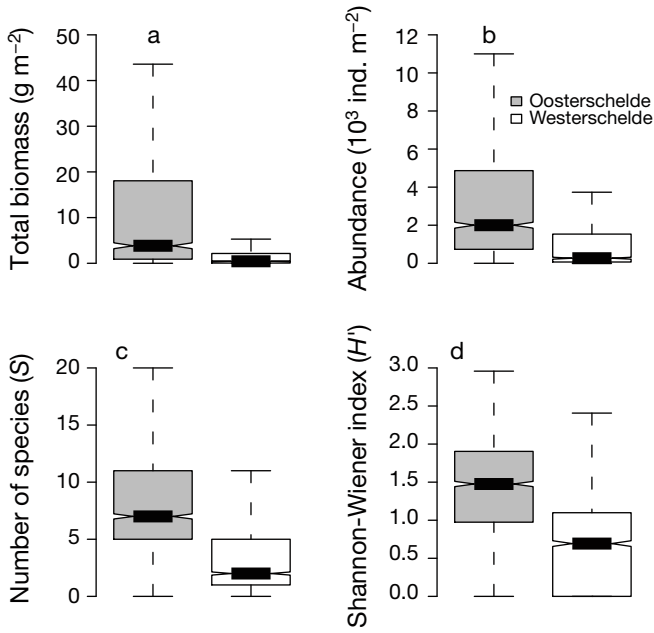


Fig. 5. Community indices: (a) total biomass, (b) abundance, (c) number of species, (d) Shannon-Wiener diversity index. The bottom and the top of the boxes are the 25th and the 75th quantiles of the distribution, the band near to the middle of the boxes is the 50th. The bottom and the top whiskers are the 2.5th and the 97.5th quantiles

Biotic variables

Total biomass, abundance, number of taxa per sample and Shannon-Wiener diversity were different between the basins (Fig. 5). Comparing the median values, the benthic communities of the Westerschelde reached only 11% of the biomass (Westerschelde 435 mg m<sup>-2</sup> versus Oosterschelde 3832 mg m<sup>-2</sup>), 13% of the abundance (266 versus 2000 individuals m<sup>-2</sup>), 28% of the number of taxa (2 versus 7 taxa per site), and 42% of the diversity ( $H' = 0.53$  versus 1.27) realized in the Oosterschelde. The total biomass and abundance distributions were positively skewed, indicating a prevalence of sites with low benthic abundance and biomass.

Based on the 25 most frequent (relative number of occupied samples) taxa, clear differences in taxonomic composition were found between the benthic assemblages in both basins (Fig. 6). Only few taxa, e.g. *Arenicola marina* and *Corophium arenarium*, reached similar biomass and frequency in the Oosterschelde and the Westerschelde, while the majority were abundant in one basin and scarcely present in the other. The Westerschelde was dominated by the

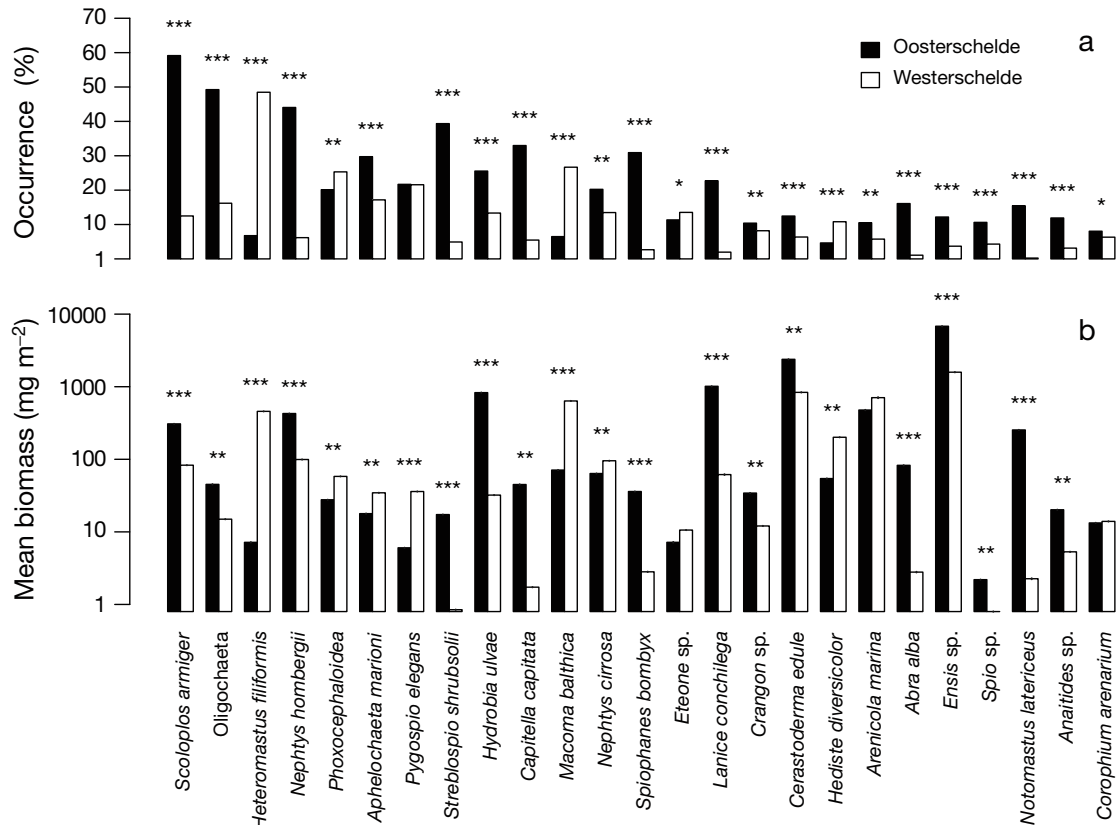


Fig. 6. Mean (a) biomass and (b) occurrence of the 25 most common taxa in the Westerschelde and the Oosterschelde. Asterisks show the significance of Fisher's test and ANOVA of differences between basins (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ )



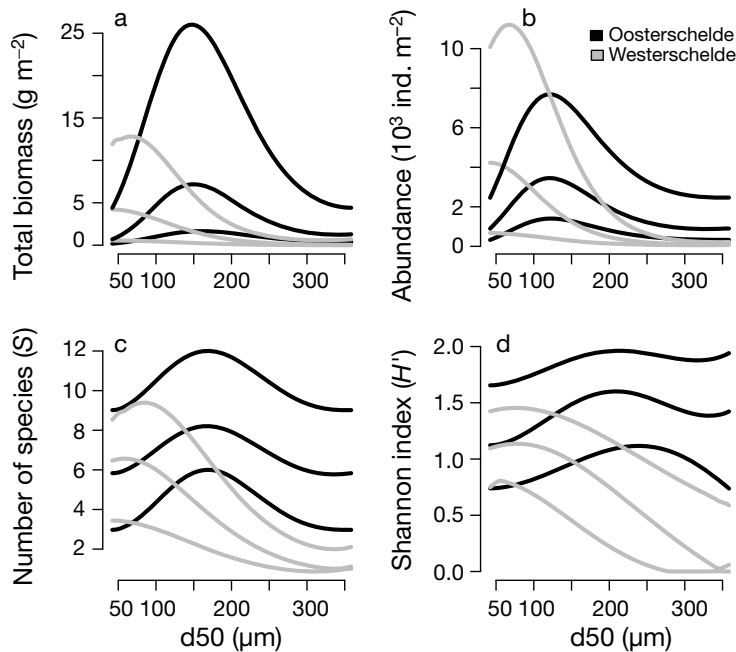


Fig. 7. Distribution of species richness, total biomass and ecological diversity along the sediment d50 gradient. The expected distributions are sketched using the 0.25, 0.5 and 0.75 quantiles

polychaetes *Heteromastus filiformis* (average [avg.] biomass 457 mg m<sup>-2</sup>, frequency 48%) and *Hediste diversicolor* (avg. biomass 201 mg m<sup>-2</sup>, frequency 11%), and the bivalve *Macoma balthica* (avg. biomass 636 mg m<sup>-2</sup>, frequency 27%); all deposit-feeding. In contrast to other tube-building worms, *Pygospio elegans* (avg. biomass 36 mg m<sup>-2</sup>, frequency 22%) attained 6× higher biomass in the Westerschelde than in the Oosterschelde (Fig. 6).

Most of the dominant taxa in the Westerschelde were almost absent in the Oosterschelde. The most common species in this basin were the scavenger-predator *Nephtys hombergii* (avg. biomass 427 mg m<sup>-2</sup>, frequency 44%) and the burrower-deposit feeder *Scoloplos armiger* (avg. biomass 308 mg m<sup>-2</sup>, frequency 59%). High biomass values were observed in the Oosterschelde for the small herbivorous gastropod *Hydrobia ulvae* (avg. biomass 829 mg m<sup>-2</sup>, frequency 26%). In the Oosterschelde, we also recorded high biomass and frequency of the tubedwelling worms *Lanice conchilega* (avg. biomass 1014 mg m<sup>-2</sup>, frequency 23%), *Spiophanes bombyx* (avg. biomass 36 mg m<sup>-2</sup>, frequency 31%) and *Notomastus latericeus* (avg. biomass 254 mg m<sup>-2</sup>, frequency 15%) and for the deposit-feeding worms *Streblospio shrubsolii* (avg. biomass 17 mg m<sup>-2</sup>, frequency 39%) and *Capitella capitata* (avg. biomass 45 mg m<sup>-2</sup>, frequency 33%).

## Non-linear quantile regression

### Community indices

Total biomass, total abundance, number of taxa and the Shannon-Wiener diversity index showed relevant variations along the sediment d50 gradient (Fig. 7). For each point along the sediment d50 gradient, the total biomass and abundance were asymmetrically distributed around the median value (i.e. the fitted median curve was closer to the 25th than to the 75th quantile curve). As effect of data heteroskedasticity, the distance between quantiles was not constant along the gradient, but decreased at less suitable conditions. The number of taxa and the Shannon-Wiener diversity distributions were homoskedastic in both basins.

In the Westerschelde, the maximum expected biomass and diversity peaked at a d50 of ~55 μm. High abundance, richness and diversity were only observed in samples with small median grain size. In the Oosterschelde, the community indices had a unimodal shape, reaching their highest values at an average d50 of ~170 μm. In contrast to the Westerschelde, benthic assemblages in very coarse sediments (d50 = 350 μm) still conserved 56% of the maximal diversity and 26% of the maximal biomass. At the lower range of d50, the Oosterschelde presented similar diversity, but lower biomass and abundance, than the Westerschelde. For a d50 up to 100 μm, the probability of measuring an elevated abundance in the Oosterschelde was consistently lower than in the Westerschelde, while the differences between the lower quantiles (observations around the 25th quantile) were less pronounced.

### Individual taxa

Predictions from nonlinear quantile regression of the biomass distribution of the 25 most common taxa in the Westerschelde and Oosterschelde basins are shown in Fig. 8. Forecasted biomass distributions tended to be unimodally distributed along the gradient. Only *Abra alba* and *Eteone* sp. (Oosterschelde) seemed to be unaffected by sediment grain size. Optima in the distributions of the different quantiles generally coincided well at approximately the same d50. The order of species preference along the sedi-

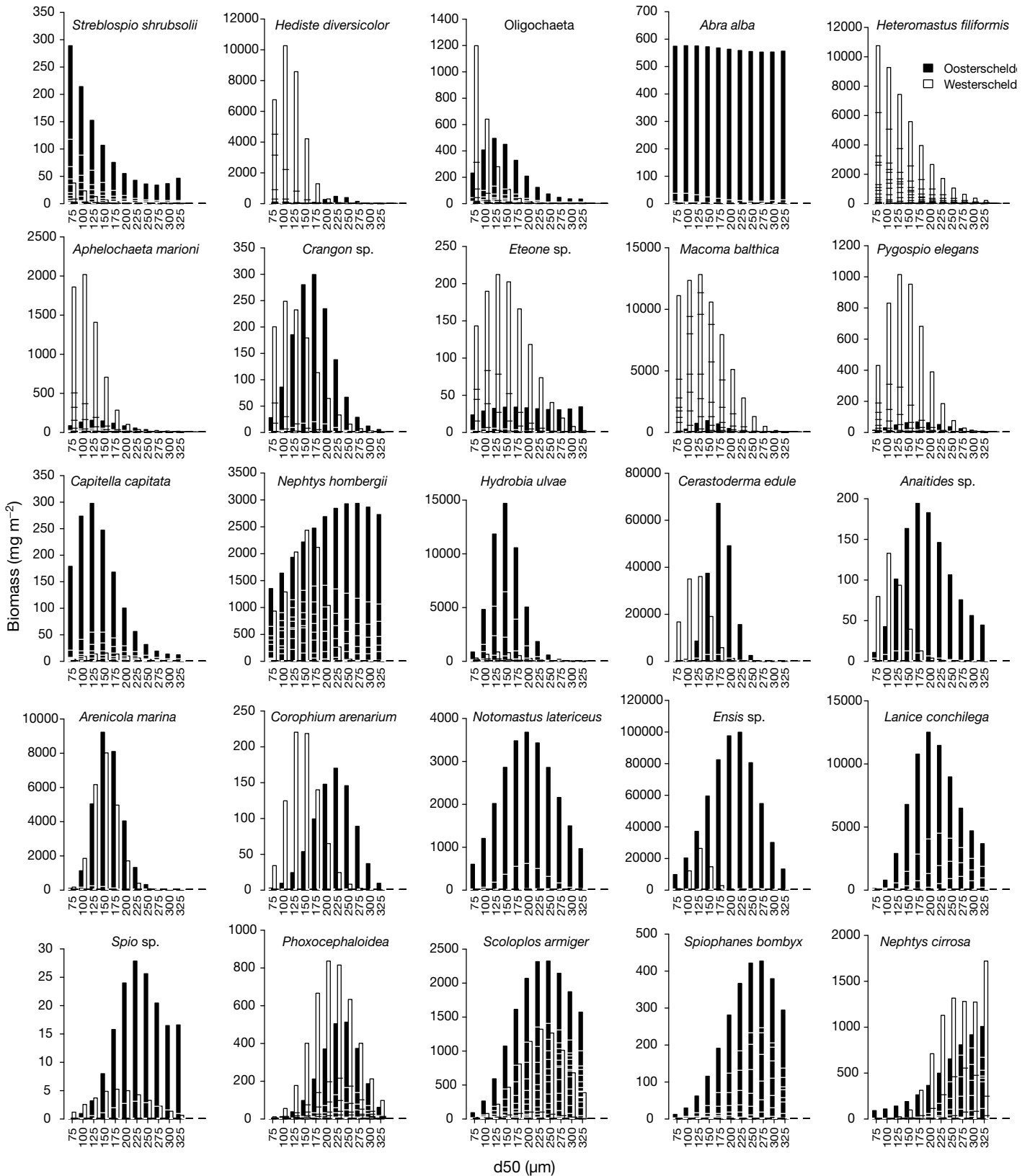


Fig. 8. Predicted d50-conditioned cumulative probability functions for the biomass distribution of the 25 most common taxa in the Oosterschelde (black) and Westerschelde (white) basins. Bars show the maximal biomass predicted for the sediment compositions indicated on the x-axis. Tick marks along the bars are drawn at intervals of 0.05 (e.g. 5% of the cases are expected to have a biomass higher than the value of the 1st tick mark from the top, 10% of the cases are expected to have a biomass higher than the 2nd tick mark from the top, and so on)

ment d50 gradient was similar in both systems (Fig. 9).

The estimated optimal values of d50 covered the full range of observed d50 values in the samples, between a minimum of 44  $\mu\text{m}$  (*Streblospio shrubsolii*) and a maximum of 319  $\mu\text{m}$  (*Nephtys cirrosa*). These extremophile species did not change preference for sediment granulometry across basins. In the Westerschelde, d50 optima were concentrated at fine granulometry. Eighty percent of the taxa had their peak abundance in the first quarter of the d50 gradient. In the Oosterschelde, the taxa were spread out much more widely over the gradient. With the exception of 4 taxa (*S. shrubsolii*, *Arenicola marina*, *Capitella capitata*, *N. cirrosa*), all taxa that occurred in both basins realized their peak at smaller grain size in the Westerschelde than in the Oosterschelde (Fig. 9). The order of species preference was roughly preserved across basins, with the exception of the broad taxonomic group of Oligochaeta, and 3 taxa that almost exclusively occurred in one of the basins and were consequently poorly estimated in the other: *Hediste diversicolor*, *Ensis* sp. and *Lanice conchilega*.

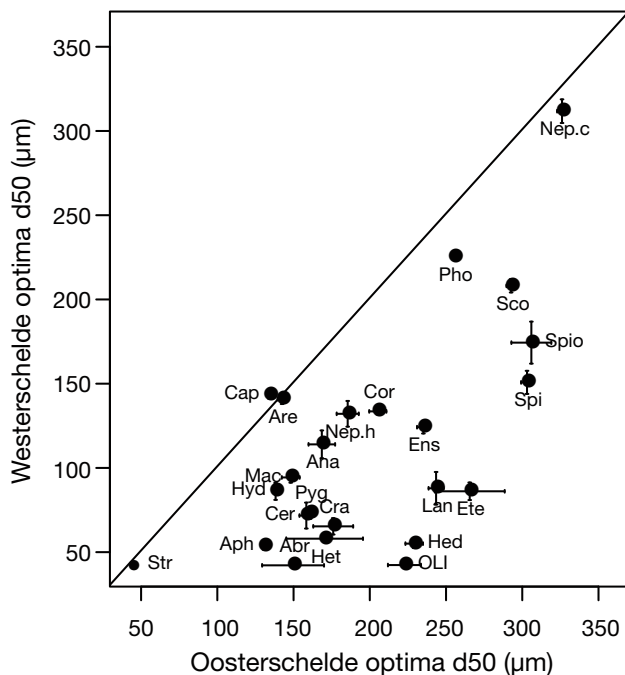


Fig. 9. Relationship between the sediment d50 optima detected separately for the 2 basins. Optima were calculated as average of the optimal d50 values detected for each quantile of the forecasted distributions (from  $\tau = 0.01$  to  $\tau = 0.99$ ). The line shows the 1:1 ratio. Error bars show the SE of optima distribution around the mean value. Abbreviations refer to the first 3 letter of the generic name of the species.

See Fig. 6 for full species names

## DISCUSSION

### Considerations about the statistical method

In this paper, we used quantile regression to study the effect of a single environmental variable (sediment grain size) in 2 systems where the covariance structure with other environmental variables was very different. There was a significant spatial gradient in salinity in the Westerschelde that was almost completely absent in the Oosterschelde. Furthermore, the correlation structure between depth, bottom shear stress and sediment grain size composition was very different between the 2 systems (Fig. 4). This comparison shows that the models can only partly be considered to be validated by the systems comparison. For several of the species, the maximum biomass modelled ( $\tau = 0.99$ ) differed largely between the 2 systems. Although the first interpretation of the maximum is that it shows the value to be reached at this grain size when all other conditions are optimal, it is clear that for many species, the conditions are never near optimal in at least 1 of the 2 systems. Thus, the maxima estimated for a particular grain size, when only using data from a single estuary, are not unconstrained. As argued before by Thrush et al. (2005), there is a strong scale dependency in species-environment relations in estuaries. In order to estimate real optimal biomass values for a particular grain size, a large number of ecosystems will have to be sampled, thus ensuring that optimal conditions for all other variables are realized. Factors may vary considerably between ecosystems, but hardly within each of the systems. It is impossible to estimate the influence of such factors when restricting the analysis (or the validation) to a single system. With the exception of a few species for which the comparison is difficult, the fitted models do yield consistent estimates of the optimal grain size of a species over the 2 systems. Often the estimate for the Westerschelde was for a somewhat lower optimal grain size, but the general ordering of the species remained the same (Fig. 9). Therefore, the relative predictions of the model (where will the species occur if it does occur at all) are much better validated by the systems comparison than the absolute predictions.

Our examples showed that the response curves for the different quantiles were similar and that the range of optima of the different quantile regression models was limited (Fig. 7). This can be ascribed to correlations (Fig. 4) and interactions between measured and unmeasured environmental variables that increase the non-linearity and decrease the hetero-

geneity of the responses (Cade et al. 2005). The similarity between quantiles is indicative of a strong correlation between the percentage of occurrence, the average and maximal biomass. Occurrence, central and maximal estimators all performed approximately equally well in describing species responses. When comparing the response models between species, we see that the majority of the single species models scored the lowest AICc when fitted on a quadratic or cubic spline (see Tables S1 & S2 in Supplement 1). Due to this relatively low degree, most fitted models had a general appearance resembling a Gaussian response model. The relative advantage of having a flexible non-parametric form for the response curve was limited, as was the relative advantage of having all percentiles fitted. However, it is reassuring that the shape and the similarity between quantiles followed from the fitting procedure itself and were not imposed *a priori*. Thus, it may be possible to simplify the fitted models without much loss of information; the presently fitted, more complicated models could form an excellent basis for the choice of a simpler model with fewer parameters.

### Community responses across basins

Benthic community distributions were very different across basins, with higher biomass, density, richness and diversity characterizing the Oosterschelde. This was partly caused by higher peaks (for biomass, species richness and diversity), but mainly by a more skewed distribution along the d50 gradient in the Westerschelde (Fig. 7). In the latter basin, high community index values were realized only at muddy sites. All measures of biological richness dropped suddenly in coarser sediments (100–250  $\mu\text{m}$ ). This pattern can be interpreted as the result of stronger limitations at the right side of the d50 gradient. Tidal variability in salinity at a particular place is responsible for a considerable part of the brackish-water depression in diversity (Attrill 2002). The high and stable salinity in the Oosterschelde might partially explain the higher species richness and diversity compared to the Westerschelde (Ysebaert et al. 2003). Higher salinity in the Oosterschelde may also explain the absence, in that system, of species that have clear preference for brackish conditions, such as *Hediste diversicolor*. However, sediment composition in the Westerschelde was homogeneously distributed along the salinity gradient. Salinity variations by themselves are therefore not a sufficient explanation for the absence of sandy substrates with

high biological richness (either in abundance, biomass or species number). In the Oosterschelde, the probability of realizing high community performances was maximal in fine sand substrates (Fig. 7). Hydrodynamic stress in these sediments was lower compared to the Westerschelde (Fig. 4), most likely explaining the higher values observed.

The responses of total biomass and total abundance (Fig. 7) differed between the 2 systems. While predicted biomass in Westerschelde was lower than in Oosterschelde over almost the entire sediment gradient, predicted abundance tended to be higher for Westerschelde in fine sediments. Especially in these muddy sediments, the fauna was mainly composed of small organisms (e.g. *Heteromastus filiformis*: average per capita biomass 0.9 mg, *Streblospio shrubsolii*: 0.1 mg) that become superabundant in fine, organic rich sediments (Pearson & Rosenberg 1978, Nilsson & Rosenberg 2000, Norkko et al. 2006). In the Oosterschelde, the maximal biomass and abundance were realized at finer granulometry (respectively at d50 of 147 and 126  $\mu\text{m}$ ) than the maximal species number and diversity (d50 of 187 and 252  $\mu\text{m}$ ). This behavior is in agreement with the Intermediate Disturbance Hypothesis: the highest diversity is reached at intermediate levels of disturbance, while the community is dominated by a few opportunistic species at high disturbance, and by a limited number of powerful competitors at low disturbance (Dial & Roughgarden 1998). The differences in individual body size and life cycle suggest that it would be interesting to repeat this analysis with a trait-based approach (e.g. Statzner et al. 2004).

### Individual taxa responses across basins

The 25 taxa analysed in detail generally showed well-defined and differentiated responses to sediment grain size, with very different height but similar position of the optimum. The different correlation structure between granulometry and other environmental factors led to a replacement of mud-preferring species by sand-preferring species when going from the Westerschelde to the Oosterschelde, rather than to a drastic shift in the sediment type occupied by a particular species. Therefore, although sediment granulometry does function as a proxy for other variables (hydrodynamic conditions in particular), it appears that the influence of sediment per se is also important, and for some species, such as *Arenicola marina*, even predominant. While the taxa largely conserve their qualitative response to sediment tex-

ture, the quantitative response is strongly dependent on subsidiary factors. The higher abundance of species tolerant of low salinity (e.g. *Heteromastus filiformis*, *Hediste diversicolor*, *Macoma balthica*) in the Westerschelde can be partially explained by the presence of a gradient from brackish to saltwater (Ysebaert et al. 2002). The same argument, however, cannot be used to explain the absence of sand-preferring species, since sandy sites are available at both low and high salinity in the Westerschelde. These patterns are more likely related to the high hydrodynamic stress typically experienced by coarser sediments in this basin. Sandy sites are not available at low current velocities, as shown in Fig. 4. It is likely that sand-preferring species cannot establish in their preferred environment due to this enhanced stress. We conclude that benthic organisms adapt their optimum with respect to granulometry only over a limited range and adapt to limitation by subsidiary factors by a drastic decrease in abundance.

The representation of the different quantiles instead of a single parameter of distributions allows clearer insights into the observed patterns. In this respect it is interesting to compare the form of the expected distributions between species, or within the same species between basins. Observing the distributions of *Heteromastus filiformis*, *Macoma balthica* and *Hediste diversicolor* in the Westerschelde (Fig. 8), it is clear that the latter species was much less abundant than the previous 2, although the higher 10% of the observations fell in a similar range. *Arenicola marina*, *Cerastoderma edule* and *Ensis* sp., i.e. taxa with high per capita biomass, had very skewed quantile distributions because rarely more than a few individuals were found at the scale of the sampling unit. Within species but between basins, it is instructive to compare the distributions of *Nephtys hombergii* and *Nephtys cirrosa*. The upper boundary was similar or even higher in the Westerschelde than in the Oosterschelde, but the quantiles falling below the maximum were much higher in the Oosterschelde. This means that a few sites had high biomass, sometimes even higher than the maximum in the Oosterschelde, but apart from these, there were hardly any sites with sub-optimal conditions and sub-maximum biomass (Fig. 8).

The systematic shift of optima towards larger d<sub>50</sub> in the Oosterschelde with respect to the Westerschelde (Fig. 9) is likely related to the milder habitat conditions associated with coarse sediment (Fig. 4). Thus, it can be considered an emergent property of the local covariance structure of environmental variables rather than an intrinsic propriety of the sedi-

ment composition gradient. Biotic components are also involved in determining the local combination of relevant environmental variables, and they can co-explain part of the observed variation between basins. As an example, *Heteromastus filiformis* is a subsurface deposit feeding polychaete well adapted to muddy and organic-rich (oxygen-poor) sediment thanks to high oxygen-affinity haemoglobin. This species exhibits exceptionally high densities in the muddiest bottoms of some estuaries, while it is generally scarce in others (Fig. 6) (Thrush et al. 2005). The abundance of *H. filiformis* is mostly controlled by predation from *Nephtys hombergii* (Beukema et al. 2000). High biomass of the latter species was exceptional in the Westerschelde, and much rarer than in the Oosterschelde (Fig. 8). The combination of physical (low salinity in part of the estuary, hydrodynamic stress also relevant to muddy sites) and biological conditions (intense predation at only a few sites) in the Westerschelde led to a striking dominance of *H. filiformis*.

### Management considerations

To our knowledge, the recent coastal engineering works in the Oosterschelde are the main cause of the decoupling of the association between stronger hydrodynamics and coarse sediment (Louters et al. 1998), while the deepening of the Westerschelde channels probably had the opposite effect. Our analysis suggests that the benthic community composition and structure were affected by these human interventions. Anthropogenic alterations of the Oosterschelde, although leading to management problems in conservation of the eroding intertidal flats, have also probably led to improved species richness and biomass of the benthos. Anthropogenic alteration of the Westerschelde may have restricted benthic life in sandy sediments due to enhanced hydrodynamic stress (De Vriend et al. 2011). Further analyses, including historical data, should confirm these observations and help lead to proper management strategies for both basins.

### CONCLUSIONS

Non-linear quantile regression is a versatile and precise method for summarizing the response of species to an environmental gradient, as shown by our analysis of the responses of benthic macrofauna to sediment granulometry. The consistency of the

description of species responses was validated by the comparison of 2 adjacent but dynamically different ecosystems. We showed that species responses are somewhat variable between systems, but that in general, species have a clear preference for a particular type of environment. In our study, the 2 ecosystems differed in that sandy habitats were richest in density and biomass in one system, while muddy habitats were richest in the other system. This was due to a different covariance structure with other variables; in particular, hydrodynamic stress due to anthropogenic changes. We showed that this primarily leads to a shift in species between systems. We propose that a multivariate version of quantile regression might be able to better resolve these patterns.

*Acknowledgements.* We thank the Monitor taskforce (NIOZ-Yerseke) and Rijkswaterstaat for sample collection and processing, W. Sistermans and V. Escavage for the help in managing the dataset, M. Eelkema for the DELFT 3D modeling and 3 anonymous reviewers for their constructive criticism. This work was funded by the Ecoshape/Building with Nature project. At the time of starting this project, NIOZ-Yerseke belonged to the Netherlands Institute of Ecology.

#### LITERATURE CITED

- Akaike H (1981) A new look at the statistical model identification. *Cc/Eng Tech Appl Sci, Citation Classic*(51):22.
- Allen JR (1985) Field measurement of longshore sediment transport sandy hook, New Jersey, USA. *J Coast Res* 1: 231–240
- Anderson MJ (2008) Animal-sediment relationships revisited: Characterising species' distributions along an environmental gradient using canonical analysis and quantile regression splines. *J Exp Mar Biol Ecol* 366:16–27
- Attrill MJ (2002) A testable linear model for diversity trends in estuaries. *J Anim Ecol* 71:262–269
- Austin MP (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecol Modell* 200:1–19
- Barsky BA, Thomas SW (1981) Transpline—a system for representing curves using transformations among 4 spline formulations. *Comput J* 24:271–277
- Beukema JJ, Essink K, Dekker R (2000) Long-term observations on the dynamics of three species of polychaetes living on tidal flats of the Wadden Sea: the role of weather and predator-prey interactions. *J Anim Ecol* 69:31–44
- Bijker WE (2002) The Oosterschelde storm surge barrier—a test case for Dutch water technology, management, and politics. *Technol Cult* 43:569–584
- Burnham KP, Anderson DR (2004) Multimodel inference: Understanding AIC and BIC in model selection. *Sociol Methods Res* 33:261–304
- Butman CA (1987) Larval settlement of soft-sediment invertebrates—the spatial scale of pattern explained by active habitat selection and the emerging role of hydrodynamical processes. *Oceanogr Mar Biol Annu Rev* 25:113–165
- Cade BS, Noon BR (2003) A gentle introduction to quantile regression for ecologists. *Front Ecol Environ* 1:412–420
- Cade BS, Terrell JW, Schroeder RL (1999) Estimating effects of limiting factors with regression quantiles. *Ecology* 80: 311–323
- Cade BS, Noon BR, Flather CH (2005) Quantile regression reveals hidden bias and uncertainty in habitat models. *Ecology* 86:786–800
- Chapman MG, Tolhurst TJ, Murphy RJ, Underwood AJ (2010) Complex and inconsistent patterns of variation in benthos, micro-algae and sediment over multiple spatial scales. *Mar Ecol Prog Ser* 398:33–47
- Curry KJ, Bennett RH, Mayer LM, Curry A, Abril M, Biesiot PM, Hulbert MH (2007) Direct visualization of clay microfabric signatures driving organic matter preservation in fine-grained sediment. *Geochim Cosmochim Acta* 71:1709–1720
- Day JW, Hall CAS, Kemp WM, Yanez-Arancibia A (1989) *Estuarine ecology*. Wiley, New York, NY
- De Vriend HJ, Wang ZB, Ysebaert T, Herman PMJ, Ding P (2011) Eco-morphological problems in the Yangtze Estuary and the Western Scheldt. *Wetlands* 31:1033–1042
- Degraer S, Verfaillie E, Willems W, Adriaens E, Vincx M, Van Lancker V (2008) Habitat suitability modelling as a mapping tool for macrobenthic communities: an example from the Belgian part of the North Sea. *Cont Shelf Res* 28:369–379
- Dial R, Roughgarden J (1998) Theory of marine communities: the intermediate disturbance hypothesis. *Ecology* 79:1412–1424
- Downes BJ (2010) Back to the future: Little-used tools and principles of scientific inference can help disentangle effects of multiple stressors on freshwater ecosystems. *Freshw Biol* 55:60–79
- Eelkema M, Wang ZB, Stive MJF (2012) Impact of back-barrier dams on the development of the ebb-tidal delta of the Eastern Scheldt. *J Coast Res* 28:1591–1605
- Elith J, Leathwick J (2007) Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Divers Distrib* 13:265–275
- Ellis J, Ysebaert T, Hume T, Norkko A and others (2006) Predicting macrofaunal species distributions in estuarine gradients using logistic regression and classification systems. *Mar Ecol Prog Ser* 316:69–83
- Franklin J (2009). *Mapping species distributions—spatial inference and prediction*. Ecology, biodiversity and conservation. Cambridge University Press, Cambridge
- Gray JS (1974) Animal-sediment relationships. *Oceanogr Mar Biol Annu Rev* 12:223–261
- Heip CHR, Goosen NK, Herman PMJ, Kromkamp J, Middelburg JJ, Soetaert K (1995) Production and consumption of biological particles in temperate tidal estuaries. *Oceanogr Mar Biol Annu Rev* 33:1–149
- Herman PMJ, Middelburg JJ, van de Koppel J, Heip CHR (1999) Ecology of estuarine macrobenthos. *Adv Ecol Res* 29:195–240
- Herman PMJ, Middelburg JJ, Heip CHR (2001) Benthic community structure and sediment processes on an intertidal flat: results from the ecoflat project. *Cont Shelf Res* 21:2055–2071
- Hurvich CM, Tsai CL (1990) The impact of model selection on inference in linear regression. *Am Stat* 44:214–217
- Koenker R (1994) Quantile smoothing splines. *Biometrika* 81:673–680

- Koenker R (2010) quantreg: quantile regression. R package version 4.50. Available at <http://CRAN.R-project.org/package=quantreg>
- Koenker R, Basset G (1978) Regression quantiles. *Econometrica* 46:33–50
- Koenker R, Hallock K (2001). Quantile regression. *J Econ Perspect* 15(4):143–156. Annual Meeting of the Allied-Social-Science-Association, New Orleans, LA, Jan 2001
- Koenker R, Machado JAF (1999) Goodness of fit and related inference processes for quantile regression. *J Am Stat Assoc* 94:1296–1310
- Koenker R, Xiao Z (2006) Quantile autoregression. *J Am Stat Assoc* 101:980–990
- Kraufvelin P, Sinisalo B, Leppakoski E, Mattila J, Bonsdorff E (2001) Changes in zoobenthic community structure after pollution abatement from fish farms in the Archipelago Sea (N. Baltic Sea). *Mar Environ Res* 51:229–245
- Kraufvelin P, Perus J, Bonsdorff E (2011) Scale-dependent distribution of soft-bottom infauna and possible structuring forces in low diversity systems. *Mar Ecol Prog Ser* 426:13–28
- Lesser GR, Roelvink JA, van Kester JATM, Stelling GS (2004) Development and validation of a three-dimensional morphological model. *Coast Eng* 51:883–915
- Louters T, van den Berg JH, Mulder JPM (1998) Geomorphological changes of the Oosterschelde tidal system during and after the implementation of the delta project. *J Coast Res* 14:1134–1151
- Mayer LM (1994) Surface-area control of organic carbon accumulation in continental shelf sediments. *Geochim Cosmochim Acta* 58:1271–1284
- Merckx B, Steyaert M, Vanreusel A, Vincx M, Vanaverbeke J (2011) Null models reveal preferential sampling, spatial autocorrelation and overfitting in habitat suitability modeling. *Ecol Modell* 222:588–597
- Nilsson HC, Rosenberg R (2000) Succession in marine benthic habitats and fauna in response to oxygen deficiency: analysed by sediment profile-imaging and by grab samples. *Mar Ecol Prog Ser* 197:139–149
- Norkko A, Rosenberg R, Thrush SF, Whitlatch RB (2006) Scale- and intensity-dependent disturbance determines the magnitude of opportunistic response. *J Exp Mar Biol Ecol* 330:195–207
- Pearson TH, Rosenberg R (1978) Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanogr Mar Biol Annu Rev* 16:229–311
- R Development Core Team (2010). R: a language and environment for statistical computing. R foundation for Statistical Computing, Vienna
- Snelgrove PVR, Butman CA (1994) Animal–sediment relationships revisited: cause versus effect. *Oceanogr Mar Biol* 32:111–177
- Statzner B, Doledec S, Hugueny B (2004) Biological trait composition of European stream invertebrate communities: assessing the effects of various trait filter types. *Ecography* 27:470–488
- Sutherland TF, Amos CL, Grant J (1998) The effect of buoyant biofilms on the erodibility of sublittoral sediments of a temperate microtidal estuary. *Limnol Oceanogr* 43:225–235
- Terrell JW, Cade BS, Carpenter J, Thompson JM (1996) Modeling stream fish habitat limitations from wedge-shaped patterns of variation in standing stock. *Trans Am Fish Soc* 125:104–117
- Thompson P, Cai Y, Moyeed R, Reeve D, Stander J (2010) Bayesian nonparametric quantile regression using splines. *Comput Stat Data Anal* 54:1138–1150
- Thomson JD, Weiblen G, Thomson BA, Alfaro S, Legendre P (1996) Untangling multiple factors in spatial distributions: lilies, gophers, and rocks. *Ecology* 77:1698–1715
- Thrush SF, Hewitt JE, Norkko A, Nicholls PE, Funnell GA, Ellis J (2003) Habitat change in estuaries: Predicting broad-scale responses of intertidal macrofauna to sediment mud content. *Mar Ecol Prog Ser* 263:101–112
- Thrush SF, Hewitt JE, Cummings V (2004) Muddy waters: elevating sediment input to coastal and estuarine habitats. *Front Ecol Environ* 2:299–306
- Thrush SF, Hewitt JE, Herman PMJ (2005) Multi-scale analysis of species–environment relationships. *Mar Ecol Prog Ser* 302:13–26
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *J Roy Stat Soc B Met* 58:267–288
- van der Wal D, Herman PMJ, Forster RM, Ysebaert T and others (2008) Distribution and dynamics of intertidal macrobenthos predicted from remote sensing: response to microphytobenthos and environment. *Mar Ecol Prog Ser* 367:57–72
- van Ledden M, van Kesteren WMG, Winterwerp JC (2004) A conceptual framework for the erosion behaviour of sand–mud mixtures. *Cont Shelf Res* 24:1–11
- van Prooijen B, Winterwerp JC (2010) A stochastic formulation for erosion of cohesive sediments. *J Geophys Res C* 115:C01005
- White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48:817–838
- Wolff WJ (1983) Estuarine benthos. In: Ketchum B (ed) *Ecosystems of the world: estuaries and enclosed seas*. Elsevier, New York, NY, p 121–132
- Ysebaert T, Meire P, Herman PMJ, Verbeek H (2002) Macrobenthic species response surfaces along estuarine gradients: prediction by logistic regression. *Mar Ecol Prog Ser* 225:79–95
- Ysebaert T, Herman PMJ, Meire P, Craeymeersch J, Verbeek H, Heip CHR (2003) Large-scale spatial patterns in estuaries: estuarine macrobenthic communities in the Schelde estuary, NW Europe. *Estuar Coast Shelf Sci* 57:335–355

*Editorial responsibility: Hans Heinrich Janssen, Oldendorf/Luhe, Germany*

*Submitted: February 29, 2012; Accepted: October 11, 2012  
Proofs received from author(s): January 22, 2013*