Continental Shelf Research 63 (2013) 69–84

Contents lists available at SciVerse ScienceDirect

Continental Shelf Research

journal homepage: www.elsevier.com/locate/csr

Research papers

Simulating sediment–water exchange of nutrients and oxygen: A comparative assessment of models against mesocosm observations

Robin F. Wilson, Katja Fennel*, J. Paul Mattern

Department of Oceanography, Dalhousie University, Halifax, NS, Canada B3H 4J1

ARTICLE INFO

Article history:
Received 30 October 2012
Received in revised form
30 April 2013
Accepted 4 May 2013
Available online 14 May 2013

Keywords:
Diagenetic model
Sediment remineralization
Parameter optimization

ABSTRACT

How to represent nutrient fluxes resulting from organic matter remineralization in sediments should be an important consideration when formulating a biogeochemical ocean model. Here representations ranging from simple parameterizations to vertically resolved diagenetic models are compared against a comprehensive, multi-year data set from a mesocosm eutrophication study. Observations of sediment–water fluxes of nutrients and oxygen and measurements of the state of the overlying water column were made over 2.5 years in nine mesocosms, six of which received geometrically increasing loads of inorganic nutrients. These observations are used here to force and optimize two simple parameterizations of sediment oxygen uptake, one representative two-layer diagenetic model and one representative multi-layer diagenetic model. In cross-validation experiments the predictive ability of these different representations is compared. The main results are that the optimized multi-layer model fits the observations best and also proved to be the most parsimonious, while the two-layer model failed the cross-validation indicating that it is prone to over-fitting and was less parsimonious even than one of the simpler functional oxygen flux models. We recommend that sediment models that are candidates for inclusion in a biogeochemical model be assessed through a process of optimization and cross-validation as we have done here.

1. Introduction

Remineralization of organic matter in marine sediments is a key component of the global cycles of biogeochemically reactive elements (Berner, 1980; Burdige, 2006). In estuarine and continental shelf systems the return of inorganic nutrients to the water column that results from sediment remineralization is important in maintaining primary production and the coincident uptake of dissolved oxygen by sediments can contribute significantly to the development of hypoxic conditions (Kemp et al., 1992; Peña et al., 2010; Fennel et al., 2013). Sediments are important sites of denitrification (including sediments that underlie well-oxygenated bottom waters), which affects the supply of bioavailable nitrogen locally on timescales of months to years (Seitzinger and Giblin 1996) and the global inventory of bioavailable nitrogen on geological timescales of 1000 s to millions of years (Fennel et al., 2005). Quantification of sediment remineralization fluxes and their accurate representation in biogeochemical models is thus of considerable interest.

Direct observation of sediment remineralization fluxes is difficult. Fluxes across the sediment–water interface can be measured by chamber incubations at the sea floor, by column incubations in the lab or with the help of eddy flux correlation techniques. Profiles of solid and pore water constituents can be measured by coring or with the help of microelectrodes. However, the direct measurement of organic matter remineralization, which occurs through various pathways ranging from aerobic remineralization and denitrification to sulfate reduction, is not possible (Boudreau, 1997). Furthermore, organic matter remineralization is known to be variable in space and time, especially in estuarine and shelf systems, and depends in non-linear ways on a number of factors such as the rate of organic matter deposition, temperature, bottom water concentrations of nutrients and oxygen and bioturbation (Burdige, 2011). Local, regional and global estimates of sediment remineralization rates and of the resulting sediment–water fluxes thus have to rely on a combination of measurements and numerical models (Berner, 1980; Boudreau 1997, 2000).

Consideration of sediment remineralization in biogeochemical models is especially important in regional models describing estuarine and shelf systems and in global ocean models that are used to simulate processes over geologic timescales; however, sediment remineralization is not always represented realistically (see review by Soetaert et al., 2000). Following Soetaert et al. (2000) the numerical treatment of sediment remineralization in biogeochemical models can be categorized as follows: (1) in the simplest approach constant oxygen and nutrient fluxes are assigned at the sediment–water interface or sediment fluxes are ignored

* Corresponding author. Tel.: +1 902 494 4526; fax: +1 902 494 3877.
E-mail address: Katja.Fennel@dal.ca (K. Fennel).

0278-4343/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.csr.2013.05.003

altogether and instead deep water concentrations are assigned; (2) *sediment parameterizations* predict sediment fluxes as a function of nutrient and oxygen concentrations in overlying bottom waters, but do not represent the sediments with any time-dependent state variables (e.g., Middelburg et al., 1996; Fennel et al., 2006; Hetland and DiMarco, 2008); (3) *depth-integrated diagenetic models* represent the sediment in one or two layers, each layer typically describing a distinct chemical environment (e.g., Emerson et al., 1984; DiToro, 2001; Bohlen et al., 2012); and (4) *depth-resolved diagenetic models* divide the sediment into many vertical layers with the aim of resolving profiles of solid and pore water constituents (e.g., Dhakar and Burdige, 1996; Soetaert, 1996a,1996b; Kelly-Gerreyn et al., 1999; Katsev et al., 2007; Dale et al., 2011). Only the latter two types, the depth-integrated and depth-resolved diagenetic models, store concentrations of solid and dissolved sediment constituents as state variables and are able to account for the effects of environmental history on remineralization. Diagenesis is defined as "the sum total of processes that bring about changes in a sediment or sedimentary rock, subsequent to deposition" (Berner, 1980, p. 3). Diagenetic models thus describe transport and reaction processes, which can be the result of biological or physical phenomena (Boudreau, 1997).

Which of the above approaches is chosen to represent sediment remineralization in a biogeochemical model should be motivated by the importance of benthic–pelagic coupling in the system to be studied, although the computational effort has to be considered as well; for instance, sediments in estuarine and continental shelf systems are known to account for a large fraction of total system respiration (Kemp et al. 1992) and can act as an important sink for bioavailable nitrogen through denitrification (Seitzinger et al., 2006; Rao et al., 2007), which can greatly affect primary production and even air-sea fluxes of CO_2 (Fennel et al., 2008). Thus the use of constant sediment–water fluxes or constant bottom water concentrations (category 1 from above) should be ruled out for these systems. In systems where bottom waters are well oxygenated and bottom water conditions are stable, simple parameterizations (category 2) may be appropriate for describing sediment–water exchange fluxes, while systems with variable conditions, especially those prone to hypoxic and anoxic bottom waters (see, e.g., Katsev et al., 2007), may require diagenetic models (categories 3 and 4). In 3-dimensional models the desired realism in representing processes underlying sediment remineralization has to be balanced by computational feasibility. Diagenetic models are computationally much more costly than the simpler parameterizations, which is one of the reasons they are seldom used in 3-dimensional biogeochemical models. Furthermore, layered diagenetic models are much more computationally demanding than vertically integrated diagenetic models.

The overarching objective of this study is to aid in the decision making process of how sediment remineralization should be represented during implementation of biogeochemical models. This study presents an assessment of the predictive abilities of a range of sediment formulations (representing categories 2 to 4 from above) against a comprehensive set of sediment–water flux measurements. More specifically, two simple parameterizations of sediment oxygen uptake representing category 2 (Hetland and DiMarco, 2008; Murrell and Lehrter, 2011) and two representative examples of diagenetic models, the two-layer model by DiToro (2001) representing category 3 and the multi-layer model by Soetaert (1996a,1996b) representing category 4, are compared against the same observational data set. The data set was obtained as part of an eutrophication experiment carried out at the Mesocosm Experimental Research Laboratory (MERL) at the University of Rhode Island, USA, over a period of 2.5 years (Nixon et al., 1984; Oviatt et al., 1986). Observation types used here include sediment–water fluxes of nitrate, ammonium, phosphate and oxygen and concentrations of the same species and temperature in the overlying water.

First, depositional fluxes of organic matter had to be parameterized since these were not observed directly. Four different parameterizations, one with constant fluxes and three dependent on overlying biomass concentrations, were formulated, optimized and compared. The parameterization that resulted in the best fit between model and observations was then used for the remainder of the experiments.

Second, model parameters were optimized by fitting the models to the observations. This was a necessary step because complex models depend on a number of parameters, many of which are difficult or impossible to measure directly. Formal optimization removes biases that could result from subjective model tuning and is key for meaningful model comparisons (Friedrichs et al., 2007). During the optimization process, which is analogous to non-linear least squares regression, the model parameters are systematically varied in order to minimize the misfit between model output and corresponding observations (Berg et al., 1998; Friedrichs et al., 2007; Bagniewski et al., 2011).

Finally, a cross-validation analysis was carried out in order to determine if any of the models are over-fitting the observations. Over-fitting is an undesirable feature that can occur when a model contains too many free parameters. On the one hand a model with many free parameters (or degrees of freedom) might easily be adapted to a particular observation set, but at the risk of fitting parameters to meaningless variation (noise) in the observations. Over-fitting thus may decrease the model's ability to predict an independent set of observations. On the other hand, a model with too few free parameters, although not as prone to over-fitting, may not be able to represent the dynamics of the system and thus could generate a poor fit. The objective is therefore to find the most parsimonious model. The results of the cross-validation will be discussed.

2. Materials and methods

2.1. Notation

The notation adopted in this text is as follows: J represents fluxes across the sediment–water interface with positive values indicating fluxes out of the sediment; r is a specific rate constant; R is a reaction rate; R^* is a vertically integrated reaction rate; and k is used to represent half saturation constants.

2.2. Dataset

Observations from a mesocosm eutrophication experiment, conducted between 1981 and 1983 at MERL to study the effects of nutrient loading on coastal ecology and sediment biogeochemistry (Nixon et al., 1984; Oviatt et al., 1986), are used as model inputs and for model optimization and evaluation. Full details of the experiment including the resulting observations are available in Frithsen et al. (1985a, 1985b, 1985c). Only the key features are repeated briefly here.

Six of the nine mesocosms were subjected to a wide range of nutrient loads (ammonium, phosphate and silicate were added weekly in the stoichiometric ratio 12.8 N: 1.0 P: 0.91 Si); the remaining three were used as controls (referred to as C1, C5 and C8). The first of the six mesocosms (referred to as $1\times$) received 7.57 mmol N/day, and each one thereafter (referred to as $2\times$, $4\times$, $8\times$, $16\times$, $32\times$) received double the nutrient load of the preceding mesocosm so that the magnitude of nutrient inputs increased geometrically between mesocosms. The stoichiometry of the nutrient additions resembles that of sewage and the nutrient

concentrations in the 32× mesocosm resembles those of the heavily eutrophied New York Harbor.

The cylindrical mesocosms (7 m deep, 1.5 m in diameter, open at the top) were engineered to mimic conditions in the adjacent Narragansett Bay. Each mesocosm contained sediment (40 cm in depth) that was transplanted from the bay. Mesocosm temperatures were equilibrated by heat exchangers with adjacent bay water, and bay water was pumped through the tanks at a turnover rate close to that of the bay (~27 days). Tidal mixing was simulated with an automatic stirrer on a 6-h schedule with mixing rates chosen to match the concentration of resuspended sediment in the mesocosm with that of the bay (~3 mg/L).

Throughout the experiment, monthly measurements of water properties (O_2, NO_3, NH_4, PO_4, temperature, chlorophyll and zooplankton biomass) were taken and benthic chambers were used to measure the exchange fluxes of O_2, NO_3, NH_4, and PO_4 between the sediments and the overlying water column. The evolution of temperature and of O_2, NO_3, NH_4, and PO_4 concentrations is shown in Figs. 1 and 2.

2.3. Reaction rate estimates

Since direct measurements of benthic reaction rates are not available from this experiment, rates of carbon remineralization, denitrification, and nitrification were estimated assuming that fluxes of oxygen, ammonium and nitrate are in steady state as in Fennel et al. (2009). The three known fluxes between sediment and overlying water column are the oxygen flux J_{O_2} (μmol O_2/cm^{2}d), the ammonium flux J_{NH_4} (μmol N/cm^{2}d) and the nitrate flux J_{NO_3} (μmol N/cm^{2}d). The four unknown, vertically integrated reaction rates are the total rate of carbon remineralization R^*_{met} (μmol C/cm^{2}d), the rate of carbon remineralization via denitrification R^*_{dnf} (μmol C/cm^{2}d), the rate of carbon remineralization via the sum of all other forms of remineralization R^*_{aer} (μmol C/cm^{2}d)and the rate of nitrification R^*_{nit} (μmol O_2/cm^{2}d). Assuming that the composition of organic matter follows Redfield stoichiometry (106C:16N:1P), the following equilibrium relationships result:

$$R^*_{met} = R^*_{aer} + R^*_{dnf}$$

$$J_{O_2} = -\frac{1\ \text{mol O}_2}{1\ \text{mol C}} R^*_{aer} - R^*_{nit}$$

$$J_{NH_4} = \frac{16\ \text{mol N}}{106\ \text{mol C}} R^*_{met} - \frac{1\ \text{mol N}}{2\ \text{mol O}_2} R^*_{nit}$$

$$J_{NO_3} = \frac{1\ \text{mol N}}{2\ \text{mol O}_2} R^*_{nit} - \frac{4\ \text{mol N}}{5\ \text{mol C}} R^*_{dnf}.$$

The rate R^*_{aer} represents the oxygen sink due to aerobic respiration as well as the reoxidation of reduced species (H_2S, CH_4, etc.) that were produced by sulphate reduction, methane production, or metal oxide reductions. Given the four equations above and the three known fluxes, the four unknown rates were solved for as follows:

$$R^*_{met} = -\left(212\frac{\text{mol C}}{\text{mol O}_2} J_{O_2} - 159\frac{\text{mol C}}{\text{mol N}} J_{NH_4} + 265\frac{\text{mol C}}{\text{mol N}} J_{NO_3}\right)/236$$

$$R^*_{aer} = -\left(43\frac{\text{mol C}}{\text{mol O}_2} J_{O_2} - 106\frac{\text{mol C}}{\text{mol N}} J_{NH_4} - 20\frac{\text{mol C}}{\text{mol N}} J_{NO_3}\right)/59$$

$$R^*_{dnf} = -\left(40\frac{\text{mol C}}{\text{mol O}_2} J_{O_2} + 265\frac{\text{mol C}}{\text{mol N}} J_{NH_4} + 345\frac{\text{mol C}}{\text{mol N}} J_{NO_3}\right)/236$$

$$R^*_{nit} = -(16 J_{O_2} + 106\frac{\text{mol O}_2}{\text{mol N}} J_{NH_4} + 20\frac{\text{mol O}_2}{\text{mol N}} J_{NO_3})/59$$

The resulting reaction rates are later compared to the corresponding rates simulated by the diagenetic models (see discussion in Section 4.2 below and Figs. 6, 5 and 11).

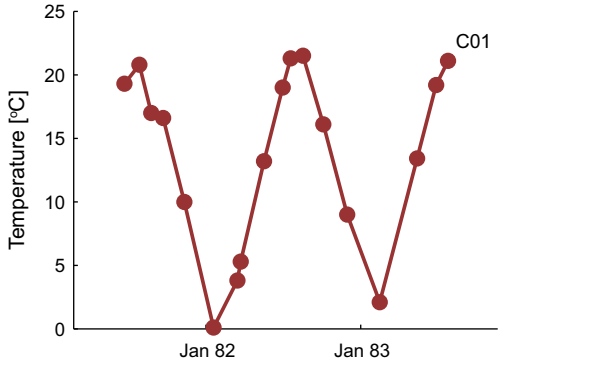

Fig. 1. Water temperature in control mesocosm C01. The temperature in the other mesocosms is virtually identical.

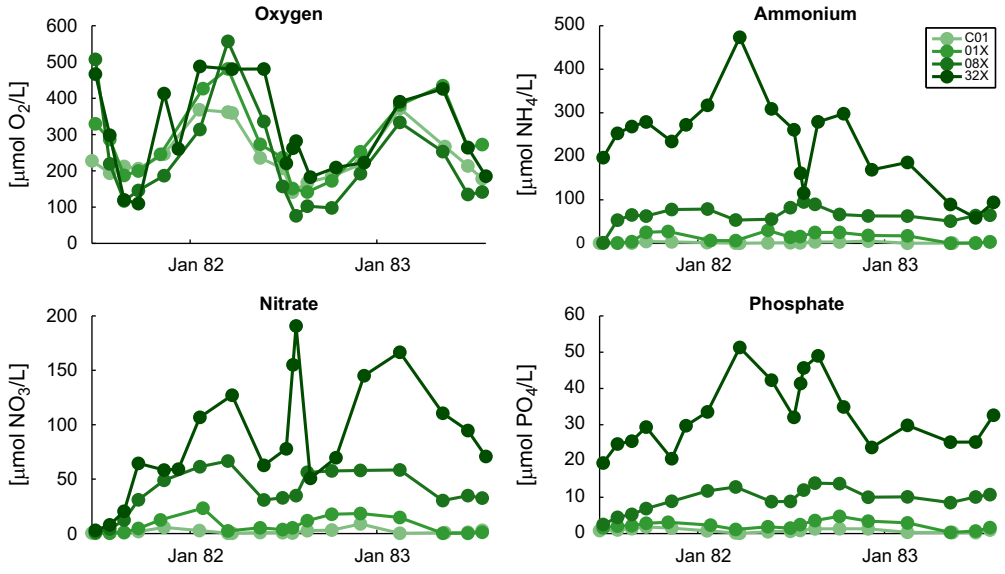

Fig. 2. Concentration of dissolved oxygen, ammonium, nitrate and phosphate in control mesocosm C01 and mesocosms 1×, 8× and 32×.

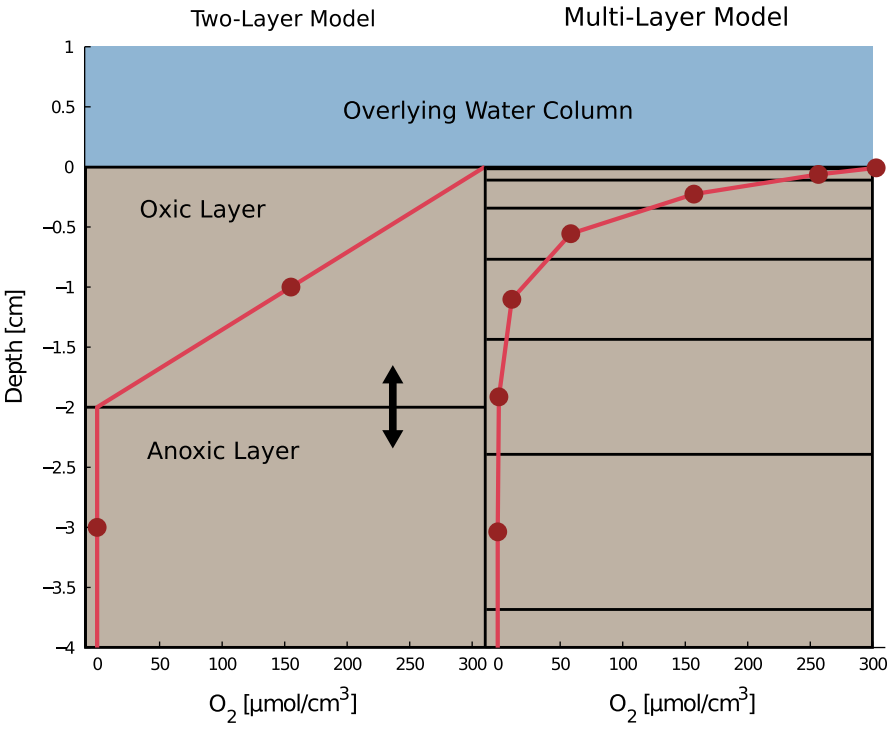

Fig. 3. Spatial representation of the two- and multi-layer models. Red dots represent typical oxygen concentrations for each layer; red lines indicate implied oxygen profiles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
State variables of the two-layer model.

State variable	Description	Units
$\mathbf{C_i}$	Solid-volume concentration of organic carbon of lability class i ($i=1$: very labile; $i=2$: labile, $i=3$: inert)	$\frac{\mu mol\ C}{L_s}$
$\mathbf{NH_3}$	Concentration of total inorganic ammonia (dissolved and solid phases)	$\frac{\mu mol\ NH_3}{L}$
$\mathbf{NO_3}$	Concentration of total dissolved inorganic nitrate and nitrite	$\frac{\mu mol\ NO_3}{L}$
$\mathbf{PO_4}$	Concentration of total dissolved inorganic phosphate (dissolved and solid phases)	$\frac{\mu mol\ PO_4}{L}$
$\mathbf{H_2S}$	Concentration of total hydrogen sulfide (dissolved gas and solid phases)	$\frac{\mu mol\ H_2S}{L}$
$\mathbf{CH_4}$	Concentration of total methane (dissolved gas and solid phases)	$\frac{\mu mol\ CH_4}{L}$
bnthstr	Biological stress index. A time-dependent, oxygen-sensitive parameter responsible for adjusting bioturbation rates in suboxic environments	–

2.4. Simple oxygen flux parameterizations

Two empirical oxygen flux parameterizations were assessed. These parameterizations use only one or two parameters (p_1 and p_2) and require only one or two variables as input. The first parameterization (referred to as ML11) was suggested by Murrell and Lehrter (2011) and assumes that oxygen flux into the sediment $J_{O_2,ML}$ increases linearly with oxygen concentration in the overlying water, i.e.

$$J_{O_2,ML} = -p_1[O_2]$$

The second parameterization (referred to as HD08) was used in Hetland and DiMarco (2008) and uses overlying water oxygen concentration and temperature as inputs. This parameterization assumes that oxygen uptake by the sediment $J_{O_2,HD}$ is proportional to overlying oxygen concentrations only at low concentrations and that metabolic rates and hence oxygen uptake accelerate with increasing temperatures according to

$$J_{O_2,HD} = -p_1 2^{\frac{T}{10}}\left(1-\exp\left(-\frac{[O_2]}{p_2}\right)\right)$$

here T represents temperature in °C. Temperature dependence (i.e. the factor $2^{T/10}$) follows the "Q_{10}" rule such that an increase by 10° roughly doubles rates of reaction.

2.5. Two-layer diagenetic model

The two-layer diagenetic model (DiToro, 2001) simulates three carbon remineralization reactions (sulphate reduction, denitrification, and methane production), three oxidation reactions of pore water solutes (nitrification, sulfide oxidation and methane oxidation) and diffusive and bioturbative transports between the two vertical layers (a thin aerobic layer on the surface and a thicker anaerobic layer below). These two layers provide a coarse resolution for the depth profiles of solid and dissolved chemical species (Fig. 3) and differ in their biochemical representations in terms of processes and process rates. Aerobic remineralization is not explicitly represented in this model, but oxygen uptake by the sediment results from the three oxidation reactions mentioned above. Since organic matter respiration in coastal sediments is known to be dominated by anaerobic respiration (e.g. Burdige, 2006) this is not an unreasonable assumption. The model does not include dissimilatory nitrate reduction to ammonium or anammox. The model state variables are listed in Table 1.

The model represents carbon respiration with a specific rate constant $r_{\mathrm{met},l}$ and temperature adjustment factor $\theta_l$for each of the three carbon lability classes $[C]_l$ where $l=1,\ldots,3$. The combined respiration rate is

$$R_{\mathrm{met}} = \sum_l r_{\mathrm{met},l}\theta_l^{T-20}[C]_l$$

The temperature adjustment factors increase the reaction rate by slightly over a factor of 2 for every 10 °C increase. Remineralized ammonium and phosphate are released when carbon is consumed according to Redfield stoichiometry.

Denitrification rates are calculated separately in each layer, independent of the carbon respiration R_{met}. Two denitrification rates, $R_{\mathrm{dnf},i}$ (μmol NO3/cm^{3}d), with $i=1$, 2 for the aerobic and anaerobic layers, are calculated as

$$R_{\mathrm{dnf},i} = r_{\mathrm{dnf},i}\theta_{\mathrm{dnf}}^{T-20}[NO_3]_i$$

where $r_{\mathrm{dnf},i}$ are rate parameters (1/d) and $\theta_{\mathrm{dnf}}^{T-20}$ is a dimensionless temperature adjustment factor.

The model assumes that carbon remineralization R_{met} is larger than the carbon-based denitrification rate at all times and that leftover carbon remineralization occurs either by sulphate reduction or methane production. Sulphate reduction is energetically more favorable, and usually sufficient sulphate is present to support it. The model calculates sulphate availability from the salinity of the overlying water column, and in the rare situation that this supply is insufficient, all remaining carbon consumption leads to methane production.

Nitrification only takes place in the aerobic model layer according to the equation

$$R_{\mathrm{nit}} = r_{\mathrm{nit}}\theta_{\mathrm{nit}}^{T-20}\frac{[\mathrm{NH_4}]_1}{k_{\mathrm{nit,NH_4}}\theta_{k,\mathrm{nit,NH_4}}^{T-20}+[\mathrm{NH_4}]_1}\frac{[\mathrm{O_2}]_1}{k_{\mathrm{nit},O_2}+\mathrm{O_2}}$$

where R_{nit} is the nitrification rate (μmol N/cm^{3}d), r_{nit} is the rate constant (1/d), $[\mathrm{NH_4}]_1$ is the aerobic layer's ammonium concentration (μmol N/cm^3), k_{nit,O_2} and $k_{\mathrm{nit,NH_4}}$are Michaelis–Menten parameters for oxygen and ammonium respectively (μmol/cm^3), and $\theta_{\mathrm{nit}}^{T-20}$ and $\theta_{k,\mathrm{nit,NH_4}}^{T-20}$ temperature-dependence factors.

Hydrogen sulfide oxidation is governed by the equations

$$R_{\mathrm{H_2S}} = r_{\mathrm{H_2S}}\,\theta_{\mathrm{H_2S}}^{T-20}\frac{[\mathrm{O_2}]_1}{K_{\mathrm{H_2Sox},O_2}}[\mathrm{H_2S}]$$

and

$$R_{\mathrm{S}} = r_{\mathrm{S}}\,\theta_{\mathrm{H_2S}}^{T-20}\frac{[\mathrm{O_2}]_1}{K_{\mathrm{H_2Sox},O_2}}[\mathrm{S}]$$

where $R_{\mathrm{H_2S}}$ and R_{S} are the sulfide oxidation rates for aqueous and solid phases, respectively, (units μmol O$_2$/cm^{3}d), $r_{\mathrm{H_2S}}$ and r_{S} govern the reaction rate (units μmol O_2/d cm^3), [H$_2$S] and [S] are the concentrations of aqueous hydrogen sulfide and solid metal sulfides (units μmol S/cm^3), $K_{\mathrm{H_2S},O_2}$ is a scaling factor (μmol O_2/cm^3), and $\theta_{\mathrm{H_2S}}^{T-20}$ is a temperature adjustment factor.

Methane oxidation rates, $R_{\mathrm{CH_4,ox}}$ (units μmol C/cm^{3}d) are found using

$$R_{\mathrm{CH_4,ox}} = r_{\mathrm{CH_4,ox}}\theta_{\mathrm{CH_4,ox}}^{T-20}\frac{[\mathrm{O_2}]_1}{k_{\mathrm{CH_4,ox}}+[\mathrm{O_2}]_1}[\mathrm{CH_4}]$$

where $r_{\mathrm{CH_4,ox}}$ is the specific rate constant, $\theta_{\mathrm{CH_4,ox}}^{T-20}$ is a temperature adjustment factor and $k_{\mathrm{CH_4,ox}}$ is a Michaelis–Menten parameter.

Some model state variable concentrations (most notably PO_4) exist in two phases, aqueous and solid, which are diffusively transported according to different parameters. In each layer the concentrations of both phases are redistributed every time step to ensure equilibrium according to $\Pi_i = C_{i,\mathrm{s}}/C_{i,\mathrm{aq}}$ where $C_{i,\mathrm{s}}$ and $C_{i,\mathrm{aq}}$ are the concentrations of the solid and aqueous phases of chemical species i. In the case of PO_4 the partitioning parameter $\Pi_{\mathrm{PO_4}}$ is sensitive to oxygen according to $\Pi_{\mathrm{PO_4}} = \pi_1\pi_2^{([\mathrm{O_2}]/[\mathrm{O_2}]_{\mathrm{crit}})}$when$[\mathrm{O_2}] > [\mathrm{O_2}]_{\mathrm{crit}}$ and $\Pi_{\mathrm{PO_4}} = \pi_1$ when $[\mathrm{O_2}] \le [\mathrm{O_2}]_{\mathrm{crit}}$ The two-layer model assumes that oxygen concentrations achieve steady state at every time step and iteratively solves for the steady state oxygen demand and concentrations. All of the sediment source and sink terms are calculated during this iterative process. For more details see DiToro (2001) and Wilson (2011).

2.6. Multi-layer diagenetic model

The multi-layer diagenetic model of Soetaert (1996a,1996b) divides the sediment into multiple layers in order to explicitly resolve the depth profiles of chemical species (Fig. 3). The original model includes organic carbon, O_2, NO_3, NH_4, and ODU (oxygen demand units, i.e. any highly reduced chemical such as H_2S) as state variables, but was modified for this study to also include solid and aqueous phases of PO_4. The model calculates rates of carbon metabolism (aerobic, denitrification and anaerobic), porewater oxidation (nitrification and oxidation of ODUs), and diffusive and bioturbative fluxes. Similar to the two-layer model, dissimilatory nitrate reduction and anammox are not included. The time- and space-dependent concentrations of all model state variables (listed in Table 2) are described by differential equations that are solved by finite differences approximations. Thus, model concentrations are not necessarily in equilibrium at all times.

Organic carbon is represented by two lability classes, one more and one less reactive, and is respired with a decay rate of

$$R_{\mathrm{met}} = \sum_l r_{\mathrm{met},l}[C]_l\theta_{\mathrm{met},l}^{T-20}$$

where $r_{\mathrm{met},l}$ is the maximum rate constant for carbon lability class l and $\theta_{\mathrm{met},l}$ is a temperature adjustment factor.

Similar to the two-layer model the total carbon consumption rate is calculated for each layer and partitioned between three

Table 2
State variables of the multi-layer model.

State variable	Description	Units
C	Solid-volume concentration of organic carbon of lability class i ($i=1$: inert; $i=2$: labile)	$\frac{\mu\mathrm{mol\ C}}{\mathrm{L_s}}$
O_2	Porewater concentration of dissolved oxygen	$\frac{\mu\mathrm{mol\ O_2}}{\mathrm{L_{pw}}}$
NH_3	Porewater concentration of dissolved inorganic ammonia	$\frac{\mu\mathrm{mol\ NH_3}}{\mathrm{L_{pw}}}$
NO_3	Porewater concentration of dissolved inorganic nitrate and nitrite	$\frac{\mu\mathrm{mol\ NO_3}}{\mathrm{L_{pw}}}$
ODU	Oxygen demand units, a placeholder for one or more reduced end-products of an anaerobic Ćmetabolism (i.e. H_2S, CH_4, metal sulfides etc.). Transported as an aqueous phase within the porewater	$\frac{\mu\mathrm{mol\ O_2}}{\mathrm{L_{pw}}}$
PO_{4pw}	Porewater concentration of dissolved inorganic phosphate	$\frac{\mu\mathrm{mol\ PO_4}}{\mathrm{L_{pw}}}$
PO_{4s}	Solid-volume concentration of inorganic phosphate precipitate	$\frac{\mu\mathrm{mol\ PO_4}}{\mathrm{L_s}}$

remineralization pathways (aerobic, denitrification and other anaerobic) according to the availability of reactants by calculating limitation terms for each metabolism as

$$lim_{\mathrm{aer}} = \frac{[O_2]}{[O_2] + k_{\mathrm{aer},O_2}},$$

$$lim_{\mathrm{dnf}} = \frac{[NO_3]}{[NO_3] + k_{\mathrm{dnf},NO_3}} \frac{k_{\mathrm{dnf},O_2}}{[O_2] + k_{\mathrm{dnf},O_2}},$$

$$lim_{\mathrm{anox}} = \frac{k_{\mathrm{anox},NO_3}}{[NO_3] + k_{\mathrm{anox},NO_3}} \frac{k_{\mathrm{anox},O_2}}{[O_2] + k_{\mathrm{anox},O_2}}$$

and

$$\sum lim = lim_{\mathrm{aer}} + lim_{\mathrm{dnf}} + lim_{\mathrm{anox}}$$

where k_{aer,O_2}, k_{dnf,O_2}, k_{dnf,NO_3}, k_{anox,NO_3}, and k_{anox,O_2} are Michaelis–Menten parameters. The model assigns the fractions of decomposed carbon, $R_{\mathrm{met}}(lim_{\mathrm{aer}}/\sum lim)$, $R_{\mathrm{met}}(lim_{\mathrm{dnf}}/\sum lim)$ and $R_{\mathrm{met}}(lim_{\mathrm{anox}}/\sum lim)$ to aerobic remineralization, denitrification and other anaerobic remineralization, respectively.

Pore water oxidization reactions are controlled with a maximum rate constant and a Michaelis-Menten parameter. For nitrification

$$R_{\mathrm{nit}} = r_{\mathrm{nit}}[NH_4] \frac{[O_2]}{[O_2] + k_{O_{2,\mathrm{nit}}}},$$

where r_{nit} is the maximum nitrification rate and $k_{O_{2,\mathrm{nit}}}$ the half-saturation constant. Oxidation of ODU is handled the same way as

$$R_{\mathrm{ODU_{ox}}} = r_{\mathrm{ODU_{ox}}}[\mathrm{ODU}] \frac{[O_2]}{[O_2] + k_{\mathrm{ODU_{ox}}O_2}}$$

The multi-layer model was modified to additionally represent the dynamics of PO_4. When organic matter is remineralized, PO_4 is released according to the Redfield ratio. As in the two-layer model, PO_4 is represented in solid and aqueous forms, and re-equilibrated every time step to satisfy $\Pi_{PO_4} = PO_{4s}/PO_{4aq}$ where PO_{4s} and PO_{4aq} are the concentrations of solid and aqueous phosphate, and Π_{PO_4} is the equilibrium constant. This constant is sensitive to oxygen and evaluated every time step using three parameters, π_1, π_2 and $[O_2]_{\mathrm{crit}}$ as

$$\Pi_{PO_4} = \pi_1 \pi_2^{([O_2]/[O_2]_{\mathrm{crit}})} \text{when} [O_2] > [O_2]_{\mathrm{crit}}$$

and

$$\Pi_{PO_4} = \pi_1 \text{ when} [O_2] \leq [O_2]_{\mathrm{crit}}.$$

2.7. POM flux parameterizations

During the eutrophication experiment no observations of the depositional flux of particulate organic matter (POM flux) were made. Since the depositional flux is a necessary input for both layered models it was parameterized. Four parameterization methods were tested:

- Method A follows DiToro (2001) in assigning a constant POM flux to each mesocosm over the entire 2.5 year simulation, i.e.

  $J_{\mathrm{POM}} = \mathrm{const.}$

- In method B the POM deposition scales in direct proportion to chlorophyll concentrations in the overlying water, i.e.,

  $J_{\mathrm{POM}} = p_{\mathrm{chla}}[\mathrm{Chl}]$,

  where p_{chla} is a constant.
- In method C the POM flux varies linearly with the chlorophyll of diatoms and other phytoplankton, approximating the relative fractions from algal abundance counts as

  $$J_{\mathrm{POM}} = \frac{p_{\mathrm{dia}}A_{\mathrm{dia}} + p_{\mathrm{other}}(A_{\mathrm{tot}} - A_{\mathrm{dia}})}{A_{\mathrm{tot}}},$$

  where A_{dia} is the diatomaceous algal count, and A_{tot} the total algal count and p_{dia} and p_{other} are constants.
- In method D the POM flux was constructed from a linear combination of chlorophyll-*a* and zooplankton biomass as

  $J_{\mathrm{POM}} = p_{\mathrm{chl}}[\mathrm{Chl}] + p_{\mathrm{zoo}}[\mathrm{zoo}]$,

  where again p_{chl} and p_{zoo} are constants.

The constants in these parameterizations were determined by nonlinear optimization as described below.

2.8. Optimization methods and experiments

The misfit between models and observations was quantified with a cost function,

$$F(\vec{p}) = \sum_{m=1}^{M} \frac{1}{D} \sum_{d=1}^{D} \frac{1}{S_d} \sum_{i=1}^{I_m} \frac{(X_{m,d,i}^{\mathrm{mod}}(\vec{p}) - X_{m,d,i}^{\mathrm{obs}})^2}{\sigma_{m,d}^2},$$

where $X_{m,d,i}^{\mathrm{obs}}$ are the available observations ($i = 1, \ldots I_m$) for mesocosm m and observation type d, $X_{m,d,i}^{\mathrm{mod}}(p)$ are the corresponding model outputs which depend on the input parameter set p; $M = 9$ is the number of mesocosms, $D = 4$ is the number of different data types $d \in \{J_{O_2}, J_{NO_3}, J_{NH_4}, J_{PO_4}\}$, I_m is the number of observations for each data type and mesocosm, S_d is a weight for each data type and $\sigma_{m,d}^2$ is the uncertainty of a given observation. The weights S_d can be thought of as normalization between data types (accounting for different units and magnitudes of fluxes) and were chosen so that each data type contributes equally to the cost function value of the baseline simulation of the two-layer model. Inclusion of the uncertainties $\sigma_{m,d}^2$ guarantees that uncertain observations contribute less to the cost function than more precise observations. This ensures, for example, that the highly uncertain nitrate, ammonium and phosphate fluxes in the most eutrophic mesocosms ($16\times$ and $32\times$) are not overrepresented. The weights and uncertainties are given in Table 3.

The cost function $F(\vec{p})$ can be algorithmically minimized with respect to the parameter set $\vec{p}$ (also referred to as optimization) using a number of different techniques. The parameter set that represents the minimum of the cost function (i.e. results in the smallest misfit between observations and model) is referred to as the optimal parameter set. The cost function value of the optimal parameters is a useful diagnostic for comparing the different

Table 3
Weights and uncertainties used for cost calculation according to mesocosm and datatype. Uncertainties are from Kelly et al. (1985).

Datatype (d)	S_d	$\sigma^2_{C01,d}$	$\sigma^2_{C05,d}$	$\sigma^2_{C08,d}$	$\sigma^2_{01X,d}$	$\sigma^2_{02X,d}$	$\sigma^2_{04X,d}$	$\sigma^2_{08X,d}$	$\sigma^2_{16X,d}$	$\sigma^2_{32X,d}$
J_{O_2}	53.	0.053	0.065	0.053	0.221	0.054	0.122	0.218	1.097	1.116
J_{NO_3}	30.	0.036	0.036	0.036	0.028	0.072	0.120	1.164	2.328	3.377
J_{PO_4}	107.	0.009	0.009	0.010	0.012	0.010	0.009	0.125	0.086	0.058

models. The uncertainty in the optimal cost function value due to observational errors was estimated with the help of a Monte Carlo analysis as in Bagniewski et al. (2011). For this purpose the observations were perturbed by adding random noise to each data point drawn from normal distributions with a mean of zero and a standard deviation equal to the standard error of the corresponding observation. 1000 such randomly perturbed data sets were created, the corresponding cost function values computed and the resulting standard deviations calculated. These standard deviations are reported along with the optimal cost function values. Baseline parameter values (i.e. the initial guess for optimization) for the two- and multi-layer models were taken from their original publications (Soetaert et al., 1996b; DiToro, 2001).

In all optimizations time-dependent model simulations were performed repeatedly for each mesocosm over the 2.5-year duration of the experiment. In these simulations each mesocosm was initialized with the final state of a multi-year simulation for control mesocosm C1 in order to ensure that a dynamic steady state had established.

First, the depositional flux parameterizations (methods A through D) were optimized using a gradient descent algorithm. All methods (A–D) were optimized for the two-layer model. For methods B, C and D two sets of optimizations were performed, one where the same parameters were applied across all mesocosms (referred to as joint optimizations) and one where each mesocosm had its own parameter set (referred to as individual optimizations). Note that method A assumes a different constant for each mesocosm and thus should be considered an individual optimization. As will be described below, methods C and D performed poorly and thus were not used further. Then the depositional flux parameterizations A and B were optimized (the latter individually and jointly) for the multi-layer model using an evolutionary algorithm (Mattern, 2008; Wood et al., in review). The best depositional flux parameterization was chosen and used in the remaining optimization experiments for the two-layer and multi-layer models.

In preparation for optimization of the model parameters, the curvature of the cost function with respect to changes in parameter values was analyzed surrounding the baseline parameter set and the ten most sensitive model parameters were determined. These 10 model parameters along with the nine POM flux input parameters were then optimized to fit the model-predicted sediment fluxes to observed oxygen, ammonium, nitrate and phosphate fluxes. These optimizations used the genetic algorithm for both, the two-layer and the multi-layer model. In addition the two simple oxygen flux parameterizations were optimized using a standard nonlinear least squares routine.

3. Results

3.1. Depositional flux optimizations

The depositional flux of particulate organic matter was parameterized according to four different methods that are defined in Section 2.7 and were optimized for the two-layer model as described in Section 2.8. Parameterizing the POM flux as a constant (method A) resulted in a minimum cost of 0.940 ± 0.036, which is the smallest misfit achieved by any parameterization. The other three parameterizations (B, C and D) estimate POM fluxes using a qualitatively different approach, scaling the POM flux by water column biomass observations. The minimum costs for the joint optimizations of methods B, C and D are 1.71 ± 0.053, 1.51 ± 0.047 and 1.51 ± 0.047, respectively. The minimum costs for the individual optimizations of methods B, C and D are 1.13 ± 0.041, 1.10 ± 0.040 and 1.37 ± 0.045, respectively. As expected, the jointly optimized parameterizations produced larger costs than the individual optimizations, which have more free parameters and thus more degrees of freedom that can be utilized in the optimization process. The resulting overall mean POM fluxes are shown in Fig. 4 and their parameter values are listed in Table 4.

When comparing the optimal depositional fluxes for the different parameterization methods it appears that some methods are more sensitive to nutrient load than others. According to method A depositional fluxes vary by less than a factor of 2 between the least eutrophic 1x and the most eutrophic 32× mesocosms (Fig. 4), which seems unrealistic given that the latter receives 32 times more nutrients. Jointly optimized flux parameterizations B, C, and D showed a stronger sensitivity with depositional fluxes in the 32× mesocosm being 8-fold larger than in 1×. This is a direct consequence of the increase in water column biomass along the eutrophication gradient; biomass is 8-fold larger in the 32× mesocosm compared to the controls. However, when these same parameterizations were individually optimized, their sensitivity to eutrophication more closely matched that of method A with POM fluxes varying by little over a factor of 2 between mesocosms.

Method A was always optimized individually, requiring a total of nine parameters. Since jointly implemented methods B, C, and D used only one or two parameters, it is possible that method A leads to a smaller misfit between model and observations simply because of over-fitting, and that the models with fewer parameters are actually to be preferred. More generally stated, the objective of the parameterizations was not necessarily to produce

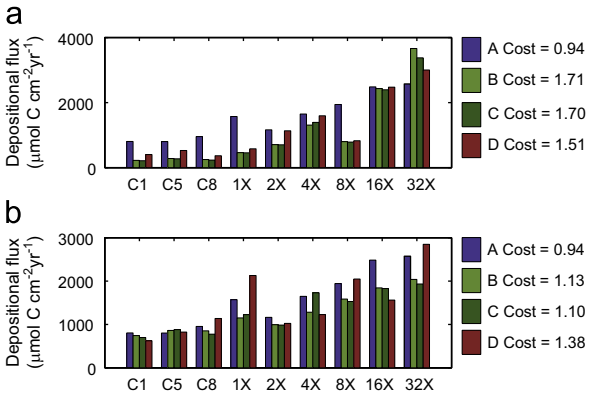

Fig. 4. Mean depositional flux predicted by optimized parameterization A and by jointly (a) and individually (b) optimized methods B, C and D.

Table 4
Parameter values for optimized depositional flux parameterizations A, B, C and D (A was optimized individually, the other three jointly). Unit of p_{const} is μmol C/cm^2 yr. Unit of p_{chla}, p_{dia} and p_{other} is (μmol C/cm^2yr) × (L/μg Chla). Unit of p_{zoo} is (μmolC/cm^2yr) × (L/μgZoopl).

Method	A	B	C		D	
Parameter	p_{const}	p_{chla}	p_{dia}	p_{other}	p_{dia}	p_{zoo}
Joint	–	3.38	2.408	0.251	2.53	0.23
C01	862.1	11.23	11.50	5.45	−0.19	0.65
C05	876.7	10.39	12.36	4.53	−0.18	0.64
C08	1014.2	11.57	11.54	4.47	0.16	1.51
01×	1783.4	8.43	10.91	4.51	0.12	2.12
02×	1192.5	4.74	4.79	−4.07	1.28	0.30
04×	1703.1	3.31	7.33	4.09	2.24	0.14
08×	2429.8	6.71	2.13	4.35	0.84	1.80
16×	2462.2	2.56	1.66	3.54	1.07	0.26
32×	2667.0	1.88	8.20	2.64	0.93	1.18

Table 5
Two-layer model f-tests comparing depositional flux parameterization A against the simpler parameterizations B_{joint}, C_{joint}, and D_{joint}.

Compare method A against	B_{joint}	C_{joint}	D_{joint}
***f*-score**	974.01	838.61	838.52
Critical value	1.95	2.02	2.02

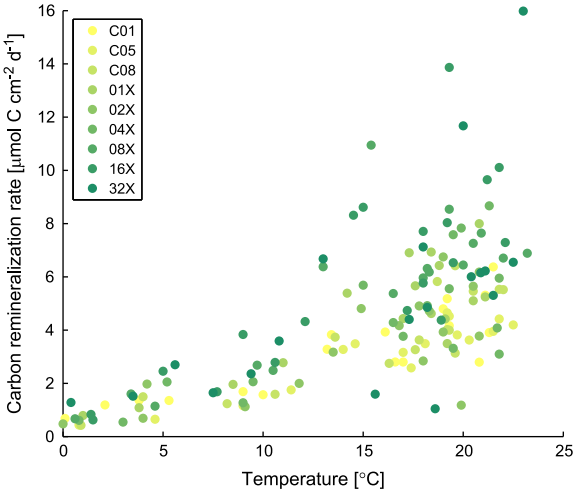

Fig. 5. Rate of carbon remineralization in the sediment plotted over temperature. Rates were calculated as described in Section 2.3.

the best fit, but to find the most parsimonious one. The relative parsimony of method A against the other methods was tested with the help of an f-test, which can check if additional parameters are justified (Soetaert, 1996a,1996b; Berg et al., 1998). An f-value, calculated from the cost values and numbers of parameters for each parameterization pair, was compared against the critical value of its f-distribution. An f-value that exceeds the critical value indicates that the extra parameters are justified by the improved fit. The results of the f-tests are shown in Table 5. For all three pairs, the f-value was much larger than the corresponding critical value, indicating that the extra parameters of method A produced a statistically significant improvement in the model-data fit.

It should be noted that optimization of the flux parameterizations did not always generate meaningful parameter values. Individual optimizations of methods C and D sometimes resulted in negative parameters implying a correlation between the presence of water column phytoplankton or zooplankton, and the *removal* of organic matter from the sediments. This is clearly in contradiction to the rationale of the parameterizations and illustrates that although individually optimized methods C and D produced smaller cost functions than the joint parameterizations, this was only a result of the optimizer exploiting the extra degrees of freedom.

All the optimizations described so far were for the two-layer model. In addition, the depositional flux parameterization A and joint and individual parameterization B was optimized for the multi-layer model. (Since methods C and D performed poorly in the optimization of the two-layer model, it was not used further in the multi-layer model.) Method A generated a cost of 0.78 ± 0.036,

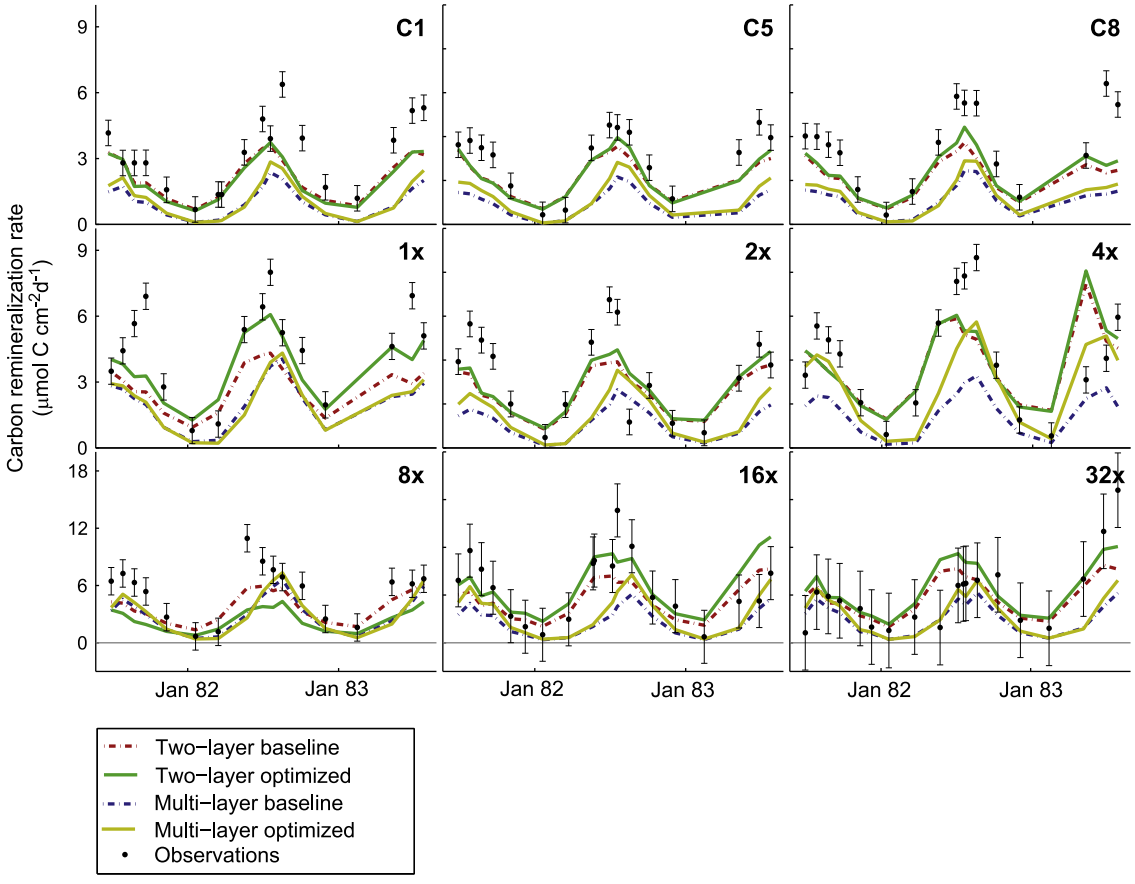

Fig. 6. Carbon remineralization rates of the baseline and optimized two-layer and multi-layer models and reaction rate estimates based on the observations (as described in Section 2.3). The simulated rates were integrated vertically (from sediment surface to 40 cm depth).

again smaller than the individually optimized method B, which scored 0.89 ± 0.040. Jointly optimized method B performed worse, scoring 0.99 ± 0.043.

Since method A yields the best model-data fits (smallest cost function values) for both models and does not appear to overfit the observations based on the *f*-tests it was used to generate depositional fluxes for both sediment models for the remainder of this study.

3.2. Sediment model parameter optimizations

The two-layer and the multi-layer diagenetic models were optimized as described in Section 2.8. The cost function was scaled such that the initial cost for the two-layer model (i.e. the cost for the baseline parameters set) is equal to 1. Optimal cost function values smaller than 1 thus indicate a better fit than the two-layer model with baseline parameters. The optimized parameters are listed in Supplementary Tables A and B.

For its optimal parameter set the cost function value of the two-layer model is 0.86 ± 0.039, which is significantly larger than the corresponding cost function value of the multi-layer model of 0.71 ± 0.036. (We consider two cost function values to be significantly different if they are outside of each other's uncertainty range.) The cost contributions of the nitrate and phosphate fluxes are not significantly different between both models with a nitrate flux contribution of 0.15 ± 0.03 and 0.13 ± 0.02 and a phosphate contribution of 0.23 ± 0.02 and 0.21 ± 0.02 for the two-layer and multi-layer models, respectively. However, the multi-layer model has a significantly smaller cost contribution from the oxygen flux of 0.17 ± 0.008 compared to the two-layer model with 0.23 ± 0.009 and a slightly smaller cost contribution from the ammonium flux of 0.20 ± 0.01 compared to the two layer model with 0.24 ± 0.02.

Since the sediment–water fluxes of oxygen and inorganic nutrients are largely driven by remineralization of organic carbon in the sediments, the simulated carbon remineralization rate is shown for the baseline and optimized simulations in Fig. 5 along with the observation-based estimates (see Section 2.3). Carbon remineralization shows a pronounced seasonal cycle, the amplitude of which increases along the eutrophication gradient. The seasonal cycle appears to be driven by temperature as illustrated in Fig. 6, where the observation-based estimates of carbon remineralization are plotted over temperature. Despite constant depositional fluxes of organic matter, the simulations reproduce the seasonal cycle (carbon remineralization is temperature-dependent in both models) and the increase in maximum fluxes across the eutrophication gradient (changes in mean remineralization fluxes roughly scale with changes in depositional fluxes). The only systematic difference between both models is that the optimized multi-layer model consistently predicts smaller rates than the two-layer model (except for brief periods in the 4 × and 8 × mesocosms).

Simulated and observed sediment–water fluxes of oxygen and inorganic nutrients are shown in Figs. 7–10 for the baseline and optimized simulations of both sediment models. The observed and simulated oxygen fluxes (Fig. 7) are anti-correlated with the carbon remineralization rates with largest sediment uptake of oxygen occurring in summer. The simulated oxygen fluxes are remarkably similar for both models, except for the 8 × mesocosm, where the multi-layer model agrees notably better with the observations.

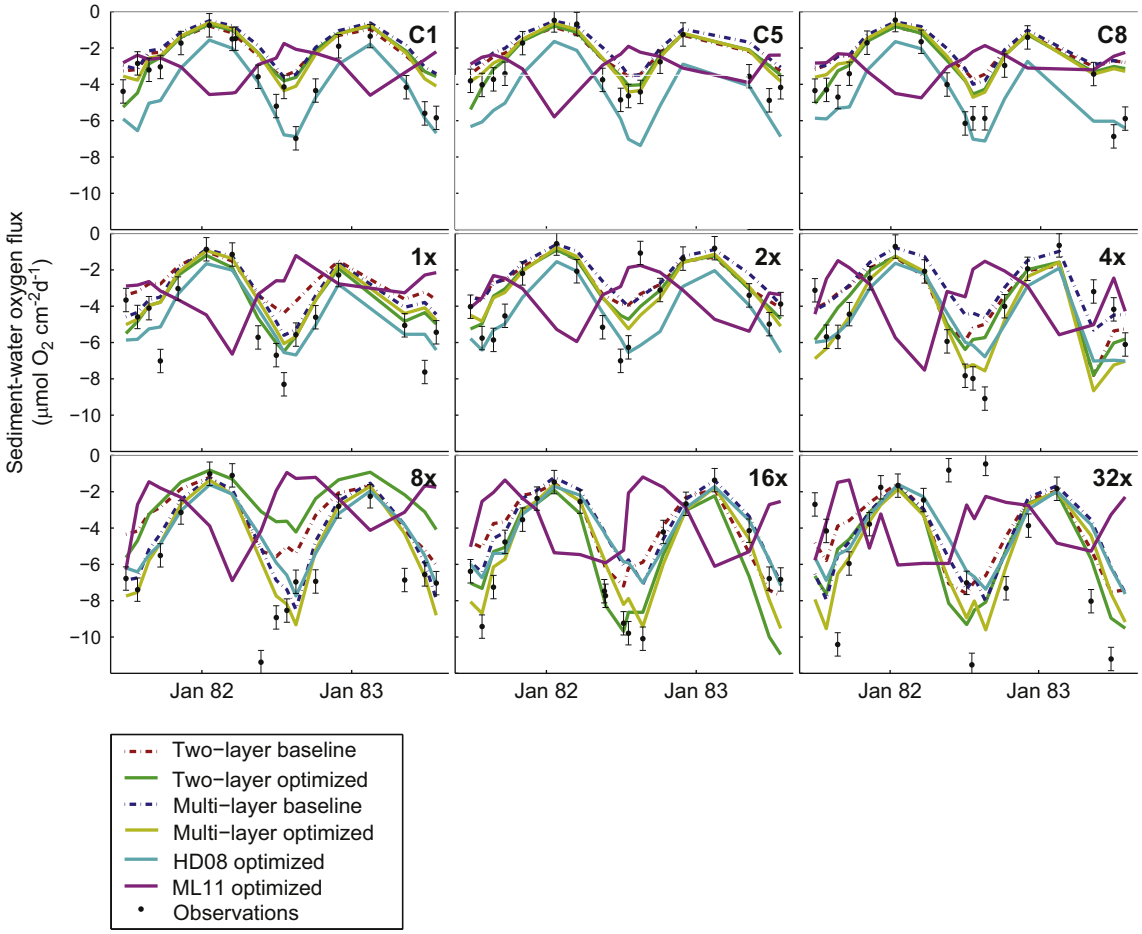

Fig. 7. Observed and simulated benthic oxygen fluxes. Simulated fluxes from baseline and optimized layered models and from optimized parameterizations are shown.

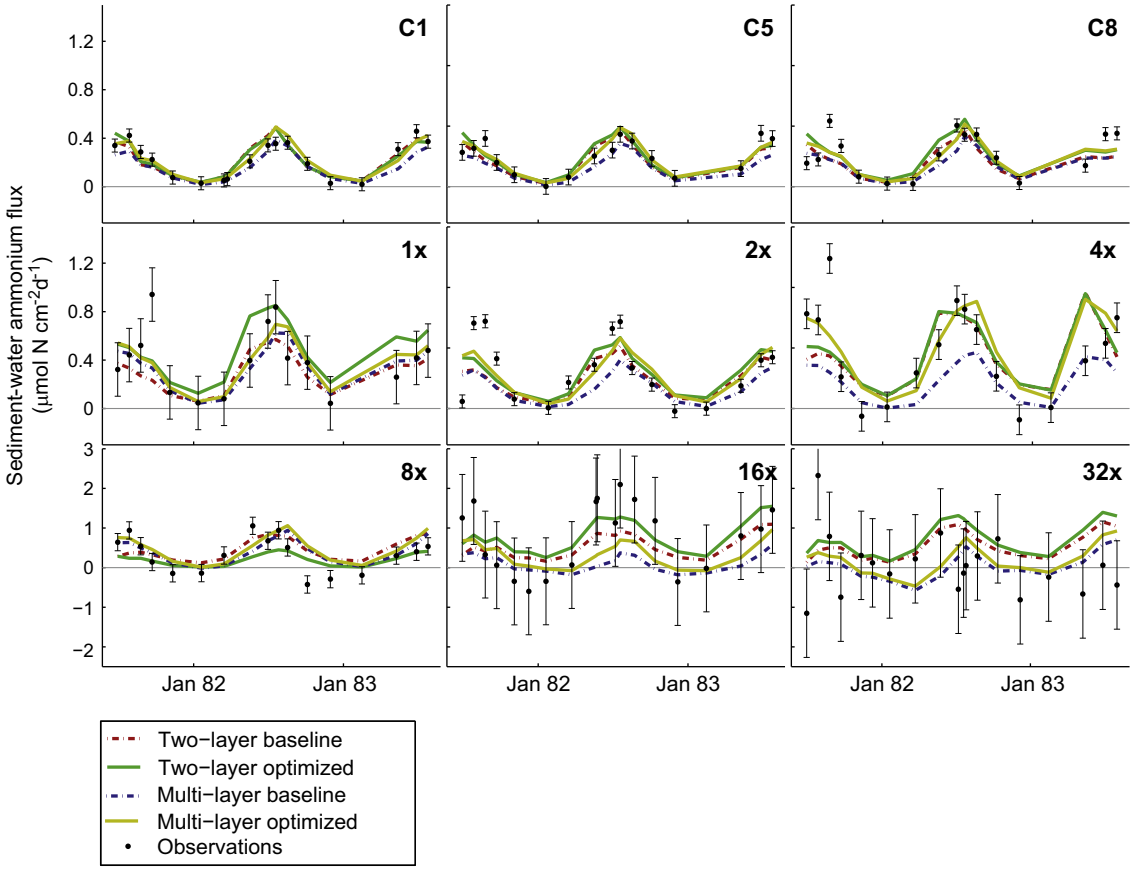

Fig. 8. Sediment–water fluxes of ammonium from the baseline and optimized two-layer and multi-layer models and observations.

The observed and simulated ammonium fluxes (Fig. 8) exhibit the same seasonal variations as the carbon remineralization rates and are generally directed out of the sediment. In winter, ammonium uptake by the sediment is observed in the $8\times$, $16\times$ and $32\times$ mesocosms, although the uncertainties on these fluxes are much larger for the latter two. While both optimized models successfully capture the magnitude and seasonality of the observed fluxes, only the multi-layer model reproduces the negative fluxes in the most eutrophic mesocosms.

Observed phosphate fluxes (Fig. 9) generally follow the same seasonal signal as carbon remineralization and in the less eutrophic mesocosms both sediment models capture this behavior well. In the $16\times$ and $32\times$ mesocosms, there appear to be some abrupt changes in the observed fluxes that both optimized models represent differently.

Observed nitrate fluxes (Fig. 10) show the least distinctive pattern with relatively constant fluxes out of the sediments in the controls and $1\times$ mesocosm and fluxes fluctuating around zero in the more eutrophic mesocosms. Both models reproduced these fluxes well as indicated by the relatively small cost contribution of the nitrate fluxes to the overall cost (see above).

For completeness sake we show the simulated and observation-based denitrification rates in Fig. 11. Both models tend to underestimate the amplitude of seasonal variations in the observation-based estimates and underestimate the maxima in summer in the controls and $1\times$ to $4\times$ mesocosms. In the $8\times$ to $24\times$ mesocosms the rate estimates come with large uncertainties; given these both models agree well with the observation-based estimates. One small but systematic difference between both models, that is notable in the controls and the $1\times$ to $4\times$ mesocosms, is that the multi-layer model represents the shape of the annual cycle correctly with largest denitrification fluxes in the summer, while the two-layer model is predicting the largest fluxes in winter. However, the magnitude of this signal is small.

3.3. *Oxygen flux parameterizations*

Two simple oxygen flux parameterizations were fit to the observations: ML11 which simply assumes a linear dependence of sediment oxygen uptake on oxygen concentrations in overlying water, and HD08, which assumes a saturating response of sediment oxygen uptake and a temperature dependence. The resulting oxygen fluxes are shown in Fig. 7 along with the model-simulated and observed fluxes.

The oxygen cost contributions for the parameterizations are 0.62 ± 0.01 for ML11 and 0.22 ± 0.01 for HD08. The misfit of the HD08 parameterization is larger than the multi-layer model (0.17 ± 0.008), but similar to the two-layer model (0.23 ± 0.009). However, the misfit of the ML11 parameterization is large. As can be seen in Fig. 7, the seasonal cycle of oxygen fluxes predicted by this parameterization is out of phase with that in the observations.

The reason why ML11 is ill suited for the MERL system is that its oxygen concentrations are typically well above 100 mmol O_2 m^{-3} and thus in the range where one would not expect oxygen limitation of the sediment oxygen consumption rate. Furthermore, sediment oxygen uptake rates in the mesocosms are smallest when oxygen concentrations are highest (Fig. 12), which is a direct consequence of the seasonal variations described above (in summer carbon remineralization rates are highest while oxygen solubility and thus dissolved oxygen concentrations are lowest).

The HD08 parameterization produces a much better fit simply because of its temperature dependence. The parameter controlling the oxygen dependence is inconsequential and essentially unconstrained by the observations (Fig. 12).

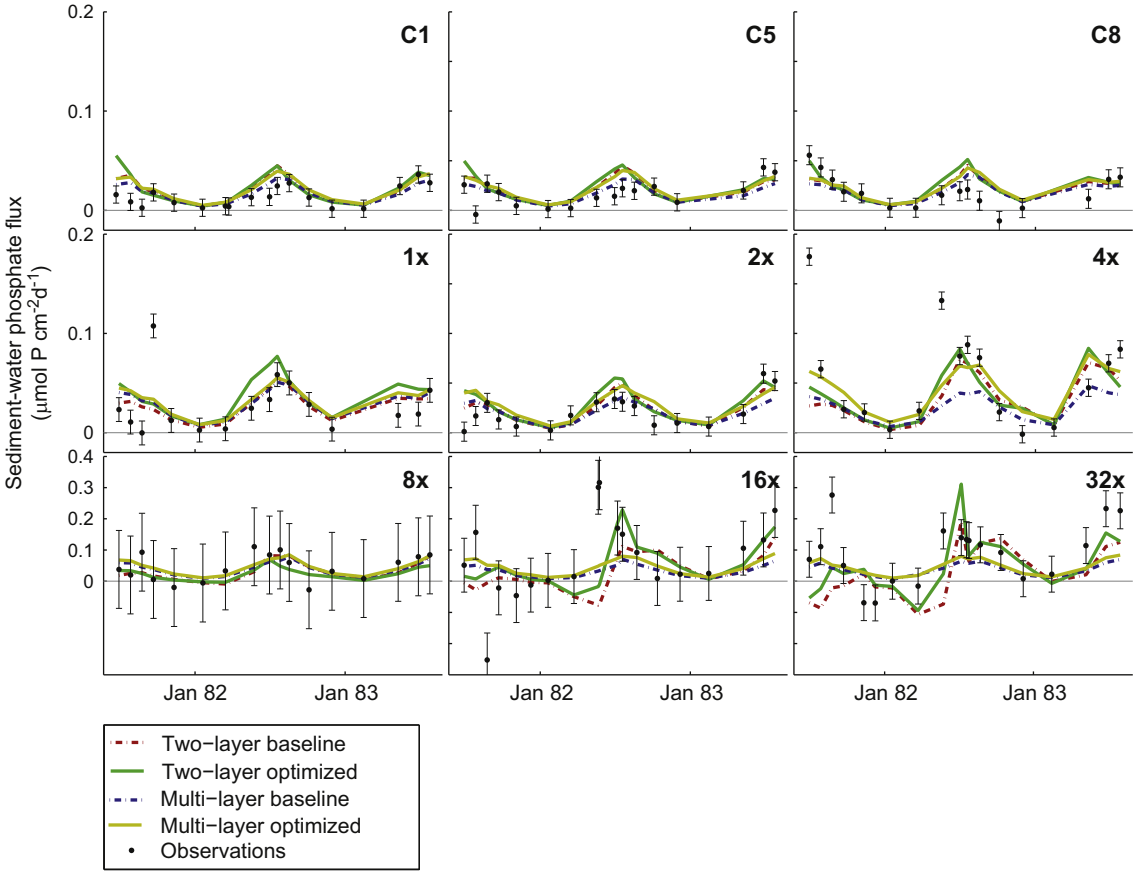

Fig. 9. Like Fig. 8 but for phosphate flux.

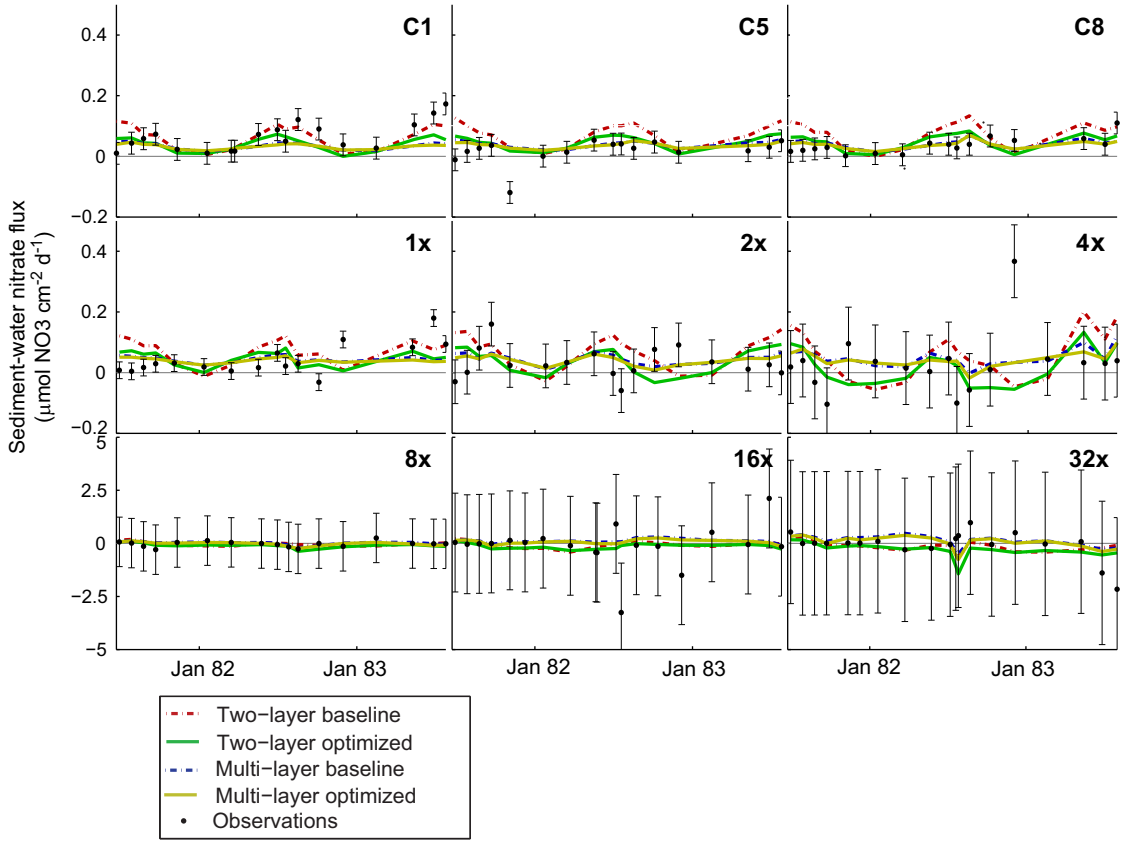

Fig. 10. Like Fig. 8 but for nitrate flux.

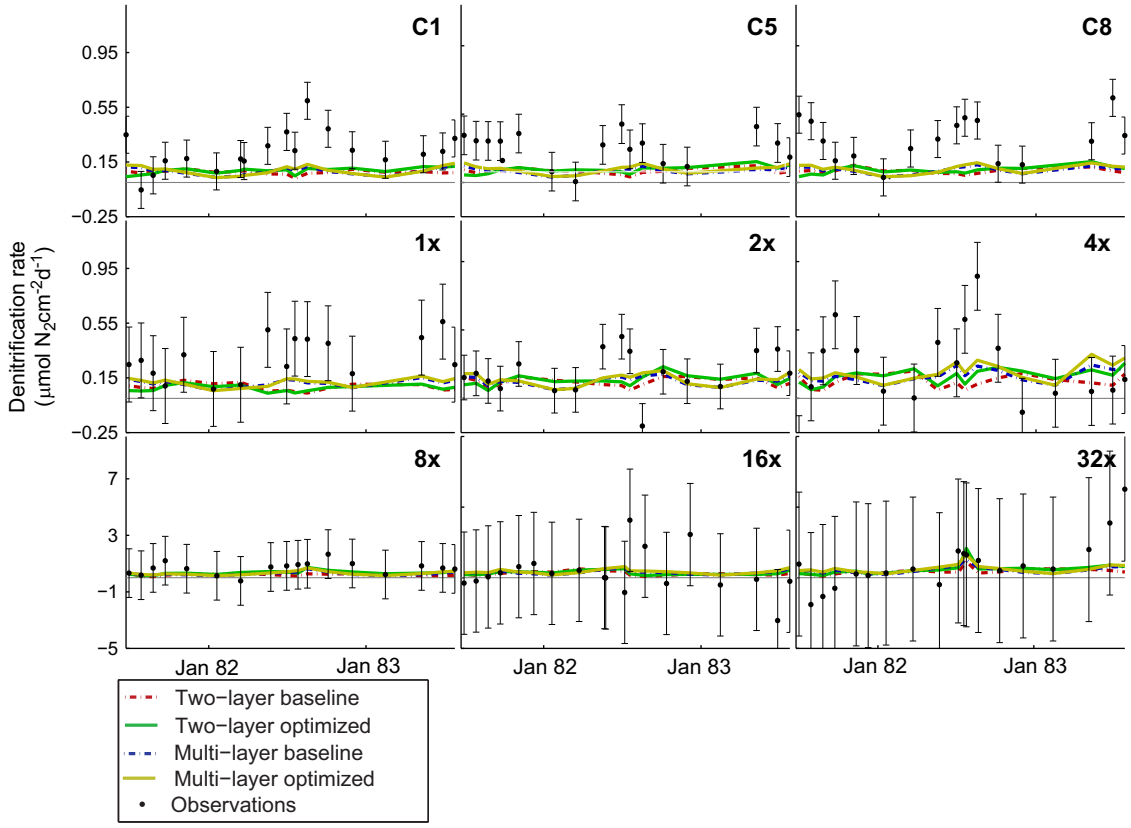

Fig. 11. Denitrification rates of the baseline and optimized two-layer and multi-layer models and rate estimates based on the observations (as described in Section 2.3). The simulated rates were integrated vertically (from sediment surface to 40 cm depth).

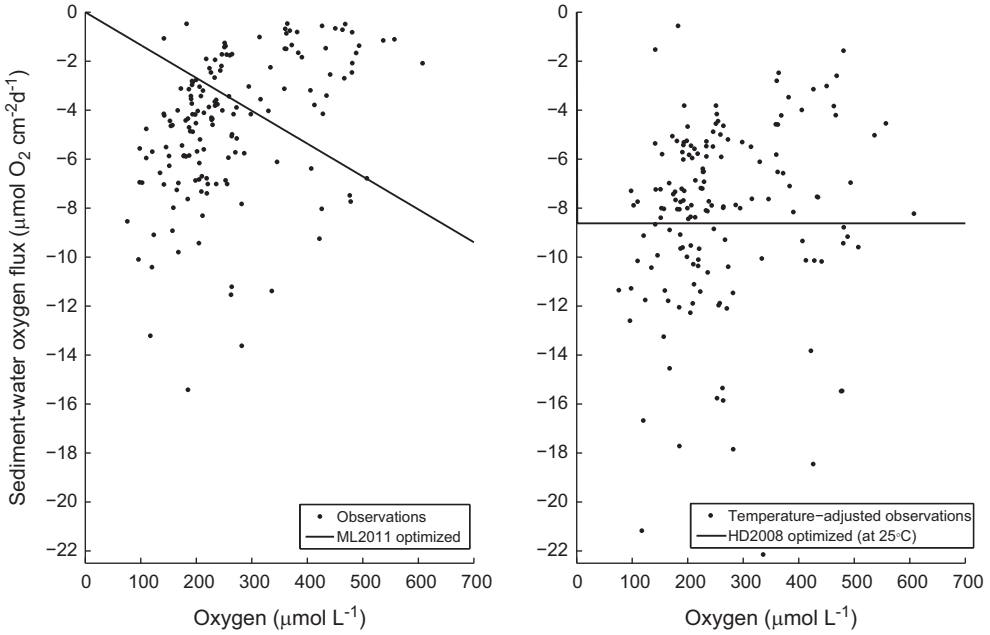

Fig. 12. Optimal sediment oxygen flux parameterizations in comparison with observations. The temperature-independent ML11 parameterization is shown in the left panel; the temperature-dependent HD08 parameterization is shown for 25 °C along with observed fluxes that were adjusted to the same temperature.

3.4. Cross-validation

Cross-validation, an important step in model optimization, aims to asses how generally applicable an optimized model is and to ensure that the reduction in misfit between model and observations is not a result of over-fitting to meaningless variation. Ideally, optimizing a model to one set of observations will improve its predictive power such that when applied to a new, independent set of observations the model predictions are improved (when compared to the unoptimized model), or at least that the optimization had no negative effect.

In the cross-validation experiments carried out here, both sediment models were fit to a randomly chosen subset of five mesocosms. Then the resulting parameter set was used within all

mesocosms and the resulting cost function value calculated. Ten such optimizations were performed and the resulting cost function values were averaged. The multi-layer model generated the smaller average total cost of 0.72 ± 0.03, while the two-layer model averaged at 1.32 ± 0.04, suggesting that the optimization of the latter resulted in over-fitting.

A similar cross-validation experiment was performed for the HD08 parameterization, resulting in an averaged oxygen flux cost of 0.31 ± 0.01. In comparison, the cross-validation experiments for the layered models generated oxygen flux cost contributions of 0.18 ± 0.01 for the multi-layer model and 0.33 ± 0.01 for the two-layer model.

4. Discussion

4.1. Parameterizations of depositional flux

Deposition of organic matter to the sediments is one of the major processes connecting water column and sediment biogeochemistry. The deposition of organic matter through sinking is an important input variable for the models optimized in this study driving many if not all of the diagenetic processes represented. Unfortunately, the depositional flux is difficult to observe directly (no direct observations are available for the MERL system) and had to be parameterized for this study.

Since the depositional flux of POM is ultimately derived from biomass within the water column, it is reasonable to expect that the former scales with the latter. Following this line of reasoning, we formulated flux parameterizations that generate a depositional flux in proportion to proxies of water column biomass, specifically, chlorophyll a, diatom abundance and/or zooplankton concentrations (methods B, C and D). These parameterizations are consistent with the treatment of sinking in conventional water-column biological models, which use similar linear or quadratic relationships, and with observations that benthic microbial activity is stimulated immediately by the input of sinking phytoplankton during the spring bloom, e.g., by Graf et al. (1982) in the Baltic Sea. By positing a linear relationship between depositional flux and water-column biomass, two outcomes were expected: first, since water column biomass grows with the availability of inorganic nutrients, the largest depositional fluxes should occur in the more eutrophic mesocosms; second, seasonal variation in water-column productivity should result in seasonal variations of the depositional flux.

As expected, when optimized jointly the parameterizations B, C and D roughly scale with nutrient load across the eutrophication gradient. An improved fit was achieved in individual optimizations where these same methods assigned a different parameter set for each mesocosm (i.e. when individually implemented); however, the resulting depositional fluxes were far less sensitive to nutrient loading. Method A assigned a constant depositional flux to each mesocosm and generated a similar eutrophication-insensitive result. Method A also resulted in the best fit overall. Thus the parameterization that best fit the data was the least sensitive to nutrient loading and exhibits no seasonal variation whatsoever. While this was contrary to our expectation, other studies, e.g. Laursen and Seitzinger (2002) when measuring benthic fluxes in the shallow Mid-Atlantic Bight, have found no pronounced seasonal variations either.

Neither of these divergences from expectation appears to be the result of poorly tuned sediment model parameters (when we included parameters in the optimization process, we found the same result). It is also unlikely that the results are due to a systematic bias in either of the models, because both the two-layer model and the multi-layer model generated a similar result (although the possibility that both models share the same bias hasn't been eliminated).

It is worth noting that sediment was intentionally resuspended in the mesocosms by movement of the integrated stirring plungers (see Section 2.2), and that this removal process was not simulated in either model. In effect the resuspension returns organic matter to the water column where it may remineralize outside of the sediment. If higher nutrient loads indeed induced larger depositional fluxes, this may have been counteracted by larger resuspension of easily erodible organic carbon from the sediment. Thus, for a more eutrophic mesocosm, proportionally less carbon would have been remineralized in the sediments.

4.2. Optimized model results

The multi-layer model produced smaller model-data misfits than the two-layer model, possibly due to its higher spatial resolution and fewer parameter redundancies. Another difference between both models is that fluxes simulated by the multi-layer model are not necessarily in steady state, while the two-layer model solves for steady-state fluxes at each time step. However, we tested how far the multi-layer model fluxes differ from their steady-state fluxes and found the differences to be negligible. Since the pattern of cost contribution by observation type and mesocosm are similar for both sediment models, it is difficult to identify a specific reason for why the multi-layer model did better.

There are reasons to expect that the models might behave similarly as both simulate the same chemical processes (albeit sometimes under different labels). The two-layer model simulates the anaerobic sulphate metabolism, and the oxidation of H_2S, while the multi-layer model simulates an analogous anaerobic metabolism, generating ODUs which can be oxidized in aerobic layers. Only the multi-layer model explicitly describes an aerobic metabolism, but this distinction should have little effect on the simulated sediment-water fluxes as both add any outfluxes of H_2S/ODUs to their oxygen influx estimates. Since this conversion consumes the same amount of oxygen per mole of decomposed carbon, it effectively erases any stoichiometric distinction between the aerobic and anaerobic remineralization pathways and directly links the oxygen flux to the rate of carbon remineralization.

Similarities between both models' cost contributions may also result from the stability of the oxygen concentrations in the mesocosms. Had low oxygen conditions occurred, this would likely have induced different non-linear responses in both models. For instance, the two-layer model represents oxygen sensitivity in denitrification rates by providing different denitrification parameters to the upper and lower layers. The multi-layer model, rather than using different parameter values in different layers, uses a universal parameterization to inhibit denitrification rates when oxygen is present.

None of the multi-layer model's denitrification parameters were included in the optimization, while two of the three two-layer model parameters were included. Despite this disadvantage, the multi-layer model's denitrification output was within error of that of the two-layer model. Given its finer vertical resolution the multi-layer model is better able to estimate the overlap between oxygen and nitrogen distributions for denitrification rate calculations. The two-layer model is less effective in determining the overlap between oxygen and nitrate since it resolves neither the concentration nor distribution of oxygen and instead has to rely on different rate parameters for the aerobic and anaerobic sediment layers.

Denitrification is often classified based on the source of nitrate as either “coupled” denitrification, which consumes nitrate that was produced by nitrification in the sediments, and “direct” denitrification, which consumes nitrate that has diffused into the

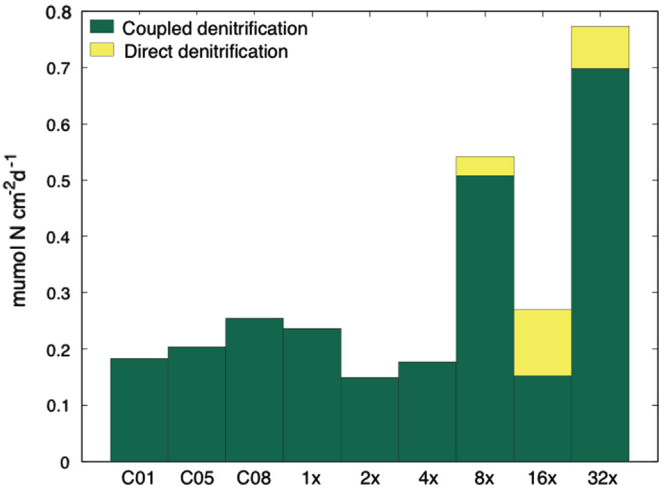

Fig. 13. Estimated denitrification rates in MERL mesocosms (rates were estimated as described in Section 2.3). Direct denitrification occurs only in the eutrophic mesocosms.

sediments from the overlying water. Since the treatment mesocosms received a constant supply of inorganic nitrogen, it is reasonable to expect more diffusion of nitrate into the sediments in the more eutrophic mesocosms, resulting in more direct denitrification. Observation-based estimates (Section 2.3) of coupled and direct denitrification are given in Fig. 13 and show that direct denitrification is only taking place in the three most eutrophic mesocosms, accounting for only about 20% of their total denitrification. In order for direct denitrification to take place in the models, nitrate needs to diffuse into the sediment. Both models produce the necessary influxes in the eutrophic mesocosms.

Interestingly, ammonium fluxes into the sediments were observed during winter in the eutrophic mesocosms. Only the multi-layer model was able to re-create a negative flux in the 16 × and 32 × mesocosms. In this case ammonium concentrations were greater in the overlying water column than in the sediment and ammonium diffused into the sediment as a result.

The production of inorganic phosphate is directly coupled to carbon remineralization. It is interesting, then, that the sediment models show pronounced differences in their carbon remineralization rates (Fig. 6), while the misfits of phosphate outfluxes are within error of one another (Fig. 9). Since phosphate is never consumed in the sediments, its chemical dynamics depends only on carbon remineralization and the partitioning of phosphate between the dissolved and solid phase. While both models have the ability to simulate oxygen-sensitive sorbtion of PO_4, only the two-layer model exhibits examples of this behavior in the 16 × and 32 × mesocosms (see Fig. 9). Since the eutrophic mesocosms have large uncertainties and contribute little to the cost function, the latter difference in model behaviors is of low significance overall.

4.3. Oxygen flux parameterizations

As discussed earlier, complexity is not always a desirable model characteristic, since over-fitting becomes a possibility and computational effort increases. Representing a simpler form of sediment model, two oxygen flux parameterizations were fit to the observations using non-linear regression and their resulting output compared to that of the layered models. The fit produced by these simple parameterizations was inferior to that of the multi-layer model, but not the two-layer model.

While there should obviously exist a relationship between oxygen flux into the sediment and oxygen concentration in overlying water, such that the flux decreases as oxygen approaches zero, the observations indicate an increase of sediment oxygen uptake with decreasing oxygen concentrations (Fig. 12). Correlation does not imply causation however, as this trend is the result of seasonality in the two signals. High overlying oxygen concentrations during winter and spring (driven by increased solubility and the spring diatom bloom) coincide with relatively lower carbon respiration rates and oxygen demand while the lowest overlying oxygen concentrations coincide with high oxygen demand during the summer. As a result, Murrell and Lehrter's (2011) parameterization (ML11), which assumes a positive linear correlation, produces oxygen fluxes that are out of phase with the observations (Fig. 7). The Hetland and DiMarco (2008) parameterization (HD08), which is dependent on temperature and oxygen, is much more successful, albeit entirely due to its ability to vary fluxes with temperature. The oxygen dependency contributed nothing to this parameterization's predictions, as the optimization process completely eliminated its oxygen sensitivity (Fig. 12), effectively reducing the HD08 parameterization to a single Arrhenius factor.

Interestingly, when considering the cost function error resulting from observational uncertainty, the oxygen fluxes produced by the HD08 parameterization are on par with the two-layer model, which along with the multi-layer model all share a very similar temperature-dependence in their calculation of carbon metabolism rates. It is worth noting that while the HD08 model was fit only to oxygen fluxes, the layered models were also fit to other observations, requiring some compromise of their oxygen fluxes for the sake of fitting other outputs. Furthermore, the time-invariant POM flux inputs may have nullified a significant advantage of the layered models, as they no longer were able to exploit the full dynamics of carbon storage, but instead accumulated and metabolized carbon according to temperature trends only (it is even possible that the temperature-dependence concealed the seasonality in depositional fluxes discussed above in Section 4.1). That the simple HD08 parameterization can generate oxygen fluxes on par with the two-layer model demonstrates that simple models can be valuable, given the right conditions. This model also demonstrates the critical importance of representing temperature sensitivity within the sediments, as the temperature-dependence alone is able to explain the majority of the variation within oxygen fluxes.

4.4. Cross-validation

Cross-validation is a way to assess a model's predictive capability by quantifying its ability to explain new observations without any adjustment of its parameters. This was done by optimizing the models in randomly chosen subsets of mesocosms and then using the resulting parameter set to run the models in all mesocosms. By comparing the average cost of different models one can gauge which is best suited to general application.

At the outset of our study we expected two outcomes of the cross-validation: (1) for all models, optimization leads to an improvement in predictive skill for the system considered; and (2) predictive skill of a model scales with its computational cost, in other words, the multi-layer model should do best and the simple parameterizations do worst. Our results are not consistent with these expectations.

While the multi-layer model proved the most versatile by maintaining the smallest average cost, the two-layer model did not stay below its baseline cost value of 1.0, scoring an average cost of 1.32 ± 0.05. This means that the un-optimized baseline version of the two-layer model was more generally applicable than the optimized version and indicates that the two-layer model is prone to over-fitting. Even more telling is that the two-layer model's oxygen flux results in a larger cost contribution than the HD08 oxygen flux parameterization in the cross-validation.

It is difficult to identify which aspect specifically leads to its apparent over-parameterization, but it is worth noting that the

two-layer model parameterizes a number of phenomena that the multi-layer model simulates more explicitly. For example, the two-layer model produces an oxygen-sensitive rate of denitrification simply by using different parameters for the aerobic and anaerobic layers, while the multi-layer model uses the same parameterization in all layers. In this respect the two-layer model is at a disadvantage since it uses more parameters and is less able to resolve the vertical distribution of oxygen.

5. Conclusions

In a quantitative comparison of different parameterizations of organic matter deposition the parameterization that prescribes a constant flux for each mesocosm performed better than any of the time-varying, biomass-dependent parameterizations tested. Contrary to expectation, the optimal constant depositional fluxes varied by less than a factor of 2 between the least eutrophic and the most eutrophic mesocosm, despite the fact that the latter received 32 times the nutrients of the former. Possibly higher deposition of organic matter in the more eutrophic mesocosms was partly counteracted by larger resuspension of freshly deposited material.

The mesocosm observations show a pronounced temperature dependence of sediment oxygen uptake (and by implication organic matter remineralization) and an increase in sediment oxygen uptake with decreasing oxygen concentration. Of the two parameterizations of sediment oxygen uptake tested, the one that includes temperature dependence produced a good fit, similar to one of the diagenetic models. Since the mesocosms were always well-oxygenated oxygen dependence of these parameterizations turned out to be a hindrance in fitting the observations. The parameterization that assumed dependence on oxygen only produced the poorest fit; the oxygen dependence in the other parameterization was minimized during optimization producing essentially a temperature-based parameterization.

Of the two diagenetic models analyzed, the more spatially resolved and computationally expensive model fit the observations better. The fit in oxygen fluxes is similar between the two-layer model and the simple temperature-dependent parameterization, while the multi-layer model produced a better fit.

In cross-validation experiments the multi-layer model displayed the best predictive skill while the two-layer model essentially failed the cross-validation; its ability to represent the observations degraded even in comparison to the simple parameterization and to its unoptimized baseline parameter set. This indicates that the two-layer model is prone to over-fitting, which will result in a degradation of the model's predictive ability.

When considering different sediment model candidates for inclusion in a biogeochemical model we recommend that the candidate models be evaluated against each other through a process of optimization and cross-validation as done here. Ideally the models should be optimized with and evaluated against an observational data set that represents the system of interest. (In practice this is easier said than done, since comprehensive sediment–water flux data sets are sparse.)

Further, we recommend that simple parameterizations be considered as a viable option for representing sediments in biogeochemical models. It was shown here that, contrary to expectation, the simple, temperature-dependent parameterization of sediment oxygen consumption has better predictive skill than the two-layered diagenetic model in describing oxygen consumption by the sediment. In terms of computational demands the former is trivial while the latter requires a substantial investment.

Lastly, we would like to caution model users against the temptation to judge models solely based on the biogeochemical processes that they represent. Even though both diagenetic models used here describe almost the same set of processes they do so using different parameterizations and assumptions (some, but not all, of these differences follow from their different spatial framework) and one of them (the multi-layer model) performs well in cross-validation while the other (the two-layer model) essentially fails. Even though two models may describe the same set of processes, different choices of how they are represented can lead to very different predictive abilities.

Acknowledgments

We would like to thank Damian Brady, Dominic DiToro, Jeremy Testa and Mike Kemp for many constructive discussions about the two-layer model. We also thank Bernie Boudreau and Paul Hill for comments on Robin Wilson's MSc thesis, which formed the basis for this manuscript. This work was supported by the NOAA grant NA07NOS4780191 and is NOAA CHRP contribution number 179.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.csr.2013.05.003.

References

Bagniewski, W., Fennel, K., Perry, M., D'Asaro, E., 2011. Optimizing models of the North Atlantic spring bloom using physical, chemical and bio-optical observations from a Lagrangian float. Biogeosciences 8, 1291–1307.

Berg, P., Risgaard-Petersen, N., Rysgaard, S., 1998. Interpretation of measured concentration profiles in sediment pore water. Limnology and Oceanography, 1500–1510.

Berner, R.A., 1980. Early Diagenesis: A Theoretical Approach. Princeton University Press.

Bohlen, L., Dale, A., Wallmann, K., 2012. Simple transfer functions for calculating benthic fixed nitrogen losses and C:N:P regeneration ratios in global biogeochemical models. Global Biogeochemical Cycle 26GB3029.

Burdige, D.J., 2006. Geochemistry of Marine Sediments. Princeton University Press.

Burdige, D.J., 2011. Estuarine and coastal sediments—coupled biogeochemical cycling. In: Wolanski, E., McLusky, D.S. (Eds.), Treatise on Estuarine and Coastal Science, vol. 5. Academic Press, Waltham, pp. 279–316.

Boudreau, B., 1997. Diagenetic Models and Their Implementation: Modelling Transport and Reactions in Aquatic Sediments. Springer, New York.

Boudreau, B., 2000. The mathematics of early diagenesis: from worms to waves. Reviews of Geophysics 38, 389–416.

Dale, A., Sommer, S., Bohlen, L., Treude, T., Bertics, V.J., Bange, H., Pfannkuche, O., Schorp, T., Mattsdotter, M.E.-K., Wallmann, K., 2011. Rates and regulation of nitrogen cycling in seasonally hypoxic sediments during winter (Boknis Eck, SW Baltic Sea): sensitivity to environmental variables. Estuarine, Coastal and Shelf Science 95, 14–28.

Dhakar, S., Burdige, D., 1996. A coupled, non-linear, steady state model for early diagenetic processes in pelagic sediments. American Journal of Science 296, 296–330.

DiToro, D., 2001. Sediment Flux Modeling. Wiley-Interscience.

Emerson, S., Jahnke, R., Heggie, D., 1984. Sediment–water exchange in shallow water estuarine sediments. Journal of Marine Research 42, 709–730.

Fennel, K., Follows, M., Falkowski, P.G., 2005. The co-evolution of the nitrogen, carbon and oxygen cycles in the Proterozoic Ocean. American Journal of Science 305, 526–545.

Fennel, K., Wilkin, J., Levin, J., Moisan, J., OReilly, J., Haidvogel, D., 2006. Nitrogen cycling in the Middle Atlantic Bight: results from a three-dimensional model and implications for the North Atlantic nitrogen budget. Global Biogeochemical Cycles 20, GB3007, http://dx.doi.org/10.1029/2005GB002456.

Fennel, K., Wilkin, J., Previdi, M., Najjar, R., 2008. Denitrification effects on air–sea CO_2 flux in the coastal ocean: simulations for the Northwest North Atlantic. Geophysical Research Letters 35, L24608, http://dx.doi.org/10.1029/2008GL036147.

Fennel, K., et al., 2009. Modeling denitrification in aquatic sediments. Biogeochemistry 93, 159–178.

Fennel, K., Hu, J., Laurent, A., Marta-Almeida, M., Hetland, R., Sensitivity of hypoxia predictions for the Northern Gulf of Mexico to sediment oxygen consumption and model nesting, Journal of Geophysical Research-Oceans, 118, 2013, http://dx.doi.org/10.1002/jgrc.20077.

Friedrichs, M., et al., 2007. Assessment of skill and portability in regional marine biogeochemical models: role of multiple planktonic groups. Journal of Geophysical Research 112, C08001.
Frithsen, J., A. Keller, M. Pilson, Effects of Inorganic Nutrient Additions in Coastal Areas: A Mesocosm Experiment Data Report. Volume 1. MERL Series, Technical Report. University of Rhode Island, 1985a.
Frithsen, J., A. Keller, M. Pilson, Effects of Inorganic Nutrient Additions in Coastal Areas: A Mesocosm Experiment Data Report. Volume 2. MERL Series, Technical Report. University of Rhode Island, 1985b.
Frithsen, J., A. Keller, M. Pilson, Effects of Inorganic Nutrient Additions in Coastal Areas: A Mesocosm Experiment Data Report. Volume 3. MERL Series, Technical Report. University of Rhode Island, 1985c.
Graf, G., Bengtsson, W., Diesner, U., Schulz, R., Theede, H., 1982. Benthic response to sedimentation of a spring phytoplankton bloom: process and budget. Marine Biology 67 (2), 201–208.
Hetland, R., DiMarco, S., 2008. How does the character of oxygen demand control the structure of hypoxia on the Texas–Louisiana continental shelf? Journal of Marine Systems 70, 49–62.
Katsev, S., Chaillou, G., Sunby, B., Mucci, A., 2007. Effects of progressive oxygen depletion on sediment diagenesis and fluxes: a model for the Lower St. Lawrence River Estuary. Limnology and Oceanography 52, 2555–2568.
Kelly, J., Berounsky, V., Nixon, S., Oviatt, C., 1985. Benthic–pelagic coupling and nutrient cycling across an experimental eutrophication gradient. Marine Ecology Progress Series 26, 207–219.
Kelly-Gerreyn, B.A., Hydes, D.J., Trimmer, M., Nedwell, D.B., 1999. Calibration of an early diagenesis model for high nitrate, low reactive sediments in a temperate latitude estuary (Great Ouse, UK). Marine Ecology Progress Series 177, 37–50.
Kemp, W., Sampou, P., Garber, J., Tuttle, J., Boynton, W., 1992. Seasonal depletion of oxygen from bottom waters of Chesapeake Bay: roles of benthic and planktonic respiration and physical exchange processes. Marine Ecology Progress Series. Oldendorf 85, 137–152.
Laursen, A.E., Seitzinger, S.P., 2002. The role of denitrification in nitrogen removal and carbon mineralization in Mid-Atlantic Bight sediments. Continental Shelf Research 22, 1397–1416.
Mattern, J. P., Ensemble-based Data Assimilation for a Physical–biological Ocean Model Near Bermuda, Master's Thesis. Universität zu Lübeck, 2008.
Middelburg, J., Soetaert, K., Herman, P., Heip, C., 1996, 661-673.. Denitrification in marine sediments: a model study. Global Biogeochemical Cycles 10.
Murrell, M., Lehrter, J., 2011. Sediment and lower water column oxygen consumption in the seasonally hypoxic region of the Louisiana continental Shelf. Estuaries and Coasts 34, 912–924.
Nixon, S.W., Pilson, M.E.Q., Oviatt, C.A., Donaghay, P., Sullivan, B., Seitzinger, S., Rudnick, D., Frithsen, J., 1984. Eutrophication along a coastal marine eutrophication gradient—an experimental study using the MERL mesocosms. In: Fasham, M.J.R. (Ed.), Flows of Energy and Materials in Marine Ecosystems. Plenum Press, pp. 105–135.
Oviatt, C., Keller, A., Sampou, P., Beatty, L., 1986. Patterns of productivity during eutrophication: a mesocosm experiment. Marine Ecology Progress Series 28, 69–80.
Peña, M.A., Katsev, S., Oguz, T., Gilbert, D., 2010. Modeling dissolved oxygen dynamics and hypoxia. Biogeosciences 7, 933–957.
Rao, A.M.F., McCarthy, M.J., Gardner, W.S., Jahnke, R.A., 2007. Respiration and denitrification in permeable continental shelf deposits on the South Atlantic Bight: rates of carbon and nitrogen cycling from sediment column experiments. Continental Shelf Research 27, 1801–1819.
Seitzinger, S., Giblin, A., 1996. Estimating denitrification in North Atlantic continental shelf sediments. Biogeochemistry 35, 235–260.
Seitzinger, S., Harrison, J., Böhlke, J., Bouwman, A., Lowrance, R., Peterson, B., To- bias, C., Drecht, G., 2006. Denitrification across landscapes and waterscapes: a synthesis. Ecological Applications 16, 2064–2090.
Soetaert, K., Herman, P., Middelburg, J., 1996a. A model of early diagenetic processes from the shelf to abyssal depths. Geochimica et Cosmochimica Acta 60, 1019–1040.
Soetaert, K., Herman, P., Middelburg, J., 1996b. Dynamic response of deep-sea sediments to seasonal variations: a model. Limnology and Oceanography 41, 1651–1668.
Soetaert, K., Middelburg, J., Herman, P., Buis, K., 2000. On the coupling of benthic and pelagic biogeochemical models. Earth-Science Reviews 51, 173–201.
Wilson, R.F., Comparative Assessment of Two Diagenetic Models, M.Sc. Thesis. Dalhousie University, ⟨http://hdl.handle.net/10222/14386⟩, 2011.
Wood, C.C., Martin, A.P., Statham, P.J., Kelly-Gerreyn, B.A., Modelling macro nutrients in shelf sea sediments: fitting model output to experimental data using a genetic algorithm, Journal of Soils and Sediments in review.