Below I've replied to your comments, and outlined where the manuscript text has been changed.

Experimental part:

The experiment was either not well designed or not clearly described. First, why a 5-cm core thickness was chosen? Did this influence the experiment result (leading to the depletion of Fe(III) and near zero Alk production)? How well does this simulate natural processes? The result shows that DIC flux was nearly constant (except when OM was added on day 70), indicating a similar overall lability of the OM or a constant OM decomposition rate over the experiment period. Thus, it appears what was depleted was Fe(III).

This study follows up observations from the field Cook 2004 – the aim was not to simulate specific conditions, but investigate the physical explanation for the phenomenon. We have now indicated in the text that this "was chosen to represent the diagenetically active depth.." The agreement with the previous field data shows that this does not alter the outcome of the experiment.

Second, the analytical method for Fe analysis was mentioned in the method section (p.5449), but the Fe (III or total?) content of the sediments was never reported (or that was buried somewhere) nor the [Fe(III)] change over the experiment period. Later, in p.5453, line 22-24, the authors speculated without any data support—"The most likely explanation for this observation is that there was an initial formation of a large pool of iron sulfides. Initially, a large pool of Fe(III) is formed during sediment pretreatment, which is subsequently gradually depleted and turned into metal sulfides."

Unfortunately we don't have an initial measurement of Fe other than an observed colour change from black to brown as it was reoxidised, and back to black as it was reduced. It is well established in the literature that Fe(III) will be produced under the initial conditions we subjected the sediment to, so in the absence of this data we rely on our observations and well established biogeochemical concepts to explain what is happening.

Third, I initially didn't know it was a 6- and 6-hour cycle but figure it out later in the model part (in p.5452, line 10). Maybe an illustration graph of how the experiment was setup and run will help. It appears to me that the authors have little interest to presenting the experiment.

We have included a figure to describe the experimental setup. The 6 hour cycle is described more explicitly in section 2.1.

Model part:

understand that this paper applies an existing diagenetic model that comes from a leading modeling group. However, as it deals with the CO2-carbonate system, I am still shocked to see there is no description on how the authors dealt with the production/consumption and diffusion of various inorganic carbon species and other acid-base species aspects of the sediment diagenetic processes. This weakness was reflected not only in the description of the model but also in how the model was set up and what parameters are need to be measured in the experiment. For example, what are exactly included in their Alk definition? Were HPO42- and DOM considered as acid-base species in the Alk equation (see p. 442 in Cai et al. 2010)?

DOM and HPO4 have been shown to contribute to alk fluxes, but reactions involving S (and it interactions with Fe) are the major processes controlling the alkalinity balance within the sediment, and hence these are the only processes we considered. We justify this on the basis that these are the major alkalinity producing and consuming species. Although previous models of TA have included HPO_4^{2-} and DOM, only these were included as they were deemed to be the most important alkalinity producing and consuming reactions. The good agreement between the model, lab experiments, and previous field observations support this assumption. We have mentioned these species in the discussion:

"Dissolved organic matter and HPO_4^{2-} have been shown to contribute to alkalinity fluxes (Cai et al., 2010), but reactions involving sulfur (and it interactions with iron) are the major processes controlling the alkalinity balance within the sediment, and hence only these processes are considered."

We have now discussed the acid-base species clearly in the methods section:

"The set of acid-base reactions governing the pH dynamics in the pore water includes the carbonate, borate, sulfide and water equilibria. The impact of dissociation reactions involving phosphate, ammonium, silicate and dissolved organic compounds was assumed negligible. The associated total alkalinity and total species are hence defined as:

$$A_T = \left[HCO_3^- \right] + 2\left[CO_3^{2-} \right] + \left[B(OH)_4^- \right] + \left[HS^- \right] + \left[OH^- \right] - \left[H^+ \right]$$

$$SumCO2 = \left[CO_2 \right] + \left[HCO_3^- \right] + \left[CO_3^{2-} \right]$$

$$SumBOH3 = \left[B(OH)_3 \right] + \left[B(OH)_4^- \right]$$

$$SumH2S = \left[HS^- \right] + \left[H_2S \right]$$

The associated equilibrium constants were calculated as a function of temperature and salinity using AquaEnv, a dedicated R-package for acid-base and CO2 system calculations (Hofmann et al., 2010). Specifically, for the carbonate equilibria, we used the relationships provided by Millero et al. (2006)."

The second thing I am not happy with the paper is there is no information on species profiles and production/consumption rates that the model generated internally. When we read such an experiment-model comparison paper, the readers not only are interested in knowing the result (CO2 flux during exposure < DIC flux during inundation), we also want to know how the CO2- DIC-Alk fluxes are facilitated. What are the reduced species that are oxidized (at what depths) that also produces proton and become an Alk sink? For an example, in most sediment pore waters, TA and DIC values/concentrations are similarly elevated (as both are dominated by HCO3 -) in a few examples that both were measured (see for example Cai et al. 2006), why DIC flux appears so much higher than the Alk flux? In an early paper (Cai et al. 2000), I have shown that within the sediment, inorganic carbon diffusion is still facilitated by HCO3 - as it is the dominated species even at a fairly low pH. However, once HCO3 - diffuses to the

We have now shown profiles of CO_2 HCO_3^-. CO_3^{2-}, pH, HS^- and Fe(III) under high and low Fe conditions. We have included some discussion on the relevance of these figures with regard to how Fe is influencing pH, and alkalinity flux, and CO_2 flux out of the sediment. controlling alkalinity. The relevant text from the discussion is quoted below:

"For simulations with little or no Fe, there is a higher temporary sink of HS^- in the sediment pore waters during intense alkalinity production (Fig. 7), as insoluble iron sulfides cannot form. Subsequent alkalinity consumption occurs as the HS^- is re-oxidised. The effect of the reoxidation of reduced sulfur can be seen in Figure 7, which shows the simulation with no FeOOH having lower pH, and thus higher CO_2 concentration, than the simulation with 95 μmol Fe g^{-1} dry sediment. The alkalinity concentration (as indicated by HCO_3^- and CO_3^{2-}) throughout the FeOOH containing profile is also considerably higher."

In our experiments, denitrification was unlikely to be a significant source of alkalinity generation because of the lack of NO_3^- in the overlying water of our incubations. Consequently, we did not model the effect of denitrification on alkalinity in this model. Coupled nitrification-denitrification would not result in significant net alkalinity generation (Hu and Cai, 2011) We have described this in the text.

The study by (Cook et al. 2009) was on carbon flows through bacteria and microphytobenthos. There was no explicit study of the carbonate system and it is not relevant to this work.

The authors think that there was initial a large pool of Fe(III), which was consumed and resulted in negligible Alk production by days 60 days (p.5453, line 23). Then, why Alk production increases after OM addition at day 70 without adding reactive Fe(III)? Need to explain this.

We have included some profiles (FeOOH profile before and after organic matter pulse) at your suggestion which we feel clarifies this phenomenon. We agree that this will help our discussion significantly. The added text regarding this is quoted below:

" The profile of FeOOH in the top 0.5 cm of sediment changed dramatically upon the addition of the organic matter pulse to the top 0.2 cm (Fig. 5). The profile taken two tidal cycles before the pulse shows the FeOOH reaching ~15 µmol Fe g^{-1} dry sediment in the surface layer, and reducing with increasing depth. One tidal cycle after the addition of organic matter, the FeOOH profile is oxidising, with the concentration in the top 0.2 cm falling to a maximum of 2 µmol Fe g^{-1} dry sediment in the surface layer. Below 0.2 cm, the FeOOH profile is similar between the two times, as the organic matter concentration has not changed in this domain.

"The reduction of Fe in the surface of the sediment leads to a transient alkalinity flux and associated increase in the discrepancy between CO_2 flux and DIC flux (Fig. 5). This observation is seen in both the model and the laboratory incubations, although there is a slight offset due to the fact that there is a delay in the maximum respiration rate in the laboratory incubations compared to the model."

Specific comments:

p.5446, Line 24-25, the argument that DIC is hard to measure than CO2 is odd as one can argue the other way. For example, though both DIC and CO2 are measured in an incubation simulation system (core covered by water or air), it is a lot easier to simulate water flow (with reasonable stir) and measure the right DIC flux than to simulate air flow and measure the correct CO2 flux. It's good to know that most people study tidal mudflats measures the CO2 gaseous flux not DIC, but I can also give you one opposite example from Sapelo Island marshes (see Table 2 in Cai et al. 1999, L&O).

We have included a citation of your example of DIC measurements.

p.5447, line 15-16, not sure why bio-irrigation cease upon emersion (submerge the sediment core under water).

We did not include Bioturbation in these simulations, as we sieved the sediment anyway. (We have also edited previous statements which suggested that we included bioturbation.)

p.5447, line 20-28, what about acidification via oxidation of reduced species? This can decrease Alk and increase CO2 at the sediment surface.

p.5447, line 20-28, We have shown and discussed this in the profiles we've added (discussed above)

p.5448, line 11, change sediment to sediments; line 12 delete sediment and change was to were.

We have fixed the English problem here.

Discussed earlier.

Throughout the text we have replaced emersion with exposure for clarity. This will clarify hopefully any confusion. Also, an illustration of the experiment is added as figure 1.

We have added these details:
"They were stored in glass vials and refrigerated at 4 °C for less than one week before analysis. The concentration was determined using a modified Gran titration (Almgren et al., 1983). pH was determined at 20 °C using a pH electrode (Hach PHC301 connected to an HQ40d meter), calibrated with NBS buffers."

Discussed previously.

We did not include Bioturbation in these simulations, as we sieved the sediment anyway. (We have also edited previous statements which suggested that we included bioturbation.)

When k>1, CO_2 concentration is the most important factor in controlling flux. PV wasn't determined in the cores. The point is that CO_2 exchange will never be limited by transport, rather the production rate.

We feel 'extreme' is probably a fine word to describe what we mean, but we have clarified the sentence by changing it to "the extreme".

Covered by new figure of profiles and their discussion (as discussed earlier)

We did not consider nitrogen containing species in this model as we assumed their effect would be negligible when compared to the processes we have modelled (see earlier response)

Agreed, we have now explicitly stated in the text that when k>1, CO_2 concentration is the most important factor in controlling flux.(see section 4.2 on new manuscript)

The model run showed transient spikes last ~30 min after inundation, and we took the average to simulate our measurements. We have clarified this in the text.

As discussed earlier.

Thank you for your thoughtful comments,

P. Faber and co-authors