Answer to reviewer #1

We wish to thank the reviewer for the critical remarks brought to the diatom database paper presented for ESSD. We agree with the remark on Table 4 that :

1. It is a very rough estimation and reflects all the listed biases discussed in the paper
2. coastal and open ocean data could have been seperated to investigate whether estimates with the coastal areas are in any ways different than open ocean alone data
3. Integrated BSi data available for many ocean provinces could have been compared for further quantitative comparison in this paper. Originally, the focus was on C biomass, but indeed adding derived BSi calculation and confronting them to actual measurements should be a valuable addition to the paper.

Our response is as follows :

Using a Matalb routine, each geographical point was associated to a bottom depth. Then, datapoints associated to depths<100 m were considered "coastal" data, while the rest were considered "open ocean" data. As the resolution of the grid is quite coarse, some approximation were made in the bottom depth/lat/long association, which were visually corrected by replacing all sampling site on a global map indicating topography. 58 stations which were wrongly attributed to the "open ocean" were reattributed to the pool of coastal ocean data points, as they clearly were above the 100 m isobath.

The database thus comprises 3826 open ocean sites and 552 coastal ocean sites. Then the same calculations were done as previously, but were run first on the entire dataset, then on the data set consisting of open ocean data alone, as running the 3 X 3 binning procedure on coastal data alone would not be adequate (too little spatial coverage). The difference between the two datasets thus reflects the weight of coastal data. The conversion factor used to convert C biomass to Si biomass has been averaged to Si:C=0.093, considering the average between Si-stressed diatoms (0.056, DeLaRocha et al., 2010) and Si-replete diatoms (0.130, Brzezinski et al., 2011).

Three new tables were hence generated. Table 4 now lists all mean integrated BSi data (0-200m) for each ocean (geometric and arithmetic means) and puts these data in perspective with other regional data obtained from direct measurements from the literature as suggested by the reviewer. This table shows that a high degree of variability in the literature integrated BSi data can be obtained and that they are most of the time much larger than the 3-4 mmol m^{-2} for open ocean estimate. However, upon recalculation following the indications given above, we obtain for mean geometric ΣBSi over 200 m) – recap of Table 4 for our study :

	All data	Open ocean data only	Δwithout coastal data
Arctic	4.6 mmol m^{-2}	9.9 mmol m^{-2}	+215%
Atlantic	3.4 mmol m^{-2}	3.3 mmol m^{-2}	-3%
Pacific	8.0 mmol m^{-2}	7.1 mmol m^{-2}	-11%
Indian	29.1 mmol m^{-2}	26.9 mmol m^{-2}	-7%
Southern O.	4.0 mmol m^{-2}	4.4 mmol m^{-2}	+10%

These results show :

1. That mean ΣBSi for all data are not that different than the range of 3-4 mmol m^{-2} cited by the reviewer for open ocean, 10-50 mmol m^{-2} for Sargasso Sea bloom events or 6-10 mmol m^{-2} for the Pacific. Data calculated here for the Arctic, Atlantic, Pacific and Southern Ocean combined range between 3.3 and 9.9 mmol m^{-2} which seems consistent with the reviewer's estimate. Only the Indian Ocean exhibits a higher value of 26.9 mmol m^{-2}, probably skewed by very high biomass data in the Arabian Sea, and particularly on the Kerguelen Plateau, which is included in the Indian Ocean. Another estimate of 2 to 26 mmol Si m^{-2} is given for HNLC and oligotrophic regions in Adjou et al. (2011), which is also consistent with the values obtained here.
2. Surprisingly, the coastal data do not weigh too much in the global budgets, since taking coastal data points out only induces a reduction of -3 to -11 % in the mean ΣBSi for the Atlantic, Pacific and Indian Ocean, depending on the ocean considered. The only notable exception is the Arctic, where looking at open ocean data alone without the coastal points actually <u>increases</u> the mean ΣBSi by 215%. Since the coastal shelf North of Russia is quite broad, it seems that there are major differrences in biomass between very coastal points and more offshore data. The effects of fresh water inputs from the many rivers on the continental shelf might explain the lowest diatom biomass there. The Southern Ocean shows a moderate increase of 10% in Si biomass when removing coastal data points, but the ice coverage limit, different from the coastal bathymetry is a likely cause for this increase.

We have compiled these results in 3 new tables (splitting the information contained in the previous Table 4) : Table 4 now indicates the Si budget per oceanic province and compares these estimates with literature data. Table 5 and 6 give the global ocean estimates for all data and open ocean data alone, with global C biomass, global Si biomass and Si turnover rates estimates, which range between 0.15 and 0.19 d^{-1} (considering the geometric mean).

Finally, an error in the TABASCO cruise coordinates, initially located in the Indian Ocean, but which should have been in the Southern Pacific, has been identified and corrected in the database (a new version has been resubmitted), and Figures 2, 4, 5 as well as Table 4 have been modified accordingly.

A new version of the manuscript taking these modifications into account (database, Figures, Tables and last paragraph of the discussion) has been resubmitted.