

## Overview and general comments:

The manuscript of Moras et al. reports results from a laboratory study on magnesium hydroxide additions to seawater, examining the effects of mineral grain size, seawater salinity, and dissolved organic carbon concentrations on the dissolution of magnesium hydroxide and secondary calcium carbonate precipitation. Their results demonstrated that secondary CaCO<sub>3</sub> precipitation was delayed the longest when using medium grain size Mg(OH)<sub>2</sub>, and they also found that secondary precipitation was accelerated at lower concentrations of Mg and DOC, both of which are known CaCO<sub>3</sub> precipitation inhibitors. This is an important study with implications for optimizing the efficacy of ocean alkalinity enhancement.

The authors' conclusions regarding the effects of grain size, Mg, and DOC on delaying secondary precipitation are generally supported by the data, and the manuscript was straightforward to follow. Some major issues that need to be addressed include providing more details on the carbonate chemistry calculations (including uncertainty estimates—especially for  $\Omega$ , which is an indicator for the secondary precipitation threshold), clearly distinguishing between calculated and measured CO<sub>2</sub> parameters, and the unexplained DIC trends in some of the data (which the authors attribute to ingassing, but does not seem to be well-supported by the data). Additionally, the presentation of the some of the figures and tables can be improved. With minor revisions, I recommend publication.

## Major Comments:

### Carbonate Chemistry Calculations:

The authors need to provide full details on the carbonate chemistry calculations, such as the choice of constants used in the calculations (Lines 145-146). Additionally, uncertainties should be estimated and reported for calculated CO<sub>2</sub> parameters using the uncertainty propagation routines available on CO<sub>2</sub>SYS. The uncertainty estimates are especially important when reporting values of  $\Omega$ , which undoubtedly will inform future research on critical thresholds for CaCO<sub>3</sub> precipitation. Without uncertainty estimates, the results in this study cannot be interpreted. In Table 1, the calculation uncertainty for  $\Omega$  should be combined in quadrature with the standard deviation of the replicates.

Thank you for the comment. We indeed forgot to mention the constants used. This has been added in lines 154-156 (of the **TRACK CHANGES** document). Similarly, the uncertainties for calculated CO<sub>2</sub>SYS parameters have been added to the graphs and in table 1.

#### Calculated vs. Measured CO<sub>2</sub> Parameters:

The text and figures are often confusing as to which CO<sub>2</sub> parameters were measured and calculated. Furthermore, the authors do not always specify which input parameters were used in the calculations. In Lines 127-135, it appears that most of the DIC and TA values in the Mg and DOC experiments were calculated values (unlike in the salinity and grain size experiments where each experiment had 9 to 10 DIC and TA measurements). Fig. 5, however, does not make it clear which of the  $\Delta$ DIC and  $\Delta$ TA values were measured and calculated (and the input parameters used). Similarly, in Line 133, “the difference between estimated maximum TA and final measured TA” does not specify how TA was calculated.

Thank you for the comment. In the caption of Figure 5, the term “reconstructed” will be replaced to “calculated”. Regarding the second part of the comment the information is given at line 136, where we specify that **“a sample for TA and DIC was also taken before Mg(OH)<sub>2</sub> addition and at the end of each treatment”**, suggesting that these TA and DIC data have been measured the same way than other TA and DIC samples, and used for the reconstruction.

#### pH measurements on the “free” scale:

The authors report pH measurements on the “free” scale based on pH electrode measurements calibrated with Metrohm buffer solutions (pH 4, 7, and 9). As the calibration solutions are low-ionic strength solutions, there will be large liquid junction potential errors when making measurements in seawater solutions at high ionic strength, and it is difficult to estimate the uncertainty. While the electrode measurements may still be useful in reporting changes in pH, it seems problematic to use the data to calculate other CO<sub>2</sub> parameters in seawater. One solution may be to calibrate the electrode measurements against pH calculated from measured DIC and TA whenever these measurements coincide. The authors should also make clear which parameters were calculated with pH and how the uncertainty in the pH measurements may affect the interpretation of their results.

Thank you for the comment. To perform accurate CO<sub>2</sub>SYS calculations, the scale on which the pH is measured needs to be known. Based on the buffer used, the pH probe was calibrated on the free scale. This was further confirmed by comparing the measured pH with estimated pH<sub>F</sub> from CO<sub>2</sub>SYS using measured TA and DIC. After reviewing all data, the measured pH was identical to the calculated pH<sub>F</sub> from CO<sub>2</sub>SYS. Therefore, this allowed us to report the pH values as pH<sub>F</sub>. Regarding the difference between calculated and measured parameters, this information is given in paragraphs 2.3 and 2.4. All samples for TA and DIC were measured, except when clearly stated as for the Mg and DOC experiments (lines 137 to 143).

Maximum  $\Omega_A$  attained (Line 288): The authors attribute differences in the maximum  $\Omega_A$  attained in their  $Mg(OH)_2$  experiments and previous studies with  $Ca(OH)_2$  to differences in the starting water composition. It is difficult for the reader to confirm this without information provided on the starting water composition. I suggest visualizing the results of this and previous studies with  $Ca(OH)_2$  on a DIC and TA diagram with  $\Omega_A$  isolines, showing the initial water composition and final composition at maximum  $\Omega_A$ .

Thank you for pointing this out, we agree that this can be hard for the reviewer. We have reviewed the text, and it now reads (lines 299-303): **“Such a difference is likely due to the difference in the starting conditions and experimental settings. In the experiments shown here, the starting  $\Omega_A$  was ~2.8 while it was about ~2.5 in Moras et al. (2022). This is explained by the difference in the starting water composition and salinity, ultimately affecting the final  $\Omega_A$  despite similar TA increases. Furthermore, higher amounts of  $Mg(OH)_2$  were added compared to Moras et al. (2022), leading to a higher  $\Omega_A$  and a higher theoretical  $\Delta TA$ , if no early  $CaCO_3$  nucleation occurred.”**

Secondary  $CaCO_3$  precipitation thresholds:

The authors chose  $\Omega_A \sim 9$  based on their hypothesis that the secondary precipitation threshold would be higher when adding  $Mg(OH)_2$  compared to the threshold when adding  $CaO$  and  $Ca(OH)_2$ . However, the threshold was not directly investigated in this study. It would be worth discussing how future research might better constrain the threshold for  $Mg(OH)_2$  additions and how secondary precipitation might be mitigated when adding  $Mg(OH)_2$  vs.  $CaO$  and  $Ca(OH)_2$ . Is it more advantageous to optimize the grain size of the feedstock or to stay below the critical  $\Omega$  threshold?

Thank for this comment. We agree that the threshold not being directly investigated, a discussion would be a good addition to the manuscript. We added (lines 404-408): **“Such pattern was also observed for Ca-rich minerals as well, but at lower  $\Omega_A$ . While further research is required to precisely determine the critical  $\Omega_A$  for both Ca- and Mg-rich minerals, the longer time for  $CaCO_3$  runaway precipitation to be initiated and the overall higher  $\Omega_A$  may suggest that  $Mg(OH)_2$  is a safer alkaline feedstock for OAE. Nevertheless, staying under critical  $\Omega_A$  should remain the primary target of any OAE”.**

Regarding the second part of the question, we believe that staying below the critical  $\Omega_A$  threshold is and remain the more suitable option. However, most OAE research stresses out the importance of efficient grinding. This manuscript helps understanding that while dissolution rates can be effectively increased with smaller grain size, the higher surface area is a problem that need to be monitored as well.

CO<sub>2</sub> ingassing in experiments:

In many of the experiments (see Fig. 2, 4, 5, Line 187, Line 227), the data show the DIC increasing after the rapid drop in DIC and TA from CaCO<sub>3</sub> precipitation, yet the cause of this trend is not well-explained. In Fig. 1, the data show the pH of the small grain size experiment continuously decreasing after reaching its maximum, which the authors attribute to ingassing. Yet, this was not apparent in the medium and large grain size experiments. In Sect. 4.1, the authors claim that ingassing is likely occurring in the other experiments as well, but masked by the stronger pH increase during dissolution. However, their explanation seems to contradict Fig. 1, which shows the strongest pH increase during dissolution for the smallest grain size.

Thank you for the comment. We agree with the points the reviewer made and edited the text accordingly (lines 291-292 and 338-345). We explained that the CO<sub>2</sub> ingassing, while occurring in all experiments, is only clearly visible in the smaller grain size as in the other experiments, the slower dissolution led to a longer pH<sub>F</sub> increase overtime. We therefore believe that over the last 2-4 hours of pH<sub>F</sub> measurements, the changes appear stable while it is likely that the pH<sub>F</sub> goes down due to ingassing while increasing due to the last grains of Mg(OH)<sub>2</sub> dissolving overtime.

What was the pCO<sub>2</sub> of the seawater samples relative to the laboratory atmospheric pCO<sub>2</sub>, and is this consistent with ingassing?

The starting pCO<sub>2</sub> was consistent in all experiments, with a starting pCO<sub>2</sub> of about 420.61 ±28.59 μatm across experiments. We added this information in lines 84-85. Yes, pCO<sub>2</sub> was consistently lower than atmospheric CO<sub>2</sub> at the end of the experiments, with:

- Salinity 35: 389.51 ±10.01 μatm
- Salinity 28: 258.51 ±8.30 μatm
- Salinity 20: 116.45 ±15.90 μatm
- Large grain size: 312.24 ±26.15 μatm
- Medium grain size: 389.51 ±10.01 μatm
- Small grain size: 307.27 ±4.87 μatm

One weakness in the experimental design was that there were no control samples (subjected to the same headspace conditions from the periodic sampling but not the Mg(OH)<sub>2</sub> addition) from which the magnitude of gas exchange can be directly assessed. The authors should address how they can be certain that the changes in DIC not attributable to CaCO<sub>3</sub> precipitation are necessarily due to gas exchange.

Thank you for the comment. We understand the reviewer worries, however, any decrease in DIC could only be linked to CaCO<sub>3</sub> precipitation as the seawater pCO<sub>2</sub> never exceeded the atmospheric one. Therefore, if DIC is removed from the water column, it cannot be from degassing but only through CaCO<sub>3</sub> precipitation. Similarly. Any increase in DIC in seawater can only come from CO<sub>2</sub>, as carbonates do not dissolve in seawater. Furthermore, the lower seawater pCO<sub>2</sub> coupled to the increase in DIC can only come from CO<sub>2</sub> ingassing, rounding up the carbonate chemistry changes.

## Minor Comments:

Line 21: Change “higher alkalinity/pH” to “higher alkalinity and pH.”

This has been edited.

Line 83: What was the laboratory pCO<sub>2</sub>?

This was changed (lines 84-85) to **“This ensured comparable starting conditions for the various experiments, with a calculated starting pCO<sub>2</sub> of 420.6 ±28.6 μatm in all experiments.”**

Line 101: Ω<sub>A</sub> ~ 9 is not the threshold for CaCO<sub>3</sub> precipitation when adding CaO and Ca(OH)<sub>2</sub>. It would be clearer to explain early on that Ω<sub>A</sub> ~ 9 was selected based on the suspicion that the precipitation threshold for CaCO<sub>3</sub> precipitation when adding Mg(OH)<sub>2</sub> may be higher than the Ω<sub>A</sub> ~ 7 threshold observed when adding CaO and Ca(OH)<sub>2</sub>. This was explained later in Lines 304-305.

This was added lines 101-104: **“For all experiments, Mg(OH)<sub>2</sub> additions were adjusted to yield an Ω<sub>A</sub> of ~9 (Table 1) to allow for a significant TA increase and secondary CaCO<sub>3</sub> precipitation, based on previously found thresholds for CaO and Ca(OH)<sub>2</sub>, and with the assumption that the CaCO<sub>3</sub> inhibition role of Mg<sup>2+</sup> requires a higher Ω<sub>A</sub> for CaCO<sub>3</sub> precipitation within days (Moras et al., 2022)”**

Line 109: Please reference the specific figures in “see figures.”

This was changed (line 116) to **“(see figures 2 and 4)”**

Lines 110-111: What was the sample size for the TA and DIC measurements? The description in these two lines sounds like ~150 mL of sample was used for each TA and DIC measurement. However, if 10 samples were taken for TA and another 10 samples for DIC measurements (as suggested in Line 111), there would not be enough sample (1.5 L total). It sounds like most of the sample was consumed by the end of the experiment. Please clarify the measurement process and address how the increasing headspace over the course of the experiment may contribute to the instability of the DIC.

We understand the comment. However, the first two samples were taken from the reaction vessel with a floating lid, described in lines 89-90, where 1.5L of seawater was used. The whole content of the beaker was then transferred to 1L Schott bottles. However, while these bottles are labelled as

1L, they actually can contain slightly more, with about 1.2L when filled to the top. These 1.2L allowed us to collect up to 8 more samples for TA and DIC analyses.

Table 1: The organization of the table needs to be improved for clarity.

The Table 1 has been reworked and cleared from unused information. The table is divided in 4 blocks, with the experiment details, starting conditions, conditions at the peak of  $\Delta TA$  and final conditions.

Firstly, the initial carbonate composition (e.g., DIC and TA) of the samples should be reported, not just the changes in DIC and TA.

Thank you for the comment, these values are now reported in table 1 for further clarity.

The first three columns after “Experimental details” pertain to the dissolution phase of the experiment and the next three columns pertain to the secondary precipitation phase. It would be helpful to label these columns as such.

Thank you for the comment, we have added three headings: “Starting Conditions”, “Conditions after Full Dissolution”, and “End Conditions”.

Without reading the manuscript text, “Days of stable TA” is an unclear description. Consider including a short explanation in the table caption.

The following statement has been added to the caption (line 123): **““Days of stable TA” encompasses the time between maximum  $\Delta TA$  recorded and the start of  $CaCO_3$  runaway precipitation.”**

Line 176: The increase in TA for the small grain size experiment was ~428  $\mu mol/kg$  according to Table 1.

Thank you for pointing this out, however, we decided to report rounded number to the closest multiply of 10, i.e., 430 instead of 427.48, 530 instead of 530.27 and 350 instead of 351.25. for consistency, we preferred keeping the value as “430”.

Fig. 2: Please specify which input parameter pairs were used to calculate DIC and TA from pH.

Thank you for the comment, we have added this information at lines 154-156.

### Panel figures (Fig. 2, Fig. 4, and Fig. 5):

In general, the panel figures need to be improved for visibility as the font size on the axes may be too small.

The figures have been reprocessed and edited to increase the font size.

Why are the pH measurements not included?

pH<sub>F</sub> measurements were of prime interest at the start of the experiments (as per Figures 1 and 2). However, for CaCO<sub>3</sub> precipitation, TA and DIC measurements are more suitable since while CaCO<sub>3</sub> precipitate, pH<sub>F</sub> would decrease. At the same time, CO<sub>2</sub> ingassing can still occur, enhancing this pH<sub>F</sub> decrease. Therefore, the use of pH<sub>F</sub> for the second part of the experiment was not deemed relevant and TA and DIC were preferred overall.

Line 244: The text should be 120 μmol kg<sup>-1</sup> DOC.

Thank you for pointing this out. The text has been edited accordingly (line 254).

Fig. 5. Please explain in the caption what the error bars represent in this and other figures.

Thank you for the comment. The following statement has been added to the caption (line 267-268): **“Values reported in the [Mg] Effect graphs represent the average of triplicate experiments run at salinity 20 and salinity 20 + MgCl<sub>2</sub>, with standard deviations represented by the error bars.”**

Line 273: Reference Fig. 1.

Thank you for pointing this out. The text has been edited accordingly (line 284).

Line 275: The text should be “If CaCO<sub>3</sub> precipitation.”

Thank you for the comment, this has been edited (line 286).

Fig. 6: How were the TA data normalized?

Thank you for the comment. The data were normalised by reporting the changes in calculated TA (using the starting measured TA, DIC and salinity conditions, and the pH<sub>F</sub> measurements to reconstruct TA changes) from the start of the experiments over the first 30 minutes. This allowed to remove the effect of salinity and the different starting conditions (particularly visible in figure 3 due to low salinity) for better comparison.

Line 380-382: The mechanism of CaCO<sub>3</sub> precipitation in relation to grain size was not demonstrated in this study, so this conclusion should be stated as a hypothesis rather than a finding.

Thank you for pointing this out. The text has been edited accordingly (line 404-407): **“A major finding was that two processes seem to occur during CaCO<sub>3</sub> precipitation in relation to grain size, one where the higher surface area of smaller particles could increase precipitation rates, while the second could may maintain a higher pH around larger particles due to a larger diffusive boundary layer compared to smaller particles, which increased precipitation rates.”**