Interactive comment on “Technical Note: Highly precise quantitative measurements of total dissolved inorganic carbon from small amounts of seawater using a common gas chromatographic system: an alternative method compared to established detection systems” by T. Hansen et al.

T. Hansen et al.

thansen@geomar.de

Received and published: 14 June 2013

Interactive comment on “Technical Note: Highly precise quantitative measurements of total dissolved inorganic carbon from small amounts of seawater using a common gas chromatographic system: an alternative method compared to established detection systems” by T. Hansen et al.

We thank both referees for their positive and constructive reviews which will help us
to improve our manuscript entitled” Highly precise quantitative measurements of total dissolved inorganic carbon from small amounts of seawater using a common gas chromatographic system: an alternative method compared to established detection systems” Please find below our answers to your general and specific comments.

General comments by Anonymous Referee #1

The authors of this technical note present a method to measure dissolved inorganic carbon (CT) by gas chromatography. For the growing field of research on ocean acidification effects on marine organisms, an establishment of the described method could mean a big improvement, as it is relatively cheap to establish (gas chromatography systems are available in most institutions) and requires only a small sample volume for the measurements. Even though I am not an expert in gas chromatography, the technical details of the method seem to be well described and I have no reason to question the applicability for CT measurements.

1. There is only one larger point that should be addressed before publishing this manuscript: I am missing a comprehensive table comparing the three different methods you compare. Please provide a table with information on range of CT values used to verify the method, accuracy, precision, and overall uncertainty (next to the other two, this is also an important measure of measurement quality!) for your method as well as for the coulometric- and infrared detection method.

We will add a table comparing the methods. The table will contain the used volume injected to perform a CT measurement for each method, the total sample volumes required after standard operation protocol 1 (Dickson & Goyet 1994, Handbook of methods for the analysis of the various parameters of the carbon dioxide system in seawater) and the estimated measurement uncertainties.

2. If not too much extra work it would also be good to have some samples measured with all three methods. Currently, the number of measurements done to test your method is rather low, a method comparison could also help improving this aspect.
We agree with the reviewer that a direct comparison analysing CT samples using other methods would improve the method validation. We plan an intercomparison program between different labs.

3. Please check the grammar as well as the use of different tenses throughout the manuscript, maybe with the help of a native speaker. Especially in the methods section some parts are written in present and others in past tense, but those parts should be in the same tense.

We will rewrite the method section and check the entire manuscript in order to use the same tense (past tense).

Specific comments by Anonymous Referee #1

1. p. 4449 l.5: Please delete the “The” in front of “coulometric- and infrared. . .”. We will do this

2. p. 4440, l.20-21 and throughout the whole manuscript: What you refer to as “Dickson standard” should be called “certified reference material (CRM)” and should be referenced (e.g. Dickson et al. 2003 in Marine Chemistry) or described as produced by Prof. Dickson’s laboratory at Scripps.

We will replace “Dickson standard” with “certified reference material (CRM)” and use the corresponding reference. Moreover, in the revised version we will re-define the previous term “internal lab standard” by CT lab standard.

3. p. 4441, l.5-8: References are missing. We will add references:


4. p. 4441, l.9: This sounds as there would be a logical relationship between the small volume and the need for filtration. Furthermore, high biomass should generally be avoided in OA experiments (see Guide for Best Practices in Ocean Acidification Research and Data Reporting by Riebesell et al 2010).

We will clarify this misunderstanding. We will explain that there is an increasing request for processing CT measurements from biological experiments very often providing small sample volumes (e.g. Matthiessen, B., Eggers, S. L., and Krug, S. A.: High nitrate to phosphorus regime attenuates negative effects of rising pCO2 on total population carbon accumulation, Biogeosciences, 9, 1195-1203, 10.5194/bg-9-1195-2012, 2012). Though we agree that high biomass in general should be avoided in OA experiments, samples derived from biological experiments inherently contain biomass
which requires sample-filtration before a CT measurement can be conducted. We will clarify the misimpression of a logical relationship between small sample volumes and the need for filtration.

5. p. 4441, l.20: This should read “sample and acid addition”.
We will do this

6. p. 4441, l.28: Reference for standard deviation is missing.
We will add references:


7. p.4441, l.17-p4442, l. 2: The paragraph about GC measurements seems the interrupt the description of commonly used CT-measurements. Also, you continue to write about GC in the next paragraph (p.4442, l. 19 ff.). I suggest moving these two paragraphs together.

We started the paragraph with reporting about the history of CT measurements because GC technique quantifying CT from open ocean seawater samples was applied before coulometric and later NDIR detection was established. In order to avoid the interruption, we will re-arrange paragraphs and start with established methods to determine CT and continue with the evolution of CT measurements using the GC technique.

8. p.4442, l.3-5: Wording: If coulometric and infrared methods replaced CG measurements in the 70ties and 80ties, this is not “currently”.
We will delete “currently”
9. p.4442, l.8: Please combine the two brackets.
We will do this

10. p.4442 l.20-28: replace “DIC” by “CT”
We want to use “DIC” for “dissolved inorganic carbon” when measurements were applied for stable isotope ratio analysis only. When talking about quantitative DIC analytics, we use CT-measurements.

11. p.4445 l.7: The first sentence is lacking some information. Please specify what you mean, for instance by repeating part of the heading of this paragraph.
We will write now: In order to prepare a measurement chamber for CT measurements...

12. p.4446, l.9 & p.4447, l.22: A value can either be “accepted” or “true” (even though I would try to avoid talking about the “truth” in the context of measurements). I would rather call it the “assigned value”.

We will follow the recommendations of both reviewers. In the revised version we will use the term “certified reference value (CRM)” instead of “Dickson standard”. For the validation of CT measurements we will use the term “estimated measurement uncertainty” given by a standard deviation instead of using the term “precision”.

13. p.4448, l.23-25: You mention the need to filter samples from phytoplankton cultures several times throughout the manuscript but it never gets properly discussed nor do your data provide information on a) how large the impact of a filtration step can be (you would need measurements before and after filtration from abiotic samples with different CT); b) why this problem is solved better with your method compared to others. If you want this to be part of the manuscript, spend some more time discussing this issue.

We reported about CT measurements using filtrated and non-filtrated certified reference material (CRM) to evaluate this effect (page 4446 section 2.7.2). We will improve
the manuscript sections in order to strengthen the parts dealing with sample filtration.

14. p4449, l.18: Please add some more information to this sentence, e.g. “. . . the typically small sample amount available for CT measurements”.

We will rewrite this sentence and add more information about such experimental microcosm designs which typically give small sample amounts for CT measurements. This is due to the fact that today often applied multi-factorial experimental designs result in high numbers of replicates. This in turn leads to high numbers of small experimental units from which only small samples volumes are available.

15. p.4449 l.20-22: This sentence basically repeats what has been said above.
We will delete this sentence

16. p.4449 l.23-25: You use the words “potential”, “might” and “potentially” in one sentence for describing one single point. If you think this can be an issue, you can also be a bit braver here.
We will rewrite this sentence and delete the words “potential” and “potentially”

17. p.4449 l.26: “. . . e.g. due to manipulated seawater” should read “. . . e.g. in manipulated seawater” or “. . . g.e. when working with manipulated seawater”.
We will write now: “This could be particularly critical when using extremely high CT ranges in manipulated seawater”

18. p.4451, l.20: In your conclusion you suggest sub-sampling CRMs to decrease costs of CT measurements. To be able to rely on sub-sampling, one would need data showing that sub-samples are stable over time (i.e. no gas exchange with the surrounding atmosphere).

We agree with the reviewer that we do not presented data about the long term stability of collected sub-samples from CRM. However, we cited Brandes, 2009 and Spötl, 2005; page 4450, line 5-7, using same kind of sealing investigating stability of stable
isotope ratio of dissolved inorganic carbon in the course of time. Results indicated that such samples were stable over month.

19. Figure 3: Is the CT on the x-axis the assigned / target-value for the respective volume or measurement results? Please clarify.

We will clarify this misunderstanding. CT-Values on x-axis were assigned values for the respective injected volume taken from CRM and CT lab standard (sodium carbonate solution). Y-axis represents measured peak area by GC for the respective volume. See additional response from authors to Prof. Dr. A. Dickson, special comments: #10. p. 4447, section 3

General comments by invited referee Prof. Dr. Andrew Dickson

1. The authors describe a GC-based approach to measuring total dissolved inorganic carbon in seawater. The GC approach is similar to that used for at least 30 years to measure CO2 in air (i.e., using an acidic drying agent, together with hydrogen and a nickel catalyst to ensure reduction of CO2 to methane and subsequent detection with an FID); the use of a cold-trap to "focus" the CO2 from an acidified seawater sample is essentially that of Wong (1970), though – it appears – automated. The authors’ contribution has been to put a system together from commercially available components that achieves this measurement relatively straightforwardly.

We added information about earlier scientific work using GC technique to quantify CT from seawater samples. We will additionally cite Wong (1970) in the method section regarding the application of traps to separate carbon dioxide from impurities and water. However, we do not agree with the reviewer’s comment that our contribution can simply be cut down on putting together commercially available components. The applied components are used for many other analytical purposes (e.g. Kaminski et al. 2003; “Determination of carbon monoxide, methane and carbon dioxide in refinery hydrogen gases and air by gas chromatography”).
2. They also indicate that – with care – it may be practical to filter samples for this measurement (a practice that is usually not done for high-quality measurements). So what might make this a "better" method? The prime aspect that the authors emphasize is that it provides a good estimate of total dissolved inorganic carbon from a relatively small sample of seawater (10 mL). This is – in part – because typically only 0.200 mL of sample are injected into the stripping chamber for analysis; the balance of the sample acts to mitigate the likelihood of significant CO2 loss / gain occurring during handling and to allow for replication, if desired. I should point out here that two commercially available systems for measurement of the total dissolved inorganic carbon in seawater (the Airica from Marianda, and the AS-C3 from Apollo SciTech Inc.) also claim in their specifications that small sample sizes (1 mL) are used, and one explicitly states that "less than 10 mL of sample is required." (This contrasts to the comment – p. 4449, line 19 – suggesting that at least 250 mL are needed.)

We tried to explain in the introduction that there is an increasing scientific interest to investigate the influence of seawater acidification on biological processes. Such samples often contain a non negligible quantity of biomass which needs filtration before a CT measurement can be performed. We showed that sample filtration as applied in this manuscript does not affect the estimated measurement uncertainty.

We will add following information: Commercial manufactures like “Marianda, Kiel, Germany” and “Apollo SciTech Inc, Bogart, USA.” providing CT analyzers with NDIR technique. These analyzers also require small amounts of sample to perform a CT measurement. However, no detailed statements are given by the manufactures about (i) the total sample volume required during a sample collection (e.g. SOP 1, Dickson & Goyet 1994) supplying these technique and (ii) the volume necessary for rinsing procedure between alternating high and low CT measurements.

3. The authors also suggest that cost is a significant advantage of their system (noting the significant cost of the LiCor 7000 for IR systems – though it is practical to use the much cheaper LiCor 840 with only some degradation of precision – Ludger
Mintrop, Marianda, personal communication). I should like to take this opportunity to point out that choosing a system based on acquisition cost alone is ill-advised; it is recommended that the full costs – including the (usually dominant) personnel costs for operation and maintenance of the equipment, in addition to the acquisition and operating costs – be compared. (I note that the analysis time for a single sample – excluding time for changing samples – is 14.5 min: Fig. 2.)

We will discuss the costs for the hardware, operation and maintenance of our method in comparison to other established methods in more detail. We will add following information: “Costs of analytical hardware for our GC method is lower compared to other established techniques. The subdivision of 500mL CRM into approximately 20 “reference standards”, 10 mL each, for GC measurements and the accumulation of only small amounts of toxic waste during the sample measurements additionally reduce the total analytical costs. However, maintenance of the GC system with the gases helium and hydrogen, and the comparatively long measurement time of 14.5 min for performing a CT measurement which results in higher personnel costs, potentially compensate for this advantage.”

4. However, it is the description of the method itself, and the characterization of its performance, that I feel still need some improvement. The instrumental description is quite vague and ill-organized, possibly making it difficult for other laboratories to replicate as it neither provides a clear statement of the components used, nor does it guide other labs as to the component specifications that are key to achieving the results detailed here. Thus I believe that this section (describing the system) could benefit from careful planning and reorganization to assist such readers.

We will specify and reorganize the instrumental description. In particular we will add more information about used components to guide other users.

5. Also, the statement is made that: "Method validation was done in agreement with the terminology for accuracy and precision (RSC, 2003)." I disagree. That – admittedly...
short – document differs from the authors in a number of key aspects. For example, it states: "Because precision depends on the conditions of measurement, the conditions must be specified when referring to an estimate of precision." The authors do not do this. (I feel that their precisions are all estimated under repeatability conditions and would be better estimated as a pooled value.) Furthermore, the RSC document is at pains to state, "an estimate of uncertainty is probably the most appropriate means of expressing the accuracy of results" and also "An estimate of uncertainty must incorporate what is known about random effects and what is known about uncertainties associated with systematic effects on the measurement process." Granted this is non-trivial (see for example "Quantifying Uncertainty in Analytical Measurements", EURACHEM/CITAC Guide CG 4, 2012), but would be a much more useful guide to the method than the (I suspect over-optimistic) comment that: "Our technique is able to provide . . . an accuracy of $\pm 1.2$ mol/kg . . ." Thus, although I believe that this is a potentially useful approach, particularly as there is a need for measurements of total dissolved inorganic carbon in seawater that can achieve adequate quality from a small original sample volume, I feel it could have been presented far more effectively, and that it should not be published in its current form.

We agree with the reviewer and will accordingly revise this part of our method section. We will replace the term “precision” with the term “estimate of measurement uncertainty” for presented analytical data following the guideline recommended above. Authors will mention known aspects of random and systematic effects which can influence our measurements.

Special comments by Prof. Dr. Andrew Dickson

1. p. 4441, line 3. I would dispute that the need is for "precise" measurements, instead I would assert we need measurements of known "uncertainty" that can therefore be compared reliably both within and between laboratories.

We will use the term “estimated measurement uncertainty” to verify our method.
2. p. 4442. A variety of "precisions" are stated. I am not convinced that they are really comparable, as I would note that the numbers given for the precision of coulometric analysis are those found from repeated analyses of separate bottles of CO2 in seawater reference materials over a period of months, whereas the precision in Stoll et al. seems to be a repeatability (based on a single analyzer run).

We will consider different information about presented NDIR detection and continuous flow method (Stoll et al.) to summarize "precisions" as estimated measurement uncertainties in a table together with the coulometric method. For this, see also response to general comment #1 by reviewer #1.

3. p. 4444, section 2.2 I am surprised (given my own experience with using Na2CO3 as a calibration standard) that it was not dried (at >270 °C) nor is any comment noted about care with regard to static when weighing or transferring. I also could not find a Sartorius SR2 balance with a simple web search, so feel it would be desirable to indicate the balance accuracy for this 4 mg weighing.

We will add the information about the preparation of a sodium carbonate calibration standard (CT lab standard). Here, we followed the recommendation after SOP 2; “Handbook of methods for the analysis of the various parameters of the carbon dioxide system in sea water”, Annexe B, authored by Dickson and Goyet 1994. We could not report about static when weighing or transferring dried sodium carbonate. This was probably due to using fine granulated sodium carbonate (Suprapur®, 99,999 %, Merck, Darmstadt, Germany). Moreover we used a metal spoon to transfer sodium carbonate into tin caps. We will add and correct detailed information about the used type of balance and its “accuracy”. Type of balance is actually SC2, called micro balance, Company SARTORIUS; readability: 0.1 µg; reproducibility: 0.25 µg.


4. p. 4444, section 2.3. "samples were collected by hydrostatic pressure into 10 mL vials"—It is not clear what this means. Perhaps pneumatic pressure was used to force
the sample through the tube?

Authors will rewrite the collection procedure of sub-samples from a 500 mL CRM bottle (batch 108, CT = 2022.7 ± 0.45 µmol kg-1). We will write now: “The samples were forced into 10mL vials by gravity. In order to initiate the flow, a single-use syringe was used to draw the samples into a Tygon® plastic tube.”

5. p. 4445, line 4. The use of 20 microliters of saturated HgCl2 solution in a 10 mL sample will change the total dissolved inorganic carbon by about 0.2% (4 micromol per kg at a level of 2000 micromol/kg). Was this dilution adjusted for?

We agree that a mistake happened. Authors will correct for this. Final HgCl2 concentration was adjusted for ~0.02% like recommended by SOP 1; Guide for water sampling, Dickson&Goyet 1994. We will write now: “The filtered samples were poisoned with 2 µl of a saturated HgCl2 solution and immediately closed with a headspace <1 % taking butyl rubber-septa and stored at 4 °C in darkness until usage.”

6. p. 4445, line 20. "After the gas pressure inside the measurement chamber was close to zero . . ." – Surely not! Is this the amount of CO2 remaining in the measurement chamber?

We agree with the reviewer. We will rewrite this sentence and state that the pressure inside the measuring chamber was close to ambient pressure.

7. p. 4446, line 12-13. "... were always blank corrected (5 micromol ± 1.8)." – Should this not be micromol/kg? Also, it is not clear what the relevance of the phrase "200 microliters of DI water which was used for the preparation of the sodium carbonate solution" is. Does mean that that solution has a necessarily higher "blank"?

Authors will write “µmol kg-1”. We will rewrite the phrase about the measurement of 200 µL of DI water and explain it in more detail. Each measurement using the CT lab standard (sodium carbonate solution) contained an additional blank by the injection of 200 µL, which was generated by using DI water as the solvent to prepare the sodium
carbonate solution. This measured blank was taken into account when calculating the CT.

8. p. 4446, line 13-14. "sample temperature and salinity were taken into account" I’d guess this is for a density estimate, but it doesn’t say so.

Yes, sample temperature and salinity were taken into account to estimate the density. We will rewrite this sentence.

9. p. 4446, section 2.7.1 suggests that different volumes of both the internal standard and the "Dickson standard" were injected to achieve a range of values; however, the first sentence of section 3 implies that "different concentrations" of internal standard were used together with "different volumes" of Dickson standard. Which is correct?

The range of different values of both standards was achieved by injecting different volumes. We will correct this mistake and rewrite the first sentence of section 3.

10. p. 4447, section 3. I found this confusing. The first paragraph seems to describe the data of Fig. 3, providing the least-squares parameters describing the line which is, apparently, based on two separate Na2CO3 solutions and the Dickson standard. However, the statement regarding the "precision" seems to be based on Na2CO3 solutions alone. Also, I am not clear as to the significance of the "intercept" (equivalent to about 10 micromol/kg?), particularly in the light of the comment about blank correction on the previous page.

Authors will rewrite the section 3 and separately present least-square parameters of injected volumes using CRM and sodium carbonate as CT lab standard. We will discuss the significance of these intercepts regarding blank estimation.

11. The second paragraph then describes Fig. 4, with the mean of the 9? measurements on the "original" (not subsampled but from the same bottle?) Dickson standard seeming to be identical to the certified value. Is this a lucky accident or were the data adjusted to indicate that the other handling methods only result in small average de-
viations? Also where does the number chosen as an "accuracy" come from? The average deviation of all these 27 data from the certified value? As the RSC document the authors quote is at pains to state, "an estimate of uncertainty is probably the most appropriate means of expressing the accuracy of results".

Authors will rewrite both paragraphs and clarify misunderstandings; e.g. the CRM was repeatedly (3x) measured and the resulted GC response in peak area was arithmetically averaged and set as “true value (batch 108, CT = 2022.7 µmol kg-1) ± our measured standard deviation. Additionally, four sets of triplicate measurements from filtered and non-filtered sub-sampled (10 mL) CRMs, respectively, were performed over a course of four days. This resulted in 12 measurements for filtrated and non-filtrated CRMs, respectively. In total, this gave 27 measurements (3 CRM; 12 sub-CRM; 12 filtered sub-CRM). In the revised version we will present the data in a table instead of in figure 4. As a proxy for the estimated measurement uncertainty, we will report the standard deviations for each set of the triplicate measurements, and as requested by the reviewer, as pooled standard deviations across all measurements of sub CRMs and filtrated sub-CRMs, respectively.

12. p. 4448, Paragraph 3 of section 3 discusses Fig. 5, again treating its "precision" as a separate concept for statistical analysis. I really believe that all the replicate data given in this paper should be combined to provide a single estimate of the measurement repeatability (a pooled standard deviation). I do not believe that any of the estimates given are really significantly different one from another as the individual numbers of degrees of freedom are typically small. (You could confirm my assertion with a series of F-tests.)

As requested we will present average measurement values ± standard deviations from the three different phytoplankton cultures (i.e with ‘low’, ‘intermediate’, and ‘high’ CT) in a table. Additionally we will provide the pooled standard deviations across all measurements. We do not really understand the comment regarding applying a series of F-tests because the previously provided but criticized ANOVAs were per definition
based on F-statistics.

13. p. 4448, line 22. The assertion is made that other methods have not been "validated" over an equivalent concentration range. This is not really very convincing. First, Goyet & Snover based their results on a calibration curve over a wide range of total inorganic carbon concentrations; second, even the coulometric method has been shown to be linear over a larger range than stated here, and it too is an integrative method thus is unlikely not to work just as well (in absolute terms) at lower concentrations.

Authors will correct their mistake and delete this sentence. We will cite Goyet & Hacker 1991 and Goyet & Snover 1993 who presented a calibration curve from measured standard solutions (sodium carbonate) covering a CT range of 0-2500 µmol kg.

14. Fig 2. I’d have the time axis in minutes (with one decimal place); this seems to be the inherent time step.

We will set the time axis of Fig 2 in minutes with one decimal place.

15. Fig. 3 A plot like this is an almost useless approach to displaying good-quality linear data where the size of the symbol significantly exceeds the measurement repeatability. I suggest plotting the deviations from a straight line as a function of total inorganic carbon concentration, possibly scaled to indicate the discrepancy in concentration units.

We will replace Fig 3 with a figure which separately shows least-square parameters of injected volumes for sub-CRM and CT lab standard (sodium carbonate solution), respectively. Each data set will consist of 6 single measurements. We will present the data like Goyet & Snover 1993 to indicate the linearity over a wide CT range. Previously included triplicate measurements and its resulting standard deviations of the second CT lab standard will be presented in an extra table.

16. Flgs 4 & 5 & 6 would be even clearer as tabular material (and then some of the text might also not be needed). Fig 6. It seems a pity that the two parts of this plot are not on the same scale as this confuses any comparison of precision.
We will present measured data from Fig 4, Fig 5 and Fig 6 as tabular material and rewrite the result section. See our response to your comments #11 and #12 above.

Finally, I don’t believe there is a 1:1 correspondence between citations, and the bibliographic list of references. The authors should check this more carefully.

We will check correspondence between citations and bibliographic list of references.

Interactive comment on Biogeosciences Discuss., 10, 4439, 2013.