**Earth System
Science Data
Discussions**

Interactive
Comment

Interactive comment on "Data recovery of A06 and A07 WOCE cruises" *by* N. M. Fajar et al.

Anonymous Referee #1

Received and published: 12 December 2011

Fajar and colleagues attempt to "recover" problematic inorganic carbon system data from the A06 and A07 WOCE cruises.

[Here are comments on these data sets, according to Wanninkhof et al. 2003: Total Alkalinity (TA): "Based on nine crossovers, the offsets and scatter are such that we do not recommend including the data from A06 or A07 in the synthesis product." Dissolved Inorganic Carbon (DIC): "The 11 crossovers showed high variability and often large differences along with large standard deviations, suggesting significant variability over the density range investigated. The MLR indicated that these cruises had biases of 11 to 15 μmol/kg with large standard deviations compared with other cruises in this region. These results suggest significant station-to-station differences and possible variability with depth. Therefore, these cruises are not recommended for use."]

Fajar et al. believe they can "fix" these data through a series of assumptions.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

I will comment on what I believe Fajar et al. did based upon multiple readings of the manuscript, however, it should be pointed out that the writing itself is of rather poor quality and is difficult to follow. For example, how is one to interpret the following sentence (Sect. 4.2 para. 1 lines 20-21)?: "However, idem that happened in AT crossover analysis, not all the cruises have given an offset." This is but one of many intelligibility issues that crop up within the manuscript.

Regarding their analysis:

First, they use a "3-DwMLR" technique to generate new TA data: "the unreliable quality AT data of A06 and A07 has been replaced by the 3-DwMLR method..." The assumption here is that one can infer TA values based on measured oxygen, nutrients, etc. I see this as a misuse of MLR techniques. These techniques are useful for detecting offsets/biases between datasets, and for detecting individual outliers. MLR results are not a substitute for measured data. TA data may be inaccurate because of calibration offsets/biases, in which case a single linear correction may be applied to a dataset; they are also subject to random errors, which cannot be corrected. Had Fajar et al. applied a single linear correction to the dataset, I could have considered their approach plausible, but what they appear to have done is to replace measured data with MLR results that better fit their view of how TA should behave.

Second, they use these "inferred" alkalinity data to generate "new" DIC data, assuming that the DIC/TA ratio is accurate despite the fact that neither the DIC nor the TA measurements are "reliable." What is the basis for accuracy of DIC:TA when neither value alone is considered accurate? Some discussion of the respective standards used to calibrate DIC and TA measurements is warranted here; otherwise, it is hard to see why should we expect the respective errors to covary.

Third, they derive "new" pH data; it is still unclear to me whether they used CO2SYS to calculate these from their "new" TA and DIC data, or whether they used an MLR approach with the raw pH data. Either way, the authors would like us to accept that

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

these "new" pH values are somehow more "realistic" than the measured pH values be-
cause of the resulting "realistic" profiles; but, "smoothness" of a profile does not equate
with accuracy. I am especially troubled that no information is available on the temper-
ature at which the original pH data were measured. Who made these measurements?
Can't they be asked what they did? Have the Chief Scientists of these cruises been
consulted?

The authors themselves admit (Sect. 3.1 para. 2 lines 4-5, Sect. 4.2 para. 1 line
1, Sect. 4.3 para. 1 lines 4-5, respectively) that the TA, DIC and pH data from these
cruises are unreliable. I am not convinced that their reworking of these three unreliable
datasets should yield reliable results. It is probably best that these data are treated
exactly as they were by Wanninkhof et al. (2003): as unsuitable for use in global
compilations.

Interactive comment on Earth Syst. Sci. Data Discuss., 4, 99, 2011.