

# Reply to comments of Mario Hoppema (Topic Editor), received on 12 Mar 2020

First, we thank Mario Hoppema for serving as Topic Editor and for the thorough evaluation of our manuscript. Thank you for all these valuable comments that highly improve our manuscript.

We answer each comment point-by-point in the following text. A revised manuscript with marked changes based on these comments can be found later in this document. Furthermore, all lines named in this answer refer to this changed and marked manuscript.

**Comment #1:** P1, L3 „The work described here focuses on impacts of ...” Work cannot focus on something, but you yourself can. Please correct the sentence.

**Answer # 1:** Modified the sentence (see line 3)

**Comment #2:** P1, L6 did (instead of does). Appears to fit better in such a sentence. If you need to use “does”, the sentence ending must be changed.

**Answer #2:** Changed into ”did” (see line 6)

**Comment #3:** P1, L6 I do not see the use of the word “subsequent” here.

**Answer #3:** The word ”subsequent” refers to the development of the method, which had to be done before the analysis. But for the whole understanding of the abstract, it is not necessary. Consequently, we leave it out (see line 6).

**Comment #4:** P1, L11 “However, it could not be fully overcome.” What could not be overcome? Please clarify in the text.

**Answer #4:** The ”it” means the drift behavior of the analyzer. We clarified this in the text (see line 11).

**Comment #5:** P1, L19-20 the purest BCG that is purchasable.

**Answer #5:** Changed from ”which” to ”that” (see line 20).

**Comment #6:** P1, L22-23 “not only the biogeochemical processes but also the uptake, transport and accumulation of anthropogenic CO<sub>2</sub> in the ocean” I think these three processes could all be taken as biogeochemical processes. Please modify text.

**Answer #6:** Changed the text (see lines 23 - 24).

**Comment #7:** P1, L23 parameters or variables?

**Answer #7:** They are variables. We changed the text with refer to this (see line 24).

**Comment #8:** L25 I think “corresponding” is not appropriate here.

**Answer #8:** As ”corresponding” is not necessary for the understanding of the sentence, we leave it out (see line 25).

**Comment #9:** L27 pH measurements have been common all the time. I think “prominent” fits better here.

**Answer #9:** Modified the sentence by using ”prominent” (see line 28).

**Comment #10:** L28 I think you mean “decades” instead of “centuries”. Ocean carbon observations are not that old.

**Comment #10:** Yes, we meant ”decades”. Changed the word (see line 29).

**Comment #11:** L38 “have revealed”, not reveals

**Answer #11:** Changed into ”have revealed” (see line 39).

**Comment #12:** L64 please define LC here

**Answer #12:** A definition for ”LC” has been included (see line 64).

**Comment #13:** L119 Would it be possible to use a different word than “calibrated”, since you do not mean calibrated here, do you? Or possibly rewrite the sentence completely.

**Answer #13:** You are right, it is not a real calibration. It is a sample volume determination for the inner tubing of the analyzer, as this is the only unknown variable within the whole  $A_T$  determination with the HydroFIA TA. We added a proper explanation for ”calibrated” (see lines 121 - 123).

**Comment #14:** L128 accounted for instead of considered?

**Answer #14:** Changed ”considered” to ”accounted for” (see line 131).

**Comment #15:** L159 delete: There,

**Answer #15:** Deleted (see line 160).

**Comment #16:** L160 delete: However, (there is no contradiction here)

**Answer #16:** Deleted (see line 164).

**Comment #17:** L173 delete: respectively

**Answer #17:** Deleted (see line 177).

**Comment #18:** L194-195 This sentence is hard to read. I suggest to change it to: All AT measurements took into account the relative uncertainty of the analyzer, determined as 0.08 % (Seelmann et al., 2019).

**Answer #18:** Changed the sentence following the suggestion (see line 198 - 199).

**Comment #19:** L201 Many sentences here begin with “However”. Maybe change this one?

**Answer #19:** Rearranged the sentence without ”However” (see line 206).

**Comment #20:** L203-204 Maybe the authors can indicate what they consider to be

acceptable for a drifting and changing value of AT. This would give some guidance to readers/users of the system.

**Answer #20:** We added information to this (see lines 214 - 217).

**Comment #21:** L209 change to: on a monthly basis

**Answer #21:** Changed (see line 214).

**Comment #22:** L227 The phrase “For better presentation” reads odd. It is not just a matter of presentation that you are performing this. Please rephrase.

**Answer #22:** The sentence was rephrased (see line 235 - 238).

**Comment #23:** L228 delete: At the time of this study, (this is superfluous)

**Answer #23:** Deleted (see line 237).

**Comment #24:** L242 “In contrast, ”low-purity” indicators behaved totally different.” “in contrast” and “totally different” say exactly the same. Please rephrase.

**Answer #24:** Deleted ”In contrast” as it is not necessary (see line 252).

**Comment #25:** L257 Could the authors come up with a different word than “calibration”? I think the reader might get confused by bringing it this way. If it is a real calibration, then the use of quotation marks is not necessary.

**Answer #25:** It is not a real calibration. It is an internal seawater volume determination, which is also explained before (Method section) and directly afterwards in the manuscript. The operation software of the HydroFIA calls the procedure ”calibration” and we adapted this term. However, we added a statement that this is not a real calibration in order to inform the reader about this (see lines 121 - 123 and 267 - 269). We hope this is sufficient. Besides, there is exactly the same situation for the SOMMA system for DIC measurements, where the CRM measurements are essentially performed to ”calibrate” the to-deliver-volume of the pipette (plus any additional smaller uncertainties such as the performance to ”electron counting” in the coulometer). But this is just for your information.

**Comment #26:** Table 4 Both the precision and accuracy are given. In some cases the accuracy is better than the precision. I think this is not according to statistical principles. In the requirements for open-cell titrators according to Dickson et al 2007, one can see how it should work: the accuracy has a clearly higher absolute value (often double as large) than the precision. The reason is that for determining the accuracy, one needs to take into account the precision of the measurement of the reference material (or standard) used to determine the accuracy, and of the measurement itself. Please explain and correct your results.

**Answer #26:** We equate the bias with the accuracy (requirement of  $\pm 2 \mu\text{mol kg}^{-1}$  by Dickson et al 2007 = ”overall bias” (see page 11)). The combination of accuracy and precision is then called uncertainty. But you are right, sometimes uncertainty and accuracy are the same, and a combination out of bias and precision. This depends on

the definitions. However, our "accuracy" values are the bias between the mean of the measured values and the true value (in our case the RMSE of the linear regression), as given in the Guide of Dickson et al 2007 on page 11 for standard open cell titrators: "...overall bias of less than 2  $\mu\text{mol kg}^{-1}$ ...". However, we replaced the word "accuracy" with "bias" in order to avoid confusion about the used terms (see lines 15, 147, 244, 246, 248, 258, 303 and table 4). We hope that is sufficient.

**Comment #27:** L272 appropriately, not probably

**Answer #27:** Changed (see line 283).

**Comment #28:** L294 delete However, (not necessary here, and not at the start of a paragraph)

**Answer #28:** Deleted "However" and rearranged the sentence (see line 305).

**Comment #29:** L320 delete: In this study, (superfluous)

**Answer #29:** Deleted (see line 313).

**Comment #30:** As to the Conclusions section, this is much like a summary. Although this cannot be avoided to some extent, the entire structure of this section in the present manuscript is built around this summary. See for example the last sentence of the section in L341 "Furthermore, no characterization of the purified BCG was carried out." This may well be deleted. Please restructure the Conclusions section and incorporate more real conclusions, which the manuscript definitely has to offer.

**Answer #30:** We modified the Conclusion Section to a more concluding part with recommendations for the reader/user instead of a summarizing part (see Conclusion section).

**Comment #31:** L367 CO<sub>2</sub> with subscript

**Answer #31:** Changed (see line 389)

# Impact of impurities in bromocresol green indicator dye on spectrophotometric total alkalinity measurements

Katharina Seelmann<sup>1</sup>, Martha Gledhill<sup>1</sup>, Steffen Aßmann<sup>2</sup>, and Arne Körtzinger<sup>1,3</sup>

<sup>1</sup>GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

<sup>2</sup>Kongsberg Maritime Contros GmbH, Kiel, Germany (at the time of this study)

<sup>3</sup>Christian-Albrechts-Universität zu Kiel, Kiel, Germany

**Correspondence:** Katharina Seelmann (kseelmann@geomar.de)

**Abstract.** Due to its accurate and precise character, the spectrophotometric pH detection is a common technique applied in measurement methods for carbonate system parameters. However, impurities in the used pH indicator dyes can influence the measurements quality. ~~The~~ During our work described here ~~focuses~~, we focused on impacts of impurities in the pH indicator dye bromocresol green (BCG) on spectrophotometric seawater total alkalinity ( $A_T$ ) measurements. In order to evaluate the extent of such influences, purified BCG served as a reference. First, a high-performance liquid chromatography (HPLC) purification method for BCG was developed as such a method ~~does~~ did not exist at the time of this study. ~~A subsequent~~ An analysis of BCG dye from four different vendors with this method revealed different types and quantities of impurities. After successful purification,  $A_T$  measurements with purified and unpurified BCG were carried out using the novel autonomous analyzer CONTROS HydroFIA® TA. Long-term measurements in the laboratory revealed a direct influence of impurity types and quantities on the drift behavior of the analyzer. The purer the BCG, the smaller was the  $A_T$  increase per measurement. The observed drift is generally caused by deposits in the optical pathway mainly generated by the impurities. However, ~~it~~ the analyzers drift behavior could not be fully overcome. Furthermore, we could show that a certain impurity type in some indicator dyes changed the drift pattern from linear to non-linear, which can impair long-term deployments of the system. Consequently, such indicators are impractical for these applications. Laboratory performance characterization experiments revealed no improvement of the measurement quality (precision and ~~accuracy~~ bias) by using purified BCG as long as the impurities of the unpurified dye do not exceed a quantity of 2 % (relationship of peak areas in the chromatogram). However, BCG with impurity quantities higher than 6 % provided  $A_T$  values, which failed fundamental quality requirements. Concluding, to gain optimal  $A_T$  measurements especially during long-term deployments, an indicator purification is not necessarily required as long as the purchased dye has a purity level of at least 98 % and is free of the previously named impurity type. Consequently, high-quality  $A_T$  measurements do not require pure but the purest BCG ~~, which~~ that is purchasable.

## 1 Introduction

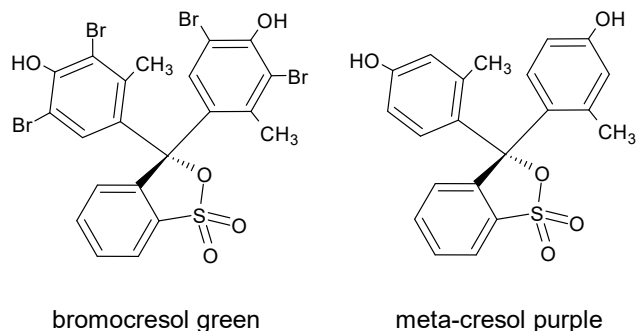
Global observations of the marine carbonate system are of high importance to understand ~~not only the biogeochemical processes but also the uptake, transport and accumulation of anthropogenic CO<sub>2</sub>~~ biogeochemical processes in the ocean effected by anthropogenic CO<sub>2</sub>. The measurable key ~~parameters~~ variables characterizing the ocean carbon cycle are pH, total alkalinity

25 ( $A_T$ ),  $p\text{CO}_2$ , and total dissolved inorganic carbon ( $C_T$ ). Due to their ~~corresponding~~ thermodynamic relationships, it is only necessary to measure two of these four parameters for a full characterization of the marine carbonate system (Millero, 2007). Traditionally,  $A_T$  and  $C_T$  were the preferred parameters for this purpose when measuring discrete samples. However, more recently, pH measurements have become more ~~common~~ prominent within the oceanographic communities. During ~~centuries~~ decades of ocean carbon observations, several analytical methods have been established, ranging from manual bench top  
30 systems for laboratory work via at-sea flow-through analyzers to in situ sensors. Among all these available methods, spectrophotometric pH determination techniques using sulphonephthalein indicator dyes are described as simple, fast, and precise (e.g. Clayton and Byrne, 1993; Tapp et al., 2000; Bellerby et al., 2002; Aßmann et al., 2011). They have been utilized in marine research especially for ocean carbon observations since the late 1980's (Robert-Baldo et al., 1985; Byrne, 1987; Byrne and Breland, 1989; King and Kester, 1989). Since Breland and Byrne (1993) showed that the sulphonephthalein indicator dye  
35 bromocresol green (BCG) is suitable for seawater pH determination in the pH range 3.4 to 4.6, it has been used in several spectrophotometric  $A_T$  measurement systems with comparable precision and accuracy as traditional methods (Yao and Byrne, 1998; Li et al., 2013; Seelmann et al., 2019).

Investigations of Yao et al. (2007) on seawater pH measurements with the most common indicator dye, meta-cresol purple (mCP) from different vendors ~~reveals~~ have revealed different types and quantities of light-absorbing impurities. These impuri-  
40 ties can contribute to pH offsets of up to 0.01 pH units. To overcome the uncertainties caused by indicator impurities, Liu et al. (2011) developed a preparative high-performance liquid chromatography (HPLC) method to purify mCP and characterized this purified dye. Furthermore, to produce large batches of purified mCP, Patsavas et al. (2013a) developed a flash chromatography (FC) method resulting in a 3.5 times increased yield per run. However, not all users of spectrophotometric seawater pH measurement systems are able to purify or to purchase purified mCP. Therefore, Douglas and Byrne (2017) published a  
45 mathematical correction for accurate pH measurements using unpurified mCP.

In order to apply these findings to spectrophotometric  $A_T$  measurements, Nand and Ellwood (2018) described a simple colorimetric method for determining seawater  $A_T$  using purified bromophenol blue (BPB) as pH indicator dye. However, at the time of this study, there are no comparable detailed investigations on how indicator impurities in BCG may influence spectrophotometric seawater  $A_T$  measurements.

50 Since our previous work dealt with an open-cell single-point titration analyzer with spectrophotometric pH determination using BCG as indicator dye (Seelmann et al., 2019), we investigated the influences of any impurities in BCG from different vendors in comparison to purified BCG as reference. Hence, the first step of this study was to develop a purification method for BCG. Due to similarity in the chemical structure of BCG and mCP (see Fig. 1) and the available facilities in our laboratory, we decided to develop an HPLC analysis and purification method for BCG based on the mCP purification method published by  
55 Liu et al. (2011). Once the developed method was sufficient for BCG purification, a small batch of purified BCG was produced. Following this, comparative experiments were carried out with a novel autonomous analyzer for seawater  $A_T$  using purified and unpurified BCG in order to evaluate the influence of impurities in the indicator dye.



**Figure 1.** Chemical structure of bromocresol green and meta-cresol purple

## 2 Materials and methods

### 2.1 HPLC method

#### 60 2.1.1 Reagents and instrumentation

The BCG indicator (as sodium salt) was obtained from the following vendors: Acros Organics, Alfa Aesar, Carl Roth, and TCI. The solvents used in the HPLC purification were water (H<sub>2</sub>O), acetonitrile (ACN), and trifluoroacetic acid (TFA). The ACN (HPLC grade) and the TFA (Purity: ≥ 99.9 %) were obtained from Fisher Scientific and Carl Roth, respectively.

A Shimadzu [LC-liquid chromatography \(LC\)](#) system performed both the analytical and preparative chromatography. This system included an auto-sampler (SIL-10ADvp) (only for analytical mode), a isocratic preparative LC pump (LC-8A), an isocratic analytical HPLC pump (LC-10ADvp), a column oven (CTO-10ASvp), a single channel UV-VIS detector (SPD-10Avp), and an LC controller (SCL-10Avp).

The Primesep B2 HPLC columns were obtained from SIELC Technologies. This is a reverse-phase column with embedded basic ion-pairing groups that retains analytes by reverse-phase and ion-exchange mechanisms. For developing the purification method, an analytical Primesep B2 column (4.6 x 250 mm, particle size: 5 μm) was chosen. The purification was performed by a preparative Primesep B2 column (21.2 x 250 mm, particle size: 5 μm). Analytical separations were performed at 25 °C, but preparative chromatography was undertaken at room temperature.

#### 2.1.2 Method development

The method development included the optimization of the mobile phase composition for BCG separation on the chosen column and was performed in analytical mode. For this purpose, a 10 mg mL<sup>-1</sup> BCG solution of each vendor was prepared in the mobile phase and 20 μL were injected. One HPLC run with a flow-rate of 1.5 mL min<sup>-1</sup> took 60 min and was monitored using the UV-VIS detector at 280 nm. The optimal mobile phase composition was determined by systematically changing the concentrations of the solvents starting from the conditions described by Liu et al. (2011). There, the mobile phase composition

was 70:30 ACN:H<sub>2</sub>O (volume:volume) with 0.05 % of TFA. Afterwards, the ACN and TFA concentration were increased by 5 % and 0.05 % increments, respectively, until the mobile phase consisted of 85 % ACN and 0.2 % TFA. One BCG injection was done per mobile phase composition each followed by a blank run. Blank runs were carried out by injecting the mobile phase as sample.

### 2.1.3 Comparison of BCG from different vendors

Once the optimal mobile phase composition was found, we tested BCG from different vendors for impurity types and quantities. For that, a BCG solution of each vendor was prepared and analyzed as described in Sect. 2.1.2 with the optimal mobile phase composition. To quantitatively compare the purity of each dye, we defined the relative purity of BCG at 280 nm wavelength ( $P_{BCG}$ ), which was calculated as follows:

$$P_{BCG} = \frac{A_{BCG}}{\sum_{i=1}^n A_i} \times 100\% \quad (1)$$

where  $A_{BCG}$  is the area of the BCG peak,  $n$  is the number of peaks, and  $A_i$  is the area of the  $i^{\text{th}}$  peak.

### 2.1.4 Purification of BCG

The purification was performed by the LC system in preparative mode. A 7.5 mg mL<sup>-1</sup> BCG solution was prepared in the mobile phase and 10 mL were injected onto the preparative column. Impurities were separated by isocratic flow (flow rate 31.2 mL min<sup>-1</sup>) with 75:25:0.1 ACN:H<sub>2</sub>O:TFA as mobile phase. The pure BCG was collected manually in a round bottom flask at its retention time of about 52 min. Approximately 90 % of the mobile phase was removed from the BCG eluate using a rotary evaporator, with the final 10 % left to evaporate in a dark open box at room temperature. The pure crystalline dye was transferred to a brown flask for further experiments.

In order to verify the success of the purification, the purified BCG was analyzed by the analytical HPLC procedure as described in Sect. 2.1.2.

## 2.2 Total alkalinity measurements

### 2.2.1 Reagents and instrumentation

Total alkalinity measurements were performed using the novel autonomous analyzer CONTROS HydroFIA<sup>®</sup> TA (Kongsberg Maritime Contros GmbH, Kiel, Germany). Its measurement principle is based on a single-point open-cell titration of the seawater sample with subsequent spectrophotometric pH detection using BCG as indicator (Brelund and Byrne, 1993; Yao and Byrne, 1998; Li et al., 2013; Seelmann et al., 2019). The seawater sample was titrated with 0.1 mol kg<sup>-1</sup> hydrochloric acid (HCl) obtained from Carl Roth and constantly temperature controlled to 25.0 ± 0.1 °C by the systems internal heat exchanger.

The  $A_T$  value of the sample was calculated by the following general equation:

$$\frac{-V_{sw} \times \rho_{sw} \times A_T + V_t \times \rho_t \times C_t}{V_{sw} \times \rho_{sw} + V_t \times \rho_t} = [H^+]_F + [HF] + [HSO_4^-] + [HI^-] \quad (2)$$



where  $V_{sw}$  and  $V_t$  are the volumes of the seawater sample and the added titrant (HCl and BCG solutions), respectively, and  $\rho_{sw}$  and  $\rho_t$  are the densities of the seawater sample and the added titrant, respectively.  $C_t$  is the acid concentration in the combined  
110 titrant solution.  $[H^+]_F$  is the free concentration of hydrogen ions, and  $[HI^-]$  is the concentration of the protonated (i. e. acidic) form of BCG.  $[HF]$  and  $[HSO_4^-]$  are the concentrations of hydrogen fluoride and the bisulfate ion in the seawater sample.  $[H^+]_F$ , or  $pH_F$ , in the sample-titrant mixture is measured spectrophotometrically. Following Breland and Byrne (1993) and Yao and Byrne (1998),  $pH_F$  is described by

$$pH_F = 4.4166 + 0.0005946 \times (35 - S_{mix}) + \log \left( \frac{R - 0.0013}{2.3148 - R \times 0.1299} \right) \quad (3)$$

115 where  $S_{mix}$  is the salinity of the sample-titrant mixture, and  $R$  is the ratio between the absorbances at 444 and 616 nm.

Certified reference material, CRM, (batch 160,  $A_{T,reference} = 2212.44 \pm 0.67 \mu\text{mol kg}^{-1}$ ) was obtained from A. G. Dickson at the Scripps Institution of Oceanography of the University of California, San Diego. The seawater for ~~the-our~~ experiments was prepared by diluting ~~concentrated seawater solutions called a commercially available 8.33-fold concentrate of seawater~~ ("Absolute Ocean" ~~obtained from ATI Aquaristik, ATI Aquaristik~~) with deionized water. Its absolute  $A_T$  value was not impor-  
120 tant as it was only used for mimicking semi-continuous measurement conditions between the references. All total alkalinity measurements were carried out in an air-conditioned laboratory and after the system was "calibrated" with a freshly opened CRM. However, the "calibration" routine conducted by the CONTROS HydroFIA® TA is not a calibration in a true sense. It rather serves the determination of the exact sample volume by utilizing a one point CRM measurement. The seawater sample volume is the only unknown variable of the absolute  $A_T$  determination method (Seelmann et al., 2019).

## 125 2.2.2 Long-term measurements

For the long-term measurements,  $0.002 \text{ mol kg}^{-1}$  BCG solutions were prepared from unpurified BCG (from different vendors) and purified BCG and used as indicator dye in the analyzer. The unpurified dyes (sodium salts) were dissolved in deionized water (DI-water). The purified dye was dissolved in DI-water with sodium hydroxide (NaOH) as additive. The exact amount of NaOH was calculated by the molar ratios and molar masses of BCG and NaOH. This transferred the pure BCG to its  
130 sodium salt and improved its solubility. For both unpurified and purified BCG solutions the ionic strength was kept very low (only created by the dissolved BCG sodium salt itself) in order to realize high concentrations of BCG stock solution. However, the dilution of the sample seawater by the added BCG and HCl solution was ~~considered~~ accounted for in the  $A_T$  calculation procedure.

The prepared seawater sample ( $\approx 25 \text{ L}$ ) was measured more than 300 times with a measuring interval of 15 min, which  
135 took about four days. For monitoring the drift, a freshly opened CRM was measured at the beginning and at the end of this experiment, as well as daily in between. Each of these CRM measurements consisted of five consecutive single measurements.

## 2.2.3 Standard addition experiment

In order to evaluate the impact of impurities on the measuring performance of the system, we carried out a standard addition experiment with each unpurified and the purified BCG. This experiment is the standard validation procedure for evaluating the

140 performance of the analyzer under laboratory conditions. Therefore, a seawater sample (with relatively high  $A_T$ ) was titrated with an HCl solution ( $0.1 \text{ mol kg}^{-1}$ ) to lower its  $A_T$  in five steps. The titration was carried out by adding different precisely known volumes of HCl to a known volume of seawater resulting in five seawater samples with stepwise decreased  $A_T$ . The theoretical  $A_T$  ( $A_{T,\text{titrated}}$ ) was calculated from the volumes of added acid and seawater, the concentration of the acid, and the original  $A_T$  of the seawater. To determine the practical  $A_T$  ( $A_{T,\text{measured}}$ ), each of these samples was repeatedly measured  
145 with the analyzer ( $n = 5$ ).

The precision was determined by averaging the standard deviation ( $\sigma$ ) of each sample measurement. The root mean square error (RMSE) of the linear regression after plotting  $A_{T,\text{measured}}$  vs.  $A_{T,\text{titrated}}$  gave us information about the ~~measuring accuracy~~bias of the method. It was calculated by

$$\text{RMSE} = \pm \sqrt{\frac{1}{n} \times \sum_{i=1}^n (A_{T,\text{fitted},i} - A_{T,\text{measured},i})^2} \quad (4)$$

150 where  $n$  is the number of samples,  $A_{T,\text{fitted},i}$  is the  $i^{\text{th}}$   $A_T$  value calculated with the linear regression equation with  $A_{T,\text{titrated},i}$  as  $x$  variable, where  $A_{T,\text{measured}}$  is the average of the five repeatedly measured  $A_T$  values of each titrated seawater sample. Slope and intercept of this regression were important for the evaluation of linearity and sensitivity. Within the standard validation procedure of the analyzer, these terms must fulfill within their uncertainties the following requirements: Slope = 1; intercept = 0.

## 155 3 Results and discussion

### 3.1 HPLC separation and purification of BCG

#### 3.1.1 Method development

Table 1 summarizes the influence of the different mobile phase compositions on the BCG separation. For saving solvents and time, it is important to keep the time of each HPLC run under 60 min, but, at the same time, with an optimal separation of  
160 BCG from its impurities. Hence, the optimal separation of BCG was achieved with 75:25:0.1 ACN:H<sub>2</sub>O:TFA as mobile phase. ~~There, the~~The pure BCG was eluted from the column as fast as possible (retention time: 52 min) with the best dye-impurity separation.

#### 3.1.2 Comparison of BCG from different vendors

Figure 2 shows the resulting analytical HPLC chromatograms. There, BCG from different vendors shows different types and quantities of impurities. ~~However, the~~The retention time of the pure BCG was 52 min in all chromatograms. Another similarity between all chromatograms was the cluster of several peaks around 3-5 min. Only the peak areas of these peaks strongly differed. As there are no peaks at these retention times in the blank chromatogram, this peak cluster had to be caused by the indicator and not by the solvent. BCG from Acros Organics, Alfa Aesar and Carl Roth showed an intensive peak around

**Table 1.** Mobile phase compositions and their impact on the BCG separation

Mobile phase composition:			BCG separation:	
ACN (%)	H <sub>2</sub> O (%)	TFA (%)	BCG peak (min)	Sufficient separation of impurities
70	30	0.05	no elution <sup>a</sup>	-
70	30	0.10	60	no <sup>b</sup>
70	30	0.15	no elution <sup>a</sup>	-
70	30	0.20	no elution <sup>a</sup>	-
75	25	0.10	52	yes
80	20	0.10	56	yes
85	15	0.10	60	no <sup>b</sup>

<sup>a</sup>Within 60 min run time<sup>b</sup>Impurities found in subsequent blank run**Table 2.** Summary of analytical HPLC of unpurified BCG from different vendors

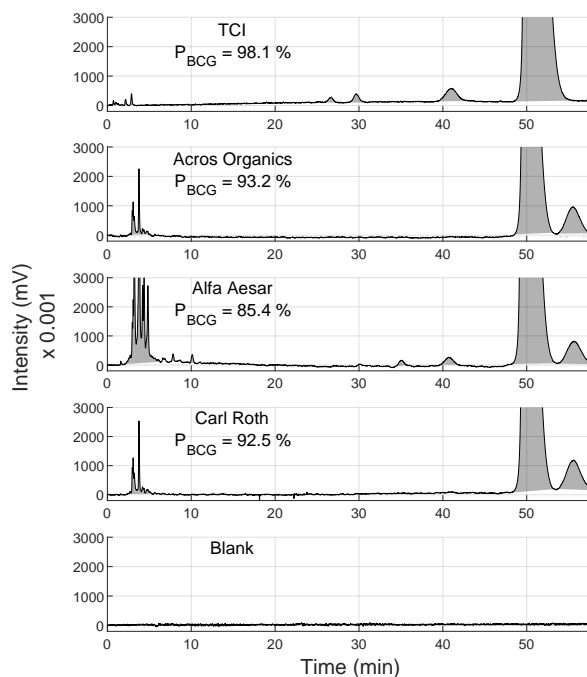
	Acros Organics	Alfa Aesar	Carl Roth	TCI
Number of peaks	3	7	3	5
$P_{BCG}$ (%)	93.2	85.4	92.5	98.1

58 min, which is not present in the BCG from TCI. However, BCG from TCI showed three other small peaks around 26 min, 29 min, and 42 min. Alfa Aesar BCG also showed the 42 min peak in addition to small peaks around 7 min, 10 min, and 35 min. These various quantities of total absorbance at 280 nm resulted in different  $P_{BCG}$ . The calculated  $P_{BCG}$  for each vendor (following Equ. 1) are summarized in Table 2. It has to be taken into account that these quantities are only valid when using an UV detector. Other detectors may result in different purity levels.

### 3.1.3 Purification of BCG

In order to test the effectiveness of the purification method, we decided to purify the least pure BCG from Alfa Aesar. Furthermore, to produce the most pure dye, also BCG from TCI was chosen for purification. The obtained yields were between 60 % and 70 % for both BCGs with around 50 mg purified BCG recovered per injection.

Figure 3 shows the analytical HPLC chromatograms of purified TCI BCG, and Alfa Aesar BCG, respectively. Both chromatograms still show the peak cluster around 3-5 min, but with much smaller areas, especially with purified Alfa Aesar BCG. Furthermore, the 42 min and 58 min peaks of Alfa Aesar BCG could not be totally removed. However, the purity of TCI BCG, and Alfa Aesar BCG improved to 99.6 %, and 99.3 %, respectively. The results are summarized in Table 3.

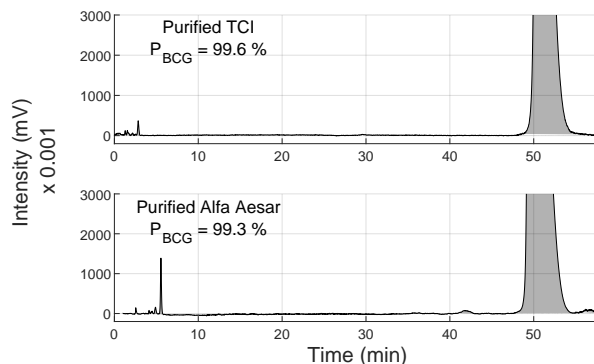


**Figure 2.** Analytical HPLC chromatograms of unpurified BCG from different vendors with their  $P_{BCG}$  and a chromatogram from a solvent injection without BCG (Blank). All peaks are highlighted with gray background color.

**Table 3.** Summary of analytical HPLC of purified BCG

	Purified from	
	TCI	Alfa Aesar
Number of peaks	2	4
$P_{BCG}$ (%)	99.6	99.3

Since the relative purity of Alfa Aesar BCG was improved from 85.4 % to 99.3 %, the success of the purification was proven. Hence, the HPLC purification method developed here is considered sufficient for the nearly full removal of impurities from BCG.



**Figure 3.** Analytical HPLC chromatograms of purified BCG from TCI, and Alfa Aesar with their  $P_{\text{BCG}}$ . All peaks are highlighted with gray background color.

## 185 3.2 Total alkalinity measurements

### 3.2.1 Long-term measurements

During previous studies with the CONTROS HydroFIA<sup>®</sup> TA analyzer, we found that a linear drift to higher  $A_T$  values appears to be the typical behavior of the system (Seelmann et al., 2019). We also found out that the drift is caused by deposits in the optical pathway. As a result, the light intensity decreases and therefore the absorbances at 444 and 616 nm (wavelengths where the acid and base form of BCG have their absorbance maxima) changes in a certain ratio so that the  $A_T$  values increases per measurement. In the present study we wanted to examine the impact of BCG impurities and the usage of purified BCG, respectively, on the drift behavior of the system.

In order to evaluate the drift of the system supposedly caused by impurities of the BCG indicator dye, the bias between the measured  $A_T$  value and the reference  $A_T$  value of the CRM ( $\Delta A_T = A_{T,\text{measured}} - A_{T,\text{reference}}$ ) was plotted vs. the measurement counter. Figure 4 shows the results for  $A_T$  measurements with purified and unpurified TCI BCG, as well as unpurified BCG from Alfa Aesar and Acros Organics. Measurements with purified and unpurified TCI BCG resulted in a linear drift to higher values with the regression equation  $y = (0.0193 \pm 0.0009) \times x + (-0.18 \pm 0.16)$ , and  $y = (0.0317 \pm 0.0004) \times x + (-0.16 \pm 0.10)$ , respectively. However, unpurified Acros Organics and Alfa Aesar BCG showed a non-linear drift to higher values. All  $A_T$  measurements took ~~the analyzers relative uncertainty determined with~~ into account the relative uncertainty of the analyzer, determined as 0.08 % (Seelmann et al., 2019) ~~into account. However, Fig. Figure~~ 4 does not show these uncertainties as they are too small for the scaling of the y-axis.

One important outcome of this experiment is, that the magnitude and shape of the drift directly depends on the purity of the used BCG. The drift caused by purified TCI BCG is reduced by  $0.0124 \mu\text{mol kg}^{-1}$  per measurement with respect to unpurified TCI BCG. This indicates that the drift of the system must be primarily caused by impurities of the BCG indicator and not by the indicator itself as hypothesized in our previous study (Seelmann et al., 2019). However, there is still a remaining

small drift component even with the most pure TCI BCG. Hence, BCG purification appears to significantly reduce but not completely eliminate the observed system drift. ~~However~~For resetting the drift, a flush with ethanol or isopropyl alcohol removes any impurity deposits in the optical pathway caused by the indicator dye~~and, therefore, resets the drift~~. The frequency of these cleanings during long-term deployments can be reduced by using purer dye. But finally, the user of the CONTROS HydroFIA<sup>®</sup> TA analyzer decides the cleaning interval as its frequency depends on the certain application of the system and how often measurements are conducted. Furthermore, there is a dependency on the measured water matrix as well, e.g. high turbidity coastal water requires more often cleanings than open ocean water. We can only make recommendations based on our experiences with the analyzer. During our field deployments of the analyzer (not part of this study), we ran a cleaning procedure using ethanol right before a new "calibration" of the system with CRM. As our analyzer measured around 1000  $A_T$  values per month, we carried out an ethanol flush with a subsequent calibration on monthly ~~bases~~-basis. We also experienced that the subsequent drift correction is entirely manageable up to a maximum  $\Delta A_T$  of approximately  $30 \mu\text{mol kg}^{-1}$  (as observed during our field deployments, not part of this study). This  $\Delta A_{T, \text{max}}$  can be used as a guidance for determining the cleaning frequency.

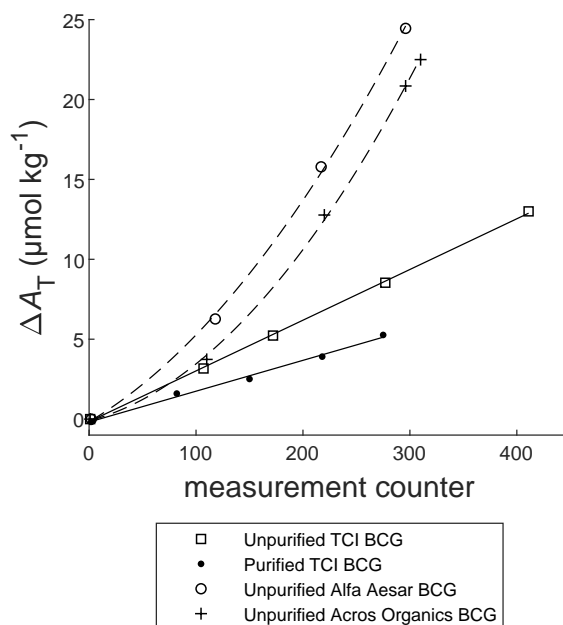
Another important outcome is that the shape of the drift differed with the amount of impurities. Below a certain purity grade (between 93.2 % and 98.1 %), the drift behavior appears to change from linear to non-linear. However, for unattended long-term installations of the CONTROS HydroFIA<sup>®</sup> TA analyzer it is highly preferable to have a linear drift. Under this condition, the correction during the post-processing of the data is easier and the necessary reference measurements can be reduced to a pre- and post-deployment measurement. Furthermore, the upper limit of the analyzer's working range will be reached faster with a non-linear increase of the  $A_T$  values per measurement. Hence, there is the risk, that the measured  $A_T$  values are rendered useless towards the end of a long-term deployment.

Due to the nearly similar drift behavior of Acros Organics and Alfa Aesar BCG, we also hypothesize that the observed non-linear behavior was mainly caused by the impurity with the retention time around 58 min, which is only present in BCG from Acros Organics, Carl Roth, and Alfa Aesar. Additional tests with the Carl Roth indicator supported the hypothesis (results not shown). This certain impurity might be a molecule with a higher adsorption tendency to the glass wall of the cuvette compared to other impurities. If the used indicator dye contains this impurity type, the magnitude and shape of the drift is mainly driven by the presence of this molecule than by the BCG purity itself. As a consequence, the usage of BCG indicators containing this impurity should be avoided especially during long-term deployments.

Additional to the impacts on the drift, we also experienced, that the frequency of system cleanings has to be increased when using BCG with low purity. For unattended long-term deployments, this must be taken into account.

### 235 3.2.2 Standard addition experiment

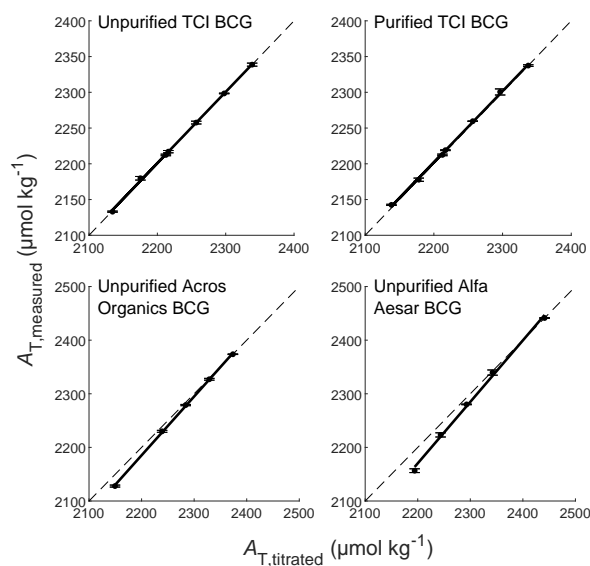
~~For better presentation, the results and discussions for the various BCG types are divided into two groups: 1) For~~ After this experiment was conducted, we experienced that "high-purity" BCG ( $P_{\text{BCG}} > 98 \%$ ) , and 2) and "low-purity" BCG ( $P_{\text{BCG}} < 94 \%$ ) . ~~At the time of this study, we showed different results. Hence, we decided to divide the results and discussion~~



**Figure 4.** Bias ( $\Delta A_T$ ) between measured  $A_T$  and reference value of the CRM as a function of the measurement counter of the CONTROS HydroFIA<sup>®</sup> TA analyzer, where filled circles, open squares, crosses, and open circles represent the average of five repeated measurements made with purified BCG (TCI), unpurified TCI BCG, unpurified Acros Organics BCG, and unpurified Alfa Aesar BCG, respectively. The solid lines are the linear regressions of the associated measurement points. The dashed lines represent a non-linear regression.

[section of this experiment into two groups. Which type of BCG belongs to which group can be found in Table 4.](#) We cannot say anything about the behavior of BCG with  $98\% > P_{BCG} > 94\%$ , because none of the tested dyes fall in this range.

The results of the standard addition experiment carried out with purified and unpurified TCI BCG ("high-purity" BCG) are shown in Fig. 5. By plotting  $A_{T,measured}$  vs.  $A_{T,titrated}$ , purified and unpurified TCI BCG show a linear equation of  $y = (0.996 \pm 0.013) \times x + (11 \pm 29)$ , and  $y = (0.997 \pm 0.012) \times x + (7 \pm 26)$ , respectively. Both correlations satisfy the quality requirements (slope = 1, intercept = 0) within their uncertainty, and they were statistically indistinguishable. Hence, the sensitivity and linearity of these measurements are considered satisfactory. The evaluation of precision and [accuracy bias](#), which is summarized in Table 4 revealed no significant differences between measurements. Furthermore, both [accuracy biases](#) were in full agreement with previous laboratory performance characterizations of the system (Seelmann et al. (2019):  $\pm 1.0 \mu\text{mol kg}^{-1}$ ) and with the requirements of Dickson et al. (2007) for standard open-cell  $A_T$  titrators for which an [accuracy-overall bias](#) of  $\pm 2 \mu\text{mol kg}^{-1}$  is required. However, the requirement for precision (standard open-cell titrator:  $\pm 1 \mu\text{mol kg}^{-1}$ ) were not fully achieved, but both results are entirely comparable to our previous laboratory performance characteristic (Seelmann et al.



**Figure 5.**  $A_{T,\text{measured}}$  as a function of  $A_{T,\text{titrated}}$  of each titration step measured with purified and unpurified TCI BCG as well as unpurified BCG from Acros Organics and Alfa Aesar. The black filled circles represent the average of five repeated measurements for each sample with their standard deviations ( $\sigma$ ) as errorbar. The black solid lines indicate the linear fit of the data points. The black dashed lines indicate the 1-to-1 line of these plots.

(2019):  $\pm 1.5 \mu\text{mol kg}^{-1}$ ). Consequently, above a relative purity grade of 98 % no negative influence of indicator impurities on the measurement performance of the analyzer could be identified.

In-contrast, "low-purity" indicators behaved totally-completely different. The results of the standard addition experiment carried out with unpurified BCG from Acros Organics and Alfa Aesar are shown in Fig. 5 with their linear equations of  $y = (1.097 \pm 0.013) \times x + (-228 \pm 29)$ , and  $y = (1.147 \pm 0.036) \times x + (-352 \pm 82)$ , respectively. Clearly, these correlations were not satisfactory, and statistically different to the correlation of purified BCG. Hence, these "low-purity" dyes do not show the sensitivity and linearity behavior that is required for most accurate measurements with the analyzer. Table 4 shows that measurements with Acros Organics BCG ( $P_{\text{BCG}} = 93.2 \%$ ) still fell within acceptable ranges regarding precision and accuracy bias requirements. However, measurements using Alfa Aesar BCG ( $P_{\text{BCG}} = 85.4 \%$ ) did not meet the quality requirements.

Summing up, we can state that the uncertainty of  $A_T$  measurements only deteriorates significantly for a BCG purity grade below 94 %. Indicator dyes with  $P_{\text{BCG}} > 98 \%$  provide  $A_T$  measurements with a quality comparable to these measured with purified BCG. These findings partially support other studies dealing with different purified pH indicators for spectrophotometric pH (Yao et al., 2007; Liu et al., 2011; DeGrandpre et al., 2014; Lai et al., 2016) and  $A_T$  measurements (Nand and Ellwood, 2018). There, indicator purification always led to an improvement in measurement precision. Under the scope of this study, we proved that purification of BCG is not necessary to improve the quality of the  $A_T$  measurements with the CONTROS



**Table 4.** Precision and ~~accuracy~~-bias of unpurified and purified BCG

	"High-purity" BCG:		"Low-purity" BCG:		Requirements for standard open-cell titrators <sup>a</sup>	Typical performance of the analyzer <sup>b</sup>
	unpurified TCI	purified TCI	unpurified Acros Organics	unpurified Alfa Aesar		
Precision $\sigma$ ( $\mu\text{mol kg}^{-1}$ )	$\pm 1.6$	$\pm 1.5$	$\pm 1.4$	$\pm 2.7$	$\pm 1$	$\pm 1.5$
<del>Accuracy</del> - <u>RMSE</u> <u>Bias (RMSE)</u> ( $\mu\text{mol kg}^{-1}$ )	$\pm 1.2$	$\pm 1.1$	$\pm 1.7$	$\pm 5.3$	$\pm 2$	$\pm 1.0$

<sup>a</sup>Dickson et al. (2007); <sup>b</sup>Seelmann et al. (2019)

HydroFIA<sup>®</sup> TA analyzer as long as the used BCG is purer than 98 %. The reason for the deviation to other studies lies in the measuring principle of the system. Before starting measurements with newly prepared solutions, it is obligatory to "calibrate" the system by measuring a CRM. ~~During this "calibration", the~~ However, this procedure is not a calibration in the real sense, as the method has an absolute character. During this routine, the exact volume of the analyzers internal seawater sample loop,  $V_{\text{SW}}$  is determined being the only unknown variable within this method. Hence, all inevitable uncertainties (including impurities of the indicator) are combined in  $V_{\text{SW}}$  and thereby taken into account for subsequent  $A_{\text{T}}$  measurements. The present results prove that this ~~"calibration"~~ procedure is able to compensate any influences of indicator impurities on the ~~precision and accuracy~~ measurement quality up to an impurity level of 2 %. Consequently, the usage of "low-purity" BCG is not recommended.

## 4 Cost-benefit analysis

### 275 4.1 Measurements with purified vs. unpurified BCG

This study proves that an HPLC purification of BCG is entirely feasible. But is the purification of BCG worths the efforts and costs involved? To answer this question we compare the costs incurred and the benefits gained for  $A_{\text{T}}$  measurements with the CONTROS HydroFIA<sup>®</sup> TA analyzer. Due to the relatively long HPLC run time of 60 min and a flow rate of  $31.2 \text{ mL min}^{-1}$ , the purification method needs about 1.5 L of ACN per run (including pre- and post-flushes). To carry out the long-term and standard addition experiment for this study (around 500 measurements), approximately 144 mg of purified BCG were needed. Hence, with a yield of around 50 mg pure BCG per purification run, a minimum of three injections was necessary. However, for long-term measurement campaigns with the analyzer, the typical volume of BCG solution is 500 mL, which is sufficient for at least 2300 measurements. This would need 700 mg of purified BCG, which corresponds to a minimum of 14 purification runs and 21 L of ACN. ACN with HPLC grade is a relatively expensive chemical, and it must be ~~probably~~-appropriately disposed. This causes additional costs. Furthermore, the whole purification process takes about a full working day per run.

Rough calculations on the actual costs per measurement with the CONTROS HydroFIA<sup>®</sup> TA analyzer revealed that indicator purification would approximately double measurement costs. The calculation for measurements with unpurified indicator are

based on ready-to-use 500 mL cartridges for both chemicals (HCl and BCG) ordered from Kongsberg Maritime Contros GmbH without any preparation effort for the user.

290 To overcome this high increase in measurement costs, there could be the possibility to develop a flash chromatography (FC) purification method for BCG as described for mCP by Patsavas et al. (2013a) to increase the yield of purified dye per run. According to the method description in this publication (Patsavas et al., 2013a), the solvent consumption of both methods (FC and HPLC) per purification run is approximately the same. Provided the FC method would increase the yield 3.5 times as it was described for mCP, only 4 injections will be necessary to produce enough purified BCG for a long-term deployment  
295 with at least 2300 measurements. The estimated measurement costs for such a FC method would be approximately a third of these for measurements using BCG purified by HPLC. Hence, if BCG purification would be necessary, the FC method would be the more cost-effective choice. However, it has to be taken into account that the calculations for these measurement costs (especially for the FC method) are just theoretically estimated and may differ from reality depending on availability of resources and equipment. Furthermore, a FC purification method for BCG is so far not developed and validated, which means  
300 additional costs and working time.

Finally, if we compare the purified BCG with "high-purity" BCG like from TCI, the only benefit gained from the purification is a reduced drift per  $A_T$  measurement. However, as long as the drift pattern is linear, its actual slope is irrelevant as it can be easily corrected by regular reference measurements. Furthermore, there is no improvement in the measurement quality (precision and ~~accuracy~~bias) as long as the impurity level is 2 % or below. Since the drift behavior cannot fully overcome, it  
305 seems not worth the effort to purify BCG for  $A_T$  measurements with the CONTROS HydroFIA<sup>®</sup> TA analyzer.

~~However, the~~The types and quantities of impurities can nevertheless have a strong influence on measurement quality in unattended long-term applications of the system as it was shown before (e.g. change of the drift behavior, non fulfillment of the quality requirements). Hence, the purity of the used BCG is not unimportant at all. To achieve the best long-term measurement experience with the analyzer it is not necessary to use purified BCG, as the purest available indicator (e.g. BCG from TCI)  
310 generate fully satisfying quality results. Users of the CONTROS HydroFIA<sup>®</sup> TA should take the consequences of indicator impurities into account when choosing their BCG supplier. From this perspective, it would be beneficial to invest into higher purity indicator avoiding the issues described above. If applicable, an HPLC analysis of the used indicator following the here described analytical method can show any types and quantities of impurities. However, if there is no HPLC available, long-term laboratory measurements as described here can help to evaluate whether the purchased indicator is suitable or not by evaluating  
315 the drift behavior. As there could be batch to batch variability in purity, the drift pattern should be also assessed for each batch of BCG provided by the same supplier.

## 4.2 BCG characterization

Most of the studies dealing with purification of indicator dyes for spectrophotometric seawater pH measurements conducted a subsequent characterization of the purified indicator (e.g. Liu et al., 2015; Patsavas et al., 2013b; Nand and Ellwood, 2018).  
320 Due to impurity impacts, coefficients and constants of purified indicators may be different to these of unpurified dyes. During our work with purified BCG, we decided to forgo of an indicator characterization. There were two reasons for this decision:

1. Li et al. (2013) investigated the impact of different BCG characteristics found in the literature on spectrophotometric  $A_T$  measurements and concluded that the influences are insignificant also with regard to possible impurities. They justified this conclusion with the calibration of the system using CRM. The CONTROS HydroFIA<sup>®</sup> TA analyzer follows a similar measurement principle as the analyzer described by Li et al. (2013) and also conducted a calibration routine. Therefore, any uncertainties regarding the coefficients are taken into account for subsequent measurements.
2. The measurement quality using both purified and unpurified "high-purity" BCG were fully satisfying and met the quality requirements for  $A_T$  measurements. Furthermore, both uncertainties did not significantly differ from each other.

Finally, we concluded that a characterization of purified BCG would not improve the measurement quality at all and therefore decided to not conduct it.

## 5 Conclusions

~~In this study, we~~ We successfully developed an HPLC purification method for BCG and subsequently tested the impact of using the purified and unpurified dye on measurements with a novel autonomous analyzer for seawater  $A_T$ , the CONTROS HydroFIA<sup>®</sup> TA.

~~HPLC analyses revealed that types and quantities of the impurities differed for each vendor. We tested indicator dyes from four different vendors with a resulting BCG purity ranging from 85.4~~

~~Taking all the achieved results into account, we conclude that a purification of BCG is not strictly recommended to carry out high-quality measurements with the CONTROS HydroFIA<sup>®</sup> TA analyzer. But the usage of "high-purity" BCG ( $P_{BCG} > 98\%$  to 98.1, e.g. from TCI) is highly recommended to avoid a non-linear drift behavior and resulting loss of measurement quality as it was observed with "low-purity" BCG ( $P_{BCG} < 94\%$ . After HPLC purification, the purity was improved to between 99.3 and 99.6, depending on BCG vendor.~~

~~Long-term measurements with the total alkalinity analyzer in the laboratory revealed that the systems drift behavior was strongly related to the purity of the used indicator and the type of containing impurities. The purer the BCG, the smaller was the  $A_T$  drift per measurement. However, the drift behavior cannot be overcome by purifying the used indicator dye. Furthermore, BCG containing a certain impurity type with a retention time of around 58 min showed a~~ Users of the CONTROS HydroFIA<sup>®</sup> TA analyzer should take these recommendations into account, if they want to prepare the BCG solution on their own. A preceding HPLC analysis of the indicator dye would be the preferred approach to test the BCG purity and avoid a loss of analytical performance. BCG indicator dyes showing a relatively big peak after the BCG peak should be avoided, because their usage results in a non-linear drift behavior, which can cause several problems during long-term measurement campaigns by pattern. It must be noted that modified HPLC methods (e.g. exceeding the upper working range limit, difficult data correction during post-processing, higher cleaning frequency or total loss of measuring capacity during long campaigns. Therefore, indicators containing this impurity should not be used for with a different mobile phase composition, detector or column) may result in altered peak patterns or relative BCG purities. However, not every user of the CONTROS

HydroFIA® TA has the facilities for such HPLC analyses. In case of any doubts, the compatibility of the purchased BCG can be easily tested by applying the laboratory long-term deployments of the analyzer.

Laboratory performance characterization experiments with purified and unpurified BCG that only contains up to 2 measurement experiment explained in this study. Dyes resulting in a linear drift pattern can be used without any concern providing the cleaning intervals are performed regularly to limit the absolute drift to  $< 30 \mu\text{mol kg}^{-1}$ . Furthermore, the measurement quality should be monitored on regular basis, especially if the BCG solution decomposes over time. These tests should be also conducted with a new batch of BCG from the purification on either linearity and sensitivity or the measurement quality (precision and accuracy) of the system. The obligatory "calibration" routine of the analyzer after each change of indicator dye solution was sufficient to compensate any impurity effects. If the dye contains more than 6 impurities, the measurements do not fulfill the linearity and sensitivity requirements. With even higher amounts of impurities, the analyzer cannot meet the quality requirements as they are demanded within the oceanographic communities (Diekxson et al., 2007; Seelmann et al., 2019).

Taking all these results into account, a cost-benefit analysis revealed that a purification of BCG is not strictly recommended to carry out high-quality measurements with the used analyzer. But the usage of "high-purity" BCG (e.g. from TCI) is highly recommended to avoid the previously identified issues. Furthermore, no characterization of the purified BCG was carried out same vendor, because there could be batch-to-batch variabilities in purity.

*Author contributions.* KS: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - Original Draft, Visualization; MG: Methodology, Resources, Writing - Review & Editing; SA: Conceptualization, Methodology, Writing - Review & Editing; AK: Conceptualization, Resources, Writing - Review & Editing, Supervision, Funding acquisition

*Competing interests.* At the time of this study the co-author S. Aßmann was an employee of Kongsberg Maritime Contros GmbH (Kiel, Germany), which commercialized the used analyzer. However, its measurement principle follows conceptual design and dedicated studies performed during his doctorate in our working group. Apart from this there are no conflicts of interests to declare.

*Acknowledgements.* This work was supported by the European Union's Horizon 2020 Research and Innovation Programme (AtlantOS) (grant number: 633211).

## References

- 380 Aßmann, S., Frank, C., and Körtzinger, A.: Spectrophotometric high-precision seawater pH determination for use in underway measuring systems, *Ocean Science*, 7, 597–607, <https://doi.org/10.5194/os-7-597-2011>, 2011.
- Bellerby, R. G., Olsen, A., Johannessen, T., and Croot, P.: A high precision spectrophotometric method for on-line shipboard seawater pH measurements: the automated marine pH sensor (AMpS), *Talanta*, 56, 61–69, [https://doi.org/10.1016/S0039-9140\(01\)00541-0](https://doi.org/10.1016/S0039-9140(01)00541-0), 2002.
- Breland, J. A. and Byrne, R. H.: Spectrophotometric procedures for determination of sea water alkalinity using bromocresol green, *Deep Sea Research Part I: Oceanographic Research Papers*, 40, 629–641, [https://doi.org/10.1016/0967-0637\(93\)90149-W](https://doi.org/10.1016/0967-0637(93)90149-W), 1993.
- 385 Byrne, R. H.: Standardization of standard buffers by visible spectrometry, *Analytical Chemistry*, 59, 1479–1481, <https://doi.org/10.1021/ac00137a025>, 1987.
- Byrne, R. H. and Breland, J. A.: High precision multiwavelength pH determinations in seawater using cresol red, *Deep Sea Research Part A. Oceanographic Research Papers*, 36, 803–810, [https://doi.org/10.1016/0198-0149\(89\)90152-0](https://doi.org/10.1016/0198-0149(89)90152-0), 1989.
- Clayton, T. D. and Byrne, R. H.: Spectrophotometric seawater pH measurements: total hydrogen ion concentration scale calibration of m-cresol purple and at-sea results, *Deep Sea Research Part I: Oceanographic Research Papers*, 40, 2115–2129, [https://doi.org/10.1016/0967-0637\(93\)90048-8](https://doi.org/10.1016/0967-0637(93)90048-8), 1993.
- DeGrandpre, M. D., Spaulding, R. S., Newton, J. O., Jaqueth, E. J., Hamblock, S. E., Umansky, A. A., and Harris, K. E.: Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies, *Limnology and Oceanography: Methods*, 12, 830–839, <https://doi.org/10.4319/lom.2014.12.830>, 2014.
- 395 Dickson, A. G., Sabine, C. L., and Christian, J. R.: Guide to best practices for ocean CO<sub>2</sub> measurements, *PICES Special Publications* 3, 2007.
- Douglas, N. K. and Byrne, R. H.: Achieving accurate spectrophotometric pH measurements using unpurified meta-cresol purple, *Marine Chemistry*, 190, 66–72, <https://doi.org/10.1016/j.marchem.2017.02.004>, 2017.
- King, D. and Kester, D. R.: Determination of seawater pH from 1.5 to 8.5 using colorimetric indicators, *Marine Chemistry*, 26, 5–20, 400 [https://doi.org/10.1016/0304-4203\(89\)90061-3](https://doi.org/10.1016/0304-4203(89)90061-3), 1989.
- Lai, C.-Z., DeGrandpre, M. D., Wasser, B. D., Brandon, T. A., Clucas, D. S., Jaqueth, E. J., Benson, Z. D., Beatty, C. M., and Spaulding, R. S.: Spectrophotometric measurement of freshwater pH with purified meta-cresol purple and phenol red, *Limnology and Oceanography: Methods*, 14, 864–873, <https://doi.org/10.1002/lom3.10137>, 2016.
- Li, Q., Wang, F., Wang, Z. A., Yuan, D., Dai, M., Chen, J., Dai, J., and Hoering, K. A.: Automated Spectrophotometric Analyzer for Rapid 405 Single-Point Titration of Seawater Total Alkalinity, *Environ. Sci. Technol.*, 47, 11 139–11 146, <https://doi.org/10.1021/es402421a>, 2013.
- Liu, X., Patsavas, M. C., and Byrne, R. H.: Purification and characterization of meta-cresol purple for spectrophotometric seawater pH measurements, *Environ. Sci. Technol.*, 45, 4862–4868, <https://doi.org/10.1021/es200665d>, 2011.
- Liu, X., Byrne, R. H., Lindemuth, M., Easley, R., and Mathis, J. T.: An automated procedure for laboratory and shipboard spectrophotometric measurements of seawater alkalinity: Continuously monitored single-step acid additions, *Marine Chemistry*, 174, 141–146, 410 <https://doi.org/10.1016/J.MARCHEM.2015.06.008>, 2015.
- Millero, F. J.: The Marine Inorganic Carbon Cycle, *Chemical Reviews*, 107, 308–341, <https://doi.org/10.1021/cr0503557>, 2007.
- Nand, V. and Ellwood, M. J.: A simple colorimetric method for determining seawater alkalinity using bromophenol blue, *Limnology and Oceanography: Methods*, 16, 401–410, <https://doi.org/10.1002/lom3.10253>, 2018.

- Patsavas, M. C., Byrne, R. H., and Liu, X.: Purification of meta-cresol purple and cresol red by flash chromatography: Procedures for ensuring accurate spectrophotometric seawater pH measurements, *Marine Chemistry*, 150, 19–24, <https://doi.org/10.1016/j.marchem.2013.01.004>, 2013a.
- Patsavas, M. C., Byrne, R. H., and Liu, X.: Physical–chemical characterization of purified cresol red for spectrophotometric pH measurements in seawater, *Marine Chemistry*, 155, 158–164, <https://doi.org/10.1016/j.marchem.2013.06.007>, 2013b.
- Robert-Baldo, G. L., Morris, M. J., and Byrne, R. H.: Spectrophotometric determination of seawater pH using phenol red, *Analytical Chemistry*, 57, 2564–2567, <https://doi.org/10.1021/ac00290a030>, 1985.
- Seelmann, K., Aßmann, S., and Körtzinger, A.: Characterization of a novel autonomous analyzer for seawater total alkalinity: Results from laboratory and field tests, *Limnology and Oceanography: Methods*, 17, 515–532, <https://doi.org/10.1002/lom3.10329>, 2019.
- Tapp, M., Hunter, K., Currie, K., and Mackaskill, B.: Apparatus for continuous-flow underway spectrophotometric measurement of surface water pH, *Marine Chemistry*, 72, 193–202, [https://doi.org/10.1016/S0304-4203\(00\)00081-5](https://doi.org/10.1016/S0304-4203(00)00081-5), 2000.
- 415 Yao, W. and Byrne, R. H.: Simplified seawater alkalinity analysis: Use of linear array spectrometers, *Deep Sea Res., Part I*, 45, 1383–1392, [https://doi.org/10.1016/S0967-0637\(98\)00018-1](https://doi.org/10.1016/S0967-0637(98)00018-1), 1998.
- 420 Yao, W., Liu, X., and Byrne, R. H.: Impurities in indicators used for spectrophotometric seawater pH measurements: Assessment and remedies, *Marine Chemistry*, 107, 167–172, <https://doi.org/10.1016/j.marchem.2007.06.012>, 2007.