Cup of Carbon: Smartphone-based analysis of dissolved organic carbon in water for use in citizen science

Michael R Muir, Adrian M Bass, Kenny Galt, Kerry Morrison , Lewis Robertson , Emily Taylor

Supplementary Information

SI. 1 Image of the Cup of Carbon “shiny” computer app

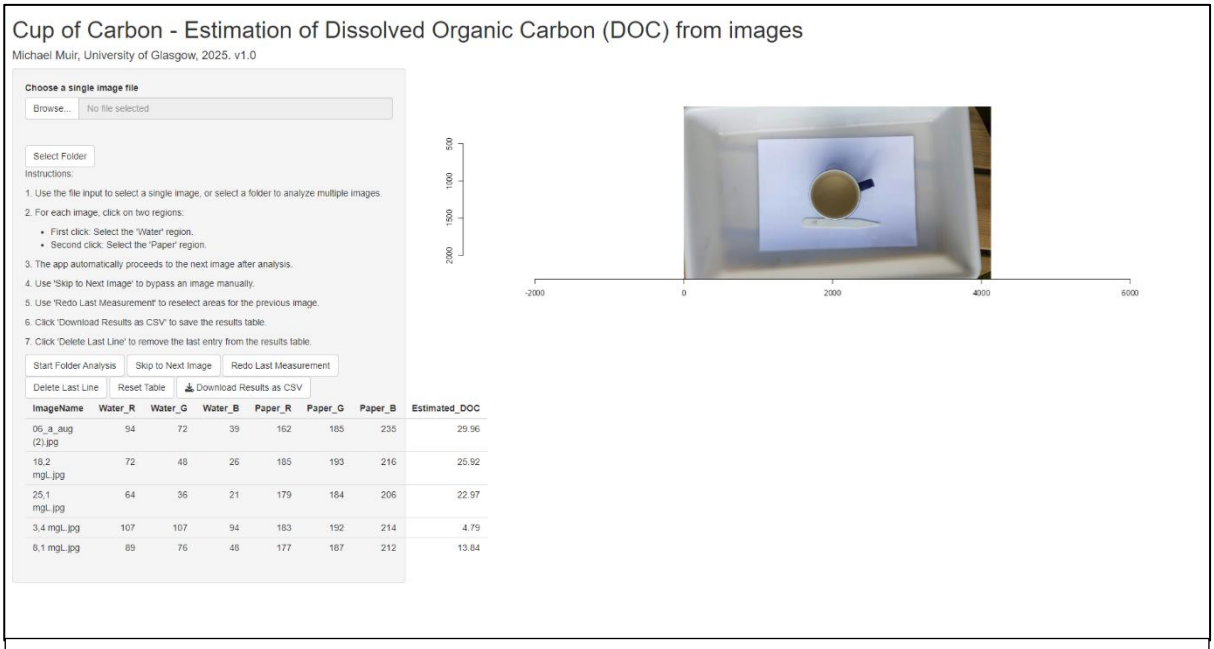

Figure S1. Screenshot of the Cup of Carbon “shiny” computer app. The app allows a user to upload images of a cup containing water sample or navigate to a folder of images. The user then clicks the center of the cup and a section of the white paper background. The app stores the RGB data for the portions of the image clicked by the user in a table, and Estimates the DOC concentration based on the built in calibration.

SI. 2 Verification of water DOC concentration by TOC analysis

The DOC concentration of samples was determined directly from UV/Vis spectra using the dual wavelength model of Carter et al., (2012), which estimates the DOC concentration of samples based on the absorbance at 270 and 350 nm using a model parametrised with DOC and absorbance data from ~1700 surface water samples from Europe and North America. Eleven samples collected in August were also analysed using a total organic carbon (TOC) analyser (Thermolox, Sercon) to ensure the accuracy of the UV/Vis modelled DOC concentrations. The results of the UV/Vis method showed excellent agreement with the TOC analyser, with a regression coefficient of $R^2 = 0.9987$ (figure S2) and Lin's Concordance Correlation Coefficient (CCC) of r_c=0.997 (Lin, 1989; Muir & Innes, 2024). The UV/Vis method gave results which were slightly higher than those measured by TOC, with an average recovery of 104.30 %, however all the DOC concentrations measured by the UV/vis method were within the 95 % confidence intervals (CI) reported by Carter et al., (2012), of 0.9 mg L^{-1} for concentrations from 0 – 5 mg L^{-1}, 2 mg L^{-1} for concentrations from 5 – 20 mg L^{-1}, and 4 mg L^{-1} for concentrations from 20 – 80 mg L^{-1} (table S1).

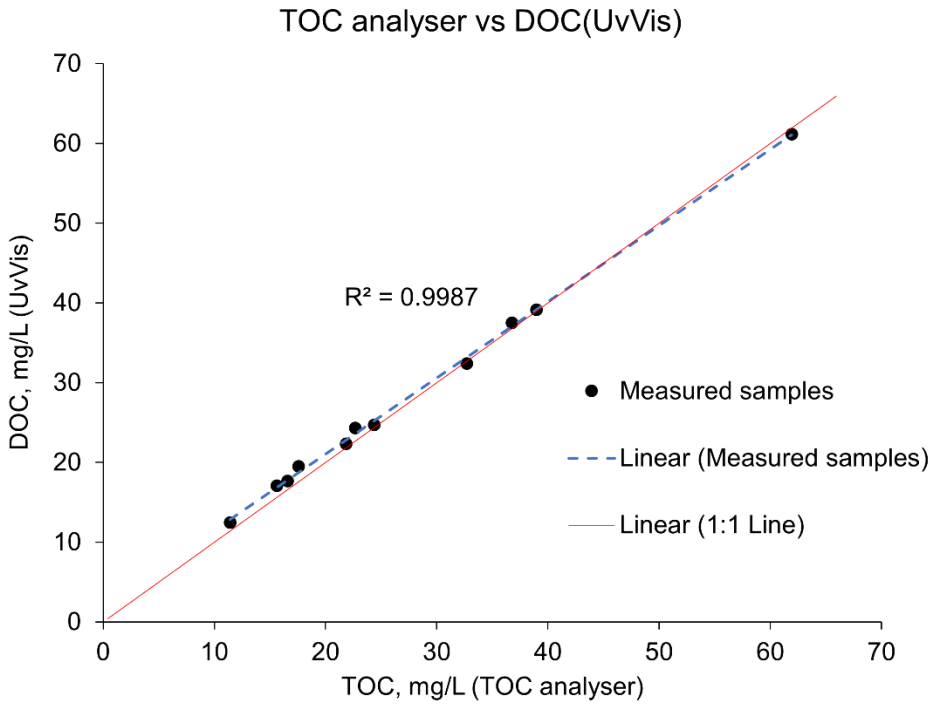

Figure S2. Scatter plot showing the comparison of the DOC concentration measured by the direct UV/Vis spectrophotometer method compared to the reference TOC analyser method. The 1:1 line is shown in red, and the blue dashed line shows the linear regression between the two methods.

Table S1. Results of the analysis of DOC by the UV/Vis and TOC analyser methods

TOC analyser (mg L^{-1})	**DOC, UV/vis (mg L^{-1})**	**Difference (mg L^{-1})**	**95 % CI from Carter et al., (2012)**	**Recovery (%)**
16.57	17.67	1.10	± 2 mg L^{-1}	106.61
11.41	12.48	1.07	± 2 mg L^{-1}	109.40
17.59	19.50	1.91	± 2 mg L^{-1}	110.88
21.84	22.35	0.52	± 4 mg L^{-1}	102.38
32.70	32.42	-0.28	± 4 mg L^{-1}	99.13
22.68	24.33	1.65	± 4 mg L^{-1}	107.26
15.64	17.05	1.42	± 2 mg L^{-1}	109.06
36.78	37.50	0.71	± 4 mg L^{-1}	101.94
38.99	39.16	0.17	± 4 mg L^{-1}	100.44
61.92	61.15	-0.77	± 4 mg L^{-1}	98.76
24.38	24.73	0.35	± 4 mg L^{-1}	101.44
			Average	**104.30 %**
			Standard deviation	**4.43 %**

SI. 3 Comparison between the DOC measurement in filtered and unfiltered samples

Suspended particulate matter (turbidity) can also influence water colour (Jerome et al., 1994), therefore filtration (0.45 µm syringe filters, VWR) followed by UV/Vis absorbance analysis was applied to test the potential influence of filtration on the DOC concentration and absorbance characteristics of the collected water samples. Figure S3 shows example absorbance spectra from the water sample collected at site 5 in March 2023. The black solid line shows the absorbance of the unfiltered sample over a wavelength range of 800 – 200 nm, and the dashed red line shows the absorbance of the same sample after filtration (0.45 µm).

The filtered water spectrum shows only very moderate changes compared to the unfiltered water spectrum, indicating that the light absorbing material responsible for the dark colour of the water is predominantly truly dissolved in the solution phase. These spectra are typical of all the samples analysed as part of this project which showed minimal differences between the spectra of filtered and unfiltered samples.

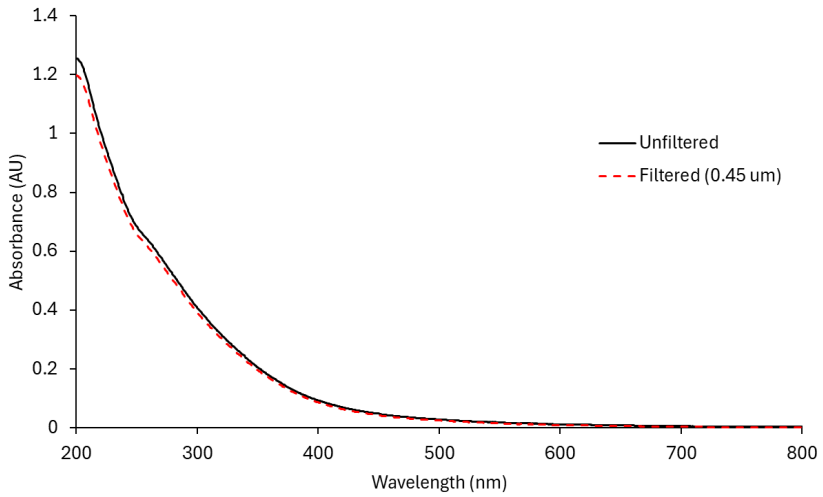

Figure S3. Example absorbance spectra of the unfiltered and filtered water sample collected from sample site 5 in March 2023. The black solid line shows the absorbance spectrum of the unfiltered sample, and the red dashed line shows the spectrum of the filtered sample. The x axis shows wavelength in nm and the y axis shows the absorbance in absorbance units (AU).

Figure S4 shows the comparison between the DOC concentration estimated for filtered vs unfiltered samples using the two-component model of Carter et al., (2012). The filtered samples generally show a very minor changes in DOC compared to the unfiltered samples. However, there is one exception to this which is the sample collected at site 5 in August 2023 (highlighted with the arrow on figure S2) which showed a decrease in DOC concentration from 34.6 mg L^{-1}, to 26.3 mg L^{-1} after filtration. The analysis of DOC concentration in unfiltered vs filtered samples showed that the DOC concentration changed by an average of only 0.08 ± 1.04 mg/L or 1.36 ± 6.20 % after filtration. The water samples had very low influence from suspended particulate material, therefore the DOC is predominantly in the truly dissolved fraction, and the colour of the water can be reliably attributed to the DOC concentration rather than suspended particulate matter.

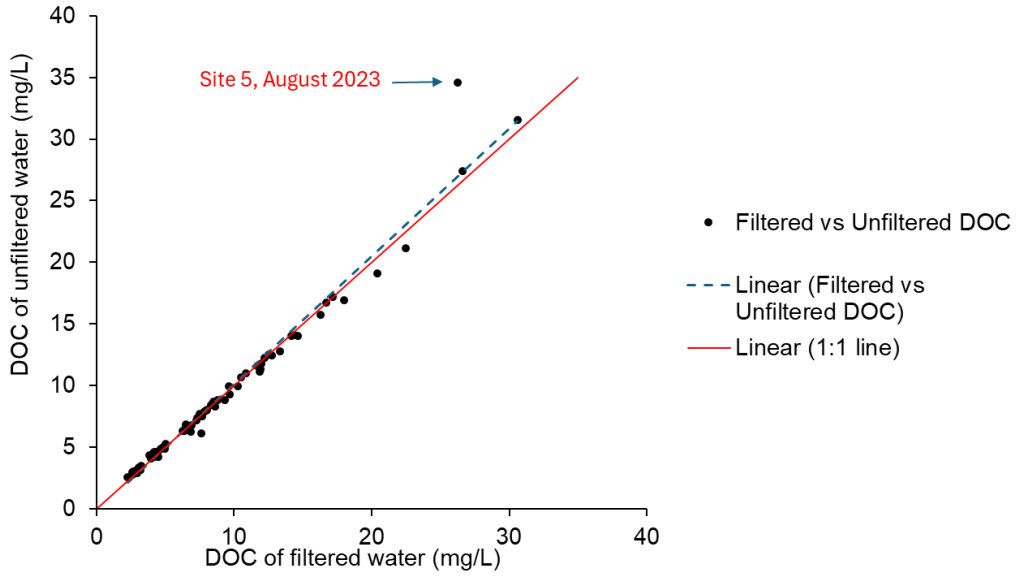

Figure S4. Scatter plot showing the comparison of the DOC concentration in filtered (0.45 µm syringe filter) vs unfiltered water samples measured by the direct UV/Vis spectrophotometer method of Carter et al. (2012). The 1:1 line is shown in red, and the blue dashed line shows the linear regression between the filtered and unfiltered samples.

SI. 4 Boxplots of the blank, medium and high DOC samples analysed during method development

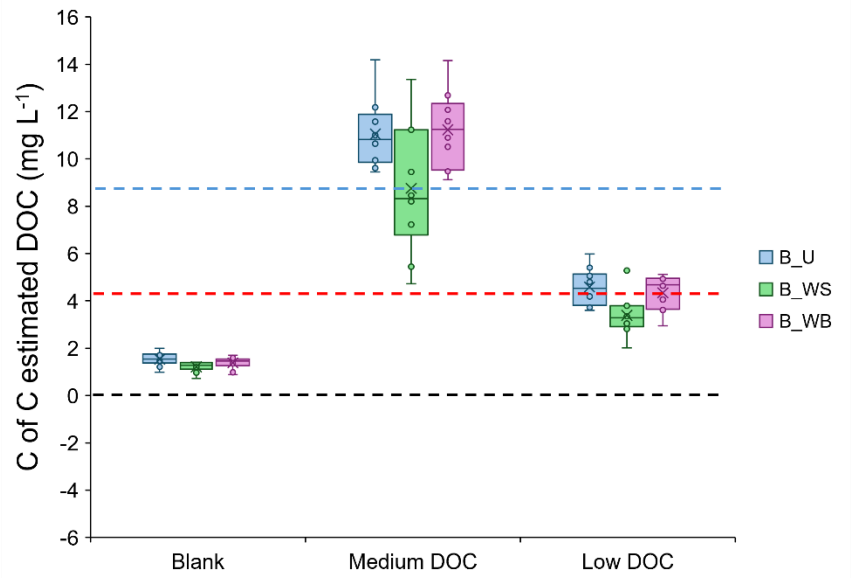

Figure S5. Boxplots showing the results of replicate measurements of blank (0 mg L^{-1}, represented by the black dashed line), medium (8.63 mg L^{-1}, represented by the blue dashed line) and low (4.27 mg L^{-1}, represented by the red dashed line) concentration DOC samples using the B_U (blue), B_{WS} (green), and B_{WB} (pink) digital image analysis options.

Table S2. Results of the ANOVA and post-hoc Tukey's HSD test to compare the differences between the blank, Low and medium DOC samples using each of the DIA options (B_U, B_{WS}, and B_{WB}).

Method	Comparison	Mean difference	95% CI (lwr, upr)	p-value	Sig.
B_U	Low DOC - Blank	3.06	1.99 – 4.13	<0.001	***
B_U	Medium DOC – Blank	9.52	8.44 – 10.59	<0.001	***
B_U	Medium DOC – Low DOC	6.46	5.39 – 7.53	<0.001	***
B_{WS}	Low DOC - Blank	2.17	0.37 – 3.97	0.016	*
B_{WS}	Medium DOC – Blank	7.54	5.74 – 9.34	<0.001	***
B_{WS}	Medium DOC – Low DOC	5.37	3.57 – 7.17	<0.001	***
B_{WB}	Low DOC - Blank	2.96	1.81 – 4.11	<0.001	***
B_{WB}	Medium DOC – Blank	9.85	8.70 – 10.99	<0.001	***
B_{WB}	Medium DOC – Low DOC	6.89	5.74 – 8.03	<0.001	***

B_U ANOVA: F = 252.3, p<0.001.
B_{WS} ANOVA: F = 57.3, p<0.001
B_{WB} ANOVA: F = 238, p<0.001

Values are mean differences with 95% confidence intervals and adjusted p-values.
Significance codes: *** p < 0.001; ** p < 0.01; * p < 0.05.

SI. 5 Output from the Linear Mixed Model comparison between the results of different users with the B_U and B_{WB} digital image analysis options

The LMM analysis of the DOC estimates using the B_U and B_{WB} DIA options are shown in tables S3 and S4. Both of these DIA options led to the same interpretation, with phones D, F and H showing significant differences from Phone A with $p<0.05$, except that the significant difference in Phone F using B_U was higher at $p<0.01$.

Both of these LMM options showed greater variation between phones than the B_{WS} DIA option with 3 phones showing significant differences, compared to only 1 using B_{WS} (table 3, main text), which lends further credibility to the novel B_{WS} DIA option as an appropriate choice for correcting variation between images.

Table S3. Results of LMM analysis using the B_U DIA option to compare the variation in estimates of DOC between phones. DOC estimated by B_U was the response variable, with "Phone" as a fixed factor and Sample Site as a random factor. The estimated intercepts of Phones B - H are compared against Phone A for significant difference, while the intercept of Phone A is compared to 0.

Phone	Intercept ± SE	df	t value	p value	significance
A	8.27 ± 1.37	27.93	6.025	1.73e-06	***
B	1.31 ± 1.21	93.13	1.083	0.28151	
C	-0.91 ± 1.40	93.63	-0.653	0.51525	
D	-2.83 ± 1.37	93.63	-2.063	0.04192	*
E	-2.22 ± 1.40	93.76	-1.583	0.11639	
F	5.52 ± 1.79	94.14	3.079	0.00272	**
G	2.86 ± 2.01	94.05	1.420	0.15900	
H	11.66 ± 4.62	94.46	2.521	0.01337	*

SE = Standard Error
df = Degrees of freedom
Significance codes: ***p<0.001; **p<0.01, *p<0.05

Table S4. Results of LMM analysis using the B_{WB} DIA option to compare the variation in estimates of DOC between phones. DOC estimated by B_{WB} was the response variable, with "Phone" as a fixed factor and Sample Site as a random factor. The estimated intercepts of Phones B - H are compared against Phone A for significant difference, while the intercept of Phone A is compared to 0.

Phone	Intercept ± SE	df	t value	p value	significance
A	8.96 ± 1.41	25.59	6.370	1.03e-06	***
B	1.52 ± 1.15	93.33	1.320	0.1901	
C	-0.90 ± 1.33	93.73	-0.677	0.5002	
D	-3.00 ± 1.31	93.73	-2.293	0.0241	*
E	-1.58 ± 1.34	93.84	-1.182	0.2400	
F	4.09 ± 1.71	94.16	2.390	0.0189	*
G	2.26 ± 1.92	94.07	1.178	0.2418	
H	9.35 ± 4.41	94.39	2.117	0.0369	*

SE = Standard Error
df = Degrees of freedom
Significance codes: ***p<0.001; **p<0.01, *p<0.05

SI 6 Comparison of images taken by Phone A and Phone D

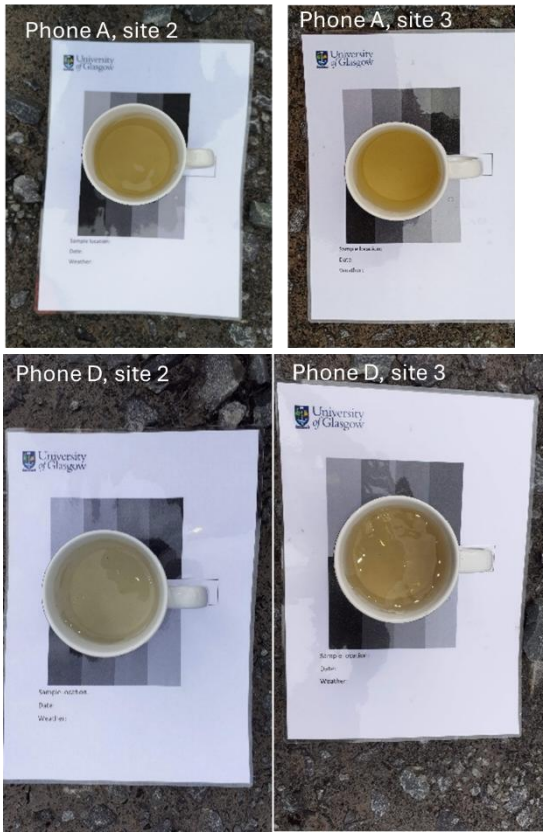

Figure S6. Comparison of the images collected by phone A (top two images) and phone D (bottom 2 images), highlighting that phone D had the flash on during image acquisition, as is evident from the reflections on the water and paper background in the images.

Carter, H. T., Tipping, E., Koprivnjak, J. F., Miller, M. P., Cookson, B., & Hamilton-Taylor, J. (2012). Freshwater DOM quantity and quality from a two-component model of UV absorbance. *Water Research*, *46*(14), 4532–4542. https://doi.org/10.1016/j.watres.2012.05.021

Jerome, J. H., Bukata, R. P., & Whitfield, P. H. (1994). Colours of natural waters: 1. Factors controlling the dominant wavelength. *Northwest Science*, *68*(1), 43–52.

Lin, L. I.-K. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, *45*(1), 255–268. https://www.jstor.org/stable/2532051

Muir, M. R., & Innes, A. (2024). Comparison of test strip, conductivity, and novel smartphone digital image colorimetry methods for field assessment of soil chloride and salinity. *Analytical Methods*, *16*, 5571–5583. https://doi.org/https://doi.org/10.1039/D4AY00991F