Assessing the Accuracy of Nitrate Concentration Data for Water Quality Monitoring Using Visual and Cell Phone Quantification Methods

MELISSA TOPPING (iD)

ALAN KOLOK (iD)

Author affiliations can be found in the back matter of this article

]u[ubiquity press

ABSTRACT

In this methodological project we tested the accuracy of two systems used to quantify results obtained using the Hach© nitrate strip for water quality volunteer monitoring programs. The test strip determines nitrate concentration in accordance with the Lambert-Beer law as increased nitrate concentrations result in greater color intensity on the strip's sampling pad. In this study, first-time volunteers estimated nitrate concentrations with the test strip, either visually or using the Deltares Nitrate App, a smartphone application that uses the phone's camera as a spectrometer. Results from two different series of tests indicate that volunteers using visual methods produce the more accurate results. Although cell phone apps might have the potential to increase data quality for colorimetric assays such as the one employed by the Hach© nitrate test strip, the current technology was not an improvement relative to visual interpolation.

INTRODUCTION

Cultural eutrophication, or the enrichment of water bodies with nitrogen and phosphorous from anthropogenic sources, is a major water quality problem worldwide (Carpenter et al. 1998; Malone and Newton 2020), with eutrophic waterways increasing in frequency and severity (Malone and Newton 2020). Unfortunately, many monitoring agencies are facing resource constraints in the form of personnel, time, and budget limitations that prevent them from adequately addressing pressing environmental issues (Conrad and Hilchey 2011; Wyeth et al. 2019) such as cultural eutrophication. The recruitment of citizen scientists has been suggested as one method to supplement agency monitoring (Cohn 2008; Hadj-Hammou et al. 2017; Thornhill, Chautard and Loiselle 2018; Wyeth et al. 2019); however, if such data are to be trusted by researchers and regulatory agencies, it is necessary that volunteers produce data that is as accurate as possible (Jollymore et al. 2017).

One tool that is currently available for public evaluation of nitrate concentration is the Hach© Nitrate strip. The test strip is a colorimetric assay that is a modification of the Griess reaction (Nelson, Kurtz, and Bray 1954), performing according to the Beer-Lambert law, such that an increase in color intensity is proportional to an increase in nitrate concentation. As such, the reported accuracy of the strip is primarily limited by the sensor that quantifies the change in intensity, in this case the human eye, not by the chemical reaction within the strip itself. In this methodological paper, we tested the accuracy of the human eye versus a camera-based smartphone application in judging the intensity of color of nitrate strips.

Citizen scientists have commonly quantified results obtained with the Hach© Nitrate test strip visually (Loperfido et al. 2010; Muenich et al. 2016; Ali et al. 2019). Recently, it has been proposed that cell phones can increase the quality of data collected by citizen scientists (Burke et al. 2006), which may include increasing the accuracy of data collected using the Hach© Nitrate strip. This assertion is supported by the fact that technological advancements in smartphone cameras now allow them to be used as relatively low-cost spectrometers (McGonigle et al. 2018).

The objective of this work was to assess the accuracy of citizen scientists quantifying results from the Hach© Nitrate test strips with and without the addition of the Deltares smartphone application. In one set of tests, results from volunteers who visually estimated two different nitrate concentrations were compared with the results of volunteers who used the Deltares smart phone app—a platform designed by Deltares, a surface and subsurface water research institute (*https://www.deltares.nl/en/*). In a second series of tests, volunteers quantified nitrate concentration in a number of solutions that spanned the range that the Hach© Nitrate strip can perceive (0–50 mg/L or ppm NO_3-N). Results suggest that, in both cases, the phone app did not increase the accuracy of the results.

METHODS

RECRUITMENT OF CITIZEN SCIENTISTS

Citizen scientists were recruited from various populations through the coordination of 12 separate testing events in Idaho and Washington. These events were hosted on university campuses, at local scientific meetings, and with high school groups from February 2019 to February 2020. In a manner consistent with Ali et al. 2019, participating volunteers had varied skill levels and backgrounds (*Table 1*). Furthermore, while water resource professionals (experts) were included in the population of volunteers, it was not expected that their findings would be more accurate than volunteers with less experience in the water resources field (Ali et al. 2019). The events ranged in participation from 3 to 23 volunteers, with a total of 142 citizen scientists participating in the testing.

NITRATE TEST STRIPS

Volunteers measured nitrate concentrations using Hach© test strips. When quantified visually, these test strips have been used in a variety of citizen science monitoring programs (Loperfido et al. 2010; Muenich et al. 2016) and have been validated previously in laboratory-controlled experiments (Ali et al. 2019).

SAMPLE PREPARATION

Each testing event required volunteers to quantify nitrate concentrations in prepared spiked deionized water samples. All the nitrate samples were prepared using KNO_3 and preserved with sulfuric acid. New solutions were made for each of the 12 testing events, and each concentration was confirmed analytically either by the University of Idaho's analytical laboratory or through use of an in-house discrete analyzer (Seal AQ400: Method: EPA-114-C).

STUDY DESIGN

Two experiments were conducted to address the objectives of this study. The first aimed to assess the accuracy of data produced by volunteers using app and visual methods when evaluating two different laboratory-prepared nitrate concentrations. The second experiment quantified a continuum of laboratory-prepared nitrate concentrations with the intent of evaluating how well visual volunteers and those using the Deltares Nitrate App could quantify the results.

Experiment 1: humans versus smartphone app

To accomplish the goal of the first experiment, 132 volunteers were provided a nitrate sample and instructed to quantify the concentration either visually or using the app. The app volunteers used an iPad equipped with the Deltares Nitrate App to quantify the concentration of their sample. The visual volunteers quantified their sample visually using the colorimetric scale provided by the test strip instructions. All of the volunteers conducted their testing indoors, under similar lighting conditions, because the Deltares Nitrate App can be influenced by poor or inconsistent light. Volunteers were given a water sample and written, but not verbal, instructions to follow. The written instructions emphasized among other things that strip intensity was time sensitive, and that failure to accurately control the timing of the incubation could lead to inaccurate results.

When quantifying the nitrate strip, visual volunteers assigned the color to one of the distinct color categories provided on the side of the bottle. In contrast, volunteers quantifying the strip using the app generated concentrations along a continuous scale. To directly compare the results from both tools, it was necessary that categorical data and continuous data were collected by volunteers using both the app and visual methods. Consequently, in an initial battery of tests, some of the volunteers that collected their data visually were asked to interpolate between the different categorical bins and generate an integer concentration between 1 and 50 ppm, while the remaining volunteers categorized their data into one of the discrete bins. Consistently, half of the app volunteers recorded continuous data, while the remaining volunteers put their data into one of the bins as discrete categories indicated on the Hach© bottle.

Of the recruited volunteers, 66 app and 66 visual samplers participated in this part of the project (*Table 2*). All of the volunteers were asked to quantify the concentration of a water sample, which was either 2 or 15 ppm NO_3. 2 ppm was chosen because it was a categorical option for those testing visually, while 15 ppm was chosen because it was equidistant between two categorical options: 10 and 20 ppm. Relative to the 2 ppm solution, responses were assumed to be accurate if the result was between

TESTING EVENT	DATE	SAMPLE SIZE	VOLUNTEER TYPE
ORED staff	2/26/19	23	University of Idaho staff
Idaho commons	3/22/19	23	University of Idaho college students, staff
ORED open house	4/4/19	11	General college population
Spokane River forum	4/16/19	16	Water professionals, educators, general public
Columbia High School	5/28/19	15	High school students
Palouse Basin aquifer committee meeting	10/10/19	13	General public, water professionals, students
OurGem symposium	11/6/19	9	General public, water professionals
Idaho Water Institute symposium	11/12/19	3	Water resources graduate students, faculty
Idaho commons	12/6/19	9	University of Idaho college students
Idaho water quality workshop	2/11/20	10	Water professionals, students, general public, faculty
Continuum test: visual samplers	6/24/19	5	Idaho Water Institute staff and interns
Continuum test: app samplers	1/24/20	5	Water resources graduate students

Table 1 Organizations that participated in testing events with their corresponding test dates and sample sizes.

Notes: ORED = Office of Research and Economic Development.

NITRATE CONCENTRATION (PPM)	TOOL	CATEGORICAL INSTRUCTIONS N =	CONTINUOUS INSTRUCTIONS N =
2	App	17	16
	Eye	18	16
15	App	17	16
	Eye	17	15

Table 2 Breakdown of volunteers using each quantification method using either categorical or continuous instructions to test their sample.

the two flanking categories (i.e., 1.1 to 4.9 ppm). Values below 1.1 ppm were considered underestimates, and values above 4.9 ppm were considered overestimates. Relative to 15 ppm, solution responses were assumed to be accurate if the result fell between the two categories that flanked 15 (i.e., 10 or 20 ppm). Values below 10 ppm were considered underestimations, and those above 20 ppm were considered overestimates.

Experiment 2: quantifying a continuum of nitrate concentrations

The second experiment was conducted to understand how each tool performed when tested on a continuum of nitrate samples. Ten individuals were tasked with quantifying a continuum of 25 randomized nitrate samples (*Table 1*). Half of these samplers were instructed to visually quantify their test strip and categorize their samples into the discrete bins on the Hach© bottle. The other half of the samplers were instructed to continually quantify their 25 samples using the Deltares Nitrate App. The 25 samples began at 1 ppm and increased to 50 ppm nitrate on every odd value.

STATISTICAL ANALYSIS

In the first experiment, Fisher's exact tests were used to determine if data type had any impact on the accuracy of results, and to determine whether there was any difference in accuracy between the app and visual volunteers when combining results across both concentrations.

The second experiment was conducted with the intent of understanding how each analytic tool performed when quantifying a continuum of nitrate samples. It was necessary to transform the continuous samples used by the volunteers using the app into the same categories that the visual volunteers used. To do this, the continuous values were binned into corresponding categories as per Muenich et al. (2016) who similarly compared continuous lab samples to categorical field samples (*Table 3*). These data

HACH© TEST STRIP SCALE (PPM)	CONTINUOUS SAMPLES ASSIGNED TO EACH CATEGORY (PPM)
0.0	n/a
1.0	1
2.0	3
5.0	5, 7
10.0	9, 11, 13, 15
20.0	17, 19, 21, 23, 25, 27, 29, 31, 33, 35
50.0	37, 39, 41, 43, 45, 47, 49

Table 3 Test strip categories used by visual volunteers and the ranges of continuous nitrate samples placed in those bins.

were plotted and analyzed statistically using Spearman's correlation to determine the relationship between the binned continuum samples and the volunteers' recorded categories.

For the app volunteers, the responses did not need binning because both data types were continuous. The volunteer data were plotted against the true nitrate sample concentrations and fitted with a linear regression model.

All statistical analyses for this project were preformed using either JMP (v. 14.0) or Microsoft Excel (v. 16.33) software with $\alpha = 0.05$.

RESULTS

EXPERIMENT 1: HUMANS VERSUS SMARTPHONE APP

Visual volunteers

Thirty-four volunteers quantified a 2-ppm nitrate solution using visual methods. Of the 34 volunteers, 18 estimated concentration by category, and 16 estimated concentration by extrapolation to a continuous scale. Categorical volunteers were accurate 89% of the time, and continuous volunteers were accurate 75% of the time. The proportions of accurate to inaccurate results were not significantly different (Fisher's exact test, p = 0.3872) between the two quantification methods. Interestingly, all the volunteers that were incorrect overestimated the concentration. The app volunteers produced a wider range of overestimation values (5–36 ppm) than the visual volunteers, who were closer to the true concentration (5 ppm).

Thirty-two volunteers quantified a 15-ppm nitrate solution using visual methods. Of the 32 volunteers, 17 estimated concentration by category, and 15 estimated concentration by extrapolation to a continuous scale. Continuous volunteers were accurate 80% of the time, and categorical volunteers produced 100% accurate results. The proportions of accurate to inaccurate results were not significantly different (Fisher's exact test, p = 0.0917) between the two quantification methods. In contrast to the samplers testing at 2 ppm, continuous volunteers both underestimated and overestimated the nitrate concentration.

When the results from all 66 volunteers were combined and analyzed together, there were no significant differences in accuracy between the visual volunteers who estimated nitrate concentration by category relative to those who estimated concentration by interpolation to a continuous scale (Fisher's exact test, p = 0.719).

App volunteers

Thirty-three volunteers quantified a 2-ppm nitrate solution using the Deltares Nitrate App, with 17 volunteers

who estimated concentration by category, and 16 who estimated concentration by extrapolation to a continuous scale. Categorical volunteers, who were instructed to estimate their sample to the nearest concentration bin, were accurate 52% of the time, and continuous volunteers, who were allowed to interpolate their sample concentration, were accurate 75% of the time. There were no significant differences in accuracy between the categorical and continuous results (Fisher's exact test, p = 0.2818).

Thirty-three volunteers quantified a 15-ppm nitrate solution using the Deltares Nitrate App. Of the 33 volunteers, 17 estimated concentration by category, and 16 estimated concentration by extrapolation to a continuous scale. Continuous volunteers were accurate 44% of the time, and categorical volunteers produced accurate results 24% of the time. There were no significant differences in accuracy between the categorical and continuous results (Fisher's exact test, p = 0.2818).

When the results from all 66 volunteers were combined and analyzed together, there were no significant differences in accuracy between the app users who estimated nitrate concentration using categories relative to those who estimated concentration by interpolation to a continuous scale (Fisher's exact test, p = 0.1387).

Visual versus app volunteers

To compare the two analytic tools, all the responses were pooled into accurate response or inaccurate response from the 66 visual and 66 app volunteers, regardless of data type or concentration. The proportions of accurate to inaccurate responses were determined for each tool and then were analyzed using a Fisher's exact test. The results indicate that volunteers using visual methods are statistically more likely to be accurate than their app-testing counterparts (Fisher's exact test, p < 0.00001) (*Figure 1*). The data were further broken down into the proportion of accurate to inaccurate responses at the two concentrations, and were then analyzed using a Fisher's exact test. The findings indicate that at 2 ppm, results from the visual and app volunteers were not statistically different (Fisher's exact test, p = 0.1036) from each other, whereas at 15 ppm, the visual volunteers were statistically more likely to be accurate than their app-testing counterparts (Fisher's exact test, p < 0.00001*).

EXPERIMENT 2: QUANTIFYING A CONTINUUM OF NITRATE CONCENTRATIONS
App volunteers
The second experiment was conducted with the intent of understanding how each analytic tool performed when quantifying a continuum of nitrate samples. The continuous data produced by the app volunteers were plotted against the true nitrate sample concentration (*Figure 2*). A linear regression model explained more than 75% of the total variation in the data ($y = 1.0369x + 4.9569$, $R^2 = 0.77$, p = 0.0011*).

Regression residuals were calculated by subtracting the volunteer's continuous response from the true nitrate concentration. Of the 125 concentration estimates, 18 of the residuals were underestimations, 14 were accurate, and 93 were overestimations. These observations were compared against a uniform proportion of expected results (33% for each category). A chi-square test of independence was then performed on these proportions to examine the relationship between the sign of regression residuals
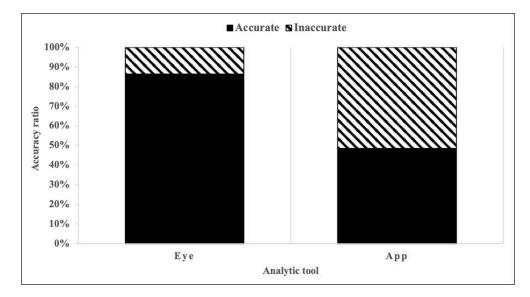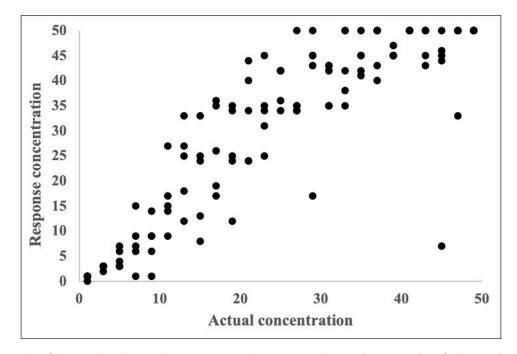

Figure 1 Graphical representation of the proportions of accurate and inaccurate responses for both analytic tools.
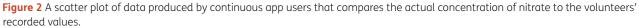
(positive, negative, or zero) and the expected values. The relationship between these variables was significant ($X^2 =$ 42.12, p < 0.00001*), indicating that there is a statistically significant relationship between sign and expected values. Most of the residuals were negative, indicating that the app tends to overestimate.

Visual volunteers

To determine the accuracy of volunteers visually measuring nitrate across a continuum of concentrations, the continuous sample concentrations were binned into one of the six existing Hach© categories as per Muenich et al. 2016 (*Table 3*, *Figure 3*). A comparison was then made of the categorical results that the volunteers generated to the actual concentrations after the data were binned (*Figure 4*). There was a strong positive correlation (Spearman rank correlation, $\rho = 0.8735$, p < 0.0001*) between the true binned concentrations and the categorical estimates of the volunteers.

Regression residuals were calculated by taking the volunteer's response category from the true concentration bin. These data were plotted against the actual concentrations. For the lowest four categories, not including zero (1–4), residuals were off by only one category. The
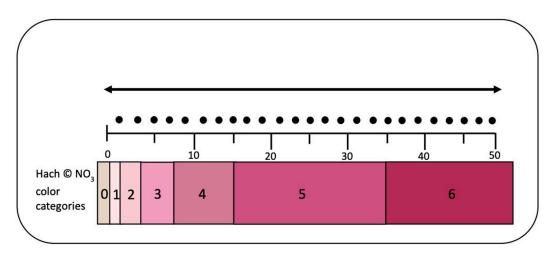

Figure 2 A scatter plot of data produced by continuous app users that compares the actual concentration of nitrate to the volunteers' recorded values.

Figure 3 The dots represent continuous sample concentrations and their designated Hach© category bins. For example, concentrations that fall between 15.1 and 35.9 ppm (or mg/L) would be binned into category 5.
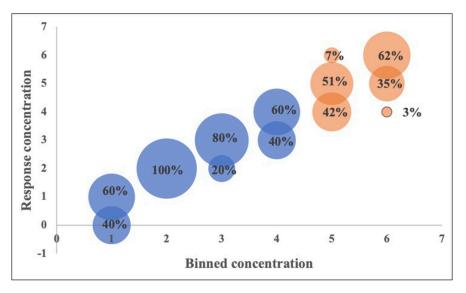
Figure 4 The relationship between nitrate concentrations and corresponding response categories produced by visual volunteers. The blue color corresponds to the samples in the lower range with smaller residuals and the orange corresponds to the samples in the higher range with residuals across three categories. The size of the circles provides a visual approximation of the proportion of samples that fall within each response concentration. In each case the total proportion equals 100%. The response concentration categories are defined by Hach as: 0 = 0ppm, 1 = 1ppm, 2 = 2ppm, 3 = 5 ppm, 4 = 10 ppm, 5 = 20 ppm, and 6 = 50 ppm.

higher concentration bins displayed higher residual ranges, indicating that as nitrate concentration increased, so did the range of categories that were recorded by the volunteers.

Categorical responses were broken into two groups—orange, which corresponded to 0–15 ppm (binned concentrations 0–4), and blue, which corresponded to 17–49 ppm (binned categories 5 and 6), in the same manner as with the continuous analysis. The Spearman's rank correlation was statistically significant (ρ = 0.8523, p < 0.0001* for the categories between 0–4 as well as categories 5 and 6 (ρ = 0.6431, p < 0.0001* *Figure 4*).

DISCUSSION

The objective of this project was to assess the accuracy of citizen scientists measuring nitrate concentrations using the Hach© Nitrate test strips with and without the addition of the Deltares smartphone application. The results do not suggest that using a cell app increases the accuracy of first-time volunteers.

In the first experiment, volunteers that quantified the concentration on their nitrate strips by eye were more accurate than the volunteers who used the app (*Figure 1*). Because all of the volunteers were first-time volunteers, it is possible that an increase in volunteer experience may have increased data accuracy (see Kosmala et al. 2016).

It was also observed that volunteers using both the app and visual quantitative methods tended to overestimate nitrate concentrations. These findings are consistent with those of Ali et al. (2019), who suggested that

improper timekeeping may have been responsible for the overestimations of their volunteers. Given that the Deltares Nitrate App has a built-in timer, it is possible that the timer actually worked against the volunteers, as it was noticed that some volunteers hesitated between immersing the strip and observing the time on the continually rolling timer. Other factors, such as lighting variations or the angles at which the device was positioned might also be responsible for the overestimations. Cell phone apps that generate continuous results from colorimetric assays can be biased due to lighting variations, angles, and device type (Shen et al. 2012; Yetisen et al. 2014; Karlsen and Dong 2015). The findings from this study suggest that the Deltares Nitrate App might also be sensitive to changes in ambient lighting, which could be problematic for volunteers recording data in the field under varying weather conditions and light intensities.

In the second experiment, volunteers were tasked with quantifying 25 randomized water samples that ranged from 0 to 50 ppm nitrate. Volunteers that visually quantified the strips categorized their results into one of the seven concentration bins as per the Hach© instructions, whereas volunteers using the app produced data on a continuous scale ranging from 0 to 50 ppm. Results from both groups increased in variation with increasing sample concentration. For both groups, concentrations between 1 and 15 ppm (categories 0–4) experienced lower variability, and estimates of concentrations between 17 and 49 ppm (categories 5 and 6) were decidedly more variable.

In the second experiment, the volunteers that used the app were prone to overestimation, which was

consistent with the results from the first test. Unlike the first round of testing, these volunteers were more experienced with the platform after testing 25 consecutive samples, so inexperience with timekeeping is less likely responsible for their overestimations. Instead, these overestimations are likely the result of the app software consistently overestimating the test results. In contrast to the overestimations produced by app volunteers, the volunteers who visually estimated their 25 samples were more likely to underestimate. These findings are a bit more difficult to explain, as both Ali et al. (2019) and our findings from the first test indicate that novice volunteers tend to overestimate. Once again, these were not inexperienced volunteers, as they tested 25 samples in a row, so inaccuracies due to timekeeping errors were likely not the explanation for these findings. These results could be due to difficulties perceiving slight chromatic color changes between the higher categories of 10, 20, and 50, which are less stark than the color changes for the lower ranges.

Citizen scientists benefit from the use of cell phone apps, as they gain a powerful analytic tool right in their hands that allows for the incorporation of GPS information and rapid data transmission to be combined with human observation (Burke et al. 2006). If these apps are to be useful outside a controlled setting, they must be flexible enough to accommodate for external variabilities (Karlsen and Dong 2015; Shen et al. 2012; Yetisen et al. 2014) and must be approachable for first time users. Unfortunately, the results from this study suggest that further refinement of the tool will be necessary for cell phone apps to reach their full potential relative to crowdsourced data recovery.

DATA ACCESSIBILITY STATEMENTS

The data used in this research project has not been made available but could be made available upon request.

ACKNOWLEDGEMENTS

FUNDING INFORMATION

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

MT and AK conceived and designed the experiments in this study. MT analyzed the data and wrote the first drafts of the manuscript. MT and AK agree with the manuscript results and conclusions, jointly developed the structure and arguments for the paper, and made critical revisions and approved final versions of the text. Both authors reviewed and approved the final manuscript.

AUTHOR AFFILIATIONS

Melissa Topping   *orcid.org/0000-0002-7053-2585*
University of Idaho, US

Alan Kolok   *orcid.org/0000-0002-1585-6034*
University of Idaho, US

REFERENCES

Ali, JM, Sangster, JL, Snow, DD, Bartelt-Hunt, SL and **Kolok, AS.** 2019. Assessing the accuracy of citizen scientist reported measurements for agrichemical contaminants. *Environmental Science & Technology. American Chemical Society*, 53(10): 5633–5640. DOI: *https://doi.org/10.1021/acs.est.8b06707*

Burke, JA, Estrin, D, Hansen, M, Parker, A, Ramanathan, N, Reddy, S and **Srivastava, MB.** 2006. Participatory sensing. *World Sensor Web Workshop '06 at SenSys '06*, October 31, 2006, Boulder, Colorado, USA.

Carpenter, SR, Caraco, NF, Correll, DL, Howarth, RW, Sharplet, AN and **Smith, VH.** 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications,* 8(3): 559–568. *http://www.jstor.org/stable/2641247*. DOI: *https://doi.org/10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2*

Cohn, JP. 2008. Citizen science: can volunteers do real research?. *BioScience*, 58(3): 192–197. DOI: *https://doi.org/10.1641/B580303*

Conrad, CC and **Hilchey, KG.** 2011. A review of citizen science and community-based environmental monitoring: Issues and opportunities. *Environmental Monitoring and Assessment*, 176(1–4): 273–291. DOI: *https://doi.org/10.1007/s10661-010-1582-5*

Hadj-Hammou, J, Loiselle, S, Ophof, D and **Thornhill, I.** 2017. Getting the full picture: assessing the complementarity of citizen science and agency monitoring data. *PLoS ONE*, 12(12): 1–18. DOI: *https://doi.org/10.1371/journal.pone.0188507*

Jollymore, A, Haines, MJ, Satterfield, T and **Johnson, MS.** 2017. Citizen science for water quality monitoring: data implications of citizen perspectives. *Journal of Environmental Management*. Elsevier Ltd, 200(2017): 456–467. DOI: *https://doi.org/10.1016/j.jenvman.2017.05.083*

Karlsen, H and **Dong, T.** 2015. A smart phone-based robust correction algorithm for the colorimetric detection of Urinary Tract Infection. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. IEEE, November 2015: 1251–1254. DOI: *https://doi.org/10.1109/EMBC.2015.7318594*

Loperfido, JV, Beyer, P, Just, CL and **Schnoor, JL.** 2010. Uses and biases of volunteer water quality data. *Environmental Science and Technology*, 44(19): 7193–7199. DOI: *https://doi.org/10.1021/es100164c*

Malone, TC and **Newton, A.** 2020. The globalization of cultural eutrophication in the coastal ocean: causes and consequences. *Frontiers in Marine Science*, 7(August): 1–30. DOI: *https://doi.org/10.3389/fmars.2020.00670*

McGonigle, AJS, Wilkes, TC, Pering, TD, Wilmott, JR, Cook, JM, Mims III, FM and **Parisi, AV.** 2018. Smartphone spectrometers. *Sensors*, 18(1): 1–15. DOI: *https://doi.org/10.3390/s18010223*

Muenich, RL, Peel, S, Bowling, LC, Heller Haas, M, Turco, RF, Frankenberger, JR and **Chaubey, I.** 2016. The Wabash sampling blitz: a study on the effectiveness of citizen science. *Citizen Science: Theory and Practice.* Ubiquity Press, 1(1), p. 3. DOI: *https://doi.org/10.5334/cstp.1*

Nelson, JL, Kurtz, LT and **Bray, RH.** 1954. Rapid determination of nitrates and nitrites. *Analytical Chemistry*, 26(6): 1081–1082. DOI: *https://doi.org/10.1021/ac60090a041*

Shen, L, Hagen, JA and **Papautsky, I.** 2012. Point-of-care colorimetric detection with a smartphone. *Lab on a Chip*, 12(21): 4240–4243. DOI: *https://doi.org/10.1039/c2lc40741h*

Thornhill, I, Chautard, A and **Loiselle, S.** 2018. Monitoring biological and chemical trends in temperate stillwaters using citizen science. *Water (Switzerland)*, 10(7). DOI: *https://doi.org/10.3390/w10070839*

Wyeth, G, Paddock, LC, Parker, A, Glicksman, RL and **Williams, J.** 2019. The impact of citizen environmental science in the United States. *Environmental Law Reporter; GWU Law School Public Law Research Paper No. 2019*-44, 49(3).

Yetisen, AK, Martinez-Hurtado, JL, Garcia-Melendrez, A, Vasconcellos, FC and **Lowe, CR.** 2014. A smartphone algorithm with inter-phone repeatability for the analysis of colorimetric tests. *Sensors and Actuators, B: Chemical*. Elsevier B.V., 196: 156–160. DOI: *https://doi.org/10.1016/j.snb.2014.01.077*

Citizen Science: Theory and Practice is a peer-reviewed open access journal published by Ubiquity Press.