Evaluating the accuracy of two in-situ optical sensors to estimate DOC concentrations for drinking water production

Citation for the original published paper (version of record):

N.B. When citing this work, cite the original published paper.
Evaluating the accuracy of two \textit{in situ} optical sensors to estimate DOC concentrations for drinking water production\dagger

S. Hoffmeister, a\textsuperscript{,}b K. R. Murphy, c C. Cascone, d a J. L. J. Ledesma d,ad and S. J. Köhler\textsuperscript{a}

Two \textit{in situ} optical sensors, a single-excitation fluorescence-based sensor (fDOM) mounted on a multi-parameter EXO2 sonde (YSI), and a stand-alone, multispectral absorbance-based instrument (spectrolyser, scan Messtechnik GmbH), were evaluated for their capability to (i) estimate river dissolved organic carbon (DOC) concentrations and (ii) provide oversight of drinking water production. The sensors were deployed between March and November 2017 in the river Fyris, which drains a mixed forested and agricultural 2003 km\textsuperscript{2} catchment and serves as a drinking water source by managed aquifer recharge. Grab samples were collected every 2 to 3 weeks and compared with logged sensor data collected at 15 minute intervals. The fDOM probe signal was used to estimate DOC concentrations in the range of 10.4 to 24.4 mg L\textsuperscript{−1} using linear regression ($R^2 = 0.71$, RMSE = 2.5 mg L\textsuperscript{−1}), after correction for temperature, turbidity and inner-filter effects. Temporal changes in DOC character associated with the mixed land use landscape, as indicated by optical indices, reduced this sensor accuracy for estimating DOC concentration. Nevertheless, humic substance concentrations, the fraction of DOC that is preferentially removed during artificial infiltration, were well captured. The spectrolyser signal was used to establish a 2-component partial least square model that captured DOC fluctuations from 10.2 to 29.4 mg L\textsuperscript{−1} ($R^2 = 0.92$; RMSE = 1.3 mg L\textsuperscript{−1}). This multiple-wavelength model (220 to 720 nm) effectively handled the changes in DOC composition while accurately estimating DOC concentrations. This study explores the advantages and limitations of optical sensors for their use in managed aquifer recharge and drinking water production in relation to DOC levels.

Water impact

Tracking fast and unpredictable changes in source water quality is crucial for optimized drinking water production. Current increased development of \textit{in situ} water quality sensors can help anticipate those changes, but their implementation is still constrained by uncertainties. Studies comparing different sensor techniques are often conducted in the laboratory. Here we examine the application of two different sensors under field conditions.

1. Introduction

Northern Europe experienced an extremely dry summer in 2018, which highlighted the vulnerability of drinking water supplies and the challenges of suppliers in the face of a changing climate.\textsuperscript{1} Expected changes in meteorological conditions (e.g. heat waves, sudden heavy precipitation events) will lead to greater variation in the quantity and quality of surface waters, which are used by 50% of Nordic drinking water producers.\textsuperscript{2} Due to this, and decreased acid deposition, surface waters are experiencing “browning” which manifests as increased dissolved organic carbon (DOC) concentrations and associated yellow to brown colour.\textsuperscript{3–5} DOC is of concern to drinking water producers as it is linked to unwanted colour, taste, odour and risk of (i) bacterial growth in the distribution network, (ii) membrane fouling, and (iii) disinfection by-products posing potential health threats.\textsuperscript{6–9} In combination with greater water demand due to a growing population and expanding cities, browner water requires more costly treatment, potentially including higher coagulant doses,\textsuperscript{10} retrofitted treatment steps,\textsuperscript{11} and improved pre-treatment.
In Sweden and other Scandinavian countries, artificial infiltration of surface water through managed aquifer recharge (MAR) is often employed as an efficient pretreatment process that can help to smooth out fluctuations in raw water quality and quantity. Additional benefits include low and stable temperature and natural removal of colour, pathogens, nutrients and metals, reducing the need for posterior treatment inside drinking water facilities. However, large fluctuations in infiltration water quality can threaten proper MAR functionality. High DOC concentrations in the source water could exceed the natural removal capacity and lead to clogging issues, disrupting the process and reducing overall water quality. Hence, monitoring of DOC concentrations becomes necessary to adjust infiltration fluxes accordingly and to minimize risks and costs.

Fluorescence and absorbance spectroscopy have been used in a large number of studies to track water quality changes, mainly related to changes in DOC character. Analysing fluorescence excitation-emission matrices (EEM) is a very powerful tool to quantify subtle changes in DOC character that may correlate to macroscopic parameters such as molecular weight, but EEMs are data-intensive. Optical indices can also be useful tools to understand DOC dynamics and some can be rapidly obtained with online sensors that measure only restricted wavelengths. Specific ultra violet absorbance (SUVA) is widely used in the drinking water sector to predict the presence of hydrophobic substances. Humic substances (HS), detected by e.g. liquid chromatography-organic carbon detection LC-OCD, are a fraction of DOC that are well-removed during coagulation treatment. SUVA and HS/DOC are reportedly good indicators of treatability.

DOC concentrations and character vary on a range of timescales, including seasonal, daily, and sub-daily through changing source areas and in-stream processes. High-frequency sensors can record rapid changes in water quality that might be missed by a program involving only intermittent (grab) sampling. Collecting real-time high-frequency measurements based on absorbance or fluorescence of the ingoing raw water will increase the efficiency of MAR management processes by allowing an immediate response to peaks in turbidity and DOC concentrations, which includes shutting down the MAR to protect the aquifer much earlier than what is the case today.

Due to differences in technical functioning and sensitivity, it is presently unclear whether fluorescence- and absorbance-based sensors provide comparable water quality information, and whether they are worth incorporating into a MAR management scheme. Optical measurements are influenced by the sample matrix, and corrections for temperature quenching, turbidity and inner-filter effects (IFE) need to be performed prior to analysis of the data. In this study, we evaluated whether two optical sensors (one fluorescence-based and one absorbance-based) can provide accurate DOC concentration estimates and beneficial additional information for drinking water production compared to intermittent grab samples. Specifically, we aimed to answer two questions: 1) how accurate are the two different optical sensors in estimating DOC concentrations, and 2) how useful is their application for drinking water production. Additionally, we investigated DOC and catchment dynamic-related processes from the knowledge gained during the measurement period from the combination of sensors and grab samples.

2. Material and methods

2.1. Site description

The river Fyris, hereafter Fyrisån, is the main raw water source for two drinking water treatment plants servicing the population of Uppsala, Sweden. It flows along 80 km from North to South (Fig. 1) and discharges into lake Ekoln, which is part of lake Mälaren (Sweden’s third largest lake with 1090 km²), draining into the Baltic Sea. The annual specific discharge is 204 ± 63 mm per year (1995–2010). Fyrisån is typically dominated by low runoff during summer and high runoff during autumn (a consequence of autumn rainfalls and reduced evapotranspiration) and spring (a consequence of snowmelt). It is one of the largest contributors of DOC per unit of area of all Mälaren tributaries, with an average load of 3.3 ± 1.3 g C m⁻² per year (1995 to 2010).

The Fyrisån catchment covers 203 km² with elevations ranging from 15 to 115 m and is mostly covered by forest (63%, mostly coniferous trees), followed by agricultural lands (32%). Urban areas make up about 3% and wetlands/lakes add up to about 2%. Precipitation is on average 557 ± 87 mm per year (1981–2010) and the mean annual temperature is 6.1 ± 1.0 °C (1981–2010). When the river water level falls below a threshold, water from the adjacent lake Tämnaren and river Vendelån (each contributing with about 0.6–0.7 m³ s⁻¹) is pumped into the river to maintain a minimum water level. The main study site, Storvad, is located ca. 12 km upstream of the city Uppsala, and about 19 km from the river outlet (Fig. 1). At this location the river water is most representative for the raw water inside the drinking water treatment plant.

2.2. Grab sampling and laboratory measurements

In total, 19 grab samples were collected at the Storvad site between March and November 2017 with an interval of 2 to 3 weeks. Filtered samples (pre-burned GFF, nominally 0.7 µm) were analysed for absorbance, fluorescence and DOC concentration in the laboratory. Measurements of absorbance spectra and fluorescence EEMs were performed with an Aqualog (Horiba Jobin Yvon) and a 1 cm quartz cuvette (excitation wavelength range: 250 to 450 nm, emission 212 to 620 nm, at 2 nm intervals). The EEMs were corrected within the associated Aqualog® software FluorEscence™ for IFE of the EEMs using the EEM of Milli-Q and Rayleigh masking to remove primary and secondary Rayleigh peaks. Intensities were converted to Raman units (RU) by dividing the measured fluorescence by the Raman peak area of pure water. In the
following, the term “corrected” EEMs refers to the EEMs treated following this procedure. For comparison of laboratory data with in situ measurements, the RU units were converted to QSU by simple linear regression. DOC concentration was measured using a Shimadzu TOC-VCPH carbon analyser by catalytic combustion with a precision of about 0.2 mg L\(^{-1}\). For quality assurance, standards and Milli-Q water were measured with each batch of samples (absorbance: \(K\)-phthalate (10 mg L\(^{-1}\)), DOC: EDTA 10 ± 0.15 mg L\(^{-1}\)).

2.3. In situ measurements (fluorescence and absorbance sensors)

A fluorescence-based sensor tracking fluorescent dissolved organic matter, fDOM on EXO2 (YSI),\(^42\) and an absorbance-based sensor, s:can spectro:lyser (scan Messtechnik GmbH),\(^43\) were programmed to record data at 15 min intervals. The sensors were deployed inside the pumping station at Storvad which provided electricity and protection from external threats (e.g. weathering, vandalism). River water, pumped into the pumping station, passed through the measuring path of the spectrolyser before draining into an overflow bucket, into which the EXO2 was immersed (Fig. A1, ESI†).

2.3.1. fDOM sensor. The EXO2 probe is a battery-supplied multi-parameter 6-port device with anti-fouling wiper. The device used here was equipped with sensors that measure pH, conductivity/temperature (\(T_{\text{meas}}\)), turbidity (\(\text{Turb}_{\text{meas}}\)) and fDOM (\(\text{fDOM}_{\text{meas}}\)). Additionally, a wiper was installed centrally to wipe the sensors before each measurement. Calibrations were performed according to the manufacturer manual guidelines.\(^42\) The fDOM sensor is an optical fluorescence sensor with an excitation wavelength at 365 ± 5 nm, and emission wavelength at 480 ± 40 nm. A two-point calibration was performed, using pure water (i.e. 0 µg C L\(^{-1}\)) as the first calibration standard, and a 300 µg C L\(^{-1}\) quinine sulphate solution as the second standard, to produce data in quinine sulphate equivalent units (QSU).

2.3.2. Absorbance sensor. The spectrolyser measured absorbance spectra ranging from 220 to 720 nm at 2.5 nm intervals and had a path length of 0.5 cm. Automatic cleaning was conducted twice per day by pressurized air supplied from an oil-free compressor. Manual cleaning with a smooth brush and lint-free cloths took place when collecting grab samples. Connection to the mains was required both for the compressor and the device itself.

2.4. Calibration datasets

Field experiments were undertaken for in situ calibration of turbidity, absorbance and temperature effects on fDOM prior to deployment. To acquire correction functions valid for a wide range of measurements, data of higher and lower concentrations than the expected measured time series had to be generated. Both sensors were placed into a shallower part of the river (ca. 30 cm) around 19 km downstream of the main study site at a nearly-constant temperature of about 10 °C. After a five minute period for adaptation and recording of stable data, turbidity was increased by manually stirring the water and thereby dispersing the sediment near the sensor. The stirring was stopped after five minutes and data logging
continued until turbidity decreased to approximately pre-experiment conditions. As the calibration data set was recorded in less than an hour under calm weather conditions, DOC concentrations were assumed to be stable (11.4 ± 0.1 mg L⁻¹) and all observed changes of fDOM_{meas} were assigned to the introduced variations of turbidity from manual stirring thus facilitating the calibration procedure. During these experiments, data were recorded every 30 seconds by the EXO2 and every minute by the spectrolizer. A grab sample was taken adjacent to the sensors for laboratory analyses of filtered aliquots for DOC, absorbance and fluorescence using the Aqualog. The EXO2 device was exposed to the same water sample after overnight storage in a cooling room (6 °C) for measurements at low temperatures. The lab experiments for calibrating the temperature dependency of the optical sensors were undertaken at a fixed DOC concentration (11.4 mg L⁻¹).

Dilution experiments were used to investigate concentration-related biases in the in situ fDOM signal and to develop a correction function for the field measurements. For this, a river sample with known DOC concentration (26 mg L⁻¹) was diluted five times in series to obtain samples with the same organic matter composition but with different DOC concentrations and absorbance. For each dilution, fDOM was measured with the sensor attached to the EXO2 and absorbance and DOC were analysed in the laboratory.

2.5. Supplementary catchment data

One site within the Fyrisån catchment (Klastorp, Fig. 1) has been sampled monthly as part of the long-term Swedish environmental monitoring program. Due to its relatively close positioning to the Storvad study site (ca. 7 km downstream), Klastorp data were used to complement the study dataset. Analogously to the samples from Storvad, samples from Klastorp were analysed in the laboratory for absorbance, fluorescence and DOC concentration, and additionally as a routine of the monitoring program among others for silica and filtered iron concentrations and absorbance at 420 nm (Abs_{254}) (Fig. A2, ESIF). Furthermore, data from another EXO2 sensor 9 km downstream of Storvad (Bärbyleden, Fig. 1) were available for comparison for part of the study period, producing overlapping sensor data series (data not shown). Water level data were obtained from Uppsala University from a pressure transducer installed in the city centre, approximately 14 km downstream of the study site. Hourly discharge was calculated based on an existing rating curve for this river section.

2.6. Data treatment

2.6.1. Laboratory fluorescence metrics. For each of the grab samples analysed in the laboratory, three optical indices were calculated from the corrected EEMs: humification index (HIX), fluorescence index (FI) and freshness index (β : α). Definitions and interpretations of these indices are presented in Table 1. Additionally, intensities of T and C peaks were located in the EEMs and the resultant ratio of T/C calculated. The T/C peak region is sensitive to changes in the DOC character and can thus provide important information for understanding both the DOC quality and the sensor’s response.

2.6.2. Laboratory absorbance metrics. SUVA values were calculated from absorbance measurements at 254 nm (Abs_{254} [cm⁻¹]) and DOC concentrations. Abs_{254} was utilized to approximate concentrations of HS [mg L⁻¹]. The equation is based on LC-OCD measurements sampled in both Fyrisån and along the infiltration area.

\[ \text{HS} = 1.1 + 19.2 \text{Abs}_{254} \] (1)

2.6.3. In situ fluorescence (fDOM) corrections and DOC estimation. The EXO2 fDOM sensor measures at similar wavelengths as the defined peak C (i.e. excitation: 365 ± 5 nm, emission: 480 ± 40 nm) and hence captures the humic-like peak (Table 1). Fluctuations in turbidity and temperature interfere with the fluorescence sensor signal. Thus, the fDOM_{meas} signal from the EXO2 sensor was corrected for the effects of turbidity, temperature and IFE using a modified version of the method presented by Downing et al.

Instrument-specific functions for the correction of turbidity and IFE were obtained by using the calibration

---

Table 1: Definition and interpretation of established optical indices used in this study. A = area under the curve. Em = emission. Ex = excitation. Numbers refer to wavelength in nm unless otherwise specified. References provided in the text.

<table>
<thead>
<tr>
<th>Index</th>
<th>Definition [No. in nm (if not other indicated)]</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIX</td>
<td>Em: A(345 – 480)/[A(300 – 345) + A(435 – 480)]</td>
<td>Degree of humification; high HIX = more humified material</td>
</tr>
<tr>
<td></td>
<td>Ex: 254</td>
<td>Source; gradient from higher FI (more microbial) to smaller FI (more terrestrial)</td>
</tr>
<tr>
<td>FI</td>
<td>Em: 470/520</td>
<td>Age; higher β : α = larger contribution of more freshly produced DOC</td>
</tr>
<tr>
<td></td>
<td>Ex: 370</td>
<td>Tryptophan-like peak</td>
</tr>
<tr>
<td>β : α</td>
<td>Em: 380/max(420–435)</td>
<td>Humic-like peak</td>
</tr>
<tr>
<td>Peak T</td>
<td>Em: 275</td>
<td>Tryptophan-like peak</td>
</tr>
<tr>
<td></td>
<td>Ex: 340</td>
<td>Humic-like peak</td>
</tr>
<tr>
<td>Peak C</td>
<td>Em: 450</td>
<td>Humic-like peak</td>
</tr>
<tr>
<td></td>
<td>Ex: 350</td>
<td>Humic-like peak</td>
</tr>
<tr>
<td>T/C Intensity (peak T)/intensity (peak C)</td>
<td>Ratio indicating shift from tryptophan-like to humic-like DOC</td>
<td></td>
</tr>
<tr>
<td>SUVA</td>
<td>Absorbance at 254 nm/DOC</td>
<td>DOC quality as indicator of carbon aromaticity</td>
</tr>
</tbody>
</table>

---

[Note: The table and equation are placeholders for the actual data and calculations presented in the original text.]
dataset (Table A1, Fig. A3, ESI†). Eqn (2a) and (b) describe the underlying terms involved for calculating corrected values (fDOMcorr) from fDOMmeas in two steps.

\[
fDOM_{\text{corr, temp}} = fDOM_{\text{meas}} + \frac{(T_{\text{meas}} - 25)}{\rho_d(FNU)} \tag{2a}
\]

\[
fDOM_{\text{corr}} = \frac{fDOM_{\text{corr, temp}}}{r_d(A_{254})} + fDOM_{\text{corr, temp}} \tag{2b}
\]

where fDOMcorr_temp is the temperature and turbidity corrected fDOM, Tmeas is the measured water temperature, \(\rho\) is the temperature-specific fluorescence coefficient, and \(r_p\) and \(r_d\) are instrument-specific correction functions for suspended particulates and dissolved substances (scattering and absorbance) both varying from 0 to 1. For details, readers are referred to Downing et al.32 The temperature-specific fluorescence coefficient \(\rho\) was determined by linear regression of fluorescence intensities against temperature for two different temperatures (6 and 17 °C). Eqn (2a) relates fDOMmeas to a reference temperature of 25 °C.

The correction function for turbidity is based on the relationship between turbidity and fluorescence observations at fixed temperature and DOC concentration (calibration dataset).

In this study, the IFE correction factor \(r_d\) (eqn (2a) and (b)) was obtained by following an absorbance-based approach50 described in detail in Kothawala et al.51 This method is advantageous as no further measurements are necessary. The correction is based on available data (dilution experiments, calibration dataset) using the absorbance at the same wavelength pair as excitation and emission of the fDOM sensor. Thus, correction for IFE was performed according to Kothawala et al.51

\[
P_{\text{IFE Corr}} = P_{\text{Obs}} 10^{0.5(A_{254}+A_{420})} \tag{3}
\]

where \(F\) is the fluorescence signals of the excitation and the emission wavelength \(\lambda\), \(A\) is the respective \(A_{254}\), and superscripts Corr and Obs refer to corrected and observed signal, respectively (the observed signal in our case is the temperature- and turbidity-corrected fDOM time series, i.e. \(fDOM_{\text{corr, temp}}\)).

Subsequently, for both \(r_d\) and \(r_p\), the fluorescence data were translated into fDOM attenuation [%] (i.e. percentage difference between pre- and post-correction fDOM is reported) and expressed as a function of \(A_{254}\) or turbidity \((r_d, r_p\) in eqn (2a) and (b)). For instance, an attenuation of 20% translates into a \(r_d\) of 0.8, i.e. no attenuation is expressed by a \(r_d\) of 1. The fDOMcorr was calculated according to eqn (2b).

Finally, a DOC model was obtained by linear regression between fDOMcorr and observed DOC concentrations. The DOC concentrations estimated from the fDOMcorr measurements are referred to as DOCAbsMod.

2.6.4. In situ absorbance corrections and DOC estimation.

To account for baseline shifts in the spectrolyser dataset, the mean absorbance from 700 to 732.5 nm was subtracted from the full spectrum. The baseline-corrected absorbance data was used to predict DOC concentrations using a partial least square (PLS) regression model. The values from the 214 wavelengths were all used as the explanatory variables (predictors) in the model. In a PLS, \(R^2\) Y indicates how much variation in the response variables is explained by the predictors, \(Q^2\) indicates the predictive ability of the model (derived by subtracting the ratio between predicted residual sum of squares and sum of squares from one), and the variable important for the projection (VIP) indicates the relative importance of the predictors (VIP > 1 indicate important variables).52 PLS analyses are appropriate in the case of overfitting (i.e. number of predictor variables is higher than the number of observations) and strong collinearity between predictors.53 A model with 2 latent variables was established and statistical coefficients were used to analyse the model’s performance. The DOC concentrations estimated from the spectrolyser measurements are referred to as DOCAbsMod.

2.6.5. Statistical analyses.

The strength of correlations between estimated and observed DOC concentrations was assessed using the coefficient of determination \(R^2\) and the root mean square error RMSE.

The significance of trends of fluorescence indices based on a significance level of 0.05 was evaluated using the Mann–Kendall test,54 which is a non-parametric test known for giving robust results for evaluating temporal trends in water quality-related parameters.55,56

3. Results

3.1. DOC concentrations and character measured in the laboratory

Dry conditions were present for most of the study period, with 304 mm of precipitation from May to September producing an average flow of 1.6 m³ s⁻¹. The precedent snow melt episode led to a maximum discharge of 11.6 m³ s⁻¹ on March 23rd, whereas a rain storm at the end of October (28th) produced a maximum discharge of 16.1 m³ s⁻¹ (Fig. 2). These
High-frequency measurements were collected from March to November with only short interruptions. Power outages and technical problems (e.g. software issues) led to gaps in the spectrolayer time series during March and April. Generally, weather conditions were reflected by the high-frequency measurements with stable observations of turbidity, conductivity and temperature in the months May to September and extreme values of turbidity and fDOM_meas in March and October. Systematic shifts in the data due to periodical manual cleaning were not apparent.

3.2. fDOM sensor. The instrument-specific correction functions were derived from the calibration data sets, which contained 57 data points mostly covering the data range of fDOM_meas (68 to 138 QSU) and turbidity (22 to 47 FNU) measurements. The majority (61.2%) of fDOM_meas were within this range, as well as most of the elevated turbidity values (at low turbidity levels, signals are less prone to particle influences). Based on the observed slopes of fDOM_meas versus temperature, a coefficient of −0.012 °C⁻¹ was derived. This corresponds to a decrease of 1.2% in the fDOM_meas signal per 1 °C increase in water temperature. fDOM_corr signal had lower variance than fDOM_meas (Fig. A4, ESI†). The temperature correction led to a mean correction of 18% (ratio only temperature corrected fDOM_corr_temp/ fDOM_meas) and was strongest in March when temperatures were lowest. Applying the turbidity correction increased fDOM_meas during high turbidity events (e.g. see Fig. A4, ESI†: mid-March, mid-October). Multiplication with IFE correction function shifted the signal upwards.

As visible in Fig. A4, ESI†, fDOM_corr compares to the pattern of laboratory DOC in spring and autumn, but in

<table>
<thead>
<tr>
<th>Date</th>
<th>HIX</th>
<th>FI</th>
<th>β×α</th>
<th>SUVA</th>
<th>DOC [mg L⁻¹]</th>
<th>HS [mg L⁻¹]</th>
<th>HS/DOC</th>
<th>T/C</th>
<th>fDOM lab [QSU]</th>
<th>fDOM_corr EXO2 [QSU]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar 09</td>
<td>0.94</td>
<td>1.50</td>
<td>0.57</td>
<td>3.34</td>
<td>11.9</td>
<td>8.7</td>
<td>0.73</td>
<td>0.26</td>
<td>169</td>
<td>146</td>
</tr>
<tr>
<td>Mar 16</td>
<td>0.92</td>
<td>1.51</td>
<td>0.56</td>
<td>3.35</td>
<td>10.7</td>
<td>8.0</td>
<td>0.75</td>
<td>0.28</td>
<td>156</td>
<td>136</td>
</tr>
<tr>
<td>Mar 23</td>
<td>0.94</td>
<td>1.48</td>
<td>0.54</td>
<td>3.35</td>
<td>13.1</td>
<td>9.9</td>
<td>0.76</td>
<td>0.21</td>
<td>185</td>
<td>170</td>
</tr>
<tr>
<td>Mar 31</td>
<td>0.93</td>
<td>1.47</td>
<td>0.54</td>
<td>3.85</td>
<td>12.5</td>
<td>10.3</td>
<td>0.82</td>
<td>0.22</td>
<td>183</td>
<td>171</td>
</tr>
<tr>
<td>Apr 05</td>
<td>0.93</td>
<td>1.48</td>
<td>0.55</td>
<td>3.58</td>
<td>13.0</td>
<td>10.0</td>
<td>0.78</td>
<td>0.22</td>
<td>180</td>
<td>171</td>
</tr>
<tr>
<td>Apr 18</td>
<td>0.94</td>
<td>1.45</td>
<td>0.56</td>
<td>3.16</td>
<td>13.6</td>
<td>9.4</td>
<td>0.69</td>
<td>0.27</td>
<td>160</td>
<td>147</td>
</tr>
<tr>
<td>May 05</td>
<td>0.93</td>
<td>1.46</td>
<td>0.55</td>
<td>3.24</td>
<td>14.3</td>
<td>9.5</td>
<td>0.66</td>
<td>0.26</td>
<td>157</td>
<td>147</td>
</tr>
<tr>
<td>May 15</td>
<td>0.93</td>
<td>1.45</td>
<td>0.56</td>
<td>3.14</td>
<td>12.2</td>
<td>8.5</td>
<td>0.70</td>
<td>0.30</td>
<td>141</td>
<td>132</td>
</tr>
<tr>
<td>Jun 22</td>
<td>0.91</td>
<td>1.06</td>
<td>0.59</td>
<td>2.96</td>
<td>11.4</td>
<td>7.6</td>
<td>0.67</td>
<td>0.41</td>
<td>125</td>
<td>116</td>
</tr>
<tr>
<td>Aug 03</td>
<td>0.91</td>
<td>1.47</td>
<td>0.61</td>
<td>2.78</td>
<td>13.2</td>
<td>8.2</td>
<td>0.62</td>
<td>0.45</td>
<td>126</td>
<td>109</td>
</tr>
<tr>
<td>Aug 14</td>
<td>0.91</td>
<td>1.50</td>
<td>0.61</td>
<td>2.52</td>
<td>13.9</td>
<td>7.8</td>
<td>0.56</td>
<td>0.53</td>
<td>109</td>
<td>106</td>
</tr>
<tr>
<td>Sep 08</td>
<td>0.90</td>
<td>1.50</td>
<td>0.64</td>
<td>2.28</td>
<td>13.9</td>
<td>7.2</td>
<td>0.52</td>
<td>0.46</td>
<td>111</td>
<td>93</td>
</tr>
<tr>
<td>Sep 21</td>
<td>0.89</td>
<td>1.48</td>
<td>0.64</td>
<td>2.60</td>
<td>13.1</td>
<td>7.6</td>
<td>0.58</td>
<td>0.48</td>
<td>112</td>
<td>97</td>
</tr>
<tr>
<td>Oct 02</td>
<td>0.90</td>
<td>1.48</td>
<td>0.63</td>
<td>2.70</td>
<td>12.5</td>
<td>7.6</td>
<td>0.61</td>
<td>0.43</td>
<td>118</td>
<td>104</td>
</tr>
<tr>
<td>Oct 09</td>
<td>0.90</td>
<td>1.50</td>
<td>0.63</td>
<td>2.09</td>
<td>16.3</td>
<td>7.6</td>
<td>0.47</td>
<td>0.46</td>
<td>117</td>
<td>100</td>
</tr>
<tr>
<td>Oct 13</td>
<td>0.93</td>
<td>1.54</td>
<td>0.59</td>
<td>3.23</td>
<td>15.6</td>
<td>10.8</td>
<td>0.69</td>
<td>0.25</td>
<td>187</td>
<td>181</td>
</tr>
<tr>
<td>Oct 16</td>
<td>0.94</td>
<td>1.50</td>
<td>0.54</td>
<td>3.35</td>
<td>24.8</td>
<td>17.1</td>
<td>0.69</td>
<td>0.22</td>
<td>301</td>
<td>272</td>
</tr>
<tr>
<td>Oct 23</td>
<td>0.94</td>
<td>1.50</td>
<td>0.53</td>
<td>3.39</td>
<td>22.6</td>
<td>15.8</td>
<td>0.70</td>
<td>0.21</td>
<td>246</td>
<td>248</td>
</tr>
<tr>
<td>Nov 06</td>
<td>0.95</td>
<td>1.51</td>
<td>0.51</td>
<td>3.34</td>
<td>26.6</td>
<td>18.2</td>
<td>0.68</td>
<td>0.18</td>
<td>289</td>
<td>293</td>
</tr>
</tbody>
</table>

*a* Calculated from eqn (1),
contrast to the stable DOC observations from May to September, it decreased steadily (from 160 QSU to 83 QSU) leading to deviations between the two signals. Correlation between measured DOC and the fDOMcorr resulted in a $R^2$ of 0.71 and a RMSE of 2.5 mg L$^{-1}$ (Fig. 3 and Table 2).

Despite the observed trends for the summer period in HIX, SUVA and $\beta$; $\alpha$, none of the indices described well the overall variation in the fDOMcorr signal. HIX ($R^2 = 0.69$) and $\beta$; $\alpha$ ($R^2 = 0.65$) correlated better with fDOMcorr than SUVA ($R^2 = 0.41$) and FI ($R^2 \leq 0.1$). The fDOMcorr deviations were best-explained by the HS concentrations ($R^2 = 0.96$).

**3.2.2. Absorbance sensor.** A 2-component PLS model (Table A2, ESI†) was established based on the measured absorbance spectra, which explained 92% of the variation in DOC concentrations ($i.e. R^2Y = 0.92$) with a high predictive ability ($Q^2 = 0.85$). Thus, the estimated DOC based on the model correlated well with the measured DOC ($R^2 = 0.92$, RMSE = 1.3 mg L$^{-1}$, Fig. 3) and captured well the DOC dynamics (Fig. 2). Most of the 214 wavelengths were important predictors in the PLS model ($VIP > 1$), except for the higher wavelengths (larger than ca. 530 nm) (Fig. A5, ESI†). The estimated DOC from the spectroyler did not significantly correlate with any of the optical indices (HIX, FI, $\beta$: $\alpha$ and SUVA) during the sampling events (all $R^2 < 0.25$, $p > 0.01$, Table 2), but a significant correlation with HS was apparent ($R^2 = 0.86$, $p < 0.001$).

**4. Discussion**

**4.1. Sensor evaluation – comparison of DOC estimations**

Both sensors delivered estimates of DOC concentration in the form of high-frequency time series, however, their success at predicting laboratory-based DOC measurement differed. Although both instruments required data post-processing, this was most onerous for the fDOM sensor, which required compensation for temperature, turbidity and IFE, including dedicated experiments to establish sensor- and site-specific corrections. Under the conditions of our study, the spectroyler resulted in better estimates of DOC concentration than the fDOM sensor. The RMSE for the fDOM sensor estimation (2.5 mg L$^{-1}$) was greater than that for the spectroyler (1.3 mg L$^{-1}$). The absorbance-based DOC model was satisfactory, confirming the suitability of the modelling approach and the stability of the spectroyler over time. PLS regression from absorbance data has successfully been used previously for calculating DOC concentrations.$^{57,58}$ On the other hand, in contrast to earlier studies that reported online fDOM to be a reliable proxy for DOC concentration (e.g. $R^2 = 0.79,59$ $R^2 = 0.96,60$ $R^2 = 0.93$ and 0.85$^{17}$), we observed a period of relatively poor prediction between May and September.

Ruling out errors in setup and data treatment is essential before attempting to explain the differences between observed and estimated DOC by natural processes. Differences between the predictions by the two sensors cannot be explained by instrument setup since flow cell construction and installation were similar for both sensors (Fig. A1, ESI†). Another fDOM sensor installed 9 km downstream at Bärbyleden (Fig. 1), recorded similar fDOM as the one analysed in depth here, as demonstrated by linear correlation with a $R^2$ of 0.99 between the fDOMdata datasets of the two sensors (data not shown). The fDOM data corrections employed here included a combination of instrument- and site-specific factors, which complicates direct comparisons with other studies that used slightly different approaches in settings with different specific characteristics. However, molecular fluorescence of humic-like organic matter usually drops about 1% °C$^{-1}$ increase in temperature$^{31,61}$ and our site-specific temperature correction factor of $-0.009$ °C$^{-1}$ matches this and other published values.$^{34,62}$ Downing et al.$^{32}$ reported that about 10 to 19% of the fDOM signal was attenuated at a turbidity level of 50 FNU. Our results similarly showed an attenuation of 15% at 50 FNU. Downing et al.$^{32}$ also reported that correction of the IFE at an absorbance of 0.5 at 254 nm led to about 14 to 28% attenuation. Our correction resulted in an attenuation of 14% at this absorbance. Finally, comparing the field fDOMcorr signal with laboratory fDOM measurements (fDOM lab) from grab samples revealed that, after normalization, the observed dynamics in these two signals were almost identical ($R^2 = 0.95$).
Our results do indicate that the fDOM$_{corr}$ signal was influenced by changes in DOC character in Fyrisån between March and September. As DOC quantity was stable during long periods of the late spring and summer, we are confident that the measured spectroscopic properties reflected changes in DOC character. The observed trends of HIX, SUVA and $\beta : \alpha$ from April to September indicate a shift in the DOC quality. During low-flow summer conditions, riverine DOC is composed of less-humified material with a larger contribution of more freshly produced DOC (lower HIX, higher $\beta : \alpha$) compared to high-flow conditions. The aromaticity (SUVA) decreased simultaneously and a shift occurred from humic-like (peak C) to protein-like (peak T) organic matter. Underlying reasons are a combination of several processes taking place in both the soil and the water column. On the one hand, during the summer months, low discharge and longer residence times trigger microbial in-stream processes that can lead to the production of microbially-derived DOC. Elevated temperatures enhance microbial activity and the photosynthetic activity of microbially-derived DOC. Elevated temperatures enhance microbial activity and the photosynthetic activity of diatoms; evidence of the latter is seen in the decreasing silica concentrations in nearby Klastorp (Fig. A2, ESI†). Similar in-stream processes affecting DOC are typically observed in Mediterranean catchments where dry summer conditions occur frequently.\(^6\) On the other hand, a change of the DOC source from the catchment can be considered likely. Our records from the monitoring station at Klastorp indicate that during high discharge, and thus during periods with elevated groundwater tables, absorbance, filtered iron and, consequently, SUVA are increased (Fig. A2 and A6, ESI†), which can be explained by the mobilization of water from shallow riparian soil levels. This is in accordance with the dominant source layer concept.\(^6\) At low discharge, groundwater tables are low leading to lateral water inputs from deeper, mineral-like soil layers, which inherit DOC that is less aromatic, more microbially derived, and more recently produced than that from shallow organic riparian layers.\(^10\) At deeper soil layers, the character of DOC changes due to the selective removal of terrestrial carbon during soil passage, whereby DOC attaches to mineral surfaces or flocculates with iron or aluminium hydroxides.\(^5\)

From March to September, we observed lower concentrations of HS and fresher DOC (Table 2), probably due to release of degraded organic matter of the riparian vegetation and from high in-stream processing of fresh organic matter. Changes in the ratio of fDOM$_{corr}$ per DOC can be a consequence of these processes (Fig. 4). Furthermore, the pattern of the T/C ratio supported the DOC compositional variations. As the signal decreased during the summer months, a shift from the humic-like to the tryptophan-like peak occurred, indicating a shift from allochthonous terrestrial organic matter to autochthonous material, hence greater degree of microbial activity. Monitoring data from the Klastorp site support this conclusion, highlighting the typical annual pattern of high flow, high silica under cold periods and low flow, warm temperature and biological in-stream processing of silica in summer (Fig. A2, ESI†).

We hypothesize that earlier observed good fits between fDOM and DOC were established in surface waters with a smaller fraction of agricultural land use and/or less temporally-variable DOC character, e.g. in lakes\(^32,60,65\) that represent more stable environments than our river water. As such, care needs to be taken when using a fDOM sensor for estimating DOC concentrations in lotic waters.

### 4.2. Application in drinking water treatment

Next to practicalities such as acquisition cost, maintenance and power requirements, the best choice of sensor depends on the monitoring aim and the catchment under investigation. The required accuracy of DOC concentration estimations differs depending on the application of the sensors. For illustration, if during the drinking water production process 3 to 4 mg L$^{-1}$ were removed, then a DOC estimation error of 1.5 mg L$^{-1}$ would not be satisfactory. However, for deciding on temporary interruptions of the water intake due to a high concentration event, it might still fulfill its purpose. For drinking water applications, the fDOM sensor might suffice if the main objective is to adjust chemical dose for coagulation processes. The HS, representing the organic matter fraction that can be removed by flocculation,\(^6\) showed here a very good correlation with the fDOM$_{corr}$ ($R^2 = 0.96$). However, complex site- and instrument-specific corrections need to be applied to the data recorded by the fDOM sensor.

Grab sampling is also needed in order to maintain the sensors as well as to improve the correction functions. If measurement of several parameters is intended, it might be justified to invest in a full-spectrum sensor and to calibrate it to parameters of interest (e.g. HS fraction, DOC, turbidity). The spectrolyser would be a more reliable choice for predicting DOC, especially in flashy catchments. This study indicates that greater precision in DOC concentration estimates can be achieved if information from a broad range of wavelengths is available as compared to a single excitation sensor.

Sensors offer well-known advantages over intermittent samples, especially the possibility of reacting immediately to changes in the raw water quality, allowing for real-time adjustment of processes such as MAR. Grab samples are prone to miss certain events leading to false assumptions about DOC dynamics, as exemplified here by comparing linearly interpolating DOC concentrations of the grab samples with the estimated DOC concentrations from the spectrolyser, where systematic differences of up to ±4 mg L$^{-1}$ are present at several periods of days (Fig. A7, ESI†). At drinking water facilities, tracking and responding to changes in DOC character can reduce costs of chemicals and increase membrane lifetimes by reducing backwashing frequency. Once reliable routines for treating sensor data are developed, personnel costs will diminish due to a reduced need for grab samples and laboratory analyses.
5. Conclusions

In this study, a fluorescence- and an absorbance-based sensor were evaluated in relation to their accuracy for estimating DOC concentrations and in their applicability for drinking water production. Both sensors were found to be suitable supportive tools for MAR depending on the water quality parameters of greatest interest at the treatment plant.

Optical indices derived from grab samples indicated temporal changes in the DOC character throughout the measuring period in this mixed forested and agricultural landscape. These changes impacted the accuracy of the compared sensors differently. The EXO2 fDOM sensor provided periodically inaccurate predictions of DOC; however, if the objective is to monitor the easily-removed DOC fraction (i.e. HS), then the fDOM sensor is a suitable solution assuming that a convenient post-processing data pipeline can be developed. Further confirmation of the stability of the fDOM–HS relationship is needed. In contrast, the multiple-wavelength model based on spectrolyser absorbance spectra reliably compensated for changes in DOC composition. Expanding the dataset to measure more data under extreme conditions will improve the prediction of the full range of DOC concentrations.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was part of two projects supported by VINNOVA (2017-03729) and SVU (SVU 16-103). KRM was supported by the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS grant 2013-1214). JLJL was supported by the Swedish Research Council (FORMAS grant 2015-1518) and the Spanish Government through a Juan de la Cierva grant (FJCI-2017-32111). We thank the Department of Earth Sciences at Uppsala University for providing water level measurements and the respective discharge curve and Don Pierson for loaning one of the instruments. We thank Uppsala Vatten for providing access to the site and constructive feedback.

References


Environmental Science: Water Research & Technology