Research Article | Vol. 27, No. 16 | 5 Aug 2019 | OPTICS EXPRESS 22339
Optics EXPRESS

Impact of fluorescence on Raman remote sensing of temperature in natural water samples

ANDRÉA DE LIMA RIBEIRO,[1,*] CHRISTOPHER ARTLETT,[1,2] PENELOPE A. AJANI,[3] CARO DERKENNE,[1] AND HELEN PASK[1]

[1]*MQ Photonics Research Centre, Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*
[2]*Defence, Science and technology Group, Maritime Division, Eveleigh, NSW 2015, Australia*
[3]*University Technology Sydney, Ultimo, NSW 2007, Australia*
[*]*andrea.delimaribeiro@mq.edu.au*

Abstract: A comprehensive investigation into the impact of spectral baseline on temperature prediction in natural marine water samples by Raman spectroscopy is presented. The origin of baseline signals is investigated using principal component analysis and phytoplankton cultures in laboratory experiments. Results indicate that fluorescence from photosynthetic pigments and dissolved organic matter may overlap with the Raman peak for 532 nm excitation and compromise the accuracy of temperature predictions. Two methods of spectral baseline correction in natural waters are evaluated: a traditional tilted baseline correction and a new correction by temperature marker values, with accuracies as high as ± 0.2°C being achieved in both cases.

1. Introduction

The use of Raman spectroscopy to predict water temperature was first proposed in [1,2] and the potential for extending this to determine depth-resolved temperature profiles using LIDAR methods was investigated in [3,4]. The applications for such knowledge are extensive, and include making predictions about underwater communication, validating hydrologic and climate change models, and obtaining habitat information. In principle, the methods could be compatible with airborne, surface, land-based or even underwater platforms. Since the comprehensive studies of several decades ago, there have been a modest number of studies that have advanced the field [5–7], accompanied by large advances in the sensitivity of photomultipliers, spectrometers and numerical statistical methods.

In 2015, we reported our first work in this field, harnessing these advances to obtain high quality Raman spectra, using statistical methods to identify the spectral parameters most sensitive to temperature change, and systematically predicting the accuracy with which temperature could be predicted by means of a temperature marker. In this case the temperature marker was the ratio of Raman signal intensities at two particular frequencies within the broad Raman band associated with OH stretching [8]. The method is commonly known as the "two-colour method", and the marker is commonly called the "two-colour ratio". It was found that water temperature could be predicted with an accuracy better than ± 0.2 °C in the case of pure (reverse-osmosis) water, and informed the design of a simple two-channel apparatus that used pulsed excitation, high fidelity filters and photomultipliers in proof of principle experiments to predict the temperature of tap water in a 1 m long cell to within ± 0.5 °C. That was a first step towards optical instrumentation for vertically profiling water temperature in natural environments.

A key challenge to implementing the two-colour method for temperature determination in real environments is the presence of additional materials such as dissolved organic matter

#366647 https://doi.org/10.1364/OE.27.022339
Journal © 2019 Received 20 May 2019; revised 8 Jul 2019; accepted 9 Jul 2019; published 23 Jul 2019

Optics EXPRESS

(DOM), phytoplankton and particulates. These constituents exist in different concentrations across locations, varying within the water column and over time according to tides, seasons and climatic factors. Furthermore, they have the potential to fluoresce at wavelengths that overlap the Raman band, absorb Raman photons, and scatter both excitation and Raman photons. DOM fluorescence has a short lifetime [9], exhibiting two characteristics components: from 300 to 350 nm (associated with protein decomposition) and from 400 to 580 nm associated with humic substances of organic origin, also known as *Gelbstoff* [10]. Phytoplankton fluorescence signals are typically around 685 nm due to the presence of chlorophyll-a (Chl-a), a mandatory photosynthetic pigment found in all phytoplankton species, and its intensity is proportional to phytoplankton concentration [11]. Depending on the species, physiological state of the cell and environmental conditions the concentration of accessory pigments (e.g. chlorophyll-b,c,d, carotenoids, phycobilins) may increase in the cell, giving rise to fluorescence at additional wavelengths.

Simulations in [12] explored the spectral overlap between the Raman signal and laser-induced fluorescence in natural marine waters, within an excitation range of 510-570 nm. The authors identified a distortion on the Raman signal due to Chl-a fluorescence, compromising the water temperature prediction methods based on Raman spectroscopy and proposed a fluorescence subtraction approach. Following on, the authors of [13] used a tunable laser (480 to 530 nm) to investigate laser-induced fluorescence in natural waters from Chesapeake Bay. Three main spectral features were identified in addition to the Raman signal which occurred at a fixed separation (centred around 3400 cm^{-1}) from the excitation wavelength: the Chl-a fluorescence band at 685 nm, a weaker phytoplankton band at 730 nm, and low-level fluorescence from DOM at shorter wavelengths. With regard to temperature prediction, it was found that shorter excitation wavelengths (480 nm) generated Raman signals that were more likely to overlap with DOM fluorescence, while the Raman signals generated by longer wavelength excitation (530 nm) were more likely to overlap with Chl-a fluorescence signals and that these overlaps could compromise temperature prediction. The same author subsequently broadened his study to investigate the temperature dependence of the Raman and fluorescence signals [14]. It was observed that, within the excitation range from 490 to 520 nm, the Raman signal counts were higher for shorter-wavelength laser sources (blue) and reduced towards longer excitation (green). By comparison, any dependence of Chl-a and DOM fluorescence on excitation wavelength was found to be insignificant. For excitation within this range, there was minimal overlap of Raman and the Chl-a fluorescence signals at 685 nm, and that baseline subtraction could be used prior to applying the two-colour method. When a longer excitation wavelength of 535 nm was used, significant overlap between Raman and Chl-a fluorescence bands was observed and this was problematic for temperature prediction.

Next we note an essential difference between our approach, and that in most other works where the Raman spectra are decomposed into a series of Gaussians, the heights and/or widths of which are then analysed to determine water temperature. In this work, as in [8] we simply integrate the Raman signal that falls within pairs of selected spectral channels of typically 200 cm^{-1} width. While the studies in [12–14] found the optimal wavelength for remote sensing of temperature to be between 488 nm and 520 nm, there are few commercial laser sources available in this band. On the other hand, 532 nm lasers are widely available and are typically used for laser bathymetry, and accordingly it is important to quantify the extent to which DOM and phytoplankton fluorescence might impact on the accuracy of temperature prediction in natural water samples by using 532 nm excitation, and to explore strategies for making corrections using our two-channel approach.

In the present study, we collected unpolarised Raman spectra for natural water samples from five locations around Sydney Harbour, which typically show the Raman band at 3400-3600 cm^{-1} superimposed on a background signal, which we show is largely due to fluorescence. We have carried out a controlled experiment in which we grew 3 common

Optics EXPRESS

species of phytoplankton and measured the concentrations that gave rise to a baseline large enough to perturb the Raman signal. These concentrations were within the normal range of concentrations along the coast of eastern Australia. We propose two methods for correcting for fluorescence and present the accuracy with which temperature can be predicted with and without baseline correction, considering the implications of our findings for developing future methods that are less susceptible to the presence of fluorescing matter.

2. Experimental details

2.1 Raman spectral measurements and analysis

Natural water samples were collected from various locations around Sydney. These include Manly Beach, which is outside the Harbour, Clontarf and Sugarloaf Bay, which are located in Middle Harbour approximately 4 km and 7 km respectively from the Harbour entrance, and Rhodes which is located in the main harbour approximately 20 km from the Harbour entrance. Water samples from Rose Bay were of particular importance to this study. These samples were collected from deep waters in an open part of the harbour, approximately 2 km from its entrance and investigated after being filtered and UV treated. Water samples from Manly Dam, a freshwater body of approximately 2000 ML located in an urban bush reserve were investigated, along with pure (Milli-Q) laboratory water.

Raman spectra were collected for each sample within a few hours of being collected, and all the data was collected, using methods that have been described in detail in [8]. Briefly the spectrometer used was an Enwave EZRaman-I, a dispersive Raman spectrometer having a spectral resolution of 8 cm^{-1} and using 30 mW continuous-wave (CW) laser at 532 nm for excitation (Fig. 1). The unpolarised Raman signal was detected using a 180° backscattering geometry, and wavelength calibration of the spectrometer was carried out using an acetonitrile (CH_3CN) reference sample. Spectral data were smoothed with the Savitsky-Golay algorithm to reduce noise (2nd order, 25-point window). The spectrometer integration time was typically 30 seconds and each spectrum shown is an average of 3 acquisitions to improve consistency. Each sample was conditioned inside a quartz cuvette (pathlength of 10 mm) and stepped through a range of temperatures from 12 to 33°C, with a waiting time of several minutes allowed after reaching each set point to enable the water sample to reach thermal equilibrium. The reference temperature was measured using a temperature probe within the QNW QPod2e, which has a specified accuracy of ± 0.2°C.

The unpolarised Raman spectra were analysed using the "two-colour" method, in which the temperature marker is the ratio of the Raman signals in spectral bands on either side of the isosbestic point (point of equal scattering) of the OH stretching band. This ratio is found to be proportional to water temperature. In this work, as in [8], the ratio is determined by integrating the Raman signals that correspond to the spectral bands of interest. The spectral resolution (8 cm^{-1}) was substantially lower than the channel widths over which Raman signals are integrated (200 cm^{-1}). The analysis method involves carrying out a linear least squares regression, using MatlabR2017b, of the two-colour ratio against the measured reference temperature to yield a temperature-predictive model. Each combination of wavenumber pairs produced a Root Mean Squared Temperature Error (RMSTE), the set of which was used as a parameter for estimating the accuracies of temperature predictions in various water samples.

Optics EXPRESS

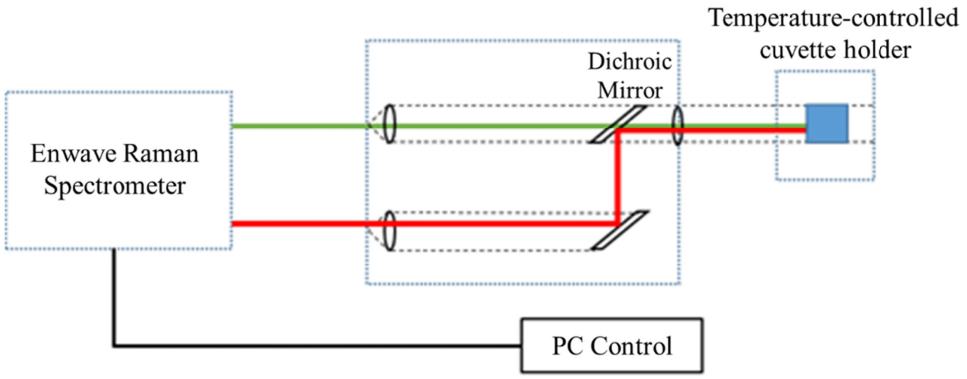

Fig. 1. Experimental setup for collection of Raman spectra. The green line depicts the 532 nm excitation beam, and the red line depicts the Raman return.

2.2 Phytoplankton growth and spectral analysis

Non-axenic clonal cultures of *Synechococcus* Red, *Synechococcus* Green, *Nannochloropysis* sp., *Ditylum brightwelli*, *Dunallella tertiolecta*, *Ostreopsis siamensis* and *Rhodomonas salina* were obtained from collections held at Macquarie University and University of Technology Sydney and maintained in 2 x 200 mL culture flasks (replicates) containing F/2 media [15]. Cultures were grown under 24 hr LED lighting at 20°C and monitored using light microscopy. One mL of cultivated culture from each strain was transferred into fresh media every two weeks to maintain healthy and exponentially growing cultures over the duration of the study.

To investigate the extent to which phytoplankton fluorescence overlaps the water Raman spectra, three species were chosen: *Synechococcus* Red, *Synechococcus* Green, *Nannochloropysis sp*., These three cover the major naturally fluorescing pigments found in most phytoplankton species [16], and are also relatively small in relation to the excitation volume in the Raman spectrometer. Raman spectra were first recorded for an f/2 growth medium sample, and then a pipette was used to add drops of each phytoplankton culture. Raman spectra were recorded after each drop, and an average of 10 spectra was taken for each sample. Two samples for each species were fixed using Lugol solution, and the concentration of phytoplankton in each sample was then measured by counting under a microscope.

3. Background signals, their implication for determining temperature, and their origin

Figure 2 shows two examples of temperature-dependent unpolarised Raman spectral data, corresponding to the OH stretching band, acquired for water samples from coastal locations: Rose Bay, Clontarf, and Rhodes. The Raman signal is given in terms of signal counts registered by the spectrometer (i.e. the CCD counts corrected for grating and detector spectral response) and have not been normalised. In the case of the Rose Bay water sample shown in Fig. 2(a), which had been filtered and UV treated, the temperature-dependence is well-defined in terms of an isosbestic at ~3422 cm^{-1}. However, the same cannot be said for the other examples, which lack a clear isosbestic point and exhibit higher baseline levels. It was observed that the baseline could be reduced by passing the water sample through filter paper and that it could be increased by stirring the water sample, both of which suggest that the baselines arise at least partially from suspended material. The samples analysed here were neither filtered nor stirred.

Optics EXPRESS

Maps depicting the accuracy with which water temperature can be predicted, as a function of the positioning of the "low shift" and "high shift" spectral bands, were generated from the measured Raman spectra and are shown in Fig. 2. Fairly wide spectral channels (200 cm^{-1}) were used in the analysis, motivated by our end goal being to implement the methods we develop in the field, rather than the laboratory. As found in [8] the optimal channel width is a compromise between temperature sensitivity and signal strength. In practice, its selection will depend on factors such as the laser source, receiver and detector characteristics. The RMSTE maps in Fig. 2 show that it is straightforward to predict the water temperature of the Rose Bay water (filtered and UV treated) with a relatively high accuracy of ± 0.3°C, and there is a reasonable amount of flexibility when selecting the optimal spectral channels. In the case of the water sample from Rhodes, for which a considerable baseline is present, the accuracy was somewhat lower at around ± 1.0°C and moreover there is a narrow window for selecting the optimal spectral channels.

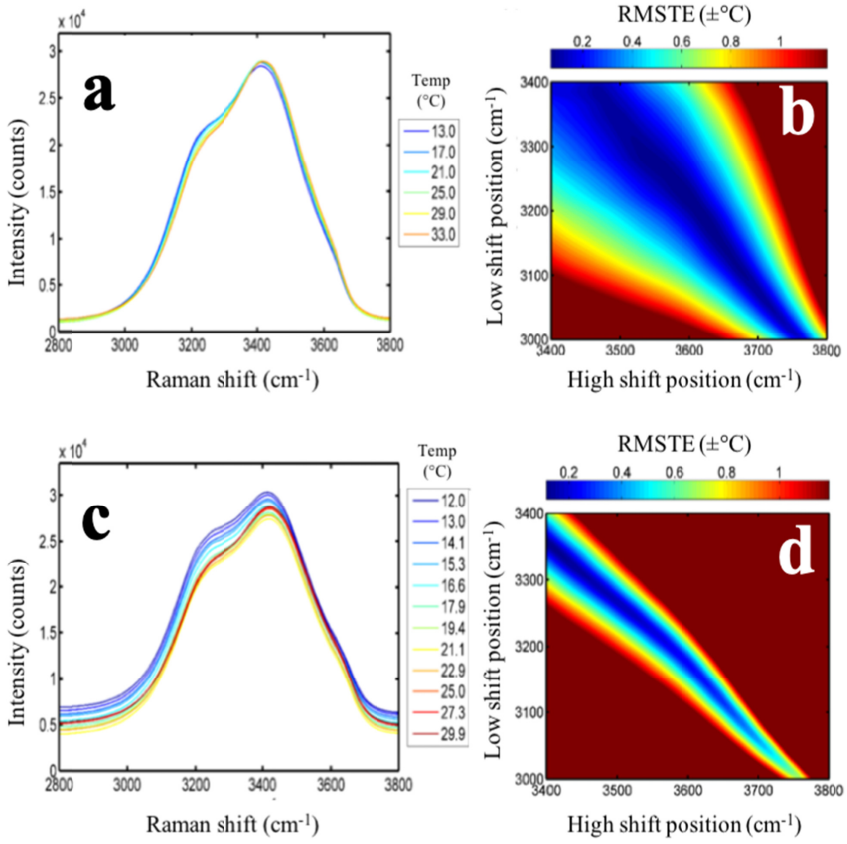

Fig. 2. (a) and (c) show unpolarised Raman spectra for Rose Bay and Rhodes samples, as a function of temperature; (b) and (d) show associated maps of RMSTE computed for all pairs of wavenumber channel centres (200 cm^{-1} channel width).

The presence of background signals in Raman spectroscopy is very well-known and is frequently due to fluorescence from various constituents of the sample which interact with the excitation laser. In the case of natural waters, several studies [13,14] have shown that fluorescence can arise from organic molecules excited at 532 nm, specifically DOM and Chl-a. Understanding the origin of the baselines apparent in our spectra is highly desirable in terms of developing a strategy for correction and ultimately for increasing the accuracy with which temperature can be predicted. To this end, we sought to understand which portions of

Optics EXPRESS

the spectra gave rise to higher variances between waters collected at different locations, by means of a spectral Principal Component Analysis (PCA) using the Single Value Decomposition algorithm, a non-iterative algorithm with optimum performance for spectral data. PCA is a technique able to summarize large multi-dimensional data sets, such as spectral data, into fewer dimensions of variability called Principal Components (PC). PCs are orthogonal to each other and represent axes of maximum variability between samples *i.e.*, PC-1 always accounts for the maximum variability, PC2 accounts for most of the variability not explained by PC-1 and so on. Interpretation of PCA results require prior knowledge regarding potential spectral signatures in the sample, as a PC might represent residual variations with no spectral significance [17,18]. In our study, two data sets were selected to perform the PCA: a calibration set comprised an average of Raman intensities at different temperatures for each location (13°C, 17°C, 25°C, 29°C and 33°C) and a validation set of Raman intensities measured for 21°C. Both data sets were normalised by the peak aiming to reduce the temperature-dependent effect around the OH stretching band and isosbestic points. The UnscramblerX, version 10.5, a software from CAMO, was used for the PCA.

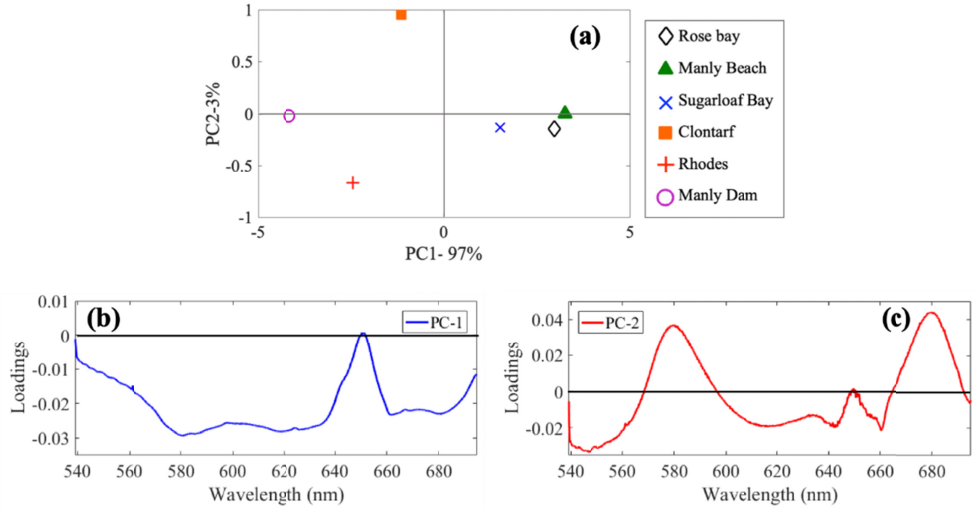

Fig. 3. Overview of the PCA findings. (a) Scores plot: the distance from the origin depicts the extent to which PC-1 and PC-2 contributes to the spectral variance. (b) PC-1 and (c) PC-2 loadings plots: these indicate which spectral regions contribute most to the observed variability for each PC.

The PCA of Raman spectra generated a model explaining 99.5% (98.6%) of spectral variation among locations for the calibration (validation) data by using two principal components: PC-1 and PC-2. Rhodes and Clontarf (solid symbols) variances were associated with both PC1 (x axis) and PC2 (y axis); Rose Bay, Sugarloaf Bay, Manly Beach and Manly Dam (hollow symbols) had their variances explained by PC1 alone (Fig. 3(a)).

Loadings are estimations of how much each variable (in this study, the intensity of the Raman signal at a given wavelength) contributes to variability on each PC. In order to understand the origin of spectral variabilities from PCA findings, the loadings plot (Fig. 3(c)) need to be interpreted along with the scores plot (Fig. 3(a)). For a given PC, positive (negative) values on the scores plot exhibit higher (lower) than average values for wavelengths with positive (negative) loadings. Negative scores have lower (higher) than average values for wavelengths with positive (negative) loadings.

PC-1 explained 97% of the total variance in the general model and its loadings plot is shown in Fig. 3(b). There were two domains of variability, with negative loadings values for wavelengths from 560 nm to 635 nm, corresponding to the spectral region of DOM

Optics EXPRESS

fluorescence; and from 665 nm to 700 nm, including the Chl-a fluorescence peak at 680 nm. Negative loadings at these spectral regions indicate higher than average values for samples with negative scores for PC-1 (Fig. 3(a)), represented by Rhodes, Clontarf and Manly Dam. Ultimately, these samples exhibited higher than average variance at spectral regions with known spectral signatures associated with fluorescence of common optically active constituents in natural waters. Rose Bay, Sugarloaf Bay and Manly Beach had positive PC-1 scores.

PC-2 accounted for 3% of total modelled variance among locations and its loadings plot show well-defined negative peaks at around 580 nm, area of *Gelbstoff* fluorescence and 680 nm (Chl-a) (Fig. 3(c)). Negative loading peaks were associated with high variabilities for negative scores, here attributed to Rhodes (Fig. 3(a)), indicating higher than average values for fluorescence on this sample. Conversely, the analysis indicated lower than average values for fluorescence at these well-defined peaks in the Clontarf sample. Rose Bay, Sugarloaf Bay, Manly Beach and Manly Dam samples exhibited PC-2 scores close to zero, indicating PC-2 doesn't explain significant spectral variances on these samples.

In summary, the PCA analysis reveals systematic but complex differences between the water samples investigated. It is interesting to note that including a freshwater sample (Manly Dam) didn't compromise the model performance. It may be indicative that, when temperature variation is excluded, background fluorescence seems to be more important than salinity effects on spectral variance among these natural water samples.

Having established that fluorescence largely accounts for the background effects we observe, we sought to explore the extent to which fluorescence from different species of phytoplankton might interfere with measurements of the Raman spectra for natural waters, and to quantify the concentration of phytoplankton that might substantially modify the Raman signal. To do this we selected species of phytoplankton that have a range of characteristic pigments that are broadly representative of the major phytoplankton groups such as zeaxanthin, phycocianin, phycoerythrin, carotene and violaxathin. These were *Synechococcus* Red, *Synechococcus* Green, *Nannochloropysis sp.*, *Ditylum Brightwelli, Dunallella tertiolecta, Ostreopsis siamensis* and *Rhodomonas salina*. Dilute samples of each were placed in a Varian Eclipse fluorescence spectrometer and emission and excitation spectra recorded. Each species, with the exception of *Synechococcus* Red, when excited at 532 nm exhibited fluorescence with a broad (typically 10-30 nm full width at half maximum) peaks in the range 660-685 nm. When excited at the shorter wavelength of 473 nm, *Synechococcus* Red and *Rhodomonas salina* exhibited additional peaks at 560 and 590 respectively. The fluorescence from phytoplankton has been widely studied and our observations are consistent with the literature [16].

Having observed the overlap between phytoplankton fluorescence and the Raman signal, we set about investigating what phytoplankton concentrations would cause significant distortion to the Raman signal. In order to do this, Raman spectra were recorded while small amounts of phytoplankton were added to the f/2 growth medium. The intensity of the Raman signal exhibited considerable variation after the phytoplankton was added, and we hypothesise that this variation was due to phytoplankton drifting across the excitation volume.

Selected fluorescence spectra obtained using 532 nm excitation are shown in Fig. 4. Figure 4(a) shows the Raman spectrum obtained for the f/2 growth medium, and it can be seen that the Raman band is clean, with relatively low background. Substantial changes were observed as phytoplankton was added, with the Raman band then being superimposed upon a fluorescence pedestal. For *Nannochloropsis* and *Synechococcus* Green, the fluorescence was peaked around 680 nm, while for *Synechococcu*s Red, substantial fluorescence occurred around 575 nm. The corresponding spectra are shown in Fig. 4, along with the corresponding phytoplankton cell counts that were determined using the counting procedure outlined earlier. The concentrations are comparable to those found in nature, and thus it is clear that

Optics EXPRESS

the presence of phytoplankton in natural water samples can potentially compromise the efficiency with which temperature can be determined via Raman spectroscopy.

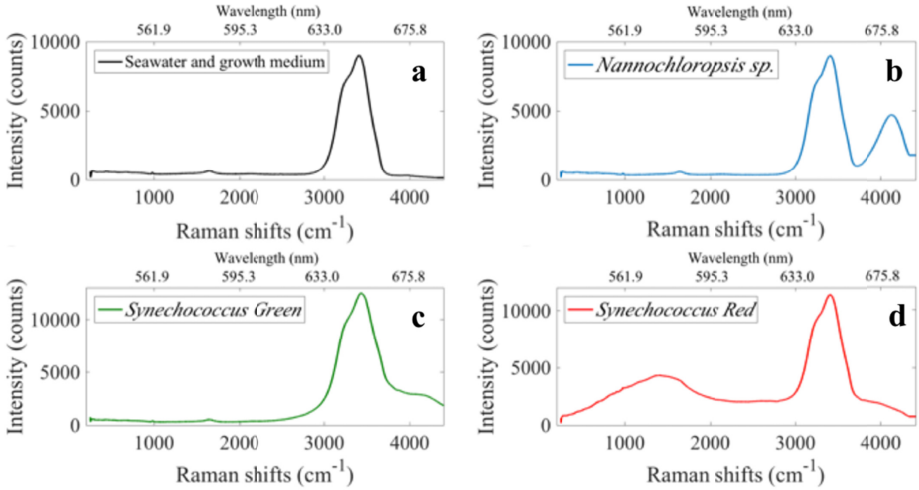

Fig. 4. Raman spectra obtained for autoclaved seawater with f/2 growth medium with (a) no added phytoplankton, (b) *Nannochloropsis* at 9.14 x10^3 cells per mL, (c) *Synnecococcus* Green at 7.78x10^4 cells per mL, and (d) *Synnecococcus* Red at 1.23x10^5 cells per mL.

4. Improving RMSTE by post-acquisition baseline correction

On the basis of the foregoing measurements and analysis, we attribute the background signals observed for water samples to fluorescence, predominantly from DOM and phytoplankton. It has been further established that naturally occurring phytoplankton concentrations are sufficient to significantly perturb the Raman spectra collected using excitation at 532 nm. In [14] it was suggested that the optimal excitation wavelength for Raman excitation was below 520 nm, however 532 nm lasers are more readily available and already used for other oceanographic measurements. Next, we will explore baseline correction techniques which make it feasible predicting water temperature with high accuracy using the two-colour method in natural waters.

It is apparent from the temperature maps of Fig. 2 that fluorescence background signals reduce the accuracy with which temperature can be determined, and the optimal centre wavelength for the two channels is substantially constrained. Here we analyse the accuracy with which temperature can be predicted from the raw Raman spectra, and present two effective methods of correction post acquisition. The channels selected for two-colour prediction analysis are centred at 3200 cm^{-1} and 3600 cm^{-1}, a compromise based on the RMSTE maps for all the water samples investigated here. These channels are depicted in Fig. 5 and will be referred to as “low shift” and “high shift” channels hereafter.

Optics EXPRESS

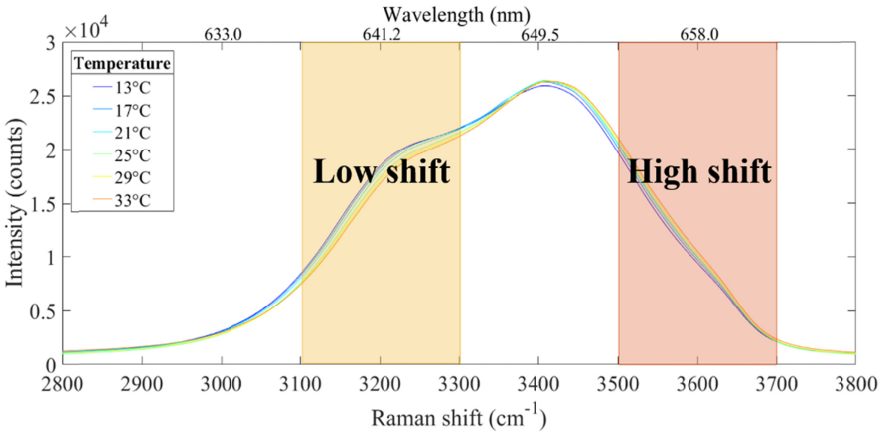

Fig. 5. Integration channels (200 cm^{-1} bandwidth) superimposed on unpolarised Raman spectra for Rose Bay sample.

RMSTEs for uncorrected Raman signal of various water samples were determined simply by integrating the signals within the low shift and high shift bands, taking their ratios and making a linear regression against reference temperature. Accuracies varied from ± 0.3°C to ± 2.6°C, with lowest RMSTE values found for Rose Bay and higher values for freshwater samples from Manly Dam and the inner harbour location of Rhodes. The association between larger baselines and lower RMSTEs reinforces the notion that fluorescence compromises the accuracy of temperature prediction. It can be seen by inspecting the data in Fig. 7(a) that there are considerable differences in the slopes and the actual values of the two-colour markers for water samples from different locations.

Next, we present two methods by which this accuracy was found to be improved: tilted baseline correction and an entirely new method: correction by temperature marker values.

4.1. Method 1: tilted baseline correction

Tilted baseline is a background correction technique established on the premise that the measured Raman signal should be close to zero on either side of the OH stretching band. For this study, points at 2800 cm^{-1} and 3750 cm^{-1} for each spectrum were used to define a "tilted baseline" that was subtracted from each spectrum to yield new, "baseline-corrected" spectra which were then analysed according to the usual two-colour method.

Figure 7(b) shows the predicted RMSTEs for the various water samples after applying the tilted baseline method. When compared to uncorrected RMSTE data all locations showed lower RMSTEs as a result of the tilted baseline correction. The technique has proven to be a simple data processing technique for reducing the impact of fluorescence on Raman spectroscopy water temperature prediction and inspection of Fig. 7(b) shows that while the tilted baseline correction has been effective in making the slopes more uniform, there is still considerable variation in the values of the two-colour markers. It was this observation, that fluorescence affects both slopes and marker values, that led us to propose a second correction method. It should be noted, however, that there is no physical reason to assume the baseline is in fact linear.

4.2. Method 2: correction by temperature marker values

Previously, it was observed that Rose Bay samples are the least impacted by fluorescence signals and have very low baseline. These samples had been subject to filtration and UV treatment leading to low concentrations of phytoplankton, and due to the location being close to the harbour mouth, DOM concentration is also expected to be low. Hence, we assume the

Optics EXPRESS

temperature markers calculated by two-colour analysis for this location represent a "standard" set of temperature markers for Sydney Harbour seawater, which are minimally influenced by background fluorescence.

We now present an entirely new approach to baseline correction based on the differences in temperature markers between natural water samples and the Rose Bay standard. Our premise (for the purpose of this correction method) is that the natural water sample comprises water that is similar to that at Rose Bay, plus additional constituents that fluoresce. Therefore, the two-colour markers for the natural water sample deviate from the two-colour markers for Rose Bay ratio by an amount due to that fluorescence (Eq. (1)).

$$R_{N_{T1}} = R_{RB_{T1}} + \Delta R_{Fl}, \tag{1}$$

where $R_{RB_{T1}}$ indicates the two-colour marker for the Rose Bay standard at temperature T_1, $R_{N_{T1}}$ is the ratio for a natural water sample and ΔR_{Fl} represents the influence of fluorescence on the marker for natural water samples.

As established earlier, the distribution of fluorescence signals across the Raman spectrum of natural waters is not homogenous. Accordingly, we introduce two fluorescence variables into the model: X (corresponding to fluorescence in the low shift band) and Y (corresponding to fluorescence in the high shift band), as indicated in Fig. 6 and in Eq. (2).

$$R_{N_{T1}} = \frac{B_{N_{T1}}}{A_{N_{T1}}} = \frac{B_{RB_{T1}} + Y}{A_{RB_{T1}} + X} \tag{2}$$

with $B_{N_{T1}}$ representing the integrated signal for the high shift channel (200 cm^{-1} width and centred at 3600 cm^{-1}) and $A_{N_{T1}}$ representing the integrated signal for the low shift channel (200 cm^{-1} width and centred at 3200 cm^{-1}).

Two parameters were extracted from the measured spectra to solve Eq. (2) for X and Y values (Eqs. (3) and (4)):

$$\Delta R_{Fl} = R_{N_{T1}} - R_{RB_{T1}} \tag{3}$$

and

$$C_{T1} = \frac{R_{N_{T1}}}{R_{RB_{T1}}}. \tag{4}$$

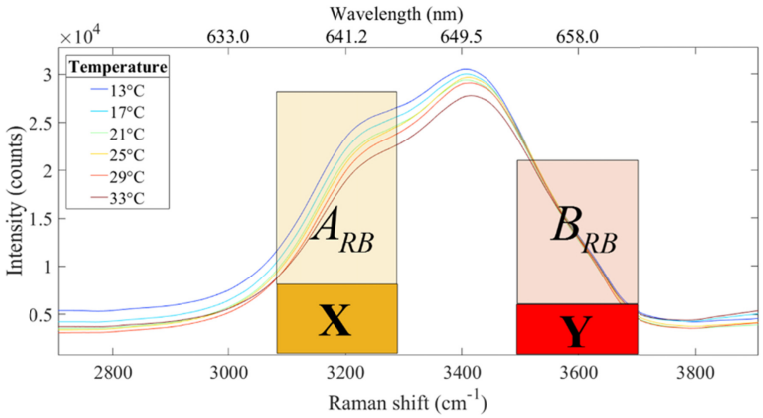

Fig. 6. Depiction of the premise underlying correction method 2. The Raman spectra for Clontarf within a channel comprises Raman photons plus fluorescence photons.

Optics EXPRESS

Some readers may wonder whether it would be more convenient to find X and Y from X = A_N-A_{RB} and Y = B_N-B_{RB}. The signals collected in individual channels may vary according to the apparatus used, variations in laser power etc. The use of ratios for makes our methods much more robust and compatible with implementation in the field.

For each water sample considered here, the Y values were considerably smaller than the X values. Y values were found to be independent of temperature, allowing for the use of an average value when calculating the new temperature marker. Unexpectedly, and for reasons which are not clear, X values showed a systematic variability with temperature. We believe that photobleaching of the fluorescing components in the focal volume may be responsible, but are yet to test this hypothesis. Therefore temperature-dependent X values were used when calculating the new temperature marker (*i.e* $X_{T1}, X_{T2} \cdots X_{Tn}$). By this process, corrected data sets were calculated, and are shown in Fig. 7(c). Predicted values for RMSTEs after correction by ratios, alongside the RMSTE for uncorrected data, are shown in Fig. 7(d). Method 2 was effective in that the corrected data for different location exhibited similar slopes and similar values, with improved temperature predictions in all water samples. RMSTEs after correction by method 2 ranged from ± 0.2°C to ± 0.3°C, therefore the method appears to be very effective and can in principle be applied to any water sample. In the case of our freshwater sample (Manly Dam), reverse-osmosis water was used as the "standard".

The RMSTEs for uncorrected data and the data after correction by either method 1 or method 2 are presented in Fig. 7(d) for the water samples from different locations. Both methods had the effect of improving the quality (R^2 value) and the linear fit (Figs. 7(a) - 7(c)), and both methods were effective in substantially increasing the accuracy with which temperature can be predicted. Method 2 was more effective than method 1 for Clontarf and Manly Beach samples.

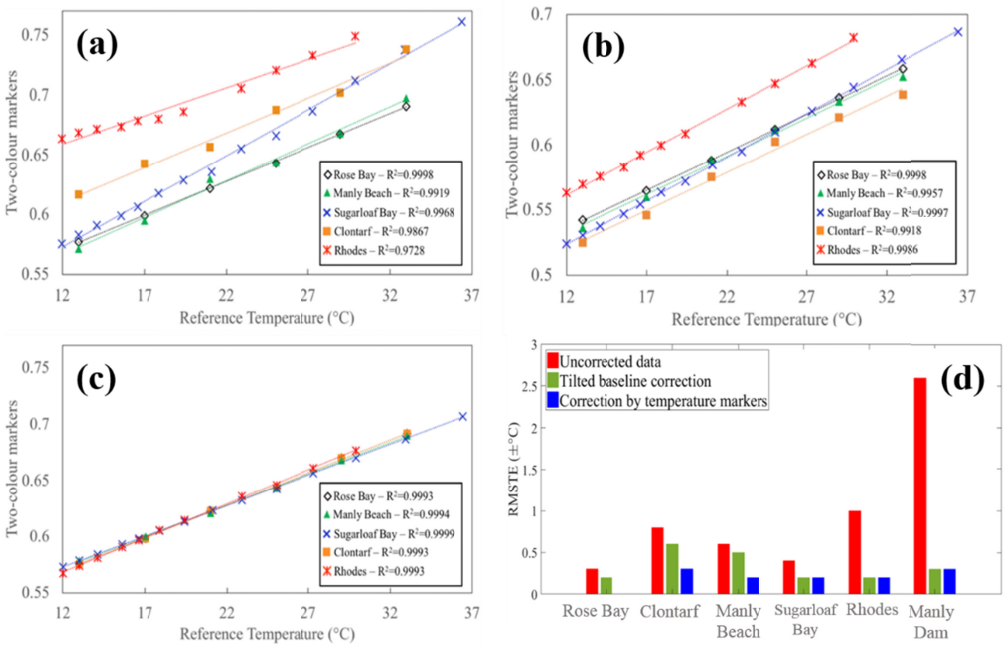

Fig. 7. Two-colour markers are plotted as a function of reference temperature for (a) uncorrected data; (b) data corrected by method 1; (c) data corrected by method 2; (d) presents the RSMTEs for samples from different locations with and without correction.

Optics EXPRESS

5. Discussion

The marine environment is known for its complex and dynamic nature, especially in coastal areas, resulting in high variability of water properties such as active optical components, salinity and vertical thermal stratification. For these methods presented here to be truly useful, they need to be applicable to a range of water samples in which fluorescing material may be present, and this paper is intended to advance that cause.

We have taken a qualitative approach to establishing that the baseline found in our Raman spectra for natural water samples arise mainly from DOM and Chl-a fluorescence. Furthermore, we have measured the phytoplankton concentrations that give rise to substantial fluorescence which overlaps with the Raman spectral band.

We have demonstrated two methods for baseline correction which are effective in increasing the accuracy with which temperature can be determined in natural waters. In terms of ease of field implementation, both methods present distinctive challenges. The tilted baseline correction method would require an additional two channels for collecting the very weak signals arising from fluorescence on either side of the Raman band, thereby reducing the signal-to-noise ratio for each of the main channels. A second drawback of this method is that there is no physical reason to expect the fluorescence spectrum to be linear around the Raman feature, and this is the fundamental assumption of the method

Method 2 does not require special filters or impact directly on SNR but it does require a database of temperature-dependent ratios as standards for comparison. In principle, ratios for all Jerlov water types could be obtained and used for calibration; alternatively, a local water sample can be collected and analysed to determine the correction factors required for analysis.

The work presented here provides proof of concept and future work will test the viability of this promising method. To this end, we have recently reported [19] a 4-channel Raman spectrometer that uses pulsed excitation and fast detection by photomultipliers. It has been used, in the laboratory, to predict water temperature of 3 natural water samples with accuracies in the range ± 0.4 °C to ± 0.8 °C. That recent work is an important step in validating the approach outlined in this manuscript, because it proves that it is not necessary to acquire the full Raman spectrum in order to predict temperature. We anticipate that the correction methods described here could lead to improved accuracies in predicting temperature, and also expand the range of waters that can be covered by a single model.

6. Conclusion

The findings in this paper help us to move one step forward our final goal: a LIDAR-compatible custom-built multichannel Raman spectrometer able to measure depth-resolved temperature data *in situ* as firstly proposed in [8].

We have shown that Raman spectra collected from natural water samples around Sydney Harbour exhibit background signal levels that adversely affect the accuracy with which temperature can be determined. We have shown the background signals arise from fluorescence, allegedly from DOM and Chl-a, and quantified the phytoplankton concentrations that cause distortion of the OH stretching band. We have proposed two methods of baseline corrections that are effective in improving temperature accuracy and considered how they could be implemented in the field. There is scope for a systematic fluorescence and Raman spectroscopic study that considers a wider range of samples which are characterised in terms of DOM concentration, chlorophyll-a concentration and salinity.

Funding

Australian Research Council (FT120100294).

Optics EXPRESS

Acknowledgments

Dr. Andréa de Lima Ribeiro and Dr. Christopher Artlett gratefully acknowledge receipt of Macquarie University iMQRES/MQRES PhD scholarships. Prof. Helen Pask gratefully acknowledges receipt of an Australian Research Council Future Fellowship (project number FT120100294).

References

1. G. E. Walrafen, "Raman spectral studies of the effects of temperature on water structure," J. Chem. Phys. **47**(1), 114–126 (1967).
2. G. E. Walrafen, M. R. Fisher, M. S. Hokmabadi, and W. H. Yang, "Temperature dependence of the low- and high-frequency Raman scattering from liquid water," J. Chem. Phys. **85**(12), 6970–6982 (1986).
3. C. H. Chang and L. A. Young, *Seawater Temperature Measurement from Raman Spectra* (National Technical Information Services, 1972).
4. D. A. Leonard, B. Caputo, and F. E. Hoge, "Remote sensing of subsurface water temperature by Raman scattering," Appl. Opt. **18**(11), 1732–1745 (1979).
5. B. Breschi, G. Cecchi, L. Pantani, V. Raimondi, D. Tirelli, and G. Valmori, ""Measurement of water column temperature by Raman scattering," EARsel Adv. Remote Sens. **1**(2), 131–134 (1992).
6. Z. Liu, J. Zhang, W. Chen, X.-S. Huang, and J. Ma, "Remote sensing of subsurface water temperature using Raman LIDAR," Proc. SPIE **1633**, 321–329 (1992).
7. A. F. Bunkin, V. K. Klinkov, V. N. Lednev, D. L. Lushnikov, A. V. Marchenko, E. G. Morozov, S. M. Pershin, R. N. Yulmetov, A. F. Bunkin, V. K. Klinkov, V. N. Lednev, D. L. Lushnikov, A. V. Marchenko, E. G. Morozov, S. M. Pershin, and R. N. Yulmetov, "Remote sensing of seawater and drifting ice in Svalbard fjords by compact Raman lidar," Appl. Opt. **51**(22), 5477–5485 (2012).
8. C. P. Artlett and H. M. Pask, "Optical remote sensing of water temperature using Raman spectroscopy," Opt. Express **23**(25), 31844–31856 (2015).
9. R. M. Measures, W. R. Houston, and D. G. Stephenson, "Laser induced fluorescent decay spectra - a new form of environmental signature," Opt. Eng. **13**(6), 494–501 (1974).
10. K. Kalle, "The problem of the Gelbstoff in the sea," Oceanogr. Mar. Biol. Annu. Rev. **4**, 91–104 (1966).
11. C. D. Mobley, *Light and Water: Radiative Transfer in Natural Waters* (Academic, 1994).
12. J. E. James, C. S. Lin, and W. P. Hooper, "Simulation of laser-induced light emissions from water and extraction of Raman signal," J. Atmos. Ocean. Technol. **16**(3), 394–401 (1999).
13. C. Lin, "Tunable laser induced scattering from coastal water," IEEE Trans. Geosci. Remote Sens. **37**(5), 2461–2468 (1999).
14. C. S. Lin, "Characteristics of laser-induced inelastic-scattering signals from coastal waters," Remote Sens. Environ. **77**(1), 104–111 (2001).
15. R. R. L. Guillard, "Culture of phytoplankton for feeding marine invertebrates," in *Culture of Marine Invertebrate Animals*, W. L. Smith and M. H. Chanley, eds. (Springer, 1975), pp. 29–60.
16. S. Roy, C. Llewellyn, E. S. Egeland, and G. Johnsen, eds., *Phytoplankton Pigments* (Cambridge University, 2011).
17. H. E. Nystad, "Comparison of principal component analysis and spectral angle mapping for identification of materials in Terahertz transmission measurements," Master thesis, Norwegian University of Science and Technology (2015).
18. H. Abdi and L. J. Williams, "Principal component analysis," Wiley Interdiscip. Rev. Comput. Stat. **2**(4), 433–459 (2010).
19. A. de Lima Ribeiro, C. Artlett, and H. Pask, "A LIDAR-compatible, multichannel Raman spectrometer for remote sensing of water temperature," Sensors (Basel) **19**(13), 2933 (2019).