Received: 30 September 2020 | Revised: 19 August 2021 | Accepted: 20 August 2021

DOI: 10.1002/hyp.14357

SCIENTIFIC BRIEFING

HP TODAY WILEY

High-frequency multi-solute calibration using an in situ UV–visible sensor

Juan Pesántez[1] | Christian Birkel[2] | Giovanny M. Mosquera[1,3] | Pablo Peña[1] | Viviana Arízaga-Idrovo[1] | Emma Mora[1] | William H. McDowell[4] | Patricio Crespo[1]

[1]Departamento de Recursos Hídricos y Ciencias Ambientales & Facultad de Ingeniería, Universidad de Cuenca, Cuenca, Ecuador

[2]Department of Geography and Water and Global Change Observatory, University of Costa Rica, San José, Costa Rica

[3]Instituto Biósfera, Universidad San Francisco de Quito USFQ, Diego de Robles y Vía Interoceánica, Quito, Ecuador

[4]Department of Natural Resources and the Environment, University of New Hampshire, Durham, New Hampshire, USA

Correspondence
Juan Pesántez, Departamento de Recursos Hídricos y Ciencias Ambientales & Facultad de Ingeniería, Universidad de Cuenca, Av. 12 de Abril, Cuenca, Ecuador.
Email: juanp.pesantezv@ucuenca.edu.ec

Funding information
This research was funded by the IAEA research contract 22906 and Central Research Office (DIUC) of the Universidad de Cuenca in the framework of the project 'Evaluation of Non Stationary Hydrological Conditions in the Andean Páramo'.

Abstract

Monitoring the temporal variation of solute concentrations in streams at high temporal frequency can play an important role in understanding the hydrological and biogeochemical behaviour of catchments. UV–visible spectrometry is a relatively inexpensive and easily used tool to infer those concentrations in streams at high temporal resolution. However, it is not yet clear which solutes can be modelled with such an in-situ sensor. Here, we installed a UV–visible spectrometer probe (200–750 nm) in a high-altitude tropical Páramo stream to record the wavelength absorbance at a 5-min temporal resolution. For calibration, we simultaneously sampled stream water at a 4-h frequency from February 2018 to March 2019 for subsequent laboratory analysis. Absorbance spectra and laboratory-determined solute concentrations were used to identify the best calibration method and to determine which solute concentrations can be effectively inferred using in situ spectrometry through the evaluation of six calibration methods of different mathematical complexity. Based on the Nash–Sutcliffe efficiency (NSE) and Akaike information criterion metrics, our results suggest that multivariate methods always outperformed simpler strategies to infer solute concentrations. Eleven out of 21 studied solutes (Al, DOC, Ca, Cu, K, Mg, N, Na, Rb, Si and Sr) were successfully calibrated (NSE >0.50) and could be inferred using UV–visible spectrometry even with a reduced daily sampling frequency. It is worth noting that most calibrated solutes were correlated with wavelengths (WLs) in the low range of the spectra (i.e., UV range) and showed relatively good correlation with DOC. The latter suggests that estimation of metal concentrations could be possible in other streams with a high organic load (e.g., peat dominated catchments). In situ operation of spectrometers to monitor water quality parameters at high temporal frequency (sub-hourly) can enhance the protection of human water supplies and aquatic ecosystems as well as providing information for assessing catchment hydrological functioning.

KEYWORDS
calibration, high frequency monitoring, Páramo, solutes, UV–visible spectrometry, water quality

1 | INTRODUCTION

Monitoring the temporal variation of solutes (nutrients and metals) in streamflow plays an important role in understanding the hydrological and biogeochemical behaviour of catchments when used as tracers in hydro-geochemical models (Birkel et al., 2017; Correa et al., 2019; McGuire & McDonnell, 2015). Observations of these solutes as water quality parameters are also fundamental to investigate the dynamics of instream biogeochemical processes, as well as hydrological connectivity in ecosystems (Lloyd et al., 2016; McDowell & Asbury, 1994; McGuire & McDonnell, 2010). Understanding these processes plays a critical role in human and ecosystem health, particularly when related to the use of water for human consumption (Kaushal et al., 2018; Lin et al., 2002; Nimick et al., 2003) and the protection and preservation of aquatic habits (Schmutz & Sendzimir, 2018; Soulsby et al., 2016).

Nevertheless, despite the advantages of monitoring nutrients and metals in stream water, their application at high temporal frequency and over longer periods in hydrological, biogeochemical, and water quality studies is still limited to selected experimental sites (Green et al., 2020; Heathwaite & Bieroza, 2021; Vorobyev et al., 2019; Wymore et al., 2018). This mainly arises due to the high financial cost associated with field sampling and laboratory analyses of water samples. These constraints usually limit the monitoring of solute concentrations at most sites to a few times during the year or specific streamflow conditions, hampering the capacity to understand in-stream changes in hydrological and biogeochemical conditions, and medium- to long-term changes driven by the effects of global environmental stressors such as changes in land use and climate.

There is a growing interest in exploring techniques to estimate stream water solute concentrations at a high temporal frequency and over longer periods (Cook et al., 2017; Etheridge et al., 2014; Ruhala & Zarnetske, 2017; Thomas & Causse, 2017). Among the techniques used for monitoring solute concentration in situ and in-stream are ion chromatography (IC), which is only used under controlled conditions (von Freyberg et al., 2017), and UV–visible spectrometry. Determining the concentrations of metals represents a particular challenge as their measurement requires the use of expensive complexometry or reagents making them unsustainable long-term (Kulkarni et al., 2015; Xu et al., 2014), or IC (von Freyberg et al., 2017), which in remote areas is not feasible to maintain. In contrast, UV–visible spectrometry can be relatively easily deployed and operated anywhere for long-term and high-resolution monitoring. UV–visible spectrometry has been previously used to estimate in situ stream water nutrient concentrations such as carbon, nitrogen, and phosphorus by various authors (Waterloo et al. (2006); Koehler et al. (2009); Avagyan et al. (2014); Etheridge et al. (2014); Lopez-Kleine and Torres (2014); Huebsch et al. (2015); Cook et al. (2017); Ruhala and Zarnetske (2017).

UV–visible spectrometry allows inferring the concentration of solutes with molecular structures known to absorb light at specific wavelengths, such as the NO_3^- ion, that can require site-specific corrections to obtain maximum accuracy (e.g., Snyder et al., 2018). Although UV–visible sensors are widely used to estimate DOC, DOM, and NO_3^- in situ, there are still considerable uncertainties regarding their use to estimate other solute concentrations (Huebsch et al., 2015; Ruhala & Zarnetske, 2017; Snyder et al., 2018; Thomas & Burgess, 2007). It is unclear whether a spectrometer can be used more generally under different environmental conditions and to estimate a broad range of solutes for which no specific absorption wavelength is known (McIntyre et al., 1982). Despite the advantages of optical techniques (in particular UV–visible spectrometry), we are not aware of studies that have inferred metal concentrations in situ using spectrometry. Prior studies analysing metals through spectrometry were carried out under controlled conditions in the laboratory applying a reagent to identify the solute of interest (Aguerssif et al., 2008; Garcia Rodriguez et al., 1998; Kulkarni et al., 2015; Zhou et al., 2019). Hence, there is a need to explore methods that use spectrometry to infer the concentration of a variety of metals and nutrients in situ with relatively low maintenance costs.

Past studies proposed different calibration methods using UV–visible spectrometry (Cook et al., 2017; Garcia Rodriguez et al., 1998; Thomas & Burgess, 2007; Wei et al., 2018). The simplest approach uses a single absorbance at a determined wavelength. However, such simple approaches may lose important information from analysing wavelengths of suspended particles or other types of interferences in the water (Shi et al., 2020, 2021; Snyder et al., 2018). More complex calibration methods are based on multivariate statistics that consider all or selected information recorded in the UV–visible spectrum (e.g., the 5–10 most important WLs using the PLS method used by manufacturers for calibration [Langergraber et al., 2003]). Multivariate methods considering a larger amount of information from the UV–visible spectrum are thus likely to take into account the aforementioned interferences. Simple calibration methods do not require much mathematical computation, but may not be as accurate as multivariate methods (Carré et al., 2017; Peacock et al., 2014; Thomas & Burgess, 2007; van den Broeke, 2007). Although different calibration methods were previously tested and compared, a more complete comparison of the spectrum range for each method is still needed to enhance our ability to conduct in situ monitoring of stream water solutes at high temporal frequency using spectrometry.

The purpose of this scientific briefing is to analyse and compare different calibration methods for inferring in situ solute concentrations (nutrients and metals) using data from a UV–visible spectrometer at an experimental catchment in a high-elevation Páramo ecosystem that is representative of catchments with organic-rich stream waters in other climates and geomorphic settings.

2 | MATERIALS AND METHODS

2.1 | Study site

The study was conducted at the Zhurucay Ecohydrological Observatory (Figure 1), southwest of the city of Cuenca in southern Ecuador (3°4′S, 79°14′W). The observatory is on the west slope of the Andean mountain range at 3400–3900 m a.s.l. The climate in the study region

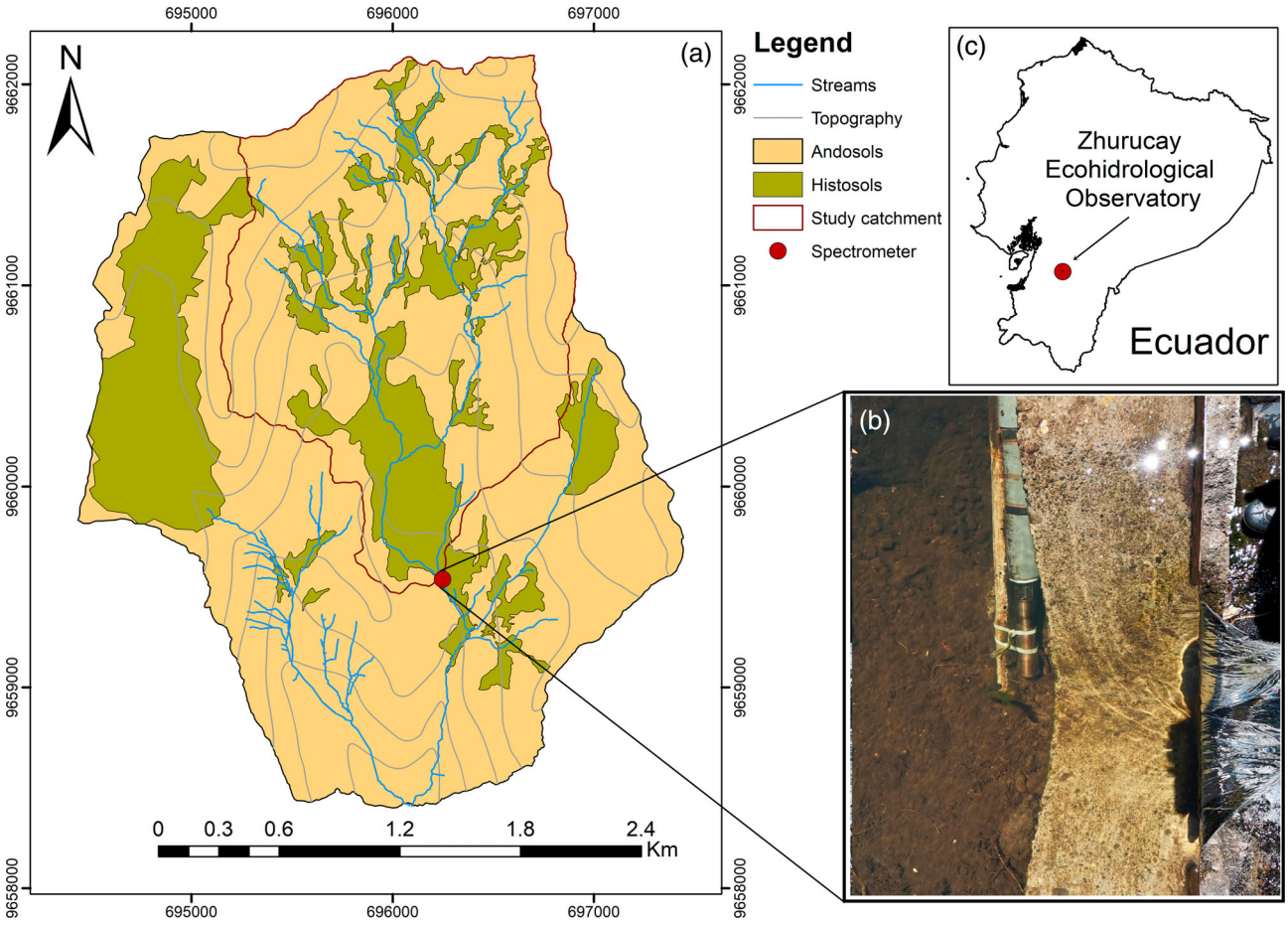

FIGURE 1 (a) Soil type map of the Zhurucay Ecohydrological Observatory showing the position of the spectrometer installed, (b) spectrometer installed in the river and (c) location of the observatory within Ecuador

is mainly influenced by air masses originating from the Amazon forest (Esquivel-Hernández et al., 2019). Mean annual temperature and precipitation during the period December 2017 to March 2019 were 6.0°C and 1300 mm at 3779 m a.s.l., respectively. The maximum observed temperature was 15.9°C and the minimum −2.4°C (Córdova et al., 2015). The study catchment is part of the observatory with a drainage area of 3.28 km^2 and ranges from 3676 to 3900 m a.s.l. with an average slope of 18%. The catchment geology is composed of the Quimsacocha formation and mostly basaltic flows of plagioclases, feldspars, and andesitic pyroclastic Miocene or younger rocks in turn deposited on Palaeozoic and Mesozoic rocks (Beate, 1999). The catchment represents typical Páramo (tropical alpine) ecosystem characteristics and is mainly covered with tussock grasses (*Calamagrostis* sp. and *Festuca* sp.) covering Andosols (74% of the catchment) on the hillslopes. Cushion plants cover Histosols (22% of the catchment) in the riparian areas, and in some plains within the catchment (Mosquera et al., 2015). A small area (4%) is covered with Leptosols. Andosols and Histosols are organic-rich soils with high acidity (e.g., pH = 4.7) (Quichimbo et al., 2012). Extensive volcanic ash deposits result in high concentrations of aluminium, iron, and other metals in the soil (Buytaert, Sevink, et al., 2005; Buytaert, Wyseure, et al., 2005), which are probably continuously exported to the streams (Correa et al., 2017, 2019). These metals are associated with organic components forming organo-metallic complexes in the soil and are then released to the streams depending on the hydrometeorological conditions (Buytaert, Sevink, et al., 2005; Correa et al., 2019; Pesántez et al., 2018).

2.2 | In situ monitoring and water sampling

We operated a UV–visible (200–750 nm) Spectrometer Probe V2 (Spectrolyser, s::can Messtechnik GmbH, Vienna, Austria) in the stream at the catchment outlet (Figure 1) to record the wavelength absorbance at a 5-min temporal resolution from February 2018 to March 2019. In order to record reliable wavelength measurements a cleaning system with a Compressor V4.0 (Spectrolyser, s::can Messtechnik GmbH, Vienna, Austria) was installed together with the spectrometer and activated every 15 min. In addition, manual cleaning of the monitored stream section and the probe was carried out twice a week to avoid sediment and algae accumulation. A time series analysis was performed to identify inconsistencies in the data resulting from the cleaning process (as a result 20% of the data was not included in the analysis). A Conductivity-Pressure Smart Sensor

(CT2X, Seametrics, Washington) was used to measure water level and electrical conductivity (EC) data. Water level data were converted into discharge using the Kindsvater–Shen equations (U.S. Bureau of Reclamation, 2001) using the salt dilution method (Moore, 2004).

Samples for laboratory analyses were collected using a Portable Discrete Water Sampler (PVS4120D, Campbell Scientific, Inc., Utah; 4–6-h frequency) at the same site. The samples were taken during March 2018 to March 2019 for DOC and TNb and March 2018 to September 2019 for the rest of solutes. Samples from the autosampler were retrieved twice a week. The samples were filtered using 0.45 μm polypropylene membrane filters (Puradisc 25PP Whatman, Inc., Clifton, NJ) for metals and dissolved organic carbon. An extra unfiltered sample for total nitrogen (TNb) was taken. All the samples were stored in high-density polyethylene bottles which were washed before sampling in a thermodisinfection washer (GW1160, Smeg S.p. A, Guastalla [RE], Italy) and rinsed three times with type II distilled water and subsequently three times with type I distilled water. The bottles for metal analysis were washed in an acid washing device (traceClean, Milestone) using purified nitric acid. Water samples for metal analyses were acidified with ultrapurified nitric acid (pH < 2) on the same sampling collection day to avoid trace metal precipitation and adsorption. All samples were analysed within 2 months of field sampling. Water samples for DOC and TNb analysis were analysed the day after sampling or stored in the freezer (<−4°C) until analysis.

2.3 | Laboratory analyses

We determined solute concentrations – to be later compared to UV–visible absorbance inferred ones – in approximately 1800 water samples for carbon and nitrogen and around 1150 samples for the rest of the solutes (i.e., the complete dataset) at the Water and Soils Chemical Analysis Laboratory of the Department of Water Resources and Environmental Sciences of the University of Cuenca. Nineteen of the solutes targeted in this study (Al, As, B, Ba, Ca, Cd, Cr, Cu, Fe, K, Mg, Na, Pb, Rb, Si, Sr, V, Y, and Zn) were analysed with an ICP-MS Perkin Elmer 350X (Shelton, CT). All these analyses used calibration standards of Inorganic Venture, Inc., ISO 9001 registered and ISO 17025 accredited. We performed five calibration curves, each with six standard dilutions, prepared by weight/weight dilutions in Type I distilled water. Each calibration curve was used for five different groups of targeted compounds based on the range of concentrations of each group of solutes. This was done to avoid interferences between the metals due to their concentration ranges from mg/L to ng/L.

Dissolved organic carbon (DOC), measured as non-purgeable organic carbon (NPOC), and total nitrogen (TNb) concentrations were determined by combustion using a Vario TOC cube (Elementar, Germany). The standards were Potassium hydrogen phthalate BioXtra ≥99.95% (Sigma-Aldrich, MO) for NPOC, and Sodium Nitrate ReagentPlus ≥99.0% (Sigma-Aldrich, MO) and Ammonium Chloride ≥99.5% (VWR Analytical, PA) for TNb.

In all the analyses, the Pearson coefficient (r) of the calibration curves must always be greater than 0.99 to guarantee accuracy. The method detection limit (DL) was calculated as three times the standard deviation of the blank (Ellison & Williams, 2012). All the results obtained are the average of three repeated analyses. QA and QC analysis were performed every 10 samples using two standard dilutions for metals and one for NPOC and TNb. The acceptance criteria (accuracy) were: less than 16% of the relative standard deviation in mg/L, less than 45% in μg/L, and less than 60% in ng/L, according to the Horwitz equation (Thompson, 2004). If these criteria were not met for a certain interval, the corresponding samples were reanalyzed.

2.4 | Multi-solute calibration methods

To determine the stream water solutes that could be inferred using UV–visible spectrometry, we compared solute concentrations estimated using this technique against those determined for the 4-h field collected samples via laboratory. To infer solute concentrations, we used data sets for each solute (WLs vs. laboratory-determined concentrations for each solute) after quality control (due to the cleaning processes of the probe, system failures, or any problems related to storage, handling or analysis of the samples). Thus, we used 1571 observations for DOC and N, 784 observations for Al, and around 1100 observations for the rest of the solutes. Calibration of the UV–visible spectrometry based concentrations was conducted using different methods. As an initial inspection of the spectral behaviour, the entire wavelength (WL) absorbance range (200–750 nm) provided by the probe was tested at 2.5 nm intervals. The spectra (in the UV–visible range) recorded at the exact time of sample collection were compared against the laboratory-determined concentration of each solute in a simple linear regression analysis. This allowed us to identify the optimal WLs or peaks (those with the highest R^2) and the range of optimal WLs ('optimal ranges') in the spectra in which some of the variance ($R^2 > 0.15$) in the solute concentrations was explained by the absorbance at the WL or WLs analysed (Falk & Miller, 1992; Rights & Sterba, 2019).

We then divided each solute data set into calibration and validation subsets. These data sets were divided into six subsets and we randomly selected a 30% sample of each of them to compose the validation subset (Arsenault et al., 2018; Fernandez-Palomino et al., 2020; Tantithamthavorn et al., 2017). We used the *dplyr* package (Wickham et al., 2020) of the R statistical language (R Core Team, 2020) for subsetting and random sampling. Validation data were used to estimate the prediction error after calibration. The remaining 70% of the data was used for evaluating direct and indirect calibration methods to find the most suitable one for inferring the concentrations of each solute (for calibration and validation time series please refer to Appendix S2 and S3):

1. Direct calibration methods used the absorbance at one or more WLs directly against the laboratory-measured concentrations.
2. Indirect calibration methods were based on dimension reduction or multivariate statistics.

The direct methods included the single wavelength calibration, the ratio calibration, and the area under the spectra calibration. The identified optimal WL was tested against their respective solute concentrations as the single WL calibration method. The absorbance ratios between all possible pairs of WLs absorbance (each 2.5 nm within the whole measurement range of the spectrometer) were tested against each set of solute concentrations as the ratio WL method (Thomas & Burgess, 2007). The area under the UV–visible spectra provided by the sensor was used as a combination of all the WL absorbance values to infer the solute concentrations (Thomas & Burgess, 2007; Wang & Hsieh, 2001).

The indirect or multivariate methods included the principal component regression method (hereafter referred to as PCR) and the partial least squares regression method (hereafter referred to as PLSR). For these methods we used all the absorbance values at each WL recorded in the UV–Vis range to infer solute concentrations. Both methods are based on dimension reduction of the explanatory variables (spectral absorbance reduction) through Principal Component Analysis (PCA). The main difference is that the PLSR method takes into account the response variable (observed data) when maximizing the criteria for the dimensionality reduction, while the PCR maximizes the criteria extracting the most relevant information from the explanatory variables without taking into account the relationship between the predictive variables (absorbance in this case) with the response variable (Singh & Sarkar, 2018). The PCR and PLSR methods have been shown to improve the prediction estimation of water samples with interferences (e.g., in situ stream water samples). The improvement is attributed to the simultaneous use of different spectra WLs absorbance (Geladi & Kowalski, 1986; Haaland & Thomas, 1988; Olivieri, 2018). These calculations were performed using the *pls* R package (Mevik et al., 2020).

2.5 | Multi-solute calibration model performance

Three quantitative metrics were applied to evaluate the performance of the aforementioned methods in calibration and validation and to objectively select the best performing and most parsimonious model: Nash–Sutcliffe efficiency in calibration (NSEc) and validation (NSEv; Nash & Sutcliffe, 1970); normalized root mean square error in calibration (nRMSEc) and validation (nRMSEv); and Akaike's Information Criterion in calibration (AICc) and validation (AICv; Akaike, 1974). Use of nRMSE normalization around the mean concentration was chosen to avoid bias among solutes because of the different concentration ranges or scales. A nRMSE closer to zero indicates a better model fit. AIC represents the level of model parsimoniousness aiding model selection, as it penalizes the number of parameters used. A lower AIC value indicates the most parsimonious model. The NSE was used to determine the solutes that can be adequately inferred using UV–visible spectrometry. NSE values >0.65 were considered as a good fit and NSE > 0.50 as an acceptable fit to the laboratory-measured concentrations (Moriasi et al., 2007). For our purpose, we accepted a calibration method as reasonable for NSE > 0.50.

In order to evaluate the effect of the sampling resolution to infer solute concentrations we calibrated and evaluated all the models using five data sets with different 'sampling' resolutions. The resulting paired data sets included: two samples per day (∼527 observations for DOC and N, 264 for Al and ∼413 for the rest of solutes), one sample per day (∼261 observations for DOC and N, 132 for Al and ∼206 for the rest of solutes), two samples per week (∼74 observations for DOC and N, 39 for Al and ∼60 for the rest of solutes), one sample per week (∼38 observations for DOC and N, 21 for Al and ∼31 for the rest of solutes), and a biweekly sample (∼18 observations for DOC and N, 10 for Al and ∼15 for the rest of solutes).

Spearman correlation analysis was used to identify relationships among the laboratory-determined concentrations of the studied solutes. Relationships among solutes could identify co-linear solutes as potential proxies (predictors) of the concentrations of other ones. A k-means cluster method was used to order the solutes that were most related to each other using the *corrplot* R package (Wei & Simko, 2017). Finally, we tested different regression types to account for potentially strong linear and/or non-linear relationships between laboratory determined DOC concentrations and the laboratory determined concentrations of the rest of solutes. This analysis was carried out to evaluate whether DOC laboratory-determined concentrations could be used as a surrogate to calibrate other solute concentrations without the need of the full WL spectra.

3 | RESULTS

3.1 | Analysing absorbance against solute concentrations

Several of the analysed solutes showed a relationship between laboratory-determined concentrations and specific WLs (Figure 2). The solutes showing the highest R^2 values (Ca, Mg, Na, Si, and Sr; $0.40 < R^2 < .62$) exhibited peak absorbance values at low WL values, generally between 200 and 210 nm. In addition, some solutes presented the highest R^2 values (0.62 for Al, 0.74 for DOC, and 0.63 for Cu) at higher WL values (230 nm for Al and DOC and 352.5 nm for Cu). This showed the importance of the UV range (200–400 nm) measured by the spectrometer for calibration. Only As, B, Fe, and V had peaks at longer WL values (720, 557.5, 527.5, and 720 nm respectively). However, their R^2 values were the lowest of the analysed solutes ($R^2 < 0.18$).

3.2 | Multi-solute calibration performance

3.2.1 | Direct solute calibration methods

Applying a simple linear regression model, DOC (NSEc = 0.74; NSEv = 0.70; WL = 230), showed a good fit (NSE ≥ 0.65) between the observed and modelled concentrations in calibration and validation (Figure 3 and Table 1). Al (NSEc = 0.62; NSEv = 0.58;

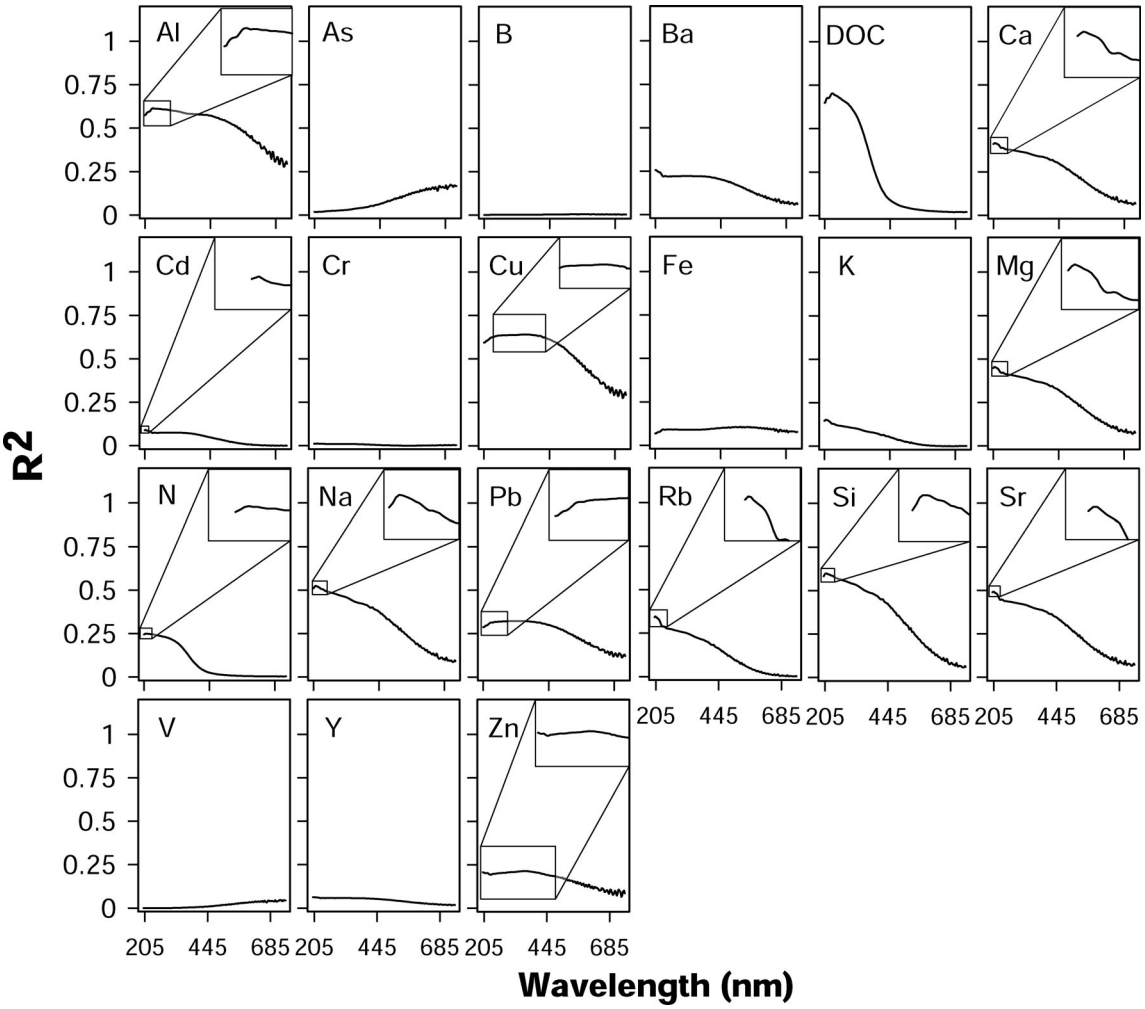

FIGURE 2 Coefficient of determination (R^2) variation between solute concentrations and absorbance at each wavelength (WL) analysed in the UV–visible spectra range. The inset shows the peak R^2 for the WL with the best correlation between solute concentration and absorbance

WL = 230 nm), Cu (NSEc = 0.63; NSEv = 0.65; WL = 352.5 nm), Si (NSEc = 0.56; NSEv = 0.60; WL = 210 nm), and Na (NSEc = 0.54; NSEv = 0.51; WL = 207.5 nm) gave an acceptable fit (i.e., 0.50 < NSE < 0.65) in calibration and validation. For the other solutes, the NSE was below 0.50 and considered not sufficient to be calibrated with the single WL method. For the solutes that can be calibrated using this method the nRMSE was between 20 and 33% in calibration and validation, respectively (Figure 3).

The results of the ratio method were similar to those obtained with the single WL method (Table 1). DOC (NSEc = 0.73; NSEv = 0.75; ratio WL = 210 / 202.5 nm) and Si (NSEc = 0.76; NSEv = 0.72; WL ratio 200/210 nm) presented a good fit. Likewise, Cu, Mg, Na, Rb, Sr, and Si had an acceptable fit. The NSEs of Mg, Rb, and Sr increased in calibration and validation from non-acceptable (NSE < 0.50) using the single WL methods to acceptable (0.53 < NSE <0.63) using the ratio method (Table 1). These ratio method results are in accordance with the single WL method and with the spectra behaviour related to the solutes (Section 3.1.) that showed a preference for the lower WLs. The nRMSE decreased by around 2.2% compared to the single WL method for As, B, Ba, Ca, Cd, Cr, K, Mg, N, Rb, Si, Sr, V and Zn, and increased for Al, DOC, Cu, Na, Pb and by around 1.6% (Figure 3). The last direct method used was the area under the spectra. Figure 3 also shows that the area under the spectra method reported similar NSE values compared to the single WL and ratio methods. The errors were also similar in relation to the other direct calibration methods.

The AIC values in both calibration and validation (Figure 3e–f) mostly coincided with the trends in NSE and nRMSE values for all direct methods, especially for those in which the calibration was possible. The AIC values of the single WL and ratio methods were higher (indicating less parsimoniousness) compared to the area under the spectra method. The results from the single wavelength and the ratio method showed similar AIC values.

3.2.2 | Indirect solute calibration methods

The indirect or multivariate PCR and PLSR methods were used to calibrate solute concentrations using spectrometry. The PCR showed an

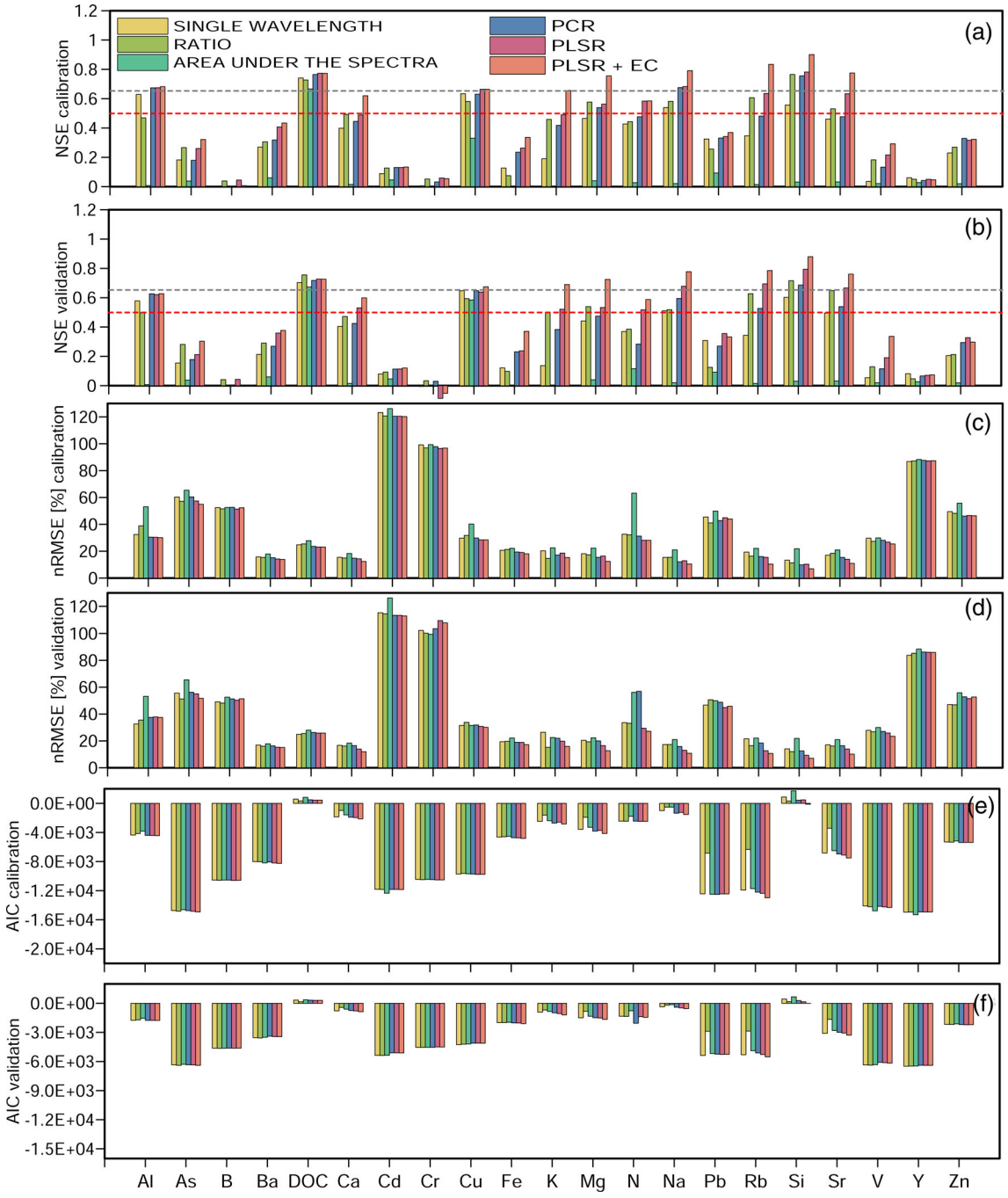

FIGURE 3 The Nash–Sutcliffe efficiency (NSE) in the calibration (a) and validation (b) datasets, the normalized root mean square error (%) (nRMSE) in the calibration (c) and validation (d) datasets, and the Akaike information criteria (AIC) in the calibration (e) and validation (f) datasets for each of the calibration methods applied in this study. In subplots (a) and (b) the red line represents an acceptable fit (i.e., 0.50 < NSE < 0.65) and the grey line represents a good fit (NSE ≥ 0.65)

improvement in the NSE and nRMSE compared to the direct methods for the calibration and validation datasets. Nevertheless, the improvement was not sufficient to infer the concentrations of additional solutes than those found with the direct methods (Figure 3a–d).

Use of PLSR improved the NSE compared to PCR for all solutes. The improvement allowed inferring the concentrations of N and Rb with an acceptable fit (NSEc: $N = 0.58$; Rb = 0.63, NSEv: $N = 0.51$; Rb = 0.69). Cu and Na improved significantly from an acceptable to a good fit to high goodness of fit (NSEc: Cu = 0.66; Na = 0.68, NSEv: Cu = 0.63; Na = 0.67). The nRMSE also improved for all solutes compared to all other methods (Figure 3a–d).

The available EC measurements were included as an additional predictor in the method that yielded the best results (i.e., PLSR) over the entire range of absorbance under the different WLs. The latter was because conducting salt species represented by EC can be an indicator of the ionic materials that were calibrated (Thomas &

Solutes	Best single WL (nm)	Best WL ratios (nm)			
	230	200/210	200/212.5	210/202.5	210/200
Al	**0.62 (0.58)**	**0.46 (0.49)**	**0.62**	**0.62**	**0.26**
As	720	445/457.5	557.5/640	567.5/637.5	557.5/605
	0.18 (0.15)	0.26 (0.28)	0.24	0.24	0.26
B	557.5	662.5/667.5	662.5/667.5	530/527.5	662.5/710
	0 (0)	0.03 (0.04)	0.04	0.03	0.04
Ba	200	620/662.5	657.5/690	432.5/435	432.5/437.5
	0.27 (0.21)	0.30 (0.29)	0.30	0.30	0.30
DOC	230	210/202.5	200/210	212.5/202.5	210/200
	0.74 (0.70)	0.73 (0.75)	0.71	0.71	0.72
Ca	205	205/200	207.5/200	210/200	212.5/200
	0.40 (0.40)	0.49 (0.47)	0.39	0.37	0.35
Cd	202.5	510/502.5	507.5/502.5	507.5/505	525/502.5
	0.08 (0.07)	0.12 (0.09)	0.06	0.05	0.12
Cr	202.5	685/642.5	685/630	685/607.5	685/605
	0 (0)	0.05 (0.03)	0.05	0.05	0.05
Cu	352.5	200/217.5	200/220	200/215	200/222.5
	0.63 (0.65)	0.58 (0.59)	0.57	0.56	0.57
Fe	527.5	707.5/680	707.5/692.5	657.5/677.5	485/482.5
	0.13 (0.12)	0.07 (0.09)	0.11	0.1	0.13
K	207.5	337.5/340	330/340	330/342.5	452.5/457.5
	0.19 (0.14)	0.45 (0.50)	0.35	0.34	0.33
Mg	205	205/200	207.5/200	210/200	202.5/200
	0.46 (0.44)	0.57 (0.53)	0.5	0.47	0.47
N	207.5	210/200	207.5/200	212.5/200	200/210
	0.43 (0.37)	0.44 (0.38)	0.27	0.26	0.27
Na	207.5	205/200	207.5/200	210/200	212.5/200
	0.54 (0.51)	0.59 (0.51)	0.59	0.56	0.54
Pb	342.5	200/220	200/222.5	200/217.5	200/225
	0.32 (0.31)	0.25 (0.12)	0.28	0.28	0.28
Rb	202.5	387.5/325	387.5/330	387.5/327.5	387.5/320
	0.34 (0.34)	0.60 (0.62)	0.48	0.48	0.48
Si	210	200/210	200/207.5	200/210	200/212.5
	0.56 (0.60)	0.76 (0.72)	0.68	0.66	0.64
Sr	202.5	205/200	207.5/200	210/200	202.5/200
	0.46 (0.49)	0.53 (0.64)	0.42	0.40	0.37
V	720	460/480	655/210	655/212.5	235/245
	0.03 (0.05)	0.18 (0.13)	0.09	0.09	0.17
Y	202.5	625/605	600/580	652.5/605	727.5/675
	0.06 (0.08)	0.05 (0.05)	0.06	0.05	0.04
Zn	357.5	405/400	520/505	405/402.5	442.5/435
	0.22 (0.20)	0.26 (0.21)	0.25	0.25	0.25

TABLE 1 Best single wavelength (WL) and WL ratios correlated to the laboratory-measured concentration of each solute. Below each predictor (best single WL or best WL ratio) we show the NSE in calibration. The NSE in validation for the best single WL and best ratio methods inference is in parentheses

Abbreviation: NSE, Nash–Sutcliffe efficiency.

Causse, 2017). When EC was included as a predictor in the PLSR method (hereafter referred to as PLSR+EC method), the results improved for all solutes (mean improvement: NSEc = + 0.06; NSEv = +0.06; nRMSEc = − 1.36%, nRMSEv = −1.35%; Figure 3a–d). PLSR+EC also allowed calibration of two additional solutes, Ca and K, with an acceptable fit (NSEc: Ca = 0.62; K = 0.66,

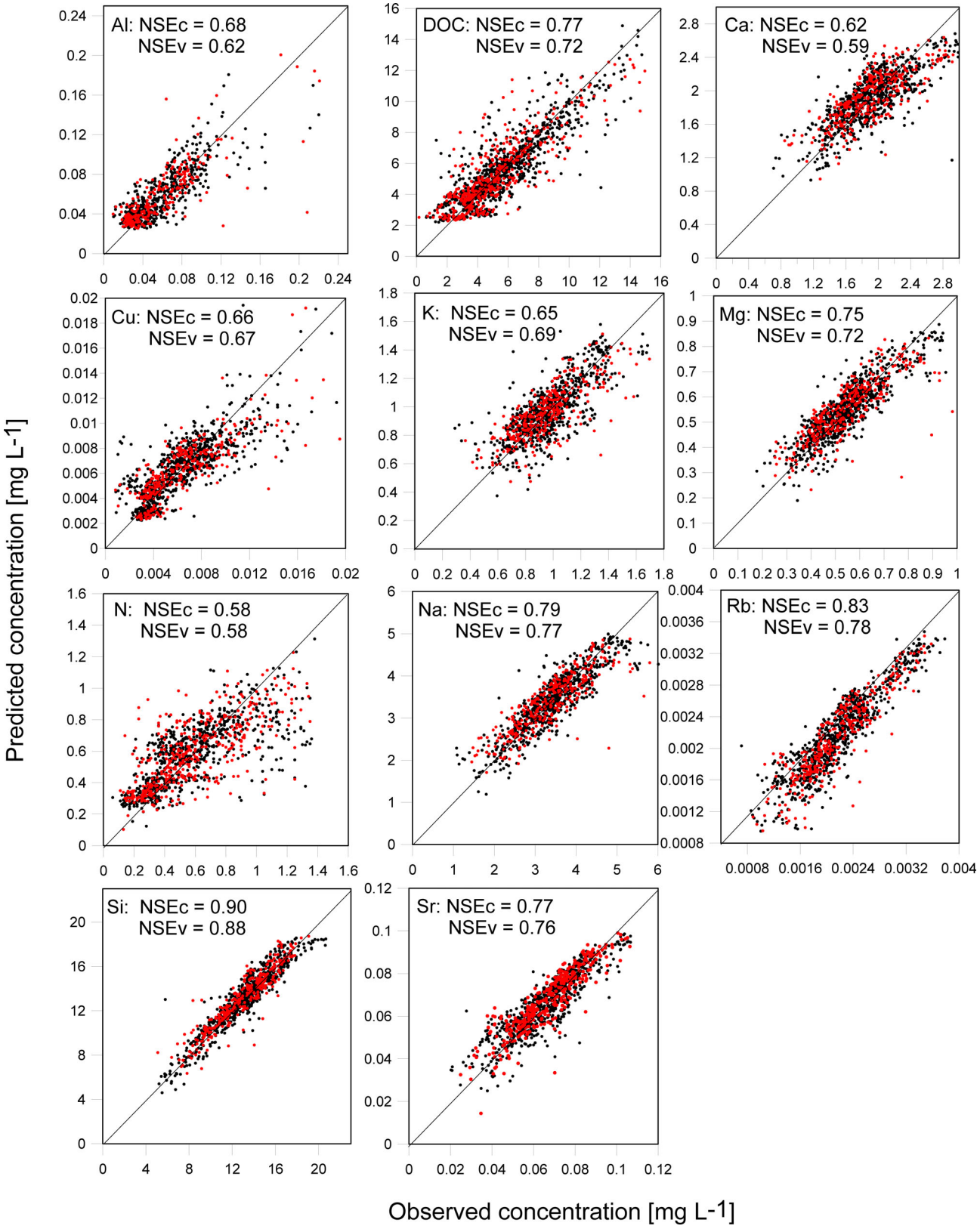

FIGURE 4 *X-Y* scatter plots showing the 1:1 relation between the inferred (predicted) and laboratory-determined (observed) concentrations of the 11 solutes that were calibrated using the PLSR + EC method with the UV–visible spectrometer. The black points represent the dataset used in calibration and the red points the dataset used for validation

NSEv: Ca = 0.60; K = 0.70). These solutes could only be calibrated with the PLSR when EC was included. Solutes that could be calibrated using PLSR + EC were: Al, DOC, Ca, Cu, K, Mg, N, Na, Rb, Si, and Sr. For these solutes, scatter plots of laboratory-determined (observed) versus inferred (predicted) concentrations showed a consistent correlation along the 1:1 line (Figure 4). The AIC for the indirect methods in both calibration and validation also supported the findings of the goodness of fit and error metrics, showing that the indirect methods outperformed the direct methods to infer solute concentrations (most parsimonious model: PLSR + EC; Figure 3).

To investigate the influence of sampling frequency on solute concentration inference, we evaluated the NSE and nRMSE results during validation (Figure 5). For most solutes, results did not change notably compared to the use of samples collected every 4 h until a frequency of two samples per week was used (NSE > 0.50; nRMSE <33%). The only exceptions were Al, Ca, and DOC for which one sample per day was necessary and N for which two samples per day were necessary to reach at least an acceptable fit (i.e., NSE > 0.5).

3.2.3 | Relationship between laboratory-determined DOC and solutes concentrations

We found similarities among the optimal WL values and optimal WL ranges and all the observed laboratory-determined solute concentrations (including those that could not be calibrated). The solutes that presented an acceptable to good fit and can be inferred by spectrometry (NSE > 0.50; Figure 3) were significantly correlated with DOC (Spearman correlation >|0.50|; *p*-value <0.001; Figure 6). Furthermore, most solutes that could not be inferred showed very little correlation with DOC (Spearman correlation <|0.30|). Only Ba (Spearman correlation = −0.49; *p*-value <0.001) and *Y* (Spearman correlation = 0.51; *p*-value <0.001) were acceptably correlated with DOC (Figure 6, Appendix S1), but could not be calibrated using the PLSR +EC method (NSEc: Ba = 0.27; Y = 0.06, NSEv: Ba = 0.21; *Y* = 0.08).

In addition, we compared the laboratory-determined DOC concentrations against the laboratory-determined concentrations of each successfully inferred solute, in order to evaluate the value of the absorbance in the spectra for calibration compared to only inferring concentrations using DOC as predictor. The mean of the difference in terms of NSE, nRMSE, and AIC metrics comparing the results obtained through the regression analysis and the PLSR + EC method for the solutes that could be inferred using spectrometry resulted in 0.39 for NSE, −9.74% for nRMSE, and −1598.80 for AIC. Even the best regressions were outperformed in terms of NSE, nRMSE, and AIC by the PLSR+EC for all calibrated solutes (see Appendix S4).

4 | DISCUSSION

We compared laboratory-determined solute concentrations with the WL absorbance of a UV–visible sensor and found that the highest R^2 values coincided with the lowest WLs. This outcome for DOC is in accordance with other studies that found a high correlation between the absorbance at low WLs (in UV range) and the concentrations of organic compounds, particularly for DOC (Cook et al., 2017; Peacock et al., 2014; Tunaley et al., 2017; Wei et al., 2018). Our results showed peak R^2 values for metals at low WLs (similar or even lower than those for DOC, i.e., >230 nm). These findings differ from studies that analysed metals using complexometry or reagents in which the peak depended on the reagent used (Aguerssif et al., 2008; Zhou et al., 2019).

Comparing the direct calibration methods (Table 1), most values for the best ratios (those that can be calibrated with the method) were in the UV range. These findings and those for DOC are consistent with those of Peacock et al. (2014) who found the best WL to calibrate DOC was 230 nm. The latter could be attributed to similar or even higher DOC concentrations in the stream (0–40 mg/L). Furthermore, our findings present different optimal WLs and ratios for metals (between 200 and 222.5 nm; Table 1), except for K and Rb (between

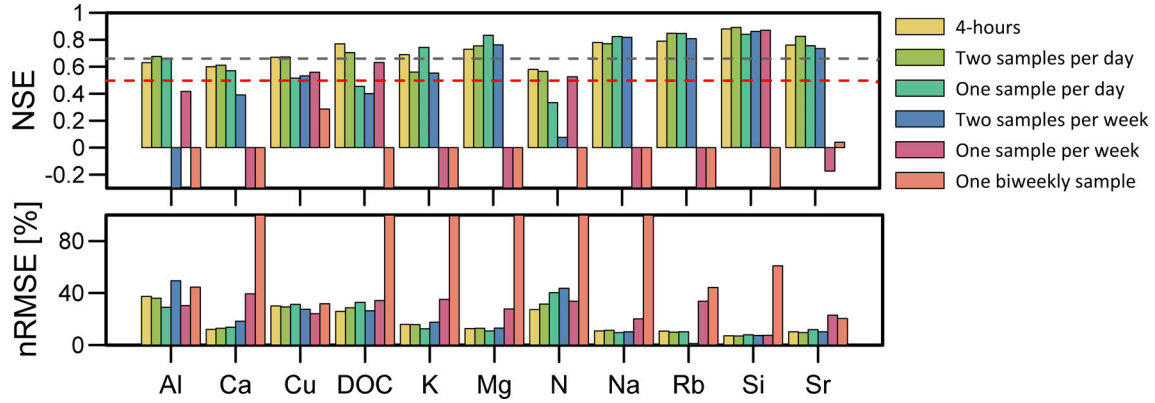

FIGURE 5 Nash–Sutcliffe efficiency (NSE) (a) and the normalized root mean square error (%) (nRMSE) (b) datasets in the validation for different sampling resolution data sets. In (a) the red line represents an acceptable fit (i.e., 0.50 < NSE < 0.65) and the grey line represents a good fit (NSE ≥ 0.65)

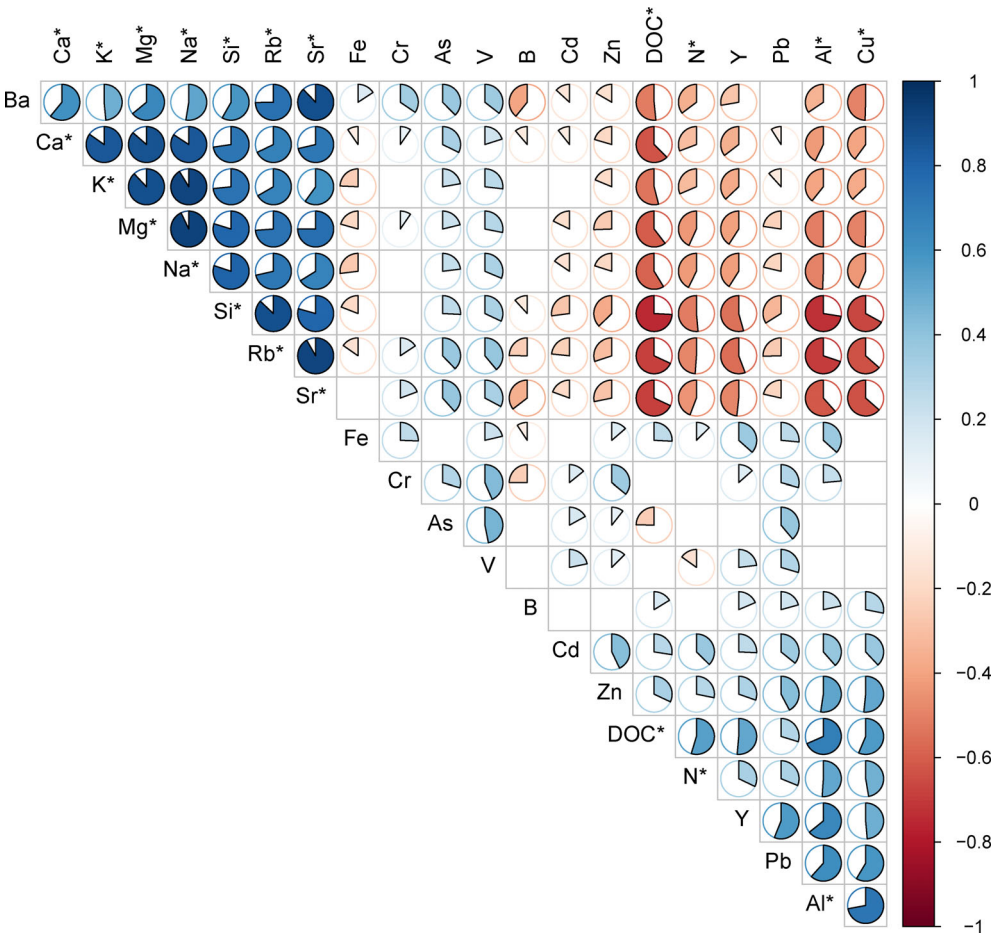

FIGURE 6 Spearman correlation analysis between pairs of laboratory-determined (observed) solute concentrations. Circles are shown only when the correlation is statistically significant (p-value <0.05). Colour intensity and the area of each circle represent the absolute value of the corresponding correlation coefficient. Positive correlations are displayed in blue and negative correlations in red. (*) Satisfactorily calibrated solutes

330 and 390 nm) compared to those studied elsewhere (Aguerssif et al., 2008; Zhou et al., 2019). The latter is related to the fact that metals were previously analysed using complexometry and the WL for the calibration depends on the reagent used (Aguerssif et al., 2008; Garcia Rodriguez et al., 1998; Kulkarni et al., 2015). The direct calibration results for various metals at the same time have not been previously described for in situ analysis in tropical regions. Similar NSE values and errors to the single WL and the ratio methods were found for the area under the spectra method for both nutrients and metals. This result is similar to previous studies that found the area under the spectra was a good predictor of organic matter compared to single WL or ratio methods (Peacock et al., 2014; Wang & Hsieh, 2001).

Indirect multivariate methods showed an important improvement in calibration compared to the direct methods. These results support the findings from previous studies that recommended the use of the PLSR method to infer solutes concentrations (Aguerssif et al., 2008; Carré et al., 2017; Chen et al., 2021; Garcia Rodriguez et al., 1998; Khatri et al., 2020; Xu et al., 2014). This is likely because using the whole spectrum allows accounting for interferences (e.g., suspended solids) in light absorbance of stream water (Shi et al., 2020, 2021). Further, the use of EC as a predictor for solute concentrations can be beneficial as it accounts for the conductive nature of the calibrated solutes (Thomas & Burgess, 2007). This physicochemical parameter can be monitored in situ at a high temporal frequency and relatively low cost (Mosquera et al., 2018). It is worth highlighting that this technique worked even to a low sampling resolution (daily or twice weekly) which supports the use of spectrometry to calibrate and subsequently infer important solute concentrations to achieve improved understanding of catchment functioning (e.g., Correa et al., 2016, 2019).

Finally, we analysed the possibility of calibrating models to infer metal solutes and found high correlations between DOC and other solutes, indicating that DOC can be a proxy for metals. These similarities could be attributed to a complexation phenomenon between metals and carbon which likely results in a specific water coloration (Nieder & Benbi, 2008), increasing or decreasing the absorbance under different wavelengths and helping to determine their concentrations. In the Páramos, a complexation phenomenon between carbon and metals (organocomplexation) was reported earlier (Buytaert et al., 2006; Correa et al., 2019), and thus could be the factor that allows the determination of some metals. Complexation was also reported to allow identification of metals by spectrometry under controlled laboratory conditions (Brown et al., 2005; Kulkarni et al., 2015; Thomas & Burgess, 2007). Interestingly, some solutes (Sr, Rb, Si, Na, Mg, K, and Ca) that were not positively correlated (Appendix S2–S4) or collocated with DOC, could be adequately calibrated. Here, the influence of water from mineral soil horizons increases (Correa et al., 2017, 2019), resulting in decreased DOC concentrations from

shallower horizons. Therefore, these solutes can be inferred. Despite this relationship between DOC and water source, inferring solutes concentrations directly from DOC had a lower performance than the PLSR + EC method using the entire UV–visible spectrum (see Appendix S4). Without the information provided by the complete UV–visible spectra, the calibration of such solutes that are not related positively to DOC would not have been possible (see Appendix S4). In addition, it indicated that the DOC by itself would not be a predictor of the rest of the solutes that were possible to infer. Although the extrapolation of our correlation analysis to systems with low DOC concentrations but relatively high trace metal concentrations may not be possible, our findings for the Páramos with organic-rich soils and stream waters are likely to be representative of other peatland catchments in temperate and sub-alpine regions of, for example, the USA, UK and Germany (Aitkenhead & McDowell, 2000; Birkel et al., 2017; Dawson et al., 2008), respectively. Our findings contribute to advancing the quantification of solutes used as tracers to improve understanding of water quality dynamics at a high temporal frequency in ecosystems dominated by organic rich soils.

5 | CONCLUSIONS

We assessed the use of in stream UV–visible absorbance spectrometry sensors to infer solutes (e.g., nutrients and metals) concentrations at high temporal resolution using calibration methods of different mathematical complexity. We found that in addition to DOC and nitrogen some metals can be inferred in situ at our study catchment. The latter can be likely attributed to a complexation phenomenon between fluvial carbon and metals. Such complexation provides novel research avenues to infer solutes in situ, particularly in organic-rich peatland dominated ecosystems. The solutes whose concentrations could be satisfactorily inferred (i.e., Al, DOC, Ca, Cu, K, Mg, N, Na, Rb, Si, and Sr) yielded the highest performance when calibrated using multivariate statistical methods, particularly the PLSR + EC method. It is important to highlight that using EC permitted calibration of Ca and K, solutes that could not have been calibrated without including EC as a predictor in the PLSR method. Evaluating different sampling resolutions for calibration showed that similar results could be obtained when using two samples per week for most solutes in comparison to the original 4-hourly frequency of the original dataset presented in the study, with only a few exceptions, namely Al, Ca, DOC, and N, which require a finer sample frequency (i.e., daily sampling). Lower sampling frequencies for calibration could broaden the use of UV–visible sensors to sites with similar conditions (organic-rich ecosystems). Our findings indicate that calibrating in-stream UV–visible spectrometers to infer solute concentrations can be valuable for hydrology, biogeochemistry, and water quality field studies as it could potentially allow monitoring concentrations of several solutes, including metals, at high temporal frequency (e.g., 5-min resolution). A further study could assess different sites or ecosystem conditions (e.g., low organic load), which represent a limitation of this study, in order to generalize the calibration methods to infer solutes in stream waters.

ACKNOWLEDGMENTS

This manuscript is an outcome of the University of Cuenca's Doctoral Program in Water Resources and Master Program in Ecohydrology. We thank INV Metals S.A. staff for their assistance in the logistics during field work at the Zhurucay Experimental Observatory. We also thank Comuna Chumblín Sombrederas (Azuay, San Fernando) for their logistical support and access to the community land reserve. Thanks are also due to the researchers and students of the Department of Water Resources and Environmental Sciences at the University of Cuenca who supported the collection of samples. Giovanny M. Mosquera is supported by a Postdoctoral Fellowship from Universidad San Francisco de Quito. We acknowledge the constructive criticism of four anonymous reviewers and the editor Jim Buttle, who helped improve the quality of our paper.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request.

ORCID

Juan Pesántez https://orcid.org/0000-0003-4640-4556
Christian Birkel https://orcid.org/0000-0002-6792-852X
Giovanny M. Mosquera https://orcid.org/0000-0002-4764-4685
William H. McDowell https://orcid.org/0000-0002-8739-9047
Patricio Crespo https://orcid.org/0000-0001-5126-0687

REFERENCES

Aguerssif, N., Benamor, M., Kachbi, M., & Draa, M. T. (2008). Simultaneous determination of Fe(III) and Al(III) by first-derivative spectrophotometry and partial least-squares (PLS-2) method – Application to post-haemodialysis fluids. *Journal of Trace Elements in Medicine and Biology*, *22*(3), 175–182. https://doi.org/10.1016/j.jtemb.2007.12.004

Aitkenhead, J. A., & McDowell, W. H. (2000). Soil C: N ratio as a predictor of annual riverine DOC flux at local and global scales. *Global Biogeochemical Cycles*, *14*(1), 127–138. https://doi.org/10.1029/1999GB900083

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Arsenault, R., Brissette, F., & Martel, J. L. (2018). The hazards of split-sample validation in hydrological model calibration. *Journal of Hydrology*, *566*, 346–362. https://doi.org/10.1016/j.jhydrol.2018.09.027

Avagyan, A., Runkle, B. R. K., & Kutzbach, L. (2014). Application of high-resolution spectral absorbance measurements to determine dissolved organic carbon concentration in remote areas. *Journal of Hydrology*, *517*, 435–446. https://doi.org/10.1016/j.jhydrol.2014.05.060

Beate, B. (1999). *Stratigraphy of the Quimsacocha volcanic Centre, Azuay province, southern Ecuador*. Quito.

Birkel, C., Broder, T., & Biester, H. (2017). Nonlinear and threshold-dominated runoff generation controls DOC export in a small peat catchment. *Journal of Geophysical Research: Biogeosciences*, *122*(3), 498–513. https://doi.org/10.1002/2016JG003621

Brown, A. E., Zhang, L., Mcmahon, T. A., Western, A. W., & Vertessy, R. A. (2005). A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, 310, 28–61. https://doi.org/10.1016/j.jhydrol.2004.12.010

Buytaert, W., Deckers, J., & Wyseure, G. (2006). Description and classification of nonallophanic andosols in south Ecuadorian alpine grasslands (páramo). *Geomorphology*, *73*(3–4), 207–221. https://doi.org/10.1016/j.geomorph.2005.06.012

Buytaert, W., Sevink, J., De Leeuw, B., & Deckers, J. (2005). Clay mineralogy of the soils in the south Ecuadorian páramo region. *Geoderma*, *127*(1–2), 114–129. https://doi.org/10.1016/j.geoderma.2004.11.021

Buytaert, W., Wyseure, G., De Bièvre, B., & Deckers, J. (2005). The effect of land-use changes on the hydrological behaviour of Histic Andosols in South Ecuador. *Hydrological Processes*, *19*(20), 3985–3997. https://doi.org/10.1002/hyp.5867

Carré, E., Pérot, J., Jauzein, V., Lin, L., & Lopez-Ferber, M. (2017). Estimation of water quality by UV/Vis spectrometry in the framework of treated wastewater reuse. *Water Science and Technology*, *76*(3), 633–641. https://doi.org/10.2166/wst.2017.096

Chen, X., Yin, G., Zhao, N., Gan, T., Yang, R., Xia, M., Feng, C., Chen, Y., & Huang, Y. (2021). Simultaneous determination of nitrate, chemical oxygen demand and turbidity in water based on UV–Vis absorption spectrometry combined with interval analysis. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *244*, 118827. https://doi.org/10.1016/j.saa.2020.118827

Cook, S., Peacock, M., Evans, C. D., Page, S. E., Whelan, M. J., Gauci, V., & Kho, L. K. (2017). Quantifying tropical peatland dissolved organic carbon (DOC) using UV-visible spectroscopy. *Water Research*, *115*, 229–235. https://doi.org/10.1016/j.watres.2017.02.059

Córdova, M., Carrillo-Rojas, G., Crespo, P., Wilcox, B., & Célleri, R. (2015). Evaluation of the Penman-Monteith (FAO 56 PM) method for calculating reference evapotranspiration using limited data. *Mountain Research and Development*, *35*(3), 230–239. https://doi.org/10.1659/MRD-JOURNAL-D-14-0024.1

Correa, A., Breuer, L., Crespo, P., Célleri, R., Feyen, J., Birkel, C., Silva, C., & Windhorst, D. (2019). Spatially distributed hydrochemical data with temporally high-resolution is needed to adequately assess the hydrological functioning of headwater catchments. *Science of the Total Environment*, *651*, 1613–1626. https://doi.org/10.1016/j.scitotenv.2018.09.189

Correa, A., Windhorst, D., Crespo, P., Célleri, R., Feyen, J., & Breuer, L. (2016). Continuous versus event-based sampling: How many samples are required for deriving general hydrological understanding on Ecuador's páramo region? *Hydrological Processes*, *30*(22), 4059–4073. https://doi.org/10.1002/hyp.10975

Correa, A., Windhorst, D., Tetzlaff, D., Crespo, P., Célleri, R., Feyen, J., & Breuer, L. (2017). Temporal dynamics in dominant runoff sources and flow paths in the Andean Páramo. *Water Resources Research*, *53*, 5998–6017. https://doi.org/10.1002/2016WR020187

Dawson, J. J. C., Soulsby, C., Tetzlaff, D., Hrachowitz, M., Dunn, S. M., & Malcolm, I. A. (2008). Influence of hydrology and seasonality on DOC exports from three contrasting upland catchments. *Biogeochemistry*, *90*(1), 93–113. https://doi.org/10.1007/s10533-008-9234-3

Ellison, S. L. R., & Williams, A. (2012). *Quantifying uncertainty in analytical measurement*. 3rd edition. Teddington, UK: Eurachem/CITAC. http://dx.doi.org/10.25607/OBP-952

Esquivel-Hernández, G., Mosquera, G. M., Sánchez-Murillo, R., Quesada-Román, A., Birkel, C., Crespo, P., Célleri, R., Windhorst, D., Breuer, L., & Boll, J. (2019). Moisture transport and seasonal variations in the stable isotopic composition of rainfall in Central American and Andean Páramo during El Niño conditions (2015–2016). *Hydrological Processes*, *33*(13), 1802–1817. https://doi.org/10.1002/hyp.13438

Etheridge, J. R., Birgand, F., Osborne, J. A., Osburn, C. L., Burchell, M. R., & Irving, J. (2014). Using in situ ultraviolet-visual spectroscopy to measure nitrogen, carbon, phosphorus, and suspended solids concentrations at a high frequency in a brackish tidal marsh. *Limnology and Oceanography: Methods*, *12*, 10–22. https://doi.org/10.4319/lom.2014.12.10

Falk, R. F., & Miller, N. B. (1992). *A primer for soft modeling*. University of Akron Press.

Fernandez-Palomino, C. A., Hattermann, F. F., Krysanova, V., Vega-Jácome, F., & Bronstert, A. (2020). Towards a more consistent eco-hydrological modelling through multi-objective calibration: A case study in the Andean Vilcanota River basin, Peru. *Hydrological Sciences Journal*, *66*, 59–74. https://doi.org/10.1080/02626667.2020.1846740

Garcia Rodriguez, A. M., Garcia De Torres, A., Cano Pavon, J. M., & Bosch, O. C. (1998). Simultaneous determination of iron, cobalt, nickel and copper by UV- visible spectrophotometry with multivariate calibration. *Talanta*, *47*(2), 463–470. https://doi.org/10.1016/S0039-9140(98)00157-X

Geladi, P., & Kowalski, B. R. (1986). Partial least-squares regression: A tutorial. *Analytica chimica acta*, *185*(9), 1–17. https://doi.org/10.1016/0003-2670(86)80028-9

Green, M. B., Pardo, L. H., Bailey, S. W., Campbell, J. L., McDowell, W. H., Bernhardt, E. S., & Rosi, E. J. (2020). Predicting high-frequency variation in stream solute concentrations with water quality sensors and machine learning. *Hydrological Processes*, *35*(1), e14000. https://doi.org/10.1002/hyp.14000

Haaland, D. M., & Thomas, E. V. (1988). Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry*, *60*(11), 1193–1202. https://doi.org/10.1021/ac00162a020

Heathwaite, A. L., & Bieroza, M. (2021). Fingerprinting hydrological and biogeochemical drivers of freshwater quality. *Hydrological Processes*, *35*(1), 1–17. https://doi.org/10.1002/hyp.13973

Huebsch, M., Grimmeisen, F., Zemann, M., Fenton, O., Richards, K. G., Jordan, P., Sawarieh, A., Blum, P., & Goldscheider, N. (2015). Technical Note: Field experiences using UV/VIS sensors for high-resolution monitoring of nitrate in groundwater. *Hydrology and Earth System Sciences*, *19*(4), 1589–1598. https://doi.org/10.5194/hess-19-1589-2015

Kaushal, S. S., Gold, A. J., Bernal, S., Johnson, T. A. N., Addy, K., Burgin, A., Burns, D. A., Coble, A. A., Hood, E., Lu, Y., Mayer, P., Minor, E. C., Schroth, A. W., Vidon, P., Wilson, H., Xenopoulos, M. A., Doody, T., Galella, J. G., Goodling, P., … Belt, K. T. (2018). Watershed 'chemical cocktails': Forming novel elemental combinations in Anthropocene fresh waters. *Biogeochemistry*, *141*, 281–305. https://doi.org/10.1007/s10533-018-0502-6

Khatri, P., Gupta, K. K., & Gupta, R. K. (2020). A review of partial least squares modeling (PLSM) for water quality analysis. *Modeling Earth Systems and Environment*, *0123456789*, 703–714. https://doi.org/10.1007/s40808-020-00995-4

Koehler, A. K., Murphy, K., Kiely, G., & Sottocornola, M. (2009). Seasonal variation of DOC concentration and annual loss of DOC from an Atlantic blanket bog in South Western Ireland. *Biogeochemistry*, *95*(2), 231–242. https://doi.org/10.1007/s10533-009-9333-9

Kulkarni, S., Dhokpande, S., & Kaware, J. (2015). A review on spectrophotometric determination of heavy metals with emphasis on cadmium and nickel determination by U.V. spectrophotometry. *International Journal of Advanced Engineering Research and Science*, *2*(9), 35–38 Retrieved from https://www.semanticscholar.org/paper/A-Review-on-Spectrophotometric-Determination-of-on-Kulkarni-Dhokpande/e008e44994ef31fec28922f5a6cc08a746f9b91a

Langergraber, G., Fleischmann, N., & Hofstädter, F. (2003). A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. *Water Science and Technology*, *47*(2), 63–71. https://doi.org/10.2166/wst.2003.0086

Lin, S., Hsieh, I. J., Huang, K. M., & Wang, C. H. (2002). Influence of the Yangtze River and grain size on the spatial variations of heavy metals

and organic carbon in the East China Sea continental shelf sediments. *Chemical Geology*, *182*(2–4), 377–394. https://doi.org/10.1016/S0009-2541(01)00331-X

Lloyd, C. E. M., Freer, J. E., Johnes, P. J., & Collins, A. L. (2016). Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments. *Science of the Total Environment*, *543*, 388–404. https://doi.org/10.1016/j.scitotenv.2015.11.028

Lopez-Kleine, L., & Torres, A. (2014). UV-vis in situ spectrometry data mining through linear and non linear analysis methods. *Dyna*, *81*(185), 182. https://doi.org/10.15446/dyna.v81n185.37718

McDowell, W. H., & Asbury, C. E. (1994). Export of carbon, nitrogen, and major ions from three tropical montane watersheds. *Limnology and Oceanography*, *39*(1), 111–125. https://doi.org/10.4319/lo.1994.39.1.0111

McGuire, K. J., & McDonnell, J. J. (2010). Hydrological connectivity of hillslopes and streams: Characteristic time scales and nonlinearities. *Water Resources Research*, *46*, 1–17. https://doi.org/10.1029/2010WR009341

McGuire, K. J., & McDonnell, J. J. (2015). Tracer advances in catchment hydrology. *Hydrological Processes*, *29*(25), 5135–5138. https://doi.org/10.1002/hyp.10740

McIntyre, J. F., Foley, R. T., & Brown, B. F. (1982). Ultraviolet spectra of aluminum salt solutions. *Inorganic Chemistry*, *21*(3), 1167–1172. https://doi.org/10.1021/ic00133a056

Mevik, B.-H., Wehrens, R., & Liland, K. H. (2020). Pls: Partial least squares and principal component regression Retrieved from https://cran.r-project.org/package=pls

Moore, R. B. (2004). Introduction to salt dilution gauging for streamflow measurement part 2: Constant-rate injection. *Streamline Watershed Management Blulletin*, *8*(1), 11–15.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers*, *50*(3), 885–900. https://doi.org/10.13031/2013.23153

Mosquera, G. M., Lazo, P. X., Célleri, R., Wilcox, B. P., & Crespo, P. (2015). Runoff from tropical alpine grasslands increases with areal extent of wetlands. *Catena*, *125*, 120–128. https://doi.org/10.1016/j.catena.2014.10.010

Mosquera, G. M., Segura, C., & Crespo, P. (2018). Flow partitioning modelling using high-resolution isotopic and electrical conductivity data. *Water (Switzerland)*, *10*(7), 1–23. https://doi.org/10.3390/w10070904

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I-a discussion of principles*. *Journal of Hydrology*, *10*, 282–290 Retrieved from https://hydrology.agu.org/wp-content/uploads/sites/19/2016/04/NashSutcillfe1.pdf

Nieder, R., & Benbi, D. K. (2008). *Carbon and nitrogen in the terrestrial environment*. Springer. https://doi.org/10.1017/CBO9781107415324.004

Nimick, D. A., Gammons, C. H., Cleasby, T. E., Madison, J. P., Skaar, D., & Brick, C. M. (2003). Diel cycles in dissolved metal concentrations in streams: Occurrence and possible causes. *Water Resources Research*, *39*(9), 2-1–2-17. https://doi.org/10.1029/2002WR001571

Olivieri, A. C. (2018). *Introduction to multivariate calibration*. Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-319-97097-4

Peacock, M., Evans, C. D., Fenner, N., Freeman, C., Gough, R., Jones, T. G., & Lebron, I. (2014). UV-visible absorbance spectroscopy as a proxy for peatland dissolved organic carbon (DOC) quantity and quality: Considerations on wavelength and absorbance degradation. *Environmental science*, *16*, 1445–1461. https://doi.org/10.1039/c4em00108g

Pesántez, J., Mosquera, G. M., Crespo, P., Breuer, L., & Windhorst, D. (2018). Effect of land cover and hydro - meteorological controls on soil water DOC concentrations in a high - elevation tropical environment. *Hydrological Processes*, *32*(17), 2624–2635. https://doi.org/10.1002/hyp.13224

Quichimbo, P., Tenorio, G., Borja, P., Cárdenas, I., Crespo, P., & Célleri, R. (2012). Efectos sobre las propiedades físicas y químicas de los suelos por el cambio de la cobertura vegetal y uso del suelo: Páramo de Quimsacocha al sur del Ecuador. *Suelos Ecuatoriales*, *42*(2), 138–153.

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org

Rights, J. D., & Sterba, S. K. (2019). Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. *Psychological Methods*, *24*(3), 309–338. https://doi.org/10.1037/met0000184

Ruhala, S. S., & Zarnetske, J. P. (2017). Using in-situ optical sensors to study dissolved organic carbon dynamics of streams and watersheds: A review. *Science of the Total Environment*, *575*, 713–723. https://doi.org/10.1016/j.scitotenv.2016.09.113

Schmutz, S., & Sendzimir, J. (2018). In S. Schmutz & J. Sendzimir (Eds.), Challenges in Riverine Ecosystem Management. *Marine ecosystem management*. (pp. 1–17).

Shi, Z., Chow, C. W. K., Fabris, R., Liu, J., & Jin, B. (2020). Alternative particle compensation techniques for online water quality monitoring using UV–vis spectrophotometer. *Chemometrics and Intelligent Laboratory Systems*, *204*, 104074. https://doi.org/10.1016/j.chemolab.2020.104074

Shi, Z., Chow, C. W. K., Fabris, R., Zheng, T., Liu, J., & Jin, B. (2021). Evaluation of the impact of suspended particles on the UV absorbance at 254 nm (UV254) measurements using a submersible UV-Vis spectrophotometer. *Environmental Science and Pollution Research*, *28*(10), 12576–12586. https://doi.org/10.1007/s11356-020-11178-0

Singh, M., & Sarkar, A. (2018). Comparative study of the PLSR and PCR methods in laser-induced breakdown spectroscopic analysis. *Journal of Applied Spectroscopy*, *85*(5), 962–970. https://doi.org/10.1007/s10812-018-0746-x

Snyder, L., Potter, J. D., & McDowell, W. H. (2018). An evaluation of nitrate, fDOM, and turbidity sensors in New Hampshire streams. *Water Resources Research*, *54*(3), 2466–2479. https://doi.org/10.1002/2017WR020678

Soulsby, C., Birkel, C., & Tetzlaff, D. (2016). Modelling storage-driven connectivity between landscapes and riverscapes: Towards a simple framework for long-term ecohydrological assessment. *Hydrological Processes*, *30*(14), 2482–2497. https://doi.org/10.1002/hyp.10862

Tantithamthavorn, C., McIntosh, S., Hassan, A. E., & Matsumoto, K. (2017). An empirical comparison of model validation techniques for defect prediction models. *IEEE Transactions on Software Engineering*, *43*(1), 1–18. https://doi.org/10.1109/TSE.2016.2584050

Thomas, O., & Burgess, C. (2007). *UV-visible spectrophtometry of water and wastewater*. Elsevier.

Thomas, O., & Causse, J. (2017). *From spectra to qualitative and quantitative results*. Elsevier B.V. https://doi.org/10.1016/B978-0-444-63897-7.00002-0

Thompson, M. (2004). The amazing Horwitz function. *AMC Technical Brief*, *17*(17), 1–2. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+amazing+Horwitz+function#0

Tunaley, C., Tetzlaff, D., & Soulsby, C. (2017). Scaling effects of riparian peatlands on stable isotopes in runoff and DOC. *Journal of Hydrology*, *549*, 220–235. https://doi.org/10.1016/j.jhydrol.2017.03.056

US Department of the Interior Bureau of Reclamation. (2001). *Water measurement manual*. US Government Printing Office Washington, DC.

van den Broeke, J. (2007). On-line and in-situ UV/Vis spectroscopy: Real time multi parameter measurements with a single instrument. *AWE International*, *9*, 55–59.

von Freyberg, J., Studer, B., & Kirchner, J. W. (2017). A lab in the field: High-frequency analysis of water quality and stable isotopes in stream water and precipitation. *Hydrology and Earth System Sciences*, *21*(3), 1721–1739. https://doi.org/10.5194/hess-21-1721-2017

Vorobyev, S. N., Kolesnichenko, L. G., Shirokova, L. S., & Karlsson, J. (2019). Biogeochemistry of dissolved carbon, major, and trace elements during spring flood periods on the Ob river. *Hydrological Processes*, *33*(11), 1579–1594. https://doi.org/10.1002/hyp.13424

Wang, G., & Hsieh, S. (2001). Monitoring natural organic matter in water with scanning spectrophotometer. *Environment International*, *26*, 205–212. https://doi.org/10.1016/S0160-4120(00)00107-0

Waterloo, M. J., Oliveira, S. M., Drucker, D. P., Nobre, A. D., Cuartas, L. A., Hodnett, M. G., Langedijk, I., Jans, W. W. P., Tomasella, J., de Araújo, A. C., Pimentel, T. P., & Múnera Estrada, J. C. (2006). Export of organic carbon in run-off from an Amazonian rainforest Blackwater catchment. *Hydrological Processes*, *20*, 2581–2597. https://doi.org/10.1002/hyp.6217Export

Wei, H., Yu, H., Pan, H., Gao, H. (2018). *Application of UV-visible absorption spectroscopy combined with two-dimensional correlation for insight into DOM fractions from native halophyte soils in a larger estuarine delta* (pp. 14197–14205)

Wei, T., & Simko, V. (2017). R package 'corrplot': Visualization of a correlation matrix. Stat Retrieved from https://github.com/taiyun/corrplot

Wickham, H., François, R., Henry, L., & Müller, K. (2020). Dplyr: A grammar of data manipulation Retrieved from https://cran.r-project.org/package=dplyr

Wymore, A. S., Potter, J., Rodríguez-Cardona, B., & McDowell, W. H. (2018). Using in-situ optical sensors to understand the biogeochemistry of dissolved organic matter across a stream network. *Water Resources Research*, *54*(4), 2949–2958. https://doi.org/10.1002/2017WR022168

Xu, D., Fan, W., Lv, H., Liang, Y., Shan, Y., Li, G., Yang, Z., & Yu, L. (2014). Simultaneous determination of traces amounts of cadmium, zinc, and cobalt based on UV-vis spectrometry combined with wavelength selection and partial least squares regression. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *123*, 430–435. https://doi.org/10.1016/j.saa.2013.12.086

Zhou, F., Li, C., Zhu, H., & Li, Y. (2019). A novel method for simultaneous determination of zinc, nickel, cobalt and copper based on UV–vis spectrometry. *Optik*, *182*, 58–64. https://doi.org/10.1016/j.ijleo.2018.12.159

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Pesántez, J., Birkel, C., Mosquera, G. M., Peña, P., Arízaga-Idrovo, V., Mora, E., McDowell, W. H., & Crespo, P. (2021). High-frequency multi-solute calibration using an in situ UV–visible sensor. *Hydrological Processes*, *35*(9), e14357. https://doi.org/10.1002/hyp.14357