Use of near-infrared spectroscopy on predicting wastewater constituents to facilitate the operation of a membrane bioreactor

Kim, Sang Yeob; Ćurko, Josip; Gajdoš Kljusurić, Jasenka; Matošić, Marin; Crnek, Vlado; López-Vázquez, Carlos M.; Garcia, Hector A.; Brdjanović, Damir; Valinger, Davor

DOI
10.1016/j.chemosphere.2021.129899

Publication date
2021

Document Version
Final published version

Published in
Chemosphere

Citation (APA)
Kim, S. Y., Ćurko, J., Gajdoš Kljusurić, J., Matošić, M., Crnek, V., López-Vázquez, C. M., Garcia, H. A., Brdjanović, D., & Valinger, D. (2021). Use of near-infrared spectroscopy on predicting wastewater constituents to facilitate the operation of a membrane bioreactor. *Chemosphere*, *272*, Article 129899. https://doi.org/10.1016/j.chemosphere.2021.129899

Chemosphere 272 (2021) 129899

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Use of near-infrared spectroscopy on predicting wastewater constituents to facilitate the operation of a membrane bioreactor

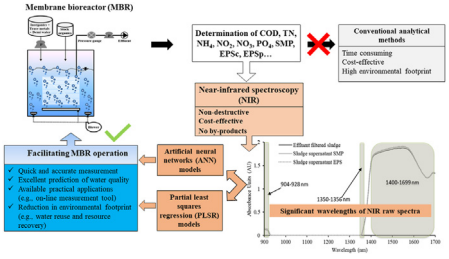
Sang Yeob Kim [a, b], Josip Ćurko [c, *], Jasenka Gajdoš Kljusurić [c], Marin Matošić [c], Vlado Crnek [c], Carlos M. López-Vázquez [a], Hector A. Garcia [a], Damir Brdjanović [a, b], Davor Valinger [c]

[a] Department of Water Supply, Sanitation and Environmental Engineering, IHE Delft Institute for Water Education, Westvest 7, 2611AX, Delft, the Netherlands
[b] Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, the Netherlands
[c] Faculty of Food Technology and Biotechnology, University of Zagreb, Pierottijeva 6, 10000, Zagreb, Croatia

HIGHLIGHTS

- Conventional analytical measurements in wastewater are costly and labour intensive.
- NIR spectroscopy results were modelled to predict wastewater constituents.
- Excellent prediction of wastewater constituents was obtained with both models.
- Membrane foulants (EPS & SMP) were successfully determined by NIR spectroscopy.
- Developed procedure requires translucent samples and an expert interpreter.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 14 September 2020
Received in revised form
15 January 2021
Accepted 2 February 2021
Available online 8 February 2021

Handling Editor: Derek Muir

Keywords:
Near-infrared spectroscopy
Wastewater analyses
Wastewater treatment
Membrane bioreactor
Soluble microbial products
Extracellular polymeric substances

ABSTRACT

The use of near-infrared (NIR) spectroscopy in wastewater treatment has continuously expanded. As an alternative to conventional analytical methods for monitoring constituents in wastewater treatment processes, the use of NIR spectroscopy is considered to be cost-effective and less time-consuming. NIR spectroscopy does not distort the measured sample in any way as no prior treatment is required, making it a waste-free technique. On the negative side, one has to be very well versed with chemometric techniques to interpret the results. In this study, filtered and centrifuged wastewater and sludge samples from a lab-scale membrane bioreactor (MBR) were analysed. Two analytical methods (conventional and NIR spectroscopy) were used to determine and compare major wastewater constituents. Particular attention was paid to soluble microbial products (SMPs) and extracellular polymeric substances (EPSs) known to promote membrane fouling. The parameters measured by NIR spectroscopy were analysed and processed with partial least squares regression (PLSR) and artificial neural networks (ANN) models to assess whether the evaluated wastewater constituents can be monitored by NIR spectroscopy. Very good results were obtained with PLSR models, except for the determination of SMP, making the model qualitative rather than quantitative for their monitoring. ANN showed better performance in terms of correlation of NIR spectra with all measured parameters, resulting in correlation coefficients higher than 0.97 for training, testing, and validation in most cases. Based on the results of this research, the

* Corresponding author.
E-mail address: josip.curko@pbf.unizg.hr (J. Ćurko).

https://doi.org/10.1016/j.chemosphere.2021.129899

combination of NIR spectra and chemometric modelling offers advantages over conventional analytical methods.

1. Introduction

Biological wastewater treatment systems based on the activated sludge concept have become widely used to treat wastewater containing biodegradable pollutants. Since continuous monitoring is required to maintain a stable wastewater treatment process, many analytical methods (e.g., ion chromatography, gas chromatography, titrimetric or colourimetric absorption photometry, and others) have been used over the years to identify and quantify various chemical compounds commonly found in raw and treated wastewater. However, the listed methods are quite time-consuming; they require specialised laboratory equipment and involve complex sample preparation. Although the analysis time for the determination of the main chemical parameters in raw and treated wastewater (e.g., chemical oxygen demand (COD) and constituents (total nitrogen (TN)), NH_4-N, NO_2-N, NO_3-N, total phosphorus (TP), PO_4-P, etc.) has been significantly reduced by the development of test kits (e.g., Hach Lange Cuvette Tests), such application is much more costly than the traditional methods (e.g., titration method), and thus increases the operating costs of a wastewater treatment plant (WWTP) (Li et al., 2018).

Membrane bioreactor (MBR) technology, which integrates membrane filtration with the activated sludge process, has become an established technology for wastewater treatment. MBR systems produce a clarified and disinfected effluent with a very low biochemical oxygen demand (BOD) and free of suspended solids. MBR systems present several other advantages including handling higher mixed liquor suspended solids (MLSS) concentrations than conventional activated sludge (CAS) systems, operating at high solid retention times (SRTs), producing low amounts of sludge, handling high organic and shock loads, and satisfying the most challenging wastewater discharge standard (Henze et al., 2008; Kim et al., 2019). One of the major disadvantages of MBR technology is membrane fouling and associated activities for the cleaning and replacement of membranes (Bagheri et al., 2019). Soluble microbial products (SMP) and extracellular polymeric substances (EPS) are the major causes of membrane fouling. The SMP consists of soluble organic compounds generated as a consequence of the bacterial metabolism and bacterial autolysis. These compounds can also be present in influent wastewater (Soh et al., 2020). The EPS are highly hydrated and usually characterised by a mixture of high molecular weight polymers. Proteins, polysaccharides, humic acids, lipids, and nucleic acids, include the major constituents making up the EPS (Frølund et al., 1996). The EPS can be further classified into soluble EPS (sometimes also referred to as SMP (Laspidou and Rittmann, 2002)) and bound EPS. Besides, the EPS is subdivided into the carbohydrate fraction of the EPS (EPS_c) and the protein fraction of the EPS (EPS_p). The analytical determinations of SMP and EPS are usually time-consuming, and they require both analytical skills, as well as specific analytical laboratory equipment (Nielsen and Jahn, 1999). Moreover, there are different methods for the determination of EPS, which introduces additional difficulties in comparing results generated and reported by different analytical methods (Comte et al., 2006; D'Abzac et al., 2010). Therefore, the development of alternative methods for the determination of EPS and SMP is urgently needed and desirable.

In this context, the application of near-infrared (NIR) spectroscopy has enabled the development of new analytical methods to predict results, replacing conventional analytical methods. The basic principle of NIR spectroscopy is that a sample under investigation is irradiated with NIR radiation, usually between 750 and 2500 nm, and the reflected or transmitted radiation is recorded in the form of a spectrum (Prieto et al., 2017). Compared to traditional analytical methods for monitoring wastewater treatment compounds, NIR spectroscopy is faster and avoids the use of toxic or corrosive chemical reagents (Pereira et al., 2019). Therefore, the particular advantages of NIR spectroscopy include (i) the non-destructive and non-invasive characteristics of such analytical method; (ii) the cost-effectiveness of the determination; and (iii) the absence of residues to be disposed of after the analytical determinations have been carried out (Xie et al., 2016). NIR spectroscopy analyses can be affected by non-linearities in experimental conditions (e.g., temperature, agitation, aeration, dispersive light, etc.); however, all of these parameters can be well controlled during the performance of the analysis (Dias et al., 2008; Prieto et al., 2017).

The main limitation inherent in NIR determinations is that the results (expressed as absorbance) themselves are not directly comparable to those obtained by conventional analytical methods. That is, to interpret the results derived from NIR spectroscopy, chemometric techniques such as multiple linear regression analysis (MLRA), principal component analysis (PCA) and partial least squares regression (PLSR) must be performed, followed by chemometric modelling (Xu et al., 2008). The chemometric techniques and modelling have become an integral part of spectral data analysis along with the pre-processing of NIR spectra (Gajdoš Kljusurić et al., 2017), which aims to remove physical interferences in the spectra to improve subsequent multivariate regression, classification model, or exploratory analysis (Rinnan et al., 2009). NIR spectroscopy monitors the vibrations of molecules in the NIR region and generates a large amount of data. The chemometric techniques and models use various tools to find the best possible connection in such large datasets relating the information contained in the spectra with the associated specific analytical parameters to be determined. If the observed correlation between the dataset and the parameters to be determined is acceptable, then a certain pattern of behaviour of a measured parameter could be established. The methodology requires calibrating and validating the model before predicting the results of the specific analytical parameters to be determined. There are different tools used in that modelling part, such as PCA, principal component regression, PLSR, and artificial neural networks (ANN), among others. PCA and PLSR are among the most commonly used calibration methods to determine linear relationships between the target parameter and the intensity of spectral absorption bands (Han et al., 2020). In recent years, a combination of PCA and ANN has shown great potential for high non-linearity applications.

The use of NIR spectroscopy as an alternative analytical method for the determination of wastewater constituents is relatively new. Takamura et al. (2002) was one of the first references to introduce this concept. The authors investigated a relationship between NIR spectra and the presence and concentration of organic matter in municipal wastewater expressed as total organic carbon (TOC), COD, and BOD. Suehara et al. (2007) also evaluated the use of NIR

spectroscopy to determine the concentrations of urea, solids, and oil of a wastewater from a biodiesel fuel production plant with satisfactory accuracy. In addition, other researchers have investigated the use of NIR spectroscopy to predict the concentration of relevant municipal and industrial wastewater treatment parameters such as total suspended solids (TSS), BOD, COD, TOC, TN, and TP (Dahlbacka et al., 2014; Inagaki et al., 2010; Pan et al., 2012; Pan and Chen, 2012; Pascoa et al., 2008; Yang et al., 2009). Although some studies have used NIR spectroscopy to determine wastewater treatment-related constituents, there is no information in the literature on the application of the NIR analytical method to determine wastewater parameters for MBR systems. MBR systems produce treated wastewater by filtering the wastewater through a micro/ultrafiltration membrane. Thus, MBR effluent is free of suspended solids and a suitable media for NIR spectral analyses, which need to be conducted in homogenous media. In addition, there is no literature on analytical determination with NIR spectroscopy of SMP and EPS constituents, known to cause membrane fouling in the MBR system.

The objective of this research was to evaluate the use of NIR spectroscopy followed by chemometric modelling (PLSR and ANN models) for the determination of MBR relevant wastewater treatment constituents including COD, TN, NH_4-N, NO_2-N, NO_3-N, PO_4-P, SMP, EPS_c, EPS_p. In addition, this study evaluated the capabilities of NIR spectroscopy to detect changes in the operational performance of the MBR system by interpreting the NIR spectra of the wastewater samples.

2. Materials and methods

2.1. MBR experiment

A lab-scale MBR treating synthetic wastewater was operated continuously for 80 days. The MBR system was first provided with fine bubble diffusers, and later with a pressurised aeration system (PAS) to supply dissolved oxygen. The lab-scale PAS consisted of a pressurised chamber in which the mixed liquor stream was recirculated from the bioreactor by a peristaltic pump getting exposed to pure oxygen at high-pressure conditions; the pure oxygen was supplied by an oxygen cylinder providing the desired pressure in the pressurised chamber. The pressurised chamber had a total volume of 2.75 L, and it was occupied with 0.55 L of mixed liquor at an operational pressure of 689 kPa. The MBR was inoculated with fresh mixed liquor taken from a full-scale municipal WWTP. The two experimental setups of the MBR with the two different aeration systems are shown in Fig. 1. The configuration of the MBR was the same for both systems; only the aeration systems provided to the MBR were different. The MBR was operated for 40 days with the fine bubble diffusers. Then, the diffusers were replaced by PAS and the MBR was operated for the subsequent 16 days. Then the MBR was again operated again with fine bubble diffusers for 18 days. Finally, PAS was reintroduced, and the MBR was operated for another six days. The bioreactor had a total volume of 30.6 L ($16 \times 25.5 \times 75$ cm) with a working volume of 6.5 L. The MBR was equipped with a flat-sheet membrane (XJ3 module by Kubota made of chlorinated polyethylene with an effective filtration area of 0.11 m^2 and nominal pore size of 0.4 μm). The bioreactor base was fitted with a stone coarse bubble diffuser (Uxcell, model number: US-SA-AJD-231698, Hong Kong) to provide shear, thus cleaning the membrane surface.

The MBR was fed with synthetic wastewater as described in Table 1. The feed has COD, TN, and TP concentrations of 1140 mg/L, 52.3 mg/L, and 11 mg/L, respectively. Two piston pumps were provided to extract the permeate out from the bioreactor through the membrane and to add the synthetic wastewater. A hydraulic retention time of 4.0 h and an SRT for the sludge of 10 days were set in the MBR system. The MBR system operated at steady-state.

During the lab-scale MBR operation, samples were regularly collected to determine the concentration of the following parameters in the treated effluent (permeate) COD, TN, NH_4-N, NO_2-N, NO_3-N, and PO_4-P, and in the filtered mixed liquor SMP and EPS. The analytical determinations were carried out using conventional analytical techniques as well as NIR spectroscopy.

The samples collected from the lab-scale MBR included: (i) the treated effluent/permeate (referred as the effluent filtered sludge (EFS)); and (ii) the supernatant of the centrifuged sludge referred to as: (a) the sludge supernatant for the determination of SMP (SS-SMP), and (b) the sludge supernatant for the determination of EPS (SS-EPS) (depending on the method of sample preparation as explained in section 2.2.2). Thirteen samples were taken during the MBR experiment. Eight samples were collected when the MBR system was operated with the fine bubble diffuser aeration system (abbreviated MBR-DIFF), and five samples were collected when the MBR system was operated with the PAS (abbreviated MBR-PAS). The summary of the abbreviation used for sampling purposes is presented as follows:

- MBR-DIFF: MBR system provided with fine bubble diffusers
- MBR-PAS: MBR system provided with PAS
- EFS: MBR effluent
- SS-SMP: Sludge supernatant for SMP measurement
- SS-EPS: Sludge supernatant for EPS measurement

2.2. Analytical determination

2.2.1. Chemical analyses – conventional analytical methods

Conventional analytical determinations for COD, NH_4-N, NO_2-N, NO_3-N, TN, and PO_4-P were conducted with Hach Lange Cuvette Tests (LCK 238, 303, 304, 314, 339, 341, 342, 350, 514).

2.2.2. SMP and EPS analyses – conventional analytical methods

The SMP and EPS analytical determinations were carried out following the method described by Le-Clech et al. (2006). Mixed liquor samples containing a volume of 60 mL were centrifuged for 5 min at 5,000g using a Rotina 35 centrifuge (Hettich, Germany). The supernatant was then filtered through a 1.2 μm Minisart® syringe filter (Sartorius, Germany). The filtrate represented the SMP solution referred to as SS-SMP. The SMP concentration was determined by analysing the TOC content using a TOC analyser (TOC-5000A, Shimadzu, Japan) and determining the UV_{254} absorbance by a spectrophotometer UNICAM Helios Beta (Thermo Fisher Scientific, USA) following the method reported by Jarusutthirak and Amy (2006).

After removing the supernatant from the sample for the SMP determination, the remaining pellet retained at the bottom of the centrifuge tube was re-suspended with demineralised water. The mixture was then heated for 10 min at a temperature of 80 °C in a water bath (Memmert, Germany), which was then centrifuged for 10 min at 7,000g. The supernatant was then filtered through the 1.2 μm filter, which represented the EPS solution referred to as SS-EPS. The EPS can be further classified into carbohydrate EPS (EPS_c), and protein EPS (EPS_p). The EPS_c was determined by photometric methods performing H_2SO_4/phenol oxidation followed by a colourimeter method using a DR3900 spectrophotometer (Hach, USA) (Lowry et al., 1951). The EPS_p was determined by applying the Folin–Ciocalteu method with a bovine serum albumin solution as the standard (Dubois et al., 1956).

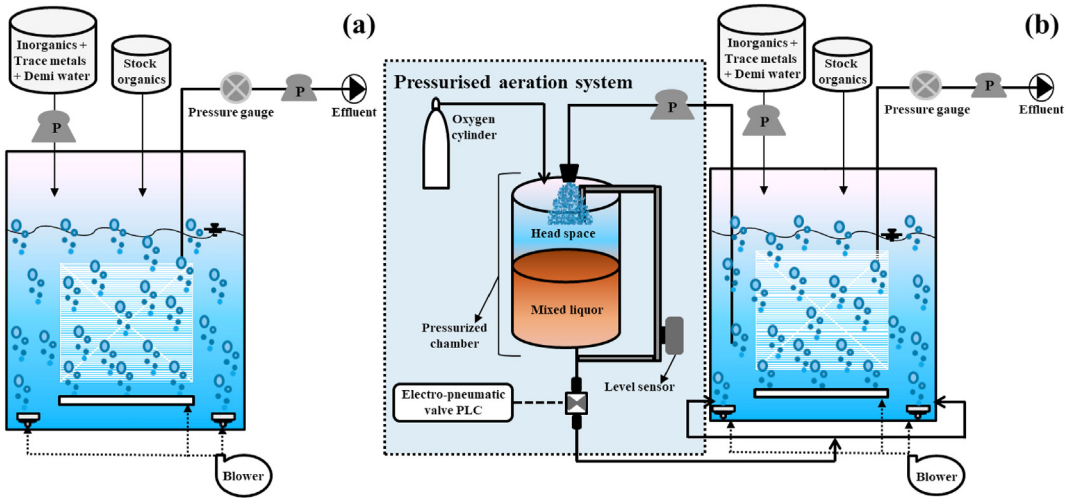


Fig. 1. Experimental setup of the lab-scale MBR with two aerations systems: (a) bubble diffusers and (b) PAS.

Table 1
Characterization of the synthetic wastewater reaching the MBR system.

Chemical compounds	Concentration (mg/L)	Chemical compounds	Concentration (mg/L)
$C_6H_{12}O_6$	421.88	$FeCl_3 \cdot 6H_2O$	19.36
$C_2H_3NaO_2$	571.28	$C_{10}H_{14}N_2Na_2O_8 \cdot 2H_2O$	30.00
Peptone	260.00	$MnCl_2 \cdot 4H_2O$	0.74
Yeast	40.00	$ZnSO_4 \cdot 7H_2O$	2.50
NH_4Cl	65.69	$CuSO_4 \cdot 5H_2O$	0.61
KH_2PO_4	48.33	$CoCl_2 \cdot 6H_2O$	2.09
$NaHCO_3$	251.95	$Na_2MoO_4 \cdot 2H_2O$	0.26
$CaCl_2$	40.37	H_3BO_3	0.13
$MgSO_4$	65.65	$NiSO_4 \cdot 7H_2O$	0.29

2.2.3. NIR spectroscopy

The measurements of the spectra were performed using a NIR spectrophotometer NIR128L-1.7 (Control Development, South Bend, Indiana, USA) provided with a control development software Spec32 using a halogen light source (HL-2000) for the wavelength range of $\lambda = 904–1699$ nm. The complete NIR instrument setup has been previously described in a previous study carried out by Bicanic et al. (2015). Ten consecutive determinations were recorded for every sample across the entire spectral range of the instrument. For all samples, no prior treatment was used and 2 mL of sample was transferred to a 10 mm quartz cuvette placed in a holder to avoid any light interference during the measurement. The recorded spectra, referred to as the raw spectra in the text, were not pre-processed to alleviate the entire subsequent chemometric processing. Also, no mechanical or chemical pre-treatment of the samples was required prior to performing the NIR spectroscopy measurements.

2.3. Data modelling

2.3.1. Principal component analysis

PCA was used both to identify patterns and to highlight similarities and differences between data from each set of the experiments. Furthermore, PCA was used to extract as much significant information as possible from a large data set and put it into a more suitable form for further calculations. Since the data table contained 896 values for each recorded spectra (considering all evaluated wavelengths), the physico-chemical properties of the samples could be hidden in such a large amount; therefore, it was necessary to reduce as much as possible the number of variables needed for further calculation without losing significant information. Also, it is important to select the key wavelengths in the NIR spectrum; especially those associated with the vibration of molecules in the observed wavelength range. These wavelengths are more significantly related to the physicochemical parameters of interest and they have to be chosen before conducting the PCA analysis. Then after conducting the PCA analysis those physico-chemical parameters will be preserved in factors. The relevance of the selected number of factors that will be observed was estimated based on the percentage of the entire number of factors describing the variation in the observed dataset with the final goal to capture the highest possible percentages of the variation (Bicanic et al., 2015). The raw unprocessed spectra were used to perform PCA using Unscrambler® X 10.4 software (CAMO Software, Norway).

2.3.2. PLSR modelling

To predict the values of the measured parameters in the samples from the NIR spectra, PLSR modelling was applied. The PLSR model is one of the most popular linear calibration methods used in quantitative NIR data analysis (Alexandrino and Poppi, 2013; Huang et al., 2013). The PLSR model simultaneously reduces the amount of data (dimension reduction of the NIR spectra matrix) while performing the regression analysis. The PLSR methodology prediction is achieved by extracting from the predictors, a set of orthogonal factors called latent variables, which have the best predictive power (Abdi, 2003). PLSR modelling was carried out on coordinates of NIR spectra factors obtained by PCA using Unscrambler® X 10.4 software (CAMO software, Norway). Model performance was assessed by evaluating the square of the

correlation coefficients of calibration (R^2_C) and validation (R^2_V), the root mean square error of calibration (RMSEC), the root mean square error of prediction (RMSEP), the standard error of prediction (SEP), the residual predictive deviation (RPD), and the ratio of error range (RER). The RPD represents the ratio of the standard deviation of the original data to SEP, while the RER represents the ratio of the range of the original data to SEP. A good NIR calibration model exists if the value of the correlation coefficient is above 0.81, while the minimum values for both RPD and RER are above 2.5. In terms of a model that is excellent and can be used for any application, the correlation coefficient value should be above 0.98 with RPD and RER values above 8.1 and 41 respectively (Mangalvedhe et al., 2015).

2.3.3. Artificial neural network modelling

To predict the values of the measured parameters in the samples from the NIR spectra, ANN modelling was applied. Multiple layer perceptron networks were developed using Statistica 10.0 software (StatSoft, USA). Based on the PCA results, coordinates containing the first five factors from PCA were selected and used as input variables related to the percentage of total variance. These first five factors explained 99.9% of the total variance. The results of analytical tests related to the concentrations of COD, TN, NH_4-N, NO_2-N, NO_3-N, PO_4-P, SMP, EPS_c and EPS_p were used as output values. The ANN training was performed by separating the data into training, testing, and validation at 60:20:20 ratio. A back-error propagation algorithm available in Statistica 10.0 (Stat-Soft, USA) was applied for the training the model. Model performance was evaluated using R^2 and root mean square error (RMSE) coefficients for training, testing and validation.

3. Results and discussion

The samples obtained from the MBR system were analysed by both conventional analytical methods and NIR spectroscopy. The result of NIR spectroscopy is a whole spectrum and not just a single value as in conventional analytical methods. Therefore, the NIR results are needed to be further processed by the use of models. In this research, PLSR and ANN were selected as processing models to investigate whether they can model the NIR spectra to predict values obtained by conventional analytical methods and to determine which model, PLSR or ANN performed more accurately. The MBR system was equipped with two completely different aeration systems that exposed the biomass to different conditions, namely pressure and shear, which in turn could have affected the biological performance of the system. For this reason, the NIR spectra were analysed to evaluate the ability of the developed models to predict the origin of the samples; i.e., to evaluate whether the models can relate the samples to the specific aeration technology used and to detect changes in the operational performance of the MBR system.

3.1. Conventional analytical methods

The values of the parameters determined by conventional analytical methods are presented in Table 2. The minimum, maximum, average, and standard deviation values were given for the 13 samples analysed EFS, SS-SMP and SS-EPS. The composition of the effluent from the MBR system indicated relatively low COD, phosphate, and ammonia concentrations as expected from a state-of-the-art biological wastewater treatment system (i.e., the MBR), which utilised organic matter, nitrogen, and phosphorus for bacterial growth. The concentrations of ammonia, nitrate, and nitrite indicated that complete nitrification was achieved in the MBR system (i.e., the ammonia originally present in the influent wastewater was converted almost completely to nitrate). The centrifuged samples of sludge from which both EPS and SMP were determined, exhibited a relatively high concentration of such microbial by-products; this was also expected considering the relatively high sludge concentration in the MBR system, of approximately 15 g/L.

3.2. Partial least squares discriminant analysis

NIR determinations were made by measuring 10 spectra for each of the 13 samples analysed for EFS, SS-SMP, and SS-EPS, yielding a total of 390 spectra. Since different samples often resulted in similar trends and sometimes almost identical spectra, there was the need to detect fingerprint within the spectra that could be used for distinguishing the individual parameters to be determined from the analysed samples. Fig. 2 shows the combined NIR average raw spectra for the three types of analysed samples (EFS, SS-SMP, and SS-EPS). In the wavelength range from 928 to 1350 nm, it was not possible to distinguish major differences between the spectra. This was also confirmed by performing factor analysis (FA), so this part of the spectra was not used for further analyses and calculations. Selecting or rejecting a particular region of wavelengths cannot simply rely on visual inspection. Even the smallest differences can sometimes hide relevant information. Therefore, to find the fingerprint for the subsequent chemometric analysis (PLSR and ANN), the average spectra for each sample (EFS, SS-SMP and SS-EPS) was obtained from all the original raw spectra recorded. On these average spectra for all the samples, FA in the observed wavelength region (λ = 904–1699 nm) with a weight setpoint of 0.8 was performed. This is the simplest way to obtain relevant part or parts of spectra containing significant information about the samples. As presented in Fig. 2, three different significant wavelength regions were identified: (i) 904–928 nm which detects the third overtone C-H vibrations and the second overtone region for C-H stretches, (ii) 1350–1356 nm, and (iii) 1400–1699 nm related to the first and second overtone region for O-H and N-H, as well as the first overtone region for C-H (Eldin, 2011). For those three regions, it is visible that the highest differences accrue in the third region (1400–1699 nm) where average EFS spectra have the lowest absorbance and average SS-EPS spectra have the highest

Table 2
The concentration range of measured parameters using conventional analytical methods.

Sample	EFS						SS-SMP	SS-EPS	
Parameter	COD (mg/L)	TN(mg/L)	NH_4-N(mg/L)	NO_2-N (mg/L)	NO_3-N (mg/L)	PO_4-P (mg/L)	SMP (L/mg-m)	EPS_c (mg/L)	EPS_p (mg/L)
Minimum	17.3	12.8	0.07	0.001	2.7	0.1	0.08	340.8	539.4
Maximum	60.5	41.4	3.25	0.14	36.0	8.2	0.58	648.9	1336.6
Average	37.7	26.7	0.21	0.05	22.0	4.1	0.31	484.2	825.5
St.dev.	12.4	6.2	0.56	0.04	6.7	2.1	0.18	104.2	245.9

St.dev: standard deviation.

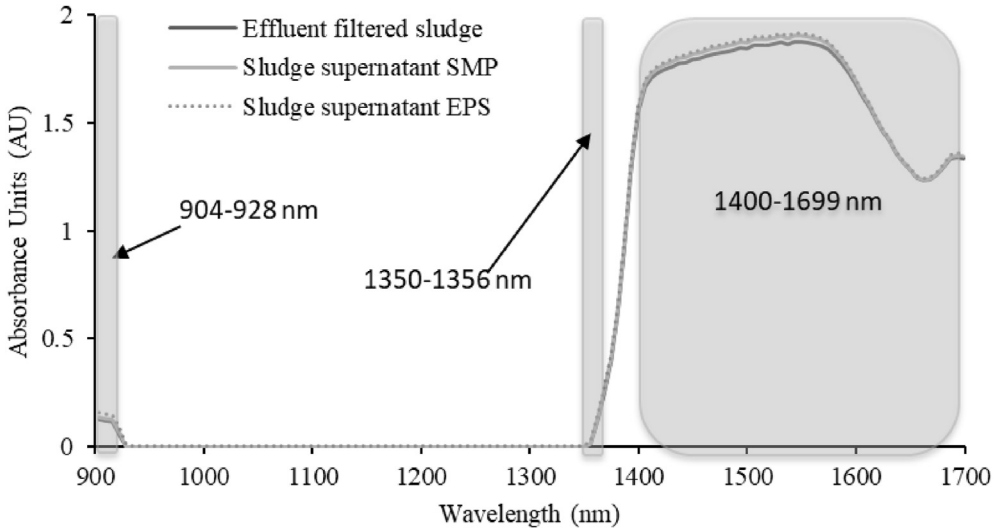


Fig. 2. Near-infrared average raw spectra of the EFS, SS-SMP and SS-EPS with the significant wavelengths.

absorbance. The same is visible for the first region (904–928 nm) and the lowest distinction between different spectra of samples visible in the second region (1350–3356 nm).

After unravelling the relevant region of the spectra that could provide significant information about the samples, a partial least squares discrimination analysis (PLS-DA) was carried out to determine if the analysed spectra could distinguish variations between the different evaluated samples (EFS, SS-SMP, and SS-EPS). Moreover, since the samples also come from the MBR system operating with different sources of aeration systems (MBR-DIFF and MBR-PAS), the PLS-DA was also carried out to evaluate the capacity of the analysed spectra and also to distinguish variations among those samples. PLS-DA has enabled the development of a non-invasive on-line monitoring method using analytical results in a way that it allows running processes to be corrected if needed. It also shortens the time used in each step and increases the quality of the end-product (analytical result) (Alcalà et al., 2010). The PLS-DA was able to distinguish from 39 tested samples and determine which sample belonged to a different group (EFS, SS-SMP or SS-EPS). For instance, the PLS-DA could completely distinguish the EFS samples from the SS-SMP and SS-EPS samples for all the evaluated samples. Similarly, the same situation was observed for all the SS-SMP samples and all the SS-EPS samples. Furthermore, the PLS-DA was able to distinguish differences between the samples taken from the MBR system when provided with the fine bubble diffusers (24 MBR-DIFF samples out of the 39 samples analysed) from those taken with the MBR system provided with the PAS (15 MBR-PAS samples out of the 39 samples analysed) with 100% separation efficiency.

The next step was to investigate the quantitative relation between the NIR spectra and the evaluated parameters in the EFS (COD, TN, NH_4-N, NO_2-N, NO_3-N, PO_4-P), SS-SMP and SS-EPS (EPS_c, EPS_p) samples. To evaluate such quantitative relationship, PLSR models were applied. The NIR spectra of the samples were used as the inputs and the parameters measured with conventional analytical methods as the outputs. The results of the PLSR models are presented in Table 3.

Three different PLSR models were performed for each of the evaluated samples (EFS, SS-SMP and SS-EPS). The first PLSR model was performed on all the NIR spectra including all the samples taken from the MBR system regardless of the aeration system used (indicated as Combined MBR-DIFF and MBR-PAS in Table 3). The second and third PLSR models were performed on samples taken from the MBR system provided with the bubble diffusers or with the PAS, respectively. The data matrix of the PLSR model consisted of 429 rows and 341 columns. Each of the EFS, SS-SMP and SS-EPS samples, represented one-third of the rows, respectively. The PLSR model was conducted in two steps as follows: first, the model needed to be trained, and later validated. Every second NIR spectra (out of the 10 spectra carried out per sample – EFS, SS-SMP, and SS-EPS) were used for training the model with the corresponding measured and evaluated parameter. The remaining NIR spectra were used for validating the model with the corresponding measured parameter. Before conducting the PLSR models, all the NIR spectra were pre-treated by applying the standard normalised variate (SNV) method. The PLSR models were subsequently calibrated and validated. For evaluating the model, standard parameters as described by Yang et al. (2016) were used, including RMSEC, R^2_C, RMSEP, SEP, R^2_V, RPD, and RER. RMSEP and SEP, as the simplest and most efficient indicators, to determine the uncertainty in the NIR predictions. Good models are obtained when: (i) the RMSE and RMSEP coefficients (which represent the prediction error of the model) are as low as possible; (ii) the RPD coefficient ranges between 2.5 and 3 or higher (Nicolai et al., 2007); and (iii) the RER coefficient ranges between 2.5 and 10 or higher (AACC method 39-00). The study of Yang et al. (2016) suggested that a model with an RPD in the range of 2.5–3 and more than 3 provides good and excellent prediction, respectively. Based on the RPD values in our study, almost all the models showed either very good or excellent results in terms of correlating the NIR spectra with the measured parameters.

When comparing the efficiency of the model for different aeration systems highest values of R^2 for validation of 0.933–0.980 and R^2 values of 0.966–0.990 for calibration were observed in the treated effluent/permeate (EFS) for fine bubble diffuser aeration system (MBR-DIFF) samples with RPD values ranging from 4.060 to 7.425 and RER values ranging from 10.349 to 19.974. The lowest efficiency of the model was observed for sludge supernatant for determining (SS-SMP) for combined spectra of MBR-DIFF and MBR-PAS SMP samples with R^2 value of 0.803 for validation and R^2 value of 0.896 for calibration with RPD and RER values of 2.338 and 6.589, respectively. It is also important to notice that the highest efficiency of the model was achieved for EPS_p samples for MBR-PAS aeration system of sludge supernatant for determining EPS with R^2 value of

Table 3
Results of the PLSR models (RMSE, RMSEP, SEP, R^2_C, R^2_V) and model efficiency parameters (RPD, RER) in different samples obtained from two different aeration systems used in lab-scale MBR.

Sample	Treatment	Parameter	Calibration		Validation			RPD	RER
			RMSE	R^2_C	RMSEP	SEP	R^2_V		
EFS	Combined MBR-DIFF and MBR-PAS	COD	2.774	0.962	3.015	3.022	0.926	3.785	13.136
		TN	2.119	0.924	2.410	2.418	0.854	2.740	10.257
		NH_4-N	0.007	0.904	0.008	0.008	0.818	2.416	8.645
		NO_2-N	0.008	0.960	0.010	0.010	0.922	3.785	13.813
		NO_3-N	1.431	0.949	1.644	1.644	0.900	3.267	10.052
		PO_4-P	0.379	0.932	0.443	0.442	0.869	2.904	11.998
	MBR-DIFF	COD	1.447	0.990	1.753	1.705	0.980	7.425	19.974
		TN	1.059	0.983	1.161	1.167	0.967	5.886	19.751
		NH_4-N	0.004	0.966	0.004	0.004	0.933	4.060	10.349
		NO_2-N	0.006	0.986	0.006	0.006	0.971	6.291	19.230
		NO_3-N	1.089	0.975	1.335	1.335	0.951	4.794	12.379
		PO_4-P	0.268	0.975	0.342	0.336	0.950	4.773	15.779
	MBR-PAS	COD	0.935	0.983	1.454	1.457	0.965	5.948	14.964
		TN	1.697	0.817	2.540	2.516	0.668	1.918	4.392
		NH_4-N	0.006	0.901	0.007	0.007	0.812	2.551	5.950
		NO_2-N	0.003	0.986	0.004	0.004	0.971	6.516	16.769
		NO_3-N	0.954	0.935	1.133	1.144	0.874	3.119	7.126
		PO_4-P	0.295	0.794	0.357	0.360	0.630	1.811	4.307
SS-SMP	Combined MBR-DIFF and MBR-PAS	SMP	0.067	0.896	0.008	0.076	0.803	2.338	6.589
	MBR-DIFF	SMP	0.028	0.976	0.034	0.094	0.952	1.757	4.828
	MBR-PAS	SMP	0.032	0.658	0.047	0.047	0.918	3.853	8.901
SS-EPS	Combined MBR-DIFF and MBR-PAS	EPS_c	33.040	0.916	36.610	36.608	0.839	2.582	8.415
		EPS_p	64.402	0.960	69.243	69.413	0.922	3.711	11.484
	MBR-DIFF	EPS_c	27.980	0.947	26.950	27.051	0.897	3.314	10.924
		EPS_p	52.058	0.971	60.011	59.960	0.944	4.433	10.105
	MBR-PAS	EPS_c	19.301	0.961	24.347	24.490	0.924	4.023	9.354
		EPS_p	6.369	0.994	12.826	12.745	0.988	10.213	27.585

0.988 for validation and R^2 value of 0.994 for calibration with RPD and RER values of 10.213 and 27.585, respectively. For that parameter, the model was applicable for on-line monitoring with NIR spectroscopy. Although most of the parameters had high RPD values, some parameters like NH_4-N for EFS samples (Combined MBR-DIFF and MBR-PAS), TN and PO_4-P for EFS samples (MBR-PAS) had values of 2.416, 1.918 and 1.811, respectively. For SS-SMP samples (both in Combined MBR-DIFF and MBR-PAS and in MBR-DIFF) the SMP exhibited a lower RPD value of 2.338 and 1.757, respectively thus limiting the model for their monitoring to qualitative rather than quantitative. These lower values could be attributed to the nature of PLSR models, which are used for linear modelling and the trend of data showed nonlinear behaviour. Further investigation is needed to test how suspended matter influences PLSR models.

Although we have not found experiments similar to ours in the literature that use NIR spectroscopy and PLS modelling to measure constituents characteristic of MBR, some studies have attempted to use NIR spectroscopy in combination with PLS modelling to predict certain parameters in wastewater. Inagaki et al. (2010) worked with samples from a sewage treatment facility in Nagoya, Japan. They measured TP, TN, BOD, TOC and turbidity with correlation coefficients for cross-validation of 0.79, 0.78, 0.83, 0.79, and 0.75, respectively. Furthermore, Dahlbacka et al. (2014) tried to predict COD in wastewater from pulp and paper mill. Since the values of R^2 for validation and R^2 for validation in PLS model were 0.76 and 0.19 respectively, their model was not suitable for on-line predictions. One of the highest R^2 values for cross-validation was recorded in the work of Pascoa et al. (2008) where the wastewater treatment process was conducted in a custom-designed sequential batch reactor. Using NIR and PLS modelling for total solids, TSS and COD, the R^2 values were 0.92, 0.91, 0.87 with RER values of 15.6, 15.8, 9.8 and RPD values of 3.48, 3.54, 2.75 respectively.

3.3. Artificial neural networks

Although very good predictions were observed for PLSR models, the scope of this research was also to compare the predictions of NIR spectra with ANN models. The first step for developing the ANN models involved performing PCA on all spectra for all samples. The NIR spectra used in this work contained 793 wavelengths. Therefore, to reduce the amount of data that could later be used to run the ANN models, PCA was performed to obtain coordinates of factors (scores). PCA has been used in several studies to scale down NIR spectroscopy wavelengths to match the column number of output variables used for the ANN model (Allouche et al., 2015; Candolfi et al., 1999; Dou et al., 2007; Gorry, 1990). This reduces the training time of the ANN model. The first five factors obtained by PCA, which explained 99.99% of the variance, were used as inputs for the ANN analyses; the same parameters as in the PLSR models were used as outputs including COD, TN, NH_4-N, NO_2-N, NO_3-N, PO_4-P from the EFS samples, SMP from the SS-SMP samples, and EPS_c and EPS_p from the SS-EPS samples. The development of ANN, similar to human brain, requires the use of a large amount of data for training purposes. From the collected NIR spectra, the largest percentage of data needed to be used for training, while the rest of the data was split between the testing of the learned knowledge and the validation of the trained and tested knowledge. The ANN model was then performed by splitting the experimental data. For instance, out of the 10 NIR spectra obtained per each of the evaluated samples, six of them were used for training, two of them for testing, while the remaining two for validation. One hidden layer was used for the ANN development, and the number of neurons in the hidden layer was set to a range from 3 to 11. The very same procedure as used in the PLSR models was performed regarding the use of the data from the EFS, SS-SMP and SS-EPS samples to obtain the ANN models. As described in Table 3, the three sets of samples (EFS, SS-SMP, and SS-EPS) were evaluated. For each of the samples evaluated, one network was selected for all samples without

Table 4
Characteristic of the ANN models developed for the prediction of parameters.

Sample	Treatment	Network structure	Training perf. (R^2)	Training Error (RMSE)	Testing perf. (R^2)	Testing Error (RMSE)	Validation perf. (R^2)	Validation Error (RMSE)	Hidden activation	Output activation
EFS	Combined MBR-DIFF and MBR-PAS	5-11-6	0.955	0.018	0.959	0.017	0.942	0.026	Exponential	Logistic
	MBR-DIFF	5-10-6	0.997	0.002	0.996	0.003	0.994	0.004	Tanh	Logistic
	MBR-PAS	5-8-6	0.994	0.005	0.987	0.023	0.956	0.036	Tanh	Logistic
SS-SMP	Combined MBR-DIFF and MBR-PAS	5-9-1	0.943	0.006	0.925	0.009	0.919	0.008	Tanh	Identity
	MBR-DIFF	5-5-1	0.996	4E-1	0.994	0.001	0.997	0.001	Tanh	Logistic
	MBR-PAS	5-10-1	0.991	0.001	0.992	0.001	0.971	0.005	Tanh	Logistic
SS-EPS	Combined MBR-DIFF and MBR-PAS	5-9-2	0.984	0.003	0.949	0.010	0.958	0.007	Tanh	Logistic
	MBR-DIFF	5-6-2	0.980	0.008	0.937	0.014	0.979	0.007	Tanh	Logistic
	MBR-PAS	5-9-2	0.997	0.001	0.997	0.001	0.984	0.005	Tanh	Identity

considering the aeration source (Combined MBR-DIFF and MBR-PAS); a second and third network were selected when the MBR system had bubble diffusers and with PAS, respectively. Since for each step multiple ANN were tested, the performance of the ANN models was evaluated based on both the R^2 and RMSE coefficients for the training, testing and validation phases; the selected ANN are shown in Table 4.

The criteria for evaluating the performance of the models were as follows: R^2 values below 0.70 indicated that the model could only distinguish low-medium-high values; R^2 values between 0.70 and 0.90 indicated that the models could be considered accurate, and R^2 above 0.90 indicated a good model (Liu et al., 2011; Urbano-Cuadrado et al., 2004). The results obtained for the ANN models for all the cases showed very good correlations between the NIR spectra and the investigated parameters. On the ANN modelling, different neural networks were used as can be seen from the network structure in Table 4. For example, under the EFS sample for MBR-DIFF treatment, the ANN model presented five inputs, 10 neurons in a hidden layer, and six outputs, which gave the best performance in terms of training test and validation with R^2 values of 0.997, 0.996 and 0.994, respectively. The five inputs refer to the first five PCA factors of the NIR spectra related to the six outputs (evaluated/measured parameters: COD, TN, NH_4-N, NO_2-N, NO_3-N, PO_4-P), through the 10 neurons. Table 5 describes the correlation between the NIR spectra and the specific evaluated/measured parameter obtained by the ANN models. A very good correlation between the NIR spectra and the measured parameter was obtained for all the evaluated parameters, including both aeration systems. For instance, the ANN model for the previous example with a network structure of 5-10-6 for the EFS sample when the MBR was provided with the bubble diffuser aeration exhibited an R^2 validation performance of 0.996 for COD, 0.999 for TN, 0.997 for NH_4-N, 0.993 for NO_2-N, 0.993 for NO_3-N, and 0.989 for PO_4-P.

The results showed a very good correlation between the values predicted by the ANN model and the measured water quality parameters by conventional analytical techniques for all constituents determined in the analysed EFS, SS-SMP, and SS-EPS samples. Moreover, good correlations were also obtained regardless of the aeration system in the MBR system. The ANN models had R^2 values above 0.9 for all measured parameters for the training, testing, and

Table 5
ANN models for the prediction of the evaluated parameters.

Sample	Treatment	Parameter	Training perf. (R^2)	Testing perf. (R^2)	Validation perf. (R^2)
EFS	Combined MBR-DIFF and MBR-PAS	COD	0.989	0.973	0.972
		TN	0.976	0.986	0.969
		NH_4-N	0.990	0.971	0.990
		NO_2-N	0.993	0.991	0.994
		NO_3-N	0.994	0.986	0.987
		PO_4-P	0.989	0.965	0.979
	MBR-DIFF	COD	0.998	0.995	0.996
		TN	0.997	1.000	0.999
		NH_4-N	0.997	0.996	0.997
		NO_2-N	0.995	0.996	0.993
		NO_3-N	0.998	0.995	0.993
		PO_4-P	0.996	0.995	0.989
	MBR-PAS	COD	0.979	0.984	0.989
		TN	0.997	0.972	0.981
		NH_4-N	0.991	0.997	0.988
		NO_2-N	0.994	0.999	0.998
		NO_3-N	0.969	0.957	0.946
		PO_4-P	0.978	0.983	0.969
SS-SMP	Combined MBR-DIFF and MBR-PAS	SMP	0.984	0.965	0.910
	MBR-DIFF	SMP	0.996	0.994	0.997
	MBR-PAS	SMP	0.991	0.992	0.971
SS-EPS	Combined MBR-DIFF and MBR-PAS	EPS_c	0.977	0.952	0.941
		EPS_p	0.992	0.946	0.975
	MBR-DIFF	EPS_c	0.981	0.891	0.968
		EPS_p	0.979	0.983	0.989
	MBR-PAS	EPS_c	0.995	0.996	0.973
		EPS_p	0.999	0.997	0.996

validation phases of the NIR spectra. Only one R^2 value below 0.9, of 0.891, was observed in the testing phase for EPS_c in the SS-EPS samples when the MBR system was provided with the bubble diffusers. However, this sample had training and validation R^2 values of 0.981 and 0.968, respectively.

3.4. Practical application

Based on the obtained results for all the evaluated parameters, modelled NIR spectra can be used as a reliable alternative for monitoring COD, TN, NH_4-N, NO_2-N, NO_3-N, PO_4-P, SMP, EPS_c, EPS_p in MBR wastewater samples for samples that do not contain suspended matter and an expert is available to model the results of NIR spectroscopy.

In this regard, NIR spectroscopy may be able to overcome the limitations of conventional analytical methods for the determination of specific wastewater treatment constituents such as microbial products like SMP and EPS. In terms of MBR operation, the ability to quickly and accurately measure microbial products associated with membrane fouling could lead to more sustainable operation with fewer chemicals for cleaning, reducing its environmental footprint and supporting its application in water reuse and resource recovery. In addition, the ability of NIR to be used as an on-line measurement tool allows the operator of an MBR treatment to obtain important information about the performance of the treatment in near real-time.

4. Conclusions

NIR spectroscopy of wastewater samples combined with chemometric modelling using PLSR and ANN has been very successful in predicting several important water quality parameters in wastewater treatment. In particular, the determination of SMP, EPS_c and EPS_p was extremely important as these compounds are highly associated with membrane fouling in MBR systems. ANN models achieved better performance in terms of correlation of NIR spectra with all measured parameters compared to PLSR. This approach of processing a large amount of spectroscopy data through chemometric modelling led to a promising strategy for wastewater treatment monitoring. The models distinguished between different samples that came from the same wastewater treatment provided with different aeration systems that had affected the water quality and sludge properties. This feature cannot be achieved by conventional analyses, and it seems that a lot of insight can be gained by chemometric analysis of NIR spectra.

Credit author statement

Sang Yeob Kim: Methodology, validation, formal analysis, investigation, visualisation, and writing (original and revised manuscript). Josip Ćurko: Methodology, validation, analysis, supervision, and writing (original and revised manuscript). Jasenka Gajdoš Kljusurić: Supervision, visualization, and writing (original and revised version). Marin Matošić: Supervision, conceptualisation, visualization, and writing (original and revised version). Vlado Crnek: Methodology, validation, formal analysis, investigation. Carlos M. Lopez-Vazquez: Supervision, visualization, and writing (original version). Hector A. Garcia: Methodology, validation, analysis, supervision, and writing (original version). Damir Brdjanović: Conceptualization, supervision, funding acquisition. Davor Valinger: Conceptualization, methodology, validation, analysis, supervision, and writing (original and revised manuscript).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

List of abbreviations

ANN	artificial neural networks
BOD	biochemical oxygen demand
COD	chemical oxygen demand
CAS	conventional activated sludge
EFS	effluent filtered sludge
EPS	extracellular polymeric substances
EPS_c	carbohydrate fraction of the extracellular polymeric substances
EPS_p	protein fraction of the extracellular polymeric substances
FA	factor analysis
MBR	membrane bioreactor
MBR-DIFF	membrane bioreactor provided with fine bubble diffusers
MBR-PAS	membrane bioreactor system provided with pressurised aeration system
MLSS	mixed liquor suspended solids
MLRA	multiple linear regression analysis
NIR	near-infrared
PLSR	partial least squares regression
PLS-DA	partial least squares discrimination analysis
PAS	pressurised aeration system
PCA	principal component analysis
RER	ratio of error range
RPD	residual predictive deviation
RMSE	root mean square error
RMSEC	root mean square error of calibration
RMSEP	root mean square error of prediction
R^2_C	square of the correlation coefficients of calibration
R^2_V	square of the correlation coefficients of validation
SS-EPS	sludge supernatant for determination of extracellular polymeric substances
SS-SMP	sludge supernatant for determination of soluble microbial products
SRTs	solid retention times
SMP	soluble microbial products
SEP	standard error of prediction
SNV	standard normalised variate
TN	total nitrogen
TOC	total organic carbon
TP	total phosphorus
TSS	total suspended solids
WWTP	wastewater treatment plant

References

Abdi, H., 2003. PLS-Regression; Multivariate analysis. In: Lewis-Beck, M., Bryman, A., Futing, T. (Eds.), Encyclopedia of Social Sciences Research Methods. SAGE publications, Thousand Oaks (CA), USA, pp. 792–795.

Alcalà, M., Blanco, M., Bautista, M., Gonzàlez, J.M., 2010. On-line monitoring of a granulation process by NIR spectroscopy. J. Pharmaceut. Sci. 99 (1), 336–345.

Alexandrino, G.L., Poppi, R.J., 2013. NIR imaging spectroscopy for quantification of constituents in polymers thin films loaded with paracetamol. Anal. Chim. Acta 765, 37–44.

Allouche, Y., Funes López, E., Beltrán Maza, G., Jiménez Márquez, A., 2015. Near infrared spectroscopy and artificial neural network to characterise olive fruit and oil online for process optimisation. J. Near Infrared Spectrosc. 23, 111–121.

Bagheri, M., Akbari, A., Mirbagheri, S.A., 2019. Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: a critical review. Process Saf. Environ. Protect. 123, 229–252.

Bicanic, D., Streza, M., Dóka, O., Valinger, D., Luterotti, S., Ajtony, Zs, Kurtanjek, Z., Dadarlat, D., 2015. Non-destructive measurement of total carotenoid content in processed tomato products: infrared lock-in thermography, near-infrared spectroscopy/chemometrics, and condensed phase laser-based photo-acoustics-pilot study. Int. J. Thermophys. 36 (9), 2380–2388.
Candolfi, A., Maesschalck, R.De, Jouan-Rimbaud, D., Hailey, P.A., Massart, D.L., 1999. The influence of data pre-processing in the pattern recognition of excipients near-infrared spectra. J. Pharmaceut. Biomed. Anal. 21, 115–132.
Comte, S., Guibaud, G., Baudu, M., 2006. Relations between extraction protocols for activated sludge extracellular polymeric substances (EPS) and EPS complexation properties: Part I. Comparison of the efficiency of eight EPS extraction methods. Enzym. Microb. Technol. 38 (1–2), 237–245.
D'Abzac, P., Bordas, F., Van Hullebusch, E., Lens, P.N., Guibaud, G., 2010. Extraction of extracellular polymeric substances (EPS) from anaerobic granular sludges: comparison of chemical and physical extraction protocols. Appl. Microbiol. Biotechnol. 85 (5), 1589–1599.
Dahlbacka, J., Nyström, J., Mossing, T., Geladi, P., Lillhonga, T., 2014. On-line measurement of the chemical oxygen demand in wastewater in a pulp and paper mill using near infrared spectroscopy. Spectr. Anal. Rev. 2, 19–25.
Dias, A.M.A., Moita, I., Páscoa, R., Alves, M.M., Lopes, J.A., Ferreira, E.C., 2008. Activated sludge process monitoring through in situ near-infrared spectral analysis. Water Sci. Technol. 57 (10), 1643–1650.
Dou, Y., Zou, T., Liu, T., Qu, N., Ren, Y., 2007. Calibration in non-linear NIR spectroscopy using principal component artificial neural networks. Spectrochim. Acta, Part A 68, 1201–1206.
Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., Smith, F., 1956. Colorimetric method for determination of sugars and related substances. Anal. Chem. 28, 350–356.
Eldin, A.B., 2011. Near infrared spectroscopy. In: Isin Akyar, Dr (Ed.), Wide Spectra of Quality Control (InTechOpen, Rijeka, Croatia).
Frølund, B., Palmgren, R., Keiding, K., Nielsen, P.H., 1996. Extraction of extracellular polymers from activated sludge using a cation exchange resin. Water Res. 30 (8), 1749–1758.
Gajdoš Kljusurić, J., Valinger, D., Jurinjak Tušek, A., Benković, M., Jurina, T., 2017. Application of near infrared spectroscopy (NIRs), PCA and PLS models for the analysis of dried medicinal plants. Science within Food: Up-to-date advances on research and educational ideas/Méndez-Vilas, A. Formatex Research Center 28–35.
Gorry, P.A., 1990. General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. Anal. Chem. 62 (6), 570–573.
Han, X., Lü, E., Lu, H., Zeng, F., Qiu, G., Yu, Q., Zhang, M., 2020. Detection of spray-dried porcine plasma (SDPP) based on electronic nose and near-infrared spectroscopy data. Appl. Sci. 10, 2967.
Henze, M., Van Loosdrecht, M.C.M., Ekama, G.A., Brdjanovic, D., 2008. Biological Wastewater Treatment: Principles, Modelling and Design. IWA publishing, London, UK.
Huang, Y., Cao, J., Ye, S., Duan, J., Wu, L., Li, Q., Min, S., Xiong, Y., 2013. Near-infrared spectral imaging for quantitative analysis of active component in counterfeit imidacloprid using PLS regression. Optik 124, 1644–1649.
Inagaki, T., Shinoda, Y., Miyazawa, M., Takamura, H., Tsuchikawa, S., 2010. Near-infrared spectroscopic assessment of contamination level of sewage. Water Sci. Technol. 61, 1957–1963.
Jarusutthirak, C., Amy, G., 2006. Role of soluble microbial products (SMP) in membrane fouling and flux decline. Environ. Sci. Technol. 40 (3), 969–974.
Kim, S.Y., Garcia, H.A., Lopez-Vazquez, C.M., Milligan, C., Livingston, D., Herrera, A., Matosic, M., Curko, J., Brdjanovic, D., 2019. Limitations imposed by conventional fine bubble diffusers on the design of a high-loaded membrane bioreactor (HL-MBR). Environ. Sci. Pollut. Control Ser. 26 (33), 34285–34300.
Laspidou, C.S., Rittmann, B.E., 2002. A unified theory for extracellular polymeric substances, soluble microbial products, and active and inert biomass. Water Res. 36, 2711–2720.
Le-Clech, P., Chen, V., Fane, A.G., 2006. Fouling in membrane bioreactors used in wastewater treatment. J. Membr. Sci. 284, 17–53.
Li, J., Luo, G., He, L.J., Xu, J., Lyu, J., 2018. Analytical approaches for determining chemical oxygen demand in water bodies: a review. Crit. Rev. Anal. Chem. 48 (1), 47–65.
Liu, F., He, Y., Wang, L., Sun, G., 2011. Detection of organic acids and pH of fruit vinegars using near-infrared spectroscopy and multivariate calibration. Food Bioprocess Technol. 4 (8), 1331–1340.
Lowry, O., Rosebrough, N., Farr, A., Randall, R., 1951. Protein measurement with the folin phenol reagent. J. Biol. Chem. 193, 265–275.
Mangalvedhe, A.A., Danao, M.-G.C., Paulsmeyer, M., Rausch, K.D., Singh, V., Juvik, J.A., 2015. Anthocyanin determination in different hybrids using near infrared spectroscopy. ASABE Annual International Meeting Paper 3–14.
Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. Postharvest Biol. Technol. 46 (2), 99–118.
Nielsen, P.H., Jahn, A., 1999. Extraction of EPS. In: Microbial Extracellular Polymeric Substances: Characterization, Structure and Function. Springer, Berlin, Germany.
Pan, T., Chen, Z., 2012. Rapid determination of chemical oxygen demand in sugar refinery wastewater by short-wave near infrared spectroscopy. Adv. Mater. Res. 549, 167–171.
Pan, T., Chen, W., Huang, W., Qu, R., 2012. Model optimization for near-infrared spectroscopy analysis of chemical oxygen demand of wastewater. Key Eng. Mater. 500, 832–837.
Pascoa, R.N.M., Lopes, J.A., Lima, J.L.F.C., 2008. In situ near-infrared monitoring of activated dairy sludge wastewater treatment processes. J. Near Infrared Spectrosc. 16, 409–419.
Pereira, C.S., Cunha, S.C., Fernandes, J.O., 2019. Prevalent mycotoxins in animal feed: occurrence and analytical methods. Toxins 11, 290.
Prieto, N., Pawluczyk, O., Dugan, M.E.R., Aalhus, J.L., 2017. A review of the principles and applications of near-infrared spectroscopy to characterize meat, fat, and meat products. Appl. Spectrosc. 71 (7), 1403–1426.
Rinnan, A., van den Berg, F., Balling Engelsen, S., 2009. Review of the most common pre-processing techniques for near-infrared spectra. Trends Anal. Chem. 28 (10), 1201–1222.
Soh, Y.N.A., Kunacheva, C., Webster, R.D., Stuckey, D.C., 2020. Identification of the production and biotransformational changes of soluble microbial products (SMP) in wastewater treatment processes: a short review. Chemosphere 251, 126391.
Suehara, K.I., Owari, K., Kohda, J., Nakano, Y., Yano, T., 2007. Rapid and simple determination of oil and urea concentrations and solids content to monitor biodegradation conditions of wastewater discharged from a biodiesel fuel production plant. J. Near Infrared Spectrosc. 15 (2), 89–96.
Takamura, H., Miyamoto, H., Mori, Y., Matoba, T., 2002. Evaluation of drainage by near infrared spectroscopy. *Near Infrared Spectroscopy*. Proceedings of the 10th International Conference 405–408.
Urbano-Cuadrado, M., Luque de Castro, M.D., Pérez-Juan, P.M., García-Olmo, J., Gómez-Nieto, M.A., 2004. Near infrared reflectance, spectroscopy and multivariate analysis in enology - determination or screening of fifteen parameters in different types of wines. Anal. Chim. Acta 527 (1), 81–88.
Xie, L., Chen, M., Ying, Y., 2016. Development of methods for determination of aflatoxins. Crit. Rev. Food Sci. Nutr. 56 (16), 2642–2664.
Xu, L., Zhou, Y.P., Tang, L.J., Wu, H.L., Jiang, J.H., Shen, G.L., Yu, R.Q., 2008. Ensemble preprocessing of near-infrared (NIR) spectra for multivariate calibration. Anal. Chim. Acta 616, 138–143.
Yang, Q., Liu, Z., Yang, J., 2009. Simultaneous determination of chemical oxygen demand (COD) and biological oxygen demand (BOD5) in wastewater by near-infrared spectrometry. J. Water Resour. Protect. 4, 286–289.
Yang, Z., Li, K., Zhang, M., Xin, D., Zhang, J., 2016. Rapid determination of chemical composition and classification of bamboo fractions using visible–near infrared spectroscopy coupled with multivariate data analysis. Biotechnol. Biofuels 9 (35).