Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation

Julien Petit, Nassim Ait-Mouheb, Sílvia Mas García, Maxime Metz, Bruno Molle, Ryad Bendoula

Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation

Julien Petit, Nassim Ait-Mouheb, Sílvia Mas García, Maxime Metz, Bruno Molle, Ryad Bendoula

Research Paper

Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation

Julien Petit [a,b,*], Nassim Ait-Mouheb [a], Sílvia Mas García [b], Maxime Metz [b], Bruno Molle [a], Ryad Bendoula [b]

[a] INRAE, UMR GEAU, University of Montpellier, 361 Rue Jean François Breton, 34090 Montpellier, France
[b] INRAE-SupAgro, UMR ITAP, 361 Rue Jean François Breton, 34090 Montpellier, France

ARTICLE INFO

Article history:
Received 16 April 2021
Received in revised form
8 July 2021
Accepted 11 July 2021
Published online xxx

Keywords:
Drippers
Physical and chemical clogging
Absorbance spectroscopy
Partial least squares regression (PLSR)
Partial least squares with discriminant analysis (PLS-DA)
Optical coherence tomography (OCT)

Drip irrigation is one of the most efficient irrigation techniques, but it is susceptible to dripper clogging. This study proposes a novel and non-destructive method based on visible and near infrared (Vis/NIR) spectroscopy coupled with chemometric methods for the discrimination and thickness estimation of physical and chemical fouling in drip-irrigation systems. Four representative materials linked to physical and chemical clogging (kaolin, bentonite, sand and calcium carbonate) at different thicknesses were selected to illustrate the potential of the approach. Partial least squares regression (PLSR) and its modification partial least squares with discriminant analysis (PLS-DA) were selected for the modelling of clogging materials. The PLS-DA model was able to predict with 96.97% accuracy all classes of materials. The PLSR models were able to estimate fouling thickness with relative prediction errors comprised between 134 μm and 164 μm. This difference appears mainly to be due to the physical properties of the selected materials. This prediction accuracy enabled the estimation of the clogging thickness between 10 and 21% of dripper channel coverage depending on the dripper channel section and the material under study. The proposed method offers an appropriate approach for clogging studies in drip irrigation systems that could be transferred to field applications.

1. Introduction

Drip irrigation can offer up to 90% application efficiency and is one of the most effective irrigation techniques (Camp, Lamm, Evans, & Phene, 2000). It offers a high irrigation uniformity, which leads to improved crop yields (Wang, Gong, Di, Yu, & Zhao, 2013). In addition, drip irrigation is well adapted to using reclaimed water and for fertigation as it limits contact between irrigation water and the plant (Lamm, Ayars, &

* *Corresponding author*. INRAE, UMR GEAU, University of Montpellier, 361 Rue Jean François Breton, 34090 Montpellier, France. Fax: +33 499612882.
E-mail address: julien.petit@inrae.fr (J. Petit).
https://doi.org/10.1016/j.biosystemseng.2021.07.013

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013

Nomenclature

CR	classification rate
LV	latent variables
MIR	mid-infrared
NIR	near infrared
OCT	optical coherence tomography
PLS	partial least squares
PLS-DA	partial least squares with discriminant analysis
PLSR	partial least squares regression
RMSEC	root mean square error in calibration
RMSECV	root mean square error in cross-validation
RMSEP	root mean square error in prediction
SG	Savitzky–Golay
SNV	standard normal variate
UV	ultra violet
Vis	visible

Nakayama, 2006). However, drip irrigation emitters are prone to be clogged, which reduces both irrigation performance and uniformity (Lamm et al., 2006; Niu, Liu, & Chen, 2013).

There are three main types of emitter clogging: physical, chemical and biological (Bucks, Nakayama, & Gilbert, 1979; Adin & Sacks, 1991). Physical clogging is caused by the deposition of solid particles present in the irrigation water (clays, silts, fine sands) (Niu et al., 2013; Oliveira et al., 2017). Chemical clogging is associated with the precipitation of insoluble materials stemming from the reaction between cations (Ca^{2+}, Mg^{2+} and Fe^{2+}) and anions (CO_3^{2-}, HCO_3^- and SO_4^{2-}) (Hills, Nawar, & Waller, 1989; Rizk, Ait-Mouheb, Bourrié, Molle, & Roche, 2017). Biological clogging is due to bacterial development that forms a matrix called a biofilm (Lequette, Ait-Mouheb, & Wéry, 2020; Qian et al., 2017; Yan et al., 2009). Moreover, these three clogging types can occur simultaneously, forming a more complex clogging, particularly for treated wastewater irrigation (Rizk, Ait-Mouheb, Molle, & Roche, 2019; Tarchitzky, Rimon, Kenig, Dosoretz, & Chen, 2013; Yan et al., 2009).

Depending on the main type of clogging observed, different maintenance operations should be performed (Lamm et al., 2006). Upstream filtration and hydraulic flushes are usually performed to prevent physical clogging (Puig-Bargués, Lamm, Trooien, & Clark, 2010), acid injections are used to avoid chemical clogging (Hills et al., 1989) and chlorination is the most common operation used to eliminate biological development (Lequette et al., 2021). Therefore, detecting and identifying these clogging types is crucial for drip irrigation system management. Clogging detection and identification under field conditions remains problematic. Only indirect methods based on emitters' discharge or drip lines pressure are used to estimate clogging development (Camargo, Molle, Tomas, & Frizzone, 2014). However, these methods do not give information on the clogging nature and are unable to detect clogging at its early stages.

Concerning the function of the drippers used, internal flow channel depths generally vary between 800 μm and 1330 μm (Lequette et al., 2020; Yang et al., 2020), which significantly limits the number of methods that can be used to determine the fouling levels *in situ*. However, some methods that could cope with such scales appear in the literature. The electrical sensors studied by Pavanello et al. (2011) and Turolla, Di Mauro, Mezzera, Antonelli, and Carminati (2019) are suitable for detecting biofilm and calcium carbonate. Nevertheless, these sensors require deposition of the clogging material on a steel surface. These conditions do not accurately represent the clogging mechanisms found with plastic drip irrigation emitters. Vibration sensors, developed by Pereira, Mendes, and Melo (2008), are useful for biofilm detection, and the ultrasonic sensors proposed by Mizrahi et al. (2012) can detect calcium carbonate precipitates in membrane filtration. However, both these solutions are only suitable for one specific clogging material. The optical sensors developed by Mendret, Guigui, Schmitz, Cabassud, and Duru (2007) can quantify the thickness of the clogging material in membrane filtration. Nevertheless, this sensor is unable to discriminate the materials constituting the membrane clogging. Therefore, none of these devices are well suited for clogging studies in drip irrigation, and for the selection of a proper cleaning method such as chlorination or acid injections.

In this context, absorbance spectroscopy could be used for non-destructive and *in situ* measurements for dripper clogging detection in order to classify clogging type and to estimate fouling thickness. Indeed, Tarchitzky et al. (2013) have already highlighted the relevance of mid-infrared (MIR) spectroscopy for the analysis of biological and physical clogging. However, MIR equipment would not appear suitable for on-site application mainly because of its complexity and cost (Viscarra Rossel, Walvoort, McBratney, Janik, & Skjemstad, 2006; Haas & Mizaikoff, 2016). In the ultra violet/visible/near infrared (UV/Vis/NIR) region, spectroscopy can be used for low-cost *in situ* monitoring as was demonstrated by Zimmerleiter et al. (2020). Moreover, over the past few decades, the use of absorbance spectroscopy in the development of non-destructive, *in situ*, low-cost and rapid measurement has increased in different research areas such as water and soil science (Adamchuk, Hummel, Morgan, & Upadhyaya, 2004; Berho, Pouet, Bayle, Azema, & Thomas, 2004; ViscarraRossel et al., 2006, Zimmerleiter et al., 2020). This technique is usually associated with models developed using chemometric methodologies to predict the characteristics of analysed samples.

In general, chemometric methods can be used to solve problems in qualitative/quantitative and discrimination analysis (Xie et al., 2015; Lopo, Teixeira dos Santos, Páscoa, Graça, & Lopes, 2018). Among them, partial least squares regression (PLSR) (Geladi & Kowalski, 1986) and its modification partial least squares with discriminant analysis (PLS-DA) (Wold et al., 1984) are the most commonly used. However, no chemometric works have been reported in the literature for the characterisation of clogging in drip irrigation systems.

In this work, PLS-DA is used to discriminate different types of clogging materials and PLSR to perform a thickness estimation of these clogging materials. The main goal of this study was to test the applicability of Vis/NIR spectroscopy combined with chemometric methods as a mean to detect the type and estimate the thickness of physical and chemical clogging in drip irrigation systems. For this purpose, four

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013

clogging materials: kaolin, bentonite and sand (responsible for physical clogging) and calcium carbonate (responsible for chemical clogging), were tested.

The proposed methodology attempts to provide a method that could be potentially adapted to field conditions for acquiring knowledge of the clogging mechanism and their impact on drip emitters. This methodology could be used to plan maintenance operations and help to select the most suitable form of maintenance, and hence ensure functional and durable drip irrigation systems.

2. Material and methods

2.1. Samples preparation

The characteristics of the clogging materials used to prepare the samples in this study are summarised in Table 1.

As can be seen in Table 1, the studied materials represent both physical and chemical clogging. Clays and sand are associated with physical clogging (Adin & Sacks, 1991; Oliveira et al., 2020) whilst calcium carbonate is a major component of chemical clogging (Hills et al., 1989; Rizk et al., 2019). The two clays investigated (kaolin and bentonite) were chosen because they have different swelling properties (Luckham & Rossi, 1999), and thus have different impacts on dripper clogging (Ait-Mouheb et al., 2019; Oliveira et al., 2020).

In order to prepare the samples, the clogging material was placed in a 40-mm diameter and 12-mm high cylindrical glass cup. Samples were constituted using one clogging material among the four tested. Different thicknesses of clogging material were prepared ranging from 50 μm to 1000 μm with theoretical steps depending on the particle size of the clogging material (<10 μm for clays and calcium carbonate and >50 μm for sand). Solutions containing clogging material and demineralised water were prepared at different volume concentration to constitute different thicknesses after decantation. The solution was added to the cup so that the total height of the sample reached approximatively 1000 μm. This sample height was chosen to imitate the channels dimensions of the drippers (between 800 μm and 1330 μm). Optical coherence tomography (OCT) images were used to obtain the thickness reference value of the prepared samples. Three Vis/NIR spectra around 2 mm diameter were acquired on each sample. A total of 94 spectra corresponding to samples with different thicknesses were analysed: 31 from kaolin, 23 from bentonite, 31 from calcium carbonate and 9 from sand. The difference in number of samples for each material is caused by the particle size for sand and by the swelling properties for bentonite that limited the number of samples with different thicknesses.

2.2. OCT image measurement

2.2.1. Image sampling and analysis

For each sample, three-dimensional OCT measurements were acquired using a Thorlabs GANYMEDE II OCT (LSM03 lens, lateral resolution: 8 μm; Thorlabs GmbH, Lübeck, Germany). The axial voxel size in water (n = 1.333) of GANYMEDE II was 2.1 μm and the central wavelength was 930 nm. The OCT acquisition covers an area of 0.74 cm^2 that was centred according to the area of spectral acquisition.

2.2.2. Image processing

Firstly, 3-D OCT datasets were processed in Fiji (running on ImageJ version 1.50 b, Schindelin et al., 2012). Secondly, an in-house code was used to detect the pixels associated with the fouling particles using MATLAB R2018r (MathWorks, Natick, MA, USA, version 2018b). A threshold (adapted to all datasets) was applied to binarise the dataset and the region above the threshold intensity losses during simple penetration. The intensity contrast between the fouling pixels and the remaining parts (water) was deliberately high to facilitate allocation of these last areas to the back-ground (black). Thus, in the studied OCT acquisitions, the threshold values vary between 0 and 60 with noise intensity of between 0 and 15 (water) and a signal related fouling of between 25 and 60. In addition, suspended solids and outliers were removed by applying the 'clean and majority functions' of MATLAB.

In order to represent the spectral acquisition spot, the obtained OCT image was cropped to obtain a 2-mm-by-2-mm square surface. For each sample, the thickness value was calculated by averaging the thickness over this 4-mm^2 area. Next, a standard deviation was also calculated for each sample, which represents thickness variations in the area covered by the spectral acquisition.

2.3. Spectroscopy measurement

To assess the composition and the clogging thickness of the sample, absorbance spectroscopy was used. Figure 1 shows the schematic diagram of the spectroscopy system. A tungsten-halogen source (Ocean Optics HL-200-FHSA) was coupled with a 910 μm core diameter optical fibre (Thorlabs FG910LEC, NA 0.22, Thorlabs, Newton, NJ, USA) and a biconvex lens (Thorlabs LB1723-B, f = 60 mm, d = 50.8 mm), forming a ≈2 mm spot on the surface of the sample. The distance between the lens and the sample was equal to 5 cm. The sample backscattered light was collected using a biconvex lens (Thorlabs LB1092, f = 15 mm, d = 12.5 mm), coupled with a 910 μm core diameter optical fibre (Thorlabs FG910LEC, NA 0.22). This fibre was connected to a spectrometer (LabSpec1, ASD Boulder, CO, USA). The spectral range of measurement

Table 1 – Characteristics of the clogging materials used.

Sample	Particle size	Clogging type	Origin
Kaolin	<10 μm	Physical	Honeywell Fluka, Charlotte, NC, USA
Bentonite	<10 μm	Physical	VWR, part of Avantor, Radnor, PA, USA
Sand Fontainebleau	50 μm–300 μm	Physical	VWR, part of Avantor, Radnor, PA, USA
Calcium Carbonate ($CaCO_3$)	<10 μm	Chemical	Fisher Scientific, Loughborough, Leics, UK

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013

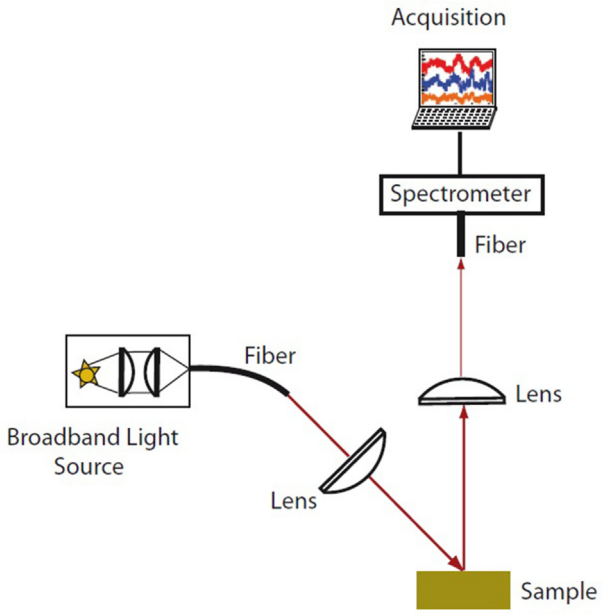

Fig. 1 – Schematic diagram of the spectroscopy system.

extended from 700 nm to 2200 nm with a 1 nm step, and a resolution of 3 nm for the range 700 nm–1000 nm and 10 nm for the range 1000 nm–2500 nm. A constant angle of 40° was maintained between the excitation and collection arms. This angle was chosen to optimise the intensity of the reflected beam and to avoid specular reflection.

For each sample, the intensity of the reflected light ($I(\lambda)$) was measured. Dark current ($I_n(\lambda)$) i.e. signal without light, was recorded from all measured spectra and then subtracted. A white reference (SRS99, Spectralon®) ($I_0(\lambda)$) was measured to standardise spectra and prevent nonlinearities of all the instrumentation components (light source, lens, fibres and spectrometer). From these measurements, a reflectance ($R(\lambda)$) was calculated for each sample, as follows:

$$R(\lambda) = \frac{I(\lambda) - I_n(\lambda)}{I_0(\lambda) - I_n(\lambda)} \quad (1)$$

2.4. Data analysis

2.4.1. Spectral pre-processing

The recorded signal could be distorted by scattering effects due to the presence of particles, of difference sizes, in the measured media. Data pre-processing is usually required to enable the extraction of chemical/physical information and attenuate undesirable signal contributions. In the present work, spectra were first centred by subtraction of the arithmetic mean. Next, commonly used method for spectral pre-processing were applied: Savitzky–Golay (SG) (Savitzky & Golay, 1964) and combination of SG with standard normal variate (SNV) scaling (Barnes, Dhanoa, and Lister (1989)).

SG is a digital filter that uses the linear least squares technique for data smoothing. The smoothing parameters include the size of smoothing window and the polynomial degree. These parameters are selected so that the smoothing process retained the spectral features contained in the original spectra. Applying SNV, the individual mean value is subtracted from each spectrum and the result is divided by the standard deviation over all variables for the given sample. Consequently, scattering effects on the measurement were reduced.

2.4.2. Chemometrics

The main goal was to build multivariate models able to predict the class-type (y response variables), and to estimate the thickness of clogging materials as a function of measured spectra (X block of predictor variables). PLSR and its modification PLS-DA were selected for the modelling of clogging materials.

PLSR (Geladi & Kowalski, 1986) was used to estimate the thickness of the different clogging materials investigated in this study. This regression technique consists in using an appropriate set of latent variables (LVs) to describe the maximum correlation between the predictors and response variables, whereas PLS-DA was applied to discriminate between the different types of clogging materials. This method (Wold et al., 1984) finds the components or LVs, and seeks to optimally discriminate different groups of samples (clogging materials) from their spectra (X block), according to their maximum covariance with a target class defined in the y data block. Once the PLSR and PLS-DA models are built and validated, they can be used to predict the clogging thickness and the class-type of unknown samples, respectively.

Before using the above-mentioned methods, the dataset was divided into two sets of samples: training and independent test sets. The training sets were used to build and internally validate the models and, thus, should be comprised of representative samples. The test sets were left for external validation and were not used to build the models. To implement the PLSR models, the datasets of the different clogging materials were divided into a training and an independent test set with a ratio of 2/3 and 1/3 respectively as detailed in Table 2. This division was made assuring a similar distribution of thickness values in both sets. In the case of PLS-DA model the training sets associated to all four materials were merged to constitute the training set, and the same process was applied for the test set.

Cross-validation is a useful tool for internal model validation because it removes random sample subsets from the training set before building the model to be used to perform sub-validation tests where the samples used for validation were not used to construct the model. Cross-validation is also used to determine the number of latent variables needed to build the PLSR and PLS-DA models. The number of LVs to be included in the model is the result of a compromise between

Table 2 – Number of samples in the training and independent test set for each model.

Model	Number of samples		
	Training set	Test set	Total
PLS-DA (all samples)	61	33	94
PLSR on kaolin	20	11	31
PLSR on bentonite	15	8	23
PLSR on calcium carbonate	20	11	31

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013

maximum efficacy and a minimum number of LVs. In the case of the PLSR models, the leave-one-out cross-validation was selected as a strategy to select the number of LVs. In this process, a single thickness value was used to internally validate the training model built with the remaining thickness values. This was repeated such that each thickness value in the original dataset was used only once as the validation data. For the PLS-DA model, one sample associated to each class (four samples in total) were used for internal validation. This step was repeated so that all samples from the largest class were used for cross-validation.

2.4.3. *Figures of merit*

The following figures of merit have been calculated to evaluate and validate the PLSR models: root mean square error in calibration (RMSEC), root mean square error in cross-validation (RMSECV) and root mean square error in prediction (RMSEP) for PLSR. For the PLS-DA model, classification rates (CR) were calculated to evaluate and validate the model. The CR is the percentage of individuals that are attributed to their actual classes by the model. The confusion matrix was also calculated in PLS-DA for visualising classification results. A confusion matrix of classification indicates how many samples were classified in each class. It carries information about the predicted and actual classifications of samples, with each row showing the instances in an actual class, and each column representing the instances in a predicted class.

All calculations were performed using R software (R Core Team, 2017), using the *rnirs* package (Lesnoff, 2021) for calculation and the *ggplot2* (Wickham, 2009) and *gridExtra* (Auguie, 2017) packages for figures production.

3. Results and discussion

3.1. *Visual observation of absorbance spectra*

The raw spectra from the samples of each studied material at the different clogging thickness are presented in Fig. 2.

Two relevant peaks around 1450 nm and 1950 nm were observed in all spectra. These peaks can be attributed to the first overtone of the O–H stretching vibration and the combination of the O–H stretching and H–O–H bending vibrations of water spectrum respectively (Büning-Pfaue, 2003; Phelan, Barlow, Kelly, Jinguji, & Callis, 1989). No other relevant spectral signatures were observed, indicating that water has a major impact on these spectra. According to Ben-Dor and Banin (1995) and Xie et al. (2015), clay materials such as kaolin and bentonite show absorbance peaks of approximately 1400, 1850, 1900, 2090, 2200 and 2500 nm and calcium carbonates show peaks at 1450, 1650 and 2300 nm. However, the water bands cover these wavelengths, making it impracticable the visually observe the characteristic spectral signatures of clays and calcium carbonates.

A difference in global spectrum intensities was observed for all clogging materials. This difference is a clear indication of scattering effects. These effects were more significant in kaolin and calcium carbonate samples. Kaolin and calcium carbonate structures are scarcely influenced by the presence of water. Therefore, their particles could reflect light, thus engendering strong scattering effects. Conversely, bentonite is a swelling clay (Luckham & Rossi, 1999) and thus, its structural properties significantly change in the presence of water. When bentonite swells, it usually form a very homogenous structure that is less sensitive to scattering effects. Finally, sand in water decant to form a uniform layer that may account for the absence of significant scattering effects in its spectra.

A strong background noise caused by the strong absorption of water around 1450 nm and in the 1800–2200 nm region was observed in all spectra. Since the noise and the scattering effects before observed are usually detrimental for further modelling, the data was pre-processed as discussed in section 2.4.1.

A Savitzky–Golay filter was used to reduce noise. For all samples, the smoothing window was set at 5 nm to keep the original shape of the spectra and a second order polynomial was used for calculation. For calcium carbonate, SNV was applied after the SG filter to reduce the scattering effect. The SNV method was not used on kaolin spectra because most of the scattering effects were corrected after centring the spectra. After pre-processing, the global shapes of the spectra were retained, as only the noise and the scattering effects were reduced (see Fig. 3).

As stated above, visual inspection of those spectra was insufficient to discriminate between the different types of clogging materials due to the significant impact of water on their spectra. Therefore, multivariate analysis methods were required to predict the nature of these clogging materials. Moreover, these methods will also help estimate the clogging thickness of these materials when coupled with the thickness measurements obtained using OCT images.

3.2. *Discrimination between the different types of clogging material using PLS-DA*

As mentioned previously, PLS-DA was performed to discriminate the four clogging materials studied: kaolin, bentonite, calcium carbonate and sand. The classification rate (CR) is used as a quality criterion of the developed model (see section 2.3 for more detailed information about model evaluation and model development). Nine LVs were chosen, giving the highest CR value (92.06%) after internal cross-validation.

Once the optimal number of LVs was selected, the calculation of the PLS-DA model on the entire training set was carried out selecting nine LVs. In Table 3, figures of merit obtained with the training set and in cross-validation are collected. All clogging materials were correctly classified into their corresponding classes with an accuracy rate of over 92% in both calibration and cross-validation. The CR was lower in cross-validation than the one obtained with the training set. As the cross-validation step is performed using a leave-one-out algorithm, it would appear that some specific samples in the training set were not well handled by the model.

Although these results with high CR in training and cross-validation appear encouraging, the true predictive performance of the model requires assessment using an independent test set. Figures of merit from the independent test set confirm the same classification performance (see Table 3) with a CR of 96.97%. The CR of the independent test set is higher than the cross-validation CR indicating that the samples in the independent test set are generally well discriminated by the

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013

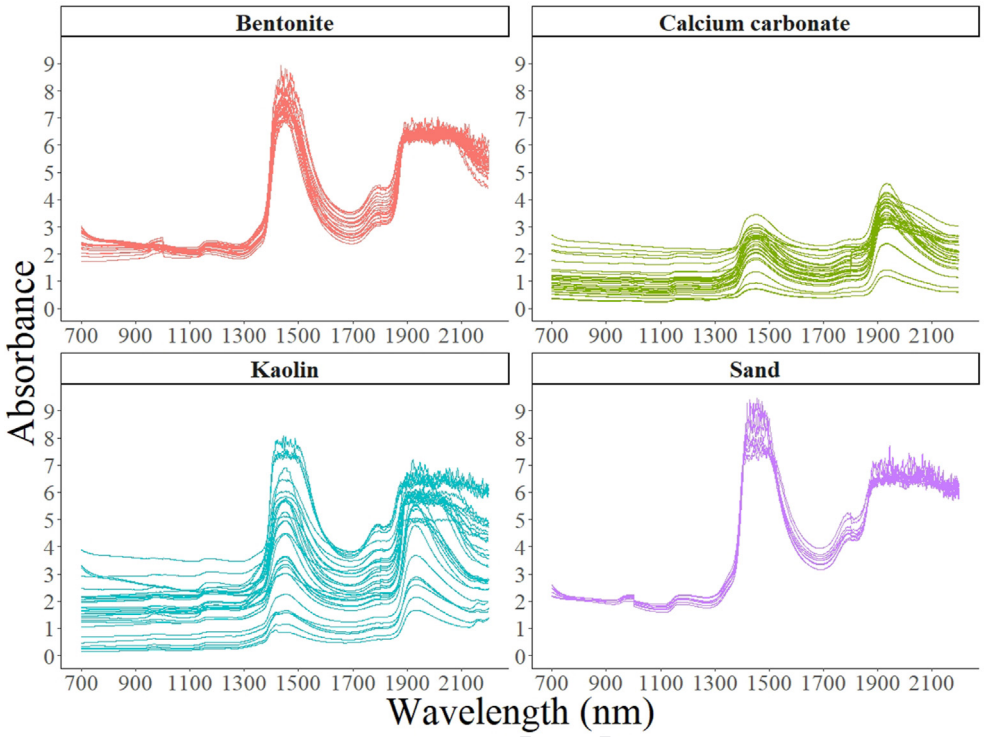

Fig. 2 – Raw absorbance spectra of samples of kaolin, calcium carbonate, bentonite, and sand at different clogging thicknesses.

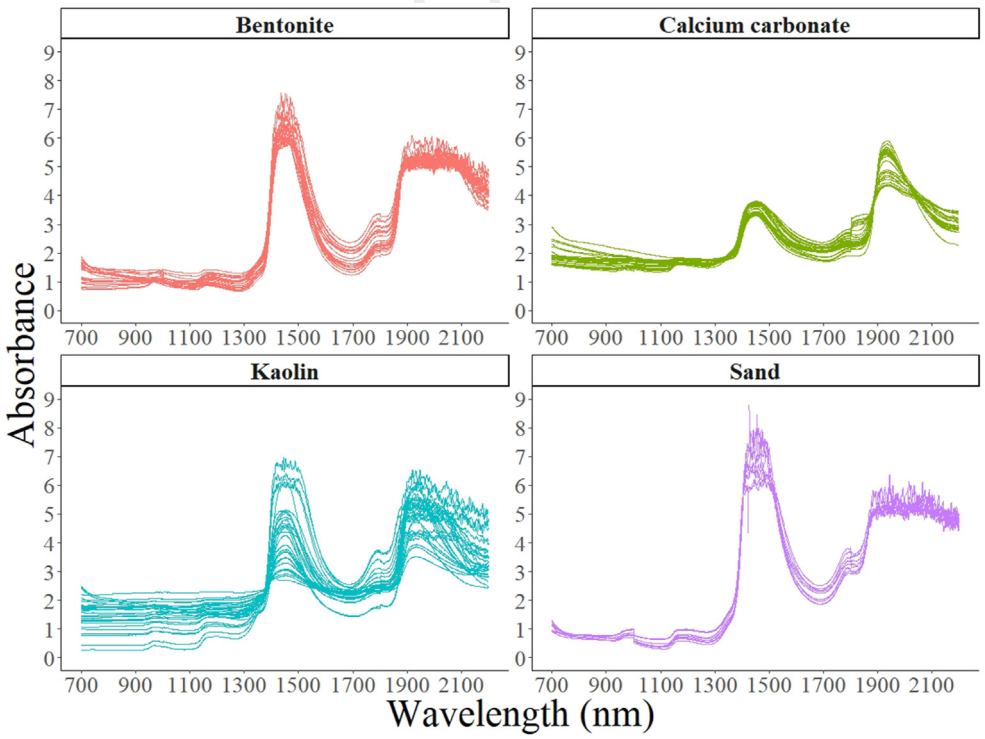

Fig. 3 – Pre-processed absorbance spectra of samples of kaolin, calcium carbonate, bentonite, and sand at different clogging thicknesses.

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013

Table 3 – Classification rates (CR) of PLS-DA model calculations.

LV	Training set CR (%)	Cross-validation CR (%)	Test set CR (%)
9	98.41	92.06	96.97

model compared to some samples in cross-validation. Classification results obtained in the test set are satisfactory compared to those obtained in the literature: Xie et al. (2015) and Lopo et al. (2018) obtained CR of 90% when discriminating clays, sand and calcium carbonate in soil samples. This means that the PLS-DA model can be considered suitable for sample classification for the four studied classes.

Furthermore, it is remarkable that almost all samples are correctly classified in their corresponding classes, as can be identified in the confusion matrix of clogging class prediction (see Table 4).

It can be observed that only one kaolin sample was not well classified, as it is mistakenly attributed to the calcium carbonate class. This confusion could be attributed to the fact that these two classes show similar spectra, especially in the 1700 nm–2200 nm region (see Figs. 2 and 3). No misclassification between kaolin and bentonite clays was observed. Bentonite is known to be a swelling clay and kaolin a more inert clay. Therefore, since the samples contain water, the swelling properties of bentonite appear to make a difference in the spectra and thus in the prediction model. The two clays have a different impact on dripper clogging (Ait-Mouheb et al., 2019; Oliveira et al., 2020), and being able to discriminate between the two is of major interest for clogging studies. Similarly, the ability to discriminate chemical materials such as calcium carbonate from physical materials is of great interest in drip irrigation. Moreover, the presence of water in all samples and its effect on the spectra does not seems to alter the discrimination ability of this model. This aspect is also important for further application on drip irrigation where the equipment always contains water, even when it is no longer in operation (Tarchitzky et al., 2013).

In conclusion, these figures of merit demonstrate that this analysis strategy can be applied to complex multi-class problems related to clogging research. It should be mentioned that the number of samples were limited, and larger datasets would be required to build more robust PLS models. In particular, the number of sand samples was not sufficient to consider the model efficient for classification of sand samples. Nevertheless, the model presented appears to be very efficient in predicting the clogging nature among the four studied materials. These results are very encouraging regarding the potential of Vis/NIR spectroscopy in combination with PLS-DA models to differentiate this type of materials. Therefore, this methodology can be proposed as an effective tool for the comprehensive study of clogging monitoring. Furthermore, this approach could also be useful for the maintenance of drip irrigation systems as it would help to select the most suitable cleaning protocol depending on the clogging type detected (Lamm et al., 2006).

3.3. *Estimation of clogging material thickness with PLSR*

This section shows the results of the PLSR models for estimating the clogging thickness of the studied materials. PLSR method was not applied to sand due to the limited number of available sand samples.

3.3.1. *Thickness measurements*

As explained in section 2.2, the average thickness and its standard deviation were calculated for each sample. The thickness of the clogging materials in the different samples measured with OCT ranging between 80 μm and 1050 μm. The deviation in thickness measurements should be taken account when evaluating the predictive accuracy of PLSR models.

3.3.2. *PLSR results*

Individual PLSR models were developed for each clogging material. The number of LVs chosen and the figures of merit associated to the calibration (training set), cross-validation and prediction (independent test set) are shown in Table 5 (see section 2.4.3. for more details about figures of merit for PLSR).

For the bentonite and calcium carbonate models, the RMSEC and RMSECV were close, so the models show stability. For kaolin, the difference is higher as the data may have outlier spectra and the robustness of the model may be lower. For the

Table 5 – Figures of merit of the calibration and validation steps of the PLSR models developed for thickness estimation of kaolin, bentonite, calcium carbonate and sand deposits.

Class	LVs	Training set	Cross-validation	Test set
		RMSEC[a] (μm)	RMSECV[b] (μm)	RMSEP[c] (μm)
Kaolin	4	186	274	159
Bentonite	2	121	180	132
Calcium carbonate	4	91	154	164

[a] Root mean square error in calibration.
[b] Root mean square error in cross-validation.
[c] Root mean square error in prediction.

Table 4 – Confusion matrix of clogging class prediction (33 individuals).

n = 33		Predicted classes				Well predicted classes (%)
		Kaolin	Bentonite	Calcium carbonate	Sand	
Actual classes	Kaolin	10	0	1	0	90,9%
	Bentonite	0	8	0	0	100%
	Calcium carbonate	0	0	11	0	100%
	Sand	0	0	0	3	100%

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013

three different materials, the number of LVs was low, indicating that the variability between the samples –caused by variations in thickness – is well handled by the models, and that the models were therefore robust.

Once a model was developed, it was applied to the independent test set. The RMSEP values indicate the prediction error of the model. For all samples, the values of RMSEC and RMSEP were close, indicating that the models were stable. In this study, dripper channel depths ranged between 800 μm and 1330 μm (Lequette et al., 2020; Yang et al., 2020), thus the obtained RMSEP values represent between 10 and 21% of the channel height. This means that the current PLSR models can detect and estimate clogging thickness starting at 10% or 21% of clogging depending on the clogging nature and the size of the dripper channel. At such early stages of clogging development, it is nearly impossible to detect the clogging as the discharge of the emitters does not vary (Lequette et al., 2021; Oliveira et al., 2020), and that makes current clogging estimation based on discharge measurement inefficient. Therefore, this method could allow early clogging detection, and thus greatly help in planning maintenance operations.

For all clogging materials, the predicted thickness values obtained with the PLSR models were close to the observed values (see Fig. 4). For kaolin, the deviation between the regression line and the identity line was greater for thicknesses >500 μm. This indicates that kaolin thickness predictions tend to be less accurate for medium and high thicknesses. This may be caused by the low number of samples with high thicknesses in the training set. For bentonite, the prediction seems to be accurate for all thicknesses. However, for this material, the number of spectra is lower, thus reducing the possibility of incorrect prediction. Then, for calcium carbonate, two samples with thicknesses of below 300 μm are far from the identity line, indicating a significant difference between prediction and OCT measurements. It is remarkable that these samples have a high deviation on thickness measurement with OCT. Thus, a high heterogeneity in sample thickness have a negative effect on the thickness predicted using this model.

The ability of the PLSR models to predict the thickness of kaolin, bentonite and calcium carbonate in the presence of water based on Vis/NIR absorbance spectra is encouraging. Increasing the number of samples in the training set would improve the robustness of the models and lead to more accurate thickness estimation. Indeed, increasing the number of samples would reduce the prediction error and thus clogging detection limits. The presence of water does not seem to have a significant effect on thickness prediction, and this confirms the relevance of this method for drip irrigation. Therefore, the use of spectroscopy coupled with multivariate analysis for clogging thickness estimation seems to be a very promising method. However, a larger number of samples is required to obtain better results.

To summarise, the combination of Vis/NIR spectroscopy and chemometric methods has proven to be efficient at detecting and distinguishing the studied materials. However, some further aspects should be taken into consideration to guarantee its feasibility in field applications.

In future works, the physical size of the spectral measurement sensor must be reduced to better fit emitter dimensions. Presently, the spectra acquisition covers an area of

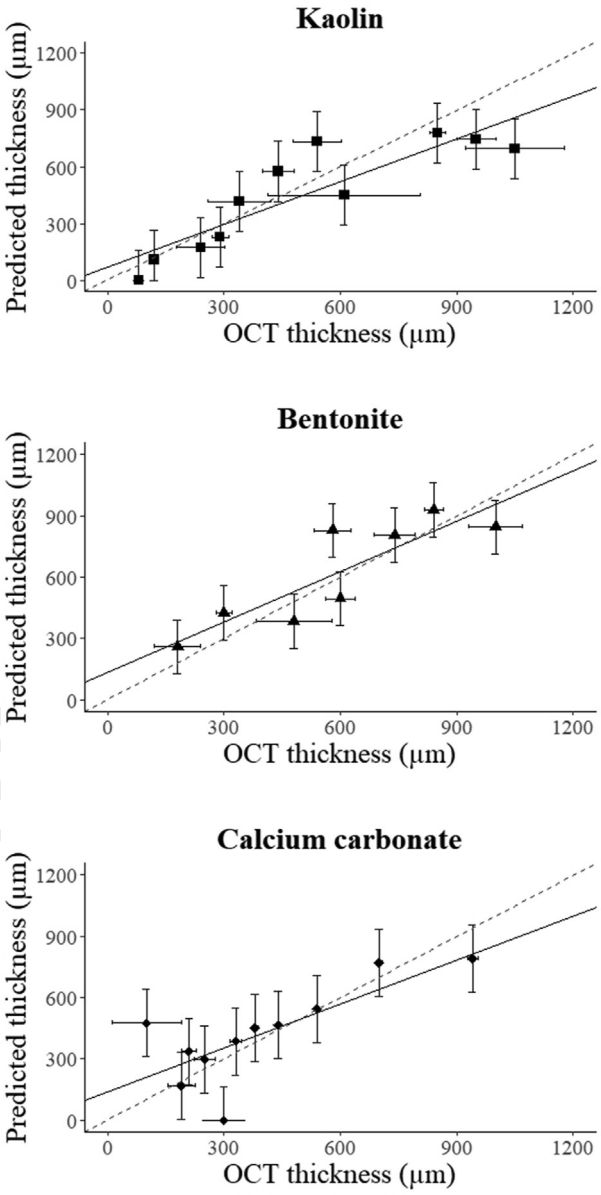

Fig. 4 – Thicknesses measured using OCT vs predicted thicknesses obtained using the PLSR models on kaolin, bentonite, calcium carbonate and sand in the independent test set. Regression line (black) and identity line (red) are shown. Error bars are obtained from the standard deviation in OCT thickness and from the RMSEP values for predicted thickness.

4 mm^2 and is therefore too large for a real application. The location of the measurement should be also considered as clogging is always a heterogeneous phenomenon that is highly dependent on channel design. Moreover, the detection of biological clogging still requires investigation. This cause of clogging is widely observed when treated wastewater is used (Tarchitzky et al., 2013; Rizk et al., 2019; Lequette et al., 2020) and with surface water as the presence of physical and chemical clogging can favour the attachment of biofilm in the dripper.

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013

4. Conclusions

This work confirms the potential of Vis/NIR spectroscopy combined with chemometric methods for clogging studies related to drip irrigation systems. PLSR and PLS-DA methods proved to be useful tools for the detection and the discrimination of four clogging materials associated to physical and chemical clogging. The established PLS-DA model was able to predict with 96.97% accuracy the clogging type from among the four studied materials. Results from PLS models showed that clogging thickness can be estimated when at least 10%–21% of the total depth of the dripper channel is filled with clogging material. These results are promising as they could be used to define a threshold above which the user can be notified that they should launch a maintenance operation. This is a service that no method currently offers.

The conclusions reached in this work justify further efforts in evaluating the feasibility of this approach in the field. Nevertheless, Vis/NIR spectroscopy spectra coupled with the proposed chemometric analysis methods has already proven to be a valuable tool for the monitoring of clogging in drip irrigation systems from a control perspective.

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.

Acknowledgements

Q3 The authors gratefully acknowledge the financial support of INRAE Montpellier and Region Occitanie. We also wish to thank the people working in the optical and biochemical laboratories at INRAE Montpellier.

REFERENCES

Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture, 44*(1), 71–91. https://doi.org/10.1016/j.compag.2004.03.002

Adin, A., & Sacks, M. (1991). Dripper-clogging factors in wastewater irrigation. *Journal of Irrigation and Drainage Engineering, 117*(6), 813–826. https://doi.org/10.1061/(ASCE)0733-9437(1991)117:6(813)

Ait-Mouheb, N., Schillings, J., Al-Muhammad, J., Bendoula, R., Tomas, S., Amielh, M., et al. (2019). Impact of hydrodynamics on clay particle deposition and biofilm development in a labyrinth-channel dripper. *Irrigation Science, 37*(1), 1–10. https://doi.org/10.1007/s00271-018-0595-7

Auguie, B. (2017). *gridExtra: Miscellaneous functions for "grid" graphics*. https://CRAN.R-project.org/package=gridExtra.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy, 43*(5), 772–777. https://doi.org/10.1366/0003702894202201

Ben-Dor, E., & Banin, A. (1995). Near-infrared analysis as rapid method to simultaneously evaluate several soil properties. *Soil Science Society of America Journal, 59*(2), 364–372. https://doi.org/10.2136/sssaj1995.03615995005900020014x

Berho, C., Pouet, M.-C., Bayle, S., Azema, N., & Thomas, O. (2004). Study of UV-vis response of mineral suspensions in water. *Colloids and Surfaces, 248*(1–3), 9–16. https://doi.org/10.1016/j.colsurfa.2004.08.046

Bucks, D. A., Nakayama, F. S., & Gilbert, R. G. (1979). Trickle irrigation water quality and preventive maintenance. *Agricultural Water Management, 2*(2), 149–162. https://doi.org/10.1016/0378-3774(79)90028-3

Büning-Pfaue, H. (2003). Analysis of water in food by near infrared spectroscopy. *Food Chemistry, 82*(1), 107–115. https://doi.org/10.1016/S0308-8146(02)00583-6

Camargo, A. de, Molle, B., Tomas, S., & Frizzone, J. A. (2014). Assessment of clogging effects on lateral hydraulics: Proposing a monitoring and detection protocol. *Irrigation Science, 32*(3), 181–191. https://doi.org/10.1007/s00271-013-0423-z

Camp, C. R., Lamm, F. R., Evans, R. G., & Phene, C. J. (2000). Subsurface drip irrigation: Past, present and future. In *Proceedings of the fourth decennial irrigation symposium* (pp. 363–372).

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

Haas, J., & Mizaikoff, B. (2016). Advances in mid-infrared spectroscopy for chemical analysis. *Annual Review of Analytical Chemistry, 9*(1), 45–68. https://doi.org/10.1146/annurev-anchem-071015-041507

Hills, D. J., Nawar, F. M., & Waller, P. M. (1989). Effects of chemical clogging on drip-tape irrigation uniformity. *Transactions of the ASAE, 32*(4), 1202–1206. https://doi.org/10.13031/2013.31135

Lamm, F. R., Ayars, J. E., & Nakayama, F. S. (2006). *Microirrigation for crop production: Design, operation, and management* (1st ed.). Amsterdam, Boston: Elsevier.

Lequette, K., Ait-Mouheb, N., Adam, N., Muffat-Jeandet, M., Bru-Adan, V., & Wéry, N. (2021). Effects of the chlorination and pressure flushing of drippers fed by reclaimed wastewater on biofouling. *The Science of the Total Environment, 758*, 143598. https://doi.org/10.1016/j.scitotenv.2020.143598

Lequette, K., Ait-Mouheb, N., & Wéry, N. (2020). Hydrodynamic effect on biofouling of milli-labyrinth channel and bacterial communities in drip irrigation systems fed with reclaimed wastewater. *The Science of the Total Environment, 738*, 139778. https://doi.org/10.1016/j.scitotenv.2020.139778

Lesnoff, M. (2021). *R package rnirs: Dimension reduction, regression and discrimination for chemometrics*. https://github.com/mlesnoff/rnirs.

Lopo, M., Teixeira dos Santos, C. A., Páscoa, R. N. M. J., Graça, A. R., & Lopes, J. A. (2018). Near infrared spectroscopy as a tool for intensive mapping of vineyards soil. *Precision Agriculture, 19*(3), 445–462. https://doi.org/10.1007/s11119-017-9529-2

Luckham, P. F., & Rossi, S. (1999). The colloidal and rheological properties of bentonite suspensions. *Advances in Colloid and Interface Science, 82*(1), 43–92. https://doi.org/10.1016/S0001-8686(99)00005-6

Mendret, J., Guigui, C., Schmitz, P., Cabassud, C., & Duru, P. (2007). An optical method for in situ characterization of fouling during filtration. *AIChE Journal, 53*(9), 2265–2274. https://doi.org/10.1002/aic.11257

Mizrahi, G., Wong, K., Lu, X., Kujundzic, E., Greenberg, A. R., & Gilron, J. (2012). Ultrasonic sensor control of flow reversal in RO desalination. Part 2 Mitigation of calcium carbonate

scaling. *Journal of Membrane Science, 419–420*, 9–19. https://doi.org/10.1016/j.memsci.2012.05.026

Niu, W., Liu, L., & Chen, X. (2013). Influence of fine particle size and concentration on the clogging of labyrinth emitters. *Irrigation Science, 31*(4), 545–555. https://doi.org/10.1007/s00271-012-0328-2

Oliveira, F. C., Lavanholi, R., Camargo, A. de, Ait-Mouheb, N., Frizzone, J. A., Tomas, S., et al. (2020). Clogging of drippers caused by suspensions of kaolinite and montmorillonite clays. *Irrigation Science, 38*(1), 65–75. https://doi.org/10.1007/s00271-019-00652-4

Oliveira, F. C., Lavanholi, R., Camargo, A., Frizzone, J. A., Ait-Mouheb, N., Tomas, S., et al. (2017). Influence of concentration and type of clay particles on dripper clogging. *Irrigation & Drainage Systems Engineering, 6*(1), 184–189. https://doi.org/10.4172/2168-9768.1000184

Pavanello, G., Faimali, M., Pittore, M., Mollica, A., Mollica, A., & Mollica, A. (2011). Exploiting a new electrochemical sensor for biofilm monitoring and water treatment optimization. *Water Research, 45*(4), 1651–1658. https://doi.org/10.1016/j.watres.2010.12.003

Pereira, A., Mendes, J., & Melo, L. F. (2008). Using nanovibrations to monitor biofouling. *Biotechnology and Bioengineering, 99*(6), 1407–1415. https://doi.org/10.1002/bit.21696

Phelan, M. K., Barlow, C. H., Kelly, J. J., Jinguji, T. M., & Callis, J. B. (1989). Measurement of caustic and caustic brine solutions by spectroscopic detection of the hydroxide ion in the near-infrared region, 700-1150 nm. *Analytical Chemistry, 61*(13), 1419–1424. https://doi.org/10.1021/ac00188a023

Puig-Bargués, J., Lamm, F. R., Trooien, T. P., & Clark, G. A. (2010). Effect of dripline flushing on subsurface drip irrigation systems. *Transactions of the ASABE, 53*(1), 147–155. https://doi.org/10.13031/2013.29513

Qian, J., Horn, H., Tarchitzky, J., Chen, Y., Katz, S., & Wagner, M. (2017). Water quality and daily temperature cycle affect biofilm formation in drip irrigation devices revealed by optical coherence tomography. *Biofouling, 33*(3), 211–221. https://doi.org/10.1080/08927014.2017.1285017

R Core Team. (2017). *R: A language and environment for statistical computing*. https://www.R-project.org/.

Rizk, N., Ait-Mouheb, N., Bourrié, G., Molle, B., & Roche, N. (2017). Parameters controlling chemical deposits in micro-irrigation with treated wastewater. *Journal of Water Supply: Research & Technology - Aqua, 66*(8), 587–597. https://doi.org/10.2166/aqua.2017.065

Rizk, N., Ait-Mouheb, N., Molle, B., & Roche, N. (2019). Treated wastewater reuse in micro-irrigation: Effect of shear stress on biofilm development kinetics and chemical precipitation. *Environmental Technology*, 1–11. https://doi.org/10.1080/09593330.2019.1625956

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares Pro,cedures. *Analytical Chemistry, 36*(8), 1627–1639.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods, 9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Tarchitzky, J., Rimon, A., Kenig, E., Dosoretz, C. G., & Chen, Y. (2013). Biological and chemical fouling in drip irrigation systems utilizing treated wastewater. *Irrigation Science, 31*(6), 1277–1288. https://doi.org/10.1007/s00271-013-0406-0

Turolla, A., Di Mauro, M., Mezzera, L., Antonelli, M., & Carminati, M. (2019). Development of a miniaturized and selective impedance sensor for real-time slime monitoring in pipes and tanks. *Sensors and Actuators B: Chemical, 281*, 288–295. https://doi.org/10.1016/j.snb.2018.10.107

Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance. *Geoderma, 131*(1–2), 59–75. https://doi.org/10.1016/j.geoderma.2005.03.007

Wang, J., Gong, S., Di, X., Yu, Y., & Zhao, Y. (2013). Impact of drip and level-basin irrigation on growth and yield of winter wheat in the North China Plain. *Irrigation Science, 31*(5), 1025–1037. https://doi.org/10.1007/s00271-012-0384-7

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. https://CRAN.R-project.org/package=gridExtra.

Wold, S., Albano, C., Dunn, W. J., Edlund, U., Esbensen, K., Geladi, P., et al. (1984). Multivariate data analysis in chemistry. In B. R. Kowalski (Ed.), *Chemometrics: Mathematics and statistics in chemistry* (pp. 17–95). Dordrecht: Springer International Publishing. https://doi.org/10.1007/978-94-017-1026-8_2.

Xie, H., Zhao, J., Wang, Q., Sui, Y., Wang, J., Yang, X., et al. (2015). Soil type recognition as improved by genetic algorithm-based variable selection using near infrared spectroscopy and partial least squares discriminant analysis. *Scientific Reports, 5*(1), 1–10. https://doi.org/10.1038/srep10930

Yan, D., Bai, Z., Rowan, M., Gu, L., Shumei, R., & Yang, P. (2009). Biofilm structure and its influence on clogging in drip irrigation emitters distributing reclaimed wastewater. *Journal of Environmental Sciences, 21*(6), 834–841. https://doi.org/10.1016/S1001-0742(08)62349-9

Yang, B., Wang, J., Zhang, Y., Wang, H., Ma, X., & Mo, Y. (2020). Anti-clogging performance optimization for dentiform labyrinth emitters. *Irrigation Science, 38*(3), 275–285. https://doi.org/10.1007/s00271-020-00671-6

Zimmerleiter, R., Leiss-Holzinger, E., Wagner, E., Kober-Rychli, K., Wagner, M., & Brandstetter, M. (2020). Inline biofilm monitoring based on near-infrared spectroscopy with ultracompact spectrometer technology. *NIR News, 31*(7–8), 9–13. https://doi.org/10.1177/096033602097871

Please cite this article as: Petit, J et al., Potential of visible/near infrared spectroscopy coupled with chemometric methods for discriminating and estimating the thickness of clogging in drip-irrigation, *Biosystems Engineering*, https://doi.org/10.1016/j.biosystemseng.2021.07.013