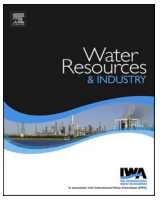
Water Resources and Industry 30 (2023) 100215

Contents lists available at ScienceDirect

Water Resources and Industry

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



Capability evaluation of real-time inline COD detection technique for dynamic water footprint management in the beverage manufacturing industry

Xinyue Cui [a,b], D. Patrick Webb [a,*], Shahin Rahimifard [a]

[a] Centre for SMART, Wolfson School of Mechanical and Manufacturing Engineering, Loughborough University, Leicestershire, LE11 3TU, UK
[b] School of Water, Energy and Environment, Cranfield University, Cranfield, UK

ARTICLE INFO

Keywords:
Dynamic water footprint (WF)
Moving window partial least squares (mwPLS)
UV-VIS
Bland-altman plot
Food and drink manufacturing
Wastewater

ABSTRACT

This paper reports the development of a real-time inline Chemical Oxygen Demand (COD) detection technique in a beverage manufacturing plant in England and the evaluation of its capability for dynamic Water Footprint (WF) management. The inline technique employed Ultraviolet–Visible (UV-VIS) spectroscopy and Moving Window Partial Least Squares (mwPLS), which was then applied to calculating Grey WF for the production activities in the plant, referred to here as WFrt. A traditional offline COD measurement method was also utilised for the Grey WF calculation, to act as the reference method, referred to here as WFtrad. In a method-comparison study (Bland-Altman Plot), the results showed that WFrt detected the order of magnitude variation of WFtrad, and WFtrad was on average between 0.897 and 1.243 times WFrt with no systematic bias. This indicates that WFrt may be used for both short-time frame (minutes to hours) WF monitoring and long-term (weeks to months) analysis of trends and the effect of WF optimisation strategies.

1. Introduction

Global water demand by 2050 is expected to be 1.57 times that of the year 2000, during which time the water demand of manufacturing industry is expected to increase by as much as 400% [1]. Sharply increased water utilisation would lead to a more significant wastewater generation. Although there is limited data on the global amount of industrial wastewater generation and discharge, it can be stated that in the EU, industry is the leading polluter, and that manufacturing is the single largest source of wastewater generation [2]. Nearly 80% of global wastewater from industry and municipal sources is directly discharged into the environment without any treatment, and the situation is even worse in countries with underdeveloped economies [3], resulting in a considerable environmental burden and a hidden danger to human health. As a significant water consumer and wastewater generator, industry should be aware of its impact on the local environment and contribute to water sustainability.

A Water Footprint (WF) serves as an indicator to quantify the direct (operational) and indirect (supply-chain) water consumption for a product, a production process, or a business. It is subdivided into Green WF, Blue WF, and Grey WF [4]. Green WF refers to the rainwater consumption during plant growth. Blue WF covers the volume of fresh surface water or groundwater consumed in human activities. Grey WF is defined as the freshwater volume required for assimilating the pollutant load in the wastewater to meet the local

* Corresponding author.
E-mail addresses: xinyue.cui@cranfield.ac.uk (X. Cui), d.p.webb@lboro.ac.uk (D.P. Webb), s.rahimifard@lboro.ac.uk (S. Rahimifard).

https://doi.org/10.1016/j.wri.2023.100215
Received 8 June 2022; Received in revised form 1 June 2023; Accepted 2 June 2023
Available online 10 June 2023

ambient water standards. The Grey WF attributed to a source is determined by the critical pollutant, which is the pollutant requiring the largest degree of dilution among all pollutants in the wastewater [5]. It is easier and more efficient for an industrial entity to start a water management programme by considering direct WF (operational WF) as they can directly control and manage the WF in their factories, rather than supply chain WF. Operational WF is further divided into two categories: WF directly related to the product production (e.g., product ingredient, cleaning, heating or cooling and steaming) and WF for overhead activities or facilities (e.g., staff kitchen, office, transport, toilet, etc.) supporting product production.

WF has been widely employed as a metric to guide water management efforts in various industry sectors [5–10] However, these WF studies have been carried out based on published databases and historical records. Consequently, there will be a significant time lag between any changes made in response to the analysis, and the water usage, resulting in an extra but avoidable burden on local water resources. One of the main obstacles to reducing this lag is that the data collection process is cumbersome. To address this issue, a Cyber-Physical System (CPS) architecture has been proposed to dynamically assess and monitor water sustainability related to manufacturing production [11]. Here dynamically is taken to mean the ability to measure and respond to changes on both short-term and long-term scales. To achieve this, real-time measurement capability is required. The proposed CPS consists of the physical world (production activities), interfaces (sensors and actuators), the cyber world (Cloud storage & Cloud computing and WF monitoring) and decision-making support (WF optimisation and water management plan). Smart water and wastewater sensors collect the required data directly from production activities and then automatically transmit the data to the Cloud services for further classification, storage, WF calculation and sustainability evaluation. The results from Cloud computing can trigger an alarm if the calculated production WF falls out of the pre-set WF range. The Cloud computing results are also transmitted to a real-time dashboard visualisation to support visibility for management personnel of the water sustainability of operations.

The next step in developing the proposed CPS is to identify and evaluate the feasibility of using inline technologies to act as the interfaces with the production system. For instance, real-time data collection on the concentration of the pollutants required to calculate the Grey WF, such as Biochemical Oxygen Demand (BOD) or its proxy Chemical Oxygen Demand (COD), relies on the availability of an appropriate inline sensor. Current practice is typically to measure COD for purposes of compliance with discharge permissions to a high degree of accuracy using sampling and offline laboratory measurement, which is a time-consuming procedure.

In principle, one can obtain the Grey WF variation with time by regular sampling, offline chemical COD methods and data on flow volumes. However, this is not feasible in a real factory because of the effort involved. The proposed method using an inline COD detection technique and simultaneous flow volume measurement is more viable but may be less accurate. The question being addressed in this paper is whether a particular proposed inline COD technique produces sufficiently reliable results for Grey WF management. To answer this question, comparison is made with traditional offline chemical COD measurement as the reference method, using a Bland-Altman (B-A) Plot. The B-A Plot is a statistical method that was developed to handle this kind of question. Hence in this paper the development of a real-time inline COD detection tool to support Grey WF calculation and subsequent method-comparison study based on the B-A Plot to evaluate its desired capability is reported. This proposed inline sensing technique is expected to detect the order of magnitude variation for short-time frame (minutes to hours) monitoring of WF and lack systematic bias in the measurement for long time frame (weeks to months) analysis of trends and the effect of WF reduction strategies. The evaluation is carried out using samples from a beverage manufacturing plant.

2. Materials and methods

2.1. Overview of the beverage manufacturing plant

The selected beverage manufacturing plant for the study is located in the middle of England, where various products are manufactured and bottled, such as fruit cordials, fruit juice, sparkling drinks, ginger beer etc. Production wastewater of this plant mainly comes from the processing room and production lines. More specifically, raw material washing, cleaning for production hygiene and product quality checking all contribute to the production Grey WF of the plant. The processing room is designed to process raw materials and prepare ready-to-bottle beverages for the production lines. Processes include material washing, infusion, sugar dissolving, syrup mixing, and beverage refrigeration. Two production lines operate in the plant, both of which contain four functional zones (see Fig. 1): de-pallet, bottling, pasteurisation, and packing.

In addition, checkpoints are located in the bottling, pasteurisation and packing zones to guarantee product quality at each stage. The bottling checkpoint procedure is to take at least one bottled beverage out of the production line every 15 min to examine its cap quality and tightness, taste, Brix, pH and temperature. The pasteurisation checkpoint procedure is to take out any broken or defective

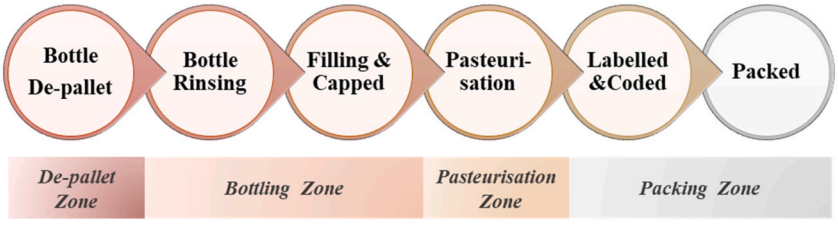


Fig. 1. Production line of the selected beverage manufacturing plant.

bottles after pasteurisation. All bottled products taken out from these two checkpoints are emptied into the drain, which is the main source of high COD concentrations in the plant effluent. The packing checkpoint aims to ensure that the label and package meet the requirements and no beverage waste is generated at this checkpoint.

In the selected plant, most of the production effluent by volume comes from washing and cleaning to keep hygienic conditions for processing and production, including use of Clean-in-Place (CIP). CIP is an automatic method to clean the internal surfaces of processing equipment without dismantling [12], including pre-rinse, caustic soda wash, intermediate rinse, Peracetic Acid (PAA) wash, and final rinse. CIP is utilised for cleaning tanks, mixers, filters, and containers in the processing room and is also employed for the production lines just before the daily production begins or between the production runs of different types of products. After the daily production is completed, a 12-min water rinse is executed for the pipes, tanks, and fillers.

2.2. Research steps

The study was conducted according to the procedures shown in Fig. 2, which comprised three steps: (1) development of the real-time inline COD detection technique, (2) developed technique's application and Grey WF computation, and (3) method-comparison study.

Step 1 aimed to develop a real-time COD detection tool by combining Ultraviolet–Visible (UV-VIS) spectroscopy and Partial Least Squares (PLS) regression. The employed UV-VIS spectroscopy refers to the absorption of radiation in the partial ultraviolet (190 nm–380 nm) and visible regions (380 nm–780 nm) of the electromagnetic spectrum [13]. Although the COD detection tool was developed with spectra obtained from offline samples in this study, regression modelling with UV–Vis spectra is inline capable [14,15], and as such has been reported in many academic studies due to its environmentally friendly, fast, non-destructive and reagent-free nature [16–19]. In this stage, wastewater samples at fixed time intervals were collected from the effluent outlet of the beverage manufacturing plant. The volume of wastewater generated during the periods between the times at which samples were collected was recorded simultaneously. Afterwards, a traditional offline COD measurement and UV-VIS spectrum measurement for each collected sample were carried out. Some samples were reserved for Step 2, and the remaining samples were utilised to establish regression models by PLS, after excluding outliers. Subsequently, a variable selection tool - Moving Window PLS (mwPLS) – was employed to optimise the PLS models. Then, the most optimal model was selected for the next step.

The 2nd step was to obtain the reserved samples' COD from the traditional offline chemical assay and the developed inline technique using the chosen PLS model. And then, values of Grey WF were calculated for each sample using COD values obtained from the traditional offline chemical assay and the developed inline technique and the measured wastewater volumes.

In Step 3, a method-comparison study based on the Bland-Altman (B-A) Plot was conducted to study the agreement between the two sets of values for Grey WF from Step 2. A series of screening statistical tests were required before using the B-A plot method, which were a correlation study, a normality test and a one-sample T-test. The following sections, 2.3 to 2.7, give a more detailed introduction and explanation of the methods used in the above steps.

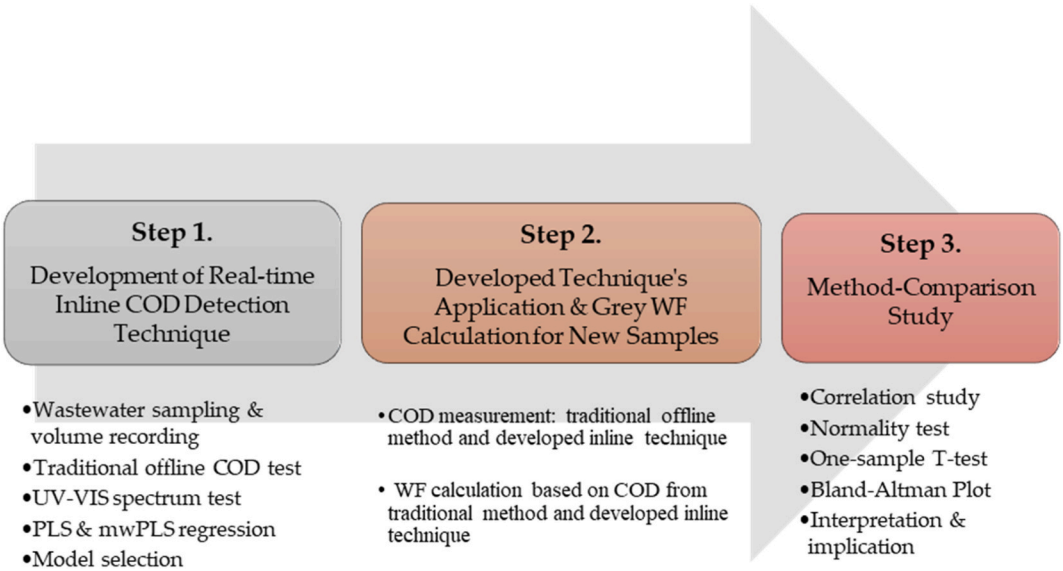


Fig. 2. Research procedures of this study.

2.3. Sampling and experimental procedures

2.3.1. Sample collection and pre-treatment

All wastewater from the production lines and the processing room flows through an outlet into a 'buffer' pit outside the plant. When reaching a pre-set level that triggers the pump operation, the wastewater in the pit is automatically pumped out to a large storage tank onsite. When the tank gets full, a wastewater treatment company comes to collect it. The sampling location was at the pit, and the wastewater was collected from the outflow stream at the outlet mentioned above, using a long probe with a receptacle attached to the end. The collected samples were decanted into 350 ml PET bottles with PP caps. The sample collection and transport followed the quality assurance and quality control procedures in accordance with ISO 5667-14 [20]and ISO 5667-3 [21]. Nine visits were made for sampling on specific production days between November 2020 and July 2021, on which Elderflower Cordial and Elderflower Sparkling drinks were bottled in production line 1 and line 2, respectively. A total of 180 time-series samples were collected, at a time interval of 15 min. A sample preservation treatment following the standard ISO 15705 [22] was applied to stabilise the samples.

Once brought to the laboratory, chloride test strips were utilised to test the concentration range of chloride ions (Cl^-) in the samples, which is required not to exceed 1000 mg/L when using the ISO 15705 ST-COD method [22,23]. The results indicated that of the 180 samples, the Cl^- concentration of 158 were within 1000 mg/L, and 22 were over 1000 mg/L. The over-range 22 samples were diluted by deionised water and retested for Cl^- concentration. The pH of samples was adjusted to be under 2 using sulfuric acid (5 mol/L) before being stored in a refrigerator at a temperature of between 2 °C and 8 °C.

2.3.2. Experimental set-up and procedures

The maximum turnaround time between sample collection and completion of all measurement procedures was 72 h. The traditional offline COD measurement utilised commercial chemical test cuvettes (LCI400, HACH) based on ISO 15705, a thermostat (LT 200, HACH) and a portable VIS spectrophotometer (DR 1900, HACH). Samples with COD exceeding the LCI400 range of 1000 mg/L were diluted by deionised water. The actual COD values were calculated by multiplying the diluted sample COD value by its dilution factor.

The sample spectrum measurement was performed by a double-beam UV-VIS spectrometer (Lambda 35, PerkinElmer) and UV WINLAB software (PerkinElmer). In this study, the spectral range was 190 nm–780 nm, and the light path length was 10 mm. 2 nm was used for the slit width. The data interval of the spectrum was 1 nm, and the scanning speed was 480 nm/min. The solution for baseline correction and reference utilised deionised water. The dilution factor of the diluted samples can be set in UVWINLAB software. If the collected sample was diluted with a sample-to-water ratio of 1:20, the dilution factor of this sample was 21. Then, UVWINLAB software provided the spectrum adjusted with the dilution factor. i.e., the obtained spectrum was for the original sample without dilution.

2.4. Methods for development of real-time COD detection technique

2.4.1. Dataset division

In principle, samples used for PLS regression should cover the full range of content likely to be seen in the plant effluent during regular operation. As it was not feasible within the resource available to the study to carry out sufficient sampling to ensure this was the case, sampling was carried out only on days when Elderflower Cordial and Elderflower Sparkling drinks were being bottled in an attempt to reduce variability in the samples. The collected 180 samples were divided into two groups, Group A and B, for different

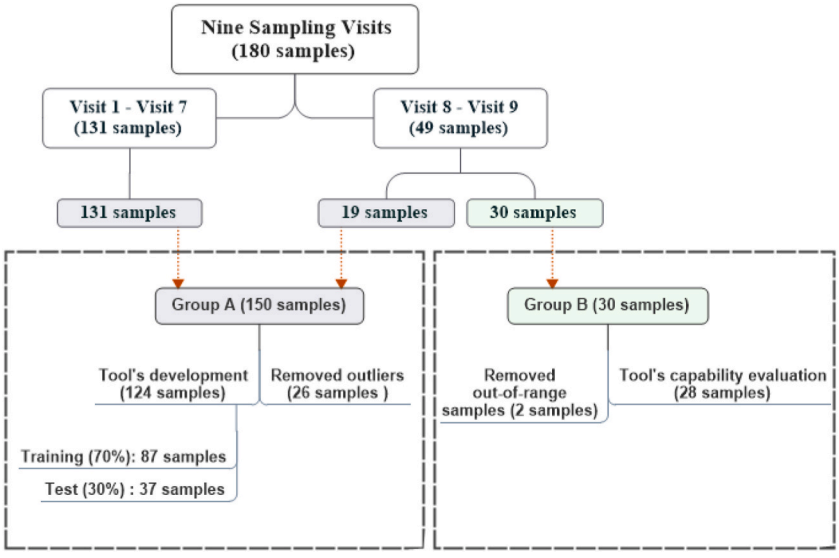


Fig. 3. Dataset division of collected samples for technique development and capability evaluation.

purposes. Samples in Group A were used to develop the inline COD detection technique in Step 1 (as shown in Fig. 2), while samples in Group B were reserved for Grey WF calculation (Step 2) and the method-comparison study (Step 3). The initial plan was to assign all 131 samples from Visit 1 - Visit 7 to Group A and all 49 samples from Visit 8 - Visit 9 to Group B. However, considering the content variation of wastewater samples in each visit, the plan was adjusted, as shown in Fig. 3. More detailed explanations for the content variation are provided in section 3.1.

In addition to all samples in the first seven visits, around 40% of samples in the last two visits, i.e., 19 samples, were added to Group A to develop the tool. In total 150 samples were in Group A and 30 in Group B. In Group A, after removing 26 outliers, the remaining 124 samples were further divided into a training dataset and a test dataset with a ratio of 70%:30%. i.e., 87 samples for model training and 37 samples for model test. The training dataset was used to find the relationship between absorbance (predictors) and COD (responses) and generate a well-fitted PLS model with good predictive capability for new wastewater samples from the same source. In this stage, the number of latent variables (LVs), the employed variables (spectral ranges), and the coefficient matrix of the PLS model were defined. Test datasets were held to assess the model's performance developed by the entire training dataset. In Group B, 2 samples with COD values exceeding the range used for the tool's development were removed from the capability evaluation, so 28 samples were finally used in the method-comparison study.

2.4.2. Partial Least Squares (PLS)

In this study, PLS regression was applied to establish the relationship between COD and absorbance in the UV-VIS spectrum (190 nm–780 nm). PLS is a robust statistical tool correlating high dimensional predictor variables (X) and one or several response variables (Y). It has been widely used in water quality measurement of river water, groundwater, coastal water, and wastewater, usually combined with different spectroscopy techniques, e.g., UV-VIS spectroscopy, Near Infrared (NIR) spectroscopy and fluorescence spectroscopy [24]. To build the calibration, a series of water samples are collected, which should cover the range of variation of water contents from the source. For each sample, COD is measured from the HACH ST-COD test, and the relevant spectrum is obtained from the UV-VIS spectrometer. PLS regression is then applied. This technique does not directly build the relationship between Y and X, but firstly extracts the latent variables (LVs) from Y and X, which are also called Y-scores (U) and X-scores (T), respectively. Then, T is used to predict U, and then predicted U gives the predicted results for the related Y. The extracted T and U should respectively represent the most variation of X and Y, which means that the extracted T and U keep the information and structure of X and Y to the greatest extent. Moreover, the extracted T should have the best explanation of U, i.e., T and U have the highest correlation [25]. The PLS regression model describes the relationship between Y and X as shown in Eq. (1):

$$Y_{(n\times m)} = X_{(n\times p)} \times B_{(p\times m)} + E_{(n\times m)} \tag{1}$$

where, $Y_{(n\times m)}$ is the response matrix with a size of n x m, $X_{(n\times P)}$ is the predictor matrix with a size of n x p, $B_{(p\times m)}$ is the coefficient matrix with a size of p x m, and $E_{(n\times m)}$ is the error matrix with a size of n x m. In this study, each element of Y is the COD value of a sample, each row of X is the ordered set of absorbance measurements in the spectrum of a sample, n is the number of samples utilised in the model, $m = 1$, and p is the number of spectral data points in the wavelength range used for the model. MATLAB software was employed for the PLS regression, in which the wavelength range and number of LVs need to be identified. The full spectrum (190 nm–780 nm) was first used to establish the model (global model) and act as a baseline. Moving Window PLS (mwPLS) was used for the variable selection. The number of LVs was decided according to the Root Mean Square Error of Cross-Validation (RMSEcv). The model performance was evaluated by its Correlation Coefficient of Training (Rtrn), Root Mean Square Error of Training (RMSEtrn), Correlation Coefficient of Test (Rtst) and Root Mean Square Error of Test (RMSEtst).

2.4.3. Moving window PLS (mwPLS)

mwPLS is an efficient variable selection tool for PLS regression [26]. In mwPLS, the wavelength window size needs to be defined, which determines the number of spectral data points contained in the window. Then, a series of PLS models are successively established based on the spectral information in the window that moves from the beginning position to the end position of the whole spectrum [27]. The most informative wavelength range is located according to the mwPLS performance of cross-validation. In this study, mwPLS was performed by means of the algorithm contained in the iToolbox [28] and purpose-written MATLAB code. A total of 11 window sizes were investigated separately, in which the window size increases sequentially from 21 to 121 at an interval of 10. There were 591 spectral points between 190 nm and 780 nm. The window moved from the first point (190 nm) to the last point (780 nm), generating 591 spectral intervals and developing 591 models based on these intervals. In the end, a total of 6501 local mwPLS models were developed. This study used MATLAB code to automatically obtain the fitted local models with RMSEcv better than the global model. Afterwards, the models with Rtrn or Rtst below 0.7 were filtered out. Then the remaining models were sorted by Rtst from largest to smallest. Finally, a model was selected from the shortlist by considering Rtst, RMSEtst and the absolute difference, |RMSEtst-RMSEtrn|, as a percentage of RMSEtrn simultaneously to find the best compromise.

In the mwPLS function provided by iToolbox [28], it is worth noting that for the same window size, the number of spectral points included in the moving window could be different, which is because the window spectral range depends on the position where the moving window is with respect to the minimum or maximum wavelength in the complete spectral data set. The relationship between the Window Size (WS), Window's Position (WP) and the selected spectral range in this study are described as follows. For $1 \le WP \le (WS-1)/2$, the spectral range is from the 1st spectral point to the number $(WP+(WS-1)/2)$ point on the whole spectrum with 591 spectral points. For $(WS-1)/2 < WP \le 591-(WS-1)/2$, the spectral range covers the spectral points from $WP-(WS-1)/2$ to $WP+(WS-1)/2$ on the spectrum. If $591-(WS-1)/2 < WP \le 591$, the spectral range selects from $WP-(WS-1)/2$ to the 591st

spectral point on the spectrum.

2.5. Method for Grey WF accounting

In the usual definition, the 'wastewater' in Grey WF accounting refers to the effluents finally discharged into the environment, which means that the Grey WF can be assumed as zero if the effluents from production are adequately treated before draining into the natural water body. However, the Grey WF calculated in this study is considered to be that of the wastewater directly from production activities before any treatment. It is considered that WF calculated in this way from untreated sewage directly reflects the degree of water pollution created by production activities and consequently encourages manufacturers to find out how to decrease wastewater generation rather than rely on wastewater treatment [11]. In addition, wastewater treatment has an environmental burden associated with it, which would increase the whole system WF if properly accounted for. In this study, the time interval of the Grey WF of production is the same as the time interval of samples, i.e., 15 min. The Grey WF with COD as the critical pollutant can be calculated by Eq. (2) [4].

$$\text{Grey WF (L / 15mins)} = \frac{C_{ww} \times V_{ww} - C_{Fw} \times V_{Fw}}{C_{max} - C_{nat}} = \frac{L_p}{C_{max} - C_{nat}} \quad (2)$$

where C_{ww} (mg/L) is the COD concentration of the production wastewater, V_{ww} (L/15mins) is the production wastewater volume generated in 15 min, C_{Fw} (mg/L) is the COD concentration of the freshwater used in production, V_{Fw} (L/15mins) is the freshwater volume used for production in 15 min. L_p (mg/15mins) is the COD load generated from the production in 15 min. C_{Fw} can be assumed as zero, therefore, $C_{Fw} \times V_{Fw}$ can be ignored, i.e., $L_p = C_{ww} \times V_{ww}$. C_{max} (mg/L) is the maximum acceptable concentration of COD for the local ambient water. According to the discharge limit for COD required by the UK government, $C_{max} = 125$ mg/L was used [29]. C_{nat} (mg/L) is the natural COD concentration in the local ambient water body. Generally, the concentration of COD in unpolluted surface water is less than or equal to 20 mg/L [30]. Thus, $C_{nat} = 20$ mg/L was used for the following Grey WF calculation.

In this study, the volume of wastewater from the outlet, V_{ww} (L/15mins), was not directly available as no volume sensor was installed in the outlet and pit. However, the tank was equipped with a volume monitor sensor, from which the wastewater volume in the tank can be read at every sampling event and after every pump operation. Then, V_{ww} (L/15mins) can be estimated based on the volume changes in the tank, time intervals between pump operation and each sample collection, and the pre-set water volume that triggers pump operation. The estimation premise was that the flow rate was uniform in the studied time unit. The Grey WF of the reserved time-series samples from the last two visits were calculated based on the estimated V_{ww} and COD from the two COD detection methods: the traditional COD measurement and the developed real-time COD detection method, referred to as Traditional Grey WF (WFtrad) and Real-Time Grey WF (WFrt), respectively.

2.6. Methods for the method-comparison study: Bland-Altman (B-A) plot

The traditional methods used for performance evaluation, such as correlation coefficient (R) and coefficient of determination (R squared), are not sufficient in this application [31]. This is because R and R squared only indicate the correlation strength and common proportion of variance between two variables, respectively. Thus, the Bland-Altman (B-A) Plot [32], also known as the Tukey mean-difference plot, was employed for the method-comparison study. In the past decades, the B-A plot has been broadly applied to assess whether a (new) method or piece of equipment can replace a traditional one, especially in the medical area [33]. The classic B-A method plots the difference between the two methods (Y-axis) against the average value of the two methods (X-axis). 95% of Limit of Agreements (LOAs), including the upper LOA and the lower LOA, serve as two reference lines on the Y-axis, between which 95% of differences fall. It is worth mentioning that the difference plotted on Y-axis in B-A Plot can be the absolute difference (WFtrad – WFrt),

Table 1
Summary of statistical parameters required by the B-A Plot following Ludbrook [35] and Bland and Altman [32].

Parameters	Equation	Equation No.
Sample Number (n)	–	–
Degrees of Freedom (df)	df = n-1	(4)
t value (t) of two-tailed t-distribution (p = 0.05, df = n-1)	–	–
Mean Difference ($\overline{d}$)	$\overline{d} = \left(\sum_{k=1}^{n} D_k^{a*}\right) \Big/ n$	(5)
Standard Deviation (s)	$s = \sqrt{\left(\sum_{k=1}^{n} (D_k - \overline{d})^2\right) \Big/ n-1}$	(6)
95% Confidence Interval (CI) of $\overline{d}$	$CI = \overline{d} \pm t \times \sqrt{s^2/n}$	(7)
95% Limits of Agreement (LOAs)	$LOAs = \overline{d} \pm t \times \sqrt{1 + 1/n} \times s$	(8)
95% Tolerance Limits (TLs) with 95% Confidence	$\overline{d} \pm k_7^{b*} \times s$	(9)

[a]* D_k is the difference (absolute difference or percentage difference or ratio) between WFtrad and WFrt, or the difference between the natural logarithm of WFtrad and WFrt of sample k.
[b]* k_7 factor is taken from the Geigy Scientific Tables [35,36], used for obtaining the 95% TLs with 95% confidence.

a ratio (WFtrad/WFrt) or a relative difference, such as percentage difference (Eq. (3)), between the two methods.

$$\text{Percentage difference} = \frac{\text{WFtrad} - \text{WFrt}}{\frac{1}{2}\ (\text{WFtrad} + \text{WFrt})} \times 100\% \tag{3}$$

In this study, either the ratio or percentage difference between WFtrad and WFrt was considered to be more appropriate than the absolute difference for the B-A plot. This approach was chosen because the relative difference can provide a more objective evaluation than the absolute difference for this study where order of magnitudes variation in COD values can be expected. For instance, an absolute difference of 100 L/15mins between 200 L/15mins and 300 L/15mins is more striking than a difference of 100 L/15mins between 20,000 L/15mins and 20,100 L/15mins. Whether using the absolute difference, ratio (WFtrad/WFrt) or percentage difference, statistical screening tests must be applied prior to carrying out a B-A Plot.

To carry out the B-A Plot properly, it is also necessary to check whether the difference (absolute difference, ratio or percentage difference) changes with magnitude (the averages of each measurement pair). In this study, SPSS Statistics was used to conduct a linear regression on the difference and the average, in which the p-value, coefficient and constant of the regression equation were calculated. The p-value is used to test the null hypothesis that "there is no statistically significant difference between 0 and the coefficient of the regression line" by comparing it with the default alpha level of 0.05. When the calculated p-value is larger than 0.05, the null hypothesis is accepted, i.e., the linear relationship between difference and average can be neglected. In this case, the B-A Plot uses the statistical parameters and equations (Eq. (4) – (9)) displayed in Table 1. However, if the null hypothesis is rejected, i.e., the difference has a relationship with magnitude, then logarithmic (log) transformations of both measurements should be subjected to regression analysis instead, and then used in the B-A plot if the null hypothesis can be accepted. In the case where the log transformations fail, a regression approach can be taken to calculate a new LOA. More details of this approach can be found in the literature [34,35].

This study employed a more conservative Eq. (8) [36] suggested by Ludbrook [35] to calculate LOAs for the sample number (n) less than 60, instead of using Eq. (10) proposed by Bland and Altman [32].

$$\text{LOAs} = \text{Mean Difference}\ (\overline{d}) \pm 1.96 \times \text{Standard Deviation}\ (s) \tag{10}$$

For the same reason, this study applies Eq. (9) suggested by Ludbrook [35] to calculate 95% Tolerance Limits (TLs) with 95% confidence. This equation includes the factor k_7 obtained from the Geigy Scientific Tables [35,36], which allows the size of n to be considered. Thus, the 95% TLs with 95% confidence for the population are adopted in this study as the preferred and most conservative parameters to indicate the interchangeability of the two methods.

The most significant advantage of the B-A Plot is simplicity and visualisation. It clearly quantifies the difference distribution range and agreement degree between the two methods within the measured magnitudes. At the beginning of this B-A Plot study, a criterion was defined to determine whether the degree of agreement between WFtrad and WFrt can be accepted: this is that the calculated TLs should imply that WFrt is able to detect at least the order of magnitude WF variation for a short-time frame (minutes to hours) monitoring and produce a relatively accurate estimate of cumulative Grey WF for long time frame (weeks to months) analysis of trends and the effect of WF optimisation strategies.

2.7. Methods for the screening analyses preparatory to the B-A plot

2.7.1. Correlation study

The first assessment was to check how close the correlation was between WFtrad and WFrt. For this purpose, the Pearson Correlation Coefficient (R) for linear relationship strength was utilised, which is defined as Eq. (11) [37].

$$R = \frac{cov(X, Y)}{\sigma_X\ \sigma_Y} \tag{11}$$

where, $cov(X, Y)$ is the covariance between two variables X and Y, σ_X and σ_Y are the standard deviation of variable X and variable Y, respectively. The range of R is between −1 and 1. If the absolute value of R equals 1, $|R| = 1$, a perfect correlation between two variables is indicated, i.e., $|X| = |Y|$. $|R| = 0$ reveals no correlation between the two variables. $0 < |R| < 0.3$, $0.3 < |R| < 0.5$ and $0.5 < |R| < 1$ shows the correlation between the two variables is weak, moderate and strong, respectively [38]. If R indicates a weak correlation between two variables, there is no further need for an agreement study. When R suggests a relatively strong correlation, then performing an agreement study by the B-A Plot method can be considered. In this study, R was calculated between WFtrad and WFrt by SPSS Statistics software. At the same time, the p-value was calculated to test the null hypothesis that the correlation between these two variables is not statistically significant. When the calculated p-value is larger than the chosen alpha level of 0.05, the null hypothesis is accepted, i.e., the correlation is not statistically significant. Otherwise, the null hypothesis is rejected, i.e., the correlation is statistically significant.

2.7.2. Normality test

To perform a B-A Plot, the difference (absolute difference or percentage difference, or ratio) between WFtrad and WFrt needs to be normally distributed. The well-known Shapiro-Wilk test [39] is a practical normality test for small sample sizes, being applicable even for sample sizes of less than 20. It was employed in this study to check the normality of the difference using SPSS Statistics software. The null hypothesis for the test is that the tested data is from a normal distribution. The objective is to accept or reject the null hypothesis by comparing the p-value to the chosen alpha level of 0.05. When the calculated p-value exceeds 0.05, the null hypothesis is

accepted, i.e., the tested data is from a normal distribution. Otherwise, the null hypothesis is rejected. In the case of rejection, the test can be performed again for the difference between the natural logarithmic transformations of WFtrad and WFrt, i.e., ln (WFtrad) and ln (WFrt), and if normality is found, the natural logs would also be used for the one-sample T-test and B-A Plot [34].

2.7.3. One-sample T-test

When the difference values (absolute difference or percentage difference, or ratio) are normally distributed, the one-sample T-test is then performed using SPSS statistics software to examine the statistical significance between the mean value of difference and zero [40,41]. The null hypothesis is that there is no statistical significance between them, which is accepted when the p-value is above 0.05. If the null hypothesis is rejected (p-value <0.05), it is unnecessary to use the B-A Plot method for the agreement study as the T-test will have already indicated a fixed bias between WFtrad and WFrt.

3. Results and discussion

3.1. UV-VIS spectrum measurement

Fig. 4 shows the absorption spectra (190 nm–780 nm) with dilution factors of the collected samples from every visit. It can be clearly seen that the most striking absorbance variation always appears in the range of 190 nm–400 nm in all measured spectra. In general, COD detection utilises the wavelength range from 250 nm to 380 nm, within which organic molecules absorb UV light [42]. The absorbance (Abs.) range and spectrum shape vary from visit to visit, which means these samples contain compounds of different types and concentrations. This is because the production activities in the processing room are flexible, vary daily, and are not always related to the bottling schedule. There could be preparation activities (e.g., material washing, infusion, mixing etc.) for other products, or CIP operation for containers/mixers/tanks used for different solutions or other activities. So that was why, as stated in section 2.4.1,

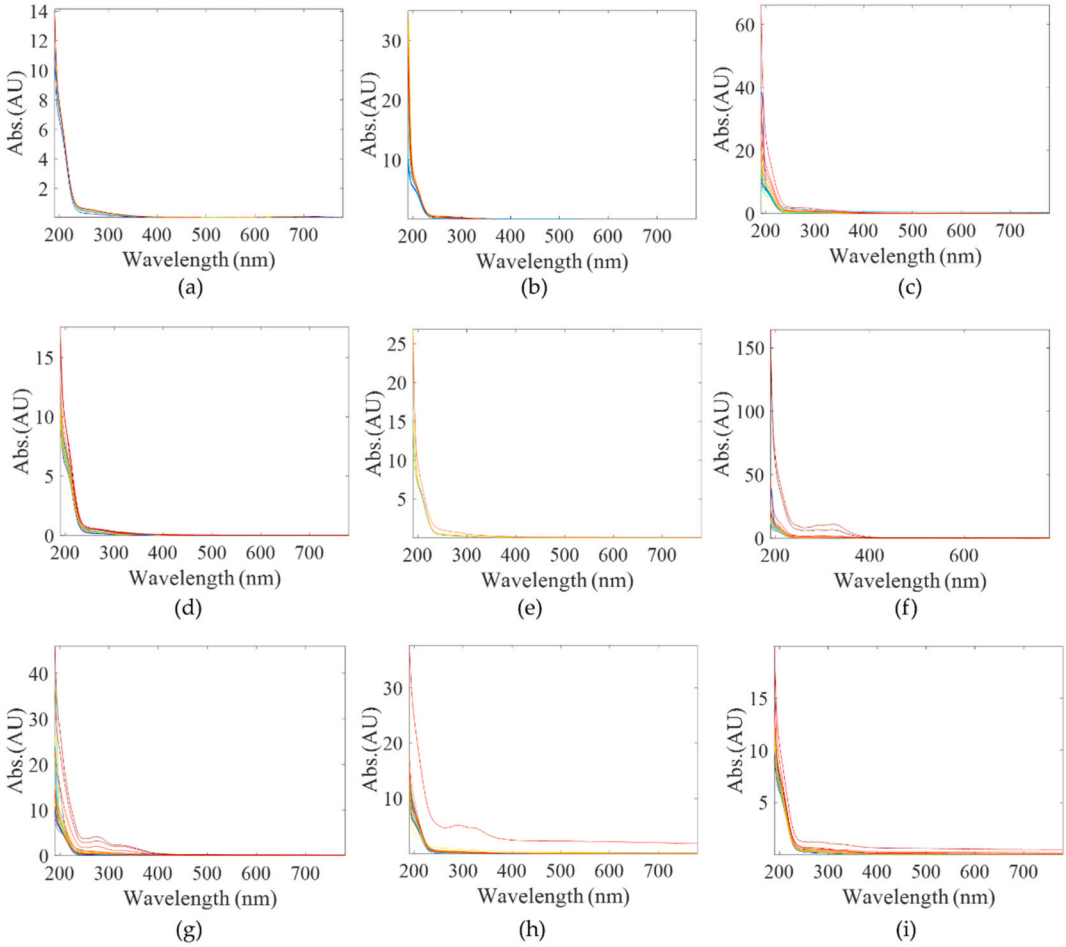


Fig. 4. UV-VIS spectrums of wastewater samples in nine sampling visits: (a) Visit 1 on 05/11/2020; (b) Visit 2 on 04/12/2020; (c) Visit 3 on 18/03/2021; (d) Visit 4 on 15/04/2021; (e) Visit 5 on 11/06/2021; (f) Visit 6 on 18/06/2021; (g) Visit 7 on 21/06/2021; (h) Visit 8 on 02/07/2021; (i) Visit 9 on 09/07/2021.

a portion of samples from Visit 8 and Visit 9 were included in Group A to ensure the developed model covers the information of samples collected in the last two visits.

3.2. COD experiment results

Table 2 summarises the essential statistical characteristics of the COD results of the training and test datasets. They had similar ranges (training: 630 mg/L-10988 mg/L; test: 693 mg/L-10659 mg/L), mean values (training: 4435 mg/L; test: 4330 mg/L) and median values (training: 4221 mg/L; test: 4158 mg/L). This indicates that the test datasets can provide an unbiased and reasonable assessment for the model fitted on the training dataset.

From the box plots (Fig. 5), it is more apparent that the training dataset had a similar distribution to the test dataset. The box plot shows that 75% of COD values fall in the range between the minimum and the third quartile (Q3), i.e., 630 mg/L to 5943 mg/L for training data and 693 mg/L to 5917 mg/L for test data, while the remaining 25% of data falls into the whisker range between Q3 and maximum, i.e., 5943 mg/L to 10988 mg/L for the training dataset and 5917 mg/L to 10659 mg/L for the test dataset. It is evident that for both datasets, the COD values in the whisker range (Q3 to maximum) have a more considerable variance than the rest of the data in the set, i.e., the variance of the upper 25% of the data is as great as that of the other 75% for each set.

3.3. PLS regression results

PLS regression was first carried out on the full spectrum (190 nm–780 nm) to establish a global model as a reference. mwPLS regression was then carried out, from which a total of 1638 models with RMSEcv better than the global model were preliminarily taken out using MATLAB codes. Only 9 out of the 1638 models produce Rtrn and Rtst above 0.7. Finally, the optimal mwPLS model with 5 LVs and a spectral interval from 234 nm to 304 nm was selected based on the principle stated in section 2.4.3. This selected optimal mwPLS was generated when the window (window size = 71) moved to the 80th window location.

As an illustration, Fig. 6 displays the RMSEcv of the 591 mwPLS models generated for a window size of 71. The horizontal dotted line (RMSEcv = 2406 mg/L) represents the RMSEcv of the global model. In this case, there were 154 data points below the dotted line, i.e., 154 mwPLS models' RMSEcv were better than the global model. The study did not directly select the model with the smallest RMSEcv (1825 mg/L) at the 123rd window (277nmnm to 347 nm), but further selection criteria for model stability were applied to avoid overfitting. The effect of overfitting can be seen in Table 3, showing that that the mwPLS model based on the 123rd window had a test performance (Rtst = 0.643 and RMSEtst = 2252 mg/L) far inferior to its training performance (Rtrn = 0.813 and RMSEtrn = 1428 mg/L) and was therefore probably overfitted. For the selected mwPLS, the test performances (Rtst = 0.746 and RMSEtst = 1980 mg/L) were almost as good as that of the training (Rtrn = 0.746 and RMSEtrn = 1630 mg/L), indicating that this model has a very low probability of overfitting and that its performance is relatively stable, i.e., the model performance does not change significantly with different new input data. It can also be seen that the chosen model has greatly improved performance compared to the global model employing the full measurement spectrum. Rtrn (0.274) and Rtst (0.582) increased to 0.746. RMSEtrn and RMSEtst dropped around 31% and 8.5%, to 1630 mg/L and 1980 mg/L, separately. Fig. 7 visualises the fitting performance of the model as a scatter plot.

3.4. Developed technique's application and Grey WF computation

The traditional offline method and the developed real-time inline capable technique were used to measure the COD of the reserved 30 samples (trial) in Group B. The trial group's COD values by the traditional offline method ranged from 924 mg/L to 14,847 mg/L, out of the range of the training dataset (630 mg/L to 10,988 mg/L) and test dataset (693 mg/L to 10,659 mg/L) used in developing the inline technique. This is more apparent in Fig. 8, displaying overlapping histograms of the traditional COD of the three datasets.

Two of the trial samples fall into the range of 13,000 mg/L to 15,000 mg/L, which the training and test ranges do not cover. Accordingly, these two samples were excluded from the method-comparison study in section 3.5. Moreover, only relatively few samples of the training (8 samples) and test (3 samples) sets fall in the range of 8000 mg/L to 11,000 mg/L. Therefore, the accuracy of the developed technique may not be guaranteed for the samples' COD that fall into this range. Table 4 lists the wastewater volume changes (ΔV), COD values obtained from the two methods, and the calculated Grey WF (WFtrad and WFrt).

The time variation of both Grey WF values over the sampling period on the two visit dates is shown in Fig. 9. It can be seen qualitatively that the WFrt can be used to detect the general trend in WFtrad. The rising trend of Grey WF from 13:00 on 02/07/2021 (Fig. 9 (a)) and 12:40 on 09/07/2021 (Fig. 9 (b)) was because production had finished and the water rinse cycle for the pipes, tanks, and fillers started. In this stage, the highly concentrated organic beverage that remained in the fillers and pipes was rinsed out utilising a large volume of water.

Table 2
COD experiment results: statistical characteristics of training and test datasets.

Application	Dataset	Number	Wastewater Samples' COD (mg/L)			
			Min.	Max.	Mean	Median
Tool's Development	Training	87	630	10988	4435	4221
	Test	37	693	10659	4330	4158

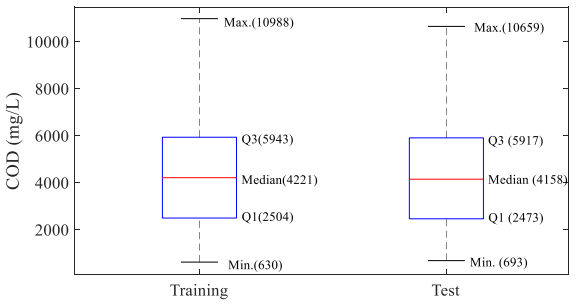


Fig. 5. Box plot: COD results of training and test datasets.

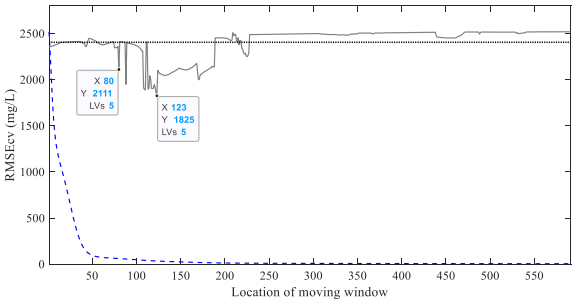


Fig. 6. RMSEcv of mwPLS model at different window locations (window size = 71). The horizontal dotted line represents RMSEcv of the global model, RMSEcv = 2406 mg/L. The blue dash line is the scaled mean spectrum (AU) of the training dataset. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Performance of PLS model (full spectrum) and selected mwPLS model.

Model	Window Location	Spectral Interval (nm)	LVs	RMSEcv (mg/L)	Training		Test	
					Rtrn	RMSEtrn (mg/L)	Rtst	RMSEtst (mg/L)
Global model	–	[190:780]	2	2406	0.274	2356	0.582	2163
mwPLS	123	[277:347]	5	1825	0.813	1428	0.643	2252
mwPLS (selected)	80	[234:304]	5	2111	0.746	1630	0.746	1980

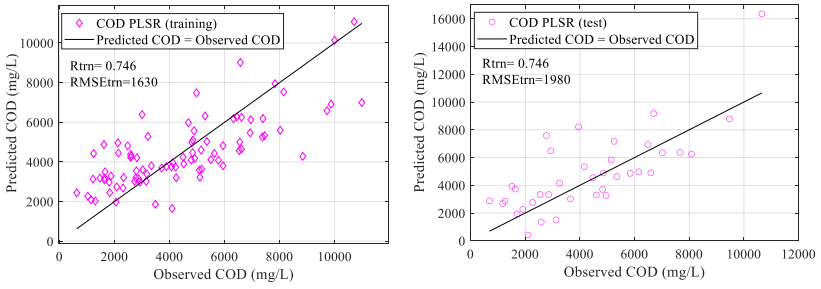


Fig. 7. Scatter Plot: Performance of selected mwPLS: (a) training; (b) test. The solid line is the line of equality.

In Fig. 9 (a), the gaps between WFtrad and WFrt at two consecutive time points (12:00 and 12:15) were unusually large, where WFrt were much lower than the WFtrad, leading to the two methods providing an opposite WF trend between 11:45 and 12:15. It was because the developed inline technique underestimated the COD of 9912 mg/L at 12:00 (sample 2021-07-02-P21) and 8022 mg/L at 12:15 (sample 2021-07-02-P22) to be 3632 mg/L and 4704 mg/L, respectively. Two possible reasons for the discrepancy can be identified. As stated at the beginning of this section, the samples used in the technique development were relatively few in the range of

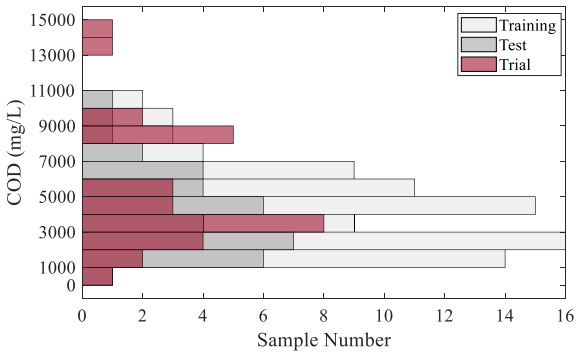


Fig. 8. Histogram of training, test, and trial datasets.

Table 4
Grey WF results based on traditional offline COD measurement method and developed real-time inline COD detection technique.

No.	Sample ID (Year-Month-Day-Number)	Time	Volume Changes (ΔV)	COD (mg/L)		Grey WF (L/15mins)	
				Traditional Method	Real-time Technique	Traditional Method (WFtrad)	Real-time Method (WFrt)
1	2021-07-02-P11	09:30	507	5376	4489	25958	21676
2	2021-07-02-P12	09:45	399	9261	6514	35192	24754
3	2021-07-02-P13	10:00	778	8085	8415	59906	62353
4	2021-07-02-P14	10:15	390	2583	4179	9594	15521
5	2021-07-02-P15	10:30	232	3024	3257	6682	7197
6	2021-07-02-P16	10:45	575	13818	5506	75670	30151
7	2021-07-02-P17	11:00	226	8295	5245	17854	11289
8	2021-07-02-P18	11:15	492	4179	2897	19582	13572
9	2021-07-02-P19	11:30	964	924	1781	8483	16348
10	2021-07-02-P20	11:45	1962	1491	2933	27860	54801
11	2021-07-02-P21	11:55	1051	9912	3632	99214	36354
12	2021-07-02-P22	12:10	1063	8022	4704	81213	47624
13	2021-07-02-P23	12:25	2447	3150	3702	73410	86283
14	2021-07-02-P24	12:40	1375	2877	3184	37675	41694
15	2021-07-02-P25	12:55	842	4746	6380	38058	51160
16	2021-07-02-P26	13:10	1678	4284	4127	68462	65948
17	2021-07-09-P10	09:40	1233	5355	3236	62883	37996
18	2021-07-09-P11	09:55	710	3885	3766	26270	25468
19	2021-07-09-P12	10:10	823	3906	3416	30616	26774
20	2021-07-09-P13	10:25	988	8337	7833	78447	73708
21	2021-07-09-P14	10:40	677	3150	5040	20310	32498
22	2021-07-09-P15	10:55	895	2667	2920	22733	24893
23	2021-07-09-P16	11:10	1698	1764	2041	28526	33008
24	2021-07-09-P17	11:25	1396	3024	3258	40205	43312
25	2021-07-09-P18	11:40	1396	2772	4125	36854	54842
26	2021-07-09-P19	11:55	1665	3822	1281	60606	20316
27	2021-07-09-P20	12:10	1001	5145	3897	49049	37150
28	2021-07-09-P21	12:25	758	3003	3667	21679	26474
29	2021-07-09-P22	12:40	877	8001	4882	66827	40772
30	2021-07-09-P23	12:55	2418	14847	8902	341905	205009

8000 mg/L to 11,000 mg/L. Moreover, the sampling day of 02/07/2021 was during the elderflower harvest season. Wastewater from elderflower infusion, such as infusion tank rinsing around noon, also contributed to the higher COD value. However, this operation was not covered in the previous sampling visits. It can also be seen that WFrt was lower than the WFtrad at 10:45 on 02/07/2021 (sample 2021-07-02-P16) and 12:55 on 09/07/2021 (sample 2021-07-09-P23). These are the samples which, as previously stated, have levels of COD outside of the training range and are excluded from the following comparison study.

3.5. Method-comparison study for these two methods

3.5.1. Screening tests preparatory to Bland and Altman (B-A) plot

- Correlation study

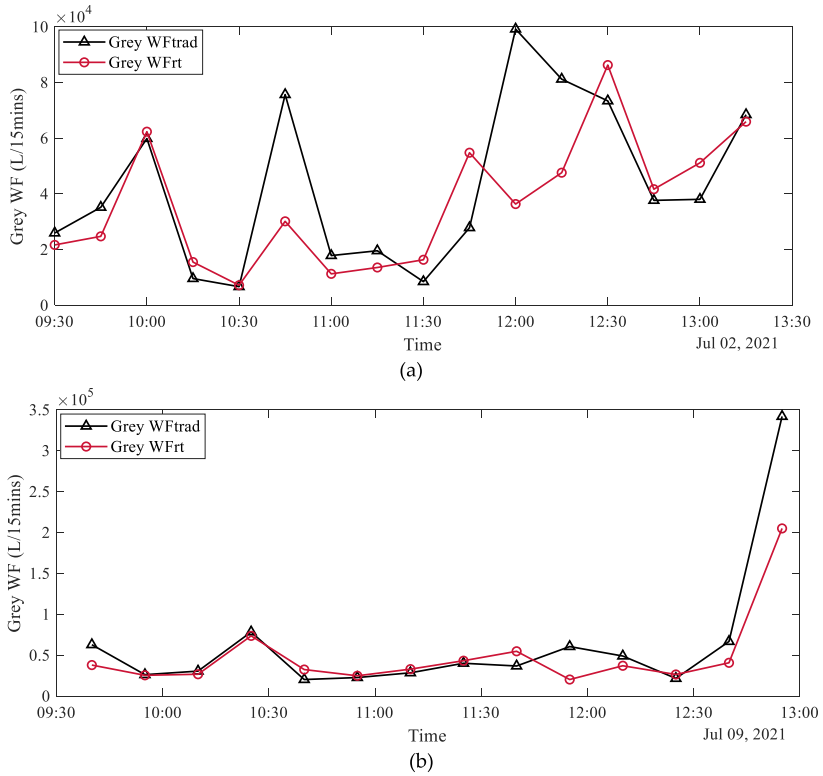


Fig. 9. Grey WF versus sampling time based on two COD detection methods: (a) Grey WF between 09:30 and 13:15 on 02/07/2021; (b) Grey WF between 09:40 and 12:55 on 09/07/2021.

Calculation using SPSS Statistics returned the result of the Pearson correlation study for WFtrad and WFrt, in which $R = 0.661$ and p-value <0.001. As $0.5<R<1$, this indicates that they have a strong correlation. The p-value was much lower than 0.05, so the null hypothesis that "the correlation was not statistically significant" was rejected, i.e., the correlation between WFtrad and WFrt was statistically significant. Thus, the B-A Plot can be employed to explore the agreement between WFtrad and WFrt further.

- Normality test

As highlighted in section 2.6, the ratio (WFtrad/WFrt) or percentage difference between WFtrad and WFrt was preferred for this B-A Plot study. From Fig. 10, it is apparent that the ratio (WFtrad/WFrt) was not normally distributed, but the percentage difference between WFtrad and WFrt was normally distributed, which was further verified by the Shapiro-Wilk test.

The test result for the ratio (WFtrad/WFrt) showed that the p-value was '< 0.001', i.e., far below 0.05. Thus, the null hypothesis of "tested data was from a normal distribution" was rejected. By contrast, the p-value of the percentage difference was 0.565, exceeding 0.05, and then the null hypothesis in this case was accepted. Therefore, the B-A Plot utilised the percentage difference between WFtrad and WFrt, referred to as DIFF. Although the ratio cannot be directly used in the B-A Plot, Eq. (12) enables the conversion between the percentage difference (DIFF) and ratio (WFtrad/WFrt).

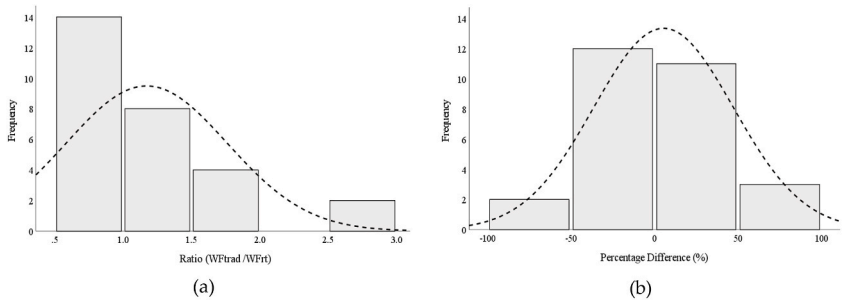


Fig. 10. Histogram plot of the distribution of ratio and percentage difference between WFtrad and WFrt: (a) ratio; (b) percentage difference (DIFF).

$$Ratio = \frac{200 + DIFF\ (\%)}{200 - DIFF(\%)} \tag{12}$$

- One-sample T-test

The results of the one-sample T-test for DIFF showed that the p-value (Sig.) was 0.5, greater than 0.05. So the null hypothesis of "no statistical significance between mean DIFF and zero" was accepted. In this case, a B-A Plot can be carried out further to study the agreement between the two WF methods.

3.5.2. Bland and Altman (B-A) plot

The average of each measurement pair for each trial sample, i.e., <WFrt, WFtrad>, referred to as AVG, was plotted on the X-axis of the B-A Plot, with DIFF of each pair plotted on the Y-axis. SPSS was used to conduct linear regression for the plot to check whether there is a dependence of DIFF on AVG. The best fit line is described by Eq. (13).

$$DIFF = -16.229 + 0.001 \times AVG \tag{13}$$

As the calculated p-value of 0.168 was greater than 0.05, the null hypothesis that "there is no statistically significant difference between 0 and the coefficient (0.001)" was accepted, i.e. the slope of the regression line can be taken to be 0 meaning that there is no systematic bias in DIFF. Therefore, the statistical parameters required by the B-A Plot can be calculated using the equations shown in Table 1. Accordingly, the calculated Mean DIFF ($\overline{d}$), 95% Limits of Agreement (LOAs), 95% Confidence Interval (CI) of $\overline{d}$ and 95% Tolerance Limits (TLs) with 95% confidence are listed in Table 5. In addition, the corresponding ratios are also presented in the same table.

The implications of the B-A Plot (Fig. 11) are discussed in the following paragraphs. To give further insight into the relationship between the two methods, the Y-axis values from the B-A Plot, including the statistical parameters, have been transformed to the ratio (WFtrad/WFrt) using Eq. (12) and plotted in Fig. 12.

In Fig. 11, the first parameter to examine is the Mean DIFF ($\overline{d}$) = 5.4% with 95% CI, highlighted as the shaded areas with dark Grey, from −10.8% to 21.6%. i.e., with 95% confidence that the true value of DIFF is between −10.8% and 21.6%. The corresponding view in Fig. 12 shows that the mean ratio between WFtrad and WFrt was 1.056, with on average the value of the ratio being between 0.897 and 1.243. Importantly the line of equality (DIFF = 0 or Ratio = 1) was within the CI of $\overline{d}$, which also verified "no statistical significance between mean DIFF and zero" from the previous one-sample T-test. If the CI of $\overline{d}$ does not include the line of equality, the conclusion can be drawn that there is a systematic difference (fixed bias) between methods, i.e. the measured value of one method is always higher or lower than that of the other method [31].

The purpose of the B-A Plot is to visualise the relationship between the distribution of DIFF and the 95% LOAs. The 95% LOAs are shown as the blue dash-dot lines in Fig. 11, ranging from −82.0% (Lower LOA) to 92.8% (Upper LOA), within which 95% of values of DIFF should fall, i.e., the DIFF between the two methods is expected to be within −82.0%–92.8% in 95% of cases. For the 28 trial samples in this study, 27 DIFF values, around 96.4%, fell within the range. The corresponding ratios in Fig. 12 show that the corresponding expectation for the ratio WFtrad to WFrt is 0.418–2.731.

As mentioned before, the TLs are the preferred and most conservative parameters to indicate the interchangeability of the two WF methods. Thus, the 95% TLs with 95% confidence for the population have a more extensive range than LOAs. In Fig. 11, the TLs are shown as two Grey dash lines where the covered area is shaded by light Grey, ranging from −102.6% (Lower TL) to 113.4% (Upper TL). It indicates that there is 95% confidence that 95% of the value of DIFF for any future measurement of WFrt and WFtrad will be between −102.6% and 113.4%. This corresponds to WFtrad being 0.322 to 3.621 times WFrt. The implications of this conclusion for the suitability of the developed inline technique for measuring Grey WF are discussed in Section 4.

Table 5
Statistical parameters required by the B-A Plot.

Parameters	Plot Values	Corresponding Ratios
Sample Number (n)	28	–
Degrees of Freedom (df)	27	–
t value (t) of two-tailed t-distribution (p = 0.05, df = n-1)	2.052[a*]	–
k_7 factor	2.582[b*]	–
Mean Difference ($\overline{d}$)	5.4%	1.056
Standard Deviation (s)	41.8%	–
95% Confidence Interval (CI) of $\overline{d}$	[-10.8%, 21.6%]	[0.897, 1.243]
95% Limits of Agreement (LOAs)	[-82.0%, 92.8%]	[0.418, 2.731]
95% Tolerance Limits (TLs) with 95% Confidence	[-102.6%, 113.4%]	[0.322, 3.621]

[a*] t value (p = 0.05, df = 27) = 2.052 was obtained from the two-tailed t-distribution table [43].
[b*] k_7 factor was obtained from the Geigy Scientific Tables [35,36].

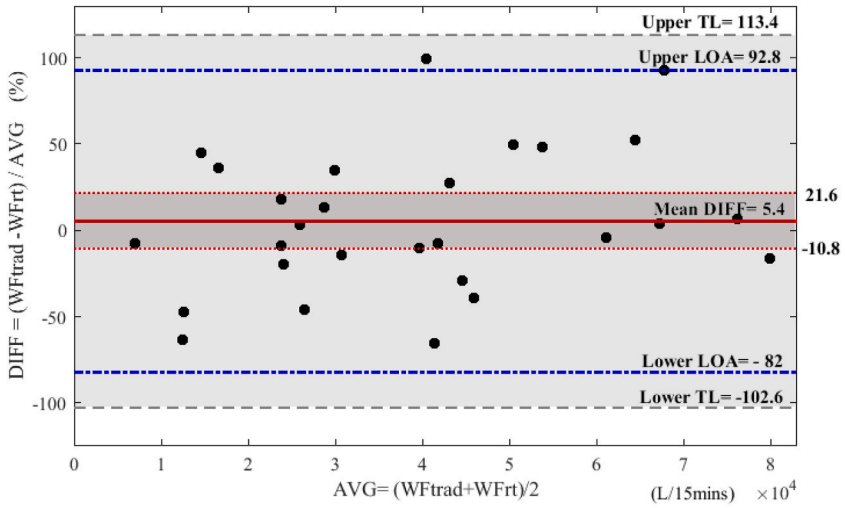


Fig. 11. Bland-Altman plot: Diff against AVG

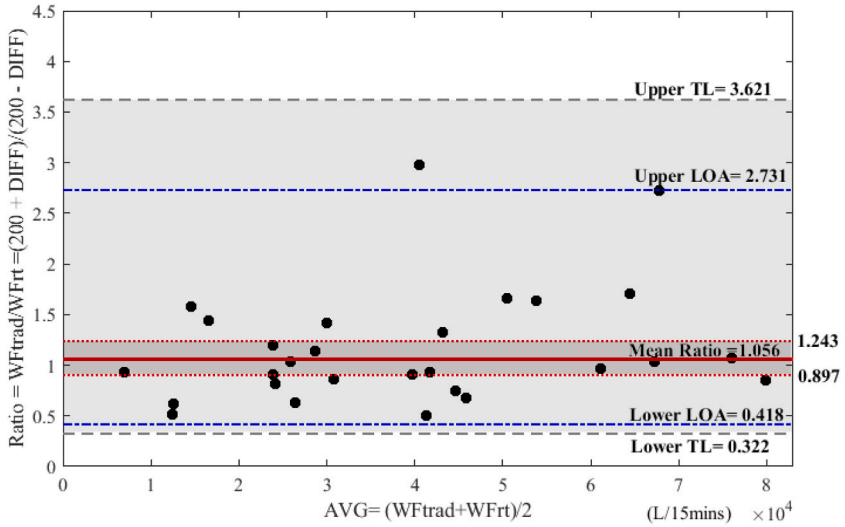


Fig. 12. Bland-Altman Plot: corresponding Ratio against AVG.

4. Implications of the method-comparison study results for use of the developed inline technique

The method-comparison study results indicate that this real-time method is suitable for use in two ways for water management and improving water sustainability: short-term (minutes to hours) monitoring and long-term (weeks to months) optimisation.

4.1. Short-term monitoring

It is asserted that in selecting a technique suitable for WF management, a high level of accuracy can be sacrificed for the advantages of dynamic measurement, and that the ability to detect WF variations within an order of magnitude is sufficient for the short-time frame (minutes to hours) management of WF. The B-A Plot has shown that WF can be determined within a factor of 0.322–3.621 using the proposed inline technique. This is sufficient to detect abnormal changes in WF, such as the Grey WF increasing sharply within a short time, going far out of the pre-determined range of acceptability of WF, or staying at a high level for a long time. Abnormal changes in Grey WF may reflect something going wrong in production. One example is that the conductivity sensors installed in pipes of the CIP system malfunction, which can lead to a premature discharge of caustic soda or PAA, which should have been recycled and reused, into the effluent system. The associated abnormal change of Grey WF would be detected, and staff would be alerted to check, find and address the root causes of the change. Another example in beverage manufacturing plants is when machine malfunctions lead to defective packaged beverages, e.g., the fillers fail to fill the bottle to the pre-set volume, or cappers are damaged or fail to apply caps

appropriately. Defective beverage products will be emptied and drained into the effluent system at the checkpoints, which rapidly and considerably raises the COD concentration of the production wastewater, leading to a sharply increased Grey WF. The frequency of abnormal changes that appears during the Grey WF monitoring represents the malfunction rate of these machines, providing a tool for process and machine health monitoring and water management.

4.2. Long-term optimisation

The lack of systematic bias seen in this developed real-time method (WFrt) indicates that a relatively accurate estimate of cumulative Grey WF can be obtained over the long term. The Grey WF obtained from WFrt will be within a factor of 0.897–1.243 times the real value that could be expected if a programme of frequent wet chemical method measurements of COD were to be possible and undertaken. Thus, WFrt is suitable for a long-time frame (weeks to months) analysis of trends and the effect of WF reduction strategies. It's worth noting that the traditional Grey WF accounting method is not suitable for this due to the limited sampling frequency and sample size, and time-consuming nature of laboratory analysis. Consequently, it is unable to effectively sample the nature and volume of wastewater generated from a manufacturing plant. The wastewater flow generally depends on production schedule, and so may vary irregularly on time scales of hours to minutes, especially where a plant produces various products and has multiple production lines.

Dynamic WF monitoring can enable companies to comprehensively understand water utilisation and wastewater generation in their production. Then, WF can be optimised by the PDCA cycle (Plan–Do–Check–Adjust) [44]. Moreover, the PDCA cycle can be combined with a simulation model of the production facility to test different what-if scenarios, allowing testing of strategies without disturbing the current manufacturing system, significantly reducing the time, cost and risk of the implementation of new strategies [44,45]. WF optimisation strategies should consider reduction of town water usage, wastewater generation reduction and pollutant reduction. Reduction of water usage can start with identifying opportunities for reduction in the water required for processes, for water reuse and for water recycling [46,47]. Examples of water reduction are preventing over-cleaning in the CIP by employing real-time monitoring sensor techniques [48], lowering the flow rate for floor cleaning, avoiding high-pressure water unless necessary, and replacing an outdated machine that consumes overmuch water. Water reuse can involve employing the degraded water from bottle rinse or raw materials washing for other activities with low water quality requirements, such as floor cleaning, truck washing etc. Moreover, water from bottle rinse or other processes can be recovered, treated, and then used for other purposes, such as cooling towers, hot wells, pasteurisers etc. All approaches used in water-saving also lessen wastewater discharge.

In addition to the quantity reduction of water use and wastewater discharge, decreasing pollutants in the wastewater from its source is also critical to Grey WF optimisation. The pollutant source and types vary by industry. For the food and beverage manufacturing industry, pollutants mainly are organic matters from remnant food or beverages on processing equipment or floors and remaining or spent detergent from all kinds of cleaning to maintain hygienic conditions, resulting in a high concentration of BOD and COD. For instance, changing food processing methods or ingredients in a food manufacturing plant may reduce the remaining food in a cooking tank, resulting in less organic matter in the cleaning effluents. Minimising detergent amounts in the cleaning processes, such as minor adjustments for the CIP system in the initial rinse time, detergent solution conductivity, final rinse time, and detergent recovery values, can significantly reduce the water consumption, detergent consumption and effluent discharge [49].

4.3. Research contributions

As pointed out in the introduction, conventional WF studies look at average WF value over a period of years by analysing data from published databases and historical records [6,8,9]. As also pointed out, this approach does not strongly support practical interventions by individual supply-chain actors because of the time lags and effort involved in collecting the data. This research has shown that the proposed CPS system [11] using an optical detection method can support continuous WF monitoring and therefore provide the required practical tool for an active approach to WF reduction.

An additional contribution is related to the fact that in most WF studies [7,8,50], supply-chain WF generally received more attention than operational WF. These studies exemplify the general tendency in the literature to concentrate on supply-chain WF rather than operational WF. This tendency occurs because the former is much higher than the latter, especially for the food and beverage manufacturing industry, where the dominant WF is from the crop-raising phase in agriculture. While such studies are useful for identifying supply chain hot spots for WF generation, they are less helpful in identifying potential operational changes to reduce WF by individual actors and resolving any effects of such changes within a reasonable time frame. The optical detection method verified in this study can support the proposed CPS system [11] to remedy this lack.

In operational WF studies reported in the literature [7,50–52], while the definition of Grey WF used nominally quantifies the impacts caused by the effluents finally discharged into the environment, the assumption is adopted that Grey WF can be neglected on the grounds that the operational effluent would be treated to meet emission standards onsite or by a downstream wastewater treatment plant. However, no matter which wastewater treatment is employed, they all have an invisible WF and burden on the environment due to the energy, materials and water consumption during the treatment processes. Given that avoiding waste is better than treating it, the lack of attention to this invisible environmental burden cannot help industrial actors trace the roots of pollution roots or encourage them to reduce the WF caused by wastewater at its source rather than rely on wastewater treatment. This study addresses this gap by proposing a different view of attributing Grey WF to wastewater at source and so providing data helpful in managing untreated wastewater directly from production-related activities.

4.4. Research limitations and future work

This study likely represents a worst performance of the method for real-time monitoring of WF than might be achievable in practice. This is due to the limitations of the study, i.e., the real-time COD detection tool and its performance in the method-comparison study are limited by the relatively small number of samples used in the PLS regression. This, in turn, is a result of the amount of effort required to manually collect the samples. Better performance could be achieved if more samples were collected to cover as wide a range as possible of production conditions for Elderflower Cordial and Elderflower Sparkling. Future development of this industrial system would see the use of inline in-situ optical sensor devices to measure the required UV-VIS spectra, which could rapidly gather a much larger data set and permit improved regression models for better performance.

The conclusions reached in the present study warrant further and wider investigation of the potential of the method for real-time WF monitoring. For example, the main pollutants and the most critical pollutant of effluent vary by industry. For the beverage industry, the main pollutants are types of organic matter. The Grey WF calculation in this study assumes that COD is the critical pollutant. More water quality parameters related to organic matter, such as Total Organic Carbon (TOC), could also be measured and used in the Grey WF calculation. Then the final Grey WF would be determined by the pollutant with the most significant Grey WF among all pollutants in the wastewater.

5. Conclusion

Sharp global increases in industrial water demand and wastewater generation put an enormous burden on water sustainability. In a previous study, a proposed CPS was described which could enable an industrial actor to dynamically manage water sustainability. The system requires appropriate real-time inline techniques to support data collection for WF monitoring and management. This paper evaluated the capability of a real-time inline capable COD monitoring technique to fulfil this role. The inline tool was developed utilising PLS regression applied to UV-VIS spectroscopy with samples of wastewater obtained from a beverage factory. mwPLS was employed to optimise the PLS model by selecting the most effective and informative spectral interval. The mwPLS model selected has 5 LVs based on a spectral interval of 234 nm–304 nm, of which the Rtrn = 0.746, Rtst = 0.746, RMSEtrn = 1630 mg/L, and RMSEtst = 1980 mg/L. The selected model serving as a COD detection tool was applied to calculate the Grey WF associated with production activities in the beverage manufacturing plant. The traditional offline COD measurement method was also employed for the Grey WF calculation acting as a reference method. Then a method-comparison study employing a Bland and Altman (B-A) Plot was carried out to check the agreement between WFtrad (reference method) and WFrt (developed method). The results showed that WFtrad was between 0.418 and 2.731 times WFrt in 95% of the cases for the 28 samples in the study. Following statistical analysis of the sampling error and variance in the underlying population, it was calculated that there is 95% confidence that 95% of WFtrad was between 0.322 and 3.621 times WFrt for both the current set and underlying population of samples. It was argued that the ability of WFrt to detect the order of magnitude variation in Grey WF is sufficient for short-time frame (minutes to hours) monitoring of WF. i.e., a high level of accuracy can be sacrificed for the advantages of real-time measurement. There was also observed to be no fixed bias between WFtrad and WFrt, and WFtrad was on average between 0.897 and 1.243 times WFrt, indicating a relatively accurate estimate of cumulative Grey WF can be obtained over the long term. It was considered that this is sufficient for a long-time frame (weeks to months) analysis of trends and the effect of WF optimisation strategies.

It is concluded that this proposed inline COD technique is a promising candidate to support the previously proposed CPS for dynamic Grey WF management in the manufacturing industry production. Such a system would provide a practical tool for individual actors in a supply chain to address the impact of water use on the environment. This contrasts with the current approach reported in the literature, which is dependent on databases and surveys and tends to concentrate on supply-chain-related impact, not operational impact. In addition, a modified Grey WF definition has been used whereby pollution can be considered at source in accordance with good sustainability principles.

Data access statement

The experimental data cited in this article is available through the Loughborough University Research Repository (https://repository.lboro.ac.uk/).

Author statement

Xinyue Cui: Methodology, Formal analysis, Investigation, Writing – Original Draft. D. Patrick Webb: Conceptualisation, Writing – Review & Editing, Supervision. Shahin Rahimifard: Supervision, Funding acquisition.

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.

Data availability

Data available through the Loughborough University data repository

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Centre for Innovative Manufacturing in Food [Reference: EP/K030957/1]. The authors thank Belvoir Farm (England) and their staff, who supported data and sample collection in our research, especially Mr Frank Fitzgibbons and Mr Marcin Wilk.

References

[1] OECD, "OECD environmental outlook to 2050,", Dec. 08, 2021. [Online]. Available, https://read.oecd-ilibrary.org/environment/oecd-environmental-outlook-to-2050_9789264122246-en#page1, 2012.
[2] WWAP, "The United Nations world water development report, 2017: wastewater: the untapped resource,", UNESCO (2017).
[3] WWAP, The United Nations World Water Development Report 2021: Valuing Water, UNESCO, Paris, 2021. Dec. 08, 2021. [Online]. Available, https://unesdoc.unesco.org/ark:/48223/pf0000375724.
[4] A.Y. Hoekstra, A.K. Chapagain, M.M. Aldaya, M.M. Mekonnen, The Water Footprint Assessment Manual: Setting the Global Standard, Earthscan, London, 2011.
[5] P.W. Gerbens-Leenes, A.Y. Hoekstra, R. Bosman, "The blue and grey water footprint of construction materials: steel, cement and glass,", Water Resour. Ind. 19 (Jun. 2018) 1–12, https://doi.org/10.1016/j.wri.2017.11.002.
[6] E. Bonamente, F. Scrucca, S. Rinaldi, M.C. Merico, F. Asdrubali, L. Lamastra, "Environmental impact of an Italian wine bottle: Carbon and water footprint assessment,", Sci. Total Environ. 560–561 (Aug. 2016) 274–283, https://doi.org/10.1016/j.scitotenv.2016.04.026.
[7] G. Rivas Ibáñez, J.M. Molina Ruíz, M.I. Román Sánchez, J.L. Casas López, "A corporate water footprint case study: the production of Gazpacho, a chilled vegetable soup,", Water Resour. Ind. 17 (Jun. 2017) 34–42, https://doi.org/10.1016/j.wri.2017.04.001.
[8] E. Owusu-Sekyere, H. Jordaan, H. Chouchane, "Evaluation of water footprint and economic water productivities of dairy products of South Africa,", Ecol. Indicat. 83 (Dec. 2017) 32–40, https://doi.org/10.1016/j.ecolind.2017.07.041.
[9] T. Ratchawat, S. Panyatona, P. Nopchinwong, A. Chidthaisong, S. Chiarakorn, "Carbon and water footprint of Robusta coffee through its production chains in Thailand,", Environ. Dev. Sustain. 22 (3) (Mar. 2018) 2415–2429, https://doi.org/10.1007/s10668-018-0299-4.
[10] G. Skouteris, et al., "Water footprint and water pinch analysis techniques for sustainable water management in the brick-manufacturing industry,", J. Clean. Prod. 172 (Jan. 2018) 786–794, https://doi.org/10.1016/j.jclepro.2017.10.213.
[11] X. Cui, "Cyber-Physical System (CPS) Architecture for Real-Time Water Sustainability Management in Manufacturing Industry," in *Procedia CIRP*, Elsevier, Jan. 2021, pp. 543–548, https://doi.org/10.1016/j.procir.2021.03.074.
[12] M. Walton, "Plant and Equipment: In-Place Cleaning," in *Encyclopedia of Dairy Sciences: Second Edition*, Elsevier, 2011, pp. 283–285, https://doi.org/10.1016/B978-0-12-374407-4.00416-7.
[13] Iso, ISO 20473:2007: optics and photonics - spectral bands, Int. Organiz. Standard. (2007).
[14] Real Tech Inc, "BOD/COD probe BA-X series,", Real Tech. Inc. (2020). https://realtechwater.com/w_p/wp-content/uploads/2020/08/BOD_COD-Probe-BA-X-Specification-Sheet-Real-Tech-1.pdf. (Accessed 2 February 2022).
[15] s::can GmbH, "spectro::lyser V3,", s::can GmbH (2020). https://www.s-can.at/wp_contents/uploads/2021/11/spectrolyser_v3_ww_en-1.pdf. (Accessed 2 February 2022).
[16] D. Carreres-Prieto, J.T. García, F. Cerdán-Cartagena, J. Suardiaz-Muro, "Wastewater quality estimation through spectrophotometry-based statistical models,", Sensors 20 (19) (Oct. 2020) 5631, https://doi.org/10.3390/s20195631.
[17] R.S. Brito, H.M. Pinheiro, F. Ferreira, J.S. Matos, N.D. Lourenço, "In situ UV-Vis spectroscopy to estimate COD and TSS in wastewater drainage systems,", Urban Water J. 11 (4) (May 2014) 261–273, https://doi.org/10.1080/1573062X.2013.783087.
[18] B. Chen, H. Wu, S.F.Y. Li, "Development of variable pathlength UV-vis spectroscopy combined with partial-least-squares regression for wastewater chemical oxygen demand (COD) monitoring,", Talanta 120 (2014) 325–330, https://doi.org/10.1016/j.talanta.2013.12.026.
[19] A. Torres, J.-L. Bertrand-Krajewski, "Partial Least Squares local calibration of a UV–visible spectrometer used for in situ measurements of COD and TSS concentrations in urban drainage systems,", Water Sci. Technol. 57 (4) (Mar. 2008) 581–588, https://doi.org/10.2166/wst.2008.131.
[20] Iso, ISO 5667-14:2014 Water quality — sampling — Part 14: guidance on quality assurance and quality control of environmental water sampling and handling, Int. Organiz. Standard. (2014).
[21] Iso, ISO 5667-3:2018 Water quality — sampling — Part 3: preservation and handling of water samples, Int. Organiz. Standard. (2018).
[22] Iso, ISO 15705:2002 Water quality — determination of the chemical oxygen demand index (ST-COD) — small-scale sealed-tube method, Int. Organiz. Standard. (2002).
[23] "Working Hach, Procedure LCI400 COD - ISO 15705- LCI 400 chemical oxygen demand (COD),", HACH (2019). https://uk.hach.com/cod-cuvette-test-iso-15705-0-1000-mg-l-o/product-downloads?id=26370268931.
[24] P. Khatri, K.K. Gupta, R.K. Gupta, "A review of partial least squares modeling (PLSM) for water quality analysis,", Mod. Earth Syst. Environ. 7 (2) (Jun. 2021) 703–714, https://doi.org/10.1007/s40808-020-00995-4.
[25] R.D. Tobias, "An introduction to partial least squares regression,". http://support.sas.com/techsup/technote/ts509.pdf, 1995.
[26] L.L. Wang, et al., "A selective review and comparison for interval variable selection in spectroscopic modeling,", Chemometr. Intell. Lab. Syst. 172 (2018) 229–240, https://doi.org/10.1016/j.chemolab.2017.11.008, no. September 2017.
[27] J.H. Jiang, R. James, B.H.W. Siesler, Y. Ozaki, "Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data,", Anal. Chem. 74 (14) (2002) 3555–3565, https://doi.org/10.1021/ac011177u.
[28] L. Nørgaard, "iToolbox Manual,", Aug. 12, 2022. [Online]. Available, http://www.models.kvl.dk/sites/default/files/iToolbox_Manual.pdf, April, 2013.
[29] Environmental Agency, "Waste water treatment works: treatment monitoring and compliance limits,", Environ.Agenc. (2019). https://www.gov.uk/government/publications/waste-water-treatment-works-treatment-monitoring-and-compliance-limits/waste-water-treatment-works-treatment-monitoring-and-compliance-limits. (Accessed 27 May 2021).
[30] D.V. Chapman, Water Quality Assessments - A Guide to Use of Biota, Sediments and Water in Environmental Monitoring, second ed., E&FN Spon, London, 1996. Dec. 03, 2021. [Online]. Available, https://apps.who.int/iris/bitstream/handle/10665/41850/0419216006_eng.pdf?sequence=1&i.
[31] D. Giavarina, "Understanding Bland altman analysis,", Biochem. Med. 25 (2) (2015) 141–151, https://doi.org/10.11613/BM.2015.015.
[32] J. Martin Bland, D.G. Altman, "STATISTICAL methods for assessing agreement between two methods of clinical measurement,", Lancet 327 (8476) (1986) 307–310, https://doi.org/10.1016/S0140-6736(86)90837-8.
[33] R. Zaki, A. Bulgiba, R. Ismail, N.A. Ismail, "Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review,", PLoS One 7 (5) (May 2012), e37908 https://doi.org/10.1371/JOURNAL.PONE.0037908.
[34] J.M. Bland, D.G. Altman, "Measuring agreement in method comparison studies,", Stat. Methods Med. Res. 8 (2) (1999) 135–160, https://doi.org/10.1177/096228029900800204.

[35] J. Ludbrook, "Confidence in Altman-Bland plots: a critical review of the method of differences,", Clin. Exp. Pharmacol. Physiol. 37 (2) (Feb. 2010) 143–149, https://doi.org/10.1111/j.1440-1681.2009.05288.x.
[36] C. Lentner, K. Diem, J. Seldrup, Geigy Scientific Tables. Basle : CIBA-GEIGY, 1982.
[37] Karl Pearson, "Note on regression and inheritance in the case of two parents,", Proceed.Royal Soc. London, Royal Soc. (1895) 240–242.
[38] Spss, "SPSS tutorials :Pearson correlation,". https://libguides.library.kent.edu/spss/pearsoncorr#cite_cohen, 2022. (Accessed 8 April 2022).
[39] S. Shapiro, M.B. Wilk, "An analysis of variance test for normality,", Biometrika 52 (3/4) (1965) 591–611.
[40] A. Ross, V.L. Willson, "One-Sample T-Test," in Basic and Advanced Statistical Tests, SensePublishers, Rotterdam, 2017, pp. 9–12, https://doi.org/10.1007/978-94-6351-086-8_2.
[41] SPSS, "SPSS one sample T-test tutorial,", SPSS (2021). https://www.spss-tutorials.com/spss-one-sample-t-test/. (Accessed 30 November 2021).
[42] L. Rieger, G. Langergraber, M. Thomann, N. Fleischmann, H. Siegrist, "Spectral in-situ analysis of NO2, NO3, COD, DOC and TSS in the effluent of a WWTP,", Water Sci. Technol. 50 (11) (2004) 143–152, https://doi.org/10.2166/wst.2004.0682.
[43] MedCalc Software Ltd, "t-distribution table (two-tailed)." https://www.medcalc.org/manual/t-distribution-table.php (accessed December. 25, 2021).
[44] C.N. Johnson, "The benefits of PDCA,", Qual. Prog. 49 (1) (2016) 45.
[45] M. Raza, A.A. Malik, A. Bilberg, "PDCA integrated simulations enable effective deployment of collaborative robots: case of a manufacturing SME,", in: Procedia CIRP, Elsevier B.V., 2021, pp. 1518–1522, https://doi.org/10.1016/j.procir.2021.11.256.
[46] M. Sachidananda, D.P. Webb, S. Rahimifard, "A concept of water usage efficiency to support water reduction in manufacturing industry,", Sustainability 8 (12) (2016) 1–15, https://doi.org/10.3390/su8121222.
[47] C. Davis, E. Rosenblum, Sustainable Industrial Water Use: Perspectives, Incentives, and Tools, IWA Publishing, London, 2021, https://doi.org/10.2166/9781789060676.
[48] J. Escrig, E. Woolley, S. Rangappa, A. Simeone, N.J. Watson, "Clean-in-place monitoring of different food fouling materials using ultrasonic measurements,", Food Control 104 (Oct. 2019) 358–366, https://doi.org/10.1016/j.foodcont.2019.05.013.
[49] L. do C. Pereira, É. Hansen, "Wastewater reuse in the clean in place process of a beverage industry,", ITEGAM- J.Eng. Technol. Ind. Appl.(ITEGAM-JETIA) 6 (21) (2020) https://doi.org/10.5935/2447-0228.20200006.
[50] A.E. Ercin, M.M. Aldaya, A.Y. Hoekstra, "Corporate water footprint accounting and impact assessment: the case of the water footprint of a sugar-containing carbonated beverage,", Water Resour. Manag. 25 (2) (Jan. 2011) 721–741, https://doi.org/10.1007/s11269-010-9723-8.
[51] S.A. Ene, C. Teodosiu, B. Robu, I. Volf, "Water footprint assessment in the winemaking industry: a case study for a Romanian medium size production plant,", J. Clean. Prod. 43 (Mar. 2013) 122–135, https://doi.org/10.1016/j.jclepro.2012.11.051.
[52] E. Bonamente, F. Scrucca, F. Asdrubali, F. Cotana, A. Presciutti, "The water footprint of the wine industry: implementation of an assessment methodology and application to a case study,", Sustainability 7 (9) (Sep. 2015) 12190–12208, https://doi.org/10.3390/su70912190.