Highlights

Intelligent optical fiber probe for rapid beer quality assessment

Felipe Oliveira Barino, João Victor Calderano, Renato Luiz Faraco-Filho, Alexandre Bessa dos Santos

- Novel optical fiber sensor for beer quality assessment;
- Beer brands and lots clustered by linear discriminant analysis and density-based spatial clustering;
- Label-free grouping and detection;
- Real-time out-of-spec sample detection capability.

Intelligent optical fiber probe for rapid beer quality assessment

Felipe Oliveira Barino[a,*], João Victor Calderano[a], Renato Luiz Faraco-Filho[a] and Alexandre Bessa dos Santos[a]

[a] *Universidade Federal de Juiz de Fora, Rua José Lourenço Kelmer, São Pedro, Juiz de Fora, 36036 900, Minas Gerais, Brazil*

ARTICLE INFO

Keywords:
optical fiber sensor
food quality
artificial intelligence
clustering

ABSTRACT

This paper presents a novel optical fiber sensor for beer quality assessment. The sensor is based on a modal Mach-Zehnder interferometer (MZI) incorporating a long period grating (LPG) pair. The MZI is a highly sensitive device capable of measuring the optical properties of beer samples to identify their brands and lots. We demonstrated the performance of the MZI by applying it to three commercial beer brands with different lots. Employing linear discriminant analysis (LDA), Jeffries-Matusita distance, and density-based spatial clustering of applications with noise (DBSCAN), we identified similarities and differences among beer samples. Our results demonstrate the MZI's capability to separate beer brands and lots. Additionally, we detected a discrepancy in one of the lots, indicating a potential quality issue. This method offers a non-destructive, real-time, and low-cost solution for beer quality monitoring that aligns with the principles of Industry 4.0. Our research contributes significantly to the ongoing efforts to enhance the efficiency and effectiveness of quality assessment in the beer industry.

1. Introduction

The history of fermented foods (such as cheese, wine, miso, and beer, for example) is closely intertwined with the evolution of human civilization. These foods were intentionally developed to enhance stability, preservation, and functionality. Interestingly, the discovery of fermentation processes occurred simultaneously across the globe (Tamang et al., 2020). Within the expansive realm of fermented foods, beer stands out as a significant contributor to the global food industry, with its historical origins tracing back to ancient Mesopotamia around 4000 BCE (Damerow, 2012).Since then, manufacturing processes have undergone significant transformations. The rise of industrial and large-scale manufacturing has led to several paradigm shifts.

Currently, we are witnessing a new industrial revolution, enabled by the advent of artificial intelligence and the ability to collect and process massive amounts of data. This revolution, known as Industry 4.0, integrates the data collection capabilities of the Internet of Things (IoT) with the processing power of machine learning and cloud computing into production facilities. The key advantages of this revolution include smart manufacturing, real-time automatic decision making, predictive maintenance, traceability, enhanced automation, and improved quality monitoring (Cañas et al., 2021; Lasi et al., 2014; Zhou et al., 2015).

In the food industry, traditional quality assessment methodologies are often time-consuming, labor-intensive, and destructive. However, the technologies associated with Industry 4.0 hold great promise for overcoming these challenges (Hassoun et al., 2023a,b). The beer industry has begun to embrace Industry 4.0 technologies to elevate the quality and consistency of its products. The utilization of sensors has emerged as a transformative technology in beer production (Benadouda et al., 2023; Viejo and Fuentes, 2020).

To incorporate the new data-driven approaches, food sensing needs further improvement. Most of the research efforts in beer sensing, until 2023, have focused on toxin detection and ethanol quantification (Benadouda et al., 2023). Examples include detecting contamination (Xu et al., 2020), sugar content (Jaywant et al., 2022), and ethanol content (Tsai et al., 2007). These metrics and other food quality analytical techniques coupled with artificial intelligence, have improved the limitations of traditional quality assessment methodologies (Arslan et al., 2021).

For example, Ghasemi-Varnamkhasti et al. (2012) applied dimensionality reduction to bioelectric tongue sensor signals as means to classify beer types and aging process using neural networks, a similar procedure was also employed for beer type classification in (Blanco et al., 2015). An alternative to the complexity of electrical tongues is optical

*Corresponding author

felipebarino@gmail.com (F.O. Barino); felipe.barino@ufjf.br (F.O. Barino); alexandre.bessa@ufjf.br (A.B.d. Santos)

ORCID(s): 0000-0003-3433-0887 (F.O. Barino); 0009-0005-5605-6788 (J.V. Calderano); 0000-0002-4684-0094 (R.L. Faraco-Filho); 0000-0003-2742-7664 (A.B.d. Santos)

sensing, such as spectroscopy (Mignani et al., 2013). Optical sensing improves signal quality and efficiency. To further improve the apparatus and the industrial implementation of such sensing technology, optical fiber sensing is promising. Optical fiber sensors are highly reliable, easy to install, robust, noise-resilient, and small (Gangwar et al., 2023; Méndez and Morse, 2011; Rajan and Iniewski, 2017).

Given these advantages, Tai et al. (2016) proposes a fiber-tip sensor to measure the refractive index and evaporation rate to evaluate alcoholic beverage quality. Plastic optical fibers were also used to quantify alcohol in beverages, offering a cost-effective optical fiber sensor (Yunus et al., 2023). These examples illustrate the versatility of optical fiber sensor technology. Among the several fiber sensing technologies, the LPG can be highlighted as an optimal chemical sensing platform. Several sensitive materials can be introduced into the fiber to adapt the sensor to a wide range of measurands and chemicals, thus forming a versatile platform for chemical sensing (Esposito, 2021; Cai et al., 2023). Additionally, the sensor can be tailored to further improve its sensitivity to the external media (Colaco et al., 2016; Navruz et al., 2018).

LPGs have also been used to assess the quality of beverages. An LPG sensor was able to discriminate between different types of beverages based on the change in refractive index of the sensitive film (Korposh et al., 2014). And a modal Mach-Zehnder interferometer (MZI) has been used to measure beer bitterness by refractive index change between known and unknown beer samples (Filho et al., 2024). These sensors focus on specific beverage aspects (measurands) such as alcohol, bitterness, or sugar. Such metrics are indeed crucial for quality control and ensuring product reproducibility. However, in the production environment, a simple question must be answered swiftly: "Is this beer conformant?" Analyzing individual metrics and comparing them to standards can be time-consuming, potentially leading to material loss.

This question can be addressed more efficiently through artificial intelligence, either by fusing data from multiple sensors or by using a single multi-parameter sensor coupled with AI. This paper explores the latter approach, proposing the use of a twin-LPG modal Mach-Zehnder interferometer (MZI), similar to the one used by Filho et al. (2024). The interferometric fringes are highly sensitive and provide several features to be analyzed. The on-line self-calibration setup methodology presented by Filho et al. (2024) showed the possibility of using such sensitivity to address beer bitterness, considering a single resonant dip. However, to this date the rest of the interferometric pattern remained unused. Note that the previous work considered only one resonant wavelength and a direct correlation to bitterness. In this work we improve robustness by incorporating more interferometric information to cluster commercial beer brands and lots. Unlike previous methods, our approach simplifies the experimental apparatus. Our proposal uses a single fiber sensor to address brand conformity and determine if a sample is similar to the other samples of the same brand. This approach needs a single compact probe that can be far apart from optoelectronics, due to the low attenuation of optical fibers. This makes our method efficient and easy to implement, aligning with the principles of Industry 4.0. We believe that our research contributes significantly to the ongoing efforts to enhance the efficiency and effectiveness of quality assessment in the beer industry.

2. Principle of operation

LPGs and MZIs are two important devices in optical fiber sensing technology. An LPG is manufactured within the optical fiber by varying its refractive index and/or geometry (Bock et al., 2007; Vengsarkar et al., 1996; Yin et al., 2014). It acts as an optical filter, attenuating part of the light at a given wavelength. This attenuation is caused by optical scattering through the optical fiber's cladding. The periodic structure of LPGs causes the light guided inside the core to couple to cladding modes at specific wavelengths given by:

$$\lambda_{res}^{m} = \left(n_{eff,co} - n_{eff,cl}^{m} \right) \Lambda \tag{1}$$

where λ_{res}^{m} is the m-th wavelength that light was coupled (resonant wavelength), $n_{eff,co}$ and $n_{eff,cl}$ are the core and cladding mode's effective refractive index, m denotes the order of the cladding mode, and Λ is the grating period (generally hundreds of microns for LPGs).

Since the attenuation around λ_{res}^{m} is caused by scattering at the cladding and surrounding media interface, the refractive index of the surrounding media affects the scattering characteristic. Furthermore, the $n_{eff,cl}^{m}$ is also affected by the surrounding refractive index (SRI), inducing a change in λ_{res} (Shu et al., 2002):

$$\frac{\lambda_{res}}{n_{sur}} = -\lambda_{res} \frac{\frac{\lambda_{res}}{\Lambda}}{n_{eff,co} - n^m_{eff,cl}} \frac{u_m^2 \lambda_{res}^3 n_{sur}}{8\pi r_{cl}^3 \left(n_{eff,co} - n^m_{eff,cl}\right) \left(n_{cl}^2 - n_{sur}^3\right)^{\frac{3}{2}}} \tag{2}$$

where, u_m is the zeroth-order Bessel function m-th root, $r_c l$ and $n_c l$ are the cladding radius and refractive index, respectively. And for that matter, LPG-based SRI probe has been extensively reported in the literature (Chen et al., 2016; Chiang et al., 2000; Gan et al., 2022; Li et al., 2011; Liu et al., 2023; Qi et al., 2014; Shen et al., 2017; Tan et al., 2014).

An MZI can have several topologies, but essentially, it is an optical device that splits light into two and recombines them. When light is recombined, an interference pattern emerges due to the optical path difference of different wavelengths. Consequently, an even-so-small change in the optical path changes the interference pattern dramatically.

For an MZI based on two identical LPGs, the interference occurs between the core and cladding modes. One LPG couples light to cladding modes and the second LPG recouples the cladding modes to the core, as shown in Figure 1. The difference between the core and cladding effective refractive indices induces an optical path difference, with phase difference given by (Dong et al.; Lee and Nishii, 1999):

$$\theta = \frac{2\pi}{\lambda} \left(n_{eff,co} - n^m_{eff,cl}\right) L \tag{3}$$

where λ is the wavelength and L is the interference length. Discarding the fiber dispersion, the distance between two adjacent minimum transmission dips at the LPG-based MZI fringes is given by:

$$\Delta\lambda = \frac{\lambda^2}{L \left(n_{eff,co} - n^m_{eff,cl}\right)} \tag{4}$$

Hence, SRI changes the impact of the fringe pattern in both position and spacing. Consequently, the LPG-based MZI results in a highly sensitive and information-rich sensor due to the fringe variation and spectral complexity (en Fan et al., 2011; Vasconcelos et al., 2019). These advantages can also be extrapolated to other in-fiber MZI topologies, such as LPG-bitaper (Xue et al., 2013) and LPG-taper (Filho et al., 2023).

The interaction between the beer and the evanescent field caused by mode coupling is the core principle of the proposed sensor. Variations in the beer's composition, including alcohol content, sugar content, and bitterness, for example, alter the light propagation due to the different SRIs of these parameters' combination, causing fringe variation that can be mapped to the sample characteristics.

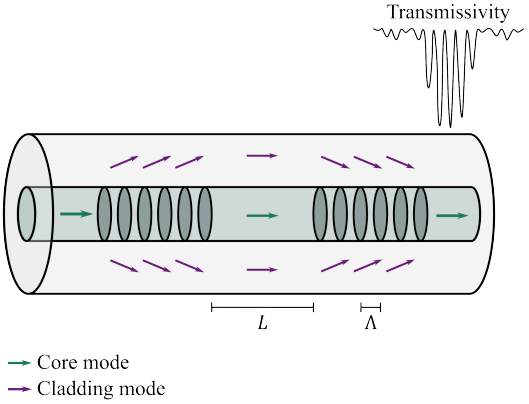

Figure 1: Schematic of an LPG-based MZI.

3. Methods

We manufactured the LPG-based MZI sensor, measured the sensor transmission spectrum for eight different beer samples and processed the spectrum to group the beer samples. The optical setup for acquiring the signals consisted of a broadband ELED source from OptoLink, the proposed optical probe, and a ThorLabs OSA203 Fourier transform optical spectrum analyzer (OSA).

The optical probe was manufactured in-house using two micro-tapered long-period fiber gratings (MT-LPGs) spaced by $L = 13 \pm 1$mm. The two gratings were created by periodic tapering a SMF 28 optical fiber by means of arc discharges. We employed the same manufacturing setup and equipment as described by (Filho et al., 2023, 2024), using a modified fusion splicing machine coupled to a controllable linear stage. The LPGs were manufactured with 20 micro-tapers spaced by $500 \pm 1\mu$m (10 mm total length), using the same arc intensity and arc time. However, due to variability in the splicing machine's arc and errors in the translation stage, no two sensors are exactly alike. Consequently, the interference pattern observed at the MZI transmission is affected.

Spectra of the manufactured sensor are shown in Figure 2. The optical fiber probe's spectrum in both water and beer can be seen. Note the interference pattern obtained for the MZI sensor, with high-visibility fringes at the 1480 nm to 1580 nm spectral portion. Although the fringes are not as clear as the ones shown in Figure 1 due to manufacturing variability, one can see the spectra shifts significantly with changes in the SRI, as sample changes from water to beer. For further processing, not all spectral data was considered, so we focused on the high visibility portion of the spectrum. To compress even more information, due to fringe sensitivity to refractive index, we extracted the resonance dips positions, as illustrated by the inset (zoom in at the 1450-1600 nm).

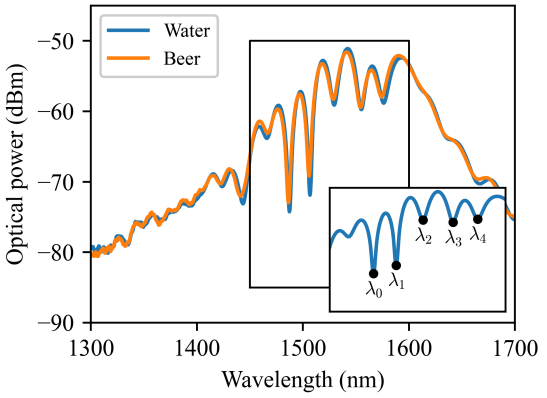

Figure 2: MZI spectra for different surrounding. Inset is a closeup on the resonant dips considered in this work.

To acquire information from beer samples (namely, MZI's transmission spectrum), we filled the probe surroundings with the beer under test. A water-detergent mixture was used to clean the probe between measurements. To accommodate the liquid samples, we used a watch glass fixated at a custom 3d-printed holder. The optical fiber sensor was glued to the watch glass using two component epoxy glue to ensure minimal strain variation. The sensor fixation and sample management scheme are shown in Figure 3.

The beer samples used in this work were all Pilsen, from different brands of the same brewing company. We tested three different brands and three to four different manufacturing lots (depending on the brand). For simplicity, we named the samples using a letter followed by a number. The letter indicates the beer brand, whereas the number represents the lot. For example: sample A1 is beer brand A, lot 1.

The acquired spectra were processed to group beer brand and lot. The spectral processing methodology involved spectral feature extraction followed by dimensionality reduction. The spectral feature extraction was performed by extracting the position of most varying and relevant fringes in the transmission spectrum. These were the dips roughly at 1480 nm, 1505 nm, 1525 nm, 1550 nm, and 1580 nm. To extract the dip position, we fitted a modified Lorentzian function to each fringe dip:

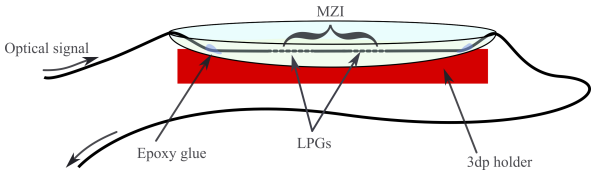

Figure 3: Sensor fixation and sample management scheme.

$$T_{Lorentzian}(\lambda) = -a\left(1 + \left(\frac{\lambda - \lambda_{res}}{\frac{w}{s\sqrt{\left|\frac{a}{3}-1\right|}}}\right)^2\right)^{-1} - b \tag{5}$$

where b represents the LPG's insertion loss, and the resonant dip's parameters are depth (a), center (λ), width (w), and bias (b). The fringe centers were named as $\lambda_i, i \in \{0, 1, \dots 4\}$. The inset in Figure 2 shows the detected dips, thus their x-axis value is equal to λ_i.

Dimensionality reduction was carried by linear discriminant analysis (LDA), as previously reported for beer classification (Blanco et al., 2015; Ghasemi-Varnamkhasti et al., 2012). LDA has been used for decades to promote data separation and classification, including on spectral data (Cao et al., 2023; Mallet et al., 1996). And has also been coupled to clustering algorithms for sample characterization (Novák et al., 2017). In this work we used LDA as a dimensionality reduction tool, whereby finding the mean vector for each sample and projecting the original data onto a new subspace one can effectively reduce the number of features while preserving the essential information for class discrimination. This is achieved by choosing the subspace that maximizes the separation between different classes, as described in (Hastie et al., 2009).

The dataset used in this work was constructed from these spectral measurements. A total of 3673 data points were acquired from 8 different beer samples. These data points were obtained by 8 measurements for each sample, each lasting roughly 1 minute at a 1 Hz OSA sampling rate (60 spectra per sample per repetition). Hence, eight samples and eight replicates, i.e., a measurement consisted of placing a beer sample (B2, for example) into the glass container shown in Figure 3, collecting the MZI spectra for about 1 minute, removing the sample, cleaning the sensor with a water-detergent mixture, and rinsing the sensor with water. This procedure was repeated eight times per sample for all eight beer samples (A, B1-3, and C1-4). After data collection, the dataset consisted of 3673 observations, each with 5 features corresponding to the five resonant dips. The following steps were conducted using sklearn API (Buitinck et al., 2013). The five wavelengths extracted by Lorentzian fitting were standardized then we performed dimensionality reduction using the LDA to reduce the five resonant dips values to two principal components.

To investigate sample separation between two beer brands, we used a logistic regression model (DeMaris, 1995). The model is suitable due to the simplicity and linear separability of this two-brand analysis. This step was taken as proof of concept and to show that although beer samples are similar, compared to water samples, two different brands can be identified. Then, we moved towards multi-lot identification. Note that in this paper, machine learning models and techniques were merely tools, and well-stablished methods are employed, so we won't focus on their operating principles and mathematics. Moreover, our focus is to validate the optical probe technology as a transducer for beer quality assessment. For that matter, we consider that simply obtaining separable enough data from distinct beers is a success. We state that because the possibility to group/distinguish similar/different samples implies that in the near future one could apply this technology to the quality control by identifying samples that are sufficiently far from normality.

To evaluate the sample separation and validate the use of the proposed sensor for beer sample discrimination, we considered Welch's t-test on each resonant wavelength, for each pair of beer samples. The test was used to evaluate if the sensor response is different for different beer samples. Furthermore, we also used the LDA projections to evaluate distance. For such, we calculated the Jeffries-Matusita distance (JMDist) between each pair in other to estimate sample separation and evaluate the use of the proposed sensor for beer sample identification. The JMDist is a statistical measure

used to determine the dissimilarity between two probability distributions, in this case the LDA projections for each beer sample. The JMDist can be calculated based on the Bhattacharyya distance (B) (Bhattacharyya, 1943):

$$\text{JMDist} = 2\left(1 - e^{-B}\right) \tag{6}$$

It is important to highlight that the goal of this study was not to rigorously evaluate the performance of the machine learning models themselves, but rather to demonstrate the potential of the optical sensor for beer separation. Consequently, we opted to not fully classify beer brands but to preprocess and cluster the data and analyze the results considering the brand and lot of each sample. Although directly classifying beers concerning brands and lots seems very appealing, we believe that unsupervised clustering might be more useful for real-world scenarios. Indeed, several papers suggest the use of clustering algorithms for real-time anomaly detection (Ahmad et al., 2017; Habeeb et al., 2022; Burbeck and Nadjm-Tehrani, 2007; Li et al., 2021; Mazarbhuiya and Shenify, 2023). The main principle is to automatically group unlabeled data and observe whether the same brands are grouped together. If a given manufacturing lot gets clustered outside its brand cluster, a manufacturing anomaly may have occurred. This approach is simple, suitable for industry, does not need labeled datasets as it grows with the use of the system, and serves as a simple and effective alert during production. Figure 4 shows the step-by-step procedure to identify a beer lot as conform or non-conform. The characterization of a cluster as a typical brand cluster is defined by a voting system, i.e., if more than one cluster is formed for a single beer brand, the cluster with more lots is considered the typical (standard or conform) cluster.

Hence, we measured the beer samples' separability using histograms and clustering to give a glimpse into the future applications of the proposed technology. We used density-based spatial clustering of applications with noise (DBSCAN) (Ester et al., 1996) to cluster data because the number of clusters is automatically set. The algorithm also deals well with noise and arbitrary shaped clusters. Furthermore, we used the silhouette metric to evaluate the cluster separation (Rousseeuw, 1987).

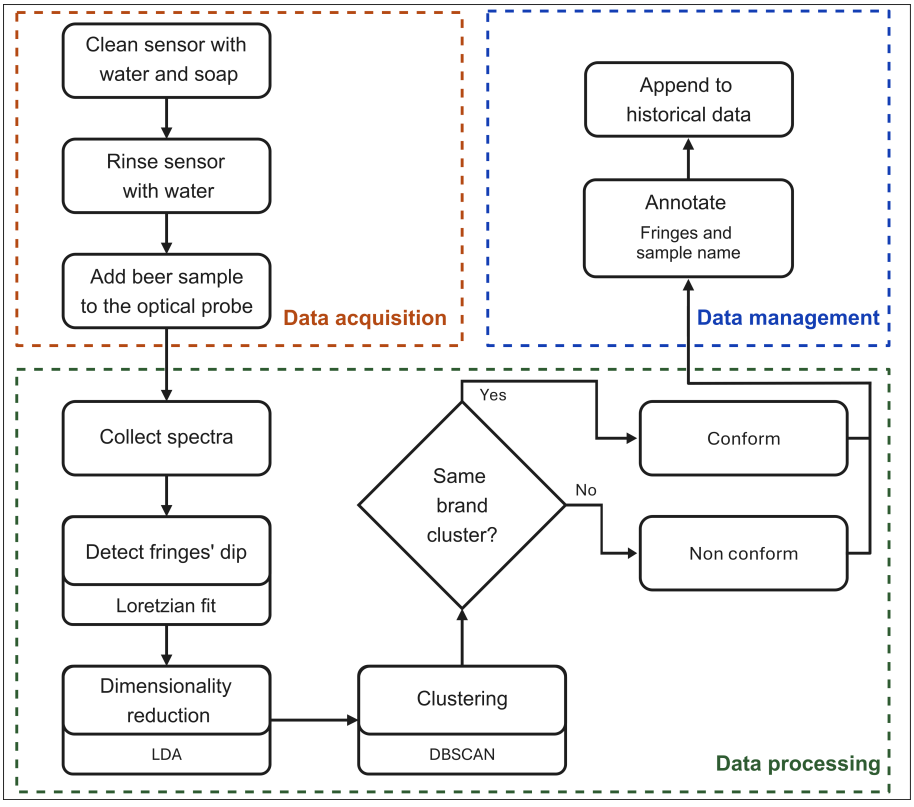

Figure 4: Flowchart for beer quality assessment using the proposed sensor.

To validate the clustering approach, two additional beer samples were measured (A2 and C5). The data was processed as shown in Figure 4. Furthermore, to fully validate the sensor probe and the methodology proposed we made an iterative data ablation test. A data ablation test is a test in which the model, or method, is tested on a subset of the initial dataset. Hence, in this case, the validation started with a single lot per beer brand and a beer sample was added at each step, simulating real-time data collection as illustrated in Figure 4. Note that as new samples are measured, these were added to the historical data.

In this validation test, LDA was fitted using historical data (past measurements) and the DBSCAN considered all data (historical data and the new sample). Based on the clustering results, the lot could be identified as conformant or not. Because of the growing historical data, the principal components shift, and the clusters become more meaningful, emphasizing the scalability of the unsupervised approach considered. Additionally, a voting system was considered to determine which of the two clusters representing the same brand is non-conforming. I.e., redundant clusters with less lots were considered non-conforming.

4. Results

First, let us analyze the brand separation by comparing brands A and B. Figure 5 shows the two components obtained by LDA for water, beer A and beer B. One could note the clear separation between water and beer and a close separation between beer A and beer B. Excluding water from the classification task, we were able to perfectly separate the two beer brands by logistic regression, see Figure 6. This is a promising result for the proposed system, as it shows that one could differentiate between two distinct beer brands using the sensor. Indeed, we estimated the JMDist to be 1.558 arb. units between these two brands, indicating great separability, as this metric is asymptotic to 2 when signatures are completely different.

We then expanded our analysis to include more samples from beer B and a third brand, incorporating information about known manufacturing lots. To evaluate the separability of the beer samples, we analyzed the distribution of the resonant wavelengths, λ_i. Figure 7 shows the probability density function for each resonant dip, λ_i. Density was obtained by Gaussian kernel density estimation using 1 nm bandwidth, which matches the 99% interval for independent measurements of the resonant dips of the sensor. Note that the amplitude and bandwidth of the density peaks are related to the distribution of the resonant wavelengths for each sample, not the resonance strength and bandwidth of the resonant peaks shown in Figure 2.

One could observe the resonant dips' variability among the beer samples. Note that brand B is well defined within λ_i, $i = 0, \ldots, 3$, the first four wavelength dips considered, regarding its mean but the wider bandwidth with lower peak amplitudes observed for samples of beer B indicates a greater variability in the resonant wavelengths for those samples. Whereas brand A could be easily confused with brand C for these wavelengths but could easily be separated considering λ_4. Among the lots, for brand B all seems close, with minor deviations observed considering λ_4. For brand C, however, sample C1 seems far from the other for all wavelengths. At first glance, λ_4 appears to be the most effective

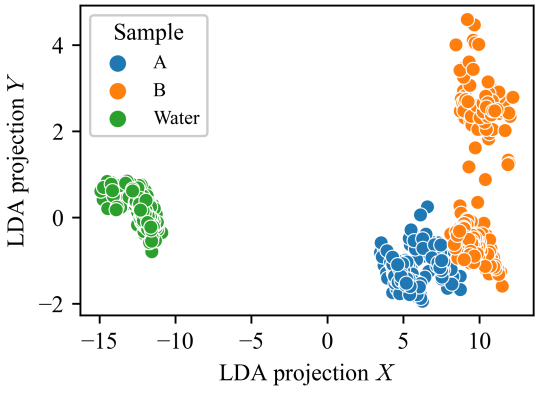

Figure 5: Sensor response to water and two beer samples, after LDA preprocessing.

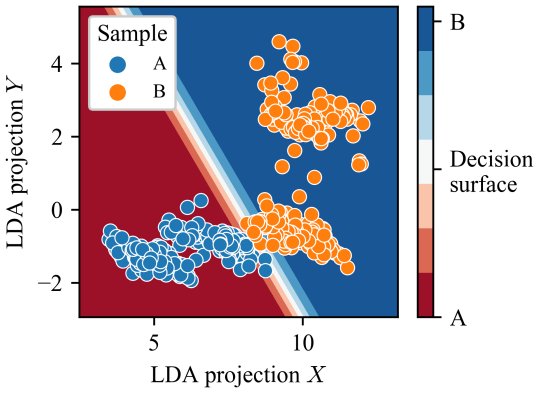

Figure 6: Logistic regression to classify two beer brands.

feature for separating all samples. To evaluate the difference between each resonant wavelength for each group (brand and lot combination), we performed a Welch's t-test for each group combination considering the individual resonant dip. The results are summarized in Figure 8, where equal means are shown in red, and support that λ_4 has more group separability. However, it can even separate different lots of the same brand, so the brand grouping would be lost. Note also that most equality happened for the same brands (except between A and C3 at λ_3), which is desirable for brand grouping.

Applying the LDA to this data and reducing the five variables to two, one could better visualize the brand and lot separation, see Figure 10. The results shown in this figure corroborates with the findings discussed concerning the raw resonant wavelength analysis. Note that brand B is closely packed for all lots, with minor variations. Brand A and B could be separated, as seen before. Moreover, brand C lot were closely packed, except for lot 1.

These results showed the optical probe ability to separate the brands and lots. The brand separation is quite visible and clear, whereas lot discrepancies were also seen but with less separation. Note that the C1 discrepancy may indicate that certain parameters of this particular lot deviate significantly from the expected characteristics of this brand. Hence, the clear separation of beer brands in the LDA space and the successful clustering of similar beers validate the system's ability to group similar beers and separate dissimilar ones, hence, the clustering step should correctly group the same beer brands to further detect non-conformity.

To evaluate separability numerically, we applied the JMDist to each pair of samples to: measure intra-brand similarity and inter-brand separation. The results are summarized in Fig. 10, where the dashed lines represent the samples of the same brand, so the intra-brand values (inside the dashed squares) should be clear and outside darker to indicate proper brand separation/characterization. The results showed that Brand A seems to be very well separated from all other brands, with JMDist close to 2 in almost all cases. Within brand B, there's some variation between the

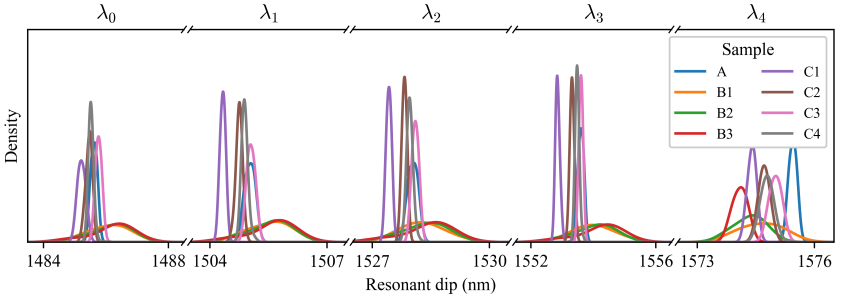

Figure 7: Optical probe's resonant dips distribution among the beer samples.

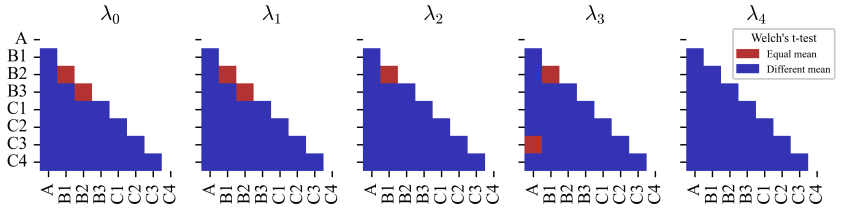

Figure 8: Results for Welch's t-test on each group pair ($\alpha = 0.05$). Blue: different mean and red: equal mean.

lots. B1 and B2 have a lower JMDist (0.95), indicating more similarity compared to B1 and B3 (1.43) or B2 and B3 (0.99). And for Brand C we found a noticeable difference from C1 to C2, C3, and C4 (1.77, 1.86, and 1.88). However, the distance between C2, C3, and C4 suggest similarity.

Finally, to assess the conformity of measured beer samples, we applied the DBSCAN clustering method to segment the collection of data. As illustrated in Figure 11, the resulting clusters effectively grouped the beer samples. The silhouette score of 0.70 indicated strong cluster separation, confirming the probe's ability to distinguish between different beer brands and lots. To further validate the clustering results, we leveraged our knowledge of the samples' true brand assignments and calculated the accuracy and F1-score. The analysis yielded an accuracy of 84% and an F1-score of 72%. It is important to note that these metrics were primarily influenced by the unique cluster created for sample C1, which, as previously discussed, exhibited significant deviations from its brand's mean characteristics. This deviation highlights the sensitivity of the optical probe in detecting potential quality issues. Following the flowchart in Figure 4, an alarm could be generated to flag sample C1 as potentially non-conformant. Overall, the proposed optical probe demonstrates its potential as a valuable tool for beer quality control by effectively segmenting brands and lots with good separability and identifying samples that deviate from their brand's expected characteristics.

Finally, we used the two new beer samples (A2 and C4) and performed a data ablation test, simulating the approach illustrated in Figure 4. We initiated the test with a single lot per beer brand and added a new sample iteratively. So, the validation test started with samples A, B1, and C1; and the samples were added in the following order: B2, C2, B3, C3, C4, A2, and C5. Note that A2 and C5 were not processed before, so the test truly captures the on-line measurement capability. Figure 12 shows the validation procedure, with more emphasis on data management than showed in Figure 4, and the results obtained for each iteration (new added sample). From this figure one could fully understand the optical fiber sensor application in depth and identify the beer classification iteratively.

The algorithm starts and as B2 is added, one could see that this sample is conforming to the B1 measurements, so this lot is conformant regarding the historical data. The next iteration starts by adding data from C2 and DBSCAN creates a new cluster, so C1 and C2 are different enough to be separated into different clusters. This would create a

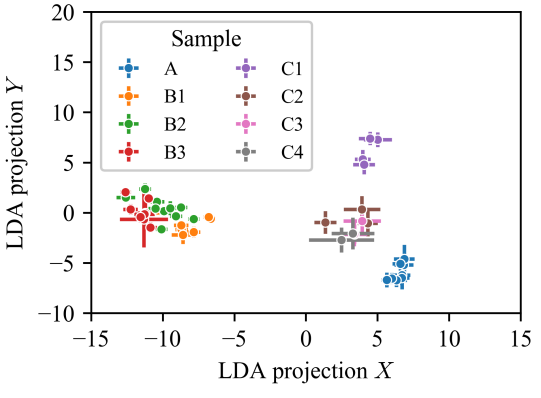

Figure 9: Optical probe results after dimensionality reduction.

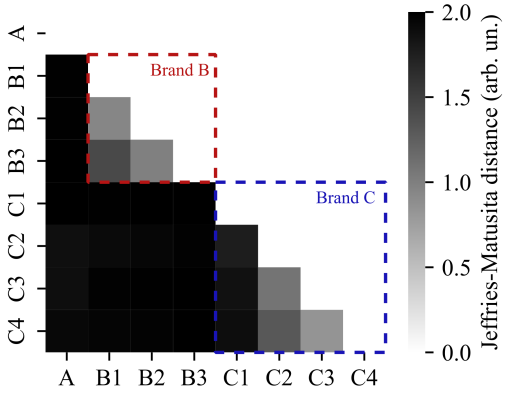

Figure 10: Jeffries-Matusita distance (JMDist) between pairs of beer samples from different brands and lots.

production line alarm, to evaluate these lots. Note, however, that such alarm has little confidence as with only two samples and two clusters not even a voting system could decide which sample is non-conform.

For the next iteration B3 is added and stated as conform, due to the clustering alongside B1 and B2. As C3 was added, it was clustered alongside C2, indicating (by two votes against one) that C1 is non-conforming. Then, C4 was added and clustered with C2 and C3, further indicating that C1 is non-conforming. At this point C4 goes to the historical data and all samples previously studied are used for the LDA dimensionality reduction. Hence, the plot at this stage should be equal to the one shown in Figure 11.

Indeed, after introducing the new A2 beer sample, the clustering results were consistent with those shown on Figure 11. The A2 beer sample was grouped with the existing sample A, indicating conformity with the previous lot of beer A. Finally, the new sample C5 was evaluated and clustered with samples C2-C4, confirming it as a conformant lot. For the last steps, the clusters shape remained close to the shown in Figure 11, since A2 and C5 were conformant, i.e., within existing clusters.

In summary, the validation results of the optical probe demonstrate the efficacy of the unsupervised approach for distinguishing beer brands and lots. The iterative data ablation test confirmed the probe's capability to identify conforming and non-conforming lots in real-time, thus proving to be a valuable tool for beer quality control. The clustering method and the LDA analysis both revealed clear separations between the brands and detected significant

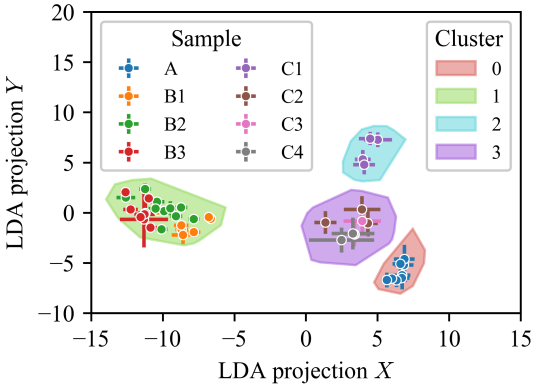

Figure 11: Clustering results.

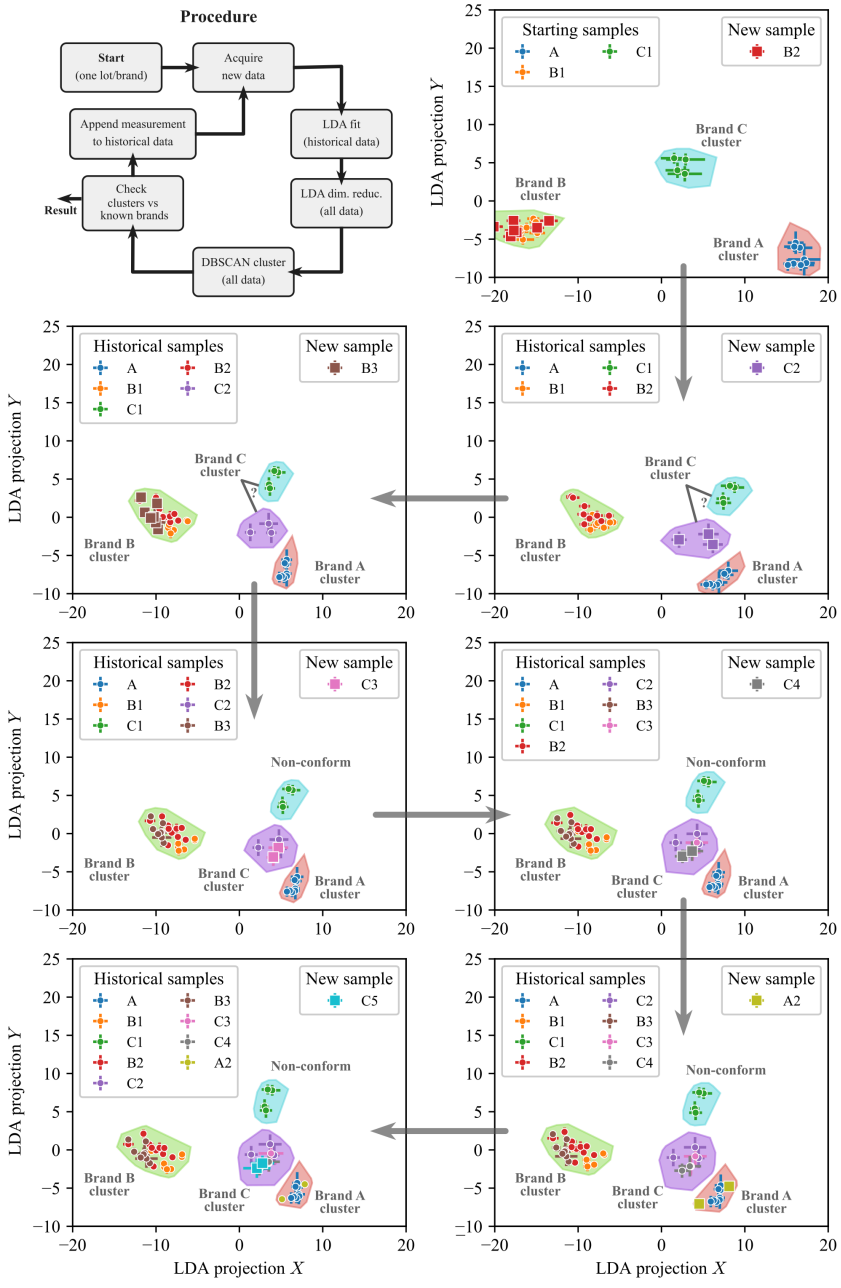

Figure 12: Validation of the intelligent optical fiber probe with iterative (and additive) data ablation.

deviations within lots, emphasizing the probe's sensitivity and reliability. These findings underscore the potential of this optical fiber sensing technology in enhancing quality assessment processes in the beer industry.

5. Conclusions

In this paper, we proposed a novel optical fiber sensing technology for beer quality assessment based on a modal Mach-Zehnder interferometer (MZI). The MZI is a simple and compact probe that can measure the optical properties of beer samples and identify their brands and lots. We demonstrated the performance of the MZI by applying it to three

commercial beer brands with different lots. We used linear discriminant analysis (LDA) and logistic regression to classify the beer samples based on the resonant wavelengths of the MZI. We also used density-based spatial clustering of applications with noise (DBSCAN) to segment the beer samples into clusters. Our results showed that the MZI coupled to LDA was able to separate the beer brands and lots with high accuracy and precision. We also detected a discrepancy in one of the lots, indicating a potential quality issue. The proposed methodology for real-time detection was also validated with data ablation and two additional measurements. The identification of unexpected clusters, i.e. more than one cluster per beer brand could be used to trigger a production line alarm, so the samples involved could be deeply evaluated in the lab. Moreover, the use of a voting system could further enhance the alarm system, by triggering the alarm for the lot with less samples from the same brand.

Our method offers a non-destructive, real-time, and low-cost solution for beer quality monitoring that aligns with the principles of Industry 4.0. We believe that our research contributes significantly to the ongoing efforts to enhance the efficiency and effectiveness of quality assessment in the beer industry. Validation test illustrates the robustness and scalability of the approach based on similarity and clustering, as historical data only improves the sensor detection. Recalling the case where C2 was added and two clusters were created for brand C, until more information regarding this brand was obtained, an uncertain state happened. As soon as C3 was considered, a voting system could establish C1 as non-conformant. This illustrates that upon production line implementation, the intelligent optical probe would improve its performance by itself, generating more meaningful alarms overtime, confirming the method scalability. Although the proposed sensor cannot identify the nonconformity or any beer parameter, it is rapid and useful for production line alarm and rapid assessment of an alteration, so that lab tests could be implemented as soon as possible to reduce production loss.

CRediT authorship contribution statement

Felipe Oliveira Barino: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. **João Victor Calderano:** Data curation Investigation, Methodology. **Renato Luiz Faraco-Filho:** Data curation, Investigation, Methodology. **Alexandre Bessa dos Santos:** Methodology, Supervision, Validation, Project administration, Funding acquisition.

References

Ahmad, S., Lavin, A., Purdy, S., Agha, Z., 2017. Unsupervised real-time anomaly detection for streaming data. Neurocomputing 262, 134–147. doi:10.1016/J.NEUCOM.2017.04.070.

Arslan, M., Tahir, H.E., Zareef, M., Shi, J., Rakha, A., Bilal, M., Xiaowei, H., Zhihua, L., Xiaobo, Z., 2021. Recent trends in quality control, discrimination and authentication of alcoholic beverages using nondestructive instrumental techniques. Trends in Food Science Technology 107, 80–113. doi:10.1016/j.tifs.2020.11.021.

Benadouda, K., Sajid, S., Chaudhri, S.F., Tazally, K.J., Nielsen, M.M.K., Prabhala, B.K., 2023. Current state of sensors and sensing systems utilized in beer analysis. Beverages 9, 5. doi:10.3390/beverages9010005.

Bhattacharyya, A., 1943. On a measure of divergence between two statistical populations defined by their probability distribution. Bulletin of the Calcutta Mathematical Society 35, 99–110.

Blanco, C.A., Fuente, R.D.L., Caballero, I., Rodríguez-Méndez, M.L., 2015. Beer discrimination using a portable electronic tongue based on screen-printed electrodes. Journal of Food Engineering 157, 57–62. doi:10.1016/J.JFOODENG.2015.02.018.

Bock, W.J., Chen, J., Mikulic, P., Eftimov, T., 2007. A novel fiber-optic tapered long-period grating sensor for pressure monitoring. IEEE Transactions on Instrumentation and Measurement 56, 1176–1180. doi:10.1109/TIM.2007.899904.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., Varoquaux, G., 2013. API design for machine learning software: experiences from the scikit-learn project, in: ECML PKDD Workshop: Languages for Data Mining and Machine Learning, pp. 108–122. doi:10.48550/arXiv.1309.0238.

Burbeck, K., Nadjm-Tehrani, S., 2007. Adaptive real-time anomaly detection with incremental clustering. Information Security Technical Report 12, 56–67. doi:10.1016/J.ISTR.2007.02.004.

Cai, J., Liu, Y., Shu, X., 2023. Long-period fiber grating sensors for chemical and biomedical applications. Sensors 23, 542. doi:10.3390/s23010542.

Cao, Z., Zhang, S., Liu, Y., Smith, C.J., Sherman, A.M., Hwang, Y., Simpson, G.J., 2023. Spectral classification by generative adversarial linear discriminant analysis. Analytica Chimica Acta 1261, 341129. doi:10.1016/J.ACA.2023.341129.

Cañas, H., Mula, J., Díaz-Madroñero, M., Campuzano-Bolarín, F., 2021. Implementing industry 4.0 principles. Computers Industrial Engineering 158, 107379. doi:10.1016/j.cie.2021.107379.

Chen, Y., Tang, F., Bao, Y., Tang, Y., Chen, G., 2016. A fe-c coated long-period fiber grating sensor for corrosion-induced mass loss measurement. Optics Letters 41, 2306. doi:10.1364/OL.41.002306.

Chiang, K.S., Liu, Y., Ng, M.N., Dong, X., 2000. Analysis of etched long-period fibre grating and its response to external refractive index. Electronics Letters 36, 966. doi:10.1049/el:20000701.

Colaco, C., Caldas, P., Villar, I.D., Chibante, R., Rego, G., 2016. Arc-induced long-period fiber gratings in the dispersion turning points. Journal of Lightwave Technology 34, 4584–4590. doi:10.1109/JLT.2016.2540678.

Damerow, P., 2012. Sumerian beer: The origins of brewing technology in ancient mesopotamia. Cuneiform Digital Library Journal 2012. URL: https://cdli.mpiwg-berlin.mpg.de/articles/cdlj/2012-2.pdf.

DeMaris, A., 1995. A tutorial in logistic regression. Journal of Marriage and the Family 57, 956. doi:10.2307/353415.

Dong, X., Su, L., Shum, P., Chung, Y., Chan, C.C., . Wavelength-selective all-fiber filter based on a single long-period fiber grating and a misaligned splicing point doi:10.1016/j.optcom.2005.07.075.

Esposito, F., 2021. (INVITED) Chemical sensors based on long period fiber gratings: A review. Results in Optics 5, 100196. doi:10.1016/j.rio.2021.100196.

Ester, M., Kriegel, H.P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise, in: KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pp. 226–231. URL: https://cdn.aaai.org/KDD/1996/KDD96-037.pdf.

en Fan, Y., Zhu, T., Shi, L., Rao, Y.J., 2011. Highly sensitive refractive index sensor based on two cascaded special long-period fiber gratings with rotary refractive index modulation. Applied Optics 50, 4604. doi:10.1364/AO.50.004604.

Filho, R.L.F., Barino, F.O., Calderano, J., Ítalo Fernando Valle Alvarenga, Campos, D., dos Santos, A.B., 2023. In-fiber mach–zehnder interferometer as a promising tool for optical nose and odor prediction during the fermentation process. Optics Letters 48, 3905. doi:10.1364/OL.486742.

Filho, R.L.F., Barino, F.O., Calderano, J.V., Campos, D., Ítalo Fernandes Valle Alvarenga, Manhas, A.W.P., dos Santos, A.B., 2024. Novel beer bitterness measurement instrument using optical fiber sensor. Journal of Food Engineering 383, 112246. doi:10.1016/j.jfoodeng.2024.112246.

Gan, W., Li, Y., Liu, T., Yang, Y., Song, B., Dai, S., Xu, T., Wang, Y., Lin, T.J., Zhang, P., 2022. Rapid and sensitive detection of ammonia in water by a long period fiber grating sensor coated with sol-gel silica. Optics Express 30, 33817. doi:10.1364/OE.472205.

Gangwar, R.K., Kumari, S., Pathak, A.K., Gutlapalli, S.D., Meena, M.C., 2023. Optical fiber based temperature sensors: A review. Optics 2023, Vol. 4, Pages 171-197 4, 171–197. doi:10.3390/OPT4010013.

Ghasemi-Varnamkhasti, M., Rodríguez-Méndez, M.L., Mohtasebi, S.S., Apetrei, C., Lozano, J., di, H., Razavi, S.H., de Saja, J.A., 2012. Monitoring the aging of beers using a bioelectronic tongue. Food Control 25, 216–224. doi:10.1016/J.FOODCONT.2011.10.020.

Habeeb, R.A.A., Nasaruddin, F., Gani, A., Amanullah, M.A., Hashem, I.A.T., Ahmed, E., Imran, M., 2022. Clustering-based real-time anomaly detection—a breakthrough in big data technologies. Transactions on Emerging Telecommunications Technologies 33, e3647. doi:10.1002/ETT.3647.

Hassoun, A., Jagtap, S., Garcia-Garcia, G., Trollman, H., Pateiro, M., Lorenzo, J.M., Trif, M., Rusu, A.V., Aadil, R.M., Šimat, V., Cropotova, J., Câmara, J.S., 2023a. Food quality 4.0: From traditional approaches to digitalized automated analysis. Journal of Food Engineering 337, 111216. doi:10.1016/j.jfoodeng.2022.111216.

Hassoun, A., Jagtap, S., Trollman, H., Garcia-Garcia, G., Abdullah, N.A., Goksen, G., Bader, F., Ozogul, F., Barba, F.J., Cropotova, J., Munekata, P.E., Lorenzo, J.M., 2023b. Food processing 4.0: Current and future developments spurred by the fourth industrial revolution. Food Control 145, 109507. doi:10.1016/j.foodcont.2022.109507.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Reduced-Rank Linear Discriminant Analysis. 2 ed.. Springer. pp. 113–19.

Jaywant, S.A., Singh, H., Arif, K.M., 2022. Sensors and instruments for brix measurement: A review. Sensors 22, 2290. doi:10.3390/s22062290.

Korposh, S., Selyanchyn, R., James, S., Tatam, R., Lee, S.W., 2014. Identification and quality assessment of beverages using a long period grating fibre-optic sensor modified with a mesoporous thin film. Sensing and Bio-Sensing Research 1, 26–33. doi:10.1016/j.sbsr.2014.06.001.

Lasi, H., Fettke, P., Kemper, H.G., Feld, T., Hoffmann, M., 2014. Industrie 4.0. WIRTSCHAFTSINFORMATIK 56, 261–264. doi:10.1007/s11576-014-0424-4.

Lee, B.H., Nishii, J., 1999. Dependence of fringe spacing on the grating separation in a long-period fiber grating pair. Applied Optics 38, 3450. doi:10.1364/AO.38.003450.

Li, B., Jiang, L., Wang, S., Tsai, H.L., Xiao, H., 2011. Femtosecond laser fabrication of long period fiber gratings and applications in refractive index sensing. Optics Laser Technology 43, 1420–1423. doi:10.1016/j.optlastec.2011.04.011.

Li, J., Izakian, H., Pedrycz, W., Jamal, I., 2021. Clustering-based anomaly detection in multivariate time series data. Applied Soft Computing 100, 106919. doi:10.1016/J.ASOC.2020.106919.

Liu, T., Li, Y., Dai, X., Gan, W., Wang, X., Dai, S., Song, B., Xu, T., Zhang, P., 2023. Simultaneous detection of temperature, strain, refractive index, and ph based on a phase-shifted long-period fiber grating. Journal of Lightwave Technology 41, 5169–5180. doi:10.1109/JLT.2023.3254550.

Mallet, Y., Coomans, D., Vel, O.D., 1996. Recent developments in discriminant analysis on high dimensional spectral data. Chemometrics and Intelligent Laboratory Systems 35, 157–173. doi:10.1016/S0169-7439(96)00050-0.

Mazarbhuiya, F.A., Shenify, M., 2023. A mixed clustering approach for real-time anomaly detection. Applied Sciences 2023, Vol. 13, Page 4151 13, 4151. doi:10.3390/APP13074151.

Mignani, A.G., Ciaccheri, L., Mencaglia, A.A., Ottevaere, H., Baća, E.E.S., Thienpont, H., 2013. Optical measurements and pattern-recognition techniques for identifying the characteristics of beer and distinguishing belgian beers. Sensors and Actuators B: Chemical 179, 140–149. doi:10.1016/J.SNB.2012.10.029.

Méndez, A., Morse, T.F., 2011. Specialty optical fibers handbook. Elsevier.

Navruz, I., Ari, F., Bilsel, M., AL-Mashhadani, Z.A., 2018. Enhancing refractive index sensitivity using micro-tapered long-period fiber grating inscribed in biconical tapered fiber. Optical Fiber Technology 45, 201–207. doi:https://doi.org/10.1016/j.yofte.2018.07.018.

Novák, M., Palya, D., Bodai, Z., Nyiri, Z., Magyar, N., Kovács, J., Eke, Z., 2017. Combined cluster and discriminant analysis: An efficient chemometric approach in diesel fuel characterization. Forensic Science International 270, 61–69. doi:10.1016/J.FORSCIINT.2016.11.025.

Qi, L., Zhao, C.L., Yuan, J., Ye, M., Wang, J., Zhang, Z., Jin, S., 2014. Highly reflective long period fiber grating sensor and its application in refractive index sensing. Sensors and Actuators B: Chemical 193, 185–189. doi:10.1016/j.snb.2013.11.063.

Rajan, G., Iniewski, K. (Eds.), 2017. Optical Fiber Sensors. CRC Press. doi:10.1201/b18074.

Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20, 53–65. doi:10.1016/0377-0427(87)90125-7.

Shen, F., Wang, C., Sun, Z., Zhou, K., Zhang, L., Shu, X., 2017. Small-period long-period fiber grating with improved refractive index sensitivity and dual-parameter sensing ability. Optics Letters 42, 199. doi:10.1364/OL.42.000199.

Shu, X., Zhang, L., Bennion, I., Shu, X., Zhang, L., Bennion, I., Shu, X., Zhang, L., Bennion, I., 2002. Sensitivity characteristics of long-period fiber gratings. Journal of Lightwave Technology 20, 255–266. doi:10.1109/50.983240.

Tai, Y.H., Pan, M.Y., Lin, E.H., Huang, D.W., Wei, P.K., 2016. Quality detection of alcoholic beverages using optical fiber tips. IEEE Sensors Journal 16, 5626–5631. doi:10.1109/JSEN.2016.2571307.

Tamang, J.P., Cotter, P.D., Endo, A., Han, N.S., Kort, R., Liu, S.Q., Mayo, B., Westerik, N., Hutkins, R., 2020. Fermented foods in a global age: East meets west. Comprehensive Reviews in Food Science and Food Safety 19, 184–217. doi:10.1111/1541-4337.12520.

Tan, Y., Ji, W., Mamidala, V., Chow, K., Tjin, S., 2014. Carbon-nanotube-deposited long period fiber grating for continuous refractive index sensor applications. Sensors and Actuators B: Chemical 196, 260–264. doi:10.1016/j.snb.2014.01.063.

Tsai, Y.C., Huang, J.D., Chiu, C.C., 2007. Amperometric ethanol biosensor based on poly(vinyl alcohol)–multiwalled carbon nanotube–alcohol dehydrogenase biocomposite. Biosensors and Bioelectronics 22, 3051–3056. doi:10.1016/j.bios.2007.01.005.

Vasconcelos, H.C.A.S.G., de Almeida, J.M.M.M., Saraiva, C.M.T., da Silva Jorge, P.A., Coelho, L.C.C., 2019. Mach–zehnder interferometers based on long period fiber grating coated with titanium dioxide for refractive index sensing. Journal of Lightwave Technology 37, 4584–4589. doi:10.1109/JLT.2019.2912829.

Vengsarkar, A.M., Lemaire, P.J., Judkins, J.B., Bhatia, V., Erdogan, T., Sipe, J.E., 1996. Long-period fiber gratings as band-rejection filters. Journal of lightwave technology 14, 58–65. doi:10.1109/50.476137.

Viejo, C.G., Fuentes, S., 2020. Low-cost methods to assess beer quality using artificial intelligence involving robotics, an electronic nose, and machine learning. Fermentation 6, 104. doi:10.3390/fermentation6040104.

Xu, Z., Luo, Y., Mao, Y., Peng, R., Chen, J., Soteyome, T., Bai, C., Chen, L., Liang, Y., Su, J., Wang, K., Liu, J., Kjellerup, B.V., 2020. Spoilage lactic acid bacteria in the brewing industry. Journal of Microbiology and Biotechnology 30, 955–961. doi:10.4014/jmb.1908.08069.

Xue, X., Zhang, W., Zhang, H., Gao, S., Geng, P., Bai, Z., Li, J., 2013. Simultaneous measurement of temperature and refractive index using a simplified modal interferometer composed of tilted long-period fiber grating and fiber bitaper. Measurement Science and Technology 24, 065103. doi:10.1088/0957-0233/24/6/065103.

Yin, G., Wang, Y., Liao, C., Zhou, J., Zhong, X., Wang, G., Sun, B., He, J., 2014. Long period fiber gratings inscribed by periodically tapering a fiber. IEEE Photonics Technology Letters 26, 698–701. doi:10.1109/LPT.2014.2302901.

Yunus, M., Aziz, S., Tang, B., Usman, Y.H., Irsan, I., Kurniasari, S., 2023. Imperfection method based on optical fiber for alcohol content detection sensor. Jurnal Pijar Mipa 18, 260–264. doi:10.29303/jpm.v18i2.4605.

Zhou, K., Liu, T., Zhou, L., 2015. Industry 4.0: Towards future industrial opportunities and challenges, in: 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), IEEE. pp. 2147–2152. doi:10.1109/FSKD.2015.7382284.