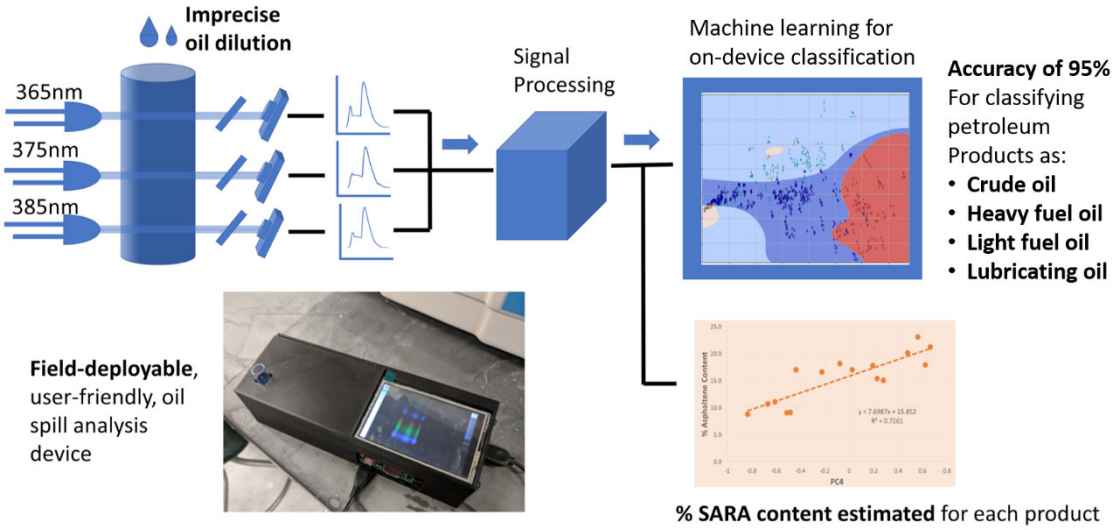
Handheld UV fluorescence spectrophotometer device for the classification and analysis of petroleum oil samples

Matthew V. Bills[a], Andrew Loh[b], Katelyn Sosnowski[a], Brandon T. Nguyen[a], Sung Yong Ha[b], Un Hyuk Yim[b,*], and Jeong-Yeol Yoon[a,*]

[a] Department of Biomedical Engineering, The University of Arizona, Tucson, Arizona 85721, United States

[b] Korea Institute of Ocean Science and Technology, Geoje-si, Gyeongsangnam-do 53201, Republic of Korea

ABSTRACT: Oil spills can be environmentally devastating and result in unintended economic and social consequences. An important element of this concerted effort includes the ability to rapidly classify and characterize oil spill samples, preferably on-site. An easy-to-use, handheld sensor is developed and demonstrated in this work, capable of classifying oil spills rapidly on-site. Our device uses the computational power and affordability of a Raspberry Pi microcontroller and a Pi camera, coupled with three ultraviolet light emitting diodes (UV-LEDs), a diffraction grating, and collimation slit, in order to collect a large data set of UV fluorescence fingerprints from various oil samples. Based on a 160-sample (in 5x replicates each with slightly varied dilutions) database this platform is able to classify oil samples into four broad categories: crude oil, heavy fuel oil, light fuel oil, and lubricating oil. The device uses principal component analysis (PCA) to reduce spectral dimensionality (1203 features) and support vector machine (SVM) for classification with 95% accuracy. The device is also able to predict some physiochemical properties, specifically saturate, aromatic, resin, and asphaltene percentages (SARA) based off linear relationships between different principal components (PCs) and the percentages of these residues. Sample preparation for our device is also straightforward and appropriate for field deployment, requiring little more than a Pasteur pipette and not being affected by dilution factors. These properties make our device a valuable field-deployable tool for oil sample analysis.



Keywords: ultraviolet light emitting diode; fluorescence spectroscopy; Raspberry Pi; saturate, aromatic, resin, and asphaltene contents; oil spill; support vector machine.

* Corresponding authors. E-mail addresses: uhyim@kiost.ac.kr (U.H.Y.); jyyoon@email.arizona.edu (J.-Y.Y.).

1. INTRODUCTION

Oil spills can have substantial environmental, economic, and societal hazards especially in the ocean (Lee et al., 2017; Yim and Short, 2017; Brussaard et al., 2016). Interactions of spilled oils with biota and their habitats are complex, often impacting animal lifecycles, disrupting marine mating and feeding cycles, and effecting fishing and tourism industries (Yim and Short, 2017). Additionally, oil spills are very harmful to marine vertebrates and are known to penetrate the plumage of birds and the fur of aquatic mammals, undermining their insulating and buoyancy properties. Many birds ingest oil as they try to clean their plumage, which can poison them (Bergeon Burns et al., 2014).

Oil spills can arise from many sources ranging from the extraction and transportation of petroleum, to accidents and routine occurrences during fueling and shipping of any marine vessel including but not limited to tankers (Talley et al., 2004; ITOPF, 2019; United States Census Bureau, 2019). Additionally, the severity of an oil spill does not always correlate to the size of the spill. Smaller oil spills are still serious environmental hazards that are often overlooked in part due to a lack of manpower and the difficulty of classifying oil spills on-site. While prevention is preferred, rapid response and informed clean-up protocols significantly improve the outcome and reduce the effect of oil spills.

An important element of this concerted effort includes the ability to rapidly classify and characterize oil spill samples, preferably on-site. Such classification of oil type and identification of its specific residues are important parts of any oil spill investigation (Yim and Short, 2017; Stout, 2016; Suneel et al., 2014; Wang et al., 2006; Wang, 2002; Zakaria et al., 2000; Federici and Mintz, 2014; Han and Clement, 2018). The most commonly used methods for classifying/identifying these oil spill samples and obtaining their chemical fingerprints are gas chromatography with mass spectrometry detection (GC/MS) or flame ionization detection (GC/FID), as well as thin layer chromatography with flame ionization detection (TLC/FID) (Yim and Short, 2017; Stout, 2016; Wang et al., 2006, Wang, 2002; Han and Clement, 2018). This battery of tests can be used to obtain a great deal of information about oil spill samples including: the relative abundance of saturate hydrocarbons, aromatic hydrocarbons, resins, and asphaltene molecules (SARA); the relative abundances of differently sized hydrocarbons; specific residues, etc. These advanced techniques are invaluable for identifying oil spill samples. However, these tests require time, resources, and highly trained technicians. Additionally, the samples need to be delivered to and tested in advanced laboratory settings (Yim and Short, 2017; Han and Clement, 2018). As a result, it becomes impractical to test thousands of field samples from a large number of small oil spills or to repeatedly run tests following a large oil spill as is often required. As a result, there is a clear need for more cost-effective, rapid, emergency field investigation strategies and monitoring methods (Yim and Short, 2017; Han and Clement, 2018).

In investigating oil spill samples from the Deepwater Horizon's (DWH) spill, Han and Clement demonstrated six sensory inspection techniques that were shown to be useful for characterizing tar-ball samples obtained from that specific oil spill sample. Those tests included a strength test, shine test, sand test, stickiness test, color test, and odor test (Han and Clement, 2018). These tests lack rigorous controls and are dependent on the subjectivity of the individual performing the test. Additionally, the results of

these tests do not readily lend themselves to identifying a new unknown sample, as they were designed to specifically identify and characterize the DWH spill.

The goal of this work was to develop an accurate portable device that can characterize and classify unknown petroleum samples by comparing the fluorescence produced by oil samples excited by UV-A light sources.

Our device makes use of the fluorescent properties of oil samples through developing a handheld spectrophotometer. It measures fluorescence from 350 nm to 750 nm using three ultraviolet light emitting diodes (UV-LEDs) as light sources (365 nm, 375 nm, and 385 nm), a Raspberry Pi microcontroller, and a Pi camera. Digitalized light intensity values are collected for each integer wavelength value, generating 401 data points from 350 to 750 nm. As three different spectra are collected per sample using three different UV-LEDs, a total of 1203 features are collected per sample. Measurements are repeated five times per sample, each time imprecisely varying the dilution slightly to test robustness of the device. A total of 160 oil samples are tested to build a single data set, corresponding to 1203 features x 5 replicates x 160 samples = 962,400 data points in our model. This high dimension was reduced to 10 dimensions using principal component analysis (PCA) to aid in easier visualization and more accurate classification of the data. All 160 oil samples are broadly classified into one of four major categories using support vector machine (SVM) with a radial basis function (RBF) kernel: 1) crude oil, 2) heavy fuel oil, consisting of marine fuel oil (MFO) and Bunker C, 3) light fuel oil, consisting of all other fuel oils ranging from gasoline, diesel, Bunker A, marine gas oil (MGO), and up to Bunker B, and 4) lubricating oil.

Correlating the principal components with physiochemical properties of oil samples, including the percent contents of saturated hydrocarbons, aromatic hydrocarbons, resins, and asphaltene molecules (SARA) obtained by TLC/FID, we aim to predict with some accuracy the percent contents of SARA for tested oil samples. Fig. 1 graphically depicts the overall procedure of our method.

Information obtained using our method and device provides less detailed information about the sample make-up than traditional analytical chemistry methods including gas and thin layer chromatography (GC/FID and TLC/FID). However, traditional analytical techniques require highly trained personnel, significant time (both for transportation and testing), very expensive equipment, and significant maintenance costs and time. Our device was assembled for approximately 750 USD (the majority of which was due to purchasing expensive high-power UV-LEDs for which substitutions could likely be made) excluding labor cost, this is significantly lower than comparable analytical devices cost. Additionally, samples can be prepared easily by eyeball by non-scientific personnel using no more than a Pasteur pipette. Diluted samples can be analyzed to obtain sample classification and predicted SARA content in under 30 seconds, including time for the user to input a sample name and known properties. All data collection/processing, machine learning classification, and SARA content prediction are conducted within the device using a custom-built Python code loaded on the Raspberry Pi microcontroller. Neither additional computer nor internet/cellular data connection are necessary. It is reasonable to assume our device might be employed in concert with pre-existing methods providing rapid on-site analysis of oil samples to inform clean-up protocols and aid in identifying culpable parties. If additional information was desired more precise analytical methods could still be employed.

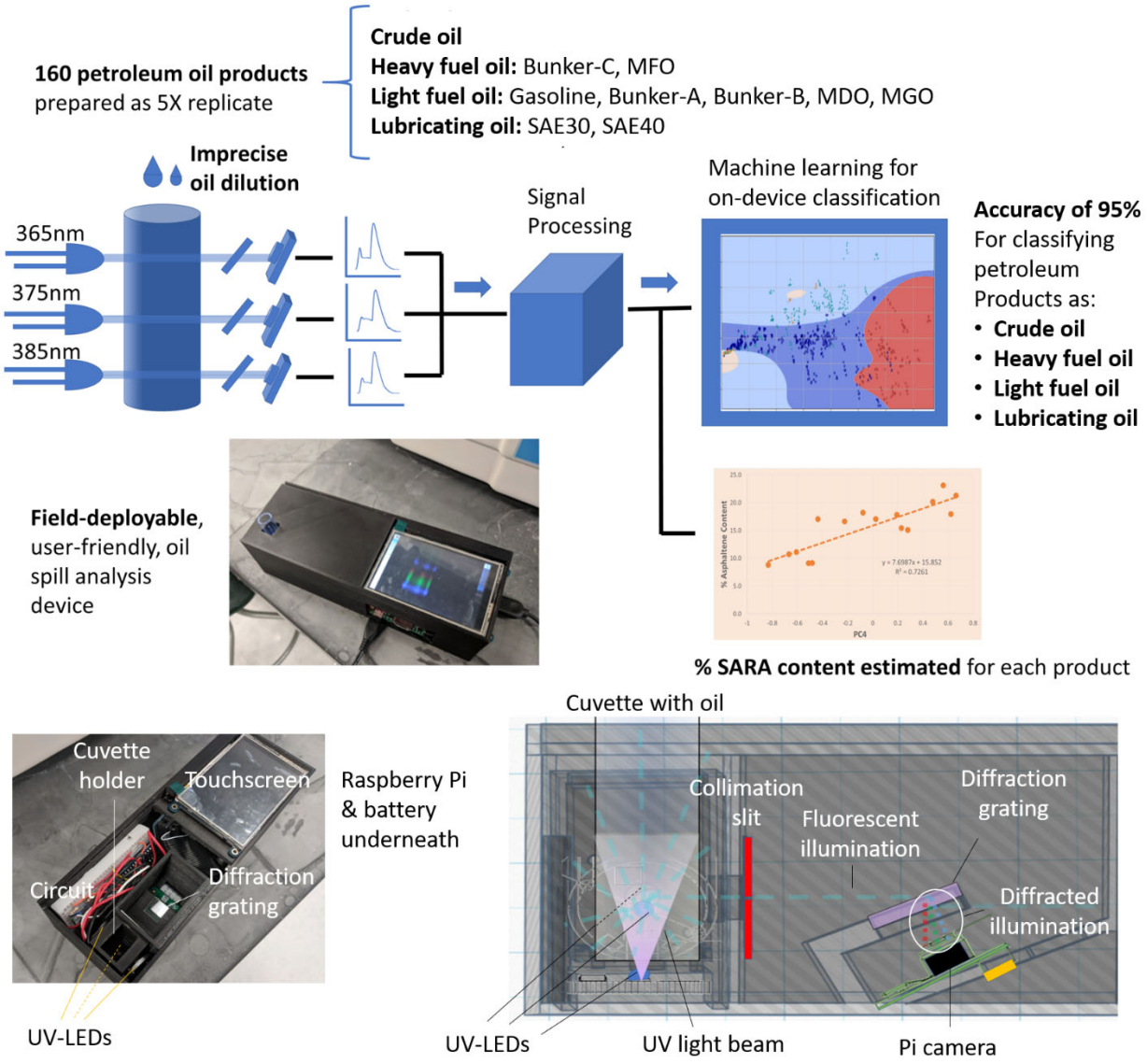


Fig. 1. Schematic illustration of the method and the device. MFO = marine fuel oil; MDO = marine diesel oil; MGO = marine gas oil.

2. MATERIALS AND METHODS

2.1. Device design

Device housing was created using SolidWorks 2013 x64 Edition (Dassault Systèmes, Vélizy-Villacoublay, France) and 3D-printed in ABS (acrylonitrile-butadiene-styrene) polymer using Ultimaker 3 (Ultimaker, Utrecht, Netherlands). The housing can be subdivided into four major areas: 1) housing for the Raspberry Pi microcontroller; 2) a 3D-printed cuvette holder with mounts for three UV-LEDs; 3) an optical path with mounts for a collimation slit diffraction grating and a camera mount; 4) a region for device circuitry and a battery (Fig. 1).

Raspberry Pi microcontroller (Raspberry Pi 3B+; Raspberry Pi Foundation, Cambridge, UK) was housed in the first area, which has an impressive computation power (requiring no external computer), and could communicate with UV-LEDs, circuit, Pi camera, and a touchscreen. Two additional hats were designed and used to integrate with the Raspberry Pi to accommodate the followings: 1) a universal power supply (UPS) as a battery (Raspi UPS HAT V1.1; Geekworm, Shenzhen, Guangdong, China), made out of 3.7 V 2600 mAh lithium ion battery, enabling the device to be used portably without requiring constant connection to a wall outlet; 2) a touchscreen (PiTFT 2.8; Adafruit Industries, New York, NY, USA) to aid in the portability of the final device in lieu of a monitor and a mouse.

The cuvette holder is a separate 3D-printed component made out of ABS that slides firmly into place in the housing (Fig. 1). The top of the housing aligns with the sample loading port where cuvette is inserted. On the side if the loading port nearest the optical path, there is a window that aligns with the collimation slit. While the side opposite the collimation slit is blank, the remaining three other sides (left, right, and bottom) all contain the mounts for the UV-LEDs emitting fluorescence excitations peaked at 365 nm, 375nm, and 385 nm, respectively (part numbers: M365D2, M375D4 and M385D2; Thorlabs, Inc., Newton, NJ, USA).

The optical path passes through a 3D-printed ABS slit with the final dimension of 0.23 mm x 5.6 mm slit. The light travels 23 mm before interacting with a 1200 line/mm diffraction grating (GT13-12; Thorlabs, Inc.) positioned at 75° relative to the grating normal, splitting and diffracting the light. The Pi camera (Camera V2.1; Raspberry Pi Foundation) is positioned parallel to the grating so that the camera lens is positioned 3 mm underneath the diffraction grating to collect the light intensities (350–750 nm) from both the first and second order diffractions.

2.2. Circuit design

The UV-LEDs toggle on and off one at a time using the general-purpose input/output (GPIO) pins from the Raspberry Pi. The problem we encountered was that the output from the GPIO pins was too weak (3.3 V with 16 mA restricted current) to run the UV-LEDs. Raspberry Pi does have a 5 V voltage output pin that can provide up to 1 A, which is appropriate to power the UV-LEDs, although it cannot be toggled. The voltage signal from the GPIO output pin was utilized as a control signal, through feeding it into a transistor to toggle the 5 V voltage output, enabling this switch-like behavior. A couple of diodes and resistors are also used to complete this circuit (Supplementary Figure S1).

2.3. Oil samples

Crude and marine fuel oils (MFOs) used in this study were purchased from SGS Korea (Republic of Korea), and gasoline, diesel, lubricating, and bunker oils were collected from the Korean Coast Guard marine vessels throughout Republic of Korea.

2.4. Physicochemical testing with TLC/FID

Saturate, aromatic, resin, and asphaltene (SARA) contents were measured by thin layer chromatography with a flame ionization detector (TLC/FID; Iatroscan Mk IV; Iatroscan Labs Inc., Tokyo, Japan). Briefly, 1 µL aliquots in dichloromethane (DCM) were spotted at the base of a silica gel sintered glass rod, known as a chromarod (Chromarod S III; Iatron Laboratories) using a micropipette. Before sample spotting, the chromarods were blank scanned to clean the rods and check for contamination. Blank analyses of the rods were devoid of any detectable signal. After sample spotting, the rods were

successively developed in hexane (30 min elution time), toluene (15 min), and 95/5 DCM/methanol (2.5 min). Between each development step, the rods were dried for 1 min at 200 mbar using low-pressure nitrogen. Rods were analyzed in a TLC/FID analyzer at a scan speed of 30 seconds per chromarod (Stephens, 2004). The operating conditions of the FID were maintained at a flow of 160 mL min^{-1} H_2 and 2 L min^{-1} air. Values that were obtained from each sample was based on sample triplicates, and the mean values of all samples and instrumental replicates were within the acceptable relative standard deviation (RSD) values.

2.5. Sample preparation

Initially, our device was tested using 10 samples. Once our device showed promise, it was further tested on 160 oil samples, five replicates each. Depending on sample clarity, samples were diluted in ≥ 99.5% pure dichloromethane (DCM) using one of the following two methods: Opaque samples were diluted until transparent with just a light shade of coloration to it. Transparent samples were diluted roughly ten-fold. This dilution process is relatively rapid and simple enough that most untrained personnel can perform it easily. As the dilutions are made intuitively, the exact dilution factors varied not only from sample-to-sample, but also within the replicates. This is intentional to improve robustness of our device. Final volumes ranging from 1-3 mL were added by Pasteur pipette to borosilicate cuvettes for analysis.

2.6. Comparison to commercial spectrophotometer

Fluorescence spectra of the ten samples described previously in section 2.5 were obtained using a commercial spectrophotometer (USB 4000; Ocean Optics, Dunedin, Florida, United States), coupled with the same UV-LEDs housed in our device.

2.7. Python script

We designed a script in Python to be run within our device to enable easy data collection and sample identification. A key element toward making this device accessible to the end user was designing a friendly graphical user interface (GUI) that is easy to navigate and directs the user how to use the device. The GUI was designed using the tkinter library (available at https://docs.python.org/3/library/tkinter.html).

2.8. Calibration of device

The integrated Python script calibrates the device each time it is run in order to ensure that all data is processed identically. This is an important step to relate the pixel location in the acquired image to the wavelength of a spectrum and to normalize the spectra regardless of the dilution factor. Each time a sample is run the device is calibrated using a two-step calibration method. The end result of these calibration steps is to determine which pixels from the raw images will be used to generate the output data representing spectra from 350 nm to 750 nm.

The first step of calibration is achieved by asking the user to load a blank sample of DCM (used to dilute samples). Images of the fluorescence from a single sample with each of the three UV-LEDs are collected and the grayscale of each of those images are overlaid together. The user is then asked to choose the left and right boundaries for the data they are interested in. This is done to remove large

regions of “black space” in the image that corresponds to areas that fall outside the path of the light beam and therefore do not contribute to the sample’s fluorescent spectral response.

The second calibration step is completely automated and determines which rows of pixels correspond to integer values of wavelengths from 350 nm to 750 nm. In this step, the 365-nm illuminated image and the 375-nm illuminated image are captured from the blank sample (DCM) and further split into their red and green channels. The images are cropped using the user specified boundaries from the first calibration step mentioned above. The images are then converted to a NumPy array using the NumPy library (https://numpy.org/). Each row is averaged together to yield average intensity observed at that specific wavelength, and those average values are appended to a list representing the average intensity value at each row of pixels in the cropped image. This list was then put through the SciPy.signal.find_peaks algorithm that finds peaks from a continuous list of data setting “distance” (between peaks) to 200 and (minimum) prominence to 1. This identifies only the most prominent peaks, their position and intensity. We expect the first order diffraction spectra to show up brighter and earlier. Also, the DCM spectra is known to exhibit two peaks in the green channel with the 365-nm and 375-nm illuminations: the first peak corresponds to a spike in UV leakage at 390 nm, and the second peak corresponds to a very weak elevation in green intensity at 515 nm. We only expect a single peak in the red channel with the 365-nm and 375-nm illuminations that corresponds to UV leakage at 390 nm. Most digital cameras recognize very short blue wavelength color (close to 400 nm) not only in red channels but also in blue channels, to mimic the human eye’s perception of this color as “violet.” The pixel number is obtained for each of the points and they are all plotted and processed with a linear regression from the SciPy statistics library (https://docs.scipy.org/doc/scipy/reference/stats.html) and the R-squared value, slope, and intercept are all noted. If the R-squared value is lower than 0.99 an error appears on the GUI, advising the user to rerun the blank sample. This occurs if the 365-nm and 375-nm light sources cannot agree where those peaks should be. The device then checks to make sure that both 350 nm spectra are in range by making sure that the intercept is less than or equal to 350 nm. If this is not the case the device throws an error, asking the user to check the positioning of the optical components.

A list of desired pixels that correspond to integer values of wavelengths from 350 to 750 nm is then generated by iteratively running wavelengths through the linear equation generated from the linear regression and appending pixel values to a new list. This list corresponds to the pixel numbers that most closely match the integer wavelength values in the range of 350 to 750 nm and is applied to all samples that are collected immediately following calibration. Figure 2 graphically illustrates this calibration procedure for one of the three UV-light sources obtained from a representative DCM sample.

Following calibration, when the user is ready to collect data, the user inputs a name for the sample they are testing. The device illuminates the sample at each of the three wavelengths and captures images during each illumination. These images are then gray-scaled, cropped based off the user-defined boundaries, and converted into a NumPy array. The rows of the array are averaged together, and the pixel values that correlate to integer wavelength values from 350 nm to 750 nm based off of the array collected during calibration are appended to a list that contains the sample name, regression slope, and regression intercept. The lists for the 365-nm, 375-nm and 385-nm UV-LED excitations are appended to the list with the sample name, and all data is added as a row to a comma separated text (CSV) document that contains all samples that are collected from the device.

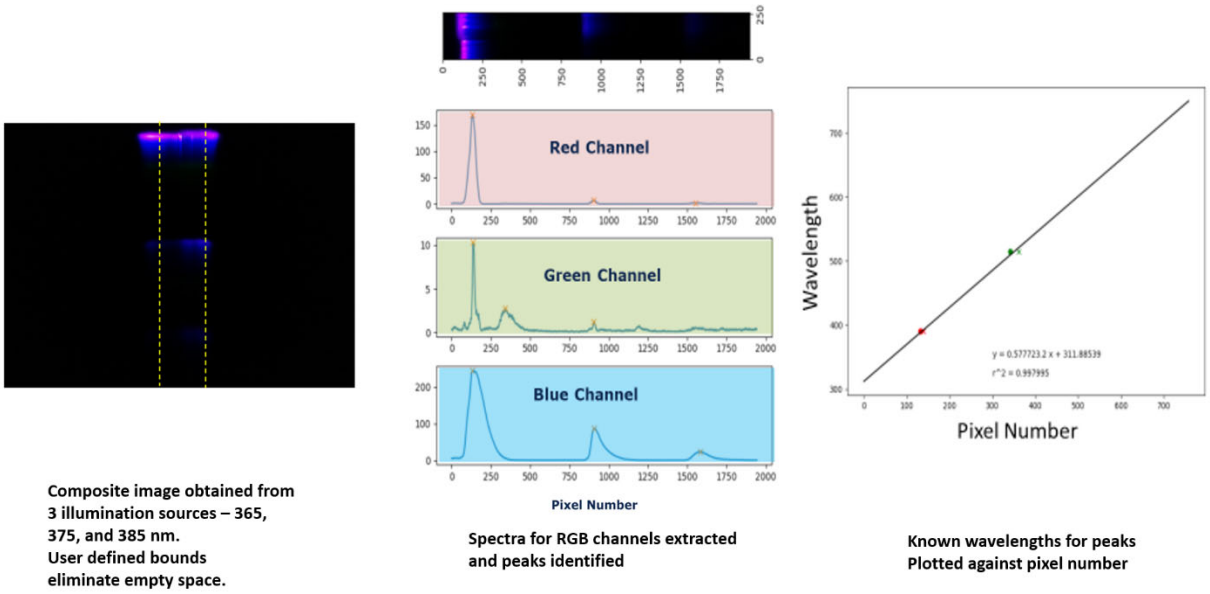


Fig. 2. Calibration procedure for the device.

2.9. Model validation

Data obtained from the device was initially plotted in Microsoft Excel for visual inspection. The data was then processed using principal component analysis (PCA) from the sci-kit learn library (https//scikit-learn.org/stable/) to reduce dimensionality from 1203 features to 10. Sample classification was also completed using the sci-kit learn function for support vector machine (SVM) with a radial basis function (RBF) kernel. SVM parameters C and gamma were optimized by grid-search, where we made use of the sci-kit learn function for kfold validation setting n_splits equal to three and iteratively validating our data for different values of C (.001, .01, .1, 1, 10, 100, and 1000) and gamma (.001, .01, .1, 1, 10, 100, and 1000). Linear and polynomial kernels were also tested on both the raw data and PCA transformed data. We also tested K-Nearest Neighbor, Random Forest, TanH-MLP, and Relu-MLP classifiers but settled on SVM classification using an RBF kernel of PCA transformed data, partially due to computation time and also because it yielded one of the better classification rates.

This classification is also integrated into the Python script so that sample classification is immediately returned to the end user.

2.10. Training and testing

To ensure accuracy and reproducibility, the 160 oil samples (with five replicates each) were subdivided into two categories: 1) testing dataset and 2) training dataset. The testing dataset was obtained by randomly selecting 15% of the data and setting it aside to test our training model. The training dataset consisted of the remaining 85% of the data and was analyzed by SVM using the parameters that were identified during the model validation stages of data analysis. The accuracy of our method was than analyzed using the testing data set. The results were reported as a confusion matrix.

2.11. Comparison to physicochemical properties

To demonstrate that our fluorescence spectra data set have a relationship to the physical properties of the oil samples, we plotted the first four principal components (which each corresponded to at least 2% of the data, while the other four represented less than 2%) against the sample's percentile contents of the following structural elements: saturated hydrocarbons, aromatic hydrocarbons, resin compounds, and asphaltene hydrocarbons (SARA), obtained by TLC/FID as mentioned above.

A predicted percentile classification of these samples is also integrated into the device following the initial classification process.

3. RESULTS AND DISCUSSION

3.1. Initial results with 10 oil samples

Initially, our device was tested on a limited data set in order to see whether it was able to classify between different sample types and group similar samples together. This limited data set included: 1) three crude oil samples, two of which were very similar (from Iraq and Iran) and one that was quite different (from Australia); 2) two marine fuel oil (MFO) samples, which are the most common heavy fuel oil used by marine vessels; 3) three lighter fuel samples – Bunker A, marine gas oil (MGO), and marine diesel oil (MDO); and 4) two lubricating oil samples, one fresh and one used. The spectra we collected of these samples allows for the visual identification and classification of a number of samples (Fig. 3). For example, it is immediately apparent that the lubricating oils tend to exhibit lower florescence at most wavelengths, but this is especially true at higher wavelengths. The light fuel samples tend to exhibit less florescence at high wavelengths but are not as attenuated as lubricating oils. Heavy fuel oils (MFOs) behave similarly. Two crude oil samples from Middle Eastern countries behave similarly, while the other crude oil sample has a unique profile of its own. The MFO samples differ from the crude oil samples notably in the low wavelength regions where they exhibit less florescence.

Comparison with spectra obtained using a commercial spectrophotometer (Supplementary Figure S2) produces trends that are consistent with those obtained using our device (Fig 3). However, a notable discrepancy can be observed between 420 nm and 480 nm where our device exhibits little to no variation between samples and the commercial spectrophotometer exhibits robust variation. The lack of data provided by our device for those wavelengths is almost certainly a product of blue channel pixels in the Pi camera being over-saturated, generating flat responses after calibration. Such lack of data can be explained by the fact that they are close to the UV LED excitations (365, 375 and 385 nm), which are not at all monochromatic and leak into blue color. However, even without usable data from those wavelengths, sufficient fluorescence characteristics could be obtained in the longer wavelength blue color (480 – 500 nm), green color (500 – 550 nm), and yellow/red color (> 550 nm), which are sufficiently far away from the UV LED excitations. Our device still produces usable consistent spectra that correspond to sample variation.

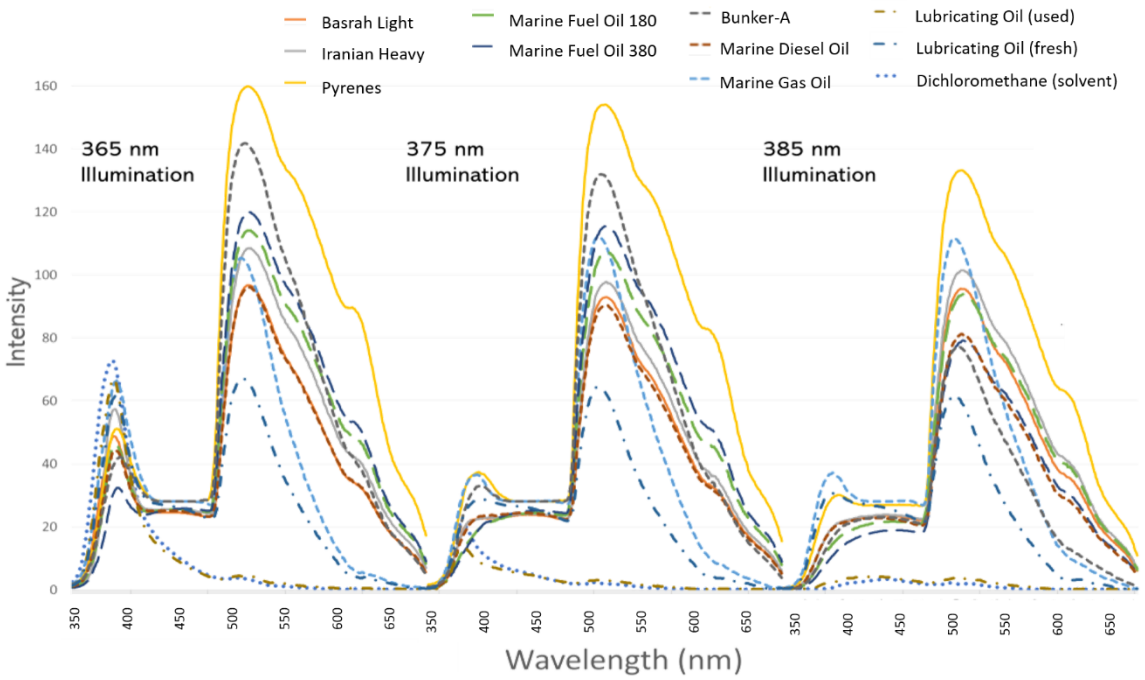


Fig. 3. Spectra obtained from 10 different oil samples and DCM. Each plot is the average intensity obtained from 10 replicates. Plot is divided into three regions each corresponding with a different excitation wavelength (365, 375, and 385 nm). DCM is plotted in dots. Crude oil samples are plotted in solid lines: Basrah Light (Iraq), Iranian Heavy, and Pyrenees (Australia). Heavy marine fuel oils (MFOs) are plotted in long dashes: MFO 180 and MFO 380, all originated from GS Caltex (Republic of Korea). Light marine fuel oils are plotted in small dashes: Bunker A, marine diesel oil (MDO) originated from SK Energy (Republic of Korea), marine gas oil (MGO) originated from GS Caltex. Lubricating oils (used and fresh) are plotted in dash-dots. DCM is plotted in dots.

3.2. Developing a training database

A training/testing database of 160 samples was collected using our device and the protocol provided in this work. This database included 82 crude oil samples, 16 heavy fuel oil samples, 47 light fuel oil samples ranging from the samples as light as gasoline all the way up to a single sample of Bunker B, and finally 13 lubricating oil samples. The spectral data from these samples was collected and visually inspected, however, because of the sheer number of samples it was impractical to visualize the raw spectra. To aid in sample visualization and later to assist in classification, samples were transformed to 10 dimensions using PCA. PCA weighting scores were plotted against wavelength for the first four principal components (PCs) to visualize which portions of the spectra give rise to each PC. Figure 4 shows the results.

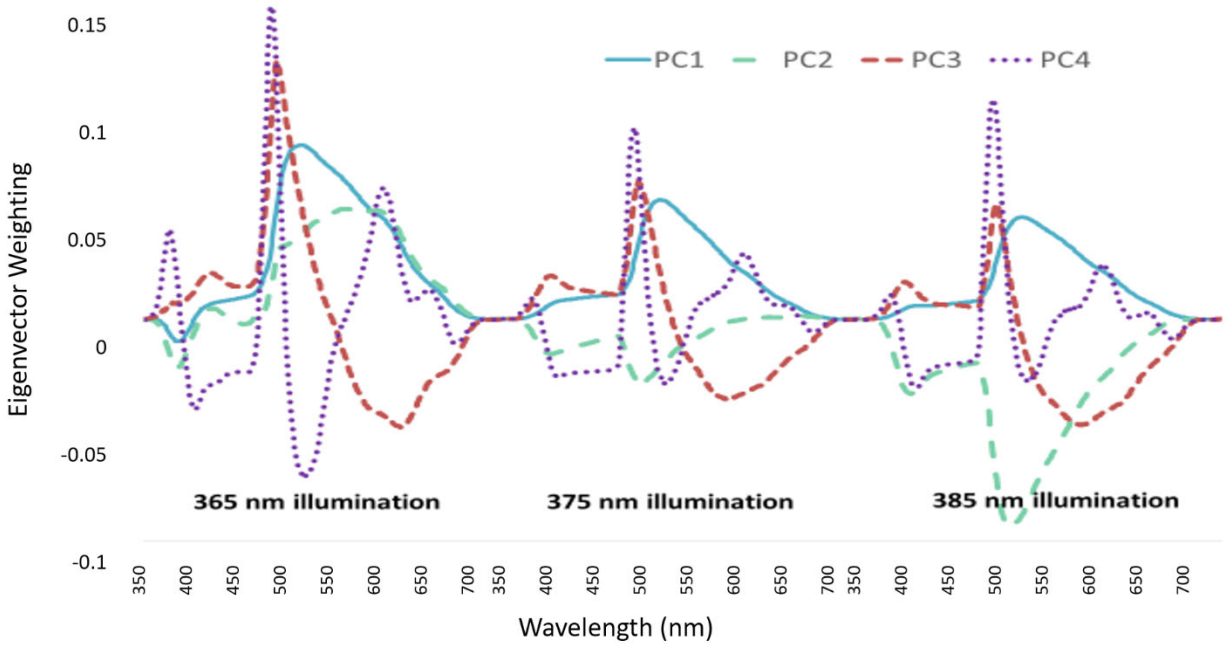


Fig. 4. PCA loading plot for the first four PCs of the 160-sample dataset.

All four PCs exhibited sufficiently significant wavelength-dependent behaviors, while the remaining PCs did not. This figure is also useful for visualizing which wavelengths contribute to which PCs.

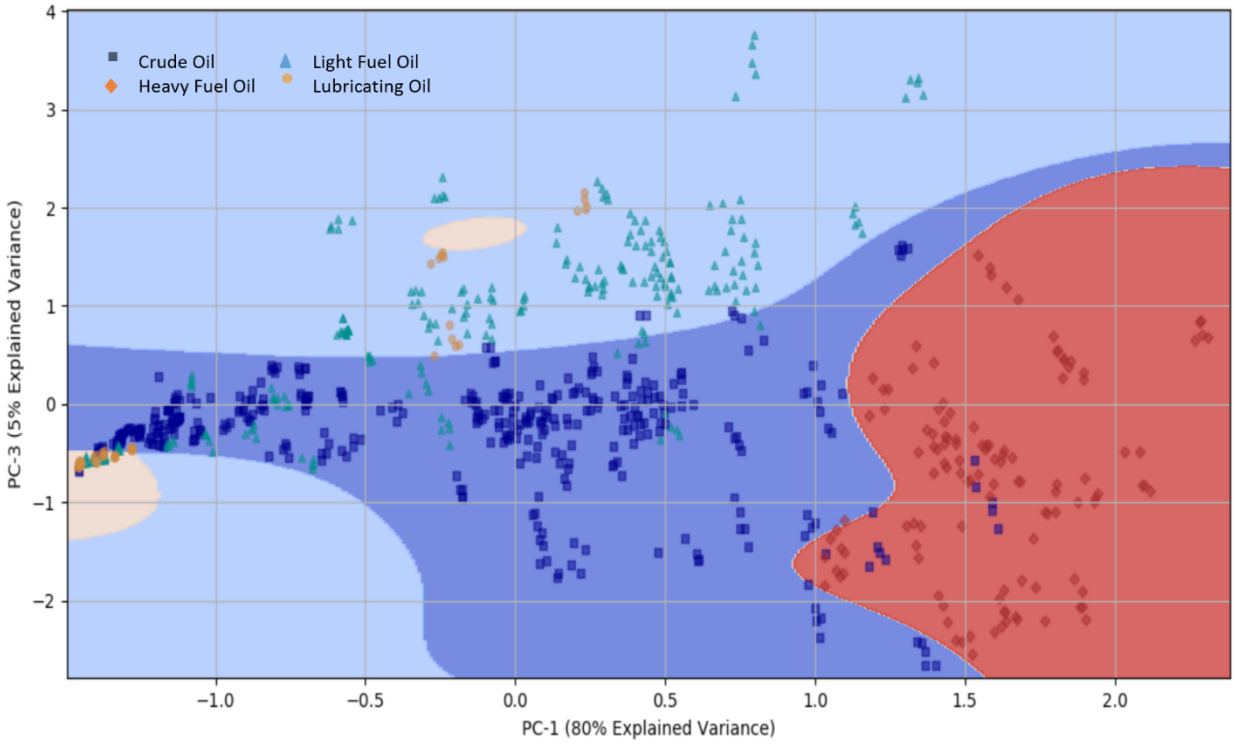
3.3. Sample classification

Samples were broadly classified into one of four oil types: 1) crude oil, 2) heavy fuel oil, consisting of Bunker C and marine fuel oils (MFOs), 3) light fuel oil, consisting of all other fuel oils, and 4) lubricating oil. Samples were visualized by PCA and classified by SVM using an RBF kernel. SVM parameters C and gamma were set to 100 and 0.1 respectively based on the findings of a grid-search. PCA rather clearly separates the different classes. The two principal components that seemed to demonstrate the clearest separation were PC-1 and PC-3, which together accounted for 85% of the variance in the data (Fig. 5 in the top). However, the SVM classification algorithm we employed makes use of all 10 principal components in our data.

Omitting 15% of our data to build a training model and using those 15% to test our model, we were able to a produce a confusion matrix (Fig. 5 in the bottom). Overall, we demonstrated a 95% correct classification into one of the four broad oil classifications.

Samples were further analyzed to determine whether samples that were purchased commercially differed from those that were obtained from marine vessels. Our 160-sample included both commercial and marine vessel samples within both the light fuel oil samples and heavy fuel categories. No significant variability was measured between samples of commercial vs. marine vessel origin (Supplementary Figure S3). However, commercially purchased heavy fuel oil samples did exhibit some interesting trends, most notably, a broader standard deviation and shifted average for the first four principal components relative to heavy fuel oil samples obtained from marine vessels. The most

reasonable explanations for this shift are as follows: Heavy fuel oil samples obtained from marine vessels are typically heated to improve flow, which could result in degradation of fluorescent compounds, resulting in a less varied spectrum. Alternatively, heavy marine fuel samples from vessels could have been treated with unknown chemicals either during use, collection, or transport. Regardless of the variation between samples based on collection origin, our model is still capable of classifying samples by oil type.



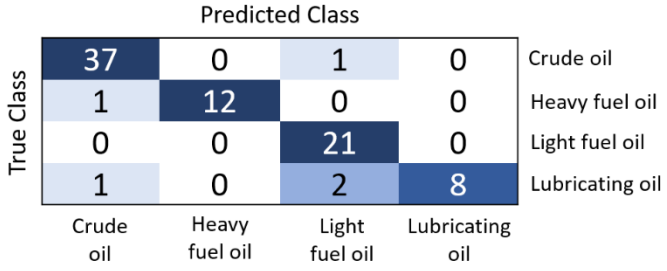


Fig. 5. Top: PCA plot and SVM decision boundaries for the 160-samples (5x replicates) training data set. Samples were categorized by SVM using an RBF kernel. Training data was classified into one of four broad categories: crude oil, heavy fuel oil (consisting of Bunker C and MFOs), light fuel oil (consisting of all other fuel oils), and lubricating oil. Bottom: Confusion matrix of testing dataset consisting of 15%of 160 sample training/testing dataset), showing 95% correct classification rate.

There are a couple regions in our PCA plots with notable sample overlap. The most prominent of these regions on Fig. 5 exists in the crude classification range where PC-1 is less than -0.2. Looking at the raw data from these samples we see this region corresponds to the samples that exhibit low florescence, calling into question our device's ability to properly classify these particular samples. As a

result, we are considering including a warning for the user that the classification predicted in this region has a reasonable chance of being a misclassification.

3.4. Physicochemical properties

Physiochemical significance of the principal components (PCs) was investigated by plotting the PCs against the percent content of saturates, aromatics, resins, and asphaltenes (SARA). Not all of these properties are predictable from each sample using our device, however, we are able to predict many of these properties for most samples. For example, crude samples largely follow a linear trend with asphaltene content for PC-1, although outliers do exist suggesting that the PC-1 does not exclusively represent the asphaltene content (Fig. 6A).

Light fuel oil follows linear trends between PC-1 and aromatic and saturate contents (Fig. 6B). However, this trend is largely a product of the light fuel oils being grouped into two discrete populations.

The heavy fuel oil samples follow the most significant linear trends with asphaltene contents and aromatic contents and PC-4 (Fig. 6C). These parameters also show less statistically significant trends with other principle components.

Additionally, all samples have trends between the asphaltene content and PC-3 (Fig. 6D). There also seems to be a trend for all samples (, except for lubricating oil, between the asphaltene content and PC-3, although this trend is somewhat tenuous.

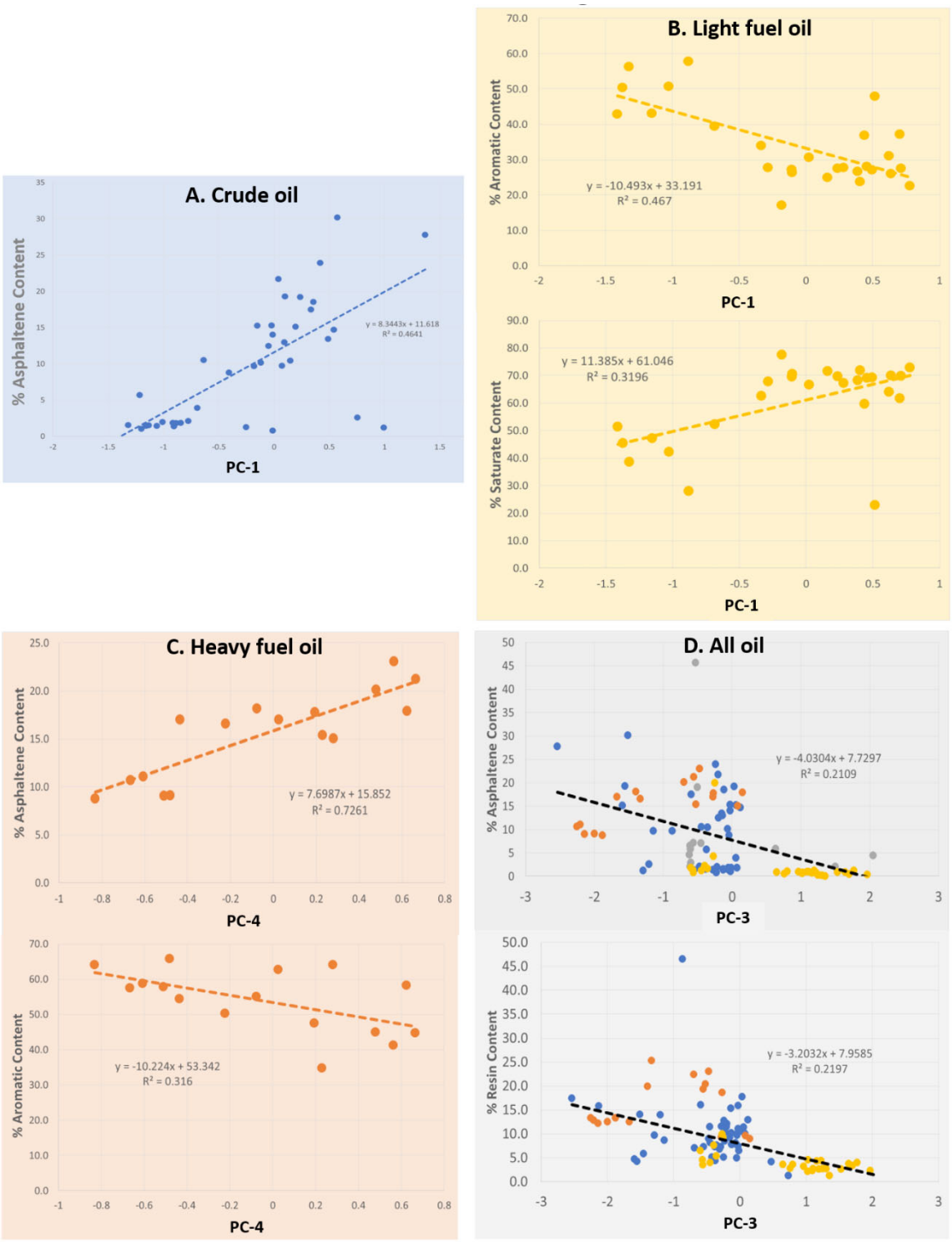


Fig. 6. Oil samples' physicochemical properties as predicted by PCA. A) Crude oil samples: asphaltene content vs. PC-1. B) Light fuel oil samples: aromatic and saturate contents vs. PC-1. C) Heavy fuel oil samples: asphaltene and aromatic contents vs. PC-4. D) All oil samples (except lubricating oils for the resin content plot): asphaltene and resin contents vs. PC-3. Crude oil = blue, light fuel oil = yellow, heavy fuel oil = orange, and lubricating oil = gray.

4. CONCLUSION

In conclusion, we designed a portable device appropriate for on-site oil testing. The device is capable of broadly classifying oil samples into one of four categories: crude oil, heavy fuel oil (consisting of Bunker C and marine fuel oils), light fuel oil (consisting of all other fuel types), and lubricating oil, with a 95% accuracy. Additionally, data classification into more discrete categories seems likely with more samples; however, some of the smaller datasets were too small to result in a more refined classification. Additionally, we can predict with some accuracy a sample's physiochemical content with regards to saturate, aromatic, resin and asphaltene (SARA) content percentages. The sample classification, as well as any pertinent percent compositions of SARA contents, is displayed to the user along with the predicted accuracy of those findings.

This platform promises to be a useful test for the rapid classification of oil samples in part due to its portability and simplicity of use. It should serve as a useful platform for regulatory agencies performing environmental forensic studies following an oil spill and during clean-up.

A great deal of future work is planned to improve, test, validate, and move the device from the benchtop to the field. The first step we plan to make is to improve the spectra obtained over the 420 – 480 nm range, through selectively attenuating the saturated signals. Additionally, we have plans to produce up to 10 of these devices to employ in a field setting. Towards that goal we will need to first test the batch reproducibility of our device. Once these steps have been taken, we plan to use the device to collect and test data from thousands of additional field samples and further verify and refine our reported method. Additionally, we have plans to produce up to 10 of these devices to employ in a field setting. Towards that goal we will need to first test the batch reproducibility of our device. Once these steps have been taken, we plan to use the device to collect and test data from thousands of additional field samples and further verify and refine our reported method.

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.

Author contributions

The overall concept was jointly conceived by J.-Y.Y. and U.H.Y, with input from M.V.B. The handheld spectrophotometer was designed, assembled and fabricated by M.V.B. with assistance from K.S. and B.T.N. and input from J.-Y.Y., U.H.Y. and A.L. The assay protocol was developed by M.V.B. and J.-Y.Y. All oil samples were collected and analyzed by A.L. with input from U.H.Y. and S.Y.H. Assay data were collected by M.V.B. and A.L. with assistance from K.S. and B.T.N. Machine learning analysis was conducted by M.V.B. Subsequent machine learning code and GUI were developed by M.V.B. and K.S. Data analysis was conducted by M.V.B. and J.-Y.Y. The manuscript was written by M.V.B. with assistance from J.-Y.Y. and input from all other others. All authors have given approval to the final version of the manuscript.

Acknowledgements

This research was a part of the project titled "Development of Advanced Oil Fingerprinting system (PN67490)" funded by the Korea Coast Guard, Republic of Korea.

Appendix A. Supplementary data

Supplementary data to this article can be found online.

References

Bergeon Burns, C.M., Olin, J.A., Woltmann, S., Stouffer, P.C., Taylor, S.S. 2014. Effects of oil on terrestrial vertebrates: Predicting impacts of the Macondo blowout. 64, 820-828. doi:10.1093/biosci/biu124

Brussaard, C.P.D., Peperzak, L., Beggah, S., Wick, L.Y., Wuerz, B., Weber, J., Arey, J.S., van der Burg, B., Jonas, A., Huis0man, J., van der Meer, J. R. 2016. Immediate ecotoxicological effects of short-lived oil spills on marine biota. Nat. Commun. 7, 11206. doi:10.1038/ncomms11206

Federici, C., Mintz, J. 2014. Oil properties and their impact on spill response options -- literature review. CNA: Alexandria. https://www.bsee.gov/sites/bsee.gov/files/osrr-oil-spill-response-research/ 1017aa.pdf, Accessed October 23, 2019.

Han, Y., Clement, T.P. 2018. Development of a field testing protocol for identifying Deepwater Horizon oil spill residues trapped near Gulf of Mexico beaches. PLOS ONE 13, e0190508. doi:10.1371/journal.pone.0190508

ITOPF. 2019. Past Spill Statistics. https://www.itopf.org/knowledge-resources/data-statistics/statistics, Accessed October 23, 2019.

Lee, Y.G., Garza-Gomez, X., Lee, R.M. 2018. Ultimate costs of the disaster: Seven years after the Deepwater Horizon oil spill. J. Corp. Account. Financ. 29, 69-79. doi:10.1002/jcaf.22306

Stephens, F.L., 2004. Thin layer chromatography - flame ionization detection analysis of in-situ petroleum biodegradation, M.S. thesis. Texas A&M University: College Station. https://oaktrust.library.tamu.edu/handle/1969.1/1039, Accessed October 23, 2019.

Stout, S.A. 2016. Oil spill fingerprinting method for oily matrices used in the Deepwater Horizon NRDA. Environ. Forensics 17, 218-243. doi:10.1080/15275922.2016.1177759

Suneel, V., Vethamony, P., Naik, B.G., Kumar, K.V., Sreenu, L., Samiksha, S.V., Tai, Y., Sudheesh, K. 2014. Source investigation of the tar balls deposited along the Gujarat coast, India, using chemical fingerprinting and transport modeling techniques. Environ. Sci. Technol. 48, 11343-11351. doi:10.1021/es5032213

Talley, W.K., Jin, D., Kite-Powell, H. 2004. Post OPA-90 vessel oil spill differentials: Transfers versus vessel accidents. Marit. Policy Manag. 31, 225-240. doi:10.1080/ 0308883042000209571

United States Census Bureau. 2019. Oil Spills in U.S. Water--Number and Volume. https://www.allcountries.org/uscensus/ 390_oil_spills_in_u_s_water.html. Accessed October 23, 2019.

Wang, Z. 2002. Using multiple criteria for fingerprinting unknown oil samples having very similar chemical composition. Environ. Forensics 3, 251-262. doi:10.1006/enfo.2002.0098

Wang, Z., Stout, S.A., Fingas, M. 2006. Forensic fingerprinting of biomarkers for oil spill characterization and source identification. Environ. Forensics 7, 105-146. doi:10.1080/ 15275920600667104

Yim, U.H., Short, J. 2017. Marine environmental emergencies in the North Pacific Ocean: Lessons learned from recent oil spills. Arch. Environ. Contam. Toxicol. 73, 1-4. doi:10.1007/s00244-017-0416-7

Zakaria, M.P., Horinouchi, A., Tsutsumi, S., Takada, H., Tanabe, S., Ismail, A. 2000. Oil pollution in the Straits of Malacca, Malaysia: Application of molecular markers for source identification. Environ. Sci. Technol. 34, 1189-1196. doi:10.1021/es990950o