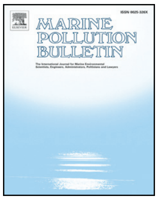
Marine Pollution Bulletin xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Marine Pollution Bulletin

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



Oil species identification technique developed by Gabor wavelet analysis and support vector machine based on concentration-synchronous-matrix-fluorescence spectroscopy

Chunyan Wang [a,b,c,d], Xiaofeng Shi [c], Wendong Li [c], Lin Wang [a], Jinliang Zhang [b], Chun Yang [d], Zhendi Wang [d]

[a] *Department of Physics and Electronic Science, Weifang University, Weifang 261061, China*
[b] *College of Resources Science and Technology, Beijing Normal University, Beijing 100875, China*
[c] *Optics and Optoelectronics Laboratory, Ocean University of China, Qingdao 266100, China*
[d] *Emergencies Science and Technology Section(ESTS), Science and Technology Branch, Environment Canada, 335 River Rd., Ottawa, Ontario K1A 0H3, Canada*

ARTICLE INFO

Article history:
Received 23 February 2015
Received in revised form 31 December 2015
Accepted 5 January 2016
Available online xxxx

Keywords:
Concentration-synchronous-matrix-fluorescence (CSMF)
Gabor wavelet analysis
Partial least square analysis (PLS)
Principle component analysis (PCA)
Supported vector machine (SVM)
Oil spill

ABSTRACT

Concentration-synchronous-matrix-fluorescence (CSMF) spectroscopy was applied to discriminate the oil species by characterizing the concentration dependent fluorescence properties of petroleum related samples. Seven days weathering experiment of 3 crude oil samples from the Bohai Sea platforms of China was carried out under controlled laboratory conditions and showed that weathering had no significant effect on the CSMF spectra. While different feature extraction methods, such as PCA, PLS and Gabor wavelet analysis, were applied to extract discriminative patterns from CSMF spectra, classifications were made via SVM to compare their respective performance of oil species recognition. Ideal correct rates of oil species recognition of 100% for the different types of oil spill samples and 92% for the closely-related source oil samples were achieved by combining Gabor wavelet with SVM, which indicated its advantages to be developed to a rapid, cost-effective, and accurate forensic oil spill identification technique.

1. Introduction

Oil spill occurs every day worldwide and causes extensive damage to marine life, terrestrial life, human health, and natural resources (Wang and Fingas, 2003).

Disasters like the Gulf of Mexico (USA) oil spill and Dalian (China) oil spill are guaranteed to hit the headlines, while small oil spills often go unnoticed and unreported. But with high rate of occurrence and low rate of source identification, these small-scale incidents could be just as damaging to the environment as large spills (O'Hara et al., 2009).

Tiered analytical approaches form the basis of modern oil hydrocarbon fingerprinting. Initial screening of oil samples is generally the first tier in fingerprinting protocols to determine the oil type and narrows down the number of possible matches for a spill sample (Christensen and Tomasi, 2007.). Fluorescence spectroscopy (Bugden et al., 2008; Zhou et al., 2013a, 2013b), infrared spectroscopy (Fernández-Varela et al., 2009; Fresco-Rivera et al., 2007), and gas chromatography/flame ionization detector (GC/FID) (Jagoš et al., 2014; Wang et al., 2009) are the main techniques for initial screening of oil samples.

As for detailed fingerprinting, gas chromatography/mass spectrometry (GC/MS) (Yang et al., 2009; Wang et al., 2002) analysis, which has a high resolving power and is able to separate and quantify thousands of individual hydrocarbon compounds, constitutes a cornerstone of modern oil spill fingerprint and has been taken as the most reliable method to collect forensic evidence (Fernández-Varela et al., 2008).

Nevertheless, it is broadly accepted that no single method can undoubtedly assign a hydrocarbon spilled into the environment to its potential sources because each method has its bottlenecks to break through (Fernández-Varela et al., 2010).

Fluorescence-based techniques feature high sensitivity, good diagnostic potential, relatively simple instrumentation, easy sample preparation, and less affected by weathering. But for crude oil samples, usually the shapes of the spectra are more sensitive to the concentration than the oil species, thus the routine fluorescence fingerprinting techniques, such as excitation–emission-matrix fluorescence (EEMF) (Divya and Mishra, 2007; Zhou et al., 2013a, 2013b) and synchronous fluorescence (SF) (Abbas et al., 2006), have relatively lower accuracy without consideration of the nonlinear effect between the shape of spectra and the certain concentration values. For complex mixtures such as petroleum related samples, the discrepancy of the fluorescence spectra between different concentration values is more obvious than between the different but closely related source oil. The results of oil species identification based on fluorescence spectra can be distorted by the inevitable errors of the solution concentration values. In particular, the concentration values

http://dx.doi.org/10.1016/j.marpolbul.2016.01.001
0025-326X/© 2016 Elsevier Ltd. All rights reserved.

Please cite this article as: Wang, C., et al., Oil species identification technique developed by Gabor wavelet analysis and support vector machine based on concentration-synchro..., Marine Pollution Bulletin (2016), http://dx.doi.org/10.1016/j.marpolbul.2016.01.001

of the oil spill samples are always difficult to determine in the marine environment due to the seawater adulteration. That is the main reason why the fluorescence spectroscopy hasn't been employed independently and broadly into the oil spill identification even though the petroleum related samples have abundance of PAHs with high fluorescence emission and weathering resistance. Fortunately more and more researchers have noticed this phenomenon and have been probing into the ways of solving this problem (Patra and Mishra, 2001; Patra, 2003; Patra et al., 2012; Smith and Sinski, 1999).

By introducing concentration value as a new dimension to the synchronous fluorescence spectroscopy, the Concentration-synchronous-matrix-fluorescence (CSMF) (Wang et al., 2010a, 2010b) spectroscopy is formed, which expanded the fluorescence information from lower ring to higher ring PAHs at different levels of concentration. These concentration dependent spectra can accumulate the differences of the species and better reflect the chemical components information of oil. Here, the discrimination criterion using for oil species identification is the red shift tendency of CSMF along with the increase of concentration values rather than a certain synchronous fluorescence spectrum at certain fixed concentration value. This CSMF method eliminates the errors of concentration uncertainty inevitably caused by seawater adulation.

Preliminary studies of the comparison of the identification accuracy of SFS, EEM and CSFM have indicted that this newly CSMF spectroscopy

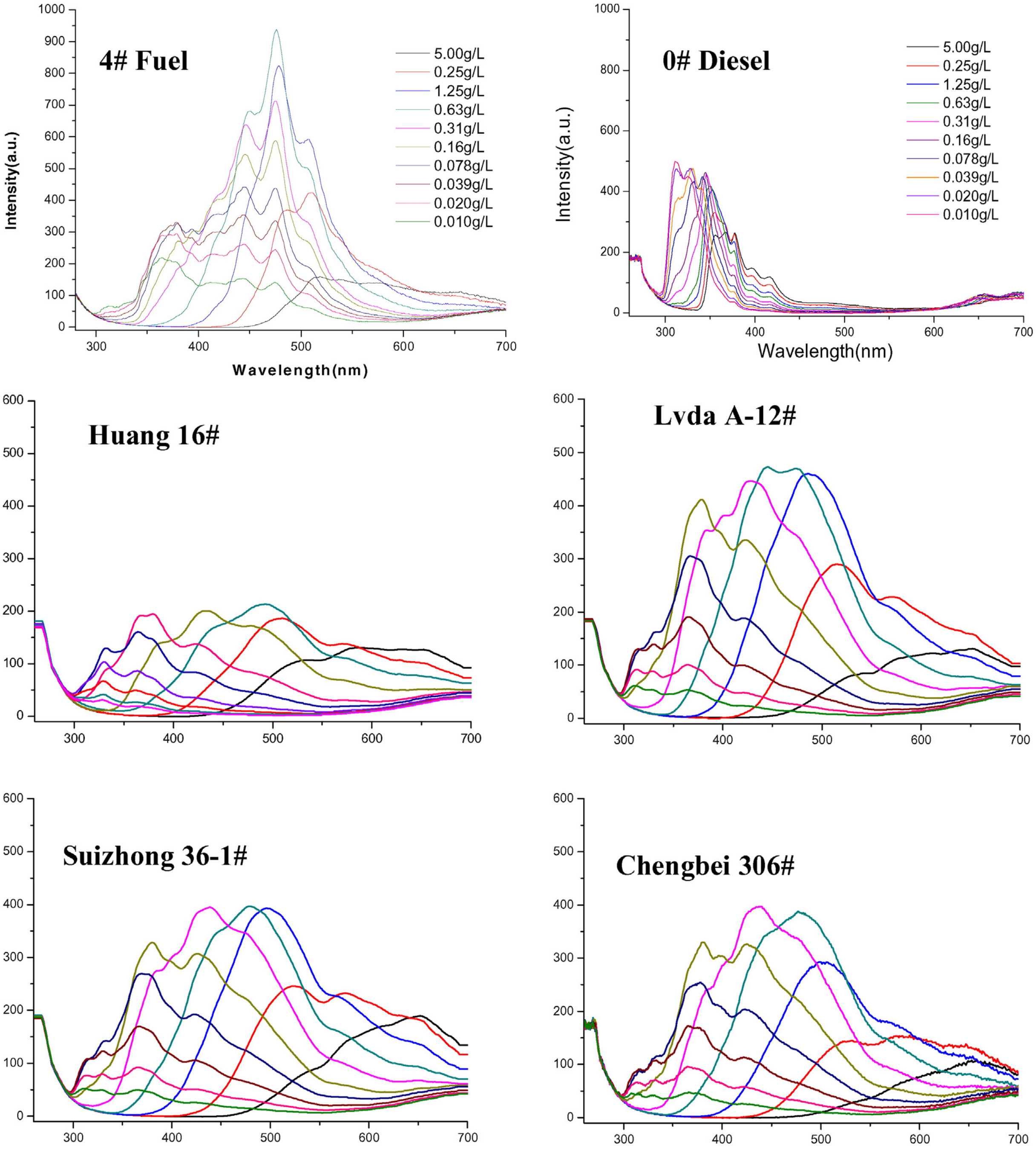


Fig. 1. The respective concentration dependent synchronous fluorescence spectra of petroleum related samples (4# Fuel, 0# Diesel, Huang 16#, Chengbei 306#, Lvda A-12#, Suizhong 36–1#).

Please cite this article as: Wang, C., et al., Oil species identification technique developed by Gabor wavelet analysis and support vector machine based on concentration-synchro..., Marine Pollution Bulletin (2016), http://dx.doi.org/10.1016/j.marpolbul.2016.01.001