Journal of Hydrology 537 (2016) 221–233

Contents lists available at ScienceDirect

Journal of Hydrology

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

Classification of river water pollution using Hyperion data

Soumyashree Kar [a], V.S. Rathore [a,*], P.K. Champati ray [b], Richa Sharma [b], S.K. Swain [c]

[a] Department of Remote Sensing, Birla Institute of Technology, Mesra, Ranchi, Jharkhand 835215, India
[b] Indian Institute of Remote Sensing, #4, Kalidas Road, Dehradun, Uttarakhand 248001, India
[c] Central Instrumentation Facility, Birla Institute of Technology, Mesra, Ranchi, Jharkhand 835215, India

ARTICLE INFO

Article history:
Received 25 December 2015
Received in revised form 18 March 2016
Accepted 21 March 2016
Available online 29 March 2016
This manuscript was handled by Geoff Syme, Editor-in-Chief

Keywords:
Hyperspectral remote sensing
Spectral band depth analysis
Image classification
Image spectra
River water
Metal pollutants

SUMMARY

A novel attempt is made to use hyperspectral remote sensing to identify the spatial variability of metal pollutants present in river water. It was also attempted to classify the hyperspectral image – Earth Observation-1 (EO-1) Hyperion data of an 8 km stretch of the river Yamuna, near Allahabad city in India depending on its chemical composition. For validating image analysis results, a total of 10 water samples were collected and chemically analyzed using Inductively Coupled Plasma-Optical Emission Spectroscopy (ICP-OES). Two different spectral libraries from field and image data were generated for the 10 sample locations. Advanced per-pixel supervised classifications such as Spectral Angle Mapper (SAM), SAM target finder using BandMax and Support Vector Machine (SVM) were carried out along with the unsupervised clustering procedure – Iterative Self-Organizing Data Analysis Technique (ISODATA). The results were compared and assessed with respect to ground data. Analytical Spectral Devices (ASD), Inc. spectroradiometer, FieldSpec 4 was used to generate the spectra of the water samples which were compiled into a spectral library and used for Spectral Absorption Depth (SAD) analysis. The spectral depth pattern of image and field spectral libraries was found to be highly correlated (correlation coefficient, $R^2 = 0.99$) which validated the image analysis results with respect to the ground data. Further, we carried out a multivariate regression analysis to assess the varying concentrations of metal ions present in water based on the spectral depth of the corresponding absorption feature. Spectral Absorption Depth (SAD) analysis along with metal analysis of field data revealed the order in which the metals affected the river pollution, which was in conformity with the findings of Central Pollution Control Board (CPCB). Therefore, it is concluded that hyperspectral imaging provides opportunity that can be used for satellite based remote monitoring of water quality from space.

1. Introduction

Water management initiatives require identification, estimation and assessment of several pollutants deteriorating the water quality. Due to rapid growth of population and industrialization, the domestic and industrial effluents often meet a river course untreated thereby increasing the toxicity levels of river water. It has been found that in the Indian Sub-Continent region such polluted rivers are quite significant on the eastern coasts (except the Himalayan region where most of the rivers are perennial and highly discharged ones during monsoon) than the western coasts (Prasad and Singh, 2010). The Yamuna river is an eastern flowing important river of India and is the most polluted river. Therefore, there is an urgent need to monitor and assess the river water quality to understand the impact of development and industrialization.

Satellite remote sensing has been in use since long for water quality assessment based on several water quality parameters. Space-borne MODIS data have been found useful to identify the spatial and seasonal distribution of chlorophyll concentrations (identified in the visible region of the spectrum) across the global oceans (Dasgupta et al., 2009). Singh et al. (1997) have also identified changes in and around a lake using the 400–900 nm range of the CASI (Compact Airborne Spectrographic Imager) and associated the radiance anomalies to the atmospheric aerosol. However, the most significant technology development in remote sensing is the launch of hyperspectral sensors, which provide us a spectral hyper-cube enabling the distinction, identification and classification of spectrally unique elements. These imaging spectrometers or hyperspectral sensors also provide a unique combination of both spatially and spectrally contiguous images of the earth's surface unavailable from other sources for monitoring Water Quality

* Corresponding author. Mobile: +91 9431382641.
E-mail addresses: karsoumya23@gmail.com (S. Kar), vsrathore@bitmesra.ac.in (V.S. Rathore), champati_ray@iirs.gov.in (P.K. Champati ray), richa@iirs.gov.in (R. Sharma), sanjayarkl@gmail.com (S.K. Swain).

http://dx.doi.org/10.1016/j.jhydrol.2016.03.047
0022-1694/

Parameters (WQPs) (Jupp et al., 2002; Ritchie et al., 2003). Hyperspectral sensors provide a means of extended hyperspectral mapping capability to areas not accessible to multispectral sensors; like information on chemical composition of minerals or the types of algae or Total Suspended Solids (TSS) present in water (Brando and Dekker, 2003).

Although designed as a technical demonstration for land applications, a paradigm development in hyperspectral remote sensing was realized when Hyperion was tested over a range of water targets in Eastern Australia (Brando and Dekker, 2003). The shorter the wavelength the higher is the penetration in clear water, therefore, both wavelength and water characteristics are key to the process of water quality assessment using spectral techniques (Abou El-Magd and El-Zeiny, 2014). As a result, many studies have been carried out using hyperspectral imaging techniques to establish WQPs such as estimation of chlorophyll, turbidity, Total Dissolved Solids (TDS), Colored Dissolved Organic Matter (CDOM) and Secchi Disk Depth (SDD) (Brando and Dekker, 2003; Chakravortty and Chakrabarti, 2011; Bhatti et al., 2010a; Bhatti et al., 2010b). Koponen et al. (2002) have also performed lake water classification for water quality parameters like SDD, turbidity and chlorophyll a. However, all these techniques have exploited only the electromagnetic spectrum range of 400–900 nanometers (nm), whereas Hyperion sensor provides data at 10 nm spectral resolution covering a wider spectral range from 357 to 2577 nm, enabling identification of several spectral absorption features. The Hyperion data has sufficient sensitivity to detect optical water quality (Shafique et al., 2001; Brando and Dekker, 2003), besides its overwhelming significance in geologic mapping and mineral abundance estimation. Therefore, it has prompted to carry out both qualitative and quantitative analysis of the Short Wave Infrared (SWIR) range of the target spectra for possible applications in water resources (Kruse et al., 2003a, 2003b; van der Meer, 2004). Moreover, it has been clearly stated in Hakvoort et al. (2002) that water quality parameters can be successfully retrieved from remotely sensed data using the subsurface reflectance spectra.

The use of hyperspectral satellite imagery for detecting river water pollution is quite challenging due to its low signal-to-noise ratio (Jupp et al., 2002). Hence, a detailed knowledge of water quality parameters and spectral information is required, involving both in-situ measurements and hyperspectral remote sensing techniques (Richardson and LeDrew, 2006; Chakravortty and Chakrabarti, 2011). Spectral library and band-depth analysis have revealed that quantitative assessment of target can be done which extracts sub-pixel information as against pixel based classification (Chakravortty and Chakrabarti, 2011). Since the amount of absorption at a certain wavelength depends on the quantity of the absorbing feature, the pattern of change in spectral depth values and concerned chemical composition can be correlated. It has been observed that the spectral signatures of the sampling points strongly correlate with the optically active constituents present within the water body (Bhatti et al., 2010a). A detailed analysis of the remotely sensed hyperspectral data has also helped to develop band ratios and regression algorithms that is based on good correlation between spectral signatures and the water quality parameters like Total Suspended Solids (TSS) and Colored Dissolved Organic Matter (CDOM) (Gitelson et al., 1993; Bhatti et al., 2010b).

The present study envisages the detection of metal pollutants using Hyperspectral data in river water with the spectral analysis of absorption features beyond the optical range. To achieve the objective, various image processing techniques were applied. Then, the image classifications (SVM and ISODATA) based on both spectral and spatial properties were carried out to classify the water surface. The approach involves the use of per-pixel supervised procedures which reduce data for further analysis with enhanced contrast among different classes (Shwetank et al., 2012). SAM target finder using BandMax being an enhanced algorithm over SAM was preferred, because it identifies the bands that are best able to distinguish the targets from background (Parks, 2006). SVM classified the image using regions of interest while ISODATA iteratively clustered spectrally identical classes. Band depth analysis was employed to identify the metal pollutants, and map the classes to different levels of pollution.

2. Study area and dataset

An 8 km stretch of the river Yamuna extending between coordinates 25°21′21.26″N, 81°38′32.20″E to 25°21′1.73″N, 81°41′53.73″E near Allahabad city was selected for the study (Fig. 1). The river starts its journey from the glacier Yamunotri flowing downstream through major cities (Delhi, Mathura, Agra, etc.), where it collects most of the toxic pollutants i.e. 90% of its solid wastes, until it joins the river Ganga near Allahabad (Karsauliya, 2013). It is the longest tributary of the river Ganga and travels around 1127 km before joining it in Allahabad, India. It has a catchment area of 366,223 km^2. Almost 60 million people of 10 major cities on its bank and nearby villages depend on its water. According to the Centre for Science and Environment, approximately 75–80% of the river's pollution is the result of raw sewage, industrial runoff and the garbage thrown into the river and it totals over 3 billion liters of waste per day (Misra, 2010).

The study area is characterized by extensive crop lands and very sparse settlement along the banks of the river Yamuna. The excessive pollution and alarming toxicity in the river water, rules it out from underwater life or domestic usage, deeming it fit only for recreation and industrial cooling purposes (with pH between 6.0 and 8.5) thereby belonging to class E (Misra, 2010; CPCB, 2008a). On realizing the growing pollution level in the river Yamuna, water quality monitoring of it was started in 1976 by CPCB with 18 stations (CPCB, 2008b). Subsequently, it has been found that the presence of toxic and heavy metal pollution like Fe, As, Cd, Hg, Zn, Ca, Mg which have been identified in the river water and can be primarily attributed to factors like industrial mining, salinity due to irrigation and urban pollution (CPCB, 2008a, 2008b).

Along the river course in the study area the major sources of pollutants are primarily identified as the discharge of domestic wastes through drainage channels, sand mining and urban pollution shown in Fig. 1b.

3. Materials and methods

3.1. Materials

In the present analysis L1R and L1T Hyperion datasets (ID EO1H1430422013074110KF) with path and row numbers 143 and 42 respectively, 30 m spatial and 10 nm spectral resolution covering 357–2577 nm wavelength range were downloaded from http://earthexplorer.usgs.gov/ (Fig. 2a and b). L1R Hyperion image of March 15, 2013 was used for image analysis, while L1T served as the reference image for geo-registering the raw data (L1R). 30 Ground Control Points (GCPs) were used for image-to-image georegistration using nearest neighbor re-sampling method. With a spatial distribution of approximately 0.5 km, the Root Mean Square (RMS) error was found to be 1.186263. Ten sampling locations were identified on the image considering the reflectance characteristics of the image and spacing between the sample locations to ensure proper distribution. Water samples were collected by following standard sample collection protocol and guidelines given in Indian Standards methods IS: 3025 (Part-1) and American Public Health Association (APHA) 22nd edition for determination of