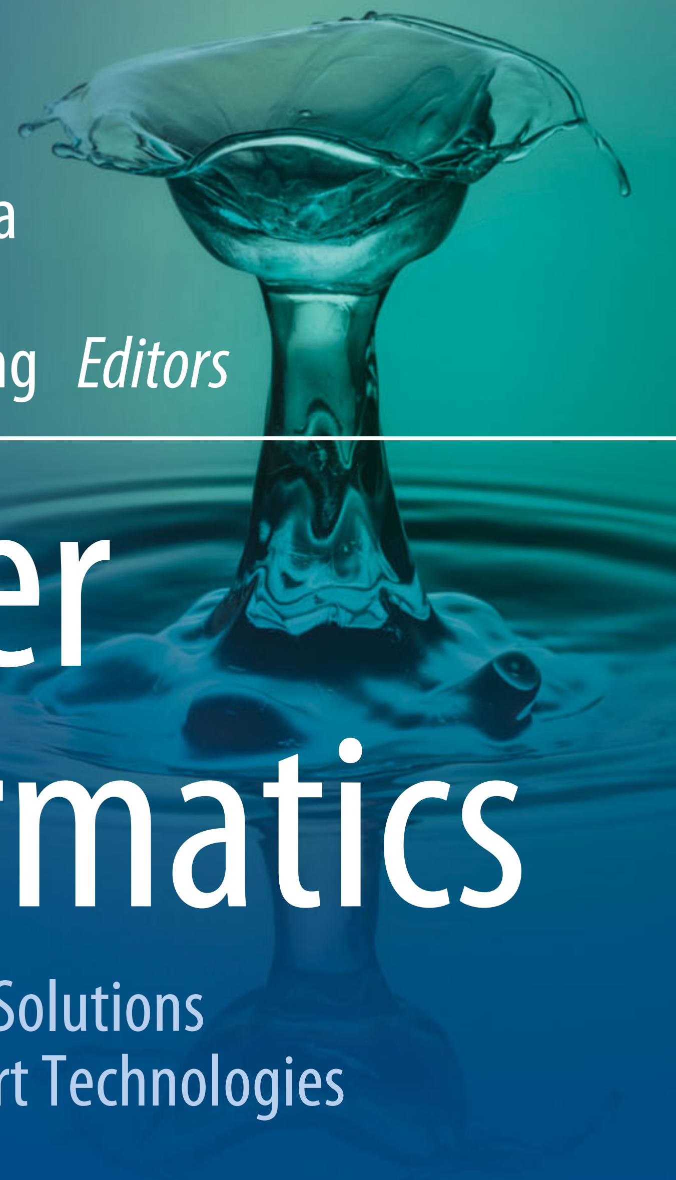
Water Informatics for Water Resource Management



Supreeti Kamilya
Arindam Biswas
Sheng-Lung Peng *Editors*

Water Informatics

Challenges and Solutions Using State of Art Technologies

Springer

Water Informatics for Water Resource Management

Series Editors

The primary focus of the book series is to understand digital technologies and solutions thereby enabling them to correlate the solutions with their region-specific problems and finally apply them to the Water Resource Management. This will help to generate novel research and innovative solutions. This domain of Water Informatics is growing at a very fast rate and its potential has been realized since it harnesses the use of digital technologies to leverage the applications offered for Water Resource Management.

The book series is tailored to meet the following learning objectives:

- Learn Machine Learning Techniques for data-driven modeling for management of Water Utilities.
- Learn IOT technologies for Smart Water Management.
- Learn Artificial Neural Networks for decision making in Water Utilities.
- Learn Risk Analysis, Fuzzy Logic and management of uncertainty in Water Utilities.
- Learn Block-Chain Technology for effective allocation of Water Resources.
- Learn Cyber-Security essential techniques for Smart Water Services.
- Learning Evolutionary Methods for Water Management.

Supreeti Kamilya · Arindam Biswas ·
Sheng-Lung Peng
Editors

Water Informatics

Challenges and Solutions Using State of Art Technologies

Springer

Editors
Supreeti Kamilya
Department of Computer Science
and Engineering
Birla Institute of Technology Mesra
Ranchi, India

Arindam Biswas
Department of Mining Engineering
Kazi Nazrul University
Asansol, West Bengal, India

Sheng-Lung Peng
Department of Creative Technologies
and Product Design
National Taipei University of Business
Taipei, Taiwan

ISSN 2731-9105 ISSN 2731-9113 (electronic)
Water Informatics for Water Resource Management
ISBN 978-981-97-1517-6 ISBN 978-981-97-1518-3 (eBook)
https://doi.org/10.1007/978-981-97-1518-3

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Water holds immense significance in human civilization, playing a pivotal role since its inception, particularly in the realm of agriculture. The availability of water has been a foundational element for the development of human societies. Its importance extends to a multitude of applications, from household needs to large-scale industrial operations. Furthermore, advancements in technology have greatly facilitated various water treatment methods. In this context, the role of computations becomes paramount. Consequently, data related to water are essential for these computations and the broader spectrum of water treatment processes. This data serves as a critical tool for analyzing various water-related issues, including the assessment of water quality in both surface and groundwater sources, as well as the utilization of water in diverse fields and sectors.

Water informatics is a multidisciplinary field focused on the gathering, examination, and retention of water-related data. With the recent advancements in computer technology, researchers worldwide have been able to offer a range of solutions for managing water resources in various domains, including hydrology, oceanography, and meteorology. Since the early 2000s, there has been a substantial increase in the availability of water science data on the Internet. Utilizing this data allows for the generation of valuable insights pertaining to cutting-edge water resource technologies. These water-related challenges encompass issues like waterlogging, groundwater contamination, flood prediction and mitigation, water quality monitoring, water body identification, and more. The copious amounts of data accessible online are harnessed by various tools, such as machine learning, deep learning, the Internet of Things (IoT), cellular automata, soft computing, and more. These tools analyze the data and provide solutions to address these water-related challenges.

The book's objective is to leverage existing data resources and address challenges related to surface and groundwater using cutting-edge technologies. It encompasses the study of water informatics, which may include the utilization of satellite imagery data at various spatial resolutions to detect water bodies. The studies on water body extraction methods include NDWI, MNDWI, machine learning classifiers, and industrial IoT-enabled techniques. The book covers a wide spectrum of concerns. Surface water is critical for human sustainability, while underground water management is equally vital for human survival and various industries. The monitoring of surface and underground water bodies is essential for pollution control, preserving the environment and ecology. The book also addresses a mathematical tool, cellular automata-based technology that can be applied for regulation of water demand and water flow prediction. Additionally, it delves into the simulation of water distribution system using deep learning approaches. It addresses water infrastructure planning based on social media data. In essence, this book provides valuable guidance to researchers interested in tackling a diverse array of water-related challenges and harnessing emerging technologies in the modern era of computation.

Ranchi, India — Dr. Supreeti Kamilya
Asansol, India — Dr. Arindam Biswas
Taipei, Taiwan — Dr. Sheng-Lung Peng

Contents

Editors and Contributors

About the Editors

Dr. Supreeti Kamilya has been working as an Assistant Professor in the Department of Computer Science and Engineering since December 2021. She received her master's degree in computer science and Engineering from IIEST, Shibpur with 1st rank in 2015. She completed her Ph.D. from IIEST Shibpur in 2021. She works on various research topics such as Theoretical Computer Science (Cellular Automata), Chaos Theory, Machine Learning, and Deep Learning. She has several journal and conference publications with International repute. She was awarded the Visvesvaraya Fellowship during her Ph.D. She was a Visiting Fellow at the University of Turku, Finland. She was awarded best paper award different international conferences. She has acted as a Reviewer of different reputed journals.

Dr. Arindam Biswas received M.Tech. degree in Radio Physics and Electronics from the University of Calcutta, India, in 2010, and a Ph.D. from NIT Durgapur in 2013. He has worked as a Post-doctoral Researcher at Pusan National University, South Korea, with a prestigious BK21PLUS Fellowship, Republic of Korea. He has received a DST-JSPS Invitation fellow at RIE, Japan, a DST-ASEAN Invitation Fellow at Duy Tan University, Vietnam, and Taylors University, Malaysia, Visiting Fellow to the Department of Electrical and Computer Engineering, National University Singapore. He has worked as a Specially Appointed Associate Professor (Visiting) at the Research Institute of Electronics, Shizuoka University, Japan. Presently, he is working as an Associate Professor in the School of Mines and Metallurgy at Kazi Nazrul University, WB, India where he has been actively engaged in teaching, research, and administration. He has 53 technical papers in different journals and 55 conference proceedings eight authored books, 18 edited volumes, and 10 book chapters with international repute.

Sheng-Lung Peng is a Professor at the Department of Creative Technologies and Product Design, and the Dean of the College of Innovative Design and Management at National Taipei University of Business in Taiwan. He received a Ph.D. degree from the Computer Science Department of National Tsing Hua University in Taiwan. He is an honorary Professor at Beijing Information Science and Technology University and a visiting Professor at Ningxia Institute of Science and Technology in China. He is also an adjunct Professor at National Dong Hwa University in Taiwan and, Mandsaur University and Kazi Nazrul University in India. Dr. Peng has edited several special issues in journals, such as *Frontiers in Public Health*, *Journal of Internet Technology*, *IEEE Internet of Things Magazine*, *Computers and Electrical Engineering*, Sensors, and so on. His research interests are algorithm design in the fields of artificial intelligence, bioinformatics, combinatorics, data mining, and networking.

Contributors

Pelin Angin Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

Arunima Asthana Department of Management, Birla Institute of Technology, Mesra, Ranchi, India

Ravi Kiran Bagadi Department of Computer Science & Engineering, Gitam (Deemed to Be University), Visakhapatnam, Andhra-Pradesh, India

Bannishikha Banerjee Amity Institute of Information Technology, Amity University, Kolkata, India

Tanmoy Kr. Banerjee Department of Management, Birla Institute of Technology, Mesra, Ranchi, India

Ranjan Kumar Behera Department of Computer Science and Engineering, Indian Institute of Information Technology, Ranchi, Jharkhand, India

Debnath Bhattacharyya Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India

Partha Sarathi Bishnu Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India

Shuvendu Chakraborty St. Xavier's University, Kolkata, West Bengal, India

Sonia Djebali Léonard De Vinci, Research Center, Paris La Défense, France

Osvaldo Gervasi Department of Mathematics and Computer Science, University of Perugia, Perugia, Italy

Dipayan Hazra Amity Institute of Information Technology, Amity University, Kolkata, India

E. Stephen Neal Joshua Department of Computer Science & Engineering, Gitam (Deemed to Be University), Visakhapatnam, Andhra-Pradesh, India

Tai-hoon Kim School of Electrical and Computer Engineering, Yeosu Campus, Chonnam National University, Yeosu-Si, Jeollanam-Do, Republic of Korea

Anisha Kumari Department of Computer Science and Engineering, National Institute of Technology, Rourkela, India

Sriparna Mandal Department of Computer Science and Technology, Indian Institute of Engineering Science and Technology, Howrah, West Bengal, India

N. Marline Joys Kumari Department of Computer Science and Engineering, Anil Neerukonda Institute of Technology and Sciences, Visakhapatnam, AP, India

Jit Mukherjee Birla Institute of Technology Mesra, Mesra, India

Partha Kumar Mukherjee The Neotia University, Sarisha, West Bengal, India

Soumyadeep Paty Kazi Nazrul University, Asansol, West Bengal, India

P. K. Paul Department of CIS, and Information Scientist (Offg.), Raiganj University, Raiganj, West Bengal, India

T. Pavan Kumar Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India

K. Prasad Department of Computer Science & Engineering, Gitam (Deemed to Be University), Visakhapatnam, Andhra-Pradesh, India

Bibhudatta Sahoo Department of Computer Science and Engineering, National Institute of Technology, Rourkela, India

Dhrubasish Sarkar Supreme Institute of Management and Technology, Hooghly, West Bengal, India

Sutapa Sarkar Brainware University, Kolkata, West Bengal, India

P. Srinivas Department of Computer Science and Engineering, Malla Reddy Engineering College, Hyderabad, TG, India

N. Thirupathi Rao Department of Computer Science and Engineering, Vignan's Institute of Information Technology, Visakhapatnam, Andhra Pradesh, India

Identifying the Changes of Mine Water Bodies from Landsat 8 OLI Images in Automated Manner: A Case Study in Jharia, India

Jit Mukherjee

Abstract Identifying and monitoring water bodies have been active research areas because of their multi-fold effects on the environment and society. Water bodies are detected in the literature by several indexes such as normalized difference water index (NDWI), modified normalized water body index (MNDWI), automated water extraction index (AWEI), and others from multi-spectral satellite images. Identifying and separating different types of water bodies using such indexes have been found complex due to their multiple homogeneous features. The high mineral abundance of the surrounding regions can be the distinguishing attribute of a mine water body. This idea has been used in the past to separate mine water bodies from other kinds of water bodies. However, a certain limitation has been reported as other high mineral abundance regions can be present close to a water body, which are not associated with mining. Further, these mine water bodies change frequently due to their uses. Monitoring such water bodies has several applications in the mining industry, water pollution, and health. Hence, automated detection of the changes in mine water bodies needs more extensive attention. In this work, this research gap has been addressed. First, mine water bodies are separated. Further, their translation, rotation, and shearing changes are computed using the coherent point drift technique.

Keywords Mine water bodies · Change detection · Hausdorff distance · Coherent point drift · Landsat 8 · AWEI

1 Introduction

Water is the essential conduit through which our civilization has grown. Rivers and other sources of water sustain our life. Thus, the management of water bodies has been an essential part for a long period. Among different kinds of water bodies, mine water bodies become perilous to the environment because of acid mine drainage, heavy metal contamination and leaching, and pollution by chemical agents. Fresh

J. Mukherjee (✉)
Birla Institute of Technology Mesra, Mesra, India
e-mail: jit.mukherjee@bitmesra.ac.in

S. Kamilya et al. (eds.), *Water Informatics*, Water Informatics for Water Resource Management, https://doi.org/10.1007/978-981-97-1518-3_1

water is heavily used in the mining process. Mine affluent and seepage from different mining regions especially tailing dams may directly merge into rivers and other water bodies, whose water is being used on a daily basis by civilians. Such water pollution has endless effects on human health, society, and the environment. Hence, monitoring of mine water bodies has multi-fold research challenges. In the past, such mine water bodies were monitored physically. With the new innovation in satellite imaging and remote sensing, monitoring of land classes in remote manners has become a new standard.

Satellite image-based water body detection is a long-discussed research problem. Water bodies are detected using the spectral responses of different bands which emphasize open water features, such as the Normalized Difference Water Index (NDWI) (Gao 1996; McFeeters 1996), Modified Normalized Difference Water Index (MNDWI) (Xu 2006), and Automated Water Extraction Index (AWEI) (Feyisa et al. 2014). These indexes have been used in the literature for different applications of wetland quantification (Rebelo et al. 2009), flood monitoring (Chignell et al. 2015), shallow water detection (Eugenio et al. 2015), and others. However, water bodies in mining regions cannot be separated from other water bodies using these indexes exclusively (Mukherjee et al. 2018, 2019b). The significant characteristic of a water body inside mining regions is its surrounding areas, which are mining regions. In this work, primarily surface mines are considered. Surface mining is a widely used excavation technique. In such cases, a shallow ore deposit is mined by removing the upper layer of the earth's surface. Excavated minerals and residual portions of the earth's surface are dumped nearby the mining region. Different land classes of a surface mine region have been detected in the literature by supervised and semi-supervised techniques (Karan et al. 2016; Petropoulos et al. 2013). However, the detection of mine water bodies needs to be explored further (Mukherjee et al. 2019b). A water body in such a region must have a high amount of minerals in the surrounding. This idea is explored in different works to detect water bodies inside mining regions from mid-resolution satellite images. There are different spectral indexes, which detect hydrothermally altered rock such as clay mineral ratio, iron oxide ratio, ferrous mineral ratio, and others (Drury 1993). In Mukherjee et al. (2018), water bodies are detected using NDWI and they are treated independently. Using a threshold value over the mean of clay mineral ratio values in the surrounding region of a detected water body, water bodies inside mining regions are detected. The separation of water bodies inside mining regions from other water bodies is automated in Mukherjee et al. (2019b) by using a K-Means clustering over the feature space of clay mineral ratio and iron oxide ratio of every water body and their surrounding regions. The coal mine index enhances the concept of clay mineral ratio (Mukherjee et al. 2019c). The coal mine index has been found to be effective to detect different land classes of surface coal mine regions without labeled datasets (Mukherjee et al. 2019d, 2022a, 2021). The coal mine index (CMI) has also shown substantial accuracy in other mining regions (Mukherjee et al. 2022b). In Mukherjee et al. (2019a), an automated technique has been applied to identify water bodies inside mining regions using CMI and K-Means clustering. However, if there is any region, which has a high mineral quantity next to a water body but not a mine water body, it can be falsely detected by this technique (Mukherjee et al. 2019a). A river sandbank can be an example of such misclassification. Automated monitoring of such water bodies needs automated detection of mine water bodies first. Hence, there is a significant research gap in the automated monitoring of water bodies inside mining regions.

Here, automated detection of water bodies inside mining regions is addressed first. Mostly, river sandbank regions are detected as false positives (Mukherjee et al. 2019a). This issue is addressed by separating river regions from the detected open waters. Rivers are higher in length and they may show higher sinuosity. These characteristics of rivers are used to separate them prior to identifying mine water bodies from other water bodies through mineral abundances. These water bodies are considered as shapes that change over time. There are several shape features in the literature, such as contour-based features (Adamek and Connor 2003), aspect ratio as a feature (Omachi and Omachi 2008), Euler number (Humberto et al. 2014), Hu moments (Zhang et al. 2015), Zenerik moments (Li et al. 2008), and others. Here, the shapes of water bodies do not follow a rigid body transformation and can have a high number of data points. Automated detection of change of shapes using such techniques has several associated issues. Hausdorff distance can be a feasible solution as the number of points in both the shapes need not have to be the same (Yu et al. 2009). However, it does not provide direction of change. Procrustes analysis has been widely used in the literature (Badawi-Fayad and Cabanis 2007) but mostly uses one to one correspondence. Iterative closest point (Du et al. 2010) and coherent point drift (Fan et al. 2022) algorithms work for shapes with different number of points and provide direction of changes. Hence, in this work, first, changes in water bodies are computed using Hausdorff distance. If it provides significant changes, their transformations in terms of translation, rotation, and scaling are computed using the coherent point drift algorithm.

2 Background Techniques

The primary background techniques used in this work are discussed below.

2.1 Hausdorff Distance

Hausdorff distance is a measure to quantify the dissimilarity between two sets of points. Let the boundaries of two shapes be represented as two sets Ω and ξ. From a point in set Ω, minimum distance from each point of set ξ is computed. The maximum of these minimum distances indicates the largest separation of these two sets. Let $\Omega = \omega_1, \omega_2, \omega_3, \ldots, \omega_n$ and $\xi = \varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_m$. The one-sided Hausdorff distance is computed as Eq. (1)

$$\delta_H(\Omega, \xi) = \max_{\omega\in\Omega} \min_{\varepsilon\in\xi} |\omega - \varepsilon| \tag{1}$$

The two-sided Hausdorff distance is computed as Eq. (2).

$$d_H(\Omega, \xi) = \max(\delta_H(\Omega, \xi), \delta_H(\xi, \Omega)) \tag{2}$$

In shape feature analysis of images, Hausdorff distance is extensively used (Huttenlocher et al. 1993).

2.2 Coherent Point Drift

Coherent point drift is primarily used for point cloud registration. A correspondence of two set points is achieved, and thereafter the transformation that maps one set of points to another is computed (Myronenko and Song 2010). The transformation is modeled by estimating a Gaussian mixture model representation. It offers robustness to noise and outliers, and it can handle point clouds with different number of points. Further it allows non-rigid transformations (Myronenko and Song 2010). Here, coherent point drift is used for two-dimensional data. It provides a transformation matrix, i.e. $T = \begin{bmatrix} R_{11} & R_{12} & t_1 \\ R_{21} & R_{22} & t_2 \\ 0 & 0 & 1 \end{bmatrix}$. The bottom row, i.e. $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$, is added for homogeneous coordinates. The transformation matrix can also be defined as $T = \begin{bmatrix} s * \cos(\theta) & -s * \sin(\theta) & t_1 \\ s * \sin(\theta) & s * \cos(\theta) & t_2 \\ 0 & 0 & 1 \end{bmatrix}$. Here, s and θ represent scaling and rotation angle respectively. A scaling factor <1 enlarges the shape, while a scaling factor between 0 and 1 shrinks the shape. (t_1, t_2) represent the translation in the x and y directions, respectively. The direction of change can be computed as $\theta = \arctan(\frac{R_{12}}{R_{11}})$. $\theta > 0$ represents counterclockwise rotation, and $\theta < 0$ represents clockwise rotation.

2.3 Connected Component Analysis

Connected component analysis groups similar pixels using their pixel connectivity. It scans through the image pixel-wise. The scanning starts from top to bottom and left to right. The process groups the connected regions of the adjacent pixels which share the same intensity value. As water bodies in mining regions have distinguishing features of mining regions in surroundings, they are needed to be treated individually which is studied by connected component analysis.

2.4 Spectral Indexes

In multi-spectral images, different band combinations, namely spectral indexes are used to detect different land classes. A typical spectral index is defined in Eq. (3):

$$\phi(\lambda_1, \lambda_2) = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \tag{3}$$

Here, λ represents reflectance value in a band. Normalized difference water index (NDWI) is one such spectral index used to detect water bodies. Higher values of NDWI preserve water content. There are two variants. One, i.e. $\phi(\lambda_{\text{NIR}}, \lambda_{\text{SWIR-I}})$, has been proposed to detect the water content of leaves (Gao 1996). Another variant is $\phi(\lambda_{\text{Green}}, \lambda_{\text{NIR}})$, which detects water content in water bodies (McFeeters 1996). Modified normalized difference water index (MNDWI), i.e. $\phi(\lambda_{\text{Green}}, \lambda_{\text{SWIR-I}})$, enhances the performance of NDWI by emphasizing open water features more (Singh et al. 2015). However, these indexes use empirical threshold values. An automated water extraction technique is proposed in Feyisa et al. (2014), where positive values defined water bodies. It proposes two techniques namely AWEI_{nsh} and AWEI_{sh} as defined in Eq. (4):

$$\begin{aligned} \text{AWEI}_{\text{nsh}} &= 4 \times (\lambda_{\text{Green}} - \lambda_{\text{SWIR-I}}) - (0.25 \times \lambda_{\text{NIR}} + 2.75 \times \lambda_{\text{SWIR-II}}) \\ \text{AWEI}_{\text{sh}} &= \lambda_{\text{Blue}} + 2.5 \times \lambda_{\text{Green}} - 1.5 \times (\lambda_{\text{NIR}} + \lambda_{\text{SWIR-I}}) - 0.25 \times \lambda_{\text{SWIR-II}} \end{aligned} \tag{4}$$

AWEI_{nsh} removes dark build surfaces in urban regions, and AWEI_{sh} eliminates shadowy pixels (Feyisa et al. 2014). Clay mineral ratio, i.e. $\frac{\lambda_{\text{SWIR-I}}}{\lambda_{\text{SWIR-II}}}$, a geo-physical index, detects hydrothermally altered rocks containing clay and alunite (Drury 1993). Coal mine index, i.e. $\phi(\lambda_{\text{SWIR-I}}, \lambda_{\text{SWIR-II}})$, extends the idea of clay mineral ratio. Lower values of the coal mine index have been found to be effective to identify mining regions (Mukherjee et al. 2019c, 2022b). Here reflectance values in blue, green, near infra-red, short wave infra-red one, and short wave infra-red two bands are denoted by λ_{Blue}, λ_{Green}, λ_{NIR}, $\lambda_{\text{SWIR-I}}$, and $\lambda_{\text{SWIR-II}}$, respectively.

3 Methodology

The process flow of the proposed technique to detect changes in water bodies at mining regions by an automated manner is shown in Fig. 1. The process flow is divided into two segments (Fig. 1). First, water bodies at mining regions are separated in an automated manner. Next, these water bodies at mining regions are treated in a temporal manner. Satellite images in two different times are registered through their geographic metadata using QGIS. In the final stage, the direction of changes of these water bodies are studied using Hausdorff distance and coherent point drift.

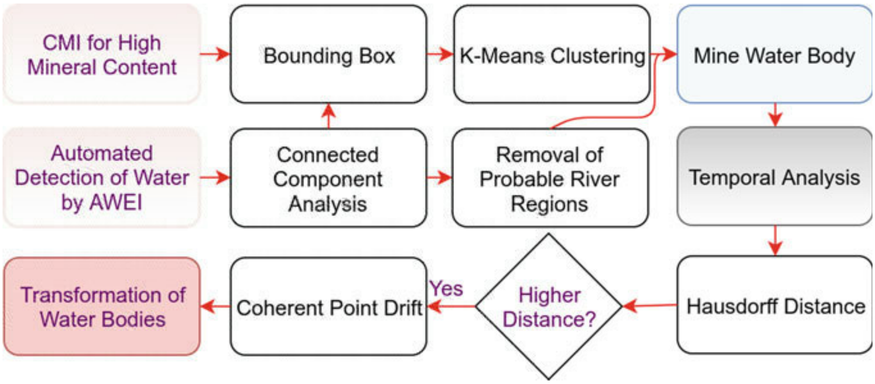


Fig. 1 Flow diagram to detect changes in water bodies at mining regions by automated manner

3.1 Separation of Mine Water Bodies

The proposed technique follows a similar workflow as discussed in (Mukherjee et al., 2019a, b). Mine water bodies are those water bodies in the vicinity of a mining region. This is treated as the primary hypothesis here. Hence, first, water bodies are detected in an automated manner using AWEI. These detected water bodies contain heterogeneous water bodies such as lakes, rivers, dams, and others along with water bodies at mining regions. Further, it may have different misclassified water body regions. Hence, each water body is needed to be treated individually to remove false positives. The connected component analysis is applied here to treat them individually. As the surrounding regions are considered, a bounding box is computed over each connected component. Water bodies inside mining regions are smaller, and a bounding box may not capture the land class properties of the surroundings properly. A padding of five pixels on each side is added to each bounding box. Further, the CMI values of these bounding boxes are studied through K-Means clustering with two classes. As lower values of CMI preserve mining regions, the cluster having lower values are detected as water bodies at mining regions.

However, this can detect any water body, which has a high mineral content region nearby but they are not mining region (Mukherjee et al. 2019a, b). A river sandbank has the same characteristics. A land class may have high mineral quantity and is closer to a river. Hence, to remove such regions, probable river regions are removed after analyzing the bounding boxes. A river has a higher stretch. Further, a river may show sinuosity. Hence, it is assumed that the area of the bounding box of a river is much higher than its connected component. Let the connected component and bounding box of a water body be W_c and W_b. A water body is considered as a river when $Wb > 4 \times W_c$. These regions are discarded. First, CMI values of all the water bodies are studied. Mean values of the bounding boxes of such water bodies are analyzed by K-Means clustering where $k = 2$. This produces two clusters. The cluster center which has lower values is preserved. The data points in this cluster are treated as water bodies at mining regions. The water bodies which satisfy the criteria of $Wb > 4 \times W_c$ are detected as rivers. However, a few river trails may appear isolated. Hence, the water bodies are treated with morphological dilation. Then, the criteria of $Wb > 4 \times W_c$ are applied. Further, the outcome is treated with morphological erosion. These are treated as the detected river regions. Detected lower CMI regions close to these are discarded.

3.2 Changes in Mine Water Bodies

Mine water bodies from two different periods are studied to detect the changes in an automated manner. The contours of these water bodies are detected, and these points are treated as the shape points of these water bodies. Here one to one correspondence is assumed for every mine water body. The cases where two water bodies have merged into one or one water body has been divided into multiple water bodies are not considered. The contour points are treated as shape points of two sets. For a contour in a past image, the closest contour in the next image is considered. Hausdorff distance provides the largest deviation of two sets of points. An empirical threshold based on this Hausdorff distance has been used here to remove such water bodies, where changes are less. The water bodies which have high changes through Hausdorff distance are further treated with coherent point drift. Coherent point drift is primarily a point cloud matching algorithm. It provides the translation, rotation, and shearing changes of two sets of points. Thus, the direction of changes in these water bodies is also perceived in an automated manner.

4 Region of Interest and Specification of Satellite Image Data

Landsat 8 images of 2017, which are accessed from USGS earth explorer, are used here for experimentation. Landsat 8 has two instruments such as operational land imager (OLI) and thermal infra-red (TIRS). Landsat 8 images consist of 9 multi-spectral bands from 0.43 to 1.38μm along with 2 thermal bands (10.60–12.51μm). The spatial and temporal resolution of these bands are 30 m (except the panchromatic image) and 16 days, respectively. Here, top of the atmosphere reflectance (TOA) values are computed from the metadata and $L1$ images. These images are orthorectified. As multi-spectral images cannot penetrate clouds, images with $<10\%$ cloud cover are considered here. The Jharia coal field (JCF) is a prominent coal mining region of India in the state of Jharkhand (latitudes 23°38′ N, and 23°50′ N and longitudes $86°07'E$, and $86°30'E$) as shown in Fig. 2. This region has heterogeneous land classes such as forest, rivers, dams, crop fields, urban lands, and mining regions. Figure 2 (right) shows a high-resolution Google Earth image of the study area. The sickle-shaped

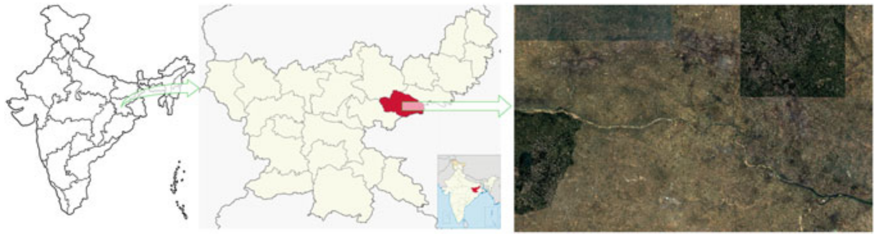


Fig. 2 The region of interest: Jharia coal field

area in Fig. 2 (right) is the JCF region. The Damodar river flows bisecting the study area (right image of Fig. 2). There are other different rivers and canals with smaller widths such as the Jamunia river that can be spotted from the north and the Ijri river at the bottom of the study area meeting the Panchet dam. High-resolution Google Earth images are considered here for validation. Ground truth regions are marked by visual observations and expert opinions from these high-resolution Google Earth images. Landsat 8 and these Google Earth images are registered using QGIS, and ground truth regions are used for validation.

5 Results and Discussions

A sample outcome of the proposed technique is shown in Fig. 3. Figure 3a shows the outcome of AWEI. AWEI extracts water from satellite images in an automated manner which does not require manual interpretation like other water indexes such as NDWI and MNDWI. There are two variations of AWEI, i.e. $\mathrm{AWEI_{sh}}$ and AWEI_{nsh}. AWEI_{nsh} removes dark build surfaces in urban regions, and $\mathrm{AWEI_{sh}}$ eliminates shadowy pixels (Feyisa et al. 2014). It has been observed that in JCF regions, $\mathrm{AWEI_{sh}}$ generates better outcomes (Mukherjee et al. 2019a). Hence, in this work, $\mathrm{AWEI_{sh}}$ has been used. Figure 3 shows the water bodies detected by AWEI. These water bodies contain different types of water body regions such as mine water bodies, lakes, swamps, dams, rivers, and narrow rivers. It is difficult to separate such water bodies using water indexes exclusively. CMI extends the concept of clay mineral ratio. CMI has been found instrumental to detect mining regions, especially coal mining regions (Mukherjee et al. 2019c, 2022b). The CMI values in the JCF region are shown in Fig. 3b. It can be observed that the sickle-shaped area in the JCF has been enhanced. The mineral content in the surrounding region is considered here as the primary distinguishing factor to separate water bodies in mining regions from other water bodies. It is treated as the primary hypothesis for detecting water bodies at mining regions (Mukherjee et al. 2018, 2019b). In the past, water bodes inside mining regions have been separated from other water bodies using CMI values (Mukherjee et al. 2019a). The detected water bodies are treated individually. To perceive the abundance of minerals, the surrounding of each connected component needs to be analyzed. Hence, bounding boxes are computed. Due to the structure of a connected component, in a few cases, a simple bounding box may not capture the mineral abundance in the surroundings. Hence, each of these bounding boxes has a padding of five pixels on each side. CMI values of these bounding boxes are studied. The mean of these values is computed. These mean values of every bounding box are treated with a K-Means clustering algorithm for two cluster points. Lower values of CMI preserve regions with high mineral abundance. Hence, connected components which are associated with the cluster center with lower values are preserved. These regions are considered as the detected mine water body regions. However, these techniques may become error-prone if there is a high mineral abundance region, which is not a mining region close to a water body. River sandbank regions have been mostly found to be detected as false positives for this limitation. The water body close to a river sandbank is a river. This limitation is addressed in this work by detecting the probable river regions. Rivers have higher stretches. The connected component, which has a very high value, is likely to be a river. Further, a river does not always follow a straight path. Rivers have bends. Thus, the bounding box of a river is highly likely to have a bigger area than the connected component, i.e. the area of the river path. This is unlikely for smaller lakes like mine water bodies. Further, due to the occlusion of other land classes, narrow width, presence of prominent river sandbank, and other reasons, a river may appear as isolated trails than a continuous one. In such cases, the area covered by the bounding box may not be adequately higher than the connected component. Thus, the water bodies are treated with morphological dilation such that those isolated trails may get connected. Thereafter, the criteria of $Wb > 4 \times W_c$ is considered and probable river regions are detected. The probable river regions are shown in Fig. 3d. Next, regions close to the probable river regions which have lower CMI values, i.e. high mineral abundance, are discarded. As the bounding box of a probable river covers a large area, here, high mineral abundance regions are treated individually rather than analyzing the probable river regions by connected components. Figure 3e and f shows the outcome of the proposed technique in December and January of 2017, respectively. Two samples are chosen from Fig. 3e and f for the change detection analysis. Figure 3g and h shows one such sample of mine water bodies in January and December. Figure 3i and j shows the second sample of mine water bodies in January and December. It can be observed that the mine water bodies shown in Fig. 3i and j have significant changes than the mine water bodies in Fig. 3g and h. Different spectral indexes such as clay mineral ratio, a combination of clay mineral and iron oxide ratio, and coal mine index have been studied in the literature to identify mine water bodies. The clay mineral ratio provides an accuracy of 80% in detecting mine water bodies (Mukherjee et al. 2018). The combination of clay mineral and iron oxide ratio separates mine and non-mine water bodies with the precision, recall, and F_1 score of 71.36, 69.39, and 70.36%, respectively (Mukherjee et al. 2019b). CMI detects water bodies in mining regions by an automated technique with precision, recall, and $F1$ score of 87.46, 65.89, and 75.16%, respectively (Mukherjee et al. 2019a). The precision and recall of the proposed technique to detect mine water bodies have been found to be 89.21, and 64.48%, respectively. The precision has increased as a few river sandbank regions

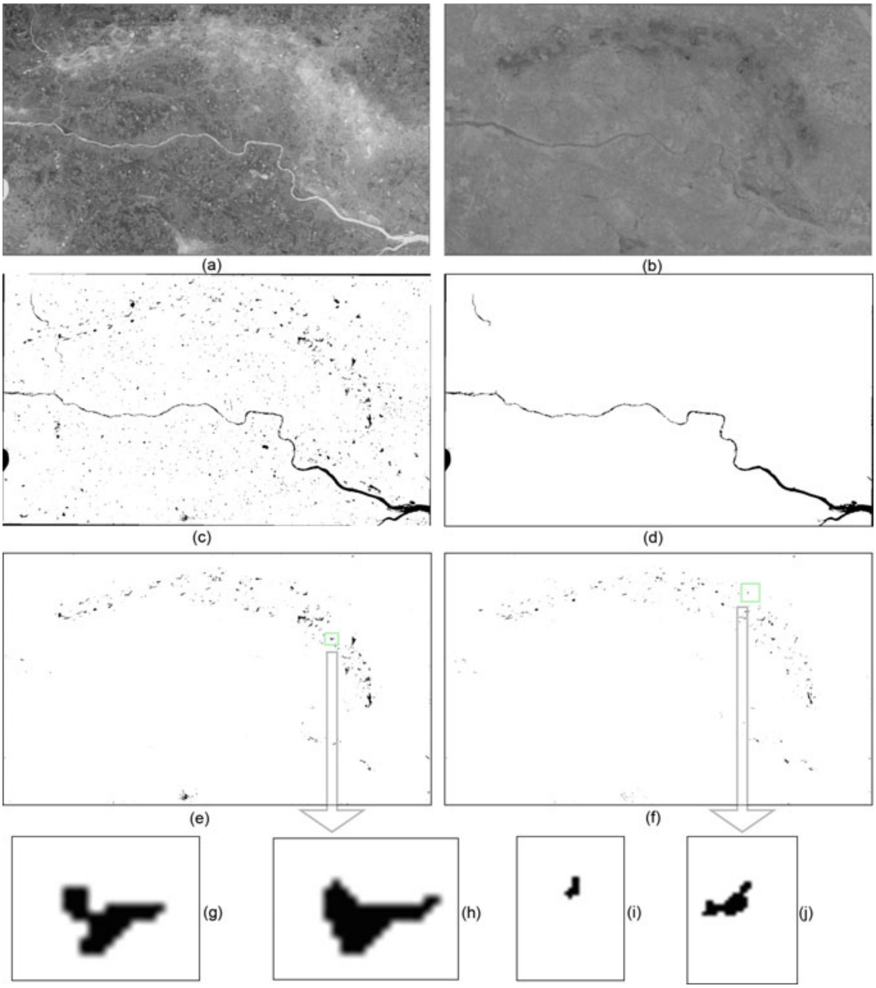


Fig. 3 Results: **a** AWEI Image; **b** CMI Image; **c** Detected water bodies from AWEI; **d** Probable rivers; **e** Mine water bodies in December 2017; **f** Mine water bodies in January 2017; **g** A sample mine water body in January; **h** The same water body of (**g**) in December; **i** A sample mine water body in January; **j** The same water body of **i** in December

have been discarded whereas, a few true positive mining regions in the vicinity of a river are also discarded. Hence, the recall has decreased. This is treated as one of the future directions of this work.

Water bodies in January and December of 2017 are considered here for change detection. The change of these water bodies is computed using two automated techniques, such as Hausdorff distance and coherent point drift. Higher Hausdorff distance indicates higher changes in shapes as shown in Table 1. Table 1 shows the image of the mine water body in January and December, the Hausdorff distance,

Table 1 A few sample mine water bodies in January and December along with Hausdorff distance and Transformation matrix from coherent point drift

January	December	Hausdorff distance	Transformation	θ	Scale
		10.2956	$\begin{matrix} 0.6174 & -0.7866 & 2193.89 \\ -0.7866 & -0.6174 & 6166.13 \\ 0 & 0 & 1 \end{matrix}$	−51 967°	1.001
		2.236	$\begin{matrix} -0.991 & -0.133 & 23496.24 \\ 0.1332 & -0.991 & -5629.27 \\ 0 & 0 & 1 \end{matrix}$	7.62°	1.0006

the transformation matrix, the direction of rotational change, and the scaling factor. The mine water bodies in the first row have gone through a wide variety of changes. Thus the Hausdorff distance of these images is large. Further, the changes in the second row of Table 1 are less than in the first row. Hence, the Hausdorff distance of these images is less. The transformation matrix is shown in the fourth column of Table 1. The last column of the transformation matrix shows the translation in X and Y directions. The degree of rotational change has been computed and it is shown in Table 1. The images in the first row show clockwise and the images in the second row show counterclockwise rotation. In both images, the mine water body has expanded. Thus the scaling factor has been found to be >1. Hence, the proposed technique can compute the direction of change and quantify the change in translation, rotation, and scaling. In Sarp and Ozcelik (2017), mine water body changes are detected using support vector machines through spectra water indexes and structural similarity index. This study used satellite image interpretation and geographic information systems. Different techniques of glacial lake expansion are discussed in Ahmed et al. (2021). Water spread mapping of multiple lakes is studied using different water indexes in Deoli et al. (2021). In Ali et al. (2019), changes of urban water bodies are quantified through water indexes. However, works regarding the change detection of water bodies in mining regions have been limited. The proposed technique detects the changes in mine water bodies without a labeled dataset in an automated manner using Landsat images explicitly. However, the technique assumes a one-to-one mapping of mine water bodies. Whereas the mine water bodies are very dynamic in nature, they can go through frequent changes. One mine water body can split into several other water bodies and different water bodies can merge into one. This is considered as a special case and need further attentions. It is considered as one of the future directions of this work. The proposed technique has been applied to coal mine regions with TOA reflectance values. Further experimentation regarding