

Advances in Intelligent Systems and Computing 769

Herwig Unger · Sunantha Sodsee
Phayung Meesad *Editors*

Recent Advances in Information and Communication Technology 2018

Proceedings of the 14th International
Conference on Computing and
Information Technology (IC2IT 2018)

 Springer

Advances in Intelligent Systems and Computing

Volume 769

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within “Advances in Intelligent Systems and Computing” are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Chairman

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in

Members

Rafael Bello Perez, Universidad Central “Marta Abreu” de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu

Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchado@usal.es

Hani Hagrass, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk

László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu

Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu

Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlm@mail.nctu.edu.tw

Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au

Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org

Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br

Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl

Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

More information about this series at <http://www.springer.com/series/11156>

Herwig Unger · Sunantha Sodsee
Phayung Meesad
Editors

Recent Advances in Information and Communication Technology 2018

Proceedings of the 14th International
Conference on Computing and Information
Technology (IC2IT 2018)

 Springer

Editors

Herwig Unger
LG Kommunikationsnetze
Fern Universität in Hagen
Hagen, Germany

Phayung Meesad
Faculty of Information Technology
King Mongkut's University of Technology
North Bangkok
Bangkok, Thailand

Sunantha Sodsee
Faculty of Information Technology
King Mongkut's University of Technology
North Bangkok
Bangkok, Thailand

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-93691-8 ISBN 978-3-319-93692-5 (eBook)
<https://doi.org/10.1007/978-3-319-93692-5>

Library of Congress Control Number: 2018947344

© Springer International Publishing AG, part of Springer Nature 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG
part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This proceedings volume contains the papers of the 14th International Conference on Computing and Information Technology (IC2IT 2018) held on July 5–6, 2018, in Chiang Mai, Thailand. IC2IT is a platform for researchers to meet and exchange knowledge in the field of computer and information technology. The participants in IC2IT present their current research new findings and discuss with partners to seek new research directions and solutions as well as cooperation.

Springer has published the proceedings of IC2IT in its well-established and worldwide-distributed series on Advances in Intelligent and Soft Computing, Janusz Kacprzyk (Series Editor). This year there were total 88 submissions from 16 countries. Each submission was assigned to at least three program committee members, and at least two members must accept a paper in order to include in the proceedings. The committee accepted 33 papers for presenting at the conference and publishing the papers in the proceedings. The contents of the proceedings are divided into subfields: Data Mining, Machine Learning, Natural Language Processing, Image Processing, Network and Security, Software Engineering, and Information Technology.

We would like to thank all authors for their submissions. We would also like to thank all the program committee members for their support in reviewing assigned papers and in giving good comments back to the authors to revise their papers. In addition, we would like to thank all university partners in both Thailand and overseas for academic cooperation. Special thanks also go to the staff members of the Faculty of Information Technology at King Mongkut's University of Technology North Bangkok who have done many technical and organizational works. Without the painstaking work of Dr. Watchareewan Jitsakul, the proceedings could not have been completed in the needed form at the right time.

Finally yet importantly, we would like to thank all the speakers and audiences for their contributions and discussions at the conference that made the conference a success. We hope that the proceedings IC2IT will be a good source of research papers for future references with the state of the art.

April 2018

Herwig Unger
Sunantha Sodsee
Phayung Meesad

Organization

Program Committee

| | |
|---------------------|---------------------------------|
| M. Aiello | Uni. Stuttgart, Germany |
| S. Auwatanamongkol | NIDA, Thailand |
| G. Azzopardi | Uni. Groningen, The Netherlands |
| T. Bernard | Syscom CReSTIC, France |
| C. Biemann | TU, Darmstadt |
| S. Boonkrong | SUT, Thailand |
| S. Butcharoen | TOT, Thailand |
| M. Caspar | Braunschweig, Germany |
| T. Chintakovid | KMUTNB, Thailand |
| K. Chochiang | PSU, Thailand |
| K. Dittakan | PSU, Thailand |
| T. Eggendorfer | HS Weingarten, Germany |
| M. Ghazali | UTM, Malaysia |
| M. Hagan | OSU, USA |
| W. Halang | FernUni, Germany |
| C. Haruechaiyasak | NECTEC, Thailand |
| U. Inyaem | RMUTT, Thailand |
| W. Jirachiefpattana | NIDA, Thailand |
| W. Jitsakul | KMUTNB, Thailand |
| S. Jungjit | TSU, Thailand |
| M. Kaenampornpan | MSU, Thailand |
| M. Komkhao | RMUTP, Thailand |
| A. Kongthon | NECTEC, Thailand |
| P. Kropf | UniNe, Switzerland |
| M. Kubek | FernUni, Germany |
| K. Kyamakya | AAU, Austria |
| U. Lechner | Uni. München, Germany |
| H. Lefmann | TU Chemnitz, Germany |

| | |
|-------------------------|--------------------|
| P. Meesad | KMUTNB, Thailand |
| W. Nadee | PSU, Thailand |
| A. Nanthaamornphong | PSU, Thailand |
| K. Nimkerdphol | RMUTT, Thailand |
| M. Phongpaibul | TU, Thailand |
| J. Polpinij | MSU, Thailand |
| W. Pongphet | NIDA, Thailand |
| N. Porrawatpreyakorn | KMUTNB, Thailand |
| P. Prathombutr | DEPA, Thailand |
| P. Saengsiri | TIRTR, Thailand |
| P. Sanguansat | PIM, Thailand |
| T. Siriborvornratanakul | NIDA, Thailand |
| M. Sodanil | KMUTNB, Thailand |
| S. Sodsee | KMUTNB, Thailand |
| V. K. Solanki | CMRIT, India |
| W. Sriurai | UBU, Thailand |
| T. Sucontphunt | NIDA, Thailand |
| W. Tang | CityUni, Hong Kong |
| C. Thaenchaikun | PSU, Thailand |
| J. Thaenthong | PSU, Thailand |
| D. Thammasiri | NPRU, Thailand |
| N. Theeraampornpunt | PSU, Thailand |
| J. Thongkam | MSU, Thailand |
| N. Tongtep | PSU, Thailand |
| D. Tutsch | UNIWUP, Germany |
| H. Unger | FernUni, Germany |
| V. Wattanasoontorn | PSU, Thailand |
| M. Weiser | OSU, USA |
| W. Werapun | PSU, Thailand |
| N. Wisitpongphan | KMUTNB, Thailand |
| K. Woraratpanya | KMITL, Thailand |
| P. Wuttidittachotti | KMUTNB, Thailand |

Organizing Partners

In Cooperation with

King Mongkut's University of Technology North Bangkok (KMUTNB)

FernUniversitaet in Hagen, Germany (FernUni)

Chemnitz University, Germany (CUT)

Oklahoma State University, USA (OSU)

Edith Cowan University, Western Australia (ECU)

Hanoi National University of Education, Vietnam (HNUE)

Gesellschaft für Informatik (GI)
Maharakham University (MSU)
Ubon Ratchathani University (UBU)
Kanchanaburi Rajabhat University (KRU)
Nakhon Pathom Rajabhat University (NPRU)
Phetchaburi Rajabhat University (PBRU)
Rajamangala University of Technology Krungthep (RMUTK)
Rajamangala University of Technology Thanyaburi (RMUTT)
Prince of Songkla University, Phuket Campus (PSU)
National Institute of Development Administration (NIDA)
Sisaket Rajabhat University (SSKRU)
Council of IT Deans of Thailand (CITT)
IEEE CIS Thailand

Contents

Data Mining

| | |
|---|----|
| Combining Multiple Features for Product Categorisation by Multiple Kernel Learning | 3 |
| Chanawee Chavaltada, Kitsuchart Pasupa, and David R. Hardoon | |
| ARMFEG: Association Rule Mining by Frequency-Edge-Graph for Rare Items | 13 |
| Pramool Suksakaophonng, Phayung Meesad, and Herwig Unger | |
| Hybrid Filtering Approach for Linked Open Data Exploration Applied to a Scholar Object Linking System | 23 |
| Siraya Sitthisarn and Nariman Virasith | |
| Sugarcane Yield Grade Prediction Using Random Forest with Forward Feature Selection and Hyper-parameter Tuning | 33 |
| Phusanisa Charoen-Ung and Pradit Mittrapiyanuruk | |
| The Efficacy of Using Virtual Reality for Job Interviews and Its Effects on Mitigating Discrimination | 43 |
| Rico A. Beti, Faris Al-Khatib, and David M. Cook | |

Natural Language Processing

| | |
|--|----|
| Short Text Topic Model with Word Embeddings and Context Information | 55 |
| Xianchao Zhang, Ran Feng, and Wenxin Liang | |
| Temporal Analysis of Twitter Response and Performance Evaluation of Twitter Channels Using Capacitor Charging Model | 65 |
| Sirisup Laohakiat, Photchanan Ratanajaipan, Krissada Chalermsook, Leenhapat Navaravong, Rachanee Ungrangsi, Aekavute Sujarae, and Krissada Maleewong | |

| | |
|--|-----|
| On Evolving Text Centroids | 75 |
| Herwig Unger and Mario Kubek | |
| Thai Word Safe Segmentation with Bounding Extension for Data Indexing in Search Engine | 83 |
| Achariya Klahan, Sukunya Pannoi, Prapeepat Uewichitrapochana, and Rungrat Wiangsripanawan | |
| Robust Kurtosis Projection Approach for Mangrove Classification | 93 |
| Dyah E. Herwindiati, Janson Hendryli, and Sidik Mulyono | |
| Machine Learning | |
| Diabetes Classification with Fuzzy Genetic Algorithm | 107 |
| Wissanu Thungrut and Naruemon Wattanapongsakorn | |
| Decision Fusion Using Fuzzy Dempster-Shafer Theory | 115 |
| Somnuek Surathong, Sansanee Auephanwiriyaikul, and Nipon Theera-Umpon | |
| Enhanced Manhattan-Based Clustering Using Fuzzy C-Means Algorithm | 126 |
| Joven A. Tolentino, Bobby D. Gerardo, and Ruji P. Medina | |
| The Cooperation of Candidate Solutions Vortex Search for Numerical Function Optimization | 135 |
| Wirote Apinantanakon, Siriporn Pattanakitsiri, and Pochra Uttamaphant | |
| Integrating Multi-agent System, Geographic Information System, and Reinforcement Learning to Simulate and Optimize Traffic Signal Control | 145 |
| Hoang Van Dong, Bach Quoc Khanh, Nguyen Tran Lich, and Nguyen Thi Ngoc Anh | |
| Image Processing | |
| Digital Image Watermarking Based on Fuzzy Image Filter | 157 |
| K. M. Sadrul Ula, Thitiporn Pramoun, Surapont Toomnark, and Thumrongrat Amornraksa | |
| Image Watermarking Against Vintage and Retro Photo Effects | 167 |
| Piyantart Chotikawanid, Thitiporn Pramoun, Pipat Supasirisun, and Thumrongrat Amornraksa | |
| Performance Analysis of Lossless Compression Algorithms on Medical Images | 177 |
| I Made Alan Priyatna, Media A. Ayu, and Teddy Mantoro | |

| | |
|--|-----|
| Brain Wave Pattern Recognition of Two-Task Imagination by Using Single-Electrode EEG | 187 |
| Sararat Wannajam and Wachirawut Thamviset | |
| Isarn Dharma Handwritten Character Recognition Using Neural Network and Support Vector Machine | 197 |
| Saowaluk Thaiklang and Pusadee Seresangtakul | |
| Network and Security | |
| Subset Sum-Based Verifiable Secret Sharing Scheme for Secure Multiparty Computation | 209 |
| Romulo L. Olalia, Jr., Ariel M. Sison, and Ruji P. Medina | |
| The Semantics Loss Tracker of Firewall Rules | 220 |
| Suchart Khummanee | |
| Cryptanalysis of a Multi-factor Biometric-Based Remote Authentication Scheme | 232 |
| Sirapat Boonkrong | |
| Haptic Alternatives for Mobile Device Authentication by Older Technology Users | 243 |
| Kulwinder Kaur and David M. Cook | |
| The New Modified Methodology to Solve ECDLP Based on Brute Force Attack | 255 |
| Kritsanapong Somsuk and Chalida Sanemueang | |
| Information Technology | |
| Generations and Level of Information on Mobile Devices Usage: An Entropy-Based Study | 267 |
| Charnsak Srisawatsakul and Waransanang Boontarig | |
| Requirement Patterns Analysis and Design of Online Social Media Marketing for Promoting Eco-Tourism in Thailand | 277 |
| Orasa Tetiwat, Vatcharaporn Esichaikul, and Ranee Esichaikul | |
| The Empirical Discovery of Near-Optimal Offline Cache Replacement Policies for Nonuniform Objects | 286 |
| Thepparit Banditwattanawong | |
| Organic Farming Policy Effects in Northern of Thailand: Spatial Lag Analysis | 297 |
| Ke Nunthasen and Waraporn Nunthasen | |
| Approaching Mobile Constraint-Based Recommendation to Car Parking System | 306 |
| Khin Mar Win and Benjawan Srisura | |

Software Engineering

| | |
|--|------------|
| A Comparative Study on Scrum Simulations | 317 |
| Sakgasit Ramingwong, Lachana Ramingwong, Kenneth Cosh, and Narissara Eiamkanitchat | |
| Transformation of State Machines for a Microservice-Based Event-Driven Architecture: A Proof-of-Concept | 327 |
| Roland Petrasch | |
| Knowledge Discovery from Thai Research Articles by Solr-Based Faceted Search | 337 |
| Siripinyo Chantamunee, Chun Che Fung, Kok Wai Wong, and Chaiyun Dumkeaw | |
| Author Index | 347 |



Robust Kurtosis Projection Approach for Mangrove Classification

Dyah E. Herwindiati¹(✉), Janson Hendryli¹, and Sidik Mulyono²

¹ Faculty of Information Technology, Tarumanagara University, Jakarta, Indonesia

{dyahh, jansonh}@fti.untar.ac.id

² The Indonesian Agency for the Assessment and Application of Technology (BPPT), Central Jakarta, Indonesia

sidik.mulyono@bppt.go.id

Abstract. Mangroves are coastal vegetations that grow at the interface between land and sea. It can be found in tropical and subtropical tidal areas. Mangrove ecosystems have many ecological roles spans from forestry, fisheries, environmental conservation. The Indonesian archipelago is home to a large mangrove population which has enormous ecological value. This paper discusses mangrove land detection in the North Jakarta from Landsat 8 satellite imagery. One of the special characteristics of mangroves that are distinguishing them from another vegetation is their growing location. This characteristic makes mangrove classification using satellite imagery non trivial task. We need an advanced method that can confidently detect the mangrove ecosystem from the satellite images. The objective of this paper is to propose the robust algorithm using projection kurtosis and minimizing vector variance for mangrove land classification. The evaluation classification provides that the proposed algorithm has a good performance.

Keywords: Kurtosis · Mangrove · Multivariate · Projection pursuit Robust · Satellite imagery · Vector variance

1 Introduction

In this paper, we discuss a robust kurtosis projection minimizing vector variance for mangrove detection. Robust kurtosis projection minimizing vector variance is a procedure which combines two advantages of both kurtosis projection pursuit and robust estimation minimizing vector variance. Projection pursuit is a technique that explores a higher to lower dimensional space by examining the marginal distributions of low dimensional linear projections.

Robust statistics has been developed since almost sixty years ago. Huber [16] introduced the robust estimator because the assumptions of normality, linearity, and independence stucked on the classic estimation are frequently not satisfied. The major goal of robust statistics is to develop the statistics measures that are robust against one or more outliers hidden in the dataset [9]. There are several outlier definitions; the word “outlier” is closed to the word “inconsistent”. Hawkins [6] defined an outlier as an

observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism. Since the 20th century, the robust method has been used by scientists to eliminate the influence of anomalous observations in numerous applications.

The robust kurtosis projection minimizing vector variance method was proposed by Herwindiati et al. [4]. The algorithm works in two stages. In the first stage, it has been proposed the projection approach finding the orthonormal set of all vectors that maximize the kurtosis of the projected standardized data. This approach improves on the slow convergence rate proposed by Friedman [11]. In the second stage, we estimate robust covariance matrix minimizing vector variance to identify the data contamination which is the labeled outlier.

Robust minimum vector variance (MVV) is a measure minimizing vector variance to obtain the robust estimator. Robust MVV was proposed by [3]. The multivariate dispersion vector variance (VV) is a measure of dispersion. Geometrically, VV is a square of the length of the diagonal of a parallelotope generated by all principal components [13]. Robust kurtosis projection has a good performance and robust to detect data contamination in a low, moderate, high, even very high percentage.

Mangroves are coastal vegetations that grow at the interface between land and sea [12] and can be found in tropical and subtropical tidal areas. Mangrove ecosystems have many ecological roles; spans from forestry, fisheries, and environmental conservation [1]. The Indonesian archipelago is home to a large mangrove population which has enormous ecological values.

The input data of this research is derived from Landsat 8 satellite imagery. One of the special characteristics of mangroves that are distinguishing them from another vegetation is their growing location. Detecting mangrove might not be enough by focusing only on NIR, red, and green spectral band from satellite imagery, but also the existence of the water, particularly the sea, around the ecosystem. Because of that fact, existing approaches for detection using vegetation indices cannot effectively detect the mangrove ecosystem. We need an advanced method that is able to confidently detect the mangrove ecosystem from satellite images.

The objective of this paper is to propose a robust algorithm using kurtosis projection and minimizing vector variance for mangrove detection. The case study is the area of coastal and waters of North Jakarta, Indonesia.

2 Case Study - Mangrove

Mangroves are coastal vegetations that grow at the interface between land and sea [12] and can be found in tropical and subtropical tidal areas. Mangrove ecosystems have many ecological roles, spans from forestry, fisheries, and environmental conservation [1]. Nutrient cycling and fish spawning grounds are other services provided by the mangrove ecosystem [5]. The erosion of coastal areas can also be prevented by mangrove forests. Lately, mangrove ecosystems have also been studied as the object of conservation for reducing the effects of the tsunami in Asia [8].

The Indonesian archipelago is home to a large mangrove population which has enormous ecological value. The total area of mangrove forests in Indonesia accounts for 49% of mangroves in Asia, followed by Malaysia (10%) and Myanmar (9%) [2]. In spite of that, it is estimated that the area of mangrove forests in Indonesia has been reduced by about 120,000 hectares from 1980 to 2005, mainly due to the changes in land use for agriculture. In 2007, the ministry of forestry released a report stated that 70% of the total area of the mangrove ecosystem in Indonesia, which is 7,758,410.595 hectares, are damaged. Conservation of mangroves in Indonesia should be prioritized for the benefits of the environment and ecosystems [5].

Some of the mangrove conservation areas in Indonesia are in the northwest of Java, particularly in the North Jakarta, Tangerang, and Bekasi region. Figure 1 shows Google Earth image at North Jakarta, Indonesia. According to the Department of Forestry West Java province, 1,000 hectares of the mangrove forest at Jakarta bay in 1977 have been reduced to only 200 hectares. The mangrove forest at the area faces serious threats from urban development and waste, in addition to the lack of conservation plans from the government.

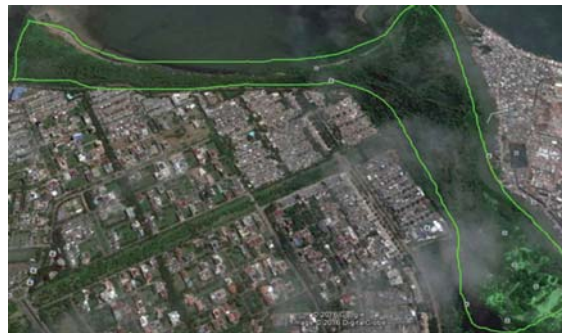


Fig. 1. The area of mangrove forest at North Jakarta, Indonesia

The near-infrared (NIR), red, and green spectral bands are often used for vegetation detection from satellite images. The normalized difference vegetation index (NDVI), which is the ratio of the difference of the red and NIR spectral band divided by their sum, is widely used in remote sensing studies of vegetation [14]. However, the characteristics of mangrove vegetation might be different, so that some vegetation indices may not be able to detect it well. One of the special characteristics of mangroves is their growing location, that is the coastal areas. Detecting mangrove might not be enough by focusing on NIR, red, and green spectral band, but also the existence of the sea around the ecosystem. Furthermore, the location of mangrove can coincide with another green vegetation, as in Fig. 2, and make it harder for remote sensing methods to correctly detect it.

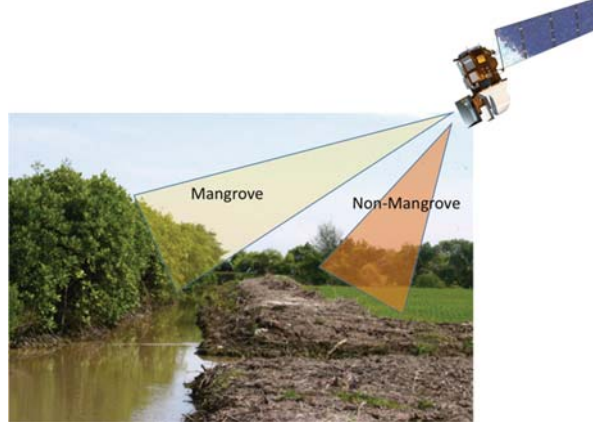


Fig. 2. The challenge of mangrove detection from satellite imagery

3 The Kurtosis Projection Approach Using Minimum Vector Variance

The kurtosis coefficient is a measure of how peaked or flat the distribution is. The dataset with high kurtosis tends to have heavy tails and sharp density peak near the center point. Kurtosis can be formally defined as the standardized fourth population moment about the mean.

$$K = \frac{E(X - \mu)^4}{\left(E(X - \mu)^2\right)^2} = \frac{\mu_4}{\sigma^4} \quad (1)$$

Projection pursuit [11] is a technique aiming at identifying low dimensional projections of data that reveal interesting structures. The framework of projection pursuit is formulated as an optimization problem with the goal of finding projection axes that minimize or maximize a measure of interest called projection index. Projection pursuit is a technique that explores a high dimensional data by examining the marginal distributions of low dimensional linear projections. The two basic components of projection pursuit are its index and its algorithm [15].

The algorithm of robust kurtosis projection minimizing vector variance (MVV) works in two stages. In the first stage, the projection approach finds the orthonormal set of all vectors that maximize the kurtosis of the projected standardized data. The algorithm of projection is inspired by Pena and Prieto [7]. In the second stage, we estimate robust covariance matrix MVV to estimate spectral characteristic of mangrove

Let $X = (x_1, x_2, \dots, x_n)'$ be a data matrix of size $n \times p$ as the observation result of p variables to n individual objects. The orthogonal projection of each observation results on to one dimensional space spanned by d is $y_i = d'x_i$ where d is a unit vector in \mathbb{R}^p .

As the projected data is written as $Y = (y_1, y_2, \dots, y_n)$, the kurtosis of the projected data is formulated as in Eq. 2.

$$K = \frac{\frac{1}{n} \sum_{i \in \mathbb{N}_n} (y_i - t)^4}{s^4} \quad (2)$$

where $t = \frac{1}{n} \sum_{i \in \mathbb{N}_n} y_i$ and $s^2 = \frac{1}{n} \sum_{i \in \mathbb{N}_n} (y_i - t)^2$ are the sample mean and the sample of the data Y respectively. It should be noted that s^4 is the square of s^2 and \mathbb{N}_n is the set of all natural numbers less than or equal to n .

Centering and scaling transformation of Y gives a new data $Z = (z_1, z_2, \dots, z_n)$ where for all $i \in \mathbb{N}_n$, $z_i = \frac{y_i - t}{s}$. Since the sample mean of Z is 0 and the sample variance of Z is 1, the kurtosis of Z can be formulated as in Eq. 3.

$$K = \frac{1}{n} \sum_{i \in \mathbb{N}_n} z_i^4 \quad (3)$$

The sample mean t and the covariance matrix S of the data matrix X can be defined as in Eqs. 4 and 5. Then, the kurtosis K can also be written as in Eq. 6.

$$t = \frac{1}{n} \sum_{i \in \mathbb{N}_n} x_i \quad (4)$$

$$S = \frac{1}{n} \sum_{i \in \mathbb{N}_n} (x_i - t)(x_i - t)' \quad (5)$$

$$K = \frac{1}{n} \sum_{i \in \mathbb{N}_n} \left(\frac{d'(x_i - t)}{\sqrt{d'Sd}} \right)^4 \quad (6)$$

Consider the objection function f defined as in Eq. 7. We assume that κ_1 is the eigen value of $\frac{1}{n} \sum_{i \in \mathbb{N}_n} (d'y_i)^2 y_i y_i'$ and d is the corresponding eigenvector. Multiplying Eq. 7 by d' from the left gives us the following result in Eq. 8.

$$f(d) = \frac{1}{n} \sum_{i \in \mathbb{N}_n} (d'y_i)^4 - \lambda(d'd - 1) \quad (7)$$

$$\kappa_1 = d' \left(\frac{1}{n} \sum_{i \in \mathbb{N}_n} (d'y_i)^2 y_i y_i' \right) d = \frac{1}{n} \sum_{i \in \mathbb{N}_n} (d'y_i)^4 = K \quad (8)$$

So the unit vector d that maximizes K is the eigenvector of $\frac{1}{n} \sum_{i \in \mathbb{N}_n} (d'y_i)^2 y_i y_i'$ corresponding to the maximum eigenvalue of that matrix. We will call such eigen vector by d_1 . The Fig. 3 illustrates the good performance of kurtosis projection from $X = (x_1, x_2, \dots, x_n) \in \mathbb{R}^3$ to d_1 .

The measures of kurtosis relate to the fourth moment of the data and emphasize the tails of the distribution. This paper use the algorithm of robust projection of maximizing kurtosis which was discussed by [4] for classification of mangrove. The good

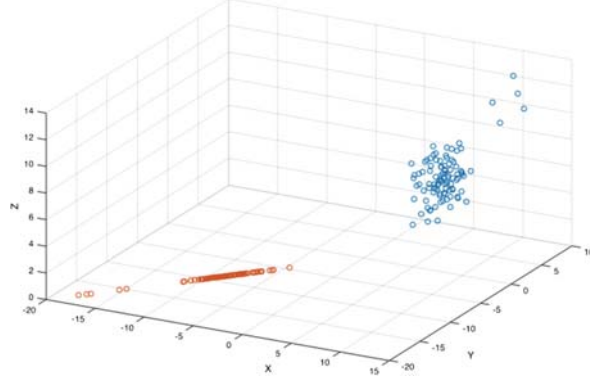


Fig. 3. Illustration of kurtosis projection based on simulation

performance of robust approach will be applied for detection and mapping of mangrove land in Jakarta.

The kurtosis projection approach is used in the initial step of robust kurtosis projection minimizing vector variance. Suppose a matrix data $X = (x_1, x_2, \dots, x_n)'$ of size $n \times p$, t is the sample mean, and S is the sample covariance matrix of X , the kurtosis projection algorithm can be derived as follows:

1. Standardize the matrix X such that the projected data has mean 0 and variance 1.

$$y_i = S^{-\frac{1}{2}}(x_i - t) \quad (9)$$

2. Find the first unit vector d_1 as the eigenvector of $\frac{1}{n} \sum_{i \in \mathbb{N}_n} (d' y_i)^2 y_i y_i'$ corresponding to the maximum eigenvalue of that matrix.
3. For $k = 2, 3, \dots, p$, find the k -th unit vector d_k as the eigenvector of $\left(I - \sum_{j=1}^{k-1} d_j d_j' \right) \frac{1}{n} \sum_{i \in \mathbb{N}_n} (d' y_i)^2 y_i y_i'$ corresponding to the maximum eigen value of that matrix. The output is an orthonormal set of all vectors that maximize the kurtosis d_1, d_2, \dots, d_p .
4. Find the projection $z_i = d' y$.

Kurtosis projection provides excellent results for detection of multivariate labeling outlier with very small and small contamination in the dataset, but it fails on a moderate percentage of data contamination. To improve the performance, the algorithm of robust kurtosis projection minimizing vector variance was proposed by [4].

Robust minimum vector variance was proposed by Herwindiati [3] for identifying multivariate outlier labeling. MVV estimator is the robust high breakdown point, i.e. $\frac{n-2(p-1)}{2n}$. Geometrically, vector variance (VV) is a square of length of paralleloptope diagonal generated by all principal components of X [14].

Assume a random sample X_1, X_2, \dots, X_n of a p -variate, the estimator MVV is the pair (T_{MVV}, C_{MVV}) having minimum square of diagonal length of the parallelogram. The robust MVV algorithm was described by [3].

$$T_{MVV} = \frac{1}{n} \sum_{i \in H} X_i \quad (10)$$

$$C_{MVV} = \frac{1}{n} \sum_{i \in H} (X_i - T_{MVV})(X_i - T_{MVV})^T \quad (11)$$

Robust kurtosis projection minimizing vector variance is an effective method to identify the multivariate outliers in the high dimension. In the simulation experience, robust kurtosis projection using MVV has a breakdown point nearing 0.5. The measure indicates that the robust kurtosis projection method provides results that can be trusted until the half of the data is contaminated [5].

The algorithm of robust kurtosis method can be briefly explained as follows:

1. Find the orthonormal set of all vectors that maximizing kurtosis $\{\vec{d}_1, \vec{d}_2, \dots, \vec{d}_p\}$.
2. Compute MVV robust estimators: the location and scale estimator.

The detailed computation for MVV robust estimator can be found in [3–5].

4 The Mangrove Classification

The supervised mangrove classification will be done in two processes: training and testing process. To conduct the training process, the algorithm of robust kurtosis projection is used for the estimation of mangrove and non-mangrove's spectral.

The mangrove and non-mangrove area for the training and testing purposes can be located visually from Google Earth; and by using the Global Mapper and ENVI software, the Landsat images can be cropped according to the coordinates. The data retrieved from the satellite images are the spectral values of each pixel.

The mangrove vegetation is detected from the Landsat 8 satellite imagery which can be freely downloaded from <http://glovis.usgs.gov>. For the model training process, we use North Jakarta area which is cropped using the Global Mapper software. Subsequently, 300 points are selected by visual inspection of the area from Google Earth which proportionally represents mangrove, water area, and soil. Meanwhile, for testing and evaluating the trained model, we use the Greater Jakarta area. Therefore, the advantage of our model is that it can learn from a small training dataset size of 300 locations to detect much larger area (more than a million pixels in the testing data). The illustration of the training data can be seen in Fig. 4.

The classification step is conducted by using the spectral reference in the training step. For each satellite image pixels from the training data, our model learns the characteristics of mangrove vegetation, water, and soil.

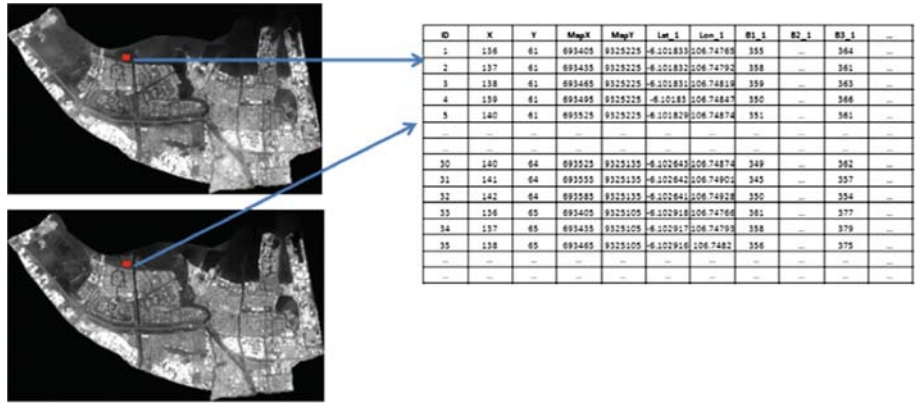


Fig. 4. Illustration of cropping data for training process from the Landsat 8 satellite imagery (band 1 and band 3)

The spectral references are very useful for mangrove classification process. Let W_1, W_2, \dots, W_n be the pixels of the North Jakarta area having p -variables. The mapping of mangrove is conducted by checking the similarity between each pixel W_k , for $k = 1, 2, \dots, n$, and the spectral references. Each pixel is classified into one of the three classes based on the value of the smallest distance. The pixel of W_k is classified as mangrove if its distance with the mangrove spectral reference is the smallest one. Figure 5 shows the mapping result of mangrove detection in North Jakarta.



Fig. 5. The mapping of mangrove in the north of Jakarta on 2015

The classification result is evaluated in two experiments. Google Earth imagery and our ground truth are used as the evaluation data. In the first experiment, we choose two areas that are classified as mangrove and non-mangrove. We compare the results of our classification model to the actual land cover. The visualization of this experiments can be seen in Fig. 6.



Fig. 6. Evaluation of the first experiments on the performance of mangrove classification using robust kurtosis projection

In the second experiment, we compare two results of classification, the classification based on the robust kurtosis projection and the classification using the classical method. Figure 7 illustrates the comparison. The failure of detection is found in the classical method; the existing mangrove land is detected as non mangrove land at the coordinates (Lat: 6.111449; Long: 106.768597). This tells us that robust approach has better performance than the classical method, see Fig. 7.

The arithmetic mean \bar{X} is often considered as a location estimator in the classical method. The arithmetic mean is the sum of a collection of numbers divided by the number of numbers in the collection. In this research we build the classical distance for classical detection of mangrove land. The distance measures the similarity of an every pixel to the arithmetic mean \bar{X} . The classical distance is very sensitive to an anomolous observation or an outlier. The occurrence of one or more outliers shifts the mean vector toward outliers and the covariance matrix becomes to be inflated.

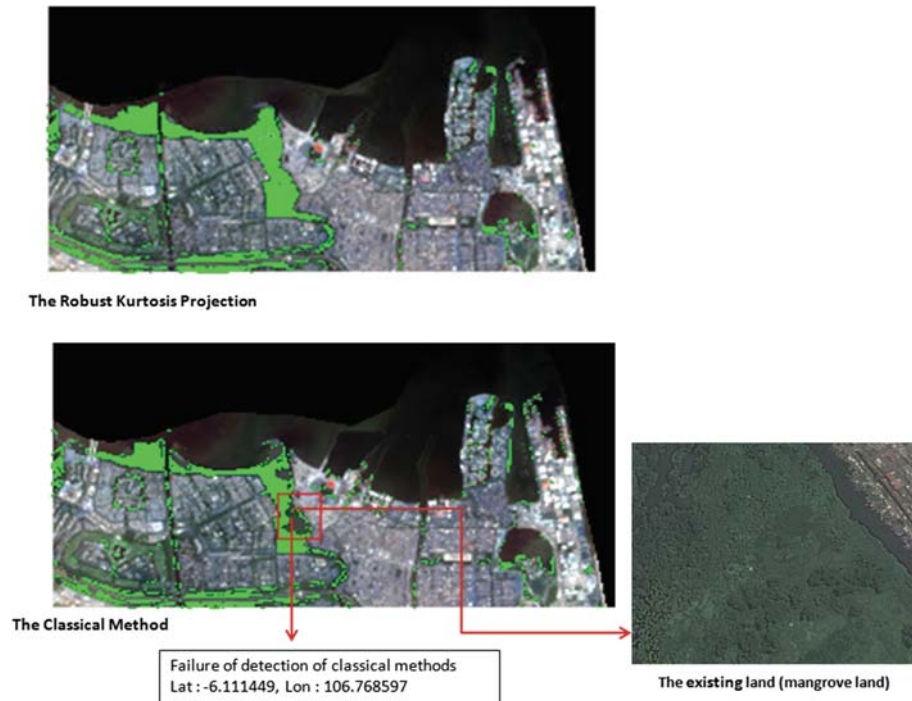


Fig. 7. The performance comparison of robust kurtosis projection performance and classical method for mangrove classification

5 Remark

The experiments and the evaluations prove that the robust kurtosis projection is an effective method that can be considered for mangrove classification. The process of mangrove land detection is not trivial. The mangrove land detection is different from the detection of green land. From the research experience, we need an advanced method that can confidently detect the mangrove ecosystem from satellite imagery. The special characteristics of mangrove is its growing location that can coincide with other green vegetation. This characteristic causes failure in mangrove detection. The empirical results provide strong evidence that the robust kurtosis projection has good performance for mangrove classification.

This study is our preliminary research. Regarding to the study results, we will use the robust kurtosis projection method for our further research; that is the mapping of mangrove land in the Central Java Province Indonesia.

References

1. Vaiphasa, C.: Remote sensing techniques for mangrove mapping. Unpublished thesis, Wageningen University (2006)
2. Center for International Forestry Research. Mangroves: a global treasure under threat. <http://blog.cifor.org/31178/indonesian-mangroves-special-fact-file-a-global-treasure-under-threat>
3. Herwindiati, D.E., Djauhari, M.A., Mashuri, M.: Robust multivariate outlier labeling. *Commun. Stat. Simul. Comput.* **36**(6), 1287–1294 (2007)
4. Herwindiati, D.E., Sagara, R., Hendryli, J.: Robust kurtosis projection for multivariate outlier labeling. In: *International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pp. 97–102. IEEE, Depok (2015)
5. Murdiyarto, D., Purbopuspito, J., Kauffman, J.B., Warren, M.W., Sasmito, S.D., Donato, D. C., Manuri, S., Krisnawati, H., Taberima, S., Kurnianto, S.: The potential of Indonesian mangrove forests for global climate change mitigation. *Nat. Clim. Change* **5**(12), 1089–1092 (2015)
6. Hawkins, D.M.: The feasible solution algorithm for the minimum covariance determinant estimator in multivariate data. *Comput. Stat. Data Anal.* **17**(2), 197–210 (1994)
7. Peña, D., Prieto, F.J.: Multivariate outlier detection and robust covariance matrix estimation. *Technometrics* **43**(3), 286–310 (2001)
8. Danielsen, F., Sørensen, M.K., Olwig, M.F., Selvam, V., Parish, F., Burgess, N.D., Hiraishi, T., Karunakaran, V.M., Rasmussen, M.S., Hansen, L.B.: The Asian Tsunami: a protective role for coastal vegetation. *Science* **310**(5748), 643 (2005)
9. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A.: *Linear models: robust estimation*. In: *Robust Statistics: The Approach Based on Influence Functions*, pp. 307–341. Wiley (1986)
10. Colwell, J.E.: Vegetation canopy reflectance. *Remote Sens. Environ.* **3**(3), 175–183 (1974)
11. Friedman, J.H.: Exploratory projection pursuit. *J. Am. Stat. Assoc.* **82**(397), 249–266 (1987)
12. Kathiresan, K., Bingham, B.L.: Biology of mangroves and mangrove ecosystems. *Adv. Mar. Biol.* **40**, 81–251 (2001)
13. Djauhari, M.A.: Improved monitoring of multivariate process variability. *J. Qual. Technol.* **37**(1), 32–39 (2005)
14. Brown, M.E., Pinzón, J.E., Didan, K., Morisette, J.T., Tucker, C.J.: Evaluation of the consistency of long-term NDVI time series derived from AVHRR, SPOT-vegetation, SEAWIFS, MODIS, and Landsat ETM+ sensors. *IEEE Trans. Geosci. Remote Sens.* **44**(7), 1787–1793 (2006)
15. Huber, P.J.: *Robust Statistics: Wiley Series in Mathematical Statistics*. Wiley, New York (1980)
16. Huber, P.J., et al.: Robust estimation of a location parameter. *Ann. Math. Stat.* **35**(1), 73–101 (1964)
17. U.S.G.S.: Landsat 8 data user handbook. <https://landsat.usgs.gov/sites/default/files/documents/Landsat8DataUsersHandbook.pdf>