Big Data Analytics for Predicting Coral Bleaching in Samui Island Area, Suratthani Province

Tanatpong Udomchaipitak

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Applied Mathematics and Computing Science

Prince of Songkla University

2022

Big Data Analytics for Predicting Coral Bleaching in Samui Island Area, Suratthani Province

Tanatpong Udomchaipitak

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Applied Mathematics and Computing Science

Prince of Songkla University

2022

Thesis Title	Big Data Analytics for Predicting Coral Bleaching in Samui Island Area, Suratthani Province
Author	Mr.Tanatpong Udomchaipitak
Major Program	Applied Mathematics and Computing Science

Major Advisor

....

(Assoc. Prof. Dr. Supattra Puttinaovarat)

Co-advisor

....

(Asst. Prof. Dr. Nathaphon Boonnam)

Examining Committee:

....Chairperson

(Assoc. Prof. Dr. Jirapond Muangprathub)

....Committee

(Assoc. Prof. Dr. Supattra Puttinaovarat)

....Committee

(Asst. Prof. Dr. Nathaphon Boonnam)

....Committee

(Asst. Prof. Dr. Jinda Kongcharoen)

....Committee

(Asst. Prof. Dr.Teerayut Horanont)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Science Degree in Applied Mathematics and Computing Science.

....

(Asst. Prof. Dr. Thakerng Wongsirichot)

Acting Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgment has been made of any assistance received.

……………………………………………………………Signature

(Assoc. Prof. Dr. Supattra Puttinaovarat)

Major Advisor

……………………………………………………………Signature

(Asst. Prof. Dr. Nathaphon Boonnam)

Co-advisor

……………………………………………………………Signature

(Mr. Tanatpong Udomchaipitak)

Candidate

I hereby certify that this work has not been accepted in substance for any degree, and is not being currently submitted in candidature for any degree.

……………………………………………………………Signature

(Mr. Tanatpong Udomchaipitak)

Candidate

ชื่อวิทยานิพนธ์	การวิเคราะห์ข้อมูลขนาดใหญ่สำหรับการทำนายปะการังฟอกขาวในพื้นที่ เกาะสมุยจังหวัดสุราษฎร์ธานี
ผู้เขียน	นายธนัชพงศ์ อุดมไชยพิทักษ์
สาขาวิชา	คณิตศาสตร์ประยุกต์และวิทยาการคำนวณ
ปีการศึกษา	2565

บทคัดย่อ

งานวิจัยนี้เป็นการศึกษาการฟอกสีปะการังโดยรวบรวมข้อมูลอุณหภูมิน้ำทะเลและความเป็นกรดของน้ำทะเล ผ่านเทคโนโลยีอินเทอร์เน็ตของสรรพสิ่งร่วมกับ LoRa และการเก็บข้อมูลดาวเทียมเพื่อทดสอบประสิทธิภาพของแบบจำลองต่าง ๆ สำหรับการทำนายการฟอกขาวของปะการังบริเวณเกาะสมุย จังหวัดสุราษฎร์ธานี ซึ่งประกอบด้วย SVM, Naive Bayes, Logistic Regression Model และนำเสนอข้อมูลโดยใช้เทคนิคการวิเคราะห์เชิงพื้นที่ ผลการศึกษาและการทดสอบ พบว่า พารามิเตอร์ที่สำคัญสามประการสำหรับการพัฒนาอุปกรณ์ LoRa ได้แก่ Spreading Factor, Bandwidth และ Code Rate พารามิเตอร์ที่สำคัญที่สุดในการตั้งค่าและส่งผลต่อค่า RSSI คือ Spreading Factor หลังจากการศึกษาและทดสอบ LoRaผู้วิจัยได้ทดสอบประสิทธิภาพของแบบจำลองในชุดข้อมูลต่าง ๆ พบว่าแบบจำลอง SVM มีความแม่นยำที่ดีกับข้อมูลจากทุ่นที่ติดตั้งอุปกรณ์อินเทอร์เน็ตของสรรพสิ่ง นอกจากนี้ประสิทธิภาพของแบบจำลองได้รับการทดสอบโดยใช้เทคนิค Split test และ K Fold Cross Validation หลังจากทดสอบประสิทธิภาพของแบบจำลองเสร็จเรียบร้อย ได้ทำการนำเสนอแผนภาพที่รวบรวมจากแหล่งข้อมูลต่าง ๆ แสดงให้เห็นความเสี่ยงของการฟอกขาวของปะการังในช่วงระยะเวลาการเก็บรวบรวมข้อมูล แหล่งข้อมูลที่แสดงให้เห็นความเสี่ยงอย่างชัดเจนที่สุด คือ แหล่งข้อมูลจากทุ่นที่ติดตั้งอุปกรณ์อินเทอร์เน็ตของสรรพสิ่ง

Thesis Title Big Data Analytics for Predicting Coral Bleaching in Samui Island Area, Suratthani Province

Author Mr. Tanatpong Udomchaipitak

Major Program Applied Mathematics and Computing Science

Academic Year 2022

ABSTRACT

This research investigates coral bleaching by collecting sea temperature data and seawater acidity. Through the technology of the Internet of Things together with LoRa and satellite data acquisition. We test the effectiveness of predicting coral bleaching in the Samui island area of Suratthani province which consists of an SVM, Naive Bayes, Logistic regression model, and data visualization using spatial analysis techniques. The result of this study was found that three important parameters for the development of LoRa devices are spreading factor, bandwidth, and code rate. The most important parameter to set up and affect the RSSI value is the spreading factor. After we studied and tested the system with LoRa, we tested the effectiveness of the model in various datasets. We found that the SVM model had accurate on data from a pontoon. The model's performance was tested using two techniques: split test, and K fold cross validation. We visualize of information gathered from various sources shows the risk of coral bleaching during the data collection period. The most clearly shown source is the pontoon source.

ACKNOWLEDGMENT

I would like to express my sincere thanks to my thesis advisor, Assoc. Prof. Dr. Supattra Puttinaovarat and Co-advisor, Asst. Prof. Dr. Nathaphon Boonnam for their teaching, support everything, the better suggestion, and the most important thing is the encouragement that both teachers have given me all along. This thesis may not be completed without all support from them throughout the course of this research.

I would like to thank my family, especially my mother for always supporting and encouraging me. Thank you for the opportunity to give the little bird an unconditional flight in the vast world. My mother did her job perfectly without giving up on any obstacles, for this reason, I was never discouraged when I encountered obstacles during the course of my thesis.

I would like to express my sincere thanks to Asst. Prof. Dr. Suwat Jutapruet, Asst. Prof. Dr. Rattanasak Hama, Mrs. Pattaraphon Hama, and all my friends for the support in the installation of equipment in Koh Samui, which is a very important process for the completion of this thesis, and finally, I would like to thank all lecturers from Master of Science in Applied Mathematics and Computing Science for their advice and support in various matters thus making this thesis work perfectly.

Thank you to the Marine and Coastal Resources Office 4, Surat Thani (MACRO 4) and Koh Tan Pakarang Beach for providing the facility and equipment for us to install the equipment in the case study area. which is very important for this thesis.

I would like to thank Prince of Songkla University This research was funded by Prince of Songkla University, Thailand. And I would like to thank Assoc. Prof. Dr. Seppo Carrila and the Publishing Clinic for assistance in preparing the manuscript. Finally, I would like to thank the personnel of the Faculty of Science and Industrial Technology who assisted in various matters that made this thesis work flawless and complete.

Tanatpong Udomchaipitak

CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (CONT.)

LIST OF PAPER AND PROCEEDINGS

1. Nathaphon Boonnam, Tanatpong Udomchaipitak, Supattra Puttinaovarat, Thanapong Chaichana, Veera Boonjing, and Jirapond Muangprathub, "Coral Reef Bleaching under Climate Change: Prediction Modeling and Machine Learning", MDPI (Sustainability), vol.14, no.10, pp.1-13, 19 May 2022.
2. Tanatpong Udomchaipitak, Nathaphon Boonnam, Supattra Puttinaovarat, and Paramate Horkaew, "Forecast Coral Bleaching by Machine Learnings of Remotely Sensed Geospatial Data", International Journal of Design & Nature and Ecodynamics, vol.17, no.3, pp.423-431, June 2022.
3. Tanatpong Udomchaipitak, Nathaphon Boonnam, and Supattra Puttinaovarat, "An Experimental Study of RSSI for LoRa Technology in Different Bandwidths", The 37th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC).

Chapter 1

Introduction

1.1 Rationale

Corals are large marine ecosystems where various animals use them as a source of food and habitat. Corals are also the spawning grounds of marine life. It is a natural wave break that helps reducing the wave's intensity before they reach shore, which is more effective than man-made breakwaters. The structure of the coral is calcium carbonate or limestone formed by the deposition of sediments that are washed by seawater. Forming corals then create tissues opening up different types of algae to provide their energy. Because corals and algae coexist independently, in other words, corals are algae habitats and algae provide power to form corals ecosystem. In addition to the benefits of coral reefs for marine life, it can be also humans' benefit in terms of tourism as well. Due to the situation the problem of global warming having a wide impact on the world, it spreads to the underwater world which is the habitat and food source of marine life. It also affects the corals that host them with problems caused by rising global temperatures and being coral bleaching, respectively a phenomenon has been occurred for many years decades, or maybe even centuries, have come into existence in this world. The bleaching of corals is caused by factors in the coral environment that are not suitable for it. This stresses the corals and affects the algae within the coral, which are then pushed out of their original habitats to find new habitats. When the algae that live in the coral leave can see it as its skeleton. Coral bleaching occurs that the corals do not have the energy to sustain life. As a result, during the entire bleaching period, corals will have lower immunity than normal and this can lead to the loss of beautiful corals as shown in Fig. 1.

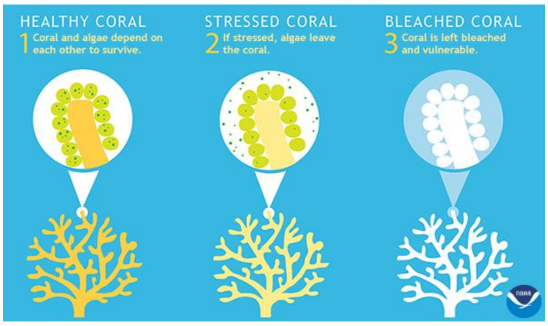


Fig. 1. Coral bleaching (https://oceanservice.noaa.gov/facts/coral_bleach.html., 2021).

Therefore, it is the researcher for the source or factors that affect coral bleaching. Many scientists have helped to find various factors that affect coral reefs. There are many factors causing bleaching. But the main factors that are recognized as the factors affect the coral. However, the main factor are the seawater temperature and acidity factor.

- Seawater temperature

  Changes in the global temperature environment result in changes in warmer sea temperatures when the temperature of the warm water exceeds appropriate (S. Hafoud et al., 2022). This puts stress on the coral, which in turn causes the coral to expel the algae that live from it. This allows us to see the coral's color fade to white or coral bleaching. Corals normally survive during bleaching and will gradually die if their environment does not improve. Still, warming of the sea is not all the reason for coral bleaching (Oceanservice.noaa.gov., 2021).

- Seawater acidity

  The acidity of the seas is caused by changes in global temperature changes as a result of the increase in Carbon dioxide (CO_2) in the atmosphere. This causes ocean acidification because corals are very sensitive to ambient sea conditions, which can be very catastrophic. This is the result of human action that produces carbon dioxide in the atmosphere. Combining carbon dioxide with water and carbonates produces bicarbonate, which if too much of it can be toxic to living organisms as shown in Fig 2 (Katharina E. F. et al., 2020., NOAA, 2020.).

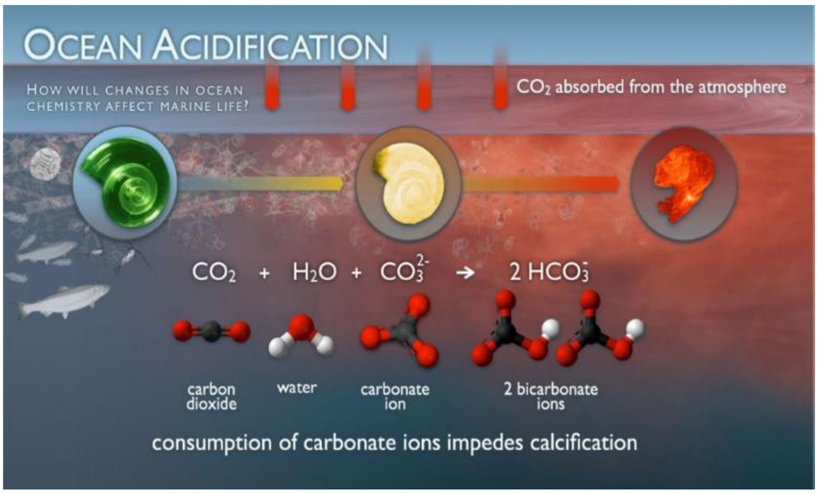


Fig. 2. Seawater acidity (https://www.noaa.gov/).

The problem of coral bleaching that is well-known and a widely used in case study area is The Great Barrier Reef (GBR), also namely as coral country. It is the largest coral reef in the world, located in Australia. This 2,300-kilometer-long shallow reef, there is home to thousands of fish, invertebrates, algae, reptiles, birds, and algae (Jeremiah G. et. al., 2015). The first recorded record of coral bleaching as a result of temperature factors occurred in 1929 and continued to occur in coral bleaching. In 2020 there was another bleaching that greatly increased the area of damage. Compared to past bleaching, this is important evidence to suggest that temperature factors affect coral bleaching as shown in Fig. 3.

Fig. 3. The Great Barrier Reef's annual bleaching comparison. (https://globalvoices.org/).

From the reports of coral bleaching as we mentioned above, this causes a lot of damage to the ecosystem looking back in Thailand, where there are many coral habitats, especially in the southern part of the country, especially Suratthani province, which has many coral-related attractions. So, we define the area of Koh Samui - Suratthani Province. It is a data repository for use in predicting results in a report to make humans aware of the value of nature and prevent coral bleaching in those study areas. There will be a lot of damage to the marine ecosystem and may affect the business and tourism in the area as well.

In this study and preparation of this thesis, the intention was to study coral bleaching in coral reef areas in order to determine the causes or factors that may contribute to coral bleaching. Once the cause or factor is clearly known, various sensor devices will be studied in order to harvest information from those factors. We find the most effective model to be used. The area used as a case study area is the Koh Samui area, Suratthani province. The reason for choosing the Koh Samui area in Suratthani Province is because of the Koh Samui is an area that can be linked to islands or tourist attractions such as Koh Phangan and many other islands, each island is inhabited by coral reefs and in the past 2019, coral bleaching in the area of Koh Samui has begun to occur as shown in Fig. 4. It has been hypothesized as to what factors have caused this incident.

Fig. 4. An aerial photograph of the coral that has begun to bleach in the Koh Samui area. (https://marinegiscenter.dmcr.go.th/gis/).

From the source mentioned above, that is why it is very important that we study the factors affecting coral bleaching in the case study area. In particular, sea temperature and seawater acidity are the main causes that affect corals greatly. Therefore, we are interested in studying data collection technologies and techniques with the aim of targeting the two main impact factors mentioned above. To use the data from these factors to come up with the most suitable prediction model for predicting coral bleaching in this thesis.

1.2 Objectives

1. To study suitable and high-performance models for predicting the bleaching corals in the Samui Island area, Suratthani Province.
2. To study the structure of Pontoon Monitoring Equipment for the marine data collection system.
3. To compare the predictive accuracy between the data derived from Pontoon Monitoring Equipment with obtained data from NOAA and the Google Earth Engine.

1.3 Outcomes

1. The most suitable and effective model for predicting coral bleaching in Samui Island area, Suratthani Province.

2. It enables to devise an internet of things equipment used for marine data collection.
3. The results were obtained by comparing the accuracy of the data from these data sources.

1.4 Scope of the Research

This research covers the study of the appreciate model. Designing and building interactive devices were used to store data, which consists of factors such as seawater temperature, seawater acidity, etc. We bring the obtained data to predict and compare the accuracy of two other data sources, NOAA and Google Earth Engine. They are the sources with the most accurate predictions. The results obtained from that dataset are compiled into a report where puts the device to store data. We study the factors that cause the beginning of coral bleaching in the area of Koh Samui as shown in Fig. 5.

Fig. 5. Aerial photography of Koh Samui Island, Suratthani Province.

(https://marinegiscenter.dmcr.go.th/gis/).

1.5 Research Methodology

1.5.1 Define research problem

In this thesis, the motivation came from the online media in the context of undersea exploration. In addition to the debilitating of the dangers of the sea with only animals, there is no mention of the habitats of those organisms, such as corals, where problems with corals have been around for a long time, but it has not yet been solved seriously, which is why the researcher came up with this thesis topic to make humans conscious of the value of natural resources.

1.5.2 Review the literature

In the first part of research, the corals were fed in three different temperature groups (Melissa S. Roth et al., 2012), in the first group with submerged water temperatures that corals could live in. It is the part of the coral group that is in the controlled water temperature which is the right one for the corals and the last group is the group where the corals are in warm water temperature. A total of 20 days in the first trial, starting from 0-5 days, can clearly see the local coral transformation in the warm and cold-water population. However, there is one sentence that can be easily summarized in a research paper saying: "Cold water temperatures in the short term are more harmful to corals than the temperature of the algae. Heated water for a short time, but the temperature of hot water for a long time is more dangerous than cold water temperature for a long time." When there are different causes as for the bleaching of corals, it is necessary to have a device for measuring marine measurements, which in (Jeremiah G. Plass-Johnson et al., 2015), the Internet of Things in examining the marine environment, which in the research work, has proposed a system of architecture in the system as follows:

- The awareness and action layer are responsible for collecting information.
- The data transmission layer is responsible for transmitting data through the network.
- The pre-processing layer cleans the dataset before it is processed.
- The application layer is responsible for serving the needs of the user.
- The business layer is responsible for building a business model and reviewing the four previous layers to improve and maintain user privacy.

In addition, it describes the nodes to monitor the environment with four main parts:

- Mote part: consists of different sensors that collect information to the system dividing into two types: internal sensors and chemical sensors.
- Microcontroller part: the processing of the data values from the mote section.
- The signal transmission part: the transmission of communication signals via radio waves or transmission of data via satellites and radars.

- Power supply part: consists of an energy storage and management device.

All of the above will work together in a systematic manner to help the marine environment monitoring system operated at its best (Guobao Xu et al., 2019). When the data has been harvested, it comes to the end of which is the prediction of the data that has been harvested. We will study predicting about coral bleaching (Scott Wooldridge and Terry Done., 2004) that takes different factors. It is possible to produce coral bleaching through the use of multiple sources and help in this prediction, but the factors that are important and supporting for each other It is the thermal stress factor and the community factor of the coral. Experiments show suitable results after several experiments. The model used in this research is the Bayesian belief network model, and the PCA technique was used to help in clustering this prediction. All coming from global warming is the skill of humans who do not value natural resources.

1.5.3 Formulate hypotheses

Determine the factors cause coral bleaching in order to collect the data from the above factors for analysis to obtain the prediction results of all the sources to compare the accuracy of which sources have the same values. The most accurate from to make a report.

1.5.4 System Design

Determine the area used in the study or experiment. In order to provide enough equipment to be used in the survey for this study, when the equipment is fully procured, these devices will be floated at the specified point for the device to collect data into the prepared database system. When there are data values, simply predict them in multiple models to determine the correctness. Data collected from other sources is also processed in that model. To compare the accuracy of each source to how accurate they are select the prediction results obtained from the predicted data source with the highest accuracy to prepare a report. By checking the data value of the device through ThingSpeak as shown in Fig. 6.

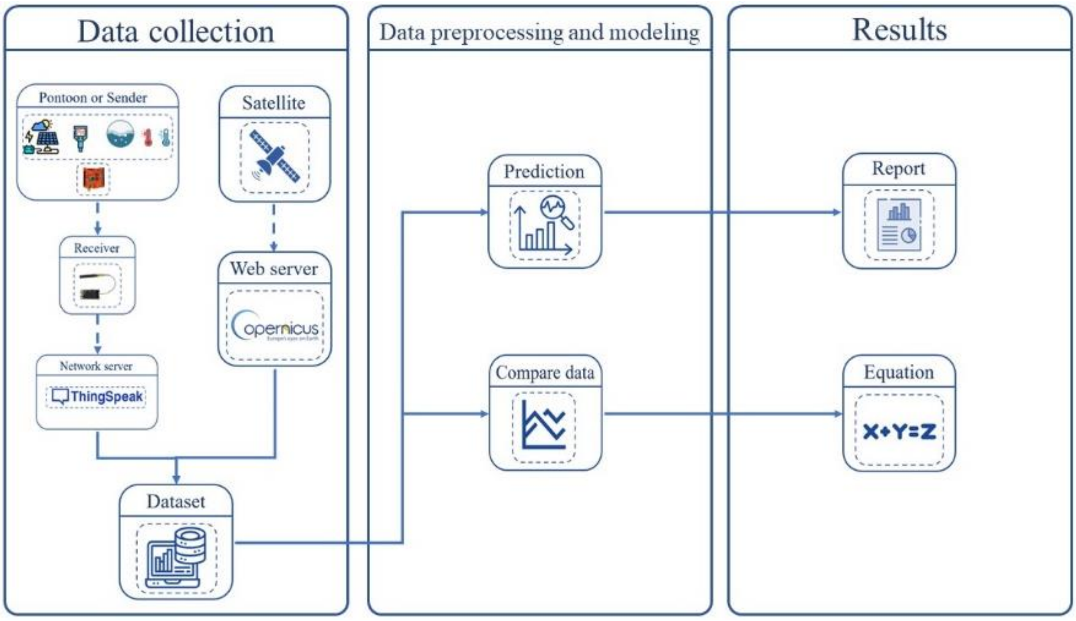


Fig. 6. Operation flow diagram.

1.5.5 System Development

When data is obtained from a variety of factors, the sensor data is collected in conjunction with the LoRa technology developed to provide stability while inside the buoy. The various data will be stored in a large amount and waiting to be synthesized the information that has come out for use the next step is to predict future coral bleaching based on the current data with various models. The data used in this thesis may be partially from satellites.

1.5.6 Analyze and Interpret data

Analysis of data obtained from sensors as we mentioned in the preceding section, it is the process of introducing data into different processes, starting with data purification to prevent any other predictive errors. This is an important part of data analysis. When the information is obtained and then we find the grouping of the data, we get from our sensor data collection in this step, bringing it in to know the group of data. In other words, we get interpretation of the coral state of each data group. Since we study the model having the highest predictive accuracy to be used in further research. The models of interest in the experiment were used to compare the best models such as:

A decision tree (Harsh H. Patel and Purvi Prajapati, 2018) makes up of multiple nodes: Root nodes, branches, and leaves, the results are in the leaf node, with the root node being the primary node of all nodes and the top-tier node of the tree. A decision tree is a tree where each node represents an attribute, each branch represents a decision (rule), and each one represents a result. Since the decision tree mimics the human mind, it is easy to pick up information and interpret it. Making it easier to understand, decision trees can resolve situations both continuously and discrete. When it comes to decision trees, an algorithm named ID3 is the fundamental algorithm used to create structured decision trees using news theory and measured values to decide which variables to use in the prediction or data type. In addition to ID3, there is an algorithm called CART that can be used for both tasks. Classification or the use of principle regression is to create a branch that selects one characteristic and then divides the data into two parts according to the criterion of that particular trait as shown in Fig. 7.

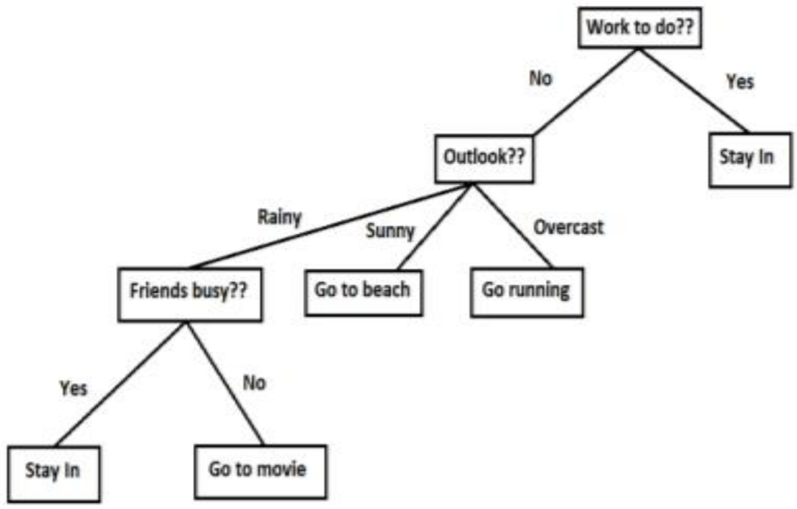


Fig. 7. Example of decision tree result (Harsh H. Patel and Purvi Prajapati, 2018).

1.5.7 Thesis preparation, defense, and improvement

Summary of key factors contributing to coral bleaching to get a device checks those factors. The data was then obtained from the equipment at sea. We make predictions of situations that may cause coral bleaching in the study area for thesis preparation, defense, and improve the thesis documentation.

Chapter 2

Literature Review

2.1 Coral Bleaching

Coral is an ecosystem very important to the marine environment because it is a habitat and refuge for animals which spread widely distributed on the coast and each ocean, it is attracting the attention of tourists. Many corals live together, we call the coral reef ("Bleaching" is the bleaching of the tissues of coral that are home to algae by bleaching as a result of algae habitats not being suitable, make them look for new habitats. So, we can see the coral "skeleton"). The biggest of coral reef lives in Australia, namely the great barrier reef. The effect of climate change on the environment makes them sick. The problem of sick coral occurs in coastal areas around the world such as the great barrier reef areas, etc. NOAA report about coral's sickness has look white color all. Problem that coral change color to white was known as coral bleaching. Since 2016 – 2020, NOAA report coral bleaching was found more every year as shown in Fig. 8.

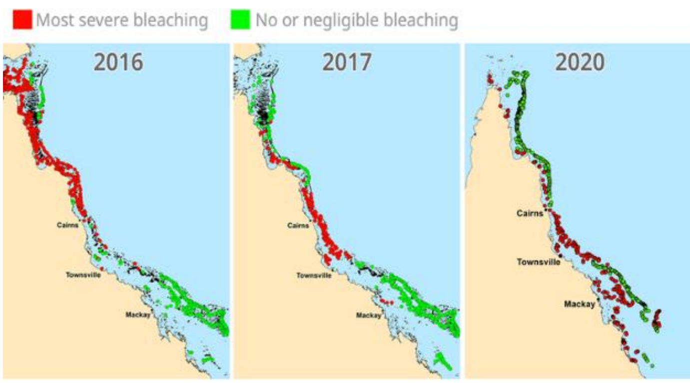


Fig. 8. Report Coral bleaching in the great barrier reef from NOAA (https://www.climateaction.org/).

In the part of coral bleaching have been more than 100 years ago. We found first recorded evidence of coral bleaching in the year 1929, the first and worst bleaching occurred around the world. This resulted in 16% of coral deaths worldwide, of which 80% of all coral in the western Indian ocean was almost 50% bleached, which resulted in 95% of the above 50% bleached corals. The reasons cause coral bleaching resulted from the main factors are as follows:

2.1.1 Seawater temperature

The seawater temperature is the main factor in the effect of climate change. It makes including more the temperature every year. The reason when the temperature of the sea change makes algae in coral leave away is due to, they need the right temperature. Melissa S. Roth et al. (2012) learn about coral reef testing by allowing them to live in water with different temperatures, for a period of more than two weeks. In range of the first week was found algae in coral (warm water) reduce and more over time. Until finish, the experiment was found algae in coral (warm water) lost until only the skeleton or coral bleaching while algae in other groups also live with corals, other groups of algae live with corals. Although the number of algae is not the same because coral in cool water will react during the first period of the experiment. So, it can be concluded that the algae of coral that live in cool water can live for a long term on the other hand, the short term the algae will react with water cool by off from coral. But the algae of coral that live in warm water can live for a short term and the same it cannot live in warm water for a long-term as shown in Fig. 9.

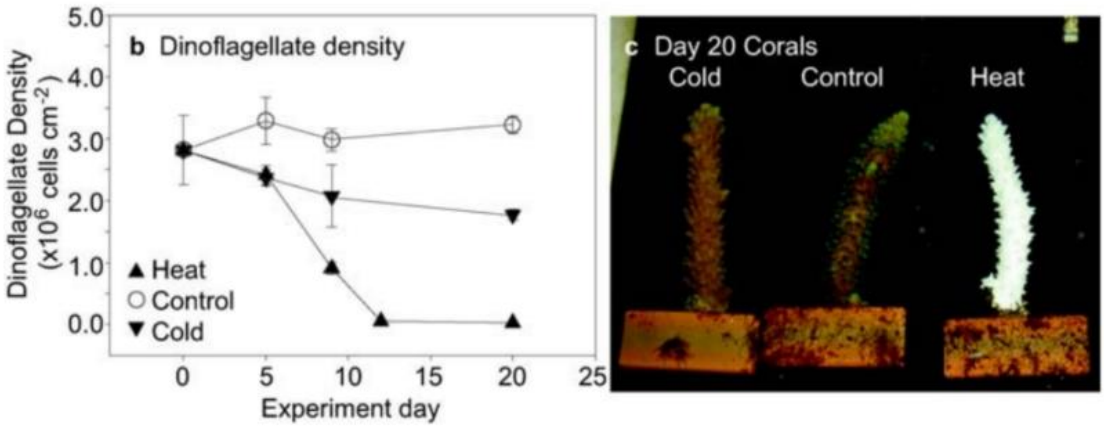


Fig. 9. Experiment's corals living in different temperature waters (Melissa S. Roth et al., 2012).

While the Department of Marine and Coastal Resources (DMCR) specify about coral bleaching, normal coral in Thailand can live in water temperature 28 – 29 degree Celsius. But water temperature increases between 30 – 31 degrees Celsius or more than for a period of time 3 – 4 weeks or more, it can bleach (DMCR, 2017; Marine Knowledge Hub, 2018).

2.1.2 Seawater acidification

The acidification of the sea can change with sea surface temperature (SST), which causes coral bleaching. Joost W. van Dam et al. (2011) studied about pollution on coral reef. The chemical in water may attack damage to coral reef, but it depends on persistence of chemical and environment in which it is contaminated. Especially the chemical used for the extermination of insects in the agricultural industry is endosulfan, which can contaminate water bodies and maintain the condition of its and flow into the sea. In addition to these chemicals, there are many chemicals that can affect coral reefs such as chlorpyrifos etc. It may be had another chemical can attack coral suck as oil. All of the chemicals listed above It can have a huge impact on corals. If there is a leak (Joost W. van Dam et al., 2011). Department of Marine and Coastal Resources (DMCR) provides knowledge about discharge of wastewater into the sea in Andaman coastal area, Thailand that can attack coral to bleach or die. It causes the degradation of coral reefs under the sea. This problem, that they are most often found in large community areas near the coast, does not have a wastewater treatment system before releasing it into natural water sources (DMCR., 2017). In coral bleaching, there are usually two main coral bleaching mechanisms: Inhibition of photosynthesis: a problem that happened to algae due to stimulated by various factors in its environment which reaction with algae to weak that effect to coral making algae cannot generate energy to enough for corals (Jeremiah G. Plass-Johnson et al., 2015) and Oxidation stress: it born from creation radicals too fast and too much due to sea temperature includes more, that make radicals dominate antioxidant of algae cause a lot of damage (Jeremiah G. Plass-Johnson et al., 2015).

2.1.3 Another factor

Dive event of human due to present it is popular event for travelers who like challenges, to meet the beautiful creatures under the sea, to research, or another, etc. Dive events, therefore, are considered reason damage for coral because some divers may be careless about their equipment, due to a dive event having much equipment. So, equipment-intensive dives such as SCUBA should pay more attention to equipment that can damage corals. Suchai Worachananant et al. (2008) studied about damage for SCUBA divers to coral reef. In their research effect from SACUBA divers, with their equipment such as FIN (Equipment for moving underwater). A FIN can make damage more than you know, which effect to coral may be from inattentive them. Department of Marine and Coastal Resources (DMCR) comments about the danger of diving no matter skin dive or SCUBA due to tourism may tread coral, which it can make damage to coral (DMCR., 2017).

2.2 Spatial Analysis

Spatial analysis is the process of manipulating data in a spatial form. Such analysis is usually carried out by a Geographic Information System (GIS), which provides tools to assist spatial analysis and computation in the form of statistics (QGIS, 2021). The reason to use spatial in this thesis because it includes many techniques to interesting such as interpolation techniques, IDW techniques etc. It can analysis the data for prediction and visualization in a map format. To show the change in spatial is, therefore, a technique that plays an important role in the study of changes in another area (GeoffBoeing, 2021) which have a benefit to this thesis for prepared and visualization data. It is considered another important part of this thesis study. Spatial analyzes were used experimentally for testing the coral bleaching hypothesis with nitrogen and heat stress as factors of bleaching in two genera of corals (*Pocillopora* and *Acropora*). Through the prediction of the Bayesian hierarchical model, the results showed that both corals responded in a similar pattern to the spatial difference in nitrogen content. Coral bleaching is more intense in the presence of nitrogen (Mary K. Donovan et al., 2020). In predicting coral bleaching, several techniques are used to assist in the prediction, for example using the Interpolation For improving the incoming data to optimize the prediction to the best accuracy (Alejandra Virgen Urcelay, 2021). Interpolation It is a spatial technique for interpolation that is widely used in science. It is applied in the form of a weather simulation. Its libraries have been developed over the years, with them being modified in linear, bifacial, multilinear, and conservative data formats. The interpolation technique is often applied to information about resources on Earth, such as energy or water. Interpolation conservatives are often given special attention for jobs in various fields. For this reason, it is a method that is used more widely than others (Alexander Pletzer and Wolfgang

Hayek, 2018). As mentioned above, spatial have many techniques to use. So, in this thesis we used some technique in prepared and visualization. We used a total of 2 technique include:

- Interpolation

Spatial interpolation is the process of using known values to estimate other points. Such as precipitation mapping because in many countries found that the location of rain measurement stations regularly and sufficiently of course, in the non-installed areas of the station, no data will be available. So, spatial interpolation is possible, it was brought in to solve these problems. This technique is therefore suitable for a variety of tasks such as estimating rainfall. Accumulation of snow, etc., by interpolation It can be divided into several techniques (QGIS., 2021), especially, Inverse distance weighted (IDW) interpolation is a technique used to measure and correct the spatial data of a data point in order to provide it with another nearby point. The influence of IDW will be depending on the distance of the points. If it is a nearby point, it will be affected by IDW more than the point farther away (ArcGIS., 2021) as shown in Fig. 10.

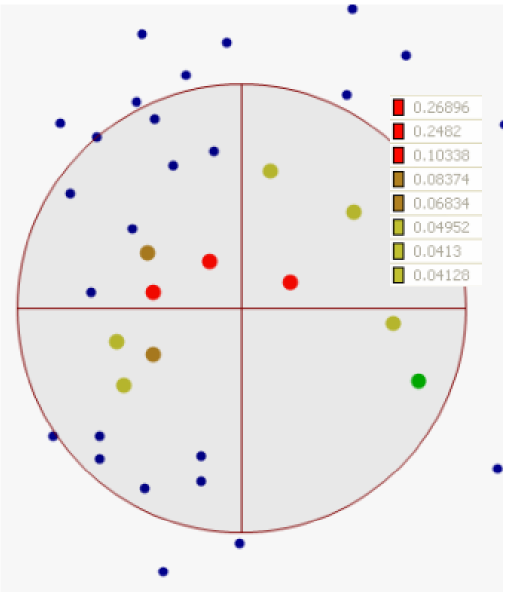
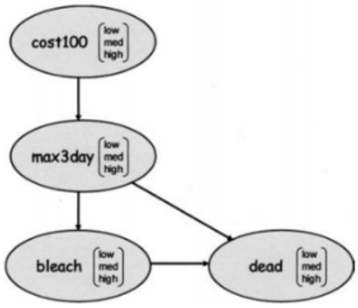


Fig. 10. Inverse distance weighted (IDW) interpolation (https://pro.arcgis.com/).

It shows that the points near the midpoint are influenced by the IDW Interpolation maximally and decrease as the distance travels.

2.3 Data Prediction

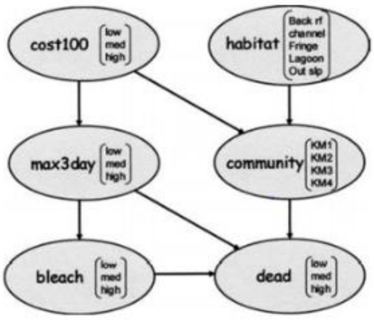
As part of the prediction, M. Ateweberhan and Tim R. McClanahan (2010) investigated the relationship between past sea surface temperature variability and coral mortality caused by climate change in the Indian Ocean. In 1998, an area of the West Indian Ocean was killed, which generated huge field data. In the same year, 36 major coral reef sites were collected by collecting data from JCOMM-SST as an intergovernmental agency of technical experts that provides mechanisms for international coordination of maritime and meteorological observations, data management and services by combining technological expertise and capacity building. Capabilities of the Meteorological and Oceanic Community Instead, it extracted data from NOAA, a satellite observation review, and used the data to perform a mixed multiple regression analysis to obtain average coral cover results in climate variability events in 2016. When collecting the values according to the determined factors or meeting the factors mentioned in the above research, those factors information was used in the field of prediction, which (Scott Wooldridge and Terry Done, 2004). A study was conducted on learning to predict large-scale coral bleaching from past events by Bayesian methods using remote sensing data. Information on coral bleaching sites and institutions conducting research or study on coral Inside, it discusses coral bleaching as a result of heat stress and the ecological effects of bleaching. Coral researchers conduct a coral reef ecosystem study aiming to understand as much as possible about past and future bleaching events processes to make reasonable predictions. In addition, a case study of the Great Barrier Reef (GBR) bleaching in the area has been described as an anomaly caused by heat, bleaching and coral death. which noted that There are living corals that should either be bleached or dead. Researchers have pointed out that survival rates tend to be tides. They also discussed the cooling mechanism of water called Thermocline (Harsh H. Patel and Purvi Prajapati, 2018). If diving to the Thermocline layer, there will be a sudden change in water temperature and the deeper you dive, the temperature will be much lower than the surface temperature. It is considered part of helping to cool the water in the layers in which the corals live. As for prediction, many variables were used. For example, the water surface temperature used in this study used a maximum temperature in summer (max 3 days), for example. The Bayesian Belief Network (BNN) model, in which the input data for the predictions, were obtained from a total of 150 sites where data were collected in this field using PCA techniques to help in grouping the data in the experiment, various variables were randomly sampled to find the correlation and had the highest predictive accuracy. Combining heat stress and community indicators will provide better predictions for coral mortality. than using a single variable of heat stress.



Predicted Coral Mortality			Observed	Predicitive Rate
Low	Medium	High		
13	18	5	Low	13/36 = 0.36
6	**46**	14	Medium	46/66 = 0.70
1	15	**32**	High	32/48 = 0.67

Fig. 11. Results and predictive factors, set A (Wooldridge and Terry Done, 2004).

This is the data from the factors as shown in Fig. 11 on the left, namely cold-water flow factors. 3-day peak temperatures in summer, bleaching, mortality, community and habitat were used to predict Set A.



Predicted Coral Mortality			Observed	Predicitive Rate
Low	Medium	High		
26	7	3	Low	26/36 = 0.72
11	**46**	9	Medium	46/66 = 0.70
1	13	**34**	High	34/48 = 0.71

Fig. 12. Outcome and predictive factors, set B (Wooldridge and Terry Done, 2004).

This is the data from the factors as shown in Fig. 12 on the left, namely cold-water flow factors. Summer 3-day peak temperatures, bleaching, and mortality were used to predict set B from Fig. 11 and 12. The predictions A and B were significantly different from both A and B predictions. And B uses the same data set but uses different variables as in the example above. There is also research (Krisanadej J. et al., 2020) that has investigated the restoration of coral fish and coral reef communities after severe global warming. In 2010, Thailand mentioned the complexity of the marine ecosystem and its high fish diversity. There are about 5,000 species of fish in the coral reef, and this coral is a habitat. Fish nursery or maybe a source of escape for fish. In the past few decades, coral bleaching has occurred in the Khon Kae Bay area. In 2010, 90% of the damage was sustained after fish population surveys were carried out in the area of Racha Yai Island during 2010. Once the data was collected, a linear regression analysis was performed to obtain information on coral recovery and fish populations in 2013 – 2019.

2.4 Big data analytics

Predictions about the information we gathered in the data collection step. We then proceeded with the prediction based on the data analysis cycle which consisted of six main steps - Data discovery, Data preparation, Planning of data models, the building of data models, communication of results, and operationalization as shown as Fig. 13.

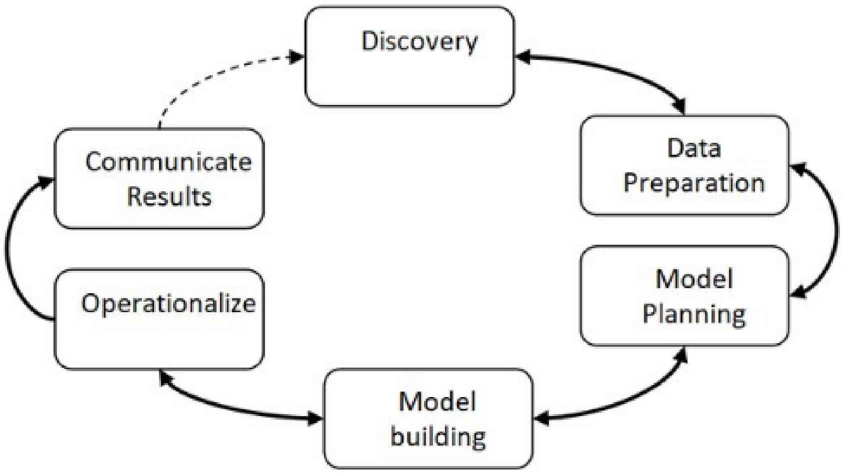


Fig. 13. Data analytics life cycle.

We clean data in order to verify the integrity of the entire dataset following a data preparation procedure, which is a procedure for verifying the loading and transformation of data, in this case, abnormal data or the missing value is generated by sending data between LoRa nodes, so we need to convert the loaded data from the network server layer to remove it from all data. The next part is planning the modeling. It is a process for a conceptualization of good models and how data are related to each other. to apply to our study The model is then created and validated against the prepared data as planned in the previous step. After this is over, we communicate with each other about the problem of analysis. So that all operations will be successful or not, when everything is done, a summary of the results is reported before being put into practice. (Sridevi Bonthu and Hima Bindu, 2018; Jirapond Muangprathub, 2018). To prove the relationship between the variables that we are interested in, the correlation technique was applied to find the relationship from the two variables by using Pearson's correlation method. It calculates the variance of the two variables and divides the two by the product of the deviation standard to get results in the form of numbers that can be interpreted shown as Eq. (1).

$$\text{Pearson's correlation coefficient} = \frac{\text{covariance}(X, Y)}{\left(\text{stdv}(X) * \text{stdv}(Y)\right)}. \qquad (1)$$

where x denotes seawater temperature and y is pH of seawater. If the result of the equation is close to 1, it means that the relationship between the variables is strongly correlated, if the result is negative, the relationship between the two variables is negative (-) that mean the relation of both is negative relation and too of the result is positive (+), then the relationship of the two is definitely positive. But if the result is expressed as 0, then the two variables are not related to each other (Philip M. Sedgwick., 2012). And we are interested in many models to be applied in this study. But there have few good models for predicting bleaching. Therefore, we studied a total of 3 models to be tested in this study.

- Naïve Bayes

Naïve Bayes is a simplified algorithm that works based on Bayes law with hypotheses which more rigorous than previously perceived terms, Bayes tends to provide accurate classification of various types of data coupled with control over its performance, thus Bayes are widely used in a wide range of applications. It is an algorithm classified in the unsupervised machine learning group, so users do not need data for identifying groups of data to train the model. (Geoffrey I. Webb, 2017). Naïve Bays have also been applied in a research paper by Nathaphon Boonnam et al. (2022), where it has been used to test its effectiveness in predicting coral bleaching in southern Thailand. With the ability of Naïve Bayes to use the probability of being a member class. And because the model has the ability to reach high accuracy and good speed, it is suitable for application in big data analysis, but there is still some caution in its work, as Naive Bayes often considers all predictors to be independent variables. which rarely happens in real life (Pavan Vadapalli.,2021).

- Support vector machine (SVM)

It is a technique that has been developed with good data classification performance. SVM is a linear classification technique in which it is classified into a variety of formats but is mainly separated into linear SVM and non-linear SVM. In which linear SVM operation is a powerful technique for working on high dimensional data applications. Such as document classification, ambiguity correction, etc. The results obtained between both SVMs from the same set of data showed that the resulting accuracy was similar, but the nonlinear SVMs took more time to train than the linear ones. (Mayank Arya Chandra and S. S. Bedi, 2018; Vinod Kumar Chauhan et al., 2019; Nathaphon Boonnam et al., 2022) However, it is worth considering how SVM is applied to various applications as SVM is an ideal model for small to medium-sized data operations. It is therefore very challenging to use such a model in data-intensive tasks such as analyzing big data too complicated (Dhiraj K., 2019) as shown in Fig. 14.

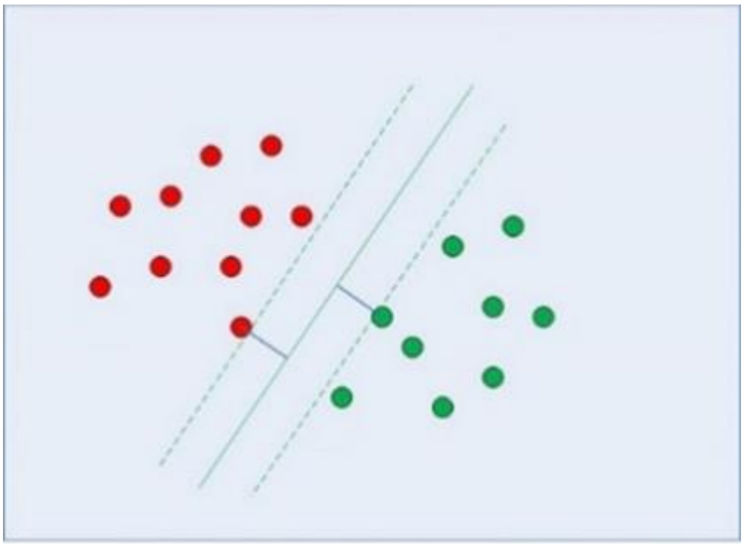


Fig. 14. SVM algorithm

(https://stackabuse.com/implementing-svm-and-kernel-svm-with-pythons-scikit-learn/).

- Logistic regression

In machine learning logistic regression, it is one of the supervised machine learning models, which are models used for estimating the probabilities of an event. It has been applied to life sciences for decades. Such a model has a unique predictive nature, that is, its predictive outcome is born or not born only this is one of the charms of this model. The nature of the work of this model (Michael P. LaValley, 2008; Saishruthi Swaminathan, 2018) for logistic regression, should be taken when predicting continuous functions. Because such models are only suitable for predicting discrete functions as shown in Fig. 15.

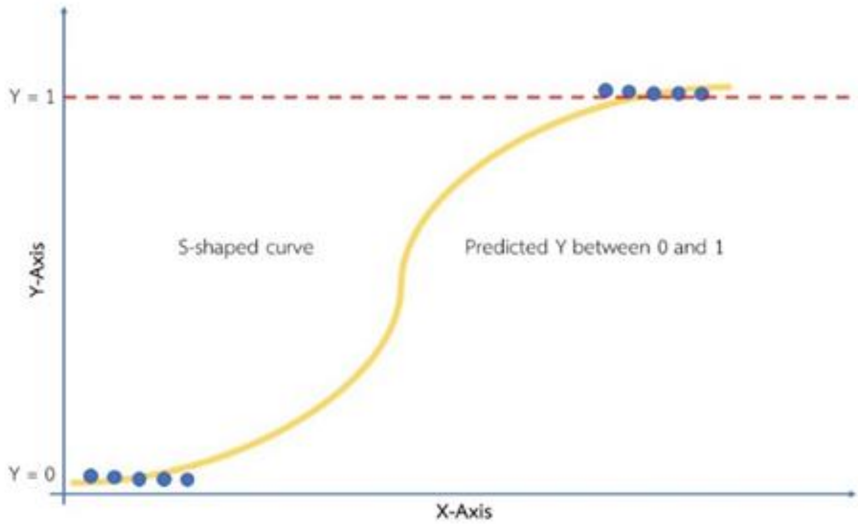


Fig. 15. Logistic regression (Amiya Ranjan Rout., 2020)

In addition to the models mentioned above We applied the predictive performance testing techniques of the three models. By using 2 testing techniques as follows:

- Split test

The split test technique divides the data from the entire dataset into two groups for training and model performance testing. In general, the data division ratio is usually in the range of 70: 30 or 80: 20. For this method of splitting the data, sometimes in the test model the data characteristics of the model training group are similar to that of the test data groups second, which is a model training group. That may result in terms of the effectiveness of predictions may be exaggerated. Therefore, it may be a limitation of the split test (Ramesh Medar et al., 2017; Eakasit Pacharawongsakda et al., 2015) as shown in Fig. 16.

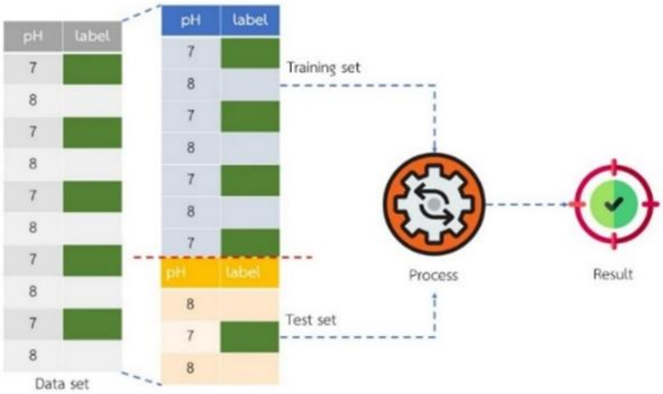


Fig. 16. Split test.

- K-Fold Cross validation

It is one of the most popular techniques in research that tests the effectiveness of models. Because the results obtained are quite reliable. Due to the data resampling method to assess the generalization ability of predictive models and to prevent overfitting and its work will split data into subgroups and the number of groups comes out in the form of k-values (Daniel Berrar., 2018, Eakasit Pacharawongsakda et al., 2015) as shown in Fig. 17.

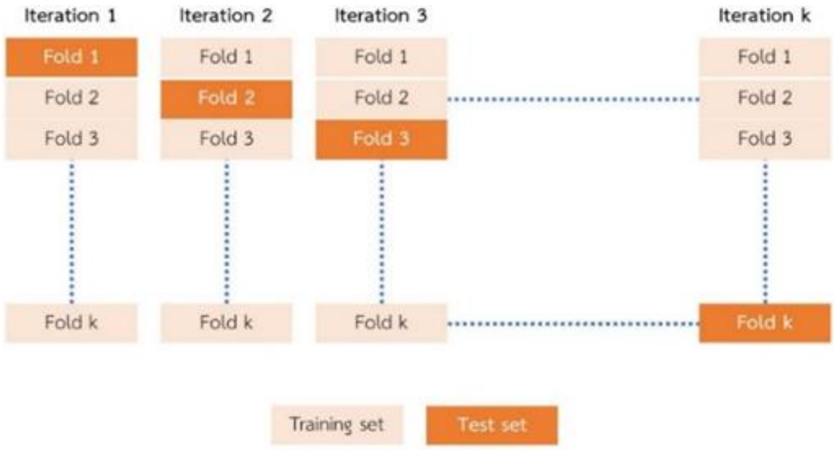


Fig. 17. K fold cross validation.

There are also variables that we will see in the process of displaying the results of testing the predictive performance of various models: Standard Deviation (Std.) is the standard deviation, which shows how much the value is different from the mean (DataStar., 2022), precision is the value for comparing the error from the labels measured by the amount of data entered the results show (Michael Kane., 2009), recall the model's ability indicators used to find labels within the dataset (Will Koehrsen., 2021), f1-score is a harmonic mean of precision and recall (Joos Korstanje., 2021), support variables representing the amount of data, accuracy used for various decisions (Editorial Staff., 2022), macro avg averages calculated metrics separately for accuracy. For each label (Ajitesh Kumar., 2022), the weight avg the label priority is calculated in a different dataset (Akhilesh Ganti., 2022).

2.5 Related Technology

Coral bleaching can occur anywhere on earth as corals are sensitive to environmental factors. Therefore, there must be a system to measure the quality of the environment in which corals live in order to find ways to protect them. It combines Internet of things (IoT) science with LoRa technology, sensors, and floating platforms to measure the quality of seawater inhabited by a wide variety of corals. In addition to the pontoon that we installed in the sea; data received from various satellites is also collected. To be used to compare the performance of the two systems, this topic will be divided into 2 topics as follows. Collecting data via pontoon needs to be a device that will be used to collect various data under the sea and send it to the network server. To use those data for further prediction by most devices that used as a device that can be found easily. To facilitate development and repair when equipment is damaged for all system components, there are 3 main parts as shown in Fig. 18.

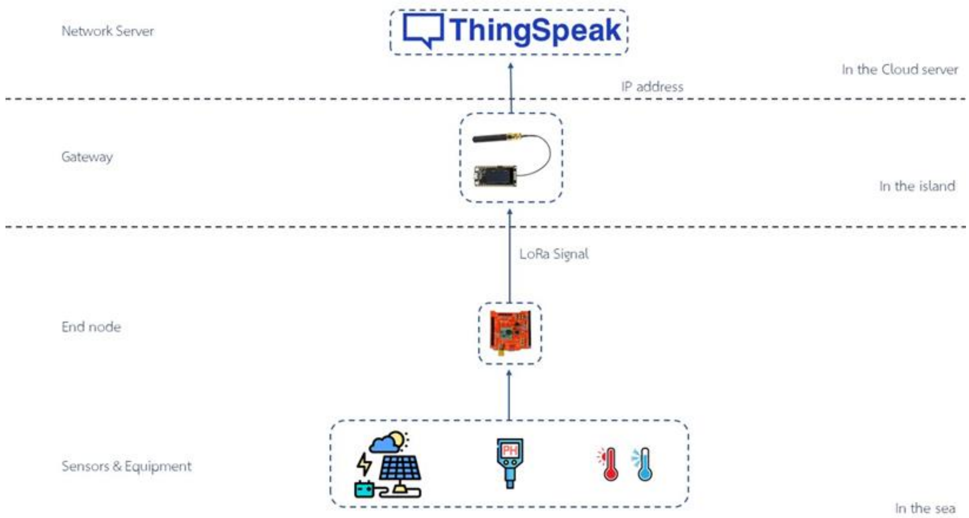


Fig. 18. An overview of the entire system.

From Fig. 18, we can see 3 main parts used in the data collection but this part we will split into 2 parts: Hardware and Software which both have the same aim, that is data collection from seawater in each factor (Seawater temperature, Seawater acidification). As seen in Fig.18, it has a LoRa device. This study applied LoRa technology to the transmission of data from sea for the prediction of the risk of coral bleaching. In 2 parts mentioned above will be shown below:

LoRa is a long-range and low-power communication system. It communicates wirelessly to the gateway connected to the Internet and acts as a forwarding message between these end devices and a central network server. It is one called Low-Power Wide Area Networks (LPWAN). LoRa can be applied in IoT applications, showing the efficiency of transmitting data over several distances, and in terms of power, it is a low-power technology compared to that. It is used to transmit data by using the LoRa technology as technology. It is essential to plan the settings and the features of the work to be used well before being used (Alexandru Lavric., 2019.) because LoRa uses a signal modulation technique and sends it to the gateway with a low frequency. It means the influence that affects the signal, whether the natural environment will interfere with or absorb the transmitted signal. In general, we can see the effect of canopy openness, Sky view factor, buildings, etc. All factors or influences affect the communication between the device or the device and the server (Riccardo Berto et al., 2021, Bilguunmaa Myagmardulam et al., 2021.) that can go wrong which effects are obtained by setting the three primary parameters or the influence of the signal received by the barriers. Moreover, LoRa is a technology worth exploring, brings many applications such as oceanography to attach a buoy for measuring its parameters. In this study, we will deploy LoRa on a sailboat to monitor a wide range of ship traffic near the coast (Lorenzo Parri et al., 2020, Ramon Sanchez-Iborra et al., 2018). LoRa signaling experiments are performed in the hilly terrain of the tall forest cover as signal transmission (Bilguunmaa Myagmardulam et al ., 2021.). Many of the above studies have confirmed the excellent performance of LoRa, whether it is transmitted over longer distances or lower power consumption than other technologies. Therefore, we can use LoRa to develop him with various tasks. We only need to be careful in setting LoRa and consider the influences that will affect the signal to make it work total efficiently (Alexandru Lavric ., 2019.). There are only three main parameters that affect the signal strength between the receiver and the transmitter namely: (Emanuele Goldoni et al., 2018, Nikola Jovalekic. et al., 2018, Riccardo Marini. et al., 2021.) code rate (CR) is set as in Eq. (2) (Emanuele Goldoni et al.), where FEC is to increase the number of bits used to transmit a signal. It reduced the amount of bit errors that can occur in the transmission. In addition, the CR can also be used to calculate the bit rate or the data rate, which is a different data transfer rate depending on the setting parameters of spreading factor (SF), bandwidth (BW), we can calculate bit rate (R_b) as in Eq. (3). From the above equation, we notice that the above equation is the equation shown to see the relationship between the three parameters, that is, setting parameters will affect the result of bit rate or data rate (Van Dai Pham et al., 2019.). Suppose we have data of variable bit rate or data rate. In that case, we can take the variable SF calculated mainly in LoRa technology withstands CR signals, broadcast time increased energy consumption decreased data rate (Riccardo Berto et al., 2021, Seungku Kim et al., 2020.) by calculating the SF from Eq. (4). Notice that SF refers to the number of chirps per symbol and ranges between 7 to 12. The SF is also a variable used to find the symbol rate (R_s) by calculating together with BW. The BW used in LoRa communication technology comprises three signal waves, namely 125, 250, and 500 kHz. In some either regions, or continents, or each depending on the regulations of each country, the BW is usually the same as chirp rate (R_c), transmission rate, and is used as a variable to find the time of symbol (T_s), as in the Eq. (5) (Emanuele Goldoni et al., 2018; Nikola Jovalekic et al., 2018; Tanatpong Udomchaipitak et al., 2022):

$$CR = \frac{4}{4+n}; \ \ n \in 1, 2, 3, 4, \qquad (2)$$

$$R_b = SF * CR * R_s, \qquad (3)$$

$$SF = log_2\left(\frac{R_c}{R_s}\right), \qquad (4)$$

$$T_s = \frac{2^{SF}}{BW}. \qquad (5)$$

Typically, LoRa transmission efficiency is measured at the receiving node in the form of an RSSI or Received Signal Strength Indicator. That is intimates the signal strength value or the ability to receive signals from one node to another (Emanuele Goldoni et al., 2018; LoRa Developer portal, 2021). We can calculate the RSSI value from Eq. (6). The equation variables definitions are A is the receiving power in dBm, d is the distance between the source node and the destination node, and n is the loss parameter. The loss parameter is a value that results from the barriers that come together with the transmission distance that greatly influence the signal quality impact that occurs in each area, in each work area, and will have different effects due to different obstacles or terrain (Ana Elisa Ferreira et al., 2020; Reza Firsandaya Malik et al., 2019):

$$RSSI = -(10n \log_{10} d - A). \qquad (6)$$

After studying the LoRa section in this study, we found that the parameter that affects the RSSI the most is the Spreading factor in signal transmission and signal resistance, i.e. The higher the SF setting, the greater the transmission distance, but at the same time, the noise tolerance decreases, which may damage the transmitted data or damage the data. can be lost It also results in less data transmission compared to the lower SF settings. After our study, LoRa technology investigated the properties of various sensor devices to collect marine data. In the transmission of data (hardware), we will divide it into 2 parts: sender and gateway, which are detailed below.

- Sender

The Pontoon used this time is a buoy built from 6 inches and 4 inches PVC pipe to install a variety of devices that will be used to collect marine values as shown in Fig. 19.

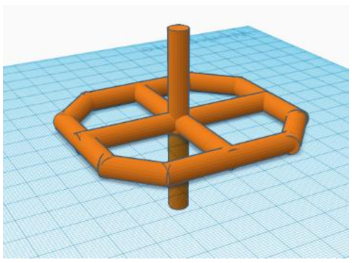


Fig. 19. Pontoon structure.

The pontoon features a large (6 inches) main tube in the center of the pontoon, a tube used to collect the internal equipment and a support tube (4 inches) that surrounds and supports the tube main. When combining the two will make the Pontoon work more balanced. When used with all devices, they are shown below.

Dragino LoRa shield for Arduino is a microcontroller board developed to transmit data over long distances. With low data rates High anti-interference while using low current the LoRa modulation technique is patented by Semtech, the SX1276 chip is capable of sensitivity greater than -148 dBm when combined with a power amplifier of +20 dBm, making it ideal for Will be used for jobs with long distances. This device operates on a frequency of 915 mHz, according to the office of the National Broadcasting and Telecommunications Commission (NBTC) defined regulations. Dragino LoRa shield for Arduino the Dragino LoRa shield can be powered by 3.3 or 5v for peripheral devices and can also be used for 915 mHz, 868 mHz, 433 mHz (factory configuration). Compatible with Arduino Leonardo, Uno, Mega, DUE devices and connect external antenna via I-Pex connector (Dragino.,2021) as shown in Fig. 20.

Fig. 20. Dargino LoRa Shield for arduino (https://wzper.my/).

Dragino LoRa Shield for Arduino is used to transmit data received from sensors located inside the buoy to the LoRa device onshore which is responsible for receiving the data and send it to the network server to collect the collected data. used to predict the future.

The DHT 22 Module is a digital sensor capable of measuring temperature and humidity with an accuracy of between ±0.5℃ or ±1%RH with an operating voltage of 3.5v to 5.5v in the range. The temperature can be detected in the range of -40 ℃ to 80 ℃ and the humidity can be detected in the range from 0%RH to 100%RH (BOGDAN Mihai. , 2016) as shown in Fig. 21.

Fig. 21. Temperature and humidity sensor DHT22 (https://www.jd.co.th/).

The reason why the DHT22 sensor was used to collect this research data is because the DHT22 is a sensor used to measure surface temperature around a buoy, and the DHT22 is also an available sensor. Easy installation and maintenance

pH sensor is a sensor capable of measuring the pH of liquids in the range from 0 to 14 pH. It features 5v operating electricity and can operate in the temperature range of 0 ℃ to 60 ℃ in response time duration ≤ 1 min and detection accuracy of ± 0.1pH at 25 ℃ (RFROBOT., 2020) as shown in Fig. 22.

Fig. 22. pH sensor (https://www.ubuy.td/).

The pH sensor used in conjunction with this buoy was due to the unstable pH factor of seawater in the algae range. Zooxanthellae want, it may result in the corals being cleared and eventually resulting in algae. Zooxanthellae Escape from coral tissues, eventually causing coral bleaching.

Waterproof temperature Sensor is a wired temperature sensor that is covered with a metal tip. Which uses a current in the range of 3.0 to 5.5v. The temperature supported by the sensor is in the range of -55 ℃ to 125 ℃. It is compatible with a variety of microcontroller boards such as Arduino, Micro bit and Raspberry Pi, etc. The sensor has a high value. Detection accuracy is in the range of ±0.5℃ from -55℃ to +125℃ (Analogread., 2021) as shown in Fig. 23.

Fig. 23. Waterproof temperature sensor (https://www.arduitronics.com/).

The reason why waterproof temperature sensors are introduced is because temperature factors are the most affecting coral reefs because rising sea temperatures may affect algae. Zooxanthellae escape from the mouth of the nest that is the original habitat Thus, we can clearly see the spine of the coral.

- Gateway

As for Gateway, it connects the end node within the buoy to the network server, which is responsible for transmitting and receiving data from the end node by receiving the LoRa signal and sending the data to the network server via IP transmission. Set in order to store data from the end node to the network server.

The TTGO LoRa32 microcontroller board is a development board based on the ESP32 with the LoRa chip (Chip SX1276) and displays the results via a 0.96-inch OLED screen. Both the chip and the screen on the ESP32 will work or communicate via GPIO. On the board where the board is used, the frequency is not set at the factory like the Dragino LoRa Shield for Arduino but when used in Thailand, it has been configured according to the requirements of the office of the National Broadcasting and Telecommunications Commission (NBTC). According to NBTC's specifications, it can be used to transmit data over LoRa signals. The operating voltage ranges from 1.8 to 3.7v, the transmit power is 20 dBm, with a frequency error of ±15 KHz. It works in the range of -40 °C to 85 °C, as shown in Fig. 24.

Fig. 24. TTGO LoRa32 (https://th.aliexpress.com/)

Since Board TTGO LoRa32 is based on ESP32, which has a part to connect to the internet via wifi, it is suitable for application as a gateway in this research. There is also an OLED display, which we can develop to show the status of various system operations through the screen. To facilitate the administration of the administrator (Arduitronics.,2021). The next section will be about the software we use for LoRa system development, LoRa data collection, and satellite data collection.

In software part, Arduino IDE is a program that we use for developing sensor devices that are built into a pontoon. Arduino IDE or Arduino Integrated Development Environment is a program used for developing various Arduino related devices. of the work of the library, which has various functions that make the work of developers faster (Arduino.,2022) And in this study, since we have adopted LoRa technology, it is necessary to extract a library named LoRa, which is a library used for data transmission in LoRa signals, which will support the chip. The development of Semtech (SX1276/77/78/79) installed on various LoRa devices (Arduino LoRa.,2022)

In the operation of a pontoon, there is also a connection to the network sever layer in order to collect the information obtained from the pontoon installed at sea. Which the network server part is for collecting data from the end node within pontoon and sending it to the gateway for the Network server to collect the data. In this research, we chose ThingSpeak as a network server to collect data because this is a software that facilitates data access. Data extraction data logging through this API is why we have chosen ThingSpeak as a network server to collect data before it can be processed further. It helps to collect data on the marine environment according to the various sensors installed. To wait for the end of the collection of such data to cleaning data again in the prediction section (ThingSpeak, 2021). After the development of data collection with various sensor devices. Next will be the software or web service that we use for collecting seawater parameters obtained from satellites. In part collecting data from satellites this is the extraction of information collected within the website over time and compiled into a dataset. It is used to predict the likelihood of bleaching in the study area. The websites used in this study is Google Earth Engine, which is a platform for analysis in scientific and visualization (Google Earth Engine., 2021). The parameters obtained from the Google earth engine only sea water temperature (SWT). To find the error of the values between the pontoon and satellite. This includes modeling the predictions of the two data sources that differentiate the accuracy between the two data sources as to which data source provides the best accuracy for this research.

Chapter 3

Materials and Methods

Data collection in our study area. We have divided the seawater data collection into two parts: collection of seawater data from sensors inside the buoy and seawater data collection from satellite loading. The reason for collecting data from two sources is to find the error of the two data to get the result in the form of equations from both data sources. In addition to obtaining data for the prediction of various models and finally for displaying it in the form of a map.

3.1 Study area

Installing a pontoon in the sea was a challenge for this study. Due to various natural and legal restrictions, we have therefore designed a grant suitable for attaching the existing Department of Marine and Coastal Resources (DMCR) pontoon in the study area. Used to collect various parameters at diving points in the Koh Taen area. Which is the island near Koh Samui as shown in Fig. 25.

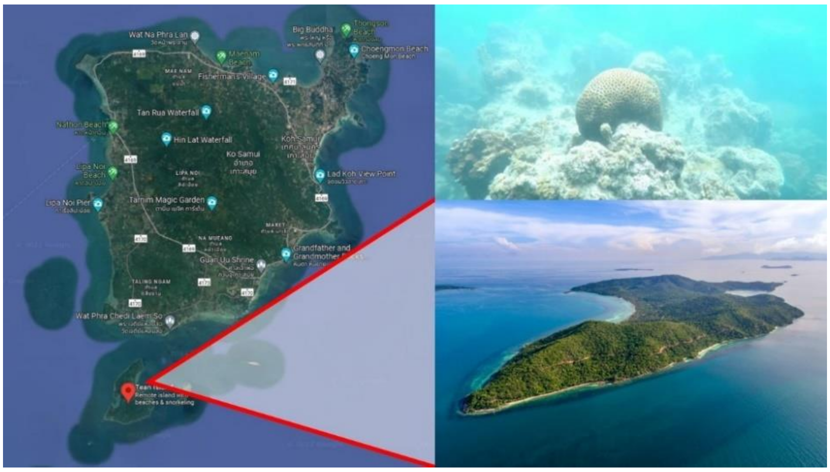


Fig. 25. Koh Tan Island.

Koh Taen is an island located in the southern part of Koh Samui. Which is an educational tourist attraction comprising two beaches: "Ao ok" is an educational site for shallow water coral reefs, that is our educational site, and "Ao Tok" is an educational site for mangrove forests. Traveling from Koh Samui by long-tail boat from the Thong Krut Bay area takes 30 minutes (tourismthailand, 2021).

3.2 Hardware Part

The installation of a pontoon in such areas is to attach to the existing pontoon, which is a pontoon used for tying boats. If tourists want to dive, the boat must be tied to such a pontoon only due to cannot leave the anchor, because the area is a coral reef which can cause damage. So, the development of our pontoon aims to install and prevent seawater damage to the sensor devices inside. For the sensor devices to work at their best, the internal devices installed in the two pontoons are of the same design. All devices are described in detail in the previous chapter. An overview of the sensor system inside the pontoon is shown in Fig. 26.

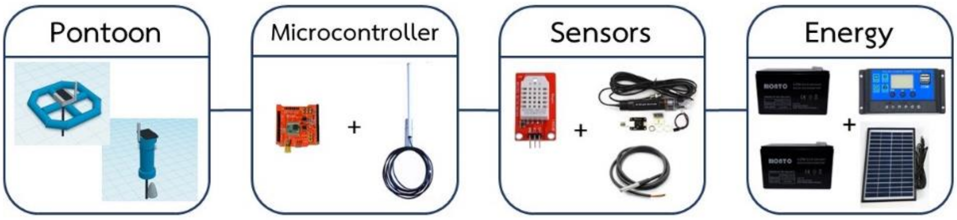


Fig. 26. Overview of the sensor system inside the pontoon.

For the operation of the various devices used for marine data collection in our study area. We are divided into four parts: pontoon, microcontroller, sensors, and energy where all three electronic devices are contained within a pontoon. Our pontoon developments are designed in two configurations to contain all devices: pontoon capsule and pontoon eight square or satellite shape. With both pontoons, we have brought it to the real area. In practical use, the capsule form is a difficult-to-install model because the installation requires materials that are heavier and more mass than water to help. While octagonal pontoon is an upgraded version of the capsule to make it easier to install and float. It can live in a monsoon sea condition, and it is the format that we use all the time and complete the data collection. The part of the microcontroller board and accessories is the part that we have tried and evaluated before going to the real area because transmission of information from the sea to the shores of the islands with the receiver station installed is very difficult to signal. Therefore, in this study, it was necessary to replace the antenna from the traditional antenna with a larger and more powerful antenna. For transmitting the signal back to the station in order to attach the sensors to the microcontroller board, we have installed 3 sensor devices, 2 of which are sensors for measuring seawater and 1 sensor that is installed inside the pontoon to check the air inside the pontoon, whether it facilitates the operation of the microcontroller board or not, and another function is the resemble of the black box in case of seawater leaks inside and no signal from the pontoon is sent. It can be checked how in the last transmission the weather inside the pontoon was. Finally, the energy part is an extremely important part of energy extraction. It is the part that produces the energy which is obtained from the solar cell and distributes the energy to feed it. System in pontoon for sending data back to shore. After the sensors have collected data from the seawater, the system sends the data back to shore via LoRa signals to receiver stations located at the shore. As for the function of the receiver station, it is responsible for receiving the signal from the pontoon and sending the data to the network server to collect all data as shown in Fig. 27.

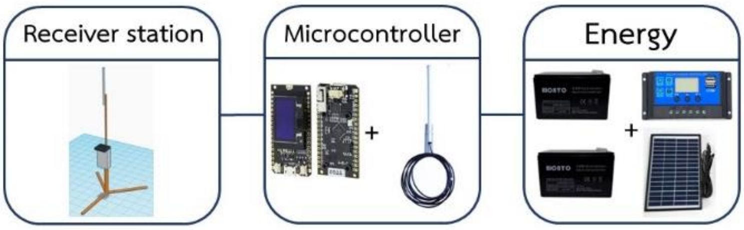


Fig. 27. Receiver station.

The operation of the receiver station consists of 3 parts: a 3-meter steel column with base, microcontroller, and energy. The 3-meter steel column is used for installing a waterproof box that contains various electronic devices and antenna, the reason why we have to use a pole with a height of 3 meters is to use for installing antennas that are used to transmit signals and also help reduce data noise because the area around the side is full of signal equipment that may cause noise to data such as trees, banners, etc. The next part will be part of the microcontroller in this study, we used the TTGO LoRa32 OLED board as a receiver node instead of the more expensive gateway. Its highlight is an OLED display that can be configured to display data received from a pontoon in real time or a point to check the status of the pontoon if the pontoon has leaked water. In addition to its long-distance transmission and display, the TTGO LoRa32 OLED can also connect to Wi-Fi signals from its built-in module. This means that it can send the information received from the pontoon to the network server itself, and of course in the case of long-distance transmission and there are many obstacles like this. We had to adjust its antennas to prevent signal loss while we were collecting data, and the last part was the most important part of the whole system, energy, which was the part that powers the entire system in the study. These properties as shown in Table 1.

Table 1. Features of all devices.

Device	Features
Dragino LoRa shield for Arduino	Compatible with 3.3v or 5v I/O Arduino Board Frequency Band: 915 MHZ/868 MHZ/433 MHZ Low power consumption Compatible with Arduino Leonardo, Uno, Mega External Antenna via I-Pex connector
TTGO LoRa32 OLED	Chip ESP32 Operating Temperature -40 to 85 degree celsius OLED screen display OLED 0.96 Inch Frequency Band: 915 MHZ/868 MHZ/433 MHZ
Antenna	Gain: 7 dBi Max, power(W): 100 Cable On Antenna: 3 Metes Weight (g.): 600 g
DHT22	Operating Voltage: 3.5V to 5.5V. Temperature Range: -40 to 80 degrees Celsius Humidity Range: 0% to 100% Accuracy: ±0.5°C and ±1%
Analog pH Sensor / Meter Pro Kit V2	Supply Voltage: 3.3~5.5V Measurement Accuracy: ±0.1@25 degree Celsius Cable Length: 500cm Detection Range: 0~14 Probe Life: 7*24hours >0.5 years (depending on the water quality)
Waterproof 1-Wire DS18B20 Compatible Digital temperature sensor	temperature range: -55 to 125 degrees Celsius ±0.5°C Accuracy from -10°C to +85°C Usable with 3.0V to 5.5V power/data

Battery	Dry battery 12V/8AH/20HR Weight 2.1 Kg. Charging voltage 13.5-13.8 Volt Continuous use at 0.4A for 20 hours
Solar cell charger controller	Control Charger, the device can choose to use the voltage according to the contractor, 12v 24v automatically. Withstands a maximum current of 10A (actual use should be sensitive to 20%)
	It can set the time to turn on and off the lamp or load the usage by yourself. Users can adjust their own suitable charging settings. There is an LCD display screen showing the status, easy to use. There are 2 USB ports for charging your phone.
Solar cell	Maximum Power output (Pm): 30W Maximum Operating Voltage (Vmp): 12V Maximum Power Current (Imp): 2.5A Open Circuit Voltage (Voc): 14.16V

In addition to the properties mentioned devices in the microcontroller of both parts also need to set the parameters of LoRa so that the system can transmit data over long distances without problems or with minimal signal transmission problems as shown in Table 2.

Table 2. Setting parameter.

Device	Parameter	Value
Dragino LoRa shield for Arduino & TTGO LoRa32 OLED	Spreading Factor (SF)	12
	Bandwidth (BW)	250
	Coderate (CR)	4/5

From the settings of both devices, it is set up to be used for long-distance transmission using the maximum performance supported by the device and compliant. After receiving the data from the devices in the pontoon, the receiver station will continue to send the data to the network server layer ThingSpeak through the API key installed during development. microcontroller device the highlight of ThingSpeak is accessibility and easy to understand and can also distribute the information that we have collected for others to access. After the development and testing of the pontoon system that we will use for marine data collection are complete. So, we have installed our pontoon in our study area by installing pontoon. Our pontoon is installed with the Department of Marine and Coastal Resources (DMCR) pontoon. The reason why we have to install it in this way is that the documentation process with government agencies takes a long time and there is a complexity between the organizations. It is therefore necessary to request assistance from the Department of Marine and Coastal Resources (DMCR) for the joint pontoon installation. The installation characteristics are shown as Fig. 28.

Fig. 28. Installation of pontoon.

3.3 Software Part

Collecting data via satellite, we did this through a web service using the Google earth engine, which collects geographic data captured by satellites for this study. It has a catalog of various geographic information for visitors to choose from. In our study, we selected a data catalog in the surface temperature category, which contains data obtained from satellite sensors that collect surface heat data on the earth, whether on land or sea surface. Within the surface temperature group, many different data sets are available, such as analysis of climate occurring on land or measurement of sea surface temperatures. For our study, we chose a dataset that surveyed and collected data on sea surface temperature data operated by major organizations, that is NOAA (NOAA: National Oceanic and Atmospheric Administration is an agency that enriches life through science. Their reach goes from the surface of the sun to the depths of the ocean floor as they work to keep the public informed of the changing environment around them (NOAA., 2021)). It is a survey and measurement of various data occurring in seawater with data available in the period from 1981-09-01 to 2022-09-04 (Earth engine data catalog). Compared with the data obtained from pontoons installed in the period 2022-03-05 to 2022-03-17, we used that data range to collect data from the Google earth engine using the same coordinates as the installation of the pontoons (9.3848, 99.9513), which can view results from Google earth engine processing, and can also be exported as a .CSV file for use in various tasks. especially the analysis and prediction of results from various models as shown in Fig. 29.

Fig. 29. Results from Google earth engine processing.

The collected data from the satellite in .CSV file format is ready for processing in the next step, but before we can use both parts of the data to make predictions, we should clean the data first to minimize any errors follow to the part of data preparation. This may occur in the process of prediction and follow to data analytics life cycle.

3.4 Big Data Analytics

After we have collected the data from both sources, we clean the data received from both sources, especially the data received from the pontoon because during the storage period there may be high waves which may cause internal datasets to contain noise-affected data. In our data cleansing as part of the data preparation process, our data cleansing process removes any data or rows containing noise from our dataset, which may make performance tests erroneous or impossible to predict. But before we do that test. We must first find the relationship between the two variables (Temperature of seawater, acidity of seawater) within the pontoon dataset to come to a conclusion about the relationship between the two variables. After that, the effectiveness of the three models (Naïve Bayes, SVM, and Logistic regression) was tested to determine the most effective model for predicting coral bleaching in the Koh Samui area. Suratthani Province Two of the above techniques (Split test, K fold cross validation) were used, in particular the K-value. In the K-fold cross-validation test, we tested this technique using K-values from 2-10 to find the K value that gives the best accuracy.

Chapter 4

Result and Discussion

4.1 Data Prediction

After collecting data on various marine factors that could result in coral bleaching, we cleaned the data obtained because the LoRa signal transmission was noisy, so before we could use the data it was necessary to go clean data, which is in the process of data preparation before it can be used with various models. After the clean data is complete, we have visualization Initially, which came out in the form of a graph of the two factors that were examined as shown in Fig. 30.

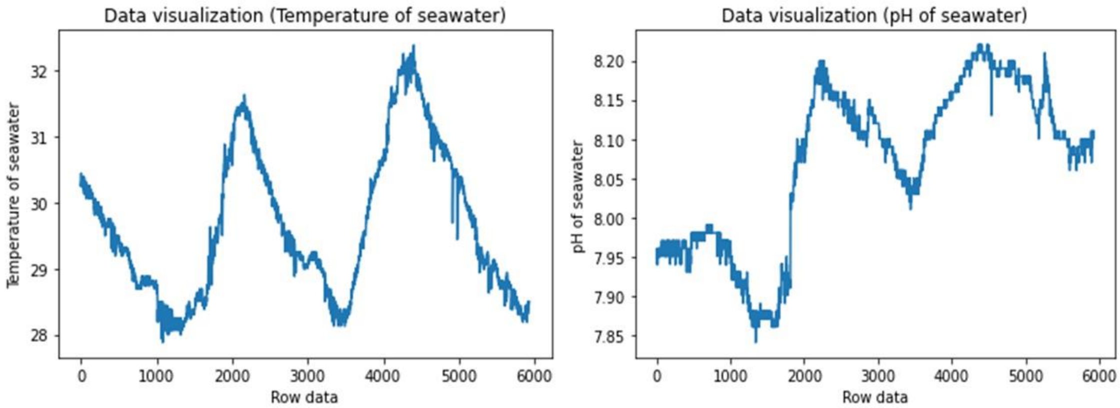


Fig. 30. Data visualization of both factors.

At the beginning of our prophecy, we set up labels to use for that case division of 1 and 2, where label 1 denotes unfavorable seawater factors that may affect corals that may cause coral bleaching., and label 2 means the water factor is in normal conditions and can dive. In order to divide our labels, we have to put all our criteria into 2 criteria: first If the sea temperature is above or equal to 30 degrees Celsius will be categorized in label 1, if not on label 2, and second, if the pH of seawater that falls below 8 is classified as label 1, but if not, it is classified as label 2, and during that time if one of the events in the data meets one of our criteria, it is also labeled 1. For example, at 12:35 the temperature of seawater is 30.3, and seawater pH is 8 = label 1. Conversely, at 12:35 the temperature of the seawater is 28.5, and the seawater pH is 7.7 = label 1, etc., as shown in Fig. 31.

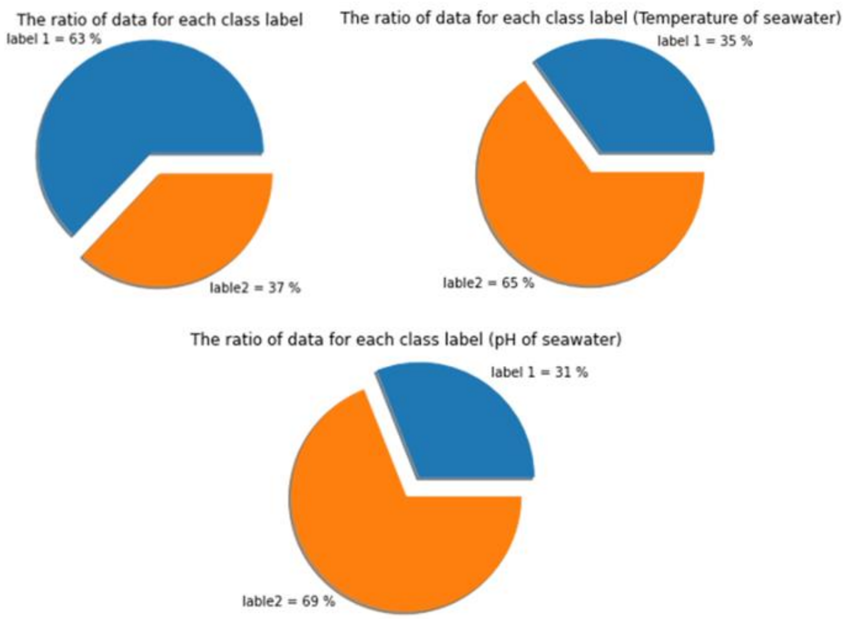


Fig. 31. label chart in each factor and summary.

The ratios we show are all ratios that we have defined in the data preparation stage, with a total of 3 groups: Class label overall of all data, Class label of the temperature of seawater, and Class label of pH of seawater observed. That the label summary graph (top left) has a number of label 1 different from the rest of the graph in accordance with the above example, i.e., if a variable exceeds a certain value, it will be placed in label 1 in label summary. After label our data we found the correlation between the two variables (seawater temperature and pH of seawater) And after processing the correlation of both variables by Pearson's correlation method, 0.683 was obtained. The two outputs are in the same direction and since their values are directed towards 1 and greater than 0.5, we can conclude that the two variables are related as observed on the graph as shown in Fig. 32.

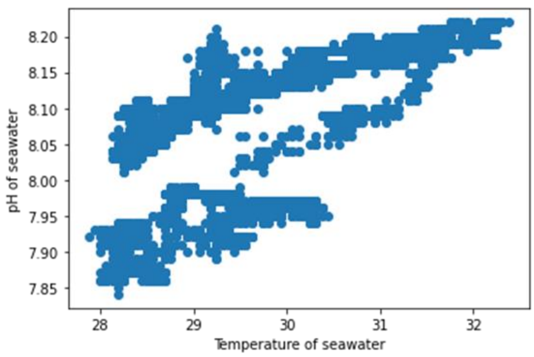


Fig. 32. Chart of Pearson correlation result.

The reason why we chose to correlate the two variables by Pearson's Correlation method because the evaluation of the visualization data above (Fig. 30) showing the increase and decrease of the two values leads us to be curious to what extent the two variables are related. The results showed us that the two variables were clearly positively correlated and that they were strong. After going through the clean data process and correlating the two variables from Pearson's Correlation, it was time to test the predictive performance of each model with both testing techniques. The results are shown as Fig. 33.

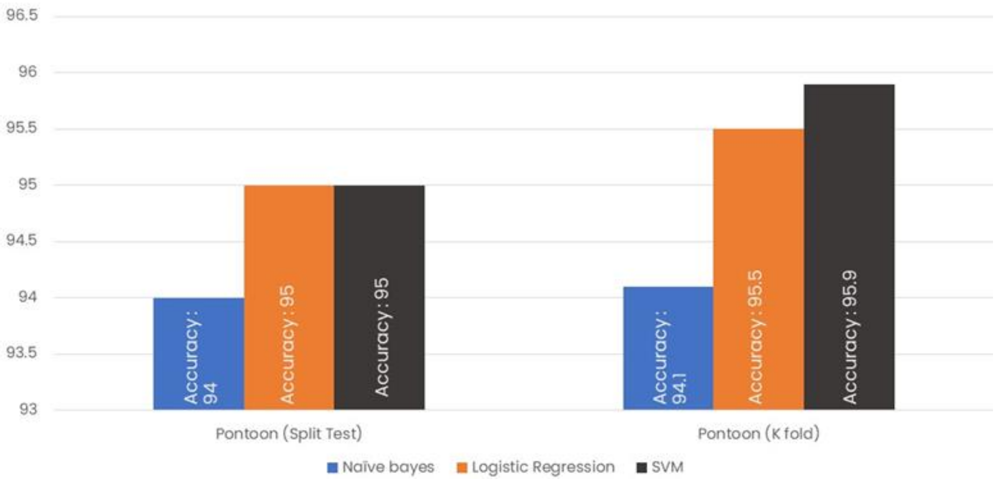


Fig. 33. Accuracy in each model (Pontoon dataset).

Based on the results we have shown above, we tested the performance of all three models using two different testing techniques: Split test, and K fold cross validation. The results were obtained derived from the split test technique, there are two different models with the same accuracy, SVM and Logistic regression were slightly better, leading us to conclude that to test the performance of the model using the split test technique, the SVM and Logistic regression model has the best value for testing in this technique. Next, we tested all models using the K-fold cross-validation technique. In our test this time we tested values of K from 2 to 10 to determine the K value. Can test and get the best accuracy and then we will find the average value in each model as a result of this test, we can see that the SVM model has slightly higher accuracy than other models. It is therefore concluded that in testing the model's performance using the K fold cross-validation technique, the SVM model had the best value.

After collecting sea surface temperature data from the data catalog of NOAA obtained from the google earth engine, we found that collecting data from satellites is a limitation to certain data collection. Since the satellite data collection has a limitation in the orbit of the satellite that needs to orbit the earth all the time, the amount of data obtained from this data collection is much lower than the amount of data obtained from pontoons as shown in Fig. 34.

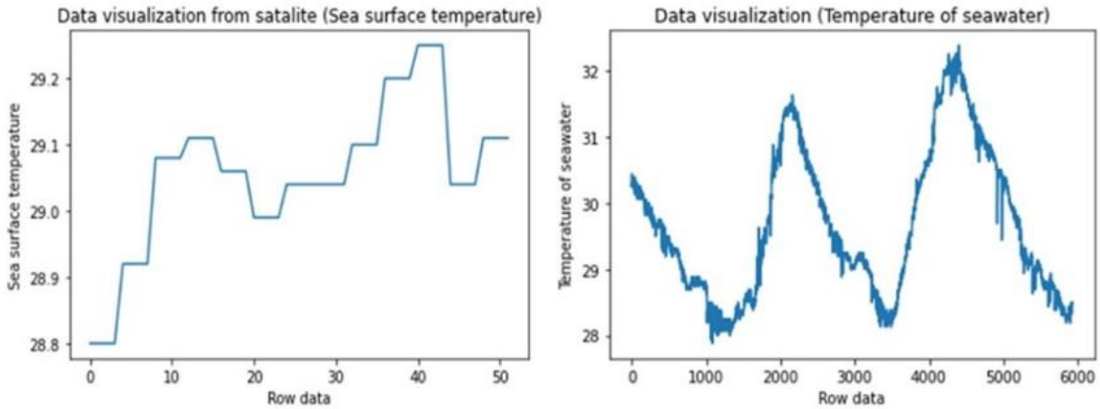


Fig. 34. Both datasets when compared.

Comparing the two data, the satellite-derived data represents only 0.86% of the data we collect from the pontoon, and the next drawback is that the satellite-derived values do not contain any range of seawater temperature above or equal to 30 degrees Celsius are a stressful range for corals, which can have a big impact. We, therefore, came up with the idea of using spatial analysis tools to solve this problem. By using ArcMap 10.3 program for this operation. For this ArcMap implementation, we have two tasks to perform: segmentation for data label classification and visualization in the form of a map. The reason why we had to segment in ArcMap because the satellite data was insufficient when labeling the data, making predictions impossible if the dataset's label only had one label. We have therefore worked on the segmentation of such data using Quantile techniques segmentation is a technique where each layer contains the same amount of data, where the number of data greater than and less than the median is equal as shown in Fig. 35.

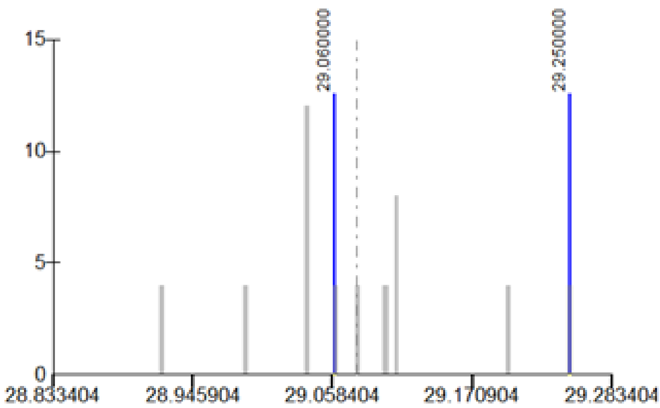


Fig. 35. Result from using Quantile techniques.

The above results showed that the data segments were divided into two intervals after applying the Quantile technique to segment the data, thus knowing the criteria for our label classification. From collected data from satellites as the results show, label 1 is in the range of 29.07 to 29.25, and label 2 is in the range of 28.8 to 29.06. Make the dataset available for re-testing the model's performance. In our test we used the same model and testing techniques as applied to the pontoon-derived dataset, the only difference being that in the satellite dataset there was no pH of seawater factors or variables, because the Advanced Very High-Resolution Radiometer (AVHRR) measurement of factors or variables in the Earth's atmosphere is not yet qualified to monitor seawater pH, so testing the predictive performance of each algorithm requires only one factor. After testing the predictive efficiency of the three models, the results can be shown in the test shown as Fig. 36.

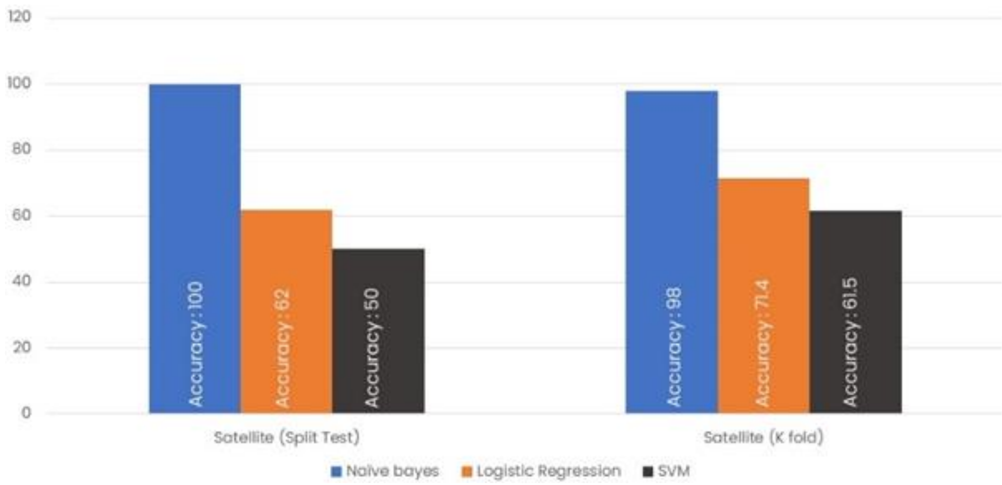


Fig. 36. Accuracy in each model (Satellite dataset).

The results of testing for accuracy and other values from the two testing techniques using the values collected from satellites show a huge difference in the decrease in accuracy and increase. As demonstrated by the figures in this part, we expect that the relatively small amount of data may cause the case studies in the training segment to be less or similar to the test datasets. So much so that the results in the Naïve Bayes model are so accurate that it's hard to believe in some parts. In a split test scenario with a Naïve Bayes model in this low dataset, it is not suitable for prediction. Therefore, we compared the two remaining models, concluding that the results from the test in such low-data cases logistic regression were more efficient and reliable than the other models. Performance testing from the K fold cross-validation technique is performed by dividing data into multiple subgroups based on the given K values and then we average the result accuracy in each model same as the pontoon dataset. Which has the highest accuracy of the Naïve Bayes model compared to the other two. Therefore, in this testing step, we conclude that the model performance test by the K fold cross-validation technique has the best value of the Naïve Bayes model.

- Equation

After we tested the predictive performance of various models and made visualization from both data already. We have studied the creation of equations for calculating the temperature of seawater variables from another regression. The results are shown as Eq. (7) - (8):

$$\mathrm{T_p} = 0.9947\mathrm{T_s} \quad (7),$$

$$\mathrm{T_s} = 1.0048\mathrm{T_p} \quad (8)$$

where $\mathrm{T_p}$ is result of sea water temperature value from pontoon,

$\mathrm{T_s}$ is result of sea water temperature value from satellite.

From the results shown in the form of both equations, the equations analyzed from linear regression with an R-square of 0.995 of both equations were tested error. The results calculated from the two equations compared to the values obtained from the actual data collection showed an average error of 1.98% for both equations by calculating from Eq. (9)

$$\text{error}(\%) = \frac{|\text{R}-\text{E}|}{\text{E}} \times 100 \quad (9).$$

where R is result value from the data in the real experiment,

E is calculated result value of the equation.

It is an equation for calculating percentage error from calculating through experimental results and calculated results. For calculating the percentage error between the results obtained from the two data sources. We conducted a virtual experiment by using the above equation to calculate seawater temperature from equation 8 using the variables obtained from the pontoon to obtain the results in the form of satellite data in order to obtain a new data set for testing. Test model performance again, but this time we combine the new data with the old data using the old data as a training dataset and the new data as the data for testing the efficacy of both models using a test technique. Two techniques as before shown as Fig. 37.

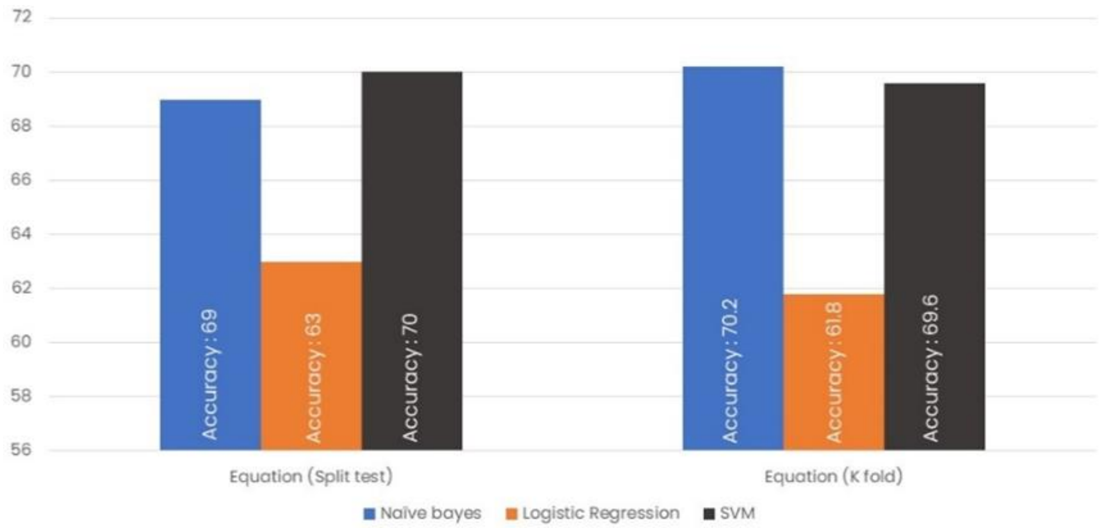


Fig. 37. Accuracy in each model (Equation dataset).

From the results obtained from the data-computed predictive performance tests of Equation 8, in the split test, it can be seen that the accuracy is the highest among all model comparisons. In conclusion, the model with the highest accuracy is the SVM model, and the next section is the K-fold cross-validation technique. From testing the model when comparing the average accuracy of all models, Naïve Bayes is the highest among all model performance tests using the K fold cross-validation technique. Based on the results of our model prediction performance test. It can be seen that our predictive performance has decreased accuracy compared to pontoon accuracy, which is presumed to be due to insufficient training data for model training. The accuracy value decreases when applied to the model in this virtual experiment. By comparing the overall accuracy of all the model performance tests of the two techniques in both datasets(pontoon and satellite), we can come up with a model suitable for predicting bleaching coral in the Koh Samui area of Suratthani Province. In this comparison, in addition to the accuracy of the three models, we also used the RMSE value or Root Mean Square Error, which is used to measure the model error in quantitative data prediction (James Moody., 2019) shown as Tabel 3.

Tabel 3. Accuracy and RMSE in each model (Pontoon & Satellite).

	Spilt test				**K fold cross validation**			
Model	Pontoon		Satellite		Pontoon		Satellite	
	Accuracy	RMSE	Accuracy	RMSE	Accuracy	RMSE	Accuracy	RMSE
SVM	95	0.213	50	0.707	95.9	0.979	61.5	0.784
Naïve bayes	94	0.251	100	0	94.1	0.970	98	0.980
Logistic regression	95	0.227	62	0.612	95.5	0.977	71.4	0.784

Based on the efficacy evaluation of the models we conducted in predicting coral bleaching in the Koh Samui area. Surat Thani Province we found that the most efficient model out of the three was the SVM model,This is because the RMSE value obtained on the split side of the split test is lower than that of the other models, and although in the K fold test, the SVM has a higher RMSE than the other models, but in comparison it was found in The RMSE values of the other models were similar and when tested in the K fold, the RMSE was higher than the split test when compared to the RMSE evaluated in both datasets. For compared between datasets derived from pontoon and satellite will find that datasets that optimize all mock tests are pontoon-derived datasets. After that, we have done a visualization of the results obtained from this prediction through the ArcMap program as shown in Fig. 38

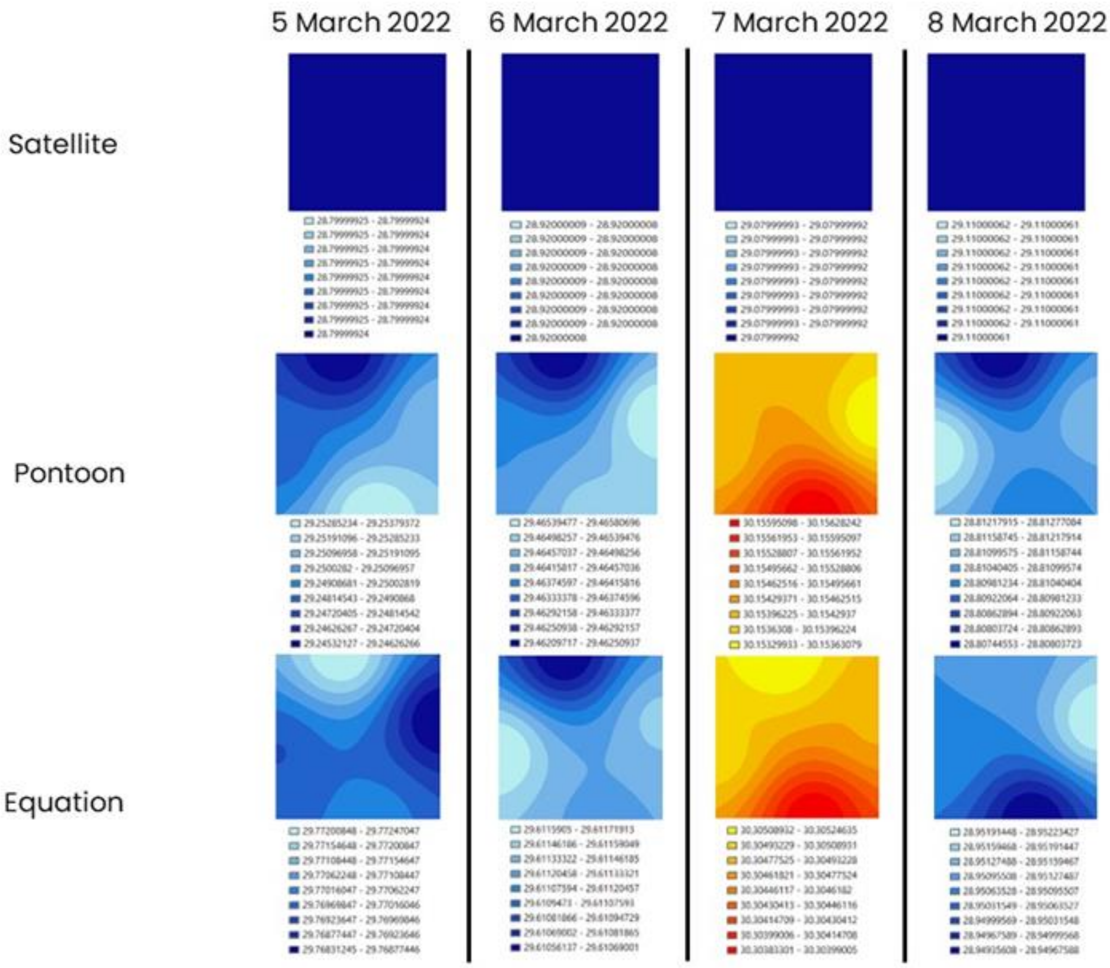


Fig. 38. Data visualization in each example date.

The visualization results of ArcMap's Interpolation (IDW) technique show that the resolution of the results expressed in terms of the results obtained from the satellite dataset is less than that of the pontoon dataset and equation in at the same time when observing the change in sea temperature, it can be seen that During the 8th March, corals are subject to bleaching damage. This is because on Day 7 March the water temperature was relatively warm (>30C), whereas on Day 8 March experienced a sharp drop in temperature just across the day (<28C). The experiment of Melissa S. Roth et al. is consistent with their experiment that sudden water cooling over a short period of time can harm corals.

Chapter 5

Conclusion and Future Work

For this study of predicting coral bleaching by applying Internet of Things technology (IoT). The goal is to study pontoon monitoring devices that are used to store various parameter data with new technology involved in the work. Testing the performance of various models suitable for coral bleaching forecasting and calculating equations derived from satellite and sensor data. In order to keep the beautiful resources under the sea for humans to study and research this wonder forever.

In this experiment, we studied the construction of a pontoon monitoring system by applying LoRa technology to create a wireless connection to the Internet for collecting marine data to predict coral bleaching in Koh Taen, an island in the south of Koh Samui. And there is a diving point, which is a tourist attraction that can attract tourists from many nations to see the beauty under the sea. A total of three models were used to test the predictive performance in this study. In our study, data from two sources, the pontoon dataset and the NOAA dataset of the google earth engine, were used in our study, and the results after model performance testing in each dataset were obtained with a model suitable for predicting coral bleaching. In our study area the SVM model, the dataset with the most accurate and reliable predictive results was the pontoon dataset, as the system has a more granular collection of parameters compared to the NOAA dataset of the google earth engine. Visualization comes out in the form of a map and classification (satellite data) with the ArcMap program and finally an equation derived from the comparison of data from two sources with an R-square of 0.995 and a percentage error of 1.98%. For this study, we have encountered a number of obstacles, most often due to unforeseen circumstances, during data collection from seawater using a pontoon, as our pontoon creations take a long time to make and test. Until our data collection began during the storm of Koh Samui, there were many obstacles and accidents with our pontoon, such as severe sea swells that caused the pontoon's rope to tie the boat of the Department of Marine and Coastal Resources (DMCR) while our pontoon is tethered. That is an event make has severely damaged the pontoon system, etc.

Finally, future work may be integrated into the part of bringing the material studied in this time together with tourism of Koh Samui and nearby islands as a dive point recommendation system. whose objectives are: The first is to recommend tourists to dive in beautiful spots. The second place is indirect coral restoration. In the event that the parameters of the seawater are not suitable, the coral will be stressed that we cannot know how much more impact if tourists enter that location. And finally, the application of image processing technology water into the system to increase the efficiency of work.

BIBLIOGRAPHY

S. Hafoud, K. Boutoial, A. Oussama, and F. Kzaiber. (2020). IMPACT OF CLIMATE CHANGE ON THE SEASURFACE TEMPERATURE OF DAKHLA BAY-MOROCCO. International Journal of Advanced Research in Engineering and Technology (IJARET), 11(11), 1039-1052, DOI: 10.34218/IJARET.11.11.2020.096.

NOAA. What is coral bleaching? Available online: https://oceanservice.noaa.gov/facts/ coral_bleach.html (accessed on 2 June 2021).

Katharina E. Fabricius, Craig Neill, Erik Van Ooijen, Joy N. Smith, and Bronte Tilbrook. (2020). Progressive seawater acidification on the Great Barrier Reef continental shelf. Scientific Reports, 10, 1-15, https://doi.org/10.1038/s41598-020-75293-1.

NOAA. Ocean acidification Available online: https://www.noaa.gov/education/resource-collections/ocean-coasts/ocean-acidification (accessed on 9 June 2021).

Jeremiah G. Plass-Johnson, Ulisse Cardini, Nanne van Hoytema, Elisa Bayraktarov, Ingo Burghardt, Malik S. Naumann, and Christian Wild. (2015). Coral Bleaching (Chapter 9) C- and N-fixation of coral reefs in response to environmental factors (CANCOR), 1, 118-135.

Melissa S. Roth, Ralf Goericke and Dimitri D. Deheyn. (2012). Cold induces acute stress but heat is ultimately more deleterious for the reef-building coral Acropora yongei. Scientific Reports, 2, 2-5.

Guobao Xu, Yanjun Shi, Xueyan Sun and Weiming Shen. (2019). Internet of Things in Marine Environment Monitoring. MDPI Sensors, 19(7), 1-10, https://doi.org/10.3390/s19071711.

Scott Wooldridge and Terry Done. (2004). Learning to predict large-scale coral bleaching from past events: A Bayesian approach using remotely sensed data, in-situ data, and environmental proxies. Coral Reefs, 23, 96–108.

BIBLIOGRAPHY (CONT.)

Harsh H. Patel and Purvi Prajapati. (2018). Study and Analysis of Decision Tree Based Classification Algorithms. International Journal of Computer Sciences and Engineering, 6, 2018, 74-78.

Harsh H. Patel and Purvi Prajapati. (2018). Study and Analysis of Decision Tree Based Classification Algorithms. International Journal of Computer Sciences and Engineering, 6, 2018, 74-78.

Department of Marine and Coastal Resources (DMCR). ปะการังฟอกขาว Available online: https://km.dmcr.go.th/c_254/d_17309 (accessed on 6 December 2017).

Marine Knowledge Hub. ปรากฏการณ์ปะการังฟอกขาวในทะเลไทย Available online: http://www.mkh.in.th/index.php?option=com_content&view=article&id=269&Itemid=236 &lang=th (accessed on 6 December 2018).

Joost W. van Dam, Andrew P. Negri, Sven Uthicke, and Jochen F. Mueller. (2011). Chemical Pollution on Coral Reefs: Exposure and Ecological Effects. Ecological Impacts of Toxic Chemicals, Ecological Impacts of Toxic Chemicals, 187-211.

Suchai Worachananan, R.W. (Bill) Carter, Marc Hockings, and Pasinee Reopanichkul. (2008). Managing the Impacts of SCUBA Divers on Thailand's Coral Reefs, Journal of Sustainable Tourism, 16(6), 645-663, DOI: 10.1080/09669580802159677.

QGIS. Spatial Analysis (Interpolation) Available online: https://docs.qgis.org/2.18/en/docs (accessed on 20 December 2021).

Geoff Boeing. (2021). Spatial information and the legibility of urban form: Big data in urban morphology, International Journal of Information Management, 56, 1-9.

BIBLIOGRAPHY (CONT.)

Mary K. Donovan, Thomas C. Adam, Andrew A. Shantz, Kelly E. Speare, Katrina S. Munsterman, Mallory M. Rice, Russell J. Schmitt, Sally J. Holbrook, and Deron E. Burkepile. (2020). Nitrogen pollution interacts with heat stress to increase coral bleaching across the seascape, BIOLOGICAL SCIENCES, 117(10), 5351-5357.

Alejandra and Virgen Urcelay. (2021). Assessing the extent of global mass coral bleaching with an updated database, 1-85.

Alexander Pletzer and Wolfgang Hayek. (2018). Mimetic Interpolation of Vector Fields on Arakawa C/D Grids, Monthly Weather Review, 147(1), 3-16.

ArcGIS. How inverse distance weighted interpolation works Available online: https://pro.arcgis.com/en/pro-app (accessed on 26 December 2021).

M. Ateweberhan and Tim R. McClanahan. (2010). Relationship between historical sea- surface temperature variability and climate change-induced coral mortality in the western Indian Ocean. Marine Pollution Bulletin 60 ,2010, 964–970.

Scott Wooldridge and Terry Done. (2004). Learning to predict large-scale coral bleaching from past events: A Bayesian approach using remotely sensed data, in-situ data, and environmental proxies. Coral Reefs, 23, 96–108.

Harsh H. Patel and Purvi Prajapati. (2018). Study and Analysis of Decision Tree Based Classification Algorithms. International Journal of Computer Sciences and Engineering 6, 2018, 74-78.

Krisanadej Jaroensutasineea, Sirirat Somchueaa, and Mullica Jaroensutasineea. (2020). Coral and reef fish community recovery following the 2010 extreme ocean warming event (mass bleaching event) at Thailand. Research article,1-8.

Sridevi Bonthu and Hima Bindu. (2018). Review of Leading Data Analytics Tools, International Journal of Engineering & Technology, 7(3), 10-15.

BIBLIOGRAPHY (CONT.)

Jirapond Muangprathub. (2018). Practical Big data analysis, Chapter 2 Data analysis life cycle.

Philip M. Sedgwick. (2012). Pearson's correlation coefficient, BMJ Clinical Research 345, 4483-4483

Geoffrey I Webb. (2017). Naïve Bayes, Encyclopedia of Machine Learning and Data Mining, 1-4.

Nathaphon Boonnam, Tanatpong Udomchaipitak, Supattra Puttinaovarat, Thanapong Chaichana, Veera Boonjing and Jirapond Muangprathub. (2022). Coral Reef Bleaching under Climate Change: Prediction Modeling and Machine Learning, Sustainability, 14(10), 1-13.

Pavan Vadapalli. Naive Bayes Explained: Function, Advantages & Disadvantages, Applications in 2022 Available online: https://www.upgrad.com/blog/naive-bayes-explained/ (1 September 2022)

Mayank Arya Chandra and S. S. Bedi. (2018). Survey on SVM and their application in image classification, International Journal of Information Technology, 13(2), 1-11.

Mayank Arya Chandra and S. S. Bedi. (2018). Survey on SVM and their application in image classification, International Journal of Information Technology, 13(2), 1-11.

Vinod Kumar Chauhan, Kalpana Dahiya, Anuj Sharma. (2019). Problem formulations and solvers in linear SVM: a review, Artificial Intelligence Review, 52(2), 803-855.

Dhiraj K. Top 4 advantages and disadvantages of Support Vector Machine or SVM Available online: https://dhirajkumarblog.medium.com/top-4-advantages-and-disadvantages-of-support-vector-machine-or-svm-a3c06a2b107 (accessed on 10 January 2022).

Michael P. LaValley. (2008). Logistic Regression, Circulation, 117(8), 2395–2399.

Saishruthi Swaminathan. Logistic Regression — Detailed Overview Available online: https://towardsdatascience.com/logistic-regression-detailed-overview-46c4da4303bc (accessed on 10 January 2022).

BIBLIOGRAPHY (CONT.)

AmiyaRanjanRout. Advantages and Disadvantages of Logistic Regression Available online: https://www.geeksforgeeks.org/advantages-and-disadvantages-of-logistic-regression/ (1 September 2022)

Ramesh Medar, Vijay Rajpurohit, B. Rashmi. (2017). Impact of Training and Testing Data Splits on Accuracy of Time Series Forecasting in Machine Learning, 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), 1-6.

Daniel Berrar. (2018). Cross-Validation, Reference Module in Life Sciences, 1-8

Eakasit Pacharawongsakda. การแบ่งข้อมูลเพื่อนำมาทดสอบประสิทธิภาพของโมเดล Available online: https://th.linkedin.com/pulse (accessed on 10 January 2022).

DataStar. How to Interpret Standard Deviation and Standard Error in Survey Research Available online: https://www.greenbook.org/marketing-research/how-to-interpret-standard-deviation-and-standard-error-in-survey-research-03377?fbclid=IwAR3qsMuCEmfPo-_dPEwGpvx-Uggz6vTDxr36JVO9bA1htvYAIycXAHTRvf4 (accessed on 20 January 2022).

Michael Kane. (2009). The Precision of Measurements, Applied Measurement in Education, 9, 355-379.

Will Koehrsen. When Accuracy Isn’t Enough, Use Precision and Recall to Evaluate Your Classification Model Available online: https://builtin.com/data-science/precision-and-recall (accessed on 20 January 2022).

Editorial Staff. What is Data Accuracy? (How to improve data entry accuracy) Available online: https://www.erp-information.com/dataaccuracy.html#:~:text=Data%20accuracy%20 means%20data%20without,is%20to%20iit%20true%20value.%E2%80%9D (accessed on 20 January 2022).

BIBLIOGRAPHY (CONT.)

Ajitesh Kumar. Micro-average & Macro-average Scoring Metrics – Python Available online: https://vitalflux.com/micro-average-macro-average-scoring-metrics-multi-class-classification-python/ (accessed on 20 January 2022).

Akhilesh Ganti. What Is Weighted Average? Available online: https://www.investopedia.com/terms (accessed on 20 January 2022).

Alexandru Lavric. (2019). LoRa (Long-Range) High-Density Sensors for Internet of Things, Journal of Sensors, 2019(4), 1-9.

Riccardo BertoPaolo, Paolo Napoletano, Marco Savi. (2021). A LoRa-Based Mesh Network for Peer-to-Peer Long-Range Communication, MDPI (sensor), 21(13), 1-12.

Bilguunmaa Myagmardulam, Ryu Miura, Fumie Ono, Toshinori Kagawa, Lin Shan, Tadachika Nakayama, Fumihide Kojima, and Baasandash Choijil. (2021). Performance Evaluation of LoRa 920 MHz Frequency Band in a Hilly Forested Area, MDPI (electronics), 10, 1-12.

Giacomo Peruzzi, Alessandro Pozzebon, Stefano Parrino, Stefano Parrino, Lorenzo Parri. (2020). Offshore LoRaWAN Networking: Transmission Performances Analysis Under Different Environmental Conditions, IEEE Transactions on Instrumentation and Measurement, 70, 1-10.

Ramon Sanchez-Iborra, Ignacio G. Liaño, Christian Simoes, Elena Couñago, and Antonio F. Skarmeta. (2018). Tracking and Monitoring System Based on LoRa Technology for Lightweight Boats, MDPI (electronics), 8(15), 1-18.

Tanatpong Udomchaipitak, Nathaphon Boonnam, and Supattra Puttinaovarat. (2022). An experimental study of RSSI for LoRa technology in Different Bandwidths, 2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), 728-731

BIBLIOGRAPHY (CONT.)

Emanuele Goldoni, Luca Prando, Anna Vizziello, Pietro Savazzi, and Paolo Gamba. (2018). Experimental data set analysis of RSSI-based indoor and outdoor localization in LoRa networks, Internet Technology Letters, 2(1), 1-6.

Nikola Jovalekic, Vujo Drndarevic, Ermanno Pietrosemoli, Iain Darby, and Marco Zennaro. (2018). Experimental Study of LoRa Transmission over Seawater, MDPI (sensor), 18, 2-23.

Van Dai Pham, Truong Duy Dinh, and Ruslan Kirichek. (2019). Method for Organizing Mesh Topology based on LoRa Technology, 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), pp. 1-6.

Seungku Kim, Heonkook Lee, and Sungho Jeon. (2020). An Adaptive Spreading Factor Selection Scheme for a Single Channel LoRa Modem, MDPI (sensor), 20(4), 1-13.

LoRa Developer portal. Documentation Available online: https://lora-developers.semtech.com/documentation (accessed on 30 September 2021).

Ana Elisa Ferreira, Fernando M. Ortiz, Luís Henrique M. K. Costa, Brandon Foubert, Ibrahim Amadou, and Nathalie Mitton. (2020). A study of the LoRa signal propagation in forest, urban, and suburban environments, Annals of Telecommunications, 75, 333-351.

Reza Firsandaya Malik, Andhika Rizky Perdana, Rido Zulfahmi and Huda Ubaya. (2019). Dual-Channel LoRa Gateway using Channel Assignment on Raspberry Pi, Journal of Physics, 1196, 1-10.

Dragino. Arduino Shield featuring LoRa technology Available online: https://www.dragino.com/products/lora/item/102-lora-shield.html (accessed on 6 October 2021).

Mihai Bogdan. (2016). How to Use the DHT22 Sensor for Measuring Temperature and Humidity with the Arduino Board, sciendo, 68, 22-25, https://doi.org/10.1515/aucts-2016-0005.

BIBLIOGRAPHY (CONT.)

RFROBOT. Gravity: Analog pH Sensor / Meter Pro Kit V2 Available online: https://www.dfrobot.com/product-2069.html (accessed on 6 September 2021).

Analogread. Temperature sensor DS18b20 Waterproof Available online: https://www.analogread.com/product/1355/temperature-sensor-ds18b20-waterproof (accessed on 6 September 2021)

Arduitronics. TTGO SX1276 LoRa32 915MHz Available online: https://www.arduitronics.com/product/4134 (accessed on 6 October 2021).

Arduino. Arduino Integrated Development Environment (IDE) v1 Available online: https://docs.arduino.cc/software/ide-v1/tutorials/arduino-ide-v1-basics (accessed on 16 June 2021).

Arduino LoRa. LoRa Available online: https://www.arduino.cc/reference/en/libraries/lora/ (accessed on 16 June 2021).

Thingspeak. Learn More About ThingSpeak Available online: https://thingspeak.com/ (accessed on 18 March 2021).

Google earth engine. FAQ Available online: https://earthengine.google.com/faq/ (accessed on 20 September 2021).

Tourismthailand. เกาะแตน Available online: https://thai.tourismthailand.org/Attraction/เกาะแตน (accessed on 10 September 2021).

NOAA. About our agency Available online: https://www.noaa.gov/about-our-agency (accessed on 27 September 2021).

James Moody. What does RMSE really mean? Available online: https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48e (accessed on 21 January 2022).

APPENDIX

APPENDIX A: Publication at the MDPI Sensor

MDPI (Sensor), vol.14, no.10, pp.1-13, 19 May 2022.

sustainability

MDPI

Article

Coral Reef Bleaching under Climate Change: Prediction Modeling and Machine Learning

Nathaphon Boonnam [1], Tanatpong Udomchaipitak [1], Supattra Puttinaovarat [1], Thanapong Chaichana [2], Veera Boonjing [3] and Jirapond Muangprathub [1,4,*]

[1] Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Surat Thani 84000, Thailand; nathaphon.b@psu.ac.th (N.B.); 6340320504@psu.ac.th (T.U.); supattra.p@psu.ac.th (S.P.)
[2] College of Maritime Studies and Management, Chiang Mai University, Samut Sakhon 74000, Thailand; thanapong@wavertree.org
[3] Department of Computer Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand; veera.bo@kmitl.ac.th
[4] Integrated High-Value of Oleochemical (IHVO) Research Center, Surat Thani Campus, Prince of Songkla University, Surat Thani 84000, Thailand
* Correspondence: jirapond.m@psu.ac.th; Tel.: +66-887-539-041

Abstract: The coral reefs are important ecosystems to protect underwater life and coastal areas. It is also a natural attraction that attracts many tourists to eco-tourism under the sea. However, the impact of climate change has led to coral reef bleaching and elevated mortality rates. Thus, this paper modeled and predicted coral reef bleaching under climate change by using machine learning techniques to provide the data to support coral reefs protection. Supervised machine learning was used to predict the level of coral damage based on previous information, while unsupervised machine learning was applied to model the coral reef bleaching area and discovery knowledge of the relationship among bleaching factors. In supervised machine learning, three widely used algorithms were included: Naïve Bayes, support vector machine (SVM), and decision tree. The accuracy of classifying coral reef bleaching under climate change was compared between these three models. Unsupervised machine learning based on a clustering technique was used to group similar characteristics of coral reef bleaching. Then, the correlation between bleaching conditions and characteristics was examined. We used a 5-year dataset obtained from the Department of Marine and Coastal Resources, Thailand, during 2013–2018. The results showed that SVM was the most effective classification model with 88.85% accuracy, followed by decision tree and Naïve Bayes that achieved 80.25% and 71.34% accuracy, respectively. In unsupervised machine learning, coral reef characteristics were clustered into six groups, and we found that seawater pH and sea surface temperature correlated with coral reef bleaching.

Keywords: coral reef bleaching; climate change; machine learning; sustainable management; predictive model

check for updates

Citation: Boonnam, N.; Udomchaipitak, T.; Puttinaovarat, S.; Chaichana, T.; Boonjing, V.; Muangprathub, J. Coral Reef Bleaching under Climate Change: Prediction Modeling and Machine Learning. *Sustainability* **2022**, *14*, 6161. https://doi.org/10.3390/su14106161

Academic Editors: Rezzy Eko Caraka, Eng. Anjar Dimara Sakti, Youngjo Lee and Toni Toharudin

Received: 14 March 2022
Accepted: 16 May 2022
Published: 19 May 2022

1. Introduction

Coral reefs are important and valuable ecosystems to protect underwater life and coastal areas. Moreover, the presence of many coral reefs supports commercial livelihood for fishing and tourism careers. Healthy coral reefs attract many tourists to eco-tourism under the sea. In Thailand, the income derived from the marine tourism industry in southern Thailand drives the GDP growth of country [1]. For this reason, if the level of marine fertility situation can be predicted, it can lead to an assessment of fishing and tourism careers subsistence. It also is an able to organize zoning tourism for further conservation further. Thus, this research presents a modeling and predictive approach to guideline the healthy coral reefs with consideration of coral reef bleaching. There are many

APPENDIX A: Publication at the MDPI Sensor (Cont.)

causes of coral reef bleaching, but the higher ocean temperatures from climate change is the leading cause of bleaching and elevated mortality rates [2,3]. The current global temperature has an increasing trend, with simultaneous slight warming of the ocean waters. Warmer water is a major contributor to coral bleaching [4–7]. The coral bleaching effect is caused by climate change and other processes, such as runoff and pollution (storm generated precipitations can rapidly dilute ocean water and runoff can carry pollutants that can bleach near-shore corals), overexposure to sunlight (when temperatures are high, high solar irradiance contributes to bleaching in shallow-water corals), and extreme low tides (exposure to the air during extreme low tides can cause bleaching in shallow corals). Thus, the sea surface temperatures are the main factor used for the predictive model in this work.

Another contributing factor is wind speed, which has an interaction effect with the seawater temperature. When the wind speed is low, a large amount of solar energy can penetrate the water surface, and this increases the water temperature [8]. Therefore, when corals are exposed to very high sunlight in combination with low wind speeds, the algae in the corals can be harmed by the sunlight, making coral bleaching more likely [9,10]. In addition to the factors mentioned above, many other factors affect coral reefs, such as the levels of chemical contaminants in the sea that largely have negative effects. Some of these are natural and some are anthropogenic. Contaminants can cause a change in pH, which can harm the corals. Coral tissue that is irritated tends to contract and form a gel layer [11,12]. Thus, an explicit and measurable expression of pH is chosen as one factor for this proposed modeling.

Predictive models are widely used for prevention and protection activities [10,13–16]. In particular, machine learning techniques have been applied to predict the bleaching of coral reefs. These techniques determine statistical relationships and patterns in a dataset by fitting a limited predefined model structure, or by statistical inference [13]. Most studies have used machine learning on remote-sensed imagery and in geospatial image processing for predicting coral reef bleaching. This approach is advantageous because it enables data visualization for monitoring coral reef bleaching. However, this technique requires costly data processing when an organization or a government agency provides big data for a retrospective study. Studies have reviewed data collected for coral reefs and coral bleaching [17,18]. Only a few studies have used machine learning techniques to predict and model collected data. Thus, this current study applied machine learning techniques, both supervised and unsupervised learning.

In this study, various alternative models were studied to select the model giving the highest accuracy. The study used three types of models, namely, Naïve Bayes, SVM, and Decision tree. The Naïve Bayes model makes probabilistic predictions of a classification label. Coral damage levels resolve classification problems based on findings are also predicted by a trained SVM. It is a model uses coefficients to create cluster boundaries. A decision tree model can be easily interpreted by people as a set of simple rules at the decision nodes. In this research, the natural factors that change over time are sea water temperature and wind speed. A further factor that may be influenced naturally or by anthropogenic effects is the pH of water. The retrospective data for these were collected from various databases for a period of 5 years. The study area is made up of coastal and island areas in southern Thailand, and the data were provided from the Department of Marine and Coastal Resources [19].

This article is organized as follows: Section 2 reviews the study factors of coral reef bleaching. Section 3 provides the materials and methods used for predictive modelling. Section 4 presents the proposed methodology Section 5 presents the results and discussion, and finally, Section 6 concludes the article.

2. Related Works

In this section, we address the study factors of coral reef bleaching. Coral bleaching is mainly caused by stress due to climate change, in which corals expel symbiotic algae to

APPENDIX A: Publication at the MDPI Sensor (Cont.)

turn white. Coral may bleach for other reasons, like extremely low tides, pollution, or too much sunlight. For this reason, many studies have addressed the factors effect to coral reef bleaching [4–7,20,21]. Melissa et al. [20] and Steve et al. [21] studied the temperature factor affecting coral bleaching. In an experimental study, samples of Acropora Formosa were taken to study the coral bleaching effects of temperature. The coral samples were divided into three treatment groups, with some placed in cool water, some in temperature-controlled water, and some in water with elevated temperature. The duration of the study was 20 days and some changes were observed on day 5 of the study, suggesting a decrease in coral algal density in the corals having cool or warm water. However, the corals in controlled temperature water had an increase in the algae inhabiting the coral. Afterwards, in the study from day 10 onwards, it became apparent that the warm temperature caused continually decreasing coral algal density, and this was a major contributor to coral bleaching. This is because the algae that live in corals escape from their original habitat to find a new habitat with a suitable water temperature. Thus, the temperature factor has a significant effect on coral bleaching.

In addition, climate change that leads to warming the ocean leads to higher sea levels and changes ocean conditions due to decreases in the salinity. Thus, sea water conditions such as turbidity and pH of water have implications for coral bleaching risk. The turbidity is partially caused by wind speed.

Paparella et al. [22] studied the effect of wind speed and daily temperature change. They found that both factors were highly correlated and led to effects on coral bleaching. The faster winds caused cooling, with the magnitude of temperature decline increasing with wind speed [8]. Recent researchers found that the pH change led to changes in calcium carbonate content of corals [11,12]. Tresguerres et al. [23] studied maintaining a steady pH to function all metabolic processes. They found that acid–base homeostasis mechanisms affect coral physiological responses. Understanding the physiological interactions between temperature stress and acid–base homeostasis is critical for predicting coral performance and acclimatization potential in a changing environment. However, little is known about how climate change affects acid–base homeostasis in corals. The potential effects of heat stress on acid–base regulation pose a particular challenge for maintaining coral calcification, as biomineralization is highly pH-dependent [24,25]. For these reasons, this work selects three mentioned factors to model forecasting on coral reefs bleaching.

3. Materials and Methods

Coral reefs are crucial for maintaining diverse ecosystems in the sea. Many studies have investigated the major causes of coral bleaching in various areas of the world [20]. Modern technologies such as machine learning can be used in such investigations. Machine learning is a part of artificial intelligence that automates analytical model building by using data, and automatically builds predictive models without being explicitly programmed for that task, requiring only little human involvement [26–28]. Machine learning is finding its way into every facet of not only society but also the natural world [6,29]. Most studies on coral reefs have applied machine learning to detect the health of coral reefs [29–32]. Machine learning algorithms learn to make decisions or predictions based on data. If the level of the marine fertility situation can be predicted, it can lead to an assessment of fishing and tourism careers subsistence. These algorithms can be traditionally classified into three main categories based on learning feedback [28,33]: supervised learning, unsupervised learning, and reinforcement learning. In this study, we focused on supervised and unsupervised learning.

Supervised learning is used to predict the level of coral bleaching by learning from previous information. Moreover, this approach has many algorithms that can be used to build the predictive model. Thus, this work selected three popular algorithms based on different construction to fit model. Namely, Naïve Bayes, SVM, and decision tree are built based on probability, functional, and information gain theory, respectively. On the other hand, unsupervised learning is used to cluster coral reef bleaching that has led

APPENDIX A: Publication at the MDPI Sensor (Cont.)

to organized zoning of the level of healthy coral reefs. Moreover, the association rule based on unsupervised learning is used to discover the relationship of the study coral bleaching factors.

3.1. Supervised Learning

Supervised learning algorithms can be used to build mathematical models of a set of data that contains both inputs and target outputs as training data (i.e., inputs and known outputs caused by them) [33]. The training data set is used to build a representative model that has learned the relationship between the input and output. The trained or fitted model is then used to evaluate its performance in test data. After testing, when an unseen case is fed into the system, the model can be used to predict the expected output. Recently, algorithms to train models have been developed for many approaches. We describe the following supervised learning algorithms for classification tasks that are applied in this study: Naïve Bayes, Support Vector Machine (SVM), and Decision Tree models. These procedures will use the R program as a management and analytical tool.

1. Naïve Bayes

Naïve Bayes or Bayesian classifier can be used to predict class membership probabilities, i.e., the probability that a given sample belongs to a particular class [34]. This type of model also can reach a high accuracy and speed when applied to large databases. Moreover, training the model requires only a small number of exemplars to learn the model parameters. The principles of classification are based on the Bayes theorem in conditional probability:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

where $P(H)$ and $P(X)$ are the probabilities of observing H and X without regard of each other, while $P(H|X)$ is the conditional probability of H given X, $P(X|H)$ is the conditional probability of X given H, and $P(X|H)/P(X)$ is called the likelihood ratio or Bayes factor. The workflow is summarized as follows:

- Represent each exemplar with a parameter vector $X = (x_1, x_2, \ldots, x_n)$, where x_i is factor attribute i.
- Compute probability of class label among the m classes $C_1, C_2, \ldots, C_m$ according to Bayes theorem

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)}$$

where class label in this work is the condition of level coral reefs.

- Compute the probability of each attribute for all class with $P(X|C_i)P(C_i)$ where $P(X|C_i)$ is calculated from product rule for independent events

$$P(X|C_i) = \prod_{k=1}^{n} P(x_k|C_i)$$

- Classify an unknown case X to the class C_i which gives the maximal $P(X|Ci)P(Ci)$.

2. Support Vector Machine

SVM is another widely used supervised learning method that reduces the empirical risk while maximizing the margin from a separating hyperplane to the separated classes [35]. Basically, SVMs are linear classifiers when they use linear kernel functions, which find a hyperplane to separate two classes of data. Linear kernel functions work effectively if classes are linearly separated. Nonlinear separation can be done with other kernel functions. Many mapping functions are available, including linear, polynomial, and radial basis functions (RBF). Polynomial and RBF kernel functions are commonly used depending on the training dataset. SVM maps the training exemplar to a point in high-dimensional space to separate the clusters with a hyperplane that maximizes the margin gaps to the two categories, while mapping back to the original lower dimensional space would show

APPENDIX A: Publication at the MDPI Sensor (Cont.)

this as nonlinear separation. SVM supports two-class classification. To manage multi-class classification, there are various strategies, such as one-against-rest and error-correcting output coding. The advantage of SVM is that it is based on theoretical mathematics and often provides high-performance classification for both high- and low-dimensional data.

3. Decision Tree

Decision tree is a learning model that performs classification through binary branchings at decision nodes. In the tree, each internal node denotes an attribute and a threshold, and each branch represents a value range of the attribute, while the final leaf nodes hold the class labels called. This model was applied using the algorithm in [34], To classify an unknown sample, the attribute values of the sample are tested against the decision tree, and the path is traversed from the root to a leaf node, which shows the class label call for that sample or exemplar [34]. Decision trees can be easily converted into classification rules, taken from the binary decision nodes. Thus, an unseen case without a class label can be classified simply by comparing attribute values with the nodes of the decision tree. The advantages of a decision tree are intuitively appealing knowledge expression, simple implementation, and high classification accuracy. Thus, we selected this model to compare its performance with the other candidate models.

3.2. Unsupervised Learning

An unsupervised learning algorithm learns patterns from an unlabeled set of inputs (there is no target output). Practically, this requires finding patterns, structures, or knowledge from unlabeled data. In this study, we used k-means clustering and association rules for unsupervised learning.

1. Clustering

Clustering is a method often used in exploratory analysis. Closely similar cases or exemplars are assigned to the same cluster that should differ from objects in the other clusters. A popular method is the k-means clustering that we use in this study. The k-means algorithm takes the training dataset and the number of clusters k as required inputs. The steps in the k-means algorithm are as follows:

- Select the value of k to estimate group of data. Next, each group of k will be guessed for the centroids.
- Calculate the distance data between the centroids k and data point for all centroids.Next, the minimum distance value of centroid k will be chosen to assign the first k clusters.
- Calculate the new centroid for each clusters from Step 2.
- Repeat Steps 2 and 3 until the algorithm converges to an answer.

An advantage of the k-means algorithm is that it is easy to use and the interpretation of cluster results is also easy. From the k-means algorithm, the k clusters are groups of similar data, but the choice of k should be determined by testing several alternatives. Commonly, the within sum of squares (WSS) metric is used to select the value of k. For WSS, the sum of squares is calculated for squared distance between each data point and the closest centroid, and these are compared over a range of k values.

2. Association Rules

Association rules record the associations or correlations among a large set of data [35]. The process of rule generation consists of two main steps: finding all frequent items and creating rules for the frequent items. In the first step, each record of data is counted for the item frequency, and a cut-off is based on a predetermined minimum support. Later, the frequent data sets are converted to association rules. The rules must satisfy both minimum support and minimum confidence.

In this study, we used the Apriori algorithm to generate association rules, in the WEKA software. Association rules are formed using if/then statements that determine the logical relationships in coral reef bleaching data. These rules summarize the relationships between causal factors and the condition of coral reefs. The minimum support and confidence used in this study were 0.80 and 0.95, respectively.

APPENDIX A: Publication at the MDPI Sensor (Cont.)

4. Research Methodology

In this study, the modeling and prediction of coral reef bleaching associated with climate change had three phases: data preparation, data modeling, and evaluation and deployment. The data preparation phase provided the data set in the first step in order to build the predictive model using machine learning. Next, the data modeling phase involved the fitting predictive model. The supervised machine learning i.e., Naïve Bayes, SVM, and decision tree, was used to predict level of coral damage from previous data. Likewise, unsupervised machine learning was used to cluster and recover the bleaching factors relationship with k-mean cluster and association rule, respectively. The final phase is the evaluation and deployment mention validation methods. Each phase was implemented as described below.

4.1. Data Preparation

This study examined the bleaching of corals in the southern marine area of Thailand. This area is rich in marine nature and has become a major tourist attraction of the country. In Thailand, the income derived from the marine tourism industry drives the GDP growth of country. Coral regions are ecological areas with high biodiversity. However, the occurrence of coral reef bleaching indicates serious damage to those ecosystems. Thus, this study tested modeling and predicting of coral reef bleaching to prepare guidelines for managing or preventing this phenomenon. We used data from the Department of Marine and Coastal Resources [19] that collects statistical information on marine resources in Thailand, for the years from 2013 to 2018. This study focused on three factors that affect coral bleaching: pH, sea surface temperature, and wind speed. We used these factors to model and predict the level of coral reef bleaching.

The data were collected by performing a staff survey in each quarter to explore the level of coral reef bleaching, where the level classification of damage was based on the ratio of total area of coral to area of coral bleaching, in each locality. This study examined 287 coral areas in the southern marine area of Thailand (Figure 1; border shading shows the coastal and island areas).

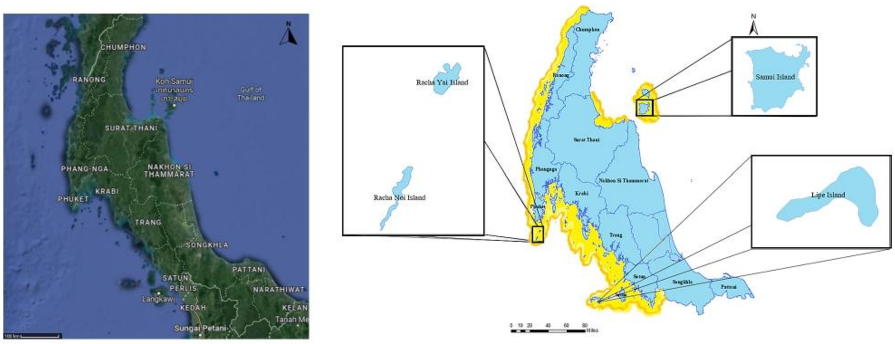
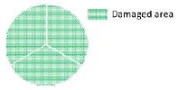
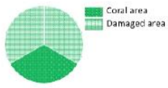
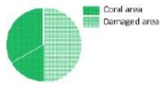
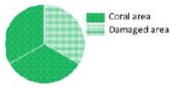
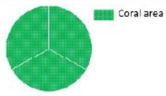


Figure 1. The area of study.

The overall condition of coral reefs was defined in five levels: completely damaged coral, damaged coral, moderately luxuriant coral, luxuriant coral, and perfectly luxuriant coral. These are also the targeted class labels in classifier models. The meaning of each level is shown in Table 1. The data set contained the coral colony count showing bleaching. (i.e., the percentage of coral reef recorded as bleached), which is an indication of the relative area of bleaching (Table 1).

APPENDIX A: Publication at the MDPI Sensor (Cont.)

Table 1. The condition of level of coral reef.

Classificaton	Percentage Coral Damage of Total Coral (%)	Represented Areas
Completely damaged	100% of total coral	
Damaged	75% of total coral	
Moderately luxuriant	50% of total coral	
Luxuriant	25% of total coral	
Perfect luxuriant	All corals were undamaged.	

The obtained data set was prepared for further analysis by removing unnecessary columns i.e., I.D. for each analysis. Missing values were handled by truncating the data in that record. Afterwards, the data were consolidated into forms appropriate for data analysis. Finally, we stored the output data in .csv format for the next step.

4.2. Data Model

The planning process was based on the study of algorithms used for the classification of coral bleaching conditions. In this study, we modeled with and compared the performances of three classifier types: Naïve Bayes, SVM, and decision tree. Figure 1 shows the study area. Moreover, we applied clustering to a group of similar factors and then used the association rule to determine dependent factors to suggest relevance of data used to examine coral reef bleaching.

Figure 2 shows steps including data segmentation, modeling, testing, and benchmarking. This study performed modeling in two parts: supervised learning (classification task) and unsupervised learning.

In the supervised learning part, Naïve Bayes, SVM, and decision trees were trained in order to select the best model for classifying and predicting the risk of coral bleaching, based on the aforementioned three causal factors. Naïve Bayes was used to categorize or classify coral data groups by using probability principles. Next, SVM was used to categorize coral condition in the areas of this study. Finally, a decision tree could analyze the data in the form of a tree diagram demonstrating the role of each causal factor or condition in causing the bleaching of corals. For this part, we divided the data into two sets: training and test sets (Training set: 70 and Test set: 30). The training dataset was used to fit each model to these data, while the test data were used to measure classification accuracy of each trained model in new data not shown to them earlier. Afterwards, the model showing the highest performance in test data was chosen for subsequent studies.

APPENDIX A: Publication at the MDPI Sensor (Cont.)

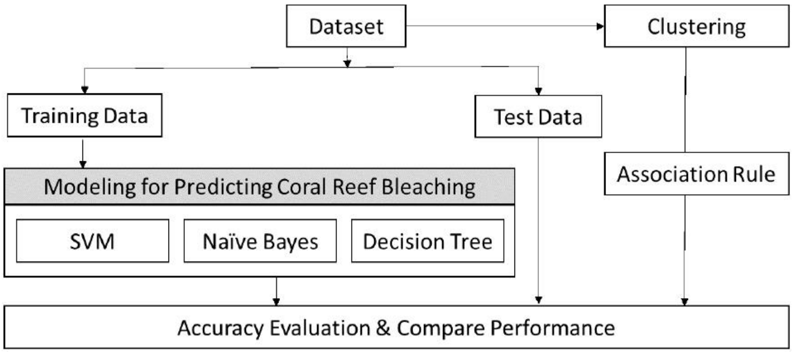


Figure 2. Overview of the proposed modeling.

In the unsupervised learning part, clustering was first used to group similar cases. This step applied the k-means clustering. Next, association rule learning was applied to determine the associations or correlations among a large set of data items. To determine factors contributing to coral bleaching in the southern sea regions of Thailand, each group of clusters was formulated to examine causes of coral bleaching.

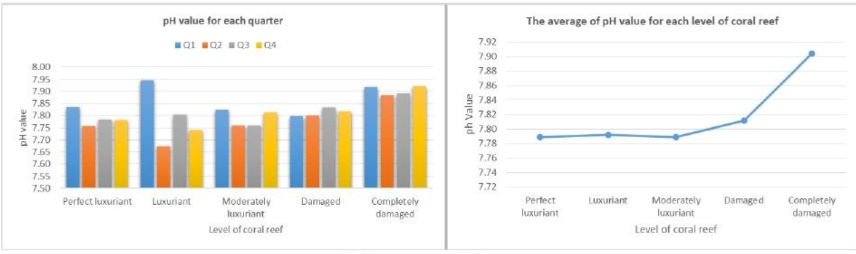
4.3. Evaluation and Deployment

Estimating classifier accuracy is crucial because it indicates how reliably it correctly calls the labels in future data. Thus, this study used 10-fold cross-validation to assess the classifier accuracies. Moreover, we used another accuracy indicator of the model performance namely the kappa coefficient. Its value is between 0 and 1, where 0 indicates no agreement of classifier calls with reference data, and 1 indicates perfectly identical calls with true labels. Thus, a larger kappa coefficient is better. The performance of the model was evaluated based on both classification accuracy and kappa coefficient. To examine the performance of clustering and of the association rules, within sum of squares (WSS) was used to find the best clustering and supply the grouped data. Afterwards, each group with similar factors was examined using an association rule with a minimum confidence threshold of 95%.

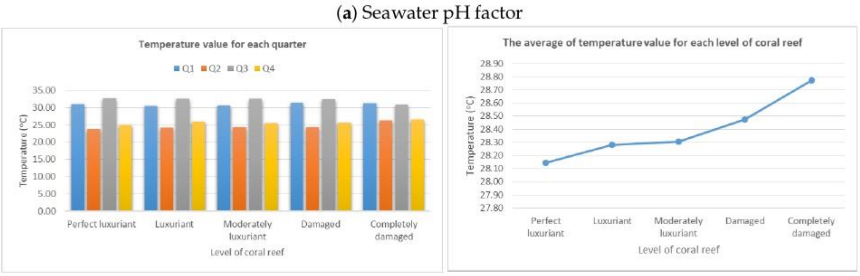
5. Results and Discussion

In this study, we used collected data to analyze coral reef bleaching in relation to climate change by considering three candidate causal factors: seawater pH, water temperature, and wind speed. The results of the preliminary analysis are shown in Figure 3, where (a), (b), and (c) depict findings for seawater pH, temperature, and wind speed, respectively. The time interval was divided into four quarters (Q). Each quarter on the left side of the figure shows the time interval for affecting the coral reef. The reason we collected four quarters of data is because the changes in the weather during each quarter results in changes in the parameters measured in the sea, which in turn can result in coral bleaching each quarter. The average across all quarters on the right side of figure shows that each of these factors damages coral reefs when the level is elevated.

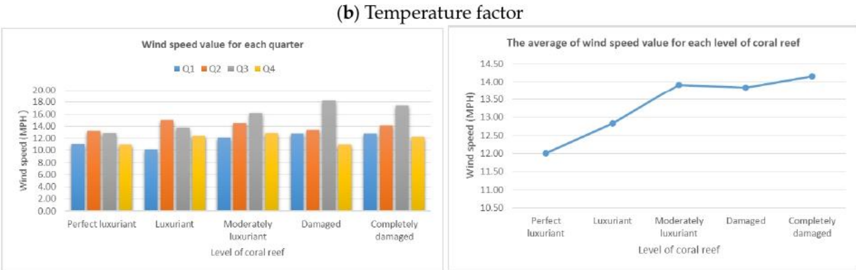
APPENDIX A: Publication at the MDPI Sensor (Cont.)



(**a**) Seawater pH factor



(**b**) Temperature factor



(**c**) Wind speed factor

Figure 3. The factors affecting the level of coral reef from years 2013 to 2018.

This study modeled and predicted coral reef bleaching by using machine learning techniques, including both supervised and unsupervised learning. First, supervised learning of a classification task was used to build models for predicting severity level of coral reef bleaching in a future situation. Then, unsupervised learning based on clustering and association rules was applied to a group of similar situations to assess correlations among each group of coral reef bleaching data. This study used the R program as a management and analytical tool with R-3.5.1 for Windows and RStudio Version 1.1.456. The results of supervised learning are listed in Table 2.

Table 2. The classification accuracy results of each model.

Model	Accuracy (%)	Kappa Coefficient
Naive Bayes	71.34	0.620
SVM	88.85	0.851
Decision Tree	80.25	0.735

APPENDIX A: Publication at the MDPI Sensor (Cont.)

The results show that the SVM model was the most accurate in classifying corals in the southern marine area of Thailand, having an accuracy of 88.85%, followed by the decision tree model with an accuracy of 80.25% and the Naïve Bayes model with an accuracy of 71.34%. This indicates that all three model types were effective in coral bleach classification in the sea area of southern Thailand. This is in accordance with some previous studies [26,36] that have used the Naïve Bayes model to predict the risk posed to the health and resilience of the coral reef system from adverse effects of climate change and harmful human activities, and the possible success of adaptation strategies. Thus, on the basis of these results, it is possible to examine relationships and conditions that cause the white foaming phenomenon in corals. These models can be used to model and predict coral reef bleaching during climate change. Later on, the data set was used to cluster groups of records with similar properties. WSS was used to find a suitable number of clusters for k-means clustering as shown in Figure 4.

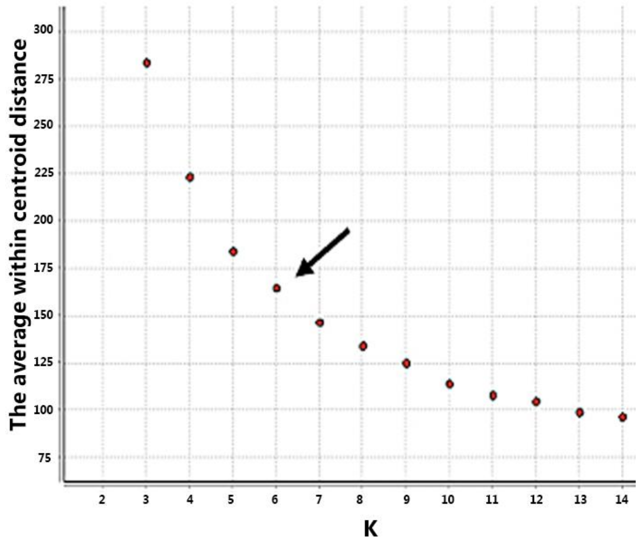
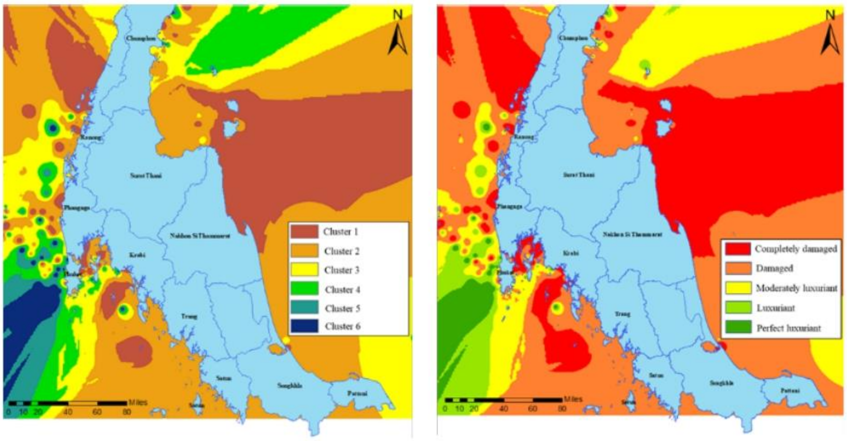


Figure 4. The results of WSS across k-values.

As shown in Figure 4, the k-means algorithm was run for $k = 2, 3, \ldots, 15$. The WSS was computed to determine a suitable k, the number of clusters. WSS was smaller with more clusters, allowing better splitting of the data for higher similarity within clusters. WSS declined significantly as the k value increased from 1 to 2. Another substantial reduction in WSS occurred at k value of 6; thus, from this analysis we selected $k = 6$. The process of determining the optimum value of k is known as finding the elbow in the WSS curve.

When we compared the classification accuracy of the condition of coral reef to a previous study [19] based on zoning features, we found that the proposed model provides more details of the condition as shown in Figure 5.

APPENDIX A: Publication at the MDPI Sensor (Cont.)



(**a**) The proposed model (**b**) The previous work

Figure 5. Comparison of the proposed model with general clustering for data from the Department of Marine and Coastal Resources. (Department of Marine and Coastal Resources, Thailand, 2020).

We found that using the presented model could provide details of the area that should be monitored and maintained because of relatively high coral fertility. Each group was discovered with the association rules from the Apriori algorithm. We found that pH depends on sea surface temperature. If the sea surface temperature increases, then the pH reduces [37–40]. Due to current climate change, it is likely that sea acidity and its temperature may continue to increase in the future [41].

In summary, this work applied the supervised machine learning with finding the fitted predictive model to predict the level of healthy coral reefs. The SVM model is the best model to classify bleaching status when we know the value of pH, sea surface temperature, and wind speed. In addition, the unsupervised machine learning using k-mean cluster lead to the explicit zoning of coral reefs bleaching. This method also discovered the relationship among the study factors of coral reef bleaching using the association rule approach. This discovery knowledge can be used as a guideline for the protection of coral reefs.

6. Conclusions

Coral reef bleaching is an important sign of marine ecosystem destruction, which affect subsistence and businesses in the marine aspect. Although climate change is the main cause and is unmanageable in a short time, the protection of coral reefs can still be obtained to prevent further damage. The past information can be used to guide coral safeguarding for the future. Thus, machine learning is suitable for use in decision support for coral reef protection. In this study, we demonstrated building a model for predicting the bleaching of coral reef by using machine learning. We applied three supervised learning algorithms, namely, Naïve Bayes, SVM, and decision trees, to select from these the model with the best classification performance. The SVM presented the best performance for classifying the level of coral reefs bleaching with 88.85% accuracy. The classifier calls indicated severity of bleaching. The developed model could be used to predict the level of coral reef bleaching under climate change. Unsupervised learning was used to obtain knowledge from previous coral reef bleaching data. The level of coral reef bleaching should be classified into six levels based on retrospective data. In addition, we found that the pH level is associated with the sea surface temperature. Thus, this study provides additional evidence that machine learning models are a viable and useful approach to monitoring and analyzing coral reef bleaching under climate change.

APPENDIX A: Publication at the MDPI Sensor (Cont.)

Author Contributions: Conceptualization, J.M., N.B.; Data curation, T.U.; Formal analysis, T.U., N.B., S.P. and J.M.; Investigation, T.U. and N.B.; Methodology, J.M., N.B.; Project administration, J.M., N.B.; Software, T.U. and J.M.; Supervision, N.B. and J.M.; Validation, N.B., S.P., T.C. and V.B.; Visualization, S.P.; Writing—original draft, T.U. and J.M.; Writing—review & editing, N.B., S.P., T.C. and V.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: shorturl.at/cvGLR.

Acknowledgments: The authors are deeply grateful to the Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Thailand. Moreover, the authors would like to thank the Department of Marine and Coastal Resources(DMCR), Thailand to support data and knowledge. This research was financially supported by Prince of Songkla University, Surat Thani Campus and Hat Yai Campus for research grants. The authors also gratefully acknowledge the helpful check of the English language by Seppo Karrila.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kasikornbank. Trends in the Thai Tourism Industry in 2019. IndustryBusiness. Available online: https://www.kasikornbank.com/international-business/th/Thailand (accessed on 28 September 2019).
2. Jeremiah, G.P.; Ulisse, C.; Nanne, V.H.; Elisa, B.; Ingo, B.; Malik, S.N.; Christian, W. Coral bleaching. In *Environmental Indicators*, 2nd ed.; Armon, R.H., Hänninen, O., Eds.; Springer Science+Business Media: Bremen, Germany, 2015; Chapter 9, pp 117–146. [CrossRef]
3. Ankeny, R.A.; Leonelli, S.; Matlin, K.; Maienschein, J. From the Beach to the Bench: Why Marine Biological Studies. In *Using Repertoires to Explore Changing Practices in Recent Coral Research*, 3rd ed.; Matlin, K.S., Maienschein, J., Ankeny, R.A., Eds.; University of Chicago Press: Chicago, IL, USA, 2020; pp. 249–270.
4. Fulton, C.J.; Berkström, C.; Wilson, S.K.; Abesamis, R.A.; Bradley, M.; Åkerlund, C.; Barrett, L.T.; Bucol, A.A.; Chacin, D.H.; Chong, S.K.M.; et al. Macroalgal meadow haitats support fish and fisheries in diverse tropical seascapes. *Fish Fish.* **2020**, *21*, 700–717. [CrossRef]
5. Lin, T.H.; Akamatsu, T.; Sinniger, F.; Harii, S. Exploring coral reef biodiversity via underwater soundscapes. *Biol. Conserv.* **2021**, *253*, 108901. [CrossRef]
6. Dahal, S.; Schaeffer, R.; Abdelfattah, E. Performance of Different Classification Models on National Coral Reef Monitoring Dataset. In Proceedings of the IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 27–30 January 2021; pp. 662–666. [CrossRef]
7. Baird, M.E.; Green, R.; Lowe, R.; Mongin, M.; Bougeot, E. Optimising cool-water injections to reduce thermal stress on coral reefs of the Great Barrier Reef. *PLoS ONE* **2020**, *15*, e0239978. [CrossRef] [PubMed]
8. NOAA. Coral Bleaching. Available online: https://www.coral.noaa.gov/education/bleaching-background.html (accessed on 18 April 2021).
9. Malcolm, J.S.; Michael, P.L.; Paul, L.J. Effects of ultraviolet radiation on corals and other coral reef organisms. *Glob. Chang. Biol.* **1996**, *2*, 527–545. [CrossRef]
10. Masona, R.A.B.; Skirvingb, W.J.; Dove, S.G. Integrating physiology with remote sensing to advance the prediction of coral bleaching events. *Remote Sens. Environ.* **2020**, *246*, 111794. [CrossRef]
11. Sylvain, A.; Ben, P.H.; Shigeki, W.; Koetsu, K.; Marco, M.; Kazuo, I.; Jason, M.H. Ocean acidification drives community shifts towards simplified non-calcified habitats in a subtropical temperate transition zone. *Sci. Rep.* **2018**, *8*, 11354. [CrossRef]
12. Tianran, C.; Shu, L.; Kefu, Y. Macrobioerosion in Porites corals in subtropical northern South China Sea: A limiting factor for high-latitude reef framework development. *Coral Reefs* **2013**, *32*, 101–108. [CrossRef]
13. Çınar, Z.M.; Abdussalam, N.A.; Zeeshan, Q.; Korhan, O.; Asmael, M.; Safaei, B. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. *Sustainability* **2020**, 12, 8211. [CrossRef]
14. Pernice, M.; Hughes, D.J. Forecasting global coral bleaching. *Nat. Clim. Chang.* **2019**, *9*, 803–804. [CrossRef]
15. Vercelloni, J.; Liquet, B.; Kennedy, E.V.; González-Rivero, M.; Caley, M.J.; Peterson, E.E.; Puotinen, M.; Hoegh-Guldberg, O.; Mengersen, K. Forecasting intensifying disturbance effects on coral reefs. *Glob. Chang. Biol.* **2020**, *26*, 2785–2797. [CrossRef]
16. Liu, G.; Eakin, C.M.; Chen, M.; Kumar, A.; Cour, J.L.D.L.; Heron, S.F.; Geiger, E.F.; Skirving, W.J.; Tirak, K.V.; Strong, A.E. Predicting Heat Stress to Inform Reef Management: NOAA Coral Reef Watch's 4-Month Coral Bleaching Outlook. *Glob. Chang. Future Ocean.* **2018**, *5*, 57. [CrossRef]
17. Tebbett, S.B.; Bellwood, D.R. Algal turf productivity on coral reefs: A meta-analysis. *Mar. Environ. Res.* **2021**, *168*, 105311. [CrossRef] [PubMed]

APPENDIX A: Publication at the MDPI Sensor (Cont.)

18. Schoepf, V.; Jung, M.U.; McCulloch, M.T.; White, N.E.; Stat, M.; Thomas, L. Thermally variable, macrotidal reef habitats promote rapid recovery from mass coral bleaching. *Front. Mar. Sci.* **2020**, *7*, 245. [CrossRef]
19. Marine, Coastal and Mangrove Resources Research and Development Institute; Department of Marine and Coastal Resources. Survey and Assessment Report of the Status and Potential of Marine and Coastal Resources, Coral and Sea Grass in 2013, 2014, 2015, and 2018. Available online: https://www.dmcr.go.th (accessed on 16 March 2021).
20. Melissa, S.R.; Ralf, G.; Dimitri, D.D. Cold induces acute stress but heat is ultimately more deleterious for the reef-building coral Acropora yongei. *Sci. Rep.* **2012**, *2*, 240.
21. Doo, S.S.; Leplastrier, A.; Graba Landry, A.; Harianto, J.; Coleman, R.A.; Byrne, M. Amelioration of ocean acidification and warming effects through physiological buffering of a macroalgae. *Ecol. Evol.* **2020**, *10*, 8465–8475. [CrossRef]
22. Paparella, F.; Xu, C.; Vaughan, G.O.; Burt, J.A. Coral bleaching in the Persian/Arabian Gulf is modulated by summer winds. *Front. Mar. Sci.* **2020**, *6*, 205. [CrossRef]
23. Tresguerres, M.; Clifford, A.M.; Harter, T.S.; Roa, J.N.; Thies, A.B.; Yee, D.P.; Brauner, C.J. Evolutionary links between intra- and extracellular acid–base regulation in fish and other aquatic animals. *J. Exp. Zool. Part A Ecol. Integr. Physiol.* **2020**, *333*, 449–465. [CrossRef]
24. Barott, K.L.; Venn, A.A.; Thies, A.B.; Tambutté, S.; Tresguerres, M. Regulation of coral calcification by the acid-base sensing en-zyme soluble adenylyl cyclase. *Biochem. Biophys. Res. Commun.* **2020**, *525*, 576–580. [CrossRef]
25. Innis, T.; Allen-Waller, L.; Brown, K.T.; Sparagon, W.; Carlson, C.; Kruse, E.; Huffmyer, A.S.; Nelson, C.E.; Putnam, H.M.; Barott, K.L. Marine heatwaves depress metabolic activity and impair cellular acid–base homeostasis in reef-building corals regardless of bleaching susceptibility. *Glob. Chang. Biol.* **2021**, *27*, 2728–2743. [CrossRef]
26. Wooldridge, S.; Done, T. Learning to predict large-scale coral bleaching from past events: A Bayesian approach using remotely sensed data, in-situ data, and environmental proxies. *Coral Reefs* **2004**, *23*, 96–108. [CrossRef]
27. Sen, P.C.; Hajra, M.; Ghosh, M. Supervised Classification Algorithms in Machine Learning: A Survey and Review In *Emerging Technology in Modelling and Graphics*, 2nd ed.; Mandal J., Bhattacharya D., Eds.; Springer: Singapore, 2020; Volume 937, pp. 99–111. [CrossRef]
28. Verbraeken, J.; Wolting, M.; Katzy, J.; Kloppenburg, J.; Verbelen, T.; Rellermeyer, J.S. A survey on distributed machine learning. *ACM Comput. Surv.* **2020**, *53*, 1–33. [CrossRef]
29. Gonzalez-Rivero, M.; Beijbom, O.; Rodriguez-Ramirez, A.; Bryant, D.E.; Ganase, A.; Gonzalez-Marrero, Y.; Herrera-Reveles, A.; Kennedy, E.V.; Kim, C.J.; Lopez-Marcano, S.; et al. Monitoring of coral reefs using artificial intelligence: A feasible and cost-effective approach. *Remote Sens.* **2020**, *12*, 489. [CrossRef]
30. Raphael, A.; Dubinsky, Z.; Iluz, D.; Netanyahu, N.S. Neural Network Recognition of Marine Benthos and Corals. *Diversity* **2020**, *12*, 29. [CrossRef]
31. Raphael, A.; Dubinsky, Z.; Iluz, D.; Benichou, J.I.; Netanyahu, N.S. Deep neural network recognition of shallow water corals in the Gulf of Eilat (Aqaba). *Sci. Rep.* **2020**, *10*, 12959. [CrossRef]
32. Hu, Z.; Hu, J.; Hu, H.; Zhou, Y. Predictive habitat suitability modeling of deep-sea framework-forming scleractinian corals in the Gulf of Mexico. *Sci. Total Environ.* **2020**, *742*, 140562. [CrossRef]
33. Murphy K.P. *Machine Learning: A Probabilistic Perspective*, 1st ed.; MIT Press: Cambridge, MA, USA; London, UK, 2012; ISBN 0262304325/9780262304320.
34. Han, J.; Pei, J.; Kamber, M. *Data Mining: Concepts and Techniques, Morgan Kaufmann*; Elsevier: San Francisco, MA, USA, 2017; ISBN 0123814804/9780123814807.
35. Xie, J.; Yu, F.R.; Huang, T.; Xie, R.; Liu, J.; Wang, C.; Liu, Y. A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 393–430. [CrossRef]
36. Carriger, J.F.; Yee, S.H.; Fisher, W.S. An introduction to Bayesian networks as assessment and decision support tools for managing coral reef ecosystem services. *Ocean. Coast. Manag.* **2019**, *177*, 188–199. [CrossRef]
37. Keith, A.H. The temperature dependence of pH in surface seawater. *Deep-Sea Res.* **1998**, *45*, 1919–1929.
38. Worachananant, S.; Carter, R.W.; Hockings, M.; Reopanichkul, P. Managing the Impacts of SCUBA Divers on Thailand's Coral Reefs. *J. Sustain. Tour.* **2008**, *16*, 645–663. [CrossRef]
39. Humphreys, A.F.; Halfar, J.; Ingle, J.C.; Manzello, D.; Reymond, C.E.; Westphal, H.; Riegl, B. Effect of seawater temperature, pH, and nutrients on the distribution and character of low abundance shallow water benthic foraminifera in the Gala'pagos. *PLoS ONE* **2018**, *13*, e0202746. [CrossRef] [PubMed]
40. Bahr, K.D.; Tran, T.; Jury, C.P.; Toonen, R.J. Abundance, size, and survival of recruits of the reef coral Pocillopora acuta under ocean warming and acidification. *PLoS ONE* **2020**, *15*, e0228168. [CrossRef] [PubMed]
41. Johnston, N.K.; Campbell, J.E.; Paul, V.J.; Hay, M.E. Effects of future climate on coral-coral competition. *PLoS ONE* **2020**, *15*, e0235465. [CrossRef] [PubMed]

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics.

International Journal of Design & Nature and Ecodynamics, vol.17, no.3, pp.423-431.

IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

International Journal of Design & Nature and Ecodynamics
Vol. 17, No. 3, June, 2022, pp. 423-431
Journal homepage: http://iieta.org/journals/ijdne

Forecast Coral Bleaching by Machine Learnings of Remotely Sensed Geospatial Data

Tanatpong Udomchaipitak[1], Nathaphon Boonnam[1], Supattra Puttinaovarat[1*], Paramate Horkaew[2]

[1] Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Surat Thani 84000, Thailand
[2] School of Computer Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

Corresponding Author Email: supattra.p@psu.ac.th

https://doi.org/10.18280/ijdne.170313

Received: 3 February 2022
Accepted: 22 May 2022

Keywords:
coral bleaching, machine learning, forecast, remote sensing, cloud

ABSTRACT

With the rapid changes in Earth climates, coral bleaching has been spreading worldwide and getting much severe. It is considered an imminent threat to marine animals as well as causing adverse impacts on fisheries and tourisms. Environmental agencies in affected regions have been made aware of the problem and hence starting to contain coral bleaching. Thus far, they often rely on conventional site survey to determine suitable sites to intervene and commence coral reef reviving process. With the recent advances in remote sensing technology, sea surface temperature (SST), acquired by satellites, has become a viable delegate to coral bleaching. Predicting coral bleaching based solely on SST is limited, as it is only one of many determinants. In addition, areas with different SST levels also exhibit different bleaching characteristics. Hence, area specific models are important for appropriately monitoring the events. Thus far, forecasting the bleaching based on SST alone has limited accuracy, because other disregarded factors are found equally influential. These are turbidity, salinity, and wind speed. Taken into account these geospatial factors, this paper evaluates different machine learning (ML) algorithms, on forecasting coral bleaching levels. Compared with official survey data, it was found that random forest (RF) gave the most accurate results, with accuracy and Kappa of 88.24% and 0.83, respectively. To further assist involved agencies in making data driven solutions to this problem, mapping forecasted by RF were visualized on a web application, implemented with Python and the most recent web frameworks and database systems. The proposed scheme could be extended to modelling coral bleaching in other areas, hence greatly reducing delayed in data acquisition and survey costs.

1. INTRODUCTION

Coral is a creature that greatly contribute to the ecosystem, because they are habitations of plants and animals, underwater attractions, important sources of medical research, and key elements in reducing oceanic wave violence. Coral bleaching has adverse impacts not only on coral ecosystem, but also on fishes and other marine animals, inhabiting along coral reefs. As a consequence, it also undermines livelihoods of fisherman and the local community [1-3]. Moreover, coral bleaching also rapidly deteriorates seaside exuberance, inevitably affecting local tourisms [4-6], and the benefits of otherwise undamaged marine ecosystem, such as retarding tidal wave and preventing coastal erosion [7-9]. Thus, detecting coral bleaching plays a major part in its monitoring, reviving, and devising preventive measures against bleached coral reefs.

One of the most common problems affecting a large number of corals is coral bleaching. It is defined as a phenomenon that corals become white or paler due to the loss of Zooxanthellae, which in turn is resulted from too unsuitable states for seaweed to survive. According to the related studies, it was revealed that, on one hand, the factors causing coral bleaching include exceedingly high sea temperatures [10-16], salinity [17-19], turbidity [20-22], and human doings, e.g., releasing wastes into seas, littering on beaches, and discharging wastewater. On the other hand, the factor that reduces sea temperatures, and hence helping corals to recover from bleaching and then to survive is wind speed [23-25].

The south of Thailand is not only attracting tourists worldwide, but also a location, maintaining abundant marine resources. Unfortunately, there are a number of areas in the region, which is currently affected by coral bleaching. Besides, environmental agencies have no means nor any technology to efficiently monitor coral bleaching. There are several most recent studies on the contributing factors of coral bleaching, that suggest applying remote sensing (RS) to acquire *in situ* these elementary factors, i.e., sea surface temperature (SST), salinity, turbidity, and wind speed. This is mainly because surveys and records from satellites are generally available over periods of time and can be utilized as inputs to coral bleaching, forecasting and monitoring. It has been shown elsewhere that a popular and highly accurate technique for similar tasks is machine learning (ML) [26-29]. Therefore, this paper presents a novel ML based method for forecasting coral bleaching by using RS. Its main objective was focused on monitoring the events. It will be later elucidated in this paper by experiments that, with the proposed method, manpower and onsite visit, required by conventional scheme, can be effectively reduced. As such, delayed data acquisition, and unnecessarily high survey cost, in each particular area can be effectively avoided.

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics. (Cont.)

2. LITERATURE REVIEW

Extensive works have been carried out in an attempt to forecast and monitor corals bleaching. Their analyses were mainly based on Geographical Information System (GIS) and or RS. The existing works can be categorized by factors being considered into two main groups. The former consists of those that take solely SST [10-16] into account, while those in the latter combine it with other factors (e.g., wind speed and water turbidity, etc.) in their analyses [30, 31]. However, based on the approach taken, these groups can be further divided into those detecting and monitoring corals bleaching by using satellite image processing [32, 33], by statistical analyses of SST, recorded in the photos archived in the National Oceanic and Atmospheric Administration (NOAA) repository [12-16], and by different ML algorithms, namely, random forest [30] and Bayesian [16]. Critical discussion and detailed insights into these methods are given as follow.

Investigations by Brown et al. and Van Hooidonk revealed the correlations between coral bleaching and SST levels, obtained from NOAA. Based on these correlations they could forecast and monitor the events in specific areas [12, 13]. By taking a similar approach, Pernice and Hughes developed a global-wide coral bleaching monitoring and warning system [14]. With this system, degree heating week (DHW) was analysed from NOAA and then used to monitor coral bleaching in real-time. Conventionally, DeCarlo presented the result of coral bleaching correlated to only higher SST. Therefore, a framework based on advanced statistical approach was proposed to improve the accuracy of forecast [15]. However, detailed quantitative evaluation on its accuracy was not presented. Other studies tackled the issue by using ML algorithms. Based on SST accumulation within 4-8 weeks period, Lachs et al. generated a Bayesian hierarchical model that was able to forecast coral bleaching at specified areas, with 7.9% higher accuracy than the traditional methods [16]. Motivated by this improvement, it is believed that ML based methods could well overcome conventional statistics. These methods, however, did not consider any other factors apart from SST.

By taking into account other determinants, Knudby et al. presented the methodology for mapping coral reef resilience indicators using field and remotely sensed data [30]. Therein, indicators related to both coral bleaching coral reef resilience were analysed by using remote sensing imagery at very high resolution, i.e., IKONOS and QUICKBIRD. Beside SST, those indicative factors were stress-tolerant taxa, coral generic diversity, fish herbivore biomass, fish herbivore functional group richness, live coral, and crustose, and coralline algae. An ML algorithm called random forest (RF) and Gaussian process regression (Kriging interpolation) were compared. It was found that, mapping of coral bleaching and reef resilience were clearly presented at higher resolution, thanks to that of satellite base images. Additionally, Aslam et al. presented the methodology for mapping of coral bleaching susceptibility was also studied [31], by using multi-criteria analysis (MCA) and GIS. In their work, other factors, i.e., wind velocity, photosynthetic active radiation (PAR), aragonite saturation state, bathymetry, and reef slope, were considered with SST. Based on these factors, locations on susceptibility mapping were classified into 3 levels, i.e., low, moderate, and high risks. Their experiments demonstrated that MCA and GIS played important parts in highly efficient data analyses and hence enhancing the mapping accuracy. Like their many preceding works, however, numerical accuracy assessment was not presented.

Image processing was found a viable tool for detecting and monitoring coral bleaching in many studies. For example, Xu et al. presented the method for examining Sentinel-2 satellite images at different durations. The images were analysed by using Pseudo invariant features (PIFs) and depth invariant indices (DII) methods. Subsequently, the processed pixels were classified by support vector machine (SVM). Their experiments indicated that coral bleaching could be examined at a particular area with 88.9% accuracy on average [32].

Last but not least, GIS is considered a driving technology in various fields, e.g., geoscience, environment, and public health. Levine and Feinholz presented the method to inspect coral reef management. It was used for displaying the coastal and oceanic data in Hawaii on graphical maps for coral reef management [33]. The system was able to present relevant spatial information to government agencies and thus assisted them to make data-driven managerial decisions.

From the above review on coral bleaching forecast and monitoring schemes, it could be drawn that there remain some areas that need improving and issues need to be addressed, especially, the accuracies of detecting and predicting coral bleaching. The above studies have come to a similar conclusion that using SST alone was inadequate. However, there were only a few studies that attempts to combine it with other factors in their analyses. Moreover, not only that the combinations were not explicitly standardized, but also that accuracy assessments were not reported. Nonetheless, the results obtained by applying ML algorithms on remotely sensed data were very promising, despite some limitations. To address these issues, this paper presents a method for making coral bleaching forecast, based on satellite image processing and ML algorithms. Its merit was elucidated by demonstrating a web application for monitoring coral bleaching events at targeted areas.

3. MATERIALS AND METHODS

The design and development of the proposed system were divided into 3 key processes as illustrated in Figure 1. They were spatial data query, coral bleaching forecast by using ML, and visualization of forecast data. The description of each process is given in detailed in the subsections to follow. The subsequent experiments were carried out on 2 study areas, which were Chumphon and Surat Thani provinces. In these areas, there were various extents of coral bleaching, including non-existent. With these as the reference, the constructed forecasting model was assessed by its accuracy on anticipating the extents of coral bleaching.

3.1 Spatial data query

This process queried surveyed data recorded by satellites. They were stored on the Google cloud repository, called Earth Engine Data catalogue. From this platform, a range of remotely sensed data, such as satellite images, terrains, and climate and weather information acquired by meteorological satellites, could be accessed via trivial Python scripts. In this study, the data involved were SST, turbidity, salinity, and wind speed, whose characteristics and sources are provided in Table 1. Shown in Figure 2a is an example of SST mapping, where different levels of temperature are represented in standard

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics. (Cont.)

false colors. Specifically, high, moderate, and low temperatures within a given range, were shown in various shades of red to purple, green, and blue, respectively. Figure 2b illustrated an example of turbidity ranging between –0.5 and 0.0. Likewise, each pixel is false colored, with respect to its relative levels. Locations with high turbidity close to 0.0, for examples, were plotted in purple, whilst those on the other end were in blue. Based on this representation, the lower the values, the clearer the water in that area. An example of wind speed is displayed in Figure 2c. Unlike previous mapping schemes, its false color palette is discrete, ranging from blue at the lower end speed, to green, orange, and red at the other. All these spatial data were then fed into the forecast process by using ML algorithms.

The RS data considered in this study contained factors with different spatial resolution, depending on their availability. To normalize the resolution of all relevant layers to 30 meters, Kriging algorithm was employed to interpolate the data, prior to feature extraction. The features considered in this study were selected from those suggested in related works, and those reportedly causing coral bleaching at the study area in the past. These factors were also validated against the created model.

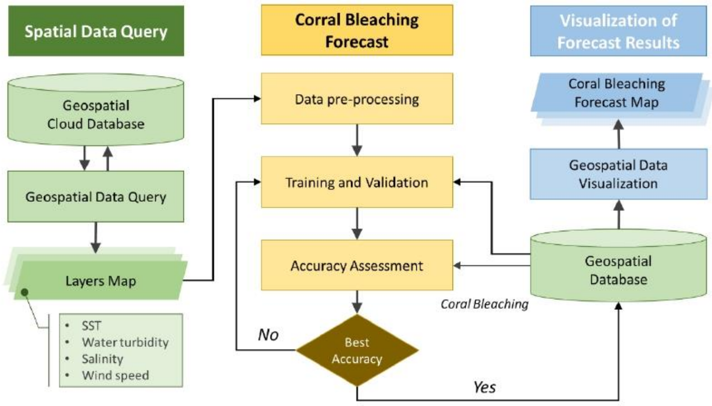


Figure 1. Conceptual diagram of the proposed system

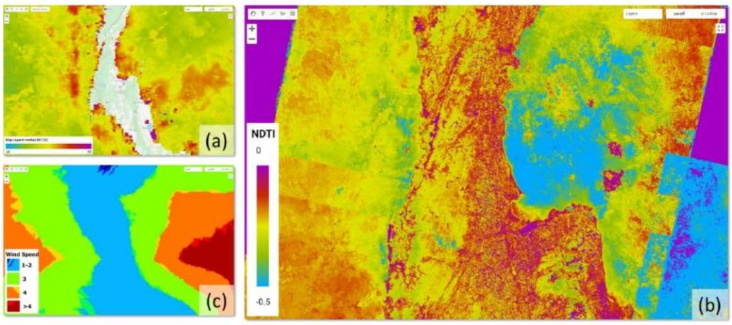


Figure 2. Examples of sea surface temperature (a), water turbidity (b), and wind speed (c)

Table 1. Detailed characteristics of spatial data and their sources

Data	Type	Source	Period
Sea Surface Temperature (SST)	Spatial	GEE (MODIS and NOAA)	Summer, 2020
Water-Turbidity	Spatial	GEE (Landsat 8)	Summer, 2020
Salinity	Spatial	GEE (HYCOM: Hybrid Coordinate Ocean Model)	Summer, 2020
Wind Speed	Spatial	GEE (NOAA CDR)	Summer, 2020

Note: GEE stands for Google Earth Engine data catalogue.

3.2 Coral bleaching forecast by using ML

This subsection explains the construction of forecasting model. The model received inputs from GEE repository via spatial queries, and was trained by onsite survey data, explored, and recorded by the Department of Marine and Coastal Resources (DMCR), Thailand. These data were fetched and stored in the local geospatial database, for training and testing

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics. (Cont.)

purposes. In this study, four ML algorithms were developed and benchmarked. They were multi-layer perceptron (MLP) artificial neural network (ANN), random forest (RF), decision tree (DT), and radial basis function (RBF) ANN. On constructing an ML model 149 records were learned. Each resultant model was then assessed by means of 10-fold cross validation. Another set of 51 unseen records was used for testing, to ensure that the model was not overfitted by training data. The model that gave the highest accuracy would then be chosen for producing the forecast maps in the study areas, which were also stored in the local database.

On accessing forecasting accuracy, standard metrics were calculated for each model. They were overall accuracy, root mean squared error (RMSE), and Kappa. The validation was made against the ground truths, i.e., survey data, obtained from DMCR.

During the ML process, thematic features were extracted from a number of instances from the geospatial database and tagged with their geolocations (latitude and longitude) and acquisition time. Each model was trained with identical set of data in turns, using their default configurations, and the one that yielded the best accuracy was chosen. The results were visualized in two modes, i.e., per location and as an image of the region of interest, showing the bleaching levels.

3.3 Visualization of forecast results

The forecasts made by the best performing model were visualized on an interactive map embedded on an inhouse web application. The application was mainly written in Python and involved various software libraries and database operating systems, i.e., Google Earth Engine (GEE), QGIS, QGIS2Web, PostgreSQL, and PostGIS. With this application, forecast map depicts a selected area on a web canvas, whose pixels were colour-coded by bleaching level at that location. Accordingly, users could access this information online and rely on the forecasts to monitor coral bleaching, and to make informed decision or to support administrative actions on resolving environmental issues, related to coral bleaching at a particular area.

4. RESULTS AND DISCUSSION

4.1 Results

The experimental results are divided into 3 parts, i.e., the ground truth data of coral bleaching obtained DMCR, numerical assessments, and bleaching forecasts made by the best performing ML. The reference maps of study areas are depicted in Figure 3a. The coloured star shape markers represent different levels of bleaching. Dark green stars indicates that the area has perfect corals with no bleaching. Light green, yellow, orange, and red stars, indicate those with low, medium-low, medium-high, and high levels of coral bleaching, respectively. The rest depicts the maps of four selected islands, namely, Kho Samui (3b), Koh Phangan (3c), and Kho Mae Ko and Kho Wua Ta Lap (3d).

Table 2. 10-fold cross validation results, obtained by different ML models

Machine Learning	Accuracy	RMSE	Kappa
ANN (MLP)	88.07	0.21	0.85
Random Forest	97.25	0.08	0.96
Decision Tree	94.49	0.14	0.93
ANN (RBF)	90.83	0.18	0.88

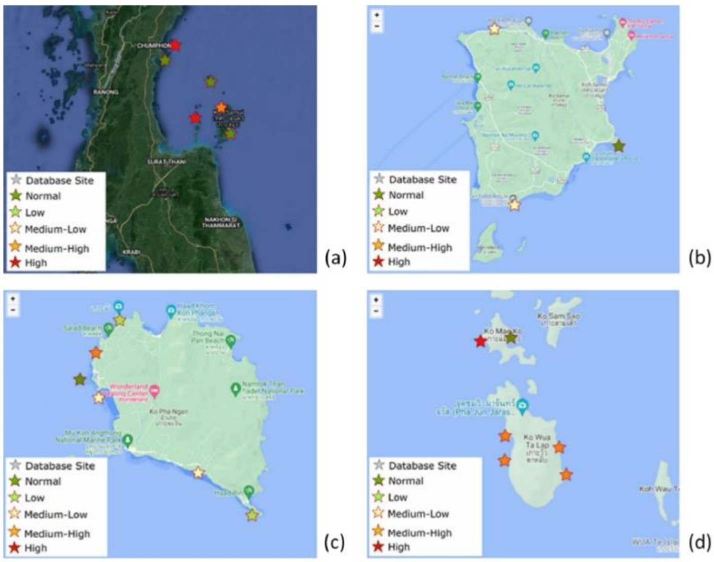


Figure 3. Reference coral bleaching data obtained from DMCR, showing the entire study area (a) and four selected islands (b), (c) and (d)

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics. (Cont.)

Once training was completed, forecasts made by each model were evaluated by 10-fold cross validation. The ML models considered in this study were ANN (MLP), RF, DT, and ANN (RBF). The assessment results, i.e., accuracy, RMSE, and Kappa, are lists in Table 2 and plotted in Figure 4. According to these results, it was revealed that RF performed best in terms of accuracy at 97.25%. It was followed by DT, ANN (RBF), and ANN (MLP), whose accuracies were 94.49%, 90.83%, and 88.07%, respectively. This finding was also confirmed by measuring the RMSE and Kappa. In these aspects, RF also gave the least RMSE (0.08) and the highest Kappa (0.96), compared to its counterparts. Figures 4(a), 4(b), and 4(c) graphically compare these ML algorithms with respect to each metric, i.e., overall accuracy, RMSE, and Kappa, respectively.

Among these model candidates, RF yielded the highest accuracy, with respect to 10-folds cross validation. Its forecast would thus be preferred for the subsequent process. However, to ensure that it was not overfitted, another validation was performed on RF model, based on 51 unseen locations. In this experiment, it could forecast the bleaching levels at 88.24% accuracy, with Kappa = 0.83 and RMSE = 0.25. Therefore, it is safe to conclude that the RF model maintained its high performance, and hence was the most suitable predictor for coral bleaching.

Coral bleaching levels predicted by RF at a selected site, based on the factors listed in the Table 1 is illustrated in Figure 5. Its dense map was produced by Kriging interpolation. In this map, red brown, brown, and light-yellow pixels, represent locations with high, medium high, low levels of coral bleaching. While green ones represent those with perfect corals or those without bleaching. By visual inspection, comparing the forecasted bleaching levels (false color pixels) with against onsite surveys by DMCR (colored stars), it was found that they were highly correlate. For example, the northern part of the study is Koh Ngam Yai, Chumphon province. The forecast in this area appeared in red brown, indicating high level of bleaching. This conformed to that marked by survey data as red star-shaped markers. Another example is Koh Samui, located in Surat Thani province. This area had varying levels of bleaching, from low (green) to medium-low (light yellow). By slight difference, it was marked by green star-shaped markers.

A more localized forecast mapping at two selected islands is shown in Figures 6 and 7. The former depicts the results at Koh Wua Ta Lap, while the latter does at Koh Mae Ko. In these figures, stripped pattern indicates targeted coral reefs (a), which are also overlaid on the map rendered with their bleaching levels (b). It is apparent in Figure 6 that most of Koh Wua Ta Lap areas had medium-high to high levels of bleaching. However, the north and the northeast of the island exhibited perfect coral to low coral bleaching. With similar observation, Figure 7 shows that Koh Mae Ko had the extent of medium-high to high bleaching were approximately the same as that of none to low bleaching. Specifically, the former laid on the west, the north, the northeast and the southwest, while the latter did on the east, the northeast, and the southeast of the island.

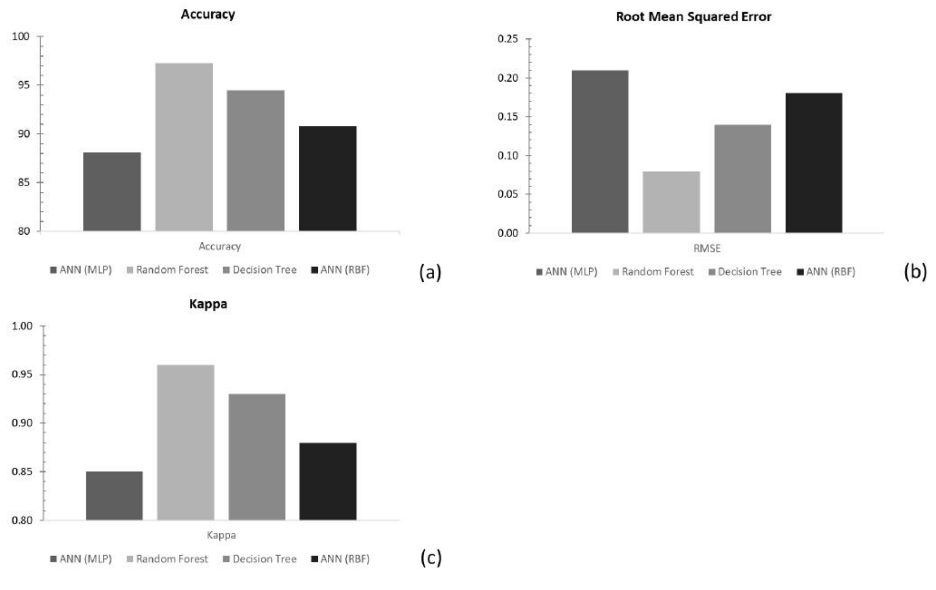


Figure 4. Comparison between different ML algorithms with respect to overall accuracy (a), RMSE (b), and Kappa (c)

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics. (Cont.)

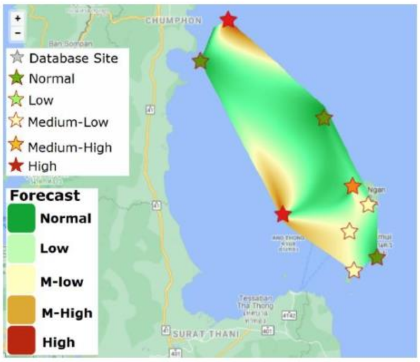
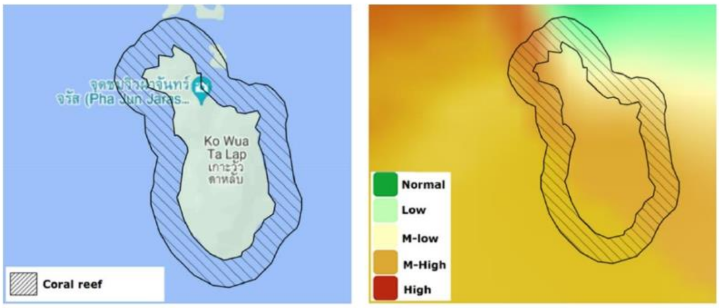
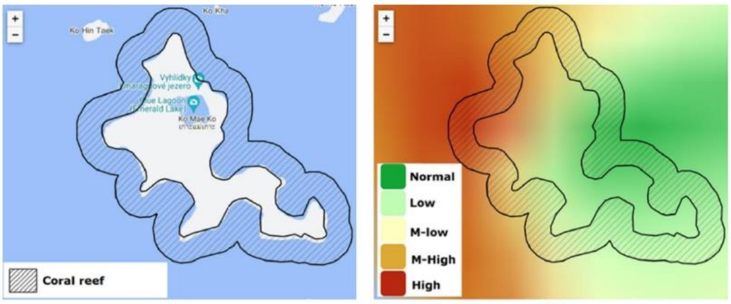


Figure 5. A selected example of forecasted coral bleaching map



(a) (b)

Figure 6. Localized forecasted coral bleaching levels at Koh Wua Ta Lap, showing the targeted coral reefs (a) overlaid on the forecast map (b)



(a) (b)

Figure 7. Localized forecasted coral bleaching levels at Koh Mae Ko, showing the targeted coral reefs (a) overlaid on the forecast map (b)

Closer inspections on SST data, acquired between 2011 – 2021, also revealed that an area with coral bleaching generally had higher SST than that without. For instance, in May 2020, the former had SST than the latter by 2–3°C. When used as a casual factor in previous studies, however, it is not so discriminative as those combined in this study.

Finally, the forecasted coral bleaching, made the by RF algorithm, could be visualized on an interactive map, by using the developed web application, as illustrated in Figure 8. All main functions discussed above were implemented and built in. With this web application, an involved party, could make online spatial queries on coral bleaching at an area of interest, to monitor bleaching and also to prepare appropriate actions to resolve environmental related issues in the area.

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics. (Cont.)

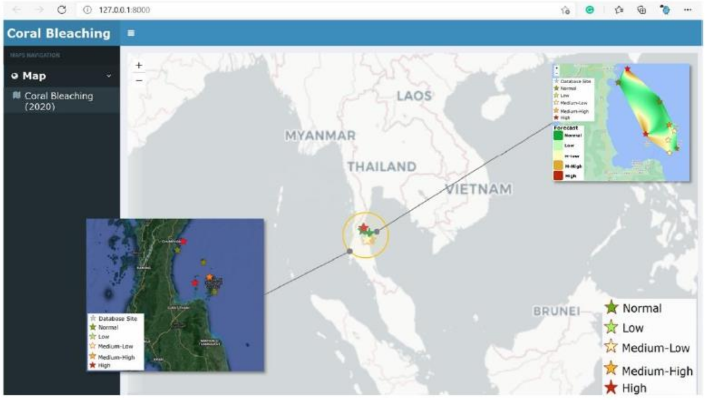


Figure 8. The starting page of the prototype web-based map visualization application. The insets depict the main built-in functions for coral bleaching forecast

4.2 Discussion

Unlike a few related works that addressed similar issues, the present study compared state-of-the-art ML algorithms, and discovered that, based on remotely sensed data, i.e., SST, turbidity, salinity, and wind speed, the RF algorithm performed best. Specifically, it could forecast coral bleaching at 97.25% and 88.24% accuracy, based on 10-fold cross validation and that on a set of unseen data, respectively. Compared with previous studies that typically relied only on SST [12-16], the present results also took into account other remotely sensed factors, resulting in much accurate forecasts. Although there were a few studies that analysed multiple factors [30, 31, 34], neither accuracy on coral bleaching analysis nor monitoring were reported. Furthermore, we also developed a web-based application that was able to make spatial queries to remotely sensed cloud repository, hosted by GEE, by using well-known Python language and the most recent web frameworks. It was demonstrated that users could visualize forecasted map of coral bleaching at specific areas, interactively, without the need for onsite surveys.

5. CONCLUSIONS

This paper presents a novel ML based method for forecasting coral bleaching levels, based on SST, turbidity, salinity, and wind speed. These geospatial factors were queried from GEE data catalogue by using Python scripts and peripheral web frameworks. Herein, experimental results indicated that RF could forecast coral bleaching in the study area most accurately, compared to other ML algorithms. The resultant forecasts were stored in local geospatial database. As such, they were accessible via a web application. Information interactively presented on this application is valuable to involved agencies in making data-driven managerial decisions and devising appropriate solutions to coral bleaching problems.

Thus far, RS remains limited, since it is unable to acquire human activities, such as littering and discharging wastewater into the ocean, both from industrial plants and households. The solution to these shortcomings has not been addressed in the present study. However, a dedicated geospatial platform could be implemented and employed to gather these activities, so as to be included in forecasting model, making it more accurate.

Future directions worth explored include using Internet of Things (IoT), equipped with sensors to acquire more localized and up-to-date data, e.g., SST and turbidity. It is believed that with more sensors being installed, the precision of a forecast map can be improved, both spatially and temporally.

ACKNOWLEDGMENT

The authors would like to thank the Google and the Department of Marine and Coastal Resources of Thailand for providing the remotely sensed data used in preparation of the paper.

The work was financially supported by Prince of Songkla University, Surat Thani Campus, Graduate School, Prince of Songkla University Surat Campus.

REFERENCES

[1] Hoegh-Guldberg, O., Poloczanska, E.S., Skirving, W., Dove, S. (2017). Coral reef ecosystems under climate change and ocean acidification. Frontiers in Marine Science, 4: 158. https://doi.org/10.3389/fmars.2017.00158

[2] Ouédraogo, D.Y., Delaunay, M., Sordello, R., et al. (2021). Evidence on the impacts of chemicals arising from human activity on tropical reef-building corals; a systematic map. Environmental Evidence, 10(1): 1-18. https://doi.org/10.1186/s13750-021-00237-9

[3] Pendleton, L., Comte, A., Langdon, C., et al. (2016). Coral reefs and people in a high-CO_2 world: where can science make a difference to people? PloS One, 11(11): e0164699. https://doi.org/10.1371/journal.pone.0164699

[4] Caparrós-Martínez, J.L., Martínez-Vázquez, R.M., de Pablo Valenciano, J. (2022). Analysis and global research trends on nautical tourism and green coastal infrastructures: The case of coral reefs and seagrass meadows. Environmental Sciences Europe, 34(1): 1-13. https://doi.org/10.1186/s12302-022-00614-2

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics. (Cont.)

[5] Spalding, M., Burke, L., Wood, S.A., Ashpole, J., Hutchison, J., Zu Ermgassen, P. (2017). Mapping the global value and distribution of coral reef tourism. Marine Policy, 82: 104-113. https://doi.org/10.1016/j.marpol.2017.05.014
[6] Phillips, W.N. (2015). Tourism threats to coral reef resilience at Koh Sak, Pattaya bay. Environment and Natural Resources Journal, 13(1): 47-60. Retrieved from https://ph02.tci-thaijo.org/index.php/ennrj/article/view/83867.
[7] Escudero, M., Reguero, B.G., Mendoza, E., Secaira, F., Silva, R. (2021). Coral reef geometry and hydrodynamics in beach erosion control in North Quintana Roo, Mexico. Frontiers in Marine Science. https://doi.org/10.3389/fmars.2021.684732
[8] Zhao, M., Zhang, H., Zhong, Y., et al. (2019). The status of coral reefs and its importance for coastal protection: a case study of Northeastern Hainan Island, South China Sea. Sustainability, 11(16): 4354. https://doi.org/10.3390/su11164354
[9] Reguero, B.G., Beck, M.W., Agostini, V.N., Kramer, P., Hancock, B. (2018). Coral reefs for coastal protection: A new methodological approach and engineering case study in Grenada. Journal of Environmental Management, 210: 146-161. https://doi.org/10.1016/j.jenvman.2018.01.024
[10] Dao, H.N., Vu, H.T., Kay, S., Sailley, S. (2021). Impact of seawater temperature on coral reefs in the context of climate change. A case study of Cu Lao Cham–Hoi an biosphere reserve. Frontiers in Marine Science, 1057. https://doi.org/10.3389/fmars.2021.704682
[11] Brown, B.E., Dunne, R.P., Somerfield, P.J., Edwards, A.J., Simons, W.J.F., Phongsuwan, N., Naeije, M.C. (2019). Long-term impacts of rising sea temperature and sea level on shallow water coral communities over a~ 40 year period. Scientific Reports, 9(1): 1-12. https://doi.org/10.1038/s41598-019-45188-x
[12] Van Hooidonk, R., Huber, M. (2009). Quantifying the quality of coral bleaching predictions. Coral Reefs, 28(3): 579-587. https://doi.org/10.1007/s00338-009-0502-z
[13] Strong, A.E., Liu, G., Skirving, W., Eakin, C.M. (2011). NOAA's coral reef watch program from satellite observations. Annals of GIS, 17(2): 83-92. https://doi.org/10.1080/19475683.2011.576266
[14] Pernice, M., Hughes, D.J. (2019). Forecasting global coral bleaching. Nature Climate Change, 9(11): 803-804. https://doi.org/10.1038/s41558-019-0593-7
[15] DeCarlo, T.M. (2020). The past century of coral bleaching in the Saudi Arabian central Red Sea. PeerJ, 8: e10200. https://doi.org/10.7717/peerj.10200
[16] Lachs, L., Bythell, J.C., East, H.K., Edwards, A.J., Mumby, P.J., Skirving, W.J., Guest, J.R. (2021). Fine-tuning heat stress algorithms to optimise global predictions of mass coral bleaching. Remote Sens. 2021, 13(14): 2677. https://doi.org/10.3390/rs13142677
[17] Samlansin, K., Chawakitchareon, P., Rungsupa, S. (2020). Effects of salinity and nitrate on coral health levels of Acropora sp. Thai Environmental Engineering Journal, 34(1): 19-26.
[18] Kuanui, P., Chavanich, S., Viyakarn, V., Omori, M., Lin, C. (2015). Effects of temperature and salinity on survival rate of cultured corals and photosynthetic efficiency of zooxanthellae in coral tissues. Ocean Science Journal, 50(2): 263-268. https://doi.org/10.1007/s12601-015-0023-3
[19] Gegner, H.M., Ziegler, M., Rädecker, N., Buitrago-López, C., Aranda, M., Voolstra, C.R. (2017). High salinity conveys thermotolerance in the coral model Aiptasia. Biology Open, 6(12): 1943-1948. https://doi.org/10.1242/bio.028878
[20] O'Leary, M., Morgan, K., Browne, N.K. (2021). Turbid coral reefs: past, present and future-a review. Diversity, 13(6): 251. https://doi.org/10.3390/d13060251
[21] Morgan, K.M., Perry, C.T., Johnson, J.A., Smithers, S.G. (2017). Nearshore turbid-zone corals exhibit high bleaching tolerance on the Great Barrier Reef following the 2016 ocean warming event. Frontiers in Marine Science, 4: 224. https://doi.org/10.3389/fmars.2017.00224
[22] Hughes, T.P., Kerry, J.T., Álvarez-Noriega, M., Álvarez-Romero, J.G., Anderson, K.D., Baird, A.H., Wilson, S.K. (2017). Global warming and recurrent mass bleaching of corals. Nature, 543(7645): 373-377. https://doi.org/10.1038/nature21707
[23] Paparella, F., Xu, C., Vaughan, G.O., Burt, J.A. (2019). Coral bleaching in the Persian/Arabian gulf is modulated by summer winds. Frontiers in Marine Science, 6: 205. https://doi.org/10.3389/fmars.2019.00205
[24] Barnes, B.B., Hallock, P., Hu, C., Muller-Karger, F., Palandro, D., Walter, C., Zepp, R. (2015). Prediction of coral bleaching in the Florida Keys using remotely sensed data. Coral Reefs, 34(2): 491-503. https://doi.org/10.1007/s00338-015-1258-2
[25] Yee, S.H., Barron, M.G. (2010). Predicting coral bleaching in response to environmental stressors using 8 years of global-scale data. Environmental Monitoring and Assessment, 161(1): 423-438. https://doi.org/10.1007/s10661-009-0758-3
[26] Horkaew, P., Puttinaovarat, S., Khaimook, K. (2015). River boundary delineation from remotely sensed imagery based on SVM and relaxation labeling of water index and DSM. Journal of Theoretical & Applied Information Technology, 71(3): 376-386.
[27] da Silveira, C.B.L., Strenzel, G.M.R., Maida, M., Gaspar, A.L.B., Ferreira, B.P. (2021). Coral reef mapping with remote sensing and machine learning: A nurture and nature analysis in marine protected areas. Remote Sensing, 13(15): 2907. https://doi.org/10.3390/rs13152907
[28] Puttinaovarat, S., Horkaew, P. (2019). Deep and machine learnings of remotely sensed imagery and its multi-band visual features for detecting oil palm plantation. Earth Science Informatics, 12(4): 429-446. https://doi.org/10.1007/s12145-019-00387-y
[29] Puttinaovarat, S., Horkaew, P., Khaimook, K., Polnigongit, W. (2015). Adaptive hydrological flow field modeling based on water body extraction and surface information. Journal of Applied Remote Sensing, 9(1): 095041. https://doi.org/10.1117/1.JRS.9.095041
[30] Knudby, A., Jupiter, S., Roelfsema, C., Lyons, M., Phinn, S. (2013). Mapping coral reef resilience indicators using field and remotely sensed data. Remote Sensing, 5(3): 1311-1334. https://doi.org/10.3390/rs5031311
[31] Aslam, K., Khalil, R.Z., Zafar, S., Malik, S., Basit, A. (2019). Modeling coral reef susceptibility using GIS multi-criteria analysis. In 2019 Sixth International Conference on Aerospace Science and Engineering (ICASE), pp. 1-5.

APPENDIX A: Publication at the International Journal of Design & Nature and Ecodynamics. (Cont.)

https://doi.org/10.1109/ICASE48783.2019.9059119

[32] Xu, J., Zhao, J., Wang, F., Chen, Y., Lee, Z. (2021). Detection of coral reef bleaching based on sentinel-2 multi-temporal imagery: Simulation and case study. Frontiers in Marine Science, 8: 268. https://doi.org/10.3389/fmars.2021.584263

[33] Levine, A.S., Feinholz, C.L. (2015). Participatory GIS to inform coral reef ecosystem management: Mapping human coastal and ocean uses in Hawaii. Applied Geography, 59: 60-69. https://doi.org/10.1016/j.apgeog.2014.12.004

[34] Almutalabi, S.N., Alzuhairi, M., Hashim, F.A. (2021). Two stages thermal and catalytic cracking of polyethylene terephthalate to fuel production. International Journal of Design & Nature and Ecodynamics, 16(6): 725-732. https://doi.org/10.18280/ijdne.160615

APPENDIX A: Publication at the 37th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC).

An Experimental Study of RSSI for LoRa Technology in Different Bandwidths

1st Tanatpong Udomchaipitak
Faculty of Science and Industrial Technology
Prince of Songkla University, Surat Thani Campus
Suratthani, Thailand
6340320504@psu.ac.th

2nd Nathaphon Boonnam
Faculty of Science and Industrial Technology
Prince of Songkla University, Surat Thani Campus
Suratthani, Thailand
nathaphon.b@psu.ac.th

3rd Supattra Puttinaovarat
Faculty of Science and Industrial Technology
Prince of Songkla University, Surat Thani Campus
Suratthani, Thailand
supattra.p@psu.ac.th

***Abstract*— LoRa technology is applied through a modulation technique for transmitting data at a longer distance than other technologies such as Wi-Fi, Bluetooth, cellular, etc. LoRa is a standard technology for low power technology and many parameters. There are several equations in each parameter for calculation. Therefore, we compared experimental RSSI and computed RSSI in areas with various signal barriers. In our estimation, we found the n-value from the RSSI equation suitable in this area is 4-6. In our experiment, we computed the percentage of error as an indicator to find the error of both RSSI with SF12 and n=5. We found the percentage of error value of the experiment at 9.56% and 8.63% as BW125 and BW250, respectively.**

Keywords— LoRa technology, Received signal strength indicator, Bandwidth, Percentage error

I. Introduction

LoRa technology is now considered a technology with low-frequency transmission capabilities. Thus, it has the longest transmission distance and the lowest power consumption compared to other technologies such as Wi-Fi or Bluetooth. LoRa has been applied to many IoT applications [1]. Because of the low frequency of LoRa signal transmission, the distance between the sender node and the receiver node may be interfered with during signal transmission. This impediment may reduce the signal strength at the receiver node itself. It is attenuated due to the absorption of the obstacles transmitted by the LoRa signal [2,3].

Many researchers can apply LoRa technology to a wide variety of applications, whether it is the application of LoRa in the work of tracking sailing ships in the coastal area. The experiment found that some blind spots prevented the transmitted signal from reaching the receiver node [4]. Alternatively, the application of LoRa technology collects parameters in the sea and gathers them back to the shore [5]. In addition to various IoT applications, there have been experiments on topology. The suitable LoRa topology is usually of the star type. However, the application of LoRa to mesh topology is to study how LoRa works in such a topology. Every node is defined as both a receiver and a sender on its own [5,6]. The signal quality is determined by received signal strength indicators (RSSI). There are many things causing the RSSI to be reduced. Whether it is the natural environment or buildings, it is surely encountered for each type with a different signal absorption rate [2,3].

This paper studies the RSSI values in LoRa technology using two parameters to determine their tolerances between the calculated RSSI (CRSSI) and the experimental RSSI (ERSSI). The experiments were conducted at various signal barriers and in highlands near hills, which were particularly difficult for LoRa technology, with the factory providing essential experimental equipment. The first author can perform this practical task with financial support at Prince of Songkla University, Surat Thani Campus. We also would like to express our thanks for the reinforcement in this experiment.

II. Related work

The main parameters of LoRa are bandwidth (BW), spreading factor (SF), and code rate (CR). The transmission distance and the transmission time [6,7]. In LoRa, we can calculate the main parameter or use the main parameter to calculate the other parameters as follows: in (1), SF determination using n_{chirps}, which is the number of chirps per symbol, and the emblem contain 2^{SF}chirps. The chirp rate is usually the same as the BW. We can use it to determine the duration of the symbol, as in (2). For CR, we can calculate the rate. As in (3), and when taking all three of them into mathematical operations, we can obtain the bit rate from (4) [7,8]. In addition to the parameters used in the settings, another parameter can be considered a parameter that gives a value and can be interpreted in various ways, for example, signal strength or the distance between the receiver node and the sender node [9], namely the received signal strength indicator (RSSI). We can calculate the RSSI value from (5)

$$SF = \log_2 n_{chirps} \tag{1}$$

$$T_s = \frac{2^{SF}}{BW} \tag{2}$$

$$CR = \frac{4}{4+n}, n \in \{1,2,3,4\} \tag{3}$$

$$R_b = SF \frac{BW}{2^{SF}} CR \tag{4}$$

$$RSSI = -(10n \log_{10} d - A) \tag{5}$$

where $A = 1$ represents the power received in dBm when the distance between the transmitter and receiver antennas is measured in meters, d the distance between the sender and receiver nodes, and n the environment's loss parameter or loss exponent. In general, the loss parameter is a value resulting from the barriers. They come together with the transmission distance that greatly influences signal quality. The impact in each work area has different effects due to other obstacles or terrain [9,10].

978-1-5386-5541-2/18/$31.00 ©2018 IEEE

APPENDIX A: Publication at the 37th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC). (Cont.)

In the LoRa experiment [3], they studied LoRa transmission in the hill area. As an experiment, the same, but their antenna size is adjusted more than ours. As a result, the LoRa can transmit far and well in our experiments, and the equipment may be more expensive due to changing the antenna [11].

III. MATERIALS AND METHODS

A. Experiment Area

Our experiment is a task in the highlands near the hills; characteristics of the area were used in the experiment. They were covered by nature and buildings, as shown in Fig. 1. It was a massive challenge in our test using all the base equipment provided by the factory.

Fig. 1. Experiment area.

We tested in a highland area of a slight hill interspersed with the presence of high buildings. It results in data transmission over longer distances, which may lead to a sharp drop in RSSI values. We have also added to the challenge of this experiment by testing the transmission distance in LoRa by transmitting the values in a terrestrial way. We do not customize the antennas; all are factory equipment. It makes for a highly challenging experiment for our study.

B. Experiment an equipment

In this experiment, we used a commercially available microcontroller board. The device was already equipped with a LoRa chip for LoRa operation as the TTGO LoRa32 OLED board as shown in Fig. 2.

Fig. 2. TTGO LoRa32 OLED.

Semtech TTGO LoRa32 SX1276 OLED is a simple LoRa device with customizable capabilities to support both receiver and sender node functions. This development leads to reduced signal interference as well as signal loss prevention. The board also contains other communication technologies such as Wi-Fi, Bluetooth, etc. In addition, there is a 0.96-inch OLED display on board and the central processing unit of the board. It uses the ESP32 chip to process data with other devices on the board. Overall, the board can operate in various climatic regions around the world and support regional frequencies [12].

C. Settings Parameter

To set the parameters for this experiment, we have different parameters through the LoRa library of the Arduino IDE program, a library for transmitting devices that use LoRa signals. It can be supported by various Semtech chips such as SX1276/77/78/79 [13]. In our settings, we used the parameters and equipment in this experiment. We set the SF to 12 to test the most extended transmission using the factory base equipment. The highest SF is required. The limitation on the BW part is only available to two values, 125 and 250 kHz, so we can only perform two experiments. At the same time, the frequency of 915 MHz was used in this experiment to correspond with the frequency setting from the factory and keep LoRa working at its best. For CR, we set the value to 4/5 to be suitable for the LoRa wireless area network protocol and in accordance with the LoRa recommendations alliance. Then, in the settings in the Arduino IDE, we adjusted the value at the sender node to high power at 20 mW to find the best performance for the alliance [14]. Finally, LoRa is to install a factory-received antenna with an antenna size of 1 dBi. After setting up the device, we studied whether an RSSI trend would occur by using polynomial regression to analyze the RSSI if we adjusted the device to be more efficient.

In our experiment, we located the receiver node on the ground floor (0 meter high) of the building. We mounted the sender node to the vehicle to move away from the receiver node, stopping at a predetermined point along the path through it to observe and note the RSSI values obtained at each end. We perform periodic movements and pauses to record the ERSSI value (every 100 m) due to we expect the ERSSI to change significantly in our experimental area and with the limitations of the experimental area covered by natural barriers, buildings, and hills, LoRa transmits ground-level data. cannot transmit signals over long distances. And then, our experiment loop continues until the signal from the receiver node and sender node is disconnected, thus pausing each experiment. In addition, we calculate the RSSI value in (5) showing the possibility that will occur in this experiment. We repeated it in the same as the actual experiment by substituting the distance values into the calculation formula. Table I shows the recorded results obtaining from this calculation with respect to the distance.

TABLE I. CRSSI RESULT

Distance	1	100	200	300	400	500
CRSSI	13	-69	-79	-86	-91	-94

We obtained the result from the calculation formula by substituting the variables within the equation using $n = 4$. Due to the nature of the experimental area, it is part of the obstacles in the building [15]. The variable A is given a value of 13, which is a conversion from (6). It changes from mW to dBm [16], and the distance projected in this section is set at 500 m. Based on this calculation formula, we get the power emission to happen when the receiver and sender are connected.

APPENDIX A: Publication at the 37th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC). (Cont.)

$$P_{(dBm)} = -10\log_{10}\frac{P_{(mW)}}{1mW}$$

The output in dBm is obtained, and the result is used in the CRSSI. After we finished computing CRSSI from (5), the results were compared with the collected ERSSI to determine the discrepancy between the two results.

After we collected the RSSI values, which were obtained from both experiments, we hence computed the difference between the two outputs in the form of

$$\%erorr = \frac{|ERSSI - CRSSI|}{ERSSI}100$$

We can find the percentage error of both RSSI values in each range by positively difference of $ERSSI$ and $CRSSI$ value and dividing by the $ERSSI$.

IV. Result

In an actual experiment to analyze the RSSI values, we collected a total of two samples. Each cycle changes one key parameter, BW, to study the differences between BW. Fig. 3 shows the average results obtained from the actual experiment conducted in the figures of the graphs and polynomial regression. The polynomial graph is derived from the equations obtained from both experiments, through polynomial regression, to use the equation to find trends in the RSSI values from (8) and (9):

$$y_{BW125} = 0.0016x^2 - 0.4793x - 75.77 - 2x^{(-6)}, \quad (8)$$

$$y_{BW250} = 0.0016x^2 - 0.4889x - 75.48 - 2x^{(-6)}. \quad (9)$$

The equation is obtained for polynomial regression for determining RSSI from both BW, with (8) being the equation at BW=125 and (9) being the equation at BW=250, respectively. Both equations are derived from polynomial regression and compared with other RSSI values as shown in Fig. 3.

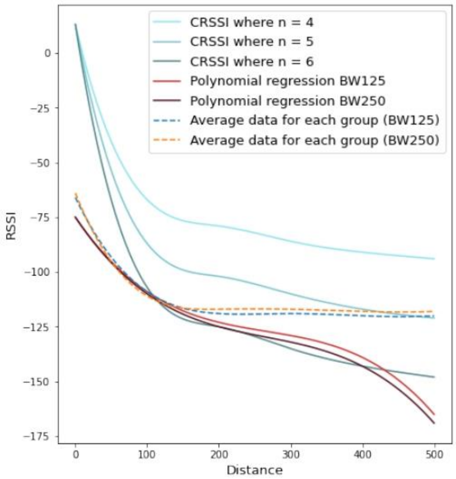


Fig. 3. Results from calculations compared with real experiments.

Fig. 3 shows a sharp decline in the RSSI value as the movement through the signal obstacle is more likely to result in a range of 0-100. Although the graph obtained from the results of the RSSI formula has not yet been created, it is immediately apparent that the $n = 4$ representation is not suitable for use in this experiment. Values in the range of 4-6 are the lost parameters to find the n value closest to the ERSSI. After seeing the most comparable RSSI value, we calculated the RSSI error in the following order, represented by the light blue line in Fig. 3. It depicts the likelihood that the $n = 5$ values are closest to the experimentally obtained values of the three n-values and both dashed lines obtained from ERSSI; for each line, we average all of the data and draw conclusions from each line. Additionally, both red lines tend to occur if obstacles increase the signal between the receiver and sender node while it can still transmit data. We obtained these results from polynomial regression statistical calculations to study the effects of RSSI values. In this case, both nodes can transmit data to each other. There are more obstacles than this experiment, and the results likely follow the trend of polynomial regression. We collected the RSSI values from both the experimental and the calculated ones. The RSSI values from both sources were used to calculate the error values from (7) as shown in Table II.

TABLE II. RSSI ERROR RESULT

RSSI error (%) depend on distance (m.)								
BW	n	100	200	300	400	500	AVG	SD
125	4	40	34	28	24	21	29.40	6.86
	5	22	15	8	2	0.8	9.56	8.01
	6	4	4	12	19	23	12.40	7.71
250	4	39	32	27	22	20	28.00	6.90
	5	21	12	6	0.8	2	8.63	7.43
	6	3	6	14	21	25	13.80	8.42

The experiment results demonstrate the superior $n = 5$ for this study, as there is a minimum data error of 8.36–9.56% over the BW 250 and BW 125 ranges, respectively. But if we compare ERSSI with the RSSI polynomial result by equation (7) as shown in Table III.

TABLE III. RSSI ERROR RESULT (COMPARED POLYNOMIAL)

BW	100	200	300	400	500	AVG	SD
125	2.6	2.5	7.5	15.8	37.5	13.18	13.09
250	0.9	3.3	11.8	21.1	43.2	16.06	15.32

Although the mean error was lower than in Table II, the calculated results of (8) and (9) were derived from the ERSSI data collected through polynomial regression. A similar study area or experiment may use these above equations to calculate RSSI. But the disadvantage of such polynomial regression equations is that results with longer distances may cause many more errors.

APPENDIX A: Publication at the 37th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC). (Cont.)

V. CONCLUSION AND DISCUSSION

In this experiment, we found an error between the CRSSI and ERSSI values with polynomial regression and experimentation to determine the n -values or equations suitable for calculating the RSSI values in a region similar to our experimental result. From this experiment, we can conclude that the CRSSI value in such a distinct area, the equation (5) must be calculated using $n = 5$. On the other hand, although the error range of 0-200 m, $n = 6$ is better, overall LoRa is a remarkable technique for transmitting long-distance signals. Therefore, as a whole, $n = 5$ is still a good n value for this experiment. Our investigation found many obstacles to signal and did not extend the antenna more, making it non-transmit so far. Therefore, LoRa cannot transmit data in this experiment at a longer distance than in the other articles. As well as [3], we found a similar problem with the most significant obstacle, which is many buildings. It creates a signal that cannot be transmitted over a long distance. Because of the equipment installed at ground level, the signal was attenuated when passing through an obstacle. This is different from installing the device in a high place to transmit better signals. In the case of this experiment, we can study the selection of variables in the equation of each parameter to find the percentage of error arising from actual experiments and values obtained from RSSI calculations only. More generally, we can collect more data from the investigation to synthesize the system of equations more precisely. For future work, we will expect to use the regression of signals to find the occurrence of an RSSI, which may give better results for the upcoming RSSI forecast.

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.

[2] A. Shrivastava. "Chapter 3 : Plastic Properties and Testing," Introduction to Plastics Engineering, pp. 49-110, 2018.

[3] M. Bilguunmaa, R. Miura, F. Ono, T. Kagawa, L. Shan, T. Nakayama, F. Kojima, and B. Choijil. "Performance Evaluation of LoRa 920 MHz Frequency Band in a Hilly Forested Area," Electronics, vol. 10, no. 4, pp. 1-12, 2021.

[4] S. I. Ramon, I. G. Liaño, C. Simoes, E. Couñago, and A. F. Skarmeta. "Tracking and monitoring system based on LoRa technology for lightweight boats," Electronics, vol. 8, no. 1, pp. 1-18, 2019.

[5] N. Arago, E. Galido, M. V. C. Padilla, L. M. Tulin, G. C. Virrey, and J. Velasco. "Long Range Communication Technology for Weather Buoy," International Journal of Emerging Trends in Engineering Research, vol. 8, pp. 4399-4404, 2020.

[6] B. Riccardo, P. Napoletano, and M. Savi. "A LoRa-Based Mesh Network for Peer-to-Peer Long-Range Communication," Sensors, vol. 21, no. 13, pp. 1-12, 2021.

[7] G. Emanuele, L. Prando, A. Vizziello, P. Savazzi, and P. Gamba. "Experimental data set analysis of RSSI - based indoor and outdoor localization in LoRa networks," Internet Technology Letters, vol. 2, no. 1, pp. 1-6, 2019.

[8] A. Lavric. "LoRa (long-range) high-density sensors for internet of things," Journal of Sensors, vol. 2019, pp. 1-9, 2019.

[9] K. H. Lam, C. C. Cheung, and W. C. Lee, "RSSI-based LoRa localization systems for large-scale indoor and outdoor environments," IEEE Transactions on Vehicular Technology, IEEE, pp. 1-13, 2019.

[10] R. F. Malik, A. R. Perdana, R. Zulfahmi, and H. Ubaya. "Dual-Channel LoRa Gateway using Channel Assignment on Raspberry Pi," Journal of Physics: Conference Series , pp. 1-9, 2019.

[11] N. Jovalekic, V. Drndarevic, E. Pietrosemoli, I. Darby, and M. Zennaro. (2018). "Experimental Study of LoRa Transmission over Seawater," Sensor, pp. 333-351, 2018.

[12] "TTGO LoRa32 SX1276 OLED Board with Arduino IDE," microcontrollerslab, Available: https://microcontrollerslab.com/ttgo-lora32-sx1276-oled-board-pinout-getting-started-with-arduino-ide/. [Accessed December 27, 2020].

[13] S. Mistry, "LoRa," Arduino, Available: https://www.arduino.cc/reference/en/libraries/lora/. [Accessed December 27, 2020].

[14] LoRa-developers,"Documentaton," Available: https://lora-developers.semtech.com/documentation/tech-papers-and-guides/lora-and-lorawan/. [Accessed December 27, 2020].

[15] M. Jorge, R. Abrishambaf, T. Gomes, P. Gonçalves, J. Cabral, A. Tavares, and J. Monteiro. "Path loss exponent analysis in wireless sensor networks: Experimental evaluation," 2013 11th IEEE international conference on industrial informatics (INDIN), IEEE, pp. 54-58, 2013.

[16] U. Keller and L. Gallmann. "Ultrafast Laser Physics," Avilable: http://www. ulp. ethz. ch/education/ultrafastlaserphysics/3_Dispersion_compensation. pdf, [Accessed January 15, 2021].

VITAE

Name Mr.Tanantpong Udomchaipitak

Student ID 6340320504

Educational Attainment

Degree	Name of Institution	Year of Graduation
B.Sc (Information Technology)	Prince of Songkla University	2019

Scholarship Awards during Enrolment

1. Graduate tuition research fund fiscal year 2020 from Prince of Songkla University

.

List of Publication and Proceeding

Nathaphon Boonnam, Tanatpong Udomchaipitak, Supattra Puttinaovarat, Thanapong Chaichana, Veera Boonjing, and Jirapond Muangprathub. (2022). Coral Reef Bleaching under Climate Change: Prediction Modeling and Machine Learning, MDPI (Sensor), vol.14, no.10, pp.1-13.

Tanatpong Udomchaipitak, Nathaphon Boonnam, Supattra Puttinaovarat, and Paramate Horkaew. (2022). Forecast Coral Bleaching by Machine Learnings of Remotely Sensed Geospatial Data, International Journal of Design & Nature and Ecodynamics, vol.17, no.3, pp.423-431.

Tanatpong Udomchaipitak, Nathaphon Boonnam, and Supattra Puttinaovarat. (2022). An Experimental Study of RSSI for LoRa Technology in Different Bandwidths, The 37th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC) at Phuket, Thailand, pp.728-731.