Forecast Coral Bleaching by Machine Learnings of Remotely Sensed Geospatial Data

Tanatpong Udomchaipitak¹, Nathaphon Boonnam¹, Supattra Puttinaovarat¹,², Paramate Horkaew²

¹Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Surat Thani 84000, Thailand
²School 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 tourism. Environmental agencies in affected regions have 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 tourism [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.
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 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
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.

![Spatial Data Query](image1)

**Spatial Data Query**

- Geospatial Cloud Database
- Geospatial Data Query
- Layers Map
  - SST
  - Water turbidity
  - Salinity
  - Wind speed

![Corral Bleaching Forecast](image2)

**Corral Bleaching Forecast**

- Data pre-processing
- Training and Validation
- Accuracy Assessment

![Visualization of Forecast Results](image3)

**Visualization of Forecast Results**

- Coral Bleaching Forecast Map
- Geospatial Data Visualization

![Figure 1. Conceptual diagram of the proposed system](image4)

**Figure 1.** Conceptual diagram of the proposed system

![Figure 2. Examples of sea surface temperature (a), water turbidity (b), and wind speed (c)](image5)

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

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Data Type</th>
<th>Source</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sea Surface Temperature (SST)</td>
<td>Spatial</td>
<td>GEE (MODIS and NOAA)</td>
<td>Summer, 2020</td>
</tr>
<tr>
<td>Water-Turbidity</td>
<td>Spatial</td>
<td>GEE (Landsat 8)</td>
<td>Summer, 2020</td>
</tr>
<tr>
<td>Salinity</td>
<td>Spatial</td>
<td>GEE (HYCOM: Hybrid Coordinate Ocean Model)</td>
<td>Summer, 2020</td>
</tr>
<tr>
<td>Wind Speed</td>
<td>Spatial</td>
<td>GEE (NOAA CDR)</td>
<td>Summer, 2020</td>
</tr>
</tbody>
</table>

**Table 1.** Detailed characteristics of spatial data and their sources

**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
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>
<thead>
<tr>
<th>Machine Learning</th>
<th>Accuracy</th>
<th>RMSE</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN (MLP)</td>
<td>88.07</td>
<td>0.21</td>
<td>0.85</td>
</tr>
<tr>
<td>Random Forest</td>
<td>97.25</td>
<td>0.08</td>
<td>0.96</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>94.49</td>
<td>0.14</td>
<td>0.93</td>
</tr>
<tr>
<td>ANN (RBF)</td>
<td>90.83</td>
<td>0.18</td>
<td>0.88</td>
</tr>
</tbody>
</table>

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

Figure 3. Reference coral bleaching data obtained from DMCR, showing the entire study area (a) and four selected islands (b), (c) and (d)
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 listed 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](image-url)  
Figure 4. Comparison between different ML algorithms with respect to overall accuracy (a), RMSE (b), and Kappa (c)
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.
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


