

Cloud component removal for shallow water depth retrieval with multi-spectral satellite imagery

Po-Yao Tsou, Peter Tian-Yuan Shih*, and Tee-Ann Teo

Department of Civil Engineering, National Chiao Tung University, Hsinchu, Taiwan

Article history:

Received 16 April 2018

Revised 10 August 2018

Accepted 10 September 2018

Keywords:

Linear spectral un-mixing, Optical properties, Bathymetry

Citation:

Tsou, P.-Y., P. T.-Y. Shih, and T.-A. Teo, 2019: Cloud component removal for shallow water depth retrieval with multi-spectral satellite imagery. *Terr. Atmos. Ocean. Sci.*, 30, 467-480, doi: 10.3319/TAO.2018.09.10.01

ABSTRACT

Water depth and the topography under water provide important information for near shore human activities. With the intensification of international territory concerns, bathymetric mapping is also gaining attention. Optical satellite imagery is an efficient tool for estimating shallow water depth as compared to the traditional field surveying because of its wide coverage area. On the other hand, the existence of cloud and haze contaminates the spectral signatures, which introduces errors to the depth data retrieved. In this research, the contaminated pixels are treated as a mixture of water and cloud. Linear Spectral Unmixing (LSU) procedure is applied for estimating the cloud abundance in mixed pixels. The cloud component is then removed with a linear function and the “purified” water component used for depth retrieval. In this research, water depth is estimated with two methods, namely, artificial neural network (ANN) and physical model. The former demands in-situ bathymetric samples for training, the latter requires site information of inherent optical properties (IOPs) and apparent optical properties (AOPs). The experiments reveal that retrieving depth with ANN generates better results than the physical model, but with a few extremely large errors. As for mixed pixels, the error of depth estimation becomes higher as cloud abundance increases. The precision of depth retrieval is higher for mixed pixels at the reef flat level (within 10 m in depth) than those in the lagoon (about 20 m in depth), and the precision generally agrees with those retrieved from water pixels without cloud or haze.

1. INTRODUCTION

The water depth and underwater topography are essential geospatial information for coastal engineering, aquaculture, fishing industry, and other marine applications. Traditionally, the water depth requires on-site surveying such as shipborne sonar (Sound Navigation and Ranging) and airborne bathymetry lidar (Light Detection and Ranging). The shipborne sonar are classified into single-beam and multi-beam echo sounders. The principle of sonar is to measure the water depth by transmitting acoustic waves from transmitter to seafloor. On the other hand, the bathymetry lidar determines the water depth from green and near-infrared lasers. These two lasers are used for collecting waveform features from sea floor and sea surface, respectively. Then, the waveforms are analyzed for deriving the sea-surface and/or

seafloor topography. Instead of active sensor, another approach is to use passive remote sensing for the determination of water depth, for example, optical satellite image for shallow water. The idea is to utilize the characteristics of attenuation of visible light through water. Therefore, the attenuation of visible light is a function of the properties of water column in which light passing through. To compare these three approaches, the measurement of sonar is deeper than two other approaches. But shipborne sonar is usually difficult to operate in shallow water region because of potential navigation risk and narrow swath. For shallow water region, the airborne bathymetry lidar system is more flexible than shipborne sonar system as the aircraft acquire data effectively in a flying platform. As the water transparency is the major constraint for laser and optical image in water depth measurements. Hence, these two systems are usually applied on coastal zone.

Retrieving water depth from satellite optical images

* Corresponding author
E-mail: tyshih@mail.nctu.edu.tw

features means a large coverage at a lower cost. Although disadvantages of this method include lower accuracy and a limited depth range, valuable information for understanding the sea bottom topography that is likely complimentary to the traditional acoustic sounding techniques can still be gathered. For these optical images, cloud and haze are common sources of disturbance. This study investigates the mitigation of haze disturbance to the depth retrieval with linear un-mixing schemes.

The relationship between depth and the radiance observed has been long established. The approaches to observing this relationship include using single band and band ratio. Polcyn et al. (1970) reported the ratio between two bands decreases with the increase of water depth. Lyzenga (1978) proposed depth retrieval algorithms based on Beer's Law with single band and applied them to airborne multispectral imagery. Lyzenga (1981, 1985), Stumpf et al. (2003), and Kao et al. (2009) are examples of much related research along this line. Neural network algorithms are also applied in the depth retrieval modelling, Sandidge and Holyer (1998) applied to hyperspectral image. Ceyhun and Yalçın (2010) applied MLP neural network for modeling ASTER and QuickBird images with depth. Gholamalifard et al. (2013) applied Back-Propagation neural network for Landsat 5 TM with depth. Moses et al. (2013) and Mohamed et al. (2015) are other examples on using neural network. In this study, both the neural network and semi-analytical modeling schemes are applied.

2. MATERIAL

The Dongsha Atoll is located in the north of South China Sea. It is one of the national parks of Taiwan, which provides rich bio-diversities of marine life. The geographic location of Dongsha Atoll is between 20°35' to 47' north latitude and 116° 42' to 55' east longitude. It is about 445 km from Kaohsiung. The shape of Dongsha Atoll is like a full moon, and the diameter of the ring-shaped atoll is about 25 km. The reef crest is about 46 km long and 2 km wide. The greatest depth of the lagoon reaches 20 m in the center of the atoll. Because the water clarity is high and the location is remote, retrieving water depth from satellite images presents an economical and effective observation scheme.

DigitalGlobe Inc. launched the WorldView series satellites. Currently, there are four high-resolution satellites in this satellite constellation. WorldView-1 provides a 0.5-m panchromatic image; WorldView-2 provides 0.5-m panchromatic and 2-m multispectral images cover eight different wavelengths; WorldView-3 provides a standard panchromatic, 8-band multispectral, 8-band short-wave infrared (SWIR), and 12-band CAVIS imagery in different spatial resolutions. WorldView-4 collects a panchromatic image at the 0.31-m resolution and 4-band multispectral images at 1.24-m resolution.

A scene of WorldView-2 image (WV-2, Fig. 1a) and a scene of WorldView-3 (WV-3, Fig. 1b) of the area are used. The WV-2 image dimension is 12589 × 11588 pixels, corresponding to 25 km × 25 km on the ground. The image is acquired on July 3rd, 2010. The WV-3 image dimension is 1325 × 1270 pixels, equivalent to 2.65 km × 2.54 km on the ground, acquired on April 5th, 2015. The area of the WV-3 image is located in the red box as shown in Fig. 1. The ground truth is derived from the airborne bathymetric lidar mission with an AHAB HawkEye II (Shih et al. 2011). The survey mission was held in September of 2010 and the grid resolution of the resampled bathymetric model is 8 m. The spatial resolution of panchromatic and multispectral bands are 0.5 and 2 m respectively for WV-2, and 0.3 and 1.2 m for WV-3. As shown in Table 1, there are eight bands. In this study, the standard Blue, Green, Red, and Near-infrared multispectral packages were available and the Blue and Green bands are both utilized for the depth retrieval.

3. METHODOLOGY

The workflow of depth retrieval includes three processes, image pre-processing for converting digital numbers into reflectance, cloud component removal, and finally the depth retrieval. In the first process, ATCOR module in PCI Geomatica 2016 is applied for the reflectance conversion. Meanwhile, the image is resampled into 8 m with pixel aggregation.

The second process is to remove the cloud components based on linear un-mixing. The cloud area is identified first; subsequently, two schemes are applied. The first scheme is a simple one based on applying a threshold to the infrared band. This is based on the fact that water would absorb infrared and result in dark reflectance. The other scheme is utilizing Haze Optimized Transformation (HOT) proposed by Zhang et al. (2002). HOT was developed and validated with Landsat TM imagery. The clear-sky vector and haze vector, which are orthogonal to each other, are derived from TM1 (blue) and TM3 (red) bands. Three thresholds, $\mu + \sigma$, $\mu + 1.5\sigma$, and $\mu + 2\sigma$, are applied, producing three masks for each image. Where μ and σ are the mean and standard deviation of pixel value, respectively. The pixel value could be reflectance of near-infrared or HOT. These three thresholds are simply applied to distinguish cloud and non-cloud pixels, separately. For example, if pixel value is larger than $\mu + \sigma$, then, this pixel is treated as a cloud pixel. The three masks from near-infrared and three masks from HOT were then examined manually in the best mask selection. In the experiments conducted, these masks separated water and non-water area effectively. However, the area with cloud coverage is frequently mixed with white sand on Dongsha Island and the breaking wave. These areas were manually removed.

Linear un-mixing is based on the concept of endmembers. That is, every pixel is either pure or a linear mixing of

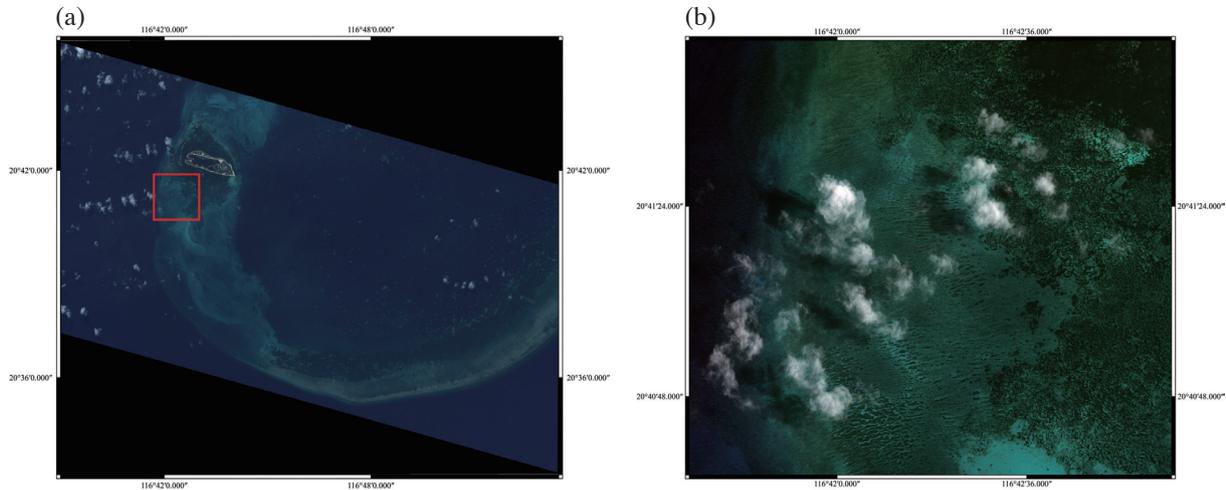


Fig. 1. (a) Dongsha Atoll, WorldView-2, 2017-07-03. (b) Dongsha Atoll, WorldView-3, 2015-04-05.

Table 1. Spectral ranges of WorldView-3 sensors.

Band	Spectral range (nm)
Coastal	450 - 800
Blue	400 - 450
Green	450 - 510
Yellow	510 - 580
Red	585 - 625
RedEdge	630 - 690
NIR1	705 - 745
NIR2	770 - 895

a set of end-members.

$$R_{ij} = E f_{ij} + \varepsilon_{ij} \quad (1)$$

Where R_{ij} is the reflectance at pixel (i, j) , with a dimension of $L \times 1$; E is the spectral feature of endmembers, the dimension is $L \times K$; f_{ij} is the abundance of each endmembers of the pixel, the dimension is $K \times 1$; ε_{ij} is the noise. Huguenin et al. (2004) applied this concept to purifying the pixel for depth retrieval. For the computation of the abundance matrix in un-mixing, the number of end members should be less than the bands. In this study, there are four multispectral bands in the WorldView images obtained. Four endmembers, cloud and three water components with different substrates, are chosen. ISODATA is applied for classifying the water area without cloud components into three classes. The average spectra of each classes are used as the spectra of endmember. With the selected endmembers, the abundance is derived. Then, the cloud component is removed with the following equation,

$$W = [P - f_c C] / (1 - f_c) \quad (2)$$

Where, P is the spectra of the mixed pixel; W is the spectra of water; C is the spectra of cloud; f_c is the fraction of cloud in this pixel.

The third process is the depth retrieval. Two schemes are applied, the back-propagation neural network (BP) and the semi-analytical model. While ground truth is required for the training of BP; optical property parameters are required for the semi-analytical model. BP is a well-established neural network scheme. Gholamalifard et al. (2013) indicated that the artificial neural networks method performing the highest correlation ($r = 0.94$) when compared to principal components analysis and single band algorithms. Besides, the three-layered back-propagation network can approximate most polynomial functions and suitable for the prediction from visible bands to bathymetry. According to the previous study, this study utilized a three-layered back-propagation network. In the input layer, three units correspond to the reflectance of R, G, and B bands. There are six neurons in the hidden layer, and one unit in the output layer also known as the water depth. Log sigmoidal function is used as the transfer in the hidden layer, while linear transfer function is used in the output layer.

The other scheme for depth retrieval is analytical model. The property of both water itself such as water reflectance, attenuation coefficient, and the type of substrate would affect the reflectance collected via remote sensors. Lee et al. (2001) proposed a semi-analytical model for describing the relationship between above surface reflectance, water, and substrate. The model is shown in Eqs. (3) and (4).

$$R_{rs} \approx \frac{0.5r_{rs}}{1 - 1.5r_{rs}} \quad (3)$$

$$r_{rs} \approx r_{rs}^{dp} \left(1 - \exp \left\{ - \left[\frac{1}{\cos(\vartheta_s)} + \frac{D_u^C}{\cos(\vartheta)} \right] kH \right\} \right) + \frac{1}{\pi} \rho \exp \left\{ - \left[\frac{1}{\cos(\vartheta_s)} + \frac{D_u^B}{\cos(\vartheta)} \right] kH \right\} \quad (4)$$

Where, $D_u^C \approx 1.03(1 + 2.4u)^{0.5}$, $D_u^B \approx 1.04(1 + 5.4u)^{0.5}$, and $u = b_b/(a + b_b)$, the meaning of parameters in Eqs. (3) and (4) are shown in Table 2.

Liew and He (2008) proposed a similar but simpler model. While the model of Lee et al. (2001) is developed for hyperspectral images, Liew and He (2008) is designed for the green and red bands of SPOT-5. The attenuation coefficient of deep water in the Liew and He (2008) model is derived with the absorption coefficient and the backscattering coefficient from Lee et al. (2002) model. The unknown parameters are reduced to water depth and the bottom reflectance, and solved with two bands. Because the penetration capability of the red band is limited, Liew and He (2008) model is targeted for water depths of less than 3 m. This study followed the process of Liew and He (2008) but utilizes the more detailed function from Lee et al. (2001), i.e., Eqs. (3) and (4).

Blue and green bands from WorldView are used as observation material. Water surface reflectance R_{rs} is derived through application of atmospheric correction. Deep-water reflectance r_{rs}^{dp} is obtained from the image. Zenith angle of the sun ϑ_w and viewing angle of sensors ϑ are obtained from the satellite image metadata. The attenuation coefficients of blue and green light is obtained from IOCCG (International Ocean Color Coordination Group) dataset (IOCCG 2006).

The semi-analytical model to determine water depth is shown in Eq. (3). The subsurface water reflectance r_{rs} in Eq. (3) can be substituted by Eq. (4). All the sensor-dependent parameters (e.g., zenith and viewing angles) derived from metadata and deep-water reflectance from blue and green images are substituted into physical model. The unknown parameters (i.e., water depth H and bottom albedo ρ) can be calculated based on the principle of least squares. The IOCCG data set (a total of 500 points) contains both inherent (IOP) and apparent (AOP) optical properties, including the absorption coefficient of pure seawater a_w , backscattering coefficient of pure seawater b_{bw} , absorption coefficient of phytoplankton pigments a_{ph} , absorption coefficient of detritus/mineral a_{dm} , backscattering coefficient of detritus/mineral b_{bdm} . Spectral wavelength ranges from 400 - 700 nm at steps of 10 nm. This dataset uses phytoplankton concentration $[C]$ as a variable, ranges from 0.03 - 30.0 $\mu\text{g l}^{-1}$, in 20 layers, 25 sets in each layer.

The central wavelength of WorldView blue and green bands is 480 and 545 nm, respectively. The absorption and backscattering coefficients for these two wavelengths are extracted from IOCCG synthetic dataset. Together with the image reflectance, water depth is derived with the semi-

analytical model and evaluated with ground truth. The best coefficients are then selected for the depth retrieval. Due to the consideration of water clarity in Dongsha Atoll, the first 15 layers, including pure seawater, rather than all 20 layers, are computed. The average of 25 sets in each layer is taken as the attenuation coefficient of that layer. The resulting coefficients are listed in Table 3.

The depth derived from the semi-analytical model should be the absolute value theoretically, because the ground truths were used. In practice, systematic biases exist. This bias is then adjusted with land/water boundary lines, where the depth of water should be zero. In this study, this waterline is extracted with NIR band by thresholding and using a Sobel operator for line extraction.

4. THE EXPERIMENTS

With the WorldView images and procedure described above, the experiments on depth retrieval are conducted. Both the semi-analytical model and BP regression approach are performed for validating the effectiveness of the cloud removal procedure.

4.1 Cloud Masking

There are two schemes for masking cloud areas, direct thresholding with NIR and thresholding after HOT. Figure 2 shows the sub-scenes of WV-2 result from the first method and Fig. 3 from the second. Manual interpretation confirmed that direct thresholding with $(\mu + 1.5\sigma)$ and $(\mu + 2\sigma)$ provides better results than the HOT approach. The result from WV-3 scene is the same but less significant. The direct thresholding with NIR image is then selected and is followed by manual editing for refinement such as removing the breaking wave area. The threshold chosen is $(\mu + 1.5\sigma)$ for the least editing required.

4.2 Cloud Component Removal

This study determined four endmembers through the cloud masking and water substrate classification. These four endmembers (i.e., cloud, substrates 1 to 3) are shown as Table 4. This study combines endmembers (E) and mixed reflectance (R) to determine the abundance (f) of each endmembers of the pixel using linear un-mixing model [Eq. (1)]. The cloud abundance of a WV-2 sub-scene is shown in Fig. 4a and WV-3 in Fig. 4b. The highest value is approximately 0.49 (49%) for WV-2 and 1 (100%) for WV-3. The distribution agrees with the visual impressions shown below. After cloud component removal, the resulting images are shown in Fig. 5. Only the clouds inside the atoll, the study area, are processed and removed. The cloud outside the atoll remains. The Cloud removal result from the WV-3 scene is shown in Fig. 6. The disturbance from

cloud components are significantly reduced and the texture of substrates can be clearly observed. Examining carefully, there are still some cloud effects left, particularly on the edge of the cloud where the cloud is thin, but these remnants are not obvious.

4.3 Depth Retrieval with Neural Network

There are about three million grid points in the digital bathymetric model applied for investigation. Random sampling is applied in selecting 1000 grid points for the training. In order to account for the substrate types, the WV-2 image is classified into six classes with ISODATA. The result is shown in Fig. 7.

Plotting the randomly selected sampling points with classes as the axis of abscissa, depth as the vertical axis, Fig. 8 indicates that the samples are reasonably distributed in the combination of substrate types and depth ranges. Substrate 1 has 332 samples, 226 for substrate 2, 134 for substrate 3, 111 for substrate 4, 124 for substrate 5, and 73 for substrate 6. The depth ranges from 0 - 45 m for substrate type 1, 2, 3, which are corresponding to the deeper area inside lagoon. Substrate type 4, 5, and 6 are corresponding to shallower water. The depth is subsequently less than 10 m.

The BP neural network is trained with the depth and WV-2 image reflectance of randomly selected 1000 samples. The constructed model is then applied to estimate depth with both WV-2 and WV-3 images. The training and validation

Table 2. The parameter in the semi-analytical model.

Parameter	Description
R_{rs}	Above-surface remote sensing reflectance
r_{rs}	Subsurface remote sensing reflectance
r_{rs}^{dp}	Remote sensing reflectance for deep water
D_u^C	Optical path-elongation factor for scattered photon from water column
D_u^B	Optical path-elongation factor for scattered photon from water bottom
ϑ_s	Subsurface solar zenith angle
ϑ	Viewing angle
ρ	Bottom albedo
k	Attenuation coefficient
H	Water depth

Table 3. The coefficients of 15 layers.

[C]	480 nm		545 nm	
	Absorption a	Backscatter b_b	Absorption a	Backscatter b_b
0	0.01270000	0.001698900	0.05195000	0.000999345
0.03	0.02089596	0.003225972	0.05431798	0.002257068
0.05	0.02413064	0.003738796	0.05519632	0.002698724
0.07	0.02719744	0.004367980	0.05606174	0.003266780
0.10	0.03379752	0.004752224	0.05902766	0.003589304
0.15	0.03777720	0.005379860	0.06028880	0.004199552
0.20	0.04781844	0.005614964	0.06345974	0.004423208
0.30	0.05771264	0.007199072	0.06704276	0.005842460
0.50	0.08551636	0.009958872	0.07877340	0.008473350
0.70	0.10438128	0.010626892	0.08857622	0.009179984
1.00	0.13120812	0.012861188	0.09501068	0.011408436
1.50	0.19048800	0.017107552	0.12470846	0.015525784
2.00	0.19304640	0.018210316	0.12631864	0.016704478
3.00	0.27081280	0.019985104	0.15801500	0.018702958
5.00	0.34714160	0.028513440	0.19578580	0.027271360

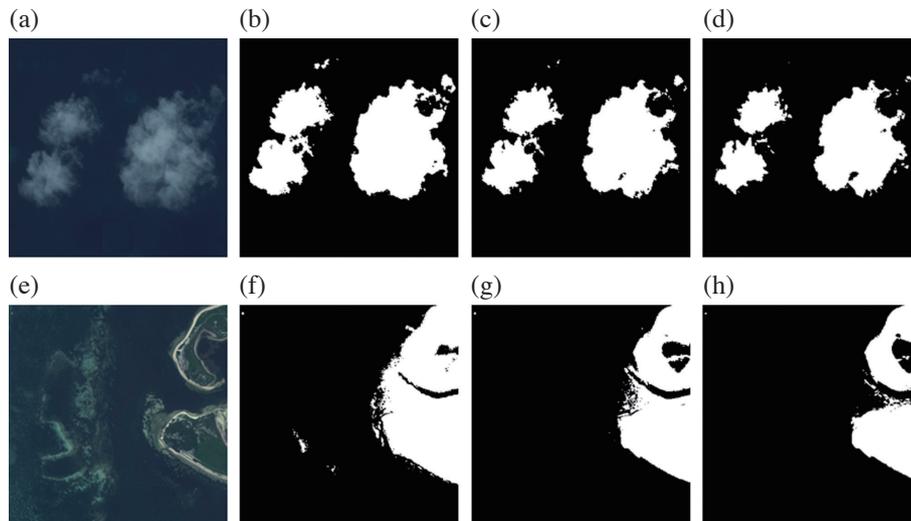


Fig. 2. Cloud mask with direct thresholding, sub-scenes of WV-2, (a) (e) the true color image; (b) (f) NIR ($\mu + \sigma$); (c) (g) NIR ($\mu + 1.5\sigma$); (d) (h) NIR ($\mu + 2\sigma$).

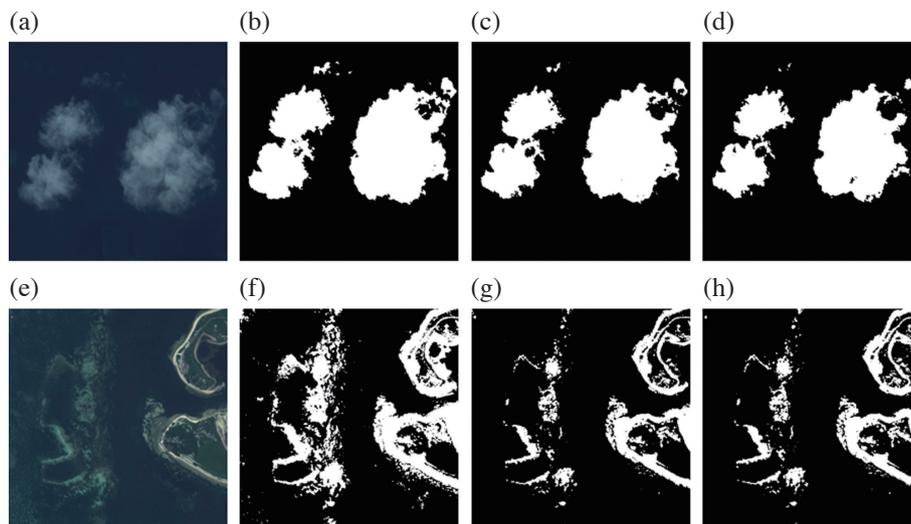


Fig. 3. Cloud mask thresholding after HOTS, sub-scenes of WV-2, (a) (e) the true color image; (b) (f) NIR ($\mu + \sigma$); (c) (g) NIR ($\mu + 1.5\sigma$); (d) (h) NIR ($\mu + 2\sigma$).

Table 4. The endmembers E for linear un-mixing.

		Band 1 (478.3 nm)	Band 2 (545.7 nm)	Band 3 (658.8 nm)	Band 4 (832.9 nm)
WV2	Cloud	0.331372	0.310115	0.278355	0.281826
	substrate 1	0.050445	0.039589	0.014943	0.022746
	substrate 2	0.077279	0.072141	0.019162	0.024029
	substrate 3	0.108760	0.105438	0.022221	0.025948
WV3	Cloud	0.559959	0.568018	0.549883	0.523428
	substrate 1	0.038311	0.022660	0.012651	0.021222
	substrate 2	0.057343	0.053939	0.020646	0.023372
	substrate 3	0.092521	0.095148	0.032143	0.026000

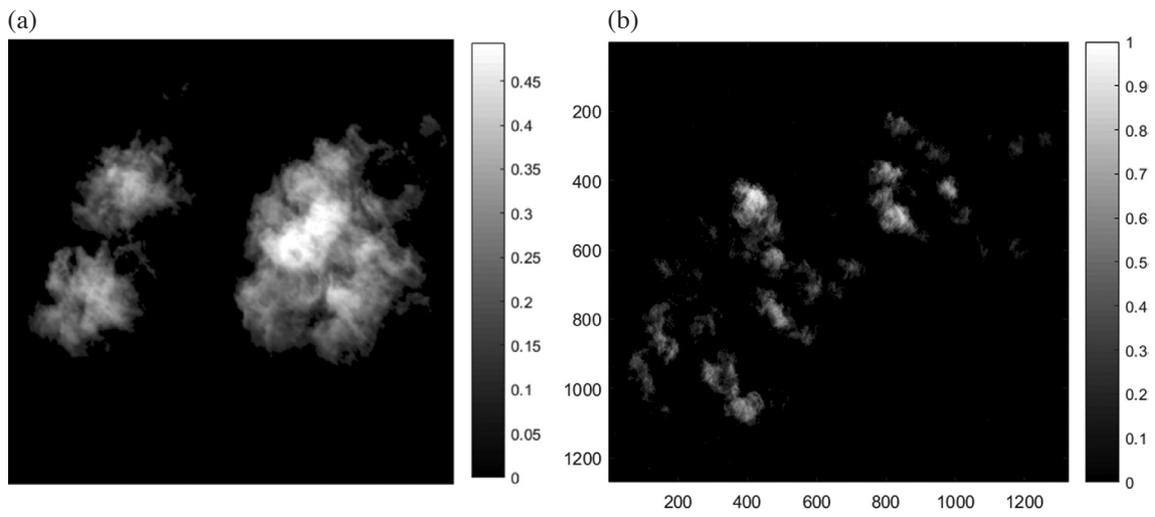


Fig. 4. Cloud abundance, (a) WV-2 sub-scene; (b) WV-3 sub-scene.

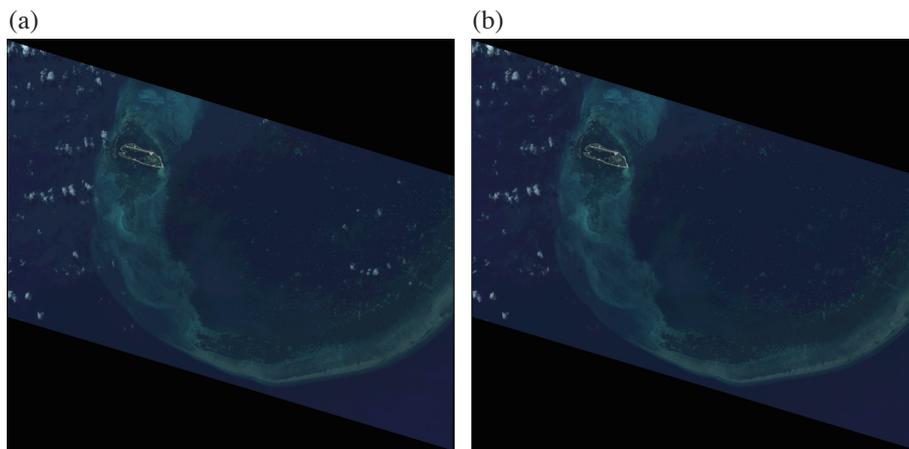


Fig. 5. Cloud removal, WV-2, (a) before; (b) after.

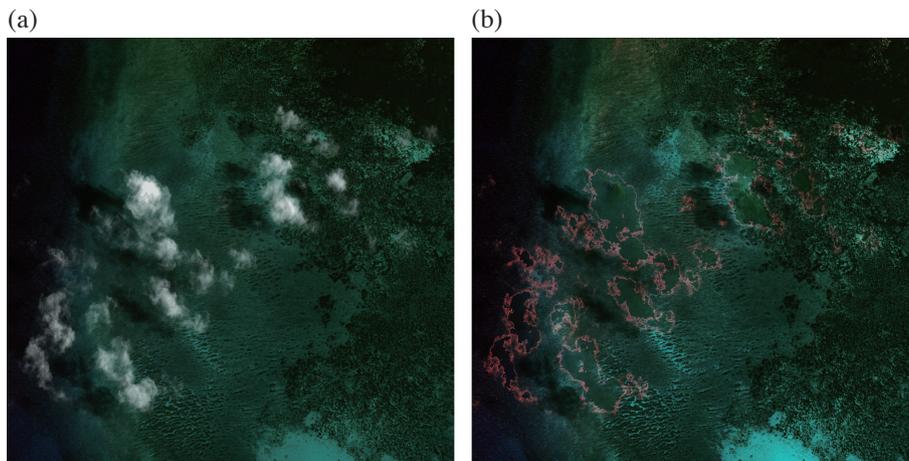


Fig. 6. Cloud removal, Wv-3, (a) before; (b) after.

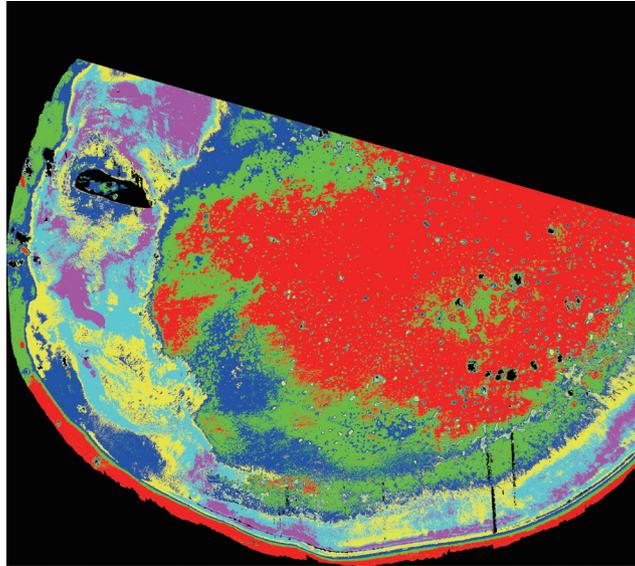


Fig. 7. Substrate classification with ISODATA, WV-2.

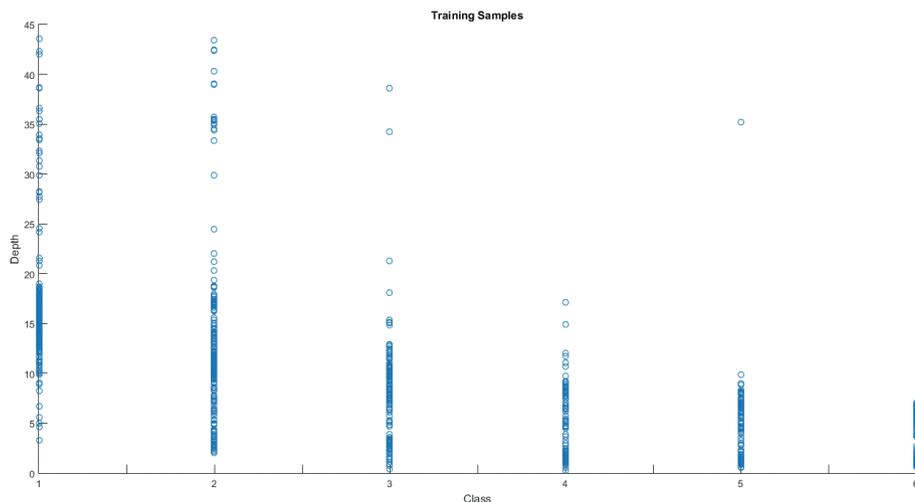


Fig. 8. The distribution of the depth and substrate type training samples.

performance is shown in Fig. 9. Best performance reached at the 30th epoch with an MSE value of 6.5 m. Among the 1000 samples, 80% were used as training subset. The correlation R between output and target is 0.92869. 10% were used for validation with a correlation of 0.9293. 10% were used for testing with a correlation of 0.94401. The correlation of the complete sample set is 0.93106.

After the cloud component removal, the estimated water depth from WV-2 and its discrepancy with the ground truth is shown in Fig. 10. The error indices derived for the complete image, including cloud influenced area, are, average error is 0.21 m, standard deviation 2.79 m, RMSE 2.80 m, extreme errors are 150.69 and -77.70 m. This data is clearly normally distributed, as seen from the histogram of errors. As a concluding remark, the overall accuracy is better than

3 m, but with large outliers. The error indices derived for the pure water area and the cloud-influenced area are summarized in Tables 5 and 6, respectively.

From Table 5, the positive correlation between the RMSE of depth estimation and the depth of the water is quite clear. When the depth exceeds 30 m, the mean error is above -8 m and no longer meaningful. This generally agrees with the penetration capability of light in water. For depth 0 - 15 m, the RMSE is 2.00 m, equivalent to 13.3% of the depth. The RMSE is 2.81 m with a ratio of 9.36% for the 15 - 30 m range.

The statistics for cloud-influenced areas is summarized in Table 6. In general, the RMSE increases with the cloud abundance. The exception is the 0 - 10% category. This can be explained by the depth range that reaches a height of

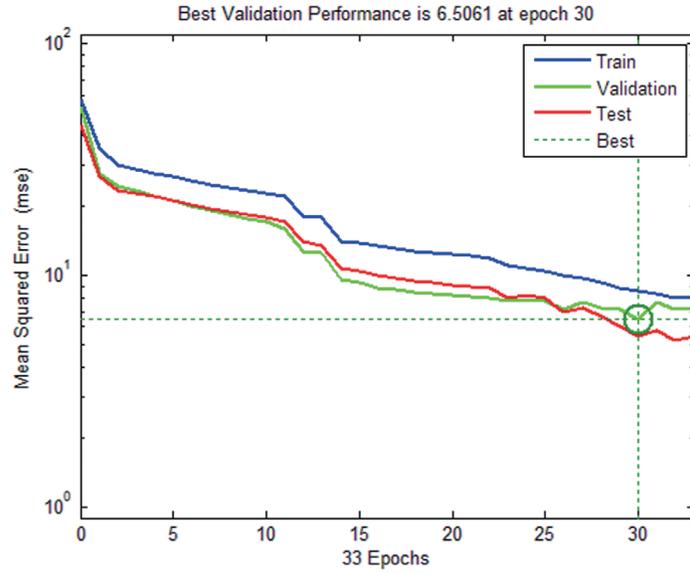


Fig. 9. MSE curve of training.

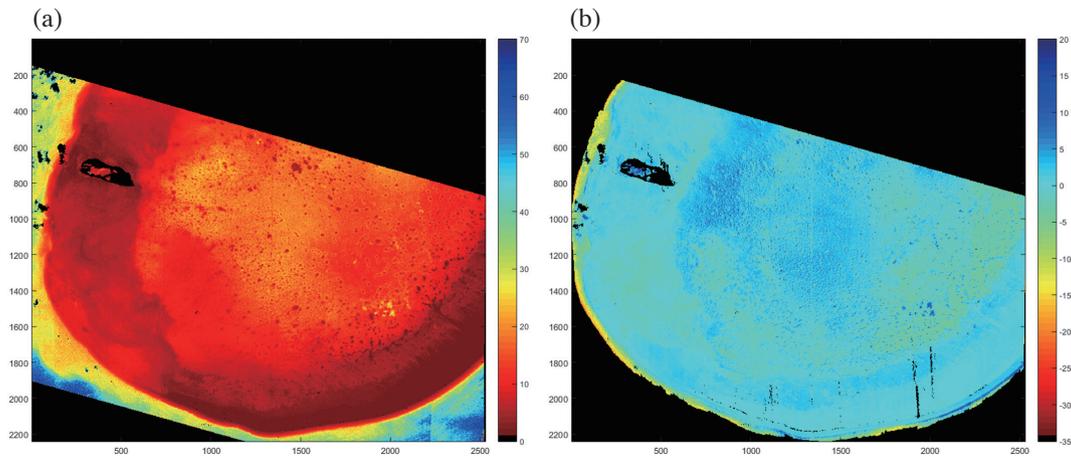


Fig. 10. WV-2 result from BP, (a) depth retrieved; (b) error.

Table 5. Error indices of pure water area, BP depth estimation, WV-2.

Depth (m)	Mean Error (m)	Stan. Dev. (m)	RMSE (m)	Maximum (m)	Minimum (m)
0 - 15	0.52	1.93	2.00	150.69	-77.70
15 - 30	0.36	2.79	2.81	69.35	-43.76
30 - 48	-8.46	5.13	9.89	40.86	-53.45

Table 6. Error indices of cloud influenced area, BP depth estimation, WV-2.

Cloud (%)	Mean Error (m)	Stan. Dev. (m)	RMSE (m)	Maximum (m)	Minimum (m)	Reference Depth (m)
0 - 10	-0.63	3.01	3.08	14.65	-32.90	0.6 - 43.8
10 - 20	0.72	2.44	2.54	13.38	-7.69	0.7 - 18.5
20 - 30	3.49	3.82	5.18	19.00	-4.92	1.7 - 18.7
30 - 40	5.43	4.87	7.30	23.66	-3.37	1.9 - 18.7
40 - 49	11.69	6.11	13.24	25.03	-2.45	13.0 - 17.8

43.8 m and is the largest of the five categories. The RMSE value reflects that uncertainty of the estimate is a function of both the depth and the cloud abundance. When limiting depth is less than 20 m, the RMSE is 2.82 m for 0 - 20% cloud abundance, which is about 14% of the depth.

The depth retrieved and its error from WV-3 are shown in Fig. 10. The trace of cloud coverage is visible in both, indicating the incompleteness of the cloud removal scheme. Overall, the mean error is 0.16 m, standard deviation 2.79 m, RMSE 2.79 m, error ranges from 41.59 m to -16.43 m. This experiment demonstrated that the retrieval model established with training samples from a WV-2 image could yield depth estimation from another image of WV-3, with RMSE equivalent to those yielded from the original training image.

The error statistics from the pure water area and cloud-influenced area of the WV-3 case are summarized in Tables 7 and 8, respectively. Due to the nature of the site, the water is relatively shallower nearby the Dongsha island. However, the trend that the RMSE is positively correlated to the depth is still clear. For areas with depth ranges from 0 - 15 m, the RMSE is 1.54 m, about 10% of the depth, which is slightly lower than the 13% in WV-2 case. But, this is reversed in the 15 - 30 m depth range.

The influence of depth and cloud abundance to the depth uncertainty is reflected in Table 8. There are much fewer samples in the WV-3 case than the WV-2. In particular, there are only 6 samples for 0 - 10% cloud abundance and 15 samples for the 90 - 100% category. This should be taken into consideration when interpreting the indices. In

comparing the WV-3 case with WV-2, the cloud removal scheme performs better in the WV-2 case than in WV-3. This may result from the site characteristics. In WV-3 scene, the sea bottom of the cloud-influenced area has good reflectance due to the substrate type, which is also an important factor of the depth estimation quality.

4.4 Depth Retrieval with Semi-Analytical Model

In this processing scheme, the reflectance from the blue and green bands of the WorldView images are the input observation. The 15 sets of attenuation coefficients derived from the absorption and backscatter coefficient with the use of IOCCG synthetic dataset, as listed in Table 3, are applied to retrieve the depth. Then, the depth of each set is evaluated with the aid of reference depth obtained from airborne bathymetric lidar. The set of coefficients, which produces the best agreement with ground truth, is selected. This procedure may be avoided if the in situ measurements of the water property parameters are available.

IOCCG dataset uses phytoplankton concentration [C] as a variable, ranges from 0.03 - 30.0 $\mu\text{g l}^{-1}$, in 20 layers, with 25 sets in each layer. Starting with [C] = 0, the depth at waterline is -4.59 m, which is above water. The trend is consistent at [C] = 0.5 with depth of -6.08 m, then reverses at [C] = 5.0 with depth of -0.65. The final retrieved depth is adjusted with waterline constraints. The RMSEs of depth from these 15 sets are all larger than 6 m before applying waterline constraints. With waterline constraints, the RMSE

Table 7. Error indices of pure water area, BP depth estimation, WV-3.

Depth (m)	Mean Error (m)	Stan. Dev. (m)	RMSE (m)	Maximum (m)	Minimum (m)
0 - 15	0.72	1.36	1.54	4.74	-7.13
15 - 30	-3.28	5.83	6.69	36.31	-14.79
30 - 38	-4.07	5.80	7.09	20.53	-16.43

Table 8. Error indices of cloud influenced area, BP depth estimation, WV-3.

Cloud (%)	Mean Error (m)	Stan. Dev. (m)	RMSE (m)	Maximum (m)	Minimum (m)	Reference Depth (m)
0 - 10	-0.77	5.54	5.60	1.89	-12.06	3.5 - 30.0
10 - 20	-0.45	1.95	2.00	25.06	-7.90	3.0 - 25.0
20 - 30	0.11	2.61	2.62	41.59	-4.96	3.0 - 22.4
30 - 40	0.54	3.47	3.51	36.53	-3.02	3.2 - 21.6
40 - 50	0.00	1.93	1.93	38.19	-1.56	3.7 - 19.5
50 - 60	-0.26	1.09	1.13	10.41	-2.22	4.1 - 12.1
60 - 70	-0.84	0.66	1.07	1.55	-2.52	4.3 - 6.6
70 - 80	-1.04	1.20	1.59	3.13	-3.36	4.3 - 6.7
80 - 90	-1.42	1.95	2.42	5.88	-4.10	5.0 - 6.7
90 - 100	-0.44	2.20	2.25	3.97	-3.09	6.2 - 6.6

of $[C] = 0.30$ dataset reduces to 3.88 m, the lowest among the 15. Consequently, the coefficients corresponding to $[C] = 0.30$, i.e., 480 nm: $a = 0.05771264$, $b_b = 0.07199072$ for wavelength 480 nm, blue band; and $a = 0.06704276$, $b_b = 0.05842460$ for 545 nm, green band, are applied in the model for retrieval. The waterline offset is -5.4650 m.

The depth retrieved from this semi-analytical model with waterline constraint and its discrepancy between the reference dataset is shown in Fig. 11. The overall mean error is -0.07 m, standard deviation 3.89 m, RMSE 3.89 m, error range from 19.74 to -34.65 m. Assessing pure water areas and cloud influenced areas separately, the error indices are summarized in Tables 9 and 10. The overall performance of this semi-analytical model based approach is inferior to the one obtained with BP (RMSE 2.80 m). In the 0 - 15 m depth range without cloud influence, RMSE 2.22 m from the semi-analytical model is close to the 2.0 m from BP. The other interesting finding is that the error ranges of the semi-analytical model are smaller than the one with BP. In other words, the depth retrieved with this model has less probability of large errors as compared to the BP approach.

The retrieved depth and error from WV-3 are shown in Fig. 12. The overall mean error is 1.06 m, standard deviation 2.91 m, RMSE 3.10 m, and error ranges from 17.70 to -18.45 m. The performance is also inferior to the BP approach, which has RMSE value 2.79 m. Regarding the cloud removal, the trace of cloud in Fig. 12 is much less noticeable than Fig. 13 produced with the BP approach. Error indices for the pure water and cloud influenced areas are

summarized in Tables 11 and 12. The observations on the performance are consistent with the WV-2 case.

4.5 Discussions

4.5.1 The Comparison of Two Different Approaches in Depth Retrieval

This study applied two methods for water depth retrieval from optical satellite images. The semi-analytical model contains several important physical parameters in depth retrieval while the BP neural network implies non-linear function to model the relationship between reflectance and water depth. The BP neural network highly relies on the quality of input data. The ground truth of this study was acquired by high precision bathymetric lidar. The numerical analysis indicated that the BP neural network provides better overall accuracy than the semi-analytical model. However, BP neural network has larger maximum depth error when compared to semi-analytical method. In the visual analysis, the trend of water depth from semi-analytical model is more stable than BP neural network.

4.5.2 The Effect of Water Depth

The actual water depth is one of the constraints for the optical image in water depth retrieval. The experiment analyzed the accuracies in different depths. The results demonstrated that the deeper water depth induces larger depth error via optical satellite image. To consider the water depth

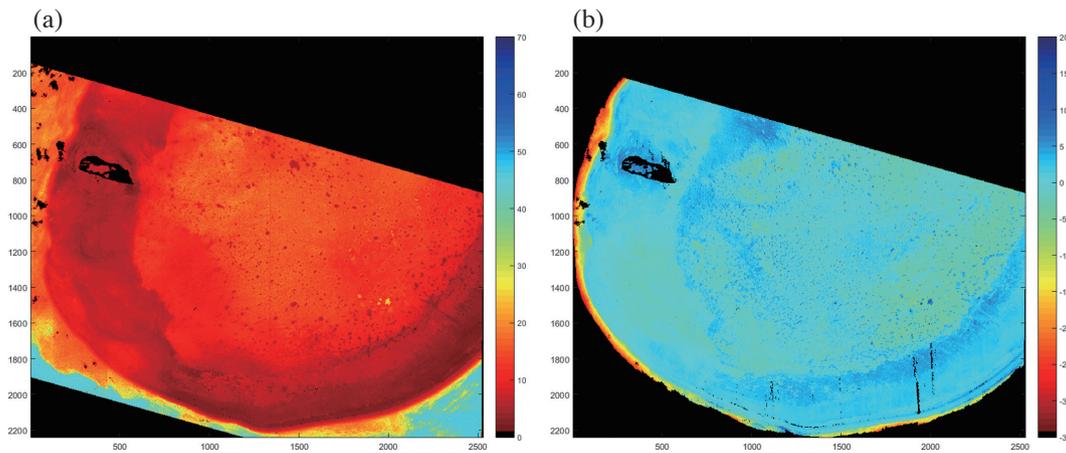


Fig. 11. WV-2 result from model with waterline constraint, (a) depth retrieved; (b) error.

Table 9. Error indices of pure water area, model depth estimation, WV-2.

Depth (m)	Mean Error (m)	Stan. Dev. (m)	RMSE (m)	Maximum (m)	Minimum (m)
0 - 15	1.03	1.97	2.22	11.39	-10.57
15 - 30	-1.59	2.42	2.89	19.74	-22.13
30 - 48	-16.03	6.87	17.44	15.70	-34.65

Table 10. Error indices of cloud influenced area, model depth estimation, WV-2.

Cloud (%)	Mean Error (m)	Stan. Dev. (m)	RMSE (m)	Maximum (m)	Minimum (m)	Reference Depth (m)
0 - 10	-1.49	3.99	4.26	11.58	-31.08	0.6 - 43.8
10 - 20	0.68	3.05	3.13	14.09	-5.90	0.7 - 18.5
20 - 30	4.49	3.57	5.74	18.93	-4.97	1.7 - 18.7
30 - 40	7.25	3.59	8.09	19.47	-2.37	1.9 - 18.7
40 - 49	10.74	2.81	11.16	16.16	2.16	13.0 - 17.8

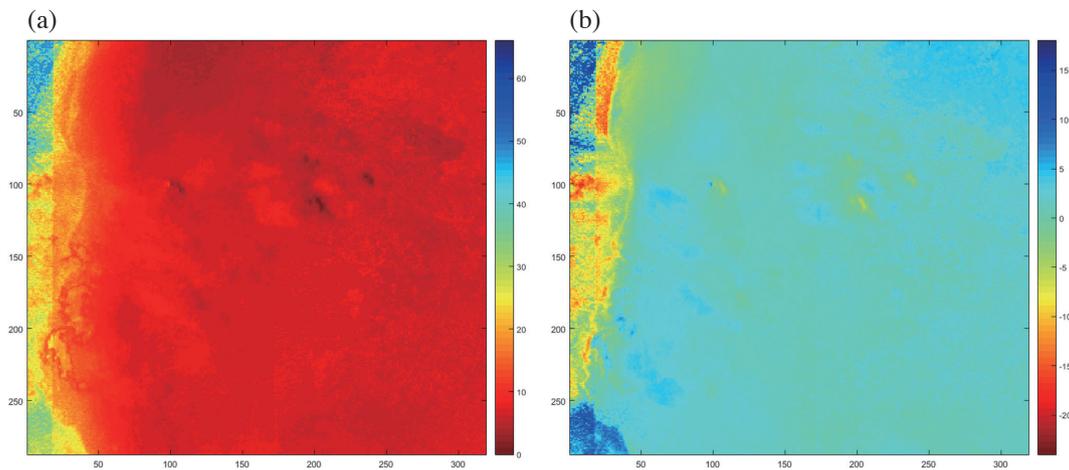


Fig. 12. WV-3 result from model with waterline constraint, (a) depth retrieved; (b) error.

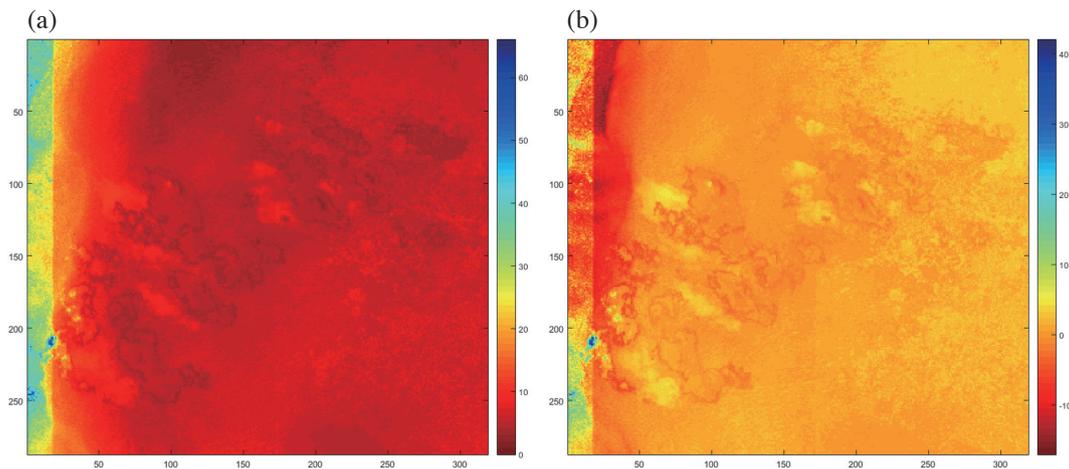


Fig. 13. WV-3 result from BP, (a) depth retrieved; (b) error.

Table 11. Error indices of pure water area, model depth estimation, WV-3.

Depth (m)	Mean Error (m)	Stan. Dev. (m)	RMSE (m)	Maximum (m)	Minimum (m)
0 - 15	1.66	1.11	2.00	5.63	-4.32
15 - 30	-2.87	5.49	6.19	15.73	-13.92
30 - 38	-3.50	8.27	8.98	17.70	-18.45

Table 12. Error indices of cloud influenced area, model depth estimation, WV-3.

Cloud (%)	Mean Error (m)	Stan. Dev. (m)	RMSE (m)	Maximum (m)	Minimum (m)	Reference Depth (m)
0 - 10	0.37	6.21	6.22	3.32	-12.28	3.5 - 30.0
10 - 20	1.15	1.34	1.77	4.48	-9.15	3.0 - 25.0
20 - 30	1.26	0.95	1.58	4.61	-5.09	3.0 - 22.4
30 - 40	1.22	1.15	1.68	6.45	-1.39	3.2 - 21.6
40 - 50	0.86	1.33	1.59	6.93	-1.95	3.7 - 19.5
50 - 60	0.49	1.49	1.57	7.35	-2.10	4.1 - 12.1
60 - 70	-0.57	1.30	1.42	3.53	-2.97	4.3 - 6.6
70 - 80	-1.16	1.88	2.21	5.38	-4.92	4.3 - 6.7
80 - 90	-1.86	2.55	3.16	6.98	-6.00	5.0 - 6.7
90 - 100	-1.63	1.68	2.38	2.93	-3.71	6.2 - 6.6

in the pure water region less than 30 m, the root-mean-squared errors of both methods is better than 2.89 m, which is equivalent to 10% relative error in WV2. In summary, the overall accuracy of the water depth retrieval using very high resolution multispectral satellite image reached 10%.

4.5.3 The Effect of Cloud Coverage

This study also considers the effect of cloud abundance and utilizes linear un-mixing for the cloud component removal for improving the quality of depth retrieval. The higher cloud abundance depicts the lower water abundance, and consequently, the retrieval water depth in cloud influenced area showed lower accuracy when compared to pure water region. To consider the cloud abundance less than 20%, the RMSE of both methods is better than a 3.13 m while the reference depth is less than 20 m. But for the 50% cloud abundance, the RMSE reached 13.24 and 11.16 m for semi-analytical and BP models. In summary, the cloud abundance should be considered in the water depth retrieval from the optical image.

5. CONCLUSIONS AND FUTURE WORKS

The depth retrieval scheme proposed in this study includes cloud component removal with linear un-mixing algorithm. With the experiment focused on two WorldView images and validated with bathymetric data obtained from airborne bathymetric lidar survey, the cloud removal procedure is shown to be most effective. Two depth retrieval approaches are then implemented. While the overall performance of BP neural network is better than the semi-analytical model, the model has less probability in producing large errors than the BP. The cloud removal procedure is even more effective in the semi-analytical model than with BP.

Three error factors are identified: the cloud abundance, the depth, and the substrate type. The agreement between the estimated and ground truth is better with less cloud

abundance. It is likewise clear that the error increases with the depth. Regarding the substrate type, the one providing higher reflectance performs better.

In this experiment, the conversion model established with the WV-2 image is applied to another image successfully. This may imply that a model established at a given time could be applied to other images collected at different dates providing the environment factors such as water property remain unchanged. Reference bathymetric data were used in both depth retrieval schemes. While the training is indispensable for the BP approach, there is possibility that no in situ depth observation is required for the semi-analytical model approach. In case the in situ information of phytoplankton concentration [C] is available, ground truth may not be required.

In this study, the proposed method only focused on cloud area without cloud shadow. As the cloud shadow will also affect the result of water depth determination, future work will also focus on cloud shadow removing. Once the cloud regions have been detected, it is possible to identify the possible cloud shadow region based on the azimuth and elevation angles of satellite and sun from image metadata (Fisher 2014). It can reduce the impact of cloud shadow and consequently improve the quality of water depth from an optical satellite image.

Acknowledgements The authors wish to thank Prof. Hsuan Ren, Center for Space and Remote Sensing Research, National Central University, Taiwan, for his constructive suggestions during the research.

REFERENCES

- Ceyhun, Ö. and A. Yalçın, 2010: Remote sensing of water depths in shallow waters via artificial neural networks. *Estuar. Coast. Shelf Sci.*, **89**, 89-96, doi: 10.1016/j.ecss.2010.05.015. [[Link](#)]

- Fisher, A., 2014: Cloud and Cloud-Shadow Detection in SPOT5 HRG Imagery with Automated Morphological Feature Extraction. *Rem. Sens.*, **6**, 776-800, doi: 10.3390/rs6010776. [[Link](#)]
- Gholamalifard, M., T. Kutser, A. Esmaili-Sari, A. Abkar, and B. Naimi, 2013: Remotely Sensed Empirical Modeling of Bathymetry in the Southeastern Caspian Sea. *Rem. Sens.*, **5**, 2746-2762, doi: 10.3390/rs5062746. [[Link](#)]
- Huguenin, R. L., M. H. Wang, R. Biehl, S. Stoodley, and J. N. Rogers, 2004: Automated Subpixel Photobathymetry and Water Quality Mapping. *Photogramm. Eng. Rem. Sens.*, **70**, 111-123, doi: 10.14358/PERS.70.1.111. [[Link](#)]
- IOCCG, 2006: Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. In: Lee, Z.-P. (Ed.), Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada, 122 pp.
- Kao, H. M., H. Ren, C. S. Lee, C. P. Chang, J. Y. Yen, and T. H. Lin, 2009: Determination of shallow water depth using optical satellite images. *Int. J. Remote Sens.*, **30**, 6241-6260, doi: 10.1080/01431160902842367. [[Link](#)]
- Lee, Z., K. L. Carder, R. F. Chen, and T. G. Peacock, 2001: Properties of the water column and bottom derived from Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data. *J. Geophys. Res.*, **106**, 11639-11651, doi: 10.1029/2000JC000554. [[Link](#)]
- Lee, Z., K. L. Carder, and R. A. Arnone, 2002: Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Appl. Optics*, **41**, 5755-5772, doi: 10.1364/AO.41.005755. [[Link](#)]
- Liew, S. C. and J. He, 2008: Uplift of a Coral Island in the Andaman Sea Due to the 2004 Sumatra Earthquake Measured using Remote Sensing Reflectance of Water. *IEEE Geosci. Rem. Sens. Lett.*, **5**, 701-704, doi: 10.1109/LGRS.2008.2004208. [[Link](#)]
- Lyzenga, D. R., 1978: Passive remote sensing techniques for mapping water depth and bottom features. *Appl. Optics*, **17**, 379-383, doi: 10.1364/AO.17.000379. [[Link](#)]
- Lyzenga, D. R., 1981: Remote sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and Landsat data. *Int. J. Remote Sens.*, **2**, 71-82, doi: 10.1080/01431168108948342. [[Link](#)]
- Lyzenga, D. R., 1985: Shallow-water bathymetry using combined lidar and passive multispectral scanner data. *Int. J. Remote Sens.*, **6**, 115-125, doi: 10.1080/01431168508948428. [[Link](#)]
- Mohamed, H., A. Negm, M. Zahran, and O. C. Saavedra, 2015: Assessment of Artificial Neural Network for bathymetry estimation using High Resolution Satellite imagery in Shallow Lakes: Case Study El Burullus Lake. Eighteenth International Water Technology Conference, IWTC18 Sharm ElSheikh, Egypt.
- Moses, S. A., L. Janaki, S. Joseph, J. P. Gomathi, and J. Joseph, 2013: Lake Bathymetry from Indian Remote Sensing (P6-LISS III) satellite imagery using artificial neural network model. *Lakes & Reservoirs: Research & Management*, **18**, 145-153, doi: 10.1111/lre.12027. [[Link](#)]
- Polcyn, F. C., W. L. Brown, and I. J. Sattinger, 1970: The Measurement of Water Depth by Remote Sensing Techniques. Willow Run Laboratories, the University of Michigan, Ann Arbor, 47 pp.
- Sandidge, J. C. and R. J. Holyer, 1998: Coastal Bathymetry from Hyperspectral Observations of Water Radiance. *Remote Sens. Environ.*, **65**, 341-352, doi: 10.1016/S0034-4257(98)00043-1. [[Link](#)]
- Shih, P. T.-Y., D. Arumugam, and S.-W. Shyue, 2011: Bathymetric Lidar Survey of Penghu Islands and Dongsha Atoll Using an Ellipsoidal Height System for Bathymetric Mapping In Shallow Waters and Difficult-to-Navigate Environments. *Sea Technology*, **52**, 42-45.
- Stumpf, R. P., K. Holderied, and M. Sinclair, 2003: Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnol. Oceanogr.*, **48**, 547-556, doi: 10.4319/lo.2003.48.1_part_2.0547. [[Link](#)]
- Zhang, Y., B. Guindon, and J. Cihlar, 2002: An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. *Remote Sens. Environ.*, **82**, 173-187, doi: 10.1016/S0034-4257(02)00034-2. [[Link](#)]