

Estimating and mapping chlorophyll-*a* concentration in Pensacola Bay, Florida using Landsat ETM+ data

LUOHENG HAN* and KAREN J. JORDAN

Department of Geography, University of Alabama, Tuscaloosa, AL 35487-0322, USA

The purpose of this study was to develop algorithms for estimating chlorophyll-*a* concentration in Pensacola Bay using Landsat 7 ETM+ data. The techniques used were band ratioing and regression modelling. Pensacola Bay is located on the west end of the Florida panhandle. As one of 39 estuaries located on the Gulf of Mexico, Pensacola Bay is impacted largely by rivers. The Landsat ETM+ data were first geometrically rectified. Then brightness values were converted to reflectance through the radiometric correction process. For the regression models, logarithmically transformed chlorophyll-*a* was used as the dependent variable. Single bands, band ratios and logarithmically transformed band ratios were the independent variables. R^2 values were computed and evaluated. Results from the study indicate that the ratio of ETM+ 1/ETM+ 3 was the most effective in estimating chlorophyll-*a*. Using this model a chlorophyll-*a* map was generated for Pensacola Bay.

1. Introduction

Estuaries are dynamic environments generally defined as semi-enclosed coastal bodies of water that have a free connection to the sea and within which seawater mixes with fresh water resulting from land drainage (Pritchard 1967). The estuarine environment is deemed one of the most productive and sensitive marine ecosystems (Baban 1997, Harding *et al.* 2002). These environments fluctuate in physical, chemical and biological conditions due to changes in fresh water input, tidal regimes, temperature, salinity, seasonal variability and other chemical and physical factors as well as human input from a multitude of effluents stemming from various human activities such as agricultural, industrial and municipal wastes (Kennish 1986).

Phytoplankton are floating or drifting single-cell algae that are primarily transported by water motion (Kennish 1986, Day *et al.* 1989). These organisms are found in all estuarine waters and contribute greatly to overall primary production. Due to the significant role that phytoplankton play in marine habitats, they are used as indicators of health in systems such as estuaries. Phytoplankton contain chloroplasts, which absorb and use the underwater light to fix carbon in the form of carbohydrate. Among the chloroplast pigments, chlorophyll-*a* is common to all phytoplankton, although two major colour phytoplankton groups – green and brown – also contain chlorophylls-*b* and -*c*, respectively (Boney 1988). Thus, chlorophyll-*a* is an indicator of the abundance of phytoplankton in the water. The pronounced scattering/absorption features of chlorophyll-*a* are: strong absorption

*Corresponding author. Email: lhan@bama.ua.edu

between 400–500 nm (blue) and at 680 nm (red), and reflectance maximums at 550 nm (green) and 700 nm (near-infrared (NIR)) (Han 1997).

Remote sensing techniques have been applied to measure chlorophyll-*a* by researchers. There are some algorithms that have been developed in this endeavour. Among these algorithms, band ratioing has proven to be advantageous because it tends to allow compensation for variations from atmospheric influences (Jensen 2005). In addition, the scattering and absorption characteristics of chlorophyll-*a* can be studied when more than one band is used (Dekker *et al.* 1991). A basic principle of using band ratios is to select two spectral bands that are representative of absorption/scattering features of chlorophyll-*a* (Gin *et al.* 2002). The previous studies have also indicated that the wavelength range for characterizing chlorophyll-*a* is between 400 nm and 900 nm. Therefore, the four bands which are mostly associated with chlorophyll-*a* are the blue, green, red and NIR bands (Han *et al.* 1994, Han and Rundquist 1997, Gin *et al.* 2002). The decision involving which bands to use is not always straightforward (Jensen 2005). Previous work has utilized the band ratio method for various types of sensors such as spectroradiometers (Dekker *et al.* 1991, Mittenzwey *et al.* 1992, Gitelson *et al.* 1996, Han 1997, Gin *et al.* 2002) and airborne and satellite sensors from narrow band to broadband sensors such as AISA (Airborne Imaging Spectroradiometer for Applications), MODIS (Moderate Resolution Imaging Spectroradiometer), Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) (Dekker *et al.* 1991, Baban 1997, Zhang *et al.* 2003, Chang *et al.* 2004). In addition, studies have been carried out in various geographical locations as well as environmental settings ranging from lakes of various trophic status to coastal lagoons, bays and estuaries.

Several satellite sensing systems were designed specifically for monitoring chlorophyll-*a* in ocean water, such as Coastal Zone Color Scanner (CZCS, 1978–1986) and the Sea-viewing Wide Field of view Sensor (SeaWiFS) (Jensen 2000). But they are mostly useful for deep ocean (Case I) waters. Landsat TM/ETM+ data are useful for assessing estuarine systems for several reasons. The data are economic, routinely available and archived. Although the spectral resolution of Landsat TM/ETM+ are modest in quantifying chlorophyll-*a*, the spatial resolution and coverage are adequate for monitoring estuaries. Landsat TM/ETM+ in conjunction with *in situ* water sampling provides the means to establishing a relationship between satellite-derived reflectance values and chlorophyll-*a* concentrations. The temporal and spatial distribution of chlorophyll-*a* can therefore be mapped.

The purpose of this study is to test and evaluate various band ratios of Landsat 7 ETM+ bands 1–4 and map the concentration of chlorophyll-*a* using the best band ratio for Pensacola Bay, Florida.

2. Methodology

2.1 Study area

The Pensacola Bay estuarine system is located on the west end of the Florida panhandle and is one of 39 estuaries located on the Gulf of Mexico (figure 1). Covering 373 km², the bay receives fresh water input from three main rivers: the Escambia, Blackwater and Yellow Rivers. The bay has an average depth of 6 m, with the bottom sediments predominantly made up of silts and sands (Schroeder

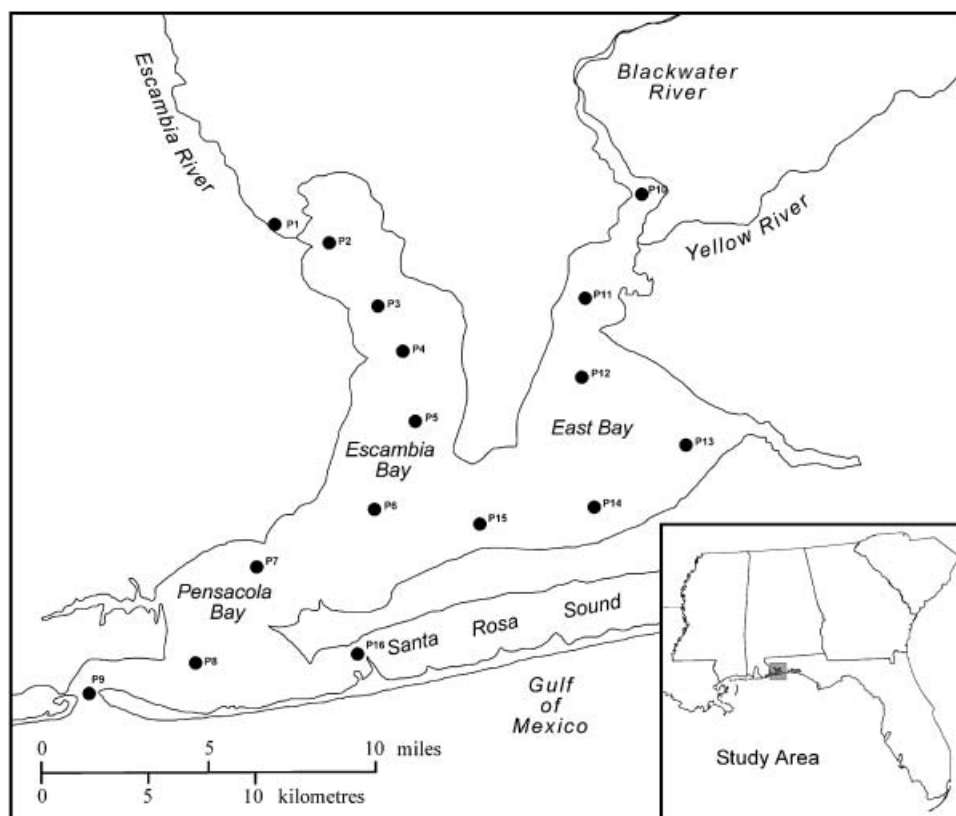


Figure 1. Map of Pensacola Bay study area and sampling sites.

and Wiseman 1999). Pensacola Bay is distinguished as a partially stratified, microtidal, drowned river valley system (Schroeder and Wiseman 1999, Murrell *et al.* 2002). More than 18 000 km² of watershed drain into the Pensacola Bay estuarine system, encompassing a multitude of land uses such as mixed forests, cropland, pastures, urban and industrial areas that introduce a high variability in water quality. Pensacola Bay, like most estuaries, plays an important role in supporting various marine lives.

2.2 Landsat 7 ETM+ data used

A Landsat 7 ETM+ scene (path: 20/row: 39; ID: 7020039000214050) of 20 May 2002, was used for this study. The image was acquired under a clear sky and windy conditions. The wind speed during the data collection was 24 km h⁻¹. In addition to a 15 m panchromatic band, Landsat ETM+ includes seven multi-spectral bands with 30 m spatial resolution for all bands except band 6, which is a thermal infrared (TIR) band with 60 m spatial resolution. The visible bands are ETM+1 (blue: 0.45–0.51 μ m), ETM+2 (green: 0.525–0.605 μ m), and ETM+3 (red: 0.63–0.690 μ m). The NIR band is ETM+4 (0.75–0.90 μ m) and the mid-infrared bands are ETM+5 (1.55–1.75 μ m) and ETM+7 (2.09–2.35 μ m). The TIR band is ETM+6 (10.40–12.50 μ m). The panchromatic band covers 0.52–0.90 μ m.

2.3 *In situ* data collection

In situ data collection consisted of water sampling and hyperspectral reflectance data collection. Both were conducted simultaneously on board the same boat on 14 and 15 May 2002, which was prior to the date when the Landsat ETM+ data were acquired. Although it is preferable to have both the *in situ* data collection and satellite overpass coincide, given that no major weather events occurred between 14 and 20 May 2002 and the tidal variability for the system was minimal, it was believed that the water quality parameters would remain relatively stable. Some previous research has produced fair results with much larger temporally mismatched satellite data and collected *in situ* data (Baban 1997).

The 16 sampling stations in the study are evenly distributed along Escambia Bay and East Bay (figure 1) with the depths ranging from 2.1 m to 11.3 m (table 1). These stations were selected and sampled monthly by the Gulf Ecology Division (GED) of the US EPA. For this study, stations P1 to P9 were sampled on the first day (May 14) whereas P10 to P16 were sampled on the second day (May 15). Each sampling station was located using the Global Positioning System (GPS) installed on the EPA water sampling boat with a positional error of less than 10 m. At each station, water samples were collected at the surface (at 0.5 m depth) waters. In addition to chlorophyll-*a*, there were more than 20 water quality parameters collected such as water temperature, salinity, pH, nitrate, phosphorus and other chemical measures. This paper focuses on estimating chlorophyll-*a* using Landsat ETM+ data only.

The hyperspectral data were collected with a hand-held spectroradiometer during the water sampling. The analysis of the data and results are reported in a separate paper.

2.4 *Image data processing*

ERDAS Imagine 8.6 (Leica Geosystems, Atlanta, USA), a digital image processing and GIS software, was used to process the Landsat ETM+ data. The image was first geometrically rectified to UTM (Universal Transverse Mercator) projection

Table 1. Geographical locations of the 16 sampling sites.

Station	Latitude	UTM Northing (m)	Longitude	UTM Easting (m)	Depth (m)
P1	30°33.133 N	3379993.2	87°12.037 W	480760.6	3
P2	30°32.368 N	3378579.8	87°09.663 W	484546.8	2.4
P3	30°30.953 N	3375962.2	87°08.577 W	486280.1	2.1
P4	30°29.621 N	3373498.2	87°07.828 W	487479.2	2.4
P5	30°27.409 N	3369413.8	87°07.925 W	487319.9	3.7
P6	30°24.908 N	3364796.0	87°08.937 W	485694.4	5.8
P7	30°23.064 N	3361399.8	87°12.585 W	479848.8	10.7
P8	30°20.549 N	3356761.5	87°14.785 W	476315.2	11.3
P9	30°19.659 N	3355132.8	87°18.510 W	470343.8	6.7
P10	30°34.137 N	3381832.5	86°59.931 W	500110.5	2.7
P11	30°31.253 N	3376505.8	87°01.670 W	497329.8	2.1
P12	30°28.812 N	3371997.8	87°01.985 W	496824.9	2.4
P13	30°26.949 N	3368556.8	86°58.600 W	502240.0	2.7
P14	30°25.690 N	3366231.8	87°01.041 W	498333.9	3.4
P15	30°25.051 N	3365055.0	87°54.200 W	491324.4	6.4
P16	30°20.623 N	3356884.0	87°09.575 W	484662.5	5.8

(Zone 16; Datum: WGS84). The image was rectified using USGS 7.5 minute quadrangle topographic maps. More than 30 ground control points (GCPs) were selected from both the image and topo maps. For the spatial interpolation portion of the geometric correction, a first-order, affine transformation was used. The RMS error was less than 0.5 pixel. For intensity interpolation, the nearest-neighbour interpolation was adopted in order to preserve the pixel values in the scene (Jensen 2005). Next, radiometric correction of the Landsat ETM+ image was performed based on the Chavez (1996) COST method. Brightness values (BVs) were converted to at-sensor reflectance values in this process.

The geometrically and radiometrically corrected Landsat ETM+ image was used in the analysis. A 3×3 window was established around each sampling pixel based on the UTM coordinates determined with a GPS during the water sampling. The mean reflectance of the 3×3 window was extracted and used in the modelling process. There are two primary reasons for using a 3×3 window instead of a single pixel: (1) because of possible errors in the geometric correction and the dynamics of water bodies, the use of a pixel window compensates for errors due to disparity of coordinates that may occur by using a single pixel; and (2) the pixel windows give an estimate of spatial variability. This technique was used widely in other previous research efforts (Baban 1997, Woodruff *et al.* 1999, Braga *et al.* 2003) and the size of a 3×3 window was suggested as an optimal one (Reddy 1997).

The band ratios among the first four ETM+ bands as proposed and tested in the literature were computed (Gitelson *et al.* 1996). In the regression models established, the logarithmically transformed chlorophyll-*a* concentration was used as a dependent variable (Chang *et al.* 2004). The three types of independent variables were tested: reflectance of a single band, logarithmically transformed band ratios, and ratios of logarithmically transformed single band. R^2 values were computed. From the best results, a map was generated showing the chlorophyll-*a* distribution and concentration in Pensacola Bay.

3. Results and discussion

Table 2 shows the laboratory-measured chlorophyll-*a* concentration for each sampling station. The mean chlorophyll-*a* concentration of 16 sampling stations was $7.41 \mu\text{g l}^{-1}$ and a relatively high standard deviation ($5.89 \mu\text{g l}^{-1}$) was recorded, which indicated the spatial variability of chlorophyll-*a* concentration. Escambia Bay had higher chlorophyll-*a* concentration (sampling stations P1–P9) than East Bay (sampling stations P10–P16). Higher concentrations of chlorophyll-*a* were also found in the upper part of the estuary, e.g. P2, P3, P10 and P11, where rivers tend to bring in nutrients. The highest concentration ($23.23 \mu\text{g l}^{-1}$) was recorded at station P2, which is located near the mouth of Escambia River while the lowest amount of chlorophyll was found at station P9 ($1.14 \mu\text{g l}^{-1}$) that is situated at the corridor of Pensacola Bay to the Gulf of Mexico.

Table 3 shows the average reflectance computed from a 3×3 window surrounding each sampling station. These values were used in building band ratio models. The first type of models that were established were the single band regression models, with reflectance and logarithmically transformed reflectance being the independent variable respectively, and logarithmically transformed chlorophyll-*a* concentration being the dependent variable (Chang *et al.* 2004). The low R^2 values indicated that there was little or no correlation between the reflectance of a single band and chlorophyll-*a* (table 4).

Table 2. Laboratory-measured chlorophyll-*a* concentrations.

Station	Chlorophyll- <i>a</i> ($\mu\text{g l}^{-1}$)
P1	8.99
P2	23.23
P3	18.20
P4	12.07
P5	7.12
P6	4.08
P7	3.32
P8	2.52
P9	1.14
P10	6.90
P11	7.32
P12	4.70
P13	4.70
P14	5.72
P15	4.85
P16	3.70
Mean	7.41
Std. Dev.	5.89

Table 3. Averaged reflectance from a 3×3 window for each sampling station.

Station	ETM + 1	ETM + 2	ETM + 3	ETM + 4
P1	0.1253	0.0982	0.0751	0.0680
P2	0.1374	0.1117	0.0932	0.0767
P3	0.1392	0.1134	0.0907	0.0776
P4	0.1370	0.1113	0.0863	0.0727
P5	0.1391	0.1109	0.0832	0.0713
P6	0.1423	0.1128	0.0838	0.0734
P7	0.1424	0.1113	0.0841	0.0767
P8	0.1412	0.1083	0.0800	0.0741
P9	0.1448	0.1090	0.0794	0.0780
P10	0.1312	0.1021	0.0802	0.0728
P11	0.1334	0.1050	0.0820	0.0717
P12	0.1397	0.1127	0.0889	0.0733
P13	0.1438	0.1172	0.0909	0.0762
P14	0.1487	0.1237	0.1001	0.0811
P15	0.1428	0.1127	0.0884	0.0788
P16	0.1403	0.1084	0.0781	0.0683

The next group of regression models tested were ones with chlorophyll-*a* (again logarithmically transformed) as the dependent variable and band ratios as the independent variable (table 5). As expected, the association between the dependent and independent variables became stronger as compared with the single band model

Table 4. R^2 for single band models.

Regression models	ETM + 1	ETM + 2	ETM + 3	ETM + 4
$\log(chl-a) = y_0 + a * b_j$	0.21	0	0.14	0.001
$\log(chl-a) = y_0 + a * \log(b_j)$	0.21	0	0.14	0.001

Table 5. R^2 for band ratio models.

Regression models	1/2	1/3	1/4	2/3	2/4	3/4
$\log(chl-a) = y_0 + a \cdot (b_j/b_k)$	0.28	0.45	0.23	0.49	0.01	0.24
$\log(chl-a) = y_0 + a \cdot \log(b_j/b_k)$	0.45	0.58	0.16	0.55	0	0.42
$\log(chl-a) = y_0 + a \cdot (\log b_j / \log b_k)$	0.57	0.67	0.26	0.58	0.01	0.41

(also table 4). Notice that the best regression model was the one with the form of

$$\log(chl-a) = y_0 + a \cdot (\log b_j / \log b_k) \quad (1)$$

Where $chl-a$ is chlorophyll- a concentration ($\mu g l^{-1}$); b_j and b_k are Landsat ETM+ bands 1–4.

Another observation that can be made from table 5 is that the ratio between ETM+1 and ETM+3 seemed to be the best of all six ratios. Figure 2 summarizes the statistical results of using the regression model described in equation (1) and the ratio of ETM+1/ETM+3. The R^2 value was 0.67 ($p < 0.0001$) and the standard error of estimate was 0.19 ($1.55 \mu g l^{-1}$). The principle behind this ratio is that both bands correspond to chlorophyll- a absorption. As chlorophyll- a increased, the reflectance in both ETM+1 (blue) and ETM+3 (red) decreased. But the rate of decrease in ETM+3 was faster than the one in ETM+1. Since the reflectance of ETM+3 (0.63 – $0.69 \mu m$) is also affected by inorganic suspended sediments and dissolved organic matter, the ratio of ETM+1 and ETM+3 may work effectively in estimating chlorophyll- a only when the chlorophyll- a concentration is higher than a certain level and the turbidity of water is relatively low. Gitelson *et al.* (1996) found the ratio of TM3/TM1, a reciprocal ratio of TM1/TM3, was effective in estimating chlorophyll- a concentrations greater than $3 mg m^{-3}$. A correlation coefficient, R , greater than 0.74 between TM3/TM1 and chlorophyll- a was produced.

Using this model, a chlorophyll- a concentration map was produced using ERDAS Imagine software (figure 3). The map characterizes the spatial pattern of chlorophyll- a in Pensacola Bay. Using chlorophyll- a as an indicator of phytoplankton, the map indicated that higher abundance of phytoplankton occurred in the upper part of Escambia Bay, where Escambia River enters the bay. A similar pattern was also found in East Bay. Higher chlorophyll- a concentrations were also

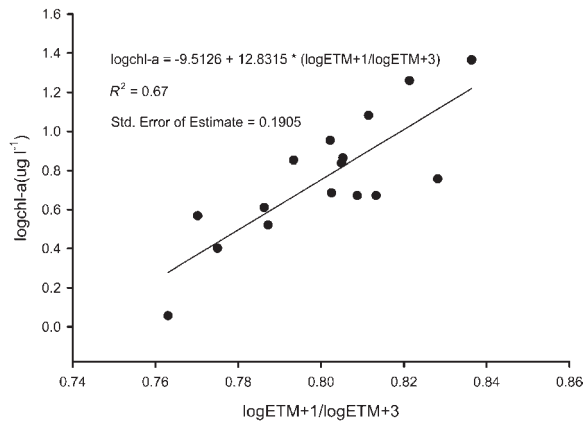


Figure 2. The regression model using the ratio of ETM+1/ETM+3.

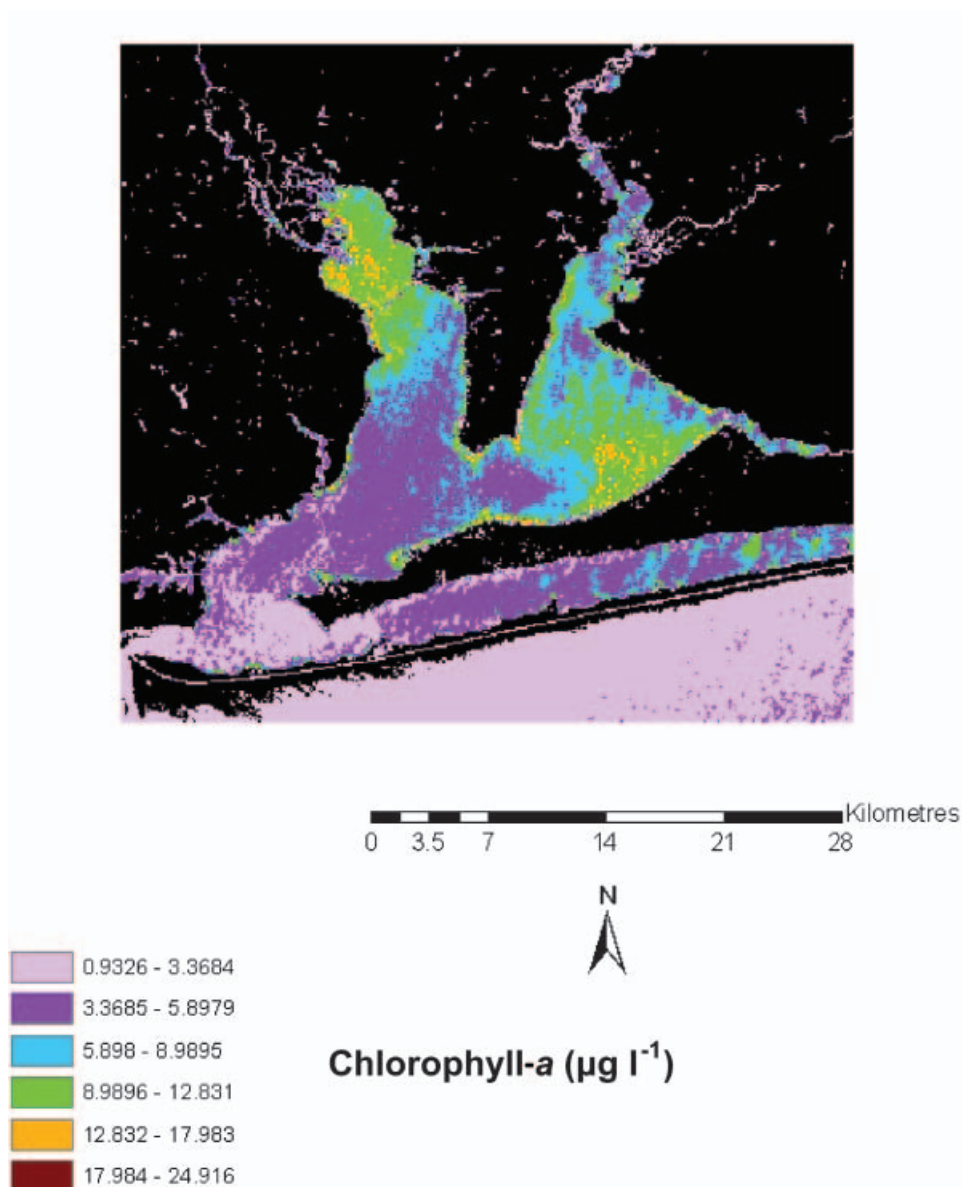


Figure 3. The chlorophyll-*a* map of Pensacola Bay.

found along the shore lines. It is believed to indicate the abundance of certain types of benthic algae.

4. Conclusions

Chlorophyll-*a* is an indicator of the abundance of phytoplankton, which make an important contribution to overall primary productivity of coastal waters. Therefore, using remote sensing techniques to estimate and map chlorophyll-*a* concentration is a significant undertaking for improving the monitoring and assessment of water quality in estuaries.

Although the spectral resolution of Landsat 7 ETM+ is relatively low in estimating chlorophyll-*a* compared with some of the other sensing systems, the ratio of ETM + 1/ETM + 3 seemed to offer a practical solution in estimating and mapping chlorophyll-*a*. There are, however, limits to this study, such as the fact that it was based the data collected on only one occasion, and 16 sampling stations.

Acknowledgements

This research has been supported by a grant from the US Environmental Protection Agency's Science to Achieve Results (STAR) Estuarine and Great Lakes (EaGLe) program through funding to the CEER-GOM, US EPA Agreement R82945801. Although the research described in this article has been funded wholly or in part by the United States Environmental Protection Agency, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. The Gulf Ecology Division of the US EPA provided water quality data and facilitated the spectral data collection. The authors thank Yongtao Luo for providing expertise in image processing and modelling.

References

- BABAN, S.M.J., 1997, Environmental monitoring of estuaries; estimating and mapping various environmental indicators in Breydon Water Estuary, U.K., using Landsat TM Imagery. *Estuarine, Coastal, and Shelf Science*, **44**, pp. 589–598.
- BONEY, A.D., 1988, *Phytoplankton* (London: Edward Arnold).
- BRAGA, C.Z.F., VIANNA, M.L. and KJERFVE, B., 2003, Environmental characterization of a hypersaline coastal lagoon from Landsat-5 Thematic Mapper data. *International Journal of Remote Sensing*, **24**, pp. 3219–3234.
- CHANG, K.W., SHEN, Y. and CHEN, P.C., 2004, Predicting algal bloom in the Techi reservoir using Landsat TM data. *International Journal of Remote Sensing*, **25**, pp. 3411–3422.
- CHAVEZ, P.S. Jr., 1996, Image-based atmospheric corrections-revisited and revised. *Photogrammetric Engineering and Remote Sensing*, **62**, pp. 1025–1036.
- DAY, J.W. Jr., HALL, C.A.S., KEMP, W.M. and YÁNEZ-ARANCIBIA, A., 1989, *Estuarine Ecology* (New York: John Wiley & Sons).
- DEKKER, A.G., MALTHUS, T.J. and SEYHAN, E., 1991, Quantitative modeling of inland water quality for high-resolution mss systems. *IEEE Transactions on Geoscience and Remote Sensing*, **29**, pp. 89–95.
- GIN, K.Y., KOH, S.T. and CHAN, E.S., 2002, Application of spectral signatures and colour ratios to estimate chlorophyll in Singapore's coastal waters. *Estuarine, Coastal and Shelf Science*, **55**, pp. 719–728.
- GITELSON, A.A., YACOBI, Y.Z., KARNIELI, A. and KRESS, N., 1996, Reflectance spectra of polluted marine waters in Haifa Bay, Southeastern Mediterranean: features and application for remote estimation of chlorophyll concentration. *Israel Journal of Earth Science*, **45**, pp. 127–136.
- HAN, L., 1997, Spectral reflectance with varying suspended sediment concentrations in clear and algal-laden waters. *Photogrammetric Engineering and Remote Sensing*, **63**, pp. 701–705.
- HAN, L. and RUNDQUIST, D.C., 1997, Comparison of NIR/Red ratio and first derivative of reflectance in estimating algal-chlorophyll concentration: a case study in a turbid reservoir. *Remote Sensing of Environment*, **62**, pp. 253–261.
- HAN, L., RUNDQUIST, D.C., LIU, L.L. and FRASER, L.N., 1994, The spectral responses of algal chlorophyll in water with varying levels of suspended sediment. *International Journal of Remote Sensing*, **15**, pp. 3707–3718.

- HARDING, L.W. Jr., MALLONE, M.E. and PERRY, E.S., 2002, Toward a predictive understanding of primary productivity in a temperate, partially stratified estuary. *Estuarine, Coastal and Shelf Science*, **55**, pp. 437–463.
- JENSEN, J.R., 2000, *Remote Sensing of the Environment, An Earth Resource Perspective* (Upper Saddle River, New Jersey: Prentice Hall).
- JENSEN, J.R., 2005, *Introductory Digital Image Processing, A Remote Sensing Perspective* (Upper Saddle River, New Jersey: Prentice Hall).
- KENNISH, M.J., 1986, *Ecology of Estuaries, Volume II, Biological Aspects* (Boca Raton: CRC Press).
- MITTENZWEY, K.H., ULLRICH, S., GITELSON, A.A. and KONDRATIEV, K.Y., 1992, Determination of chlorophyll *a* of inland waters on the basis of spectral reflectance. *Limnology and Oceanography*, **37**, pp. 147–149.
- MURRELL, M., STANLEY, R., LORES, E., DiDONATO, G., SMITH, L. and FLEMER, D., 2002, Evidence that phosphorus limits phytoplankton growth in a Gulf of Mexico estuary: Pensacola Bay, Florida, USA. *Bulletin of Marine Science*, **70**, pp. 155–167.
- PRITCHARD, D.W., 1967, What is an estuary, physical viewpoint. In *Estuaries*, G.H. Lauf (Ed.), pp. 3–5 (Washington D.C.: American Association for the Advancement of Science, Publ 83).
- REDDY, M.A., 1997, A detailed statistical study on selection of optimum IRS LISS pixel Configuration for development of water quality models. *International Journal of Remote Sensing*, **18**, pp. 2559–2570.
- SCHROEDER, A. and WISEMAN, A., 1999, Geology and hydrodynamics of estuaries. In *Biogeochemistry of Gulf of Mexico Estuaries*, T.S. Bianchi, J.R. Pennock and R.R. Twiley (Eds), chap. 1, pp. 3–22 (New York: John Wiley & Sons).
- WOODRUFF, D.L., STUMPF, R.P., SCOPE, J.A. and PAERL, H.W., 1999, Remote estimation of water clarity in optically complex estuarine waters. *Remote Sensing of Environment*, **68**, pp. 41–52.
- ZHANG, Y., KOPONEN, S.S., PULLIAINEN, J.T. and HALLIKAINEN, M.T., 2003, Application of empirical neural networks to chlorophyll *a* estimation in coastal waters using remote optosensors. *IEEE Sensors Journal*, **3**, pp. 376–382.