OCEAN COLOUR REMOTE SENSING OF CASE 2 WATERS USING AN OPTIMISED NEURAL NETWORK

SAUMI SYAHREZA

UNIVERSITI SAINS MALAYSIA

2016

OCEAN COLOUR REMOTE SENSING OF CASE 2 WATERS USING AN OPTIMISED NEURAL NETWORK

by

SAUMI SYAHREZA

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

MARCH 2016

ACKNOWLEDGEMENTS

In the name of Allah, Most Gracious

Most Merciful First and foremost, I give all praises and adoration to Allah SWT for granting me good health, patience, and wisdom to complete this study. I pray for peace and for a blessing on all His noble Prophets and Messenger, and in particular on the last of them all, the blessed Prophet Muhammad SAW.

I am highly grateful to my supervisors, Professor Mohd Zubir Mat Jafri and Associate Professor Lim Hwee San from School of Physics, Universiti Sains Malaysia, for their consistent guidance, time devotion, and supports throughout my study. I have been honoured to carry out this work under their instructions. Their passion in education and borderless academic has enabled me to accomplish this project and establish of my scientific expertise.

Also, I am grateful to the members of our research group Messrs.' Rifhan, Haidar, Sinan, Ahmed Nazri, Beh, Tan, E.K Makamah. All your contributions to this work are highly appreciated.

I would also like to thank all staff (administrative, academic, laboratory, technical and technical) and students from School of Physics, Universiti Sains Malaysia, especially the staff of Engineering Physics and Geophysics Laboratory for their cooperation and assistance during the field work, Messrs. Yaakob Othman and Shahil Ahmad Khosani. Also, other members of the academic staff of the School of Physics, Associate Professor Dr. Khiruddin Abdullah, Dr. Md. Noordin Abu Bakar and Dr. Ahmad Fairuz Bin Omar. My appreciations also come to Dr. R. Richter of DLR - German Aerospace Centre, Weslling, Germany for ATCOR 2/3 software used in this research. Too many whom I cannot mention one by one in this limited space.

I say thank you. I also particularly grateful for all the financial support and discussions I received from (1) DIKTI Scholarship-Batch 3, Ministry of Education, Indonesia; (2) Research University-Postgraduate Research Grant Scheme (RU-PRGS) USM, Grant number: 1001/PFIZIK/8442020; (3) BBNAD assistant, Aceh Government, Indonesia; (4) PORSEC and INCOIS for the opportunity to attend the workshop and conference on active microwave and ocean colour remote sensing at Kochi, Kerala-India; (5) GISTDA-ISPRS Student Consortium, AARS, and Burapha University, for attended the 8th Summer School 2012 on "Advance Remote Sensing for Coastal Zone Monitoring and Disaster Management" at Thailand; (6) Yangtze Environmental Specimen Bank-Tongji University for attended the training and conference on regional pollution and environmental changes at Shanghai, China.

My deepest appreciations go to my family for their understanding, support and encouragements, My Mother, my dearest wife (dr. Rina), my lovely boy (Naufal) and my siblings, my elder sister Elfiana, S.Kep, Adi, S.T and Budi, S.T. I also appreciate all the support and contributions of my colleagues in place of work; Dr. M Syukri Surbakti, Dr. Nazli, M.Si, T. Khairuman, M.Si, and Nasrullah Zaini, M.Sc. May Allah reward you in the manifold for your contributions. Also, very special thanks to all my friends whom I cannot mention one by one in this limited space. Finally, I particularly thank the Syiah Kuala University, Aceh for the grant to pursue my doctoral program.

SAUMI SYAHREZA

2015

TABLES OF CONTENTS

LIST OF TABLE

Page

LIST OF FIGURES

LIST OF SYMBOL

Symbols	Descriptions
α_o, α_k, α_j and β	The coefficients derived from regression analysis
$\beta(\theta,\lambda)$	Solid angle
$\beta(\psi,\lambda)$	Spectral volume scattering function
ρ	Internal Fresnel reflectance
$\rho(\lambda)$	Spectra reflectance
$\bar{\rho}$	Air-water Fresnel reflection
σ	Slope
ψ	Angle
$\Delta\Omega$	Spectral radiant intensity
ΔA	Area
ΔV	Volume of water
Δr	Thickness of water
ΔV	Volume of water
$\Phi_a(\lambda)$	Incident power-absorb
$\Phi_i(\lambda)$	Incident power
$\Phi_s(\lambda)$	Incident power- scatter
$\Phi_t(\lambda)$	Incident power- transmit
$\Phi_{my}(\lambda)$	Spectral radiant power
$\Phi_s(\lambda, \psi)$	Scattered out of the beam at an angle ψ
a'_{ph},	Specific absorption of phytoplankton
a'_s	Specific absorption of suspended matter
a'_y	Specific absorption of yellow substance
a_k	Activation potential
a_{ph}	Absorption coefficient of phytoplankton
a_s	Absorption coefficient of suspended matter
a_w	Absorption coefficient of absolutely pure water
a_y	Absorption coefficient of yellow substance
$a(\lambda)$	Absorption coefficient
$A(\lambda)$	Spectral absorptance
$b_b(\lambda)$	Backscattering coefficient
$b_f(\lambda)$	Forwards scattering coefficient
$b(\lambda)$	Scattering coefficient
b'_{ph},	Specific backscattering of phytoplankton
b'_s	Specific backscattering of suspended matter
b_k	Bias
b_w	Water backscattering coefficient
$B(\lambda)$	Scatterance
$B(\theta, \lambda)$	Fraction of incident power
$c(\lambda)$	Beam attenuation coefficient
$C_{b,b}$	Covariance of each bands
$Cal(\lambda)$	Calibration factor
e	Coefficient

E	error
E_d	Downwelling plane irradiances
E_i	Incident irradiance
E_I	Energy of wavelength reflected
E_R	Energy of wavelength incident
E_u	Spectral upwelling
F(y), $F(\alpha)$	Activation function
$F(a)$	Logistic activation function
I_s	Intensity
L	Water leaving radiance, Radiance
n	Refractive index
Q	Q Factor
$\bar{r}$	Water-air reflection
R_{rs}	Remote sensing reflectance
R(0-).	Below water reflectance
R(0+),	Above-water reflectance
$R(z;\lambda)$	Spectral irradiance reflectance
R_{ird}	Irradiance reflectance just beneath the water surface
$R_{XY,}\, r$	Pearson correlation
$R_{XY,}\, R^2$	Coefficient of determination
$S(\lambda)$	Radiance from reference panel
$T(\lambda)$	Transmittance
$w_{k1,}\, w_{km}$	Weights
x′	normalized input value
x_i, x_m	Input signal
y_k	Output neurons
$\hat{y}$, Y	Parameters of ocean colour products
z, w	Depth

LIST OF ABBREVIATION

ALOS	Advanced land observing satellite
ANNs	Artificial neural networks
AOPs	Apparent optical properties
ATCOR-2	Atmospheric correction-2
AVNIR	Advanced visible and near-infrared radiometer
BPNN	Back propagation neural network
CA–NN	Correlation analysis-neural network
C_{chl}	Chlorophyll-a concentration
C_s	Suspended sediment concentration
CZCS	Coastal zone colour scanner
DEM	Digital elevation model
ETM	Enhanced Thematic Mapper (ETM)
HR	High resolution
IOPs	Inherent optical properties
IVS	Input variable selections
JAXA	Japan aerospace exploration agency
Landsat TM	Land remote sensing satellite-Thematic Mapper
MLPNN	Multilayer perceptron neural network
MNLR	Multivariate nonlinear regression
MODIS	Moderate resolution imaging spectroradiometer
MSS	Multi spectral scanner
MVLR	Multivariate linear regression
MVR	Multivariate regression
MVRM	Multivariate regression model
NASA	National aeronautic space of America
NN	Neural network
NTU	Nephelometric turbidity units
OCTS	Ocean colour and temperature scanner
OM	Optical model
OR	Optical regression
Pan	Panchromatic
PC	Personal computer
PCA-NN	Principal correlation-analysis neural network
PLS	Partial least square
POLDER	Polarization and directionality of Earth's reflectances
RBV	Return beam vidicon
RMSE	Root mean square error
SDD	Secchi disk depth
SeaWiFS	Sea-viewing Wide Field-of-view sensor
SLR	Simple linear regression
SLRTM	Shuttle radar topography mission
TSS	Total suspended solid
TURB	Turbidity
VIF	Variance inflation factor

PENDERIAAN JAUH WARNA LAUT BAGI PERAIRAN KES 2 MENGGUNAKAN RANGKAIAN NEURAL YANG DIOPTIMUMKAN

ABSTRAK

Kajian ini memberi tumpuan pada pembangunan algoritma baru untuk mendapatkan produk warna laut di kawasan perairan kes 2 menggunakan model rangkain neural (NN) dari pelbagai jenis data penderiaan jauh sebagai input. Model NN dan parameter latihan dioptimumkan dengan input yang dipilih berdasarkan analisis korelasi (CA) dan analisis komponen utama (PCA). Di pesisiran pantai Kelantan, penggunaan data spektra pantulan in situ dan simulasi penderiaan jauh satelit telah dikaji untuk menganggar dua parameter kejelasan iaitu kekeruhan (TURB) dan cakera kedalaman Sechhi (SDD). Data simulasi Landsat TM dan AVNIR-2 diuji berdasarkan pengukuran spektra pantulan in situ menggunakan ASD spectroradiometer. Keputusan menunjukkan bahawa data simulasi Landsat TM dan AVNIR-2 membenarkan tafsiran TURB dan SDD. Di kawasan pesisiran pantai Pulau Pinang, penggunaan data satelit penderiaan jauh tunggal dan gabungan pelbagai tarikh telah dikaji untuk menganggar sediment terampai (C_s) dan kepekatan klorofil (C_{chl}). Pengukuran sampel air pelbagai tarikh telah dibuat selari dengan perolehan data satelit Landsat TM dan AVNIR-2 di lokasi terpilih dari Februari 1999 hingga Mac 2011. Data pantulan dari imej Landsat TM dan AVNIR-2 yang diperoleh menggunakan perisian ATCOR-2/3 diambil pada setiap lokasi sampel air dan data pantulan tersebut diperiksa dengan pelbagai algoritma. Walaupun hubungan yang signifikan telah dikesan antara nilai-nilai pantulan terhadap C_s dan C_{chl} apabila menggunakan model regresi linear berganda dan optikal, namun aplikasi NN

menunjukkan prestasi yang lebih baik untuk pemodelan dalam kajian ini. Keputusan menunjukkan bahawa ketepatan penganggaran untuk ciri-ciri dua produk warna laut menggunakan NN adalah jauh lebih baik daripada algoritma empirikal dan optikal. Keputusan juga menunjukkan bahawa NN yang dioptimumkan berdasarkan analisis korelasi (CA) dan analisis komponen utama (PCA) boleh meningkatkan penganggaran ciri-ciri produk warna laut. Dengan menggunakan lima pembolehubah bebas (TM1, TM2, TM4, TM5 dan TM7) sebagai input, model rangkain neural (PCA-NN) dari data Landsat TM gabungan pelbagai tarikh menunjukkan sedikit peningkatan prestasi pencarian C_s and C_{chl} (R^2 = 0.93 dan R^2 = 0.92 untuk RMSE = 22.38 μg/L dan RMSE = 1.29 μg/L) berbanding dengan model NN dan CA-NN. Sebaliknya, model CA-NN dengan empat pembolehubah bebas (AV1, AV3, AV4 dan AV3/AV2) daripada data AVNIR-2 gabungan pelbagai tarikh menunjukkan prestasi ramalan yang lebih baik (R^2 = 0.93 dan R^2 = 0.92 untuk RMSE = 22.38 μg/L dan RMSE = 1.29 μg /L) berbanding dengan model NN dan PCA-NN. Walaupun PCA-NN menggunakan data AVNIR-2 gabungan pelbagai tarikh telah berkurang sedikit kuasa generalisasi, namun pengoptimuman model PCA-NN menunjukkan prestasi ramalan yang lebih baik berbanding NN model. Oleh itu, ia memungkinkan untuk membangunkan algoritma warna laut di mana CA dan PCA digunakan sebagai kaedah pemilihan input pembolehubah pada NN untuk pemerhatian penderiaan jauh.

OCEAN COLOUR REMOTE SENSING OF CASE 2 WATERS USING AN OPTIMISED NEURAL NETWORK

ABSTRACT

This study focused on the development of the new algorithm for retrieving ocean colour products of Case 2 water types using the neural network (NN) model and multiple types of remotely sensed data as inputs. The NN model architecture and training parameters were optimised, with inputs being selected based correlation analysis (CA) and principal component analysis (PCA). In Kelantan coastal waters, the use of in situ reflectance spectra and simulated satellite data for estimation of two water clarity parameters namely turbidity (TURB) and Secchi disk depth (SDD) have been studied. The simulated Landsat TM and AVNIR-2 data were tested against in situ reflectance spectra measurements using ASD Spectroradiometer. The results show that the simulated Landsat TM and AVNIR-2 data enables the interpretation of TURB and SDD. In Penang coastal area, the use of single and multitemporal remote sensing data for estimation of C_s and C_{chl} has been studied. Multidate in-situ water sample measurements concurrent with Landsat TM and AVNIR-2 satellite data were obtained in selected locations from February 1999 to March 2011. The irradiance reflectances of Landsat TM and AVNIR-2 derived by ATCOR-2/3 software from the water sampling sites were extracted and examined with numerous algorithms. Although significant correlation was detected between reflectance values of C_s and C_{chl} when using the multivariate, and optical model, however the application of NN appears to produce superior performance in modelling transfer function in this study. The results show that the estimation accuracy for characteristics of two ocean colour

products using neural network is much better than those empirical and optical algorithms. The results also indicated that NN based on CA and PCA can improve the estimation of these characteristics. Using five independent variables (TM1, TM2, TM4, TM5 and TM7) as inputs, PCA-NN model from multitemporal Landsat TM data has shown slightly improved the retrieval performance of C_s and C_{chl} (R^2 = 0.93 and R^2 = 0.92 for RMSE = 22.38 mg/L and RMSE = 1.29 μg/L) in comparison with the NN and CA-NN model. On contrary, the CA-NN model with four independent variables (AV1, AV3, AV4 and AV3/AV2) of multitemporal AVNIR-2 data has shown better prediction performance for C_s and C_{chl} with (R^2 = 0.93 and R^2 = 0.92 for RMSE = 22.38 mg/L and RMSE = 1.29 μg/L) in comparison with the NN and PCA-NN model. Although the PCA-NN using the multitemporal AVNIR-2 data has a slightly reduce generalization power, however the optimization of the PCA-NN model has demonstrated better prediction performance in comparison with the NN model. Therefore, it may be possible to develop ocean colour algorithms in which CA and PCA are used as methods for selecting input variables to a NN for remote sensing observation.

CHAPTER 1

INTRODUCTION

1.1 Research background

Coastal areas are the important ecological system that has a strategic value for many nations. Coastal areas are generally defined as the transition area between terrestrial and sea, the landward boundary covering part of the dry and wetland that is still affected by the oceanography properties at the sea surface, such as wind, tidal, seawater intrusion in the coastal, which characterised by a typical vegetation. On the other hand, the boundary of the coastal areas to the seaward reaches the outer limit of the continental shelf, where the characteristics of its waters are still influenced by a natural process that occurs in the terrestrial environment, such as sedimentation and flow of fresh water on estuarine, or process caused by human activities such as deforestation and contamination.

The coastal area is a very complex, dynamic environment and it is the most productive area of the global ocean (Chen-Tung Arthur et al., 2003; Muller-Karger et al., 2005). It is also the area for processing of nutrients and dissolved materials derived from the terrestrial environments to the coastal ocean. A large number of biological, geological, chemical, and physical processes occur over a multi-time and space scales. In addition, the optical properties of coastal waters also have particular complexity. It is characterised by the propagation and utilization of light in the water body varies over time and space scales conforming to changes in concentrations of optically active ocean components, such as phytoplankton, suspended sediments and coloured dissolved organic matter (Miller et al., 2005). Currently, the coastal areas are facing great pressure from various excessive human activities, either directly or

indirectly, starting with the intense concentration of population, overfishing, and mineral extraction, changing sediment load, coastal deforestation, pollution, tourism, global warming and sea level rise. These pressures continue to increase, and will threaten the future of countless ecosystems. Many attempts have been made by researchers in studying coastal waters to obtain a better understanding of earth system processes for climate change or environmental change research for integrated coastal management**.** Unfortunately, the dynamic nature of coastal waters made most measurements and traditional data collection becoming ineffective (Miller et al., 2005). In addition, the quality and quantity of the measurements can have different in every moment. This task would cost relatively much and requires a long time. It becomes more uneconomical when making detailed measurements on the variability of various coastal processes for large coverage simultaneously.

In contrast, space-based remote sensing offers unique large-scale synoptic information to address the complex environment of coastal waters. The periodic satellite overpass enables the routine and cost-effective collection of various observations over large areas and often inaccessible locations from the coast and adjacent waters in a short period of time. This set of satellite observations meets the needs of many users. Coastal managers, fishermen, and environmental scientists are some of the many groups that take advantage of oceanographic products derived from remote sensing observations (Brown et al., 2005).

1.1.1 Coastal ocean colour

Since 1970, there has been important progress in the oceanic observations. Spacecraft and sensors are exclusively developed to the study of ocean colour, for example, CZCS, OCTS, SeaWiFS, POLDER, MODIS, Landsat TM/ETM,

ALOS/AVNIR-2, IKONOS, Etc. The goal of these remote sensors are to obtain information on processes such as the distribution and dissemination of dissolved organic matter, identifying phytoplankton, biogeochemical assessment, understanding changes in benthic communities in clear, shallow tropical coral reefs, and monitoring coastal water quality (Bukata, 2005; Muller-Karger et al., 2005).

Many efforts have been made to estimate and map coastal and ocean colour products using these ocean colour sensors (Garcia et al., 2005; Miller & McKee, 2004; Reinart & Reinhold, 2008; Tilstone et al., 2015; G. Wu et al., 2014; M. Wu et al., 2009). Investigations also suggest that the high resolution satellite imageries such as Landsat TM/ETM, ALOS/AVNIR-2, and SPOT may be more useful to gain information in coastal waters (Alashloo et al., 2013; Allan et al., 2011; Bonansea et al., 2015; Pattiaratchi et al., 1994; Sakuno et al., 2014; G. Wu et al., 2013).

The observations of ocean colour are particularly important because they are useful to estimate the spatial distribution of routine and the rate of change in a variety of organic carbon in the ocean. It is related to living and natural waste materials and with coloured dissolved the organic matter (CDOM). These data are valuable for the coastal zona applications including managing and monitoring of suspended sediments and dissolved materials linked to the discharge of rivers and resuspension from the bottom. These tasks can be accomplished by retrieving inherent optical properties of the water constituents such as suspended matter and chlorophyll from ocean surface reflectance measurements (Muller-Karger et al., 2005).

1.1.2 Case 1 and Case 2 waters

Physical oceanographers have developed an extensive theoretical basis for the assessment of ocean hue and parameters that influence it. Pure water itself

absorbs and scatters light, but otherwise the colour of ocean waters depends on the water constituent concentrations that absorb and scatter light. The constituents of concern are phytoplankton, organic and inorganic particulate materials (detritus and inorganic sediment), and coloured dissolved organic matter (CDOM, yellow substance, or Gelbstoff) that result from significant quantities of terrigenous materials (Woodruff et al., 1999). In shallow coastal waters, reflected light from the bottom can affect the hue of water (Shubha Sathyendranath, 2000).

In general, coastal areas or in lakes are often designated as "Case 2" waters. It is recognized that Case 2 waters are more complex rather than Case 1 waters in their constituent compositions and optical properties. On the other hand, Case 1 waters usually refer to open sea are when phytoplankton concentrations mainly responsible for ocean colour. Therefore, the interpretation of an optical signal from Case 2 waters can be rather difficult due to those optical complexities (Shubha Sathyendranath, 2000). Panda et al. (2004) explained that the complex optical situation in coastal waters and water dynamics are more difficult to have linear relationship. Ritchie et al. (2003) found that the linear relationship between the concentration of suspended sediments and sea water reflectance in the range between 0 and 50 mg/L. In other studies, the relationship found between reflectance spectra from datasets typical coastal waters and the suspended sediments was strongly nonlinear, even when suspended sediment range between 0.5–10 mg/L (Froidefond et al., 1993). Y. Zhang et al. (2002); Panda et al. (2004) found that the fit function between suspended sediment and reflectance measured by satellite data is only linear between 1–20 mg/L.

1.1.3 Case 2 ocean colour algorithms

Many studies have tried to introduce new approaches to improve the performance of Case 2 ocean colour algorithms (Ali & Ortiz, 2014; Doxaran et al., 2002; Islam et al., 2001; Ritchie et al., 1990; Ritchie et al., 2003; Teodoro et al., 2007). For instance, Doxaran et al. (2002) proposed a relationship between the visible and near-IRed band reflectance of satellite data and suspended sediment concentrations. Ritchie et al. (1990); Ritchie et al. (2003) developed an empirical algorithm to estimate the suspended sediments and chlorophyll concentration from remote-sensing measurements. Band ratio algorithms were proposed (Doxaran et al., 2002; M. Zhang et al., 2010), though band ratio algorithms sometimes can improve suspended sediment inversion accuracy, but not significantly (Sakuno et al., 2014; Y. Zhang et al., 2003). Multivariate algorithms (Ali & Ortiz, 2014; Su et al., 2008; Vignolo et al., 2006), and new estimation algorithms such as polynomial, logarithmic, power and exponential (Forget & Ouillon, 1998; Islam et al., 2001; Ma & Dai, 2005) were also proposed.

Cipollini et al. (2001); Y. Zhang et al. (2002) use artificial neural network (ANN) to retrieve the ocean colour products from hyperspectral and multispectral data. Gan et al. (2004) compared the results of linear regression, non-linear regression, and an ANN to determine the seawater optically active parameters using Landsat TM data. They conclude that the spectral reflectance relationship is highly non-linear spatially and temporally, even non-linear regressions are not sufficient for modelling nonlinear transfer functions. For study of the non-linear behaviour to extract the concentration of water constituents from ocean colour, the ANN probably more potential than the regressions models, partially because of its parallel computing structure that arises from neurons being interconnected by a network (Gan

et al., 2004). In another part of NN study, Y. Zhang et al. (2002) compared the results of multivariate linear regression (MLR) and NN model to retrieve surface water quality using combined satellite optical and microwave data. The results show that the estimation accuracy of surface water using NN is much better than those from MLR analysis. The ANNs have been widely applied for optimal retrieval of water constituents from ocean colour in Case 2 waters using space platform (Chebud et al., 2012; X. Chen et al., 2004; Dzwonkowski & Yan, 2005; Gross et al., 1999; Gross et al., 2000; Song et al., 2012; Tanaka et al., 2004; T.-S. Wang et al., 2008).

However, most of the artificial neural networks (ANN) studies cited above have involved the input from all channels of satellite data for the evaluation of the ocean colour. While other studies have formed a relatively small number of input parameters yet with more than a hidden layer were used. In a case study based on remote sensing data, hyperspectral and multispectral remote sensing are receiving the reflected energy from the terrestrial surface at different intervals of the electromagnetic spectrum with specific wavelengths. The information from this wavelength range is stored at independent bands (Estornell et al., 2013). Due to the specific nature of the certain information received by the remote sensor (Shubha Sathyendranath, 2000) causing two bands located closely each other are highly correlated, whereas highly correlated input can create misperception to the NN (Walczak & Cerpa, 1999), or it can create redundancy, which results in low predictive power (May et al., 2011).

In this current study, an investigation was conducted to develop new algorithms for more accurate estimation of water constituents from ocean colour in Case 2 waters using remotely sensed data. The main difference between conventional NN algorithms and the NN approach is that the NN model is fewer complexes,

requires a smaller number of input parameters and hidden layer units, and most importantly, input parameters are pre-determined using CA and PCA. Therefore, the ANN approach can be implemented on ocean colour products using the inputs available from remote sensing data.

1.1.4 Case studies

This thesis presents two case studies to illustrate and prove the potential use of optimised algorithms for studying ocean colour remote sensing of Case 2 at different sites and different ocean colour parameters: C_s, C_{chl}, TURB, and SDD. The case studies do not represent all possible uses of developed algorithm in Case 2 waters but are typical of some problems encountered by oceanographers and other researchers. Besides, the choice of case studies was also based on in situ field data available and author's personal experience in study areas.

1.2 Research problems

As mentioned in the introduction, the water dynamics is very complex (Case 2) to have a linear correlation between the ocean colour products and the satellite spectral signatures, because there is significant scattering (in the visible light and near-IRed) from the coastal waters containing high C_s and C_{chl}. In addition, the complex optical situation of coastal waters may confound the reflectance signal (Ali & Ortiz, 2014). Therefore, the development of methods to characterise ocean colour in coastal areas has been a great challenge topic and required new approximation algorithms.

Previous studies in the coastal area of Penang (Case 2) have developed an optical algorithm and explored the possibilities of using remotely sensed data for the

determination of C_s and C_{chl}, with varying degrees of success (Abdullah et al., 2002; Asadpour et al., 2011; Jafri et al., 2003; Lim et al., 2008) These studies have shown that three bands (band 1–3) of data from various satellite sensors contain adequate information to conduct quantitative evaluations for both of this constituent. However, due to limited only to bands 1, 2 and 3 used to develop the model. Therefore, the estimation algorithms still need to be enhanced by using another band combination, it based on detecting multicollinearity between independent bands by methods of examining the variance inflation factors (VIF).

The data availability by in situ measurements of any single sampling event to develop such models for describing the variation of water constituent from ocean colour of Case 2 waters is limited in this study, particularly in spatial terms, whereas, a conceptual model requires sufficient data on the problems entity that can use to build the model, to develop mathematical and logical relationships, and to test the model's underlying assumptions (Sargent, 2005). Bruzzone et al. (1999) also described that the value of such a probability may remarkably affect the prediction result. Furthermore, appropriately testing model requires splitting the data into a calibration and validation, which further reduce the amount of data. Therefore, the study of integration potential of multi-date satellite data to obtain better predictions of the C_s and C_{chl} (multitemporal) uses a time series of 30-m resolution images from the Landsat TM and 10-m resolution from the ALOS/AVNIR-2 are investigated.

The NN models have been extensively applied for the evaluation of ocean colour in Case 2 waters by using remotely sensed data (J. Chen et al., 2014; Gross et al., 2000; Ioannou et al., 2013; M. Wu et al., 2009). The NN training using in situ data has become a good method for these purposes. However, one of the essential issues in applying NN model for the analysis of nonlinear transfer function is the

input variables selection problem (Giordano et al., 2014). There are no such assumptions made regarding the structure of the NN model. In contrast, the input variable selection to the neural and developing models based on the available data (May et al., 2011). Therefore, the algorithms can still to be refined by identifying the optimal form of NN models.

The analysis of spectral characteristics is the important part in the remote sensing technologies. The main advantage of this method is to optimise the combination of bands, which may help to decrease the difficulty in selection of bands and development of contextual algorithm. Therefore, it is necessary to study the spectral characteristics of water constituents from ocean colour in Case 2 waters. On the other hand, remote sensing offers some advantages such as good spatial and temporal coverage and the possibility of measuring many lakes simultaneously. With remote sensing, some others parameters of water quality, such TURB and SDD may influence ocean colour could potentially be assessed using a remotely sensed data. Therefore, in order to study the usefulness of remote sensing data for water quality measurements, the use of spectroradiometer in-situ data and simulated satellite data are important to be investigated, especially in areas that have not done specific research on water quality using satellite data.

1.3 Research objectives

The main objective of this research is to develop an optimised neural network algorithm for ocean colour products estimation in Case 2 waters using multitemporal remotely sensed data as inputs. The specific objectives that summarize the importance of this research are as follow:

1. To examine the relationship between satellites spectral reflectance and water constituents such as C_s and C_{chl} from ocean colour from numerous time measurements.
2. To improve the accuracy of model-based optical for interpretation of C_s and C_{chl} from ocean colour in Case 2 waters, based on detecting less-multicollinearity between independent bands by the methods of examination the VIF.
3. To characterise the spectral reflectance of Case 2 waters, which is characterised by heterogeneous TURB and SDD as the optical properties that influences ocean colour and aims to optimise the combination of bands that could aid the quantification both variables using simulated satellite data.

1.4 Research scopes

This study primarily focuses on the development of a new algorithm for estimating concentrations of the substances in the water column, such as C_s, C_{chl}, TURB and SDD that influence on ocean colour in Case 2 waters. As case studies are in Penang and Kelantan coastal waters, both of these coastal areas are the subject of substantial water quality, originating from rivers and various human activities. It confirms that the study areas can be classified into Case 2 waters.

The different types of data formats (e.g., single and multitemporal satellite images, in situ spectral reflectance and ground truth data) are examined in numerous transformations such as simple regression (SLR) model multivariate regression model (MVRM) optical model (OM) and NN model. This study only exploits two types of satellite images: the Landsat TM 30-m and the AVNIR-2 10-m resolution image, where overpasses both satellite was concurrent with field data collection.

The new approach is the optimization of NN model from multi-temporal remote data as input. The architecture and training parameters are optimised with designing effective as possible the input variables and hidden layer units, thus resulting model with high predictive power. In other sections, classical regression analysis and optical model often employed to assess coastal waters were also enhanced the accuracy of the estimation through the selection of the input variables, based on detecting less-multicollinearity between independent bands by methods of examining the VIF.

1.5 Research novelties

Many studies have proposed algorithms between satellite data and in-situ water constituents from ocean colour in Case 2 waters using regression models and optical models. In this study, we demonstrated the algorithms of using time series Landsat TM, ALOS/AVNIR-2 and in-situ spectral reflectance data to characterise the dynamics of ocean colour products in the Penang and Kelantan coastal waters. The novelty of this study is that we developed optimised NN model to determine C_s, C_{chl}, TURB and SDD based on examining potentially less-multicollinearity between independent bands. In generally, initial determinations of less-multicollinearity between independent bands have provided the ability improving the accuracy of the estimates.

The developed NN algorithms are relatively different in the input variable selection when compared with another NN algorithms used in remote sensing for ocean colour studies in Case 2 waters. In another developed NN models, there are no such assumptions made regarding the structure of the NN model, in which the input variables are selected based on available data from the remote sensor. Therefore, this

optimization method can be considered as a novelty approach to determine the water constituents from ocean colour in Case 2 waters of using remote sensing data (satellite sensor or reflectance spectra) due to the input variables to the NN architecture selected based on CA or PCA.

1.6 Outline of the thesis

The discussion of research in this thesis is well-organized in seven chapters. This research focuses on developing multiple ocean colour algorithms of Case 2 waters. Different instruments were used to extract the C_s, C_{chl}, TURB and NTU. This case study is focused on two study areas that are Penang and Kelantan coastal areas. Because complexity the results and discussion of this study, therefore this structure are made in three chapters.

Chapter 1 introduces the fundamental aspect of the research, research scopes, and objectives, including the thesis outline. This chapter sets up a preliminary understanding of the research activities and a brief methodological perspective. The description of the study area is given in Chapter 3.

Chapter 2 presents a literature review of previous research regarding ocean colour remote sensing. This includes the application theories of optical oceanography for the remote sensing system, such as IOPs and their relation to the water constituents. This chapter also explains the algorithms that have been developed for ocean colour products, especially for the coastal area or Case 2 type waters.

Chapter 3 elaborates on the stages of implementation that have been applied for the entire research. It begins with the methods used to collect data, type of data used, supporting equipment, the procedure of implementation and the research steps. Also, this chapter describes the variables which are included in the research and

research analysis. Experimental methodology, including equipment, pre-processing, developed algorithms and a flow chart of the research steps. The chapter continued by elucidating the distinctive approaches in executing the research objectives.

In Chapter 4, retrieval of ocean colour products from every single and multitemporal remote sensing data and in-situ observation is discussed. This chapter starts with the characteristics analysis of various C_s and C_{chl}. This chapter is to develop empirical algorithms for estimating C_s and C_{chl} in the Penang coastal water from Landsat TM and AVNIR-2 data.

Chapter 5 focuses on the estimation of the TURB and SDD from simulated satellite data (e.g., Landsat TM and AVNIR-2) by using different algorithms. This chapter is to optimise the combination of bands which could aid the quantification of TURB and SDD in Case 2 waters using Landsat TM and AVNIR-2 data.

Chapter 6 discusses the results of retrieving C_s and C_{chl} using model-based approaches from single and multitemporal satellite data. Next, the best retrieval algorithms using multitemporal data are applied to the Penang coastal area and comparing the performance both of them in mapping the distribution of C_s and C_{chl} is described in this chapter. Finally, chapter 7 concludes this thesis with a discussion of the main results and finding of all previous chapters and possible suggestions for future research and developments.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The colour of the water is one of the physical phenomena which occur because the processes of the scattering and absorption of visible light by pure water, inorganic and organic, particulate and dissolved, as well as the material are present in the water (Shubha Sathyendranath, 2000) The researchers recognise that these substances can vary independently of each other in Case 2 waters, thus, must consider each of the separations. In addition, the colour of Case 2 waters can be affected by bottom reflectance in the shallow area, clear waters.

The objective of this chapter is to present a literature review related to remote sensing of optically complex waters which is usually encountered in coastal regions. Water masses of this kind are often appointed as Case 2 waters. Also, much intention is given to discussing the principal and basic concepts to understand how to express the remotely detected signal as a function of the concentrations of the various substances present in the water column which is strongly related to this research. Finally, an overview of approaches and algorithms that are recently developed to derive relevant oceanic quantities from satellite ocean colour data from Case 2 waters is described. This section provides a comprehensive overview of the two major groups of algorithms: Empirical approaches and model-based approaches.

2.2 Ocean colour

One of the natural phenomena that we often encounter around us is ocean colour (Figure 2.1). Physically, the ocean looks blue is due to absorption and

scattering of light. It is similar to the scattering phenomenon in the sky, but for the clear ocean, water absorption is the much larger aspect than the scattering.

Figure 2.1. Ocean hue is due to the phenomenon of the absorption and scattering of sunlight (Source: http://science.nasa.gov).

Red and other parts of the visible light are more strongly absorbed by clear water when compared to the blue part, the red light and others are absorbed rapidly in the ocean leaving blue. In details, the longest wavelengths of visible light such as red, yellow and green are absorbed by water molecules in the ocean. When sunlight reaches the ocean, some of the light parts are redirected directly, but most of it penetrate to the ocean surface and interacts with the water molecules that encounters. Red, orange, yellow and green wavelengths of light are absorbed so that the remaining light is structured of the shorter wavelength blues and violet.

Although almost all sunlight that enters the ocean is absorbed, however, this phenomenon is not entirely applicable in the coastal area. In the coastal area, the scattering of light will increase due to the presence any suspended particles. Runoff

from rivers, which convey a high concentration of mineral suspended solid and nutrient, resuspension of sand and silt from the bottom by waves, tides, and the wind and a number of another substance, which is alter the ocean colour of the coastal areas. On the other hand, some particles such as microscopic living plant matter (commonly referred to as phytoplankton or algae) also contain the substances that absorb the certain wavelength of light, which can change its characteristic. A detail description of absorption of light by oceans with various types of the particles and other components can be found in (Wozniak & Dera, 2007).

2.3 Optical oceanography

The solutes and particles of natural waters (fresh and saline) are both optically essential and very variable in type and concentration. As a consequence, the optical properties of natural waters show larger temporal and spatial variations and rarely resemble pure water. Huge variations in the optical properties of natural waters are the main challenge to researchers who desire information about the aquatic ecosystem. However, the merging between constituent properties and optical properties presented that optical measurements can be used to infer information about this ecosystem. Definitely, the relationship between optical properties and chemical, biological and geological constituents of natural waters has asserted the important role of optics in aquatic research.

In a large scale, optical properties of water are divided into two mutually exclusive classes: the inherent and apparent. The inherent optical properties (IOPs) are the properties of the medium itself that do not affect by ambient light field. IOPs are the underlying parameters of hydrological optics. The inherent optical properties commonly vary in two terms, the absorption coefficient (coefficient $a(\lambda)$), which

controls the exponential degree of deterioration of flux per unit of path length of light in the medium, and per unit incident flux, because the absorption process, and the scattering coefficient (coefficient $b(\lambda)$), which defines equally of the exponential rate of the flux of deterioration due to scattering (Bhatti, 2008; Froidefond et al., 2002).

The apparent optical properties (AOP's) are those properties that depend both on the IOP's and on the light field in which they are measured (Mobley, 1994). The most widely-used AOPs are the irradiant reflectance, the remote sensing reflectance, and the attenuation coefficients. The used AOP related to the study are described below. In the ocean colour remote sensing, the integration of in situ data, optical models and remotely sensed data have been widely employed by researchers to develop the relationship between water quality parameters and the optical properties of water bodies. Therefore, it is important to discuss the theory and measurement of inherent and apparent optical properties of oceanic waters.

2.3.1 Inherent optical properties

Based on the earlier description, it is clear that the IOPs determine the optical properties of natural waters in a form suited to the needs of radiative transfer theory. Thus, IOPs are defined here. When a photon interacts with matter one of two things can happen: The absorption and scattering. The absorption is the processes when the energy of a photon being converted to another form such as heat. Scattering is the processes when the energy of the photon changes direction. Both phenomena can be deduced mathematically as follow:

Consider a small volume ΔV of water, of a thickness Δr, illuminated by a collimated beam of monochromatic light of spectral radiant power $\Phi_{my}(\lambda)$ in unit W nm^{-1}, schematically depicted in Figure 2.2. Some part of the incident power $\Phi_i(\lambda)$ is

absorbed within the volume of water. Some part $\Phi_s(\lambda,\psi)$ is scattered out of the beam at an angle ψ, and the remaining power $\Phi_t(\lambda)$ is transmitted through the volume with no change in direction. Let $\Phi_s(\lambda)$ be the total power that is scattered into all directions.

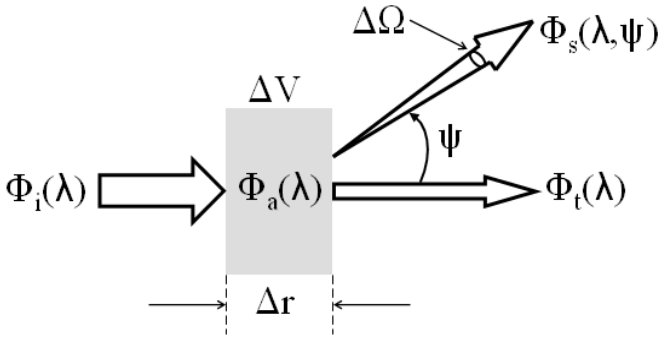

Figure 2.2. The geometry used to define inherent optical properties (Source: Mobley, 1994, 2004).

Moreover, by assuming that there are no photons undergo a change in wavelength during the scattering process. By the law of conservation of energy,

$$\Phi_i(\lambda) \equiv \Phi_a(\lambda) + \Phi_s(\lambda) + \Phi_t(\lambda) \tag{2.1}$$

The spectral absorptance $A(\lambda)$ is the fraction of incident power that is absorbed within the volume:

$$A(\lambda) = \frac{\Phi_a(\lambda)}{\Phi_i(\lambda)} \tag{2.2}$$

In the same manner, the spectral scatterance $B(\lambda)$ is the fractional part of the incident power that is scattered out of the beam in all directions,

$$A(\lambda) = \frac{\Phi_s(\lambda)}{\Phi_i(\lambda)} \tag{2.3}$$

and the spectral transmittance $T(\lambda)$, the fraction of the incident power that passes through the volume without interacting with the medium, is

$$T(\lambda) = \frac{\Phi_t(\lambda)}{\Phi_i(\lambda)} \tag{2.4}$$

$$A(\lambda) + B(\lambda) + T(\lambda) = 1 \tag{2.5}$$

The Inherent optical properties generally apply in optical hydrology is the spectral absorption and scattering coefficients, which are respectively the spectral absorptance and scatterance per unit distance in the medium. Based on Figure 2.2, and assuming that the thickness Δr is approaching zero. The absorption coefficient $a(\lambda)$ is then defined as:

$$a(\lambda) \equiv \lim_{\Delta r \to 0} \frac{\Delta A(\lambda)}{\Delta r} = \frac{dA(\lambda)}{dr} \qquad (m^{-1}) \tag{2.6}$$

Likewise the scattering coefficient b(λ) is defined as

$$b(\lambda) \equiv \lim_{\Delta r \to 0} \frac{\Delta B(\lambda)}{\Delta r} = \frac{dB(\lambda)}{dr} \qquad (m^{-1}) \tag{2.7}$$

The spectral beam attenuation coefficient $c(\lambda)$ is defined as

$$c(\lambda) \equiv a(\lambda) + b(\lambda) \tag{2.8}$$

The angular distribution of the scattered power, with $B(\psi, \lambda)$ being the fraction of incident power scattered out of the beam through an angle θ into a solid angle ΔΩ centred on θ, as shown in Figure 2.2. The angle ψ is called the scattering angle; its values lay in the interval $0 \leq \psi \leq \pi$. Then the angular scatterance per unit distance and unit solid angle $\beta(\psi, \lambda)$, is:

$$\beta(\psi, \lambda) = \lim_{\Delta r \to 0} \lim_{\Delta \Omega \to 0} \frac{\Phi_s(\psi, \lambda)}{\Phi_i(\lambda) \Delta r \, \Delta \Omega} \tag{2.9}$$

The spectral power scattered into the given solid angle ΔΩ is the spectral radiant intensity scattered into direction ψ times the solid angle:

$$\Phi_s(\psi, \lambda) = I_s(\psi, \lambda) \Delta \Omega \tag{2.10}$$

Moreover, if the incident power $\Phi_i(\lambda)$ falls on an area ΔA, then the corresponding incident irradiance is:

$$E_i(\lambda) = \Phi_i(\lambda)\Delta A \tag{2.11}$$

And $\Delta V = \Delta r \Delta A$ is the volume of water that is illuminated by the incident beam. The expression becomes

$$\beta(\psi,\lambda) = \lim_{\Delta r \to 0} \frac{I_s(\psi,\lambda)}{E_i(\lambda)\Delta V} \tag{2.12}$$

This form of $\beta(\psi, \lambda)$ suggests the name, spectral volume scattering function and the physical interpretation of scattered intensity per unit incident irradiance per unit volume of water Integrating $\beta(\psi, \lambda)$ over all directions (solid angles) gives the total scattered power per unit incident irradiance and unit volume of water, in other words the spectral scattering coefficient;

$$b(\lambda) = 2\pi \int_0^{\pi} \beta(\psi,\lambda) \sin\psi \, d\psi \tag{2.13}$$

This integration is often divided into forward scattering, $0 \leq \psi < \pi/2$, and backward scattering, $\pi/2 \leq \psi \leq \pi$ parts.

$$b = b_b + b_f \tag{2.14}$$

$$b_f(\lambda) = 2\pi \int_0^{\pi} \beta(\psi,\lambda) sin\psi \, d\psi \tag{2.15}$$

$$b_b(\lambda) = 2\pi \int_0^{\pi} \beta(\psi,\lambda) sin\psi \, d\psi \tag{2.16}$$

2.3.2 Apparent optical properties

One of the main goals of optical oceanography is to learn something about the constitution of water, such as the concentration of suspended sediment and chlorophyll of optical measurement. This means, by measuring the absorption coefficient and volume scattering function, then the information about the optical

properties of water bodies will be obtained, because the IOP_S can provide information about the type and concentration of water constituents. However, in the early days of optical oceanography, is it difficult to measure in situ IOP_S. On the other hand, it is relatively easy to measure radiometric variables such as upwelling and downwelling irradiances plane. This led to the use of the optical properties (AOPs) more dominant than IOP_S in describing the optical properties of most of the water bodies. Appropriate AOPs can provide useful information about the bodies of water, for example, the type and concentration of water constituents, from simple measurements, were made in the light field.

2.3.2.1 Irradiance reflectance

In generally, the reflection characteristic of earth surface features is the dimensionless ratio of the radiance emittance of an object and the irradiance. When measured as a function of wavelength it is called the spectral reflectance, and defined mathematically as (Lillesand et al., 2014):

$$\rho(\lambda) = \frac{E_R(\lambda)}{E_I(\lambda)} \times 100 \qquad (2.17)$$

where the ρ (λ) is expressed as a percentage, E_R is the energy of wavelength reflected from the object, and E_I is the energy of wavelength incident upon the object. A graph of spectral reflectance of an object is called the spectral reflectance curve. This configuration spectral curve gives information about the spectral characteristics of the object. It is necessary in the determination of the wavelength region (s) in which the remotely sensed data is acquired for a particular application.

In the water bodies, reflectance *R* is treated as an apparent optical property (AOP). Irradiance reflectances commonly used AOP because they are essential to remote sensing of the oceans. Algorithm developed to correlate irradiance reflectance

R to quantities such as suspended sediment and chlorophyll concentration (Gordon & Morel, 2012; M. Zhang et al., 2010). In ocean colour remote sensing, the spectral irradiance reflectance is symbolized as (Mobley, 1994), is defined as the ratio of the spectral upwelling to downwelling plane irradiances:

$$R(z;\ \lambda) = \frac{E_u(z;\ \lambda)}{E_d(z;\lambda)} \tag{2.18}$$

$R\ (z;\lambda)$ is often assessed in the water just below the surface and denote this depth by $z = w = 0$. The illustrated of irradiance reflectance is shown in Figure 2.3.

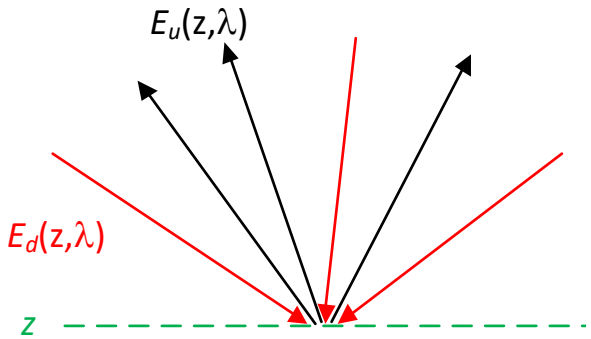

Figure 2.3. Illustration of irradiance reflectance (Redrawn from: http://www.oceanopticsbook.info).

Another reflectance used for remote sensing of ocean colour is the remote sensing reflectance (R_{rs}). More lately, the R_{rs} has become the AOP of another alternative for remote sensing of ocean properties (Mobley et al., 2005; O'Reilly et al., 1998). The spectral remote-sensing reflectance R_{rs} is defined as:

$$R_{rs}(\theta, \phi, \lambda) = \frac{\mathrm{L}(\theta, \phi, \lambda)}{\mathrm{E_d}(\theta, \phi, \lambda)} \quad (\mathrm{Sr}^{-1}). \tag{2.19}$$

R_{rs} (as illustrated in Figure 2.4) is evaluated using L and E_d in the air just above the surface of the water, and L is often indicated as the water leaving radiance. The reason that R_{rs} has replaced R for remote sensing is due to less sensitive to the environmental conditions such as sun angle or sky condition. However,

determination of R_{rs} is more complicated than R. It requires another parameter and different sensor, which must be accurately calibrated (Froidefond et al., 1999; Mobley et al., 2011).

Although Rrs is currently used for most remote sensing application of ocean colour, yet, is still commonly measured in water studies and is widely used. For example, inverting R spectra to acquire absorption and scattering properties (Roesler & Perry, 1995).

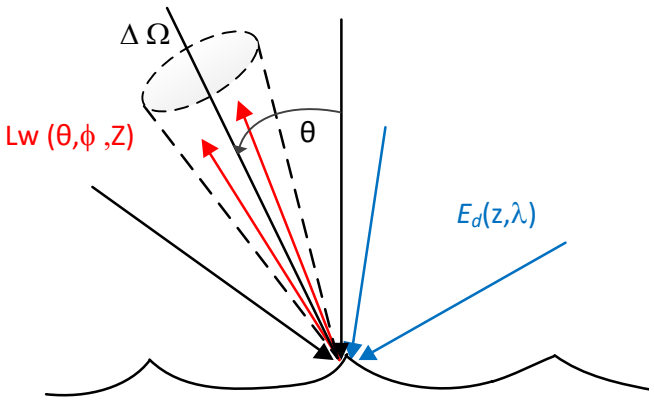

Figure 2.4. Illustration of remote sensing reflectance (Redrawn from: http://www.oceanopticsbook.info).

2.4 Optical constituents of the ocean

It is important to have some information on the composition of those waters. Natural waters contain a constant size distribution of particles starting from water molecules of size ~0.1 nm, to small organic molecules of size ~1 nm, to large organic molecules of size ~100 nm, to a huge size such as whales and submarines of sizes ~10 m to ~100 m. Thus, clearly, the water is composed fully of particles. However, the constituents of natural waters have traditionally been divided into "dissolved" and "particulate" matter, of organic and inorganic origin, living and non-living (Mobley, 1994).

Constituents in water, which affects its optical properties, are traditionally grouped into the following: (1) Pure water, both fresh and saline (water + inorganic dissolved material). (2) Micro-plant (phytoplankton). (3) Colour dissolved organic material (CDOM). (4) Non-phytoplankton organic particles (sometimes referred to as detritus or tripton). (5) Inorganic particles, and (6) Bubbles. Following below three of six constituents in water that affect its optical properties are discussed in this study.

The IOPs mentioned above have provided an understanding of the aquatic environment and implies that optical measurements can be used to deduce information of optically active constituents. The water column is assumed to be homogeneous in terms of optical properties. Optical remote sensing is likely to become quite challenging when the optical properties of the upper water column significantly known. The method discussed and applied here relies on the concepts of model an inversion technique, leading to analytical algorithms.

2.5 Statistical techniques

Input variable selection (IVS) methods are important in finding the optimal functional form of statistical models. The task of choosing the input variables is generally in the development of all statistical models and is largely dependent on the discovery of relationships within the data available to identify the appropriate predictors of the model output. However, in the case of ANN, and other similar data-driven statistical modelling approaches, there are no such assumptions made regarding the structure of the NN model. In contrast, the IVS and developing models based on the available data (May et al., 2011).