Retrieval of Oceanic Constituents with Artificial Neural Network Based on Radiative Transfer Simulation Techniques

by

Tinglu Zhang

Dissertation zur Erlangung des akademischen Grades des Doktors der Naturwissenschaften am Fachbereich Geowissenschaften der Freien Universität Berlin

Berlin, September 2003
1. Gutachter: Prof. Dr. Jürgen Fischer
2. Gutachter: Prof. Dr. Hermann Kaufmann

Tag der Disputation: 29 October, 2003
ACKNOWLEDGEMENTS

I am most fortunate to have had a chance to carry out my thesis project under the supervision of Prof. Dr. Jürgen Fischer. I am grateful to him not only providing me with invaluable guidance and suggestions, but also a top-notch research environment.

I would like to express my thanks here to Dr. Frank Fell who has helped me to finish this thesis. He not only gave me invaluable guidance and suggestions throughout my study, but also gave me many invaluable comments and suggestions for the draft of this thesis, as well as improved on language for it.

I also thank Dr. Rene Preusker for allowing me to use his powerful ANN program, as well as for his many helps in computer.

I thank Thomas Schröder for his helps in atmospheric correction.

My special thanks go to Dr. Michael Schaal e and Ms. Gisela Braß for helping me in managing my living. Due to their helps, I have a quiet, comfortable, convenient living environment.

I give my heartfelt thanks to all colleagues who have helped me in my study and living.

I thank the following institutions for providing me with the financial support: the China Scholarship Council (May 1999–April 2000), BMBF (Aug. 2001 – June 2003), and Institute for Space Sciences of Free University Berlin (July 2003 – Oct. 2003).

I also acknowledge all the people and institutes who collected the data sets COASTLOOC, SeaBAM, and PMNS which I have used in this thesis.

Finally, I give my thanks to my wife Hongyan Qu and my son Shutian Zhang for their support and understanding during my study.
ABSTRACT

In this thesis, a method for the retrieval of oceanic constituents from ocean colour in Case I and Case II waters is reported. The method is derived from radiative transfer simulations and subsequent application of Artificial Neural Network (ANN) techniques. Three applications of this method are presented in this thesis. Firstly, an ANN-based algorithm is developed for the retrieval of the pigment concentration in Case I waters from the remote sensing reflectance just above sea surface. The performance of the algorithm is assessed by comparing it to the in-situ measurement data sets SeaBAM and COASTLOOC. The results show that the performance of the ANN-based retrieval scheme is comparable to the most successful empirical algorithms such as OC4. Secondly, an ANN-based algorithm is developed for the retrieval of oceanic constituents concentrations (CHL, SPM and CDOM) in Case II waters from the hemispherical reflectance just below sea surface. The performance of the algorithm is assessed by comparing it to the in situ measurement data sets COASTLOOC, PMNS. The results show that the performance of the ANN-based retrieval scheme is better than that of the empirical algorithms developed by PMNS. Thirdly, an ANN-based algorithm is developed for the retrieval of oceanic constituents concentrations (CHL, SPM and CDOM) in Case II waters from MERIS imagery. This algorithm has the capability to deal with various atmospheres from weakly to strongly absorbing aerosols. Applying this algorithm to MERIS images taken over the North Sea and the China Seas, reasonable results were obtained except for the highly turbid areas in the China Seas.
# TABLE OF CONTENTS

TABLE OF CONTENTS .................................................................................................. I  
LIST OF TABLES ......................................................................................................... IV  
LIST OF FIGURES ....................................................................................................... VI  
LIST OF SYMBOLS ...................................................................................................... XI  
LIST OF ACRONYMS ................................................................................................ XIII  

CHAPTER 1 INTRODUCTION ...................................................................................... 1  
1.1. Background ............................................................................................................ 1  
1.2. Retrieval of Oceanic Constituents from Ocean Colour Measurements  
at Sea Level .............................................................................................................. 3  
1.2.1. The Forward Problem ...................................................................................... 3  
1.2.2. The Inverse Problem ........................................................................................ 4  
1.3. Retrieval of Oceanic Constituents from Ocean Colour Measurements Taken at  
Top of Atmosphere ................................................................................................. 6  
1.4. Objectives and Outline ......................................................................................... 7  

CHAPTER 2 THEORETICAL BACKGROUND ................................................................ 9  
2.1. Concepts and Definitions ....................................................................................... 9  
2.1.1. Radiometric Concepts ...................................................................................... 9  
2.1.2. Inherent Optical Properties .............................................................................. 9  
2.1.3. Apparent Optical Properties ........................................................................... 11  
2.2. Optical Properties of Sea Water ........................................................................... 11  
2.2.1. Pure Seawater ................................................................................................ 12  
2.2.2. Phytoplankton ................................................................................................ 12  
2.2.3. Suspended particulate matter ......................................................................... 15  
2.2.4. CDOM (Yellow Substance) ........................................................................... 17  
2.2.5. Classification of Waters ................................................................................. 18  
2.3. Radiative Transfer Simulations ............................................................................ 18  
2.4. Artificial Neural Network .................................................................................... 20  

CHAPTER 3 RETRIEVAL OF THE PIGMENT CONCENTRATION  
IN CASE I WATERS ................................................................................................... 22  
3.1. Introduction .......................................................................................................... 22  
3.2. Data Sets ............................................................................................................... 23  
3.2.1. Training Data ................................................................................................ 23  
3.2.2. Validation Data: the SeaBAM Data Set ......................................................... 24  
3.2.3. Test Data: the COASTLOOC Data Set ......................................................... 26  
3.3. ANN-based Pigment Retrieval in Case I Waters ................................................. 26
3.3.1. Parameters affecting the performance of an ANN ........................................ 26
3.3.2. ANN Training ............................................................................................... 27
3.3.3. Determining ANN Architecture and Noise Adding for Optimal Pigment
      Retrieval ........................................................................................................ 28
3.3.4. Best ANN for Pigment Retrieval ................................................................. 29
3.4. Evaluating the Performance of the ANN-based Pigment Retrieval Algorithm ... 30
  3.4.1. Assessing the Performance of the Trained ANN .......................................... 30
  3.4.2. Comparison with Existing Empirical Algorithms ....................................... 32
  3.4.3. Resistance against Noise ............................................................................ 34
3.5. Conclusions ...................................................................................................... 35

CHAPTER 4 MODELLING THE BACKSCATTERING PROBABILITY OF
      MARINE PARTICLES IN CASE II WATERS .................................................. 38
  4.1. Background .................................................................................................... 38
  4.2. The COASTLOOC Data Set .......................................................................... 38
  4.3. Inherent Optical Properties of Oceanic Constituents in Case II Waters .......... 41
  4.4. Comparing in-situ Data and Simulations of the Marine Light Field Using
      Petzold’s Phase Function for Marine Particles .............................................. 42
  4.5. Modelling the Marine Particle Phase Function in Case II Waters ................. 43
  4.6. Discussion .................................................................................................... 48

CHAPTER 5 RETRIEVAL OF OCEANIC CONSTITUENTS IN CASE II
      WATERS ........................................................................................................... 51
  5.1. Introduction ................................................................................................... 51
  5.2. Data Sets ....................................................................................................... 52
  5.2.1. Training Data ............................................................................................ 52
  5.2.2. In-situ Measurement Data Sets ................................................................. 55
  5.2.3. Data Processing ....................................................................................... 55
  5.3. Retrieval of the Oceanic Constituents with ANN .......................................... 57
  5.3.1. Artificial Neural Network ......................................................................... 57
  5.3.2. ANN Training .......................................................................................... 58
  5.3.3. Determining ANN Architecture and Noise Adding for Optimal Oceanic
         Constituent Retrieval .................................................................................. 58
  5.4. Evaluating the Performance of the ANN-based Oceanic Constituents Retrieval
       Algorithms .................................................................................................... 63
  5.4.1. Assessing the Performance of the Trained ANN ....................................... 63
  5.4.2. Comparison with Existing Retrieval Algorithms ..................................... 68
  5.5. Conclusions .................................................................................................. 69

CHAPTER 6 RETRIEVAL OF OCEANIC CONSTITUENTS CONCENTRATIONS
      IN CASE II WATERS FROM MERIS DATA .................................................. 70
  6.1. Background .................................................................................................. 70
6.2. Simulated Data Sets ............................................................................................. 70
  6.2.1. Atmosphere .................................................................................................... 70
  6.2.2. Water ............................................................................................................. 73
  6.2.3. Sea Surface State .......................................................................................... 74
  6.2.4. Ranges of the Related Parameters ................................................................. 74
  6.2.5. RT Simulations .............................................................................................. 75
  6.2.6. Creation of ANN Training Data Set .............................................................. 76

6.3. Artificial Neural Networks ................................................................................... 77
  6.3.1. Structure of the ANN for Retrieval of the Constituents ................................ 77
  6.3.2. Performance of the ANN-based Algorithm with Simulated Data ................. 77

6.4. Application to the MERIS Imagery ..................................................................... 84

6.5. Conclusion ............................................................................................................ 92

CHAPTER 7 SUMMARY ……………………………………………………...……..93

7.1. Conclusion ............................................................................................................ 93

7.2. Future Work ......................................................................................................... 93

REFERENCES ............................................................................................................... 95
## LIST OF TABLES

3.1 Parameterisations of inherent optical properties of pure sea water and constituents of Case I waters ................................................................. 24

3.2 Data sources and characteristics of the SeaBAM data set, reproduced after *O’Reilly et al.* [1998] ............................................................................................. 25

3.3 Regional distribution of the COASTLOOC stations located in Case I waters ...... 26

3.4 Performance with regard to pigment retrieval of spectral band combinations used as input to ANNs. The performance is given for the most successful of the 16 ANNs trained for each combination of input parameters, characterised by the number of neurons in the hidden layer and the noise level .......................................................................................................... 28

3.5 Empirical ocean colour algorithms based on remote sensing reflectance .......... 33

3.6 Performance of the ANN-based pigment retrieval scheme as compared to selected empirical algorithms when applied to the SeaBAM and COASTLOOC data sets ............................................................................................ 33

3.7 Performance against noise of the ANN-based pigment retrieval scheme as compared to empirical algorithms. All algorithms were applied to SeaBAM remote sensing reflectance ratios to which different levels of noise have been added. The following acronyms are used: POL = POLDER, CCO = CalCOFI two-band cubic, MO3 = MOREL-3, MO4 = MOREL-4, OC2, OC2B and OC4 = Ocean Chlorophyll algorithms 2, 2B and 4 ................................................ 35

4.1 Information on the COASTLOOC data set ............................................................. 40

4.2 Parameterisation of inherent optical properties of pure sea water and constituents of Case II waters ................................................................. 42

4.3 Inputs for Mie calculations ...................................................................................... 49

5.1 Ranges of oceanic constituent concentrations ......................................................... 52

5.2 Characteristics of *in situ* data sets ....................................................................... 56

5.3 Performance with regard to pigment retrieval of spectral band combinations used input to ANNs (Case II waters) ......................................................... 59
5.4 Performance with regard to SPM retrieval of spectral band combinations used input to ANNs (Case II waters) ................................................ 60
5.5 Performance with regard to CDOM retrieval of spectral band combinations used input to ANNs (Case II waters) ................................................ 61
5.6 Performance of the selected optimal ANNs for the retrievals of three oceanic constituents ..............................................................64
5.7 (a) Performance of the ANN-based pigment retrieval algorithms as compared to the PMNS algorithm .............................................................68
5.7 (b) Performance of the ANN-based SPM retrieval algorithms as compared to the PMNS algorithm .............................................................68
5.7 (c) Performance of the ANN-based CDOM retrieval algorithms as compared to the PMNS algorithm .............................................................68
6.1 Aerosol models used in this study.............................................................71
6.2 Size distributions of various aerosols used in this study .........................72
6.3 Ranges of oceanic constituents ...............................................................74
6.4 Ranges of atmospheric components .......................................................75
6.5 Ranges of geometric and other parameters.............................................75
6.6 Spectral channels and their bandwidth used in this study .......................75
6.7 The noise level added to training data inputs ............................................76
6.8 Performance of the ANN-based algorithm for the retrieval of the oceanic constituents as well as the optical thickness of aerosols .........................79
6.9 Aerosol mixtures for testing the performance of the trained ANN .............84
LIST OF FIGURES

1.1 Sketch of different origins of light received by space-borne sensor .........................2

2.1 Absorption coefficient of pure seawater ($\lambda=380$-709 nm from 
Pope and Fry (1997), $\lambda > 709$ nm from Hale and Querry (1973)) ......................13

2.2 Scattering coefficient of pure seawater, from Morel (1974) .................................13

2.3 Phase function of pure seawater, from Morel (1974) ............................................14

2.4 Absorption coefficient of endogenous particles for CHL=1 mg m$^{-3}$, based on (2.18) ........................................................................................................15

2.5 Absorption coefficient of exogenous particles for SPM=3 g m$^{-3}$, 
based on Eq. (2.22) and (2.23) .................................................................................16

2.6 Measured by Petzold (1972) volume scattering function from 
three natural waters and derived by Mobley et al. (1993) 
particles phase function, all at $\lambda=514$ nm ..........................................................17

2.7 Absorption coefficient of CDOM, for $a_g(440)=1.0$ m$^{-1}$ and $s_f=0.017$ ...............18

2.8 Architecture of the Multi-Layer-Perceptron used in this thesis .........................21

3.1 Scatter plot showing the performance of the ANN for the synthetic 
training data set. The target pigment concentration designates 
the pigment concentration used as input to the RT simulations. 
The dashed lines indicate the factor 2 error margin ....................................................31

3.2 Comparison of ANN-derived pigment concentration to 
Corresponding in-situ measurements for the SeaBAM (A) 
and COASTLOOC (B) data sets. The dashed lines indicate the 
factor 2 error margin .................................................................................................32

3.3 Performance of the ANN-based pigment algorithm presented 
in this study as compared to selected empirical algorithms 
from O'Reilly et al. [1998], using square of correlation coefficient 
$r^2$ (A) and root mean square error RMSE (B) of derived vs. actually measured pigment concentration as measure of success. 
The algorithms were applied to the SeaBAM (x) and 
COASTLOOC (◊) data sets ....................................................................................34
3.4 Performance against noise of the ANN-based algorithm presented in this study and selected empirical algorithms, using square of correlation coefficient $r^2$ (A) and root mean square error RMSE (B) of derived vs. measured pigment concentrations as measure of success ..........................................................36

4.1 The locations of stations visited during COASTLOOOC cruises, adapted from Babin (2000) ..................................................39

4.2 Comparison between the simulated hemispherical reflectances just below the sea surface and the corresponding COASTLOOOC data. The simulations have been performed assuming a constant back scattering probability of 1.81% ..................................................44

4.3 The model for the back scattering probability of marine particles in Case II waters as function of the ratio of POM to SPM and wavelength........45

4.4 The relationship between $a_y(443)+a_{ph}(443)$ and the concentration of particulate organic matter ..........................................................46

4.5 The model for the back scattering probability of marine particles in Case II waters as function of the ratio of $a_y(443)+a_{ph}(443)$ to SPM and wavelength ..........................................................46

4.6 Comparisons between the simulated hemispherical reflectances and the corresponding COASTLOOOC data. The simulations have been performed with the new model for the back scattering probability of marine particles proposed in this study ...........................................47

4.7 Comparison of statistical parameters between the simulations with the new model and 1.81% back scattering probability ..................................48

4.8 Comparison between the new model proposed in this study and the Mie calculations ...........................................................................50

5.1 Scatter plots of SPM against CHL (top), CDOM against CHL (middle), and CDOM against SPM (bottom). Data set collected during COASTLOOOC cruises ..................................................54

5.2 Scatter plot showing the performance of the ANN-based pigment
retrieval algorithm for the synthetic training data set (A), validation
data set (B), and test data (C). The dashed lines indicate the factor 2
error margin ........................................................................................................65

5.3 Scatter plot showing the performance of the ANN-based SPM
retrieval algorithm for the synthetic training data set (A), validation
data set (B), and test data (C). The dashed lines indicate the factor 2
error margin ........................................................................................................66

5.4 Scatter plot showing the performance of the ANN-based CDOM
retrieval algorithm for the synthetic training data set (A), validation
data set (B), and test data (C). The dashed lines indicate the factor 2
error margin ........................................................................................................67

6.1 Phase functions of the aerosol models used in this study ......................73

6.2 Normalised attenuation coefficients of the aerosol models used in this study ....73

6.3 Single scattering albedo of the aerosol models used in this study .............74

6.4 Scatter plot showing the performance of the ANN-based algorithm for
the Rayleigh scattering correction. (A), (C) and (E) for the training data set
at three wavelengths: 412 nm, 560 nm and 885 nm, respectively. (B), (D)
and (F) for test data set at the same wavelengths .............................................78

6.5 Scatter plot showing the performance of the ANN-based algorithm
for pigment retrieval: training data set (top left panel, TRAIN), test
data sets for seven different atmospheres (TEST1~TEST7). The
target pigment concentration designates the pigment concentration used
as input to the RT simulations. The dashed lines indicate the factor 3
error margin ........................................................................................................80

6.6 Scatter plot showing the performance of the ANN-based algorithm
for SPM retrieval: training data set (top left panel, TRAIN), test data
sets for seven different atmospheres (TEST1~TEST7). The target
SPM concentration designates the SPM concentration used as input
to the RT simulations. The dashed lines indicate the factor 3 error margin ..........81

6.7 Scatter plot showing the performance of the ANN-based algorithm
for CDOM retrieval: training data set (top left panel, TRAIN), test
data sets for seven different atmospheres (TEST1~TEST7). The
target CDOM absorption coefficient designates the CDOM absorption coefficients used as input to the RT simulations. The dashed lines indicate the factor 3 error margin .......................................................................................82

6.8 Scatter plot showing the performance of the ANN-based algorithm for aerosol optical thickness retrieval: training data set (top left panel, TRAIN), test data sets for seven different atmospheres (TEST1~TEST7). The target aerosol optical thickness designates the aerosol optical thickness used as input to the RT simulations. The dashed lines indicate the factor 2 error margin .....................................................................................................83

6.9 RGB image of MERIS over the North Sea (22 March, 2003). The black line indicates the transect used for further analysis ...........................................................................86

6.10 Pigment distribution derived from MERIS imagery for the North Sea (22 March, 2003) .................................................................................................................86

6.11 SPM distribution derived from MERIS imagery for the North Sea (22 March, 2003) .................................................................................................................87

6.12 CDOM distribution derived from MERIS imagery for the North Sea (22 March, 2003) .................................................................................................................87

6.13 Aerosol distribution derived from MERIS imagery for the North Sea (22 March, 2003) .................................................................................................................88

6.14 Distributions of pigment, SPM, CDOM and aerosol derived from MERIS imagery for the North Sea along the specific transect (22 March, 2003) .................................................................................................................88

6.15 RGB image of MERIS over the China Seas (16 April, 2003). The black line indicates the transect used for further analysis .........................................................89

6.16 Pigment distribution derived from MERIS imagery for the China Seas (16 April, 2003) .................................................................................................................89

6.17 SPM distribution derived from MERIS imagery for the China Seas (16 April, 2003) .................................................................................................................90

6.18 CDOM distribution derived from MERIS imagery for the China Seas (16 April, 2003) .................................................................................................................90
6.19 Aerosol distribution derived from MERIS imagery for the China Seas
(16 April, 2003) ........................................................................................................91

6.20 Distributions of pigment, SPM, CDOM and aerosol derived from
MERIS imagery for the China Seas along the specific transect
(16 April, 2003) ........................................................................................................91
# LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>Inherent Optical Properties</strong></td>
</tr>
<tr>
<td>$a(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Absorption coefficient</td>
</tr>
<tr>
<td>$a_w(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Absorption coefficient of pure seawater</td>
</tr>
<tr>
<td>$a_y(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Absorption coefficient of yellow substance</td>
</tr>
<tr>
<td>$a_{nap}(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Absorption coefficient of non-chlorophyllous particles</td>
</tr>
<tr>
<td>$a_{ph}(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Absorption coefficient of phytoplankton</td>
</tr>
<tr>
<td>$a_p(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Absorption coefficient of particles</td>
</tr>
<tr>
<td>$b(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Scattering coefficient</td>
</tr>
<tr>
<td>$b_w(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Scattering coefficient of pure seawater</td>
</tr>
<tr>
<td>$b_p(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Scattering coefficient of particles</td>
</tr>
<tr>
<td>$b_0(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Back scattering coefficient</td>
</tr>
<tr>
<td>$\tilde{b}_b(\lambda)$</td>
<td>%</td>
<td>Back scattering probability, defined here as ratio of back scattering coefficient to total scattering coefficient</td>
</tr>
<tr>
<td>$c(\lambda)$</td>
<td>m$^{-1}$</td>
<td>Beam attenuation coefficient</td>
</tr>
<tr>
<td>$\omega_b(\lambda)$</td>
<td>1</td>
<td>Single scattering albedo</td>
</tr>
<tr>
<td>$\beta(\theta, \lambda)$</td>
<td>m$^{-1}$ sr$^{-1}$</td>
<td>Volume scattering function</td>
</tr>
<tr>
<td>$\tilde{\beta}(\theta, \lambda)$</td>
<td>sr$^{-1}$</td>
<td>Scattering phase function</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Radiometric Quantities</strong></td>
</tr>
<tr>
<td>$L(\lambda, \theta_s, \theta_v, \phi_v)$</td>
<td>W m$^{-2}$ sr$^{-1}$</td>
<td>Radiance</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Apparent Optical Properties</strong></td>
</tr>
<tr>
<td>$E(\lambda, \theta_s)$</td>
<td>W m$^{-2}$</td>
<td>Irradiance</td>
</tr>
<tr>
<td>$E_d(\lambda)$</td>
<td>W m$^{-2}$</td>
<td>Downward irradiance</td>
</tr>
<tr>
<td>$E_u(\lambda)$</td>
<td>W m$^{-2}$</td>
<td>Upward Irradiance</td>
</tr>
<tr>
<td>$R_{RS}(\lambda, \theta_s, \theta_v, \phi_v)$</td>
<td>sr$^{-1}$</td>
<td>Remote sensing reflectance</td>
</tr>
<tr>
<td>$R(\lambda, \theta_s)$</td>
<td>1</td>
<td>Irradiance reflectance</td>
</tr>
<tr>
<td>$Q(\lambda, \theta_s, \theta_v, \phi_v)$</td>
<td>sr</td>
<td>Factor describing the bidirectional character of the light field</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Atmosphere and aerosol properties</strong></td>
</tr>
<tr>
<td>$\tau_r(\lambda)$</td>
<td>1</td>
<td>Optical thickness due to Rayleigh scattering</td>
</tr>
<tr>
<td>Symbol</td>
<td>Unit</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>$\tau_\alpha(\lambda)$</td>
<td>1</td>
<td>Optical properties due to aerosol scattering and absorption</td>
</tr>
<tr>
<td>$RH$</td>
<td>percents</td>
<td>Relative humidity</td>
</tr>
<tr>
<td>$Chl$</td>
<td>mg m$^{-3}$</td>
<td>Pigment concentration, defined here as sum of chlorophyll and phaeopigment concentration</td>
</tr>
<tr>
<td>$SPM$</td>
<td>g m$^{-3}$</td>
<td>Suspended matter concentration</td>
</tr>
<tr>
<td>$CDOM$</td>
<td>m$^{-1}$</td>
<td>Absorption coefficient of Coloured dissolved organic matter</td>
</tr>
</tbody>
</table>

**Geophysical properties**

- $\theta_s$ degrees: Sun zenith angle
- $\theta_v$ degrees: Observation zenith angle
- $\phi_v$ degrees: Observation azimuth angle
- $\lambda$ nm: Wavelength

**Others**

- $p$ hPa: Pressure at sea level
- $w$ m s$^{-1}$: Wind speed
- $f$ 1: Proportionality factor
- $r$ 1: Pearson’s correlation coefficient
- $s_c$ 1: Sigmoidal function
- $c_t$ 1: Temperature constant in sigmoidal function
## LIST OF ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOP</td>
<td>Apparent Optical Properties</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>CDOM</td>
<td>Coloured Dissolved Organic Matter</td>
</tr>
<tr>
<td>CNES</td>
<td>Centre National d Etudes Spatiales</td>
</tr>
<tr>
<td>CNSA</td>
<td>China National Space Administration</td>
</tr>
<tr>
<td>COASTLOOC</td>
<td>Coastal Surveillance Through Observation of Ocean Colour</td>
</tr>
<tr>
<td>COCTS</td>
<td>China Ocean Colour and Temperature Scanner</td>
</tr>
<tr>
<td>CZCS</td>
<td>Coastal Zone Colour Scanner</td>
</tr>
<tr>
<td>DLR</td>
<td>German Aerospace Center</td>
</tr>
<tr>
<td>DOM</td>
<td>Dissolved Organic Matter</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>GLI</td>
<td>Global Imager</td>
</tr>
<tr>
<td>IOP</td>
<td>Inherent Optical Properties</td>
</tr>
<tr>
<td>ISPO</td>
<td>Indian Space Research Organization</td>
</tr>
<tr>
<td>KARI</td>
<td>Korean Aerospace Research Institute</td>
</tr>
<tr>
<td>MERIS</td>
<td>Medium Resolution Imaging Spectrometer (ESA)</td>
</tr>
<tr>
<td>MLP</td>
<td>Multi-Layer Perceptron</td>
</tr>
<tr>
<td>MISR</td>
<td>Multi-angle Imaging SpectroRadiometer</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectrometer (NASA)</td>
</tr>
<tr>
<td>MOMO</td>
<td>Matrix Operator Model</td>
</tr>
<tr>
<td>MOS</td>
<td>Modular Optoelectric Scanner</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>NASDA</td>
<td>National Space Development Agency of Japan</td>
</tr>
<tr>
<td>OCM</td>
<td>Ocean Colour Monitor</td>
</tr>
<tr>
<td>OCTS</td>
<td>Ocean Colour and Temperature Scanner</td>
</tr>
<tr>
<td>OSMI</td>
<td>Ocean Scanning Multispectral Imager</td>
</tr>
<tr>
<td>PMNS</td>
<td>Particulate Matter North Sea</td>
</tr>
<tr>
<td>POLDER</td>
<td>Polarization and Directionality of the Earth's Reflectances</td>
</tr>
<tr>
<td>POM</td>
<td>Particulate Organic Matter</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>RT</td>
<td>Radiative Transfer</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>RTC</td>
<td>Radiative Transfer Calculation</td>
</tr>
<tr>
<td>S-GLI</td>
<td>Second generation GLobal Imager</td>
</tr>
<tr>
<td>SeaBAM</td>
<td>SeaWiFS Bio-optical Algorithm Mini-Workshop</td>
</tr>
<tr>
<td>SeaBASS</td>
<td>SeaWiFS Bio-optical Archive and Storage System</td>
</tr>
<tr>
<td>SeaWiFS</td>
<td>Sea-viewing Wide Field-of-view Sensor (NASA)</td>
</tr>
<tr>
<td>SPM</td>
<td>Suspended Particulate Matter</td>
</tr>
<tr>
<td>TOA</td>
<td>Top of Atmosphere</td>
</tr>
<tr>
<td>VIIRS</td>
<td>Visible Infrared Imager Radiometer Suite</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

1.1. Background

Chlorophyll, suspended particulate matter and coloured dissolved organic matter are optically significant constituents in ocean which can be detected from ocean colour measurement. Chlorophyll as the most important photosynthetic pigment is an index of phytoplankton biomass which acts as the first link in the marine food chain. As a result, it plays a key role in the ecology of the marine ecosystem, and changes in their patterns of distribution and abundance can have significant impact on the entire ecosystem. In addition, Phytoplankton have a major role in the global carbon cycle [Falkowski, 1994]. During photosynthesis, phytoplankton remove carbon dioxide from sea water, and release oxygen as a by-product. Transport of suspended particulate matter (SPM) determines the evolution of the coastline, the deposition and erosion of the beaches and is thus a process of primary importance in coastal engineering. Besides, SPM is an important factor determining water quality. Its presence affects water quality by reducing the light available to aquatic vegetation and by providing a substrate for the transport of phosphate, ammonium, heavy metals, and some pathogenic bacteria [Luoma, 1989]. Coloured dissolved organic matter (CDOM) represents the optically active fraction of the bulk dissolved organic matter (DOM). CDOM in marine environment, especially in estuaries and the coastal area, where the concentration of CDOM is high, plays an important role in a number of biological and chemical processes [Mopper et al., 1991; Siegel et al., 1996, 2002; Moran et al., 1997], including global carbon cycling, functioning of microbial food webs, and penetration of sunlight into seawater.

In marine water studies, satellite remote sensing represents the most suitable technique for large-scale, long-term and continuous monitoring of bio-geochemical or physical parameters. Ocean colour remote sensing is an important technique to obtain the optical properties and oceanic constituents in the upper ocean layer. In the past twenty years, especially in the recent years, a number of ocean colour sensors have been launched [IOCCG, 2003]. The CZCS (1978-1986) is the earliest of all ocean-colour satellite sensors. Next, a series of increasingly-sophisticated instruments, such as MOS (DLR, Germany), OCTS (NASDA, Japan), POLDER (CNES, France), SeaWiFS (NASA, USA), MODIS (NASA, USA), MISR (NASA, USA), OCM (ISPO, India), GLI (NASDA, Japan), OSMI (KARI, Korea), COCTS (CNSA, China), MERIS (ESA, Europe), and POLDER-2 (CNES, France), have been launched between 1996 and 2002. More ocean colour instruments are scheduled to be launched in the future, such as S-GLI (NASDA, Japan) and VIIRS (USA).
In general, ocean colour remote sensing is one of the passive remote techniques. The sensor, mounted on a satellite, an aircraft or other remote platform, detects the radiometric flux at several selected wavelengths in the visible and near-infrared domains. The signal received by the sensor is determined by different processes in the water, as well as in the atmosphere (as shown in Figure 1.1). 1. scattering of sunlight by the atmosphere, 2. reflection of direct sunlight at the sea surface, 3. reflection of sunlight at sea surface, and 4. light reflected within the water body. Only the portion of the signal originating from the water body contains information on the water constituents; the remaining portion of the signal, which takes up more than 80 % of the total signal, has to be assessed precisely to extract the contribution from the water body [Morel, 1980].

There are two strategies to derive oceanic constituents from the signal of ocean colour sensor at top of atmosphere (TOA), a one-step method and a two-step method. For the traditionally used two-step method, the water leaving radiance (or reflectance) is firstly derived from the signal at TOA (this procedure is called ‘atmospheric correction’), and then oceanic constituents are

Ocean colour Sensor

Figure 1. Sketch of different origins of light received by space-borne sensor
retrieved from water leaving radiance (or reflectance). For the one-step method, oceanic constituents are directly derived from the signal at TOA. Since the one-step method has good performance for dealing with strongly absorbing aerosols, it has been paid more attentions recently [Gordon et al., 1997; Chomko and Gordon, 1998; Li et al., 2002]. The one-step method assumes that radiative transfer in the ocean and atmosphere is coupled. The oceanic constituents and aerosol properties are simultaneously derived from satellite measurements at TOA by using the entire spectrum available to ocean colour instruments. In this thesis, the focus is on two issues: a) to retrieve oceanic constituents from ocean colour measurements at sea level, b) to retrieve oceanic constituents from ocean colour measurements at top of atmosphere with the one-step method.

1.2. Retrieval of Oceanic Constituents from Ocean Colour Measurements at Sea Level

There are three major issues in the retrieval of oceanic constituents from ocean colour:

- How to quantify the relationship between optically significant oceanic constituents and inherent optical properties (IOPs)?
- How do IOPs determine ocean colour?
- How to obtain oceanic constituents from ocean colour measurements?

The first two issues are the so called ‘forward problem’, and the last issue is the so called ‘inverse problem’.

1.2.1. The Forward Problem

The forward problem is solved by radiative transfer theory. Radiative transfer theory describes the relationship between the IOPs of the oceanic constituents and the ocean colour. Based on radiative transfer theory, two different approaches relating the ocean colour to IOPs have been developed: one analytical and one numerical. The mostly used analytical expression relates the hemispherical reflectance $R$ just below the sea surface to the absorption coefficient $a$ and back scattering coefficient $b_b$ and was introduced by Gordon and Brown (1973):

$$R = f \frac{b_b}{a + b_b}$$

The proportionality factor $f$ varies between approx 0.3 to 0.5, depending on the ambient light field and the optical properties of water [Morel and Gentili, 1993].

Another analytical expression relating the remote sensing reflectance to the IOPs of oceanic constituents was derived by Lee et al. (1994):
where \( t \) is the transmittance of the air-sea interface, \( Q \) is the upwelling irradiance-to-radiance ratio, which is a function of the solar zenith angle and optical properties of water [Morel and Gentili, 1993], and \( n \) is the real part of the refractive index of seawater.

The numerical approach is based on simulations of radiative transfer. It allows to include all factors determining the ocean colour, i.e. IOPs, rough sea surface, observation geometry, inelastic scattering processes, etc. and has a potential for the development of more advanced retrieval methods. Another advantage is to avoid errors due to eventually poor approximation of the factor \( Q \) and the parameter \( f \).

A prerequisite for the numerical approach is the availability of bio-optical models relating IOPs and the actual constituents concentrations. These models are statistical expressions which are built up from concomitant in-situ measurements of IOPs and the corresponding constituents. Although the development of IOP models has made progress in recent years, there are still some of them whose accuracy is not sufficient for the development of oceanic constituents retrieval methods. The state-of-art in bio-optical modelling is as follows:

(a) some bio-optical models have been obtained from large global data sets, e.g., the absorption coefficients of phytoplankton [Bricaud et al., 1995; Bricaud et al., 1998] and CDOM [Bricaud et al., 1981] in the open ocean, as well as the scattering coefficient of phytoplankton and associated particles [Gordon and Morel, 1983; Loisel and Morel, 1998] also in the open ocean.

(b) Some bio-optical models have been developed for specific seas, e.g., the absorption coefficients of particles and CDOM, as well as the particles scattering coefficient in European coastal waters [Babin, 2000].

(c) Models for the phase function or the back scattering probability of marine particle in Case I waters are available [Zhang et al., 2003]. For Case II waters, such generic models are not available. It is one aim of this thesis to contribute to the development of a phase function model for Case II waters.

The uncertainties of bio-optical models are one of the major causes for errors of the retrieval of oceanic constituents.

1.2.2. The Inverse Problem

The determination of the oceanic constituents from ocean colour is a parameter estimation problem, where a set of parameters \( \mathbf{C} = \{ c_i, i = 1, ..., I \} \) are estimated from a set of measurements \( \mathbf{R} = \{ r_j, j = 1, ..., J \} \). The functional relationship between measurements and parameters can be expressed as:

\[
R_{rs} = \frac{ft^2}{Qn^2} \frac{b_h}{a + b_h},
\]

(1.2)
\[ R = g(C) . \quad (1.3) \]

Inverting Equation (1.3), one obtains the set of parameters \( C \) from the set of measurements \( R \):

\[ C = g^{-1}(R) . \quad (1.4) \]

In the frame of this thesis, \( C \) represents three different oceanic constituents: pigment, suspended particulate matter and coloured dissolved organic matter, while \( R \) is either the remote sensing reflectance, defined as the ratio of water leaving radiance to downwelling irradiance or the hemispherical reflectance, defined as the ratio of upwelling to downwelling irradiance, at sea level in \( J \) spectral channels.

If \( g \) would be a linear function, one could derive the inverse function \( g^{-1} \), and such obtain the oceanic constituents from the measured spectral reflectance. However, the functional relationship between the oceanic constituents and the resulting reflectance is complex and non-linear. It is therefore mostly impossible to achieve an analytic inversion of \( g \). The traditional way to overcome this problem is to make assumptions on the functional form of \( g^{-1} \) and then to solve Equation (1.4) by regression techniques or other statistical methods. However, it is often difficult to find the most appropriate functional form for \( g^{-1} \), which has direct implications on the accuracy of the retrieved constituent concentrations. In order to derive the oceanic constituents, a number of methods have been developed.

Based on statistical regression techniques, several representations of \( g^{-1} \) have been developed for Case I waters [see compilation in O'Reilly et al., 1998]. These empirical algorithms relate the water leaving reflectance at two or more wavelengths to the pigment concentration. They are still taken for the most successful operational methods to derive the oceanic constituents in Case I waters [O'Reilly et al., ibid]. But they are not valid for Case II waters.

A number of algorithms have been derived from Equation (1.1) and (1.2), using different inverse techniques. Carder et al. [1999] used an algebraic method to derive chlorophyll concentration. Bukata et al. [1981], Roesler and Perry [1995], Garver and Siegel [1997] and Lee et al., [1999] applied a non-linear optimisation method as the inversion technique for oceanic constituents retrieval.

In recent years, Artificial Neural Networks (ANN) have been increasingly applied to remote sensing data from ocean observing instruments, among those scatterometers and ocean colour sensors [see, e.g. Thiria et al., 1993; Keiner and Brown, 1999; Schiller and Doerffer, 1999; Gross et al., 2000]. ANN techniques are well suited for solving non-linear problems [Thiria et al., ibid]. No assumptions on the functions \( g \) or \( g^{-1} \) defined in Equations (1.3) or (1.4) are required. A number of studies have shown that ANN techniques have a good potential to derive the water constituents both in Case I and Case II waters [Buckton et al., 1999; Schiller and Doerffer, ibid; Gross et al., ibid]. Compared to the empirical and semi-analytical methods actually employed, they are less sensitive to noise. Furthermore, although the training of the ANN requires
considerable computational effort, its application is very fast. Therefore, ANN techniques are a promising method to derive oceanic constituents from ocean colour data.

Data of different origin may be used for the training of ANNs: "real" data which are obtained from in-situ measurements, and synthetic data which are obtained from numerical simulations. Obviously, it would be desirable to train ANNs entirely with real measurements. However, the number of data sets combining constituent concentrations and concomitant measurements of the oceanic light field is still rather limited. One of the most complete databases for Case I waters, SeaBAM (SeaWiFS Bio-optical Algorithm Mini-Workshop), contains just 900 data sets located in Case I waters. This is on one hand caused by rather strict measurements protocols that shall be applied to obtain high quality in-situ data [Mobley, 1999, Fargion et al., 2000], it is on the other hand a consequence of the difficult measurement conditions encountered in the marine environment. Besides, the pigment concentration in the SeaBAM data is inhomogeneously distributed: in more than 65.0 % of the cases, the pigment concentration is contained between 0.07 and 0.7 mg m$^{-3}$, while only 5.7% of the measurements have been taken at pigment concentrations above 5.0 mg m$^{-3}$. Such inhomogeneously distributed data sets are not well suited for the training of ANNs since they may lead to undesirable “overfitting” effects, meaning that a trained ANN gives good results where training data have been dense and bad results where training data have been sparse. There are more different constituents in Case II waters than in Case I waters, and the dynamic ranges of constituent concentrations are wider [IOCCG, 2000]. This makes it much more difficult to build up a data set of in-situ measurements that fulfils the requirements for ANN training. Radiative transfer (RT) simulations offer the opportunity to provide training data with a denser and more homogeneous distribution of the relevant parameters. A prerequisite for this is that the IOPs of the water constituents required as input to the RT simulations are well representing the conditions to which the derived ANN is later on applied.

1.3. Retrieval of Oceanic Constituents from Ocean Colour Measurements Taken at Top of Atmosphere

Traditionally, the retrieval of oceanic constituents is performed by a two-step process: atmospheric correction followed by a bio-optical algorithm to obtain the desired parameters.

The atmospheric correction algorithms which are commonly used are based on ‘the black pixel assumption’ [Gordon and Clark, 1981; Gordon, 1997; Siegel et al., 2000]. These algorithms were primarily designed for clear deep ocean areas. The information about atmospheric aerosols is derived from channels in the red and near-infrared (above 670 nm), where the water leaving radiance is close to zero. The derived aerosol information is extrapolated towards the visible channels and the atmospheric contribution is calculated and removed for full
For the turbid coastal environment, the ocean can no longer assumed to be black in the red and near-infrared because of strong back scattering by suspended materials. Under these conditions, ‘the black pixel assumption’ is no longer valid for deriving information on atmospheric aerosols. As a result, the algorithms developed for applications to clear ocean waters cannot be easily modified to retrieve water leaving radiance from remote sensing data acquired over the coastal environments.

Besides, even in the open ocean, the commonly used algorithms for atmospheric correction fail in the presence of strongly-absorbing aerosols [Gordon et al., 1997; Chomko and Gordon, 1998; Li et al., 2002]. If the aerosol is strongly absorbing, due to soot or dust component, the visible reflectance can not be derived from the NIR reflectance [Gordon et al., 1997]. The size distribution of strongly absorbing aerosols can be similar to that of the weakly absorbing aerosols typically present over ocean. Since the spectral variation of aerosol scattering depends mostly on the aerosol size distribution and only weakly on the refractive index, the spectral variation of scattering in the NIR is not sufficient to distinguish between weakly and strongly absorbing aerosols. Furthermore, the strongly absorbing aerosols (soot or dust) are coloured, i.e., their absorption is a function of wavelength [Nakajima et al., 1989]. Even if it was possible to estimate the absorption characteristics of these strongly absorbing aerosols in the NIR, the absorption in the visible could not be obtained by extent. Gordon et al. [1997] proposed a one-step algorithm to simultaneously determine aerosol properties and pigment concentration in Case I waters, which uses all the spectral bands of the sensor and can be applied to deal with weakly or strongly absorbing aerosols. Chomko and Gordon [2001] and Chomko et al. [2003] further extended this method. The look-up table or optimisation procedure was employed to retrieve the desired parameters in these studies. The limit of these processing methods is slowly computing, and not well suited to be used as an operation algorithm.

1.4. Objectives and Outline

As outlined in the previous sections, there is a need for a scheme to retrieve the oceanic constituents from reflectance at sea level or TOA.

The objective of this thesis is to contribute to the development of fast, accurate and robust algorithms for retrieval of oceanic constituents in Case I and Case II waters. To approach this objective, the following work has been done:

(1). Development of an ANN based on Radiative Transfer Calculations (RTC) for retrieval of the pigment concentration from remote sensing reflectance just above the sea surface in Case I waters;

(2). Modelling of the back scattering probability for marine particles in Case II water using the COASTLOOC in-situ measurements;
(3). Development of an ANN based on RTC for retrieval of the oceanic constituent concentration (CHL, SPM and CDOM) from hemispherical reflectance just below the sea surface in Case II waters;

(4). Development of an ANN based on RTC for retrieval of the oceanic constituent concentration (CHL, SPM and CDOM) from MERIS top of atmosphere measurements.

This thesis is structured as follows: after the introduction, the theoretical background is introduced in Chapter 2. Chapter 3 describes how the retrieval algorithms for pigment concentration in Case I waters were developed. Chapter 4 describes how the back scattering probability for marine particles in Case II waters was modelled. In Chapter 5, an ANN-based scheme for retrieval of oceanic constituents in Case II is derived using the hemispherical reflectance just below the sea surface as input. In Chapter 6, a scheme is proposed to retrieve the oceanic constituents in Case II waters from MERIS measurements data at top of atmosphere. In the last chapter, the summary of this investigation is given.
Chapter 2
THEORETICAL BACKGROUND

2.1. Concepts and Definitions

2.1.1. Radiometric Concepts

Two basic radiometric parameters are commonly used to describe the distribution of light field: radiance and irradiance. The definitions of the radiometric parameters used in this thesis are as follows [Kirk, 1984; Mobley, 1994]:

Radiance \( L(\theta, \phi, \lambda) \): The radiant power in a beam per unit solid angle per unit area perpendicular to the beam per unit wavelength interval.

Irradiance \( E(\lambda) \): The radiant power per unit area per unit wavelength interval.

Upward irradiance \( E_u(\lambda) \): The upward directed radiant power per unit area onto a downward facing horizontal surface.

Downward irradiance \( E_d(\lambda) \): The downward directed radiant power per unit area onto an upward facing horizontal surface.

\( E_u(\lambda) \) and \( E_d(\lambda) \) can be derived by the integration of \( L(\theta, \phi, \lambda) \) over appropriate angles:

\[
E_u(\lambda) = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi/2} L(\theta, \phi, \lambda) \cos \theta \sin \theta \, d\theta \, d\phi, \quad (2.1)
\]

\[
E_d(\lambda) = \int_{\phi=0}^{2\pi} \int_{\theta=\pi/2}^{\pi} L(\theta, \phi, \lambda) \cos \theta \sin \theta \, d\theta \, d\phi, \quad (2.2)
\]

where \( \theta \) is the zenith angle, and \( \phi \) is the azimuth angle.

2.1.2. Inherent Optical Properties

There are two types of optical properties of water: the inherent optical properties (IOPs) and the apparent optical properties (AOPs). IOPs solely depend on the type and the concentration of substance present in water, are independent of the ambient light field within the water. AOPs depend both on the type and the concentration of substance present in water and on the geometric structure of the ambient light field. AOPs can be derived from the measurements taken by an ocean colour sensor. The radiative transfer theory provides the connection between the AOPs and the IOPs. In this subsection, IOPs are introduced, and AOPs will be described in the next subsection.

The fundamental IOPs are the absorption coefficient and the volume scattering function. Other IOPs, such as the scattering coefficient, the beam attenuation coefficient and the single-scattering albedo, can be derived from these two fundamental IOPs.
The spectral absorption coefficient \( a(\lambda) \) is defined as the spectral absorptance per unit distance of photon travel in a dielectric medium.

The spectral scattering coefficient \( b(\lambda) \) is defined as the spectral scatterance per unit distance of photon travel in a dielectric medium.

The spectral beam attenuation coefficient \( c(\lambda) \) is defined as the sum of \( a(\lambda) \) and \( b(\lambda) \).

The spectral volume scattering function \( \beta(\theta, \lambda) \) is defined as the ratio of the scattered spectral radiance to the incident spectral irradiance per unit volume.

Integrating \( \beta(\theta, \lambda) \) over all directions gives the spectral scattering coefficient. It can be expressed as:

\[
b(\lambda) = 2\pi \int_{0}^{\frac{\pi}{2}} \beta(\theta, \lambda) \sin(\theta) d\theta \quad (2.3)
\]

The spectral scattering coefficient can be divided into two parts: the forward spectral scattering coefficient, \( b_f \), relating to light scattered from the beam in a forward direction, and the spectral back scattering coefficient, \( b_b \), relating to light scattered from the beam in a backward direction. They are expressed as,

\[
b_f(\lambda) = 2\pi \int_{0}^{\frac{\pi}{2}} \beta(\theta, \lambda) \sin(\theta) d\theta \quad (2.4)
\]

\[
b_b(\lambda) = 2\pi \int_{\frac{\pi}{2}}^{\pi} \beta(\theta, \lambda) \sin(\theta) d\theta \quad (2.5)
\]

The spectral phase function, \( \tilde{\beta}(\theta, \lambda) \) is defined as the ratio of the spectral volume scattering function to the spectral scattering coefficient. Hence, integral of \( \tilde{\beta}(\theta, \lambda) \) over all solid angles is equal to 1.

Forward scattering probability \( \tilde{b}_f(\lambda) \) and back scattering probability \( \tilde{b}_b(\lambda) \) are defined as,

\[
\tilde{b}_f(\lambda) \equiv \frac{b_f(\lambda)}{b(\lambda)} , \quad (2.6)
\]

\[
\tilde{b}_b(\lambda) \equiv \frac{b_b(\lambda)}{b(\lambda)} . \quad (2.7)
\]

Another inherent optical property commonly used is the spectral single scattering albedo, \( \omega_0(\lambda) \), defined as the ratio of the spectral scattering coefficient to the spectral beam attenuation coefficient \( c(\lambda) \):

\[
\omega_0(\lambda) \equiv \frac{b(\lambda)}{c(\lambda)} , \quad (2.8)
\]

where \( c(\lambda) = a(\lambda) + b(\lambda) \).

The single scattering albedo is also known as the probability of photon survival.
2.1.3. Apparent Optical Properties

Two apparent optical properties are commonly used to represent the intrinsic colour of the ocean: hemispherical reflectance $R(\lambda)$ and remote sensing reflectance $R_{\text{RS}}(\lambda)$.

**The spectral hemispherical reflectance** is defined as the ratio of spectral upward to downward irradiance:

$$R(\lambda, z) \equiv \frac{E_u(\lambda, z)}{E_d(\lambda, z)}$$

(2.9)

$R(\lambda, z)$ is often evaluated in the water just below the surface ($z=0^-$).

**The spectral remote sensing reflectance** $R_{\text{RS}}$ is defined as the ratio of the water leaving radiance to the downward irradiance just above the sea surface ($z=0^+$):

$$R_{\text{RS}}(\theta, \phi, \lambda) \equiv \frac{L_w(\theta, \phi, \lambda)}{E_d(\lambda)}$$

(2.10)

The spectral remote sensing reflectance specifies that portion of downward light incident onto the water surface which is returned through the surface into direction $(\theta, \phi)$.

**The Q factor** is defined as the ratio of upwelling irradiance to upwelling radiance just below the sea surface:

$$Q = \frac{E_u}{L_u}$$

(2.11)

If the $L_u$ distribution were isotropic that is independent on the direction of propagation, $Q$ would equal $\pi$. Actually, $Q$ is a function of solar zenith angle, observation angle, as well as the wavelength [Morel and Gentili, 1993, 1996].

2.2. Optical Properties of Sea Water

In order to study the absorbing and scattering properties of natural waters, it is necessary to know the composition of those waters.

Constituents of natural waters are traditionally divided into particulate matter and dissolved matter. When filtering a water sample, everything that passes through a filter whose pore size is on the order of 0.4 $\mu$m is called dissolved matter, and everything retained on the filter is called particulate matter. The composition of particulate matter is very complex. It is divided into organic and inorganic particles, or living and non-living particles, or endogenous and exogenous particles for different purposes. In this study, suspended particulate matter is divided into endogenous and exogenous particles. The endogenous particles consist of phytoplankton and other associated particles, such as detritus, bacteria produced locally in a water body. The exogenous particles consist of sediments brought by rivers, re-suspended from the shallow bottom, and dust-settled from the atmosphere.
The dissolved organic matter is mostly produced during the decay of plant matter. The coloured fraction of it consists mostly of various humic and fulvic acids and is referred to as coloured dissolved organic matter (CDOM), or yellow matter, gilvin, and gelbstoff, because of the properties of spectral absorption. Based on the different sources, it is divided into endogenous and exogenous CDOM. The endogenous CDOM is associated with biological activity in ocean, while the exogenous CDOM comes from land drainage, and is mainly produced by the decay of terrestrial vegetation. The concentrations of exogenous CDOM are mostly higher in the coastal waters influenced by river discharge.

The absorption coefficient of water \( a(\lambda) \) can be separated into individual contributions by different water constituents:

\[
a(\lambda) = a_w(\lambda) + a_{p1}(\lambda) + a_{p2}(\lambda) + a_{y1}(\lambda) + a_{y2}(\lambda),
\]

where the subscripts w, p1, p2, y1 and y2, respectively, indicate pure sea water, endogenous particles, exogenous particles, endogenous CDOM, and exogenous CDOM.

The total scattering coefficient of natural water is modelled as the sum of the scattering coefficients of pure seawater and particles:

\[
b(\lambda) = b_w(\lambda) + b_p(\lambda)
\]

The phase function of natural water depends on the phase function of the various scattering substances and their relative contribution to the total scattering coefficient:

\[
\tilde{\beta}(\lambda) = \frac{(b_w(\lambda) \times \tilde{\beta}_w(\lambda) + b_p(\lambda) \times \tilde{\beta}_p(\lambda))}{(b_w(\lambda) + b_p(\lambda))}
\]

### 2.2.1. Pure Seawater

The absorption coefficient of pure seawater \( a_w(\lambda) \) has been measured with sufficient accuracy for most applications. The Pope and Fry (1997) absorption values in the wavelength range 380 nm to 709 nm are widely used. For wavelengths above 709 nm, the Hale and Querry (1973) absorption values are commonly taken. They are shown in Figure 2.1.

The scattering coefficient of pure seawater \( b_w(\lambda) \) (Figure 2.2) is taken from Morel (1974):

\[
b_w(\lambda) = 0.00288\left(\frac{\lambda}{500}\right)^{-4.32}
\]

The phase function \( \tilde{\beta}_w(\theta, \lambda) \) (Figure 2.3) was given by (Morel, 1974):

\[
\tilde{\beta}_w(\theta) = 0.06225(1 + 0.835 \cos^2 \theta)
\]

### 2.2.2. Phytoplankton

Phytoplankton cells are strong absorbers of visible light and therefore play a major role in determining the absorption of natural waters. The spectral absorption properties of phytoplankton have been widely studied [Prieur and Sathyendranath, 1981; Morel, 1988; Carder et al., 1991; Hoepffner and Sathyendranath, 1993; Bricaud et al., 1995; Bricaud et al., 1998; Kondratyev et
al., 1998]. These studies show that the phytoplankton absorption has several general features: there are distinct absorption peaks around 440 nm and 675 nm; there is relatively little absorption between 550 and 650 nm.

Figure 2.1. The absorption coefficient of pure seawater ($\lambda$=380-709 nm from Pope and Fry (1994), $\lambda > 709$ nm from Hale and Querry (1973)).

Figure 2.2. The scattering coefficient of pure seawater, from Morel (1974)
Bricaud et al. (1995) developed an empirical relationship between the chlorophyll-a specific absorption coefficient of phytoplankton $a_{ph}^*(\lambda)$ and the chlorophyll concentration from a data set including 815 spectra determined with the wet filter technique in different regions of the world ocean (covering the range of chlorophyll concentration 0.02-25 µg/l). It is expressed as:

$$a_{ph}^*(\lambda) = A(\lambda) < \text{chla} >^{-B(\lambda)}, \quad (2.17)$$

where $A(\lambda)$ and $B(\lambda)$ are positive, wavelength-dependent parameters.

Based on 1116 spectra of endogenous particles, Bricaud et al. (1998) derived a power function to represent the relationship between the absorption coefficient of endogenous particles (i.e. phytoplankton plus detritus particles) and the chlorophyll concentration:

$$a_p(\lambda) = A_p(\lambda) < \text{CHL} >^{-E_p(\lambda)}, \quad (2.18)$$

where $A_p(\lambda)$ and $E_p(\lambda)$ are positive, wavelength-dependent parameters.

Figure 2.2 shows the spectral absorption coefficient of endogenous particles for CHL =1 mg m$^{-3}$, based on Eq. (2.18).

Several models are available for the total scattering coefficient $b_p(\lambda)$ of endogenous particles (Gordon and Morel, 1983; Loisel and Morel, 1998).

Based on a large dataset (N = 2,787) made up of recent measurements of the beam attenuation coefficient at 660 nm and of the chlorophyll concentration in Case I waters, Loisel and Morel (1998) proposed a power law relating the pigment concentration to the beam attenuation coefficient:

$$c_p(660) = \alpha [\text{CHL}]^\beta \quad (2.19)$$
At 660 nm, about 97% of the particle beam attenuation coefficient is caused by scattering [Loisel and Morel, ibid]. At this wavelength, the scattering coefficient may therefore be assumed identical to the beam attenuation coefficient. Recent in-situ measurements show that the wavelength dependence of scattering by marine particles is low or is even not observed [Barnard et al., 1998; Gould et al., 1999; Babin et al., 2000]. Therefore, the scattering of marine particles is assumed to be independent of wavelength within the considered domain, expressed as:

\[
b_p = A_{bp} [CHL]^{B_{bp}},
\]

(2.20)

where the coefficients \(A_{bp}=0.252\) and \(B_{bp}=0.635\) were derived from 610 measurements in the homogeneous surface layer taken in the tropical and subtropical Atlantic and Pacific Oceans and the Mediterranean Sea [Loisel and Morel, ibid.]. In this study, Eq. (2.20) is taken as the model for the scattering coefficient of endogenous particles.

The model of Gordon and Morel [1983] is given by:

\[
b_p(\lambda) = 0.3 <chl >^{0.62} \left( \frac{550}{\lambda} \right)
\]

(2.21)

Figure 2.4. The absorption coefficient of endogenous particles for CHL=1 mg m\(^{-3}\), based on Eq. 2.18.

2.2.3. Suspended particulate matter

Based on the absorption properties of nonalgal particles (NAP) measured at about 350 stations in various coastal waters around Europe including the English Channel, Adriatic Sea, Baltic Sea, Mediterranean Sea, and North Sea, absorption spectra of these particles were well described by an exponential function [Babin, 2000]:
\[ a_{\text{nap}}(\lambda) = a_{\text{nap}}(443) \exp^{-s(\lambda-443)}, \quad (2.22) \]
\[ a_{\text{nap}}(443) = 0.0216 < SPM >^{1.0247}, \quad (2.23) \]

where \( s \) is empirically determined, an average value of 0.0123 nm\(^{-1}\) with SD = 0.0013 nm\(^{-1}\). SPM is the concentration of total suspended matter, in g m\(^{-3}\). Figure 2.5 shows the spectral absorption coefficient of exogenous particles for SPM = 3 g m\(^{-3}\), based on Eq. (2.22) and (2.23).

The scattering coefficient of total suspended particulate matter in Case II waters (endogenous plus exogenous particles) \( b_p(\lambda) \) can be expressed as [Babin, 2000]:
\[ b_p(\lambda) = b_p(550) \]
\[ b_p(550) = 0.5 < SPM > \]

The above relationships were derived based the measurements taken at 250 different coastal sites located in the English Channel, Adriatic Sea, Baltic Sea, Mediterranean Sea, and North Sea.

Figure 2.5. The absorption coefficient of exogenous particles for SPM= 3 g m\(^{-3}\), based on Eq. (2.22) and (2.23).

Although commercial instruments for measuring the angular distribution of scattered light in situ are meanwhile becoming available [Pegau et al., 2001], measurements of the volume scattering function or the back scattering probability for marine particles are still scare. The phase function for marine particles still mostly used in ocean optics has been derived by Mobley et al. (1993) based on measurements taken in San Diego Harbour and San Diego Bay [Petzold, 1972]. Figure 2.6 shows the volume scattering functions for three oceanic waters [Petzold, 1972] and particles phase function derived by Mobley et al. (1993), all at \( \lambda = 514 \) nm.
Figure 2.6. Volume scattering function from three natural waters measured by Petzold (1972) and particles phase function derived by Mobley et al. (1993), all at $\lambda=514$. The solid line represents clear water, the dotted line represents coastal water, the dash dot line represents turbid water, and the long dashes line represents particles phase function.

The particles phase function derived by Mobley et al. (1993) is characterised by a constant back-scattering probability of 1.8%. It has been recognised that it is not realistic either in Case I waters or in Case II waters. The phase functions of particles in marine environment should have large variability. The shape of the phase function depends on the refractive index and the size distribution of particles. There is a high diversity of the refractive index and the size distribution, depending on the geological origins of the inorganic particles or on the proportion of organic to inorganic particles, especially in coastal waters. It is recognised that uncertainties of the phase function is one of the largest sources of errors, when deriving retrieval methods from RT simulations.

2.2.4. CDOM (Yellow Substance)

CDOM absorbs very strong in the blue domain, but its absorption decrease rapidly with increasing wavelength. Absorption coefficient of CDOM is reasonably well modelled as (Bricaud et al., 1981):

$$a_y(\lambda) = a_y(440)e^{-S_y(\lambda-440)},$$  

where $S_y$ is empirically determined. Many researchers have reported that the average value of $S_y$ is around 0.014 nm$^{-1}$ (Bricaud et al., 1981, Kishnio et al., 1984). For the GOMEX and COLOR cruises, an average value of 0.017 nm$^{-1}$ was obtained (Carder et al., 1991). From the
COASTLOOC data, an average value of 0.0176 nm\(^{-1}\) was derived (Babin, 2000). Figure 2.7 shows the spectrum of CDOM with \(S_y=0.0176\) nm\(^{-1}\) and \(a_y(440)=1.0\) m\(^{-1}\) based on Eq. (2.26).

CDOM is assumed to be a pure absorber.

### 2.2.5. Classification of Waters

A classification of ocean waters was introduced by Morel and Prieur (1977), refined later by Gordon and Morel (1983). It is very useful to deal with the retrieval of ocean colour. By definition, ocean waters are divided into two types: Case I and Case II. Case I water are those waters for which phytoplankton and associated products (detritus, CDOM, bacteria) play a dominant role in the variation of optical properties of waters. Case II waters are those waters whose optical properties are influenced not only by phytoplankton and associated products, but also by other substance, such as exogenous particles and exogenous CDOM.

Based on the above definition, Case I water consists of the following optically significant constituents: Phytoplankton and associated particles (endogenous particles), and associated CDOM (endogenous CDOM). Beside the constituents included in Case I waters, Case II waters also consist of exogenous particles, discharged by rivers, or re-suspended from shallow bottom, and the exogenous CDOM from land drainage.

![Figure 2.7. Absorption coefficient of CDOM, for \(a_y(440)=1.0\) m\(^{-1}\) and \(s_y=0.017\) nm\(^{-1}\)](image)

#### 2.3. Radiative Transfer Simulations

Assuming a plane-parallel medium in the ocean-atmosphere system, with a constant input of monochromatic unpolarized radiation at top of atmosphere, the radiative transfer equation may be written as [Smith, 1974]:

\[
\text{Smith, 1974}
\]
\[
\frac{dL(z, \theta, \phi)}{dr} = -c(z)L(z, \theta, \phi) + L^*(z, \theta, \phi) + L^S(z, \theta, \phi) \tag{2.27}
\]

The term on the left is the rate of change of radiance with distance, \( r \), along the path specified by zenith and azimuthal angles \( \theta \) and \( \phi \), at depth \( z \). The net rate of change is the resultant of three processes: loss by attenuation along the direction of travel (\( c(z) \) being the value of the beam attenuation coefficient at depth \( z \)), gain by elastic scattering of the light initially travelling in other direction \((\theta', \phi')\) into the direction \((\theta, \phi)\), and gain by inelastic scattering into the direction \((\theta, \phi)\). The elastic scattering term is given by

\[
L^*(z, \theta, \phi) = \int_0^{2\pi} \beta(z, \theta, \phi; \theta', \phi') L(z, \theta', \phi') d\omega(\theta', \phi') \tag{2.28}
\]

where \( \beta(z, \theta, \phi) \) is the volume scattering function of the medium at depth \( z \), and \( \omega \) is the solid angle.

Because of the mathematical complexity of Eq. (2.27), it must be solved numerically for any realistic situations. MOMO is one of the methods for the above purpose [Fischer and Grassl, 1984; Fell and Fischer, 2001]. All radiative transfer simulations in this study are performed with the radiative transfer model MOMO. The input needed for MOMO are the IOPs of the atmospheric and oceanic constituents, namely beam attenuation coefficient, single scattering albedo and phase function. This code is based on the Matrix-Operator Method and allows the simulation of the ambient light field in the coupled atmosphere-ocean system. The code is described in detail elsewhere [Fell and Fischer, 2001] and just a short introduction into its characteristics will be presented here. The main advantage of the Matrix-Operator method is its efficiency with regard to the simulation of light propagation in optically thick media [Plass et al., 1973] and is therefore particularly suited for the development of remote sensing methods regarding the retrieval of water constituents. The code calculates the azimuthally resolved radiance at a discrete number of solar incident and observation zenith angle. The required vertical profiles of the atmospheric and water constituents are introduced through an appropriate number of homogeneous plane parallel layers. Absorption by atmospheric gases is modelled using the modified \( k \)-distribution approach from Bennartz and Fischer [2000]. Reflection at the rough sea surface is modelled according to the statistical description of the wave facet distribution derived by Cox and Munk [1954], transmission through the rough sea surface is treated using an approximation introduced by Fell and Fischer [2001]. Alternatively, the sea surface may be assumed as flat. Two inelastic scattering processes have been incorporated into the MOMO code: sun-stimulated chlorophyll fluorescence [after Gordon, 1976] and Raman scattering [based on Mobley, 1994]. Polarisation, CDOM fluorescence and wind direction dependent effects are not taken into account.
The MOMO code has been thoroughly tested and validated through a number of measures [Fell, 1997; Montagner, 1999, Fell and Fischer, 2001]: it has been compared to analytical solutions of the radiative transfer equation for specific simple problems, it has been successfully applied to the canonical problems of oceanic radiative transfer defined in Mobley et al. [1993] and has been compared to other codes dealing with the radiative transfer in the atmosphere-ocean system on occasion of a model comparison exercise executed in the frame of the MERIS algorithm development activities. The observed deviations between the MOMO output and the analytical solutions are mostly below 0.1%, the differences between the compared codes range between <0.1% up to about 5%, depending on the difficulty of the treated problem. These results confirm the suitability of MOMO for the execution of the presented study.

2.4. Artificial Neural Network

A neural network derives its computing power through its massively parallel distributed structure and its ability to learn and therefore to generalise. A neural network is composed of a number of neurons, which are arranged in different network layers and are connected by links. Each link has a numeric weight associated with it. Weights are the primary means of long-term storage in neural networks, and learning usually takes place by updating the weights (Rumelhart and McClelland, 1986).

In this thesis, a particular class of ANNs is used, the so-called Multi-Layer Perceptrons (MLPs), to derive the oceanic constituents from ocean colour at the sea level or at top of atmosphere. The MLP used within this study consists of three layers: input layer, one hidden layer and output layer. It has been proved that a neural network with one hidden layer can approximate any continuous function to any arbitrary accuracy given sufficient hidden nodes [Cybenko, 1989; Masters, 1993]. A bias parameter is added to both input and hidden layer (Figure 2.8). The neurons in the hidden layer are connected with every neuron in the input and output layers. The transfer of information through the MLP can be described by the following equation:

$$\tilde{O} = s_c(W^\text{HO} \times s_c(W^{\text{HI}} \times \tilde{I})),$$

where $\tilde{I}$ is the input to the MLP, consisting of appropriately pre-processed reflectance values as well as other auxiliary parameters. The weight matrix $W^{\text{HI}}$ contains the weights of all connections between input and hidden layer whereas the weight matrix $W^{\text{HO}}$ contains the weights of all connections between hidden and output layer. At each neuron of hidden and output layer, transition of information is done through the non-linear sigmoidal function, defined by:

$$s_c(x) = \frac{1}{1 + \exp(-c_f x)},$$

where $c_f$ is a constant.
where $c_i$ is called the temperature constant. The output of the sigmoid defined in Equation (2.30) is confined to the interval $[0, 1]$. Training of the ANN is done through the back-propagation method, which is a supervised learning technique that compares the responses of the output neurons with the known true responses (training data), traces the errors back through the ANN and readjusts the randomly initialised elements of the weight matrices in order to minimise the multidimensional error function:

$$E = f(W_{IH}, W_{HO})$$

(2.31)

In order to test if a trained ANN is able to generalise the functional relationship describing the investigated physical system and has not just found a specific approximation of the offered training data (so called "overfitting", which is comparable to fitting few data with a high-order polynomial), it is applied to another data set (test data) independent of the training data. Only if the absolute error of the ANN output for the test data is within the desired range, the training is deemed successful.

---

![Figure 2.8. Architecture of the Multi-Layer-Perceptron used in this thesis.](image)
Chapter 3

RETRIEVAL OF THE PIGMENT CONCENTRATION IN CASE I WATERS

3.1. Introduction

It has been recognised that Artificial Neural Network (ANN) techniques have a good potential to derive the water constituents both in Case I and Case II waters [Keiner and Brown, 1999; Schiller and Doeffer, 1999; Gross et al., 2000]. ANNs have several distinct advantages. First, ANNs may be used to approximate the non-linear relationship between observations and target parameter(s) without explicitly knowing their direct functional dependence. Second, ANNs have the ability to generalise even in the presence of noisy data. Third, once an ANN is created, the processing time for parameter retrieval is short as compared to the inverse modelling techniques.

The number of in-situ data sets combining pigment concentration and concomitant measurements of the oceanic light field is still rather limited. Besides, the distribution of pigment concentrations in the available data sets is rather inhomogenous and thus does not meet the requirement for successful training of an ANN. Therefore, the training data used in this study is obtained from simulations of radiative transfer in the atmosphere-ocean system. RT predicts the light field based on physical models describing the interaction of matter and light. If all processes relevant to the radiative transfer in the atmosphere-ocean system were known, the light field could be predicted for arbitrary combinations of water constituents concentrations, sea surface state, atmospheric composition and observation geometry. The key problem here is to exactly define the inherent optical properties (IOPs) of sea water and to relate these to the parameters of interest (such as the pigment concentration). The IOPs of pure sea water are meanwhile well known. Well established parameterisations also exist for the absorption coefficient of marine particles in Case I waters as function of the pigment concentration and the absorption of coloured dissolved organic matter (CDOM). Recently, a new model for the phase function of marine particles was derived by optimising the agreement between RT simulations of the remote sensing reflectance with the corresponding SeaBAM data [Zhang et al., 2003]. This availability allows to simulate the light field in Case I waters as required for this study.

In the present work, a method is proposed for the retrieval of the pigment concentration in Case I waters on a global scale, based on Artificial Neural Network (ANN) techniques. Input to the presented method is the spectral remote sensing reflectance just above the sea surface. The resistance against atmospheric correction errors which is required for methods working at sea
level is assured by adding the previously determined appropriate noise level to the data set used for the training of the ANNs.

3.2. Data Sets

Three different data sets are used in this study, relating remote sensing reflectance (SeaBAM) or hemispherical reflectance (COASTLOOC) to the pigment concentration:

1. a synthetic data set ("training data") derived from RT simulations to train the different ANNs,
2. the SeaBAM in-situ data ("validation data") to evaluate the performance of each individual ANN and such to identify the most appropriate one with respect to pigment retrieval,
3. the COASTLOOC in-situ data ("test data") to assess in how far the ANN-based pigment retrieval scheme is applicable to data other than those contained in the SeaBAM data set.

These three data sets are described in more detail in the following subsections.

3.2.1. Training Data

The synthetic data set used for the training of ANNs was created using the computer code MOMO [Fell and Fischer, 2001] which has been described in Section 2.3. In order to speed up the radiative transfer simulations and to make it convenient to compare simulations with in-situ measurements, a number of assumptions and simplifications were made:

- atmospheric Rayleigh scattering according to a constant surface air pressure of 1013.25 hPa,
- atmospheric scattering by maritime aerosol assuming a constant aerosol optical depth of 0.1 at 550 nm,
- no atmospheric absorption,
- flat air-sea interface,
- no vertical stratification of the ocean,
- no inelastic scattering inside the water body (i.e. no Raman scattering, no chlorophyll-a fluorescence, no CDOM fluorescence).

For any study based on RT simulations of the oceanic light field, detailed knowledge on the IOPs of the oceanic constituents is required. The IOP models used in this study are summarised in Table 3.1. Based on the above assumptions and the selected IOP models, simulations of remote sensing reflectance at nadir direction were made for:

- 18 solar zenith angles between 0° and 87°,
- 5 wavelengths: 412, 443, 490, 515, and 555 nm,
- 300 pigment concentrations between 0.025 and 35 mg m⁻³.
In the SeaBAM data set (see below), no information is given on the solar zenith angles at which measurements have been taken. Therefore, the simulated remote sensing reflectances were averaged over solar zenith angles between 0° and 70° to construct a synthetic data set where the relationship between \( R_{RS} \) and pigment concentration is independent on the viewing geometry. The wavelengths chosen for the simulations correspond to the central wavelengths of spectral channels commonly used for space-borne or in-situ ocean colour measurements. The method derived herein may therefore be applied to most ocean colour instruments. A logarithmic distribution of the pigment concentrations has been selected for the simulations, thus, each order of magnitude within the considered pigment concentration range is represented with a similar number of cases.

### Table 3.1. Parameterisations of inherent optical properties of pure sea water and water constituents used of Case I waters.

<table>
<thead>
<tr>
<th>Constituents</th>
<th>IOP</th>
<th>Parameterisation or measurement</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pure sea water</td>
<td>Absorption</td>
<td>Directly measured</td>
<td>Pope and Fry [1997]</td>
</tr>
<tr>
<td></td>
<td>Total scattering</td>
<td>( b_w(\lambda) = 0.00288(\frac{\lambda}{500})^{-4.32} )</td>
<td>Morel [1974]</td>
</tr>
<tr>
<td></td>
<td>Phase function</td>
<td>( p_w(\lambda) = 0.06225(1 + 0.835 \cos^2 \theta) )</td>
<td>Morel [1974]</td>
</tr>
<tr>
<td>Marine particulate matter</td>
<td>Absorption</td>
<td>( a_p(\lambda) = A_{wp}(\lambda) [Chl]^b_{wp}(\lambda) )</td>
<td>Bricaud et al. [1998]</td>
</tr>
<tr>
<td></td>
<td>Total scattering</td>
<td>( b_p(\lambda) = A_{bp}[Chl]^{b_{bp}}, )</td>
<td>Loisel and Morel [1998]</td>
</tr>
<tr>
<td></td>
<td>Phase function</td>
<td>( p_p = f(Chl, \lambda) )</td>
<td>Zhang et al. [2003]</td>
</tr>
<tr>
<td>CDOM</td>
<td>Absorption</td>
<td>( a_y(\lambda) = a_y(443) e^{-S_y(\lambda-443)} ), ( S_y = 0.014 )</td>
<td>Bricaud et al. [1981]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( a_y(443) = 0.00348 + 0.5 \ast a_p(443) )</td>
<td>Zhang et al. [2003]</td>
</tr>
</tbody>
</table>

### 3.2.2. Validation Data: the SeaBAM Data Set

The performance of each individually trained ANN is assessed using the SeaBAM data set. The SeaBAM data set is a compilation of measurements and derived parameters taken at 919 stations on occasion of various oceanographic campaigns. 900 stations are considered as being located in Case I waters. The SeaBAM data set covers a pigment range between approx. 0.03 and
30.0 mg m$^{-3}$, and provides remote sensing reflectance values into nadir direction at six wavelengths (412, 443, 490, 515, 555, 670 nm). It is considered to contain the most comprehensive information on remote sensing reflectance values and concomitant pigment concentrations in Case I waters on a global scale. Detailed information on the SeaBAM data set is shown in Table 3.2. It is publicly available from the SeaBASS (SeaWiFS Bio-optical Archive and Storage System) web site (http://seabass.gsfc.nasa.gov).

Table 3.2. Data sources and characteristics of the SeaBAM data set, reproduced after O’Reilly et al. [1998].

<table>
<thead>
<tr>
<th>Campaign name</th>
<th>Provider / PI</th>
<th>Area</th>
<th>Time</th>
<th>Data sets</th>
<th>Water type</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBOP92-93</td>
<td>D. Siegel</td>
<td>Sargasso Sea</td>
<td>monthly, 1992-1993</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>BBOP94-95</td>
<td>D. Siegel</td>
<td>Sargasso Sea</td>
<td>monthly, 1994-1995</td>
<td>67</td>
<td>1</td>
</tr>
<tr>
<td>WOCE</td>
<td>J. Marra</td>
<td>50°S-13°N, 88°-91°W, 10°S-30°N, 18°-37°W</td>
<td>March 1993, April 1993</td>
<td>70</td>
<td>1</td>
</tr>
<tr>
<td>EQPAC</td>
<td>C. Davis</td>
<td>0.140°W</td>
<td>March, Sept. 1992</td>
<td>126</td>
<td>1</td>
</tr>
<tr>
<td>NABE</td>
<td>C. Davis</td>
<td>46°N, 19°W</td>
<td>April 1989</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>NABE</td>
<td>C. Trees</td>
<td>46°-59°N, 17°-20°W</td>
<td>May 1989</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>CALCOFI</td>
<td>G. Mitchell</td>
<td>California Current</td>
<td>Aug. 1993-July 1996</td>
<td>303</td>
<td>1</td>
</tr>
<tr>
<td>MOCE 1</td>
<td>D. Clark</td>
<td>Monterey Bay</td>
<td>Sept. 1992</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>MOCE 2</td>
<td>D. Clark</td>
<td>Gulf of California</td>
<td>April 1993</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>North Sea</td>
<td>R. Doerffer</td>
<td>55°-52°N, 0°-8°E</td>
<td>July 1994</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Chesapeake Bay</td>
<td>L. Harding</td>
<td>~37°N, 75°W</td>
<td>April, July 1995</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Canadian Arctic</td>
<td>G. Cota</td>
<td>~74.38°N, 95°W</td>
<td>Aug. 1996</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>AMT</td>
<td>G. Moore</td>
<td>50°N-50°S</td>
<td>Sept. 1995, April 1996</td>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>919</td>
</tr>
</tbody>
</table>
3.2.3. Test Data: the COASTLOOC Data Set

The COASTLOOC (Coastal Surveillance Through Observation of Ocean Colour) data set relates the subsurface hemispherical reflectance at 13 wavelengths between 412 and 865 nm to concomitantly measured IOPs and water constituents concentrations [Babin, 2000]. Most of the 424 stations visited during the COASTLOOC campaigns have been gathered in Case II waters in different European coastal areas, except for 93 stations located in Case I waters in the Atlantic Ocean and Mediterranean Sea. The measurements at 67 of the 93 Case I water stations fulfilled the requirements in terms of completeness and accuracy, and were retained for further processing in the frame of algorithm development for Case I waters. Table 3.3 provides more information on this data set.

<table>
<thead>
<tr>
<th>Campaign</th>
<th>PI</th>
<th>Area</th>
<th>Time</th>
<th>Case I stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>COASTLOOC 1</td>
<td>M. Babin</td>
<td>NE Atlantic Ocean</td>
<td>April 1997</td>
<td>34</td>
</tr>
<tr>
<td>COASTLOOC 3</td>
<td>M. Babin</td>
<td>Northern Adriatic Sea</td>
<td>July, August 1997</td>
<td>6</td>
</tr>
<tr>
<td>COASTLOOC 4</td>
<td>M. Babin</td>
<td>Mediterranean Sea</td>
<td>October 1997</td>
<td>12</td>
</tr>
<tr>
<td>ALMOFRONT 2</td>
<td>H. Claustre</td>
<td>Mediterranean Sea (Golfe du Lion)</td>
<td>January 1998</td>
<td>41</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>93</td>
</tr>
</tbody>
</table>

3.3. ANN-based Pigment Retrieval in Case I Waters

In the present study, a three-layer MLPs ANN is used to derive the functional relationship between the spectral remote sensing reflectance into the nadir direction and the pigment concentration. The architecture of the ANN has been described in Section 2.4.

3.3.1. Parameters affecting the performance of an ANN

The performance of a trained MLP depends mainly on three limiting factors: information content of the offered input, number of the neurons in the hidden layer, noise level added to the training data set. This is elucidated in the following subsections.
3.3.1.1. Information content of the offered input

To determine which and how many spectral bands or band ratios are best suited for pigment retrieval, nine combinations of input data (listed in Table 3.4) were tested. Of the nine listed cases, the first three are combinations of remote sensing reflectances at different wavelengths. The other six cases are combinations of remote sensing reflectance ratios. The dimensionality of the input data determines the number of neurons in the input layer.

3.3.1.2. Number of the neurons in the hidden layer

How many hidden neurons should be used? Generally speaking, if too few hidden neurons are used, high training error and high generalisation error may be obtained due to underfitting and high statistical bias. If too many hidden neurons are used, low training error may result, but there may be high generalisation errors due to overfitting and high variance [Geman et al., 1992]. Besides, the training process becomes very slow. In this study, the performance of ANNs with 4, 9, 20, and 30 hidden neurons has been tested.

3.3.1.3. Noise level added to the training data set

Through adding noise to synthetic training data, the robustness of the trained ANN with respect to noisy input data is increased. However, adding too much noise will obviously lead to a high generalisation error [Koistinen and Holmström, 1992]. To determine the appropriate noise level, 5 %, 10 %, 20 %, and 30 % noise was added to the synthetic training data used as input to the ANNs (RRS or RRS ratios). Since no specific noise model was available, a simple approach was chosen: the noise in different samples and in different elements of a sample is assumed to be independent, its distribution is assumed to be random and to have zero mean.

3.3.2. ANN Training

A synthetic data set for ANN training has been generated from the RT simulations outlined in Section 3.2. The synthetic data set is composed of 300 remote sensing reflectance spectra, corresponding to 300 pigment concentration values. Using a higher number of samples to train the ANNs did not significantly improve their performance. The main reason for this is the relative simplicity of the IOP models which are all parameterised as function of the pigment concentration alone in combination with the constant observation geometry.

For each of the nine input combinations listed in Table 3.4, training data sets of four different noise levels (5%, 10%, 20%, 30%) were used to train ANNs with 4, 9, 20, or 30 neurons in the hidden layer. This gives a total of $9 \times 4 \times 4 = 144$ combinations, each of which is represented by an individually trained ANN. One specific training set has been derived for each combination of input parameters and noise level, giving a total of $9 \times 4 = 36$ training data sets. Each training data set is then used to train the 4 ANNs characterised by their different number of hidden neurons.
Table 3.4. Performance with regard to pigment retrieval of spectral band combinations used as input to ANNs.

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Input Neurons in hidden layer</th>
<th>Noise level (%)</th>
<th>Training data (synthetic)</th>
<th>Validation data (SeaBAM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RMSE</td>
<td>r^2</td>
</tr>
<tr>
<td>1</td>
<td>R\textsubscript{RS}412, R\textsubscript{RS}443, R\textsubscript{RS}490, R\textsubscript{RS}515, R\textsubscript{RS}555</td>
<td>4 30</td>
<td>0.156</td>
<td>0.971</td>
</tr>
<tr>
<td>2</td>
<td>R\textsubscript{RS}443, R\textsubscript{RS}490, R\textsubscript{RS}515</td>
<td>4 30</td>
<td>0.199</td>
<td>0.953</td>
</tr>
<tr>
<td>3</td>
<td>R\textsubscript{RS}443, R\textsubscript{RS}490</td>
<td>4 30</td>
<td>0.210</td>
<td>0.948</td>
</tr>
<tr>
<td>4</td>
<td>R\textsubscript{RS}412/R\textsubscript{RS}555, R\textsubscript{RS}443/R\textsubscript{RS}555, R\textsubscript{RS}490/R\textsubscript{RS}555, R\textsubscript{RS}515/R\textsubscript{RS}555</td>
<td>20 5</td>
<td>0.0983</td>
<td>0.988</td>
</tr>
<tr>
<td>5</td>
<td>R\textsubscript{RS}443/R\textsubscript{RS}555, R\textsubscript{RS}490/R\textsubscript{RS}555, R\textsubscript{RS}515/R\textsubscript{RS}555</td>
<td>30 5</td>
<td>0.0787</td>
<td>0.993</td>
</tr>
<tr>
<td>6</td>
<td>R\textsubscript{RS}443/R\textsubscript{RS}555, R\textsubscript{RS}515/R\textsubscript{RS}555</td>
<td>30 5</td>
<td>0.0781</td>
<td>0.993</td>
</tr>
<tr>
<td>7</td>
<td>R\textsubscript{RS}443/R\textsubscript{RS}555, R\textsubscript{RS}490/R\textsubscript{RS}555</td>
<td>4 20</td>
<td>0.127</td>
<td>0.981</td>
</tr>
<tr>
<td>8</td>
<td>R\textsubscript{RS}443/R\textsubscript{RS}555</td>
<td>9 10</td>
<td>0.0924</td>
<td>0.990</td>
</tr>
<tr>
<td>9</td>
<td>R\textsubscript{RS}490/R\textsubscript{RS}555</td>
<td>9 5</td>
<td>0.0886</td>
<td>0.982</td>
</tr>
</tbody>
</table>

3.3.3. Determining ANN Architecture and Noise Adding for Optimal Pigment Retrieval

In order to find the ANN best suited for pigment retrieval, the ANN pigment forecasts were compared to the SeaBAM data by two error measures. First, root mean square error (RMSE) defined by:

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} [\log_{10}(CHL_{i}^D) - \log_{10}(CHL_{i}^M)]^2},
\]  

where \(CHL\) represents the pigment concentration, and the superscripts \(D\) and \(M\) indicate derived and measured values. Second, the square of the Pearson’s correlation coefficient \(r^2\). The pigment concentrations are log-transformed prior to calculating RMSE and correlation coefficients. The optimum number of hidden neurons and the appropriate noise level with regard to the above two error measures are given in Table 3.4 for each of the nine input combinations.

Using the remote sensing reflectance at five wavelengths between 412 and 555 nm as input to the ANN (case No. 1) results in \(r^2 = 0.835\) between the ANN pigment forecast and the corresponding SeaBAM values (RMSE = 0.271). Using one or more remote sensing reflectance
ratios as input to the ANN (cases No. 4-9), the $r^2$ values are in any case above 0.919, and the RMSE values below 0.167. The performance of the ANNs using remote sensing reflectance ratios as input is thus considerably higher than that of ANNs using absolute remote sensing reflectance values. The explanation for this effect is that spectrally correlated noise is partly cancelled out through division of the remote sensing reflectance at two wavelengths.

Of the six combinations of remote sensing reflectance ratios taken as input to the ANN, the best results (lowest RMSE and highest $r^2$) were obtained using the three remote sensing reflectance ratios $R_{RS443}/R_{RS555}$, $R_{RS490}/R_{RS555}$, and $R_{RS515}/R_{RS555}$ as input (case No. 5). Interestingly, this case is not the one with most spectral ratios as input (case No. 4). Even the performance of case No. 8 with only one spectral ratio ($R_{RS443}/R_{RS555}$) as input is better than that of case No. 4 using four spectral ratios, and close to that of case No. 6 using three spectral ratios. Generally speaking, the higher the number of the spectral bands, the more information is available, and the higher should be the accuracy of the ANN forecast. On the other hand, inclusion of noisy data will reduce the ANN forecast performance.

### 3.3.4. Best ANN for Pigment Retrieval

From the above results, the following optimum ANN architecture has been selected: three layers, three neurons (plus one bias parameter) in the input layer, 30 neurons (plus one bias parameter) in the hidden layer, one neuron in the output layer. The appropriate noise level to be added to the synthetic data is 5%. Due to the chosen architecture, the input to the neural network is given by the four element vector:

$$ I = [R_1, R_2, R_3, B], \quad (3.5) $$

where the bias parameter $B$ is always set to one, and the elements $R_1$, $R_2$, and $R_3$ are given by:

$$ R_1 = 0.05 + 0.305 \times (\ln (R_{RS443}/R_{RS555}) + 0.829), \quad (3.6a) $$
$$ R_2 = 0.05 + 0.411 \times (\ln (R_{RS490}/R_{RS555}) + 0.473), \quad (3.6b) $$
$$ R_3 = 0.05 + 0.805 \times (\ln (R_{RS515}/R_{RS555}) + 0.382). \quad (3.6c) $$

Appropriate pre-processing is essential for a successful training of ANNs: the transformations (3.6a)-(3.6c) have the aim a) to represent the roughly linear dependence of the logarithms of colour ratio and pigment concentration and b) to map the input data to an interval confined to the range [0.05, 0.95]. Choosing this interval instead of [0.0, 1.0] allows for pigment retrieval even if the input values are slightly outside the range of the synthetic training data set. However, such extrapolation should be handled with great care. The temperature constant of the sigmoid was given a value of $c_t = 1.0$. The elements of the weight matrices

$$ W^{HH} = \begin{pmatrix} W^{HH}_{1,1} & \cdots & W^{HH}_{1,4} \\ \cdots & \cdots & \cdots \\ W^{HH}_{31,1} & \cdots & W^{HH}_{31,4} \end{pmatrix}, $$

(3.7a)
and

\[ W^{HO} = \begin{pmatrix} W_{1,1}^{HO} & \ldots & W_{1,31}^{HO} \end{pmatrix}, \]

(3.7b)

the one-element output vector \( \bar{O} = [O_1] \) of the ANN is re-transformed into the pigment concentration by:

\[ Chl = \exp ((O_1 - 0.05) / 0.124 - 3.71) . \]

(3.8)

3.4. Evaluating the Performance of the ANN-based Pigment Retrieval Algorithm

3.4.1. Assessing the Performance of the Trained ANN

The potential of the selected ANN for pigment retrieval from real measurements is assessed in three steps by applying it a) to the synthetic data ("training data") used for the ANN training, b) to the SeaBAM data ("validation data") used to determine the ANN architecture and noise level to be added to the input data, and c) to the COASTLOOC data ("test data") which have not been used for the ANN development.

3.4.1.1. Performance with respect to the training data

The ANN pigment concentration forecasts using simulated remote sensing reflectance ratios as input data are compared to the pigment concentrations used as input for the corresponding RT simulations. As shown in Figure 3.1, the inversion is successful with regard to the synthetic training data set \( (r^2 = 0.993, \ RMSE = 0.0787 \) for the log-transformed pigment concentration). Figure 3.1 also reveals a limitation of the synthetic training data set: For small pigment concentrations, the dependence of the simulated remote sensing reflectance on the pigment concentration is weak. The ANN may therefore not distinguish between pigment concentrations less than approx. 0.06 mg m\(^{-3}\), although the training data comprises pigment concentrations between 0.025 and 35 mg m\(^{-3}\).

3.4.1.2. Performance with respect to the validation data

In a second step, the ANN is applied to the in-situ measurements of the remote sensing reflectance contained in the SeaBAM data set (900 stations in Case I waters, Figure 3.2A). Here again, the pigment concentrations derived from the SeaBAM reflectance data agree well (as compared to empirical algorithms, see section 3.4.2) with the corresponding in-situ measurements \( (r^2 = 0.934, \ RMSE = 0.148 \) for the log-transformed pigment concentration). This is not surprising, since the SeaBAM data have a) been used to derive the back scattering model for marine particles used for the RT simulations [Zhang et al., 2003], and b) to select the most appropriate ANN architecture and noise level.
3.4.1.3. Performance with respect to the test data

In a third step, the ANN is applied to 67 COASTLOOC Case I reflectance spectra (Figure 3.2B). Here, it is assumed that the ratio of the hemispherical reflectance at two wavelengths is identical to the corresponding remote sensing reflectance ratio. Satisfactory (again as compared to the empirical algorithms) performance ($r^2 = 0.892$, RMSE = 0.219 for the log-transformed pigment concentration) is observed, even though the COASTLOOC data are totally independent from the SeaBAM data and have not been used in any respect to derive the ANN. This is a hint that the retrieval method presented herein has a potential to be applied on a regional or global scale. However, ANN-derived pigment concentrations from the COASTLOOC reflectance data seem to be systematically overestimated by about 30 %. Since a similar effect is also observed when applying empirical algorithms, the observed behavior appears to be data driven. The instrumental problems are mainly suspected to be reason behind these deviations, since the radiometer used for all COASTLOOC and ALMOFRONT campaigns showed a significant degradation in several spectral channels which might not have fully been compensated by the applied correction procedure. Besides, systematic differences between corresponding spectral ratios of the remote sensing reflectance and the hemispherical reflectance might also contribute to the observed deviations.
3.4.2. Comparison with Existing Empirical Algorithms

In order to further evaluate the performance of the trained ANN, it is compared to the most successful empirical algorithms compiled in O’Reilly et al. [1998], listed here in Table 3.5. The algorithms to be compared are applied to the reflectance data from 900 SeaBAM Case I stations. The performance of the different algorithms is given in Table 3.6 and is depicted in Figure 3.3. The ANN has the highest $r^2$ and lowest RMSE of all compared algorithms. The relative success of the ANN-based pigment retrieval is partly explained by the fact that the underlying IOP models represent the SeaBAM data well. A further reason is that it uses more spectral information than do the empirical algorithms. The performance of an ANN using one colour ratio
\((R_{RS443}/R_{RS555})\) slightly drops \((r^2 = 0.929\) instead of \(r^2 = 0.934\), \(\text{RMSE} = 0.156\) instead of \(\text{RMSE} = 0.148\)); however, it is still above that of the corresponding empirical algorithms (e.g., MOREL-3: \(r^2 = 0.918\), \(\text{RMSE} = 0.175\), OC2B: \(r^2 = 0.924\), \(\text{RMSE} = 0.156\)).

Table 3.5. Empirical ocean colour algorithms based on remote sensing reflectance [O'Reilly et al., 1998].

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Functional form</th>
<th>Band ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>POLDER</td>
<td>(C = 10^{(a_0 + a_1R + a_2R^2 + a_3R^3)})</td>
<td>(R = \log_{10}(R_{RS443}/R_{RS555}))</td>
</tr>
<tr>
<td>CalCOFI 2-band cubic</td>
<td>(C = 10^{(a_0 + a_1R + a_2R^2 + a_3R^3)})</td>
<td>(R = \log_{10}(R_{RS490}/R_{RS555}))</td>
</tr>
<tr>
<td>MOREL-3</td>
<td>(C = 10^{(a_0 + a_1R + a_2R^2 + a_3R^3)})</td>
<td>(R = \log_{10}(R_{RS443}/R_{RS555}))</td>
</tr>
<tr>
<td>MOREL-4</td>
<td>(C = \exp^{(a_0 + a_1R + a_2R^2 + a_3R^3)})</td>
<td>(R = \log_{10}(R_{RS490}/R_{RS555}))</td>
</tr>
<tr>
<td>OC2</td>
<td>(C = 10^{(a_0 + a_1R + a_2R^2 + a_3R^3) + a_4})</td>
<td>(R = \log_{10}(R_{RS490}/R_{RS555}))</td>
</tr>
<tr>
<td>OC2B</td>
<td>(C = 10^{(a_0 + a_1R + a_2R^2 + a_3R^3) + a_4})</td>
<td>(R = \log_{10}(R_{RS443}/R_{RS555}))</td>
</tr>
<tr>
<td>OC4</td>
<td>(C = 10^{(a_0 + a_1R + a_2R^2 + a_3R^3) + a_4})</td>
<td>(R = \log_{10}(R_{RS443}/R_{RS490}/R_{RS510})/R_{RS555})</td>
</tr>
</tbody>
</table>

Table 3.6. Performance of the ANN-based pigment retrieval scheme as compared to selected empirical algorithms when applied to the SeaBAM and COASTLOOC data sets.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>SeaBAM Case I (N = 900)</th>
<th>COASTLOOC Case I (N = 67)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>(r^2)</td>
</tr>
<tr>
<td>ANN</td>
<td>0.148</td>
<td>0.934</td>
</tr>
<tr>
<td>POLDER</td>
<td>0.256</td>
<td>0.919</td>
</tr>
<tr>
<td>CalCOFI 2 band cubic</td>
<td>0.186</td>
<td>0.918</td>
</tr>
<tr>
<td>MOREL-3</td>
<td>0.175</td>
<td>0.918</td>
</tr>
<tr>
<td>MOREL-4</td>
<td>0.198</td>
<td>0.906</td>
</tr>
<tr>
<td>OC2</td>
<td>0.162</td>
<td>0.918</td>
</tr>
<tr>
<td>OC2B</td>
<td>0.156</td>
<td>0.924</td>
</tr>
<tr>
<td>OC4</td>
<td>0.151</td>
<td>0.928</td>
</tr>
</tbody>
</table>
3.4.3. Resistance against Noise

In order to compare the resistance against noise of the trained ANN to that of the empirical algorithms, random noise of 0–30% was added to the SeaBAM remote sensing reflectance ratios before offering them as input to the pigment retrieval algorithms. Comparing RMSE and $r^2$ (Table 3.7), one observes that the ANN-based algorithm has a higher performance against noise than do have the empirical algorithms (Figure 3.4). This is a fundamental advantage when applying algorithms to real measurements.
Table 3.7. Performance against noise of the ANN-based pigment retrieval scheme as compared to empirical algorithms. All algorithms were applied to SeaBAM remote sensing reflectance ratios to which different levels of noise have been added. The following acronyms are used: POL = POLDER, CCO = CalCOFI two-band cubic, MO3 = MOREL-3, MO4 = MOREL-4, OC2, OC2B and OC4 = Ocean Chlorophyll algorithms 2, 2B and 4 [O’Reilly et al., 1998].

<table>
<thead>
<tr>
<th>Noise</th>
<th>ANN</th>
<th>POL</th>
<th>CCO</th>
<th>MO3</th>
<th>MO4</th>
<th>OC2</th>
<th>OC2B</th>
<th>OC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 %</td>
<td>RMSE</td>
<td>0.148</td>
<td>0.256</td>
<td>0.186</td>
<td>0.175</td>
<td>0.198</td>
<td>0.162</td>
<td>0.156</td>
</tr>
<tr>
<td></td>
<td>$r^2$</td>
<td>0.934</td>
<td>0.919</td>
<td>0.918</td>
<td>0.918</td>
<td>0.906</td>
<td>0.918</td>
<td>0.924</td>
</tr>
<tr>
<td>5 %</td>
<td>RMSE</td>
<td>0.150</td>
<td>0.258</td>
<td>0.189</td>
<td>0.177</td>
<td>0.203</td>
<td>0.167</td>
<td>0.158</td>
</tr>
<tr>
<td></td>
<td>$r^2$</td>
<td>0.932</td>
<td>0.917</td>
<td>0.915</td>
<td>0.916</td>
<td>0.901</td>
<td>0.913</td>
<td>0.921</td>
</tr>
<tr>
<td>10 %</td>
<td>RMSE</td>
<td>0.155</td>
<td>0.263</td>
<td>0.197</td>
<td>0.181</td>
<td>0.213</td>
<td>0.179</td>
<td>0.164</td>
</tr>
<tr>
<td></td>
<td>$r^2$</td>
<td>0.927</td>
<td>0.912</td>
<td>0.906</td>
<td>0.911</td>
<td>0.890</td>
<td>0.901</td>
<td>0.915</td>
</tr>
<tr>
<td>20 %</td>
<td>RMSE</td>
<td>0.174</td>
<td>0.281</td>
<td>0.227</td>
<td>0.200</td>
<td>0.251</td>
<td>0.226</td>
<td>0.186</td>
</tr>
<tr>
<td></td>
<td>$r^2$</td>
<td>0.910</td>
<td>0.892</td>
<td>0.871</td>
<td>0.891</td>
<td>0.850</td>
<td>0.850</td>
<td>0.893</td>
</tr>
<tr>
<td>30 %</td>
<td>RMSE</td>
<td>0.197</td>
<td>0.310</td>
<td>0.271</td>
<td>0.227</td>
<td>0.304</td>
<td>0.264</td>
<td>0.220</td>
</tr>
<tr>
<td></td>
<td>$r^2$</td>
<td>0.888</td>
<td>0.861</td>
<td>0.818</td>
<td>0.859</td>
<td>0.790</td>
<td>0.811</td>
<td>0.856</td>
</tr>
</tbody>
</table>

3.5. Conclusions

In this study, a method for pigment retrieval from ocean colour in Case I waters has been derived and evaluated. The retrieval method is derived by applying ANN techniques to a set of remote sensing reflectance spectra typical of Case I waters, which have been obtained from RT simulations and the IOP models compiled in Table 3.1.

The ANN employed in this study has three layers: one input layer consisting of three neurons (plus one bias parameter), one hidden layer consisting of 30 neurons (plus one bias parameter), and one output layer consisting of one neuron. The three neurons in the input layer correspond to the remote sensing reflectance ratios $R_{RS443}/R_{RS555}$, $R_{RS490}/R_{RS555}$, and $R_{RS515}/R_{RS555}$. Applying the trained ANN to the Case I SeaBAM data gives a correlation between predicted and measured pigment concentrations of $r^2 = 0.934$ and a root mean square error of RMSE = 0.148; applying it to the Case I COASTLOOC data which have not been used to derive the phase function model of marine particles used for the RT simulations, results in $r^2 = 0.892$ and RMSE = 0.219.

The performance of the ANN-based pigment retrieval scheme is comparable to the most successful empirical algorithms: applying e.g., the SeaWiFS algorithm OC4 to the SeaBAM data set gives $r^2 = 0.928$, and RMSE = 0.151 as compared to $r^2 = 0.934$ and RMSE = 0.148 for the
Figure 3.4. Performance against noise of the ANN-based algorithm presented in this study and selected empirical algorithms, using square of correlation coefficient $r^2$ (A) and root mean square error RMSE (B) of derived vs. measured pigment concentrations as measure of success.

ANN-based algorithm. The resistance against noise of the ANN-based algorithm is superior as compared to the empirical algorithms: adding 20 % random noise to the SeaBAM remote sensing reflectance ratios before offering them to the pigment retrieval algorithms, $r^2$ of OC4 drops from 0.928 to 0.886, while only going down from 0.934 to 0.910 for the ANN-based algorithm. Correspondingly, the RMSE of OC4 is increased from 0.151 to 0.192, while only rising from 0.148 to 0.174 for the ANN-based algorithm. It is this resistance against noisy input data which make ANN-based pigment retrieval schemes a very suitable tool for studying the marine
environment from space-borne Earth Observation data. This is especially true considering the fact that the presented method in principle allows the inclusion of viewing geometry, sun position, rough air-sea interface, etc. In fact, all processes may be considered for which a physical or statistical model exists that can be integrated into the RT code.
Chapter 4

MODELLING THE BACKSCATTERING PROBABILITY OF MARINE PARTICLES IN CASE II WATERS

4.1. Background

In order to develop new methods to retrieve bio-geochemical or physical parameters of the marine environment from ocean colour data, either based on semi-analytic models or radiative transfer simulations, parameterisations of the inherent optical properties (IOPs) of the various constituents present in sea water are required. Such parameterisations are meanwhile rather well established for Case I waters [Bricaud et al., 1981; Bricaud et al., 1998; Loisel and Morel, 1998; Zhang et al., 2003]. The constituents present in Case II waters are characterised by a larger variability as compared to Case I waters. This variability as well as the difficulties encountered with in-situ determination of bio-optical parameters are the main reasons for the actual lack of bio-optical models for Case II waters. During the COASTLOOC (Coastal Surveillance Through Observation of Ocean Colour) project [Babin, 2000], bio-optical models of the absorption of non-chlorophyllous particles and coloured dissolved organic matter (CDOM) as well as marine particles scattering were developed from measurements taken in European coastal waters. One objective of this study is to develop a bio-optical model of the backscattering probability of marine particles which shall allow to simulate the light field in Case II waters with the accuracy required for the development of retrieval schemes for oceanic constituents. The model was derived by optimising the agreement between radiative transfer simulations of the hemispherical reflectance just below the sea surface and the corresponding COASTLOOC data.

4.2. The COASTLOOC Data Set

The COASTLOOC [Babin, 2000] data set has been used to derive the model for the backscattering probability of marine particles presented in this study. The COASTLOOC data set relates the subsurface hemispherical reflectance at 13 wavelengths between 412 nm and 865 nm to a variety of IOPs and water constituents concentrations. Most of the 424 stations visited during COASTLOOC have been gathered in Case II waters, except for 93 stations located in Case I waters in the Atlantic Ocean and Mediterranean Sea. The locations of stations visited during all COASTLOOC cruises are shown in Figure 4.1. The measurements taken in the Adriatic Sea, Baltic Sea, English Channel and North Sea were used for the development of the backscattering model. The measurements at 196 of the 243 stations in these areas fulfilled the requirements in terms of completeness and accuracy. More information on COASTLOOC is given in Table 4.1.
Figure 4.1. The locations of stations visited during COASTLOOC cruises, taken from Babin [2000].

Absorption spectra of particles were measured using the modified Filter-Transfer-Freeze (FTF) method, which eliminates the path length amplifications effect occurring with measurements on GF/F filters [Allali et al., 1995]. The water sample was filtered onto 0.4µm polycarbonate membrane filters. The optical density of the particles was measured from 350 to
800 nm by a Perkin-Elmer Lambda 19 double-beam spectrophotometer. Then absorption coefficients of particles were obtained from the optical densities. The distinction of absorption spectra of particles \( a_p \) into algal \( a_{p1} \) and non-algal \( a_{p2} \) components was performed by numerical decomposition \cite{Bricaud and Stramski, 1990}.

Table 4.1. Information on the COASTLOOC data set

<table>
<thead>
<tr>
<th>Campaign</th>
<th>Provider / PI</th>
<th>Area</th>
<th>Time</th>
<th>Case II stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>COASTLOOC 1</td>
<td>M. Babin</td>
<td>North Sea</td>
<td>April 1997</td>
<td>12</td>
</tr>
<tr>
<td>COASTLOOC 2</td>
<td>M. Babin</td>
<td>Mediterranean Sea</td>
<td>July 1997</td>
<td>15</td>
</tr>
<tr>
<td>COASTLOOC 3</td>
<td>M. Babin</td>
<td>Northern Adriatic Sea</td>
<td>July, August 1997</td>
<td>34</td>
</tr>
<tr>
<td>COASTLOOC 4</td>
<td>M. Babin</td>
<td>Mediterranean Sea (Golfe du Lion)</td>
<td>October 1997</td>
<td>36</td>
</tr>
<tr>
<td>COASTLOOC 5</td>
<td>M. Babin</td>
<td>English Channel, North Sea</td>
<td>May 1998</td>
<td>56</td>
</tr>
<tr>
<td>COASTLOOC 6</td>
<td>M. Babin</td>
<td>English Channel, North Sea, Baltic Sea</td>
<td>Sep. 1998</td>
<td>178</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>331</td>
</tr>
</tbody>
</table>

Regarding the measurement of CDOM absorption, the particulate material was firstly removed from seawater samples by 0.4\(\mu\)m membrane filters. Then the optical density of the filtrate were measured from 350 to 800 nm by a Perkin-Elmer-Lambda-19 double-beam spectrophotometer. Finally, absorption coefficients of particles were derived from optical densities \cite{Ferrai and Dowell, 1998}.

The spectral absorption, \( a(\lambda) \), and beam attenuation, \( c(\lambda) \), coefficients were measured using an ac-9 profiler (WET Labs) deployed either from a ship or from a helicopter. The particle scattering coefficient, \( b_p(\lambda) \), is then obtained directly from \( c(\lambda)-a(\lambda) \).

Suspended particulate matter was measured as the weight of material (dried at 70°C) collected on pre-weighted GF/F filters. Pigment concentrations were measured using high pressure liquid chromatography (HPLC) techniques, using the procedure described by Vidussi et al. [1996]. The concentration of particulate organic carbon (POC) was determined using a Carlo-Erba-NCS-2500 elemental analyzer. The concentration of particulate organic matter (POM) was obtained using the following statistical relationship \cite{van Raaphorst and Melschaert, 1996}:

\[
POM = 2.6 \times POC
\]  (4.1)
Downward and upward irradiance at 13 wavelengths was measured by a SPMR radiometer (Satlantic Inc.). The subsurface hemispherical reflectance was obtained by the division of the extrapolated upward and downward irradiance just below sea surface.

4.3. Inherent Optical Properties of Oceanic Constituents in Case II Waters

Oceanic constituents in Case II waters can be divided into 5 groups [Babin, 2000]: pure seawater, phytoplankton and associated particles, endogenous CDOM, exogenous particles, and exogenous CDOM. The corresponding absorption and scattering coefficients are, respectively, \( a_w \) and \( b_w \), \( a_{p1} \) and \( b_{p1} \), \( a_{p2} \) and \( b_{p2} \), and \( a_y \). The total absorption coefficient in Case II waters is composed of the contributions of all absorbing substances:

\[
a = a_w + a_{p1} + a_{p2} + a_y
\]

(4.2)

In practice, the contributions of endogenous CDOM and exogenous CDOM (\( a_{y1} \) and \( a_{y2} \)) to total CDOM absorption can not be separated from each other. Therefore, \( a_y \) is used as the sum of \( a_{y1} \) and \( a_{y2} \). So, Eq. (4.2) can be modified as:

\[
a = a_w + a_{p1} + a_{p2} + a_y
\]

(4.3)

The total scattering coefficient in Case II waters is composed of the contributions of all scattering substances:

\[
b = b_w + b_{p1} + b_{p2}
\]

(4.4)

where the subscripts \( p1 \) and \( p2 \), respectively, indicate endogenous and exogenous particles. The scattering of CDOM is assumed as zero. Actually, the contributions of endogenous and exogenous particles to total particles scattering can not be separated from each other. Therefore, \( b_p \) is used as the sum of the \( b_{p1} \) and \( b_{p2} \). So, the Eq. (4.4) can be rewritten as:

\[
b = b_w + b_p
\]

(4.5)

The IOPs models used in this chapter are listed in Table 4.2. More information on these models has been given in Section 2.2. Some of the IOPs models, such as for pure sea water and the absorption coefficient of phytoplankton, are commonly used. Other IOPs models, such as for the absorption coefficient of non-algal particles (exogenous particles) and yellow substance, as well as scattering coefficient of particles, were derived from the COASTLOOC data set.

The atmosphere has been considered in the simulations in order to provide a realistic diffuse light field incident on the sea surface. Since variations of the atmospheric composition have little influence on the oceanic reflectance, an atmosphere with Maritime aerosol and Rayleigh scattering for a constant surface air pressure was deemed sufficient for the execution of the present study.
Table 4.2. Parameterisation of inherent optical properties of pure sea water and constituents of Case II waters.

<table>
<thead>
<tr>
<th>Constituents</th>
<th>IOP</th>
<th>Model / Measurement</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pure sea water</td>
<td>Absorption</td>
<td>Directly measured</td>
<td>Pope and Fry [1997]</td>
</tr>
<tr>
<td></td>
<td>Scattering</td>
<td>$b_w(\lambda) = 0.00288\left(\frac{\lambda}{500}\right)^{-4.32}$</td>
<td>Morel [1974]</td>
</tr>
<tr>
<td></td>
<td>Phase function</td>
<td>$p_w(\lambda) = 0.06225(1 + 0.835 \cos^2 \theta)$</td>
<td>Morel [1974]</td>
</tr>
<tr>
<td>Particulate matter</td>
<td>Absorption</td>
<td>$a_p(\lambda) = A_p(\lambda)[Chl]^{B_p(\lambda)}$</td>
<td>Bricaud et al. [1998]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$a_{p2}(\lambda) = a_{p2}(443)e^{-Sp(\lambda-443)}$, $a_{p2}(443) = A_{p2} &lt; SPM &gt;^{B_{p2}}$, $S_p=0.0122$, $A_{p2}=0.0216$, $B_{p2}=1.0247$</td>
<td>Babin [2000]</td>
</tr>
<tr>
<td></td>
<td>Scattering</td>
<td>$b_p(\lambda) = A &lt; SPM &gt;$, $A=0.5$</td>
<td>Babin [2000]</td>
</tr>
<tr>
<td></td>
<td>Back scattering probability</td>
<td>$\tilde{b}_b = f(r, \lambda)$</td>
<td>To be developed in this thesis</td>
</tr>
<tr>
<td>CDOM</td>
<td>Absorption</td>
<td>$a_y(\lambda) = a_y(443)e^{-S_y(\lambda-443)}$, $S_y=0.0176$</td>
<td>Babin [2000]</td>
</tr>
</tbody>
</table>

4.4. Comparing *in-situ* Data and Simulations of the Marine Light Field Using Petzold’s Phase Function for Marine Particles

The computer code MOMO [*Fell and Fischer*, 2001] was used to create the radiative transfer simulations. The assumptions and simplifications for the simulations are the same as used in Section 3.2.

Based on the above assumptions and the chosen IOP models in Table 4.2, simulations of the hemispherical reflectance just below the sea surface were made for:

- 196 combinations of constituent concentrations corresponding to those measured at the 196 COASTLOOC stations mentioned above,
- 8 wavelengths: 411, 443, 490, 532, 559, 619, 665, and 705 nm,
- 17 solar zenith angles between 0° and 87°.

The particle phase function most often used for radiative transfer calculations of the marine light field has been derived by *Mobley et al.* [1993] based on measurements on water samples obtained in waters ranging from turbid to oligotrophic [*Petzold*, 1972]. This particle phase function, referred to as Petzold’s phase function, is characterised by a back scattering probability
of 1.81 %. As shown below, using one single phase function for marine particles does not allow to accurately simulate the in-water light field over larger wavelengths and constituents concentration ranges. Using absorption and scattering coefficients of pure seawater, particles, and CDOM as defined in Table 4.2, as well as Petzold’s phase function as input to the radiative transfer simulations, simulated and measured hemispherical reflectance just below the sea surface compare as follows (Figure 4.2):

1. There is significant scatter around the 1:1 line of simulations against measurements;
2. The differences between measurements and simulations decrease with increasing wavelength;
3. The simulated reflectances in the Baltic Sea are generally overestimated, except for the simulations at 705 nm.

### 4.5. Modelling the Marine Particle Phase Function in Case II Waters

Assuming that absorption and total scattering coefficients of CDOM and marine particles contained in the COASTLOOC data set are comparably well represented by the corresponding models developed by Babin [2000], the observed differences between measured and simulated spectral reflectance must primarily be associated with the phase function employed, largely characterised by the back scattering probability. In order to reconcile simulations and measurements, a new back scattering probability model is proposed for marine particles in Case II waters based on the Fournier-Forand phase function [Fournier and Forand, 1994], in the form given by Haltrin [1998]. The new backscattering model will be used to select the appropriate phase function for RT simulations. In practice, the back scattering probability model was derived through the following steps:

1. A set of 40 Fournier-Forand phase functions with back scattering probabilities ranging from 0.2 % to 5.0 % was established;
2. For each of 196 constituents combinations actually measured, the hemispherical reflectances at 8 wavelengths for the 40 Fournier-Forand phase functions mentioned above have been calculated;
3. The most appropriate phase function for the specific constituents combination and wavelength in terms of minimum difference between simulated and measured hemispherical reflectance was determined;
4. The back scattering probability corresponding to the selected phase function is related to wavelengths and the ratio of POM to SPM (see below).
Figure 4.2. Comparison between the simulated hemispherical reflectances just below the sea surface and the corresponding COASTLOOC data. The simulations have been performed using Petzold’s phase function (corresponding to a constant back scattering probability of 1.81% ).
A model for relating the ratio of POM to SPM and wavelength to \( \tilde{b}_b(\lambda) \) was derived from the results of step (4), expressed as:

\[
\tilde{b}_b(r, \lambda) = 0.244 + 1.82 \times 10^{-3} \times \lambda + 0.5 \left[ 1 - \frac{1}{1 + \exp[-10 \times (r - 0.5)]]} \right],
\]

where \( r = \text{POM/SPM} \), and \( \lambda \) is wavelength. Eq. (4.6) may be applied within the ranges \( 0.1 \leq r \leq 1 \) and \( 411 \leq \lambda \leq 705 \) nm. If \( r \leq 0.1 \) then \( \tilde{b}_b(r, \lambda) = \tilde{b}_b(0.1, \lambda) \); if \( r \geq 1 \) then \( \tilde{b}_b(r, \lambda) = \tilde{b}_b(1, \lambda) \). The resulting model is plotted in Figure 4.3.

![Figure 4.3](image)

Figure 4.3. The model for the back scattering probability of marine particles in Case II waters as function of the ratio of POM to SPM and wavelength.

Theoretically, the back scattering probability of marine particles mainly depends on the ratio of organic to inorganic particles and wavelengths. Marine particles consist of two types: organic particles and inorganic particles. The scattering properties of these two types are different from each other. Organic particles are characterised by a lower refractive index, which results in smaller back scattering probabilities. Inorganic particles are characterised by higher refractive index, which results in larger back scattering probabilities. This is the reason that the particle back scattering probability is related to the ratio of POM to SPM.

In Eq. (4.6), \( r \) is a variable. Therefore, if this model is applied to radiative transfer simulations, there is one more variable besides CHL, SPM and CDOM. However, this will make it more complex to retrieve oceanic constituents. In order not to increase the complexity of oceanic constituents retrieval, a replacement of POM was found in this thesis, which is a combination of CHL and CDOM.

During the COASTLOOC cruises, the concentration of particulate organic matter (POM) was determined. In order to investigate the correlation between POM and other variables, the concentration of POM was related to \( a_v(443) \), \( a_v(443) + a_{ph}(443) \) (shown in Figure 4.4), and \( a_{ph}(443) \), respectively. These relationships are expressed as:
\[
\log(POM) = 0.587 + 0.859 \times \log(a_y(443)) \quad (r^2 = 0.689, N = 283) \quad (4.7)
\]
\[
\log(POM) = 0.457 + 0.995 \times \log(a_{ph}(443) + a_y(443)) \quad (r^2 = 0.789, N = 258) \quad (4.8)
\]
\[
\log(POM) = 0.707 + 0.875 \times \log(a_{ph}(443)) \quad (r^2 = 0.627, N = 269) \quad (4.9)
\]

From the above results, it can be seen that there is a statistical significant relationship between POM and \(a_y(443) + a_{ph}(443)\). Therefore, \(a_y(443) + a_{ph}(443)\) through Eq. (4.8) is used to represent POM in Eq. (4.6). The resulting model of the backscattering probability is plotted in Figure 4.5.

Figure 4.4. The relationship between \(a_y(443) + a_{ph}(443)\) and the concentration of particulate organic matter. The black line was plotted based on Eq. (4.8).

Figure 4.5. The model for the backscattering probability of marine particles in Case II waters as function of the ratio of \(a_y(443) + a_{ph}(443)\) to SPM and wavelength.

When applying this new model to the RT simulations (Figure 4.6), the differences between the simulations and the measurements are significantly reduced as compared to RT simulations based on Petzold’s phase function (Figure 4.7).
Figure 4.6. Comparisons between the simulated hemispherical reflectances and the corresponding COASTLOOC data. The simulations have been performed with the new model for the back scattering probability of marine particles proposed in this study.
4.6. Discussion

The new model for the backscattering probability of marine particles may be qualitatively interpreted by Mie theory. The higher the ratio of $a_s(443) + a_{ph(443)}$ to SPM, the higher is the proportion of POM to SPM, which results in smaller backscattering probability. In the Baltic Sea, the proportion of POM to SPM is very high [Babin et al., 2003], marine particles should therefore have lower backscattering probability.

However, due to the lack of information on size distribution and complex refractive index, it is impossible to quantitatively validate the presented backscattering probability model. Here, it
was examined if this model is representative of the conditions in natural marine environments using known information on size distribution and complex refractive index by Mie theory.

It was assumed that particles in natural marine environments consist of two types: organic and inorganic particles. The particle size distribution in marine environments is commonly described using a Junge size distribution (a power law model), with slope parameters typically ranging between 3 and 5 [Bader, 1970; Sheldon et al., 1972; Jonasz, 1983]. Here, minimal and maximal size of marine particles are set equal to 0.02 and 200 µm, respectively. The contribution from particles beyond this range to the total scattering can be ignored [Morel and Ahn, 1991; Stramski and Kiefer, 1991; Ulloa et al., 1994; Babin et al., 2003]. For the real part of the relative refractive index of mineral particles, a typical value of 1.17 has been chosen [Lide, 2001]. The real part of the relative refractive index of organic particles is set to 1.115. It is higher than that of the living algae (1.05) [Brown & Gordon, 1974; Morel & Ahn, ibid; Stramski & Kiefer, 1991]. Aas [1996] found that the most important factor for the variation of the refractive index is the algal water content. For the algal dry mass, the average relative refractive index is about 1.15. Organic particles consist of living algal particles and detrital organic particles. Detrital organic particles are characterised by lower water content. Therefore, the real refractive index of organic particles is higher than that of living algal particles (1.05), but lower than that of the algal dry mass (1.15). The imaginary refractive indexes for organic particles and inorganic particles are taken from Babin et al. [2003]. The back scattering probabilities calculated by the new model and estimated by Mie theory are depicted in Figure 4.8 for three wavelengths. The inputs for Mie calculations are listed in Table 4.3. As shown in Figure 4.8, a good agreement between the model proposed in this study and Mie calculations was obtained. The model for the back scattering probabilities of marine particles in Case II waters proposed in this study is therefore deemed to be representative of the conditions in natural marine environments.

Table 4.3. Inputs for Mie calculations.

<table>
<thead>
<tr>
<th>Particle Compound</th>
<th>Size distribution</th>
<th>slope</th>
<th>Size range</th>
<th>Refractive index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organic particles</td>
<td>Power-law</td>
<td>3.6</td>
<td>0.02~200 µm</td>
<td>1.115-0.0025<em>a</em>ph(λ)</td>
</tr>
<tr>
<td>Inorganic particles</td>
<td>Power-law</td>
<td>3.9</td>
<td>0.02~200 µm</td>
<td>1.170-0.0045<em>a</em>nap(λ)</td>
</tr>
</tbody>
</table>
Figure 4.8. Comparison between the model of the back scattering probability proposed in this study (Equation 4.6) and the Mie calculations. The inputs of the Mie calculations are listed Table 4.3.
Chapter 5

RETRIEVAL OF OCEANIC CONSTITUENTS IN CASE II WATERS

5.1. Introduction

Algorithms for operational retrieval of chlorophyll concentration in Case I waters from satellite ocean colour data are now available. These algorithms often fail in Case II waters. The following facts make the retrieval of oceanic constituents in Case II waters more difficult than in Case I waters:

(1). There are more constituents present in Case II waters. Ocean colour at any wavelength of interest can not be related directly to any one single constituent.

(2). Some of the IOPs of the various constituents which influence ocean colour are similar. For example, the absorption spectra of both CDOM and non-chlorophyllous particles can be modelled by a similar exponential function; the absorption coefficients of phytoplankton and CDOM decrease both from about 440 nm to 550 nm. These phenomena may result in similar ocean colour for different combinations of the constituents concentrations. In these cases, ambiguities may result when retrieving the constituents concentrations from ocean colour.

(3). The compositions of exogenous particulate matter and exogenous CDOM in Case II waters vary strongly with time and region. This means that the IOPs of exogenous particulate matter and exogenous CDOM may greatly vary from time to time and region to region.

(4). There is strong scattering of SPM. If the concentrations of SPM is high, it dominates the water-leaving radiance and may overshadow the contributions of other constituents to the measured signal.

In the following, a method is proposed for the retrieval of the oceanic constituent concentration in Case II waters, based on Artificial Neural Network (ANN) techniques. Input to the presented method is the spectral hemispherical reflectance just below the sea surface. A synthetic data set from radiative transfer simulations is used for the training of an ANN. As mentioned in Chapter 1, a prerequisite for this is that the Inherent Optical Properties (IOPs) of the water constituents required as input to the RT simulations are well representing the conditions to which the trained ANN is later on applied.

As shown in Chapter 4, using a newly developed model of the back scattering probability for marine particles together with other bio-optical models developed from the COASTLOOC data set, the simulated hemispherical reflectance just below the sea surface agrees well with the
corresponding *in-situ* measurements. The IOPs models used herein are therefore deemed to satisfactorily represent the situation encountered in European coastal waters during the COASTLOOC campaigns. This justifies to use the synthetic data set as training data in this study.

5.2. Data Sets

There are two different kind of data used in this study: *in-situ* measurement and RT simulations. In both cases, the data sets relate hemispherical reflectance to the three oceanic parameters: pigment concentration, total suspended particulate matter concentration and the absorption coefficient of CDOM at 443 nm. Based on their individual role in this study, the three different data sets are referred to as:

1. *training data*: a synthetic data set obtained from RT simulations used to train the different ANNs,
2. *validation data*: one part of the *in-situ* measurement data that have been used to develop the model of $b_h(\lambda)$. This data is used to evaluate the performance of each individual ANN and such to identify the most appropriate one with respect to the retrieval of specific oceanic constituents,
3. *test data*: the other part of the *in-situ* measurement data. This data is used to assess in how far the ANN-based oceanic constituent retrieval scheme is applicable to independent data.

These three data sets are described in more detail in the following.

5.2.1. Training Data

The synthetic data set used to train the ANNs for oceanic constituent retrieval was created using the computer code MOMO [Fell and Fischer, 2001]. The IOP models of the oceanic constituents are described in Chapter 4. Based on these models, the IOPs of sea water as required for the RTC can be obtained for given concentrations of oceanic constituents.

It is well known that MLPs, the type of ANN used in this study, has good performance for interpolation, but should not be used for extrapolation. Therefore, it must be made sure that the ranges of the parameters in the training data cover the actual ranges observed in the marine environment of interest. In this study, the ranges of the three components are listed in Table 5.1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Concentration Unit</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHL</td>
<td>mg/m(^3)</td>
<td>0.05</td>
<td>50</td>
</tr>
<tr>
<td>SPM</td>
<td>g/m(^3)</td>
<td>0.05</td>
<td>100</td>
</tr>
<tr>
<td>$a_y(443)$</td>
<td>m(^{-1})</td>
<td>0.005</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 5.1. Ranges of oceanic constituent concentrations
When simulating the radiative transfer in Case II waters, it is often assumed that the concentration of each of the constituents is independent from the other constituents. However, as can be seen from Figure 5.1 displaying COASTLOOC measurements, there is a certain degree of covariance between SPM and CHL, CDOM and CHL, as well as CDOM and SPM also in Case II waters. For a given concentration of one constituent, the other two constituents vary no more than 2 orders of magnitude. This fact may be used to greatly reduce the number of radiative transfer calculations used for the development of Case II algorithms. It allows to identify combinations of the oceanic constituents which are very unlikely to occur in the natural environment and therefore need not to be modelled. Again from the COASTLOOC data, upper and lower boundaries were defined for each constituent concentration in relation to the other constituents (see also Figure 5.1):

\[
\begin{align*}
\text{SPM against CHL:} \\
& spm_l(chl) = 0.12 \times CHL^{0.77} \quad (5.1) \\
& spm_h(chl) = 5.89 \times CHL^{0.77} \quad (5.2)
\end{align*}
\]

\[
\begin{align*}
\text{CDOM against CHL} \\
& cdom_lc(chl) = 0.0079 \times CHL^{0.75} \quad (5.3) \\
& cdom_hc(chl) = 0.25 \times CHL^{0.75} \quad (5.4)
\end{align*}
\]

\[
\begin{align*}
\text{CDOM against SPM} \\
& cdom_l_s(spm) = 0.0063 \times SPM^{0.9} \quad (5.5) \\
& cdom_h_s(spm) = 0.44 \times SPM^{0.9} \quad (5.6)
\end{align*}
\]

The combinations of oceanic constituent concentrations used for radiative transfer simulations were selected according to the following steps. Here, a logarithmic distribution of the three oceanic constituent concentrations was applied, so that each order of magnitude is represented with a similar number of cases:

1. CHL is randomly selected within the range 0.05 and 50 mg/m³.
2. For the selected CHL, SPM is randomly selected between the lower and upper boundaries defined by Equations (5.1) and (5.2).
3. For the selected SPM, CDOM at 443 nm is selected between the lower and upper boundaries defined by Equations (5.3) and (5.4), as well as (5.5) and (5.6).
4. Steps (1)–(3) are repeated until a sufficient number of combinations of the oceanic constituents has been generated.

The objective of this approach is to potentially reduce the likelihood for ambiguous solutions for the retrieval of the three constituents by reducing the ranges of the constituent distributions.

Based on the above strategy and using the IOP models defined in Chapter 4, simulations of the hemispherical reflectance just below the sea surface were made for:

- 1000 combinations of three oceanic constituents,
- 8 wavelengths: 411, 443, 490, 509, 559, 619, 665, 705 nm,
• 17 solar zenith angles between 0° and 87°.

Besides, the same assumptions and simplifications were made as specified in Section 3.2.

Figure 5.1. Scatter plots of SPM against CHL (top), CDOM absorption against CHL (middle), and CDOM absorption against SPM (bottom). All data set collected during the COASTLOOC cruises.
5.2.2. In-situ Measurement Data Sets

The in-situ measurement data used in this study come from the following data sets: COASTLOOC (Coastal Surveillance Through Observation of Ocean Colour) [Babin, 2000] and PMNS (Particulate Matter North Sea) [Shimwell et al., 1995]. Detailed information on the COASTLOOC data set has been given in Chapter 4.

During the PMNS programme [Shimwell et al., 1995], IOPs, AOPs and water quality parameters were measured in the southern North Sea during five cruises. This region is commonly described as the Rhine region of fresh water influence. The waters in this area are characterised as Case II waters.

Table 5.2 provides more information on the in-situ measurement data sets chosen for this study. As shown in Table 5.2 (a) and (b), the selected in-situ measurement data sets are divided into two parts: validation data and test data. The validation data (which have been used to develop the back scattering model for marine particles in Chapter 4) consist of seven subsets of COASTLOOC data. The test data consist of the PMNS data and three subsets of the COASTLOOC data. The radiometric parameter specified in the COASTLOOC and PMNS data sets is the hemispherical reflectance just below the sea surface. All validation data set were obtained in Case II waters. The test data were partly obtained in CASE I waters, partly in Case II waters. Only chlorophyll concentration is available in the subset COAST_9. For the other subsets, the concentrations of all three oceanic constituents are available.

5.2.3. Data Processing

The in-situ data used in this study stem from different sources and differ in a number of details. The following processing procedures were applied in order to generate a consistent data set.

(1). Conversion of reflectance at 532 nm to 509 nm

As shown in Table 5.2, some subsets contain data which are not available at some wavelengths. For subsets COAST_1, COAST_3, COAST_5, COAST_8 and COAST_9 there is no reflectance available at 509 nm, but at 532 nm. Subsets COAST_2, COAST_4, COAST_6 and COAST_7 contain reflectance measurements at both 509 nm and 532 nm. A statistical regression between the reflectances at these two wavelengths yields the following equation:

\[ R(509) = -0.00523 + 0.779[R(532)]^{0.922} \]  \( r^2 = 0.985, N = 180 \)  \( (5.7) \)

Hence, \( R(509) \) and \( R(532) \) are highly correlated. Therefore, Equation (5.7) was applied to derive the reflectances at 509 nm in subsets COAST_1, COAST_3, COAST_5, COAST_8, and COAST_9.
Table 5.2 (a) The characteristics of *in-situ* data sets

<table>
<thead>
<tr>
<th>Data set</th>
<th>N</th>
<th>R</th>
<th>Locations</th>
<th>Water type</th>
<th>Available concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CHL</td>
<td>SPM</td>
</tr>
<tr>
<td>Validation data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COAST_1</td>
<td>35</td>
<td>R(0-)</td>
<td>Adriatic Sea</td>
<td>2</td>
<td>yes</td>
</tr>
<tr>
<td>COAST_2</td>
<td>57</td>
<td>R(0-)</td>
<td>Baltic Sea</td>
<td>2</td>
<td>yes</td>
</tr>
<tr>
<td>COAST_3</td>
<td>17</td>
<td>R(0-)</td>
<td>English channel</td>
<td>2</td>
<td>yes</td>
</tr>
<tr>
<td>COAST_4</td>
<td>51</td>
<td>R(0-)</td>
<td>English channel</td>
<td>2</td>
<td>yes</td>
</tr>
<tr>
<td>COAST_5</td>
<td>10</td>
<td>R(0-)</td>
<td>North sea</td>
<td>2</td>
<td>yes</td>
</tr>
<tr>
<td>COAST_6</td>
<td>64</td>
<td>R(0-)</td>
<td>North sea</td>
<td>2</td>
<td>yes</td>
</tr>
<tr>
<td>COAST_7</td>
<td>9</td>
<td>R(0-)</td>
<td>North Sea</td>
<td>2</td>
<td>yes</td>
</tr>
<tr>
<td>Test data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COAST_8</td>
<td>48</td>
<td>R(0-)</td>
<td>Mediterranean Sea</td>
<td>2</td>
<td>yes</td>
</tr>
<tr>
<td>COAST_9</td>
<td>34</td>
<td>R(0-)</td>
<td>Mediterranean Sea</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td>COAST_10</td>
<td>28</td>
<td>R(0-)</td>
<td>Atlantic Ocean</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td>PMNS</td>
<td>131</td>
<td>R(0-)</td>
<td>North Sea</td>
<td>2</td>
<td>yes</td>
</tr>
</tbody>
</table>

Table 5.2 (b) The characteristics of *in-situ* data sets

<table>
<thead>
<tr>
<th>Data set</th>
<th>Wavelength of reflectance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>411 443 490 --- 532 559 619 665 705</td>
</tr>
<tr>
<td>Validation data</td>
<td>COAST_1</td>
</tr>
<tr>
<td>COAST_2</td>
<td>411 443 490 509 532 559 619 665 705</td>
</tr>
<tr>
<td>COAST_3</td>
<td>411 443 490 --- 532 559 619 665 705</td>
</tr>
<tr>
<td>COAST_4</td>
<td>411 443 490 509 532 559 619 665 705</td>
</tr>
<tr>
<td>COAST_5</td>
<td>411 443 490 --- 532 559 619 665 705</td>
</tr>
<tr>
<td>COAST_6</td>
<td>411 443 490 509 532 559 619 665 705</td>
</tr>
<tr>
<td>COAST_7</td>
<td>411 443 490 509 532 559 619 665 --- 705</td>
</tr>
<tr>
<td>Test data</td>
<td>COAST_8</td>
</tr>
<tr>
<td>COAST_9</td>
<td>411 443 490 --- 532 559 619 665 705</td>
</tr>
<tr>
<td>COAST_10</td>
<td>411 443 490 509 532 556 --- 665 705</td>
</tr>
<tr>
<td>PMNS</td>
<td>412 443 490 513 --- 559 622 665 701</td>
</tr>
</tbody>
</table>

(2). Conversion of reflectance at 665 nm to 619 nm

For subsets COAST_7 and COAST_10, the reflectance at 619 nm is not available. A statistical regression between the reflectances at 619 nm and 665 nm in subsets of COAST_1, COAST_2, COAST_3, COAST_4, COAST_5, COAST_6, COAST_8 and COAST_9 has been obtained, expressed as:
\[
\log[R(619)]=0.1106+0.957\log[R(665)]
\]  
\[r^2=0.991, \text{N}=315\]  

\[\text{R(619) and R(665) are also highly correlated. Thus, Eq. (5.8) was applied to derive the reflectance at 619 nm in subsets COAST_7, COAST_10. Since the correlation between R(619) and R(665) \(r^2=0.963, \text{N}=315\) is weaker than that between }\log[R(619)]\text{ and }\log[R(665)], \text{ the log scale was here used instead of the linear scale.}\]

(3). Conversion of the absorption coefficient of CDOM at 380 nm to 443 nm

The PMNS data set indicates the absorption coefficient of CDOM at 380 nm instead of 443 nm. The following equation was used to convert the absorption coefficient of CDOM from wavelength 380 nm to 443 nm [Bricaud et al., 1981]:

\[a_y(443)=a_y(380)e^{-S_y(\lambda_1-380)}\]  

\[\text{where } S_y=0.0176, \text{ and } \lambda_1=443 \text{ nm.}\]

(4). Conversion of the chlorophyll-a concentration to pigment concentration

The PMNS data set comprises the concentration of chlorophyll-a instead of the pigment concentration. The following relationship was used to convert from chlorophyll-a concentration to pigment concentration [O’Reilly et al., 1998]:

\[[\text{pigment}]=1.34\times[\text{chl a}]^{0.983}\]  

5.3. Retrieval of the Oceanic Constituents with ANN

5.3.1. Artificial Neural Network

In this study, the Multi-Layer-Perceptron (MLP) is used to approximate the relationship between ocean colour and the concentrations of the oceanic constituents in Case II waters. The theory of MLP is described in detail in Section 2.4. It consists of three layers: input layer, one hidden layer and output layer. A bias parameter is added both to the input layer and to the hidden layer.

In this study, three different MLPs are used to retrieve pigment, SPM and CDOM separately, rather than using one single MLP to retrieve the three constituents at one time. Thus, in the output layer, there is one neuron which corresponds to one of the three constituents concentrations. Although more effort is required for the training process, it is of advantage to have one trained MLP for each constituent since it allows to individually optimise constituent retrieval (see Section 5.3.3).

To determine which and how many spectral bands or band ratios are best suited for the retrieval of CHL, SPM and CDOM, 13 combinations of input data (listed in Tables 5.3, 5.4 and 5.5, respectively, for the retrieval of CHL, SPM and CDOM) were tested. Of the 13 listed cases, the first four are combinations of absolute reflectance values at different wavelengths. The other
nine cases are combinations of reflectance ratios. The dimensionality of the input data determines the number of neurons in the input layer.

The optimal number of neurons in the hidden layer depends on various factors. The determination of the number of neurons in the hidden layer is described below.

5.3.2. ANN Training

The performance of each of the algorithms based on the inputs listed in Tables 5.3, 5.4 and 5.5 depends mainly on the following two parameters: neurons in hidden layer and noise level added to the training data set. The same procedures as described in Chapter 3 were used to determine the optimal number of hidden neurons and appropriate noise level. To find the optimal number of hidden neurons, the performance of ANNs with 6, 12, 20, and 30 hidden neurons were tested. To determine the appropriate noise level, 10%, 20%, 30%, and 40% noise was added to the synthetic training data used as input.

A synthetic data set for ANN training has been generated from the RT simulations outlined in Section 5.2. The synthetic data set is composed of 1000 hemispherical reflectance spectra, corresponding to 1000 combinations of three constituent concentration values.

For each of the 13 input combinations listed in Tables 5.3, 5.4 and 5.5, there are five training data sets which correspond to five different noise levels of 0%, 10%, 20%, 30%, 40%, respectively, and four ANNs with 6, 12, 20, or 30 neurons in the hidden layer. These five training data sets were used to train the four different ANNs, thus a total of $5 \times 4 = 20$ trained ANNs were obtained, which are the candidates of the corresponding algorithm.

5.3.3. Determining ANN Architecture and Noise Adding for Optimal Oceanic Constituent Retrieval

In order to find the ANN best suited for a specific oceanic constituent retrieval from all the candidate ANNs constructed as described in section 5.3.2, the ANN forecasts were compared to the ‘validation data’ by two error measures: root mean square error (RMSE) (defined in Section 3.3) and the square of the Pearson’s correlation coefficient $r^2$. The optimum number of hidden neurons and the appropriate noise level with regard to the above two error measures are also given in Tables 5.3, 5.4 and 5.5 for each of the retrieval algorithms of three oceanic constituents. From the results shown in Tables 5.3, 5.4 and 5.5, the following conclusions can be drawn:

1. Regarding the retrieval of pigment concentrations, the best results (lowest RMSE and highest $r^2$) were obtained using the seven reflectance ratios as input (case No. 6 in Table 5.3). In this case, the $r^2$ value is 0.729, and the RMSE value is 0.274. It is better than that of case No. 1 (RMSE=0.285, and $r^2=0.707$) which has highest performance of the cases using absolute reflectance value as input.
Table 5.3. Performance with regard to pigment retrieval of spectral band combinations used input to ANNs

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Input Neurons in Hidden layer</th>
<th>Noise level (%)</th>
<th>Training data (N=1000)</th>
<th>Validation data (N=205)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RMSE</td>
<td>$r^2$</td>
</tr>
<tr>
<td>1</td>
<td>R411, R443, R490, R510, R559, R619, R665, R705</td>
<td>12</td>
<td>30</td>
<td>0.240</td>
</tr>
<tr>
<td>2</td>
<td>R411, R443, R490, R510, R559, R619, R665, R705, 0S</td>
<td>6</td>
<td>40</td>
<td>0.268</td>
</tr>
<tr>
<td>3</td>
<td>R411, R443, R490, R510, R559, R665</td>
<td>6</td>
<td>30</td>
<td>0.280</td>
</tr>
<tr>
<td>4</td>
<td>R412, R443, R490, R510, R559, R665, 0S</td>
<td>6</td>
<td>30</td>
<td>0.276</td>
</tr>
<tr>
<td>5</td>
<td>R443/R411, R490/R411, R510/R411, R559/R411, R619/R411, R665/R411, R705/R411</td>
<td>6</td>
<td>40</td>
<td>0.302</td>
</tr>
<tr>
<td>6</td>
<td>R411/R443, R490/R443, R510/R443, R559/R443, R619/R443, R665/R443, R705/R443</td>
<td>12</td>
<td>40</td>
<td>0.290</td>
</tr>
<tr>
<td>7</td>
<td>R411/R490, R443/R490, R510/R490, R559/R490, R619/R490, R665/R490, R705/R490</td>
<td>6</td>
<td>40</td>
<td>0.272</td>
</tr>
<tr>
<td>8</td>
<td>R411/R559, R443/R559, R490/R559, R510/R559, R619/R559, R665/R559, R705/R559</td>
<td>12</td>
<td>40</td>
<td>0.286</td>
</tr>
<tr>
<td>9</td>
<td>R411/R665, R443/R665, R490/R665, R510/R665, R559/R665, R619/R665, R665/R665, R705/R665</td>
<td>6</td>
<td>40</td>
<td>0.285</td>
</tr>
<tr>
<td>10</td>
<td>R411/R443, R490/R443, R510/R443, R559/R443, R665/R443</td>
<td>30</td>
<td>40</td>
<td>0.325</td>
</tr>
<tr>
<td>11</td>
<td>R411/R559, R443/R559, R490/R559, R510/R559, R665/R559</td>
<td>30</td>
<td>40</td>
<td>0.327</td>
</tr>
<tr>
<td>12</td>
<td>R411/R665, R443/R665, R490/R665, R510/R665, R559/R665</td>
<td>6</td>
<td>40</td>
<td>0.319</td>
</tr>
<tr>
<td>13</td>
<td>R443/R559, R490/R559, R510/R559</td>
<td>6</td>
<td>40</td>
<td>0.396</td>
</tr>
</tbody>
</table>
Table 5.4. Performance with regard to SPM retrieval of spectral band combinations used input to ANNs

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Input Neurons</th>
<th>Neurons in Hidden layer</th>
<th>Noise level (%)</th>
<th>Training data (N=1000)</th>
<th>Validation data (N=218)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RMSE</td>
<td>r²</td>
</tr>
<tr>
<td>1</td>
<td>R411, R443, R490, R510, R559, R619, R665, R705</td>
<td>6 0</td>
<td>0.0576 0.994</td>
<td>0.201 0.772</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>R411, R443, R490, R510, R559, R619, R665, R705, 0S</td>
<td>6 0</td>
<td>0.0299 0.998</td>
<td>0.212 0.750</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>R411, R443, R490, R510, R559, R665</td>
<td>12 10</td>
<td>0.0726 0.991</td>
<td>0.215 0.756</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>R411, R443, R490, R510, R559, R665, 0S</td>
<td>6 0</td>
<td>0.0268 0.999</td>
<td>0.223 0.739</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>R443/R411, R490/R411, R510/R411, R559/R411, R619/R411, R665/R411, R705/R411</td>
<td>6 20</td>
<td>0.239 0.903</td>
<td>0.242 0.654</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>R411/R559, R443/R559, R490/R559, R510/R559, R619/R559, R665/R559, R705/R559</td>
<td>12 20</td>
<td>0.221 0.916</td>
<td>0.244 0.660</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>R411/R665, R443/R665, R490/R665, R510/R665, R559/R665, R619/R665, R665/R559, R705/R665</td>
<td>6 20</td>
<td>0.222 0.916</td>
<td>0.225 0.693</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>R443/R411, R490/R411, R510/R411, R559/R411, R665/R411,</td>
<td>12 20</td>
<td>0.247 0.895</td>
<td>0.232 0.657</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>R411/R559, R443/R559, R490/R559, R510/R510, R665/R559,</td>
<td>6 20</td>
<td>0.233 0.906</td>
<td>0.221 0.693</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>R411/R665, R443/R665, R490/R665, R510/R665, R559/R665,</td>
<td>6 20</td>
<td>0.228 0.911</td>
<td>0.228 0.672</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>R510/R665, R559/R665, R619/R665, R705/R665</td>
<td>12 20</td>
<td>0.228 0.911</td>
<td>0.242 0.675</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>R559/R510, R619/R510, R665/R510, R705/R510</td>
<td>12 20</td>
<td>0.230 0.909</td>
<td>0.233 0.666</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>R490/R619, R510/R619, R559/R619</td>
<td>12 20</td>
<td>0.245 0.897</td>
<td>0.244 0.630</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.5. Performance with regard to CDOM retrieval of spectral band combinations used input to ANNs

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Input</th>
<th>Neurons in Hidden layer</th>
<th>Noise level (%)</th>
<th>Training data (N=1000)</th>
<th>Validation data (N=214)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RMSE</td>
<td>$r^2$</td>
</tr>
<tr>
<td>1</td>
<td>R411, R443, R490, R510, R559, R619, R665, R705</td>
<td>20</td>
<td>40</td>
<td>0.153</td>
<td>0.941</td>
</tr>
<tr>
<td>2</td>
<td>R411, R443, R490, R510, R559, R619, R665, R705, 0S</td>
<td>30</td>
<td>40</td>
<td>0.154</td>
<td>0.940</td>
</tr>
<tr>
<td>3</td>
<td>R411, R443, R490, R510, R559, R665</td>
<td>30</td>
<td>40</td>
<td>0.160</td>
<td>0.936</td>
</tr>
<tr>
<td>4</td>
<td>R412, R443, R490, R510, R559, R665, 0S</td>
<td>30</td>
<td>40</td>
<td>0.163</td>
<td>0.933</td>
</tr>
<tr>
<td>5</td>
<td>R443/R411, R490/R411, R510/R411, R559/R411, R619/R411, R665/R411, R705/R411</td>
<td>20</td>
<td>40</td>
<td>0.227</td>
<td>0.870</td>
</tr>
<tr>
<td>6</td>
<td>R411/R490, R443/R490, R510/R490, R559/R490, R619/R490, R665/R490, R705/R490</td>
<td>30</td>
<td>40</td>
<td>0.220</td>
<td>0.878</td>
</tr>
<tr>
<td>7</td>
<td>R411/R559, R443/R559, R490/R559, R510/R559, R559/R559, R619/R559, R665/R559, R705/R559</td>
<td>12</td>
<td>40</td>
<td>0.206</td>
<td>0.893</td>
</tr>
<tr>
<td>8</td>
<td>R411/R665, R443/R665, R490/R665, R510/R665, R559/R665, R619/R665, R665/R665, R705/R665</td>
<td>12</td>
<td>40</td>
<td>0.218</td>
<td>0.880</td>
</tr>
<tr>
<td>9</td>
<td>R443/R411, R490/R411, R510/R411, R559/R411, R619/R411, 0S</td>
<td>12</td>
<td>40</td>
<td>0.206</td>
<td>0.892</td>
</tr>
<tr>
<td>10</td>
<td>R411/R559, R443/R559, R490/R559, R510/R559, R559/R559, R619/R559, R665/R559</td>
<td>6</td>
<td>40</td>
<td>0.217</td>
<td>0.881</td>
</tr>
<tr>
<td>11</td>
<td>R411/R665, R443/R665, R490/R665, R510/R665, R559/R665, R619/R665</td>
<td>20</td>
<td>40</td>
<td>0.222</td>
<td>0.875</td>
</tr>
<tr>
<td>12</td>
<td>R443/R411, R490/R411, R510/R411, R559/R411, R619/R411</td>
<td>30</td>
<td>40</td>
<td>0.208</td>
<td>0.891</td>
</tr>
<tr>
<td>13</td>
<td>R443/R411, R490/R411, R510/R411, R559/R411</td>
<td>6</td>
<td>40</td>
<td>0.225</td>
<td>0.872</td>
</tr>
</tbody>
</table>
(2). Regarding the retrieval of SPM concentrations, the performance of the ANNs using absolute reflectance values as input is significantly higher than that of ANNs using the reflectance ratios. The best results (lowest RMSE and highest $r^2$) were obtained using the eight reflectances as input (case No. 1 in Table 5.4). In this case, the $r^2$ value is 0.772, and the RMSE value is 0.201.

Besides, the performance of case No. 1 and case No. 2 using eight reflectances as input is slightly better than that of case No. 3 and case No. 4 using six reflectances as input. This means that providing additional information at wavelengths of 619 nm and 709 nm can improve the accuracy of the SPM concentration retrieval.

(3). Regarding the retrieval of CDOM absorption, there is no significant difference of the ANNs performance between using absolute reflectance values and the reflectance ratios as input. The lowest RMSE was obtained in case No. 13 (in Table 5.5) using the four reflectance ratios as input (RMSE=0.161, and $r^2=0.790$), while the highest $r^2$ was obtained in case No. 4 using six reflectances plus solar zenith angle as input (RMSE=0.189, and $r^2=0.835$). By a comprehensive consideration, case No. 4 was taken as the best case for the retrieval of CDOM absorption.

Besides, the performance of case No. 3 and case No. 4 using six reflectances as input is slightly better than that of case No. 1 and case No. 2 using eight reflectances as input. This means that providing additional information at wavelengths of 620 nm and 709 nm does not improve the accuracy of the CDOM absorption retrieval, but in contrast reduces it. The reason for this is that the reflectances at 620 nm and 709 nm are highly correlated with the reflectance at 665 nm, and the information of the reflectance at 665 nm is sufficient for the retrieval of CDOM. Using noisy information at these wavelengths (620 nm and 709 nm) will therefore increase the error of the retrieval CDOM absorption coefficient.

(4). It has been shown in Chapter 3 that the performance of pigment retrieval algorithms based on ratio input are much better than that of algorithms based on absolute reflectance input in Case I waters. The reason is that spectrally correlated noise is partly cancelled out through division of the reflectances at two wavelengths. From the results of this chapter, however, the algorithm best suited for the retrieval of pigment concentration uses the reflectance ratios as input, while the algorithms best suited for the retrieval of SPM and CDOM use the absolute reflectance values as input. The observed behaviour may be explained in the following way. The absorption spectra of CDOM as well as SPM are characterised as an exponential function. Thus, the contribution of CDOM absorption or SPM absorption to the reflectance for different channels acts as a similar way. Therefore, through the division of the reflectances at two wavelengths, on the one hand, the spectrally correlated noise can be partly cancelled out, on the other hand, some useful information for the retrieval of CDOM or SPM may also be partly removed. However, the absorption spectra of the pigment are
characterised by significant difference over the whole spectral domain. Therefore, the contribution of the pigment absorption to the reflectance may not be significantly reduced through division of the reflectances at two wavelengths, while only the spectrally correlated noise can be partly removed.

(5). For each of the retrieval of the three constituents, the noise levels added to training data set are significantly different for optimal retrieval of three oceanic constituents. To sum up, the optimal retrieval of SPM requires little noise added to training data set. However, the optimal retrieval of CHL and CDOM requires more noise added to the training data set. The reason behind the observed behaviour may be as follows. The concentration of SPM is much more sensitive to the variation of reflectance spectra than that of the other two constituents, because of its strong scattering. Therefore, the number of the training cases required for retrieval of SPM is relative small. While pigment concentration and absorption coefficient of CDOM are less sensitive to the variation of reflectance spectra, because of influence of the strong scattering of SPM, as well as their influence from each other. Therefore, the number of the training cases required for the retrieval of pigment and CDOM is relative large. In principle, a sufficiently large set of training cases is necessary to get a good generalisation. If there are no enough number of training cases, a good generalisation can also be obtained by adding an appropriate noise level to the training data set. This technique has been commonly used to reduce the effort of the creation of training data and ANN training process. Generally speaking, the less training cases are available, the more noise is needed. In this study, 1000 training cases for retrievals of all three constituents were used. Therefore, the noise level adding to training data set for the optimal retrieval of pigment and CDOM should be larger than that for the optimal retrieval of SPM.

5.4. Evaluating the Performance of the ANN-based Oceanic Constituents Retrieval Algorithms

5.4.1. Assessing the Performance of the Trained ANN

The potential of the selected optimal ANNs (listed in Table 5.6) for each of three constituents from real measurements is assessed in three steps by applying it a) to the synthetic data (“training data”) used for the ANN training, b) to “validation data” used to determine the ANN architecture and noise level to be added to the input data, and c) to “test data” which have not been used for the ANN development.

(1). Performance with respect to the training data

The ANN forecasts of each individual oceanic constituent concentration based on the simulated hemispherical reflectance are compared to the corresponding oceanic constituent
concentrations used as input for the RT simulations. As shown in Table 5.6, as well as in Figures 5.2 (A), 5.3 (A) and 5.4 (A), respectively, for CHL, SPM, and CDOM, the inversion is successful with regard to the synthetic training data set.

Table 5.6. Performance of the selected optimal ANNs for the retrievals of three oceanic constituents

<table>
<thead>
<tr>
<th>Constituent</th>
<th>Input Neurons</th>
<th>Noise level (%)</th>
<th>Training data N=1000</th>
<th>Validation data N=205</th>
<th>Test data N=163</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHL</td>
<td>R411/R443, R490/R443, R510/R443, R559/R443, R619/R443, R665/R443, R705/R443</td>
<td>12 40</td>
<td>0.290 0.889</td>
<td>0.274 0.729</td>
<td>0.339 0.860</td>
</tr>
<tr>
<td>SPM</td>
<td>R411, R443, R490, R510, R559, R619, R665, R705</td>
<td>6 0</td>
<td>0.0576 0.994</td>
<td>0.201 0.772</td>
<td>0.338 0.910</td>
</tr>
<tr>
<td>CDOM</td>
<td>R411, R443, R490, R510, R559, R665, 0S</td>
<td>30 40</td>
<td>0.163 0.933</td>
<td>0.189 0.835</td>
<td>0.279 0.769</td>
</tr>
</tbody>
</table>

(2). Performance with respect to the validation data

In a second step, the ANNs for the retrieval of the different constituent retrievals are applied to the ‘validation data’ consisting of in-situ measurements. The retrieval results for pigment, SPM and CDOM are also listed in Table 5.6. The results are depicted in Figures 5.2 (B), 5.3 (B) and 5.4 (B) for CHL, SPM and CDOM, respectively. The concentrations of the three constituents derived from the ‘validation data’ agree well with the corresponding in-situ measurements. This is not surprising, since the ‘validation data’ have a) been used to derive the back scattering model for marine particles used for the RT simulations (see Chapter 4), and b) been selected to identify the most appropriate ANN architecture and noise level.

(3). Performance with respect to the test data

In a third step, the ANNs for the retrieval of the different constituents are applied to the second set of in-situ measurement data set (‘test data’). The retrieval results for pigment, SPM and CDOM are also listed in Table 5.6. The results are depicted in Figure 5.2 (C), 5.3 (C) and 5.4 (C) for CHL, SPM and CDOM, respectively. Satisfactory performance is observed even though the ‘test data’ are totally independent from the ‘validation data’ and have not been used in any respect for the development of the ANN.
Figure 5.2. Scatter plot showing the performance of the ANN-based pigment retrieval algorithms for the synthetic training data set (A), validation data set (B), and test data (C). The dashed lines indicate the factor 2 error margin.
Figure 5.3. Scatter plot showing the performance of the ANN-based SPM retrieval algorithms for the synthetic training data set (A), validation data set (B), and test data (C). The dashed lines indicate the factor 2 error margin.
Figure 5.4. Scatter plot showing the performance of the ANN-based CDOM retrieval algorithms for the synthetic training data set (A), validation data set (B), and test data (C). The dashed lines indicate the factor 2 error margin.
5.4.2. Comparison with Existing Retrieval Algorithms

In order to further evaluate the performance of the trained ANNs, it was compared to the PMNS empirical algorithms listed in Table 5.7(a), (b) and (c). The PMNS algorithms [Shimwell et al., 1995] were developed for the retrieval of CHL, SPM and CDOM retrieval, and are based on data collected in the southern North Sea. For each of three constituents, there are two retrieval algorithms based on different band ratios. These algorithms were applied to the reflectance data from the ‘validation data’ and ‘test data’. The performance of the algorithms for the three oceanic constituents is also given in Table 5.7(a), (b) and (c). The ANN has the highest $r^2$ and lowest RMSE of all compared algorithms for the three constituents. The relative success of the ANN-based retrieval is partly explained by the fact that the underlying IOP models represent the in-situ measurement data well. Further reason is that it uses more spectral information than do the empirical algorithms PMNS.

Table 5.7 (a). Performance of the ANN-based pigment retrieval algorithms as compared to the PMNS algorithms

<table>
<thead>
<tr>
<th>Name of algorithm</th>
<th>Algorithm form</th>
<th>Validation data</th>
<th>Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>$r^2$</td>
</tr>
<tr>
<td>PMNS_1</td>
<td>3.4*[R(510)/R(560)]$^{-3.65}$</td>
<td>0.369</td>
<td>0.575</td>
</tr>
<tr>
<td>PMNS_2</td>
<td>22.3*[R(665)/R(705)]$^{-2.85}$</td>
<td>0.441</td>
<td>0.351</td>
</tr>
<tr>
<td>ANN</td>
<td>ANN</td>
<td>0.274</td>
<td>0.729</td>
</tr>
</tbody>
</table>

Table 5.7 (b). Performance of the ANN-based SPM retrieval algorithms as compared to the PMNS algorithms

<table>
<thead>
<tr>
<th>Name of algorithm</th>
<th>Algorithm form</th>
<th>Validation data</th>
<th>Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>$r^2$</td>
</tr>
<tr>
<td>PMNS_1</td>
<td>12.4*[R(412)/R(665)]$^{-1.0}$</td>
<td>0.741</td>
<td>0.588</td>
</tr>
<tr>
<td>PMNS_2</td>
<td>53.1*[R(560)/R(620)]$^{-2.58}$</td>
<td>0.517</td>
<td>0.681</td>
</tr>
<tr>
<td>ANN</td>
<td>ANN</td>
<td>0.201</td>
<td>0.772</td>
</tr>
</tbody>
</table>

Table 5.7 (c). Performance of the ANN-based CDOM retrieval algorithms as compared to the PMNS algorithms

<table>
<thead>
<tr>
<th>Name of algorithm</th>
<th>Algorithm form</th>
<th>Validation data</th>
<th>Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>$r^2$</td>
</tr>
<tr>
<td>PMNS_1</td>
<td>0.76*[R(490)/R(665)]$^{-0.83}$</td>
<td>0.396</td>
<td>0.694</td>
</tr>
<tr>
<td>PMNS_2</td>
<td>0.72*[R(520)/R(620)]$^{-1.05}$</td>
<td>0.407</td>
<td>0.714</td>
</tr>
<tr>
<td>ANN</td>
<td>ANN</td>
<td>0.163</td>
<td>0.835</td>
</tr>
</tbody>
</table>
5.5. Conclusions

In this study, a methodology for the retrieval of three constituents from ocean colour in Case II waters have been derived. The retrieval method is derived by applying ANN techniques to a set of hemispherical reflectance spectra typical of Case II waters, which have been obtained from RT simulations.

Three ANN-based algorithms were obtained in this study for the retrievals of CHL, SPM and CDOM, respectively. Each individual ANN has three layers: one input layer, one hidden layer and one output layer. A bias parameter is added both to the input layer and to the hidden layer. The output layer consists of one neuron which corresponds to one of the three constituent concentrations. The number of neurons in the input layer and in the hidden layer which was determined in terms of the optimal retrieval of constituents is different for these algorithms. For the optimal retrieval of pigment, the ANN has seven neurons in the input layer corresponding to the seven reflectance ratios, and 12 hidden neurons. For the optimal retrieval of SPM, the ANN has eight neurons in the input layer corresponding to the eight reflectances, and 6 hidden neurons. For the optimal retrieval of CDOM, the ANN has seven neurons corresponding to six reflectances and solar zenith angle, and 30 hidden neurons.

Applying the three trained ANNs either to the validation data, or to the test data which have not been used to derive the back scattering probability model of marine particles used for the RT simulations, the results for the retrieval of all three constituents are satisfactory. For example, for retrieval of pigment concentration, applying the algorithm to the validation data gives a correlation between predicted and measured pigment concentrations of $r^2 = 0.729$ and RMSE = 0.274; applying it to the test data, results in $r^2 = 0.860$ and RMSE = 0.339. For the retrieval of SPM concentration, applying the algorithm to the validation data gives a correlation between predicted and measured pigment concentrations of $r^2 = 0.772$ and RMSE = 0.201; applying it to the test data, results in $r^2 = 0.910$ and RMSE = 0.338. For the retrieval of CDOM, applying the algorithm to the validation data gives a correlation between predicted and measured pigment concentrations of $r^2 = 0.835$ and RMSE = 0.189; applying it to the test data, results in $r^2 = 0.769$ and RMSE = 0.279.

The performance of the ANN-based retrieval scheme is generally better than that of the empirical algorithms PMNS. For example, for the retrieval of pigment concentration, applying the PMNS algorithm PMNS_1 to the test data set gives $r^2 = 0.753$, and RMSE = 0.505 as compared to $r^2 = 0.860$ and RMSE = 0.263 for the ANN-based algorithm. For the retrieval of SPM concentration, applying the PMNS algorithm PMNS_1 to the test data set gives $r^2 = 0.855$, and RMSE = 0.390 as compared to $r^2 = 0.910$ and RMSE = 0.338 for the ANN-based algorithm. For the retrieval of CDOM, applying the PMNS algorithm PMNS_1 to the test data set gives $r^2 = 0.764$, and RMSE = 0.450 as compared to $r^2 = 0.769$ and RMSE = 0.262 for the ANN-based algorithm.
6.1. Background

As described in Chapter 1, there are two challenges for atmospheric correction which need to be resolved. One concerns the atmospheric correction over turbid waters, the other concerns the atmospheric correction for atmospheres with strong-absorbing aerosols over both clear and turbid waters.

In Chapters 3 and 5, the oceanic constituents concentrations (CHL, SPM and CDOM) are successfully derived from ocean colour measurements at the sea level with ANN based on RT simulations. In this chapter, the same methodology is used to derive the oceanic constituents concentrations from ocean colour measurements at the top of atmosphere.

In this study, the light field in the atmosphere-ocean system at top of atmosphere is simulated by the radiative transfer code MOMO. The inherent optical properties of the oceanic constituents are the same as used in Chapter 5. In the atmosphere, various aerosols are considered, including maritime, continental, soot, dust as well as H$_2$SO$_4$. The simulated light field at the top of atmosphere is taken as the training data set for the ANN used for the retrieval of oceanic constituents. Input to the ANN is the remote sensing reflectance and other auxiliary parameters.

6.2. Simulated Data Sets

In the frame of this thesis, simulated data sets are used to derive the algorithms for the retrieval of oceanic constituent concentrations from reflectances at top of atmosphere. The inputs of radiative transfer simulations: IOPs and vertical distributions of the constituents in ocean and atmosphere, are described in the following subsections.

6.2.1. Atmosphere

The following atmospheric constituents are considered: air (molecular scattering) and aerosols (scattering and absorption). Other atmospheric constituents (clouds, absorbing gases, rain) are neglected.
6.2.1.1. Rayleigh Scattering

The vertical profile of molecular scattering is taken from Elterman (1968). The total optical thickness at a specific pressure \( p \) is computed from the approximation given by Hansen and Travis (1974):

\[
\tau_r = \frac{P}{P_0} (8.524 \times 10^{-3} \lambda^{-4} + 9.63 \times 10^{-5} \lambda^{-6} + 1.1 \times 10^{-7} \lambda^{-8}),
\]

where \( P_0 = 1013.25 \) hPa

6.2.1.2. Aerosols

The basic constituents of aerosols are the following:
(a). oceanic particles (sea salt solution in water),
(b). water soluble particles,
(c). dust-like particles,
(d). Asian dust particles,
(e). soot particles,
(f). sulphuric acid solution in water.

Five aerosol models have been defined (listed in Table 6.1). The corresponding size distributions are listed in Table 6.2 [Shettle and Fenn, 1979; WCRP, 1986; Fukushima and Toratani, 1997].

<table>
<thead>
<tr>
<th>Aerosol model</th>
<th>Constituent</th>
<th>Proportion (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maritime</td>
<td>Rural aerosol mixtures (70% water soluble, and 30% dust-like particles) Oceanic</td>
<td>99</td>
<td>Shettle and Fenn, 1979</td>
</tr>
<tr>
<td>Continental</td>
<td>Water soluble Dust-like Soot</td>
<td>93.876 2.27*10^-4 6.123</td>
<td>WCRP, 1986</td>
</tr>
<tr>
<td>Dust</td>
<td>Asian dust</td>
<td>100</td>
<td>Fukushima and Toratani, 1997</td>
</tr>
<tr>
<td>Soot</td>
<td>Soot</td>
<td>100</td>
<td>WCRP, 1986</td>
</tr>
<tr>
<td>H2SO4</td>
<td>75% solution of sulphuric acid in water</td>
<td>100</td>
<td>WCRP, 1986</td>
</tr>
</tbody>
</table>

The log-normal distribution is expressed as:
The modified Gamma function is expressed as:

\[
\frac{dN(r)}{d(\log(r))} = \frac{\log^2(r/r_b)}{2[\log(\delta)]^2} e^{-\frac{\log^2(r/r_b)}{2[\log(\delta)]^2}}
\]  

(6.2)

The modified Gamma function is expressed as:

\[
\frac{dN(r)}{dr} = a r^\alpha e^{-b r^\gamma}
\]  

(6.3)

Here \(N(r)\) is the number of particles having radii smaller than \(r\), \(r_b\) represents the mode radius, and \(\delta\) represents the standard deviation.

Based on the above parameters, phase function (Figure 6.1), extinction coefficient (Figure 6.2), and single scattering albedo (Figure 6.3) are computed from Mie theory.

The vertical distribution of the aerosol is assumed as follows:

1. Boundary layer (0 ~ 2 km): maritime, soot,
2. Tropospheric layer (2 ~ 12 km): dust, continental
3. Stratospheric layer (12 ~ 50 km): H\(_2\)SO\(_4\).

Table 6.2  Size distributions of various aerosols used in this study

<table>
<thead>
<tr>
<th>Type</th>
<th>Component</th>
<th>Size distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log-normal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(r_0)</td>
<td>(\sigma)</td>
</tr>
<tr>
<td>Maritime_1</td>
<td>Rural_RH70</td>
<td>0.02846</td>
</tr>
<tr>
<td>Maritime_1</td>
<td>Oceanic_RH70</td>
<td>0.2041</td>
</tr>
<tr>
<td>Maritime_2</td>
<td>Rural_RH80</td>
<td>0.03274</td>
</tr>
<tr>
<td>Maritime_2</td>
<td>Oceanic_RH80</td>
<td>0.3180</td>
</tr>
<tr>
<td>Maritime_3</td>
<td>Rural_RH95</td>
<td>0.03884</td>
</tr>
<tr>
<td>Maritime_3</td>
<td>Oceanic_RH95</td>
<td>0.3803</td>
</tr>
<tr>
<td>Maritime_4</td>
<td>Rural_RH99</td>
<td>0.05215</td>
</tr>
<tr>
<td>Maritime_4</td>
<td>Oceanic_RH99</td>
<td>0.7505</td>
</tr>
<tr>
<td>Continental</td>
<td>Water soluble</td>
<td>0.005</td>
</tr>
<tr>
<td>Continental</td>
<td>Dust-like</td>
<td>0.5</td>
</tr>
<tr>
<td>Continental</td>
<td>Soot</td>
<td>0.0118</td>
</tr>
<tr>
<td>Dust</td>
<td>Asian dust</td>
<td>0.40</td>
</tr>
<tr>
<td>Soot</td>
<td>Soot</td>
<td>0.0118</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>gamma modified</th>
<th>(a)</th>
<th>(\alpha)</th>
<th>(\gamma)</th>
<th>(b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H(_2)SO(_4)</td>
<td>324</td>
<td>1</td>
<td>1</td>
<td>18</td>
</tr>
</tbody>
</table>
6.2.2. Water

The considered water constituents and their IOPs models are the same as those described in Chapter 4. The following assumptions for the radiative transfer simulations in ocean are made:

- no vertical stratification of the ocean,
- the ocean is of infinite depth,
- no inelastic scattering inside the water body (i.e. no Raman scattering, no chlorophyll-a fluorescence, no CDOM fluorescence).

![Figure 6.1. Phase functions of the aerosol models used in this study](image1)

![Figure 6.2. Normalised attenuation coefficients of the aerosol models used in this study](image2)
6.2.3. Sea Surface State

The rough sea surface is considered. The effects of the air-sea interface on the light field are introduced in the radiative transfer simulations by applying the statistic model of the wave surface distribution by Cox & Munk (1954), assuming a gaussian distribution of waves and a variable wind speed. The wind direction is not considered.

6.2.4. Ranges of the Related Parameters

Ranges of the oceanic constituent concentrations, optical thickness of various aerosols models at 550 nm, and other parameters used for the radiative transfer calculations are shown in Tables 6.3, 6.4 and 6.5, respectively.

![Figure 6.3. Single scattering albedo of the aerosol models used in this study](image)

The wavelengths used in this study (Table 6.6) are the same as MERIS’s wavelengths, but the following wavelengths are not included: 681.25 nm (chlorophyll fluorescence peak), 753.75 nm (cloud), 760.6 nm (O₂ absorption), 865 nm (water vapour reference), and 900 nm (water vapour absorption).

<table>
<thead>
<tr>
<th>Table 6.3 Ranges of the oceanic constituents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Pigment</td>
</tr>
<tr>
<td>SPM</td>
</tr>
<tr>
<td>aₜ(443)</td>
</tr>
</tbody>
</table>
Table 6.4 Ranges of the atmospheric components

<table>
<thead>
<tr>
<th>Aerosol model</th>
<th>Variable</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maritime</td>
<td>$\tau_a(550)$</td>
<td>0.01</td>
<td>0.5</td>
</tr>
<tr>
<td>Dust</td>
<td>$\tau_a(550)$</td>
<td>0.01</td>
<td>0.5</td>
</tr>
<tr>
<td>Soot</td>
<td>$\tau_a(550)$</td>
<td>0.001</td>
<td>0.1</td>
</tr>
<tr>
<td>Continental</td>
<td>$\tau_a(550)$</td>
<td>0.01</td>
<td>0.5</td>
</tr>
<tr>
<td>H2SO4</td>
<td>$\tau_a(550)$</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>Total aerosol</td>
<td>$\tau_a(550)$</td>
<td>0.065</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 6.5. Ranges of geometric and other parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solar zenith angle ($\theta_s$)</td>
<td>degree</td>
<td>0.08</td>
<td>75.7</td>
</tr>
<tr>
<td>Obv. zenith angle ($\theta_v$)</td>
<td>degree</td>
<td>0</td>
<td>41</td>
</tr>
<tr>
<td>Obv. azim. angle ($\phi_v$)</td>
<td>degree</td>
<td>0</td>
<td>180</td>
</tr>
<tr>
<td>Surface air pressure (p)</td>
<td>hPa</td>
<td>1000</td>
<td>1040</td>
</tr>
<tr>
<td>Wind speed (w)</td>
<td>m/s</td>
<td>1</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 6.6 Spectral channels and their bandwidths used in this study

<table>
<thead>
<tr>
<th>MERIS Chan. No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>12</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center $\lambda$</td>
<td>412.5</td>
<td>442.5</td>
<td>490</td>
<td>510</td>
<td>560</td>
<td>620</td>
<td>665</td>
<td>709</td>
<td>779</td>
<td>885</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>10</td>
</tr>
</tbody>
</table>

6.2.5. RT Simulations

The concentrations of the oceanic constituents for the RT simulations were selected according to the procedures described in Chapter 5. The optical thickness of the aerosols, the pressure at sea level, as well as the wind speed were selected randomly within the ranges defined in Tables 6.4 and 6.5. The relative humidity of the maritime aerosol was randomly distributed among the four values 70%, 80%, 95% and 99%.

Simulations of the remote sensing reflectance at the top of atmosphere were made for:
- 14 solar zenith angles between $0.08^\circ$ and $75.6^\circ$,
- 9 observation zenith angles between $0^\circ$ and $41.4^\circ$,
- 25 observation azimuth angles between $0^\circ$ and $180^\circ$. 

75
• 10 wavelengths: 412.5, 442.5, 490, 510, 560, 620, 665, 709, 779 and 885 nm,
• 2000 combinations of concentrations of three oceanic constituents, optical thickness of four aerosol models, pressures, as well as wind speed.

6.2.6. Creation of ANN Training Data Set

Following the simulation strategy described in section 6.2.5, for each of 2000 combinations, there are 14×9×25=3150 spectra which correspond to different observation geometries. Therefore, we have 2000×3150=6300000 spectra which can be used for ANN training. In practice, this number is too large to train ANNs due to limitations in memory and computing power. In this study, 20 of the 3150 spectra for each of the 2000 combinations were randomly taken as the training data.

Through adding noise to synthetic training data, the robustness of the trained ANN with respect to noisy input data is increased. In the present study, measurement errors exist more or less in all input parameters, e.g., radiometric calibration errors. Thus, it is necessary to add the appropriate noise to the data set used for the training of the ANNs to resist these measurement errors. The noise levels for different input parameters used in this study are listed in Table 6.7.

Table 6.7 The noise level added to training data inputs. The following acronyms are used: R1 to R10 represent the remote sensing reflectances at 10 wavelengths at TOA, ‘W’ represents the wind speed, ‘a’ represents the angles defined in Table 6.5, ‘P’ represents the pressure at sea level.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R6</th>
<th>R7</th>
<th>R8</th>
<th>R9</th>
<th>R10</th>
<th>W</th>
<th>a, P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise (%)</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>1.1</td>
<td>1.6</td>
<td>10</td>
<td>0.5</td>
</tr>
</tbody>
</table>

In order to reduce the complexity, a Rayleigh scattering correction was applied to the reflectance at the top of atmosphere. Here, no interaction between aerosol scattering and Rayleigh scattering was assumed. The algorithm for the Rayleigh scattering correction uses an ANN trained with radiative transfer calculations. The algorithm was developed according to the following procedures:

1. Assuming black ocean and a rough sea surface for a pure Rayleigh atmosphere, 200 simulations corresponding to 200 different air pressures within the range 900 hPa to 1040 hPa were made.
2. 150 of the 200 simulations were used to train a MLP ANN, the other 50 simulations were taken as validation data. The ANN consists of three layers: input layer, one hidden layer and output layer. The input layer consists of 6 neurons which correspond to the following parameters: pressure at the sea level, wind speed, as well as four geometric parameters: μ, x,
y, z (see below). The hidden layer consists of 30 neurons. The output layer consists of 10 neurons which correspond to remote sensing reflectances at TOA for 10 wavelengths.

(3). The ANN forecasts of remote sensing reflectances due to the contribution of Rayleigh scattering at top of atmosphere were compared to the corresponding reflectances resulting directly from RT simulations. The comparison results were shown in Figure 6.4 for training data set and validation data set. As shown in Figure 6.4, the remote sensing reflectances from the contributions of the Rayleigh scattering at TOA can be successful derived by the trained ANN.

6.3. Artificial Neural Networks

6.3.1. Structure of the ANN for Retrieval of the Constituents

In the present study, Multi-Layer Perceptrons (MLPs) ANN are used for the retrieval of the oceanic constituents from TOA data. The architecture of MLPs has been described in Section 2.4.

The retrieval of the oceanic constituents is based on the information contained in 15 parameters: Rayleigh-corrected remote sensing reflectances in 10 spectral channels at TOA: \( R_{rs}(412) \), \( R_{rs}(443) \), \( R_{sp}(490) \), \( R_{rs}(510) \), \( R_{rs}(560) \), \( R_{rs}(620) \), \( R_{rs}(665) \), \( R_{rs}(709) \), \( R_{rs}(779) \), \( R_{rs}(885) \), wind speed, as well as the geometric parameters \( \mu, x, y, \) and \( z \), defined by Schröder (2003, personal communication):

\[
\mu=\cos(\theta_v) \\
x=\sin(\theta_v)\cos(\phi) \\
y=\sin(\theta_v)\sin(\phi) \\
z=\cos(\theta_v)
\]

A principal component analysis (PCA) was used to decorrelate all input parameters. As a result, the dimensionality of inputs to the ANN is reduced.

The output layer consists of 4 neurons which correspond to three oceanic constituents concentrations (CHL, SPM and CDOM) and total aerosol optical thickness at 550 nm. The hidden layer consists of 50 neurons.

6.3.2. Performance of the ANN-based Algorithm with Simulated Data

6.3.2.1. Performance with respect to the training data

40000 samples which have been generated according to the procedure described in Section 6.2.7 were used to train the ANN. The ANN-derived parameters were compared to the corresponding parameters used as input for the radiative transfer simulations (Figures 6.5, 6.6, 6.7 and 6.8 top of left panel for CHL, SPM, CDOM and \( \tau_a(550) \), respectively). The calculated RMSE
and Pearson's correlation coefficient are listed in Table 6.8. As shown in these figures and Table 6.8, the inversion for the four parameters is successful with regard to the simulated training data.

![Scatter plot showing the performance of the ANN-based algorithm for the Rayleigh scattering correction.](image)

**Figure 6.4.** Scatter plot showing the performance of the ANN-based algorithm for the Rayleigh scattering correction. (A), (C) and (E) for the training data set at three wavelengths: 412 nm, 560 nm and 885 nm, respectively. (B), (D) and (F) for test data set at the same wavelengths.
Table 6.8. Performance of the ANN-based algorithm for the retrieval of the oceanic constituents as well as the optical thickness of aerosols

| Data set | CHL  |  | SPM  |  | CDOM |  | Aerosol |  |
|----------|------|  |      |  |      |  |         |  |
|          | $r^2$ | RMSE (%) | $r^2$ | RMSE (%) | $r^2$ | RMSE (%) | $r^2$ | RMSE (%) |
| Train    | 0.926 | 73.1 | 0.956 | 48.6 | 0.895 | 65.4 | 0.961 | 13.1 |
| Test1    | 0.929 | 87.5 | 0.955 | 48.2 | 0.897 | 64.9 | 0.962 | 12.8 |
| Test2    | 0.936 | 61.2 | 0.963 | 41.7 | 0.916 | 84.7 | 0.933 | 16.8 |
| Test3    | 0.931 | 77.2 | 0.963 | 42.7 | 0.915 | 72.0 | 0.961 | 13.4 |
| Test4    | 0.928 | 73.9 | 0.965 | 44.1 | 0.907 | 64.9 | 0.940 | 16.5 |
| Test5    | 0.929 | 63.9 | 0.967 | 40.2 | 0.906 | 70.8 | 0.883 | 25.0 |
| Test6    | 0.924 | 73.5 | 0.966 | 42.1 | 0.910 | 73.2 | 0.961 | 15.0 |
| Test7    | 0.935 | 54.8 | 0.965 | 39.6 | 0.909 | 56.6 | 0.855 | 48.4 |

6.3.2.2 Performance with respect to the test data

In order to examine the performance of ANN-based algorithms for various atmospheric studies which may occur in the real environment, seven test data sets which correspond to different aerosol mixtures in the atmosphere were simulated (listed in Table 6.9). The inherent optical properties of the oceanic constituents used for simulations of test data sets are the same as used for the simulations of the training data set. Each test data set consists of 500 combinations of concentrations of three oceanic constituents, optical thickness of aerosols, air pressures, as well as wind speeds. The trained ANN was applied to these seven test data sets. The results are also, respectively, shown in Figures 6.5, 6.6, 6.7 and 6.8 for CHL, SPM, CDOM and $\tau_a(550)$, as well as listed in Table 6.8. Based on these results, the following conclusions can be drawn:

1. With the trained ANN, the three oceanic constituents concentrations can be successfully derived from simulated reflectances at TOA for atmospheres ranging from weakly to strongly absorbing aerosols.

2. With the trained ANN, the total optical thickness can be successfully derived for most cases except for TEST5 and TEST7. For both TEST5 and TEST7, the derived $\tau_a(550)$ is underestimated. TEST5 represents the most strongly absorbing aerosols, and TEST7 (Urban aerosol) is not included in the simulations of training data set. Although the retrieval of the total optical thickness for these two cases has a large error, it does not significantly impact the retrieval of oceanic constituents concentrations.
Figure 6.5. Scatter plot showing the performance of the ANN-based algorithm for pigment retrieval: training data set (top left panel, TRAIN), test data sets for seven different atmospheres (TEST1~TEST7). The target pigment concentration designates the pigment concentration used as input to the RT simulations. The dashed lines indicate the factor 3 error margin.
Figure 6.6. Scatter plot showing the performance of the ANN-based algorithm for SPM retrieval: training data set (top left panel, TRAIN), test data sets for seven different atmospheres (TEST1~TEST7). The target SPM concentration designates the SPM concentration used as input to the RT simulations. The dashed lines indicate the factor 3 error margin.
Figure 6.7. Scatter plot showing the performance of the ANN-based algorithm for CDOM retrieval: training data set (top left panel, TRAIN), test data sets for seven different atmospheres (TEST1~TEST7). The target CDOM absorption coefficient designates the CDOM absorption coefficients used as input to the RT simulations. The dashed lines indicate the factor 3 error margin.
Figure 6.8. Scatter plot showing the performance of the ANN-based algorithm for aerosol optical thickness retrieval: training data set (top left panel, TRAIN), test data sets for seven different atmospheres (TEST1~TEST7). The target aerosol optical thickness designates the aerosol optical thickness used as input to the RT simulations. The dashed lines indicate the factor 2 error margin.
### Table 6.9. Aerosol mixtures for testing the performance of the trained ANN

<table>
<thead>
<tr>
<th>Test Data</th>
<th>Aerosol</th>
<th>Vertical Profile</th>
<th>Test Data</th>
<th>Aerosol</th>
<th>Vertical Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test1</td>
<td>Maritime</td>
<td>0~2 km</td>
<td>Test5</td>
<td>Maritime</td>
<td>0~2 km</td>
</tr>
<tr>
<td></td>
<td>Soot</td>
<td>0~2 km</td>
<td></td>
<td>Soot</td>
<td>0~2 km</td>
</tr>
<tr>
<td></td>
<td>Continental</td>
<td>2~12 km</td>
<td></td>
<td>H2SO4</td>
<td>12~50 km</td>
</tr>
<tr>
<td></td>
<td>Asian dust</td>
<td>2~12 km</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H2SO4</td>
<td>12~50 km</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test2</td>
<td>Maritime</td>
<td>0~2 km</td>
<td>Test6</td>
<td>Maritime</td>
<td>0~2 km</td>
</tr>
<tr>
<td></td>
<td>H2SO4</td>
<td>12~50 km</td>
<td></td>
<td>Asian dust</td>
<td>2~12 km</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12~50 km</td>
</tr>
<tr>
<td>Test3</td>
<td>Maritime</td>
<td>0~2 km</td>
<td>Test7</td>
<td>Urban</td>
<td>0~2 km</td>
</tr>
<tr>
<td></td>
<td>Continental</td>
<td>2~12 km</td>
<td></td>
<td>H2SO4</td>
<td>12~50 km</td>
</tr>
<tr>
<td></td>
<td>H2SO4</td>
<td>12~50 km</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test4</td>
<td>Maritime</td>
<td>0~2 km</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soot</td>
<td>0~2 km</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian dust</td>
<td>2~12 km</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H2SO4</td>
<td>12~50 km</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 6.4. Application to the MERIS Imagery

MERIS (Medium Resolution Imaging Spectrometer) has been launched on board the earth observation satellite ENVISAT on March 1, 2002. It has the following main specifications: spatial resolution of 300 m in full resolution mode, revisit period of 2-3 days, 15 programmable spectral bands with high radiometric sensitivity. These specifications are very suitable for monitoring water properties in coastal waters.

Two MERIS images were selected, which are located in the North Sea and the China Seas. Both the North Sea and the China Seas belong to Case II waters. The North Sea is greatly influenced by the Elbe, Weser, Ems, Rhine, and the Thames rivers. As a result, the concentrations of both SPM and CDOM are very high. Besides, the phytoplankton concentration is usually high due to a steady nutrient supply from the rivers and the atmosphere. The China Seas are greatly influenced by the discharge of Yellow River and Yangzi River. These two rivers are famous for their high SPM concentration: SPM values can exceed 100 mg/m³ at the estuary of the two rivers [Hu, 2000; Jiang et al., 2002; Wu et al., 2003]. Besides, aerosols over the China Seas are characterised by their high variability in type and concentrations because of the presence of Asian dust particles as well as soot which originates from the main land of China [Fukushima and Toratani, 1997; Li et al., 2002].
The ANN-based algorithm described in Section 6.3 was applied to the two selected images (Figures 6.9 and 6.15). The pigment concentration, SPM concentration, CDOM absorption coefficient at 443 nm, and total optical thickness of aerosols at 550 nm have been calculated. The results for the North Sea are shown in Figures 6.10, 6.11, 6.12 and 6.13 for CHL, SPM, CDOM and $\tau_a(550)$, respectively. The distribution of the derived parameters in the North Sea along a transect is depicted in Figure 6.14. Corresponding results for the China Seas are shown in Figures 6.16, 6.17, 6.18 and 6.19. The distribution of the derived parameters in the China Seas along a transect is depicted in Figure 6.20.

Since simultaneous in-situ measurements in these areas for the two selected images are not available, the derived results from the algorithm developed in this study can not directly be validated. Here, the following criteria to assess whether the derived results are reasonable.

1. A low covariance between the distributions of SPM and aerosols.

By visual inspection, it becomes clear that the structures of the SPM distribution (Figure 6.11 and 6.17 for the North Sea and the China Seas, respectively) are different from the structures of the aerosol distribution (Figure 6.13 for the North Sea, Figure 6.19 for the China Seas), except for the area of Subei shoal and the mouth of Yellow River in the China Seas, where the concentrations of SPM are very high. This phenomenon can also clearly be seen from the distributions of SPM and aerosol along the same transect (Figure 6.14 for the North Sea, Figure 6.20 for the China Seas).

2. The agreement of the distributions of the derived parameters with known distributions.

In the southern North Sea, the structure of the SPM distribution is very similar to that derived from CZCS imagery [Doeffer and Fischer, 1994]. A water band with high SPM concentration stretches from the Thames estuary at the southeast edge of Great Britain into the North Sea. Besides, CDOM shows the expected distribution in the mouth of the Elbe River, where a high CDOM band was found.

In the China Seas, SPM distribution shows the expected pattern in the mouth of Yellow River, in the Subei shoal, in the centre of Yellow Sea, as well as in the Bohai Strait. In the mouth of Yellow River, SPM concentration is very high due to the discharge of inorganic particles. In the Subei shoal, SPM concentration is also very high due to the re-suspended sediments. The centre of Yellow Sea is very clear and SPM concentrations are frequently found to be less than 0.5 g/m$^3$ [Ahn et al., 2001]. In the strait of the Bohai Sea, SPM concentration is relatively low due to the influence of Yellow Sea, from which a band of low SPM concentration stretches into the Bohai Sea [Jiang et al., 2002]. CDOM distributions in the Bohai Sea and around the mouth of Yangzhi River are also within the expected pattern. In the Bohai Sea, CDOM is always high due to the discharge of Yellow River and waste water from factories. Around the mouth of Yangzhi River, input of riverine waters causes high CDOM concentrations. The CHL concentrations in the Bohai Sea and around the outer path of Yangzhi River plume are typically
Figure 6.9. RGB image of MERIS above the North Sea (22 March, 2003). The black line indicates the transect used for further analysis.

Figure 6.10. Pigment distribution derived from MERIS imagery for the North Sea (22 March, 2003)
Figure 6.11. SPM distribution derived from MERIS imagery for the North Sea (22 March, 2003)

Figure 6.12. CDOM distribution derived from MERIS imagery for the North Sea (22 March, 2003)
Figure 6.13. Aerosol distribution derived from MERIS imagery for the North Sea (22 March, 2003)

Figure 6.14. Distributions of pigment, SPM, CDOM and aerosol derived from MERIS imagery for the North Sea along the specific transect (22 March, 2003)
Figure 6.15. RGB image of MERIS above the China Seas (16 April, 2003). The black line indicates the transect used for further analysis

Figure 6.16. Pigment distribution derived from MERIS imagery for the China Seas (16 April, 2003)
Figure 6.17. SPM distribution derived from MERIS imagery for the China Seas (16 April, 2003)

Figure 6.18. CDOM distribution derived from MERIS imagery for the China Seas (16 April, 2003)
Figure 6.19. Aerosol distribution derived from MERIS imagery for the China Seas (16 April, 2003).

Figure 6.20. Distributions of pigment, SPM, CDOM and aerosol derived from MERIS imagery for the China Seas along the specific transect (16 April, 2003).
higher than that in Yellow Sea. This distribution is also found in Figure 6.16. The range of derived SPM concentration is within the range known for the China Seas. The range of derived CHL concentration is also within the known range, except for the highly turbid waters. In the highly turbid waters, the derived CHL concentration is unreasonably high: around the Subei shoal and the mouth of Yangzi River, the derived CHL concentration exceeds 30 mg m\(^{-3}\). In fact, although there is a rich nutrient supply in these areas, the availability of light is reduced due to the high SPM concentration. As a result, photosynthesis is reduced and the CHL concentration should be comparably low. The observed values in these areas are between 0.1 ~ 10 mg m\(^{-3}\) [Hu, 2000].

Judging from the above two criteria, the derived results seem to be reasonable except for the CHL retrieval in highly turbid waters of the China Seas. A possible reason for this is that the IOP models and the constraint conditions for the ranges of CHL, SPM and CDOM [Eq.(5.1) ~ (5.6)], used for development of the algorithm, are optimised for European waters. They may not be valid for some areas of the China Seas.

6.5. Conclusion

In this study, a methodology for the retrieval of three constituents from ocean colour at TOA above Case II waters have been derived. The retrieval method is derived by applying ANN techniques to a set of remote sensing reflectance spectra at TOA typical of Case II waters, which have been obtained from RT simulations.

The ANN employed in this study is a MLP with three layers: one input, one hidden and one output layer. A bias parameter is added both to the input layer and to the hidden layer. The input layer consists of 15 neurons which correspond to Rayleigh-corrected remote sensing reflectances of 10 wavelengths at TOA and other 5 auxiliary parameters. The hidden layer consists of 50 neurons. The output layer consists of four neurons which correspond to the concentrations of CHL, SPM and CDOM, as well as the total optical thickness of aerosols.

The ANN-based algorithm developed in this study seems to be a promising technique for the retrieval of oceanic constituents in Case II waters from ocean colour measurements at TOA. The examination of the performance of the ANN-based algorithm shows that it can deal with various atmospheres from weakly to strongly absorbing aerosols. Applied to the MERIS imagery of the North Sea and the China Seas, the results appears reasonable except for the highly turbid areas in the China Seas. Up to now, the algorithm has been tested for only a few images, and the validation is preliminary. In future, a more thorough validation of this algorithm using simultaneous in-situ measurements in different waters should be carried out. Besides, in order to optimise this method for the China Seas, precise information on the optical properties of oceanic constituents in the China Seas are required.
Chapter 7

SUMMARY

7.1. Conclusion

In this thesis, a methodology for the retrieval of oceanic constituents from ocean colour in Case I and Case II waters has been derived. The retrieval method is derived by applying ANN techniques to a set of reflectance spectra typical of Case I and Case II waters, which have been obtained from RT simulations. The results show that this method is a promising technique for the retrieval of oceanic constituents in Case I or Case II waters from ocean colour measurements at sea level or at TOA. The conclusions may be summarised as follows:

1. The method is successfully applied to derive the pigment concentration in Case I waters from remote sensing reflectances just above the sea surface. It compares favourably with commonly used empirical pigment retrieval schemes.

2. A bio-optical model of the back scattering probability of marine particles in Case II waters is proposed in order to simulate the light field in Case II waters with sufficient accuracy for the development of retrieval schemes for oceanic constituents.

3. The method is successfully applied to derive oceanic constituents concentrations (CHL, SPM, and CDOM) in European Case II waters from hemispherical reflectances just below the sea surface.

4. When the method is applied to derive oceanic constituents concentrations (CHL, SPM, and CDOM) from the MERIS imagery at TOA above the North sea and the China Seas, plausible results are obtained. However, more work is required to validate the method for this specific application.

7.2. Future Work

1. With more accurate knowledge about the IOPs of oceanic constituents, the IOP models of oceanic constituents used in RT simulations of this thesis should be updated to increase the retrieval accuracy, especially for Case II waters.

2. More information on the IOPs in the China Seas should be obtained to solve the problems which have been met when applying the algorithm to the MERIS imagery above the China Seas.

3. The algorithm developed in Chapter 5 should be applied to more areas (outside the European waters) to check its performance for a global scale.

4. The algorithm developed in Chapter 6 for MERIS imagery will be widely validated in different areas with the simultaneous in situ measurements.
(5). Since the variation of the IOPs in the different areas in Case II waters are large, specific algorithms for individual areas should be developed to increase the retrieval accuracy.
REFERENCES


Babin, M., Coastal surveillance through observation of ocean colour (COASTLOOC), *Final Report, Project ENV4-CT96-0310*, 233 pp., Laboratoire de Physique et Chimie Marines, Villefranche-sur-mer, France, 2000.


