Retrieval of the diffuse attenuation coefficient $K_d(\lambda)$ in open and coastal waters using a neural network inversion

Jamet, C., H., Loisel and D., Dessailly
Laboratoire d'Océanologie et de Géosciences
32 avenue Foch, 62930 Wimereux, France
cedric.jamet@univ-littoral.fr

3rd CoastColour User Consultation Meeting
19-20 October 2011
Lisbon, Portugal
• Diffuse attenuation coefficient $K_d(\lambda)$ of the spectral downward irradiance plays a critical role:

- Heat transfer in the upper ocean (Chang and Dickey, 2004; Lewis et al., 1990; Morel and Antoine, 1994)
- Photosynthesis and other biological processes in the water column (Marra et al., 1995; McClain et al., 1996; Platt et al., 1988)
- Turbidity of the oceanic and coastal waters (Jerlov, 1976; Kirk, 1986)
• $K_d(\lambda)$ is an apparent optical property (Preisendorfer, 1976) $\Rightarrow$ varies with solar zenith angle, sky and surface conditions, depth

• Satellite observations: only effective method to provide large-scale maps of $K_d(490)$ over basin and global scales

• Ocean color remote sensing: vertically averaged value of $K_d(490)$ in the surface mixed layer
State-of-the-art (1/2)

• One Step Empirical relationships:

  - **NASA Meris algorithm (Werdell, 2009):**

    \[ K_d(490) = 10^{(-0.8515 - 1.8263 X + 1.8714 X^2 - 2.4414 X^3 - 1.0690 X^4)} + 0.0166 \]

    with \( X = \log_{10}(R_{rs}(490)/R_{rs}(560)) \)

  - **Alternative algorithm (Kratzer, 2008):**

    \[ K_d(490) = \exp(-1.03 \cdot \log(R_{rs}(490)/R_{rs}(620)) - 0.18) + 0.0166; \]
• Two-step empirical algorithm with intermediate link
  – Morel, 2007:
    • chl-a = \(10^{(0.4502748 - 3.259491 \times X + 3.522731 \times X^2 - 3.359422 \times X^3 - 0.949586 \times X^4)}\)
    with \(X = \max(Rrs(443), Rrs(490), Rrs(510))\)
    • \(Kd(490) = 0.0166 + 0.07242[chl-a]^{0.68955}\)
Way to improve the estimation

- Use of artificial neural networks → Multi-Layer Perceptron (MLP)
  - Purely empirical method
  - Non-linear inversion
  - Universal approximator of any derivable function
  - Can handle “easily” noise and outliers
  - Taking more spectral information

- Method widely used in atmospheric sciences but rarely in spatial oceanography
Principles of NN

• A MLP is a set of interconnected neurons that is able to solve complicated problems
• Each neuron receives from and send signals to only the neurons to which it is connected

• Applications in geophysics:
  - Non-linear regression and inversion (Badran and Thiria, J. Phys. IV, 1998; Cherkassky, Neural Networks, 2006)

Advantages:
• Universal approximators of any non-linear continuous and derivable function
• Multi-dimensional function
• More accurate and faster in operational mode

Limits and drawbacks:
• Need adequate database
• Learning phase is time consuming
• Number of hidden layers and neurons unknown: need to determine them
Dataset

• **Learning/testing datasets → Calibration of the NN**
  - NOMAD database (Werdell and Bailey, 2005):
    - 337 set of \((R_{rs}, K_d(\lambda))\) per wavelength
  - IOCCG synthetical dataset (http://ioccg.org/groups/lee.html):
    - 1500 set of \((R_{rs}, K_d(\lambda))\) per wavelength
    - Three solar angles: 0°, 30°, 60°

• **80% of the entire dataset randomly taken for the learning phase** (e.g., determination of the optimal configuration of the artificial neural networks)

• The rest of the dataset used for the validation phase
Architecture of the Multi-Layered Perceptron:
Two hidden layers with 7 neurons on the first layer and 4 on the second layer
Dataset

- **Learning/testing datasets → Calibration of the NN**
  - NOMAD database (Werdell and Bailey, 2005):
    - 337 set of (Rrs,Kd(λ)) per wavelength
  - IOCCG synthetical dataset (http://ioccg.org/groups/lee.html):
    - 1500 set of (Rrs, Kd(λ)) per wavelength
    - Three solar angles: 0°, 30°, 60°

- 80% of the entire dataset randomly taken for the learning phase (e.g., determination of the optimal configuration of the artificial neural networks)

- The rest of the dataset used for the validation phase
<table>
<thead>
<tr>
<th></th>
<th>RMS (m(^{-1}))</th>
<th>Relative error (%)</th>
<th>Slope</th>
<th>(r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN</td>
<td>0.110</td>
<td>10.09</td>
<td>1.0</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Statistics on the test dataset
Comparison with other methods

• **COASTLOOC DATABASE (Babin et al., 2003)**
  - Observations in European coastal waters between 1997 and 1998
  - Entirely independent dataset from NOMAD and IOCCG
  - $K_d(490)$ ranging from 0.023 m$^{-1}$ and 3.14 m$^{-1}$ with a mean value of 0.64 m$^{-1}$
  - Nb total data: 132

• **Comparison of $K_d(490)$**
<table>
<thead>
<tr>
<th></th>
<th>Werdell</th>
<th>Morel</th>
<th>Kratzer</th>
<th>NN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RMS</strong></td>
<td>1.204</td>
<td>0.732</td>
<td>0.846</td>
<td>0.212</td>
</tr>
<tr>
<td><strong>Relative error (%)</strong></td>
<td>48.81</td>
<td>43.17</td>
<td>124.48</td>
<td>25.23</td>
</tr>
<tr>
<td><strong>Slope</strong></td>
<td>0.24</td>
<td>0.12</td>
<td>0.49</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Intercept</strong></td>
<td>0.34</td>
<td>0.28</td>
<td>0.76</td>
<td>0.16</td>
</tr>
<tr>
<td><strong>r</strong></td>
<td>0.13</td>
<td>0.19</td>
<td>0.40</td>
<td>0.94</td>
</tr>
</tbody>
</table>
Conclusions and Perspectives

• On the used dataset:
  - **Net overall improvement** of the estimation of the Kd(λ)
  - **Same quality** for the very low values of Kd(490), i.e. < 0.2 m\(^{-1}\)
  - **Huge improvement** for the greater values, especially for very turbid waters (K_d(490) > 0.5 m\(^{-1}\))

• **Will be freely available at:**
<table>
<thead>
<tr>
<th>SeaWiFS</th>
<th>412</th>
<th>443</th>
<th>510</th>
<th>555</th>
<th>670</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>0.379</td>
<td>0.249</td>
<td>0.227</td>
<td>0.196</td>
<td>0.206</td>
</tr>
<tr>
<td>Relative error (%)</td>
<td>31.57</td>
<td>26.08</td>
<td>31.87</td>
<td>22.34</td>
<td>15.70</td>
</tr>
<tr>
<td>Slope</td>
<td>1.02</td>
<td>0.88</td>
<td>0.68</td>
<td>0.64</td>
<td>0.67</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.15</td>
<td>0.18</td>
<td>0.16</td>
<td>0.12</td>
<td>0.29</td>
</tr>
<tr>
<td>r</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Statistical results for a SeaWiFS Kd from COASTLOOC database
Acknowledgments

• CNRS and INSU for funding
• Marcel Babin for providing the COASTLOOC database
• IOCCG for providing the synthetical database
• NASA for providing the NOMAD database