Retrieval of the diffuse attenuation coefficient Kd(λ) 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

Purpose of the study (1/2)

- **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)

Purpose of the study (2/2)

- $K_d(\lambda)$ is an apparent optical property (Preisendorfer, 1976) → 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):

    - Kd(490)= $10^{(-0.8515 - 1.8263X + 1.8714X^2 - 2.4414X^3 - 1.0690X^4)} + 0.0166$

    with X=log10(Rrs(490)/Rrs(560))

  - Alternative algorithm (Kratzer, 2008)

    - Kd(490)= exp(-1.03*log(Rrs(490)/Rrs(620))-0.18) + 0.0166;

State-of-the art (2/2)

- Two-step empirical algorithm with intermediate link
  - Morel, 2007:
    - chl-a= $10^{(0.4502748-3.259491*X+3.522731*X^2-3.359422*X^3-0.949586*X^4)}$

      with X=max(Rrs(443),Rrs(490),Rrs(510))
    - Kd(490)=0.0166 + 0.07242[chl-a]$^{0.68955}$

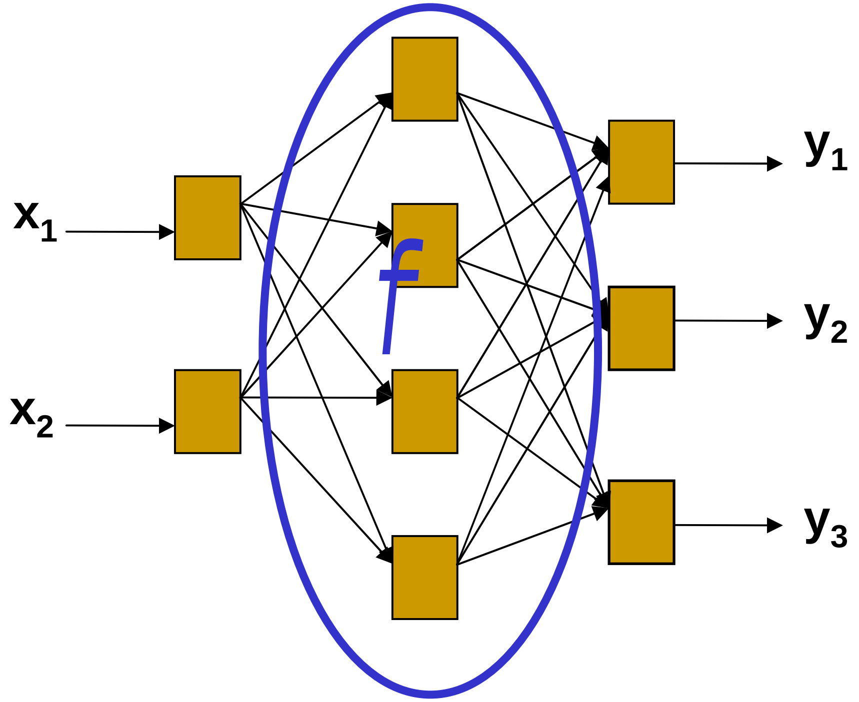
Way to improve the estimation

- 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)
  - Statistical analysis of dataset (Hsieh, W.W, Rev. Geophys., 2004)

Advantages:

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

Limits and drawbacks:

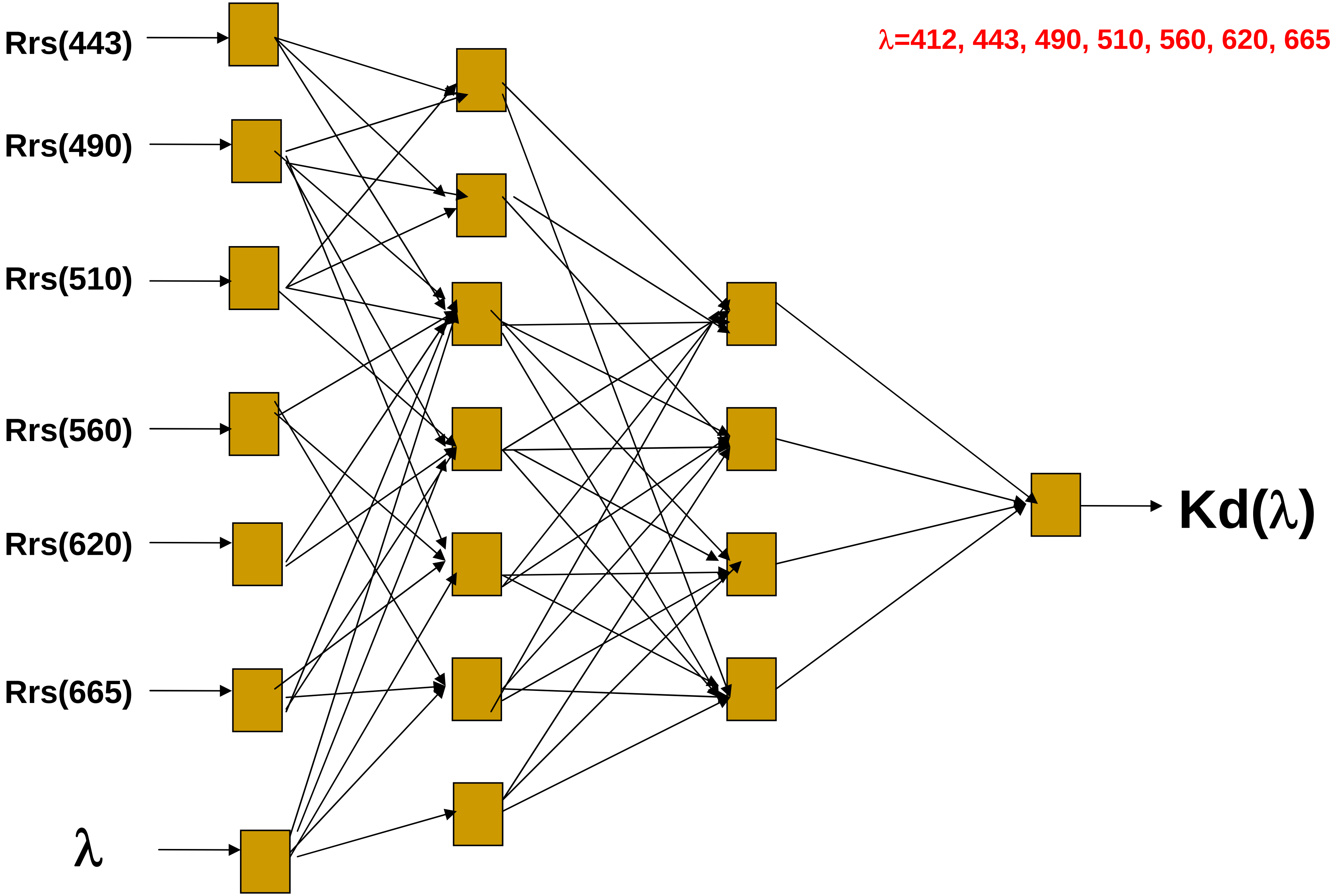
- Need adequate database
- Learning phase is time consuming
- Number of hidden layers and neurons unknown: need to determine them

Dataset

- **Learning/testing datasets → Calibration of the NN**
  - NOMAD database (Werdell and Bailey, 2005):
    - 337 set of (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



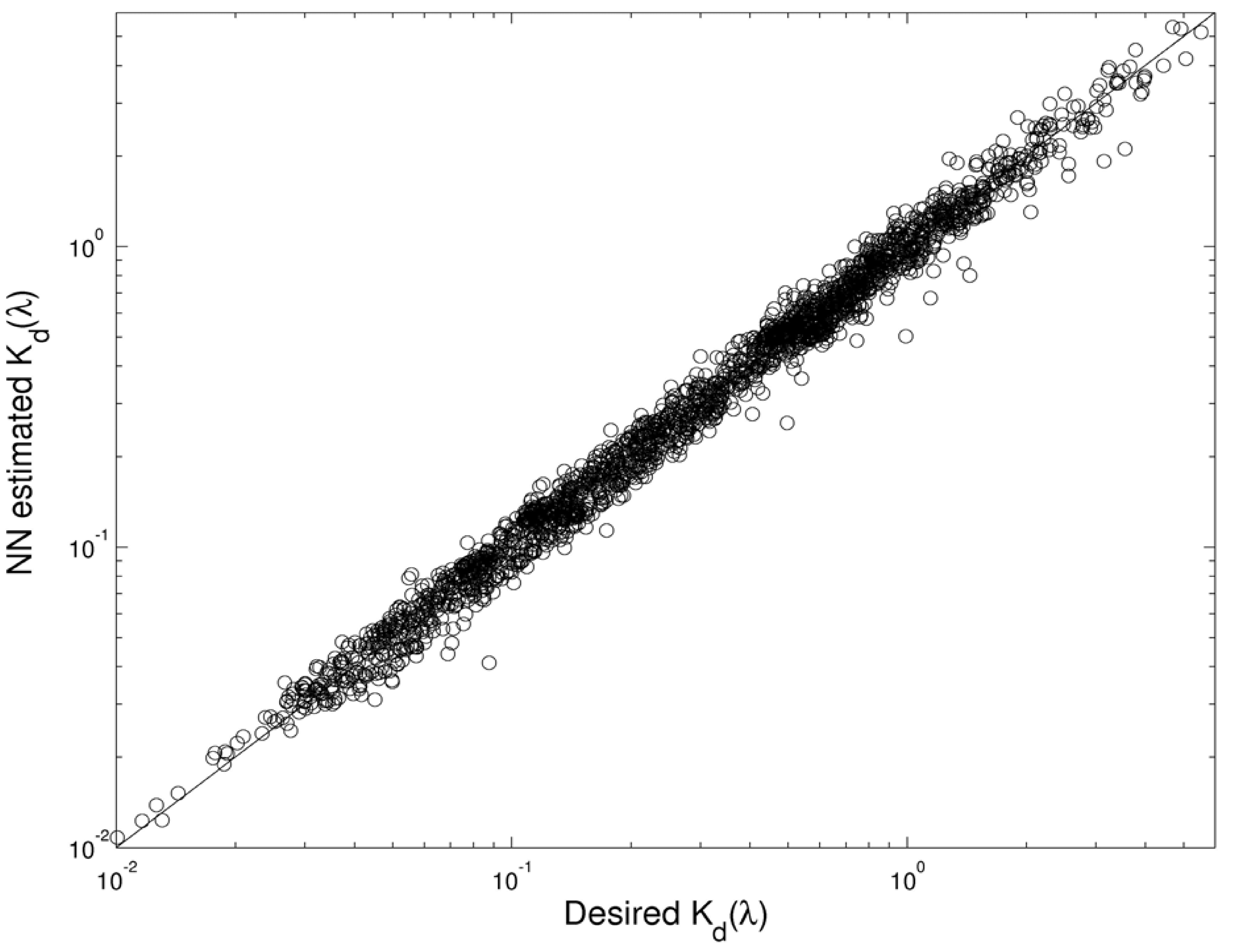
Architecture of the Multi-Layered Perceptron:
Two hidden layers with 7 neurons on the first layer and 4 on the second layer

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



NN estimated K_d(λ)
Desired K_d(λ)
10^-2
10^-1
10^0

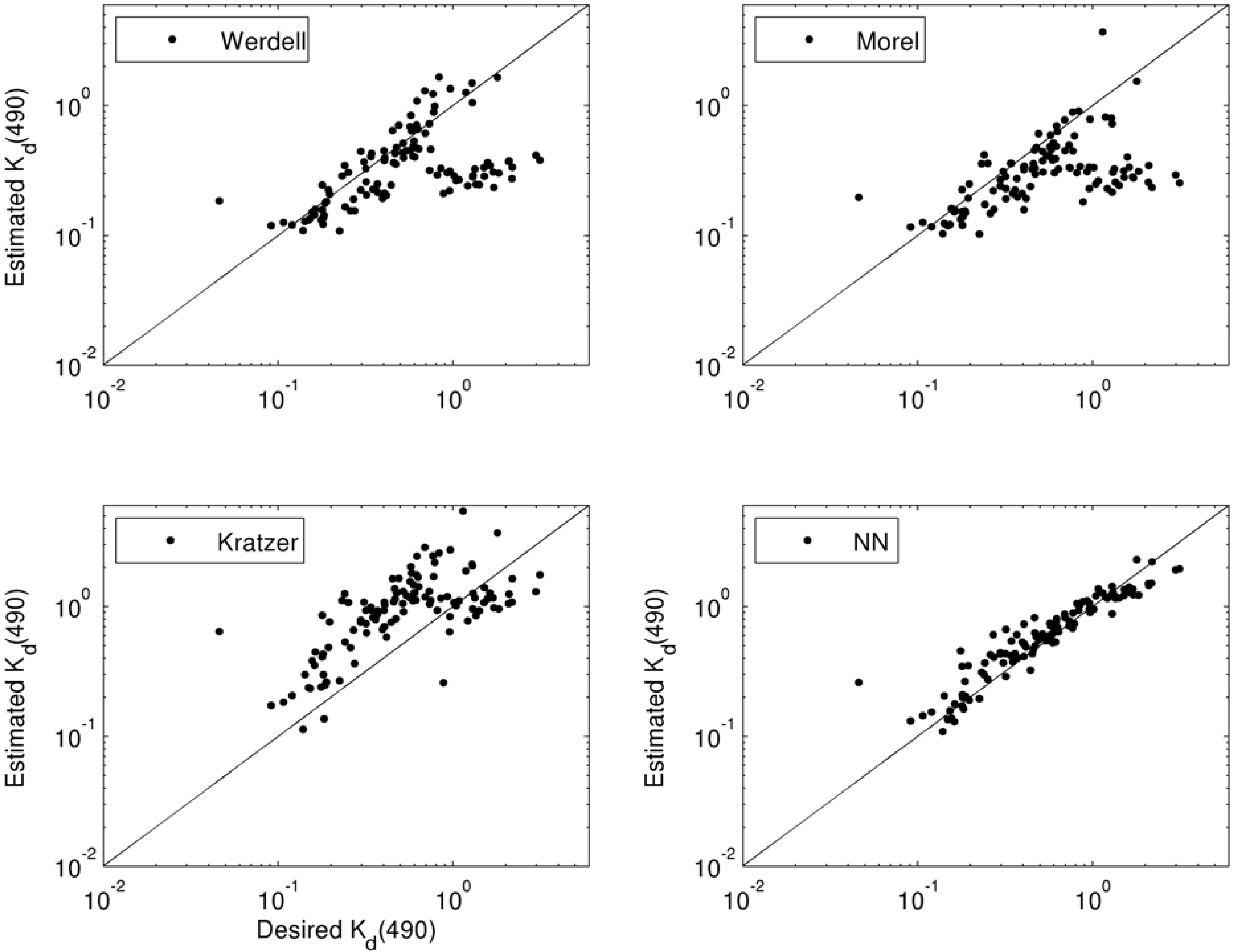
	RMS (m^{-1})	Relative error (%)	Slope	*r*
NN	**0.110**	**10.09**	**1.0**	**0.98**

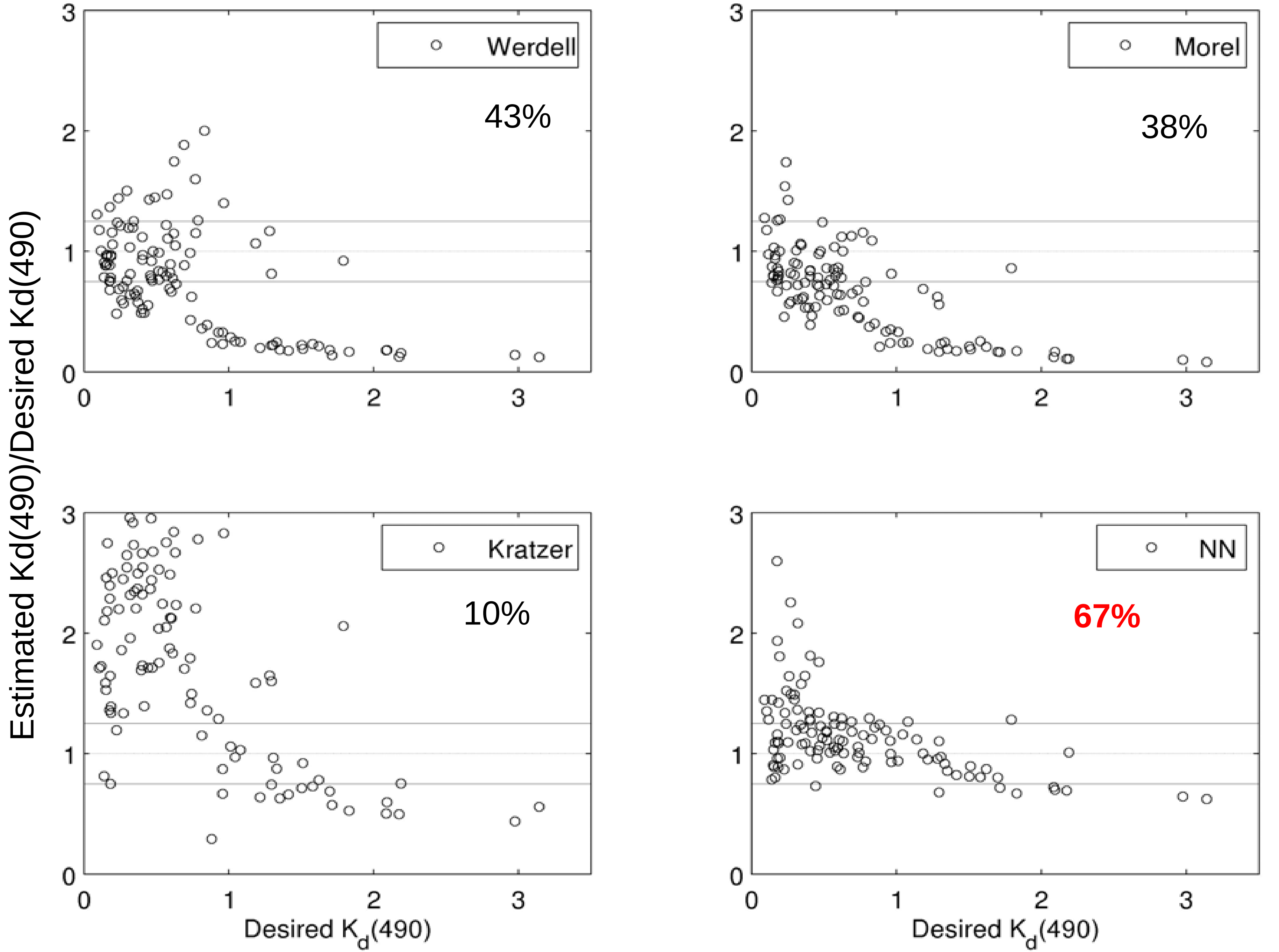
Statistics on the test dataset

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**
  - **Kd(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 Kd(490)**


Werdell
Morel
Kratzer
NN
Estimated $K_d(490)$
Desired $K_d(490)$
10^{-2}
10^{-1}
10^{0}



Estimated Kd(490)/Desired Kd(490)
Werdell
43%
Morel
38%
Kratzer
10%
NN
67%
Desired K_d(490)
Desired K_d(490)
0
1
2
3

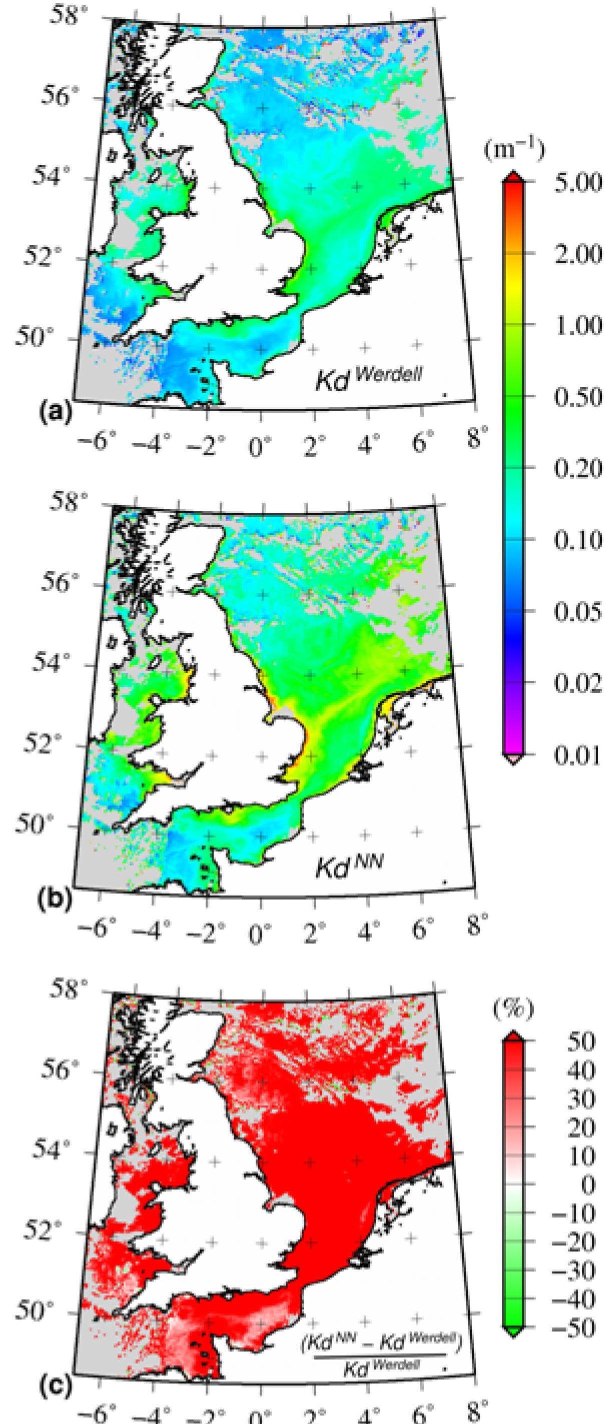
	Werdell	*Morel*	*Kratzer*	*NN*
RMS	1.204	0.732	0.846	**0.212**
Relative error (%)	48.81	43.17	124.48	**25.23**
Slope	0.24	0.12	0.49	**0.79**
Intercept	0.34	0.28	0.76	**0.16**
r	0.13	0.19	0.40	**0.94**

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:
  - **http://log.univ-littoral.fr/oceano/**

SeaWiFS	*412*	*443*	*510*	*555*	*670*
RMS	0.379	0.249	0.227	0.196	0.206
Relative error (%)	31.57	26.08	31.87	22.34	15.70
Slope	1.02	0.88	0.68	0.64	0.67
Intercept	0.15	0.18	0.16	0.12	0.29
r	0.95	0.95	0.95	0.94	0.87

Statistical results for a SeaWiFS Kd from COASTLOOC database


58°
56°
54°
52°
50°
Kd Werdell
(a)
−6°
−4°
−2°
0°
2°
4°
6°
8°
(m⁻¹)
5.00
2.00
1.00
0.50
0.20
0.10
0.05
0.02
0.01
Kd NN
(b)
(%)
50
40
30
20
10
0
−10
−20
−30
−40
−50
(Kd NN − Kd Werdell) / Kd Werdell
(c)

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