



Earth Observation Laboratory

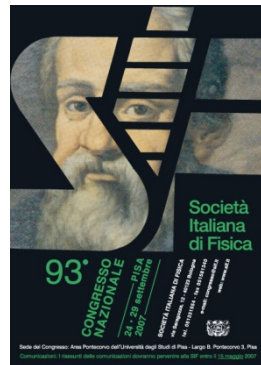
PhD Program in GeoInformation

DISP - Tor Vergata University of Rome

Algoritmi di Rete Neurale per l'Inversione del Dato Satellitare

Autori:

P. Sellitto, E. Angiuli, G. Licciardi, A. Minchella



SOMMARIO (1/2)

Il considerevole incremento dell'informazione raccolta dai molteplici sensori a bordo delle attuali piattaforme satellitari ha incoraggiato la sperimentazione di nuove tecniche di processamento dati veloci ed affidabili.

Gli algoritmi di rete neurale sono stati usati negli ultimi anni in numerose aree delle scienze ambientali, della climatologia e del telerilevamento.

Le loro applicazioni all'osservazione della Terra sono legate alla classificazione, modellizzazione, inversione del dato, previsione.

In questa corrispondenza sono presentate alcune attività legate all'impiego delle reti neurali per classificazione ed inversione del dato satellitare con applicazione all'osservazione dell'atmosfera, degli oceani e della superficie terrestre.

SOMMARIO (2/2)

- **INTRODUZIONE**

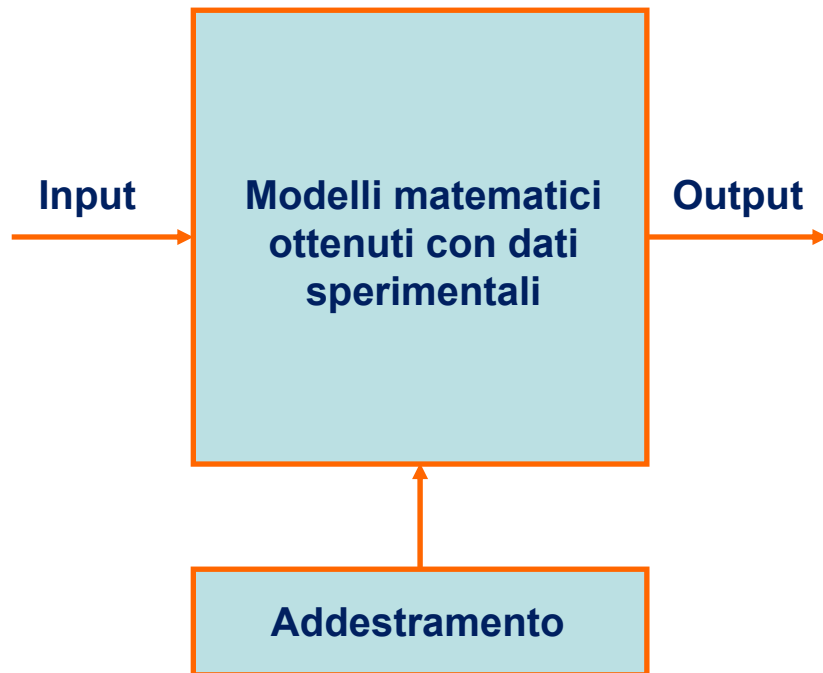
- I. TEORIA DELLE RETI NEURALI
- II. LE RETI NEURALI E IL TELERILEVAMENTO SATELLITARE

- **ALCUNI ESEMPI**

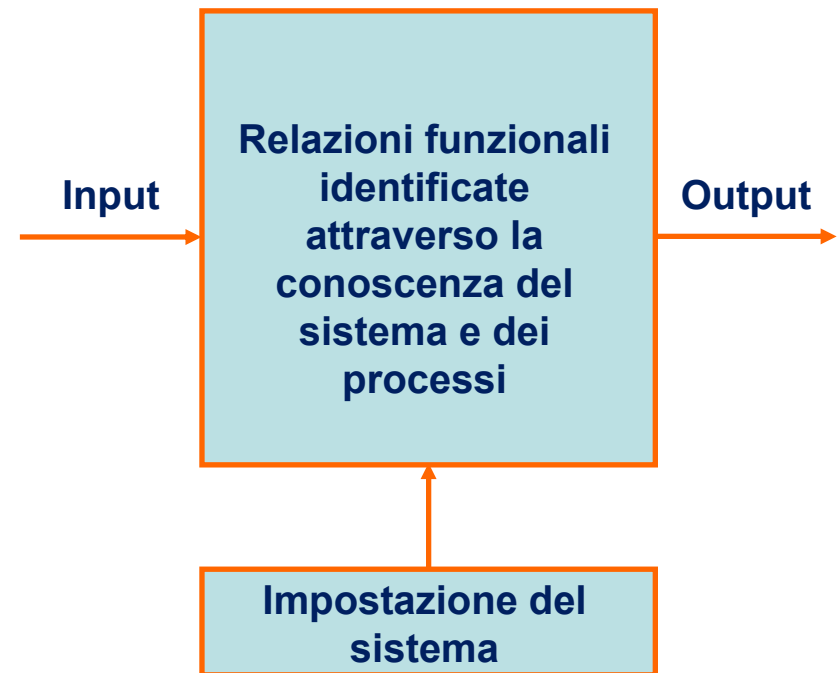
- I. TELERILEVAMENTO DI OZONO DA SENSORI UV/VIS
- II. INVERSIONE DEL DATO SAR POLARIMETRICO-INTERFEROMETRICO
- III. OIL SPILL DETECTION
- IV. CLASSIFICAZIONE/LAND-USE SU AREE URBANE DA SENSORI MULTISPETTRALI
- V. INDIVIDUAZIONE DI HARMFUL ALGAE BLOOM (HAB) DA SENSORI MULTISPETTRALI

Le reti neurali (NN) appartengono alla categoria delle tecniche di *soft computing*. Identificano le dipendenze tra variabili di **input** e variabili di **output** attraverso dei **dataset di addestramento**. Dopo l'addestramento le NN rappresentano delle funzioni multivariate non lineari.

NN – Learning/soft computing



Algoritmi standard - Hard computing



Le NN risolvono problemi di **regressione** e **classificazione** variando dei parametri (**pesi**, **w**) che controllano il modo in cui *imparano* mentre scorrono i dati di addestramento. Le performance del modello sono misurate da una funzione di errore **$E=E(w)$** .

E dipende in maniera non lineare dai pesi **w** e l'ipersuperficie non ha necessariamente un unico e garantito minimo.

Il problema dell'addestramento di una NN cade nella categoria dei problemi di ottimizzazione non lineare.

INTRODUZIONE - TEORIA DELLE RETI NEURALI (3/10)

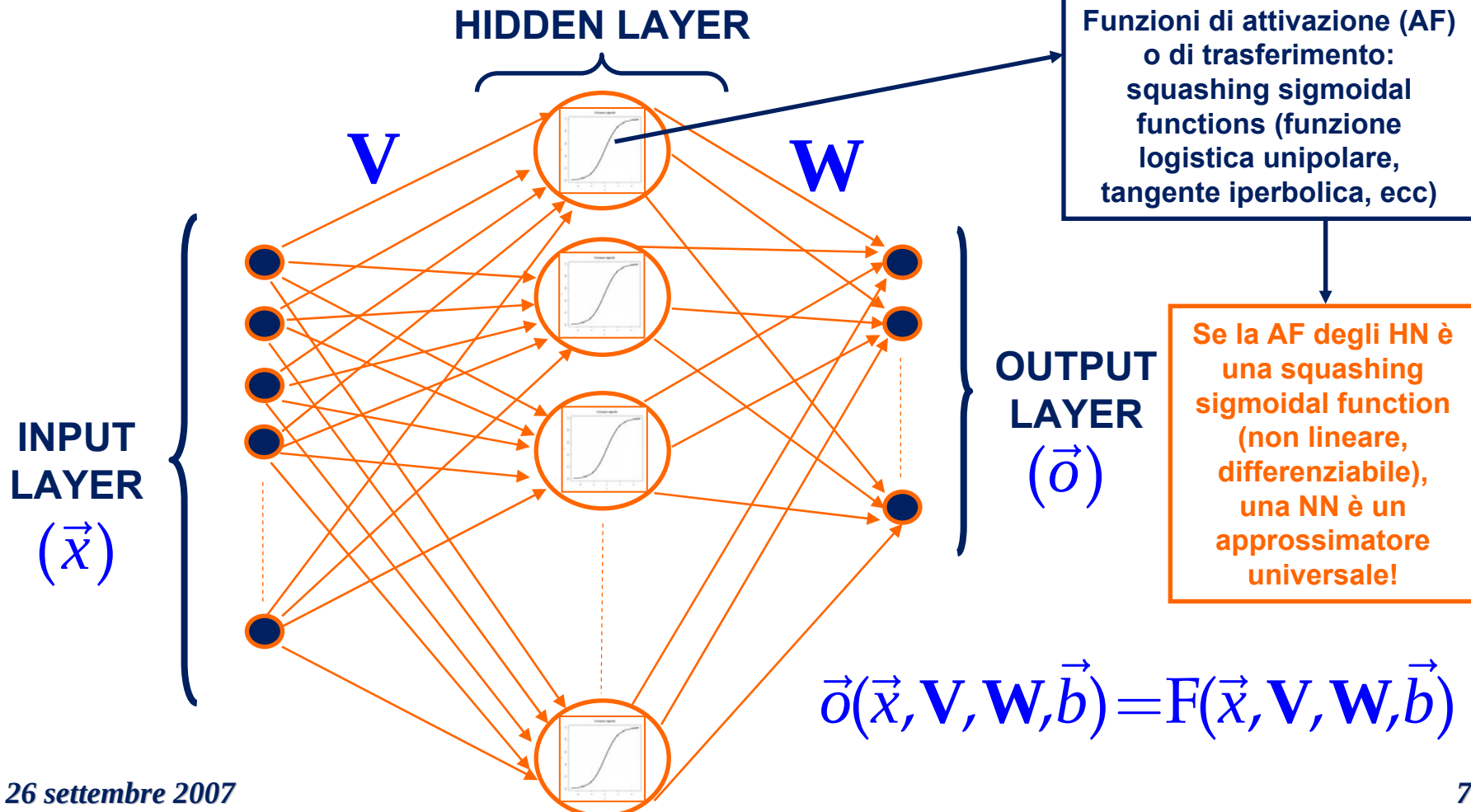
Vantaggi		Svantaggi
Hanno la proprietà di imparare dai dati (misure), imitando le funzioni dell'apprendimento umano/animale	→	Necessitano addestramento, che può essere un processo lungo
Possono approssimare ogni funzione multivariata non lineare	→	Sono soggette a cattiva generalizzazione
Non richiedono conoscenza del processo da modellizzare	→	Non palesano le relazioni tra i parametri fisici che descrivono il processo
Sono robuste alla presenza di dati rumorosi		
Hanno struttura parallela, facilmente implementabile		
Le stesse reti possono risolvere problemi diversi	→	Ci sono poche indicazioni a priori rispetto al tipo di rete da usare e la sua ottimizzazione (numero di neuroni hidden, algoritmo di training, ecc)

INTRODUZIONE - TEORIA DELLE RETI NEURALI (4/10)

Una NN è costituita da unità di processamento, i **neuroni**, connessi tra loro da sinapsi di peso **w**.

I neuroni sono solitamente organizzati in strati o layer.

Feedforward NN con addestramento supervisionato



Per un output layer monodimensionale:

$$o(\vec{x}, \mathbf{V}, \vec{W}, \vec{b}) = F(\vec{x}, \mathbf{V}, \vec{W}, \vec{b}) = \sum_{j=1}^{\bar{J}} w_j \sigma_j(\vec{v}_j \cdot \vec{x} + b_j)$$

Se si considera il vettore $\vec{b}$ come ulteriore e ultima colonna della matrice $\mathbf{V}$:

$$o = \vec{W} \cdot \sigma(\mathbf{V}\vec{x})$$

Ci sono due aspetti importanti da sottolineare:

- il problema della rappresentazione: cosa possono rappresentare questi modelli?
- il problema dell'apprendimento: come si può far rappresentare a questi modelli quello che vogliamo?

Cosa possono rappresentare le NN?

Problema dell'approssimazione di una funzione

$$f(\vec{x}) \rightarrow f_a(\vec{x}, \mathbf{W})$$

Le NN sono degli approssimatori universali!

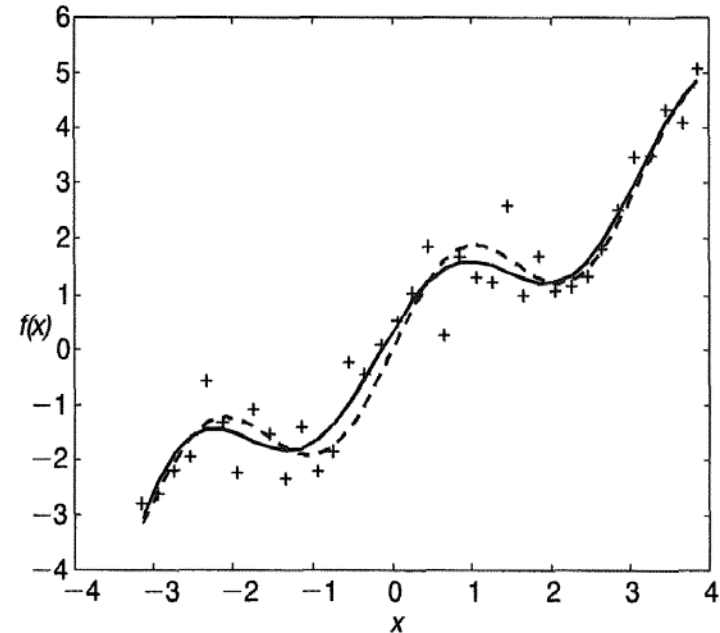
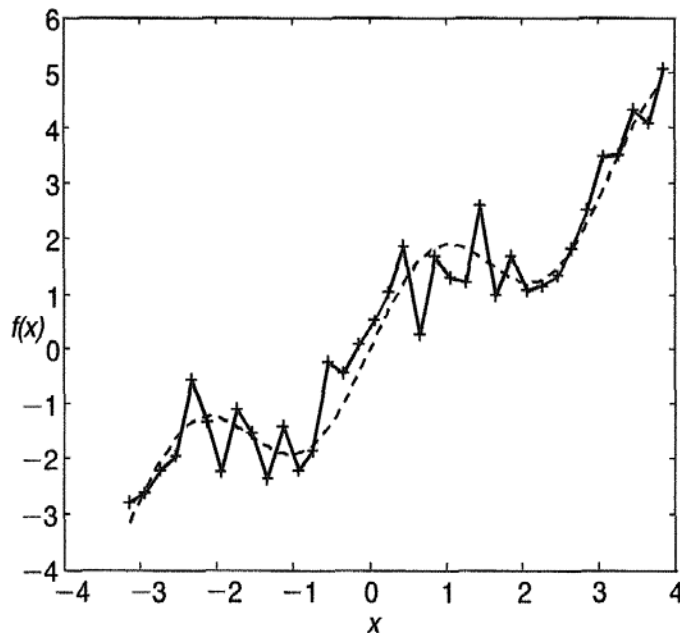
Il teorema di Weierstrass può essere esteso a combinazioni di funzioni sigmoidali – Teorema di Cybenko-Hornik-Funahashi:

$$f_a(\vec{x}, \mathbf{W}) = \sum_{j=1}^{\bar{J}} w_j \sigma_j(\vec{v}_j \cdot \vec{x} + b_j) \text{ è denso in } C[I_d], I_d \equiv [0, 1]^d$$

$$\forall \vec{x} \in I_d, f(\vec{x}) \in C[I_d], \varepsilon > 0 \quad \exists f_a(\vec{x}, \mathbf{W}):$$

$$|f_a(\vec{x}, \mathbf{W}) - f(\vec{x})| < \varepsilon$$

L'ordine della funzione approssimante (numero di neuroni hidden) deve essere abbastanza elevato da seguire l'andamento dei punti sperimentali e abbastanza limitato da non produrre *overfitting* e filtrare il rumore:



Come possiamo fare in modo che le NN rappresentino quello che vogliamo?

Problema della ottimizzazione non lineare

Si definisce una norma che misuri la bontà dell'approssimazione:

$$E = \sum_{p=1}^{\bar{p}} \left[\vec{d}_p - f_a \left(\vec{x}_p, \mathbf{W} \right) \right] = \sum_{p=1}^{\bar{p}} \left[\vec{d}_p - \vec{o}_p \right]$$

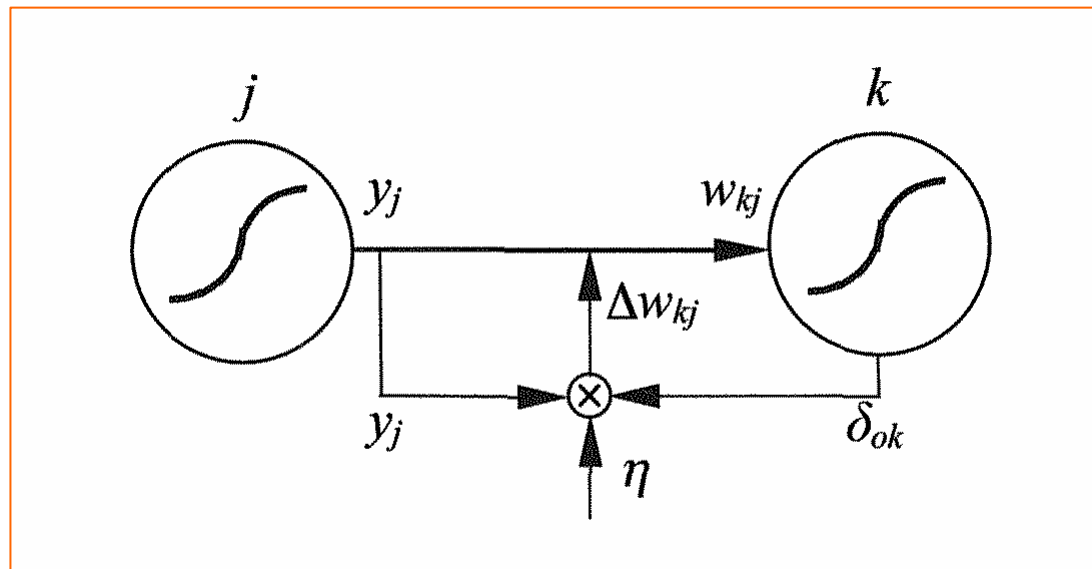
La funzione di errore si può minimizzare mediante adattamento dei pesi $\mathbf{W}$ attraverso apprendimento basato sul gradiente. La direzione locale del vettore gradiente (-) della funzione errore indica la massima diminuzione dell'ipersuperficie definita da $E(\mathbf{W})$:

$$\vec{g} = \nabla E(\mathbf{W}) = \begin{pmatrix} \frac{\partial E}{\partial w_{11}} \\ \vdots \\ \frac{\partial E}{\partial w_{NM}} \end{pmatrix}$$

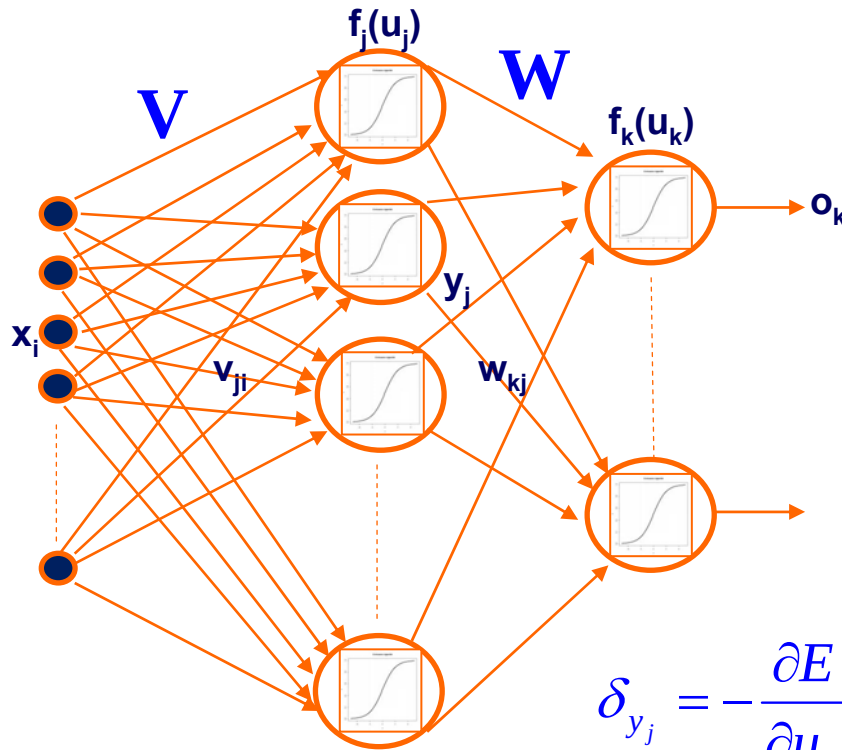
I pesi W vengono *aggiustati* iterativamente secondo lo schema:

$$\mathbf{W}_{it} = \mathbf{W}_{it-1} + \eta_i \Delta \mathbf{W}_{it} \left(\vec{g}(\mathbf{W}_{it-1}) \right)$$

Algoritmo di Error Backpropagation



Delta rule generalizzata: aggiustamenti di W per mezzo del segnale di errore (misura del gradiente dell'errore) *retropropagato*



$$E = \frac{1}{2} \sum_{k=1}^{\bar{k}} (d_k - o_k)^2$$

$$\delta_{o_k} = -\frac{\partial E}{\partial u_k} = (d_k - o_k) \mathbf{f}'(u_k)$$

$$\delta_{y_j} = -\frac{\partial E}{\partial u_j} = -\sum_{k=1}^{\bar{k}} (d_k - o_k) \mathbf{f}'(u_k) \frac{\partial u_k}{\partial y_j} = -\sum_{k=1}^{\bar{k}} \delta_{o_k} w_{jk}$$

$$\Delta w_{kj} = \eta \delta_{o_k} y_j \Rightarrow w_{kj} (it) = w_{kj} (it-1) + \eta \delta_{o_k} y_j$$

$$\Delta v_{ji} = \eta \delta_{y_j} x_i \Rightarrow v_{ji} (it) = v_{ji} (it-1) + \eta \delta_{y_j} x_i$$

Negli anni '90 il settore delle NN era maturo per applicazioni in vaste aree della scienza e della tecnologia. Un numero significativo di applicazioni NN è stato sviluppato nel campo delle geoscienze: telerilevamento satellitare e non, meteorologia, numerical modeling, ecc.

L'incremento di dati da sensori satellitari (incremento di risoluzione spaziale e temporale) ha richiesto algoritmi di elaborazione e assimilazione rapidi (ed accurati).

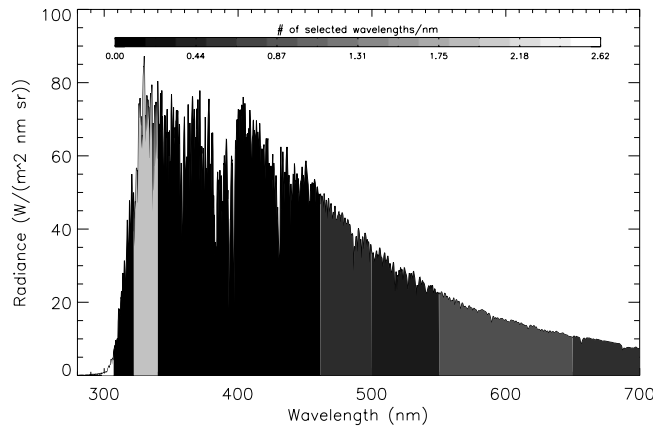
Le NN possono essere uno strumento utile in questo contesto:

$$\vec{y} = M(\vec{x}), \vec{x} \in \mathfrak{R}^n, \vec{y} \in \mathfrak{R}^m$$

converte radianze, coefficienti di backscattering, le *features*, ecc, in parametri geofisici o in *classi*.

PURPOSE:

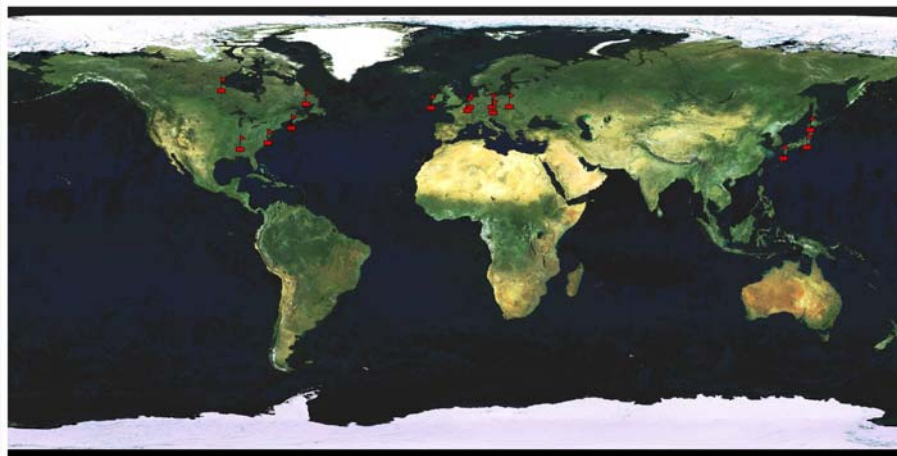
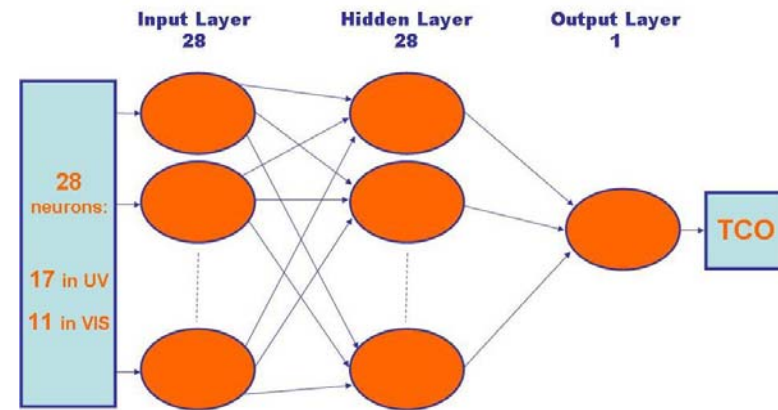
To develop **fast** and **reliable** Neural Networks (NNs) algorithm for direct tropospheric ozone retrieval (e.g., Tropospheric Column Ozone - TCO) from UV/VIS Earth's radiance spectra measured from satellite. ESA-Envisat SCIAMACHY used as test sensor.



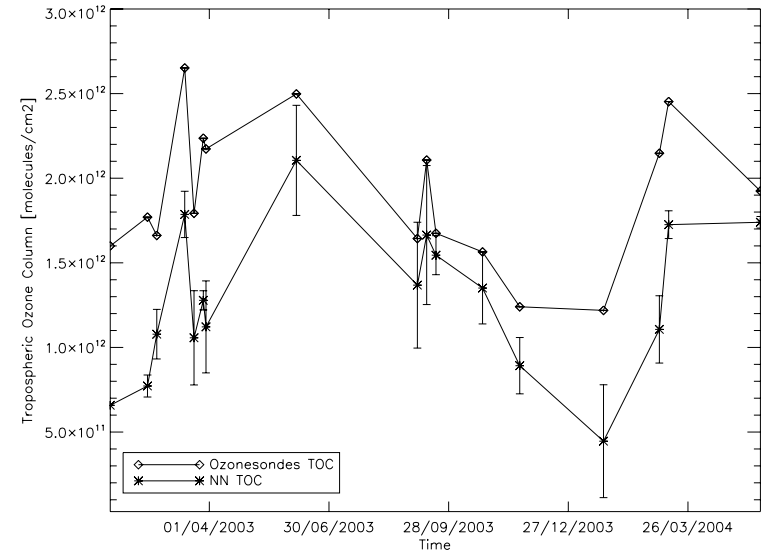
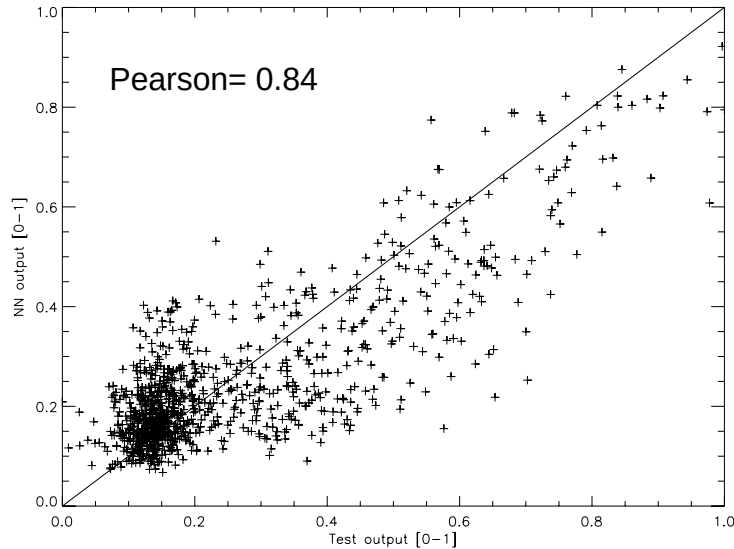
✓ UV sub-band 298-307 nm most informative;
✓ VIS density of selected wavelengths consistent with Chappuis ozone bands structure.

OPERATIVE STEPS:

- wavelength selection;
- topology definition;
- experimental dataset preparation;
- test and validation.



- set of measured radiances extracted from **SCIAMACHY Level 1b** data;
- satellite data matched with ozonesondes measurements from **WOUDC** and **SHADOZ** databases;
- matching distance selected as 150 km (poles) and non-linearly varying with latitude until 400 km (equator);
- matching time distance 12 h;
- only northern mid-latitudes stations (30°N-60°N).



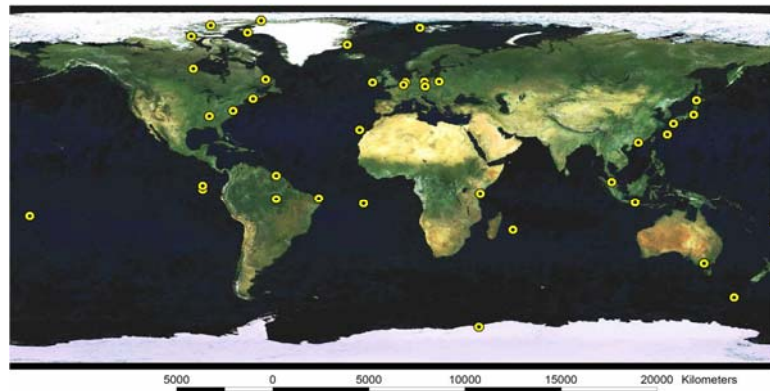
- ✓ independent test dataset to verify NN's reliability on simulated data: reasonably good correlation between observed and retrieved TOCs - UV/VIS NN improve of over 10% the retrieval capability in terms of Pearson coefficient with respect to the UV-only case;
- ✓ retrieved TOCs follow time trend of reference ozonesondes data at some reference mid-latitudes sites (Debilt, The Netherlands and Uccle, Belgium overpass data).

ESA Cat-1 projects 2930/4069,
PI F. Del Frate, Co-investigator P. Sellitto

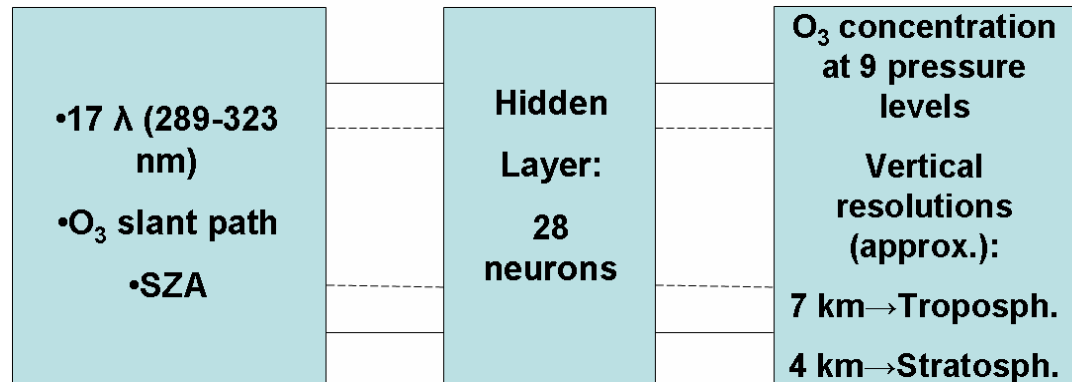
REFERENCES:

- [1] F. Del Frate et al. "Design of Neural Network algorithms for the retrieval of tropospheric ozone from satellite data", Proceedings of Envisat Symposium 2007, Montreux, Switzerland.
- [2] P. Sellitto et al. "Dedicated Neural Networks algorithms for direct estimation of tropospheric ozone from satellite measurements", Proceedings of IGARSS 2007, Barcelona, Spain.

Neural Networks algorithms for the estimation of atmospheric ozone from satellite data: ozone profiles

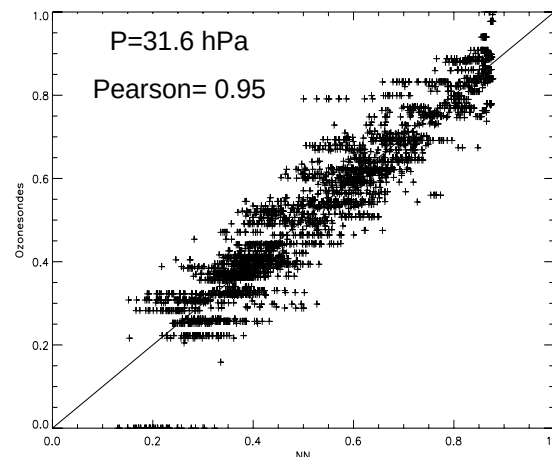
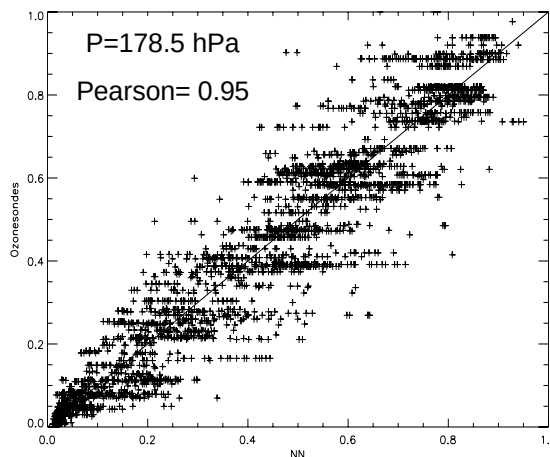


- NNs trained and tested with two standard ozonesondes databases: WOUDC/SHADOZ;
- a heterogeneous dataset is expected to represent the global ozone behaviour;
- test satellite measurements: ESA-Envisat SCIAMACHY channel 1 and 2.



inputs chosen by means of sensitivity analyses:

- satellite reflectance data in a relatively narrow UV interval;
- the ozone slant path obtained with the higher selected wavelengths;
- the solar zenith angle relatively to the measurement.



ESA Cat-1 projects 2930/4069, PI F. Del Frate, Co-investigator P. Sellitto

REFERENCE: P. Sellitto et al. "Neural Networks algorithms for the retrieval of ozone concentration profiles from Envisat-SCIAMACHY measurements", Proceedings of Envisat Symposium 2007, Montreux, Switzerland.

PolInSAR (L band SAR) (1/5)

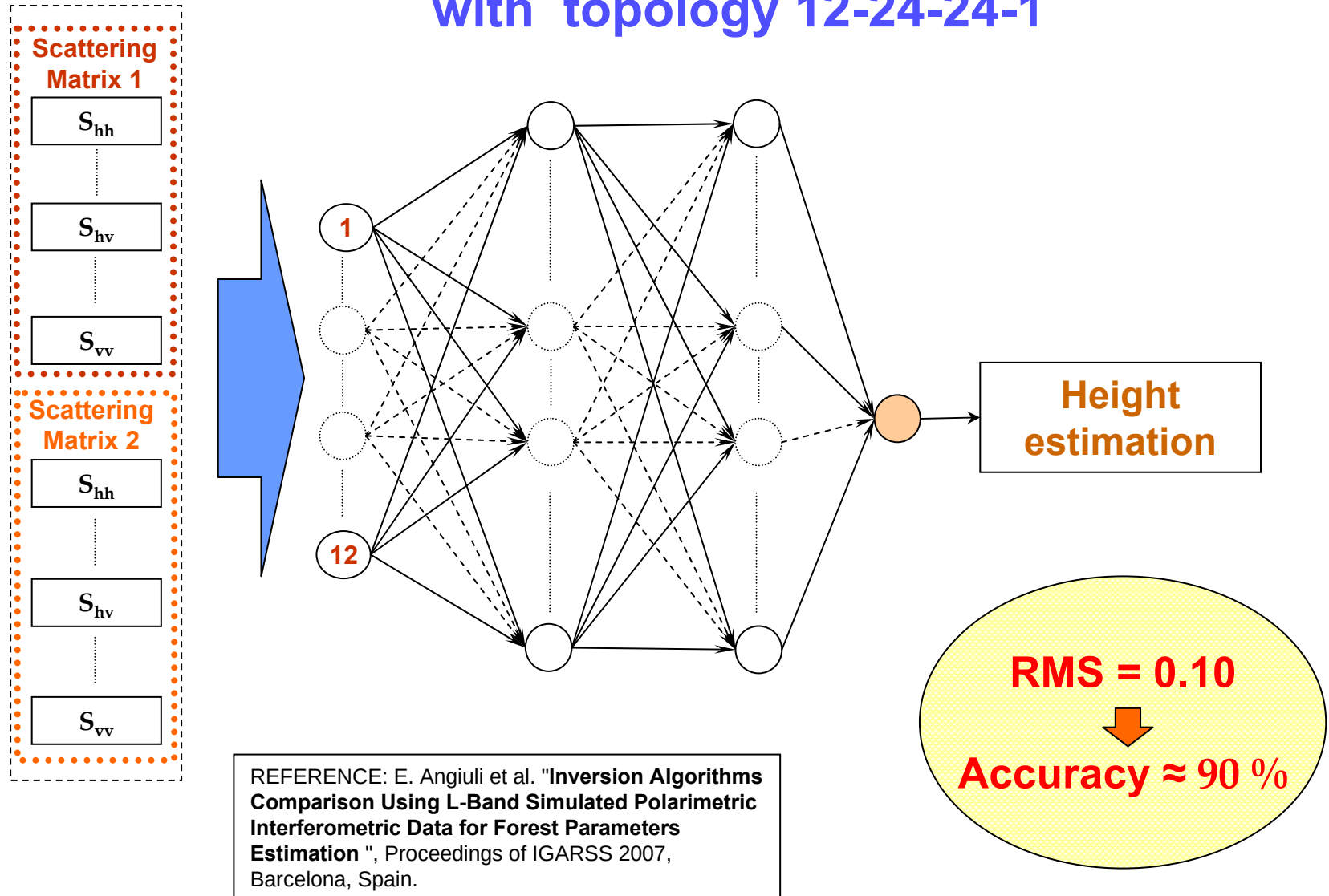
Polarimetric SAR Interferometry (PolInSAR) has been proposed as a tool for forest parameters retrieval.

- **Polarimetry:** yields information on backscattering mechanisms.
- **Interferometry:** yields phase information (height).
- **Polarimetry + interferometry :** yield height info for each backscattering mechanism.

Vegetation height and density can be estimated from interferometric phase and coherence.

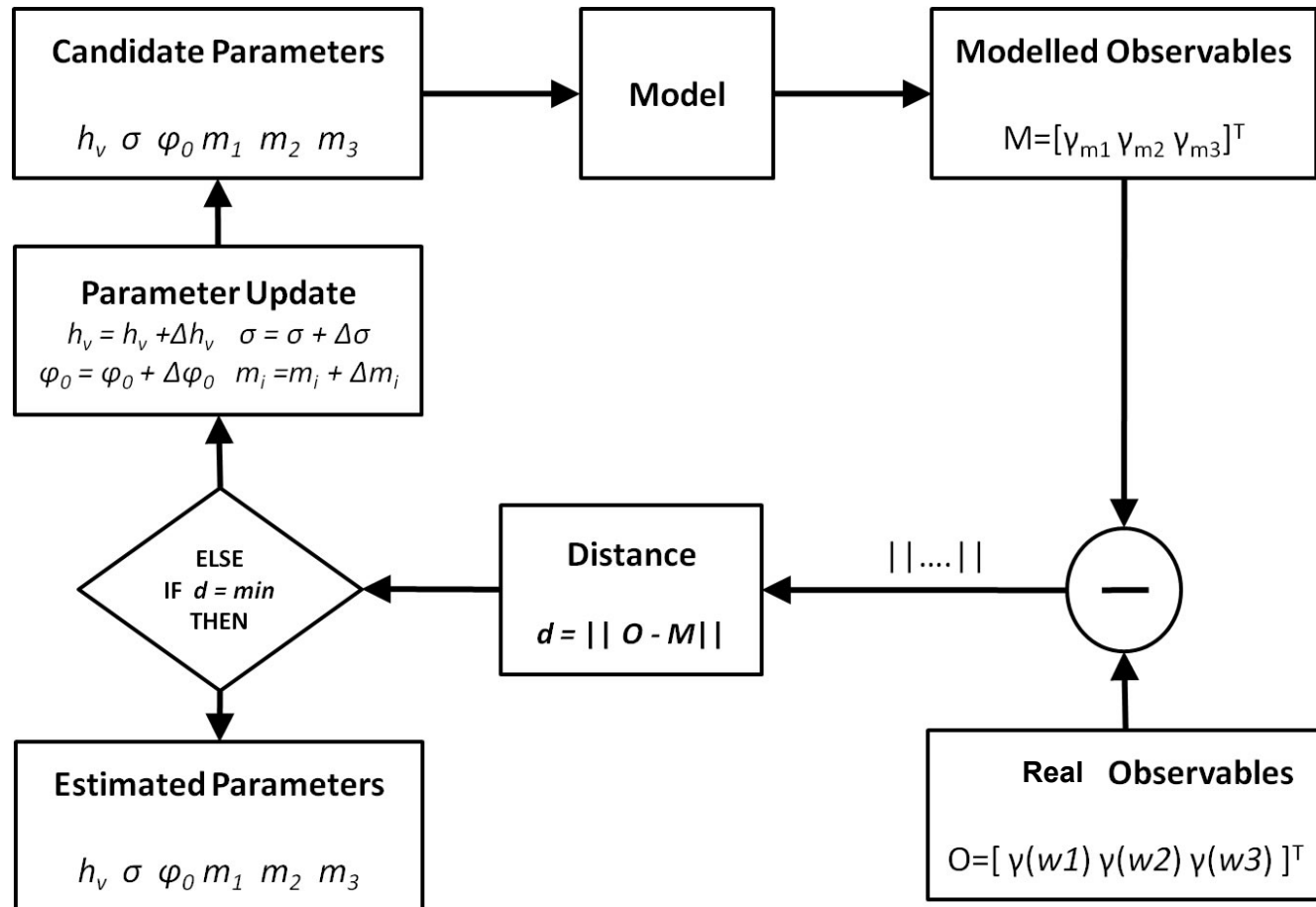
POL-INSAR Neural Network method (2/5)

MLP (Multi Layer Perceptron) network with topology 12-24-24-1

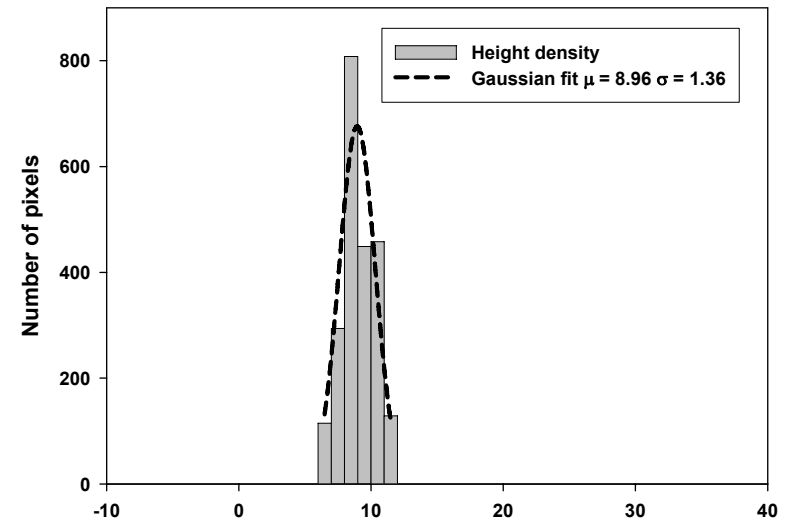
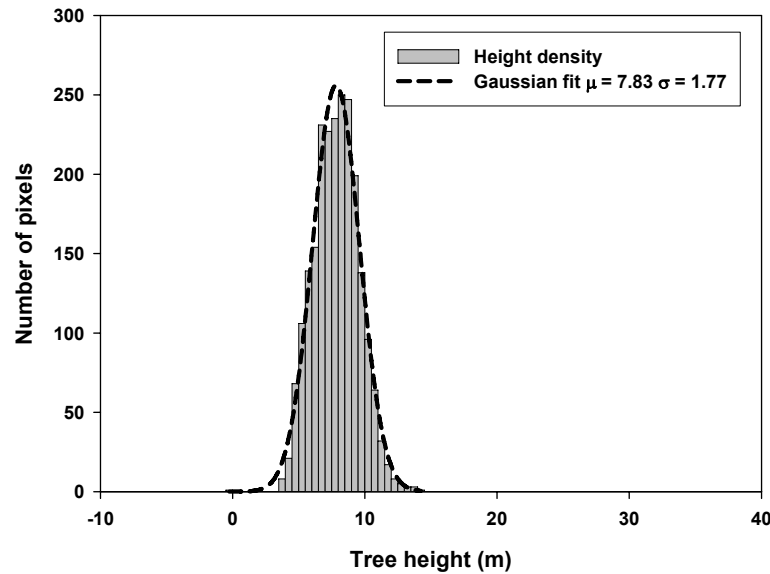


Optimal Estimation approach (3/5)

Optimal estimation (OE) minimizes difference between model predictions and observations in a least-square sense.



Inversions methods comparison (4/5)



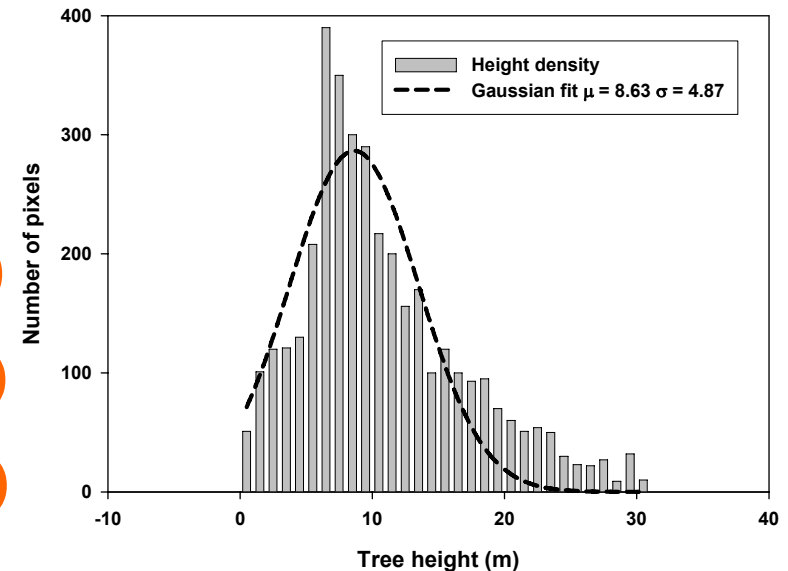
“True” tree height: 8 [m]

Retrieved tree height:

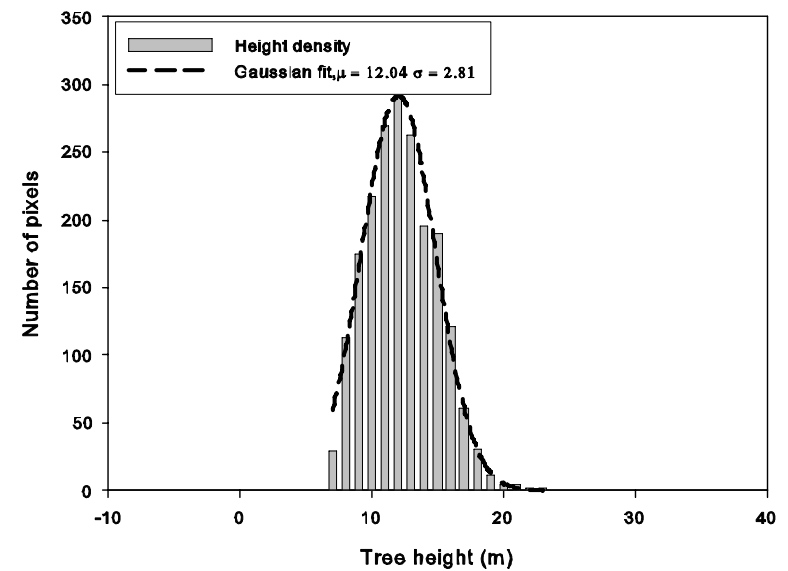
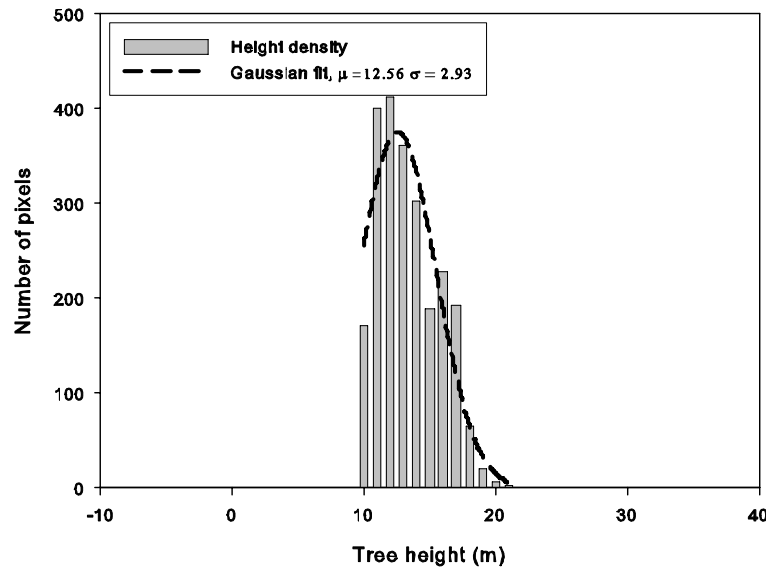
“manual” OE ≈ 7.83 [m] (97 %)

“automatic” OE ≈ 8.96 [m] (89 %)

Neural Network ≈ 8.63 [m] (92 %)



Inversions methods comparison (5/5)



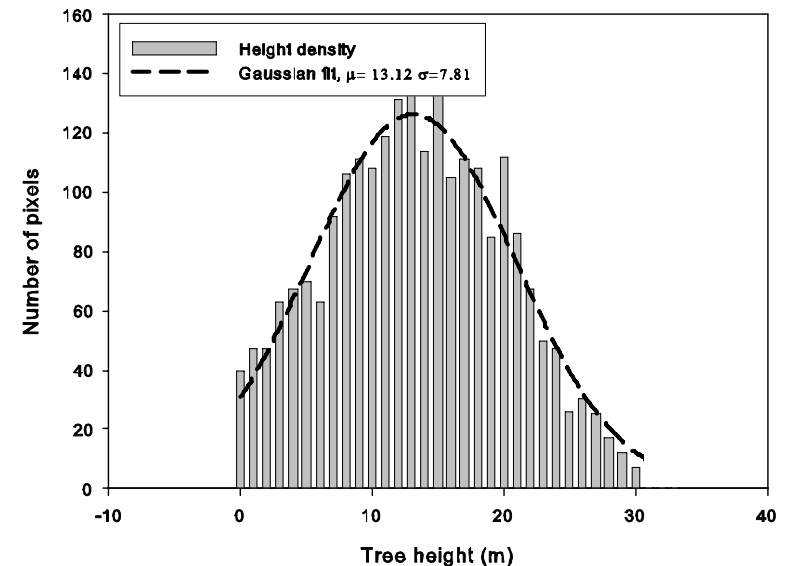
“True” tree height: 14 [m]

Retrieved tree height:

“manual” OE ≈ 12.6 [m] (89 %)

“automatic” OE ≈ 12.0 [m] (86 %)

Neural Network ≈ 13.0 [m] (93 %)

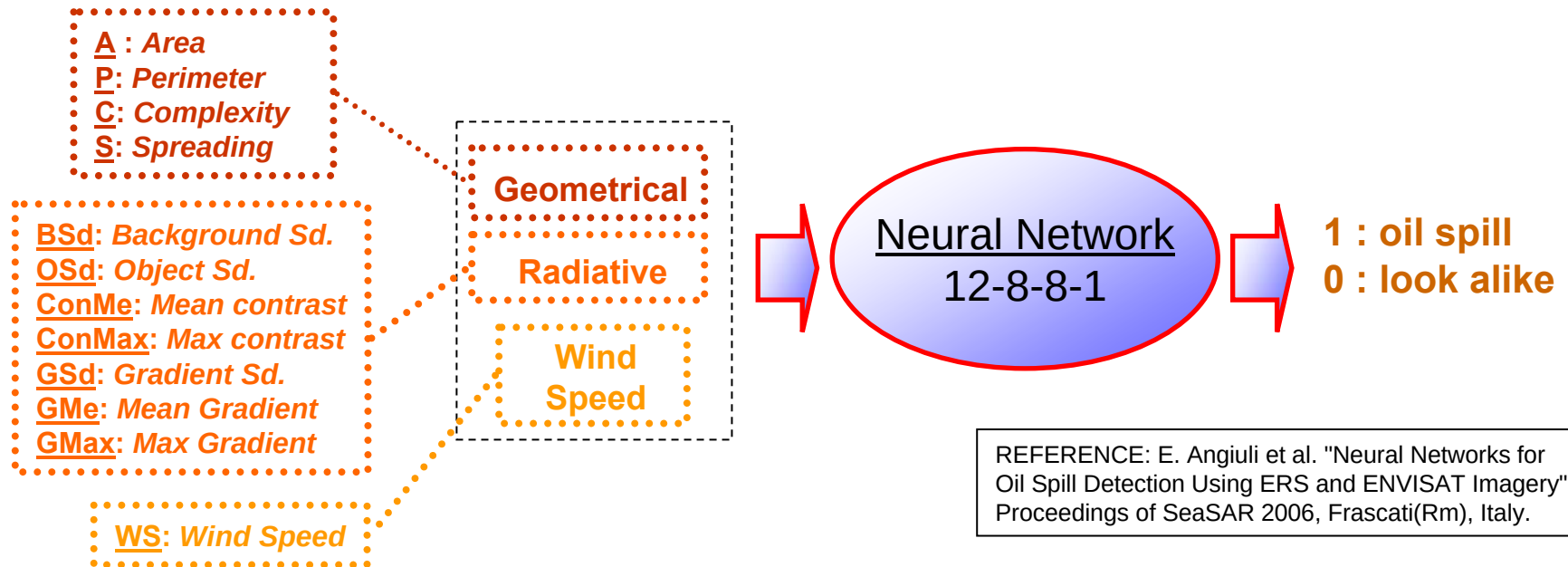


Oil spill classification by means of Neural Networks

It is necessary to extract from C band SAR images some salient features which characterize the selected dark spot ("slicks") and that will be included in the neural network input vector.

The features can be of three different types:

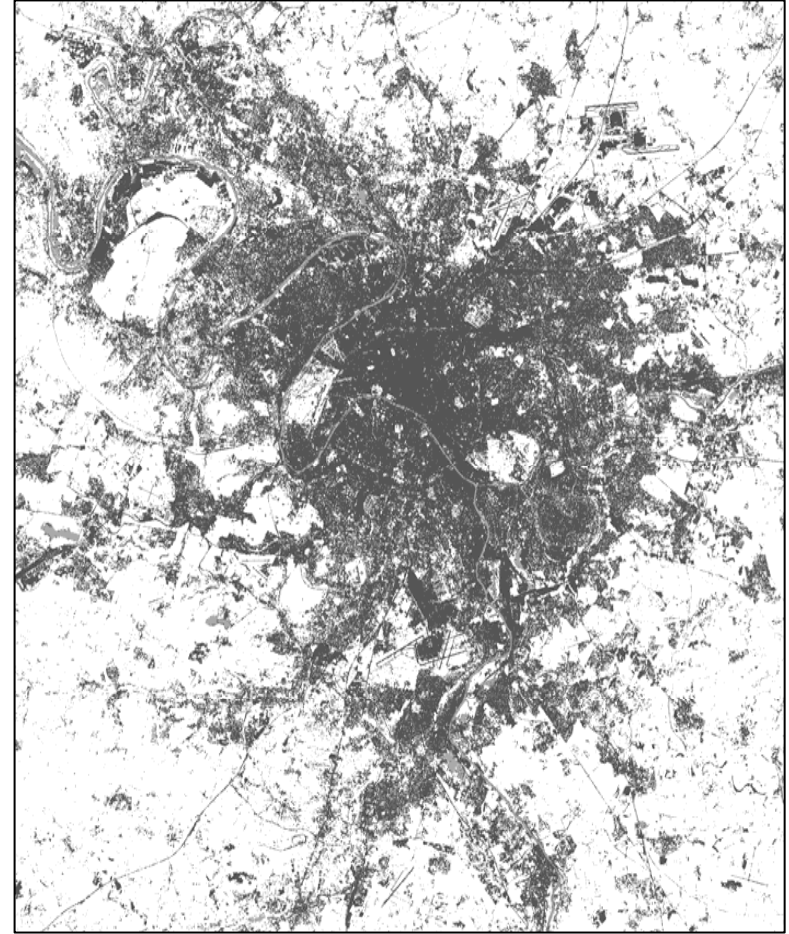
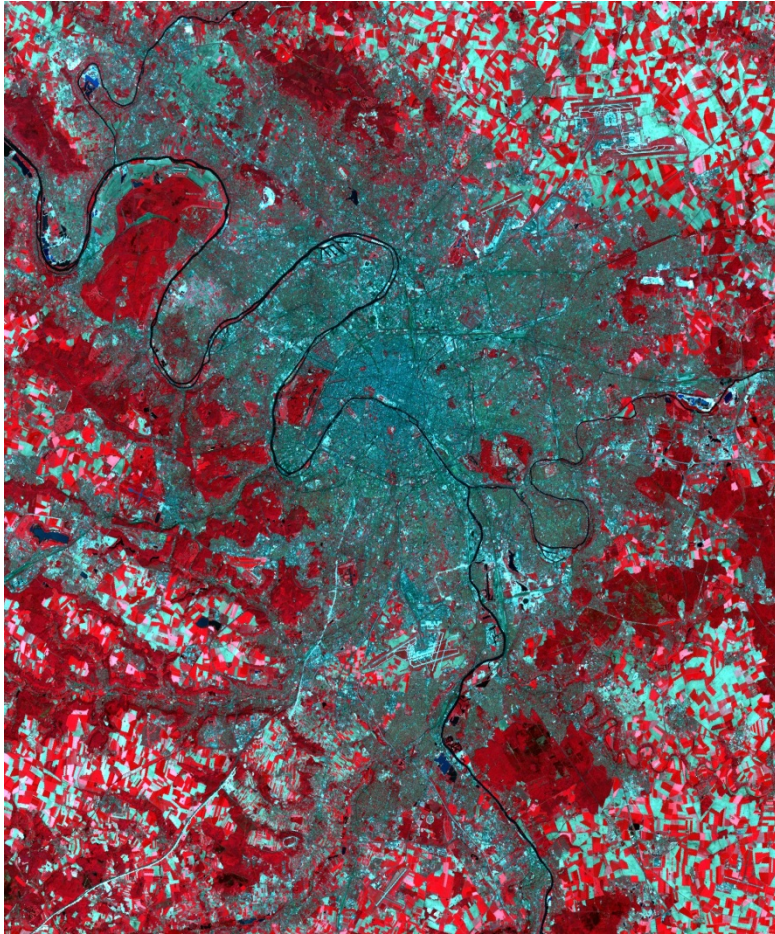
- Containing information on the backscattering intensity gradient along the border of the dark spot.
- Containing information on the backscattering intensity values in the dark spot and/or in the background.
- Containing information on the geometry and the shape of the dark spot.



REFERENCE: E. Angiuli et al. "Neural Networks for Oil Spill Detection Using ERS and ENVISAT Imagery", Proceedings of SeaSAR 2006, Frascati(Rm), Italy.

Neural Network Algorithm has led to an accuracy of 88%

Straightforward spectral data (Landsat channel 1-5 and 7)



Soil sealing mapping

Features approach (2/5)

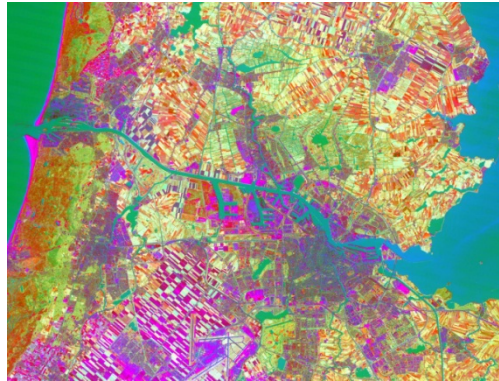
Starting from the straightforward spectral approach we tried to enlarge the number of identifiable classes, using combinations of the Landsat bands to emphasize some specific features.

This process might be thought of in three parts:

- Understanding the relationships among the spectral bands for the classes of interest;**
- Compressing the n spectral bands of information into a manageable number of features;**
- Extracting physical characteristics from these features.**

Considered features (3/5)

Principal comp.



Spectral



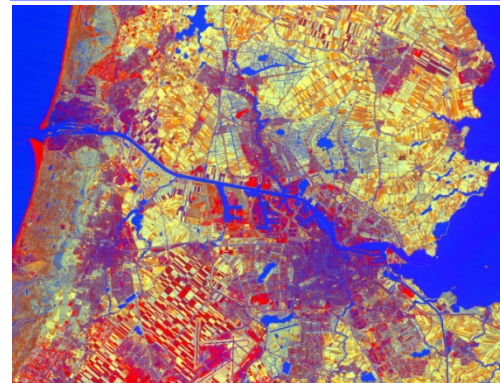
Texture



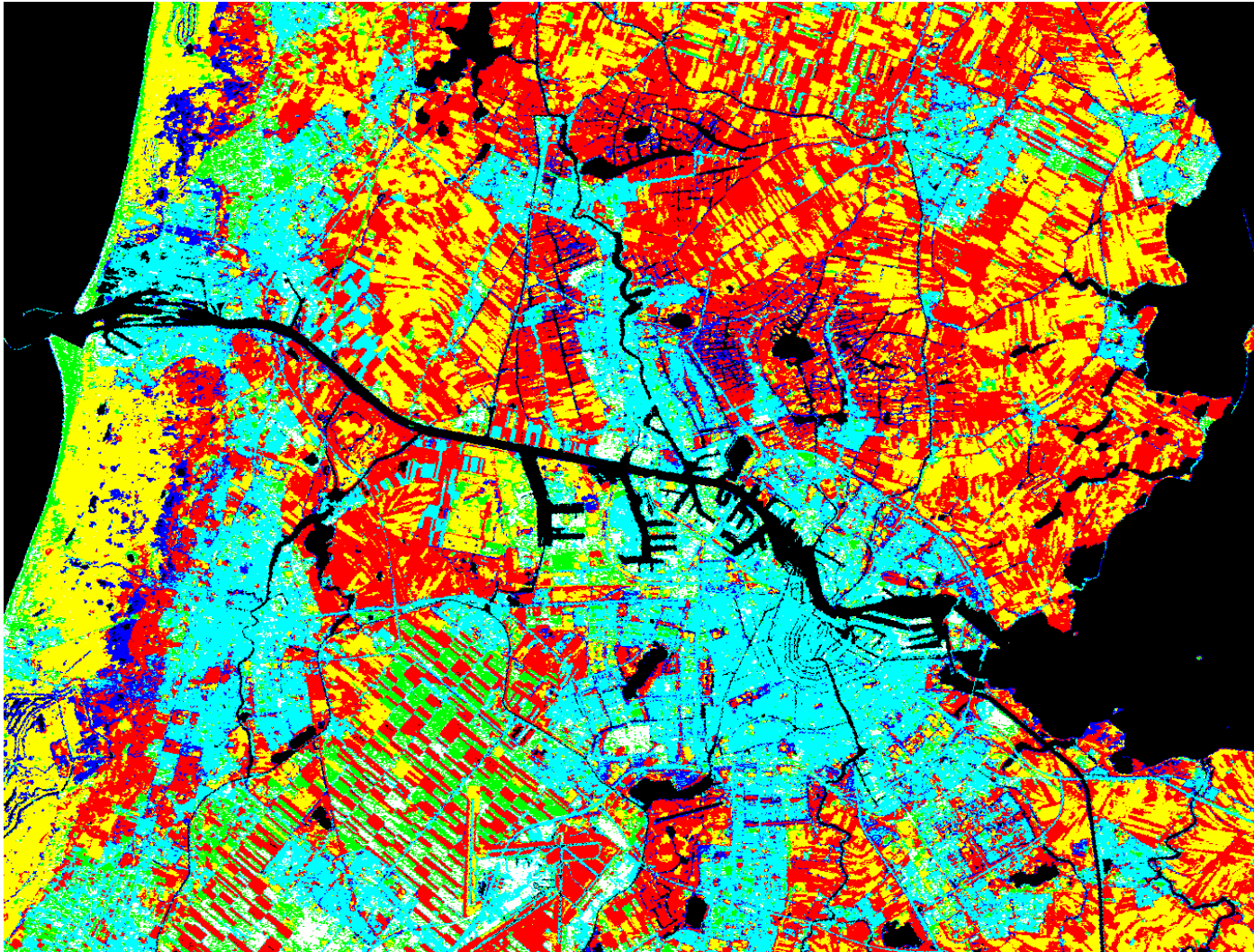
NDVI



Tasseled Cap



Results: Amsterdam TC (4/5)

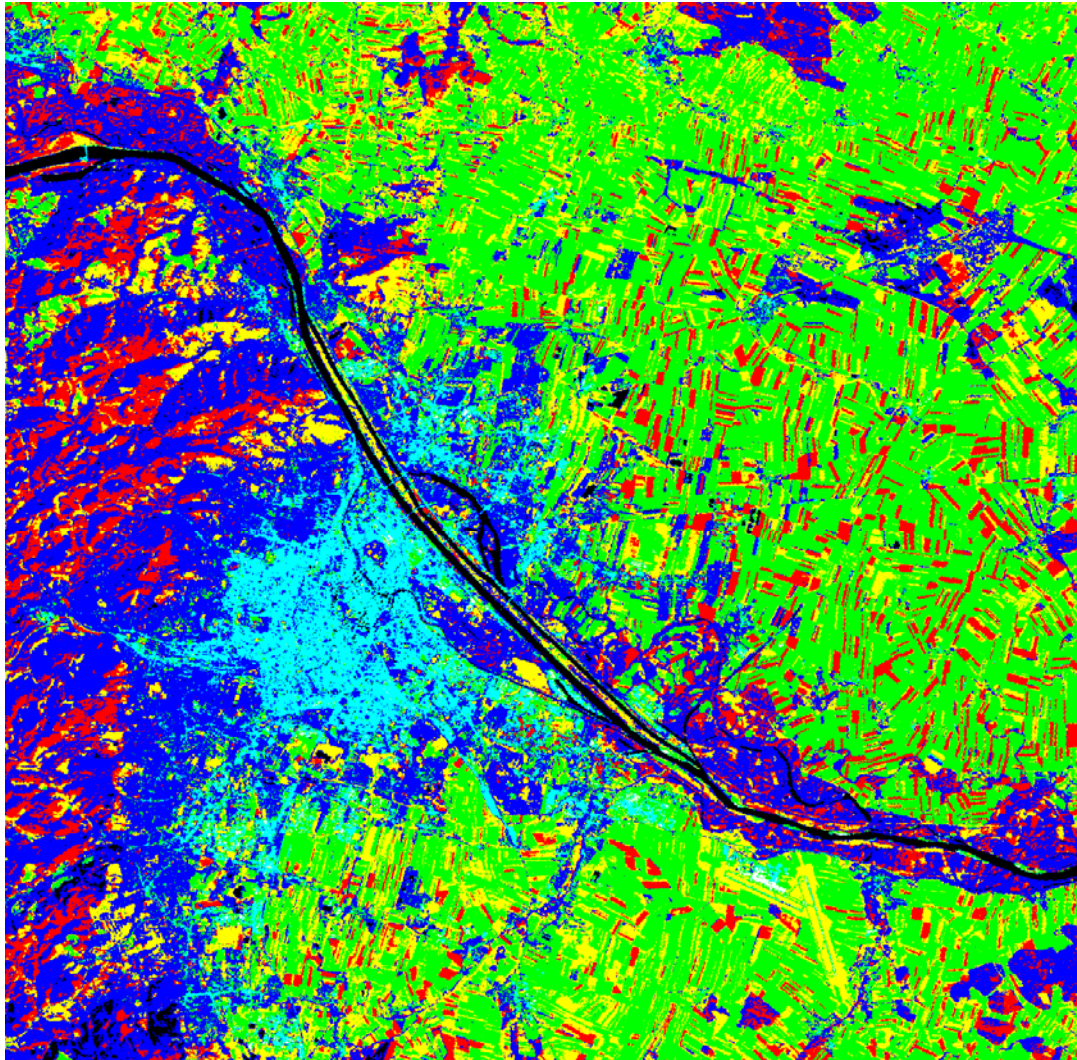


	LEG END
BARE SOIL	Green
CROPLAND	Red
GRASSLAND	Yellow
FOREST	Blue
HIGH DENSITY	White
LOW DENSITY	Cyan
WATER	Black

ACCURACY:

95%

Results: Wien TC (5/5)

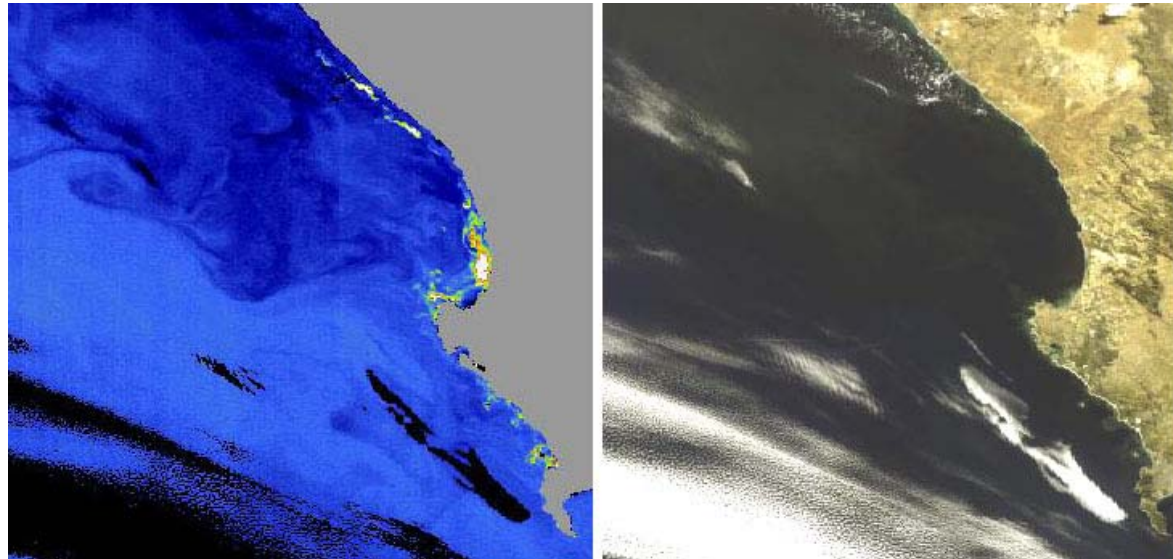


	LEG END
BARE SOIL	GREEN
CROPLAND	RED
GRASSLAND	YELLOW
FOREST	BLUE
HIGH DENSITY	WHITE
LOW DENSITY	CYAN
WATER	BLACK

ACCURACY:

92%

During the last few decades, an increasing number of harmful algal blooms (HABs) have been reported worldwide, and received a lot of attention due to their negative impact on the marine environment in which they occur. HAB events are characterised by two main features: they are caused by microalgae and they have a negative impact on human health and activities. So, because of the transient and often unpredictable nature of HABs, their improved detection and monitoring that could reduce the impact of the blooms, are technically challenging. The aim is to create an instrument supporting the monitoring of the water quality detecting the HABs in oceanic and coastal waters.

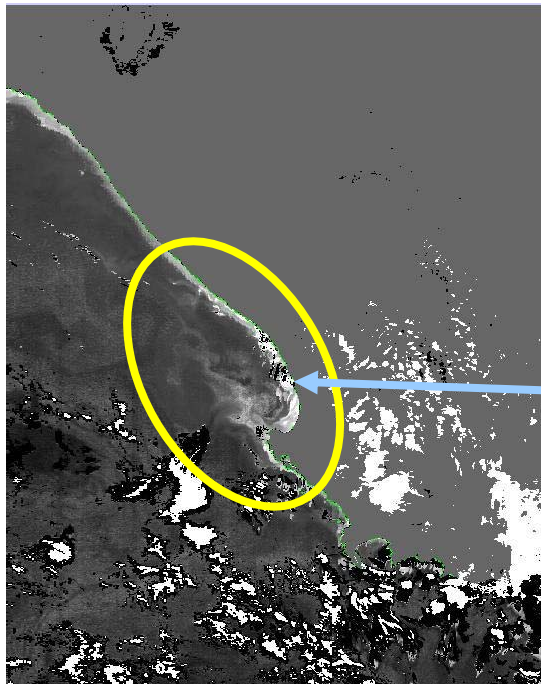


Right: MERIS image of Benguela Gulf area (SA)

Left: Corresponding HAB concentration in the same area

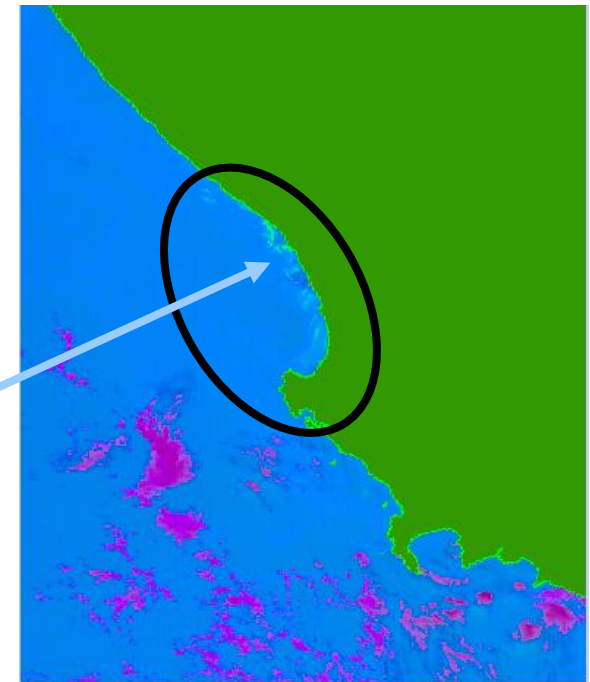
Topology: 9-32-32-1

INPUTS – MERIS bands
1-8 (620-754 nm)



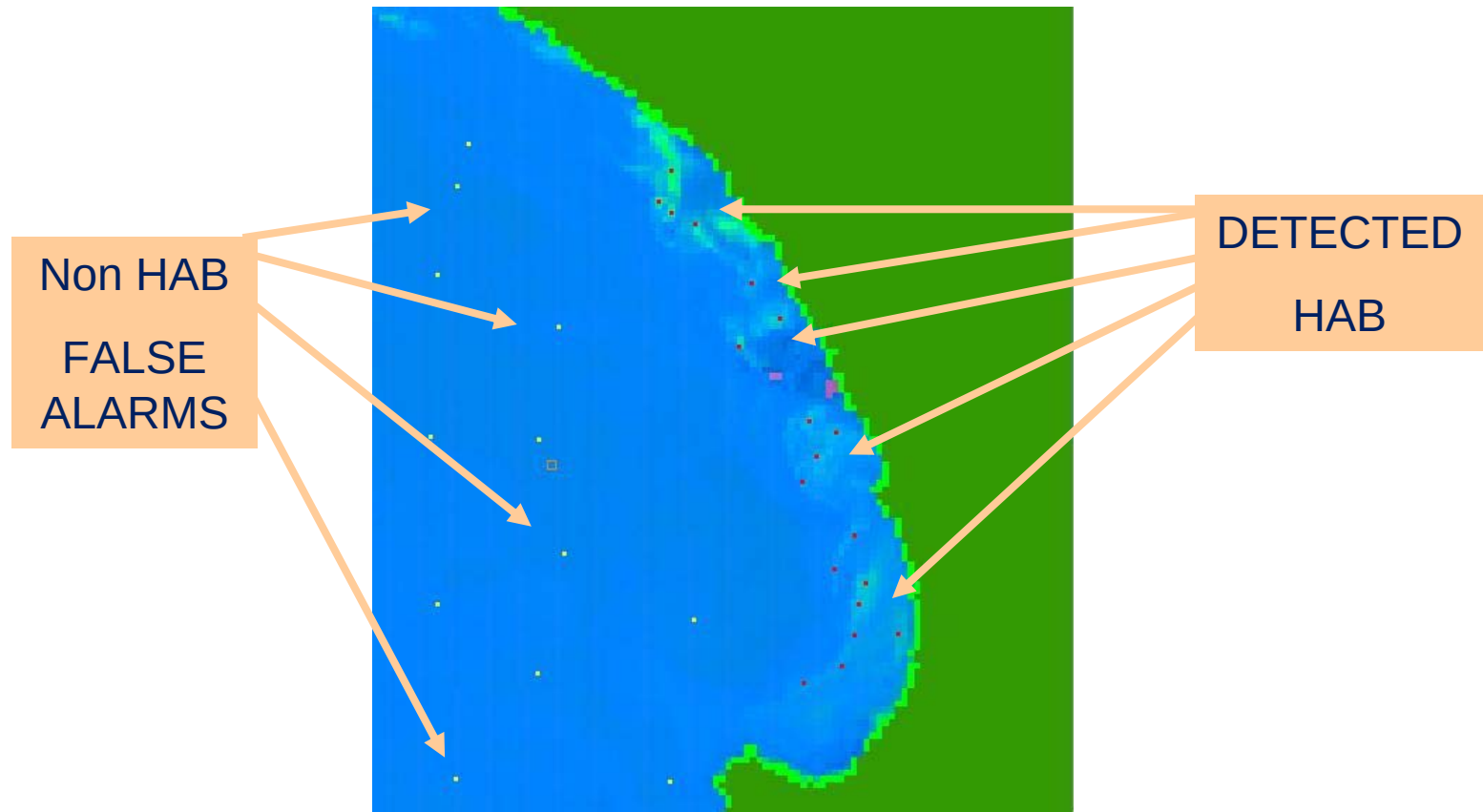
MERIS L2 product –
Chlorophyll

HABs
AREA



FLH

RESULTS: comparison
with ancillary data (buoys)



- Le NN sono uno strumento utile al telerilevamento e alla geoinformazione, alternativo ai modelli classici di inversione e classificazione – potente quando non si ha la necessità di conoscere nei dettagli il sistema studiato;
- a volte permettono di avere informazioni sulle relazioni funzionali tra grandezze o sulla fisica dei processi in studio;
- alcune applicazioni mostrano la potenza di questi algoritmi;
- **non sono una scatola magica**: per molte applicazioni rimane preferibile usare i modelli non-neurali!

GRAZIE PER L'ATTENZIONE



www.disp.uniroma2.it/eolab

sellitto@disp.uniroma2.it