

Statistical Assessment of Radiometric Measurements From Autonomous Systems

Davide D'Alimonte and Giuseppe Zibordi

Abstract—*In situ* autonomous systems are commonly used for the collection of measurements for the vicarious calibration of satellite data and the successive validation of derived products. However, the use of autonomous systems creates the need of assessing the quality of the large volume of collected data. Within the framework of ocean color activities, this work investigates the consistency of normalized water leaving radiance spectra produced from measurements taken with an above-water autonomous system installed on an oceanographic tower. The study has shown the need of addressing the problem under two different levels of inference. The first level, so-called *self-consistency*, has demonstrated the capability of identifying spectra with a low statistical representativeness within the dataset itself. The second level, so-called *relative-consistency*, has provided the possibility of evaluating whether a spectrum is relatively consistent to a reference set of quality-assured data.

Index Terms—Neural network, ocean color, water leaving radiance.

I. INTRODUCTION

IN *SITU* autonomous systems are commonly used for the collection of a large amount of measurements to support Earth observation programs. Within the framework of ocean color activities, autonomous *in situ* measurements for the calibration and validation (cal/val) of space data can be produced with in-water radiometers operated on buoys, or with above-water autonomous radiometers deployed on platforms like oceanographic towers, large buoys, or ships. Measurement sites characterized by almost stable environmental conditions are particularly suitable for the vicarious calibration of space sensors, while sites representing different environmental regimes are more appropriate for the validation of space derived products. In both cases, autonomous systems open the issue of how assessing the quality of the large volume of collected data.

This challenging problem is here investigated through the implementation of advanced statistical methods and their successive application to an existing set of radiometric data produced with an above-water autonomous system installed on an oceanographic tower. To ensure the quality of these measurements collected in daylight conditions with time intervals which may vary from a few minutes up to tens of minutes, it is necessary to detect and remove data affected by cloud perturbations, measurement artifacts, and poor instrument performance. The consistency of radiometric data is here addressed under two different levels of

inference. The first level, so-called *self-consistency*, evaluates whether a radiance spectrum is anomalous with respect to all the available spectra. The second level, so-called *relative-consistency*, evaluates whether a spectrum is consistent *relatively* to a reference set of independent and quality-assured spectra.

The *self-consistency* analysis is undertaken exploiting a *recursive* modeling of radiance spectra. For this purpose, an *auto-associative* neural network (NN) is trained using the same spectra as input and output. This study introduces a new application of the auto-associative NN already used as a feature extraction procedure [1], [2], and in checking the efficiency of sensors in engine control systems [3]. As for any function regression model, the effectiveness of the auto-associative NN depends on the density of the data. Thus, to increase the effectiveness of the whole *self-consistency* scheme, the auto-associative NN is complemented with an estimate of the data density to detect anomalous spectra that may not be identified with a function regression approach.

The *relative-consistency* analysis is undertaken with a *novelty detection* scheme [4] by modeling the distribution of the quality-assured spectra projected onto the two-dimensional (2-D) space generated with the auto-associative NN. This feature extraction approach has been adopted by observing that the auto-associative NN can be viewed as a non-linear generalization of a principal component analysis [5]–[7] which was shown to explain most of the variance of the seawater radiometric spectra through the first two principal components [8].

II. DATASETS

The *normalized water leaving radiance*, L_{wn} , (i.e., the radiance leaving the sea surface and normalized with respect to the downward irradiance [9]) is a fundamental radiometric quantity for ocean color cal/val activities. In this study, the methods for the quality assurance of L_{wn} spectra are applied on the basis of two datasets. The first consists of above-water measurements performed with two Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Photometer Revision for Incident Surface Measurement (SeaPRISM) autonomous systems sequentially operated in a three-year period. The second dataset results from in-water optical profiles individually executed on ideal measurement conditions (low cloudiness and sun not perturbed by clouds) with the Winched Stabilized Profiling Environmental Radiometer (WiSPER). Both datasets were produced in the northern Adriatic Sea at the Acqua Alta Oceanographic Tower (AAOT) and were extensively used to support ocean color development and validation activities within the framework of the Coastal Atmosphere and Sea Time Series (CoASTS) program [10]. These data benefit from the unique condition offered by the measurement site located eight nautical miles off the Venice Lagoon (45° 19' N, 12° 30' E) of representing both

Manuscript received April 7, 2005; revised August 3, 2005.

D. D'Alimonte is with the Neural Computing Research Group, School of Engineering and Applied Science, Aston University, Birmingham, B4 7ET U.K. (e-mail: d.dalimonte@aston.ac.uk).

G. Zibordi is with the Inland and Marine Waters Unit, Institute for Environment and Sustainability, Joint Research Centre, 21020 Ispra (VA), Italy (e-mail: giuseppe.zibordi@jrc.it).

Digital Object Identifier 10.1109/TGRS.2005.862505

Case-1 and Case-2 water types with approximately one third of the observations pertaining to Case-2 [11].

A. Radiometric Data From Autonomous Systems

The L_{wn} spectra from the autonomous above-water radiometer were produced with SeaPRISMs [manufactured by CIMEL (Paris, France)] included in the Aerosol Robotic Network (AERONET) of sun photometers [12]. The latter guarantees support from the AERONET program for real-time data handling in addition to regular system calibration. The SeaPRISM measurements are taken with a full angle field of view of 1.5° in eight spectral bands at center-wavelengths relevant for the atmospheric aerosol and water vapor monitoring, and for ocean color applications (i.e., 412, 440, 500, 555, 675, 870, 940, 1020 nm). With the exception of the water vapor channel centered at 940 nm, all bands are 10 nm wide.

The SeaPRISM sea-viewing scenario produces the data required for determining L_{wn} through sequential and repeated measurements of sea-radiance at viewing angle $\theta = 40^\circ$ from nadir and relative azimuth angle $\phi = 90^\circ$ with respect to the sun azimuth, and sky-radiance at viewing angle $\theta' = 140^\circ$ (with $\theta' = 180^\circ - \theta$) and relative azimuth angle $\phi = 90^\circ$. The SeaPRISMs operated at the AAOT were programmed to perform measurement sequences every 30 min during daylight conditions. Each sequence is comprising of 11 sea- and 3 sky-radiance observations per spectral channel and lasts approximately 6 min. An assessment of the accuracy of the SeaPRISM data, together with the description of the data processing used for producing L_{wn} , was published elsewhere using data collected under various environmental conditions [13]. The expected uncertainty in SeaPRISM L_{wn} is 5% from 412–555 nm, and 12% at 675 nm [14].

The SeaPRISM data used in this study were processed with the AERONET code implemented in agreement with the published method [13]. This allowed for a basic quality assurance including: 1) removal of measurement sequences with incomplete data records; 2) filtering out measurement sequences characterized by high variability in repeated sea and sky measurements; and 3) cloud screening according to a scheme exclusively based on sun-photometric measurements (i.e., cloud-perturbed observations are determined using triplets of direct sun irradiance measurements and daily occurrence of aerosol optical thickness retrievals [15]). However, this basic protocol may fail to detect spectral anomalies due to sparse clouds, the presence of occasional obstacles in the proximity of the instrument, or the poor performance of the radiometer in some channels.

B. Reference Set of Quality Assured Data

Quality-assured L_{wn} spectra were derived from in-water profiles collected with the WiSPER system. The WiSPER upwelling radiance profiles are taken in seven spectral bands 10 nm wide at center-wavelengths relevant for ocean color applications (i.e., 412, 443, 490, 510, 555, 665, and 685 nm). The WiSPER radiance sensor, an OCR-200 radiometer manufactured by Satlantic, Inc. (Halifax, Canada), is installed on a custom-built profiling rig whose rigidity and stability is maintained by two taut wires anchored between the deployment platform of the AAOT and a weight on the sea bottom. Above-water downward irradiance data, collected at the same time as the in-water radiometric data, are used to correct these latter for illumination changes during casts. This correction

is performed by normalizing the in-water data with respect to the time-correspondent above-water downward irradiance, and multiplying the resulting values by the downward irradiance at a reference time generally coinciding with the start of the cast. The overall processing applied to produce L_{wn} spectra from WiSPER profile data relied on assessed protocols [13]. The expected uncertainty in WiSPER L_{wn} data is 5% [14]. The set of quality-assured WiSPER data is composed of 244 spectra collected in the period January 1999–February 2001 (the same dataset was also used to investigate optical profiling protocols in coastal waters [16]). An independent study [13] has shown that time-coincident above- and in-water L_{wn} spectra produced with the SeaPRISM and WiSPER systems, exhibit spectrally averaged percent differences within 5%.

III. METHODS

The spectral measurements used in this analysis are pre-processed, and afterward their consistency is assessed by identifying anomalous spectra without accounting for any extra information obtainable from the quality-assured data (*self-consistency*). A novelty detection approach is also investigated on the basis of statistical information derived from the reference set of quality-assured data (*relative-consistency*).

A. Data Preprocessing

The study is undertaken considering SeaPRISM L_{wn} data at the center-wavelengths 412, 440, 490, 555, 675 nm and WiSPER L_{wn} data at 412, 443, 490, 555, 665 nm. The L_{wn} value at 490 nm, not directly measured by the current series of SeaPRISMs, is determined by linearly interpolating L_{wn} at 440 and 500 nm. The L_{wn} spectra are then rescaled with respect to the value measured at 555 nm. The choice of 555 nm as normalizing wavelength is mostly supported by its use, in combination with different center-wavelengths [17], [18], to discriminate optically significant seawater components.

The appropriateness of interpolating L_{wn} at 490 nm is supported by an analysis of SeaPRISM and WiSPER band ratios between 490 and 555 nm, which has shown an average difference of 3.9% [14]. This difference is comparable to the uncertainty induced in SeaPRISM L_{wn} data by environmental perturbations (as determined on the basis of a 3% uncertainty independently affecting the L_{wn} values at 490 and 555 nm [13]).

The effects of slight shifts in SeaPRISM and WiSPER L_{wn} center-wavelengths (i.e., 440 versus 443 nm and 675 versus 665 nm, respectively) are neglected. This is supported by comparisons between SeaPRISM and WiSPER L_{wn} exhibiting average differences generally lower than their respective uncertainties [13].

Tables I and II summarize the statistics of L_{wn} spectra produced with the SeaPRISM systems identified by their AERONET ID number (i.e., 139 and 176). Table III refers to the set of quality-assured L_{wn} spectra produced with the WiSPER system.

B. Self-Consistency Classification Scheme

The *self-consistency* approach relies on the hypothesis that the autonomous sensor is mostly well performing and anomalous spectra are occasional. Let us indicate the L_{wn} variable with x ; the true, but unknown, L_{wn} with \tilde{x} ; and the corresponding actual (i.e., measured) L_{wn} with \hat{x} . Now, assume that

TABLE I
MINIMUM, MAXIMUM, MEAN, AND STANDARD DEVIATION (MIN, MAX, μ , AND σ , RESPECTIVELY) OF THE 951 L_{wn} SPECTRA PRODUCED WITH SEAPRISM-139 SYSTEM (L_{wn} DATA ARE IN UNITS OF MILLIWATTS PER SQUARE CENTIMETER PER MICRON PER STERADIAN)

Wavelength [nm]	412	440	490	555	675
Min	0.18	0.31	0.53	0.20	-0.05
Max	2.55	3.12	4.28	3.95	0.98
μ	0.80	0.96	1.30	1.13	0.16
σ	0.31	0.40	0.58	0.60	0.12

the actual L_{wn} is affected by a Gaussian additive noise, δ , with **O** mean and spherical covariance $\Sigma = \sigma^2 \mathbf{I}$

$$\hat{\mathbf{x}} = \tilde{\mathbf{x}} + \delta \quad (1)$$

$$\delta \in \mathcal{N}(\mathbf{0}, \Sigma) \quad (2)$$

where \mathbf{I} is the identity matrix and \mathcal{N} indicates the multivariate normal distribution. The assumption of additive noise with zero mean and normal distribution is supported by the following elements.

- 1) The uncertainty affecting the SeaPRISM and WiSPER radiometric data investigated in this work has three main sources: absolute calibration, environmental variability, and corrections for off-nadir view (specific for SeaPRISM data) and for shading perturbations (specific for WiSPER data) [13]. The uncertainty affecting the absolute calibration coefficients is a multiplicative factor. The environmental effects are a source of additive noise. The uncertainty produced by the off-nadir and shading corrections is a composition of both multiplicative and additive contributions to the measurement uncertainty. An assumption of overall additive noise was thus made considering that an error budget estimate has shown the environmental effects as the most significant noise source for both the SeaPRISM and WiSPER measurements [13].
- 2) Several and independent factors concur to generate the environmental noise (i.e., wave perturbations, changes in the water column and in the illumination conditions). Although these factors may not independently produce a normal distribution of uncertainties, their joint effect tends to be normal for the central limit theorem.¹
- 3) Finally, the comparison of SeaPRISM and WiSPER L_{wn} has not shown the existence of any relevant bias (i.e., above the expected radiometric uncertainties [13]) between the two.

By indicating with $p(\mathbf{x} | \hat{\mathbf{x}})$ the conditional probability density function and with E the expectation, it is natural to express the unknown $\tilde{\mathbf{x}}$ as

$$\begin{aligned} \tilde{\mathbf{x}} &= E[\mathbf{x} | \hat{\mathbf{x}}] \\ &= \int \mathbf{x} p(\mathbf{x} | \hat{\mathbf{x}}) d\mathbf{x}. \end{aligned} \quad (3)$$

¹The central limit theorem can be summarized as follows: "The distribution of a sum tends to be normal, even when the distribution of the data from which the sum is computed is non-normal."

TABLE II
AS IN TABLE I, BUT FOR THE 1424 L_{wn} SPECTRA PRODUCED WITH SEAPRISM-176 SYSTEM

Wavelength [nm]	412	440	490	555	675
Min	-0.05	0.09	0.31	0.20	-0.12
Max	2.67	3.43	4.08	3.78	0.75
μ	0.83	1.02	1.37	1.11	0.14
σ	0.32	0.43	0.58	0.57	0.11

TABLE III
AS IN TABLE I, BUT FOR THE 244 L_{wn} SPECTRA PRODUCED WITH THE WISPER SYSTEM

Wavelength [nm]	412	443	490	555	665
Min	0.42	0.54	0.68	0.47	0.04
Max	1.97	2.72	3.57	2.99	0.85
μ	0.84	1.09	1.45	1.18	0.17
σ	0.32	0.47	0.63	0.59	0.14

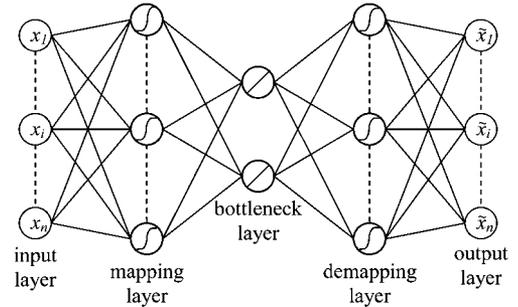


Fig. 1. Schematic representation of the auto-associative NN. The bottleneck layer, made by two nodes, has been used for the reduced dimensional representation of the input L_{wn} spectra.

Equation (3) also enables to estimate σ^2 from the experimental data

$$\sigma^2 \approx \frac{1}{MN} \sum_{i=1}^N \|\hat{\mathbf{x}}_i - E[\mathbf{x} | \hat{\mathbf{x}}_i]\|^2 \quad (4)$$

where N is the number of samples in the dataset and M is the number of input wavelengths. This study relies on the auto-associative regression model described in Section III-B to estimate $E[\mathbf{x} | \hat{\mathbf{x}}]$.

It is stressed that the effectiveness of a regression approach depends on the data density: the estimation of the conditional expectation becomes less accurate for isolated or anomalous spectra which may not be identified by the auto-associative model. To address this issue, the present work combines the data regression with a non-parametric estimation of the data density based on the *k*-nearest-neighbors algorithm (KNN), presented in Section III-B2. Accounting that function regression and density estimation complement each other, their combined use improves the performance of the *self-consistency* approach.

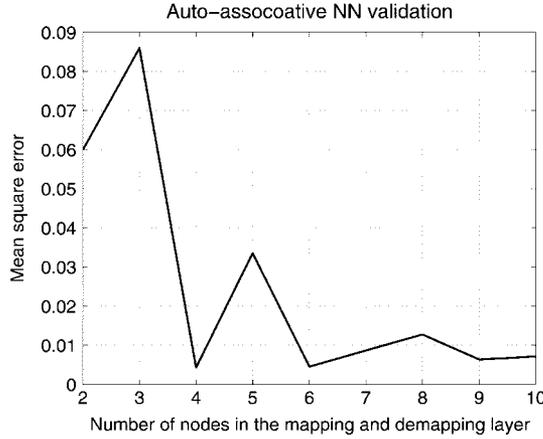


Fig. 2. Dependence of the model performance on the architecture of the auto-associative NN (number of nodes in the mapping and demapping layers). For each configuration, the auto-associative NN has been trained five times, each time restarting the training process. The effect of each model architecture is here summarized averaging the performance of the auto-associative NN with respect to the validation data.

1) *Auto-Associative Neural Network*: The five-layer feed-forward auto-associative network is illustrated in Fig. 1. Specifically:

- 1) The first layer is the auto-associative model input, i.e., the actual L_{wn} spectrum; the fifth layer is the corresponding conditional expectation.
- 2) The second and the fourth layers are called *mapping* and *demapping* layers, respectively. Both these layers have an arctangent transfer function that determines the non-linear nature of the model.
- 3) The third layer is referred to as *bottleneck*, and is made by a smaller number of nodes with respect to the model input. The transfer function of the third layer is the identity function.

The auto-associative NN model has been implemented using the neural network toolbox of MATLAB. The model training is performed with the Levenberg–Marquard algorithm to set the parameters of the NN minimizing the sum of squared errors (SSE) between the actual L_{wn} and the corresponding model output. Half available spectra (i.e., 1309 out of 2618 including data from both quality-assured and autonomous system measurements) were randomly sampled for the model training, the remaining half for the validation. Testing models with two to ten nodes in the mapping and demapping layers allowed to select the architecture with six nodes (see Fig. 2).

Notice that the relatively large number of model parameters (i.e., $90 = 5 * 6 + 7 * 2 + 3 * 6 + 7 * 4$) is compensated by the size of the training dataset, and this reduces the risk of overfitting. It was found that outliers and local minima of the learning cost function can occasionally affect the representation of spectra at the bottleneck layer (see Section IV). For this reason, once the number of nodes has been defined, an additional set of model training (restarting the training process and resampling the training data) has been used to set the final auto-associative NN for the operational classification task. Fig. 3 shows the corresponding data partitioning into training and validation sets, and the resulting model generalization capability.

Each actual L_{wn} is classified on the basis of the squared difference between itself and the corresponding NN output. When

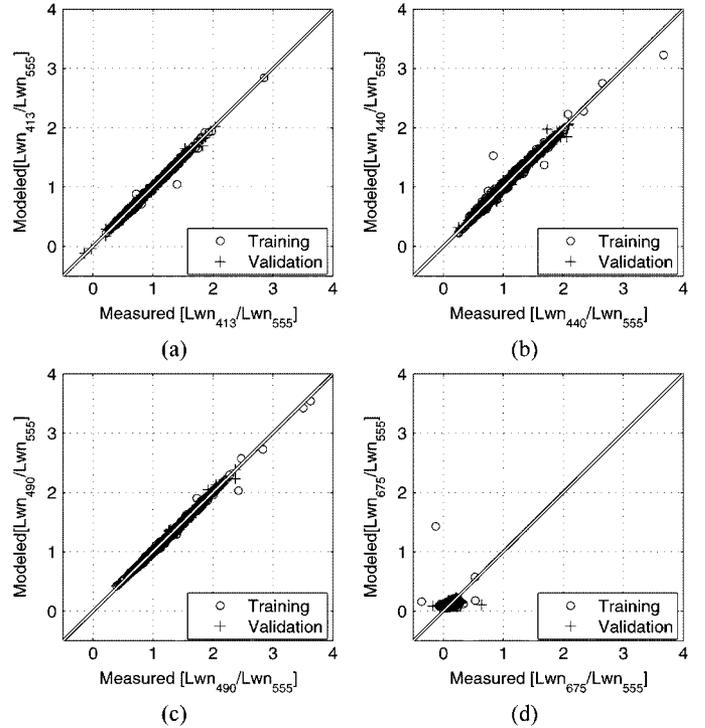


Fig. 3. Scatter plot of the measured L_{wn} band ratios versus the corresponding values at the output of the auto-associative NN. The mean square error of training and validation data are 0.0051 and 0.0024, respectively. This result supports the validity of the architecture selected. In fact, although in this case a relevant part of the anomalous spectra was randomly included in the training set, the learning process of the auto-associative NN was not driven by these data.

this difference overcomes significantly σ^2 [see (4)], it means that the input spectrum presents some anomaly. Notice that this approach may also allow for detecting anomalies at specific individual center-wavelengths, or set of center-wavelengths. Finally, it is outlined that the auto-associative NN learns to represent expected spectra from the training data. Thus, the presence of spectra affected by systematic errors in the training dataset would compromise the identification of anomalous spectra affected by the same systematic errors.

2) *K-Nearest-Neighbors Density Estimation*: The regression effectiveness of the auto-associative NN depends on the data density. In order to identify anomalous spectra untraceable with the auto-associative NN, the *K-nearest-neighbors* algorithm is additionally used.

To illustrate the *K-nearest-neighbors* algorithm [19], consider that the probability, P , for a data point \mathbf{x} to fall in a region \mathcal{R} is

$$P = \int_{\mathcal{R}} p(\mathbf{x}') d\mathbf{x}' \quad (5)$$

where $p(\mathbf{x})$ is the probability density function. So, if N data points are independently drawn from $p(\mathbf{x})$, the expected number of points falling in \mathcal{R} is $K = N \cdot P$. Assuming that $p(\mathbf{x})$ does not vary significantly over the region \mathcal{R} , and indicating with V the volume of \mathcal{R} , it follows that

$$p(\mathbf{x}) \simeq \frac{K}{NV}. \quad (6)$$

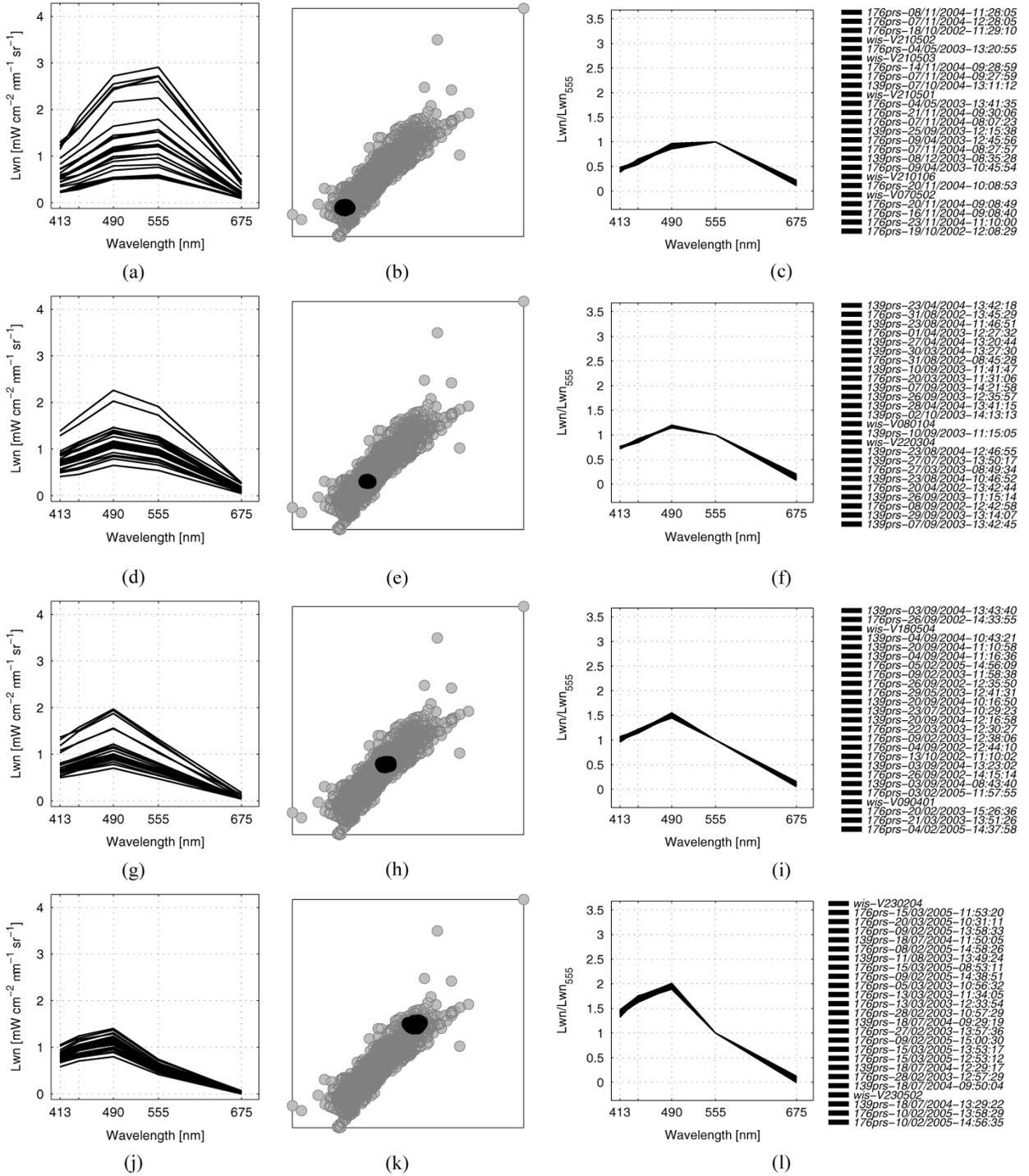


Fig. 4. Two-dimensional latent representation of L_{wn} spectra through the bottleneck of the auto-associative NN. The panels in the first row illustrate a set of points chosen in the bottom left region of the 2-D latent space (central panel), together with the corresponding radiometric data in physical units (left panel) and rescaled values (right panel). The same scheme is used in the subsequent rows of panels to show the topographic nature of this 2-D representation.

The *K-nearest-neighbors* algorithm used in this study identifies the region \mathcal{R} with a hypersphere centered in \mathbf{x} . Initially the radius of the hypersphere is zero and $p(\mathbf{x})$ is estimated from (6) increasing the radius until k points are found. L_{wn} spectra are classified as *self-inconsistent* when the radius of the hypersphere containing the closest k spectra overcomes a threshold.

C. Relative-Consistency Classification Scheme

The distribution of the quality-assured data is modeled with a Gaussian mixture model (GMM) on the basis of the features extracted from the original spectra by the bottleneck of the auto-associative NN. *Novel* spectra are those with a low probability of being represented within the quality-assured data [20], and are here classified as *relative-inconsistent* (these would include those spectra not detected by the auto-associative NN because affected by systematic errors).

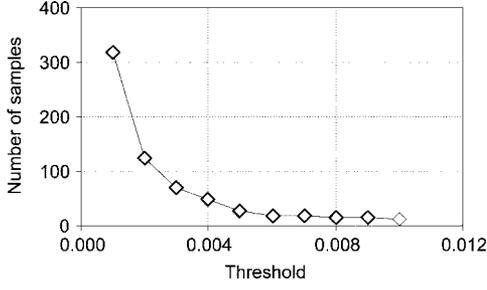


Fig. 5. Sensitivity of the *self-consistency* scheme to the threshold t of the auto-associative NN. For the purpose of this study an empirical value $t = 0.006$ has been used, although a smaller value could be required for specific applications.

1) *Gaussian Mixture Model*: The GMM represents complex data distributions through a linear combination of simpler Gaussian density functions

$$p(\mathbf{x}) = \sum_k P(k)p(\mathbf{x} | k) \quad (7)$$

where $P(k)$ is the *mixing coefficients* of the k th *kernel*

$$p(\mathbf{x} | k) = \frac{1}{(2\pi)^{d/2} |\Sigma_k|^{1/2}}$$

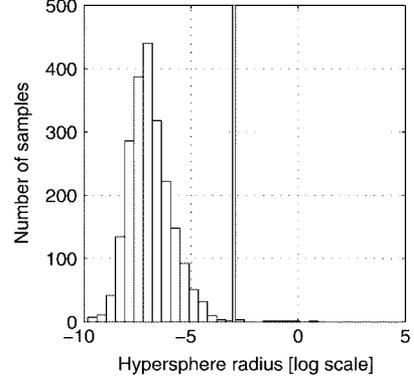


Fig. 6. Distribution of the hypersphere radius with $k = 3$ used for the KNN identification of *self-inconsistent* spectra. The vertical line corresponds to the threshold radius of 0.04 chosen to identify anomalous spectra.

$$\times \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k) \right\} \quad (8)$$

with $\boldsymbol{\mu}_k$ the center and Σ_k the covariance matrix of the kernel, d the dataset dimensionality (for the application considered in this study, $d = 2$), and T indicates the matrix transpose. Although the kernel functions are quite simple, their combination in a mixture representation allows for the modeling of any continuous distribution [19]. The GMM can be efficiently trained through the expectation–maximization (EM) algorithm [21]. The NETLAB toolbox [22] for MATLAB has been used for the GMM numerical implementation.

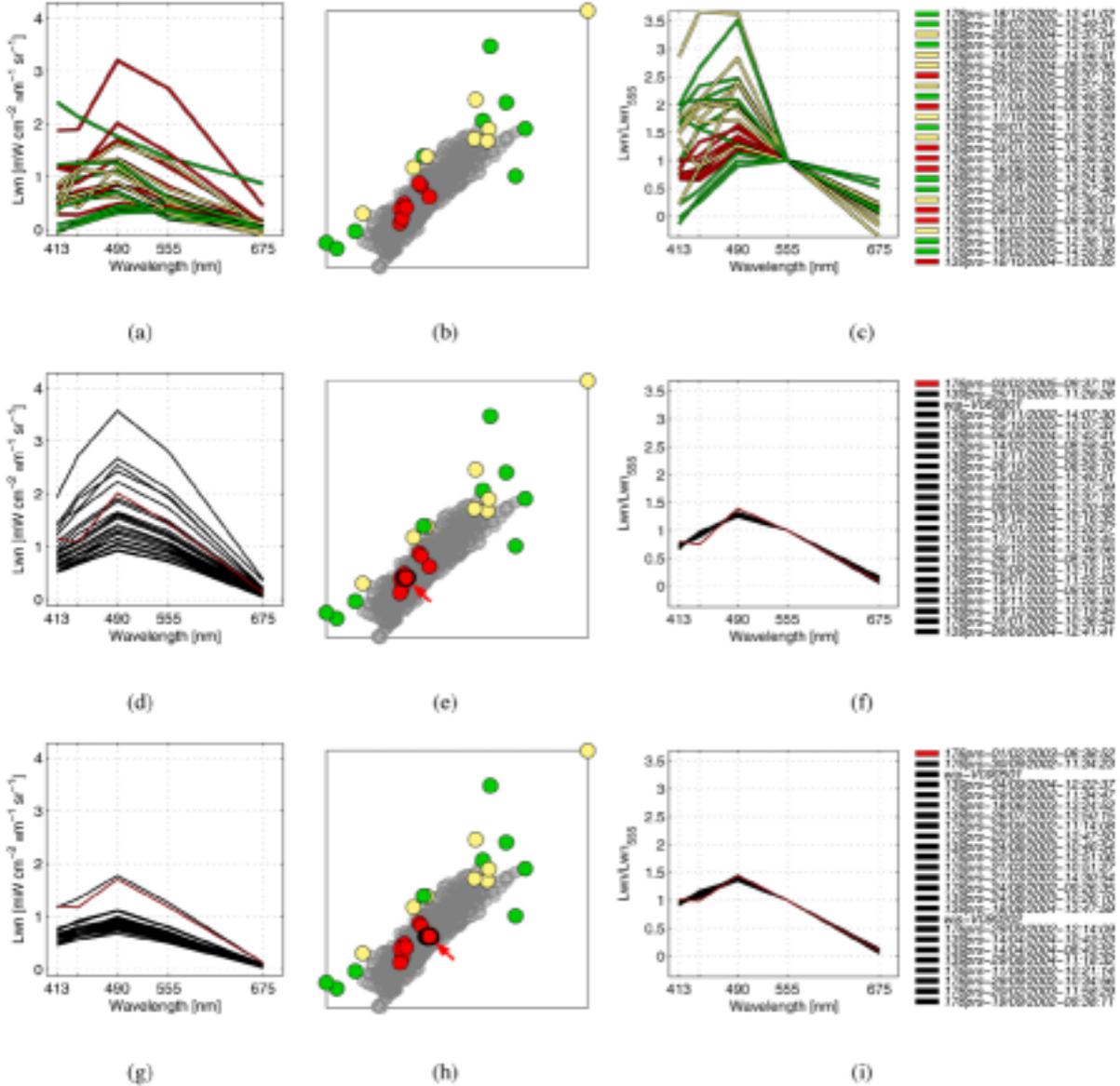


Fig. 7. *Self-inconsistent* L_{wn} spectra are highlighted in the panels of the top row. Red points correspond to spectra that overcome the threshold of the auto-associative analysis. Green points refer to spectra detected by the KNN scheme. When the same spectrum violates both criteria, the corresponding latent point is colored in yellow. The panels in the second and third row highlight the effectiveness of the auto-associative approach in detecting minor spectral anomalies.

IV. RESULTS AND DISCUSSION

The panels of the central column of Fig. 4 show the projection of the L_{wn} spectra onto the 2-D *latent plane*² generated by the bottleneck layer of the auto-associative NN (i.e., each neuron of the bottleneck layer corresponds to an axis of this plane). Four sets of points are identified in different regions of the latent space. Each set corresponds to a different row of panels: selected points are highlighted in black in the central panel, rescaled spectra projected onto these points are in the right panel; corresponding spectra in physical units are in the left panel. The panels of the first row illustrate a set of points chosen in the bottom left region of the 2-D latent space. The L_{wn} spectra cor-

²The projection of L_{wn} spectra onto the latent space is shown without providing axis labels and ranges because of the arbitrary rescaling performed in the bottleneck layer of the auto-associative NN. A similar scheme is also applied visualizing data features extracted with other models (see for instance NeuroScale [23], Generative Topographic Mapping [24], and Self Organizing Map [25]).

responding to these data points are characterized by a maximum at 555 nm. The subsequent rows of panels show how the spectral properties change across the data distribution (i.e., from the bottom left to the top right of the latent plane). It can be observed that data points in the central region of the latent plane are related to spectra with maximum at 490 nm; while points in the top left region are linked to spectra with a relative minimum and maximum at 440 and 555 nm, respectively. In all cases, points that are close in the latent space correspond to specific spectral patterns and refer to different bio-optical conditions characterized by an increasing concentration of optically significant constituents (mostly Chlorophyll *a*) from the bottom to the top case, respectively.

A. Self-Consistency

The spectral *self-consistency* is assessed: 1) applying a threshold t to the difference $1/M \|\hat{x} - E[x | \hat{x}]\|^2$ with M the

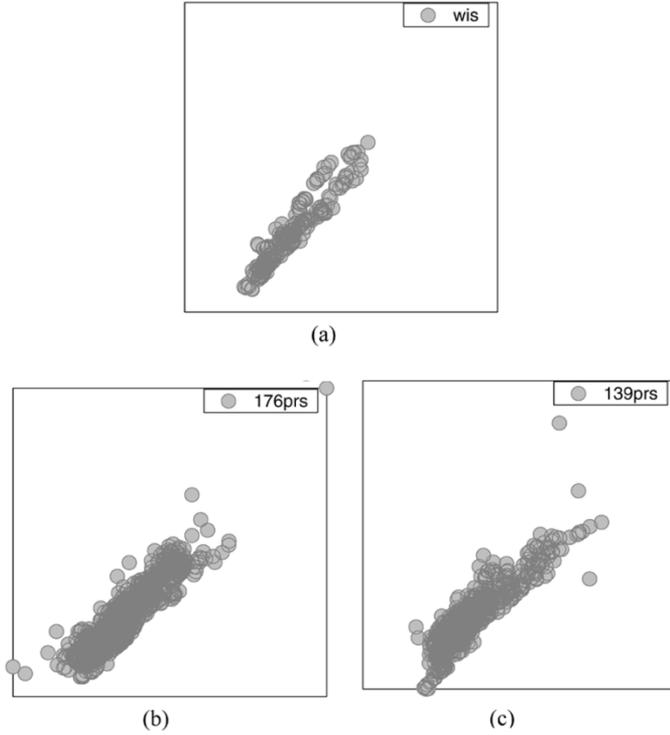


Fig. 8. Projection of L_{wm} spectra data onto the 2-D latent space. Each panel highlights the distribution of the radiometric data from a different sensor.

number of sampling wavelengths, $\hat{\mathbf{x}}$ and $E[\mathbf{x} | \hat{\mathbf{x}}]$ the auto-associative NN model input and output, respectively [see (4)]; and 2) setting the k and r thresholds for the outliers detection scheme.

The sensitivity of the *self-consistency* scheme to the threshold t of the auto-associative NN is presented in Fig. 5. Results show that the number of spectra identified as inconsistent slowly increases from 12 to 19 with t decreasing from 0.010 to 0.005, while it largely increases for $t < 0.005$. For the purpose of this study an empirical value $t = 0.006$ has been used, although a smaller value could be required for specific applications.

The k -nearest-neighbors algorithm, also used to identify *self-inconsistent* spectra, requires quantifying the k parameter and the radius r of the hypersphere. A spectrum is here assumed *self-inconsistent* if it does not have at least two other “very close” spectra. Hence, $k = 3$, with r used to quantify “very close.” The value of r was defined on the basis of the distribution of the hypersphere radius presented in Fig. 6. Here, each hypersphere is centered on a different data point and the radius is such that it allows for including the two closest data points in the hypersphere itself. An anomalous tail starts approximately when the value of the hypersphere is approximately 0.04. So, an L_{wm} spectrum is considered anomalous when the hypersphere radius is more than 0.04.

Notice that thresholds are dimensionless in both cases because data are preprocessed rescaling the L_{wm} spectra with respect to the value measured at 555 nm (see Section III). Also, the present analysis does not account for the dependence of the standard deviation of the radiometric data on the sampling wavelengths (see Tables I–III). This is supported by the normalization of L_{wm} spectra: the standard deviation of the rescaled spectra is between 0.14 and 0.15 for all wavelengths, with the exception of 440 nm for which it is 0.1.

TABLE IV
SUMMARY OF L_{wm} SPECTRA IDENTIFIED AS *SELF-INCONSISTENT*. THE THRESHOLD APPLIED TO THE AUTO-ASSOCIATIVE MODEL IS 0.006. OUTLIERS ARE IDENTIFIED SETTING THE KNN PARAMETERS $k = 3$ AND $r = 0.04$, RESPECTIVELY

L_{wm} spectra identifier	Auto-associative scheme	Outlier detection scheme
139prs-03/01/2004-13:48:06	*	
139prs-11/09/2004-08:40:58	*	
139prs-16/10/2004-13:08:55	*	
176prs-01/01/2003-09:48:21	*	
176prs-01/02/2003-08:38:52	*	
176prs-09/02/2003-10:38:03	*	
176prs-16/06/2003-13:24:40	*	
176prs-03/02/2005-09:37:19	*	
139prs-25/02/2004-12:37:04	*	*
139prs-25/07/2004-09:29:36	*	*
139prs-17/10/2004-12:29:29	*	*
176prs-25/09/2002-12:36:07	*	*
176prs-14/02/2003-14:56:51	*	*
176prs-16/02/2005-14:57:55	*	*
176prs-27/02/2005-09:36:43	*	*
176prs-27/02/2005-09:57:22	*	*
139prs-18/07/2003-12:49:51		*
139prs-30/08/2003-13:24:36		*
139prs-30/08/2003-13:45:18		*
139prs-30/01/2004-10:36:23		*
176prs-18/12/2002-13:41:02		*
176prs-01/01/2003-08:27:48		*
176prs-01/01/2003-08:48:26		*
176prs-10/02/2003-14:59:36		*
176prs-16/02/2005-12:38:19		*

Data classified as *self-inconsistent* in the original higher dimensional spectral space are highlighted in the top row panels of Fig. 7 through the same latent map used in the previous section. The red points in the top central panel correspond to spectra that violate the auto-associative constraint while green points refer to the KNN scheme. When the same spectrum violates both criteria, the corresponding latent point is colored in yellow.

A simple visual inspection confirms the presence of significant anomalies in those spectra which violate the KNN scheme, or both the KNN and auto-associative schemes. Differently, spectra violating the auto-associative scheme only may present minor anomalies. A quantitative summary of results from the applied statistical methods is presented in Table IV. Out of the 25 spectra classified as *self-inconsistent*, eight do not satisfy the auto-associative requirements, nine the KNN scheme, and eight both criteria. For illustrative purposes, the second and third row panels of Fig. 7 detail the spectral shape of the points nearby those indicated by the red arrows. In both cases the anomalous spectrum exhibits an inconsistent value at 440 nm.

The former results indicate that the use of reference data is not strictly required to assess the *self-consistence* of L_{wm} spectra

collected by the autonomous system, provided these data do not include measurements affected by systematic errors. Still, statistics from reference data—when available—can offer additional quality assurance elements further on increasing the confidence on the investigated data (such a capability has relevance in the production of very high quality data for vicarious calibration processes).

B. Relative-Consistency

Latent points resulting from the projection of the quality-assured spectra from the WiSPER system (see Section II-B) are shown in Fig. 8(a), while points resulting from the projection of the spectra from the two SeaPRISM systems (see Section II-A) are displayed in Fig. 8(b) and (c). The WiSPER latent points exhibit a narrow and almost regular distribution. This is likely to be explained by the occasional collection, generally on ideal conditions, of these quality-assured measurements. Since the axes of all these three plots have the same scale (not shown), an appreciable similarity of mean location, spread, and orientation characterizes the WiSPER and SeaPRISM data distributions. The SeaPRISM spectra projected onto latent regions with a low density of points are probably related to extreme environmental conditions encountered during the continuous operation of the autonomous systems.

After modeling the distribution of the reference dataset with the GMM, the spectral *relative-consistency* is defined on the basis of the probability of finding the spectrum to be classified within the distribution of the quality-assured data. In the central column of Fig. 9, data points are highlighted with different colors for different degrees of novelty. Data points with more than 50% probability of being included within the distribution of the quality-assured data are indicated in yellow. Similarly, points in green, dark green and red correspond to threshold probabilities of 25%, 20%, and 10%, respectively. Additionally, Fig. 9 details three sets of points from latent regions with highly *relative-inconsistent* spectra. Each set of points corresponds to a peculiar bio-optical condition, as hereafter described.

The top row panels mostly display data from a single event occurring in May-June 2004 and characterized by the presence of relatively high pigment concentration ranging from 3 to approximately 25 mg/m³. This event, considered quite exceptional during the three-year SeaPRISM deployment at the AAOT, was also confirmed by the SeaWiFS satellite data and, on a lesser extent, by *in situ* measurements of pigment concentration [14].

The second row panels mostly present data observed in the months of July-August and exhibiting features typical of relatively low pigments concentration (i.e., below 0.2 mg/m³) already observed during summer seasons [18].

Finally, the third row panels illustrate data collected in the months of October-January and exhibiting features determined by intermediate pigments concentration but with increased (when compared with the former two cases) colored dissolved organic matter and non-pigmented particles of riverine origin mostly occurring during fall-winter seasons [18].

It is relevant to note that the outlier spectra appearing in the second and third row of Fig. 9 were previously flagged as *self-inconsistent* (see Fig. 7 and Table IV).

assessed with a novelty detection approach only (i.e., modeling the distribution of quality-assured samples and then looking for outliers). In fact, data not exhibiting any spectral anomaly, but not represented within the set of quality-assured measurements, would be misclassified: they would appear novel without being anomalous. In addition, the natural variability of L_{wn} makes difficult the identification of spectra with minor anomalies as novel data. Consequently, instead of relying on a single scheme, the present study has investigated the consistency of L_{wn} spectra at two different level of inference, so-called *self-consistency* and *relative-consistency*. Specifically, an auto-associative NN was jointly used with a KNN algorithm to assess the *self-consistency* of L_{wn} spectra without accounting for any extra information obtainable from quality-assured data. Independently, a novelty detection approach was applied to identify *relative-consistent* spectra explicitly on the basis of the quality-assured data.

By assuming that the intrinsic dimensionality of the modeled spectra is smaller than the number of sampling wavelengths, the *relative-consistency* was defined on the basis of the features extracted from the bottleneck of the auto-associative model. To illustrate that the original data structure is preserved at the NN bottleneck outputs, this study has shown how points close in the latent space correspond to similar spectra. Results have thus highlighted how different latent regions can be related to different bio-optical conditions. This “topographic nature,” which is expected as long as the continuity between spectra can be captured by the auto-associative NN, additionally supports the validity of the implemented model.

Due to the impossibility of giving an universal definition of anomalous spectrum, the effectiveness of the proposed quality assurance scheme has been *a posteriori* assessed through a visual inspection of the classification results. Specifically, the *self-consistency* scheme has illustrated the capability of filtering out different percentages of anomalous spectra as a function of the thresholds used for the auto-associative NN and KNN algorithms. For instance, changing from 0.01 to 0.001 the threshold applied to the auto-associative NN allowed for the removal of percentages of *self-inconsistent* spectra increasing from approximately 1% to above 10% of the total measurements. It is recalled that this relatively small number of abnormal spectra depends on the basic quality assurance applied during the dataset creation (i.e., the L_{wn} data used in this study were already screened for major artifacts).

It is finally pointed out that the possibility of detecting anomalous spectra with the *self-consistency* analysis, and of additionally identifying spectra mostly comparable to reference measurements (*relative-consistency* analysis), allows for a cross-screening of L_{wn} spectra produced from autonomous systems. This is of extreme importance for the production of very high quality data suitable for the vicarious calibration of space sensors.

V. SUMMARY AND CONCLUSION

The consistency of large volumes of normalized water leaving radiance spectra, L_{wn} , from autonomous systems cannot be as-

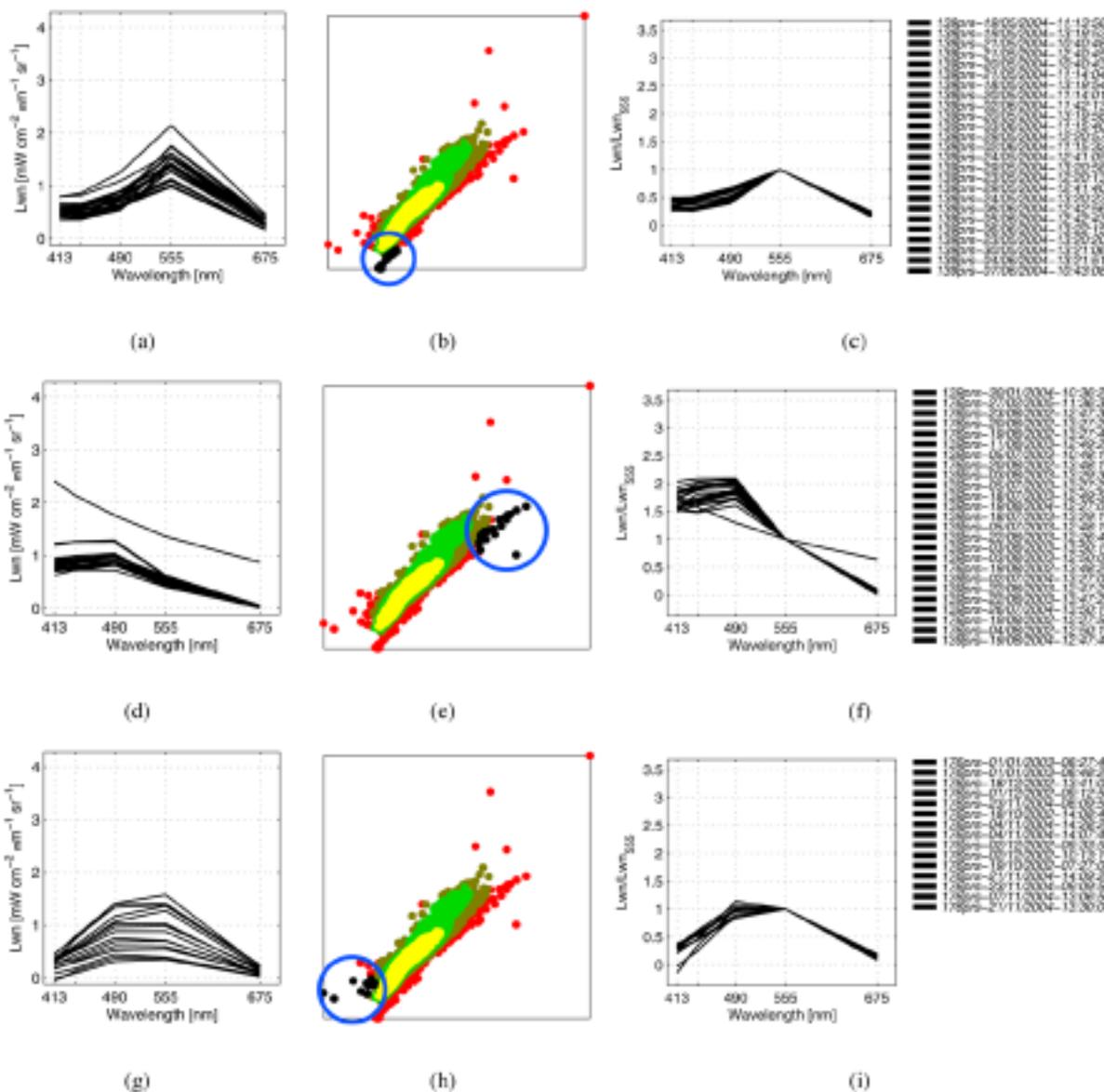


Fig. 9. Novelty detection screening of *self-consistent* L_{wm} spectra based on the GMM algorithm. Yellow points correspond to spectra with probability higher than 50% of being represented in the distribution of the quality-assured data. Similarly, points in green, dark green, and red correspond to a threshold probabilities of 25%, 20%, and 10%, respectively. The three sets of points highlighted with the blue circles, one for each set of row panels, correspond to different bio-optical conditions (see text).

ACKNOWLEDGMENT

B. Holben and I. Slutsker are acknowledged for the support provided in managing the SeaPRISM data through AERONET.

REFERENCES

- [1] L. Boaz, G. Hugo, A. Mayer, and D. Its'hak, "A comparative study of neural network based feature extraction paradigms," *Pattern Recognit. Lett.*, vol. 20, pp. 7–14, 1999.
- [2] G. Kerschen and J.-C. Golinval, "Feature extraction using auto-associative neural networks," *Smart Mater. Struct.*, vol. 13, pp. 211–219, 2004.
- [3] D. L. Mattern, L. C. Jaw, T. H. Guo, R. Graham, and W. McCoy, "Using neural networks for sensor validation," NASA Lewis Res. Center, Cleveland, OH, TM-1998-208483, 1998.
- [4] C. M. Bishop, "Novelty detection and neural network validation," in *IEEE Proc. Vision and Image and Signal Processing*, vol. 141, 1994, pp. 217–222.
- [5] P. Baldi and K. Hornik, "Neural networks and principal component analysis: Learning from from examples without local minima.," *Neural Netw.*, vol. 2, pp. 53–58, 1989.
- [6] H. Bourlard and Y. Kamp, "Auto-association by multilayer perceptrons and singular value decomposition.," *Biol. Cybern.*, vol. 59, pp. 291–294, 1988.
- [7] M. A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks.," *Neural Comput.*, vol. 9, no. 7, pp. 1493–1516, 1991.
- [8] T. Aarup, N. Holt, and N. K. Højerslev, "Optical measurements in the north sea—Baltic transition zone. Statistical analysis of bio-optical data from the eastern North Sea and the Skagerrak and the Kattegat," *Cont. Shelf Res.*, vol. 16, pp. 1355–1377, 1996.
- [9] A. Morel and J. L. Muller, "Normalized water leaving radiance and remote sensing reflectance: Bidirectional reflectance and other factors," NASA Goddard Space Flight Center, Greenbelt, MD, TM-2002-210004/Rev. 3, vol. 1, pp. 183–210, 2002.

- [10] G. Zibordi, J.-F. Berthon, B. Bulgarelli, D. D'Alimonte, D. van der Linde, F. Mélin, and C. Targa, "Ocean color validation activities at the AAOT in the northern adriatic sea," *Int. J. Remote Sens.*, vol. 25, pp. 1533–1537, 2004.
- [11] J.-F. Berthon and G. Zibordi, "Bio-optical relationships for the northern adriatic sea," *Int. J. Remote Sens.*, vol. 25, pp. 1527–1532, 2004.
- [12] B. N. Holben, T. F. Eck, I. Slutsker, D. Tanré, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov, "AERONET—A federated instrument network and data archive for aerosol characterization," *Remote Sens. Environ.*, vol. 66, pp. 1–16, 1998.
- [13] G. Zibordi, F. Mélin, S. B. Hooker, D. D'Alimonte, and B. N. Holben, "An above-water system for the validation of ocean color radiance data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 2, pp. 402–414, Feb. 2004.
- [14] G. Zibordi, F. Mélin, and J.-F. Berthon, "A time-series of above-water radiometric measurements for coastal water monitoring and remote sensing product validation," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 120–124, Jan. 2006.
- [15] A. Smirnov, B. N. Holben, T. F. Eck, O. Dubovik, and I. Slutsker, "Cloud screening and quality control algorithms for the AERONET database," *Remote Sens. Environ.*, vol. 73, pp. 337–349, 2000.
- [16] G. Zibordi, D. D'Alimonte, and J.-F. Berthon, "An evaluation of depth resolution requirements for optical profiling in coastal waters," *IEEE Trans. Geosci. Remote Sensing*, vol. 21, no. 7, pp. 1059–1073, Jul. 2004.
- [17] J. E. O'Reilly, S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, and C. R. McClain, "Ocean color chlorophyll algorithms for SeaWiFS," *J. Geophys. Res.*, vol. 103, pp. 24 937–24 953, 1998.
- [18] J.-F. Berthon, G. Zibordi, J. Doyle, S. Grossi, D. van der Linde, and C. Targa, "Coastal atmosphere and sea time series (CoASTS): Data analysis," NASA Goddard Space Flight Center, Greenbelt, MD, TM-2002-206892, vol. 20, 2002. SeaWiFS Postlaunch Technical Report Series, pp. 1–25.
- [19] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.
- [20] D. D'Alimonte, F. Mélin, G. Zibordi, and J.-F. Berthon, "Use of the novelty detection technique to identify the range of applicability of empirical ocean color algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2833–2843, Dec. 2003.
- [21] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc. B*, vol. B39, no. 1, pp. 1–38, 1977.
- [22] I. T. Nabney, *Netlab: Algorithms for Pattern Recognition*. Berlin, Germany: Springer-Verlag, 2001.
- [23] D. Lowe and M. E. Tipping, "Neuroscale: Novel topographic feature extraction using RBF networks," in *Advances in Neural Information Processing Systems*, M. C. Mozer, M. I. Jordan, and T. Petsche, Eds. Cambridge, MA: MIT Press, 1997, vol. 9, pp. 543–549.
- [24] C. M. Bishop, M. Svensen, and C. K. Williams, "GTM: The generative topographic mapping," *Neural Comput.*, vol. 10, no. 1, pp. 215–234, 1997.
- [25] T. Kohonen, *Self-Organizing Maps*. Berlin, Germany: Springer-Verlag, 1995.



Davide D'Alimonte received the laurea degree in physics from the University of Turin, Turin, Italy, in 1997, and the Ph.D. degree in oceanography from the University of Southampton, Southampton, U.K., in 2004.

From 1999 to 2003, he was with the Joint Research Centre of the European Commission, Ispra, Italy, involved in ocean color activities for the determination of optically significant seawater constituents in complex coastal waters using neural network methods. From 2003 to 2005, he was a member of the Neural

Computing Research Group, University of Aston, Birmingham, U.K., where his research work focused on the mathematical modeling and statistical analysis of DNA microarray data, and ocean color applications. Since April 2006, he has been a member of the Joint Center for Earth Systems Technology, University of Maryland, Baltimore County.



Giuseppe Zibordi received the laurea degree in physics from the University of Modena, Modena, Italy, in 1981.

He was a Researcher with the Italian National Research Council, Modena, Italy, from 1984 to 1992, where his work focused on quantitative remote sensing of coastal and polar regions. Since 1993, he has been with the Joint Research Centre of the European Commission, Ispra, Italy, working on ocean color development and validation activities. His research interests include remote sensing techniques

for the determination of atmospheric and marine optical properties, protocols for *in situ* measurements, and methods for the absolute radiometric calibration of optical instruments.