УДК 551.465

© *Дж.Пеннучи, А.Альварес, Ч.Триз*, 2012
Центр подводных исследований НАТО, Ла Специя, Италия
pennucci@nurc.nato.int

СПУТНИКОВЫЙ МЕТОД, ОСНОВАННЫЙ НА КОВАРИАЦИИ, ДЛЯ ПОДДЕРЖКИ ДЕЯТЕЛЬНОСТИ АЭРОНЕТ – ВЕРИФИКАЦИЯ ДАННЫХ ПО ЦВЕТУ ОКЕАНА

Цель работы – определить места в исследуемой области океана, где натурные измерения для калибровки/верификации (Cal/Val) в течение различных периодов времени обеспечивают наибольшее улучшение радиометрической точности и достоверности результатов, получаемых по спутниковым данным. Представлен метод объединения спутниковых изображений с данными натурных измерений и выработки наилучшей стратегии проведения натурных измерений, подходящей для проведения спутниковой калибровки/верификации. Эта методология использует спутниковые данные, чтобы построить ковариационную матрицу, содержащую информацию о пространственно-временно́й изменчивости в исследуемой области. Ковариационная матрица используется в Байесовском методе для объединения спутниковых и натурных данных и создания продукта с наименьшими ошибками. Наилучшее место для калибровки/верификации находится с помощью метода оптимального планирования (A-оптимальный план), который минимизирует расчетную дисперсию объединенного продукта.

Ключевые слова: спутниковые изображения, натурные измерения, калибровка/верификация, объединенный продукт.

With the increasing availability of satellite time-series imagery of sea surface temperature (SST) and ocean color (e.g. from Moderate Advanced Very High Resolution Radiometer-MODIS), it has become possible to monitor temporal and spatial variation of coastal and open waters. These data improve our view of the ocean when compared to the very limited spatial sampling offered by *in situ* observations (e.g. ships and moorings). Generally, merging remote sensing data with *in situ* measurements has become a standard procedure to increase the quality of satellite derived products. Conventionally, covariance analysis is applied to oceanographic and meteorological data sets to decompose space and time distributed data into modes ranked by their temporal variance, while optimum sampling analysis is applied to find adequate number and allocation of *in situ* data to improve satellite quality by reducing the overall observational error. In this paper, different fields needed for implementing such concepts are studied and presented.

These methodologies were implemented on several available data sets (e.g. using satellite MODIS time series and optical *in situ* platform, such as the AErosol RObotic NETwork – AERONET and the Marine Optical Buoy-MOBY platforms).

Methods. *Theoretical Approach.* Time series of satellite images can be employed to build a covariance matrix encoding the spatio-temporal variability of an area of interest. *In situ* observational resources can then be adaptively distributed following a covariance-oriented criterion, to assign the best value of the *in situ* observed field at grid points of a regular grid coincident with the centers of satellite pixels. The best *in situ* location can be found implementing optimum design procedure, such as A-optimum design.

Covariance – Consider a generic time-series $\{\psi(x, y, t_i)\}_{i=1}^{N}$ measured from satellite with a given observational error. The dataset is a three-dimensional grid that depends on longitude (x), latitude (y) and time (t). Alternatively, this gridded data can be reshaped into a two-dimensional grid of M rows by N columns:

$$T_{NM} \equiv T(N,M),$$

where M represents the number of spatially distributed points (the product of x by y) and N represents the number of points over time (t). Using this representation, the covariance matrix (C) can be numerically evaluated by multiplying T by its transpose:

$$C \propto (T_{NM})' \cdot T_{NM}.$$

Because the covariance matrix C is derived from satellite observations, it will contain contributions from the sensor noise (σ^2_{sat}). For this reason we have studied a methodology to remove the impact of the sensor noise (supposing that σ^2_{sat} is known *a priori*) on the covariance matrix. To achieve this, we have decomposed C in two orthogonal matrixes that verify the following equation:

$$C \cdot V = V \cdot D. \tag{1}$$

These matrices are the eigenvalues (D) and eigenvectors (V) of C; in particular, D is the *canonical form* of C (a diagonal matrix with C's eigenvalues on the main diagonal), while V is the *modal matrix* (its columns are the eigenvectors of C). Assuming the sensor noise as a white noise stochastic process, its impact on the covariance matrix C is limited to the diagonal terms. These characteristics make it possible to remove the sensor noise by using the eigenvalues of C; therefore the matrix D will be modified as follow:

$$D = \left(D - \left(\frac{\sigma^2_{sat}}{M} \right) I \right),$$

where I is the MXM identity matrix. Finally, replacing the negative eigenvalues with zeros, the new covariance matrix can be evaluated using the formula:

$$C = VDV'. \tag{2}$$

Merging procedure – Merging remote sensing data with *in situ* measurement is a standard procedure that allows increasing the quality of the satellite-derived products. The idea is to study the spatial-temporal variability of the satellite data and to distribute the *in situ* sample over the image following the covariance criterion. Therefore, once the covariance C has been obtained from eq.(2), a new field, merging *in situ* and satellite data, is retrieved maximizing the following probability distribution:

$$P(\psi_K) \propto \exp[-(\psi_{obs} - H\psi_k)^T \textstyle\sum_{obs}^{-1} (\psi_{obs} - H\psi_K) - (\psi_K - \overline{\psi})^T C^{-1} (\psi_K - \overline{\psi})]$$

Where ψ_K represents the vector of pixel values, ψ_{obs} is the observation vector, H is the observation matrix, $\sum_{obs}$ is the observation error matrix and $\overline{\psi}$ is the average field. The first part in the exponential represents the likelihood density while the second product of matrices represents the *a priori* probability. The merging procedure is performed maximizing the *a posteriori* probability distribution; therefore the best estimation is represented by the field ψ_{merged} that verifies:

$$\psi_{merged} = \arg\min_{\psi_k} \left((\psi_{obs} - H\psi_k)^T \textstyle\sum_{obs}^{-1} (\psi_{obs} - H\psi_k) - (\psi_k - \overline{\psi})^T C^{-1} (\psi_k - \overline{\psi}) \right). \tag{3}$$

The solution of eq.(3) represents the merged image, fig.1 show an example of merging using SST AVHRR image and an *in situ* track.

Optimum design and Uncertainty Index – Sampling strategies of *in situ* observational resources driven by a design principle called A-optimality could substantially improve the accuracy of the final blended products. The scope of A-optimal designs is to minimize the variance of the estimated field with respect the sample locations. This optimal criterion will select locations in regions with low uncertainty and large spatial representation. Like other standard variance-oriented criteria in optimal experimental design, a covariance model must be known *a priori*. *In situ* observational resources could be adaptively distributed following the variance-oriented criterion, to assign the best values of the *in situ* observed fields at grid points of a regular grid coincident with the center of satellite pixels. This procedure would ensure the optimality of merged products for limited *in situ* observational resources on the basis of an Uncertainty Index (UI). The implementation of this technique was initially performed using a Genetic Algorithm (GA) that minimizes the process of natural evolution. This algorithm is iteratively used to search the best *in situ* position minimizing the variance of the retrieved solutions. The optimization problem was also investigate using a Simulated annealing (SA) strategy that is a generic probabilistic metaheuristic for the global optimization problem of locating a good approximation to the global optimum of a given function in a large search space. The retrieved optimization results from GA and from SA are comparable; therefore the SA method results more efficient in term of computation (faster) than GA.

a *b*

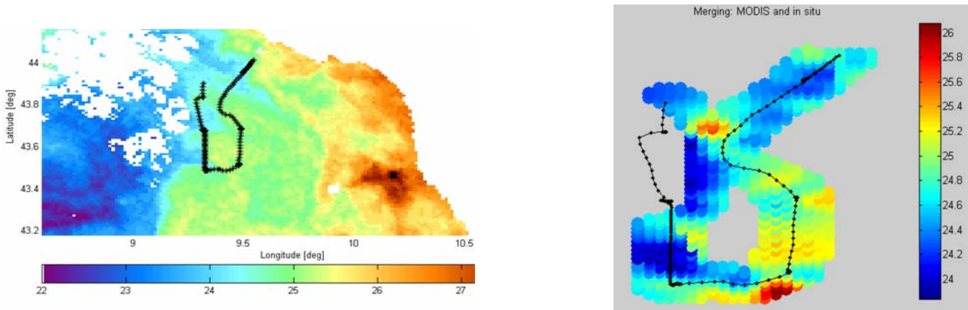


Fig.1. AVHRR SST image and *in situ* ship data (black dots) – *a*; colored circles display the field resulting from merging ship and satellite observations of the sea surface temperature – *b*.

Results. Implementation of the procedures described above focuses on the AErosol RObotic NETwork (AERONET) sites. AERONET program is a federation of ground-based remote sensing aerosol networks established by NASA and LOA-PHOTONS (CNRS) and other national collaborators. The program provides a long-term, continuous and readily accessible public domain database of aerosol optical, mircrophysical and radiative properties for aerosol research and characterization, validation of satellite retrievals, and synergism with other databases [3, 4]. In particular we focus the attention on the Venice Acqua Alta AERONET site that is located 13 KM of the cost of the Venetian lagoon (as in fig.2).

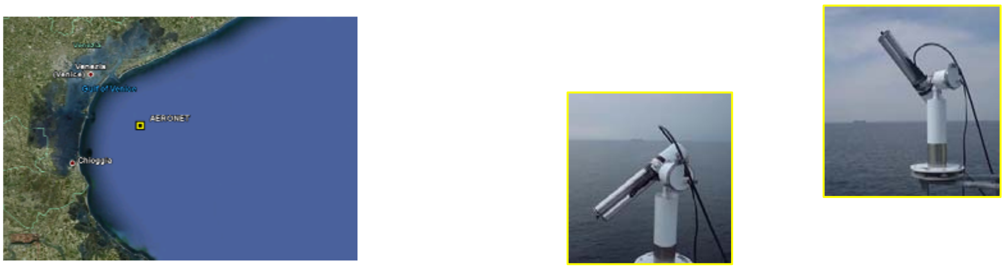
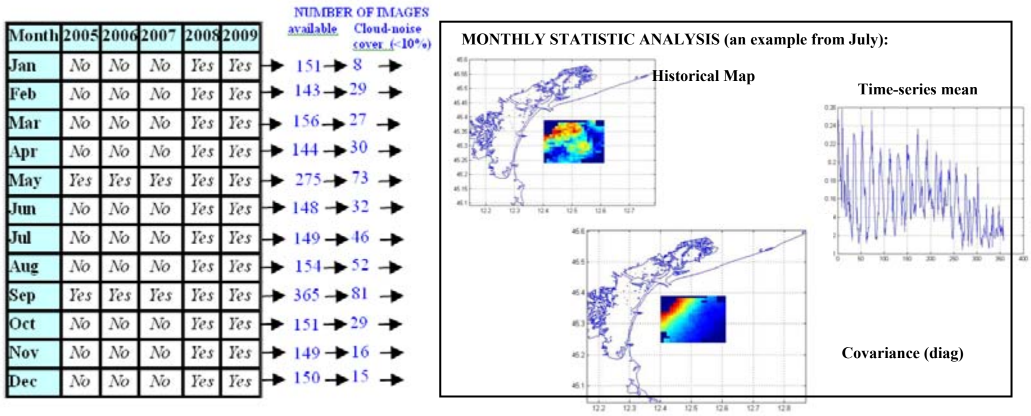


Fig.2. The northern-East Italy coast region, showing the location of the AERONET site.

To evaluate the AERONET *in situ* position in terms of uncertainly we have implemented the proposed procedure on monthly satellite time-series that have been retrieved using several Moderate Advanced Very High Resolution Radiometer (MODIS) images acquired on the area of interest with 1 km at nadir of ground resolution. In particular we have performed the following steps.

Monthly time series have been created using about two acquisitions per days during the period from January 2005 to December 2009, for a total of 2135 images. All the «clear» images (438) were processed focusing the attention on a box area of 30 by 30 km around the AERONET site (this size was fixed for the convenience of computation and analysis). Each monthly time series were arranged into a two-dimensional array $T(x, t)$, where x and t are the spatial and temporal indices. Because the data retrieved from MODIS are much more dense in space than in time ($x>>t$), the covariance matrix was evaluated implementing eq.(1) for each month. To represent the monthly statistic analysis we have evaluate the mean of each time series and we also define a «Historical Covariance Map» (C_{HIST}) that represents the pixel standard deviation of considered time-series, as resumed in fig.3. Using this technique we produce twelve C_{HIST} that have been used to calibrate *in situ* data without satellite acquisitions but taking into account an «satellite statistical behavior».



Month	2005	2006	2007	2008	2009	NUMBER OF IMAGES available	Cloud-noise cover (<10%)
Jan	No	No	No	Yes	Yes	151	8
Feb	No	No	No	Yes	Yes	143	29
Mar	No	No	No	Yes	Yes	156	27
Apr	No	No	No	Yes	Yes	144	30
May	Yes	Yes	Yes	Yes	Yes	275	73
Jun	No	No	No	Yes	Yes	148	32
Jul	No	No	No	Yes	Yes	149	46
Aug	No	No	No	Yes	Yes	154	52
Sep	Yes	Yes	Yes	Yes	Yes	365	81
Oct	No	No	No	Yes	Yes	151	29
Nov	No	No	No	Yes	Yes	149	16
Dec	No	No	No	Yes	Yes	150	15

Fig.3. Data processing chain to perform a statistical analysis on the available time series.

A maximum covariance value (P_{max} = $1/e$ ~ max probability) is defined and just the covariance-pixels lower than P_{max} were considered. This procedure allows retrieving an Uncertainty Map (fig.4, *a*) that represents the reduction of the satellite uncertainly (error) in a particular area.

The *in situ* observational resources were then adaptively distributed following the variance-oriented criterion, to assign the best values of the *in situ* observed fields at grid points of a regular grid coincident with the centres of satellite pixels. We define an Impact Map as the result of the merging between an *in situ* (with hypothetic latitude longitude position) and the Uncertainty Map.

As showed in fig.4, *b*, *c* the Impact Map depends on the *in situ* position. In order to have number that represents this variation, we have defined an Impact Index that takes into account of the effective error reduction (due to the *in situ*) in a fixed area. In particular, we define an Image retrieved from the difference between the Uncertainty Map (fig.5), we consider a box-area (10×10 km) around the *in situ* (*A* pixels) and we fix a threshold (*B*) with the following value:

$$\mathrm{B} = \mathrm{Uncertainty\_map}\left(\mathrm{lon}_{\mathrm{in\ situ}}, \mathrm{lat}_{\mathrm{in\ situ}}\right) - 3\sigma_{\mathrm{Satellite\_error}} \,.$$

If N represents the number of pixels >B, the Impact Index can be defined as N/A*100. Knowing the Impact Index of each possible *in situ* location allows one to identify which is the best placement in terms of Cal/Val activities.

Conclusion. We have presented a procedure for merging satellite data with *in situ* measurements to increase the quality of satellite derived products. This methodology is used to define the location where *in situ* data should be collected in order to determine the uncertainty of using these data for calibration and validation of satellite products. Satellite products include Sea Surface Temperature, Ocean Color products of water leaving radiance, chlorophyll, inherent and apparent properties (retrieved from AVHRR and MODIS satellite sensors). *In situ* measurements can be obtained from moorings (such as AErosol RObotic NETwork-

AERONET and/or Marine Optical Buoy-MOBY), from ships or from autonomous vehicles (such as Autonomous Underwater vehicle and/or Gliders).

a *b* *c*

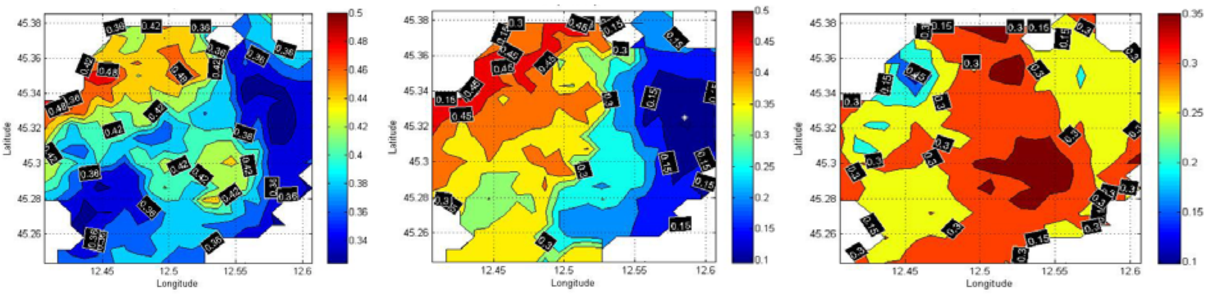


Fig.4. Uncertainty Index map (*a*); Impact Index Map merging with an *in situ* placed in a low uncertainty area (43.325N and 12.583E) (*b*); Impact Index Map merging with an *in situ* in placed in a low uncertainty area (43.325N and 12.583E) (*c*).

a *b*

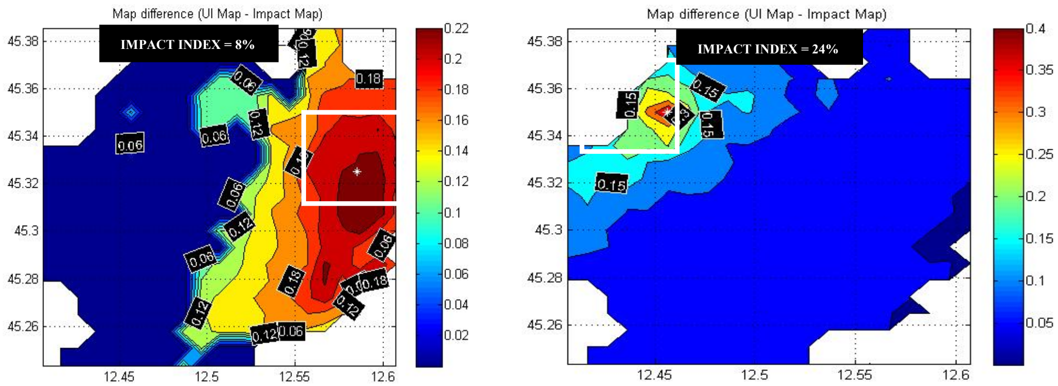


Fig.5. Image Difference and Impact Index merging with an *in situ* placed in a low uncertainty area (43.325N and 12.583E) (*a*); Image Difference and Impact Index merging with an *in situ* placed in a high uncertainty area (43.325N and 12.583E) (*b*).

We also present results using MODIS time-series images and AERONET-OC in the Venice (Acqua Alta) site. The covariance matrix of the time-series was used in a Bayesian framework to estimate the best *in situ* location for Cal/Val efforts using a Simulated Anealling Algorithm. In particular, the covariance has been evaluated using the available monthly time-series MODIS acquisitions from 2005 to 2009. The resulting Historical Maps have been used to calibrate *in situ* data position without satellite acquisitions but taking into account on the "satellite statistical behaviour".

We express special thanks to the Naval Research Laboratory, Stennis Space Center (NRL-SSC) for support to NURC in the participation of the project «Execution of a Calibration/Validation (Cal/Val) Plans for Oceans Environmental Data Records (EDRS) for the Visible Infrared Spectrometer (VIIRS) sensor on the Preparatory Project (NPP) and National Polar Orbiting Operational Environmental Satellite System (NPOESS)».

References

1. *Pennucci G., Alvarez A., Trees C.* Study of covariance as a function of cloud cover, NURC report (Under Progress), December 2010.
2. *Lagerloef G.S.E., Bernstein R.L.* Empirical Orthogonal Function Analysis of Advanced Very High Resolution Radiometer Surface Temperature Patterns in Santa Barbara Channel // J. of Geophysical Research. June 15, 1988. V.93, N C6. P.6863–6873.
3. *Zibordi G.B.* et all. AERONET-OC: A Network for the Validation of Ocean Color Primary Products // J. Atmos. Oceanic Technol., 26, 1634–1651, DOI: 10.1175/2009JTECHO654.1, 2009.
4. http://aeronet.gsfc.nasa.gov/.

Статья поступила в редакцию 18.04.2012 г.