A Grid Enabled Look-Up Table for Aerosol Optical Thickness Estimation on Coastal Water

Andrea Guerriero§, Raffaella Matarrese*, Alberto Morea*, Khalid Tijani*
§Department of Electric and Electronic, Politecnico di Bari
*Department of Physics, Politecnico di Bari
Via Re David 200, 70100 Bari (BA)

Abstract: - Coastal area represents one of the more delicate and complex relation between natural environment and human activities. Remote sensing, offering considerable understanding of the temporal dynamics of bio-physical factors at different scale, represents a solution for monitoring the ecosystem. In particular marine phenomena, i.e. Sea Surface Temperature, chlorophyll concentration etc. can be related to the sea color observed from satellites. However standard satellite products, such as chlorophyll concentrations maps calculated from MODIS data, are useful only for oceanic water and cannot be directly extended to coastal water. For coastal water quality evaluation, specific inversion procedures should be used. Some of these procedures are based on look-up table generation that are computational intensive. In this paper a Grid based approach is proposed for LUT generation. This approach, producing accurate estimation of Aerosol Optical Thickness, represents a useful system to give fast and accurate answers without demand large economic investments.

Key-Words: - Aerosol Look-up table, Grid computing, chlorophyll retrieval, remote sensing.

1 Introduction

Coastal area represents one of the more delicate and complex relation between natural environment and human activities. The increase, in the last few decades, of a culture that holds in greater consideration the environment, induced several Governments to promote programs to protecting and conserving the environment and in particular coastal area and water. These programs aim to monitoring the main environment parameters and to identify systems capable to give fast answers without demand large economic investments.

Satellite remote sensing, offering the possibility of covering a large spatial area with a high temporal frequency, represents the solution for monitoring the ecosystem, offering considerable understanding of the temporal dynamics of bio-physical factors at different scale.

In particular, marine phenomena, i.e. Sea Surface Temperature, chlorophyll concentration etc. can be related to the sea color observed from satellites [1-5]. However standard satellite products are optimized only for oceanic water and cannot be directly extended to coastal water. This is due to the differences between the two types of waters. In *case I* waters (oceanic) the characteristics of the components in the water allow to simplify the problem of the parameters extraction, because their strong correlation (the phytoplankton become the dominant parameter), the low suspended sediment presence (that simplify the atmospheric correction of the images with the validity of the assumption of black pixel), the homogeneity of the aerosol on large areas, the water deep and the distance from the coast. Each of these features introduces a complication in coastal areas study, *case II* waters, especially concerning the atmospheric corrections to apply on the images.

The approaches generally used by the oceanographers for the identification of water components, resolve the radiative transfer equation for the entire system atmosphere-sea. Several numerical models have been proposed (i.e. Neural networks [6], Monte Carlo [7]), to calculate water leaving reflectances, chlorophyll concentrations, suspended sediments and yellow substance. Such approaches, because the differences between the two types of waters, result suitable for oceans, but, for coastal waters, do not supply good control on the error sources that can influence the final result.

A different approach consists in splitting the atmosphere-water system, retrieving the most important parameter that links both systems, the Water Leaving Reflectance [8]. This means removing all the atmosphere contributions, by the mean of an accurate retrieval of the Aerosol Optical Thickness (AOT), the most delicate unknown describing the atmosphere, very changeable in coastal waters.

For standard products of satellite data, a Look Up Table approach is followed [9-12]. Usually a huge look up table, which is a matrix of dependency variable corresponding to a set of combinations of independent variable values: Sensor-Target-Solar geometric conditions, ground surface reflectance, aerosol and atmospheric condition is built in order to retrieve Aerosol Optical Thickness. Several atmospheric Radiative Transfer Model have been proposed to produce LUT, i.e. 6S [12] and Modtran [11]. In [13] 6S code has been used to build a LUT of 91125 input combination sets, representing the geometrical conditions, atmospheric conditions, aerosol model and concentration, spectral band and ground reflectance. The building a so huge LUT, that requires a very large amount of computational time, is appropriate only if the LUT will be always the same for each pixel in the processed image, but this generalization produces a less accurate estimation of the AOT.

In this paper, with the aim of creating thematic maps of chlorophyll, a LUT (with less input parameters than the LUT usually used), is built for each pixel of the image. In this way it is possible to include in the process all the actual geometric conditions of the pixel, avoiding the interpolation that comes using a larger LUT, reducing the error sources in the algorithm [14].

LUT generation for AOT retrieval for each pixel is computational intensive, about 20 minutes per pixel on a P4 PC, requiring several days to obtain data for a single small image.

In this paper we evaluate a Grid based approach to Aerosol LUT generation for a small image. Grid technology, easily providing powerful computational resources and efficient distributed data management, could represent a solution capable to give fast answers without demand large economic investments.

The paper is structured as follows. Section 2 describes LUT generation for coastal area, section 3 outlines the main characteristics of the Grid infrastructure and services realized and finally, section 4 concludes the paper discussing the system performances and future improvements.

2 LUT generation

In this work water leaving reflectances for 850 pixel of MODIS sensor (http://modis.gsfc.nasa.gov), are calculated improving the algorithm both for oceanic and coastal waters, and a specific inversion procedure for coastal waters is adopted [9], allowing the evaluation of the optical parameters concentrations. To obtain a more accurate value of the water leaving reflectance, atmospheric correction should account the Aerosol Optical Thickness (AOT) and its size distribution corresponding to every pixel for the selected image.

A look-up table with a value of apparent reflectivity for 45450 input parameters (nine size-distribution models for the aerosol particle, arranged percentage-wise between 0 and 100, ten values of AOT and five MODIS bands) is built. The minimization of the error on the estimated apparent reflectivity done by the 6S [12] code for the atmospheric correction, gives a good estimation of the mix between the size-distribution models and the corresponding value of AOT.

The 6S Vectorial code is used in this work. The version 1.0 Beta is now available to the public (source code available at ftp://kratmos.gsfc.nasa.gov/pub/eric/6S/ and web interface at http://6s/ltdri.org). A careful examination of its accuracy has been conducted both by comparing to other radiative code under special reference conditions and by applying it to actual remote sensing data for which independent surface measurements could be compared against [14].

In our approach the information about the actual geometric conditions of the pixel for which compute the LUT, its reflectance in the solar spectral region and the values of pressure, temperature, concentration of water vapour and ozone, are obtained by MOD03, MOD021 and MOD09CRS, MODIS standard products [15].

The knowledge of the real values of these parameters, allows the size reduction of the LUT and a more accurate apparent reflectivity evaluation. In order to avoid each machine in the grid structure to load a full MODIS image (2030x1354 pixels, about 350Mb), a pre-processing step is carried out, extracting all the information for a pixel and writing them to an ASCII input file. In this way 25 pixels for 34 images are distributed on Grid nodes.

3 The Grid Infrastructure

Currently, Grid technology by Internet interconnection can provides “virtual” supercomputers with a complete integration of heterogeneous computing systems, data resources and software tools that can be used by researcher and scientists of one or more virtual organization.

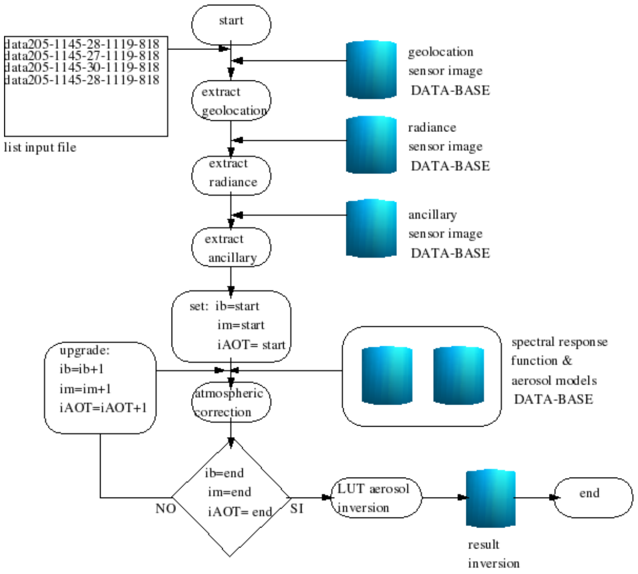

Fig.1 Aerosol LUT generation.

Dynamic Grid virtual organization members can share their resources by standard interfaces, moreover “this sharing is, necessarily, highly controlled, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the condition under which sharing occurs” [16] . Globus Toolkit [17] is the most widely used middleware for building Grid system and applications. This is a collection of open source software services and libraries that enable distribute infrastructures and applications. It allows interoperation capabilities between different software applications running on different platforms, standardizes the message exchange making large use of Web Services [18].

Our application involves data management, with access to huge quantities of data distributed between different sites. Several GT4 services implement useful mechanisms for data access, for example: GridFTP to realize high performance and secure file transfer and RLS to maintain information about the location of replicated files and datasets. For our application we use an open source relational database, MySQL, as data repository. To improve the system performances data are distributed among three computational nodes, therefore the RLS is used to find in which node data are located.

To find a copy of a data file, a request is sent to the data management system, the Lightweight Data Replicator (LDR). LDR queries the Replica Location Service to find a local copy of the file; if data file is not found locally, the RLS tells the data management system where the file exists in the Grid.

The AOT evaluation services, process selected data without heavy raw data transfer and make the whole processing more user “friendly”.

The processing system proposed in this work, is realized on a Grid infrastructure formed by 19 heterogeneous nodes, localized between the Taranto and Bari locations of Politecnico di Bari, interconnected by a VPN. Linux Fedora Operating system and Globus toolkit 4.0.3 middleware have been installed on each node, providing libraries and tools to use and realize Grid services. Cl21.tabavpn.it is the Grid certification node, it signs and trusts the certificate requests. It is also the MySQL server. TaGIIS2.tabavpn.it is the GIIS node, par02.tabavpn.it and cl21.tabavpn.it are two storage nodes.

A Graphical User Interface allows data exploration, selection and processing, managing Grid resources and services in a totally transparent way. The user can select the ROI, the EOS satellite, MODIS product, date or period of interest and search for available data. From retrieved data, the actual geometric, reflectance and atmospheric values are extracted and the AOT evaluation service can be invoked to build the LUT for single pixel or for small areas. Each LUT building requests is submitted to a distinct Grid computing nodes for simultaneous execution, reducing the total execution time.

4 Performance evaluation

Our experiments were carried out on a Grid infrastructure composed by 19 heterogeneous P4 PC.

In the first test, from a dataset composed by a MOD03, a MOD02 and a MOD09CRS MODIS product, the values corresponding to a chosen pixel have been used to build a LUT. The number of pixels and computational nodes have been incremented to 4, 8 and 16 and the parallel execution time, reported in Tab. 1, remained almost the same. To reduce the difference in the computational time evaluation, due to different pixel values, the same pixel has been used for all the evaluations.

Dataset	Nodes	Sequential Execution Time (minutes)	Parallel Execution Time (minutes)
2	1	48	48
4	2	96	49
8	4	192	49
16	8	384	48

Tab.1 the parallel execution time does not change increasing, at the same time, dataset and computational nodes.

Table 2 shows the results of the second test. A dataset corresponding to 16 pixels has been used to measure the speedup and efficiency of the proposed Grid approach. The workload has been partitioned between 1, 2, 4, 8 or 16 node, obtaining always an excellent efficiency ratio.

Dataset	Nodes	Parallel Execution Time (minutes)	speedup	Efficiency %
16	1	378	1.00	100.00
16	2	189	2.00	100.00
16	4	128	2.95	73.8
16	8	48	7.85	98.43
16	16	24	15.75	98.41

Tab.2 speedup and efficiency of the proposed Grid approach.

Table 3 shows the results of the last test, a comparison between sequential execution time (about 22 days) and the Grid approach (20 hours) in a real application to build the LUT for 34 areas of 25 pixel each.

In fig.2 the workload distribution for the last test is shown.

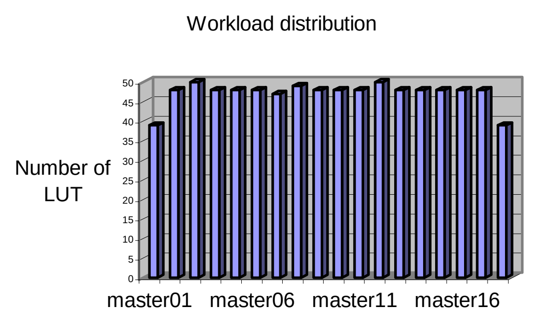

Fig.2 Workload distribution

The GRID offers capabilities for distributed computation, but has severe limitations in EO data computing, given the very heavy storage requirements for images. For this reason it is not possible to exchange image data between different nodes. The transfer time and cost will largely offset the advantage deriving from distributed computation.

A reasonable configuration for EO data computation in a GRID environment can be described as follows:

- The image database should be distributed and properly ordered among the different computation centres belonging to the grid.
- The number of nodes for each centre should be sized in order to offer adequate computational power to process locally EO images.
- In order to guarantee the privacy and security of the data and of the systems, and to reduce the number of IPs needed, the GRID could be built with a number of private networks (VPN) with private IP, protected by tunnels and accessible by only one or very few WEB Portals or input computers, using SSH and Certification authority.

In this work we propose and tested a Grid approach to generate thematic map of chlorophyll from MODIS data. The tests shown that the proposed approach offers all expected capabilities and flexibilities of EO GRID environment.

Computing time for one pixel	28 min
Total number of processed pixel	850
Sequential computing time	538.5 hours about 22 days
Grig computing elements	19
Total Grid Computing time	20 hours

Tab.3 comparison between sequential execution time and grid based approach for a real application.

5 Acknowledgment

This work is partially supported by the “New. Ton.” Project, co-financed by the ERDF through the INTERREG IIIB ARCHIMED Programme A.1.180 (http://www.newton.interreg.net)

References:

[1] Matarrese R., Vermote E., Kearney M., “Impact of Hurricane Isabel on the water properties of the Chesapeake Bay area” *Proc. Of Hurricane Isabel in Perspective Conference, November 15-17, 2004, Linthicum Heights, Maryland.*

[2] Matarrese R., De Pasquale V., Guerriero L., Morea A., Pasquariello G., Umgiesser G., Scroccaro I., Alabiso G. “Comparison Between Remote Sensed Data and in situ Measurements in Coastal Waters: the Taranto Sea Case.” *Chemistry and Ecology, vol. 20, No. 3, pp. 225-237, June 2004.*

[3] Matarrese R., De Pasquale V., Pasquariello G., Chiaradia M.T., Tijanj K., “Chlorophyll-a concentration measure in coastal water using MERIS and MODIS data”, IEEE *IGARSS 2004, Proc. Of Igarss 2004, September 20-24, 2004, Anchorage, Alaska.*

[4] De Pasquale V., Matarrese R., Morea A., Chiaradia M. T., Pasquariello G., "Retrieval chlorophyll-a concentration in the Taranto coastal areas using remote sensed data", *IEEE IGARSS 2003, Proc. of IGARSS 2003, July 21-25 2003, Toulouse, France.*

[5] Matarrese R., De Pasquale V., Guerriero L., Morea A., Pasquariello G., Scroccaro I., "Application of dataset from atmospheric and oceanic EO satellites for coastal water studies", *IEEE IGARSS 2003, Proc. of IGARSS 2003, July 21-25 2003, Toulouse, France*

[6] Yuanzhi Zhang, Sampsa S. Koponen, Jouni T. Pulliainen, Matrtti T. Hallikainen, "Application of Empirical Neural Networks to Clorophyll-a Estimation in Coastal Waters Using Remote Optosensors", *IEEE Sensors Journal, Vol 3, N4, August 2003.*

[7] Gen-Tao Chiang, Martin Dove, Stuart Ballard, Charles Bostater, Ian Frame," A Grid Enable Monte Carlo Hyperspectral Syntetic Image Remote Sensing Model (GRID-MCHSIM) for Coastal Water Quality Algorithm", *Proc. Of SPIE Vol. 6360 636009-1 2006.*

[8] Matarrese Raffaella, PhD Thesis, University of Bari, "Application of remote sensing and numerical techniques GIS integrated to monitoring coastal waters", May 2003.

[9] Kaufman, Y.J., Tanre, D., Remer, L., Vermote, E., Chu, A., Holben, B., „Operational Remote Sensing of Tropospheric Earosol Over Land from EOS Moderate Resolution Imaging Spectroradiometer", *Journal of Geophisics Research, Vol. 102, N. D14, 1997.*

[10]Mobley, C., Sundman, L.K., Davis, C.O., Montes, M., Bissett, W.P., "A Look Up Table Approach to Inverting Remotely Sensing Ocean Color Data", *Ocean Optics XVI, Santa Fe, New Mexico, 2002.*

[11] Anderson, G.P., Kneizys, F.X., Chetwynd, J.H., Wang, J., Hoke, M.L., Rothman, L.S., Kimball, L.M., McClatchey, R.A.", FASCODE/MODTRAN/LOWTRAN: Past/Present/Future, *18th Annual Review Conference on Atmospheric Transmission Models, 1995.*

[12] Vermote, E., Tanre, D., Deuze, J.L., Herman, m., Morcrette, J.J., "Second Simulation of the Satellite Signal in the Solar Spectrum (6S)", *User guide, 1995.*

[13] Jiakui Tang et al., " A new approach to generate the look-up table for aerosol remote sensing on Grid Plattaform ",Geoscience and Remote Sensing Symposium, 2004. IGARSS apos;04. *Proceedings. 2004 IEEE International Volume 5, Issue , 20-24 Sept. 2004 Page(s): 2949 - 2952 vol.5*

[14] S. Y. Kotchenova, E. F. Vermote, R. Matarrese, and F. J. Klemm, Jr., "Validation of a New Vector Version of the 6S Radiative Transfer Code for Atmospheric Correction of MODIS Data (Part I)", Applied Optics, Vol. 45, Issue 26, pp. 6762-6774.

[15] Modis design concept: http://daac.gsfc.nasa.gov/MODIS/products.shtml

[16] I. Foster, C. Kesselman, S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations", *International J. Supercomputer Applications, 15(3), 2001.*

[17] http://www.globus.org/toolkit/docs/4.0/

[18] I. Foster, H. Kishimoto, A. Savva, D. Berry, A. Djaoui, A. Grimshaw, B. Horn, F. Maciel, F. Siebenlist, R. Subramaniam, J. Treadwell, J. Von Reich, "The Open Grid Services Architecture, Version 1.0.", *Informational Document, Global Grid Forum (GGF), January 29, 2005.*

[19] Open Grid Services Architecture–Data Access and IntegrationProject(OGSA-DAI),http://www.ogsadai.org.uk/