Sensor Web and Data Mining Approaches for Harmful Algal Bloom Detection and Monitoring in the Gulf of Mexico Region

Balakrishna Gokaraju, Surya S. Durbha, Roger L. King, and Nicolas H. Younan
Department of Electrical and Computer Engineering
GeoResources Institute (GRI)
Mississippi State University, Mississippi State, MS 39762-9571, USA
balu@gri.msstate.edu, suryad@gri.msstate.edu, rking@engr.msstate.edu, and younan@ece.msstate.edu

Harmful Algal Blooms (HABs) belong to the class of phytoplankton but are toxic in nature and occur due to a combination of physical forces and fast growth blooms. In the Gulf of Mexico, HABs are known as "red tides" because they frequently color coastal and ocean waters with a red color. Economic effects of HABs in the U.S. are at least $82 million annually. The Governor's action plan for the Gulf of Mexico emphasizes the need for improving the detection and forecasting of harmful algal blooms (HABs) in the Gulf of Mexico and to also better understand the public health and socio-economic effects of the bloom events. The HAB forecasting system from NOAA (http://tidesandcurrents.noaa.gov/hab/) provides valuable information on HAB's, the HAB SOS (http://habsos.noaa.gov/) combines a interactive map with a geographic information system (GIS) that collects, stores, and displays various data layers for target user groups. The target user groups are Federal, state, and local resource and environmental managers and scientists.

However, the lack of standardization of the data exchange mechanism to encapsulate, store, exchange and configure these data sets and the information that goes into the monitoring and forecast is an impediment for its wider (large scale) integration into relevant regional wide decision support systems. Hence, there is a need to improve the capabilities of Gulf-wide monitoring networks to provide the bloom information needed for local coastal managers to minimize HAB effects.

To establish such a data exchange mechanism and facilitate the sharing of sensor data, observations and measurements, model and analysis tools, we intend to adapt the Services Oriented Architecture (SOA) of the OGC Sensor Web Enablement Framework (http://www.opengeospatial.org/projects/groups/sensorweb). Sensor webs are distributed, organized systems of sensor, computing and storage nodes, interconnected by a communications framework that behaves as a single, coherent system. The HAB's Virtual Sensor Web (HABSW) will be an intelligent and integrated virtual observation network that provides timely, on-demand data and analysis to users.

In addition, sensor web data mining based on machine learning approaches are used to investigate the relationships between the various elements of the information that go into the analysis and forecast of HAB's. In particular the continuous ocean color data (i.e. chlorophyll 'a', chlorophyll pigment, chlorophyll anomaly, attenuation constant K_d, Chl-fuorescence line height, carbon dissolved organic matter and sea surface temperature) that is available from SeaWIFS, MODIS and MERIS sensors and other relevant *In situ* data is be used to implement a time series classification of satellite ocean color data using support vector machines along with a kernel principal components (KPCA) technique for feature reduction. The Methodology consists of applying various feature extraction algorithms to obtain features from the ocean color data,

followed by a Kernel Principal Component Analysis (KPCA) driven feature reduction technique. The support vector machine is used as the classifier. The most common constraint associated in classification of HABs is the sediments resuspension along the coast. This is addressed by using spatial filters or by area-perimeter ratios as they have unique spatial feature that are different from HABs. The *In situ* data is used for validation. Various statistical measures are used to evaluate the retrieval accuracy.