Consistent Ocean Color and its Assimilation in Ocean Models

Vladimir Krasnopolsky[1], Sudhir Nadiga[2], Eric Bayler[3], Avichal Mehra[1], David Behringer[1]
[1]NWS/NCEP/EMC, [2]IMSG at NWS/NCEP/EMC, [3]NOAA/NESDIS/STAR
Email: vladimir.krasnopolsky@noaa.gov; Phone: 301-683-3742

Satellite remote-sensing of ocean color (OC) parameters provides a means for broadly observing the foundation of the biological component of the world's oceans. Consequently, this data needs to be exploited for the analysis and prediction of ocean bio-physical processes and initiating the biogeochemical path, through primary productivity and associated processes and natural cycles, to ocean ecological forecasts. Operational integration/assimilation of OC fields (chlorophyll, Kd490, KdPAR) into NOAA's operational ocean models has three fundamental requirements/conditions: 1) gaps in the observations need to be addressed, both in the current instance and for extended gaps; 2) the data being assimilated must have a long data record for establishing a robust statistical database that spans multiple seasons; and 3) for assimilation, the data must be for a predicted parameter.

In previous work [1], we demonstrated that a neural network (NN) technique can successfully fill both short and small gaps (several days and several grid points), as well as extended gaps (several months and global) in satellite OC measurements. In this work, we show that the other two principal requirements can also be satisfied using the NN technique.

Consistent Ocean Color

Three major OC data sets exist, produced by the SeaWiFS (09/1997 – 12/2010), MODIS (07/2002 – present), and VIIRS (1/2011 – present) sensors. These three data sets have different error statistics; therefore, it is not a simple task to integrate them into a single consistent long-term data set. One possible approach is examined here. Using three years of VIIRS data (the most accurate and recent measurement), we trained an ensemble of NNs. Each of the NN ensemble members performs a mapping of relevant ocean variables (SST, SSH, and the upper-ocean portions of vertical temperature and salinity profiles) to the logarithm of chlorophyll-a concentration, *C*, which can be expressed as:

$$LC = \log_{10} C = NN(SST, SSH, \vec{T}, \vec{S}, lat, lon, doy) \qquad (1)$$

where $\vec{T}$ and $\vec{S}$ are the upper-ocean portions of the temperature and salinity profiles, and *doy* is the day of the year. Using a logarithm of *C* as the NN output, rather than *C*, produces a more accurate NN approximation and extrapolation of VIIRS data. When training the NNs, the mean square error function is used; however, this error function is optimal for normally distributed data. Chlorophyll data have an almost log-normal distribution (see Fig. 1); thus, $\log_{10} C$ is nearly normally distributed (Fig.1). Using three years of VIIRS data for training with $\log_{10} C$ as the NN output produced an ensemble of NNs capable of stable long-range extrapolation of chlorophyll values.

For signatures of upper-ocean dynamics this study employs satellite-derived surface variables (sea-surface temperature (SST), sea-surface height (SSH)), and gridded ARGO salinity and temperature profiles of the top 75m depth. Chlorophyll fields from NOAA's operational Visible Imaging Infrared Radiometer Suite (VIIRS) are used. The NNs are trained using data for three years (2012 through 2014) and assessed for a period of 10 years (2005 through 2014). To reduce noise in the data and to obtain a stable computation of the NN Jacobian for sensitivity studies and data assimilation [2], an ensemble of NNs was constructed. Results are assessed using the

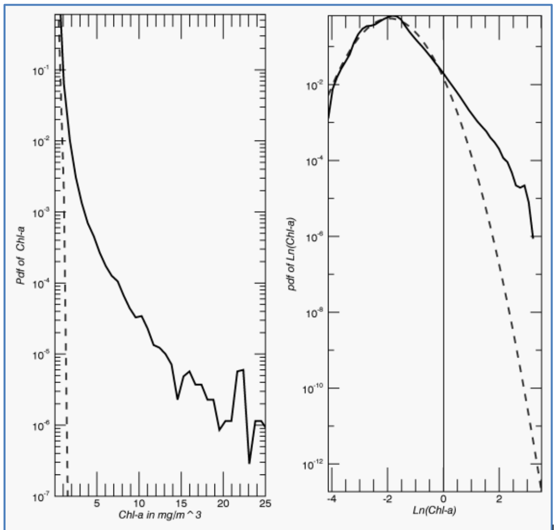

Fig.1 Probability density function for C (left panel and for $\ln C$ (right panel) shown by solid lines. In both panels, dashed lines show normal distribution with the same means and standard deviations.

root-mean-square error (RMSE) metric and cross-correlations between observed chlorophyll fields and NN output. Chlorophyll measurements from the three different OC sensors (SeaWiFS, MODIS, and VIIRS) available during the validation period were used. Fig. 2 presents the validation results. The correlations between the NN-generated and observed C decrease slightly from ~0.85 to ~ 0.75 when moving away from the training interval (2012 through 2014). RMS differences, however, do not change significantly. Results for all three used satellite sensors are very consistent. It means that NN generated C can serve as consistent long term OC data for different uses, including assimilation in to ocean models.

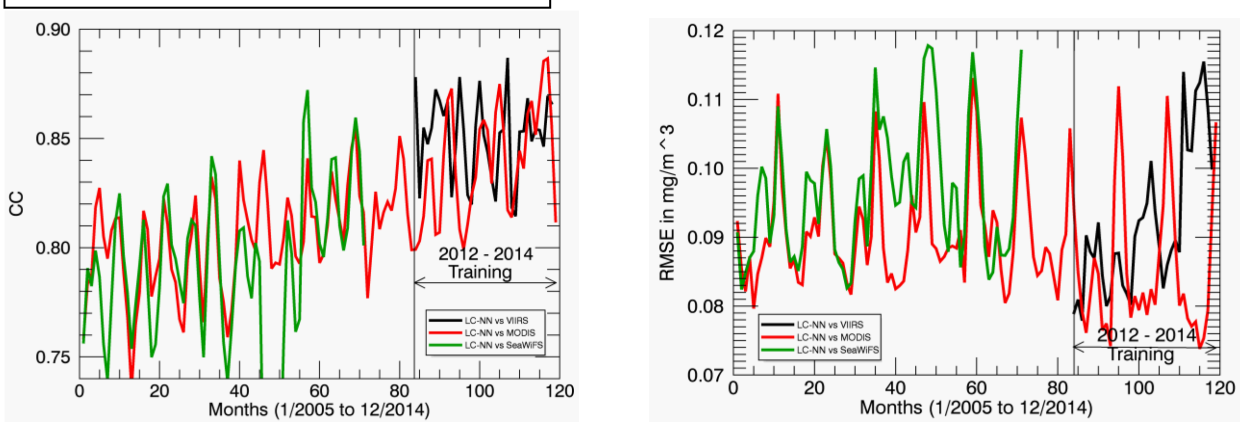

Fig.2 Correlation (left panel) and RMS differences (right panel) between C produced by ensemble of NNs (trained on 3 years of VIIRS data) and C observed by VIIRS (black curves)), MODIS (red curves), and SeaWiFS (green curves).

Assimilating ocean color parameters into ocean models

OC parameters are not prognostic variables in current oceanic models; therefore, OC assimilation requires the coupling of a biochemical component or introducing an observation operator relating C to ocean prognostic variables into the data assimilation system. The NN presented in Eq. (1) can serve as such an operator. The Jacobian of NNs (Eq. 1) can relate innovations in C to innovations in ocean prognostic variables in the data assimilation system [2].

References

[1] Krasnopolsky V., S. Nadiga, A. Mehra, E. Bayler, and D. Behringer, 2015. "Neural Network Technique for Filling Gaps in Satellite Measurements: Application to Satellite Ocean Color Observations", Computational Intelligence and Neuroscience, Article ID 923230, December, http://www.hindawi.com/journals/cin/aa/923230/

[2] Krasnopolsky V., 2013, "The Application of Neural Networks in the Earth System Sciences. Neural Network Emulations for Complex Multidimensional Mappings", Springer, 200 pp.