U. S. Department of Commerce
National Oceanic and Atmospheric Administration
National Weather Service
National Centers for Environmental Prediction

Office Note 487
http://doi.org/10.7289/V5/ON-NCEP-487

Optimization of a Neural Network-Based Biological Model for Chlorophyll-a Concentration in the Upper Ocean

Vladimir Krasnopolsky[1], Sudhir Nadiga[2], Avichal Mehra[1], Eric Bayler[3], and Hae-Cheol Kim[2]

[1]NWS/NCEP/EMC, [2]IMSG at NWS/NCEP/EMC, [3]NESDIS/STAR

February 2017

MMAB Contribution No. 330

Email: Vladimir.Krasnopolsky@noaa.gov; Phone: 301-683-3742

Contents

Abstract

Integrating/assimilating satellite ocean color (OC) fields (chlorophyll-a, Kd_{490}, Kd_{PAR}) in NOAA's operational ocean models requires scientifically consistent and robust techniques to obtain long (several years long) time series of OC fields. In addition, in order to dynamically take into account the OC signal in ocean and coupled climate models, a biological model for components of OC is required. In this work, we introduce one possible approach based on a Neural Network (NN) technique, linking Chl-a variability — which is primarily driven by biological processes — with the physical processes of the upper ocean, using NN-based biological model for Chl-a. A NN method for correlating satellite OC fields with other assimilated satellite and *in situ* observations: a) instigates fewer assimilation errors (since the inputs to the NN are already being assimilated), b) reduces reliance on sparse *in situ* OC observations, and c) provides a dynamical feedback between biological and physical processes in the upper ocean in ocean and coupled climate models. In this study, satellite-derived surface variables — sea-surface temperature (SST), sea-surface height (SSH), and sea-surface salinity (SSS) fields — and gridded ARGO salinity and temperature profiles from 0 – 75m depth are employed as signatures of upper-ocean dynamics. OC Chl-a fields from NOAA's operational Visible Imaging Infrared Radiometer Suite (VIIRS) are used, as well as MODIS and SeaWiFS Chl-a concentrations. Different methods of optimization of the NN technique are investigated. Results are assessed using the root-mean-square error (RMSE) metric and cross-correlations between observed OC fields and NN output. To reduce the impact of the noise in the data and to obtain a stable computation of the NN Jacobian, an ensemble of NN with different weights is constructed. The results for the ensemble mean are compared with those for a single NN. This study demonstrated that the NN technique provides an accurate, computationally cheap method to generate long (up to 10 yearlong) time series of consistent Chl-a concentration, which are in good agreement with Chl-a data observed by different OC sensors during this period. It is noteworthy that a single NN (or a single NN ensemble) is capable of generating OC fields all over the globe (at all grid points of the global grid). Also, the accuracy of NN prediction does not deteriorate during the validation period: NN trained on three years of data (2012 and 2013) performs well during the ten years (2005 - 2014) validation period. These results demonstrate a very good generalization ability of the NN both in terms of spatial and temporal generalization. It means that the NN-based empirical biological model for Chl-a can be used in the oceanic models and coupled climate prediction system to dynamically take into account biological processes in the upper ocean.

I. Introduction

Operational integration/assimilation of ocean color (OC) fields (chlorophyll, Kd490, Kd_{PAR}) into ocean models has a significant positive impact on predictive skills of climate models (Anderson *et al*. 2009, Ballabrera-Poy *et al*. 2007, Morel and Antoine 1994, Murtugudde *et al.* 2002, and Zhang *et al.* 2009). For successful assimilation, the OC data have to satisfy three fundamental requirements/conditions. First, gaps in the observations, spatial and temporal, need to be filled; second, data assimilation must be for a predicted parameter (prognostic variable) or a parameter explicitly related to a prognostic variable, and third, the data being assimilated must have a long data record to facilitate compilation of a robust statistical database spanning multiple seasons.

In our previous works (Krasnopolsky et al. 215 and Nadiga et al. 2016) we addressed the first and second items. A new, based on neural network (NN) technique, approach was developed that allowed us to fill large spatial (up to the global size) and temporal (up to yearlong) gaps in Joint Polar Satellite System (JPSS) Visible Infrared Imaging Radiometer Suite (VIIRS) satellite OC chlorophyll-a (Chl-a) fields. This NN also provides linkages between a signature of biological processes, i.e. satellite-derived OC fields, and signatures of upper-ocean physical processes, i.e. prognostic variables. Thus, the NN can be used as an empirical ecological model to relate Chl-a to prognostic variables in ocean model in data assimilation systems.

In this study, in Section II, we investigated different approaches to optimizing the NN technique. We evaluate the impact of (i) extending the training set, (ii) optimization of NN inputs, and (iii) optimization of outputs on the performance of the NN, emulating Chl-a concentration, and on ability of the NN to generate a long consistent time series of Chl-a concentrations. In Section III, we present some discussion and conclusions.

II. Optimization of NN performance.

In this study, as in our previous works (Krasnopolsky 2013, Krasnopolsky et al. 215, and Nadiga et al. 2016), we used NN multilayer perseptron,

$$y_q = a_{q0} + \sum_{j=1}^{k} a_{qj} \cdot \tanh(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i); \quad q = 1,2,\dots,m \qquad (1)$$

where x_i and y_q are components of the NN input and output vectors X and Y, respectively, a and b are fitting parameters (NN weights).

Several options are available for optimization of NN performance:

- The NN training set can be extended to expose NN to greater variety of input patterns. An independent validation set can be extended to better understand limitations of trained NN and improve the NN performance during the next NN training. Also, the validation data set can be enriched by data for the same parameter measured by different instruments/sensors to evaluate cross sensors consistency of the data simulated by the NN.
- The set of NN inputs can be optimized by removing some not important inputs. The importance of inputs can be determined by sensitivity study (e.g., Krasnopolsky et al. 215) that reveals the sensitivity of outputs to different inputs.
- A transformation of output variable can be performed to reconcile statistical properties of outputs with the error function that is used for training.

All aforementioned options are investigated in the following Sections.

II.1 Extending training and enriching validation sets

Fig. 1 shows the satellite and *in situ* data available for NN training and validation. The daily VIIRS data on 1° by 1° global grid are available during the period of three years (2012 to 2014). In our previous study (Krasnopolsky et al. 2015), we used first two years (2012 and 2013 or 730 days) of daily data (~ 20,000,000 grid points or records) for NN training and test. The data for 2014 (365 days) were left as a validation set for validating trained NNs and estimating its prediction (generalization) capability. It was shown that two years of training data are sufficient

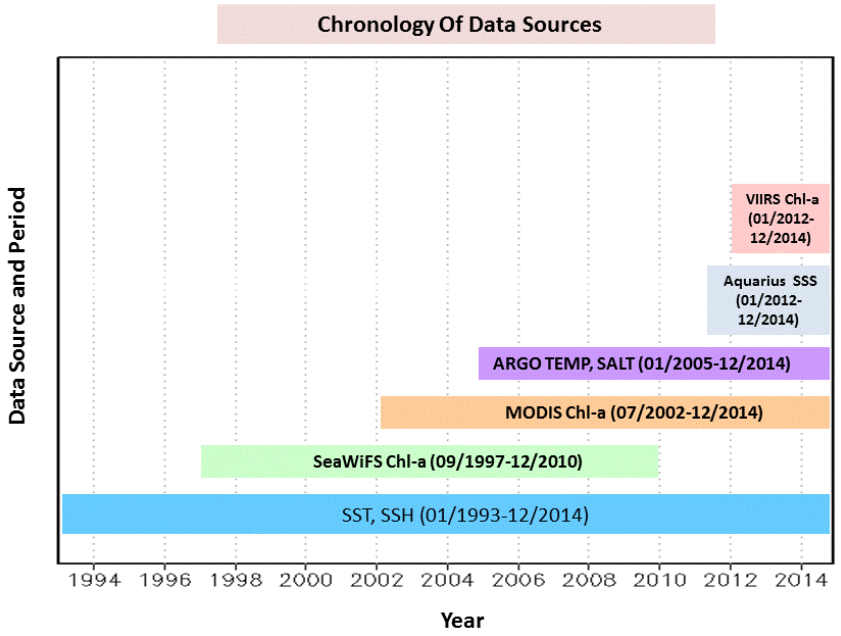

Figure 1. Satellite and in situ data available for NN training and validation.

for at least one year of accurate forecast of Chl-a concentration. In this study, we use all three years of daily VIIRS data for training and testing sets. For validation the MODIS (ten years, from 2005 to 2014) and SeaWiFS (six years, from 2005 to 2010) data were used. The time period starting from 2005 was selected because the ARGO profiles that are mostly important inputs of the NN, simulating Chl-a concentration (Krasnopolsky et al. 2015, Nadiga et al. 2016), are available starting 2005. We used NOAA SSH and SST fields and NASA Aquarius mission SSS fields. The SSH/SST/SSS fields were spatially and temporally averaged in the same way

as Chl-a data. Thus, daily averaged and gridded (on 1° × 1° grid) data were used for the NN training. Results are assessed using the mean error (bias), root-mean-square error (RMSE), and cross-correlations between observed and NN generated OC. To reduce the impact of the noise in the data and to calculate the NN Jacobian for sensitivity studies, an ensemble of NNs with different weights was developed.

To support using SeaWiFS and MODIS data for validation of NNs trained on VIIRS data and to evaluate the differences between the two sensors, we compared MODIS data with VIIRS data during the period of their overlap (2012 to 2014) and MODIS data with SeaWiFS data during the period of their overlap (2005 to 2010). The results are shown in Figures 2 to 4.

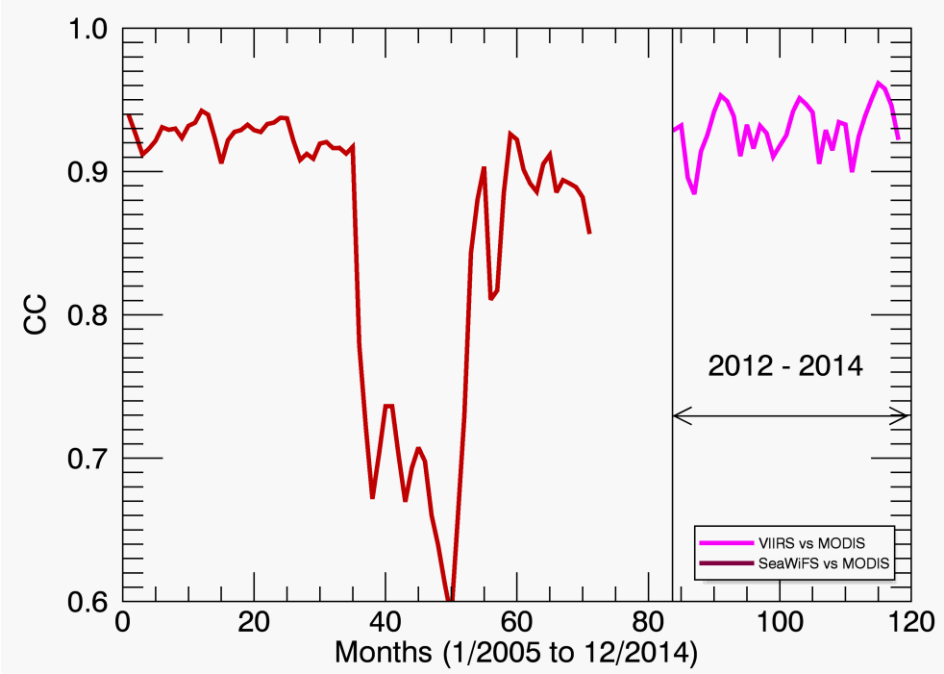

Figure 2. Correlation between MODIS and VIIRS and MODIS and SeaWiFS data.

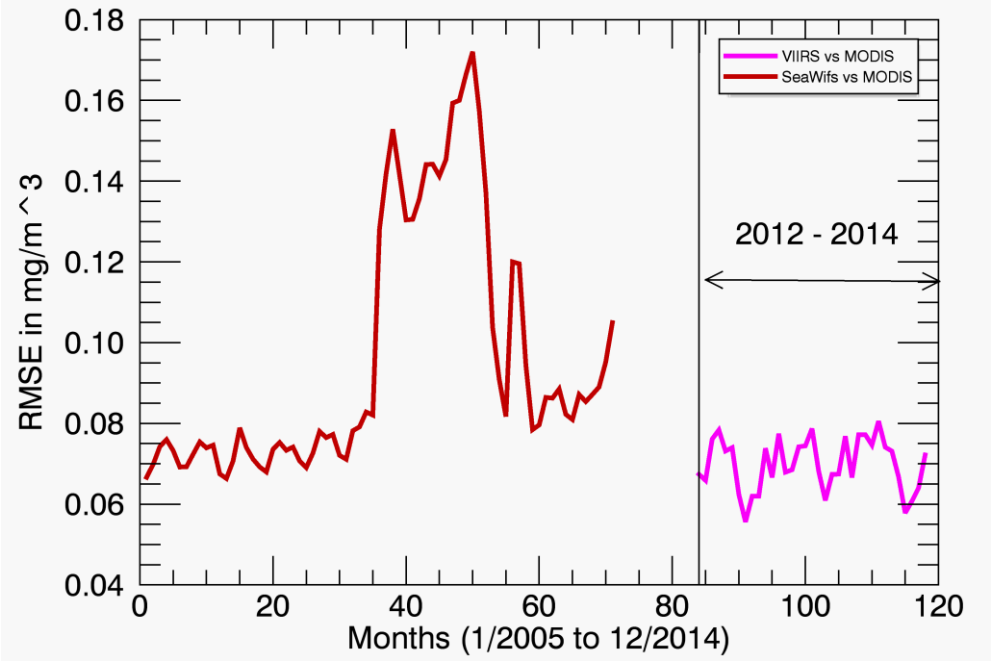

Figure 3. RMS differences between MODIS and VIIRS and MODIS and SeaWiFS data.

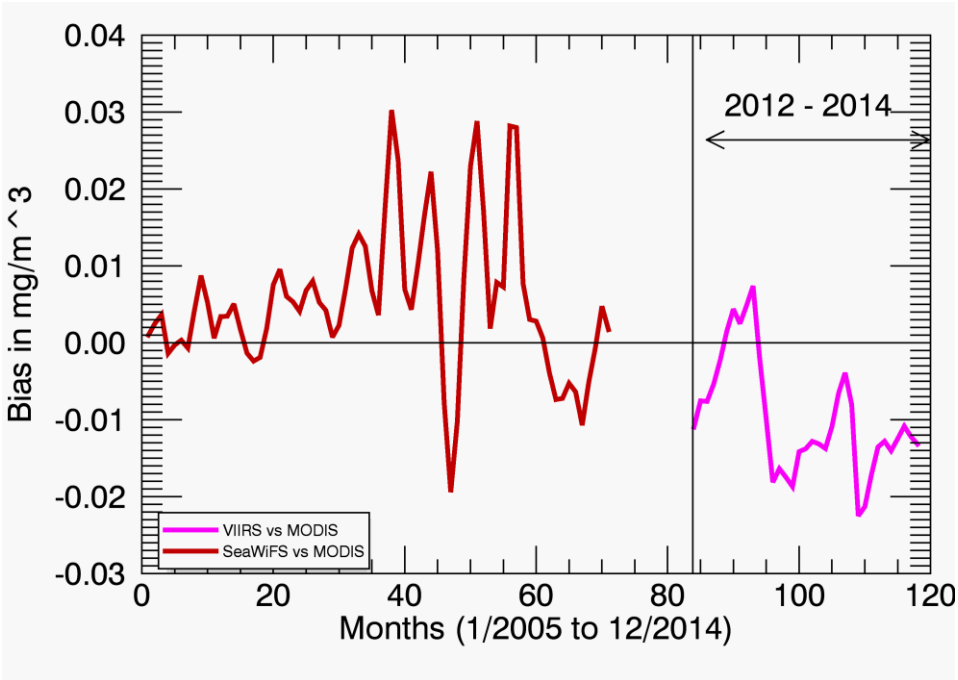

Figure 4. Mean differences between MODIS and VIIRS and MODIS and SeaWiFS data.

Fig. 2 demonstrates a high level of correlation (> 0.9) between VIIRS and MODIS data. The correlation between MODIS and SeaWiFS has the same order of magnitude everywhere except several months in 2008 when SeaWiFS has problems.

Figs. 3 and 4 provide estimates of RMS and mean differences between the sensors that will be used below for comparison. Comparison of sensors shows that they are close enough except for periods of SeaWiFS problems and can be used for validating the NN ensembles, taking into account the aforementioned estimates of their differences.

Finally, two NN ensembles, each one consisting of six NNs ensemble members, was trained; one ensemble, using two years of daily VIIRS data and another one, using three years of data. All ensemble members have the same architecture: 23 inputs, 30 hidden neurons in one hidden layer, and 1 output; the NN ensemble members were trained using different initial values for NN weights, a_{ij} and b_{ij}, in (1). All following NN results correspond to the ensemble averages. Fig. 5 shows Taylor Diagram (Taylor 2001) for the ensemble trained, using three years of data. This figure shows that, as we demonstrated in our previous work (Krasnopolsky et al. 215), the ensemble members have a significant spread and the ensemble average improves results as compared with a single ensemble member. Actually, seven ensemble members were trained (as shown in Fig. 5); however, the member #1 was removed as outlier.

Fig. 6 shows correlation of NN simulated Chl-a with MODIS data for two NN ensembles, one trained with two year training set and another one – with three year training set. The figure demonstrates that three years of data provide sufficient information for training NN ensemble, and that the performance of the NN ensemble (trained on three years of data) only slowly

deteriorates during seven years validation period (2005 to 2011). The NN ensemble trained on two years of data does not demonstrate sufficient stable level of performance during this period.

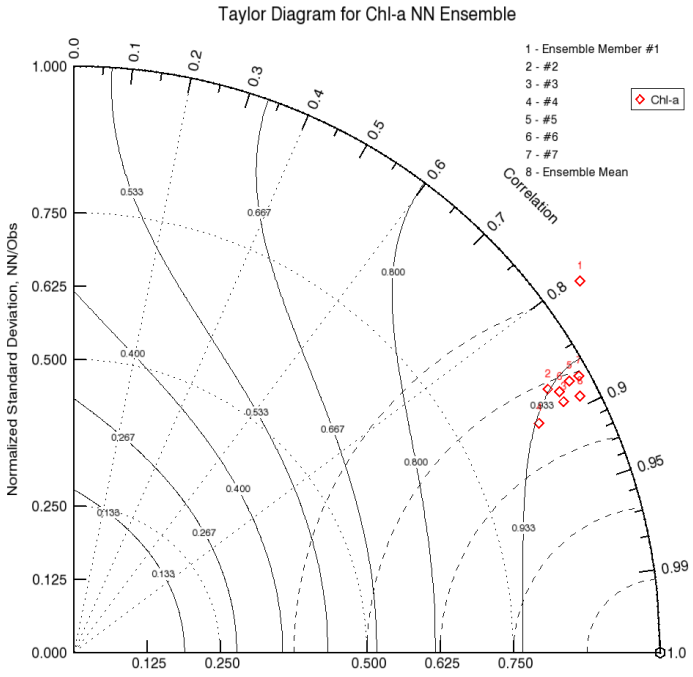

Figure 5. Taylor Diagram for the ensemble trained, using three years of data.

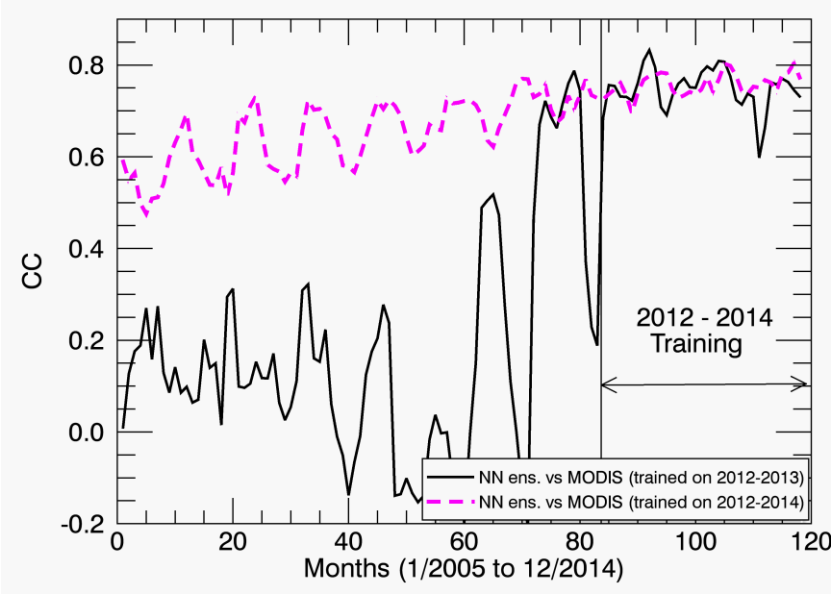

Figure 6. Correlation between Chl-a simulated by two NN ensembles. Black curve shows results for NN ensemble trained on two years (2012 and 2013) and pink curve – for NN ensemble trained on three (2012 to 2014) of dayly VIIRS data.

We compared NN ensemble trained, using three years of daily VIIRS, with MODIS and SeaWiFS data. The daily VIIRS and NN data were averaged to get monthly data and compared them with monthly MODIS and SeaWiFS data. Fig. 7 shows comparisons for monthly mean Chl-a concentrations. Fig.8 shows comparisons of correlations between monthly mean Chl-a concentrations simulated by NN ensemble and VIIRS, MODIS, and SeaWiFS data.

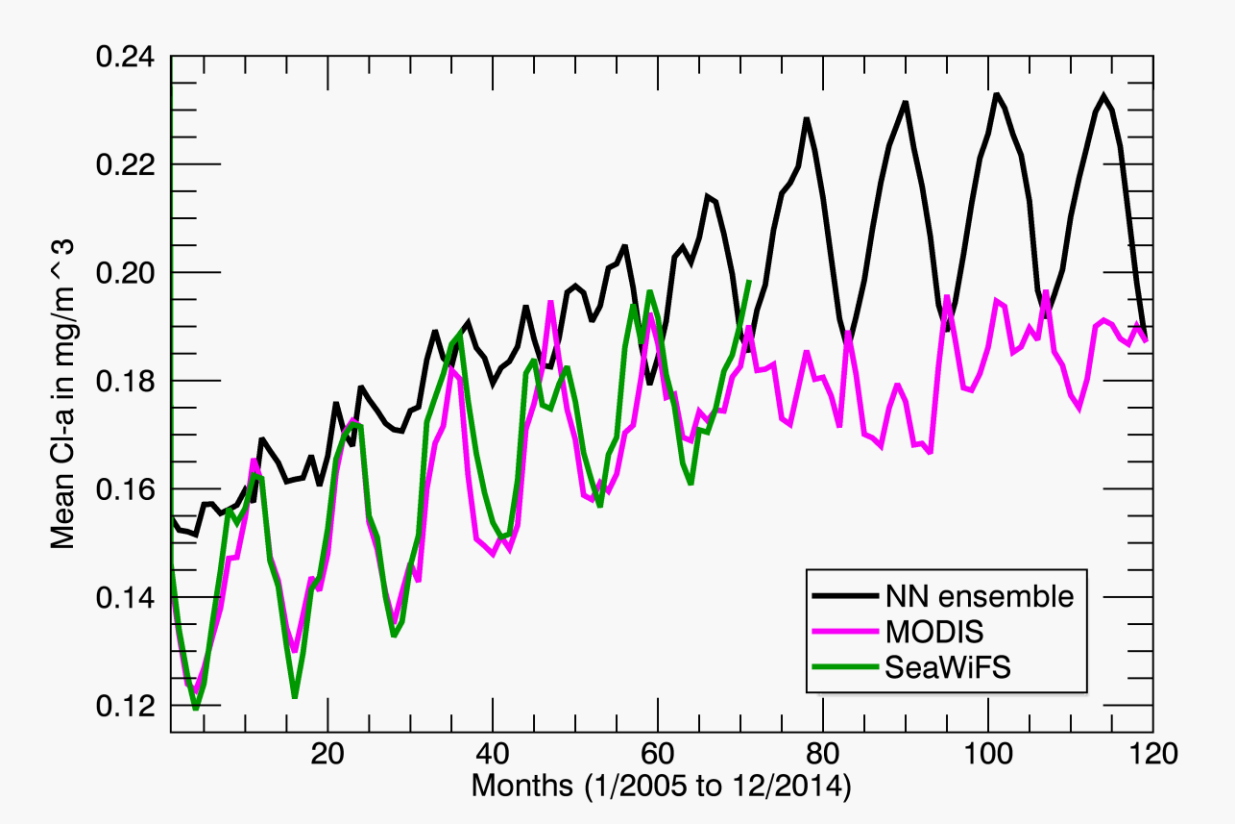

Figure 7. Monthly mean Chl-a concentrations simulated by NN ensemble (black curve), pink curve – MODIS data, and grin curve – SeaWiFS data.

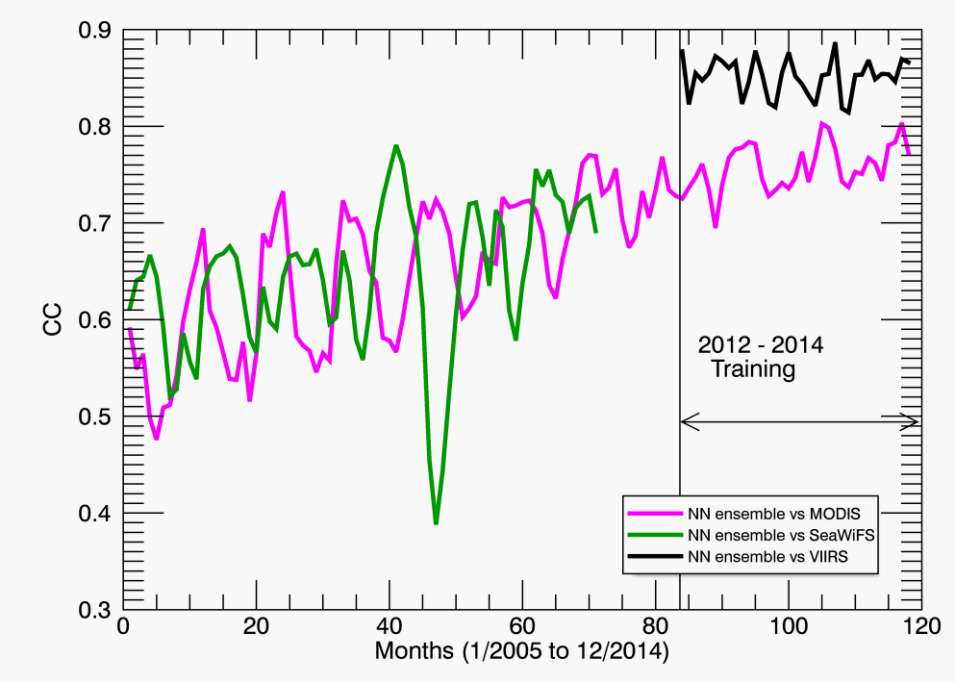

Figure 8. Correlation between monthly mean Chl-a concentrations simulated by NN ensemble and VIIRS data (black curve), pink curve – with MODIS data, and grin curve –with SeaWiFS data.

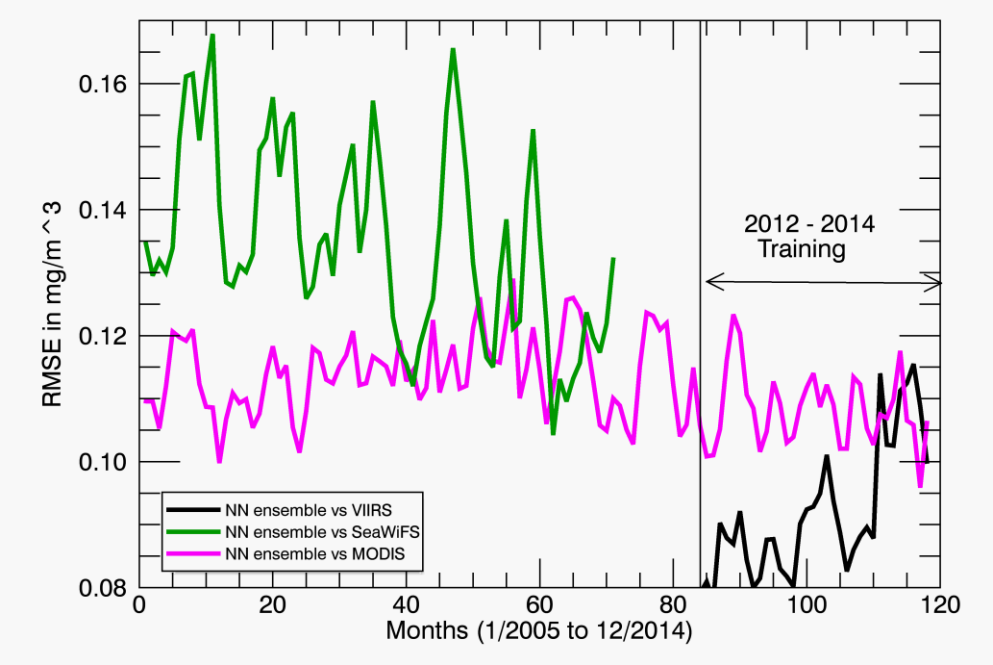

Figure 9. RMSE in monthly mean Chl-a concentrations simulated by NN ensemble vs. VIIRS data (black curve), pink curve – vs. MODIS data, and grin curve – vs. SeaWiFS data.

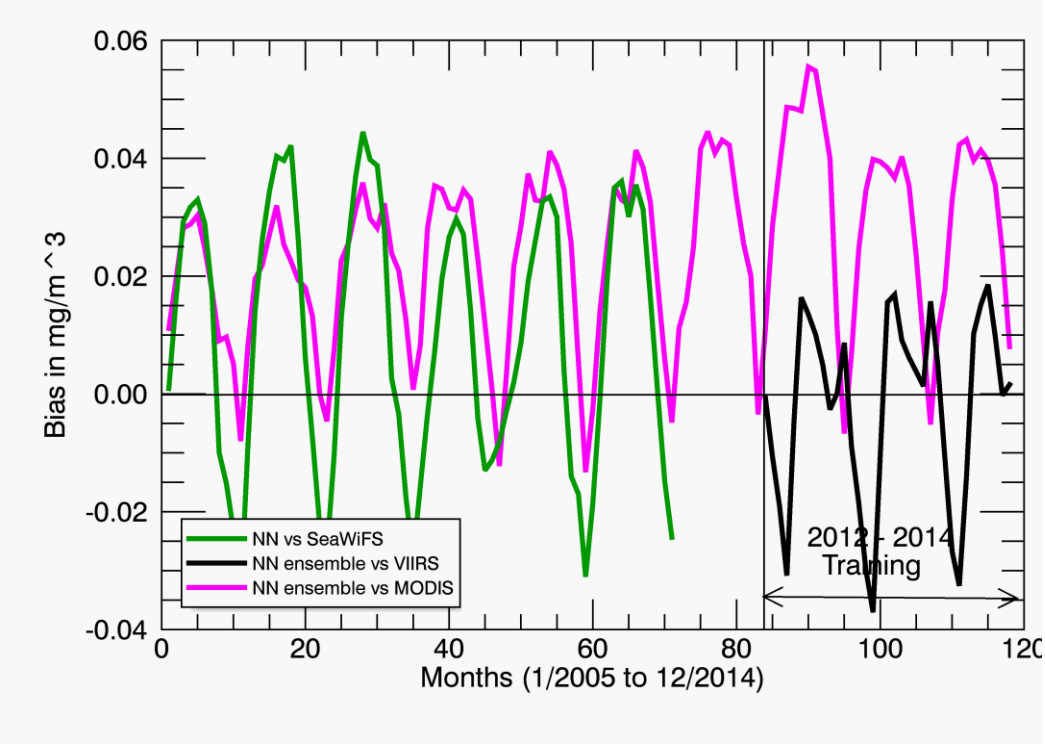

Figure 10. Bias in monthly mean Chl-a concentrations simulated by NN ensemble vs. VIIRS data (black curve), pink curve – vs. MODIS data, and grin curve – vs. SeaWiFS data.

Figs. 9 and 10 show comparisons of RMSE and bias in monthly mean Chl-a concentrations simulated by NN ensemble vs. VIIRS, MODIS, and SeaWiFS data. The presented comparisons show that the ensemble of NNs trained on three years of daily VIIRS data performs satisfactory and simulates Chl-a data that are in good agreements with the Chl-a concentrations observed by three different satellite during ten year period. Correlations and differences between NN simulated data and three satellite derived data sets are similarly small.

II.2 Optimizing inputs and outputs

Table 1. A Table of Emulating Neural Network Inputs and Outputs

Input #	Variable	Units	Input	Size
1	Year		yr	1
2	Day of the year		$\sin(\frac{2 \cdot day \cdot \pi}{366.})$	1
3	Day of the year		$\cos(\frac{2 \cdot day \cdot \pi}{366.})$	1
4	Longitude		$\sin(\ lon\ )$	1
5	Longitude		$\cos(\ lon\ )$	1
6	Latitude		$\sin(\ lat\ )$	1
7	Sea surface temperature [SST]	°C	SST	1
8 – 14	Argo salinity [S(z)]	PSS	sal	7
15 – 21	Argo temperature [T(z)]	°C	temp	7
22	Depth	m	depth	1
23	Sea surface height [SSH]	m	SSH	1
Total	All Inputs			23
Output #	**Variable**	**Units**	**Output**	**Size**
1	Chlorophyll-a	Mg/m^3	Ln(Chl-a)	1

The inputs of the NN used in this study have been slightly changed as compared with our previous works (Krasnopolsky et al. 215 and Nadiga et al. 2016) in accordance with the sensitivity study performed there. The surface salinity data have been removed from the NN input vector and ocean depth has been added to better resolve Chl-a variations with the change of the ocean depth. The final set of the NN inputs is presented in Table 1.

II.3 Optimizing the error function

In our works the mean square differences between the training data $\{Y_i\}_{i=1,\ldots,N}$ and NN outputs are used as the error function (2). This function is minimized in the process of the NN training.

$$E = \frac{1}{N}\sum_{i=1}^{N}\left[Y_i - NN\left(X_i\right)\right]^2 \qquad (2)$$

The minimization of this error function is equivalent to the basic statistical maximum likelihood principle only if probability density function (pdf) of outputs is normal (Krasnopolsky 2013). It means that if the pdf is not normal, the error function does not deliver the optimal parameters for the trained NN and this NN does not provide the optimal (the best) approximation for the training set. Therefore, to get the best results, our goal should be to make the pdf as close to normal as possible.

Fig. 11 shows two pdfs. The left panel show the Chl-a pdf (solid line). It is very far from normal (actually, it is log-normal); it has a very long tail as compared with the normal distribution (dashed line). The right panel shows pdf for the logarithm of Chl-a (Ln(Chl-a)). This function is almost normal. Thus, if in eq. (2) we use not Y_i but Ln(Y_i), the error function (2) becomes almost optimal and the NN, which now generates Ln(Chl-a), becomes almost the best approximation for the data.

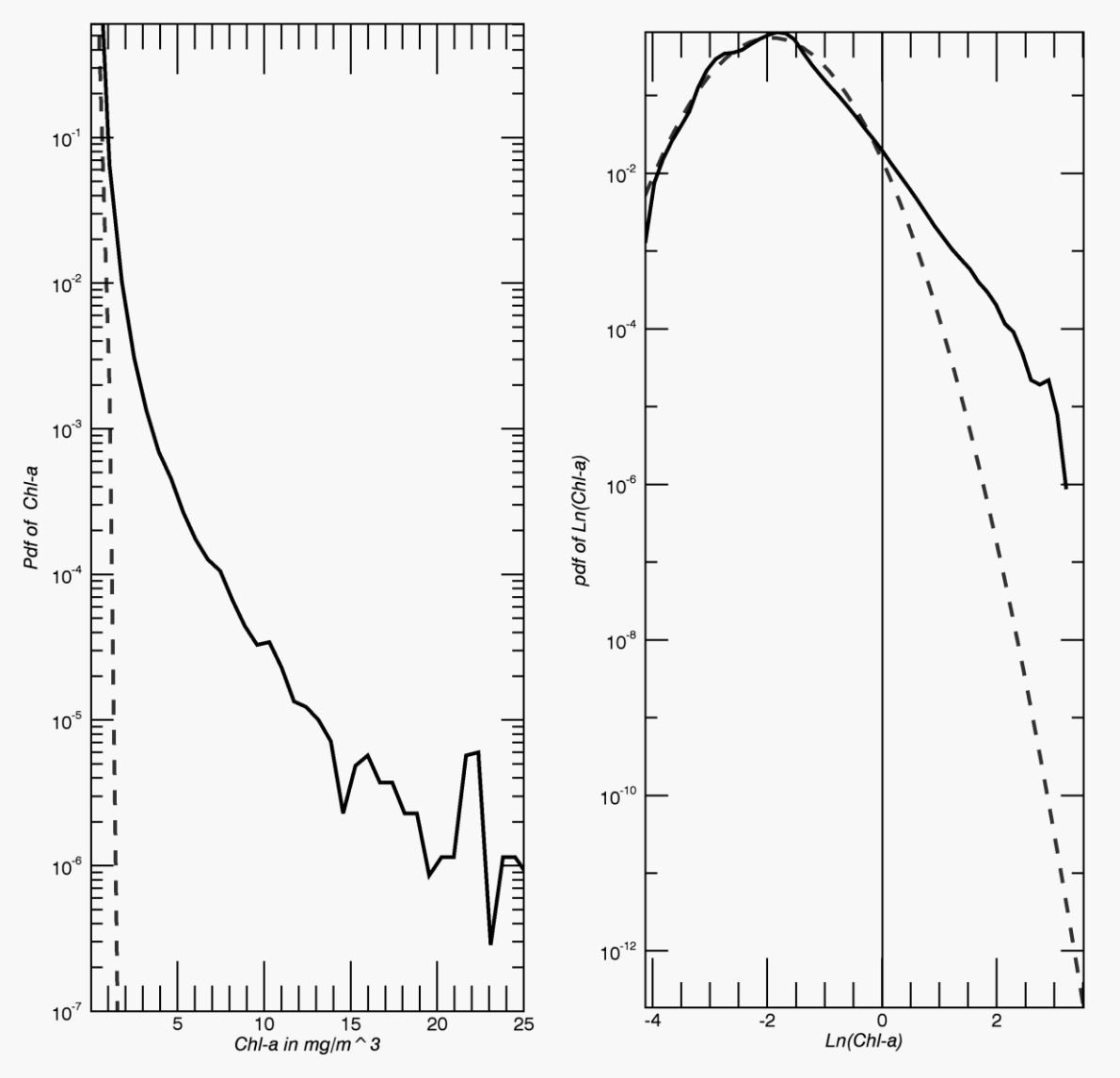

Figure 11. Pdfs of Chl-a (left panel) and Ln(Chl-a) (right panel) are shown by solid lines. For each case, the normal distributions with the same mean value and standard deviation are shown by dashed lines.

To illustrate advantages of using for NN training the error function (2), containing logarithms of the Chl-a, that is, of using NNs that have a logarithm of Chl-a as the output, an ensemble of NNs with Ln(Chl-a) as the output was trained, using three years of VIIRS data (2012 to 2014). Then the ensemble was applied to the independent test sets and monthly means were calculated. Figures 12 – 14 show comparisons of correlations and error statistics (RMSEs and biases) for two NN ensembles with: (1) ensemble member NNs with the Chl-a output (solid lines) and (2)

ensemble member NNs with the Ln(Chl-a) output (dashed lines). Ensemble means of both ensembles are compared with VIIRS (black), MOSID (red) and SeaWiFS (green) observations.

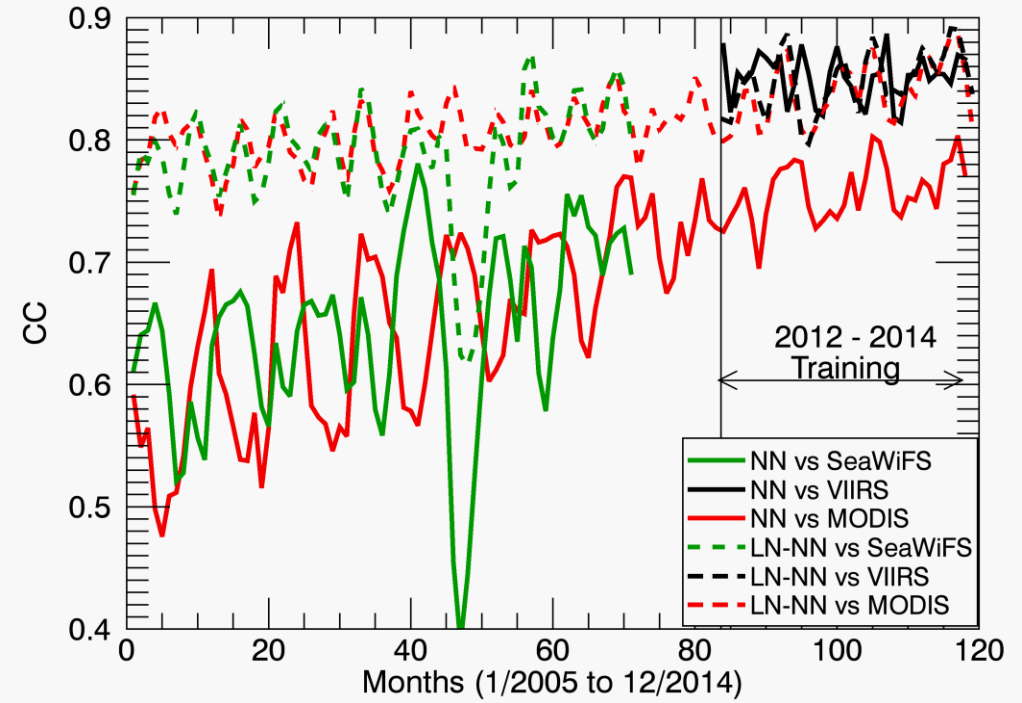

Figure 12. Correlation between monthly mean Chl-a concentrations simulated by NN (solid curves) and LN-NN (dashed curves) ensembles and VIIRS (black curves), MODIS (red curves), and SeaWiFS (grin curves) data.

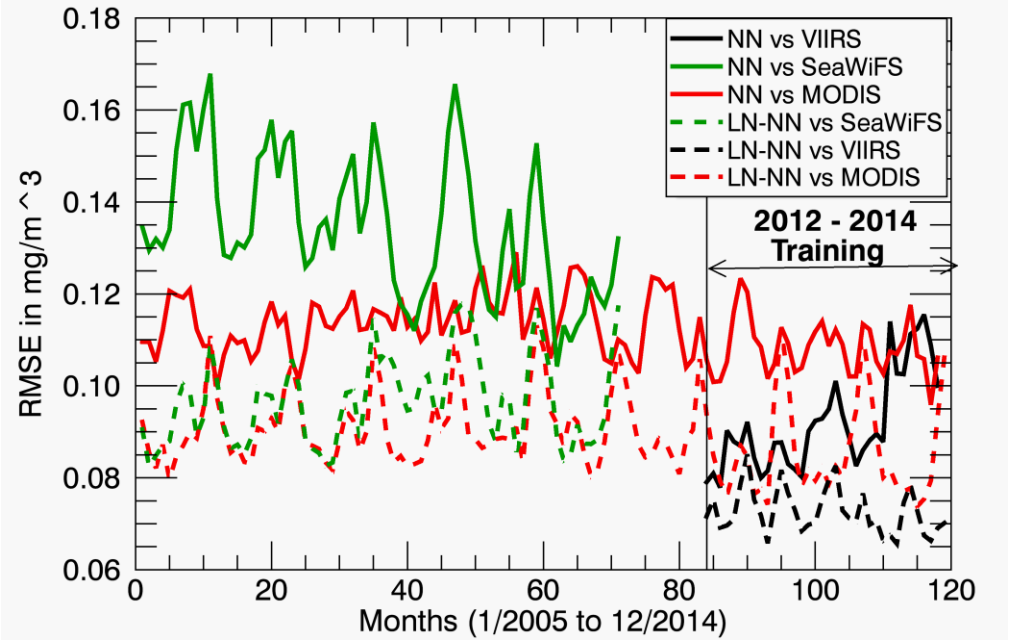

Figure 13. RMSEs of monthly mean Chl-a concentrations simulated by NN (solid curves) and LN-NN (dashed curves) ensembles vs VIIRS (black curves), MODIS (red curves), and SeaWiFS (grin curves) data.

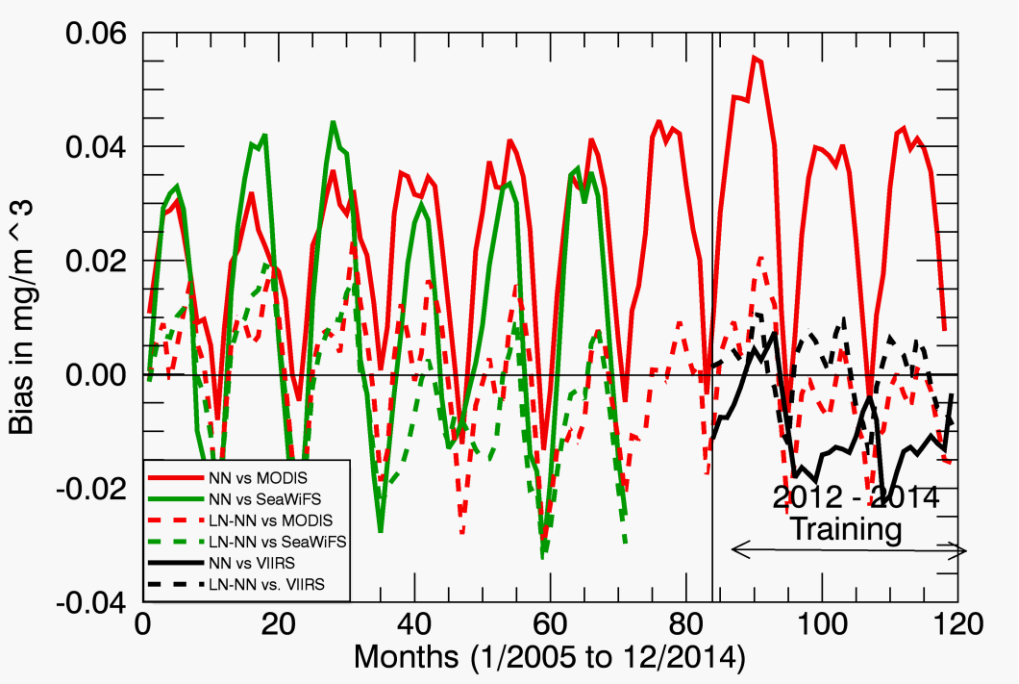

Figure 14. Biases of monthly mean Chl-a concentrations simulated by NN (solid curves) and LN-NN (dashed curves) ensembles vs VIIRS (black curves), MODIS (red curves), and SeaWiFS (grin curves) data.

Figs. 12 - 14 demonstrate significant improvements in the quality of simulated Chl-a concentrations with the transition from NN to LN-NN. Correlations between simulated and observed data become significantly higher and decreases significantly slower with increasing distance from the training period (Fig. 12). RMSEs (Fig. 13) and biases (Fig. 14) diminish significantly. The amplitude of the annual cycle decreases very significantly in all aforementioned statistics.

II.4 Global Spatial Maps of Root-Mean-Square Error (RMSE) and Cross-Correlations for NN and LN-NN

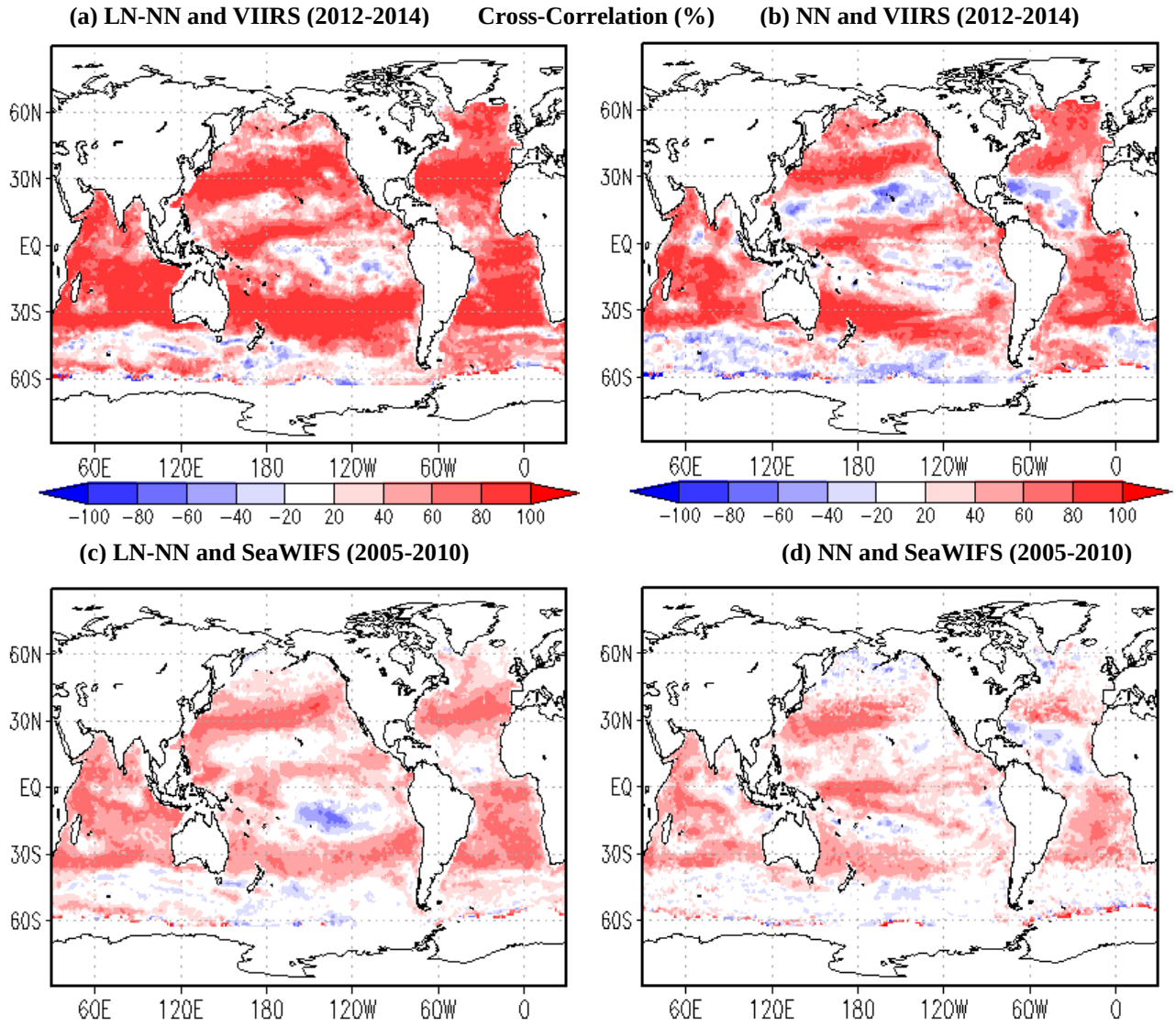

Figure 15. Cross-correlation of monthly mean: (a) LN-NN simulated chl-a concentrations and VIIRS chl-a fields **for 2012-2014; (b) NN simulated chl-a concentrations and VIIRS chl-a fields for 2012-2014; (c) LN-NN simulated** chl-a concentrations and SeaWiFS chl-a fields for 2005-2010; and (d) NN simulated chl-a concentrations and SeaWiFS chl-a fields for 2005-2010.

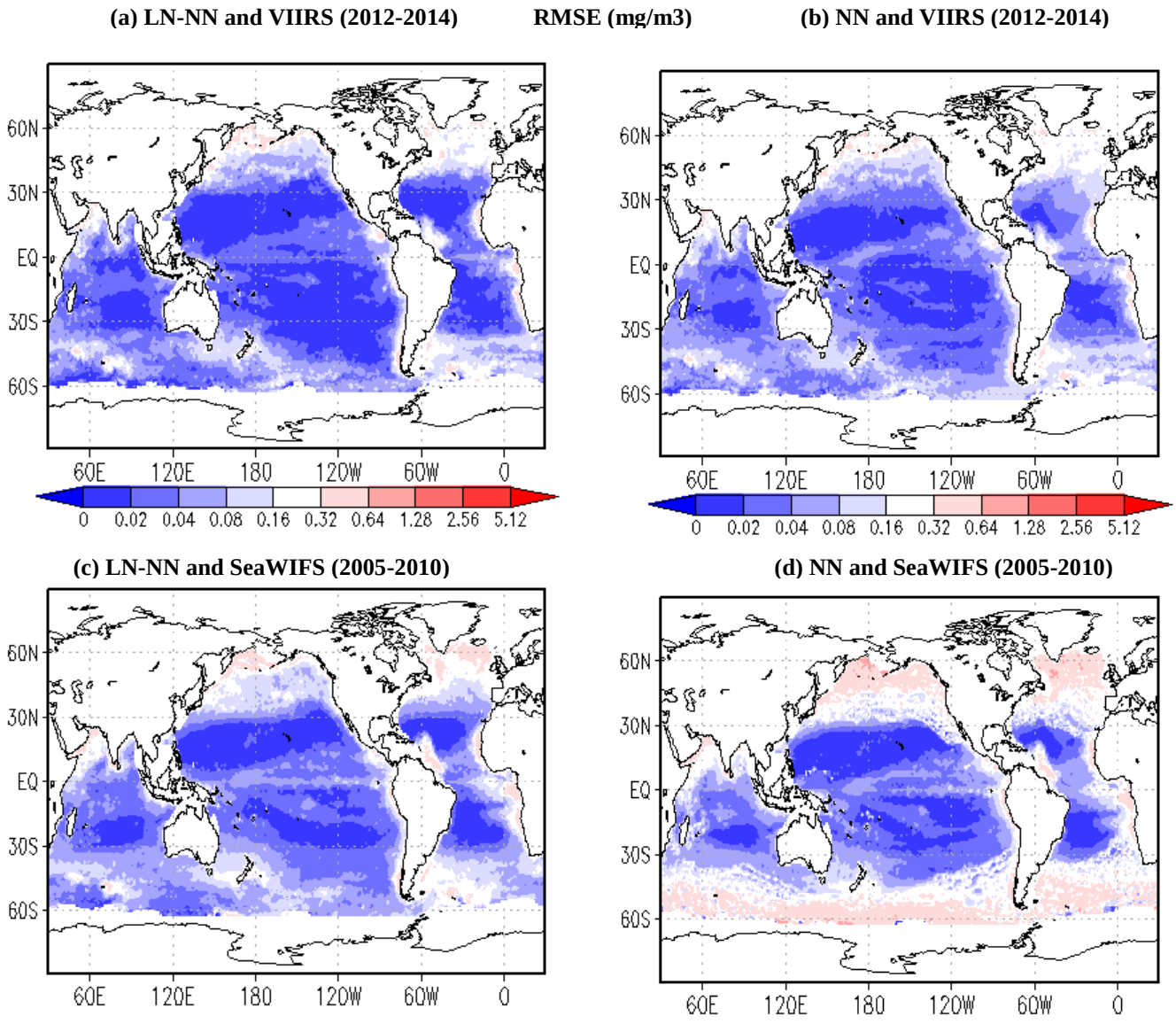

Figure 16. Root-mean-square errors (RMSE) of monthly mean: (a) LN-NN simulated chl-a concentrations compared to VIIRS chl-a fields for 2012-2014; (b) NN simulated chl-a concentrations compared to VIIRS chl-a fields for 2012-2014; (c) LN-NN simulated chl-a concentrations compared to SeaWiFS chl-a fields for 2005-2010; and (d) NN simulated chl-a concentrations compared to SeaWiFS chl-a fields for 2005-2010.

The NN has difficulty in certain regions (Figs 15), possibly because high spatial gradients and temporal variability in satellite chl-a values are not adequately sampled by the inputs (SST, SSS, T, and S). It is clear that the cross-correlation values are much improved for LN-NN with respect to VIIRS (Fig. 15a) and SeaWiFS (Fig. 15c) as compared with NN (Fig. 15b and Fig.

15d) in most regions in the global oceans. Figure 16 indicates that spatial plot of LN-NN with respect to VIIRS observations (Fig. 16a) is quite similar to that with respect to SeaWiFS (Fig. 16c) in the equatorial and tropical oceans, but in the polar oceans, the RMSE is larger with respect to SeaWiFS observations. Also, the RMSE of LN-NN with respect to VIIRS (Fig. 16a) and SeaWIFS (Fig. 16c) is lower in all regions of the global oceans than the corresponding values for NN (Fig. 16b and Fig. 16d, respectively).

III. Discussion and Conclusions

In our previous works (Krasnopolsky et al. 2015 and Nadiga et al. 2016), we introduced a new approach. The approach is based on a NN technique; it relates a biological parameter, the Chl-a concentration, to the physical processes of the upper ocean. We developed a NN, which maps satellite-derived surface variables — sea-surface temperature (SST), sea-surface height (SSH), and sea-surface salinity (SSS) fields — and some *in situ* observations (upper layers of Argo salinity and temperature profiles), to the Chl-a concentration. A concise summary of this study is: **we have developed a NN-based empirical biological model for Chl-a**. However, the NN model developed in the previous study had limited predictive skills.

In this work, we evaluated several methods of optimization of the developed NN-based model to improve the predictive skills. Thus, we (1) develop an empirical biological model for Chl-a capable of long term (several years) prediction of global Chl-a fields and (2) develop a NN capable of simulating a long term (up to 10 years, 2005 to 2014) global data set for Chl-a, which is consistent with the data from all three OC sensors (SeaWiFS, MODIS, and VIIRS). Results

are assessed using the mean error (bias), root-mean-square error (RMSE), and cross-correlations between observed and NN generated Chl-a concentrations.

The coarse spatial and temporal resolution of the data limited types of features that can be observed in NN-generated Chl-a fields. As was shown, global and mesoscale features are represented sufficiently well in the NN-estimated OC fields; however, to generate finer scale features, the NN has to be trained on data with finer resolution.

This study demonstrated that the NN technique provides an accurate, computationally cheap method to generate long (up to 10 year long) time series of consistent Chl-a concentration, which are in good agreement with Chl-a data observed by different OC sensors during this period. It is noteworthy that a single NN (or a single NN ensemble) is capable of generating OC fields all over the globe (at all grid points of the global grid). Also, the accuracy of NN prediction does not deteriorate during the validation period: NN trained on three years of data (2012 and 2013) performs well during the ten years (2005 - 2014) of the validation period. These results demonstrate a very good generalization ability of the NN both in terms of spatial and temporal generalization. It means that the NN-based empirical biological model for Chl-a can be used in the ocean models and coupled climate prediction systems (Saha 2013) to dynamically take into account biological processes in the upper ocean.

The proposed method accurately estimates the seasonal cycle and large-scale spatial patterns in satellite derived Chl-a fields. It best reproduces Chl-a variability in the major ocean gyres in the mid-latitudes. The largest errors are found in areas where the spatial scales of variability are small and the variability is large, e.g., continental shelves, coastal regions, marginal seas, etc. In these areas the OC and other satellite-derived data have the highest level of noise and Chl-a

variability is high. Thus, removing data points with Chl-a > 1 (less than 1% of points) improves NN performance due to reducing the noise in input and output data. Also, even if the data were not noisy in these areas, the amount of data is very small and not sufficient for NN training here.

The NN approach successfully eliminates the systematic component of the noise (bias). To reduce the random component of the noise, an NN ensemble was trained. Our future objective envisioned is a weighted blend of NN estimates and near-real-time (NRT) VIIRS data for ocean model initialization and assimilation for:

- nowcasts and one-to-two week ocean forecasts by NOAA's operational Real-Time Ocean Forecast System (RTOFS), and
- reanalysis, establishing the best ocean initial conditions by NOAA's operational seasonal-interannual Coupled Forecast System Reanalysis (CFSR)/Global Ocean Data Assimilation System (GODAS).

References:

- Anderson, W., Gnanadesikan, A., and Wittenberg, A., 2009, Regional impacts of ocean color on tropical Pacific variability, *Ocean Sci., 5*, 313-327.

- Ballabrera-Poy, J., R. Murtugudde, R.-H Zhang, and A. Busalacchi, 2007, Coupled Ocean-Atmosphere Response to Seasonal Modulation of Ocean Color: Impact on Interannual Climate Simulations in the Tropical Pacific, *J. Clim., 20*, 353-374.

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

- 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/

- Morel, A., and D. Antoine, 1994, Heating rate within the upper ocean in relation to its bio-optical state, *J. Phys. Oceanogr., 24*, 1652-1665.

- Murtugudde, R., Beauchamp, J., McClain, C.R., Lewis, M., and Busalacchi, A., 2002, Effects of penetrative radiation on the upper tropical ocean circulation, *J. Clim., 15*, 470-486.

- Nadiga, S., Krasnopolsky V., Bayler E., Kim H.-C., Mehra, A., and BehringerD., Neural Network Technique for Gap-Filling Satellite Ocean Color Observations, NCEP Office Note No. 483, 2016, doi:10.7289/V5/NCEP-ON-483.

- Saha, S., and co-authors, 2013, The NCEP climate forecast system version 2, *J. Clim.*,doi:, 10.1175/JCLI-D-12-00823.1
- Taylor, K.E. (2001). "Summarizing multiple aspects of model performance in a single diagram". J. Geophys. Res. 106: 7183–7192. doi:10.1029/2000JD900719
- Zhang, R.-H.,, Busalacchi, A., Wang, X., Ballabrera-Poy, J., Murtugudde, R., Hackert, E., Chen, D., 2009, Role of ocean biology-induced feedback in the modulation of El Nino-Southern Oscillation, *Geophys. Res., Lett.,* ., 36, L03608.