Construction of a Geospatial Predictor by Fusion of Global and Local Models

Bo Han[1], Slobodan Vucetic[1], Amy Braverman[2], Zoran Obradovic[1]

[1]Center for Information Science and Technology
Temple University
1805 N. Broad Str., Philadelphia, PA 19122
{hanbo,vucetic,zoran}@ist.temple.edu

[2]Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109-8099
Amy.J.Braverman@jpl.nasa.gov

Abstract – *Geospatial data collected by remote sensing instruments are characterized by substantial variations in attribute values and relationships over space and time, posing great challenges to develop models with maximum predictive power. In this paper, we propose an approach in which global and local models are constructed, and predictions made by properly weighting their outputs. The algorithm is evaluated on aerosol optical thickness prediction using four consecutive MISR data sets collected in 2002 over the continental US. Results show that while the R^2 accuracy of the ANN global and local models are at most 0.25 and 0.4 respectively, the fusion model is significantly more successful, achieving R^2 accuracy above 0.50. In addition, accuracy improvements differ by spatial location, the largest being in the western US, and the smallest being in the east. This could be exploited to further improve the fusion algorithm.*

Keywords: geophysical retrievals, aerosols, regression, heterogeneous data, model fusion.

1 Introduction

Geophysical data collected through remote sensing are increasingly used in Earth sciences applications [10]. The size, quality, complexity and variability of atmospheric data collected by satellite instruments create many challenges in data analysis and modeling. The challenge addressed in this article is related to supervised learning from heterogeneous data characterized by large variations of observed attribute statistics, and by target functions that depend on unobserved attributes. We consider the problem of using remotely sensed attributes to predict geospatial parameters of interest. The specific objective of this study is prediction of atmospheric aerosol information from radiances observed by satellite instruments.

Prior related research has focused on construction of global prediction models from available labeled data. For example, a linear regression model was developed to measure correlation between dust concentrations and mean monthly values of aerosol optical thickness (AOT) on a global scale [11]. While such global approaches are convenient and use data collected over the entire Earth to learn complex models, the global predictors might fail to fully explain properties of specific spatial regions. Such problems were observed in various studies including exploration of the relationship between remote sensing visible bands and surface reflectance. For instance, it was found that surface reflectance, which is a critical parameter in the accurate derivation of optical thickness over land, varied substantially from one type of vegetation to another [12].

For heterogeneous data with varying characteristics over different local regions, an alternative solution is to construct a number of local predictors, each specific to a given spatial area. Such an approach is employed in the operational algorithms of MISR (the Multi-angle Imaging SpectroRadiometer) [8], a satellite instrument collecting aerosol data for NASA since early 2000. MISR is unique because it views Earth from nine different viewing angles and in four spectral bands simultaneously. Most sensors of this type use only one view angle. MISR measures reflected solar radiation in (9 x 4 =) 36 channels, and complex, computationally intensive algorithms, called retrieval algorithms, are used to convert the radiances into aerosol optical thickness (AOT) measurements. AOT is a measure of the attenuation of solar energy as photons travel through a column of atmosphere. It is an important quantity in the study of climate and climate change. MISR uses different algorithms for retrievals over different surface types [2, 9]. Region-specific approaches to retrieval of aerosol properties were also considered in [4]. While development of local models addresses the data heterogeneity problem, the scarcity of locally-specific annotated data could raise issues related to the choice of model complexity and overfitting control in supervised learning.

In this paper, we propose a fusion modeling approach where both global and local models are constructed, and a prediction obtained by weighting their outputs. We also develop a procedure for determining optimal weighting factors that minimize the mean squared prediction error over a specific region.

Our method is evaluated by performing six AOT prediction experiments using MISR data over the continental US during four time periods of 16-day duration in 2002. We obtained MISR radiance and AOT data from NASA's Langley Research Center [5] for the periods from 07/01/2002 – 09/02/2002. In each experiment, data collected in the previous period(s) are used to train global, local, and fusion models. These

0-7803-9286-8/05/$20.00 ©2005 IEEE

models are applied to predict AOT values in the following 16-day period. We used artificial neural networks to construct both global and local models. Results indicate that the fusion model is more accurate than either local or global models alone. In addition, spatial analysis of prediction errors reveals that AOT over the western US are more difficult to predict than those in the east. Nevertheless, in both areas it is evident that prediction accuracy improves with the increase of the training data set size.

2 Methodology

2.1 Problem Definition

Given a spatial data set D with N training examples, each example $d_i \in D$ is represented by a pair (x_i,y_i), where y_i is the target attribute and $x_i = [x_{i1}...x_{iM}]$ is an M-dimensional vector of attributes derived from the observed radiance information. The objective is to construct an accurate predictor $f(x;\beta)$ of target attribute y by optimizing parameters β from a training set D, such that the mean squared error is minimized:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - f(x_i,\beta))^2 . \quad (1)$$

2.2 Global, Local, and Fusion Models

For data D partitioned into L spatial regions, the objective of local modeling is to construct L region specific predictors. Given a training data subset T_j of size N_j representing the local region j, j = 1...L, the corresponding local model is learned by optimizing parameters of function f_j that minimize the mean squared error over the local training data T_j.

Local modeling allows us to exploit specific dependencies existing in spatially constrained regions. However, to avoid local model overfitting due to scarcity of training data in local regions, one is forced to rely on insufficiently expressive models that might overlook important nonlinear local relationships. This issue is addressed by a global modeling approach that learns a single global model f_G from the whole data set D. However, global models can fail to exploit the intricacies of the heterogeneous spatial data.

In view of this trade-off between local and global model benefits, we propose a fusion predictor g_j for each local region j, j = 1...L:

$$g_j(x) = \alpha_j f_G(x,\beta_G) + (1-\alpha_j) f_j(x,\beta_j), \quad (2)$$

where the fusion parameters α_j, j = 1...L, are constrained to lie in the interval [0, 1].

In (2), the global model f_G aims at discovering spatially-independent properties of the whole data set by adjusting parameters β_G. Local model f_j aims to discover locally specific properties of the data by adjusting parameters β_j. We used feedforward artificial neural networks (ANN's) with a single layer of hidden neurons to construct both global and local models. In the fusion predictor g_j, the fusion parameters α_j, j = 1...L, are used to find the optimal trade-off between the global and local models that maximizes the prediction accuracy over a particular local region.

2.3 Optimization of the Fusion Parameters

The fusion parameter α_j in equation (2) is optimized to minimize mean squared prediction error of fusion model g_j defined as

$$\mathrm{MSE}_j = \sum_{i=1}^{N_j}(y_i - (\alpha_j f_G(x_i,\beta_G) + (1-\alpha_j) f_j(x_i,\beta_j)))^2 . \quad (3)$$

It is easy to show that the optimal choice of α_j is

$$\alpha_j = \frac{\sum_{i=1}^{N_j}(y_i - f_j(x_i,\beta_j))\cdot(f_G(x_i,\beta_G) - f_j(x_i,\beta_j))}{\sum_{i=1}^{N_j}(f_G(x_i,\beta_G) - f_j(x_i,\beta_j))^2} \quad (4)$$

If $\alpha_j < 0$, its value is set to zero, and if $\alpha_j > 1$, its value is set to one.

It is worth noting that the value of the fusion parameter is a very informative piece of information about the nature of a local region. If α_j is near 1, the properties of the local area do not differ from the global properties, while if α_j is near 0, the local area is very different from the overall data distribution.

3 Experimental Results

3.1 Data Set

The MISR instrument aboard NASA's Terra satellite consists of nine cameras [1, 3, 6]. Terra is in polar orbit, and so MISR sees 36 (nine angles and four wavelengths) pieces of information for each north-to-south swaths of the Earth on each Terra daytime half-orbit. The central camera points directly downward, four cameras point forward and down along the flight path, and four point afterward and down. The pointing angles are 0°, ±26.1°, ±45.6°, ±60.0°, and ±70.5° relative to nadir. Each camera measures radiances in four spectral bands: blue, green, red and near-infrared. The band spectral shapes are approximately Gaussian and centered at 446, 558, 672, and 866nm, respectively. Thus, each 1.1 km pixel is described by 36 radiance measurements. Terra has a repeat cycle of 16 days, meaning that every 16 days the ground track repeats. In total, there are 233 distinct MISR orbit paths in the cycle, each covering about 360 km wide scanning swath.

In our experiments, we used four consecutive 16-day cycles of MISR Level 1B2 radiance data and MISR Level 2 aerosol data [5] over the continental US. There are 47 paths covering the continental US, but data from only 33 paths were used in this study as only these paths were available for all 4 cycles (see Table 1). The radiance product includes 36 radiance measurements per data point

with 1.1km × 1.1km resolution. The MISR Level 2 aerosol product provides AOT information at 17.6k × 17.6km resolution.

To merge the radiance and AOT data, each radiance attribute is averaged over each 17.6k × 17.6km block represented by a single AOT value. Before averaging, the MISR quality flag is used to identify and remove pixels with non-valid radiance information. Since AOT is not provided for regions covered with clouds, such regions are not used.

Table 1. Data set summary

	Time Period	Number of points (over 33/47)
Cycle 1	2002-7-1 – 2002-7-16	30530/45448
Cycle 2	2002-7-17 – 2002-8-1	32374/49224
Cycle 3	2002-8-2 – 2002-8-17	22395/35903
Cycle 4	2002-8-18 – 2002-9-2	20079/32394

3.2 Experimental Design

Given spatial coordinates, 36 radiance attributes, and AOT values obtained in previous cycles, our task is to predict AOT on the subsequent cycles. We designed the following six groups of experiments to compare prediction accuracy of our fusion approach against those of global and local models alone.

E1. Cycle 1 training data are used to predict optical thickness in cycle 2;
E2. Cycle 1 training data are used to predict optical thickness in cycle 3;
E3. Cycle 1 training data are used to predict optical thickness in cycle 4;
E4. Cycle 1 and 2 training data are used to predict optical thickness in cycle 3;
E5. Cycle 1 and 2 training data are used to predict optical thickness in cycle 4;
E6. Cycle 1, 2 and 3 training data are used to predict optical thickness in cycle 4.

The experimental procedure consists of the following steps:

STEP 1. Merge all data points from training cycles into a global data set T.
STEP 2. Randomly select 10,000 examples from T to construct a global model f_G.
STEP 3. For each path j, j = 1...33, merge the corresponding data points into a local data set T_j. Divide T_j evenly into two subsets T_{j1} and T_{j2}.
STEP 4. By using data in T_{j1}, construct a local model f_j.
STEP 5. Use T_{j2} to validate and compute the fusion parameter α_j.
STEP 6. Use global, local, and fusion models to predict AOT values for each path of the test cycle(s).
STEP 7. Return to STEP 3 by proceeding with the data from the next path.

In the experiments, we used feedforward neural networks trained by the backpropagation algorithm to construct global and local models. Each neural network had an input layer with 36 radiance attributes, a single hidden layer with 10 hidden units, and a single output unit.

3.3 Accuracy Results

The results of comparing the overall accuracy achieved by global, local and fusion model are reported in Table 2. The accuracy is measured by the mean squared error (MSE) and by the coefficient of determination R-squared (R^2) $R^2 = 1 - MSE/Var(AOT)$, where Var(AOT) is the variance of AOT values on test data.

Table 2. Prediction accuracy on the test cycles

E1: Predict cycle 2 by training on cycle 1			
Predictor	Global	Local	Fusion
MSE	0.0189	0.0187	0.0147
R^2	0.1391	0.1491	0.3286

E2: Predict cycle 3 by training on cycle 1			
Predictor	Global	Local	Fusion
MSE	0.0251	0.0180	0.0153
R^2	0.0800	0.3404	0.4550

E3: Predict cycle 4 by training on cycle 1			
Predictor	Global	Local	Fusion
MSE	0.0156	0.0111	0.0100
R^2	0.0240	0.2916	0.3227

E4: Predict cycle 3 by training on cycles 1+2			
Predictor	Global	Local	Fusion
MSE	0.0198	0.0164	0.0139
R^2	0.2910	0.4117	0.5016

E5: Predict cycle 4 by training on cycles 1+2			
Predictor	Global	Local	Fusion
MSE	0.0137	0.0113	0.0089
R^2	0.1373	0.2856	0.4411

E6: Predict cycle 4 by training on cycles 1+2+3			
Predictor	Global	Local	Fusion
MSE	0.0119	0.0095	0.0084
R^2	0.2521	0.3998	0.4686

The fusion models outperformed local and global models in all of the six experiments. In addition, by comparing the accuracies on Cycle 4 in experiments E3, E5, and E6, as well as accuracies on Cycle 3 in experiments E2 and E4, it is evident that increase of the training data set size results in more accurate global, local, and fusion models.

Accuracies of global, local, and fusion models on individual paths for all of the six experiments E1-E6 are shown at Fig. 1. Prediction accuracies of all models are higher in the eastern paths than in the western paths of the study area . This is consistent with previously published observations [7] comparing the AOT retrieved by MISR with AOT retrieved from ground-based instruments: western regions had lower correlation coefficients than eastern ones.

To compare the accuracy of two models M1 and M2, we present the logarithms of their MSE ratios: log (MSE(M1)/MSE(M2)). Fig. 2-7, show this ratio for each orbit path, Scores below 0 indicates that the model M1 is

more accurate than M2, while scores above 0 indicates that the model M2 is more accurate than M1.

Top, middle and bottom panels in Fig. 2-7 correspond to pairwise comparisons of log MSE accuracies for fusion vs. local, fusion vs. global and local vs. global models in experiments E1-E6. It is interesting to note that the performances of global and local models were not stable from experiment to experiment. However, the fusion model outperforms both global and local models over most paths indicating that the approach may be successful regardless of the complexity of local aerosol properties.

4 Conclusions

In this study, we proposed a fusion approach that improves accuracy of geospatial predictors by appropriately weighting global and local models. The approach takes advantage of large global data sets, but also exploits more specific spatial properties at local sites. The approach is evaluated on AOT prediction using four MISR datasets representing four consecutive 16-day cycles during 2002 over the continental US. The results of six predictive experiments provide evidence that the fusion approach achieved higher overall accuracy than either local or global models alone. In addition, our analysis reveals that the fusion model outperforms both global and local models over most paths, and that prediciton of AOT for paths in the western US is more difficult than prediction of AOT in the east. The fusion model outperformed both global model and local models in most paths individually. This suggests that the fusion approach can improve the prediction accuracy at multiple scales by taking optimal advantage of both global and local information.

Acknowledgments

We thank Ralph A. Kahn and John Martonchik, MISR Science Team members at the Jet Propulsion Laboratory, for their generous help with understanding of the domain. We also thank Atmospheric Sciences Data Center at NASA Langley Research Center for their support in collection of MISR data and Qifang Xu at Temple University for data management and preprocessing help.

References

[1] Graham W. Bothwell, Earl G. Hansen, Robert E. Vargo and Kyle C. Miller. The Multi-angle Imaging SpectroRadiometer science data system, its products, tools, and performance. *IEEE Trans. Geosci. Remote Sens.,* 40 (7): 1467-1476, 2002.

[2] Carol J. Bruegge, Nadine L. Chrien, Robert R. Ando, David J. Diner, Mark C. Helmlinger, Stuart H. Pilorz, and Kurtis J. Thome. Early validation of the Multi-angle Imaging SpectroRadiometer (MISR) radiometric scale. *IEEE Trans. Geosci. Remote Sens.,* 40 (7): 1477-1492, 2002.

[3] David J. Diner, Jewel C. Beckert, Terrence H. Reilly, Carol J. Bruegge, James E. Conel, Ralph A. Kahn, John V. Martonchik, Thomas P. Ackerman, Roger Davies, Siegfried A. W. Gerstl, Howard R. Gordon, Jan-Peter Muller, Ranga B. Myneni, Piers J. Sellers, Bernard Pinty, and Michel M. Verstraete,.Multi-angle Imaging SpectroRadiometer (MISR) Instrument Description and Experiment Overview. *IEEE Trans. Geosci. Remote Sens.,* 36(4): 1072-1087, 1998.

[4] S.-C. Tsay, M. D. King, Y. J. Kaufman, J. R. Herman, and J. C. Wei. Aerosol Properties over Bright-Reflecting Source Regions. *IEEE Transaction on geoscience and remote sensing,* vol 42, No. 3, 2004.

[5] http://delenn.gsfc.nasa.gov/~imswww/pub/imswelcome/.

[6] Jet Propulsion Laboratory, California Institute of Technology, Multi-angle Imaging Spectro-Radiometer Data Product Specifications, April 2003.

[7] Yang Liu, Jeremy A. Sarnat, Brent A. Coull, Petros Koutrakis and Daniel J. Jacob1.Validation of Multiangle Imaging Spectroradiometer (MISR) aerosol optical thickness measurements using Aerosol Robotic Network (AERONET) observations over the contiguous United States. *Journal of Geophysical Research,* Vol. 109, D06205, doi: 10.1029/2003JD003981, 2004.

[8] John V. Martonchik, David J. Diner, Kathleen A. Crean, and Michael A. Bull. Regional Aerosol Retrieval Results From MISR. *IEEE Transactions on geoscience and remote sensing,* vol. 40, No. 7, July 2002.

[9] John V. Martonchik, David J. Diner, Bernard Pinty, Michel M. Verstraete, Ranga B. Myneni, Yuri Knyazikhin, and Howard R. Gordon. Techniques for the Retrieval of Aerosol properties Over Land and Ocean Using Multiangle Imaging. *IEEE Trans. Geosci. Remote Sens.,* 36 (4): 1212-1227, 1998.

[10] V. Ramanathan, P. J. Crutzen, J. T. Kiehl and D. Rosenfeld. Atmosphere – 'Aerosols, climate, and the hydrological cycle'. *Science,* 294(5549): 2119-2124, 2002.

[11] A. Smirnov, B. N. Holben, D. Savoie, J. M. Prospero, Y. J. Kaufman, D. Tanre, T. F. Eck and I. Slutsker. Relationship between column aerosol optical thickness and in situ ground based dust concentrations over Barbados. *Geophysical Research Letters,* 27(11): 1643-1645, June, 2000.

[12] Guoyong Wen, Si-Chee Tsay, Robert F. Cahalan and Lazaros Oreopoulos1. Path radiance technique for retrieving aerosol optical thickness over land. *Journal of Geophysical Research*, vol. 104: 31321-31332, 1999.

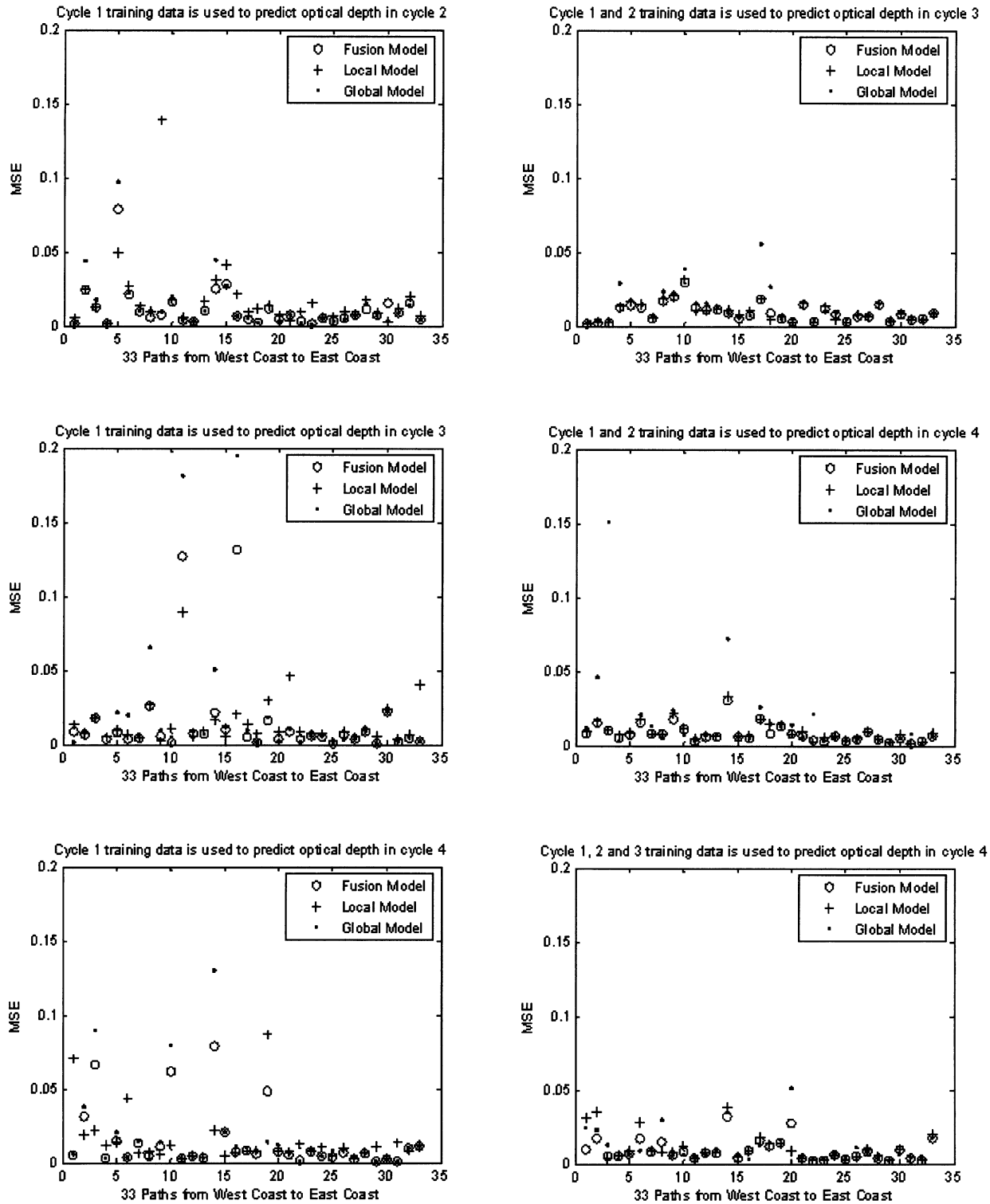
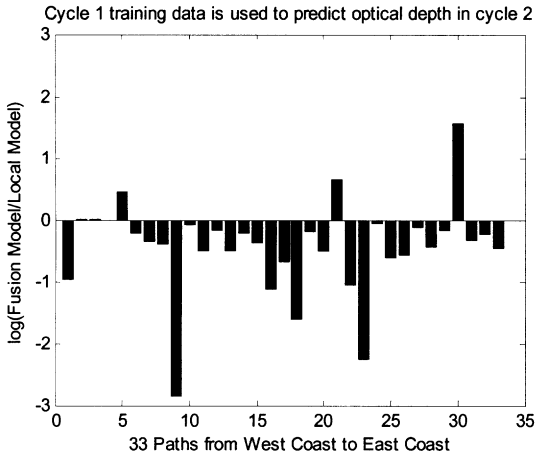
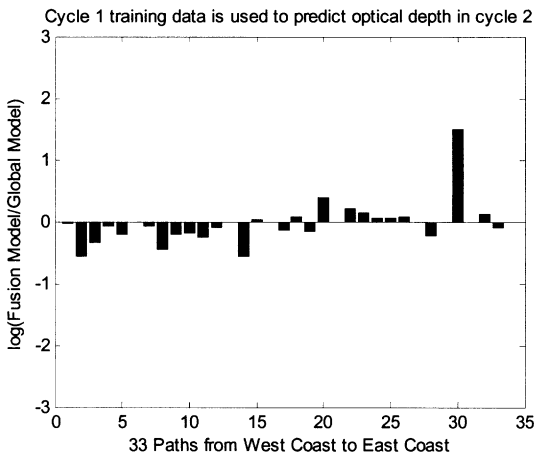


Fig.1. Comparison of MSE of three models on each path for E1 – E6 experiments





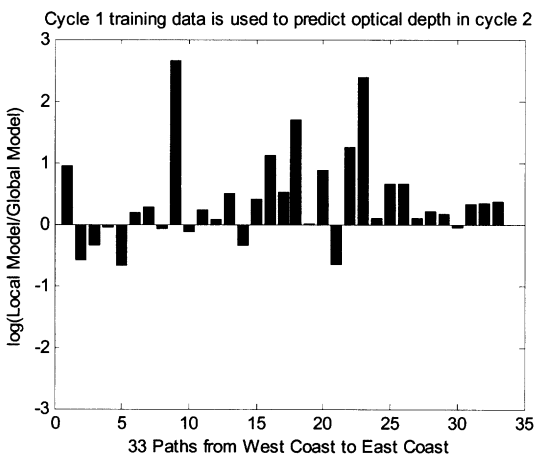
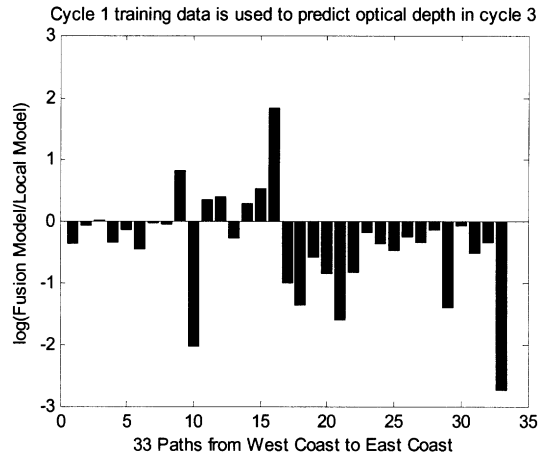
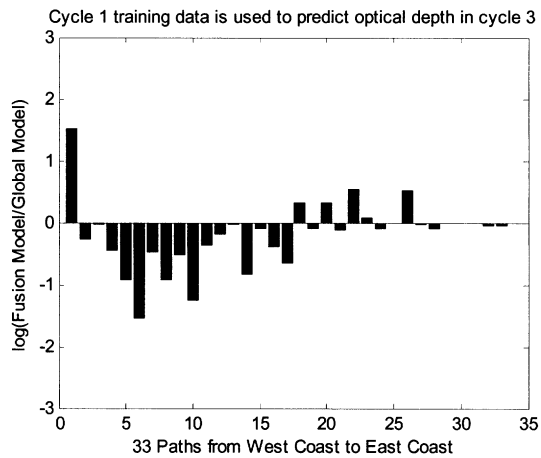


Fig. 2 : Pairwise comparison of log MSE in E1. (a) Fusion model vs. Local Model ; (b) Fusion model vs. Global Model ; (c) Local model vs. Global model





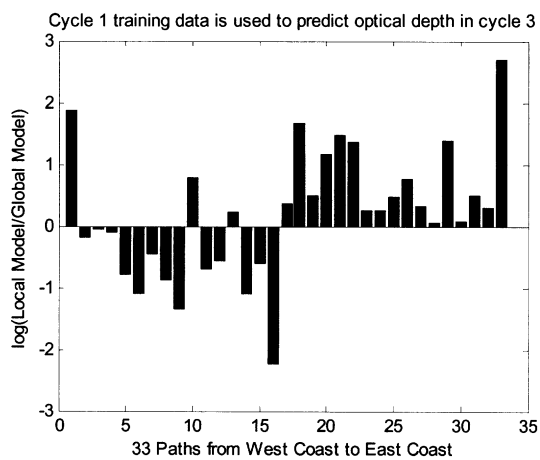
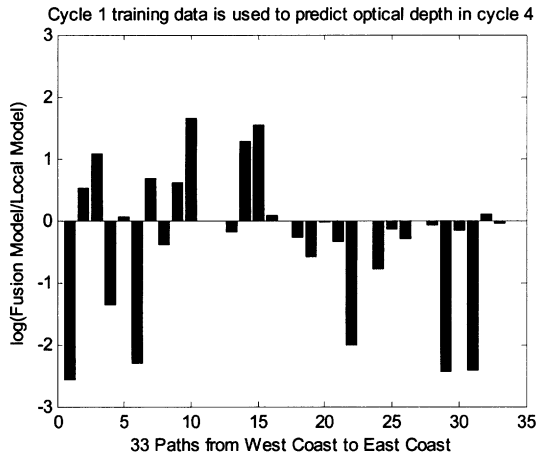
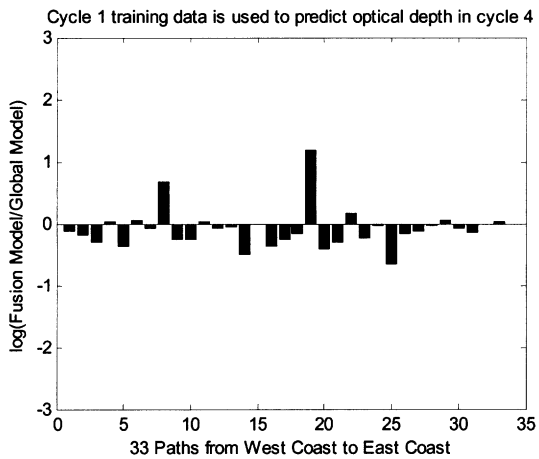


Fig. 3 : Pairwise comparison of log MSE in E2. (a) Fusion model vs. Local Model ; (b) Fusion model vs. Global Model ; (c) Local model vs. Global model





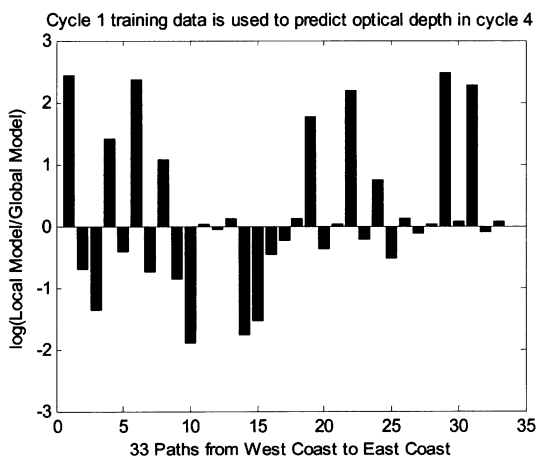
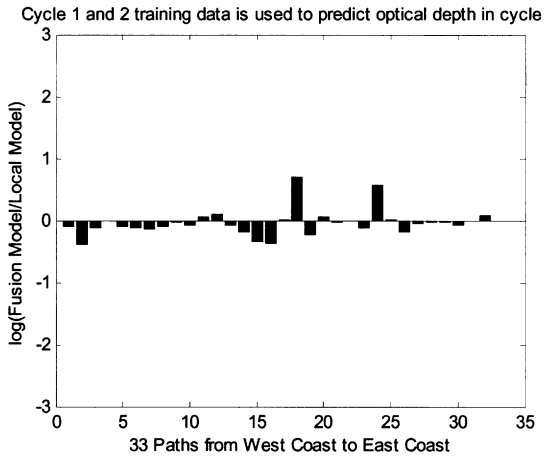
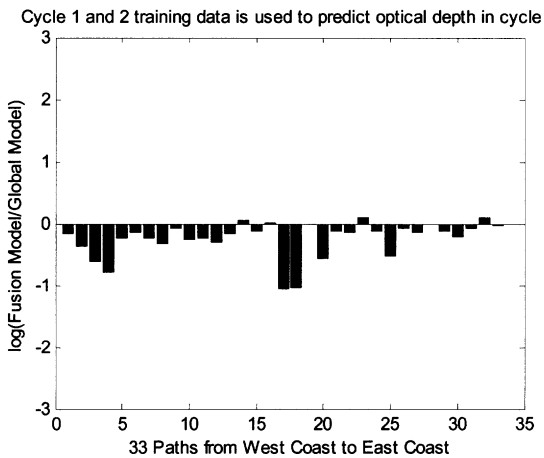


Fig. 4 : Pairwise comparison of log MSE in E3. (a) Fusion model vs. Local Model ; (b) Fusion model vs. Global Model ; (c) Local model vs. Global model





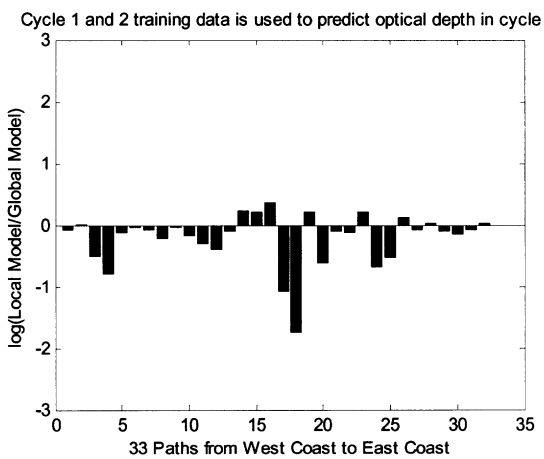
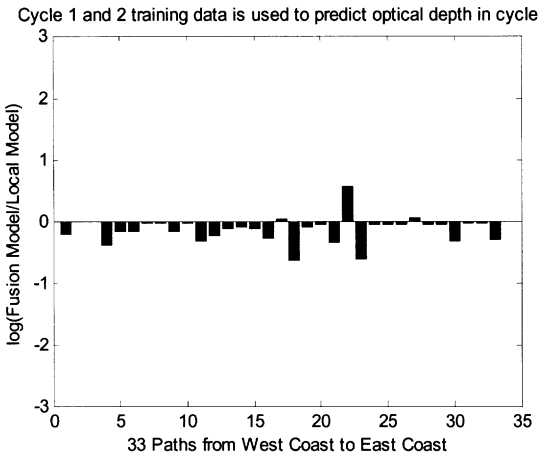
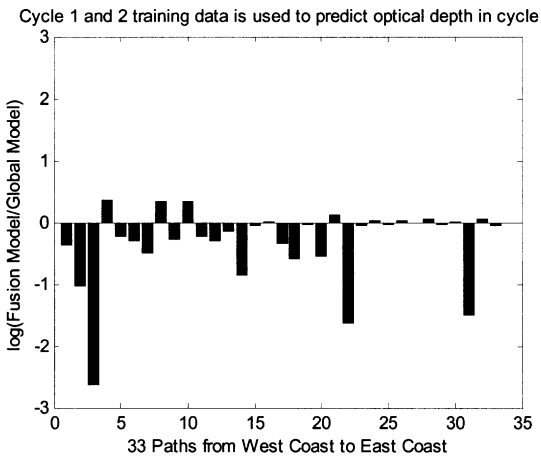


Fig. 5: Pairwise comparison of log MSE in E4. (a) Fusion model vs. Local Model ; (b) Fusion model vs. Global Model ; (c) Local model vs. Global model





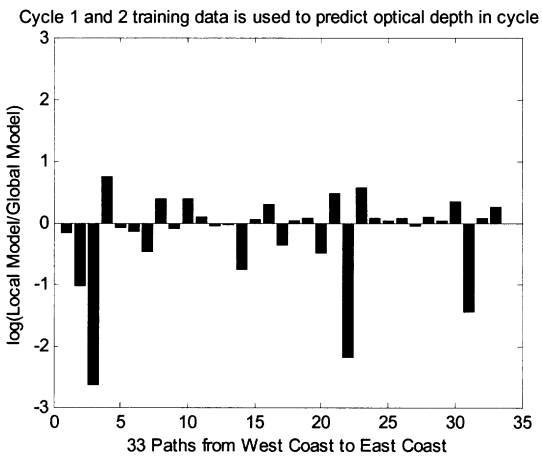
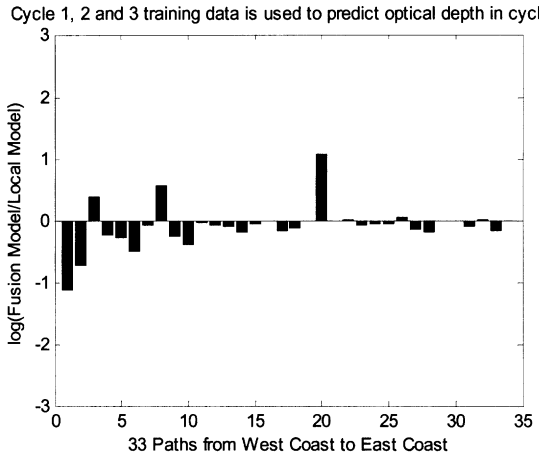
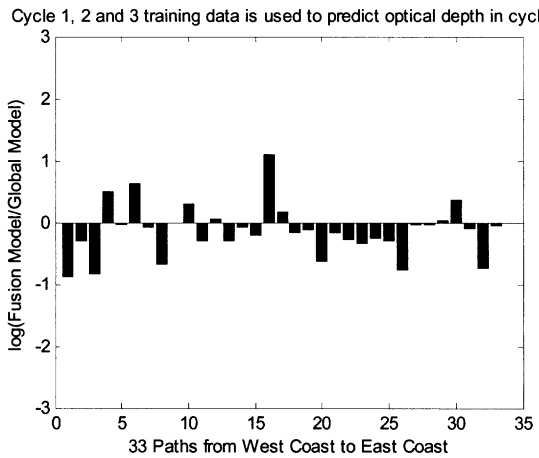


Fig. 6: Pairwise comparison of log MSE in E5. (a) Fusion model vs. Local Model ; (b) Fusion model vs. Global Model ; (c) Local model vs. Global model





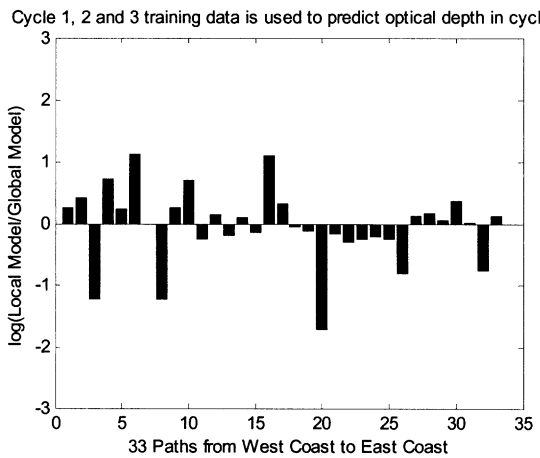


Fig. 7: Pairwise comparison of log MSE in E6. (a) Fusion model vs. Local Model ; (b) Fusion model vs. Global Model ; (c) Local model vs. Global model