Optimal Combining Data for Improving Ocean Modeling

L.I.Piterbarg
University of Southern California
Center of Applied Mathematical Sciences
Kaprielian Hall, R108, 3620 Vermont Avenue
Los Angeles, CA 90089-2532
Phone (213) 740 2459, fax (213) 740 2424, e-mail piter@usc.edu

Award # N00014-11-1-0369
http://www.onr.navy.mil/sci_tech/ocean/onrpgahm.htm

LONG-TERM GOALS

The long range scientific goals of the proposed research comprise: (1) developing rigorous approaches to optimal combining different kinds of observations (images, ADCP, HFR, glider, drifters etc) with output of regional circulation models for accurate estimating the upper ocean velocity field, subsurface thermohaline structure, and mixing characteristics (2) constructing computationally efficient and robust estimation algorithms based on alternative parameterizations of uncertainty and comprehensive testing them on synthetic data (3) processing real data in the Adriatic and Ligurian Sea (MREA coastal experiments) via new techniques

OBJECTIVES

The objectives for the first year of research were:
- Developing and testing methods for fusing HF radar data with tracer (SST, color) and/or drifter observations to improve surface velocity estimates
- Constructing and testing fusion algorithms for combining glider observations with output of high resolution circulation model
- Incorporating subgrid Lagrangian models identified via drifter data into circulation models for improving estimates of FSLE (finite size Lyapunov exponent)
- Developing theoretical approaches based on fuzzy logic to estimating oceanic parameters from small biased samples.

APPROACH

We develop theoretical approaches to the data fusion problem in context of the possibility theory (fuzzy logic) and in the framework of the classical theory of random processes and fields covered by stochastic partial differential equations. We also design computational algorithms derived from the theoretical findings. A significant part of the algorithm validation is their testing via Monte Carlo simulations. Such an approach provides us with an accurate error analysis. Together with my collaborators from Rosenstiel School of Marine and Atmospheric Research (RSMAS), Consiglio Nazionale delle Ricerche (ISMAR, LaSpezia, Italy), University of Toulon (France), and Naval Postgraduate School (Monterrey, CA) we implement the algorithms in concrete ocean models such as QG, POM, HYCOM, NCOM, and MFS, as well as carry out statistical analysis of real data sets by means of new methods.

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **30 SEP 2011**	2. REPORT TYPE	3. DATES COVERED **00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE **Optimal Combining Data for Improving Ocean Modeling**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Southern California,Center of Applied Mathematical Sciences,Kaprielian Hall, R108, 3620 Vermont Avenue,Los Angeles,CA,90080-2532**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**7**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

WORK COMPLETED

1. Developing and testing methods for fusing HF radar data with images and/or drifter observations to improve surface velocity estimates.

In now days velocity measurements by HF radars became an important tool in investigating spatial structure of surface currents in coastal regions. Still they are not perfect. Different experiments with direct velocity measurements have shown significant errors: around 15 cm/s (along the North Carolina Coast, 1997 [1]), 7-19 cm/s (along the California Coast, 2004, [1]), 6.6-11.3 cm/s (Korea/Tsushima Strait, 2006, [2]), and 6-13 cm/s (East Coast of Korea, 2007, [3]). In the last paper it was noticed that the accuracy of estimating zonal and meridional components are essentially different. Moreover, appropriate velocity estimates in a certain area can be obtained only if the area is covered at least by two radars. If only one of them works properly, then an important problem arises how to use other available information to restore the surface velocities.

We have developed two methods for combining single radar measurements with other data, first, with satellite images (color or SST) and, second, with drifter observations. The first method is based on solving multi-objective optimization problem [4] with modeling uncertainties via fuzzy sets, [5-7]. The method for incorporating drifter observations uses a simplified version of the same approach which is reduced to estimating the radial velocity component from the radar and the orthogonal component from the nearest drifter.

To test both methods we first used synthetic idealized velocity fields in the framework of twin experiments. For the second method (velocities from radar and drifters) massive Monte Carlo experiments were carried out using a synthetic velocity field obtained from a realistic circulation model.

2. Constructing and testing fusion algorithms for combining glider observations with output of high resolution circulation model.

The problem of combining glider data with model output exposes new challenges in the data fusion business such as a different resolution (for glider measurements it is essentially higher than for typical models), specific features of glider paths and uncertainty in its position, and aliasing in the case of fast variability. On the preliminary stage we ignored the temporal variability and focused on accounting for uncertainty in the glider position to develop a method for combining its measurements with model output.

The developed method combines a traditional approach of box models and a novel procedure of optimal combining data from two small biased samples suggested by PI, [8]. The method was tested on a synthetic velocity field modeling both meso and submeso scales with realistic glider trajectories.

3. Incorporating subgrid Lagrangian models and statistical information from drifter data into circulation models for improving estimates of FSLE

Three Lagrangian subgridscale (LSGS) models that had been developed on the basis of statistical considerations by PI with collaborators [9,10] were applied to tackle the multi-scale ocean transport problem. We used a hybrid approach, in which the modeled transport is based on the deterministic Lagrangian coherent structures over the mesoscale range obtained as a HYCOM output and the statistical LSGS over the submesoscale range. That approach was applied for investigating transport in the Gulf Stream region using FSLE. In now days FSLE became a popular and efficient tool for detecting barriers to transport [11], identifying hyperbolic manifolds [12], measuring local stirring [13], and identifying Lagarangian coherent structures [14]. Parameters of LSGS were estimated from drifter observations in this area. A comprehensive theoretical investigation of FSLE for simple models of turbulence was carried out to calibrate algorithms of estimating FSLE and interpret simulation results.

4. Developing theoretical approaches based on fuzzy logic to estimating oceanic parameters from small biased samples.

Sparse observations, biasness, and small samples pose serious obstacles for application of classical statistical methods in processing ocean data. An alternative approach has been developed by PI [8] for fusing such data and making inferences based on the fundamental concept of Pareto optimal solutions in multi objective optimization problems [4].
During the reported period we continued theoretical studies to enhance the developed method for estimating location parameter from several small biased samples and to compare it with other known procedures. Some new optimal properties of the method have been proven and tested by Monte Carlo simulations

RESULTS

1. The developed method of combining radar measurements and tracer observations for correcting surface velocities proved to be accurate and extremely computationally efficient in conditions of poor known sources and dissipation for the tracer. One of the experiments is illustrated in Fig.1 where a failure of the radar was simulated covering the eastern part of the region of interest.

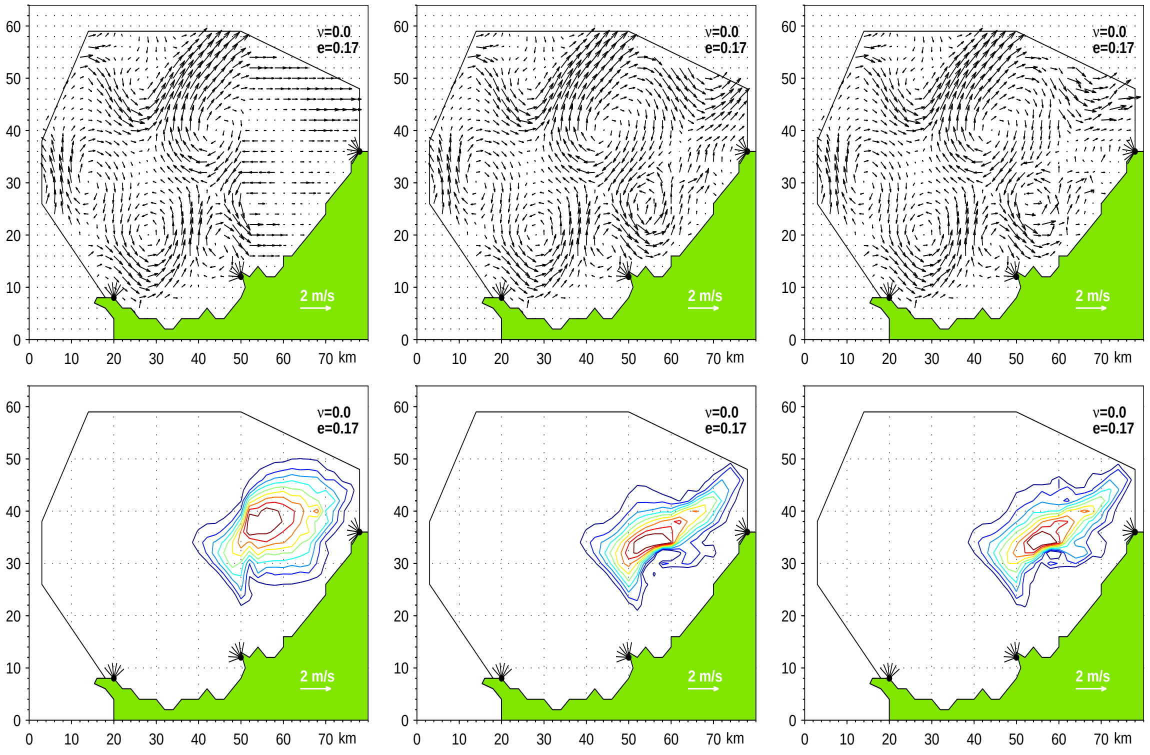
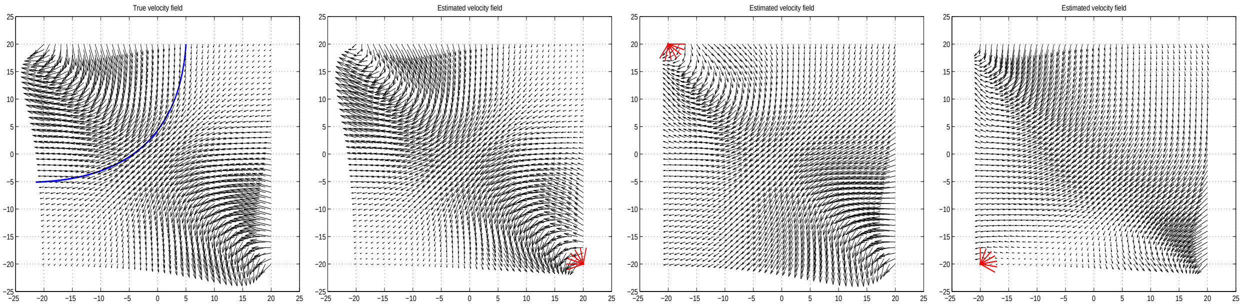


Figure 1. Modeling HF radar failure: Upper panel: 1) Radar measurements. 2) 'True' circulation 3) Estimated circulation velocity. Bottom panel 1) Tracer driven by distorted velocity field. 2) 'True' tracer distribution. 3) Estimated tracer distribution

After including synthetic sea surface temperature data the estimate was able to capture main circulation patterns of the true velocity field. In other experiments we varied different parameters such as the location of the tracer patch, its intensity, and the forcing in the transport equation. The improvement comparing to bare radar measurements varied between 63 and 14 % depending on values of the indicated parameters.

The developed procedure of optimal combining HF radar measurements of surface velocities and drifter observations turned out to be accurate enough even in the case of a single drifter. The relative estimation error δ is defined as ratio of MSE averaged over the whole region to the variance of the true velocity field also obtained by the space averaging. One of the experiments is illustrated in Fig.2 where the drifter trajectory is fixed (blue) and the radar position varies (red). From this experiment and other ones not shown here it can be seen that a higher estimation accuracy is reached when a trajectory is near orthogonal to the radar direction while in the opposite case (the trajectory is approximately parallel to the radar direction) the error becomes significant thereby depreciating drifter observations. Preliminary experiments with processing real data involving several drifters showed that the estimation procedure is efficient enough in vicinity of the drifter trajectories, but the estimate error grows fast as the distance from the drifter trajectories is increasing.



$True$ $\delta = 0.2081$ $\delta = 0.3633$ $\delta = 0.5344$

Figure 2. Dependence of the error of estimating surface velocities on the position of radar. 'True' velocity field and drifter path are shown on the leftest panel.

2. We first investigated the error s_o of interpolation of glider measurements depending on the reduced diving frequency $k = NH/L$ where L is the length of pass and H is the average depth of dives. The main finding based on experiments with synthetic temperature (density) fields is that a reasonable accuracy ($\delta < 0.4$) is reached if $k > 0.7$ and it becomes high ($\delta < 0.1$) for $k > 4$. These numbers obtained for a certain smooth enough synthetic temperature field with horizontal and vertical scales of variability comparable to L and H respectively.

On the second stage we compared the developed procedure of combining glider measurements with a hypothetical model output. In the first series of experiments the model output was assumed to be a smoothed version of the true field. In all the experiments the error of the estimate s_e turned out to be smaller than both, interpolation error s_o and the model output error s_b. One of the examples illustrating that is shown in Fig. 3 where $s_o = 0.3947, \quad s_b = 0.3680, \quad s_e = 0.2589$

Finally, we compared the error of the new method with that of traditional weighted mean (WM) of observations and model output in a different set up. Namely, both observations and model output were obtained by adding independent noises to a background with intensities σ_o and σ_m. A part of the results is presented in the table for noises with heavy tails (Cauchy distribution). As one can see the new procedure based on fuzzy sets approach significantly overperformed WM. However, for Gaussian noises the fuzzy algorithm performs only slightly better while in some cases it's even a little bit worse.

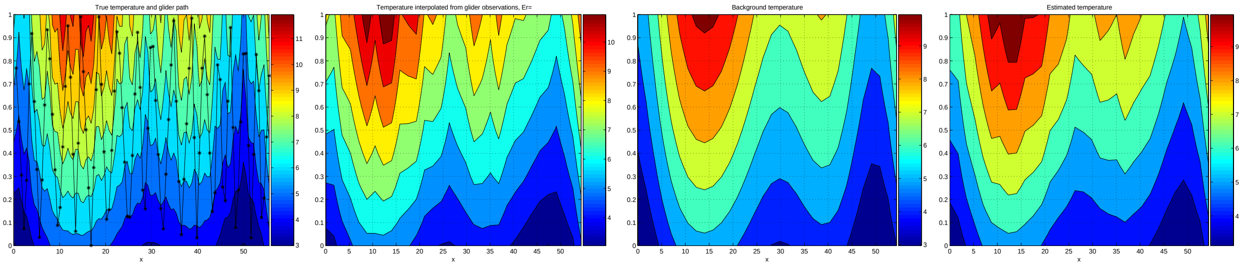


Figure 3. Fusion glider and model temperature. 1) 'True' temperature field and glider path. 2) temperature obtained by interpolation of glider measurements. 3) Background (model) temperature. 4) Estimate.

σ_m, σ_o	2, 2	2, 1	1, 2
Background error	2.8254	2.1974	0.7191
Observation error	0.8062	0.6250	0.5614
Fuzzy method error	0.1081	0.1199	0.0909
WM error	0.4379	0.4486	0.2497

3. On the theoretical side, exact expressions for FSLE $\lambda(\delta)$ were found and analyzed for several idealized models of turbulence in 1D and 2D. Among them are a random walk with discrete time and continuously distributed jumps and an isotropic Brownian flow in 2D also known as the Kraichnan flow. For the former a surprising fact is a δ^{-1} scaling for small δ in contrast to δ^{-2} well known for a random walk in continuous time (Brownian flow), for a simple random walk in discrete time, and 3D turbulence in high Reynolds number [15]. For the Kraichnan flow an exact relation is established between the scaling of $\lambda(\delta)$ and the scaling of relative dispersion in time.

The main finding in experiments with three LSGS models is that all of them turned out to be able to enhance the unrealistically low strain level of OGCMs , while it is difficult to provide indications on which one of the three parameterizations is more realistic given the sparsity of in situ data presently available for validation. It is entirely possible that the choice of the appropriate LSGS model will depend on the specific application considered, in terms of OGCM and of dynamical region. For the considered Gulf Stream region the results of [16] indicate a δ^{-1} slope in the submesoscale range which is in agreement with our theoretical and simulation results where purely stochastic LSGS models were used. This in turn suggests that drifter motion falls into the same category of random walk increments with scales of the order of 1-2 km. On the other hand, one should keep in mind that drifter launches in [11] occurred in extreme conditions during the winter season, with very high winds and waves, and were clustered in relatively small areas in the Gulf Stream and in its anticyclonic recirculation. It is therefore unclear whether or not they can be considered typical of the area and even more if they can be directly compared to model results averaged over large space and time scales.

Another important conclusion is that adapting the previously developed LSGS models to the parametrization of submesoscale dispersion by shifting the separation of scales to the radius of deformation, and evaluating their performance under a scale dependent relative dispersion metric has resulted in a substantial improvement of submesoscale Lagrangian transport.

4. We proved that using other possibility distributions (membership functions) than triangle and trapezoid ones as it was in [8] can essentially improve the estimate of a location parameter from two small biased samples. It was shown theoretically and Monte Carlo experiments that the newest version of the method outperforms not only the classical weighted mean (WM) with weights inversely proportional to variances, but also WM with weights from a wide range. A paradoxical

result was obtained implying that in terms of optimizing the bias only, the WM with equal weights is best. However, it is not the case when the mean square error is taken as an efficiency measure. Under such a condition the fuzzy estimate is still preferable.

IMPACT/APPLICATIONS

The developed methods for combining radar, tracer, and drifter observations as well as fusing glider data and model output are highly portable and computationally efficient, making them very valuable in the framework of operational strategies for rapid assessment and quick response. They have therefore some significant advantages with respect to other techniques requiring complete assimilation of the tracer information in the dynamical velocity models. These techniques, even though more powerful, require significant coding and computational time and they have to be set up in advance for the specific operational model in use. Moreover, the developed methods are resistant to uncertainties in tracer generation/dissipation , are capable to aggregate data at different resolutions, and account for sample biasness. Thus, we expect that our results will stimulate more efforts in developing fusion methods which carry no risk of ruining a model during the running time and , in addition, are well theoretically founded.

Our findings in combining OGCMs with LSGS models based on real data stimulate search for ways to accurately model submesoscale processes. In turn, appropriate modeling Lagrangian motion on submesoscales contributes to a variety of practical environmental and operational issues such as monitoring and forecasting pollutant spreading, search and rescue operations in the sea, and prediction of fish larvae.

RELATED PROJECTS

1. ”Lagrangian turbulence and transport in semi-enclosed basins and coastal regions”, ONR, PI A Griffa, RSMAS, N00014-05-1-0094
2. ”Ocean 3D+”, MURI Project, ONR N00014-11-1-0087, PIs: A. Griffa, T. Ozgokmen, I. Mezic, C. Jones , I. Rypina, S. L. Smith, L. Pratt, D. Kirwan

REFERENCES

1. Chapman, R.D. and Hans C. Graber, (1997), Validation of HF radar measurements, Oceanography, vol. 10, n.2
2. Yoshikawa, Y., A. Masuda, K. Marubayashi, M. Ishibashi, and A. Okuno (2006), On the accuracy of HF radar measurement in the Tsushima Strait, Journal of Geophysical Reseach, v. 111, C04009, 10 pp. doi:10.1029/2005JC003232
3. Na, H., K.Kim, K.Chang, (2007), Accuracy of surface current velocity measurements obtained from HF radar along the east cost of Korea, http://www.pices.int/publications/presentations/
4. Deb, K., (2001), Multi-Objective Optimization Using Evolutionary Algorithms, Willey.
5. D.Dubious and H. Prade, (1986), Possibility theory, Plenum Press, New York and London,
6. D. Dubois, H.Prade, and R.R. Yager, (1997) Fuzzy Information Engineering, John Wiley & Sons, Inc.
7. G. Shafer,(1976), A mathematical theory of evidence, Princeton University Press.
8. L.I.Piterbarg, (2011), Parameter estimation from small biased samples: statistics vs fuzzy logic, *Fuzzy Sets and Systems*, 170, 1-21

9. A.C. Haza, L.I. Piterbarg, P. Martin,T.M. Özgökmen, and A. Griffa, (2007), A Lagrangian subgridscale model for particle transport improvement and application in the Adriatic Sea using the Navy Coastal Ocean Model, Ocean Modelling, v.17, 68-91
10. L.I. Piterbarg, (2006), Parameter estimation in multi particle Lagrangian stochastic models, Monte Carlo Methods and Applications, v.12, n.5-6, 477-494
11. G. Boffetta, G. Lacorata, G. Redaelli, A. Vulpiani, (2001), Detecting barriers to transport: A review of different techniques, Physica D, 159, 58-70
12. B. Joseph and B. Legras (2002), Relation between Kinematic Boundaries, Stirring, and Barriers for the Antarctic Polar Vortex, J. Atm. Sci., v.59, pp.1198-1212
13. F. dOvidio, V. Fernandez, E. Hernandez-Garca, and C. Lopez, (2004) Mixing structures in the Mediterranean Sea from Finite-Size Lyapunov Exponents, Geoph. Res. Letters, v. 31, L17203, doi:10.1029/2004GL020328
14. E. T. Kaia, V. Rossib, J. Sudreb, H. Weimerskirchc, C. Lopezd, E. Hernandez-Garciad, F. Marsaca and V. Garonb, (2009),Top marine predators track Lagrangian coherent structures, www.pnas.org/cgi/doi/10.1073/pnas.0811034106
15. E Aurell, G Boffetta, A Crisanti, G Paladin and A Vulpiani, Predictability in the large: an extension of the concept of Lyapunov exponent, (1997), J. Phys. A: Math. Gen. 30, 1
16. Lumpkin, R., and S. Elipot (2010), Surface drifter pair spreading in the North Atlantic, J. Geophys. Res., 115, C12017, doi:10.1029/2010JC006338

PUBLICATIONS

1. L.I.Piterbarg, (2011), Parameter estimation from small biased samples: statistics vs fuzzy logic, Fuzzy Sets and Systems, 170, 1-21
2. M. Caglar, T.Bilal, L.I. Piterbarg, (2011), Lagrangian prediction and correlation analysis with Eulerian data, Turkish J. Earth Sciences, 20, 343-358
3. L.I. Piterbarg and L. Ivanov, (2011), Fuzzy-logic based algorithm for estimating circulation patterns, Current Applied Mathematics, accepted
4. L.I. Piterbarg, (2011), Finite size Lyapunov exponent for some simple models of turbulence, Applied Mathematical Modelling, under revision
5. A. Haza, T. Ozgokmen, A. Griffa, Z. Garaffo, L. Piterbarg, (2011), Parameterization of Submesoscale Transport in the Gulf Stream Region Using Lagrangian Subgridscale Models, Ocean Modelling, under revision