2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Combining networks of drifting profiling floats and gliders for adaptive sampling of the Ocean

Alberto Alvarez, Bartolome Garau and Andrea Caiti

***Abstract*— Monitoring ocean dynamics is extremely difficult due to its enormous physical dimensions and the wide range of spatio-temporal scales involved in its dynamical behaviour. It has been recently proposed that the most efficient way to monitor the ocean is through networks of small, intelligent and cheap robotic platforms. Drifting profiling floats and gliders were developed in this context. Floats move with the currents meanwhile they periodically sample the water column through controlled immersions. Conversely, gliders are underwater autonomous vehicles with controllable motion at sea. Both platforms are extensively employed in oceanography due to their high autonomy. A network called Argo of around 3000 profiling floats spreads out around the world's ocean. Glider networks are starting to settle down at smaller scale in different places.**

The advent of these networks and the still scarce resources for ocean sampling, create a demand for quantitative tools for optimizing their use. In this work, the problem of optimally merging networks of profiling floats and gliders is considered. Specifically, a genetic algorithm is employed to find optima gliders trajectories to get together an unevenly distributed network of floats the best quality of the sampled field. A measure of the quality of the oceanographic field (objective function to minimize) is defined in terms of the mean formal error obtained from an optimum interpolation scheme. Results show that the quality of the sampled field can be greatly improved by merging both networks if the resolution of glider observations is adequately selected. The spatial lag between glider observations is related to the geometry of the network of profiling floats and must be order of the grid spacing obtained from the mean data spacing of the network of floats.

I. INTRODUCTION

Marine environment is an extremely complex system, characterized by strong links between its physico-chemical processes and its biological population. The relevance of the interactions between the physical, chemical and biological fields and its high spatio-temporal variability difficult the study of the ocean: first, because they imply to simultaneously measuring the physical, chemical and biological parameters and second because these measurements must be done with an adequate spatio-temporal resolution.

Spatio-temporal resolution of ocean observations depends on the characteristics of the observing platform employed. An observing platform must be capable of mapping an ocean structure at adequate spatial resolution and faster than significant changes in the structure occur. Unfortunately, this requirement does not hold for many traditional platforms of oceanographic sampling due to physical, economic and/or operational limits of the sampling platforms. Present oceanographic surveys are fundamentally limited by too few measurements, taken too slow, at too great cost. For this reason, a new generation of robotic ocean observing platforms has been developed to sample the ocean at high spatio-temporal resolutions. Underlying this technological development is the idea that the most efficient and economic way to sample the ocean is through networks of small, intelligent and cheap ocean observing platforms [9].

Drifting profiling floats and gliders are the two newest platforms mostly employed in oceanography. This is due to their long autonomy at sea, up to a year and several months respectively. Drifting profiling floats are designed to cycle between the surface and some predetermined depth. The float spends roughly 10 to 14 days drifting at depth and returning periodically to surface to report using the Argos system, its position and information about temperature and salinity of the water column [3]. Emersion is obtained by the drifter moving oil from an internal reservoir to an external bladder, reducing the drifter's density. Conversely, a latching valve is opened to allow oil to flow back into the internal reservoir for immersion. Speed of the float is usually of several centimetres per second, depending on the current field at the drifting depth. The spatial period of a cycle is order of tens to hundred kilometres. Presently, a global array of 3000 free-drifting profiling floats at spacing of about 300 km by 300 km, allows continuous monitoring of the temperature, salinity, and velocity of the world's upper ocean.

Gliders are underwater autonomous vehicles designed to observe vast areas of the interior ocean [12]. Gliders make use of their hydrodynamic shape and small fins to induce horizontal motions while controlling their buoyancy. Buoyancy control also allows vertical motions in the water column. In summary, buoyancy changes and hydrodynamic shape allow gliders to carry out zig-zag motions between the surface and bottom of the ocean with a net horizontal displacement. Nominal horizontal speed is about 2 Km/h with a spatial cycle period of 800 m for immersion depths up to 1000 m. Unlike the case of profiling floats, networks of gliders are still incipient but first attempts to settle them down have already been done.

The advent of networks of drifting profiling floats and gliders creates a demand for quantitative tools for optimizing their use. Specifically, ways of allocating these observational resources so as to maximize the information content of the collected data are required. Previous studies have considered

A. Alvarez and B. Garau are with the Natural Resources Department, Instituto Mediterraneo de Estudios Avanzados-IMEDEA, Miquel Marques 21, 07190 Esporlas, Spain vieaaad@uib.es tomeu.garau@uib.es

A. Caiti is with the Department of Electrical Systems and Automation, Universit of Pisa, Via Diotisalvi, 2 56100 Pisa, Italy caiti@dsea.unipi.it

optimizing cast strategies for isolated networks of drifters [6] and gliders [10], [11], but the synergy of combining both observational networks has not been explored yet. Finding an optimum cast strategy to simultaneously sample the ocean with a network of profiling floats and gliders is of outstanding relevance in oceanography, due to the still scarce observational resources. It also constitutes an interesting robotic problem concerning fusion of data gathered by networks of robotic platforms, spanning different ranges of space and time variability in a spatially distributed system. Networks of profiling floats are able to capture relatively large and slow scales of ocean variability (one profile of the water column every ten to fourteen days with a spatial period of tens to hundreds kilometers) while small and fast scales are better represented by networks of gliders (continous sampling of the water column with spatial periods of few kilometers). Compatibility between the range of space and time scales spanned by the different networks is then of key relevance in the optimization problem.

This study investigates the synergy of combining sampling networks of gliders and profiling floats. To do that, a procedure to find those glider trajectories that optimize the overall sampling of the glider and profiling float networks is proposed. Specifically, a genetic algorithm (GA) is employed to find the best sampling strategy of a network of gliders in a region where profiling floats are present. The GA methods are a global optimization approach based on the idea of natural selection. They were first proposed in [7] and detailed described in [5]. GAs have already been applied successfully to the path-planning problem of underwater mobile robots [1], [13]. This paper is organized as follows. Section II defines the problem to be solved. Results are shown in Section III. Finally, Section IV gives concluding remarks.

II. Combining networks of gliders and drifting profiling floats

A set of M drifting profiling floats are considered in an ocean region. The floats are unevenly distributed over the domain of interest, as a result of the spatial variability of ocean currents that drift each float with different speed. The same cycling period is assumed for all floats. Thus, the problem of asynchronous data is not considered here. The observed ocean field is supposed stationary during the cycling period, i.e. floats provide a set of synoptic observations. Besides the profiling floats, a network of N_g gliders is operated in the region. The problem addressed here concerns planning gliders trajectories to get, together with data coming from the profiling floats, the best representation of the sampled field. Oceanographically, the quality of the sampled field is given in terms of the dynamical information that can be extracted from the measured field. This is achieved assigning, from the data gathered at arbitrarily locations, the best values at grid points of a regular grid. This is done employing an optimal interpolation scheme [2], [8] where an estimation of the observed field at position (x,y) is obtained by a linear combination of the observations:

$$F(x,y) = A^{-1}C_x d \tag{1}$$

A is the covariance matrix of observations, C_x is the covariance vector of observations with respect to the estimated field and d is the vector of observations. In practice, the covariance function is assumed to be a Gaussian function of the distance between points, $exp(-\Delta r^2/\kappa)$, with a decorrelation scale, $\sqrt{\kappa}$, depending on the dynamical aspects of the oceanographic processes present in the region. The quadratic error of the estimation (called formal error) is:

$$e^2(x,y) = < F^2 > (1 - C_x^T A^{-1} C_x) \tag{2}$$

where $< F^2 >$ is the variance of the field. The formal error depends on the number and location of observations in the spatial domain, through the covariance matrix and vectors. A reasonable election of objective function to optimize the quality of the sampled field would be to consider the spatial average of the quadratic error field (2), expressed as a percentage of the field variance:

$$J = \frac{1}{N}\sum_{i=1}^{N}\frac{e^2(x_i,y_i)}{< F^2 >} = \frac{1}{N}\sum_{i=1}^{N}(1 - C_x^T A^{-1} C_x) \tag{3}$$

With N the number of grid points of the regular grid where the field is estimated. The number N (i.e. the regular grid spacing) is fixed by the non-uniform spatial distribution of profiling floats in the network. Specifically, the so called random data spacing, Δn_r, can be used as a simple guide for determining the node spacing in the regular grid. Δn_r is the distance defined as the mean spacing derived by distributing the original number of stations uniformly over the data domain. It is given by [8]:

$$\Delta n_r = L_D \frac{1 + M^{\frac{1}{2}}}{M - 1} \tag{4}$$

Where L_D is the dimension of the domain. Typical bounds on the ratio between the grid spacing Δx and the data spacing Δn_r appear to lie in the range of 0.3-0.5 [2], [4], [8]. A ratio of 0.4 is considered here.

A GA has been considered to optimize the objective function (3). In our case, the design variables are the locations of gliders observations that define their trajectories in the domain. For each glider, the total number of observation points, N_v, is limited by the total distance the glider can cover in one cycling period of the profiling floats divided by the constant spatial lag between observations ($N_v = \frac{T_a V_g}{l}$, with T_a cycling period of floats, V_g glider speed and l spatial lag). Appropriate election of this spatial lag constitutes a goal of the present study and will be further discussed in the next sections. Considering straight line paths between two consecutive observation points, trajectories are defined by an initial point and a vector of N_v angles $[\theta_1, \theta_2, ..., \theta_{N_v}]$. With this notation, the spatial coordinates of the i-observation point are given by:

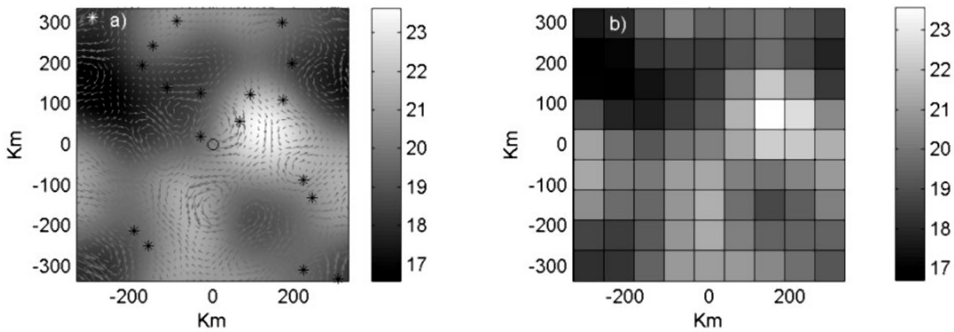

Fig. 1. a) Temperature field (^{o}C), drifting current field, floats positions (stars) and initial location of glider deployments (circle), b) temperature field (^{o}C) interpolated to the grid defined from the configuration of the network of profiling floats.

$$x_i = x_0 + \sum_{k=1}^{i} l\cos(\theta_k); \qquad y_i = y_0 + \sum_{k=1}^{i} l\sin(\theta_k) \qquad (5)$$

With (x_0, y_0) being the coordinates of the location of gliders deployment. If more than one glider is considered the vector will include the angle variables required to define the trajectories for all gliders $[\theta_{1,1}, \theta_{1,2}, ..., \theta_{1,N_v}, \theta_{2,1}, \theta_{2,2}, ..., \theta_{2,N_v} ...]$ specifying the first subindex the glider and the second the angle variable. Following the traditional approach in GAs, vectors of angle variables are coded into single binary strings to constitute a population. Given a particular population of binary strings and their performance (i.e. the value of the objective function), a new generation is obtained by crossovers between strings and mutations that flips "0"s and "1"s with certain probability. In the present case, crossover can be compared to small perturbations in the location of glider observations to explore the neighbouring space of solutions, while mutations provide a large scale exploration of the variable space.

III. RESULTS

The framework of the present study is a generic ocean domain of 672 Km by 672 Km. This physical dimension was chosen on the basis of the maximum distance covered by a glider during a cycle period of the profiling floats ($L_D = T_a V_g$, L_D is the dimension of the basin, $T_a = 14$ days is the appropiate cycling period to sample mesoscale features in the open ocean and $V_g = 48$ Km per day (Km/d) is the estimated glider speed).

A temperature field showing complex spatial variability has been randomly generated in the domain. The field has been created with a Gaussian power spectrum ($exp(-\kappa K^2/4)$ with K wavenumber) and random phases, on a grid of 40×40 points corresponding to a spatial resolution of 17.2 Km (see Fig. 1a). This produces a temperature field with Gaussian covariance function with decorrelation scale $\sqrt{\kappa}$. The value of κ is 10^4 Km^2 ($\sqrt{\kappa} = 100$ Km) corresponding to temperature structures of spatial scales of around few hundreds kilometers. These are the most energetic scales in the real ocean and conform the so called mesoscale processes.

A generic network of twenty profiling floats unevenly distributed is considered in the domain (fig. 1a). Floats locations resulted from drifting randomly distributed floats during 20 cycles (280 days) in a background current field (Fig. 1a). The current field has a characteristic spatial scale of variability of around 100 Km and maximum speeds of 8.6 Km/d. Periodic boundary conditions were considered during the drifting process to keep the number of buoys constant in the domain. Thus, for each buoy leaving the region a new buoy enters in the opposite size of the domain. The resulting float population is non-uniformly distributed, showing higher concentration northward than southward. Unevenly distributions of drifters usually occur in the real ocean, where profiling floats concentrate in areas of fluid convergence. Thus, regions of divergence are usually undersampled by networks of profiling floats.

Following (4), a regular grid with spatial resolution $\Delta x = 77.4$ Km ($\Delta x = 0.4\Delta n_r; \Delta n_r = 193.5$ Km) can be defined from this network configuration. For future comparison, Fig. 1b shows the temperature field interpolated to this grid.

Besides the floats, three gliders are considered to sample the temperature field. This is considered as a reasonable number of gliders for a usual research institute, although networks with up to twelve gliders have already been experimented [10]. Sensitivity analysis has shown that the spatial lag between glider observations should be order of the resolution of the grid defined by the network of floats. Spatial lags smaller than grid resolution, generate too close observations driving to a singular covariance matrix A that cannot be inverted. On other hand, important dynamic structures could be lost if a coarser resolution is employed. In this study, the distance between glider observations was fixed to $l = \sqrt{2}\Delta x = 109.5$ Km, but any other spatial lag of the same order than the grid resolution would be valid. Thus, trajectories are defined by six observation points. Finally, it is also assumed that gliders are initially deployed at the centre of the basin.

The GA employed to optimize the objective function (3),

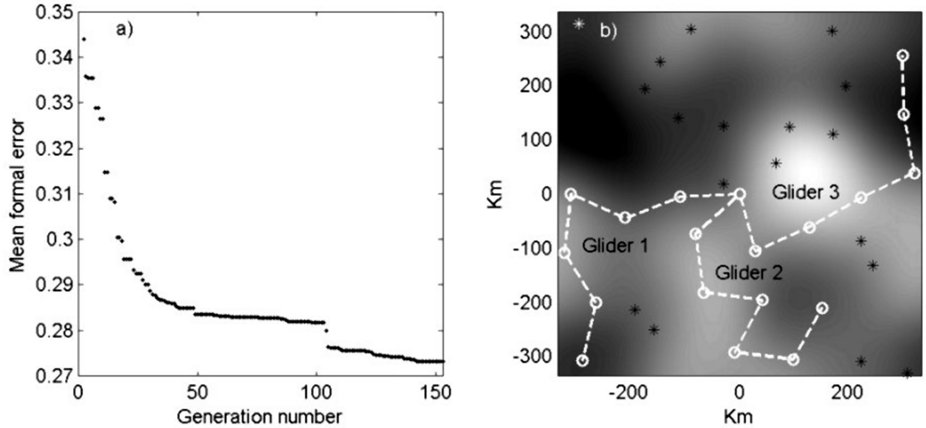

Fig. 2. a) Evolution of the objective function versus the generation number in the GA, b) optimum paths of the three gliders obtained from the GA (dashed circled lines, deployment at white circle).

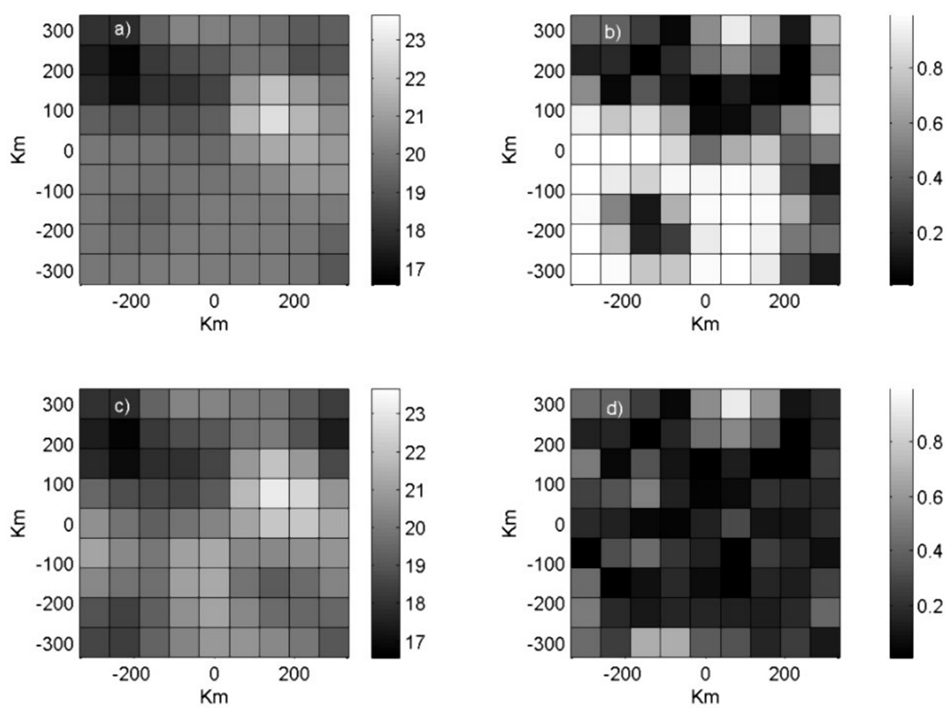

Fig. 3. Estimated field and formal error expressed in percentage of the field variance obtained from the network of profiling flows, a) and b) respectively, and obtained by optimum combination of the networks of profiling floats and gliders, c) and d).

has been initialized with a population of 300 individuals. Crossover and mutation rates were fixed to 0.8 and 0.2, respectively. Stopping criteria are given by an upper limit of 200 generations or if the algorithm is stalled for 50 generations. Fig. 2a displays the performance of the GA during the optimization process. Notice that the value of the error in the first generation would correspond to the mean formal error of a network with a total number of nodes equivalent to the number of gliders plus floats, randomly distributed. The algorithm was stalled after 153 generations. Near optimum paths for the three gliders are plotted in Fig. 2b. Essentially paths are designed to cover those areas with lowest density of profiling floats. Improvement in the interpolated field is quantitatively analysed in Fig. 3. Fig. 3a and b shows the estimated field and formal error (expressed as percentage of the field variance) respectively, when only data from profiling floats are considered. The network of profiling floats recovers part of the temperature structures at north of the domain. These are the regions where a relatively high density of floats exists. Conversely, poor estimations are

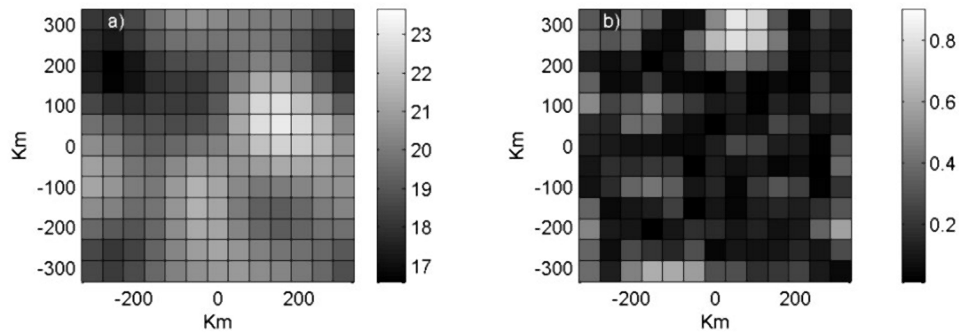

Fig. 4. Estimated field and formal error expressed in percentage of the field variance obtained from the overall network of profiling flows and gliders

achieved at the south region. Specifically, Fig. 3b shows an area of 300 Km by 500 Km southwest the domain where the formal error is maximum. The mean formal error of the temperature field estimated from the network of profiling floats is 0.59.

Fig. 3c and d display the estimated field and the formal error obtained when data from gliders following optimum paths are included. Now, the estimated field closely resembles the real field displayed in Fig. 1b. Glider data were able to substantially reduce the formal error on those areas poorly covered by the network of profiling floats, specially at southwest region. The mean formal error of the estimated field resulting from optimally combining both networks is 0.27.

The resulting overall network including observations from gliders and profiling floats is characterized by a random data spacing of $\Delta n_r = 130.12$ Km. Thus, the sampled field can be optimally interpolated to a regular grid with slightly higher resolution ($\Delta x = 52.04$ Km) than the one employed in the optimization process, Fig. 4a. The mean formal error is 0.259, Fig. 4b. Notice that this network can be defined after fixing the number of gliders observations, on the base on the random data spacing of the profiling network.

IV. Concluding remarks

This work has investigated the problem of finding optimum cast strategies to sample the ocean, combining networks of profiling floats and gliders. The motivation for this research is given by the recent emergence of these new robotic platforms, supported by the idea that using networks of distributed platforms is the most efficient way to sampling the ocean. An added motivation is the need to optimize the use of the scarce resources nowadays available to monitor the ocean.

Networks of profiling floats and gliders already exist. The former are constituted by drifting robotic platforms with uncontrollable motion, being their spatial sampling determined by the ocean currents. This induces unequal sampling rates of the ocean regions. Conversely, gliders observations can be planned a priori. For this reason, the solution proposed here to optimally combine networks of floats and gliders is to plan gliders paths in order to mitigate sampling deficiencies of the network of floats. Specifically, a GA has been employed to find those gliders paths that reduces the mean formal error when the sampled field is interpolated to a regular grid. Difficulties have been found in this procedure due to the different range of spatio-temporal scales of variability covered by both networks. Planning high resolution observations for gliders would drive to a singular covariance matrix, making interpolation unfeasible. On other hand, little benefit is gain from sampling at coarse resolutions. A sampling lag of gliders observations of the same order than the grid size based on the random data spacing of the network of floats, appeared an adequate election to make compatible both networks. Results indicate a substantial increment in the quality of the interpolated field when glider data from optimum paths are included. The procedure developed here contributes to a better use of present sampling resources of the ocean.

Acknowledgments

This work has been supported by the Spanish projects REN2003-07787-C02-01 and PR2006-0147.

References

[1] A. Alvarez, A. Caiti and R. Onken, Evolutionary path planning for autonomous underwater vehicles in a variable ocean, J. Ocean. Eng.,29, 2004, pp.418-429.

[2] S. L. Barnes, A technique for maximizing details in numerical weather map analysis, J. Appl. Meteor., 3, 1964, pp 396-409.

[3] R. E. Davis, D. C. Webb, L. A. Regier and J. Dufour, The autonomous lagrangian circulation explorer (ALACE), J. Atmos. Ocean Tech., 9,1992, pp. 264-285.

[4] C. A. Doswell, Obtaining meteorologically significant surface divergence fields through the filtering property of objective analysis, Mon. Wea. Rev., 105, 1977, pp. 885-892.

[5] D. E. Goldberg, Genetic algorithms in search, optimization and machine learning, Addison-Wesley, 1989, 432 pp.

[6] F. Hernandez, P. Y. Le Traon and N. H. Barth, Optimizing a drifter cast strategy with a genetic algorithm, J. Atmos. Ocean. Tech., 12, 1995, pp. 330-345.

[7] J. H. Holland, Adaptation in natural and artificial systems, the University of Michigan Press, 1975, 183 pp.

[8] S. E. Koch, M. DesJardins and P. J. Kocin, An interactive Barnes objective map analysis scheme for use with satellite and conventional data, J. Cli. Appl. Meteo., 22, 1983, pp. 1487-1503.

[9] R. Kunzig, A Thousand Diving Robots. Dis. Mag., 1996, 17, pp. 60-63.

[10] N. E. Leonard, D. Paley, F. Lekien, R. Sepulchre, D. M. Frantantoni and R. E. Davies, Collective motion, sensor networks and ocean sampling, Proc. of the IEEE, special issue on Networked Control Systems, 2006, in press. (www.princeton.edu/ naomi/publications/2005/IEEE05LPLSFD.html)
[11] D. L. Rudnick, R. E. Davis, C. C. Eriksen, D. M. Fratantoni and M. J. Perry, Underwater gliders for ocean research, J. Mar. Tech. Soc., 38, 2004, pp.48-59.
[12] H. Stommel, H., The SLOCUM Mission, Oceanogr., 2, 1989, pp. 22-25.
[13] K. Sugihara and J. Yuh, GA-based motion planning for underwater robotic vehicle, in Proc. 10th Int. Symp. Unmanned Untethered Submersible Technology, 1997, pp. 406-415.