Error assessment of HF radar-based ocean current measurements: an error model based on sub-period measurement variance

Kenneth E. Laws and John F.Vesecy
Baskin School of Engineering
University of California Santa Cruz
Santa Cruz, California, USA 95064
kip@soe.ucsc.edu

Jeffrey D. Paduan
Department of Oceanography, Code OC/Pd
Naval Postgraduate School
Monterey, California, USA 93943
paduan@nps.edu

***Abstract*—Data from CODAR-type ocean current sensing radar systems are used here to evaluate the performance of an error indicator provided as part of the available radar data. Investigations are based on data from pairs of radar systems with over-water baselines. Approximately year-long time series are used. The radar data are the typical hourly radial measurements provided by CODAR systems. These measurements are actually the median (or mean) of anywhere between 2 and 7 sub-hourly measurements collected by the radar system. The error indicator under examination is based on the standard deviation (std) of the sub-hourly radials, divided by the square root of the number of sub-hourly radials. These values are recorded in the hourly data files produced by recent versions of the CODAR data processing software. Examination of the model demonstrates a positive correlation between the model and the measured baseline difference std for all baseline pairs examined. The predictive capability of the error model is demonstrated by presenting its use as a data discriminator and by examination of time series of sliding boxcar samples of radar data. Baseline difference std for data rejected by a threshold based on the error model is shown to be significantly higher than for the data retained. The results presented here demonstrate potential to improve assessment of the HF radar current measurement uncertainty. Such improvement has potential to benefit all applications of HF radar data, including for example, Lagrangian particle tracking and surface current assimilation into numerical models.**

Keywords-High frequency radar, HF radar, ocean currents, current measurement, errors

I. INTRODUCTION

Maps of near surface ocean currents obtained using HF radar-based sensors have potential for assimilation into numerical circulation models (e.g., [1-3]) for surface trajectory analysis, ([4]) and for analysis of other kinematic and dynamic quantities of the vector fields, such as divergence ([5]). Divergence of the vector fields is of great interest but is particularly sensitive to effects of errors in the current vector field estimates. Predicting Surface trajectories has practical applications for search and rescue operations, contaminant tracking, oil spill mitigation and recently for evaluating connectivity of coastal marine Protected Areas [6]. All applications of HF radar measurements are to some extent limited by and would benefit from improved understanding of errors in the measurements.

Errors in HF radar measurements have been estimated through comparisons with current meter–derived velocities, drifter comparisons, radar-to-radar comparisons and simulations. Reported errors range from about 7 to 19 cm s^{-1} [7-10]. Wider ranges of errors have also been observed, 9 to 27 cm s^{-1}, see [11]. Ohlmann et al. [12], used dense arrays of drifters to evaluate differences in radar-drifter comparisons over varying time and space scales and attribute smaller observed differences, 3 to 5 cm s^{-1} to better spatial scale correspondence obtained by averaging over the region covered by the multiple drifters. Studies that employ radar-to-radar overwater baselines to estimate errors are less common. Yoshikawa et al. [13] used over water baselines and report errors of 5 to 13 cm s^{-1}.

Several simulation studies have looked at current estimation errors. Laws et al. [14] used simulations with specific current scenarios, including uniform current and current jet scenarios. Their results show that errors depend on the structure of the currents present i.e., they vary as a function of radar azimuth related to the radial component of the currents present. Toh [15] used simulation methods with specific current profiles to examine errors with CODAR SeaSonde radar systems as a function of SNR. He also examined effects of distorted antenna patterns. de Paolo, and Terrill, [16] used various simulated current scenarios and focused on limitations of CODAR's data processing methods. Laws et al. [17] used simulations to examine errors related to antenna patterns. They demonstrate, for optimal SNR conditions and a variety of ocean current and wave condition scenarios, that errors related to antenna patterns range from about 3 to 7 cm s^{-1} depending on level of distortions. They also show that radial current errors vary with azimuth depending on antenna pattern distortions, but were not able to provide a dependable characterization of the dependence. Errors in HF radar measurements are also

This work was sponsored by the *National Oceanic and Atmospheric Administration*

likely to depend on wave and other conditions as well as RF interference (when present). The errors can therefore be assumed to vary as a function of time with the changing currents and other conditions.

The desire to find a point-by-point metric for estimating uncertainty in the HF radar measurements is the primary motivation for the work presented here. The candidate metric is the std of sub-period current estimates recorded by the CODAR SeaSonde radar system and its performance is examined in the work that follows.

HF radar systems generally sample the ocean currents at a higher rate than they report them and perform some form of averaging to obtain reported estimates. CODAR SeaSondes, for example, typically sample to obtain spectra every 256 seconds and average spectra over about 10 minutes. These averaged spectra are processed to obtain radial current estimates. Typically, the average (or median) of current estimates from seven ten-minute spectra form the final radial current estimates reported at (typically) hourly intervals. There is generally some overlap in the coverage time for consecutive measurements giving a coverage time longer than the data interval.

For each range azimuth grid cell in the radar's area of coverage, anywhere from zero to seven individual estimates of the radial current may be obtained. This is because the direction finding algorithm (used by CODAR) assigns current magnitudes to radar azimuth and does not necessarily obtain for each inversion, a current estimate for each azimuthal direction within the radar's coverage. For each resolution grid cell, CODAR software reports the mean (or median) current signed magnitude, the number of current estimates obtained and the std of the individual, sub-period, current estimates (σ_{HF}). Other quantities are reported but are not discussed here.

The central question that this work seeks to resolve is whether the σ_{HF} values, can be considered an estimate of the radial current uncertainty, though the size of the statistical sample is very small (2 to 7 points) and the errors may not be Gaussian in nature. The authors are aware of Previous studies that have found no correlation between the σ_{HF} values and HF radial error estimates but results of these investigations have not been published to our knowledge. We further are not aware of any users of HF radar that actively employ σ_{HF} to estimate uncertainty in the radar measurements.

Investigations presented here use radar-to-radar over-water baselines to estimate errors. This method compares two, assumed independent ocean surface current estimates for approximately coincident patches of the ocean surface. Baseline-derived errors have the advantage over in situ current meter comparisons that the horizontal scale as well as depth mismatch between HF radar measurements and current meters, thought to be a limiting factor in the relevance of those comparisons, is avoided.

II. Error Model Evaluation

A. Baseline Error Estimation

The data sets used in this study are from radar site pairs from two different geographic locations, first, a triplet of radar systems covering part of the Channel Islands region, near Santa Barbara, operated by UCSB, and second, a pair of radar systems covering a region near Bodega Bay, north of San Francisco. Plots showing the radar locations used in the analysis are shown in Fig. 1. The data from near the Bodega Bay region cover 22 months from Jan. 2009 to Oct. 2010. The data from the radar station PTM1 cover 18 months from Jan. 2009 to July 2010. The data from the station SCI1 cover 13 months from Jan 2009 to July 2010 with a gap in the data from March 2009 to Sept. 2009. The data from the station SNI1 cover 10 months from Sept 2009 to July 2010. Temporal matchup data sets were compiled for each of the baseline pairs. The number of data points in each matchup set is given in Table I.

A single radar site only measures the component of the near surface ocean current vector parallel to the radar look direction. These single component current measurements are commonly referred to as radials since their directions are parallel lines that project radially outward from the radar system location. Errors in radar current measurements are fundamentally errors in the radial components and it is radial component errors that are examined in over-water baseline comparisons.

Ideally, comparisons should be made for HF radar range/azimuth grid cells (~3 km 5 degrees for the data used here) corresponding to perfectly overlapping patches of the ocean surface, near the midpoint and along the baseline connecting the two radar systems. Because there may be systematic bias in the measurements from either of two radar systems forming the over-water baseline, regions of resolution cells containing three range cells and 10 azimuth cells are examined here and the geometric configuration with the largest negative correlation coefficient between the observations from the two radar systems is selected for each baseline pair.

Under the assumption that the current measurements obtained by the two radar systems for the selected geometry combinations are spatially and temporally coincident and that the measurements are of the same component of the ocean current differing in relative radar look angle by 180°, the two sets of measurement would be expected, in the absence of errors, to be equal in magnitude and opposite in sign. In the absence of errors, the sum of the measurements is therefore expected to be zero. A nonzero result for the sum of the radar measurements from the two sites arises due to differences in the radial current magnitudes observed by the two sites, and so, this quantity is referred to here as the baseline difference.

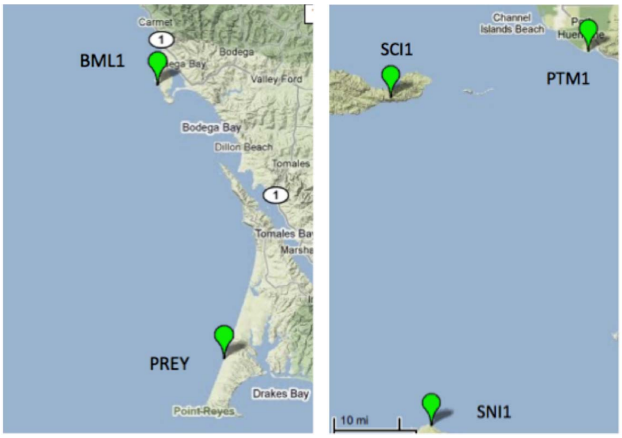

Figure 1. Radar system locations (green markers) near Bodega Bay, CA (left) and Santa Barbara (right) used for over-water baseline analysis. Panels are not to same scale. Baseline lengths are given in Table 1.

The sites used (see labels in Fig. 1), the number of data points for which data were available for each site, the baseline distance, the correlation coefficient between the two sets of observations and the std of the baseline difference are given in Table I. Selected geometries for the largest correlation magnitude baseline configuration (BML1:PREY) and lowest (SCI1:PTM1) and the corresponding radial current scatter plots are shown in Fig.2.

B. HF Radial Sub-period std

In this section we examine the relationship between the HF radial current sub-period std (σ_{HF}), described in Section I, and the errors in the hourly radial currents estimated from the measured baseline difference. Assuming normally distributed independent errors, the uncertainty in the baseline difference is given by

$$\left(\delta\Delta v_{bsl}\right)^2 = \left[\delta(v_1 + v_1')\right]^2 + \left[\delta(v_2 + v_2')\right]^2 , \qquad (1)$$

where , $v_1 = v_2$ is the component of the ocean current parallel to the radar-to-radar baseline, averaged over a region of ocean surface covered by both the resolution grid cells selected for the individual radar systems. The subscript (1,2) denotes the radar system observing the current. The additional current components, v_1' and v_2' account for differences in the error-free observations from the two radar sites. These may be due to geometrical differences between the patches of ocean observed as determined by the selection of resolution grid cells for the two radar sites or differences between the actual radar look angles and the baseline direction. Here, these contributions are assumed for simplicity to be negligible compared to the radar errors.

Ignoring limitations due to the small statistics, the uncertainty in the baseline difference measurements are expected to be related to the σ_{HF} values by

$$\left(\delta\Delta v_{bsl}\right)^2 \approx \sigma_{HF_1}^2 / N_1 + \sigma_{HF_2}^2 / N_2 , \qquad (2)$$

where N_j (j = 1,2 denotes the radar site) are the number of sub-period current estimates used in the computation the hourly radial currents. For a statistical sample of baseline difference measurements,

$$\sigma_{bsl}^2 = \left\langle \left(\delta v_{bsl}\right)^2 \right\rangle = \left\langle \sigma_{HF_1}^2 / N_1 \right\rangle + \left\langle \sigma_{HF_2}^2 / N_2 \right\rangle , \qquad (3)$$

where σ_{bsl}^2 is the predicted variance of the baseline difference and the angle brackets $\langle\ \rangle$ denote the ensemble average. This relationship is evaluated as follows using the baseline data set described in the previous section.

III. RESULTS

A. Regression Analysis

Examination of (3) was conducted for each of the baseline pairs in Table 1. The combination of radar grid cells was selected for each pair as that with the highest magnitude correlation coefficient. The correlation coefficient values, *r*, are given in the table. To evaluate (3), a 30 point sliding boxcar sampling method was used to arrange the data into ensembles. The number of points (30) was selected as a compromise between a statistically significant sample size and a large number of ensembles given the finite data set. The measured baseline difference std and the mean value of σ_{bsl} from (3) were compared for each ensemble. The correlation coefficient between baseline difference std

TABLE I. CONFIGURATION AND RESULTS

Site Pair	Baseline Information				Model Results			
	points	*Distance*	*r*	*std*	*m*	*b*	r^2	*ensembles*
		(km)		($cm\ s^{-1}$)		($cm\ s^{-1}$)		
BML1:PREY	6156	30.8	-0.89	10.2	1.62	1.0	0.52	246
SCI1:PTM1	4533	49.4	-0.62	19.4	1.93	2.0	0.46	181
SCI1:SNI1	3460	79.9	-0.85	12.2	1.35	2.8	0.44	138
SNI1:PTM1	3561	98.3	-0.69	12.9	2.05	0.0	0.29	142

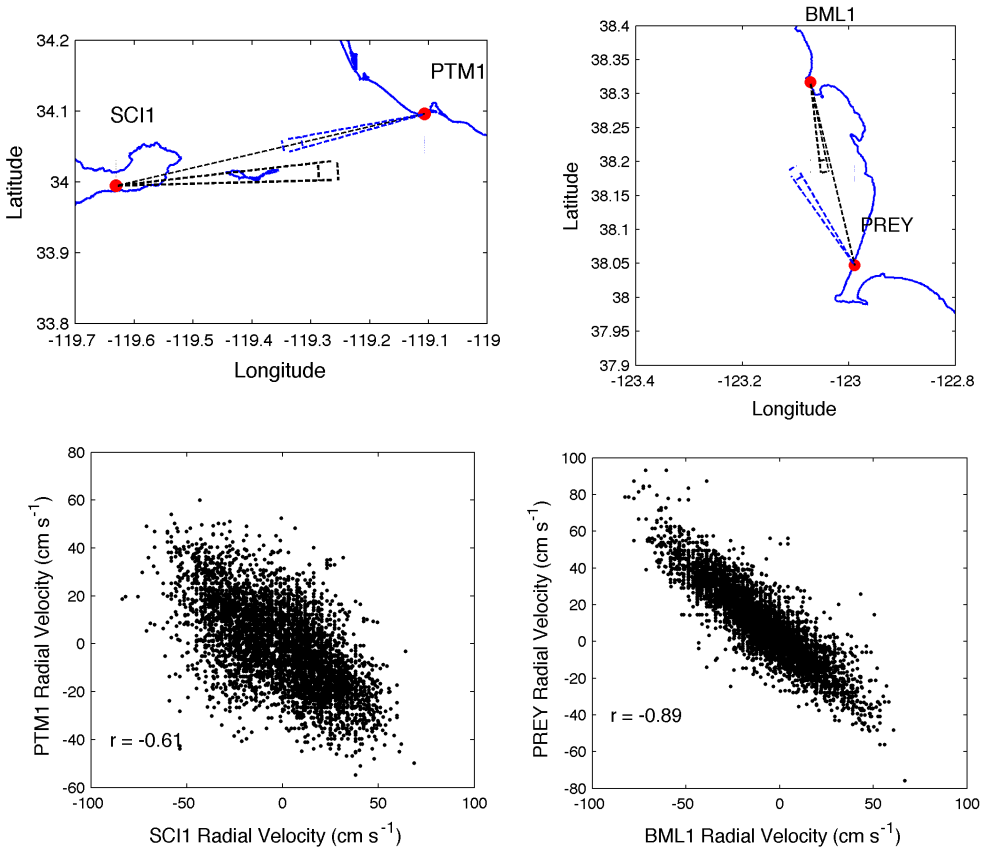

Figure 2. Selected radar resolution grid cell combinations for the SCI1:PT1 site pair (upper left) and the BML1:PREY site pair (upper right) the corresponding radial current scatter plots for are shown beneath each panel. Correlation coefficients for the geometries shown are given in the lower left corner of the scatter plots. The cases shown are for the site pairs included in this study with the best and worst correlations.

and the mean value of σ_{bsl} was computed, and a linear regression was evaluated.

Results demonstrate correlation significantly different from zero at the 1% confidence level, however, linear regression analysis yields a slope and intercept that are different than expected from (3). Examples of the scatter plots for well correlated and poorly correlated baseline site pairs are shown in Fig. 3. For these cases, the linear regression results for the slope and intercept are given on the plots. In general, the slope values obtained ranged from 1.3 to 2.0. The intercept values ranged from 0.0 to 3.9 cm s^{-1}. The coefficient of determination, r^2, relating σ_{bsl} to the baseline difference ensemble std and the number of ensembles are given in Table I for each site pair. Results indicate that σ_{bsl} accounts for about 50% of the variance over ensembles in the baseline difference std for three of the four sites and about 30% for the other.

B. Uncertainty Model

In order to develop a simple empirical model for the single-radar-site radial current uncertainty, (3) is modified by adding free parameters to adjust the resulting linear regression slope and intercept values, giving

$$\sigma_{bsl}^2 = \left\langle \left(m \cdot \sigma_{HF_1} / \sqrt{N_1} + b\right)^2 + \left(m \cdot \sigma_{HF_2} / \sqrt{N_2} + b\right)^2 \right\rangle . \quad (4)$$

The parameters m and b are derived as the values that obtain 1.0 (0.0) for the linear regression slope (intercept) respectively. The values empirically obtained for m and b are given in Table I. This model should be considered preliminary as it is based on limited data (four baseline pairs) and without theoretical justification for the need for free parameters to adjust slope and intercept values. Further investigations are needed to better understand the relationship between the σ_{HF} values and the radial current errors and to determine if the model parameters derived here apply to other radar systems and to off-baseline geometries and if they are dependent on radar range and azimuth, current conditions or other variables.

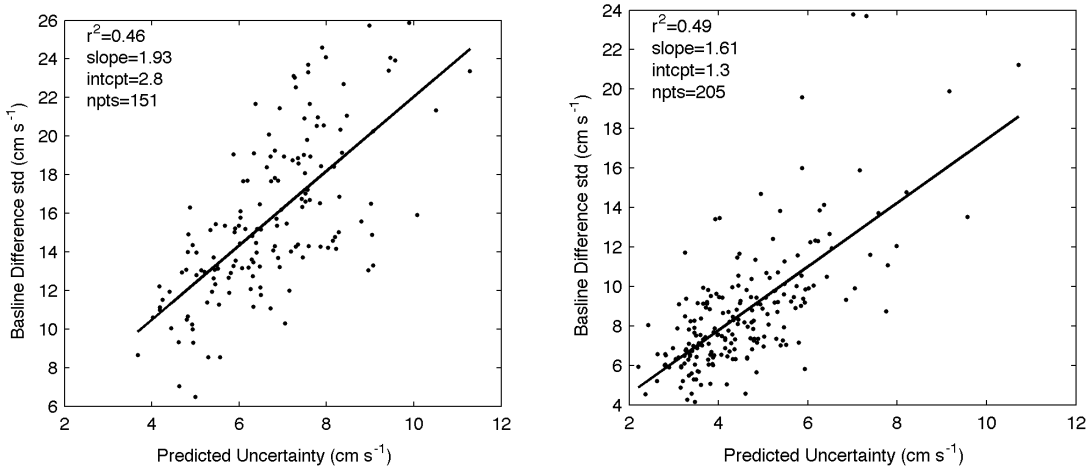

Figure 3. Examples of scatter plots and linear regression results for a site pair with poorly correlated radials (SCI1:PTM1) (left panel) and well correlated radials (BML1:PREY) (right panel) The coefficient of determination (r^2), slope, intercept and number of ensembles are included on the plots.

Time series of the boxcar sample baseline difference std and corresponding model-predicted uncertainty, for each of the four baseline pairs, are shown in Fig. 4. The plots aid in visualizing our previous assertions that uncertainty in the radar measurements fluctuates over time and that part of this fluctuation is also present in the uncertainty model (4).

As a best estimate of the parameter values, the mean of the experimental values from table I (m = 1.74, b = 1.25 cm s^{-1}) are used to construct the preliminary model for the single-site radial uncertainty,

$$\sigma_{rad} = \left(m \cdot \sigma_{HF} / \sqrt{N} + b\right) , \qquad (5)$$

and the prediction of baseline difference uncertainty,

$$\sigma_{bsl}^2 = \sigma_{rad_1}^2 + \sigma_{rad_2}^2 . \qquad (6)$$

Evaluation of the effectiveness of the model in an application for identifying data with high uncertainty is examined in the following section.

C. *Application for Data Discrimination*

To test the effectiveness of the error model as a discriminator, the baseline data set was divided on the basis of whether the model predicted baseline difference uncertainty, σ_{bsl}, estimated on a point-by-point basis using (5) and (6), was above or below a given threshold. The threshold value was allowed to vary covering the full range of predicted uncertainty for each baseline data set. The measured std of each of the divided baseline difference data sets (above and below the threshold) was computed for each threshold value and the results plotted as a function of the threshold value.

For all of the baseline pairs examined, the results (Fig. 5) show that the baseline difference std is higher for the data with predicted uncertainty above the threshold than below. For three of the four cases, the std for data above the threshold increases as the threshold increases as expected. For one case, the PTM1:SNI1 baseline pair, the std of data above the threshold increases up to a threshold level of about 18 cm s^{-1} and then decreases. This site pair data has one of the poorer radial current correlation coefficients and is the same site that had the lower value for r^2 (Table I). The reason for the threshold behavior is unknown and warrants further examination. For the BML1:PREY site pair, the data above threshold reaches a baseline difference std of nearly 3 times that of the entire data set.

IV. CONCLUSIONS

The relationship between the variance of sub-period HF radar radial current estimates and observed errors in the hourly radial current estimates has been examined. Measurements of the errors in the hourly currents are obtained from radar-to-radar over-water-baseline comparisons that, are better matched in terms of spatial scale and depth of measurement than radar-to-current meter comparisons. Results are based on relatively long time series data providing a large number (~ 3000 to 6000) of observations from four different radar site pairs. Comparisons between predicted uncertainty and measured error std were made for ensembles of data using a 30 point sliding boxcar sampling window. Results show a positive correlation between the predictions and the measurements, with the predictions accounting for about 30 to 50 % of the variance observed in the error std over the set of boxcar samples. Linear regression analysis of the relationship between the uncertainty prediction and the measured error

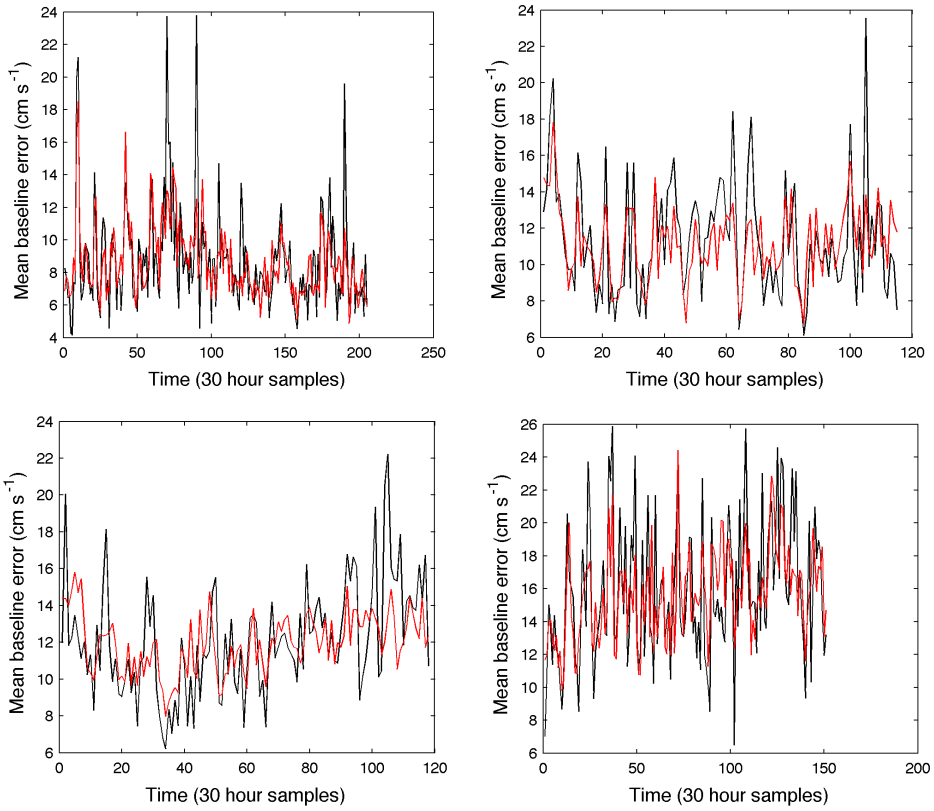

Figure 4. Time series of sliding boxcar ensembles, showing the measured baseline difference ensemble std (black line) and the model prediction for the baseline uncertainty (red line). Panels show data for the four baseline site pairs:BML1:PREY (upper left), SCI1:SNI1 (upper right), PTM1:SNI1 (lower left) and SCI1:PTM1 (lower right).

std give a slope and intercept of the best fit line to the data of 1.3 to 2.0. and 0.0 to 3.9 cm s^{-1}, respectively.

A simple model for estimating single radar site radial uncertainty has been presented. The model was calibrated using the baseline data set to find model coefficient values that adjust the slope and intercept of the linear regression line to 1.0 and 0.0, respectively, for each of the baseline site pairs. As an example for application, a more general model, using the average of the coefficient values obtained, is used as a data discriminator to divide the baseline data sets based on a threshold value. Results indicate that the simple empirical model would be effective, statistically, for identifying radial current data with high errors, but more work is needed to develop and validate a reliable error model.

ACKNOWLEDGMENT

Thanks to Jack Harlan with the *National Oceanic and Atmospheric Administration,* IOOS program, for advice and support of this work. Data used in this study were collected and archived under the Coastal Ocean Currents Monitoring Program and obtained with assistance from Brian Zelenke, with the Center for Coastal Marine Sciences, Cal Poly.

REFERENCES

[1] J.D. Paduan and I. Shulman, "CODAR data assimilation in the Monterey Bay area," *J. Geophys Res.*, 109 2004.

[2] P. R Oke, J. S. Allen, R. N. Miller, G. D. Egbert, and P. M. Kosro, "Assimilation of surface velocity data into a primitive equation coastal ocean model," *J. Geophys. Res.* 107 2002.

[3] J. L Wilkin, H. G. Arango, D. B. Haidvogel, C. S. Lichtenwalner, S. M. Glenn, and K. S. Hedström, "A regional ocean modeling system for the Long-term Ecosystem Observatory," *J. Geophys. Res.*, 110 2005.

[4] M. Roughan, "Observations of divergence and upwelling around Point Loma, California," *J. Geophys Res.*, 110 2005.

[5] D. M. Kaplan and J. Largier, "HF radar-derived origin and destination of surface waters off Bodega Bay," California, Deep Sea Research Part II: Topical Studies in Oceanography, Volume 53, Issues 25-26, The Role of Wind-Driven Flow in Shelf Productivity - Results from the Wind Events and Shelf Transport (CoOP WEST) Program, pp 2906-2930, December 2006.

[6] B. Zelenke, M. A. Moline, B. H. Jones, S. R. Ramp, G.B. Crawford, J.L. Largier, E.J. Terrill, N. Garfield, J.D. Paduan, and L. Washburn, "Evaluating connectivity between marine protected areas using CODAR high-frequency radar," *OCEANS 2009, MTS/IEEE*

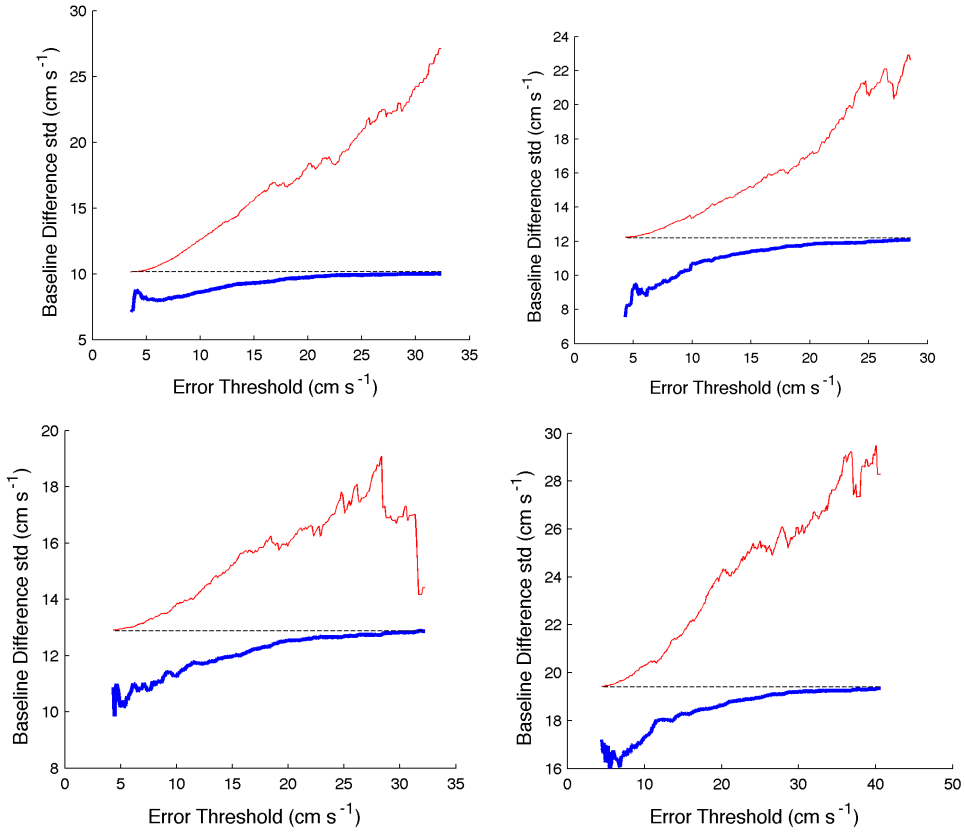

Figure 5. Measured baseline difference std for data above (red line) and below (blue line) a threshold based on model estimated uncertainty, for the baseline site pairs, BML1:PREY (upper left), SCI1:SNI1 (upper right), PTM1:SNI1 (lower left) and SCI1:PTM1 (lower right). The horizontal dashed line on each plot shows the baseline difference std for the entire data set.

Biloxi - Marine Technology for Our Future: Global and Local Challenges, vol., no., pp.1-10, 26-29 Oct. 2009.

[7] J. T.Kohut, and S. M. Glenn, "Improving HF Radar Surface Current Measurements with Measured Antenna Beam Patterns," *J. Atmos. Oceanic Technol.*, 20, pp. 1303–1316, 2003.

[8] B.M. Emery, L. Washburn, and J.A Harlan., "Evaluating radial current measurements from CODAR high-frequency radars with moored current meters," *J. Atmos. Ocean Tech.*, 21, pp. 1259-1271 2004.

[9] D. M. Kaplan, J. Largier and L. W. Botsford , " HF radar observations of surface circulation off Bodega Bay (northern California, USA)," *J. Geophys Res*, 110, 2005.

[10] J. D. Paduan; K. C. Kim, M. S. Cook, F. P. Chavez, "Calibration and Validation of Direction-Finding High-Frequency Radar Ocean Surface Current Observations," *IEEE J. Oceanic Engineering* 31, no.4, pp.862-875, Oct. 2006.

[11] R D. Chapman and H. C. Graber, "Validation of hf radar measurements," Oceanography, 10, 2 1997.

[12] C. Ohlmann, P. White, L. Washburn, B. Emery, E. Terrill, and M. Otero, "Interpretation of Coastal HF Radar–Derived Surface Currents with High-Resolution Drifter Data," *J. Atmos. Oceanic Technol.*, 24, pp. 666–680, 2007.

[13] Y. Yoshikawa, A. Masuda, K. Marubayashi, M. Ishibashi, and A. Okuno, "On the accuracy of HF radar measurement in the Tsushima Strait," *J. geophys res*., 111 2006.

[14] K. E. Laws, D. M. Fernandez and J.D. Paduan, "Simulation-based evaluations of HF radar ocean current algorithms", *IEEE J.Oceanic Engineering,* 25, 4, pp. 481–491, 2000.

[15] D. Toh, "Evaluation of Surface Current Performance by SeaSonde High Frequency Radar Through Simulations," Masters thesis, *Naval Post Graduate* school 2005, unpublished.

[16] T. de Paolo and E. Terrill, "Skill Assessment of Resolving Ocean Surface Current Structure Using Compact-Antenna-Style HF Radar and the MUSIC Direction-Finding Algorithm," *J. Atmos. Oceanic Technol.*, 24, pp. 1277–1300, 2007.

[17] K. E. Laws, J. D. Paduan and J. F. Vesecky, "Estimation and assessment of errors related to antenna pattern distortion in CODAR high-frequency radar ocean current measurements," *J. Atmos. Oceanic Technol.*, 27, 6, pp. 1029-1043 (2010).