
Article:
Wyatt, L.R. orcid.org/0000-0002-9483-0018, Mantovanelli, A., Heron, M.L. et al. (2 more authors) (2018) Assessment of surface currents measured with high-frequency phased-array radars in two regions of complex circulation. IEEE Journal of Oceanic Engineering, 43 (2). pp. 484-505. ISSN 0364-9059

https://doi.org/10.1109/JOE.2017.2704165

Assessment of surface currents measured with high-frequency phased-array radars in two regions of complex circulation

L. R.Wyatt, *Senior Member, IEEE*, A. Mantovanelli, M. L. Heron, *Fellow, IEEE*,
M. Roughan and C. R. Steinberg

Abstract—Surface velocity data from two WERA high frequency (HF) ocean radar systems, deployed as part of the Australian Integrated Marine Observing System (IMOS), are compared with near surface currents obtained from drifters and ADCPs (acoustic Doppler current profiler). We evaluate data from two contrasting locations in the first detailed evaluation of the IMOS HF radar surface velocities. HF radar measurements are generally robust but demand quality-control procedures to eliminate obvious errors and outliers that appear temporarily or systematically in the data. A number of different quality control procedures and filters are applied and assessed including Taylor diagrams, Hampel and Savitzky-Golay filters. In addition the need for and effect of averaging are discussed. The radar measurements of surface current agreed better with the near-surface drifter currents than with the subsurface ADCP currents. Nonetheless the ADCP comparisons are consistent with those previously reported in other regions. The value of the Taylor Diagram for comparing different surface current data sets and processing approaches is demonstrated. Noise levels in the radar current spectra are used to estimate the error in the measurements and in some cases, these errors were found to approach the precision of the radar estimates. Our results give guidance on the most useful temporal sampling resolution. In particular we show that, at these sites and these operating frequencies, using 10-minute sampling without further averaging does not provide additional information because the higher frequencies are dominated by noise. Averaging the radials over 30-minutes may be sufficient for many applications.

***Index Terms*— ADCP, coastal currents, Coffs Harbour, drifter, Eastern Australia, Great Barrier Reef, Hampel filter, HF radar, quality-control, Savitzky-Golay filter, Taylor Diagram, WERA**

ACORN and ANMN are Facilities of Australia's IMOS. IMOS is a national collaborative research infrastructure, supported by Australian Government. It is led by University of Tasmania in partnership with the Australian marine and climate science community Additional funding has been provided by the Australian Research Council through a LEIF grant which paid for the CBG radar purchase, James Cook University, the Australian Institute of Marine Science, the University of Queensland, the NSW State government.

L. R. Wyatt is with the School of Mathematics and Statistics, The University of Sheffield, Sheffield, S10 2TN, UK and is also an Adjunct Professor in the College of Science, Technology and Engineering, James Cook University, Townsville, QLD 4811, Australia (e-mail: l.wyatt@sheffield.ac.uk).

A. Mantovanelli was with the School of Mathematics and Statistics, The University of New South Wales, NSW 2052, Australia, the Australian Institute of Marine Science (AIMS), Townsville, QLD 4810, Australia; AIMS@JCU and The College of Science, Technology and Engineering, James Cook University, Townsville, QLD 4811, Australia(e-mail: alessandra.mantovanelli@gmail.com).

M. L. Heron is with The College of Science, Technology and Engineering, James Cook University, Townsville, QLD 4811, Australia (e-mail:mal.heron@ieee.org).

M. Roughan is with the School of Mathematics and Statistics, The University of New South Wales, NSW 2052, Australia (e-mail: mroughan@unsw.edu.au).

C. R. Steinberg is with the Australian Institute of Marine Science (AIMS), Townsville, QLD 4810, Australia (e-mail: c.steinberg@aims.gov.au).

I. INTRODUCTION

Near real-time information on the spatial variability of ocean currents is essential for coastal management (e.g., oil, pollutants and larvae tracking), navigation and search and rescue operations. High-frequency ground-wave ocean radars (HF radars) are unique in providing measurements of the two-dimensional surface current field over large coastal areas (from nearshore to 50-200 km offshore) with high temporal resolution (Paduan and Washburn [1]). HF radars are remote sensing platforms installed on land that transmit radio waves (at frequencies of 3-30MHz) in a radial path to the ocean surface and receive the backscattered signal from sea surface gravity waves of one-half of the incident radio wavelength approaching or receding with respect to the radar look direction. This is the Bragg-scattering effect first described by Crombie [2].

The Bragg-scattering generates two discrete peaks in the radar backscattered signal power (often referred to as Doppler) spectrum that are shifted (Doppler-shift) by the superimposed current field. The difference between the expected Bragg-peak (the phase speed of the Bragg-wave as determined from first order ocean wave hydrodynamics) and the observed Doppler-shift is used to calculate the speed of the radial component of the surface current moving toward or away from the receive antenna (Paduan & Washburn [1]). Radial currents produced by two or more radars, spatially separated but looking at the same patch of water at different viewing angles, can be summed to calculate the two-dimensional surface current velocity vector (Gurgel et al [3]); the separation between two radar stations determines the domain of the mapped region.

Here we evaluate the performance of two phased-array WERA HF ocean radar systems (Gurgel et al [3]; available from Helzel Messtechnik GmbH, Germany) operating on contrasting continental shelves with complex circulation, and assess the errors in current vectors after combining two radials with different errors. In addition, quality-controlled current data extracted from HF radar, acoustic Doppler current profilers (ADCP) and satellite-tracked surface drifting buoys (drifters) at the same location and instant of time are compared to cross-validate the measurements.

Bottom-mounted ADCPs provide high temporal resolution measurements of both the horizontal and

vertical current components but they are single point platforms that cannot resolve the near surface currents (top 6-15% of the depth profile) due to side-lobe contamination (Gordon [4]). Drifters provide near real-time Lagrangian measurements of the flow (integrated over the drogue depth) giving point-by-point information about the horizontal currents along the path of motion; drifter data are however sparse in time and space. Ocean HF radars provide the two-dimensional current field with high temporal (10-60 minutes) and spatial (0.3-10 km) resolutions (Gurgel et al [3]; Paduan et al. [5]) but their measurements are restricted to the surface layer thus incorporating wind and wave-driven near surface currents (Graber et al. [6]; Cosoli et al. [7]). In addition, each HF radar data point is an average over a target patch of a few square kilometres while an ADCP samples over depth cells and beam separations of a few metres.

Typically HF radar-derived surface currents are validated against subsurface ADCP currents and currents estimated from drifting buoys with both surface and subsurface drogues, despite the differences in their spatial and temporal resolutions and sampling depths. Shay et al [8] was one of the early studies of this kind and used a number of different statistics to quantify the differences between these measurements, including root-mean-square error (rms), correlation coefficients from linear regression for radials and vector components and complex correlations for the vectors. Like many other such studies, only short (1-3 month) data sets were available allowing ad-hoc methods to be used to de-spike and interpolate data for subsequent spectral analysis. Robinson et al [10] summarise a number of these validations as well as the results of their much longer (2 year) time series comparison. A simple automated de-spiking procedure was applied and least squares fitting used to obtain tidal coefficients thus avoiding the need to fill gaps. For the scalar comparisons in these studies, r^2 of 0.3 to more than 0.9 and rms differences between 0.03-0.25 $m \cdot s^{-1}$ are typical and reflect the expected spatial-temporal variability of the currents and the inherent instrumental errors and resolutions (Graber et al. [6]; Shay et al. [11]).

Although ocean radar technology has been previously validated in other locations, we provide here evaluation of the WERA radar performance at two sites along the Eastern coast of Australia subjected to complex coastal dynamics in order to provide confidence in these HF radar measurements. Many previous

studies have looked at the accuracy of the radials measured by HF radar. Here we focus on the accuracy of the data product of most interest to users, *i.e.* the vector currents. Some of the methods we have applied are not widely used in the HF radar community and part of the aim of this paper is to demonstrate their value for this application. We explore the use of the Taylor diagram as a tool to simplify the comparison of different quality control (QC) and averaging strategies by avoiding long and dense tables of comparison statistics. We test the application of filters to automatically remove outliers and smooth data sets which often contain gaps. Simple lowpass filters applied in e.g. Shay et al [12], Cosoli et al [7] can only be applied to evenly sampled data with no gaps. We use spectral methods to determine noise levels and hence limits to the accuracy of the surface current measurements.

The HF radar and ADCP data sets used here are from the southern Capricorn Bunker Group (CBG ~24° S) area of the Great Barrier Reef (GBR) and Coffs Harbour (COF ~30° S) on the coast of eastern Australia. The CBG and COF regions both exhibit complex circulations and large spatial and temporal variability of the coastal currents imposed by winds, tides, large scale circulation (East Australian Current, EAC) and topography (Mao and Luick [13]; Schaeffer et al. [14,15]; Mantovanelli et al. [16] ; Roughan and Middleton [17,18]).

Section II of the paper presents a description of the data sets followed in Section III by a description of the error analysis from radial to vector currents and the HF radar QC procedures that have been used here. In Section IV, we evaluate the effect of the QC procedure on data precision and accuracy and perform a detailed statistical comparison of the current data obtained from the different platforms (HF radar, ADCP, drifters) to assess the quality of the HF radar surface current measurements. The results are discussed in Section V and this is followed by some concluding remarks.

II. Instrumentation and Data Sets

A. *HF radar data*

A system of two high-frequency (HF) ground-wave WERA phased-array radars, operate both at the CBG (22-24° S; Fig. 1a) and COF (30-31° S; Fig. 1b), as part of the IMOS Australian Coastal Ocean Radar

Network (ACORN). The radar domains of the CBG and COF radar systems are shown respectively in Fig. 1a and Fig. 1b, including only regions with acceptable geometric dilution of precision (GDOP, see Section IIIA), and their configuration details are summarized in Table 1. The radar domain in the CBG includes the reef lagoon on the shallow continental shelf (~70 km wide and less than 50 m deep), the shelf break, which drops steeply from 50 m to 400 m depth, and a small portion of the offshore region (Fig. 1a). The COF radar domain covers the narrow continental shelf (< 30 km), the shelf-break and deeper off-shore waters (up to ~150 km from the coast; Fig. 1b). Data analyses encompassed the period between 1 November 2009 and 30 November 2010 (CBG system) and 1 September 2012 and 30 September 2013 (COF system); times are given in Coordinated Universal Time (UTC).

The radar azimuthal resolution is determined assuming a $\lambda/2$ spacing (λ is the radio wavelength) between the N receive antenna elements giving an approximate half-power beamwidth of 102/N (Skolnik [19]). This beamwidth was modified by the window used in the beam-forming and a Dolph-Chebyshev window (Harris [20]) with alpha = 3 and used for the calculation of the azimuthal resolution in Table 1. This increased the non-windowed beamwidth by a factor of 1.618 and reduced the side lobes. The maximum azimuthal range of the radars is 90° at CBG and 120° at COF (Table 1).

The radar data are processed using standard WERA software onto a rectangular grid with horizontal resolution of 4 km (CBG) and 1.5 km (COF) using a combination of digital beam-forming to grid locations and interpolation from the original range-resolved data. This has the effect of over-sampling at long ranges and under-sampling at short ranges relative to the intrinsic (azimuth and range cell) radar resolution. Radial components from two radar stations are combined to calculate the east-west (u) north-south (v) current components at each grid location (discussed below in section IIIB). Radial (every 10 minutes) and vector (obtained from 1 hour averages of the radials) data are freely available from the IMOS data portal (www.aodn.org.au) in both real time (FV00) and as Level 1 (FV01) quality-controlled (QC) NetCDF files. The analyses here use the radial FV01 files. The results are compared with the monthly-aggregated 1 hour averaged current vector FV01 files with and without taking account of the additional quality flags

contained in these files.

Errors in the determination of the radial currents may occur due to radio-wave interference, moving ship reflections, improper determination of the Bragg peaks and/or of the angle of arrival from the scattering patch through inadequate calibration of the beam pattern, azimuthal and range resolution and additional uncertainties arising during the vector mapping (Gačić et al. [21]). The HF radar data presented here were quality-controlled as described in Section III.

TABLE I
DETAILS OF THE HF RADAR SYSTEM CAPABILITIES AND CONFIGURATION FOR THE CBG AND COF SYSTEMS

	CBG	**COF**
Operating frequency (MHz)[1]	8.34	13.92
Azimuthal resolution (°)	13.75	10.36
Range cell resolution (km)	4.50	1.50
Measurement depth (m) (~ $\lambda_{radar}/8\pi$)	1.4	0.9
Sampling period each station (minutes)	5	5
Integrated sampling period (minutes)	10	10
Distance between sites (km)	138	74
Maximum configured offshore distance (km)[2]	250	150
Number of Transmitting antennas	4	4
Number of Receiving antennas (phased array)	12	16
Transmitter power (W)	30	30
Data acquisition	near real time	near real time

[1]Note that both systems have undergone a frequency change since these measurements were made, with the frequencies being 9.33(CBG) and 13.5(COF) at time of publishing.
[2] This is set to be larger than the maximum expected range for sea echoes at the given frequency.

B. ADCP data

ADCP data sets were obtained from IMOS (www.aodn.org.au) for the GBR and Coffs Harbour (CH) moorings (see locations in Fig. 1). Three ADCPs are moored in shallow waters (45-60 m) in the southern GBR, namely Heron Island North (HIN), Heron Island South (HIS) and One Tree Island (OTE), with bin sizes (1-4 m) and sampling rates (10-30 minutes) varying among the deployments (for details see Appendix 1, Table A1). The two Coffs Harbour ADCPs, located on the 70 m (CH070) and 100 m (CH100) isobaths, have bin sizes of 4 m and sampling rate of 5 minutes. All mooring stations were equipped with upward-looking four-beam broadband ADCPs (RDI Workhorse Sentinel Instruments operating at 307.2 and 614.4 kHZ; Table A1), with beam angles slanted at 20°. The set-up enabled bin-mapping correction for tilting, three beam solution and Earth coordinates (East/North/Up), with corresponding velocity components

denoted by u, v, w.

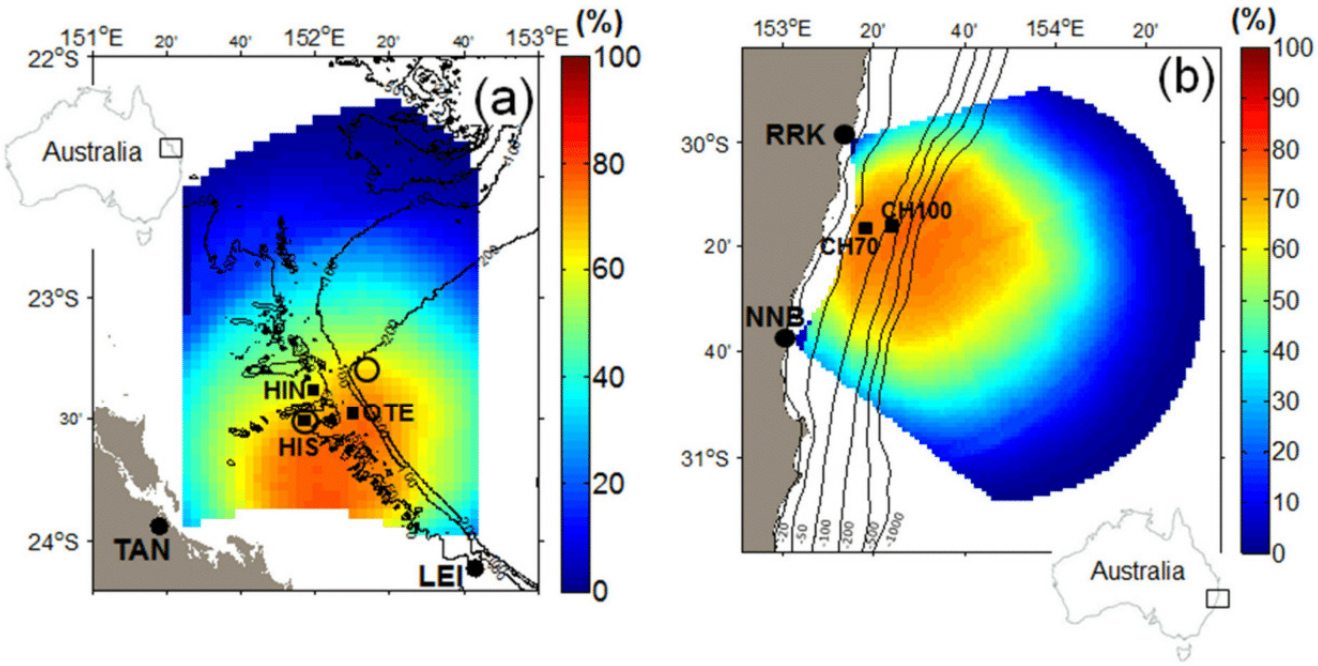

Fig. 1. Location of the radar stations (•), and the HF radar domain for surface current measurements for (a) the southern Great Barrier Reef (CBG) and (b) Coffs Harbour (COF) HF radar systems, respectively. The ADCP moorings are indicated by black squares and mooring names are labelled; the drifter clusters' release positions are indicated by circles. Spatial and temporal data coverage (%) for (a) the GBR region between 1 November 2009 and 30 November 2010 and (b) Coffs Harbour region between 1 September 12 and 31 September 2013 for areas with acceptable GDOP are indicated with the colour-shading (see colorbar).). Bathymetry is indicated by the black lines.

The ADCP data were corrected for magnetic declination and quality-controlled to remove data of poor quality i.e. (*i*) contaminated by side-lobes near the surface (first 6% of the water column), (*ii*) with low signal to noise (i.e. correlation magnitude for two or more beams < 64 counts), (*iii*) subjected to instrument tilting larger than 20° and (*iv*) identified as outlier when the measured value was larger than the time-averaged value of velocity magnitude for each depth stratum plus 5 times its standard deviation.

ADCP data were posteriorly re-mapped into uniform depth strata (as wide as the bin), taking into account the temporal changes of the bin depth in relation to the surface due to tidal variations measured by the built-in pressure sensor (Mantovanelli et al. [22]). When the bin size of a particular mooring changed during the analysed year, data were averaged over a depth stratum corresponding to the size of the largest bin for consistency.

Data from the two shallowest depth strata with more than 50% of valid observations during each

deployment after QC were used to compare with HF radar data. There were small differences in the comparisons for these two depth strata. The cases with better agreement are presented here, with centre of the depth strata at 5.5 m (HIN), 5 m (HIS), 10 m (OTE), 10 m (CH070) and 14 m (CH100). ADCP *u*-and *v*-velocity components were averaged over 30 minutes and interpolated into the same timeframe as the HF radar data for uniformity and the standard deviations of the averaged values are used as a proxy of the ADCP errors. The 30 minute averaged ADCP data are used for the comparisons with the 10 minute and 1 hour averaged radar data as well as with the 30 minute averaged radar data.

C. Drifter data

Drifters, used for comparison with the CBG radar, were built, released and monitored by the Australian Institute of Marine Science (AIMS). Four drifters separated by 1-2 km were released in a square array (less than 50 minutes apart) (*i*) on 22 April 2010 in the slope region (200-300 m water depth) and (*ii*) on 25 April 2010 in shallow shelf waters (depths < 50 m) south of Heron Island (Fig. 1a). All drifters stayed inside the HF radar domain for 16-20 days. Drifters consisted of a Davis drogue (0.45 m wide and 0.9 m long) attached to a 60 cm long PVC cylinder (that houses the electronics) linked to a surface float (diameter ~20 cm) with only 10 cm emerging in air (drag area ratio ~ 40). The entire drifter assembly occupied the first 2 m of the water column in a good approximation to the HF radar depth of measurement of 1.4 m (Table 1).

Drifters were satellite-tracked by a GPS Satellite Messenger giving a position (UTM in m, Universal Transverse Mercator projection) every 10 minutes, with nominal accuracy of 6.4 m. The actual accuracy, based on positions transmitted for 14 days by a drifter kept in a fixed location, was within ±20 m of the mean position ~80% of the time, which corresponds to an error in the speed of about ±0.03 $m \cdot s^{-1}$. Positions were interpolated into discrete time intervals of 10 minutes to match the HF radar timeframe. Drifter current velocity components (*u*, *v*) were calculated by centred difference of the Easting (Northing) UTM coordinates divided by time, and posteriorly averaged over one hour. Coarse outliers were manually removed.

III. HF RADAR QUALITY-CONTROL AND ERROR ANALYSIS

A. Radial accuracy

The Level 1 QC radial data (FV01 files) uses a 'swarming' method (Prytz [23]) to find and categorise the Bragg peaks and hence flag the radial speed and direction data. For the analysis presented here we retain data with an initial QC flag of 1 (good data) or 2 (probably good data). The standard deviations of the radials reported in the FV01 file are determined from the width of the primary Bragg peaks.

The radial velocity resolution is 8.1 $cm \cdot s^{-1}$ at COF and 13.5 $cm \cdot s^{-1}$ at CBG based on the frequency resolution in the radar Doppler spectrum (0.0075 Hz); the radial velocity accuracy is improved to 3.8 $cm \cdot s^{-1}$ (COF) and 6.3 $cm \cdot s^{-1}$ (CBG) by the standard WERA processing, which averages 5 overlapped Doppler spectra and uses a centroid method to determine the location of the Bragg peak (Barrick [24]).

In this paper we consider the impact of additional averaging by comparing (*i*) the 10-minute radial data set with (*ii*) the 30 minute running-averaged radial dataradials, which further improves the accuracies to 2.2 $cm \cdot s^{-1}$ (COF) and 3.6 $cm \cdot s^{-1}$ (CBG) and (*iii*) the 1 hour-averaged radial data (single value per hour) obtained from IMOS portal,with corresponding resolutions of 1.6 $cm \cdot s^{-1}$ (COF) and 2.6 $cm \cdot s^{-1}$ (CBG). The 30 minute running-averages on the radials are over 3 measurements stored in the middle instant of time, thus retaining the measurement discrete time interval of 10 minutes.

B. Calculation of the current velocity components and errors

The two radials are combined in the usual way (Paduan and Washburn [1]) to give the *u* (east-west) and *v* (north-south) current components.The errors in the current vectors are calculated assuming that the errors in the two radials (from the two separate radars) are uncorrelated and potentially different. The variances σ_u^2 and σ_v^2 in the *u* and *v* current components respectively can be estimated as (Wyatt [25]):

$$\sigma_u^2 = \frac{\sigma_{2a}^2 \cos^2\theta_1 + \sigma_{1a}^2 \cos^2\theta_2}{\sin^2(\theta_2 - \theta_1)} \tag{1}$$

$$\sigma_v^2 = \frac{\sigma_{2a}^2 \sin^2\theta_1 + \sigma_{1a}^2 \sin^2\theta_2}{\sin^2(\theta_2 - \theta_1)}, \tag{2}$$

where θ_1, θ_2 are the bearing angles and σ_{1a}^2 , σ_{2a}^2 are the variances of the radial speeds (*i.e.* squared values of the radial standard deviations reported at each discrete time in the FV01 files) from stations 1 and 2, respectively. For the cases where additional averaging is carried out, the averaged radial variances are the averages of the individual variances, i.e. it is assumed that there is no correlation between the data sets contributing to the average. The variance, σ_V, of the velocity magnitude, V, is given by:

$$\sigma_V^2 = \left(\frac{u}{V}\right)^2 \sigma_u^2 + \left(\frac{v}{V}\right)^2 \sigma_v^2 + 2\left(\frac{u}{V}\right)\left(\frac{v}{V}\right) C_{uv},$$

$$C_{uv} = \frac{-1}{sin^2(\theta_1-\theta_2)}(\sigma_{2a}^2 \sin\theta_1 \cos\theta_1 + \sigma_{1a}^2 \sin\theta_2 \cos\theta_2),$$

where C_{uv} is the covariance of *u* and *v* (Wyatt [24]). The square roots of the calculated variances give the standard deviations (σ_u, σ_v, σ_V) of *u*, *v* and *V* that are used to assess the measurement errors in the HF radar current vector. The same procedure was used to calculate the standard deviations in the current vectors reported in the hourly averaged IMOS files.

C. The geometric dilution of precision (GDOP)

GDOP reflects the effect of the radar geometry on the accuracy of the current vectors derived from the radials measured at the two radar sites. It is related to the angle between the radars at each location (as shown in the denominator in Equations (3) and (4)). GDOP values can be calculated separately for the *u* (GDOP east) and *v* (GDOP north) components of the current vector (following Chapman et al. [26] terminology). Here we use their modulus given by:

$$\text{GDOP} = \sqrt{\frac{2}{\sin^2\theta}}, \tag{3}$$

where θ is the intersection angle between the beams from the two radars to the cell location given by $\theta = |\theta_1 - \theta_2|$. Note that this result assumes that the errors in the two radials are the same (i.e. $\sigma_{1a} = \sigma_{2a}$ in Equations (1) and (2)), so this procedure is not exact but provides a useful indicator of the likely error. The GDOP spatial distribution for CBG (Fig. 2a) and COF (Fig. 2b) radar sites show areas of favourable geometry/higher precision (small GDOP, blue region) and areas of largest signal degradation along the antenna pair baseline and in the far field (high GDOP, red and yellow regions) for similar looking angles. The thick grey lines in Fig. 2 show locations where the intersection angle between the radial beams of the

two different radars is with the range: 30° < θ < 150 °, corresponding to GDOP ~ 3. This has been shown to be a reasonable bound for good data quality data (Shay et al. [12]; Cook and Shay [27]; Robinson et al. [10]) and is the value used in this paper.

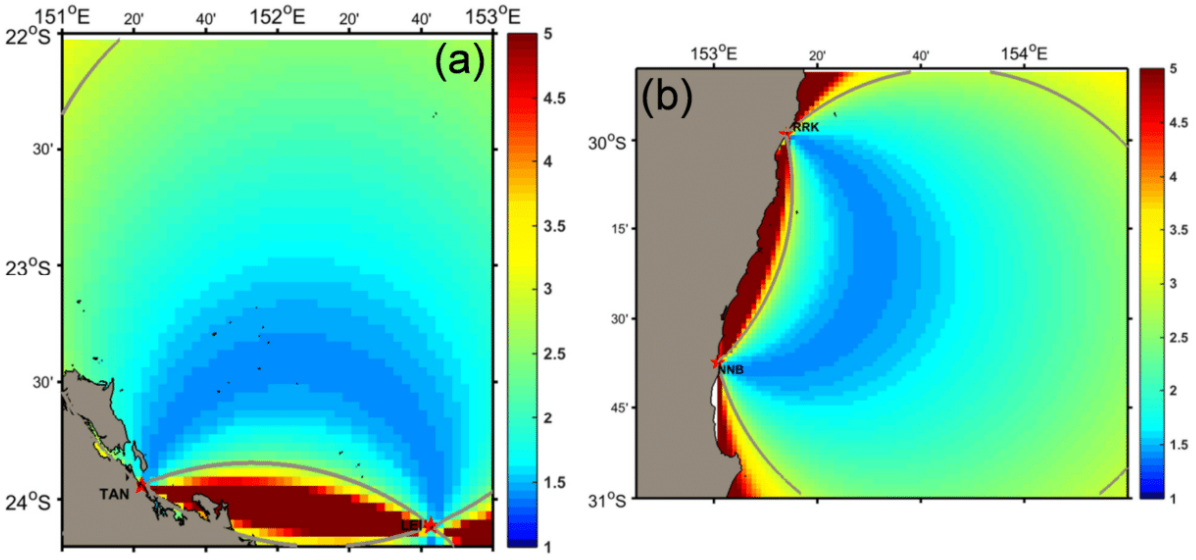

Fig. 2. Maps of geometric dilution of precision (GDOP) for (a) the CBG and (b) COF radar systems. The thick grey lines show locations where the intersection angle between the radial beams of the two different radars is within the range: 30° < θ < 150 °

D. Average current and standard deviation maps

Fig. 3 (CBG) and Fig. 4 (COF) show the standard deviations of the annual averages of the HF radar 30-minutes averaged variances of the current components (σ_u, σ_v) and current magnitude (σ_V). These show an overall decrease in the accuracy away from the radar stations (where temporal coverages are also lower; see Fig. 1a,b) and in areas with higher GDOP (see Fig. 2a,b). The standard deviations for areas with over 50% of temporal coverage (Fig. 1a,b) have the same order of magnitude as the theoretically predicted current resolutions given in Section IIIA.

Annual averages of velocity magnitude show weak currents (< 0.1 m·s^{-1}) and dominant onshore flow for the CBG region (Fig. 3d) and a predominantly southward flow in the COF region with stronger currents east of the slope under higher influence of the EAC (Fig. 4d).

E. The HF radar data quality-control and processing

In section IV, we compare a number of different QC approaches applied to the HF radar data. The QC steps adopted for the radar current data analyses starting from IMOS FV01 radials are outlined in Table 2, and those applied in the data analyses starting from the IMOS current velocity vectors are described in Table 3. The cases in Table 2 all have a time step of 10 minutes, those in Table 3 have a time resolution of 1 hour.

The QC steps in Table 2 involve: (*i*) applying GDOP (Equation (3) referred to as RAW10 in Table 2); (*ii*) using a maximum threshold for both radial and vector speed magnitude of 3 $m \cdot s^{-1}$ to remove coarse outliers and setting a minimum acceptable Bragg signal to noise ratio of 10 dB (i.e. ratio = 10 log *SNR*, where *SNR* is the value given in the NetCDF file) for each grid point (QCT10 in Table 2) to ensure a good separation between the Bragg peaks and the background noise (Fernandez et al. [28]); (*iii*) reducing spikes by applying a Matlab *hampel* filter, which calculates the median and the standard deviation of 7 data points replacing the mid-point (unless it is NaN) with the median if the mid-point varies from the median by more than 5 times the standard deviation (QCH10 in Table 2); (i*v*) starting from the 30 minute running-averaged radials (after GDOP and SNR tests, and discarding radial speeds above 3 $m \cdot s^{-1}$, as in QCT10 and QCH10 in Table 2), plus additional removal of outliers in the current components (*u*, *v*) by accepting maximum speeds less than the time-averaged value for the analysed period plus five times its standard deviation calculated separately for each velocity component(QC30, Table 2).

TABLE 2
STEPS FOR THE QC PROCEDURE STARTING WITH FV01 RADIAL DATA. THE CASES BEING COMPARED ARE LABELLED WITH THE LETTERS (B-E, H-I) FOR CONSISTENCY WITH THE RESULTS SHOWN IN FIG. 7 AND FIG. 9

Radial Steps	Vector Steps	Case
Apply GDOP	Calculate *u,v* (using Equations (1) and (2))	**RAW10 (B)**
Apply speed and SNR threshold to **B**	Calculate *u,v*	**QCT10 (C)**
Apply hampel filter to **C**	Calculate *u,v*	**QCH10 (D)**
30 minute running-averaged radial speeds and variance using **C**	Calculate *u,v* and apply current speed threshold	**QC30 (E)**
Apply 1-hour Savitzky-Golay **linear** filter to **D**	Calculate *u,v*	**SGL60 (H)**
Apply 1-hour Savitzky-Golay **quadratic** filter to **D**	Calculate *u,v*	**SGQ60 (I)**

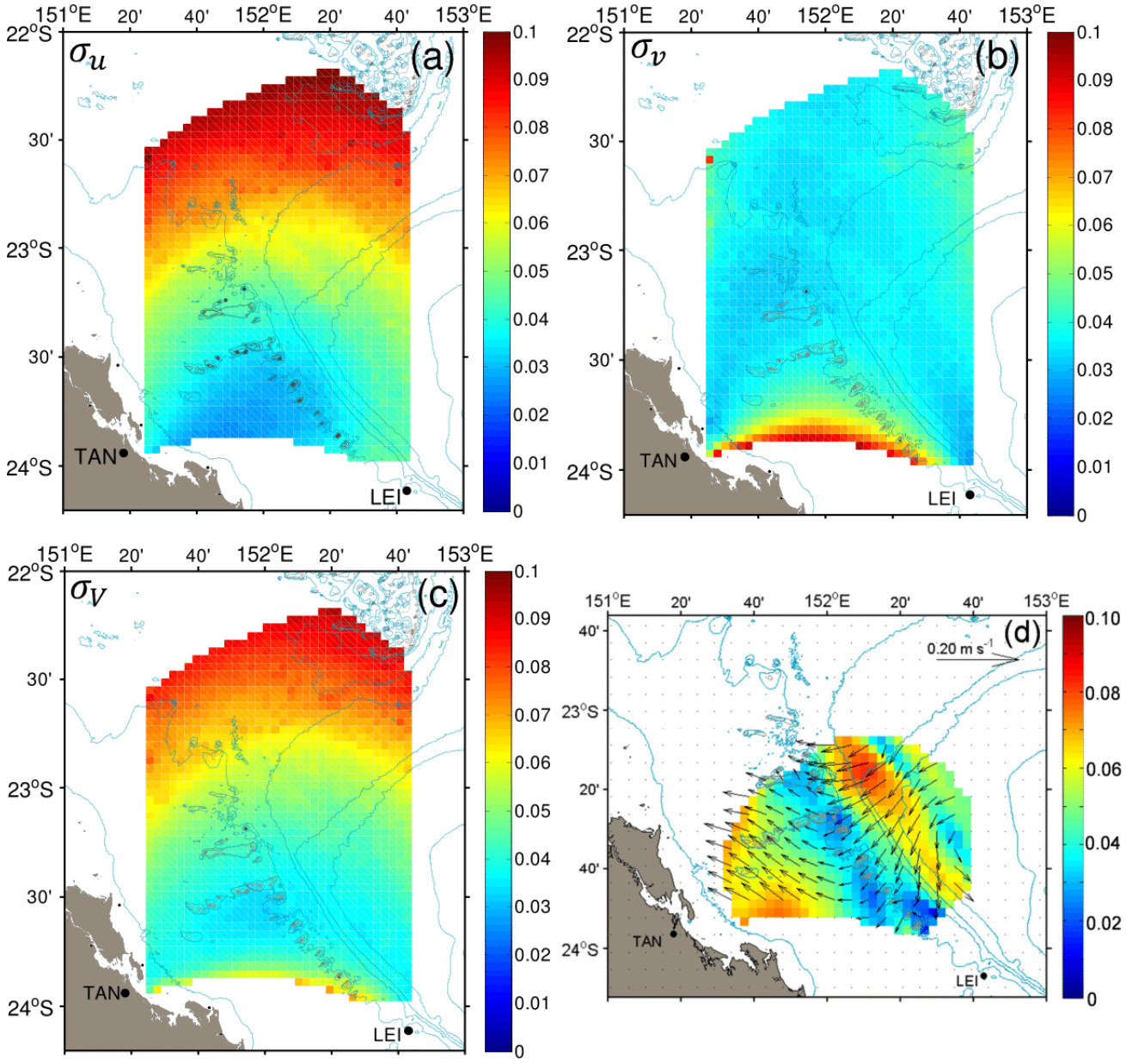

Fig. 3. Mean standard deviation calculated as the square root of the mean variance between 01/11/2009 and 31/11/2010 for the (a) *u*-component (σ_u, m·s^{-1}), (b) *v*-component (σ_v, m·s^{-1}) and (c) current absolute magnitude (σ_V, m·s^{-1}) of GBR radar system; (d) averaged current magnitude (V, m·s^{-1}; see colour bar) and direction (arrows) between 01 November 2009 and 31 November 2010 for areas with temporal coverage above 50% over the analysed year.

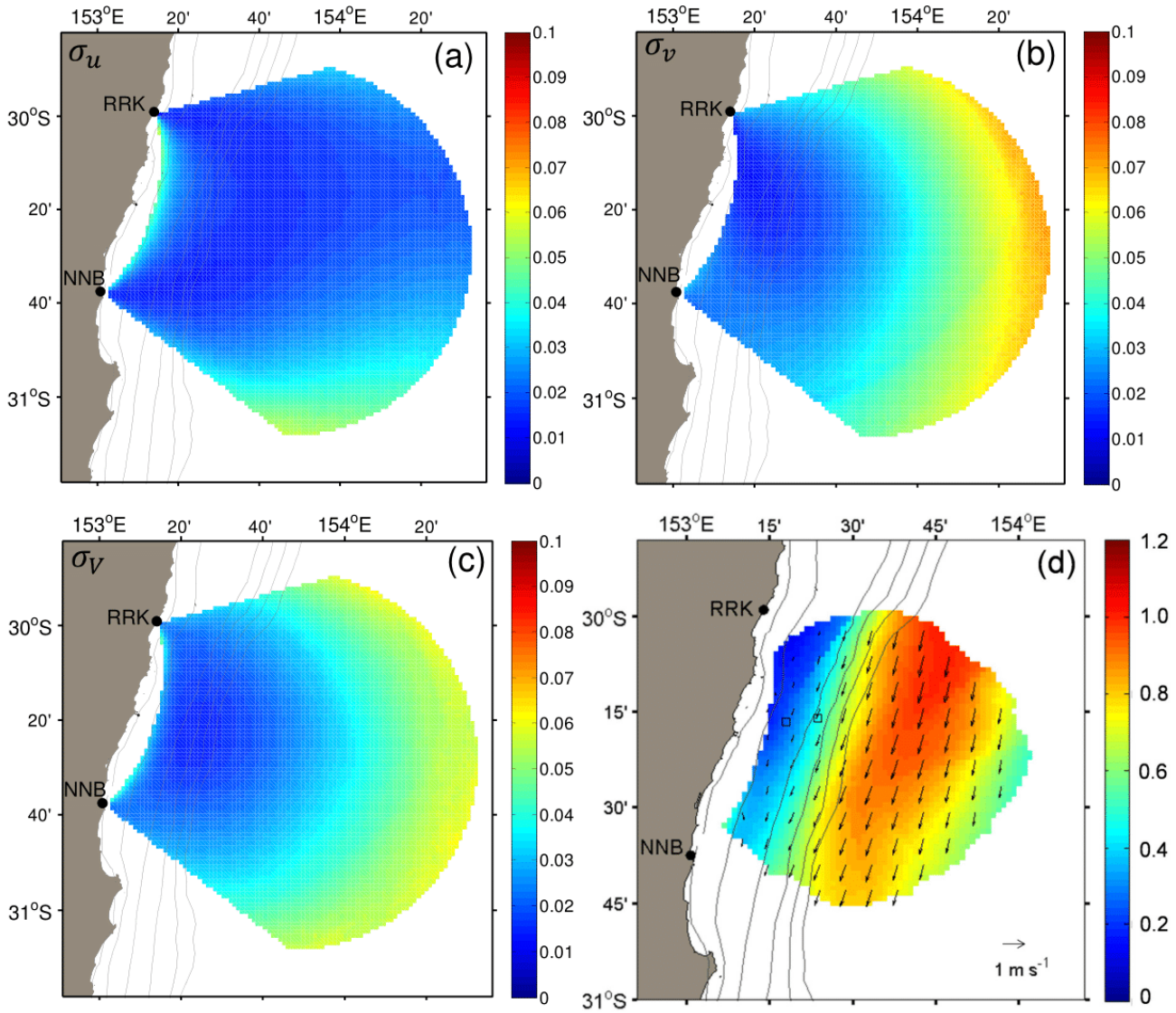

Fig. 4. Mean standard deviation calculated as the square root of the mean variance between 01/09/2012 and 30/09/2013 for the (a) *u*-component (σ_u, $^{-1}$), (b) *v*-component (σ_v, m.s^{-1}) and (c) current absolute magnitude (σ_V, m.s^{-1}) for the Coffs Harbour radar system; (d) averaged current magnitude (V, m.s^{-1}; see colour bar) and direction (arrows) between 01 September 2012 and 30 September 2013; for areas with temporal coverage above 50% over the analysed year.

We assessed here the 1 hour averaged current (u, v) data (IMOS FV01) (RAW60 in Table 3) with flags 1 or 2 (QC60 in Table 3) in the NetCDF files. These flags mean that the original radial data had signal to noise is > 8 dB, radial speeds less than 3 m.s^{-1} at COF or 2 m.s^{-1} at CBG and at least 3 such radial measurements during the hour; the vector speeds satisfy the same thresholds and the GDOP criterion is also satisfied.

TABLE 3
STEPS FOR THE QC PROCEDURE STARTING WITH IMOS CURRENT VECTORS. THE CASES BEING COMPARED ARE LABELLED WITH LETTERS (F, G) FOR CONSISTENCY WITH THE RESULTS SHOWN IN FIG. 7 AND FIG. 9.

Vector Steps	Case
Read FV01 1 hour averaged current vector data files	**RAW60 (F)**
Apply QC flags 1 and 2. These account for GDOP and signal to noise thresholds.	**QC60 (G)**

HF radar data often contain temporal and spatial gaps generated by instrument failure, weak signal-to-noise, radio frequency interferences (such as ionosphere variations) and ocean wave conditions (maximum range is reduced in high seas) and power availability (Wyatt et al. [29]). Therefore, the spatial coverage of HF radar current velocity data over time serves as a performance indicator. A higher data return and spectral quality is expected near the radar stations, with the signal availability being reduced at longer ranges because of signal propagation losses (Wyatt et al. [29]). Percentages of valid HF radar current data (over one year) for case E at CBG (Fig. 1a) and COF (Fig. 1b) regions show maximum coverages between the two radar stations and a gradual reduction of the coverage with increasing range, as expected.

The presence of temporal gaps prevents the use of many numerical filters to further smooth the data. An exception to this is the Savitzky-Golay filter (Orfanidis [30]) which can be applied to data with gaps in the time series. It also has an implementation in Matlab, although modifications were required in order to calculate the variances of the smoothed radial current estimates (see Appendix 2 for the details). We have assessed the impact of applying this filter to smooth the 10 minute (QC10) data over a 1 hour window, considering both its linear (SGL60); Table 2) and quadratic (SGQ60; Table 2) forms.

Cases QC10, QC30, QC60, SGL60 and SGQ60 in Tables 2 and 3 are the new approaches that are being assessed in this paper. We have also compared 1, 3, 7, 13, and 25 hour smoothing of the (*i*) 30-minute averages (QC30) and (*ii*) 1 hour IMOS current data (QC60) against the ADCP data with similar smoothing, with a view to the potential use of this filter for some oceanographic applications where more smoothing may be warranted (see Section IV-B-5).

IV. CROSS-VALIDATION AMONG THE DIFFERENT PLATFORMS

A. Methods for data comparison

The centred root mean square difference (E; Equation (4)) and the centred correlation coefficient (R, Equation (5)) were applied in the comparisons among the different data sets, since these parameters, together with the standard deviation of the radar time series, can be combined in one plot, namely the Taylor diagram (Taylor [31]). The Taylor diagram has not been widely used in HF radar evaluations but it has the advantage of making the comparisons between the different data sets and quality control procedures relatively straight forward. In Equations (4) and (5), f is used generically for u- and v-components of the flow and current speed (V) and the subscripts r, m are respectively used for the HF radar (f_r), and either the ADCP mooring or the drifter data (depending on the comparison) (f_m) for n = 1, 2, .. N observations in time. Similarly, the standard deviations for the u, v, V speeds of the HF radar and ADCP or drifter time series are generically represented by s_r and s_m, respectively. The parameters E and R are calculated as follows:

$$E = \left\{ \frac{1}{N} \sum_{n=1}^{N} \left[\left(f_r(n) - \overline{f_r} \right) - \left(f_m(n) - \overline{f_m} \right) \right]^2 \right\}^{\frac{1}{2}} \tag{8}$$

$$R = \frac{\frac{1}{N} \sum_{n=1}^{N} \left(f_r(n) - \overline{f_r} \right) \left(f_m(n) - \overline{f_m} \right)}{s_r s_m} . \tag{9}$$

where f is the parameter being compared (e.g. u), s is its standard deviation; subscripts r denotes the radar measurements and m the comparator. N is the number of observations in the comparison.

In addition, the bias, standard deviation, root mean square difference (rms), slope and intercept of linear regressions with 95% confidence limits and the coefficient of determination (and equivalently the correlation coefficient) were also determined.

Three methods were used for the comparison of directional and vector differences. The first method calculates the mean difference between two current measurements with its 95% confidence interval and concentration (Bowers et al. [32]). The concentration parameter is large if the spread in the directional

difference distribution is small (*i.e*, agreement is good) and vice-versa. The second method gives the circular correlation coefficient between the two directions (Fisher and Lee [33], Fisher [34]). Finally, the vector correlation was determined as described by Kundu [9].

These analyses were applied to the different mooring locations in the CBG (HIN, HIS, OTE) and COF (CH070, CH100) regions and to the drifter comparison. In addition, comparisons between ADCP and radar spectra are used (Section IV-B-3) to estimate the intrinsic uncertainty in the *u* and *v* components derived from the spectral noise levels.

B. Total current comparison

1) HF radar versus ADCP comparison for Coffs Harbour (COF)

Comparisons between HF radar and ADCP *u*- and *v*-components using the 1 hour averaged IMOS current data (QC60) at the CH070 and CH100 mooring locations are shown in Fig. 5. For this figure each pair of ADCP and radar *u*- or *v*-components was placed into 40 x 40 bins (across the speed range shown) and the numbers in each bin were counted. Bins with less than 1% of data were excluded to remove the effect of outliers. Therefore, Fig. 5 shows 99% of the data. The colour coding of the remaining bins indicates the numbers of observations in each bin normalised by the number of observations in the bin with the maximum number of data points. Lighter colours indicate bins containing larger proportions of the data.

Higher correlation coefficients were found for the *v*-component (roughly aligned with the EAC) in both locations (CH070 and CH100 moorings). Similar standard deviations were observed in the *v*- and *u*-component comparisons, except at the CH70 location. The CH70 mooring is closer to the coast and therefore the radar data may be influenced by antenna side lobe problems or small scale current structure in this location. Lower correlation coefficients for the u-components are related to the similar scatter but smaller range than the v-components.

The Taylor diagram (Fig. 6) compares the different QC procedures:(*i*) case B (red), *i.e.* RAW10, the vectors obtained from the 10 minute data with only GDOP applied; (*ii*) case C (green), *i.e.* QCT10, the 10 minute vectors after applying a 3 $m.s^{-1}$ speed and 10 dB signal-to-noise thresholds to RAW10; (*iii*) case D

(blue) QCH10 vectors after applying the hampel filter to case C; (*iv*) case E (magenta), i.e. QC30, the 30 minute averaging after applying GDOP, SNR, radial and current components' thresholds; (*v*) RAW60, vectors from the standard IMOS 1 hour averaged current data before (F, cyan), and QC60 after applying QC flags 1 and 2 (G black); (*vi*) SGL60 current vectors after Savitzky-Golay linear (H, red) and SGQ60 quadratic (I, green) filtering of case D. Note that 1-hour Savitzky-Golay filtering has also been applied to the 30 minute averages (QC30) and this is referred to in section IV-B-4. The centred rms differences and radar time series standard deviations are both normalised with respect to the standard deviations of the ADCP time series for the *u*, *v*, and speed measurements (as specified in Fig. 6).

Very little difference was found between RAW60 and QC60 at both locations (CH070 and CH100) because both GDOP and signal to noise ratio are within good ranges. The additional steps in QC60 will be more important around the edges of the radar coverage. The figure confirms the qualitative remark above that the *u*-component comparison at CH100 is better than that at CH070. The *u*-component data benefit most from the averaging (see the E-G cluster compared with B-D) because amplitudes are smaller than those for the *v*-component and hence nearer the noise level. The IMOS 1-hour current averaging (QC60) provides a small benefit over QC30 (i.e., 30 minutes average), as evidenced by the increase in the correlation coefficient, the decrease in centred rms and normalised radar standard deviation closer to 1. Therefore, G (QC60) is closer to A (point of perfect agreement) than E (QC30), but both cases (G and E) are clearly better than the 10-minute non-averaged data. (QC10, D)

The Savitzky-Golay filter aims to smooth data sets whilst retaining high-frequency content at the expense of not removing as much noise as the average of the data points over 1 hour intervals (see Fig. A2.1 in the Appendix). Therefore the 1 hour Savitzky-Golay smoothing (cases H and I) is not expected to perform as well as the 1 hour IMOS average (case G) and this can be seen in the *u*-component comparisons at both CH70 and CH100 locations (Fig. 6 left). The figure shows that the filter is not even as effective as the 30 minute average although it has some positive impact relative to the non-averaged data. Differences in variability with frequencies of about 48 cycles per day which are not suppressed as much by the Savitzky-

Golay filter may explain the difference in performance. The averaging/filtering and additional QC had little impact on the *v*-component, which is the larger of the two components at Coffs Harbour locations (CH070 and CH100; Fig. 6 middle).

The directional and vector comparisons were better at CH100 than CH070, showing higher correlation and concentration (Table 4). It is clear that the 10 minute radar data (QCT10, no average) has the worst comparison in each case (*i.e.*, the lower correlation coefficients and concentration; Table 4) although the application of the hampel filter (QCH10) provides a small improvement. The 1 hour averaged (AC60) statistics were only marginally better than the 30 minute average (QC30) statistics. The Savitzky-Golay statistics (not included in the Table) were generally between those of QCH10 and QC30 as was the case for the Taylor diagram analysis.

TABLE 4
COF DIRECTION AND VECTOR COMPARISON OF RADAR AND ADCP CURRENTS.

	CH070				**CH100**			
Case	QCT10	QCH10	QC30	QC60	QCT10	QCH10	QC30	QC60
N	22971	22971	22871	2676	34088	34088	34077	3934
Direction difference (°)	0.62	0.47	0.75	1.37	-0.68	-0.67	-0.22	0.11
± 95% CI	± 0.50	±0.49	± 0.47	± 1.34	± 0.27	± 0.26	± 0.25	± 0.72
Concentration	2.74	2.83	3.08	3.20	5.52	5.80	6.47	6.68
Complex Correlation Coefficient	0.85	0.88	0.91	0.92	0.94	0.95	0.96	0.96
Phase Difference (º)	1.81	1.75	1.20	1.82	0.03	0.03	0.02	0.20
Circular correlation	0.46	0.47	0.51	0.54	0.64	0.66	0.69	0.70

2) HF radar versus ADCP comparison for the Great Barrier Reef (CBG)

Fig. 7 shows the CBG *u*- and *v*-components of HF radar currents against the ADCP currents at the three CBG moorings, using the same processing parameters as described in Fig. 5 for COF. In all cases there was a higher correlation for the *u*-component which has the larger range of the two components in this site. The correlations for the *v*-component were significantly lower, particularly at HIN and HIS sites, closest to the reefs. Note that the *u*- and *v*-components at OTE were closer in magnitude and their statistics lie between those of the other sites.

The Taylor diagrams for CBG are presented in Fig. 8, keeping the same notation as used for COF in Fig. 6. The averaging and additional QC had little impact on the *u*-component at the CBG sites, the results of

which are rather similar to those for the *v*-component at COF and vice versa.

The directional and vector comparisons for CBG are shown in Table 5. A larger bias between the radar and ADCP current directions was observed at the HIS site as evidenced by the larger directional difference and the phase difference from the complex correlation analysis. Whether this is evidence of a surface shear or a bias in this particular ADCP measurement is not yet established. The north-south currents (*v*-component) are very small at HIS, and the main currents are aligned in the east-west direction. Therefore, small errors in the *v*-components will lead to large errors in the derived angle for both measurement systems and this could be the reason for the bias. The complex correlation coefficient is higher at HIS than at the other sites indicating that the vector currents are in better agreement with the ADCP notwithstanding the apparent bias in direction. It is clear that the 10 minute radar data has the worst comparison in each case.

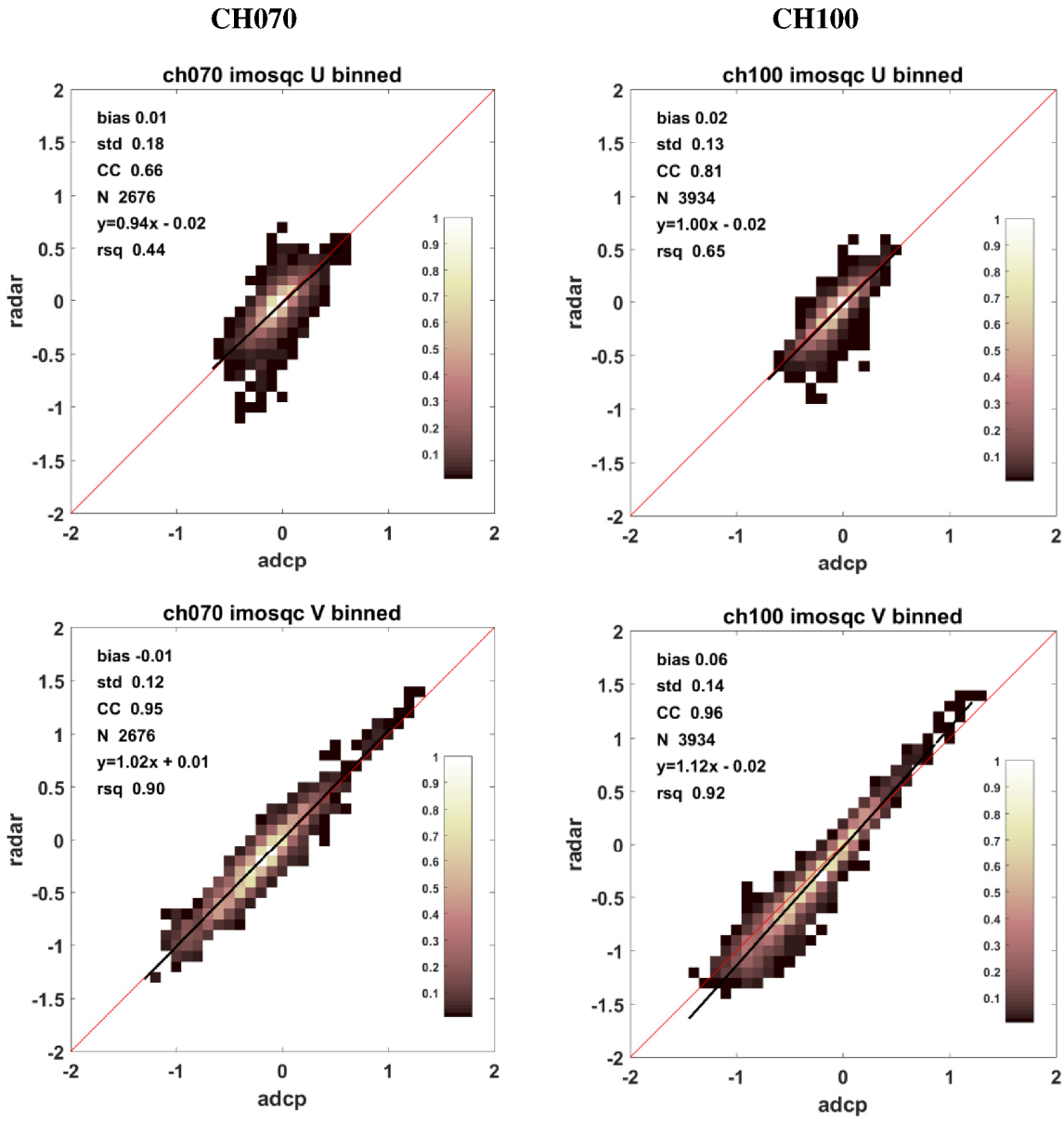

Fig. 5. Binned scatter plots of HF radar versus ADCP at Coffs Harbour (COF). The figure includes 99% of all the data. The first and second rows show respectively the *u*- and *v*-components at CH070 (left) and CH100 (right) mooring locations. The black line shows the linear regression. The colour coding of the remaining bins indicates the numbers in each bin relative to the maximum number of data points in a bin. Some of the calculated statistics are shown inside the figures: bias, std (standard deviation), CC (correlation coefficient), N (number of observations), the linear regression equation and rsq (coefficient of determination).

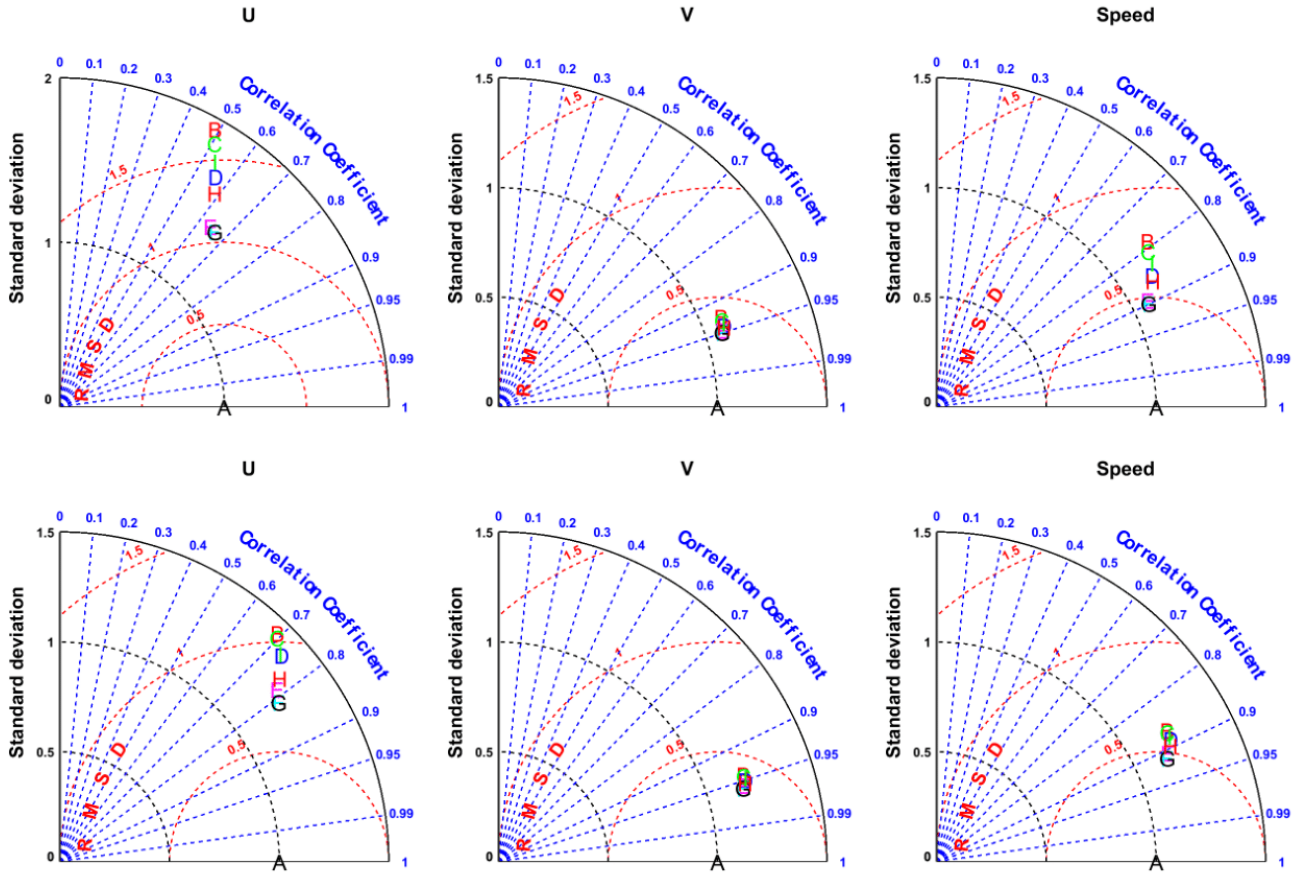

Fig. 6. Taylor diagrams showing the impact of the different QC steps for *u*, *v*, speed magnitude (*V*) for CH070 (top three figures) and CH100 (bottom three figures). A is the point of perfect agreement; B (red) is the 10-minute data with no additional QC, C (green) is the case where a 3 m.s^{-1} speed and 10 dB signal-to-noise threshold have been applied; D (blue) includes the hampel filter; E (magenta) with 30 minute averaging QC; F (cyan) with standard IMOS 1 hour averaging, G (black) with additional QC; H (red) linear and I (green) quadratic Savitzy-Golay filters. Red dashed lines show normalised centred rms difference decreasing towards A, blue dashed lines the correlation coefficient decreasing from north to east and black dashed lines the normalised radar standard deviation increasing from the origin. The normalisations used for *u*, *v*, and speed are 0.17, 0.36, 0.28 and 0.18, 0.39, 0.31 m.s^{-1} at CH070 and CH100 respectively. Note that the top left figure has a different vertical scale.

TABLE 5
CBG DIRECTION AND VECTOR COMPARISON OF RADAR AND ADCP CURRENTS. ;

	HIN			HIS			OTE		
Statistic	QCH10	QC30	QC60	QCH10	QC30	QC60	QCH10	QC30	QC60
N	32376	32375	4081	22317	22315	2771	21890	21890	2582
Direction Difference (°)	2.23 ±	1.36 ±	1.33 ±	-12.34 ±	-12.68 ±	-12.57 ±	-1.30	-1.16	-0.71
± 95% CI	0.41	0.37	1.0	0.45	0.42	1.21	±0.56	±0.52	±1.50
Concentration	2.93	3.41	3.64	3.38	3.77	3.65	2.43	2.68	2.74
Complex Correlation Coefficient	0.86	0.89	0.89	0.91	0.92	0.93	0.80	0.82	0.83
Phase Difference (°)	4.25	4.26	4.03	-10.09	-10.14	-10.23	-2.38	-2.21	-1.84
Circular correlation	0.44	0.50	0.53	0.45	0.50	0.49	0.57	0.61	0.62

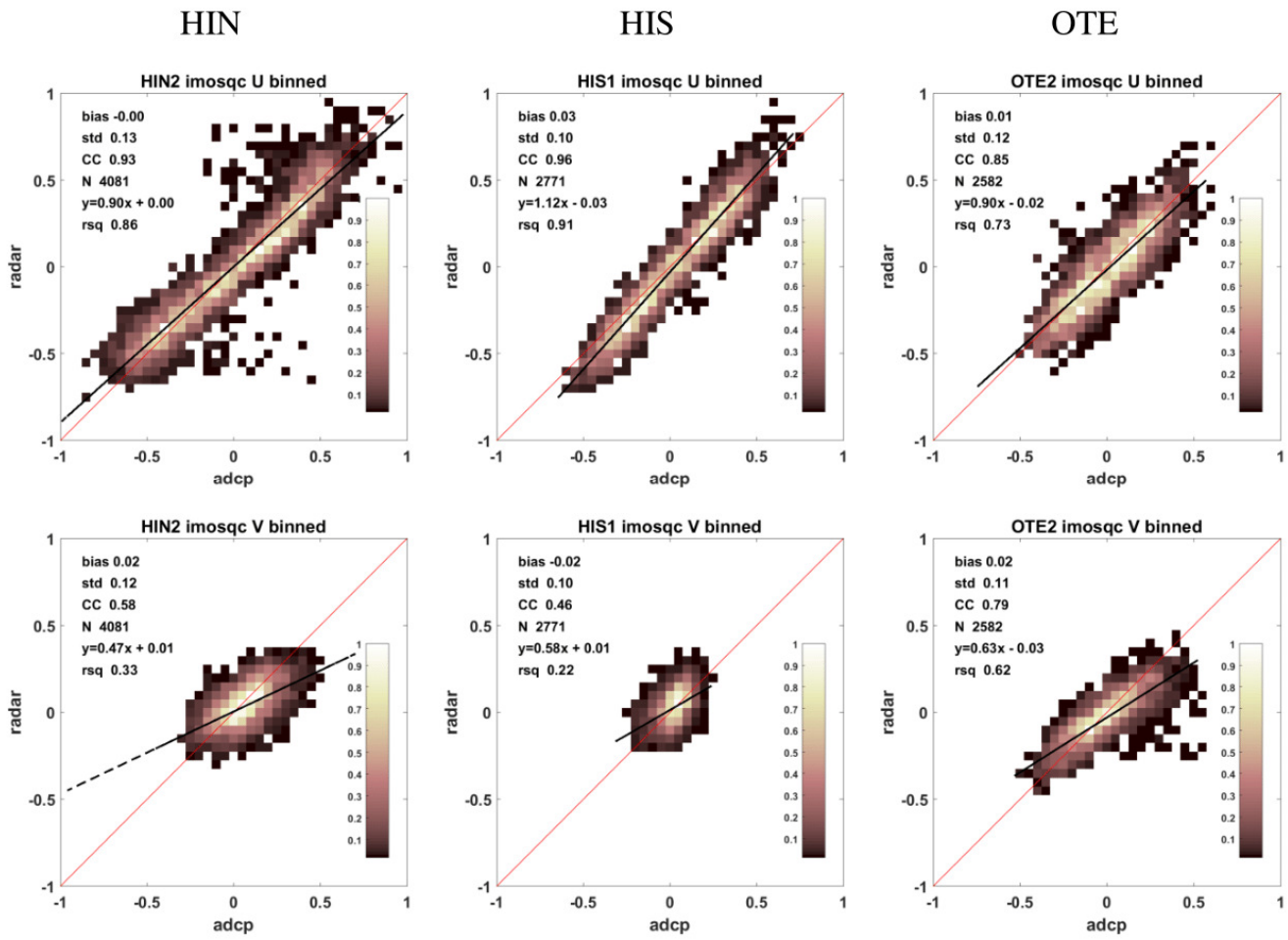

Fig. 7. Binned scatter plots of radar versus ADCP at CBG. The *u*- and *v*-components are shown respectively in the first and second rows for HIN (left), HIS (middle) and OTE (right). Colour-coding and statistics notations are as defined in Fig. 5

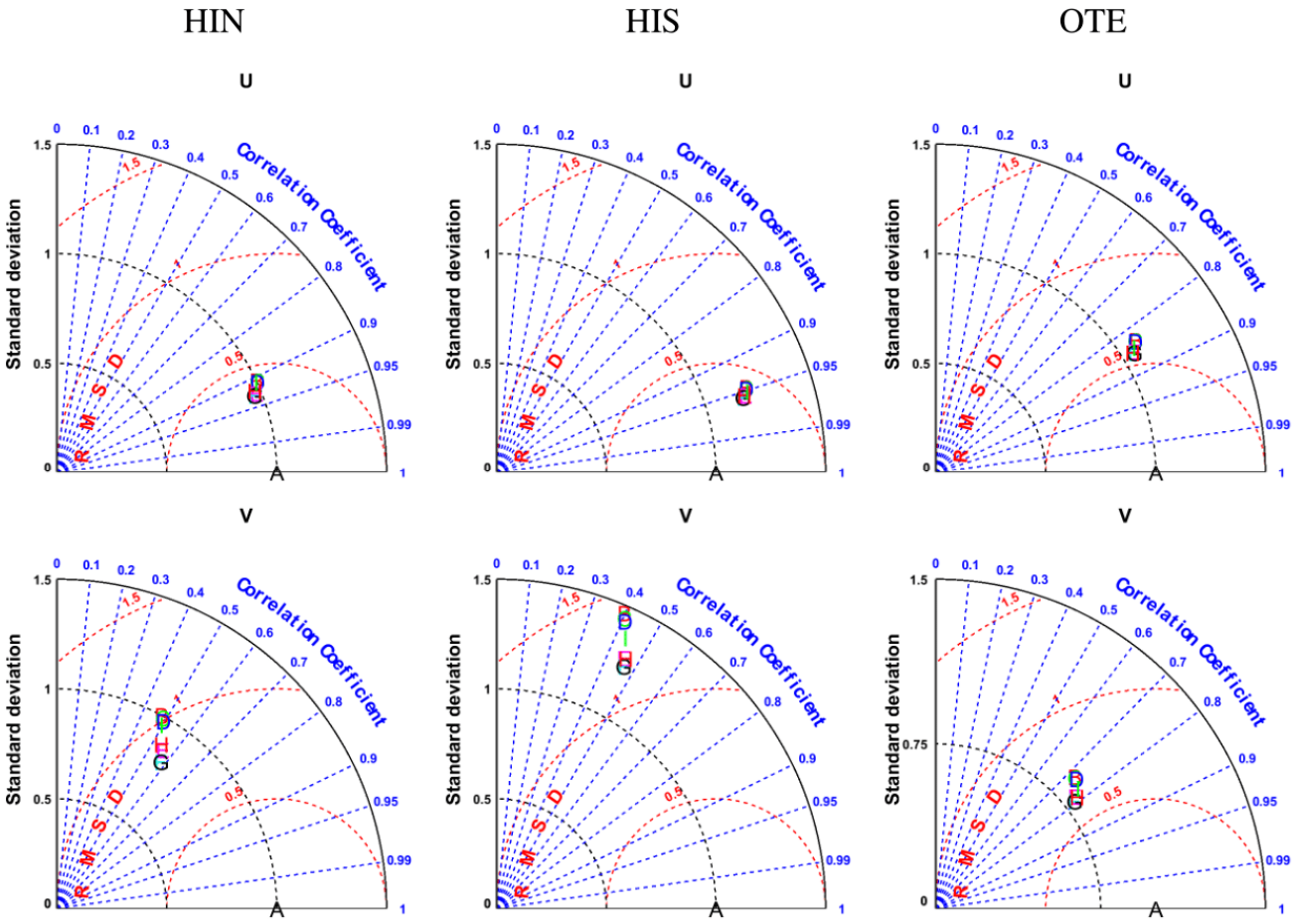

Fig. 8. Taylor diagrams showing the impact of the different QC steps for CBG *u*- (above) and *v*-(below) components. Notation is the same as Fig. 6. The normalisations used for *u* and *v* are respectively are 0.36 and 0.14 m.s^{-1} at HIN; 0.28 and 0.08 m.s^{-1} at HIS and 0.22 and 0.19 m.s^{-1}at OTE.

3) Spectral comparisons

Wyatt et al. [35] calculated spectra from HF radar current data using a time series created by filling gaps in the original data set with a tidal prediction, obtained using a least squares analysis, plus the mean of the residuals. A different approach has been adopted in this study because there is a strong wind influence in the radar data at the tidal frequencies and the radar and ADCP time series have long gaps. The full data set was split into segments with gaps no longer than 3 hours and with at least 4096 samples per segment for the 10 minutes sampling interval, and at least 768 samples, for 1-hour sampling interval. This guarantees that the analysed time periods are roughly the same (~28-32 days), giving similar spectral frequency resolutions and averaging. The short gaps (< 3 hours) were linearly interpolated. Data that did not conform to these requirements were not included in the analysis. Each segment was then processed using the Matlab *pwelch*

function (50% overlap, Hanning window) using FFT lengths of either 2048 for the 10 minutes sampling) or 256 (for the 1-hour sampling) and the resulting spectra were then averaged over all segments. The Matlab *pwelch* function is also used to calculate the confidence intervals for the spectral estimates. Note that 10 (60) minute sampling implies a Nyquist frequency of 72 (12) cycles per day.

We also used a smoothed Lomb-Scargle spectrum (Lomb [36]; Scargle [37]), which is available as *plomb* in Matlab. Smoothing of the Lomb-Scargle spectra can be achieved by applying a Bartlett window in a running average using *filter* in Matlab. This method can provide a spectrum for data with either missing samples or sampled unevenly, and therefore, avoids the need to check for continuity in the data set. The Lomb-Scargle spectrum gives similar results (not shown) to the *pwelch* method used here but the confidence intervals for the resulting spectrum are not as easy to determine.

Fig. 9 shows a *v*-component *pwelch* spectrum from COF CH100 and the *u*-component spectrum from CBG HIS, these being the larger amplitude component in each case. The other spectra (not shown) are qualitatively similar. Fig. 10 shows the same data sets but limited to the frequency range, 12 cycles per day, of the IMOS data (case G). All spectra clearly show semi-diurnal and diurnal signals; the semi-diurnal signal being particularly strong for the CBG data. At these locations these signals are both tidal and related to the wind forcing which also has significant energy at these frequencies. Note that at COF the diurnal and inertial frequencies are very similar. Fig. 10 also shows the IMOS spectra rescaled (multiplying by the 24 cycles per day sample rate divided by the 144 cycles per day sample rate of the other cases) in order to directly compare noise levels (see analysis below). We refer to Schmid [38] for a more detailed discussion about the interpretation of the Matlab *pwelch* spectrum in terms of the scaling required to get signal, noise and power levels.

The spectra at high frequencies (the noise floor) can provide an estimation of the rms fluctuations in the *u*- and *v*-components under the assumption that it is white noise. The square root of the noise power (variance) is obtained from the mean of the spectral density at the upper eighth of the frequency range multiplied by the maximum frequency (spectral bandwidth). This range was selected to avoid the ‘nulls’

associated with the smoothing applied (Appendix 2). These noise levels are plotted as dashed lines over the range 9 to 72 cycles per day in Fig. 9 and 10. A similar calculation has also been done limiting the frequency range to 10.5 – 12 cycles per day, i.e. the upper eight of the IMOS spectrum frequency range (as in Fig. 10). The results of both calculations are shown in Table 6. The 1 hour averaged noise levels are shown with a red dashed line in Fig. 9 and 10. At some locations (e.g. CBG) these two estimates are similar for the 10-minute sampled data (QCH10) and Fig 9 shows that the 10 minute spectrum is fairly flat from about a frequency of 24 cycles per day for all cases (including those not shown in the figure). This result suggests that there is little to be gained in using the data at this temporal resolution at these sites without additional averaging because the magnitude of any flow components at higher frequencies are below the noise level.

The ADCP and 30-minute radar data (QC30) estimates over the two frequency ranges are different as expected for running averages which are lowpass filters as is the Savitzky-Golay filter. In these cases, the two estimates are different as expected (Table 6) with the Savitzky-Golay filter giving similar results to those of QC30. The standard IMOS result (QC60) has lower rms fluctuations than the most of the other cases, a few of the high frequency QC30 cases are similar. For some of the cases the estimates are similar to the precision in the radial measurements discussed in Section IIIA. The estimates for the COF and CBG HIS ADCP data are similar. Those at CBG HIN and OTE are larger and the limited frequency range estimates are sometimes higher than the radar values.

TABLE 6

ESTIMATED RMS NOISE LEVEL (M/S) IN THE *U*- AND *V*-COMPONENTS AS DETERMINED FROM THEIR SPECTRA. UPPER ROW USE THE UPPER EIGHTH OF THE IMOS BANDWIDTH (12 CYCLES PER DAY) TO ESTIMATE THE RMS NOISE LEVEL, LOWER ROW USES THE UPPER EIGHTH OF THE OF THE HIGHEST FREQUENCY RANGE (72 CYCLES PER DAY).

Site	variable	ADCP	QCH10	QC30	QC60	SGL60	SGQ60
COF CH070	*u*	0.046	0.224	0.154	0.087	0.231	0.163
		0.005	0.112	0.046		0.026	0.015
	v	0.039	0.086	0.091	0.037	0.086	0.060
		0.004	0.045	0.020		0.010	0.006
COF CH100	*u*	0.048	0.139	0.125	0.064	0.1380	0.099
		0.005	0.070	0.030		0.015	0.010
	v	0.045	0.079	0.087	0.042	0.082	0.056
		0.004	0.044	0.020		0.0010	0.006
CBG HIS	*u*	0.033	0.049	0.04	0.018	0.048	0.033
		0.003	0.037	0.012		0.008	0.005
	v	0.034	0.060	0.053	0.016	0.058	0.040
		0.003	0.044	0.015		0.009	0.006
CBG HIN	*u*	0.111	0.095	0.084	0.041	0.096	0.069
		0.010	0.066	0.021		0.015	0.009
	v	0.079	0.085	0.074	0.035	0.085	0.061
		0.007	0.061	0.019		0.014	0.008
CBG OTE	*u*	0.048	0.068	0.063	0.028	0.067	0.045
		0.006	0.055	0.018		0.011	0.006
	v	0.040	0.080	0.070	0.029	0.077	0.050
		0.004	0.061	0.020		0.012	0.007

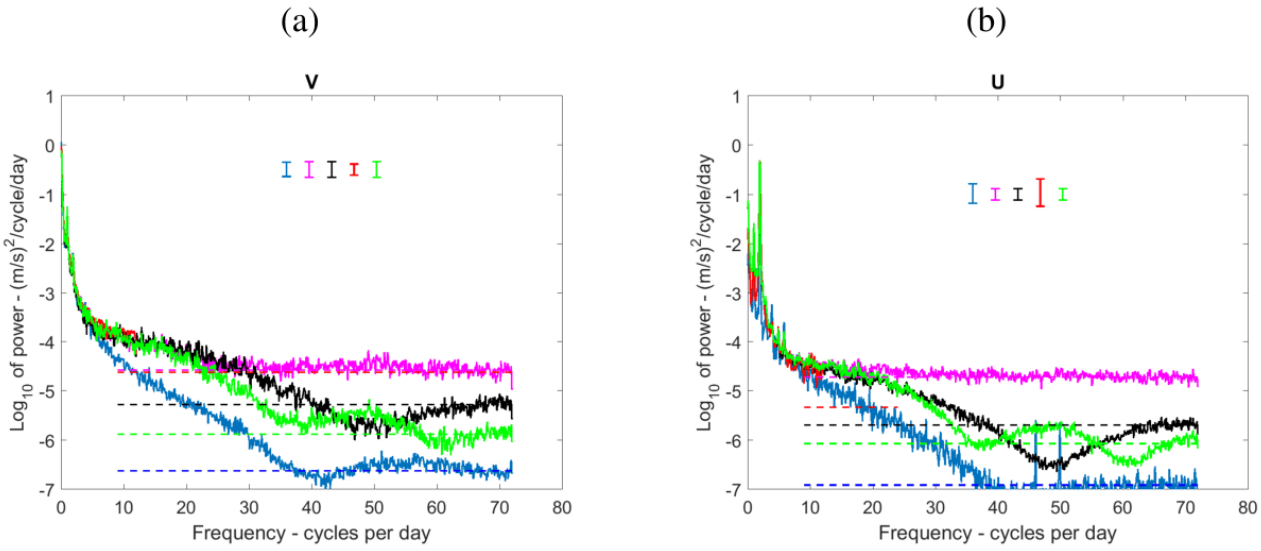

Fig. 9. Velocity component pwelch spectra for (a) the *v*-component at COF CH100 and (b) the *u*-component at CBG HIS. Blue –ADCP, magenta – 10 minute data hampel filter (QCH10); black – 30 minute QC (QC30); red – 1 hour QC IMOS data (QC60); green - Savitsky-Golay (SGQ60). 95% confidence intervals are shown using the same colour coding. Noise levels are shown with dashed lines using the same colours.

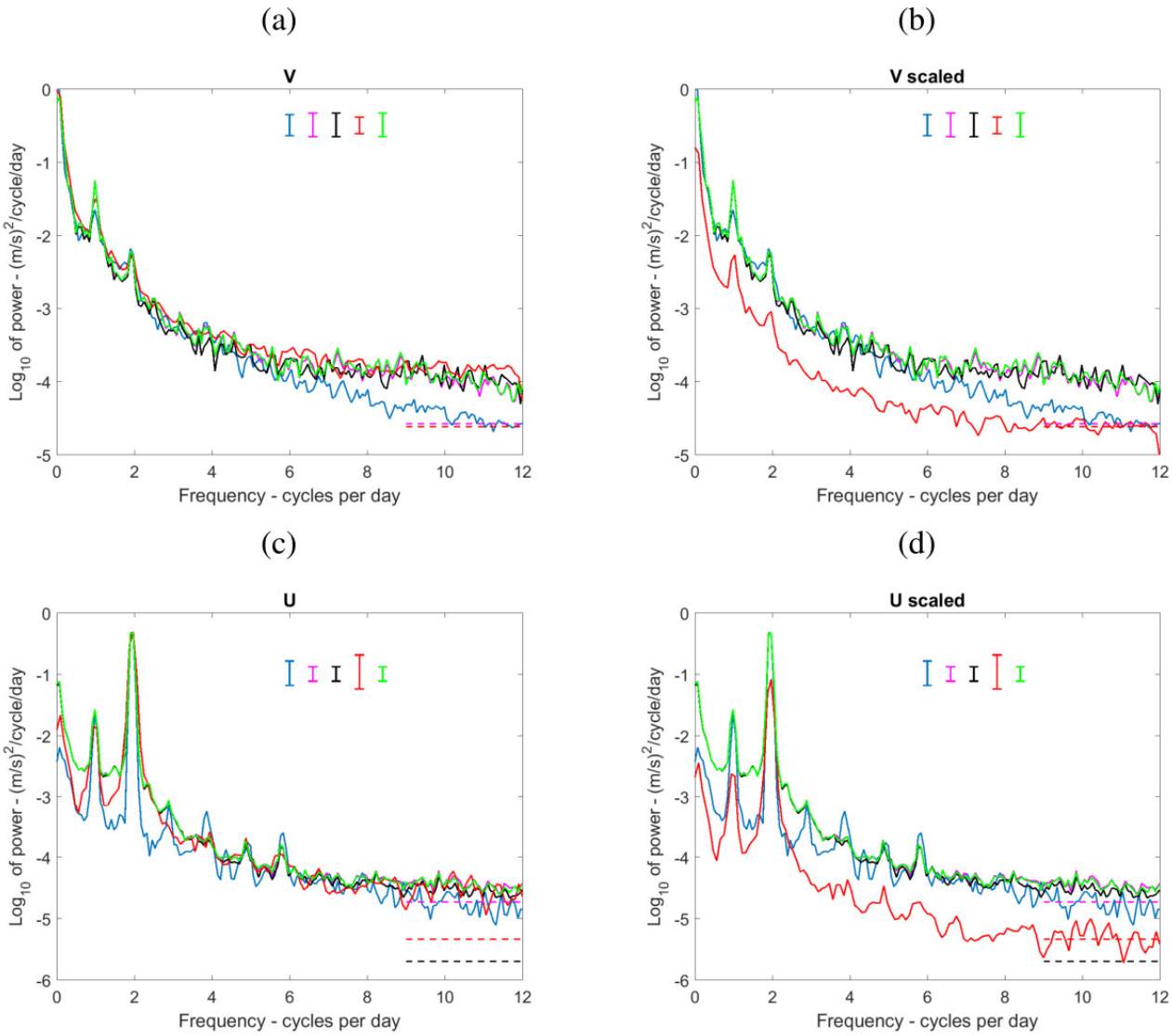

Fig. 10. Velocity component pwelch spectra for the *v*-component at COF CH100(a) not scaled and (b) scaled and for the *u*-component at CBG HIS (c) not scaled and (d) scaled. The IMOS spectra in (b) and (d) have been rescaled to enable direct comparison of the higher noise levels (dashed lines). Same colour coding as Fig. 9

4) Smoothing with the Savitzky-Golay Filter

In this section we discuss the application of the Savitzky-Golay filter to the IMOS QC data (QC60) and to the ADCP data over time windows of 1, 3, 7, 13, and 25 hours (note that the filter, as implemented in Matlab, requires an odd number of data points), as illustrated in Fig. 11 that compares time series of the *u*-components of the radar and ADCP currents. Increasing the filter order (i.e. from 1st order linear to 2nd order quadratic), retains more of the high frequency content in the data. Analyses using the 30-minute averages (QC30) as the starting point produced similar results (not shown). The same comparisons between

radar and ADCP have been made, namely: (*i*) the Taylor diagram statistics, (*ii*) the directional statistics and (*iii*) the spectral analysis.

The Taylor diagrams for the CH100 mooring at COF and the HIN mooring at CBG are shown in Fig. 12. They exhibit rather different behaviour. The COF results show a reduction in the normalised rms and increase in correlation coefficient as the smoothing period increases although the standard deviation of the radar measurements maintain a very similar ratio with that of the ADCP. Different normalisations are applied to each case (as described in Fig. 12) since the ADCP standard deviations also decrease with smoothing (albeit very slowly in the COF case). In the CBG case the longer period smoothing cases start to deviate from the ADCP measurements possibly due to the dominance of the semi-diurnal signal at this location which is being filtered out in these cases. Fig. 13 compares the 25 hour filtered time series at the HIN and OTE moorings and shows that the HIN radar *u*-component is still rather noisy in spite of the filtering. Since the amplitudes of the residual currents are small in both cases, the noisy radar data are dominating the statistics of the comparison at HIN. This noise in surface currents is likely to be due to sub-cell scale current variability at HIN which is in shallower water and closer to the reefs than OTE (Fig. 1a).

We applied three different estimates of the uncertainty in the radar measurements to all the analysis presented in this paper: (1) the measure obtained from the spectrum (the rms noise level assuming white noise), (2) the root-mean-square difference between radar and ADCP and (3) the standard deviation in the radar *u*, *v* components derived from the information supplied in the data files using the methods described in Section IIIB for the averages and Appendix 2.3 for the filter. Estimates 1 and 3 (exclusively based on the radar) are shown in Fig. 14a for these five longer period filtered cases. At COF the values relate to the *v*-component and at CBG to the *u*-component since these are the larger in both cases. The errors decrease as the length of smoothing increases, so that the five different cases are easy to identify in Fig. 14. For a given period of smoothing the linear filter (order 1 in the legend) seems to reduce both estimates more than the quadratic filter (order 2) probably because of the differences in high frequency content referred to above. The standard deviations estimated from the data (item 3 above and shown in the Fig 14a) are generally

lower at COF and higher at CBG than those obtained from the spectra. The results suggest that the COF estimates provided by the processing from radar data to radial may be a little optimistic although the estimates are clearly correlated at both sites. The time series in Fig. 13 confirms the low spectral estimates at OTE relative to those at HIN.

Estimates 1 and 2 are compared in Fig. 14b. The pattern at HIS and HIN is very different from elsewhere with an increase in the root-mean-square difference for longer smoothing periods (as observed in the Taylor diagrams and attributed to higher noise levels). The root-mean-square differences between radar and ADCP are higher than the spectral estimates indicating that there are real differences between the measurements. This is discussed further in Section V.

5) Drifter comparison

Taylor diagrams for the comparisons between the 1 h averaged drifter and radar currents extracted along the tracks of the eight drifting buoys are shown in Fig. 15, both for *u*- and *v*-components. The comparison statistics use data normalised by the standard deviation of each drifter (as specified in Fig. 15). Drifters B-E (released on the shelf; Fig. 15 and 16) showed the best agreement for both *u*- and *v*-components, the comparison being particularly good for the east-west currents (*u*-component) with the larger range. The results were similar for three of the shelf drifters (Fig. 15, drifters C, D and E). Drifter B remained close to the launch site and differences with the radar *u*-component are a little higher.

The slope drifters, released in deeper water, showed a higher variability for the *v*-component measurements than is present in the radar data (normalised radar standard deviations < 1). The slope drifters travelled along the CBG reef matrix and developed several loops around the islands, being more influenced by small scale circulations observed near the reefs (Fig. 16). In addition, they travelled more distant from the radar sites and some of this variability maybe explained by the smaller radar spatial resolution at long ranges (as shown with the black boxes in Fig. 16). The differences in the radar spatial resolutions between the shelf (smaller sampling cells) and slope (larger sampling cells) measurements may have contributed to the differences observed in the comparisons. Scatter plots of the comparison of drifter

D and for all the cases together are shown in Fig. 17. The complex correlation coefficients found for these cases were 0.97 to 0.98 for the shelf drifters and 0.93 to 0.94 for the slope drifters (not shown).

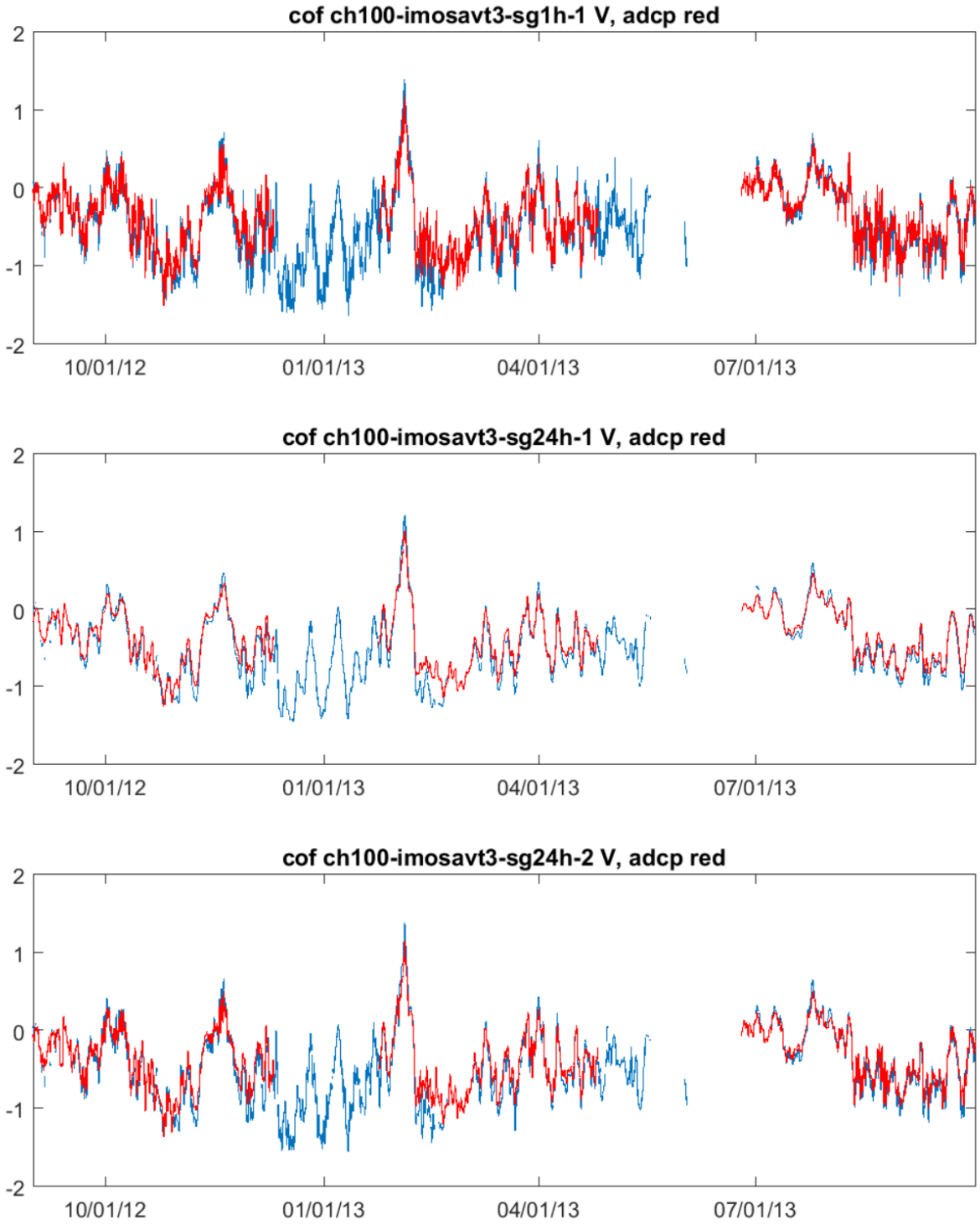

Fig. 11. Time series of the *u*-component of current at the CH100 mooring at COF for the IMOS 1-hour averaged data (upper panel); after application of the 25 hour Savitzky-Golay filter with filter order 1 (middle) and 2 (bottom). Radar data are in blue and ADCP (at 14 m depth) in red.

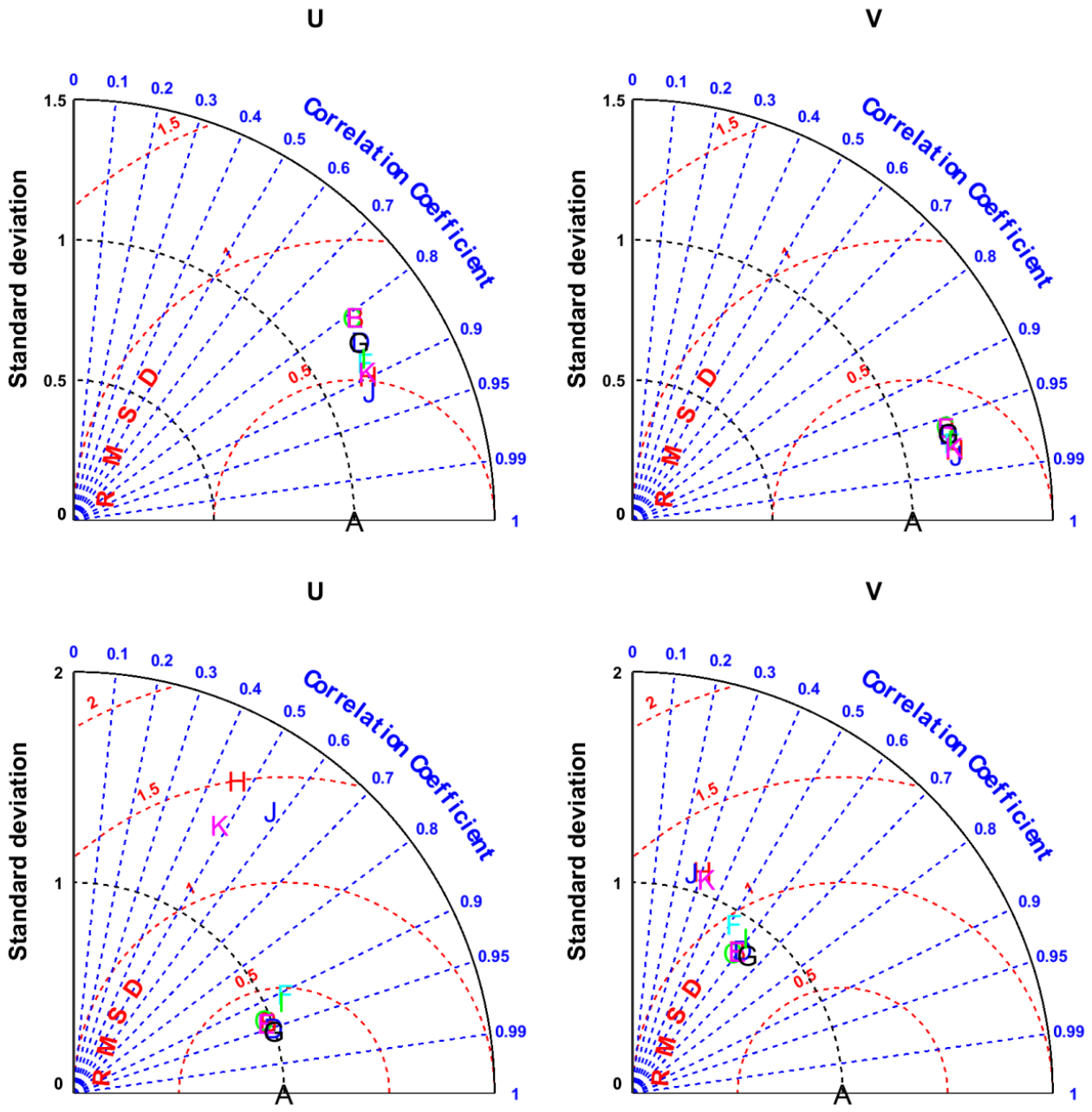

Fig. 12. Taylor diagrams showing the comparison between radar and ADCP data for different smoothing periods and filter orders denoted by the letters where B (1 h), D (3 h), F (7 h), H (13 h), J (25 h) are order 1 and C (1 h), E (3 h), G (7 h), I (13 h), K (25 h) are order 2. Upper row: CH100 at COF, lower row: HIN at CBG. Notation is otherwise the same as Fig. 6. The normalisations in m/s used at CH100 for u are 0.18(B), 0.18(C), 0.17(D), 0.18(E), 0.17(F), 0.18(G), 0.16(H), 0.17(I), 0.15(I), 0.16(J) and for *v* are 0.39(B), 0.39(C), 0.39(D), 0.39(E), 0.38(F), 0.39(G), 0.37(H), 0.39(I), 0.37(I), 0.38(J) at HIN for u are 0.35(B), 0.36(C), 0.33(D), 0.36(E), 0.21(F), 0.34(G), 0.09(H), 0.26(I), 0.07(I), 0.11(J) and for *v* are 0.14(B), 0.14(C), 0.13(D), 0.14(E), 0.10(F), 0.13(G), 0.07(H), 0.11(I), 0.07(I), 0.08(J)

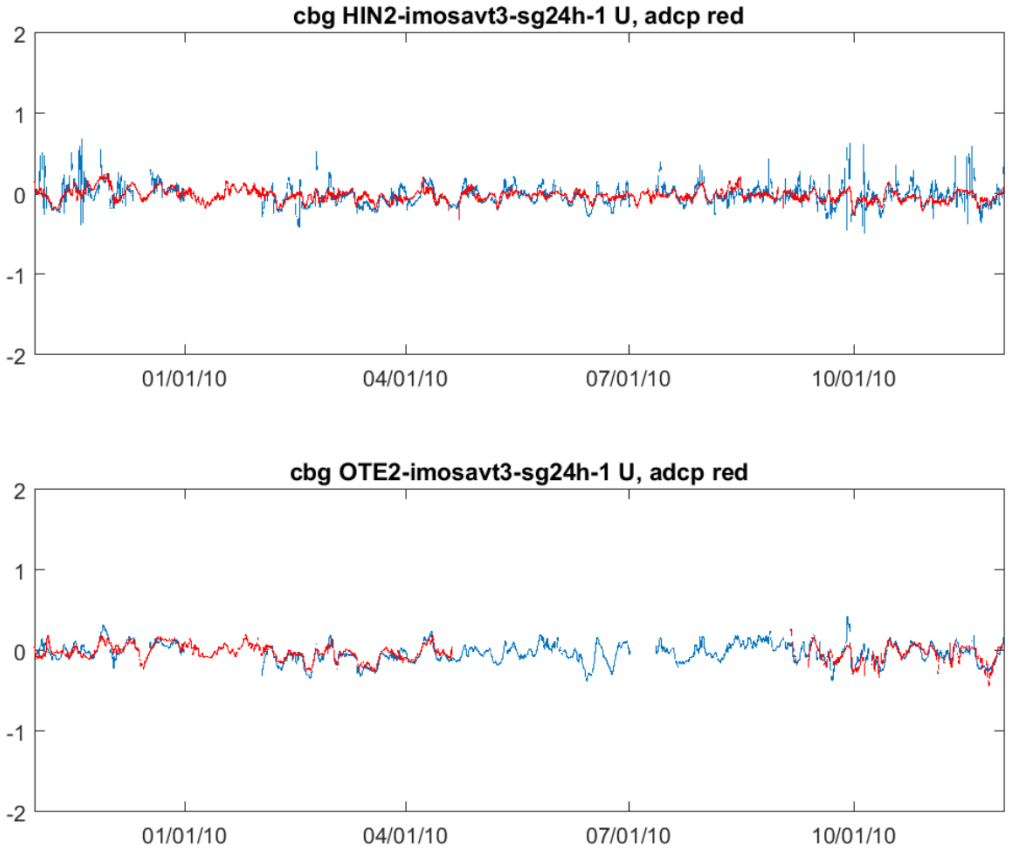

Fig. 13. Time series of the *u*- components of current at the HIN (above) and OTE2 (below) moorings at CBG after application of the 25-hour Savitzky-Golay order 1 filter. Radar data are in blue and ADCP in red

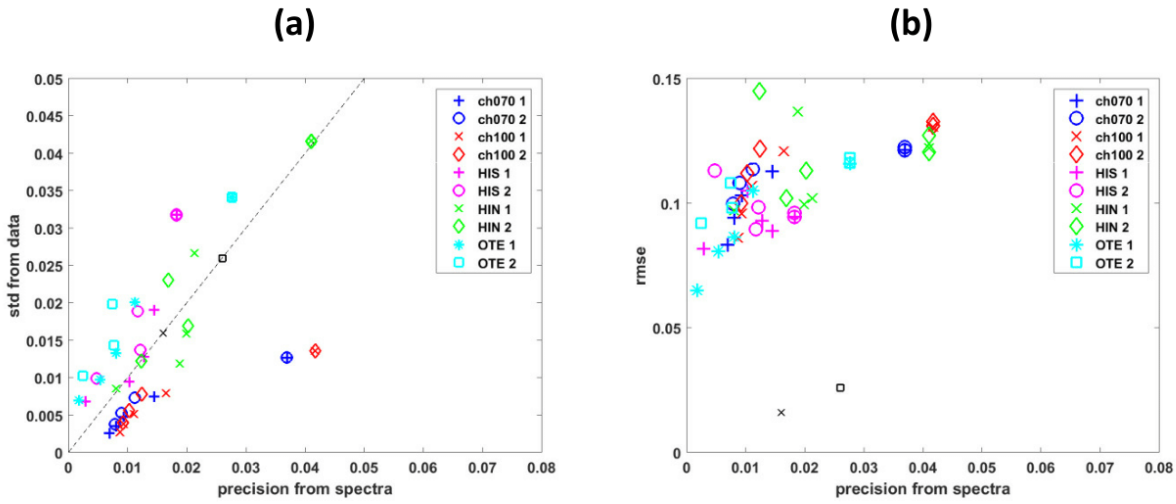

Fig. 14. Comparisons of different estimates of uncertainty in the radar measurements for the different smoothing periods (1, 3, 7, 13, 25 hours). The numbers following the mooring location in the legend refer to the order of the Savitzy-Golay filter. Spectral estimates are compared with (a) the radar standard deviation and (b) the rms difference between radar and ADCP. The smaller black symbols o and x denote the sampling variability estimate of uncertainty for 1-hour averaged radial data, discussed in section IIIA, at CBG and COF respectively. In some cases the 3 hour smoothing overlaid the 1 h smoothing point and thus there appear to be only 4 symbols for those cases on the figure. The 1:1 line (dashed) is shown in (a) for reference.

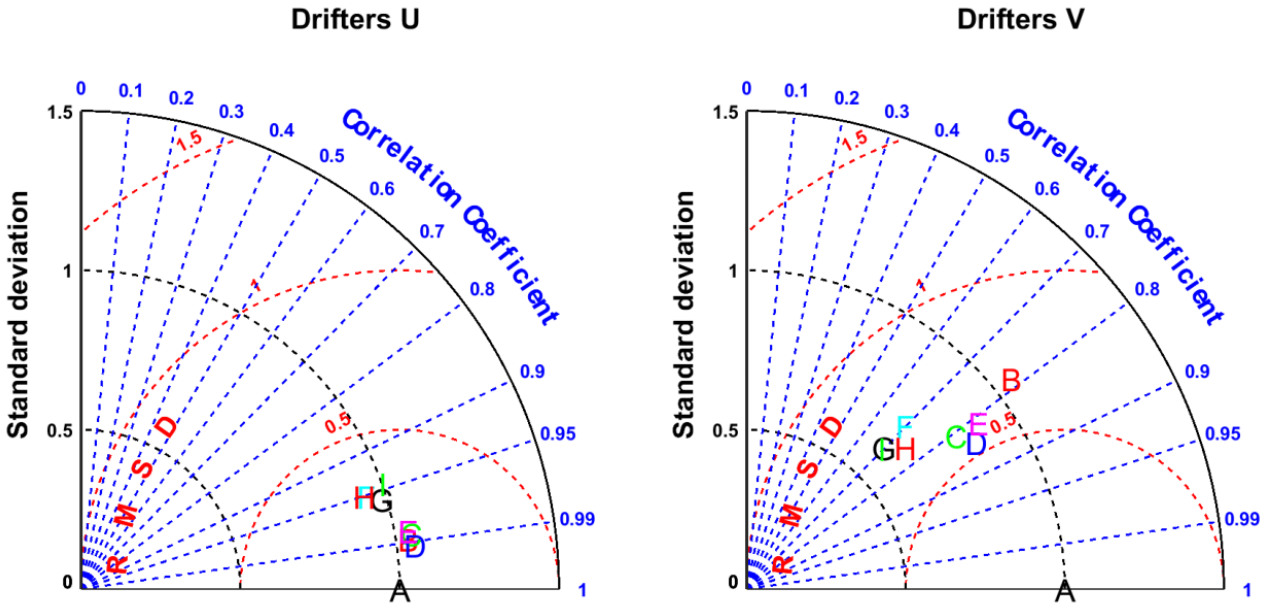

Fig. 15. Taylor diagrams showing the comparison between drifter and radar data. Notation is mainly the same as Fig. 7 except the letters now refer to different drifters B-E released on the shelf and F-I on the slope. The normalisations used are *u*: 3.4, 3.0, 3.0, 3.0, 3.6, 3.3, 4.3, 3.7 *v*: 8.0, 6.9, 6.8, 7.5, 14.9, 14.9, 14.5, 12.8 cm·s^{-1} respectively.

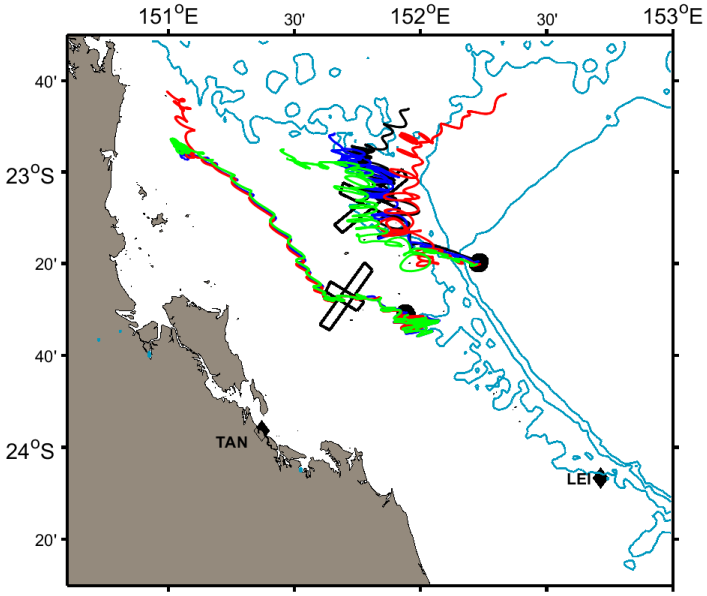

Fig. 16. Map showing the drifter tracks, 4 from each release point(●); radar sites (♦). Black boxes mark nominal radar resolution from each radar at grid cells centred in the middle of the overlap.

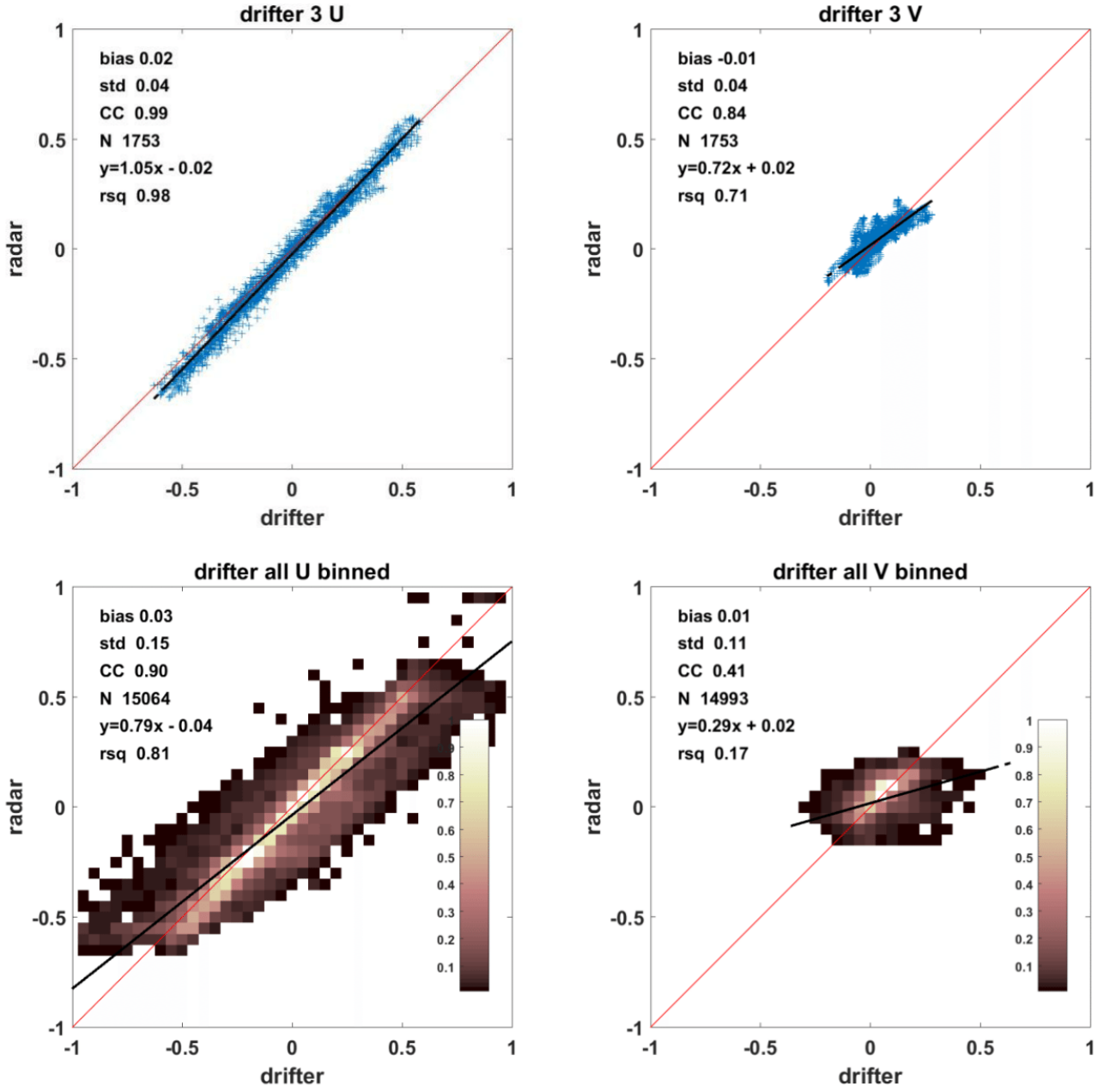

Fig. 17. Scatter plots of the comparison between drifter and radar data for case D (on the Taylor diagram above) and (as binned plots, notation as in Fig. 6) for all the data combined (below).

V. Discussion

The accuracy of HF radar current measurements at two coastal regions (Eastern Australia) with complex circulations was evaluated for the first time in this study using the Taylor diagram, the directional statistics and spectral analysis. Tidal and wind forcing dominate in the southern region of the GBR lagoon, while the

influence of the East Australian current near the coast adds complexity to the coastal circulation in the Coffs Harbour region.

Good agreement between the currents measured by different platforms was demonstrated through the Taylor diagram and additional statistics. The Taylor diagram measures performance through a combination of correlation coefficient (as high as possible), relative standard deviation (i.e variance in both measurements should be as close as possible) and rms difference (as small as possible) rather than using just one statistic. Our contention is that this gives a more complete picture of the relative merits of different data analysis methods. Nonetheless, correlation coefficients, or equivalently r^2 values, are often all that is provided in the literature so for comparative purposes we summarise these here.

Linear regressions between the HF radar and the ADCP currents had $r^2 \sim 0.73$-0.92 (Figs. 6 and 8) for the stronger current component (*v*-component in COF and *u*-component in CBG). Even higher correlations were found for the comparisons between HF radar and drifter *u*-components ($r^2 \sim 0.81$-0.98) in the Great Barrier Reef. The overall better agreement found between HF radar and drifter speeds is almost certainly due to the coincident depth at which both of these measurements were taken despite their differences in the sampling area, although discrepancies increased when drifters were travelling in areas where small scale secondary circulations (not sampled at the radar resolution) became relevant.

A lower correlation and mostly higher variance in the radar measurements was found for the weaker current component at both sites. The statistics for this component are likely to be more influenced by noise. The complex correlations between the radar and the ADCP vector current measurements lie between 0.8 and 0.96 (Tables 4 and 5) and those for the drifter comparison between 0.93 and 0.98 (Section IV-B-5). These figures provide further very good evidence of overall agreement between the different measurement systems.

The overall instrumental rms noise (based on the highest eight frequency range) had median values of 0.005 (0.02) $m.s^{-1}$ for the ADCP (radar QC30) measurements (Table 6), and the median rms difference between these two measurements is about 0.12 $m.s^{-1}$ which indicates that there are real differences between

them. Biases and variances among the different platforms occur for a number of reasons, such as instrumental errors and differences in their spatial and temporal sampling resolutions. Instrumental errors can be partially responsible for the observed RMS differences between the HF radar and ADCP measurements. For instance, comparisons between daily averaged data recorded by an ADCP with speed measurements from two mechanical current meters presented RMS differences of the order of 0.03-0.07 $m.s^{-1}$ (Plimpton et al. [39]). RMS higher differences (up to 0.5 $m.s^{-1}$) were found for instantaneous values due to reflection of acoustic energy by fish swimming near the moorings (Plimpton et al. [40]). On the other hand, errors in HF radar measurements vary over the domain with the cell area used for data averaging and also increase when the angle of intersection between two radials becomes acute (geometric degradation of precision; Section IIIC).

Although instrumental errors can be significant, differences in the sampling rate and sampled area/depth are likely to cause the biggest discrepancies among the compared platforms. These differences are especially important in coastal areas that show high temporal and spatial current variability, such as the regions under discussion here. The three compared techniques used different time resolutions: the raw HF radar outputs data every 10 minutes, ADCP data were averaged over a burst of a few minutes for sampling rates of 5 to 30 minutes (Table A1; Appendix 1); and drifter speeds were derived from two instantaneous measurements taken 10 to 20 minutes apart.

The total area sampled by each platform can also influence the measured values of speed and direction of surface currents, since surface flow varies in space due to horizontal inhomogeneities (e.g. eddies, jets, divergence, convergence and frontal areas). The HF radar system provides surface currents effectively averaged over cells of a few km^2 but varying in size within the radar domain whereas ADCP and drifters provide point-like information integrated over the sampled depth layer. Mismatching of the sampled area is likely to be the main factor leading to differences between HF radar and drifter along-shelf measurements near reef islands, since the drifters were caught in small-scale secondary circulations generated around the islands (Mantovanelli et al. [41]). The existence and importance of topographically-induced secondary

circulations, small scale eddies and upwelling around reef islands in the GBR has been extensively documented (e.g. White and Wolanski [42]) while submesoscale features (3-10 km) are relevant in the Coffs Harbour region (Mantovanelli et al. [43], Schaeffer et al [44]).

Apart from discrepancies associated with the radar spatial resolution, shallow-water effects can also bias the HF current measurements. Errors in the HF radar currents become significant in water shallower than about 5 m for a ~8 MHz radar transmit frequency (Lipa et al. [44]), as used in the GBR radar system and hence may become significant near the reefs if the radar cell includes shallow/exposed areas.

Longer period smoothing of both radar and ADCP data can be achieved through the use of the Savitzky-Golay filter which does not require uniformly sampled data. Compared with a standard averaging procedure (i.e. simple average over a time interval), this filter does not remove as much noise, hence the statistics of the comparison between radar and ADCP presented in section IV-B-2 do not benefit as much from the application of the filter as is the case for the 30 minute radial averages. This filter has the advantage of being relatively simple to apply in Matlab – albeit by slight modifications to the Matlab code in order to estimate the corresponding variances - so may be useful for longer time period smoothing especially since it has a wider passband than an equivalent moving average filter.

VI. Summary

We have shown the Taylor diagram to be a useful tool to assess instrument performance. The analyses presented here, both for the *u*- and *v*- components separately and for the full vector using the Kundu [9] complex correlation method, show that current data from the IMOS WERA HF radar systems does benefit from increased averaging and additional quality control measures, in particular averaging over more than 10 minutes. The spectral analysis suggests that the noise in the 10-minute data sets does not allow the resolution of surface current variability at this fine timescale at these sites and operating frequencies. 30-minute radial averages may be sufficient for many applications; differences in the ADCP comparison statistics between these and the 1-hour averages are quite small.

The procedures involved in generating the IMOS FV01 product could be used for the real-time data stream although they are not currently implemented. Although the FV00 (not-quality controlled) and FV01 (quality-controlled) products have not been compared, a measure of the improved performance of FV01 was provided in Wyatt [46].

The Taylor diagram assesses the performance of the radar assuming that the ADCP represents sea-truth. However differences between the two platforms are expected since the ADCP is a measurement several metres below the surface, in what could be a sheared flow, whereas the radar measurement is at the surface (averaged over a depth of about 1 m) and thus is likely to include a stronger wind driven influence. The comparisons with the drifter show a higher level of agreement with correlations of up to 0.98 and biases that are within the radar current resolution for the drifters that were released and travelled on the shelf under more spatially uniform circulation. The drifters released on the slope travelled along the reef matrix, being captured in small eddies formed on the lee of the island at spatial scales that could not be captured at long range because of the radar azimuthal resolution.

The comparison between the HF radar surface currents and those measured by ADCPs and drifters are consistent with or better than those reported elsewhere. It can be concluded that the radar currents at these two locations, and in particular the 1-hour averaged case with added QC, can be used with confidence for scientific and operational oceanography.

APPENDIX 1 ADCP DATA TABLE

TABLE A1

SUMMARY OF THE ADCP (RDI WORKHORSE SENTINEL) CONFIGURATION FOR EACH DEPLOYMENT, INCLUDING THE RECORD DURATION (FIRST AND LAST GOOD MEASUREMENTS), ADCP FREQUENCY, CONFIGURATION (NUMBER OF BINS, BIN SIZE, SAMPLING RATE AND NUMBER OF PINGS PER ENSEMBLE), SITE DEPTH AND DISTANCE OF THE INSTRUMENT FROM THE BOTTOM (OFF-BED), SITE LOCATION (LATITUDE AND LONGITUDE) AND MEAN MAGNETIC DECLINATIONS (MD). DEPLOYMENTS WITH FAILURE IN THE PRESSURE SENSOR WERE EXCLUDED. TIMES ARE IN UTC.

site	deployment duration (day/month/year)	frequency (kHZ)	Bins (size) (m)	Sample (min)	pings/ ens.	depth/ off-bed (m)	latitude/longitude	MD (°)
HIN (1)	16/10/2009-19/04/2010	614.4	50 (1)	30	170	45/ 4	23.39° S/ 151.99° E	9.71
HIN (2)	22/04/2010-30/08/2010	614.4	50 (1)	30	170	46/ 4	23.38° S/ 151.99° E	9.70
HIN (3)	02/09/2010-10/03/2011	614.4	50 (1)	10	60	44.5/ 4	23.38° S/ 151.99° E	9.70
HIS (1)	26/04/2010-31/08/2010	614.4	54 (1)	30	170	45/ 0.5	23.51° S/ 151.95° E	9.72
HIS (2)	03/09/2010-08/03/2011	614.4	28 (2)	10	60	45/ 0.5	23.51° S/ 151.95° E	9.72
OTE (1)	12/10/2009-19/04/2010	614.4	54 (1)	30	160	59/ 5	23.48° S/ 152.17° E	9.78
OTE (2)	05/09/2010-10/03/2011	307.2	16 (4)	10	30	56/ 3	23.48° S/ 152.17° E	9.76
CH70 (1)	09/07/2012-14/09/2012	307.2	18 (4)	5	50	70/ 6	30.28° S/ 153.30° E	11.89
CH70 (2)	11/09/2012-23/10/2012	307.2	18 (4)	5	50	70/ 6	30.28° S/ 153.30° E	11.89
CH70 (3)	29/10/2012-11/12/2012	307.2	18 (4)	5	50	70/ 6	30.28° S/ 153.30° E	11.89
CH70 (4)	10/12/2012-24/01/2013	307.2	18 (4)	5	50	70/ 6	30.28° S/ 153.30° E	11.89
CH70 (5)	23/01/2013-28/03/2013	307.2	18 (4)	5	50	70/ 6	30.28° S/ 153.30° E	11.89
CH70 (6)	24/06/2013-11/07/2013	307.2	18 (4)	5	50	70/ 6	30.28° S/ 153.30° E	11.89
CH100 (1)	09/07/2012-14/09/2012	307.2	26 (4)	5	50	100/ 6	30.27° S/ 153.39° E	11.92
CH100 (2)	11/09/2012-31/10/2012	307.2	26 (4)	5	50	100/ 6	30.27° S/ 153.39° E	11.92
CH100 (3)	29/10/2012-11/12/2012	307.2	26 (4)	5	50	100/ 6	30.27° S/ 153.39° E	11.92
CH100 (4)	23/01/2013-28/03/2013	307.2	26 (4)	5	50	100/ 6	30.27° S/ 153.39° E	11.92
CH100 (5)	25/03/2013-26/04/2013	307.2	26 (4)	5	50	100/ 6	30.27° S/ 153.39° E	11.92
CH100 (6)	25/03/2013-25/04/2013	307.2	26 (4)	5	50	100/ 6	30.27° S/ 153.39° E	11.92
CH100 (7)	24/06/2013-11/07/2013	307.2	26 (4)	5	50	100/ 6	30.27° S/ 153.39° E	11.92
CH100 (8)	08/07/2013-25/10/2013	307.2	26 (4)	5	50	100/ 6	30.27° S/ 153.39° E	11.91

APPENDIX 2 THE SAVITZKY-GOLAY FILTER

A. Introduction

Savitzky-Golay filters (Orfanidis [30]) can be used to smooth data sets with the aim of preserving the high-frequency content as much as possible at the expense of not removing as much noise as, say, an averaging filter. The filter works by least-squares fitting a polynomial to a selected number of samples and replacing the centre sample by the polynomial fit. This filter is implemented in Matlab either as a digital filter (*sgolayfilt*) or as an option in the general smoothing function (*smooth*). In the latter case the smoothing can be applied to data with gaps in it making it suitable for application to HF radar data.

The impact of this filter is illustrated in Fig A2.1 which shows the un-filtered spectrum of a sinusoidal signal (with period 1 hour) embedded in white noise and the spectra after a 7-point quadratic

Savitzky-Golay (red) and a 7 point moving average filter (magenta dashed) have been applied. Note that the Savitzky-Golay linear filter is the same as a moving average filter for the same number of points (Schafer, 2011) and hence the figure is also comparing the linear and quadratic Savitzky-Golay filters. It has a low frequency performance similar to the 7-point Savitzky-Golay filter. The shape of the filters at high frequencies is clear. It can be seen that the noise level (which we estimate as an average over a range of the highest frequencies assuming the spectrum is beginning to flatten) is higher for the quadratic filter but, more importantly for many applications, the range of frequencies that are minimally affected by the filter (the passband) is much broader. Higher order filters extend this range further, increasing the number of points used has the opposite effect (Orfanidis [30], Schafer, 2011). There is reduction in signal level for the sinusoidal component relative to low frequencies but nothing like the reduction in the linear filter case. The 3-point moving average filtered data (black dashed) is also shown in the figure since this is referred to in the main body of the paper.

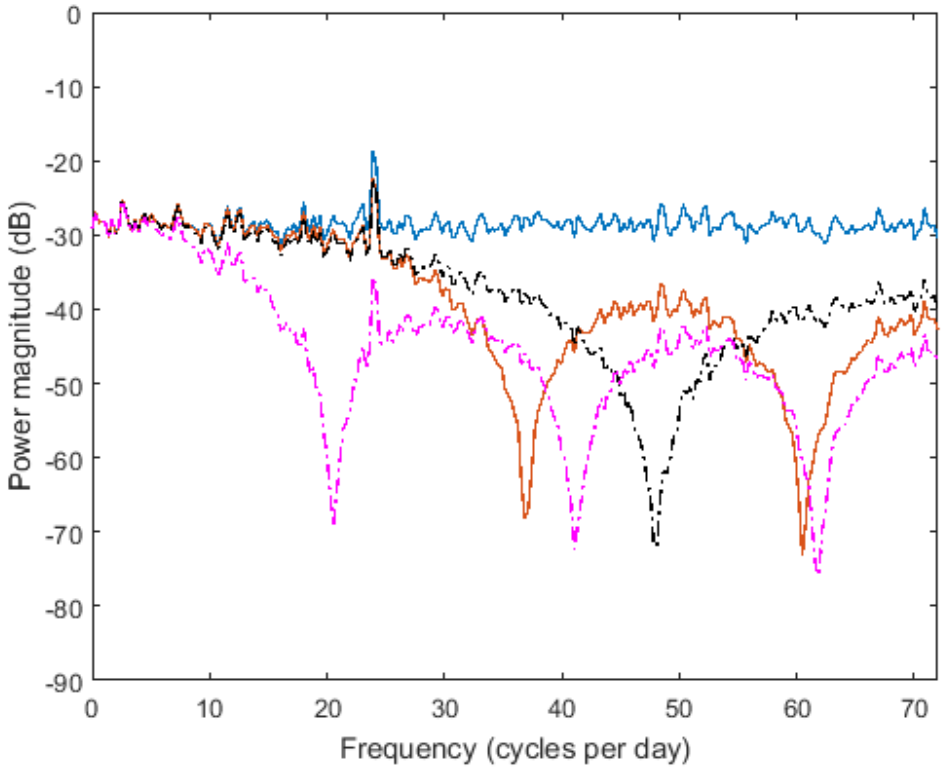

Fig. A2.1. Spectra of (blue) sinusoidal figure embedded in white noise, (brown) with application of 7-point Savitzky-Golay quadratic and (magenta dashed) linear (also 7-point moving average); (black dashed) is with the 3 point moving average.

The implementation in Matlab *smooth* is a standard least squares fit to a data set given as $(x_i, y_i), i = 1, \cdots, N$ pairs. The number of points, *2n+1*, from that set that are used for the least-squares fitting is spcified by the user. For each least-squares fit the $x_i, i = j-n, \cdots, j+n$ values are centred on the centre point of the subset, *j*, i.e. the x_i values become $x_i - x_j, i = j-n, \cdots, j+n$ with the result that the value of the polynomial at x_j is equal to the constant term.

B. Application to HF radar currents

In the application to HF radar currents, we need not just the smoothed value of the current (the constant term in the polynomial) but also the corresponding value of the variance of the smoothed current. Each current measurement has a variance, σ_i^2, associated with it (as discussed in this paper) and the impact of the variation in variance on the smoothing and of the smoothing on this variance needs to be determined. This can be done using weighted least-squares fitting. To achieve this in Matlab, the sgolay filter in *smooth* has to be modified to replace the standard least squares method with the function *lscov*. This allows both weighted least squares fitting and calculates the variances in the coefficients from the weights, in this case the variances of the un-smoothed current data which have to be added as arguments to *smooth*. The filter can introduce spurious results where there are NaNs in the data set. This is dealt with by replacing the original NaNs after the filter.

C. Propagation of errors with the Savitzky-Golay Filter

The calculation of the coefficients and their variances of a least-squares linear fit is well known, that for higher order polynomials perhaps less so. The linear (weighted and not, with variance) and quadratic (without weights or variances) cases can be found in e.g. Barlow [48]. The analysis for the weighted least squares linear and quadratic fitting are presented here for convenience. The values of the constant term in the linear and quadratic polynomials and its variance derived here have been verified to be consistent with the results obtained using Matlab.

1) Linear fit

The weighted least squares fit seeks to find a slope, *m*, and intercept, *c*, to minimise

$$E = \sum \frac{(y_i - mx_i - c)^2}{\sigma_i^2}$$

The result is obtained by differentiating *E* with respect to *m* and *c* and setting the results to *0*. The following weighted means are used in the analysis. The sums are over the *2n+1* points used in the fitting.

$$\bar{y} = \frac{\sum \frac{y_i}{\sigma_i^2}}{\frac{1}{\sigma_i^2}},\ \bar{x} = \frac{\sum \frac{x_i}{\sigma_i^2}}{\frac{1}{\sigma_i^2}},\ \overline{xy} = \frac{\sum \frac{x_i y_i}{\sigma_i^2}}{\frac{1}{\sigma_i^2}},\ \overline{x^2} = \frac{\sum \frac{x_i^2}{\sigma_i^2}}{\frac{1}{\sigma_i^2}}$$

$$\begin{aligned}\frac{\partial E}{\partial m} &= -2\sum \frac{x_i(y_i - mx_i - c)}{\sigma_i^2} \\ &= -2\left(\sum \frac{x_i y_i}{\sigma_i^2} - m\sum \frac{x_i^2}{\sigma_i^2} - c\sum \frac{x_i}{\sigma_i^2}\right) \\ &= -2\sum \frac{1}{\sigma_i^2}\left(\overline{xy} - m\overline{x^2} - c\bar{x}\right) = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial E}{\partial c} &= -2\sum \frac{(y_i - mx_i - c)}{\sigma_i^2} \\ &= -2\left(\sum \frac{y_i}{\sigma_i^2} - m\sum \frac{x_i}{\sigma_i^2} - \frac{c}{\sigma_i^2}\right) \\ &= -2\sum \frac{1}{\sigma_i^2}(\bar{y} - m\bar{x} - c) = 0\end{aligned}$$

For this analysis we are only interested in *c* which is obtained by eliminating *m* from

$$\begin{aligned}\overline{xy} - m\overline{x^2} - c\bar{x} &= 0 \\ \bar{y} - m\bar{x} - c &= 0\end{aligned}$$

by multiplying the first equation by $\bar{x}$ and the second by $\overline{x^2}$ and subtracting. This gives

$$c = \frac{\overline{x^2}\,\bar{y} - \bar{x}\,\overline{xy}}{\overline{x^2} - (\bar{x})^2}$$

Note that *c* has the same form for non-weighted least squares where the weighted means are replaced by standard means.

The variance in *c* is determined from the variances, σ_i^2 in the y_i values used as follows.

$$Var(c) = \sum \left(\frac{\partial c}{\partial y_i} \right)^2 \sigma_i^2$$

$$= \sum \left(\frac{\partial}{\partial y_i} \left(\frac{\overline{x^2} \dfrac{\sum \frac{y_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}} - \bar{x} \dfrac{\sum \frac{x_i y_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}}}{\overline{x^2} - (\bar{x})^2} \right) \right)^2 \sigma_i^2$$

$$= \sum \left(\frac{\overline{x^2} \dfrac{\frac{1}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}} - \bar{x} \dfrac{\frac{x_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}}}{\overline{x^2} - (\bar{x})^2} \right)^2 \sigma_i^2$$

$$= \frac{1}{\left(\sum \frac{1}{\sigma_i^2} \right)^2} \sum \frac{1}{\sigma_i^2} \frac{\overline{x^2}^2 + \bar{x}^2 x_i^2 - 2\overline{x^2}\, \bar{x}\, x_i}{\left(\overline{x^2} - (\bar{x})^2 \right)^2}$$

$$= \frac{1}{\sum \frac{1}{\sigma_i^2}} \frac{\overline{x^2}^2 + \bar{x}^2 \overline{x^2} - 2\overline{x^2}\, \bar{x}^2}{\left(\overline{x^2} - (\bar{x})^2 \right)^2} = \frac{1}{\sum \frac{1}{\sigma_i^2}} \frac{\overline{x^2}}{\overline{x^2} - (\bar{x})^2}$$

2) Quadratic fit

The quadratic analysis is similar. The weighted least squares fit seeks find the coefficients, a_0, a_1, a_2 to minimise

$$E = \sum \frac{\left(y_i - a_2 {x_i}^2 - a_1 x_i - a_0 \right)^2}{\sigma_i^2}$$

Some additional weighted means are required:

$$\overline{x^3} = \frac{\sum \frac{x_i^3}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}}, \quad \overline{x^4} = \frac{\sum \frac{x_i^4}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}}$$

It is then easy to derive the following from the derivatives of *E* with respect to the three variables a_0, a_1, a_2.

$$\begin{aligned}\overline{x^2 y} - a_2\overline{x^4} - a_1\overline{x^3} - a_0\overline{x^2} &= 0\\ \overline{xy} - a_2\overline{x^3} - a_1\overline{x^2} - a_0\overline{x} &= 0\\ \overline{y} - a_2\overline{x^2} - a_1\overline{x} - a_0 &= 0\end{aligned}$$

Again we are only interested in the constant term, a_0 .

Eliminate a_2 by multiplying the second equation by $\overline{x^2}$ and the third by $\overline{x^3}$ and subtracting. Similarly multiply the first equation by $\overline{x^2}$ and the second by $\overline{x^4}$ and subtract. This gives

$$\begin{aligned}\overline{x^2}\,\overline{xy} - \overline{y}\,\overline{x^3} - a_1\left(\left(\overline{x^2}\right)^2 - \overline{x}\,\overline{x^3}\right) - a_0\left(\overline{x^2}\,\overline{x} - \overline{x^3}\right) &= 0\\ \overline{x^2}\,\overline{x^2 y} - \overline{y}\,\overline{x^4} - a_1\left(\overline{x^2}\,\overline{x^3} - \overline{x}\,\overline{x^4}\right) - a_0\left(\left(\overline{x^2}\right)^2 - \overline{x^4}\right) &= 0\end{aligned}$$

To eliminate a_1 multiply second equation by $\left(\overline{x^2}\right)^2 - \overline{x}\,\overline{x^3}$ and the first equation by $\overline{x^2}\,\overline{x^3} - \overline{x}\,\overline{x^4}$ subtract and rearrange to give:

$$a_0 = \frac{\left(\overline{x^2}\,\overline{x^4} - \left(\overline{x^3}\right)^2\right)\overline{y} + \left(\overline{x^2}\,\overline{x^3} - \overline{x}\,\overline{x^4}\right)\overline{xy} + \left(\overline{x}\,\overline{x^3} - \left(\overline{x^2}\right)^2\right)\overline{x^2 y}}{\Delta}$$

where

$$\Delta = \overline{x^2}\,\overline{x^4} - \left(\overline{x^3}\right)^2 - \overline{x}\,\overline{x^4} + 2\overline{x}\,\overline{x^2}\,\overline{x^3} - \left(\overline{x^2}\right)^3$$

This expression for Δ corrects a misprint in Barlow [48] where a similar expression was derived for the un-weighted case.

The variance in a_0 is determined from the variances, σ_i^2, in the y_i values used as before.

$$Var(a_0) = \sum\left(\frac{\partial a_0}{\partial y_i}\right)^2 \sigma_i^2$$

This analysis is simplified if we substitute

$$f = \frac{\left(\overline{x^2}\,\overline{x^4} - \left(\overline{x^3}\right)^2\right)}{\Delta},\ g = \frac{\left(\overline{x^2}\,\overline{x^3} - \overline{x}\,\overline{x^4}\right)}{\Delta},\ h = \frac{\left(\overline{x}\,\overline{x^3} - \left(\overline{x^2}\right)^2\right)}{\Delta}$$

so that

$$a_0 = f\,\bar{y} + g\,\overline{xy} + h\,\overline{x^2 y}$$

Then

$$\frac{\partial a_0}{\partial y_i} = \frac{1}{\sum \frac{1}{\sigma_i^2}} \left(f \frac{1}{\sigma_i^2} + g \frac{x_i}{\sigma_i^2} + h \frac{x_i^2}{\sigma_i^2} \right)$$

Hence

$$Var(a_0) = \sum \left(\frac{1}{\sum \frac{1}{\sigma_i^2}} \left(f \frac{1}{\sigma_i^2} + g \frac{x_i}{\sigma_i^2} + h \frac{x_i^2}{\sigma_i^2} \right) \right)^2 \sigma_i^2$$

Expanding the square and writing in terms of weighted means, as was done for the linear case, the result can be written in the form

$$Var(a_0) = \frac{1}{\sum \frac{1}{\sigma_i^2}} \left(f^2 + g^2\,\overline{x^2} + h^2\,\overline{x^4} + 2f\,g\,\bar{x} + 2g\,h\,\overline{x^3} + 2f\,h\,\overline{x^2} \right).$$

ACKNOWLEDGMENT

All data are freely available on the IMOS data portal www.aodn.org.au. We acknowledge the effort of the IMOS ACORN team in operation and maintaining these two radar systems in arduous conditions. We are grateful to the anonymous reviewers who provided some useful suggestions that have improved the paper.

REFERENCES

[1] Paduan, J., and L, Washburn, 2013: High Frequency Radar Observations of Ocean Surface Currents, *Annu. Rev. Mar. Sci*., 5, pp. 17.1–17.22.

[2] Crombie, D. D., 1955: Doppler spectrum of sea echo at 13.56Mc/s. *Nature, Lond*, **175**, 681.

[3] Gurgel, K. W., Antonischki G., Essen H. H., and T. Schlick, 1999: Wellen Radar WERA: a new ground-wave HF radar for ocean remote sensing. *Coastal Engineering*, **37**, 219–234

[4] Gordon, R. L. (1996). Acoustic doppler current profiler principles of operation a practical primer. *RD Instruments, P/N*

951-6069-00, 52 p.
[5] Paduan, J. D., and L. K. Rosenfeld, 1996: Remotely sensed surface currents in Monterey Bay from shorebased HF radar (Coastal Ocean Dynamics Application Radar). *J. of Geophys. Res*., **101** (C9), 20,669-20,686.
[6] Graber, H. C., Haus, B. K., Shay, L. K., and R. D. Chapman, 1997: HF radar comparisons with moored estimates of current speed and direction: expected differences and implications. *J. of Geophys. Res*., **102**, 18,749–18,766.
[7] Cosoli, S., Mazzoldi, A., and M. Gačić, 2010: Validation of surface current measurements in the Northern Adriatic Sea from high-frequency radars. *J. Atmos. Ocean. Technol.*, **27**, (6), 908-919.
[8] Shay, L. K, Graber, H.C., Ross, D.B., Chapman, R.D. (1995): Mesoscale Ocean Surface Current Structure Detected by High-Frequency Radar, *Journal of Atmospheric and Oceanic Technology* 12(4):881-900.
[9] Kundu, P. K., 1976: Ekman veering observed near the ocean bottom. *J. of Phys. Oceanogr*., **6**, 238-242.
[10] Robinson, A. M., Wyatt L. R., and M. J. Howarth, 2011: A two year comparison between HF radar and ADCP current measurements in Liverpool Bay, *Journal of Operational Oceanography*, **4**, 33-45.
[11] Shay L. K., Lentz, S.J., Graber, H.C., Haus, B.K.(1998): Current Structure Variations Detected by High-Frequency Radar and Vector-Measuring Current Meters, *Journal of Atmospheric and Oceanic Technology*, 15(1):237-256.
[12] Shay, L. K., Martinez-Pedraja, J., Cook, T. M., Haus, B. K., and R. H. Weisberg, 2007: High-frequency radar mapping of surface currents using WERA. *J. Atmos. Oceanic Technol.,* **24,** 484–503.
[13] Mao, Y., and J. L. Luick, 2014: Circulation in the southern Great Barrier Reef studied through an integration of multiple remote sensing and in situ measurements, *J. Geophys. Res. Oceans*, **119**, doi:10.1002/2013JC009397.
[14] Schaeffer, A., Roughan, M., and B. D. Morris, 2013: Cross-Shelf Dynamics in a Western Boundary Current Regime: Implications for Upwelling. *Journal Physical Oceanography*, **43**, 1042–1059.
[15] Schaeffer, A., Roughan, M., Wood, J. E., 2014: Observed bottom boundary layer transport and uplift on the continental shelf adjacent to a western boundary current, *J. Geophys. Res.: Oceans*, 119, 1-19.
[16] Mantovanelli, A., Heron, M.L., and A. Prytz, 2010. The use of HF radar surface currents for computing Lagrangian trajectories: benefits and issues. Oceans 2010, *IEEE Xplore*, Sydney, 1-6.
[17] Roughan M., and J.H. Middleton, 2002: A comparison of observed upwelling mechanisms off the east coast of Australia, *Continental Shelf Research***, 22**, 2551-2572.
[18] Roughan M., and J.H Middleton, 2004: On the East Australian Current: Variability, encroachment, and upwelling. *Journal of Geophysical Research*, **109**, C07003 - C07003.
[19] Skolnik, M. I., 1981: Introduction to Radar Systems. *Mc-Graw Hill*.
[20] Harris F. J., ``On the use of windows for harmonic analysis with the discrete Fourier transform," *Proceedings of the IEEE*, vol. 66, pp. 51-83, 1978.
[21] Gačić,M., Kovačevića, V., Cosoli, S., Mazzoldi, A., Paduan, J.D., Mancero-Mosquera, I., Yaria, S. 2009: Surface current patterns in front of the Venice Lagoon, *Estuarine, Coastal and Shelf Science*, 82 (3), 485-494.
[22] Mantovanelli, A., Steinberg C., Heron, S. Williams, D. K., , Spagnol S., (2015), Data quality control for ADCP with pressure sensor and extraction of the topmost depth stratum. *Report Australian Institute of Marine Science (AIMS)*, 45 pp.
[23] Prytz A., 2012: The Australian Coastal Ocean Radar Network: WERA processing and quality control. http://imos.org.au/fileadmin/user_upload/shared/ACORN/ACORN_workshop_2012/Prytz.pdf last accessed 11/04/2016.
[24] Barrick, D., 1980: Accuracy of parameter extraction from sample-averaged sea-echo Doppler spectra. *IEEE Trans on Antennas and Propagation*, **28**, 1-11.
[25] Wyatt, L.R., 2014: Error analysis for ACORN HF radars, ACORN report 2014-2, available from http://imos.org.au/fileadmin/user_upload/shared/ACORN/reports/2014-2-vector-errors.pdf. Last accessed 11/04/2016
[26] Chapman R.D., L.K. Shay, H.C. Graber, J.B.Edson, and A. Karachintsev., 1997: On the accuracy of HF radar surface current measurements: Intercomparisons with ship-based sensors. *J. Geophys. Res*., **102**:18737–18748.
[27] Cook, T.M., Shay, L.K. 2002: Surface M2 tidal currents along the North Carolina shelf observed with a high-frequency radar. *J. Geophys. Res*.,107, C12, 3222, doi:10.1029/2002JC001320, 2002.
[28] Fernandez, D. M., Meadows L. A., Vesecky J. F., Teague C. C., Paduan J. D., and P. Hansen, 2000: Surface current measurements by HF radar in freshwater lakes. *IEEE J. Oceanic Eng*., **25**, 458–471.
[29] Wyatt L.R., D, Atwater, A, Mantovanelli, A, Prytz, and S, Rehder, 2013: The Australian Coastal Ocean Radar Network data availability and quality. *Proceedings 14th International Radar Symposium (IRS)*, **1**, 405-410.
[30] Orfanidis. S. J., 1996: *Introduction to Signal Processing*. Prentice Hall International Inc.
[31] Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res*., **106**, 7183-7192.
[32] Bowers, J.A., Morton I.D., and G.I. Mould, 2000: Directional statistics of the wind and waves. *Applied Ocean Research*, 33, 13-30.
[33] Fisher, N. I. and A. J. Lee, 1983: A correlation coefficient for circular data. *Biometrika*, **70**, 327-332.
[34] Fisher, N. I., 1993: *Statistical analysis of circular data*. Cambridge University Press.
[35] Wyatt L. R., J. J. Green, A. Middleditch, M D Moorhead, J. Howarth, M. Holt, and S. Keogh, 2006: Operational wave, current and wind measurements with the Pisces HF radar. *IEEE Journal of Oceanic Engineering*, 31, 819-83
[36] Lomb, N. R., 1976. Least-Squares Frequency Analysis of Unequally Spaced Data. *Astrophysics and Space Science*, **39**,

447–462.
[37] Scargle, J. D., 1982. Studies in Astronomical Time Series Analysis. II. Statistical Aspects of Spectral Analysis of Unevenly Spaced Data. *Astrophysical Journal*. **263**, 835–853.
[38] Schmid, H., 2012. How to use the FFT and Matlab's pwelch function for signal and noise simulations and measurements, report available from http://www.schmid-werren.ch/hanspeter/publications/2012fftnoise.pdf, last accessed 11/04/2015.
[39] Plimpton, P. E., Freitag, H. P. & McPhaden, M. J., 2004: Processing of subsurface ADCP data in the Equatorial Pacific. NOAA Technical Memorandum OAR, Pacific Marine Environmental Laboratory (PMEL-125, Contribution No. 2635), Washington, 40 pp.
[40] Plimpton, P. E., Freitag, H. P. & McPhaden, M. J., 1997: ADCP velocity errors from pelagic fish schooling around equatorial moorings. *J. Atmos. Ocean. Technol.*, *14*, 1212-1223.
[41] Mantovanelli, A., Heron, M.L., Heron, S.F., and C.R. Steinberg, 2012: Relative dispersion of surface drifters in a barrier reef region, *J. Geophys. Res.*, **117**, C11016, 1-15.
[42] White, L. & Wolanski, E., 2008: Flow separation and vertical motions in a tidal flow interacting with shallow-water island. *Estuarine, Coastal and Shelf Science*, 77, 457-466.
[43] Mantovanelli, A., Keating, S., Wyatt, L., Roughan, M., Schaeffer, A. (2016). Eulerian and Lagrangian characterization of two counter-rotating frontal eddies in a western boundary current. Under review *Journal of geophysical research*: *oceans*.
[44] Schaeffer, A., Gramoulle, A., Roughan, M., Mantovanelli, A. (2016) Submesoscale meanders and frontal eddies along the East Australian Current observed by HF radars. Under review *Journal of geophysical research: oceans*.
[45] Lipa, B., B. Nyden, D. Barrick, and J. Kohut, 2008: HF radar sea-echo from shallow water. *Sensors*, 8,4611–4635, doi:10.3390/s8084611.
[46] Wyatt, L.R., 2014: Performance of the ACORN HF radar systems, ACORN report 2014-1, available from http://imos.org.au/fileadmin/user_upload/shared/ACORN/reports/2014-1-radar-performance.pdf. Last accessed 11/04/2016.
[47] Schafer, R. W., 2011: What Is a Savitzky-Golay Filter? *IEEE Signal Processing Magazine*, July 2011, 111-117.
[48] Barlow, R. J., 1989: *Statistics – A guide to the use of statistical methods in the physical sciences*. John Wiley and Sons.