



**HAL**  
open science

## **Exploitation of error correlation in a large analysis validation: GlobCurrent case study**

Richard E. Danielson, Johnny A. Johannessen, Graham D. Quartly, Marie-Helene Rio, Bertrand Chapron, Fabrice Collard, Craig Donlon

### **► To cite this version:**

Richard E. Danielson, Johnny A. Johannessen, Graham D. Quartly, Marie-Helene Rio, Bertrand Chapron, et al.. Exploitation of error correlation in a large analysis validation: GlobCurrent case study. *Remote Sensing of Environment*, 2018, 217, pp.476-490. <10.1016/j.rse.2018.07.016>. <hal-04202209>

**HAL Id: hal-04202209**

**<https://hal.science/hal-04202209v1>**

Submitted on 22 Jul 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

---

## Exploitation of error correlation in a large analysis validation: GlobCurrent case study

Danielson Richard E. <sup>1,\*</sup>, Johannessen Johnny A. <sup>1</sup>, Quartly Graham D. <sup>2</sup>, Rio Marie-Helene <sup>3</sup>, Chapron Bertrand <sup>4</sup>, Collard Fabrice <sup>5</sup>, Donlon Craig <sup>6</sup>

<sup>1</sup> Nansen Environm & Remote Sensing Ctr, Bergen, Norway.

<sup>2</sup> Plymouth Marine Lab, Plymouth, Devon, England.

<sup>3</sup> Collecte Localisat Satellites, Ramonville St Agne, France.

<sup>4</sup> IFREMER, Plouzane, France.

<sup>5</sup> OceanDataLab, Locmaria Plouzane, France.

<sup>6</sup> European Space Agcy, Noordwijk, Netherlands.

\* Corresponding author : Richard E. Danielson, email address : [rickedanielson@gmail.com](mailto:rickedanielson@gmail.com)

---

### Abstract :

An assessment of variance in ocean current signal and noise shared by in situ observations (drifters) and a large gridded analysis (GlobCurrent) is sought as a function of day of the year for 1993-2015 and across a broad spectrum of current speed. Regardless of the division of collocations, it is difficult to claim that any synoptic assessment can be based on independent observations. Instead, a measurement model that departs from ordinary linear regression by accommodating error correlation is proposed. The interpretation of independence is explored by applying Fuller's (1987) concept of equation and measurement error to a division of error into shared (correlated) and unshared (uncorrelated) components, respectively. The resulting division of variance in the new model favours noise. Ocean current shared (equation) error is of comparable magnitude to unshared (measurement) error and the latter is, for GlobCurrent and drifters respectively, comparable to ordinary and reverse linear regression. Although signal variance appears to be small, its utility as a measure of agreement between two variates is highlighted.

Sparse collocations that sample a dense (high resolution) grid permit a first order autoregressive form of measurement model to be considered, including parameterizations of analysis-in situ error cross-correlation and analysis temporal error autocorrelation. The former (cross-correlation) is an equation error term that accommodates error shared by both GlobCurrent and drifters. The latter (autocorrelation) facilitates an identification and retrieval of all model parameters. Solutions are sought using a prescribed calibration between GlobCurrent and drifters (by variance matching). Because the true current variance of GlobCurrent and drifters is small, signal to noise ratio is near zero at best. This is particularly evident for moderate current speed and for the meridional current component

---

## Highlights

► Signal/noise of two ocean current datasets are compared by day of the year and speed. ► A first order autoregressive form of regression model and its solution are derived. ► Fuller's (1987) equation/measurement errors are applied to shared/unshared errors. ► Shared true variance is smaller, and equation error larger, than expected. ► Measurement error is saturated at ordinary and reverse least squares bounds.

**Keywords** : Measurement model, Ocean current, Collocation, Validation

---

## 1 **1. Introduction**

2       The idea that errors in two collocated estimates of ocean current could  
3 be independent of each other is, like geostrophy itself, both practical and in-  
4 structive. The difficult implication is that only signal (or truth) is correlated  
5 while noise (or error) is not. Considering that all measurement models are ap-  
6 proximate (Box, 1979), such a clean separation may be ideal in principle but  
7 is probably quite rare in practice. The purpose of this study is to assess the  
8 GlobCurrent analysis, but the need to accommodate cross-correlated errors

9 between GlobCurrent and drifters is not matched by an existing framework  
10 for doing so. Thus, a new measurement model is called for.

11 Although there is no evidence that ocean current signal is dictated by  
12 drifters alone, drifters are employed to refine the mean dynamic topography  
13 (MDT; Rio and Hernandez 2004; Rio et al. 2014). Thus, measurement errors  
14 may be correlated because the MDT effectively determines GlobCurrent in a  
15 time-mean sense. Measurement error is not the only type of error, however.  
16 Perhaps the simplest measurement models (including all models of this study)  
17 assume that truth and error in a dataset are *additive* and the signal in two  
18 datasets can be *linearly* related. There is growing evidence that for datasets  
19 that do not conform exactly to such assumptions, an associated *equation*  
20 *error* term needs to be considered (Fuller, 1987; Carroll and Ruppert, 1996;  
21 Kipnis et al., 1999). It is precisely because equation error may be strongly  
22 correlated that datasets should not necessarily be considered independent,  
23 even if there is no apparent physical relationship between them.

24 This study represents an experiment in ocean surface current validation  
25 that draws on advances in measurement modelling, notably in hydrology and  
26 epidemiology, but contemporary surface current validation also informs this  
27 work. Johnson et al. (2007) attribute differences between the OSCAR five-  
28 day current analysis and in situ observations in part to dynamic processes  
29 that are difficult to resolve (e.g., tropical instability waves and high latitude  
30 eddies). Additionally, although larger signal and noise are resolved by OS-  
31 CAR relative to an assimilative model, Johnson et al. highlight the existence  
32 of intrinsic challenges in capturing the meridional current near the equator  
33 and variability in both components near the poles.

34 Surface current validation by Blockley et al. (2012) and Sudre et al. (2013)  
 35 similarly acknowledge in situ error. Blockley et al. highlight differences in  
 36 the western equatorial Pacific between surface currents that they derive from  
 37 in situ observations and the FOAM assimilative model. Global correlation  
 38 between model and observations is again much better for the zonal current  
 39 component (versus meridional), especially in the tropics and north Pacific  
 40 (reduced correlation in the Atlantic is attributed to slightly greater cover-  
 41 age by eddies). Although the GECKO satellite-based analysis of Sudre et  
 42 al. finds corresponding systematic variations (by latitude and current com-  
 43 ponent), their combination of geostrophic and Ekman estimates is also sig-  
 44 nificantly correlated with in situ estimates. It is the agreement between, and  
 45 independence of, two such estimates that we wish to reconsider below.

46 It is convenient to speak of correlation either in terms of signal and noise,  
 47 or equivalently, truth and error. It is also useful to distinguish between  
 48 the (spatial or temporal) autocorrelation of a single variable and the cross-  
 49 correlation of two variables. Geophysical modelling approaches (including  
 50 this study) often assume that autocorrelation should be easy to find in high  
 51 resolution (analysis) data, and for some (in situ) collocation subset, that an  
 52 affine signal model with additive, orthogonal (or signal-uncorrelated) noise  
 53 applies. More formally, if two collocated ocean current datasets ( $I$  and  $A$ )  
 54 are divided parsimoniously into shared truth ( $t$ ) and additive error ( $\epsilon$ ) such  
 55 that

$$\begin{array}{l} \text{in situ } I = t + \epsilon_I \\ \text{analysis } A = \alpha + \beta t + \epsilon_A, \end{array} \quad (1)$$

56 then the affine signal model is a linear calibration involving an unbiased in-

57 tercept ( $\alpha$ ) and slope ( $\beta$ ) that relates signal in the two datasets by  $A_{signal} =$   
58  $\alpha + \beta I_{signal}$  (where  $I_{signal} = t$ ). The measurement model (1) is known as a  
59 regression model with errors in the variables ( $I$  and  $A$ ) but (with reference  
60 to a linear relationship between  $I_{signal}$  and  $A_{signal}$ ) no error in the equation  
61 (Fuller, 2006). Note also that cross-correlation is only expected from truth,  
62 or perhaps error, that is shared between datasets and that (1) omits a parti-  
63 tion of error into shared and unshared, or cross-correlated and uncorrelated,  
64 components.

65 If there is no obvious physical dependence between datasets, then there  
66 is no guarantee that shared error, or shared truth for that matter, exist. Be-  
67 cause the geophysical interpretation of cross-correlated error continues to  
68 evolve, this concept of sharing is at least partly unfamiliar, even in the  
69 context of two datasets (1). An established explanation in the context of  
70 three datasets (Stoffelen, 1998; O’Carroll et al., 2008) focuses on the cross-  
71 correlated part of representativeness error: it is natural for correlation to  
72 exist between two higher resolution datasets on scales that a lower resolution  
73 dataset cannot resolve, but if there is a truth that is shared by all three  
74 datasets, then by definition, this truth is also low resolution and any high  
75 resolution correlation must be considered erroneous, albeit perfectly natural.  
76 Errors of representation in geophysics (e.g., mismatches that can be written  
77 as a component of  $\epsilon_I$  or  $\epsilon_A$ , as in Gruber et al. 2016b) refer to information  
78 that is beyond some true, or target, spatiotemporal resolution limit. How-  
79 ever, if shared truth does exist, it follows that the most generic and inclusive  
80 definition of limitations in this truth is needed to define what remains in each  
81 individual dataset as error.

82       Stoffelen’s introduction of the triple collocation model provides an im-  
83 portant description, and one of the earliest quantifications, of representative-  
84 ness error (see also Vogelzang et al. 2011). Nevertheless, the triple colloca-  
85 tion model is just identified, so the parameters sought (see Appendix) are  
86 equal in number to the first and second moment equations that are available  
87 (cf. Gillard and Iles 2005). A familiar characteristic of this model (like sim-  
88 pler regression models) is its limited flexibility to identify more parameters.  
89 Hence, correlated representativeness error, and cross-correlated error in gen-  
90 eral, must either be known in advance or perhaps be justifiably small for a  
91 retrieval of the triple collocation parameters.

92       Caires and Sterl (2003) discovered a way to explore cross-correlated er-  
93 ror (between altimeters) in comparative applications of the triple collocation  
94 model. They examined significant wave height and 10-m wind speed es-  
95 timates from buoys and two altimeters, which were carefully averaged to  
96 be comparable in space and time with collocated ERA-40 estimates. Be-  
97 cause representativeness errors were reduced by this averaging, it was postu-  
98 lated that any remaining ERA-40 cross-correlated errors could be neglected  
99 if ERA-40 did not assimilate an observational dataset. A bound on cross-  
100 correlated error was then estimated for the altimeters, whose uncorrelated  
101 error was found to be relatively low when retrieved together with ERA-40  
102 rather than separately with ERA-40 and buoys. Consideration of this bound  
103 yielded an increase in altimeter error variance by a factor of two or more, but  
104 Caires and Sterl suggested that cross-correlated error may have been smaller.

105       Janssen et al. (2007) examined wave height data from two altimeters,  
106 buoys, and an ECMWF wave hindcast, and employed an iterative form of

107 orthogonal regression (Gillard and Iles, 2005) with estimates of uncorrelated  
108 error from the triple collocation model. An important acknowledgement  
109 was given of the linear calibration in (1) being a potential source of cross-  
110 correlated error (i.e., where a nonlinear signal model might be appropriate in-  
111 stead). As in Caires and Sterl (2003), it was postulated that cross-correlated  
112 errors could be neglected if data (or systematic errors) were not assimilated,  
113 but uncorrelated altimetric error was again found to be relatively low when  
114 the triple collocation model was applied to both altimeters at once. Janssen  
115 et al. proposed additional model equations (using ECMWF first guess and  
116 analysis wave products) to quantify rather than just bound most errors, but  
117 found that altimetric error, including its cross-correlated component, was  
118 small.

119 Methods of collocating buoy, radiometer, and microwave SST estimates  
120 (e.g., O'Carroll et al. 2008) also point to cross-correlated error being small,  
121 but only insofar as representativeness error is tested, as above, by paramete-  
122 ter comparisons. A novel assessment of cross-correlated error has also been  
123 given using a high resolution, rescaled in situ dataset as a proxy for truth.  
124 Yilmaz and Crow (2014) use this proxy to directly characterize terms of the  
125 triple collocation model based on soil moisture from an assimilative model  
126 and soil moisture retrievals from passive (AMSR-E) and active (ASCAT)  
127 satellites. The dependence of satellite retrievals is notable because signifi-  
128 cant cross-correlated errors are found. This study concludes that zero error  
129 cross-correlation is a tenuous assumption of the triple collocation model as  
130 its corresponding bias in parameter retrievals is systematic.

131 Contemporary calibration and validation studies have introduced a grow-

132 ing list of geophysical dataset differences, which taken together, define cor-  
 133 responding limitations on shared truth. However, perhaps the most generic  
 134 characterization of these limitations is found in the measurement modelling  
 135 literature: Fuller (1987) defines measurement error in the familiar sense of  
 136 random data departures from a linear regression solution and distinguishes  
 137 *equation error* as random departures from the linear signal model of (1),  
 138 owing to nonlinearity in the signal model of interest. Carroll and Ruppert  
 139 (1996) expose the importance of this refinement in a geophysical application  
 140 and, as noted above, Janssen et al. (2007) highlight that such nonlinearity is  
 141 a potential source of cross-correlated error.

142 The combination of measurement error and equation error is useful to bet-  
 143 ter accommodate limitations in the scope of a shared truth. With reference  
 144 to person-specific bias in epidemiology, Kipnis et al. (1999, 2002) introduce  
 145 equation error as two additional terms ( $\epsilon_{QI}$  and  $\epsilon_{QA}$ ) in (1) that lead to

$$\begin{array}{l}
 \text{in situ } I = t + \epsilon_{QI} + \epsilon_I \\
 \text{analysis } A = \alpha + \beta t + \epsilon_{QA} + \epsilon_A,
 \end{array} \tag{2}$$

146 where  $\epsilon_I$  and  $\epsilon_A$  are now random departures from a possibly nonlinear signal  
 147 model. Carroll and Ruppert (1996) note that applications of (2) have been  
 148 limited, possibly because if  $\epsilon_{QI}$  and  $\epsilon_{QA}$  are considered to be independent  
 149 of other errors, they can be recombined with  $\epsilon_I$  and  $\epsilon_A$  to yield the simpler  
 150 equation (1) with its original properties intact (Moberg and Brattström,  
 151 2011). Below, the same linear signal model as in (1) will be considered,  
 152 with shared equation error defined by  $\epsilon_{QI} = \epsilon_{QA}$  and total error involving  
 153 both shared and unshared components. In other words, equation error is  
 154 not independent so it is important to quantify this as a separate term in our

155 application of (2).

156 In addition to the interpretation of cross-correlated errors, there remains  
157 the issue of identifying solutions to increasingly sophisticated statistical mod-  
158 els. Increasing the number of collocated datasets (e.g., Janssen et al. 2007;  
159 Zwieback et al. 2012; Gruber et al. 2016a) is one approach. However, an  
160 important development in the geophysical literature is the recognition by Su  
161 et al. (2014) that three or more datasets may be unnecessary, as collocation  
162 models appear to belong to a broader family of instrumental variable  
163 regression models, and within this family, a precedent exists for using lagged  
164 variables as instruments. Following Su et al., this implies that by embracing  
165 autocorrelation, strategies should continue to emerge that depend on fewer  
166 datasets to identify a larger number of collocations and statistical model pa-  
167 rameters. By comparison with the error-in-variables model (1), the novelty  
168 of the strategy proposed below is that it also permits the retrieval of variance  
169 in shared error and, in one ocean current experiment, also equation error.

170 The present study seeks to advance measurement modelling and parame-  
171 ter identification with the benefit of error correlation. The focus is on ocean  
172 surface current validation, but general supporting concepts and terms (such  
173 as *measurement model*) are provided in the Appendix. The next section de-  
174 scribes the collocation of GlobCurrent and drifters and proposes a commonly  
175 prescribed linear relationship between them that addresses the difference in  
176 variance between these two datasets. Formulation of a measurement model  
177 that permits error correlation to be exploited is given in Section 3. We then  
178 describe the strong and weak constraints that allow a retrieval of all model  
179 parameters and assess the performance of GlobCurrent and drifter data in

180 Section 4. Throughout this paper, equal emphasis is placed on true variance  
181 and on the contributions to total error. Discussion of inferences based on the  
182 division of variance into shared truth and error are highlighted in Section 5  
183 and Section 6 contains the conclusions.

## 184 **2. Selection of a calibration**

185 We begin with the idea that GlobCurrent and drifters provide estimates  
186 of fundamentally different ocean currents, but they also provide overlapping  
187 views of a true (or target) ocean current that can be represented at 15 m  
188 below the surface on a 6-h,  $1/4^\circ$  grid. By any definition of shared truth,  
189 both GlobCurrent and drifters have errors. GlobCurrent is an analysis that  
190 linearly combines the geostrophic and Ekman components. Drifters respond  
191 locally to a combination of geostrophic, Ekman, tidal, inertial, Stokes, and  
192 wind drift processes, including (erroneous) processes on scales smaller and  
193 faster than the GlobCurrent grid can resolve. In general, such differences  
194 can be considered a mismatch in their supports (see Appendix). Nearest-  
195 neighbour collocations of drifters (whose drogues move roughly with the 15-  
196 m current) and GlobCurrent (also at 15 m, with additional samples at daily  
197 intervals) are considered below.

198 Six-hourly drifter velocity has been estimated following Hansen and Poulain  
199 (1996). We restrict attention to drifters whose continuous drogue presence  
200 was confirmed by objective or subjective means (Rio, 2012; Lumpkin et al.,  
201 2013). The resulting geographic distribution for 1993-2015 (Fig. 1) yields  
202 more than eleven million drifter and GlobCurrent zonal and meridional ve-  
203 locity estimates (Danielson 2017; a comparable number of drifters lost their

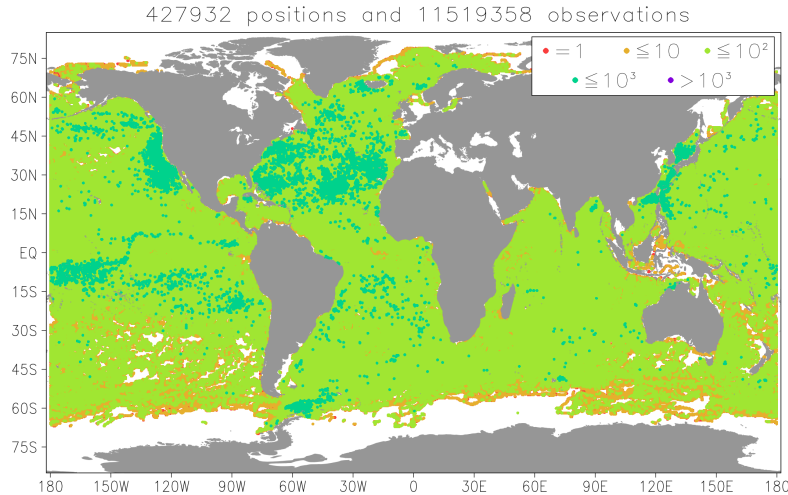


Figure 1: Number of surface drifter velocity observations between January 1993 and December 2015 (order of magnitude in colour) with drogues attached. Shown are values at the  $1/4^\circ$  resolution of the GlobCurrent grid (i.e., collocations are nearest neighbours).

204 drogues and, being more responsive to surface wind forcing, are ignored). It  
 205 is convenient to divide collocations by even and odd year, with the latter  
 206 subset permitting an independent check on calculations. Below, only the  
 207 even-year subset is discussed but the same conclusions can be obtained from  
 208 the results (available as supplementary material) of the odd-year subset.

209 Joint frequency of occurrence of current speed, including the full range  
 210 of possible linear calibrations of GlobCurrent relative to drifters, is shown in  
 211 Fig. 2. These two-dimensional histograms are rather well behaved following  
 212 removal of about 10% of the most extreme current speeds (Hubert et al.,  
 213 2012). Similar regression slopes are revealed in both the zonal and meridional  
 214 distributions. Between the bounding ordinary and reverse linear regression  
 215 reference slopes (dashed lines) is a slope defined by the ratio of total variance

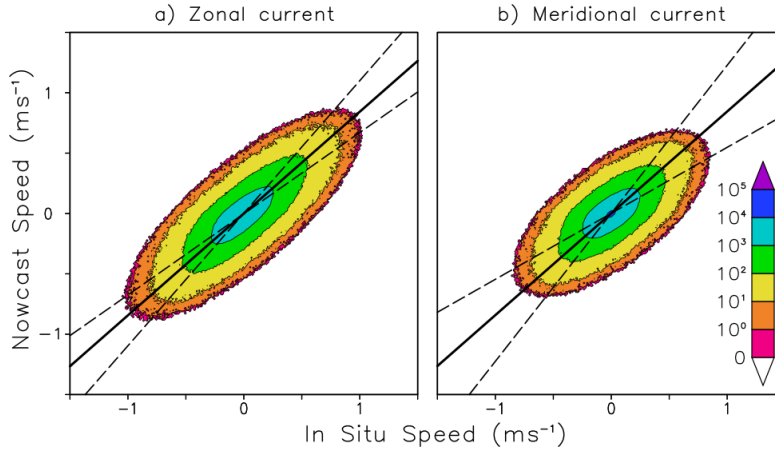


Figure 2: Two-dimensional histograms of a) zonal and b) meridional 15-m current component for 5310226 non-outlier collocations from the even years between 1993 and 2015 (approximately half the collocations of Fig. 1, after removing about 10% of these data as outliers following Hubert et al. 2012). The dashed lines are the ordinary (shallow slope) and reverse (steep slope) linear regression references for each current component. The slope of the solid line is defined by the GlobCurrent–drifter variance ratio (the same ratio for both current components; see next section). The logarithmic colourbar is number of values in  $0.01\text{-ms}^{-1}$  bins.

216 between GlobCurrent and drifters (solid line; defined in the next section).  
 217 Unfortunately, scatter away from these regression lines is a poor indication  
 218 that there might be a component of error variance that is shared between  
 219 GlobCurrent and drifters, or that total error variance might be greater than  
 220 the variance in shared truth.

221 The corresponding one-dimensional (marginal) distributions (Fig. 3) high-  
 222 light an unsurprising difference between current estimates: because drifters  
 223 capture a greater range of physical processes at higher resolution, we find  
 224 fewer low values and more high values than GlobCurrent (with an equal num-

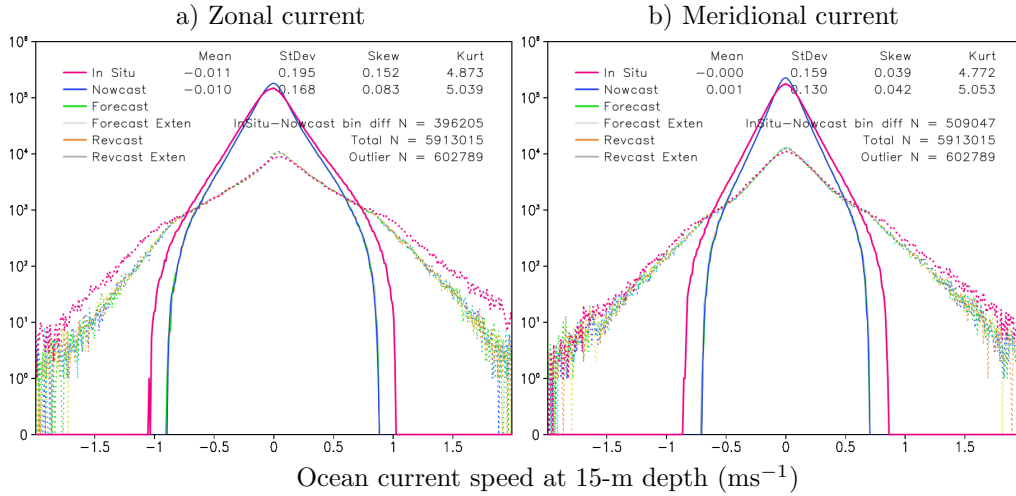


Figure 3: One-dimensional histograms of a) zonal and b) meridional 15-m current component, as in Fig. 2, but including outliers separately (dotted lines). Also shown are drifter (red) and GlobCurrent nowcast (blue), forecast (green and light grey), and reicast (orange and dark grey) histograms. Forecast and reicast data are taken one day (with extended data from two days) before and after each collocation, respectively. Statistical moments of the non-outlier in situ and nowcast distributions are given with a measure of difference between the two (i.e., one half of the in situ minus nowcast bin count difference). The logarithmic ordinate is number of values in  $0.01\text{-ms}^{-1}$  bins.

225 ber at about  $\pm 0.15 \text{ ms}^{-1}$ ). Also as expected, GlobCurrent samples at two  
 226 days (extended forecast) and one day (forecast) before each drifter (in situ)  
 227 observation, as well as one day (reicast) and two days (extended reicast)  
 228 after, have the same distribution as the GlobCurrent collocations (nowcast).  
 229 Outliers are shown separately by dotted lines in Fig. 3 and are identified  
 230 by minimizing the covariance matrix determinant for the six estimates of  
 231 zonal and meridional current (Hubert et al., 2012). Because covariance (and  
 232 skewness and kurtosis) are sensitive to outliers (McColl et al., 2014; Su et al.,  
 233 2014), collocation groups are trimmed by about 10% before other calculations

234 are performed. Often this excludes extreme values in the zonal or meridional  
 235 component and values close to zero in the opposite component.

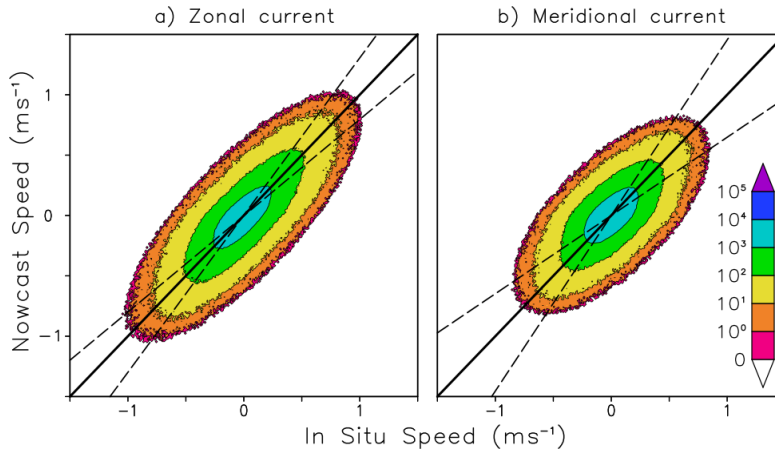


Figure 4: As in Fig. 2, but after dividing all GlobCurrent data by 0.84 (i.e., the ratio of nowcast to drifter standard deviation), where zonal and meridional components are expressed as complex numbers and the same variance match is applied to both components.

236 The distinction between cross-correlated and uncorrelated error is suf-  
 237 ficiently novel that initial solutions of (2) benefit from the assumption of  
 238 a fixed calibration that can be applied uniformly. (Subsequent work will  
 239 seek a general, varying solution, but this simplification applies to all exper-  
 240 iments below.) An assumption that would be consistent with the mismatch  
 241 in GlobCurrent and drifter *support* (rather than a bias between them) is  
 242 that both are already unbiased. However, we note in Section 4 that if cal-  
 243 ibration is bounded by ordinary and reverse linear regression (dashed lines  
 244 in Fig. 2), then this assumption would not apply to all collocation subsets.  
 245 An alternate assumption that can be applied uniformly, and whose bias is  
 246 familiar in the context of (1), is known as variance matching (Fuller, 2006;

247 Yilmaz and Crow, 2013; Su et al., 2014). This calibration is marked by a  
 248 lack of assumptions about relative error in GlobCurrent and drifters. It fixes  
 249 regression slope midway between the bounding ordinary and reverse linear  
 250 regression solutions (solid line in Fig. 2) and fixes GlobCurrent and drifter  
 251 signal-to-noise ratio (SNR) to be equal. A definition and further implications  
 252 are given in Section 3.

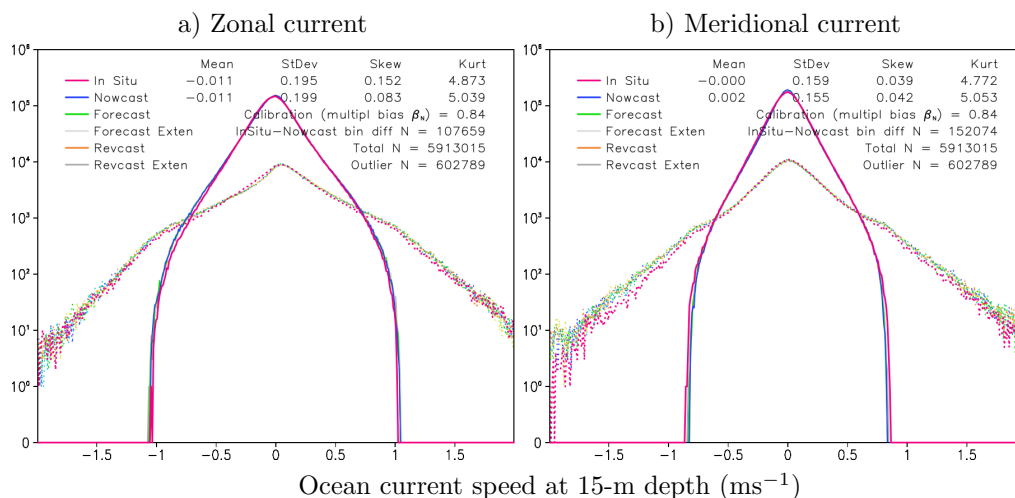


Figure 5: As in Fig. 3, but after dividing all GlobCurrent data by 0.84.

253 Figures 4 and 5 are the result of matching the variance of GlobCurrent to  
 254 that of drifters. (Simultaneous matching of the zonal and meridional com-  
 255 ponents is accomplished by expressing these two components as a complex  
 256 number.) Dividing the GlobCurrent data by a standard deviation ratio of  
 257 0.84 reduces the number of weak values and increases the number of strong  
 258 values, as expected. This calibration removes much of the cumulative dif-  
 259 ference in bin counts: from 7-8% in Fig. 3 to about 2% in Fig. 5. However,  
 260 the distinction between calibrated GlobCurrent and drifters remains, as his-  
 261 togram shape is otherwise preserved (note that skewness and kurtosis are

262 variance-normalized moments) and current direction is unchanged. More-  
 263 over, and notwithstanding important applications to assimilation and model  
 264 validation (e.g., Stoffelen 1998; Tolman 1998), this distinction would remain  
 265 at least under any affine calibration.

### 266 3. Measurement model development

267 A series of experimental models, based initially on the triple collocation  
 268 approach (Stoffelen, 1998; McColl et al., 2014) with solutions sought by the  
 269 method of moments (Gillard and Iles, 2005), have informed the measurement  
 270 model that we will focus on. The first experimental model in this series (3)  
 271 can be criticised for using extrapolated (forecast and revcast) GlobCurrent es-  
 272 timates assuming that extrapolated errors are independent. Gridded altimet-  
 273 ric data are often based on a centered span of up to 12 days of Topex/Jason  
 274 passes and a longer period for Envisat. Similarly for the Ekman (or Stokes)  
 275 current estimates from a model-based analysis, if a model has the wind front  
 276 in the wrong location or an incorrect initial storm intensity, it may retain a  
 277 consistent bias for days. Thus, the assumption of independent errors  $\epsilon$  in a  
 278 slightly modified triple collocation model,

$$\begin{aligned}
 \text{in situ } I &= t + \epsilon_I \\
 \text{forecast } F &= \alpha_F + \beta_F t + \epsilon_F \\
 \text{revcast } R &= \alpha_R + \beta_R t + \epsilon_R,
 \end{aligned} \tag{3}$$

279 can be considered experimental at best. Note that  $\alpha$ ,  $\beta$ ,  $t$ , and  $\epsilon$  are addi-  
 280 tive calibration, multiplicative calibration (or regression slope), truth, and  
 281 error, respectively, and our use of drifters as a calibration reference implies

282 that  $\alpha_I = 0$  and  $\beta_I = 1$ . Here,  $F$  and  $R$  are obtained by extrapolation of  
 283 GlobCurrent from outside a centered window of only a few days.

284 The form of (3) is recognizable in an intermediate (but still unsatisfactory)  
 285 model (4) that includes both GlobCurrent and drifter collocations ( $I$  and  
 286  $N$ ) and retains additive and multiplicative calibration parameters ( $\alpha$  and  
 287  $\beta$ ) for each GlobCurrent estimate. A notable simplification of (4) is that  
 288 extrapolation is replaced by a persistence forecast/revcast, so  $F$  and  $R$  are  
 289 just GlobCurrent samples taken one day before and after each collocation,  
 290 respectively.

$$\begin{aligned}
 \text{in situ } I &= t + \epsilon_I \\
 \text{nowcast } N &= \alpha_N + \beta_N t + \epsilon_N \\
 \text{forecast } F &= \alpha_F + \beta_F t + \epsilon_N + \epsilon_F \\
 \text{revcast } R &= \alpha_R + \beta_R t + \epsilon_N + \epsilon_R.
 \end{aligned} \tag{4}$$

291 The model (4) is overly constrained in its treatment of correlated error, how-  
 292 ever. There is no shared (equation) error between GlobCurrent and drifters  
 293 and a complete sharing of  $N$  errors in  $F$  and  $R$ . In turn, it is perhaps un-  
 294 surprising that there may be effectively no difference (in terms of physical  
 295 insight) between parameter retrievals based on (4) and ordinary and reverse  
 296 linear regression references based on  $I$  and  $N$  alone (Danielson et al., 2017).

297 Two further innovations are required to arrive at the measurement model  
 298 of interest. One is that a first-order autoregressive (AR-1) parameterization  
 299 is probably the simplest way to accommodate both GlobCurrent-drifter error  
 300 cross-correlation as well as GlobCurrent error autocorrelation. Error prop-  
 301 agation is parameterized in the same sense as it might occur in an ocean  
 302 current analysis, with observational error having its biggest impact on an

303 analysis at the time of observation, with a decreasing, but symmetric impact  
 304 at times before and after. The AR-1 form accommodates autocorrelated  
 305 errors (e.g., from altimetry) that also have a symmetric upstream and down-  
 306 stream impact (note that asymmetric error propagation may be appropriate  
 307 in some applications).

308 The second innovation, following Su et al. (2014), is that additional, or ex-  
 309 tended, samples of GlobCurrent are beneficial, assuming these remain inside  
 310 the autocorrelation envelope. The resulting model becomes

$$\begin{aligned}
 \text{in situ } I &= t + \epsilon_I \\
 \text{nowcast } N &= \alpha_N + \beta_N t + \lambda_N \epsilon_I + \epsilon_N \\
 \text{forecast } F &= \alpha_F + \beta_F t + \lambda_F (\lambda_N \epsilon_I + \epsilon_N) + \epsilon_F \\
 \text{extended forecast } E &= \alpha_E + \beta_E t + \lambda_E (\lambda_F (\lambda_N \epsilon_I + \epsilon_N) + \epsilon_F) + \epsilon_E \\
 \text{reicast } R &= \alpha_R + \beta_R t + \lambda_R (\lambda_N \epsilon_I + \epsilon_N) + \epsilon_R \\
 \text{extended reicast } S &= \alpha_S + \beta_S t + \lambda_S (\lambda_R (\lambda_N \epsilon_I + \epsilon_N) + \epsilon_R) + \epsilon_S,
 \end{aligned} \tag{5}$$

311 where Fuller's (1987) equation error, corresponding in (2) to  $\epsilon_{QI} = \epsilon_{QA}$  (Kip-  
 312 nis et al., 1999), is the shared (cross-correlated) error parameterization  $\lambda_N \epsilon_I$ .  
 313 We return to the interpretation of shared and unshared error in  $\epsilon_I$  below.  
 314 The remaining errors are uncorrelated measurement errors, also denoted in-  
 315 dividual errors:  $\epsilon_N$ ,  $\epsilon_F$ ,  $\epsilon_E$ ,  $\epsilon_R$ , and  $\epsilon_S$ .

316 A so-called INFR model, whose name is taken from the data samples  
 317 on the LHS of (4) but whose RHS is taken from (5), has parameters that  
 318 are almost identifiable (in a statistical sense). That is, one can derive 10  
 319 covariance equations (given below) but there are 11 unknown parameters.  
 320 The INFERS model (5) includes an extended forecast and reicast, which  
 321 are GlobCurrent samples two days before and after each collocation. Under

322 the assumption that GlobCurrent errors remain correlated at least over five  
323 days (e.g., as gauged by the product  $\lambda_F \lambda_E \lambda_R \lambda_S$ ), INFERS is more attractive  
324 because there are more covariance equations (21) than unknown parameters  
325 (17). (Of course, with more samples further improvement in the ratio of  
326 these numbers is possible.) Standard assumptions of no correlation between  
327 truth and error (orthogonality) and among individual errors then allow all  
328 elements of the covariance matrix to be defined by

$$\begin{aligned}
Var(I) &= \sigma_t^2 + \sigma_I^2 \\
Var(N) &= \beta_N^2 \sigma_t^2 + \lambda_N^2 \sigma_I^2 + \sigma_N^2 \\
Var(F) &= \beta_F^2 \sigma_t^2 + \lambda_F^2 \lambda_N^2 \sigma_I^2 + \lambda_F^2 \sigma_N^2 + \sigma_F^2 \\
Var(E) &= \beta_E^2 \sigma_t^2 + \lambda_E^2 \lambda_F^2 \lambda_N^2 \sigma_I^2 + \lambda_E^2 \lambda_F^2 \sigma_N^2 + \lambda_E^2 \sigma_F^2 + \sigma_E^2 \\
Var(R) &= \beta_R^2 \sigma_t^2 + \lambda_R^2 \lambda_N^2 \sigma_I^2 + \lambda_R^2 \sigma_N^2 + \sigma_R^2 \\
Var(S) &= \beta_S^2 \sigma_t^2 + \lambda_S^2 \lambda_R^2 \lambda_N^2 \sigma_I^2 + \lambda_S^2 \lambda_R^2 \sigma_N^2 + \lambda_S^2 \sigma_R^2 + \sigma_S^2 \\
Cov(I, N) &= \beta_N \sigma_t^2 + \lambda_N \sigma_I^2 \\
Cov(I, F) &= \beta_F \sigma_t^2 + \lambda_F \lambda_N \sigma_I^2 \\
Cov(I, E) &= \beta_E \sigma_t^2 + \lambda_E \lambda_F \lambda_N \sigma_I^2 \\
Cov(I, R) &= \beta_R \sigma_t^2 + \lambda_R \lambda_N \sigma_I^2 \\
Cov(I, S) &= \beta_S \sigma_t^2 + \lambda_S \lambda_R \lambda_N \sigma_I^2 \\
Cov(N, F) &= \beta_N \beta_F \sigma_t^2 + \lambda_F \lambda_N^2 \sigma_I^2 + \lambda_F \sigma_N^2 \\
Cov(N, E) &= \beta_N \beta_E \sigma_t^2 + \lambda_E \lambda_F \lambda_N^2 \sigma_I^2 + \lambda_E \lambda_F \sigma_N^2 \\
Cov(N, R) &= \beta_N \beta_R \sigma_t^2 + \lambda_R \lambda_N^2 \sigma_I^2 + \lambda_R \sigma_N^2 \\
Cov(N, S) &= \beta_N \beta_S \sigma_t^2 + \lambda_S \lambda_R \lambda_N^2 \sigma_I^2 + \lambda_S \lambda_R \sigma_N^2,
\end{aligned} \tag{6}$$

329 and

$$\begin{aligned}
Cov(F, E) &= \beta_F \beta_E \sigma_t^2 + \lambda_E \lambda_F^2 \lambda_N^2 \sigma_I^2 + \lambda_E \lambda_F^2 \sigma_N^2 + \lambda_E \sigma_F^2 \\
Cov(F, R) &= \beta_F \beta_R \sigma_t^2 + \lambda_F \lambda_R \lambda_N^2 \sigma_I^2 + \lambda_F \lambda_R \sigma_N^2 \\
Cov(F, S) &= \beta_F \beta_S \sigma_t^2 + \lambda_F \lambda_S \lambda_R \lambda_N^2 \sigma_I^2 + \lambda_F \lambda_S \lambda_R \sigma_N^2 \\
Cov(E, R) &= \beta_E \beta_R \sigma_t^2 + \lambda_E \lambda_F \lambda_R \lambda_N^2 \sigma_I^2 + \lambda_E \lambda_F \lambda_R \sigma_N^2 \\
Cov(E, S) &= \beta_E \beta_S \sigma_t^2 + \lambda_E \lambda_F \lambda_S \lambda_R \lambda_N^2 \sigma_I^2 + \lambda_E \lambda_F \lambda_S \lambda_R \sigma_N^2 \\
Cov(R, S) &= \beta_R \beta_S \sigma_t^2 + \lambda_S \lambda_R^2 \lambda_N^2 \sigma_I^2 + \lambda_S \lambda_R^2 \sigma_N^2 + \lambda_S \sigma_R^2.
\end{aligned} \tag{7}$$

330 The corresponding 17 unknowns are true variance ( $\sigma_t^2$ ), multiplicative  
331 calibration for five datasets ( $\beta_N, \beta_F, \beta_E, \beta_R, \beta_S$ ), and error variance for all six  
332 ( $\sigma_I^2, \sigma_N^2, \sigma_F^2, \sigma_E^2, \sigma_R^2, \sigma_S^2$ ). There are also five parameters that gauge GlobCurrent-  
333 drifter error cross-correlation ( $\lambda_N$  is denoted shared error fraction below) and  
334 GlobCurrent error autocorrelation ( $\lambda_F, \lambda_E, \lambda_R, \lambda_S$ ). An analytic solution of  
335 all parameters except  $\sigma_t^2$  and  $\beta_N$  is possible using (6) as a strong constraint  
336 (i.e., using all variances and the covariances involving the GlobCurrent and  
337 drifter collocations  $I$  and  $N$ ). The remaining equations (7) are denoted the  
338 autocovariance equations (i.e., covariances involving only GlobCurrent fore-  
339 cast and revcast samples *FERS*).

340 True variance ( $\sigma_t^2$ ) and multiplicative calibration or regression slope ( $\beta_N$ )  
341 between GlobCurrent and drifters are key measurement model parameters.  
342 In the context of INFERS, these are both free parameters that can be sought  
343 numerically using the autocovariance equations as a weak constraint, that is,  
344 by approaching minima in the difference between the LHS and RHS of (7).  
345 Matching GlobCurrent variance to that of drifters (as in Section 2) provides  
346 all experiments with a fixed, but approximate, slope parameter  $\beta_N$ . In other  
347 words, our focus on a search for true variance is also limited by this assump-

348 tion. It is important to note, moreover, that variance matching provides more  
 349 freedom to retrieve large cross-correlated error because it is midway between  
 350 the bounding ordinary and reverse linear regression solutions (i.e., where all  
 351 variance in either GlobCurrent or drifters is assigned to truth and the possi-  
 352 bility of cross-correlated error is excluded). It follows from this assumption  
 353 that

$$\beta_N^2 = \text{Var}(N)/\text{Var}(I) \quad \Rightarrow \quad \sigma_N^2 = \sigma_I^2(\beta_N^2 - \lambda_N^2). \quad (8)$$

354 The remaining INFERS model parameters are retrieved once a solution  
 355 for  $\sigma_t^2$  is obtained. The weakly constrained minimization of (7) is sought  
 356 between bounds for  $\sigma_t^2$  that are given by  $\text{Var}(I) = \sigma_t^2 + \sigma_I^2$  (i.e., between  $\sigma_t^2 =$   
 357 0 and the ordinary linear regression solution of  $\sigma_I^2 = 0$ ), with the additional  
 358 strong constraint that all other variances ( $\sigma_N^2, \sigma_F^2, \sigma_E^2, \sigma_R^2, \sigma_S^2$ ) also remain  
 359 non-negative. Just like the variance matched solution for  $\beta_N$ , each zonal and  
 360 meridional current component is first expressed as a complex number so that  
 361 17 parameters are identified for both components at the same time (i.e., the  
 362 covariances in (6) and (7) are real parts).

363 The remainder of this study is a diagnostic exploration of the parame-  
 364 ters obtained from (5)-(8) given surface current variations that are jointly  
 365 sampled by GlobCurrent and drifters. As required by INFERS, we also per-  
 366 form a simple check that GlobCurrent samples of truth and error (combined)  
 367 remain inside their autocorrelation envelope: for any group of collocations,  
 368 the minimum correlation between an NFERS pair (i.e., between E and S) is  
 369 expected to be larger than about 0.7. All correlation estimates are obtained  
 370 from the LHS of (6) and (7).

#### 371 4. Performance assessment

372 We introduce a retrieval of measurement model parameters for all 5310226  
373 non-outlier collocations from the even years between 1993 and 2015. This  
374 is followed by retrievals for subsets of this group as a function of day of  
375 the year and current speed. GlobCurrent and drifters appear to provide  
376 complementary information about ocean surface current. The SNR is near  
377 zero at best as variance in a shared true current tends to be smaller than the  
378 variance in total (shared and unshared) error. We also show that shared error  
379 fraction ( $\lambda_N$ ) is quite high. A posteriori, this motivates our accommodation  
380 of cross-correlated error in (5). To the extent that cross-correlated error and  
381 equation error are the same (see Section 5), an important question is raised  
382 of whether a linear signal model and additive errors for GlobCurrent and  
383 drifters can be considered robust (and by what metric). Large individual  
384 (measurement) error is consistent with GlobCurrent and drifters as offering  
385 quite noisy estimates of shared true current variability (again subject to a  
386 linear calibration). In Section 5, we find that individual error is similar to the  
387 ordinary and reverse linear regression reference solutions. In other words, it  
388 is mainly a quantification of shared/correlated truth and error that require  
389 our attention.

390 Figure 6 depicts the absolute difference in LHS minus RHS of the fore-  
391 cast and reforecast autocovariance equations (7). Differences are shown for  
392 all candidate values (i.e., true variance between zero and the variance of  
393 drifters), but model solutions are of interest only where variance is positive  
394 (unshaded region). The target minimum (open purple circle) is the average  
395 of three available local minima (i.e., no minima are associated with the ex-

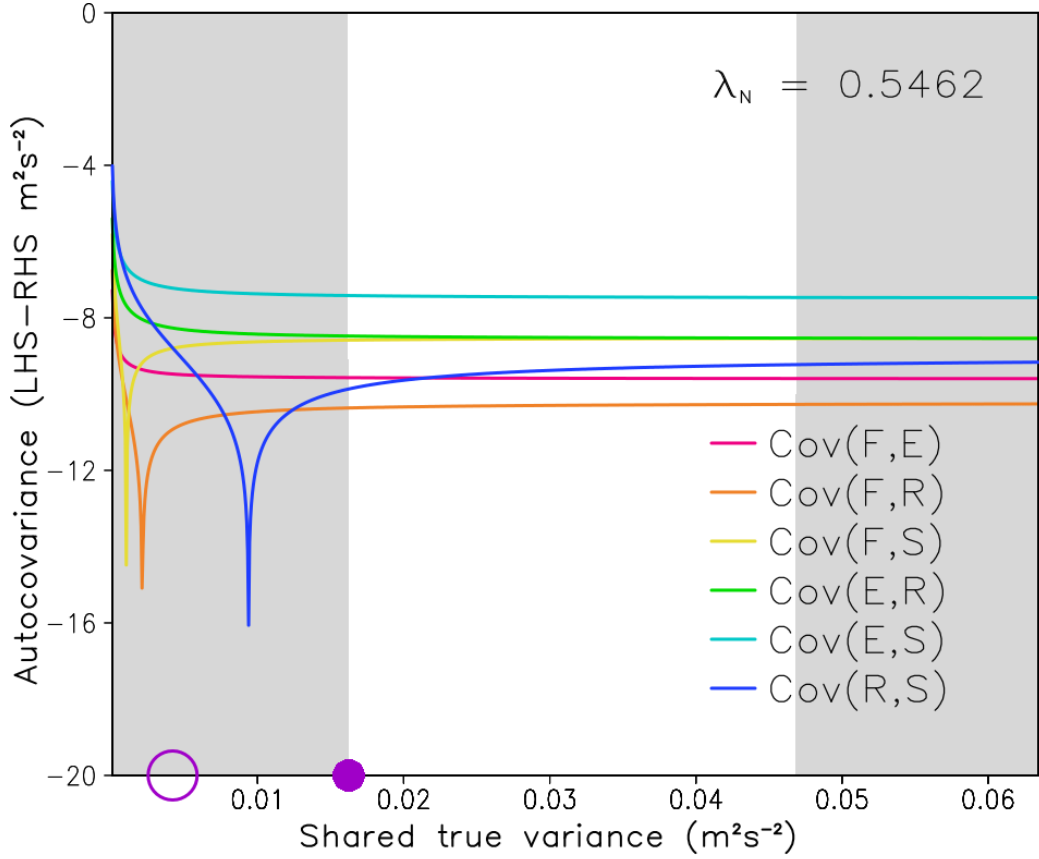


Figure 6: First demonstration of an INFERS parameter solution by weakly constrained minimization of the magnitude of differences between the LHS and RHS of the autocovariance equations (7) for the 5310226 non-outlier collocations from even years between 1993 and 2015 (roughly half of Fig. 3). The abscissa is true variance ( $\sigma_t^2$ ) in  $\text{m}^2\text{s}^{-2}$  between zero and  $\text{Var}(I)$ . The ordinate is log of absolute difference (LHS minus RHS). Grey shading denotes regions of negative error variance retrieval. Included are the target minimum (open purple circle at the average of three local minima) and the chosen minimum on the unshaded region (closed purple circle). The GlobCurrent-drifter shared error fraction ( $\lambda_N$ ) at the chosen minimum is also shown.

396 tended forecast  $E$ ). Although this target minimum is not accessible (on the  
 397 unshaded region), the chosen true variance solution is just to the right of  
 398 this locus of three minima and about the same distance from them as they  
 399 are from each other. This choice implies that at least one model variance  
 400 estimate is zero. Here, shading on the left in Fig. 6 corresponds to negative  
 401 shared true variance of the meridional current component (this is a derived  
 402 quantity that varies with  $\lambda_N$ ).

403 Whereas target solutions on the unshaded region can be seen as a re-  
 404 minder that models like (1), (2), and (5) are parsimonious (Box, 1979), the  
 405 tendency of autocovariance minima to be found on the left side of Fig. 6  
 406 may be the most important aspect of accommodating error cross-correlation.  
 407 This first demonstration indicates that true variance shared by GlobCurrent  
 408 and drifters is as small as possible (given that retrieved variance should be  
 409 positive). Visually, true and drifter error variance are the abscissa lengths  
 410 to the left and right of the closed purple dot, respectively. True variance is  
 411 thus smaller than drifter error variance when all collocations are considered.

412 Table 1 provides model parameters for the drifter (in situ) and GlobCur-  
 413 rent (nowcast) zonal ( $U$ ) and meridional ( $V$ ) current components. We find  
 414 that truth and error are of similar magnitude and that GlobCurrent and  
 415 drifters sample not only a shared truth but also shared error. However, this  
 416 truth exists only in the zonal component ( $0.127 \text{ ms}^{-1}$ ). Negligible merid-  
 417 ional amplitude ( $0.003 \text{ ms}^{-1}$ ) corresponds with a solution at the border of  
 418 the shaded region in Fig. 6. The additive calibration of GlobCurrent ( $\alpha_N$ ) is  
 419 also negligible and multiplicative calibration ( $\beta_N$ ) is prescribed by variance  
 420 matching (Fig. 5). Evidently, GlobCurrent samples are within their auto-

Table 1: Model parameters of the drifter ( $I$ ) and GlobCurrent nowcast ( $N$ ) zonal (U) and meridional (V) current components that are retrieved using 5310226 non-outlier collocations from the even years between 1993 and 2015 (cf. Fig. 3). Parameters include total standard deviation ( $\sigma$ ), true standard deviation ( $\sigma_t$ ), nowcast additive calibration ( $\alpha_N$ ), multiplicative calibration ( $\beta_N$ ), shared error fraction ( $\lambda_N$ ), individual ( $[1 - \lambda_N]^{1/2}\sigma_I$  and  $\sigma_N$ ) and total ( $\sigma_I$  and  $[\lambda_N^2\sigma_I^2 + \sigma_N^2]^{1/2}$ ) error standard deviation as in (6), signal correlation (McColl et al., 2014), and signal to noise ratio (SNR; Gruber et al. 2016b). Standard deviation and additive calibration are given in  $\text{ms}^{-1}$  and SNR is in dB.

	$\sigma$	$\sigma_t$	$\alpha_N$	$\beta_N$	$\lambda_N$	$\sigma_{indiv}$	$\sigma_{total}$	Corr	SNR
$U_I$	0.195	U: 0.127				0.100	0.148	0.652	-1.3
$V_I$	0.159					0.107	0.159	0.021	-33.6
$U_N$	0.168	V: 0.003	-0.001	0.843	0.546	0.100	0.129	0.640	-1.6
$V_N$	0.130		0.001			0.097	0.130	0.022	-33.3

421 covariance envelope as the minimum correlation for this sample is 0.91 and  
422 0.83 for the zonal and meridional current components, respectively.

423 We obtain most of the individual error terms in (5) and (6) from the model  
424 retrievals of unshared (measurement error) variance (i.e.,  $\sigma_N^2, \sigma_F^2, \sigma_E^2, \sigma_R^2$ , and  
425  $\sigma_S^2$ ). The exception is individual error for drifters ( $[1 - \lambda_N]\epsilon_I$ ), which follows  
426 from our definition of shared equation error (Kipnis et al., 1999). Diagnostic  
427 equations for shared and unshared drifter error variance can be written as  
428  $\lambda_N\sigma_I^2$  and  $(1 - \lambda_N)\sigma_I^2$ , respectively (i.e., assuming an even split of the covari-  
429 ance between equation error  $\lambda_N\epsilon_I$  and individual error  $[1 - \lambda_N]\epsilon_I$ ). Because  
430 over 50% of drifter error is shared by GlobCurrent ( $\lambda_N$ ), the percentage of  
431 total variance in (6) that is shared equation error ranges from 23% (GlobCur-

432 rent zonal component) to 55% (drifter meridional component).

433 Individual and total error variance for the zonal and meridional compo-  
434 nents are both high (Table 1). Calibration by variance matching dictates  
435 that drifter and GlobCurrent correlation with truth (McColl et al., 2014)  
436 and SNR (Gruber et al., 2016b) are roughly the same by zonal or merid-  
437 ional component (Su et al., 2014). Meridional noise dominates signal (SNR  
438 is -33dB) and even zonal noise is larger than signal (SNR < 0). Note that  
439 SNR is calculated using total error (i.e., both correlated and uncorrelated;  
440 third column from the right in Table 1). A preliminary regional assessment  
441 (not shown; GlobCurrent project document 2017) is consistent with previous  
442 studies (Johnson et al., 2007; Blockley et al., 2012; Sudre et al., 2013) in  
443 highlighting that weak meridional SNR is a characteristic of the equatorial  
444 regions.

445 Figure 7 is a second demonstration that true variance shared by GlobCur-  
446 rent and drifters is small. Parameters are retrieved as a function of day of  
447 the year, and to isolate one high latitude seasonal cycle, collocations north of  
448 15°N latitude are selected. We focus on 2385232 collocations of this northern  
449 region from even years between 1993 and 2015 (i.e., 21% of those available,  
450 using about 6000 collocations per day and applying variance matching and  
451 outlier removal at daily intervals). Figure 7 depicts solutions of true variance  
452 for an arbitrary selection of 12 days, of which eight are consistent with Fig. 6  
453 insofar as the locus of autocovariance minima (7) are at exceedingly small  
454 true variance. Only on day 240 (Fig. 7h) is true variance relatively large (as  
455 dictated by covariance involving F). An examination of all 364 days reveals  
456 a similar result: true variance is as small as possible on 250 of 339 days

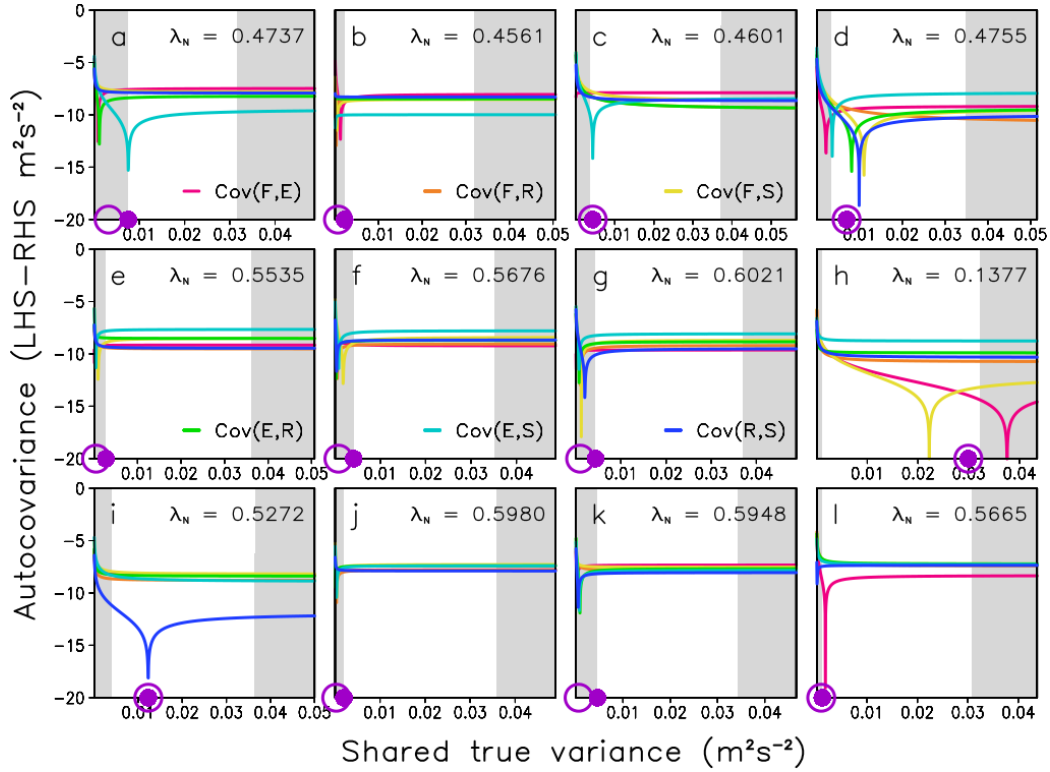


Figure 7: As in Fig. 6, but only for collocations north of 15°N on day a) 30, b) 60, c) 90, d) 120, e) 150, f) 180, g) 210, h) 240, i) 270, j) 300, k) 330, and l) 360 of the year for even years between 1993 and 2015.

457 (74%). No parameters are estimated on 25 of 364 days (7%) because no  
 458 autocovariance minima are found.

459 Figure 8 depicts the Northern Hemisphere seasonal cycle by five-day run-  
 460 ning means for the full set of INFERS model parameters. There is an annual  
 461 variation in the calibration and shared error parameters (c,d) that can be  
 462 explained by (e,j) GlobCurrent and drifter variations being slightly more  
 463 similar in amplitude toward the end of the year than at the beginning (e.g.,  
 464 solid lines tend to bracket the annual-average dashed lines in March and to

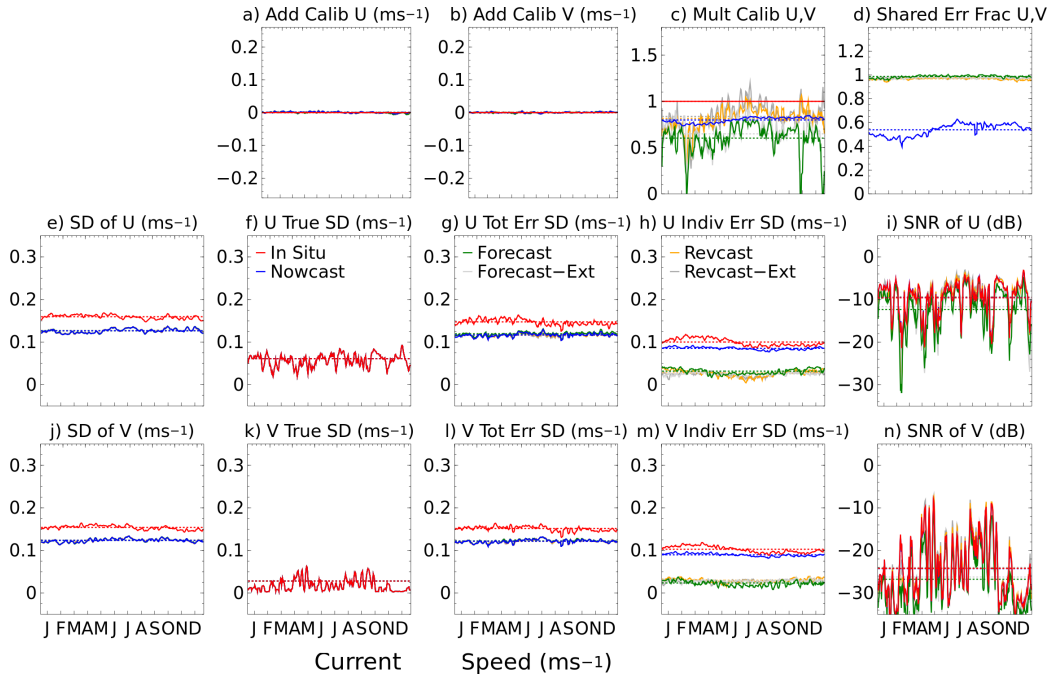


Figure 8: Retrieved model parameters as in Table 1, but for 339 days of the year using about 6000 collocations per day from north of  $15^\circ\text{N}$  and from even years between 1993 and 2015. Shown are the drifter (in situ/red) and GlobCurrent (nowcast/blue, forecast/green, revcast/orange, and extended forecast/light grey and revcast/dark grey) retrievals of a) zonal and b) meridional additive calibration ( $\text{ms}^{-1}$ ) and c) multiplicative calibration and d) shared error fraction for both zonal and meridional components, and e,j) 15-m current, f,k) shared truth, g,l) total error, and h,m) individual error standard deviation ( $\text{ms}^{-1}$ ), and i,n) signal to noise ratio (dB) for the zonal and meridional components, respectively. Solid lines are averages over five days and dashed lines are annual averages.

465 be bracketed by them in September). Of course, this similarity is largely su-  
466 perfcial, based on a consistent retrieval throughout the year of small shared  
467 truth in the zonal component (f; Fig. 7), and as in Table 1, almost no signal  
468 in the meridional component (k).

469 Drifter noise in Fig. 8 appears to be greater during spring than fall  
470 whereas GlobCurrent signal (via seasonality in multiplicative calibration)  
471 is the opposite. As a result, signal to noise ratio is higher for both GlobCur-  
472 rent and drifters in late summer compared to spring, even for the meridional  
473 current (despite its weak signal). A spatiotemporal refinement of this re-  
474 sult (with specific attention to the role of mixed layer depth) seems to be  
475 required. This same seasonality in SNR is obtained for the forecast and  
476 reicast samples, although via a different allocation of variance (i.e., with to-  
477 tal error being almost entirely defined by the GlobCurrent nowcast error).  
478 The range in multiplicative calibration (c) for the forecast and reicast data  
479 is an a posteriori justification for retaining separate calibrations in (5). All  
480 NFERS GlobCurrent samples again appear to be within their autocovariance  
481 envelope, as the minimum correlation among all days of the year is 0.88 and  
482 0.84 for the zonal and meridional current components, respectively.

483 Figure 9 is the third demonstration that true variance shared by GlobCur-  
484 rent and drifters is small. For a diagnosis of model parameters as a function  
485 of drifter current speed, we again apply variance matching and outlier re-  
486 moval (Hubert et al., 2012) as above, but to small groups of collocations.  
487 Tolman (1998) demonstrates that fine bin resolution (with sample sizes of  
488  $O[100]$ ) is useful to avoid bias in covariance estimates. Moreover, Zwieback  
489 et al. (2012) recommend at least 500 collocations based on idealized triple

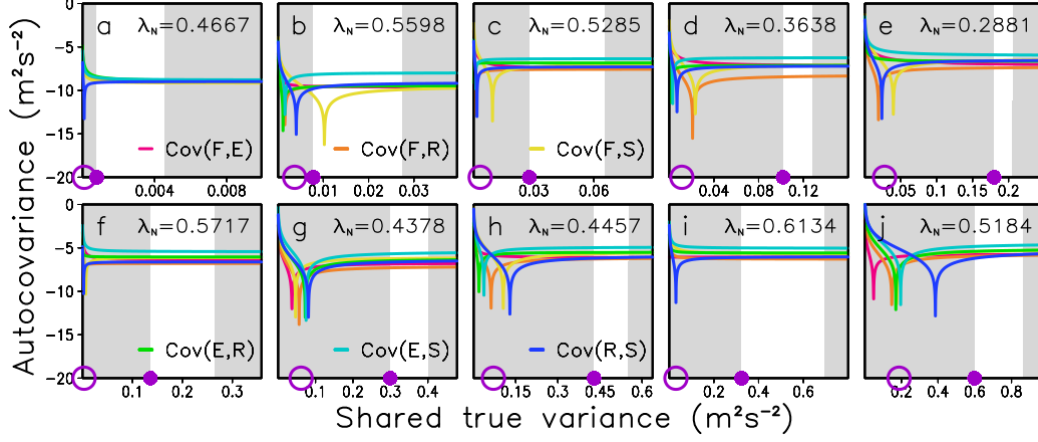


Figure 9: As in Fig. 6, but for subsets of 500 collocations whose drifter speed is nearest to a)  $0.1 \text{ ms}^{-1}$ , b)  $0.2 \text{ ms}^{-1}$ , c)  $0.3 \text{ ms}^{-1}$ , d)  $0.4 \text{ ms}^{-1}$ , e)  $0.5 \text{ ms}^{-1}$ , f)  $0.6 \text{ ms}^{-1}$ , g)  $0.7 \text{ ms}^{-1}$ , h)  $0.8 \text{ ms}^{-1}$ , i)  $0.9 \text{ ms}^{-1}$ , and j)  $1.0 \text{ ms}^{-1}$ . Note that abscissa range varies with current speed.

490 collocation simulations. Solutions of true variance are thus obtained over a  
 491 finely resolved ( $0.01\text{-ms}^{-1}$ ) range in drifter speed using 500 collocations clos-  
 492 est to each of 101 target speeds. (This sampling requires less than 1% of the  
 493 available collocations.) Individual panels in Fig. 9 are again consistent with  
 494 Fig. 6 in that all 10 loci of autocovariance minima (7) are at exceedingly  
 495 small true variance. An examination of the 101 speed bins reveals that true  
 496 variance is as small as possible for 90 of 92 bins (98%) and no parameters  
 497 are estimated for 9 of 101 bins (9%) because no autocovariance minima are  
 498 found.

499 Figure 10 illustrates the dependence of model parameters on current  
 500 speed. There are weak trends in the calibration and shared error parameters  
 501 (a-d) and strong trends in most variance parameters (e-n). As in Table 1,

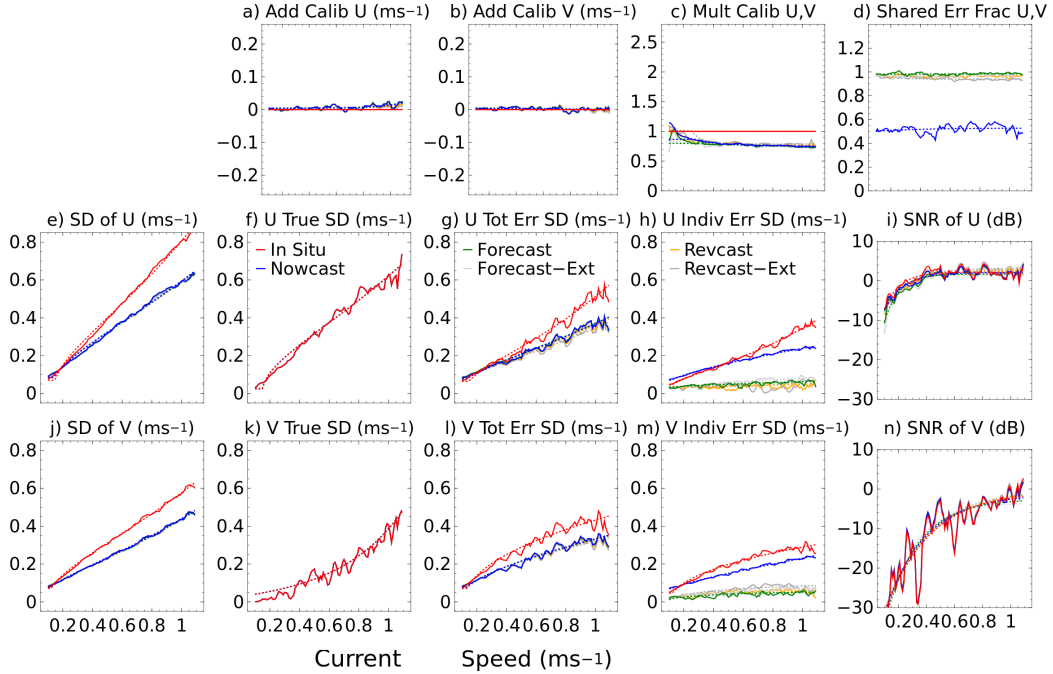


Figure 10: Model parameters as in Fig. 8, but for 92 subsets of 500 collocations whose drifter speed is nearest to target values between  $0.1 \text{ ms}^{-1}$  and  $1.1 \text{ ms}^{-1}$  at intervals of  $0.01 \text{ ms}^{-1}$  (excluding 9 solutions for which no autocovariance minima were found). Solid lines are averages of five adjacent intervals. Dashed lines are best fits of the form  $y(x) = a + be^{cx}$  (Jacquelin, 2014), but for c) multiplicative calibration, this fit ignores target values less than  $0.3 \text{ ms}^{-1}$ .

502 GlobCurrent-drifter shared error fraction ( $\lambda_N \approx 0.5$ ) is quite high, variance-  
503 matched multiplicative calibration ( $\beta_N$ ) is about 0.85 beyond  $0.3 \text{ ms}^{-1}$ , and  
504 additive calibration of GlobCurrent ( $\alpha_N$ ) is negligible. Justification for our  
505 application of variance matching throughout this study (rather than assum-  
506 ing no GlobCurrent bias) is that an upper bound on multiplicative bias, as  
507 given by reverse linear regression, falls below one at large current speed (not  
508 shown). In turn, the need to address strong current underestimation (per-

509 haps locally in time and space, but at the resolution of the GlobCurrent  
510 analysis) may continue to exist (cf. Rio et al. 2014).

511 Errors in GlobCurrent samples separated by a day are basically the same  
512 in Fig. 10g,h,l,m. The product of the forecast and reforecast shared error  
513 fraction parameters ( $\lambda_F \lambda_E \lambda_R \lambda_S$ ) is thus close to unity, which implies that  
514 GlobCurrent error is being sampled within its autocovariance envelope. In  
515 effect, this justifies the use of the extended forecast and reforecast samples in  
516 the INFERS model. Among all 92 subsets, the minimum correlation of com-  
517 bined truth and error (found at low speed between  $E$  and  $S$ ) is 0.84 and 0.76  
518 for the zonal and meridional current components, respectively.

519 Figure 10f,k reveals weak agreement between GlobCurrent and drifters on  
520 a shared truth at low current speed, but more agreement at higher current  
521 speed. This is dictated in part by current speed itself (Fig. 10e,j), but the  
522 meridional component of drifter error increases quickly with current speed  
523 (more so than the zonal component) and the opposite is the case for true  
524 variance. In contrast to negative SNR for the zonal component in Table 1,  
525 the GlobCurrent/drifter best fit SNR (Fig. 10i dashed lines; equivalent by  
526 variance matching) eventually exceed, but remain close to, 0 dB from about  
527  $0.3 \text{ ms}^{-1}$ .

528 This section constitutes an introduction to the INFERS model featuring  
529 hundreds of parameter solutions. Our experiments are thus enabled by access  
530 to millions of drifter current estimates and a GlobCurrent analysis that is  
531 about three orders of magnitude larger. This is not to say that 500 collocations  
532 is small. In many contexts, including ours, a few hundred collocations  
533 may be ample. However, with the freedom afforded by large datasets to

534 identify a range of solutions using appropriate instruments (cf. Kipnis et al.  
535 2002), comes the opportunity to better characterize shared truth and error.  
536 The next section briefly explores shared truth as an updated measure of  
537 agreement between variates and clarifies shared error as an updated measure  
538 of dependence.

## 539 **5. Discussion**

540 It is sometimes the case in geophysics that only one truth (a so-called gen-  
541 uine truth) is of interest. Implicit in this concept is the idea that truth carries  
542 no information about particular datasets, which differ only in terms of their  
543 corresponding error, and this error is intrinsic (i.e., defined without reference  
544 to another dataset). Implicit in the definition of shared truth, on the other  
545 hand, is the idea that *if shared truth exists, then it contains information about*  
546 *an overlap in data supports* (see Appendix). Beyond the scope of this paper,  
547 but notable within geophysics, are formal inference theories that concern a  
548 conjunction of information and the problem of aggregated opinion (Taran-  
549 tola, 2005). Here, it suffices to note that measurement models can provide  
550 a calibration by linear mapping, and a validation by shared/unshared error,  
551 but they can also provide a useful measure of agreement among datasets by  
552 shared truth.

553 One documented application of shared truth is an assessment by Bentamy  
554 et al. (2017) of various global ocean surface heat flux analyses. Using the  
555 INFERS model, Bentamy et al. experiment with shared truth as a metric  
556 of competitive validation (see Appendix). Following a recalibration of each  
557 gridded analysis to the same in situ reference, they observe that in situ and



568 full range of solutions consistent with (1), with familiar analytic solutions  
569 for ordinary (OLR) and reverse (RLR) linear regression being appropriate  
570 references. Solutions of the OLR and RLR models are identified by the  
571 method of moments with either drifter error ( $\epsilon_I$  for OLR) or GlobCurrent  
572 error ( $\epsilon_A$  for RLR) set to zero. INFERS estimates of truth and error from  
573 the previous section are placed alongside these two reference solutions in  
574 Fig. 11. It is notable that INFERS solutions of true standard deviation  
575 (Fig. 11a,b,e,f) are smallest. This is remarkable because the OLR and RLR  
576 references are understood to be the solutions that bound the range of true  
577 variance (and multiplicative calibration or regression slope) values that are  
578 consistent with the errors-in-variables model (1).

579 Further comparison between INFERS and the corresponding OLR and  
580 RLR reference solutions permit an interpretation of the unshared (measure-  
581 ment) errors that define much of the total error in this study. Figure 11c,d,g,h  
582 reveals that the magnitude of OLR error in GlobCurrent and RLR error in  
583 drifters appear to differ little from the unshared error shown in Fig. 8h,m  
584 and Fig. 10h,m. As noted in Section 4, some ambiguity is expected in a di-  
585 agnostic estimate of drifter unshared error, but the overlapping agreement in  
586 GlobCurrent unshared error (i.e., black dashed and blue lines) is evident for  
587 all collocation divisions. Whereas OLR and RLR impose separate assump-  
588 tions on (1) that provide hypothetical bounds on uncorrelated error, in this  
589 study a single model seems to provide both solutions.

590 Figure 11 reveals that total error in GlobCurrent and drifters can be  
591 interpreted as a combination of respective RLR and OLR upper limits in  
592 uncorrelated error. Subject to the caveat that a fixed calibration by variance

593 matching allows more freedom for shared error in our INFERS solutions, the  
594 reason that shared true variance falls outside the OLR and RLR bounding  
595 reference solutions is because not only can the INFERS model accommo-  
596 date bounds on unshared (measurement) error, as given by (1), but shared  
597 (equation) error is accommodated as well.

598 We conclude this initial characterization of model solutions by noting that  
599 shared error offers an updated measure of error dependence. It is important  
600 to recognize that any decision to exclude shared error from a measurement  
601 model, based on physical knowledge of the data alone, can always be chal-  
602 lenged. In other words, *even if there is no apparent physical relationship*  
603 *between two datasets, independence of their errors should not be presumed*  
604 *without considering that the measurement model is only an approximation*  
605 (Box, 1979). Thus, it may be appropriate to accommodate sharing even  
606 if one cannot assume that shared error (or truth) exists. More specifically  
607 (Fuller, 1987), if the model assumes that truth and error are additive with  
608 a linearly related signal, as in (1), and this might not be strictly true of  
609 the data, then some form of both equation and measurement error (2) or  
610 correlated and uncorrelated error (5) should be included.

611 Equation error and correlated error are considered to be essentially the  
612 same in this study, as we now demonstrate, but they are not strictly the same  
613 error in general. For instance, Kipnis et al. (1999) allow for correlation in  
614 both equation error and measurement error. The Introduction acknowledges  
615 that GlobCurrent and drifters also may share a component of measurement  
616 error. This is because many of the same drifters that are employed to re-  
617 fine the CNES-CLS13 mean dynamic topography (MDT; Rio and Hernandez

618 2004; Rio et al. 2014) are employed above for validation. Although INFERS  
619 provides an estimate of error correlation that may include measurement error,  
620 one option for demonstrating its interpretation as equation error is a valida-  
621 tion only after 2013. Instead, we opt to replace the CNES-CLS13 MDT in  
622 each GlobCurrent sample (NFERS) with a more approximate GOCE-only  
623 MDT (Rio et al., 2014). Drifter measurement error is thus removed from  
624 GlobCurrent and the remaining error correlation can be attributed entirely  
625 to equation error.

Table 2: As in Table 1, but for a measurement-error-independent comparison between GlobCurrent and drifters: GlobCurrent data exclude a velocity component associated with the CNES/CLS-2013 MDT and include instead a component associated with the GOCE-only geodetic MDT (Rio et al., 2014). Parameters of the drifter ( $I$ ) and GlobCurrent nowcast ( $N$ ) zonal (U) and meridional (V) current components are retrieved using 5280828 non-outlier collocations from the even years between 1993 and 2015.

	$\sigma$	$\sigma_t$	$\alpha_N$	$\beta_N$	$\lambda_N$	$\sigma_{indiv}$	$\sigma_{total}$	Corr	SNR
$U_I$	0.194	U:				0.106	0.153	0.612	-2.2
$V_I$	0.158	0.119				0.110	0.158	0.007	-43.7
$U_N$	0.161	V:	-0.001	0.818	0.517	0.101	0.128	0.604	-2.4
$V_N$	0.127	0.001	0.002			0.097	0.127	0.007	-43.5

626 Table 2 provides a comparison between GlobCurrent (GOCE-only MDT)  
627 and drifters based on 5280828 non-outlier collocations from the even years  
628 between 1993 and 2015. With some of the strongest current components (i.e.,  
629 most different in terms of MDT) again excluded as outliers, true standard de-  
630 viation in the zonal component decreases slightly ( $0.127 \text{ ms}^{-1}$  to  $0.119 \text{ ms}^{-1}$ )

631 for an MDT that lacks drifter information. Otherwise, the results of Ta-  
632 ble 1 are reproduced, including small true variance in the meridional com-  
633 ponent and a large shared error fraction ( $\lambda_N = 0.517$ ). Although this is a  
634 measurement-error-independent comparison, it is nevertheless clear that the  
635 two datasets are not independent. Shared error fraction in Table 1 is quite  
636 similar ( $\lambda_N = 0.546$ ), as is the percentage of total variance in (6) that is  
637 shared error, again ranging from 24% for the GlobCurrent zonal component  
638 to 52% for the drifter meridional component. The implication is that there  
639 is little error correlation owing to drifter measurement error in the CNES-  
640 CLS13 MDT. There is instead large error correlation owing to equation error.

## 641 **6. Conclusions**

642 This study provides an approach to the challenge of introducing and, like  
643 any other model term, identifying cross-correlated error in linear regression  
644 models such as (1). Subject to the caveat that calibration is prescribed by  
645 variance matching (rather than being jointly retrieved with shared true vari-  
646 ance), over 90% of all attempts to retrieve model parameters for GlobCurrent  
647 and drifters are successful. Perhaps the more surprising aspect is that, given  
648 two datasets, we require just a few additional samples of the GlobCurrent  
649 analysis around the time of each drifter observation. Compared to the fre-  
650 quency of these additional samples, necessary confirmation of slow changes  
651 in the evolution of GlobCurrent and its errors is also obtained.

652 Formulation of a new measurement model called INFERS (an acronym  
653 taken from data sample names) is inspired by instrumental variable regres-  
654 sion (Su et al., 2014) and specifically the triple collocation approach (Stof-

655 felen, 1998; Caires and Sterl, 2003; Janssen et al., 2007; O’Carroll et al.,  
656 2008; Vogelzang et al., 2011; Zwieback et al., 2012; McColl et al., 2014; Yil-  
657 maz and Crow, 2014; Gruber et al., 2016b). Error propagation through the  
658 data samples is modelled using a first-order autoregressive (AR-1) formula,  
659 except that propagation begins with the collocated sample equations (IN),  
660 which provide the cross-correlated error terms, and then includes a tempo-  
661 rally symmetric application of AR-1 to error autocorrelation in the remaining  
662 equations (FERS). The most direct model comparison is to solutions of the  
663 linear errors-in-variables regression model (1) because this is the same model  
664 given by the collocated sample equations (IN) if cross-correlated errors are  
665 ignored. A search for true variance in a limited parameter space of the IN-  
666 FERS model (i.e., assuming the variance matching calibration) yields values  
667 smaller than for any solution of (1), as given by ordinary (OLR) and reverse  
668 (RLR) linear regression bounds. Over three quarters of these model solu-  
669 tions (Fig. 11) support the proposition that truth and signal, as defined in  
670 the INFERS model, are small (see also Table 2 of Bentamy et al. 2017).

671 If truth is considered a shared model variable just like error (ignoring  
672 its unshared component), then shared true variance can be considered a  
673 measure of agreement between GlobCurrent and drifters. Inferences about  
674 measurement model approximations as well as overlaps in data support are  
675 then possible. While it would be unfortunate to start with a true variance  
676 that is smaller than it actually is (i.e., the variance matching calibration  
677 may yield such a bias), to start with a truth that is larger than it actually is  
678 would likely be more worrisome. This study indicates that there is a potential  
679 to overstate the agreement between GlobCurrent and drifters based on an

680 inflated true variance in the linear errors-in-variables model. Like the triple  
681 collocation model, OLR and RLR are just identified and necessarily lack  
682 a term for cross-correlated error. Because their solutions involve variance  
683 budgets with fixed total variance, as in the LHS of (6), if total error is  
684 increased by introducing a new error term (equation or correlated error),  
685 then true variance decreases by the same amount. Tables 1 and 2 reveal that  
686 roughly a quarter to a half of the total variance in GlobCurrent and drifters  
687 is shared error variance. Presumably, shared error is a first order term that  
688 could not be much larger and remain hidden. Subsequent studies are needed  
689 to confirm whether this masquerading of equation error as truth is common  
690 for other datasets and whether it should be attributed to limitations in the  
691 errors-in-variables model. However, this should *exclude prescribed calibration*  
692 and instead explore solutions in the full parameter space of the INFERS  
693 model.

694 Implications of measurement model assumptions (e.g., that truth and er-  
695 ror are additive with a linearly related signal) are discussed in geophysics  
696 (e.g., Janssen et al. 2007; Zwieback et al. 2016), and moreso in the statistical  
697 literature, where notions are established regarding how to accommodate non-  
698 linear signals in linear regression by including equation error (Fuller, 1987;  
699 Carroll and Ruppert, 1996). Furthermore, accommodation of equation er-  
700 ror and measurement error *correlation* is given in sophisticated measurement  
701 models in epidemiology (Kipnis et al., 1999, 2002). In turn, it appears that  
702 the opportunity to simultaneously identify all parameters of such models can  
703 be taken up in part by studies like this one that incorporate an experimental  
704 sampling of large datasets.

705 A sufficient number of GlobCurrent samples is taken before and after  
706 each collocation (as persistence forecasts and revcasts, respectively) so that  
707 there are more covariance equations than model parameters. Retrieval of the  
708 17 INFERS model parameters employs variance matching to first prescribe  
709 the calibration from GlobCurrent to drifters. Six autocovariance equations,  
710 involving the FERS samples, weakly constrain shared true variance and the  
711 remaining 15 covariance equations are a strong constraint on the remaining  
712 15 unknown parameters. Insofar as true variance is weakly constrained, this  
713 study avoids a common assumption that real data be cast in the form of a  
714 simple measurement model.

715 Model solutions have been examined for collocation groups numbering  
716 about six million (from eleven years), 6000 (on each day of the year in the  
717 NH), and 500 (nearest drifter speeds at  $0.01\text{-ms}^{-1}$  intervals). One must be  
718 cautious about groups of collocations both large (if in situ error is autocor-  
719 related) and small (if parameter retrievals depend on individual collocations;  
720 cf. Zwieback et al. 2012). However, for all these subsets, SNR is near zero  
721 at best because the error in GlobCurrent and drifters is high, while variance  
722 of the true current is low. There are indications that the preferentially low  
723 SNR of the meridional component is a characteristic of equatorial regions  
724 (cf. Johnson et al. 2007; Blockley et al. 2012; Sudre et al. 2013). The inter-  
725 pretation of large individual error is also interesting in that the OLR and  
726 RLR reference bounds on uncorrelated error are reached by both GlobCur-  
727 rent and drifters.

728 The last experiment of the Discussion is perhaps the most relevant for an  
729 interpretation of shared and unshared error in terms of equation and mea-

730 surement error, respectively. A measurement-error-independent comparison  
731 between GlobCurrent (using a GOCE-only MDT) and drifters permits a di-  
732 agnosis of just how large the correlation in equation error may be. There  
733 is little change in shared error fraction between the two MDT experiments,  
734 which suggests that correlated error in other comparisons of this study may  
735 be viewed as predominantly that of equation error rather than measurement  
736 error (in spite of a drifter error contribution to the CNES/CLS13 MDT).  
737 Good correspondence between equation error and correlated error provides  
738 further impetus for a review of common model assumptions.

739 The so-called genuine truth is not viewed in this study as the same true  
740 variable  $t$  that appears in most measurement models. The search for a genu-  
741 ine ocean surface current is ongoing, however, and iterative or comparative  
742 applications of a measurement model have a role to play (e.g., Bentamy et al.  
743 2017). By analogy with efforts to validate SST, surface current depth should  
744 be useful to distinguish between a slower, quasi-balanced flow and interac-  
745 tions with the atmosphere. For example, both drifters and GlobCurrent may  
746 be good references for balanced flow experiments at the equator (cf. Chan  
747 and Shepherd 2014) and at higher latitudes (cf. Penven et al. 2014). High  
748 resolution analyses are expected to grow in number, and while validation is  
749 not a prescription for finding the genuine current, there is an opportunity  
750 to quantify improvements in two or more datasets (or versions of a single  
751 dataset) against one chosen reference dataset. This study documents varia-  
752 tions in INFERS model parameters as a function of day of the year and cur-  
753 rent speed, but a high latitude flow experiment may benefit from distinctions  
754 between cyclonic and anticyclonic eddies, whereas an equatorial experiment

755 may opt to treat the zonal and meridional components separately. With a  
756 view to mapping model parameters in the dimensions of large datasets, an  
757 important challenge involves selecting subsets of collocations according to an  
758 informed physical understanding.

759 This study is a contribution to efforts of the geophysical community to  
760 construct high resolution ocean surface current analyses using assimilative  
761 numerical models and a synergy of observations (this issue). Because obser-  
762 vational coverage is sparse, especially over the ocean and in early years, a  
763 topical question remains whether to withhold reference observations from an  
764 analysis so as to later perform an independent validation. To respond to this  
765 question in the negative would imply that the same observations should be  
766 allowed to benefit both the construction of an analysis and its validation. In  
767 turn, shared signal and noise in observations and analyses need to be consid-  
768 ered and measurement models that accommodate both equation error and  
769 measurement error are called for (cf. Caires and Sterl 2003; Gruber et al.  
770 2016b). It appears that not only can a basis for understanding shared signal  
771 and noise be found in literature, but a year-on-year accumulation of geophys-  
772 ical observations and high resolution data is permitting more freedom, and  
773 slightly less parsimony, in experimental measurement modelling.

## 774 **7. Acknowledgements**

775 We are pleased to acknowledge an international effort over many years  
776 to collect, assemble, and analyze altimetric and drifter observations, as well  
777 as the support and discussions (commencing roughly in reverse time) with  
778 Bash Toulany, Will Perrie, Graham Dunn, Tim Williams, Igor Esau, Lau-

779 rent Bertino, Ad Stoffelen, Abderrahim Bentamy, Svetla Hristova-Veleva,  
780 Bryan Stiles, Zorana Jelenak, Mike Brennan, Luc Fillion, Bridget Thomas,  
781 Hal Ritchie, and Mike Dowd. The opportunity to reflect on the comments  
782 of three reviewers has been invaluable in promoting a clearer presentation.  
783 Funding for this work (again in reverse time) was from the European Space  
784 Agency via the Nansen Center and Ifremer (Data User Element’s GlobCur-  
785 rent and Support to Science Element’s Ocean Heat Flux projects, respec-  
786 tively). The first author was also supported by the U.S. National Aeronautics  
787 and Space Administration via the National Oceanic and Atmospheric Admin-  
788 istration and University Corporation for Atmospheric Research (Hurricane  
789 Science Research, Ocean Vector Winds, and Visiting Scientist programs, re-  
790 spectively), and the Canadian Space Agency via Environment and Climate  
791 Change Canada (Government Related Initiatives program).

## 792 **8. References**

- 793 Bentamy, A., Piollé, J.-F., Grouazel, A., Danielson, R. E., Gulev, S. K.,  
794 Paul, F., Azelmat, H., Mathieu, P.-P., von Schuckmann, K., Sathyen-  
795 dranath, S., Evers-King, H., Esau, I., Johannessen, J. A., Clayson, C. A.,  
796 Pinker, R. T., Grodsky, S. A., Bourassa, M., Smith, S. R., Haines, K.,  
797 Valdivieso, M., Merchant, C. J., Chapron, B., Anderson, A., Hollmann,  
798 R., Josey, S. A., 2017. Review and assessment of latent and sensible heat  
799 flux accuracy over global oceans. *Remote Sens. Environ.* 201, 196–218,  
800 doi:10.1016/j.rse.2017.08.016.
- 801 Blockley, E. W., Martin, M. J., Hyder, P., 2012. Validation of FOAM near-

- 802 surface ocean current forecasts using lagrangian drifting buoys. *Ocean Sci.*  
803 8, 551–565, doi:10.5194/os-8-551-2012.
- 804 Box, G. E. P., 1979. Robustness in the strategy of scientific model building,  
805 *Robustness in Statistics*, Academic Press, pp. 201–236 (accessed 4 August  
806 2017 at <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA070213>).
- 807 Caires, S., Sterl, A., 2003. Validation of ocean wind and wave data using  
808 triple collocation. *J. Geophys. Res.* 108, doi:10.1029/2002JC001491.
- 809 Carroll, R. J., Ruppert, D., 1996. The use and misuse of orthogonal regression  
810 in linear errors-in-variables models. *The American Statistician* 50, 1–6.
- 811 Chan, I. H., Shepherd, T. G., 2014. Diabatic balance model for the equatorial  
812 atmosphere. *J. Atmos. Sci.* 71, 985–1001, doi:10.1175/JAS-D-13-0224.1.
- 813 Danielson, R. E., 2017. Collocations of surface drifter and GlobCurrent (cen-  
814 tered five-day) 15-m zonal and meridional velocity estimates. SEANOE  
815 online data archive at <http://doi.org/10.17882/50225>.
- 816 Danielson, R. E., Johannessen, J. A., Rio, M.-H., Collard, F., Don-  
817 lon, C., Chapron, B., Quartly, G., 2017. An update on GlobCur-  
818 rent calibration and validation activities, poster presentation  
819 at the third GlobCurrent User Consultation Meeting (UCM-3)  
820 in Frascati, Italy on 21-23 March 2017 (accessed May 2017 at  
821 [https://www.researchgate.net/publication/315486597\\_Update\\_on\\_](https://www.researchgate.net/publication/315486597_Update_on_GlobCurrent_Calibration_and_Validation_Activities)  
822 [GlobCurrent\\_Calibration\\_and\\_Validation\\_Activities](https://www.researchgate.net/publication/315486597_Update_on_GlobCurrent_Calibration_and_Validation_Activities)).
- 823 Fuller, W. A., 1987. *Measurement Error Models*. Wiley, New York.

- 824 Fuller, W. A., 2006. Errors in variables. *Encyclopedia of Statistical Sciences*,  
825 S. Kotz, C. B. Read, N. Balakrishnan, B. Vidakovic and N. L. Johnson,  
826 Eds., doi:10.1002/0471667196.ess1036.pub2.
- 827 Gillard, J., Iles, T., 2005. Variance covariance matrices for lin-  
828 ear regression with errors in both variables. Cardiff School  
829 of Mathematics Technical Report (accessed August 2017 at  
830 [http://mathsdemo.cf.ac.uk/math/resources/Iles\\_Gillard\\_Tech\\_Report.pdf](http://mathsdemo.cf.ac.uk/math/resources/Iles_Gillard_Tech_Report.pdf)).
- 831 GlobCurrent project document, 2017. Fifth validation report (VAL-5).
- 832 Gruber, A., Su, C.-H., Crow, W. T., Zwieback, S., Dorigo, W. A., Wagner,  
833 W., 2016a. Estimating error cross-correlations in soil moisture data sets  
834 using extended collocation analysis. *J. Geophys. Res. Atmos.* 121, 1208–  
835 1219, doi:10.1002/2015JD024027.
- 836 Gruber, A., Su, C.-H., Zwieback, S., Crow, W. T., Dorigo, W. A.,  
837 Wagner, W., 2016b. Recent advances in (soil moisture) triple col-  
838 location analysis. *Int. J. Appl. Earth Obs. Geoinform.* 45, 200–211,  
839 doi:10.1016/j.jag.2015.09.002.
- 840 Hansen, D., Poulain, P.-M., 1996. Quality control and interpolations of  
841 WOCE-TOGA drifter data. *J. Atmos. Oceanic Technol.* 13, 900–909,  
842 doi:10.1175/1520-0426.
- 843 Hubert, M., Rousseeuw, P. J., Verdonck, T., 2012. A deterministic algorithm  
844 for robust location and scatter. *J. Comp. Grap. Stats* 21, 618–637, doi:  
845 10.1080/10618600.2012.672100.

- 846 Jacquelin, J., 2014. Régressions et équations intégrales (accessed 8 july  
847 2016 at [https://www.researchgate.net/file.postfileloader.html?id=](https://www.researchgate.net/file.postfileloader.html?id=5549ee13f15bc7863c8b456d&assetkey=as%3a273771590553603%401442283641886)  
848 [5549ee13f15bc7863c8b456d&assetkey=as%3a273771590553603%](https://www.researchgate.net/file.postfileloader.html?id=5549ee13f15bc7863c8b456d&assetkey=as%3a273771590553603%401442283641886)  
849 [401442283641886](https://www.researchgate.net/file.postfileloader.html?id=5549ee13f15bc7863c8b456d&assetkey=as%3a273771590553603%401442283641886)).
- 850 Janssen, P. A. E. M., Abdalla, S., Hersbach, H., Bidlot, J.-R., 2007. Error es-  
851 timation of buoy, satellite, and model wave height data. *J. Atmos. Oceanic*  
852 *Technol.* 24, 1665–1677, doi:10.1175/JTECH2069.1.
- 853 Johnson, E. S., Bonjean, F., Lagerloef, G. S. E., Gunn, J. T., Mitchum,  
854 G. T., 2007. Validation and error analysis of oscar sea surface currents. *J.*  
855 *Atmos. Oceanic Technol.* 24, 688–701.
- 856 Kipnis, V., Carroll, R. J., Freedman, L. S., Li, L., 1999. A new dietary  
857 measurement error model and its application to the estimation of relative  
858 risk: Application to four validation studies. *Am. J. Epidemiol.* 150, 642–  
859 651, doi:10.1093/oxfordjournals.aje.a010063.
- 860 Kipnis, V., Midthune, D., Freedman, L. S., Bingham, S., Day, N. E., Riboli,  
861 E., Ferrari, P., Carroll, R. J., 2002. Bias in dietary-report instruments and  
862 its implications for nutritional epidemiology. *Public Health Nutr.* 5(6A),  
863 915–923, doi:10.1079/PHN2002383.
- 864 Lumpkin, R., Grodsky, S. A., Centurioni, L., Rio, M.-H., Carton,  
865 J. A., Lee, D., 2013. Removing spurious low-frequency variability in  
866 surface drifter velocities. *J. Atmos. Oceanic Technol.* 30, 353–360,  
867 <http://dx.doi.org/10.1175/JTECH-D-12-00139.1>.

- 868 McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M.,  
869 Stoffelen, A., 2014. Extended triple collocation: Estimating errors and  
870 correlation coefficients with respect to an unknown target. *Geophys. Res.*  
871 *Lett.* 41, 6229–6236, doi:10.1002/2014GL061322.
- 872 Moberg, A., Brattström, G., 2011. Prediction intervals for climate recon-  
873 structions with autocorrelated noise – an analysis of ordinary least squares  
874 and measurement error methods. *Palaeogeol., Palaeoclim., Palaeoecol.* 308,  
875 313–329, doi:10.1016/j.palaeo.2011.05.035.
- 876 O’Carroll, A. G., Eyre, J. R., Saunders, R. W., 2008. Three-way error analysis  
877 between AATSR, AMSR-E, and in situ sea surface temperature observa-  
878 tions. *J. Atmos. Oceanic Technol.* 25, 1197–1207.
- 879 Penven, P., Halo, I., Pous, S., Marié, L., 2014. Cyclogeostrophic bal-  
880 ance in the Mozambique Channel. *J. Geophys. Res. Oceans* 119,  
881 doi:10.1002/2013JC009528.
- 882 Rio, M.-H., 2012. Use of altimeter and wind data to detect the anomalous loss  
883 of SVP-type drifters drogue. *J. Atmos. Oceanic Technol.* 29, 1663–1674,  
884 doi:10.1175/JTECH-D-12-00008.1.
- 885 Rio, M.-H., Hernandez, F., 2004. A mean dynamic topography computed  
886 over the world ocean from altimetry, in situ measurements, and a geoid  
887 model. *J. Geophys. Res.* 109, doi:10.1029/2003JC002226.
- 888 Rio, M.-H., Mulet, S., Picot, N., 2014. Beyond GOCE for the ocean circu-  
889 lation estimate: Synergetic use of altimetry, gravimetry, and in situ data

890 provides new insight into geostrophic and Ekman currents. *Geophys. Res.*  
891 *Lett.* 41, doi:10.1002/2014GL061773.

892 Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling  
893 and calibration using triple collocation. *J. Geophys. Res.* 103, 7755–7766,  
894 doi:10.1029/97JC03180.

895 Su, C.-H., Ryu, D., Crow, W. T., Western, A. W., 2014. Beyond triple  
896 collocation: Applications to soil moisture monitoring. *J. Geophys. Res.*  
897 *Atmos.* 119, 6419–6439, doi:10.1002/2013JD021043.

898 Sudre, J., Maes, C., Garçon, V., 2013. On the global estimates of  
899 geostrophic and Ekman surface currents. *Limnol. Ocean.: Fluids Environ.*  
900 3, doi:10.1215/21573689-2071927.

901 Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Param-*  
902 *eter Estimation.* SIAM, Philadelphia, Pennsylvania.

903 Tolman, H. L., 1998. Effect of observation errors in linear regression  
904 and bin-average analyses. *Quart. J. Roy. Meteor. Soc.* 124, 897–917,  
905 doi:10.1002/qj.49712454712.

906 Vogelzang, J., Stoffelen, A., Verhoef, A., Figa-Saldaña, J., 2011. On the  
907 quality of high-resolution scatterometer winds. *J. Geophys. Res.* 116,  
908 doi:10.1029/2010JC006640.

909 Yilmaz, M. T., Crow, W. T., 2013. The optimality of potential rescal-  
910 ing approaches in land data assimilation. *J. Hydrometeor.* 14, 650–660,  
911 doi:10.1175/JHM-D-12-052.1.

912 Yilmaz, M. T., Crow, W. T., 2014. Evaluation of assumptions in soil  
913 moisture triple collocation analysis. *J. Hydrometeor.* 15, 1293–1302,  
914 doi:10.1175/JHM-D-13-0158.1.

915 Zwieback, S., Scipal, K., Dorigo, W., Wagner, W., 2012. Structural and  
916 statistical properties of the collocation technique for error characterization.  
917 *Nonlin. Process. Geophys.* 19, 69–80, doi:10.5194/npg-19-69-2012.

918 Zwieback, S., Su, C.-H., Gruber, A., Dorigo, W. A., Wagner, W., 2016. The  
919 impact of quadratic nonlinear relations between soil moisture products on  
920 uncertainty estimates from triple collocation analysis and two quadratic ex-  
921 tensions. *J. Hydrometeor.* 17, 1725–1743, doi:10.1175/JHM-D-15-0213.1.

## 922 **9. Appendix**

923 Measurement models (defined below) are actively evolving in various  
924 fields, with geophysical applications that may be unfamiliar or are just be-  
925 ginning to have an impact. The solution of such models is called an in-  
926 verse problem (Tarantola, 2005), by contrast with evolution equations for  
927 mass, motion, and constituents as a forward model. It should be noted that  
928 longstanding experience in the geo-physical/biological/chemical communities  
929 with forward modelling and with taking high resolution (so-called longitudi-  
930 nal) observations provide the basis for estimating error autocorrelation (e.g.,  
931 using FERS). A brief clarification of other concepts relevant to this study is  
932 offered here as a complement to more formal definitions. Online sources (e.g.,  
933 Wikipedia) also provide recent and useful collaborative summaries. Concepts  
934 relevant to this study include:

- 935 • Affine calibration: synonymous with a linear calibration by intercept  
 936 ( $\alpha_N$ ) and slope ( $\beta_N$ ) parameters. Adjustment of the nowcast data  
 937 ( $N$ ) by these parameters is a good test of the retrieval method, as the  
 938 adjusted nowcast should be unbiased ( $\alpha_N \approx 0$  and  $\beta_N \approx 1$ ). Regardless  
 939 of the method, however, it is important to note that no bias correction  
 940 can fully address a mismatch in support.
  
- 941 • Autoregressive (AR) parameterization: an established expression of  
 942 information propagation; used here to encompass not just error auto-  
 943 correlation in time or space but also error cross-correlation between  
 944 two ocean current variates. The first order (AR-1) form explored here  
 945 is the simplest.
  
- 946 • Competitive validation: evaluation of two or more datasets (or versions  
 947 of a single dataset) against one chosen reference dataset, where the  
 948 metric of success is shared true variance. Even if linear calibration is  
 949 postulated (as in this study, rather than estimated from a measurement  
 950 model), removal of linear bias from one dataset has no impact on shared  
 951 truth, but this is not so for error. This approach was first attempted  
 952 by Bentamy et al. (2017) in a comparison of heat flux estimates.
  
- 953 • Footprint: target area (e.g., at the ocean surface) that contributes to  
 954 radiation received by a satellite sensor during an imaging interval. Un-  
 955 less it is possible to combine views of the same target area to synthe-  
 956 size higher resolution, the footprint often defines a support scale lower  
 957 bound.
  
- 958 • Instrumental variable: additional data is often required when the mea-

959       surement model has too many unknown parameters to estimate. A  
960       conventional instrument, following Fuller (2006), is a variable that is  
961       taken to be correlated with truth but not with error. The forecast and  
962       revcast (FERS) lagged variables, by comparison, involve correlation of  
963       both truth and error, but this is accommodated by their model equa-  
964       tions. As instruments, FERS play the required role of facilitating the  
965       identification of all model parameters.

966       • Measurement model: measurement *error* models accommodate error in  
967       all sources of information [i.e., both in the calibrated and uncalibrated  
968       data; this accommodation is known as (Fuller, 2006) an approach to  
969       errors in variables in econometrics and observation error or measure-  
970       ment error in other fields]. There is no intended distinction between a  
971       measurement model and measurement error model. The sole rationale  
972       for omitting the term “error” is that a more balanced focus on truth  
973       and error can be anticipated. In other words, a regression model is  
974       effectively a truth model as much as it is an error model. However,  
975       only if it is possible to claim that a model does not lack any broad  
976       category of error (i.e., equation error or correlated error), does it seem  
977       justifiable to explore inferences based on truth.

978       • Parsimony: synonymous with simplicity, especially in reference to mea-  
979       surement models that minimize the number of parameters to be iden-  
980       tified. That is, non-technical definitions apply (e.g., to a careful collec-  
981       tion or use of data with minimal extra assumptions).

982       • Shared variance: synonymous with correlation and involving a term

983 that appears in more than one of the measurement model equations of  
984 interest (possibly multiplied by a parameter). The concept of sharing  
985 applies to both truth and error. It is central to the idea that there can  
986 be multiple truths, with each containing information about overlap-  
987 ping data supports, and that measurement model assumptions should  
988 be considered when determining statistical independence. It should be  
989 noted that standard metrics, including the coefficient of determination  
990 or percentage of explained variance, correlation with truth (McColl  
991 et al., 2014), and SNR (Gruber et al., 2016b) are all subject to inter-  
992 pretation in terms of shared variance.

- 993 • Strong constraint: as an example, many equations of the GlobCurrent  
994 and drifter covariance matrix (6) are satisfied exactly as part of any  
995 measurement model solution (cf. weak constraint).
- 996 • Support: a characterization of the type (e.g., range or quality) of infor-  
997 mation that a given platform or instrument is sensitive to. Often this is  
998 with reference to spatial and temporal scales that can be resolved, but  
999 any information sensitivity can be included, which implies that such  
1000 information may exist as truth or perhaps as equation error, according  
1001 to the measurement model.
- 1002 • Synergy: an approach to combining information such that the whole  
1003 is more valuable and informative than the sum of individual contri-  
1004 butions. Measurement modelling is an unlikely tool to prescribe how  
1005 synergy could be achieved, but may permit the quantitative exploration  
1006 of both individual contributions and informed attempts to combine in-

1007

formation.

1008

1009

1010

1011

1012

1013

1014

1015

1016

- Triple collocation: following McColl et al. (2014), the model parameters sought are uncorrelated error variance of three independent datasets, and with one dataset as a reference, additive and multiplicative calibration of the other two. Following Stoffelen (1998), this measurement model implicitly includes cross-correlated error (e.g., representativeness error) because three different sources of information invariably have three different supports, so at least between two information sources with broader support (e.g., higher resolution), error cross-correlation would be expected.

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026

- True variance estimation: curves of the LHS-RHS of the autocovariance equations (7) are each characterized by a single localized minimum and flatness elsewhere in the range of zero to  $Var(I)$ . The present study treats each available minimum as an equally good estimate of shared true variance and their average is taken. This is in contrast to a global minimum sought using the average of all such curves. However, minima are often not overlapping so the global minimum is effectively a selection among one of the six possible minima. This implies a reliance on the accuracy of each curve in representing its own (very small) minimum value, which might be ill advised.

1027

1028

1029

1030

- Weak constraint: as an example, the autocovariance equations provide different target estimates of shared true variance that cannot all be satisfied simultaneously; a solution close to the center of the ensemble is thus adopted (cf. strong constraint).