Using Kolmogorov–Smirnov Statistics to Assess Jason, TOPEX, and Poseidon Altimeter Measurements

Sarah T. Gille
Scripps Institution of Oceanography and Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA 92093-0230

sgille@ucsd.edu

phone: 858-822-4425; fax: 858-534-9820

Submitted to *Marine Geodesy*, October 9, 2003

Short title: KOLMOGOROV–SMIRNOV STATISTICS TO ASSESS ALTIMETRY

Abstract.

The Kolmogorov–Smirnov (K–S) test is used to compare probability density functions (PDFs) of geostrophic velocities measured by the TOPEX, Poseidon, and Jason altimeters. Velocity PDFs are computed in 2.5° by 2.5° boxes for regions equatorward of 60° latitude. Although velocities measured by the TOPEX and Jason altimeters can differ, on the basis of the K–S test, the velocities are statistically equivalent during the $\sim$200 day period when the satellites followed the same orbit. Full records from TOPEX, Poseidon, and Jason show less agreement, which can be attributed to temporal variability in ocean surface velocities and differing levels of measurement noise.

Keywords: Altimetry, Kolmogorov–Smirnov Statistics, Probability Density Functions, TOPEX/Poseidon and Jason

1. Introduction

Three different satellite altimeters have provided a continuous record of sea surface height along the same ground tracks since 1992. The French-U.S. TOPEX/Poseidon satellite, launched in August 1992, carries the TOPEX and Poseidon altimeters. The newer Jason satellite, launched in December 2001, follows the original TOPEX/Poseidon ground track carrying the Poseidon-2 altimeter. Both satellites repeat their orbits every ∼9.9 days. In analyses of TOPEX and Poseidon measurements, investigators conducting studies that are sensitive to noise, sometimes reject all Poseidon data because of concerns that Poseidon appears to have higher noise levels than TOPEX [e.g. *Stammer*, 1997; *Gille and Llewellyn Smith*, 2000]. Because the Jason altimeter is an outgrowth of the Poseidon instrument, similar concerns could arise for Jason measurements. The objective of this study is to evaluate the three altimeters in order to assess whether there are statistically significant differences in the measurements that might bias studies based on more than one satellite.

Probability density functions (PDFs) of surface geostrophic velocities are used as a statistical indicator in this study. PDFs measure the empirical likelihood that a particular data value will be observed and are useful for assessing the chance of seeing extreme events that might be associated with major storms or anomolous data points. Geophysical quantities are often assumed to have Gaussian PDFs; TOPEX surface geostrophic velocities are Gaussian in most parts of the ocean, although they often indicate non-Gaussian tails. *Gille and Llewellyn Smith* [2000] found that PDFs tend to be distinctly non-Gaussian in regions associated with strong mean flows, such as occur in western boundary currents. Figure 1 **Figure 1.** shows the global velocity PDF, weighted by local variance. The global-average PDF is nearly Gaussian for small velocities but has large tails compared with a true Gaussian distribution. It shows a small dip at zero velocity, indicating that observations are slightly less likely to indicate zero velocity than a theoretical distribution might predict.

If data statistics are stationary in time, then PDFs can be used to compare observations that are collected at different points in time or space but that are expected to have similar statistics. Kolmogorov–Smirnov (K–S) statistics provide a formal measure of the likelihood that two data sets of finite size could be drawn from the same data distribution. Here they are used to assess whether PDFs from TOPEX, Poseidon, and Jason are consistent with each other. In essence, this study evaluates whether the three instruments observe small and large velocity events with equal frequency.

2. Background: Processing the Data

A number of corrected sea level anomaly products have been produced from the TOPEX/Poseidon and Jason altimeter data. However, for this study, Geophysical Data Records (GDRs) [*Benada*, 1997; *Picot et al.*, 2003] are analyzed directly. TOPEX/Poseidon data through cycle 365 and Jason data through cycle 24 were available for this study. Processing algorithms were orginally developed for geophysical analyses and are designed to retain high-wavenumber variability [*Yale et al.*, 1995]. The GDRs provide sea surface heights at a sampling frequency of 10 Hz (corresponding to a geographic distance of about 1.2 km) for TOPEX/Poseidon and 20 Hz for Jason. For this analysis, the 10-Hz measurements are filtered using a 25-point Park-McClellan low-pass filter [*Yale et al.*, 1995], and 20-Hz data are filtered with an equivalent 49-point Park-McClellan filter. Geophysical corrections are then applied, and data are stored at 5-Hz intervals. Most geophysical data corrections represent slowly varying processes, such as large-scale atmospheric effects, and are stored at 1 Hz frequency. Analyses that focus on small-scale sea surface slopes such as evaluations of small-scale sea floor bathymetry [*Smith and Sandwell*, 1997] or surface geostrophic velocities [e.g. *Gille and Llewellyn Smith*, 2000], have little sensitivity to smoothly varying corrections. Although all standard corrections are applied to the data for this analysis, only the tidal correction is expected to have much impact on the results. In this analysis, the CSR3.0 tide model is applied to TOPEX and Poseidon observations [*Eanes and Bettadpur*, 1996] and the GOT99.2 tide model to Jason observations [*Ray*, 1999].

Sea surface height anomalies are computed by removing a time-averaged along track sea surface height from individual height measurements. Height anomalies (η') are then low-pass filtered and archived at 0.5 Hz (≈12 km) resolution. Sea surface slopes between consecutive 0.5 Hz sea surface height anomalies are determined and used to compute geostrophic veocity anomalies: $v' = g/f\, \partial\eta'/\partial x$. Figure 2 shows geostrophic velocity time series from the TOPEX and Jason calibration period. These measurements span the Southern Ocean portion of track 'a003' and indicate strong variability near 55°S, which is associated with the meandering of the Antarctic Circumpolar Current. TOPEX and Jason measurements are roughly in agreement, although both show signs of anomalous events not detected by the other instrument. **Figure 2.**

For this analysis, PDFs are computed by sorting the data into 2.5° latitude by 2.5° longitude geographic boxes. Observations from ascending and descending satellite passes are combined, because previous analyses have indicated little difference in their velocity PDFs [*Gille and Llewellyn Smith*, 2000]. In the next section, PDFs from each of these geographic boxes are intercompared using K–S statistics.

3. Kolmogorov–Smirnov Statistics

K–S statistics measure the separation between cumulative distribution functions in order to evaluate whether two empirical data sets are likely to originate from the same underlying PDF. *Press et al.* [1988] provide a synopsis of the test criteria. The statistic can be strongly dependent on the number of observations available: small data sets are difficult to distinguish from each other, while large data sets that appear to have small differences often fail the K–S test.

3.1. Baseline Statistics

In this study, the K–S test was performed for pairs of geographically co-located PDFs derived from observations from the three different altimeters. PDFs are compared for each 2.5° by 2.5° box in order to evaluate geographic differences between the instruments. For the baseline tests, height anomalies were computed relative to the long-term TOPEX altimeter mean, because observed means for short records may have large statistical uncertainties. The low-pass filter applied in the final processing stage was a 140-point Gaussian filter, designed to suppress variability at frequencies greater than about 0.12 Hz. Because the data are low-pass filtered, the number of effective degrees of freedom N_{eff} is less than the total number of observations, and can be computed from the lagged covariances [e.g. *Davis*, 1976]. In this case, filtered white noise was used to estimate $N_{eff} = 0.47N$ for the standard filter, where N is the number of available 0.5 Hz observations. Thus K–S statistics were computed by selecting alternate observations. Finally, in the baseline tests summarized for each PDF, the first moment μ_1 (or mean velocity) was subtracted so that only velocity anomalies relative to the mean were considered. Data collected near the equator are included in this analysis but may be subject to high noise levels, because the Coriolis parameter f approaches zero.

Table 1 summarizes results from the K–S tests comparing velocity PDFs. Numbers indicate the percentage of boxes for which PDF comparisons fail the K–S test at the 5% level. If data behaved exactly like random samples drawn from the same distribution, approximately 5% of the samples would be expected to fail the K–S test. (Empirical tests with Gaussian white noise predict global failure rates of 5% $\pm$ 0.3% given the number of available samples.) Results show that the actual failure rate for the statistical tests varies from 2% to 28% depending on a variety of factors. **Table 1.**

Figure 3, corresponding to the baseline case in the first column of Table 1, shows the geographic distribution of PDFs that fail the K–S test. PDFs that are statistically indistinguishable are white, while regions with PDFs that differ at the 5% level are shaded gray. **Figure 3.**

Figure 3a compares PDFs from TOPEX and Jason for the overlap time period, from Jason cycle 2 through 21 (TOPEX cycles 345 to 364). Midway through Jason cycle 22, TOPEX/Poseidon was moved to an orbit designed to interleave the Jason orbits, for which insufficient data are at present available to ensure a stable mean. Over most of the ocean, Jason and TOPEX PDFs for the overlap time period are indistinguishable. Approximately 3% of the 2.5° by 2.5° boxes have PDFs that differ (gray). This indicates better agreement than the K–S statistic would predict, and suggests that Jason and TOPEX come close to measuring the same physical quantities. Figure 3a indicates that PDF mismatches can occur at all latitudes and all hemispheres, but are most likely in coastal regions. This is not surprising, since tide models are less successful in coastal areas, and mean sea surface heights may indicate substantial uncertainties.

When all available data are used, in the baseline case Jason PDFs differ from TOPEX PDFs in 15% of the available boxes, as shown in Figure 3b. This difference can be partially explained by temporal variability in PDF fields. As shown in Figure 3c, when TOPEX data from the overlap period are compared with the full record of TOPEX data, PDFs are statistically different in 6% of cases, slightly exceeding the predicted value of 5%. Approximately 1% to 2% more boxes are statistically different when Jason or TOPEX overlap data are compared with TOPEX data collected prior to the Jason launch (not shown). Thus seasonal and interannual fluctuations in ocean variability appear to change PDFs sufficiently that they cannot be assumed to be stationary in time, but not enough to explain the full difference between long-term Jason and TOPEX results. This suggests that Jason and TOPEX may have slightly different levels of measurement noise. As in Figure 3a, in Figures 3b-c, the geographic distribution of PDFs that disagree stretches across all hemispheres, with a disproportionate number of coastal points. Overall, PDFs appear more likely to disagree at mid to high latitude regions. For example, in both panels the Gulf Stream and Kuroshio Extension regions show comparatively high concentrations of gray points, which may be associated with baroclinic instability or interannual variability in these regions.

The final comparison category considers Poseidon data. Poseidon measurements are interspersed in time between TOPEX measurements through the duration of the TOPEX/Poseidon mission, so we might expect their respective PDFs to sample seasonal cycles similarly and therefore to agree. Results in Figure 3d indicate that PDFs differ in 12%

of cases. This discrepancy may be due to short-term temporal variability, but it is probably also a sign that Poseidon and TOPEX observations have different noise levels. In contrast with results in Figures 3b-c, in Figure 3d, mismatches are more likely near the tropics than in mid-latitudes. Since f is small near the equator, low-latitude velocities tend to appear noisier than high-latitude velocities, and instrumental noise is especially likely to show up at near-equator points.

Means were removed from PDFs to compute the K–S statistics in Figure 3. If data were collected primarily during a period of time when surface geostrophic velocities were unusually high or low, removing the mean might bias the statistics. Column 2 of Table 1 summarizes the statistics obtained when means are not removed from each box separately. Column 2 results indicate less agreement than Column 1 results, indicating that mean velocities in these PDFs may differ substantially.

3.2. Dependence on Spatial Filtering

Observations discussed in the previous section were filtered with a 140-point Gaussian filter, designed to retain much of the high wavenumber information available from the data. The comparisons shown in Figure 3 therefore may reflect high-wavenumber variability that most users would assume to be instrumental noise. To suppress this noise, the analysis was repeated by using a stronger 180-point Gaussian filter with a frequency cut-off approximately half that of the 140-point filter. In this case, the number of degrees of freedom is estimated to be $N_{\mathit{eff}} = 0.235N$, and K–S statistics are therefore computed using one in four velocity observations. Results shown in the fifth and sixth columns of Table 1 indicate minor differences in the fraction of PDFs that disagree.

The largest difference between strongly and weakly filtered data occurs in comparisons of Poseidon versus TOPEX observations that retain the PDF mean μ_1. In this case, applying a strong filter decreases the fraction of PDFs that fail the K–S test from 15% to 7%. This difference could represent random variability, but it may also stem from the higher noise levels reported in Poseidon observations.

Comparisons of Jason and TOPEX measurements or TOPEX and TOPEX measurements do not show a similarly strong sensitivity to filter strength. Overall, these results suggest that differences between TOPEX and Jason PDFs cannot readily be eliminated by applying stronger filters to the data.

3.3. Anomalies Relative to the Data Mean

The statistics in the first two and last two columns of Table 1 are all based on velocity anomalies computed relative to the long-term TOPEX mean. This comparison methodology is possible, because all three instruments followed the same ground tracks. In contrast, as of early 2003, the four active altimetric satellites, Envisat, Geosat Follow On (GFO), TOPEX/Poseidon, and Jason, followed four separate ground tracks. K–S statistics can in principle allow a formal comparison of data that are not precisely co-located in space or time. However, intercomparing PDFs from instruments that follow separate ground tracks is only possible if reliable mean altimetric heights can be determined for each satellite.

Columns 3 and 4 of Table 1 indicate results from the K–S test when velocity anomalies are computed relative to each satellite's long-term mean. These results are directly analogous to results from columns 1 and 2 of Table 1 and indicate that overall TOPEX better matches Jason and Poseidon when each satellite's own mean sea surface height is used. The number of PDFs that fail the K–S test still exceeds 5%, indicating that the data are not providing perfect statistical matches. This is expected given the temporal variability of the ocean. Moreover, like the anomalies, in this case the means are computed over differing time intervals and differing numbers of samples, so this may not be a perfect test. However, it suggests that the K–S test can be used successfully to measure the statistical disagreement between velocity PDFs measured along different ground tracks.

4. Comparing Variances

While the K–S test indicates whether data in two PDFs appear to be drawn from the same "true" distribution, it does not provide any information to explain why PDFs might differ. Velocity PDFs are expected to differ primarily if one data set is substantially noisier than an other. This section compares apparent noise levels by examining the variances of the observations.

Table 2.

Figure 4.

Table 2 indicates the percentage of PDFs for which one altimeter measures greater variance than the other, and Figure 4 shows the geographic distribution of the variance disagreements. If data were completely random, one satellite would have higher variance than the other in approximately 50% of the 2.5° by 2.5° PDF bins. In the baseline, lightly filtered, case (column 1 of Table 2) Jason and Poseidon both tend to have higher variance than TOPEX roughly two-thirds to three-quarters of the time. Some of this difference may be explained by the temporal sampling of the data: TOPEX data from the overlap period have higher variances than TOPEX data from the full record in 58% of the PDFs considered. However, as the K–S tests also indicated, this difference is not sufficient to explain the differences between TOPEX and the other two altimeters. Results suggest that TOPEX noise levels are consistently lower than Jason or Poseidon noise levels.

Jason and Poseidon differ in the geographic patterns of their variances, as shown in Figure 4. Jason variances are most likely to exceed TOPEX variances at high latitudes (Figures 4a-b), where high sea states may bias altimeter measurements. This suggests that Jason and TOPEX differ in their response to strong sea states. In contrast, Poseidon's variances are high compared with TOPEX near the equator (Figure 4d), where velocity variances are generally highest, and f is small so that velocities are sensitive to measurement noise. The number of cases for which Poseidon variances appear high drops significantly when data are strongly filtered (column 3 of Table 2). As the K–S tests also implied, Poseidon may experience more isolated instances of high sea surface height than does TOPEX, but that these differences may be filtered out in highly smoothed versions of the data.

5. Summary

Results from this investigation indicate that TOPEX and Jason geostrophic velocities show excellent agreement on the basis of K–S statistics when only data from the same time period are considered. The velocity PDFs agree less well when data from differing time intervals are intercompared; this is partially explained by the fact that ocean eddy statistics vary over time but may also be due to residual differences between the satellites. Overall results indicate greater agreement when mean velocities are removed from each PDF. Results are not strongly dependent on the degree of spatial filtering applied to the data. Comparisons between TOPEX and Poseidon observations indicate that disagreement is most common near the equator, where results are expected to be sensitive to noise, because f is small, and these comparisons show the clearest improvement when a stronger filter is applied. Comparisons between TOPEX and Jason do not show the same patterns.

Baseline cases were computed using velocity anomalies relative to the long-term TOPEX mean. However, long-term TOPEX, Poseidon, and Jason PDFs showed slightly greater agreement when anomalies were computed relative to the time means for each instrument. This suggests that K–S statistics should also allow plausible comparisons of data collected along different ground tracks.

Both Poseidon and Jason have higher variances than TOPEX, suggesting that the disagreements found with the K–S test are most likely attributable to higher noise levels in the Poseidon and Jason data. Together the K–S and variance statistics suggest that Poseidon is noisier than TOPEX. In contrast, Jason does not appear to experience the same noise issues as Poseidon, suggesting that instrumental effects may play only a minor role in the long-term time series that emerges from TOPEX and Jason data.

Acknowledgments. TOPEX/Poseidon Merged Geophysical Data Records and Jason Geophysical Data Records were provided by the Jet Propulsion Laboratory's Physical Oceanography Data Active Archive Center (http://podaac.jpl.nasa.gov). Funding for this study was provided by the

National Aeronautics and Space Administration through the Jason Science Working Team, JPL contract 1224031

References

Benada, J. R., 1997: Merged GDR (TOPEX/POSEIDON) Generation B (MGDR-B) User's Handbook, Tech. Rep. D-11007, Jet Propulsion Laboratory, Version 2.0.

Davis, R. E., 1976: Predictability of Sea Surface Temperature and Sea Level Pressure Anomalies over the North Pacific Ocean, *J. Phys. Oceanogr.*, *6*, 249–266.

Eanes, R., and S. Bettadpur, 1996: The CSR3.0 global ocean tide model: Diurnal and semi-diurnal tides from TOPEX/POSEIDON altimetry, Tech. Rep. CSR-TM-96=05, Univ. of Tex. Cen.t for Space Res., Austin, TX.

Gille, S. T., and S. G. Llewellyn Smith, 2000: Velocity-variance probability distribution functions from altimeter measurements, *J. Phys. Oceanogr.*, *30*, 125–136.

Picot, N., K. Case, S. Desai, and P. Vincent, 2003: AVISO and PO.DAAC User Handbook, IGDR and GDR Jason Products, Tech. Rep. JPL D-21352 and SMM-MU-M5-OP-13184-CN, Jet Propulsion Laboratory, Pasadena, CA, Edition 2.0.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1988: *Numerical Recipes in C*. Cambridge University Press, New York.

Ray, R. D., 1999: A global ocean tide model from TOPEX/POSEIDON altimetry: GOT99.2, Tech. Rep. NASA Tech. Memorandum 1999-209478, Goddard Space Flight Center, Greenbelt, MD.

Smith, W. H. F., and D. T. Sandwell, 1997: Global seafloor topography from satellite altimetery and ship depth soundings, *Science*, *277*, 1956–1962.

Stammer, D., 1997: Global characteristics of ocean variability estimated from regional TOPEX/POSEIDON altimeter measurements, *J. Phys. Oceanog.*, *27*, 1743–1769.

Yale, M. M., D. T. Sandwell, and W. H. F. Smith, 1995: Comparison of along-track resolution of stacked Geosat, ERS 1 and TOPEX satellite altimeters, *J. Geophys. Res.*, *100*, 15,117–15,127.

Sarah T. Gille, Scripps Institution of Oceanography University of California, La Jolla, CA 92093-0230

Received ______________________

Figure Captions

Figure 1. PDF computed from global TOPEX geostrophic velocity data, normalized by local variance. Comparison lines show theoretical PDFs for Gaussian and double-exponential distributions.

Figure 2. Time series of TOPEX (black) and Jason (gray) geostrophic velocity anomalies as a function of latitude across the Southern Ocean. This track (identified as a003 in this analysis) extends from 322°E at 60°S to 354°E at 30°S. TOPEX/Poseidon cycle 361 was a Poseidon cycle, and therefore for consistency, Jason cycle 18 is omitted from the K–S analysis for the overlap interval. The GDR for Jason cycle 19 was not available at the time this paper was written, so TOPEX cycle 362 was also omitted. Midway through TOPEX cycle 365, the TOPEX orbit was changed, so TOPEX cycle 365 and Jason cycle 22 are omitted from the K–S statistics.

Figure 3. Results of K–S tests for PDF similarity for (a) TOPEX compared with Jason during the overlap period only, (b) the full TOPEX mission compared with the full Jason mission, (c) the full TOPEX mission, compared with Topex from the overlap period, (d) Poseidon compared with the full TOPEX mission. White regions indicate that PDFs agree, while gray indicates that the PDFs disagree. Black denotes land or regions with insufficient altimeter observations. In all cases, the mean sea surface height from the full TOPEX mission has been removed, a 140-point Gaussian filter was applied to the observations, and the data from each 2.5° by 2.5° region were demeaned.

Figure 4. Variance comparisons for pairs of altimeter data. Gray regions indicate PDFs for which (a) Jason data has higher variance than TOPEX in the overlap time period, (b) Jason data has higher variance than TOPEX when full records are considered, (c) TOPEX data for the overlap time period only has higher variance than the full TOPEX record, and (d) Poseidon has higher variance than the full TOPEX record. White regions correspond to the first satellite indicating lower variance than the second.

Tables

Table 1. Percentage of 2.5° latitude by 2.5° longitude bins for which PDFs fail K–S test at 95% level, for different comparisons of altimetric observations. Data in the first four columns were processed with a light 140-point Gaussian filter, while the last two columns were treated with a stronger 180-point Gaussian filter, as discussed in the text. For cases using the TOPEX mean, the mean determined from 10-years of TOPEX data on fixed orbit is removed from all observations. Cases using "Own mean" remove the mean from all available TOPEX, Jason, or Poseidon observations. Finally, in cases labeled "$-\mu_1$", the mean velocity has been subtracted from observations in each 2.5° by 2.5° box, while cases labeled "w/ μ_1" retain the PDF mean.

	140-pt Filter				180-pt Filter	
	TOPEX Mean		Own Mean		TOPEX Mean	
Instrumental Records	$-\mu_1$	w/ μ_1	$-\mu_1$	w/ μ_1	$-\mu_1$	w/ μ_1
TOPEX vs Jason (overlap)	3	4			2	4
TOPEX (full) vs Jason	15	27	15	13	17	28
TOPEX (full) vs TOPEX (overlap)	6	17			10	18
TOPEX (full) vs Poseidon	12	15	9	8	9	7

Table 2. Percentage of 2.5° latitude by 2.5° longitude bins for which velocity variance from Jason, Poseidon, or an abbreviated TOPEX record exceeds TOPEX velocity variances. Since variances are normally computed relative to the data mean, μ_1 is effectively always subtracted. Filters and data means are as described in Table 1.

	140-pt Filter		180-pt Filter
Instrumental Records	TOPEX Mean	Own Mean	TOPEX Mean
Jason > TOPEX (overlap)	73		62
Jason > TOPEX	63	33	50
TOPEX (overlap) > TOPEX (full)	58		57
Poseidon > TOPEX (full)	65	45	29

Figures

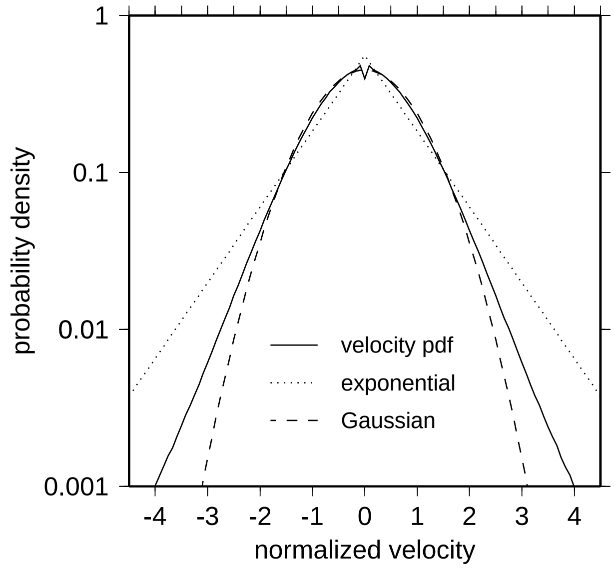

Figure 1. PDF computed from global TOPEX geostrophic velocity data, normalized by local variance. Comparison lines show theoretical PDFs for Gaussian and double-exponential distributions.

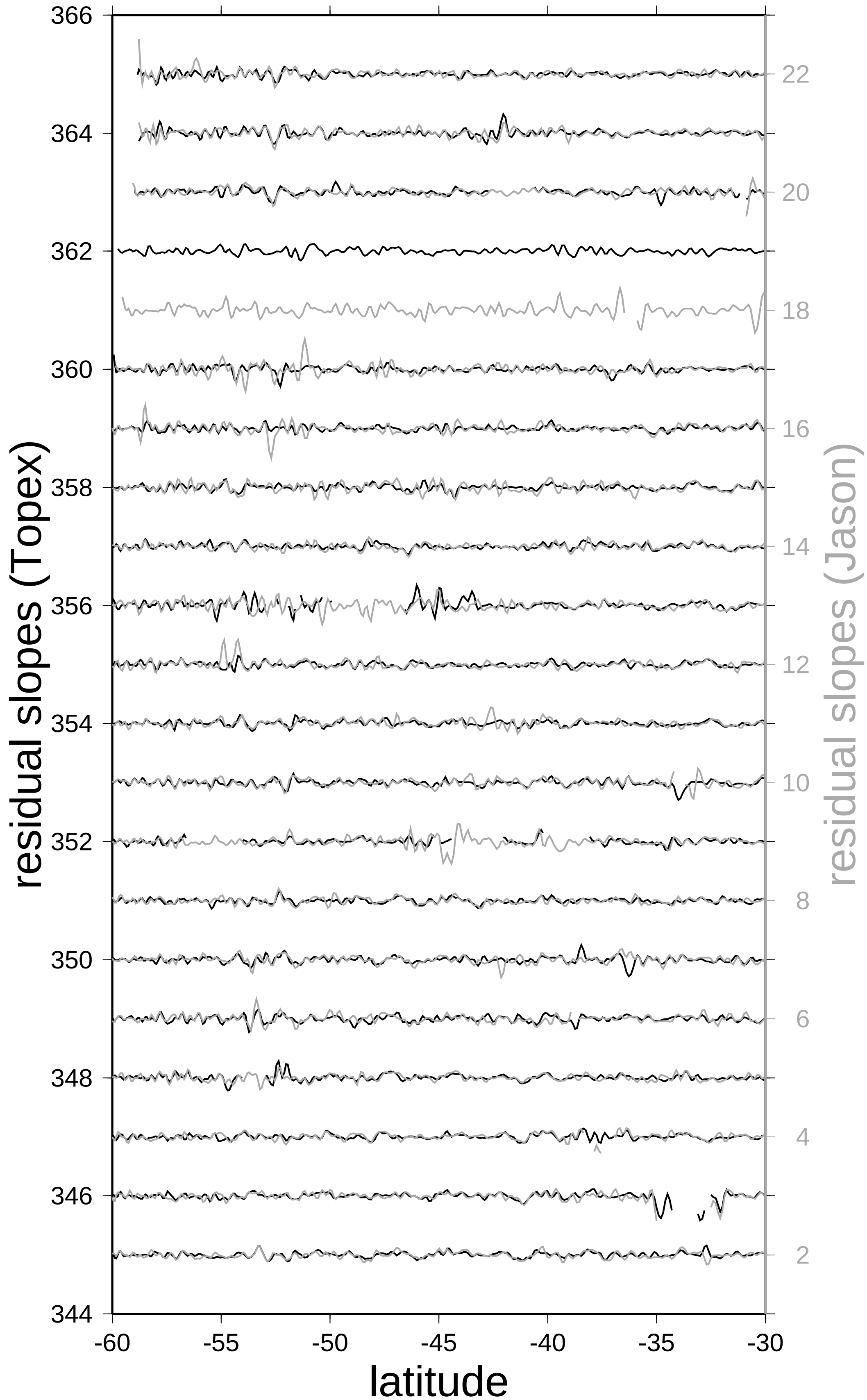

Figure 2. Time series of TOPEX (black) and Jason (gray) geostrophic velocity anomalies as a function of latitude across the Southern Ocean. This track (identified as a003 in this

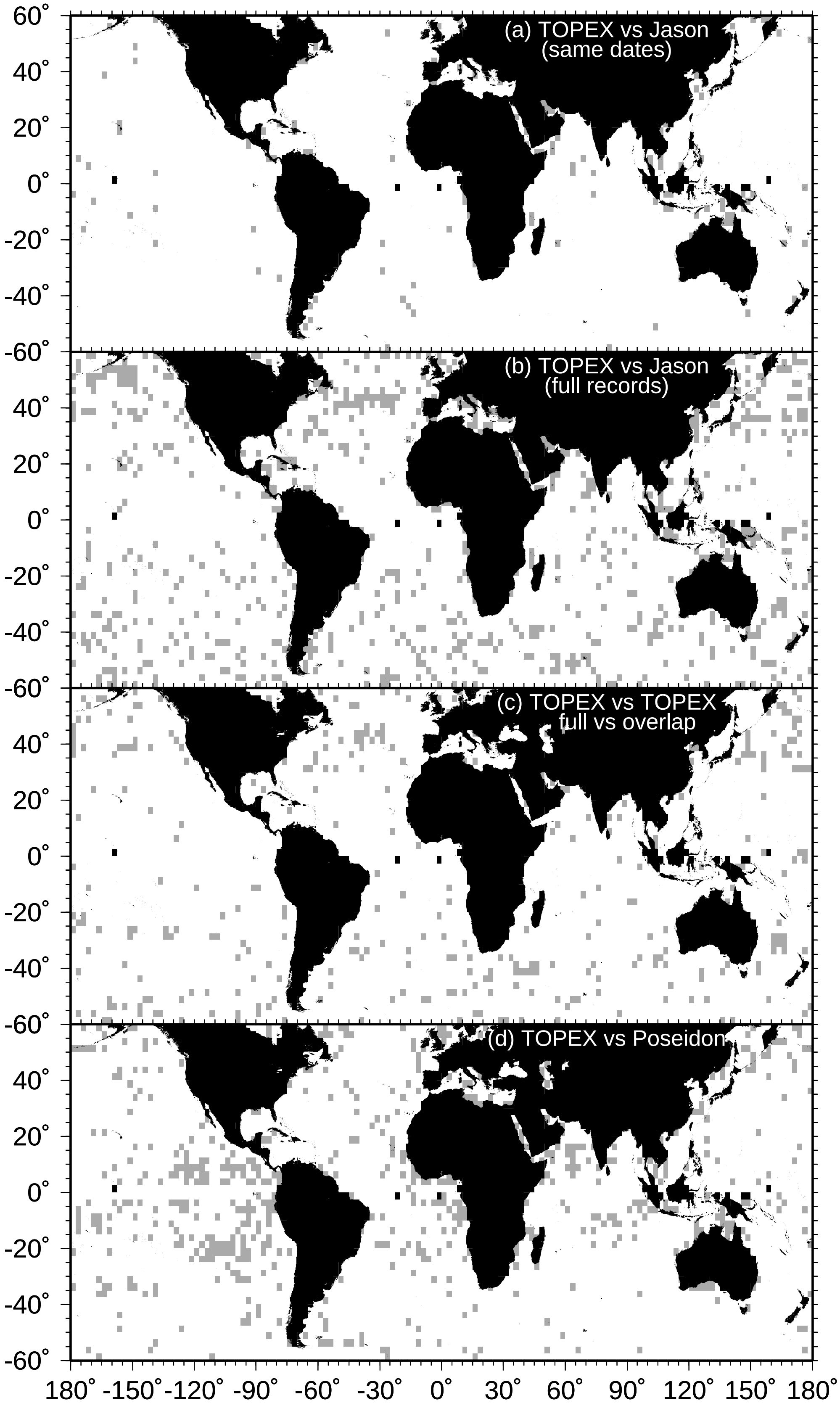

Figure 3. Results of K–S tests for PDF similarity for (a) TOPEX compared with Jason during the overlap period only, (b) the full TOPEX mission compared with the full Jason mission, (c) the full TOPEX mission, compared with Topex from the overlap period, (d) Poseidon compared

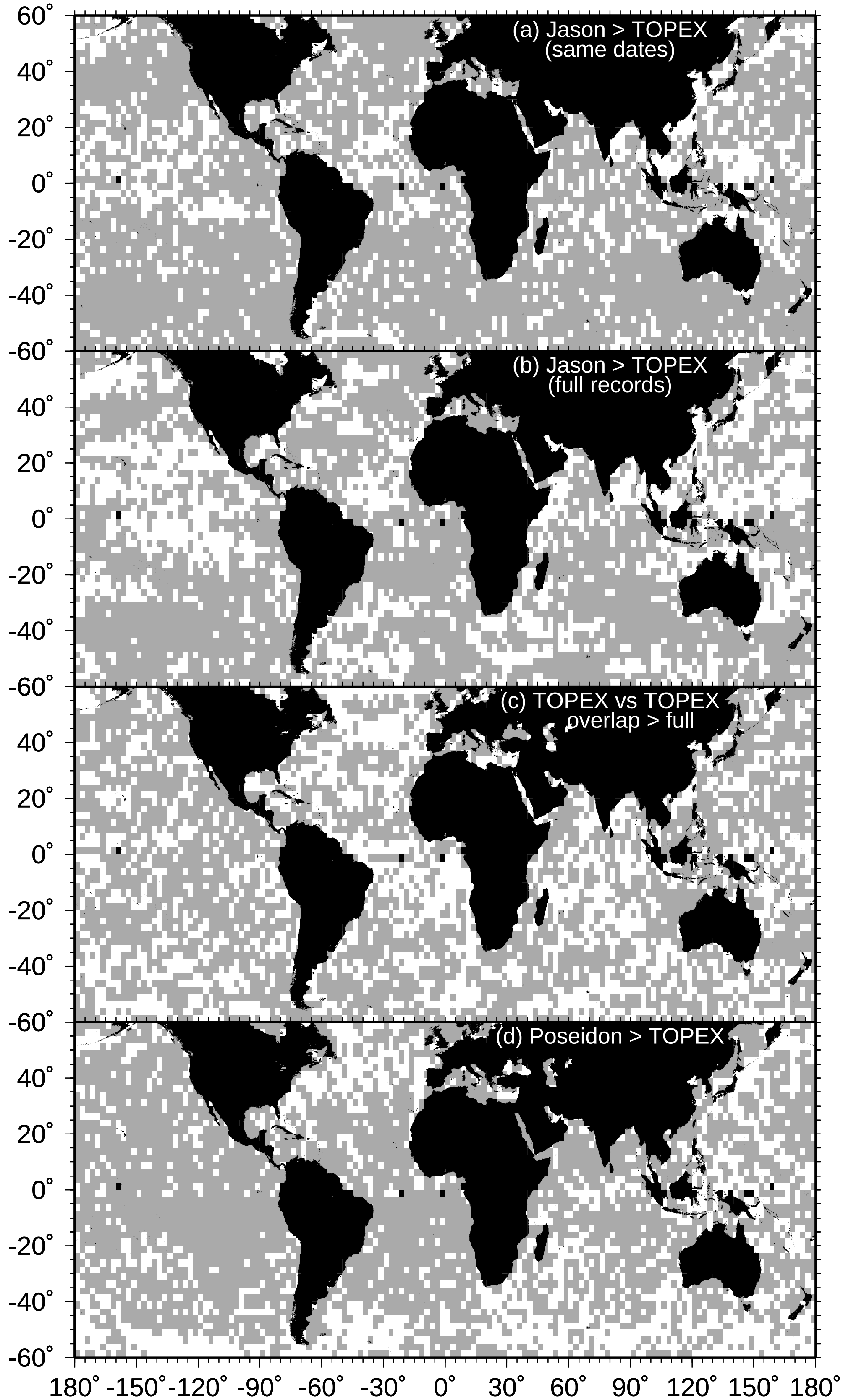

Figure 4. Variance comparisons for pairs of altimeter data. Gray regions indicate PDFs for which (a) Jason data has higher variance than TOPEX in the overlap time period, (b) Jason data has higher variance than TOPEX when full records are considered, (c) TOPEX data for