

A Comprehensive Database for Antarctic Iceberg Tracking Using Scatterometer Data

Jeffrey S. Budge and David G. Long ^{id}, *Fellow, IEEE*

Abstract—This paper describes the development of, and the methodology for, a new, consolidated Brigham Young University (BYU)/National Ice Center (NIC) Antarctic iceberg tracking database. The new database combines daily positional data from the original BYU daily iceberg tracking database derived from scatterometers, and the NIC’s weekly Antarctic iceberg tracking database derived mostly from optical and infrared sensors. Interpolation methods and statistical analyses of iceberg locations are discussed. A new, automated method of using positional data and scatterometer backscatter images to estimate sizes and rotational patterns of icebergs is also developed. This information is included in the new database.

Index Terms—Advanced scatterometer (ASCAT), backscatter, European remote sensing (ERS), icebergs, NASA scatterometer (NSCAT), OceanSat, QuikSCAT, scatterometer, Seasat scatterometer (SASS), SeaWinds.

I. INTRODUCTION

DATABASES of current and archived iceberg positions are useful for facilitating many different maritime and scientific pursuits in the Antarctic seas. Marine biologists, for example, seek current iceberg coordinates in order to study iceberg-adjacent ecosystems [1]. Current iceberg positions are used by Antarctic cruise ships, military submarines, and sailor circumnavigators to facilitate research and business, as well as to avoid collisions. Archived data have been used to study the movement and disintegration of icebergs over time [2], and to validate methods of iceberg detection through hydroacoustic sensing [3] and tracking via altimeter waveform analysis [4].

In this paper, the two databases are consolidated and enhanced; namely, the Brigham Young University (BYU) Microwave Earth Remote Sensing (MERS) Antarctic iceberg tracking database [5] and the National Ice Center (NIC) Antarctic iceberg database (www.natice.noaa.gov). The two databases are currently published in separate locations, each with a unique format. The NIC publishes a collection of coordinate points, iceberg lengths and widths, and some optical photos, using a variety of sources. The NIC database is limited to one coordinate per iceberg per week. In contrast, the BYU MERS database has a set of coordinate points for each day that any particular scatterometer is active. These data are separated by sensor and covers

Manuscript received July 10, 2017; revised October 2, 2017 and November 23, 2017; accepted December 8, 2017. Date of publication January 4, 2018; date of current version February 12, 2018. (Corresponding author: David G. Long.)

The authors are with the Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT 84602 USA (e-mail: Jeff.Budge@comcast.net; long@ee.byu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2017.2784186

TABLE I
SENSORS INCLUDED IN THE BYU DATABASE AND
THE YEARS THEY OPERATED

Sensor	Agency	Years
SASS	NASA	1978
ERS-1/2	ESA	1992–2001
NSCAT	NASA	1996–1997
QSCAT	NASA	1999–2009
ASCAT	ESA	2008–present
OSCAT	ISRO	2012–2014

from 1978 to the present day, as shown in Table I. The sensors include the European Space Agency (ESA) advanced scatterometer (ASCAT), the Indian Space Research Organization (ISRO) Oceansat-2 scatterometer (OSCAT), the NASA scatterometer (NSCAT), the European remote sensing satellites 1 and 2 (ERS-1/2), the SeaWinds on QuikSCAT scatterometer (QSCAT), and the Seasat scatterometer (SASS).

Due to differences in the organization of the NIC and BYU databases, comparing same-day data from the different sources can be difficult. Further, each sensor used by the MERS database is in a separate file and has a different mission lifetime, which introduces gaps in the archive over the life of a single iceberg as database changes from one sensor to another. To ameliorate these problems, a consolidated database has been developed to include the features of both databases. The new database is designed to simplify the long-term comparison of archived data, and to provide daily iceberg positions. To this position database, we add iceberg size and rotation information derived from scatterometer backscatter measurements. A plot of the all iceberg locations in the database is shown in Fig. 1.

The information in the new database is organized into two separate sets. Each set is arranged first by iceberg, and second by date, with information from all relevant sensors or measurements listed in the same file. At least one original measurement or an interpolated position is available for each day of an iceberg’s lifetime. The bulk of these data come from the BYU database. The NIC contributions to the database include weekly positional data as well as the names for the icebergs. NIC size estimates are used to validate the size estimates generated using the new scatterometer size estimation algorithm described in this paper.

The NIC database includes manually determined iceberg size estimates derived from optical imagery. The manual method is labor intensive and is limited by the availability and quality of cloud-free images of each iceberg. However, scatterometer backscatter measurements can be used to automatically and

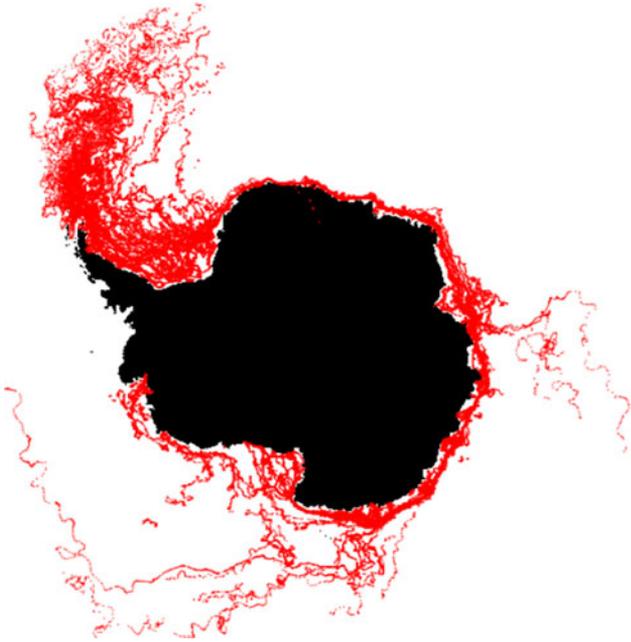


Fig. 1. Plot of all iceberg positions recorded in the consolidated database, which contains data on icebergs over 5 km^2 , from 1992 to the present. More than 90% of all icebergs, regardless of their source, pass through the Weddell Sea [5].

consistently estimate the geometrical size, shape, and area of tabular Antarctic icebergs. Such information can be useful in better understanding patterns in iceberg behavior. Further, the rotation of the iceberg can be determined. Studying this rotation can give insight into geophysical systems, which can be complicated and difficult to model by other means. Such information is expected to be useful to many different disciplines and can be used to guide decision making for economic and scientific endeavors.

In this paper, the methods by which iceberg positions are determined for each database are briefly described and the method used to create the consolidated position database is shown. Background is discussed in Section II. The process by which the position data are interpolated is discussed in Section III. Section IV describes the method by which the positional information in the compiled database is used to generate size estimates. This method is built upon in Section V to derive rotational estimates as well. Section VI considers some applications of the information in the database. Finally, Section VII concludes this paper.

II. BACKGROUND

The radar images used in this paper to determine the size, shape, and rotational motion of icebergs are created from scatterometer backscatter data. These are the same images used to determine iceberg position in the original BYU database. A scatterometer is an active space-borne microwave sensor that measures surface backscatter (σ^o) values [6]. The σ^o values depend on the geometric and dielectric properties of the surface. The magnitude of σ^o can be used to discriminate between surfaces a signal may have scattered from. Seawater, for example, generally exhibits a lower σ^o than sea ice, which generally has a lower σ^o than land. The σ^o contrast between seawater, sea ice,

and iceberg σ^o enables location and tracking of icebergs in the ocean. A complicating factor is that ocean σ^o depends on the near-surface wind speed, and thus, is highly variable in time and space.

A tabular iceberg in a microwave scatterometer image can be identified by the high σ^o values across its surface. This is due to the volume scattering of the many layers of snow and ice that constitute an iceberg. Volume scattering also reduces the variance of σ^o caused by azimuth angle variations over the different sensor passes used to create the image. Thus, the iceberg typically appears as a group of pixels with similar σ^o values [5], which enables an iceberg to be identified and tracked. This method has been used by BYU MERS to create daily position tracks of the icebergs in its database [7], which can be found at www.scp.byu.edu/iceberg.

Iceberg locations in the original MERS database are determined using daily scatterometer images of the Antarctic region from various satellite scatterometers [5]. MERS first receives scatterometer σ^o measurements, organized by satellite orbit, as L1B or SZF files. Multiple files are combined into a single enhanced resolution σ^o image using the scatterometer image reconstruction (SIR) algorithm [9], [10]. One SIR image is produced for every day of the year for which scatterometer data are available. Iceberg positions are determined by manually identifying the high backscatter regions corresponding to icebergs in these SIR images and correlating them with previously archived positions and the NIC's reported iceberg locations.

The BYU database contains information from the eight scatterometers listed in Table I. The earliest data begin with SASS in July 1978, though the bulk of the data begin in 1992 with ERS-1, and continue to the present with ASCAT. MERS position data carry the disclaimer that, because it is collected manually using moderate-resolution satellite σ^o images, its accuracy is limited to no better than plus-or-minus one pixel. Pixel resolution values vary from 2.225 to 8.9 km/pixel according to the sensor used [7].

The primary backscatter images used in this paper are Ku-band SIR images created from QSCAT data, though other sensors and bands are employed. Since it has been observed that higher frequencies produce more contrast in SIR images, Ku-band is more useful for iceberg tracking [8]. Location and contours are also generated using C-band SIR images created from ERS-1/2 and ASCAT data when QSCAT and OSCAT data are not available.

Scatterometers, as active sensors that emit radiation, do not rely on solar illumination and are able to perform consistently under nearly all weather conditions. Scatterometer σ^o images are sensitive to the roughness of the imaged surface; wind-roughened ocean surfaces, for example, exhibit a higher backscatter than a smooth, calm surface, which reduces the contrast between ocean water and land. Temperature changes also affect contrast in these images, as liquid water has a lower backscatter than sea ice or land. Summer conditions can lead to melting on the surface of an iceberg, which reduces the contrast between the iceberg and open ocean or melting sea ice. As a general rule, scatterometry produces lower resolution images than optical and I/R sensing. A typical QSCAT image of Antarctica is 3880×3880 pixels, with each pixel covering an area between 3.42 and 5.27 km^2 , see Fig. 2. A closer view of a

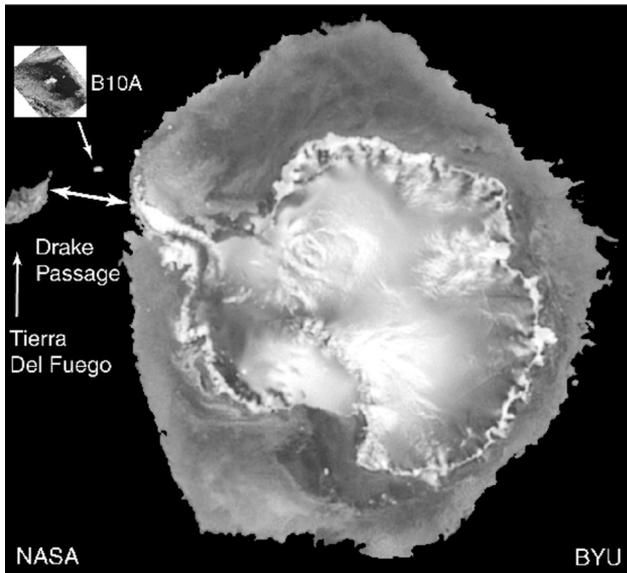


Fig. 2. Example of an enhanced resolution σ^o image of iceberg B10a on JD 205, 1999, from QSCAT data with annotation. In this image, open ocean pixels have been set to black [11].

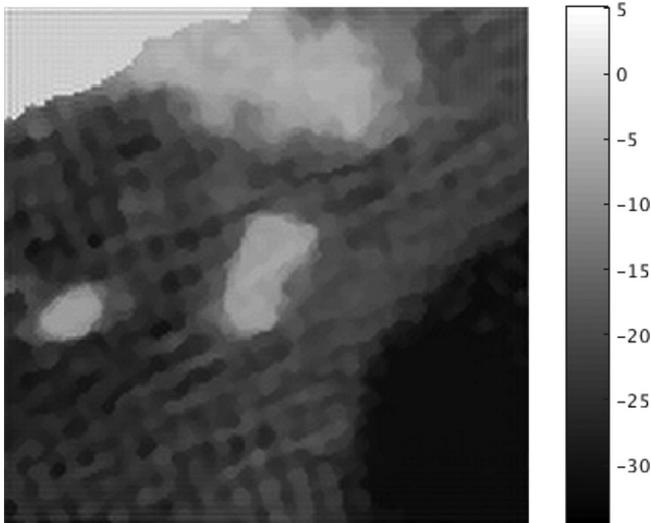


Fig. 3. Iceberg B15b, on JD 47 of 2005, in QSCAT data. This is a scatterometer σ^o image in decibel. The iceberg appears as a group of lighter pixels in the center of the 660×660 km image, centered at 64.9 S, 137.5 E. A separate iceberg is visible on the left. Open ocean is to the right. The white area in the upper left is land. The light area near the top is fast sea ice. In this particular image, the sea ice texture is an artifact of the image formation algorithm that generates high resolution images from lower resolution backscatter measurements [5], [10].

particular iceberg seen in a Ku-band backscatter image is shown in Fig. 3.

A. NIC Data Collection

The NIC is a United States governmental organization comprised of three individual agencies: the United States Navy, the National Oceanic and Atmospheric Association (NOAA), and the United States Coast Guard. Although the establishment of the NIC was not official until 1995, data collected by the Joint Ice Center [a collaboration between NOAA and the Navy (prior

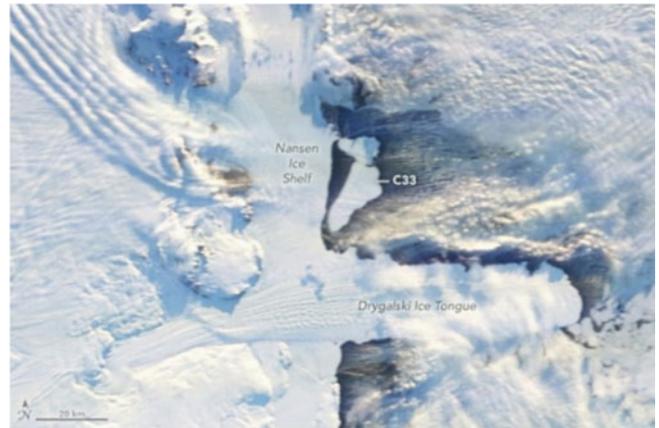


Fig. 4. Iceberg C33 as it calves from the Nansen Ice Shelf, April 7, 2016 [12], as seen in an optical image, courtesy of NASA.

to the addition of the Coast Guard)] date back to 1976 and are included in the NIC database.

About once per week, the NIC publishes the current latitudinal and longitudinal positions of most Antarctic icebergs within a specified area and above a specified size. It also publishes estimates of each iceberg's length and width, measured in nautical miles. The current data can be accessed at the NIC website www.natice.noaa.gov/Main_Products.htm. Archived data are not published online, but are available from them on request at www.natice.noaa.gov/products/south_icebergs_on_demand.html.

The NIC determines iceberg coordinates primarily from images from space-borne sensors, including the environmental satellite (EnviSat), advanced very-high-resolution radiometer (AVHRR), moderate resolution imaging spectroradiometer (MODIS), and Radarsat. A portion of the data comes from ship reports. Other contributions stem from the Defense Meteorological Satellite Program (DMSP) and Argos. The majority of sensors utilized by the NIC are passive optical or infrared (I/R) sensors.

Images collected by optical and I/R sensors are typically intuitive and relatively easy to interpret when not obscured by clouds. However, optical sensors depend on visible light from the sun; when sunlight is absorbed by clouds, or limited by changing seasonal patterns, optical sensors do not function at peak performance or produce useful images. I/R sensors are likewise affected by clouds and seasonal shifts [13]. This is particularly consequential in polar regions such as Antarctica, where the sun does not rise for multiple weeks during the winter and clouds are common. Fig. 4 shows an example of an iceberg image that, while not obscured by clouds, illustrates the potential issues in mapping white icebergs in the presence of cloud cover.

III. DATA INTERPOLATION

In the consolidated database, iceberg position reports from the NIC and original BYU databases have been combined and reformatted for both consistency and simplicity. Where needed and practical, position estimates have been interpolated to daily positions using all the available scatterometer and NIC observations. Before 1999, when QuikSCAT was launched, daily

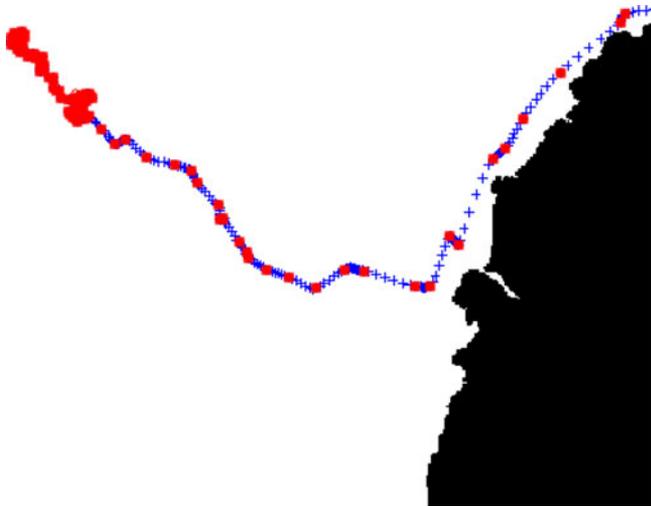


Fig. 5. Zoomed in plot of the track for iceberg B15b, which calved from B15. The iceberg leaves the coastal zone of Antarctica at 73.6 S, 23 W, near Queen Maud Land. Red stars are the observed locations. Blue crosses represent interpolated positions for days in which no data are available.

scatterometer measurements of the Antarctic region were not possible due to the coverage limitations of the instruments in orbit at the time. The limited swath width of the scatterometers operating prior to QSCAT did not provide daily coverage of the polar regions, and so, could not generate the daily backscatter images used for iceberg location tracking. In order to maintain a consistent time period between measurements in the consolidated database, some estimation of iceberg positions is necessary over short date gaps.

Iceberg positions for days without a measurement are estimated with a piece-wise cubic interpolation [14] between consecutive reported measurements. To maintain accuracy, we have chosen not to interpolate iceberg positions between measurements gathered more than two weeks apart. Interpolated position data are provided alongside the original data in the compiled database. An illustrative example is provided in Fig. 5.

To ensure accurate interpolation, an aggregate of all available recorded sensor measurements for a given track is used to determine the curve of best fit between the two endpoints of any data gaps in the track. Multiple sensor measurements that fall on the same day are averaged together to produce a single position for that day. The averaged positions are used in the interpolation process to ensure that the final interpolated track is as close to all reported sensor positions as possible, reducing any errors in the reported positions of the database.

In the consolidated database, the location of an iceberg on any day of its lifetime is given using this track derived from the position estimates in the original databases. With the assumption that the center of the iceberg lies close to the reported position in the database, estimates of its size and other derivative measurements are determined using scatterometer backscatter images.

IV. SIZE ESTIMATION

To estimate the size of an iceberg, we first find a contour of its surface in the σ^o image. An iceberg's contour is a closed curve that completely encompasses the iceberg as seen from

above; i.e., from a satellite's perspective. Estimating the size of an iceberg from its contour requires fitting the contour as closely as possible to the iceberg outline while removing any image pixels that do not belong to the iceberg. This involves discriminating between the iceberg's σ^o values and surrounding sea ice and ocean water, as well as land and near-by icebergs.

Over the entire polar region, some areas have σ^o values similar to the iceberg. In order to isolate an individual iceberg, we focus on an area around the iceberg, centered on its archived position. From that area, a mapping function creates a feature vector that we use to identify the pixels belonging to the iceberg of interest. A contour is then drawn around the iceberg pixels and used to compute the area of the iceberg itself, with some filtering done on the end results to improve accuracy and reduce noise. This procedure is described in detail in the following subsections.

A. Windowing

In order to isolate an individual iceberg, a windowed area is first selected around the iceberg's reported position. The windowed area is arbitrary; we have empirically determined that a 150×150 pixel (660×660 km) area is sufficient to effectively distinguish an iceberg from its surroundings while being large enough to cover the largest icebergs and minimize computation time. This enables us to represent the local backscatter image as a 150×150 matrix A of backscatter values.

The same windowing process is also completed for a separate matrix V , composed of the calculated variance for each pixel in A . The V matrix values serve to help distinguish between iceberg ice and surrounding ocean. The V values are computed by determining the spatial response function of each measurement used in the image formation algorithm [10], weighting the pixel values in the backscatter images by the spatial response function, summing the weighted values, and then, computing the variance of the difference between the summed values (which is sometimes termed the "forward projection") and the σ^o measurements. In the open ocean, wind changes from pass to pass and variation in azimuth angle cause the V value for ocean water to be large, whereas iceberg ice backscatter is temporally and azimuthally essentially constant, and so V has a smaller value. Thus, V provides discrimination capability to augment the differences in σ^o .

B. Mapping

A value mapping function is defined to enable discrimination between the iceberg σ^o and V values and other features inside the window. To find the pixels inside of the iceberg contour, all elements e in A and V are mapped according to

$$e_{ij} = |e^* - e_{ij}|^2$$

where e^* is the value of σ^o or the calculated variance averaged over the pixel at the archived position of the iceberg.

Mapping the values in this way increases the numerical distance between the iceberg e values and other values we do not want to include in the contour estimate, while also placing the archived e values of the iceberg at zero in our transformed coordinate system. An example of this is shown in Fig. 6. By extracting from each pixel a feature vector composed of the e values in the matrices A and V , and creating a covariance matrix

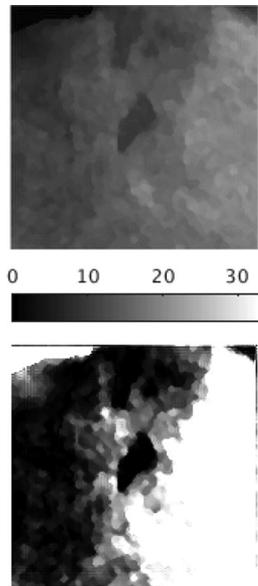


Fig. 6. Original σ^o image (top) in |dB| compared with the mapped e_{ij} values of iceberg A43c (bottom) in A from QSCAT data for JD 141 of 2001. The iceberg is in the center of each 660×660 km image. Note that the grayscale in this image is reversed so that the iceberg at the center of the image appears black.

using the feature vector values, we are able to calculate both the Euclidean and Mahalanobis distances of the pixel feature vector value from the origin. With generally little spatial variation in σ^o over the iceberg's surface, the pixels covering the iceberg can be discriminated from ocean and sea ice pixels using a simple threshold of the distances.

C. Contour Estimation

The iceberg is contoured in an iterative process using both distance metrics discussed in the previous section. The process begins by creating a separate matrix P composed of the inverse Euclidean distance values of all pixels in the windowed area. In this way, pixel values farther from the origin—the archived location of the iceberg—are weighted less heavily in P , the elements of which are considered to be related to the likelihood of a pixel belonging to the iceberg.

In order to determine which of the pixels in A and V belong to the iceberg, iceberg pixels are separated from the background image using an adaptive threshold determined by variational minimax optimization [15]. The optimization is done on an energy functional formed by an edge sensitive data term and a regularization term. All pixels in A with values greater than the threshold are flagged to create a binary image.

Once the binary image is created, image erosion and dilation steps are applied to remove any outlying flagged pixels. Binary image erosion and dilation steps remove any narrow necks that may be created by objects that are close enough to appear as a single object in the binary image [16]. In the erosion step, the area of the flagged pixels is reduced by one along the boundary of the area for each iteration. Any areas disconnected from the center of the image where the iceberg is located are then deleted with the aid of a region growing technique [16]. Then, dilation extends the area by one pixel per iteration, recovering the lost area of the iceberg.

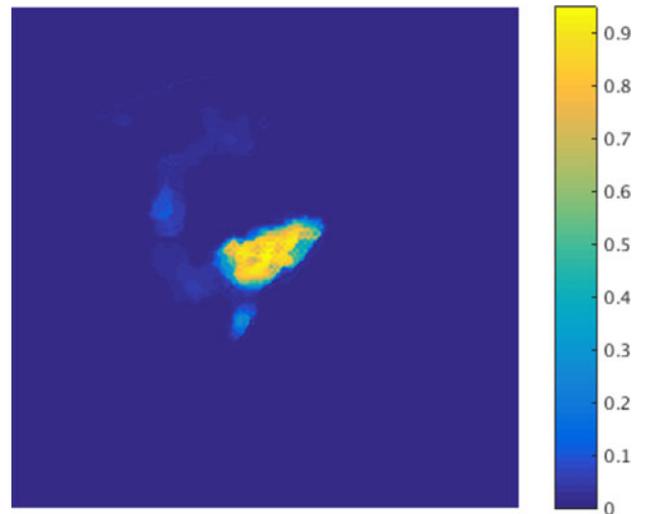


Fig. 7. Matrix P for iceberg B15b on JD 297 of 2005 (top), with its corresponding binary image (bottom), created by thresholding P . The iceberg itself has values close to one in P . Other nearby icebergs appear as groups with values less than one. Each image is 660×660 km in size.

The remaining flagged pixels are used to create a 2×2 covariance matrix R from the A and V values of each pixel. Using these values, the Mahalanobis distance of each pixel in the windowed area is calculated [17]. Mahalanobis distance is a measure of the number of standard deviations from a point to the mean of a Gaussian distribution. In this case, the distribution is centered at the origin and has a covariance of R . The inverse Mahalanobis distance values form the P matrix mentioned earlier, from which edge detection, thresholding, erosion, and dilation are applied to create the final binary image of iceberg pixels. Examples of P and the final binary image are given in Fig. 7.

A boundary line is then drawn around the iceberg pixels in order to extract the iceberg contour. The iceberg area is then computed with the aid of an array of precomputed pixel sizes, with the area value of each pixel inside the contour being added together to produce a size estimate in km^2 .

While this method removes most of the incorrectly selected pixels near the iceberg, it can fail to account for regions of pixels

belonging to other icebergs near the sides of the windowed area. These pixels are removed using an apodization window centered over the iceberg position.

D. Apodization Window

In choosing the location of the windowed area, we place the iceberg in the center of A . Any pixels far from the window center are assumed to not be part of the iceberg and are weighted in P by an apodization window, given as

$$w_{ij} = W \left(\frac{\sqrt{i^2 + j^2}}{2\sqrt{\max(i)^2 + \max(j)^2}} \right) \quad (1)$$

where

$$W(x) = \begin{cases} \frac{1}{2}1 + \cos\left(\frac{2\pi}{r}[x - \frac{r}{2}]\right), & 0 \leq x < \frac{r}{2} \\ 1, & \frac{r}{2} \leq x < 1 - \frac{r}{2} \\ \frac{1}{2}\{1 + \cos\left(\frac{2\pi}{r}[x - 1 + \frac{r}{2}]\right)\}, & 1 - \frac{r}{2} \leq x \leq 1 \end{cases} \quad (2)$$

is a tapered cosine window filter with a stretching coefficient r , determined as a ratio of the estimated contour's area to the window area, to adjust for iceberg size.

To avoid adding points to the contour that fall outside of the iceberg's shape, a simple moving average finite impulse response filter f is applied to P , where

$$q_{ij} = \sum_{k_1=-4}^4 \sum_{k_2=-4}^4 w_{ij} f(i - k_1, j - k_2) \quad (3)$$

with

$$f(i, j) = \frac{1}{16}. \quad (4)$$

The filtering smooths P , while emphasizing the iceberg pixel values.

E. Noise Filtering

Several different sources of noise affect size estimation using microwave images. Changes in temperature near an iceberg, for example, can affect the surrounding region's σ^o values due to melting or freezing sea ice [18]. At elevated temperatures, for example, sea ice or iceberg σ^o are reduced by surface melt, which reduces their contrast with surrounding ocean water. Large chunks of sea ice may have similar backscatter level to that of an iceberg undergoing surface melt, causing the sea ice to appear as part of the iceberg in a σ^o image. In addition, the quantization introduced by using discrete pixels in the image also discretizes the possible size estimates from the image. These are minimized by filtering the size estimates generated over time.

Noise sources are jointly accounted for by flagging all pixels in P within the empirically chosen value of $\pm 5\%$ of the minimax threshold function and setting these as the upper and lower bounds of possible sizes for the windowed iceberg on the given day. Size error is taken to be the difference between these bounds and the filtered size estimate for that day.

An iceberg typically only varies slowly in size through melting and shedding small fragments, with the exception of major calving events. To reduce noise and variability in the size estimates, the time series of size estimates are filtered in a

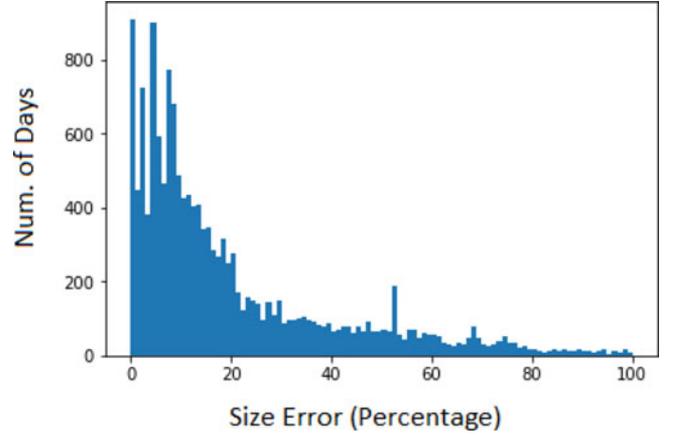


Fig. 8. Histogram of the percentage difference between 15 000 scatterometer-derived size estimates and NIC area estimate in the database. It should be noted that the original NIC size estimates are quantized to multiples of 5 km. Smaller icebergs tend to have a larger percent difference due to their smaller area, which makes single pixel misclassification a much higher percentage of the difference. The mean percentage difference is 20.1%, while the standard deviation of the percentage difference is 20.6%. More than 80% of icebergs have less than 20% difference.

two-step process. First, Hampel identification removes any outliers. Then, a Savitzky–Golay first-order filter [19] is applied to the estimates. The filtered output is on average much closer to the NIC's published sizes than the original unfiltered output.

A plot showing the percentage difference between this filtered output and NIC size estimates is shown in Fig. 8. While the difference is small for larger icebergs, small icebergs tend to have a higher percentage difference due to misclassified pixels adding a higher relative amount of area to their contours. This is reflected in high tail to the right in Fig. 8. Nonetheless, more than 80% of all comparisons have less than 20% error.

V. ROTATION ESTIMATES

From the time sequence of iceberg contours, rotational estimates are calculated. To do this, the generated contour for a particular day is compared to the previous day's contour, see Fig. 9. A least-squares ellipse is fitted to the contour [20] according to the canonical equation for a 2-D ellipse, given as

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0. \quad (5)$$

In the matrix form, this equation is

$$[x, y] \begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [D, E] \begin{bmatrix} x \\ y \end{bmatrix} + F = 0. \quad (6)$$

Eigenvalue decomposition of the first terms is used to find a rotation matrix, given as

$$\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix}^{-1}. \quad (7)$$

The rotation angle is

$$\rho = \tan^{-1} \frac{c}{s} \quad (8)$$

using the first eigenvector of the rotation matrix, which corresponds to the major axis of the ellipse. As ρ is considered to be the total angle of an iceberg from north as a function of time,



Fig. 9. Two contours of iceberg B15b on JD 071 (top) and 096 (bottom) of 2001. Note the rotation of the iceberg over the 25-day period. Each image is 660×660 km in size.

the daily rotational estimate is found as

$$\theta = \frac{d\rho}{dt} \quad (9)$$

differentiating over the period of one day, and smoothing the result with a Savitsky–Golay first-order filter.

A limitation of the method is that it does not account for single-day rotations of greater than 180° , which we believe are rare to nonexistent. Also, the contrast between an iceberg and its surroundings can distort the contour shape, which affects the accuracy of the rotation estimate. This is considered in the error bound calculation, given as follows.

A. Error Bounds

Several factors are taken into account in order to estimate error bounds of the rotation estimates, including the difference between contours, the R-square value of the fitted ellipse, the ellipticity of the least-squares fitted ellipse of each contour, and the derivative of the rotation estimate. These factors are defined as follows.

A turning function $\phi(s)$ is used to quantify the difference between two contours. $\phi(s)$ measures the change of angle ϕ along a given curve as a function of the normalized curve length s . $\phi(s)$ is derived from the difference between two contours can be taken as the L2 norm of the difference between their turning functions [21], given as

$$\epsilon = \left[\int_0^1 (\phi_1(s) - \phi_2(s + u) + \alpha)^2 ds \right]^{\frac{1}{2}} \quad (10)$$

where

$$\alpha = \int_0^1 [\phi_1(s) - \phi_2(s)] ds - 2\pi u. \quad (11)$$

The parameter u is the starting point of the turning function on the contour. Ultimately, it aligns the shape of the contour and the fitted ellipse. The parameter u is chosen by filtering and sampling each contour at the same points on $\phi_1(s)$ and $\phi_2(s)$, and taking as u the sampled point where the L1 norm between the two contours is the smallest.

Some error is introduced in the estimate of the iceberg size and rotation angle by the interpolation used to find contour points for the least-squares fitting algorithm. One measure of the goodness of fit of the ellipse to a given set of contour points is the R-square (R_s) of the fit. R-square is defined as the ratio of the sum of squares of the regression

$$\text{SSR} = \sum_{i=1}^n w_i (\hat{y}_i - \mu_y)^2 \quad (12)$$

where w_i is an array of weights applied to each contour point y_i , \hat{y}_i is the estimator of y_i , and μ_y is the mean of the y_i , and the total sum of squares

$$\text{SST} = \sum_{i=1}^n w_i (y_i - \mu_y)^2 \quad (13)$$

giving

$$R_s = \frac{\text{SSR}}{\text{SST}}. \quad (14)$$

R_s is a measure of how successful the fit is in explaining the variation of the contour. It describes the difference between the observed contour and the predicted contour from the fit, and may take on any value between zero and one. A value close to one means that the fit accounts for a greater proportion of the variance in the contour.

The ellipticity of iceberg shape also has bearing on the bounds of rotational error. Since an ellipse is fit to the generated contour, the difference between the major and minor axes of the fitted ellipse is correlated with the ability to detect rotational motion. This difference is quantified by using the eigenvalues of (7) to represent the length of the ellipse axes. The difference metric is given by

$$\lambda_d = |\lambda_1| - |\lambda_2| \quad (15)$$

where λ_1 and λ_2 are the maximum and minimum eigenvalues of the ellipse.

Because exact computation of the error bound is difficult for all cases, we resort to estimating an upper bound on the error using Monte-Carlo techniques. To determine an upper bound on the rotation estimate, thousands of Monte-Carlo simulations are run on binary images of various iceberg shapes. Binary images of a circle, along with several rectangles of different λ_d values, are also included. Monte-Carlo Gaussian noise with zero mean and σ^2 ranging from 0.01 to 0.9 is added to the binary images before the contour process to simulate noisy iceberg contours. By rotating these images over all possible angles, we find the difference e_s between the calculated rotational estimate and the modeled rotation. These data are used to empirically fit a parametric error bound to the parameters correlated with e_s . The bound shows the range of possible angles for that day,

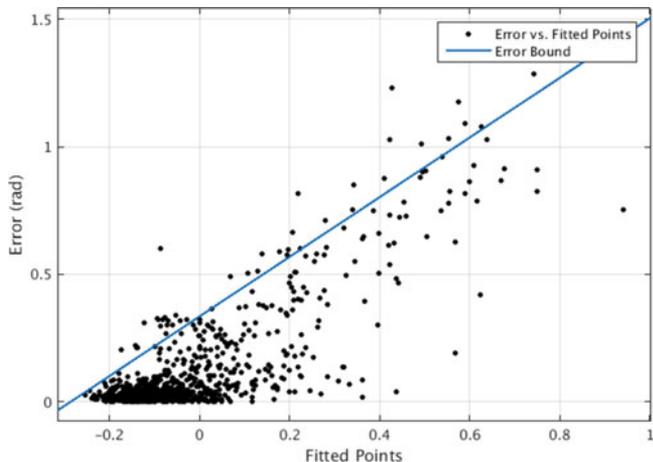


Fig. 10. Plot of the error e_s versus the sum of fitted parameters p for simulations done by rotating several iceberg images and elementary shapes. The blue line indicates the fitted line used to bound the predicted rotation error. The fit is chosen so that 90% of the simulated points fall beneath the fitted line. The error e_s has a correlation coefficient of 0.85 with p .

given the conditions described by the correlated parameters. The parametric error bound equation is found to be

$$e_\theta(p) = 0.9323p + 0.2763$$

$$p = 0.06f_s + 0.69|\theta| - 0.18\lambda_d - 0.004\epsilon - 0.04R_s \quad (16)$$

where p is the combined shape fit error, f_s is the difference between the filtered and unfiltered rotation estimates, θ is the derivative of ρ as shown in (8), λ_d is the difference between eigenvalues, ϵ is the shape error, and R_s is the R-square of the ellipse. The fit is weighted so that 90% of all simulated points fall below e_θ . A plot of e_θ versus the sum of fitted parameters p is shown in Fig. 10. We conclude from these simulations that iceberg rotational estimates made using this method are accurate to within 0.1 radians, given an accurate contour defined as a contour that yields an area within an arbitrarily chosen 5% of the NIC size estimate.

VI. DATABASE ANALYSIS

The new consolidated database is comprised of two sets of files. One contains all the original position reports from the original databases consolidated together with sizes and rotation angles determined from scatterometer data, with daily positions interpolated as needed. The second averages the position estimates to a single position per day, plus the size and rotation angle.

From daily track and size data, in this section, a simple statistical analysis is performed using the historical record. Such analysis is used to determine the movement of a typical iceberg over time, as well as information about its surroundings and calving conditions. Information about the database as a whole is used to find general trends in iceberg movement and location.

Analysis of individual icebergs in the database yields information about the average velocity and trajectory of that iceberg, as well as providing a means to determine its surface area. Over the life of an iceberg, large step reductions in area can be used to identify calving events.

Iceberg movement in aggregate is studied using the compiled dataset. We have observed, for example, that most icebergs follow a narrow counterclockwise path around the Antarctic coast.

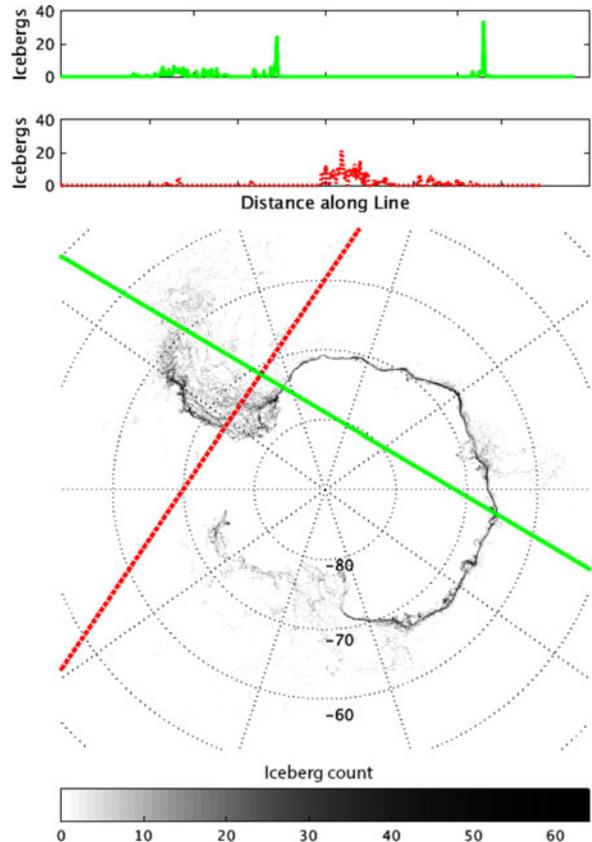


Fig. 11. (Top) Plots showing the number of icebergs moving through two line sections. The plots show the number of icebergs the database contains for each pixel along that line, starting from left of image. Note that the observed icebergs always moved counter-clockwise through these section lines. (Bottom) Histogram of the number of icebergs passing through each element of a 25×25 km pixel grid centered over the south pole. Compare with Fig. 1.

TABLE II
AVERAGE MOVEMENT PER DAY FOR ICEBERGS OF A GIVEN VALUE OF THE SURROUNDINGS MASK OVER THE ENTIRE DATASET

Mask Value	Class	Mean Daily Movement (km)
0	No data	N/A
1	Sea ice	4.83
2	Open ocean	5.02
3	Near land	1.45

This is most likely due to the Antarctic east wind drift, a near-coastal countercurrent related to the Antarctic circumpolar current or west wind drift [22]. Nearly 90% of all observed icebergs move along the coastline and eventually pass through the Weddell Sea [7]. Fig. 11 shows a density histogram of the number of icebergs that pass through each point of a 25×25 km grid over the life of the database. The near-coastal path taken by most of the icebergs is emphasized in the accompanying line plots.

We have observed from the database that iceberg velocities and trajectories are affected by the surrounding sea ice conditions. As a result, we include a “surroundings mask” in the data base. This is generated by manual examination of the radar backscatter images with one of four values assigned as shown in Table II. Differences in σ^o between an iceberg, sea ice, and open ocean are used to generate the surroundings flag value, which indicate the status of nearby land, ocean, or sea ice for

each iceberg. Flag values are assigned if more than 50% of the surrounding pixels are ocean or sea ice, or when more than 1/4 of nearby pixels are land.

Analyzing the data, we have found that icebergs traveling in sea ice tend to move more slowly than icebergs floating in the open ocean. Table II shows the average daily displacement of icebergs based on the value of the surroundings flag. The average displacements in the table confirm the high speed of icebergs in the open ocean. In contrast, near-land icebergs are often grounded or dragging the bottom, and have reduced average velocity compared to icebergs surrounded in sea ice or in the open ocean.

VII. CONCLUSION

The goal of this paper is to introduce a new, comprehensive database of iceberg location data. In this paper, we consolidate the two datasets into one comprehensive database using weekly NIC and daily BYU iceberg positional data. Consistent daily iceberg tracks are generated, and information about the surroundings of each iceberg as derived from scatterometer data are included in the new database. Using the iceberg positions, we have derived new automated daily iceberg size and rotation estimates from scatterometer σ^0 images. The new size estimates generally fall within 20% of the NIC's reported size estimates.

The consolidated database is comprised of two sets of files. One contains the original positions and interpolated daily positions, plus sizes and rotation angles and a surroundings flag manually derived from radar backscatter images. The second file contains a single position per day—when multiple sensors provide positions estimates on one day, the average of the positions is used. The new database is expected to enable easier analysis of the information that has been collected and to facilitate a deeper understanding of icebergs in the future. The consolidated database is available for download from www.scp.byu.edu/iceberg.

REFERENCES

- [1] A. O. Cefarelli, M. Vernet, and M. Ferrario, "Phytoplankton composition and abundance in relation to free-floating Antarctic icebergs," *Deep-Sea Res. II, Top. Stud. Oceanography*, vol. 58, pp. 1436–1450, 2011.
- [2] T. Scambos *et al.*, "Calving and ice-shelf break-up processes investigated by proxy: Antarctic tabular iceberg evolution during northward drift," *J. Glaciol.*, vol. 54, no. 187, pp. 179–191, 2008.
- [3] L. Evers, D. Green, N. Young, and M. Snellen, "Remote hydroacoustic sensing of large icebergs in the southern Indian Ocean: Implications for iceberg monitoring," *Geophys. Res. Lett.*, vol. 40, pp. 4694–4699, 2013.
- [4] J. Tournadre, N. Bouhier, F. Girard-Ardhuin, and F. Rémy, "Large icebergs characteristics from altimeter waveform analysis," *J. Geophys. Res., Oceans*, vol. 120, pp. 1954–1974, 2015.
- [5] K. M. Stuart and D. G. Long, "Tracking large tabular icebergs using the SeaWings Ku-band microwave scatterometer," *Deep Sea Res., II, Top. Stud. Oceanography*, vol. 58, no. 11, pp. 1285–1300, 2011, doi:10.1016/j.dsr2.2010.11.004.
- [6] F. T. Ulaby and D. G. Long, *Microwave Radar and Radiometric Remote Sensing*. Ann Arbor, MI, USA: UMP, 2014.
- [7] D. G. Long, J. Ballantyne, and C. Bertoia, "Is the number of icebergs really increasing?" *Trans. Amer. Geophys. Union*, vol. 83, no. 42, pp. 469–474, 2002.
- [8] D. G. Long, "Polar applications of spaceborne scatterometers," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 3, pp. 2307–2320, May 2017.
- [9] D. G. Long, M. R. Drinkwater, B. Holt, S. Saatchi, and C. Bertoia, "Global ice and land climate studies using scatterometer image data," *EOS Trans. Amer. Geophys. Union*, vol. 82, no. 43, p. 503, Oct. 2001.

- [10] D. S. Early and D. G. Long, "Image reconstruction and enhanced resolution imaging from irregular samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 291–302, Feb. 2001, doi:10.1109/36.905237.
- [11] Q. D. Remund and D. G. Long, "A decade of QuikSCAT scatterometer sea ice extent data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4281–4290, Jul. 2014, doi:10.1109/TGRS.2013.2281056.
- [12] "Antarctic ice shelf sheds bergs," Apr. 2016. [Online]. Available: <http://earthobservatory.nasa.gov/IOTD/view.php?id=87859&src=%C2%AADiotd>
- [13] S. Hennig, "Exploring the benefits of active vs. passive spaceborne systems," *Earth Imag. J.*, vol. 10, no. 6, Dec. 2013.
- [14] F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM J. Numerical Anal.*, vol. 17, pp. 238–246, 1980.
- [15] N. S. Baidya and R. Nilanjan, "Image thresholding by variational minimax optimization," *Pattern Recog.*, vol. 42, no. 5, pp. 843–856, 2009.
- [16] Q. P. Remund, D. G. Long, and M. R. Drinkwater, "An iterative approach to multisensor sea ice classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 4, pp. 1843–1856, Jul. 2000, doi:10.1109/36.851768.
- [17] P. C. Mahalanobis, "On the generalised distance in statistics," *Proc. Nat. Inst. Sci. India*, vol. 2, no. 1, pp. 49–55, 1936.
- [18] C. Wesche and W. Dierking, "Iceberg signatures and detection in synthetic aperture radar (SAR) images in two test regions of the Weddell Sea, Antarctica," *J. Glaciol.*, vol. 58, pp. 325–339, 2012, dx.doi.org/10.3189/2012JOG11J020.
- [19] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least square procedure," *Analytic Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964, doi:10.1021/ac60214a027.
- [20] A. W. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least squares fitting of ellipses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 476–480, May 1999.
- [21] D. Cakmakov and E. Celakoska, "Shape matching of digital curves," in *Proc. 4th Int. Symp. Image Signal Process. Anal.*, Zagreb, Croatia, 2005, pp. 457–461.
- [22] S. T. Gille, "Mean sea surface height of the Antarctic circumpolar current from GEOSAT data: Methods and application," *J. Geophys. Res.*, vol. 99, pp. 18225–18273, 1994.



Jeffrey S. Budge was born on June 27, 1990, in Logan, UT, USA. He received the B.S. and M.S. degrees in electrical engineering from Brigham Young University, Provo, UT, in 2016 and 2017, respectively. He currently works as a Radar Algorithms Engineer with Artemis, Inc., Spanish Fork, UT.



David G. Long (S'80–SM'98–F'08) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1989.

From 1983 to 1990, he worked with NASA's Jet Propulsion Laboratory (JPL), Pasadena, CA, where he developed advanced radar remote sensing systems. While at JPL, he was the Project Engineer on the NASA Scatterometer (NSCAT) project, from 1996 to 1997. He also managed the SCANSAT project, the precursor to SeaWinds in 1999 on QuikSCAT, in 2002 on ADEOS-II, and as RapidScat in 2014 on the International Space Station. He is currently a Professor with the Electrical and Computer Engineering Department, Brigham Young University (BYU), Provo, UT, USA, where he teaches upper division and graduate courses in communications, microwave remote sensing, radar, and signal processing and is the Director with the BYU Center for Remote Sensing. He is the Principle Investigator on several NASA-sponsored research projects in remote sensing. He has more than 400 publications in various areas including signal processing, radar scatterometry, and synthetic aperture radar. His research interests include microwave remote sensing, radar theory, space-based sensing, estimation theory, signal processing, and mesoscale atmospheric dynamics.

Dr. Long was the recipient the NASA Certificate of Recognition several times and is an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.