Heliyon 5 (2019) e02570

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.heliyon.com

Research Article

Applying single-image super-resolution for the enhancement of deep-water bathymetry

Kristen Nock [a], David Bonanno [a], Paul Elmore [b], Leslie Smith [a], Vicki Ferrini [c], Fred Petry [d,*]

[a] Naval Research Laboratory, Washington DC, USA
[b] Johns Hopkins University, Applied Physics Laboratory, Laurel, MD, USA
[c] Lamont-Doherty Earth Observatory Columbia University, Palisades, NY, USA
[d] Naval Research Laboratory, Stennis Space Center, MS, USA

ARTICLE INFO

Keywords:
Computer science
Earth sciences
Single-image super-resolution
Upscaling
Bathymetry

ABSTRACT

We present research using single-image super-resolution (SISR) algorithms to enhance knowledge of the seafloor using the 1-minute GEBCO 2014 grid when 100m grids from high-resolution sonar systems are available for training. We performed numerical experiments of x15 upscaling along three midocean ridge areas in the Eastern Pacific Ocean. We show that four SISR algorithms can enhance this low-resolution knowledge of bathymetry versus bicubic or Splines-In-Tension algorithms through upscaling under these conditions: 1) rough topography is present in both training and testing areas and 2) the range of depths and features in the training area contains the range of depths in the enhancement area. We quantitatively judged successful SISR enhancement versus bicubic interpolation when Student's hypothesis testing show significant improvement of the root-mean squared error (RMSE) between upscaled bathymetry and 100m gridded ground-truth bathymetry at $p < 0.05$. In addition, we found evidence that random forest based SISR methods may provide more robust enhancements versus non-forest based SISR algorithms.

1. Introduction

Given that 70% of the Earth is covered by oceans, mapping of the ocean floor is fundamental to geophysical understanding of the Earth. Quantitative measurements and maps of seafloor depth – bathymetry – and related topographic metrics is far more difficult than mapping terrestrial surfaces due to the need to use sonar instead of electro-optical techniques. Publicly available bathymetric data is sparse, however, as sonar data is slow and expensive to collect. Over 82% of Earth's ocean floor in ice-free regions remains unmapped by sonar systems (Sandwell et al., 2014; Weatherall et al., 2015) in publicly available data. As a result, we know more about the topography of Mercury, Venus, Mars, larger moons of Jupiter and Saturn – and now Pluto and Charon – than we do about Earth's surface covered by water.

What mappings that we have for the remainder of Earth's deep-water (more than 1 km depth) bathymetry in ice-free regions are only predictions of bathymetry. The conventional prediction methodology uses Newtonian gravity potential theory and decades old Fourier transforms to invert measured geoid height to predicted seafloor topography (Parker, 1972; Smith and Sandwell, 1994, 1997; Hu et al., 2015). This is an ill-posed problem. Soundings (sonar measured bathymetry) that exist within inversion areas constrains possible solutions. The current maps of world bathymetry from the General Bathymetric Charts of the Oceans (GEBCO) incorporates this satellite-altimetry predicted gridded bathymetry (Weatherall et al., 2015) at 1-minute resolution, representing low-resolution (LR) knowledge of the global ocean floor.

This practice, however, has physics-based limitations on the resolvable resolution and accuracy of potentially solutions. Despite the 1-minute GEBCO mapping, the altimetry-based portions of the grid is generally limited to 20-km resolution (Marks and Smith, 2012) (potentially 10-km with newer satellites and repeated orbits (Marks and Smith, 2016)) compared to 100-m resolution of deep-water sonar-based mapping systems. Currently, no other deep-water techniques provide 100m-resolution of seafloor bathymetry. Continued reliance on these techniques is due, in part, to the desire to obtain analytic/computation solutions based on first-principles that preserve intuition of the underlying physics and the ability to adjust approximations in the modeled conditions.

* Corresponding author.
E-mail addresses: frednavy3@gmail.com, fred.petry@nrlssc.navy.mil (F. Petry).

https://doi.org/10.1016/j.heliyon.2019.e02570
Received 19 June 2019; Received in revised form 9 August 2019; Accepted 30 September 2019
2405-8440/© 2019 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Enhancement of knowledge beyond the limitations of these methods, however, is critical to both military and civilian interests. For example, bathymetry for the area where Malaysian Airlines Flight 370 may have disappeared (Smith and Marks, 2014; Picard et al., 2017) showed nearly complete lack of sonar data at the time of the disappearance. Only the low-resolution satellite-based information existed. Furthermore, it is known that the majority of civilian air flights occur over oceanic waters that similarly are unmapped by sonar. While the Indian and Southern Oceans are now being newly surveyed as a result of the ML370 disappearance, this tragedy exemplifies how inadequate deep-water bathymetry can hinder maritime emergency, search, and recovery operations.

Thus, we turn to examining developments of new computational techniques, particularly from machine learning, that could enhance our predictions of bathymetry in these unsurveyed regions of Earth. The challenge is to maximize knowledge obtainable from LR information and available high-resolution. Super-resolution imaging (SR) is a class of techniques that enhance resolution of images by producing higher-resolution (HR) images based on given low-resolution (LR) images (Milanfar 2010). The underlying concept in SR is to use the non-redundant information in low-resolution images to produce enhanced high-resolution images. Another approach is known as single-image super-resolution (SISR) based on the concept that images usually contain redundant, repetitive content. In particular, small image patches in natural images tend to recur redundantly many times inside images, both within the same scale, and across different scales. In SISR algorithms these patches are used to enhance image resolution for natural images. These upscaling are non-unique, however, yet are satisficing for the application.

In this paper we show how to adapt single-image super-resolution techniques developed in the image-processing domain to processing deepwater (>1-km) seafloor topography. We were interested in determining whether available HR information can be leveraged to enhance LR knowledge of the seafloor. We hypothesized that SISR computer vision algorithms can enhance LR knowledge of bathymetry through upscaling. Motivating this research was success with SISR for image processing producing high-resolution details in natural images. Additionally SISR was used to enhance morphology details of galaxy images (Ball and Brunner, 2010; Schawinski et al., 2017) which have similarity to seafloor features upon treating gridded bathymetry as a gray-scale image. Additionally SISR can enable quantitative extrapolation of HR information from MBES surveys into neighboring areas covered only by predicted bathymetry. These algorithms can be trained to associate HR information onto lower-resolution versions of the surface.

Enhancement is assessed by root-mean squared error of upscaled LR bathymetry being significantly lower using SISR versus linear-invariant interpolation methods which do not leverage HR information for upscaling processes as quantified by hypothesis testing. To test this hypothesis, we performed numerical experiments of upscaling three areas of the Eastern Pacific Ocean along mid-ocean ridge systems. These areas have HR grids obtained by sonar surveys that serve for training SISR algorithms and ground-truth of upscaled LR grids for both SISR and linear-invariant interpolation methods. We tested SISR skill against our benchmark skill level of interpolation in both localized and external upscaling experiments (external upscaling here means training SISR techniques in one area to upscale other separate and distant areas). We found that four of the SISR algorithms have higher skill than interpolation under the conditions that we will discuss in this paper. It should be stressed here that the enhancement is in accuracy only, not enhancement of resolution. Being an ill-posed problem, any higher resolution fabric of the ocean floor is merely a prediction.

Section 2 of the paper provides background on previous work in topographic metrics and computer vision for digital bathymetry models and specific SISR approaches. Section 3 describes data used for experimentation and implementation details for SISR techniques. Section 4 reports on experiments that applied SISR to the bathymetry data described and discusses experimental observations including our discovery of significantly more accurate bathymetry using SISR versus bicubic interpolation. Section 5 discusses the experimental results, including conditions in which we expect SISR to be able to enhance LR bathymetry knowledge. Section 6 provides a summary and our conclusions.

2. Background

2.1. Previous work on seafloor-topography upscaling

Geophysical data, including bathymetry, is often sparse and irregularly spaced. Gridding interpolation algorithms traditionally map this sparse data to a gridded surface. One such interpolation algorithm is Splines-In-Tension (SIT) (Smith and Wessel, 1990; Wessel and Bercovici, 1998). SIT is the gridded surface generator used in Generic Mapping Tools (GMT) [Wessel et al., 2013] and widely used in the scientific community (GMT's "surface").

There are applications to undersea features of quantitative geomorphological techniques used for land data. Micallef et al. (2007) used seafloor slope, aspect, plan curvature, and profile curvature to classify distinct topographic areas of the Storegga Slide in the Norwegian Sea through a decision tree process. Another, Wormald et al. (2012), used morphometric techniques in combination with functional analysis of metrics for semi-automated quantitative classification of morphological structures on the submarine Monowai Volcano, such as lava flow fields, ridges, cones, and faults. Morphometric functions used included slope, aspect, hydrology tools, and curvature functions. These were applied to the multibeam bathymetry to aid analysis of structural features, and improve accuracy and repeatability of volcano interpretation. The functions all highlight subtle changes in the bathymetry such as directional (aspect) or the extent of change of the seafloor slope), while the hydrology tools and curvature functions identify sharp changes in bathymetry (i.e., structure edges). These approaches provide a proof of concept, i.e., evidence that different morphologies along the seafloor can be quantitatively and consistently classified based on topographic metrics.

A logical extension of seafloor metric techniques (Micallef et al., 2007) and land classification (Iwahashi and Pike, 2007; Marmion et al., 2009) is to apply machine-learning algorithms for automated classification of undersea features. Two studies used worldwide predicted bathymetry from Smith and Sandwell (1994). In one study, Gorini (2009), applied morphometric quantities in (Micallef et al., 2007) on predicted bathymetry to produce a worldwide map of nine different global geomorphology types. Gorini used the ISODATA unsupervised classification algorithm (Mather and Koch, 2011) to cluster undersea regions into nine categories. More recently, Valentine et al. (2013) used a neural network auto-encoder approach for seamount and provided automated identification of seamounts in Smith and Sandwell (1997) bathymetry with a true positive skill rate of 72.3%. How to discriminatively learn a low-dimensional feature space, in which the mapped features have small within-class scatter and big between-class separation, is a challenging problem. Zhou et al. (2019) describe a technique, compact and discriminative stacked auto-encoders for this problem.

The problems of within-class diversity and between-class similarity are still two big challenges. A method to learn discriminative CNNs to boost the performance of remote sensing image scene classification was developed by optimizing a new discriminative objective function for training (Cheng et al., 2018b). A recent study used sonar data with machine learning (Lawson et al., 2017), applying decision forests and extra-tree classifiers (Geurts et al., 2006) to identify seamounts and mid-ocean ridge segments along the East Pacific Rise (EPR). Using labeled training data based on localized statistical metrics for bathymetry and topography they achieved highest classification skill with test data from the Pacific, 97%, using the extra-tree classifier.

2.2. Single-image super-resolution

Gridded bathymetry can be treated as a continuous grayscale image for computer vision approaches from image processing for increasing image resolution (Milanfar, 2010). This problem is called "super--resolution". Early methods combined multiple frames, where each frame was slightly offset from others. Example-based methods for super-resolution (Freeman et al., 2002) led to single-image super--resolution (SISR) where image upscaling is based on a LR-HR dictionary (i.e. basis space). SISR is an ill posed problem, because many unique HR images can share a single LR image, a process that is not invertible. The SISR methods assume a mapping between the two resolutions, both for internal examples at various scales and similar external examples (Yang et al., 2014).

Next we discuss methods used in our research. An early SISR method is neighborhood embedding (NE) methods (Chang et al., 2004). This assumes that LR and HR patches lie on similarly shaped manifolds and that new pairs of data are described as a local linear combination of existing examples on the manifold. Variants of this include NE with least squares (NE-LS) and NE with the addition of nearest neighbor least squares (Yang et al., 2008) showed how to use sparse representation for super-resolution by linking a pair of LR and HR dictionaries and assuming the same few columns that approximate the LR image patch will create a HR image patch. Zeyde et al. (2010) improved on this by achieving higher quality in terms of peak signal-to-noise ratio (PSNR) and fewer artifacts. Wu and Bai (2018) used dictionary learning to produce sparse representations for seismic imaging data.

More recently (Timofte, et al., 2013) proposed two methods: (1) the anchored neighborhood regression (ANR) and (2) global regression (GR). ANR obtains comparable quality to early methods while almost an order of magnitude faster, and GR gains additional speed by sacrificing quality. Timofte et al. (2014) improved ANR with an adjusted anchored neighborhood regression (A+), similar to super-resolution by simple functions (SF) (Yang and Yang, 2013); i.e., by regressing on all training data rather than a dictionary. The improved method, termed jointly optimized regressors (JOR) (Dai et al., 2015), uses k-nearest neighbors to associate an optimal regressor for each image patch and then uses a regressor for super-resolution.

3. Design

3.1. Data

We used collocated LR and HR bathymetry data for our experiments. The LR data is 1-minute bathymetry from the 2014 GEBCO grid (Weatherall et al., 2015). The collocated HR data is from side-scan sonar and MBES datasets (Carbotte et al., 2004, http://www.marine-geo.org) and Global Multi-Resolution Topography GMRT Synthesis (Ryan et al., 2009, http://www.geomapapp.org). We considered three mid-ocean ridge regions in the Pacific Ocean. These regions are along the Juan de Fuca Ridge (JUAN) and two sections of the East Pacific Rise (EPR) Ridge, North EPR (NEPR) and South EPR (SEPR) regions. The eastern sections of the Pacific are: (1) deep enough regions for satellite altimetry to predict bathymetry; (2) the regions along ridges are shallower than other parts of

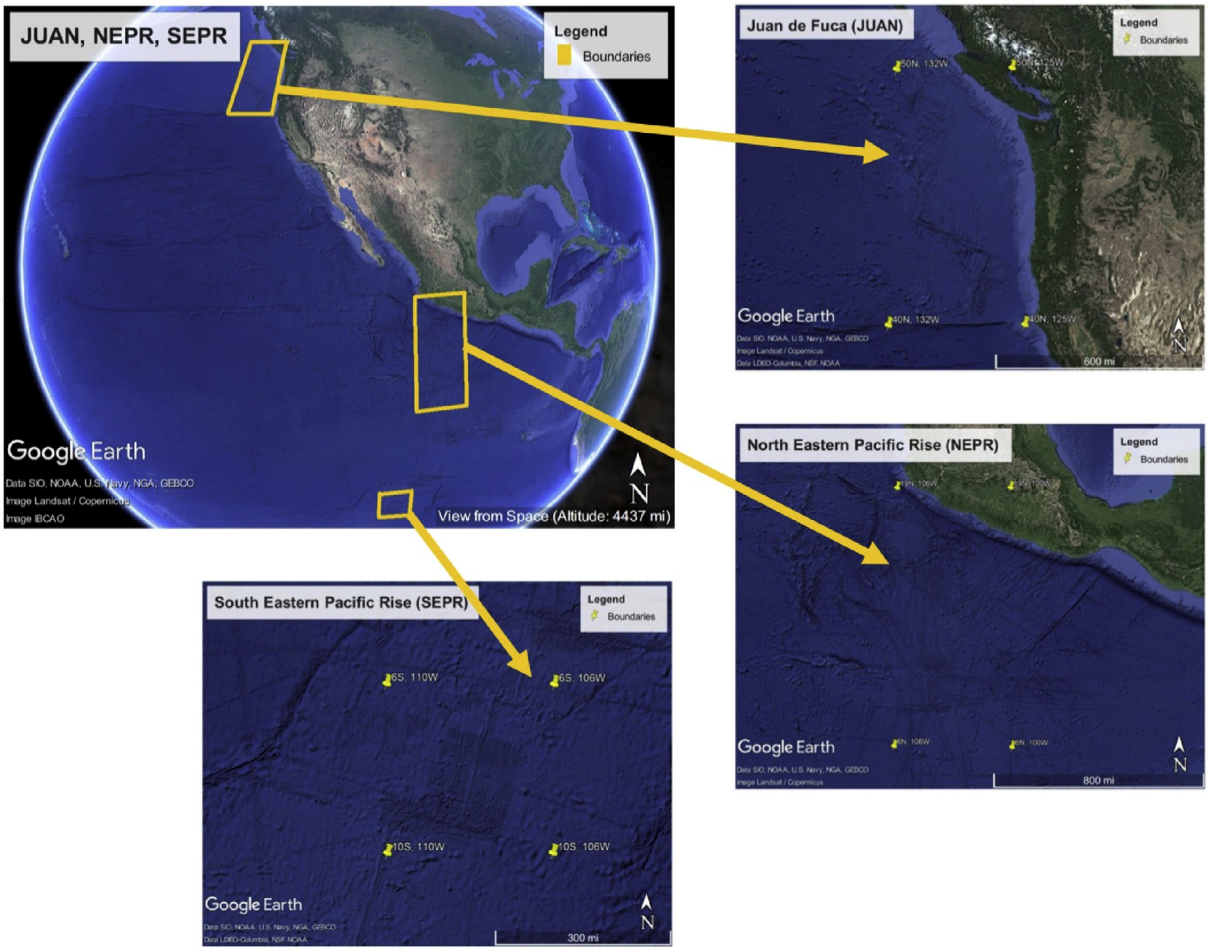

Fig. 1. Areas used in this paper from the (top row) Juan de Fuca Ridge (middle row) North East Pacific Rise Ridge (bottom row) South East Pacific Rise Ridge. Respectively, these area are labeled JUAN, NEPR, and SEPR.

the Pacific so that affects of upward continuation are lessened, and (3) available data of the regions are abundant as shown in Fig. 1. The JUAN area is an intermediate spreading center, at the boundary between Juan de Fuca Plate and Pacific plate, with a spreading rate of 56 mm/yr (Wilson, 1993). Gridded bathymetry data is available as part of GMRT Synthesis (Ryan et al., 2009), from multibeam sonar data from 36 research cruises (1981–2009). The spreading center has a well-defined ridge with a relatively broad axial valley at its crest.

The NEPR area is on the East Pacific Rise off the Pacific coast of Central America (8°N-18°N). It includes 300-m horizontal resolution gridded bathymetry, which is a combination of data from side-scan phase bathymetry sonar and multibeam sonar system surveys (Macdonald, 2012).

The SEPR area is 1700 km to the southwest from 7°-9°S and was mapped with multibeam sonar in 1991 (Cochran et al., 1993). This data has a grid resolution of 100-m and covers roughly 150 km to each side of ridge axis. Bathymetry of SEPR is broadly similar to that of the NEPR, but while depth of the ridge crest varies at the NEPR, it is relatively constant at the SEPR (Cochran et al., 1993). The data set's northern edge is approximately 1700 km southwest of the southernmost part of NEPR coverage. Fig. 2 shows for the areas of interest, the high-resolution gridded sonar data obtained from global multi-resolution topography

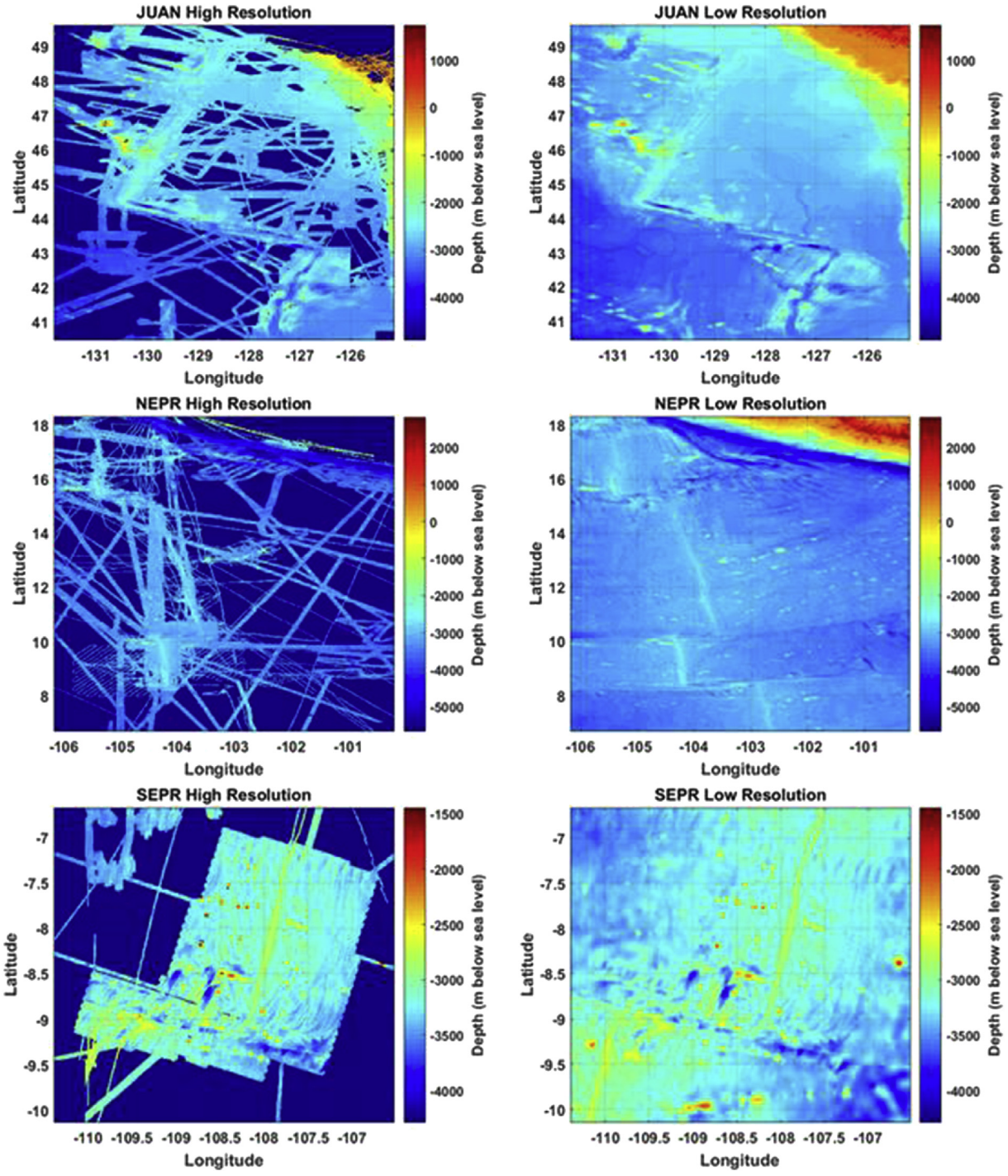

Fig. 2. In each row, the inset to the left is high-resolution gridded sonar data obtained from global multi-resolution topography (GMRT), and low resolution seafloor topography from the general bathymetric chart of the oceans (GEBCO 2014) grid is to the right.

(GMRT) (Ryan et al., 2009) is on the left, and low resolution seafloor topography from the general bathymetric chart of the oceans (GEBCO 2014) grid Weatherall et al. (2015) is to the right.

We collected corresponding HR and LR bathymetry data from these three areas, first converting LR and HR data from the angular coordinate system (latitude and longitude) to Cartesian UTM coordinates. We then upsampled the LR data from a 1/15 resolution difference of the HR to 1/6 resolution difference using bicubic interpolation (Press et al., 2007, Sec. 3.6.3). Next we aligned the three data sets to overlapping points using a triangulation-based linear interpolation. Finally, we parsed data into corresponding overlapping 10 × 10 LR blocks and 60 × 60 HR blocks (Fig. 3) resulting in 1127, 2455, and 4232 blocks respectively.

From these we plotted depth and roughness distributions (Fig. 4) for the three regions. The depth distribution, or hypsometry, came from all depth nodes with the corresponding LR (Fig. 4a) and HR (Fig. 4b) blocks. Roughness is standard deviation of depths within a patch (Fig. 4c and 4d for the LR and HR patches). We processed these through the kernel density-smoothing algorithm (Bowman and Azzalini, 1997) as implemented in the "ksdensity" function of the Matlab Statistics and Machine Learning Toolbox, (2017).

Topographically, JUAN contains shallower depths of the three distributions. NEPR and SEPR have similar depths, with a wider distribution of depths in NEPR. The JUAN and NEPR regions exhibit multimodal hypsometry, while the SEPR region is largely unimodal. Depth distributions between the LR and HR versions appear similar. Roughness for JUAN shows a largely smoothed region compared to NEPR and SEPR regions. The roughness peak for JUAN is close to zero and LR and HR distributions appear similar. The NEPR and SEPR roughness distributions are different. The peaks for both are further away from zero than the JUAN peak. The roughness also increases for both when going from the LR version to the HR version.

3.2. Single image super-resolution methods

In this paper we compare the upscaling skill of eight SISR methods against the skill of four interpolations algorithms (Table 1). Four of the methods use data invariant linear filters and include SIT 1990, SIT 1993, SIT 1998, and bicubic interpolation. The other eight methods require training from extracted features of sample coincident LR and HR patches: NE, NE + LS, NE + NNLS, Zeyde et al., ANR, GR, A+, JOR, and SRF. These SISR methods are trained with the same features and procedures. To account for the inherent differences in the bathymetry data, we treated the HR and LR blocks as a single grayscale image, in which we treated depth as the intensity value of an image. We designed the training dictionary to a size of 512 and trained using the approach presented in Zeyde et al., [2010]. During the SISR algorithm training phase, the K-SVD dictionary training algorithm used a sparsity-based minimization with a target sparsity of 20. The sparsity was a value that was designed to be changed, and appears to be application specific. In the original program, the previous author left a comment that the value of this variable is "TBD". We needed to alter the codes to perform on continuous depth maps instead of [0, 255] binned data. The original codes used were designed originally to work on images, to adapt them to allow their use on topological data we needed to change the target sparsity value.

Standard SISR techniques often involve iteratively learning a mapping between corresponding HR and LR datasets. For natural imagery, training the SISR methods begins with HR imagery, which is down sampled to create the corresponding LR image. This process inherently maintains alignment between the LR and HR data. The data is also bound between minimum and maximum intensities with pixel values

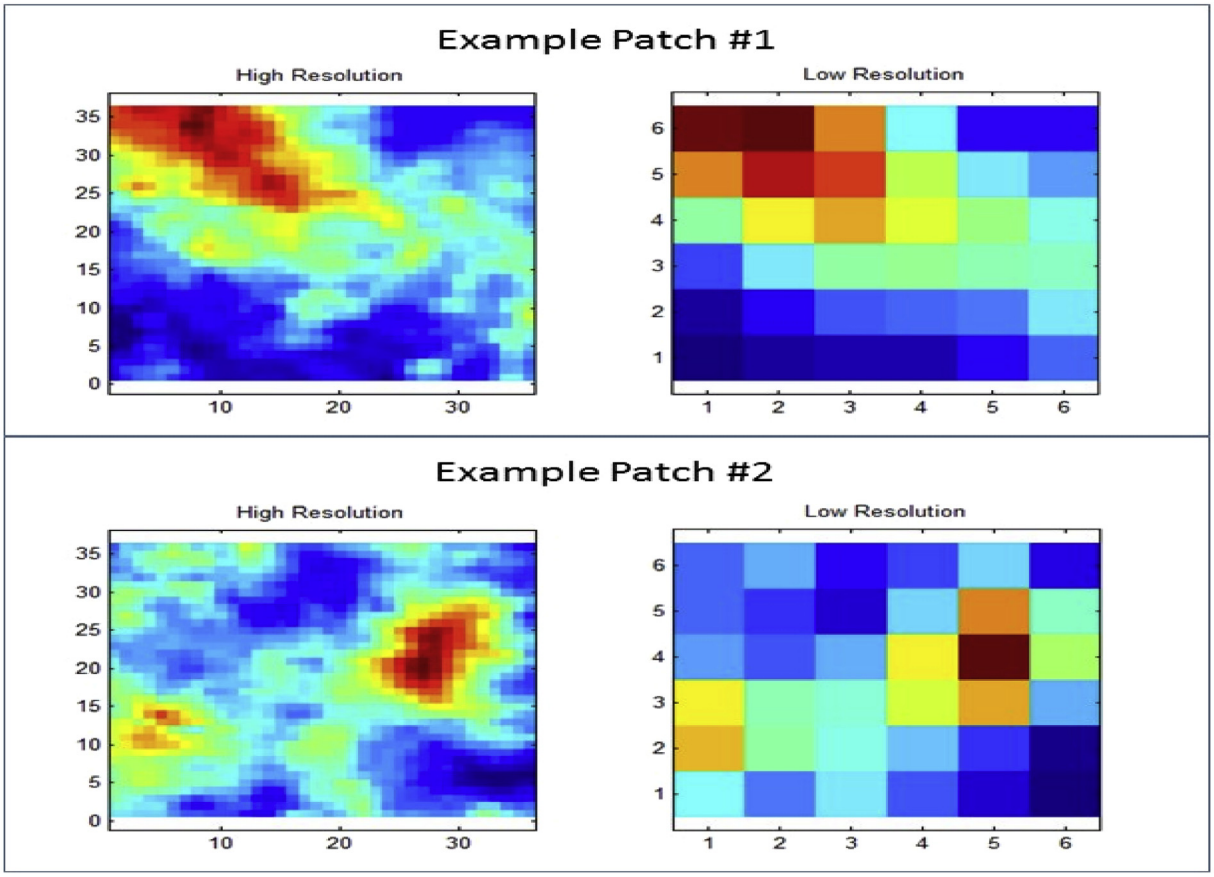

Fig. 3. Two examples of high-resolution (HR) and low-resolution (LR) patches from the SEPR data sets. There is a scale difference of x6 between the HR and LR patch data.

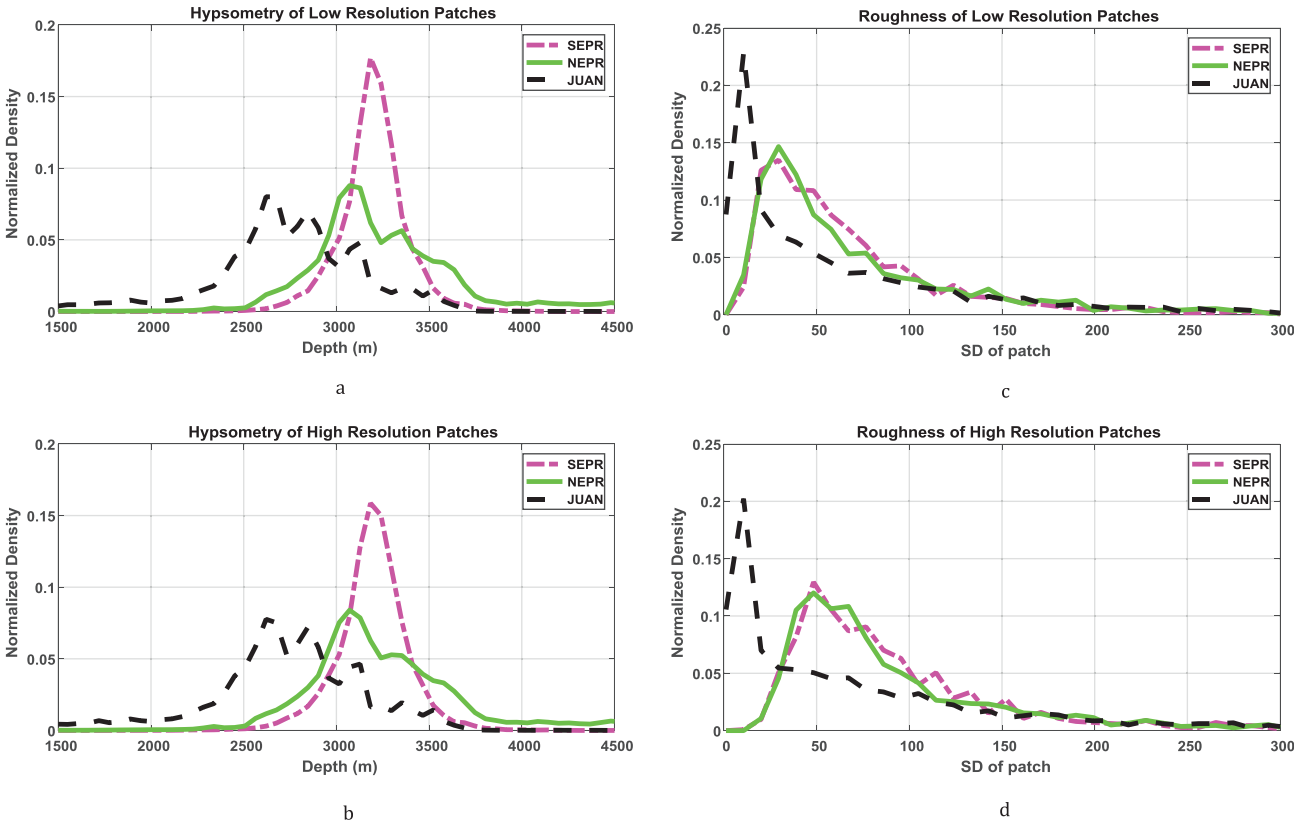

Fig. 4. Hypsomtery and roughness distributions for the JUAN, NEPR and SEPR areas. Fig. 4a and 4b are hypsomtery for the LR and HR grids, respectively. Fig. 4c and 4d are roughness distribution for the LR and HR grids.

traditionally being an 8 bit integer valued between 0 and 255. In order to apply the SISR techniques to the bathymetry data, the SISR approaches and bathymetry data required the following modifications.

3.2.1. SISR Modifications

Unlike natural images, the HR and LR data are measured from independent systems (sonar and satellite altimetry, respectively), which estimate ocean depth. The SISR approaches are customized to handle two data sources (HR and LR separately). With these changes, SISR techniques now function to map not only resolution differences, but compensate for sensor differences as well. The multiscale procedure for the A+, JOR, and SRF methods was not used in this work. As discussed, the programs used were designed originally to work on images. The original codes used a scaling feature parameter for performing bicubic interpolation on the images to make recursively scaled images. The original magnification code was set to x3. This process works better on larger images, as the scaling of both the high resolution and low resolution patches has a positive impact.

However this was not the case for our application. It created an issue of mismatched scaled high and low res patches. If we were to perform additional down sampling this would introduce undesirable noise. By removing the process we obtained more desirable results.

Specifically, JOR is set up to use 5 million training samples, 32 joint regressors learned by 20 iterations of the EM algorithm outlined in Dai et al. (2015). Furthermore, all SISR methods are set up to operate with a window size of 3×3 pixels with an overlap of two pixels.

3.2.2. Bathymetry Modifications

Bathymetry data was also modified to work with SISR algorithms. Since the SISR algorithms used were designed for natural images, training and testing data was adjusted to fit the requirements. HR and LR bathymetry blocks (represented in natural depths) were normalized between 0 and 1 by dividing recorded depths by depth of the Mariana Trench (36,200 feet deep). Normalization was applied before training of SISR methods and reversed after their application to obtain depth estimates on the same scale as original data. Also to minimize quantization errors depth values are kept in 64-bit resolution so HR and LR data can be encoded without adding significant quantization noise.

4. Results

Three sets of experiments focused on trends and limitations of applying SISR techniques to bathymetry. We compared the performance of SISR techniques quantitatively in terms of rootmean-square error (RMSE Eq. 1).

$$RMSE = \sqrt{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2 / n} \quad (1)$$

In Eq. 1 $\mathbf{y_i}$ represents ground truth soundings from MBES mapping systems, $\hat{\mathbf{y}}_\mathbf{i}$ corresponding estimates from performing SISR of coincident LR data, and $\mathbf{n}$ is the total number of individual soundings extracted from the patch in a region. Overall, we used the RMSE as a metric to compare error in estimated HR from observed ground truth.

We then used mean and standard deviation of the RMSE calculations for hypothesis testing, Student's t-test (Press et al., 2007), testing for significance between SISR methods and the lowest RMSE from an interpolation method. We rejected the null hypothesis at $p \leq 0.05$. Student's t-test indicated both significantly improved or degraded performance based on, respectively, lowered and increased RMSE when accompanied with rejection of the null hypothesis.

4.1. Internal validation

Internal validation of the twelve techniques for each of the three data sets separately used 5-fold cross-validation, where 80% of data from each

Table 1
Descriptions of Interpolation and SISR algotithms Used in Experiments.

1	**Splines in Tension (SIT) -1990** Widely used algorithm. Uses the Generic Mapping Tools (GMT) software package
2	**Splines in Tension (SIT) -1993** Uses Matlab regspline2d function. Used within open source GIS (GRASS GIS) package
3	**Splines in Tension (SIT) -1998** Uses Matlab spline2d function. Interpolates data with a cubic spline in 2 dimensions.
4	**Bicubic Interpolation (BI)** Linear space-invariant interpolation scheme. An extension of cubic interpolation for interpolating data points on a 2D grid
5	**Neighborhood Embedding + Least Square (NE + LS)** Uses unconstrained least squares to solve regression. It is based on neighbor embedding, i.e. how a feature vector corresponding to a patch can be reconstructed by its neighbors in feature space
6	**Neighborhood Embedding + NonNegative Least Square (NE + NNLS)** Very similar to NE + LS. Uses Langrange multipliers to solve problems with using standard SUM1-LS method
7	**Sparse Representation Single Image Scale Up (Zeldye)** Uses sparse-representation modeling. Assumes a local Sparse-Land model: assumes that each patch from the images considered can be well represented using a linear combination of few atoms from a dictionary.
8	**Anchored Neighorhood Regression (ANR)** ANR uses Ridge Regression to learn exemplar neighborhoods, using these neighborhoods to precompute projections to map LR patches to HR domain. This method learns sparse dictionaries and regressors, anchored to dictionary atoms
9	**Global Regression (GR)** Unlike ANR where neighborhood size is set, in GR the neighborhood coincides with whole dictionary in use. This method uses Ridge Regression/Collaborative Representation.
10	**Adjusted Anchored Neighorhood Regression (A+)** Based on ANR but instead of learning regressors on dictionary it uses full training material approach similar to methods like Simple Functions. A+ still trains the dictionary, but keeps training samples (neighborhood) after the dictionary is trained.
11	**Jointly Optimized Regressors (JOR)** Exemplar-based -input LR image decomposed to fixed size overlapping patches size. This method uses jointly optimized collection of fixed number of local regressors. Overall, the determined "most appropriate regressor" is used to resolve the HR estimate
12	**Super Resolution Forests (SRF)** Does not rely on neighborhood embedding and sparse-coding, but uses locally linear regression. This replaces the single dictionary approach of other methods with many smaller ones.

region was used for training and the remaining 20% for validating performance of each SISR method. So we trained and applied five separate sets of SISR models to each region. Only eight of twelve SISR methods in our experiments were trainable with prior extracted information from samples of coincident LR and HR patches.

The five divisions of training and testing data does not qualitatively alter the general trend of results in SISR performance. Small changes in training data does not cause a large change in model performances.

Table 2a tabulates the RMSE of multiple SISR algorithms across this five-fold cross validation. This represents overall performance of the twelve SISR methods when training is performed internally to the region. Fig. 5 displays mean RMSE and standard error of the mean, σ_M, Eq. 2:

$$\sqrt{\sigma_M} = \frac{\sigma}{N}, \tag{2}$$

where σ is standard deviation of the RMSE, and N is sample size. Table 2 tabulates the p-value of accepting the null hypothesis from Student's t-test. The bold values indicate that the SISR algorithm shows significantly lower RMSE compared to bicubic interpolation (i.e., reject the null hypothesis); italic highlights means significantly larger RMSE. A significantly lower RMSE (bold) is the desired result.

Based on five-fold cross validation, SISR algorithms generally upscale from LR to HR grids with significantly enhanced skill versus bicubic interpolation as quantified by Student's t-test for the two East Pacific Ridge data sets. Seven SISR methods pass t-test for SEPR, five SISR methods pass for NEPR. For the JUAN region, only the SRF method provides enhancement.

Additionally, we computed SISR performance on bathymetry data when all available data is mixed. Rows in Table 2 labeled ALL show RMSE and p-values from Student's t-test when all data are combined in five-fold cross validation. The same SISR that passed the t-test in NEPR pass for this ALL case, but with higher p-values than NEPR or SEPR cross-validation results alone.

4.2. External validation

For external validation, the SISR methods learn from patches in one region and then predict bathymetry in an external regions. For example, SISR techniques trained on JUAN data would be applied to enhance SEPR and NEPR. Here we used the entire data set from each of the three regions to create separate SISR models and then applied separately to the remaining two regions. In general, these tests demonstrated that trained models based on regions mapped by sonar could be extended into non-neighboring regions where only satellite altimetry predicted bathymetry is available. Fig. 6 shows bar graphs for these tests.

Note the bicubic interpolation method requires no training. All other trained SISR methods have two bars for each of two external training areas applied to the test area with the applicable RSME above each bar. There are no standard deviation bars here since predictions are for the entire test area at once. We assume that the ratio of (error of predicted depth)/(predicted depth) is equal to the fractional standard deviations of internal tests.

Table 3 a tabulates RMSE values in Fig. 6 and 3 b the p-value of accepting the t-test null hypothesis. Bold and Italic (significant improvement or degradation respectively) highlight tests in which the null hypothesis is rejected for $p < 0.05$, compared to bicubic interpolation. In addition, we highlight cases where the null hypothesis is rejected for a weaker value of $p < 0.1$ by placing numbers in parentheses.

Table 2
Internal test results: A) Root-mean squared error (rmse) values; B) Hypothesis testing p-values vs bicubic interpolation. Bold values indicate SISR algorithm shows significantly lower rmse compard to bicubic interpolation (reject null hypothesis); Italic - significantly larger rmse. Bold values are desired result.

	Bicubic	NE + LS	NE + NNS	Zeyde	ANR	GR	A+	JOR	SRF
(A)									
Juan	34.49	34.18	34.82	33.79	33.78	33.80	34.02	*36.78*	**32.37**
Nepr	46.65	**44.37**	46.26	**43.99**	**43.79**	**43.95**	45.57	45.54	**43.56**
Sepr	46.93	**42.63**	**43.98**	**42.38**	**41.58**	**41.59**	45.50	**42.60**	**42.24**
All	40.10	**38.86**	39.64	**38.74**	**38.61**	**38.77**	39.45	40.05	**37.84**
(B)									
Juan	1	0.67	0.64	0.33	0.32	0.34	0.52	*0.0029*	**0.003**
Nepr	1	**1.3E-03**	0.62	**1.7E-04**	**5.4E-05**	**1.3E-04**	0.13	0.16	**9.8E-06**
Sepr	1	**1.3E-05**	**0.005**	**3.8E-06**	**4.9E-08**	**5.3E-08**	0.16	**1.3E-05**	**1.5E-06**
All	1	**0.0096**	0.47	**0.0044**	**0.0018**	**0.0054**	0.17	0.92	**2.1E-06**

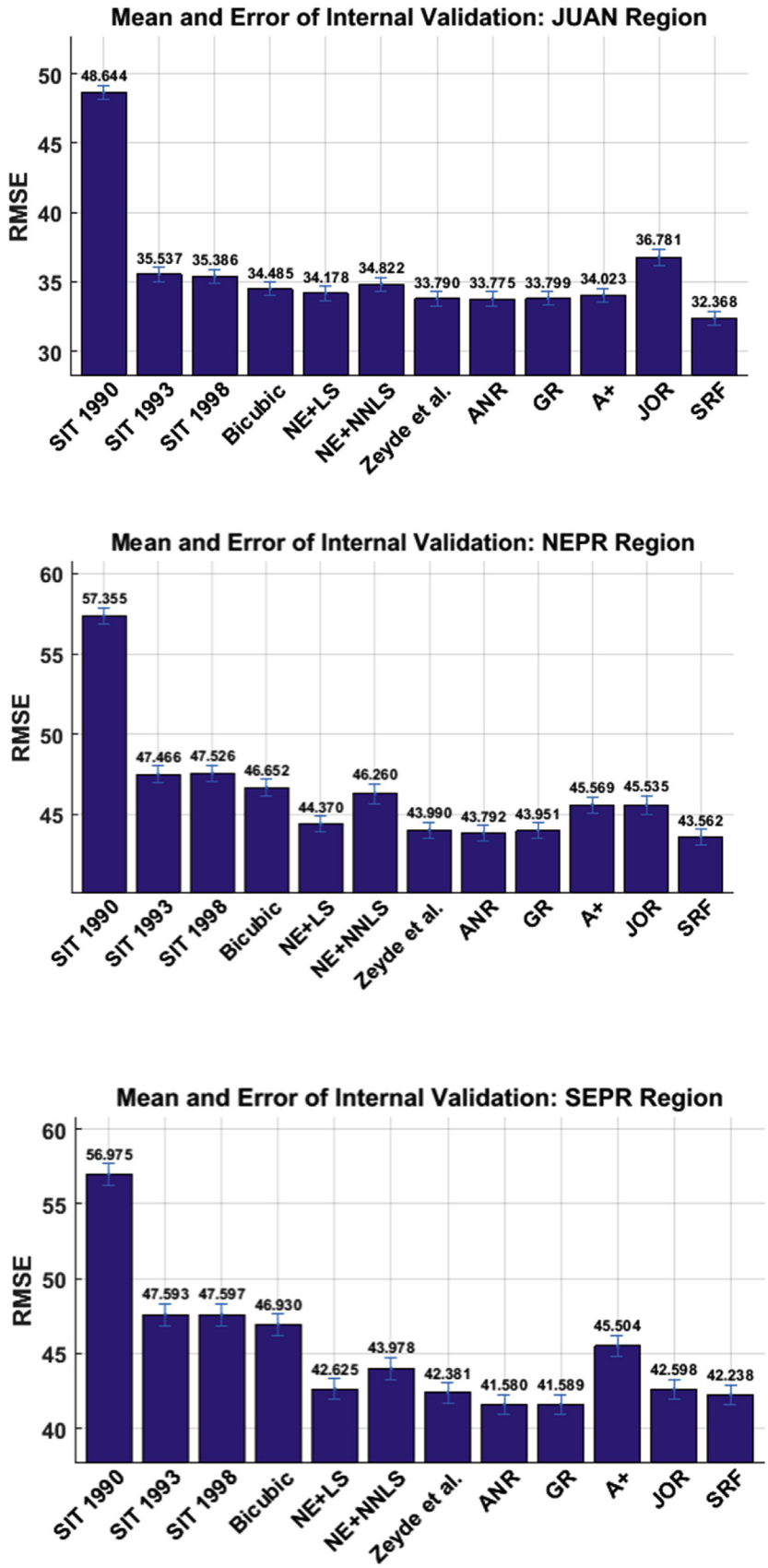

Fig. 5. Internal test results: average RMSE results in JUAN, NEPR and SEPR regions for internal five-fold tests. Lower RMSE means a better result. A bar graph of the overall average RMSE scores for all data in the region with standard error of mean given by the error bar.

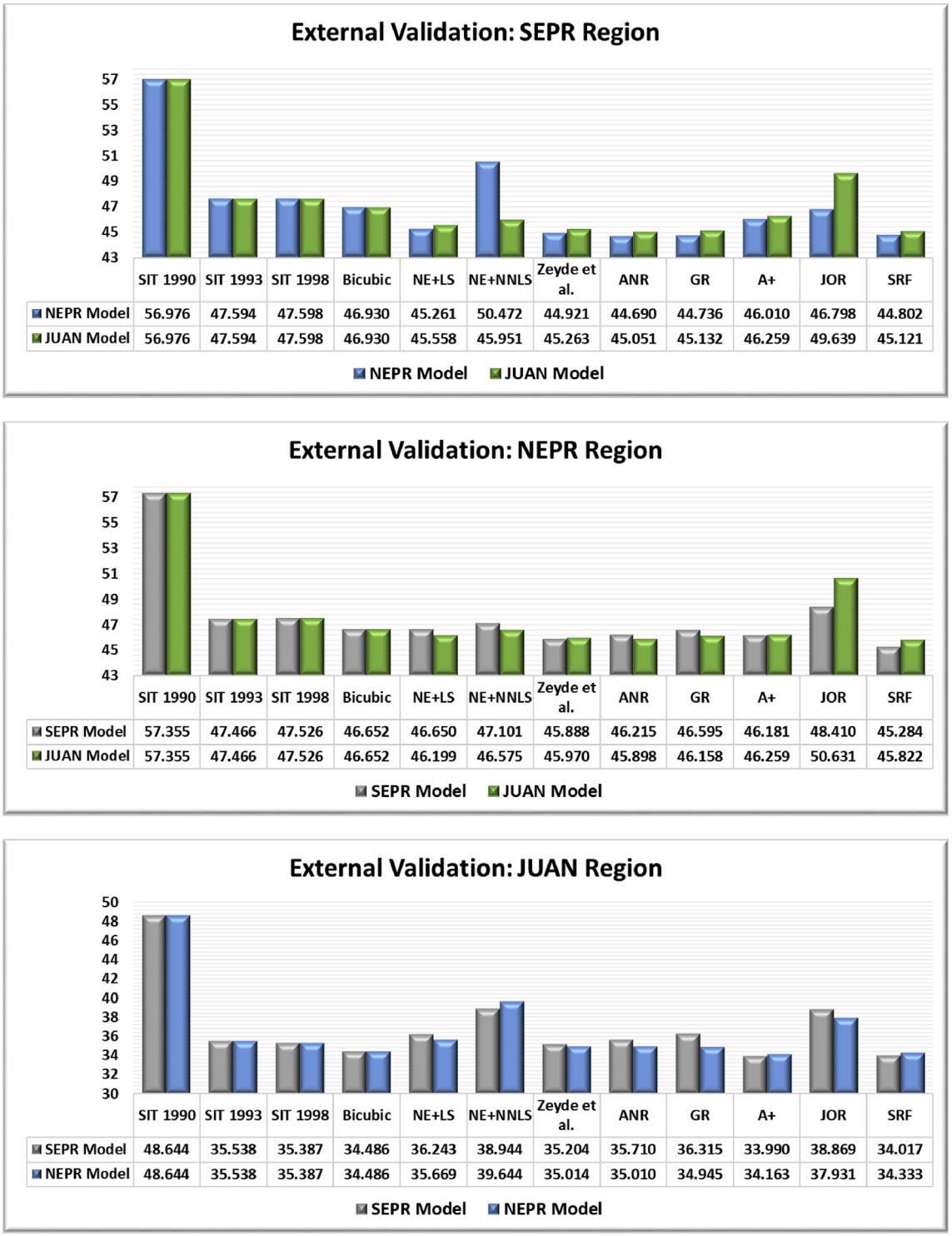

Fig. 6. External test results: Bar graphs providing the RMSE results in the JUAN, NEPR and SEPR external tests. The RMSE for the data invariant linear filters, such as bicubic interpolation, will result in one value since these interpolation methods have no training mechanism. The trained methods give varied RMSE for the two regions external to the train region.

One external case, using NEPR trained SISR algorithms, showed significantly lowered RMSE versus bicubic interpolation. In particular, the methods by Zedye, ANR, GR and SRF showed significantly lowered RMSE. For the weaker $p < 0.1$ criterion, training in JUAN provides lowered RMSE in SEPR for these same methods. In addition, training the SRF algorithm in SEPR to predict NEPR has $p = 0.052$, qualitatively close to passing at $p < 0.05$.

Experiments for data invariant linear filters alone showed that the bicubic interpolator outperforms all of the SIT algorithms. This result indicates that for the case of upscaling of gridded bathymetry (e.g., digital bathymetry models or DBMs), the bicubic method provides lower RMS error versus SIT. For data that is irregularly spaced and potentially sparse, however, SIT methods may be a better choice when using data invariant linear filters for interpolation.

Table 3
External test results: A) Rmse values; B) Hypothesis testing p-values vs bicubic interpolation. Bold values indicate SISR algorithm shows significantly lower rmse compard to bicubic interpolation (reject null hypothesis); Italic - significantly larger rmse. Bold values are desired result.

	Bicubic	NE + LS	NE + NNLS	Zeyde	ANR	GR	A+	JOR	SRF
(A)									
JUAN Test									
SEPR Models	34.49	*36.24*	*38.94*	35.20	35.71	*36.32*	33.99	*38.87*	34.02
NEPR Models	34.49	35.67	*39.64*	35.01	35.01	34.95	34.16	*37.93*	34.33
NEPR Test									
SEPR Models	46.65	46.65	47.10	45.89	46.22	46.60	46.18	*48.41*	(45.28)
JUAN Models	46.65	46.20	46.58	45.97	45.90	46.16	46.26	*50.63*	45.82
SEPR Test									
NEPR Models	46.93	45.26	50.47	**44.92**	**44.69**	**44.74**	46.01	46.80	**44.80**
JUAN Models	46.93	45.56	45.95	(45.26)	(45.05)	(45.13)	46.26	*49.64*	(45.12)
(B)									
JUAN Test									
SEPR Models	1	*0.02*	*0.02*	0.33	0.10	*0.016*	0.49	*2.9E-08*	0.52
NEPR Models	1	0.11	*8.1E-05*	0.48	0.48	0.53	0.65	*1.1E-05*	0.83
NEPR Test									
SEPR Models	1	1.00	0.55	0.28	0.54	0.94	0.51	*0.017*	0.052
JUAN Models	1	0.53	0.92	0.34	0.29	0.49	0.58	*4.8E-07*	0.24
SEPR Test									
NEPR Models	1	0.099	0.13	**0.046**	**0.026**	**0.029**	0.36	0.90	**0.032**
JUAN Models	1	0.18	0.34	0.098	0.062	0.074	0.51	*0.012*	0.069

5. Discussion

Based on the five-fold cross validation, the SISR algorithms generally upscale from LR to HR grids, with significantly enhanced skill versus bicubic interpolation as quantified by Student's t-test for the two East Pacific Ridge data sets. The performance in the JUAN region, however, has less skill, with only the SRF algorithm showing significant improvement. Examining these results with the roughness distributions in Fig. 4 shows that trained SISR models can enhance localized bathymetric estimates in areas with roughened topography and where HR multi-beam data is available in close proximity. For smoothed regions, because of lowered high frequency content, many SISR methods may provide no advantages and, in some cases, are worse predictors versus bicubic interpolation.

In addition to the roughness requirement, both the internal and external validation experiments show that there must be an adequate number of detailed exemplars from which to train. For internal tests, the ratio of exemplar to test patches is four to one. The external validation experiment of using NEPR to predict SEPR is the only external upscaling case where the SISR methods provide enhanced skill versus interpolation. The NEPR region had twice the data of the SEPR set.

As a test region, SEPR was the best region in which to predict. Using NEPR for training, upscaling using Zeyde, ANR, GR and SRF pass Student t-test at $p < 0.05$. These same methods pass using JUAN with a weaker $p < 0.1$. These results may be explainable from the wider depth ranges in the JUAN and NEPR hypsometries versus the comparatively tighter hypsometry of SEPR. Generally, diverse training sets provided enhanced skill to prediction than less diverse sets.

The range of depths in SEPR is a subset of the range of depths of NEPR and JUAN. The match of SEPR roughness makes NEPR a stronger predictor for SEPR. Likewise, because SEPR has lower diversity of depths than that of NEPR, it is a poor predictor for NEPR.

The SRF method shows some evidence of robustness versus training data available. For the case of training SRF with SEPR data to predict the NEPR, the algorithm barely fails the stronger Student's t-test of $p < 0.05$ as its $p = 0.052$ for this case; no other technique was close. Also, SRF is the only SISR method that enables enhanced cross-validation in JUAN. The relatively smaller amounts of rough examples here versus NEPR and SEPR, however, is enough for the SRF method to have enhanced skill.

6. Discussion and conclusions

We addressed the following questions: can computer vision based methods be used to enhance our LR knowledge of seafloor topography with more skill than Splines-In-Tension or bicubic interpolation alone, and under what circumstances can enhancement occur? We showed that SISR, can enhance this LR knowledge of bathymetry through upscaling. We quantitatively judged successful enhancement when RMSE of upscaled LR bathymetry is significantly lowered by using SISR methods versus linear invariant interpolation as quantified by hypothesis testing. Table 1 lists and describes SISR methods tested in this paper.

We found that multiple SISR methods (Table 2) can provide statistically significant improvement in estimating HR information. This is true within a localized region under conditions of roughness present in the seafloor, as for the NEPR and SEPR regions. There were $O(10^3)$ training exemplars; the ratio of training to testing exemplars equaled four to one. The JUAN area is smooth compared to EPR regions. Here, SISR methods offered little improvement since there is a proportionally lower amount of high-detailed seafloor from which to train.

In addition, we discovered that four SISR methods (Table 3) can significantly lower RMSE versus interpolation for the case of external prediction (training in one area to enhance LR bathymetry in a second, distant area). This case occurred for training in the NEPR to enhance the SEPR. JUAN also enhanced SEPR with weaker acceptance criteria. Meanwhile, SEPR could not be used to predict NEPR (except for the SRF method) or JUAN. NEPR and JUAN were unable to enhance each other. We interpret these results to mean that externally trained methods can enhance a LR bathymetry grid if rough topography is present in both areas and the range of depths and features in the training area contains the range of depths in the enhancement area.

We also found evidence that the SRF method may be a more robust enhancement algorithm versus the other algorithms. SRF was the only algorithm to provide internal enhancement in JUAN and upscale NEPR using SRF trained in the comparatively smaller SEPR set. This observation is aligned with two other studies (Martin et al., 2015, Lawson et al., 2017), indicating that random-forest based computer vision algorithms provide predictive skill for geophysical or topographic predictions involving the ocean floor.

SISR methods can significantly upscale LR bathymetry to provide lower RMSE versus interpolation. Certain conditions must be met to successfully implement the method. The SISR methods perform best when there is a large quantity of local MBES data available for training the SISR methods. High-detailed topographic information must be present, and the range of depths of the training region should include the range of the test region. In contrast, interpolation offers as good or better

upscaling in areas where only a few soundings of MBES data are available, or in smoothed regions.

Further investigation could focus on SISR effectiveness in other regions that are only partially mapped by deep-water multibeam systems. In this paper, we considered only three regions in the Pacific Ocean. Application of these methods in the Atlantic Ocean could be effective, given the evidence of higher SISR skill with seafloor roughness. In addition, based on the increased skill in the internal tests, these SISR methods could be applied as inpainting, the process of restoring damaged or missing parts of images (Elad, 2010, Starck et al., 2015). The trained super-resolution methods used in this paper could be trained to compensate for any missing ping data in bathymetric side-scan images; neighboring seafloor structures could be used as exemplars and enhance the bathymetric estimates of the textural structure of missing regions. Finally, increasing the amount of training data inherently alters the performance of SISR methods discussed in this paper. With more data, deep learning methods (Goodfellow et al., 2016) that require large amounts of exemplars, such as convolutional neural networks (SR-CNN) (Dong et al., 2015)], could become implementable. Such an effort would introduce a broader range of features of undersea seafloor structures. Additionally it could be possible to learn more powerful feature representation by using convolutional neural networks (Cheng et al., 2018a, b, c) and some discrimination information (Cheng et al., 2018a, b, c).

Declarations

Author contribution statement

Kristen Nock, David Bonanno: Performed the experiments; Analyzed and interpreted the data.

Paul Elmore: Conceived and designed the experiments; Analyzed and interpreted the data.

Leslie Smith: Analyzed and interpreted the data.

Vicki Ferrini: Contributed reagents, materials, analysis tools or data.

Frederick Petry: Analyzed and interpreted the data; Wrote the paper.

Funding statement

This work was supported by the U.S. Naval Research Laboratory Base program.

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Ball, N.M., Brunner, R.J., 2010. Data mining and machine learning in astronomy. Int. J. Mod. Phys. D 19 (7), 1049–1106.

Bowman, A.W., Azzalini, A., 1997. Applied Smoothing Techniques for Data Analysis. Oxford Press, New York.

Carbotte, S.M., Arko, R., Chayes, D.N., Haxby, W., Lehnert, K., O'Hara, S., Ryan, W.B.F., Weissel, R.A., Shipley, T., Gahagan, L., Johnson, K., Shank, T., 2004. New integrated data management system for Ridge2000 and MARGINS research. Eos Transactions of the American Geophysical Union 85 (51), 553.

Chang, H., Yeung, D., Xiong, Y., 2004. Superresolution through neighbor embedding. In: Proc Computer Vision and Pattern Recognition, pp. 275–282.

Cheng, G., Han, J., Zhou, P., Xu, D., 2018a. Learning rotation-invariant and Fisher discriminative convolutional neural networks for object detection. IEEE Trans. Image Process. 28, 265–278.

Cheng, G., Yang, C., Yao, X., Guo, L., Han, J., 2018b. When deep learning meets metric learning: remote sensing image scene classification via learning discriminative CNNs. IEEE Trans. Geosci. Remote Sens. 56, 2811–2821.

Cheng, G., Zhou, P., Han, J., 2018c. Duplex metric learning for image set classification. IEEE Trans. Image Process. 27, 281–292.

Cochran, J.R., Goff, J.A., Malinverno, A., Fornari, D.J., Keeley, C., Wang, X., 1993. Morphology of a superfast midocean ridge crest and flanks - the East Pacific Rise, 7-degrees-9degrees-S. Mar. Geophys. Res. 15 (1), 65–75.

Dai, D., Timofte, R., Van Gool, L., 2015. Jointly optimized regressors for image super-resolution. Comput. Graph. Forum 34 (2), 95–104.

Dong, C., Loy, C., He, K., Tang, X., 2015. Image super-resolution using deep convolutional networks. IEEE Trans. Pattern Anal. Mach. Intell. 38, 295–307.

Elad, M., 2010. Sparse and Redundant Representations - from Theory to Applications in Signal and Image Processing. Springer, New York.

Freeman, W., Jones, T., Pasztor, E., 2002. Example-based super-resolution. IEEE Computer Graphics and Applications 22 (2), 56–65.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63 (1), 3–42.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, Cambridge, Massachusetts.

Gorini, M., 2009. Physiographic classification of the ocean floor: a multi-scale geomorphometric approach. In: Geomorphometry 2009 Conference Proceedings, pp. 98–105.

Hu, M., Li, J., Li, H., Shen, C., Jin, T., Xing, L., 2015. Predicting global seafloor topography using multi-source data. Mar. Geod. 38 (2), 176–189.

Iwahashi, J., Pike, R.J., 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. Geomorphology 86 (3–4), 409–440.

Lawson, E., Smith, D., Sofge, D., Elmore, P., Petry, F., 2017. Decision forests for machine learning classification of large, noisy seafloor feature sets. Comput. Geosci. 99, 116–124.

Macdonald, K., 2012. SeaMARC II and sea beam bathymetric data compilation of East Pacific Rise and its flanks 8°-18°N (1987-1990). Integr. Earth Data Appl. (IEDA).

Marks, K.M., Smith, W.H.F., 2012. Radially symmetric coherence between satellite gravity and multibeam bathymetry grids. Mar. Geophys. Res. 33, 223–227.

Marks, K.M., Smith, W.H.F., 2016. Detecting small seamounts in AltiKa repeat cycle data. Mar. Geophys. Res. 37 (4), 349–359.

Marmion, M., Hjort, J., Thuiller, W., Luoto, M., 2009. Statistical consensus methods for improving predictive geomorphology maps. Comput. Geosci. 35 (3), 615–625.

Martin, K.M., Wood, W.T., Becker, J.J., 2015. A global prediction of seafloor sediment porosity using machine learning. Geophys. Res. Lett. 42 (–), 10,640–10,646.

Mather, P.M., Koch, M., 2011. Computer Processing of Remotely-Sensed Images an Introduction, fourth ed. WileyBlackwell, Hoboken, NJ.

Mathworks Inc., 2017. Matlab, Release 2017a. The Mathworks: Natick, Massachusetts.

Micallef, A., Berndt, C., Masson, D., Stow, D., 2007. A technique for the morphological characterization of submarine landscapes as exemplified by debris flows of the Storegga Slide. Journal of Geophysical Research-Earth Surface 112 (F2), F02001.

Milanfar, P. (Ed.), 2010. Super-Resolution Imaging. CRC Press, Boca Raton FL.

Parker, R., 1972. Rapid calculation of potential anomalies. Geophys. J. R. Astron. Soc. 31 (4), 447–455.

Picard, K., Brooke, B., Coffin, M.F., 2017. Geological insights from Malaysia Airlines flight MH370 search, *Eos*, 98. Transactions American Geophysical Union, pp. 752–761.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. MNumerical Recipes: the Art of Scientific Computing, 3ed. Cambridge University Press, Cambridge, United Kingdom.

Ryan, W., Carbotte, S., Coplan, J., OHara, S., Melkonian, A., Arko, R., Weissel, R., Ferrini, V., Goodwillie, A., Nitsche, F., Bonczkowski, J., 2009. Global multi–resolution topography synthesis. Geochem. Geophys. Geosyst. 10 (3), 47–56. http://www.geomapapp.org.

Sandwell, D.T., Muller, R.D., Smith, W.H.F., Garcia, E., Francis, R., 2014. New global marine gravity model from CryoSat-2 and Jason-1 reveals buried tectonic structure. Science 346 (6205), 65–67.

Schawinski, K., Zhang, C., Zhang, H., Fowler, L., Santhanam, G.K., 2017. Generative adversarial networks recover features in astrophysical images of galaxies beyond the deconvolution limit. Mon. Not. R. Astron. Soc. Lett. 467 (1), L110–L114.

Smith, W.H.F., Marks, K.M., 2014. Seafloor in the Malaysia Airlines flight MH370 search area. Eos, Transactions American Geophysical Union 95 (21), 173–174.

Smith, W.H.F., Sandwell, D.T., 1994. Bathymetric prediction from dense satellite altimetry and sparse shipboard bathymetry. J. Geophys. Res. B 99 (B11), 21803–21824.

Smith, W.H.F., Sandwell, D.T., 1997. Global sea floor topography from satellite altimetry and ship depth soundings. Science 277 (5334), 1956–1962.

Smith, W.H.F., Wessel, P., 1990. Gridding with continuous curvature splines in tension. Geophysics 55 (3), 293–305.

Starck, J.L., Murtagh, F., Fadill, J.M., 2015. Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis, 2ed. Cambridge University Press, Cambridge, United Kingdom.

Timofte, R., Smet, V., Gool, L., 2014. A+: adjusted anchored neighborhood regression for fast superresolution. In: Proc. Computer Vision–ACCV 2014, pp. 111–126.

Timofte, R., Smet, V., Gool, L., 2013. Anchored neighborhood regression for fast example-based super-resolution. In: Proc. IEEE Int. Conference on Computer Vision, pp. 1926–1927.

Valentine, A.P., Kalnins, L.M., Trampert, J., 2013. Discovery and analysis of topographic features using learning algorithms: a seamount case study. Geophys. Res. Lett. 40 (12), 3048–3054.

Weatherall, P., Marks, K., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J.E., Rovere, M., Chayes, D., Ferrini, V., Wigley, R., 2015. A new digital bathymetric model of the world's oceans. Earth and Space Science 2 (8), 331–345.

Wessel, P., Bercovici, D., 1998. Interpolation with splines in tension: a Green's function approach. Math. Geol. 30 (1), 77–93.

Wessel, P., Smith, W., Scharroo, R., Luis, J., Wobbe, F., 2013. Generic mapping tools: improved version released, eos. Trans. Am. Geophys. Union 94 (45), 409–410.
Wilson, D., 1993. Confidence intervals for motion and deformation of the Juan de Fuca plate. J. Geophys. Res. B 98, 16,053–16,071.
Wormald, S., Wright, I., Bull, J., Lamarche, G., Sanderson, D., 2012. Morphometric analysis of the submarine arc volcano Monowai (Tofua-Kermadec Arc) to decipher tectono-magmatic interactions. J. Volcanol. Geotherm. Res. 239, 69–82.
Wu, J., Bai, M., 2018. Incoherent dictionary learning for reducing crosstalk nosie in least-squares reverse time migation. Comput. Geosci. 114, 11–21.
Yang, C.-Y., Yang, M.-H., 2013. Fast direct super-resolution by simple functions. In: Proc. IEEE Int. Conf. on Computer Vision, pp. 561–568.
Yang, C.-Y., Ma, C., Yang, M.-H., 2014. Single-image super-resolution: a benchmark. In: Proc. Computer Vision–ECCV 2014, pp. 372–386.
Yang, J., Wright, J., Huang, T., Ma, Yi, 2008. Image super-resolution as sparse representation of raw image patches. In: Proc IEEE Conf Computer Vision and Pattern Recognition, pp. 108–117.
Zeyde, R., Elad, M., Protter, M., 2010. On single image scale-up using sparse-representations. Curves and Surfaces. Springer, Berlin Heidelberg, pp. 711–730.
Zhou, P., Han, J., Cheng, G., Zhang, B., 2019. Learning compact and discriminative stacked autoencoder for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 57 (7), 4823–4833.