Machine-Learned Mapping of Basin-wide Salt: Old Data - New Methods

Scott Morris[1], Shuang Li[2], Tony Dupont[3] and John D. Grace[4],

ABSTRACT

Machine learning techniques were applied to estimate a regional map of the depth to top of salt for the United States portion of the Gulf of Mexico (GOM). SEG-Y files for 8,043 post-stack, migrated 2D lines were used – all shot between 1981 and 1992 in 82 separate surveys. Three factors differentiate this research from previous studies of automated salt discrimination: 1. Rather than test a few lines, we assessed the scalability of unsupervised machine learning to thousands of lines, in batch, covering a basin. 2. The input data were all more than 25 years old, from disparate surveys. This tested the economic viability and technical robustness of extracting new information, via machine learning from raster and vector representations of lines, drawn from immense, old and disparate data collections, similar to those held by companies. 3. The massive volume of data prohibited checking results line-by-line. Therefore, automated methods were developed to assess confidence in the predicted salt boundaries and edit output. Our final depth to top of salt map covered 67% of the 278,000 km^2 (107,000 mi^2) study area; 17% of the original area had insufficient data coverage and analysis covering 16% of the area did not meet confidence requirements for inclusion. The final results compared well with published maps of salt and the locations of salt-trapped fields in the GOM. However, reliable mapping of salt deeper than 6 seconds two-way time could not be achieved with this data set and approach. Computing time was greater than linear in the number of lines but changes in hardware configuration could reduce run time of about three weeks to about three days.

[1] Earth Science Associates, scott@earthsci.com.

[2] Earth Science Associates and Graduate Program in Applied Mathematics, University of Southern California, lishuang@usc.edu.

[3] Earth Science Associates, tony@earthsci.com.

[4] Earth Science Associates, john@earthsci.com. Corresponding author

INTRODUCTION

Seismic technology's advance over the last three decades, particularly the transition from 2D to 3D, has seemingly rendered enormous volumes of vintage 2D data obsolete. We offer two hypotheses to challenge that conclusion. First: applying innovative machine learning techniques, in batch, produces valuable new information from these legacy assets at very low marginal costs. Second: changing the scale of intensive analysis from the prospect/field level to a whole region, or a basin, further leverages the return to investment in automated interpretation.

To test these propositions, we map basin-wide top of salt across the Gulf of Mexico (GOM). We applied a workflow of machine learning and macro-editing techniques to a set of 8,043 2D post-stack, migrated SEG-Y lines that cover approximately 278,000 km^2 (107,000 mi^2) of the northern shelf and slope. They were shot by approximately a dozen operators and contractors in 82 surveys, *all more than 25 years ago*, between 1981 and 1992. This trove of disparate data, therefore, reflects the character of many of the seismic legacy assets companies hold, as well as some of the seismic data now being released into the public domain (e.g., CDA (2018) in the United Kingdom and the United States Geological Survey (2018) in the United States).

Our central scientific enterprise, constructing a basin-wide product by machine learning of thousands of lines of seismic data, accommodating disparities in their age and provenance – while minimizing cost – required renovating previous approaches and departing from some analytic traditions.

First is the type of machine learning. Taking salt discrimination broadly as a pattern recognition problem, two genres of analysis have evolved: supervised and unsupervised. The former requires an expert interpreter to manually identify and label numerous examples of salt, non-salt and their boundaries. Algorithms are trained on the labeled subset, tested on holdout samples, perhaps refined and then applied to find salt in new data.

In a supervised system, classification rules are induced from indicators built from the library of expert-labeled, micro-scale examples. Such experiments are described by Araya-Polo, *et al* (2017), Amin, *et al* (2015), Berthelot, *et al*, (2013) and Halpert and Clapp (2008). These methods are invested with a human expert's credibility but that role bears an economic cost and the possibility of introducing biases. Supervised models have dominated the literature of machine-learned salt discrimination.

Unsupervised approaches deduce classification criteria directly from the data. The immediate objective is to reduce the dimensionality of the input data's attribute space to a smaller number (≥ 2) of interpretable clusters; for example, developing a taxonomy for surface waves (Galvis, *et al* (2017)) or identification of facies (De Matos, *et al* (2010) and Coléou, *et al*

(2003)). Reducing output clusters to the extreme of two – for salt and non-salt – is of particular value here. In this pursuit, some authors focused on models that maximally exploit the power of a single discriminator; for example, phase coherence (Shafiq, et al (2017a)). Others have statistically explored for the maximally effective ensemble of discriminators for classifying salt versus non-salt; for example, Di, *et al* (2017) and Amin, *et al* (2015).

We chose unsupervised methods and built an ensemble of discriminators. If it works, the economic return to a successful unsupervised analysis is higher than supervised methods because of the smaller time required from experts. For a basin-level *supervised* study, the size of the training set is driven even higher by the combinatorial force of the number of targets requiring exemplars (salt, non-salt, boundaries) *times* the number cases in which the seismic "background" from which they must be distinguished is materially different – simply because of survey-level acquisition and processing. An unsupervised workflow, however, is not "expertise free" – human judgement is still integral to designing processes, choosing statistical methods, establishing parameter values and thresholds; then, editing results at the survey to region (macro) level to yield final products. Even with these quality-control steps, it must be noted that our mapping yields *regional* results and is not, as constructed, a method for determining high-precision boundaries, particularly on salt flanks.

Second, the current body of published machine-learned salt discrimination techniques was uniformly developed on very small sets of real data – typically a few lines on which salt was known to exist, *a priori.*[5] These small-sample, proof-of-concept experiments were highly valuable. Over 20 years of research, they helped found a new paradigm for geophysical analysis. However, rather than locating salt on a line where it had already been found manually, we investigate the scalability and robustness of these methods and a workflow for a basin full of seismic. A central pillar of the claimed utility of machine learning is scale. So, we tested this technology at a scale where machine learning either earns its economic keep – or does not.

Finally, massive data sets always require considering computational complexity. A regional or basin-wide process must run within a practical length of time.

We address four basic questions. Is our methodology, and its component algorithms, technically effective for regional to basin-wide mapping of salt? Central to that question is the adequacy of *un*supervised classification. To the basic requirement that the method works, we add: Can this approach be generalized – scaled to batch analysis of thousands of lines – and thereby be usefully applied to a region or basin? Is this methodology robust in the face of the problems of "old" data, collected from a large number of disparate surveys? Finally, are the marginal costs to implement our approach low enough to motivate exploiting the countless,

[5] Alternatively, some research has been conducted on larger data sets; however, these were synthetic data sets, which are also useful as proof-of-concept tests but do not face the problems of application to real seismic data. See Araya-Polo, *et al* (2017) for methods developed for fault identification, tested on synthetic data, using an approach that could be applied to salt.

enormous legacy seismic assets, which once cost billions of dollars to acquire, but are now otherwise "free" but unused?

METHODOLOGY

To exhaustively partition a seismic line into two classes, say salt and non-salt, either divide it into areas, implicitly defining the boundaries between them, or, estimate the boundaries, implicitly defining the areas within them. The sets of boundary points and the points in areas are mathematical duals. The raster and vector representations of the data in a seismic line are likewise essentially dual. Although they are different objects, they are closely related and can be generally recovered from each other by involution. Yet, because they are not exact inverses, the raster version of a seismic line produces information not directly available from the vector data and *vice versa*. In the same spirit, one component of our workflow provides evidence of where salt is *not* – the complement of our target.

To maximize the information brought to bear on salt/non-salt discrimination, we applied methods that exploited *both* the raster and vector instances of the amplitude data in each seismic line. In the sense used here, raster refers to the matrix of pixels of an image of the data contained in a single seismic line. Vector refers to the representation of logical objects in the seismic data, here they represent reflectors. They are described most simply by a vector of x/y pairs describing the Cartesian location of the object on the 2D section. That location vector may be supplemented by dimensions that capture attributes at each point of the object and of the line itself.

Specifically, we developed one vector method for identification of boundaries using vector data, another four raster methods produced four statistics on the raster image of each line to classify areas and, finally, a mixed vector-raster method for identifying areas where salt (in principle) should be absent (Figure 1). Here, rasters primarily support analysis of seismic texture, the characteristic patterns of reflection amplitudes defined by the magnitude and variation of neighboring samples at a given spatial location (Gao, 2003). For this, we exploited the grey-level occurrence matrix (GLCM) methodology to help partition areas of salt and non-salt. The development of GLCM, and its fairly wide application in discrimination, is reviewed by Gao (2011) and Berthelot, et al, (2013). Our analysis focused on four salt-discriminatory GLCM statistics: contrast, homogeneity, entropy and dissimilarity. Vectors extracted from each seismic line, representing reflectors, subset by dip, strength and orientation, were employed to identify the classic image of diapiric salt – bounded by chevrons of paired, short, strong, high-dip reflectors on both sides. One method specifically looked for this geometry. The mixed raster/vector method begins with extraction of reflectors as vectors, subsets those associated with bedded rock and calculates the density of that subset to differentiate areas evidencing bedded rocks and therefore provide negative evidence on the presence of salt.

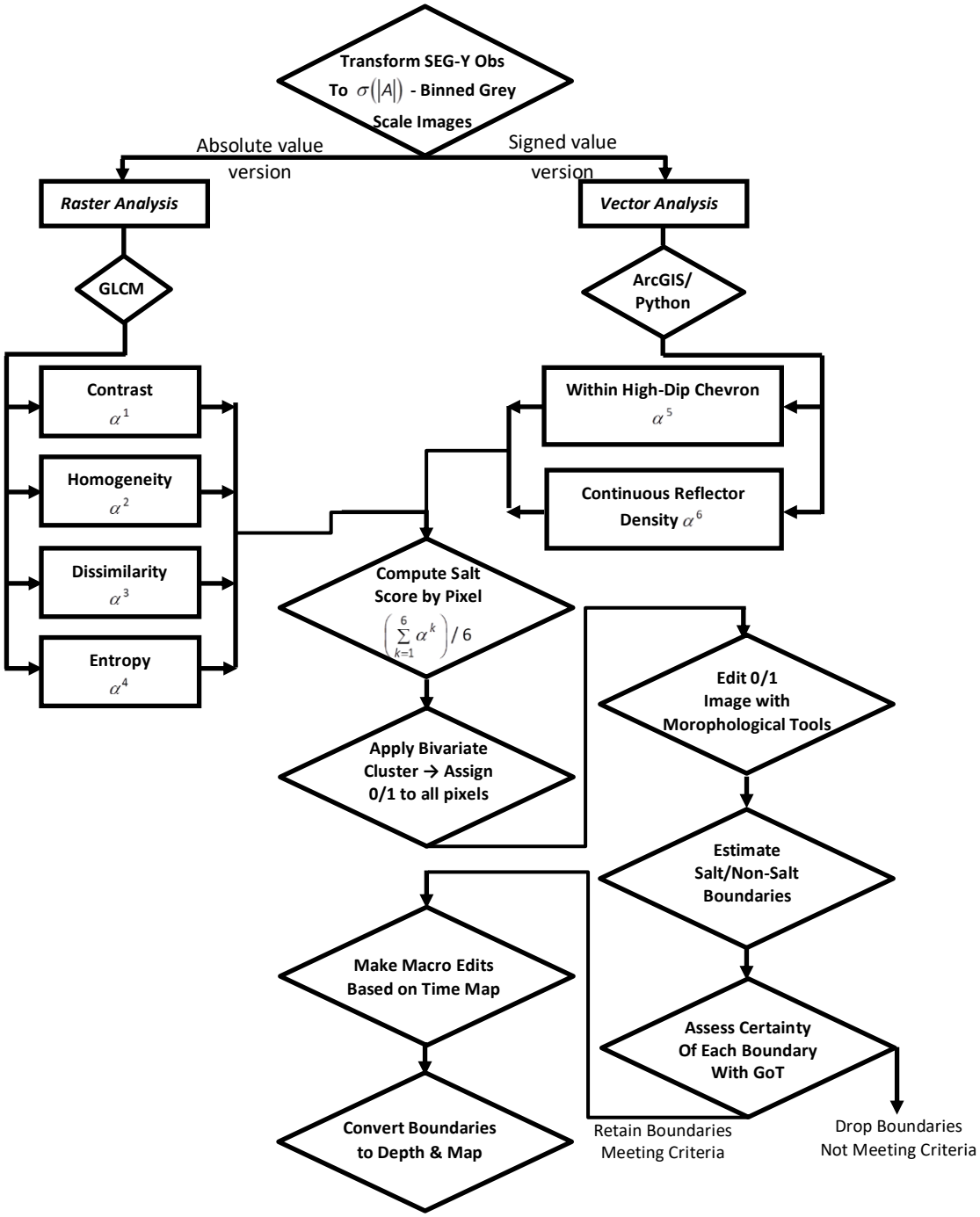


Figure 1 shows the general work flow for the application of batch discrimination of salt in 2-D seismic and the creation of the top of salt maps in time and depth. The variable $|A|$ is the absolute value of the amplitude of a seismic sample. GoT abbreviates gradient of texture.

The outputs of each of these methods were combined into a "salt score" for each pixel in a line, which was then subject to a two-group classification algorithm (Otsu, 1979). The resulting raw black-and-white image of the line was then edited with morphological tools to: 1. filter out the smallest estimated salt bodies (presumptive false positives); 2. Fill the smallest holes in estimated salt bodies (presumptive false negatives) and 3. Smooth boundaries to more closely approximate geologically acceptable interpretations (Wallet and Pepper, 2013).

The focus then shifted to evaluation of confidence in the estimated salt/non-salt boundaries using the gradient of texture (GoT) to assess each boundary in each line. GoT is commonly used in image analysis to evaluate boundaries between textures, as described by Shafiq, *et al* (2017b) and Hegazy, *et al* (2015). The average GoT of each salt polygon became a grade of confidence in the estimated boundary; this influenced the decision to retain or reject each candidate boundary. The average GoT scores for each boundary was also reflected as points along the line's navigation data coincident with the lateral extent of each estimated salt boundary and mapped.

In the final step, each line and its estimated salt boundaries were converted from time to depth using empirically estimated velocity functions. Results in depth are expected by geologists and engineers. Because of the layout of multiple lines from different surveys very close to each other, a filtering process was developed to include the most consistent data points in the gridding of the top of salt surfaces. The final products were confidence-marked top of salt maps in time and depth and 3D surfaces for use in 3D geographic information systems (GIS). The project workflow is outlined in Figure 1.

Data

Data for our experiment were drawn from publicly available seismic surveys submitted to the US government by operators in the US-federal portion of the Gulf of Mexico (United States Bureau of Ocean Energy Management (BOEM), 2018). The 8,043 2D time sections totaled approximately 450,000 km (280,000 mi) in length; they averaged 56 km (35 mi) long and typically extended to approximately 8 seconds in two-way time. Converted to images, sections averaged 2,860 x 2,082 pixels. The US government holds these records confidential for 25 years before public release. All surveys used here were shot between 1981 and 1992. They are, by any measure, "old" data.

We illustrate our methodology using part of one line, 'w-86-314', from survey '1986 namss.B-39-86-LA.msc.airgun', a 51.4 km (31.9 mi) line, with a maximum recorded time of 7.5 seconds (Figure 2a). It was shot in 1986 by Western Geophysical from northwest to southeast across of the southwestern part of the Ship Shoal protraction area of the GOM (Figure 2b). It

covers the Ship Shoal 154 field, trapped on a salt-dome, with a latest estimated ultimate recovery of 132 million barrels of oil equivalent (BOEM, 2018).

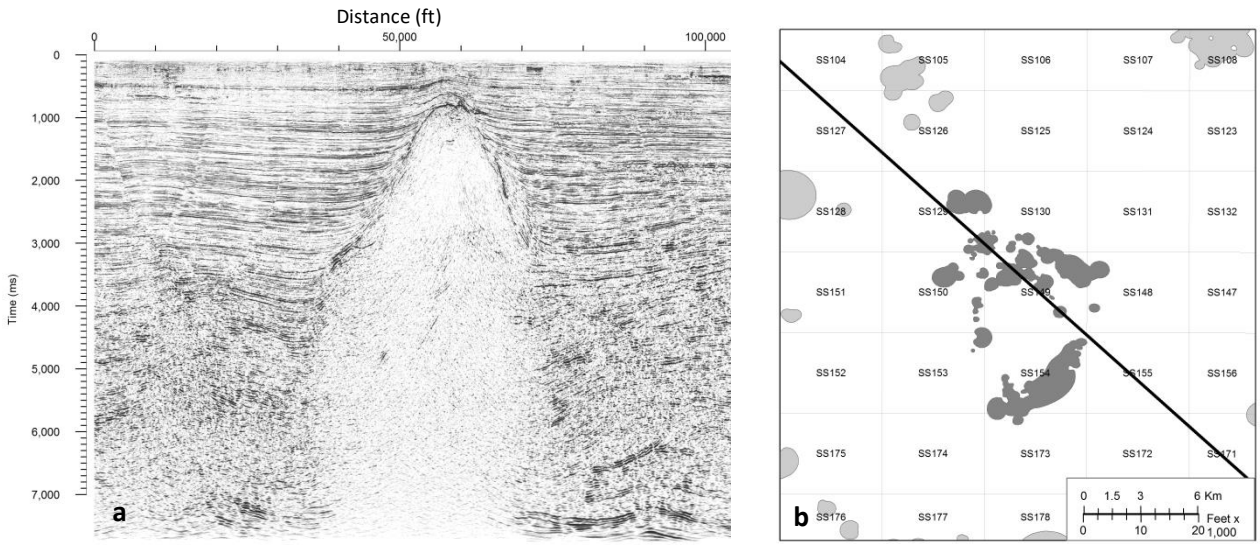


Figure 2a shows the grey-scale image of a portion of the time section of the 'w-86-314' line in survey '1986 namss.B-39-86-LA.msc.airgun' running northwest to southeast, left to right. Figure 2b maps the line and fields in the area (light grey) and the Ship Shoal 154 field (dark grey). The numbered squares are lease boundaries, SS abbreviates Ship Shoal.

Raster analysis

The amplitude records in each SEG-Y file were read and mapped to a bin based on the number of standard deviations ($|\sigma(A)|$) away from a line-wise mean value, which, in almost all cases, was zero. We binned them into a seven-level grey scale (Table 1). These boundaries reflect a non-linear transformation in which very low amplitude values (which can produce false-positive salt classifications) were given a more detailed breakdown in grey levels to help limit these errors. Conversely, extreme high absolute values of amplitude were compressed as ever higher amplitudes, beyond a point, contained less and less useful information for salt discrimination.

The seismic texture of salt (particularly in domes) is often obvious on visual inspection: reflectors in salt bodies are typically and distinguishably weak, short, chaotically oriented and sparse compared to surrounding stratified rock. Our analysis attempts to approximate the visual differentiation that comes naturally to an expert interpreter. In this, we extend previous work using GLCM analysis of seismic texture as discussed in Wu (2016), Amin (2015), Shafiq (2015), Wang (2015), Hegazy (2014) and Berthelot (2013). Our implementation employed the

GLCM library of *R,* documented by Zvoleff (2016). It consists of two steps, both executed for each pixel in each seismic section.

1. Build the GLCMs for All Pixels: Establish an examination pixel near the upper left corner of the image and define a neighborhood of *N* x *N* pixels centered on it (here, *N* = 51).[6] Then, starting in the upper left corner of the *neighborhood*, count the number of co-occurrences of grey levels in successive pairs of pixels on the top row of the neighborhood, moving right; then, drop to the row below and repeat until all rows (and pixel pairs) in the neighborhood are accounted. The process is illustrated in Figure 3. [7] For a system of *G* grey levels, this creates a *G* x *G* matrix of the empirical counts of grey-level co-occurrences for the neighborhood of one pixel (here, *G* = 7). Normalizing the count matrix by dividing each element by the total number of possible outcomes, (*N x (N-1)* = 2,550*)* produces a 7 x 7 matrix of the joint marginal probabilities of grey-level co-occurrences (p_{ij}) between the 7 grey levels (*i* and *j*) among horizontally adjacent pairs of pixels, examined row-wise left to right. This is the GLCM for a pixel.
2. Compute Statistics for Each Pixel's GLCM: A number of statistics have been applied to the GLCM to characterize the texture of a neighborhood, as described by Hall-Beyer (2017), Albregtsen (2008) and Haralick (1973). Generally, they fall into three classes: 1. probability distribution descriptors: mean, variance and energy; 2. linear correlation between the grey-level values of pairs of adjacent pixels and 3. sums of the GLCM cell values (p_{ij}), with each cell weighted by functions of the distance between the cell location in the GLCM and its principal diagonal (which contain the *G* marginal probabilities of co-occurrences of the *same* grey-level. The *GLCM* library in *R* automatically computes eight statistics.

[6] This discussion ignores the treatment of edge effects, which satisfies the sampling requirements of creating marginal probabilities in the last step of creating the GCLM.

[7] It is not the only GLCM estimable for a pixel, as the algorithm can be run horizontally, vertically or diagonally – horizontally either left-right or *vice versa*; alternatively, in the vertical and diagonal cases, up-down or *vice versa*.

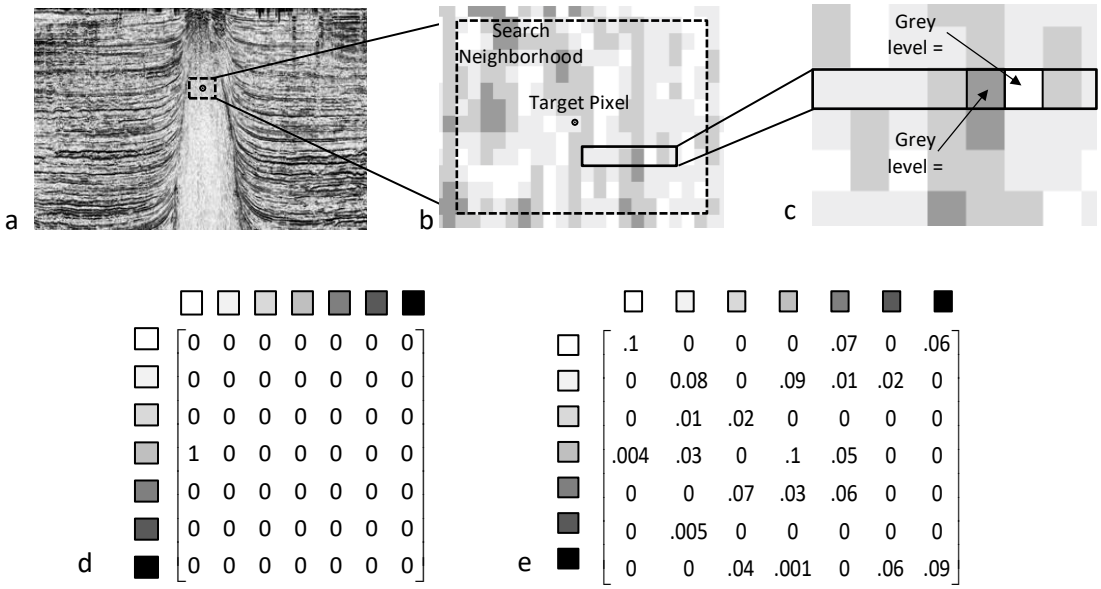


Figure 3 shows the workflow for estimation of the GLCM for a single pixel within a seismic line. Figure 3a shows a piece of the study seismic line and identifies a target pixel for the GLCM analysis (black dot) and a search neighborhood surrounding it, defined symmetrically around the target pixel (black dashed box, in our study, the dimensions of the neighborhood are 51 x 51 pixels). Figure 3b zooms in on the area within the black box in Figure 3a and identifies a piece of a row (black rectangle) along which the co-occurrences are counted row-wise between horizontally adjacent cells (pixels). Figure 3c shows, within the identified row, two horizontally adjacent cells (pixels) that are outlined – the left cell has a grey-scale value of 4 and the adjacent cell to the right has a grey-scale value of 1. Figure 3d shows the single (4 followed by 1) co-occurrence counted in Figure 3c is recorded as an entry of 1 in the 4th row and 1st column (corresponding to the grey-scale levels of the two cells counted). Figure 3e shows the GLCM after *all* pairs of horizontally adjacent cells are counted and all counts in the cells of the matrix are divided by the number of observations in the search neighborhood (in our case, 2,550 (50 x 51) to produce a matrix of marginal probabilities of the co-occurrence of all possible combinations of grey-scale values.

These become attributes of each pixel and therefore mappable for visual assessment of classification accuracy. In experiments over subsets of lines from different surveys, vintages and regions, four GLCM statistics were chosen to maximize the accuracy of classifying salt versus non-salt areas and minimize problems of false positives: contrast, α^1; homogeneity, α^2; dissimilarity, α^3 and entropy α^4 (Figures 4a-4d and Table 2).

Table 1

Transformation from SEG-Y Amplitude Values to Grey Scale Levels*

Grey-level	Bin	Grey Value
1	$\sigma(\lvert A\rvert) \leq 0.1$	0% grey (white)
2	$0.1 > \sigma(\lvert A\rvert) \leq .25$	15% grey
3	$.25 > \sigma(\lvert A\rvert) \leq .5$	30% grey
4	$.5 > \sigma(\lvert A\rvert) \leq 1$	60% grey
5	$1 > \sigma(\lvert A\rvert) \leq 1.5$	80% grey
6	$1.5 > \sigma(\lvert A\rvert) \leq 2$	90% grey
7	$\sigma(\lvert A\rvert) > 2.0$	100% grey (black)

* $|A|$ = absolute value of amplitude.
σ = standard deviation

The functions for three of the four GLCM statistics chosen are closely related. Dissimilarity and contrast vary only by the weighting factor for the distance between each probability cell (p_{ij}) in the GCLM and its principal diagonal. For dissimilarity, the weight is linear and for contrast, it is squared (exaggerating the latter). Homogeneity and contrast are inversely correlated but not exactly. Both are also separate functions of GLCM-energy, which varies with the number and the likeness of non-zero elements in the GLCM (Baraldi and Parmiggiani, 1995). Homogeneity is not an exact inverse of dissimilarity for the same reason.

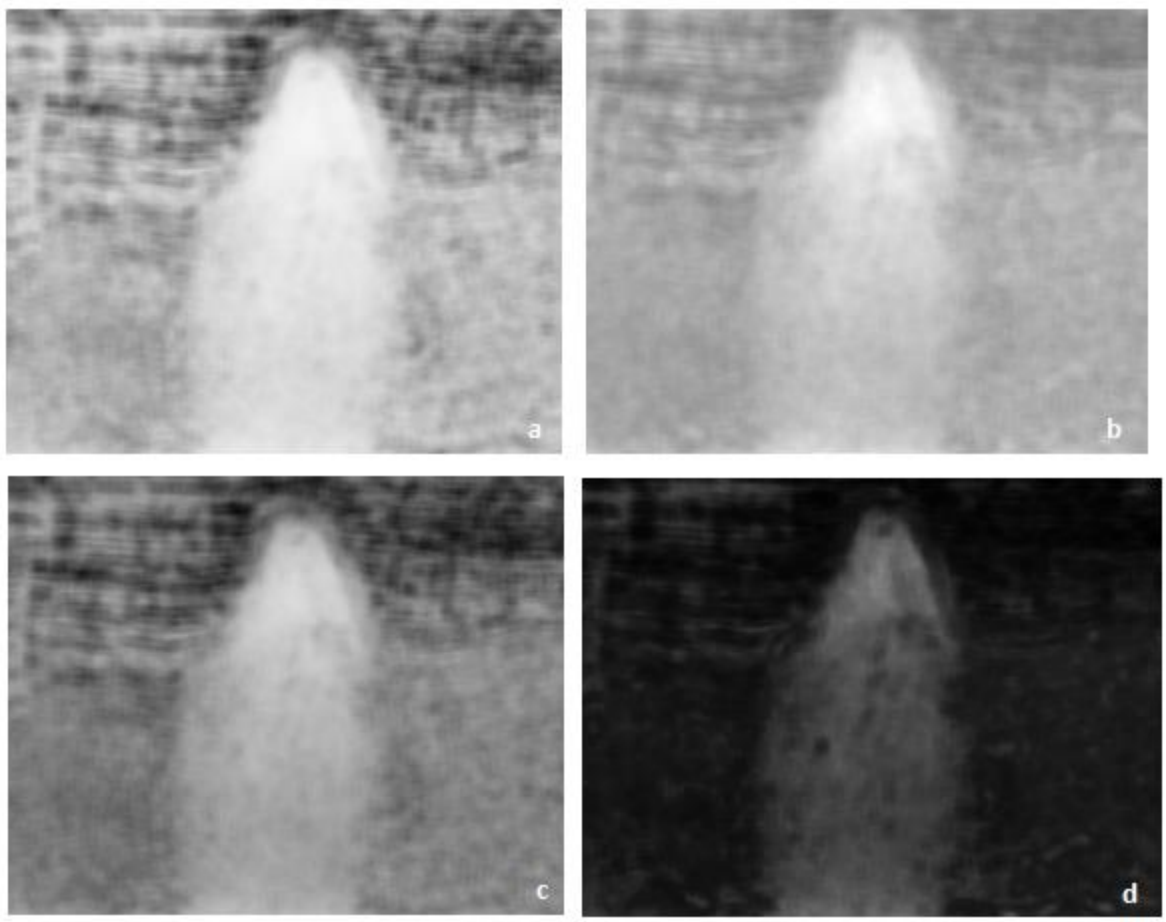
The values of each statistic in Table 2 were normalized to a [0,1] scale and all the GLCM statistics, with the exception of homogeneity, were flipped from their defining equations by defining them as ($1-\alpha_{mn}^{k}$) so they became comparable and 0 was associated with the absence of salt and 1 with its presence.

Table 2

GLCM Statistics Used*

GLCM Statistic	Equation	Measures
Contrast	$\alpha_{mn}^{1} = \sum_{ij=1}^{G} p_{ij}(i-j)^2$	Amount of grey-level variation in search neighborhood.
Homogeneity	$\alpha_{mn}^{2} = \sum_{ij=1}^{G} \left(\frac{p_{ij}}{1+(i-j)^2} \right)$	The smoothness of grey-level distribution in search neighborhood.
Dissimilarity	$\alpha_{mn}^{3} = \sum_{ij=1}^{G} p_{ij}(i-j)$	Similar to Contrast but with a smaller weight on p_{ij}.
Entropy	$\alpha_{mn}^{4} = -\sum_{i=1}^{G}\sum_{j=1}^{G} p_{ij} \ln p_{ij}$	Degree of disorder in the grey-levels of the pixels in the search neighborhood.

* α_{mn}^{k} = the kth GLCM statistic, estimated for a target pixel located at *m,n* in the seismic image
m and *n* = row and column indices of cells (pixels) in the entire seismic line image.
i and *j* = row and column indices for cells the GLCM.
G = the number of levels in the grey-scale.
p_{ij} = empirical marginal probability of co-occurrence of the ith and j^{th} grey-levels within the search neighborhood (i.e. the empirical counts of grey-level *i* horizontally adjacent to grey-level *j* for the search neighborhood, divided by the total number of possible co-occurrences (*N* x (*N*-1)).



Figures 4a – 4d show the results of the GLCM statistics: Figure 4a - contrast, Figure 4b - homogeneity, Figure 4c - dissimilarity and Figure 4d - entropy. Each has been scaled 0 – 1 (black to white), with 0 interpreted as non-salt and 1 as salt.

Vector analysis

To estimate the two vector-analysis based components of the salt score, data from the SEG-Y file were converted to vectors using the *Raster to Polygon* tool of *ArcGIS-Pro* 1.4 (ESRI, 2018) to treat each reflector as an individual feature. The tool groups all contiguous pixels of a common value into a polygon.

First, the data from the SEG-Y were divided into two sets of positive and negative amplitudes. Second, for each of those two sets, the observations were binned by standard deviation using the classes in Table 1. All observations in both the positive and negative amplitude sets $\geq 0.25\sigma$ were retained and labeled lower-intensity reflectors. Third, this process was repeated, using a cut-off $\geq 1\sigma$, to create a set that was labeled higher-intensity reflectors.

Fourth, we applied the *Raster to Polygon* tool to both sets to create polygons around contiguous pixels that could be treated as vector representations of the two classes of reflectors (the reflectors being the centerlines along the long axes of the polygons).

We developed two vector salt measures based on experiments with sample lines. The first exploits an often-observed regularity in diapiric salt on 2D seismic: salt bodies are typically bounded on both sides by relatively short, high-dip, usually intense reflectors. When the records in the salt body itself are eliminated, these bounding reflectors form chevrons going up the salt dome vertically.

In creating the first vector discriminator, the set of higher-intensity reflectors were then subset on two criteria: 1. those with the absolute value of the dip > 40° or, 2. those reflectors from the higher-intensity set that were in the longest 2.5% of the distribution of reflector lengths. The latter group was included because, on visual examination, we observed that sometimes on the boundaries of domes, very long reflectors actually extend from one side of the dome, over the top and down the other side, helping to define the boundary.

An algorithm, written in native *R*, and reading row-wise from left to the right, searched for a positively sloped reflector (with dip > 40°). Once found, the search continued on the row to the right for a "matching" highly dipping, negatively sloped or very strong, long reflector. This algorithm was then applied on the same row in the opposite direction. The pixels in between these markers (i.e., the bounding chevrons) were turned to 1 and those outside the chevrons on that line turned to 0. The analysis was repeated on the next row down until the line was processed. An example of this processing for the study line is shown in Figure 5a.

After the raw image was created, it was edited with morphological tools (see below), eliminating very small features and smoothing edges (Figure 5b). Then zeros (outside salt) and ones (inside) were assigned and this constituted the fifth salt measure, α_{mn}^{5}, for each of pixel (Figure 5c).

The final salt discriminator, α_{mn}^{6}, was based on a hybrid combination of vector and raster data. Particularly for diapiric salt on the shelf, the long, low-dip reflectors of bedded rock surrounding domes characterize where salt is *not*. To find these areas dense in long, low-dip reflectors, the set of vectors was subset to remove the shortest (the bottom 10th percentile of reflectors by length). The remaining (long) reflectors were reconverted to a raster and their density throughout the line was calculated in *R* based on a 51 x 51 moving window around each pixel. The density was rescaled to [0:1], with high densities indicating the absence of salt and low densities evidencing the presence of salt (Figure 5d).

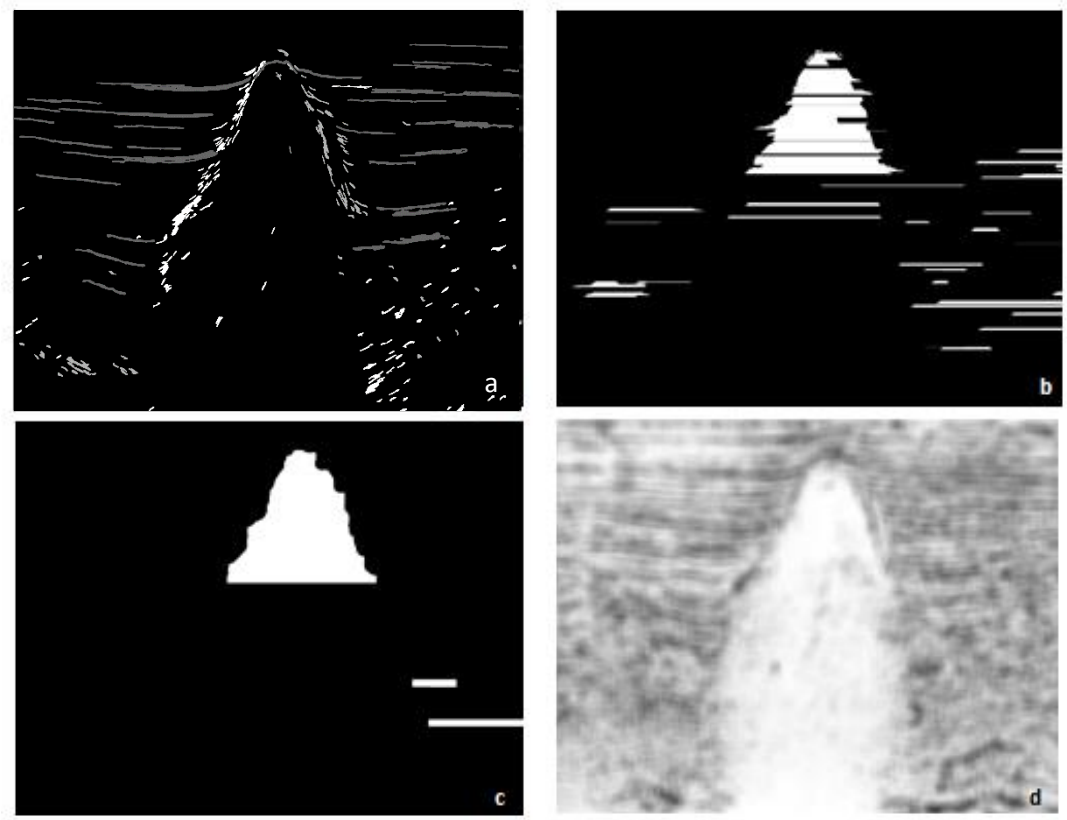


Figure 5a – 5d show the intermediate and final results of analysis of vector data. Figure 5a maps the strong reflectors in the study line. High-dip lines with positive slopes (up to the right) are shown in white; high-dip lines with negative slopes are shown in light grey. Strong low-dip lines are shown in dark grey. The assignment of the space between paired high-dip reflectors to white is shown in Figure 5b. The result of morphological editing of the Figure 5b raw image is shown in Figure 5c . Figure 5d shows the final result of the hybrid vector-raster discriminator, mapping the density of long (and generally low-dip reflectors), indicating the absence of salt.

Salt score and thresholding

Completing both the raster and vector analyses for a single seismic line produced six measures on salt for each pixel. Five of them (the four GLCM statistics, ($\alpha_{mn}^{1},\ldots,\alpha_{mn}^{4}$) and the long-reflector density raster (α_{mn}^{6}), were continuous variables from zero to one. The other vector measure, being inside or outside chevroned high-dip, short reflectors, was treated as a two-state, 0/1, categorical variable, α_{mn}^{5} .

For each pixel in the line, a "salt score" (s_{mn}) was computed as the arithmetic average of the six salt measures ($\alpha^1_{mn},\ldots,\alpha^6_{mn}$). The raw salt score for the study line is shown in Figure 6a. To partition the salt score raster into areas of "salt" and "non-salt", we applied a thresholding algorithm (Otsu (1979)). It reflects a common approach to statistical discriminant analysis: assign samples (pixels) to one of two classes to maximize the between-class variance and minimize the within-class variance of class. This can be algebraically reduced to maximizing the ratio of the between-class variance to the total variance of pixels grey-scale values over the entire image. The "raw" partitioning of the study section into salt/non-salt is shown in Figure 6b.

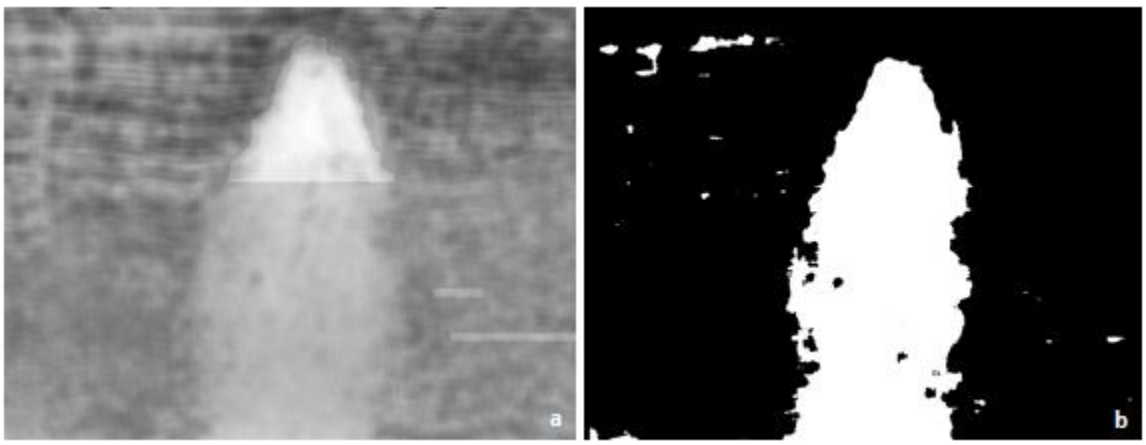


Figure 6a shows the raw salt score computed from the four raster and one pure-vector and one mixed raster/vector classification variables $(\alpha^1,\ldots,\alpha^6)$. Figure 6b shows the transformation of the raw salt score image in Figure 6a into a two-state classification through the application of the Otsu (1979) algorithm.

Morphological processing

The raw salt/non-salt partition in Figure 6b is an image in need of editing to eliminate the effects of our processing. First, remove very small estimated "salt bodies" and "clean" the boundaries of the larger estimated salt bodies, principally to remove high-frequency noise. The opening and closing tools were employed for these tasks, as described by Shafiq (2015) and Wallet and Pepper (2013), Urbach and Wilkinson (2008). The result is shown in Figure 7a. Second, fill in the small holes of "non-salt" within bodies of large salt features. All three were implemented in the *EBImage* package in *R* (Pau, 2010). After all morphological processing was complete, the final time boundaries were extracted (Figure 7b).

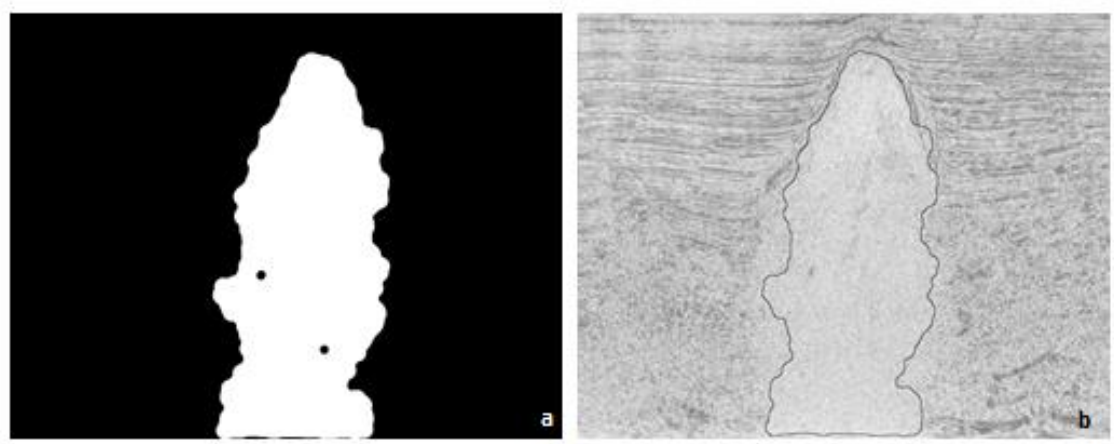


Figure 7a shows the effect of the application of the opening and closing morphological editing tools (the "fill" tool is not separately shown, because its role in this line was to simply fill the two small holes in the interior of the salt in Figure 7a). Figure 7b shows the final, post-morphological editing, salt boundary on the time image of the study line.

Macro-editing

When the analysis, through morphological processing, was complete, the results in time were mapped for a Gulf-wide top of salt map. Inspection revealed some problems at a macro level (i.e., at the survey or regional levels). These fell into four groups:

1. Local Outliers: Certain locations contain multiple seismic lines (often from disparate surveys) in very close proximity. In some of these areas, the results from one line may be different from the others, which create local outliers. Cluster analysis was performed in *ArcGIS* using the *Anselin Local Moran's I* tool (ESRI, 2018) to locate these outliers and remove them.
2. Survey-level acquisition/processing problems: In some surveys, all lines shared characteristics that interfered with our process. A few surveys were visually "too dark" relative to the others, apparently due to different handling of the range of amplitudes. Change in some of the components of the GLCM analysis "cured", or at least, alleviated some of these problems.
3. Local geologic anomalies (?) within surveys: In some cases, the morphology of "salt bodies", as mapped, seemed on inspection to be more likely expressions of geologic features that were *not* salt but interfered with our process. This occurred around the mouth of the current course of the Mississippi River; it was also mirrored at two previous Mississippi courses as they came offshore. They may be distributary mouth

bars or related deltaic facies that produced, from a salt discrimination viewpoint, false positives. The questionable data within the lines affected were removed.

4. Survey-level problems of unknown provenance: In a few cases, large areas of contiguous shallow "salt" were mapped, apparently correlated with survey boundaries. Inspection yielded no corresponding salt and the source of the irregularities could not be identified. No surgical excision or algorithm modification could suitably resolve the problem, so the surveys were dropped (e.g., from the Brazos protraction area on the Texas shelf to the south).

Measuring confidence

In analyzing more than 8,000 seismic lines, with no pretense to hand-check the quality and reliability of each one individually, communicating certainty in the analysis was critical to building a technologically useful process. Therefore, an automated assessment was applied to every salt boundary estimated in the study. At the next level of aggregation, the average certainty score for each boundary in a line was projected to the points overlying its lateral extent onto the navigation data of the line for regional mapping.

We adopted two screening criteria. The principal measure was the gradient of texture (GoT), used to assess the degree to which our estimated boundaries separated distinct textures within a line, as shown by Hegazy, et al (2015). The GoT measures the rate of a change in a diagnostic variable describing texture between adjoining parts of the image.[8]

Mechanically, a small rectangle is centered on a pixel in a boundary, with its long axis perpendicular to the boundary at that point. The rectangle is divided in half at the boundary (Figures 8a and 8b). The average pixel intensity was computed in each half of the rectangle (i.e., inside and outside the boundary) and the difference taken between the averages on the two sides of the boundary. If there is a significant difference in the pixel intensity between the two halves of the rectangle, that evidences a significant difference in seismic texture across the boundary and, therefore, it is appropriately located.

The second criterion was based on visual inspection of the estimated boundaries. Our process produced a large number of very small "salt bodies" that were very likely effects of the analysis, rather than lithologic discrimination. So, we eliminated the shortest (by pixels) 75% of the boundaries; we also removed 50% of the lowest GoT scored-boundaries.

[8] Mathematically, the GoT is a directional first difference, not a gradient, but the term captures the spirit of the analysis and is widely used.

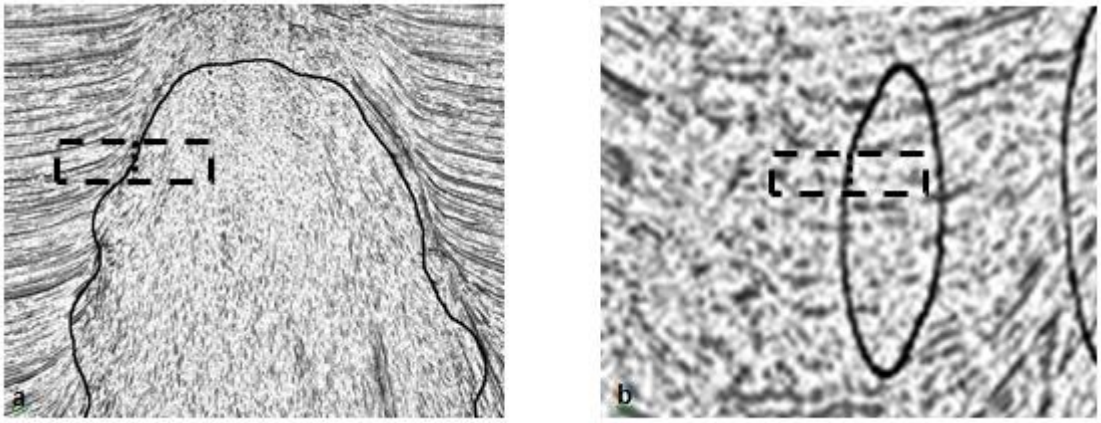


Figure 8a and 8b show the principle behind gradient of texture (GoT). In Figure 8a, a bisected rectangle is centered on a pixel in the estimated boundary of a salt dome. Using pixel intensity as the comparison variable, the intensity of pixels within the rectangle outside the boundary is clearly different than for the pixels interior to the boundary. In Figure 8b, the rectangle is again centered on the boundary. The average pixel intensity of both sides of the rectangle is nearly the same, so seismic textures are assumed to be the same and the boundary was removed.

The GoT scores for the remaining boundaries that passed these thresholds were divided by quartile for grades: "A" for the top 25% of GoT scores; "B" for the next 25% down, "C" for the second to lowest class and "D" for the bottom quartile. Finally, also based on visual inspection and mindful of the age of the seismic used in the study, boundaries estimated at depths greater than 6 seconds were dropped.

Depth conversion and mapping

The time to top of salt was extracted from the estimated salt/non-salt boundaries in each line and translated vertically to sit along the navigation line. These points were then converted to depth based on a velocity field estimated from 3,593 velocity surveys. These surveys were submitted to the US government by operators and are publicly available (BOEM, 2018). While velocity typically increases with depth, to test the simplest model of the relationship between time and depth, linear regressions (implying constant velocities at the locations of the surveys) were estimated. The average goodness-of-fit statistic (R^2) for the regressions was 0.99, which justified using that functional form and its parameters as a good first-order approximation of the local velocity field. The 3,593 estimates of average velocity were gridded for a Gulf-wide velocity map using an anisotropic ordinary kriging model (ESRI,

2018). These estimated velocities were used to convert time to top of salt to depth for every point in the study area.

Because the original data from the study included dozens of separately shot surveys, in some locations, multiple lines of different origins were in very close proximity. As the analysis of each line was performed independently, a single location (or a very small neighborhood around a single location) could have multiple depths to top of salt. If inverse distance weighting (IDW) gridding was directly applied to the union of all points, the result would be an extremely high variance surface that was much more a reflection of the diversity of the data rather than the depth of the top of salt.

Therefore, the following algorithm was adopted. First, a grid of base points, 500 m (1,640 ft) apart was established over the study area. Around each base point, all points representing the depth to the top of salt within 350 m (1,148 ft) of each point were found and grouped by their line number. Then all observations from a single line in that group were averaged. Those line-specific average depths in that neighborhood that were more than 25% deeper than the shallowest average were dropped from the calculation. In a total, 21,716 points were eliminated from the total of 241,416; then IDW was applied, with a power of 1, to estimate the depth to top of salt surface (ESRI, 2018). The final clipping rule was applied so that no part of the estimated surface extended more than 2,000 m (6,560 ft) from the nearest control.

COMPUTATIONAL COMPLEXITY

On a single desktop computer, running an Intel i5-3450, 3.10 GHz, 64-bit processor, sequential processing of 8,043 seismic lines in this study required approximately 645 hours. Raster analysis occupied roughly 62% of this time, 31% for the vector analysis and 7% for GoT, input/output and miscellaneous tasks. Our computer configuration, while not “low-end”, was specifically kept to a minimalist configuration in the spirit of measuring computational complexity for the “worst case”.

The computational complexity of processes in this workflow, for each line, were linear in (mn), the seismic image dimensions in pixels. In the case of the GLCM, the whole GLCM process had a computational complexity of $\mathcal{O}(N^2mn)$; while this expression is linear in (mn), it is not in N, which is the size of the GLCM search neighborhood.

Because of the analytic independence of processing each line, a linear gain in computation time can be made by parallelizing the project across multiple processors. Likewise, because of the centrality of matrix manipulation in our approach (particularly, GLCM), switching from standard central processing unit to a graphics processing unit (GPU)-technology based hardware platform also seems a very effective step in accelerating the analysis (perhaps as

much as 10x). Therefore, a 10-processor-GPU array could shrink computer time by a factor of as much as 50 to 100 – reducing run time from a little less than a month to, perhaps, less than three days.

RESULTS AND DISCUSSION

The principal result of the project is shown in Figure 9, a basin-wide depth to top of salt map. From it, we are in a position to judge the present research on the four criteria set forth at the outset.

First, did the workflow and its components yield technically useful results? The study area in Figure 9 is 278,000 km^2 (107,000 mi^2). Of the total study area, 17% had insufficient seismic coverage for our analysis. We removed 16% of the portion we *could* map from the final products for failure to meet our quality control thresholds. Qualitatively, we compared the depth to salt by inspection of individual lines across the basin. Misses occurred, of course. However, the final maps meet the needs of regional to basin-wide geologic mapping – one does not plan a well on a basin-wide map. The practical utility of our final mapping was dramatically improved by the systematic reporting and use of assessment of certainty in the analysis (Figure 10).

Our top of salt compared well with published regional and basin-wide work (Ganey-Curry, *et al,* 2018, Fort and Brun, 2012). A larger scale assessment of accuracy on the shelf is shown in two comparisons in Figure 11. In that map, which focuses on the central part of the GOM shelf, south of Louisiana, our estimated shallow and intermediate diapiric salt bodies are shown in deep green. Compared to these are the results of a project in the mid-1990s to hand map shallow salt across the GOM shelf (Hentz, *et al*, 1997). The shallow/diapiric salt they found is symbolized in Figure 11 by purple polygons. Finally, Figure 11, includes field outlines for oil and gas fields identified as trapped by intermediate to shallow salt. These classifications are made by the US government on the basis of data submitted to them by operators (BOEM, 2018). The outlines of those fields appear in Figure 11 as cyan stripped polygons. Our mapping of intermediate to shallow salt in this area has a very high spatial correlation with the results of *both* independent studies.

These comparisons lend support to the validity of our mapping of *top* of salt and do not provide guidance on the feasibility of mapping salt thickness or the "roots" of salt structures, particularly at substantial depths (which motivated the decision to not to map below 6 seconds). As well, while taking effective steps to reduce false positives in the analysis (mapping salt where there appears to be none). At this stage of the research, we still do not have a scale-robust method for identifying false negatives (mapping no salt where salt appears on visual examination). Very small holes (i.e., non-salt) in the middle of large salt bodies are removed by

morphological editing. However, a method of automatically identifying a seismic line with *no* salt is an important goal that has not be achieved yet.

Second, is the approach scalable? Although we made methodological additions and amendments, most of the individual salt discrimination methods used were previously introduced in the geophysical literature, originally using a very small number of lines or synthetic data. Testing the limits of their scale of application (as modified by us) was a fundamental project goal. At a basin scale, the quality of our results would have suffered had we not developed the process of macro-editing, or looking for errors on a survey or regional scale.

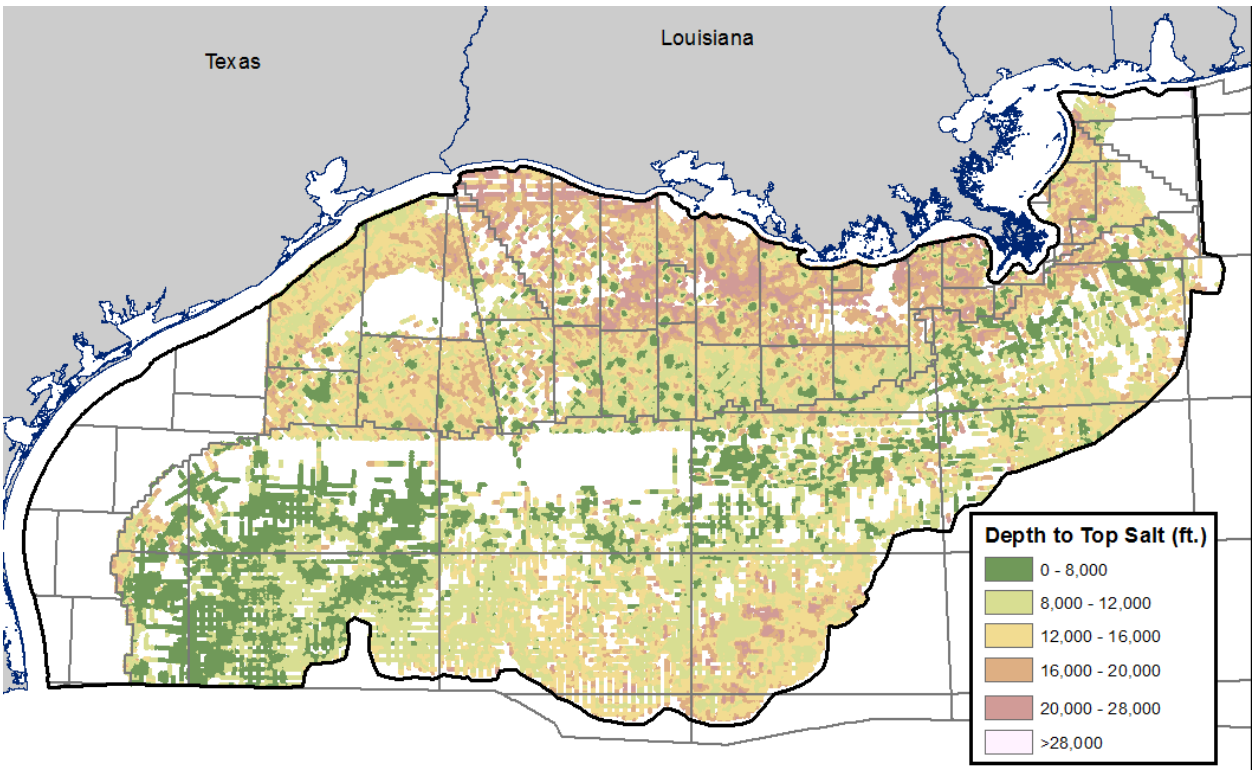


Figure 9 shows the final depth to top of salt map for the study area within the US portion of the Gulf of Mexico. The dark black line shows the limits of the study area. White areas within the study limits were not mapped either because there was no data (or insufficiently dense data) within them, or, the confidence in the estimated depth to top of salt did not meet the quality control thresholds. The grey boundaries in the map represent US government protraction areas (which organize leasing).

Nevertheless, at least through a project of 8,000+ lines, we encountered no obvious limitation on the volume of input data susceptible to a workflow of structured, unsupervised machine learning techniques. Also, with respect to scale, the practical utility of the study to

working geoscientists without an included analysis of uncertainty shown in Figure 10 would have been lower. With the basin-wide certainty mapping, as well as assessments of every individual boundary, a geologist or geophysicist can zoom-in on his or her own project area and automatically see consistent evaluation of the reliability of this work.

Third, the analysis appeared robust in the face of the diversity of 82 input surveys, acquired between 1981 and 1992. The macro-editing procedure specifically attacked survey-level problems, but with that tool in hand, the effect of variance in the input data appears to have been minimized (or at least, significantly reduced). Our macro-editing approach, however, while correcting regional errors and improving accuracy does not change the *regional* nature of our mapping or make this a high-precision estimate of boundaries (especially on the flanks of salt structures).

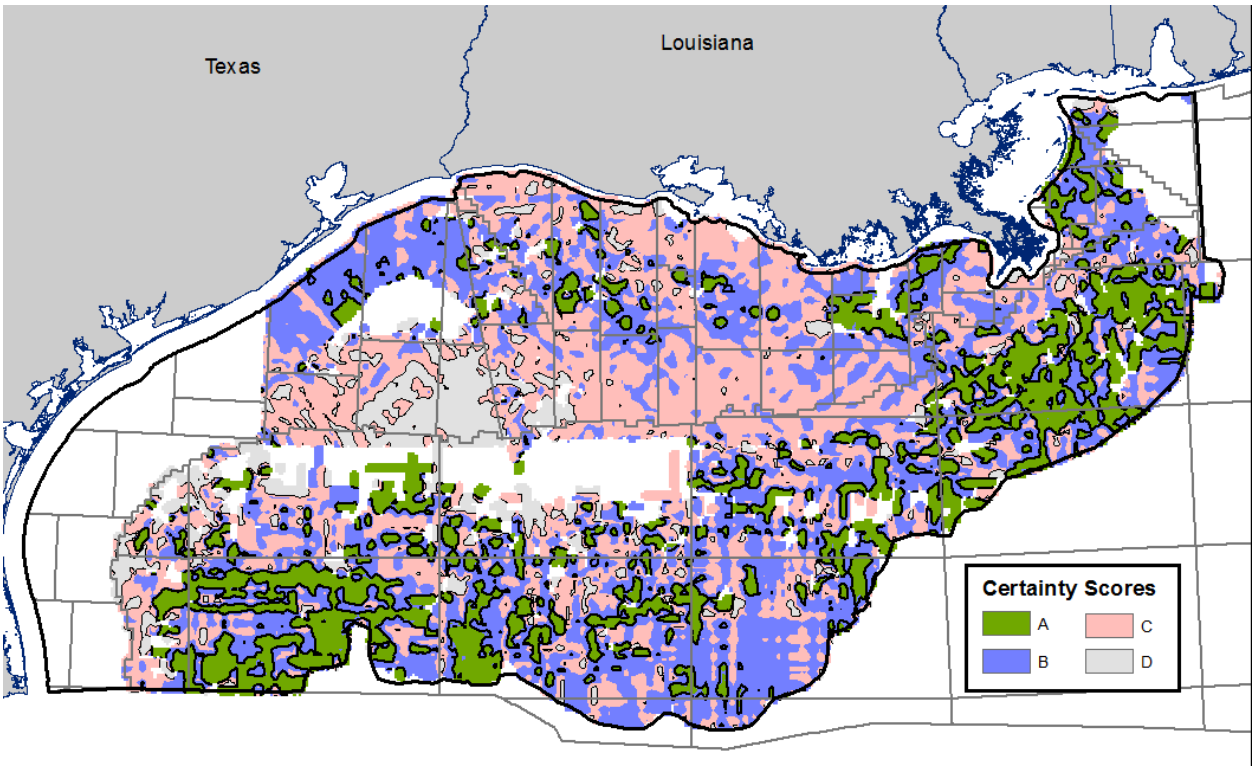


Figure 10 shows the final confidence grades assessed for each estimated salt/non-salt boundary. The assessments were restricted to the top 25% of all boundaries based on their length (eliminating small boundaries that were likely processing effects) and the top half of certainty scores, based on gradient of texture (GoT).

Of course, accuracy would be immeasurably improved had we based the project on the newest and technically homogeneous data; as well, there would have been a smaller area removed for quality control reasons. But there is nothing in the methodology that *requires* old

or diverse data, it will all work much better with the "latest and greatest" inputs. Nevertheless, using *this* data set tested the utility of a machine learning approach to exploit the *massive idle stores of older data* and that goal was accomplished with a positive outcome. That it will work with superior data is just a (very large) "up-side".

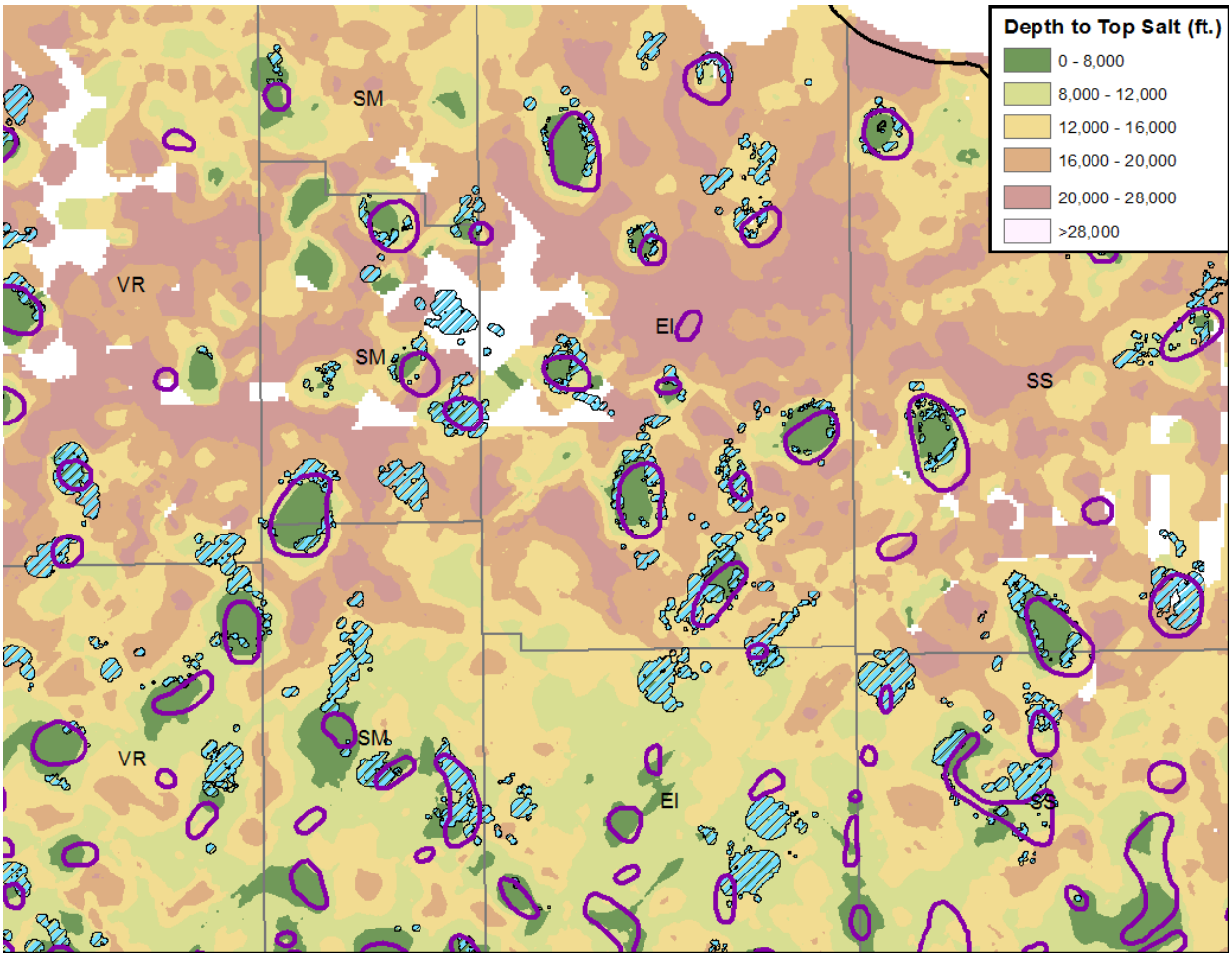


Figure 11 shows the correspondence on the shelf between salt domes estimated in this study (dark green) and two published studies. The blue striped polygons are the outlines for fields identified with shallow salt trapping (BOEM, 2018). The purple polygons are hand-mapped shallow salt (Hentz, *et al*, 1997). The labeled protraction areas are: VR=Vermillion, SM=South Marsh Island, EI=Eugene Island and SS=Ship Shoal.

The final dimension of performance was economic: does the value derived from machine-learning analysis, such as presented here, exceed the marginal cost of conducting it? Ultimately, this question must be left to companies holding seismic data resources similar to what we used (or better). However, in answering this question, realize that the original costs of

acquisition and processing of these thousands of lines were “sunk” long ago. Also, recognize that, although further research and development of our approach is definitely warranted, its forward costs are relatively low. Therefore, if this project is representative, it is hard to imagine that the economic benefit of this approach does not exceed by many times the costs of its application.

CONCLUSIONS

We purposely took a scientifically and economically important target: mapping salt. Salt is central to the habitat of hydrocarbons in the GOM and a host of highly petroliferous basins around the world. Small-scale studies had shown that salt could be located but our research extended those findings in scale, volume and difficulty of input data quality. These successes, in turn, open the door to potentially very profitable re-use of massive vintage seismic resources that today mainly occupy storage space.

When we move from the shelf to the slope, there are still extensive areas where the certainty in the analysis remains high. Indeed, there are some areas on the slope, where the horizontal, canopy-geometries dominate, that the identification of salt appears very good. Because salt bodies are thinner, there are even some cases where both the top *and bottom* of salt are clearly distinguished from the surrounding rock. However, the results also show that spatial variance in the assessment of certainty of results is higher than on the shelf. That is, on the shelf, there are broad and continuous areas receiving the same certainty score. On the slope, while some areas have high certainty scores, they tend to be smaller in area and less continuous. The variable morphology of salt on the shelf probably drives this difference and suggests that experiments with different models between these geologically disparate provinces are justified in future research.

On a more general level, it must also be admitted that, in its classic embodiment as diapirs on a wide, passive margin shelf, salt is one of the easier targets for machine learning. More nuanced features will pose stiffer challenges: faults, facies, rock properties and fluid types are all critically important. Features that synthesize both empirical geophysical properties and geologic theory, like facies, will probably raise the greatest challenges. Yet without success using these techniques, applied to salt on an industrial scale, the motivation to move on to more complex tasks would be undermined.

In testing a new technology, its limitations must be as clear as its advantages. Most broadly, nothing from our experience supports the popular notion that data may be simply dumped into machine learning algorithm and useful answers magically appear. Expertise and *many* experiments were required to design the workflow, find efficient and reliable discriminators, test the results and respond to systematic errors found. Moreover, irrespective of the provenance of input data, when scale exceeds the limit of hand checking output,

automated quality control assessment *must* accompany the results for them to be accepted and put to practical use.

This research demonstrated the value of machine learning to extract valuable regional geologic information from a massive and diverse collection of geophysical data. Equally important, the component tools we applied are broadly accessible to those who sit on the massive volumes of data to which they can be applied. After all, it is more experience in application that will best distill the technical and economic value of machine learning from the lustrous aura that surrounds it.

DATA AND CODE AVAILABILITY

All data used in this study is publicly available. The raw SEG-Y and associated SEG-P1 seismic data are available from the US Bureau of Ocean Energy Management (BOEM) via a web site of the US Geological Survey at https://walrus.wr.usgs.gov/namss/search/. The data for velocity surveys and geopolitical/regulatory data (e.g., lease blocks, protraction areas and state boundaries) are available through a web site of the BOEM, https://www.data.boem.gov/. The code used in this study is proprietary.

REFERENCES

Albregtsen, F., 2008, Statistical texture measures computed from gray level coocurrence matrices: Department of Infomatics, University of Oslo, 1-14 https://www.uio.no/studier/emner/matnat/ifi/INF4300/h08/undervisningsmateriale/glcm.pdf, accessed 20 July 2018.

Amin, A., M. Direche, T. Hegazy, Z. Wang and G. AlRegib, 2015, A novel approach for salt dome detection using a dictionary-based classifier: 85th Annual International Meeting, SEG, Expanded Abstracts, 1816-1820, http://dx.doi.org/10.1190/segam2015-5925748.1.

Araya-Polo , M., T. Dahlke , C. Frogner , C. Zhang , T. Poggio , and D. Hohl , 2017, Automated fault detection without seismic processing: The Leading Edge, **36**, 208–214.

Baraldi, A. and F. Parmiggiani, 1995, An investigation of textural characteristics associated with gray level coocurrence matrix statistical parameters: IEEE Transactions on Geoscience and Remote Sensing, **33**, 293-304.

Berthelot, A., A.H. Solberg and J.-L. Gelius, 2013, Texture attributes for detection of salt: Journal of Applied Geophysics, **88**, 52-69, https://doi.org/10.1016/j.jappgeo.2012.09.006.
CDA, accessed at www.ukoilandgasdata.com, accessed, May 2018.

Coléou, T., M. Poupon and K. Azbel, 2003, Unsupervised seismic facies classification: A review and comparison of techniques and implementation: The Leading Edge, **22**, 943-953, https://doi.org/10.1190/1.1623635.

De Matos, M.C., Marfurt K.J., P.R.S. Johann, 2010, Seismic interpretation of self-organizing maps using 2D color displays: Revista Brasileira de Geofisica, **28**, 631-642.
Di, H., M. Shafiq and G. AlRegib, 2017, Multi-attribute K-means cluster analysis for salt boundary detection: 79th Conference and Exhibition, EAGE, Extended Abstracts, https://doi.org/10.3997/2214-4609.201700915.

ESRI, 2018, ArcGIS/ArcGIS-Pro documentation, Redlands, CA. For the Raster to Polygon Tool: http://pro.arcgis.com/en/pro-app/tool-reference/conversion/raster-to-polygon.htm; for kriging and inverse distance weighting grid estimation: https://www.esri.com/en-us/arcgis/products/geostatistical-analyst/overview; for Anselin Local Moran's I tool, http://resources.esri.com/help/9.3/arcgisengine/java/gp_toolref/spatial_statistics_tools/cluste

r_and_outlier_analysis_colon_anselin_local_moran_s_i_spatial_statistics_.htm, accessed 21 June 2018.

Fort, X and J-P. Brun, 2012, Kinematics of regional salt flow in the northern Gulf of Mexico: Geological Society, London, Special Publications, **363**, 265-287, doi: https://doi.org/10.1144/SP363.12.

Ganey-Curry, P.E., editor, 2018, Gulf basin depositional synthesis (GBDS) phase XI year 20 atlas: University of Texas, Institute of Geophysics.

Galvis, I.S., Y. Villa, C. Duarte , D. Sierra , and W. Agudelo , 2017, Seismic attribute selection and clustering to detect and classify surface waves in multicomponent seismic data by using k-means algorithm: The Leading Edge, **36**, 239–248.

Gao, D., 2003, Volume texture extraction for 3D seismic visualization and interpretation: Geophysics, **68**, 1294-1302.

Gao, D., 2011, Latest developments in seismic texture analysis for subsurface structure, facies, and reservoir characterization: A review: Geophysics, **76**, no. 2, W-1-W-23.

Halpert, A. and R.G. Clapp, 2008, Salt body segmentation with dip and frequency attributes: Stanford Exploration Project, http://sepwww.stanford.edu/data/media/public/docs/sep136/adam1/paper.pdf, accessed 18 May 2018.

Hall-Beyer, M., 2017, GLCM texture: a tutorial: Department of Geography, University of Calgary, **3**, https://prism.ucalgary.ca/bitstream/handle/1880/51900/texture%20tutorial%20v%203_0%20180206.pdf?sequence=11&isAllowed=y, accessed 15 May 2018.

Haralick, R.M., K. Shanmugam and I. Dinstein, 1973, Textural features for image classification: IEEE Transactions on Systems, Man and Cybernetics, **SMC-3**, 610-621, doi: https://doi.org/10.1109/TSMC.1973.4309314.

Hegazy, T., and G. Alregib, 2014, Texture attributes for detection salt bodies in seismic data: 84th Annual International Meeting, SEG, Expanded Abstracts, http://doi.org/10.1190/segam2014-1512.1.

Hegazy, T., Z. Wang and A. Ghassan, 2015, The role of perceptual texture dissimilarity in automating seismic data interpretation: 2015 IEEE Global Conference on Signal and Information Processing, doi: https://www.doi.org/10.1109/GlobalSIP.2015.7418168.

Hentz, T.F., S.J. Seni and E.G. Wermund, Jr., 1997, Atlas of Northern Gulf of Mexico Reservoirs: Bureau of Economic Geology, Austin, TX.

Otsu, N., 1979, A threshold selection method from gray-level histograms: IEEE Transactions on Systems, Man and Cybernetics, **SMC-9**, 62-66, doi: https://www.doi.org/10.1109/TSMC.1979.4310076.

Pau , G., F. Fuchs, O. Sklyar, M. Boutros and W. Huber (2010): EBImage - an R package for image processing with applications to cellular phenotypes. Bioinformatics, 26(7), pp. 979-981, doi: https://www.doi.org/10.1093/bioinformatics/btq046.

R Core Team, 2013, R: A language and environment for statistical computing, www.R-Project.org, accessed May 2018.

Shafiq, M.A., Y. Alaudah and G. AlRegib, 2017, Salt dome delineation using edge- and texture-based attributes: 79th EAGE Conference & Exhibition 2017, doi: https://doi.org/10.3997/2214-4609.201700710.

Shafiq, M.A., , Y. Alaudah, H. Di and G. AlRegib, 2017b, Salt dome detection within migrated seismic volumes using phase congruency: SEG Technical Program Expanded Abstracts 2017, 2360-236, doi: https://doi.org/10.1190/segam2017-17443246.1.

Shafiq, M.A., Z. Wang, A. Amin, T. Hegazy, M. Diriche and G. AlRegib, 2015, Detection of salt-dome boundary surfaces in migrated seismic volumes using gradient textures: SEG Expanded Abstracts, 1811-1815, doi: https://doi.org/10.1190/segam2015-5927230.1.

United States Bureau of Ocean Energy Management, 2018, original data available through their public web site, https://www.data.boem.gov/Main/Default.aspx, accessed 21 June 2018.
United States Geological Survey, http://walrus.wr.usgs.gov/NAMSS/ , accessed 21 June 2018.

Urbach, E. R. and M.H.F. Wilkinson, M. H. F., 2008, Efficient 2-D grayscale morphological transformations with arbitrary flat structuring elements: IEEE Transactions on Image Processing, **17**, 1-8, doi: https://doi.org/10.1109/TIP.2007.912582.

Wallet, B.C. and R.E.F. Pepper, 2013, Using mathematical morphology in an attribute workflow to improve the interpretability of salt bodies in the Gulf of Mexico: 2013 SEG Annual Meeting, 1324-1328, doi: https://doi.org/10.1190/segam2013-0941.1.

Wang, Z.T., T. Hegazy, Z. Long and G. AlRegib, 2015, Noise-robust detection and tracking of salt domes in postmigrated volumes using texture, tensors and subspace learning: Geophysics, **80**, no. 6, WD101- WD11, doi: https://doi.org/10.1190/geo2015-0116.1.

Wu, X. 2016, Methods to compute salt likelihoods and extract salt boundaries from 3D seismic images: Geophysics, **81**, no. 6, doi: https://doi.org/10.1190/geo2016-0250.1.

Zvoleff, A., 2016, Calculate textures from grey-level co-occurrence matrices (GLCMs): Package 'glcm', https://cran.r-project.org/web/packages/glcm/glcm.pdf, accessed 20 June 2018.