

Statistical approach to neural network imaging of karst systems in 3D seismic reflection data

Daniel R. Ebuna¹, Jared W. Kluesner², Kevin J. Cunningham³, and Joel H. Edwards¹

Abstract

The current lack of a robust standardized technique for geophysical mapping of karst systems can be attributed to the complexity of the environment and prior technological limitations. Abrupt lateral variations in physical properties that are inherent to karst systems generate significant geophysical noise, challenging conventional seismic signal processing and interpretation. The application of neural networks (NNs) to multi-attribute seismic interpretation can provide a semiautomated method for identifying and leveraging the non-linear relationships exhibited among seismic attributes. The ambiguity generally associated with designing NNs for seismic object detection can be reduced via statistical analysis of the extracted attribute data. A data-driven approach to selecting the appropriate set of input seismic attributes, as well as the locations and suggested number of training examples, provides a more objective and computationally efficient method for identifying karst systems using reflection seismology. This statistically optimized NN technique is demonstrated using 3D seismic reflection data collected from the southeastern portion of the Florida carbonate platform. Several dimensionality reduction methods are applied, and the resulting karst probability models are evaluated relative to one another based on quantitative and qualitative criteria. Comparing the preferred model, using quadratic discriminant analysis, with previously available seismic object detection workflows demonstrates the karst-specific nature of the tool. Results suggest that the karst multiattribute workflow presented is capable of approximating the structural boundaries of karst systems with more accuracy and efficiency than a human counterpart or previously presented seismic interpretation schemes. This objective technique, using solely 3D seismic reflection data, is proposed as a practical approach to mapping karst systems for subsequent hydrogeologic modeling.

Introduction

The discipline of geophysics is well-suited to benefit from the advantages presented by advanced data analysis and machine learning techniques because it deals with large data sets representing remotely sensed physical properties, many of which are correlated. Modern computing provides the ability to rapidly sift through these data to identify complex, often nonlinear relationships that could potentially aid geologic interpretation. One of the primary tools for subsurface geologic interpretation is seismic reflection imaging, which produces data that can be transformed into a variety of useful seismic attributes (Scales, 1995; Taner, 2001). Early work in attempting to identify relationships among individual seismic attributes was performed by Aminzadeh and Chatterjee (1984) using principal component analysis (PCA) and later clustering. Other methods for combining and evaluating the correlations and relative contributions of different seismic attributes soon fol-

lowed, including simple arithmetic combination, crossplotting, regression techniques, fuzzy logic, and self-organizing maps (SOMs). More recently, neural networks (NNs) have emerged as highly effective tools for combining seismic attributes into multiattributes (or “meta-attributes”) for seismic object detection due to their high noise tolerance and ability to recognize nonlinear relationships in the data (Russell et al., 1997; Hampson et al., 2001; Aminzadeh and de Groot, 2005; Brouwer et al., 2011).

Seismic reflection multiattribute workflows have been designed to identify specific geologic features, such as faults, chimneys, salt diapirs, direct hydrocarbon indicators, and slumps, often through a relatively subjective process (de Groot et al., 2001). This study uses an objective approach to designing a supervised NN classifier for 3D imaging of karst systems, whereby statistical analyses are used to determine the optimal set of input seismic attributes, training example

¹University of California — Santa Cruz, Santa Cruz, California, USA. E-mail: debuna@ucsc.edu; jhedward@ucsc.edu.

²Pacific Coastal and Marine Science Center, U.S. Geological Survey, Santa Cruz, California, USA. E-mail: jkluesner@usgs.gov.

³Florida Water Science Center, U.S. Geological Survey, Fort Lauderdale, Florida, USA. E-mail: kcunning@usgs.gov.

Manuscript received by the Editor 1 November 2017; revised manuscript received 13 February 2018; published ahead of production 24 April 2018; published online 26 June 2018. This paper appears in *Interpretation*, Vol. 6, No. 3 (August 2018); p. B15–B35, 13 FIGS., 2 TABLES.

<http://dx.doi.org/10.1190/INT-2017-0197.1>. © 2018 Society of Exploration Geophysicists and American Association of Petroleum Geologists. All rights reserved.

locations, and the predicted minimum number of training samples required for good generalization. The statistical techniques evaluated include PCA, quadratic discriminant analysis (QDA), nonmetric multidimensional scaling (NMDS), Spearman's rho, Welch's T -test, and smooth regression via the gamma test (SRGT).

The purpose of this study is to (1) provide a new multiattribute workflow for imaging karst systems in seismic reflection data with minimal interpreter bias and (2) demonstrate the advantages of this workflow for generating high-resolution 3D geomodels of an Eocene carbonate karst system in the subsurface of the southeastern Florida Platform. The resulting karst multiattribute workflow (KMW) is a new NN seismic interpretation tool designed specifically to target the seismic response associated with the physical properties and geometries exhibited by karst systems. The performance of this approach to seismic object detection and imaging is evaluated quantitatively in terms of calculated error and accuracy, and qualitatively by comparing the spatial agreement of the model output with analog karst features interpreted elsewhere from seis-

mic reflection amplitude data. The efficacy of the KMW at identifying karst systems specifically is further demonstrated through direct spatial comparisons with the results from previously available seismic multiattribute workflows that target other features.

Geologic setting

This study is based on a 6.2 km² 3D seismic survey located on the southeastern part of the Florida Platform, which is a mostly carbonate embankment between the Gulf of Mexico and the eastern Atlantic Ocean (Figure 1). The Florida Platform developed from the Cretaceous Period to the Cenozoic Era through repeated cycles of sedimentation and subsidence, creating a series of predominantly shallow-water carbonate and evaporite deposits estimated to be approximately 6700 m thick (Hine, 2013; Cunningham et al., 2018). This study focuses on seismic reflection data that image the early to middle Eocene part of the platform composed of the Oldsmar and Avon Park Formations. These two formations are characterized by generally flat-lying carbonate depositional sequences composed of mostly vertically stacked

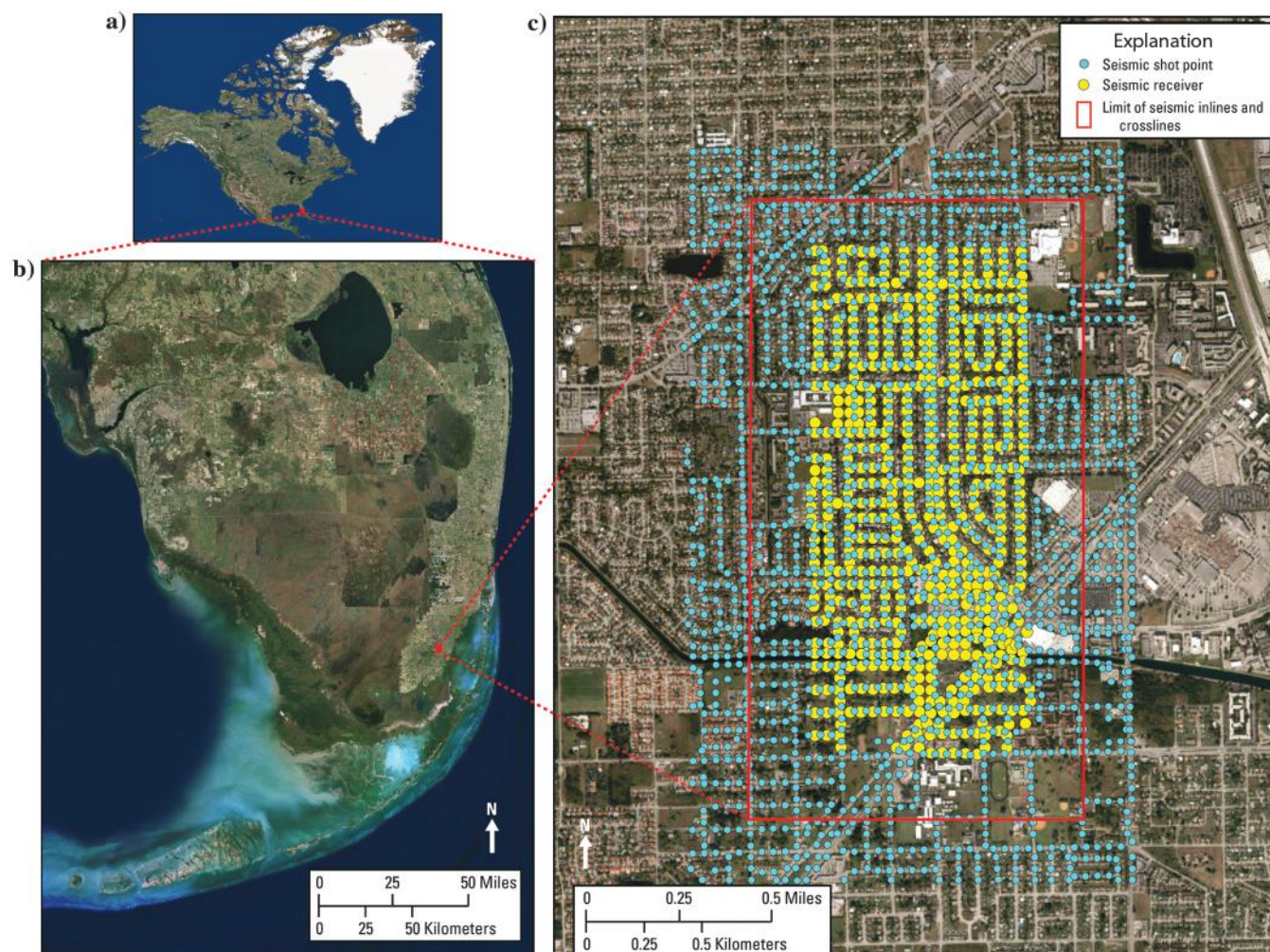


Figure 1. Location map of the study 3D seismic survey site on the southeastern part of the Florida Platform with seismic shot-points and receiver locations shown.

peritidal cycles. These carbonate rocks compose a major part of the karst Floridan aquifer system and several highly permeable water-bearing zones separated by semiconfining units (Miller, 1986; Reese and Cunningham, 2014; Cunningham, 2015; Cunningham et al., 2018). The lowermost coherent seismic reflections within the 3D seismic reflection data set correspond to the Oldsmar Formation. Evidence of faulting related to epigenic and hypogenic karst and strike-slip regional faulting has been identified in the Oldsmar and Avon Park Formations in areas nearby the study area (Reese and Cunningham, 2014; Cunningham, 2015; Cunningham et al., 2018). Regional faults, trending approximately northwest–southeast and southwest–northeast, likely influenced early vertical fluid migration and plausibly played an integral role in the spatial development of the study area karst system. A karst system is “a fluid-flow system (geohydrodynamic system) with a permeability structure evolved of initial preferential flow pathways, dominated by interconnected voids and conduits, and organized to facilitate the circulation of fluid in the down gradient direction due to the positive feedback between flow and conduit growth” (Klimchouk, 2015).

Several studies have investigated the origin of buried paleosinkholes and columnar seismic sag structures (see McDonnell, 2007) observed in seismic reflection data from onshore and offshore survey areas within the southeastern Florida Platform (Cunningham and Walker, 2009; Cunningham, 2015; Cunningham et al., 2018). The seismic sag structures are related to hypogenic karstification (Cunningham et al., 2018). Plausible hydrochemical scenarios invoking the dissolution of carbonate rock during hypogenic karstification include upward flow of mixed fresh and saline groundwater driven by Kohout convection or ascending hydrogen-sulfide-rich groundwater charged by the dissolution and reduction of calcium sulfates in the deeper formations and oxidation to produce sulfuric acid (Cunningham and Walker, 2009).

Methods

The 3D seismic reflection data used in this study were acquired in 2015 from an urban survey site in Miami-Dade County (Figure 1) as part of a southeastern Florida geologic and hydrogeologic framework study (Cunningham, 2014). Data were collected using a single truck-mounted Inova Univib seismic vibrator with a 22 s, 8–128 Hz linear sweep with four sweeps per source point and an array of Geospace GCXI autonomous nodal geophone receivers with 50.3 m spacing, as allowed by the terrain. The approximately 6.2 km² of 3D data were collected following a grid-pattern with inlines oriented north–south and binned at 25.1 m spacing. The acquisition sampling interval was 0.5 ms with a recording window of 2 s two-way traveltime (TWT), and coherent reflections discernible down to approximately 800 ms TWT.

Seismic reflection data processing included velocity analysis, NMO correction, statics, spectral balancing,

deconvolution, and poststack time migration. The resulting processed data volume has a broadband frequency spectrum ranging from 15 to 120 Hz. Dip-steered median filtering (DSMF) was performed as a data preconditioning step prior to calculating seismic attributes and extracting data for statistical evaluation. All of the interpretation processing discussed, and the associated computation times, are based on a Mac Pro 6.1 workstation with 64 GB of memory and a 3.7 GHz quad-core processor.

Geophysical detection of karst systems

Characterization of karst aquifers and reservoirs is considered especially challenging in the fields of geophysics, hydrogeology, petroleum geology, and engineering due to their highly variable and heterogeneous vuggy porosity. Vuggy porosity in carbonate rocks is visible pore space that is within or significantly larger than the grains or crystals and therefore not interparticle (Lucia, 1983). Vuggy pore space ranges in size from microscopic to cavernous. Karst systems develop through multiple mechanisms and are controlled by site-specific conditions. The source, chemistry, and flow direction of the corrosive fluids, as well as the timing of rock dissolution, subaerial exposure history, and preexisting fluid pathways such as faults and fractures are all considered major controls on the resulting character of the vuggy pore space and fluid-flow dynamics of karst systems (Hardage et al., 1996; Loucks, 1999). Regardless of the exact mechanism of formation or preexisting state, the fundamental characteristics that define karst systems as unique geologic features should theoretically also make them distinguishable in seismic reflection data. The NN multiattribute workflow method allows the interpreter to determine the approximate boundaries of a particular geologic feature or reservoir property in seismic reflection data based on particular changes in the seismic response (Aminzadeh and de Groot, 2005). Considering that karst systems are defined by fluid flow conditions and physical characteristics that are inherently different than the surrounding native formation, it follows naturally that the careful implementation of an NN multiattribute workflow, composed of the appropriate seismic attributes and training data, should be effective for mapping karst geology in high-resolution, 3D seismic reflection data.

Some investigators have suggested that although seismic reflection data are useful for identifying structural discontinuities, they are insufficient for targeting the other physical properties associated with karst geology, thus requiring the integration of other geophysical methods such as electrical resistivity (Šumanovac and Weisser, 2001; Chalikakis et al., 2011). However, seismic reflection data have the potential to reveal much more than geologic structure. Having seismic reflection data in three dimensions allows for a range of geometric seismic attributes, whereas the recorded phase, frequency spectra, and relative amplitude data permit calculation of physical attributes, such as signal, noise, and

energy (Taner, 2001). These seismic attributes can then be evaluated individually or in combination with others to locate geologic targets of interest (Hampson et al., 2001; Aminzadeh and de Groot, 2005). By means of NNs, an implicit nonlinear transformation encompassing numerous individual seismic attributes is capable of producing a single, feature-specific probability volume. To produce a karst probability volume using this method, the NN should be composed of input seismic attributes known to target physical properties and geometries that are considered characteristic of karst systems.

Geophysical noise and frequency

The heterogeneity of karst systems resulting from enhanced porosity, collapsed rock, and variable satura-

tion conditions often results in significant levels of geophysical noise in seismic reflection data (Šumanovac and Weisser, 2001). In addition, seismic frequency attenuation can be related to the microstructural characteristics of the material and its pore fill (Carcione and Picotti, 2006). Therefore, zones with karst properties would likely be accompanied by attenuation of higher frequencies in the seismic signal over very short distances. Typically, high noise levels and a significant loss of high-frequency content would seem to present serious challenges. However, in practice, the noise and frequency content characteristics can be extracted as seismic attributes and used for interpretation. Significant contrasts observed in those seismic attributes could be indicative of a change in physical properties between volumes of rock that have karst features and rock without karst features.

Geologic structure

Characteristic speleogenic features and landforms known to develop as a result of karstification should also play a significant role in the identification of these zones. Schematic models of karst systems indicate the presence of an interconnected system of subterranean vertical and lateral dissolution enlarged fluid-flow passageways (Figure 2). Regardless of whether large conduits develop or simply narrow zones of enhanced vuggy porosity and permeability promoting fluid flow, localized vertical displacements of carbonate rock are commonly associated with karst (Loucks, 1999), and where these vertical displacements have a columnar 3D geometry in seismic reflection data, they have been identified as seismic sag structures beneath the southeastern Florida Platform (Cunningham and Walker, 2009; Cunningham et al., 2018).

In 2D seismic reflection profiles, these columnar seismic sag structures are indicative of karstification that has led to suprastratal deformation (Figure 3a). Such features appear in profile seismic data as concave-upward arrangements of generally parallel reflections, are attributable to real-world features, and in many cases are supported by the presence of onlap, fill, and/or erosional seismic reflection terminations (Cunningham and Walker, 2009).

Although these karst collapse structures appear as vertical features when only 2D data are available, they generally exhibit a circular or elliptical nature (Hardage et al., 1996) when viewing the seismic data as a Z-slice or horizon in 3D data (Figure 3b). Sullivan et al. (2006) demonstrate that using seismic attributes such as coherence to bring out these features in the Z-slice was highly effective. Note that the red ellipses signifying low-similarity (low-coherence) values on the Z-slice shown in Figure 3b demonstrate excellent spatial agreement with the columnar seismic sag structures observed in

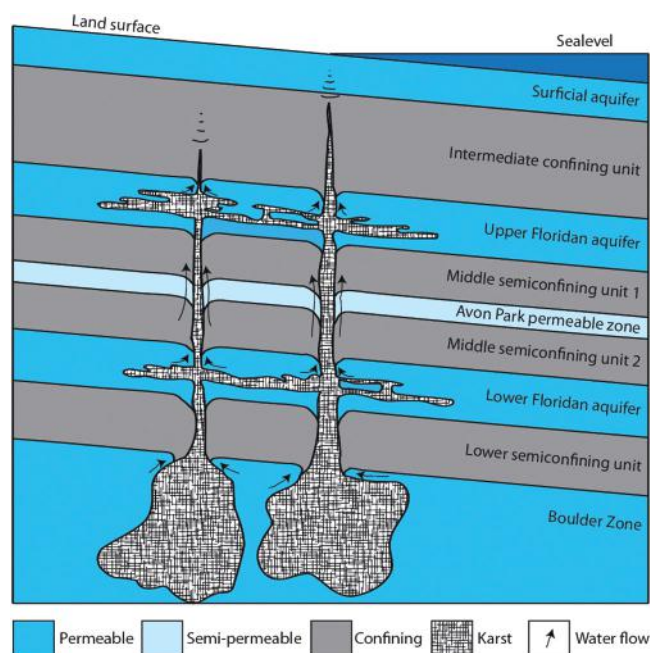


Figure 2. Schematic diagram of a hypogenic karst system (modified from Spechler, 2001).

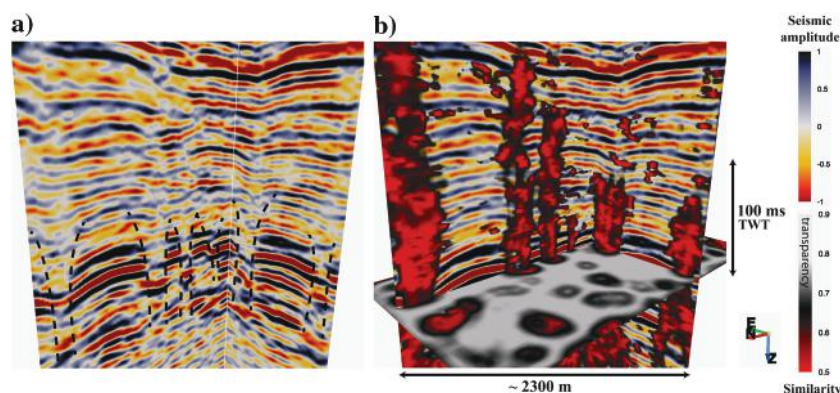


Figure 3. (a) Inline 98 and crossline 70 amplitude profiles with suspected vertical columnar seismic sag structures designated with dotted lines, (b) the same 2D profiles and a Z-slice at 615 ms TWT, each draped with a minimum-similarity (low-coherence) attribute shown in red/black with partial transparency.

2D amplitude data. The incorporation of characteristic spatial geometries that are known to be associated with these columnar karst systems should significantly contribute to their effective identification through a NN multiattribute workflow approach.

Seismic texture

Apparent textural differences observed in 2D seismic amplitude profiles could also help to identify potential zones of karst. Likely indicating common depositional histories or reservoir properties, seismic texture often serves as the basis for seismic facies classification and aids the interpreter by defining the local geometry of events. The method of converting seismic amplitude data into gray-level co-occurrence matrices (GLCMs), which are then transformed into a suite of seismic texture attributes through statistical analyses, is thoroughly described in the literature (Chopra and Alexeev, 2006; Gao, 2011; Eichkitz et al., 2015). Attributes derived from seismic texture using a GLCM are essentially measures of how frequently particular arrangements of pixel brightness values occur in a moving window. GLCM-based attributes effectively provide the interpreter access to a type of second-order statistical measure of seismic amplitude configuration (Gao, 2011) that is calculated based on spatial pixel arrangement rather than the seismic signal itself. Theoretically, such second-order textural attributes may be able to contribute quasi-independent information capable of bolstering the NN KMW. Some of the most effective statistical measures used for characterizing seismic texture are energy, entropy, contrast, homogeneity, and randomness (Chopra and Marfurt, 2005; Gao, 2011). Performing these calculations in specified directions can often enhance geologic features, and applying different pixel window sizes controls spatial resolution (Eichkitz et al., 2015). In the present study, 23 different texture attributes (statistical measures of the GLCMs) were calculated for the 16 principal spatial directions. Although there is known to be overlap among some of these texture attributes, the advanced data analysis techniques applied in this study effectively remove overly correlated attributes during dimensionality reduction.

Applying neural networks to seismic reflection data interpretation

Although machine learning in the form of NNs has become ubiquitous in the developed world (e.g., the development of self-driving cars), the geosciences have been relatively slow to adopt this technology. Some investigators have criticized NNs as a “black box” approach that is difficult to defend because the mathematics behind the decision-making processes is obscured (Aminzadeh and de Groot, 2005; Brouwer et al., 2011; Ma and Gomez, 2015). However, the rapid technological advancements made possible through the implementation of NNs in such wide a range of industries should be more than sufficient evidence for the efficacy and usefulness of these tools.

Reflection seismology is one discipline in the geosciences that has recently taken great interest in the application of NNs. Using the multitude of seismic attributes that can be extracted from 3D seismic reflection data, machine learning methods are able to perform hyperdimensional pattern recognition to identify the presence of certain geologic features. One of the more common applications of NNs in seismic interpretation is the targeting of fluid-migration pathways, which are often associated with hydrocarbon reservoirs (Meldahl et al., 1999; Ligtenberg, 2005; Kluesner and Brothers, 2016) and in some cases with aquifers (Cunningham et al., 2018). The NN multiattribute workflow developed for this purpose is referred to as the chimney cube (Meldahl et al., 1999).

There are several meta-attributes presently available, combining interpreter knowledge with multilayer perceptron NNs to recognize complex patterns across numerous seismic attributes to identify specific geologic features such as faults, fractures, chimneys, slumps, and salt bodies (Aminzadeh and de Groot, 2005). The underlying multiattribute workflows are specifically set up for the physical properties and structural geometries known to be associated with each geologic feature of interest. The most relevant physical characteristics associated with karst systems were introduced in the previous section, whereas Figure 4 demonstrates how the size, direction, and shape of attribute extraction windows should be strategically designed to target specific geologic features.

Typically, vertical structures such as chimneys will be most conspicuous when utilizing vertically stacked windows (Figure 4a), dipping faults are more likely to be identified when stepping-out diagonally (Figure 4b), and targets such as direct hydrocarbon indicators (bright spots) respond the most to horizontally adjacent extraction windows (Figure 4c) (Meldahl et al., 2001). In addition, dip-steering calculations can be used to guide the data extraction, such that each window will be centered on the reflection of interest and follow it even when dipping (Figure 4d). Because columnar karst-collapse structures are characterized by a unique combination of vertical and horizontal features, it is reasonable that a custom NN multiattribute workflow could be developed with a combination of vertical and horizontal step-outs that enhance the targeting of columnar seismic sag geometries, hence, the impetus for developing the KMW. It should be noted that the general procedure discussed here for developing this seismic object detection tool is not feature- nor site-specific; however, the final NN configuration and resulting model presented here are case-specific because the design is dependent on the data examples provided during NN training.

Methods for selecting input seismic attributes

Although the default meta-attributes (e.g., fault cube, chimney cube, and salt cube) available to us are effective tools for well-defined seismic object detection tasks, the

actual procedure for selecting the appropriate input seismic attributes that compose these tools needs to be developed and made more transparent. Although knowledge being gained through trial and error is a reasonable strategy (de Groot et al., 2001), other common practices for attribute selection can be summarized as follows:

- Use common sense and geologic knowledge to systematically choose attributes (as well as extraction windows) that will likely provide good contrast for the seismic object of interest. Some important characteristics to focus on are the size, shape, orientation/ direction, seismic texture, and induced seismic response changes (Meldahl et al., 2000; de Groot et al., 2001).
- Choose some 2D profiles that are believed to express good examples of the seismic object of interest, and use them to extract and evaluate attributes individually. Those attributes that highlight the desired feature well in 2D profiles should be used, whereas those with poor spatial agreement with the suspected seismic object may be discarded (Brouwer et al., 2011).
- Use a statistical support tool to assist with dimensionality reduction, such as PCA, Fisher's linear

discriminant analysis, or similar variants (Aminzadeh et al., 2000; de Groot et al., 2001).

- Finally, the simplest suggested method for evaluating attribute importance and potentially reducing the number of input attributes is to test run a supervised NN (using appropriate picksets) and make note of the weightings associated with each individual attribute (de Groot et al., 2004). Higher weightings indicate greater importance to the pattern recognition solution, so attributes with lower weightings would be eliminated first. The colored circles next to each attribute in the example NN training window in Figure 5 represent their relative weightings, with red circles indicating high importance and light-yellow circles indicating low importance.

The series of images in Figure 6 demonstrate the second procedure described above, in which individual attributes are evaluated visually. The geometric and physical types of seismic attributes should be overlain on amplitude profiles one at a time to determine if they effectively target the suspected seismic feature. The geometric attribute used in Figure 6c is similarity, which has been rendered transparent for values greater

than 0.7. By displaying only low values of similarity, the attribute as shown is effectively minimum similarity, signifying areas of poor seismic reflection continuity. The physical seismic attribute shown in Figure 6d is signal-to-noise ratio (S/N), and it has the same transparency gradient applied. The S/N is a useful indicator of seismic signal scattering and dampening. For example, heterogeneous lithologies with variable pore fill conditions have an abundance of potential diffraction points and signal-attenuating pore fill that can cause seismic energy to propagate chaotically, whereas in situ, well-stratified sedimentary beds transmit the seismic energy through them with much less impact on the S/N (Carcione and Picotti, 2006; Chalikhakis et al., 2011).

Note how the minimum similarity and S/N attributes in Figure 6c and 6d exhibit an excellent correlation with the seismic amplitude discontinuities and textural changes observed in the 2D profile that an interpreter would likely identify as being karst-related (Figure 6a). Figure 6b is included to show some of the picks (karst picks are represented by blue circles, and nonkarst picks are represented by yellow circles) that were chosen on the basis of these individual attributes and ultimately used for NN training.

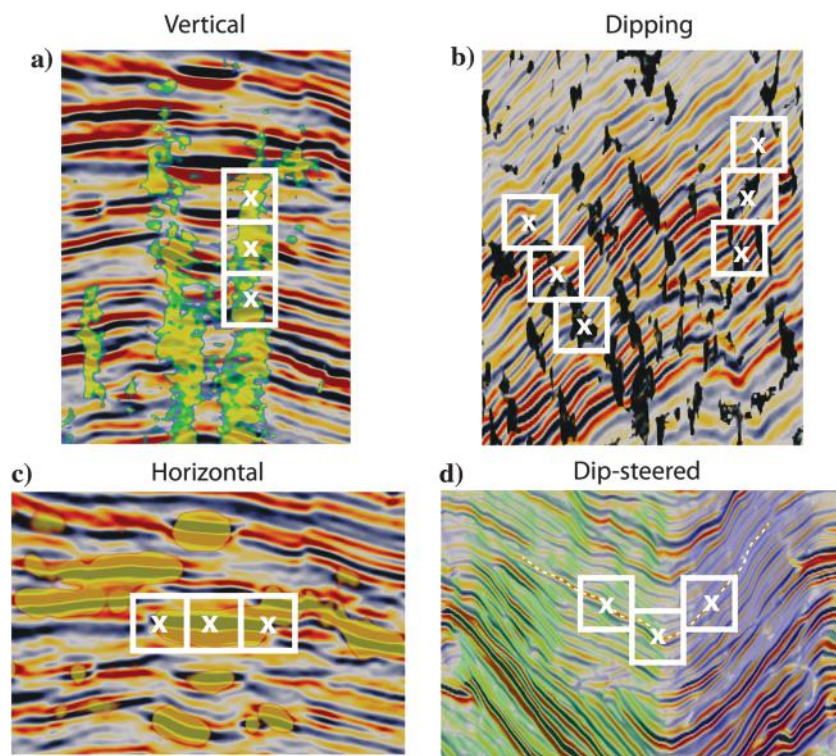


Figure 4. The size, direction, and shape of seismic attribute extraction windows are chosen specifically for the geologic feature of interest. (a) Vertically oriented data extraction and step-outs are best suited for targeting features such as seismic chimneys, (b) diagonal window step-outs are better for capturing faults, (c) horizontal data extraction windows can be used to identify seismic bright spots, which are often indicative of hydrocarbons, and (d) dip-steering can be used with a variety of geometric seismic attributes so that step-outs follow along seismic reflections rather than simple XYZ space. Seismic data examples are from Kluesner et al. (2013).

Some investigators may leave the procedure for selecting appropriate input seismic attributes as ambiguous due to the fact that NNs are highly efficient at returning a desirable result by selectively focusing on well-supported relationships, even in the presence of spurious correlations that may be found when excessive attributes are included. So in theory, given sufficiently large training data sets, all seismic attributes could be input to the NN and it would still provide a satisfactory result (Aminzadeh et al., 2000; Brouwer et al., 2011). However, this does not mean that the specific set of attributes is insignificant. Because the supervised NN is trained on example data meant to differentiate between karst and nonkarst, the meaning of the result (karst or nonkarst) will not be affected by the specific set of input attributes, only the quality will (Brouwer et al., 2011).

The possibility does exist that superfluous attributes could introduce random correlations, which may distract

the machine learning process from the more relevant ones (Kalkomey, 1997; Brouwer et al., 2011; Ma and Gomez, 2015). However, by providing large sets of example data for training, the potential for this occurring is drastically reduced (Hush and Horne, 1993; Brouwer et al., 2011). One practical concern that makes using very large sets of input attributes prohibitive is computer processing time. The machine learning process will spend time exploring the irrelevant correlations that exist among the myriad of input variables rather than focusing on the significant correlations that reflect actual geologic relationships (Aminzadeh et al., 2000; de Groot et al., 2001).

Evaluation of statistically based dimensionality reduction techniques

The present study leverages statistical techniques to ensure a data-driven, optimized method for determining input attribute sets. Although previous investigators have suggested and implemented statistical methods

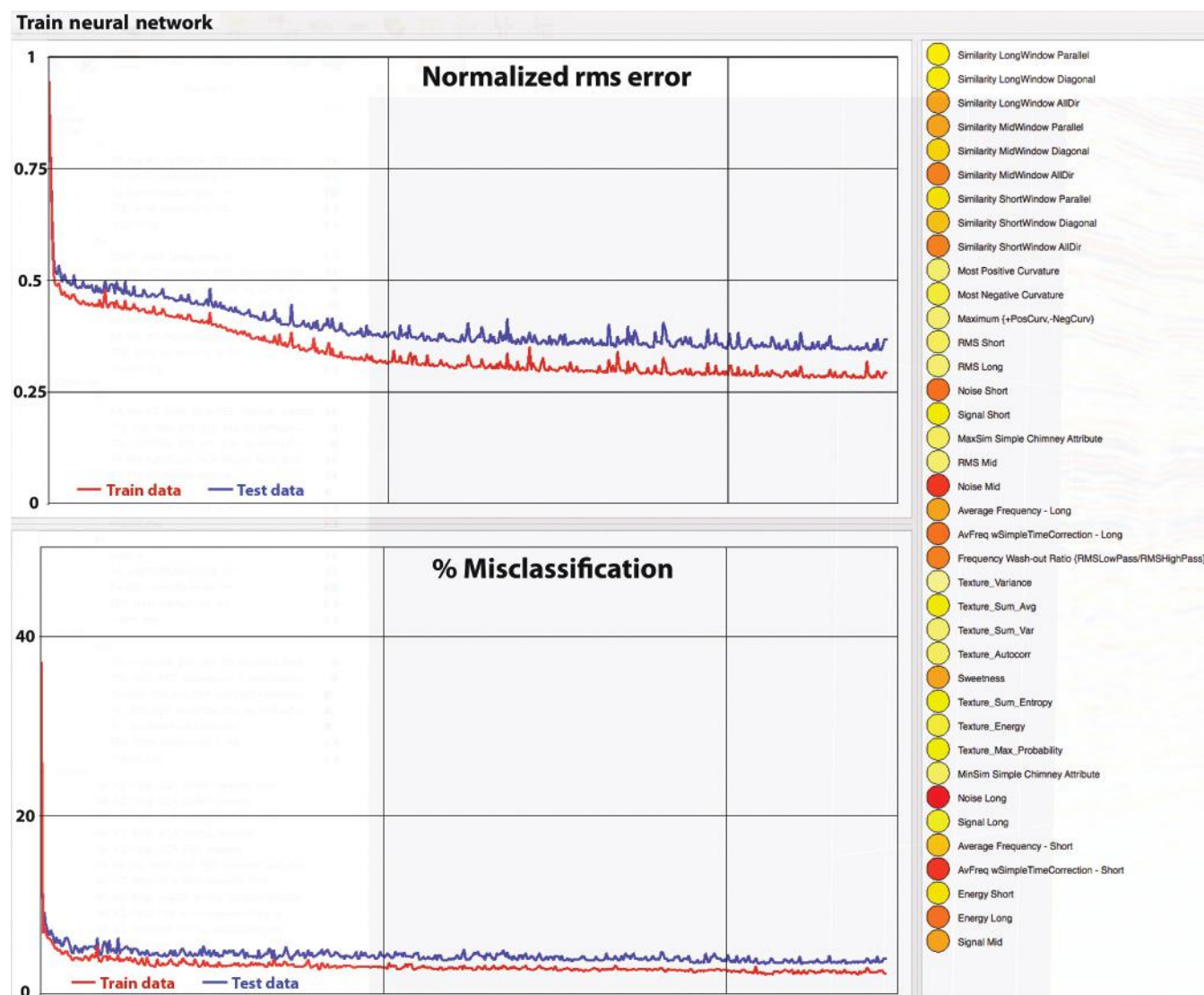


Figure 5. The NN training window displaying rms error, percent miscalculation, and the relative weightings of each attribute on the final nonlinear transformation function (red circle = high weight, light yellow/white circle = low weight).

for the purpose of seismic attribute dimensionality reduction (Aminzadeh et al., 2000; Hampson et al., 2001; Taner, 2001; Hashemi et al., 2008), this study applies a suite of statistical tools to the extracted data to present qualitative and quantitative comparisons of their performance relative to that of the initial full set of attributes.

When discussing advanced statistical analysis, the critical distinction between parametric and nonparametric statistical testing must be addressed. Parametric testing assumes that the data set being evaluated has a normal distribution, homogeneity of variances, and exhibits linear relationships (Conover and Iman, 1981). Although some of the seismic attributes being evaluated do meet these criteria, the majority do not. Therefore, nonparametric statistical testing is used when possible.

Correlation

For the purposes of dimensionality reduction, it is appropriate to remove attributes that have either near-perfect correlation with other attributes or poor correlation

with the training attribute because these superfluous variables will increase the computation time without contributing additional information. Pearson's R is a commonly used parametric test for evaluating the correlation of two variables. Its nonparametric equivalent, Spearman's rank correlation coefficient (Spearman's rho; Spearman, 1987), was applied to the data such that any attributes with statistically nonsignificant outcomes ($p > 0.05$) are eliminated, as are those with low correlation coefficients between the input and output variables (< 0.4).

Following Spearman's rho, the nonparametric variant of the T -test known as Welch's T -test (Zimmerman and Zumbo, 1993) was performed. This procedure evaluates the equality of means between two sample sets, which in this case are the karst (1) and nonkarst (0) data sets. If there is no statistically significant difference ($p > 0.05$) between the means of the two sets for a particular attribute, then it is considered a poor discriminant for identifying karst and is removed.

PCA

PCA is a common eigenvector-based multivariate analysis tool that uses orthogonal transformations to reduce a large set of variables to a smaller set that still describes the majority of the data variance of the original set (Anderson, 1984; Roden et al., 2015). This dimensionality reduction was performed by identifying the attributes with the greatest contribution to observed data variance, grouping them into components, and progressively removing the lowest scoring component groups. The rotated component matrix in Figure 7 shows the six components that were preserved during PCA, and how attributes describing similar characteristics are naturally grouped together during the orthogonal transformation process. The input seismic attributes shown in the left column of Figure 7 were assigned generic names (VAR#####) during PCA to help prevent any potential user bias. The scree plot shown in Figure 7b demonstrates how influential the first few components are relative to the higher numbered components. There are numerous philosophies regarding the cutoff point or "stopping rule" for PCA, but much of the reason for ambiguity lies in the fact that, once again, sample size and data distribution are highly influential (Peres-Neto et al., 2005). The decision to stop at six components with a total of 32 seismic attributes for this study was based on achieving greater than 95% cumulative variance explained, 0% nonredundant residuals between the observed and reproduced correlations, and the significant decrease in influence of subsequent components (as shown with the scree plot of Figure 7b). Although clearly effective at reducing dimensionality for the purposes of the study, PCA operates on the assumption of linear relationships and therefore may not be perfectly suited for evaluating many seismic attributes and their relationships because they are nonlinear (Aminzadeh and de Groot, 2005).

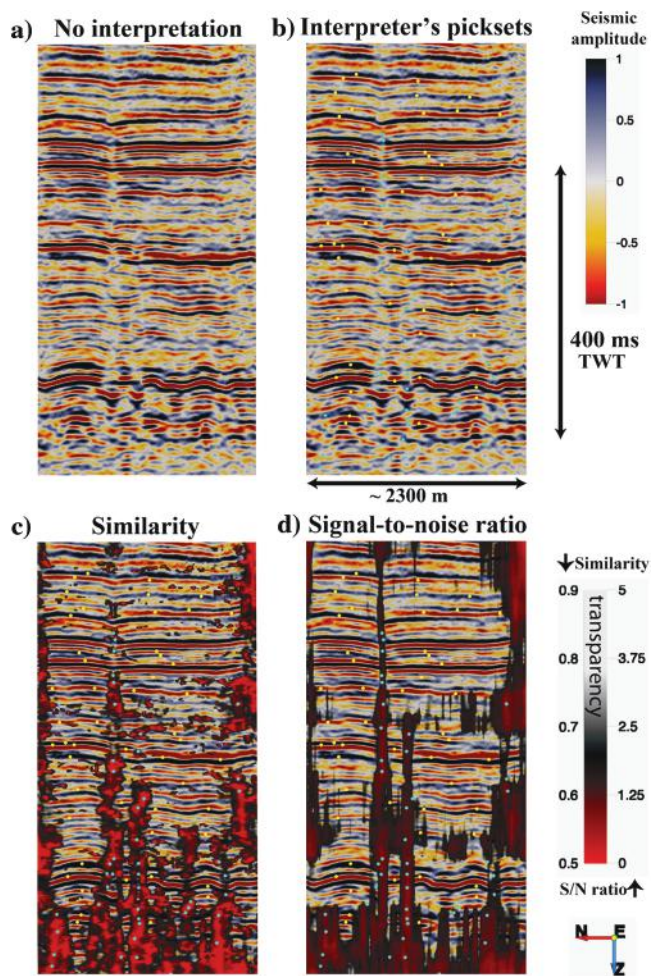


Figure 6. (a) Inline 97 seismic amplitude data, (b) with karst (blue spheres) and nonkarst (yellow spheres) picks shown, with (c) minimum similarity (geometric) and (d) signal/noise (physical) attributes demonstrating relatively good agreement.

NMDS

Given the expectation of nonlinear relationships, a similar nonparametric ordination method called NMDS may be preferable. NMDS is able to overcome the linearity assumption by iteratively arranging points to maximize rank-order correlation between real-world distance and ordination space distance (Shepard, 1980). Each variable is translated into an n -dimensional space, where between-variable distances are preserved, optimized, and scaled. These transformations are determined by minimizing loss functions known as stress and strain (Borg and Groenen, 2005). The assigned n -space coordinates can then be viewed as 2D, 3D, or even higher dimensional scatterplots such that similarly contributing variables will be represented spatially nearby one another, whereas very different variables will appear far apart (Shepard, 1980; Borg and Groenen, 2005). Much like PCA, the goal is to remove those variables whose contribution is nearly identical while maximizing the amount of data variance observed.

QDA

With discriminant analysis techniques, it is possible to directly evaluate how predictive each of the input attributes is of a predetermined or desired output such as karst (1) or nonkarst (0). Naturally, those attributes having little or no correlation with the binary output field of 1 or 0 are likely poor determinants for the identification of karst in seismic data. There are multiple variants of discriminant analysis, but when working

with suitably large sample sets having nonlinear relationships, QDA is preferable (Friedman, 1989). By evaluating the covariance matrices separately, it is possible to overcome the inequality of variances that would normally prohibit usage with nonlinear relationships. When each input attribute was evaluated in a step-wise manner for whether or not it is effective at discriminating between karst (1) and nonkarst (0), as contributing parts of the overall group of attributes, the process either keeps or rejects the attribute based on a minimization of the Wilks' lambda value, a type of measurement of differences in means among combinations of variables for separating cases into groups (Klecka, 1980). This process determined an optimal grouping of 45 attributes (out of 132) that was able to make the correct classification for 99.2% of the cases (picks). Multivariate analysis of variance and multiple (logistic) regression statistical tests were also evaluated and determined to be similar in practice, but QDA was given preference due to its greater robustness to nonnormality and nonlinearity (Klecka, 1980).

Considering that QDA deals directly with discrimination between classes (karst and nonkarst) and can be applied to nonlinear relationships, one might expect QDA to always outperform PCA; however, this may not necessarily be the case. The training sample set size is critical here because experiments have shown PCA to be superior in cases in which the number of training samples is small (Martínez and Kak, 2001). Similarly, Friedman (1989) introduces the regularized discrimi-

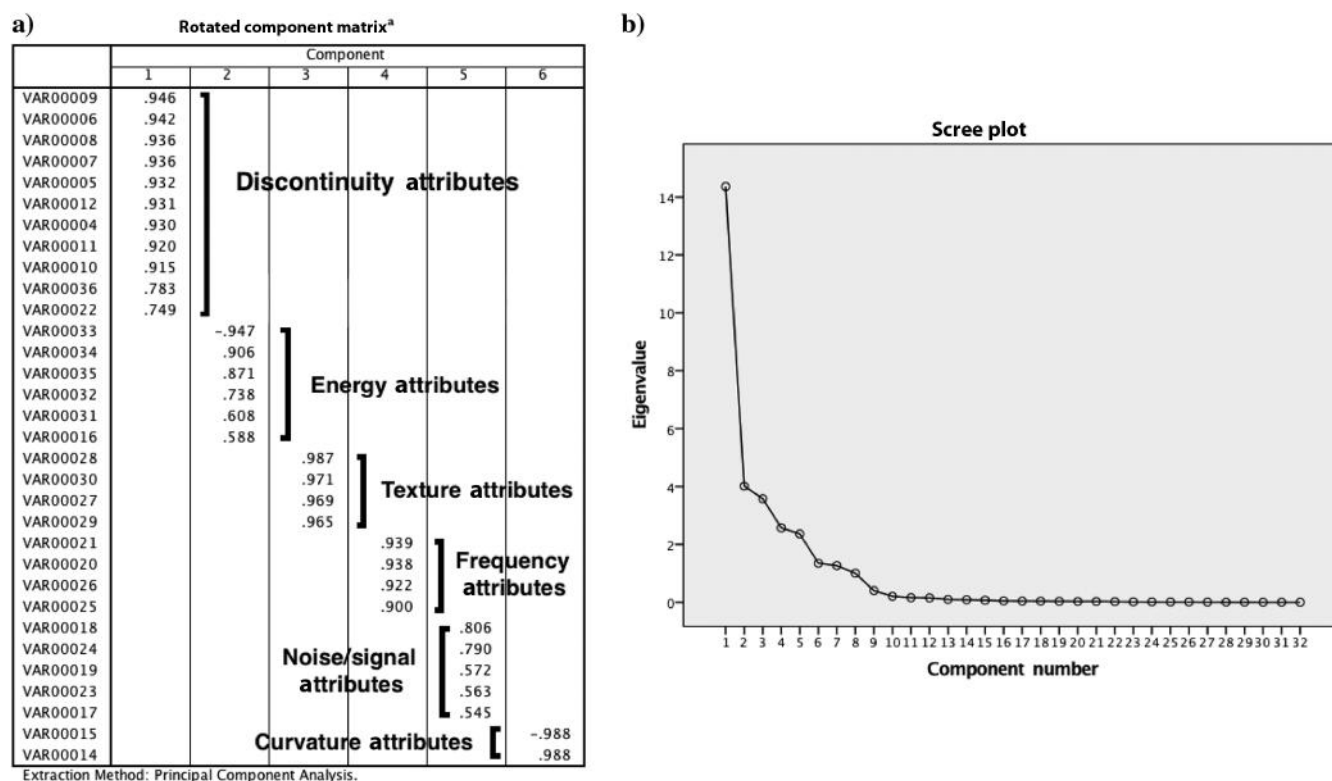


Figure 7. (a) Rotated component matrix window from PCA demonstrating the natural grouping of similar attributes as primary components and (b) scree plot showing decreasing eigenvalues associated with each component.

nant analysis (RDA) variant to correct for the exaggerated importance of low-variance classes that may occur. Although this may improve classification accuracy in some specific instances, there are many situations in which small sample sizes or equal covariance matrices can cause RDA to perform more poorly than linear discriminant analysis (LDA) or QDA (Friedman, 1989).

Gamma test

The concept of the gamma test was initially suggested in Stefánsson et al. (1997) as a nonparametric, nonlinear modeling tool. Given input and output data sets, the gamma test effectively estimates the extent to which the nonlinear model built from the input data is able to fit the provided output data, with lower gamma values indicating better agreement. This novel ability seemingly makes the gamma test an ideal approach for determining the most appropriate seismic attributes to include for NN training. Combinations of attributes with high gamma values indicate a poor fit of the inputs to the output. Conversely, low gamma values suggest that the combination of inputs is able to effectively approximate the output function with minimal errors (Iturrarán-Viveros, 2012). Gamma testing is not ideally suited for problems involving categorical output such as karst versus nonkarst (represented as 1 or 0) because a two-digit range in effective measurement is preferable, but the results should still be meaningful as long as the gamma value returned is very small (Jones, 2002; Iturrarán-Viveros and Parra, 2014).

Given the versatility and novelty of the gamma test, it was included in the evaluation of potential dimensionality reduction techniques with the understanding that spurious gamma values may result from the binary output field. In practice, SRGT is fast and effective at calculating the gamma value associated with an established set of inputs and an output; however, for dimensionality reduction, the genetic algorithm (GA) tasked with evaluating the countless possible combinations of inputs quickly becomes time-prohibitive for large sets. Although rudimentary heuristics help the GA find and evaluate the probable best candidates first (Jones, 2002), the amount of time that would be required to calculate a gamma value for each possible combination of attributes from a large initial list make it impractical for the purposes of optimized dimensionality reduction. For example, an initial list of 132 attributes would require the GA to calculate and compare each of the 2^{132} (or $5.4 \text{ E}+39$) possible combinations to determine the input set with the lowest possible associated gamma value.

It should be noted that many of these statistical tests are highly sensitive to data outliers. Therefore, careful data conditioning must be performed prior to testing, including removal of null and spuriously large values caused by step-outs beyond the survey boundaries. The removal of these significantly reduces the associated normalized root-mean-square (rms, the arithmetic mean of the squares of a set of numbers) error and percent misclassification values.

To measure the predictive performance of the NN models, a cross-validation technique (Stone, 1974) using random subsampling into two sets can be performed. During this process, all available data are randomly partitioned such that 70% of the data are used for training the model and the remaining 30% are used to calculate the error rate of that trained model. Numerous iterations (10 is common) of this prediction error estimate are performed and then averaged to give the true error estimate (Aminzadeh et al., 2000). Running multiple iterations and averaging the results ensures that the reported accuracy and error measurements are truly representative, and not one-time outliers caused by a spurious partition. Although subsampling into three sets (training, test, and validation) is often performed to ensure that overtraining does not take place, the NN training approach used here and shown in Figure 5 prevents overtraining by carefully observing the curve representing normalized rms error of the test data and stopping the process when the curve begins to diverge upward (overtrain).

Karst multiattribute workflow

An objective approach to seismic object detection, specifically targeting karst systems, has been developed in which statistical methods are applied to minimize potential interpreter bias. The workflow described below is suggested as a standardized model for NN-based 3D seismic reflection multiattribute interpretation:

- 1) Seismic reflection data preconditioning is recommended prior to applying the interpretation scheme. For the examples shown in this study, DSMF was applied to a poststack time-migrated data set.
- 2) Using knowledge of geologic properties and previous studies as a guide (Carcione and Picotti, 2006; Chalikakis et al., 2011; Roden et al., 2015), produce a list of seismic attributes that should theoretically be capable of detecting karst systems. This list should include geometric attributes (e.g., discontinuity, curvature, and dip) and physical attributes (e.g., instantaneous attributes, frequency, and rms energy). The number of attributes available to the user is virtually unlimited because time windows, lateral step-outs, and other parameters can be customized for each individual attribute depending on the scale of the feature and the nature of the seismic data being interpreted (Figure 4). For this study, an initial list of 132 individual seismic attributes was used as a starting point from which to reduce dimensionality.
- 3) Select several of those attributes and individually overlay them onto 2D amplitude profiles to evaluate the spatial agreement between the extracted attribute and what the interpreter believes to be karst. After one or more attributes have been identified as being effective for defining karst boundaries, these should be used (with transparency applied) as a helpful guide for creating two picksets — one com-

posed of locations where karst is highly suspected and the other with locations that are almost certainly not associated with karst (Figure 6).

- 4) Extract the data for the full list of attributes at the locations represented by these two new picksets, using either moving windows or finite samples depending on the seismic attribute. These data are used as the input for the interpreter's preferred statistical technique for dimensionality reduction. In this analysis, each of the statistical techniques described in the previous section were evaluated individually for accuracy and efficiency. The performance of each option as applied to the present study is summarized in Table 1.
- 5) Use the list of statistically selected attributes as the inputs for a NN, along with the full picksets created in step (3). The NN will produce a nonlinear transformation that incorporates each of the input attributes with iteratively determined contributions (weights) to the final solution through a back-propagation algorithm (Brouwer et al., 2011).
- 6) This NN multiattribute transformation is then applied to the full 3D data volume to produce probabilities ranging from 0 to 1, with discrete locations assigned values near or at 1 being confidently classified as karst and values near 0 being very unlikely to be associated with karst.

Appropriate pickset size for neural network training

As long as sufficiently large sets of representative example data are given for NN training (Hush and Horne, 1993; Brouwer et al., 2011), the resulting karst probability data volume plausibly maps the approximate spatial extent of karst systems as well as or better than a human interpreter, and in much less time. In general, a larger number of example data samples will do a better job of representing the underlying problem, and therefore lead to the NN learning a better solution, but at some point, there will be a law of diminishing returns in which the amount of time and effort committed to picking good locations is no longer justified (Hush

and Horne, 1993). One commonly referenced method for determining the minimum number of training samples required for good generalization is the Vapnik-Chervonenkis dimension (VCdim) (Hush and Horne, 1993). This value is determined based on the number of hidden layers in the NN, the number of nodes in those layers, and the nature of their connections. For a multi-layer perceptron NN with one hidden layer, such as the one being used for the present application, the lower bound of the VCdim is roughly equal to the number of connections between the inputs and the hidden layer, and the upper bound is approximately twice the number of weights in the full network. Thus, determination of the VCdim is dependent on specific knowledge of the NN architecture and makes several assumptions, including fully connected layers and nodes that use hard-limiting nonlinearities (Hush and Horne, 1993). Furthermore, results from Sontag (1989) indicate that when a multilayer perceptron applies the commonly used sigmoid activation function, the VCdim is about twice as large.

A simple alternative solution, which avoids calculation of VCdim, is to use the *M*-test. This procedure computes gamma values for an increasing number of provided data samples. The point at which the gamma statistic asymptotes are the minimum number of points required to build a model capable of effectively predicting the output. Because the data points are introduced from the full data set at random, multiple iterations of the *M*-test are run and the resulting values are averaged. Although the mathematical implementation of the gamma test and subsequent *M*-test are well beyond the scope of this study, the validity of the practice and underlying theory are well-documented elsewhere (Stefánsson et al., 1997; Jones, 2002; Iturrarán-Viveros, 2012; Iturrarán-Viveros and Parra, 2014).

Given the significant influence of training sample set size on the outcome of the NN models, an objective method for determining the appropriate number of picks to be used for training should be applied. For this study, application of the *M*-test was given preference over calculation of the VCdim due to the relative uncertainty surrounding the architecture and implementation

Table 1. Quantitative and qualitative results from statistics-based NN models (Rho/T, test of correlation combining Spearman's rho and Welch's *T*-test; PCA, principal component analysis; NMDS, nonmetric multidimensional scaling; QDA, quadratic discriminant analysis; and rms, root-mean-square).

Method	All attributes	Rho/T	PCA	NMDS	QDA	Gamma
Number of attributes	132	76	32	57	45	61
Normalized rms error (train/test)	0.15/0.18	0.16/0.20	0.22/0.26	0.16/0.18	0.16/0.19	0.18/0.20
Percent misclassification (train/test)	0.38/0.59	0.55/1.15	1.62/1.48	0.66/0.69	0.59/0.85	0.76/0.91
Gamma value	-1.31 E-3	-5.90 E-4	1.28 E-4	3.16 E-4	3.78 E-5	6.68 E-4
<i>M</i> -test (minimum number of picks)	600	850	900	650	350	900
Computer processing time	19 h 38 min	5 h 34 min	6 h 37 min	5 h 10 min	4 h 55 min	7 h 37 min
Qualitative spatial result ranking (1-6)	6	4	3	2	1	5

of the NN. According to the M -test, the minimum number of training samples required to achieve good generalization is shown in Table 1 for each of the NN models evaluated.

Optional PCA-based unsupervised neural network method of creating picksets

The workflow described above implements statistical techniques to determine the specific set of input attributes that will optimize the accuracy and computational time required for an NN-based multiattribute interpretation scheme. However, the steps in which picksets are created for NN training remain vulnerable to interpreter bias. Determining which individual attributes act as the most useful guides for creating accurate picksets can be highly subjective. For interpreters who are not confident in making the determination of karst versus nonkarst features for pickset creation, including the following optional workflow steps (in lieu of step 3 above), introduces another layer of statistical dimensionality reduction and unsupervised NN (UNN) training to further enhance interpreter objectivity:

- Create a random pickset from a data volume subset that encompasses a suspected karst feature. This pickset is automatically generated, so the

only practical constraint on the number of picks specified is the time required for data extraction at each point. Depending on the size of the volume subset and the number of seismic attributes to be evaluated, anywhere from several hundred to a few thousand picks should be sufficient without being time prohibitive.

- Perform a PCA on the data extracted from the random pickset. In the present example, each one of the 2500 randomly generated points has 132 unique data fields corresponding to the chosen attributes in the initial list. The PCA process identifies which attributes are associated with the greatest data variance. Ideally, the attributes that are preserved during PCA should have values near one end of the data range for picks coinciding with karst and the values on the opposite side of the data range should be representative of nonkarst locations.
- Some of the PCA-selected attributes may serve as helpful guides for making well-informed karst and nonkarst picksets individually, or preferably, a UNN can be trained using the PCA-selected list of input attributes and the random pickset. Two common types of UNN are SOMs and clustering, neither of which require user guidance (supervision), by definition (Roden et al., 2015). The resulting data cube highlights seismic discontinuities and textural changes that can help the interpreter create representative karst and nonkarst picksets (Figure 8). Note that the PCA-based clustering UNN multiattribute displayed with transparency in Figure 8 was produced without any interpreter input and served as a partial guide for determining the pick locations shown.

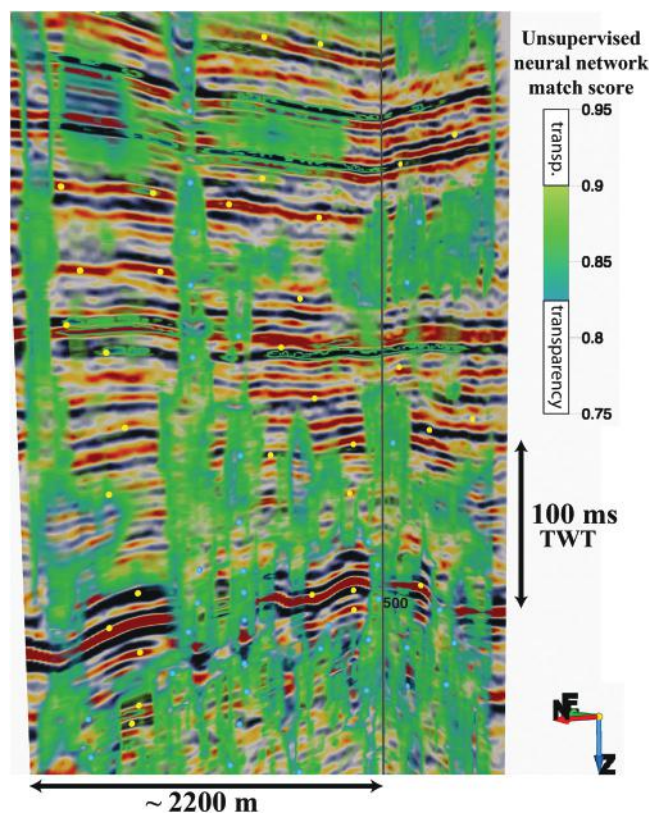


Figure 8. The (clustering) unsupervised NN results based on PCA-chosen attributes, overlain with semitransparency onto inline 97 and crossline 89, and shown with example picksets that could be created based on those UNN results (blue picks = karst; yellow picks = nonkarst).

Results

Using a 3D seismic reflection data set from southeast Florida, we developed an objective NN multiattribute workflow for mapping karst in which potential interpreter bias is minimized by applying linear and nonlinear data transformations for dimensionality reduction. PCA is a linear approach that yields a reduced set of input seismic attributes to the NN by eliminating irrelevant and overly correlated variables, while still preserving data variance. By initiating the supervised NN from an eigenspace that maximizes the separation between classes, the convergence time and accuracy of the computations are improved because the NN only needs to recognize small perturbations to the provided decision boundaries (Aminzadeh et al., 2000). The preferred nonlinear statistical method of dimensionality reduction determined in this study was step-wise QDA.

Quantitative results

Without available well-log data for groundtruthing, as is the case with this study, it is difficult to definitively identify which statistical approach to dimensionality reduction ultimately leads to the best results. From a

quantitative perspective, the accuracy and efficiency of the NN training process can be assessed in three ways: normalized rms error, percent misclassification, and total computation time. In the NN training window (Figure 5), performance can be monitored using the two graphs representing normalized rms and percent misclassification. Two curves are shown in each graph: The red line corresponds to the training data partition (70% of the total picks), and the blue line corresponds to the test data partition (30% of the total picks). Picksets are partitioned randomly into the test and training sets for each NN iteration to monitor for overfitting. During the training process, both curves should move downward and either asymptote at a stable level or begin to diverge, at which point the training should be stopped to ensure that any spurious correlations found in the training set are not allowed to propagate into the final solution.

Normalized rms error is a measure of the difference between the values predicted by a model and the actual values observed. Effectively a measure of model accuracy, rms error values between 0.8 and 0.6 are good, between 0.6 and 0.4 are excellent, and anything below 0.4 is considered perfect (OpendTect, 2016). As shown in Table 1, each of the models evaluated yielded normalized rms values that are significantly less than 0.4. The normalized rms error values associated with each model are very comparable, in general, ranging from 0.15 to 0.22 for the training sets and 0.18–0.26 for the test sets.

The percentage misclassification curves show what percentage of the training and test picksets have been misclassified as karst or nonkarst by the interpreter, relative to what the model predicts. So the percentage misclassification is strictly the number of pick locations where the NN contradicts the interpreter's original assessment, whereas the normalized rms error factors in just how strongly the given classification was refuted (OpendTect, 2016). High percentages of misclassification suggest that either the karst and nonkarst picks were poorly chosen and are not representative of the in situ geology or that the NN model is underperforming, which would also be reflected in the normalized rms error value. The percentage misclassification values for each model are less than 2% and are given in Table 1.

Based on these "perfect" scores for normalized rms error and percentage misclassification (Table 1), each of these NN models is likely to produce a high-quality interpretation of the presumed karst system. However, the results from these models are variable. Because they each incorporate specific attribute sets composed of different amounts and types of individual attributes, all of the models produce unique spatial results and have very different computation time requirements. Given that NNs are known to be highly noise-tolerant and efficient at returning a desirable result even when excessive attributes are included (assuming sufficiently large example training sets), one of the most relevant aspects of this work is the drastic reduction in computation time due to dimensionality reduction. By applying step-wise

QDA to the initial attribute set, an interpretation as good as or better than the 132-attribute model was produced in approximately one-fourth of the time. The importance of this time savings should not be underestimated, considering that it required almost 5 h of computer processing time to produce the preferred model for just the small 6.2 km² survey area. When scaling this procedure up to larger surveys, a 75% reduction in computation time can be substantial.

The complete gamma testing procedure, including GA dimensionality reduction, was applied to the initial list of attributes. The result was a list of 61 individual seismic attributes, many of which were not included in any of the other statistically based models. Considering the comparable normalized rms error and percentage misclassification values that were returned, as well as the low associated gamma value, it appears that the gamma test is effective for this application even with categorical output values. In addition, a gamma value was calculated for each of the models (Table 1). This is a useful determination of accuracy because the gamma value is a measure of how well each specific set of input seismic attributes, in combination, is able to fit the given nonlinear output function. As part of the gamma testing procedure, the *M*-test is performed to ensure that a sufficient number of data samples (picks) are included to reach an optimized solution.

Qualitative spatial agreement

For the purposes of seismic reflection data interpretation, the most important row of values in Table 1 is likely the qualitative spatial result ranking. As the name suggests, this is a purely qualitative assessment regarding which of the NN models yields the best visual accuracy when viewed as overlays on 2D seismic amplitude data and as 3D volume-rendered bodies. Spatial agreement between the resulting karst probability cubes and seismic amplitude data is arguably the most important factor because the aim of this study is to produce a semi-automated, objective approach to interpreting karst systems in 3D seismic reflection data that performs as well or better than a human counterpart. Although each of the six models produced satisfactory results, preference was given to those with (1) spatial limits coinciding precisely with reflection discontinuities, (2) low interstitial noise away from the major karst system, and (3) the fewest crossings of reflections appearing continuous in profile.

The NN probability cubes resulting from each model are shown in Figure 9 as semitransparent, light-blue-colored overlays onto Inline 97 seismic amplitude data. The areas of light blue shown on these profiles are believed to represent areas that have a high probability (>75%) of being karst. Based on the three criteria above, each one of these models is ranked in terms of spatial agreement. The model using all 132 of the seismic attributes evaluated is given the lowest ranking (6) because it appears to be systematically over-inclusive and also overly confident (only very high and very low probabilities given with little gradation) (Figure 9).

The boundaries of this interpreted system slightly overlap reflections rather than terminating at points coincident with the discontinuities. The gamma test and Spearman/Welch models seem to be similarly overestimated spatially and in terms of probability (Figure 9). The three highest ranking models were (1) QDA, (2) NMDS, and (3) PCA, all of which are comparably excellent from a qualitative perspective. Preference was given to the QDA and NMDS models over PCA because they have less interstitial noise away from the core sys-

tem and more geologically realistic boundaries. QDA was ultimately selected as the preferred model over NMDS because it exhibits a slightly more continuous, well-connected structure that appears to be corroborated by the seismic amplitude data.

Although the prevalence and nature of the secondary connecting structures vary among the models and could impact future hydrogeologic modeling, the primary frameworks of the karst system predicted by each NN iteration shown in Figure 9 are highly consistent. This

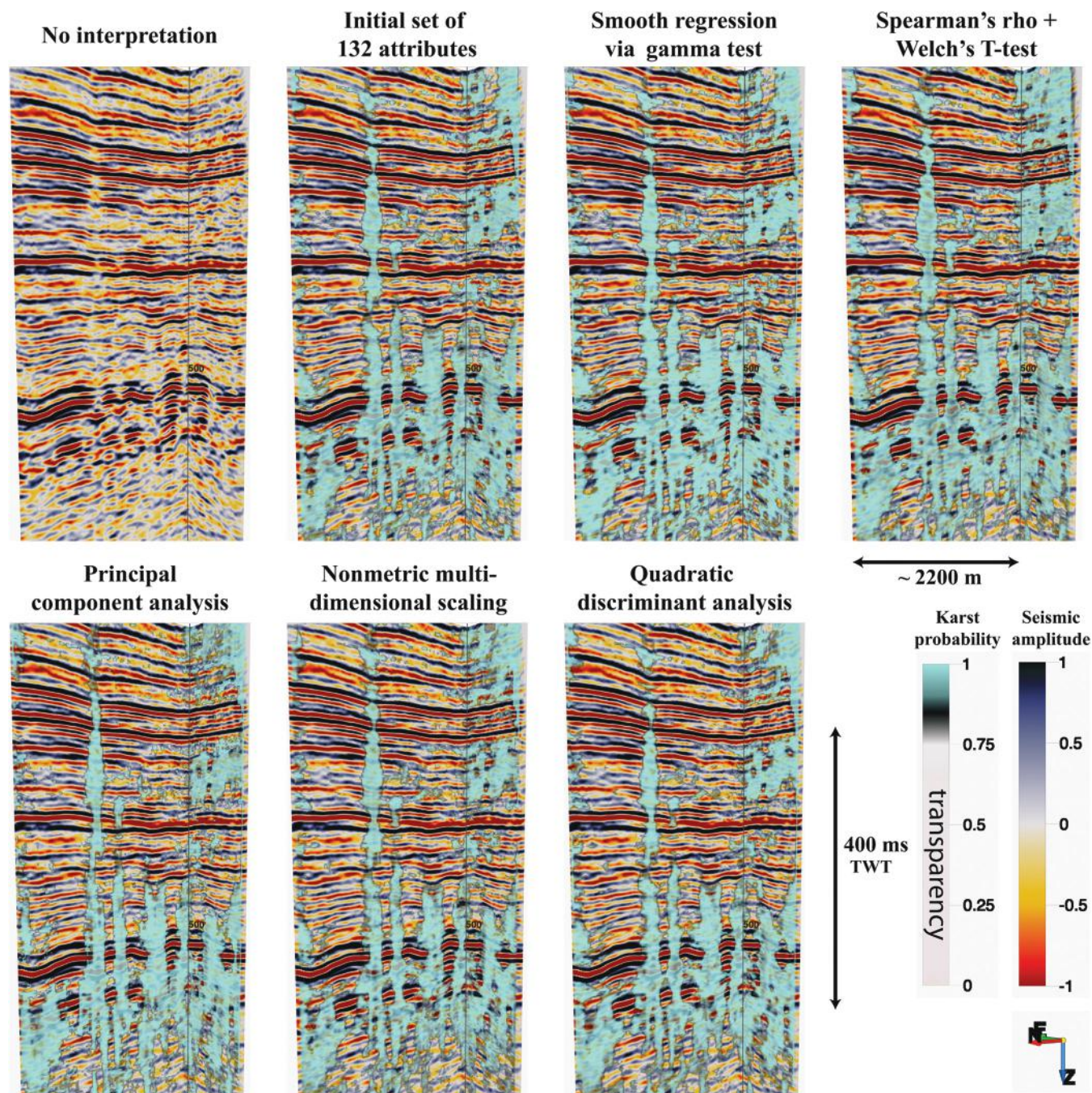


Figure 9. Spatial comparison of the six KMW models (blue overlays with partial transparency) produced using statistical dimensionality reduction techniques shown with inline 95 and crossline 90 amplitude profiles. An uninterpreted view of the same 2D amplitude profiles is shown in the upper left panel.

again supports the previous assertion that, given sufficiently large training data sets, NNs are highly efficient at producing a satisfactory result even with the presence of excessive and possibly unrelated variables.

Figure 10 provides a direct comparison of the results from the KMW (QDA-based) with the image shown in Figure 3b. The karst probability cube is overlain with transparency onto the same 2D profiles and Z-slice (with a minimum similarity attribute) as Figure 3b. This image suggests that the KMW model does in fact identify a high probability of karst coinciding with the elliptical, low-coherence (similarity) Z-slice shapes shown in Figure 3b. Furthermore, the high-probability karst zones predicted by the KMW shown on 2D profile agrees well with the approximate locations predicted in Figure 3, albeit in much more realistic configurations as opposed to idealized smooth, dotted cone shapes.

Comparison with other seismic reflection multiattribute workflows

As further demonstration that the KMW is uniquely suited to target karst specifically, the same two training picksets were applied using NN meta-attributes already available to us, known as the fault cube, advanced fault cube (AFC), chimney cube, slump cube, and salt cube. The resulting 3D probability volumes (or geobodies) were evaluated spatially for agreement with 2D seismic amplitude data (Figure 11), and it was determined that the KMW model incorporating stepwise QDA yielded the best results. The KMW approach seems to effectively balance the fault cube's bias toward targeting horizontal structures with the chimney cube's bias toward targeting vertical structures.

Once again, the fact that each of the (nonkarst-related) NN interpretations shown in Figure 11 identified roughly the same zones and structures as high-probability karst is further evidence that NNs can identify important attribute relationships and return a desirable result, even with the presence of some inappropriate attributes. Because each of the NN models were trained using picksets that distinguish karst from nonkarst, the meaning of the results was not affected by the specific input seismic attribute set, but the quality of those results was (de Groot et al., 2001). So although some of the NN results shown in Figure 11 may appear to be roughly the same as the KMW model, closer inspection reveals significant differences that would be meaningful for hydrogeologic modeling based on such input 3D geobodies.

Comparing the upper section of each image in Figure 11, it is apparent that the KMW and chimney cube models are much more laterally constrained than the others and focus primarily on the columnar seismic sag structures. Other mod-

els, such as the salt and slump cubes, indicate more significant portions of that upper section as being high-probability karst, which contradicts previous interpretations of that section everywhere being a confining unit overlying the Upper Floridan aquifer (Miller, 1986), but it could be locally leaky and semiconfining (Cunningham et al., 2018).

In addition, the cross-sectional radii of the columnar seismic sag structures for the AFC model appear to be larger than the other models, with the KMW exhibiting narrower vertical structures and the best general agreement with the reflection discontinuities observed in 2D amplitude profiles (Figure 11). This increased 3D spatial expression combined with the inclusion of regions that appear to be isolated from the primary karst system can lead to wide ranges in hydrogeologic modeling parameters between these methods. The volumes of the karst system geobodies shown in Figure 11, as well as each statistical iteration of the KMW (Figure 9), were calculated using a 3D geobody volume estimation tool, and they are listed in Table 2 for comparison. The values in Table 2 show that the choice of input attributes provided to the NN can drastically affect the karst system's estimated volume parameter, which will be an influential component of any subsequent hydrogeologic modeling of this complex system. As shown, the volume difference between the different models' results was observed to be as much as an order of magnitude.

The degree to which a karst system is interconnected can be of great importance to developing representative hydrogeologic models based on this approach. Even though many of the interpreted karst systems shown in Figures 9 and 11 appear to be very similar, there are locations throughout the data volume in which some models find connectivity whereas others indicate isolation. For example, the more horizontally constrained system resulting from the preferred QDA-KMW model exhibits less high-probability karst in general (white

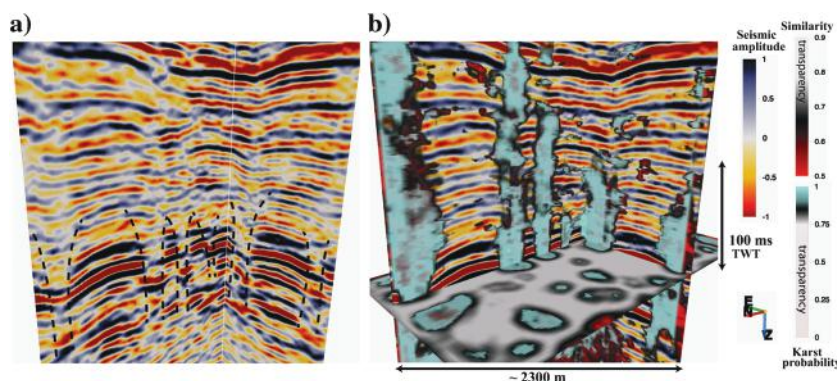


Figure 10. (a) Inline 98 and crossline 70 amplitude profiles with suspected vertical columnar seismic sag (karst collapse) structures designated with dotted lines (same as Figure 3a) and (b) same 2D profiles and a Z-slice at 615 ms TWT, with minimum similarity (low coherence) attribute shown in red/black with partial transparency (same as Figure 3b), now shown with the preferred QDA-KMW probability model overlain on top to demonstrate the corroborating, but not identical, results.

circles) and has fewer lateral connections with neighboring seismic sag structures (white arrows) than does the apparently over-inclusive AFC (Figure 12). These significant differences in system continuity were observed for each of the statistically based models as well, even though quantitative analysis shows that they are quite comparable in terms of error and accuracy. Figure 13 further demonstrates how even the three best perform-

ing models still demonstrate significant differences in lateral and vertical connectivity (white arrows).

Given that it is a physical impossibility to ever observe these subsurface conditions firsthand, regardless of how densely gridded the wells are drilled and logged, there may be no conclusive method for proving which one of the models presented in this study is most reflective of in-situ conditions. Instead, a logical and practical

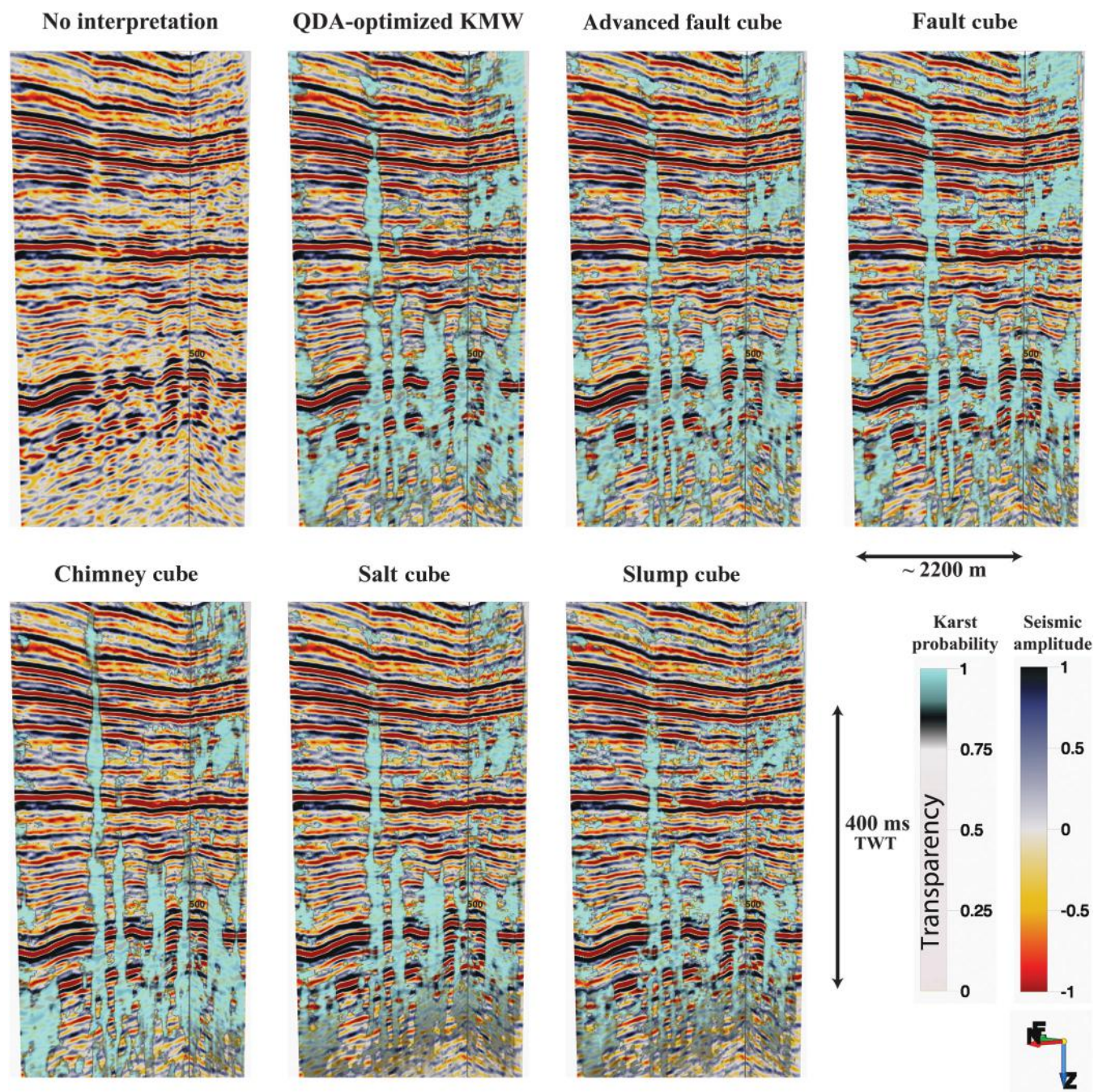


Figure 11. Spatial comparison of the karst probability volume produced using the QDA-optimized KMW, along with those produced using other available NN seismic object detection workflows (all NN trained using the same karst and nonkarst picksets). For visual clarity, the results are shown here as 2D semitransparent overlays on inline 95 and crossline 90 seismic amplitude profiles (direct comparison with Figure 9). The blue semitransparent color scale ranges from 0 to 1, with 0 indicating a low probability of karst and 1 indicating a high probability of karst. Results below approximately 0.75 have been rendered fully transparent.

application of this research would be to use each of the models to estimate ranges for what the in situ conditions are likely to be based on best- and worst-case scenarios and use that information as the basis for factors of safety.

Future work

Recently acquired well-log data from the study survey site will be used to assess the accuracy and representativeness of the interpretations resulting from the KMW. The downhole porosity and permeability measurements should be expected to increase for intervals assigned high karst probability by the KMW. In addition, zones interpreted as karst in the seismic data will be correlated back to the rock core, where available, to determine if the KMW results are supported by lithological observations.

In addition, the KMW will be applied at other 3D seismic reflection survey sites that have comparable data and geologic settings to verify that the method is effective and repeatable. The KMW will be evaluated in a similar manner for these data sets in terms of prediction error rate, spatial agreement, and correlation with downhole geophysics and core physical properties measurements.

Summary

KMW with statistical dimensionality reduction via stepwise QDA was selected as the preferred method for interpreting the approximate spatial boundaries of the presumed karst system at this particular seismic survey site. The QDA model yielded normalized rms error and percentage miscalculation values that were firmly on the low side of the ranges observed for the six statistically selected models evaluated, as well as the very lowest gamma and M -test values. Perhaps more importantly, the QDA model required the least amount of computer processing time and produced the highest ranking spatial result. Agreement between the quantitative and qualitative criteria for the best performing model (s) in this study was by no means a foregone conclusion, nor was it necessarily considered a prerequisite for meaningful results. However, it is greatly encouraging that the models determined to exhibit the best spatial interpretations are also preferred based on combined measures of error, accuracy, and gamma value. Furthermore, by minimizing the number of input seismic attributes and supervised picks used during NN training, this approach drastically reduces computer processing time.

This interpretation does not include estimates of the porosity or permeability conditions within the presumed karst system, nor does it provide theoretical

quantitative limits for these properties at which host rock should no longer be considered karstified. Instead, this study demonstrates the maximal extent to which an interpreter can image subsurface karst systems solely with 3D seismic reflection data. The minimum scale of features capable of being detected with this method is

Table 2. Approximate karst system volume based on various NN interpretation models.

NN multiattribute model name (number of attributes)	Estimated volume of 3D karst system geobody
Initial attribute list (132)	6.34 E+8 m ³
SRGT (61)	3.49 E+8 m ³
Spearman's rho/Welch's T (76)	3.06 E+8 m ³
AFC (56)	2.27 E+8 m ³
NMDS (57)	1.77 E+8 m ³
QDA (45)	1.19 E+8 m ³
Fault Cube (32)	1.18 E+8 m ³
PCA (31)	1.12 E+8 m ³
Slump Cube (11)	1.07 E+8 m ³
Salt Cube (16)	6.88 E+7 m ³
Chimney Cube (28)	6.36 E+7 m ³

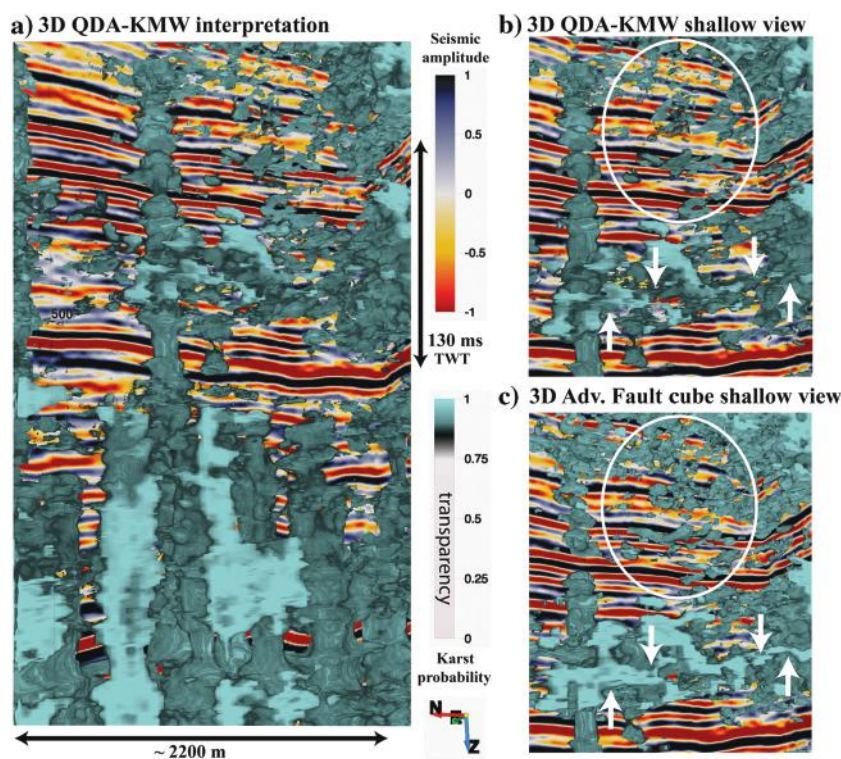


Figure 12. (a) Close-up 3D view of the preferred karst system interpretation resulting from the QDA version of the KMW (QDA-KMW) and (b and c) draw attention to spatial differences (white circle and arrows) observed between the interpretations resulting from the QDA-KMW model relative to the AFC model. Seismic amplitude profiles shown are inline 95 and crossline 90.

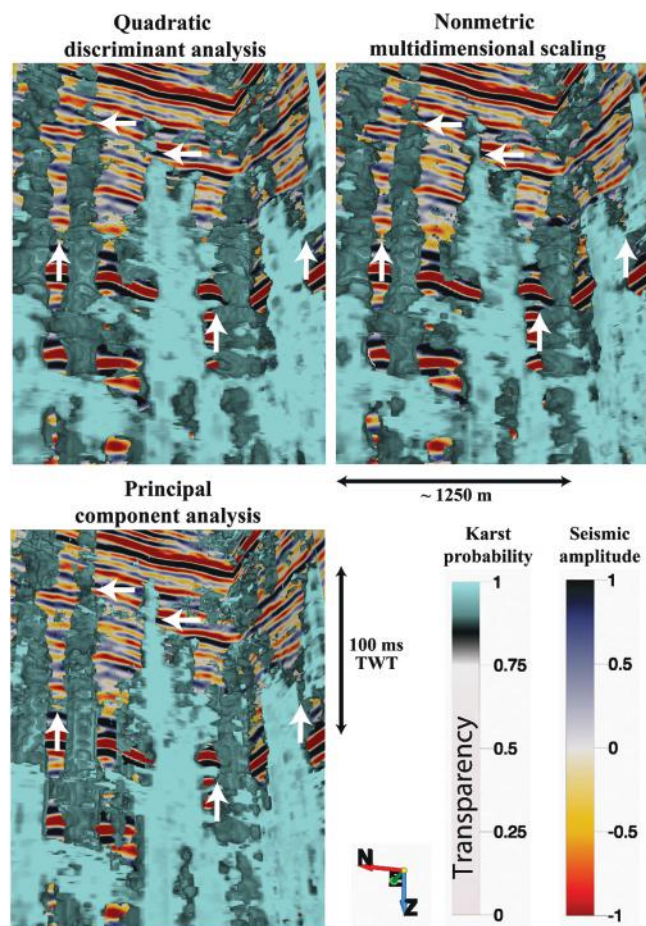


Figure 13. Cross-sectioned data volumes comparing the interconnectivity of the three top-scoring statistically based karst probability models, shown with inline 95 and crossline 90 amplitude profiles in the background and 3D geobodies cross sectioned in the foreground at inline 90 and crossline 88.

controlled by the frequency and sampling parameters set during seismic data acquisition, and some karst systems characterized by small-scale secondary vuggy porosity may not be identified. Moreover, without drilling information to provide groundtruthing, the procedure presented here does not purport to reveal a perfect representation of the in situ karst system, but it rather provides a semiautomated solution for interpreting karst in 3D seismic reflection data that performs as well or better than a human interpreter could, and in a fraction of the time.

Acknowledgments

The authors thank the CESU cooperative agreement between the United States Geological Survey (USGS) Caribbean-Florida Water Science Center and the University of California — Santa Cruz as well as Miami-Dade County for funding this research. We also thank dGB Earth Sciences and IBM for providing academic software licenses. OpendTect Pro software from dGB Earth Sciences was used for all seismic visualization,

attribute analysis, NN training, and comparisons with their existing meta-attribute tools, while the statistical techniques discussed were applied using IBM's SPSS Statistics software. We are also grateful to Cardiff University for providing open access to the winGamma software. Thanks to P. Hart as well for a constructive and thorough review. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by either the University of California or the U.S. government.

References

- Aminzadeh, F., J. Barhen, C. W. Glover, and N. B. Toomarian, 2000, Reservoir parameter estimation using a hybrid neural network: *Computers & Geosciences*, **26**, 869–875, doi: [10.1016/S0098-3004\(00\)00027-3](https://doi.org/10.1016/S0098-3004(00)00027-3).
- Aminzadeh, F., and S. Chatterjee, 1984, Applications of clustering in exploration seismology: *Geoexploration*, **23**, 147–159, doi: [10.1016/0016-7142\(84\)90028-0](https://doi.org/10.1016/0016-7142(84)90028-0).
- Aminzadeh, F., and P. de Groot, 2005, A neural networks based seismic object detection technique: 75th Annual International Meeting, SEG, Expanded Abstracts, 775–778.
- Anderson, T. W., 1984, *An introduction to multivariate statistical analysis*, 2nd ed.: Wiley.
- Borg, I., and P. Groenen, 2005, *Modern multidimensional scaling: Theory and applications*: Springer-Verlag.
- Brouwer, F. C. G., D. Connolly, and K. M. Tingahl, 2011, *A guide to the practical use of neural networks*: 31st Annual Gulf Coast Section Society of Sedimentary Geology, 440–472.
- Carcione, J. M., and S. Picotti, 2006, P-wave seismic attenuation by slow-wave diffusion: Effects of inhomogeneous rock properties: *Geophysics*, **71**, no. 3, O1–O8, doi: [10.1190/1.2194512](https://doi.org/10.1190/1.2194512).
- Chalikakis, K., V. Plagnes, R. Guerin, R. Valois, and F. P. Bosch, 2011, Contribution of geophysical methods to karst-system exploration: An overview: *Hydrogeology Journal*, **19**, 1169–1180, doi: [10.1007/s10040-011-0746-x](https://doi.org/10.1007/s10040-011-0746-x).
- Chopra, S., and K. J. Marfurt, 2005, Seismic attributes: A historical perspective: *Geophysics*, **70**, no. 5, 3S0–2S0, doi: [10.1190/1.2098670](https://doi.org/10.1190/1.2098670).
- Chopra, S., and V. Alexeev, 2006, Applications of texture attribute analysis to 3D seismic data: *The Leading Edge*, **25**, 934–940, doi: [10.1190/1.2335155](https://doi.org/10.1190/1.2335155).
- Conover, W. J., and R. L. Iman, 1981, Rank transformations as a bridge between parametric and nonparametric statistics: *The American Statistician*, **35**, 124–129, doi: [10.2307/2683975](https://doi.org/10.2307/2683975).
- Cunningham, K. J., 2014, Hydrogeologic and seismic-reflection analysis of Cretaceous formations, sub-Floridan confining unit, and Floridan aquifer system in Miami-Dade County, https://fl.water.usgs.gov/projects/cunningham_seismic_reflection/index.html, accessed 5 May 2017.

- Cunningham, K. J., 2015, Seismic-sequence stratigraphy and geologic structure of the Floridan aquifer system near "Boulder Zone" deep wells in Miami-Dade County, Florida: U.S. Geological Survey Scientific Investigations Report 2015–5013, [10.3133/sir20155013](https://doi.org/10.3133/sir20155013).
- Cunningham, K. J., and C. Walker, 2009, Seismic-sag structural systems in Tertiary carbonate rocks beneath southeastern Florida, USA: Evidence for hypogenic speleogenesis, *in* Klimchouk, A. B., and D. C. Ford, eds., Hypogene speleogenesis and karst hydrology of Artesian Basins: Ukrainian Institute of Speleology and Karstology, 151–158.
- Cunningham, K. J., J. W. Kluesner, R. L. Westcott, E. Robinson, C. Walker, and S. A. Khan, 2018, Sequence stratigraphy, seismic stratigraphy, and seismic structures of the lower intermediate confining unit and most of the Floridan aquifer system, Broward County, Florida (ver. 1.1, January 2018): U.S. Geological Survey Scientific Investigations Report 2017–5109.
- de Groot, P., H. Ligtenberg, P. Meldahl, and R. Heggland, 2001, Selecting and combining attributes to enhance detection of seismic objects: Presented at the 63rd Annual International Conference and Exhibition, EAGE.
- de Groot, P., H. Ligtenberg, T. Oldenziel, D. Connolly, and P. Meldahl, 2004, Examples of multi-attribute, neural network-based seismic object detection: Geological Society of London, Memoirs 29, 335–338.
- Eichkitz, C. G., M. G. Schreilechner, P. de Groot, and J. Amtmann, 2015, Mapping directional variations in seismic character using gray-level co-occurrence matrix-based attributes: Interpretation, **3**, no. 1, T13–T23, doi: [10.1190/INT-2014-0099.1](https://doi.org/10.1190/INT-2014-0099.1).
- Friedman, J. H., 1989, Regularized discriminant analysis: Journal of the American Statistical Association, **84**, 165–175, doi: [10.1080/01621459.1989.10478752](https://doi.org/10.1080/01621459.1989.10478752).
- Gao, D., 2011, Latest developments in seismic texture analysis for subsurface structure, facies, and reservoir characterization: A review: Geophysics, **76**, no. 2, W1–W13, doi: [10.1190/1.3553479](https://doi.org/10.1190/1.3553479).
- Hampson, D., J. Schuelke, and J. Quirein, 2001, Use of multiattribute transforms to predict log properties from seismic data: Geophysics, **66**, 220–236, doi: [10.1190/1.1444899](https://doi.org/10.1190/1.1444899).
- Hardage, B. A., D. L. Carr, D. E. Lancaster, J. L. Simmons Jr., R. Y. Elphick, V. M. Pendleton, and R. A. Johns, 1996, 3-D seismic evidence of the effects of carbonate karst collapse on overlying clastic stratigraphy and reservoir compartmentalization: Geophysics, **61**, 1336–1350, doi: [10.1190/1.1444057](https://doi.org/10.1190/1.1444057).
- Hashemi, H., D. M. J. Tax, R. P. W. Duin, A. Javaherian, and P. de Groot, 2008, Gas chimney detection based on improving the performance of combined multilayer perceptron and support vector classifier: Nonlinear Processes in Geophysics, **15**, 863–871, doi: [10.5194/npg-15-863-2008](https://doi.org/10.5194/npg-15-863-2008).
- Hine, A. C., 2013, Geologic history of Florida — Major events that formed the sunshine state: University of Florida Press.
- Hush, D. R., and B. G. Horne, 1993, Progress in supervised neural networks: IEEE Signal Processing Magazine, **10**, 8–39, doi: [10.1109/79.180705](https://doi.org/10.1109/79.180705).
- Iturrarán-Viveros, U., 2012, Smooth regression to estimate effective porosity using seismic attributes: Journal of Applied Geophysics, **76**, 1–12, doi: [10.1016/j.jappgeo.2011.10.012](https://doi.org/10.1016/j.jappgeo.2011.10.012).
- Iturrarán-Viveros, U., and J. O. Parra, 2014, Artificial neural networks applied to estimate permeability, porosity and intrinsic attenuation using seismic attributes and well-log data: Journal of Applied Geophysics, **107**, 45–54, doi: [10.1016/j.jappgeo.2014.05.010](https://doi.org/10.1016/j.jappgeo.2014.05.010).
- Jones, A. J., 2002, winGamma manual & winGamma software: Cardiff University, <http://users.cs.cf.ac.uk/O.F.Rana/Antonia.J.Jones/GammaArchive/IndexPage.htm>, accessed 19 January 2016.
- Kalkomey, C. T., 1997, Potential risks when using seismic attributes as predictors of reservoir properties: The Leading Edge, **16**, 247–251, doi: [10.1190/1.1437610](https://doi.org/10.1190/1.1437610).
- Klecka, W. R., 1980, Discriminant analysis: Sage Publications Inc.
- Klimchouk, A., 2015, The karst paradigm: Changes, trends and perspectives: Acta Carsologica, **44**, 289–313, doi: [10.3986/ac.v44i3.2996](https://doi.org/10.3986/ac.v44i3.2996).
- Kluesner, J., E. Silver, N. Bangs, K. McIntosh, J. Gibson, R. von Huene, D. Orange, and C. Ranero, 2013, High density of structurally-controlled, shallow to deep water fluid seeps imaged offshore Costa Rica: Geochemistry, Geophysics, Geosystems, **14**, 519–539, doi: [10.1002/ggge.20058](https://doi.org/10.1002/ggge.20058).
- Kluesner, J. W., and D. S. Brothers, 2016, Seismic attribute detection of faults and fluid pathways within an active strike-slip shear zone: New insights from high-resolution 3D P-Cable™ seismic data along the Hosgri Fault, offshore California: Interpretation, **4**, no. 1, SB131–SB148, doi: [10.1190/INT-2015-0143.1](https://doi.org/10.1190/INT-2015-0143.1).
- Ligtenberg, J. H., 2005, Detection of fluid migration pathways in seismic data: Implications for fault seal analysis: Basin Research, **17**, 141–153, doi: [10.1111/j.1365-2117.2005.00258.x](https://doi.org/10.1111/j.1365-2117.2005.00258.x).
- Loucks, R. G., 1999, Paleocave carbonate reservoirs: Origins, burial-depth modifications, spatial complexity, and reservoir implications: AAPG Bulletin, **83**, 1795–1834, doi: [10.1306/E4FD426F-1732-11D7-8645000102C1865D](https://doi.org/10.1306/E4FD426F-1732-11D7-8645000102C1865D).
- Lucia, F. J., 1983, Petrophysical parameters estimated from visual description of carbonate rocks: A field classification of carbonate pore space: Journal of Petroleum Technology, **35**, 629–637, doi: [10.2118/10073-PA](https://doi.org/10.2118/10073-PA).
- Ma, Y. Z., and E. Gomez, 2015, Uses and abuses in applying neural networks for predictions in hydrocarbon resource evaluation: Journal of Petroleum Science and Engineering, **133**, 66–75, doi: [10.1016/j.petrol.2015.05.006](https://doi.org/10.1016/j.petrol.2015.05.006).

- Martínez, A. M., and A. C. Kak, 2001, PCA versus LDA: IEEE Transactions on Pattern Analysis and Machine Intelligence, **23**, 228–233, doi: [10.1109/34.908974](https://doi.org/10.1109/34.908974).
- McDonnell, A., 2007, Quantifying the origin and geometry of circular sag structures in northern Fort Worth Basin, Texas: Paleocave collapse, pull-apart fault systems, or hydrothermal alteration?: AAPG Bulletin, **91**, 1295–1318, doi: [10.1306/05170706086](https://doi.org/10.1306/05170706086).
- Meldahl, P., R. Heggland, B. Bril, and P. de Groot, 1999, The chimney cube, an example of semi-automated detection of seismic objects by directive attributes and neural networks. Part 1: Methodology: 79th Annual International Meeting, SEG, Expanded Abstracts, 931–934, doi: [10.1190/1.1821262](https://doi.org/10.1190/1.1821262).
- Meldahl, P., R. Heggland, B. Bril, and P. de Groot, 2001, Identifying fault and gas chimneys using multi-attributes and neural networks: The Leading Edge, **20**, 474–482, doi: [10.1190/1.1438976](https://doi.org/10.1190/1.1438976).
- Meldahl, P., R. Heggland, P. de Groot, and B. Bril, 2000, Semi-automated detection of seismic objects by directive attributes and neural networks, method and applications: 62nd Annual International Conference and Exhibition, EAGE, Extended Abstracts, A-0.
- Miller, J. A., 1986, Hydrogeologic framework of the Floridan aquifer system in Florida, and in parts of Georgia, Alabama, and South Carolina: U.S. Geological Survey Professional Paper 1403-B, B1–B91.
- OpendTect, 2016, OpendTect Pro Plugins Documentation version 6.0, http://static.opendtect.org/images/PDF/dgb_userdoc.pdf, accessed 13 September 2016.
- Peres-Neto, P. R., D. A. Jackson, and K. M. Somers, 2005, How many principal components? Stopping rules for determining the number of non-trivial axes revisited: Computational Statistics & Data Analysis, **49**, 974–997, doi: [10.1016/j.csda.2004.06.015](https://doi.org/10.1016/j.csda.2004.06.015).
- Reese, R. S., and K. J. Cunningham, 2014, Hydrogeologic framework and salinity distribution of the Floridan Aquifer System of Broward County, Florida (No. 2014-5029): US Geological Survey Scientific Investigations Report, doi: [10.3133/sir20145029](https://doi.org/10.3133/sir20145029).
- Roden, R., T. Smith, and D. Sacrey, 2015, Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps: Interpretation, **3**, no. 4, SAE59–SAE83, doi: [10.1190/INT-2015-0037.1](https://doi.org/10.1190/INT-2015-0037.1).
- Russell, B., D. Hampson, J. Schuelke, and J. Quirein, 1997, Multiattribute seismic analysis: The Leading Edge, **16**, 1439–1444, doi: [10.1190/1.1437486](https://doi.org/10.1190/1.1437486).
- Scales, J., 1995, Theory of seismic imaging: Lecture Notes in Earth Sciences: Springer-Verlag **55**.
- Shepard, R. N., 1980, Multidimensional scaling, tree-fitting, and clustering: Science, **210**, 390–398, doi: [10.1126/science.210.4468.390](https://doi.org/10.1126/science.210.4468.390).
- Sontag, E. D., 1989, Sigmoids distinguish more efficiently than heavisides: Neural Computation, **1**, 470–472, doi: [10.1162/neco.1989.1.4.470](https://doi.org/10.1162/neco.1989.1.4.470).
- Spearman, C., 1987, The proof and measurement of association between two things: The American Journal of Psychology, **100**, 441–471, doi: [10.2307/1422689](https://doi.org/10.2307/1422689).
- Spechler, R. M., 2001, The relation between structure and saltwater intrusion in the Floridan Aquifer System, Northeastern Florida: U.S. Geological Survey Karst Interest Group proceedings, Water-Resources Investigations Report 01-4011, 25–29, http://water.usgs.gov/ogw/karst/kigconference/pdf/rms_relationintrusion.pdf, accessed 15 July 2017.
- Stefánsson, A., N. Končar, and A. J. Jones, 1997, A note on the gamma test: Neural Computing & Applications, **5**, 131–133, doi: [10.1007/BF01413858](https://doi.org/10.1007/BF01413858).
- Stone, M., 1974, Cross-validated choice and assessment of statistical predictions: Journal of the Royal Statistical Society: Series B (Methodological), **36**, 111–147.
- Sullivan, E. C., K. J. Marfurt, A. Lacazette, and M. Ammerman, 2006, Application of new seismic attributes to collapse chimneys in the Fort Worth Basin: Geophysics, **71**, no. 4, B111–B119, doi: [10.1190/1.2216189](https://doi.org/10.1190/1.2216189).
- Šumanovac, F., and M. Weissner, 2001, Evaluation of resistivity and seismic methods for hydrogeological mapping in karst terrains: Journal of Applied Geophysics, **47**, 13–28, doi: [10.1016/S0926-9851\(01\)00044-1](https://doi.org/10.1016/S0926-9851(01)00044-1).
- Taner, M. T., 2001, Seismic attributes: CSEG Recorder, **26**, 48–56.
- Zimmerman, D. W., and B. D. Zumbo, 1993, Rank transformations and the power of the Student t test and Welch t' test for non-normal populations with unequal variances: Canadian Journal of Experimental Psychology, **47**, 523–539, doi: [10.1037/h0078850](https://doi.org/10.1037/h0078850).



Daniel R. Ebuna received a B.S. (2008) in geology and geologic engineering from the Colorado School of Mines and an M.S. (2011) from the Scripps Institution of Oceanography. After working for a few years in industry as a hydrographic surveyor and geophysical technician, he joined the University of California Santa Cruz (UCSC) as a seismic interpretation specialist in 2015. His primary field of interest is multichannel seismic reflection data acquisition, processing, and interpretation with an emphasis on optimizing techniques to minimize user bias, procedural uncertainty, and data artifacts.



Jared W. Kluesner received a B.S. (2006) from Indiana State University and a Ph.D. (2011) from Scripps Institution of Oceanography. In 2012, he became a postdoctoral scholar at UCSC, where he conducted research on the geophysical detection and visualization of faults and fluid-migration pathways along the Costa Rican

margin. In 2014, he became a researcher at UCSC, and in 2015, he joined the USGS Pacific Coastal and Marine Science Center in Santa Cruz, California. His current research interests include using 3D seismic attribute analyses to better understand fluid migration at passive and active plate boundaries. In addition, he also studies and develops swath sonar and 2D/3D seismic reflection acquisition, processing, and visualization techniques.



Kevin J. Cunningham received a B.S. in geology from the University of Wisconsin-Oshkosh and an M.S. in geology from Louisiana State University. After working for eight years for Shell Oil as an exploration geologist, he received a Ph.D. (1995) in geology from the University of Kansas. Subsequently, he conducted two years of postdoctoral research at the Division of Marine Geology and Geophysics, Rosenstiel School of Marine and Atmospheric Science.

Since then, he has been a research hydrogeologist at the USGS in South Florida, where he developed the Carbonate Aquifer Characterization Laboratory. He is currently leading a project that is advancing the Cretaceous and Cenozoic sequence stratigraphic and hydrogeologic framework of southeastern Florida using mainly core, well-logs, and 2D/3D seismic reflection data.



Joel H. Edwards received a B.S. (2010) in geology from Brigham Young University and an M.S. (2013) in geology from the University of Nevada, Reno. He then worked in the renewable energy industry in Turkey and Washington DC, before moving to Santa Cruz in 2014 to pursue a Ph.D. with E. Silver and J. Kluesner. He is a Ph.D. candidate at UCSC working on 3D characterization of subduction zones and southeast Florida using swath sonar, 2D/3D seismic reflection, and well-field data sets.