MEWBOURNE COLLEGE OF EARTH AND ENERGY
SCHOOL OF GEOSCIENCES
The UNIVERSITY of OKLAHOMA

AASPI

Seismic Attributes - from Interactive Interpretation to Machine Learning

Kurt J. Marfurt (The University of Oklahoma)

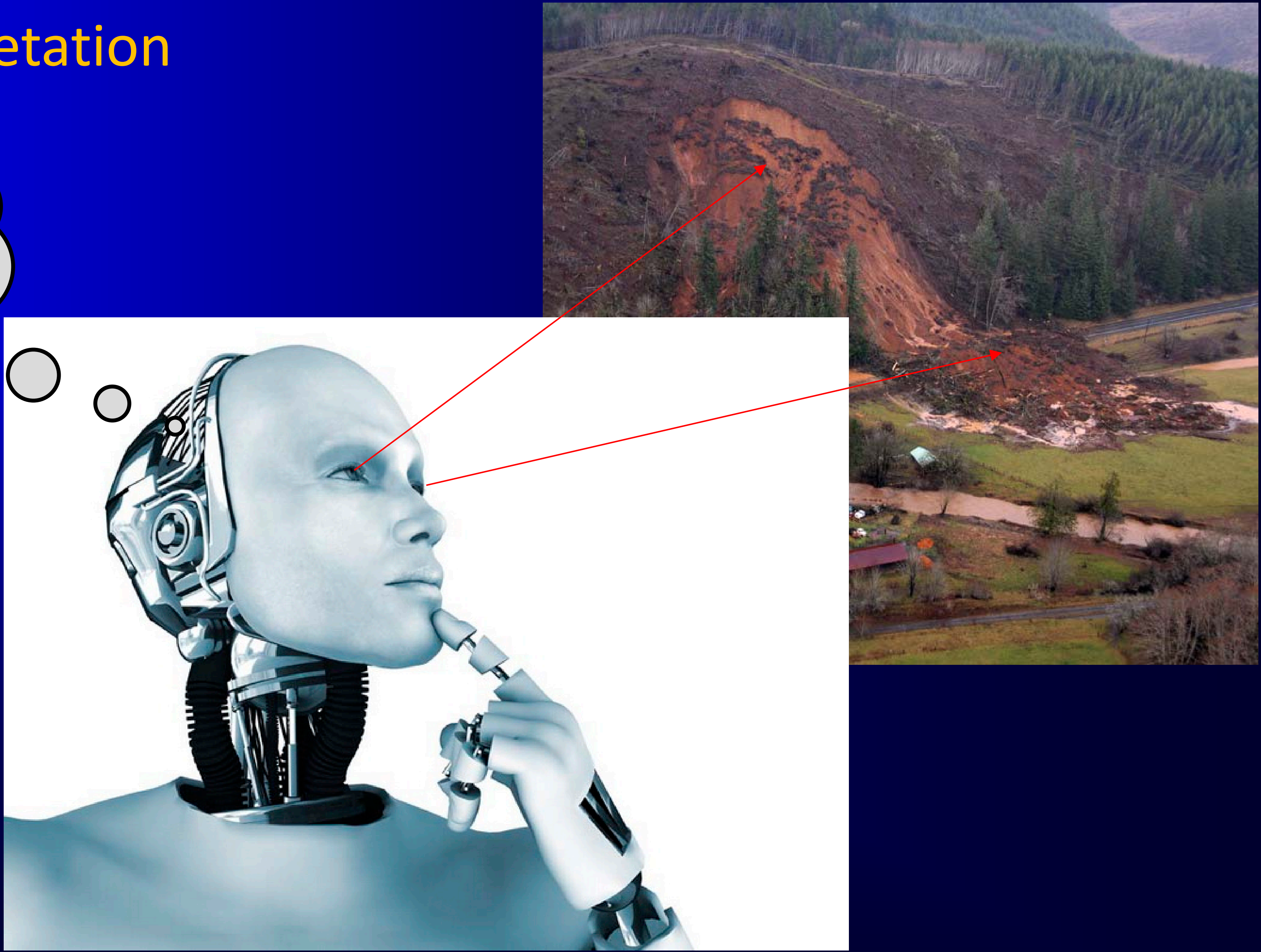
Attribute selection for seismic facies classification

Objectives in attribute selection

- Selecting attributes that are correlated to the facies of interest
- Rejecting attributes that are meaningless for the problem at hand
- Minimizing the use of redundant (or mathematically correlated) attributes

Machine-Driven Interpretation

Which attribute should I use?



Stochastic Workflow

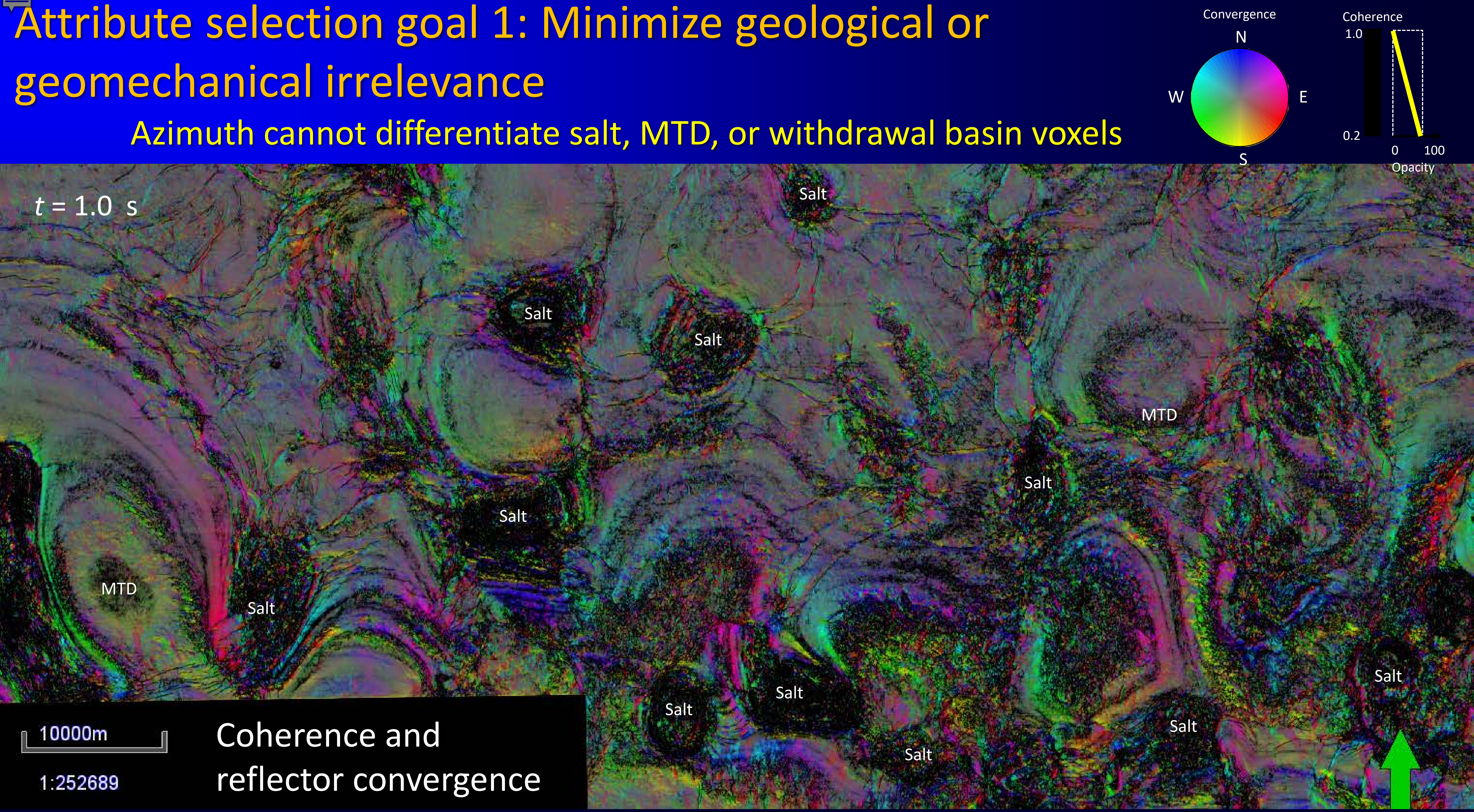
-> Try them all and see which ones correlate

Deterministic Workflow

-> Use attributes associated with a geologic model or process

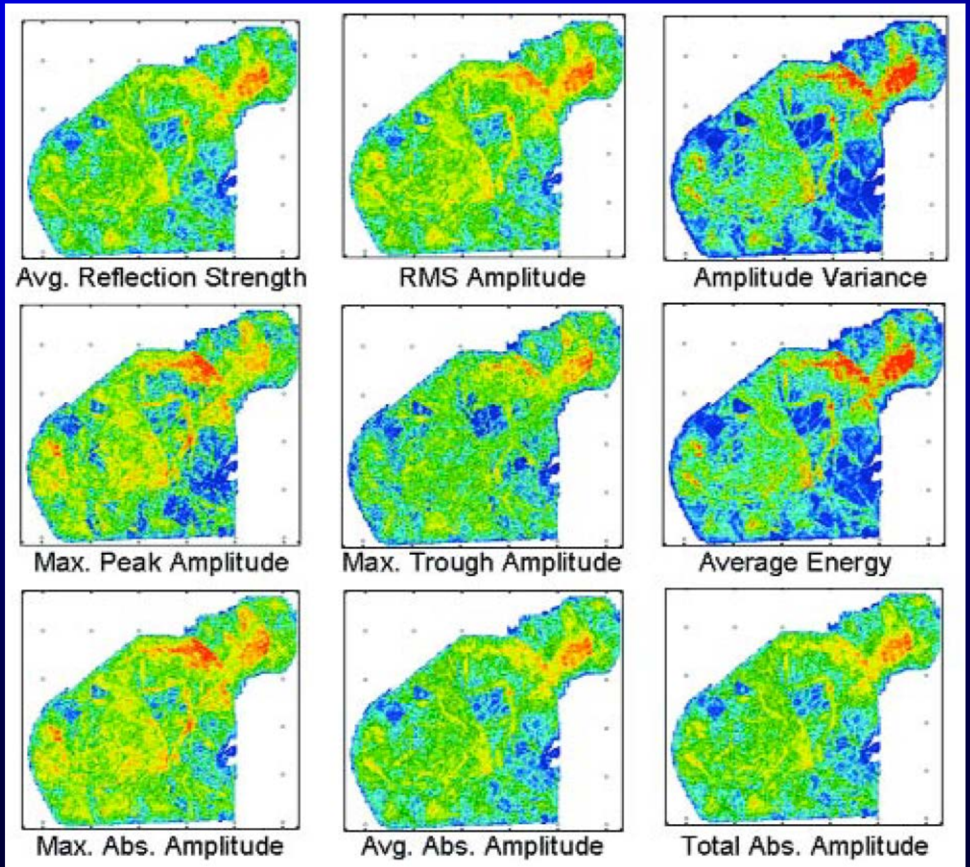
Attribute selection goal 1: Minimize geological or geomechanical irrelevance

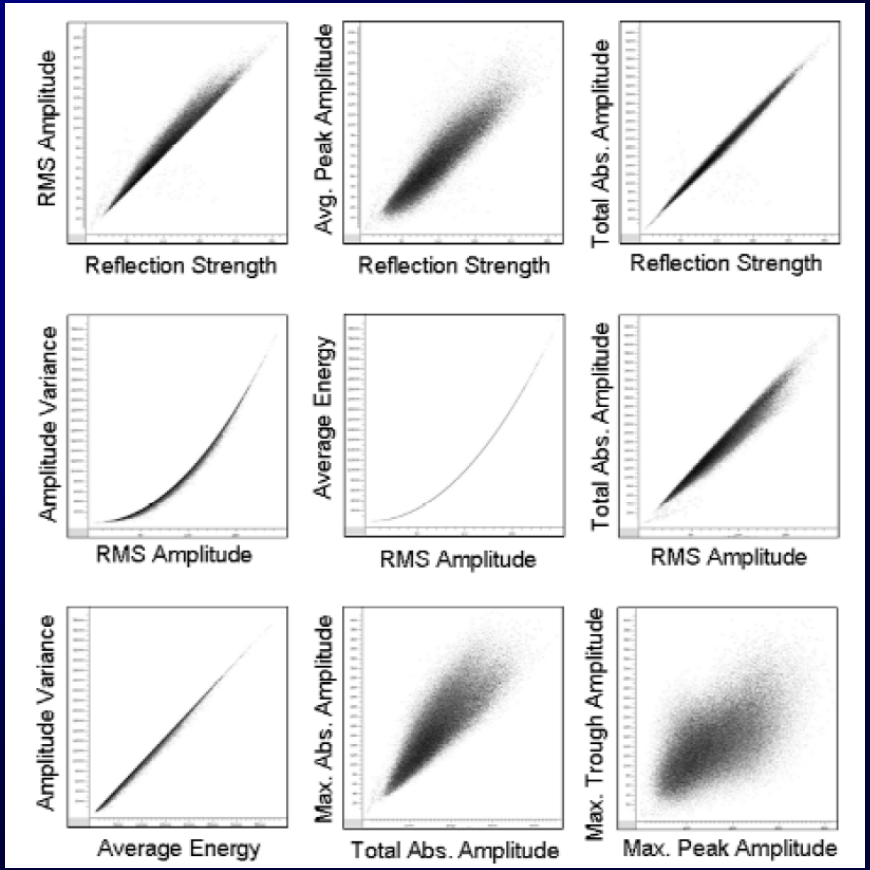
Azimuth cannot differentiate salt, MTD, or withdrawal basin voxels



Attribute selection goal 2: Minimize mathematical redundancy

- Some attributes are correlated through their algorithmic implementation and provide little useful additional information
- Instead, we want attributes that are correlated through geology or geomechanics

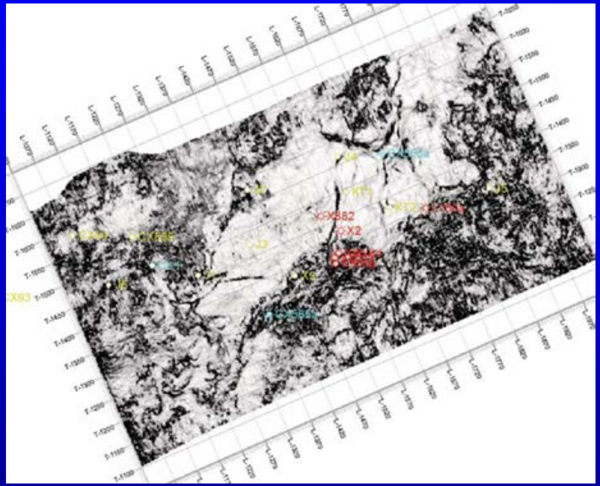




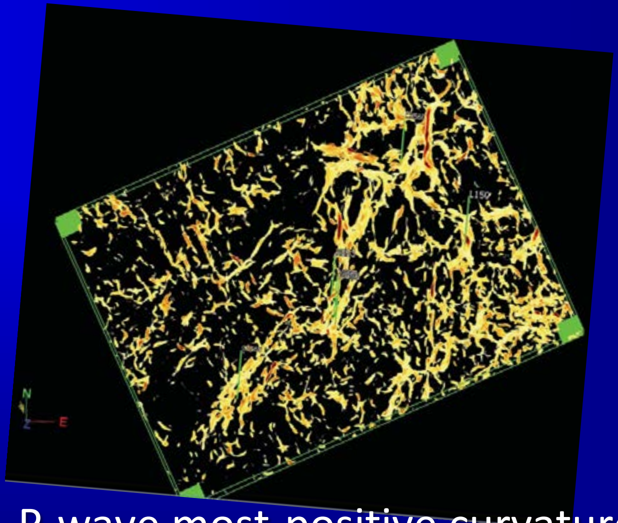
(Barnes, 2006)

Attribute selection goal 3: Maximize geological or geomechanical redundancy

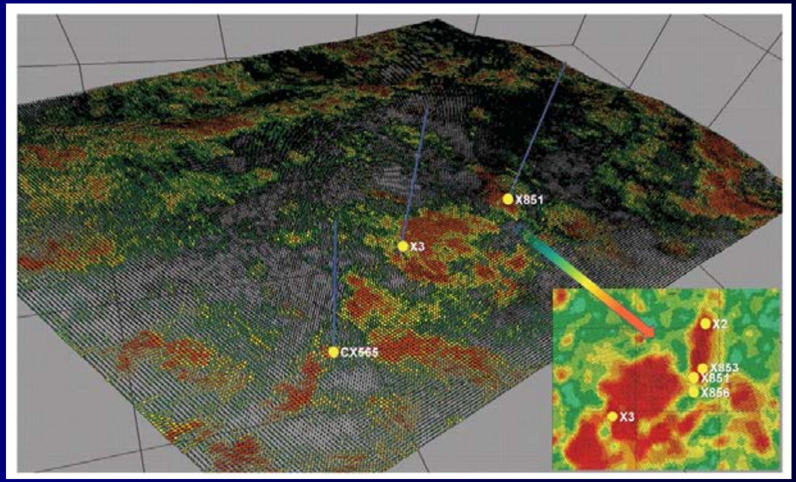
Mapping fractures in a Sichuan carbonate reservoir



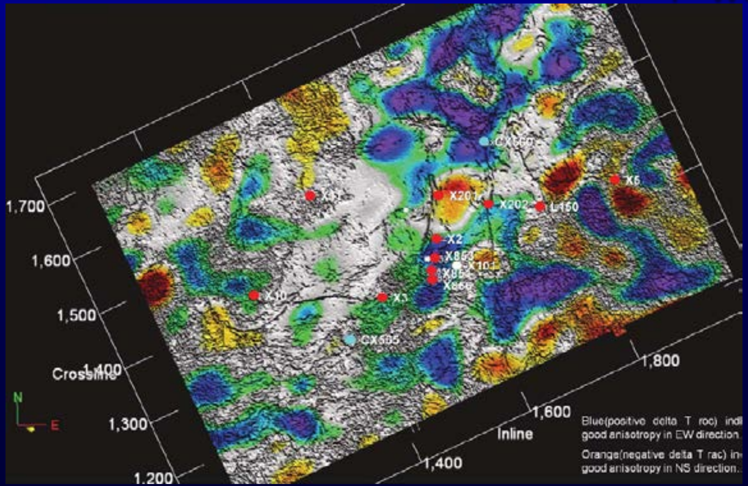
P-wave coherence



P-wave most-positive curvature



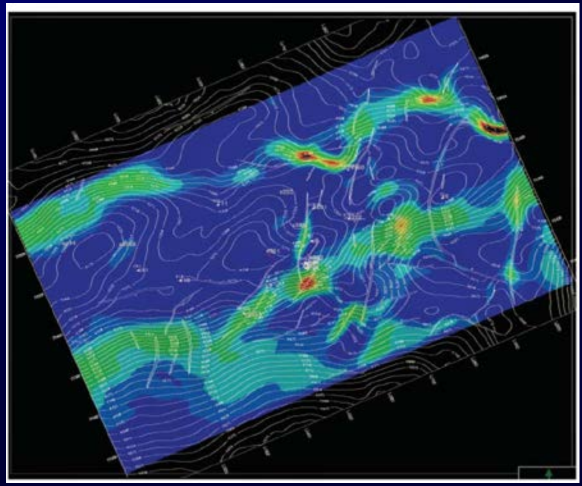
P-wave AVAz anisotropy



Shear wave time thickness difference

Section	Sand	Curvature Attribute		S-wave Splitting		P-wave Azimuthal Anisotropy	Coherency
		Class 1	Class 2	Class 1	Class 2		
Xujiahe Member 4	TX_4^8	>5	>7				<60
	TX_4^9						
Xujiahe Member 2	TX_2^2	>3	>2.6	Blue zone	Blue + green zone	>1.06	<50
	TX_2^3	>3	>2.6				
	TX_2^4	>3	>2.6	Blue zone	Blue + green zone		
	TX_2^5	>3.6	>3				
	TX_2^6	>4.6	>4			>1.09	
	TX_2^7	>6	>5				

Risk analysis table

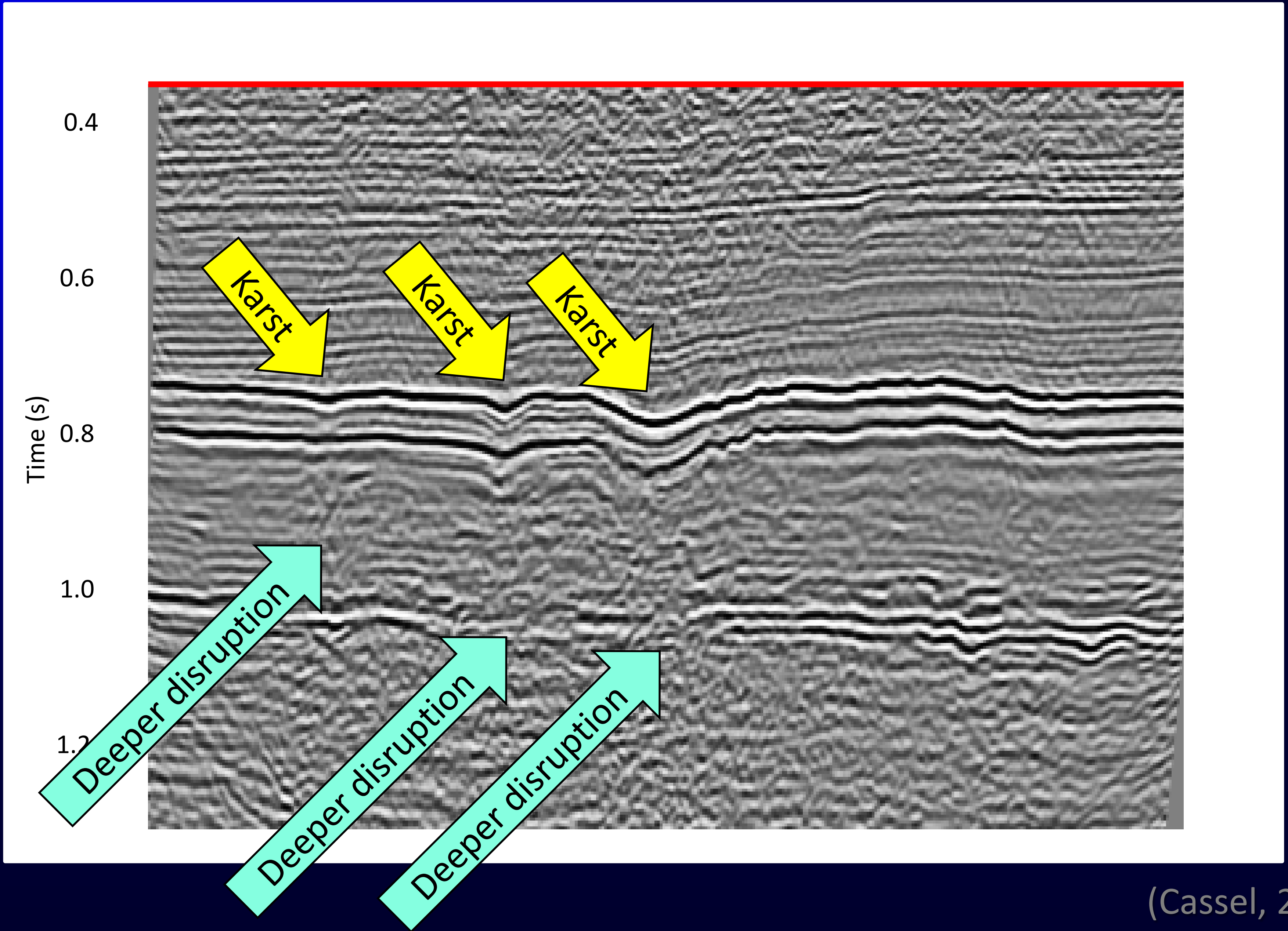


Multiattribute fracture prediction

(Jianming et al., 2009)

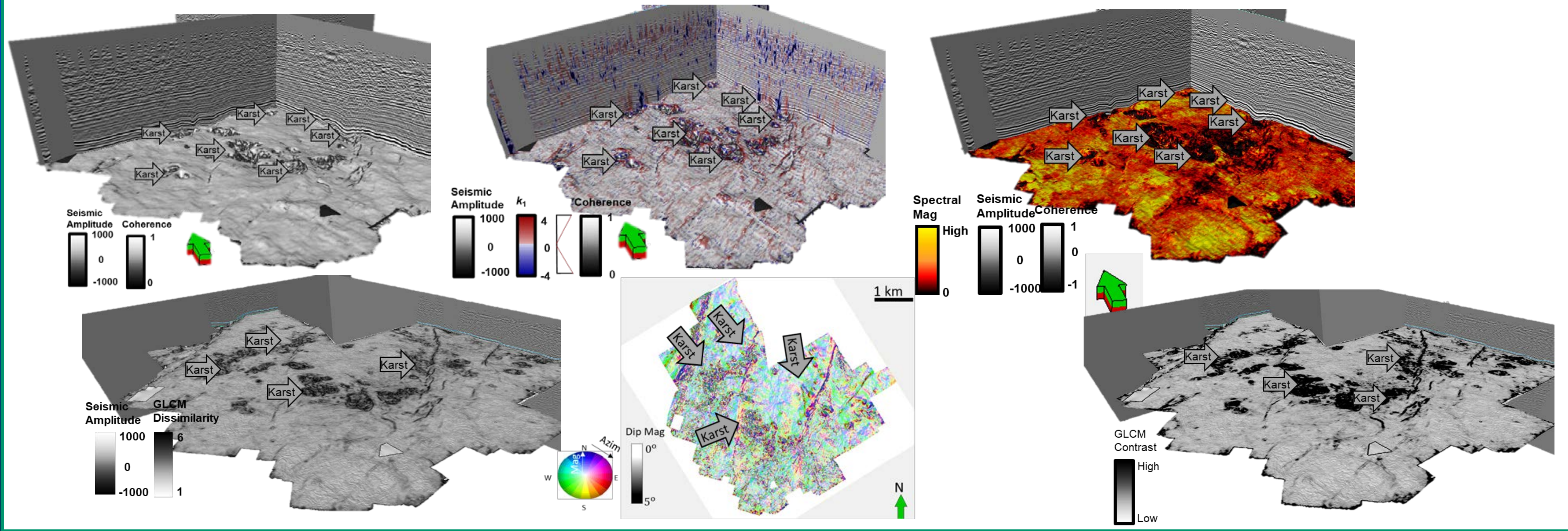
Attribute selection based on interpreter insight

Karst collapse features in the Fort Worth Basin



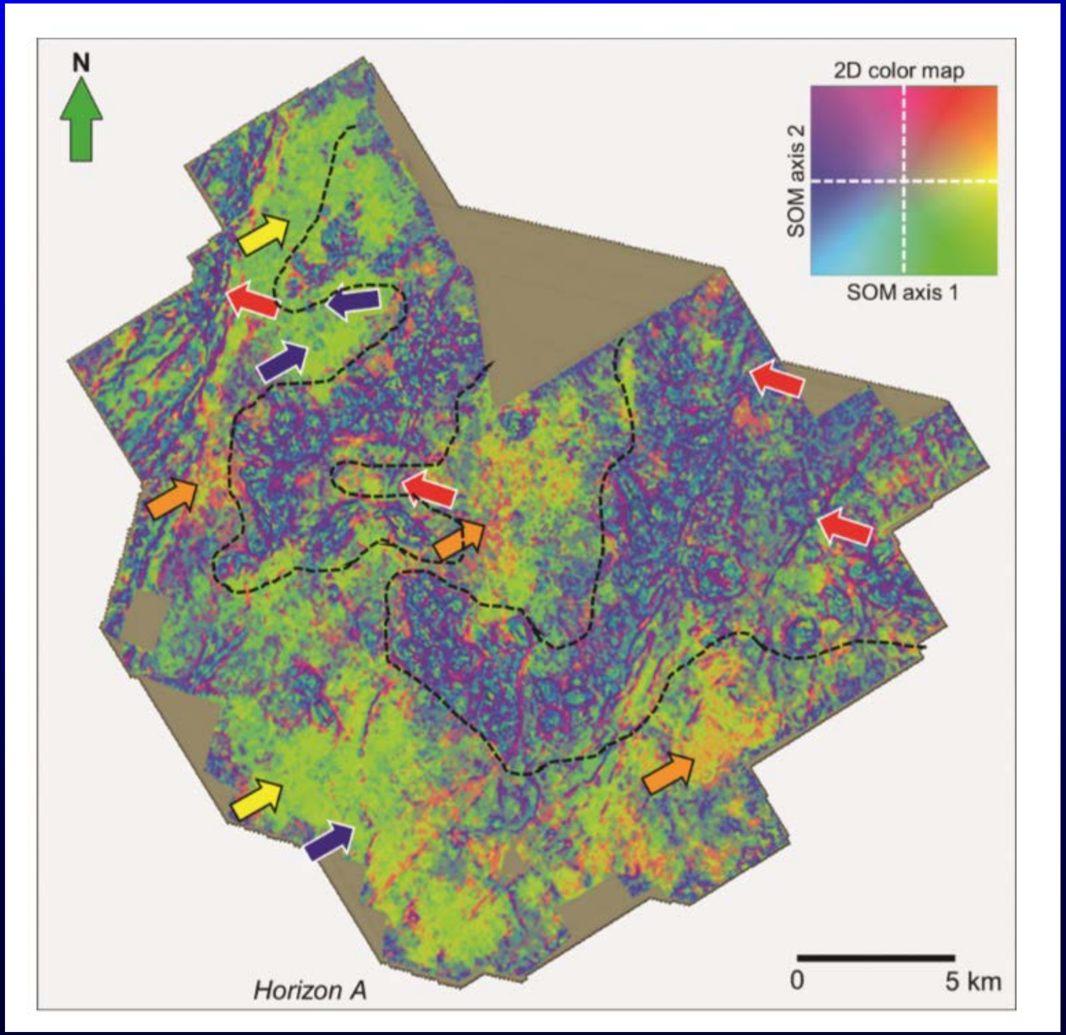
(Cassel, 2018)

Attribute selection based on interpreter insight



(Cassel, 2018)

SOM identification of karst
Ellenburger dolomite, Fort Worth Basin, TX



Attributes used:

- Coherence
- GLCM homogeneity
- Spectral Magnitude
- Amplitude Curvature e_1 and e_2
- Structural Curvature k_1 and k_2
- Peak Frequency

(Zhao et al., 2018)

Attribute selection based on interpreter insight

Geologic Feature	Visual Expression	
	Internal Reflector Pat	External Reflecto
Kar	• Piecewise continuous to • Bo • No ening • Narrow band, lower frequency	• Sediment fill on t • Di low • Ec olled • Broad band

Coherence
GLCM Entropy
GLCM Homogeneity

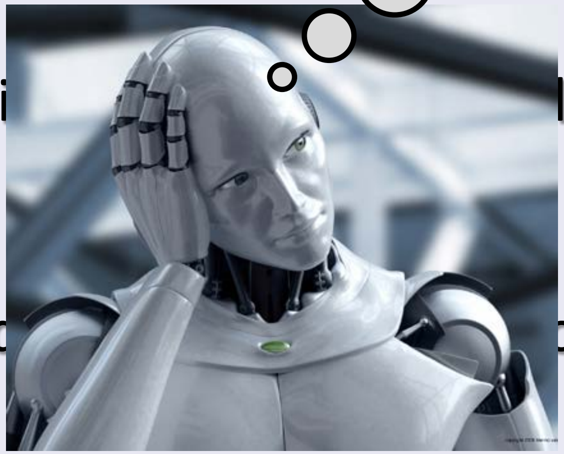
Structural curvature
Bowl shape index

Reflector convergence
Parallelism

Peak magnitude
Peak frequency
Bandwidth

Voxel-based measures:
I can do this!

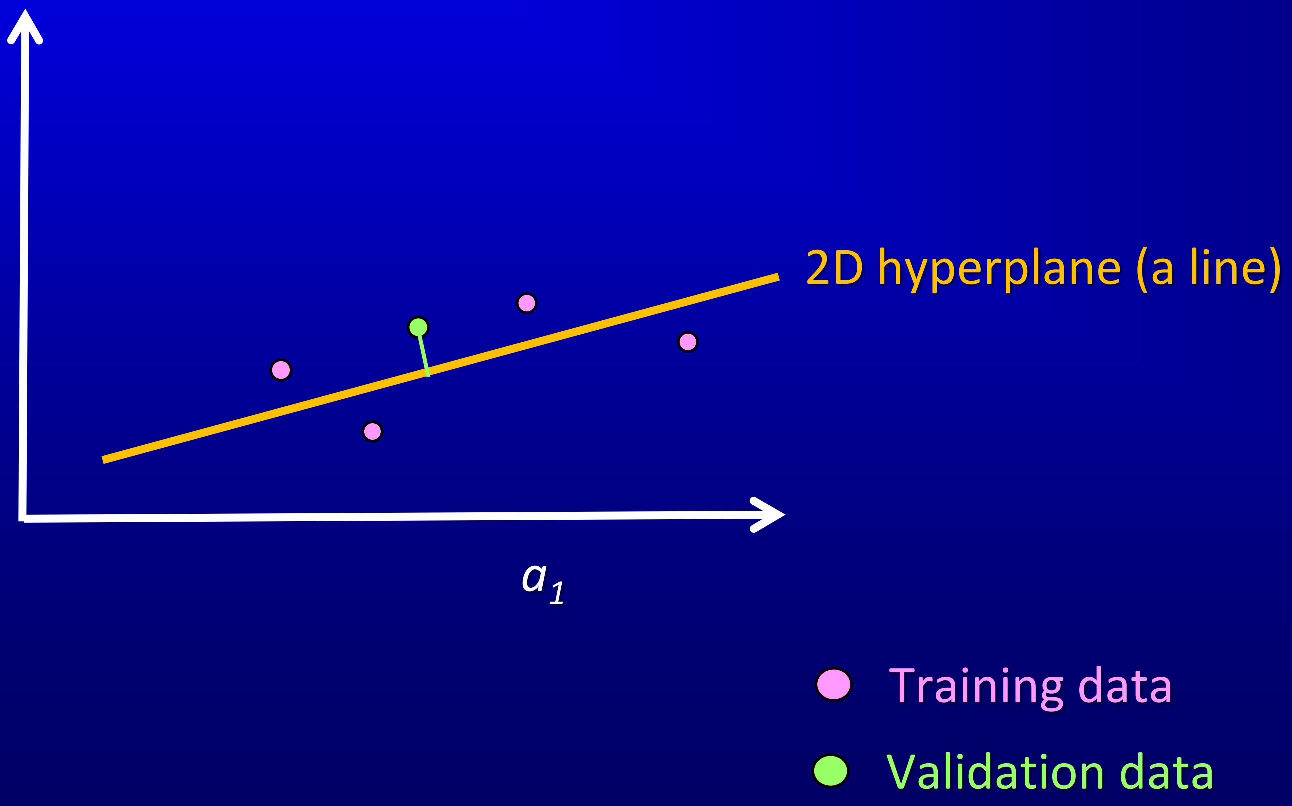
Contextual setting/scene analysis:
Geez! This is tough!



(Cassel, 2018)

Stepwise regression and the danger of too many attributes

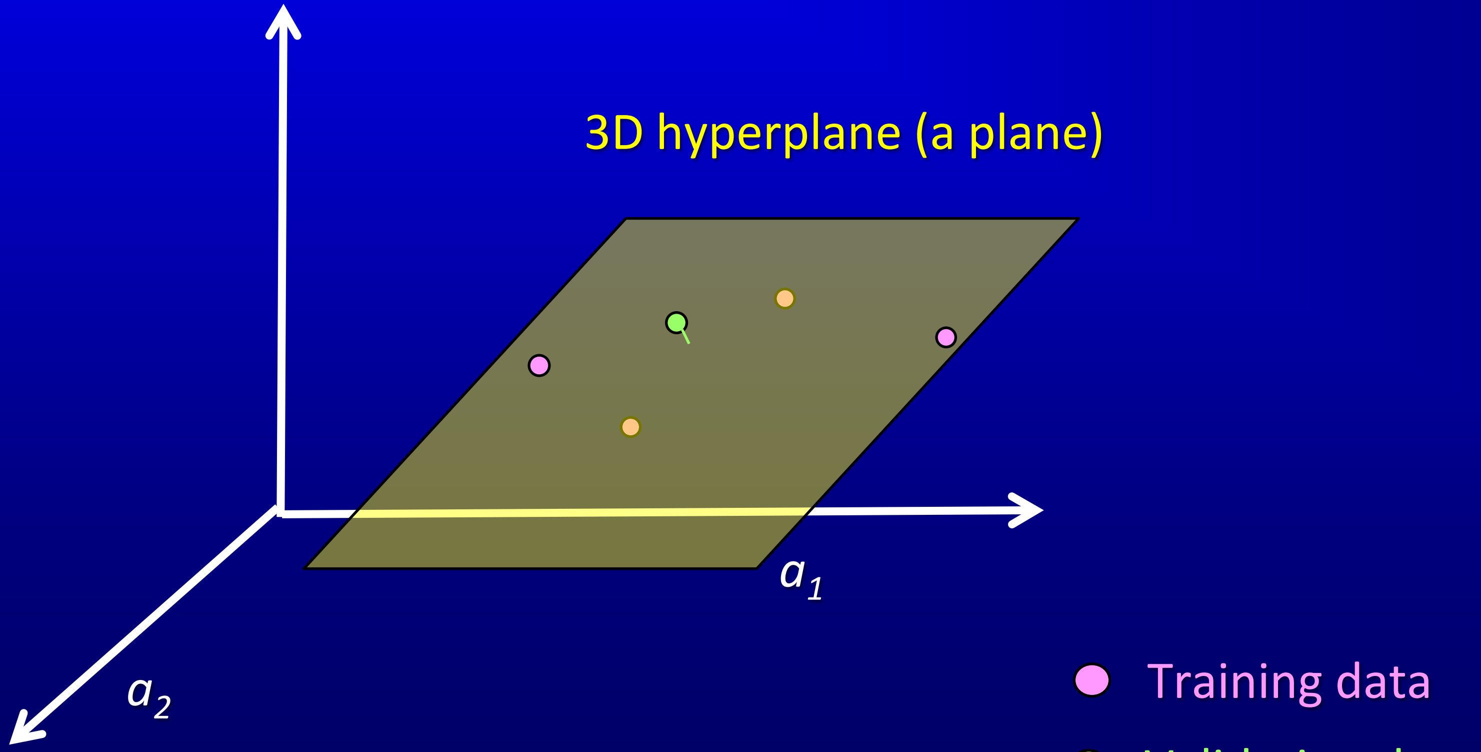
$$L = w_0 + \sum_{j=1}^{1} w_j a_j$$



(Marfurt, 2018)

Stepwise regression and the danger of too many attributes

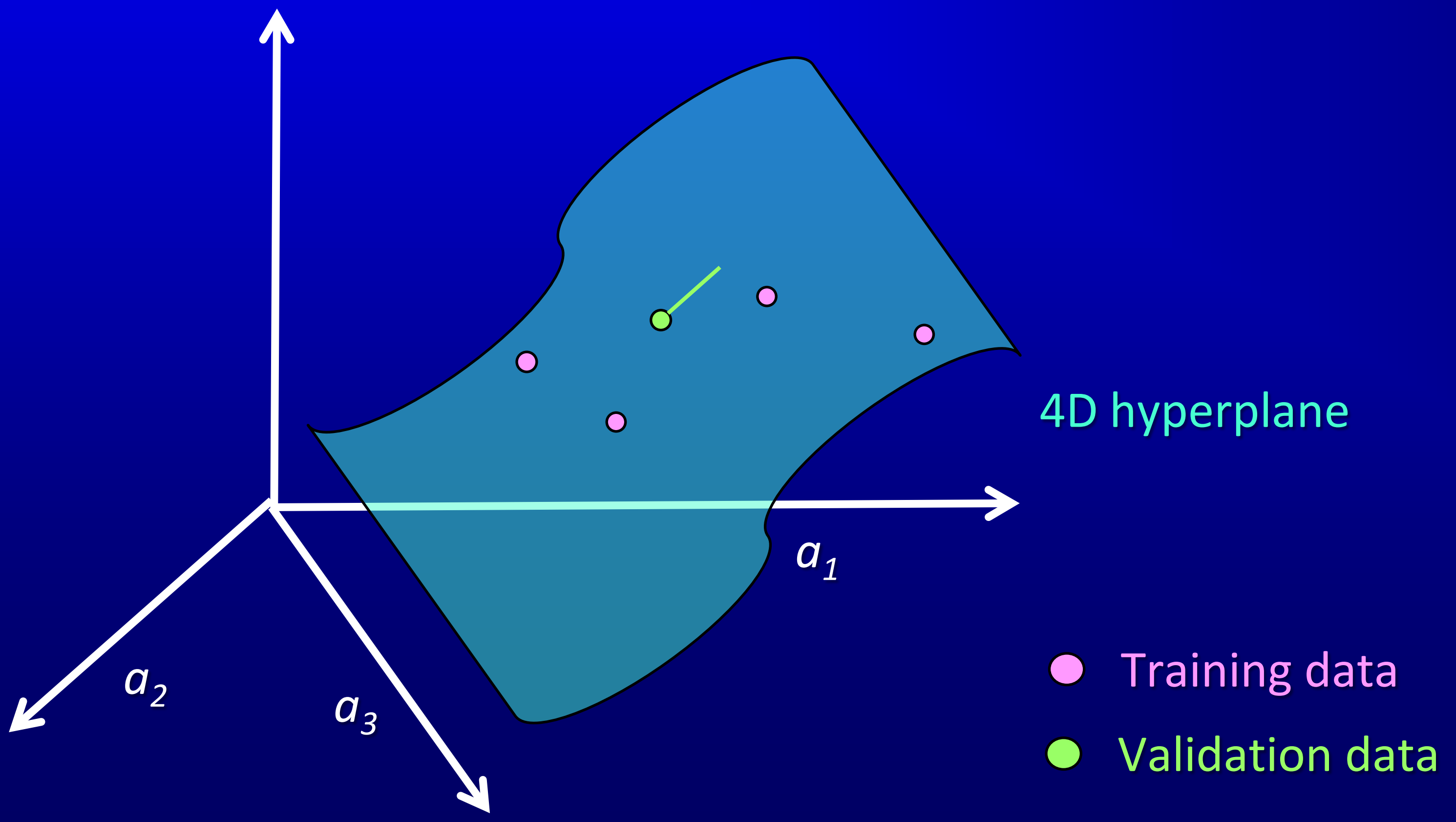
$$L = w_0 + \sum_{j=1}^{2} w_j a_j$$



(Marfurt, 2018)

Stepwise regression and the danger of too many attributes

$$L = w_0 + \sum_{j=1}^{3} w_j a_j$$



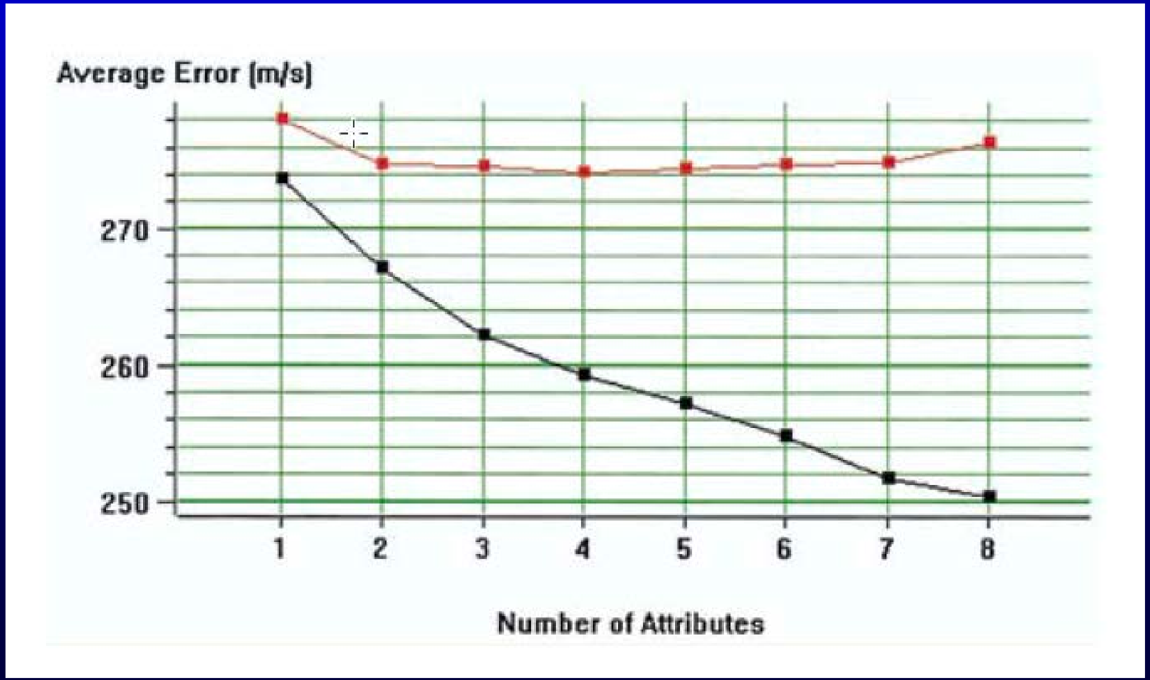
(Marfurt, 2018)

Attribute selection based on statistical analysis

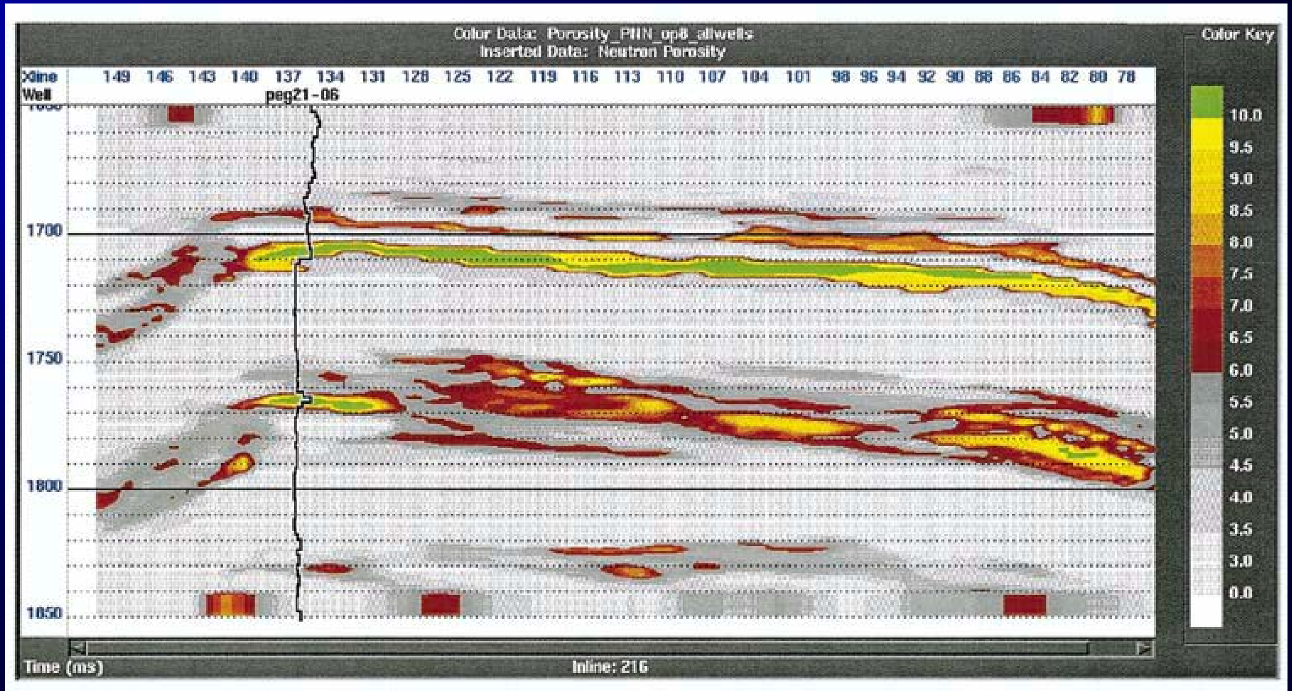
Use of multiattribute transforms to predict log properties from seismic data

Daniel P. Hampson*, James S. Schuelke‡, and John A. Quirein**

2001



Stepwise linear regression
(nonlinear transforms can be applied to the data before regression)



PNN prediction of porosity

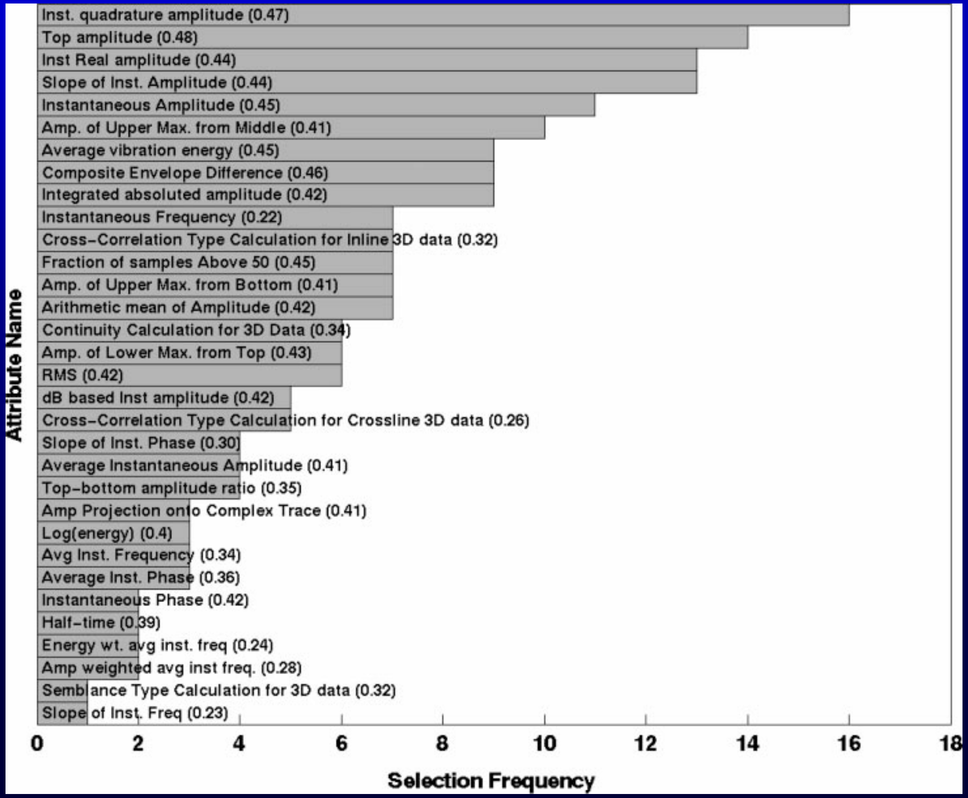
(Hampson et al., 2001)

Attribute selection based on statistical analysis

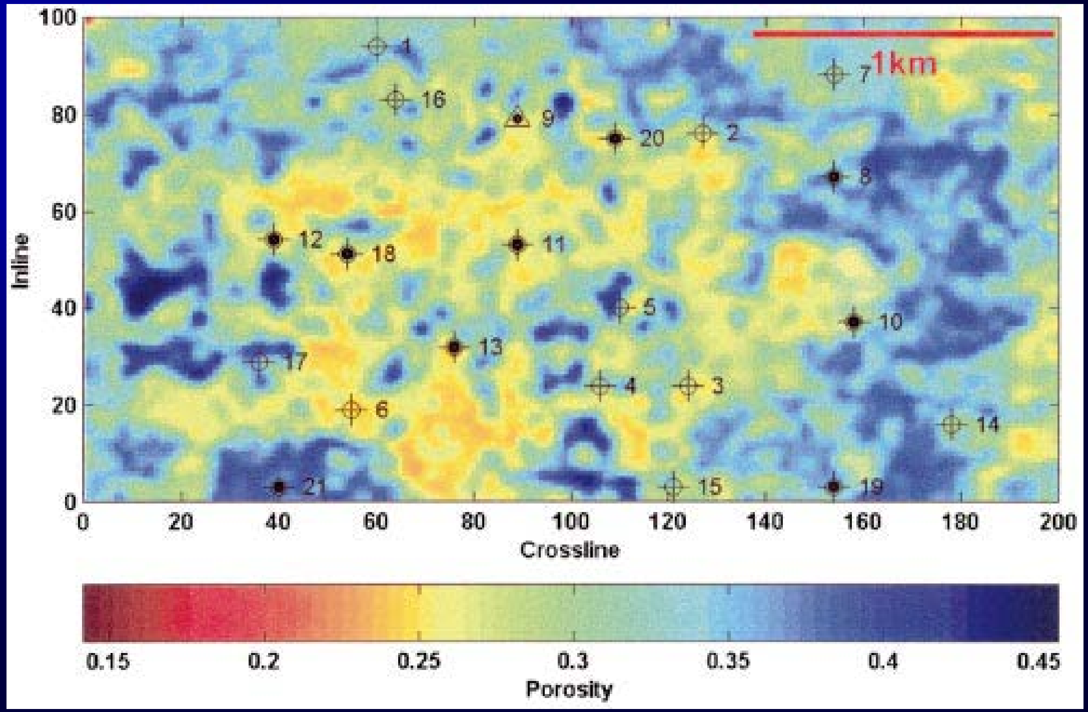
Genetic-algorithm/neural-network approach to seismic attribute selection for well-log prediction

Kevin P. Dorrington* and Curtis A. Link‡

2004



Correlation of attributes to porosity using a genetic algorithm



MLFN prediction of porosity

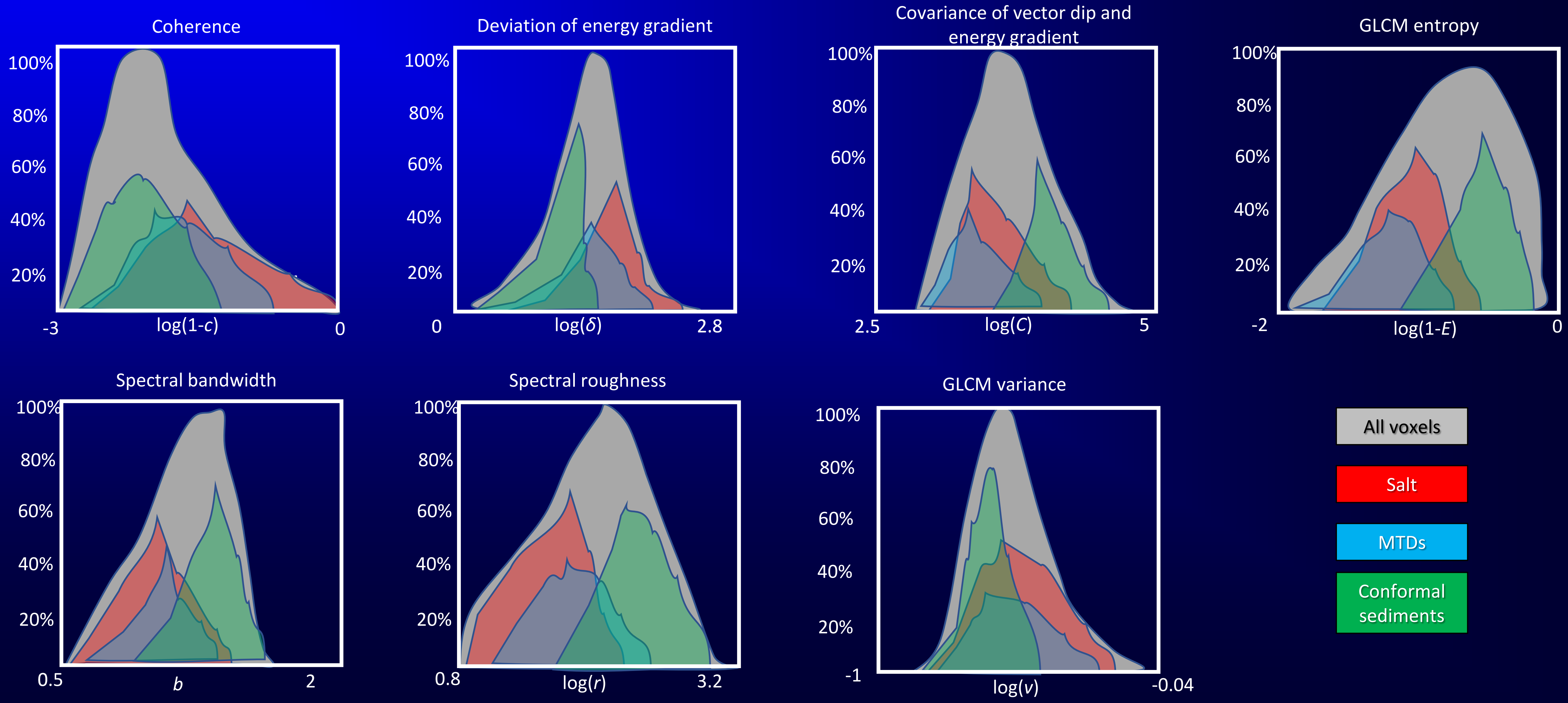
(Dorrington and Link, 2004)

Three recent workflows, all of which

- are applied to Gulf of Mexico surveys with the objective of differentiating salt, mass transport deposits, and conformal sediments
- began with a list of interpreter-supplied candidate attributes
- benefited by converting each attribute histogram to exhibit approximately Gaussian statistics
- used an exhaustive search of the best attribute combination as well as the number of attributes used

Histogram of the selected attributes

(Initialization is facilitated if the input data exhibit Gaussian distributions)



(Qi et al., 2019)

Workflow 1:

Semi-supervised classification based on generative topographic mapping

Supervision occurs in three ways:

1. Choose attributes that best separate the target facies of interest
2. Select training data biased towards the target facies of interest

Construct a manifold that fits the training data using an “unsupervised” GTM algorithm

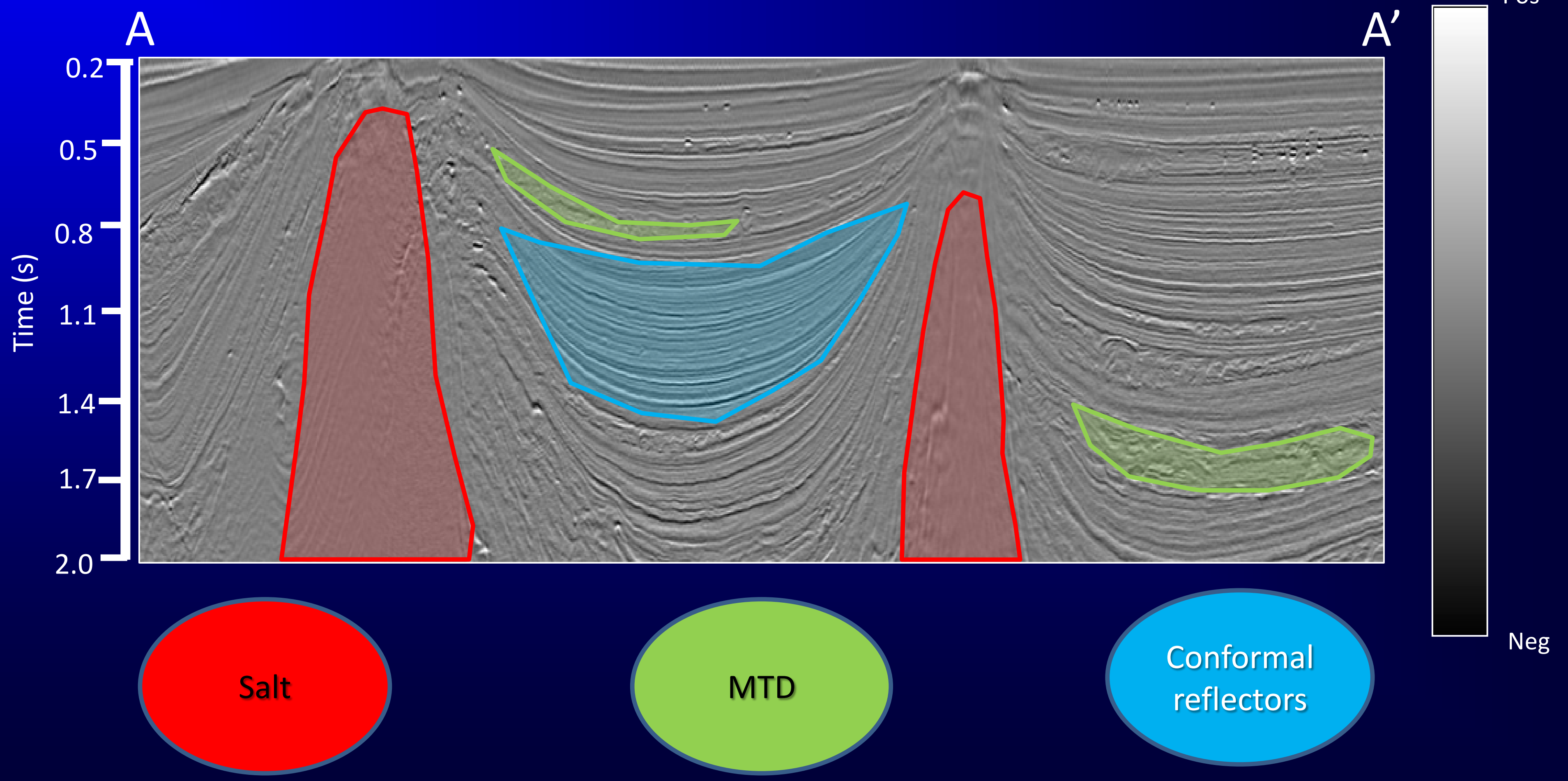
Project all data onto the manifold and corresponding latent space

3. Generate a posteriori classification based on the Bhattacharya distance of each training facies and each voxel

(Qi et al., 2019)

Step 1: Label the target features of interest on key lines and time slices

Seismic amplitude



(Qi et al., 2019)

Step 2: Use experience or geological insight to pick a suite of candidate attributes

Candidate attributes:

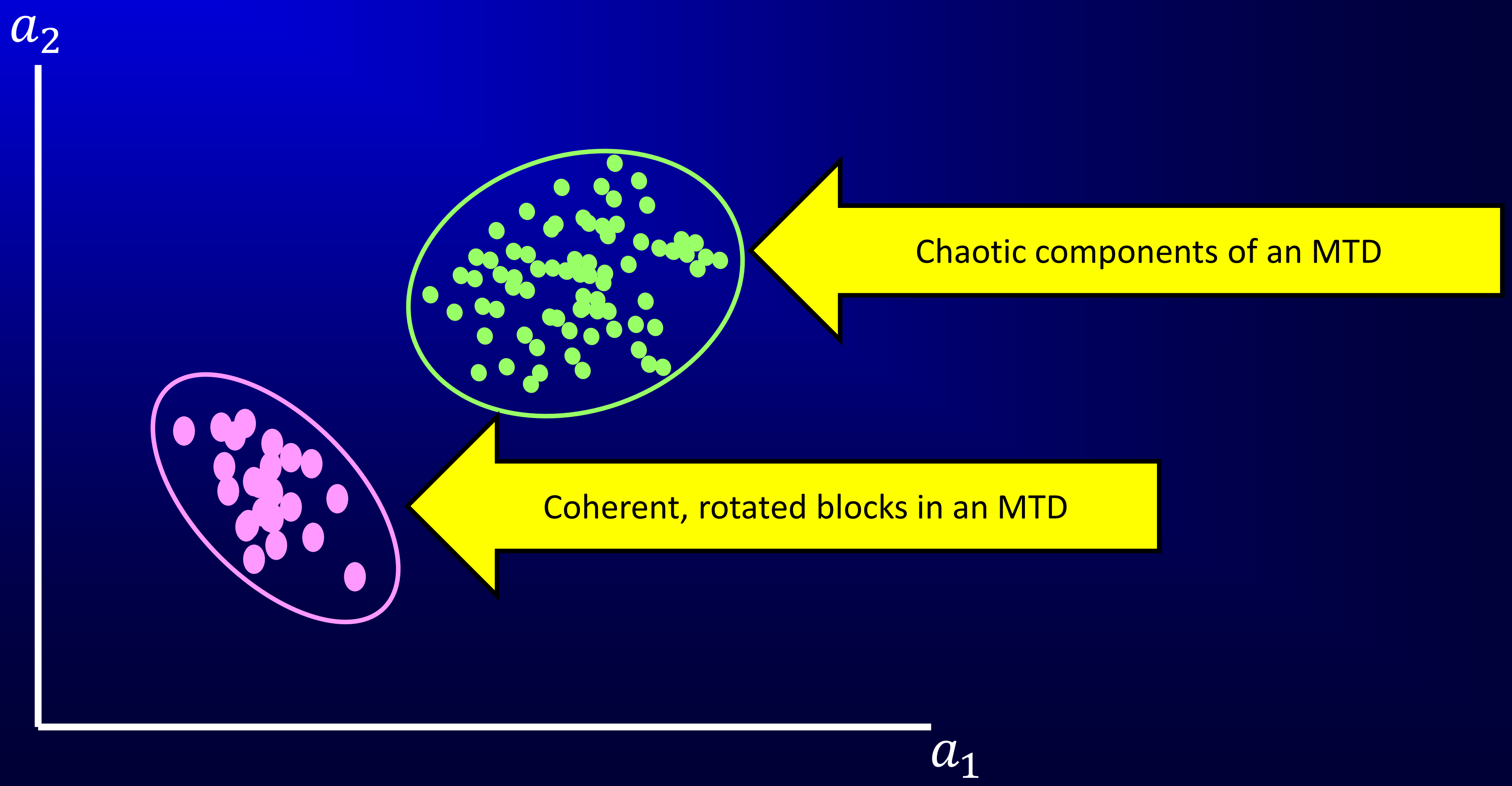
1. Coherence
2. Spectral bandwidth
3. Covariance of dip and energy
4. GLCM entropy
5. GLCM variance
6. Energy deviation
7. Spectral roughness
8. Reflector convergence
9. Dip deviation

Subset of candidate attributes	Salt	Mass transport deposit (MTD)	Conformal sedimentary background
Reflector convergence	Chaotic, nonparallel	Slumps are chaotic, rotated blocks subparallel	Parallel to subparallel
Reflector amplitude	Low reflectivity (except for noise)	Low to high based on impedance contrast within related blocks, generally low outside well imaged blocks	Low to high based on impedance contrast
Spectral response	Low to moderate spectral response	Variable frequency different in each block	Broadband response
Texture	High entropy, low homogeneity	High entropy, but may exhibit high homogeneity within a rotated block	Low entropy, high homogeneity

(Qi et al., 2019)

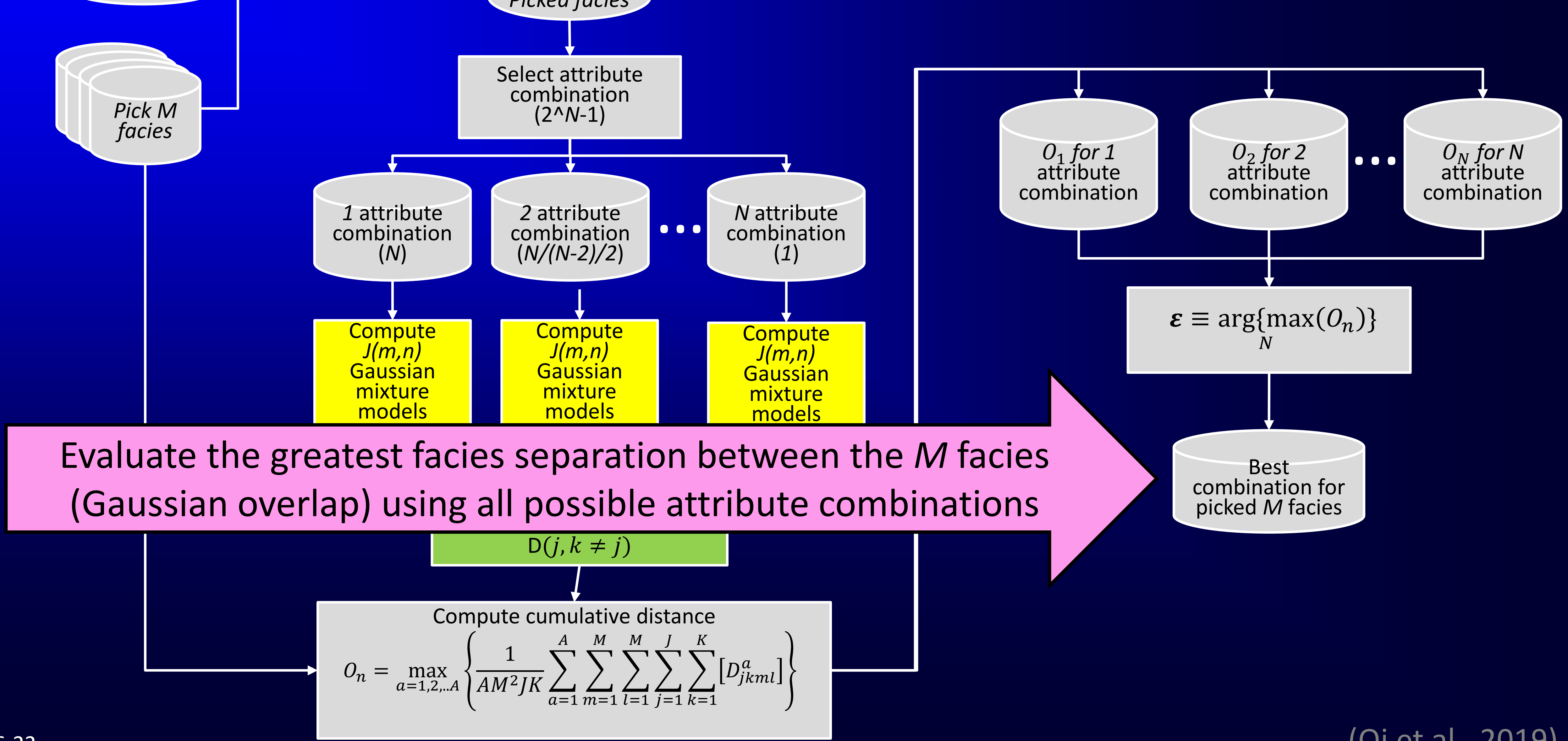
Step 3: Generate a Gaussian Mixture Model for each facies for each of the 2^N-1 attribute combinations

A given facies can be represented by one or more Gaussian distributions



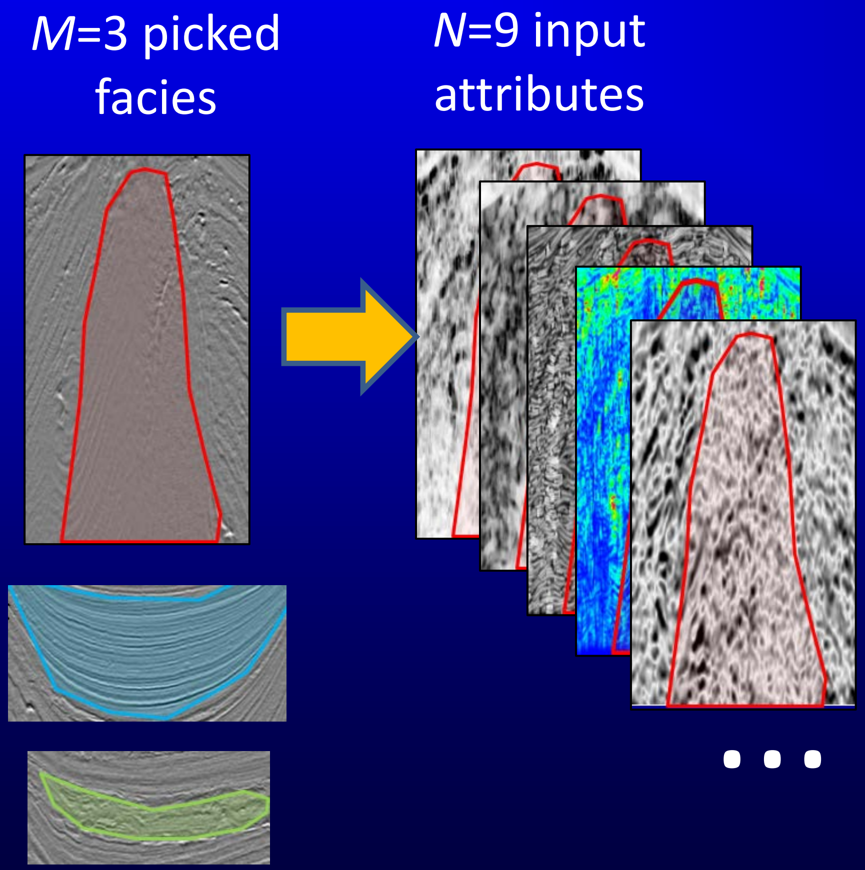
(Qi et al., 2019)

Using GMMs to select appropriate attributes

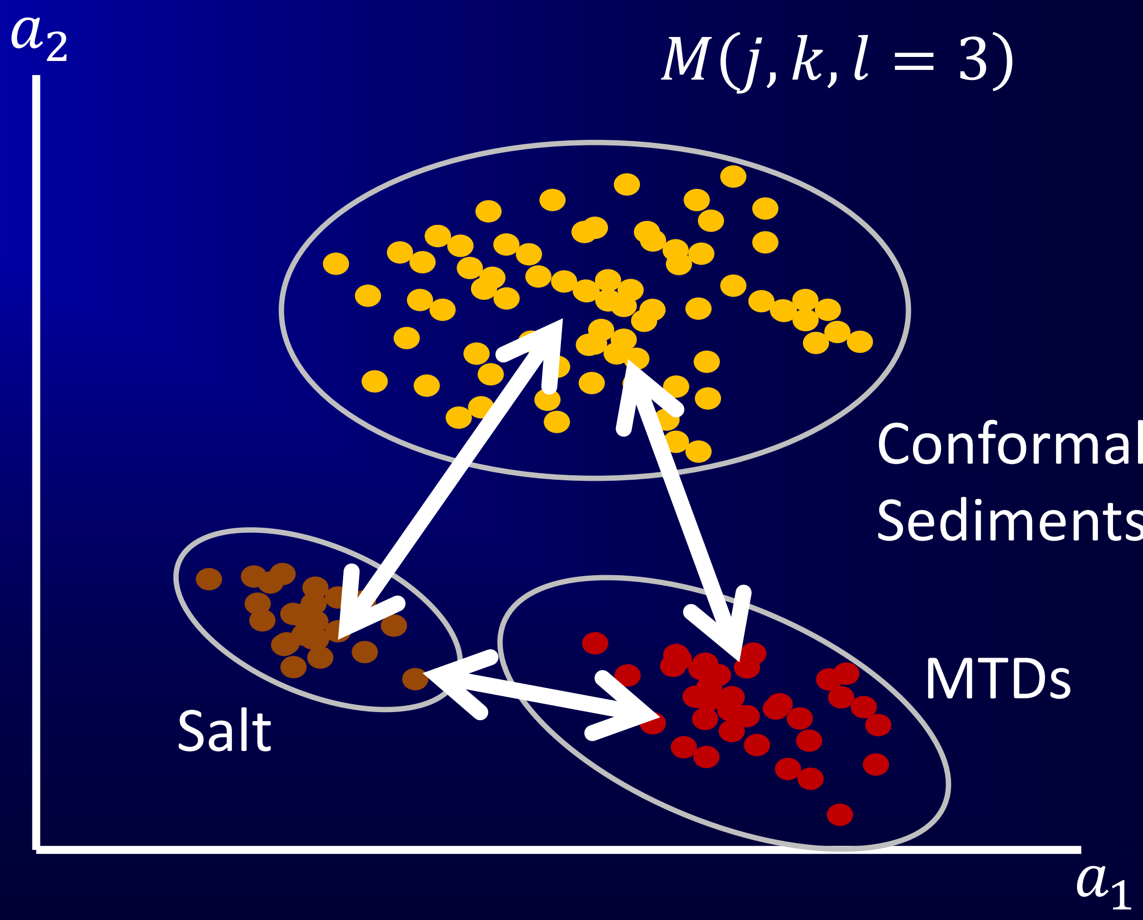


(Qi et al., 2019)

4. Choose attributes that provide the greatest cumulative Bhattacharyya distance between the Gaussian Mixture Models

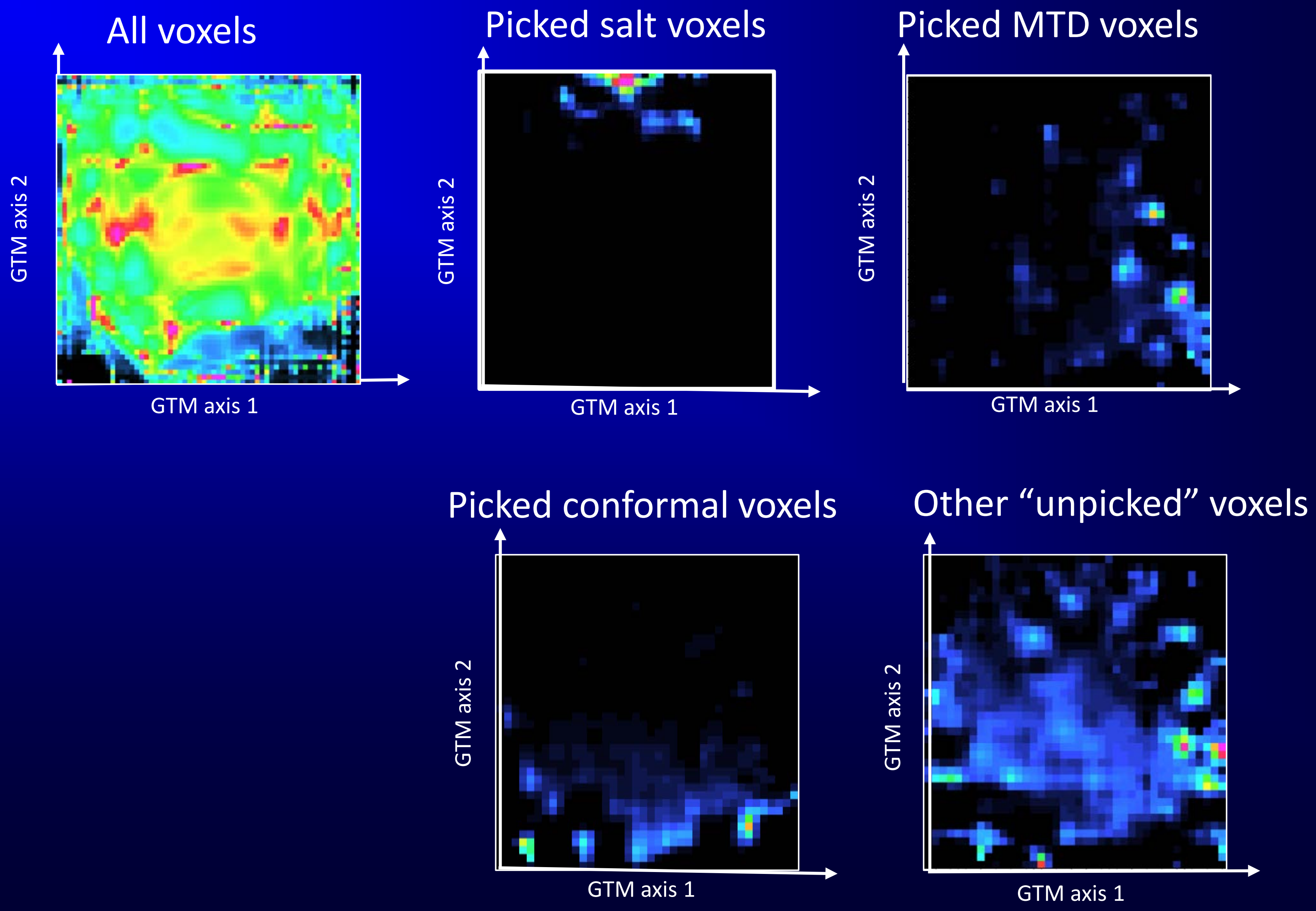


3 facies, each represented by one or more Gaussians



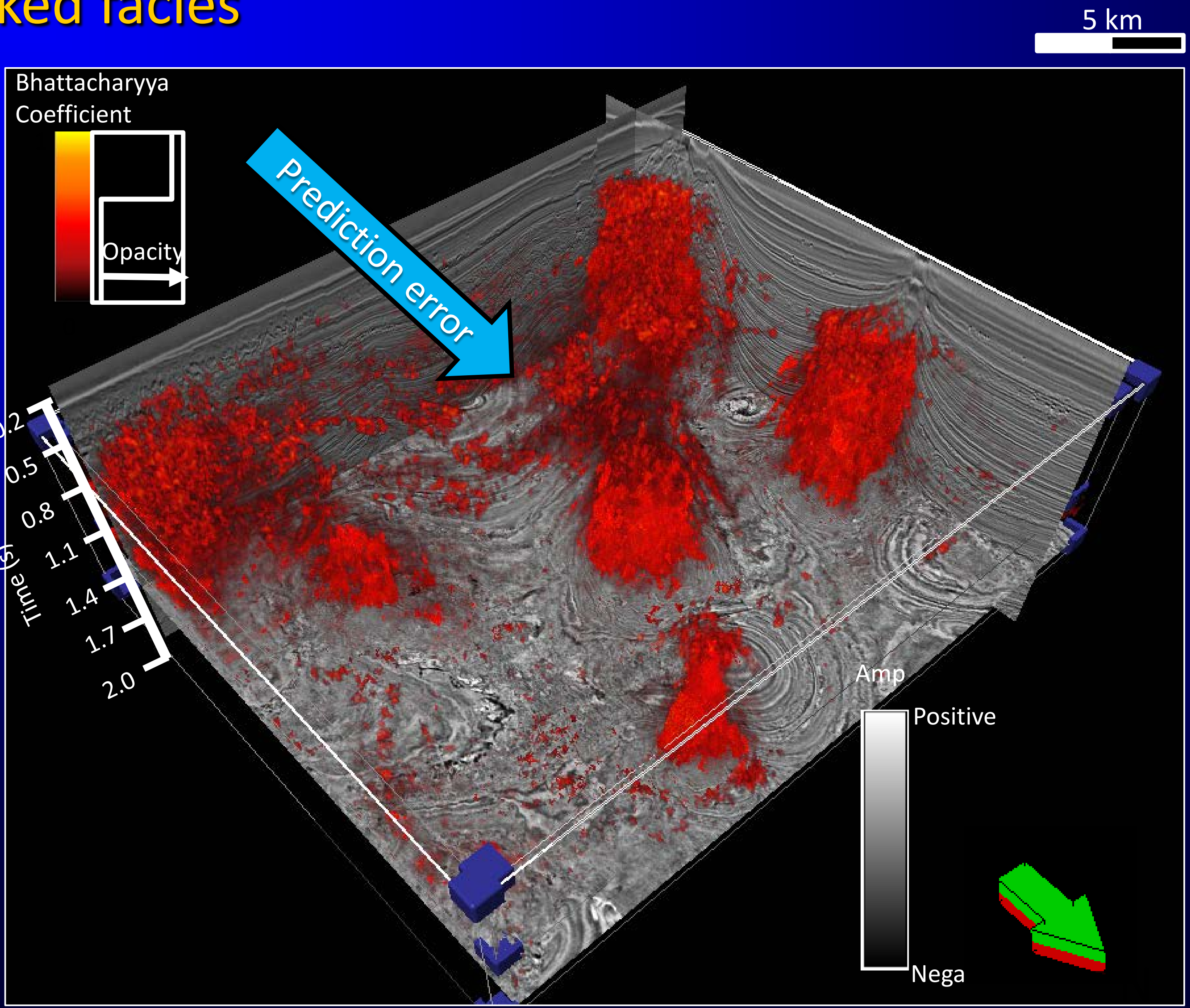
(Qi et al., 2019)

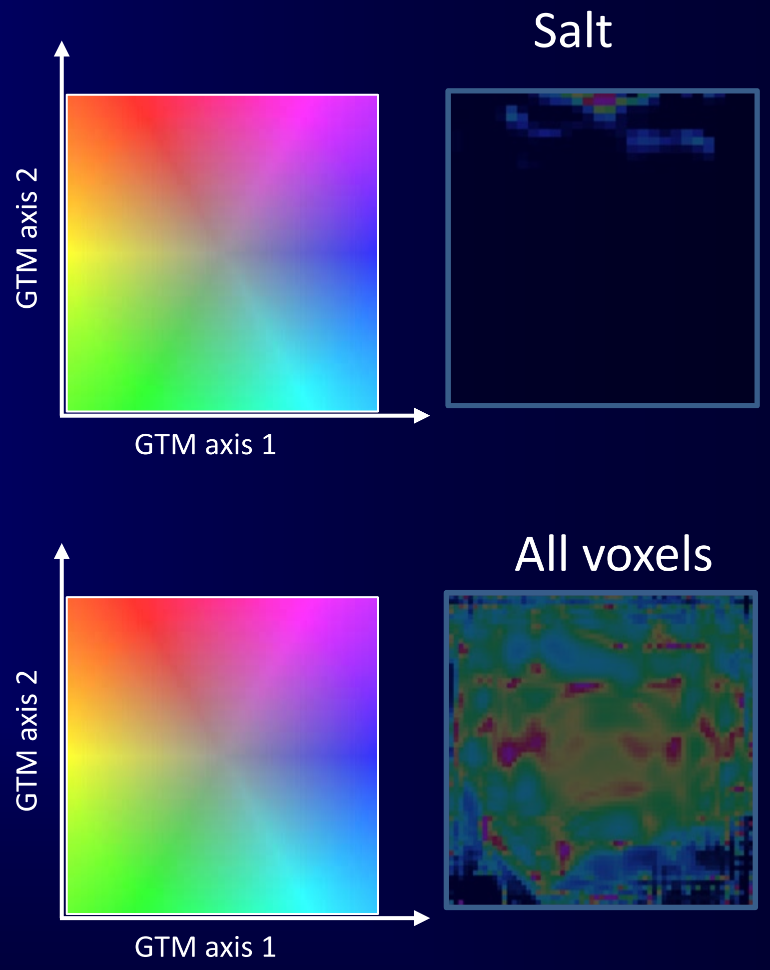
Step 5: Project all voxels onto the manifold and latent space



(Qi et al., 2019)

Step 6: Compute Bhattacharyya coefficient of each voxel with the projection of each picked facies

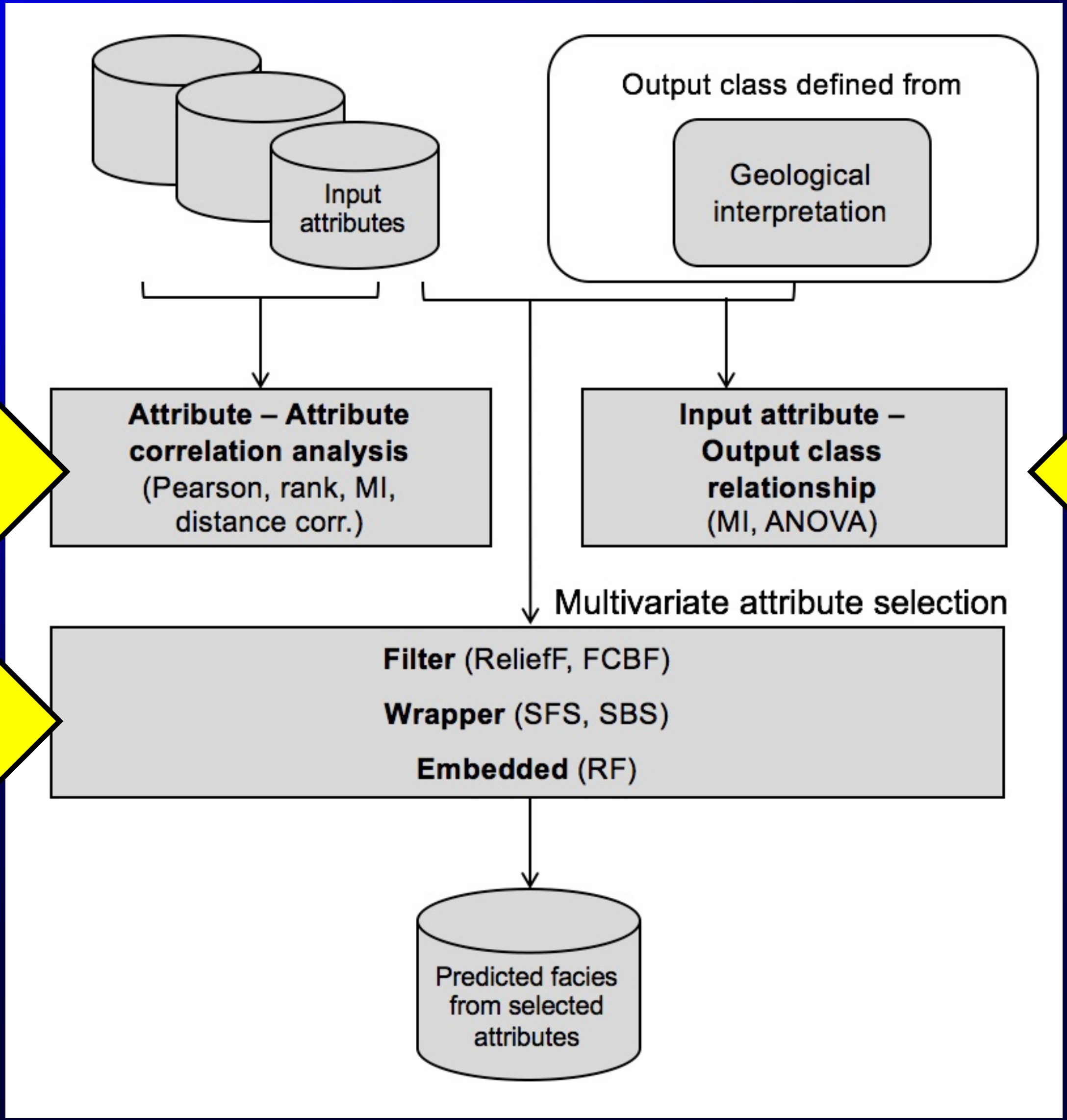




(Qi et al., 2019)

Workflow 2:
Supervised classification based on a random forest decision tree

Finding the relation between two continuous variables



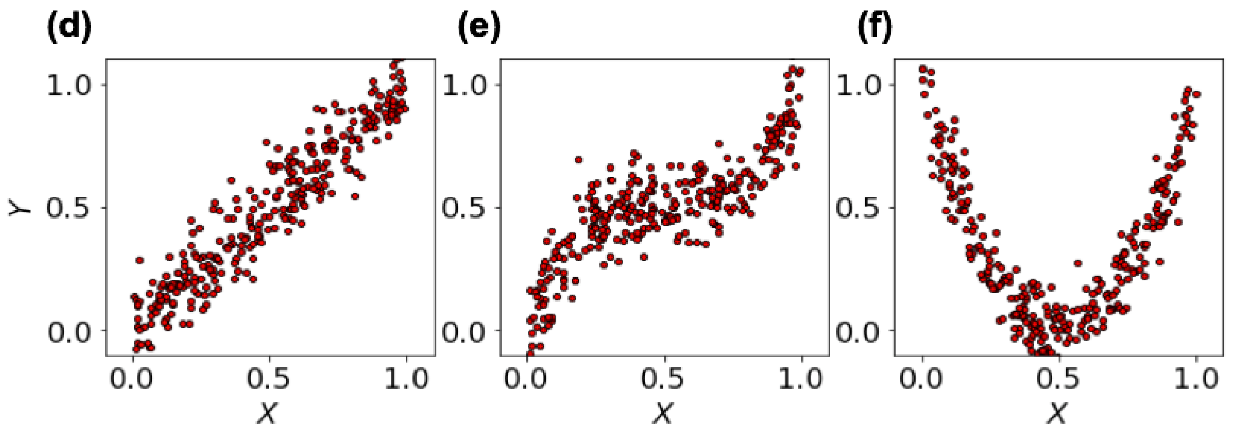
Finding the relation between a continuous and categorical variables

Determining which attribute selection workflow to use

(Kim et al., 2019)

Different types of relations between two continuous variables

Linear Monotonic | Nonlinear Monotonic | Nonlinear Nonmonotonic



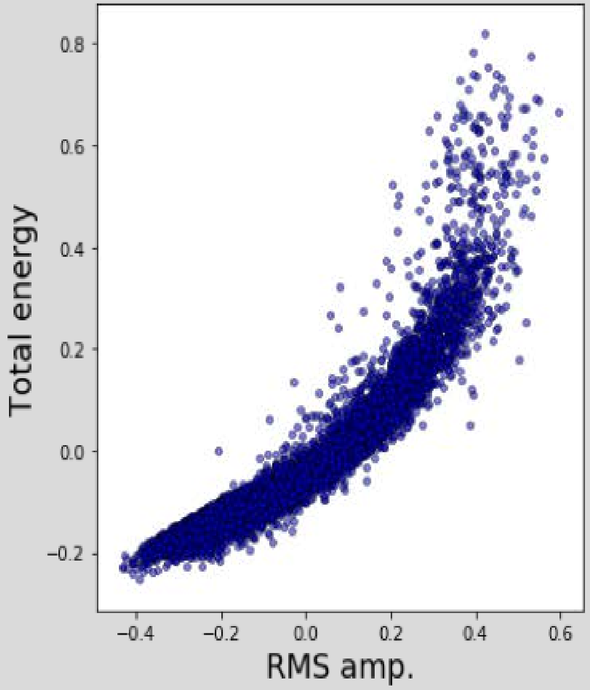
Correlation between X and Y	Pearson corr.=0.95 Rank corr.=0.95 MI=0.89 Distance corr.=0.88	Pearson corr.=0.84 Rank corr.=0.83 MI=0.62 Distance corr.=0.65	Pearson corr.=0.02 Rank corr.=-0.01 MI=0.87 Distance corr.=0.23
Regression score between X and Y	Linear Bayesian = 0.89 DNN = 0.89 RF = 0.84 SVM = 0.89	Linear Bayesian = 0.68 DNN = 0.81 RF = 0.74 SVM = 0.81	Linear Bayesian = -0.02 DNN = 0.90 RF = 0.86 SVM = 0.90

Mutual information

(Kim et al., 2019)

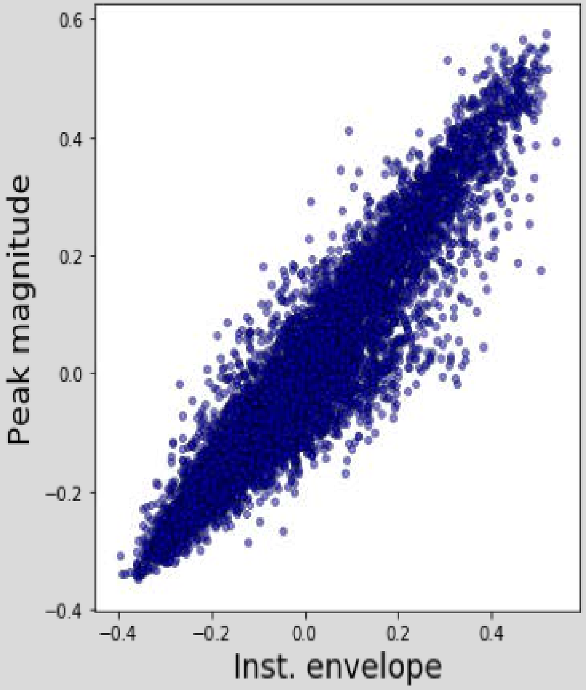
Representative relations between different attribute pairs

Highly correlated



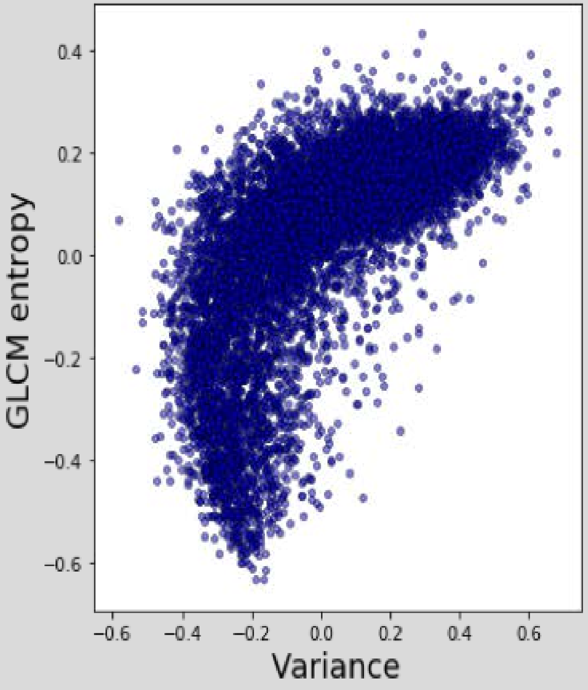
Pearson corr.=0.92
Rank corr.=0.97
MI=0.88
distance corr.=0.91

Highly correlated



Pearson corr.=0.95
Rank corr.=0.95
MI=0.75
distance corr.=0.87

Somewhat correlated



Pearson corr.=0.68
Rank corr.=0.75
MI=0.46
distance corr.=0.50

(Kim et al., 2019)

Attribute-to-attribute correlation analysis

Correlation measures	Attributes highly correlated with the other attributes (corr. coeff. > 0.6)
Pearson correlation	GLCM entropy – GLCM homogeneity (-1.0)
	Instantaneous envelope – Peak magnitude (0.96)
	RMS amplitude - Instantaneous envelope (0.93)
	RMS amplitude – Peak magnitude (0.90)
	RMS amplitude – Total energy (0.90)
	Total energy – Instantaneous envelope (0.84)
	Total energy – Peak magnitude (0.83)
	Instantaneous phase – Relative acoustic impedance (0.73)
	GLCM entropy – Chaos (0.71)
	GLCM entropy – Variance (0.70)
	Instantaneous frequency – Peak frequency (0.62)
Rank correlation	GLCM entropy – GLCM homogeneity (-1.0)
	RMS amplitude – Total energy (0.99)
	Instantaneous envelope – Peak magnitude (0.95)
	RMS amplitude – Instantaneous envelope (0.94)
	RMS amplitude – Total energy (0.93)
	RMS amplitude – Peak magnitude (0.92)
	Total energy – Peak magnitude (0.82)
	Instantaneous phase – Relative acoustic impedance (0.81)
	GLCM entropy – Variance (0.80)
	GLCM entropy – Chaos (0.71)
	Instantaneous frequency – Peak frequency (0.64)
	Instantaneous envelope – GLCM homogeneity (0.62)

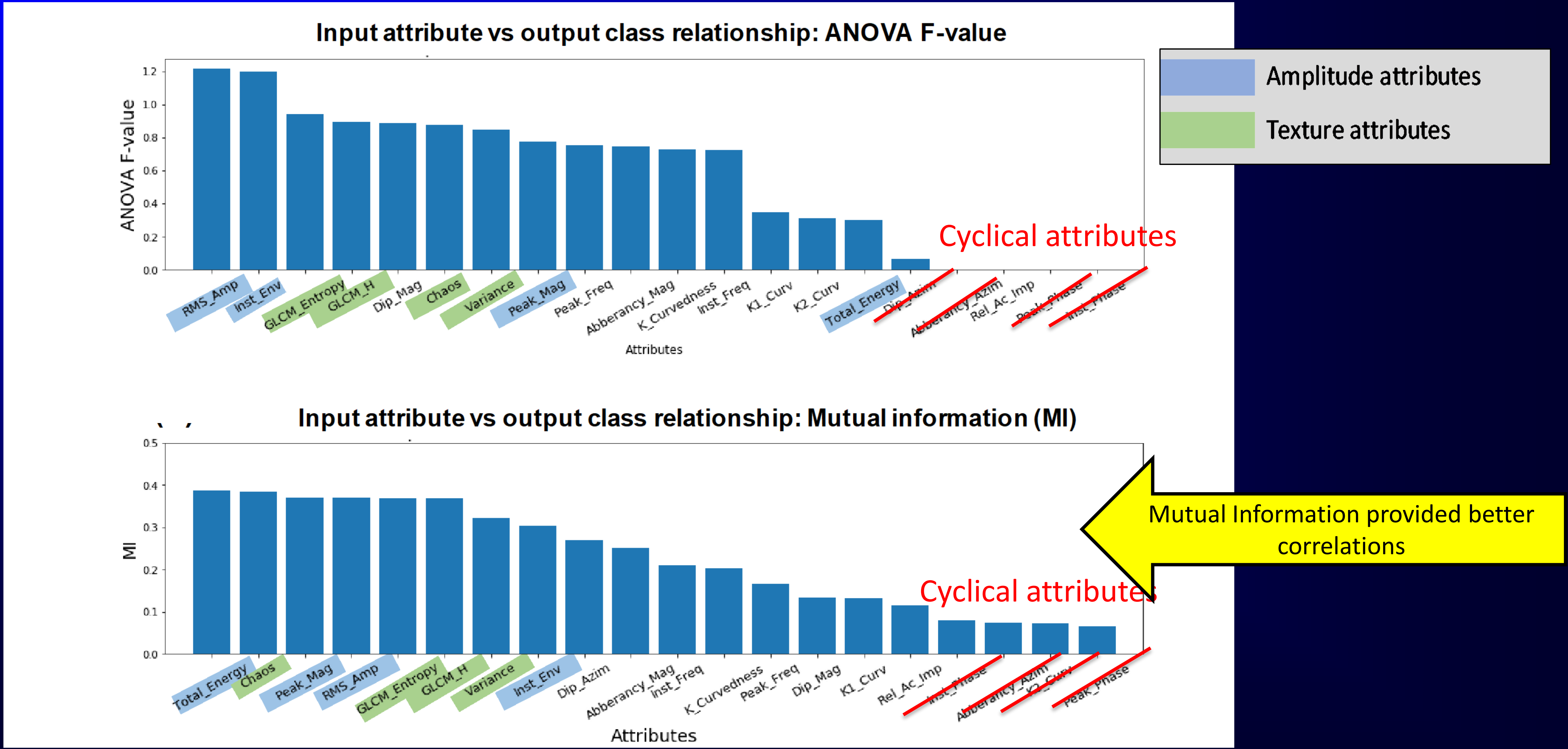
Mutual information	RMS amplitude – Total energy (0.9)
	GLCM entropy – GLCM homogeneity (0.85)
	Instantaneous envelope - Peak magnitude (0.74)
	RMS amplitude – Instantaneous envelope (0.72)
	Total energy – Instantaneous envelope (0.71)
	RMS amplitude – Peak magnitude (0.69)
	Total energy – Peak magnitude (0.68)
Distance correlation	GLCM entropy – GLCM homogeneity (0.97)
	RMS amplitude – Total energy (0.92)
	Instantaneous envelope – Peak magnitude (0.87)
	RMS amplitude – Instantaneous envelope (0.83)
	RMS amplitude – Peak magnitude (0.78)
	Total energy – Instantaneous envelope (0.78)
	Total energy – Peak magnitude (0.74)

Redundant amplitude attributes

Redundant texture attributes

(Kim et al., 2019)

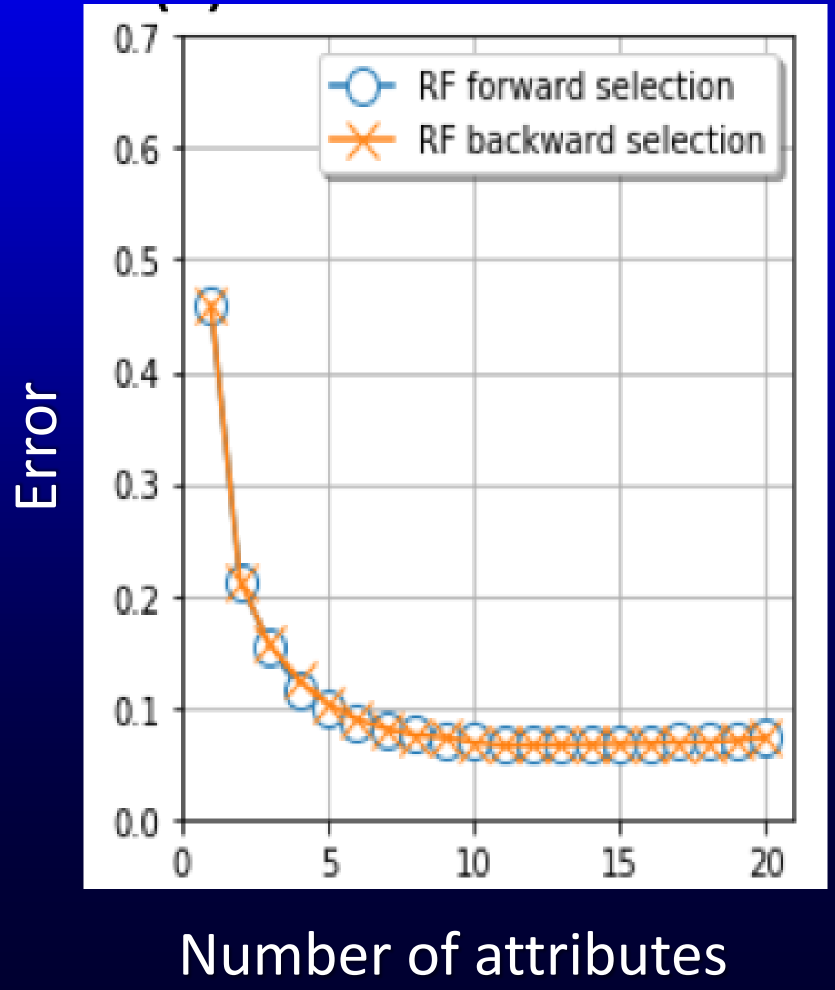
Attribute – class correlation analysis



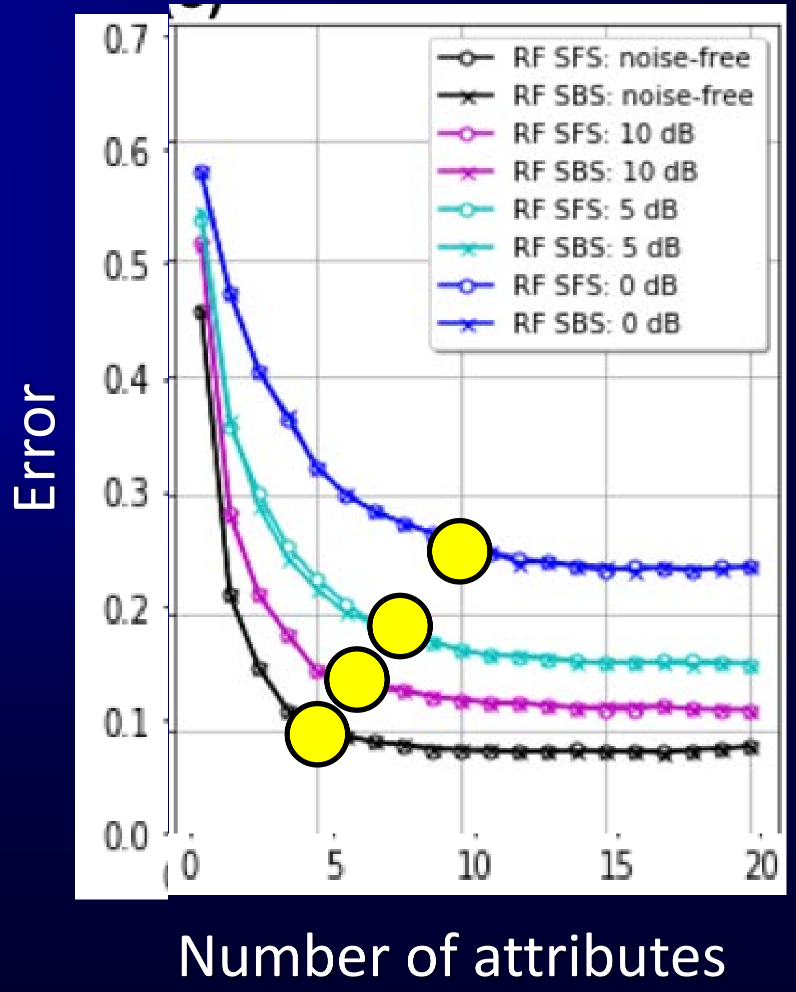
(Kim et al., 2019)

Error rate vs. the number of attributes selected for the Wrapper
(Random Forest Decision Tree Algorithm)

Noise-free attributes



Noisy attributes (Gaussian noise added)

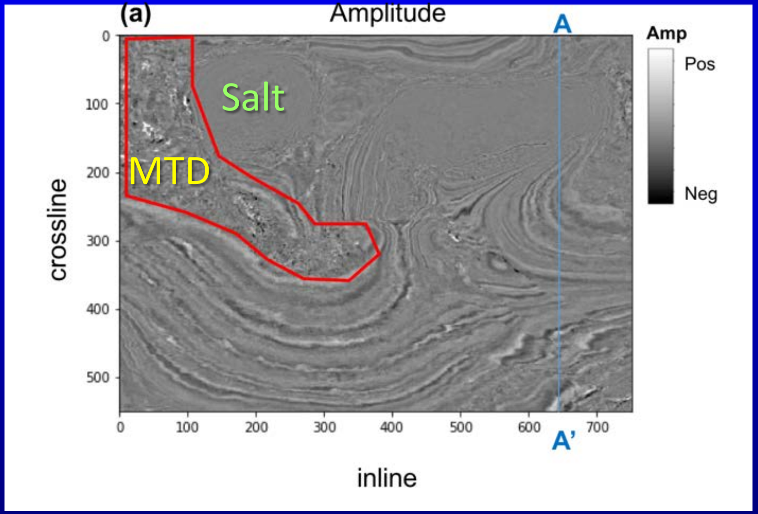


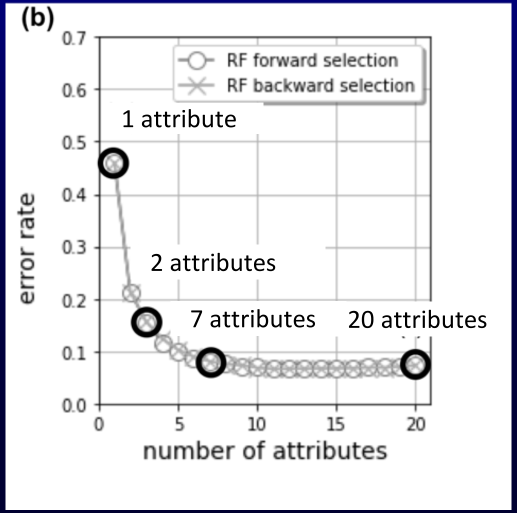
(Kim et al., 2019)

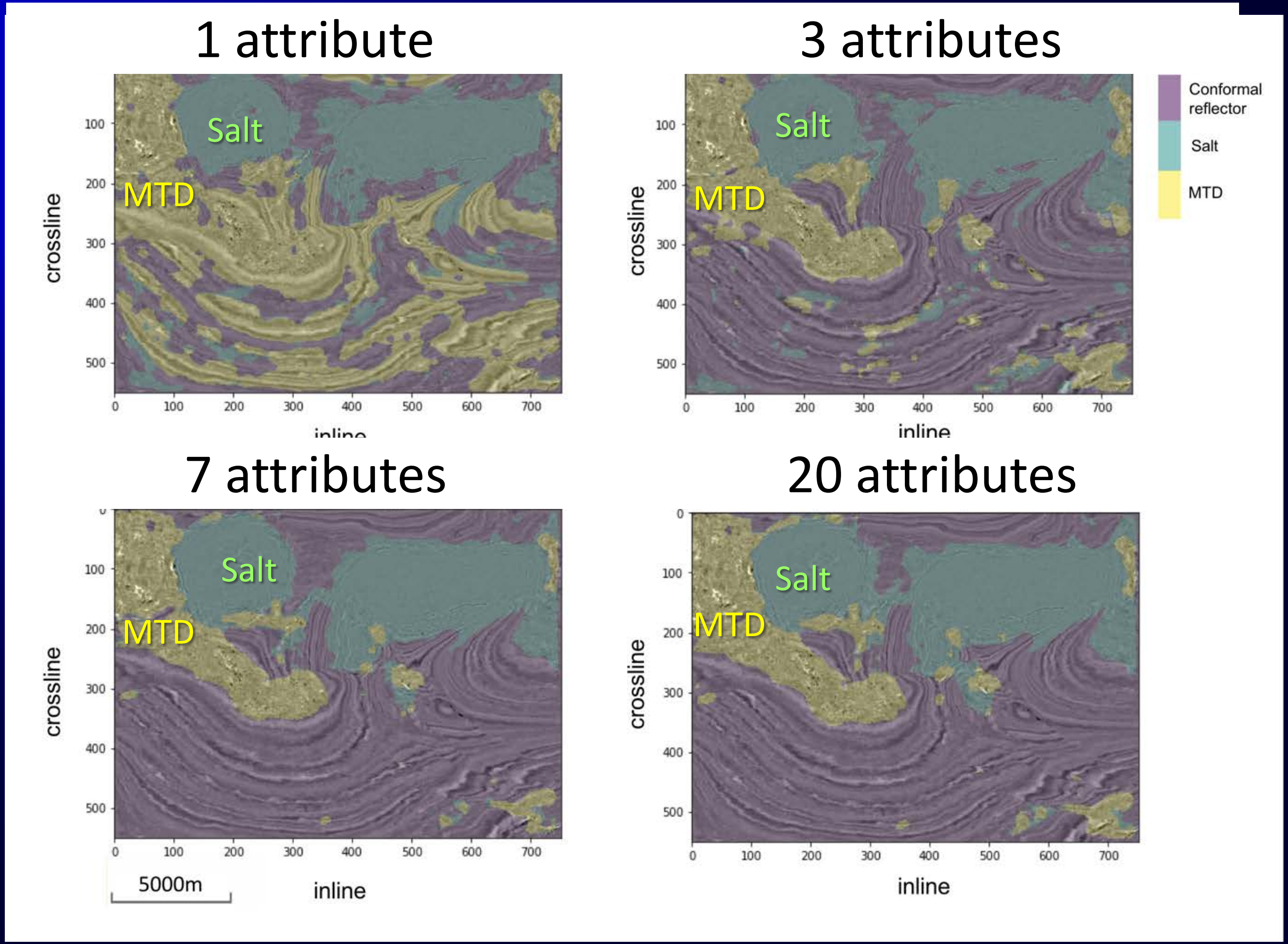
Predicted facies using subsets with different numbers of attributes

- Rank is computed using wrapper (RF) method
- Learning method for prediction: Random forest

Time slice at *t* =1.1s





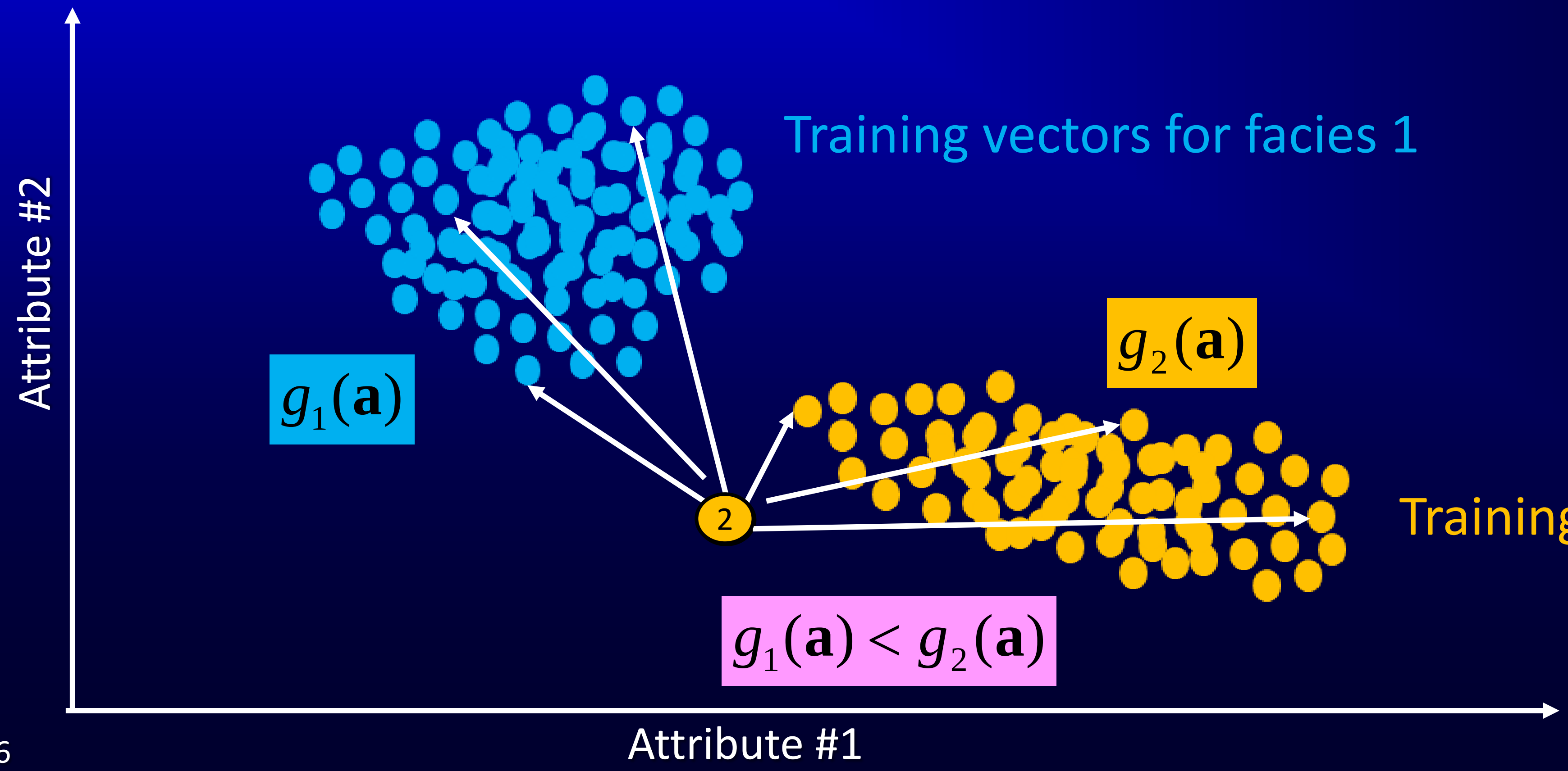


(Kim et al., 2019)

Workflow 3:
Probabilistic Neural Networks

$$g_c(\mathbf{a}) = \frac{1}{N}\sum_{n=1}^{N}\exp\left[-\sum_{m=1}^{M}\frac{(a_m - T_{nm})^2}{2r^2}\right]$$

a_m : Attribute vector at voxel to be classified
T_{nm} : Attribute training vectors
M : Number of attributes
N : Number of training vectors
r : Gaussian smoothing parameter (unknown)



(Lubo-Robles et al., 2019)

Stepwise linear regression vs exhaustive search attribute selection

Stepwise regression attribute combination	Error
a_1	E_1
a_2	E_2
a_3	E_3
a_4	E_4
$a_{1,}\ a_2$	$E_{1,2}$
$a_{1,}\ a_3$	$E_{1,3}$
$a_{1,}\ a_4$	$E_{1,4}$
a_1, a_2, a_3	$E_{1,2,3}$
a_1, a_2, a_4	$E_{1,2,4}$
a_1, a_2, a_3, a_4	$E_{1,2,3,4}$

M = 4 attributes

M+(M-1)+(M-2)+…= 10 combinations

Once attribute a_1 is chosen, it must be included in all subsequent combinations

(Lubo-Robles et al., 2019)

Stepwise linear regression vs exhaustive search attribute selection

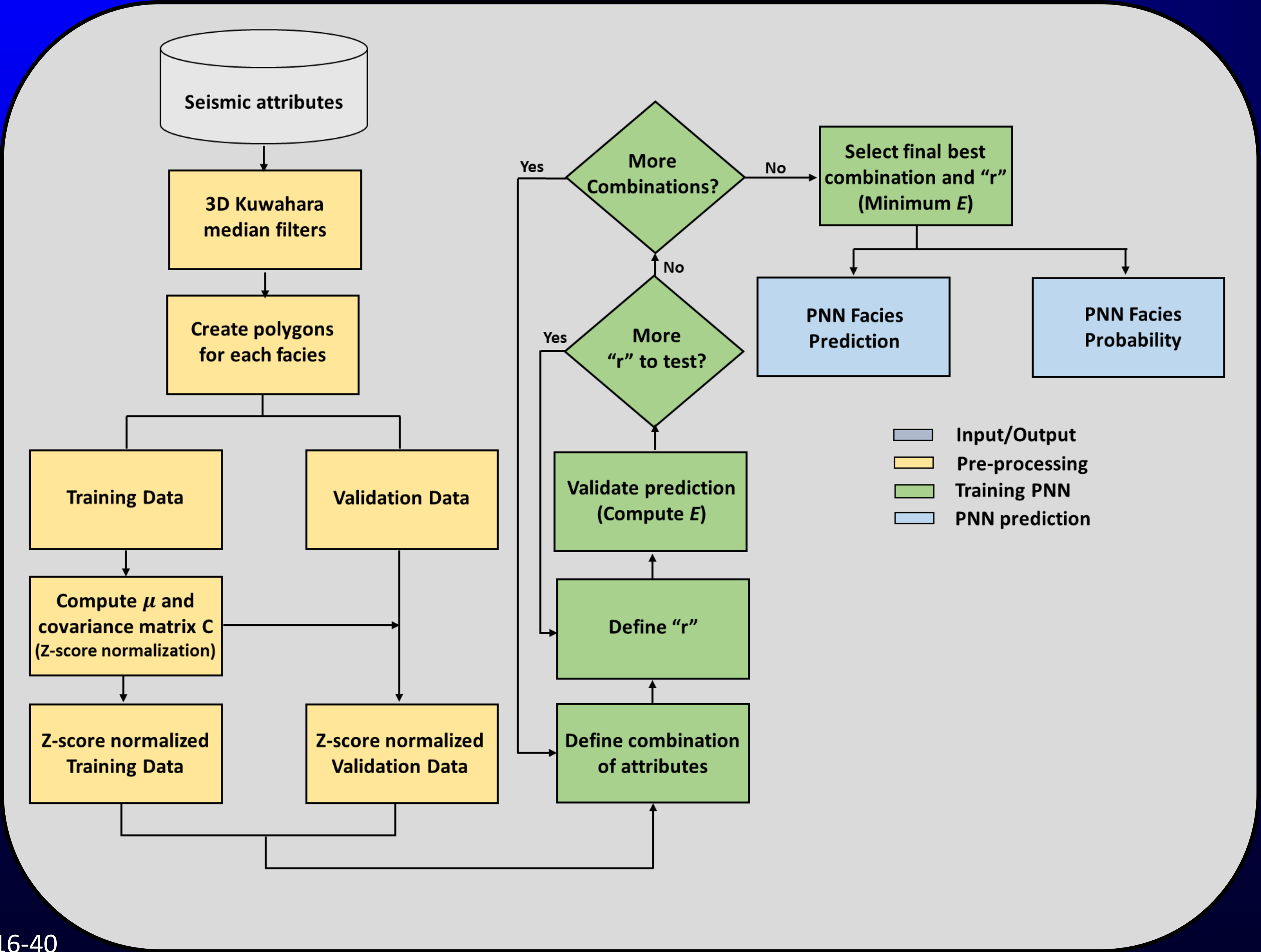
M=4 attributes

2^M-1 = 15 combinations

Exhaustive search attribute combination	Error
a_1	E_1
a_2	E_2
a_3	E_3
a_4	E_4
$a_{1,}\ a_2$	$E_{1,2}$
$a_{1,}\ a_3$	$E_{1,3}$
$a_{1,}\ a_4$	$E_{1,4}$
$a_{2\,,}a_3$	$E_{1,3}$
$a_{2\,,}a_4$	$E_{2,4}$
$a_{3\,,}a_4$	$E_{3,4}$
$a_{1,}\ a_{2\,,}a_3$	$E_{1,2,3}$
$a_{1,}\ a_{2\,,}a_4$	$E_{1,2,4}$
$a_{1,}\ a_{3,}\ a_4$	$E_{1,3,4}$
$a_{2\,,}\ a_{3\,,}a_4$	$E_{2,3,4}$
$a_{1,}\ a_{2\,,}\ a_{3\,,}a_4$	$E_{1,2,3,4}$

(Lubo-Robles et al., 2019)

Workflow 3: Probabilistic Neural Networks



Probability of a voxel being class *c* given the training data **T**

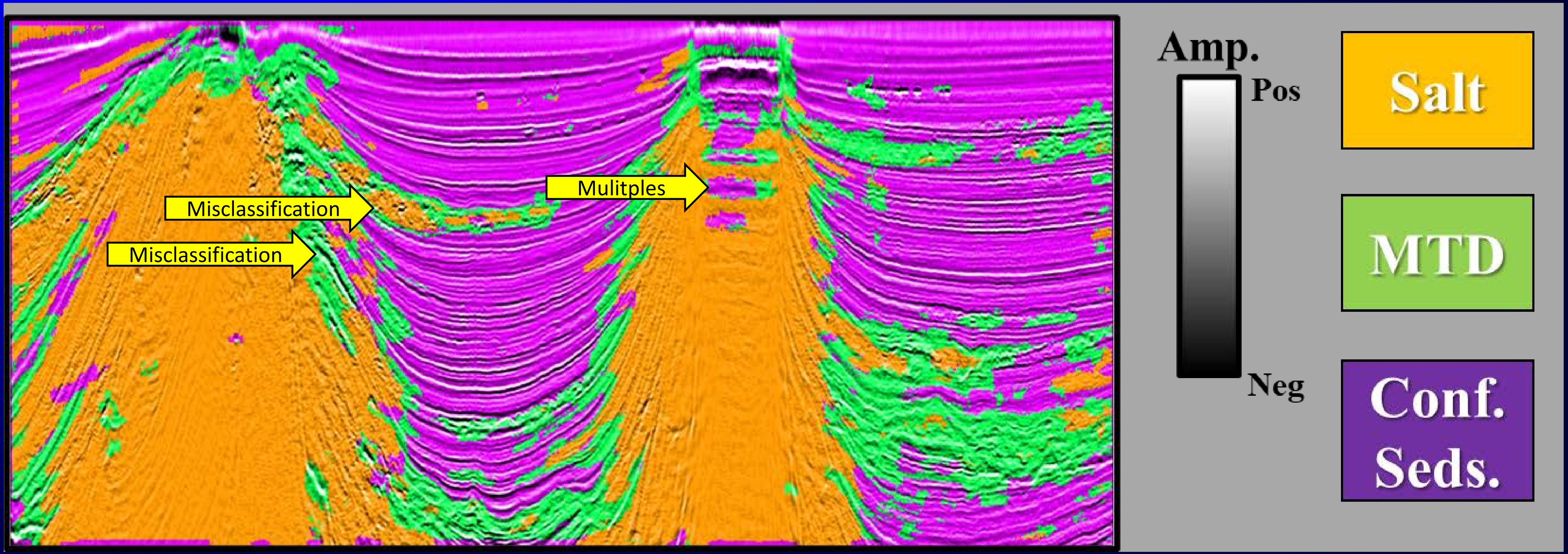
$$P(c\,|\,\mathbf{T}) = \frac{g_c(\mathbf{T})}{\sum_{j=1}^{J} g_j(\mathbf{T})}$$

6 Candidate attributes

- **Coherence**
- **GLCM contrast**
- **GLCM dissimilarity**
- **Total energy**
- **Most-positive curvature (k_1)**
- **Most-negative curvature (k_2)**

(Lubo-Robles et al., 2019)

Pitfalls and algorithm limitations



SHapley Additive exPlanations (SHAP) are recent developments that provide insight into which attributes push the classification in a given direction

(Lubo-Robles et al., 2020)

Conclusions

- Human interpreters and seismic attributes are good at identifying 2D spatial patterns; attributes can quantify patterns in 3D (and for AVAz analysis, 4D and 5D).

- Human interactive analysis is limited to about 3 volumes using RGB, CMY, or HLS color models, transparency, and/or animation. In contrast, machine learning can analyze dozens of attributes at the same time.

- Most shallow learning algorithms look for patterns occurring at the same voxel across attribute volumes; We could generalize the workflow to provide adjacent voxels as input.

- For normal amounts of training data using modern computers, an exhaustive search for the optimum number and combination of attributes is both desirable and feasible.

- I've presented three workflows for attribute selection; We do not yet know which is best for a given mapping task.

Machines vs. humans

Machines

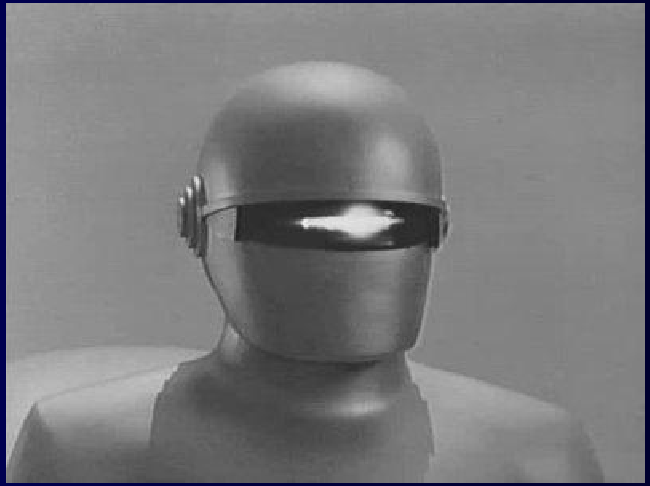
- Work 24 hours/day, 7 days/week, and require electricity and cooling
- Are consistent throughout the interpretation task
- Can analyze high dimensional data measured at each voxel
- Can quantitative evaluate the accuracy of their prediction

Humans

- Work 40-hour weeks and expect salary and benefits
- May tire or be distracted throughout the interpretation task
- Easily identify spatial patterns
- Can interpret features within the context of a geologic model

Both machines and humans

- Increase accuracy with increased experience
- Are susceptible to confirmation bias
- Will benefit by improved seismic facies databases



Some questions:

- Can using redundant attributes help with classifying noisy data?

- Is it possible to incorporate contextual information like velocity pull up or the recognition of multiples, migration artifacts, and acquisition footprint in machine learning driven facies classification?

- Can machine learning exploit constraints like the environment of deposition, tectonic style, and limits to seismic imaging routinely used by interpreters?