Information Entropy
Monte Carlo Simulation

March, 27th, 2009

Ayato Kato

Outline

- **Shannon's Information Entropy**
- **Monte Carlo Simulation**

Which attribute(s) should We Use ?

Innumerable Seismic Attributes

Type	Seismic Attribute	Major Geological significance
Conventional property	Amplitude	Lithological contrast, bedding continuity
	Interval velocity	Lithology, Porosity, Fluid Content
	Acoustic impedance	Lithology, Porosity, Fluid Content
Volume-related attribute (Multi-trace attribute)	Reflection geometry	Reservoir Architecture, Sedimentary Structure
	Trace continuity	Fault geometry, Fault distribution, Stratigraphic continuity
	Time curvature, Dip, Azimuth	Detailed Reservoir Architecture, Fault geometry, Fault distribution, Fracture density
Pre-stack attribute	AVO	Fluid Content, Lithology, Porosity
	Impedance (elastic/S-wave)	
	Poisson's ratio	
	λ, μ (Lame constant)	
	AVOZ	Fracture Orientation, Fracture Density, Fluid Content
Instantaneous attribute	Instantaneous phase	Bedding continuity
	Instantaneous Frequency	Bed thickness, lithologic contrast, fluid content
Miscellaneous attribute	Frequency	Fluid content
	Attenuation	
	Anything computed from seismic traces	???

For the purposes of predicting porosity, which attributes should we use?

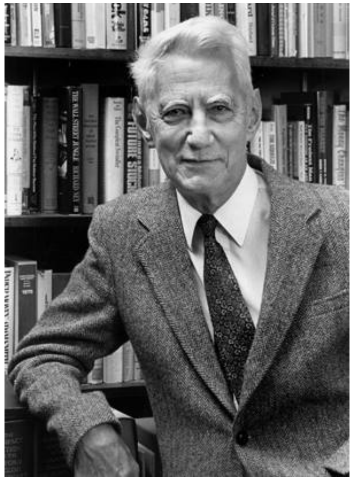
- ***AVO P_0 & G***
- ***AI & EI***
- ***$\lambda\rho$ & $\mu\rho$***
- ***Vp/Vs etc***

Shannon's Information Entropy can give us the solution quantitatively.

Shannon's Information Theory

Shannon and Weaver (1949)
"The Mathematical Theory of Communication"

Defined **Quantity of Information**



Dr. Claude Shannon

- **Information Content**
- **Information Entropy**
- **Mutual Information**

Information Content

It will snow !!

Houston

Low Probability

Very rare (1%) → Very surprised !!

(valuable information)

$I_{snow} = 6.64$

Calgary

High Probability

Very often (70%) → Not surprised

(Not-valuable information)

$I_{snow} = 0.51$

Information Content (Quantity of information)

$$\boxed{I = -\log P}$$

How surprised one would be if the event happened.

* *Base = 2*

Information Entropy

Expected Value of Information Content

Information Entropy

$$H(X) = -\sum_{i}^{n} P_i \cdot \log P_i$$

$$X = \{x_1, \cdots\cdots, x_N\}$$

- **Expected Surprise**
- Quantity of **uncertainty** associated with *P*

Houston

H=1.65

Xi	*Pi (%)*
Sunny	33
Cloudy	33
Rain	33
Snow	1

Calgary

H=1.36

Forecast is easy
Uncertainty is small

Xi	*Pi (%)*
Sunny	10
Cloudy	10
Rain	10
Snow	70

NY

H=2.00

Forecast is difficult
Uncertainty is large

Xi	*Pi (%)*
Sunny	25
Cloudy	25
Rain	25
Snow	25

Conditional Entropy

Conditional Information Entropy

$$H(X \mid A)$$

Conditional Information Entropy at a given additional information (A)

e.g. A = atmospheric pressure change data

NY

$H(X)=2.00$

Uncertainty is large

Weather	*Pi (%)*
Sunny	25
Cloudy	25
Rain	25
Snow	25

At given additional information A

$H(X|A)=1.54$

Uncertainties reduce

Weather	*Pi (%)*
Sunny	50
Cloudy	25
Rain	10
Snow	5

Mutual Information

$$I(X \mid A) = H(X) - H(X \mid A)$$

Quantity of uncertainty reduced by additional information **A**

Our Case : Porosity Prediction

Porosity Prediction by Seismic Attribute

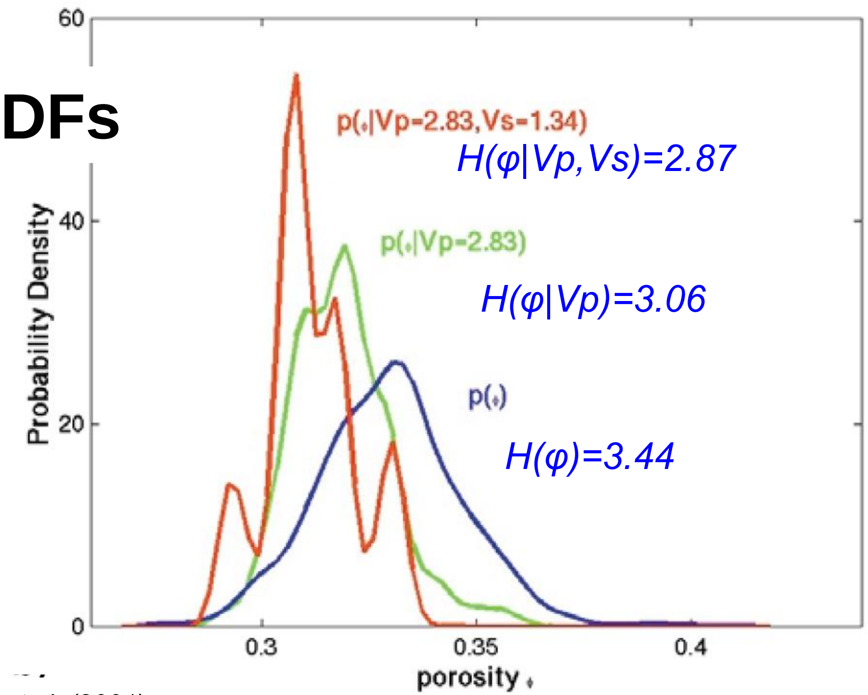
X : Porosity (continuous variable)

A : Seismic Attribute

H(X) : Information Entropy for Porosity's PDF

H(X|A) : Information Entropy for Porosity's PDF at given seismic attribute

PDFs



Adding more seismic attributes

PDF shape → Narrow, steep

Uncertainty → Decrease

In. Entropy → Decrease

Mukerji et al. (2001)

Information carried by Seismic Attributes

For example,
X = Porosity
A = Seismic Attribute

Mutual Information

$$\boxed{I(X \mid A) = H(X) - H(X \mid A)}$$

Information entropy   Conditional entropy at given A

- Mutual Information can be regarded as the reduced uncertainty by the seismic attributes.
- Thus, we should choose the one which will maximize the mutual information.

Case Studies

- North Sea Tertiary Turbiditic Reservoir

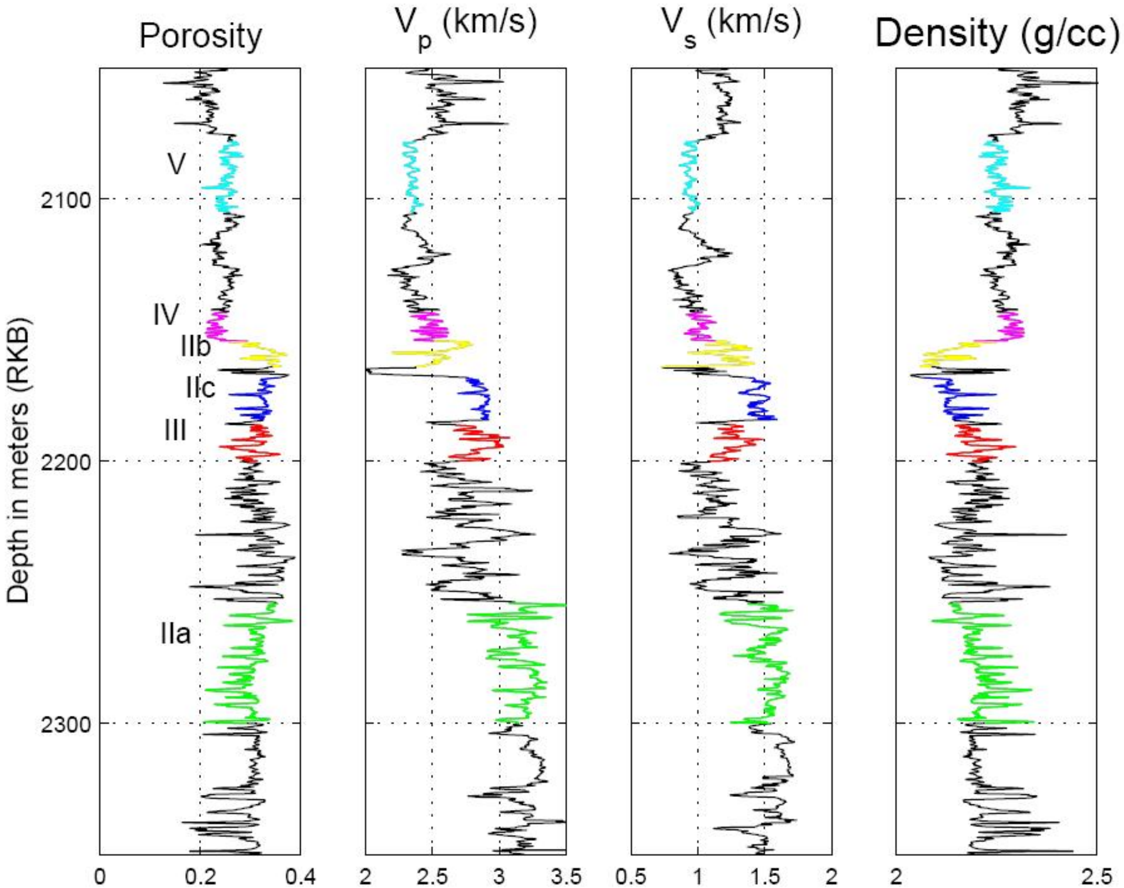
- Case I
  - Facies Identification $H\left(facies \mid attributes\right)$

- Case II
  - Pore Fluid Identification $H\left(fluid \mid attributes\right)$

- References
  - Tapan et al. (2001)
  - Takahashi et al (1999)

Well Log Data

Facies Classification in Well Log Data

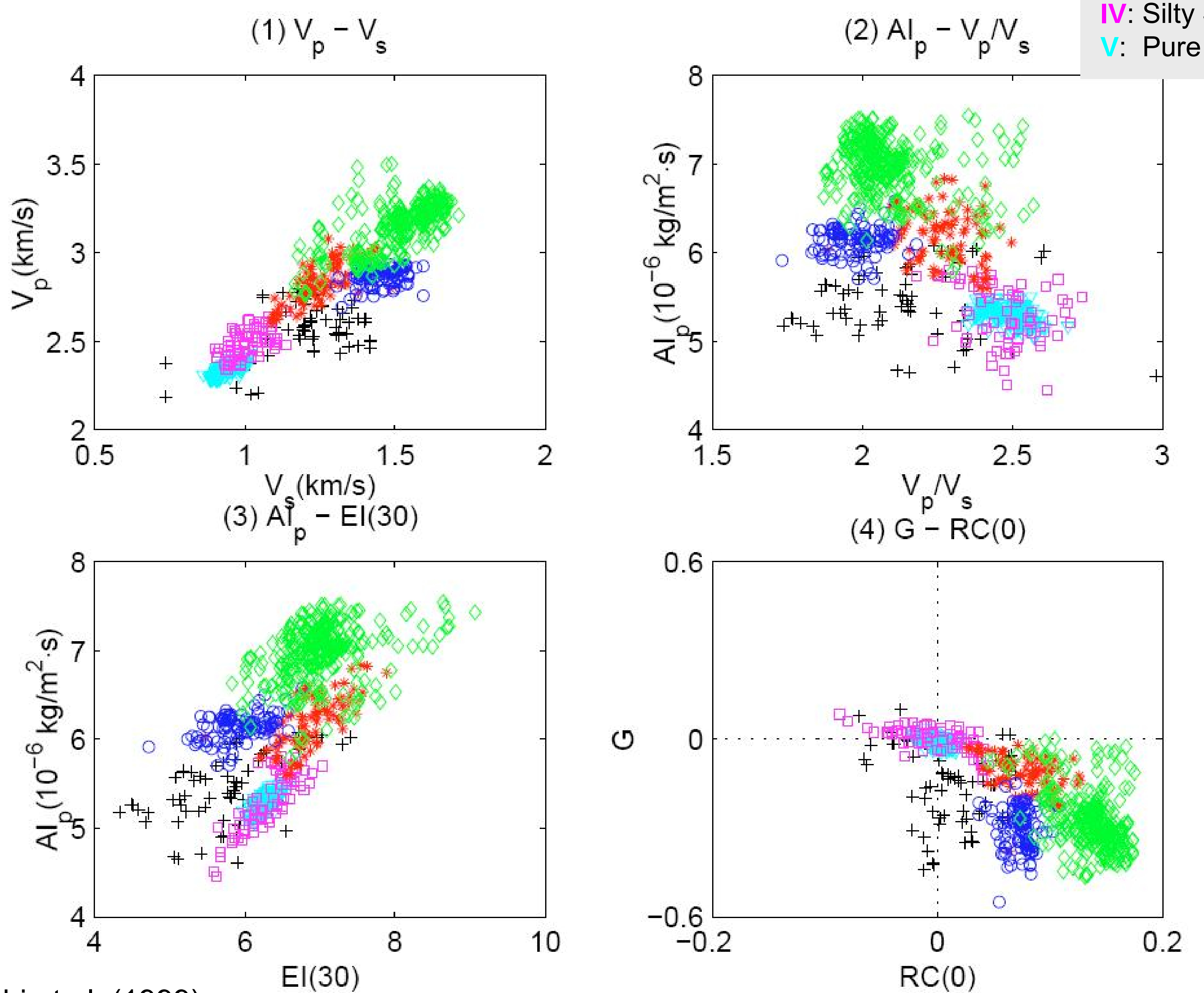


IIa: Cemented Sand
IIb: Uncololidated Sand
IIc: Laminated Sand
III: Interbedded Sand-Shale
IV: Silty Shale
V: Pure Shale

Takahashi et al. (1999)

Crossplot

IIa: Cemented Sand
IIb: Uncololidated Sand
IIc: Laminated Sand
III: Interbedded Sand-Shale
IV: Silty Shale
V: Pure Shale

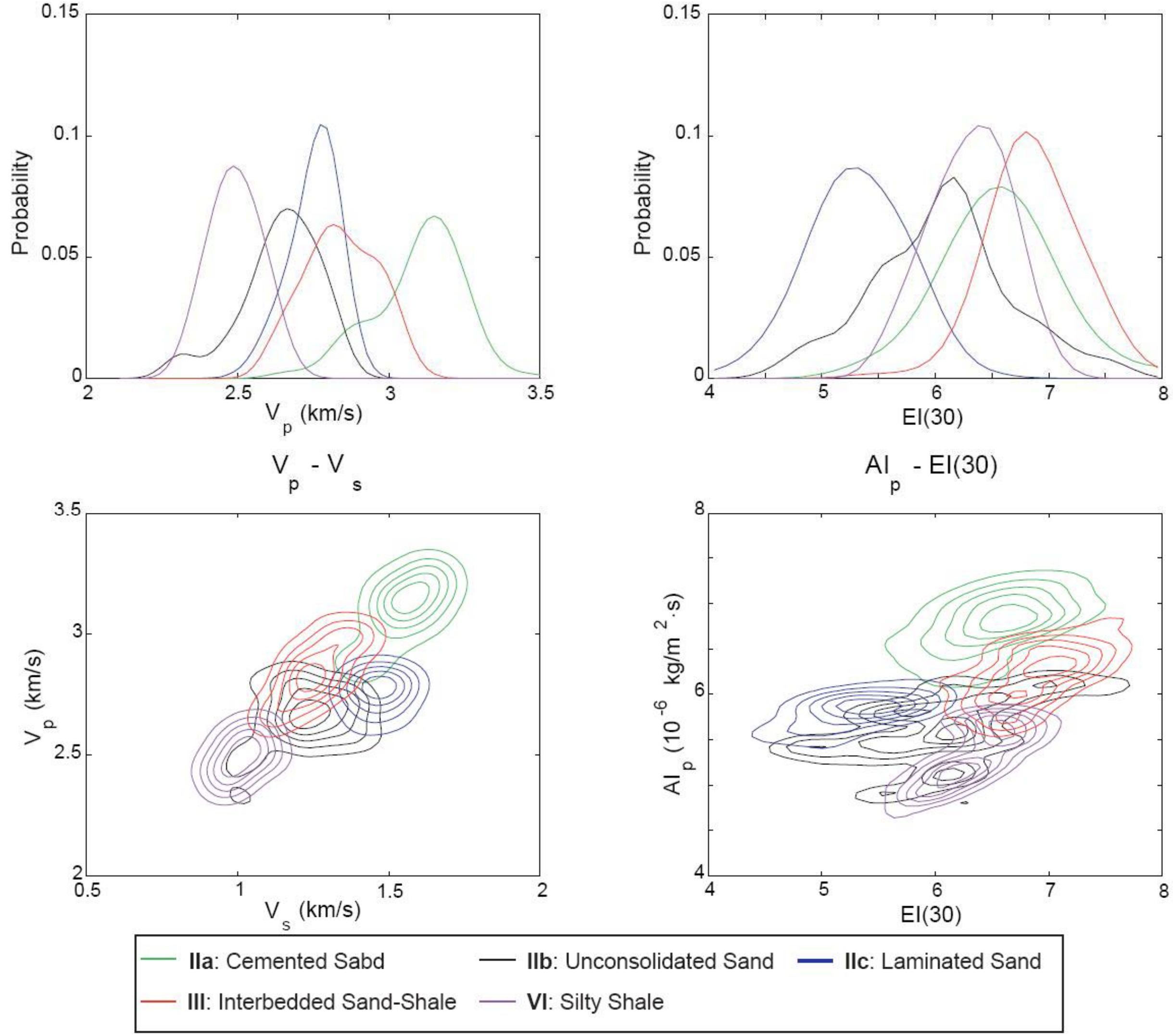


Takahashi et al. (1999)

Conditional probability distributions

Univariate

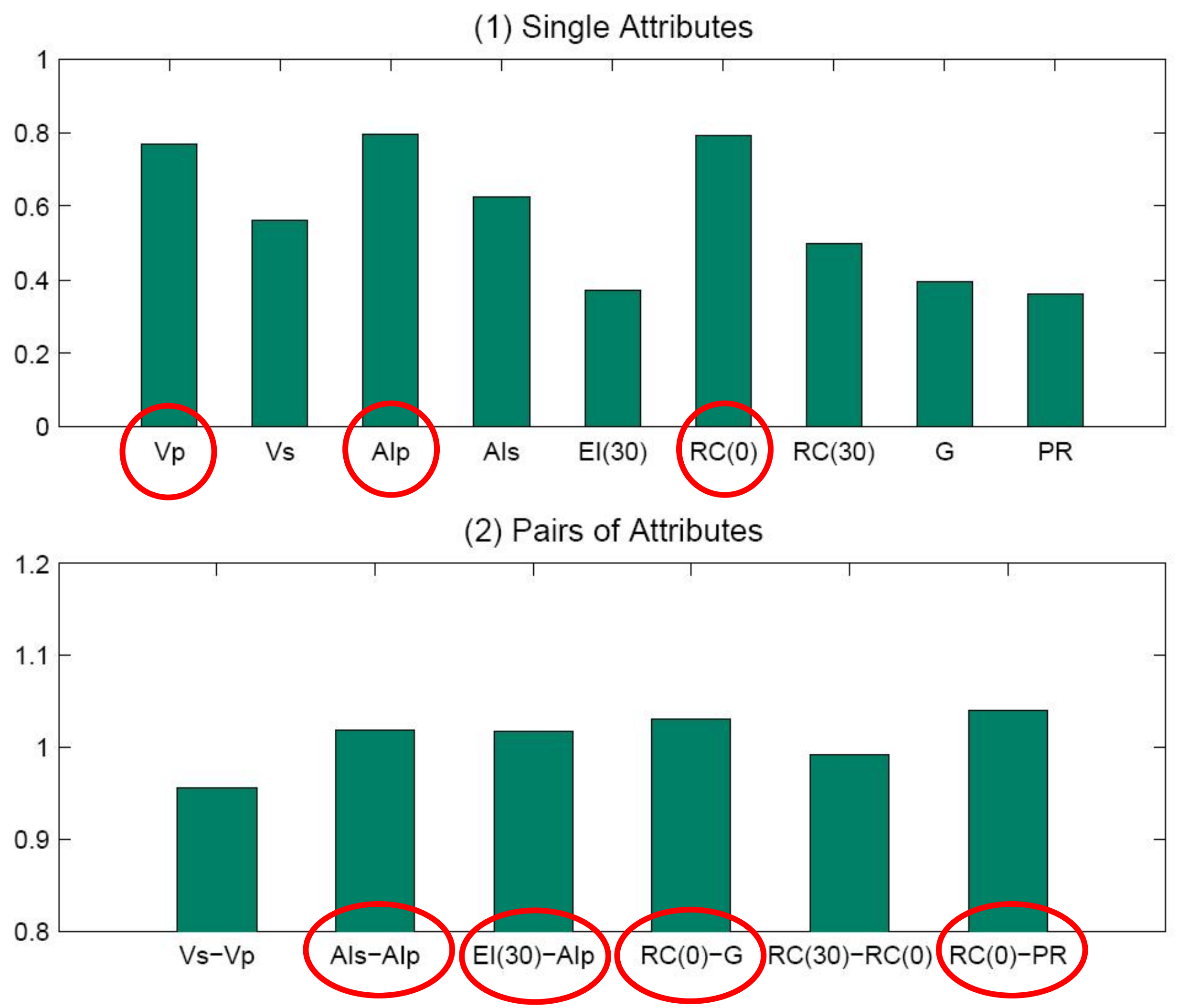
Bivariate



Takahashi et al. (1999)

Mutual Information

Information about lithofacies carried by Seismic attributes



Takahashi et al. (1999)

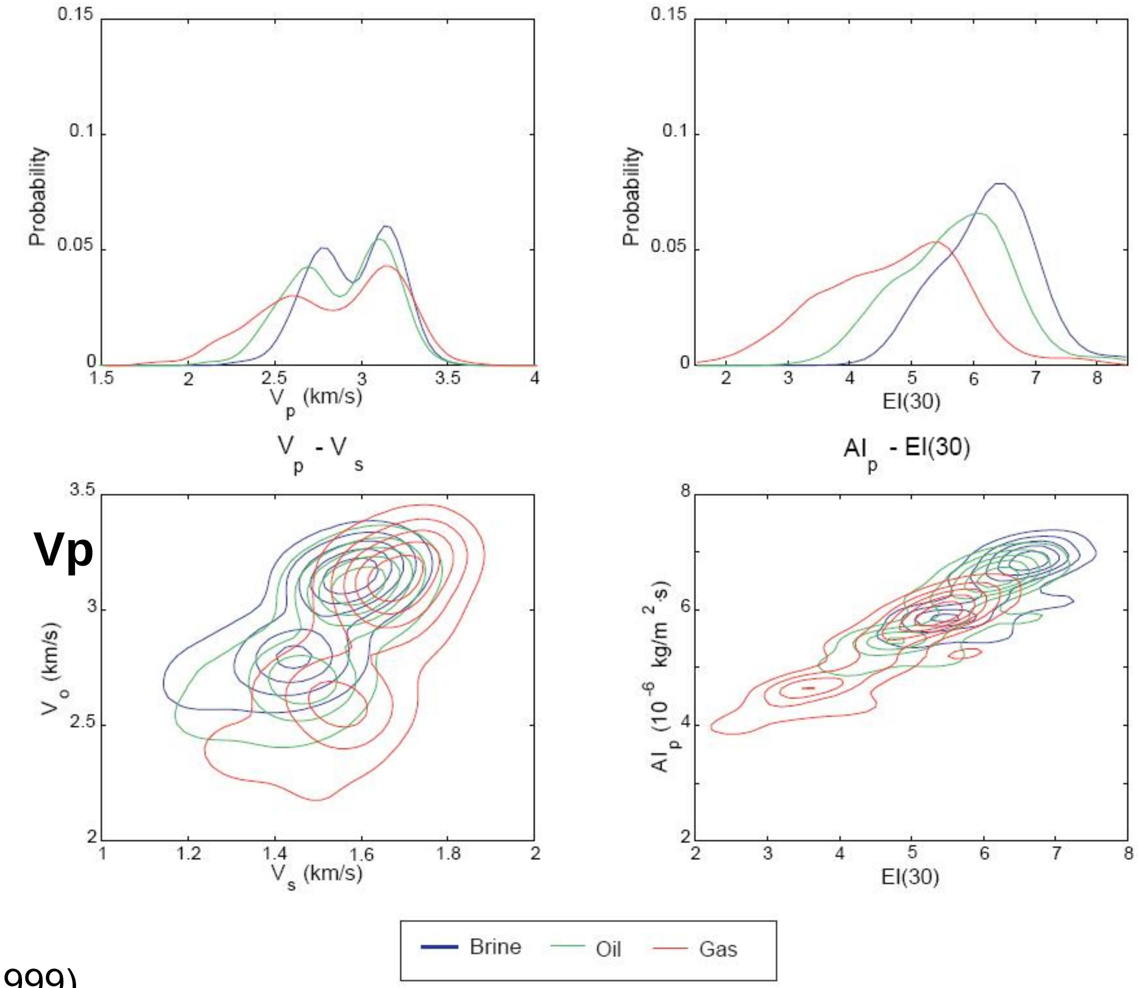
Case Study 2 (Pore Fluid)

Extended PDFs

Gassmann eq. was applied for fluid substitution in only sand reservoir

Univariate

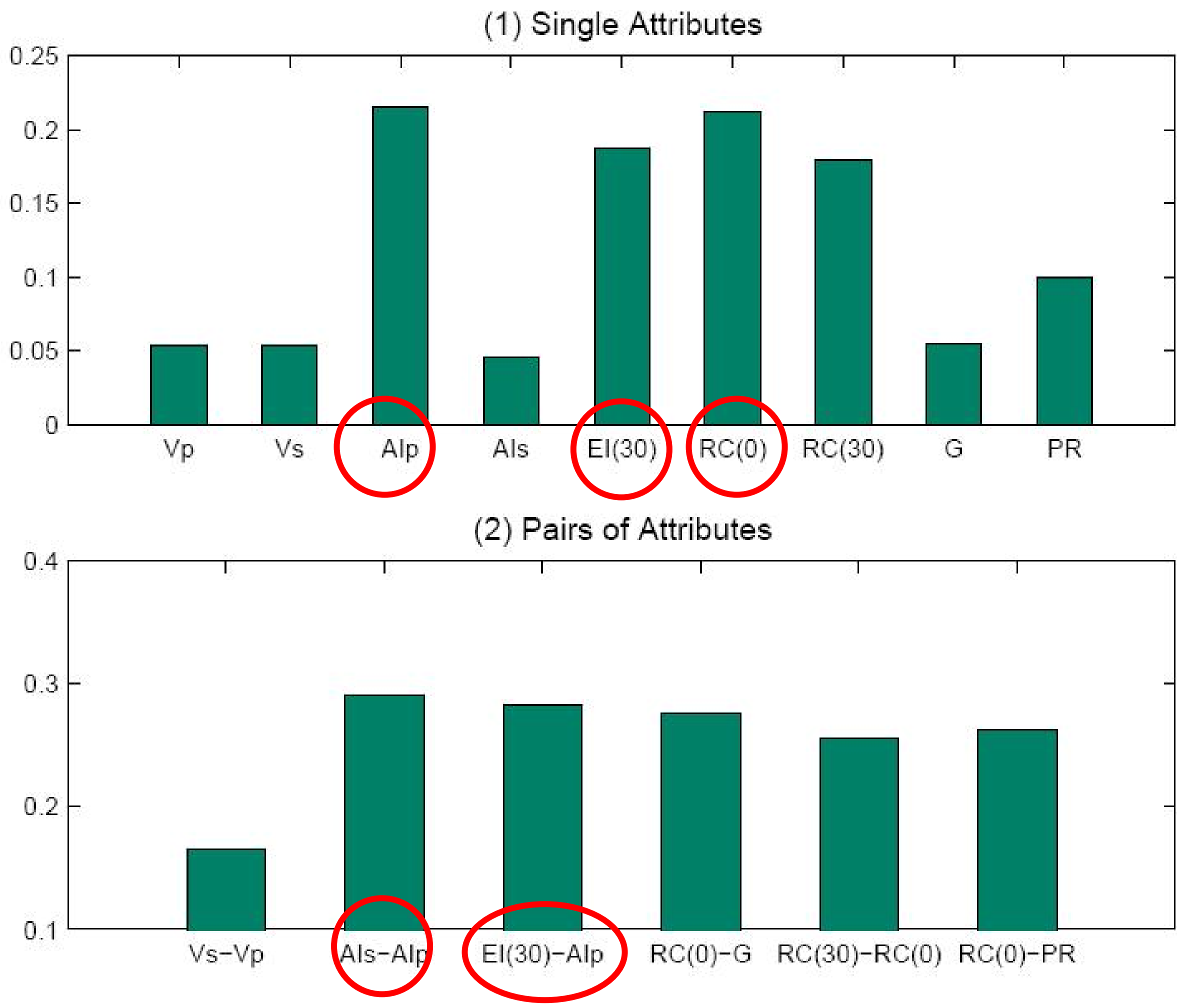
Bivariate



Takahashi et al. (1999)

Mutual Information

Information about **pore fluid** carried by Seismic attributes



Takahashi et al. (1999)

Discussions

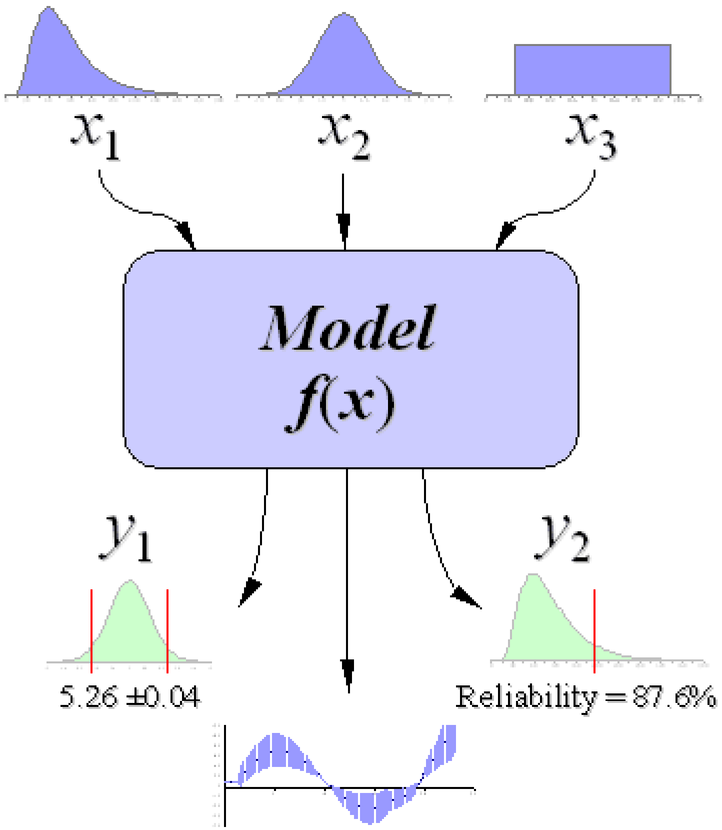
Linear measures of uncertainty, such as variance (σ^2) and covariance (σ_{12}), can be used instead of the entropy (H) ?

- Variance (covariance) can work only at limited situation
  - Parametric PDFs, such as Gaussian distribution
  - Continuous variable

- Information Entropy can work more flexibly
  - Nonparametric PDFs
  - Categorical variables (Shale, Sand)

The Entropy offers a more flexible representation of the state of information about the rock.

Monte Carlo Simulation

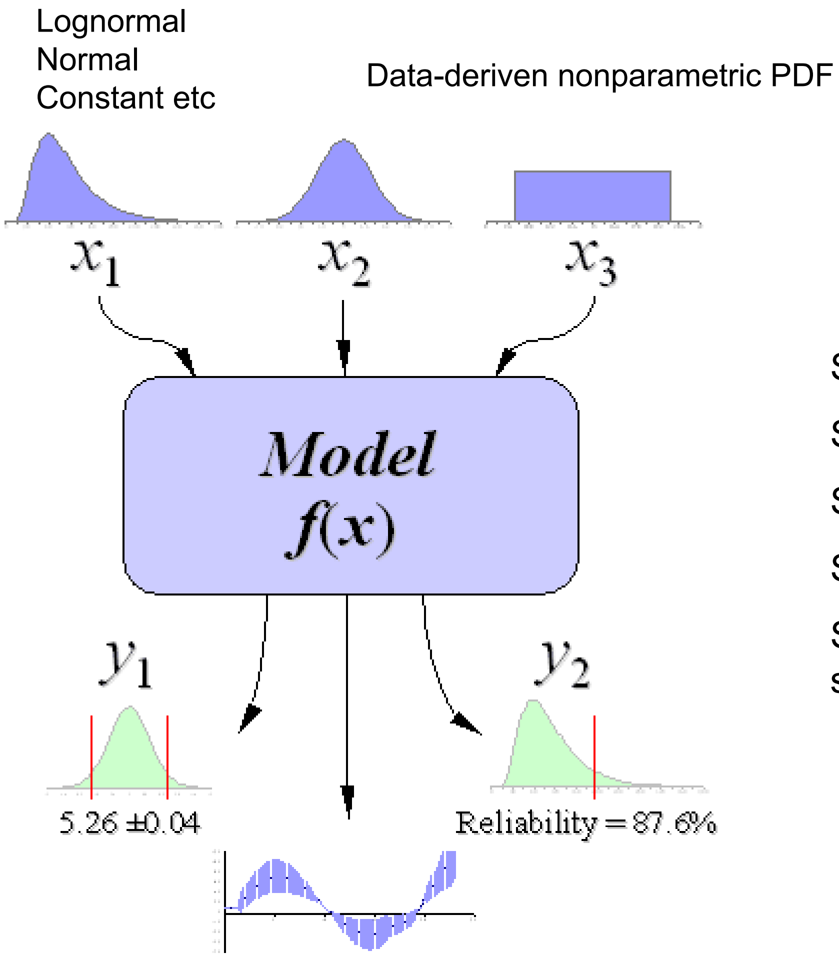


A technique using random numbers for probabilistic solution of a model

- Model is nonlinear system
- Input parameters with uncertainty
- Uncertainty analysis instead of deterministic method

http://www.vertex42.com/ExcelArticles/mc/MonteCarloSimulation.html

Monte Carlo Simulation



Step 1: Create a model, y = f(x_1, x_2, ..., x_q).

Step 2: Generate a set of random inputs, x_{i1}, x_{i2}, ..., x_{iq}.

Step 3: Use the model to obtain outputs.

Step 4: Repeat steps 2 and 3 for i = 1 to n.

Step 5: Analyze the results using histograms, summary statistics, confidence intervals, etc.

http://www.vertex42.com/ExcelArticles/mc/MonteCarloSimulation.html

Thank you for attentions