Automated Mineral Identification and Remote Sensing

Clark Glymour

Carnegie Mellon University

and

Institute for Human and Machine Cognition, University of West Florida

and

Joseph Ramsey

Carnegie Mellon University

With thanks to T. Roush, Ames; P. Gazis, Ames; and R. DeSilva, CMU.

The Goals

1. Automatically identify the qualitative—and if possible the quantitative—mineral composition of surfaces from their visible/near-infrared/infrared spectra.
2. Do it with small demands on computational space and time.
3. Do it for surfaces remote from the instrument.

Why?

1. Exploring extraterrestrial geology.
2. Analyzing earth surface composition.
3. Terrestrial industrial and scientific applications.
4. Because the instrumentation is cheap and lightweight and long used.

Relevant Instruments

Visible/Near Infrared Spectrometer—Mars, 2005 planned

Infrared Spectrometers—Most recent Mars orbiter; Earth satellites

Focus: Visible/Near Infrared Reflectance Spectroscopy

- Used in geology for over 70 years.
- Wavelengths 0.4 – 2.5 m.
- Because power spectrum of the sun changes, requires that reflected light from surface be compared with light reflected from white surface.
- Considerable laboratory and field spectra available for rocks and soils whose composition has been independently determined.

Why the Problem is Hard

- 3,000 + standard Earth minerals, but small libraries of laboratory reference spectra of pure minerals.
- Rocks and soils surfaces are typically aggregates of several minerals.
- Spectra of component minerals can combine non-linearly to produce a surface spectrum.
- Some chemically different minerals have essentially identical spectra in some wavelength ranges.

Some Proposed Techniques

- Regress the unknown sample spectrum against a linear combination of laboratory spectra using least squares or other fit criterion (Old Standby).
- Identify mineral classes by a few characteristic spectral features (Ames Expert System).
- Use linear combinations of laboratory spectra to train a neural network to identify a particular class of minerals (JPL).

Evaluating Algorithm Proposals

- Need a human expert performance baseline.
- Need comparison tests of alternative algorithms using the same test sets.
- Need a variety of test sets.
- Need to test in the field with remote unknown samples.
- *NASA seems to have no systematic procedures for the evaluation of intelligent software alternatives.*

Our Work So Far

- Established a *Human Expert Performance Baseline* using laboratory test spectra.
- Tested a wide range of machine learning algorithms on the same test data used for the human expert.
- Using field data, tested several of the best of these algorithms against human experts.
- Tested algorithms with remote sample unknowns.
- Designed automated methods for tuning search procedures to particular mineral classes.

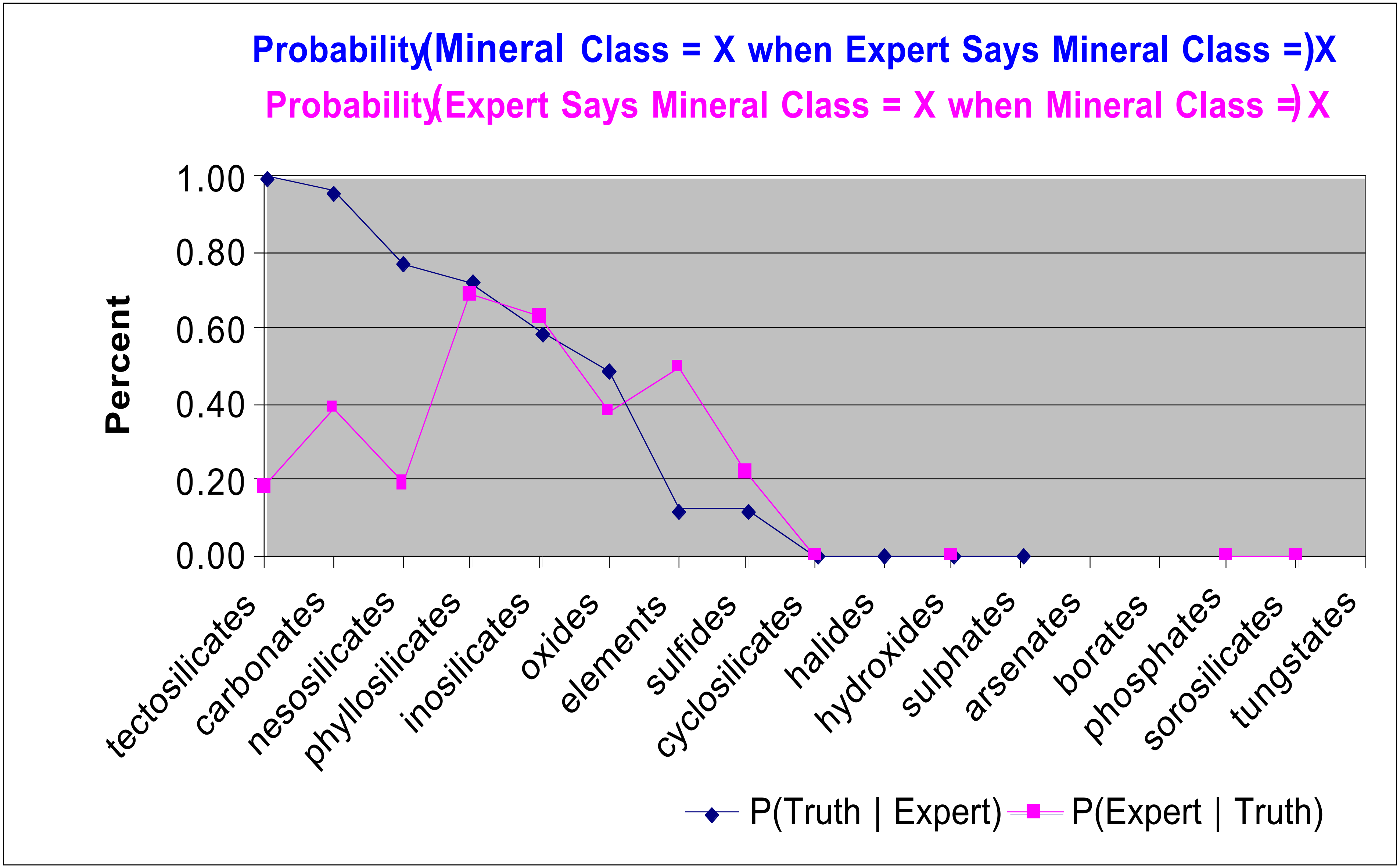
Results in Brief

In extensive tests of scores of algorithms with laboratory and field data, we have found algorithms that:

- In laboratory tests, identify a significantly larger percentage of carbonates than does a human expert from spectral data alone.
- In field tests, match the judgments of human experts who have access both to the rock sources and to the spectra.
- At the cost of slightly more false positives, identify significantly more forms of carbonate than published algorithms.
- Can be readily adapted to identify other classes of minerals.

Establishing a Human Expert Performance Baseline

- Tested the accuracy of a NASA expert (T. Roush) to detect the presence of each of 17 classes of minerals in 192 rock and soil samples (from the Johns Hopkins spectral library) using only the visible/near IR spectrum of each sample.
- Composition of test set independently estimated from laboratory petrology.
- Expert had unlimited time; access to any desired reference works. Actually took about 12 hours.

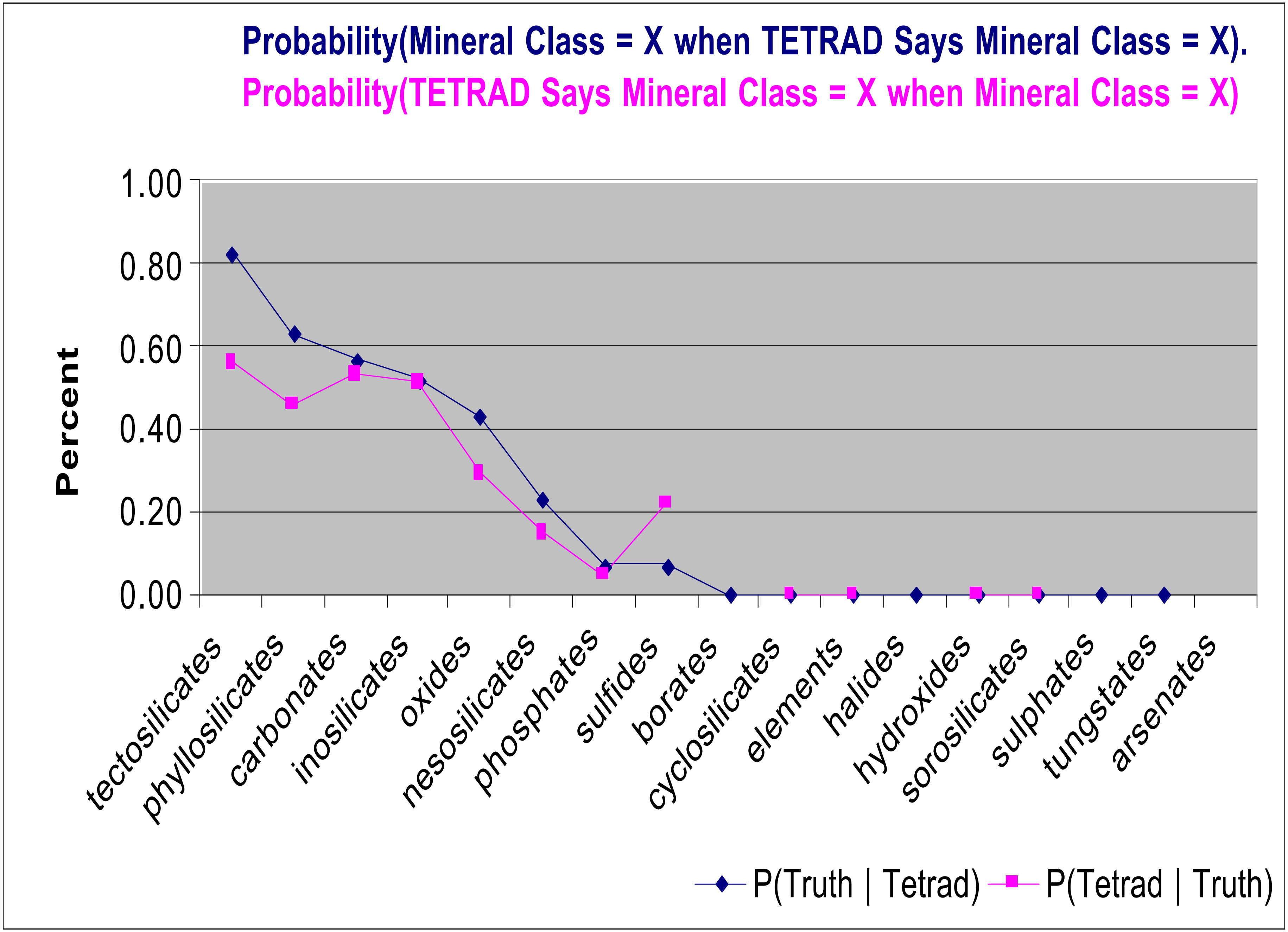

Probability(Mineral Class = X when Expert Says Mineral Class = X)
Probability(Expert Says Mineral Class = X when Mineral Class = X)
Percent
1.00
0.80
0.60
0.40
0.20
0.00
tectosilicates
carbonates
nesosilicates
phyllosilicates
inosilicates
oxides
elements
sulfides
cyclosilicates
halides
hydroxides
sulphates
arsenates
borates
phosphates
sorosilicates
tungstates
P(Truth | Expert)
P(Expert | Truth)

The Simplified TETRAD Algorithm

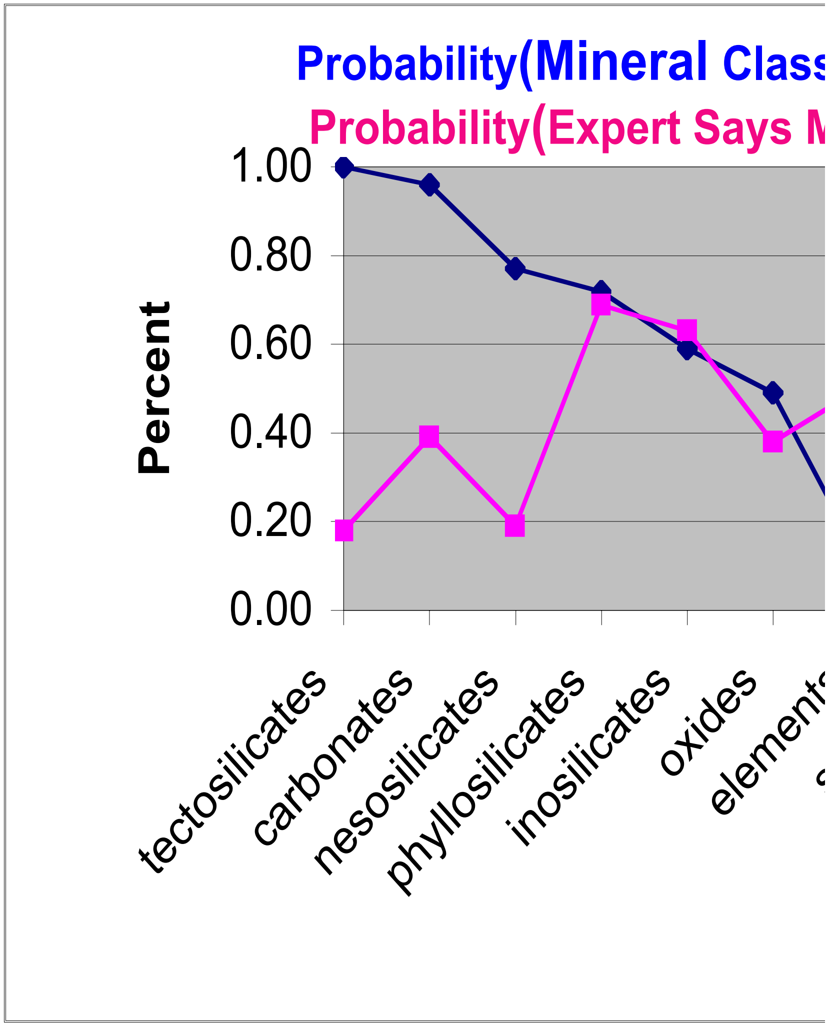
- Use JPL Library of spectra of 135 large grain powdered minerals as reference set. Order the reference set.
- Treat recorded wave lengths (frequencies) as units.
- Intensity of spectrum (at a frequency) is the only variable.
- For each JPL mineral compute the correlation of its spectrum with the unknown; eliminate the mineral if the correlation is zero.
- For each remaining ordered pair of minerals, compute the partial correlation of the spectrum of the first mineral with the unknown, controlling for the spectrum of the second mineral; eliminate the first mineral of the pair if the partial correlation is zero.
- Continue with remaining minerals, controlling for two spectra, three spectra, etc., until no further minerals are eliminated.

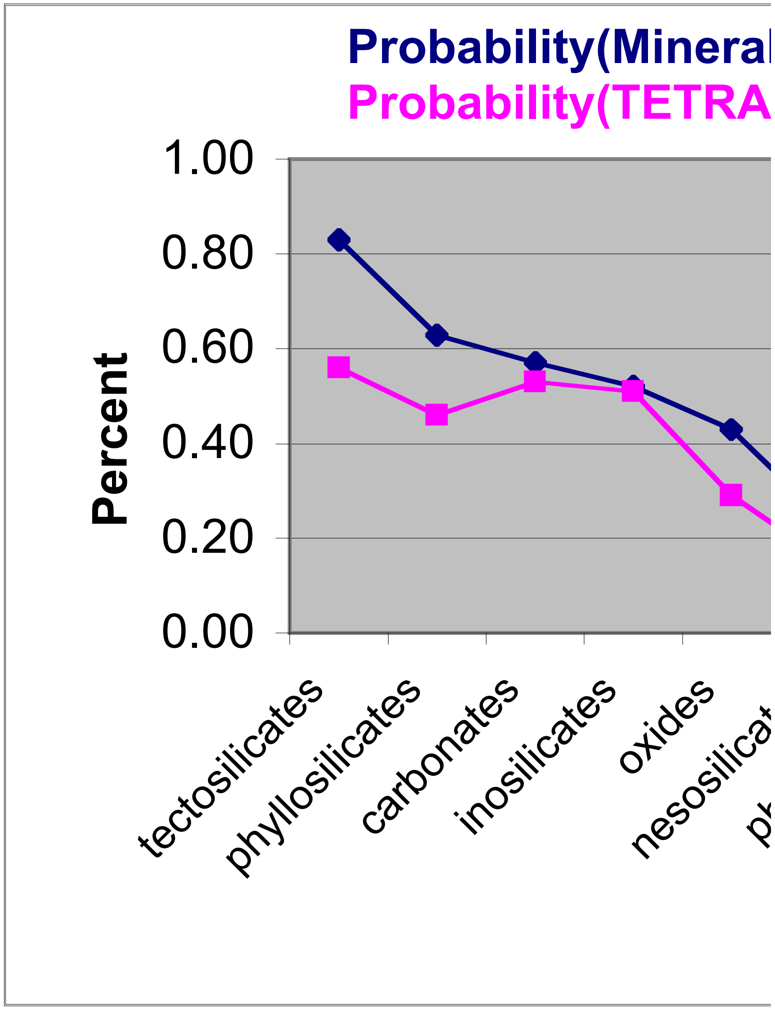
The Simplified TETRAD Algorithm

- Output of program is set of estimated mineral classes present in the sample.
- Program requires one parameter, a significance level for partial correlation tests, set by the user.
- Lower significance levels result in more cautious output.
- Significance level set at .05 in all experiments reported here, unless otherwise noted.


Probability(Mineral Class = X when TETRAD Says Mineral Class = X).
Probability(TETRAD Says Mineral Class = X when Mineral Class = X)
Percent
1.00
0.80
0.60
0.40
0.20
0.00
tectosilicates
phyllosilicates
carbonates
inosilicates
oxides
nesosilicates
phosphates
sulfides
borates
cyclosilicates
elements
halides
hydroxides
sorosilicates
sulphates
tungstates
arsenates
P(Truth | Tetrad)
P(Tetrad | Truth)

Comparing: Human Expert and the Simple TETRAD Program





Some Things We Discovered Looking at Expert and Machine Performance

- Among all of the 92 JHU rocks containing carbonates (almost half of the 192 test rocks) the expert identified *only* those that are dolomites or calcites—the two most common forms of carbonate on Earth—as in marble and limestone.
- The expert was really a "calcite or dolomite" detector, not a carbonate detector.
- The algorithm did worse for carbonates if given all of the spectral data than if given just the long wavelength end of the spectrum.

Tests of 25 +Machine Learning Algorithms For Carbonate Identification Using JHU Test Data

- Least squares multiple regression
- Least squares multiple regression for dolomite and calcite only
- Simplified TETRAD Algorithm with and without spectrum restricted to 2.0 – 2.5 μ
- Simple TETRAD for dolomite and calcite only
- MODEL 1 Commercial Program:
  - Stepwise Regression (several varieties)
  - Neural net models (several varieties)
  - Probabilistic Decision Trees

JHU TESTS: 192 Samples, 92 with Some Carbonate Content

Method	False Negatives	#Identified Correctly	#False Positives
God	0	92	0
TETRAD	54	38	20
TETRAD 2-2.5μ	45	47	16
TETRAD 2-2.5μ Cal. or Dol	54	38	3
Least Squares +	1	91	100
Least Squares + Cal or Dol	2	90	86
Model 1	56	36	37
Human Expert	68	24	1

JPL/Ames Field Tests in Silver Lake, California

- Spectra taken *in situ*, close up.
- 30 spectra taken; some spectra rejected because too noisy; 21 spectra from 21 distinct samples obtained for analysis.
- Expert geologists in the field identified samples for carbonate content by their physical appearance and their spectra.
- Laboratory analysis of composition obtained for 9 samples—agreed with field experts in all cases.

Sample Name	**Field Expert**	**TETRAD 2.0-2.5µm)**	**TETRAD (Cal, or Dol; 2.0-2.5 µm)**	**Laboratory Analysis**
Emperor #1	**C**	C	C	C (90%) NC (10%)
Emperor #2	**C**	C	C	C (90%) NC (10%)
T 103	**NC**	NC	NC	NA
T 105	**NC**	NC	NC	NA
T 106	**C**	C	C	NA
Endolith	**C**	C	C	C (93%) NC (7%)
Tubular-tabular	**NC**	NC	NC	NC (100%)
→ Arroyo disturbed	**C**	NC	NC	C (20%) NC (80%)
Arroyo undisturbed	**C**	C	C	C (25%) NC (75%)
C3PO	**C**	C	C	NA
Chewie	**NC**	C	NC	NA
Jabba	**C**	C	C	NA
Jawa	**C**	C	C	NA
Lando	**C**	C	C	C (93%) NC (7%)
Luke	**C**	C	C	NA
R2D2	**C**	C	C	C (78%) NC (22%)
Solo	**C**	C	C	NA
Tarken	**NC**	NC	NC	NA
Vader	**NC**	NC	NC	NA
Valentine	**NC**	NC	NC	NC (100%)
Yoda	**NC**	NC	NC	NA

Total Correct 19 20

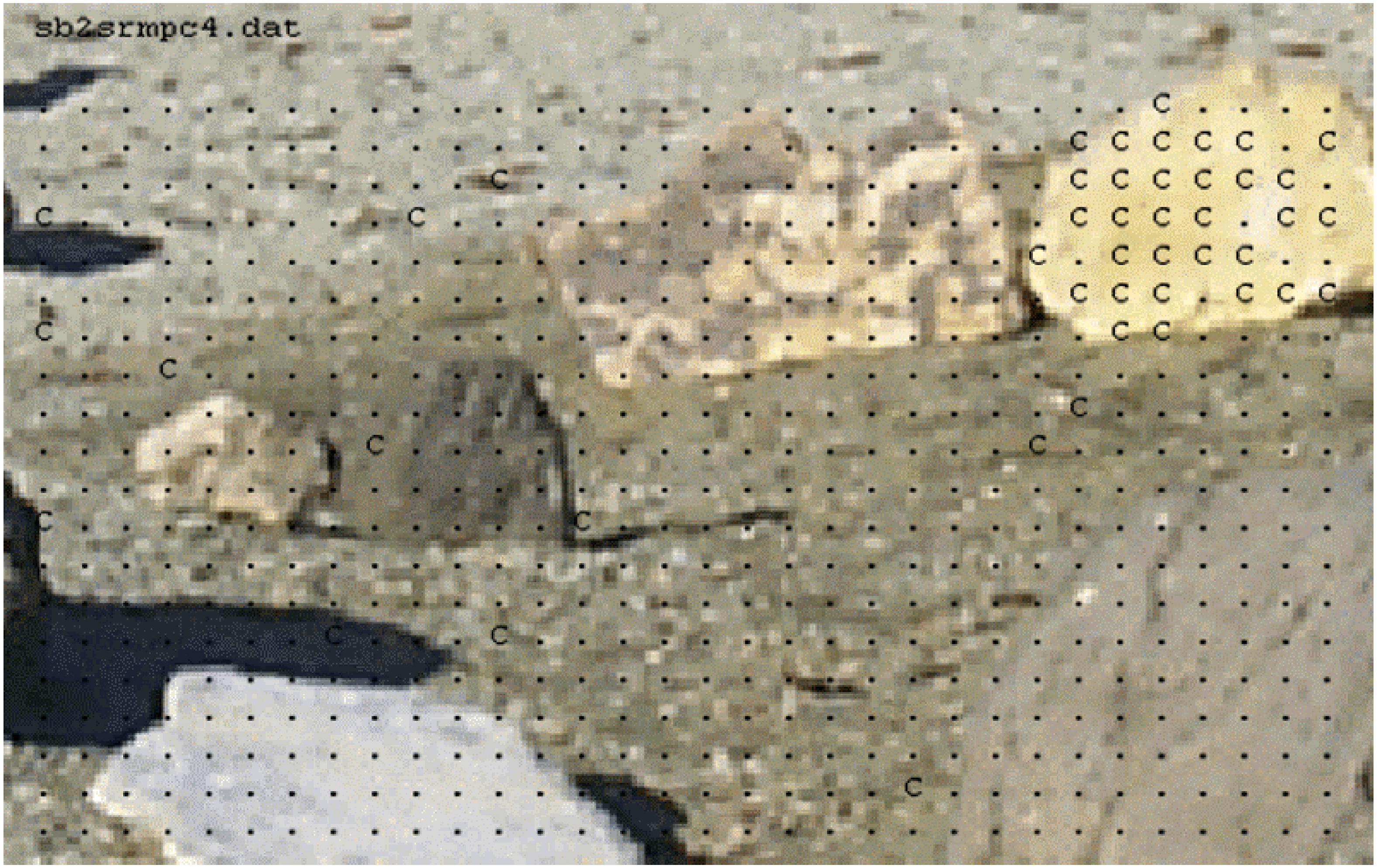
Summary Results of the JPL/AMES Field Test in Silver Lake

- Simplified TETRAD with data restricted to 2.0-2.5 μ and only reporting calcites or dolomites identifies 12 of 13 carbonates, with no false positives.
- Simplified TETRAD restricted to 2.0-2.5 μ reporting all carbonates identifies 12 of 13 carbonates, with one false positive.
- Ames Expert System, using feature detection, identifies 9 of 13 carbonates, no false positives; partial least squares does the same.
- JPL team gave unclear report, but show only 8 carbonates (Gilmore, et al., (2000). Strategies for autonomous rovers at Mars. *J. of Geophysical Research*, 105, p. 29,223-29,237).

Ames Scene Test

- Area of ~ 100 sq. feet salted with rocks of known composition, including one large carbonate, large sulphate, concrete and many non-carbonate rocks.
- Spectra taken from several meters away from the area, with white reference at nearest rock to the spectroscope.
- Sequence of spectra taken, with small field, collectively covering the entire area.
- Task: to identify the regions containing carbonate.
- Least squares, expert system, human expert, tested (Ames).
- Simple TETRAD tested with 2.0 – 2.5 μ data filter and cerrusite eliminated from reference set (because it is indistinguishable from some sulphates in that interval).

Simple TETRAD Results (Blind; .01 significance level for correlation tests)



White Rock in upper right hand corner is carbonate.

Comparisons for the Ames Scene Test

- Human expert and expert system give results similar to TETRAD
- Least squares spatters “carbonate” all over the place
- TETRAD results vary with significance level used for deciding correlations. More false positives with .05 significance level.

Ames Test of Mineral Identification with Varied Location of White Reference

- Spectra taken with white reference at target 28 feet from spectrometer; and with white reference 2 feet from spectrometer.
- Targets: granite, marble and terra cotta commercial tiles.
- 8 spectra taken of each kind of tile, with both rough and smooth surfaces, with white reference next to target
- 8 spectra taken of each kind of tile, with both rough and smooth surfaces, with white reference proximate to spectroscope.

Ames Test of Mineral Identification with Varied Location of White Reference

	Reference at Target	Reference at Instrument
Ames Expert System	2 of 8 carbonates no false positives	2 of 8 carbonates no false positives
TETRAD, 2.0 – 2.4 μ, .05 significance	7 of 8 carbonates 4 false positives	7 of 8 carbonates 1 false positive
TETRAD 2.0 – 2.4 μ .01 significance	7 of 8 carbonates 2 false positives	7 of 8 carbonates 3 false positives

Explanations

Why Does the Simple TETRAD Program with Data Filtering Outperform the Expert System on the Silver Lake Test?

1. Because the Expert System is essentially a "dolomite or calcite" detector and there are other carbonates.
2. Because the expert system looks at three lines around 2.3 m to make its decision, and the 2.0-2.5 region contains more information characteristic of carbonates.

Explanations

Why Does the Simple TETRAD Program Identify Carbonates More Accurately When Spectra Outside the 2.0 – 2.5 μ Interval Are Masked?

- Because the rest of the spectrum, 0.4 – 2.0 μ, is enormously variable for carbonates and in mixed sources may be dominated by other mineral components.
- Result: if the entire spectrum is used, the correlation of the spectrum of a reference carbonate with the spectrum of a mixed composition carbonate sample is lowered, and the algorithm makes more errors.

Explanations

Why Does Least Squares Do So Poorly in All Tests?

1. For carbonate identification, least squares (aka multivariate regression) has the same extraneous noise problems as the TETRAD algorithm outside the 2.0 – 2.5 μ region, but for statistical reasons, it cannot use the data mask.
2. Why regression can't use the data mask: In estimating the contribution of the spectrum of reference mineral M to the unknown spectrum, regression computes the partial correlation of the M spectrum and the unknown spectrum, controlling for ALL other reference spectra. But the effective sample size of the statistical significance tests is reduced by 1 for every variable controlled for. With a data mask, the effective sample size would be 0 using JPL library as reference.

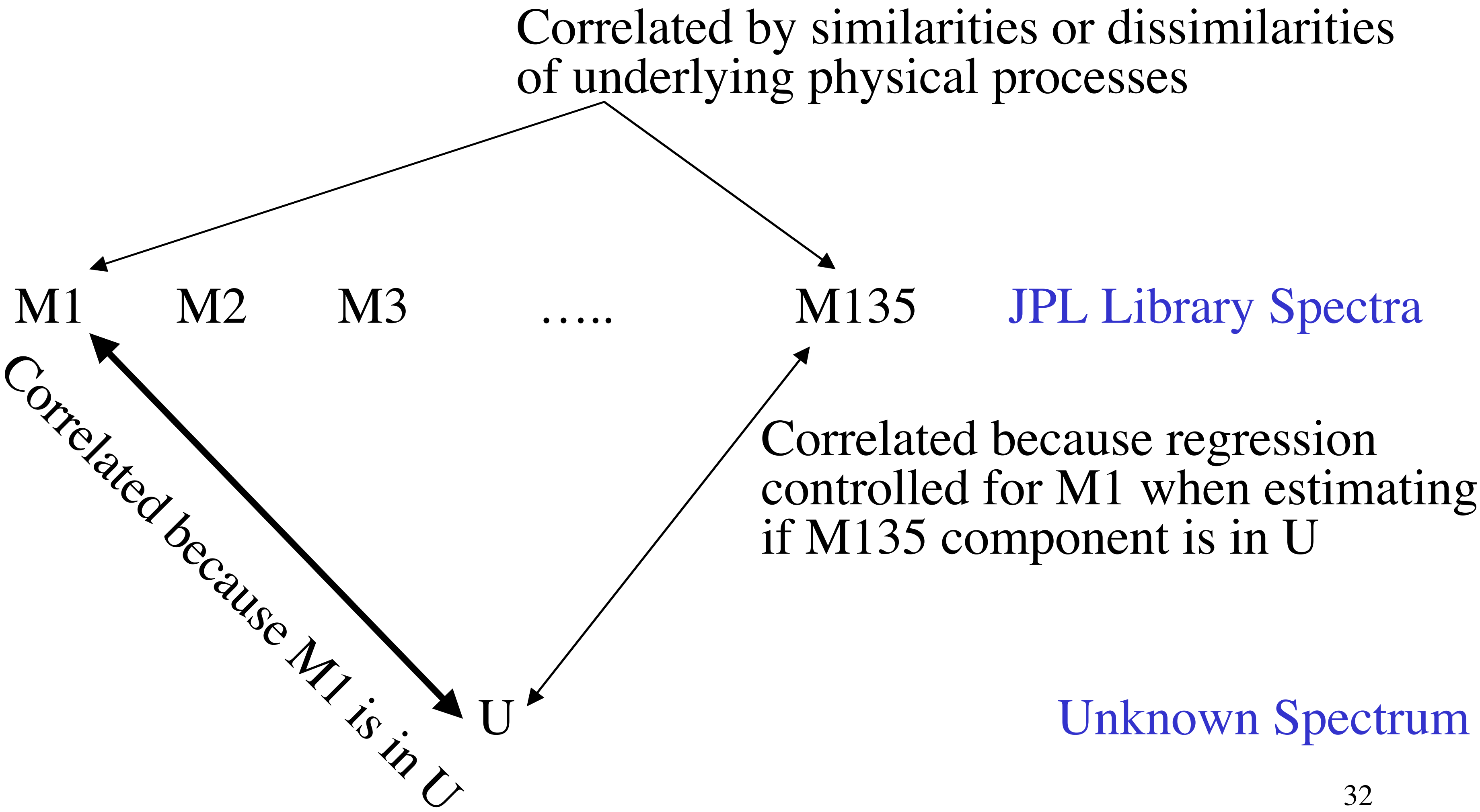
Explanations

Why Does Least Squares Do So Poorly?

1. If M1 and M2 are correlated, and M1 and U are correlated, and M2 and U and *uncorrelated*.then (depending on how the correlations come about) M2 and U may be correlated if M1 is controlled for. The partial correlation of M2 and U, controlling for M1 may be positive or negative, depending on the signs of the M1, M2 correlation and of the M1, U correlation.
2. Multivariate regression estimates the contribution of any reference mineral, e.g., M2, by computing the partial correlation of M2, U controlling for *all* other reference minerals.
3. N.B. The TETRAD algorithm minimizes controlling for other reference minerals.

Explanations

Why Does Least Squares Do So Miserably?



Correlated by similarities or dissimilarities
of underlying physical processes

M1 M2 M3 ….. M135 JPL Library Spectra

Correlated because M1 is in U

Correlated because regression
controlled for M1 when estimating
if M135 component is in U

U Unknown Spectrum

Explanations

Why Does the Simple TETRAD Program with Mask Do Better than Neural Nets?

- In principle, neural net classifiers would appear ideal for the problem.
- In practice, neural net classifiers require large training sets, and none are available.
- Synthetic training sets, produced by taking linear combinations of lab spectra of pure minerals, may be unrealistic in this spectral region.
- If unknowns contain a target mineral, e.g., a carbonate, combined with minerals not in the neural net’s training set, the neural net tends to miss the target mineral.

Problems and Prospects

- Finding data masks for other mineral classes.
- Improving the simplified TETRAD algorithm.
- The infrared.
- NASA procedures for intelligent software comparative evaluations.

Finding Data Masks:
2 Automated Methods

- Mutual information method: the intensity scale at each frequency is binned, and the information (e.g., for carbonates) computed for each frequency. Low information frequencies are masked.
- Genetic algorithm: Spectrum is divided into ten intervals, coded as genes with two alleles (corresponding to deleted/not deleted). Each genome corresponds to a mask. Genetic algorithm run with simple TETRAD algorithm used to score each mask by % of JHU carbonates correctly identified with that mask

Finding Data Masks: 2 Automated Methods

- Information method is fast but very sensitive to number of bins used
- Genetic algorithm is very slow; more accurate with finer partition of the spectrum (e.g., 10 rather than 8 genes).
- Genetic algorithm gives excellent mask for carbonates; well defined mask that works pretty well for inosilicates.
- Work remains to be done finding other mineral classes for which there are effective data masks that improve identifiability.

Improving the Simple TETRAD Algorithm

- Algorithm is low time complexity. Space requirements are essentially storage of a reference library.
- Fixed ordering of minerals can lead to errors and can be improved in reliability and speed by heuristics in Spirtes, et al., *Causation, Prediction and Search*, MIT Press, 2001.
- Algorithm can be altered to list disjunctions of two or more minerals when any of the disjuncts can equally well account for the spectra.

The Infrared

- Thermal Emission Spectrometer aboard most recent Mars orbiter.
- Generally believed spectra closer to additive in this region.
- Standard technique for identifying composition is least squares step-wise regression. (M. Ramsey)
- Procedure may be subject to same “partial correlation error” as with visible/near IR spectra and statistical problems of least squares.
- No published investigation of alternative algorithms for this spectral region.

The Final Problem: NASA

As robotic exploration becomes more autonomous, NASA mission planners will make decisions about what intelligent software to deploy for robot operations, failure detection, data analysis, and decision making.

There are many possible architectures for such intelligent software, and research on many alternatives is supported by NASA.

But there seems to be no established procedure for *comparative* testing of intelligent software, from whatever sources, before deployment decisions are made.