# Targeted Anomaly Detection in Hyperspectral Imaging for Cultural Heritage Applications – Finding a Needle in a Haystack

Neal B. Gallagher
Eigenvector Research, Inc.
nealg@eigenvector.com

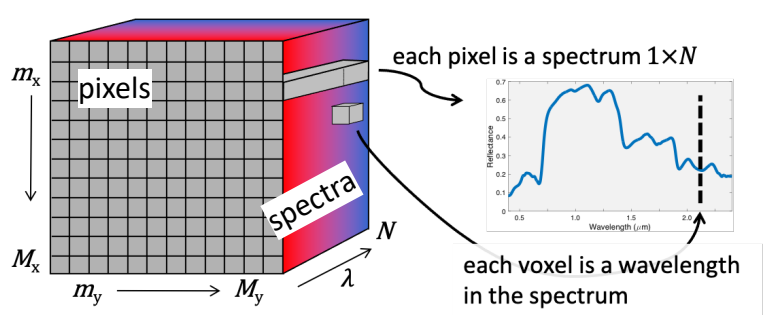


EIGENVECTOR RESEARCH INCORPORATED

1

## What is Hyperspectral Imaging?

- Hyperspectral imaging measures a spectrum at many pixels
  - Provides good sampling for exploratory analysis, detection and classification
  - Rapid and non-invasive (good for cultural heritage and archeological samples)
  - Images can be big!
- Provides
  - Spatial statistics
  - Chemical Information




https://eigenvector.com/wp-content/uploads/2020/12/IntroHSIandMIA.pdf

EIGENVECTOR RESEARCH INCORPORATED

3

# Hyperspectral (HSI) Applications

- Good for heterogeneous samples
  - Although quantities of analytes may be low on a volume basis, signal in individual pixels can be dominated by an analyte of interest
  - Good for detecting a needle in a haystack and finding rare signals
  - In contrast, destructive wet chemistry methods can be hampered by dilution effects
- Chemical Sensing, Standoff Sensing and Remote Sensing
- Many algorithms available
  - Depends on the sensing objective and how the signal manifests

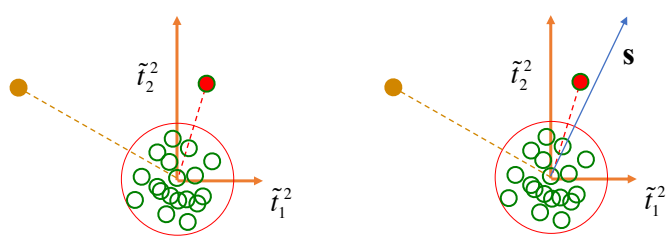
EIGENVECTOR RESEARCH INCORPORATED

4

# Detection Algorithms

- Anomaly detection finds unusual signal
- Target detection finds unusual signal that matches a specific spectrum
- Targeted anomaly detection finds unusual signal that matches a specific anomaly in an image
  - Once a strong anomaly is found, weaker anomalies that match it can be found

anomalies are found in any direction




target signal is found in a specific direction, **s**

EIGENVECTOR RESEARCH INCORPORATED

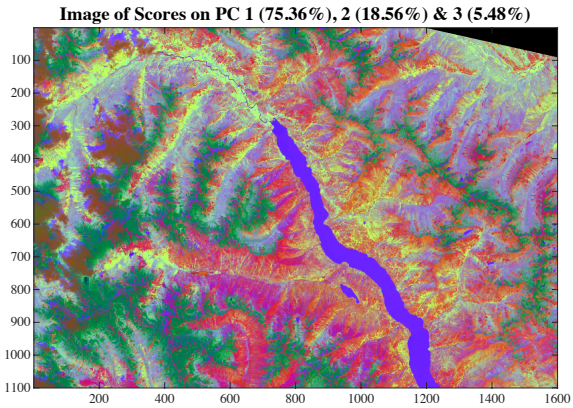
5

# Anomaly Detection Example

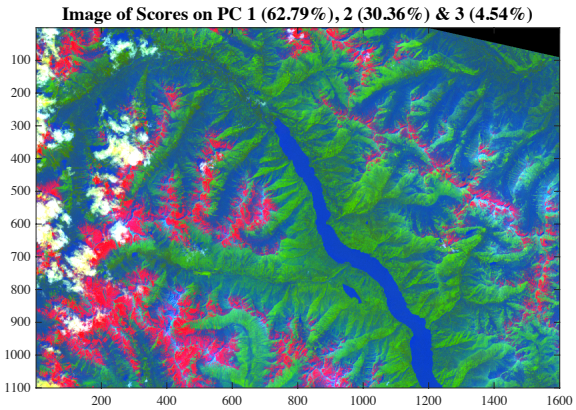
- Landsat 8 image of the NW end of Lake Chelan, WA USA
  - 1100 x 1600 pixels x 8 wavelength bands
  - 1760000 total pixels
- Principal Components Analysis (PCA)
  - Ubiquitous tool for multivariate analysis
  - Outliers can be found using:
  - Scores t
  - Hotelling's T2 (distance from data origin)
  - Q residuals (lack of fit)
  - This example modeled the entire image

6

# PCA of a Landsat 8 Image




Preprocessing: 1-norm, mean-center.
Image: autocontrasted.

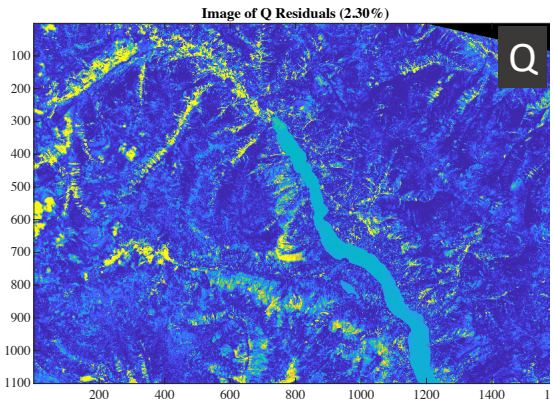


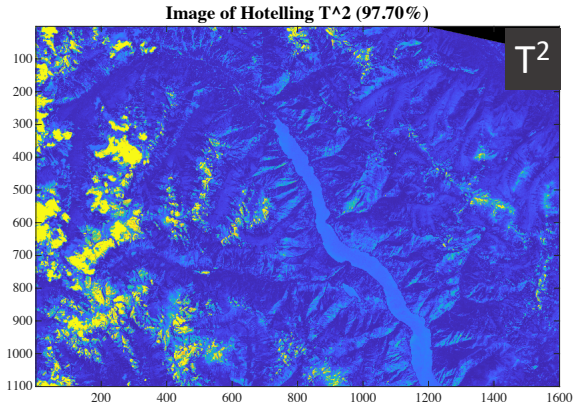

Preprocessing: autoscale.
Image: autocontrasted.

Data preprocessing can have a significant impact on the ability to detect outliers

7

# Examine Q and $T^2$ for Outliers



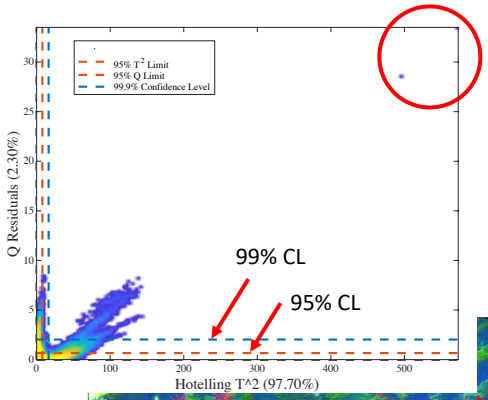




Preprocessing: autoscale.
PCA: 3 principal components.
Image: autocontrasted.

Inspection of images can show where anomalies may be but can be difficult to see small anomalies

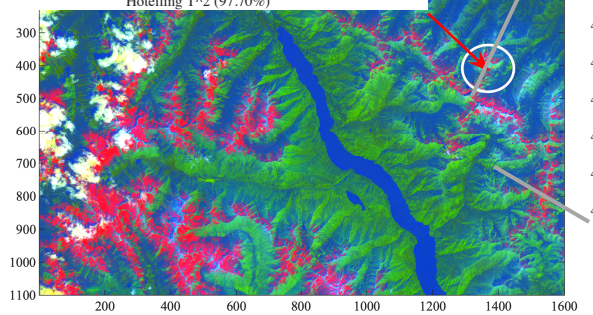
EIGENVECTOR RESEARCH INCORPORATED

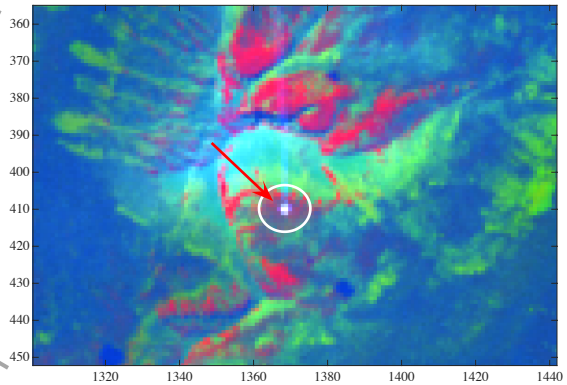
8




2 pixels at [1368 410] & [1369 410]

Plot of Q vs T2 can help identified truly unusual signal.
Changing color schemes and contrast can help explore the signal.





EIGENVECTOR RESEARCH INCORPORATED

9

# Examine Region Local to the Anomaly

Image of Scores on PC 1 (17.23%)

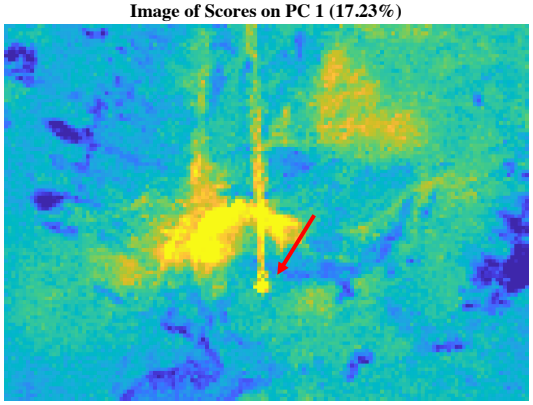
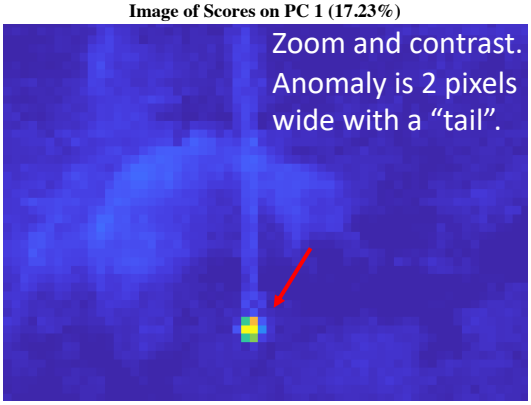
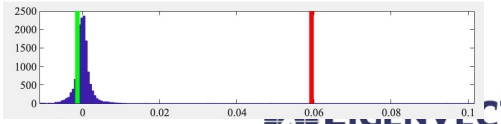


Image of Scores on PC 1 (17.23%)




Preprocessing: autoscale, glsw de-cluttering.
Image: autocontrasted.

EIGENVECTOR RESEARCH INCORPORATED

10

# Is It a Bird? A Plane?

- The pixel resolution for the image was 30 m.
  - The anomaly is ~2x2 pixels
- A 747 is 68.4 m wide and 76.3 m long
  - https://www.boeing.com/commercial/747/

Landsat 8 satellite

| Band Number | Wavelength Range (µm) | Resolution (m) |
|---|---|---|
| 1 | 0.433-0.453 | 30 |
| 2 | 0.450- 0.515 | 30 |
| 3 | 0.525- 0.600 | 30 |
| 4 | 0.630- 0.680 | 30 |
| 5 | 0.845- 0.885 | 30 |
| 6 | 1.560- 1.660 | 30 |
| 7 | 2.100- 2.300 | 30 |
| 8 (not used) | 0.500- 0.680 | 15 |
| 9 | 1.360- 1.390 | 30 |
| 10 (not used) | 10.6 -11.2 | 100 |
| 11 (not used) | 11.5 -12.5 | 100 |

EIGENVECTOR RESEARCH INCORPORATED

11

# Anomaly Detection Example

NIR reflectance image of a cellulosic swipe.

a) is an anomaly apparent?
b) where is it?
c) if seen, can the analyte be identified?

apparent anomaly observed – what is it?

Image of Scores on PC 1 (97.73%), PC 2 (1.28%) & PC 3 (0.56%)

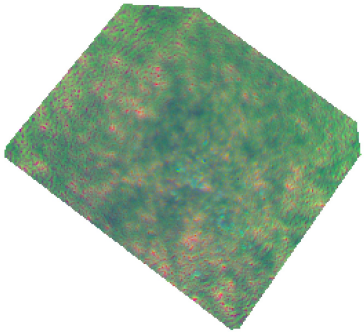
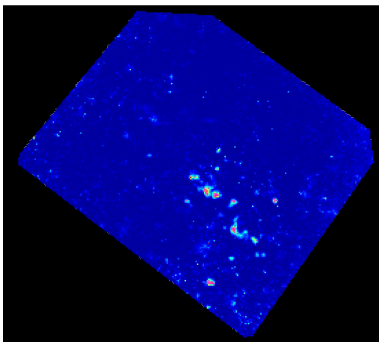


Image of Hotelling T^2 (99.83%)



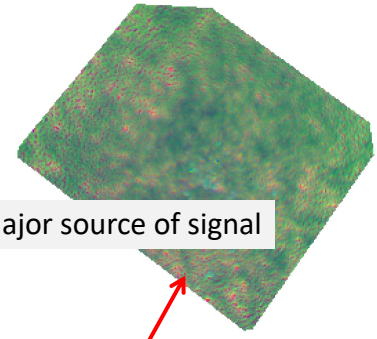
data set: courtesy OPOTEK, Inc., Carlsbad, CA

EIGENVECTOR RESEARCH INCORPORATED

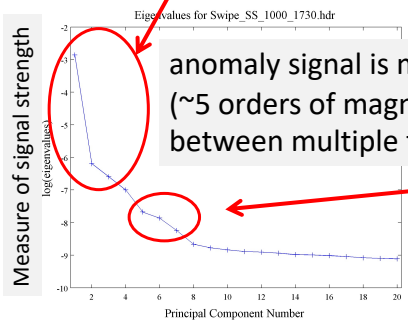
13

# Anomaly Detection Example

Image of Scores on PC 1 (97.73%), PC 2 (1.28%) & PC 3 (0.56%)



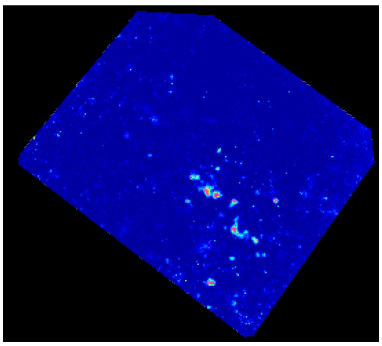
major source of signal




anomaly signal is minor (~5 orders of magnitude) and spread between multiple factors

apparent anomaly observed – what is it?

Image of Hotelling T^2 (99.83%)



EIGENVECTOR RESEARCH INCORPORATED

14

# Anomaly Detection Summary

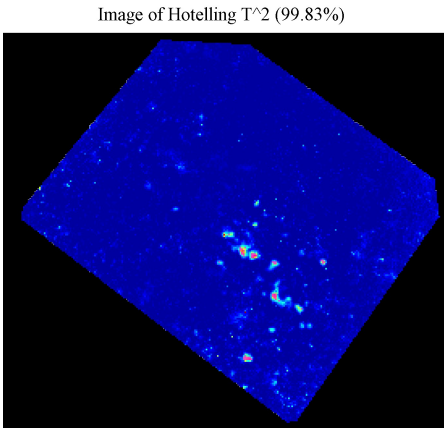
- Flexible - finds unusual signal
- Minor anomalies may be difficult to detect
  - Could appear like random noise
- Doesn't identify the signal
  - Library search
    - Spectra in library not relevant? Matrix, environment,…
    - Additional processing may be needed for impure signal to remove background

EIGENVECTOR RESEARCH INCORPORATED

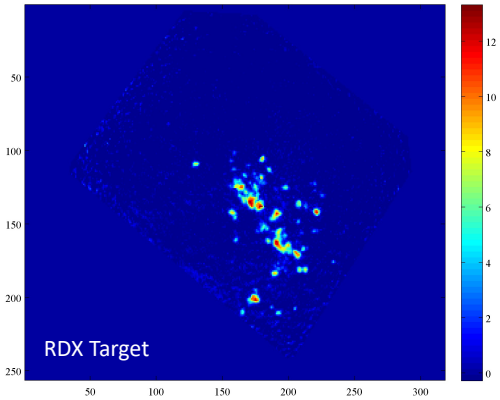
15

# Target Detection Example

Anomaly detection




Target detection
(also uses clutter suppression)

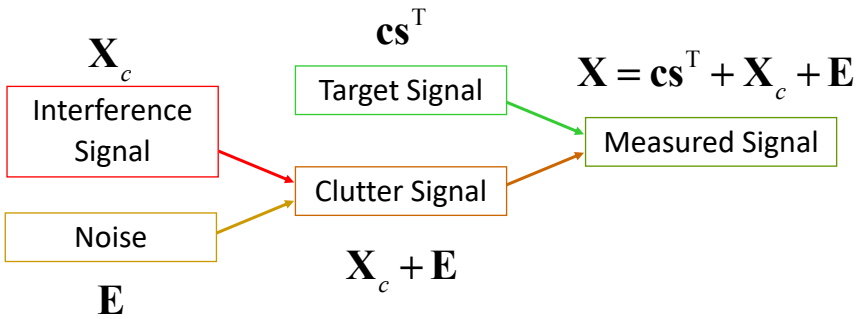



EIGENVECTOR RESEARCH INCORPORATED

16

# Measured Signal – What is Clutter?

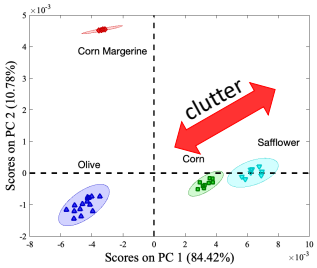
- Clutter is present in all measurements
  - clutter = interferences + noise not of interest




EIGENVECTOR RESEARCH INCORPORATED

17

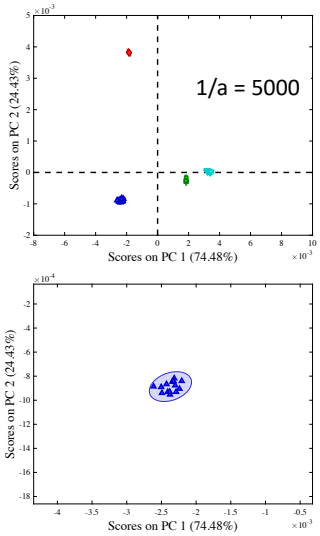
What does "decluttering" do?




generalized least squares weighting is a form of "de-cluttering" - increases inter- to intra-class variance. Increasing 1/a increases the weighting strength.

Paradigm Shift:
Maximize: signal-to-noise → signal-to-clutter

GLSW (weighting 1/a)

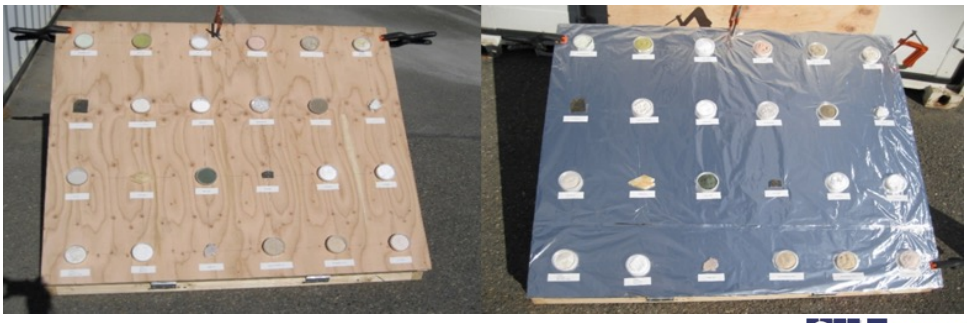



EIGENVECTOR RESEARCH INCORPORATED

18

# Target Detection Example Minerals on Plate

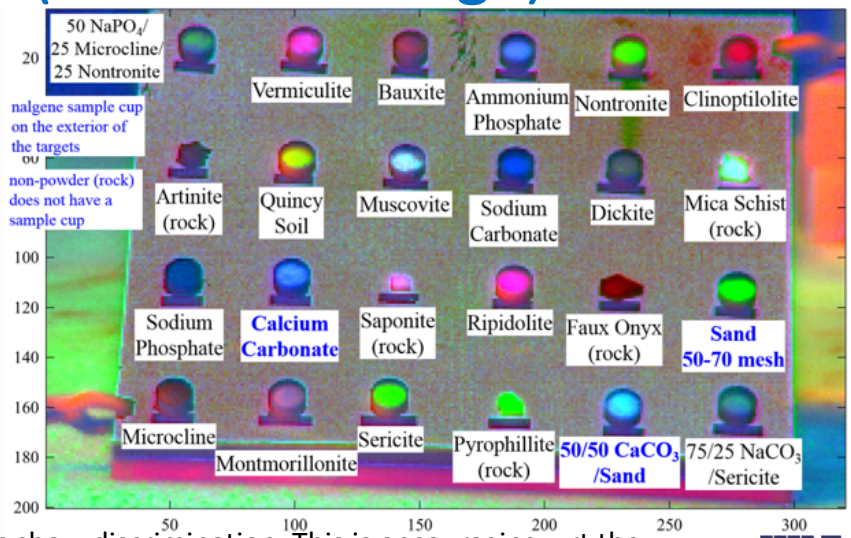
- Rock and Minerals (ground in Nalgene cup)
  - Samples: 8 inch columns, 11.3 inch rows
  - Bare plywood and aluminum foil surfaces at 45º at 14 m
  - Rudimentary atmospheric correction



EIGENVECTOR RESEARCH INCORPORATED

19

# PCA (PC Scores Image)



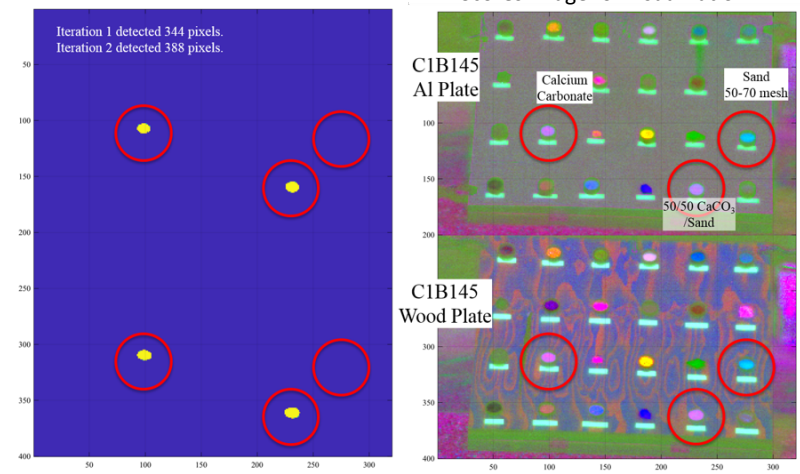

Scores images show discrimination. This is encouraging wrt the objective of detection and classification.

EIGENVECTOR RESEARCH INCORPORATED

20

# Target Detection for $CaCO_3$

Scores image for visualization

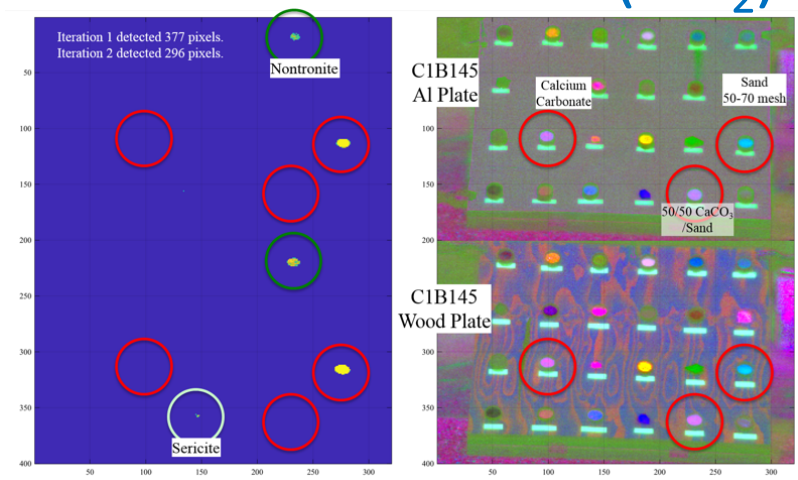



Only $CaCO_3$ is detected (including the 50/50 mixture)!

EIGENVECTOR RESEARCH INCORPORATED

21

# Target Detection for Sand ($SiO_2$)




Sand is detected (none in the 50/50 mixture with $CaCO_3$).
Nontronite is a false alarm. Sericite has a minor false alarm on the wood plate.

NVECTOR RESEARCH INCORPORATED

22

# Targeted Anomaly Detection

- Target detection finds unusual signal that looks like a target spectrum.
  - Targets are often from a library and may not correspond to a measured image
  - Requires atmospheric correction for remote sensing applications
  - Requires selectivity between targets for accurate detection and classification

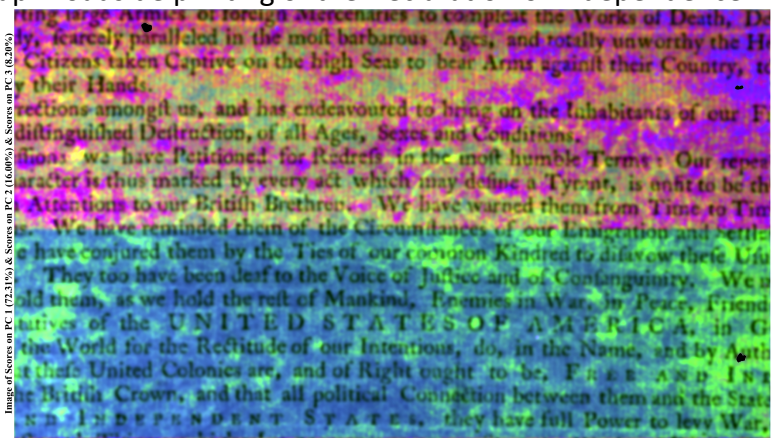
- Targeted Anomaly Detection
  - Start with library, anomaly or other knowledge of a target of interest
  - Use first detections as new target signal – it is image specific and relevant
  - Find unusual signal in an image likely related to the desired target

EIGENVECTOR RESEARCH INCORPORATED

23

# Hidden Watermark

Dunlap Broadside printing of the Declaration of Independence



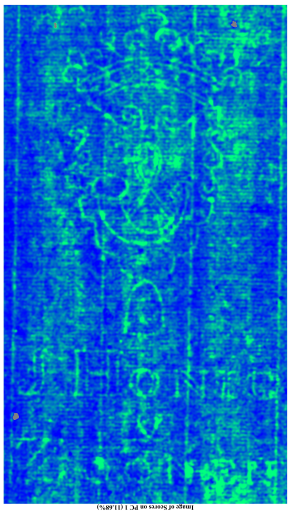

- Two major sources of clutter
- Hint of watermark used as the target

Thank you, Library of Congress
Meghan Wilson, Preservation Science Specialist
Fenella France, Chief of Preservation Research and Testing Division

EIGENVECTOR RESEARCH INCORPORATED

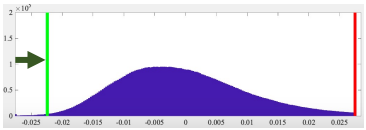
24

# Watermark



colormap winter
enhanced contrast
decluttering (1/α=100)
decluttering (1/α=10)
decluttering (1/ α=1)
decluttering (1/α=0.1)
no decluttering

- GLSW, mean-centering
  - 1/α larger is stronger de-cluttering
- Scores on PC 1
  - autocontrasted
    - mean-centered & saturated at ±2 std

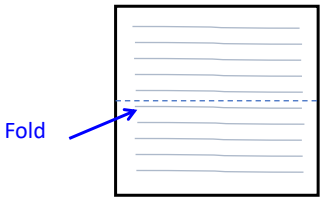


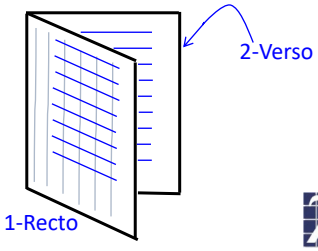
EIGENVECTOR RESEARCH INCORPORATED

25

# This 15th Century Palimpsest

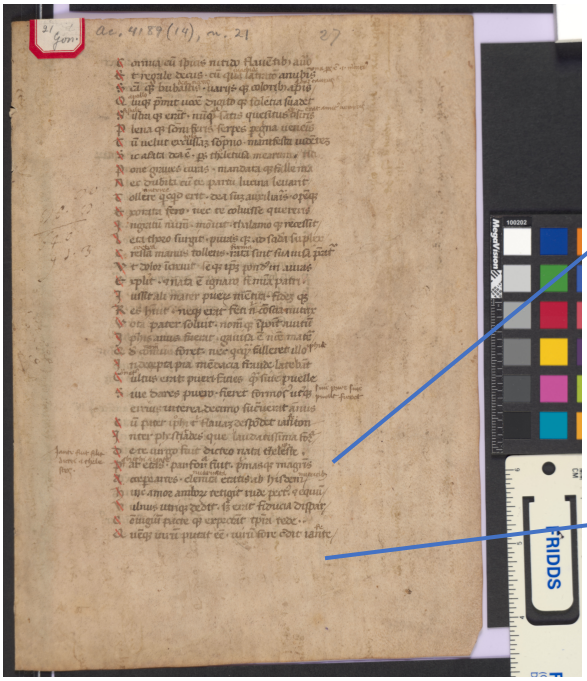
- Originally had writing on one side
- The ink was washed off
- The document was folded in half with the fold parallel to the original text yielding four sides for the new writing.
- The underwriting ended up on sides 1-Recto and 2-Verso, perpendicular to the overwriting.



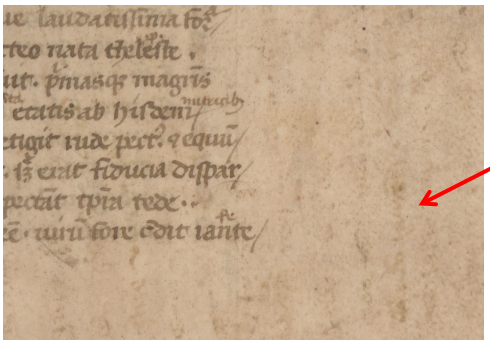




EIGENVECTOR RESEARCH INCORPORATED

26



# Hints of Older Writing

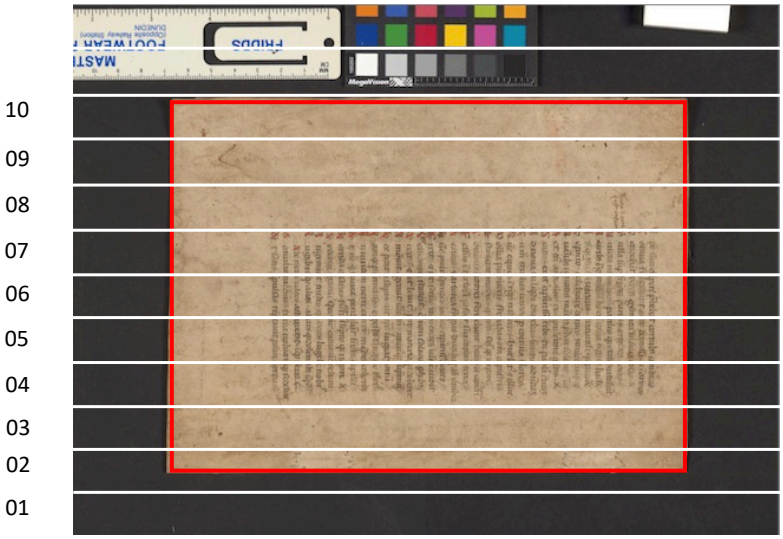


Thank you, Library of Congress
Meghan Wilson, Preservation Science Specialist
Fenella France, Chief of Preservation Research and Testing Division

EIGENVECTOR RESEARCH INCORPORATED

27

# Section into segments for Analysis




Analyzed segments 10, 09 & 08

Dahlberg, D., Gallagher, N., Wilson, M., France, F., "Revealing the Hidden Writing of a 15th Century Palimpsest Using Hyperspectral Imaging Analyzed by Principal Component Analysis and Generalized Least Squares Weighting", EAS15, Somerset, NJ, Nov 16-18 (2015). https://eigenvector.com/wp-content/uploads/2020/09/EAS15_Palimpsest-Analysis.pdf

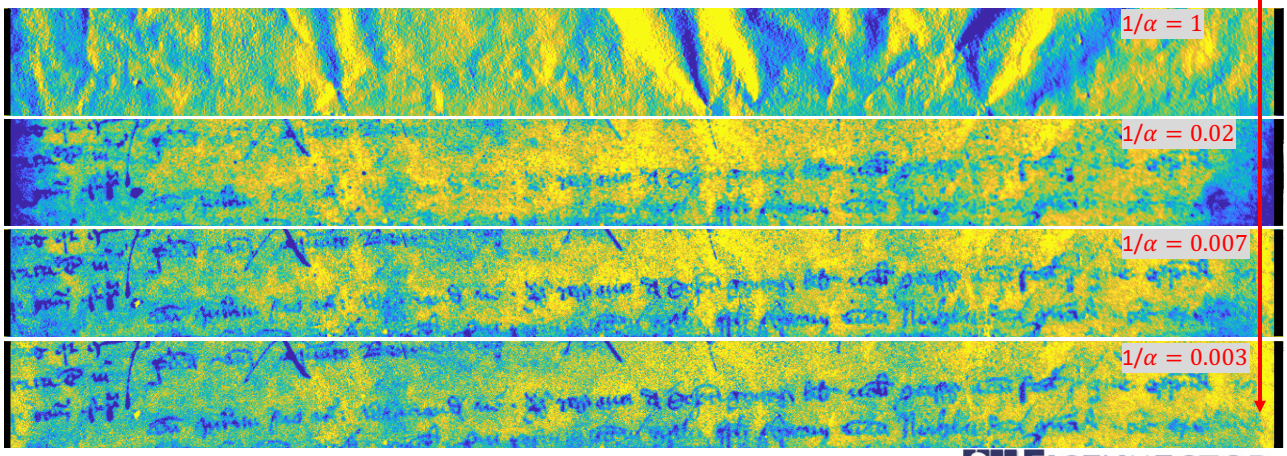
EIGENVECTOR RESEARCH INCORPORATED

28

# Section 10

Scores on PC 1

increasing de-weighting




EIGENVECTOR RESEARCH INCORPORATED

29

# Section 10

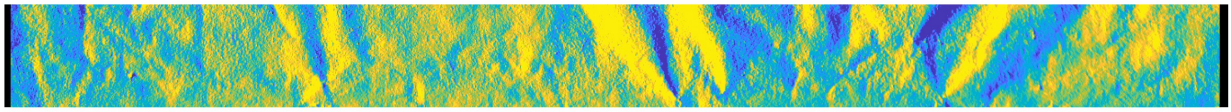
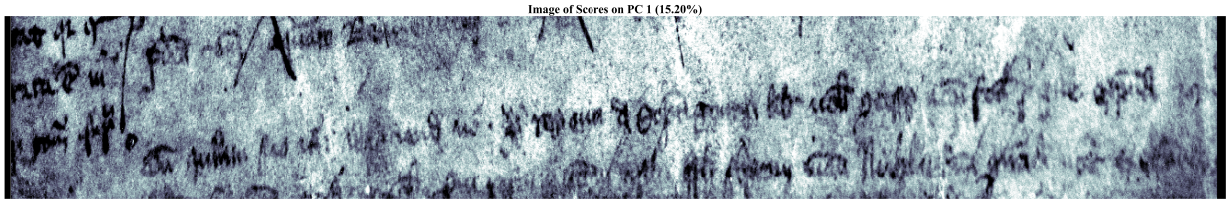


Image of Scores on PC 1 (15.20%)



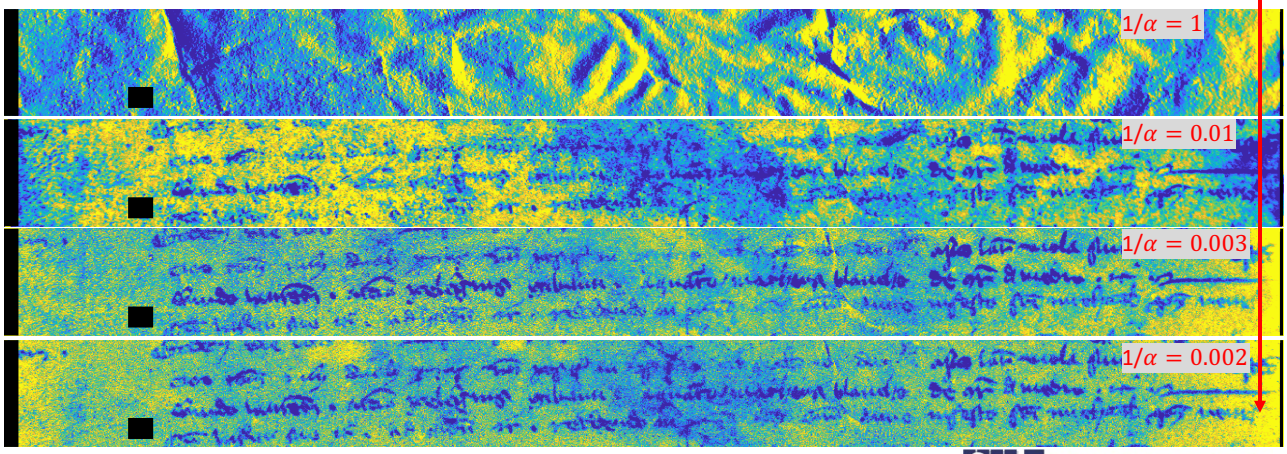
EIGENVECTOR RESEARCH INCORPORATED

30

# Section 9

Scores on PC 1

increasing de-weighting




EIGENVECTOR RESEARCH INCORPORATED

31

# Section 9

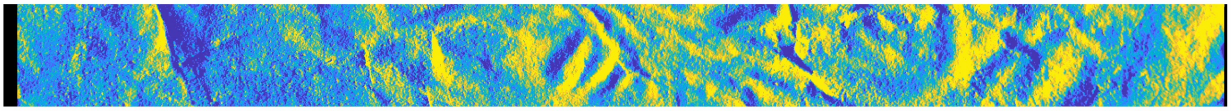
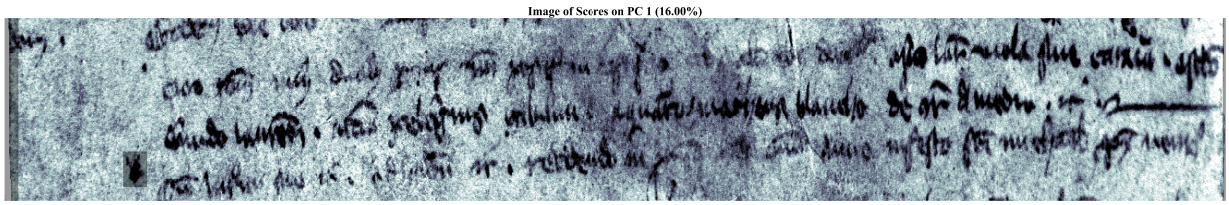


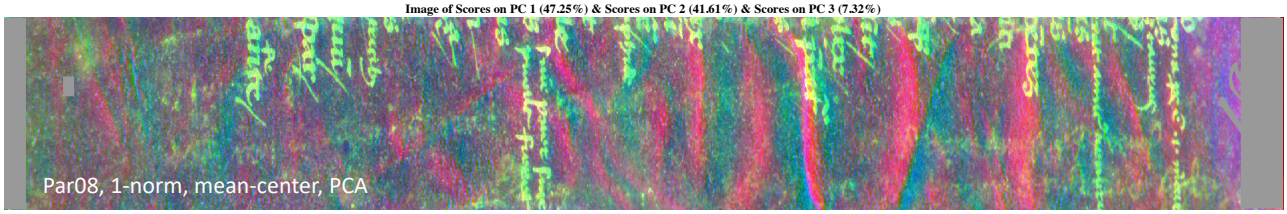
Image of Scores on PC 1 (16.00%)



EIGENVECTOR RESEARCH INCORPORATED

32

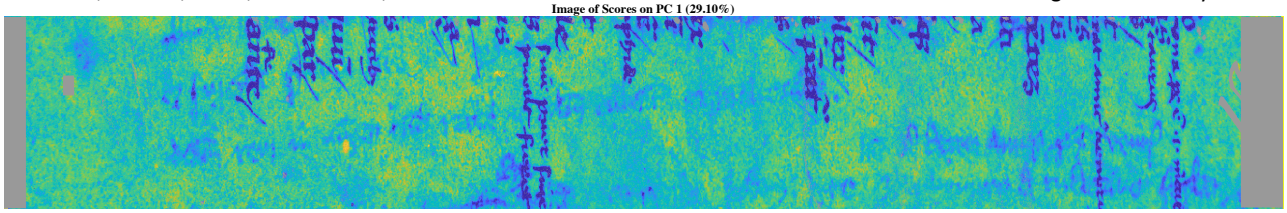
Image of Scores on PC 1 (47.25%) & Scores on PC 2 (41.61%) & Scores on PC 3 (7.32%)




Par08, 1-norm, GLSW, mean-center, PCA

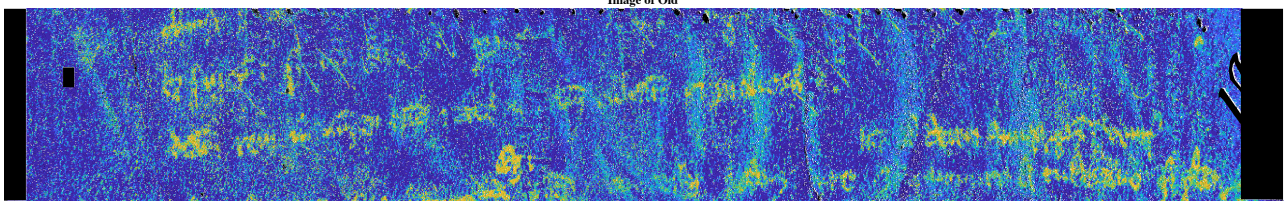
Section 8 – Some times things aren't as easy…..

Image of Scores on PC 1 (29.10%)



Par08, 1-norm, ELS, GLSW, NNLS, CLS (matched filter)

Image of Old



33

# De-Cluttering: GLSW Summary 1/2

- GLSW de-weights variations in the clutter class – e.g., paper signal.
  - Signal important to wrinkles, paper texture, uneven lighting (and overwriting) were de-weighted.
- The objective is to suppress signal not of interest (clutter) to enhance signal of interest
  - Maximize signal-to-clutter

EIGENVECTOR RESEARCH INCORPORATED

34

# De-Cluttering: GLSW Summary 2/2

- GLSW de-weighting works to remove spectral shape not magnitude
  - E.g., use row normalization to minimize magnitude effects
- The weighting finds 'directions' of correlated clutter
  - Due to interferences
  - Significant differences in noise between spectral channels
- The weighting is characterized with data that does not contain target signal

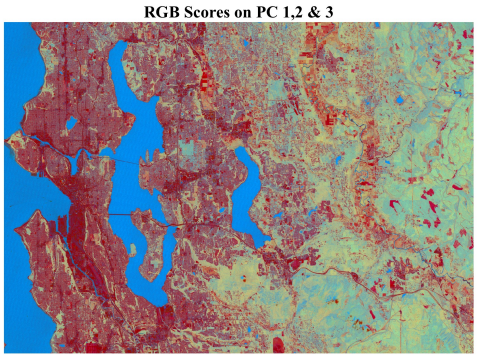
EIGENVECTOR RESEARCH INCORPORATED

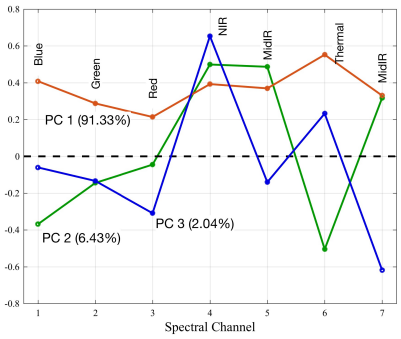
35

# PCA Scores and Loadings

Landsat of Seattle and surrounding area

**water signal is minor (it's in PC 3)**
**Can we discriminate Puget Sound from fresh water lakes?**



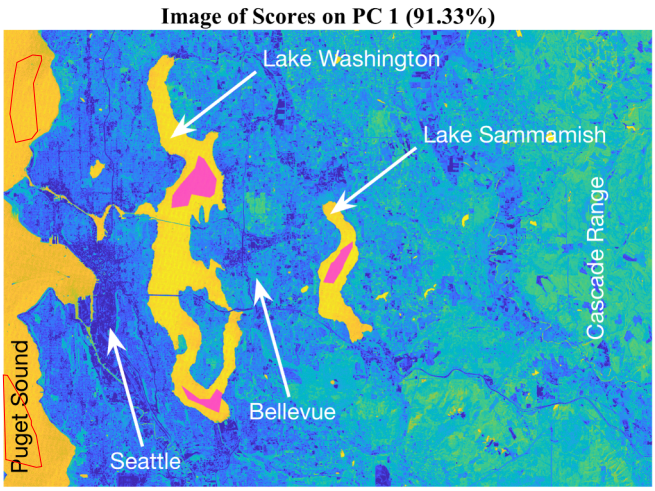




Seattle Landsat, USGS/NASA

EIGENVECTOR RESEARCH INCORPORATED

36

# Target Detection Example: Multiple Targets

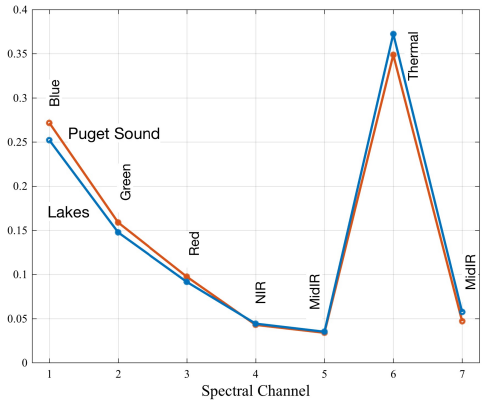



- Two-target GLS
  - Target: "Lakes" magenta
  - Interference: "Puget Sound"
- Extended Mixture Model + Generalized Least Squares
  - Can improve discrimination over the single target GLS approach

EIGENVECTOR RESEARCH INCORPORATED

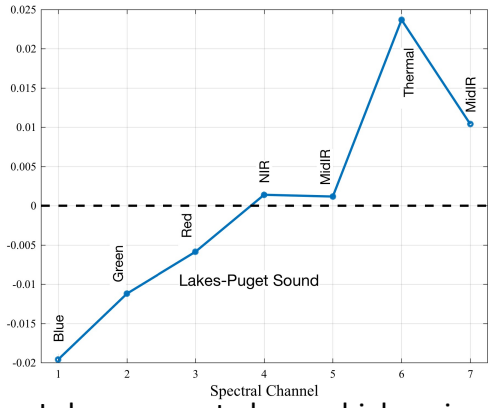
37

# Targets




The target signals are *very* similar

Differences between the two targets

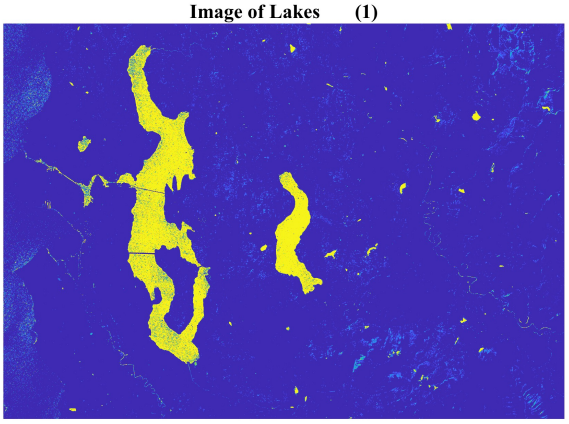



Lakes appear to have a higher signal on the Thermal channel relative to the Blue channel.

EIGENVECTOR RESEARCH INCORPORATED

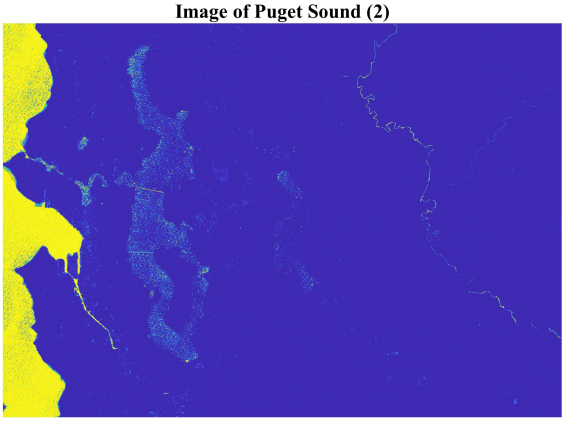
38

Image of Lakes (1)



Lakes

Image of Puget Sound (2)



Sound

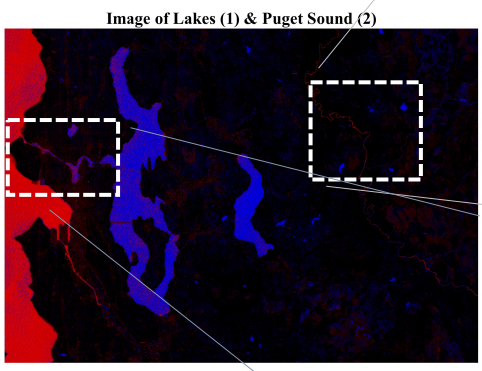
Lakes are discriminated from the Sound but ….

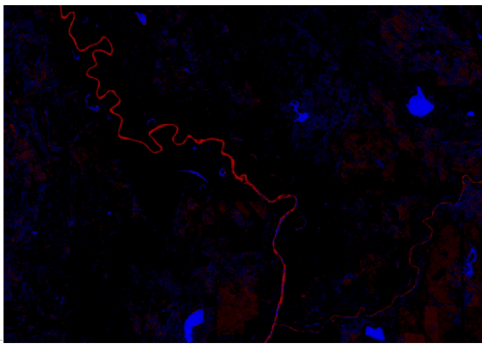
EIGENVECTOR RESEARCH INCORPORATED

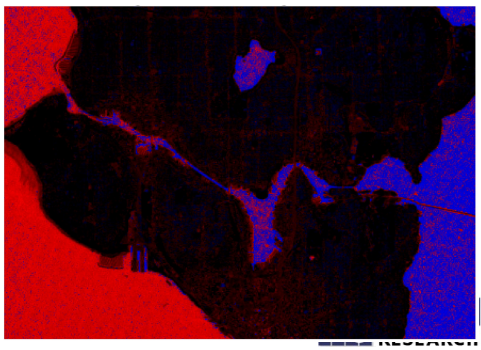
39

… It looks like temperature is a major driver for the difference.

Image of Lakes (1) & Puget Sound (2)







Discrimination is sometimes due to physics and not chemistry.

VECTOR INCORPORATED

41

# Conclusions

- The examples differed are in how "signal" and "clutter" are defined depending on the desired signal to enhance and signal to suppress
  - One man's clutter is another's target.
- Anomaly and Target Detection can be used synergistically to find signal of interest
  - signal that may be missed when used independently
  - the result is Targeted Anomaly Detection
  - very flexible approach to targeted image exploration

EIGENVECTOR RESEARCH INCORPORATED

42

# Additional Information

Myers, T.L., Johnson, T.J., Gallagher, N.B., Bernacki, B.E., Beiswenger, T.N., Szecsody, J.E. Tonkyn, R.G., Ashley M. Oeck, A.M., Su,Y.-F., Danby, T.O., "Hyperspectral Imaging of Minerals in the Longwave Infrared: The Use of Laboratory Directional-Hemispherical Reference Measurements for Field Exploration Data," *J Appl Remote Sens*, **13**(3), 034527 (2019). DOI: 10.1117/1.JRS.13.034527.

Gallagher, N.B., Shaver, J.M., Bishop,R., Roginski, R.T., Wise, B.M. "Decompositions with Maximum Signal Factors," *J. Chemometr.*, **28**(8), 663-671 (2014). DOI: 10.1002/cem.2634.

Gallagher, N.B., Kelly, J.F., Blake, T.A., "Passive Infrared Hyperspectral Imaging for Standoff Detection of Tetryl Explosive Residue on a Steel Surface," proceedings WHISPERS'10, Reykjavik, Iceland June 14-16 (2010), DOI: 10.1109/WHISPERS.2010.5594839.

Blake, T.A., Kelly, J.F., Gallagher, N.B., Gassman, P.L., Johnson, T.J., "Passive Detection of Solid Explosives in Mid-IR Hyperspectral Images," *Anal. Bioanal. Chem.* **395**(2), 337-348 (2009), DOI 10.1007/s00216-009-2907-5.

Gallagher, N.B., "Classical Least Squares for Detection and Classification," in Hyperspectral Imaging, Vol 32 (Data Handling in Science and Technology) 1st Ed, J. Manuel Amigo editor, Elsevier, 231-246, Oct 1 (2019) ISBN: 9780444639776. https://doi.org/10.1016/B978-0-444-63977-6.00011-0

Gallagher, N.B., Detection, Classification and Quantification in Hyperspectral Images using Classical Least Squares Models. In *Techniques and Applications of Hyperspectral Image Analysis*, Grahn, H.F.; Geladi, P., Eds. John Wiley & Sons: West Sussex, England, 2007; 181-201. ISBN: 978-0-470-01087-7.

EIGENVECTOR RESEARCH INCORPORATED

43