Determining the Factors that Form the Basis of the Relationship between Clinker Deposits and Archaeological Sites in Southeast Montana

Kate Clark
Jamie Hebert

Problem Statement

The ability to predict archaeological site location is not currently available in the Montana area. Recent research investigations have identified a possible correlation between Clinker deposits and archaeological site location.

Using GIS software spatial analytical capabilities, is it possible to **confirm this correlation between Clinker deposits and archaeological site location** and to introduce additional site and environmental factors into a multi-criteria model that will allow Cultural Resource Management (CRM) firms, government agencies and professional archaeologists predict where sites will be located, therefore reducing time and cost of future investigations?

Clinker Deposits

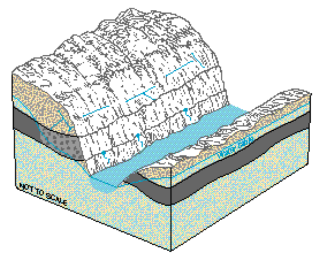


Figure 33. Clinker zones that result from burning of parts of coal beds are extremely permeable and are sources of springs in some places. Most clinker zones are unsaturated, but they form local aquifers where they are below the water table. Overlying sandstone beds may be fractured and provide conduits for recharge where the sandstone has subsided because the coal has burned.

http://capp.water.usgs.gov/gwa/ch_i/gif/I033.GIF

Archaeological Sites

http://www.blm.gov/heritage/docum/Fire/Dissertation_Buenger_files/image080.jpg

Site Type

Lithic Scatter with Stone Feature (LS_SF)

Lithic Scatter (LS)

Stone Feature (SF)

Site Subtype

Large Lithic Scatter (LLS)

Small Lithic Scatter (SLS)

Ring (R)

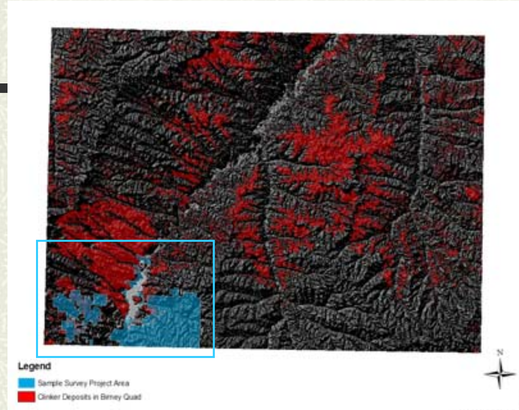
Cairn (C)

http://www.petroglyphs.us/a-csc_photo09_dummy_hunters_in_nevada.jpg

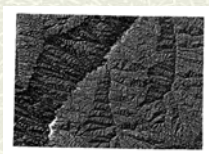
Study Area

Southeast Montana

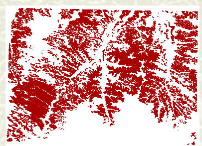
30’ X 60’
Quadrangle



GIS Layers



USGS DEM of 30’ x 60’ quadrangle of south central Montana

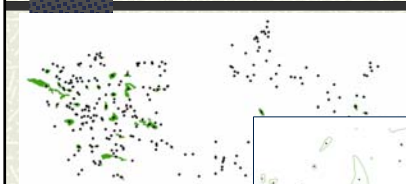


Clinker deposit polygon layer from the Montana Bureau of Mines

Archaeological site polygon layer from Ethnoscience, Inc.

Water layer from NRIS Montana State GIS Resources

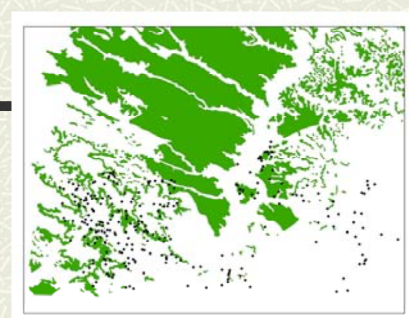
Methods



- Convert site polygons to points *(centroid)*
- Convert site points to raster

Site_Number	Type	Site_type	Site_subtype
2143/2246/2145	P	LS	SLS
24BH1005	P	LS_SF	SLS_C
24BH1006	P	SF	R

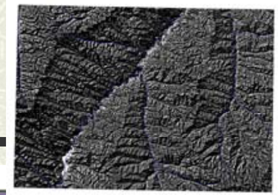
- Extract all prehistoric sites (Type P)
- Add Site_type and Site_subtype fields *(Edit session)*

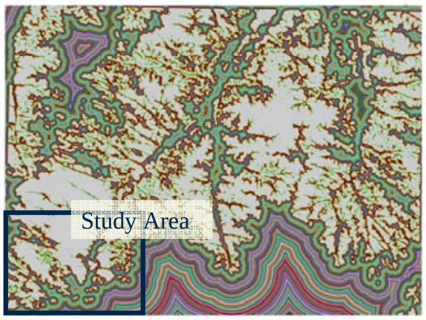


Affine Masking

- Add XY point location to sites
- Export to Excel spreadsheet
- Alter site locations
- Import altered XY site locations

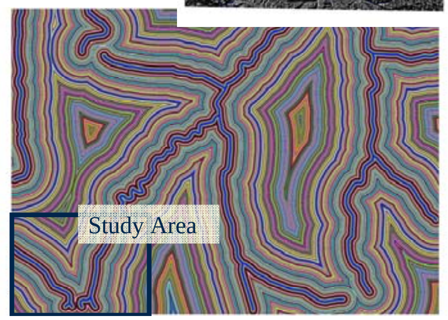
Euclidean Distance





Distance from Clinker deposits

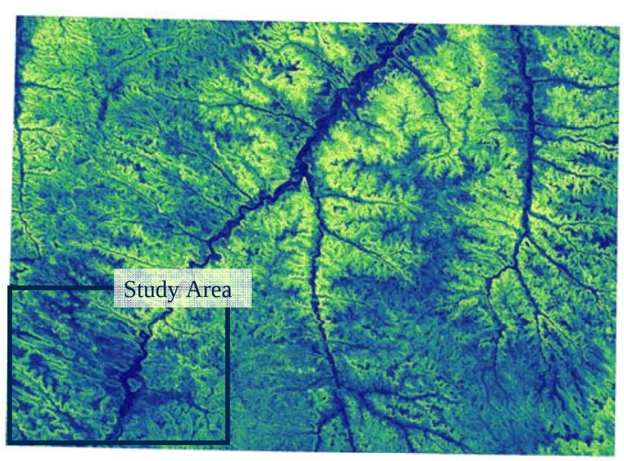
Using Raster Calculator (*Evaluate*), set extent to Study Area window



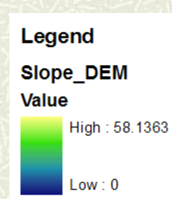
Distance from water

- *Reclass (Defined interval of 100 m increments)*
- *[Eucdist]/[sites_raster] = Sites_per_Band*
- *Convert raster to vector*
- *Spatial Join with site vector points*

Slope



- Reclass
- 10 increments



Frequency Ratio Model

Significance Levels

- >1.5 = 1
- 1.0 – 1.49 = 2
- 0.5 - 0.99 = 3
- 0.0 - 0.49 = 4

1. Raster pixel count for each distance band
2. Convert pixel count to m^2 and then convert to km^2
3. Determine % Area
4. Raster Calculator [Eucdist]/[site_raster] = Sites_Per_Band
5. Sites_Per_Band/ Total sites
6. Site_Freq/ Frac_A

DIST_BAND	PX_COUNT	AREA (KM SQ)	FRAC_A	SITES_PER_BAND	SITE_FREQ	FREQ_RATIO	SIG
0	175322	157790	26.2119	49	19	0.7	3
1	86760	78084	12.9713	62	24	1.8	1
2	63719	57347	9.5264	45	17	1.8	1
3	48395	43556	7.2355	28	11	1.5	1
4	32960	29664	4.9278	18	7	1.4	2
5	26820	24138	4.0098	9	3	0.9	3
6	19042	17138	2.8470	11	4	1.5	1
7	14971	13474	2.2383	3	1	0.5	3
8	12418	11176	1.8566	3	1	0.6	3
9	10242	9218	1.5313	2	1	0.5	3
10	9012	8111	1.3474	2	1	0.6	3
11	7481	6733	1.1185	2	1	0.7	3
12	7303	6573	1.0919	2	1	0.7	3
13	6654	5989	0.9949	1	0	0.4	4

50 Bands Total

259 Sites Total

Logistic Regression Model

SPSS Analysis

Create 259 Random Points

Dependent Variable: Site Presence (No = 0, Yes = 1)

Independent Variables: Euclidean distance from Clinker deposit (CD)

Euclidean distance to water

Slope (3x3 neighborhood)

$$Z = a + b_1X_1 + b_2X_2 + b_3X_3 + \varepsilon$$

a = .334

b_1 = -.012 (negative correlation) — X_1 = Distance from CD

b_2 = .022 (positive correlation) — X_2 = Slope

b_3 = .011 (positive correlation) — X_3 = Distance to water

Significant Variables

Distance from CD	.000
Slope	.029
Distance to water	.000

Logistic Regression Model

Logistic Regression Formula

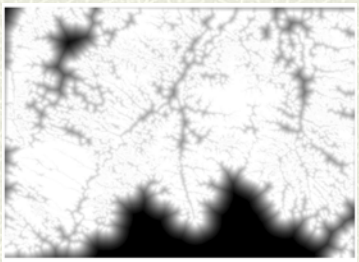
Raster Calculator

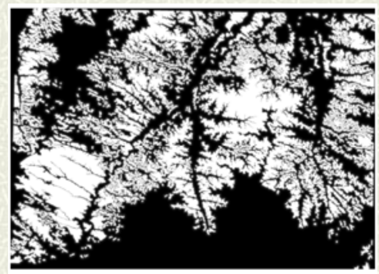
$$Z = .564 + -.008[Distance\_CD]$$

Estimated Probability of Site Occurrence

Raster Calculator

$$p = 1/(1+e^{-z})$$





Model Comparison

- Create 25 Bins
- Convert Site Frequency and p-value to 100 pt scale

BIN	Sites per BIN (Bands)	Sites per BIN (p-value)	CUM Bands	CUM p-Value	Percent Bands	Percent p-value
1	156	88	156	88	60	34
2	46	11	202	99	78	38
3	20	7	222	106	86	41
4	6	4	228	110	88	42
5	4	5	232	115	90	44
6	4	4	236	119	91	46
7	3	6	239	125	92	48
8	2	9	241	134	93	52
9	4	4	245	138	95	53
10	0	4	245	142	95	55
11	4	6	249	148	96	57
12	0	2	249	150	96	58
13	0	4	249	154	96	59
14	0	11	249	165	96	64
15	1	1	250	166	97	64
16	0	6	250	172	97	66
17	3	9	253	181	98	70
18	2	2	255	183	98	71
19	0	13	255	196	98	76
20	1	13	256	209	99	81
21	0	7	256	216	99	83
22	0	9	256	225	99	87
23	0	14	256	239	99	92
24	1	0	257	239	99	92
25	2	20	259	259	100	100

Model Verification

Frequency Ratio Model

Logistic Regression-best fit
(Frequency Ratio Model)

$y = 9.6346Ln(x) + 71.005$

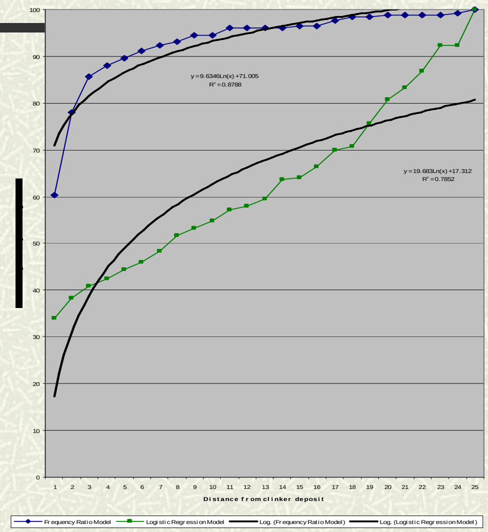
$R^2 = 0.8788$

Logistic Regression Model

Logistic Regression-best fit
(Frequency Ratio Model)

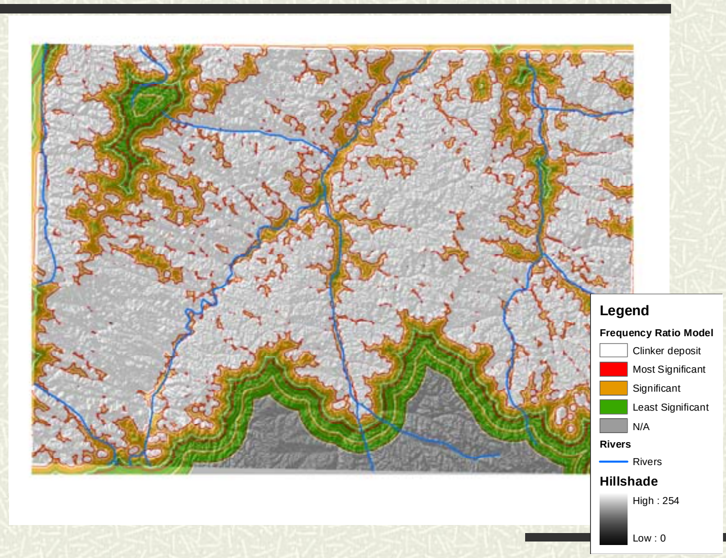
$y = 19.683Ln(x) + 17.312$

$R^2 = 0.7852$

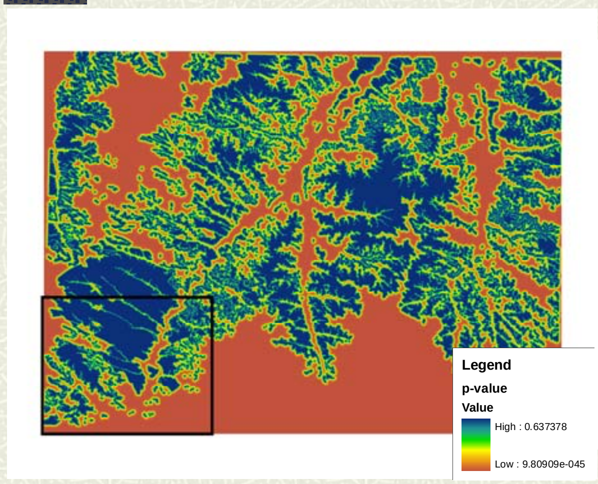


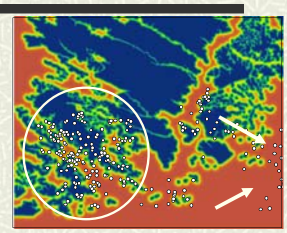
Predictive Frequency Ratio Model

Manual reclassification of Euclidean distance bands using Frequency Ratio Model Significance Values

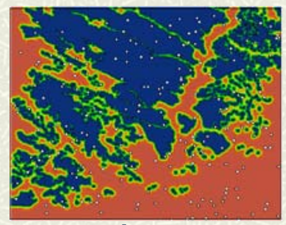


Predictive Logistic Regression Model



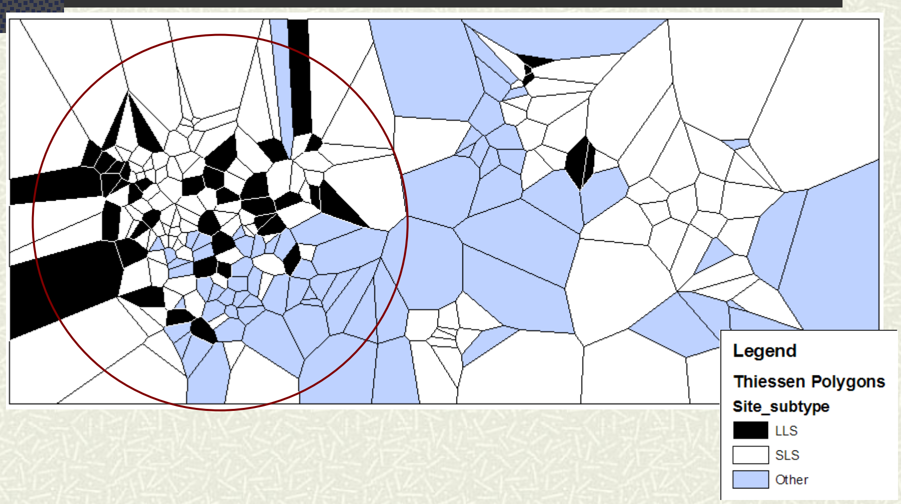


Masked Sites

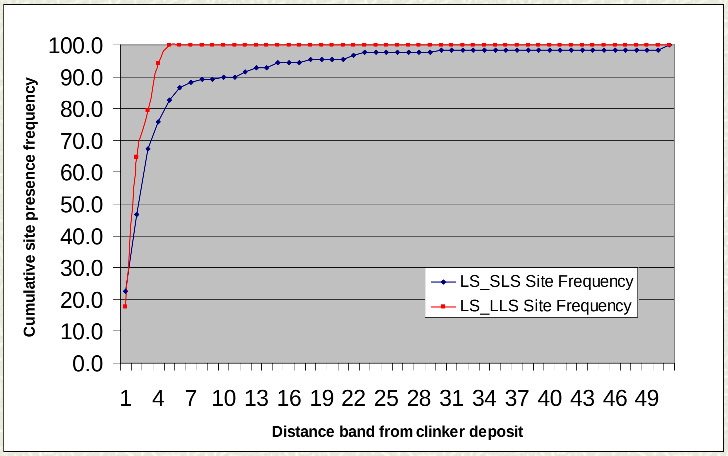


Random Points

Site Subtype Thiessen Polygons



Site Subtype Frequency Ratio Models

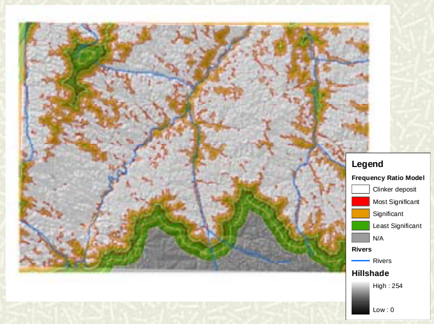


Assumptions/Errors

- Based on prior research, there is in fact a statistical correlation between archaeological sites and the Clinker deposits
- Assuming that the data from Ethnoscience, Inc., is accurate and precise
- Relatively small number of recorded archaeological site locations
- Centroids are a generalization of site location
- Slope is based on 3x3 neighborhood rather than zonal statistics
- Extent choice may produce an edge effect

Conclusions

- By comparing our models, we found that the Frequency Ratio Model is the best fit for our data.
- Future surveys in the area will determine whether this model is a viable predictor of site location.



References

- Ethnoscience Inc. (Project Director Scott Wagers)
- USGS 30m DEM <www.seamless.usgs.gov>
- Montana Bureau of Mines <http://www.mbmg.mtech.edu/gis/gis-datalinks.asp>
- Lee, S. and Pradhan, P. 2007. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. Landslides, 4: 33–41.
- NRIS Montana State GIS Resources <http://nris.state.mt.us/gis/gisdatalib/gisDataList.aspx>