ASSESSMENT OF SITE SUITABILITY FOR FRAC SAND MINING IN WINNESHIEK COUNTY

A synthesis of analysis

The University of Iowa
IOWA INITIATIVE FOR SUSTAINABLE COMMUNITIES

Smith, Austen J

Assessment of Site Suitability for Frac Sand Mining in Winneshiek County

1 INTRODUCTION & PROJECT GOAL

Winneshiek County, located in northeast Iowa, features a karst topography that gives the county its iconic natural beauty and landscape. This natural beauty, coupled with abundant cold water streams that support significant trout populations, bring many tourists to the area. When mining companies started to express interest in exploiting the easily accessible high quality quartz silica sand in the county, the county board of supervisors voted to place a moratorium on such mining activity to allow for adequate investigation and modeling of the impacts such activity would have. Due to significant industrial secrecy, the county does not have a complete understanding of where silica sand mining operations are likely to be located. Since location of likely mines is a critical component to conduct appropriate impact analysis, it is the goal of this project to determine where in the county such sand mining operations are likely to be located.

Winneshiek County has entered into a partnership with the Iowa Institute for Sustainable Communities (IISC) to conduct analysis of several current issues facing the county. The potential for large scale silica sand mining for use in hydraulic fracturing has received a high degree of public interest. Several departments at the University of Iowa are engaged with this project, conducting economic, legal, GIS, statistical, and engineering analysis of various questions surrounding the issue. The focus of this project is to integrate what has been learned from three different approaches to the question of where frac sand mines may likely be established once the moratorium on such activity is lifted. The spatial location of future mining operations is of acute interest since location has profound impacts on the inputs for other analysis.

To date three different attempts have been made to produce information to aid in prediction of future frac sand mining activity. Dr. David Bennett's GIS for Environmental Studies course conducted a laboratory exercise in which publically available data from the Iowa DNR was combined via a weighted linear combination scheme. The results of this analysis showed that the resulting likelihood predictions were quite sensitive to perturbations in the weights vector. Several assumptions about relative importance of component criteria were made due to the fact that no published transportation costs or other decision criteria could be found. The criteria weights used are arbitrarily defined, thus this analysis can be dubbed as preliminary at best.

In an attempt to more rigorously derive weights for the various model criterion, a group of graduate students who are a part of the Geoinformatics for Environmental and Energy Modeling and Prediction (GEEMaP) IGERT developed a spatially sensitive statistical model and derived beta weights from analogous data for the state of Wisconsin. The intent of this model was to improve upon the high degree of uncertainty in the criterion weights for each of the modeled criteria. After completing the first iteration of analysis, the derived relative importance of input criterions made intuitive sense with the exception of the relative weight of geology criteria. This likely occurred due to the fact that much of the study area upon which the model was developed in Wisconsin contained some form of minable sandstone. Therefore, increasing the weight of geologic criteria for Winneshiek County is advisable.

A preliminary analysis of the geological sampling point data available for Winneshiek County was conducted to get a better parameterization of the bedrock formations of interest. The preliminary analysis faced significant difficulty when interpolating the depth of the bedrock features of interest. This project extends the preliminary analysis of the geologic stratigraphic point data to improve upon earlier interpolation challenges. Through an alternative approach to geologic interpolation, two different derived data products are produced to be incorporated into a final predictive model.

Finally, an attempt is made to integrate aspects learned and improved upon from all three forms of analysis through an “informed linear combination of criterion” method. With this final method, the flexibility and option for subjective component weights is taken from the weighted linear criterion model. The degree of importance and the relative relationship amongst criterion constraints is taken from the Bayesian statistical model coefficients. Finally, significantly enhanced geological data is interpolated and the two resulting derived data products are used as new criterion components for final mine location predictions.

2 Interpolation and Criterion Synthesis - Literature Review

2.1 Bedrock Interpolation

A brief overview of the literature was conducted to understand previous approaches and expert knowledge in the area of bedrock point data interpolation. The overarching theme from the overview conducted revealed a broad range of manual, semi-automated, and automated interpolations schemes. In addition, there were many cases where the interpolation method used was highly dependent upon the available input data quality and format.

Drawing from an era before mature computerized geospatial data analysis, Horsman el al discusses how geologic characterization of a smaller site is conducted from sample point data (Horsman & Bethel, 1995). He discusses the importance of a 3-D model of geology but the methods by which interpolation is conducted are quite simple with linear or hand drawn interpolation. Also from an older era, Shaw discusses using a beta spline curve approximation method to mathematically parameterize the thickness variation between sections (Shaw, 1978).

While methods ranging from basic to mathematically complex were conducted prior to mature digital mapping analysis capabilities, since the advent of this technology others have developed more computationally intensive methods. Patel et al describe a 2D algorithm to interpolate well log data using basic inverse distance weighting or linear interpolation but also using seismic information to detect abrupt changes (Patel & McMechan, 2003). Most recently, Calcagno et al discuss an original method using a cokriging method to interpolate continuous three dimensional geology features. (Calcagno, Chiles, Courrioux, & Guillen, 2008). One key component of their method is the two dimensional potential-field which identifies interface locations between bedrock layers. Finally, Kaufmann et al make an effort to develop a model that is flexible and capable of using publically available data in a variety of forms (Kaufmann & Martin, 2008). Using the computational resources of modern digital mapping, he derives an approach that combines information from geological maps, cross-sections, digital elevation models, and borehole and outcrop descriptions.

2.2 Criterion Synthesis

The literature contains numerous examples of GIS applications to derive site suitability for a wide range of activities. During the quick review of the literature, it appeared that informed weighted linear criterion coupled with "expert knowledge" is often used in application. There are certainly examples of attempts to develop a more analytically robust method but examples of successful application of these methods are limited.

Charabi et al use a fuzzy multi-criteria evaluation method that really is extremely similar to a weighted linear criterion approach (Charabi & Gastil, 2011). The primary difference between the two methods is the fuzzy method Charabi describes integrates both Boolean numerically scaled criteria and has advantages for sensitivity analysis. Also Vasiljevic et al discuss a similar method of criterion synthesis for determining suitable placement of a landfill site in Serbia using experts from different fields and an analytic hierarchy process to derive appropriate weights (Vasiljevic, Srdjevic, Bajcetic, & Miloradov, 2011). It is clear that from this review that linear weighted criterion analysis is commonly used and accepted for use in multi-criterion decisions.

3 Geologic Interpolation Revisited

As mentioned in the project overview, one of the greatest shortcomings of past attempts to predict locations for frac sand mining in Winneshiek County was due to limited information about sandstone geology characteristics in the county. However, geological data available from the Iowa DNR is of superior quality compared to that available for other neighboring states. The lack of high quality geology data for the state of Wisconsin resulted in unintended assumptions in the original model developed by the GEEMaP group (Cowles, et al., 2014). Even though there is higher quality data easily available for Iowa, the previous linear criterion analysis only used readily available GIS layers. Specifically, a depth to bedrock raster and a shape file containing polygons of top bedrock layer type were used without consideration of their quality. While this analysis was a good first cut, there are several aspects that can be improved upon.

First, by masking the estimation area to only those locations within the St. Peter geology polygon, this assumes both that the polygon is accurate and that mining companies are unwilling to mine sites in which the formation of interest lies beneath another layer. While the depth to bedrock raster can effectively approximate the depth to the St. Peter formation for areas in which St. Peter is the top layer of bedrock, this is not possible in areas where this is not the case. Also, in previous models, no input for thickness of the St. Peter sandstone is used. This information is critical as the mining companies cost for overburden removal must be directly weighed against the economic benefit potential of extraction of the sand resource. Certainly a major component of this calculation would be the quantity of frac sand available.

To address these issues, water well and other geo referenced drilling point data in the Iowa GEOSAM database was extracted, analyzed, and interpolated. This dataset contains top and bottom elevations for each bedrock type encountered during drilling. From this information, a "depth to St. Peter" as well as a "thickness of St. Peter" raster was derived. The processing steps along with the interpolation technique used for the derivation follow.

First, Iowa stratigraphic point data is obtained from the GEOSAM database via the Iowa DNR with a 10 km buffer around the boundary of Winneshiek County. This buffer is necessary to minimize degradation of interpolation along the county boundaries. All sample point data lying within the boundary is selected for analysis. Unfortunately, since Winneshiek County shares a northern border with Minnesota, sampling data from the Iowa GEOSAM database is not available outside the state of Iowa. The Minnesota Geological Survey maintains a digitized well locations point dataset. However, further investigation into this dataset revealed that it does not contain adequate information to derive upper and lower elevations for each layer. Therefore, this analysis is limited to Iowa GEOSAM sampling points falling within the 10 km buffer of Winneshiek County with no buffer on the northern state boundary border.

The GEOSAM data is comprised of a single table with an entry for each stratigraphy unit at each location. Thus, if a single well drilled through five different formations, there would be five rows in the dataset corresponding to this site, each coded with identical location coordinates, but different stratigraphy for each row. An example of the data table structure with St. Peter stratigraphy units highlighted is shown below in figure 1.

Wnumber	Owner	Utm83_x	Utm83_y	Elevation	Bedrockdep	TotalDepth	DepthTop	DepthBot	StratUnit	UnitElev	UnitThick	Lithology
124	Decorah, City Of	597598.39	4794475.08	877	38	1250	0	18	Wisconsinan Ep.	877	18	sand and gravel
124	Decorah, City Of	597598.39	4794475.08	877	38	1250	18	36	Platteville	859	18	limestone
124	Decorah, City Of	597598.39	4794475.08	877	38	1250	36	39	Glenwood	841	3	shale
124	Decorah, City Of	597598.39	4794475.08	877	38	1250	39	145	St. Peter Ss.	838	106	sandstone
124	Decorah, City Of	597598.39	4794475.08	877	38	1250	145	201	Readstown	732	56	shale
124	Decorah, City Of	597598.39	4794475.08	877	38	1250	910	1033	Eau Claire	-33	123	shale
124	Decorah, City Of	597598.39	4794475.08	877	38	1250	1033	1250	Mt. Simon	-156	217	sandstone
181	Bakke	603812	4791597	1170	28	3300	0	28	QUATERNARY	1170	28	
181	Bakke	603812	4791597	1170	28	3300	250	288	Decorah	920	38	shale
181	Bakke	603812	4791597	1170	28	3300	288	310	Platteville	882	22	limestone
181	Bakke	603812	4791597	1170	28	3300	310	315	Glenwood	860	5	shale
181	Bakke	603812	4791597	1170	28	3300	315	405	St. Peter Ss.	855	90	sandstone
181	Bakke	603812	4791597	1170	28	3300	1240	1650	Mt. Simon	-70	410	sandstone
181	Bakke	603812	4791597	1170	28	3300	1650	3300	PRECAMBRIAN	-480	1650	sandstone
288	Iowa Conservation Commission	599079	4791966	876	15	185	0	15	QUATERNARY	876	15	
288	Iowa Conservation Commission	599079	4791966	876	15	185	15	40	Dunleith	861	25	limestone
288	Iowa Conservation Commission	599079	4791966	876	15	185	40	85	Decorah	836	45	limestone
288	Iowa Conservation Commission	599079	4791966	876	15	185	85	120	Platteville	791	35	limestone
288	Iowa Conservation Commission	599079	4791966	876	15	185	120	185	St. Peter Ss.	756	65	sandstone

Figure 1

Next, a selection was performed to choose all points containing elevation data for the St. Peter formation. This selection resulted in 368 sample points that contain the St. Peter formation. With this new subset of St. Peter data points, the unit elevation column provides a measure of the upper St. Peter elevation at each sample location. This measurement is more useful for interpolation than the "BedrockDepth" variable because it is less sensitive to changes in surface topology. With an upper elevation of the St. Peter formation, a depth below the surface can easily be derived with a high resolution digital elevation model of the county. A Kriging method of statistical interpolation included in the geo-statistical analyst toolbox is then used with the upper elevation value to interpolate the top elevation of the St. Peter formation. For this particular interpolation, simple normal score Kriging without any order trend removal was used. In addition, ten-fold cross validation was used to tune the interpolation parameters. The semivariogram and error plot associated with the interpolation are shown below in figures 2 and 3.

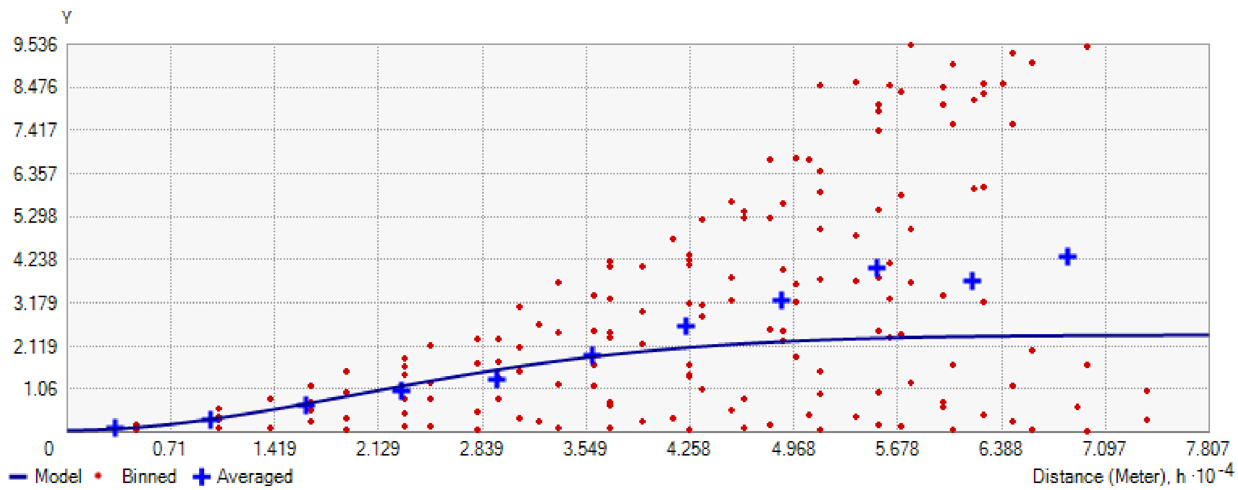

Figure 2

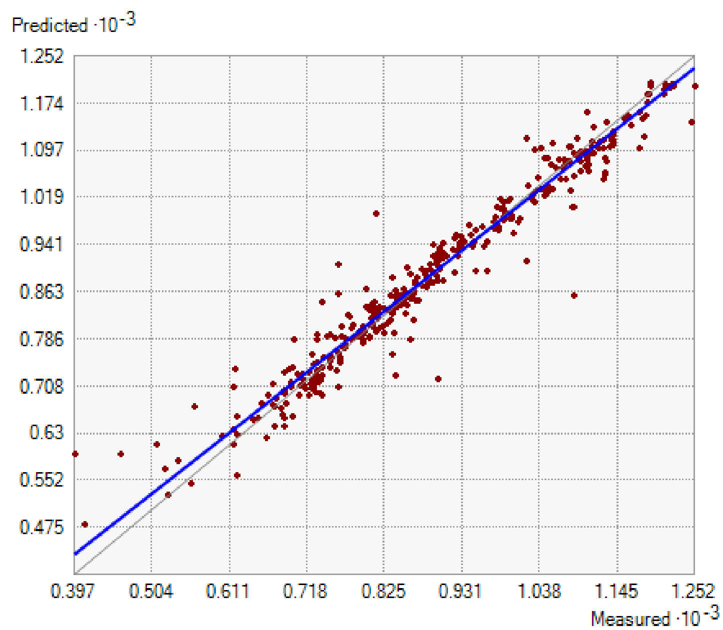

Figure 3

Based on the error plot, it is reasonable to state that the interpolation technique adequately captures the variability in the data and produces a useable surface. The resulting RMS error for this interpolation is 39.09 feet. While this amount of error is certainly not ideal, it is the best approximation we can obtain from the current sampling density. It must also be noted that this is a global error metric therefore, the error changes based on proximity to sample points. Below figure 4 the resulting derived depth to St. Peter map (depth below the surface) for the entire county. Depths are reported in feet.

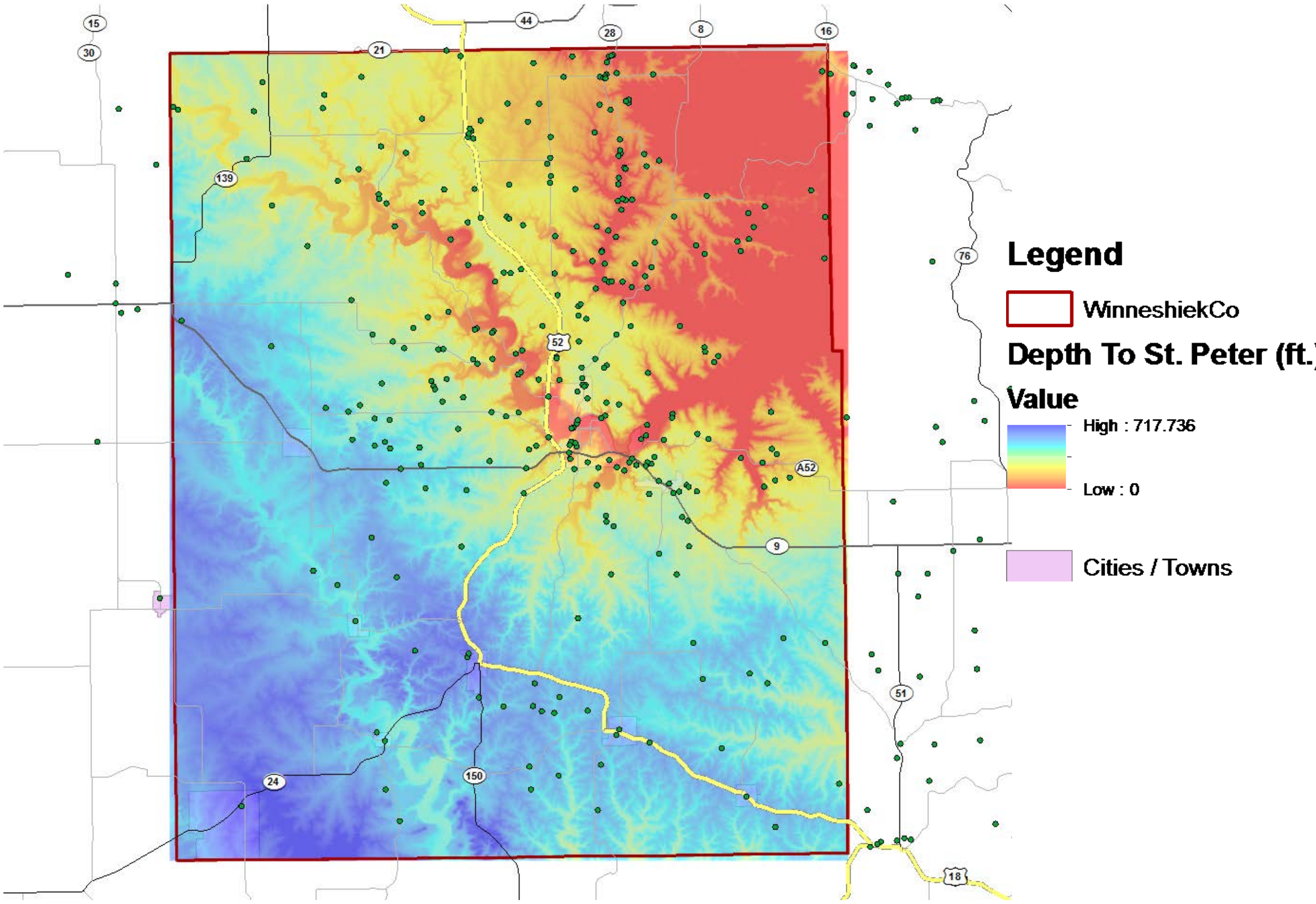

Figure 4

The economic viability of a particular mining operation is largely dependent upon the amount of overburden that must be removed to extract the sand. Based on review of available published data, analysis of existing silica sand mining operations, and discussion with miners, there seems to be a fairly universal rule of thumb that the overburden must be 50 feet or less in order to be profitable. Based on this information, raster pixels with a depth to St. Peter of 50 ft. or less are selected. The resulting area is shown in red in figure 5 below.

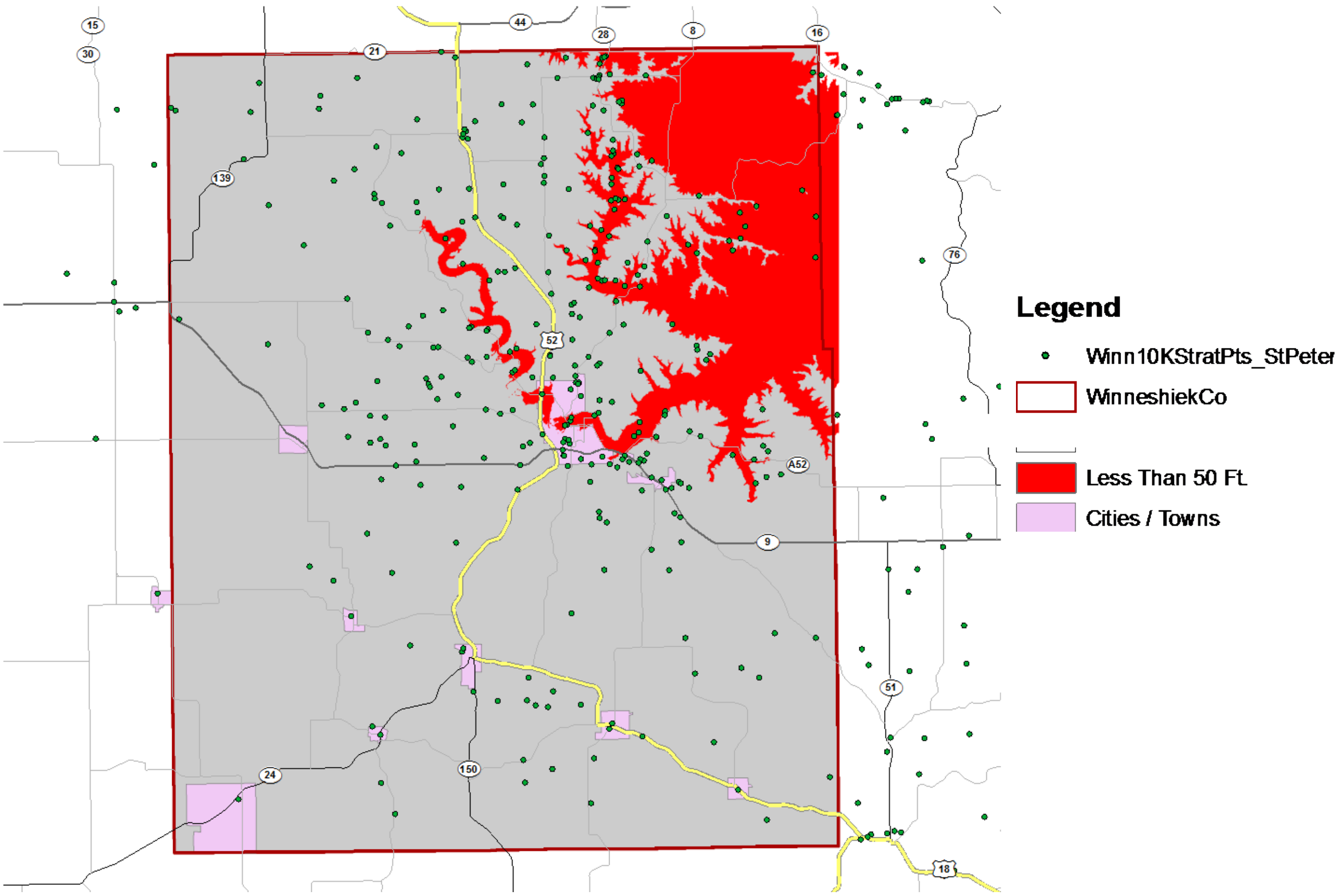

Figure 5

The long tail that extends to the left is the bottom of a riverbed and thus is not suitable for mining. Once this tail is removed, the resulting region is the area of most likely interest to frac sand mining companies. Figure 6 shows the region of interest colored by depth to the St. Peter layer.

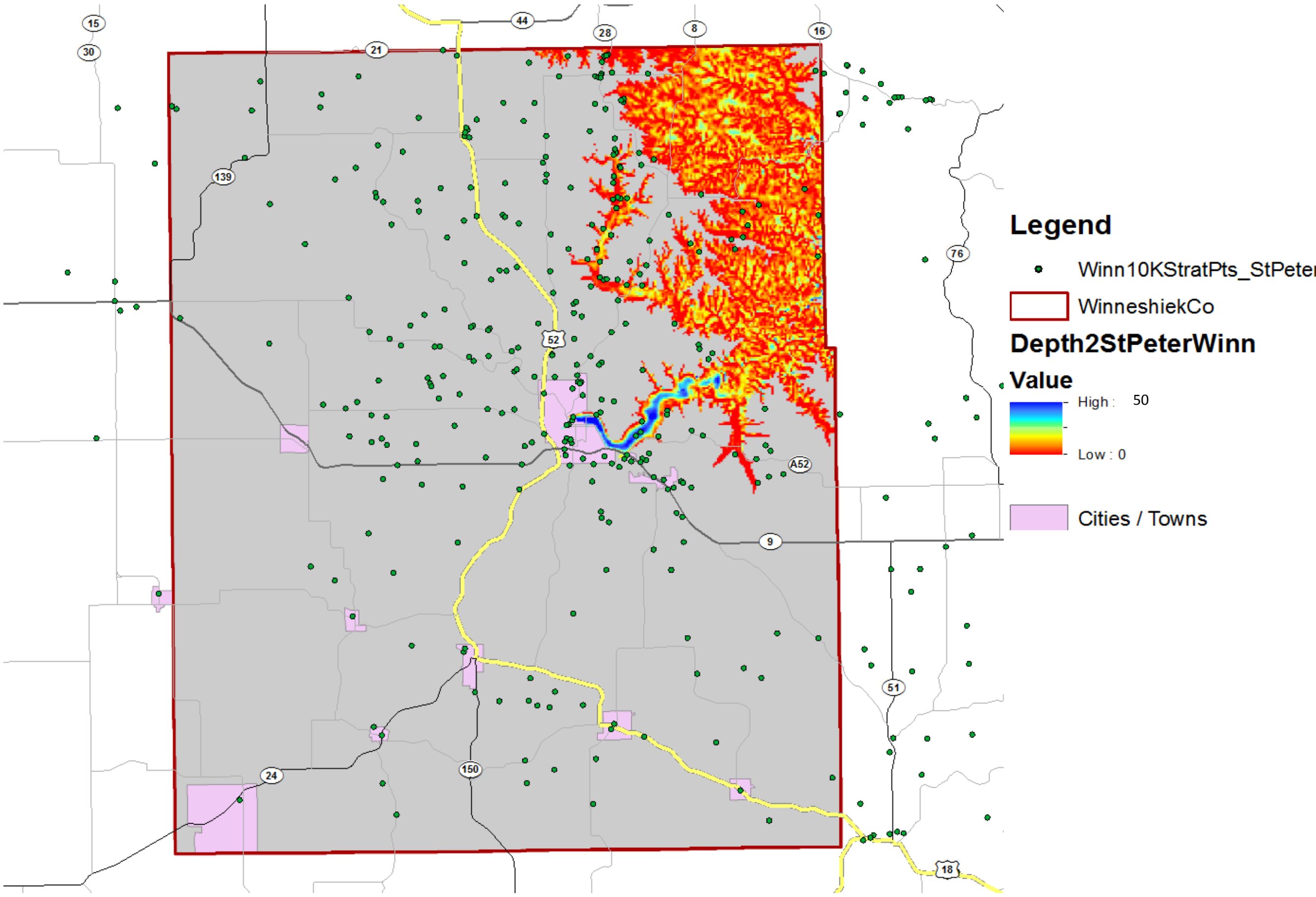

Figure 6

Next in order to get a reasonable estimate of the thickness of the St. Peter we needed to do more than simply interpolate the "UnitThickness" variable in the dataset. Any attempt to interpolate this variable reveals very low spatial autocorrelation. Upon close inspection of the dataset, many of the St. Peter sample points had lower elevations equal to the minimum elevation of the drilling. In other words, the well was not drilled all the way through the St. Peter layer, thus the lower elevation recorded was artificially high. In order to limit the effects of these points on the interpolation, a selection is performed to remove all points for which the St. Peter layer was not fully penetrated. This process removed 167 points, leaving 201 remaining for interpolating the lower depth. The distribution of complete and incomplete drillings is shown in figure 7 below.

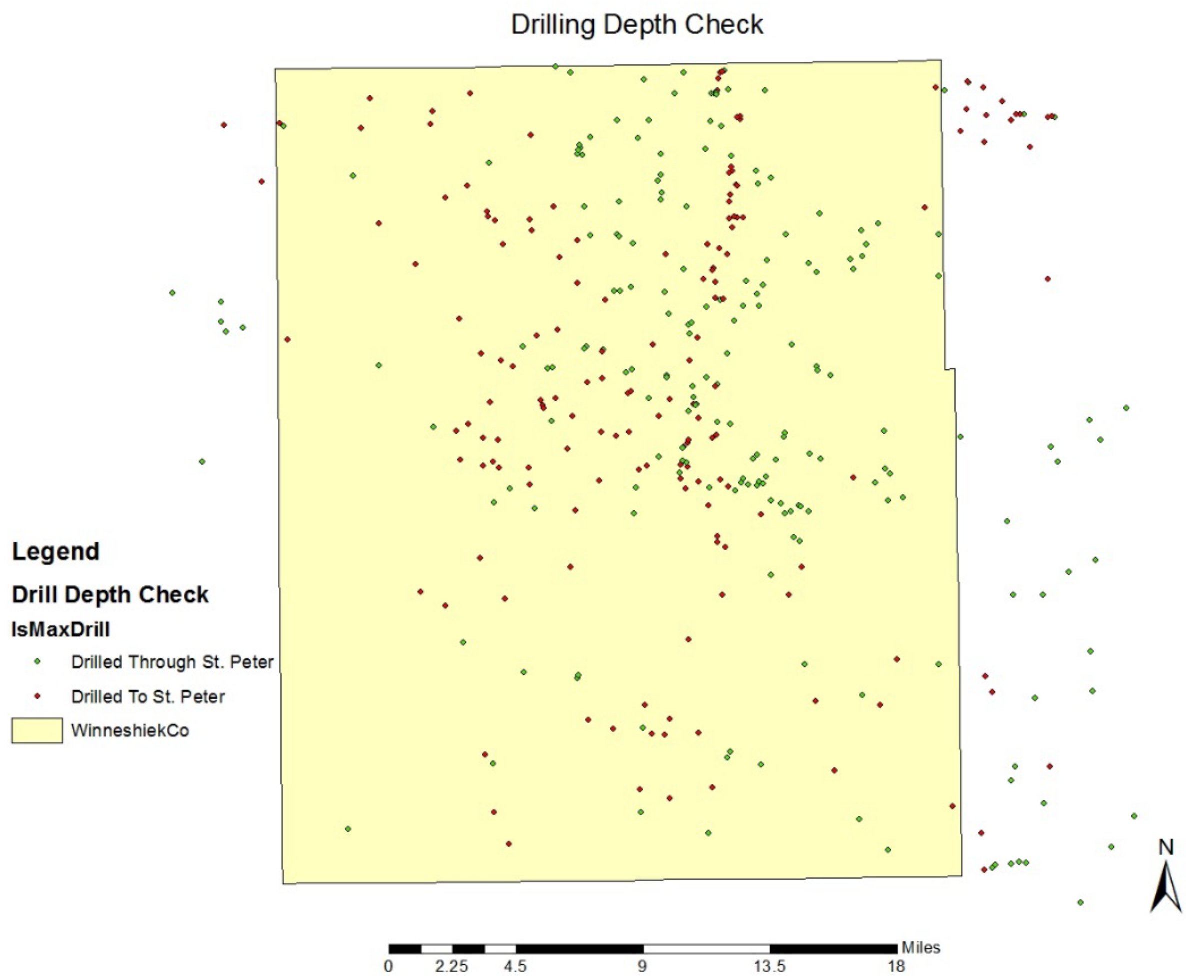

Figure 7

Even when the points without full penetration are removed, direct interpolation of the thickness values is not useful since there is very little spatial autocorrelation in these values. Next an attempt was made to use the lower elevation of the St. Peter points to interpolate a lower elevation using only points with full penetration. With a lower elevation interpolation, a thickness layer can be derived by taking the difference between the upper and lower elevation surfaces. The lower elevation measurements do exhibit a sufficient degree of spatial autocorrelation to make interpolation viable. However, after deriving the thickness layer, the predicted thickness was compared to known thicknesses at sample locations. This analysis revealed some significant discrepancies especially around the northeast corner of interest.

These discrepancies prompted discussions with geology expert, Dr. Emily Finzel, who helped formulate the next methodological iteration for deriving the St. Peter thickness. Upon closer analysis of the error at sample locations, the problem appeared to be linked to variation in the upper St. Peter elevation not captured by the interpolated upper elevation layer. This problem appeared to occur most in regions of lower surface elevations with high bluffs nearby. Dr. Finzel suggested that a layer's exposure to erosion could be assessed based on the stratigraphy unit lying directly on top of the St. Peter sandstone. Figure

8 below shows the region of interest, as well as the distribution of sample points. Those containing St. Peter sandstone are highlighted in green. The fact that many of the lower elevation sample points within the region of interest do not contain St. Peter suggests that this layer exists in the surrounding bluffs but perhaps has eroded away in the lower lying areas. Because we see many sandstone outcrops identified in this area it is reasonable to suppose that they are the result of erosion.

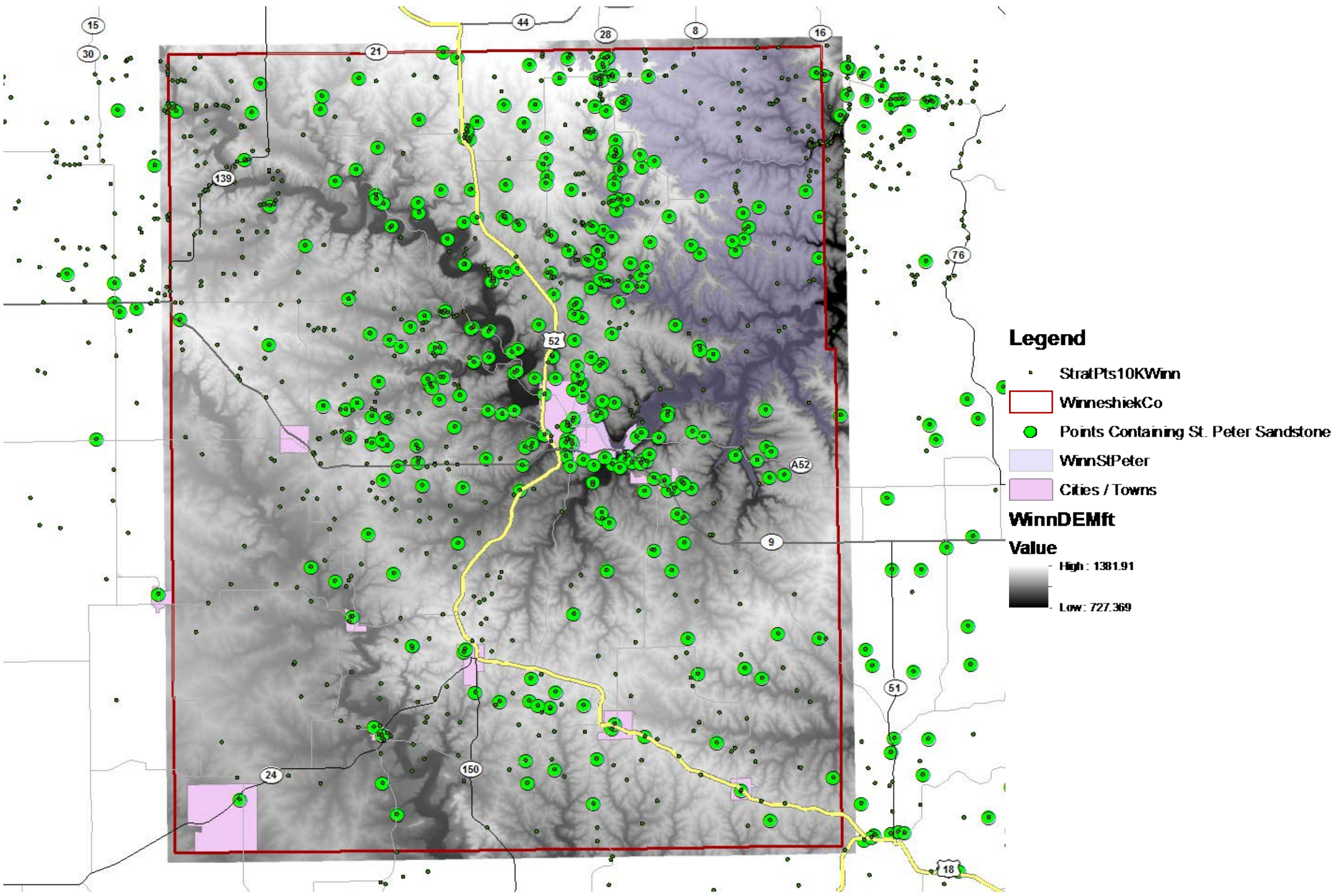

Figure 8

In order to assess exposure to erosion at each sample location, the data was restructured to include an adjacent unit column in which the stratigraphy unit directly on top of each St. Peter sample point was populated. If the unit directly above the St. Peter sandstone at a sample location is a glacial till or there is nothing above the St. Peter layer, this suggests that the St. Peter layer has likely been exposed to erosion. By selecting only points for which the unit directly above the St. Peter suggests limited erosion, a digital elevation model can later be used to correct for erosion effects. A visual diagram of this effect along with a count of the stratigraphy unit directly on top of each of the St. Peter sample points is shown in figures 9-11 below. The stratigraphy units highlighted in figure 9 in yellow are those identified by Dr. Finzel that suggest limited or no exposure to top level erosion.

Unit Code	Occurrence Count
Glenwood	206
Plattville	70
Dec./Platte./ Gllnwd./Undiff.	29
Decorah/Platteville	11
Dunleith	1
Harmony Hill	3
Noah Ck.	2
Peoria	3
QUATERNARY	14
Wisconsinan Ep.	6
Nothing on Top	32

Figure 9

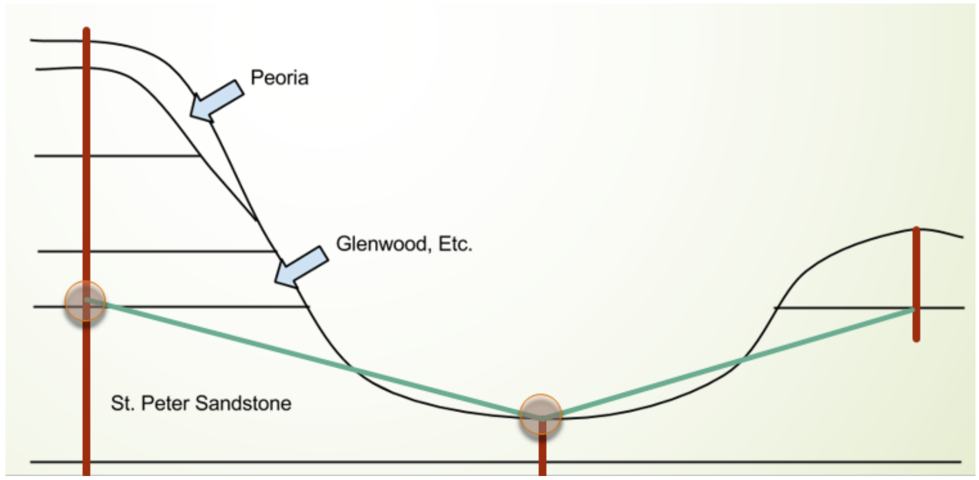

Figure 10 - The well drilling in the middle of the valley shown above has been exposed to errosion and causes noise in the upper elevation data interpolation. Such noise lowers local spatial autocorrelation and can degreade results.

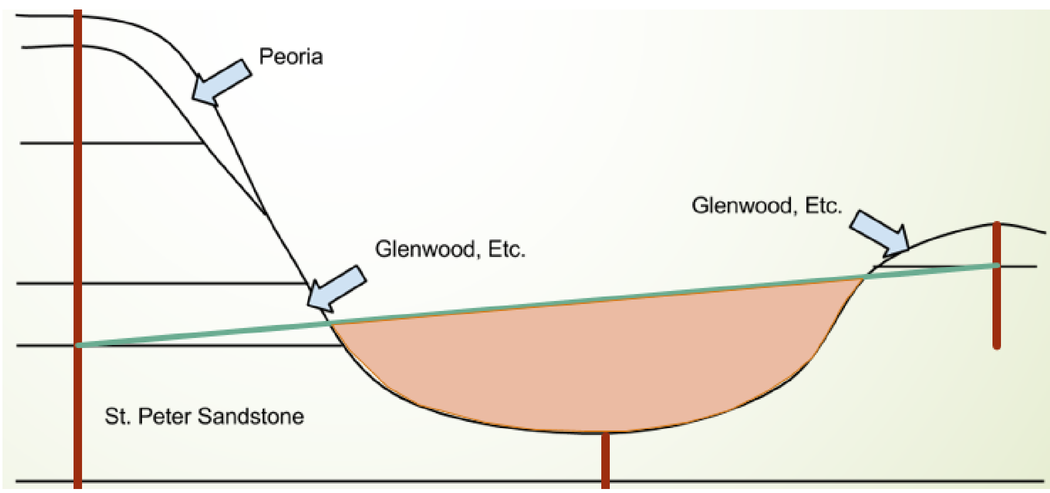

Figure 11

When the upper St. Peter elevation recorded in the center of the valley is removed from the interpolation, the upward trend in the layer elevation is better captured than with the valley point included. The shaded region between below the green line (in figure #) and the surface elevation can easily be reconciled with the use of a digital elevation model.

Using this approach, an upper elevation of St. Peter is interpolated using only sample points for which the stratigraphy unit directly above is highlighted in yellow in the table above. Next this upper elevation layer is compared to the surface elevation for each raster cell. If the interpolated upper St. Peter elevation is greater than the surface elevation, the upper St. Peter elevation is updated to the surface elevation. The lower elevation of St. Peter is interpolated from the sample locations where the St. Peter was fully penetrated. Finally, the lower elevation is subtracted from the upper elevation to create the final St. Peter thickness prediction. The predicted thickness of the St. Peter along with the observed thickness at sampled locations are shown by figure 12 below. Thicknesses are reported in feet.

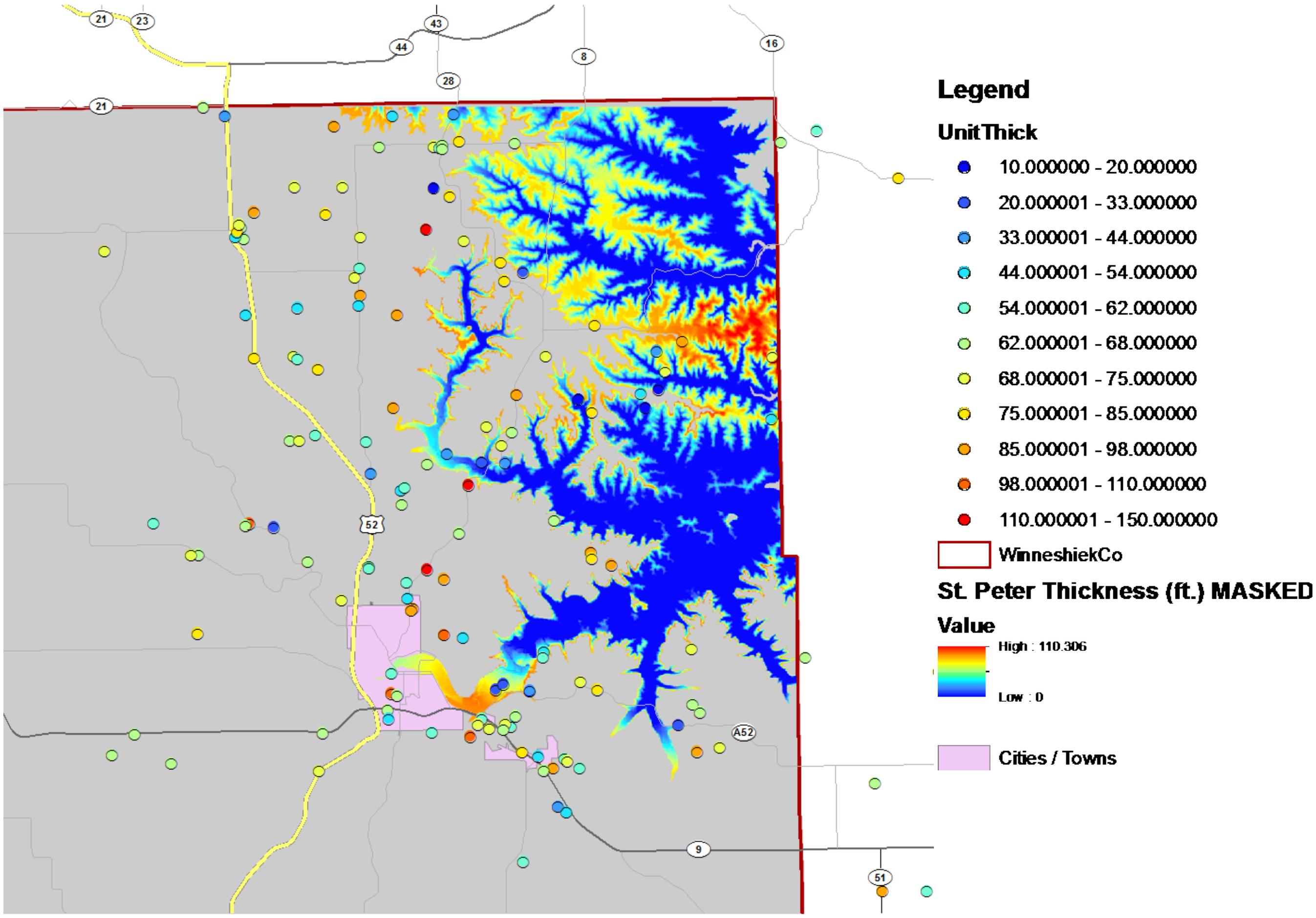

Figure 12

Figure 13 below also shows the derived St. Peter thickness but this time with the observed error at each sample location color coded. The error is denoted by "ThicknessCheck" and is presented as predicted value minus the true (observed) value.

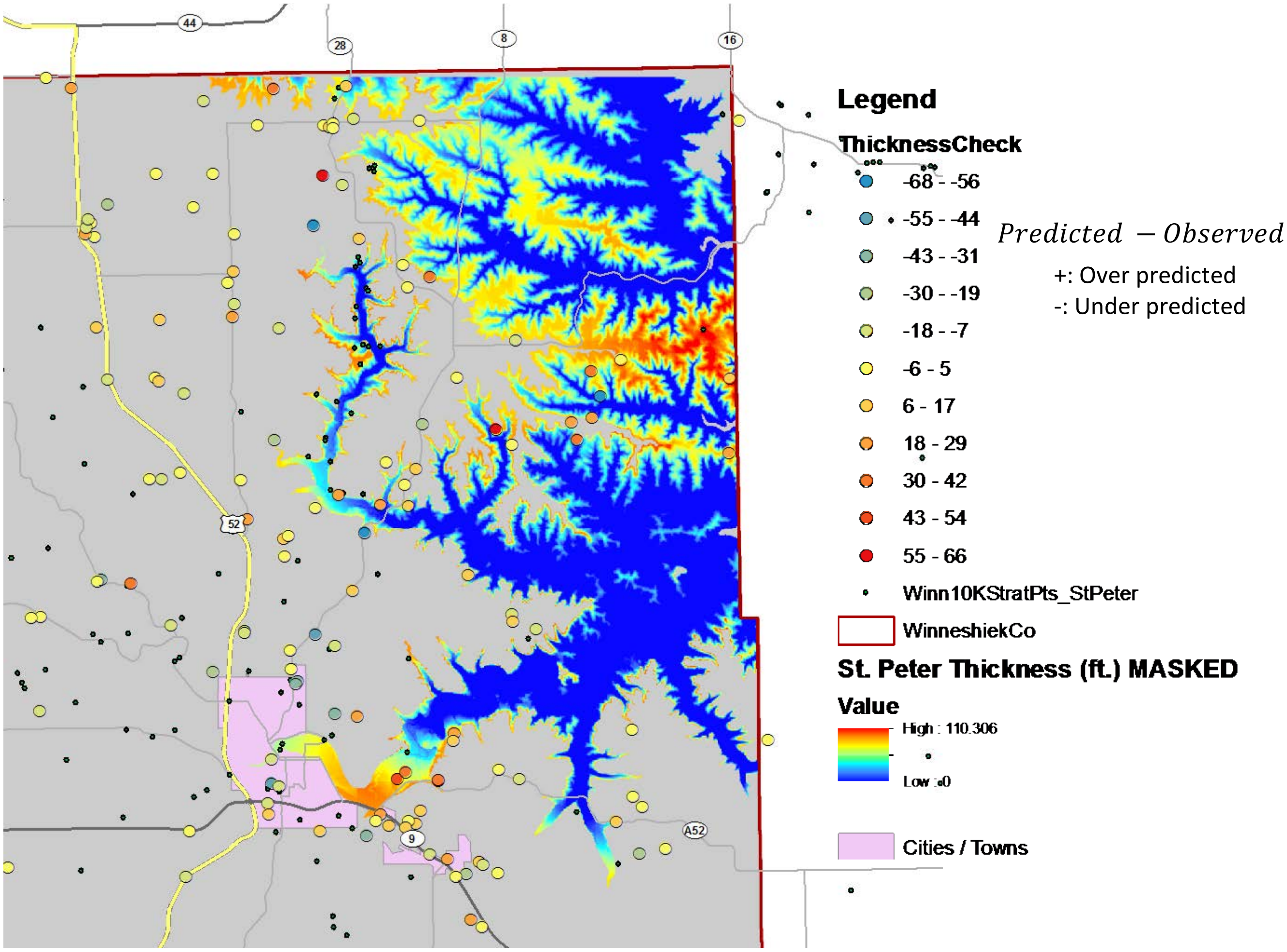

Figure 13

This visualization of the sample points enables both perception of the constraints for the interpolation and an assessment of the local accuracy of interpolation within a localized region. Finally, a histogram on the interpolation error at all sampled location is provided below in figure 14. The shape of this distribution suggests that the residual error is normally distributed. This suggests that without additional expert input, the error is due to random noise and thus the data provides a reasonable estimate given the limitations of the input data.

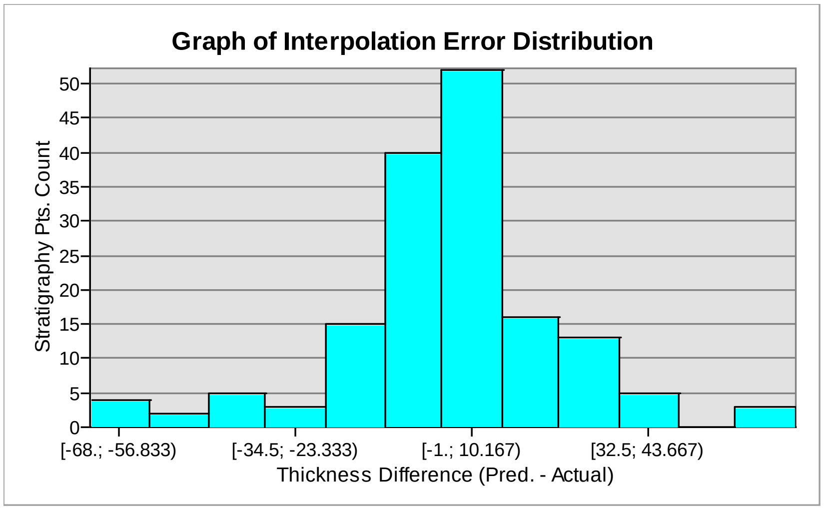

Figure 14

4 INTEGRATING ENHANCED GEOLOGY INTO MINE LOCATION PREDICTION

4.1 SELECTION OF A MASK

Since mining companies are unlikely to be interested in mining locations in which the St. Peter sandstone is over 50 feet below the surface, including locations where this criterion is not met does not make sense. Also inclusion of such area can lead to modifiable area unit problems that could lead to degradation of the analysis. Due to the fact that the St. Peter sandstone polygon provided by the Iowa DNR and the derived area where St. Peter lies within 50 feet of the surface are nearly identical, the St. Peter sandstone polygon for Winneshiek County is used as a mask for all analysis.

4.2 SELECTION OF MODEL CRITERION

The previous linear weighted approach included many criterion in the model. The selection of these criterion was done in a rather ad hoc manner. For this final analysis, included criterion are based on what was learned from the statistical model trained on Wisconsin. Given all similar datasets, the Akaike Information Criterion suggests that only three criteria have a significant impact on site selection. These criteria are depth to bedrock, distance to rail, and distance to major roads. In addition to these criteria, I postulate that the depth of the St. Peter formation in a particular location will be significant. Thus four criteria are used in the final model. The weights given in table 1 below are informed by relative magnitude of the coefficients in the previous statistical model.

Criteria	Weight
Minimize Distance to Major Roads	0.15
Minimize Distance to Rail	0.2
Minimize Depth to St. Peter Sandstone	0.3
Maximize Depth of St. Peter Sandstone	0.35

Table 1

4.3 Criterion Layers

In order to better understand resulting predictions each new criterion layer is shown in figures 15 and 16 below. The depth to the St. Peter and the thickness of St. Peter are the same as those presented in the previous section.

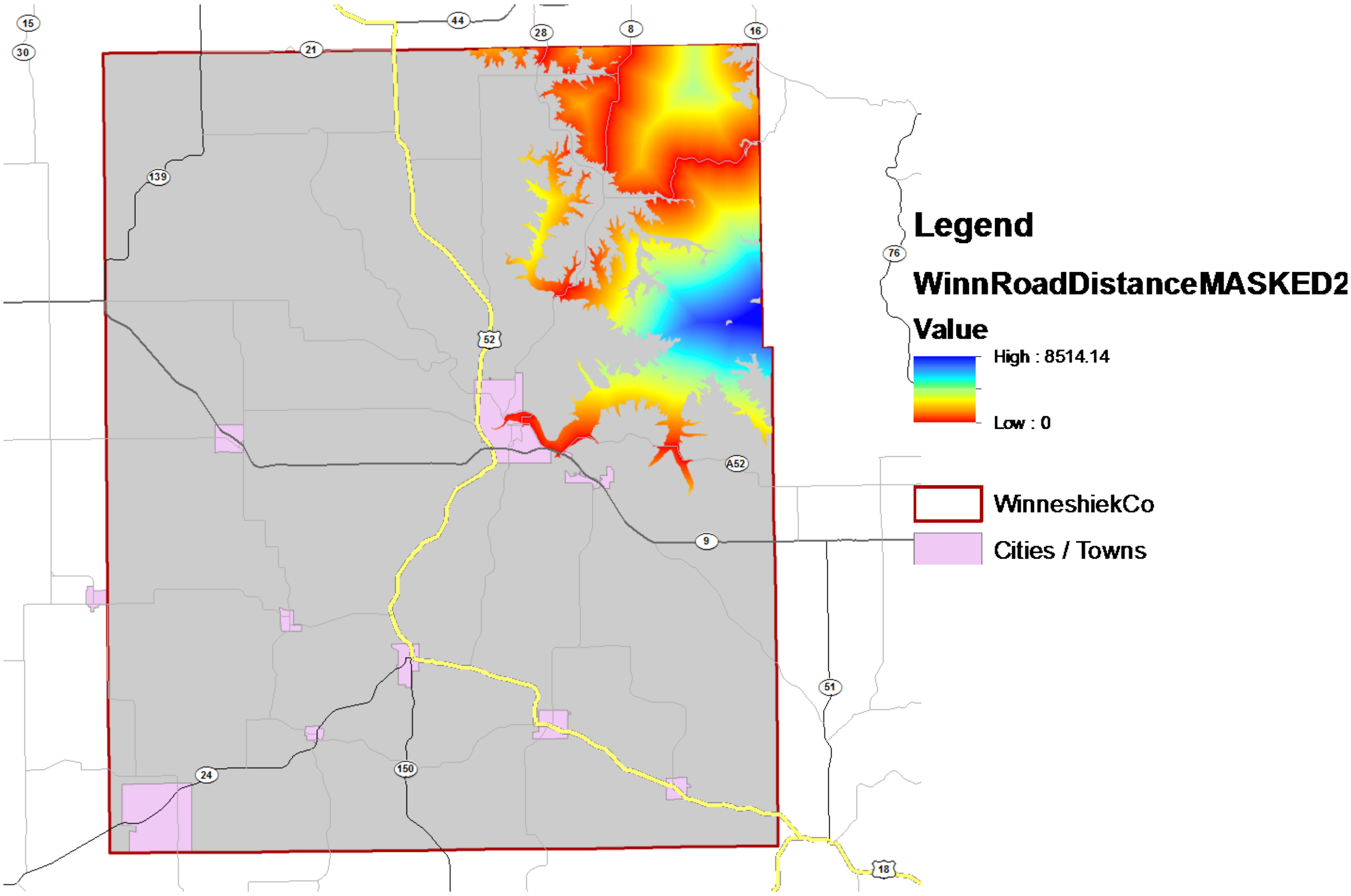

Figure 15

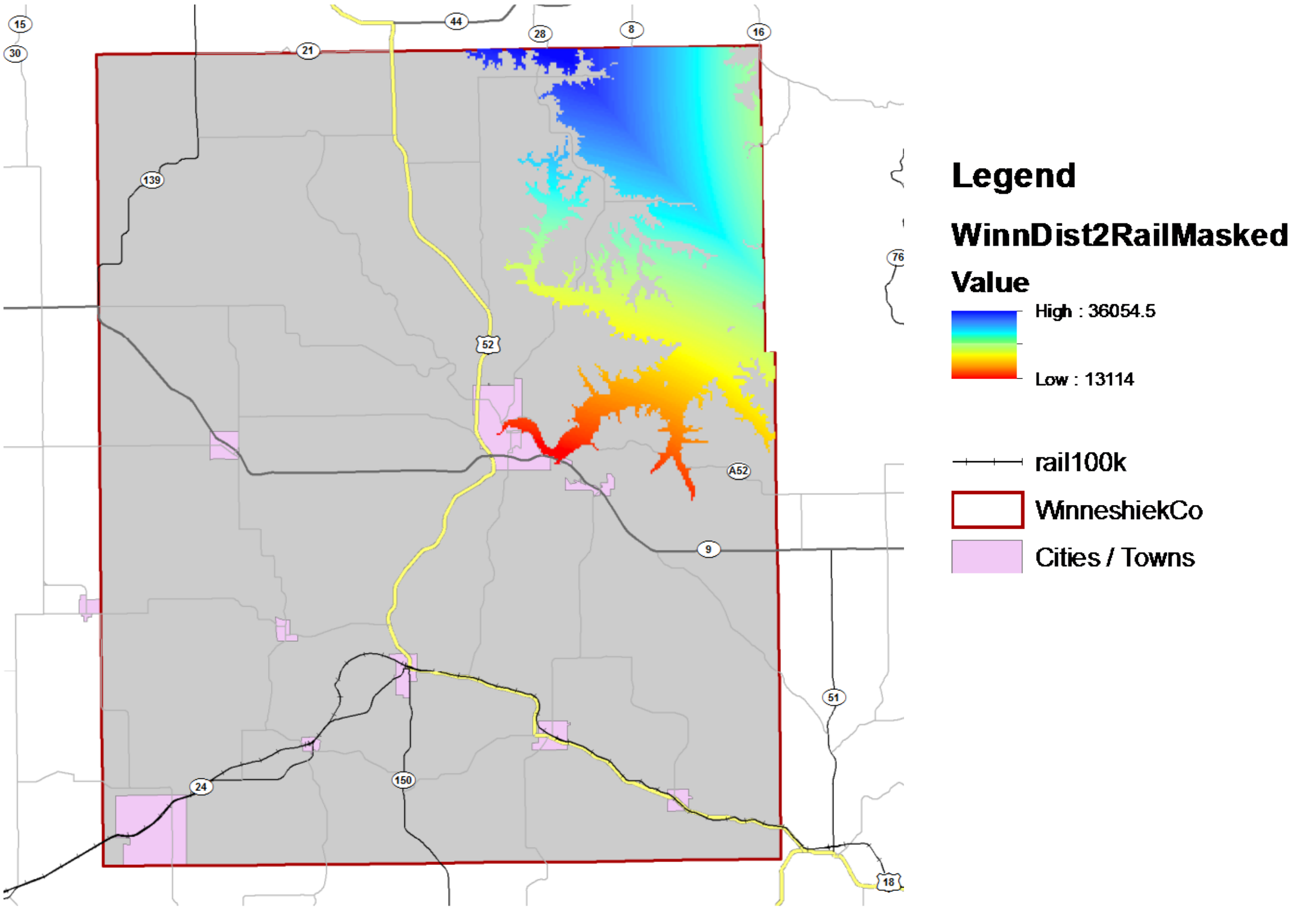

Figure 16

4.4 Final Predictions

In order it aggregate all criteria, the inputs must first be normalized. For this analysis a linear normalization scale is applied within the selected mask. The resulting prediction map using the criteria weights reported in the table 1 is shown in figure 17 below.

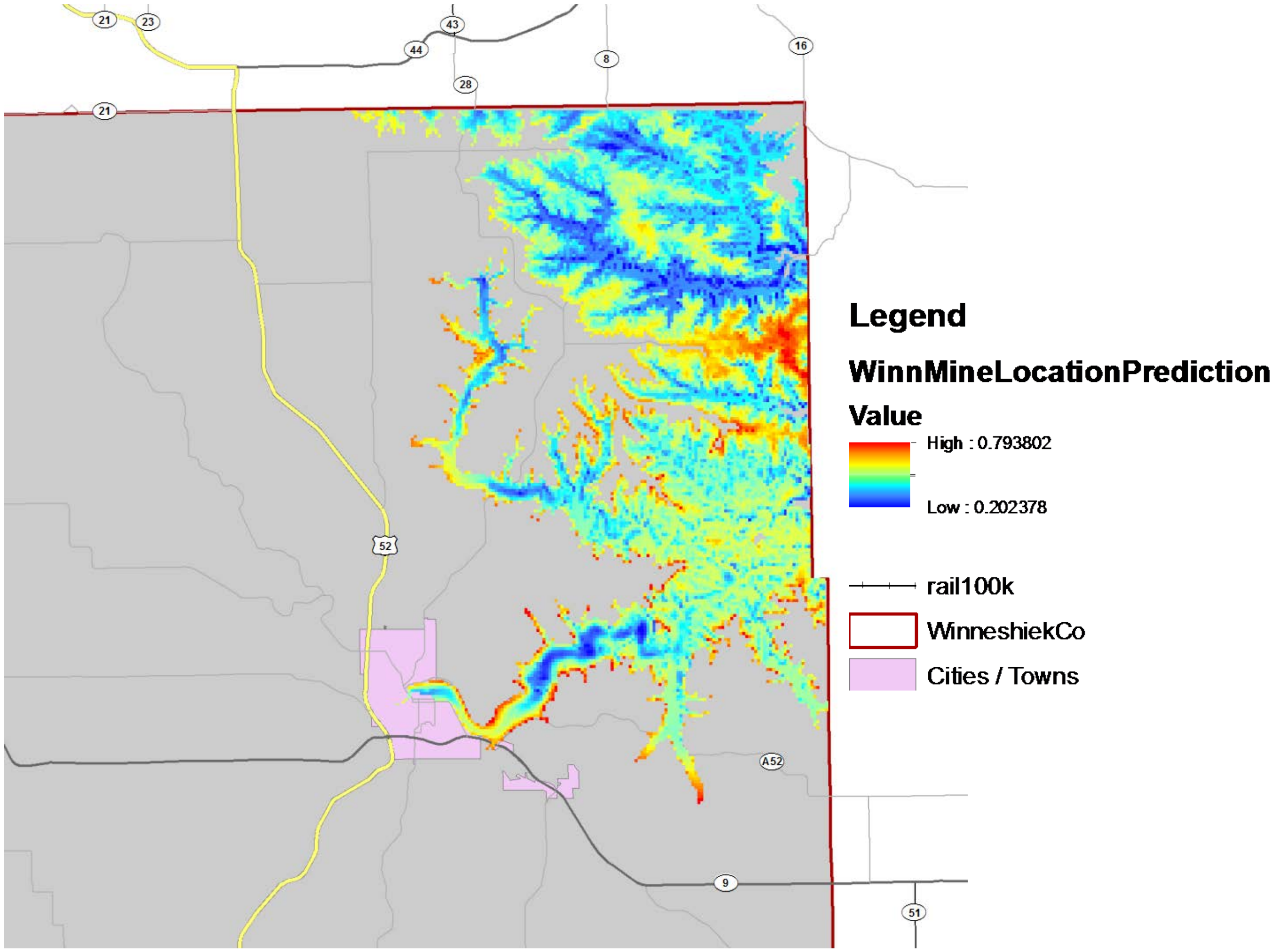

Figure 17

5 POSSIBILITIES FOR FUTURE IMPROVEMENT

One easy and likely beneficial extension of this work would be to derive similar data products for the Jordan sandstone geology in the County as it is also of interest for frac sand mining. The exclusion of the Jordan Sandstone in this analysis was primarily due to time constraints and a similar analytical approach could be followed for Jordan Sandstone to provide a more comprehensive analysis.

Another factor that is likely quite important to mining companies that is still not factored into the new derived model is a metric of sand purity and grain size. Both the grain size and the purity of the sand resource have significant economic viability implications for mining companies. The grain size has a significant impact on the expected market price of the sand. Also, the purity has implications for the degree to which the sand must be washed prior to being sold. While these are important factors to consider when placing mining operations, variability in grain size and purity is extremely variable and thus statistical methods similar to those applied in this analysis would be unlikely to provide any estimates within any useful bounds of uncertainty.

6 References

Calcagno, P., Chiles, J., Courrioux, G., & Guillen, A. (2008). Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules. *Physics of the Earth and Planetary Interiors, 171*(1), 147-157.

Charabi, Y., & Gastil, A. (2011). PV site suitability analysis using GIS-based spatial fuzzy multi-criteria evalutation. *Renewable Energy*, 2554-2561.

Cowles, K., Bales, Z., Cervantes, J., Smith, A., Landgren, J., Muñoz, C., & Young, S. (2014, 12 8). Modeling and Predicting Mine Sites for Hydraulic Fracturing Silica Sand in Northeastern Iowa. Iowa City, IA.

Horsman, J., & Bethel, W. (1995). *Methods of constructing a 3D geological model from scatter data.* Berkeley, CA: Lawrence Berkeley Lab.

Kaufmann, O., & Martin, T. (2008). 3D geological modelling from boreholes, cross-sections and geological maps, application over former natural gas storage in coal mines. *Computers & Geosciences, 24*(3), 278-290.

Patel, M. D., & McMechan, G. A. (2003). Building 2-D stratigraphic and structure models from well log data and control horizons. *Computers & Geosciences*, 557-567.

Shaw, B. (1978). Parametric interpolation of digitized log segements. *Computers & Geosciences*, 277-283.

Vasiljevic, T. Z., Srdjevic, Z., Bajcetic, R., & Miloradov, M. V. (2011). GIS and the Analytic Hierarchy Process for Regional Landfill Site Selection in Transitional Countries: A Case Study From Serbia. *Environmental Management*, 445-458.

Wu, Q., Xu, H., & Xukai, Z. (2005). An effective method for 3D geological modeling with multi-source data integration. *Computers & Geosciencse, 31*(1), 35-43.