

A Spatial Model for Protection of the Edwards Aquifer

C. Michael Parma

Abstract:

The City of San Antonio, Texas, commissioned a Scientific Evaluation Team to develop a spatial model analyzing sensitive lands for protection of the Edwards Aquifer, the primary water supply for the region. The three-county study area covered almost 1,500 square miles over the aquifer. Built upon ArcGIS ModelBuilder, Spatial Analyst, and Python scripting functionality, three submodel components were developed. Permeability was composed of slope, geology, soils, caves, and faults; Biology used vegetation and endangered karst invertebrates; and Adjacency/Watershed utilized stream orders, parcel size, and adjacency. A weighted overlay was applied to the results, producing a score raster for the entire study area. Using zonal statistics, parcel features were assigned a final score and ranked by standard deviations into eight tiers. The model took approximately four hours to analyze nineteen input datasets and produced a final parcel feature class of more than 80,000 records.

Body: **Introduction**

In May of 2000, voters of the City of San Antonio, Texas, approved Proposition 3, a 1/8 cent tax for the creation of linear parks and the purchase of 10,000 acres of land over environmentally sensitive zones of the Edwards Aquifer within Bexar County. This aquifer is the primary residential water source for 1.7 million people. The tax provided \$45 million in funds to be used for this purpose. A Scientific Evaluation Team (SET) consisting of members from various local and Federal agencies and ESRI convened to produce a model that would identify the most sensitive properties. ESRI's ModelBuilder extension for ArcView 3.x was used for this purpose. Because of limited functionality, most processes were run manually and, unfortunately, with little documentation. The final SET model in ModelBuilder consisted of a single Weighted Overlay process, the results of which were eventually assigned into prioritized tiers using a natural breaks assignment. Stone and Schindel (2002) provided an overview of the project and the development process.

In 2005, San Antonio voters approved funding for a similar program with the passage of Proposition 1. Over the next three years, the program will produce over \$100 million in funding. The scope of this project extended into two surrounding counties, Uvalde and Medina, while reevaluating the most recent condition in Bexar County. A second SET team, including some of the original members, made the decision to attempt to first replicate and then enhance the previous model using the new ModelBuilder and scripting functionalities available in the ArcGIS 9.1 environment. With the benefit of the new ModelBuilder environments and almost five years of increased computer processing power, there was great promise in producing a much stronger spatial model. The improved SET model consists of nine submodels and over 200 individual procedures generally following the same organization and flow of the original while automating the entire processes from beginning to end including data source standardization and a final standard-deviations-based tiered outcome assignments. Figure 1 provides an overview of the entire process.

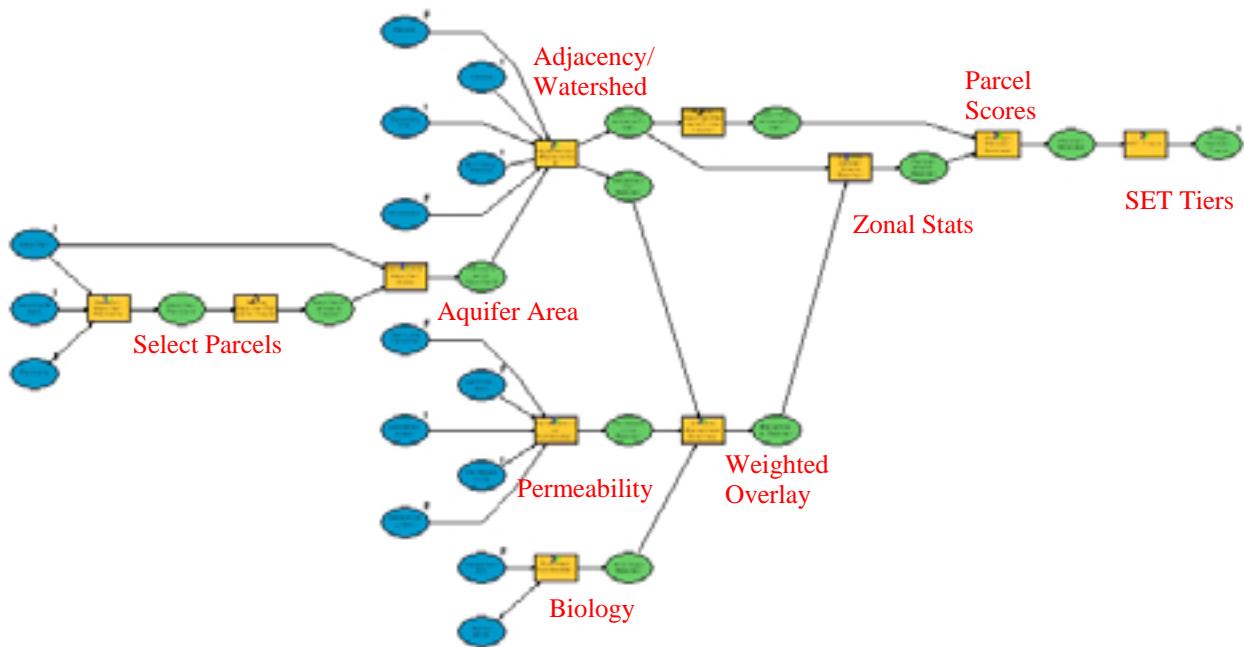


Figure 1. Overview of the SET Model

The nine submodels perform the following tasks:

1. Select all parcels in three counties which are over the Edwards Aquifer recharge and contributing zones
2. Calculate the total area over the aquifer for each parcel
3. Calculate points for biology based on vegetation and endangered species
4. Calculate points for permeability based on caves, faults, soils, slopes, and geology
5. Calculate points for adjacency/watershed based on stream order, adjacency, and parcel size
6. Calculate a weighted overlay of permeability (50%), adjacency/watershed (30%), and biology (20%)
7. Create a zonal statistics raster, calculating the average score for a given parcel
8. Assign the final SET scores to the parcel geometries
9. Calculate SET tiering (1-8) based on standard deviations and size subsets.

Several preparatory submodels are used to standardize and combine nineteen feature classes from various agencies for use as inputs in the SET Model. Although these submodels are not covered in detail in this paper, a general description is provided. Parcels from three counties were integrated in to one parcel dataset. Attribute fields were eliminated since geometries were of the primary concern and new several attribute fields were added to track the original data source in the combined dataset. Since many of the geometries in Uvalde and Medina counties had to be digitized from paper sources, large aggregate areas of small parcels were provided to minimize the deliverable time. These were attributed as “aggregate”. Areas where no where parcel data was provided were added in to the combined dataset and attributed as “missing”. Two sets of vegetation data were combined to provide a single coverage of the entire study area.

Model Processes

Selection

In the first submodel, parcels from the three counties were selected and assigned a project specific identifier. Using a spatial selection, all parcels which intersected the Edwards Aquifer recharge or contributing zones were exported to a new feature class and assigned a unique SETID.

The SET decided that as one improvement over the original model, a submodel should be created where parcels were considered not based on their size alone, but on the acreage of that parcel over the aquifer. In this way, large parcels on the edge of the recharge zone but with little area in that zone would not be given as much weight as the same sized parcel wholly over the zone. To accomplish this calculation of what was called a parcel's *aquifer area*, a union was used to combine both the parcel and aquifer geometries. The areas of each polygon with aquifer attributes were then summarized based on the SETID. The results from this summary table were then joined back to the parcel feature class and the total aquifer area for each parcel was stored.

Major submodels

The third submodel was used for considering adjacency and watershed criteria. Parcels were scored based on their aquifer area, their adjacency to open space properties, the presence higher-ordered streams, and finally, a factor to approximate the underground movement of water. The first steps of the Adjacency/Watershed (A/W) submodel gave all parcels a minimum twenty-five A/W score before unselecting all features which were flagged as coming from the "aggregate" or "missing" parcel datasets. These remaining parcels over sixty acres or over 1111 acres were then given thirty or eighty bonus A/W points. These figures were based upon a minimum sixty acres for viable biological habitat and a standard drainage area for streamhead delineation (Olivera et al., 2002).

Military bases, conservation easements, parks, and lakes were combined into a single "open spaces" dataset. Those parcels immediately adjacent (within a buffer distance of zero feet) were given thirty-five additional A/W points while those within 200 feet were given twenty-five additional A/W points. An improvement to the previous model was an attempt to allow very large parcels to effectively act as their own "open space" in which they were given bonus points for their size. It was determined that a one-acre parcel adjacent to a small park, amounting to two total acres, would receive a larger number of points than a non-adjacent 1000 acre parcel. The additional points for the acreage were an attempt to rectify this occurrence.

In a second improvement, parcels were scored on the presence of ordered streams utilizing the National Hydrography Dataset Plus (USGS and Horizon Systems, 2006) stream data. This expanded dataset provides more data on the USGS NHD data, including stream order. The SET deemed it too difficult to calculate the drainage area for an individual parcel, and so used the stream order as a surrogate for this. Parcels received additional points for the highest-ordered stream which intersects it, from ten points for a first order to fifty for a fifth order, the maximum in the study area.

A subsurface structure known as the Knippa Gap limits the eastward movement of water from western Uvalde county into the remainder of the Edwards Aquifer which supplies San Antonio (Green et al., 2006). To approximate this effect, a final multiplier was applied to the total A/W points assessed in the previous steps using what were called "Leona Factor" zones.

The area west of the estimated Knippa Gap delineation received a multiplier of one, those to the east received a two multiplier, and those further east of approximately the Uvalde/Medina county line received a multiplier of three. Figure 2 illustrates the ordered streams and the Leona Factor zones.

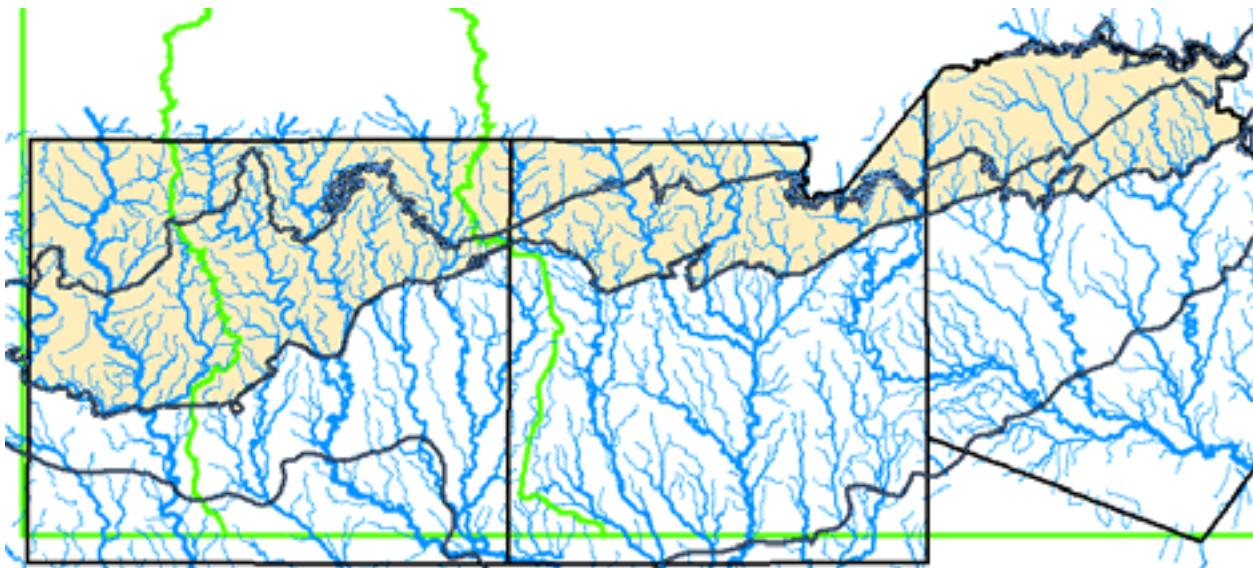


Figure 2. Three County Study Area (tan) with Streams (blue) and "Leona Factor" zones (green)

No major changes were made to the Biology or Permeability submodels, the latter provided by the USGS, other than the weighting of the vegetation and karst invertebrate layers in Biology. Due to the fact there was no documented karst invertebrates outside of Bexar County, the SET decided to reduce the karst weighting to 30% from the original 50% to eliminate an inordinate weighting on Bexar parcels. Vegetation was increased to 70%, continuing to use the same reclassifications as outlined in Stone and Schindel (2002).

The final outputs for the Permeability, Adjacency/Watershed, and Biology submodels was a raster with each cell containing a score value for the respective submodel. The scores were normalized individually against the highest possible value and returned as a percentage of that value. In that way, all three rasters had possible values between zero and 100. The normalized values served as inputs for a weighted overlay. It was this one step that constituted the “model” in Stone and Schindel (2002). The same weights for each submodel were used in this improved version: permeability – 50%; adjacency/watershed – 30%; biology – 20%.

Final scoring and tier assignments

The next three submodels represent many of the large improvements over the 2001 model, namely the assignment process for tiering the parcels. In the original, each parcel was given the value of the highest weighted overlay grid cell which lay within it. These parcel scores were then assigned a tier value based on natural breaks. A Tier 1 represented the most sensitive parcels for recharge protection, with an increasing tier ranking indicating decreasing sensitivity. It is not known how many groups were originally created, only that the top three were reported.

Originally, the situation could arise in which a large parcel with dominantly low scores and a single high score would therefore achieve an overall high score. There were no documented cases of this occurring, but it was to be avoided in the improved model. The SET decided to use an area-weighted method to determine a parcel's final score. To accomplish this, the Zonal Statistics (ZS) tool was used. This tool "calculates statistics on values of a raster within the zones of another dataset" (ESRI, 2006). The process was to calculate a mean value for all the cells within each zone (parcel). With over 80,000 parcels to evaluate, an undocumented limitation of the ZS tool was discovered. This tool requires that the zonal identifier be a small integer, less than 32,767. To overcome this limitation, a Python script was written to iteratively split the parcels into three subsets, assign a temporary process ID, convert the polygons to raster, and finally perform the zonal statistics, outputting mean parcel score value rasters for each subset. These three rasters were then mosaiced back into a single complete final parcel scores raster.

To assign the SET scores to the parcel geometries, the parcels were converted to points which were input into the Sample tool. This provided a table of the scores raster value and a coordinate value for each point. Since the parcel's SETID was not exported from the Sample tool, the Sample table was made into an XY layer and an Identity tool produced a point layer with the SETID and the mean score. This parcel and Identity layers were then joined based on SETID and the parcel scores calculated into the parcel table.

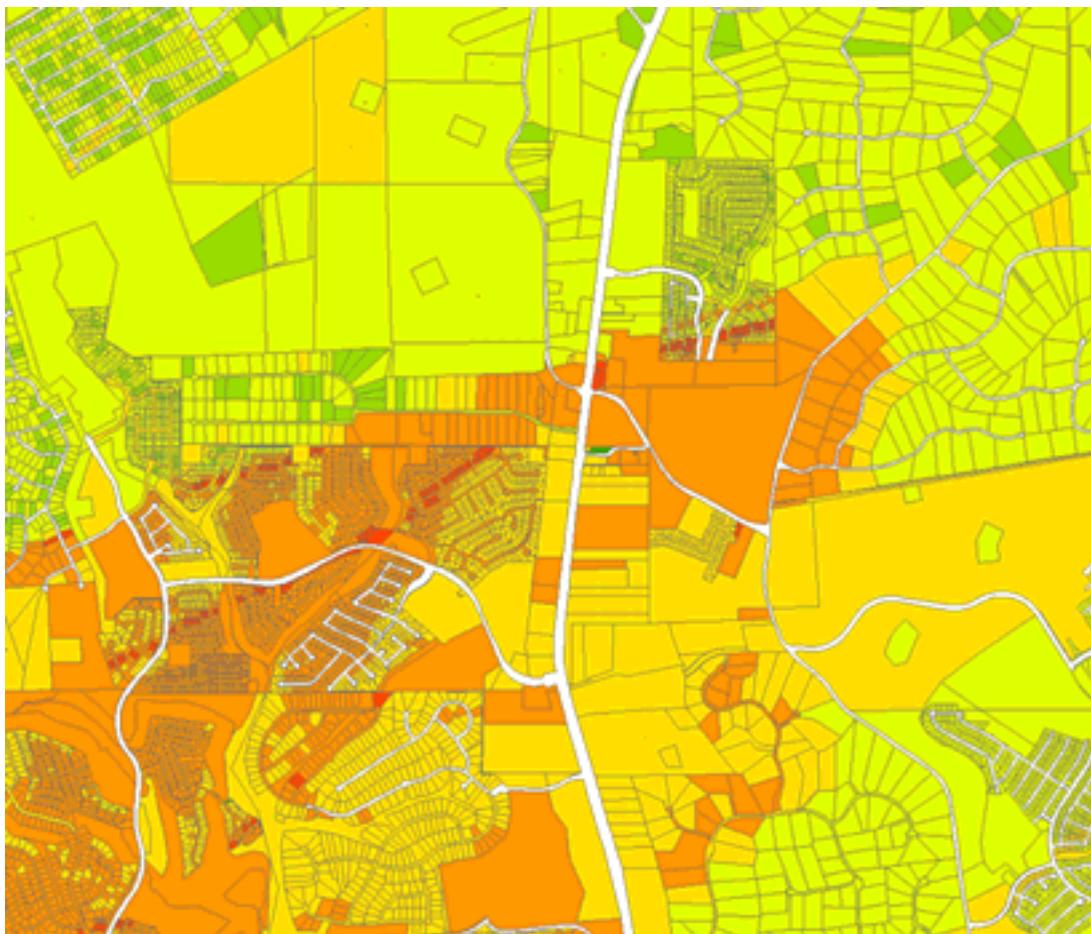


Figure 3. Example of Final Parcel SET Tiers

With the parcel scores stored in the parcel geometries, the tier assignments could finally be made. Again, the SET decided to automate the process. Whereas the natural breaks assignment process was deemed too subjective, tiering based on standard deviations of the parcel scores was implemented. The parcel scores were analyzed and the statistics written to an intermediate table. Using this table, all parcels with scores at least three standard deviations above the mean were assigned to tier one. All parcels with scores at least two standard deviations above the mean but less than three standard deviations were assigned to tier two and so on for a total of eight tiers. The final outcome of the entire model was a parcel dataset of over 80,000 records which were automatically assigned to one of eight tiers.

Figure 3 provides an example of the tier scores for a small portion of the study area. While the final tiering only considered parcels greater than fifty acres in size, all parcels remain in this classification. Due to the ongoing nature of the project and the negotiation process, information which could be used to identify parcels, has been removed and the results of all parcels (as opposed to the final fifty-acre scores) are shown.

Results

From nineteen input datasets, a single parcel dataset was created with the final SET tiering calculated. An independent committee responsible for the negotiation and purchase of lands decided that a minimum of fifty acres was required to justify the efforts on their part. The last submodel was rerun on all parcels greater than this minimum size threshold, reducing the output to less than 1,500 parcels. While 81,000 parcels in the input dataset were ranked, the minimum size allowed for a more focused consideration of 230 parcels ranked tier-3 and above.

Most of the original processes in the 2001 model were completed by hand, many requiring almost thirty-six hours to complete for one county's data. With the new software and greater hardware power, the entire improved SET model with three counties of data required only approximately four hours of run time on a Dell Precision 670 with 4GB of RAM and a dual 3.4GHz processor.

Leveraging improvements in ModelBuilder for ArcGIS 9.1, this improved version is a fully documented, replicable spatial analysis model. As part of the documentation process, a compiled Microsoft Help document was created. Given the greatly reduced run times and fully automated processes, the model can be rerun quickly as new conditions in the input datasets become available.

Acknowledgments:

The author would like to acknowledge the members of the Scientific Evaluation Team who offered substantial discussion and consideration to the development of this revised model: Joe Chapa and Susan Spegar of the City of San Antonio; George Veni of George Veni and Associates; George Ozuna and Diana Pedraza of the United States Geological Survey; and John Hoyt of the Edwards Aquifer Authority. Thanks go to Reggie Romo of ESRI-San Antonio for technical assistance on a couple of key points.

Appendices:

N/A

End Notes:

N/A

References:

1. ESRI. 2006. ArcGIS ArcInfo. [9.2].
2. Green, R. T., F. P. Bertetti, N. M. Franklin, A. P. Morris, D. A. Ferril, and R. V. Klar. 2006. Evaluation of the Edwards Aquifer in Kinney and Uvalde counties, Texas. [http://edwardsaquifer.org/pdfs/reports/AS%20Reports/Uvalde_final_19_June2006\[2\].pdf](http://edwardsaquifer.org/pdfs/reports/AS%20Reports/Uvalde_final_19_June2006[2].pdf) (last accessed 6 May 2007).
3. Olivera, F., J. Furnans, D. R. Maidment, D. Djokic, and Z. Ye. 2002. Drainage systems. In *Archydro: GIS for water resources*, ESRI Press.
4. Stone, D. and G. M. Schindel. 2002. The application of GIS in support of land acquisition of sensitive groundwater recharge properties in the Edwards Aquifer of South-Central Texas. *Journal of Cave and Karst Studies* 64(1):38-44.
5. USGS and Horizon Systems. 2006. NHDPlus User Guide. http://www.horizon-systems.com/nhdplus/data/NHDPLUS_UserGuide.pdf (last accessed 6 May 2007).

Author information:

C. Michael Parma
GIS Coordinator
City of New Braunfels, Texas
424 S. Castell Ave.
New Braunfels, TX 78130
Ph: 830/221-4000
Fx: 830/608-2112
mparma@nbtexas.org