Introduction

Project Summary

In 2014 multiple local Otsego county agencies, Otsego County Soil and Water Conservation District (SWCD), the Otsego County Planning Department (OPD), and the Otsego County Conservation Association (OCCA), came together to begin updating the current Agricultural and Farmland Protection Plan. The Otsego county plan had not been updated since 1999 and much in the agricultural sector has changed in that time despite agriculture and farms being of large importance to the economy and well-being of Otsego county. In response to this, the planning committee for the updated Agricultural and Farmland Protection Plan (AFPPUC) expressed a need for an automated system to prioritize the County's land for agricultural and farmland uses. Using the model builder application within ESRIs ArcGIS 10.5, a model was built based on the 1997 Land Evaluation and Site Assessment (LESA) system created for the Natural Resource Conservation Service (NRCS). In addition, the use of online GIS focused forums such as the GIS subreddit located within Reddit.com and the GIS Stack Exchange at stackexhange.com, were of great assistance to this project.

Model Concept and Organization

The model was planned in a similar fashion to how it is described in the 1997 LESA guide, the overall model is divided into two sub-models, the Land Evaluation section (LE) and the Site Assessment section (SA). Land Evaluation factors are soil based while SA factors are non-soil based. SA factors can be further divided into SA-1, SA-2 and SA-3 factors. SA-1 factors are any non-soil based factors that deal with limitations to agricultural productivity or farm practices; SA-2 factors deal with development pressure and/or land conversion; and SA-3 factors deal with public values, such as historic or scenic values (Pease and Coughlin, 1997). Eleven factors were chosen for this approach, consisting of LE, SA-1 and SA-2 factors. Each factor is evaluated and scored by a sequence of tools strung together within the model builder GUI. Factors are all scored on a 0 to 100-point scale, with higher scores indicating a parcel of higher priority for conservation. Each factor is then run through the weighted overlay tool, which outputs the final product. All factors use the Jenks Natural Breaks Optimization method for classification. This method seeks to minimize deviation from the mean within classes, while maximizing deviation from other class means (Jenks, 1967). The goal of the method is to determine the best configuration of values into classes for displaying on a choropleth map. Each factor receives a weight in the final weighted overlay from 0 to 100 percent.

Methodology

Data Sources and Databases

The data used in this approach was taken from multiple sources. Soil data used for LE factors comes from the NRCS soil survey database, gSSURGO. The soil survey database includes a 10M raster for the whole state of NY as well as a multitude of tabular data. Data used for SA factors was mainly sourced from the NYS clearing house, found here: https://gis.ny.gov/, and the Cornell Geospatial Information Repository, found here: http://cugir.mannlib.cornell.edu/mapsheet.jsp?code=077. These datasets include shape files for streams, wetlands, floodplains, agricultural districts, roads and municipal boundaries. Shape files for light districts, parcels receiving easements and the latest parcel boundaries (2015) were acquired from the Otsego County Real Property Tax Service and the Otsego County Planning Department. There are two file geodatabases (gdb) associated with this project, titled LESA and ModelOutputs. LESA contains all base data used to build the model. LESA also contains data that was originally part of the model but later discarded; this data may be of use to other Otsego county agencies. No feature datasets were used in LESA, only feature classes. Each was converted to NAD-83 UTM Zone 18N, the coordinate system recommended for map scales greater than 1:24000 within the extent of New York State ("Datum and Coordinate System Standard", 2017).

Use of GIS Forums

GIS-focused online forums were of great assistance when trouble-shooting problems with building the model. Online forums such as the ones used in this approach are communities of individuals with varying levels of experience with the focus of the forum, such as GIS. Forum sites typically have user accounts. Once an account is created, users can submit questions asking for help with a related topic, answer questions from other users, or share success stories and interesting projects. Two forums were frequented during the building of this model: Reddit's GIS subreddit (essentially a forum within a broader forum), found here https://www.reddit.com/r/gis/, and Stack Exchange's GIS forum, found here https://gis.stackexchange.com/. Not every answer to a problem can be found on these forums, but often one can be lead in the right direction, or find an individual who is extremely helpful and will assist in troubleshooting until the problem is solved. Lastly, users can search the databases of these forums to find older posts that may be relevant to the problem at hand.

Sequence for LE factors

Some preprocessing was required before an input for LE factors is ready for the model. First, a new shapefile, containing soil data including important farm classes, NCCPI scores and soil mapping units is created (data found in gSSURGO database). This shapefile was then clipped to the extent of Otsego County. Next a new field needed to be created to house NCCPI scores. This is due to the original field being floating point values ranging from 0.0 to 1.0. The new field was changed to double and contains the same NCCPI scores as the original, increased by a factor of 1000. This allows us to rasterize based on NCCPI scores. Using the project tool the layer is then projected in the coordinate system described above and exported to the LESA gdb, as the feature class titled Ots Soils. This edited soil feature is the input model parameter for both LE factors used in this approach. The fields for important farm classes (frmldcl) and NCCPI scores (NCCPIall) are both then rasterized using the feature to raster tool. Rasters are datasets comprised of rows and columns of cells that house values. Cell sizes can be specified by the user—here, we use a cell size of 10 for every factor. Next, the layers are reclassified, with areas of better farm classes, or higher NCCPI scores, receiving higher scores. Only four types of farm classes exist, therefore only four class breaks could be included for this factor (NCCPI scores have six class breaks). Both layers are then copied using the copy raster tool and modified as to match the format and pixel depth of the factors detailed in the following section. Lastly, both layers are inputs in the final weighted overlay. Farm classes and NCCPI scores contribute to 45 percent of a factor's score, receiving 25 percent and 20 percent respectively.

Sequence for SA factors: Overview and the Main Parcel Layer

SA factors make up the bulk of the model in this approach. Nine SA factors were chosen based on suggestions from the 1997 guide and request from the AFPPUC. SA-1 factors included: distance to wetlands, streams, floodplains, light districts, protected farmlands, parcels that have an agricultural use, and parcel size. SA-2 factors included distance to roads and agricultural districts. SA-1 factors comprise 40 percent of a cell's score. Parcel size receives a weight of 10 percent, every other SA-1 factor receives a weight of 5 percent. SA-2 factors make up the remaining 15 percent of the weight, with distance to agricultural districts receiving 10 percent and distance to roads 5 percent. Every SA factor interacts with the parcel boundary dataset for

Otsego County. This main parcel layer required some pre-processing before it was ready to be a model input. The dataset contained polygons for roads which did not include any tabular data; these were removed. In addition, any parcels within city or village boundaries were also removed as none of these parcels are desired for conservation; after removal of these polygons, 31,761 parcel features remained in the dataset. Lastly, to evaluate the distance to parcels with agricultural uses factor, a new field needed to be created within the dataset titled Distance Field, which is explained in the later sections of this report.

Sequence for SA Factors: Distance to wetlands, streams, floodplains, light districts, agricultural districts and protected farmland.

The SA factors distance to wetlands, streams, floodplains, light districts (representative of hamlets), agricultural districts and protected farmlands all follow the same sequence for generating the final output raster. First the spatial join tool, found in the spatial analysis toolbox, is used to join the LE factor being evaluated (e.g. distance light district) with the main parcel layer. Within the tool is the option to change how features between the two layers will be joined, called the match option. The setting titled CLOSEST is picked from the match option dropdown menu, which outputs a field that contains the closest target feature (main parcel layer) to the join feature (LE factor), within a specified search radius. Please note, unlike other tools within ArcMap, spatial join does not give features that fall outside of a search radius null values or no data values. Instead, the tool gives these features a value of -1. The new field is then rasterized using the feature to raster tool. The resulting raster contains no attribute table; this is problematic as later the layer will need to be reclassified. To build an attribute table two tools are needed, copy raster and build raster attribute table. Copying the raster using this tool allows us to change the pixel depth and format of this layer. 32 bit signed is chosen for pixel depth as it contains the correct range of values needed to display the data contained in the layer; the format specified is

ESRI GRID. After a modified copy has been created, the layer is run through the build raster attribute table tool, which creates a new attribute table for the layer. Next, each layer is reclassified on the same 0 to 100 scale as the LE Factors. The factors dealing with light districts, wetlands, streams and floodplains all have only two class breaks and scores, 0 and 100. Since it is preferable for agricultural operations and farmland to be outside of the specified radius, a value of -1 will receive a score 100, any value within the search radius receives a value of 0. The current search radius thresholds are 100 feet (30.48 meters) for all environmental factors (e.g. streams, floodplains, wetlands), 100 feet for distance to light districts, 500 feet (152.4 meters) for distance to agricultural districts, and a quarter mile (402.34 meters) for distance to protected farmland areas. Distance to agricultural districts and protected farmland factor scores both exist on the same 0 to 100 scale with six class breaks each. The closer a parcel within the study area is to one of these factors, the higher a score it will receive.

Sequence for SA Factors: Parcel Size

Parcel size, distance to roads and distance to parcels with agricultural uses all follow a similar but slightly different sequence than the above factors. The only difference in the sequence to generate the parcel size raster and the previously explained factors is that the field rasterized is already contained within the parcel layer. The main parcel layer contains a field titled ACRES which contains the acreage of each parcel. This factor is reclassified and scored with six class breaks, the larger the parcel is the more efficient it is to farm, and therefore larger parcels will receive higher scores than smaller ones.

Sequence for SA Factors: Distance to Parcels with an Agricultural Use.

The distance to parcels with agricultural use factor was the most difficult factor to create in this approach, due to parcels being evaluated in the main parcel layer also existing in large amounts within the factor layer. If evaluated in the same manner as the other factors, every agricultural parcel would receive a score of 100 as both the target and join features will contain that parcel. To overcome this, the generate near table tool was used. This tool can be used to generate a table that contains the closest distance to a larger number of join features, specified by the user. First, a new dataset was created titled Ots AgParcels. This dataset was created using a standard query statement to select by attribute every parcel with an agricultural use in the county, and then creating a layer from the selection, projecting the layer just as with the Ots Soil layer, and exporting to the LESA gdb. The main parcel layer and the newly created Ots AgParcels are both inputs within the generate near table tool. A search radius of a quarter mile is selected and the tool is run. Distances are recorded for each parcel evaluated from the main parcel layer. The first distance is the closest distance to any parcel that has an agricultural use, which is specified in the USERDESC field in the main parcel field. The second distance is the next closest parcel. For any parcel with an agricultural use, the second distance is the closest distance to any other parcel with an agricultural use. For a parcel with any non-agricultural use, the first closest distance will be used. Next, the table generated is joined to the main parcel layer. To achieve this, a python script was used that assigns a value from the generated table to the new field (Distance Field) within the main parcel layer; this script was sourced from the Reddit GIS forum and slightly modified to fit the pathways and field names used in the model. The script correctly applies either the first or second distance value calculated from the generate near table tool, based on the type of parcel use. The feature to raster tool is then used to rasterize this newly created field, and the raster is reclassified. This factor has six class breaks, with a parcel closer to parcels with an agricultural use receiving a higher score. One more problem needs to be addressed to create this factor: the field the raster is created from contains null values for parcels that fell outside of the search radius. Unlike the spatial join tool, generate near table does not give cells that fall outside of the search radius a value of -1; instead, they have no data. Being that for this factor, a parcel outside of the search radius will receive a score of 0, every no data value was converted to 0 using a raster calculator expression. After the expression is applied, this factor has a total of six class breaks. The closer a parcel is to another parcel that has an agricultural use, the higher its score. Finally, the layer is set an input into the final weighted overlay.

Sequence for SA Factors: Distance to Roads

The distance to roads factor follows a sequence very much the same as the factors distance to wetlands, floodplains, streams and light districts. The main difference in this sequence is that when this factor's raster is copied and modified, any values of 0 were changed to NoData. This result is perplexing as pixel depths are associated with different ranges of values that a cell can store within a raster, a value of 0 is within both the ranges of the original raster (64-bit) and the modified raster which is transformed to the standard 32-bit signed that every factor has. To rectify this issue, the same raster calculator expression used above for the distance to parcels with agricultural uses factor is used. This factor then follows the same sequence as the factors listed above. An attribute table is built and the result is reclassified with six class breaks. Again, -1 values are seen within this factors reclassify tool table; these values indicate that these cells were not within the specified search radius, 500 feet (152.4 meters). For this factor, it is preferable for a parcel to be outside of the search radius. Therefore, any cell with a value of -1 receives a score of 100 for this factor. Total this factor has six class breaks, the closer a parcel is to a state, county or federal road, the lower its score. Once reclassified, this layer is set as an input into the final weighted overlay.

Using the Model: Weighted Overlay and the Final Output

As stated above, once a factor finishes its sequence, it is then run through the weighted overlay tool found in the spatial analyst toolbox. Each factor has a different weight and therefore one factor contributes more to the overall LESA score for a cell than another. The model itself is housed within a toolbox titled LESA.tbx, which is in a directory (folder) with both databases. Once the folder is connected in ArcMap, right click the model and navigate to open, ensure all datasets and options are current and click ok. The model tool will take approximately 20 minutes to run a computer with a 2.5 Ghz CPU, 6GB DDR3 RAM, and a 5400 RPM HDD, it is not recommended to store or run the model from a flash drive unless necessary. The final output is set to a cell size of 10 and can be found in the ModelOutputs gdb; methods for changing datasets, factor scores and weights are found in the below section.

Using the Model: Updating Feature Classes, Cell Size, Search Radius, Factor Scores and Weights

Datasets, factor scoring and weights can all be modified. To change the feature classes, open the model. Options to input a dataset or shape file for each factor are present, click on the small folder graphic to the right of the model parameter in question and navigate to the dataset to be changed. Once the new datasets are selected click ok to run the model. For this model to run seamlessly the dataset must contain the correct datum, projection, spatial extent, pixel depth, and format as the rest of the datasets used in this analysis; refer to the above sections to ensure everything matches up. In addition to changing datasets (i.e. model parameters), search radius specified for a factor and the cell size of the model can also be changed; these options are found below the dataset inputs.

To change factor scores, navigate to the model location, right click, and then select edit. This brings up the model builder GUI for this model. In model builder, data inputs appear as blue circles, tools as yellow rectangles and outputs as green circles. Most factor tools are renamed to be descriptive about which factor they are associated with and which tool they are. Navigate to the tool titled reclassify. Right click the tool and open it; a small table comes up showing the scores associated with each class of values. To change the method used to create class breaks or the number of breaks, hit the classify button at the bottom right of the table. Factors that have -1 values have this value in an isolated class. If the amount of class breaks or classification method changes, these -1 values will need to be isolated again. Delete the range of values with the -1 value and add them to class within the next class break. Changing weights is like the method used for factor scores. Find the tool titled weighted overlay, it's the last tool in the flow chart. Right click to open the tool as before and there will be a table with every factor. All weights need to add up to 100 percent, to the left of the factor scores is their weight in the model, delete the entry and type the desired weight to change them. Once all changes are made, apply them, save and close the model.

References

Pease, J., & Coughlin, R. (1997). Land Evaluation and Site Assessment: A Guidebook for Rating Agricultural Lands, Second Edition. Soil And Water Conservation Society. Retrieved from https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1047455.pdf

Jenks, George F. (1967). The Data Model Concept in Statistical Mapping. International Yearbook of Cartography 7.

Datum and Coordinate System Standard. (2017). Gis.ny.gov. Retrieved 10 June 2017, from https://gis.ny.gov/coordinationprogram/workgroups/wg_1/related/standards/datum.htm