SPATIAL SIMPLIFICATION OF INPUT DATA FOR HYDROLOGIC MODELS: ITS EFFECT ON MAP ACCURACY AND MODEL RESULTS

Casson Stallings
Casson_Stallings@ncsu.edu
Computer Graphics Center, Department of Forestry

Siamak Khorram
Khorram@ncsu.edu
Computer Graphics Center, Departments of Forestry and of
Electrical and Computer Engineering

Rodney L. Huffman
Huffman@eos.ncsu.edu
Department of Biological and Agricultural Engineering
North Carolina State University
Raleigh, NC 27695

ABSTRACT

Statistical generalization can be used on map data to reduce the complexity of the data before analysis. The generalization scheme tested here calculates the dominant cover class (mode) of a theme within specified areas. In this work, the specified areas are agricultural fields, and the themes are soils, slope, and aspect. The purpose of this study was to assess the impact of using a vector-based mode amalgamation scheme on the map polygon count, map accuracy, and hydrologic model results at field and watershed scales.

The Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) model was initially run on fields using inputs based on the original coverages describing soils, slope, aspect, and field boundaries. Polygons representing one or more themes were amalgamated and the simulations were then repeated. The map accuracy of the amalgamated themes were calculated. The original model results were compared to those based on the amalgamated data.

The amalgamation technique decreased the number of polygons, but introduced substantial map errors. The soils amalgamation introduced error into most measures, but more so at the field scale than the watershed scale Amalgamations of the slope and aspect themes had less effect on the model outputs investigated. We concluded that the aggregation method and extent should depend on the model being used and on the modeling objectives.

INTRODUCTION

In using vector-based GIS it is often necessary to intersect multiple themes. Each new intersection typically increases the number of polygons in the combined coverage by many more than the added theme contained. It is likely that most vector-based GIS-model interfaces must deal with this problem as they combine information on soils, slope, watersheds, land use, and other themes. Raster-based interfaces, if they are to handle homogenous regions and use a small cell size, encounter the identical problem

Stallings *et al.* (1992) described the GIS Interface for Ground-Water Models. That GIS-model interface used multiple layers of detailed map data over a small watershed. The interface is designed to be used over larger areas, however. Application in this manner leads to large amounts of data and many polygons to process. For this and other vector-based modeling applications, each polygon must be processed individually, so that the processing (CPU) time is approximately proportional to the number of polygons. Decreasing the number of polygons required to represent the geographic data decreases the general computational effort, but at the cost of reducing the accuracy of the analysis.

Map generalization techniques can be classified according to their purpose (cartographic, statistical), domain (spatial, thematic), and data structures processed (raster, vector). Map generalization is described and reviewed by Monmonier (1983), Brassel and Weibel (1988), McMaster and Monmonier (1989), and McMaster and Shea (1992). The generalization technique investigated in this study is a statistical vector-based polygon amalgamation technique that affects both the spatial and thematic data. It is used to reduce the polygon count of the final coverages, decreasing the resulting processing effort

McMaster and Shea (1992) describe two generalization methods appropriate for reducing the polygon count of a vector-based coverage: classification and amalgamation. Classification (as described by McMaster and Shea) consists of redefining the thematic attributes into fewer classes. This can then be followed by an operation to join adjacent polygons that are no longer distinct. 'Amalgamation' is the general term McMaster and Shea use to describe the merging of adjacent polygons. Typically this is done in one of two ways. The first is to eliminate--by merging with adjacent polygons--all polygons below a set size threshold (e.g., ELIMINATE in Arc/Info, ESRI, 1992). This technique was originally used to reduce the number of slivers in intersected coverages The second is to merge adjacent polygons that share some attribute (e.g., DISSOLVE in Arc/Info, ESRI, 1992).

Map Simplification Experiences The effects of the elimination technique on map error have been investigated by Wang and Donaghy (1992). In general the increase in map error was proportional to the increase in the size threshold used. However, at very small size thresholds, many polygons could be eliminated with very little error. A similar trend was found in spatial simplification described in Stallings *et al.* (1992) In that simplification the size threshold was based on polygon sizes within agricultural fields. A one-percent threshold would result in each polygon that consisted of less than one percent of the field area being merged with an adjacent polygon This assured that despite their small size, the relatively large (percentage wise) polygons within small fields would not be decimated in the simplification process.

Most investigations of map generalizations and their effect on GIS-model results have looked at grid-cell size changes. Whede (1982) quantified the effect of cell size on map errors. He found that map errors, and the variability of the map errors, increased as cell size increased. Brown *et al.* (1993) investigated generalization of input data for the agricultural non-point-source pollution model (AGNPS). They varied grid-cell size and looked at the errors in land-cover distributions, predicted erosion, and deposition summaries and areas; they then related the errors to the fractal dimension of the inputs and to the semi-variograms of the inputs. The authors found significant errors in the model results as the grid-cell size increased beyond 120-180 meters on a side One of their conclusions was that cell size should be no larger than the lag distance of the shortest semi-variogram range. In their hydrologic modeling study, Srinivasan and Arnold (1994) used a mode-based amalgamation technique for the soils data within each subbasin. The subbasins were defined using the r.watershed program in GRASS. Within each subbasin, the mode (or dominant) soil class and land-use attributes were determined and used for simulation of the entire subbasin. The effect of this amalgamation on the model results was not tested.

Purpose and Approach The purpose of this study was to assess the impact of using a vector-based mode amalgamation scheme (Figure 1) on the map polygon count, map accuracy, and the hydrologic model results at field and watershed scales.

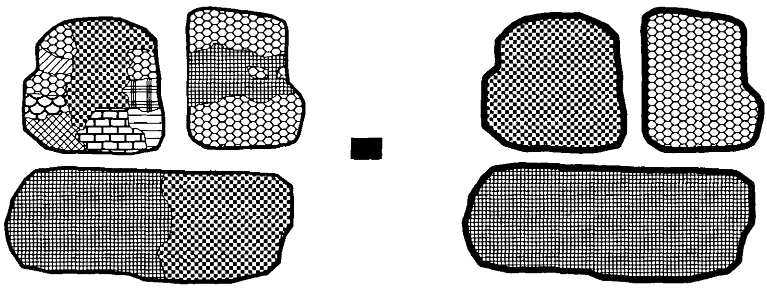

Figure 1 Application of the mode amalgamation technique to three regions (thick lines) Each class is represented by a separate shading pattern After the amalgamation each region contains only one class

Polygon coverages representing agricultural fields, soils, slope, aspect, and other pertinent factors were created and intersected to create a fundamental unit coverage (Figure 2). Each polygon in this coverage is locally unique with respect to all the pertinent attributes (e.g , field, soil, slope, aspect). A set of fields was selected at random from the study area. The coverage was simplified using mode amalgamation within each agricultural field. The amalgamations used were (1) slope, (2) aspect, (3) slope and aspect, and (4) soils, slope, and aspect (Figure 2). An amalgamated theme was reduced to one measure for each field polygon by assigning to all polygons in the field the mode (or dominant) class within the field and then merging these polygons (Figure 1). These coverages were incorporated into the GIS Interface for Ground-Water Models (Stallings *et al.*, 1992). GLEAMS simulations were run on each field using the fundamental unit coverage and four simplified versions of the coverage

The unsimplified fundamental unit coverage was used as a benchmark and is referred to as either the standard coverage or the standard fundamental unit coverage. The standard runs used this coverage to derive simulation inputs. Results based on the amalgamated themes are always compared to the results based on the standard run. Simulations used typical parameters for Duplin County, assuming that corn was grown and alachlor applied. Alachlor was the most commonly used pesticide on the fields simulated. Annual model outputs of water and alachlor percolation and runoff and sediment loss were integrated into the GIS and summarized by field before analysis

METHODS

Study Site The study site in Duplin County, North Carolina, is the 2044 ha Herrings Marsh Run Watershed This watershed is part of the larger Goshen Swamp which serves

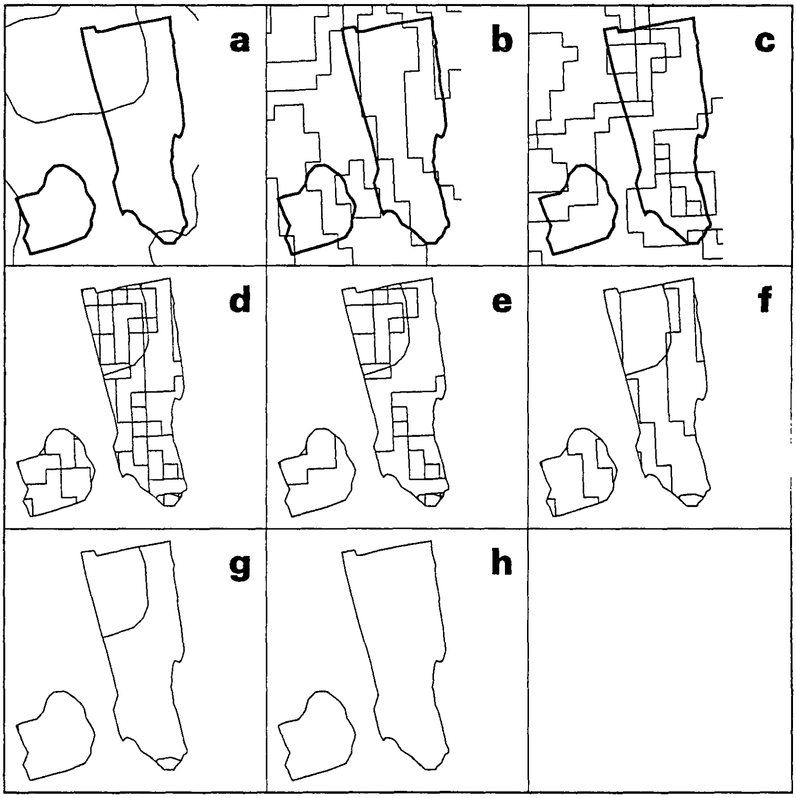

Figure 2 The original, intersected, and amalgamated themes near two of the agricultural fields Original themes agricultural fields (thick lines in a, b, and c), soils (a), slope (b), and aspect (c) Fundamental unit coverage (intersected and clipped themes) (d) Coverages resulting from amalgamations slope (e), aspect (f), slope and aspect (g), soils, slope, and aspect (h)

as the headwaters to the Northeast Cape Fear River. The watershed is in the coastal plain physiographic region. The soils are generally sandy; the major crops are soybeans, corn, cotton, tobacco, and hay.

The agricultural land is composed of 363 fields. The database contained adequate field, crop, pesticide, and soils data for 129 fields. Due to the extensive computer requirements, it was desirable to limit this study to an area represented by approximately 1000 polygons Of the 129 fields, 41 were chosen at random Intersection with the slope, aspect, and soils themes, resulted in 990 fundamental unit polygons within the fields. Of these, 147 were less than 5 square feet and were ignored in the simulation runs The area ignored was less than 0.0003% of the 41 fields. One agricultural field was represented by two distinct polygons, thus the fields are represented by 42 field polygons.

Software The GIS aspects of the work--digitizing, intersecting, mapping--primarily used ESRI's Arc/Info software, although Atlas*GIS was used for the initial digitization of the field coverages. GLEAMS version 1.8.55 (Davis *et al.*, 1990; Leonard *et al.*, 1987) was used to model water and pesticide transport.

The GIS Interface for Ground-Water Models (Stallings *et al.*, 1992) was used to overlay the coverage and run the model simulations for the watershed This software provides a linkage between Arc/Info, tabular databases, and GLEAMS. It allows one to perform GLEAMS simulations for many fields simultaneously (Stallings *et al.*, 1992). Graphical and statistical analyses were done using Splus version 3.2 (Statistical Sciences, Inc., 1993).

Spatial and Tabular Data Details of the spatial and tabular data can be found in Stallings *et al.* (1992) or Stallings (1995), they will be described here briefly. Polygon coverages describing agricultural fields, soil series, slope, aspect, and the watershed boundary were used.

The fields coverage was produced using a Zoom Transfer Scope and aerial photography. Identical cropping and pesticide practices were assumed on all fields studied. The cropping practices for corn (e.g., planting dates, field practices) were based on common usage within the county as determined by a Duplin County extension agent (Curtis Fountain, pers. comm., 1992); the typical planting date was March 15 and the harvest date was September 15 A rooting depth of 30" was used for these scenarios, therefore percolation loss values represent the loss below 30". Alachlor application was simulated on April 1 at a rate of 1.4 kg/ha. This application was based on actual alachlor applications as determined by a Cooperative Extension Service field survey of the watershed. Generalized information about Leaf Area Index (LAI) was taken from the GLEAMS manual (Davis *et al.*, 1990).

The soils coverage was digitized from a 1:24,700 prepublication map provided by the Soil Conservation Service (SCS). The SCS also provided the most recent version of their detailed soils attribute data for Duplin County. Generalized data on soil hydrologic conductivity, structure, and texture were taken from the GLEAMS manual (Davis *et al.*, 1990) and linked to the soils data.

The slope and aspect polygon coverages were based on 1:24,000 USGS digital elevation models. The DEMs were smoothed with a low-pass filter prior to calculation of the slope and aspect

The daily rainfall data for the closest weather station (Warsaw, NC) was taken from the Hydrological Information Storage and Retrieval System (Wiser, 1975). Missing values were filled with the average value for that date determined using the six years of available data (1985-1990). Monthly average high and low temperatures were also derived from the daily temperature data for this weather station. Average monthly solar irradiance was calculated based on data from the Raleigh, NC, weather station.

The corn and alachlor simulations were run on 10 years' data using identical cropping, pesticide, and 1990 weather data. After the first three years of simulation, the outputs stabilized to within three significant figures of the final values. This was also true in the Bleeker *et al.* (In Press) study

Coverage Preparation GLEAMS, at least as it is applied by the GIS-model interface, assumes polygons are homogenous with respect to the input themes. GIS intersection combines all pertinent attributes into a single coverage. The coverage breaks the fields into

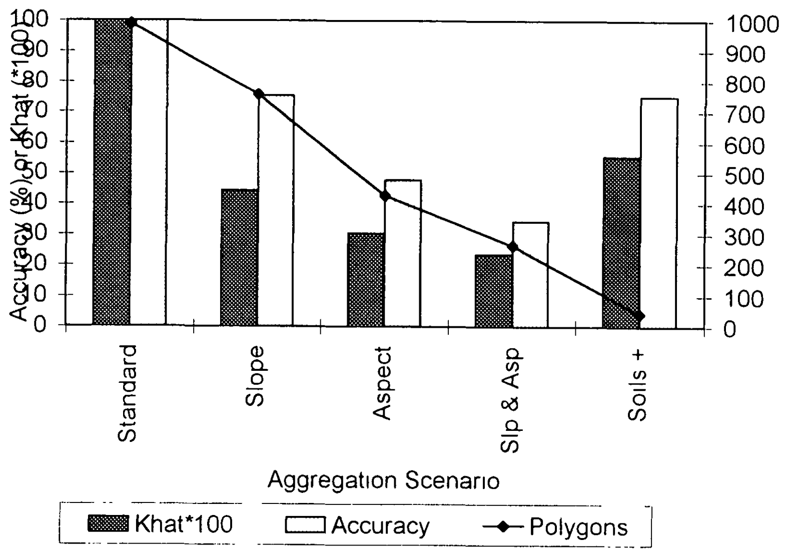

Figure 3 Change in polygon count, overall map accuracy, and Khat with different amalgamation scenarios

fundamental unit polygons that are locally unique with respect to their field, slope, aspect, and soils. The intersection and maximum overland flow calculations were performed as described in Stallings *et al.* (1992), except that the spatial-simplification routine in the GIS-model interface was not used. The coverage resulting from the intersection is the standard fundamental unit coverage; it, and the results based on it, was the benchmark against which changes were compared

Four separate amalgamations were made to the fundamental unit coverage: slope, aspect; slope and aspect; and soils, slope, and aspect (Figure 3). In these four amalgamations, one or more themes within a field were replaced with the dominant category of each theme within that field.

<u>Simulation Runs and Initial Data Analysis</u> The fundamental unit coverage and the four simplified coverages were used as input to parallel simulations using GLEAMS and the GIS-model interface Thus the slope, aspect, soil, number of polygons per field, and parameters based on polygon size and orientation (e.g., maximum overland flow length) varied with each amalgamation scenario. Inputs relating to the weather and field treatments remained constant across all polygons in each scenario.

Annual values of water and alachlor percolation and runoff and sediment loss were output and integrated into the GIS. The annual values associated with each polygon for the tenth year were extracted. Field summaries were then calculated by taking the spatially weighted average of the polygon results within each field.

STATISTICAL TESTS

Descriptive and test statistics were used to identify which amalgamation scenarios caused significant differences in the model results. Differences are always compared to model results based on the standard fundamental unit coverage; these results are considered correct. The absolute values of the results and their paired differences were judged to be non-normal based on quantile-quantile plots, so non-parametric and descriptive statistics were used for most analyses. The statistics examine results by model output parameter and amalgamation scenario at watershed and field scales. Each test produced a table of results. Conclusions are drawn based on the general trends. Watershed means were used to quantify differences in the entire population. Kendall's Tau rank correlation coefficient and the number of fields correctly classified into the top quartile are used to show the differences in field rankings. The distribution of field-specific relative errors are used to describe the differences in outputs at the field scale.

Watershed Means A t-test based on meaningfully paired differences (Steel and Torrie, 1980) was used to assess the deviation of the watershed means for each output parameter caused by each amalgamation. This test reduces the field-to-field variation. This test assumes normally distributed differences, which is not the case, so the results should be viewed with skepticism.

Kendall's Tau Rank Correlation Coefficient A likely use of the GIS-model interface is in ranking fields with respect to pesticide losses This does not require good results in the absolute sense or even a good linear correlation. Kendall's Tau (Kruskal, 1958) was used to measure the rank correlation of field specific outputs based on different map amalgamations The ranks resulting from the standard run are always used as the basis of comparison. The interpretation of Tau is similar to that for the standard Pearson sample correlation coefficient, except of course that Tau does not require a normal distribution

Relative Errors The relative error of field-specific values was calculated. Values from the standard simulation run are used as a reference. The equation is

Relative_Error = (Output_Amal. - Output_Standard) / Output_Standard

Map Accuracy Two measures are used to quantify map accuracy after amalgamation. overall map accuracy and the Khat statistic. Each is derived from an error matrix

The Khat statistic (Congalton and Mead, 1983, Hudson and Ramm, 1987) measures the amount of agreement between two maps accounting for chance agreement. Perfect agreement is indicated by a Khat of 1. One or more attributes of the standard fundamental unit coverage are used as the reference data. An error matrix is created by comparing an attribute of the fundamental unit coverage with the same attribute from a simplified coverage. A complete sample of the agricultural fields was taken The cells of the error matrix created did not represent discrete samples, but rather the area of agreement, in square feet, between the classes

RESULTS

Polygon Counts and Map Accuracy The map simplifications had the desired effect of decreasing the map polygon count (Figure 3) from 990 to as low as 42. The overall accuracy and Khat statistic decreased substantially with the amalgamations (Figure 3) The accuracy measures are based on only one attribute per map, except for the dominant slope and aspect coverage. The accuracy measures for this map consider each unique pair of

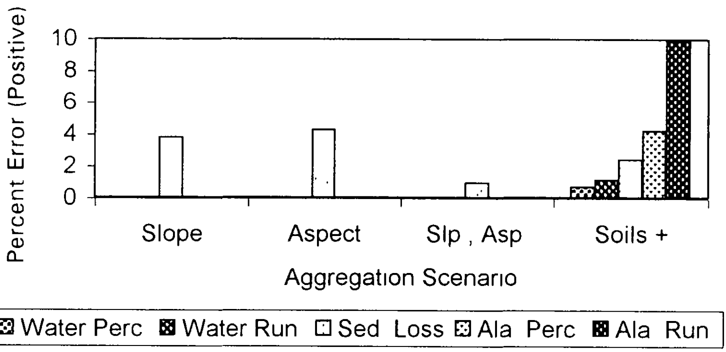

Figure 4 Percentage error in watershed means of model outputs for different amalgamation scenarios

slope and aspect as a class (e.g., slope = 1 and aspect = 2). The map accuracy for the soils simplification represents the accuracy of the soils data and ignores the fact that the slope and aspect were also amalgamated for these fields.

Watershed Means Many of the watershed means based on the amalgamation of slope and aspect were identical to the standard run means to two or three significant figures The exception was sediment loss, which varied a small amount (Figure 4). However, the runs based on the soils simplification show larger differences. None of the means reported here are significantly different from those based on the standard run. The soil simplification did cause a significant difference in soil-borne alachlor runoff loss (see Stallings (1995) for details).

Field Rankings The rank correlation coefficients comparing the standard and amalgamated themes are shown in Figure 5 Most of the Tau values showed significant departures from zero It was generally appropriate to accept the hypothesis that the results were correlated. Two Tau values are similar to zero ($P(H_0) > 0.1$), these result from the soil amalgamation

The low Tau values can be used to identify scenarios differing more from the standard

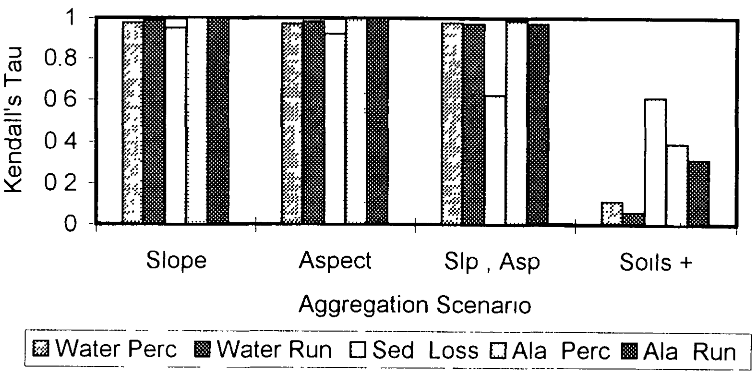

Figure 5 Kendall's Tau (rank correlation coefficient) between field results based on the standard and amalgamated data

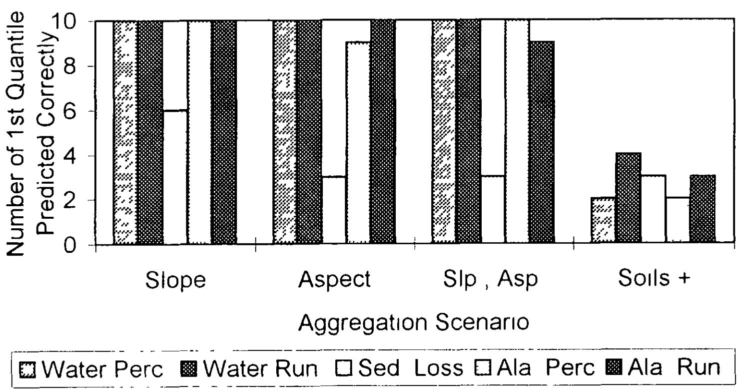

Figure 6 Number of fields correctly placed in the first quartile using results based on amalgamated data

run. The slope and aspect scenario resulted in ranks that were similar to the standard run for most output values (Tau values above 0.9). However, sediment loss shows a bigger difference (Tau values of 0.622). The soils amalgamation caused large differences in model outputs (Tau values ranging from 0.122 to 0.610).

The frequency with which the same fields were classified as being in the top quartile for different scenarios was investigated (Figure 6). The slope and aspect amalgamations caused the fewest errors, matching nine or ten of the ten fields for the majority of output parameters. Sediments loss, however, had more errors, matching only three to six of the ten fields. The soils amalgamation resulted in many errors, matching from two to four of ten fields.

Field Error The field-specific errors are perhaps the largest, since there is less averaging and abstraction to hide errors. This error is pertinent since fields are generally the smallest managed units. The error is presented as absolute relative errors (i.e , 0 10 represents a plus or minus 10% relative error). The field-specific relative errors for each model output and amalgamation scenario are shown in Figure 7. Once again the slope and aspect amalgamations introduced little error into the output values as indicated by the very narrow distributions of relative errors (left three columns). The exception once again is sediment loss (bottom row) The soils amalgamation resulted in much larger relative errors, ranging from approximately -0.8 to 0 9 (right column)

DISCUSSION

A number of factors are considered in this study the theme being amalgamated, field vs. watershed scale, absolute vs. rank results, and output parameter. The statistics used tested hypotheses related to single factors These are used to extrapolate to general trends in the data.

Clearly, the polygon count is greatly reduced by the amalgamation procedures. This results in substantial CPU savings since the CPU use is proportional to the polygon count. The map accuracy is also approximately proportional to polygon count. This relationship is hidden in Figure 3 because the stated map accuracies for the soils amalgamation do not account for the amalgamation errors of slope and aspect.

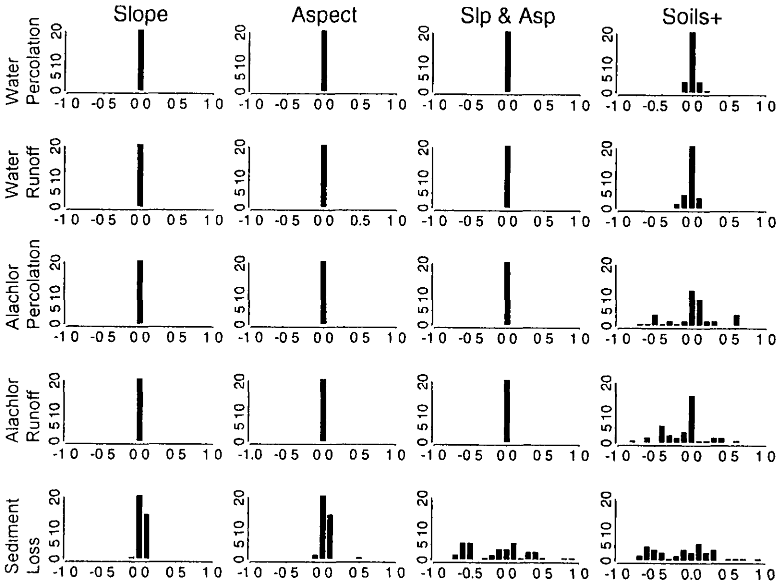

Figure 7 Histograms of field-specific relative errors for each model output and amalgamation scenario

One major advantage of the elimination-based techniques is that they can be applied incrementally. That is, the degree of generalization and map error can be controlled by changing the threshold used to select the polygons to be eliminated. The mode-amalgamation technique cannot be applied at various levels, although it can be applied to a single theme at a time, whereas the elimination techniques are generally applied to all themes at once in a combined coverage.

There does not appear to be a relationship between the accuracy of the input maps and the accuracy of the model outputs. This is due to the differing influences of the input maps. Greenland *et al.* (1985) concluded that, as a rule of thumb, Khat statistics of less than 0.90 to 0.85 indicated a map unfit for use. While the results based on soils simplifications support this, the results based on slope and aspect simplifications indicate that the sensitivity of the model to specific inputs must also be considered.

General Trends Several trends are apparent from the results It is clear that the soils amalgamation result in significantly greater error than the slope and aspect amalgamations However, the slope and aspect amalgamations do result in errors of sediment loss estimation This is due largely to the model's relative sensitivity to these inputs

The field-scale errors were greater than the watershed-scale errors. The soils amalgamation resulted in relative errors of up to 0.09 for watershed means. The same amalgamation resulted in field-specific maximum relative errors ranging from 0.177 to 0.854 (see Stallings (1995) for details) The field-scale errors tend to cancel one another in the watershed summaries, as would be expected as long as the errors caused by the amalgamation are unbiased.

The field-ranking results were similar to the field-specific results. Both the Tau values (Figure 5) and the first quartile predictions (Figure 6) show that the slope and aspect amalgamations affected model results less than the soils simplifications. The exception was sediment loss, which was affected by the slope and aspect amalgamations. The only statistically significant differences were, however, due to the soils amalgamation.

The errors due to slope and aspect amalgamations are probably acceptable in many of the contexts that GLEAMS may be used. Even when used with the best field data, GLEAMS predicts pesticide center of mass, and solute concentration distributions with depth to $\pm$ 50% (Leonard *et al.*, 1987; Pennell *et al.*, 1990). Clearly, however, some of the ranking and field-specific results are unacceptable

Limitations This study looked at only a few of GLEAMS many outputs. In particular It did not look at any daily or monthly outputs, or outputs of pesticide concentrations within the soils layers. It should also be emphasized that GLEAMS is a field-scale model. Neither GLEAMS nor the GIS-model interface accounts for routing between fields, from the edge of fields to streams, or from the bottom of the root zone to the saturated ground water. The watershed summaries are therefore indicative of mean loss at the edge of the fields and not of the amount leaving the watershed

Generality of Results The results presented here are specific to the GLEAMS model. The effects of amalgamation will be different for each model and output parameter, depending on their sensitivity to the input themes. However, most hydrologic models rely heavily on soils inputs and will probably be sensitive to changes in these inputs.

The results should be transferable geographically. The model sensitivity to different inputs will not change The relative importance of the inputs may vary geographically, however Amalgamation of the slope and aspect themes might cause more error in terrain with more relief. In areas with more heterogeneous soils, small but important soil series might be suppressed during amalgamation, causing a greater bias in the watershed means.

Future Work This method of amalgamation needs to be compared more rigorously to at least three other methods· (1) elimination of small polygons; (2) elimination of small polygons by field, and (3) amalgamation by spatially weighted means of the attributes. No single method will be best in all cases The elimination techniques have the advantage that they can be applied incrementally. The spatially weighted mean technique will probably result in smaller root-mean-square errors for numeric map data and perhaps in smaller variances in model results due to amalgamation. The mode method has the advantages that it can be applied to individual themes, it can handle class data, and it is simple

CONCLUSIONS

The slope and aspect themes are less important inputs for GLEAMS compared to the soils theme in the area investigated. They are not taken into account by GLEAMS directly when calculating either percolation or runoff volume. Slope and aspect did have an effect on outputs relating to sediment loss. An analysis of the model sensitivity as carried out by Lane and Ferrira (1980) could be used to determine which inputs are not important. Mode amalgamation of soils introduced significant error into the output parameters This will probably be true for most agricultural models predicting percolation and runoff since many use the same or similar equations to represent percolation and runoff process. Aggregating soils for model input may cause unacceptable errors in many circumstances.

In order to determine which results are acceptable, one must first determine what errors are acceptable for a particular study. Knowing this, one can determine which amalgamation scenarios are acceptable for a particular study. The error in large-area summaries due to the mode amalgamation is probably acceptable. The error in ranking and field-specific results are probably acceptable for specific output parameters and amalgamation scenarios (e.g., slope and aspect amalgamations, if percolation or runoff are the outputs of interest). In those cases where the error introduced by the amalgamation is acceptable, the technique clearly reduces the polygon count and therefore the computing effort required.

The most important conclusion of this study is that the amalgamation method and extent should be considered in context with the model and modeling objectives. At least two methods can be used to do this. In either case one must have determined the modeling objectives which model outputs are of interest, the accuracy desired, and whether site-specific values or areal summaries are of interest. The first method consists of identifying the error introduced into the input coverages by generalization and estimating how the errors will propagate through the GIS-model combination based on the model's sensitivity to the inputs derived from the coverages. Where a complete model sensitivity analysis has been done much of this analysis can be carried out with known data. The second method, used in this study, consists of implementing a pilot study. To do this one must implement the model on a representative data set. This has two advantages. It will give more exacting results for specific study locations. It will better represent the simultaneous changes of many soils parameters that occur as the areal representation of the soil series change

ACKNOWLEDGMENTS

Earlier drafts of this article were reviewed by Nancy White, David Adams, Hugh Devine, and three anonymous reviews; the article is much improved as a result. The SCS and the Duplin County office of the Agricultural Stabilization and Conservation Service supplied map and tabular data Mark Rice digitized the agricultural field boundaries. Vernon Cox, Dana Hoag, and Steve Coffey supplied agricultural and pesticide data. Staff at the Computer Graphics Center provided computer and administrative support throughout the project. This work was funded by the Cooperative State Research Service.

REFERENCES

Brassel K.E., Weibel R. (1988) A review and conceptual framework of automated map generalization. Int J Geographical Information Systems 2:229-244

Brown D.G , Bian L., Walsh S.J. (1993) Response of a distributed watershed erosion model to variations in input data aggregation levels. Computers & Geosciences 19(4):499-509

Congalton R.G., Oderwald R.G., Mead R.A. (1983) Assesing Landsat classification accuracy using discrete multivariate analysis statistical techniques. Photogram Engr and Remote Sensing 49:1671-1678

Davis F M., Leonard R.A., Knisel W.G. (1990) GLEAMS user manual version 1.8.55 USDA-ARS Southeast Watershed Research Laboratory, Tifton, Georgia, Lab Note SEWRL-030190FMD

Greenland A., Socher R.M., Thompson M.R. (1985) Statistical evaluation of accuracy for digital cartographic data bases. In: Auto-Carto 7: Digital representations of spatial knowledge. American Society of Photogrammetry, Falls Church, VA (Proceedings)

Hudson W.D., Ramm C.W. (1987) Correct formulation of the Kappa coefficient of agreement. Photogram Engr and Remote Sensing 53:421-422

Kruskal W.H (1958) Ordinal measures of association. J Amer Stat Assoc 53:814-861

Lane L.J., Ferreira V A (1980) Sensitivity analysis. In: Knisel WG (ed) CREAMS a field scale model for chemicals, runoff, and erosion from agricultural management systems. U.S. Dept of Agriculture, Washington, D.C , Conservation Research Report No. 26

Leonard R.A., Knisel W.G., Still D.A. (1987) GLEAMS: groundwater loading effects of agricultural management systems. Trans of the ASAE 30(5):1403-1418

McMaster R.B., Monmonier M. (1989) A conceptual framework for quantitative and qualitative raster-mode generalization. In GIS/LIS '89, vol Volume 2. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland (Proceedings)

McMaster R.B., Shea S.K. (1992) Generalization in digital cartography Association of American Geographers, Washington, D.C.

Monmonier M S. (1983) Raster-mode area generalization for land use and land cover maps. Cartographica 20:63-91

Pennell K D., Hornsby A.G., Jessup R E , Rao P.S.C. (1990) Evaluation of five simulation models for predicting Aldicarb and bromide behavior under field conditions Water Resources Res 26.2679-2693

Srinivasan R., Arnold J G. (1994) Integration of a basin-scale water quality model with GIS Water Resources Bulletin 30(3).453-462

Stallings C. (1995) GIS data integration for ground-water quality modeling Ph.D Dissertation, North Carolina State University. In preparation.

Stallings C , Huffman R.L., Khorram S., Guo Z (1992) Linking GLEAMS and GIS ASAE, St Joseph, Michigan (Written for the International Winter Meeting of the ASAE, Nashville, TN, 15-18 December)

Statistical Sciences (1993) S-PLUS User's Manual, Version 3.2 StatSci, a division of MathSoft, Inc., Seattle

Steele R.G D , Torrie, J.H (1980) Principles and procedures of statistics (2nd) McGraw-Hill, Inc., New York.

Wang F., Donaghy P. (1992) Assessing the impact of automated polygon elimination to overlay analysis accuracy. In. ASPRS/ACSM/RT 92 technical papers, vol 4: Remote sensing and data aquisition American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland

Whede M. (1982) Grid cell size in relation to errors in maps and inventories produced by computerized map processing. Photogram Engr and Remote Sensing 48:1289-1298

Wiser E.H. (1975) HISARS -- Hydrologic information storage and retrieval system -- reference manual. National Technical Information Service, Springfield, Virginia