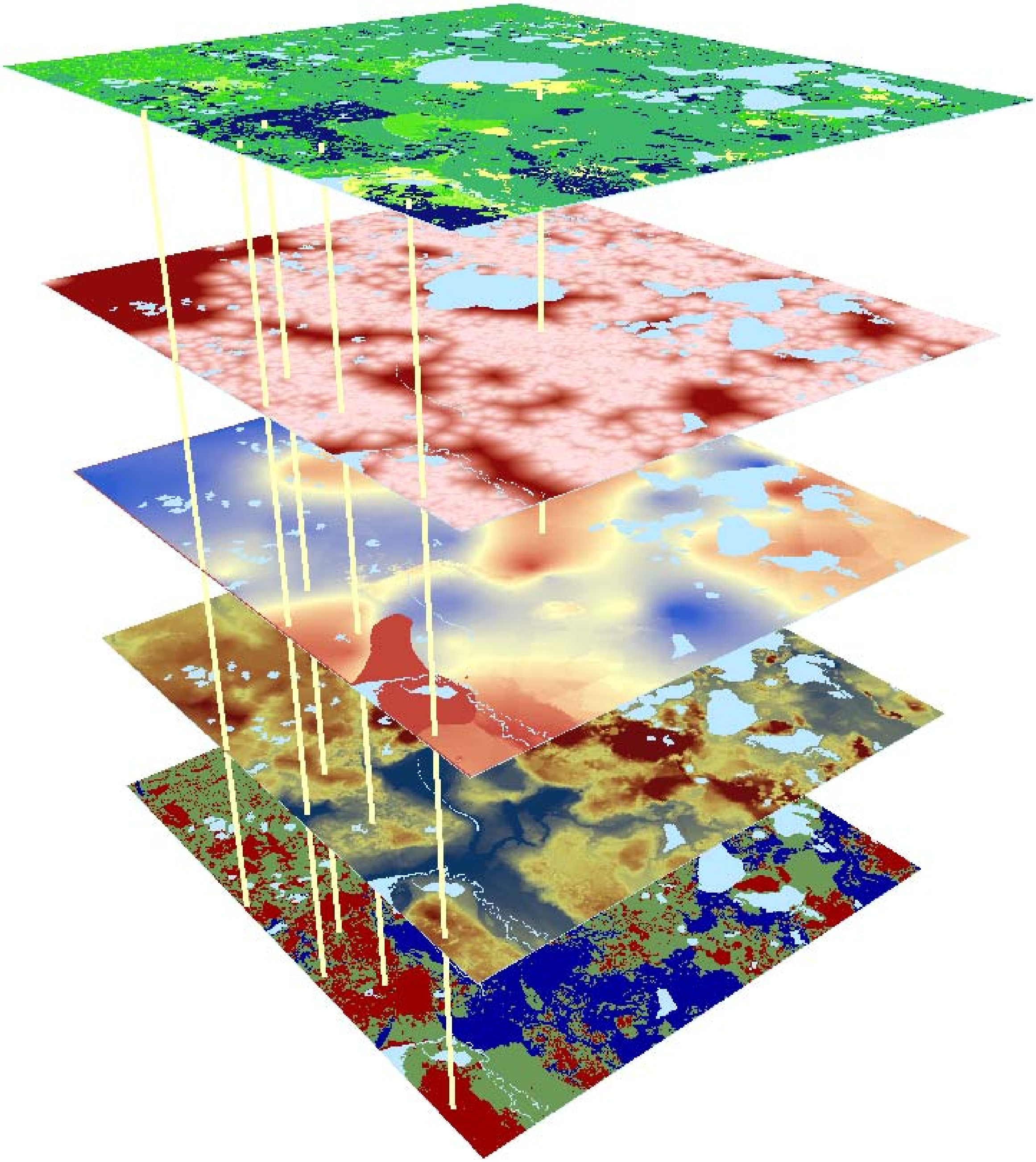

WEKIVA AQUIFER VULNERABILITY ASSESSMENT (WAVA)

Report submitted to the Florida Springs Initiative and the

Wekiva River Basin Coordinating Committee

By the Florida Geological Survey of the Florida Department of Environmental Protection

February 9, 2005

James R. Cichon, Alan E. Baker, P.G. 2324, Alex R. Wood, and Jonathan D. Arthur, P.G. 1149

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Florida's aquifer systems are of the most important and unique hydrogeologic resources in the nation. A remarkable 90% of Florida's drinking water originates from these ground-water resources (Berndt et al., 1998). More than 700 springs are known to exist in Florida, allowing a unique window into the aquifer systems (Scott, 2004). These springs have long been a resource for recreation and fresh water. Protection of Florida's springs becomes an increasingly important issue as the population of Florida continues to grow at a rate of almost 900 people per day (U.S. Census, 2005). As a result, pressures on these unique natural resources become intensified.

It is important to note that water quality of Florida's springs is a reflection of the condition of the surrounding ground-water system and not just an indication of individual spring health. Springsheds, or the spring's area of ground-water contribution, must also be taken into consideration for any vulnerability assessment if proper protection of the spring itself is to take place. An assessment of the vulnerability of the ground water which includes the springshed can therefore be a potential resource for decision making, development of rules or policies regarding environmental conservation, protection, growth management, and planning.

Due to Florida's unique hydrogeologic setting, all of Florida's ground water is vulnerable to contamination. In fact, this statement, in a more broad sense, is considered the "First Law of Ground Water Vulnerability" by the National Research Council (NRC, 1993) which states: "All ground water is vulnerable." Furthermore, the NRC defines the phrase "ground-water vulnerability to contamination" as the tendency or likelihood for contaminants to reach a specified position in the ground-water system after introduction at some location above the uppermost aquifer. In this report, a similar definition of aquifer vulnerability is adopted: The tendency or likelihood for contaminants to reach the top of the Floridan Aquifer System (FAS) after introduction at land surface.

The hydrogeology of the Wekiva River study area is characterized by moderate to no confinement and a multitude of karst features (Figure 1). Ground water recharges the FAS by infiltration through these sediments or directly through sinkholes. The Wekiva River Coordinating Committee Final Report identified numerous studies by Florida's water management districts and the United States Geological Survey (USGS) that clearly demonstrate contamination attributable to changes in land use. In other words, what occurs on the land directly and indirectly affects the quality of water moving through the subsurface conduits. Many of these conduits flow to surface-water bodies via springs.

The Florida Geological Survey (FGS) applied a modeling technique used in the statewide Florida Aquifer Vulnerability Assessment (FAVA) project to the Wekiva study area in order to delineate primary, secondary and tertiary protection zones for the FAS based on available data. This project is known as the Wekiva Aquifer Vulnerability Assessment Model or WAVA. The modeling technique used is known as Weights of Evidence (WofE), which is based entirely in a geographic information system (GIS) and is briefly described below. For a more thorough explanation of this method refer to FGS Bulletin 67, *Florida Aquifer Vulnerability Assessment: Contamination potential of Florida's principal aquifer systems* (Arthur et al., 2005, in preparation).

The software used to execute WofE and the Florida Department of Environmental Protection (FDEP) Albers projection require certain data layers to be listed in meters. However, some data layers are listed in units standard to that particular data set. For reader convenience we list the primary unit followed by the alternative unit in parentheses.

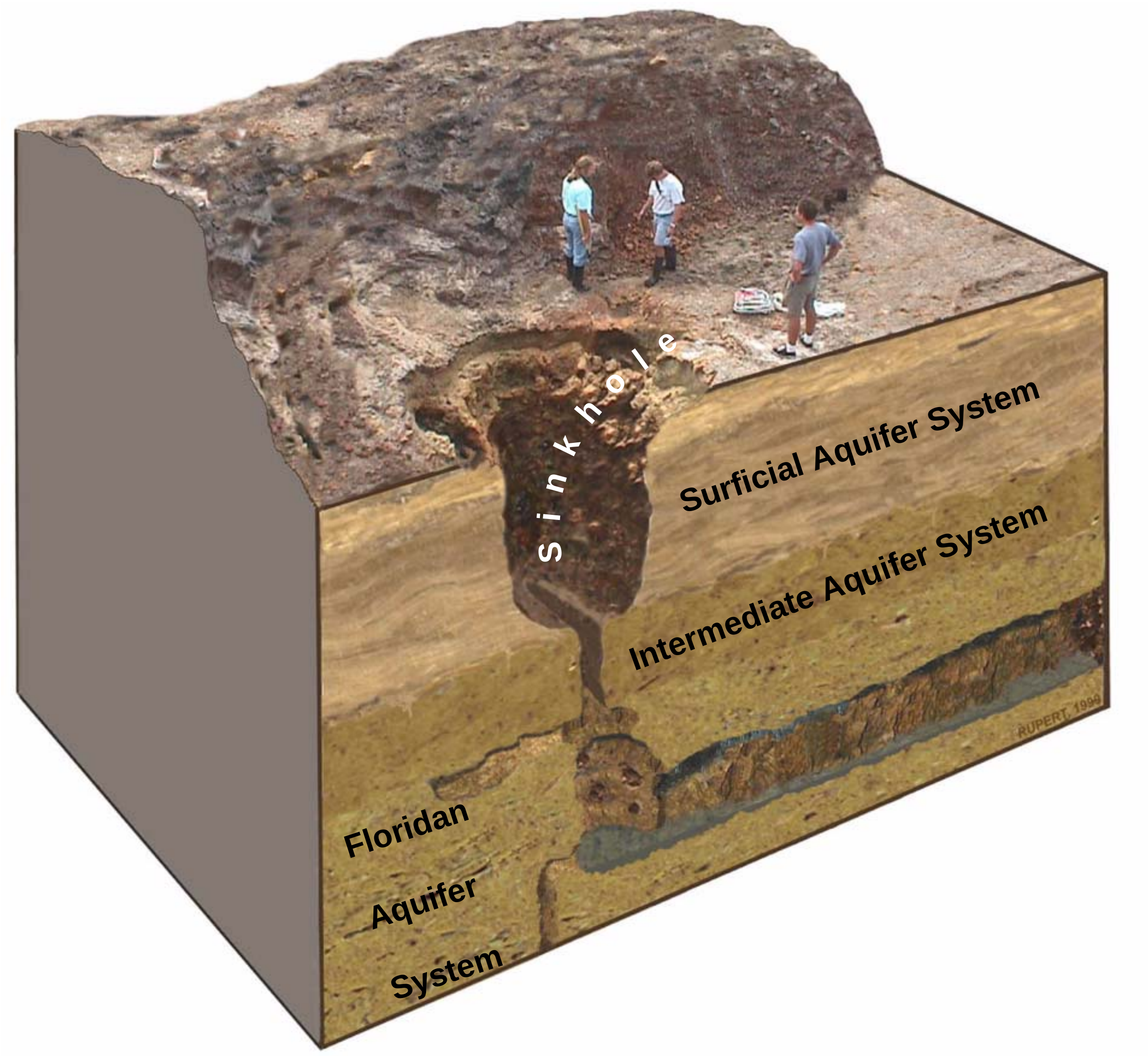

Figure 1. Hydrogeologic conceptual model of the Wekiva River study area depicting the major aquifer systems and karst features.

WEKIVA AQUIFER VULNERABILITY ASSESSMENT MODEL

Weights of Evidence

Use of WofE requires the combination of diverse spatial data which are used to describe and analyze interactions and generate predictive models (Raines et al., 2000). When applied in the WAVA project, WofE was used to generate maps of aquifer vulnerability, or response themes. These response themes were generated in the Environmental Systems Research Institute ArcView 3.2 environment. WofE was executed using the Arc Spatial Data Modeler which is available as an internet download (Kemp, et al., 2001).

A primary benefit of applying WofE to the WAVA project is that it is data-driven, rather than expert-driven. The data that "train" the model consist of known occurrences of parameters (water quality values) that reflect relative aquifer vulnerability, such as high levels of dissolved oxygen in ground-water wells. These wells were the training points used to calculate weights for laterally continuous input data layers, or evidential themes, which are then combined to yield a response theme.

When reviewing the model results, it is important to note that all aquifers, to some degree, are vulnerable to contamination from land surface. The model results simply identify those areas within the study area that are more vulnerable or less vulnerable based on the evidential themes and training points used in the model. So, the vulnerable areas are assigned values relative to each other within the study area. These "relative" values are directly transferable or correlated to another study area.

Study Area

The initial step in the development of the WAVA model was the delineation of a study area extent. The study area, composed of 30 m^2 grid cells, was used in the calculation of weights and probabilities throughout the modeling process. The extent of the study area used for input into the WAVA model was based on a combination of the Wekiva study area (as identified in the Wekiva River Basin Coordinating Committee Final Report, 2004) and the area contributing ground water to the Wekiva River (Figure 2).

Large water bodies were omitted from the WAVA model because a well would never be drilled in these areas – therefore, they would never contain a training point. Additionally, if lakes were left in the model, the surface area was increased with no chance of increasing the number of training points. This would unnecessarily bias the model. Further, large water bodies typically have no soils or other input data associated with them, thus the model output omits these areas due to lack of data or potential bias in the calculated probabilities.

Due to the geostatistical framework and evidential layers (spatial hydrogeological data) of WAVA, aquifer systems in discharge areas were sometimes predicted by the output model to be low in vulnerability, even though the discharging surface waters are highly vulnerable to contamination. These discharging surface waters are not part of the aquifer, although they can originate from it. The WAVA project was designed to focus on the ability for a contaminant to travel through soils, overburden, karst features, etc. to enter into the aquifer system. As a result, it is very important that the WAVA model never be applied to assess contamination of surface waters or discharge areas.

Training Points

Training points are locations of known measured occurrences. In mining applications for example, identified ore reserves are known occurrences. In an aquifer vulnerability assessment, wells with water quality indicative of high recharge are potential known occurrences (i.e., areas where a good connection exists between the top of the aquifer and land surface). Training points are used in WofE to calculate the following parameters: prior probability, weights for each evidential theme, and posterior probability of the response theme. Training points are converted to represent a unit area of the study area, such as a grid cell within a GIS application.

The water-quality parameter selected for the WAVA training data set was dissolved oxygen. Background levels of oxygen in the FAS are typically low where the aquifer system is not affected by activities at land surface. Therefore, where dissolved oxygen occurs at concentrations above background levels in an aquifer system, it can generally be assumed a relatively greater hydrologic connection exists between

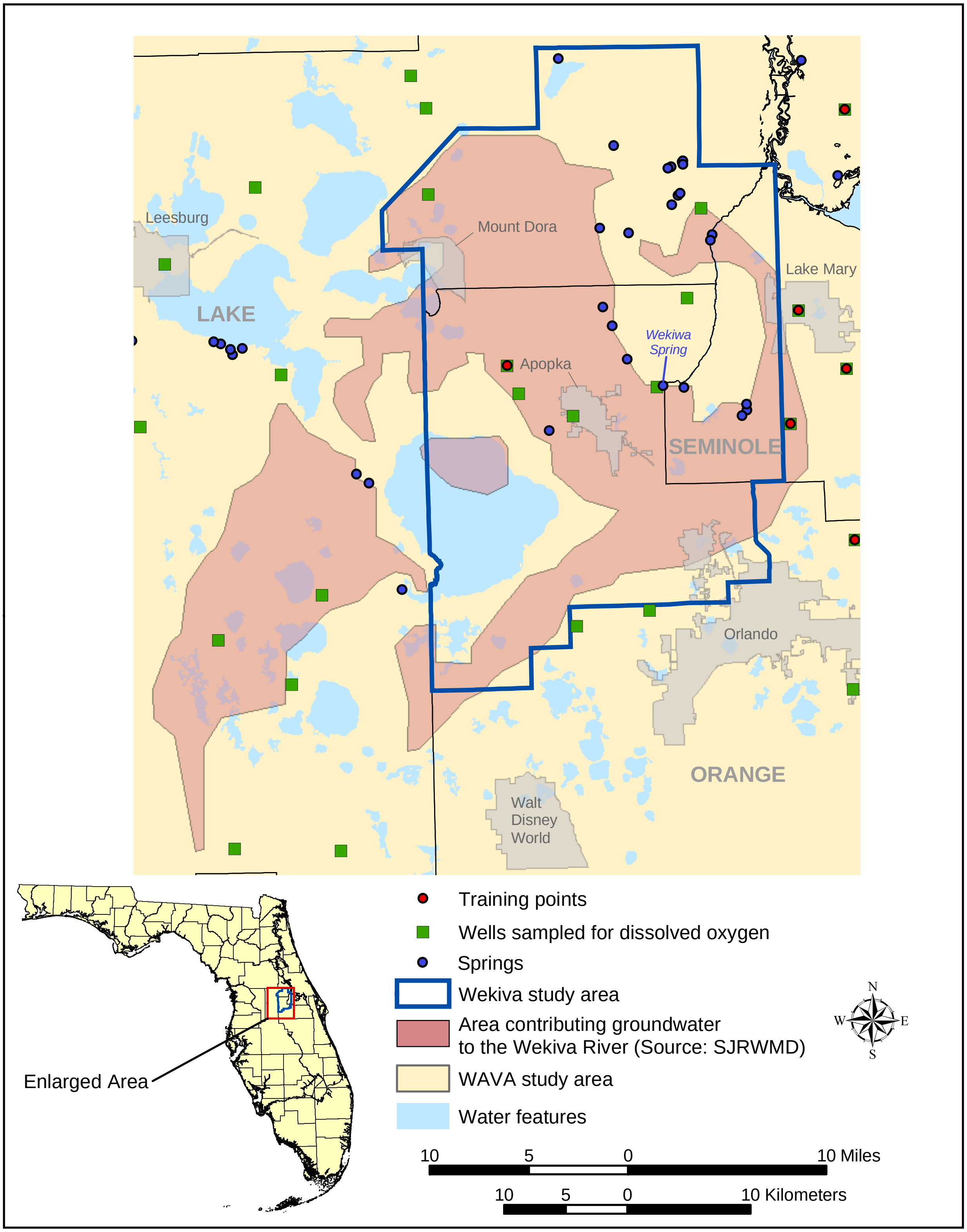

Figure 2. Extent of the WAVA study area and training point set developed based on measured dissolved oxygen. Large water bodies have been omitted from the analysis to avoid biasing the model.

land-surface activities and ground water. Dissolved oxygen analytical values served as the training point set for the WAVA model.

There were a total of 202 wells that were completed in the FAS for the WAVA study area. These wells were compiled from the FDEP Background Water Quality Monitoring Network, the FDEP STATUS network, and the St. John's River Water Management District monitoring network. Of these wells, 28 were measured for dissolved oxygen. Using statistical methods, described in Arthur et al. (2005, in preparation), two wells were identified as outliers and removed from the dataset leaving 26 wells for further analysis. The 75th percentile value was the statistical break used to identify dissolved oxygen values most indicative of a surface water influence in ground water, which is considered herein as a proxy for areas of higher recharge and vulnerability. The 75th percentile median value for dissolved oxygen concentration was 0.4 mg/L. Six wells existed in the dataset that had a measured median dissolved oxygen value greater than the 0.4 mg/L threshold. These six wells comprised the training point theme for input into the WAVA model.

The chance that a training point will occupy any given unit area within the study area, independent of any evidential theme data, is known as prior probability. The prior probability is calculated by dividing a unit area representing the training points by the total study area. In less complex terms, the prior probability is based on prior knowledge of the problem without the benefit of supporting evidence. The prior probability for the WAVA model was calculated at 0.0017. The distribution of all wells sampled for dissolved oxygen, and the wells meeting training point criteria are displayed in Figure 2.

Evidential Themes

Several evidential themes were considered for use in the WAVA model due to their potential influence on ground-water quality. They are:

- Soil permeability
- Buffered effective karst features (based on closed topographic depressions)
- Thickness of Intermediate Aquifer System (IAS)
- Hydraulic head difference between Surficial Aquifer System (SAS) and FAS
- Vertical leakage rate to and from the FAS

Vertical leakage rate to and from the FAS ultimately was not used because it was based on model simulations with a grid-cell size of 1,000 meters and there was large variability in the calculated leakage rate between adjacent grid cells. Weights were calculated for each of the other four evidential themes and generalized as discussed below.

Generalization of Evidential Themes

Evidential themes were generalized in an effort to assess which areas of the evidence shared a greater association with locations of training points. During calculation of weights for each evidential theme used in the WAVA project, a contrast value was calculated for each class of the theme by combining the positive and negative weights (positive weight – negative weight). Contrast is a measure of a theme's significance in predicting the location of training points and helps to determine the threshold or thresholds that maximize the spatial association between the evidential theme map pattern and the training point theme pattern (Bonham-Carter, 1994).

Confidence of the evidential theme equals the contrast divided by the standard deviation (a student T test) for a given evidential theme and provides a useful measure of significance of the contrast due to the uncertainties of the weights and areas of possible missing data (Raines, 1999). A confidence value of 1.282 corresponds to a 90% level of significance (see Table 1), which was the value selected as the minimum acceptable confidence level for the WAVA project evidential themes. Confidence values approximately correspond to the statistical levels of significance listed in Table 1.

Table 1. Test values calculated in WofE and their respective studentized T values expressed as level of significance in percentages.

Studentized T Value (confidence expressed as level of significance)	Test Value
99.5%	2.576
99%	2.326
97.5%	1.960
95%	1.645
90%	1.282
80%	0.842
75%	0.674
70%	0.542
60%	0.253

Contrast values were used to determine where to sub-divide evidential themes into generalized categories. The most common method of categorizing an ordered evidential theme was to select the maximum contrast as a threshold value to create a binary generalized evidential theme. In some WofE models, categorization of more than two classes may be justified (Arthur et al., 2005, in preparation). For evidential themes used for the WAVA project, a binary break was typically defined by the WofE analysis thereby creating two spatial categories: one with stronger association with the training point theme and one with weaker association with the training point theme.

Soil Permeability

Soil permeability values were obtained from the soils statewide grid layer developed for FAVA (Arthur et al., 2005, in preparation). The WAVA study area contains the SSURGO database of permeability values. This database was mapped at a scale of 1:24,000 and contains the most detailed soil permeability data. However, different permeability values can exist along some county boundaries due to the different mapping techniques used by different soil scientists. The development of this layer included the calculation of average soil permeability values for each soil horizon layer. Then, based on soil horizon thicknesses, weighted-average permeability values were calculated for the entire soil column. This allowed the generation of a data coverage of soils containing a single permeability value per soil polygon. This data coverage is displayed in Figure 3. Missing-data areas on the figure represent urban land, pits and water features.

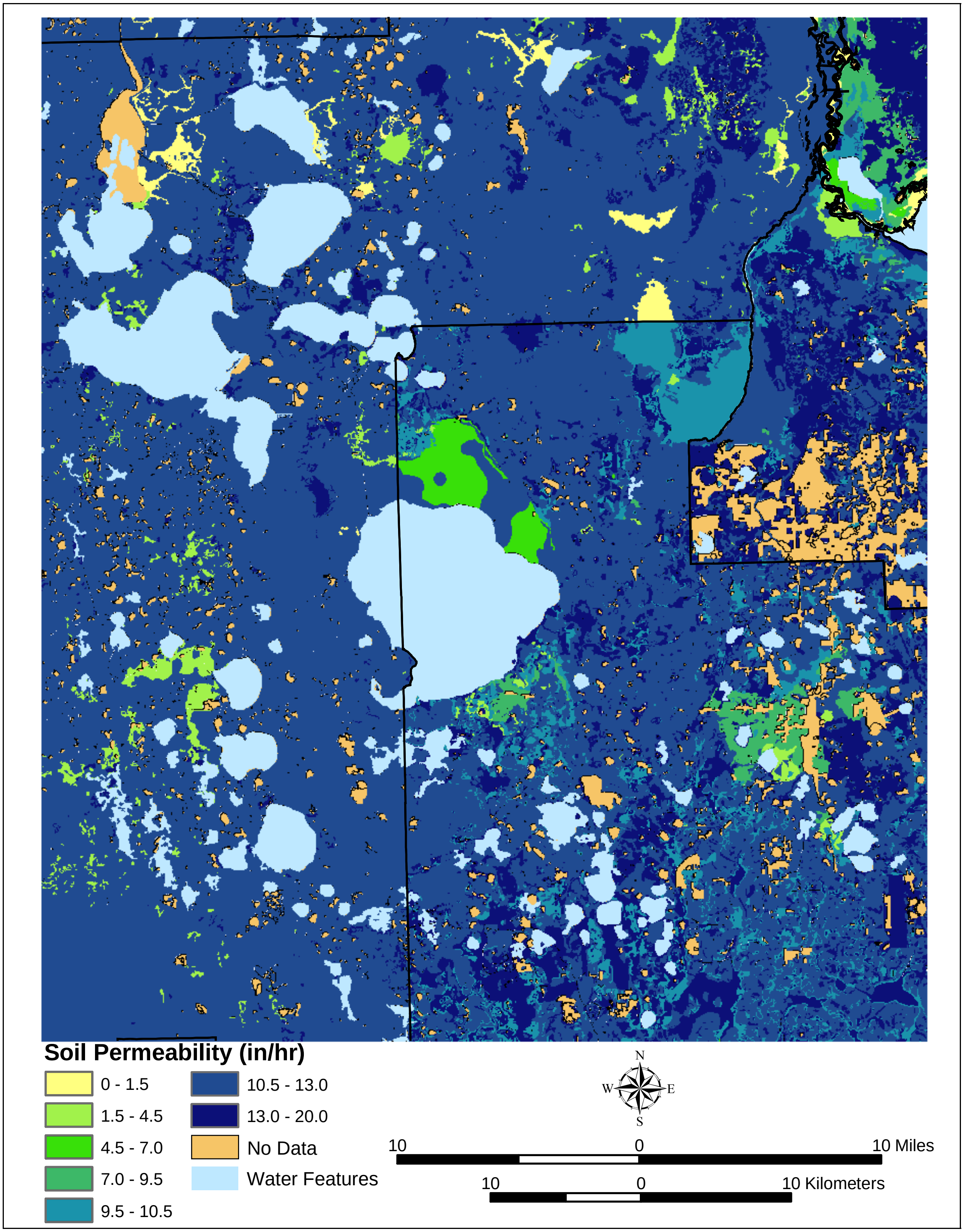

Figure 3. Soil permeability map of the WAVA study area (FGDL, 2003). Note differences across county lines due to different mapping techniques. Missing-data areas represent urban land, pits and water features.

Soil permeability is a measure of the rate at which water travels through the vadose (unsaturated) zone. Areas with high soil permeability values are normally associated with higher aquifer vulnerability. Weights were calculated for soil permeability using the cumulative descending method (refer to *Weights of Evidence Glossary*) of the WofE model technique. The highest contrast of any class was calculated at 13.1 in/hr (9.2 x 10^{-3} cm/sec) as displayed in Figure 4. Based on the calculated weights, the most appropriate break in the soil permeability evidential theme was at 13.1 in/hr (9.2 x 10^{-3} cm/sec), creating a binary generalized theme for input into the WAVA model. This contrast break indicated that values exceeding 13.1 in/hr (9.2 x 10^{-3} cm/sec) were strongly correlated with aquifer vulnerability as defined by the training point data, whereas values less than 13.1 in/hr (9.2 x 10^{-3} cm/sec) were less significant. The generalized theme is displayed in Figure 5.

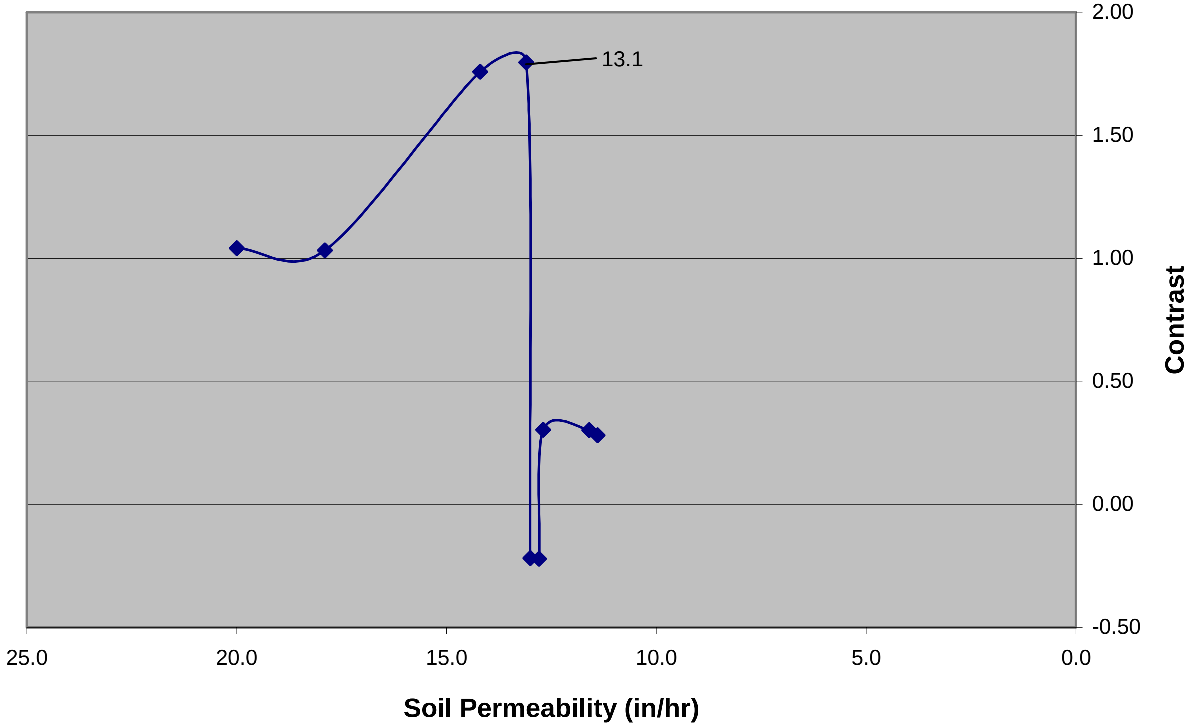

Figure 4. Soil permeability values (in/hr) plotted against contrast values calculated using WofE. The highest contrast value was calculated at 13.1 in/hr (9.2 x 10-3 cm/sec).

Buffered Effective Karst Features

To represent karst features (i.e., sinkholes) affecting vulnerability in the WAVA model study area, an effective karst GIS grid was generated based on a combination of closed topographic depressions identified in USGS 7.5-minute quadrangle maps and thickness of overburden on FAS (Figure 6). Closed topographic depressions (considered an estimation of karst feature distribution) are represented on USGS 7.5-minute maps as enclosed hachured contour lines. The effective karst feature evidential theme was developed by evaluating which of these closed topographic depressions were more likely to be hydrologically connected to the underlying FAS. This was accomplished by filtering out depressions underlain by more than 100 feet (30.5 m) of FAS overburden thickness. The 100-ft (30.5 m) threshold of overburden thickness has been used to identify karst-prone areas by Cichon et al. (2004) and Wright (1974). Though the location of training points was not used to select this filter threshold, the lack of their occurrence in areas underlain by more than 100 feet (30.5 m) of FAS overburden thickness lends support to the use of this filter.

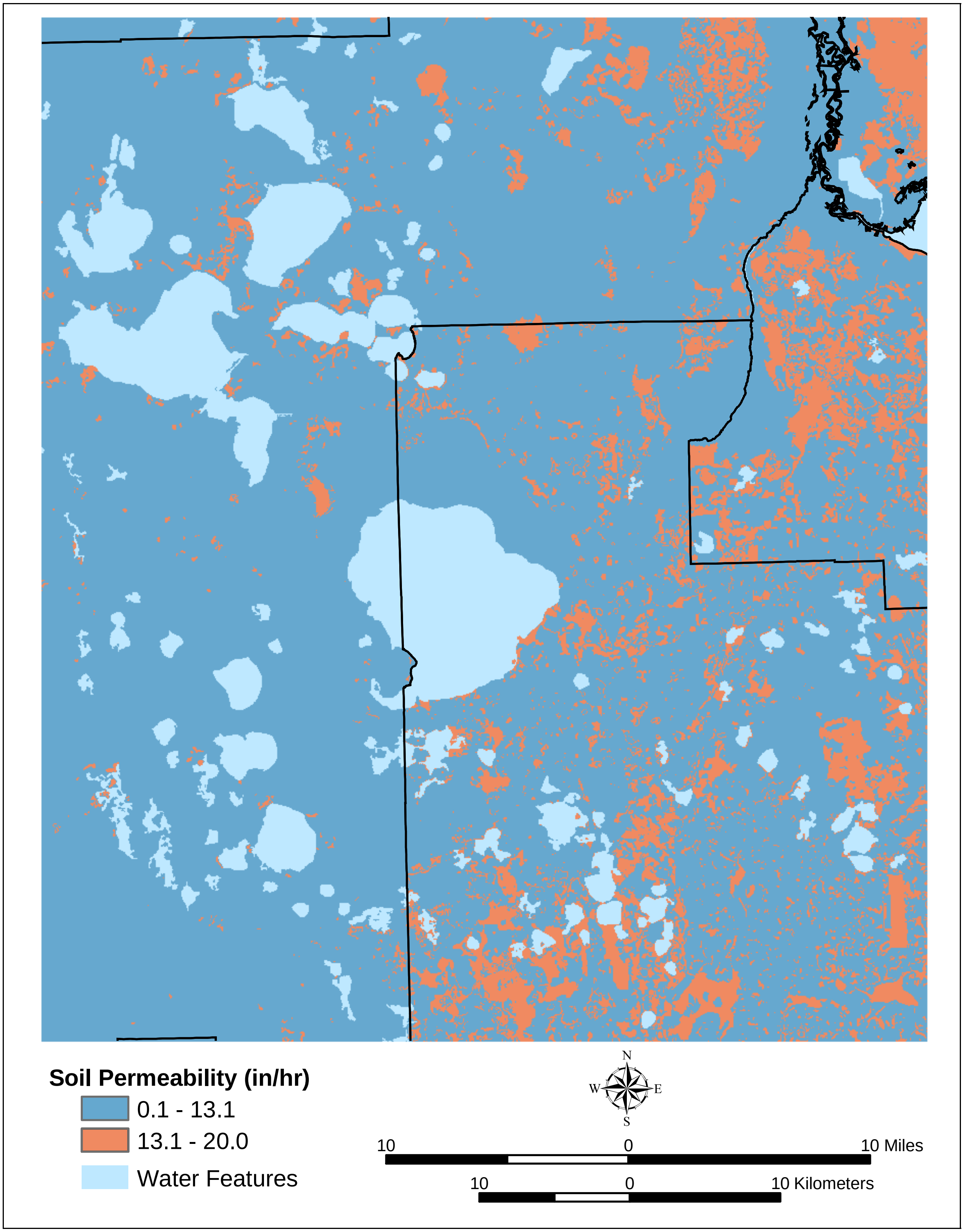

Figure 5. Map showing binary generalization of the soil permeability evidential theme, based on a contrast value of 13.1 in/hr (9.2 x 10-3 cm/sec).

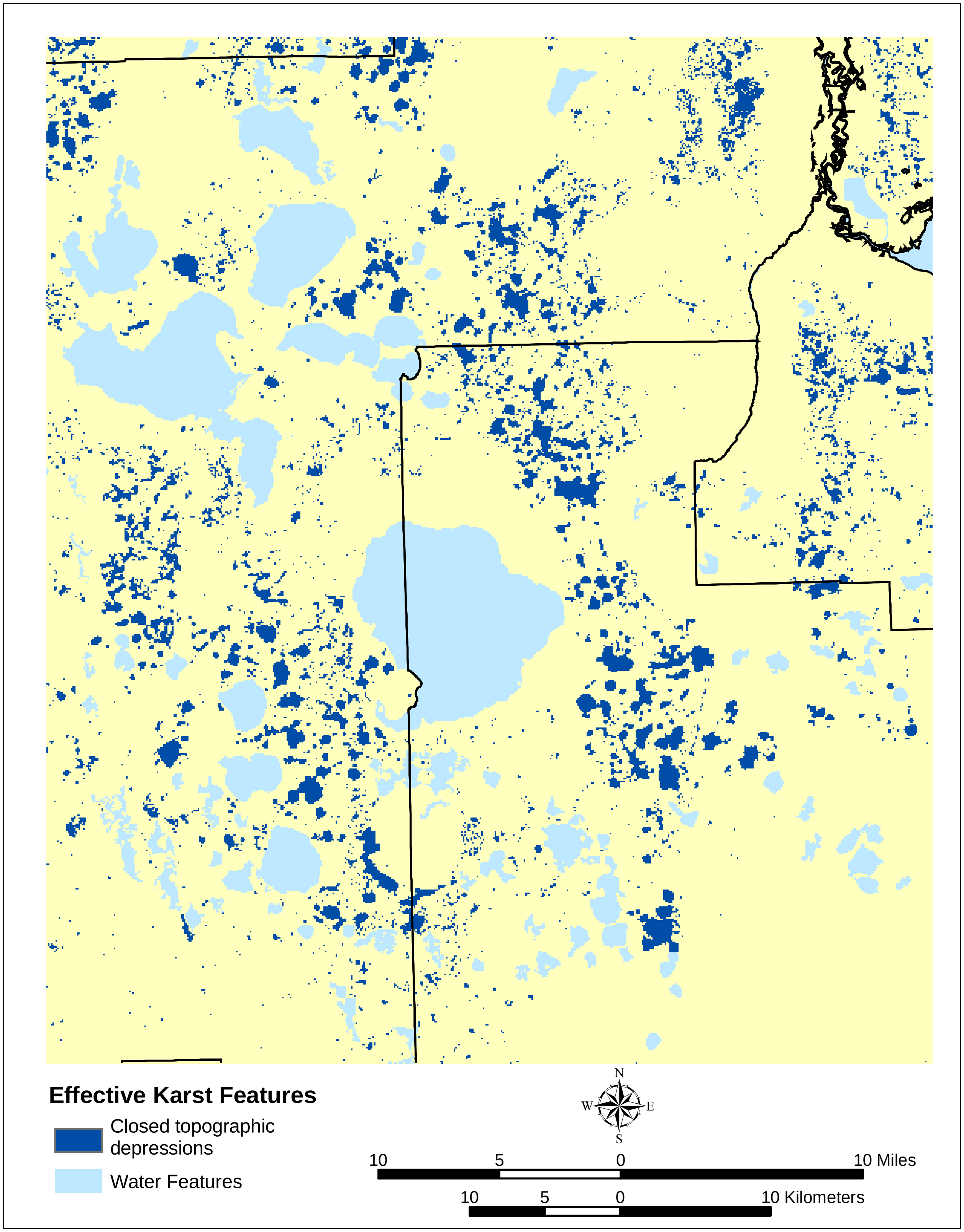

Figure 6. Closed topographic depressions underlain by less than 100 feet (30.5 m) of FAS overburden.

Areas nearer to an effective karst feature are normally associated with higher aquifer vulnerability due to the increased chance of overland flow and the infiltration into the depression. Therefore, buffer zones of 3,600-m (11,811 ft) divided into 30-m (98.4 ft) intervals were generated around each karst feature, and weights were calculated for the effective karst feature evidential theme using the cumulative ascending method. The highest contrast of any class was calculated at a distance of 60 m (196.9 ft) from a depression creating a binary generalized theme for input into the WAVA model. In other words, the analysis indicated that this threshold of proximity to karst features maximized the spatial association between the map pattern and the training point pattern. The generalized theme is displayed in Figure 7.

Intermediate Aquifer System Thickness

For the purposes of the WAVA project, "Intermediate Aquifer System," or IAS, refers to the statewide hydrostratigraphic unit (formations that can be grouped into aquifer systems or confining units) that provides variable confinement to the FAS. The initial step in the identification of the IAS was to create a database of well core and cuttings data. The well database was then used to create a hydrostratigraphic surface for the top of the IAS and the top of the FAS (which coincides with the base of the IAS). Following creation of these surface models, it was necessary to resolve the IAS with land-surface elevation in localized areas where the IAS surface interpolation extended above land surface. After the hydrostratigraphic surfaces were developed and refined, calculation of a thickness map was completed by subtraction of the IAS surface from the FAS surface. The final output yielded a continuous thickness map of the IAS (Figure 8).

Areas underlain by thinner IAS sediments are normally associated with higher aquifer vulnerability. Weights were therefore calculated for the IAS evidential theme using the cumulative ascending method. The highest contrast of any class was calculated at a thickness of 48 feet (14.6 m) creating a binary generalized theme for input into the WAVA model (Figure 9). In other words, the analysis indicated that this threshold of IAS thickness maximized the spatial association between the map pattern and the training point pattern.

The best available data indicated that the area including Wekiwa Spring is underlain by 50 feet (15.2 m) of more of IAS thickness. This data included well core and cuttings data and geophysical (gamma) well logs from FGS lithologic database and FGS Special Publication No. 50 (Davis et al., 2001). As a result, for this analysis, the spring is located outside the area which is strongly correlated with aquifer vulnerability as defined by the training point data (i.e., greater than 48 feet or 14.6 m thick). In fact, the exposed rocks within the spring pool are the Coosawhatchie Formation of the Hawthorn Group which is part of the IAS (Harley Means, personal communication, 2004).

Hydraulic Head Difference between SAS and FAS

Hydraulic head difference was calculated by subtracting the FAS 1993–1994 potentiometric surface (Sepulveda, 2002) from the water-table surface (Arthur et al., 2005, in preparation). Although more recent potentiometric surfaces for the FAS have been published, the data control in Sepulveda's (2002) map was denser and the potentiometric surface was more highly resolved. Areas with a positive hydraulic head difference value indicated that the FAS is receiving recharge, whereas areas with a negative value indicated the FAS is discharging to the overlying aquifer system (Figure 10).

Areas of greater hydraulic head difference between the water table and FAS indicate greater potential for downward recharge to the FAS, which is normally associated with higher aquifer vulnerability. Weights were therefore calculated for the hydraulic head difference evidential theme using the cumulative descending method. The highest contrast of any class was calculated at a hydraulic head difference value

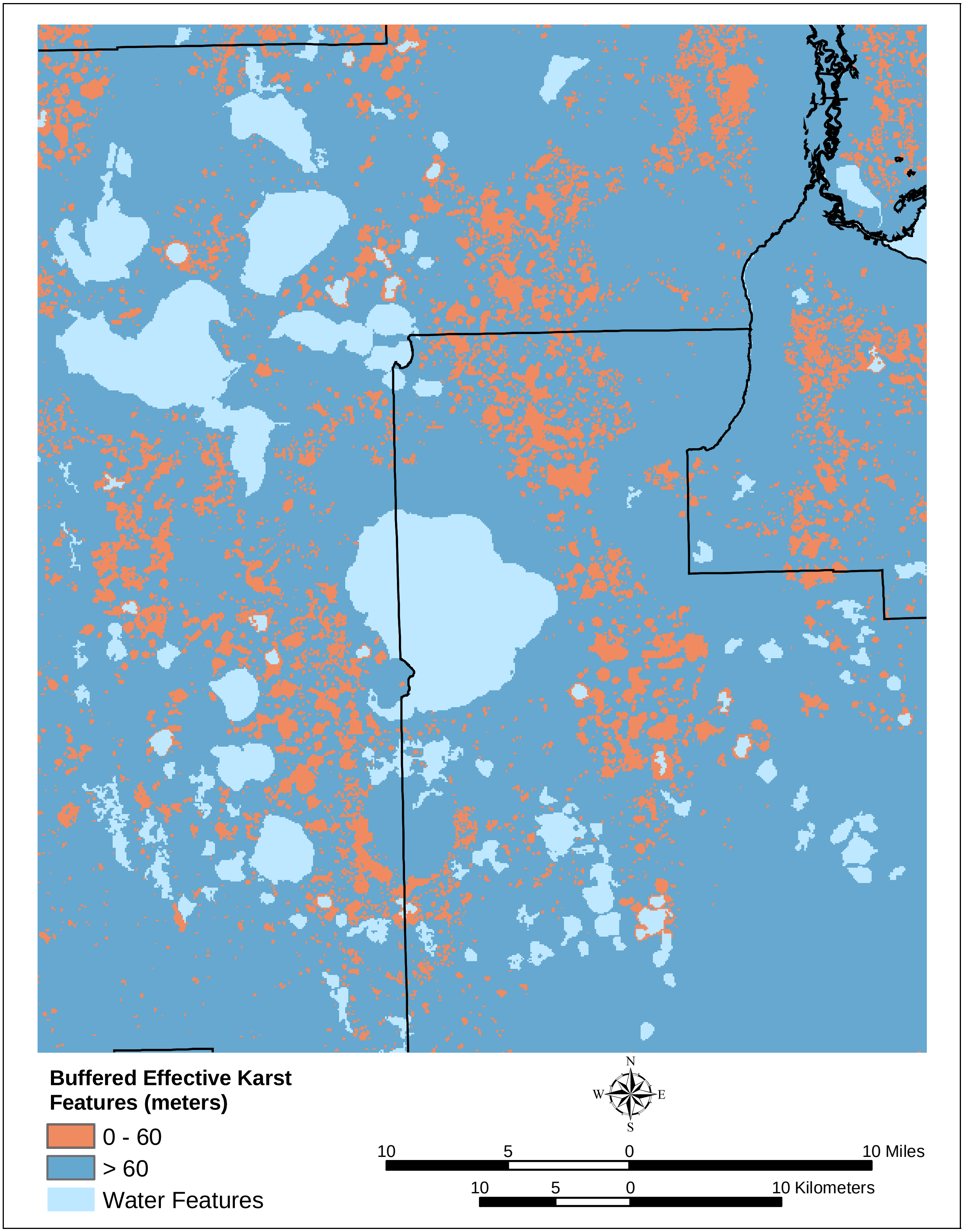

Figure 7. Map showing binary generalization of the buffered effective karst features evidential theme based on a contrast buffer distance of 60 m (196.9 ft).

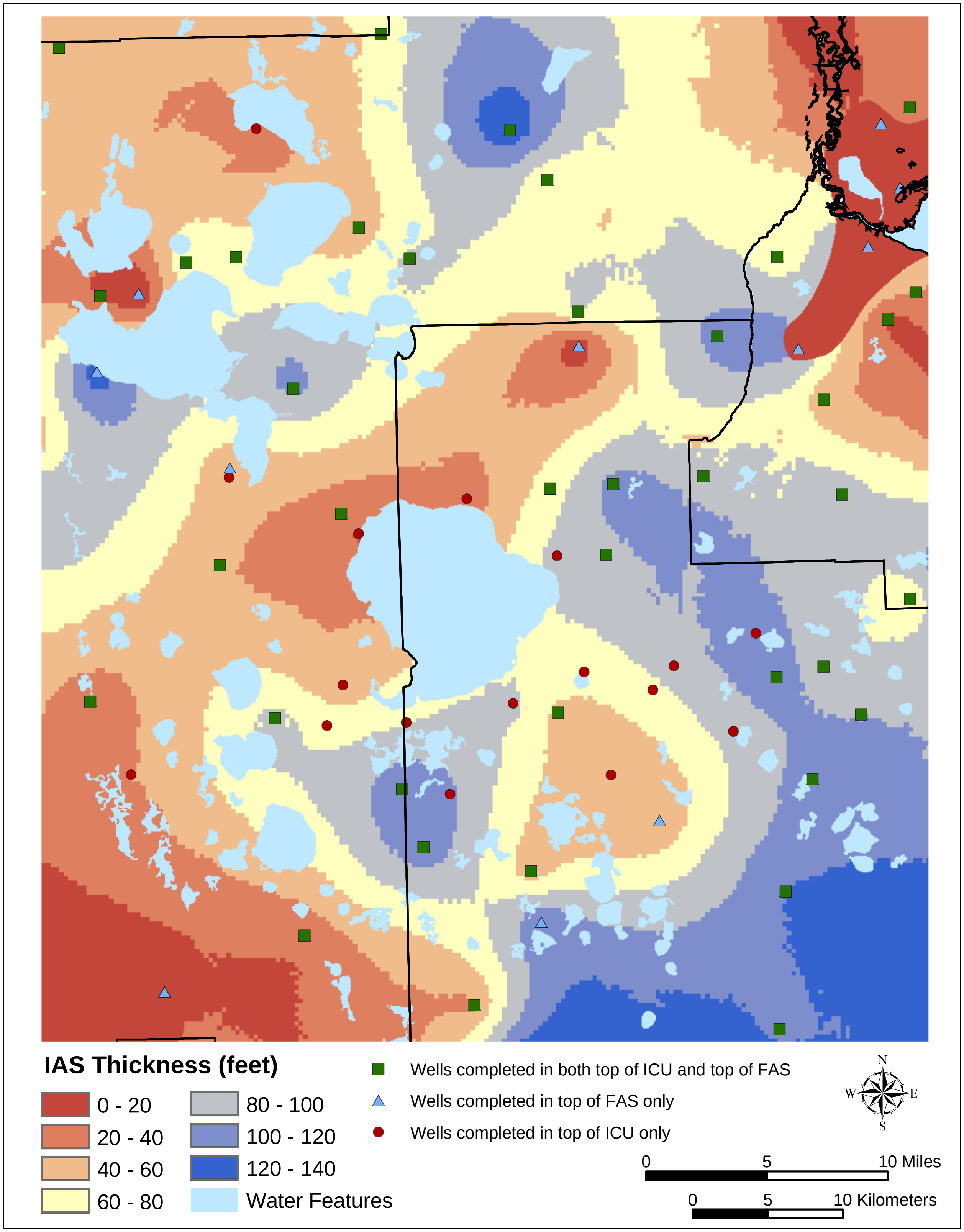

Figure 8. Thickness of IAS in feet and the distribution of wells used to determine IAS thickness. Data from FGS lithologic database combined with geophysical well logs (Davis et al., 2001) were utilized to develop this coverage.

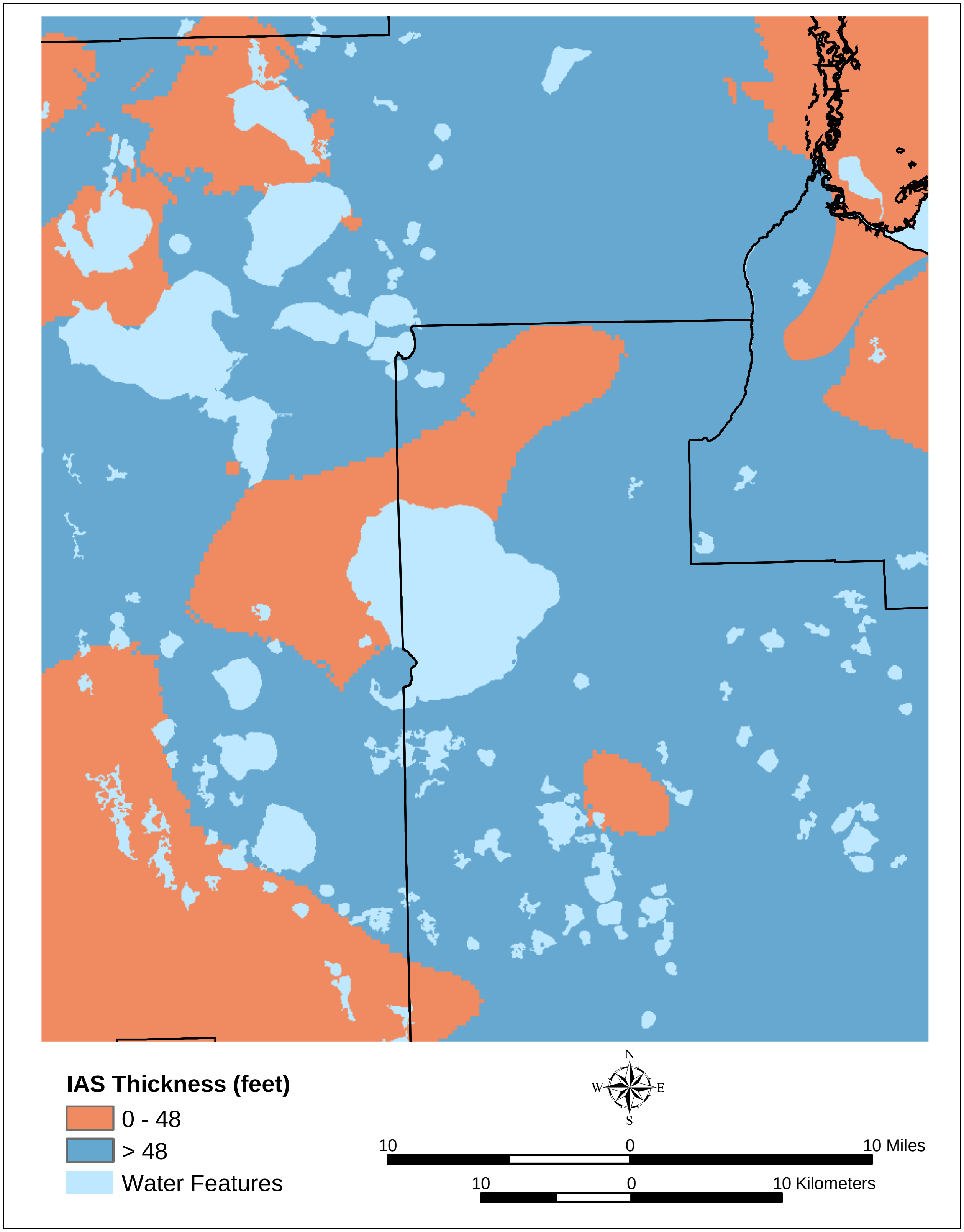

Figure 9. Map showing binary generalization of the IAS thickness evidential theme based on a contrast value of 48 ft (14.6 m).

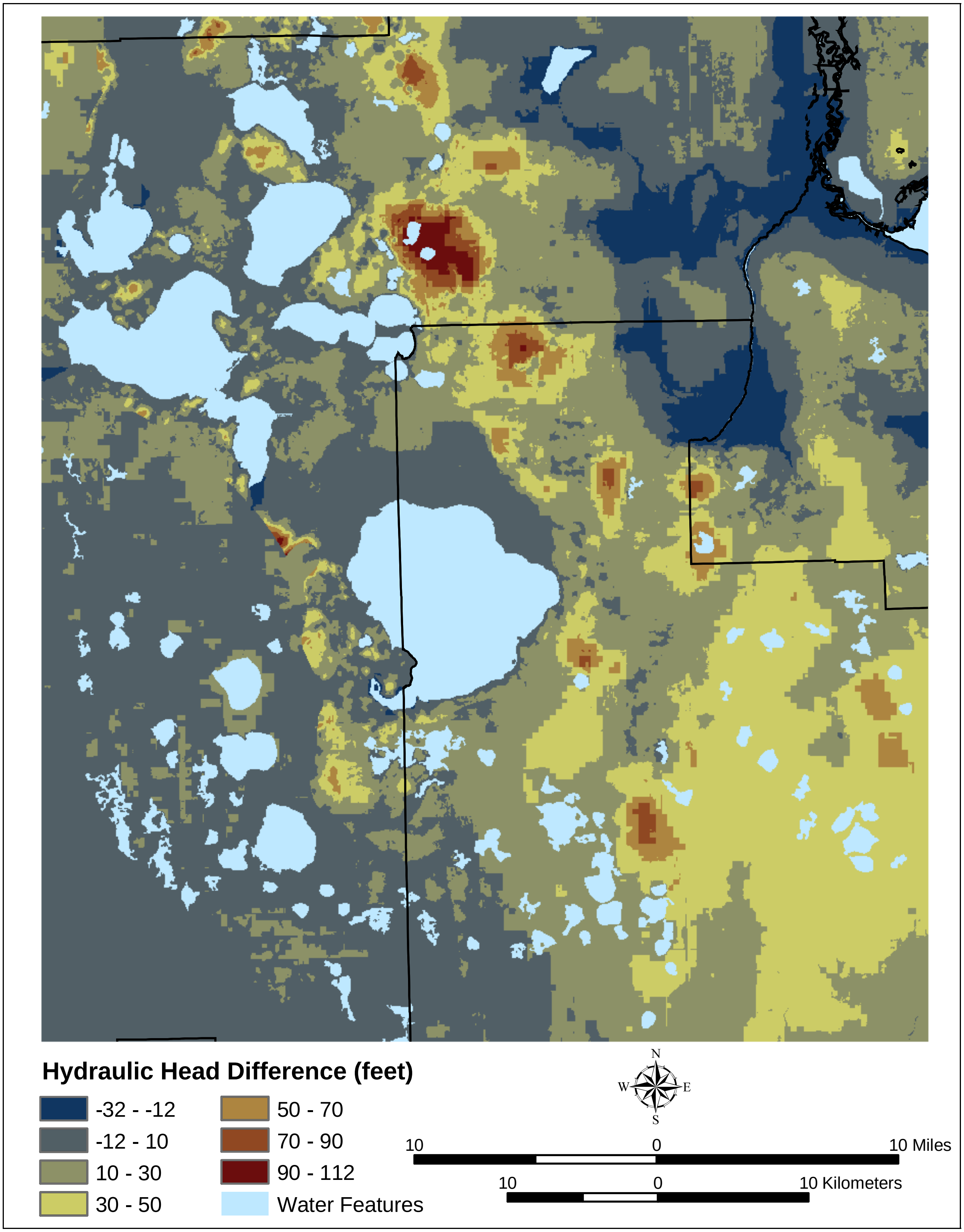

Figure 10. Hydraulic head difference between the water-table surface (Arthur et al., 2005, in preparation) and the FAS potentiometric surface (Sepulveda, 2002).

of 17 ft (5.2 m) creating a binary generalized theme for input into the WAVA model. In other words, the analysis indicated that this threshold of hydraulic head difference maximized the spatial association between the map pattern and the training point pattern. Figure 11 displays the binary generalization used for the hydraulic head difference evidential theme.

Response Theme

The WAVA response theme is an output map, calculated using WofE, showing the probability (posterior probability) that a unit area is vulnerable to contamination from land surface based on the evidence provided. A response theme is portrayed as relative vulnerability and is classified based on the inflections in a chart in which cumulative study area is plotted against posterior probability. The chart may allow for the selection of several classes of posterior probability. The more vulnerable areas correspond with higher posterior probabilities, while the less vulnerable areas are associated with lower posterior probabilities. In essence, a higher posterior probability indicates that an area is more likely to contain a training point, or more likely to be contaminated, and therefore more vulnerable to contamination from land surface.

As identified in the *Introduction*, the WAVA model was based on the modeling technique used in the FAVA project. The FAVA project involved the comparison of the vulnerability of an aquifer system over its entire extent and consolidated vulnerability into three classes (more vulnerable, vulnerable and less vulnerable). The FAVA response theme clipped to the WAVA study area (Figure 12) shows that the majority of the study area is located in the “more vulnerable” zone. A smaller part of the study area is located in the “vulnerable” zone and *none* of the study area is located in the “less vulnerable” zone. By applying the FAVA modeling technique to the WAVA study area it was possible to identify zones of vulnerability within the study area since the predicted vulnerability of the smaller area will not be masked by the predicted vulnerability of the FAS statewide extent. Note, however, that by applying the FAVA model as a baseline for the WAVA study area, all protection zones should be considered “vulnerable” to “more vulnerable” in context of the statewide relative aquifer vulnerability.

Using the four evidential themes discussed above, a response theme was generated showing the posterior probability that a unit area contained a training point based on the evidential themes provided. A conceptual model showing the association between the training points and the evidential themes is shown in Figure 13. The posterior probabilities of the response theme ranged from 0.0007261 to 0.1042028 across the model domain. Plotting posterior probability against cumulative area as a percentage (Figure 14) allowed the delineation of class breaks for display of protection zones in the final response theme (Figure 15). The breaks for these protection zones were selected where a notable stepwise increase in posterior probability relative to the cumulative study area occurred. The first break, which delineated the *tertiary protection zone* from the *secondary protection zone*, occurred at a posterior probability value of 0.000099. The *tertiary protection zone* represents approximately 25.7% of the study area. The second break delineating the *secondary protection zone* from the *primary protection zone* occurred at the next significant stepwise increase in posterior probability at a value of 0.0017, which corresponded with the prior probability. The *secondary protection zone* represents approximately 55.8% of the study area. The remainder of the study area fell into the *primary protection zone* and represents approximately 18.4% of the study area. This *primary protection zone* contains the greatest probability of containing a training point, and is therefore considered the most vulnerable to contamination from land surface.

For the purposes of this project, three classes were delineated as protection zones; however, based on the methodology used to select these classes, a third probability break could be delineated, representing approximately 3% of the study area. This area would represent the most vulnerable part of the primary zone and would be located principally in the “more vulnerable” area defined in the FAVA study.

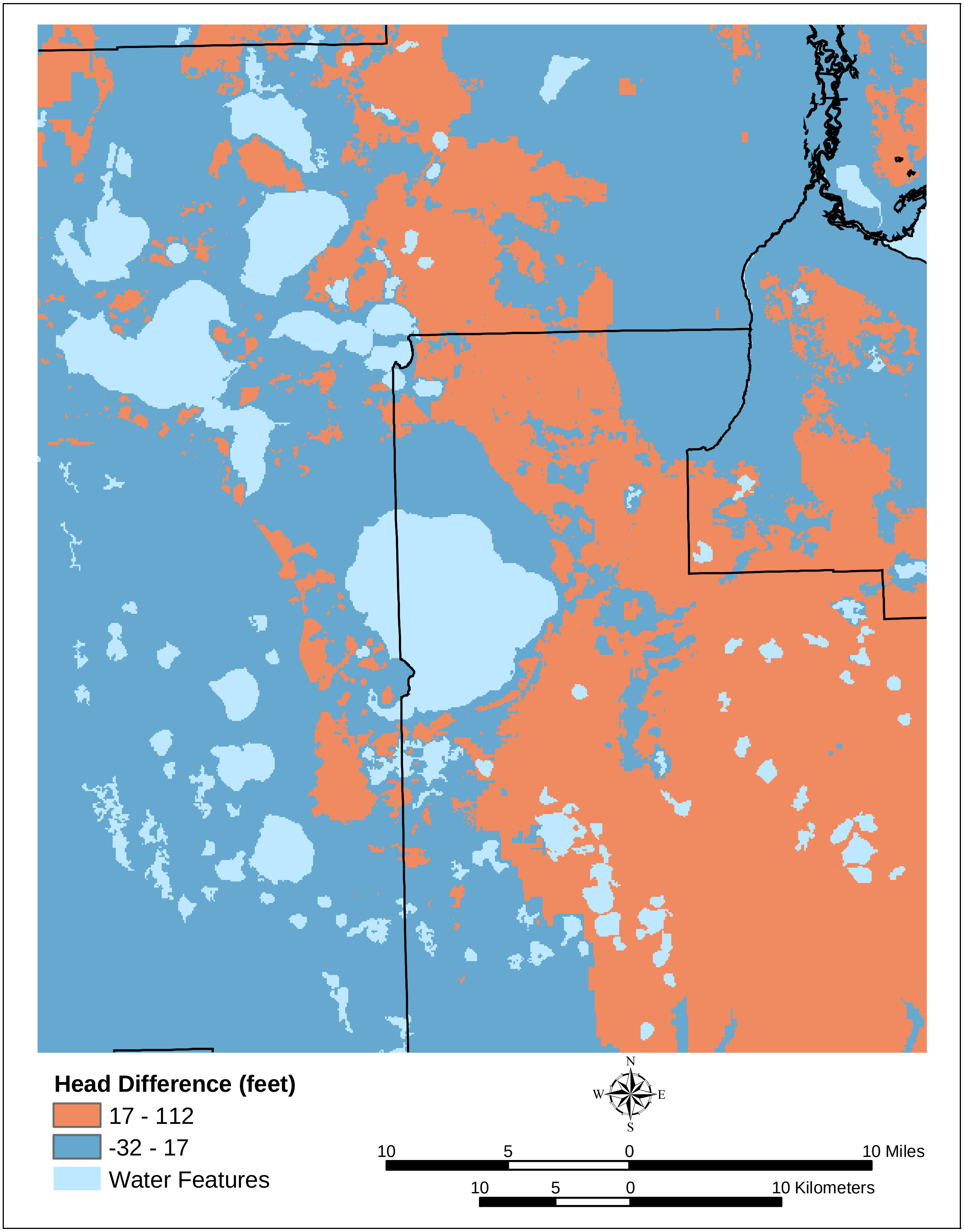

Figure 11. Map showing binary generalization of the hydraulic head difference evidential theme based on a contrast value of 17 ft (5.2 m).

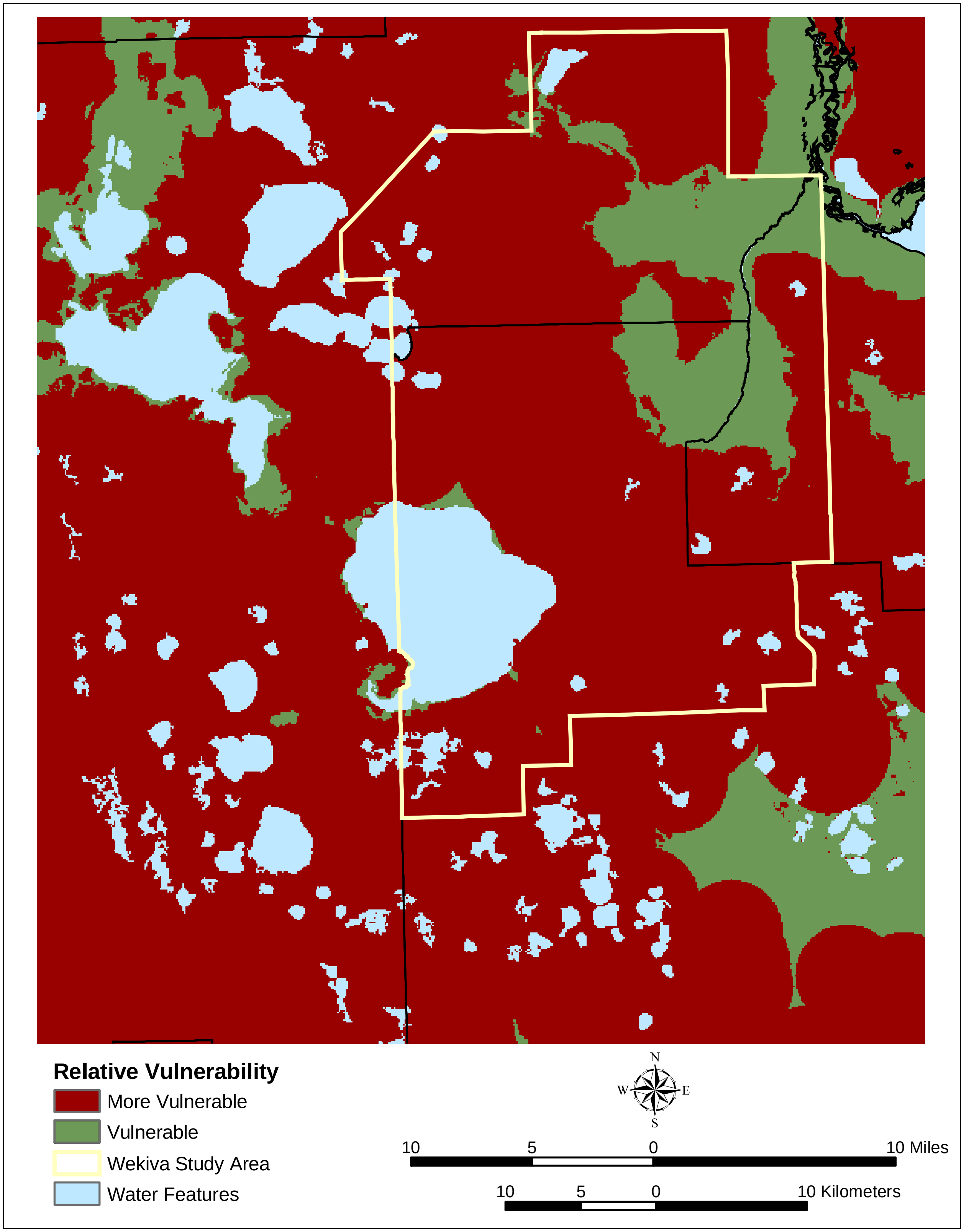

Figure 12. Relative vulnerability of the FAS FAVA model (Arthur et al., 2005, in preparation) showing zones of vulnerability based on the extent of the Floridan Aquifer System.

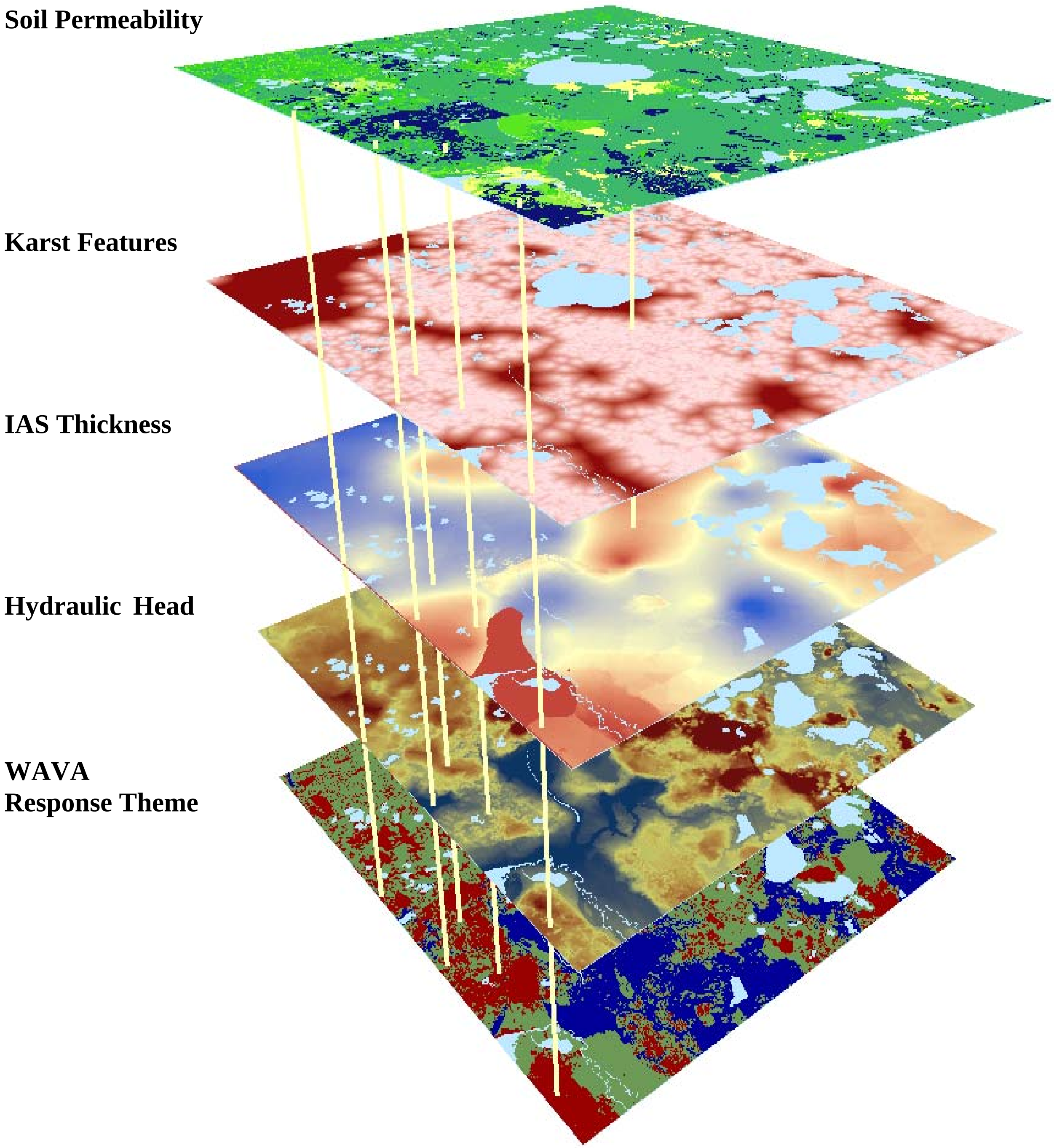

Figure 13. WofE conceptual model of the FAS. The top four layers are evidential themes and the bottom layer is the response theme divided into zones of protection. Yellow lines represent training points (wells) projected throughout the layers.

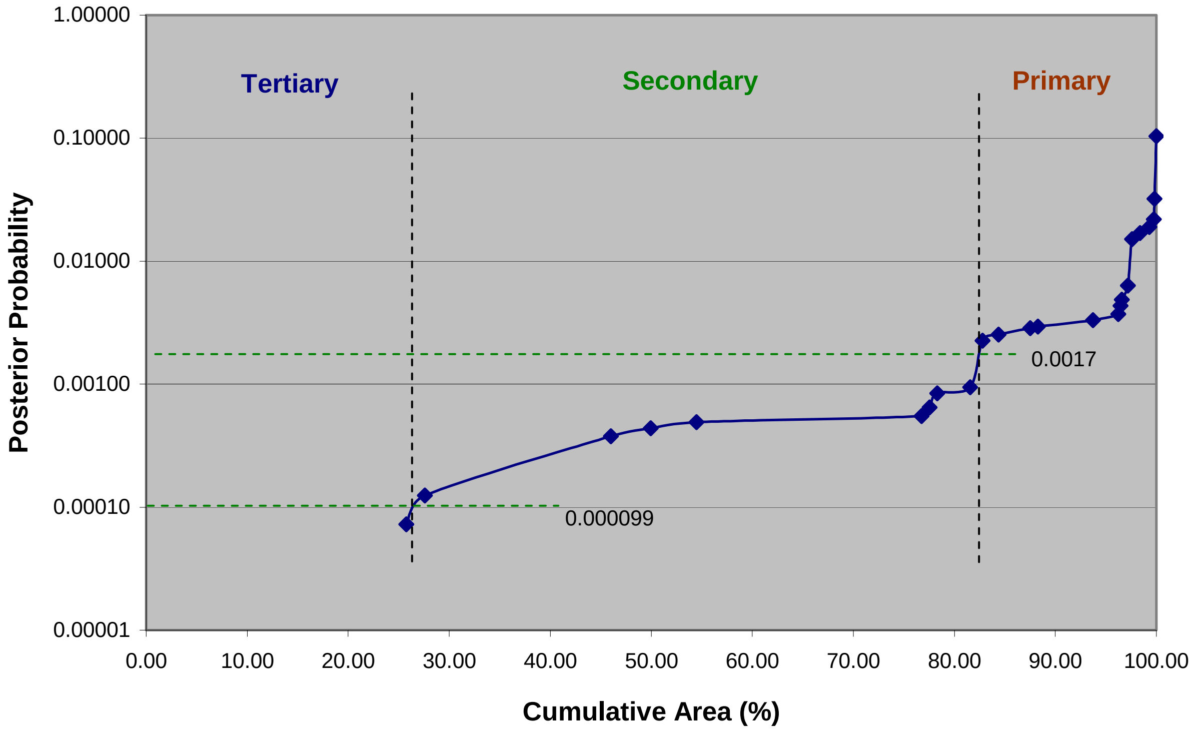

Figure 14. Class breaks, represented by green dashed lines, were placed where both a significant increase in probability and area were observed. These breaks represent the boundaries between protection zones.

The response theme indicated that the areas of highest vulnerability (high probabilities) tended to be associated with IAS sediments that are thin to absent, dense karst-feature distribution, high positive hydraulic head difference, and high soil permeability. Conversely, areas of lowest vulnerability (low probabilities) tended to be determined by thick IAS sediments, sparse karst-feature distribution, hydraulic head difference less than 17 feet (5.2 m) and low soil permeability values.

In addition to the large surface-water bodies that contain limited data, the study area is comprised of a multitude of other surface-water features (Figure 16). These features can represent areas of discharge and may be predicted with low posterior probability values. These discharging surface waters are not part of the aquifer, although they can originate from it. The WAVA project was designed to focus on the ability for a contaminant to travel through soils, overburden, karst features, etc. to enter into the aquifer system. As a result, it is very important that the WAVA model never be applied to assess contamination of surface waters or discharge areas.

As noted earlier, an assumption is made when using WofE that there is conditional independence between the layers used as predictors. Conditional independence is violated when the presence of one evidential theme influences the probability of another evidential theme. The validity of posterior probability values depends upon the degree of conditional independence calculated for the model. Evidential themes are considered independent of each other if the conditional independence value is around 1.00. For the WAVA project, appropriate conditional independence values fell within the range of 1.00 ± 0.15 (Gary Raines, personal communication, 2003). Conditional independence was calculated at 1.15 indicating minimal dependence between evidential themes.

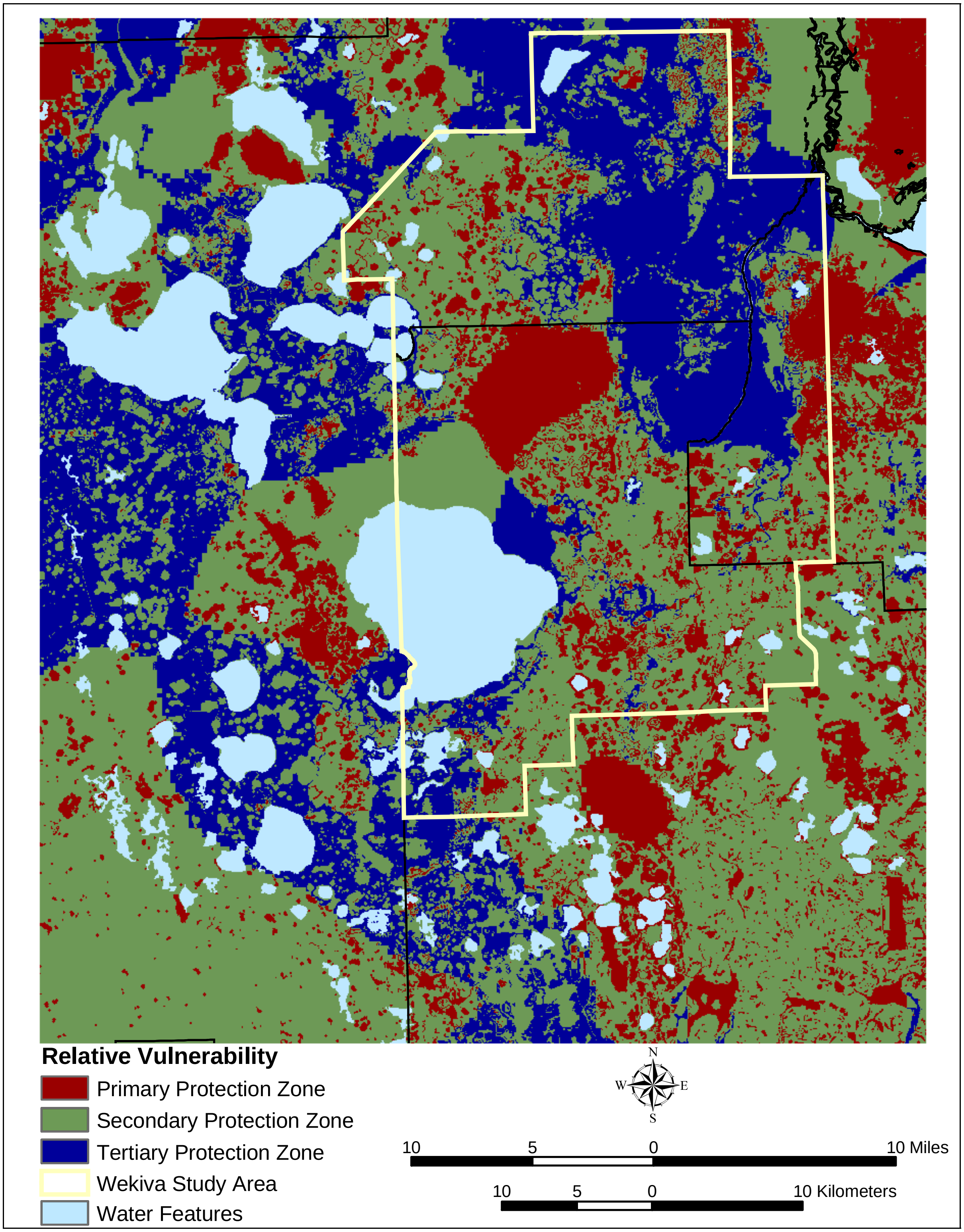

Figure 15. Relative vulnerability of the WAVA model showing primary, secondary and tertiary protection zones.

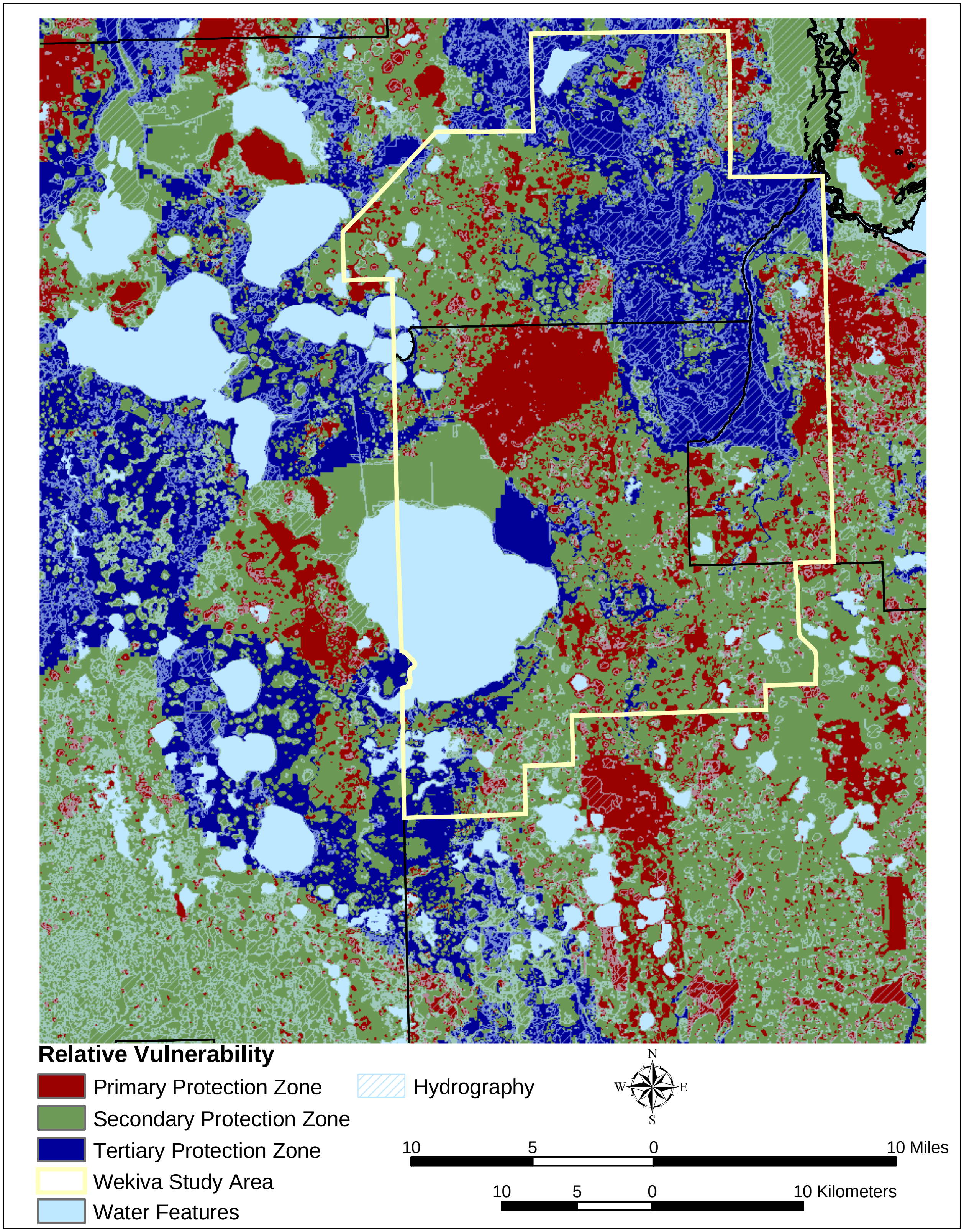

Figure 16. Relative vulnerability of the WAVA model showing primary, secondary and tertiary protection zones. Hydrography represents seasonal and permanent wetlands, ponds and small lakes.

Weights calculated for the evidential themes used in the WAVA model are included in Table 2. This table displays the evidential themes used, weights calculated for those evidential themes, as well as the contrast and confidence of the evidential theme. A positive weight indicates areas where training points were likely to occur, while a negative weight indicates areas where training points were not likely to occur. The contrast column is a combination of the highest and lowest weights (positive weight – negative weight) and is a measure of how well the generalized evidential themes predict training points. A positive contrast that is significant, based on its confidence, suggests that a generalized evidential theme is a useful predictor. The confidence of the evidential theme is the contrast divided by its standard deviation and provides a useful measure of significance of the contrast because of the uncertainties of the weights and areas of missing data (Raines, 1999).

Table 2. WofE final output table listing weights calculated for each evidential theme and their associated contrast and confidence values of the evidential themes.

Evidential Theme	W1	W2	Contrast	Confidence
Hydraulic Head Difference	0.7412	-1.2857	2.0269	1.8494
Buffered Effective Karst	1.3546	-0.5550	1.9096	2.3343
Soil Permeability	1.2564	-0.5396	1.7960	2.1955
IAS Thickness	0.8739	-0.7726	1.6465	1.8994

The hydraulic head difference evidential theme had a stronger association with the training points (i.e., highest contrast) than the other evidential themes and was therefore the primary determinant in predicting areas of vulnerability in this model. Three (buffered effective karst, confining unit thickness and soil permeability) of the evidential themes indicated where training points were more likely to occur because the positive weights (W1) were stronger (had a greater absolute value) than the negative weights (W2). Confidence values, for each evidential theme, exceeded the target value of 1.282, which corresponds to a level of significance of approximately 90% (Table 1). In fact, all confidence values for the evidential themes exceeded the 95% level of significance (1.645). For the dissolved oxygen response theme all of the training points occurred within areas of higher vulnerability, thus supporting the results of the model output.

The soil permeability evidential theme contains large areas of missing data particularly along the Orange and Seminole county border. To compensate for this in the statewide FAVA model, a nearest neighbor GIS function was employed, which was used to apply spatial statistics (Euclidean distance functions) to predict soil permeability data values for these areas. Although this worked well on a statewide scale, use of this process on a smaller area skewed the vulnerability where there were large areas of missing data. In the WAVA model for example, use of this nearest neighbor function caused these areas along the Orange and Seminole county border to be included in the primary protection zone of the response theme. However, if the nearest neighbor function was not used, the calculated response theme included these same areas in the secondary protection zone (Figure 17). Based on the uncertainty due to missing data, the nearest neighbor function was not used in the WAVA model. The response theme values in these areas were therefore based solely on IAS thickness, buffered effective karst features, and hydraulic head difference. As a result, the relative vulnerability in these areas of missing data has a higher degree of uncertainty when compared to other areas of the model, and as new data becomes available, the vulnerability of these areas can be re-assessed.

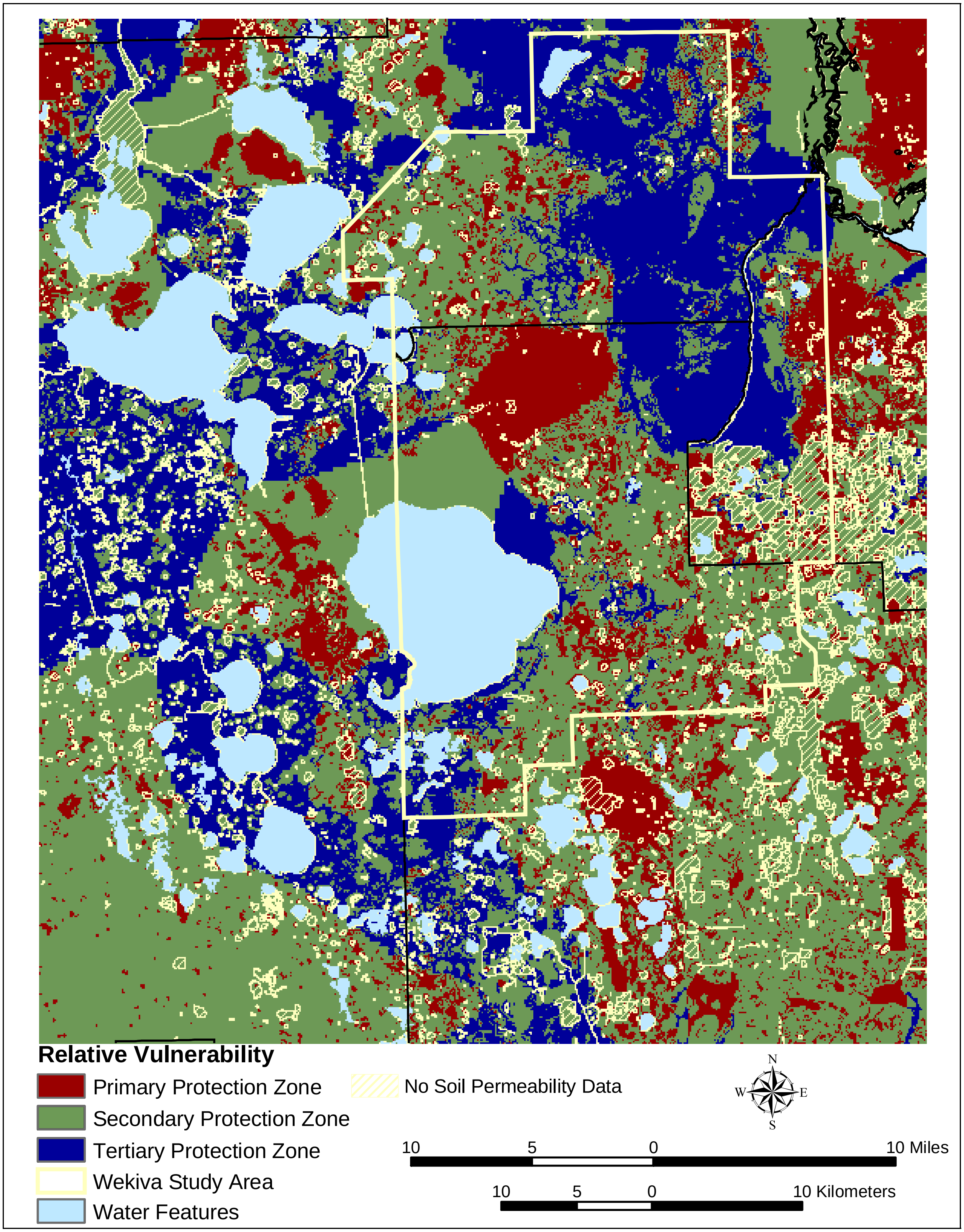

Figure 17. Relative vulnerability of the WAVA model showing areas where no soil permeability data exists. The relative vulnerability predicted in these areas could change (i.e., higher or lower) as new data becomes available.

Confidence Map

There are two types of confidence used on the WofE model. Confidence of the evidential theme, as reported in Table 2 and discussed above in *Generalization of Evidential Themes*, equals the contrast divided by the standard deviation (a student T test) for a given evidential theme (Raines, 1999). The second type of confidence can be calculated for each response theme by dividing the theme's posterior probability by its total uncertainty (standard deviation). A confidence map can be generated based on these calculations. The confidence map for the WAVA response theme is displayed in Figure 18.

Areas with a high posterior probability tend to have higher confidence values and therefore have a higher level of certainty with respect to predicting aquifer vulnerability. A small population of training points along with missing data raises the total uncertainty for the response theme which in turn lowers the confidence. A confidence (of posterior probability) map containing all evidential themes (Figure 18) can therefore contain a lower level of significance than those calculated for each separate evidential theme.

Validation and Sensitivity Analysis

For other aquifer vulnerability assessments completed by the FGS (Arthur et al., 2005, in preparation; Baker et al., 2005) models were validated by creating a training point theme based on another parameter that reflected vulnerability yet was independent of the original training point set. As mentioned above, the WAVA response theme, based on calculated weights, was a better predictor of where training points were more likely to occur because the positive weights (W1) were stronger than the negative weights (W2). In contrast, results from using a dissolved nitrogen training point set as validation for the WAVA model, indicated where training points were less likely to occur because the negative weights (W2) were stronger than the positive weights (W1). This produced a different kind of model, and, as a result, was not comparable to the original WAVA response theme. The following methods, however, did prove effective in completing reliable validation and sensitivity analyses.

Dissolved Oxygen Data vs. Posterior Probability

Posterior probability values were compared with the dissolved oxygen dataset from which the training point theme was extracted. Dissolved oxygen median concentration values were plotted versus posterior probability values (Figure 19). As expected a positive trend was observed between posterior probability and dissolved oxygen values and a significant Pearson correlation coefficient of 0.689 was calculated.

Random 75% Subset of Training Points

If the WAVA evidential themes and training points were robust (i.e., not sensitive to subtle changes in the training data set), one would expect the response theme patterns for the full training data set and a subset to be similar. For this sensitivity test, a training point theme consisting of a random subset of 75% of the original training points was generated and the model was re-executed. Response themes generated for the random subset of points were broken into the three protection zones using the methodology described in the *Response Theme* section. The subset response theme was used to evaluate the degree of correlation between the WAVA response theme and the response theme developed using the 75% subset of training points.

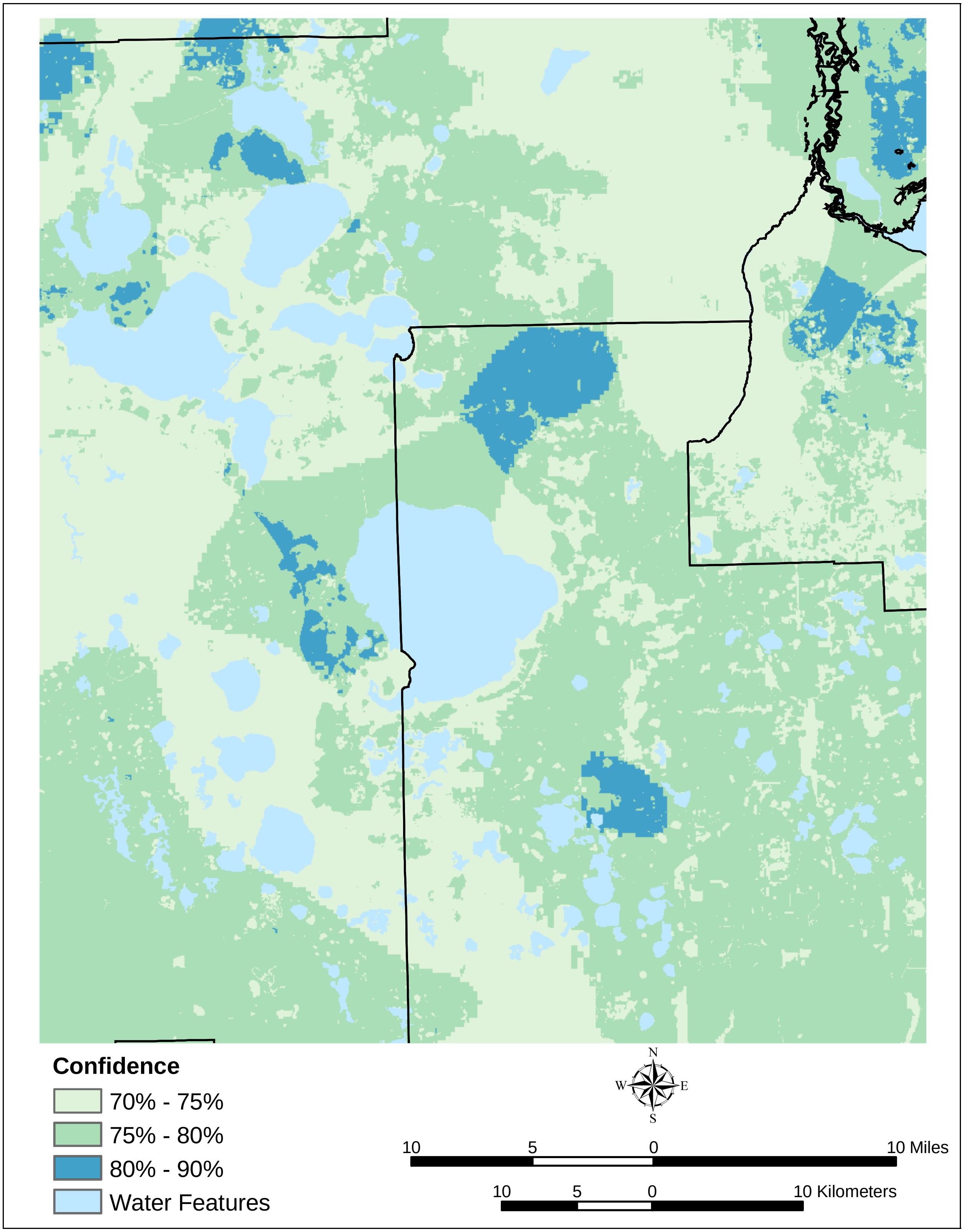

Figure 18. Distribution of confidence values calculated for WAVA response theme.

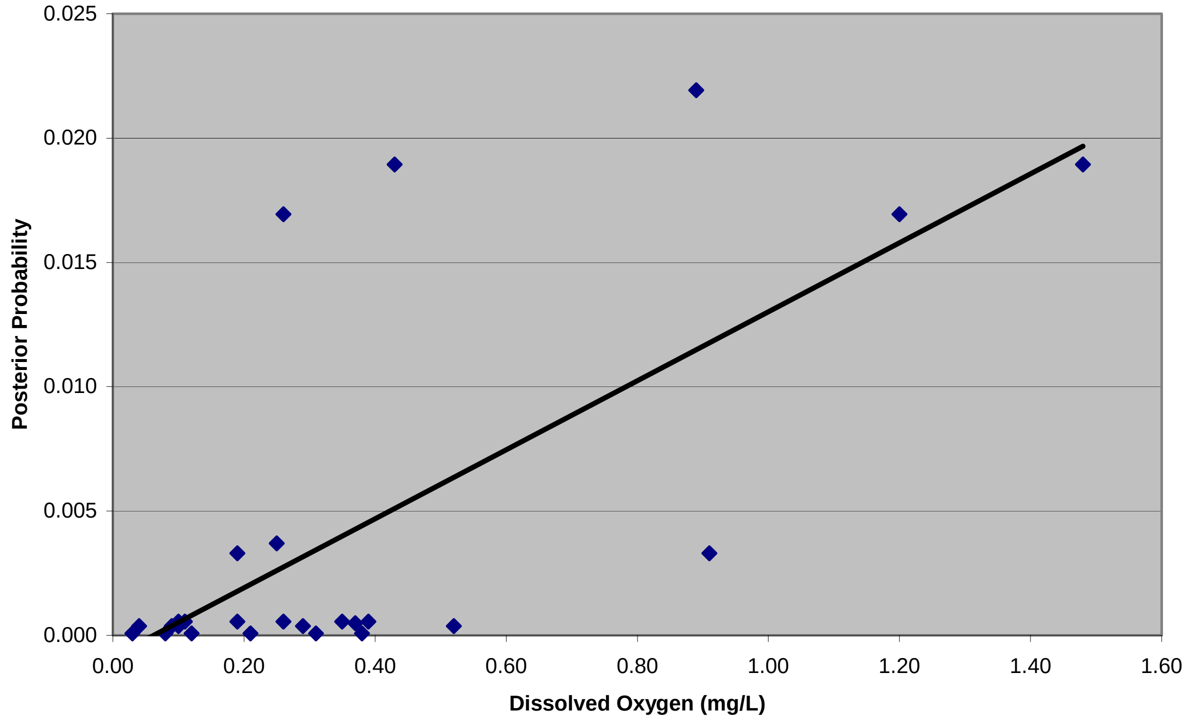

Figure 19. Posterior probability vs. dissolved oxygen median concentration values

The kappa coefficient was used to measure the amount of spatial agreement between response themes while taking into account agreement that could have occurred by chance. Additionally, conditional kappa values were calculated to determine the amount of agreement between each protection zone class of the two response themes being compared. A cross-tabulation matrix was used to classify the response themes by area (in square meters) and aided in the calculation of observed and expected proportions (i.e., agreement). Values along the diagonal in this table reflect the amount of agreement between response themes and the remaining table cells reflect where the response themes were mismatched (Table 3).

Table 3. Cross-tabulation matrix of the area in square meters per class of the WAVA response theme and 75% subset response theme. Values along the diagonal reflect the amount of agreement.

Class		75% Subset Response Theme			
		Primary	Secondary	Tertiary	Total
WAVA Response Theme	Primary	608,123,700	43,521,300	0	651,645,000
	Secondary	0	1,974,464,100	0	1,974,464,100
	Tertiary	0	0	910,533,600	910,533,600
	Total	608,123,700	2,017,985,400	910,533,600	3,536,642,700

Kappa coefficient results can range between -1 (perfect disagreement) and 1 (perfect agreement). A value of zero indicates that the agreement was no better than that expected due to chance (Bonham-Carter 1994). Kappa coefficients calculated in the WAVA project were all positive values. Positive kappa coefficients can be interpreted using Table 4.

Table 4. Kappa coefficient values and their associated interpretation (Landis and Koch, 1977).

Interpretation of kappa values for positive agreement	
Kappa	Interpretation
< 0	No agreement
0.0 – 0.19	Poor agreement
0.20 – 0.39	Fair agreement
0.40 – 0.59	Moderate agreement
0.60 – 0.79	Substantial agreement
0.80 – 1.00	Almost perfect agreement

A subset of the total dissolved oxygen training point theme was generated using a random selection process. This random subset included 75% of the original wells for a total of 5 training points and yielded a prior probability of 0.0014. Weights were then recalculated for each evidential theme, class breaks were selected, and a response theme was generated (Figure 20). The pattern of posterior probabilities was nearly identical to the original total dissolved oxygen response theme.

The kappa coefficient was used to measure the amount of spatial agreement between the random subset response theme and the WAVA response theme. The kappa coefficient between the response themes was calculated at 0.979. Additionally, conditional kappa values were calculated to determine the amount of agreement between each protection zone class of the two response themes being compared. Table 5 displays the conditional kappa coefficient between each protection zone class of the two response themes. The overall agreement between the two response themes was almost perfect as was the agreement between the protection zone classes.

Table 5. Conditional kappa coefficient values between the random 75% subset response theme and the WAVA response theme. Conditional kappa coefficient values are reported between each vulnerability class.

Agreement	Conditional Kappa (Kf) values
Primary Protection Zone	0.919
Secondary Protection Zone	~1.000
Tertiary Protection Zone	~ 1.000

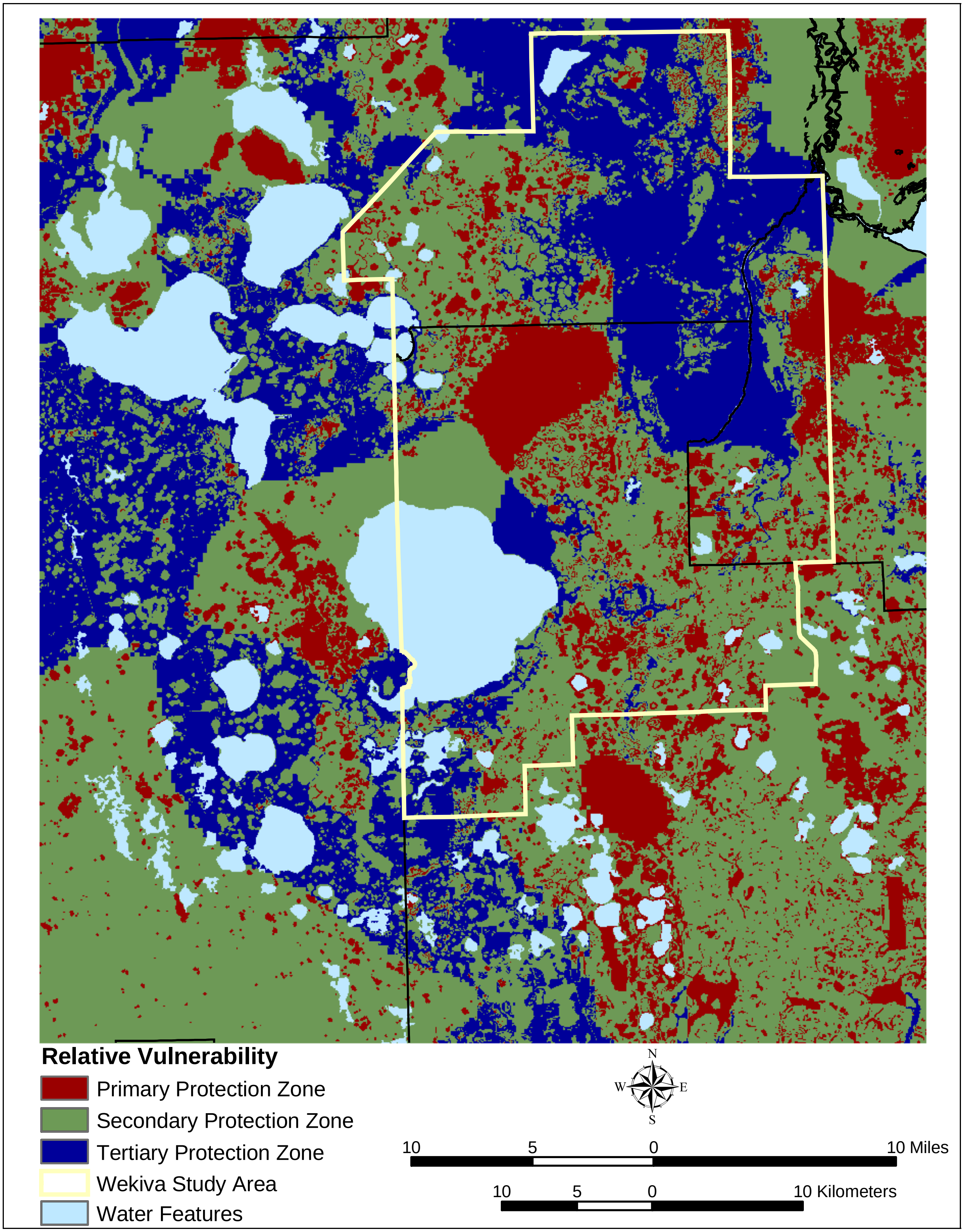

Figure 20. Relative vulnerability of the WAVA model generated using a random 75% subset of the original dissolved nitrogen training point theme.

WEIGHTS OF EVIDENCE GLOSSARY

Binary – Refers to the generalization or simplification of evidential themes or data layers. Binary layers are reclassified from the original dataset into presence/absence type themes or into two classes.

Conditional Independence – Occurs when an evidential theme does not affect the probability of another evidential theme. Evidential themes are considered independent of each other if the conditional independence value calculated is within the range 1.00 ± 0.15 (Raines, personal communication, 2003). Values that significantly deviate from this range can over inflate the posterior probabilities resulting in unreliable response themes.

Confidence of evidential theme – Contrast divided by its estimated standard deviation; provides a useful measure of significance of the contrast.

Confidence of Posterior Probability – A measure based on the ratio of posterior probability to its estimated standard deviation.

Contrast – W+ minus W- (see weights), which is an overall measure of the spatial association (correlation) of an evidential theme with the training points.

Cumulative Ascending – Calculates the cumulative weights from the first class to the last class while increasing the area. Areas nearest a training point have a stronger association, and those farthest away have a weaker association. This method is applicable for themes where the training points are mainly associated with the lower values of the evidential theme (e.g., higher vulnerability correlates with lower confinement thickness).

Cumulative Descending – Calculates the cumulative weights from the last class to the first class while increasing the area (opposite of cumulative ascending). This method is applicable for themes where the training points are mainly associated with the higher values of the evidential theme (e.g., higher vulnerability correlates with higher soil permeability).

Evidential Theme – A set of continuous spatial data that is associated with the location and distribution of known occurrences (i.e., training points); these map data layers are used as predictors of vulnerability.

Model – The characteristics of a set of training points, and the relationships of the training points to a collection of evidential themes.

Posterior Probability – The probability that a unit cell contains a training point after consideration of the evidential themes. This measurement changes from location to location depending on the values of the evidence.

Prior Probability – The probability that a unit cell contains a training point before considering the evidential themes. It is a constant value over the study area equal to the training point density (total number of training points divided by total study area in unit cells).

Response Theme – An output map that displays the probability that a unit area would contain a training point, estimated by the combined weights of the evidential themes. The output is displayed in classes of relative aquifer vulnerability or favorability to contamination (i.e., this area is more vulnerable than that area) or favorability. The response theme is the relative vulnerability map.

Spatial Data – Information about the location and shape of, and relationships among, geographic features, usually stored as coordinates and topology.

Study Area – A grid theme that acts as a mask to define the area where the model is developed and applied. It may be irregular in outline and may contain interior holes (e.g., lakes and no data areas).

Training Points – A set of locations (points) reflecting a parameter used to calculate weights for each evidential theme, one weight per class, using the overlap relationships between points and the various classes. In an aquifer vulnerability assessment, training points are wells with one or more water quality parameters indicative of relatively higher recharge which is an estimate of relative vulnerability.

Weights – A measure of an evidential-theme class. A weight is calculated for each theme class. For binary themes, these are often labeled as W+ and W-. For multiclass themes, each class can also be described by a W+ and W- pair, assuming presence/absence of this class versus all other classes. Positive weights indicate that more points occur on the class than due to chance, and the inverse for negative weights. The weight for missing data is zero. Weights are approximately equal to the proportion of training points on a theme class divided by the proportion of the study area occupied by theme class, approaching this value for an infinitely small unit cell.

REFERENCES

Arthur, J.D., Baker, A.E., Cichon, J.R., Wood, H.A.R., and Rudin, A., 2005, in preparation, Florida Aquifer Vulnerability Assessment (FAVA): Contamination potential of Florida's principal aquifer systems: Florida Geological Survey Bulletin 67.

Berndt., M.P., Oaksford, E.T., Mahon, G.L., and Schmidt., W., 1998, Groundwater, *in* Fernald, E.A., and Purdum, E. D., ed., Water Resources Atlas of Florida: Florida State University, Institute of Public Affairs, 312 p.

Cichon, J.R., Wood, H.A.R., Baker, A.E., and Arthur, J.A., 2004, Application of Geologic Mapping and Geographic Information Systems to Delineate Sensitive Karst Areas for Land-Use Decisions, American Geological Institute website, http://www.agiweb.org/environment/publications/ mapping/graphics/florida.pdf, 2004.

Davis, J., Johnson, R., Boniol, D., and Rupert, F., 2001, Guidebook to the Correlation of Geophysical Well Logs within the St. Johns River Water Management District: Special Publication No. 50, 114 p.

Florida Geographic Data Library (FGDL), GeoPlan Center, Gainesville, University of Florida, www.fgdl.org/, 2003.

Kemp, L.D., Bonham-Carter, G.F., Raines, G.L. and Looney, C.G., 2001, Arc-SDM: Arcview extension for spatial data modeling using weights of evidence, logistic regression, fuzzy logic and neural network analysis, http://ntserv.gis.nrcan.gc.ca/sdm/, 2002.

National Research Council, 1993, Ground Water Vulnerability Assessment: Predicting Relative Contamination Potential under Conditions of Uncertainty: Washington, National Academy Press, 204 p.

Raines, Gary L., 1999, Evaluation of Weights of Evidence to Predict Epithermal-Gold Deposits in the Great Basin of the Western United States: Natural Resources Research, vol. 8, no. 4, p. 257-276.

Raines, G. L., Bonham-Carter, G. F., and Kemp, L., 2000, Predictive Probabilistic Modeling Using ArcView GIS: ArcUser, v. 3, no.2, p. 45-48.

Scott, T.M., Means, G.H., Meegan, R.P., Means, R.C., Upchurch, S.B., Copeland, R.E., Jones, J., Roberts, T., and Willet, A., 2004, Springs of Florida: Florida Geological Survey Bulletin No. 66, 377 p.

Sepulveda, N., 2002, Simulation of Ground-Water Flow in the Intermediate and Floridan Aquifer Systems in Peninsular Florida: U.S. Geological Survey Water-Resource Investigation Report 02-4009, 130 p.

U.S. Census Bureau: State and County Quick Facts, 01-Feb-2005, 15:48:47 EST, ww.census.gov.

Wright, A.P., 1974, Environmental Geology and Hydrology of the Tampa Area, Florida: Florida Geological Survey Special Publication No. 19, 93 p.

Wekiva River Basin Coordinating Committee Final Report, 2004, 54 p.