USE OF REGRESSION ANALYSIS TO EVALUATE WATER LEVEL CHANGES IN THE MISSISSIPPI RIVER ALLUVIAL AQUIFER

Dean A. Pennington
Mark Stiles
Yazoo Mississippi Delta Joint Water Management District

INTRODUCTION

Agriculture and the economy of the Mississippi Delta of Mississippi have developed a critical dependence upon the limited water supplies of the region. In the last 25 years, agriculture water use has grown with the rice and catfish industries and the irrigation of cotton and soybeans. Most of the water used for agriculture is withdrawn from the Mississippi River Alluvial Aquifer. The heavy use of this resource is resulting in relatively consistent declines in some areas of the aquifer. The US Geological Survey began monitoring water levels in a network of wells located across the Delta in the early 1980s. Today water levels in the 380 well network are measured twice each year by Mississippi Department of Environmental Quality. There is now a 10 to 13 year water level history available on most of those wells.

Frequently, the data from these well measurements are used to indicate trends in water level declines across the region. Water level declines over 1 year, 10 years, or the period of record assist in understanding what is happening in the system. One method of calculating the decline over a specified period is to use the water level at the beginning of the time interval and compare it to the individual water level at the end of the interval. This is a simple and basically effective method for making this comparison but it has 3 primary limitations. These are:

1) Most of the data is ignored in the calculation. When calculating a water level trend over a long time interval with a well with a long and relatively complete history, all the data from intervening years is not used to establish the trend in water level changes.

2) Missing data for either the beginning or ending dates makes it impossible to calculate the change and eliminates that well from regional trend analysis.

3) There is considerable noise in the data due to annual differences in localized water use. Crop rotation or localized weather patterns can significantly alter the water use in the area exerting primary influence over the water levels of a monitoring well.

An alternative analysis method was evaluated for this paper that uses simple linear regression to address these limitations. Calculating a regression line that best fits the data uses all available data in the trend characterization, makes it possible to estimate a value for a missing point, and filters out much of the short term variability in the trends.

In addition to addressing the limitations of non-regression trend analysis, the regression analysis has two other advantages which are:

1) Regression results can be used to predict future water levels.

2) Annual deviations from the regression trend may be due to specific annual conditions. Differences between trend and actual values may be correlated to annual events such as weather and cropping patterns. This would assist in developing an empirical model of the aquifer trends which could be used both in making regional management decisions and in comparison analysis of theoretically based models of the aquifer system.

METHODS

Fall water level data for alluvial aquifer wells in Sunflower County were extracted from the Delta-wide alluvial aquifer water level database obtained from the Mississippi Department of Environmental Quality. Sunflower County has experienced a fairly wide-spread and consistent decline in water levels since the monitoring network was established. Only wells in Sunflower County with a minimum of 5 fall measurements and with at least 1 measurement from 1990 through 1992 were used. This represents a set of 40 wells in Sunflower County with similar characteristics to use for testing the regression method. Linear regression was performed on all fall data from 1980 through 1992 with date as the independent variable and water level below land surface as the dependent variable.

RESULTS AND DISCUSSION

Measured water levels and resulting regression lines are given for 3 wells in Figure 1. Regression line slopes, intercepts, and correlation coefficients are given for all Sunflower County monitoring wells in Table 1. The slope of the lines represents the long term average changes in water levels for each well. The average annual change from the regression analysis for all wells tested in Sunflower County was a decline of -0.43 feet per year with a standard deviation of 0.33 feet. The fit of the data to a linear pattern is fairly strong, suggesting that application of regression analysis to the water level data may be of value. Wells with poor correlation coefficients, as in Figure 1C, often have slopes near zero. Regression analysis of flat lines (zero slope) inherently produces poor correlation coefficients due to the nature of the calculations.

To evaluate the potential for regression analysis to reduce noise in data sets, the annual change in the water level from fall of 1991 to the fall of 1992 for all wells with available data was calculated from actual measurements. The results are also given in Table 1 along with the annual change (equivalent to slope) calculated for each well from the regression line. The annual decline for that year calculated from the actual measurements is -0.42 feet with a standard deviation of 0.77 feet. The regression data estimated decline was -0.43 feet with a standard deviation of 0.33 feet. The average county-wide decline predicted by the regression is only 2% greater than the decline predicted from the actual data but the standard deviation of the regression changes is less than one-half of the standard deviation of the changes calculated from the actual data. In this case, the regression effectively predicts the average change for the county with a considerably narrow population distribution as indicated by the smaller standard deviation. The narrower population distribution indicates that regression analysis does reduce some of the localized noise from the data set.

Spatial analysis of the data provides additional information. Figure 2 shows the locations of each well used in the analysis in Sunflower County with the annual changes broken into different ranges which are indicated by different symbols. The annual change for each well calculated from actual measurements is given in Figure 2A and form regression estimates in Figure 2B. Notice the elimination of triangle symbols (which represent the extremes in change) in Figure 2B. This is an example of how extreme data values, or possible fliers, can be identified and removed form a data set being used for regional trend analysis. This is frequently a problem when constructing water level decline isoline maps for the alluvial aquifer. Occasionally, a water level change calculated from measured values will be significantly different from other well changes in the area. The data must then be examined and a determination made to keep the data and include it in the regional analysis or to reject the individual data point. This is usually a subjective decision. The use of regression on the data can produce a data set with much of this complication removed. This would not eliminate the need to closely examine all data used in analysis, but it can provide a nonsubjective tool to assist the researcher or planner in making decisions about individual data points. Application of the regression analysis also increased the number of wells available for inclusion in regional analysis. Figure 2B (regression) has 40 well sites while Figure 2A (actual data) has 30. The first and last value must be available to calculate the change from the actual data. A missing data point results in fewer sites. Regression analysis provides a method to eliminate this limitation. Figure 1A gives a graphic example of how regression analysis could be used to estimate values for missing data.

Spatial analysis of the correlation coefficients is given in Figure 3. The correlation coefficients are broken into ranges and each range is represented by a different symbol. There is a general grouping of wells with poor linear relationships with water level changes in the southern portion of the county. No explanation is offered for this pattern at this time, but it suggests that there may be a fundamental difference in the relationships of water use and recharge on the alluvial aquifer in different portions of Sunflower County. This pattern provides a valuable research and management tool by grouping areas for management and study of aquifer characteristics.

Regression analysis is a useful tool in the analysis of water level changes in the Mississippi River Alluvial Aquifer, especially when used with spatial presentations of the data. Application of regressions to Delta-wide water level data should significantly simplify and improve results from the analysis of the data.

Table 1. Well identification numbers and calculations based on fall 1980 through 1992 data from Sunflower County. Calculations include slope, correlation coefficient, number of observations for regression analysis, and '91 to '92 changes calculated from actual data.

WELL ID	NUMBER OF OBSERVATIONS	CORRELATION COEFFICIENT r^2	SLOPE FEET/YEAR	ACTUAL DATA '91-'92 CHANGE FEET/YEAR
A031	9	0.48	-0.68	-1.12
A032	10	0.51	-0.55	
B003	13	0.70	-0.46	0.00
B056	13	0.17	-0.38	-0.04
B068	11	0.64	-0.99	0.30
C010	11	0.77	-0.51	-0.69
C011	12	0.49	-0.45	-0.97
E058	9	0.89	-0.46	
E062	12	0.27	-0.18	-0.83
E063	12	0.42	-0.46	-0.62
F001	10	0.92	-0.71	-0.27
F051	12	0.87	-0.53	0.09
G003	13	0.77	-0.73	0.27
G077	9	0.53	-0.78	
H067	12	0.57	-0.65	-1.78
H070	6	0.32	-0.62	
H096	10	0.13	-0.31	-0.99
J007	12	0.52	-0.40	0.12
J051	10	0.79	-0.35	-1.13
J050	10	0.91	-0.61	
K005	12	0.63	-0.87	1.20
K007	13	0.45	-0.55	-1.50
K057	8	0.71	-0.63	
L027	13	0.67	-0.61	-2.05
M038	13	0.21	-0.20	-1.55
M060	9	0.23	-0.10	-0.58
N101	11	0.64	-0.78	
N046	11	0.14	0.16	1.15
P023	13	0.13	-0.09	-0.48
P030	12	0.12	-0.12	-0.32
P031	12	0.44	-0.23	
Q004	12	0.15	-0.20	-0.49
Q008	14	0.44	-0.45	-0.19
Q068	12	0.52	-0.28	0.52
R063	12	0.54	-0.68	-0.39
S037	11	0.27	-0.22	-0.40
S040	10	0.55	-0.29	0.45
T048	13	0.44	-0.82	-0.29
T049	7	0.57	0.99	
T050	10	0.82	-0.59	
MEAN	11	0.51	-0.43	-0.42
MAX	14	0.92	0.99	1.20
MIN	6	0.12	-0.99	-2.05
STANDARD DEVIATION			0.33	0.77

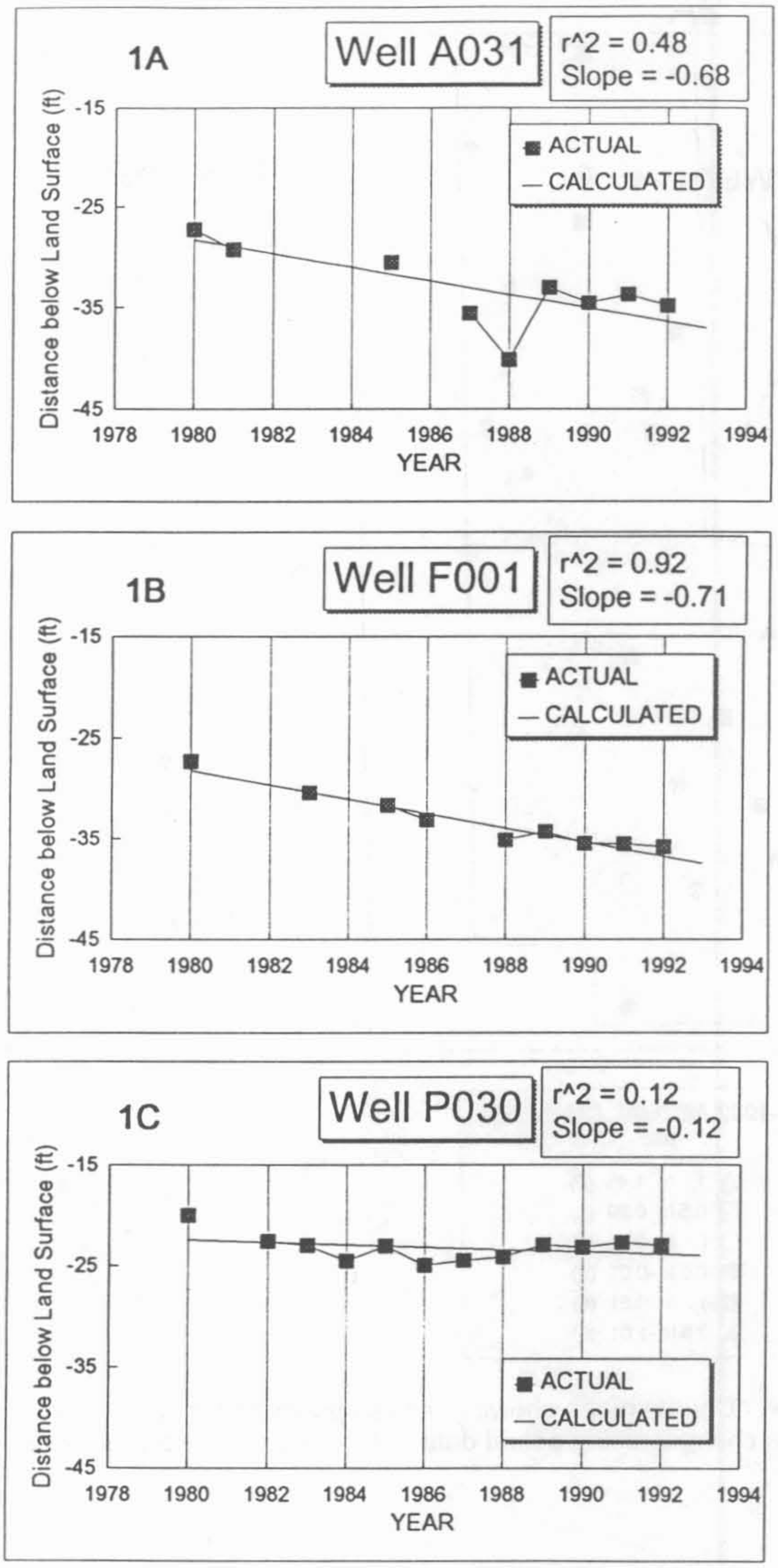


Figure 1. Measured water levels below land surface for fall data and the regression line for those measurements for the Mississippi River Alluvial Aquifer in Sunflower County. Data in 1A, 1B, and 1C represent wells with typical, best, and poorest correlation coefficients, respectively.

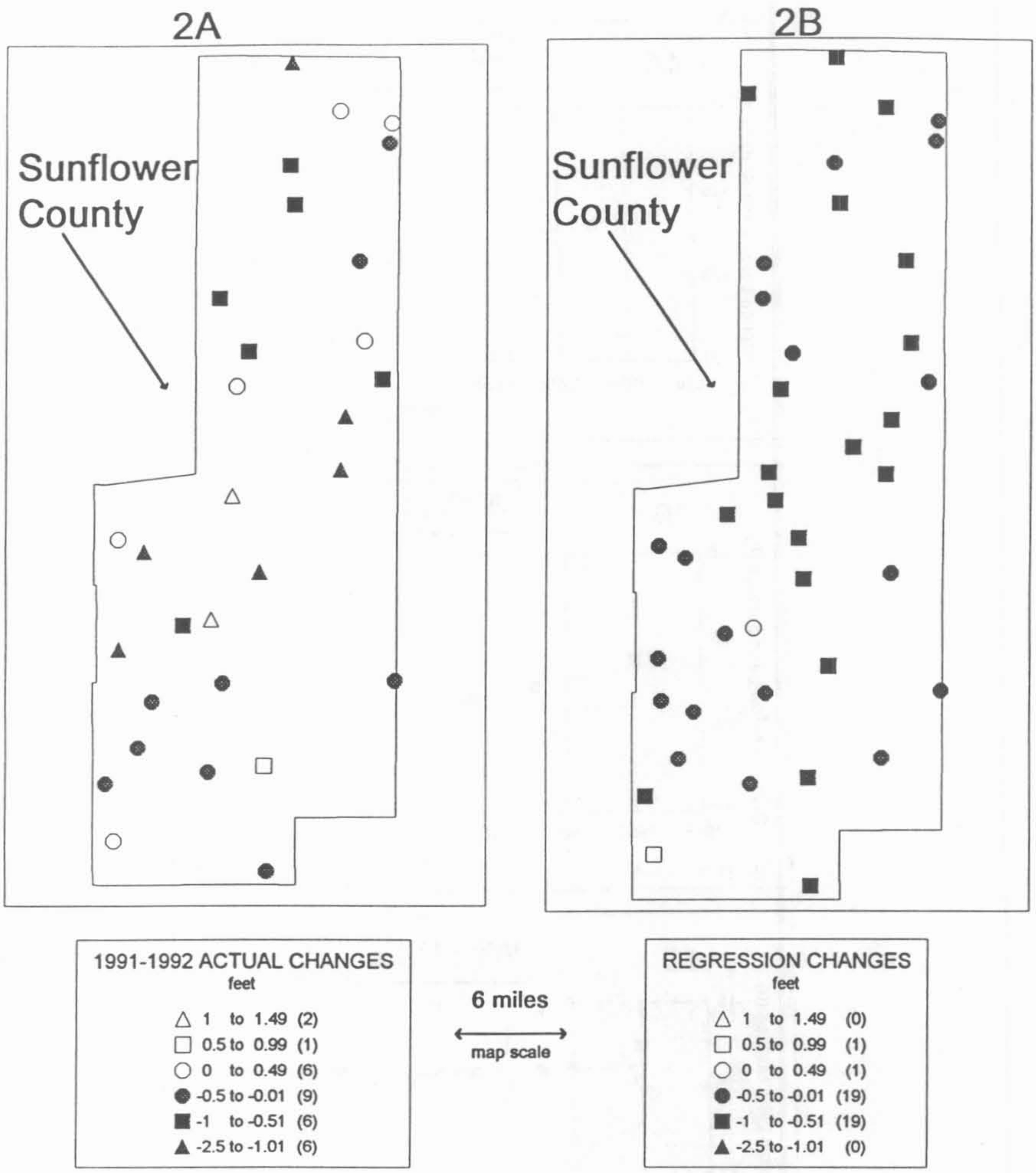
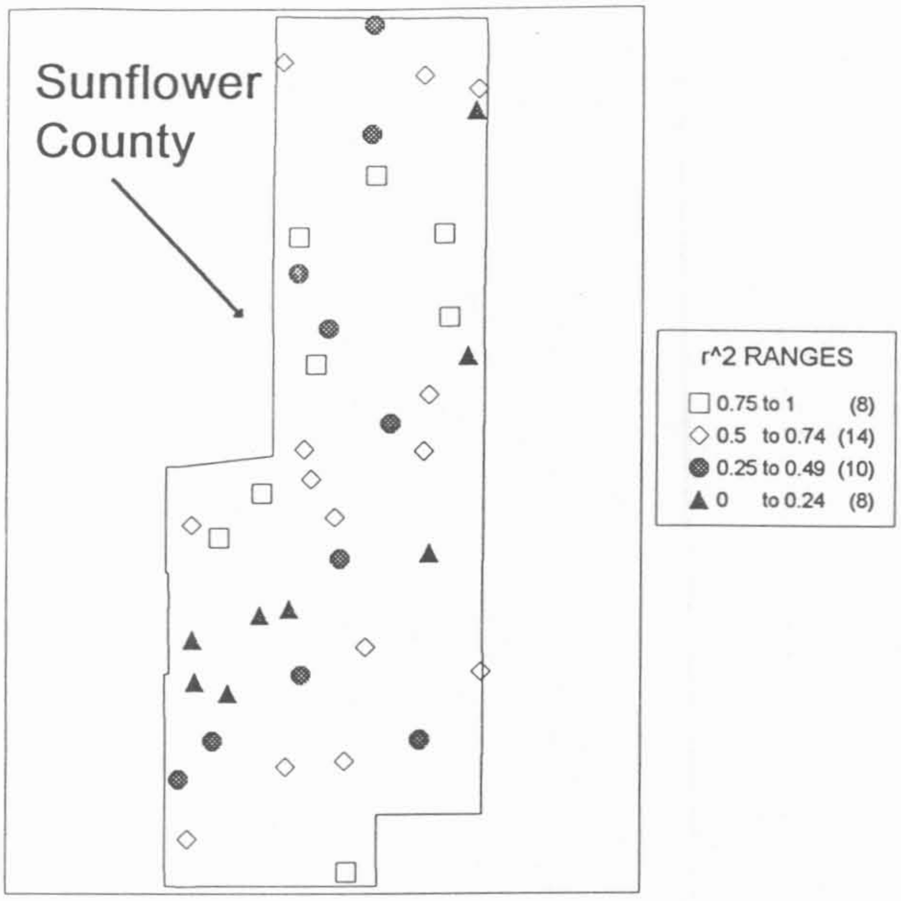


Figure 2. Sunflower County maps showing the ranges of 1 year water level changes and well locations. Map 2A shows the changes using actual data. Map 2B shows the changes using regression analysis.



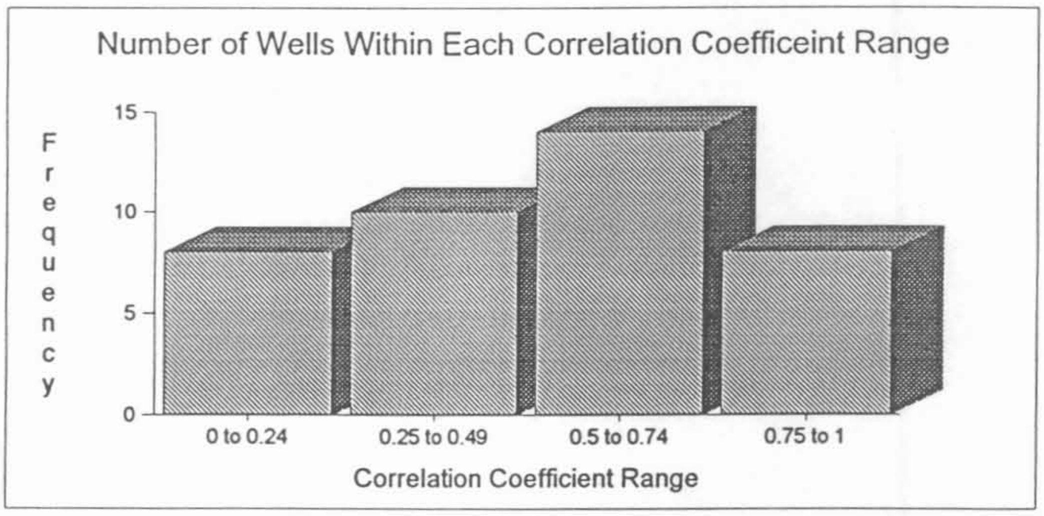


Figure 3. Sunflower County map and frequency distribution of correlation coefficients derived from linear regression calculations for fall data from 1980 through 1992.