American Journal of Environmental Engineering 2018, 8(5): 181-200
DOI: 10.5923/j.ajee.20180805.03

Can Prior Experience Provide a Means to Predict Success of Future Aquifer Storage and Recovery Systems?

Frederick Bloetscher

Florida Atlantic University, Boca Raton, FL, USA

Abstract This paper is the result of analysis data gathered from a 2013 survey of all 204 Aquifer Storage and Recovery (ASR) sites in the United States. That 2013 ASR site survey included all active, inactive, test and study sites, and collected both operational and construction details. The differences between the operational and inactive sites are of particular interest because the differences are where the most information can often be gleaned as to the potential for success of the test and study sites. The statistical analysis utilized in this analysis focused on the active and inactive sites – all sites in study mode and early stages of development were not included in the initial analysis. The intent was to determine is a predictive model for ASR success could be developed for the test and study ASR sites, as well as potential future sites. The results improve on prior papers by the author related to ASR system success and provides insight on what factors improve the likelihood of successful ASR projects. Using the results of the PCA, a linear regression model was developed for the active and inactive sites, and applied to the test and study sites to predict their likelihood of success. The results provide insight into the potential for success in the 50+ test/study sites that may be years for full development.

Keywords Groundwater Storage, Recharge, Predicting Success

1. Introduction

Water supply challenges exist throughout the world. As a result, in drought or water limited areas, the ability to store water for later use has value for sustainability of the local community. *AWWA Manual M21* [1] divides aquifer storage programs into four categories: Artificial Aquifer Creation, Aquifer Recharge, Aquifer Reclamation, and Aquifer Storage and Recovery (ASR). All of these approaches are used as part of the water supply industry to ensure that sustainable water resources are available for agricultural, environmental and urban uses. This paper focusses on the ASR portions only and utilizes the dataset developed in conjunction with AWWA Manual M-63 [2-4]. ASR is touted as a viable concept in the management of both potable and non-potable water supplies. Utilities pursue ASR programs to increase the efficiency of system operations to utilized unused water treatment plant capacity to treat water and pump it into an aquifer for later withdrawal for augmentation of water supplies at a later point of time to avoid the need to construct plants only for peak demands [2-4]. The injection applications include potable water, raw surface and groundwater, and reclaimed wastewater. The storage period can be over multiple months to allow the stored water to meet the next high demand season, an emergency such as a severe drought or during an interruption of water withdrawal due to equipment breakdown.

The concept of ASR has only been applied in the United States since the late 1960s and little development occurred until the 1990s (see Figure 1). As a result, until recently, the number of sites has been limited, and the fact that it may take 10 years to develop an operational ASR system, means that truly acquiring data has only recently become available to a number of sites. Hence, the first complete survey of ASR sites was completed in 2013, and little has changed since that time [3, 4]. Dataset was the first comprehensive analysis of the 204 sites in the US. U.S. EPA and environmental agencies in each state with ASR wells were contacted by phone or email to whether the state had such programs in place or not, and where they might be located. The list of ASR sites identified by the regulatory agencies was a critical component of the project because while prior inventories were prepared by regulatory agencies and consultants, none were complete and most excluded projects that were no longer active [5-9]. In each of these documents, the goal was to provide information on successful ASR sites as case studies and were relatively limited to a few sites as opposed to a nationwide survey (for example, AWWA [5] included only 4 sites in Florida, as opposed to 54). Hence, while AWWA [5] and Bloetscher et al [6] provided more extensive summaries that the texts by other authors, these reports were

* Corresponding author:
h2o_man@bellsouth.net (Frederick Bloetscher)
Published online at http://journal.sapub.org/ajee

also very limited in scope. No analysis of the data was conducted to identify trends, success and challenges for ASR projects. The first to analyze the successes and challenges encountered by ASR projects were Bloetscher et al [3, 4].

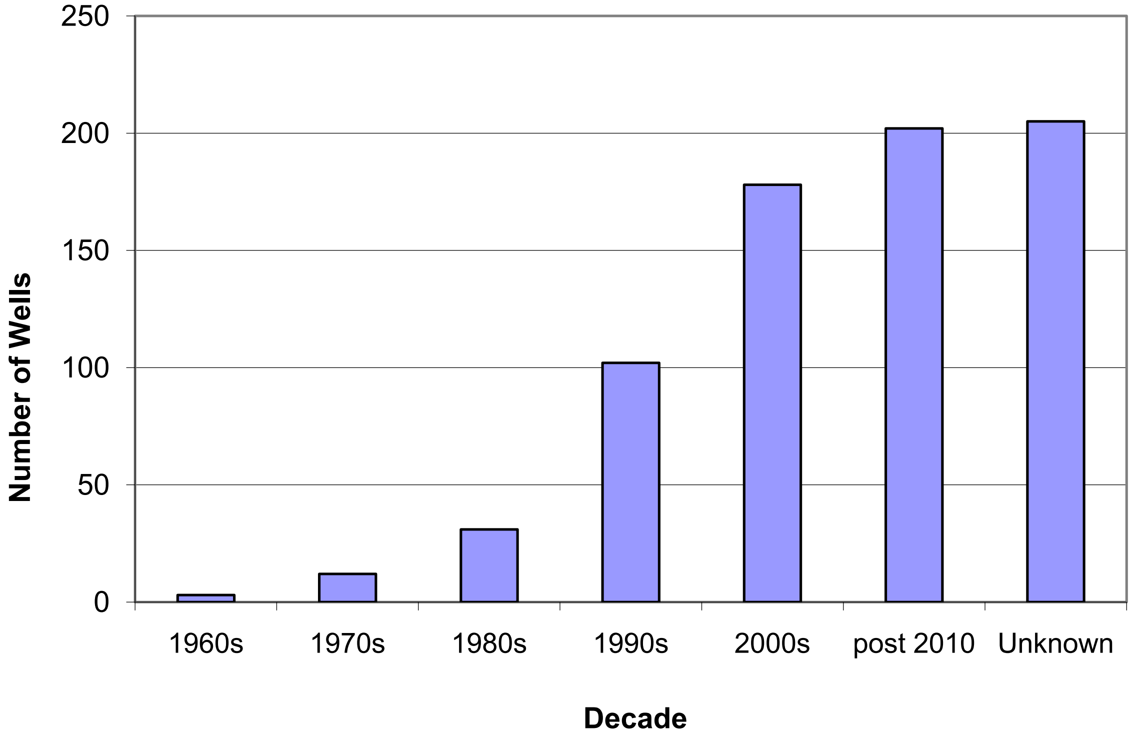

Figure 1. Cumulative ASR sites by Decade

Bloetscher et al [3, 4] outlined the basic findings of the survey, the lessons learned from the ASR survey and summarized the salient commonalities. Variables of interest were identified to account for operations, physical construction, and localized differences [2-4]. One of the issues that arose as a part of the survey was that nearly a third of these systems were not active and another third were in some phase of the testing mode. The inactive sites include a total of 220 wells that are not in use at this time. A statistical attempt was made in Bloetscher [3] to identify why ASR projects were active (or not), but no attempt was made to predict the likelihood of success of the wells in test mode. While much can be learned from successful projects, comparing the successful and inactive systems can provide insights into the criteria and process associated with the development of an ASR program and perhaps highlight factors that will suggest a high rate of success for those test and study projects.

2. Methodology

The data utilized for this analysis are noted in Table 1, which were variables extracted from the 2013 ASR site inventory [2], and then converted to numerical variables as required for the statistical methods employed (see Tables 2 and 3). Also, information was updated to reflect known changes in the ASR wells. Am0ong the issues noted was that complete information was not available for all sites and decisions needed to be made to determine is those ASR sites would be retained or the variables deleted. For example, the salinity of the injection zone is relevant when injecting fresh water into a brackish zone. Freshwater will float based on the principles of differential density, creating a challenge for recovery of the injected water. However, the dataset denoted that the majority of sites were injecting into freshwater (total dissolved solids under 1000 mg/L) except in south Florida [2]. As a result this variable was deleted as opposed to deleting several dozen sites that did not report the salinity. The decisions were important because those sites with incomplete data, or those variables that were incomplete, cannot be used in principal component analysis which would reduce the available data considerably. Likewise, the casing material was commonly not reported and the confined layer material was not well defined. These variables were also deleted to permit as many sites to remain as possible.

Table 1. Variables used in ASR Analysis

Status Operational (or not – note that test wells and well not drilled were not included in the analysis
Midwest/ Central location (TX, KS, OK, TN (Memphis))
East location (NJ, NY, DE)
Rockies location (CO, WY)
NW location (OR, WA, ID)
Southwest location (CA, NV, AZ, NM, UT)
FL location
SE location (NC, SC, VA)
West location
Estimated Start Date
Number of wells in the project
Number of Active wells
Number of inactive or abandoned wells in the project
Clogging Issues noted
Metal Leaching Issues noted
THMs/ WQ Issues noted
Low Recovery noted
Lack of Water Availability for Recharge
Unknown issues
No issues noted
Surface water sources
Reclaimed water sources
Groundwater source
Irrigation Use
Cooling Usage
Raw Water supplement
Potable Water use
Confined Aquifer formation
Alluvial formation
Limestone formation
Sand/ Sandstone formation
Basalt formation
Number of Storage Cycles
injection Capacity (MGD)
Withdrawal Capacity (MGD)
Ratio Pumping in/out
Peak Flow on Site (MGD)
Depth of well (ft)
Injection Horizon (depth in ft)
Depth of Casing (ft)
Amount of Water Stored (MG)

Table 2. Descriptive statistics of continuous variables related to the ASR sites in the United States

	Observation	Missing	Minimum	Maximum	Average	Standard Deviation
Estimated Start Date	204	0	1963	2014	1999	10
Active	75	0	1968	2010	1996	8
Inactive	55	0	1963	2007	1995	10
Test/Study	74	0	1970	2014	2004	9
Number of Active Wells	201	3	0	87	4	8
Active	75	0	0	87	6	11
Inactive	54	1	0	40	3	5
Test/Study	72	2	0	15	2	2
Number of Inactive Wells	201	3	0	40	1	3
Active	75	0	0	18	1	2
Inactive	54	1	0	40	3	5
Test/Study	72	2	0	5	0	1
Number of Monitoring Wells	203	1	0	10	1	1
Active	74	1	0	6	1	1
Inactive	55	0	0	10	1	2
Test/Study	74	0	0	7	0	1
Injection Capacity	178	26	0	15	1.4	1.9
Active	73	2	0.1	10	1.3	1.6
Inactive	49	6	0	9	1.4	1.7
Test/Study	56	18	0	15	1.5	2.4
Withdrawal Capacity	180	24	0	15	1.9	2.2
Active	74	1	0.1	10	1.9	1.9
Inactive	52	3	0	9	1.9	2
Test/Study	54	20	0	15	2.1	2.7
In/Out Ratio	174	30	0.02	5.25	0.9	0.5
Active	73	2	0.06	5.25	0.9	0.7
Inactive	48	7	0.19	1.02	0.8	0.3
Test/Study	53	21	0.02	2.5	0.8	0.4

	Observation	Missing	Minimum	Maximum	Average	Standard Deviation
Peak Flow on Site (MGD)	175	29	0	40	3.9	5.3
Active	73	2	0.1	23	4.7	4.9
Inactive	49	9	0	23.7	3.3	4
Test/Study	53	21	0	40	3.6	6.7
Amount of Water Stored (MG)	162	42	0	78,400	1,282.1	6,262.5
Active	69	6	0.2	78,400	2,166.5	9,462
Inactive	46	9	0	3,800	654	1,039.3
Test/Study	47	27	0	8,400	598.4	1,347.5
Depth of Well (ft)	181	23	33	3,882	801.3	560.5
Active	73	2	75	2,523	789.7	489.7
Inactive	52	3	33	1,770	728.3	456.1
Test/Study	56	18	50	3,882	884.2	713.6
Depth of Casing (ft)	178	26	9	3,832	594.6	489.4
Active	50	5	39	2,185	550.5	424.8
Inactive	52	3	10	1,457	561.4	371.9
Test/Study	56	18	9	3,832	680.5	636.4
Injection Horizon	159	45	7.5	1,501	225.6	234.3
Active	65	10	7.5	1,000	231	214.8
Inactive	46	9	12	1,186	200.2	228
Test/Study	48	26	21	1,501	242.7	266.3
Diameter of Casing (ft)	136	68	6	40	14.9	5.3
Active	56	19	6	26	14.5	4.2
Inactive	43	12	6	40	15.6	6.9
Test/Study	37	37	6	24	14.7	4.7
Transmissivity (gpd/sf)	127	77	0.1	620,136	75,534.6	137,647.4
Active	30	45	1	620,136	79,158.	154,511.30
Inactive	21	34	1.2	264,000	56,732.30	90,237.60
Test/Study	26	48	0.1	600,000	86,540.8	151,498.6
Total Dissolved Solids (TDS) (ppm)	66	138	50	6,000	1,563.50	1,732.50
Active	23	52	150	5,500	1,117.10	1,569.50
Inactive	28	27	140	6,000	1,896.40	1,783.40
Test/Study	15	59	50	6,000	1,631	1,837.50

Table 3. Descriptive statistics of categorical variables per ASR program status in the United States

	Observations	Missing	Basin Range	California	Mid-Atlantic	Mid-West	Pacific NW	SE Coast
Region	204	0	44	28	23	8	29	72
Active	75	0	23	14	12	3	7	16
Inactive	55	0	12	3	3	3	5	29
Test/Study	74	0	9	11	8	2	17	27
% Success of Non-Testing/Study Sites			66%	82%	80%	50%	58%	36%

	Observations	Missing	None	Clogging	Expired Permit	PWS Conversion	Recovery	Tested then Abandoned	Water Quality/ Arsenic
Issues with ASR	190	14	111	29	1	1	20	6	22
Active	72	3	56	13	1	0	0	0	2
Inactive	48	7	3	12	0	0	17	3	13
Test/Study	70	4	52	4	0	1	3	3	7
% Success of Non-Testing/Study Sites			95%	52%	100%	N/A	0%	0%	13%

	Observations	Missing	Groundwater	Industrial	Reclaimed	Surface Runoff
Water Source	204	0	41	1	28	134
Active	75	0	13	1	9	52
Inactive	55	0	17	0	7	31
Test/Study	74	0	11	0	12	51
% Success of Non-Testing/Study Sites			43%	100%	56%	63%

	Observations	Missing	Cooling	Fire	Irrigation	PWS	Raw	RSW
Water Use	203	0	5	2	28	108	52	8
Active	75	0	3	1	9	45	14	3
Inactive	55	0	0	0	5	28	18	4
Test/Study	73	0	2	1	14	35	20	1
% Success of Non-Testing/Study Sites			100%	100%	64%	62%	44%	43%

	Observations	Missing	0-1	2-5	6-10	11-20	>20
Number of Storage Cycles	204	0	38	73	28	46	19
Active	75	0	2	14	11	32	16
Inactive	55	0	8	31	10	4	2
Test/Study	74	0	28	28	7	10	1
% Success of Non-Testing/Study Sites			20%	31%	52%	89%	89%

	Observations	Missing	PVC	Fiber Glass	Stainless Steel	Steel
Casing Material	157	47	18	3	14	122
Active	67	8	5	0	7	55
Inactive	45	10	7	2	1	35
Test/Study	45	29	6	1	6	32
% Success of Non-Testing/Study Sites			42%	0%	88%	61%

	Observations	Missing	None	T&P
T&P Code	185	19	180	5
Active	74	1	73	1
Inactive	53	2	50	3
Test/Study	58	16	57	1
% Success of Non-Testing/Study Sites			59%	25%

	Observations	Missing	Alluvial	Basalt	Carbonite	Granite	Limestone	Sand	Sand/ Clay Mixture	Sandstone
Injection Formation Code	185	19	61	14	2	1	62	16	12	17
Active	70	5	31	4	1	0	13	9	6	6
Inactive	51	4	12	1	1	0	26	4	1	6
Test/Study	64	10	18	9	0	1	23	3	5	5
% Success of Non-Testing/Study Sites			72%	80%	50%	N/A	33%	69%	86%	50%

	Observations	Missing	None	Alluvial	Basalt/ Clay Mixture	Bedrock	Clay	Dolomite	Limestone	Silt	Shale/ Sedimentary
Confinement Unit Formation Code	90	114	3	2	14	3	30	13	13	7	5
Active	32	43	1	1	3	2	12	5	3	3	2
Inactive	28	27	1	0	2	0	11	3	6	3	2
Test/Study	30	44	1	1	9	1	7	5	4	1	1
% Success of Non-Testing/Study Sites			50%	100%	60%	100%	52%	63%	33%	50%	50%

2.1. PCA and FA Analysis

The factor analysis method dates from Spearman [10] and continues to develop. Today, there are two main types of factor analysis: Exploratory factor analysis (or EFA) and Confirmatory factor analysis (or CFA). EFA is used by XLStat® to reveal the possible existence of underlying factors which give an overview of the information contained in a very large number of measured variables. For EFA, the structure linking the variables is initially unknown, but the number of factors is assumed. CFA uses a method identical to EFA but the structure linking underlying factors to measured variables is assumed to be known [11].

Principal Component Analysis (PCA) is popular multivariate technical mainly used to reduce the dimensionality of *p* multi-attributes to two or three dimensions [11-13]. PCA is a special case of factor analysis (where k, the number of factors, equals p, the number of variables). While FA assumes a number of factors, PCA is used to reduce the number of variables to factor sets, while maximizing the unchanged variability in order to obtain independent (non-correlated) factors [14]. The mathematics of PCA uses an orthogonal transformation convert observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components [13-16]. PCA uses a multivariate statistical parameter called an eigenvalue, which is a measure of the amount of variation explained by each principal component. PCA summarizes the variation in a correlated multi-attribute to a set of uncorrelated components, each of which is a particular linear combination of the original variables [17]. PCA is the simplest of the true eigenvector-based multivariate analyses. A Scree Plot is a simple line segment plot that shows the fraction of total variance in the data as explained or represented by each component [18].

There are several uses for PCA, including [11]:

- The study and visualization of the correlations between variables to hopefully be able to limit the number of variables to be measured afterwards;
- Obtaining non-correlated factors which are linear combinations of the initial variables so as to use these factors in modeling methods such as linear regression, logistic regression or discriminant analysis.
- Visualizing observations in a 2- or 3-dimensional space in order to identify uniform or atypical groups of observations.

Two methods are commonly used for determining the number of factors to be used for interpreting the results: the Scree test [19] is based on the decreasing curve of eigenvalues. The number of factors to be kept corresponds to the first turning point found on the curve. However, these representations are only reliable if the sum of the variability percentages associated with the axes of the representation space are sufficiently high. If this percentage is high (for example 80%), the representation can be considered as reliable. If the percentage is reliable, it is recommended to produce representations on several axis pairs in order to validate the interpretation made on the first two factor axes.

The correlation biplot interprets the angles between the variables as these are directly linked to the correlations between the variables. The position of two observations projected onto a variable vector can be used to determine their relative level for this variable [11]. The Kaiser-Guttman rule suggests that only those factors with associated eigenvalues which are strictly greater than 1 should be kept [11]. The number of factors to be kept corresponds to the first turning point found on the curve. Crossed validation methods have been suggested to achieve this aim.

2.2. Linear Regression

Ultimately the goal is to determine if the condition has a consequence – i.e. the potential for failure. If so, one needs to know what that consequence is – in this case operation or inactive. The values were assigned for operational (1) or inactive (0) of aquifer storage units in the United States, as the dichotomous dependent variable. The impact of these factors can be developed via a linear regression model [12]. The model would be developed as follows [11, 20-21]:

$$SSI = w_1C_1+w_2C_2+w_3C_3+w_4C_4+\dots w_iC_i$$

where:

- SSI = Site success index (consequence)
- w = weighting factor
- C is condition factor

If one knows the consequence, the weights can be found:

$$f(x) = c_1x_1 + c_2x_2 + c_3x_3 + \dots + .c_nx_n$$

where the values of c_n are real numbers and

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

are the factors which are a compilation of the original carriable to maximize variance. It assumes these constraints and linear variables in the matrices are non-negative. If there are negative values, they must be made positive as follows [11]:

$$x_i^+ = \begin{cases} x_i \text{ } if \text{ } x_i \geq 0 \\ 0 \text{ } otherwise \end{cases}$$

$$x_i^- = \begin{cases} -x_i \text{ } if \text{ } x_i > 0 \\ 0 \text{ } otherwise \end{cases}$$

the linear regression model provides a mechanism to model the data to determine if differences between the active and inactive projects exists. For the existing sites, if the site was active, the consequence value was assigned a value of 1. If not, 0. As a result the hypothesis was that those sites likely to be successful if the SSI would tend toward a value of 1, and those that likely would not pan out, would trend toward 0. Note that because certain factors may have no value at present (example depth of an undrilled well), it is possible that the regression equation provides an SSI result that is greater than 1 or less than 0.

2.3. Further Data Manipulation

Because the test and study site have incomplete data, the linear regression model was re-run to include only that data that would apply to the test and study sites. For example, if no well was drilled, the casing and well depths could not be known. The revised linear regression model was used to model the test and study sites to predict the likelihood of success.

3. Results and Discussion

The states with the most ASR programs are Florida (54), followed by California, New Jersey, Arizona and Oregon (see Figure 2). However, the presence of ASR sites is not necessarily an indicator for success of ASR projects. For example, in Florida, over half the sites are not active or have wells that are no longer used. With the elimination of inactive and test sites, there are only 22 active ASR sites (as compared to 54 ASR sites) in Florida.

Table 2 outlines the descriptive statistics for descriptive statistics for all sites for the full 2013 database. Table 3 includes the categorical variables from the 2013 dataset. From the 204 sites in Tables 1 and 2, 74 were removed as a part of the process because they were in study or test mode and therefore lacked certain data that is helpful in understanding the potential for success. Removal of these sites led to Table 4 which summarizes the remaining variables. Note because PCA and FA require no missing information, the number of complete datasets was reduced to 111. Also note that the wells were grouped into regions of the country to determine if there were commonalities across different regions. The regions were similar to those proposed in Bloetscher et al [3,4], except that the Florida wells were removed.

Table 5 is a correlation analysis between variables. Significant correlations exist for:

- Sand/sandstone formations in the east
- Unconfined alluvial formations in the west/southwest
- Confined limestone formations in Florida, and
- Reclaimed water being stored for irrigation

The Scree plot [19] showed that the factors created by the eigenvalues required 11 factors to obtain 70 percent of the variance, which is a lot of factors and suggests that there is much scatter in the variables – one reason the locations were developed as a means to attempt to compare commonalities among regions and create greater degrees of correlation. The factor loadings revealed the factor loading and therefore the factor correlation with the original variables:

- F1 – relates to the location – Southwest and west locations were correlated with unconfinement and alluvial formations,
- F2 – relates to formation and water source – limestone and the use of groundwater, correlated with Florida
- F3 – relates to the number of active wells on a site (more increasing likelihood of success), depth of the well and depth of the casing
- F4 – relates to raw water as a source for the ASR wells
- F5 – relates to the number of wells (more increasing likelihood of success)
- F6 – relates to northwest wells (and basalt formations)
- F7 – relates to the number of inactive wells

All other actors had very limited factorial combinations as demonstrated by the relationships in Table 6. Each of these factors also contributed significantly to the factor loading (see Table 7).

PCA permits the use of a varimax rotation to improve correlations to explain variability. However, the varimax rotation does not significantly help to reduce the number of variables for the project but does reinforce several things:

- D1 – relates to the location – Southwest and west locations were correlated with unconfinement and alluvial formations, and not limestone
- D2 – relates to formation and water source – limestone and the use of groundwater, correlated with Florida, and differentiated from sand in the remaining southeast
- D3 –depth of the well and depth of the casing are related, perhaps weakly to recovery
- D4 – raw water and potable use from the ASR well are related.
- D5 – relates to the number of wells and number of active wells (more increasing likelihood of success)
- D6 – reinforces the relationship between northwest wells and basalt formations
- D7 – relates to active status of wells
- D8 – notes that ground and surface water system are inversely related
- D9 – relates to injection horizon and metals recovery and
- D10 – relates to withdrawal and injection capacity beings strongly correlated.

Table 8 shows that the varimax variables and their contribution to variance (see Figure 3).

The next step was to run a linear regression model in an attempt to understand if an equation could be developed to predict success. A linear regression model was run in XLSTAT®. Several variables were deleted from the original data set because they do not appear in the data for the test or study projects. The remaining variables are shown in Table 9. Table 10 shows the weight components applied to each variable (see also Figure 4). Figure 5 shows the results – predicted results for active versus non-active sites based on the predicted means (0.687 vs 0.35), and standard deviations (0.280 vs 0.180) for the active and inactive sites, respectively (note there were 58 active sites and 53 inactive sites). The model predicts these relatively well. Note that potable water use and the number of inactive wells did not contribute to the variance.

These same factors were applied to the data on the test/study wells. Table 11 outlines the variables for these test/study wells. They are similar to those of the active and inactive wells. Using the components in Table 11, Table 12 outlines the results applied to the sites (listed by state only). The factors have a range from just under zero to just over 1, as does the analysis of active and inactive wells. It appears that this model may provide useful information for likely success. Figure 6 shows that 11 of the sites have values under 0.5 (including 5 under 0.35), which means their likelihood of success is low. Twenty-three of the sties have a value greater than 0.687, which suggests that these sites are likely to have success.

Table 4. Summary Statistics for Retained Variables

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Status Operational	111	0	111	0.000	1.000	0.532	0.501
Midwest/ Central	111	0	111	0.000	1.000	0.099	0.300
East	111	0	111	0.000	1.000	0.108	0.312
Rockies	111	0	111	0.000	1.000	0.081	0.274
NW	111	0	111	0.000	1.000	0.072	0.260
Southwest	111	0	111	0.000	1.000	0.171	0.378
FL	111	0	111	0.000	1.000	0.306	0.463
SE	111	0	111	0.000	1.000	0.072	0.260
West	111	0	111	0.000	1.000	0.288	0.455
Est Start Date	111	0	111	1963.000	2011.000	1995.532	9.150
Number of wells in the project	111	0	111	0.000	87.000	4.946	9.930
Number of Active wells	111	0	111	0.000	87.000	3.550	9.461
Number of inactive or abandoned wells in the project	111	0	111	0.000	40.000	1.396	4.286
Clogging	111	0	111	0.000	1.000	0.207	0.407
Metals	111	0	111	0.000	1.000	0.081	0.274
THMs/ WQ	111	0	111	0.000	1.000	0.036	0.187
Recovery	111	0	111	0.000	1.000	0.135	0.343
Water Avail	111	0	111	0.000	1.000	0.063	0.244
unknown	111	0	111	0.000	1.000	0.144	0.353
none noted	111	0	111	0.000	1.000	0.342	0.477
surface	111	0	111	0.000	1.000	0.622	0.487
reclaimed	111	0	111	0.000	1.000	0.144	0.353
ground	111	0	111	0.000	1.000	0.243	0.431
irrigation	111	0	111	0.000	1.000	0.126	0.333
cooling	111	0	111	0.000	1.000	0.018	0.134
Raw	111	0	111	0.000	1.000	0.270	0.446
Potable	111	0	111	0.000	1.000	0.586	0.495
Confined	111	0	111	0.000	1.000	0.649	0.480
Alluvial	111	0	111	0.000	1.000	0.333	0.474
Limestone	111	0	111	0.000	1.000	0.342	0.477
Sand/ Sandstone	111	0	111	0.000	1.000	0.261	0.441
Basalt	111	0	111	0.000	1.000	0.072	0.260
Storage Cycles	111	0	111	0.000	74.000	11.090	10.358
injection Cap	111	0	111	0.100	10.000	1.416	1.692
Withdr Capacity	111	0	111	0.200	10.000	2.013	2.005
Ratio in/out	111	0	111	0.056	5.250	0.845	0.574
Peak Flow on Site (MGD)	111	0	111	0.000	714.000	10.900	67.537
Depth of well	111	0	111	40.000	2523.000	779.613	462.904
Injection Horizon	111	0	111	0.000	1186.000	189.946	200.642
Depth of Casing	111	0	111	15.000	2185.000	589.667	404.711

Table 5. Correlation Analysis of retained Variables

Variables	Status Operational	Midwest/ Central	East	Rockies	NW	Southwest	FL	SE	West	Est Start Date	Number of wells in the project	Number of Active wells	Number of inactive or abandoned wells in the project	Clogging	Metals	THMs/ WQ	Recovery	Water Avail	Unknown	None noted
Status Operational	**1**	0.009	**0.269**	0.080	-0.018	0.043	**-0.199**	0.052	-0.040	0.017	0.179	**0.273**	**-0.188**	-0.055	-0.184	**-0.206**	**-0.210**	**-0.202**	-0.077	**0.525**
Midwest/ Central	0.009	**1**	-0.115	**0.233**	-0.092	-0.151	**-0.220**	-0.092	**-0.211**	-0.016	0.084	-0.090	**0.393**	-0.021	0.122	-0.064	-0.043	-0.086	-0.050	0.078
East	**0.269**	-0.115	**1**	-0.103	-0.097	-0.158	**-0.231**	-0.097	**-0.222**	-0.052	-0.118	-0.085	-0.087	-0.106	-0.103	-0.067	-0.138	-0.090	0.022	**0.299**
Rockies	0.080	**0.233**	-0.103	**1**	-0.083	-0.135	**-0.197**	-0.083	**-0.189**	0.077	**0.199**	0.035	**0.382**	0.092	-0.088	-0.057	0.076	-0.077	-0.122	0.133
NW	-0.018	-0.092	-0.097	-0.083	**1**	-0.127	-0.185	-0.078	-0.177	0.102	-0.027	-0.031	0.007	**0.287**	0.045	-0.054	-0.110	0.071	-0.015	-0.128
Southwest	0.043	-0.151	-0.158	-0.135	-0.127	**1**	**-0.302**	-0.127	**0.714**	0.010	**0.242**	**0.255**	-0.003	**0.299**	-0.135	-0.088	-0.040	-0.020	0.018	-0.126
FL	**-0.199**	**-0.220**	**-0.231**	**-0.197**	-0.185	**-0.302**	**1**	-0.185	**-0.423**	-0.015	-0.158	-0.111	-0.121	**-0.243**	**0.232**	-0.024	**0.309**	-0.172	**0.228**	**-0.232**
SE	0.052	-0.092	-0.097	-0.083	-0.078	-0.127	-0.185	**1**	-0.177	**-0.192**	-0.044	-0.016	-0.067	-0.142	-0.083	**0.320**	-0.110	-0.072	-0.114	**0.239**
West	-0.040	**-0.211**	**-0.222**	**-0.189**	-0.177	**0.714**	**-0.423**	-0.177	**1**	0.085	**0.201**	**0.221**	-0.022	**0.263**	**-0.189**	-0.016	-0.135	**0.326**	-0.091	-0.166
Est Start Date	0.017	-0.016	-0.052	0.077	0.102	0.010	-0.015	**-0.192**	0.085	**1**	-0.107	-0.149	0.080	0.038	0.084	-0.091	0.104	0.070	-0.044	-0.123
Number of wells in the project	0.179	0.084	-0.118	**0.199**	-0.027	**0.242**	-0.158	-0.044	**0.201**	-0.107	**1**	**0.903**	**0.323**	0.066	-0.105	-0.072	-0.059	**0.313**	-0.050	-0.050
Number of Active wells	**0.273**	-0.090	-0.085	0.035	-0.031	**0.255**	-0.111	-0.016	**0.221**	-0.149	**0.903**	**1**	-0.114	-0.105	-0.098	-0.073	-0.045	**0.363**	-0.038	0.047
Number of inactive or abandoned wells in the project	**-0.188**	**0.393**	-0.087	**0.382**	0.007	-0.003	-0.121	-0.067	-0.022	0.080	**0.323**	-0.114	**1**	**0.385**	-0.028	-0.007	-0.037	-0.076	-0.032	**-0.218**
Clogging	-0.055	-0.021	-0.106	0.092	**0.287**	**0.299**	**-0.243**	-0.142	**0.263**	0.038	0.066	-0.105	**0.385**	**1**	-0.152	-0.099	-0.137	-0.133	**-0.210**	**-0.369**
Metals	-0.184	0.122	-0.103	-0.088	0.045	-0.135	**0.232**	-0.083	**-0.189**	0.084	-0.105	-0.098	-0.028	-0.152	**1**	-0.057	-0.117	0.059	-0.122	**-0.214**
THMs/ WQ	**-0.206**	-0.064	-0.067	-0.057	-0.054	-0.088	-0.024	**0.320**	-0.016	-0.091	-0.072	-0.073	-0.007	-0.099	-0.057	**1**	-0.076	-0.050	-0.079	-0.139
Recovery	**-0.210**	-0.043	-0.138	0.076	-0.110	-0.040	**0.309**	-0.110	-0.135	0.104	-0.059	-0.045	-0.037	-0.137	-0.117	-0.076	**1**	-0.103	-0.162	-0.174
Water Avail	**-0.202**	-0.086	-0.090	-0.077	0.071	-0.020	-0.172	-0.072	**0.326**	0.070	**0.313**	**0.363**	-0.076	-0.133	0.059	-0.050	-0.103	**1**	-0.106	**-0.187**
unknown	-0.077	-0.050	0.022	-0.122	-0.015	0.018	**0.228**	-0.114	-0.091	-0.044	-0.050	-0.038	-0.032	**-0.210**	-0.122	-0.079	-0.162	-0.106	**1**	**-0.296**
none noted	**0.525**	0.078	**0.299**	0.133	-0.128	-0.126	**-0.232**	**0.239**	-0.166	-0.123	-0.050	0.047	**-0.218**	**-0.369**	**-0.214**	-0.139	-0.174	**-0.187**	**-0.296**	**1**
surface	0.049	**0.197**	-0.087	0.164	**0.217**	-0.139	**-0.408**	0.146	0.045	-0.030	0.056	0.014	0.099	0.124	-0.040	0.051	-0.126	0.126	-0.156	0.054
reclaimed	-0.026	-0.136	-0.143	-0.122	-0.114	**0.494**	0.061	-0.114	**0.305**	**0.187**	0.137	0.150	-0.014	0.043	0.066	-0.079	-0.162	-0.001	**0.197**	-0.134
ground	-0.015	-0.118	**0.208**	-0.091	-0.158	**-0.258**	**0.398**	-0.077	**-0.268**	-0.091	-0.184	-0.145	-0.107	-0.134	-0.091	0.003	**0.267**	-0.147	0.006	0.078
irrigation	-0.078	-0.126	-0.132	-0.113	-0.106	**0.332**	0.101	-0.106	**0.237**	**0.219**	-0.099	-0.083	-0.048	-0.060	0.086	0.072	-0.150	-0.099	**0.230**	-0.103
cooling	0.127	-0.045	-0.047	**0.208**	-0.038	-0.062	-0.090	-0.038	0.063	0.096	-0.020	-0.029	0.019	0.098	-0.040	-0.026	-0.054	-0.035	-0.056	0.045
Raw	-0.120	-0.066	**-0.212**	-0.106	0.066	0.154	0.168	-0.091	0.105	-0.118	**0.198**	0.141	0.148	0.139	-0.106	-0.009	0.115	0.009	-0.019	-0.097
Potable	0.126	0.157	**0.293**	0.116	0.022	**-0.346**	**-0.195**	0.164	**-0.272**	-0.067	-0.106	-0.064	-0.106	-0.111	0.049	-0.034	0.012	0.068	-0.123	0.144
Confined	-0.048	-0.072	**0.256**	0.080	0.059	**-0.567**	**0.448**	0.132	**-0.781**	-0.102	-0.185	-0.175	-0.042	**-0.276**	**0.219**	-0.060	0.181	**-0.275**	0.087	0.133
Alluvial	0.013	0.085	**-0.246**	0.000	-0.049	**0.592**	**-0.428**	**-0.197**	**0.773**	0.112	**0.273**	0.182	**0.230**	**0.346**	**-0.210**	-0.034	-0.168	**0.288**	-0.073	**-0.188**
Limestone	-0.160	**-0.239**	**-0.251**	**-0.214**	**-0.201**	**-0.328**	**0.880**	0.166	**-0.459**	-0.082	-0.148	-0.092	-0.138	**-0.275**	**0.203**	0.064	**0.270**	**-0.187**	**0.190**	-0.120
Sand/ Sandstone	0.106	**0.215**	**0.585**	0.124	-0.086	**-0.216**	**-0.395**	0.072	**-0.243**	-0.066	-0.121	-0.089	-0.084	-0.152	-0.026	-0.005	-0.055	-0.154	-0.127	**0.349**
Basalt	0.122	-0.092	-0.097	0.172	**0.731**	-0.127	-0.185	-0.078	-0.177	0.091	-0.034	-0.024	-0.026	**0.201**	0.045	-0.054	-0.110	0.071	-0.015	-0.054
Storage Cycles	**0.266**	-0.167	**0.191**	-0.041	-0.097	0.084	-0.104	0.048	0.114	**-0.319**	0.159	**0.248**	-0.180	-0.097	0.017	-0.091	**-0.244**	0.098	0.046	0.174
injection Cap	0.020	0.045	-0.098	-0.107	-0.069	0.065	0.075	-0.135	0.119	0.181	-0.064	-0.041	-0.056	-0.026	-0.044	-0.074	-0.040	-0.002	-0.022	0.070
Withdr Capacity	-0.011	-0.004	-0.085	-0.156	0.078	0.076	0.034	-0.126	0.110	0.116	0.073	0.106	-0.064	-0.033	-0.100	-0.103	-0.029	0.110	0.034	0.015
Ratio in/out	0.119	-0.037	-0.077	0.016	-0.127	0.019	**0.193**	-0.120	0.028	-0.057	-0.004	-0.007	0.006	-0.007	0.060	0.007	-0.031	-0.111	0.133	-0.051
Peak Flow on Site (MGD)	0.107	-0.037	-0.045	-0.027	-0.004	-0.037	0.141	-0.028	-0.058	0.027	0.007	0.022	-0.034	-0.040	-0.038	-0.027	**0.234**	-0.020	-0.034	-0.065
Depth of well	0.023	0.127	0.019	0.071	-0.136	-0.120	0.061	-0.102	-0.096	-0.095	0.022	0.107	-0.185	**-0.225**	0.108	-0.111	**0.195**	0.064	-0.064	0.125
Injection Horizon	0.105	**0.246**	-0.118	**0.228**	-0.059	-0.124	-0.113	-0.096	0.013	-0.123	**0.232**	**0.252**	-0.018	-0.077	**0.216**	-0.026	0.146	0.057	-0.169	0.019
Depth of Casing	-0.026	0.024	0.081	-0.032	-0.126	-0.076	0.126	-0.069	-0.116	-0.047	-0.090	-0.003	**-0.203**	**-0.220**	0.016	-0.114	0.151	0.045	0.010	0.133
Amount of Water Stored (MG)	0.102	0.186	-0.056	-0.053	-0.034	**0.203**	-0.124	-0.062	0.110	-0.170	**0.455**	**0.494**	-0.032	0.006	-0.060	-0.028	-0.059	0.175	-0.068	0.035

Surface	Reclaimed	Ground	Irrigation	Cooling	Raw	Potable	Confined	Alluvial	Limestone	Sand/ Sandstone	Basalt	Storage Cycles	Injection Cap	Withdr Capacity	Ratio in/out	Peak Flow on Site (MGD)	Depth of well	Injection Horizon	Depth of Casing	Amount of Water Stored (MG)
0.049	-0.026	-0.015	-0.078	0.127	-0.120	0.126	-0.048	0.013	-0.160	0.106	0.122	**0.266**	0.020	-0.011	0.119	0.107	0.023	0.105	-0.026	0.102
0.197	-0.136	-0.118	-0.126	-0.045	-0.066	0.157	-0.072	0.085	**-0.239**	**0.215**	-0.092	-0.167	0.045	-0.004	-0.037	-0.037	0.127	**0.246**	0.024	0.186
-0.087	-0.143	**0.208**	-0.132	-0.047	**-0.212**	**0.293**	**0.256**	**-0.246**	**-0.251**	**0.585**	-0.097	**0.191**	-0.098	-0.085	-0.077	-0.045	0.019	-0.118	0.081	-0.056
0.164	-0.122	-0.091	-0.113	**0.208**	-0.106	0.116	0.080	0.000	**-0.214**	0.124	0.172	-0.041	-0.107	-0.156	0.016	-0.027	0.071	**0.228**	-0.032	-0.053
0.217	-0.114	-0.158	-0.106	-0.038	0.066	0.022	0.059	-0.049	**-0.201**	-0.086	**0.731**	-0.097	-0.069	0.078	-0.127	-0.004	-0.136	-0.059	-0.126	-0.034
-0.139	**0.494**	**-0.258**	**0.332**	-0.062	0.154	**-0.346**	**-0.567**	**0.592**	**-0.328**	**-0.216**	-0.127	0.084	0.065	0.076	0.019	-0.037	-0.120	-0.124	-0.076	**0.203**
-0.408	0.061	**0.398**	0.101	-0.090	0.168	**-0.195**	**0.448**	**-0.428**	**0.880**	**-0.395**	-0.185	-0.104	0.075	0.034	**0.193**	0.141	0.061	-0.113	0.126	-0.124
0.146	-0.114	-0.077	-0.106	-0.038	-0.091	0.164	0.132	**-0.197**	0.166	0.072	-0.078	0.048	-0.135	-0.126	-0.120	-0.028	-0.102	-0.096	-0.069	-0.062
0.045	**0.305**	**-0.268**	**0.237**	0.063	0.105	**-0.272**	**-0.781**	**0.773**	**-0.459**	**-0.243**	-0.177	0.114	0.119	0.110	0.028	-0.058	-0.096	0.013	-0.116	0.110
-0.030	**0.187**	-0.091	**0.219**	0.096	-0.118	-0.067	-0.102	0.112	-0.082	-0.066	0.091	**-0.319**	0.181	0.116	-0.057	0.027	-0.095	-0.123	-0.047	-0.170
0.056	0.137	-0.184	-0.099	-0.020	**0.198**	-0.106	-0.185	**0.273**	-0.148	-0.121	-0.034	0.159	-0.064	0.073	-0.004	0.007	0.022	**0.232**	-0.090	**0.455**
0.014	0.150	-0.145	-0.083	-0.029	0.141	-0.064	-0.175	0.182	-0.092	-0.089	-0.024	**0.248**	-0.041	0.106	-0.007	0.022	0.107	**0.252**	-0.003	**0.494**
0.099	-0.014	-0.107	-0.048	0.019	0.148	-0.106	-0.042	**0.230**	-0.138	-0.084	-0.026	-0.180	-0.056	-0.064	0.006	-0.034	-0.185	-0.018	**-0.203**	-0.032
0.124	0.043	-0.134	-0.060	0.098	0.139	-0.111	**-0.276**	**0.346**	**-0.275**	-0.152	**0.201**	-0.097	-0.026	-0.033	-0.007	-0.040	**-0.225**	-0.077	**-0.220**	0.006
-0.040	0.066	-0.091	0.086	-0.040	-0.106	0.049	**0.219**	**-0.210**	**0.203**	-0.026	0.045	0.017	-0.044	-0.100	0.060	-0.038	0.108	**0.216**	0.016	-0.060
0.051	-0.079	0.003	0.072	-0.026	-0.009	-0.034	-0.060	-0.034	0.064	-0.005	-0.054	-0.091	-0.074	-0.103	0.007	-0.027	-0.111	-0.026	-0.114	-0.028
-0.126	-0.162	**0.267**	-0.150	-0.054	0.115	0.012	0.181	-0.168	**0.270**	-0.055	-0.110	**-0.244**	-0.040	-0.029	-0.031	**0.234**	**0.195**	0.146	0.151	-0.059
0.126	-0.001	-0.147	-0.099	-0.035	0.009	0.068	**-0.275**	**0.288**	**-0.187**	-0.154	0.071	0.098	-0.002	0.110	-0.111	-0.020	0.064	0.057	0.045	0.175
-0.156	**0.197**	0.006	**0.230**	-0.056	-0.019	-0.123	0.087	-0.073	**0.190**	-0.127	-0.015	0.046	-0.022	0.034	0.133	-0.034	-0.064	-0.169	0.010	-0.068
0.054	-0.134	0.078	-0.103	0.045	-0.097	0.144	0.133	**-0.188**	-0.120	**0.349**	-0.054	0.174	0.070	0.015	-0.051	-0.065	0.125	0.019	0.133	0.035
1	**-0.526**	**-0.640**	**-0.431**	0.106	0.015	**0.249**	-0.068	0.079	**-0.298**	0.126	**0.217**	0.061	-0.063	-0.082	-0.032	-0.109	0.068	**0.206**	-0.024	-0.004
-0.526	**1**	**-0.233**	**0.848**	-0.056	-0.134	**-0.436**	**-0.235**	**0.254**	-0.026	-0.186	-0.114	-0.033	0.098	0.073	0.090	-0.038	**-0.190**	-0.168	-0.134	0.124
-0.640	**-0.233**	**1**	**-0.215**	-0.077	0.128	0.051	**0.241**	**-0.267**	**0.343**	-0.003	-0.158	-0.062	-0.014	0.033	-0.051	0.152	0.060	-0.111	0.124	-0.106
-0.431	**0.848**	**-0.215**	**1**	-0.051	**-0.231**	**-0.452**	**-0.232**	**0.249**	0.012	**-0.226**	-0.106	-0.045	0.128	-0.017	0.150	-0.040	**-0.192**	-0.166	-0.138	-0.067
0.106	-0.056	-0.077	-0.051	**1**	-0.082	-0.161	-0.042	0.048	-0.098	-0.081	**0.224**	0.117	-0.025	-0.025	-0.042	-0.006	0.068	0.166	-0.005	-0.030
0.015	-0.134	0.128	**-0.231**	-0.082	**1**	**-0.723**	-0.062	0.086	0.117	**-0.270**	0.066	0.001	0.182	**0.284**	-0.106	0.166	0.061	0.034	0.053	0.122
0.249	**-0.436**	0.051	**-0.452**	-0.161	**-0.723**	**1**	**0.224**	**-0.259**	-0.087	**0.417**	-0.048	-0.002	**-0.244**	**-0.238**	0.006	-0.122	0.057	0.036	0.047	-0.056
-0.068	**-0.235**	**0.241**	**-0.232**	-0.042	-0.062	**0.224**	**1**	**-0.921**	**0.531**	**0.309**	0.132	-0.059	**-0.198**	-0.182	0.011	0.062	0.165	-0.006	**0.192**	**-0.230**
0.079	**0.254**	**-0.267**	**0.249**	0.048	0.086	**-0.259**	**-0.921**	**1**	**-0.510**	**-0.421**	-0.123	0.042	**0.210**	**0.189**	0.017	-0.059	-0.171	0.010	**-0.200**	0.069
-0.298	-0.026	**0.343**	0.012	-0.098	0.117	-0.087	**0.531**	**-0.510**	**1**	**-0.429**	**-0.201**	-0.067	0.014	-0.018	0.164	0.129	0.039	-0.124	0.106	-0.139
0.126	-0.186	-0.003	**-0.226**	-0.081	**-0.270**	**0.417**	**0.309**	**-0.421**	**-0.429**	**1**	-0.166	0.037	**-0.201**	-0.172	-0.174	-0.067	**0.200**	0.115	0.172	0.094
0.217	-0.114	-0.158	-0.106	**0.224**	0.066	-0.048	0.132	-0.123	**-0.201**	-0.166	**1**	-0.029	-0.077	-0.019	-0.057	-0.012	-0.090	0.041	-0.123	-0.044
0.061	-0.033	-0.062	-0.045	0.117	0.001	-0.002	-0.059	0.042	-0.067	0.037	-0.029	**1**	-0.083	-0.030	0.001	-0.063	0.048	0.061	0.025	0.099
-0.063	0.098	-0.014	0.128	-0.025	0.182	**-0.244**	**-0.198**	**0.210**	0.014	**-0.201**	-0.077	-0.083	**1**	**0.778**	0.154	**0.226**	-0.055	-0.037	-0.045	-0.035
-0.082	0.073	0.033	-0.017	-0.025	**0.284**	**-0.238**	-0.182	**0.189**	-0.018	-0.172	-0.019	-0.030	**0.778**	**1**	**-0.250**	0.170	0.080	-0.049	0.116	0.112
-0.032	0.090	-0.051	0.150	-0.042	-0.106	0.006	0.011	0.017	0.164	-0.174	-0.057	0.001	0.154	**-0.250**	**1**	0.027	-0.121	0.145	**-0.210**	-0.099
-0.109	-0.038	0.152	-0.040	-0.006	0.166	-0.122	0.062	-0.059	0.129	-0.067	-0.012	-0.063	**0.226**	0.170	0.027	**1**	0.094	**0.250**	-0.017	0.002
0.068	**-0.190**	0.060	**-0.192**	0.068	0.061	0.057	0.165	-0.171	0.039	**0.200**	-0.090	0.048	-0.055	0.080	-0.121	0.094	**1**	**0.489**	**0.902**	0.021
0.206	-0.168	-0.111	-0.166	0.166	0.034	0.036	-0.006	0.010	-0.124	0.115	0.041	0.061	-0.037	-0.049	0.145	**0.250**	**0.489**	**1**	0.063	-0.003
-0.024	-0.134	0.124	-0.138	-0.005	0.053	0.047	**0.192**	**-0.200**	0.106	0.172	-0.123	0.025	-0.045	0.116	**-0.210**	-0.017	**0.902**	0.063	**1**	0.026
-0.004	0.124	-0.106	-0.067	-0.030	0.122	-0.056	**-0.230**	0.069	-0.139	0.094	-0.044	0.099	-0.035	0.112	-0.099	0.002	0.021	-0.003	0.026	**1**

Table 6. Factor Correlations (All 11)

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
Status Operational	0.000	-0.169	0.163	-0.174	0.014	-0.219	-0.275	-0.267	0.070	0.099	0.155
Midwest/ Central	-0.014	-0.160	-0.033	0.124	0.049	0.361	-0.085	-0.140	0.183	-0.172	-0.230
East	-0.134	-0.159	0.048	-0.296	0.168	-0.063	-0.019	-0.082	-0.250	-0.162	0.020
Rockies	-0.016	-0.179	-0.060	0.146	-0.081	0.217	-0.211	-0.264	-0.030	0.076	-0.024
NW	0.020	-0.086	-0.254	0.203	0.090	-0.348	-0.176	0.186	-0.137	-0.085	-0.016
Southwest	0.318	0.055	0.092	-0.128	-0.003	0.039	0.049	-0.010	-0.166	0.147	-0.018
FL	-0.193	0.385	0.050	0.103	-0.117	-0.004	-0.047	-0.027	0.006	-0.020	0.065
SE	-0.085	-0.074	-0.045	-0.108	-0.151	-0.154	0.284	-0.025	0.308	0.137	-0.193
West	0.360	-0.011	0.109	-0.077	0.030	0.053	0.164	0.122	-0.040	0.125	0.163
Est Start Date	0.067	0.059	-0.139	0.026	0.231	0.128	-0.230	0.095	0.028	-0.121	0.184
Number of wells in the project	0.169	-0.100	0.238	0.150	-0.346	-0.050	-0.072	-0.190	-0.104	-0.288	-0.095
Number of Active wells	0.137	-0.081	0.341	0.063	-0.333	-0.167	-0.082	-0.051	-0.066	-0.236	0.012
Number of inactive or abandoned wells in the project	0.089	-0.053	-0.201	0.208	-0.067	0.253	0.015	-0.328	-0.094	-0.147	-0.246
Clogging	0.183	-0.057	-0.243	0.151	0.076	-0.013	0.053	-0.141	-0.262	0.128	0.043
Metals	-0.081	0.068	-0.062	0.069	-0.147	0.135	-0.212	0.292	0.136	-0.089	-0.021
THMs/ WQ	-0.015	0.023	-0.111	-0.029	-0.116	-0.003	0.314	0.042	0.229	0.135	-0.123
Recovery	-0.109	0.120	0.094	0.202	0.056	0.189	0.114	-0.061	-0.152	0.077	0.302
Water Avail	0.127	-0.068	0.130	0.113	-0.111	-0.068	0.090	0.344	0.002	-0.331	0.153
unknown	-0.010	0.168	-0.038	-0.100	-0.090	-0.048	-0.123	0.047	-0.090	-0.116	-0.210
none noted	-0.107	-0.189	0.189	-0.229	0.137	-0.125	-0.096	-0.209	0.184	0.118	-0.080
surface	0.023	-0.313	-0.096	0.207	-0.040	-0.059	0.080	0.140	0.255	0.097	-0.053
reclaimed	0.208	0.225	-0.004	-0.255	-0.085	0.097	-0.278	0.045	-0.057	-0.033	-0.159
ground	-0.188	0.166	0.111	-0.022	0.142	-0.027	0.167	-0.225	-0.255	-0.066	0.206
irrigation	0.167	0.242	-0.082	-0.279	-0.040	0.133	-0.263	0.087	0.063	0.077	-0.135
cooling	0.029	-0.073	-0.015	0.073	0.018	-0.038	-0.202	-0.032	0.017	0.405	0.086
Raw	0.076	0.129	0.148	0.316	0.088	-0.193	0.147	-0.176	-0.106	0.113	-0.279
Potable	-0.189	-0.259	-0.074	-0.116	-0.057	0.094	0.100	0.109	0.048	-0.263	0.319
Confined	-0.367	0.046	-0.050	0.040	-0.086	-0.071	-0.146	-0.038	-0.082	-0.048	-0.131
Alluvial	0.385	-0.035	0.033	0.013	0.060	0.103	0.104	-0.014	0.035	0.009	0.109
Limestone	-0.225	0.340	0.040	0.081	-0.191	-0.072	0.052	-0.036	0.124	0.030	0.021
Sand/ Sandstone	-0.169	-0.278	0.036	-0.210	0.134	0.152	-0.005	-0.001	-0.119	-0.094	-0.135
Basalt	0.006	-0.106	-0.226	0.212	0.033	-0.349	-0.305	0.120	-0.109	0.100	0.021
Storage Cycles	0.026	-0.096	0.199	-0.146	-0.178	-0.237	-0.031	0.014	-0.030	0.113	-0.041
injection Cap	0.099	0.133	0.116	0.063	0.372	-0.066	-0.098	-0.107	0.407	-0.179	0.016
Withdr Capacity	0.096	0.101	0.194	0.129	0.392	-0.157	-0.039	-0.006	0.217	-0.258	-0.140
Ratio in/out	0.011	0.088	-0.047	-0.058	-0.215	0.068	-0.156	-0.148	0.211	0.077	0.301
Peak Flow on Site (MGD)	-0.028	0.087	0.133	0.148	0.112	-0.032	-0.070	-0.156	0.098	0.026	0.234
Depth of well	-0.127	-0.074	0.369	0.149	0.082	0.208	-0.102	0.262	-0.091	0.212	-0.115
Injection Horizon	-0.018	-0.143	0.212	0.223	-0.142	0.185	-0.197	0.021	0.140	0.162	0.147
Depth of Casing	-0.136	-0.013	0.317	0.059	0.164	0.146	-0.019	0.290	-0.173	0.162	-0.205

Table 7. Percent Contribution to the Factor

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
Status Operational	0	2.863	2.66	3.023	0.02	4.801	7.586	7.132	0.486	0.988	2.403
Midwest/ Central	0.02	2.56	0.109	1.543	0.237	13.054	0.731	1.956	3.336	2.955	5.273
East	1.784	2.522	0.23	8.762	2.819	0.398	0.038	0.677	6.253	2.622	0.039
Rockies	0.025	3.195	0.361	2.131	0.664	4.7	4.435	6.968	0.089	0.581	0.059
NW	0.039	0.732	6.457	4.117	0.817	12.113	3.11	3.456	1.882	0.726	0.025
Southwest	10.093	0.302	0.845	1.639	0.001	0.149	0.243	0.01	2.744	2.154	0.033
FL	3.712	14.795	0.255	1.068	1.376	0.002	0.217	0.074	0.004	0.039	0.416
SE	0.715	0.547	0.202	1.176	2.292	2.367	8.064	0.064	9.517	1.882	3.714
West	12.965	0.012	1.19	0.591	0.093	0.281	2.698	1.481	0.159	1.558	2.669
Est Start Date	0.452	0.354	1.943	0.066	5.358	1.647	5.31	0.901	0.077	1.471	3.371
Number of wells in the project	2.855	0.999	5.666	2.238	11.995	0.249	0.514	3.618	1.073	8.29	0.907
Number of Active wells	1.883	0.657	11.632	0.396	11.11	2.785	0.673	0.26	0.438	5.556	0.013
Number of inactive or abandoned wells in the project	0.783	0.278	4.055	4.316	0.444	6.393	0.022	10.767	0.881	2.154	6.055
Clogging	3.337	0.33	5.913	2.266	0.578	0.017	0.285	1.998	6.869	1.644	0.187
Metals	0.661	0.457	0.381	0.483	2.171	1.827	4.505	8.499	1.857	0.8	0.043
THMs/ WQ	0.022	0.054	1.236	0.087	1.352	0.001	9.833	0.173	5.254	1.815	1.507
Recovery	1.192	1.441	0.883	4.075	0.319	3.573	1.292	0.369	2.309	0.593	9.127
Water Avail	1.606	0.457	1.694	1.274	1.236	0.459	0.818	11.8	0	10.942	2.327
unknown	0.011	2.811	0.147	1.004	0.814	0.227	1.512	0.226	0.809	1.352	4.395
none noted	1.153	3.591	3.557	5.23	1.89	1.564	0.929	4.379	3.394	1.381	0.637
surface	0.051	9.799	0.918	4.281	0.16	0.35	0.635	1.965	6.512	0.948	0.277
reclaimed	4.332	5.083	0.002	6.517	0.715	0.932	7.748	0.202	0.328	0.109	2.52
ground	3.525	2.747	1.221	0.049	2.003	0.071	2.773	5.074	6.507	0.439	4.263
irrigation	2.788	5.837	0.669	7.796	0.157	1.782	6.923	0.756	0.397	0.596	1.817
cooling	0.083	0.539	0.021	0.532	0.033	0.145	4.061	0.103	0.028	16.432	0.742
Raw	0.578	1.658	2.177	9.96	0.768	3.722	2.151	3.101	1.119	1.279	7.79
Potable	3.566	6.711	0.547	1.346	0.327	0.888	0.99	1.184	0.234	6.941	10.189
Confined	13.472	0.213	0.253	0.157	0.735	0.503	2.144	0.146	0.675	0.228	1.729
Alluvial	14.825	0.119	0.106	0.018	0.365	1.057	1.075	0.019	0.125	0.008	1.196
Limestone	5.065	11.582	0.161	0.654	3.648	0.52	0.27	0.127	1.549	0.087	0.043
Sand/ Sandstone	2.848	7.743	0.129	4.429	1.8	2.313	0.002	0	1.413	0.889	1.83
Basalt	0.004	1.133	5.125	4.488	0.11	12.198	9.31	1.435	1.18	1.003	0.046
Storage Cycles	0.069	0.93	3.967	2.136	3.157	5.6	0.095	0.019	0.089	1.278	0.169
injection Cap	0.977	1.775	1.345	0.396	13.823	0.433	0.964	1.139	16.537	3.191	0.027
Withdr Capacity	0.925	1.025	3.762	1.675	15.366	2.467	0.155	0.004	4.708	6.636	1.951
Ratio in/out	0.012	0.779	0.225	0.34	4.641	0.457	2.438	2.181	4.432	0.591	9.04
Peak Flow on Site (MGD)	0.079	0.751	1.76	2.197	1.246	0.1	0.495	2.42	0.957	0.066	5.496
Depth of well	1.607	0.544	13.636	2.206	0.665	4.309	1.04	6.884	0.82	4.508	1.332
Injection Horizon	0.032	2.057	4.495	4.991	2.018	3.423	3.88	0.045	1.964	2.633	2.15
Depth of Casing	1.851	0.018	10.066	0.35	2.679	2.123	0.036	8.388	2.994	2.637	4.191

Table 8. Factors after Varimax Rotation

Component score coefficients after Varimax rotation:	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11
Status Operational	-0.004	0.026	-0.058	-0.019	0.055	0.053	0.376	0.069	0.024	0.027	0.017
Midwest/ Central	-0.054	-0.211	-0.106	0.055	0.022	-0.131	-0.116	-0.026	0.207	0.126	-0.006
East	-0.007	-0.238	-0.022	-0.005	-0.039	-0.020	0.051	0.210	-0.057	-0.041	-0.013
Rockies	-0.018	0.067	0.059	-0.068	0.058	0.010	0.147	0.007	-0.021	-0.089	0.040
NW	-0.059	-0.045	-0.041	0.027	0.019	0.367	-0.070	0.023	-0.032	0.040	-0.014
Southwest	0.136	-0.036	0.060	0.150	-0.011	-0.036	0.029	0.000	-0.062	-0.080	0.096
FL	-0.114	0.183	-0.019	0.012	0.026	-0.048	-0.047	0.108	0.032	0.003	0.020
SE	-0.075	0.069	-0.022	0.050	0.003	-0.091	0.074	-0.420	-0.154	-0.006	-0.003
West	0.237	0.064	0.062	-0.031	-0.042	-0.063	-0.002	0.010	-0.001	-0.038	-0.032
Est Start Date	0.018	0.064	0.096	-0.292	-0.021	0.132	0.030	0.036	-0.126	0.150	0.132
Number of wells in the project	-0.043	0.029	-0.040	0.001	0.361	-0.010	0.001	0.020	-0.036	-0.040	0.011
Number of Active wells	-0.034	0.059	0.014	-0.068	0.362	0.022	0.045	0.010	-0.031	-0.010	0.019
Number of inactive or abandoned wells in the project	-0.025	-0.060	-0.123	0.150	0.039	-0.071	-0.094	0.025	-0.015	-0.069	-0.017
Clogging	0.110	0.005	-0.064	0.171	-0.082	0.105	-0.002	0.118	-0.036	-0.143	-0.074
Metals	-0.075	-0.125	-0.003	0.028	-0.031	0.060	-0.181	-0.014	0.383	0.042	0.094
THMs/ WQ	0.010	0.068	-0.015	0.016	-0.056	-0.097	-0.040	-0.263	-0.012	-0.036	-0.002
Recovery	0.000	0.169	0.177	-0.066	-0.013	-0.061	-0.039	0.049	-0.162	-0.010	-0.069
Water Avail	0.068	0.081	0.041	-0.298	0.194	0.059	-0.186	0.038	-0.019	0.060	-0.091
unknown	-0.055	-0.020	-0.061	0.029	0.024	0.015	-0.085	0.092	-0.001	-0.075	0.095
none noted	-0.053	-0.080	0.001	0.002	0.000	-0.064	0.298	-0.133	-0.045	0.124	0.022
surface	0.047	0.016	0.006	-0.017	-0.034	0.064	-0.036	-0.256	0.081	0.010	-0.150
reclaimed	-0.033	-0.053	0.029	-0.025	0.067	0.021	0.018	-0.023	-0.015	-0.007	0.353
ground	-0.016	0.035	-0.031	0.018	-0.022	-0.095	0.035	0.319	-0.097	0.000	-0.133
irrigation	-0.014	-0.031	0.020	-0.042	-0.048	-0.002	0.022	-0.074	0.027	-0.014	0.347
cooling	0.058	0.143	0.123	0.004	-0.073	0.122	0.233	-0.013	0.065	-0.091	0.041
Raw	-0.018	-0.018	0.030	0.412	0.021	0.028	-0.035	-0.009	-0.006	0.069	-0.098
Potable	0.009	-0.003	-0.076	-0.345	0.035	-0.058	-0.050	0.062	-0.032	-0.026	-0.162
Confined	-0.216	-0.019	0.021	0.049	0.025	0.064	0.001	-0.025	-0.023	-0.045	0.037
Alluvial	0.215	0.044	-0.037	-0.039	-0.020	-0.068	-0.012	0.043	0.010	0.016	-0.041
Limestone	-0.136	0.231	-0.035	0.015	0.035	-0.087	-0.017	-0.064	-0.030	-0.009	-0.006
Sand/ Sandstone	-0.054	-0.314	0.075	0.006	-0.026	-0.050	-0.010	0.019	0.010	0.002	0.034
Basalt	-0.053	0.040	0.007	0.022	0.012	0.396	0.070	0.017	0.018	-0.017	0.023
Storage Cycles	0.043	-0.014	-0.026	0.116	0.036	-0.012	0.113	0.038	0.167	-0.130	-0.046
injection Cap	0.013	-0.007	-0.076	-0.016	-0.065	-0.036	0.072	-0.010	0.104	0.414	-0.016
Withdr Capacity	-0.020	-0.069	0.007	0.040	0.022	0.032	-0.010	-0.023	-0.039	0.404	-0.017
Ratio in/out	0.047	0.171	-0.167	-0.089	-0.035	-0.080	0.112	0.133	0.308	-0.070	-0.037
Peak Flow on Site (MGD)	-0.002	0.127	0.024	-0.004	0.000	-0.008	0.102	0.090	0.066	0.132	-0.073
Depth of well	0.021	-0.019	0.375	0.045	-0.017	-0.005	-0.007	-0.010	0.110	-0.032	0.016
Injection Horizon	0.022	0.046	0.119	0.022	0.007	-0.008	0.065	0.007	0.404	0.001	-0.037
Depth of Casing	0.013	-0.046	0.371	0.040	-0.023	-0.001	-0.042	-0.015	-0.086	-0.037	0.038
Amount of Water Stored (MG)	-0.054	-0.126	-0.026	0.052	0.248	-0.024	-0.068	-0.013	-0.038	0.044	0.029

Table 9. Linear Regression Model Parameters (all compete data for Active and inactive sites only)

Source	Value	Standard error	t	Pr > |t|	Lower bound (95%)	Upper bound (95%)
Intercept	-24.574	9.305	-2.641	0.010	-43.114	-6.035
Midwest/ Central	0.559	0.407	1.372	0.174	-0.253	1.370
East	0.815	0.433	1.882	0.064	-0.048	1.678
Rockies	0.246	0.371	0.665	0.508	-0.492	0.985
NW	0.015	0.459	0.032	0.975	-0.900	0.929
Southwest	-0.151	0.205	-0.738	0.463	-0.560	0.257
FL	0.684	0.504	1.358	0.179	-0.320	1.688
SE	0.645	0.474	1.360	0.178	-0.300	1.590
West	0.452	0.494	0.914	0.364	-0.533	1.436
Est Start Date	0.012	0.005	2.620	0.011	0.003	0.022
Number of wells in the project	-0.025	0.015	-1.612	0.111	-0.055	0.006
Number of Active wells	0.043	0.016	2.667	0.009	0.011	0.075
Clogging	0.189	0.195	0.970	0.335	-0.199	0.578
Metals	-0.020	0.221	-0.090	0.928	-0.461	0.421
THMs/ WQ	-0.222	0.272	-0.816	0.417	-0.763	0.319
Recovery	-0.115	0.184	-0.626	0.533	-0.481	0.251
Water Avail	-0.727	0.233	-3.117	0.003	-1.192	-0.262
unknown	0.064	0.192	0.335	0.739	-0.318	0.446
none noted	0.496	0.175	2.839	0.006	0.148	0.843
surface	-0.016	0.232	-0.067	0.946	-0.479	0.447
reclaimed	0.057	0.354	0.161	0.873	-0.648	0.762
ground	-0.109	0.230	-0.477	0.635	-0.567	0.348
irrigation	-0.284	0.288	-0.985	0.328	-0.858	0.290
cooling	-0.096	0.312	-0.308	0.759	-0.719	0.526
Raw	-0.147	0.103	-1.420	0.160	-0.353	0.059
Confined	-0.161	0.254	-0.635	0.527	-0.667	0.345
Alluvial	0.073	0.484	0.151	0.880	-0.891	1.038
Limestone	-0.201	0.488	-0.412	0.681	-1.173	0.771
Sand/ Sandstone	-0.171	0.460	-0.371	0.712	-1.086	0.745
Basalt	0.617	0.452	1.365	0.176	-0.284	1.518
Storage Cycles	0.005	0.004	1.303	0.197	-0.003	0.013
injection Cap	-0.068	0.049	-1.383	0.171	-0.167	0.030
Withdr Capacity	0.035	0.043	0.828	0.410	-0.050	0.120
Ratio in/out	0.149	0.089	1.676	0.098	-0.028	0.325
Peak Flow on Site (MGD)	0.002	0.001	2.615	0.011	0.000	0.003
Depth of well	0.000	0.000	0.096	0.924	0.000	0.000
Injection Horizon	0.000	0.000	-0.171	0.865	-0.001	0.001
Depth of Casing	0.000	0.000				

Table 10. Variables for the Active and Inactive Sites that also exist for the Test and Study Sites used in the revised Linear Regression model

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Status Operational	110	0	110	0.000	1.000	0.527	0.502
Midwest/ Central	110	0	110	0.000	1.000	0.100	0.301
East	110	0	110	0.000	1.000	0.100	0.301
Rockies	110	0	110	0.000	1.000	0.082	0.275
NW	110	0	110	0.000	1.000	0.073	0.261
SE	110	0	110	0.000	1.000	0.382	0.488
West	110	0	110	0.000	1.000	0.291	0.456
Number of wells in the project	110	0	110	0.000	87.000	4.982	9.968
Number of Active wells	110	0	110	0.000	87.000	3.573	9.501
Number of inactive or abandoned wells in the project	110	0	110	0.000	40.000	1.409	4.303
Metals	110	0	110	0.000	1.000	0.082	0.275
surface	110	0	110	0.000	1.000	0.618	0.488
reclaimed	110	0	110	0.000	1.000	0.145	0.354
ground	110	0	110	0.000	1.000	0.245	0.432
irrigation	110	0	110	0.000	1.000	0.127	0.335
Raw	110	0	110	0.000	1.000	0.273	0.447
Potable	110	0	110	0.000	1.000	0.582	0.496
Confined	110	0	110	0.000	1.000	0.636	0.483
Alluvial	110	0	110	0.000	1.000	0.336	0.475
Limestone	110	0	110	0.000	1.000	0.345	0.478
Sand/ Sandstone	110	0	110	0.000	1.000	0.255	0.438
Basalt	110	0	110	0.000	1.000	0.073	0.261
injection Cap	110	0	110	0.100	10.000	1.418	1.700
Withdr Capacity	110	0	110	0.200	10.000	2.004	2.012
Depth of well	110	0	110	40.000	2523.000	780.791	464.856
Injection Horizon	110	0	110	0.000	1186.000	190.764	201.374
Depth of Casing	110	0	110	15.000	2185.000	590.027	406.546

Table 11. Linear Regression Weights for Use in the Predictive Model

Intercept	-0.307
Midwest/ Central	0.986
East	1.416
Rockies	0.797
NW	0.626
SE	0.949
West	0.763
Number of wells in the project	-0.039
Number of Active wells	0.053
Number of inactive or abandoned wells in the project	0.000
Metals	-0.278
surface	0.190
reclaimed	0.368
ground	0.078
irrigation	-0.580
Raw	-0.439
Potable	-0.359
Confined	-0.170
Alluvial	0.176
Limestone	0.226
Sand/ Sandstone	0.186
Basalt	0.704
injection Cap	0.051
Withdr Capacity	-0.034
Depth of well	0.000
Injection Horizon	0.000
Depth of Casing	0.000

Table 12. Summary of prediction for Success for study or Test wells

State	Predicted	State	Predicted
AZ	0.49	FL	0.227
CA	0.489	IA	0.789
CA	0.559	IA	0.643
CA	0.521	KS	0.284
CA	0.085	NC	0.893
CA	0.531	NJ	1.058
CA	0.754	NJ	1.037
CA	0.559	NJ	1.067
CA	0.288	NJ	1.021
CA	0.443	NJ	0.955
FL	0.633	NJ	1.213
FL	0.878	NJ	1.067
FL	1.014	NV	0.421
FL	0.352	OK	1.306
FL	0.59	OR	0.815
FL	0.591	OR	0.995
FL	0.555	OR	1.111
FL	0.67	SC	0.481
FL	0.583	SC	1.042
FL	1.071	SC	0.751
FL	0.891	SC	0.606
FL	0.364	UT	1.33
FL	0.629	WA	0.635
FL	0.655	WA	0.622
FL	0.937	WA	0.784
FL	0.227	WY	0.403
		WY	0.419

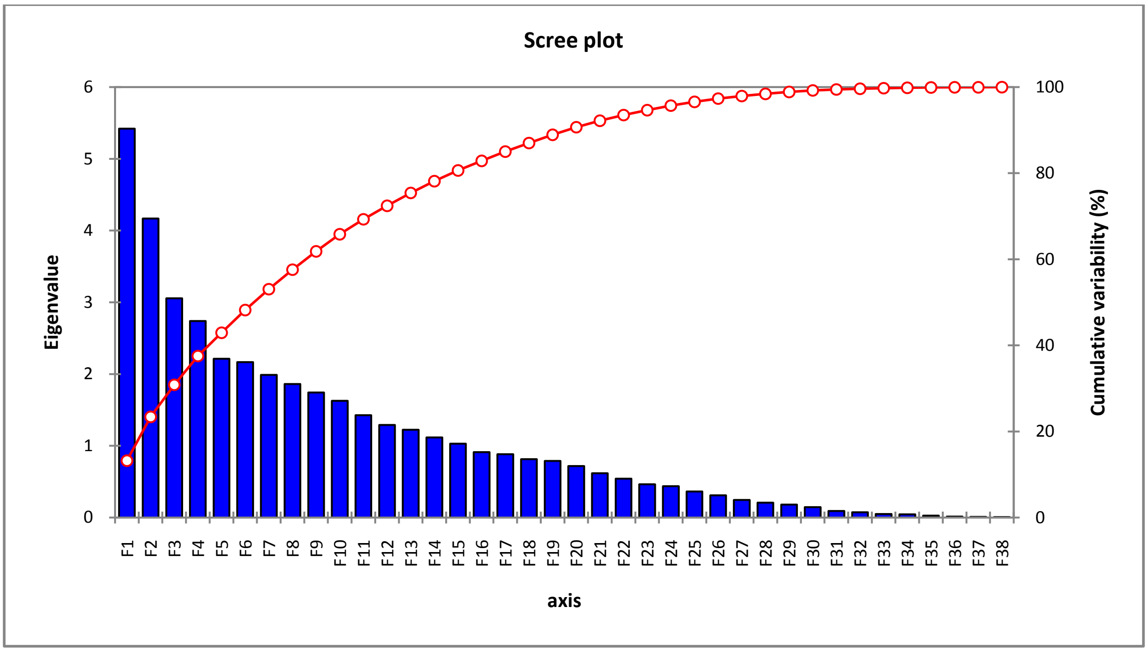

Figure 2. Scree Plot showing that 11 factors are needed to get 70 percent of variance

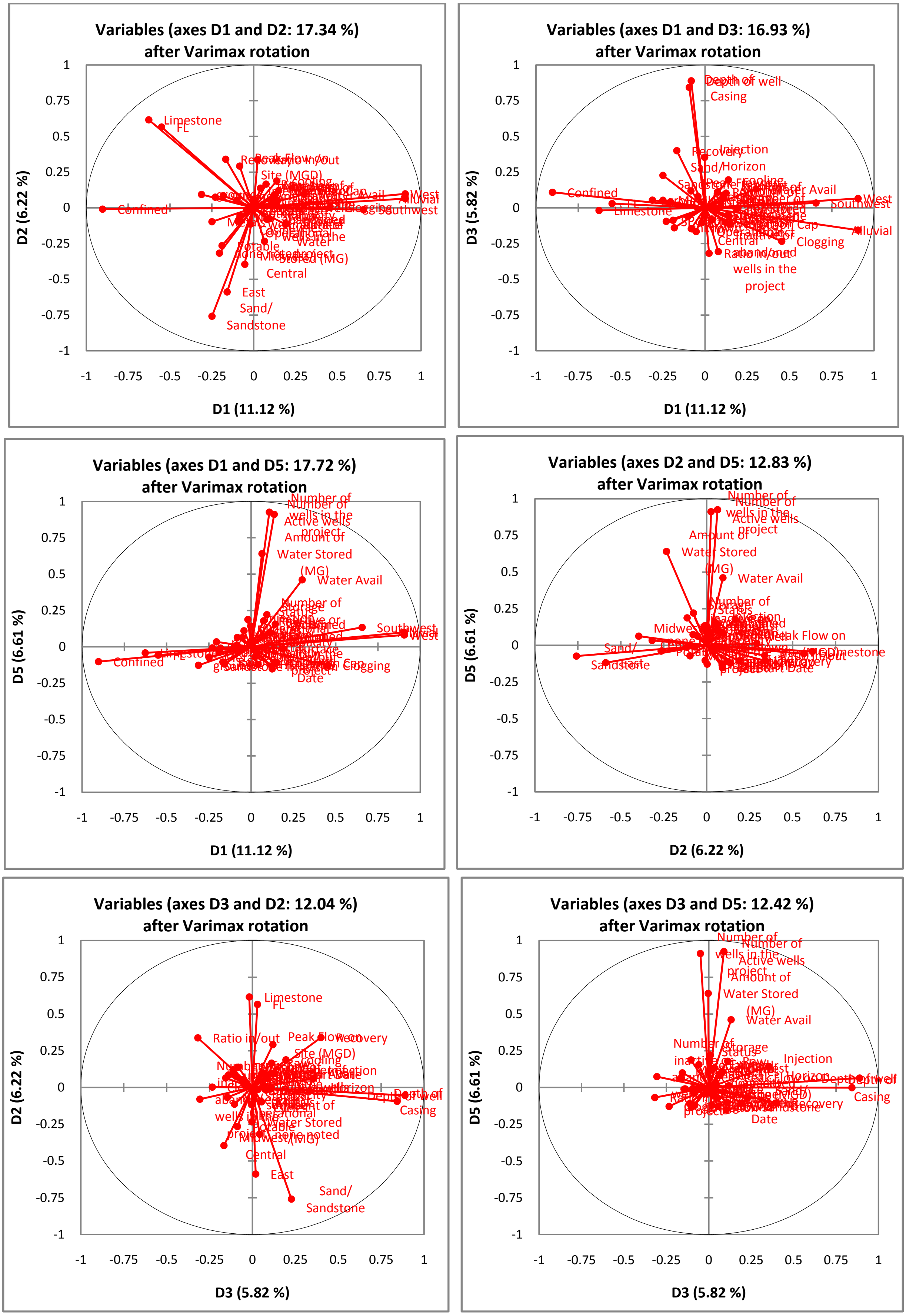

Figure 3. Varimax Plots of Factors from PCA Analysis

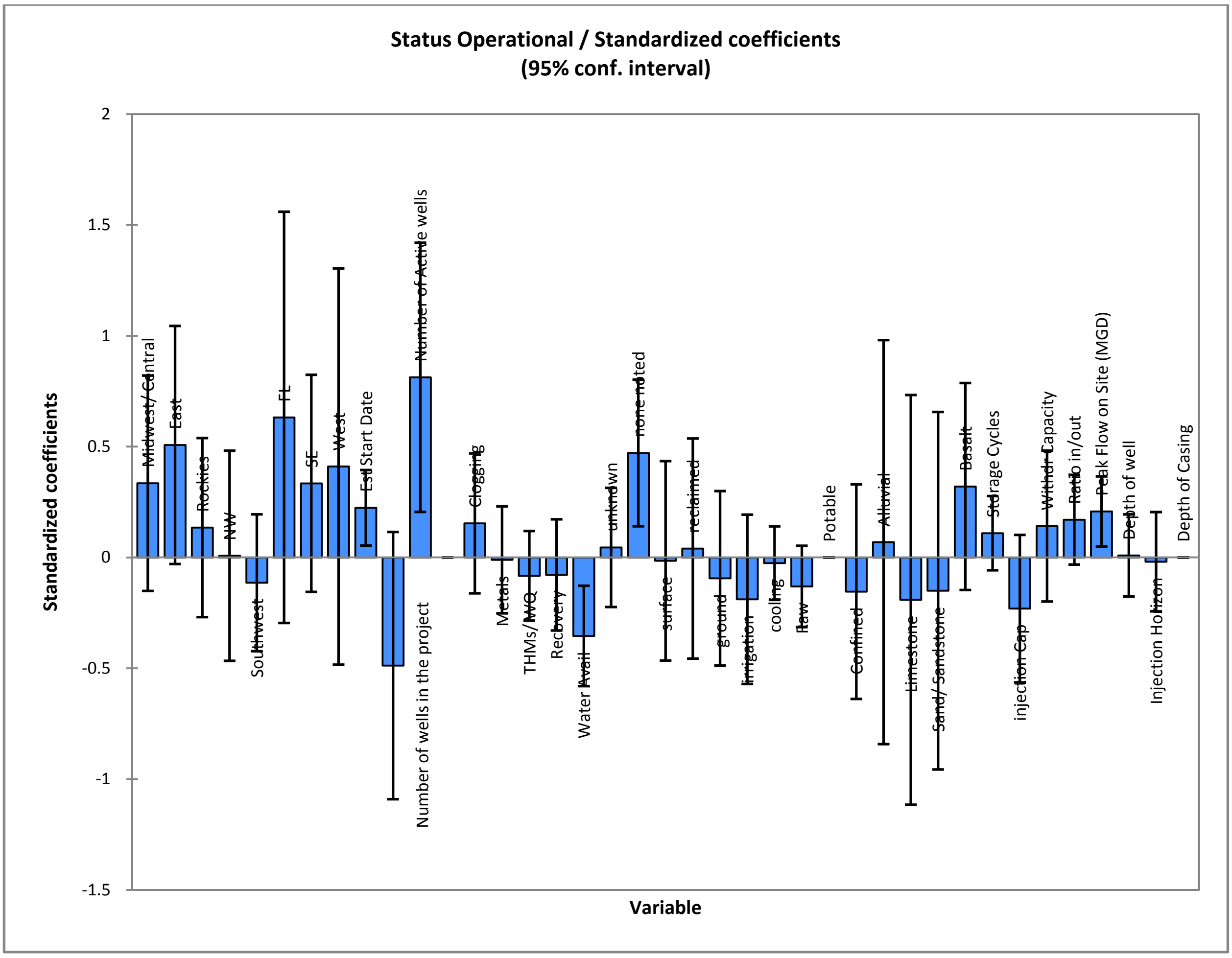

Figure 4. Perspective on Linear Regression Model Parameters

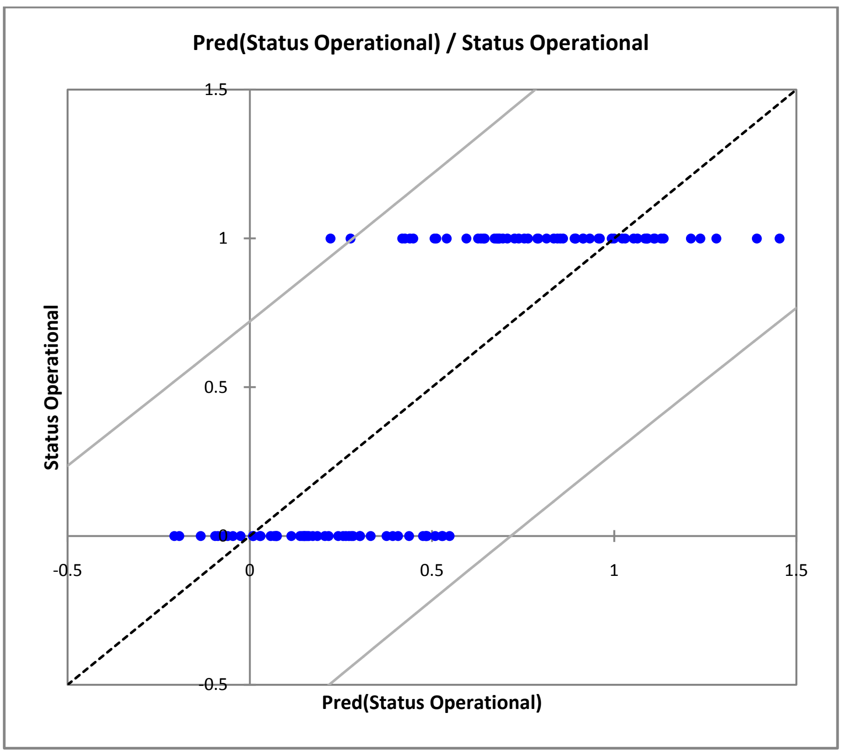

Figure 5. Correlation between predicted and actual successful wells

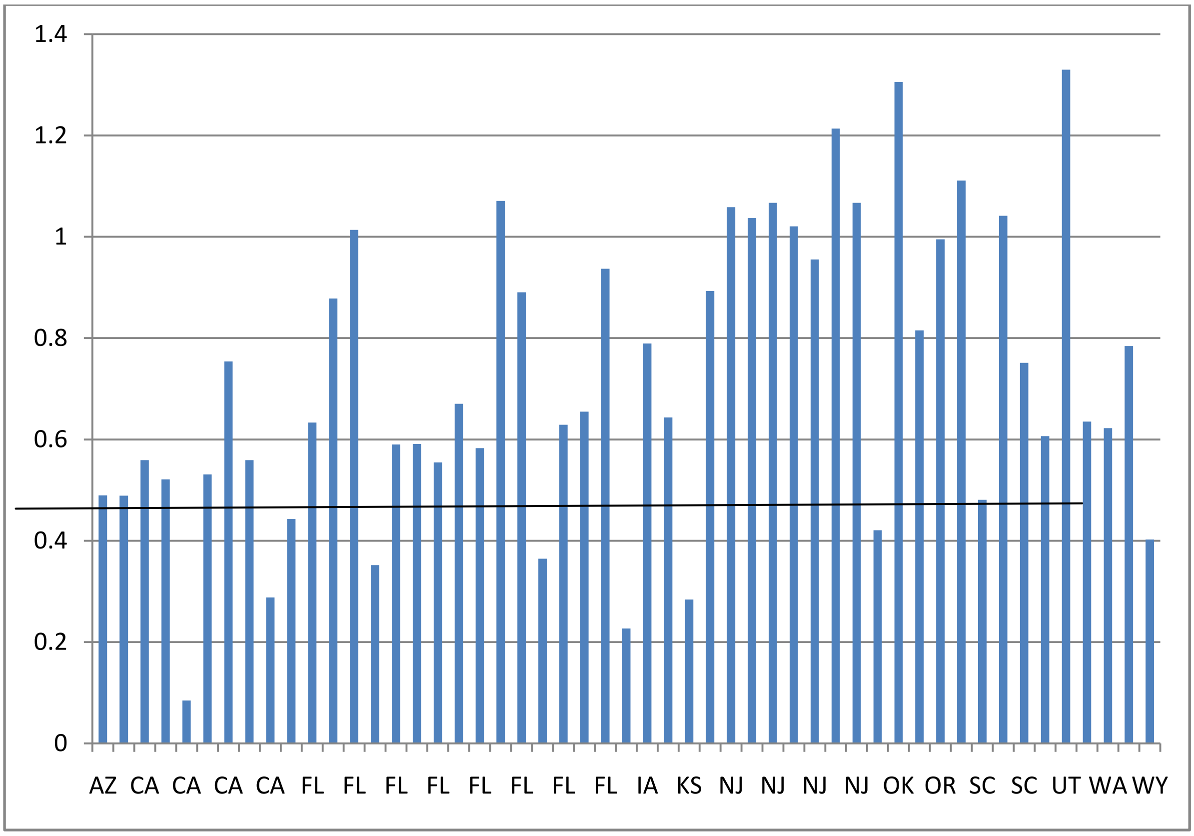

Figure 6. Prediction on success of Test wells

4. Conclusions

The goal of this paper was to use this data and apply the results from active or inactive wells to those currently in the test of study phase in an effort to determine if there was a means to predict their likely success. This paper builds on the 2013 a nationwide survey of ASR systems as discussed in Bloetscher, et al [3,4] and AWWA [2]. The data from the 2013 was analyzed using factor and principal component analysis to determine correlations and variance combinations on the data. The goal was to determine which factors correlated best as a means to determine if a useful analysis could be developed to predict success of ASR systems currently in the test phase based on the success of active ASR sites.

The results indicate that the use of PCA and linear regression can be used to project the potential for the test and study sites. Two thirds of the current 204 ASR sites are either active or inactive, and, once the data was sorted, 111 of those sites were able to be used to project future status. While the actual results may not be known for many years, the results shed light on the over 50 sites in this stage and their likelihood for success. The data suggests that about 1/3 of the wells have low likelihood for success and perhaps should not be pursued further.

Several caveats exist for this analysis. First, the regional locations ignore that geological differences can be very different between nearby sites. Some effort was made to address this issue – for example Florida (mostly limestone) was separated from the rest of the southeast that was not. Information on salinity in the injection formation would be useful as a number of people, including the author, believe this is a major barrier to success. However, the results also suggest that more complete information would be useful for further analysis. Many sites lack full information, especially those in the study phase or are prior to 1990.

REFERENCES

[1] American Water Works Association (2014). *M21 Groundwater*, AWWA, Denver, CO.

[2] American Water Works Association (2015). *M63 Aquifer Storage and Recovery*, AWWA, Denver, CO.

[3] Bloetscher, F.; Sham, Chi Ho, Ratick, S. and Danko III, J.J. 2015. Status of Aquifer Storage and Recovery In The United States – 2013, *British Journal of Science* 70 April 2015, Vol. 12 (2).

[4] Bloetscher, F.; Sham, C.H.; Danko, J.J.; and Ratick, S. (2014). Lessons Learned from Aquifer Storage and Recovery (ASR) Systems in the United States, *Journal of Water Research*, 2014, 6, 1603-1629. http://www.scirp.org/journal/jwarp http://dx.doi.org/10.4236/jwarp.2014.617146.

[5] American Water Works Association (2002). *Survey and*

Analysis of Aquifer Storage and Recovery (ASR) Systems and Associated Regulatory Programs in the United States, AWWA, Denver, CO.

[6] Bloetscher, F.; Muniz, A; and Gonzalez, A.M. (2002). Summary of AWWA Survey of ASR Practices in the United States, *Florida Section AWWA Annual Conference Proceedings*. Hazen and Sawyer, Hollywood, FL.

[7] Missimer, T. M. and Maliva, R. (2010). *Aquifer Storage and Recovery and Managed Aquifer Recharge Using Wells: Planning, Hydrogeology, Design, and Operation*, Schlumberger, Houston, TX.

[8] Pyne, R. D. G (1995). *Groundwater Recharge and Wells: A Guide to Aquifer Storage Recovery, 1st Ed.* , CRC Press, Boca Raton, FL.

[9] Pyne, R. D. G (2005). *Groundwater Recharge and Wells: A Guide to Aquifer Storage Recovery, 2nd Ed*, CRC Press, Gainesville, FL.

[10] Spearman, C. (1904). The Proof and Measurement of Association between Two Rings. *The American Journal of Psychology*, 15, 72-101. http://dx.doi.org/10.2307/1412159.

[11] Bloetscher, F. (2018). Risk and Economic Development in the Provision of Public Infrastructure, *Journal of Environmental Protection*, Vol.09 No.09(2018), Article ID:86955,18 pages. 10.4236/jep.2018.99061.

[12] Johnson, R.A. and Wichern, D.W., (1992). *Applied Multivariate Statistical Analysis*, Prentice-Hall, Englewood Cliffs, NJ.

[13] Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine*. 2 (11): 559–572. doi:10.1080/14786440109462720.

[14] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 417–441, and 498–520.

[15] Hotelling, H (1936). Relations between two sets of variates. *Biometrika*. 28 (3/4): 321–377. *doi:10.2307/2333955*.

[16] Jolliffe I.T. *Principal Component Analysis*, Series: Springer Series in Statistics, 2nd ed., Springer, NY, 2002, XXIX, 487 p. 28 illus. ISBN 978-0-387-95442-4.

[17] Pleitez Herrera, Fernando J. (2012). Predicting Removal Efficiency of Reverse Osmosis with Respect to endocrine disruptors and pharmaceuticals and personal care products. A Thesis Submitted to the Faculty of College of Engineering and Computer Science in Partial Fulfillment of the Requirements for the Degree of Master of Science, Florida Atlantic University, Boca Raton, FL.

[18] IOS (2012). Creating a Cree Plot. *Improved Outcomes Software*. http://www.improvedoutcomes.com/docs/WebSiteDocs/PCA/Creating_a_Scree_Plot.htm.

[19] Cattell R. B. (1966). The Scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245–276. http://dx.doi.org/10.1207/s15327906mbr0102_10.

[20] David A. Freedman (2009). *Statistical Models: Theory and Practice*. Cambridge University Press. p. 26.

[21] Rencher, Alvin C.; Christensen, William F. (2012), "Chapter 10, Multivariate regression – Section 10.1, Introduction", *Methods of Multivariate Analysis*, Wiley Series in Probability and Statistics, 709 (3rd ed.), John Wiley & Sons, p. 19, ISBN 9781118391679.