# VNMN Project Analysis

Tisha Hooks and Corey Smith, Winona State University Statistical Consulting Center

May 28, 2013

The database used for this analysis consisted of 675 wells across southeastern Minnesota. The nitrate levels (NO3) of each well were potentially measured at each of the following time periods:

- Time 1 – February 2008
- Time 2 – August 2008
- Time 3 – February 2009
- Time 4 – August 2009
- Time 5 – August 2010
- Time 6 – August 2011
- Time 7 – August 2012

The original data set also included a Time 0; however, this time period did not control for date and included an overwhelming number of missing values. So, for this analysis, we consider only Time Periods 1 through 7. Note that even for Times 1-7, some well samples were not submitted. Table 1 summarizes the sampling coverage by county. For example, the database includes 56 wells from Dodge County. Of these 56 wells, only 30 were sampled in Round 1, only 31 were sampled in Rounds 2 and 3, etc. Table 2 contains the same information, with the results expressed as percentages rather than counts.

*Table 1. Number of Wells Sampled in each Round by County.*

| County | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Dodge (n=56) | 30 | 31 | 31 | 29 | 29 | 28 | 27 |
| Fillmore (n=96) | 95 | 94 | 91 | 91 | 73 | 73 | 64 |
| Goodhue (n=92) | 72 | 72 | 64 | 65 | 47 | 52 | 53 |
| Houston (n=66) | 48 | 36 | 39 | 38 | 32 | 39 | 36 |
| Mower (n=88) | 58 | 55 | 51 | 51 | 43 | 40 | 36 |
| Olmsted (n=72) | 52 | 62 | 65 | 66 | 65 | 63 | 63 |
| Rice (n=67) | 63 | 63 | 60 | 52 | 53 | 50 | 51 |
| Wabasha (n=64) | 44 | 37 | 34 | 21 | 28 | 27 | 22 |
| Winona (n=74) | 57 | 60 | 59 | 58 | 50 | 53 | 54 |
| Total (n=675) | 519 | 510 | 494 | 471 | 420 | 425 | 406 |

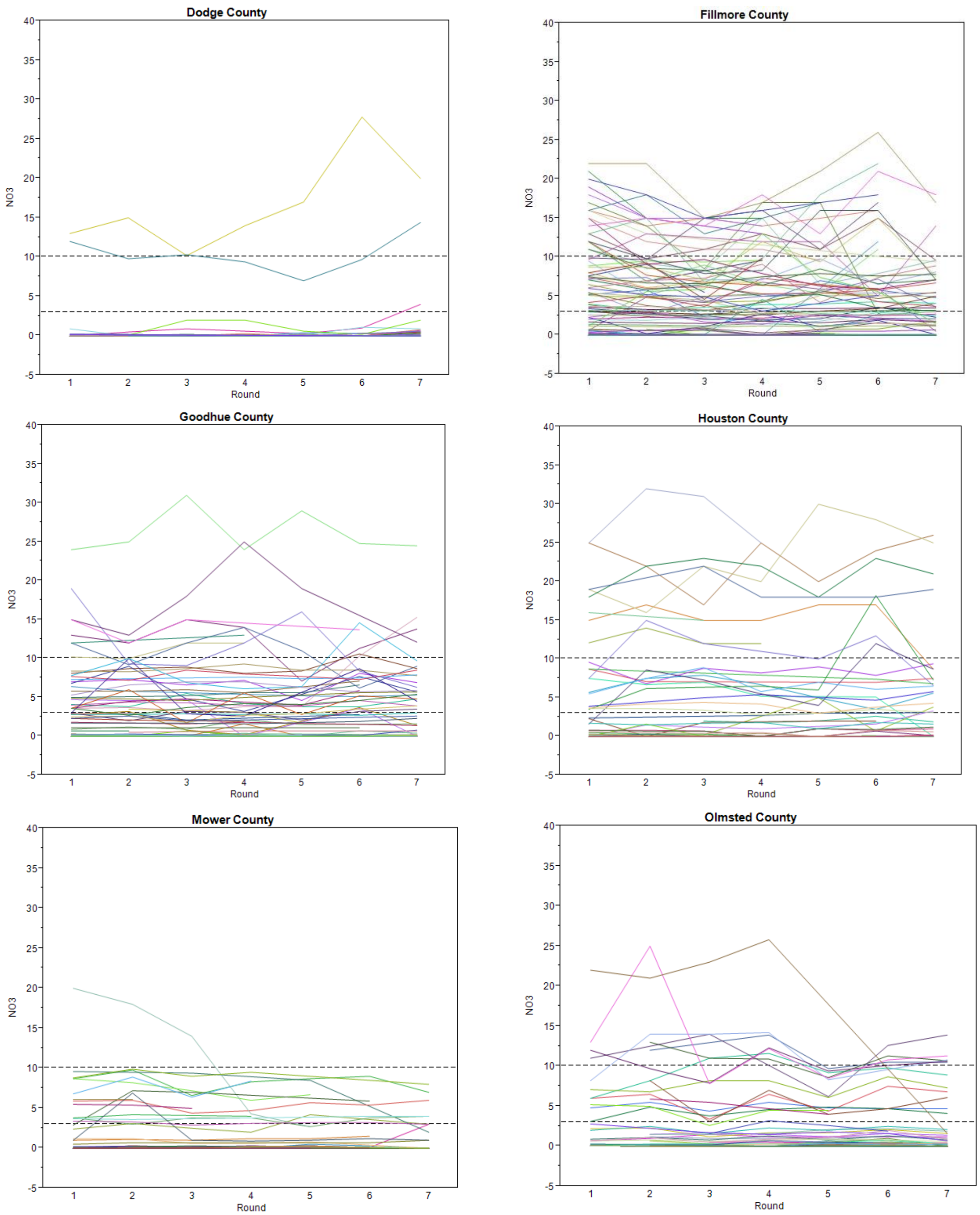
*Table 2. Percentage of Wells Sampled in each Round by County.*

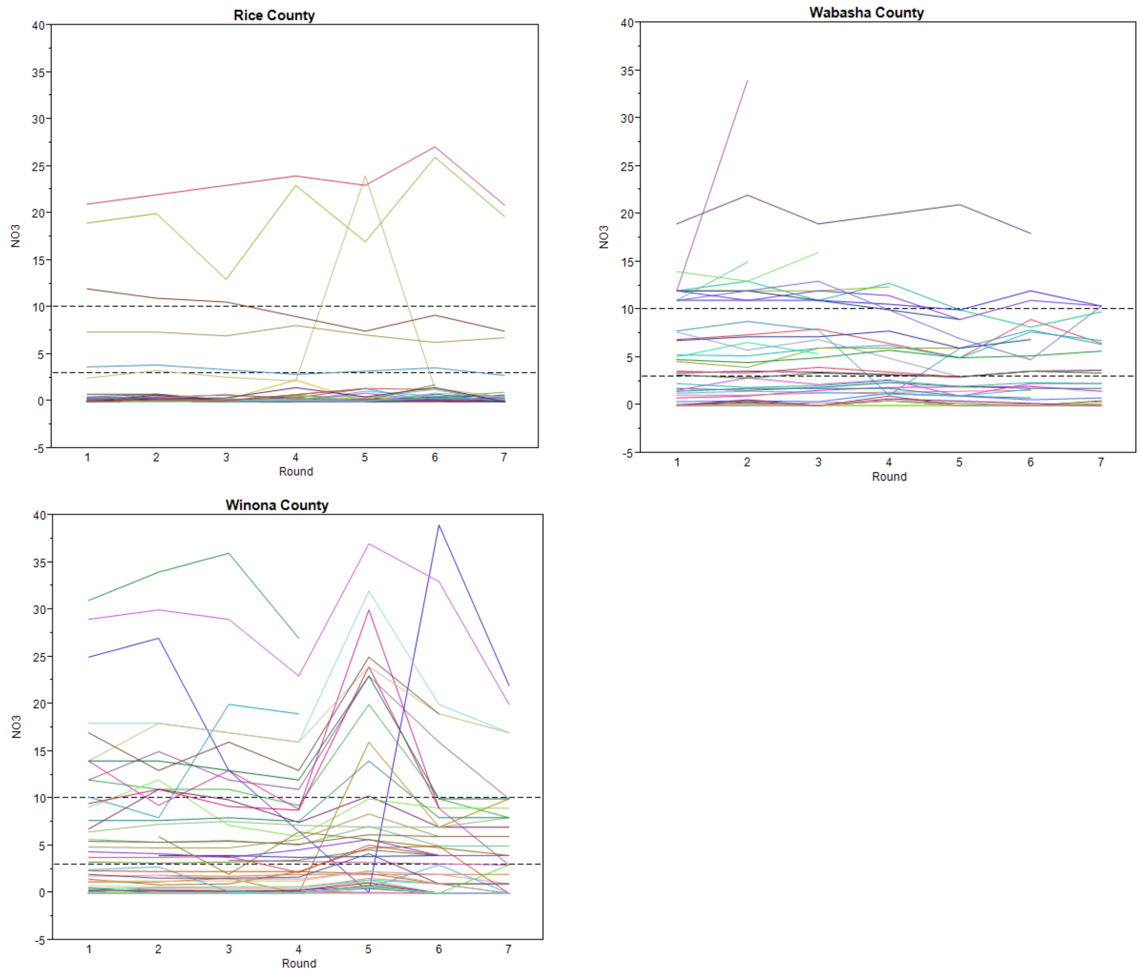
| County | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Dodge (n=56) | 54% | 55% | 55% | 52% | 52% | 50% | 48% |
| Fillmore (n=96) | 99% | 98% | 95% | 95% | 76% | 76% | 67% |
| Goodhue (n=92) | 78% | 78% | 70% | 71% | 51% | 57% | 58% |
| Houston (n=66) | 73% | 55% | 59% | 58% | 48% | 59% | 55% |
| Mower (n=88) | 66% | 63% | 58% | 58% | 49% | 45% | 41% |
| Olmsted (n=72) | 72% | 86% | 90% | 92% | 90% | 88% | 88% |
| Rice (n=67) | 94% | 94% | 90% | 78% | 79% | 75% | 76% |
| Wabasha (n=64) | 69% | 58% | 53% | 33% | 44% | 42% | 34% |
| Winona (n=74) | 77% | 81% | 80% | 78% | 68% | 72% | 73% |

**NO3 Levels Compared across County**

The plots in Figure 1 display the trends for the NO3 measurements over time for each county. Each line on these plots represents one well in that particular county. Note that some counties have many more wells than others. The two dashed horizontal lines represent NO3 measurements of 3 and 10 mg/L. From these plots, we can see that Winona County, for example, has many wells that have NO3 measurements above 10 mg/L. We see a similar trend in Fillmore County. Conversely, we see that Mower County had only one well that was ever at an elevated level, and that well is now below 10 mg/L.

*Figure 1. NO3 Measurements by County.*

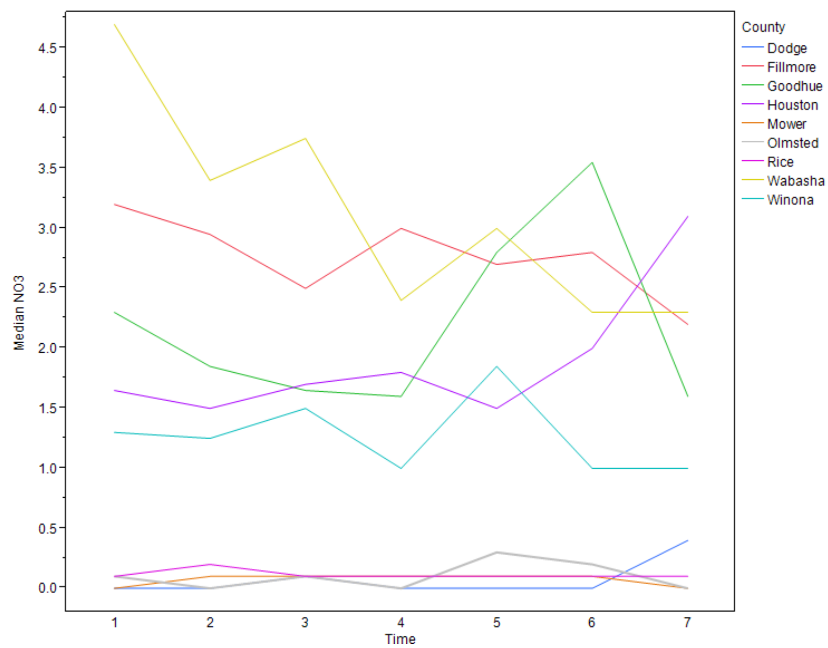





Instead of displaying the measurements for each well, Figure 2 shows the median NO3 level for each County across time (recall that the median is defined as the middle number in a list of sorted numbers). Note that the wells in Wabasha, Fillmore, Goodhue, and Houston Counties tend to have higher NO3 concentrations than do the other counties. Winona County tends to be the next highest. Dodge, Mower, Olmsted, and Rice Counties tend to have lower NO3 concentrations.

**Comment**: Values of nitrate levels reported as 0 may actually be non-detects. These were left coded as 0 for this analysis, which would not affect the median.

*Figure 2. Median NO3 Levels across Time by County.*



Tables 3 and 4 also allow for comparisons of nitrate levels across counties. Table 3 shows the percentage of wells in each county that have elevated nitrate levels (NO3 > 3 mg/L) during each of the seven rounds. Similarly, Table 4 shows the percentage of wells in each county that have nitrate levels exceeding the standard (NO3 > 10 mg/L) during each of the seven rounds.

*Table 3. Percentage of Wells in each Round with Elevated Nitrate Levels ( > 3 mg/L) by County.*

| County | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Dodge | 6.7% | 6.5% | 6.5% | 6.9% | 6.9% | 7.1% | 11.1% |
| Fillmore | 52.6% | 48.9% | 47.2% | 48.4% | 46.6% | 46.6% | 42.2% |
| Goodhue | 38.9% | 38.9% | 39.1% | 38.5% | 44.7% | 53.9% | 43.4% |
| Houston | 43.8% | 44.4% | 43.6% | 39.5% | 34.4% | 43.6% | 52.8% |
| Mower | 24.1% | 25.5% | 21.6% | 19.6% | 16.3% | 17.5% | 13.9% |
| Olmsted | 23.1% | 22.6% | 20.0% | 22.7% | 21.5% | 19.1% | 19.1% |
| Rice | 7.9% | 9.5% | 8.3% | 5.8% | 9.4% | 10.0% | 7.8% |
| Wabasha | 59.9% | 51.4% | 55.6% | 38.1% | 42.9% | 44.4% | 45.5% |
| Winona | 40.4% | 40.0% | 40.7% | 37.9% | 44.0% | 43.4% | 37.0% |

*Table 4. Percentage of Wells in each Round with Nitrate Levels Exceeding 10 mg/L by County.*

| County | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Dodge | 6.7% | 3.2% | 6.5% | 3.5% | 3.5% | 3.6% | 7.4% |
| Fillmore | 21.0% | 13.8% | 11.0% | 17.6% | 13.7% | 15.1% | 4.7% |
| Goodhue | 12.5% | 5.6% | 9.4% | 10.8% | 8.5% | 11.5% | 7.6% |
| Houston | 18.9% | 19.4% | 23.1% | 18.4% | 12.5% | 20.5% | 11.1% |
| Mower | 1.7% | 1.8% | 2.0% | 0% | 0% | 0% | 0% |
| Olmsted | 7.7% | 8.1% | 7.7% | 10.6% | 0% | 7.9% | 9.5% |
| Rice | 4.8% | 4.8% | 5.0% | 3.9% | 5.7% | 4.0% | 3.9% |
| Wabasha | 31.8% | 24.3% | 23.5% | 19.0% | 3.6% | 11.1% | 13.6% |
| Winona | 21.0% | 20.0% | 18.7% | 13.8% | 24.0% | 11.3% | 7.4% |

These trends also indicate that wells in Dodge, Mower, Olmsted, and Rice Counties tend to have lower NO3 concentrations.

**<u>Comparisons across Matrix</u>**

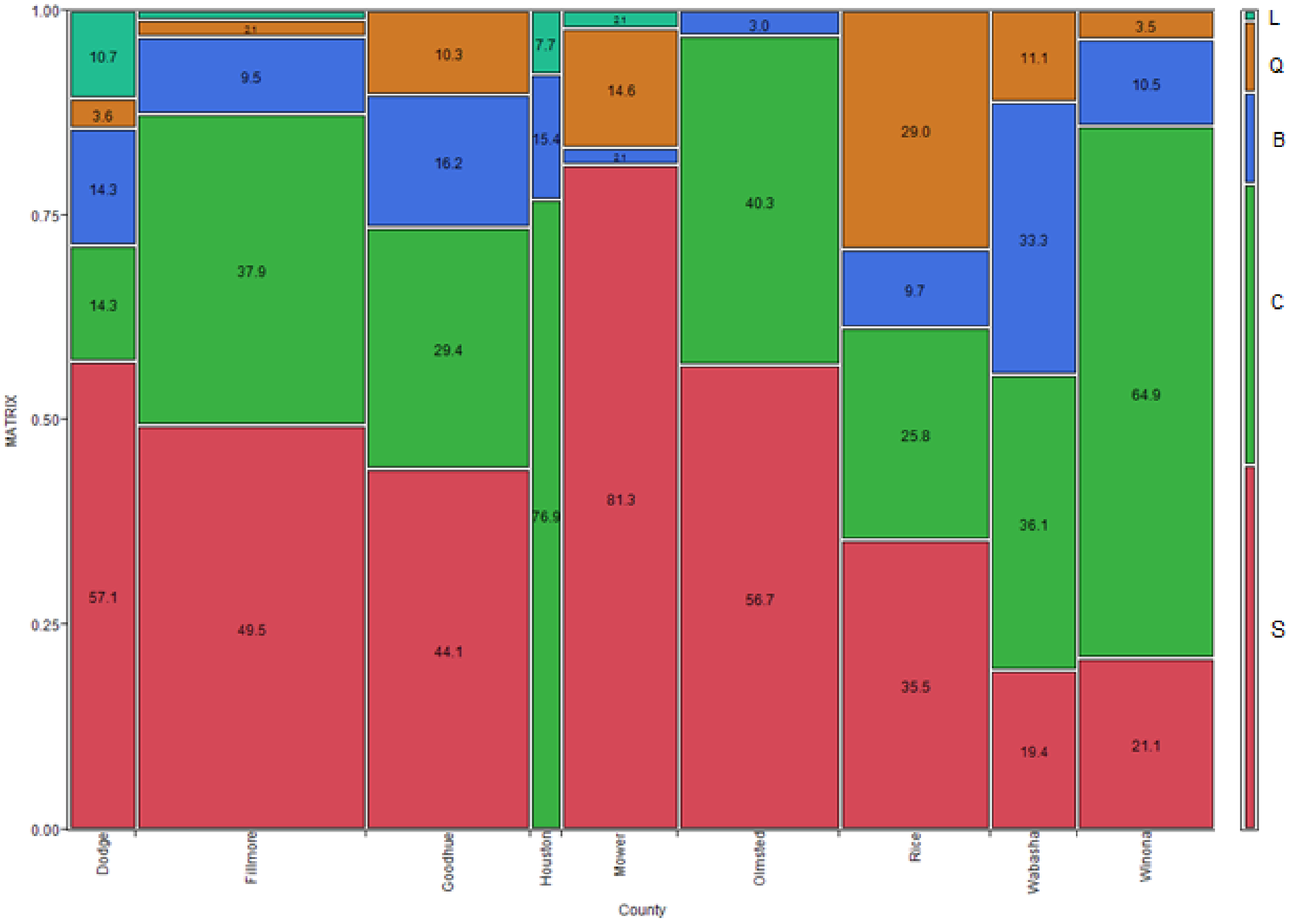
The previous summaries indicate that the NO3 levels vary across county. Note that this could be due to differences in geology from one site to the next. The land around each well in the study was classified using the following matrix types:

- B = solution weathered & clastic bedrock
- C = clastic bed rock (sandstone)
- L = low permeability material
- Q = clastic unconsolidated material
- S = solution weathered bedrock

Figure 3 shows the breakdown of these matrix types by County in a graphic known as a mosaic plot. Each column in this plot shows the percentage of wells contained in each matrix type by county. For example, of all wells in Dodge County, 57.1% are in solution weathered bedrock, 14.3% are in clastic bedrock (sandstone), 14.3% are in solution weathered & clastic bedrock, 3.6% are in clastic unconsolidated material, and 10.7% are in low permeability material.

Some columns are wider than others in this plot because the width of the column is proportional to the amount of wells with Matrix Type data available in each county. For example, for all wells with Matrix type data available, only 13 were in Houston County; conversely, 95 were in Fillmore County.

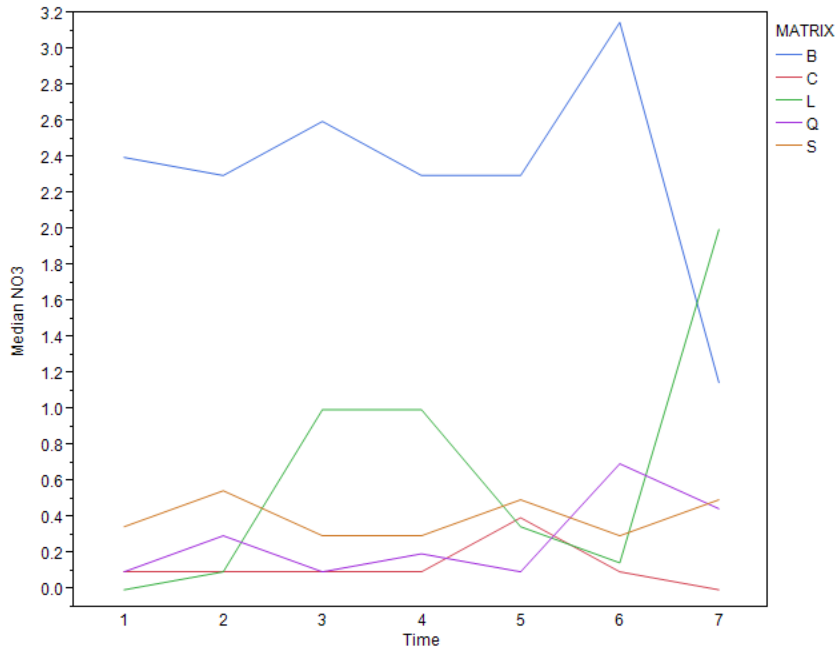
*Figure 3. The Percentage of Wells Contained in Each Matrix Type by County.*




Note that there are differences in Matrix type across counties. For example, 33.3% of the wells in Wabasha County were in Matrix type B, whereas only 2.1% of the wells in Mower County were in Matrix type B. If there are differences in NO3 levels across these Matrix types, then this could potentially help to explain the differences in NO3 levels across counties.

Figure 4 shows the median NO3 level for each matrix type across time. Note that in general, Matrix Type B (solution weathered & clastic bedrock) typically has much higher NO3 levels than the other types. For at least some time periods, Matrix Type L (low permeability material) also appears to have higher NO3 levels than Matrix Types Q, S, and C.

*Figure 4. Median NO3 Levels across Time by Matrix Type.*




<u>Statistical Comparisons</u>: Using the mixed-distribution model described in Appendix A, we first tested for differences in the probability of a well yielding a non-zero NO3 level between each of the Matrix types. The significant pairwise differences are summarized below in Table 5. Traditionally, p-values below .10 provide weak evidence for a difference, whereas p-values below .05 provide strong statistical evidence for a difference.

*Table 5. Significant Differences Across Matrix Type When Comparing the Probabilities of a Non-Zero N03 level.*

| Matrix Type Comparisons | p-value |
|---|---|
| B vs. C | < .0001 |
| B vs. Q | .0394 |
| B vs. S | .0543 |

Next, as also described in Appendix A, we tested for differences in the means of the log-transformed NO3 levels between each Matrix type. The statistically significant differences are summarized in Table 6.

*Table 6. Pairwise Differences Across Matrix Type When Comparing N03 levels.*

| Matrix Type Comparisons | p-value |
|---|---|
| B vs. C | <.0001 |
| B vs. Q | .0055 |
| B vs. S | .0277 |

The results confirm what was seen in Figure 4; the wells found in Matrix Type B (solution weathered & clastic bedrock) tend to yield higher nitrate levels than wells found in Matrix Types B, C and S.

An alternative analysis for making comparisons across Matrix Types was also completed. Instead of considering actual nitrate values, this analysis modeled the probability of a well either (1) having elevated nitrate levels (above 3 mg/L) or (2) exceeding the standard (above 10 mg/L) in each round. The percentage of wells in each of these categories for each Matrix Type are shown in Tables 7 and 8.

*Table 7. Percentage of Wells in each Round with Elevated Nitrate Levels ( > 3 mg/L) by Matrix Type.*

| Matrix Type | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| B | 47.2% | 45.1% | 47.8% | 40.9% | 44.2% | 50.0% | 34.2% |
| C | 23.0% | 20.7% | 22.0% | 18.8% | 19.9% | 20.4% | 20.5% |
| L | 33.3% | 16.7% | 33.3% | 33.3% | 16.7% | 33.3% | 40.0% |
| Q | 15% | 20.5% | 16.2% | 11.8% | 11.8% | 17.2% | 20.0% |
| S | 40.3% | 39.8% | 36.2% | 38.4% | 36.1% | 36.4% | 33.8% |

*Table 8. Percentage of Wells in each Round with Nitrate Levels Exceeding 10 mg/L by Matrix Type.*

| Matrix Type | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| B | 18.9% | 15.7% | 17.4% | 9.1% | 11.6% | 13.6% | 7.9% |
| C | 4.7% | 4.5% | 5.2% | 4.9% | 3.1% | 5.9% | 3.2% |
| L | 16.7% | 16.7% | 16.7% | 16.7% | 16.7% | 16.7% | 20.0% |
| Q | 7.5% | 7.7% | 8.1% | 8.8% | 11.8% | 6.9% | 10.0% |
| S | 19.9% | 14.3% | 12.8% | 15.9% | 10.1% | 10.5% | 8.0% |

Note that Matrix Types B and S tend to have a higher percentage of wells with nitrate values either exceeding 3 mg/L or 10 mg/L.

Statistical Comparisons:
Logistic regression models accounting for the repeated measures structure of the data was used to test whether the probability of a well either having elevated nitrate levels (above 3 mg/L) or of exceeding the 10 mg/L standard differed across Matrix Type or time. PROC GENMOD in SAS version 9 was used to fit the appropriate models.

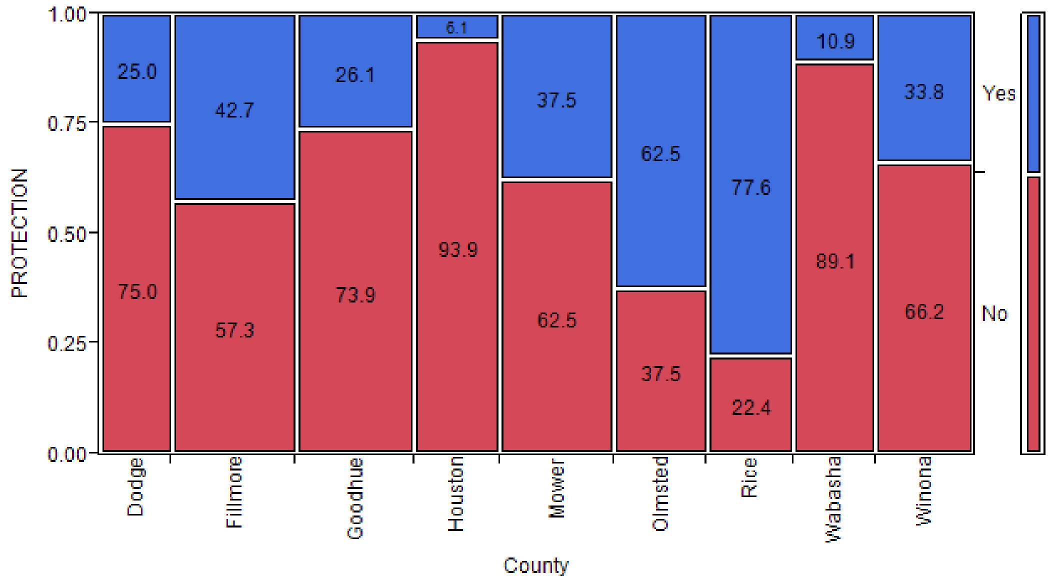
Comparing the Probability of a Well Having Elevated Nitrate Levels: The results indicate that the probability of a well having an elevated nitrate level was significantly different across Matrix Type ($p < .0001$); however, there were no differences across time ($p = .0723$). Further tests revealed which pairwise comparisons were significant: the probability was found to be significantly higher for Matrix Type B than Matrix Type C ($p = .0005$) and Matrix Type Q ($p = .0024$). Also, the probability was found to be significantly higher for Matrix Type S than Matrix Type C ($p = .0003$) and Matrix Type Q ($p = .0051$).

Comparing the Probability of a Well Exceeding the 10 mg/L Standard: The results indicate that the probability of a well having nitrate levels above 10 mg/L was significantly different across Matrix Type ($p = .0015$) and also across time ($p = .0045$). Further tests revealed which pairwise comparisons were significant: the probability of exceeding the standard was found to be significantly higher for Matrix Type B than Matrix Type C ($p = .0076$). Also, the probability was found to be significantly higher for Matrix Type S than Matrix Type C ($p = .0008$). Regarding comparisons across time, the probability was found to be significantly higher in Round 1 than in Round 2 ($p = .0177$), Round 3 ($p = .0121$), Round 5 ($p = .0017$), Round 6 ($p = .0043$), and Round 7 ($p < .0001$). Finally, the probability of exceeding the standard was found to be significantly lower in Round 7 than in Round 1 ($p=.0177$), Round 2 ($p=.0087$), Round 3 ($p = .009$), and Round 4 ($p = .0027$).

## **Comparisons across Protection Status**

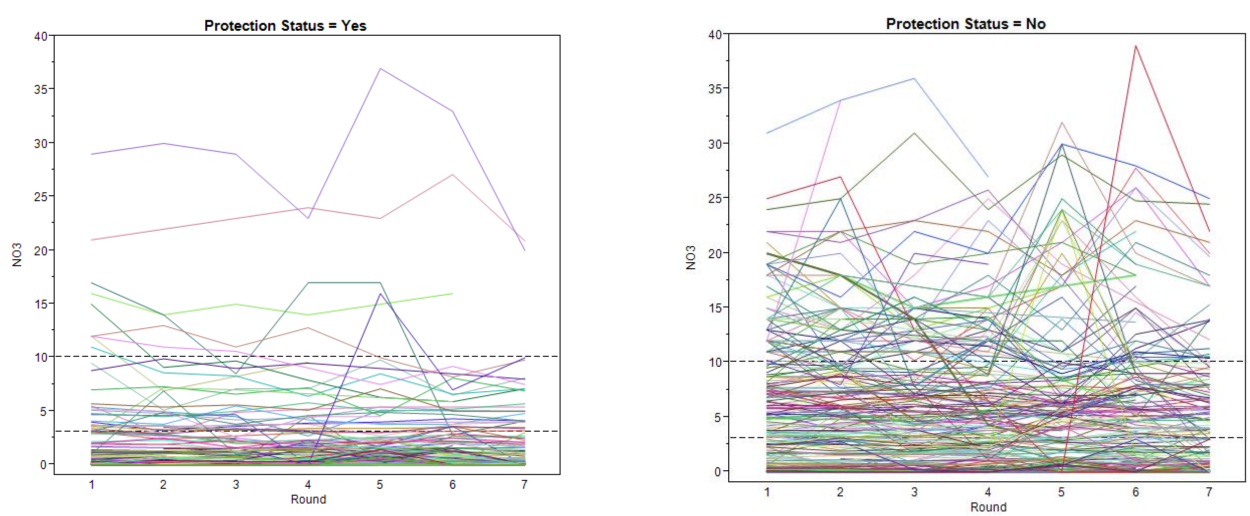
Note that the differences in NO3 levels could also be due to differences in protection status. Each well in the study was classified as either having an overlying protection layer present or not. Figure 5 shows the breakdown of Protection Status by County. For example, of all wells in Fillmore County in this data set, 43.2% are protected and 56.8% are not.

*Figure 5. The Percentage of Wells Protected by County.*




A chi-square test reveals that there are significant differences in Protection Status across County ($p < .0001$). For example, note that only 10.9% of the wells in Wabasha County are protected, whereas 62.5% of the wells in Mower County and 77.6% of the wells in Rice County are protected. This could also potentially explain why the nitrate levels in Wabasha County were higher. Figure 6 shows the trends for the NO3 measurements over time for both protected and non-protected wells.

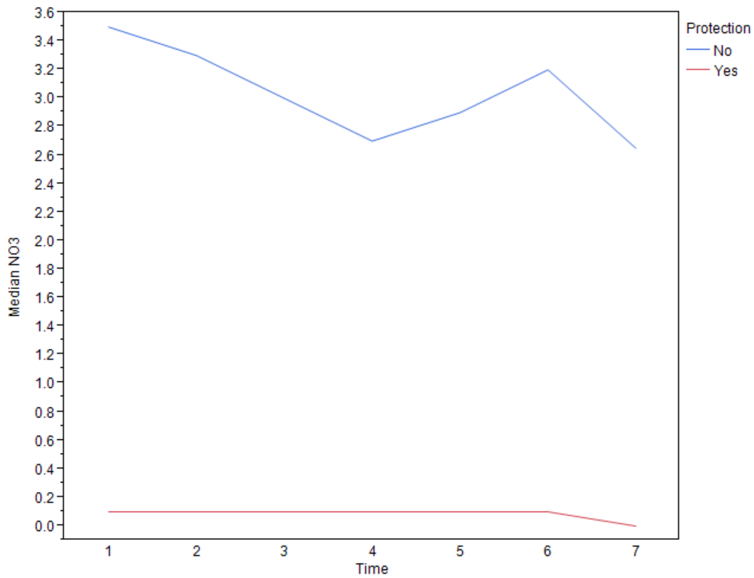
*Figure 6. NO3 Measurements by Protection Status.*



The two dashed horizontal lines represent NO3 measurements of 3 and 10 mg/L. We see that in both cases, many wells have NO3 measurements near 0. However, for the unprotected wells, we see many more wells that have measurements above 10 mg/L and above 3 mg/L when we compare them to the protected wells. For protected wells only, we see a vast majority lie below the cutoff of 3 mg/L, but when we look at the unprotected wells, we see many more wells scattered above both cutoff lines.

Figure 7 shows the median NO3 level for both protected and non-protected wells across time. Note that wells that are protected (red) have much lower median NO3 measurements than the wells that are not protected (blue).

*Figure 7. Median NO3 Levels across Time by Protection Status.*




<u>Statistical Comparisons</u>: Once again using the mixed-distribution model described in Appendix A, we first tested for differences in the probability of a well yielding a non-zero NO3 level between the protected and non-protected wells. This difference was statistically significant ($p < .0001$ ). The non-protected wells have a significantly higher probability of yielding a non-zero NO3 value than do protected wells.

Next, as described in Appendix A, we tested for differences in the means of the log-transformed NO3 levels between protected and non-protected wells. This difference was also statistically significant ($p < .0001$). That is, there is statistical evidence that that the NO3 values are significantly higher for non-protected wells.

An alternative analysis for making comparisons across Protection Status was also completed. Instead of considering actual nitrate values, this analysis modeled the probability of a well either (1) having elevated nitrate levels (above 3 mg/L) or (2) exceeding the standard (above 10 mg/L) in each round. The percentage of wells in each of these categories for each Protection Status are shown in Tables 9 and 10.

*Table 9. Percentage of Wells in each Round with Nitrate Levels Above 3 mg/L by Protection Status.*

| Protection | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Yes | 13.9% | 11.4% | 11.9% | 11.5% | 11.3% | 12.3% | 12.2% |
| No | 52.3% | 52.3% | 50.0% | 48.0% | 47.1% | 50.3% | 46.3% |

*Table 10. Percentage of Wells in each Round with Nitrate Levels Above 10 mg/L by Protection Status.*

| Protection | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Yes | 4.0% | 2.6% | 2.3% | 2.4% | 2.6% | 1.5% | 1.1% |
| No | 23.2% | 18.7% | 18.4% | 19.3% | 13.8% | 16.9% | 13.0% |

Statistical Comparisons:
Logistic regression models accounting for the repeated measures structure of the data was used to test whether the probability of a well either having elevated nitrate levels (above 3 mg/L) or of exceeding the 10 mg/L standard differed across Protection Status or time. PROC GENMOD in SAS version 9 was used to fit the appropriate models.

Comparing the Probability of a Well Having Elevated Nitrate Levels: The results indicate that the probability of a well having an elevated nitrate level was significantly different across Protection Status ($p < .0001$). Wells that are not protected have a much higher chance of having elevated nitrate levels. The probability of having an elevated nitrate level did not differ significantly across time ($p = .0680$).

Comparing the Probability of a Well Exceeding the 10 mg/L Standard: The results indicate that the probability of a well having a nitrate level exceeding the standard was significantly different across Protection Status ($p < .0001$). Wells that are not protected have a much higher chance of exceeding the standard. The probability of exceeding the standard also differed significantly across time ($p = .0058$). These results were quite similar to what was observed in the earlier analysis involving matrix types: the probability was found to be significantly higher in Round 1 than in Rounds 2, 3, 5, 6, and 7. Also, the probability of exceeding the standard was found to be significantly lower in Round 7 than in Rounds 1, 2, and 3.

**Comparisons across WellCode**

Note that the differences in NO3 levels could also be due to differences in WellCode. Each well in the study was classified as either being built to code or not. Figure 8 shows the breakdown of WellCode by County. For example, of all wells in Fillmore County in this data set, 31.3% were built to code and 68.8% were not. A chi-square test reveals that there are significant differences in WellCode across County ($p < .0001$).

*Figure 8. The Percentage of Wells Built to Code by County.*

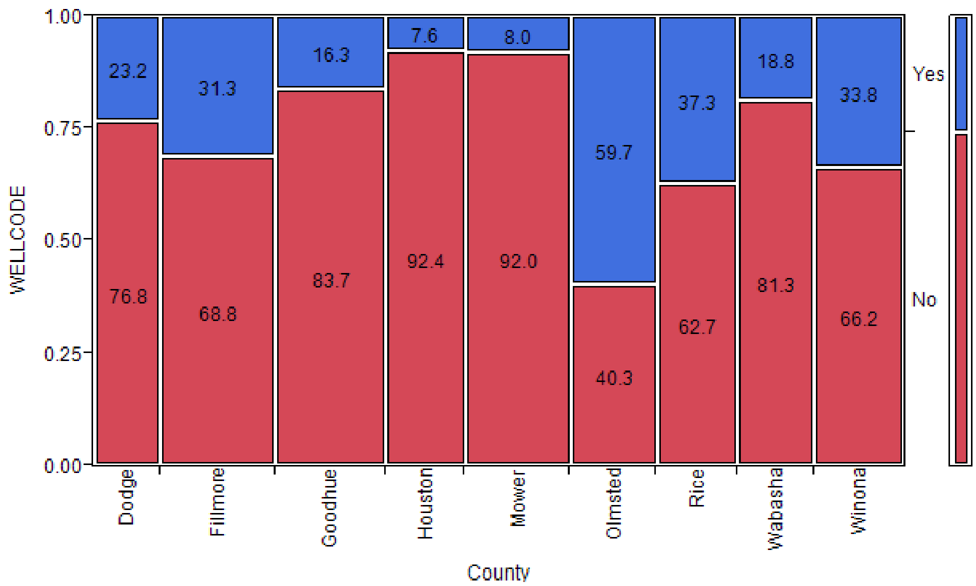
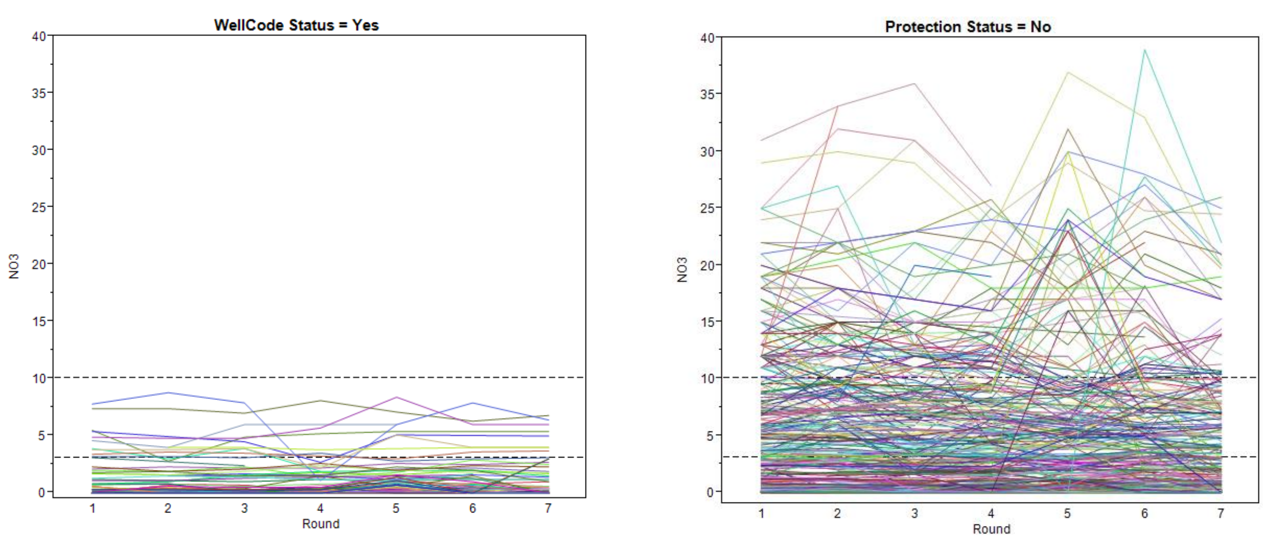



Figure 9 shows the trends for the NO3 measurements over time for both levels of WellCode status.

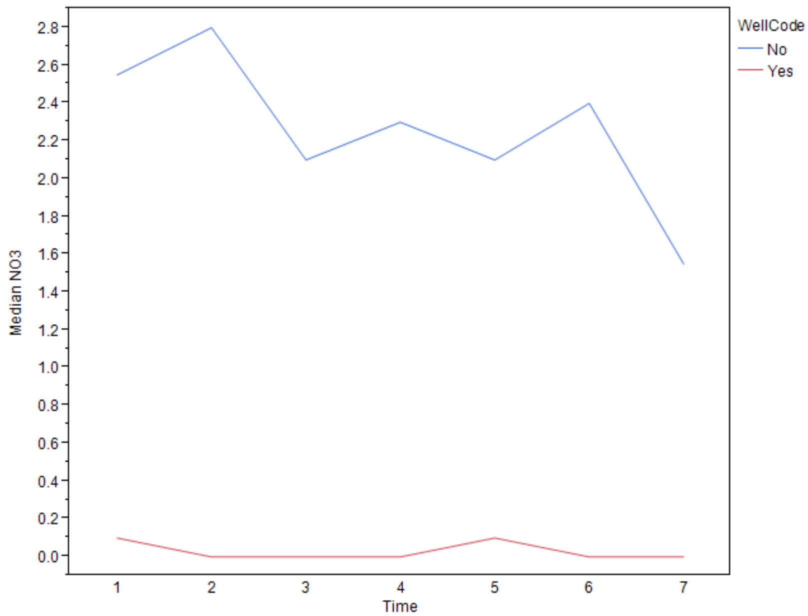
*Figure 9. NO3 Measurements by WellCode Status.*




We can see from these plots that we had many more wells that were not built to code than wells that were. We also see that when examining only wells that were built to code (Wellcode Status = Yes), we did not have any wells that were at elevated NO3 levels (above 10 mg/L). We did, however, see many wells at an elevated level when examining the wells that were not built to code.

Figure 10 shows the median NO3 level for both wells that were and were not built to code across time. Note that wells that were built to code (red) have much lower median NO3 measurements than the wells that were not (blue).

*Figure 10. Median NO3 Levels across Time by WellCode Status.*




<u>Statistical Comparisons</u>: Once again using the mixed-distribution model, we first tested for differences in the probability of a well yielding a non-zero NO3 level between the wells that were and were not built to code. This difference was statistically significant ($p < .0001$ ). The wells not built to code have a significantly higher probability of yielding a non-zero NO3 value than do wells that were built to code.

Next, we tested for differences in the means of the log-transformed NO3 levels between wells that were and were not built to code. This difference was also statistically significant ($p < .0001$). In other words, there is statistical evidence that that the NO3 values are significantly higher for wells not built to code.

An alternative analysis for making comparisons across WellCode Status was also completed. Instead of considering actual nitrate values, this analysis modeled the probability of a well either (1) having elevated nitrate levels (above 3 mg/L) or (2) exceeding the standard (above 10 mg/L) in each round. The percentage of wells in each of these categories for each Well Code Status are shown in Tables 11 and 12.

*Table 11. Percentage of Wells in each Round with Nitrate Levels Above 3 mg/L by WellCode Status.*

| Protection | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Yes | 46.7% | 46.7% | 44.6% | 44.0% | 43.8% | 47.7% | 42.4% |
| No | 7.1% | 4.9% | 6.8% | 3.2% | 4.9% | 4.8% | 6.6% |

*Table 12. Percentage of Wells in each Round with Nitrate Levels Above 10 mg/L by WellCode Status.*

| Protection | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Yes | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| No | 20.6% | 15.9% | 16.6% | 16.5% | 12.7% | 15.1% | 10.4% |

Statistical Comparisons:

Logistic regression models accounting for the repeated measures structure of the data was used to test whether the probability of a well either having elevated nitrate levels (above 3 mg/L) or of exceeding the 10 mg/L standard differed across WellCode Status or time. PROC GENMOD in SAS version 9 was used to fit the appropriate models.

Comparing the Probability of a Well Having Elevated Nitrate Levels: The results indicate that the probability of a well having an elevated nitrate level was significantly different across Well Code Status ($p < .0001$). Wells that are not protected have a much higher chance of having elevated nitrate levels. The probability of having an elevated nitrate level did not differ significantly across time ($p = .1313$).

Comparing the Probability of a Well Exceeding the 10 mg/L Standard: Because there were no protected wells with nitrate levels exceeding the standard, the logistic regression could not be fit for this case. It is clear, however, that non-protected wells are much more likely to have nitrate levels exceeding the 10 mg/L standard.

**Comparisons across Recharge2W**

Finally, note that the differences in NO3 levels could also be due to differences in whether the ground sloped toward a well or away from a well. Each well in the study was classified as either Recharge2W = Yes (when the ground sloped toward the well) or Recharge2W = No (when the ground did not slope toward the well). Figure 11 shows the breakdown of Recharge2W status by County. A chi-square test reveals that there are significant differences in Recharge2W status across County ($p < .0001$).

*Figure 11. The Percentage of Wells toward Which the Ground Slopes by County.*

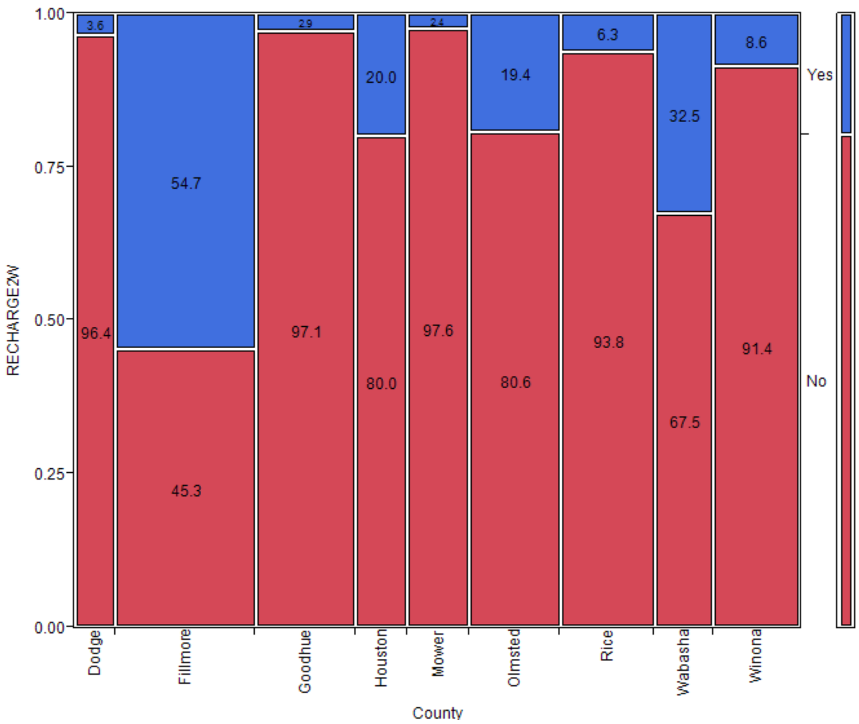
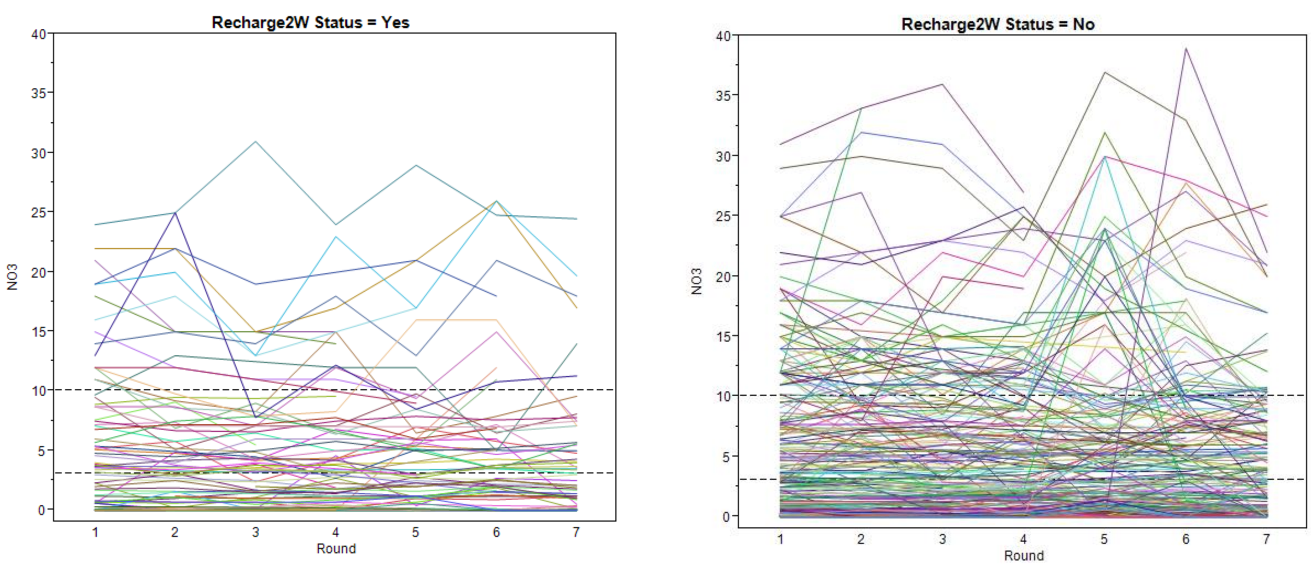



Figure 12 shows the trends for the NO3 measurements over time for both levels of Recharge2W status.

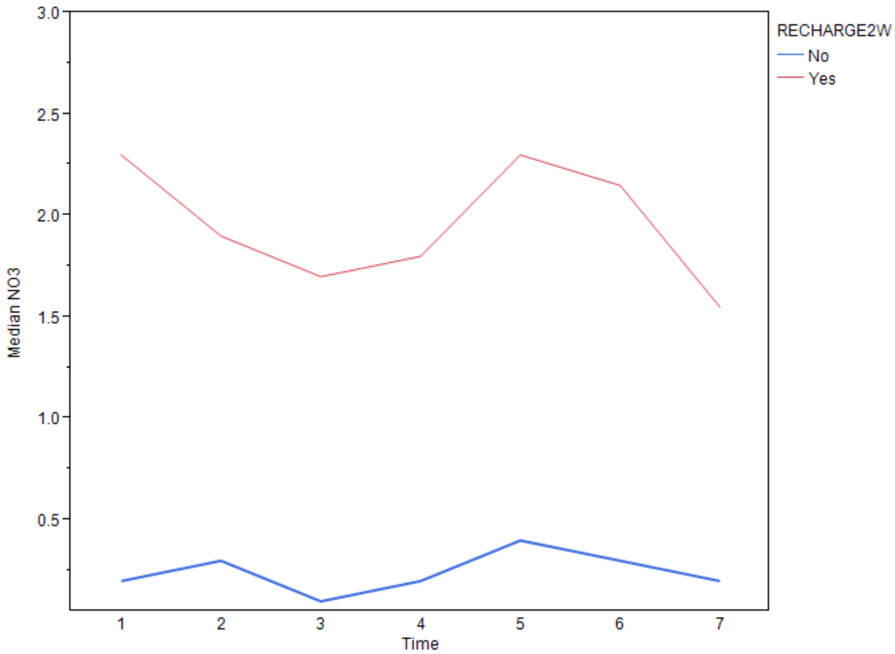
*Figure 12. NO3 Measurements by Recharge2W Status.*



We see that we have more wells without the ground sloping towards them than those that do. No general trend is evident when comparing the two levels of this variable (Recharge2w), as we see similar distributions of NO3 measurements for both cases.

This trend is more evident when we summarize the distributions with the median NO3 value. Figure 13 shows the median NO3 level for both wells that did and did not have the ground slope towards them across time. Note that wells that did not have the ground sloping towards them (blue) have much lower median NO3 measurements than the wells that did (red).

*Figure 13. Median NO3 Levels across Time by Recharge2W Status.*




Statistical Comparisons: Once again using the mixed-distribution model, we first tested for differences in the probability of a well yielding a non-zero NO3 level between the wells that did and did not have the ground sloping towards them. This difference was not statistically significant ($p = .1161$).

Next, we tested for differences in the means of the log-transformed NO3 levels between wells that did and did not have the ground sloping towards them. This difference was statistically significant ($p = .0068$). That is, there is statistical evidence that that the NO3 values are significantly higher for wells that have the ground sloping towards them.

An alternative analysis for making comparisons across Recharge2W Status was also completed. Instead of considering actual nitrate values, this analysis modeled the probability of a well either (1) having elevated nitrate levels (above 3 mg/L) or (2) exceeding the standard (above 10 mg/L) in each round. The percentage of wells in each of these categories for each Recharge2W Status are shown in Tables 13 and 14.

*Table 13. Percentage of Wells in each Round with Nitrate Levels Above 3 mg/L by* Recharge2W *Status.*

| Protection | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Yes | 47.3% | 43.0% | 44.9% | 40.2% | 38.0% | 43.2% | 39.4% |
| No | 32.0% | 31.1% | 29.8% | 28.1% | 28.6% | 29.8% | 28.1% |

*Table 14. Percentage of Wells in each Round with Nitrate Levels Above 10 mg/L by* Recharge2W *Status.*

| Protection | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| Yes | 17.6% | 12.9% | 11.2% | 14.6% | 10.1% | 13.5% | 9.1% |
| No | 14.0% | 11.4% | 11.9% | 11.2% | 8.7% | 9.5% | 6.8% |

Statistical Comparisons:

Logistic regression models accounting for the repeated measures structure of the data was used to test whether the probability of a well either having elevated nitrate levels (above 3 mg/L) or of exceeding the 10 mg/L standard differed across Recharge2W Status or time. PROC GENMOD in SAS version 9 was used to fit the appropriate models.

Comparing the Probability of a Well Having Elevated Nitrate Levels: The results indicate that the probability of a well having an elevated nitrate level was significantly different across Recharge2W Status (p = .0211). Wells that have the ground sloping towards them have a much higher chance of having elevated nitrate levels. The probability of having an elevated nitrate level did not differ significantly across time (p = .2215).

Comparing the Probability of a Well Exceeding the 10 mg/L Standard: The results indicate that the probability of a well having a nitrate level exceeding the standard did not differ significantly across Recharge2W Status (p = .4276). The probability of exceeding the standard did differ significantly across time, however (p = .0003). Once again, these results were quite similar to what was observed earlier: the probability was found to be significantly higher in Round 1 than in Rounds 2, 3, 5, 6, and 7. Also, the probability of exceeding the standard was found to be significantly lower in Round 7 than in Rounds 1, 2, and 3.

**Examining Matrix, Protection, WellCode, and Recharge2W Simultaneously**

Earlier analyses consider the factors separately. To consider these factors jointly, a single logistic regression model was fit to examine the effects of Matrix, Protection, WellCode, Recharge2W, and Round on the probability of a well either (1) having elevated nitrates (above 3 mg/L) or (2) having nitrates that exceed the standard (above 10 mg/L). This model accounted for the repeated measures nature of the data and was fit using PROC GENMOD in SAS version 9.

The results presented in Table 15 indicate that all factors were significant except for time. This indicates that each factor is still a significant predictor of the probability of having elevated nitrate levels even after adjusting for the effects of the other variables.

*Table 15. Results of Logistic Regression Model for Predicting the Probability of Nitrates Exceeding 3 mg/L.*

| Variable | p-value |
|---|---|
| Matrix | .0087 |
| Protection Status | <.0001 |
| WellCode Status | <.0001 |
| Recharge2W Status | .0004 |
| Round | .1775 |

The probability was found to be significantly higher for Matrix Type B than Matrix Type Matrix Type Q (p = .016). Also, the probability was found to be significantly higher for Matrix Type S than Matrix Type Q (p = .0007). Non-protected wells, wells not up to code, and wells with the ground sloping towards them were found to have significantly higher chances of having elevated nitrate levels than their counterparts.

A similar model was fit to predict the probability of exceeding the 10 mg/L standard. The results are presented in Table 16. Note that WellCode status was not included in this model since no protected wells exceeded the 10 mg/L standard (including this variable in the analysis would not allow for successfully fitting a logistic regression model). The results indicate that all factors were significant except for Recharge2W status. Once again, this indicates that Matrix, Protection Status, and Time are all significant predictors of the probability of having nitrate levels above 10 mg/L even after adjusting for the effects of the other variables.

*Table 16. Results of Logistic Regression Model for Predicting the Probability of Nitrates Exceeding 10 mg/L.*

| Variable | p-value |
|---|---|
| Matrix | .0016 |
| Protection Status | <.0001 |
| Recharge2W Status | .1527 |
| Round | .0152 |

The probability of exceeding the standard was found to be significantly higher for Matrix Type B than Matrix Type C (p = .0491). Also, the probability was found to be significantly higher for Matrix Type S than Matrix Type C (p = .0005). Wells not up to code and wells with the ground sloping towards them were found to have significantly higher chances of having elevated nitrate levels than their counterparts. Regarding comparisons across time, the probability was found to be significantly higher in Round 1 than in Round 2 (p = .0227), Round 3 (p = .0166), Round 5 (p = .0044), Round 6 (p = .0115), and Round 7 (p = .0005). Finally, the probability of exceeding the standard was found to be significantly lower in Round 7 than in Round 1 (p=.0277), Round 2 (p=.0347), Round 3 (p = .0319), and Round 4 (p = .0074).

**Comparisons across Aquifer**

It may also be of interest to compare nitrate levels across aquifer. Figure 14 shows the median NO3 levels across time for each aquifer. Note that there are 7 aquifers that have elevated median scores (above 3 mg/L) for one (if not all) time periods. These aquifers are as follows: CJFR (which contained only 2 wells), DCLP (which contained 7 wells), OGCM (which contained 3 wells), OPCJ (which contained 30 wells), OPGW (which contained only 2 wells), OPNR (which contained only 1 well), and OPOD (which contained 10 wells). Table 17 shows the number of wells that were measured in each round by aquifer.

Figure 14. *Median NO3 Levels across Time by Aquifer (the number of wells in each aquifer is indicated in parentheses).*

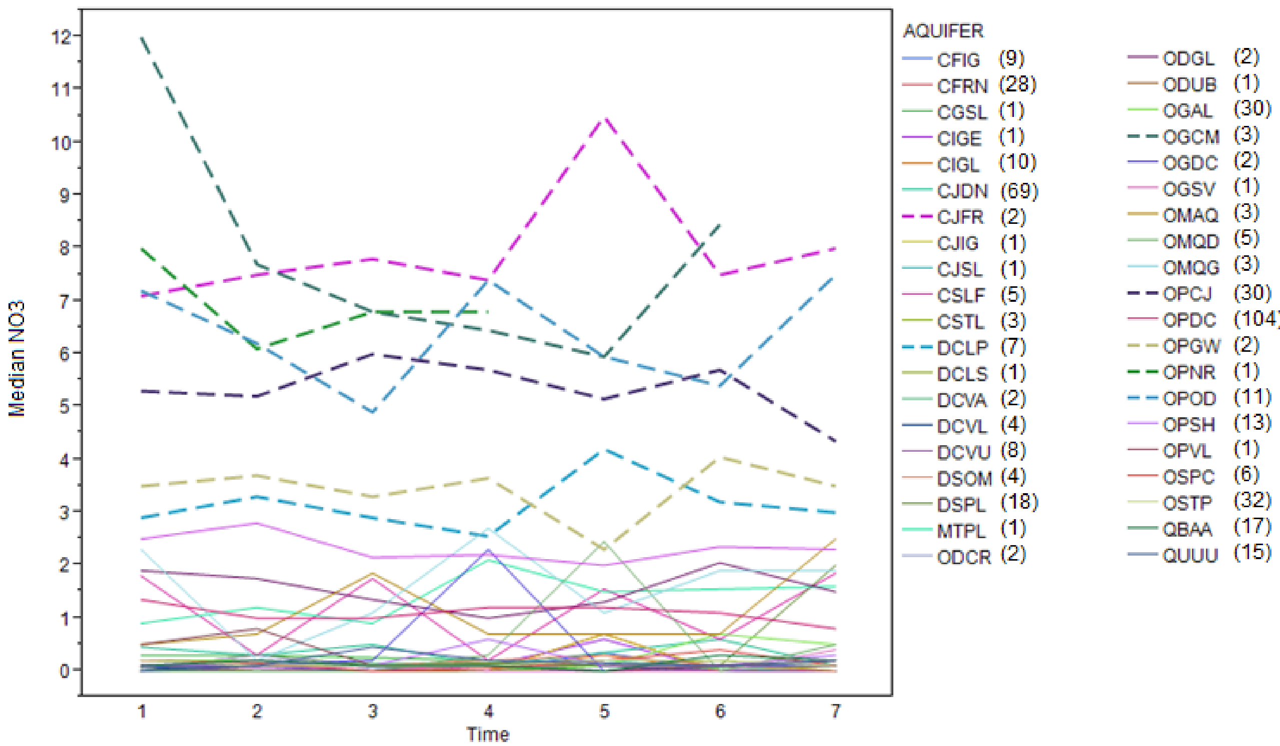



*Table 17. The Number of Wells Measured in Each Round by Aquifer.*

| Aquifer | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| CFIG | 9 | 9 | 8 | 8 | 8 | 9 | 7 |
| CFRN | 27 | 28 | 27 | 25 | 21 | 23 | 23 |
| CGSL | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| CIGE | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| CIGL | 9 | 9 | 9 | 8 | 8 | 8 | 7 |
| CJDN | 62 | 64 | 67 | 63 | 56 | 58 | 51 |
| CJFR | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| CJIG | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| CJSL | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| CSLF | 4 | 5 | 4 | 3 | 4 | 3 | 2 |
| CSTL | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

| Aquifer | Round1 | Round2 | Round3 | Round4 | Round5 | Round6 | Round7 |
|---|---|---|---|---|---|---|---|
| DCLP | 7 | 7 | 7 | 6 | 5 | 5 | 5 |
| DCLS | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| DCVA | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| DCVL | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| DCVU | 8 | 8 | 5 | 7 | 5 | 6 | 5 |
| DSOM | 3 | 4 | 3 | 3 | 3 | 3 | 3 |
| DSPL | 18 | 17 | 17 | 16 | 13 | 10 | 11 |
| MTPL | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| ODCR | 2 | 2 | 2 | 1 | 1 | 1 | 2 |
| ODGL | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| ODUB | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| OGAL | 29 | 29 | 27 | 27 | 25 | 24 | 24 |
| OCGM | 3 | 3 | 3 | 2 | 2 | 2 | 0 |
| OGDC | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| OGSV | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| OMAQ | 3 | 3 | 2 | 3 | 3 | 3 | 2 |
| OMQD | 5 | 5 | 5 | 5 | 4 | 5 | 5 |
| OMQG | 3 | 3 | 3 | 3 | 3 | 3 | 2 |
| OPCJ | 30 | 29 | 26 | 24 | 24 | 25 | 20 |
| OPDC | 94 | 93 | 93 | 88 | 82 | 78 | 71 |
| OPGW | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| OPNR | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| OPOD | 10 | 10 | 9 | 10 | 8 | 9 | 8 |
| OPSH | 12 | 13 | 11 | 11 | 13 | 13 | 12 |
| OPVL | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| OSPC | 6 | 6 | 5 | 4 | 5 | 4 | 4 |
| OSTP | 27 | 30 | 30 | 28 | 27 | 28 | 28 |
| QBAA | 16 | 16 | 15 | 14 | 15 | 11 | 12 |
| QUUU | 15 | 14 | 14 | 13 | 12 | 12 | 10 |

Appendix A. Description of Statistical Methods Used to Test for Differences Across Matrix Type

To test whether the differences in NO3 levels from one Matrix type to another are statistically significant, traditional repeated measures analyses are not appropriate because of the abundance of zeros in the data set and fact that the NO3 measurements are highly positively skewed. To account for this, we used a mixed-distribution model with correlated random effects. First, a "logistic" response component estimates the probability of obtaining a non-zero NO3 value for each Matrix type. Second, for those non-zero NO3 measurements, a "lognormal" response component estimates the mean of the log-transformed positive NO3 values and allows us to test for differences in the actual NO3 values across Matrix type.

We conducted these analyses using a SAS macro, MIXCORR, which is referenced below.

Note that similar analyses were conducted to make comparisons across Protection Status, whether the wells were up to code, and whether the ground sloped toward a well.

*Source: Tooze JA, Grunwald GK, Jones RH. Analysis of repeated measures data with clumping at zero. Stat Methods Med Res. 2002;11:341–355.*