ENVIRONMENTAL SITE CHARACTERIZATION Via ARTIFICIAL NEURAL NETWORK APPROACH

by

MAHMOUD MRYYAN

B.S., University of Missouri, 1990
M.S., Kansas State University, 1999

AN ABSTRACT OF A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Civil Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2008

Abstract

This study explored the potential use of ANNs for profiling and characterization of various environmental sites. It investigates the following environmental site profiling cases:

1. Two-dimensional and three-dimensional characterizations of a hypothetical data-rich site by various profiling methods
2. Two-dimensional characterizations of the inorganic materials (lead and copper) in soil and groundwater at a landfill site
3. Three-dimensional, time-related profiling of explosive-related contaminants (perchlorate) at the Massachusetts Military Reservation site

When examining the performance of various site profiling methodologies for a comparative analysis, a static ANN with back-propagation algorithm was used to model the environmental containment at a hypothetical data-rich contaminated site. The performance of the ANN profiling model was then compared to the following profiling models: (1) Inverse Distance to a Power, (2) Kriging, (3) Minimum Curvature, (4) Modified Shepard's, (5) Nearest Neighbor, (6) Polynomial Regression, (7) Radial Basis Function, and (8) Local Polynomial. The comparison showed that the ANN-based models proved to yield the lowest error values in the 2-D and 3-D comparison cases. The ANN-based profiling models also produced the best contaminant distribution contour maps when compared to the actual maps. Along with the fact that ANN is the only profiling methodology that allows for efficient 3-D profiling, this study clearly demonstrates that ANN-based methodology, when properly used, has the potential to provide the most accurate predictions and site profiling contour maps for a contaminated site.

ANN with a back-propagation learning algorithm was utilized in the site characterization of contaminants at the Kansas City landfill. The use of ANN profiling models made it possible to

obtain reliable predictions about the location and concentration of lead and copper contamination at the associated Kansas City landfill site. The resulting profiles can be used to determine additional sampling locations, if needed, for both groundwater and soil in any contaminated zones.

Back-propagation networks were used to characterize the MMR Demo 1 site. The purpose of the developed ANN models was to predict the concentrations of perchlorate at the MMR from appropriate input parameters. To determine the most-appropriate input parameters for this model, three different cases were investigated using nine potential input parameters.

Although the findings for seven-input and eight-input cases were somewhat comparable, the nine-input case model outperformed the seven and the eight inputs case models, therefore identifying it as the optimal ANN model for this study. It was determined that the optimal network model for the MMR perchlorate prediction model contained nine input parameters, nine hidden node, and one output parameter (9-9-1).

The ANN modeling used in this case demonstrates the neural network's ability to accurately predict perchlorate contamination using multiple variables. When comparing the trends observed using the ANN-generated data and the actual trends identified in the MMR 2006 System Performance Monitoring Report, both agree that perchlorate levels are decreasing due to the use of the ETR systems. This proves that the ETR systems were both effective and necessary for the removal of perchlorate contamination at the Demo 1 site, as demonstrated in the contour maps.

Using the knowledge obtained from the MMR perchlorate prediction model, a similar ANN with a back-propagation learning algorithm was developed to model the data importance at the Massachusetts Military Reservation site. In various testing trials, twenty-eight back-

propagation ANN models were developed, which excluded or included certain groundwater monitoring wells. These models were then used to investigate the minimum number of groundwater wells necessary to characterize the Demo 1 site accurately.

This research demonstrates the advantages of ANN site characterization modeling in contrast with traditional modeling. First, no complex mathematical formulations were developed to describe the behavior of the contaminants, and the ANN model was built up simply by training on the available laboratory/analytical data. Second, the trained-ANN model can simulate new scenarios without the need for any additional laboratory analytical-based information. Third, the developed ANN model is convenient for practical usage by either acting as a standalone simulator or by being implemented into another program (Microsoft Excel spreadsheet or Surfer Program). Fourth, flexibility and generality characterized the generated ANN-based models. Once a decision is made for what networks is to represent a site, this network can be readily used to predict the contaminant values at any desired location—this demonstrates flexibility. The only parameter a trained network needs in order to provide such predictions is the input data vector such as (x, y, z) coordinates of the point at which a prediction is desired. Generality lies in ANN's power to capture the mode of change of a contaminant's parameters based on all available data.

ENVIRONMENTAL SITE CHARACTERIZATION VIA ARTIFICIAL NEURAL NETWORK APPROACH

by

MAHMOUD MRYYAN

B.S., University of Missouri, 1990
M.S., Kansas State University, 1999

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Civil Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2008

Approved by:

Major Professor
Dr. Yacoub M. Najjar

Copyright

MAHMOUD MRYYAN

2008

Abstract

This study explored the potential use of ANNs for profiling and characterization of various environmental sites. It investigates the following environmental site profiling cases:

4. Two-dimensional and three-dimensional characterizations of a hypothetical data-rich site by various profiling methods
5. Two-dimensional characterizations of the inorganic materials (lead and copper) in soil and groundwater at a landfill site
6. Three-dimensional, time-related profiling of explosive-related contaminants (perchlorate) at the Massachusetts Military Reservation site

When examining the performance of various site profiling methodologies for a comparative analysis, a static ANN with back-propagation algorithm was used to model the environmental containment at a hypothetical data-rich contaminated site. The performance of the ANN profiling model was then compared to the following profiling models: (1) Inverse Distance to a Power, (2) Kriging, (3) Minimum Curvature, (4) Modified Shepard's, (5) Nearest Neighbor, (6) Polynomial Regression, (7) Radial Basis Function, and (8) Local Polynomial. The comparison showed that the ANN-based models proved to yield the lowest error values in the 2-D and 3-D comparison cases. The ANN-based profiling models also produced the best contaminant distribution contour maps when compared to the actual maps. Along with the fact that ANN is the only profiling methodology that allows for efficient 3-D profiling, this study clearly demonstrates that ANN-based methodology, when properly used, has the potential to provide the most accurate predictions and site profiling contour maps for a contaminated site.

ANN with a back-propagation learning algorithm was utilized in the site characterization of contaminants at the Kansas City landfill. The use of ANN profiling models made it possible to

obtain reliable predictions about the location and concentration of lead and copper contamination at the associated Kansas City landfill site. The resulting profiles can be used to determine additional sampling locations, if needed, for both groundwater and soil in any contaminated zones.

Back-propagation networks were used to characterize the MMR Demo 1 site. The purpose of the developed ANN models was to predict the concentrations of perchlorate at the MMR from appropriate input parameters. To determine the most-appropriate input parameters for this model, three different cases were investigated using nine potential input parameters.

Although the findings for seven-input and eight-input cases were somewhat comparable, the nine-input case model outperformed the seven and the eight inputs case models, therefore identifying it as the optimal ANN model for this study. It was determined that the optimal network model for the MMR perchlorate prediction model contained nine input parameters, nine hidden node, and one output parameter (9-9-1).

The ANN modeling used in this case demonstrates the neural network's ability to accurately predict perchlorate contamination using multiple variables. When comparing the trends observed using the ANN-generated data and the actual trends identified in the MMR 2006 System Performance Monitoring Report, both agree that perchlorate levels are decreasing due to the use of the ETR systems. This proves that the ETR systems were both effective and necessary for the removal of perchlorate contamination at the Demo 1 site, as demonstrated in the contour maps.

Using the knowledge obtained from the MMR perchlorate prediction model, a similar ANN with a back-propagation learning algorithm was developed to model the data importance at the Massachusetts Military Reservation site. In various testing trials, twenty-eight back-

propagation ANN models were developed, which excluded or included certain groundwater monitoring wells. These models were then used to investigate the minimum number of groundwater wells necessary to characterize the Demo 1 site accurately.

This research demonstrates the advantages of ANN site characterization modeling in contrast with traditional modeling. First, no complex mathematical formulations were developed to describe the behavior of the contaminants, and the ANN model was built up simply by training on the available laboratory/analytical data. Second, the trained-ANN model can simulate new scenarios without the need for any additional laboratory analytical-based information. Third, the developed ANN model is convenient for practical usage by either acting as a standalone simulator or by being implemented into another program (Microsoft Excel spreadsheet or Surfer Program). Fourth, flexibility and generality characterized the generated ANN-based models. Once a decision is made for what networks is to represent a site, this network can be readily used to predict the contaminant values at any desired location—this demonstrates flexibility. The only parameter a trained network needs in order to provide such predictions is the input data vector such as (x, y, z) coordinates of the point at which a prediction is desired. Generality lies in ANN's power to capture the mode of change of a contaminant's parameters based on all available data.

Table of Contents

List of Figures xvi

List of Tables xxiv

Acknowledgements xxv

Dedication xxvi

CHAPTER 1 - Introduction 1

1.1 Overview 1

1.2 Problem Statement 3

1.2.1 Two-dimensional and three-dimensional characterizations of a hypothetical data-rich site by various methods 3

1.2.2 Two-dimensional characterizations of the inorganic materials (lead and copper) in soil and groundwater at a landfill site. 4

1.2.3 Three-dimensional time-related profiling of explosive-related contaminants (perchlorate) at the Massachusetts Military Reservation site 5

1.3 Objectives 6

1.4 Organization of Dissertation 7

CHAPTER 2 - Literature Review 10

2.1 Application of Artificial Neural Networks in Environmental Site Characterization 10

2.2 Traditional Methods 10

2.3 Artificial Neural Networks Methodology 14

2.3.1 ANN in Water Contamination Profiling 15

2.3.2 ANN in Soil Contamination Profiling 17

CHAPTER 3 - Artificial Neural Networks 20
3.1 Introduction 20
3.2 Elements of an Artificial Neural Network 20
3.2.1 Neuron 21
3.2.2 Input Layer 22
3.2.3 Hidden Layer(s) 22
3.2.4 Output Layer 23
3.2.5 Connection Weights 24
3.3 Back-Propagation Neural Networks 24
3.4 Back-Propagation of Error 26
3.5 Transfer Functions 27
3.5.1 Sigmoidal Function 27
3.5.2 Hard Limiter Function 27
3.5.3 Threshold Logic Function 28
3.6 Model Implementation 28
3.6.1 Supervised Learning 28
3.6.2 Unsupervised Learning 29
3.6.3 Reinforcement Learning 30
3.6.4 Training of a Network 30
3.7 Accuracy Measures 31
CHAPTER 4 - Two-dimensional and three-dimensional characterizations of a hypothetical data-rich site by various methods 39
4.1 Introduction 39

4.2 Surfer® Profiling Methods 40
4.2.1 Inverse Distance to a Power 40
4.2.2 Kriging 41
4.2.3 Minimum Curvature 42
4.2.4 Modified Shepard's 42
4.2.5 Natural Neighbor 43
4.2.6 Polynomial Regression 44
4.2.7 Radial Basis Function 44
4.2.8 Local Polynomial 44
4.3 Two-Dimensional Case 44
4.3.1 Mathematical Equation 44
4.3.2 Databank 45
4.3.3 ANN Model Development 46
4.4 Three-Dimensional Case 49
4.4.1 Mathematical Equation 50
4.4.2 ANN Model Development 50
4.4.3 Regression Model Development 51
4.5 Results and Discussion 52
4.5.1 Comparison Using RMSE Values 52
4.5.2 Comparison Using Contour Maps 54
4.5.3 Comparison Using Forty-Five-Degree Scatter Graphs 56
4.6 Conclusion 57

CHAPTER 5 - Two-dimensional characterizations of the inorganic materials (lead and copper) in soil and groundwater at a landfill site 96
5.1 Background Information 96
5.2 Landfill Site Description 97
5.3 Site Investigation 98
5.3.1 Information Gathering 99
5.4 Database 100
5.4.1 Model Development 100
5.4.2 Databank Generation 102
5.5 Results and Analysis 103
5.5.1 Soil-Based Maps 103
5.5.2 Groundwater-Based Maps 104
5.6 Concluding Remarks 104
CHAPTER 6 - Study Area: the Massachusetts Military Reservations 120
6.1 Background 120
6.2 Study Area 120
6.2.1 Location and description of Demolition Area One (Demo 1) 121
6.2.2 Hydrology of Cape Cod 122
6.2.3 Geology of Cape Cod 123
6.2.4 History of the MMR 124
6.2.5 Explosive Related Contamination at Demo 1 125
6.3 Perchlorate 127
6.3.1 Background 127

6.3.2 Dangers 129
6.3.3 Hypothyroidism Explained 130
6.3.4 Perchlorate and the law 131
CHAPTER 7 - Three-dimensional time-related profiling of explosive-related contaminants (perchlorate) at the Massachusetts Military Reservation site 143
7.1 Introduction 143
7.2 Groundwater contamination 144
7.3 Monitoring Groundwater 145
7.4 Background of Study Area 146
7.5 Pre-Existing Data 147
7.6 Model Development 148
7.6.1 Determination of Appropriate Model Inputs 148
7.6.2 Model Training and Testing 150
7.6.3 Model Selection 151
7.7 Data Banks 152
7.8 Excel Application 152
7.9 Contour Maps 153
7.10 Concluding Remarks 154
CHAPTER 8 - ANN-Based Profiling: Data Importance 177
8.1 Introduction 177
8.2 MMR-Demo 1 Groundwater Monitoring Network 177
8.3 ANN Model Development 178
8.3.1 Determination of Appropriate Model Inputs 179

8.3.2 Data Banks and Model Selection 179
8.4 Contour Maps 181
8.5 Results and Discussion 182
8.5.1 Comparison Using ASE Values 182
8.5.2 Comparison Using Contour Maps 183
8.6 Discussion 184
8.6.1 Contour maps generated for year 2000 184
8.6.2 Contour maps generated for year 2001 185
8.6.3 Contour maps generated for year 2002 186
8.6.4 Contour maps generated for year 2003 187
8.6.5 Contour maps generated for year 2004 188
8.6.6 Contour maps generated for year 2005 189
8.7 Conclusion 189
CHAPTER 9 - Summary, Conclusions and Recommendations 206
9.1 Summary 206
9.2 Conclusions 209
9.3 Recommendations 212
References 214

List of Figures

Figure 3.1. A typical multi-layer ANN showing the input layer for ten different inputs, the middle or hidden layer(s), and the output layer having three outputs 33

Figure 3.2. The basic building block of the network system, the neuron 34

Figure 3.3. Artificial neuron in its simplest form 34

Figure 3.4. McCulloch-Pitts neuron model 35

Figure 3.5. ANN with hidden layer 35

Figure 3.6. Diagram shows a back-propagation neural network 36

Figure 3.7. Sigmoidal function 37

Figure 3.8. Hard limiter function 37

Figure 3.9. Threshold logic function 38

Figure 4.1. Location of testing and training points for the 10-ft interval case 60

Figure 4.2. All of the profiling methods and their RMSE values (2-D and 25-ft interval case) .. 61

Figure 4.3. All of the profiling methods and their RMSE values (2-D and 10-ft interval case) .. 62

Figure 4.4. Bar chart showing RMSE value for profiling methods used to predict (V) values for 3-D case 63

Figure 4.5. Baseline contour map of the pollutant V (for the 10-ft interval case), based on 10,201 actual data points 64

Figure 4.6. Contour map based on ANN model 2-2-1 for the 10-ft interval case 65

Figure 4.7. Contour map based on Radial Basis Function method for the 10-ft interval case 66

Figure 4.8. Contour map based on Kriging method for the 10-ft interval case 67

Figure 4.9. Contour map based on Modified Shepard's method for the 10-ft interval case..... 68

Figure 4.10. Contour map based on Local Polynomial method for the 10-ft interval case 69

Figure 4.11 Contour map based on Minimum Curvature method for the 10-ft interval case..... 70

Figure 4.12. Contour map based on Nearest Neighbor method for the 10-ft interval case 71

Figure 4.13. Contour map based on Polynomial Regression method for the 10-ft interval case . 72

Figure 4.14. Contour map based on Inverse Distance Method for the 10-ft interval case 73

Figure 4.15. Baseline contour map of the pollutant V (at z = 25 ft) based on 1,681 actual data points (3-D case) 74

Figure 4.16. Contour map based on ANN model 3-3-1 at z = 25 ft (3-D case)..... 75

Figure 4.17. Contour map for regression-based predicted concentration V at Z = 25 ft (3-D case)..... 76

Figure 4.18. Scatter plot of actual "V" and "V" predicted by the neural network model (2-D and 25-ft interval case) 77

Figure 4.19. Scatter plot of actual "V" and "V" predicted by the Local Polynomial Method, (2-D and 25-ft interval case)..... 78

Figure 4.20. Scatter plot of actual "V" and "V" predicted by Modified Shepard's method (2-D and 25-ft interval case)..... 79

Figure 4.21. Scatter plot of actual "V" and "V" predicted by Minimum Curvature method (2-D and 25-ft interval case)..... 80

Figure 4.22. Scatter plot of actual "V" and "V" predicted by the Kriging method (2-D and 25-ft interval case) 81

Figure 4.23. Scatter plot of actual "V" and "V" predicted by Polynomial Regression (2-D and 25-ft interval case) 82

Figure 4.24. Scatter plot of actual “V” and “V” predicted by Nearest Neighbor method (2-D and 25-ft interval case) 83
Figure 4.25. Scatter plot of actual “V” and “V” predicted by Radial Basis method (2-D and 25-ft interval case) 84
Figure 4.26. Scatter plot of actual “V” values and “V” values predicted by the Inverse Distance of Power (2-D and 25-ft Interval Case) 85
Figure 4.27. Scatter plot of actual “V” values and “V” values obtained by the Neural Network model (2-D and 10-ft interval case) 86
Figure 4.28. Scatter plot of actual “V” values and “V” values obtained by Radial Basis Function (2-D and 10-ft interval case) 87
Figure 4.29. Scatter plot of actual “V” values and “V” values obtained by Kriging Method (2-D and 10-ft interval case)..... 88
Figure 4.30. Scatter plot of actual “V” values and “V” values obtained by Modified Shepard's method (2-D and 10-ft interval case) 89
Figure 4.31. Scatter plot of actual “V” values and “V” values obtained by Local Polynomial (2-D and 10-ft interval case)..... 90
Figure 4.32. Scatter plot of actual “V” values and “V” values obtained by Nearest Neighbor (2-D and 10-ft Interval Case) 91
Figure 4.33. Scatter plot of actual “V” values and “V” values obtained by Polynomial Regression (2-D and 10-ft interval case) 92
Figure 4.34. Scatter plot of actual “V” values and “V” values obtained by Inverse Distance to a Power (2-D and 10-ft interval case) 93

Figure 4.35. Scatter plot of actual "V" values and "V" values obtained by the ANN model (3-D case) 94
Figure 4.36. Scatter plot of actual "V" values and "V" values obtained by Regression-Based Method model 95
Figure 6.1. Map of the Massachusetts Military Reservation 135
Figure 6.2. Map showing the surroundings of the MMR 136
Figure 6.3. Photograph of demolition area one at Camp Edwards 137
Figure 6.4. Illustration of Cape Cod's lenses 138
Figure 6.5. The Cape Cod glacial aquifer 139
Figure 6.6. Perchlorate users in the U.S. 140
Figure 6.7. Perchlorate in public water systems 141
Figure 6.8. National perchlorate detections as of September 2004 142
Figure 7.1. Locations of groundwater monitoring wells at Demo 1 site 159
Figure 7.2. Demo 1 site was divided both in the x and y directions at 25-ft intervals 160
Figure 7.3. Interface of Excel application for ANN profiling model 161
Figure 7.4a. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z = -50 ft 162
Figure 7.4b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z = -50 ft 162
Figure 7.4c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = -50 ft 163
Figure 7.4d. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = -50 ft 163

Figure 7.4e. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = -50 ft 164
Figure 7.4f. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = -50 ft 164
Figure 7.5a. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z = -25 ft 165
Figure 7.5b. Contour map showing the distribution of perchlorate in groundwater in 2001 at.. 165
Figure 7.5c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = -25 ft 166
Figure 7.5d. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = -25 ft 166
Figure 7.5e. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = -25 ft 167
Figure 7.5f. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = -25 ft 167
Figure 7.6a. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z =0 ft 168
Figure 7.6b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z = 0 ft 168
Figure 7.6c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 0 ft 169
Figure 7.6 d. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 0 ft 169

Figure 7.6e. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 0 ft 170
Figure 7.6f. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 0 ft 170
Figure 7.7a. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z = 25 ft 171
Figure 7.7b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z =25 ft 171
Figure 7.7c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 25 ft 172
Figure 7.7d. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 25 ft 172
Figure 7.7e. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z =25 ft 173
Figure 7.7f. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 25 ft 173
Figure 7.8a. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z = 50 ft 174
Figure 7.8b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z = 50 ft 174
Figure 7.8c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 50 ft 175

Figure 7.8d. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 50 ft 175
Figure 7.8e. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 50 ft 176
Figure 7.8f. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 50 ft 176
Figure 8.1. MMR-Demo 1 groundwater monitoring network 195
Figure 8.2a. Baseline contour map showing the distribution of perchlorate in groundwater in 2000 at Z =0 ft (Model 1) 196
Figure 8.2b. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z =0 ft and well number 165 excluded (Model 2) 196
Figure 8.2c. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z =0 ft and well number 32 excluded (Model 3) 197
Figure 8.3a. Baseline Contour map showing the distribution of perchlorate in groundwater in 2001 at Z = 0 ft (Model 1) 197
Figure 8.3b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z =0 ft and well number 165 excluded (Model 2) 198
Figure 8.3c. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z =0 ft and well number 32 excluded (Model 3) 198
Figure 8.4a. Baseline Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 0 ft. (Model 1) 199
Figure 8.4b. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 0 ft and well number 165 excluded (Model 2) 199

Figure 8.4c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 0 ft and well number 32 excluded (Model 3)..... 200
Figure 8.5a. Baseline contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 0 ft. (Model 1) 201
Figure 8.5b. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 0 ft and well number 165 excluded (Model 2)..... 201
Figure 8.5c. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 0 ft and well number 32 excluded (Model 3)..... 202
Figure 8.6a. Baseline contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 0 ft. (Model 1) 202
Figure 8.6b. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 0 ft and well number 165 excluded (Model 2)..... 203
Figure 8.6c. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 0 ft and well number 32 excluded (Model 3)..... 203
Figure 8.7a. Baseline Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 0 ft (Model 1) 204
Figure 8.7b. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 0 ft and well number 165 excluded (Model 2)..... 204
Figure 8.7c. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 0 ft and well number 32 excluded (Model 3)..... 205

List of Tables

Table 4.1 Profiling Methods and Their Corresponding RMSE Value for the 25-ft. Interval Case 58

Table 4.2 Profiling Methods for the 10 ft interval case. 59

Table 4.3 RMSE Values for Profiling Methods used to Predict (V) Values for 3D Case. 59

Table 6.1 Average Amount of Water Drawn off of the Groundwater Lenses of the Cape Cod Aquifer 133

Table 6.2 Properties of Perchlorate Compounds 134

Table 7.1 Optimal Structures of All Trial Cases 157

Table 7.2 Statistical Accuracy Measure for the Nine-Inputs Network 157

Table 7.3 Statistical Accuracy Measures for the Eight–Inputs network 158

Table 7.4 Statistical Accuracy Measures for the Seven-Inputs Network 158

Table 8.1 - Network Statistical Accuracy Measure for the Baseline 9-Input Model (Model 1). 191

Table 8.2 - Network Statistical Accuracy Measures for Model 2 191

Table 8.3 - Network Statistical Accuracy Measures for Model 3 192

Table 8.4 - Network Statistical Accuracy Measures for ANN-Based Profiling Models. 193

Acknowledgements

It is my pleasure to acknowledge the following individuals for their contributions that have helped me bring the work in this dissertation to realization

Grateful acknowledgement is due to Dr. Yacoub Najjar for his gracious support and guidance. I am deeply appreciative to him for sharing his knowledge and for his patience as a teacher. His continuous and step-by-step suggestions and recommendations at different stages made this work a triumph.

I would like to acknowledge my supervisory committee members for reviewing my work and for being in my advisory committee, namely: Dr. Esmaeily from Department of Civil Engineering, Dr. Higgins from Department of Statistics, and Dr. Starrett from Department of Civil Engineering.

I would like to acknowledge Dr. Hutchinson for serving as outside chair for my final examination.

Last but not least, I am indebted to my Wife Laura, who encouraged and supported me through these years and who is always there ready to help.

Dedication

This dissertation is dedicated to my Wife Laura, our three sons Zackery, Adam and Jordan for their support, love, and patience.

CHAPTER 1 - Introduction

1.1 Overview

Environmental site characterization is one of the most crucial and often most expensive components of any environmental remediation process. With the number of hazardous waste sites reaching over 30,000 in the U.S., it is important to have a profiling system that is both accurate and cost-effective. It is estimated that the Environmental Protection Agency (EPA) spends over $150 billion for remediation purposes while the U.S. Department of Energy and Defense will spend approximately $1 trillion over the next 20 to 30 years (Nielsen, 2005).

Environmental site characterization is the process by which a specified area is studied and evaluated for environmental contaminants. This process is the cornerstone of any project, whether it be for risk assessment, monitoring, or remediation purposes. Because of the nature of the situation, it is important that the site characterization is carefully planned and implemented. Inadequate site characterization can lead to a faulty remediation program. According to Nielsen (2005), the most common reasons for the failure of conventional methods of environmental site characterization programs are:

- Inexact or incomplete definition of site geology and hydrogeology, which results in improper positioning of monitoring wells, or selection of inefficient remediation methods.
- Poor definition of contaminant distribution, which results in placement of too few (or too many) monitoring wells to accomplish project objectives, or incomplete site cleanup.

- Inadequate collection of chemical data (i.e., incorrect analyses or wrong detection limits), resulting in monitoring for too few chemical parameters, selection of inappropriate analytical methods, or selection of an inappropriate remedial approach.

Conventional methods of environmental site characterization for subsurface assessment for remediation or monitoring purposes often involve field sampling and laboratory analyses of soil and water samples for specific contaminants species. Even though these procedures are well established and produce reliable results, they have a number of disadvantages. Among others, they are not measured in real time, and they are sometimes destructive because excavations are needed to obtain soil samples. Furthermore, the sampling and testing processes can be quite laborious and expensive. Various investigations have been carried out to develop alternative, nondestructive methods for such routine measurements. One possible method is the application of artificial neural networks (ANN) in environmental site characterization. This method has proved to be an effective modeling method for the prediction of migration paths of environmental contaminants.

A new era of engineering modeling came with the introduction of the ANN technique at the beginning of 1990s by Ghaboussi, et al. (1991). The ANN modeling approach has been receiving increasing favor in the engineering area during the last 15 years. Its massively parallel distributed structure and its ability to learn, and therefore generalize, gives the ANN-based modeling approach the following advantages over a traditional modeling approach:

1. ANN can directly learn relationships and correlations implicitly contained in the data and store the information in its connection weights, which avoids the difficulties arising from the lack of theoretical principles, understanding of the mechanisms, and formulating explicit mathematical expressions encountered in a traditional modeling approach.

2. ANN can describe highly nonlinear relationships.
3. ANN model can be easily improved by learning from new available data and expanding the application scope without creating another model as traditional modeling does.
4. No calibration test is needed.
5. ANN has the ability to extract the correct information from the noisy data and perform gracefully in case of partial damage. This feature is called fault tolerance.

However, the uses of ANN modeling for site characterization of inorganic materials such as heavy metals (i.e. lead, copper) and explosives-related contaminants (i.e. perchlorate) in water and soil have not been widely reported in the literature. For this reason, this research will explore the potential use of neural network modeling for predicting the amount and distribution of inorganic materials and explosives-related contaminants at contaminated sites.

1.2 Problem Statement

1.2.1 Two-dimensional and three-dimensional characterizations of a hypothetical data-rich site by various methods

Over the years, many methods have been developed to profile environmental contaminants in soil media and/or groundwater. These methods vary in their ability to make precise and accurate predictions. This thesis investigates the differences between the following nine profiling methodologies: Inverse Distance to a Power, Kriging, Minimum Curvature, Modified Shepard's, Nearest Neighbor, Polynomial Regression, Radial Basis Function, Local Polynomial, and ANN. Because each method uses an individualized logic, the accuracy of the methods' predicted profiles is expected to vary. To illustrate this, a hypothetical data-rich

contaminated site measuring 1000 ft in the *x* direction by 1000 ft in the *y* direction will be used for the purpose of the comparison. Accordingly, a small fraction of the available data (about 1%) is presented to each method for site profiling. A comparative study of the models' site profiling outcomes/predictions is then performed in order to assess the most accurate site profiling methodology.

1.2.2 Two-dimensional characterizations of the inorganic materials (lead and copper) in soil and groundwater at a landfill site.

At the start of any remediation process, it is important to obtain accurate, in-depth information regarding the extent of contamination at the investigated site. But this is not a simple task. Each site differs in its geological and hydrological makeup, making it very difficult to accurately predict the parameters of the contamination (Najjar, Reddi, & Basheer, 1996). Soil type, characteristics of the underlying material composition, location and depth of groundwater in relation to the investigation site, precipitation, and the topographical structure of the investigated site must be taken into consideration by the investigative team. Most site investigations begin by determining contamination by creating a grid of the area under investigation and obtaining samples at locations in both the x and y directions across the site. In theory, the most accurate information could be obtained by collecting samples at each intersecting point on the grid throughout the entire site. This, however, is not economically feasible.

In order to help investigators determine which locations should be sampled, a number of mapping methodologies have been used. Most of these methods, however, have many constraints that make it difficult to apply the model to more than one specific site. Many different variables, such as temperature, precipitation, and so on, must be assessed and entered

into the model (Rizzo & Dougherty, 1994). Because of the complexity of the required variables, the model must be recalibrated to each specific location. One model that does not require such in-depth calibration is Back-propagation Artificial Neural Networks (BPANN).

BPANN is a system that learns by example. It has the ability to take known data and find a relationship within the given parameters regardless of the complexity of the relationship between the input and output data (Basheer, 1998). In this thesis, a BPANN model will be used to predict the extent and location of lead and copper contaminants within the Kansas City landfill area.

1.2.3 Three-dimensional time-related profiling of explosive-related contaminants (perchlorate) at the Massachusetts Military Reservation site

Due to the nature of its operations, the United States military has long been a major contributor to the contamination of groundwater. With 10,444 operational ranges located in the United States and its territories, the problem of contamination within military-owned land has become worrisome. One of these ranges is Camp Edwards, located on the Massachusetts Military Reservation (MMR). Camp Edwards has been used for military mortar and artillery training exercises since the early 1900s. Because of the lack of environmental regulation prior to 1970, the use and disposal of military munitions went unmonitored, leading to the contamination of Cape Cod's primary source of drinking water, the Cape Cod Glacial Aquifer. However, in 1982, the Department of Defense (DOD) launched investigation and clean-up efforts of contaminated groundwater and soil at the MMR. The investigators found that 79 different areas on the MMR had potential environmental issues.

As of 2004, 77 out of the 79 sources have been addressed, but the cost of the investigation has been devastating (Ogden, 1999.) The fiscal year 2005 budget for clean-up

efforts at the MMR was $25.8 million with a total estimated clean-up cost of $860 million. A large portion of this expense is due to the collection and analysis of groundwater and soil samples.

The cost of sample collection and analysis can vary greatly depending upon the individual in charge of site remediation. Generally, sample locations are determined using the professional judgment of the site investigator team (Najjar, Reddi, & Basheer, 1996). Because each site is unique, there are no specific guidelines regarding the number or location of samples to be obtained at any given site. This can lead to over-sampling in uncontaminated areas, hence increasing the total cost of the remediation process. Fortunately, advances in information-based technology are allowing scientists and engineers to perform their jobs in ways that are both more time-efficient and cost-effective. One approach that is being utilized more frequently is ANN modeling. Accordingly, ANN-based profiling models, once appropriately trained, can predict areas of contamination at their specific site.

This thesis will discuss the development of a neural network model at the MMR site at Camp Edwards. By taking known data from the Demo 1 site, which is located within Camp Edwards, a model can be created to help predict the areas and extent of perchlorate contamination. This information may decrease the number of unnecessary samples collected, therefore potentially decreasing the cost of sample collection and analysis.

1.3 Objectives

The broad objective of this study is to explore the potential use of neural network modeling to predict the migration path and concentration of any environmental contaminants. The overall scope includes the following tasks:

1. Comparing the ANN profiling model performance to eight well-known profiling methodologies.
2. Develop a BPANN for modeling the extent and location of lead and copper in soil and groundwater within the Kansas City landfill area. This will be accomplished by:
    i) Collecting and analyzing samples of groundwater and soil at various locations throughout the landfill site.
    ii) Developing appropriate neural network-based profiling models using results of samples from known locations.
    iii) Developing 2D Profiles that can accurately map the spatial concentration of a specific contaminant.
3. Explore the potential use of neural network modeling for predicting the amount and distribution of perchlorate at the MMR. Experience gained from modeling the lead and copper profiles in soil and groundwater within the Kansas City landfill area will be utilized to develop initial ANN-based models for the MMR site.
4. Utilize data collected in the MMR–DEMO 1 to assess the performance of ANNs for predicting concentrations of perchlorate.
5. Evaluate various criteria for ANN model performance assessment in order to provide useful guidance to water resource managers and others assessing the applicability of a modeling strategy for highly variable water quality parameters.

1.4 Organization of Dissertation

The dissertation consists of nine chapters. Summary of each chapter is as follows:

Chapter 1-Introduction: This chapter presents a brief discussion on environmental site characterization modeling using both a traditional modeling approach and ANN modeling

approach. It addresses limitations and advantages of the traditional modeling approaches and the ANN modeling approach. It presents research objectives and organization of dissertation.

Chapter 2-Literature Review: This chapter first presents a brief literature review for several of the traditional environmental site characterization modeling approaches. The ANN material modeling approach section highlights several publications on ANN modeling that contribute significantly to the success of this research.

Chapter 3-Artificial Neural Network: This chapter presents the study of ANN-based computational algorithms and discusses issues pertaining to the development of an efficient ANN model.

Chapter 4-Two-dimensional and three-dimensional characterizations of a hypothetical data-rich site by various methods: This chapter highlights the differences between eight profiling methods and ANN methodology. This chapter uses a hypothetical data-rich contaminated site to assess the most accurate site profiling method.

Chapter 5 - Two-dimensional characterizations of the inorganic materials (lead and copper) in soil and groundwater at a landfill site: This chapter presents in detail the development of an ANN profiling system to investigate an abandoned landfill site in Kansas City, KS. The developed ANN systems will be trained on existing data and then used to predict the amounts and distribution of lead and copper contaminants within the landfill area.

Chapter 6- Study Area: the Massachusetts Military Reservations. This chapter provides background information on the MMR. This chapter also provides background information on perchlorate, including some of its chemical proprieties, health risks and regulations.

Chapter 7- Three-Dimensional Characterization of the MMR Perchlorate Contaminated Site. This chapter presents in detail the development of a neural network model at the Demo 1 site at Camp Edwards. By taking known data from the Demo 1 site, a model will be created to help predict the areas and extent of perchlorate contamination.

Chapter 8-ANN-Based Profiling: Data Importance. This chapter will explore the use of ANNs to predict the critical number of monitoring wells needed to accurately characterize the extent of the perchlorate contaminations at the Demo 1 site. The purpose of the research in this chapter is to develop a tool that can be used at some later time to evaluate the effectiveness of the current groundwater monitoring network in regards to the Demo 1 explosive-related contaminant plume.

Chapter 9-Summary, Conclusions and Recommendations: This chapter presents the important conclusions obtained from this research study and points out few recommendations for future research studies.

CHAPTER 2 - Literature Review

2.1 Application of Artificial Neural Networks in Environmental Site Characterization

Modeling contaminant behavior constitutes a critical part in the design and analysis of geoenvironmental systems. Several modeling methodologies have been (and continue to be) developed to characterize complex contaminant behavior. The role of all these modeling methodologies is to characterize multiple behaviors of contaminants such as inorganic materials (i.e., lead, copper, and perchlorate), microbial indicators of fecal contamination, agricultural contaminants and explosives-related contamination. The literature reviewed in this chapter addresses several topics related to environmental site profiling modeling and ANNs.

One of the first steps in the remediation process is to determine the characteristics of the contaminated site. This includes not only obtaining historical and geological information about the site but, most importantly, determining the location and concentrations of contaminants at the site. This is done by collecting samples at selected points throughout the area of concern and analyzing the samples to determine the concentration of the contaminants. Each sample will have its own unique set of data, which includes the location of the sample (latitude, longitude and depth) and a concentration. With this information, a detailed map of the contaminated site can be created.

2.2 Traditional Methods

Over the years, traditional methods for environmental site characterization have been relied upon to predict the amount and distribution of pollutants within the contaminated areas.

One such profiling methodology commonly utilized is the resistivity method. Rosqvist et al. (2003) employed this method for profiling leachate contamination at municipal landfill sites. This method is based on the knowledge that leakage from municipal solid waste deposits tends to have high ion concentrations and low resistivities, allowing for geoelectrical imaging. Along with soil and water analysis, this method was used to study the geoelectrical measurements at two landfill sites in South Africa and to map the leachate contamination of the sites. The geoelectrical imaging technique used in the study was described as follows:

> "The resistivity method is based on measurement of the potential distribution arising when electric current is transmitted to the underground via electrodes. The data acquisition was done as two-dimensional (2-D) resistivity imaging, using the ABEM Lund Imaging System in a version that also allowed measurements of time-domain induced polarization (IP) data in ten time windows. The system is computer controlled and consists of a resistivity-IP instrument, a relay-switching unit, four electrode cables, connectors and steel electrodes. The 2-D imaging layouts used comprise around 80 electrodes, and measurement lines can be expanded via a roll-along technique. A gradient array electrode configuration was used in order to get good resolution (Dahlin & Zhou, 2003). The measured data was processed with inverse numerical modeling (inversion) to produce model cross-sections of the resistivity and chargeability distribution of the ground using software Res2dinv."

After conducting studies at both landfill sites, the authors found that the maps of the leachate plumes obtained by the resistivity measurements were in agreement with earlier studies of the sites in terms of the extent of the plumes (Dahlin, 2004). Images created by the mapping system showed the differentiation between the layers of soil as high resistivity, intermediate

resistivity and low resistivity. The contents of each layer were confirmed by drilling surveys. The authors found the mapping of the sub-surface leachate plume migration at the landfills to be "successful" and "advantageous." They wrote:

> "The groundwater quality measurements correspond well to the geoelectrical measurements, bringing together a good picture of the extent of the leachate plumes. The leachate plumes clearly indicated by previous investigations have been confirmed by the interpretation of the resistivity data at both sites. Also, the extent of the leachate plumes as mapped by the geoelectrical imaging corresponds fairly well to the development of the leachate plumes reported in previous investigations."

From their studies, we can conclude that the resistivity method is one that could produce reliable information for the remediation of leachate in plumes while saving time and money when compared to traditional site sampling methods.

Another traditional methodology often utilized is Iso-Surface. This method can be used to create contour maps of a contaminated site. Jones and Davis (1996) explored the use of an interpolation scheme to generate a continuous 3-D function that represents a contaminated plume. The function was of the form $c=f(x,y,z)$, whereas, at any given location (x,y,z) the function returns an estimate of the concentration of the contaminant at that point. These estimated concentrations along with the sampling values could be entered into the nodes of a three-dimensional grid, which can then be used to generate a three-dimensional plot of the plume using a software program called Iso-Surface.

According to Jones and Davis (1996):

> "Iso-surfaces are the three-dimensional equivalent of two-dimensional contour lines. Just as a contour, or *iso-line,* represents a constant value of a two-dimensional

function, an iso-surface represents a constant value of a three-dimensional function. In the case of plume characterization, a threshold concentration value is selected to compute the iso-surface. The volume inside the surface represents the region where the concentration is greater than the threshold value, and the volume outside the surface represents the region where the concentration is less than the threshold value."

When discussing how the Iso-Surface program functions, Jones and Davis (1996) state: "Iso-surfaces are typically constructed using a "marching cubes" algorithm. A marching cubes algorithm constructs an iso-surface by processing each cell in the three-dimensional grid independently of the other cells. The concentration values at the eight corners of the grid are compared to the iso-value to determine if they are greater, equal to, or less than the iso-value. Depending on the status of the cell corners, a set of small triangles representing a portion of the iso-surface is created and added to a list of triangles. Once each of the cells has been processed, the iso-surface is represented by the entire set of triangles generated in this fashion."

Although programs such as Iso-Surface can be very beneficial in the mapping of contaminants, it is essential that the interpolation algorithms be fully understood to ensure that the data is interpreted correctly. Jones and Davis (1996) state:

"This interpolation process is the most critical step in the plume visualization process. In most cases, the data are sparse and the choice of interpolation scheme can have a dramatic effect on the results. There are several unique problems associated with three-dimensional interpolation of contaminant data that must be understood in order to ensure

that the resulting iso-surfaces are a reasonable and accurate interpretation of the measured concentrations. These problems include improper inference of maximum concentrations, negative concentrations, oscillations, data clustering, and problems associated with Kriging."

By being aware of such problems and making the necessary adjustments, programs such as Iso-Surface can effectively and accurately plot the areas of contamination.

2.3 Artificial Neural Networks Methodology

Conventional methods of subsurface assessment for remediation or monitoring purposes often involve field sampling and laboratory analyses of soil and water samples for specific contaminant species. Even though these procedures are well established and produce reliable results, they have a number of disadvantages. Among others, they are not measured in real time, and they are sometimes destructive because excavations are needed to obtain soil samples. Furthermore, the sampling and testing processes can be quite laborious and expensive. Various investigations have been carried out to develop alternative, nondestructive methods for such routine environmental site characterization. The application of ANNs in environmental site characterization has proved to be an effective modeling method for the prediction of migration paths of environmental contaminants. Therefore, many researchers are utilizing the ANN to predict areas and the extent of contamination at a given site.

According to Adeli (2001), ANNs are becoming increasingly useful in Civil Engineering. Adeli (2001, p.132) reviews many experiments in which ANNs have been particularly useful. For instance, in 1994, Karanthi used ANNs to successfully predict river flow (Adeli, 2001, p. 132); meanwhile, Du, et al., (1994) were able to use a back-propagation algorithm in their ANN to predict solubilization levels of heavy metals in sewer sludge (Adeli, 2001, p. 132). Then, in

1995, Grubert used an ANN to predict "the flow conditions at the interface of stratified estuaries and fords." Meanwhile, Crespo and Mora used ANNs to help estimate the flow of streams and to help predict carbon dioxide concentration (Adeli, 2001, p.133). Three years later, Thirumalaiha and Deo used the ANN to forecast river flow in real time.

In 1996, Kao and Liao used an ANN to help select solid waste sites, and in 1997 Tawfik, et al., used an ANN to model "stage-discharge relationships at stream gauging locations at the Nile River (Adeli, 2001, p.132). During the same year, Deo et al. used ANNs to predict the height of ocean waves over short periods of time.

Adeli also reports increasing use of ANNs in the late nineties. In 1996, Basheer and Najjar used the networks to "model fixed-bed absorber dynamics." Meanwhile, Rodriguez and Serodes used the back-propagation method to determine what dosage of disinfectant ought to be applied to re-chlorinated water. In the same year, Maier and Dandy used ANNs for a variety of purposes, including water salinization prediction. Meanwhile, Deo and Chaudhari used the networks to predict tides inside estuaries and bays (Adeli, 2001, p. 133).

Gangopadhyay, et al., used a graphic information system combined with a back-propagation ANN to generate profiles of subsurfaces and to identify the distribution of subsurface materials in 1999 bays (Adeli, 2001, p. 133). In 2000, Liu and James used a BPANN to "estimate the discharge capacity" in two-stage channels. Guo used the ANN to model watersheds for urban areas with small drainage areas. (Adeli, 2001, p.134).

2.3.1 ANN in Water Contamination Profiling

This following section will review the use of ANNs for water contamination profiling. Tabach used an ANN to analyze groundwater contamination in a road project (Tabach, 2007, p. 766). He did so by training the network to estimate the depth of the contaminated zone and the

volume of soil pollution infiltration. Tabach built his database by using as inputs the following: cover layer permeability, cover-layer thickness, water-table depth, and soil-pollutant contact time. His outputs, then, are the depth of contaminated soil and infiltrated pollutant quantity (Tabach, 2007, p. 767). Tabach's model based on ANNs is able to successfully assess the contamination of unsaturated soil by a trichloroethylene spill in a road accident.

Besaw (2006) used an ANN as an alternative to traditional time-consuming Kriging methods. In contrast to traditional methods, Besaw reports that ANNs are a cost-effective, reasonably accurate, and speedy alternative. Rather than relying on linear, mathematical models, Besaw creates non-linear maps of "statistical relationships between multiple variables" (2006, p. 5). Because traditional Kriging methods are complicated and time-consuming, adding input variables can greatly increase the time needed to complete analysis. In contrast, Besaw found that adding input to ANNs increases their performance.

According to Besaw, the ANN can make accurate predictions with only a limited amount of data. To demonstrate this, the team created a counterpropagation ANN, which "self-adapts to create statistical mappings of predictor and associated response vectors" (2006, p.1). Besaw and his team trained their system with previously collected hydroconductivity data, using inputs of x and y. Trained on the patterns from the hydroconductivity data, the system functions as a lookup table, and is then able to create non-linear maps to predict future events. Besaw reports the ANN to be sufficiently accurate.

Li (2006) used an ANN to address the problem of subsurface contamination by light non-aqueous phase liquids (LNAPLs). In his study, Li and his team use an ANN to simulate a hydrocarbon recovery process at a petroleum-contaminated site. Traditionally, environmental scientists use dual-phase or multiphase extraction methods to obtain samples, but such methods

are time-consuming. Therefore, Li tested ANNs as an alternative. The team used groundwater extraction rate, vacuum pressure, and saturation hydraulic conductivity as inputs and cumulative hydrocarbon recovery volume as an output. They used the data to train their network, "from implementation of a multiphase flow model for dual phase remediation process under different input variable conditions" (Li et al., 2006, p.1).

After training their model, Li and his team were able to forecast cumulative oil volume under a variety of conditions. Using some data sets for training and others for verification, they were able to ensure the accuracy of their model. Li's team found that ANN was reasonably able to detect contaminants (Li, et al., 2006, p.1).

2.3.2 ANN in Soil Contamination Profiling

The following section will review the use of ANNs for soil contamination profiling. Amegashie (2006) used an ANN to determine the locations of heavy metals, by using permittivity measurements (Amegashie, et al., 2006, p.2). Because permittivity can be measured over a large range of frequencies, large databases can be created from permittivity data, which an ANN can then use to create maps and models (Amegashie, et al., 2006, p.2).

Amegashie's team developed two models. They used one model, the ANN-M8, to detect the presence of metals. The team trained the ANN with 164 samples and designed it to output "yes" when heavy metals are present in a sample, and "no" when no metals are present. The team found that the ANN-M8 accurately classifies and detects copper, zinc and lead. Indeed, Amegashie indicates that the ANN-M8 is able to identify 46 of 52 samples lacking heavy metals and 57 of 60 samples containing heavy metals (Amegashie, 2006, p.6). To design their network, Amegashie's team used 10 neurons in the ANN-M9's input layer, eight in the hidden layer and three sets of training datasets. According to Amegashie, the ANN-M9 (created to classify the

type of each metal) is not as accurate as the ANN-M8. Nevertheless, the ANN-M9 is still able to correctly classify 13 out of 15 samples containing copper, 21 out of 25 samples of lead, and one of two samples containing zinc (Amegashie, 2006, p.6).

Juang et al. (1997, p.168) designed an ANN to establish a "realistic working profile of soil properties." Because the number of boreholes and soil tests is limited, using traditional methods can be inaccurate. Yet, Juang et al. concluded that an ANN can accurately predict SPT N values based on data from limited boreholes (1997, p. 172). To train their network, Juang's team took data from six boreholes. They used the depth of standard penetration tests (SPTs) and the locations of boreholes as input data, and SPT-N as the output (1997, p. 169). Then, they used the back-propagation method to train the system. Through trial and error, Juang et al. decided that the number of hidden neurons ought to be six. They used the Levenberg-Marquardt algorithm to train their network, thereby training until they reached their error goal (Juang, et al., 1997, p. 169).

Najjar took Juang's ideas a step further. Rather than using SPT data, Najjar chose to test the ability of the ANN to predict cone penetration test (CPT) results (Najjar 2002, p.901). Although SPTs are used more often, CPTs, according to Najjar, are used for soft clays and medium to coarse sands. These tests can provide data on the density, angle of friction, soil stratification, and bearing capacity of an area (Najjar, 2002, p. 901).

Traditionally, scientists have used mathematical models to analyze such data, but these models usually only allow the examination of one attribute at a time. This necessarily restricts data that might be useful to environmental scientists. Specifically, says Najjar, they tend to exclude what could be relevant data. Therefore, they are not suited for 3-D profiling. Najjar offers the ANN as an alternative. The ANN, he says, can show changes along different soil

stratification, while predicting more than one variable at a time. An ANN, he says, can also show how one variable affects another (Najjar, 2002, p. 902).

Najjar takes data from CPT tests related to the following attributes: friction resistance, tip resistance, and excess pore pressure. He then creates four different networks. Every network uses xyz coordinates as its inputs, but each network has different outputs. NWA's output is skin friction, NWB's is tip resistance, and NWC's is pore pressure. NWD, NWE and NWF are given different combinations of outputs to demonstrate how variables might affect one another. By experimenting in this manner, Najjar finds ANNs are able to accurately and simultaneously predict values for skin resistance and pore pressure at any given point (Najjar, 2002, p. 903).

Buzewski and Kowalkowski (2006, p.598) used "perceptrons" combined with an ANN to create a model for heavy metal transport.

For the input layers, Buzewski and Kowalkowski used the initial concentrations of metals in a contamination solution, redox potential, and pH in acid rain, soil properties and TC and IC carbon content in leakage. The outputs of the ANN are the concentrations of lead, nickel and cadmium (Buzewski & Kowalkowski, 2006, p. 598). Buzewski and Kowalkowski (2006, p. 591) use 50% of the data as learning material and 25% for validation. The success of the ANN's identification leads Buzewski and Kowalkowski (2006, p. 596) to declare that, "The ANN seems to be the future tool for modeling transport of inorganic substances in real soil profiles."

CHAPTER 3 - Artificial Neural Networks

3.1 Introduction

The concept of ANNs was first conceived in 1943 by Warren McCulloch. He was a neuroscientist that studied how the brain could produce highly complex patterns by using many basic cells that were connected together. Russell & Norvig (1995) identified the ANNs as mathematical models and algorithms designed to mimic the information processing and knowledge acquisition of the human brain.

Basheer (1998) stated that ANNs are capable of learning by example. Especially when there are highly nonlinear or complex unrecognized governing relations describing the available data sets, parameters in the data might or might not be mathematically related to each other. Fausett, (1994) indicated that an ANN has the ability to capture the relation among these parameters regardless of how strongly they are related. In this case, ANN can dynamically process and recognize complex patterns that relate the provided input data variables to the output data variables, and then precisely provides an efficient input-output mapping. ANNs are often good at solving problems that are too complex for conventional technologies (e.g., problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found).

3.2 Elements of an Artificial Neural Network

The basic "architecture" of the neural network refers to its arrangement of interconnections between the neurons and layers. Najjar, et al., (1996) categorized the typical

arrangement of the neural network as follows: neuron, input layer, hidden layer(s), output layer, and connection weights. A schematic architecture of ANN and its four parts is shown in Figure 3.1. A brief description of each part is given in this section.

3.2.1 Neuron

The basic building block of the network system is the neuron, the cell that communicates information to and from the various parts of the body (Figure 3.2 shows a neuron with its different constituents). Abrahart & See (2000) stated that that the biological neuron consists of three main parts, namely:

- A cell body called the soma
- Several spine-like extensions of the cell body called dendrites
- A single nerve fiber called the axon that branches out from the soma and connects to many other neurons

The axons and dendrites are considered to be responsible for transmitting signals to the neuron. Figure 3.3 represents an artificial neuron in its simplest form. The incoming lines in Figure 3.3 represent dendrites. Each line carries a signal from another neuron. The body represents the soma and the output represents the axon, which in its turn branches to interconnect with other neurons (Ham & Kostanic, 2000). All artificial neurons interconnect to each other form what is called an ANN. The McCulloch-Pitts neuron (Figure 3.4) is the most commonly used neuron model.

According to Ham & Kostanic (2001), each artificial neuron forms a node in the larger neural network and is constructed of the following basic elements:

- Synapses or connection links send input from one node to another in an ANN. Each synapse has its own weight or strength. A positive weight indicates an excitatory

synapse; a negative weight indicates an inhibitory one.

- An adder or linear combiner sums the weighted input signals from other nodes transmitted via the synaptic connections.
- The activation function limits the amplitude of the output of the artificial neuron. The activation functions, which are described in greater detail in section 3.6, can be continuous-values, binary (with range [0,1]) or bipolar (with range [-1,1])
- A bias (θ_i) may also be present. The bias increases or decreases the net input of the activation function.

The neuron has one or more inputs and produces one output. The inputs simulate the stimuli/signals that a neuron gets, while the output simulates the response/signal which the neuron generates. The output is calculated by multiplying each input by a different number (called weight), adding them all together, then scaling the total to a number between 0 and 1.

3.2.2 Input Layer

Anderson & McNeil (1992) indicated that the input layer is the least complex of all the layers because no mathematical calculations occur at this level. Before beginning, the number of inputs and relevance of the inputs must be decided. Inputs that are believed to have no relevance on the output should be eliminated. Thus, available input data that affect the output are fed to the network. The performance of the network depends on the number of inputs. The input layer receives and processes information and forwards it to the hidden layer.

3.2.3 Hidden Layer(s)

It is at this layer that all the calculations occur. The numbers of hidden layers vary but there is always at least one hidden layer in every network. Each layer is composed of a set of neurons. Each layer is interconnected in such a way that the first layer passes information to the second layer, the second layer to the third, and so forth (Huang & Dong, 1992). This is done via

connection weights (Figure 3.5). Connection weights connect each neuron in a certain layer to every single neuron in the next layer. The value of that weight is responsible for adjusting the output value of the neuron.

Each processing element in a specific layer is fully or partially connected to many other processing elements via weighted connections. The scalar weights determine the strength of the connections between interconnected neurons. A zero weight refers to no connection between two neurons and a negative weight refers to a prohibitive relationship. From many other processing elements, an individual processing element receives its weighted inputs, which are summed, and a bias unit or threshold is added or subtracted. The bias unit is used to scale the input to a useful range to improve the convergence properties of the neural network. The result of this combined summation is passed through a transfer function (e.g. logistic sigmoid or hyperbolic tangent) to produce the output of the processing element.

3.2.4 Output Layer

Najjar & Basheer (1996) indicated that the output layer in a network is a layer containing one or more output neurons. An output neuron will compute a value for a certain parameter or variable.

Input value to node *k*: $I_k = \sum_j W_{jk} O_j$ *3.1*

The output values O_j that leave a node j on each of its outgoing links are multiplied by a weight, wj. The input I_k to each node k in each middle and output layer is the sum of each of its weighted inputs, $w_{jk}O_j$ from all nodes j providing inputs (linked) to node k.

3.2.5 Connection Weights

Connection weights are the interconnecting links between the neurons in the layers constituting the network. Each neuron in a certain layer should be connected to every single neuron in the next layer by a connection weight. The value of that weight is responsible for adjusting the output value of the neuron. The magnitude of the weighted connection is directly related to the strength of that connection (Romaniuk, 1995). Signals travel between neurons over weighted connection links. The weight assigned to the connection is multiplied to the signal that is transmitted. Each connection link has an associated weight, which, in a typical neural network, is multiplied to the signal that is transmitted. The process of training a neural network involves the adjustment of the weights based on the given learning rule (Ham & Kostanic, 2000). The overall net input consists of the sum of the weighted connections (product of the weight times the signal).

3.3 Back-Propagation Neural Networks

Hertz & Palmer (1991) defined Back-Propagation Neural Networks (BPNN) as multi-layered, feed-forward neural networks trained using a back-propagation of error algorithm. BPNN development began in the 1970s, and it has become one of the most highly employed systems among the engineering and scientific community. Data is entered into the program to train it. Once the program has been trained, it can then be used to predict certain outcomes. This ability to predict or simulate a given situation has led to an increased application of ANNs in the areas of science and engineering (Sarle, 1994).

Figure 3.6 schematically illustrates the structure of a typical multilayer BPNN. In this figure, BPNN consists of an input layer with three neurons, a hidden layer with two neurons and output layer with two neurons. In order to calculate the output of a neuron at the output layer, the

input must pass through a Sigmoidal function (i.e., transfer function) that is the most widely used function in various BPNN applications. The produced or obtained outputs are then compared to actual outputs (i.e., target vectors) to evaluate the error. Consequently, this error is used to calculate an error function (Hanson, 1995). The resulting error function is used to propagate the error starting from the weights connected to the last layer (output layer) and backward to the input layers (that's the reason why it is called back-propagation of error) in order to modify the weights. In other words, the error generated by a network is used to adjust the weights of the connections.

The interconnection weights are not known initially, and thus are given some initial random (guess) values. The solution obtained using these weights might be far from the target values. Therefore, the correction propagates backward starting from the output layer and then from that hidden layer backward to the input layer (Hanson, 1995). The input is forwarded once again to produce new output values and consequently, new error is calculated to adjust the connection weights. The procedure of forward activation of signals and the back-propagation of errors is repeatedly carried out on all available training data sets until the error at the output side reduces to a prespecified minimum or a permissible tolerance (Najjar et al., 1997; Najjar & Zhang, 2000). The final connection weights are then stored to represent the network structure, which can be used later to predict outputs when presented with new data sets where actual output values are not available.

Feed-Forward Stage

Most applications of BPNN usually incorporate network architectures with only one single hidden layer, because one hidden layer is found sufficient in providing continuous and

nonlinear mapping between input and output patterns. When the network receives the input signals through the input nodes, the normalized values of input parameters are forwarded to the hidden layer. The same procedure is repeated from the hidden layer to the output layer. When the signals from the hidden neurons reach the output layer, the accumulated weighted signals can be obtained at output neuron. Then, the predicted normalized output(s) is (are) obtained.

3.4 Back-Propagation of Error

The objective of the training process is to adjust the connection weights in order to minimize the Averaged Square Error (ASE). This can be accomplished by utilizing the back-propagation algorithm, which provides a correction to every connection weight. Accordingly, the final ASE values are used to compare the performance of the network after every specific number of iterations; hence, the best performing network can be selected. Major features of the back-propagation algorithm are:

1. Learning occurs during a training phase.
2. Each input pattern in a training set is applied to the input units and then propagated forward.
3. The pattern of activation arriving at the output layer is then compared with the correct (associated) output pattern to calculate an error signal.
4. The error signal for each such target output pattern is then back-propagated from the outputs to the inputs in order to appropriately adjust the weights in each layer of the network.

3.5 Transfer Functions

The activation function, which is sometimes called the transfer or squashing function, is applied to the net input received by a node. Activation functions can be linear or non-linear and the output or range of the activation function is usually 0 to 1 or -1 to 1 (Haykin, 1994; Ham & Kostanic, 2001; Masters, 1993). There are several types of activation functions available for use in ANNs, and the commonly used ones are briefly discussed in this section.

3.5.1 Sigmoidal Function

Sigmoid, or S-shaped functions, are the most commonly used activation functions in ANNs (Masters, 1993; Reed and Marks, 1999; Ham and Kostanic, 2001). It is a continuous activation function, designed to respond relative to the amount of excitation received. This function is schematically shown in Figure 3.7. Mathematically, it is represented by the following equation:

$$f(input) = \frac{1}{1 + e^{-(input)}} \qquad 3.2$$

3.5.2 Hard Limiter Function

A hard limiter function can have only two values: zero or one. This type of function is used in applications where we only need ON/OFF or 1/0 outputs. This function is characterized by a threshold value, v_i. Mathematically, this function is represented as:

$$y_i = f(v_i) = \begin{Bmatrix} 1 \ if \ v_i \geq 0 \\ 0 \ if \ v_i < 0 \end{Bmatrix} \qquad 3.3$$

Figure 3.8 shows a plot of the hard limiter function.

3.5.3 Threshold Logic Function

In this type of function, the output varies between zero and one, but the relation between these two values is a linear one. The width of this interval is represented by a parameter α ; this interval starts at ϑ and has a width of $1/\alpha$ (Zupan & Gasteiger, 1993) as shown in Figure 3.9.

3.6 Model Implementation

The implementation phase of model development consists of learning or training and validation. Reed and Marks (1999) define training as the process by which the ANN adapts to learn the relationship or mapping between inputs and outputs. Learning processes consist of supervised, unsupervised, and reinforced learning and its success is typically measured by some performance metric. Validation is the testing of the model with input data that was not used to train the model in order to assess its ability to generalize the relationship between input and output data.

3.6.1 Supervised Learning

In supervised learning, the network is provided with correct answers to the problem for every input pattern. The connection weights of the network are adjusted to allow the network to produce answers as close as possible to target (teacher) answers.

With supervised learning, the ANN must be trained before it becomes useful. (Babovic & Bojkov, 2001) indicated that the training consists of presenting input and output data to the network. This data is often referred to as a training data set. That is, for each input provided to the network, the corresponding desired output set is provided as well. It is considered complete when the neural network reaches a user-defined performance level. This level indicates that the network has achieved the desired statistical accuracy as it produces the required outputs for a

given sequence of inputs. When no further learning is necessary, the weights are typically frozen for the application.

In supervised learning, there is an output or target specified for every input used in the training process. Pairs or samples are used during training input-output. The input consists of a vector of real numbers, with each element of the vector corresponding to an explanatory variable (Rojas, 1996). For example, in a site profiling modeling application, the elements of an input vector could be precipitation, groundwater elevation, and streamflow. Each input is propagated through the ANN and the model output is compared to the target data. The target data is also a vector of real numbers that gives the values of the variables being modeled by the ANN. Unless the model is perfectly trained, there will be differences between target data and the ANN output. The goal of the training process is to optimize the ANN to minimize the differences between ANN output and target data values by adjusting or updating the weights between nodes.

3.6.2 Unsupervised Learning

In unsupervised learning, during the training process no sample outputs are provided to the network against which it can measure its predictive performance for a given vector of inputs (Rojas, 1996). The network internally monitors its performance. It looks for regularities or trends in the input data set, and makes adaptations accordingly. Even without being told whether it is right or wrong, the network still must have some information about how to organize itself. This information is built into the network topology and learning rules.

3.6.3 Reinforcement Learning

The third type of learning is *reinforcement* learning. This is a special case of supervised learning in which the network is provided only with a critique on the *goodness* of network outputs for a given input pattern rather than true answers.

3.6.4 Training of a Network

The training of a network begins by:

1. Making an initial choice of the suitable neural network structure (or architecture),
2. Assigning initial random small values for the connection weights to calculate the output
3. Finally, selecting a learning rate, which can appropriately control the adjustment rate of the connection weights.

The training procedure is repeated until the actual and calculated outputs agree within some pre-determined tolerance. In other words, the network stops learning when weight adjustment produces no improvement in the output values. Training is performed in order to determine the best possible values of connection weights for further use as a prediction tool (Najjar, 1999; Najjar et al., 2000). In this research, the process of training and on-line testing was repeated thousands of times for networks with different numbers of hidden nodes. Hundreds of networks were developed and then compared in order to select the one with optimum performance.

Neural Networks can reach a least-error structure by training, using examples related to the problem under consideration. A least-error structure is the one responsible for producing outputs very close or equal to the real desired values (Jain, et al., 1996). Reasonable training input and output vectors should cover a wide range of the sampling domain. Deriving an

appropriate and representative mapping between input and output vectors reflects the effectiveness of neural networks. For proper modeling, a network should at least pass through two stages, namely training and testing stages. Selected data with their input and output values are introduced to a network (having a certain number of hidden nodes and layers) so that the network trains itself to produce output values that are as close to the real values as possible. The training is achieved by modifying the values of the connection weights. The network stops learning when adjusting the weights produces no improvement in the output values. The same network should be tested on data that was never used in training in order to verify the network's generalization capabilities. The procedures of training and testing should be repeated for networks having different numbers of hidden layers and/or hidden nodes. Changing the input parameters and the number of outputs also affects the performance of a network. This is why at this stage, we will have hundreds or thousands of networks to compare and select the one with the optimum performance.

3.7 Accuracy Measures

Generated networks are compared by their performance (i.e., accuracy) parameters. These parameters are the Averaged Square Error (ASE), coefficient of determination, known as R-square (R^2), and the Mean Absolute Relative Error (MARE %). The ASE value can be calculated by the formula:

$$ASE = \frac{\Sigma(y'-y)^2}{\#\ of\ \ data\ sets} \qquad 3.4$$

y' being the output generated by the network and y being the real value of the parameter.

The MARE value is calculated using the formula:

$$MARE(\%) = \frac{\frac{\sum |y' - y|}{y} \times 100}{(\#\ of\ outputs) \times (\#\ of\ sets)} \qquad 3.5$$

Generally, we search for the network that produces the minimum values of ASE and MARE% and the highest R^2. Testing performance parameters should be considered to select the best performing network. Training performance measure may used in special cases, if needed.

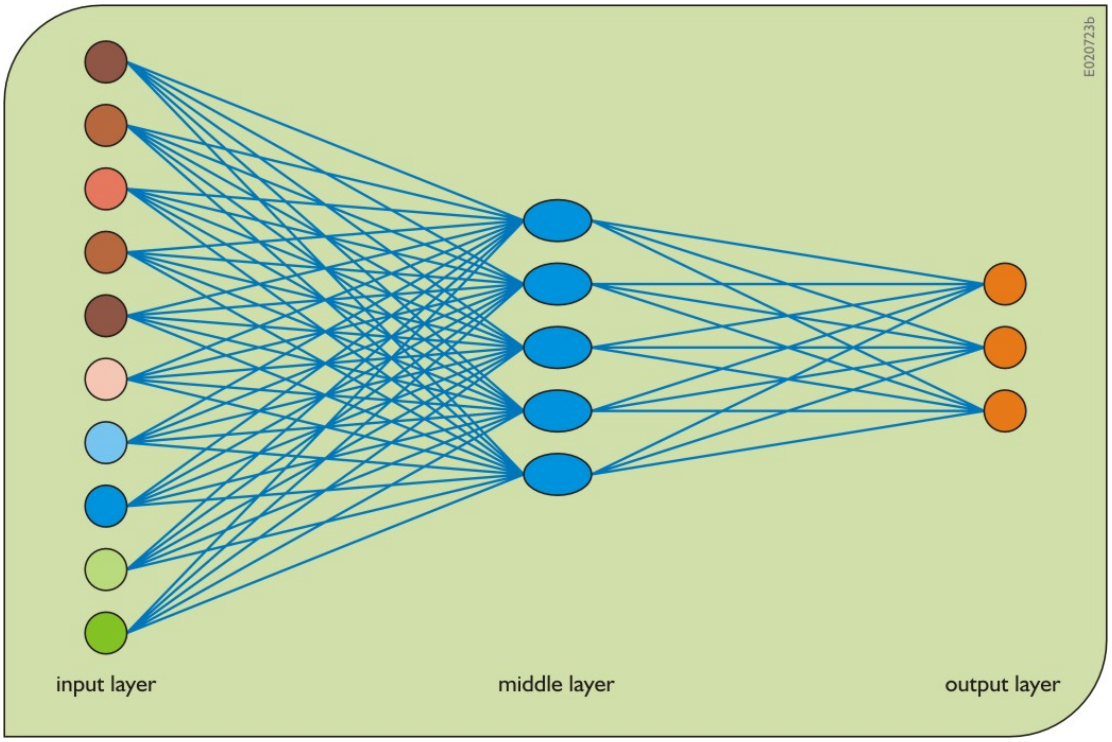


Figure 3.1. A typical multi-layer ANN showing the input layer for ten different inputs, the middle or hidden layer(s), and the output layer having three outputs

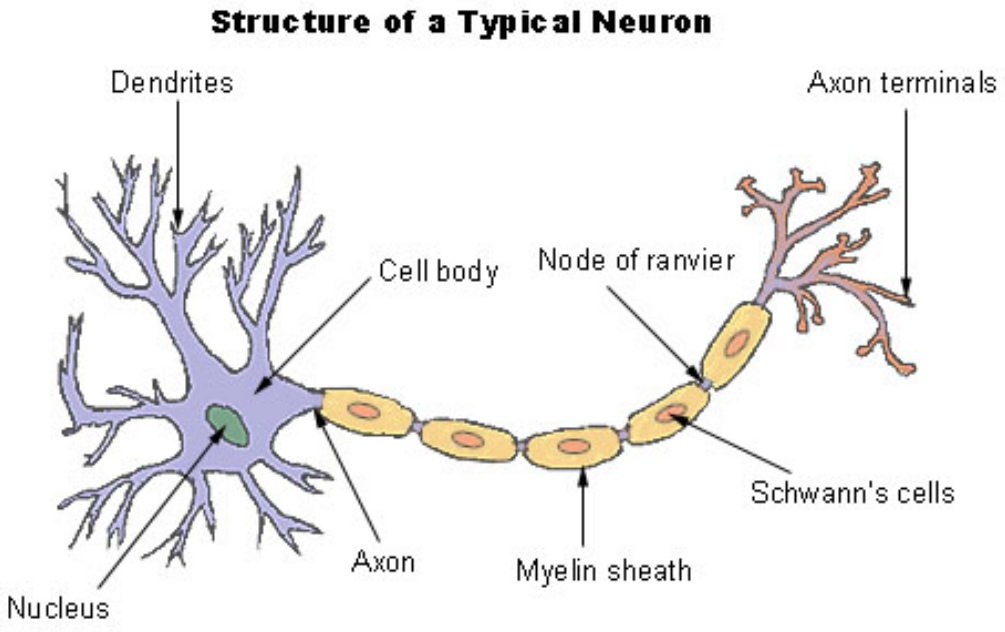


Figure 3.2. The basic building block of the network system, the neuron

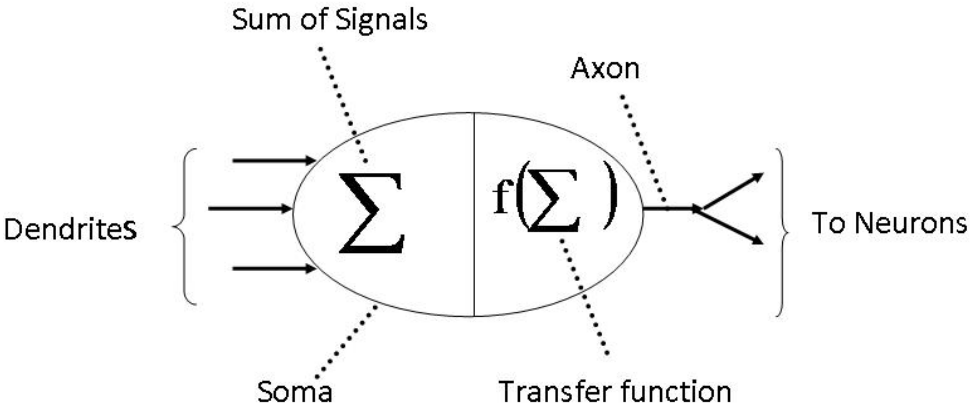


Figure 3.3. Artificial neuron in its simplest form

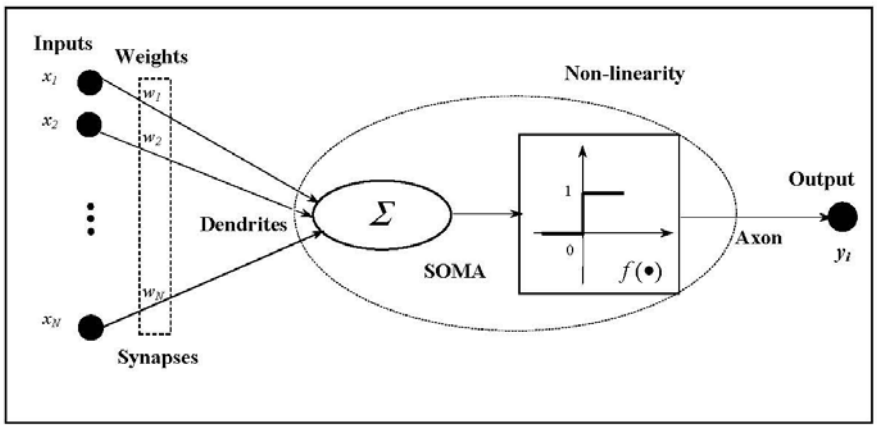


Figure 3.4. McCulloch-Pitts neuron model

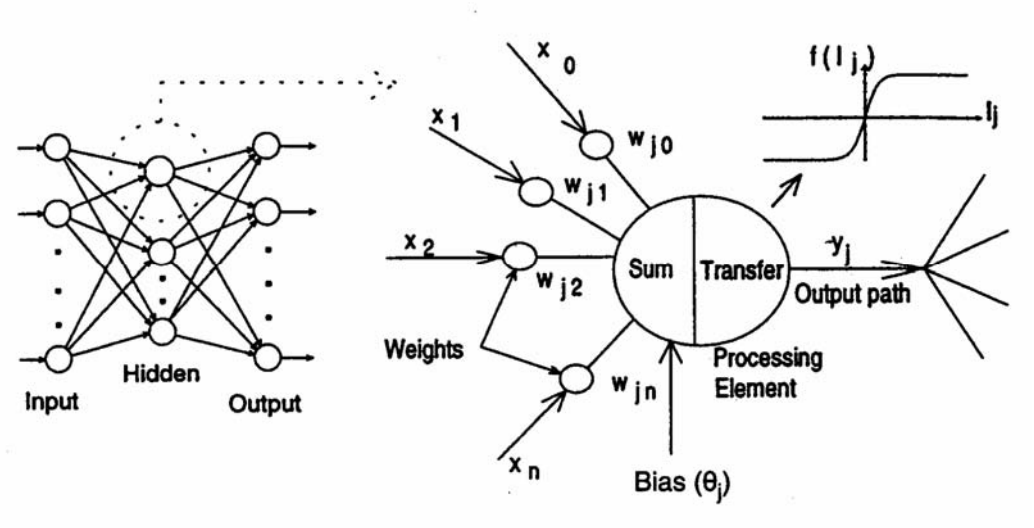


Figure 3.5. ANN with hidden layer

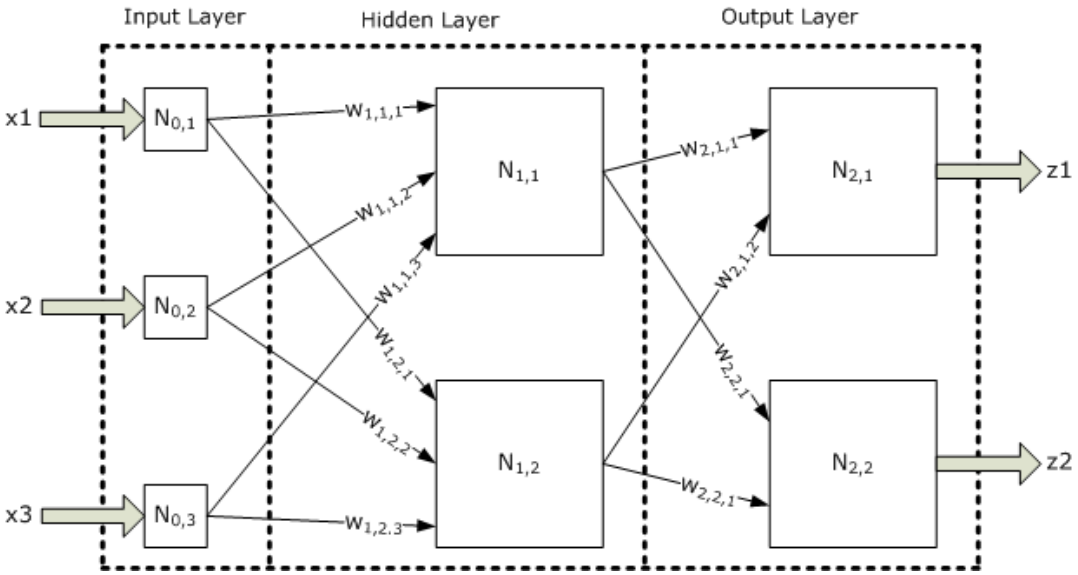


Figure 3.6. Diagram shows a back-propagation neural network

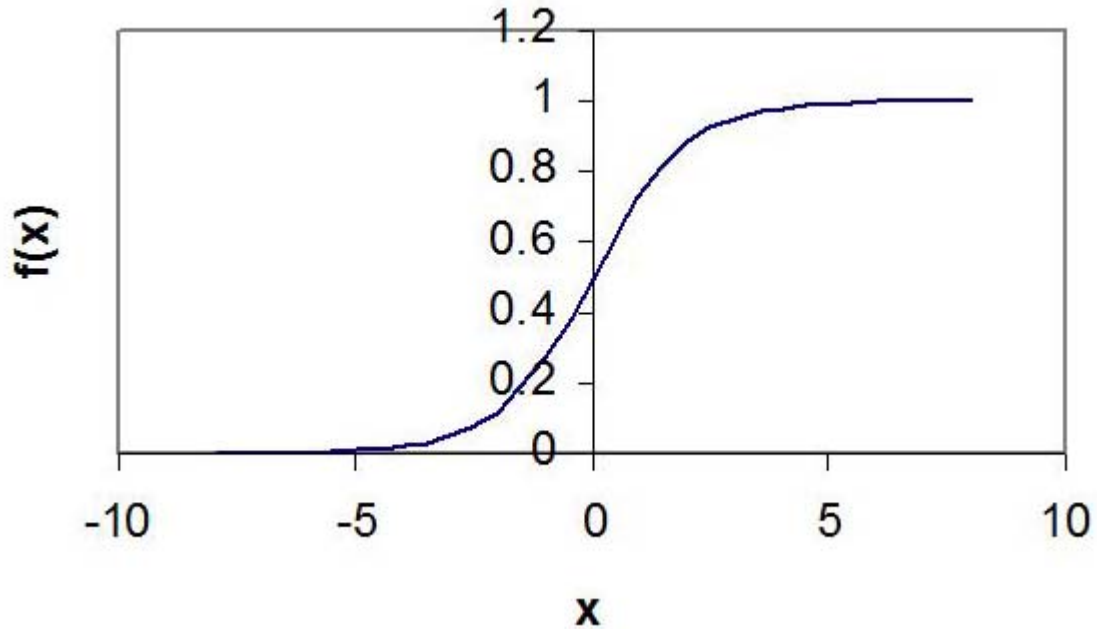


Figure 3.7. Sigmoidal function

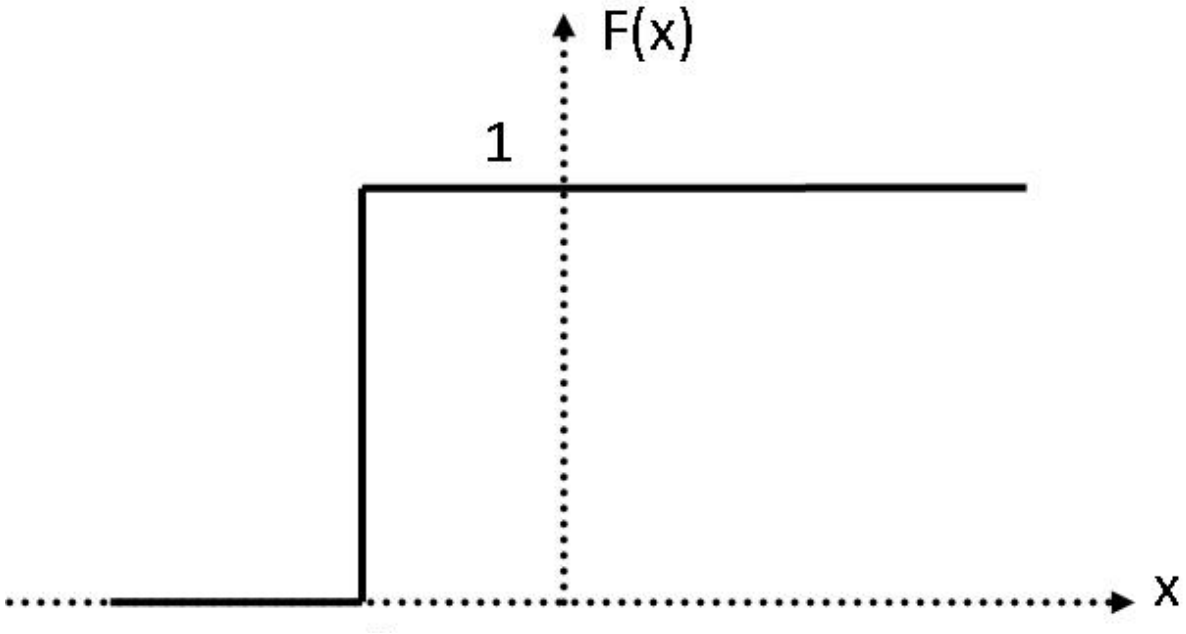


Figure 3.8. Hard limiter function

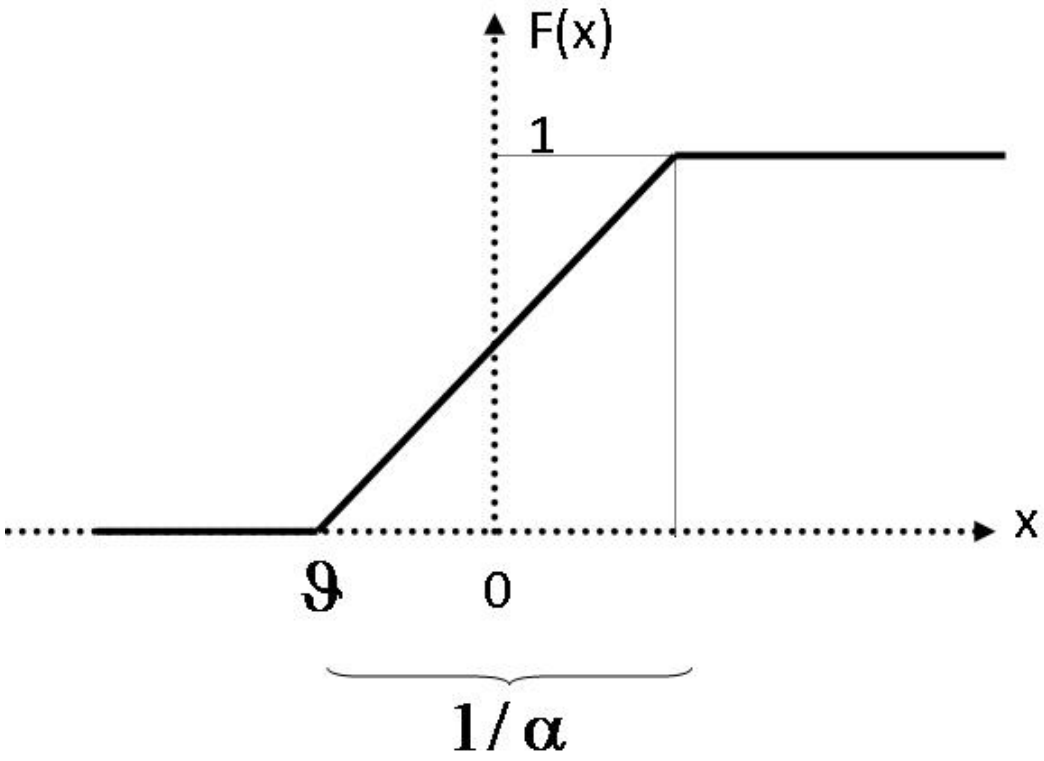


Figure 3.9. Threshold logic function

CHAPTER 4 - **Two-dimensional and three-dimensional characterizations of a hypothetical data-rich site by various methods**

4.1 Introduction

Environmental contaminants in geologic media such as soil and groundwater are of great concern in today's society. Millions of dollars are spent each year on efforts to clean up areas contaminated by pollutants from industrial and public waste such as solvents, fuels, and processing waste. Before cleanup efforts can begin, the area and extent of contamination must first be determined by using one of several profiling methodologies. In 2004, Chin, et al, stated that typically, these methodologies take known data and use specific mathematical algorithms to predict the areas and levels of contamination. In this chapter, the profiling performance of eight well-known profiling methodologies and ANNs are compared.

When trying to identify an area of contamination, the most accurate and precise means is to perform soil/groundwater sampling at designated regular intervals throughout the area in question (Najjar & Mryyan, 2005). This, however, is not a practical method due to cost and time constraints. Instead, a limited number of samples are collected throughout the area in question. Many factors determine where and how samples are collected. Generally, sample locations are determined using the professional judgment of a site investigation team. Once known data is collected, one of several profiling methodologies can be used to predict areas of contamination. Eight of the most often-used methodologies (available in Surfer® 8.0 Software: http://www.goldensoftware.com) for 2-D profiling are: Inverse Distance to a Power, Kriging, Minimum Curvature, Modified Shepard's, Natural Neighbor, Polynomial Regression, Radial Basis Function, and Local Polynomial.

According to Golden (2007), all profiling methodologies function in a similar manner, but some methodologies produce more accurate profiles than others. In order to utilize any profiling methodology, a known set of data must be present. This research used a hypothetical data-rich contaminated site scenario. Accordingly, data to compare pollution concentration profiles for 2-D and 3-D cases was generated via different profiling methodologies.

4.2 Surfer® Profiling Methods

Golden (2007) pointed out that Surfer® is a contouring and 2-D surface mapping program that quickly and easily transforms random surveying data into continuous curved face contours using interpolation. In particular, the new version, Surfer® 8.0, provides more than twelve interpolation methods, each having specific functions and related parameters. Below is a brief description of the most popular methods.

4.2.1 Inverse Distance to a Power

The Inverse Distance to a Power gridding method is a weighted average interpolator and can be either exact or smoothing. With this method, a weighting power is assigned to data that defines how the factors will decline as distance from a grid node increases. The higher the weighting power, the less effect there is on the estimation point further away from the initial grid node point. Davis (1986) found that the equation used for Inverse Distance to a Power is:

$$Z_j = \frac{\sum_{i=1}^{n} \frac{Z_i}{h_{ij^\beta}}}{\sum_{i=1}^{n} \frac{1}{h_{ij^\beta}}}$$

$$h_{ij^\beta} = \sqrt{d_{ij}^2} + \delta^2 \qquad 4.1$$

Where:

h_{ij} The effective separation distance between grid node "j" and the neighboring point "i"

Z_j The interpolated value for grid node "j"

Z_j The neighboring points

d_{ij} The distance between the grid node "j" and the neighboring point "i"

β The weighting power (the *Power* parameter)

δ The *Smoothing* parameter

Normally, Inverse Distance to a Power behaves as an exact interpolator. To calculate the grid node, data points are assigned a fractional weight and the sums of all weights are equal to 1.0. When a known point aligns with a grid node, the distance between that known point and the grid node is 0.0, and that known point is given a weight of 1.0, while all other points are given weights of 0.0. Thus, the grid node is assigned the value of the known point (Franke, 1982). One disadvantage is that the known points are not uniformly spaced among the interpolation points. Because of this, some clusters of points tend to carry an unnaturally large weight. To minimize this effect, no point is given an overpowering weight. No point is given a weighting factor equal to 1.0.

4.2.2 Kriging

Kriging is a geostatistical gridding method that has proven useful and popular in many fields. This method produces visually appealing maps from irregularly spaced data (Cressie, 1990). Kriging is a very flexible gridding method. Kriging defaults can be used to produce an accurate grid of data, or Kriging can be custom-fit to a data set by specifying the appropriate

variogram model. Within Surfer®, Kriging can be either an exact or a smoothing interpolator depending on the user-specified parameters. It incorporates anisotropy and underlying trends in an efficient and natural manner (Journel, 1989).

4.2.3 Minimum Curvature

Minimum Curvature is widely used in the earth sciences. The interpolated surface generated by Minimum Curvature is analogous to a thin, linear elastic plate passing through each of the data values with a minimum amount of bending. Minimum Curvature generates the smoothest possible surface while attempting to honor the data as closely as possible. Minimum Curvature is not an exact interpolator, however. This means that data are not always honored exactly.

Minimum Curvature produces a grid by repeatedly applying an equation over the grid in an attempt to smooth it. Each pass over the grid is counted as one iteration (Franke, 1982). The grid node values are recalculated until successive changes in the values are less than the Maximum Residuals value, or the maximum number of iterations is reached.

4.2.4 Modified Shepard's

According to Shepard (1968), the Modified Shepard's Method uses an inverse distance weighted least squares method. As such, Modified Shepard's is similar to the Inverse Distance to a Power interpolator, but the use of local least squares eliminates or reduces the appearance of the generated contours. Modified Shepard's can be either an exact or a smoothing interpolator.

Franke and Nielson (1980) state that the Modified Shepard's starts by computing a local least square fit of a quadratic surface around each observation.

4.2.5 Natural Neighbor

Sibson (1980) and (1981) reported that the Natural Neighbor gridding method is quite popular in some fields. Natural Neighbor is as simple to use as Nearest Neighbor and provides more precise results; however, it is only available for 2-D interpolations. Natural Neighbor requires that a grid be defined.

Natural Neighbor interpolation is a weighted moving average technique that uses geometric relationships in order to choose and weight nearby points. The equation for the Natural Neighbor interpolation is (Isaaks & Sirvastava, 1989):

$$\boldsymbol{G(x, y)} = \sum_{i=1}^{n} Wi \int (X_{i,} Y_{i}) \qquad 4.2$$

Where:

- G(x, y) is the natural neighbor estimation at (x, y);
- n is the number of nearest neighbors used for interpolation;
- f(x_i , y_i) is the observed value at (x_i ,y_i); and
- W_i is the weight associated with f(x_i, y_i)

According to the Surfer® Users Guide, sometimes with nearly complete grids of data, there are areas of missing data that a user might want to exclude from the grid file. In this case, the search ellipse can be set to a value so the areas of no data are assigned the blanking value in the grid file. By setting the search ellipse radii to values less than the distance between data values in the file, the blanking value is assigned at all grid nodes where data values do not exist.

4.2.6 Polynomial Regression

Polynomial Regression is used to define large-scale trends and patterns in data. Polynomial Regression is not really an interpolator because it does not attempt to predict unknown *Z* values. Several options can be used to define the type of trend surface (Draper & Smith, 1981).

4.2.7 Radial Basis Function

The Surfer® Users Guide states that the Radial Basis Function interpolation is a diverse group of data interpolation methods. In terms of the ability to fit the data and to produce a smooth surface, the multiquadric method is considered by many to be the best. Powell (1990) noted that all of the Radial Basis Function methods are exact interpolators, so they attempt to honor the data. A smoothing factor can be introduced to all the methods in an attempt to produce a smoother surface.

4.2.8 Local Polynomial

Lee and Schachter (1980) stated that the Local Polynomial gridding method assigns values to grid nodes by using a weighted least squares fit with data within the grid node's search ellipse.

4.3 Two-Dimensional Case

4.3.1 Mathematical Equation

In order to determine the distribution of contaminants at the hypothetical site, the following mathematical equation (Equation 4.3) was developed in order to produce the pollutant concentration value at any given (x, y) location:

$$V = x^{0.5} + y^{0.6} + (x^{0.3} * y^{0.2}) + 2\sqrt[4]{(x*y)} + 3\ln\left(\frac{x^{1.1} + y^{1.5} + 20{,}000}{1000}\right) \qquad 4.3$$

In Equation 4.3, V represents the contaminant concentration value. Note that x and y coordinates used in this equation refer to the x and y distances (in feet) for the associated observation point measured from a reference point (i.e., $x = 0$ ft and $y = 0$ ft).

4.3.2 Databank

Two databases containing x, y, and V values were generated for two 2-D cases at various locations across the site. The site size is 1000 ft in the x direction by 1000 ft in the y direction. To achieve this objective, the hypothetical site was divided into two grid systems as follows:

1. 25-ft interval case: In this scenario, 25-ft intervals (i.e., $\Delta x = \Delta y = 25$ ft) in both x (east) and y (north) directions were used to generate a total of 1,681 sampling points. The x and y coordinates and V values of 17 selected points (about 1% of the total sampling points), were provided for the eight profiling methodologies available in Surfer® software. Each methodology was then used to predict the corresponding contamination value (V) for the 1,681 designated x and y coordinates representing the site. The resulting data banks were processed to construct 8 contamination distribution contour maps as well as to calculate the corresponding Root Mean Squared Error (RMSE) value (using Equation 4.4) associated with each profiling methodology.
2. 10-ft interval case: Utilizing a 10-ft interval (i.e., $\Delta x = \Delta y = 10$ ft) for both x (east) and y (north) directions, it was possible to generate a total of 10,201 sampling points for the 1000 ft x 1000 ft site. Similar to the 25-ft interval case, x and y coordinates and V values of 103 selected points (about 1% of the 10,201 total sampling points) depicted in Figure 4.1 were provided for the eight profiling methodologies available in the Surfer® software. Each

methodology was then used to predict the corresponding contamination value (V) for all 10,201 designated x and y coordinates representing the site. The resulting eight data banks were processed to construct eight contamination distribution contour maps and to calculate (using Equation 4.4) the corresponding RMSE value associated with each methodology.

4.3.3 ANN Model Development

Unlike Surfer® 8.0 methodologies, the ANN-based profiling model requires the user to train or educate the network about the process it is to model. To train the network, a known set of input data along with the desired outcome is used [Dowla & Rogers (1995), Mryyan & Najjar (2005), Najjar & Itani (2000)]. The BPANN methodology [example, Najjar (1999)] using the supervised training approach can be used to train the desired ANN models to produce output values that are as close to the real values as possible via repeated modifications of the network's connection weights. This process typically continues until the error at the output layer is minimal. Once this training process is complete, the developed model can then be used for prediction tasks.

Neural Networks can reach a least-error structure by training, using examples related to the problem under consideration. A least-error structure is the one responsible for producing outputs very near or equal to the actual desired (target) values. Reasonable training input and output vectors should cover a wide range of the sampling domain. Deriving an appropriate and representative mapping between input and output vectors reflects the effectiveness of ANNs. For proper modeling, a network should pass through at least two stages, namely: training and testing (Najjar, Reddi & Basheer, 1996). In the training stage, selected data with their input and output values are introduced to a network (having a certain number of hidden nodes and layers) so that the network trains itself to produce output values that are as close to the target values as possible.

The training is achieved by modifying the values of the connection weights (Najjar & Basheer, 1996). The network stops learning when weight adjustment processes produce no improvement in the output values. The same network should be tested on data not used in training to verify its generalization capabilities. The procedure of training and testing should be repeated for various networks having different numbers of hidden layers and/or hidden nodes.

Najjar (1999) explained, when developing any ANN model, it is important to determine what input and output values will be used. For the hypothetical data-rich contaminated site case considered herein, the x and y coordinates are used as the only input values for the model. The pollution concentration value (V) is used as the output for their associated network model. In this case, x and y coordinates refer to the x and y distances (in feet) for the associated observation point, measured from a reference point (i.e., $x = 0$ ft, $y = 0$ ft).

For the 25-ft interval case (i.e., $\Delta x = \Delta y = 25$ ft), a network model was developed by using the same 17 points used by the eight Surfer® Software methodologies. For the ANN case, 12 data sets were used for training and the remaining five data sets were used for testing purposes. The best performing BPANN was determined by carrying out a number of adaptive training and online testing trials [as indicated in Najjar (1999)] in order to arrive at the least error on the testing data sets. According to Najjar (1999), overall BPANN is defined as the network yielding the least error [in terms of Averaged Squared Error (ASE) value] on the testing data sets from among all evaluated trial networks. In this case, overall BPANN was achieved at ASE value of 0.010856 and a structure topology noted as 2-3-1 (i.e., 2 inputs representing x and y coordinates, 3 hidden nodes, and 1 output denoting the associated value of the V variable). Once this network was established, it was then used to predict the V values at all 1,681 location points

for the site. The predicted values were used to construct a contamination distribution contour map and to calculate the corresponding RMSE value for this network model.

In the 10-ft interval case (i.e., $\Delta x = \Delta y = 10$ ft), a network model was developed by using about 1% of the total 10,201 data points (i.e., the same 103 data points used by the Surfer® Software methodologies). In this ANN development case, 75 data sets were used for training, and the remaining 28 data sets were used for online testing purposes (Najjar, 1999). Figure 4.1 shows the locations of the selected training and testing data points. Various ANN work by Najjar and his co-workers [Ali & Najjar (1998); Huang, et al. (2006); Mandavilli, et al. (2005); Mryyan & Najjar (2006); and Najjar & Felker (2003)], points out that it is highly imperative that the training data sets contain all data sets that have extreme attributes in terms of locations and values. Accordingly, the developed ANN model will always operate in an interpolation mode instead of an extrapolation mode. ANN-based prediction models are excellent when used in interpolation tasks, but may be unreliable when used in extrapolation tasks. Therefore, it is very important to appropriately select the distribution of the training and testing data sets. Following a strategy similar to the one used for the 25-ft interval case, the overall BPANN was achieved at an ASE value (on the testing data sets) of 0.000228 and a 2-2-1 network topology structure. Once the 2-2-1 profiling network was established, it was then used to predict the V values at all 10,201 location points for the site. The predicted values were used to construct a contamination distribution contour map and to calculate the corresponding RMSE value for this 10-ft interval case.

By comparing ASE values for the 10-ft and 25-ft interval cases, it can be observed that the ASE value for the 10-ft case was reduced by about 47-fold (i.e., 0.010856/0.000228) for the corresponding six-fold increase in data richness (i.e., 103/17). This noted behavior is expected

and logical. As more data become available, the profiling network should be able to characterize the site more accurately. Therefore, the more data that are available, the more accurate the developed profiling network will be. Moreover, it can be observed that the 10-ft interval network only needed two hidden nodes to efficiently characterize the site, in comparison with the three hidden nodes needed for the 25-ft interval network.

In order to compare (rank) the prediction accuracy of the profiling methodologies used herein, the following Root Mean Squared Error (RMSE) accuracy measure was used:

$$\text{RMSE}=\sqrt{\frac{\sum_{i}^{n}(y'-y)^2}{n}} \qquad 4.4$$

Where

n = number of data sets used

y' = the output generated by the model for the V variable

y = the actual value of the V variable

Accordingly, the best performing profiling methodology is the one yielding the least RMSE value.

4.4 Three-Dimensional Case

One of the first steps in the remediation process of any site is to determine the characteristics of the contaminated site. This includes not only obtaining historical and geological information about the site, but, most importantly, determining the locations and concentrations of contaminants in the site (Mryyan & Najjar, 2005). This is done by collecting samples at selected points throughout the area of concern and analyzing the samples to determine

the concentration of the contaminants. Each sample will have its own unique set of data that includes the location of the sample (latitude, longitude, and depth) and concentration. With this information, a detailed map of the contaminated site can be created (Najjar & Basheer, 1996).

4.4.1 Mathematical Equation

Unlike 2-D profiling methodologies available in the Surfer® Software, the ANN-based approach appropriately allows for 3-D site profiling by utilizing *x*, *y*, and *z* coordinates. In an actual field situation, samples would be collected at various locations for lab analysis in order to obtain the associated pollutant concentration values. For the purpose of this study, the following equation (Equation 4.5) was used to represent the concentration of the pollutant across the 3-D site (1000 ft x 1000 ft x 50 ft):

$$\boldsymbol{V} = x^{0.5} + y^{0.6} + z^{1.5} + (x^{0.3} * y^{0.2} + z^{0.1}) + 2(x * y * z)^{0.15} + 3\ln\left(\frac{x^{1.1} + y^{1.5} + z^{2.5} + 50{,}000}{1000}\right) \qquad 4.5$$

In this case, *x* = east, *y* = north, and *z* = depth. Accordingly, at any given location (*x*, *y*, and *z*), Equation 4.5 will produce the associated pollutant concentration value (i.e., *V* value). Note that *x*, *y*, and *z* coordinates refer to the *x*, *y*, and *z* distances (in feet) for the associated observation point measured from a reference point (i.e., *x* = 0 ft, *y* = 0 ft, and *z* = 0 ft).

4.4.2 ANN Model Development

A large database containing *x*, *y*, *z*, and associated *V* values was generated using Equation 4.5. Accordingly, the (1000 ft x 1000 ft x 50 ft) hypothetical site was divided into a grid system. Grid lines were set at 25-ft intervals for both *x* (east) and *y* (north) directions (i.e., 2-D plane), and at depths *z* = 5, 15, 25, 35, and 45 ft. A total of 1,681 sampling points were generated in this case at each depth. This produced a total of 8,405 points. In this case, *x*, *y*, and *z* coordinates were used to represent the ANN model's input nodes, while the *V* variable is used to represent

the output node. Eighty-five (85) data points (representing about 1% of the total 8,405 available data points) were selected to train and test the desired ANN model in accordance with the procedure outlined in Najjar (1999). In this case, 60 data points were used for training, while the remaining 25 points were used for online testing in order to assess the generalization capability of the trial networks. A procedure similar to the one used in the 2-D case was utilized herein to arrive at the optimal 3-D ANN profiling model. The topology structure of the 3-3-1 BPANN contained 3 input, 3 hidden, and 1 output node. This 3-3-1 BPANN model yielded an ASE on the testing data sets with a value of about 0.000300. Note that this 0.000300 ASE value is slightly higher than the one obtained for the 2-D 10-ft interval case discussed earlier. The resulting 3-3-1 ANN model was then used to predict the corresponding contamination values (V) for the 8,405 designated x, y, and z coordinates representing the site. The resulting data bank was processed to construct various contamination distribution contour maps (at z = 5, 15, 25, 35, and 45 ft) for the hypothetical contaminated site. Moreover, the resulting data bank was used to compare the ANN predicted values with the actual values at all 8,405 location points. The resulting RMSE value calculated for the developed 3-D BPANN is about 6.38%.

4.4.3 Regression Model Development

Since none of the eight Surfer®-based methodologies can perform 3-D profiling using x, y, and z, the following regression-based equation (Equation 4.6) was developed using the same 85 data points utilized in developing the BPANN model:

$$V = -4.11 + 0.0499x + 0.0766y + 7.60z \qquad 4.6$$

Where V represents the desired contaminant concentration value for given x, y, and z coordinates within the site. The regression model (Equation 4.6) was then used to predict the

corresponding contamination values (V) for the 8,405 designated x, y, and z coordinates representing the site.

The resulting data bank was processed to construct various contamination distribution contour maps (at z = 5, 15, 25, 35, and 45 ft) of the hypothetical contaminated site. Similarly, the data bank was used to compare the regression-based-model-predicted values with the actual values at all 8,405 location points. The RMSE value obtained for this case is about 17.4% .

4.5 Results and Discussion

In order to compare the performance of all methodologies utilized herein, three comparison strategies were utilized, namely:

- Comparison using RMSE values
- Comparison using contour maps
- Forty-five-degree scatter graphs

4.5.1 Comparison Using RMSE Values

4.5.1.1 Two-Dimensional Case

RMSE values obtained for both 2-D profiling cases and the nine profiling methodologies (including the ANN method) are listed in Tables 4.1 and 4.2. By examining RMSE values listed in table 4.1, it can be observed that for the 25-ft interval case, all nine methodologies attained high RMSE values. The model achieving the least RMSE value is the ANN-based model. It attains about a 19.17% RMSE. The second-most accurate methodology is the Local Polynomial, with a 19.3% RMSE. When compared to ANN performance, this represents less than 1% difference in prediction accuracy rate. The profiling methodology that produced the least

accurate profile is the Inverse Distance to a Power method. It has an RMSE value of about 42.4%. This represents more than double the RMSE value attained by the 2-3-1 ANN model.

For the 10-ft interval case (Table 4.2), all nine methodologies attained lower RMSE values than those obtained for the 25-ft interval case. This noted behavior is logical and consistent with our intuition. As more data become available, models will become more accurate. Again, the model with the least RMSE value is the ANN-based model. Its error rate is about 3.7%. The second most accurate methodology is the Radial Basis Function, with an error rate of about 4.8%. When compared to ANN performance, this represents about a 30% difference in the prediction accuracy rate. The profiling methodology that produced the least accurate profile is, again, the Inverse Distance to a Power method, with an RMSE value of about 10.4%. This represents about 2.8 times the RMSE value attained via the ANN model. The only constant in the RMSE comparison listed in Tables 4.1 and 4.2 and Figures 4.1 and 4.3 is that ANN-based profiling methodology is ranked best and the Inverse Distance to a Power methodology is ranked worst. The other eight methods seem to vary in terms of their ranking. Therefore, in order to assure that the best profiling methodology for 2-D cases is being used, the ANN-based profiling methodology is recommended.

4.5.1.2 Three-Dimensional Case

When comparing the RMSE (Table 4.3 and Figure 4.4) value obtained using the 3-D ANN-based model (with an RMSE value of 6.4%) with that obtained with the regression-based model (with RMSE value of 17.4%), the ANN model significantly outperforms the regression model. The error rate of the regression-based model is about 270% of that reported for the ANN model. Note that the same 85 data points were used to develop both models. Moreover, knowing that all eight Surfer®-based methodologies are suited only for 2-D profiling, and cannot perform

3-D profiling, makes it clear that the ANN-based methodology is the one to use for efficient 3-D profiling tasks.

4.5.2 Comparison Using Contour Maps

Contour maps were generated using the Surfer® 8.0 software program. This program was used to produce contamination concentration contour maps for the hypothetical site using the previously mentioned data banks. Contour maps were generated for the 2-D [i.e., 10-ft and 25-ft interval scenarios] and 3-D cases discussed earlier. For contour maps comparison purposes, only contour maps for the 2-D 10-ft interval case and 3-D z=25ft case will be discussed.

4.5.2.1 Two-Dimensional Case

For a visual comparison, a baseline contour map of the pollutant concentration distribution of the site based on the actual 10,201 data points was generated, as depicted in Figure 4.5. This map is used herein as a baseline contour map to compare the profiling accuracy of the nine methods listed in Table 4.2. When comparing the contour maps of the nine profiling methods (depicted in Figures 4.6 through 4.14), the ANN-based contour map (Figure 4.6) is clearly the one that most closely resembles the baseline contour map shown in Figure 4.5. Note that, as indicated in Table 4.2, the ANN-based method attained the lowest RMSE value (3.7%) among all nine profiling methods. The remaining eight methods present lesser degrees of similarity to the baseline map. Contour maps produced by the Natural Neighbor and Inverse Distance to a Power methods (Figures 4.12 and 4.14, respectively) are the worst, compared with the baseline map shown in Figure 4.5.

Four of the contour maps, produced by the Radial Basis Function, Kriging, Modified Shepard's, and Minimum Curvature methods (Figures 4.7, 4.8, 4.9, and 4.11, respectively), show in the northeast region areas of lower contamination levels where actually higher contamination levels are present. The contour map produced by the Polynomial Regression method (Figure 4.13) is inadequate, because it is unable to capture the nonlinear spatial distribution of the pollutant within the site. The Local Polynomial method seems to produce the best contour map, as depicted in Figure 4.18, among all eight Surfer®-based profiling methods, even though its RMSE value is not the minimum in this case. On the other hand, the performance of the ANN-based method is very consistent. This method produces the best contour map as well as attaining the least RMSE value among all nine methods listed in Table 4.2.

Therefore, the ANN-based method should be considered as the method of choice for any 2-D site profiling. One common observation among all models considered herein is that no model was able to accurately characterize the actual (logarithmic) behavior of the variable *V* at the south and west edges of the site. In order to account for this logarithmic behavior, more data points taken from the south and west edges must be included in the models' profile development process.

4.5.2.2 Three-Dimensional Case

The baseline contour map for the distribution of the *V* variable at $z = 25$ ft is shown in Figure 4.15. This map was generated based on 1,681 actual data points derived directly from Equation 5. The corresponding ANN-based and regression-based contour maps at $z = 25$ ft are depicted in Figures 4.16 and 4.17, respectively. The RMSE values obtained in this case (6.4% for the ANN-based model and 17.4% for the regression model (Table 4.3) are an accurate indication of the degree of agreement between the profiles presented in Figures 4.16 and 4.17, with baseline

contour maps shown in Figure 4.15. The ANN-based profile (even though it was developed utilizing no more than 1% of the available data at the z = 25 ft level) presents a reasonable agreement with the actual map. The profile generated from the regression model has a very low degree of similarity with the actual profile shown in Figure 4.15. Similar to the 2-D case, no model was able to accurately characterize the actual (logarithmic) behavior of the variable V at the south and west edges of the site. To address this profiling deficiency, far more data points (taken from the south and west edges) are needed to capture this logarithmic behavior.

4.5.3 Comparison Using Forty-Five-Degree Scatter Graphs

Forty-five-degree linear graphs were generated using the Microsoft Excel program. This program was used to produce contamination concentration linear graphs for the hypothetical site using the previously mentioned (x, y, V, and z when applicable) data banks. Graphs were generated for each of the following:

- All 2-D methodologies at:
    - 25-ft interval case (Figures 4.18 through 4.26)
    - 10-ft interval case (Figures 4.27 through 4.34)
- ANN 3-D methodology at z = 5, 15, 25, 35, and 45 ft (Figure 4.35)
- Regression Analysis at z = 5, 15, 25, 35, and 45 ft (Figure 4.36)

Each graph contains the predicted value (V) via each methodology along the y-axis and the actual data value (V) along the x-axis. These graphs provide a visual analysis of how closely the predicted values match the actual data values.

For the 25-ft case (Figures 4.18 through 4.26), it can be noted that ANN was the best fit model with R^2=0.945 (Figure 4.18). The second best fit model is the Modified Shepard's with

R^2=0.944 (Figure 4.20) and worst performing method is the Inverse Distance to a Power with R^2=0.762 (Figure 4.26).

For the 10-ft case (Figures 4.27 through 4.34), ANN attains the best fit model with R^2=0.986 (Figure 4.27), and the second best fit model is obtained using the Radial Basis Method with R^2=0.977 (Figure 4.28). The worst performing method is the Inverse Distance Method with R^2=0.889 (Figure 4.34).

For the 3D case (Figures 4.35 through 4.36), the best fit method is again the ANN with R^2=0.997 (Figure 4.35). The Regression Based Method, the alternate method in this case, attains R^2=0.979 (Figure 4.36).

Overall, considering all scatter plots presented for 2D and 3D cases, it is clear that the ANN-based models attain the best match between predicted and actual values.

4.6 Conclusion

The use of ANN-based methodology for contaminant profiling, demonstrated in this study, provided the most reliable predictions about the location and extent of contamination for the hypothetical site. The ANN-based models proved to yield the lowest RMSE values in the 2-D and 3-D comparison cases. The ANN-based profiling models also produced the best contaminant distribution contour maps and 45-degree scatter graphs for the 2-D and 3-D profiling cases. Along with the fact that ANN is the only profiling methodology that allows for efficient 3-D profiling, this study clearly demonstrates that ANN-based profiling methodology, when properly used, has the potential to provide the most accurate predictions and site profiling contour maps for a contaminated site.

Compared to the methods discussed herein, ANN-based methodology is characterized by its flexibility and generality. Its flexibility is demonstrated by its potential to accurately predict

values of a certain contaminant parameter at a specific location when only supplied with *x*, *y*, and *z* (for 3-D cases) coordinates. Its generality lies in its power to capture the mode of change in the spatial distribution of a pollutant based on all available data. Accordingly, all available data at various spatial locations can effectively be utilized by the ANN-profiling model in order to efficiently capture the spatial distribution behavior for the parameter of interest.

Table 4.1 Profiling Methods and Their Corresponding RMSE Value for the 25-ft. Interval Case

	Method	RMSE
1	ANN	19.16940%
2	Local Polynomial Contours	19.29668%
3	Modified Shepard's Method	19.77924%
4	Minimum Curvature Contours	20.39590%
5	Kriging Contours	23.66336%
6	Polynomial Regression	26.41865%
7	Nearest Neighbor	38.95251%
8	Radial Basis Function	38.95251%
9	Inverse Distance to a Power	42.38522%

Table 4.2 Profiling Methods for the 10 ft interval case.

	Method	%RMSE
1	ANN	3.716%
2	Radial Basis Function	4.795%
3	Kriging Contours	4.994%
4	Modified Shepard's Method	4.999%
5	Local Polynomial Contours	5.153%
6	Minimum Curvature Contours	5.231%
7	Nearest Neighbor	8.022%
8	Polynomial Regression	8.710%
9	Inverse Distance to a Power	10.427%

Table 4.3 RMSE Values for Profiling Methods used to Predict (V) Values for 3D Case.

Method	RMSE
ANN	**6.389%**
Regression	17.420%

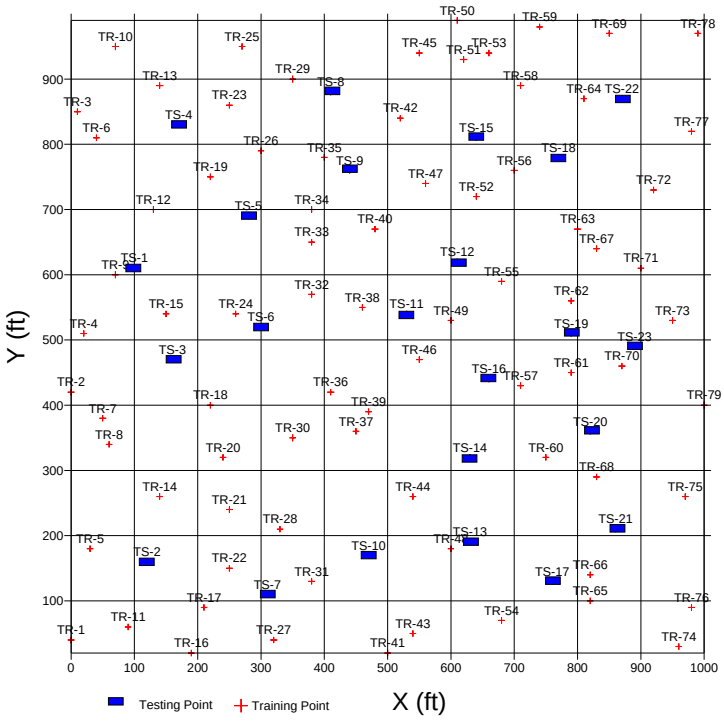


Figure 4.1. Location of testing and training points for the 10-ft interval case

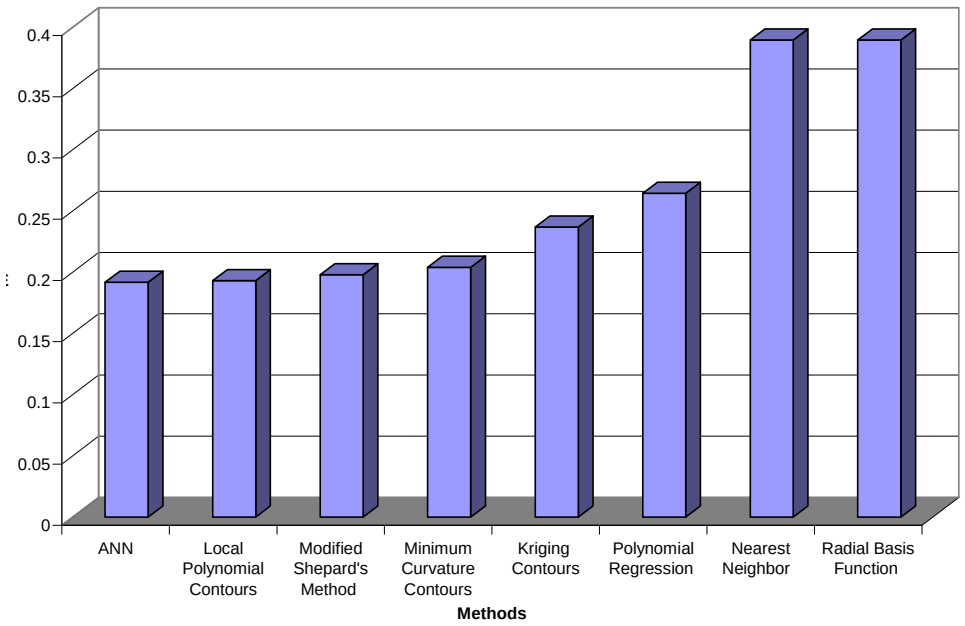


Figure 4.2. All of the profiling methods and their RMSE values (2-D and 25-ft interval case)

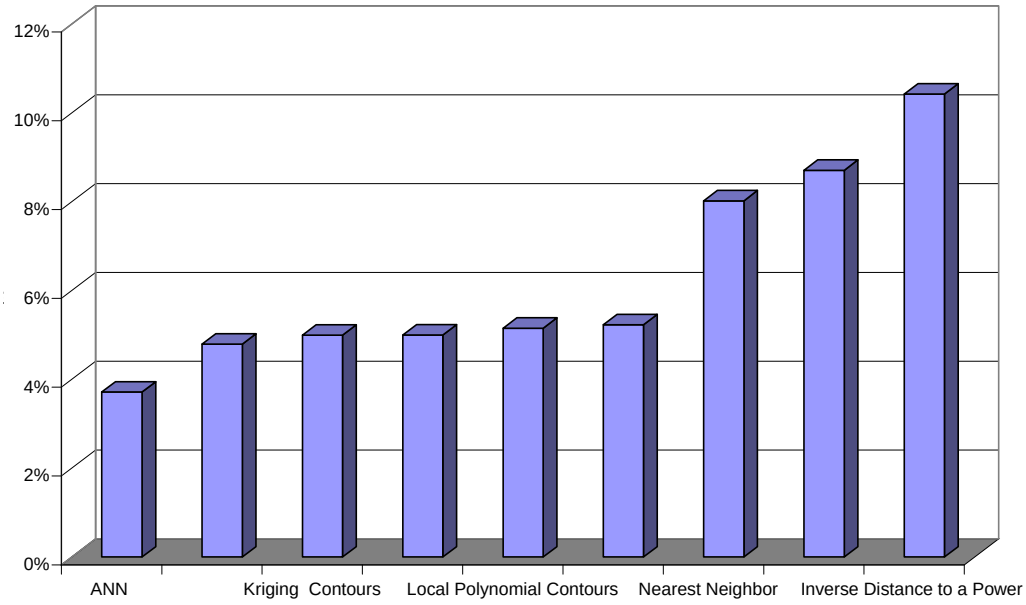


Figure 4.3. All of the profiling methods and their RMSE values (2-D and 10-ft interval case)

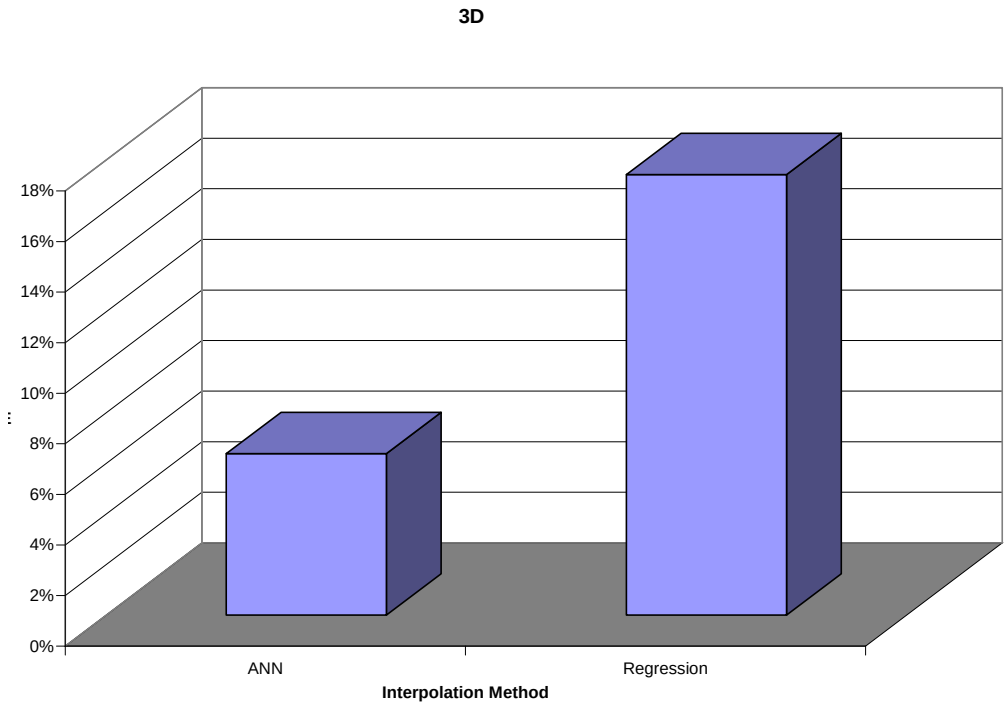


Figure 4.4. Bar chart showing RMSE value for profiling methods used to predict (V) values for 3-D case

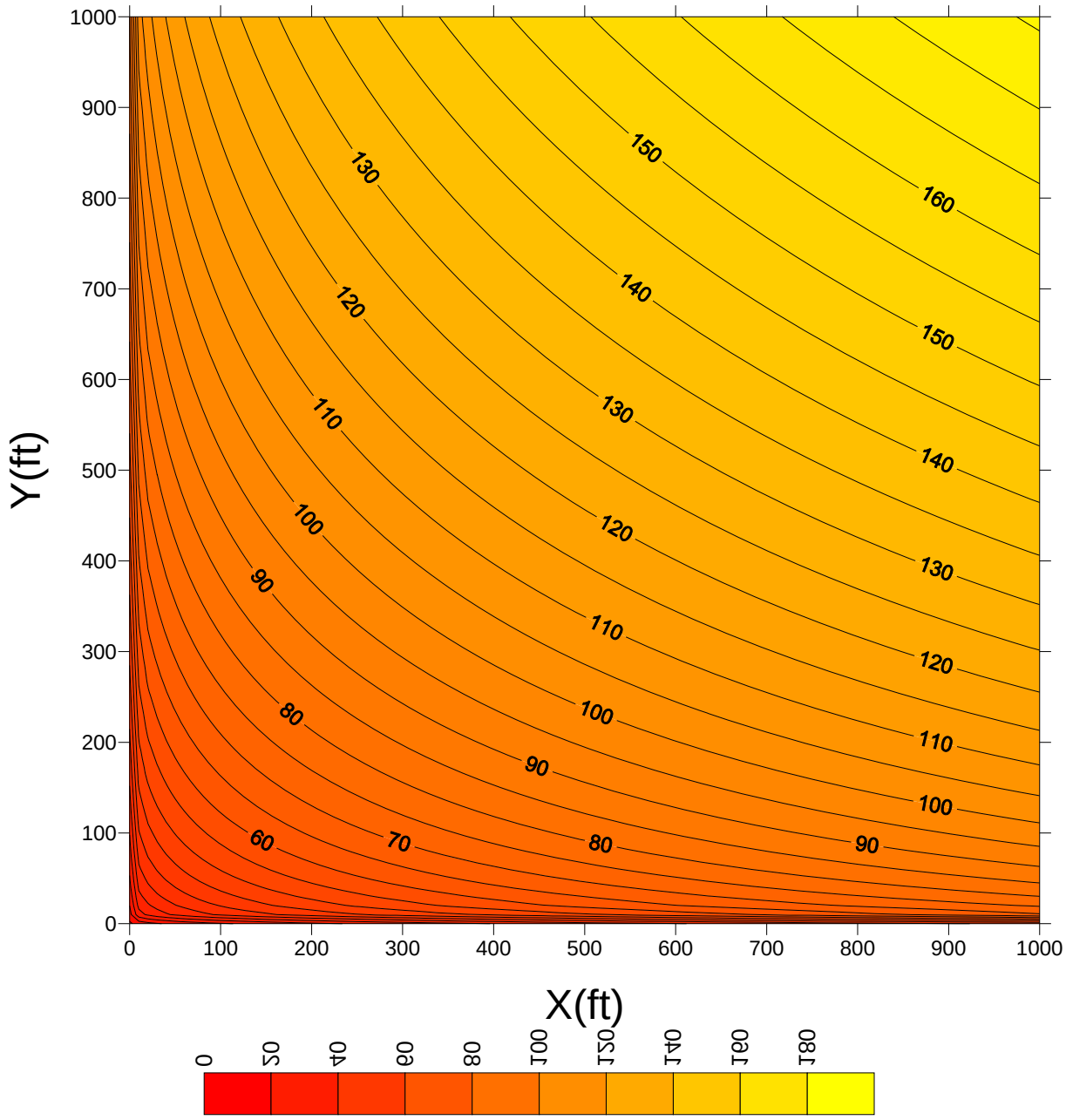


Figure 4.5. Baseline contour map of the pollutant V (for the 10-ft interval case), based on 10,201 actual data points

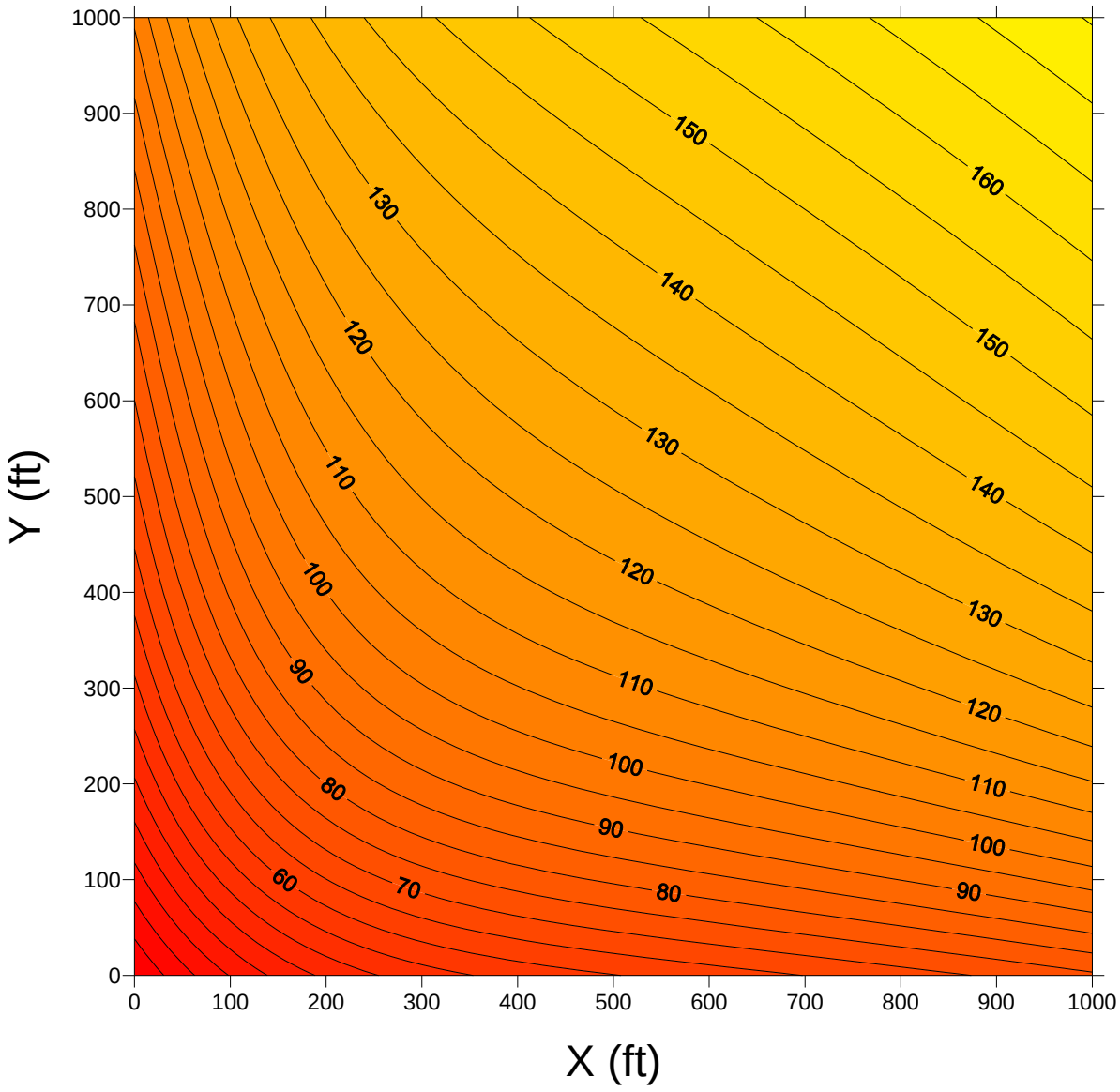


Figure 4.6. Contour map based on ANN model 2-2-1 for the 10-ft interval case

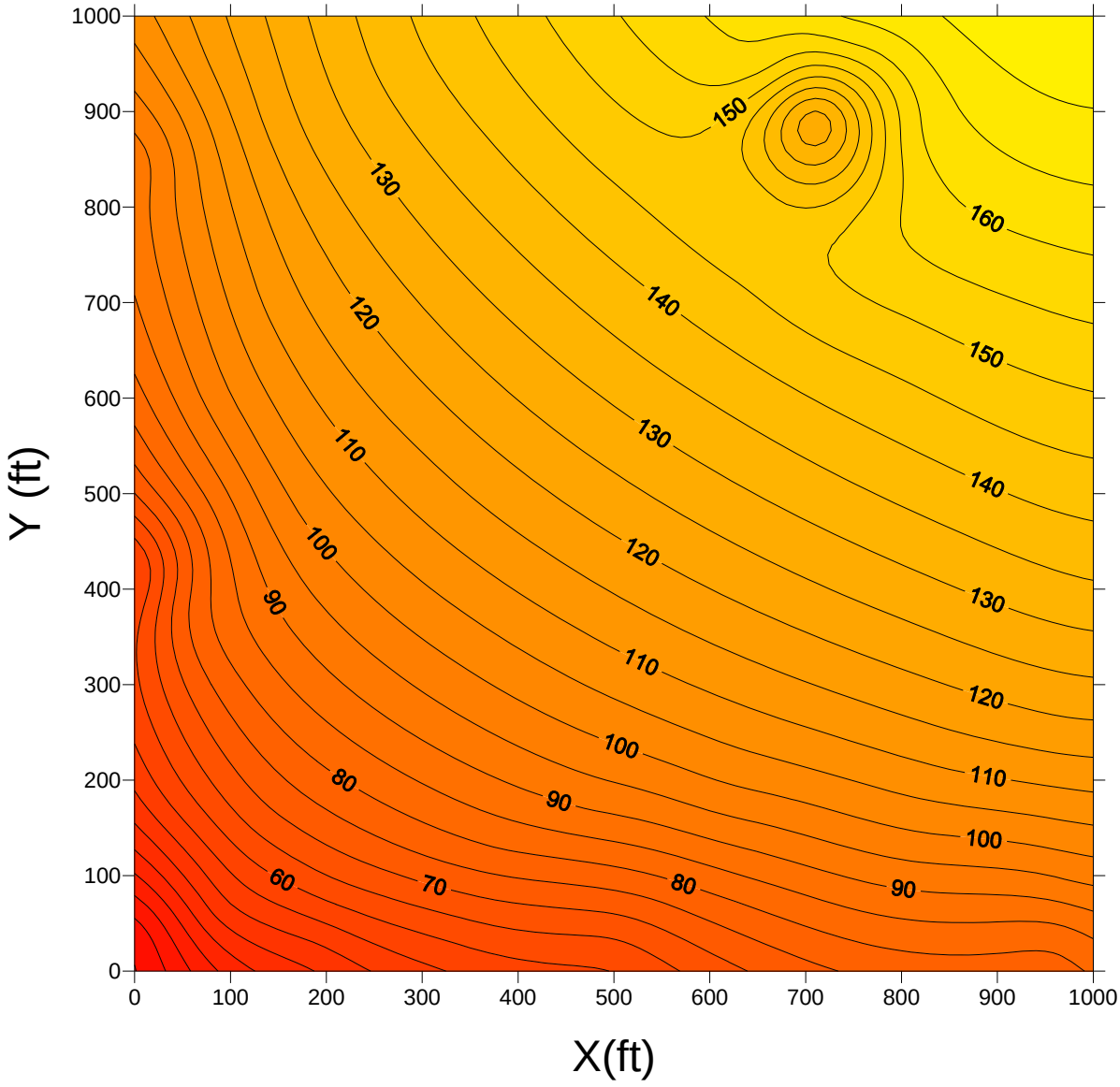


Figure 4.7. Contour map based on Radial Basis Function method for the 10-ft interval case

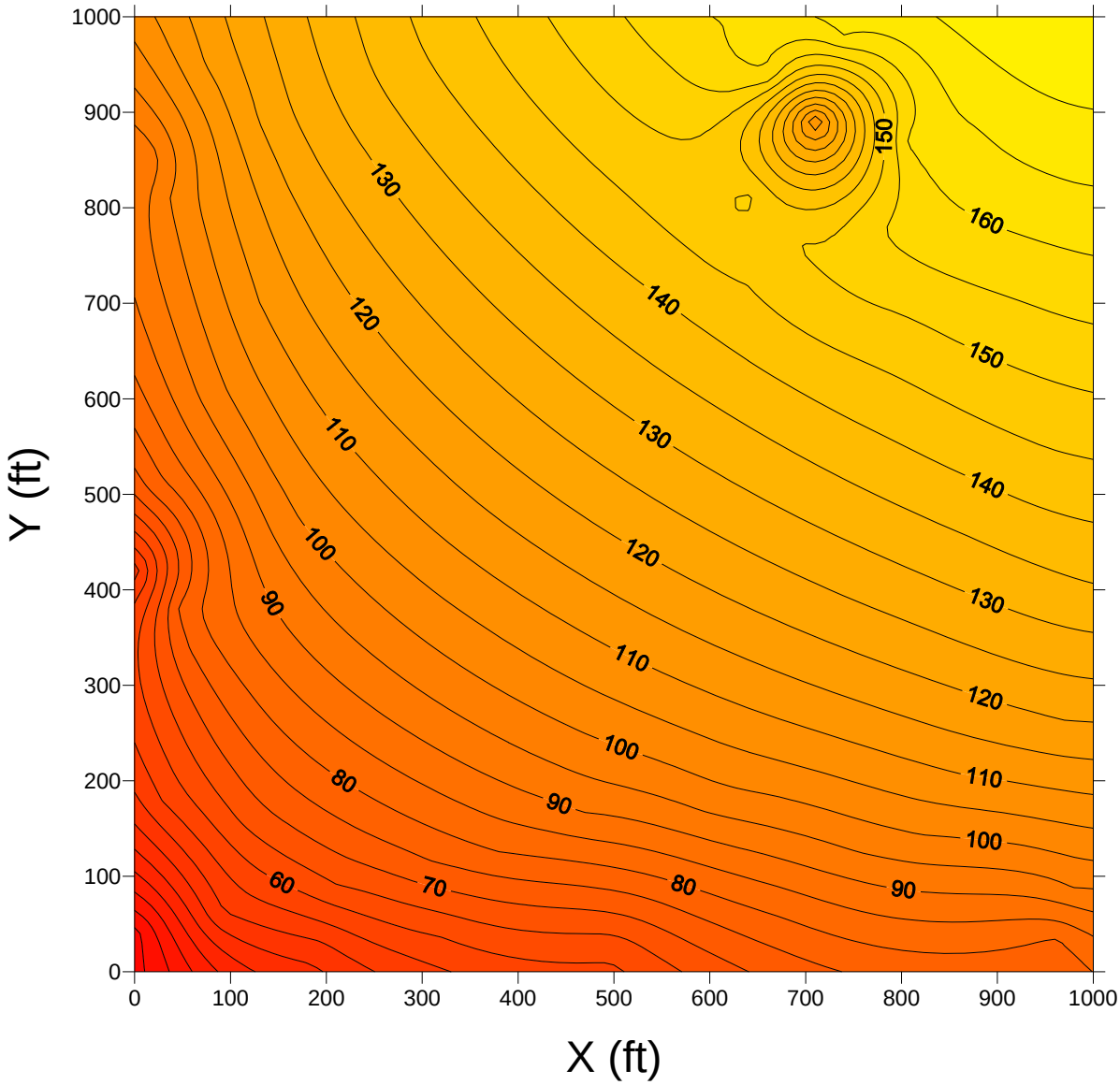


Figure 4.8. Contour map based on Kriging method for the 10-ft interval case

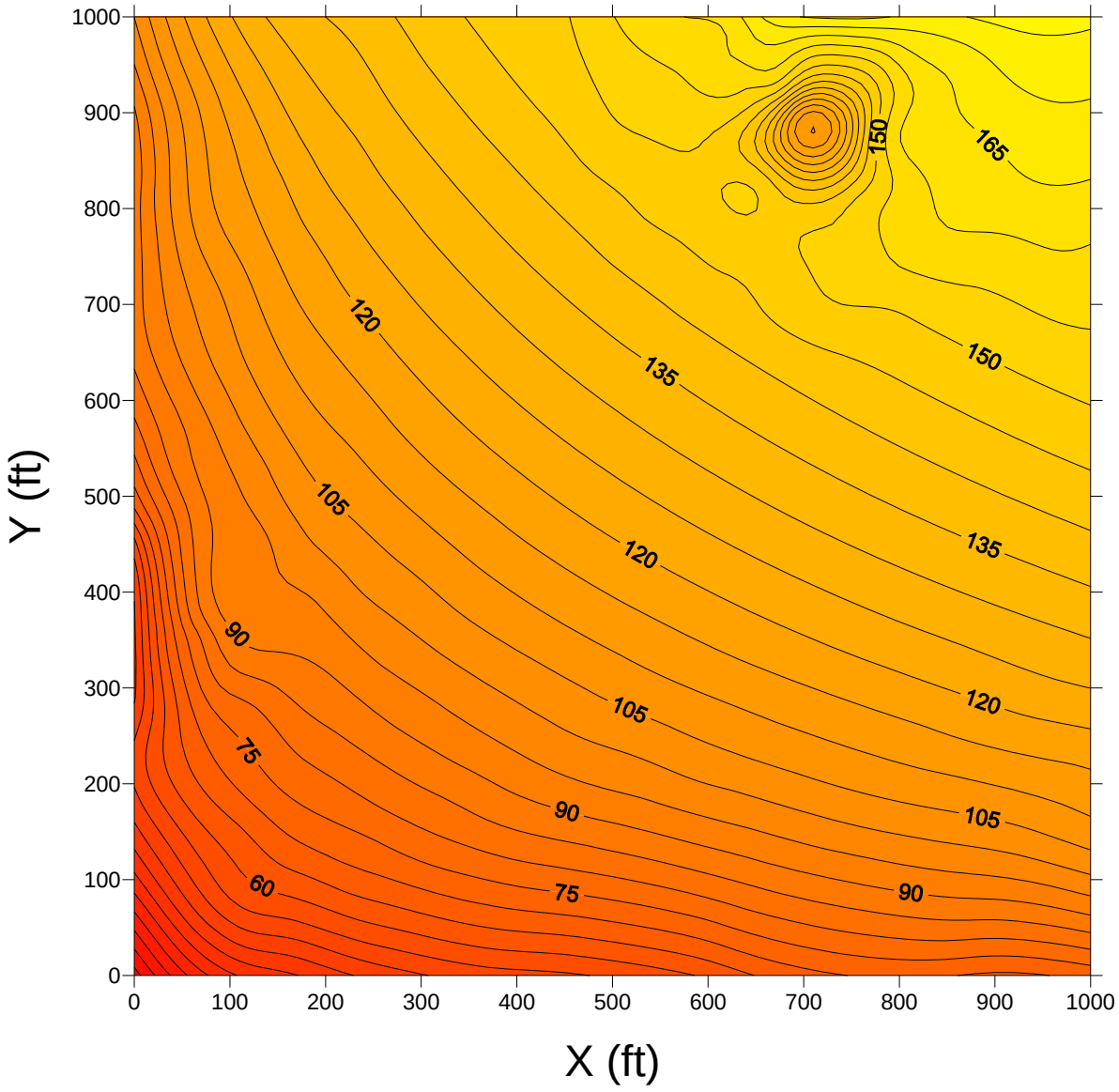


Figure 4.9. Contour map based on Modified Shepard's method for the 10-ft interval case

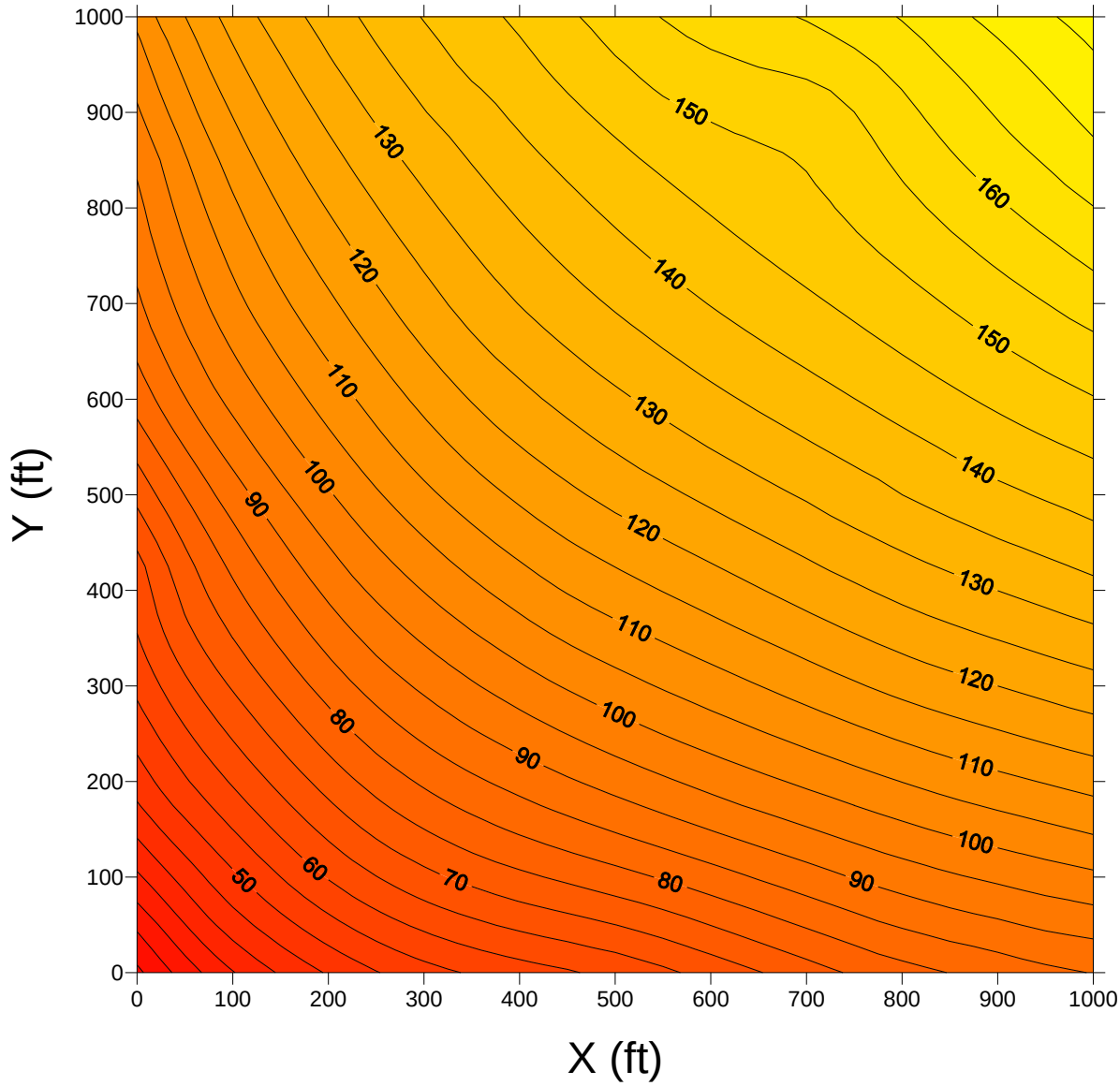


Figure 4.10. Contour map based on Local Polynomial method for the 10-ft interval case

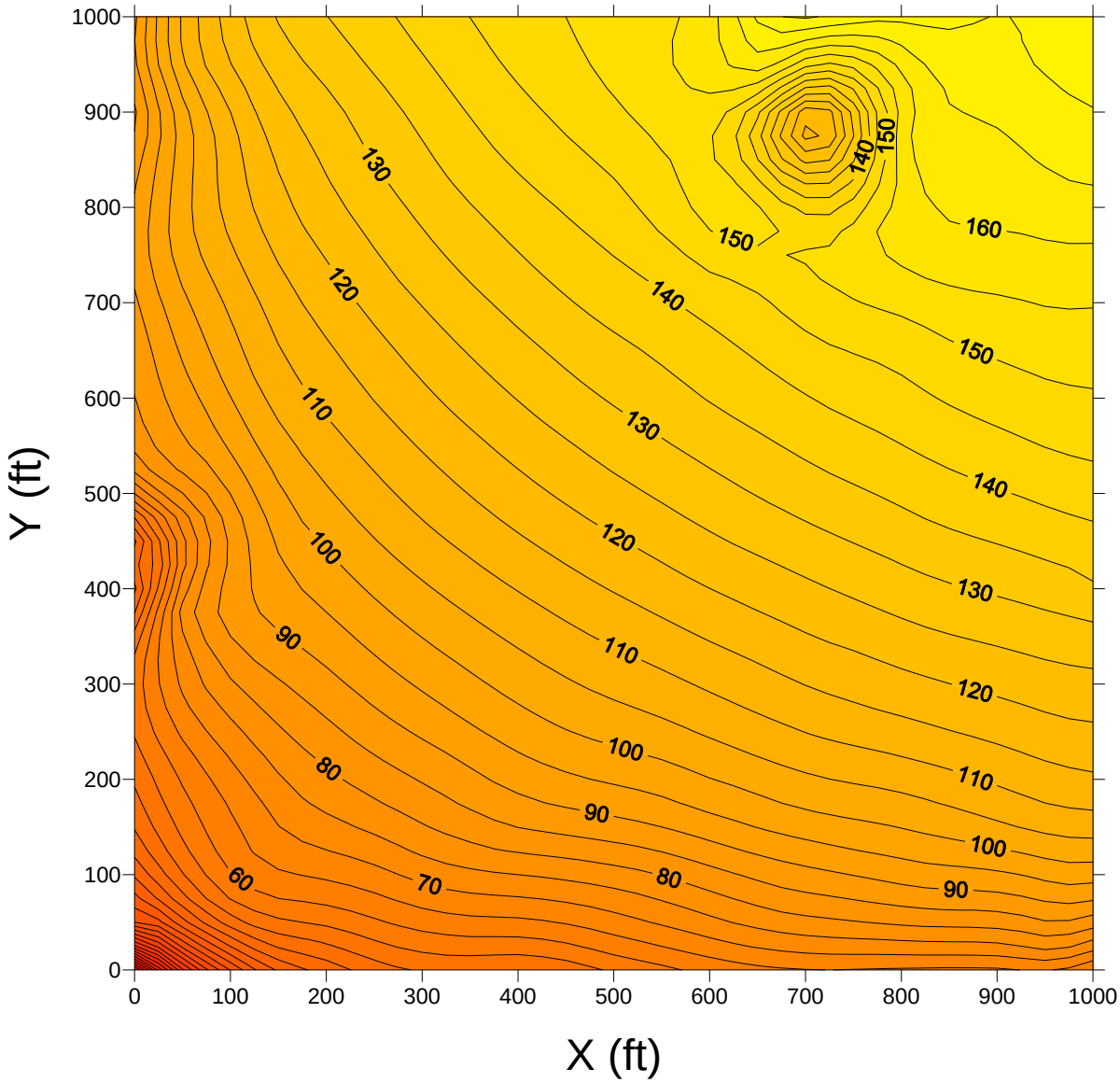


Figure 4.11 Contour map based on Minimum Curvature method for the 10-ft interval case

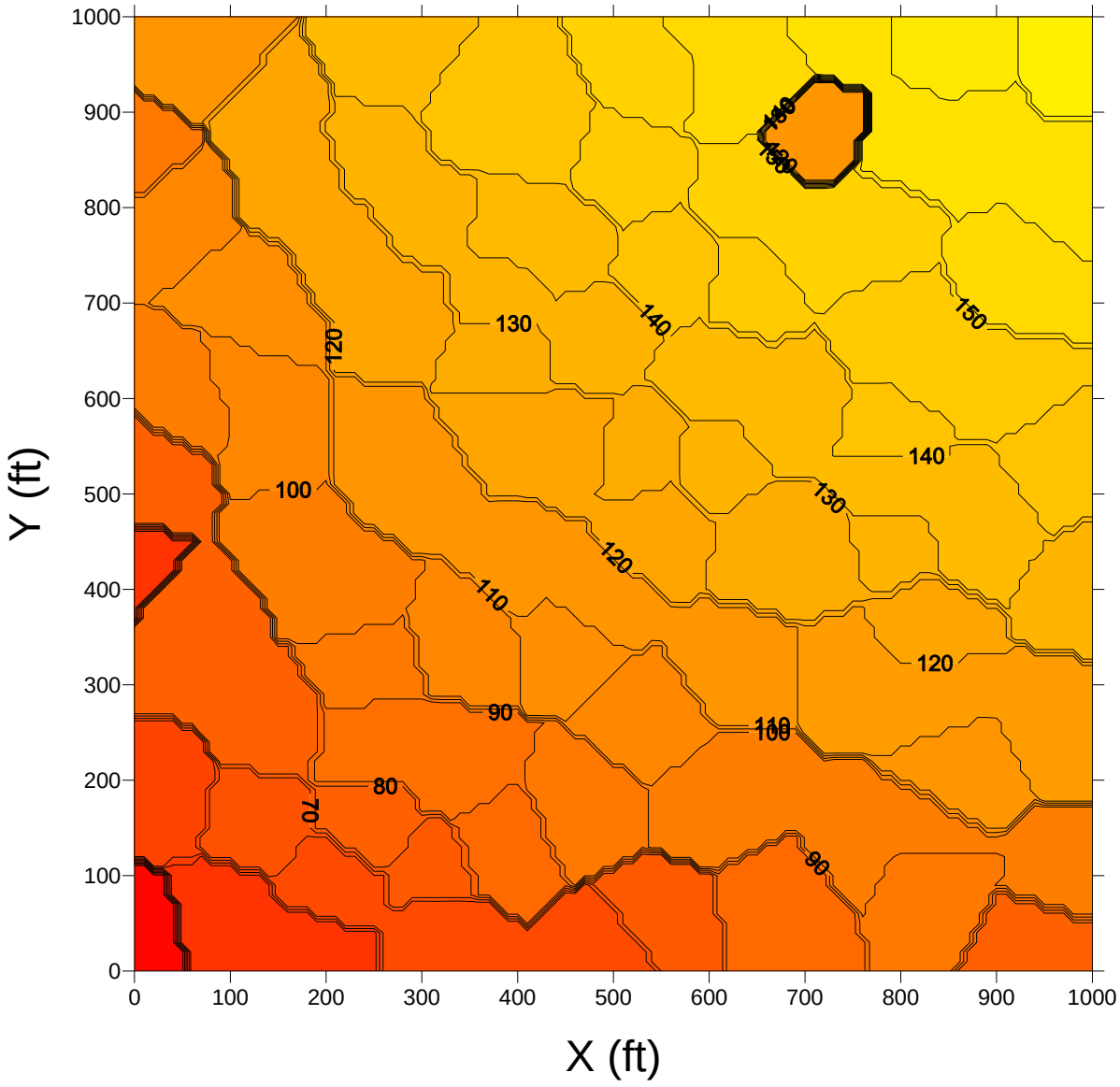


Figure 4.12. Contour map based on Nearest Neighbor method for the 10-ft interval case

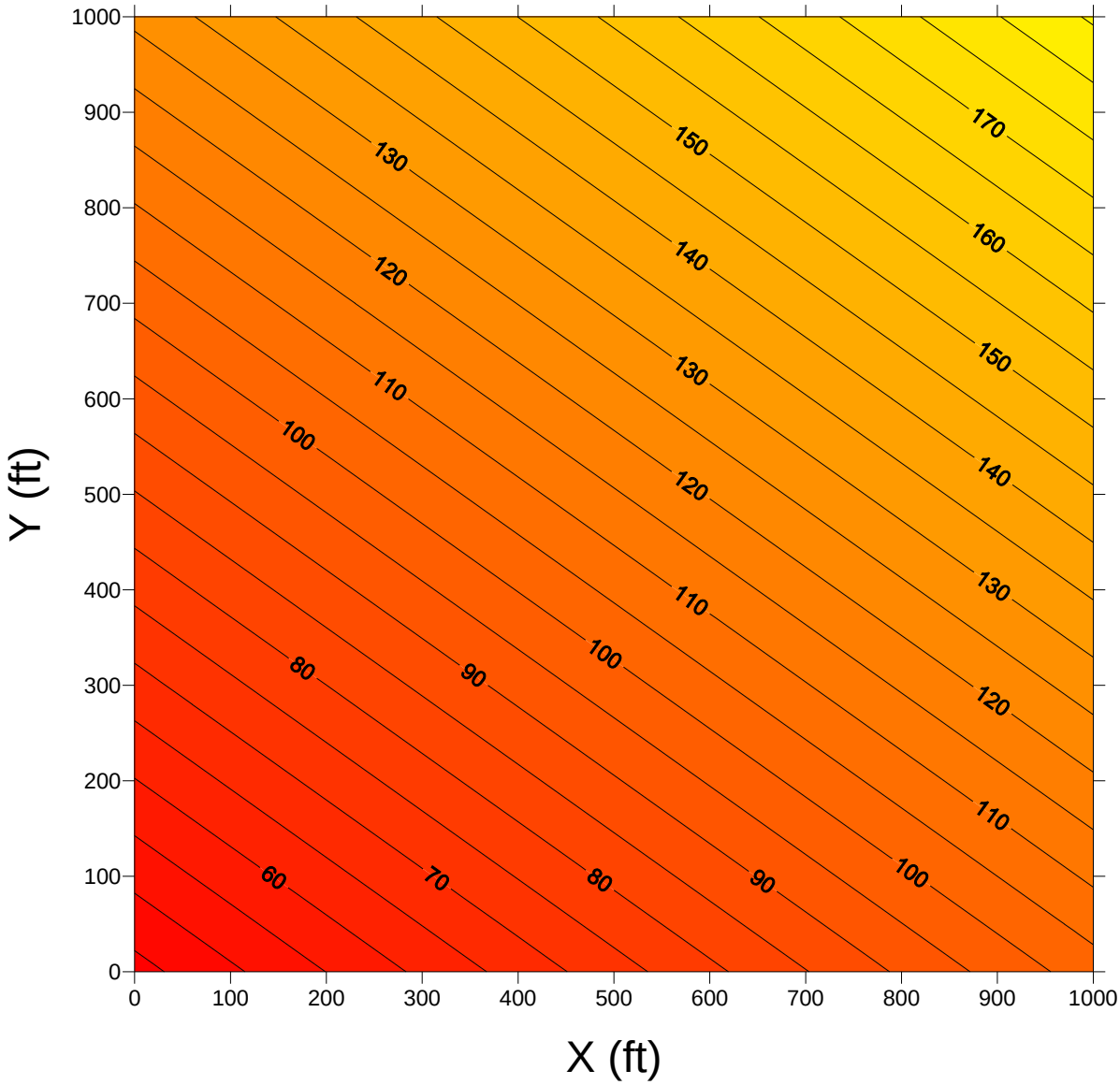


Figure 4.13. Contour map based on Polynomial Regression method for the 10-ft interval case

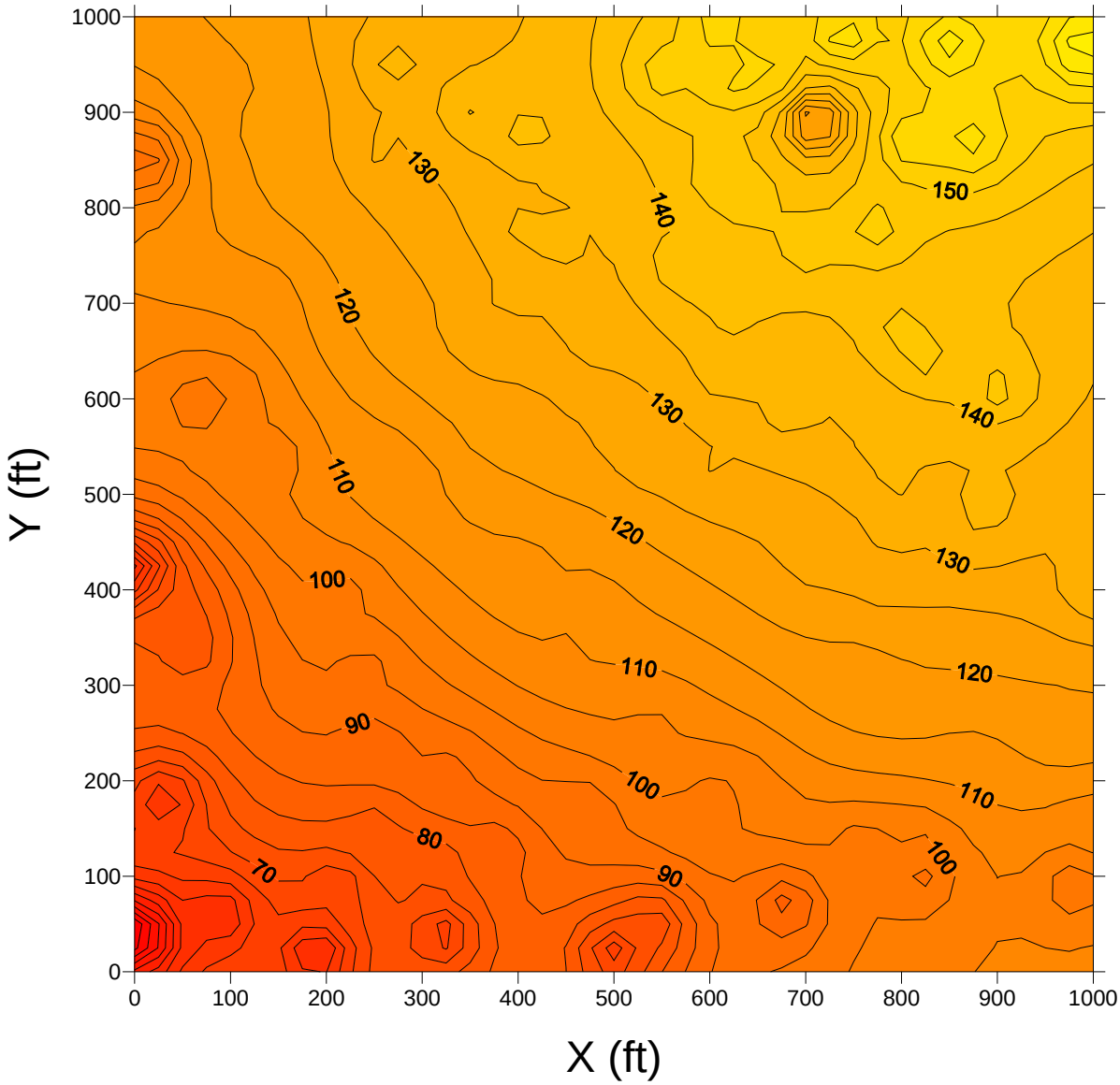


Figure 4.14. Contour map based on Inverse Distance Method for the 10-ft interval case

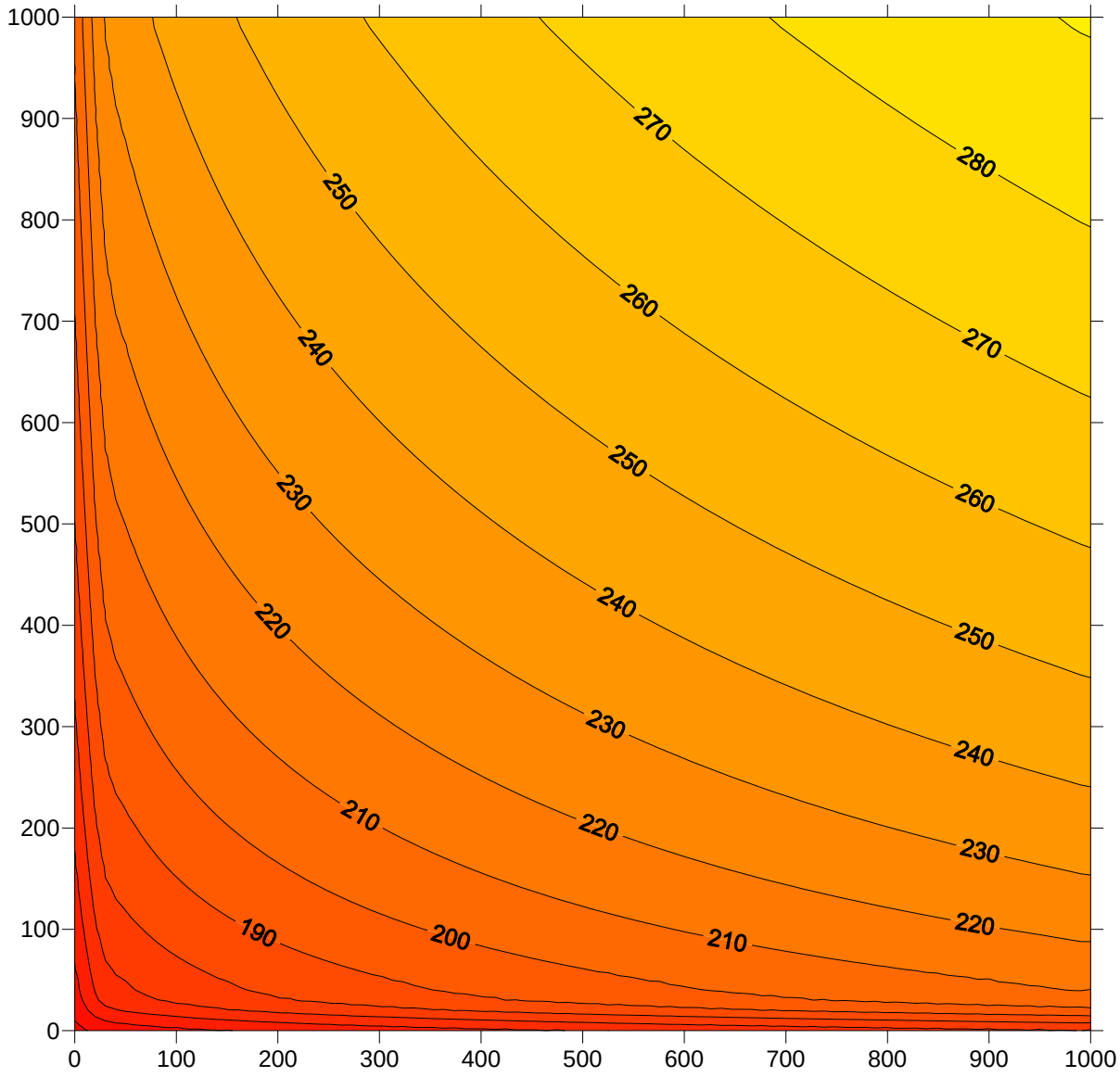


Figure 4.15. Baseline contour map of the pollutant V (at z = 25 ft) based on 1,681 actual data points (3-D case)

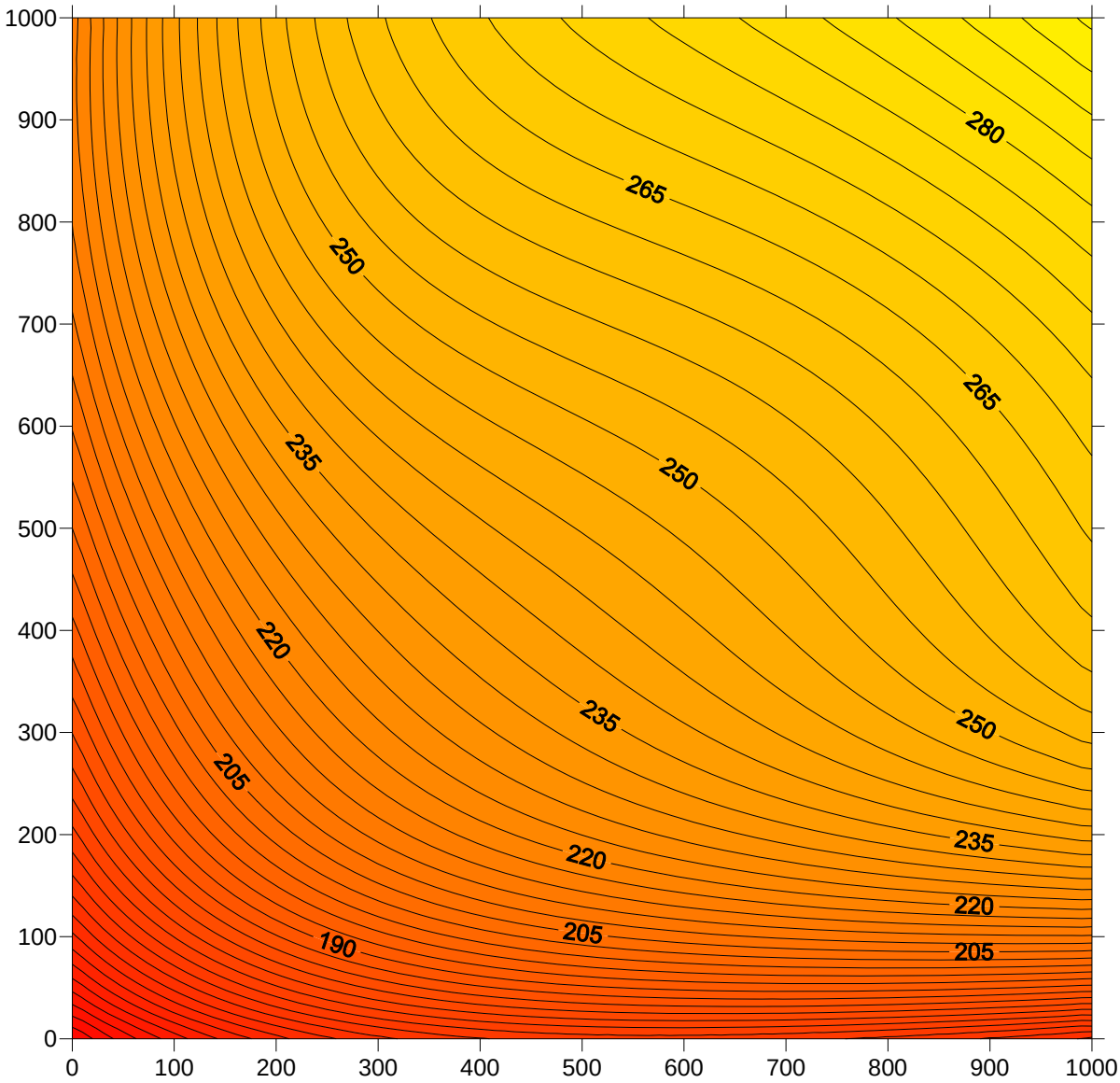


Figure 4.16. Contour map based on ANN model 3-3-1 at z = 25 ft (3-D case)

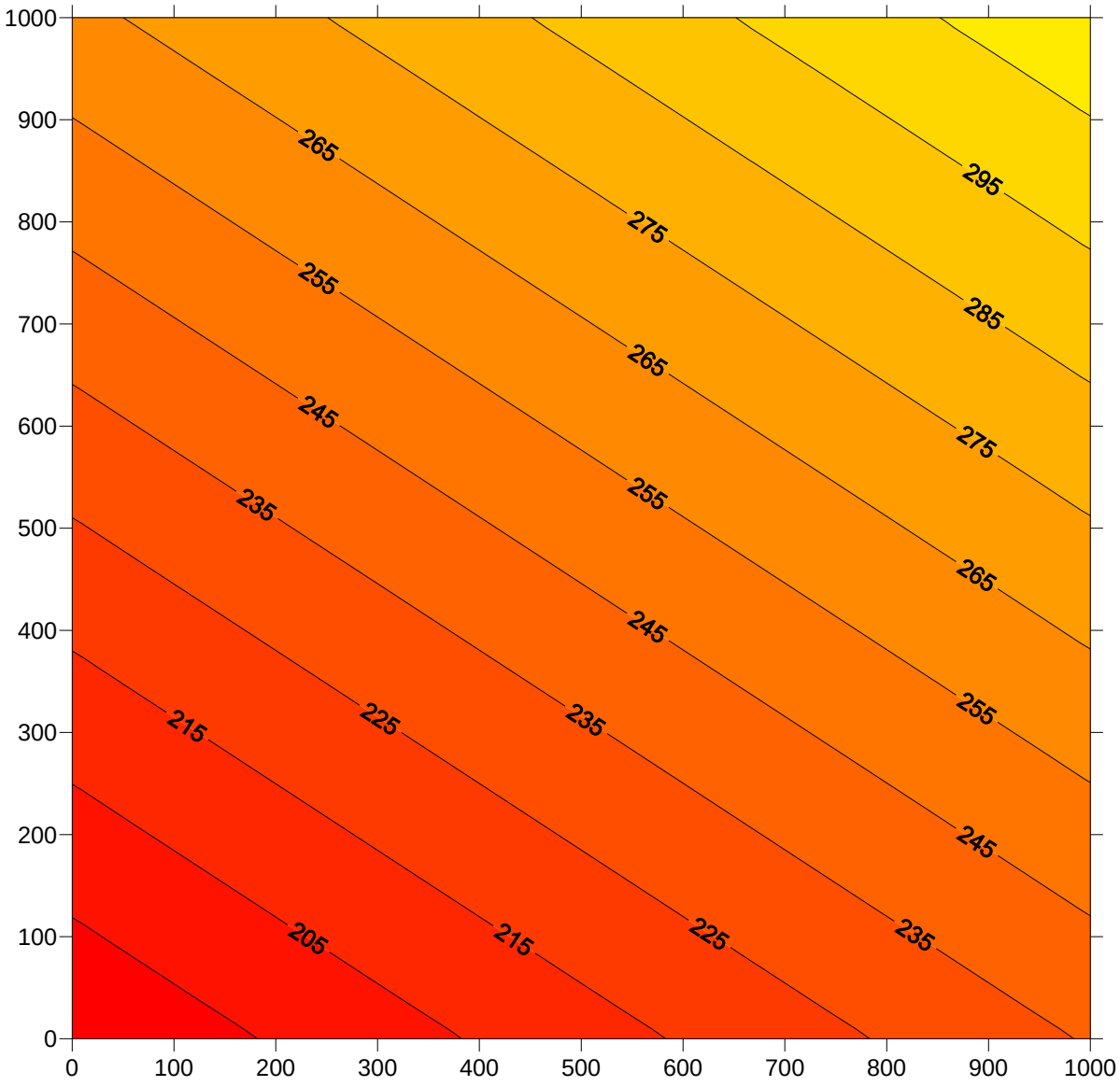


Figure 4.17. Contour map for regression-based predicted concentration V at Z = 25 ft (3-D case)

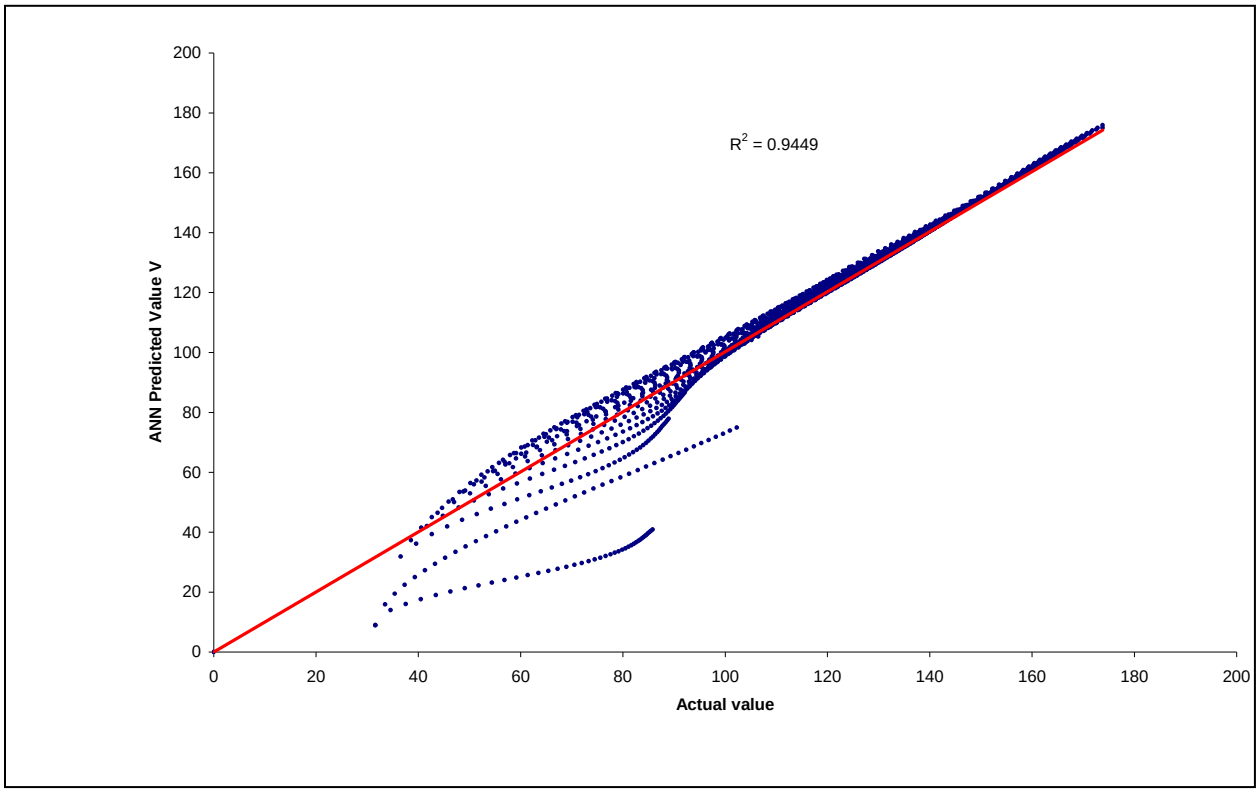


Figure 4.18. Scatter plot of actual "V" and "V" predicted by the neural network model (2-D and 25-ft interval case)

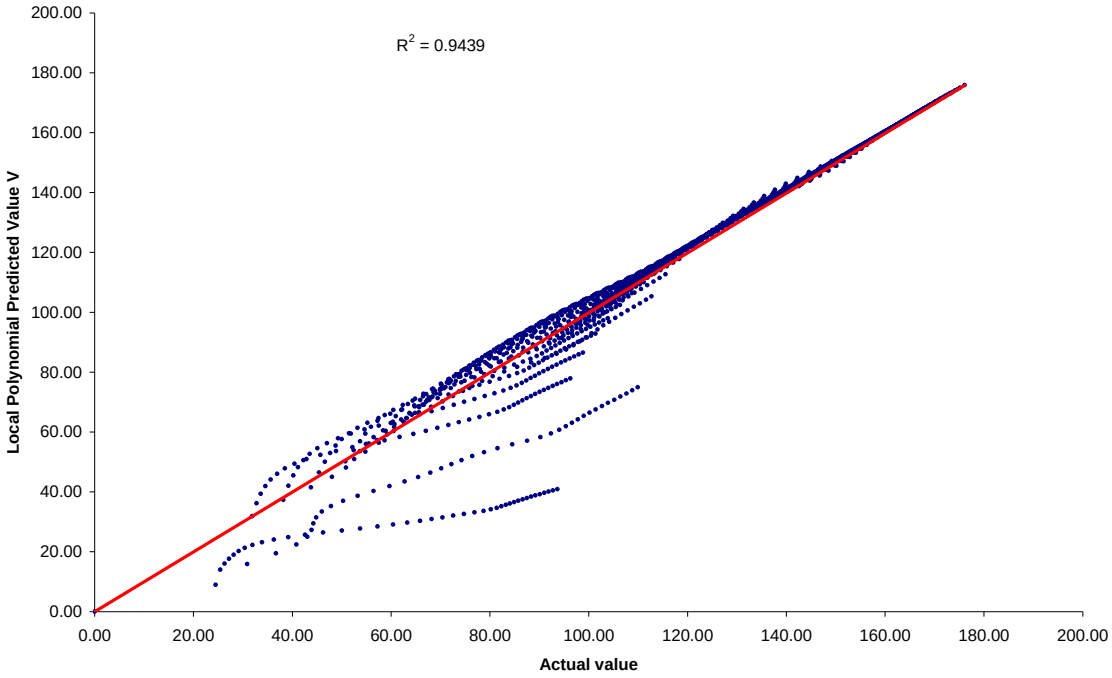


Figure 4.19. Scatter plot of actual "V" and "V" predicted by the Local Polynomial Method, (2-D and 25-ft interval case)

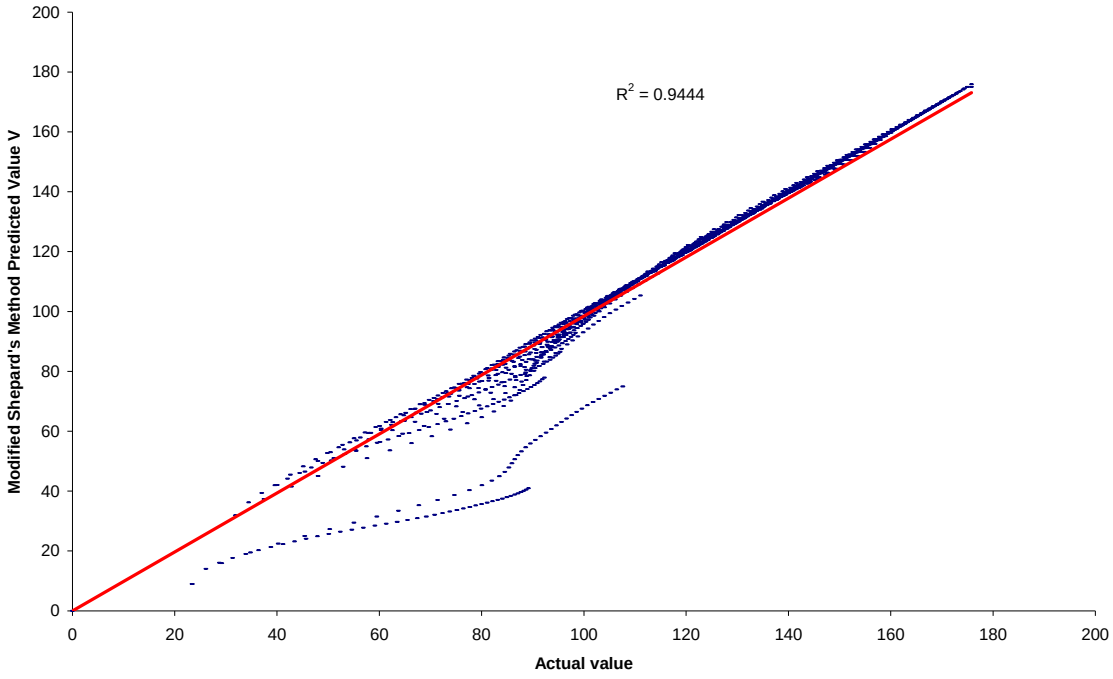


Figure 4.20. Scatter plot of actual "V" and "V" predicted by Modified Shepard's method (2-D and 25-ft interval case)

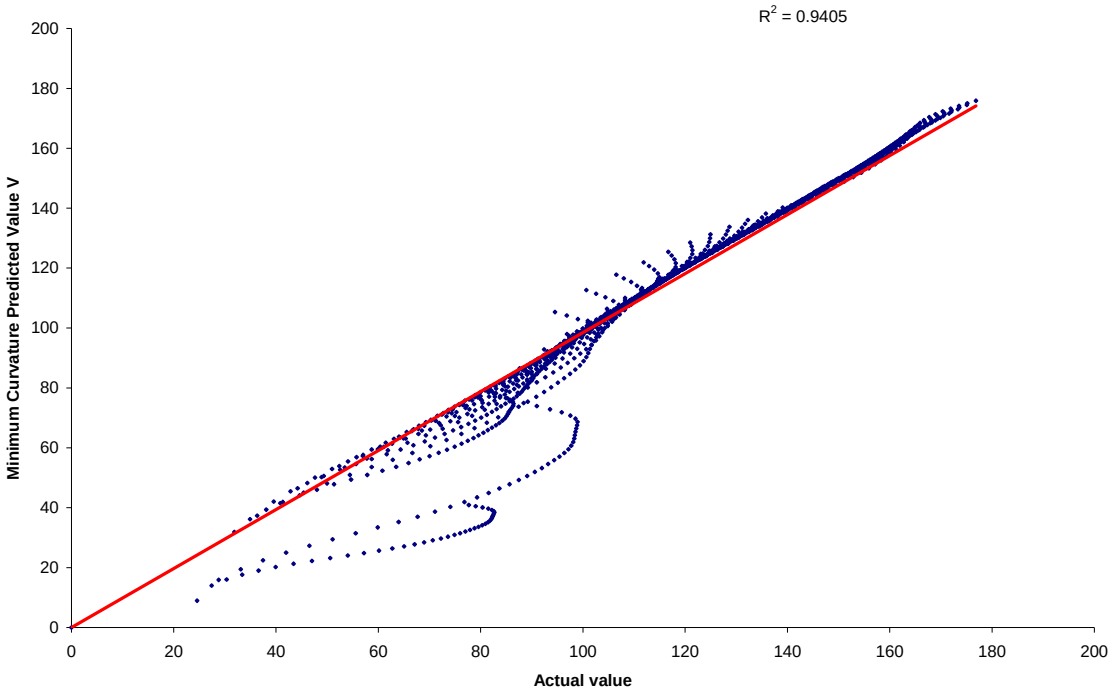


Figure 4.21. Scatter plot of actual “V” and “V” predicted by Minimum Curvature method (2-D and 25-ft interval case)

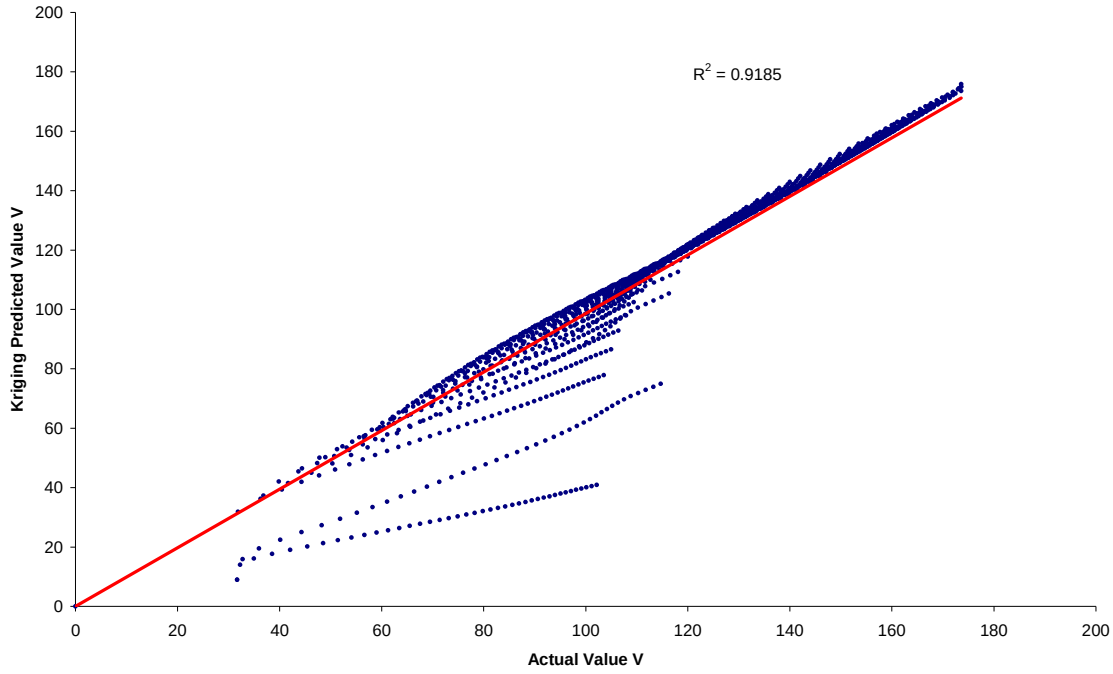


Figure 4.22. Scatter plot of actual “V” and “V” predicted by the Kriging method (2-D and 25-ft interval case)

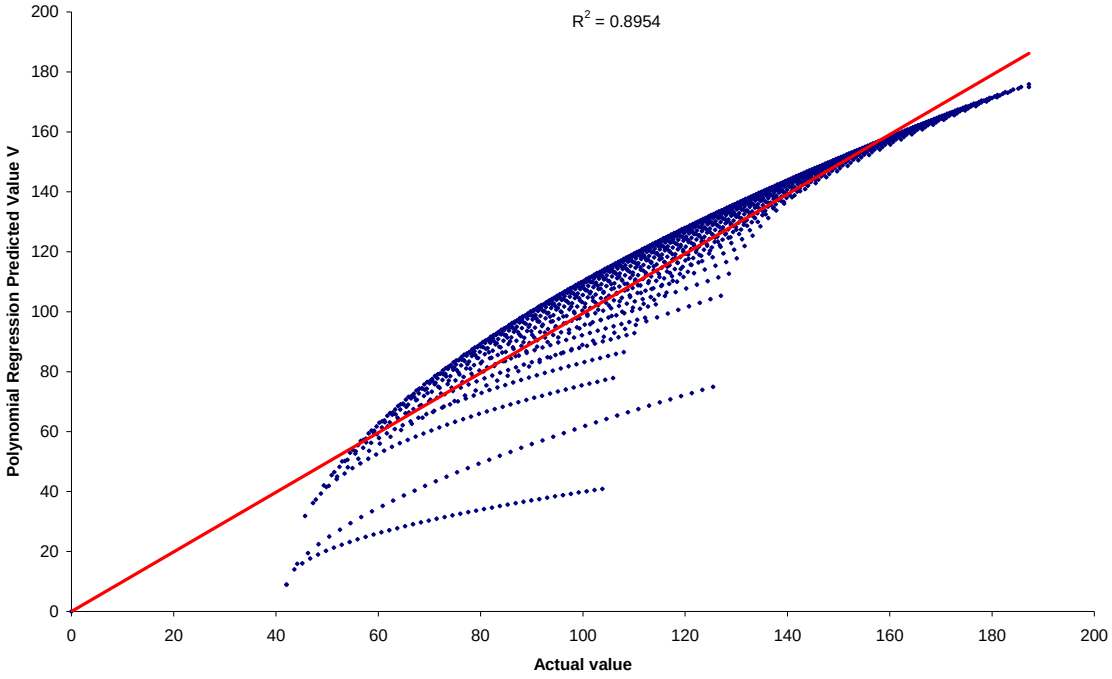


Figure 4.23. Scatter plot of actual “V” and “V” predicted by Polynomial Regression (2-D and 25-ft interval case)

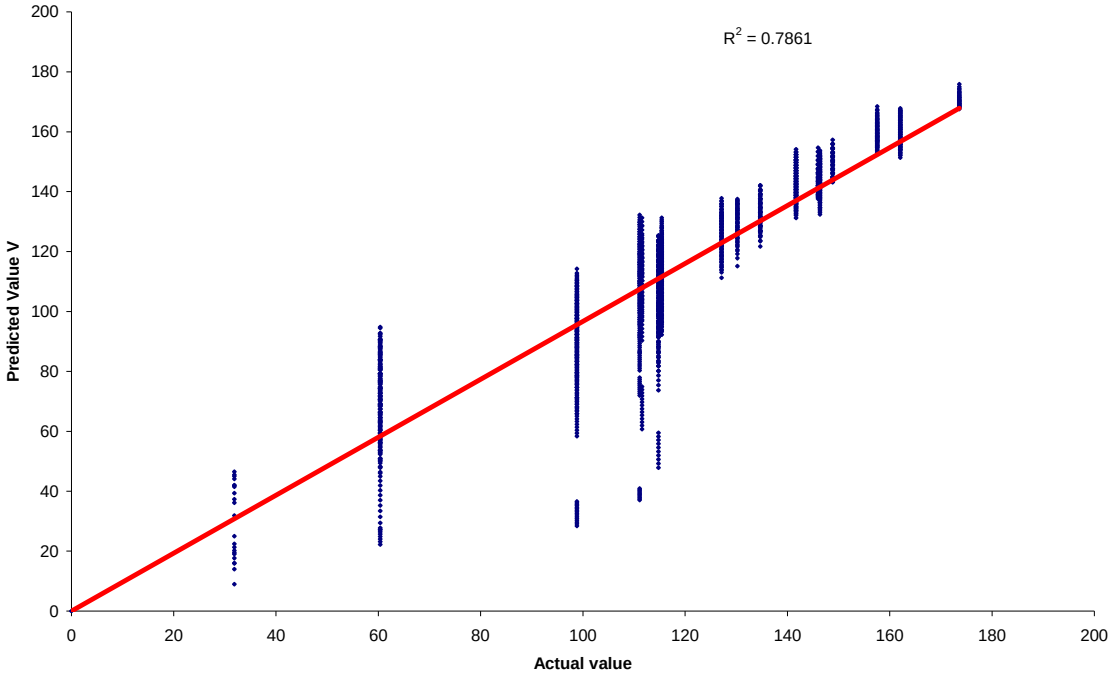


Figure 4.24. Scatter plot of actual “V” and “V” predicted by Nearest Neighbor method (2-D and 25-ft interval case)

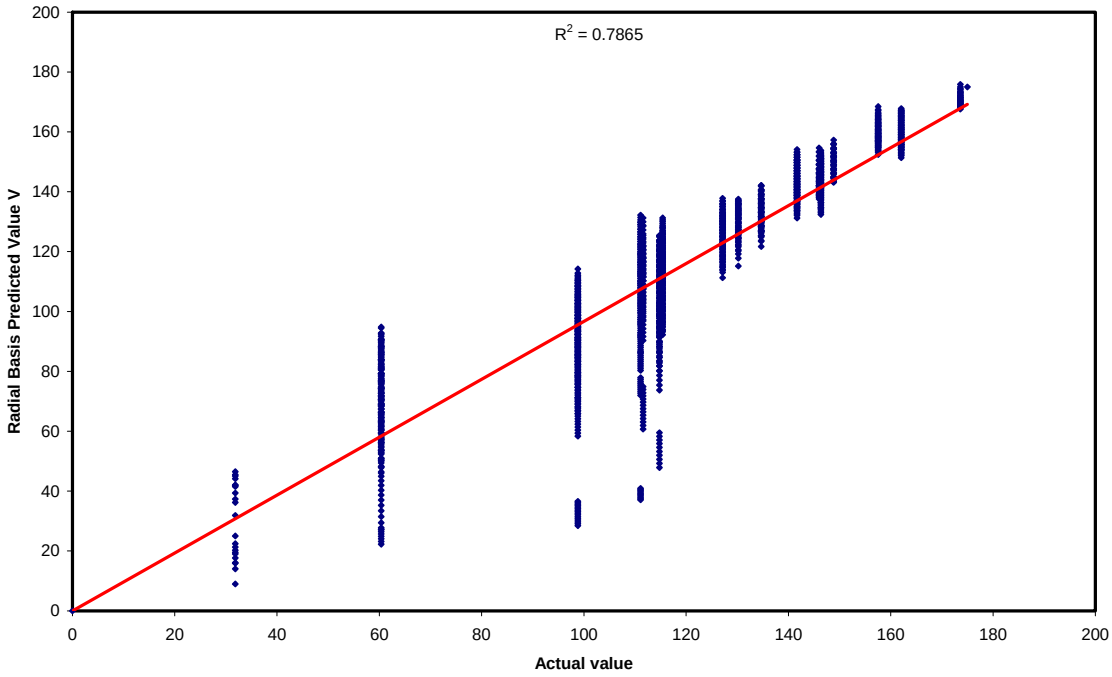


Figure 4.25. Scatter plot of actual “V” and “V” predicted by Radial Basis method (2-D and 25-ft interval case)

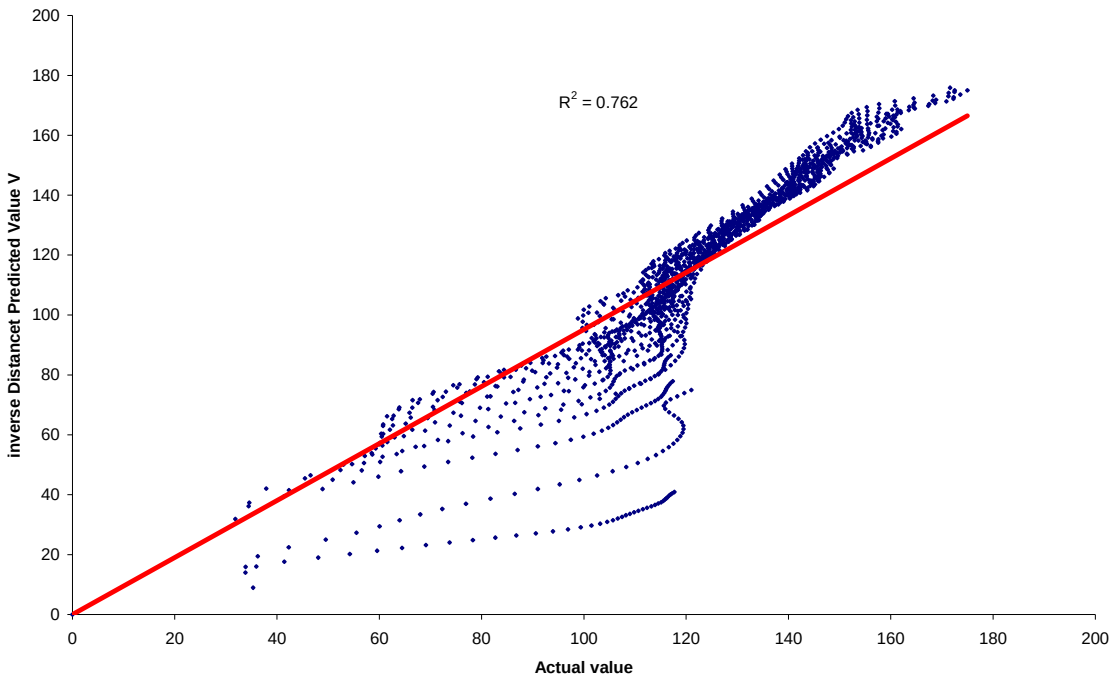


Figure 4.26. Scatter plot of actual “V” values and “V” values predicted by the Inverse Distance of Power (2-D and 25-ft Interval Case)

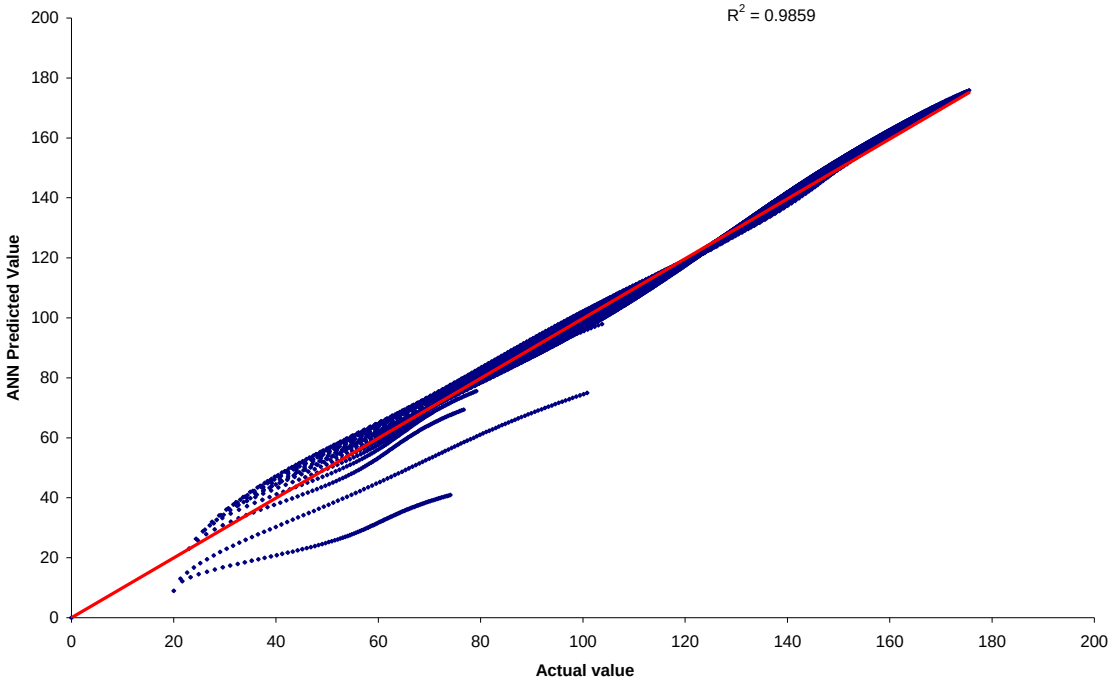


Figure 4.27. Scatter plot of actual "V" values and "V" values obtained by the Neural Network model (2-D and 10-ft interval case)

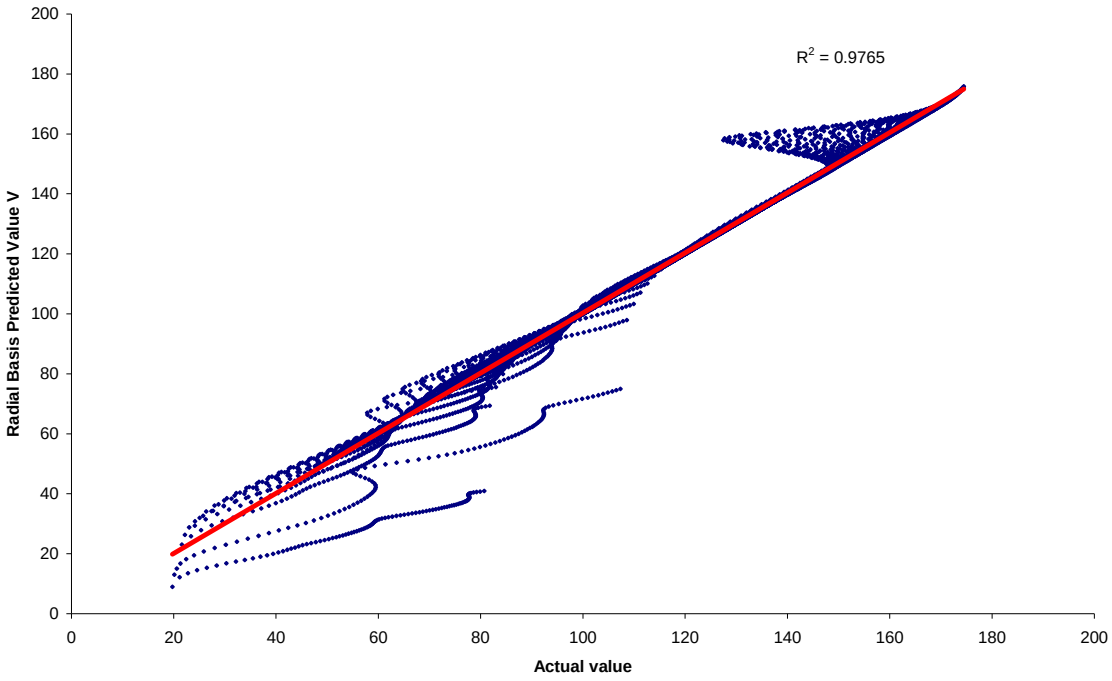


Figure 4.28. Scatter plot of actual "V" values and "V" values obtained by Radial Basis Function (2-D and 10-ft interval case)

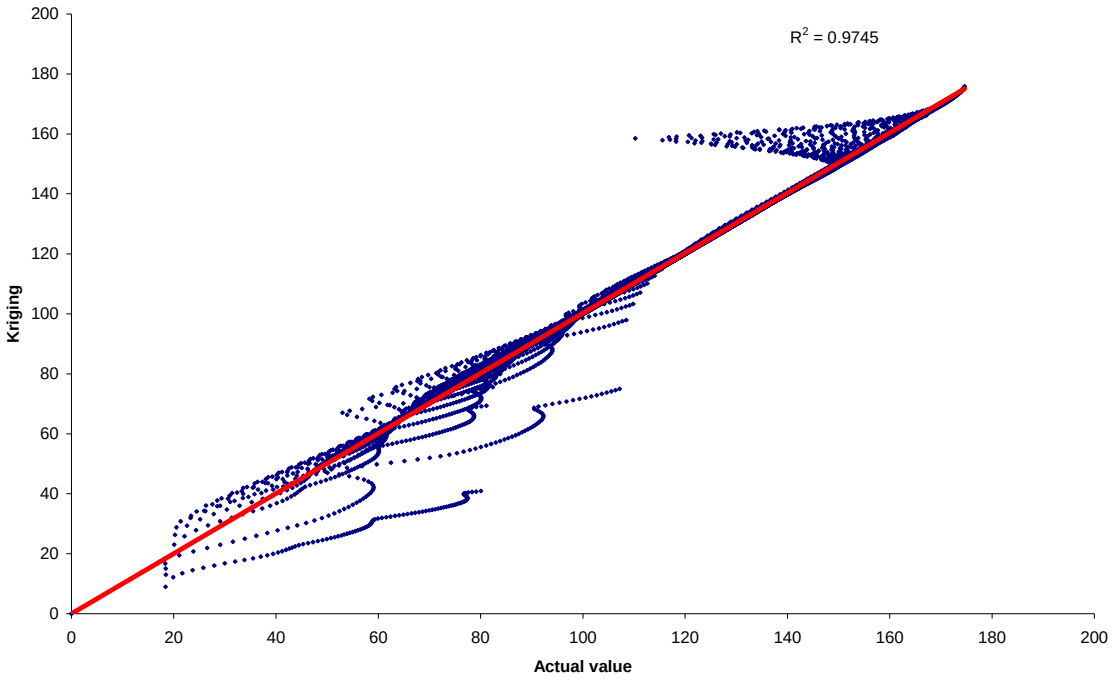


Figure 4.29. Scatter plot of actual “V” values and “V” values obtained by Kriging Method (2-D and 10-ft interval case)

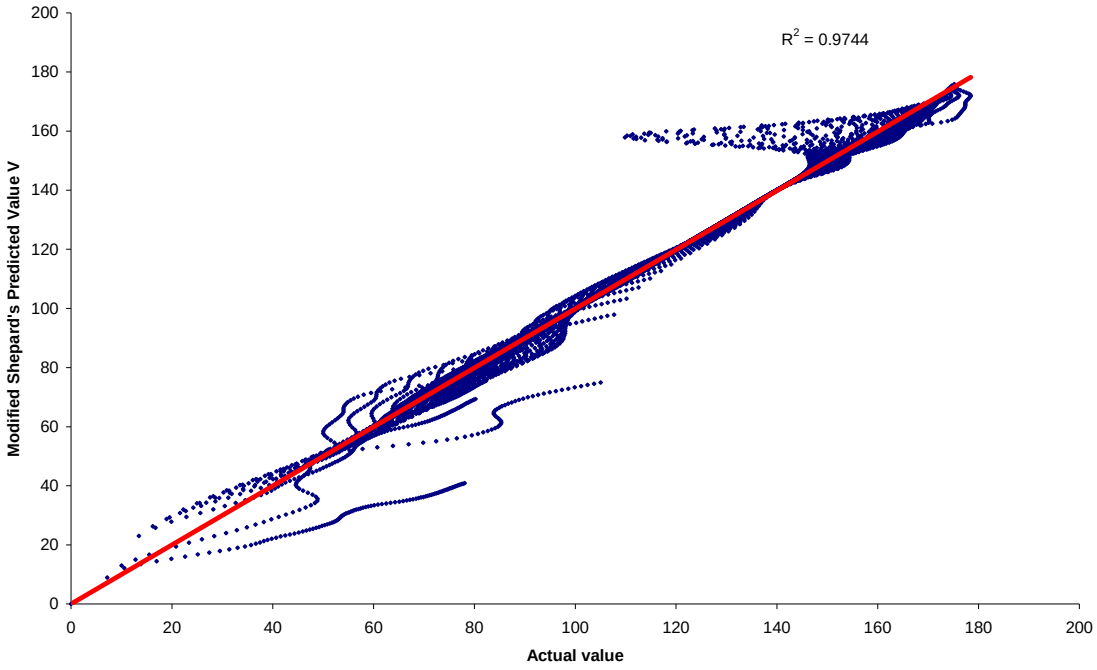


Figure 4.30. Scatter plot of actual “V” values and “V” values obtained by Modified Shepard's method (2-D and 10-ft interval case)

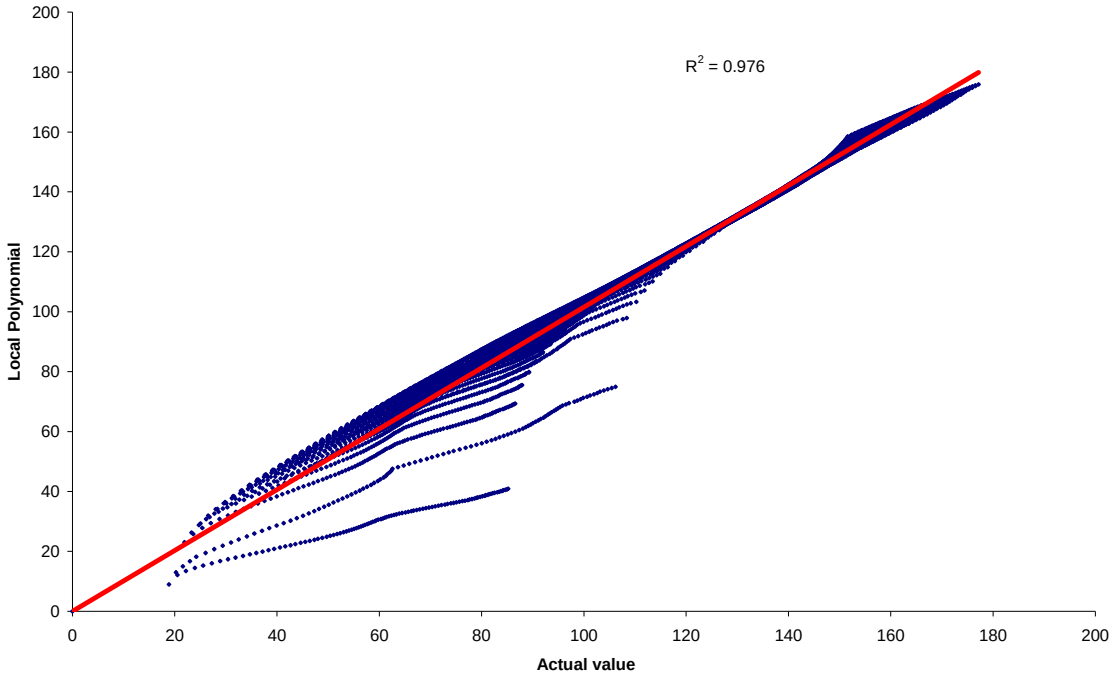


Figure 4.31. Scatter plot of actual “V” values and “V” values obtained by Local Polynomial (2-D and 10-ft interval case)

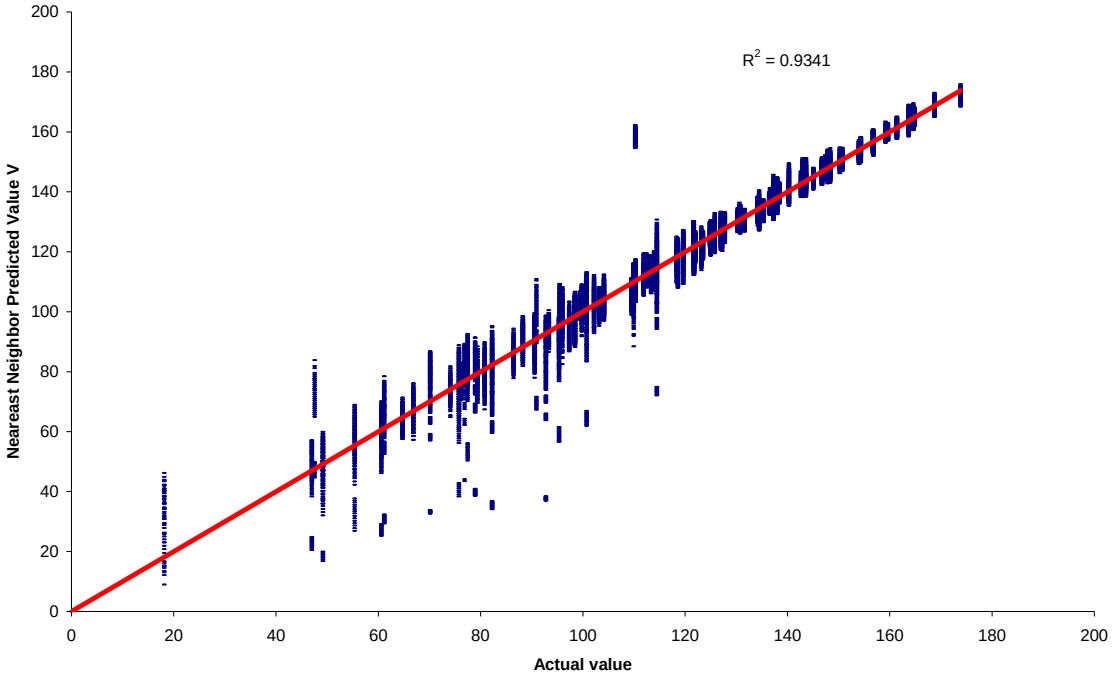


Figure 4.32. Scatter plot of actual "V" values and "V" values obtained by Nearest Neighbor (2-D and 10-ft Interval Case)

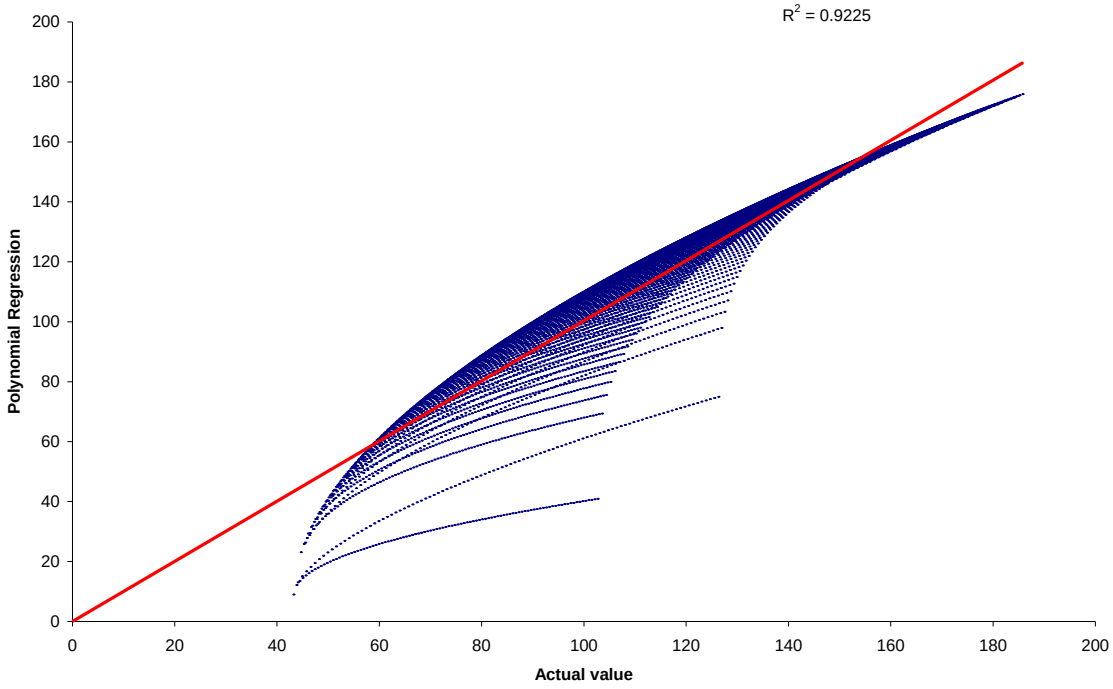


Figure 4.33. Scatter plot of actual “V” values and “V” values obtained by Polynomial Regression (2-D and 10-ft interval case)

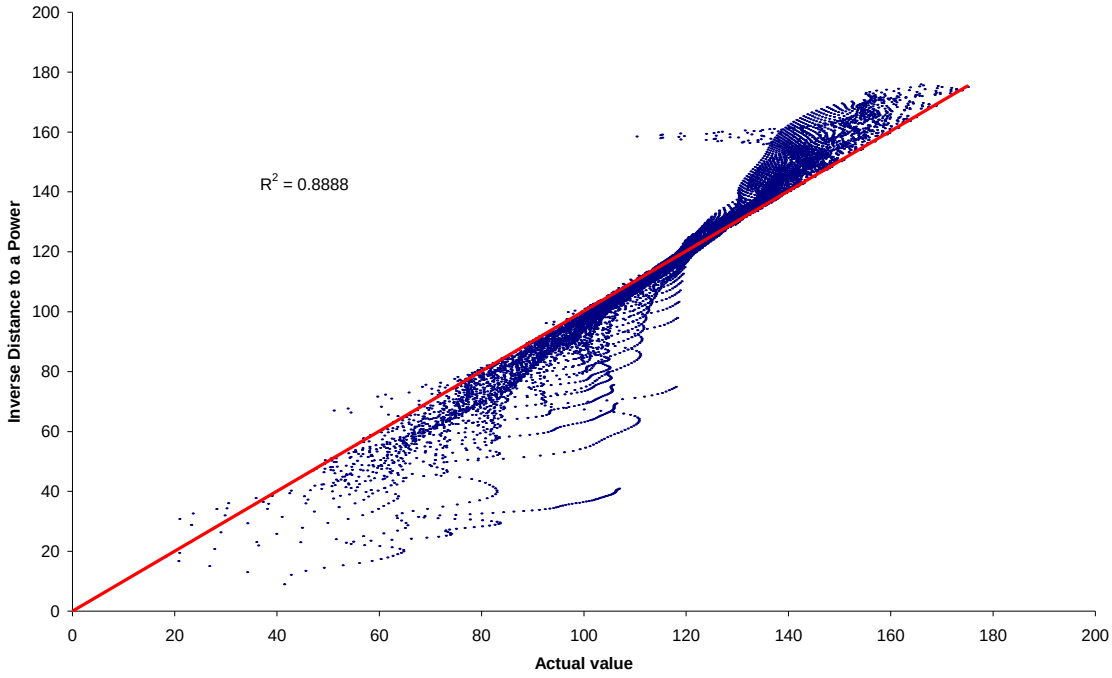


Figure 4.34. Scatter plot of actual “V” values and “V” values obtained by Inverse Distance to a Power (2-D and 10-ft interval case)

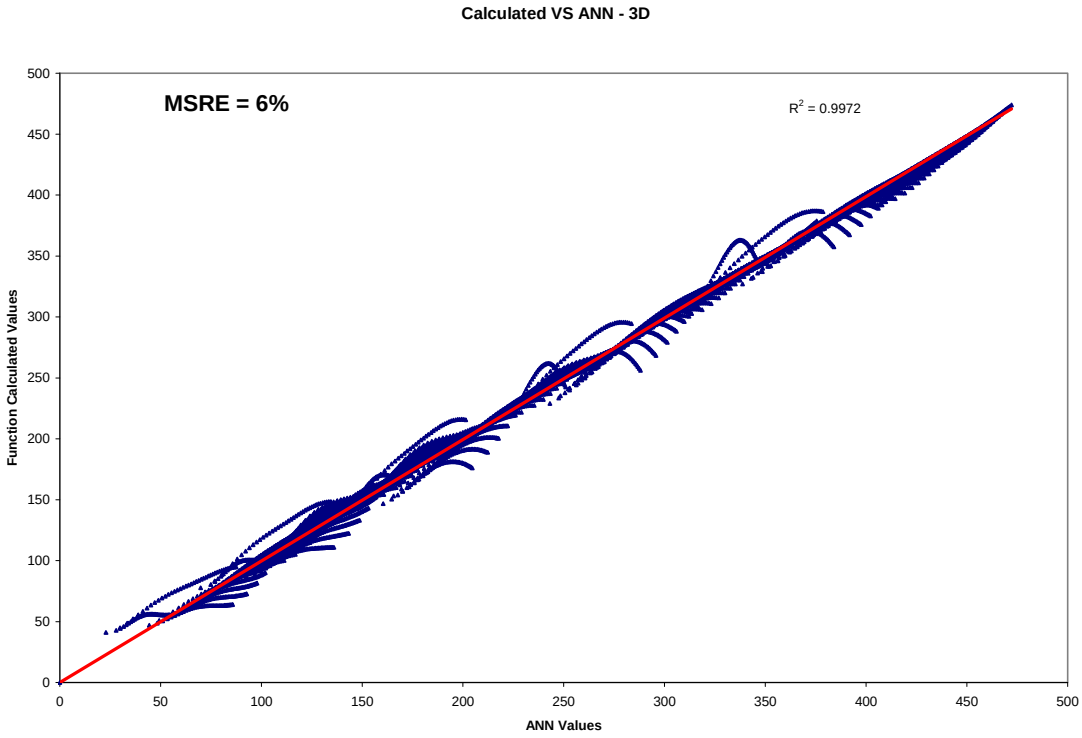


Figure 4.35. Scatter plot of actual "V" values and "V" values obtained by the ANN model (3-D case)

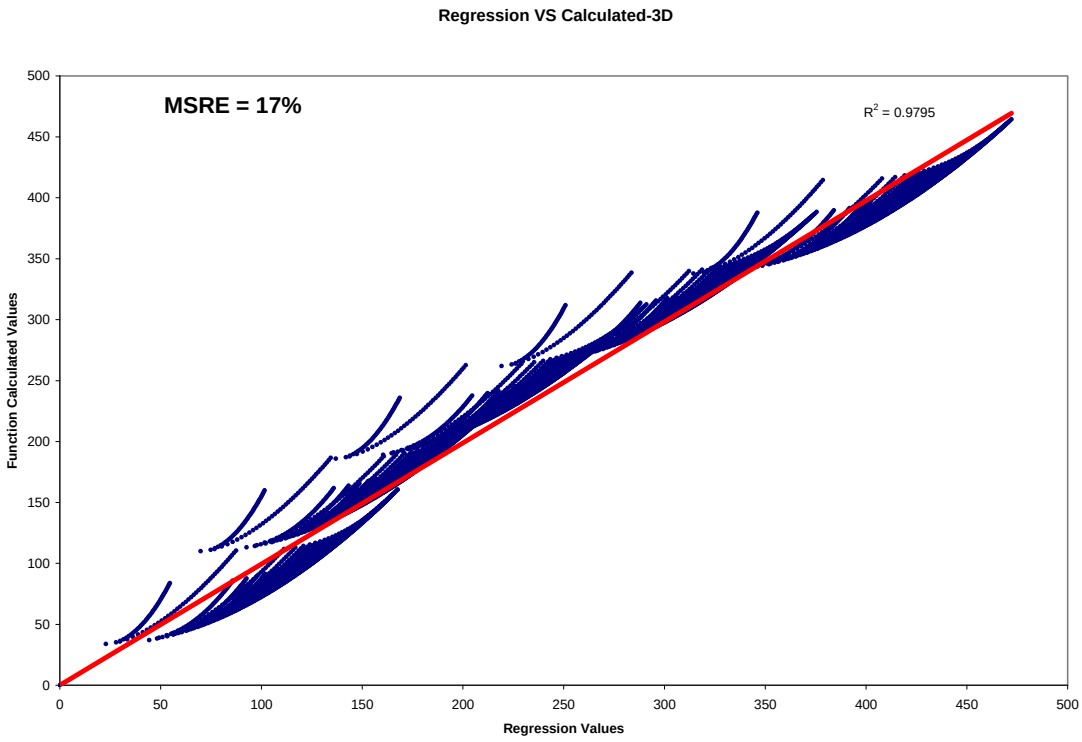


Figure 4.36. Scatter plot of actual “V” values and “V” values obtained by Regression-Based Method model

CHAPTER 5 - **Two-dimensional characterizations of the inorganic materials (lead and copper) in soil and groundwater at a landfill site**

5.1 Background Information

America is a throwaway society. With the conveniences of today's world come added waste: disposable items such as diapers, fast food packages and plastic bottles. All of these are meant to help Americans live in today's fast-paced society. In 1960, the average person produced approximately 2.7 pounds of waste per day. In 2001, this number almost doubled to approximately 4.4 pounds per day (USEPA, 2005). This adds up to over 229 million tons of waste per year, which eventually ends up in landfills around the United States.

The concept of an organized method for solid waste disposal began in the early 1960s. Prior to this, solid waste was disposed of by either burning or dumping of the waste into unregulated open dumps. The development of landfills provided a means for the disposal of solid waste, which was regulated and focused on the disposal of the waste in a manner that would minimize the negative effects on public health and the environment. By 1989, 7,379 landfills existed in the United States (BioCycle, 1999).

As time progressed, the number of landfills decreased due to the closure of older or filled landfills and the consolidation of smaller landfills. By 1999 (see Figure 5.1), the number of landfills had dropped to 2,216, while the amount of waste increased from 92 tons per day in 1989 to 300 tons per day (USEPA, 2005).

With the closure of over 5,000 landfills, new environmental regulations were necessary. In the 1980s, Federal and State regulations were enacted which required the installation of liners in landfills to help minimize the leakage of liquid waste materials into the groundwater system

(USEPA, 2005). Prior to this, landfills were unregulated and were built and operated without the environmental controls and regulations of today (Grossman, 2002). Without liners in place, precipitation and ground water will seep through the solid waste, producing contaminated water referred to as *leachate*. The leachate then seeps into the groundwater beneath the landfill, causing the groundwater to become contaminated. This contaminated groundwater is referred to as a *plume*. Because of the normal flow of the groundwater, the plume extends away from the landfill, causing further contamination. To determine the extent of the contamination, field sampling and laboratory analyses of soil and water samples must be performed.

In Kansas City, Kansas, one such landfill has been at the center of controversy for years. In 1951, this site was established to dispose of debris from the largest and most devastating flood of the century. In a ten year period, between 1953 and 1963, a prominent fiberglass company utilized the site for disposal of process wastes from the fiberglass production plant located in the Fairfax Industrial District, approximately three miles to the east. A large but unknown amount of waste fiberglass from the factory was buried in this landfill. In addition, other authorized and possibly some unauthorized dumping took place. The City of Kansas City used cinders from the local electric power plants as well as rock and soil to cover the waste material. The site was closed in 1963 and contents of the dump covered with an unknown amount of local soil (KDHE, 1989).

5.2 Landfill Site Description

According to the Kansas Department of Health and Environment (KDHE) (1996), the landfill is located near the intersection of North 18th Street and Ridge Avenue in Kansas City, Kansas. It covers an area of about 5.5 acres and is located in the uplands adjacent to the north flank of the Kansas River valley, within the city limits of Kansas City, Kansas. Drainage from

the site empties into a storm sewer, which in turn empties into an intermittent stream located adjacent to the southwest boundary of the property. Off-site drainage is south via this intermittent creek to its confluence with the Kansas River, approximately one mile south of the site.

Several years after the landfill opened, numerous complaints concerning odors and the seepage of leachate into a ravine located southwest of the landfill were reported. As a result of these complaints, the Wyandotte County Health Department recommended, in July 1958, that the landfill be closed and also that future disposal of material by the fiberglass company be within a tight clay soil in a sanitary landfill. However, both recommendations were rejected by the Kansas City officials.

Complaints revolving around the landfill continued intermittently for more than 40 years. These complaints remained unresolved, due to the fact that none of the entities involved in the landfill could agree on who was at fault for the contamination. The fiberglass company continued to utilize the landfill to dispose of their company's waste. In 1992, litigation began in an effort to resolve the issues surrounding the landfill site.

5.3 Site Investigation

In 1996, the investigation of the Kansas City landfill began. Records provided by the City and the fiberglass company showed that the following waste products had been disposed of at the landfill site (KDHE, 1996):

- Various metal sludge and grinding bottoms
- Solvents
- Phenolic resins

- Adhesives consisting of approximately 70% phenol-formaldehyde and 30% Vinsol (a by-product of turpentine formulation)
- Furnace refractory brick, demolition debris, asphalt, paper, and cinders from coal-fired power plants

The landfill was built without a protective liner, as was the practice during that era. Because of this, the previously mentioned waste products, along with other unauthorized waste products, were able to seep into the groundwater and produce leachate.

5.3.1 Information Gathering

At the start of any remediation process, it is important to obtain accurate, in-depth information regarding the extent of contamination at the investigated site (Itani & Najjar, 2000). This is typically achieved by performing a sample collection and a site investigation analysis. There are many factors that determine where and how samples are collected. Generally, sample locations are determined using the professional judgment of the site investigation team (Mryyan & Najjar, 2005). The investigation team should take into consideration the following factors when deciding upon sample locations:

- Soil type and geotechnical properties of the landfill
- Size of the landfill
- History of the site
- Site-related risk factors
- Sampling expenses

Although the most accurate results would be obtained by collecting a large number of samples from all areas of the landfill, this is not generally financially feasible (Najjar & Basheer, 1996). Alternatively, a small number of samples are normally collected at different points across

the landfill site and then sent to the laboratory for chemical analysis. This field data, along with the BPANN approach, were used herein to predict the extent and location of copper and lead contaminates within the Kansas City landfill area.

5.4 Database

In order to determine the distribution of contaminants at the Kansas City landfill site, samples of groundwater and soil were obtained at various locations throughout the landfill. Groundwater sampling was performed by drilling and installing four monitoring wells at the site. Well-related information (such as locations, depths, and so forth) are depicted in Table 5.1 and Figure 5.2. The wells were installed to detect the presence and migration of contaminants from the landfill site. Monitoring Well (MW) 1 was originally designated as the background well, while MW-2, 3 and 4 were installed in the presumed, in relation to the dump site, down-gradient direction (KDHE, 1996). Once the monitoring wells were in place, groundwater samples were collected and sent to the laboratory for analysis.

Similar to the groundwater sampling procedure discussed previously, soil sampling was conducted by collecting seven soil samples from depths varying from 0–12 inches from ground surface. Information (such as number, location, and chemical analysis results) related to the seven samples used herein in our ANN modeling task are given in Table 5.3 and 5.4 as well as Figure 5.3.

5.4.1 Model Development

In order to utilize any ANN-based model, the program must be trained or educated about the process it is supposed to model. To train the network, a known set of input data along with the desired outcome is used. The BPANN methodology/program [Mryyan & Najjar (2007); Itani

& Najjar (2000); Najjar & Basheer (1996); Itani (1996)] using the supervised training approach is used to train the desired ANN models to produce output values that are as close to the real values as possible via repeated modifications of the network's connection weights. This process continues until the error at the output layer is minimized. Once this training process has been completed, the developed model can then be used for prediction tasks. Note that the accuracy of the predicted values is dependent on the quality of the data used in the training phase. The better the quality of the training sets, the greater the accuracy of the predicted values will be. For this reason, the training sets (i.e., groundwater and soil data) used to build the desired network models were of the utmost importance in this study.

When developing any ANN model, it is important to determine what input and output values will be used (Dowla & Rogers, 1995). For the Kansas City landfill case, x and y coordinates were used as the only input values to the model. The concentration value (V) of lead or copper was used as the output for their associated network model. The x and y coordinates refer to the x and y distances for the associated observation point, measured from a reference point (i.e., x = 0, y = 0). The value of lead in soil network model was developed using five data sets for training and the remaining two data sets for testing purposes. Best network and associated number of hidden nodes were determined by training and online testing to achieve the least error on the testing data sets. Accuracy measures used in this study are:

- ASE = Averaged squared error
- MARE =Mean Absolute Relative Error and
- R^2 = Coefficient of Determination.

Best net is identified as the one having the least ASE and MARE and highest R^2. In this case, the number of hidden nodes needed to achieve this objective was found by the adaptive

training approach. The final (best performing) net contained two input nodes representing the x and y coordinates, two hidden nodes and one output node (i.e., value of lead). Similar modeling processes (i.e., training, testing and evaluation) were carried out to select the best performing network model for:

- Value of lead in groundwater table (GWT) at depth (Z) = 2 feet
- Value of lead in GWT at Z = 4 feet
- Value of lead in soil
- Value of copper in soil

For all four networks developed herein, two hidden nodes were found adequate to achieve the desired best performing net.

5.4.2 Databank Generation

A contaminated location, for the purposes of this study, is defined as any x and y coordinate location that contains lead or copper value that is higher than the Maximum Allowed Contamination Level (MACL). Table 5.5 lists the containment and their associated MACL values. A sampling location that has been observed to have a concentration value higher than MACL will be designated as a contaminated area or hot area, and therefore would require remediation. On the other hand, any location having a concentration value less than MACL value will be considered as an uncontaminated zone or safe area, and therefore requires no remediation.

Four databases containing x, y and V values were generated via the developed ANN models for each case at various locations across the site. To achieve this objective, the landfill site was divided into a grid system. In this case, grid lines were set at 10-foot intervals for both x (east) and y (north) directions. A total of 481 grid points were used for each case (See Figure

5.4). The x and y coordinates were used as input values for each of the four developed ANN profiling models. The developed models were then used to predict the corresponding contamination value for the 481 designated x and y coordinates representing the site. This produced four data banks containing 481 sets of x, y coordinates and their predicted V value for each case. The resulting data banks were processed to construct various contamination distribution maps of the landfill site.

5.5 Results and Analysis

The initial soil and groundwater samples indicated levels of copper and lead above the MACL. These findings were significant because both metals can have harmful effects on humans and animals with extended exposure. The data generated by the ANN models were processed via a software program called Surfer® 8.0. This program was used herein to produce contamination concentration contour maps for the landfill site using the previously mentioned x, y and V data banks. In doing so, two types of contamination distribution maps were produced for each case. A total of eight maps were produced for the four cases considered in this study. The produced maps can easily be used to identify hot (areas whose concentration is above MACL) and relatively safe zones (where concentration is below MACL).

5.5.1 Soil-Based Maps

Figures 5.5 through 5.8 show the generated maps for the soil case. Based on these figures, it can be observed that the greatest degree of copper and lead concentrations is located in the areas spanning the northeast corner of the landfill and down through to the southwest corner. Note that most of the original soil samples (whose concentration values are higher than MACL)

fall within the area of high contamination levels, which is predicted via the developed ANN-based profiling models.

5.5.2 Groundwater-Based Maps

Figures 5.9 through 5.12 show the generated maps for the groundwater cases. According to Figures 5.9 through 5.12, the greatest degree of lead concentration occurs diagonally from the upper northwest corner to the lower southeast corner. The entire area in the southwest corner of the landfill is contaminated. As was observed in the soil case, most of the original groundwater sample points (containing values above MACL) fall also within the contaminated area predicted by the groundwater-based network models. Comparing the groundwater contamination maps (i.e., Figures 5.9 and 5.11) obtained at depths of 2 and 4 feet, it can be observed that there is a clear difference in the spatial distribution of the contaminate. There is a lesser degree of contamination at 4 feet depth. This indicates that the majority of the contamination remains closer to the surface. This observation is of great importance, because it limits the depth that needs to be reached during the remediation process. As a result, this will yield substantial reduction in the associated cost needed for the remediation phase.

5.6 Concluding Remarks

The proper use of ANN methodology for contaminate profiling, demonstrated in this study, made it possible to obtain logical predictions about the location and extent of lead and copper contamination at the associated Kansas City landfill site. The resulting profiles can be used to determine additional sampling locations, if needed, for both groundwater and soil hot zones. Moreover, extent of remediation zones can be assessed properly, thereby reducing the associated cost needed for further sampling purposes and/or remediation tasks.

Table 5.1 Monitoring Wells Related Information

Well No.	Location	Depth of Well and Screen Interval	Well Log Description
MW-1	North side of the landfill	42.5 ft; screen depth 32–42 ft	0–33 ft misc. rock, cinder fill and fiberglass fill, 33-37ft, medium brown silty clay
MW-2	West side of landfill	42.5 ft; screen depth 27–37 ft	0–7 ft misc. rock, sand, and cinder fill, 7-42.5ft, medium blackish-brown silty clay
MW-3	Southwest side of landfill	15 ft; screen depth 5–15 ft	0–7 ft misc. rock and black cinder fill, 7-14 ft, loose, wet, black silty clay
MW-4	South side of landfill	39 ft; screen depth 29 –39 ft	0–7 ft misc. rock, cinder and trash fill. 7-39 ft. medium brown silty clay

Table 5.2 Metal Data from Groundwater Samples (Z= depth, in feet, from groundwater table level)

Metal	MW-1 (Z=2)	MW-1 (Z=4)	MW-2 (Z=2)	MW-2 (Z=4)	MW-3 (Z=2)	MW-3 (Z=4)	MW-4 (Z=2)	MW-4 (Z=4)	Unit used
Aluminum	9700	16000	40000	51000	43000	48000	120000	97000	µg/l
Antimony	1.2	2.5	1.8	1.8	2	2.4	2.6	3.4	µg/l
Arsenic	7.6	10	24	27	30	36	30	26	µg/l
Cadmium	1.5	1.6	2.9	3.6	3	3.7	5	4	µg/l
Chromium	10	59	54	68	60	63	82	35	µg/l
Cobalt	5.5	7	26	34	27	29	37	22	µg/l
Copper	14	3.3	50	63	53	57	160	130	µg/l
Iron	1200	160	11200	14700	12000	13600	36000	11000	µg/l
Lead	30	25	52	69	64	71	96	68	µg/l
Magnesium	190	70	600	780	670	690	15000	111000	µg/l
Selenium	0.8	2.3	4.5	14	3.2	2.8	7	5	µg/l
Vanadium	110	11	77	100	92	97	115	109	µg/l
Barium	650	270	290	330	460	360	2200	1500	µg/l

Table 5.3 Soil Sample Number and Location

Sample Number	Sample Location
A96015-1	West facing slope
A96016-2	West facing slope
A96017-3	West facing slope
A96018-4	Southwest corner
A96019-5	South facing slope
A96020-6	South facing slope
A96021-7	Background

Table 5.4 Metals Data from the Soil Samples in Milligrams per Kilograms (PPM or mg/kg)

Analysis	A96-1	A96-2	A96-3	A96-4	A96-5	A96-6	A96-7	Unit Used
Aluminum	27659	46034	29412	35961	35300	26410	23329	mg/kg
Antimony	5	5	5	5	5	5	5	mg/kg
Arsenic	11	17	18	23	16	12	9	mg/kg
Barium	272	433	251	118	115	254	304	mg/kg
Beryllium	3	9	7	7	7	1	1	mg/kg
Boron	33	99	88	120	115	20	22	mg/kg
Cadmium	1	1	1	1	1	1	1	mg/kg
Cobalt	16	32	23	35	34	10	13	mg/kg
Copper	33	68	54	74	66	21	26	mg/kg
Iron	57799	150282	94778	117332	119261	20765	27707	mg/kg
Lead	68	81	135	136	92	49	28	mg/kg
Magnesium	3257	3578	2338	2362	2107	4023	4728	mg/kg
Manganese	646	715	394	396	344	938	595	mg/kg

Table 5.5 Maximum Allowed Contamination Level Values

Contaminant	**MACL**
Copper in groundwater	1 PPB (Part per billion)
Copper in Soil	11PPM (Part per million)
Lead in soil	50 PPM
Lead in groundwater	5 PPB

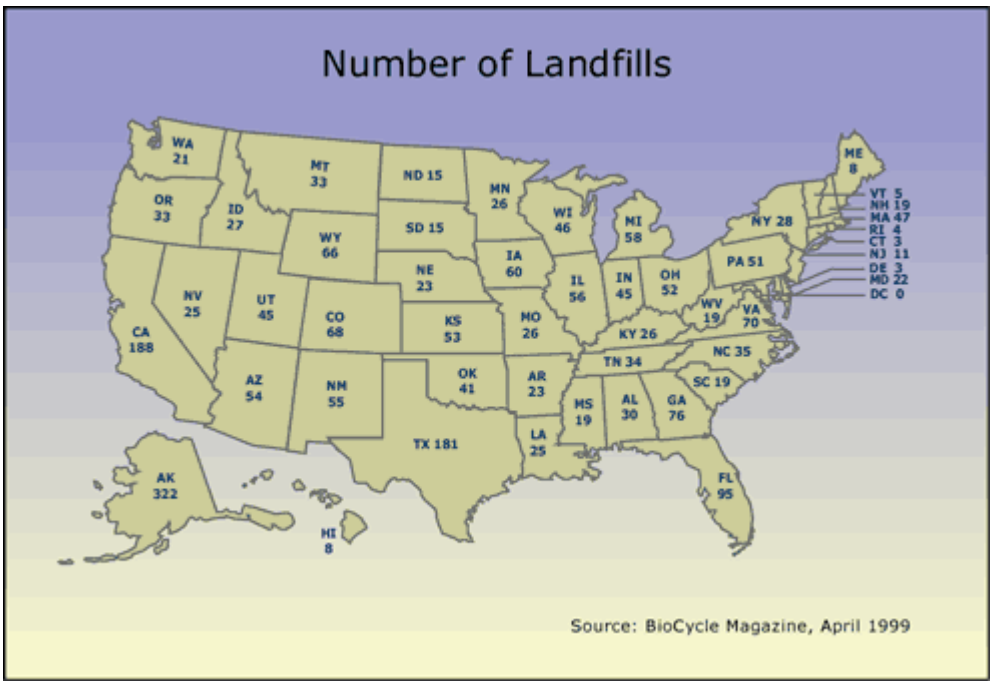


Figure 5.1. Number of landfills in the U.S. as of 1999

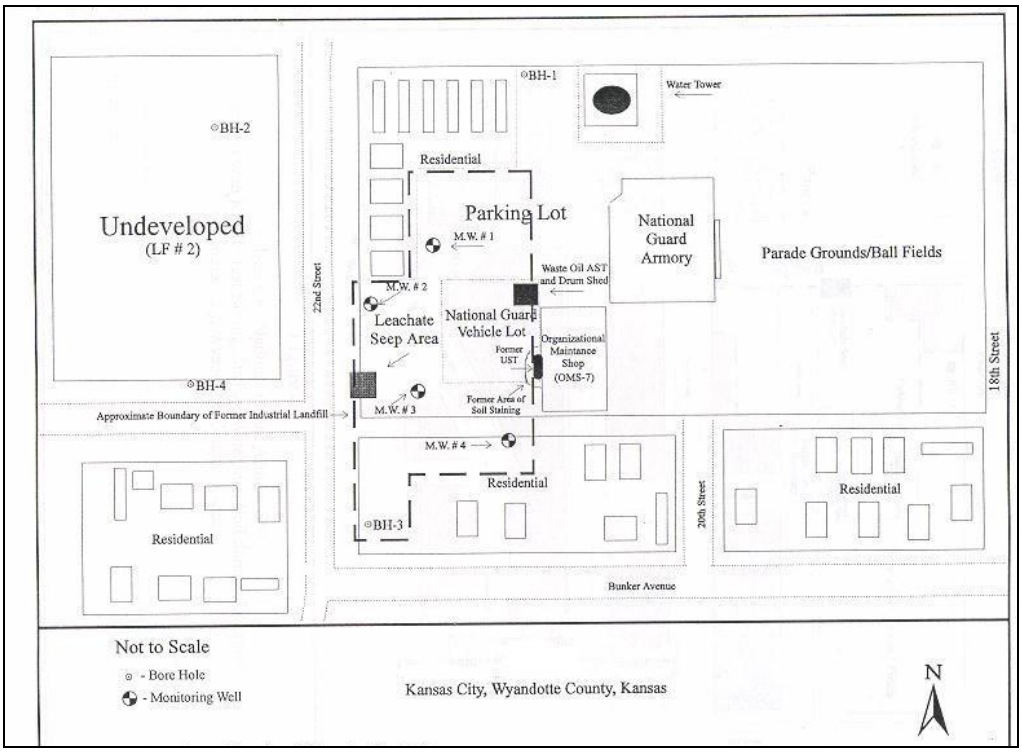


Figure 5.2. Location of monitoring wells used in this study

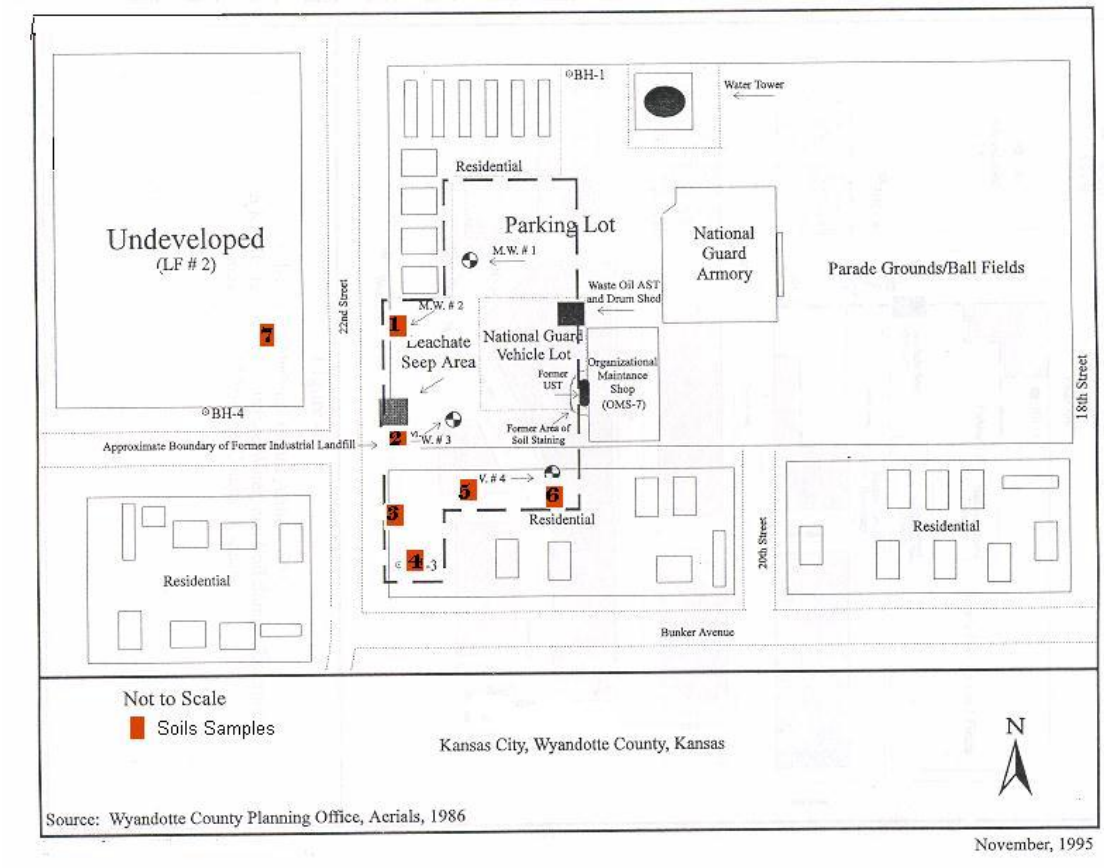


Figure 5.3. Location of soil samples used in this study

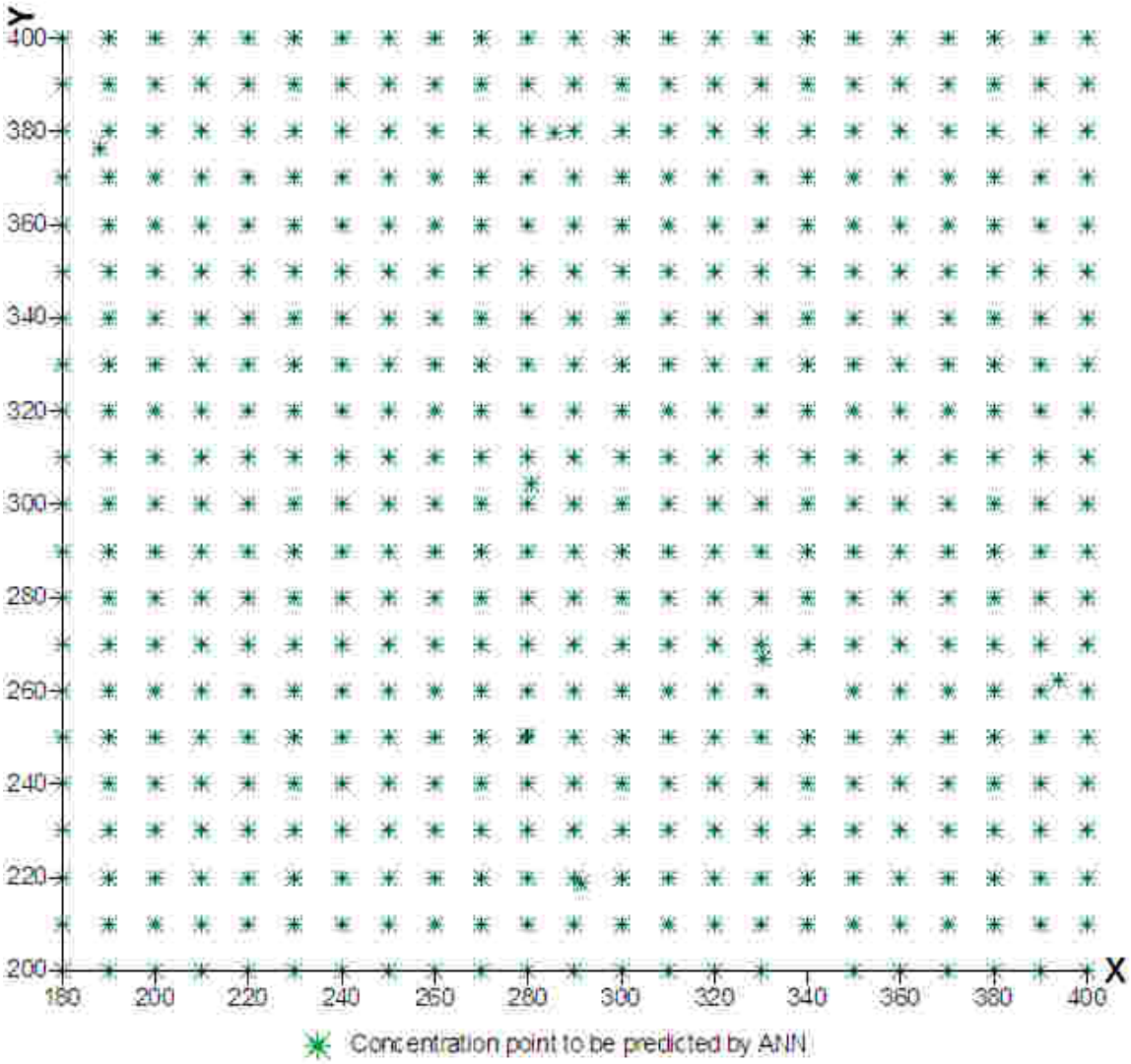


Figure 5.4. X and Y grid used by the ANN models to predict corresponding contamination value (V)

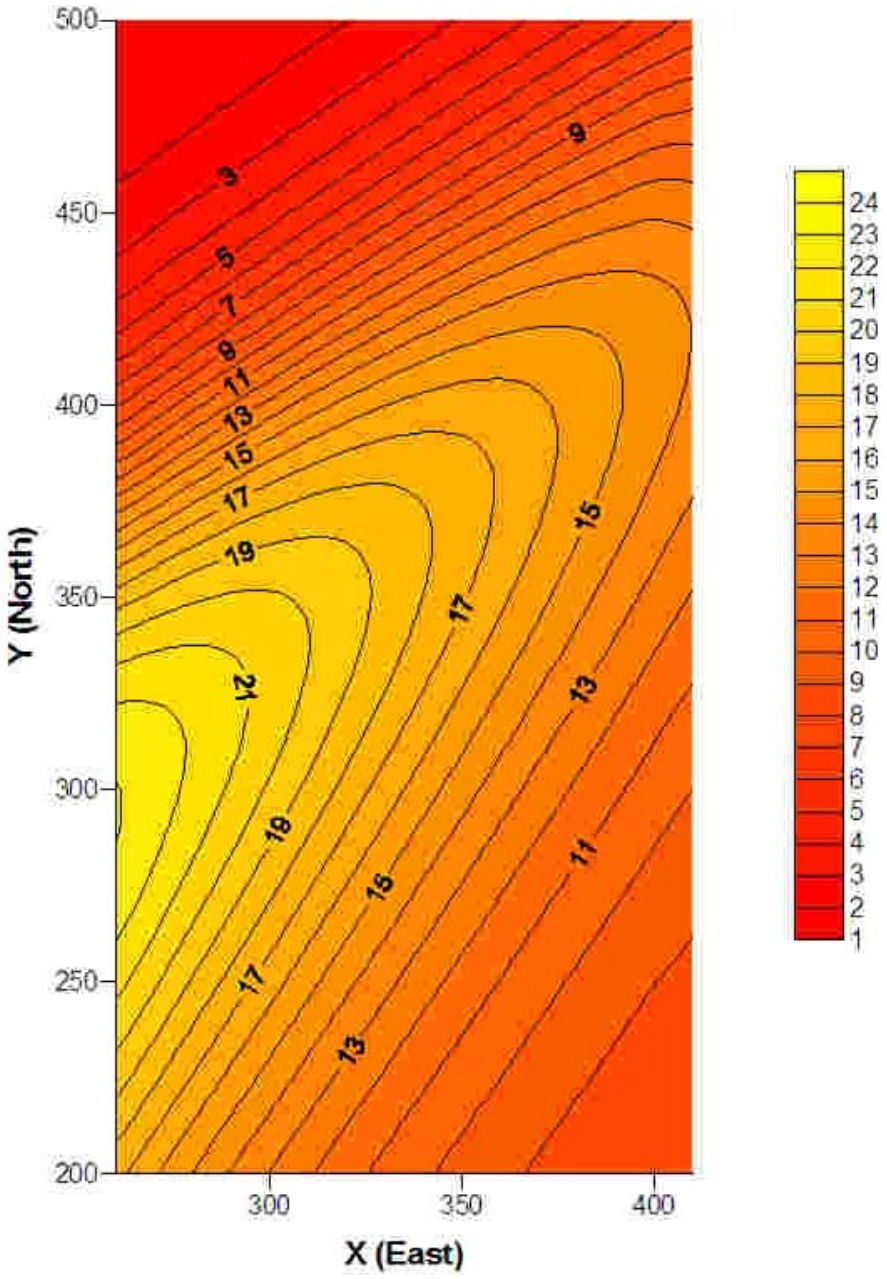


Figure 5.5. Contour map showing the distribution of copper in soil

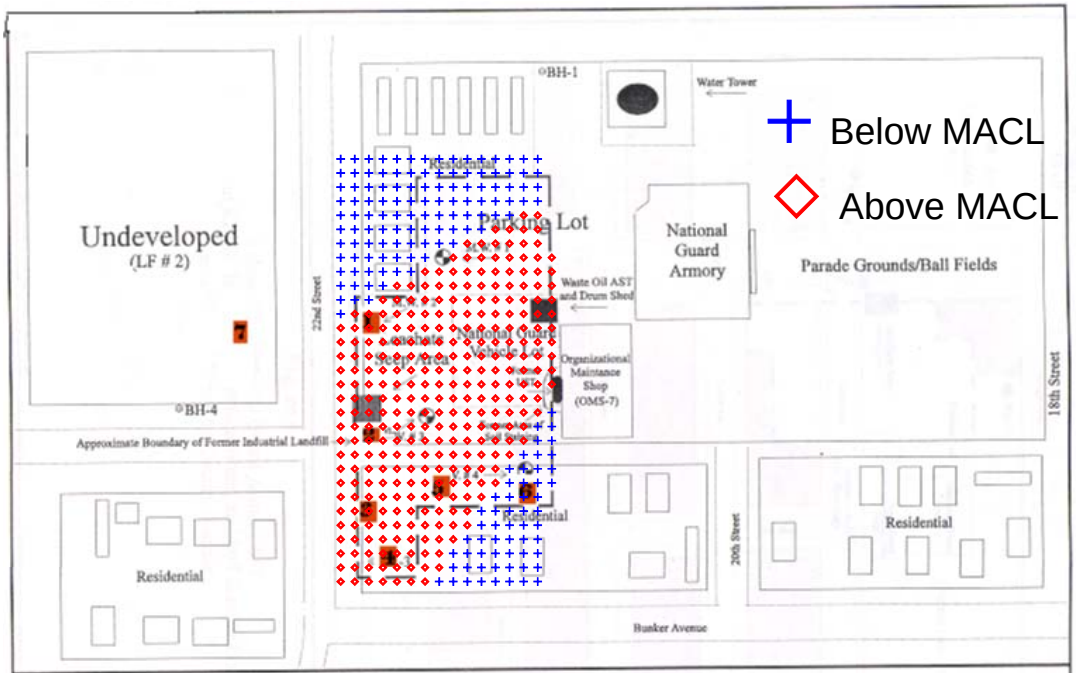


Figure 5.6. Distribution of copper in soil in regard to its MACL

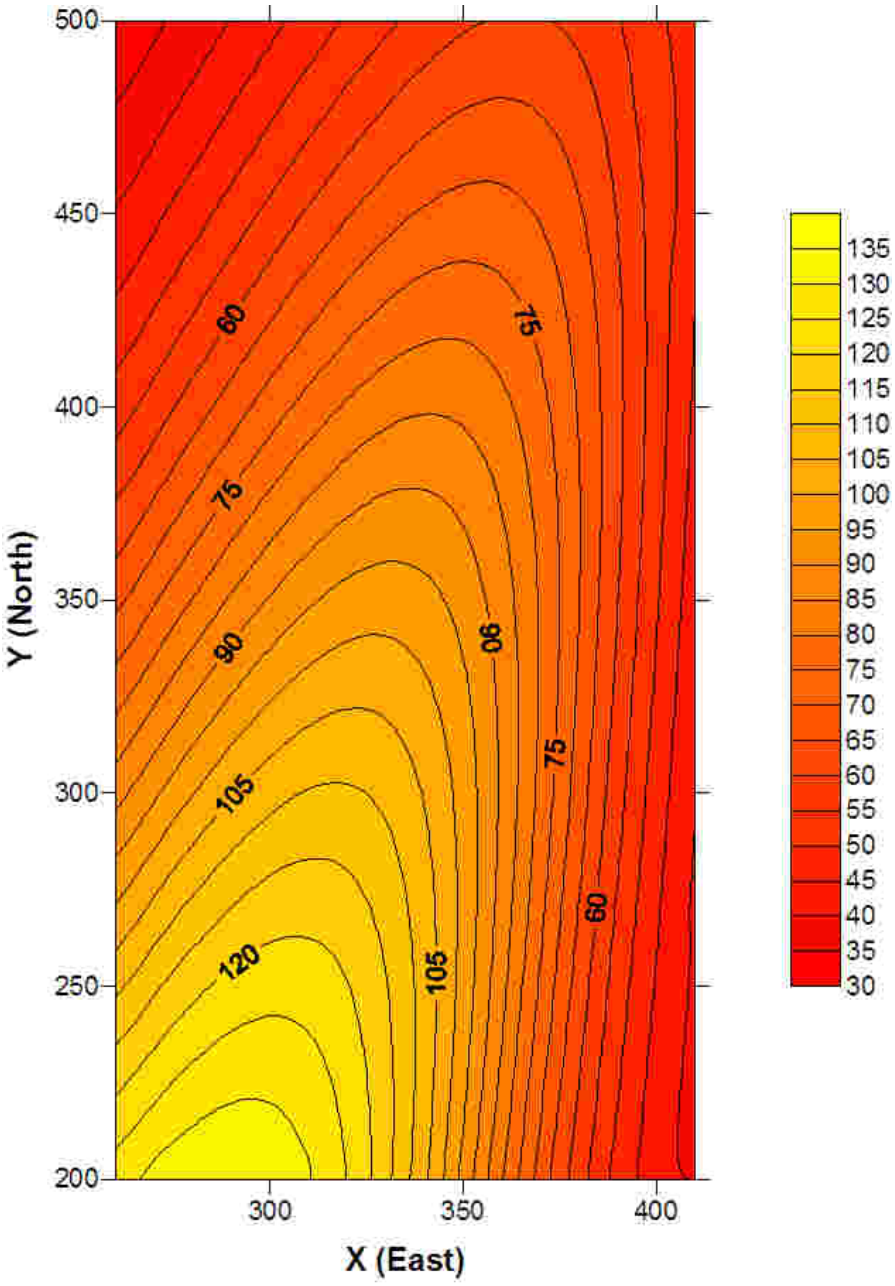


Figure 5.7. Contour map showing the distribution of lead in soil

Figure 5.8. Distribution of lead in soil in regard to its MACL

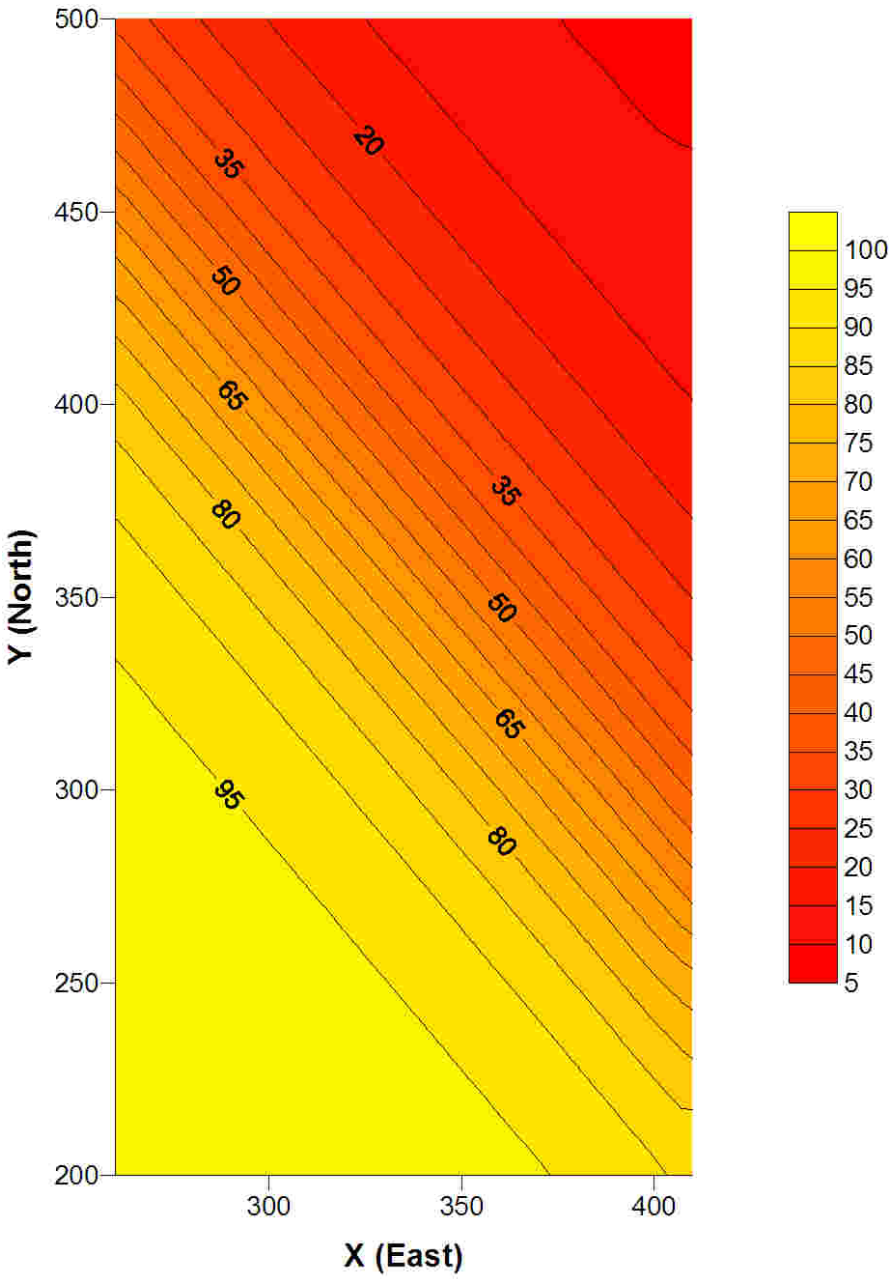


Figure 5.9. Contour map showing the distribution of lead in groundwater at Z = 2 ft

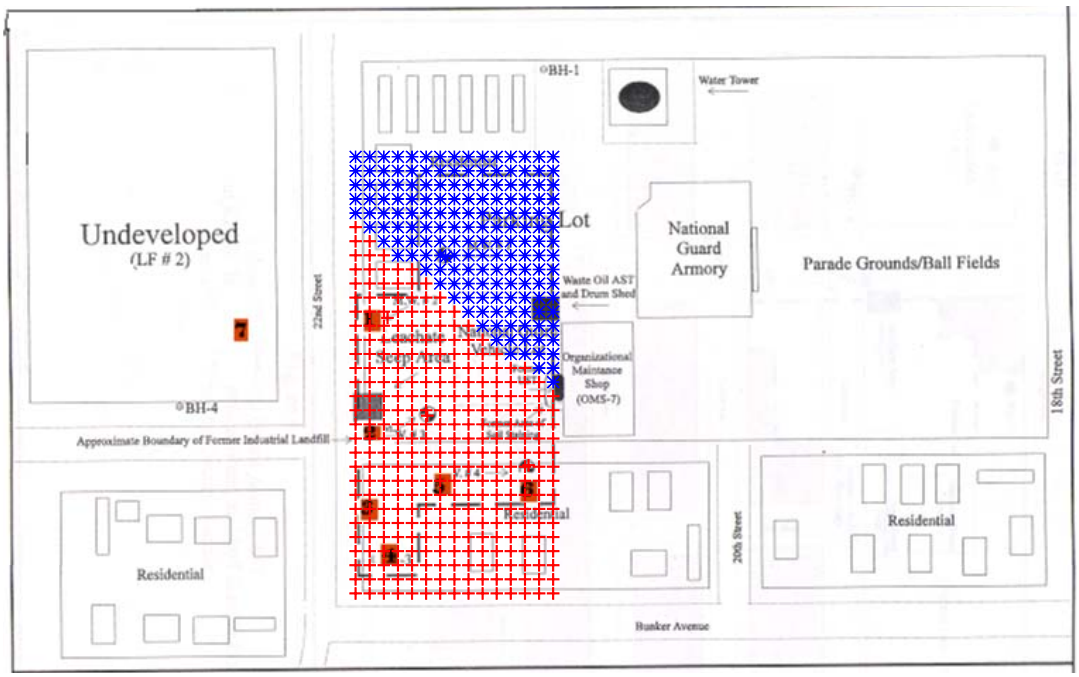


Figure 5.10. Distribution of lead, in regard to its MACL, in groundwater at Z = 2 ft

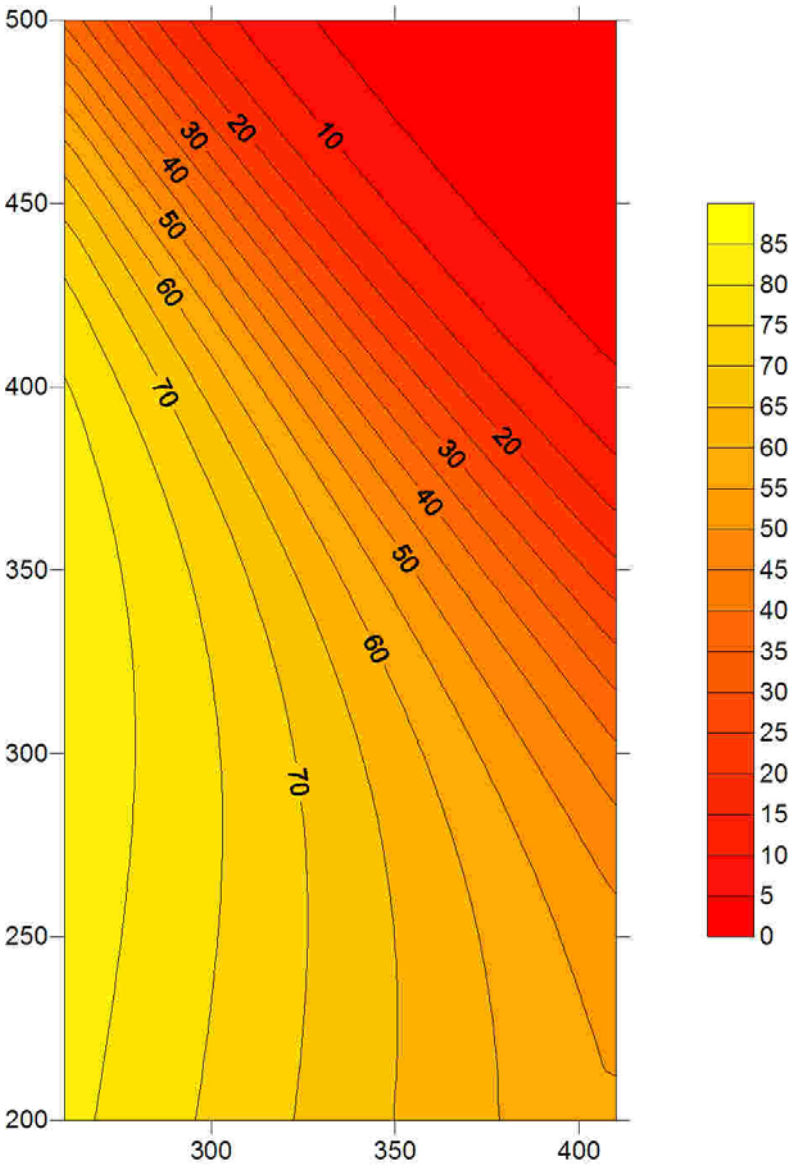


Figure 5.11. Contour map showing the distribution of lead in groundwater at Z = 4 ft

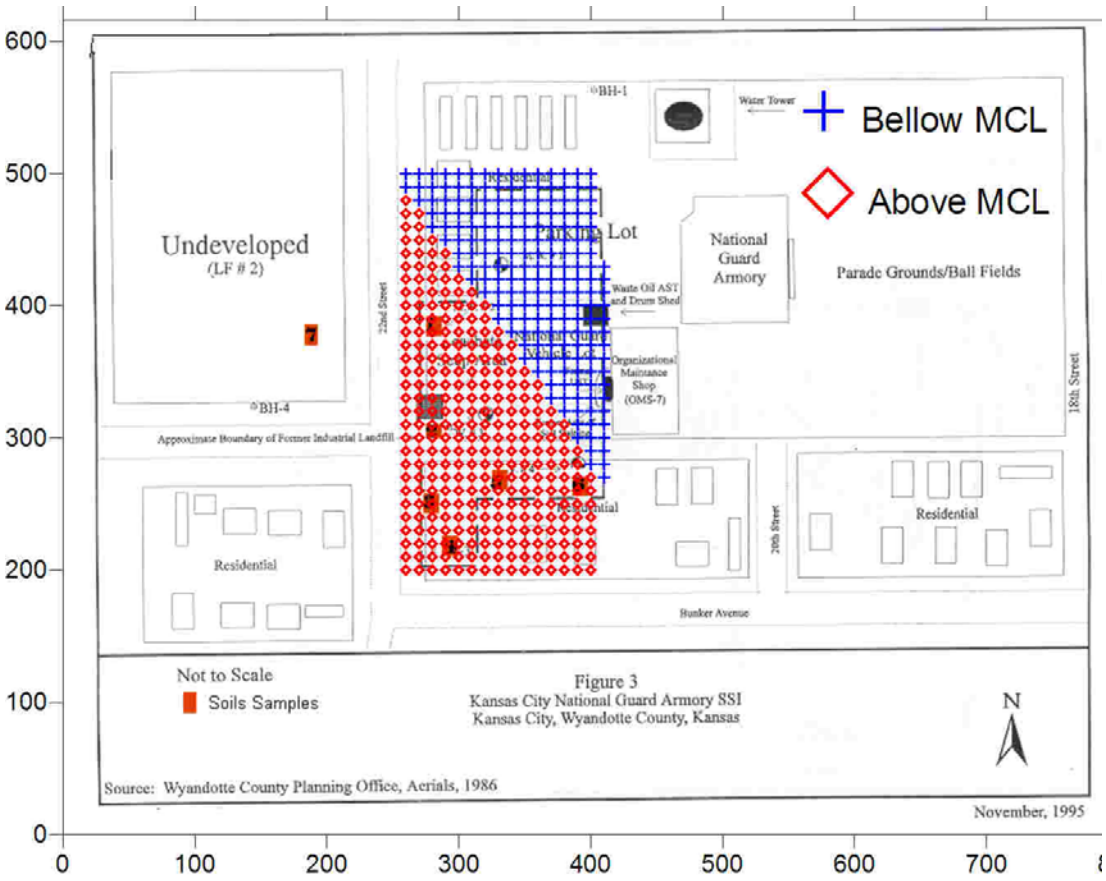


Figure 5.12. Distribution of lead, in regard to its MACL, in groundwater at Z = 4 ft

CHAPTER 6 - **Study Area: the Massachusetts Military Reservations**

6.1 Background

Due to the nature of its operations, the United States military has long been a major contributor to the contamination of groundwater. With 10,444 operational ranges located in the United States and its territories, the problem of contamination within military-owned land has become worrisome. One of these ranges is Camp Edwards, located on the Massachusetts Military Reservation (MMR). Camp Edwards has been used for military mortar and artillery training exercises since the early 1900s. Because of the lack of environmental regulation prior to 1970, the use and disposal of military munitions went unmonitored, leading to the contamination of Cape Cod's primary source of drinking water, the Cape Cod glacial aquifer. However, in 1982, the DOD launched investigation and clean-up efforts of contaminated groundwater and soil at the MMR. The investigators found that 79 different areas on the MMR had potential environmental issues.

The research discussed in this chapter and in Chapters 7 and 8 will expand on the concepts of using ANNs for contamination modeling as explored in Chapter 4. Specifically, the research will utilize explosive-related contaminant data collected in the MMR- DEMO 1, in Massachusetts, to assess the performance of ANNs for predicting concentrations of environmental contaminants, specifically perchlorate.

6.2 Study Area

According to the Air Force Center for Excellence (2004) the MMR is a military training facility located on the upper western portion of Cape Cod (see Figure 6.1), immediately south of

the Cape Cod Canal in Barnstable County, Massachusetts. It includes parts of the towns of Bourne, Mashpee, and Sandwich, and abuts the town of Falmouth (see Figure 6.2). The MMR covers nearly 21,000 acres. Ogden (1999) identified the three main areas that make up the MMR. They are the following:

- The industrial area in the southern part of the reservation where the U.S. Coast Guard, Army National Guard, and Air National Guard facilities are located.
- Aircraft runways, maintenance areas, access roads, housing, and support facilities.
- The northern 14,700-acre area, also known as Camp Edwards, which is used primarily by the Army National Guard. This area contains the 2,200-acre Impact Area, associated military training ranges, and the U.S. Coast Guard.
- The 750-acre Veterans Administration Cemetery, located in the southwestern corner of the reservation.

6.2.1 Location and description of Demolition Area One (Demo 1)

According to MAARNG (2000) Demo 1 is located on Camp Edwards, north of Forestdale Road and south of the Impact Area, near the current H Range at MMR (Figure 6.3). It is a kettle-hole of approximately 7.4 acres, with its base covering approximately 1 acre. The base is approximately 45 feet below the surrounding grade. Groundwater is located approximately 44 ft from the base of the depression (Ogden, 1999). The bottom of the kettle-hole is flat with numerous craters, so it remains wet for a great deal of the year.

The entire MMR is located over the recharge area of the Sagamore Lens (Figure 6.4), and above the sole-source aquifer supplying drinking water for the western part of Cape Cod. The Sagamore Lens is a large, 300-foot-thick layer of groundwater (Pike, 2006). In general, soils in the vicinity of MMR are sandy, permeable and permit rapid groundwater movement. The

Sagamore Lens is recharged or replenished by rainwater that seeps through the sandy soil into the aquifer.

Because Demo 1 is located directly above the Sagamore Lens, any contamination found on the site is able to leach directly into the Cape Cod aquifer. Therefore, the prospect of explosives-related contamination in Demo 1's groundwater is worrisome.

6.2.2 Hydrology of Cape Cod

Cape Cod extends approximately 40 miles into the Atlantic Ocean and has a maximum altitude of 309 feet above sea level. There is one main aquifer system that runs under Cape Cod, which is known as the Cape Cod glacial aquifer (USGS, 2002). This aquifer is an unconfined system and is therefore recharged by infiltration from precipitation (Figure 6.5). In 2005, the USGS reported that because of the geographic make-up of the cape, approximately 45% of the roughly 40 inches of yearly rainfall and snow are absorbed to recharge the aquifer. The other 55% of precipitation is evapotranspired. Less than one percent runs off directly to streams, ponds, lakes, or saltwater bodies.

According to Masterson & Portnoy (2006), the Cape Cod glacial aquifer is bounded by the ocean on three sides, with groundwater discharging into the Nantucket Sound on the south, Buzzard's Bay on the west, and Cape Cod Bay on the north. The Bass River in Yarmouth forms the eastern lateral aquifer boundary. Cape Cod aquifers are comprised of six groundwater lenses: Sagamore, Monomoy, Nauset, Chequesset, Pamet, and Pilgrim (Consortium for Atlantic Regional Assessment, 2005). These lenses are elevated areas of groundwater, which have a shape similar to a convex lens and are separated by ocean inlets or narrows that act as discharge areas. The two largest lenses are the Sagamore and Monomoy (Figure 6.4), which provide water for the majority of Cape Cod's population. Each lens is a hydraulically independent ground-flow

system, which remains steady, due to ground-water recharge and discharge. Approximately 270 million gallons of water per day flow through the six lenses combined (USGS, 2005). Water within these lenses slowly moves toward the coast at a rate of about one to two feet per day from the highest point of the water table, where it discharges into the ocean (Masterson & Portnoy, 2006).

Along with the groundwater lenses, Cape Cod's aquifer system is recharged by numerous kettle-hole ponds. Newman (2001) defined kettle-hole ponds as areas that were formed by blocks of ice stranded by retreating ice sheets and were buried under sand and gravel. Once the ice blocks melted, the sand and gravel covering and surrounding the ice blocks collapsed, leaving a depression in the ground. Over the years, these depressions filled with water and fed directly into the Cape Code glacial aquifer.

AMEC (2001) indicated that the Cape Cod glacial aquifer provides drinking water to the residents of Cape Cod with an average daily water demand of approximately 6.4 million gallons. The daily demand fluctuates depending upon the season. During the summer, the daily demand increases to approximately 10.1 million gallons per day due to tourism. During the remaining months, the daily demand lowers to 5.2 million gallons per day (USGS, 2002). Each groundwater lens of the aquifer provides a portion of the total water supply, and the main contributors are the Sagamore and Monomoy lenses (see Table 6.1).

6.2.3 Geology of Cape Cod

Cape Cod came into existence approximately 25,000 years ago with the advance and retreat of the last continental ice sheet, the Laurentide. Newman (2001) indicated that this ice sheet was divided into three lobes:

- The Buzzards Bay Moraine (BBM)

- The Sandwich Moraine (SM)
- The Mashpee Pitted Plain (MPP)

The BBM and the SM lie along the western and northern edges of western Cape Cod. The MMP lies between the BBM and SM. The majority of the MMR lies within the Mashpee Pitted Plain. MAARNG (2000) describes the geology of the MMR as follows:

> "The MPP, which consists of fine to coarse-grained sands forming a broad outwash plain, lies between the two moraines. Underlying the MPP are fine-grained, glaciolacustrine sediments and basal till at the base of the unconsolidated sediments. The BBM and SM are composed of ablation till, which is unsorted material ranging from clay to boulder size that was deposited at the leading edge of two lobes of the Wisconsonian glacier at its furthest advance. These moraines form hummocky ridges."

These materials cover a layer of bedrock that is 285 to 365 ft below ground surface and considered impermeable.

6.2.4 History of the MMR

Due to the nature of their operations, explosives-related contamination at military installations is common. Over the last several decades, the DOD has tested and fired munitions on more than 24 million acres of operational ranges. In April 2003, the DOD counted 10,444 operational ranges located in the United States and its territories. The DOD defines "operational range" as "an area used to conduct research, develop and test military munitions, or train military personnel" (U.S. Government Accountability Office, 2004).

Although many of the operational ranges are being studied for the presence of perchlorate contamination, only one installation has been studied in-depth. This is Camp Edwards, located within the MMR, in Cape Cod, Massachusetts. The MMR has been used for mortar and artillery

training since 1908. In 1935, the U.S. Army established Camp Edwards used it for military training purposes. Camp Edwards Training Ranges and Impact Area make up approximately 14,000 acres of the MMR (MAARNG, 2000). Over the years, Camp Edwards has been used for activities that include small arms, machine gun, artillery, mortar, ground to ground rocket, air to ground rocket, open burning/open detonation of explosive ordinance and pyrotechnics training. The firing of high explosive artillery rounds continued at MMR until 1989.

Otis Air Force Base was also located on the MMR and used by the Air Force from 1948 until 1973. During this period, the Air Force disposed of pollutants and hazardous materials such as petroleum products, fuels, motor oils, and cleaning solvents in landfills, drywells, sumps, and the sewage treatment plant (MAARNG, 2000). Currently, the MMR is used by the Massachusetts Air National Guard, Otis ANG Base, Massachusetts Army National Guard, Camp Edwards, the U.S. Air Force, and the U.S. Department of Agriculture.

6.2.5 Explosive Related Contamination at Demo 1

In the mid 1970s, the U.S. Army established Demo 1 as an ordnance disposal and demolition training site. The army continued to use Demo 1 until the late 1980s. During this time, various types of ordnance were used and destroyed at this site, including the following (USEPA, 1997):

- Explosive charges of plastic explosives (C-4)
- TNT (2,4,6-trinitrotoluene)
- Detonation cord with a weight limit of 40 pounds
- Bangalore torpedoes
- Claymore mines

On April 10, 1997, the USEPA ordered the Massachusetts Army National Guard to stop all training activities at Camp Edwards that release contaminants into the air, soil or water. This came after the USEPA's Feb. 27 administrative order, which required the National Guard to do the following (USEPA, 1997, p. 11):

(1) undertake a comprehensive study of groundwater related to the training range and impact area

(2) provide information to USEPA about possible contamination in the impact area

(3) develop a proposal for pollution control measures

(4) coordinate with a community-based oversight group

While clean-up efforts continue to this day, the Massachusetts Army National Guard (MANG) began their investigation of Demo 1 in 1997. Ogden (1999) reported that during MANG investigation, they recovered following items: chunks of C4 and other residual munitions, steel I-beams and plates, miscellaneous metal items, ash, burnt-out small arms cartridge casings (5.56mm, 7.62mm, 50 caliber), pyrotechnics, fuses, thermal batteries, rocket bodies, spent 20 mm practice rounds and smoke flares, hand grenades, rifle grenades, 2.36-inch rocket, 90mm dragon, TOW mortar, 81mm mortar and 4.2-inch projectile. Many of these munitions contain Hexahydro-1, 3, 5-trinitro-1, 3, 5-triazine (RDX), TNT and perchlorate (ClO_4).

6.3 Perchlorate

6.3.1 Background

Perchlorate is a naturally occurring manmade anion that has been used in the production of missile and rocket fuels as an explosive propellant for decades. Perchlorate is the salts of perchloric acid ($HClO_4$) (USEPA, 2002). When combined with ammonium, potassium, magnesium or sodium salts, they form compounds that are powerful oxidizers. Perchlorate salts are highly soluble in water and do not adhere well to minerals or organic materials. When a compound such as ammonium perchlorate is released into the environment, the ammonium portions biodegrade, but the perchlorate dissolves. This allows the perchlorate to enter surface and subsurface aqueous systems where it can remain for long periods of time (USEPA, 2007).

Ammonium perchlorate (NH_4ClO_4 or AP) is commonly used as an energetic booster in rocket fuels and potassium perchlorate ($KClO_4$ or KP) is used as a solid oxidant for rocket propulsion (Roote, 2001). Perchlorates are also found in common items such as fireworks, road flares, airbag inflators and other explosives as well as some pharmaceutical products. Perchlorates are used in over 250 types of munitions. It is the use of these munitions, along with the production of perchlorate-containing materials, which has led to the contamination of groundwater throughout the United States.

Perchlorate salts are widely used in solid rocket propellants, matches, signal flares, fireworks, explosives, additives, chemical analytical agents, automobile air bag inflators, and others by the aerospace, defense, and chemical industries (Motzer, 2001).

Perchlorate was first manufactured in commercial quantities in Masebo, Sweden in the 1890s by Stockholm's Superfostfat Fabrisk AB. Commercial production elsewhere in Europe and the United States followed shortly thereafter. In the United States, perchlorate was first

produced in 1908 by Oldbury Electrochemical plant in Niagara Falls. The mass production of perchlorate began in the 1940s during the early part of WWII (Brandhuber & Clark, 2005).

In 1945, the DOD began using the chemical in the production of military munitions items as part of its national defense system. Along with this and the advancements in NASA's aerospace program came an increased need for the production of perchlorate (ITRC, 2008). Before the 1970s, several companies existed that produced ammonium perchlorate. Between 1975 and 1998, this number dropped to only two plants: American Pacific and Kerr-McGee. In April 2003, there were more than 100 perchlorate users located in 40 states (Figure 6.6).

Although perchlorate contamination has existed for decades, widespread perchlorate contamination in the United States was not observed until after the spring of 1997, when an analytical method with a reporting limit of 4 ppb was developed. Since then, methods have been developed that can detect concentrations of 1 ppb and lower. Monitoring for perchlorate contamination has been done throughout the United States over the last several years. Indeed, ITRC (2008) states the following: "USEPA has monitored for perchlorate in public drinking water systems through the Unregulated Contaminant Monitoring Rule (UCMR) program. Under UCMR 1, detections of perchlorate were analyzed using USEPA Method 314.0, at approximately 2,800 large public water systems (see Figure 6.7) and a representative sample of 800 (out of 66,000) small public water systems. As of January 2005, perchlorate had been detected in 153 public water systems and 25 states across the United States (Figure 6.8)."

Geographically, the highest densities of perchlorate detection are in southern California, west central Texas, along the east coast between New Jersey and Long Island, and in Massachusetts (Figure 6.8). The apparent absence of perchlorate occurrence in some regions may merely be because relatively few sources have been sampled. More intensive sampling,

particularly of small systems, may detect perchlorate-contaminated drinking water sources in these regions (Brandhuber & Clark 2005). As noted in Figure 6.8, a great deal of the production of perchlorate compounds has been used in defense activities and the aerospace industry.

The USDOD (2007) reported the following: Past and present activities at DOD industry facilities that may have contributed to environmental releases of perchlorate include, but are not limited to, chemical manufacture of perchlorate materials, manufacture and maintenance of missiles, rockets, and munitions items containing perchlorate, open burning and open detonation of munitions items, the use of perchlorate-containing munitions for weapon system testing and military training (e.g., smoke grenades), ordnance testing and development, rocket motor maintenance and testing, demilitarization of perchlorate-containing munitions items using techniques such as high-pressure water jet washout of rockets and missiles containing solid propellant.

Reported perchlorate contamination in surface and ground water in the U.S. ranged from ppb (μ/L) to ppm (m/L) levels, which affects the drinking water source of 15 million people (Logan, et al., 2001). Approximately 150 perchlorate manufacturers and users have been identified in 44 states, and this number is still increasing. California, Nevada, Massachusetts and Utah are the most affected states, and 18 states, including Arizona, Texas, New York, Maryland, and Arkansas, have reported perchlorate releases (Damian & Pontius, 1999; Logan, 2001). In addition, the Colorado River, which provides drinking water and irrigation water for millions of people, currently has low levels of perchlorate from Lake Mead to Mexico (Logan, 2001).

6.3.2 Dangers

Perchlorate is highly toxic (Table 6.2 lists some properties of perchlorate compounds). When perchlorate enters the human body through drinking water or food grown in soil which is

high in perchlorate, the production of thyroid hormone is disrupted. This can create a condition called *hypothyroidism* (Urbansky, 2000). Hypothyroidism has the potential to affect metabolism and normal growth and development, which could result in brain damage. The impacts of disrupting thyroid hormone synthesis are greatest on pregnant women and their developing fetuses, infants, children, and individuals who have low levels of thyroid hormones. Impaired brain development and lower IQ are associated with children born to mothers who are iodine deficient. Tests on rats and mice at high dosage have caused benign tumor growths (Motzer, 2001). Meanwhile, scientists are concerned about the carcinogenic, developmental, reproductive, and immunotoxic effects of perchlorate as well (Nerenberg et al., 2002). Although hypothyroidism itself is treatable with medication, the secondary effects, such as brain development, are irreversible. The health effects of long-term low dose exposure are being investigated (Herman and Frankenberger, 1999; Logan, 2001).

6.3.3 Hypothyroidism Explained

Perchlorate blocks the uptake of iodide in the body. Iodide is needed to produce the two thyroid hormones triodothyronine (T3) and thyroxine (T4) (Urbansky, 2000). These hormones control physical growth and regulate the body's metabolism. If the thyroid continues to lack iodide, the pituitary gland and the hypothalamus, which regulate the thyroid hormones, will increase their own hormone production to compensate for the lowered levels of T3 and T4. Symptoms of hypothyroidism include the following (Norman Endocrine Surgery Clinic, 2005): fatigue, weakness, weight gain or increased difficulty losing weight, coarse, dry hair, dry, rough pale skin, hair loss, cold intolerance, muscle cramps and frequent muscle aches, constipation, depression, memory loss and abnormal menstrual cycles

6.3.4 Perchlorate and the law

High instances of perchlorate contamination are mainly due to the legal disposal practices of past decades that allowed unregulated waste effluents containing perchlorate into the environment (Motzer, 2001).

Perchlorate is not currently regulated in the Safe Drinking Water Act, but state advisories vary from 1-18 ppb, and Texas Commission on Environmental Quality is adopting an interim action level of 4 ppb. A federal drinking water limitation as low as 1 ppb may be adopted in the near future.

According to data compiled by the California Department of Health Services, perchlorate has been detected in 80 of 912 public water supplies tested in the state, and 292 of 5,205 private drinking water sources sampled contained measurable levels of the pollutant (CalEPA, 2002).

In 1985, the Region 9 office of the USEPA first became aware of the presence of perchlorate in wells. Due to the lack of knowledge regarding the toxicity of this chemical and absence of a valid analytical method, the USEPA focused on other more known threats. In the early 1990s, discovery of perchlorate contamination in water supplies in California continued, prompting the USEPA Superfund Technical Support Center to issue a provisional oral reference dose (RfD) for perchlorate. An oral reference dose is an estimate of the daily exposure of perchlorate to the human population (including sensitive subgroups) that is likely to be without appreciable risk of adverse effects over a lifetime (USEPA, 2002). This dose is given as the perchlorate anion since perchlorate salts readily dissolve in aqueous solutions and it is the anion that is detected in environmental samples. The basis of this RfD was an acute study in which single doses of potassium perchlorate were given to patients suffering from Graves' disease, an autoimmune condition that results in hyperthyroidism. Assuming factors of 70 kilograms of body

weight and 2 liters of water consumption per day, the USEPA converted this RfD value into a drinking water equivalent level (DWEL), stating that a range of 4 to 18 ppb of perchlorate was acceptable in drinking water. In January 1997, based on the upper level of this provisional RfD, the California Department of Health Services adopted an action level of 18 ppb for perchlorate. New York, Arizona, and Texas also initially adopted this 18 ppb action level for perchlorate in drinking water. Based on current data, in April 2007 Massachusetts has established a provisional action level of 1 μg/L for perchlorate in drinking water. Several other states, including Nevada, Maryland, and Texas, have also instituted advisory levels for the oxidant, and other states may follow suit.

In summary, perchlorate (ClO_4) in surface and groundwater has become an ever-increasing water quality concern in the United States during the last decade. Perchlorate was first detected in groundwater in California and Nevada in the early-to mid-1980s (Urbansky & Schock, 1999). However, only after the development of a more sensitive analytical method that lowered the detection limit to 1 ppb (1μg/L) in 1997, was the extent and severity of the problem gradually recognized.

Table 6.1 Average Amount of Water Drawn off of the Groundwater Lenses of the Cape Cod Aquifer

Lens	**Towns**	**Water Volume Withdrawn (mgd)**	**Area (sq mi)**	**Maximum Elevation (ft)**
Sagamore	Bourne, Sandwich Falmouth, Mashpee Barnstable, Yarmouth	25 (summer) 13 (off-season)	165	70
Monomoy	Dennis, Brewster, Harwich, Chatham, Orleans	12 (summer) 5 (off-season)	66	30
Nauset	Eastham, South Wellfleet	2.0/.8 estimate	15	15
Chequesset	Wellfleet, South Truro	1.5/.5	17	8
Pamet	Truro	1.3/.6	11	6
Pilgrim	Provincetown	N/A	4	5

Table 6.2 Properties of Perchlorate Compounds

Properties*	**Ammonium perchlorate (NH_4ClO_4)**	**Potassium perchlorate ($KClO_4$)**	**Sodium perchlorate ($NaClO_4$)**	**Perchloric acid ($HClO_4$)**
CAS#	7790-98-9	7778-74-7	7601-89-0	7601-90-3
Molecular weight	117.49	138.55	122.44	100.47
Color/form	White orthorhombic crystal	Colorless orthorhombic crystal or white crystalline powder	White orthorhombic deliquescent crystal	Colorless oily liquid
Taste/odor	Odorless	Slightly salty	Odorless	Strong odor
Density/specific gravity	1.95 g/cm3	2.53 g/cm3	2.52 g/cm3	1.77 g/cm3
Solubility	200 g/L water at 25ºC	15 g/L water at 25ºC	2096 g/L water at 25ºC	Miscible in cold water
Sorption capacity	Very low	Very low	Very low	Very low
Volatility	Nonvolatile	Nonvolatile	Nonvolatile	Volatile
Octanol/H2O partition coefficient (log Kow)	-5.84	-7.18	-7.18	-4.63
Vapor density (air = 1)	No information	4.8	No information	3.5
pH	5.5–6.5	6.0–8.5	7.0	Highly acidic

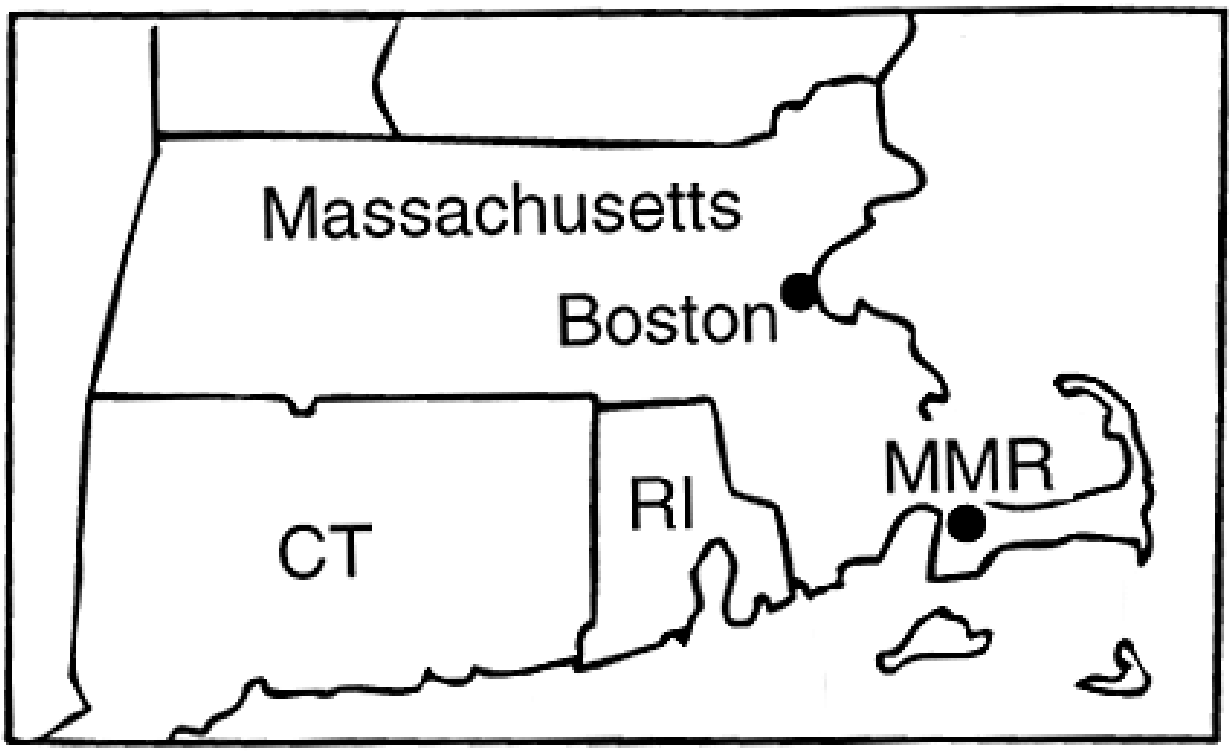


Figure 6.1. Map of the Massachusetts Military Reservation

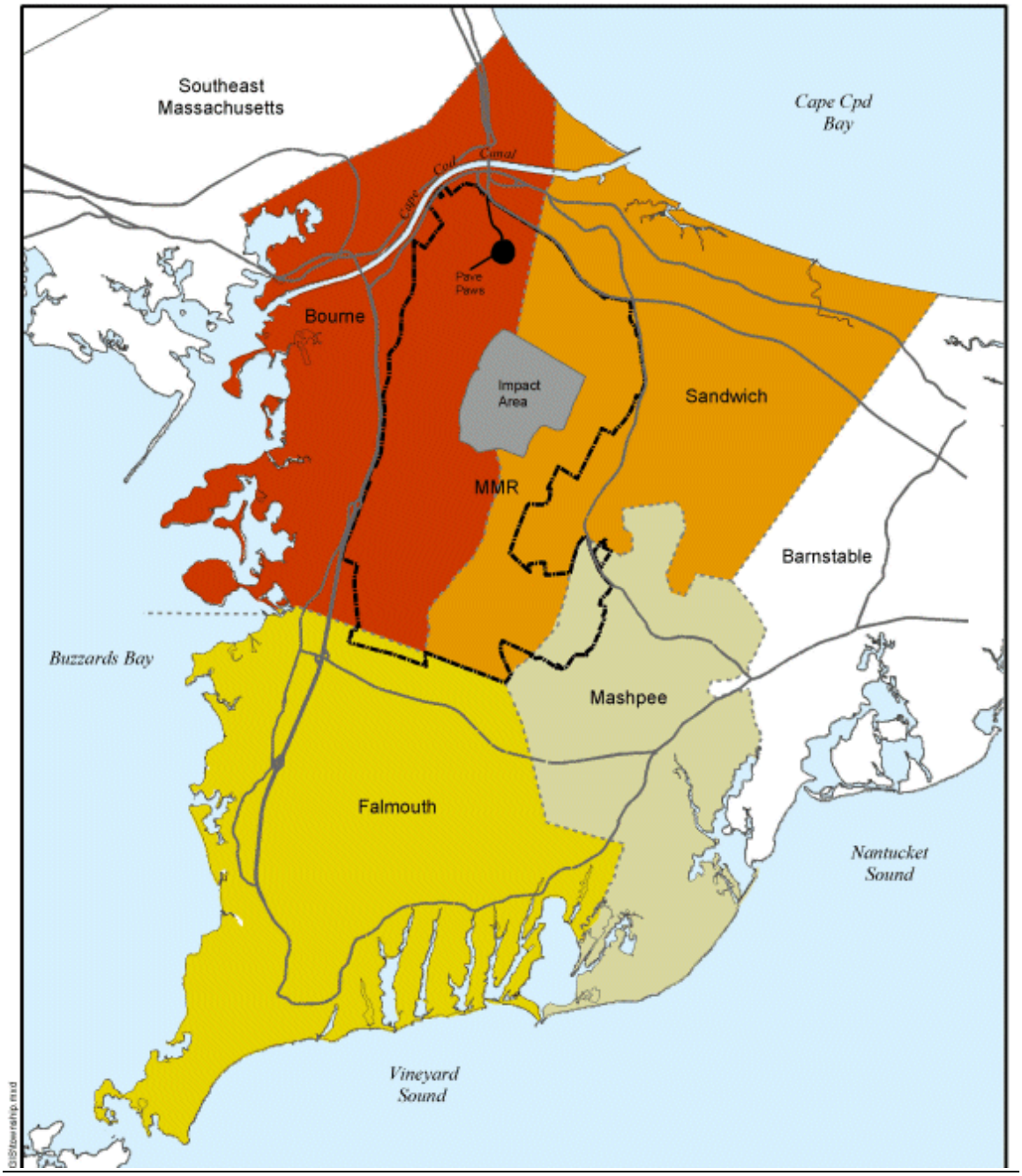


Figure 6.2. Map showing the surroundings of the MMR

Figure 6.3. Photograph of demolition area one at Camp Edwards

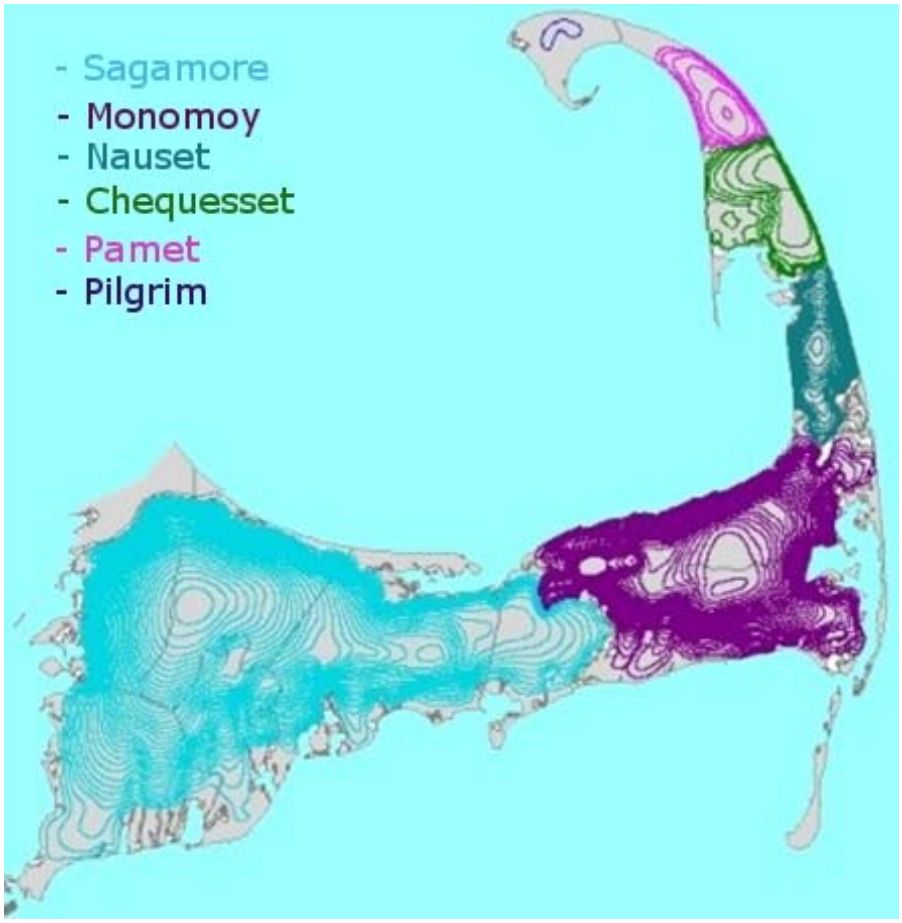
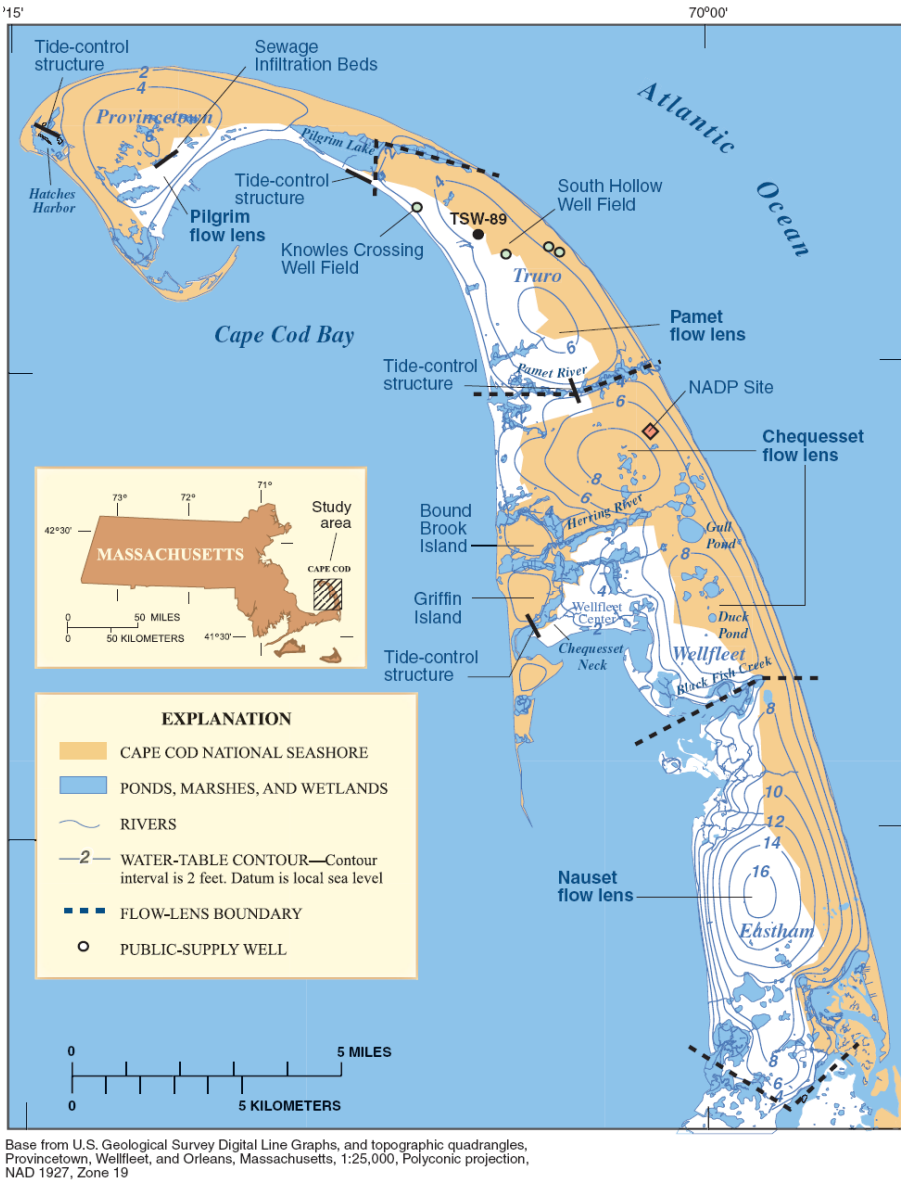


Figure 6.4. Illustration of Cape Cod's lenses



Base from U.S. Geological Survey Digital Line Graphs, and topographic quadrangles, Provincetown, Wellfleet, and Orleans, Massachusetts, 1:25,000, Polyconic projection, NAD 1927, Zone 19

Figure 6.5. The Cape Cod glacial aquifer

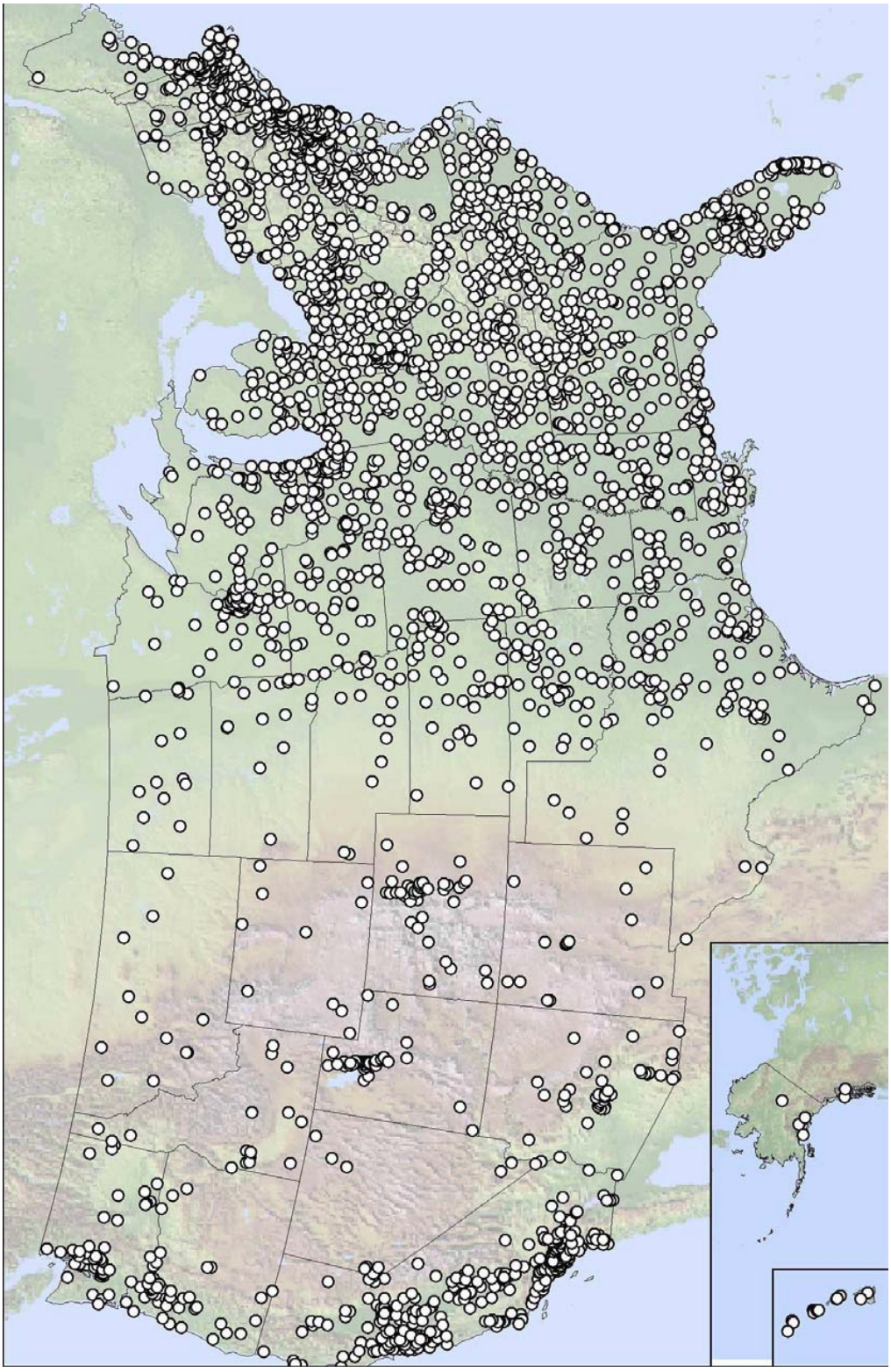


Figure 6.6. Perchlorate users in the U.S.

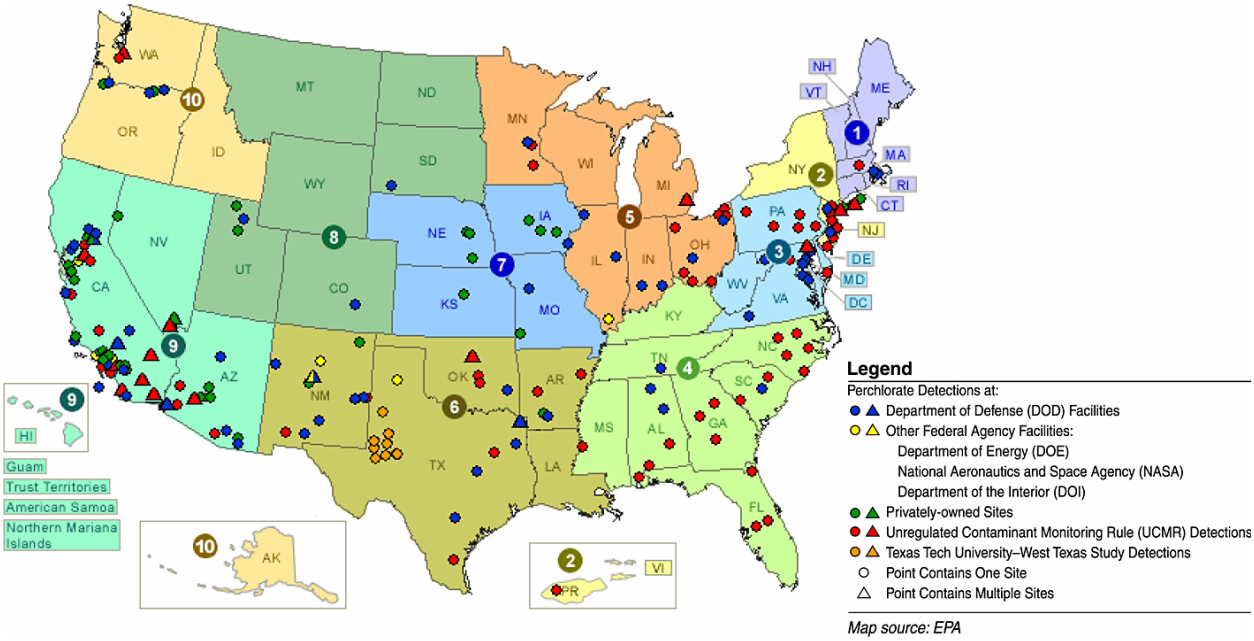


Figure 6.7. Perchlorate in public water systems

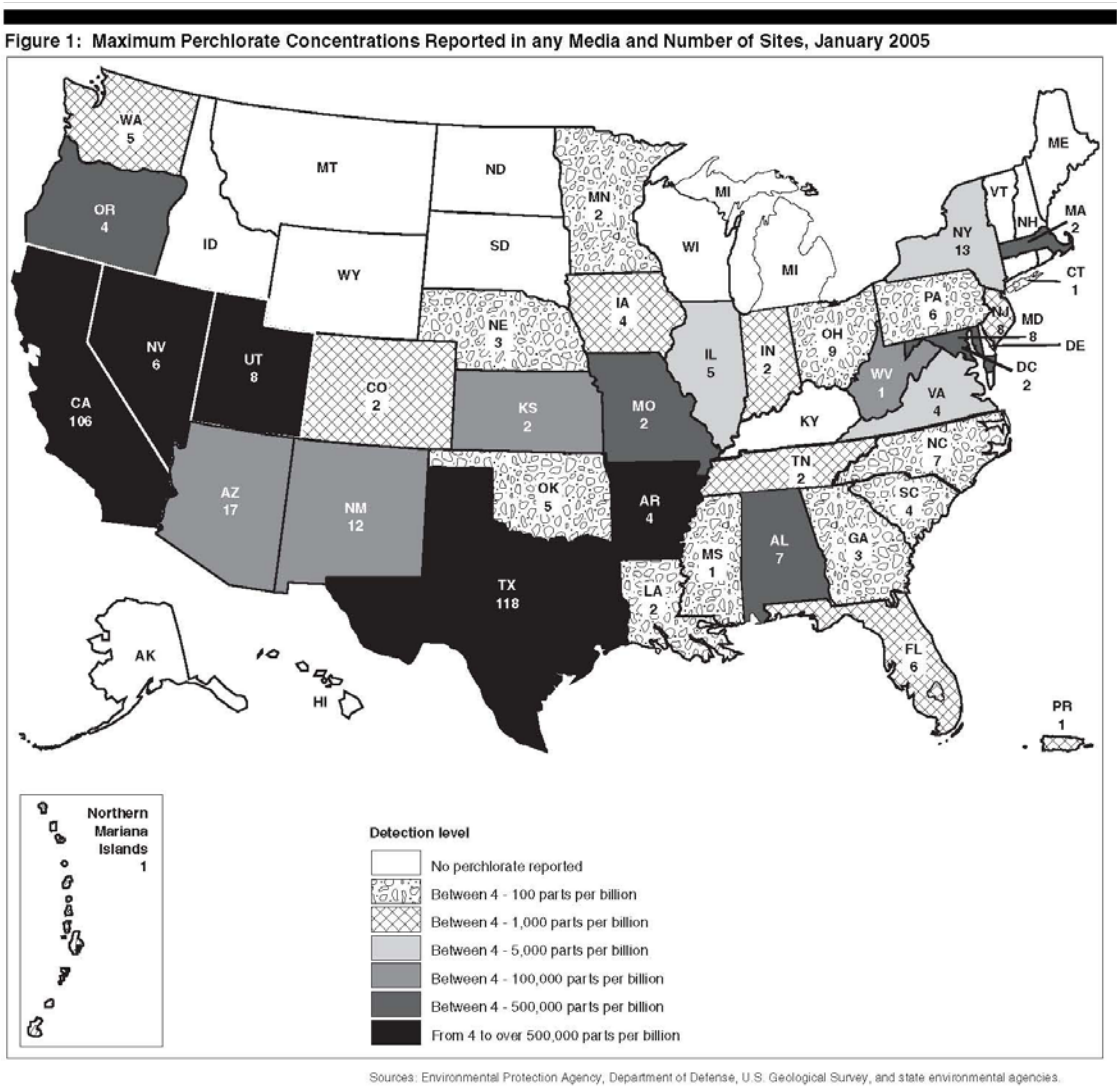


Figure 6.8. National perchlorate detections as of September 2004

CHAPTER 7 - **Three-dimensional time-related profiling of explosive-related contaminants (perchlorate) at the Massachusetts Military Reservation site**

7.1 Introduction

Testing and training ranges are essential to maintaining the readiness of the armed forces of the United States. Recently, concerns have arisen over potential environmental contamination from residues of energetic materials at impact ranges. Jenkins, et al. (2001) concluded that the current state of knowledge concerning the nature, extent, and fate of contamination is inadequate to ensure sound management of ranges as sustainable resources. The potential for environmental impacts, including contamination of drinking water supplies, mandates that the DOD demonstrate responsible management of these facilities in order to continue testing and training activities.

The application of ANNs in environmental site characterization has proven to be an effective modeling method for the prediction of migration paths of environmental contaminants (Mryyan & Najjar, 2005, Dowla & Rogers, 1995; Rizzo, et al., 1996). However, the uses of ANN modeling for the migration of explosives-related contaminants (in particular perchlorate) in water and soil have not been reported in the literature. For this reason, Chapters 7 & 8 will explore the potential use of neural network modeling for predicting the amount and distribution of perchlorate at military installations, specifically the Massachusetts Military Reservation (MMR), site described in Chapter 6.

7.2 Groundwater contamination

Groundwater contamination is not a new phenomenon. Indeed, naturally occurring inorganic contaminants, such as salts, metals, and radioactive materials have contaminated groundwater for centuries. However, the development of man-made chemicals and technological developments have introduced new ways for water to become contaminated. The leaching of chemical contaminants into the groundwater supply of the U.S. is a source of great concern to the nation. According to Groundwater Foundation (2006), 50% of the nation's population depends on groundwater for daily drinking water. When humans ingest contaminated drinking water, it can lead to such diseases as hepatitis, hypothyroidism, and cancer (Johnson & Rogers, 1995). Therefore, it is essential to take steps that ensure that America's water supply is as pollutant-free as possible. In an effort to minimize water contamination and to clean up existing contamination, Congress approved the Clean Water Act.

Sources of groundwater contamination are often difficult to describe. In industrial areas, a lack of historical information is fairly common (Gailey, et al. 1991). Land use may have changed many times: pumping and/or injection wells may have created complex and transient flow gradients, and local industries may be unable or reluctant to describe what they have contributed to contamination. Often, if the sources are unknown or hard to quantify, the best characterization or current snapshot of the plume can be achieved through soil sediment and groundwater sampling (Dougherty & Marryott, 1991).

Known sources are often classified as point or distributed sources. Hunt, et al. (1988) classified point sources as localized sources such as landfills, underground storage tanks, waste disposal wells, leaking pipelines (these may be line rather than point sources), industrial spills, and holding ponds or lagoons. Distributed sources are spread over large parts of the

contaminated area. Examples are agricultural and urban use of pesticides, fertilizers and manure, transportation chemicals (road salt, tar), sewage and septic systems, nuclear fallout, and urban storm-water runoff. Poorly constructed or abandoned wells may increase contamination from surface sources.

7.3 Monitoring Groundwater

Federal, state, and local governmental organizations collect water samples for laboratory analysis. Research organizations, community action groups, and private citizens may perform their own monitoring of groundwater. The EPA publishes standardized water quality analysis procedures and grants certification to laboratories that apply them appropriately. To collect a representative sample of aquifer water quality, monitor wells are installed. Where contamination is suspected, baseline quality is assessed by placing monitor wells upgradient and downgradient from suspected contamination (Ahlfeld, 1990). Extensive drilling, testing, and geophysical logging are usually necessary to ensure proper monitor well placement, even if the source location is known. Many factors may contribute to difficulty in intercepting contaminated groundwater, including unexpected barriers or enhancements to flow (e.g., fractures, fault zones), changes in predicted flow direction by undocumented well pumping, and denser-than-water contaminants moving downward (Bredehoeft & Young, 1983). The high costs of installing and operating monitor wells in addition to the concern for increasing contamination of aquifers at multiple levels encourages minimal monitoring.

7.4 Background of Study Area

MMR is a site that has been associated with military operations for several decades. This site was utilized for many purposes, one of which led to the presence of perchlorate in soil, sediments, surface water, and groundwater. The data contained in this chapter were obtained from this site.

Although the MMR is a large area of land, the area of concern is only 7.4 acres. This site is referred to as Demo 1 and is located on Camp Edwards, approximately 2 miles northeast of the Otis Rotary in Bourne, Massachusetts. Demo 1 is located in a natural topographic depression that covers approximately one acre at its base and lies 45 ft below the surrounding grade.

Demo 1 was established in its current location between 1986 and 1989 as a heavy demolition site. Its primary use was for training engineer and explosive ordnance disposal units and for the destruction of various types of unexploded ordnance. Such ordnance (including perchlorate) included explosive charges of C-4, 2,4,6-trinitrotoluene (TNT), and det-cord with a weight limit of 40 lbs, Bangalore torpedoes, and claymore mines (AMEC, 2001). According to Pennington, Brannon & Mirecki (2002):

> In January, 2000, the U.S. EPA Region I issued an Administrative Order for Response Action in the matter of Training Range and Impact Area, Massachusetts Military Reservation to the Massachusetts National Guard. The purpose of the Order was to require the respondents to undertake Rapid Response Actions and Feasibility Studies, Design and Remedial Actions to abate the threat to public health presented by the contamination from past and present activities and sources at and emanating from the (MMR) Training Range and Impact Area.

In 2004, the MMR implemented a program to correct the negative impacts of perchlorate contamination at the Demo 1 site. Two extraction, treatment, and recharge (ETR) systems were installed within the area of contamination. These wells pump the contaminated groundwater out, remove explosives and perchlorate contamination from the water, and then reinject the treated water back into the wells. This is done at a rate of 110 gallons per minute (USACE, 2006)**.** This process continues to date.

7.5 Pre-Existing Data

At the start of this research, perchlorate laboratory analysis data were available for the years 2000–2005. The ANN model was developed, tested, and calibrated using the available data from the MMR facility. A total of 459 samples were collected and analyzed at the Demo 1.

In accordance with the Administrative Order for Response Action, the U.S. Army Environmental Center (EC) collected water quality samples at 41 monitoring wells in 33 locations (Figure 7.1). Site investigations at and down gradient (west) of Demo 1 included the collection of approximately 650 soil samples and the installation of monitoring wells in 33 locations (USACE, 2006). Data from all monitoring wells were utilized in this study. Water quality samples were collected by the EC between March 2000 and December 2005. These samples were analyzed for over 200 compounds to include explosives, volatile organic compounds, semi-volatile organic compounds, pesticides and herbicides, polychlorinated biphenyls (PCBs), polychlorinated naphthalenes, dioxins, and heavy metals.

7.6 Model Development

Back-propagation networks were developed using the TR-SEQ1, a three-layered ANN training software package developed by Najjar (2001). The purpose of the ANN model is to predict the concentrations of perchlorate at the MMR from appropriate input parameters.

7.6.1 Determination of Appropriate Model Inputs

This section will discuss the process used to determine the selection of input parameters for the ANN models. The determination of the appropriate model inputs is a process that requires a great deal of consideration. Huang (2006) states, "Whereas in physically-based models the necessary input parameters are specified by the equations that describe the physical, chemical, or biological process being simulated, there is no such specification in ANN models." Because of this, it is imperative that there be an adequate amount of relevant input data to train the ANN model.

Based on the available MMR data, back-propagation neural network was chosen as the most-appropriate ANN for developing the site profiling prediction model. The back-propagation approach used by Mryyan & Najjar (2007 & 2005) and Dowlas & Rogers (1995) has proven successful in past environmental site profiling because of its ability to accurately predict the amount and distribution of environmental contaminants at a given site.

As mentioned in Chapter 3, the performance of a network is dependent on the following factors:

- The number of hidden layers and nodes. For ANN mapping, one hidden layer was used between the input layer and output layer. The number of hidden nodes in the hidden layer was determined by adaptive training and online monitoring of accuracy measures on the testing datasets. This was done by varying the number of initial hidden nodes, in the hidden

layer, until the network was able to best learn the patterns involved in the testing datasets. Fully connected internal structure was used (every node in one layer connects to all the nodes in the next layer). It is worth mentioning that the input domain of the network was determined based on cause-effect principle along with a trial-and-error approach, because there are no references in the literature that could provide guidance for selection of the inputs (Ali & Najjar, 1998).

- The number and type of nodes in the input layer. Given the MMR data, the following nine potential input parameters were considered:
  1. X-Coordinates (East)
  2. Y-Coordinates (North)
  3. Sample depth from sea level (Z)
  4. Groundwater elevation (G)
  5. Cumulative number of days since 1/1/2000 (T)
  6. Cumulative amount of rain since 1/1/2000 (R)
  7. Amount of water injected back to groundwater at the Pew Road well (INJ1) since April 2004
  8. Amount of water injected back to groundwater at the Perking Road well 2 (INJ2) since April 2004
  9. Amount of water injected back to groundwater at the Perking Road well 3 (INJ3) since April 2004
- In order to determine the domain of input parameters for the optimal ANN model, the effect of input parameters on the output and the performance evaluation criteria (statistical accuracy measures and graphical evaluation) are utilized to identify and distinguish the most important

parameters that contribute to the best perchlorate prediction. Various training and testing trials, eliminating different input parameters, were conducted in order to identify the most important input parameters. In the initial step, all nine potential input parameters (X, Y, G, Z, T, R, INJ1, INJ2 and INJ3) were used to develop the desired perchlorate prediction model. In the second step, one parameter was eliminated to arrive at eight inputs (X, Y, Z, T, R, INJ1, INJ2 and INJ3). Groundwater (G) parameter was not included. In the final step, seven input parameters were used (X, Y, Z, T, INJ1, INJ2 and INJ3). Rain (R) and groundwater (G) parameters were not included (See Table 7.1).

Based on previous environmental site profiling knowledge (Mryyan & Najjar, 2005; Mryyan & Najjar, 2006 and Mryyan & Najjar, 2007), input(s) with minimal impact on the accuracy of ANN model prediction were eliminated from the input domain in trial cases. If the statistical accuracy measures were improved by eliminating one input parameter, the effect of eliminating two input parameters at the same time was investigated further. This procedure was repeated until the statistical accuracy measures did not improve by eliminating more input parameters. The purpose of this procedure was to obtain the optimal ANN model—the highest statistical accuracy with the least number of input parameters. In this study, based on the two stages approach, it was determined that all of the nine potential input parameters (X, Y, G, Z, T, R, INJ1, INJ2 and INJ3) were necessary to accurately predict the amount and the distribution of perchlorate at Demo1. Tables 7.2 to 7.4 show the statistical accuracy for each of the three trial cases.

7.6.2 Model Training and Testing

In order to obtain the best perchlorate prediction model, the database used for training should represent all possible features and sub-features that the network is required to learn. This

study adopted a two-stage training methodology for every ANN model trial structure. In the first stage, the entire database was divided into training, testing and validation sub-databases at the ratio of about 50%: 25%: 25%. The training sub-database contained all the datasets with the maximum or minimum value of each input and output parameter. Using the training and testing datasets for training and testing respectively, the least-error-structure is selected based on the following statistical accuracy measures (Tables 7.2, 7.3 and 7.4): Averaged Squared Error (ASE), Mean Absolute Relative Error (MARE) and Coefficient of Determination (R^2) on the testing datasets. Then the network was trained, tested, and validated at its least-error structure with the corresponding sub-databases.

Felker (2005) reported that if the statistical accuracy measures at the least-error structure for training, testing, and validation data sets were found to be comparable, then the second stage of training is not necessary. Otherwise, this indicates that the developed net does not recognize some of the features in the database. In this case, the second stage training was carried out. In the second stage, all datasets in the database were used to re-train the least-error structure identified in stage one. A total of three cases were investigated in order to obtain the optimal ANN model for perchlorate prediction. The optimal ANN structures for all cases are listed in Table 7.1.

7.6.3 Model Selection

The best model is considered as the one with the highest statistical accuracy. Statistical (MARE, R^2 and ASE) accuracy measures were adopted to select the optimal network models. Tables 7.2 to 7.4 show the accuracy measure obtained for each trial case. The statistical accuracy measures, ASE and R^2, were improved by training on all data at its optimal ANN structure (obtained from Stage I). It was found that the seven-inputs model and the eight-inputs model are comparable; however, the nine-inputs model outperformed the seven- and the eight-inputs

models in the value of ASE for the stage II trials. It was then decided to select the nine-inputs model as the optimal ANN model for this study.

In summary, for the MMR perchlorate prediction model, it was determined that the optimal network model contained nine-inputs parameters, nine hidden nodes, and one output parameter (9-9-1). The corresponding accuracy measures for this network based on Stage II training is listed in Table 7.2.

7.7 Data Banks

Once the optimal network model was determined to represent the MMR site, the network was used to predict the values of perchlorate contamination at any desired location. The only parameter required in order to provide needed predictions is the input data vector (X, Y, G, Z, T, R, INJ1, INJ2, and INJ3).

The Demo 1 site was divided in the x, y, and z directions using $\Delta x = \Delta y = 25$ ft. The grid system generated in the (x, y) plane produced 4,527 grid points (Figure 7.2). These coordinates were used for z = -50 ft, -25 ft, 0 ft, 25 ft and 50 ft, generating a total of 22,637 grid points. For each of the generated grid points, the perchlorate concentration values were predicted using the corresponding (x, y, z) coordinates via the optimized 9-9-1 networks (Table 7.2). Predictions were made using data representing one specific date in time. This means that values for the parameters G, Z, R, INJ 1, INJ 2, and INJ3 remained constant for a specific time, although z varied for any given x and y coordinate.

7.8 Excel Application

Once the ANN model was developed, the optimal network (9-9-1) was used to create an Excel and Visual Basic software program called MMR perchlorate level determination (MMR-

PLD). To make the program user-friendly, a graphical user interface (GUI) was developed (See Figure 7.3). To find the perchlorate level at any certain x and y coordinate throughout the Demo 1 site, all the user need to do is enter the desired date (t) and desired depth (z). The program will then return the perchlorate level for that x, y and z coordinate.

7.9 Contour Maps

It is often difficult to analyze large amounts of contaminant data in relation to a specific area such as Demo 1. By taking known data and creating a contour map, a visual graph of the study area can be created. This allows for easier interpretation and comparison of given data. For this reason, contour maps of the Demo 1 site were created using version 8 of Surfer® software (2007) to assist in the visualization of perchlorate contamination from years 2000 through 2005. This was done using the results obtained from the 9-9-1 ANN model, as described previously.

The contour maps had (x, y) as a variable and (z) as a constant. Figures 7.4 through 7.8 indicate perchlorate concentrations at z = -50 ft, -25 ft, 0 ft, 25 ft and 50 ft for years 2000 to 2005. By creating contour maps for each year and each depth, trends in the concentration of perchlorate contamination over time and at different depths can be tracked easily. Such images also allow for easy identification of contaminated areas.

In all contour maps (Figures 7.4 through 7.8), a red color reflects a high value of perchlorate and a light color represents a low perchlorate value. Consequently, a red color means a high value of perchlorate concentrations above the regulatory limit of 1 part per billion, and a white color indicates no perchlorate concentration. Color ramp is used between the red and white colors to map intermediate perchlorate concentration values according to the scale shown on the figures.

As stated before, 3-D contour maps were generated at depths of -50 ft, -25 ft, 0 ft, 25 ft and 50 ft for the years 2000 through 2005. For the purposes of comparison, these depths will be divided into three categories: below sea level (-50 ft and -25 ft), sea level (0 ft) and above sea level (25 ft and 50 ft).

Before the implementation of the extraction, treatment and recharge (ETR) methods in April 2004, high levels of perchlorate were noted in groundwater samples as demonstrated in the contour maps for years 2000 through 2003. This finding is consistent with known natural migration patterns of perchlorate in groundwater. The highest levels of contamination for these years were noted below sea level, as shown in Figures 7.4a through 7.4c and 7.5a through 7.5c. At sea level and above, levels of contamination decreased but were still present, as shown in Figures 7.6a through 7.16c, 7.7a through 7.7c, and 7.8a through 7.8c.

With the implementation of the ETR methods, the pattern of perchlorate contamination changes. Groundwater that has been drawn off and treated by the ETR system no longer contains perchlorate. When this treated water is reinjected into the wells, it is reinjected below sea level. 2005, at a level 50 ft below sea level (Figures 7.4e & 7.4f), there is no evidence of perchlorate contamination. At -25 ft, levels of perchlorate contamination begin to increase (Figures 7.5e & 7.5f), with the greatest amount noted at 25 ft above sea level (Figures 7.7e & 7.7f).

7.10 Concluding Remarks

The ANN-modeling used in this research demonstrates the neural network's ability to accurately predict perchlorate contamination using multiple variables or inputs. To determine the most appropriate input parameters for this model, three different cases were investigated using nine potential input parameters.

In the initial case, all nine potential input parameters (X, Y, G, Z, T, R, INJ1, INJ2 and INJ3) were used to develop the desired perchlorate prediction model. This model produced an ASE value of 0.0025 and R^2 value of 0.607.

In the second case, one parameter (G) was eliminated to arrive at eight inputs (X, Y, Z, T, R, INJ1, INJ2 and INJ3). This model produced an ASE value of 0.0030 and R^2 value of 0.57. In the final case, two parameters (R and G) were eliminated to arrive at seven input parameters (X, Y, Z, T, INJ1, INJ2 and INJ3). This model produced an ASE value of 0.0032 and R^2 value of 0.503.

When comparing the three cases, the following was observed:

- The eight-inputs case produced an ASE value 20% greater than the nine-inputs case, while the R^2 value decreased by 6.5%.
- The seven-inputs case produced an ASE value 28% greater than the nine-inputs case, while the R^2 decreased by 17%.
- The nine-inputs models outperformed the eight-inputs models in the value of ASE for the stage II trials therefore identified the nine-input model as the optimal ANN model for this study.

Using the data generated from the 9-9-1 ANN model and the MMR-PLD Excel program, contour maps were generated and compared for levels above, at, and below sea level for the years 2000 through 2005. Contour maps generated using data prior to the implementation of the ETR (actual) system in 2004 indicate higher levels of perchlorate below sea level for the years 2000 through 2003. Contour maps generated using data after the implementation of the ETR system indicate decreased levels of perchlorate below sea level. Perchlorate levels at and above

sea level increased during this period; however, this was to be expected due to the reinjection process.

When comparing the trends observed using the ANN-generated data and the actual trends identified in the MMR 2006 System Performance Monitoring Report, both agree that perchlorate levels are decreasing due to the use of the ETR systems. This proves that the ETR systems were both effective and necessary for the removal of perchlorate contamination at the Demo 1 site, as demonstrated in the contour maps.

This study has proven that it is possible to use back-propagation ANN-modeling to accurately predict groundwater and soil contamination using limited known data. By utilizing ANN methodology, more in-depth studies of a similar site type can be performed.

Table 7.1 Optimal Structures of All Trial Cases

ANN Structure	9-inputs	8-inputs	7-inputs
Initial number of HN	1	2	2
Maximum # of iterations at optimal structure	1000	1000	4000
# of HN at optimal structure	9	8	6

- 9-inputs (all potential parameters included)
- 8-inputs (rain parameters not included)
- 7-inputs (rain and groundwater elevation parameters are not included)

Table 7.2 Statistical Accuracy Measure for the Nine-Inputs Network

9-inputs network			MARE%			R^2			ASE		
Stages	Itr	HN	Training	Testing	Validations	Training	Testing	Validations	Training	Testing	Validations
stage I-A	1000	9	116	136	NA	0.434	0.63	NA	0.0042	0.0018	NA
stage I-B	1000	9	116	NA	124	0.434	NA	0.28	0.0042	NA	0.0046
stage II	1000	9	139	NA	NA	0.607	NA	NA	0.0025	NA	NA

Table 7.3 Statistical Accuracy Measures for the Eight–Inputs network

8-inputs network			MARE%			R^2			ASE		
Stages	Itr	HN	Training	Testing	Validations	Training	Testing	Validations	Training	Testing	Validations
stage I-A	1000	8	140	138	NA	0.46	0.63	NA	0.004	0.002	NA
stage I-B	1000	8	140	NA	141	0.46	NA	0.39	0.004	NA	0.004
stage II	1000	8	118	NA	NA	0.57	NA	NA	0.003	NA	NA

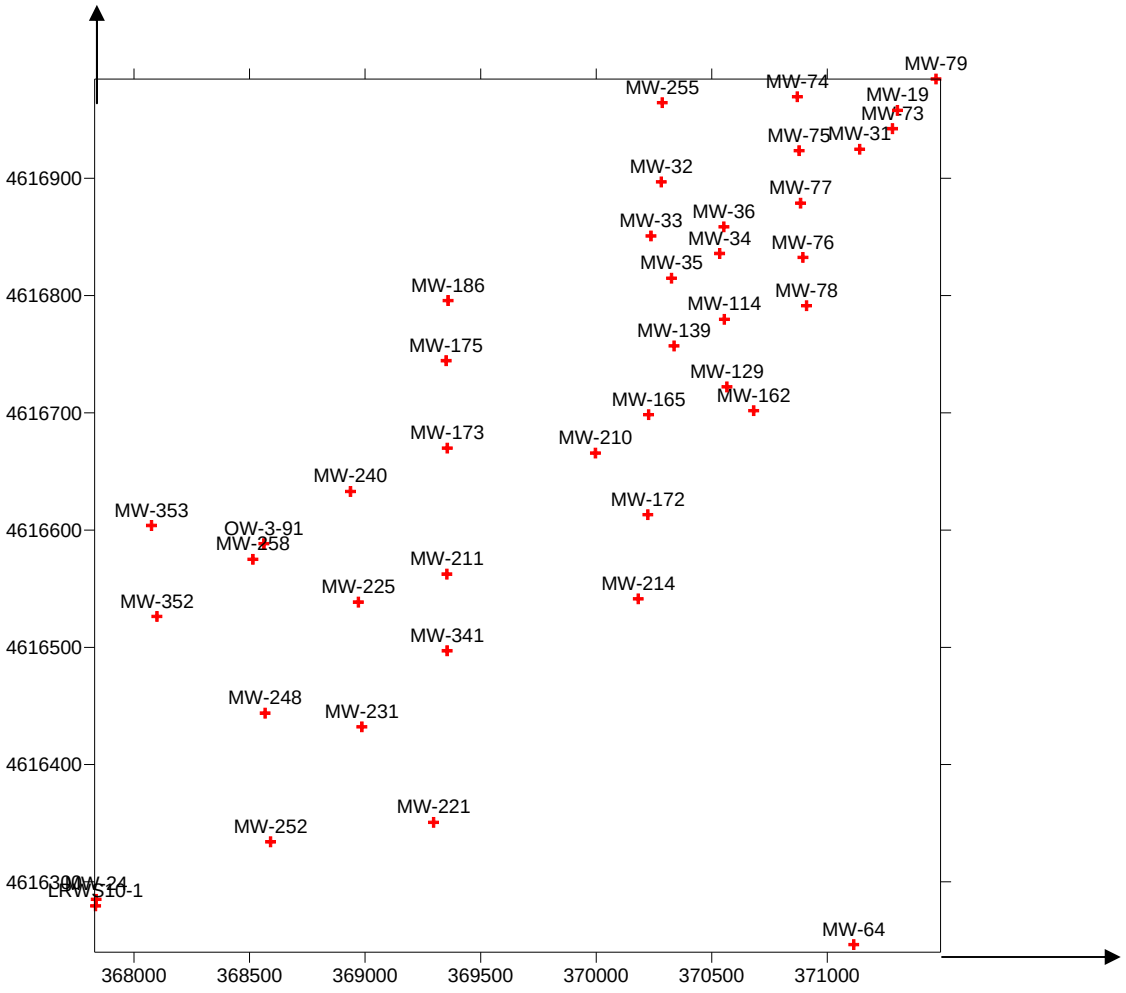
Table 7.4 Statistical Accuracy Measures for the Seven-Inputs Network

7- inputs network			MARE%			R^2			ASE		
Stages	Itr	HN	Training	Testing	Validations	Training	Testing	Validations	Training	Testing	Validations
stage I-A	4000	6	149	142	NA	0.502	0.54	NA	0.003	0.0022	NA
stage I-B	4000	6	149	NA	155	0.502	NA	0.302	0.003	NA	0.00476
stage II	4000	6	123	NA	NA	0.503	NA	NA	0.0032	NA	NA

HN: Optimal Number of Hidden Nodes
Itr: Iterations
MARE: Mean Absolute Relative Error %
R^2: Coefficient of determination
NA: Not applicable

ASE: Averaged Square Error = $\frac{\sqrt{\Sigma(\mathrm{Pr}\,edicted - Actual)^2}}{\#\,of\ data\ sets}$

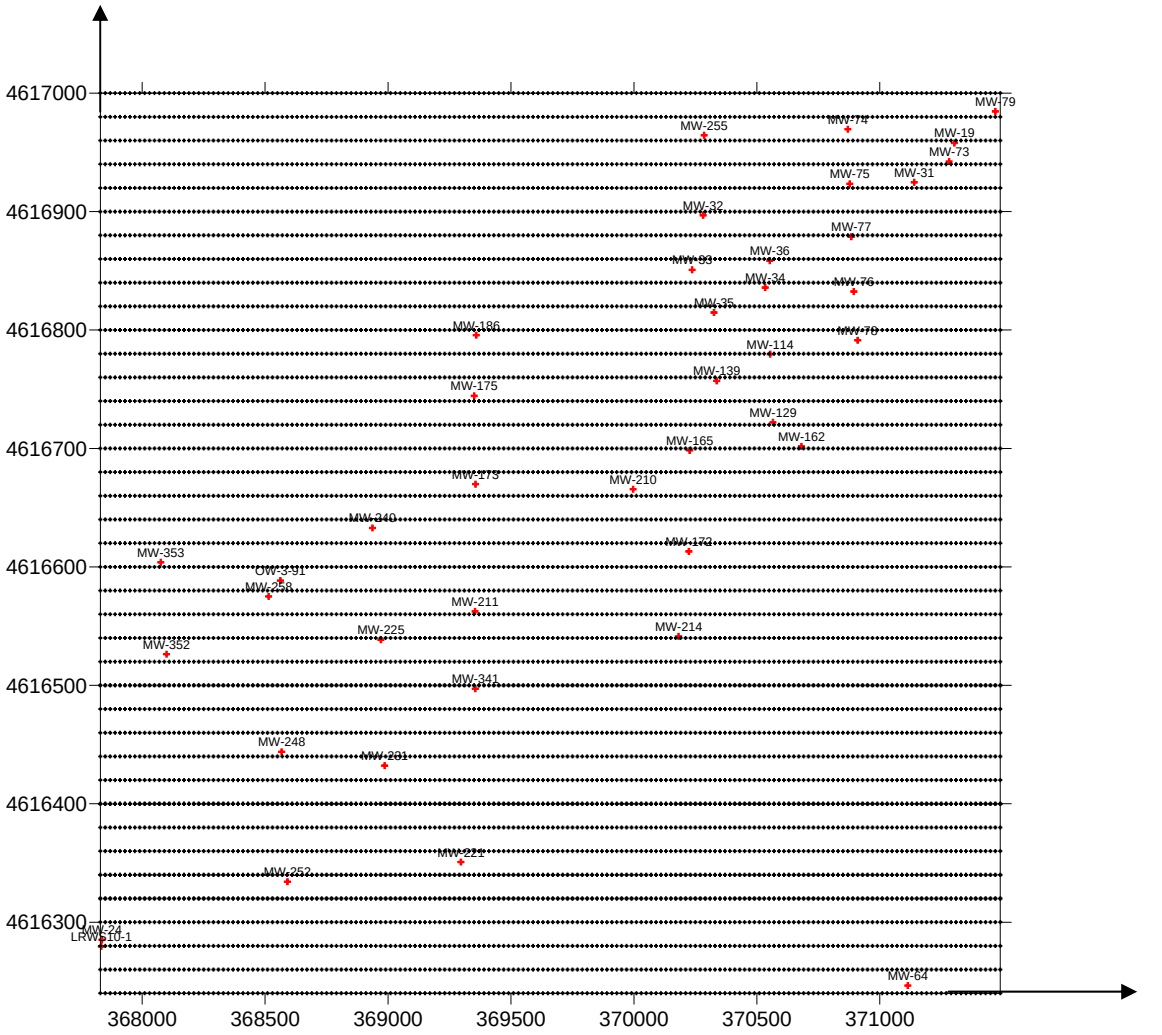
Y- Coordinates (North)



X – Coordinates (East)

Figure 7.1. Locations of groundwater monitoring wells at Demo 1 site

Y- Coordinates (North)



X-Coordinates (East)

Figure 7.2. Demo 1 site was divided both in the x and y directions at 25-ft intervals

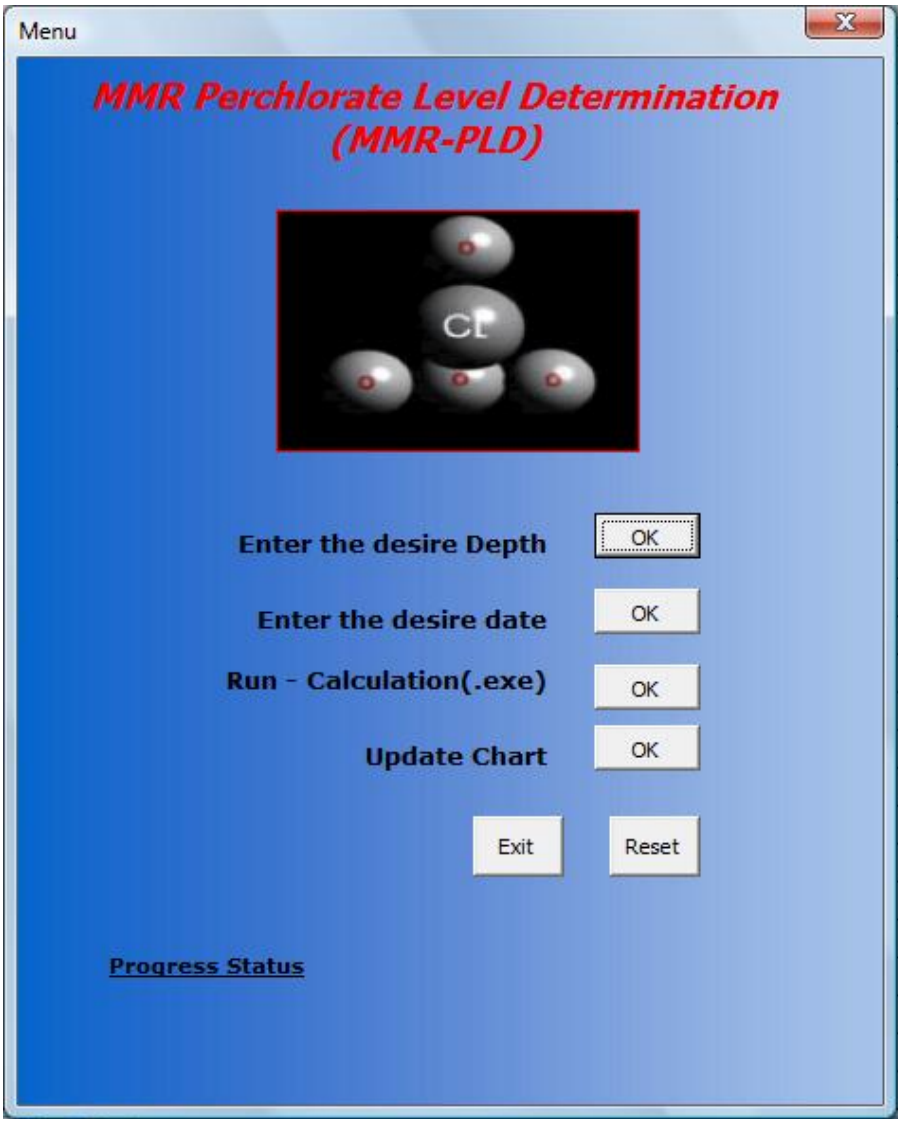


Figure 7.3. Interface of Excel application for ANN profiling model

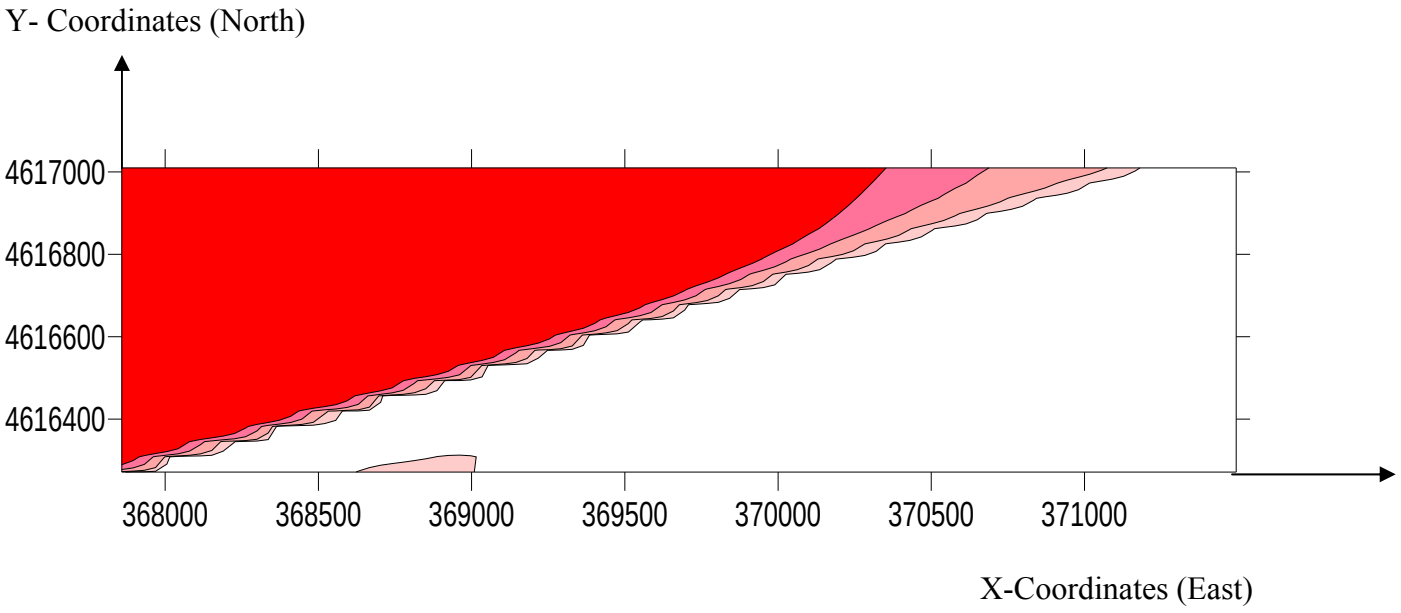


Figure 7.4a. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z = -50 ft

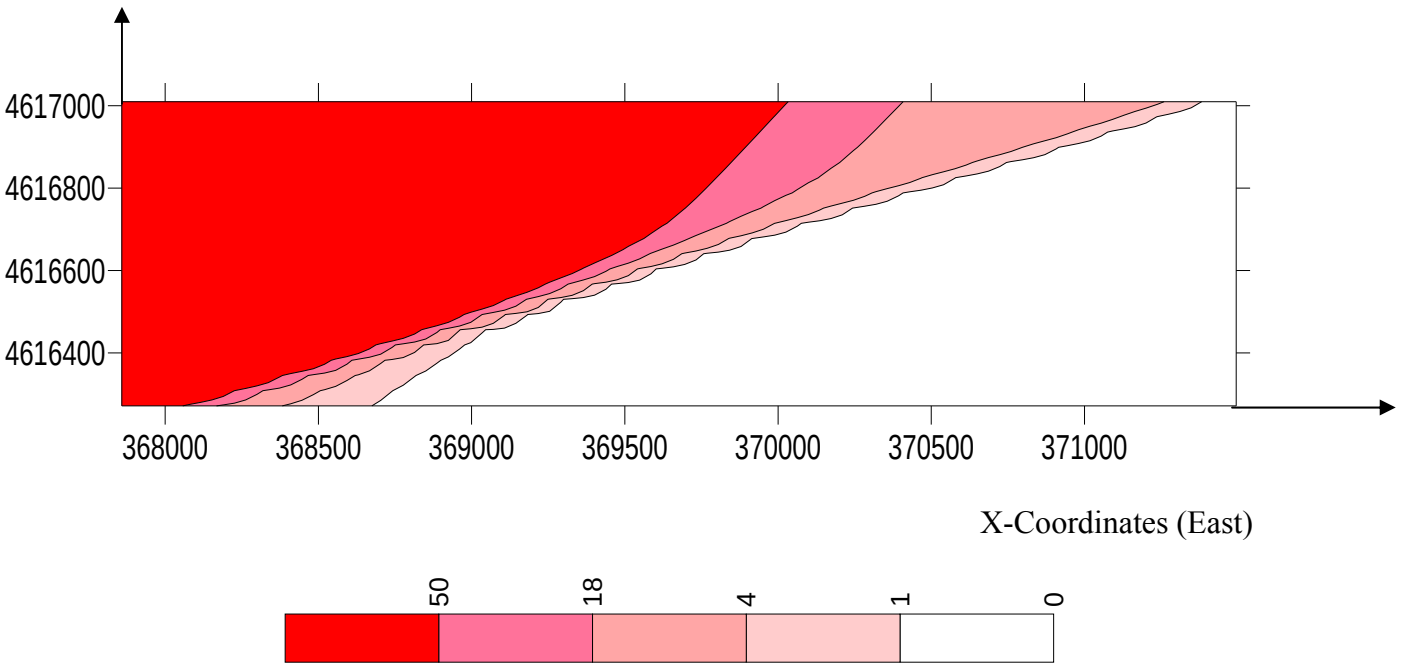


Figure 7.4b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z = -50 ft

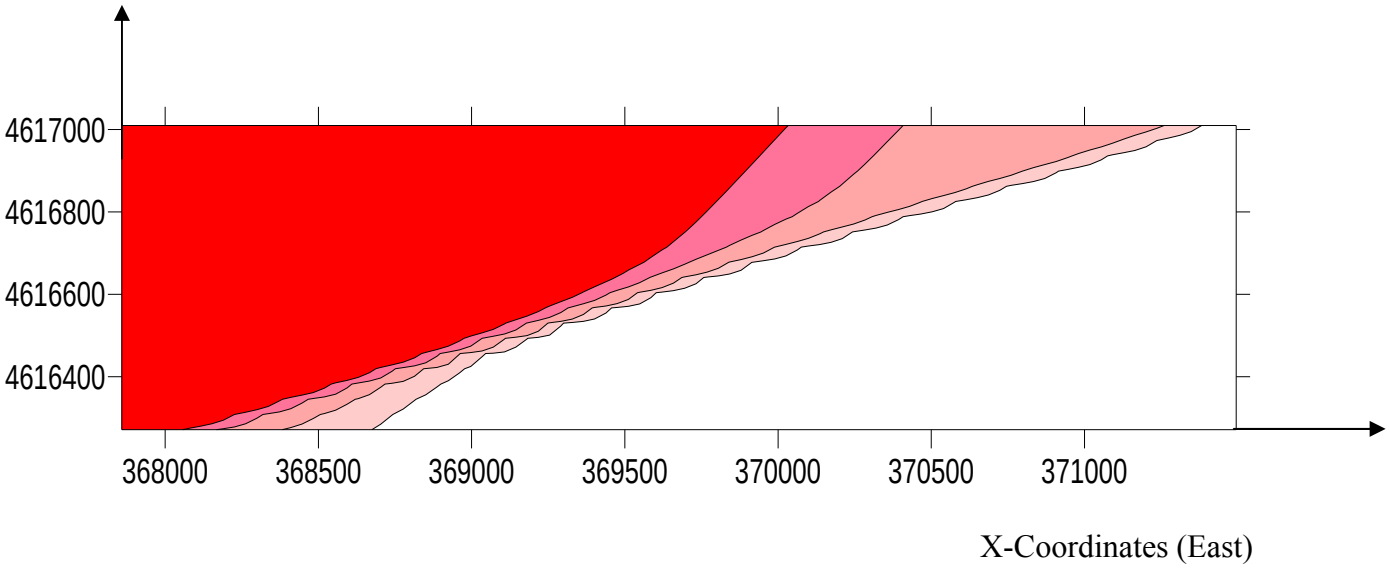


Figure 7.4c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = -50 ft

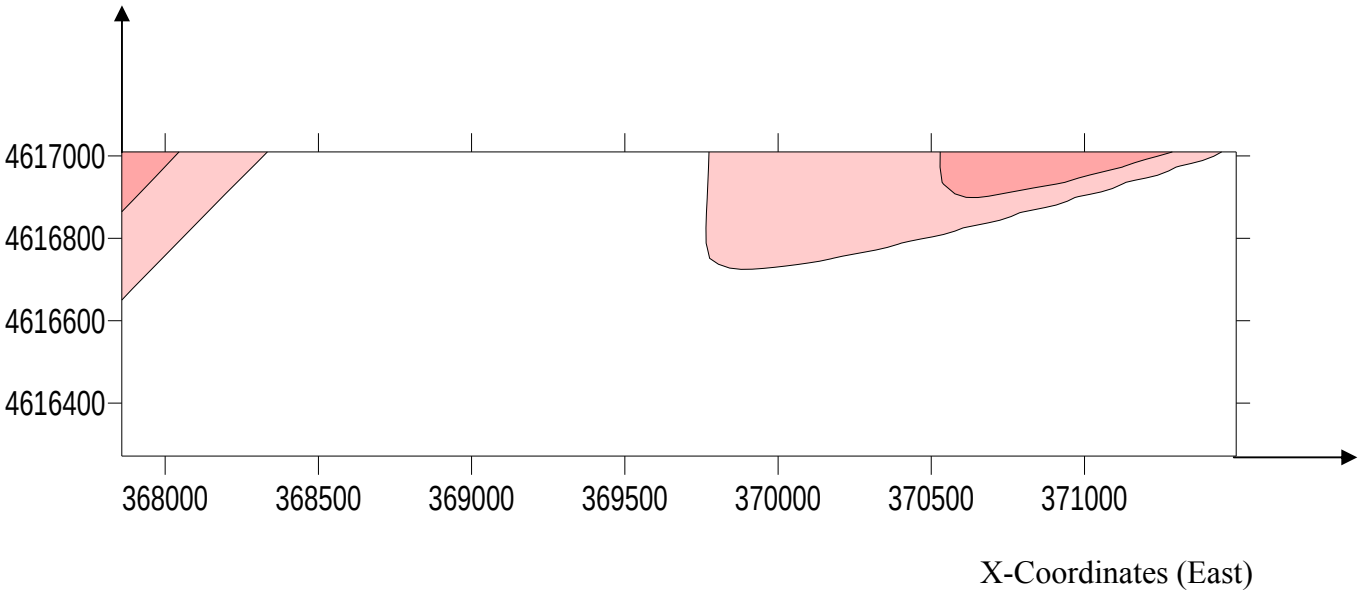


Figure 7.4d. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = -50 ft

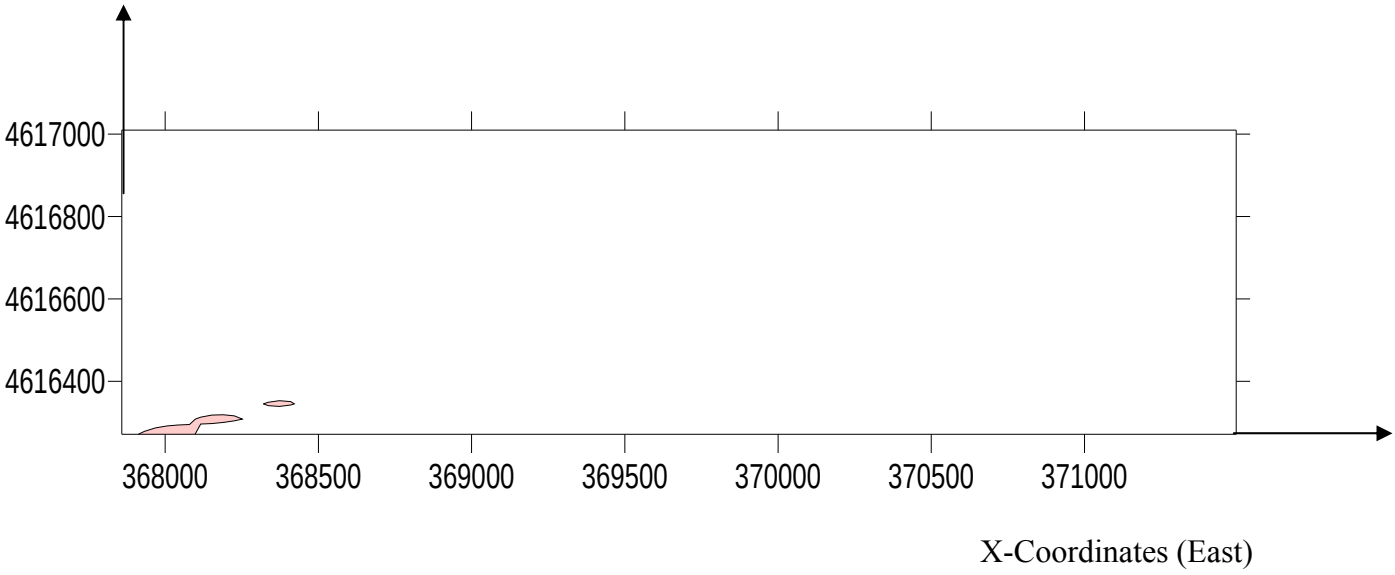


Figure 7.4e. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = -50 ft

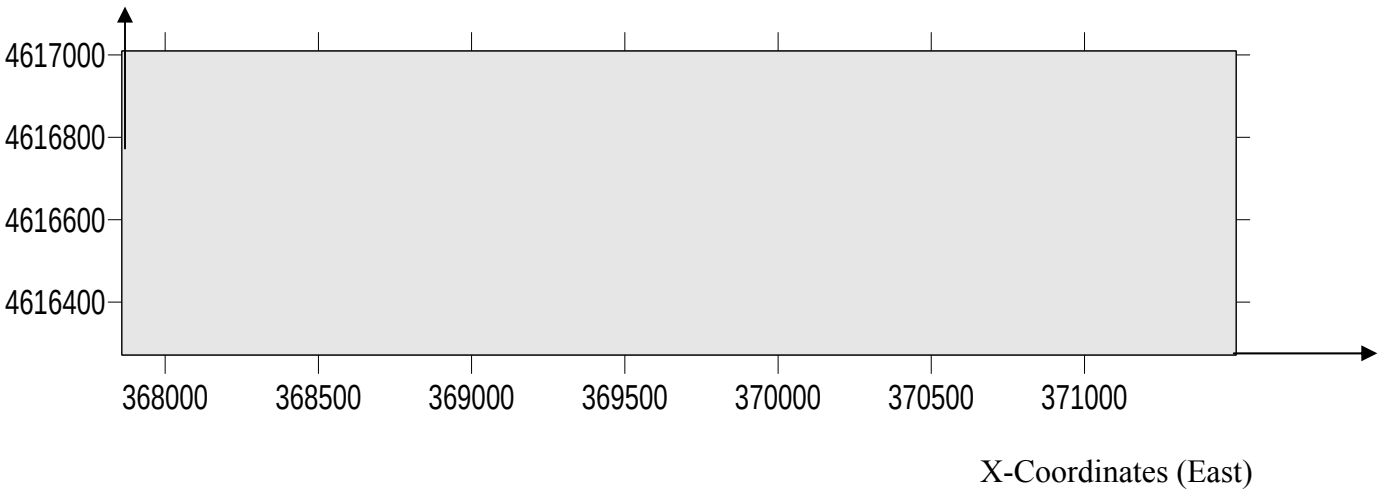


Figure 7.4f. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = -50 ft

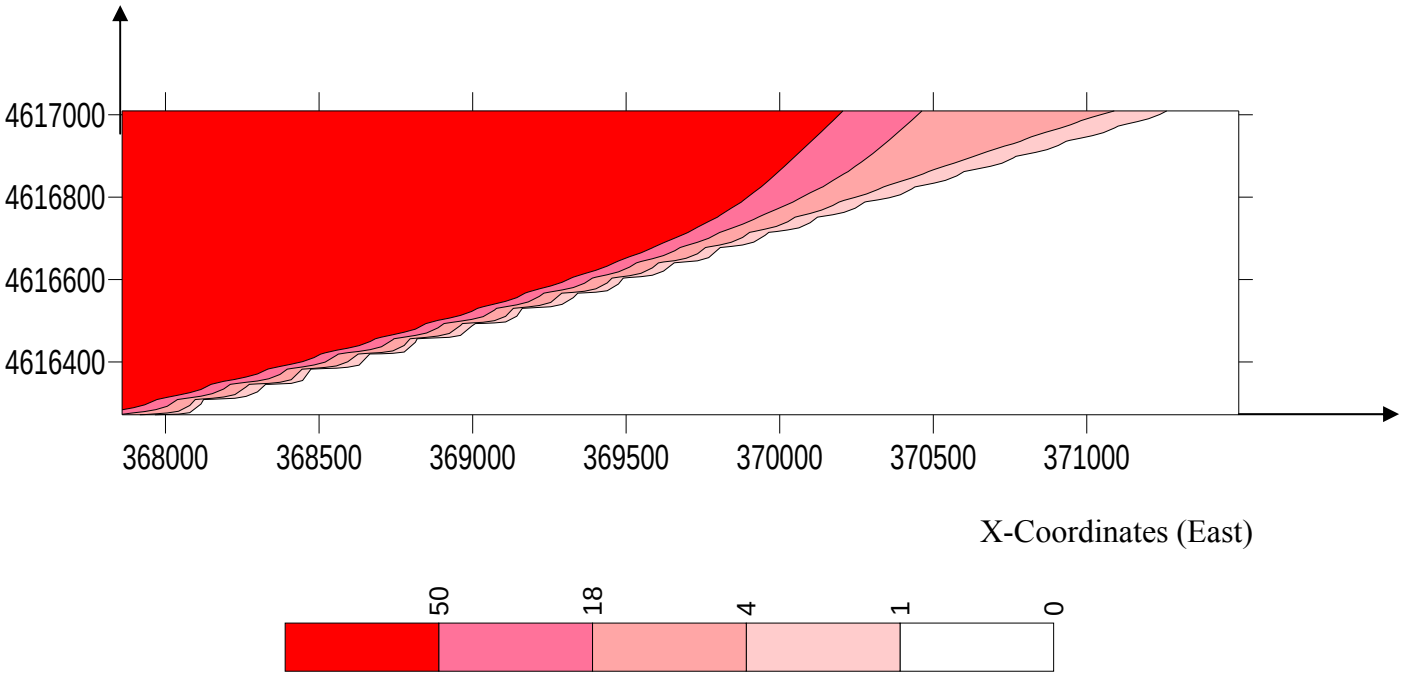


Figure 7.5a. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z = -25 ft

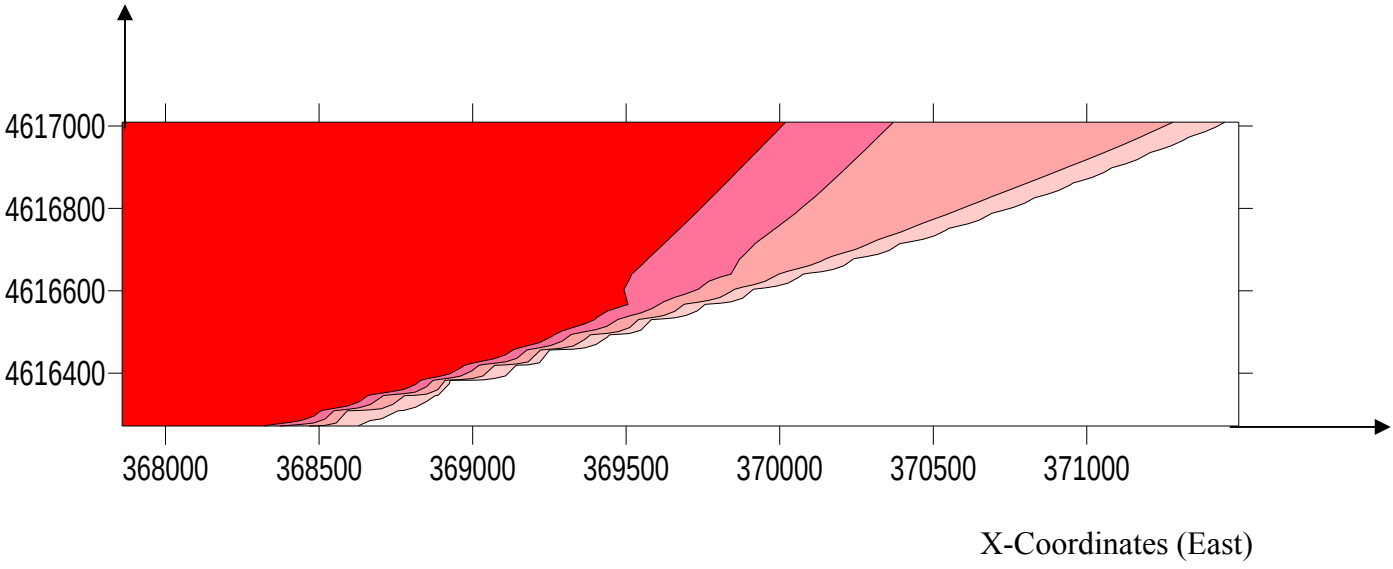


Figure 7.5b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z = -25 ft

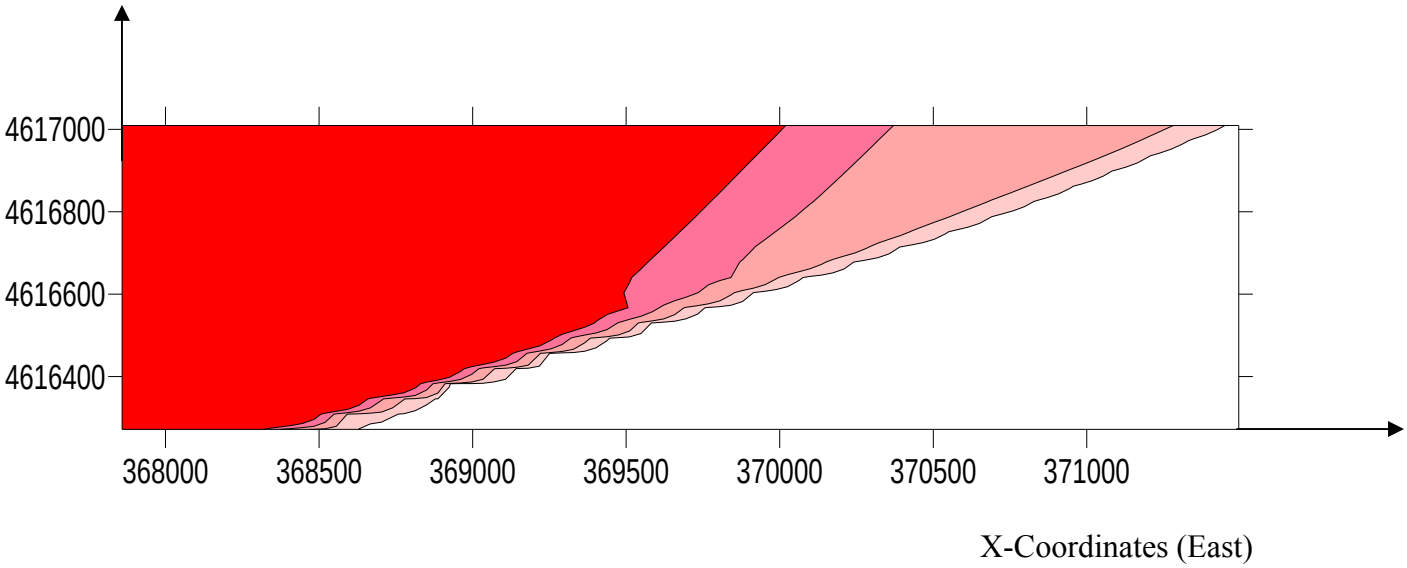


Figure 7.5c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = -25 ft

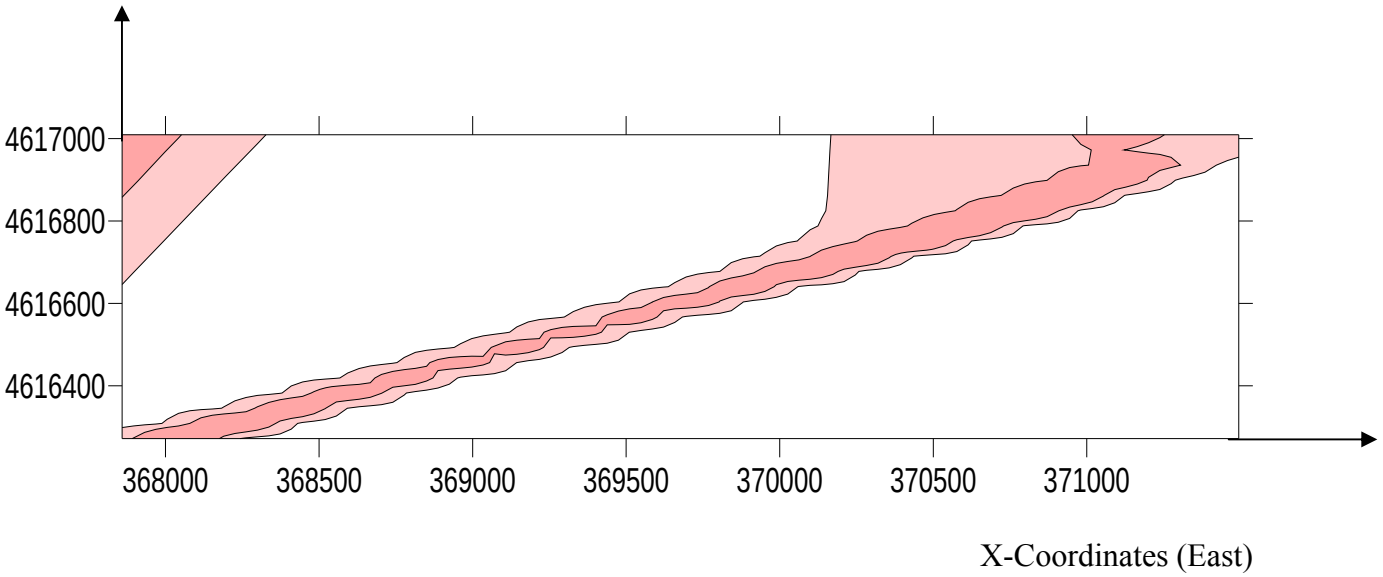


Figure 7.5d. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = -25 ft

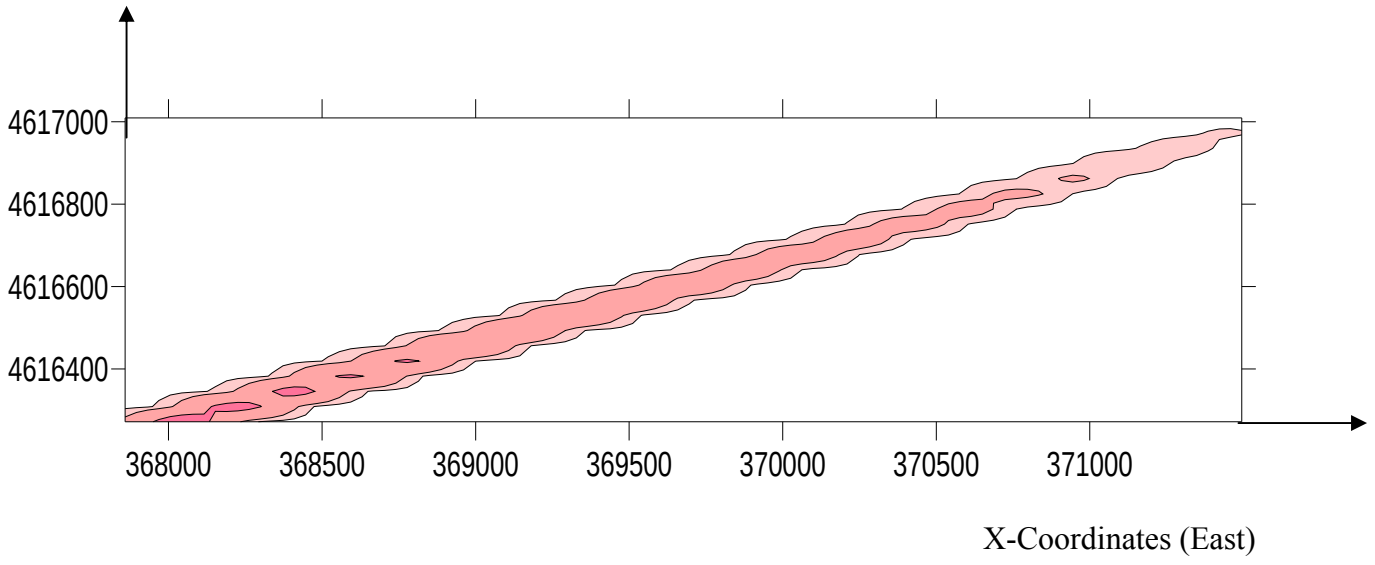


Figure 7.5e. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = -25 ft

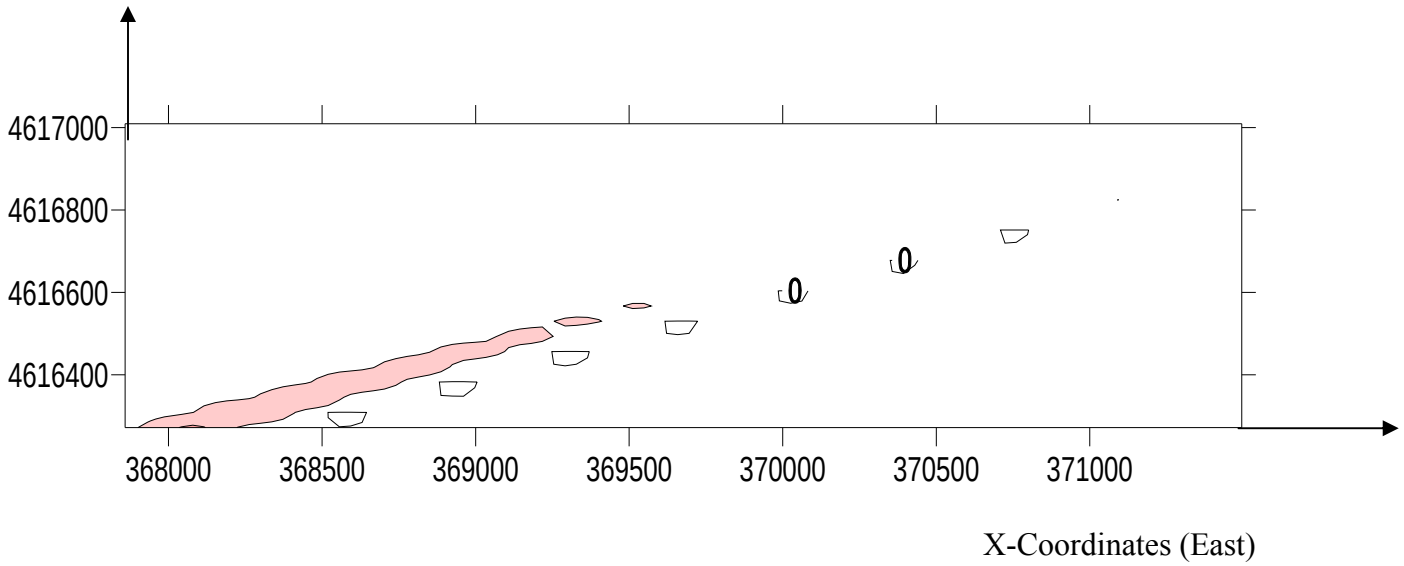


Figure 7.5f. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = -25 ft

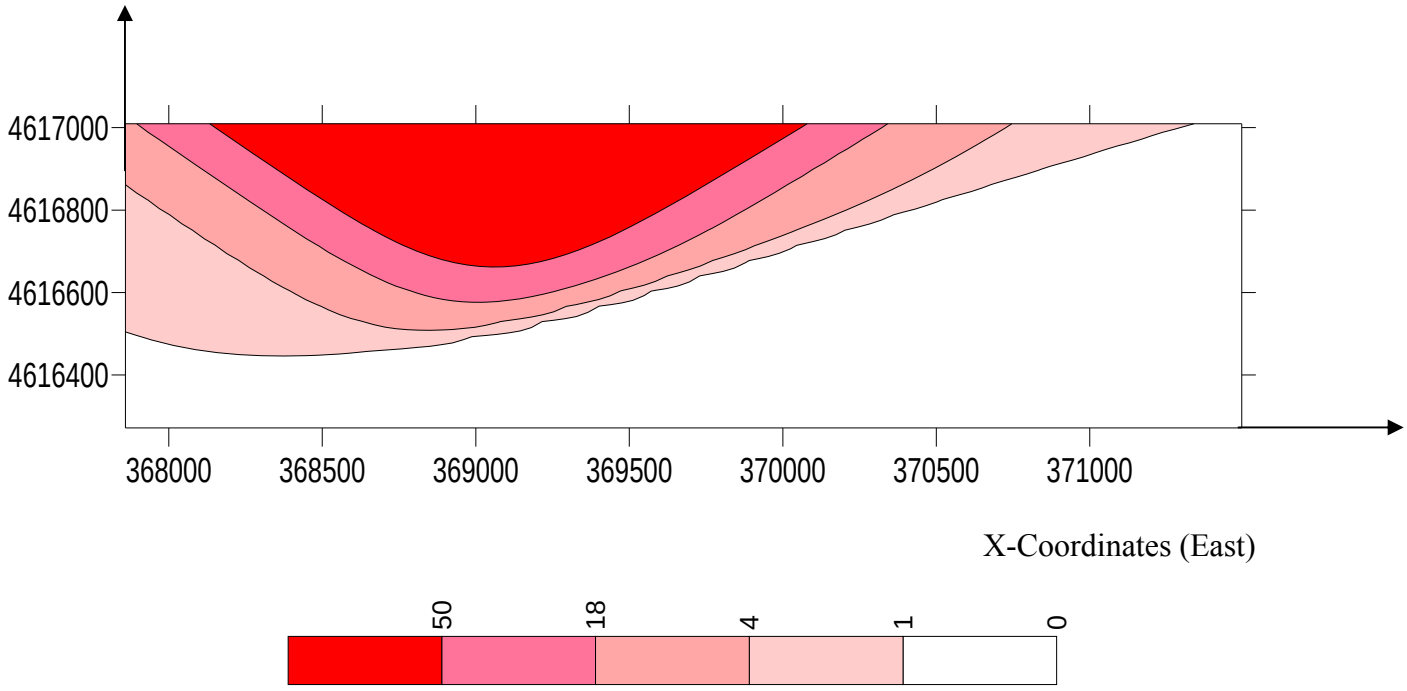


Figure 7.6a. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z =0 ft

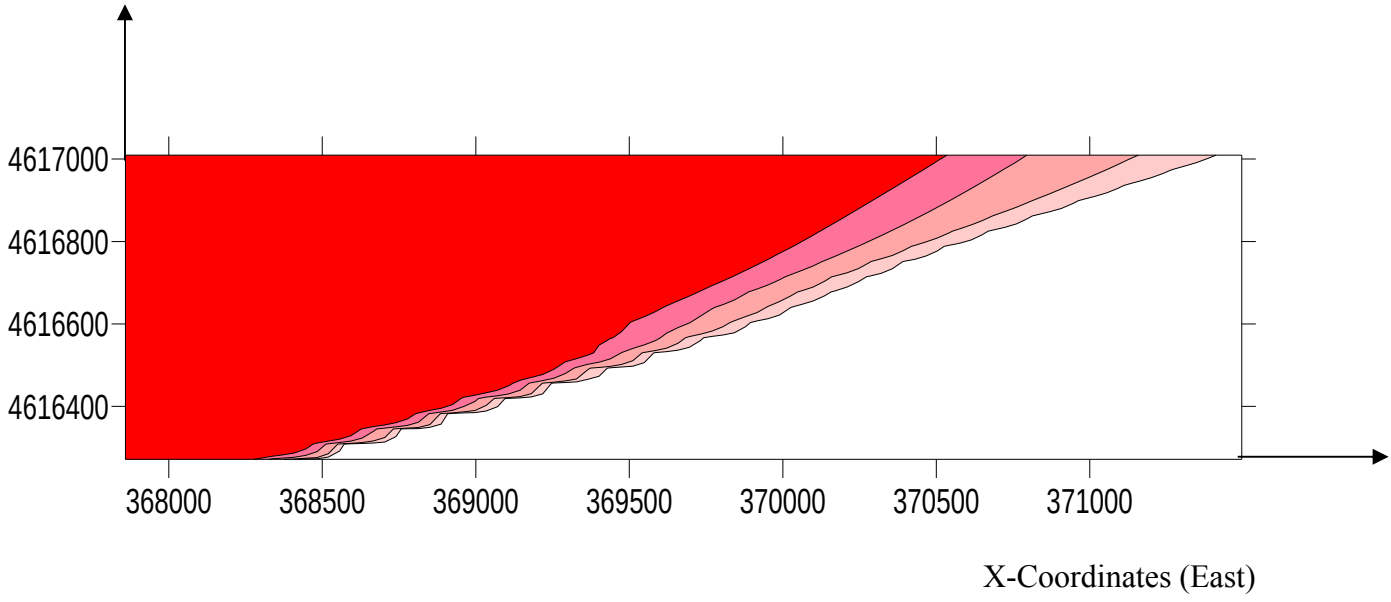


Figure 7.6b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z = 0 ft

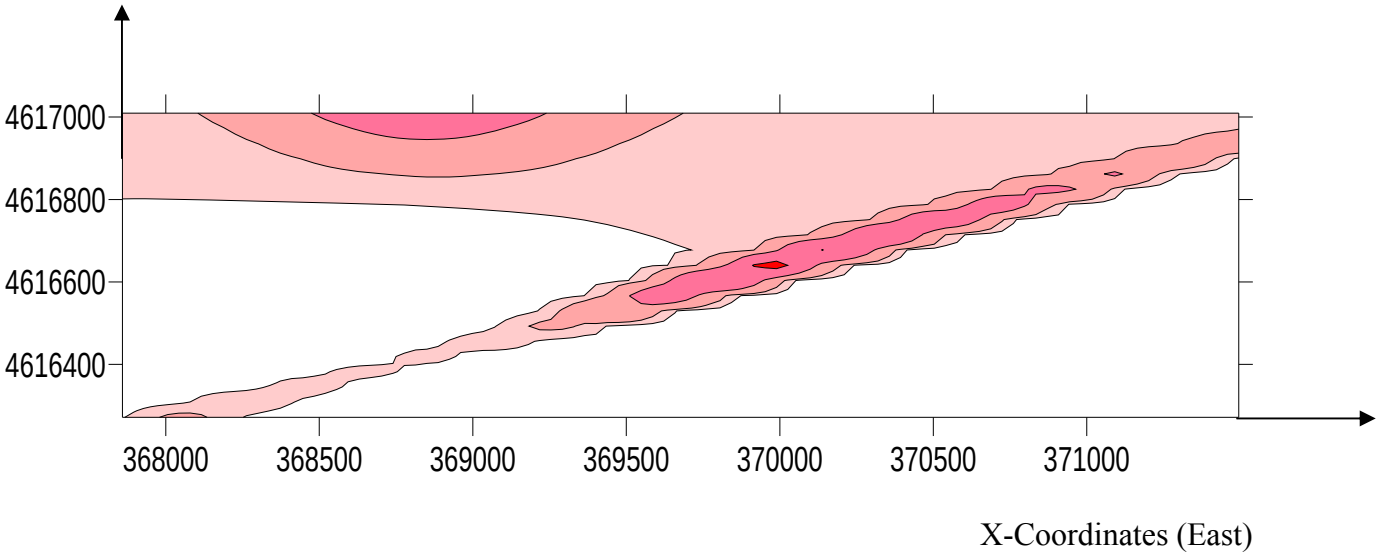


Figure 7.6c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 0 ft

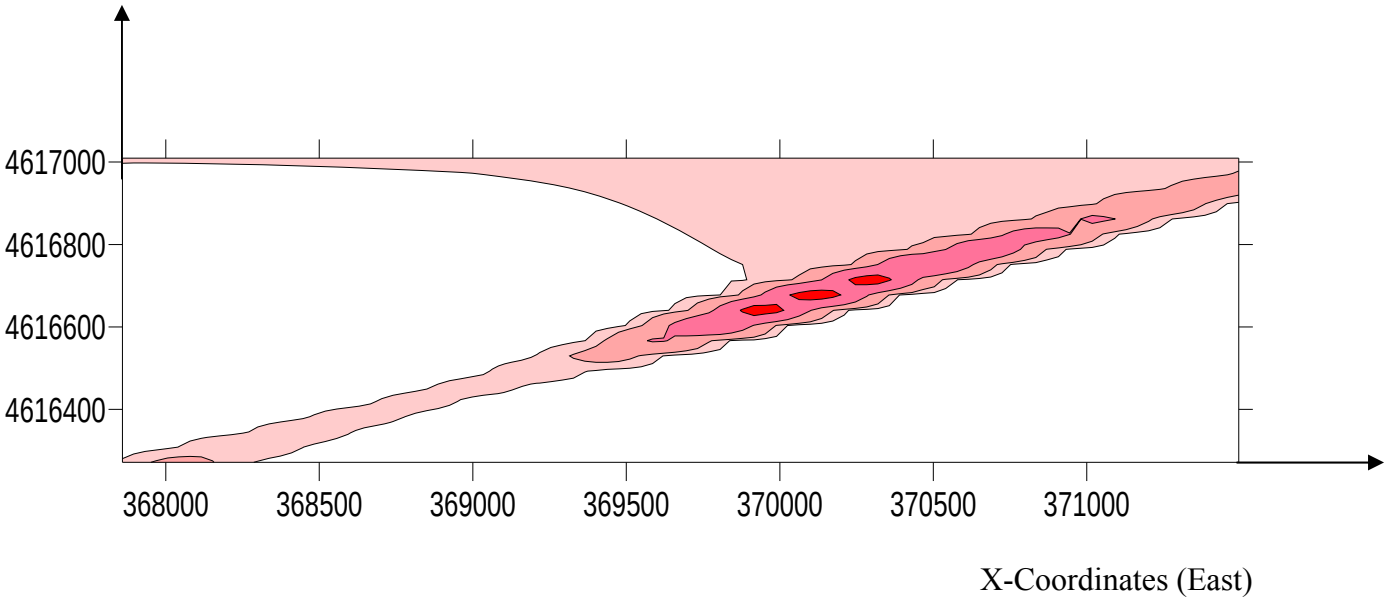


Figure 7.6 d. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 0 ft

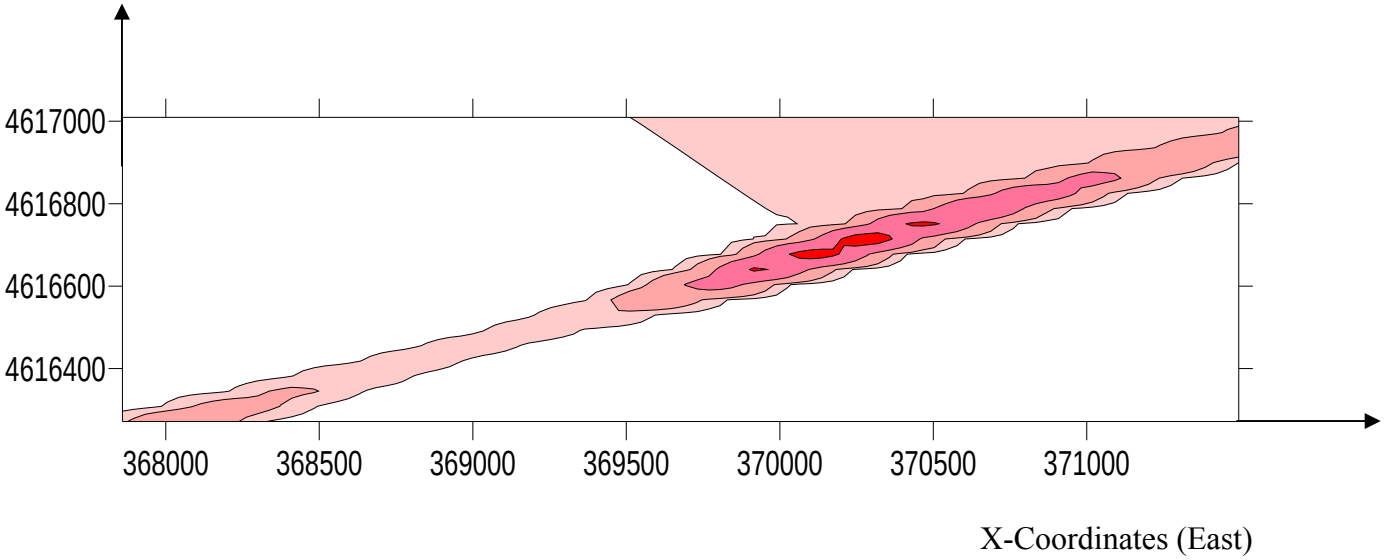


Figure 7.6e. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 0 ft

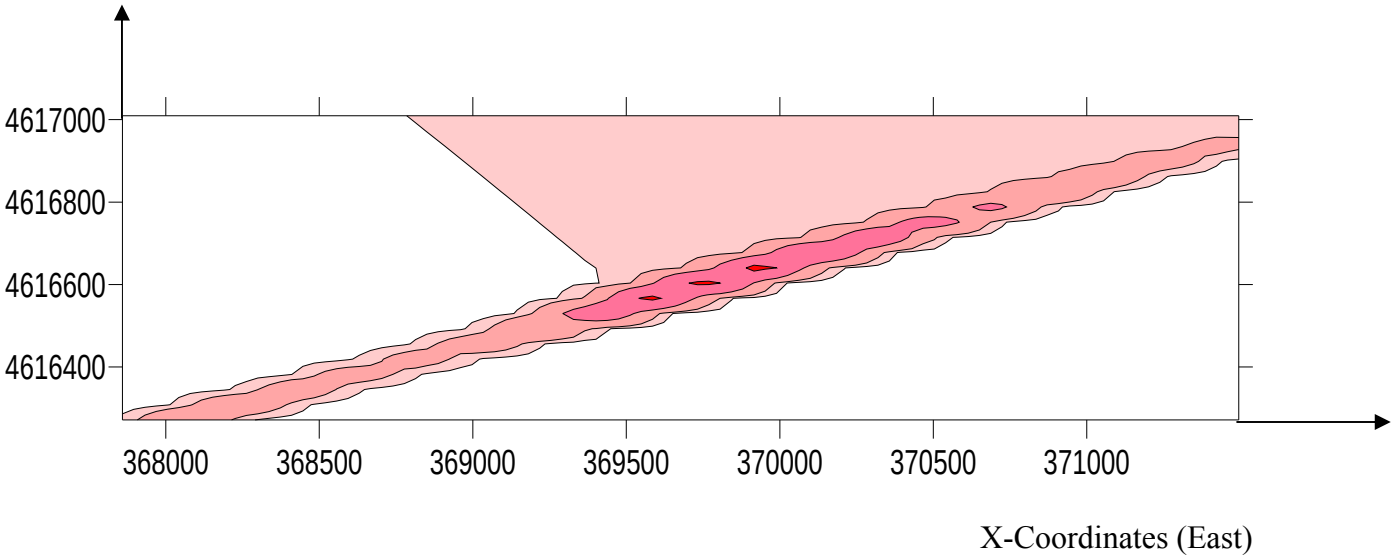


Figure 7.6f. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 0 ft

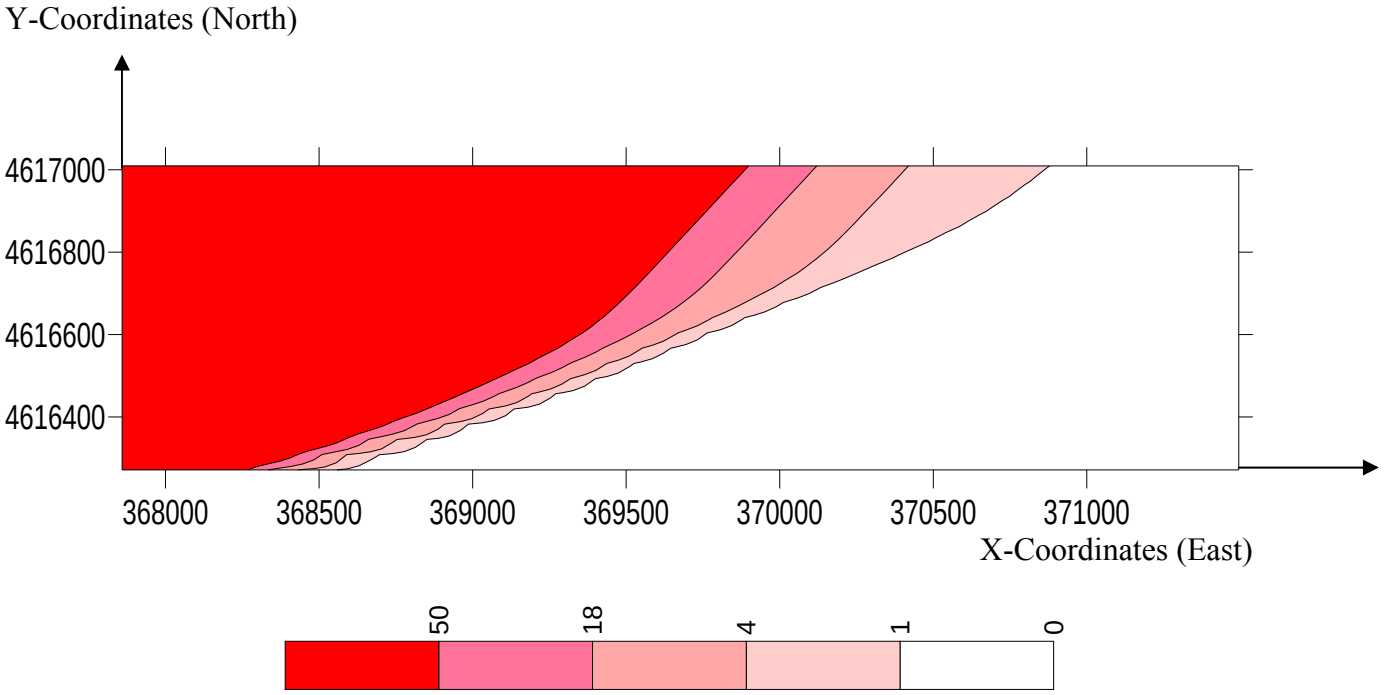


Figure 7.7a. Contour map showing the distribution of perchlorate in groundwater in 2000

at Z = 25 ft

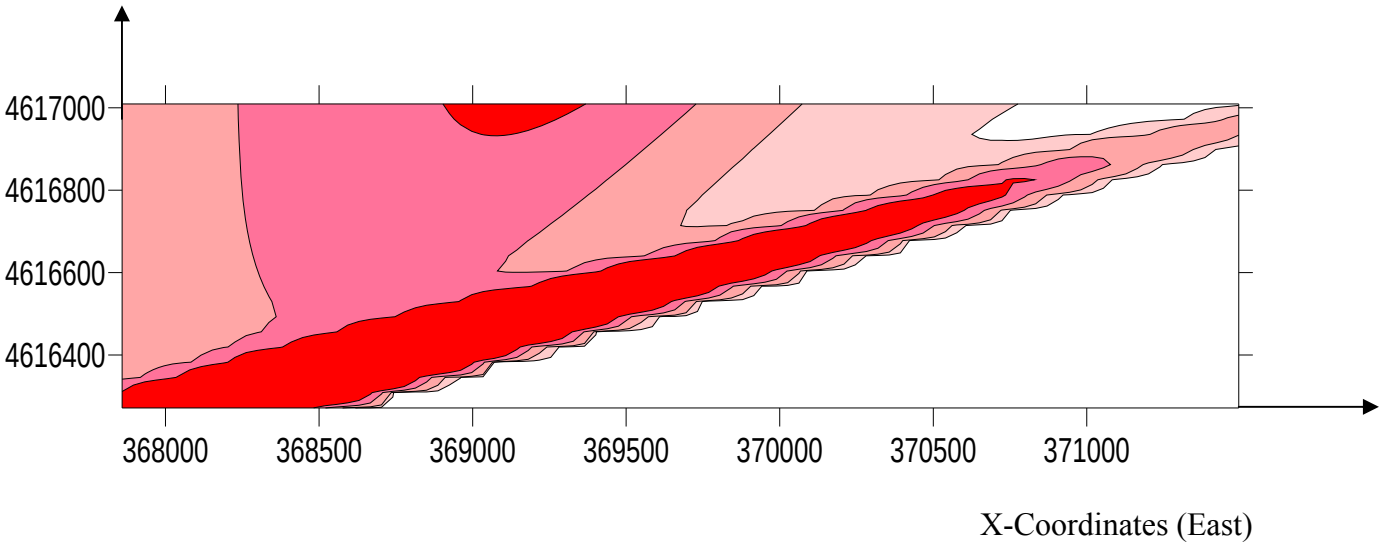


Figure 7.7b. Contour map showing the distribution of perchlorate in groundwater in 2001

at Z =25 ft

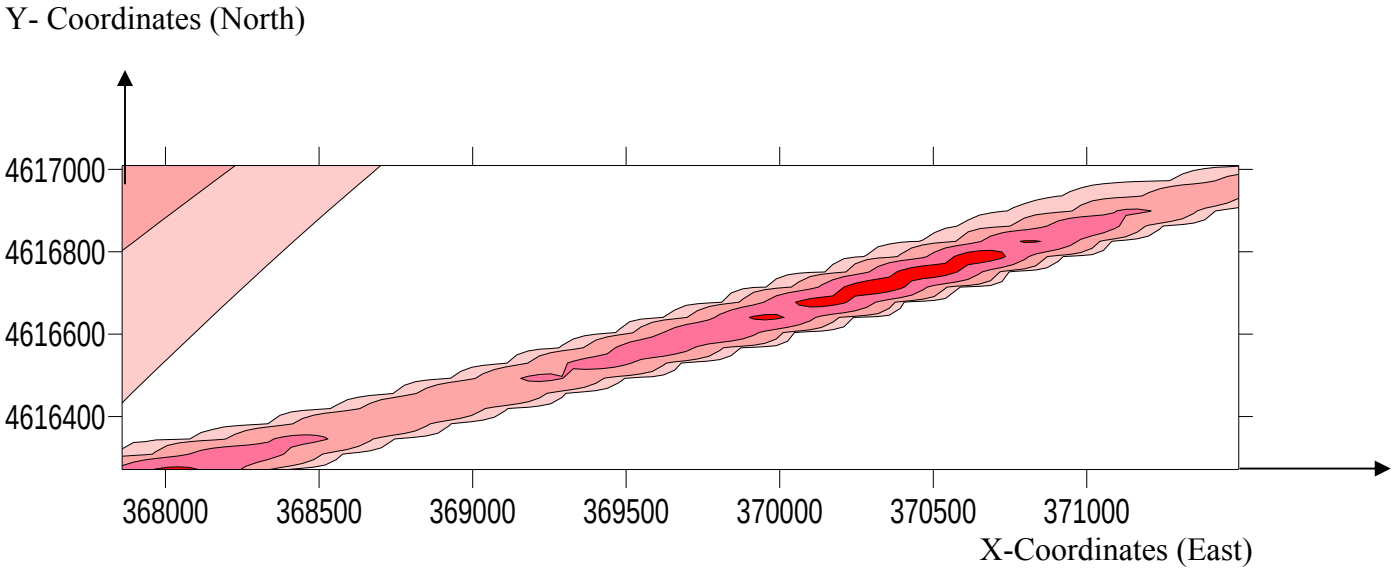


Figure 7.7c. Contour map showing the distribution of perchlorate in groundwater in 2002

at Z = 25 ft

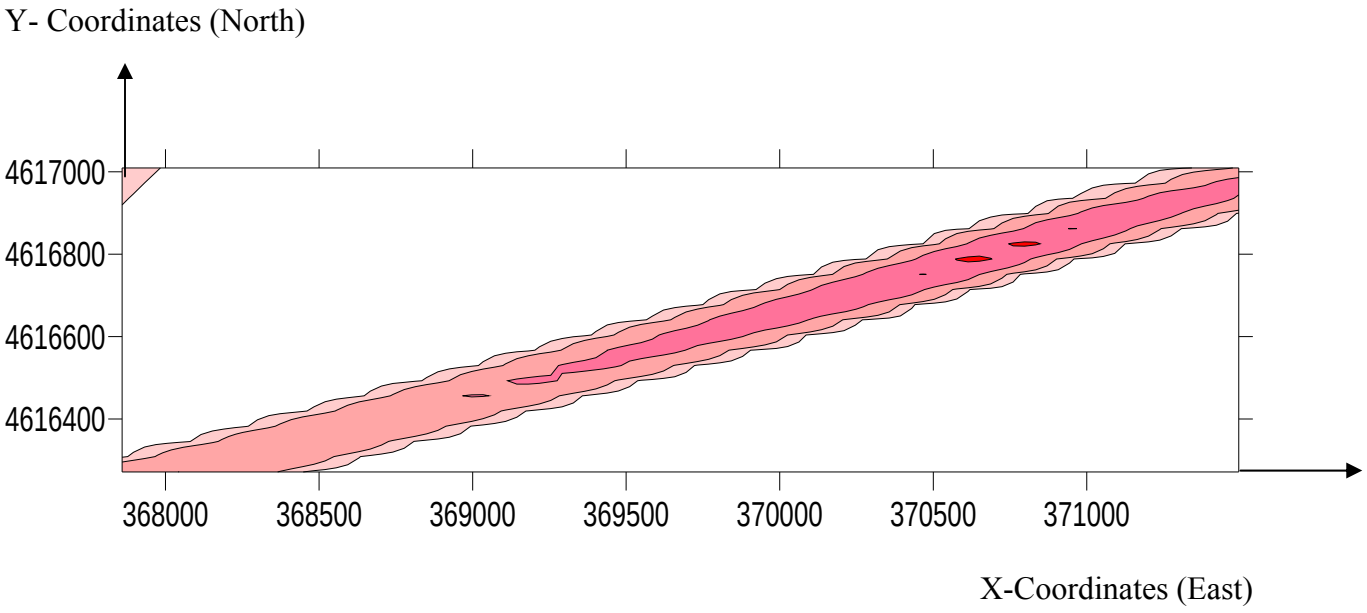


Figure 7.7d. Contour map showing the distribution of perchlorate in groundwater in 2003

at Z = 25 ft

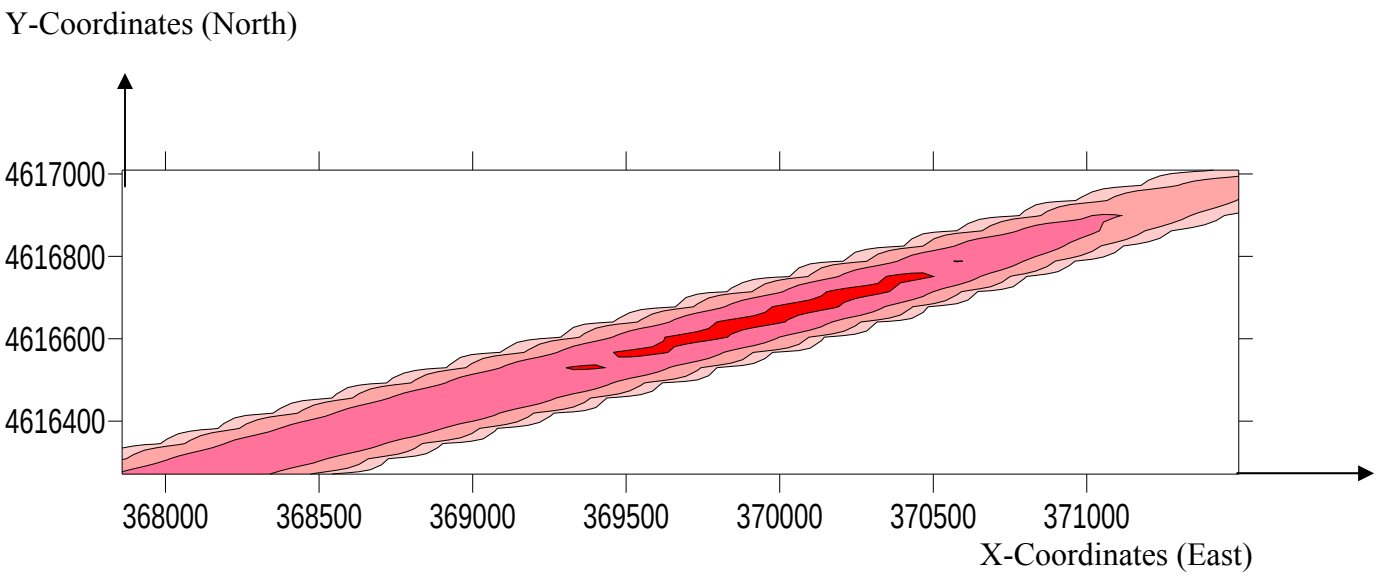


Figure 7.7e. Contour map showing the distribution of perchlorate in groundwater in 2004

at Z =25 ft

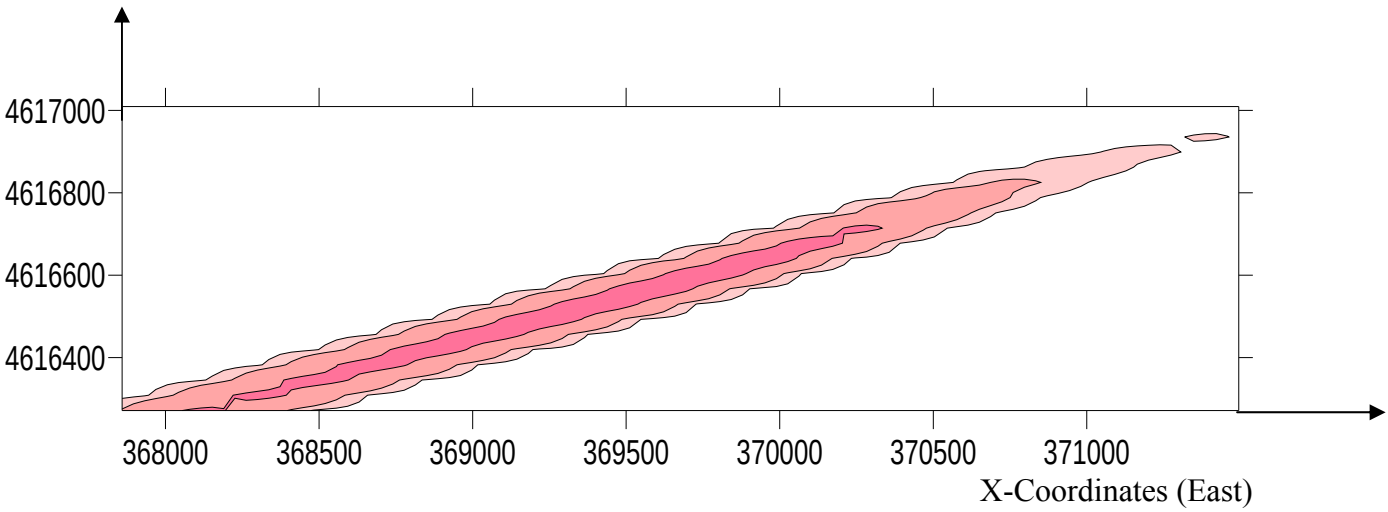


Figure 7.7f. Contour map showing the distribution of perchlorate in groundwater in 2005

at Z = 25 ft

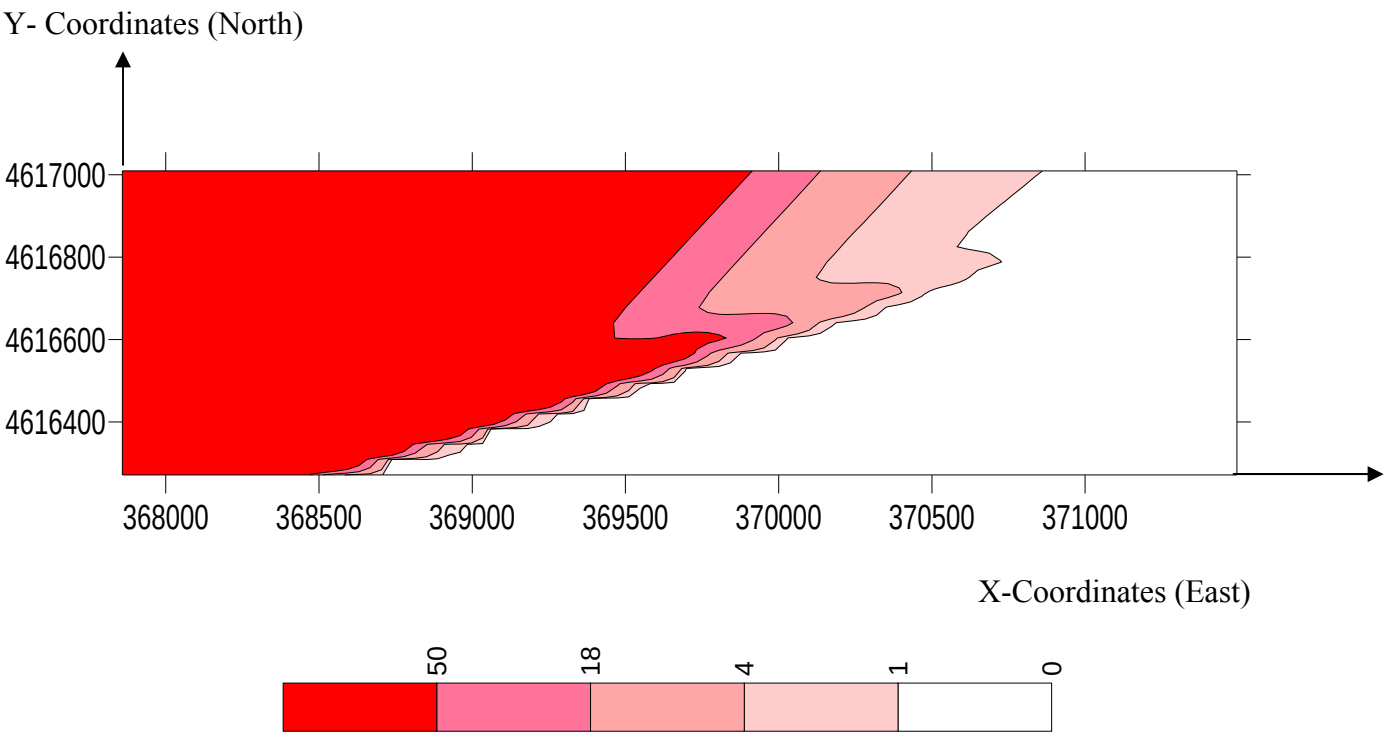


Figure 7.8a. Contour map showing the distribution of perchlorate in groundwater in 2000

at Z = 50 ft

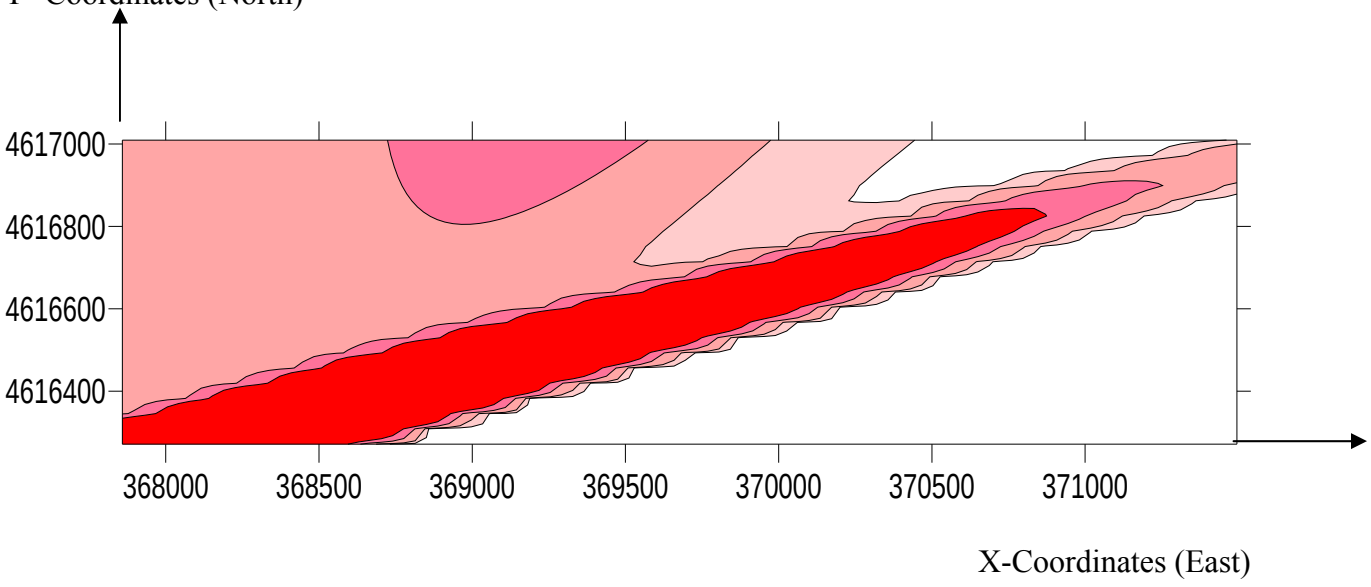


Figure 7.8b. Contour map showing the distribution of perchlorate in groundwater in 2001

at Z = 50 ft

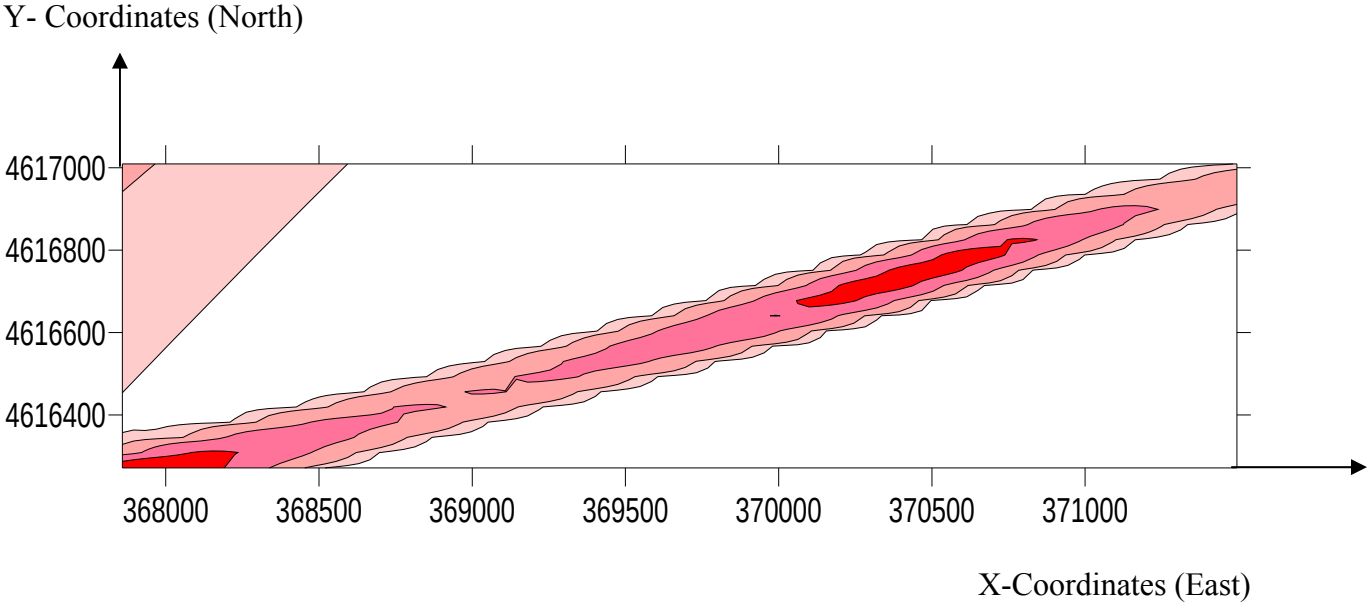


Figure 7.8c. Contour map showing the distribution of perchlorate in groundwater in 2002

at Z = 50 ft

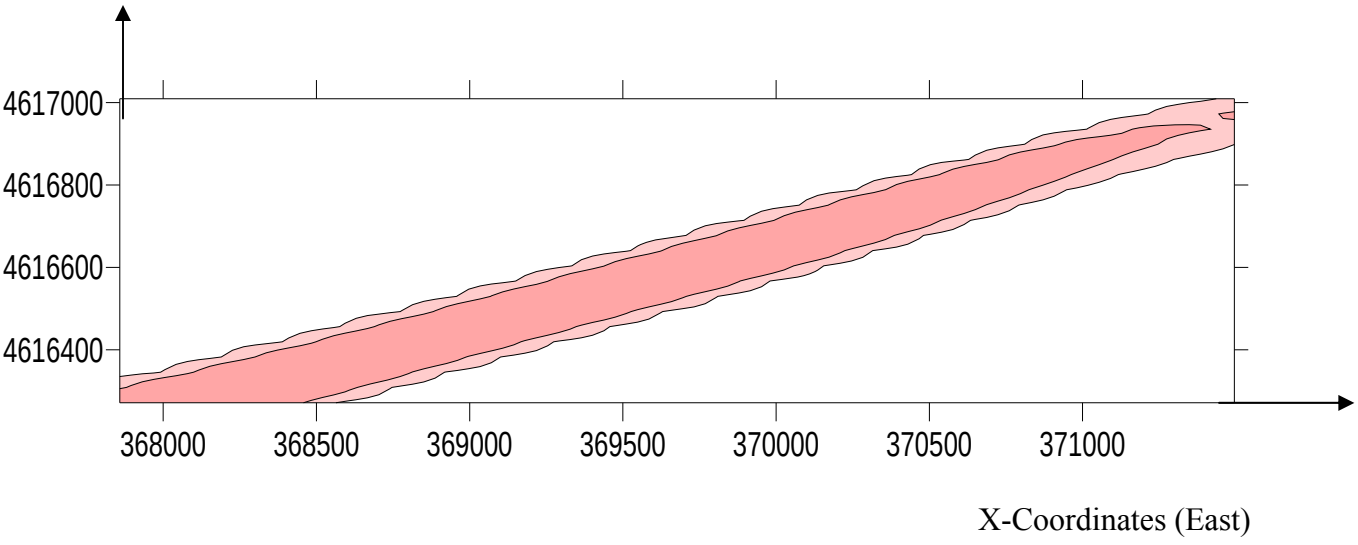


Figure 7.8d. Contour map showing the distribution of perchlorate in groundwater in 2003

at Z = 50 ft

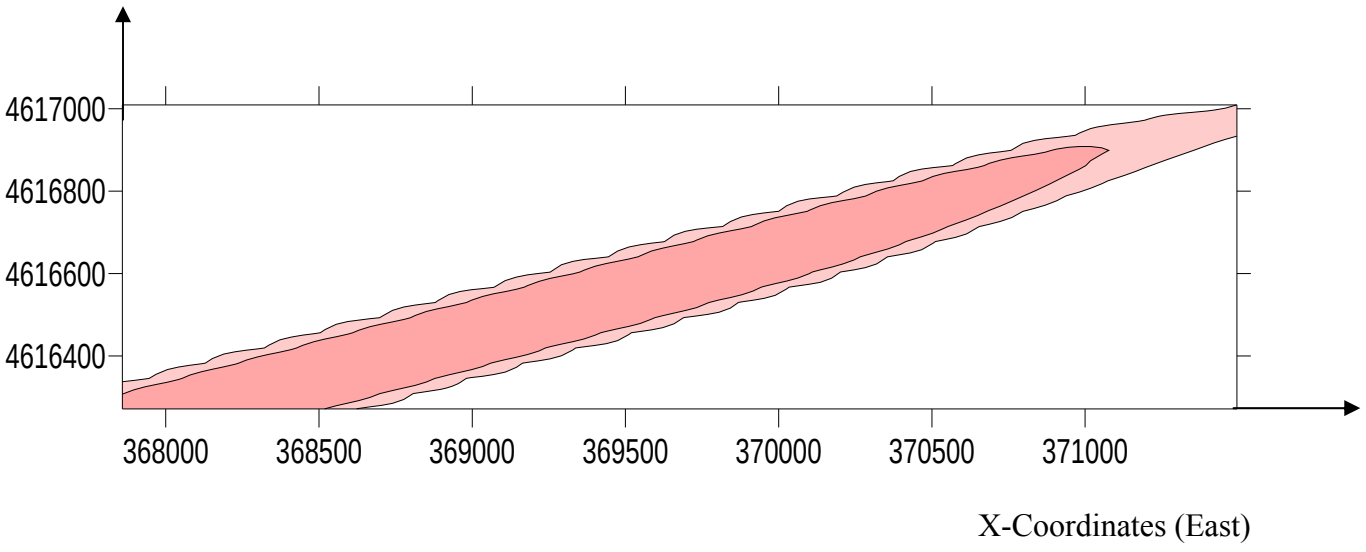


Figure 7.8e. Contour map showing the distribution of perchlorate in groundwater in 2004

at Z = 50 ft

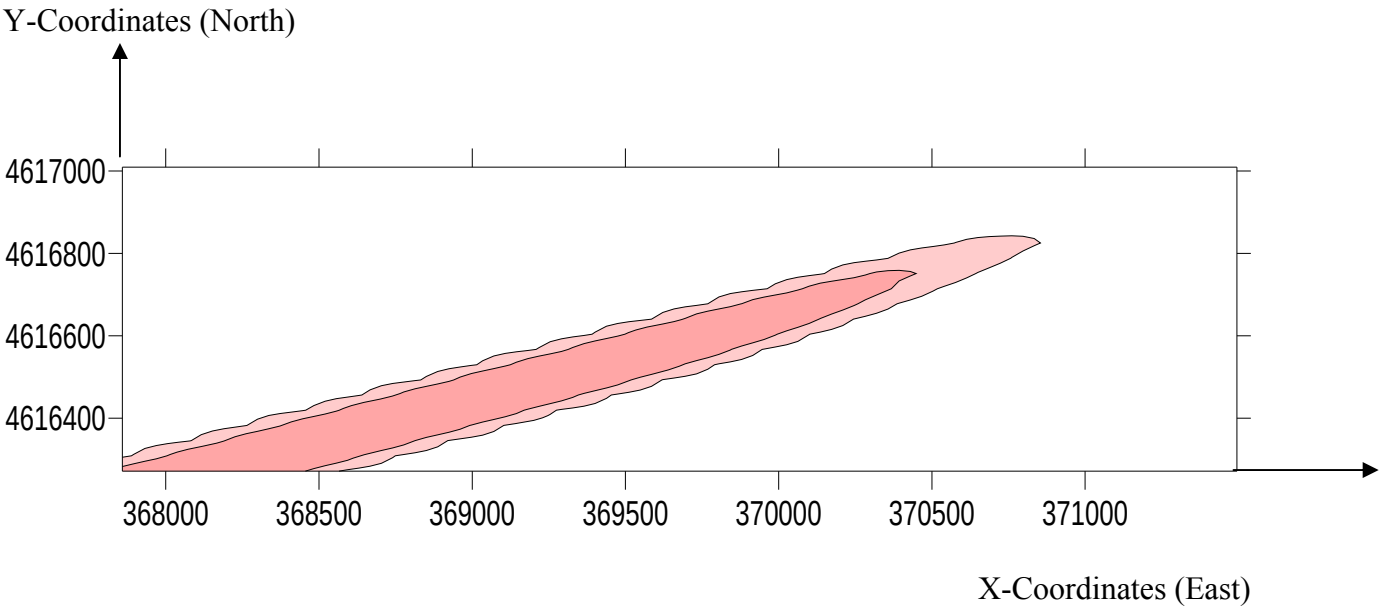


Figure 7.8f. Contour map showing the distribution of perchlorate in groundwater in 2005

at Z = 50 ft

CHAPTER 8 - **ANN-Based Profiling: Data Importance**

8.1 Introduction

Neural network models were initially formed to solve real-life problems. They were generated to function like the human brain, and consequently to produce reasonable outputs. Testing the performance capability is the first step in the development of any network that is to be used in field applications. Once a network shows a good performance for certain data, its structure might be used for other similar applications with similar databases. If the data values provided for the new application fall in the same range of the original data used for developing the network, then the developed network might be directly used with no modification. On the other hand, if the new data have a different range for their parameters, then the previously developed network could be used as a guide to develop new networks. In such a case, the inputs and outputs of the old structure might be used in developing new networks.

In Chapter 7, neural networks were used to predict the perchlorate concentration values at any desired (x, y, z) point at MMR-Demo 1. In this chapter, similar procedures will be carried out to highlight the critical number (importance) of monitoring wells needed to examine the extent of the perchlorate contamination.

8.2 MMR-Demo 1 Groundwater Monitoring Network

According to Ogden (1999), the MMR-Demo 1 groundwater monitoring network is composed of a total of 41 monitoring wells, which are installed throughout the Demo 1 site (Figure 8.1). Each monitoring well has been sampled from one to eight times depending on when it was installed. These samples were analyzed for over 200 compounds to include explosives,

volatile organic compounds, semi-volatile organic compounds, pesticides and herbicides, polychlorinated biphenyls (PCBs), polychlorinated naphthalenes, dioxins, and heavy metals (USACE, 1999). Groundwater analytical results for all sampling rounds conducted from January 2000 through December 2005 are provided in Appendix A and will be utilized in this study.

Amegashie, et al., (2006) stated that the groundwater site characterization investigation consisted of groundwater profiling and installation of monitoring wells upgradient and downgradient from contaminant source plume. Ogden (1999) reported that a total of six monitoring well triplets were installed. Well locations MW-74 through MW-78 were established along a transect perpendicular to the plume, approximately 1700 ft downgradient of the source area represented by MW-19. A total of 15 wells were installed at these five locations to delineate the lateral and vertical extent of perchlorate in groundwater. Three wells were also installed at MW-79, east of the source area, to evaluate upgradient groundwater quality. Shallow monitoring well MW-73S was installed to provide additional characterization of groundwater at the water table in the source area (Clausen, et, al. 2004).

8.3 ANN Model Development

The data of the MMR groundwater monitoring wells network (Appendix A) was utilized herein to fully investigate the data importance aspect of this research study.

Based on the available MMR-Demo 1 groundwater monitoring network data and other environmental site profiling studies, such as the ones conducted by Mryyan & Najjar in 2005, Mryyan & Najjar in 2006 and Dowla & Rogers in 1991, back-propagation neural network was chosen as the most-appropriate ANN for developing the groundwater site characterization prediction model. Back-propagation networks were developed using the TR-SEQ1, a three-layered ANN training software package developed by Najjar (2001). The purpose of the ANN

model is to predict the concentrations of perchlorate at the MMR from appropriate input parameters and to predict the impact of exclusion and/or inclusion of any monitoring wells from the Demo 1 site groundwater monitoring wells network. In this chapter, twenty-eight back-propagation neural network models were trained and tested, because the inclusion or exclusion of certain monitoring wells variables could influence the ANN model prediction results.

8.3.1 Determination of Appropriate Model Inputs

This chapter uses the same methodology used in Chapter 7 to determine the appropriate number of input parameters. As in Chapter 7, all of the nine potential input parameters (X, Y, G, Z, T, R, INJ1, INJ2 and INJ3) were necessary to accurately predict the amount and the distribution of perchlorate at Demo 1 and the minimum number of monitoring wells necessary to accurately characterize the Demo 1 site.

8.3.2 Data Banks and Model Selection

The optimal network model for the prediction of perchlorate at the MMR was determined to contain nine input parameters, nine hidden nodes, and one output parameter (9-9-1). Once this was determined, the network was used to model the impact of exclusion/inclusion of monitoring wells from the network. The only parameter required in order to provide needed predictions is the input data vector (X, Y, G, Z, T, R, INJ1, INJ2, and INJ3).

Twenty-eight models were developed by various testing trials through the exclusion or inclusion of certain groundwater monitoring wells. These models were then used to identify the minimum groundwater wells necessary to accurately characterize the Demo 1 site. A brief description of three out of the twenty-eight models developed is presented below.

Model 1. (Baseline model). The baseline model adopted the two-stage training methodology. In the first stage, which was completed in Chapter 7, [for the MMR perchlorate prediction model, it was determined that the optimal network model contained nine input parameters, nine hidden nodes, and one output parameter (9-9-1)] all 41 well data sets (459 data sets are listed in Appendix A) were used to develop the desired perchlorate prediction model. In this model, the database used for training represents all possible features and sub-features that the network is required to learn. In the first stage, the entire database was divided into training, testing and validation sub-databases at the ratio of about 50%: 25%: 25%. The training sub-database contained all the datasets with the maximum or minimum value of each input and output parameter. Using the training and testing datasets, the least-error-structure is selected based on the following statistical accuracy measures (Tables 8.1): Averaged Squared Error (ASE), Mean Absolute Relative Error (MARE) and Coefficient of Determination (R^2) on the testing datasets. Then the network was retrained at its least-error structure (defined from stage I) 9-9-1 on all data sets.

Model 2. In this model, well 165 data (20 data sets) were eliminated from the data bank to arrive at 40 wells and 439 data sets. Accordingly, the least-error-structure was determined. Following that, the second stage re-training on the least error structure (9-7-1) was conducted. Related statistical accuracy measures (for stage I and II) for model 2 are summarized in Table 8.2.

Model 3. In this model, well 32 data sets (24 data sets) were eliminated from the data bank to arrive at 40 wells and 435 data sets. Similarly, the least-error-structure (i.e., 9-8-1) was determined. Corresponding statistical accuracy measures for stage I and II are shown in Table 8.3.

For each of the three models mentioned above, once the ANN model was executed and perchlorate concentrations were predicted, contour maps were generated (Figures 8.2a through 8.2c described Model 1 through 4 for the year 2000; Figures 8.3a through 8.3c described Model 1 through 4 for the year 2001; Figures 8.4a through 8.4c described Model 1 through 4 for the year 2002; Figures 8.5a through 8.5c described Model 1 through 4 for the year 2003; Figures 8.6a through 8.6c described Model 1 through 4 for the year 2004, and Figures 8.7a through 8.7c described Model 1 through 4 for the year 2005). Also, network accuracy statistics, specifically the Averaged Squared Error (ASE), Mean Absolute Relative Error (MARE %), and Coefficient of Determination (R^2) were determined. Tables 8.4 show the network accuracy statistics for all twenty-eight groundwater ANN-based site characterization models. Note that all ASE, MARE% and R^2 values reported are the ones obtained from the stage II modeling.

8.4 Contour Maps

Once a decision is made about what network is used to represent a site, this network can be used to predict the perchlorate values at any desired location. The only thing a trained network needs in order to provide such predictions is the input data vector. In this site characterization study, the only input data needed is the (x, y, z) coordinates of the point at which a prediction is desired.

In this chapter, the Demo 1 site was divided in the x, y, and z directions using $\Delta x = \Delta y = 25$ ft. The grid system generated in the (x, y) plane produced 4,527 grid points (Figure 8.1). These coordinates were used for z = 0 ft. For each of the generated grid points, the perchlorate concentration values were predicted using the corresponding (x, y, z) coordinates via the least-error structure (optimized) networks. Predictions were made using data representing one specific date.

In all contour maps (Figures 8.2 through 8.7), a red color reflects a high perchlorate value and a light color represents a low perchlorate value. Consequently, a red color indicates a high value of perchlorate concentrations above the regulatory limit of 1 part per billion, and a white color indicates no perchlorate concentration. Color ramp is used between the red and white colors to map intermediate perchlorate concentration values according to the scale shown on the figures.

8.5 Results and Discussion

In order to compare the performance of the groundwater site characterization models utilized herein, two comparison strategies were utilized, namely:

- Comparison using ASE values
- Comparison using contour maps

8.5.1 Comparison Using ASE Values

ASE values obtained for all of the twenty-eight groundwater site characterization models (one-well exclusion cases) are listed in Tables 8.4. By examining ASE values listed in these tables, it can be observed that the exclusion of any well will have a great impact on the ANN model accuracy. Note that if the deviation is negative, then the data associated with the excluded well is considered of low importance. On the other hand, if the deviation is positive, then the data associated with the excluded well is considered of high importance. For example, excluding data of well #114 produces a deviation of -34%, while excluding data of well #162 yields about +130% deviation. These means that the data of well #114 can be considered of lower importance in comparison to data associated with well #162. Accordingly, the network that produced the least ASE value is the one excluding the well #114 data, while the highest ASE value was

obtained when the data of well # 162 was excluded. In other words, reducing our database by only 11 data sets (data of well #162) produced the worst performing model. Alternatively, reducing our database by 23 data sets (data of well # 114) actually enhanced our modeling ability. Based on this logic, it can be concluded that elimination of data associated with wells #114, 19, 129, 73, 225, & 139 actually increases the accuracy of the profiling networks. Similarly, eliminating data associated with wells # 35, 173, 165, 74 214, 75, 76, 172, 210, 250, 33, 252, 341, 78, 258, 231, and 162 will produce models with lower accuracy performance.

Based on deviation values presented in Table 8.4, we may advocate that data from few wells (such as well #114 or #19) maybe recommended to be taken out of the database. Moreover, data from such well may not need to be collected during future years (i.e., beyond 2005). These wells produce data that seems to corrupt the modeling tasks which is evident from the low ASE values obtained when their data was excluded. In contract to that, data from other wells such as well #162 and #231 seems to contribute significantly to the accuracy of the developed models. This means that their associated data is of high importance to the modeling processes. Accordingly, more data may need to be collected from these wells due to the importance of their data. Therefore, such a study will help in re-allocating future resources where more data can be collected, reduced or even eliminated.

8.5.2 Comparison Using Contour Maps

Contour maps for the Demo 1 site were generated using the Surfer® software program to assist in the visualization of perchlorate contamination from years 2000 through 2005 and the impact of data reduction via exclusion of specific monitoring wells. This was done using the results obtained from the 9-9-1 ANN model and the data banks, as described previously. Contour maps were generated for all four groundwater site characterizations models, described herein.

For a visual comparison, perchlorate pollutant concentration distribution contour maps for the site were generated. The contour maps had (x, y) coordinates as a variables and (z=0) as a constant. Figures 8.2 through 8.7 show perchlorate distributions at z = 0 ft for years 2000 to 2005 by all four models. By creating contour maps for each year, trends in the concentration of perchlorate contamination over time can be tracked easily. Such images also allow us to easily identify contaminated areas above or below the perchlorate regularly limits of 1 part per billion (PPB). These maps are used herein to compare the accuracy of the four (i.e. model 1, 2, and 3) groundwater monitoring site characterizations models described before.

8.6 Discussion

The impact of monitoring wells exclusion/inclusion on perchlorate spatial distribution for the years 2000 through 2005 obtained via models 1, 2, and 3 will be discussed herein.

8.6.1 Contour maps generated for year 2000

The contour map generated for Model 1 (baseline model) for year 2000 shows highest concentration levels of perchlorate contamination at x = 368250, y = 4616900 to about x = 370000, y = 4616600. The contour maps generated for Model 2 for the year 2000 (Figure 8.2b) show the highest concentration levels of perchlorate contamination at x = 367500, y = 4616950 to about x = 368000, y = 4617000. The contour maps generated for Model 3 for the year 2000 (Figure 8.2c) show the highest concentration levels of perchlorate contamination at x = 367500, y = 4617000 to about x = 369500, y = 4617000.

Comparing the Model 1 baseline contour map (Figure 8.2a) to the Model 2 contour map (Figure 8. 2b) reveals an obvious difference. The map for Model 2 shows only a trace of contamination in the upper northwest corner of the site, while the baseline area of high

contamination (over 50 PPB) extends almost the entire length of the site (east to west) and halfway down the site toward the south. Not only are the actual levels in Model 1 much greater in general, the high concentrations in Model 2 do not coincide with those in Model 1. When comparing the Model 1 baseline contour map (Figure 8.2a) to the Model 3 contour map (Figure 8.2c), the areas of high concentration do not overlap. In fact, Model 3 contains a large band of contamination that begins in the northeast corner of the site and extends toward the middle of the site. This shows the contamination to be on the opposite site of where the actual contamination is located.

8.6.2 Contour maps generated for year 2001

The contour map generated for Model 1 (baseline model) for year 2001 show the highest concentration levels of perchlorate contamination at x = 367750, y = 4617000 to about x = 370500, y = 4616700. The contour maps generated for Model 2 for the year 2001 (Figure 8.3b) show the highest concentration levels of perchlorate contamination at x = 368000, y = 4616400 to about x = 363720, y = 4617000. The contour maps generated for Model 3 for the year 2001 (Figure 8.3c) show highest concentration levels of perchlorate contamination at x = 369600, y = 4616600 to about x = 371500, y = 4617000.

Comparing the Model 1 baseline contour map for the year 2001 (Figure 8.3a) to the Model 2 contour map for the year 2001(Figure 8.3b) once again reveals an obvious difference. The map for Model 2 shows a band of contamination beginning in the upper northeast corner of the site that extends toward the middle of the site, while the actual (baseline) area of high contamination (over 50 PPB) extends ¾ of the length of the site from east to west while gradually extending down at a diagonal toward the southwest corner of the site. The area of contamination on Model 1 is approximately half of the entire site. Not only are the levels in

Model 1 are much greater in general, but the high concentration areas in Model 2 do not coincide with those depicted in Model 1.

When comparing the Model 1 baseline contour map (Figure 8.3a) to the Model 3 contour map (Figure. 8.3c), once again there are similar differences. Model 3 contains a large band of contamination that begins in the northeast corner of the site and extends toward the middle of the site. This shows the contamination to be on the opposite side of the site where the actual contamination is located.

8.6.3 Contour maps generated for year 2002

The contour maps generated for Model 2 for the year 2002 (Figure8.4b) show highest concentration levels of perchlorate contamination at x = 369900, y = 4616950 to about x = 371500, y = 4617000. The contour maps generated for Model 3 for the year 2002 (Figure 8.4c) show highest concentration levels of perchlorate contamination at x = 370000, y = 4616700 to about x = 372000, y = 4617000. The contour maps generated for Model 4 for the year 2002 (Figure 8.4d) show highest concentration levels of perchlorate contamination at x = 369800, y = 4616600 to about x = 371500, y = 4617000.

A comparison of the Model 1 contour map for the year 2002 (Figure 8.4a) with the Model 2 contour map for the year 2002 (Figure 8.4b) reveals some similarities. The map for Model 2 shows only a small band of contamination that begins in the upper northwest corner of the site and extends toward the middle of the site, while the model 1 high contamination area is identified as a small band that begins in the northeast upper region and extends to the center of the site. The contamination does not extend all the way to the northeast corner of the site as in Model 2. In addition, a small area of contamination is identified in the upper northwestern area of the site by Model 1 that is not present in any of the other models.

When comparing the Model 1 baseline contour map (Figure 8.4a) to the Model 3 contour map (Figure. 8.4c), the areas of high concentration are similar but do not completely overlap. As in Model 2, a small band of contamination begins in the upper northeast corner of the site and extends toward the center. This band is slightly wider than the one identified in Model 2. As in Model 2, no other area of contamination is identified (unlike in Model 1).

8.6.4 Contour maps generated for year 2003

The contour maps generated for Model 1 for year 2003 show the highest concentration levels of perchlorate contamination at x = 368250, y = 4616900 to about x = 370000, y = 4616600. The contour maps generated for Model 2 for the year 2003 (Figure 8.5b) show the highest concentration levels of perchlorate contamination at x = 367500, y = 4616950 to about x = 368000, y = 4617000. The contour map generated for Model 3 for the year 2003 (Figure 8.5c) show highest concentration levels of perchlorate contamination at x = 369500, y = 4617000 to about x = 371500, y = 4617000. The contour map generated for Model 4 for the year 2003 (Figure 8.5d) show the highest concentration levels of perchlorate contamination at x = 370500, y = 4616800 to about x = 371500, y = 4617000.

Comparing the Model 1 baseline contour map for the year 2003 (Figure 8.5a) to the Model 2 contour maps for the year 2003 (Figure 8.5b) reveals some similarities. The map for Model 2 shows only a small band of contamination that begins in the upper northwest corner of the site and extends toward the middle, while the baseline area of high contamination is identified as a small band that begins in the northeast upper region and extends to the center of the site. The contamination does not extend all the way to the northeast corner of the site as in Model 2.

When comparing the Model 1 baseline contour map (Figure 8.5a) to the Model 3 contour map (Figure. 8.5c) the areas of high concentration are similar, but do not completely overlap. As in Model 2, a small band of contamination begins in the upper northeast corner of the site and extends toward the center. This band is slightly wider than the one identified in Model 2. The contamination in Model 1 is much more centrally located than that shown in Model 3.

8.6.5 Contour maps generated for year 2004

A comparison of the Model 1 baseline contour map for the year 2004 (Figure 8.6a) to the Model 2 contour maps for the year 2004 (Figure 8.6b) reveals some similarities. The map for Model 2 shows only a small band of contamination that begins in the upper northwest corner of the site and extends toward the middle of the site, while the baseline area of contamination is identified as a small band that begins in the northeast upper region and extends to the center of the site. The contamination does not extend all the way to the northeast corner of the site as in Model 2. The highest concentrations of contamination in Model 2 are located in the uppermost northeast corner of the site, while the areas of highest contamination on Model 1 are located toward the middle of the site.

When comparing the Model 1 baseline contour map (Figure 8.6a) to the Model 3 contour map (Figure. 8.6c) it is evident that the areas of high concentration are again similar but not exact. As in Model 2, a small band of high contamination begins in the upper northeast corner of the site and extends toward the center. Unlike in Model 2, Model 3 has a much higher area of lower concentration contaminants present that extends all the way to the bottom of the site in a diagonal fashion. Although this area of lower concentration of contamination exists in Model 1, it is a much smaller area than the one represented in Model 3.

8.6.6 Contour maps generated for year 2005

Comparing the Model 1 baseline contour map for the year 2005 (Figure 8.7a) to the Model 2 contour maps for the year 2005 (Figure 8.7b) reveals only small similarities. The map for Model 2 shows only a small band of low contamination that begins in the upper northwest corner of the site and extends toward the middle of the site, while the baseline area of contamination is identified as a small band of both higher and lower concentrations of contamination that begins in the northeast upper region and extends to the center of the site. Model 2 does not contain any areas of high contamination.

When comparing the Model 1 baseline contour map (Figure 8.7a) to the Model 3 contour map (Figure. 8.7c), the areas of high concentration are again similar but not exact. Model 3 shows a small band of high contamination that begins in the upper northeast corner of the site and extends toward the center. Unlike in Model 2, Model 3 has a small area of higher concentration contaminants present in the uppermost northeastern corner of the site.

8.7 Conclusion

Contour maps generated from ANN-based models can significantly help in 3-D subsurface site visualization tasks or understand data importance for the MMR-Demo 1 Groundwater Monitoring Network. They allow for better visualization of the underlying contaminant behavior and can help interpret the significance of certain groundwater wells within the MMR Groundwater Monitoring Network. They can be used as an assistive device in the determination of the effectiveness of the existing groundwater network and the need for the addition or removal of groundwater wells.

In the previous chapters, the ANN-based modeling methodology has proven to be an effective means for the prediction of contaminants in groundwater and soil. The research in this

chapter demonstrates how ANN can be used to investigate the data importance for a specific site such as the MMR-Demo1 site.

Using the previously developed 9-9-1 ANN network, an additional thirty-two time-dependent models were developed and compared to baseline model for years 2000 through 2005. These models were developed using all known data minus the data associated with one or more groundwater wells. Contour maps were then generated and compared. After comparing the ASE values and the contour maps of Models 2 through 4 to the baseline contour map (Model 1) it is evident that data from few wells (i.e., wells # 114 & #19) may not be necessary in order to make accurate predictions regarding the extent of perchlorate contamination at the MMR-Demo 1.

When comparing contour maps for Models 2 and 3 for all years with Model 1 baseline maps, obvious differences were noted. In the earlier years, the areas of high contamination predicted via Models 2 or 3 were inverted from the actual area of contamination predicated by model 1. All baseline model maps showed significantly higher levels of contamination than those predicted by Models 2 or 3. This observation supports our belief that even the removal of data from one well may impact the site profile. Sampling errors associated with data obtained from specific wells may negatively impact the prediction accuracy which could lead to millions of dollars in wasted remediation costs. Although ANN is an accurate profiling methodology when provided with sufficient and accurate data, this study demonstrates that few of the MMR-Demo 1 groundwater monitoring wells may not be needed for an accurate site characterization assessment.

Table 8.1 - Network Statistical Accuracy Measure for the Baseline 9-Input Model (Model 1)

9-Inputs Network			MARE%			R^2			ASE		
Stages	Itr	HN	Training	Testing	Validations	Training	Testing	Validations	Training	Testing	Validations
stage I-A	1000	9	116	136	NA	0.434	0.63	NA	0.0042	0.0018	NA
stage I-B	1000	9	116	NA	124	0.434	NA	0.28	0.0042	NA	0.0046
stage II	1000	9	139	NA	NA	0.607	NA	NA	0.0025	NA	NA

HN: Optimal Number of Hidden Nodes
MARE: %Mean Absolute Relative Error
NA: Not Applicable

Itr: Iterations
R^2: Coefficient of determination

$$\text{ASE}=\frac{\sum(\text{Predicted}-\text{Actual})^2}{\#\ of\ data\ sets}$$

Table 8.2 - Network Statistical Accuracy Measures for Model 2

Model 2 Excluding well 165			MARE%			R^2			ASE		
Stages	Itr	HN	Training	Testing	Validations	Training	Testing	Validations	Training	Testing	Validations
stage I-A	1000	7	121	139	NA	0.376	0.355	NA	0.0039	0.00189	NA
stage I-B	1000	7	121	NA	143	0.376	NA	0.324	0.0039	NA	0.0039
stage II	1000	7	123	NA	NA	0.501	NA	NA	0.00296	NA	NA

Table 8.3 - Network Statistical Accuracy Measures for Model 3

Model 3 Excluding well 32			MARE%			R^2			ASE		
Stages	Itr	HN	Training	Testing	Validations	Training	Testing	Validations	Training	Testing	Validations
stage I-A	1000	8	153	183	NA	0.446	0.567	NA	0.0041	0.0021	NA
stage I-B	1000	8	153	NA	174	0.446	NA	0.267	0.0041	NA	0.0053
stage II	1000	8	170	NA	NA	0.601	NA	NA	0.0026	NA	NA

Table 8.4 - Network Statistical Accuracy Measures for ANN-Based Profiling Models.

Well Excluded	MARE%	R^2	ASE	Number of data Sets	% Deviation in ASE from Baseline model (Model 1)
Baseline Model	139	0.607	0.0025	NA	0%
114	124.9	0.7403	0.00166	23	-34%
19	105.9	0.63429	0.00181	8	-28%
129	110.1	0.75209	0.0019	32	-24%
73	114.1	0.69979	0.00197	7	-21%
225	106.5	0.75184	0.00208	9	-17%
139	105.3	0.72045	0.00209	18	-16%
211	117.3	0.65601	0.00229	13	-8%
34	112.2	0.63787	0.00238	30	-5%
32	170	0.601	0.0026	24	4%
31	119.4	0.5836	0.00282	26	13%
35	152.6	0.56111	0.00288	4	15%
173	184.3	0.55706	0.00295	13	18%
165	123	0.5017	0.00296	20	18%
74	127.5	0.5529	0.00299	11	20%
214	198	0.54355	0.00299	5	20%
75	135.7	0.54431	0.00304	22	22%
76	136.2	0.54433	0.00309	33	24%

Well Excluded	MARE%	R^2	ASE	Number of data Sets	% Deviation in ASE from Baseline model (Model 1)
172	136.7	0.53071	0.00313	12	25%
210	125.044	0.4977	0.003877	25	55%
250	125.401	0.47629	0.003894	31	56%
33	141.5	0.59699	0.00397	30	59%
252	118.325	0.46988	0.004041	23	62%
341	128.908	0.49049	0.004141	12	66%
78	152.736	0.44467	0.004163	16	67%
258	149.632	0.38847	0.004387	33	75%
231	155.208	0.38529	0.004675	29	87%
162	99.9	0.08392	0.00576	11	130%

$$\% \text{ Deviation from baseline model} = \left| \frac{ASE - 0.0025}{0.0025} \right| \times 100$$

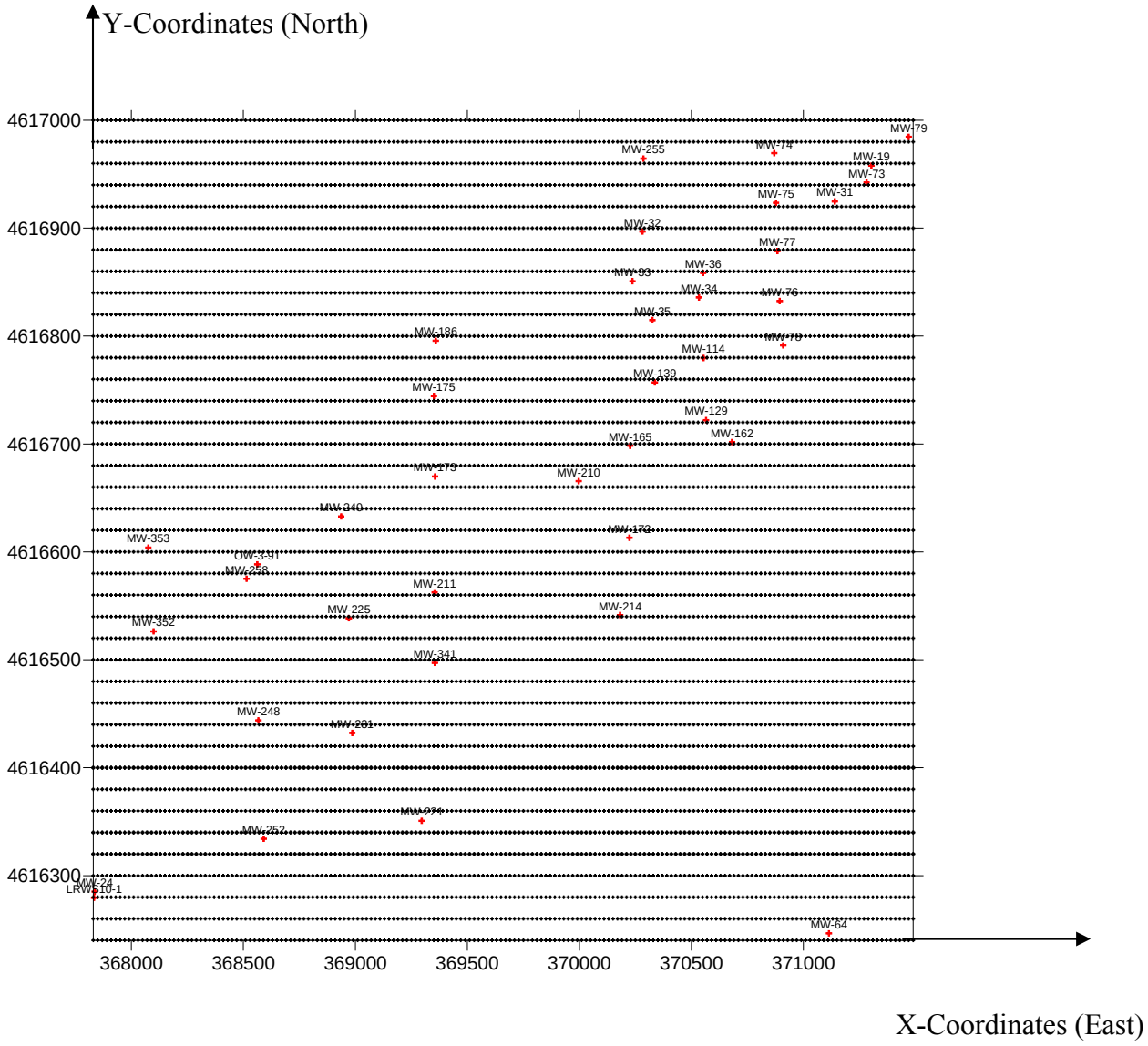


Figure 8.1. MMR-Demo 1 groundwater monitoring network

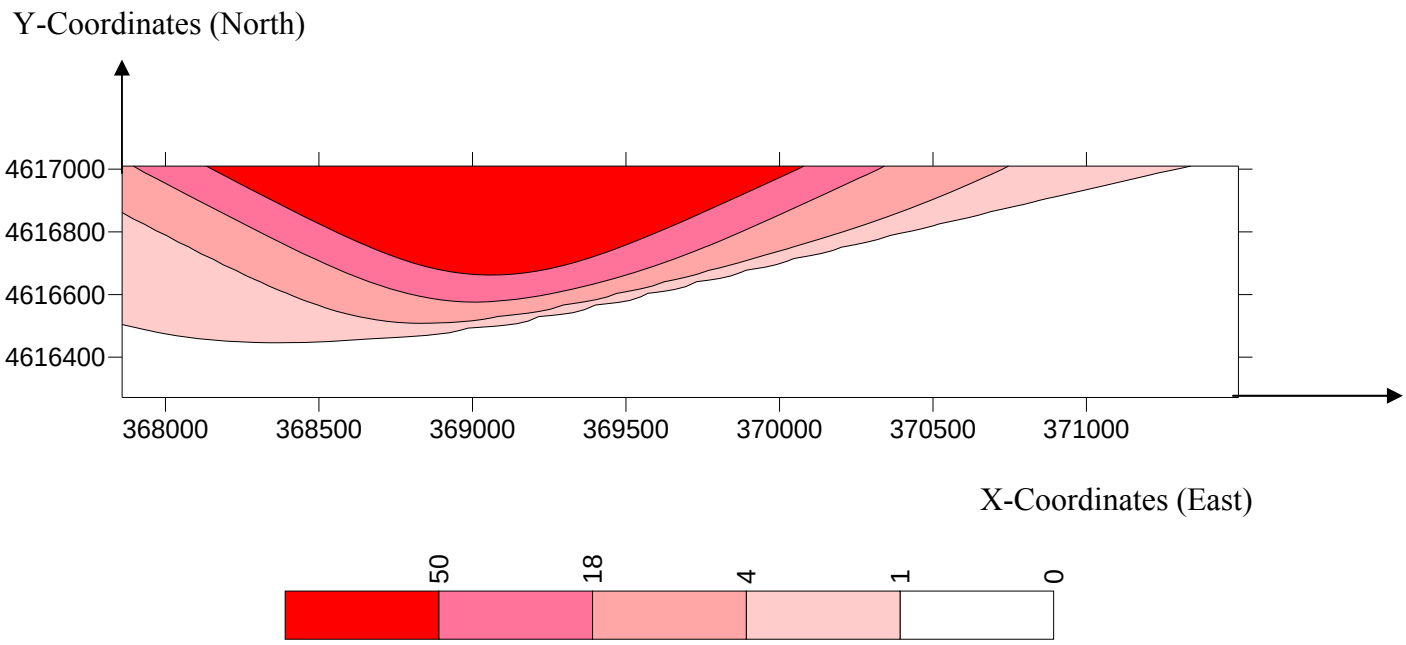


Figure 8.2a. Baseline contour map showing the distribution of perchlorate in groundwater in 2000 at Z =0 ft **(Model 1)**

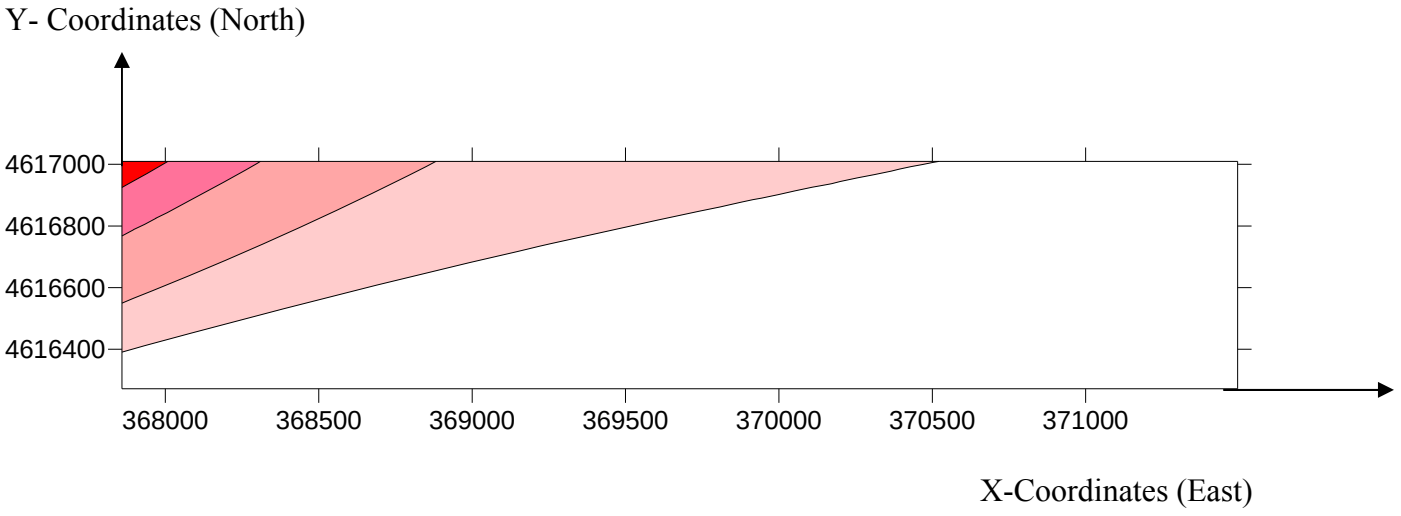


Figure 8.2b. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z =0 ft and well number 165 excluded **(Model 2)**

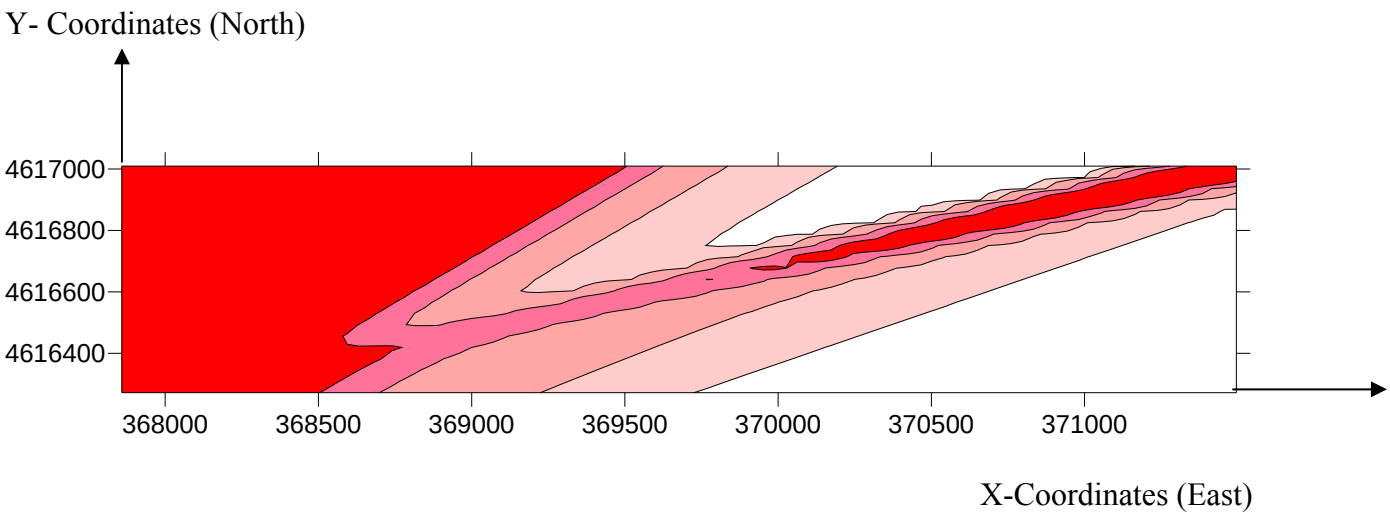


Figure 8.2c. Contour map showing the distribution of perchlorate in groundwater in 2000 at Z =0 ft and well number 32 excluded **(Model 3)**

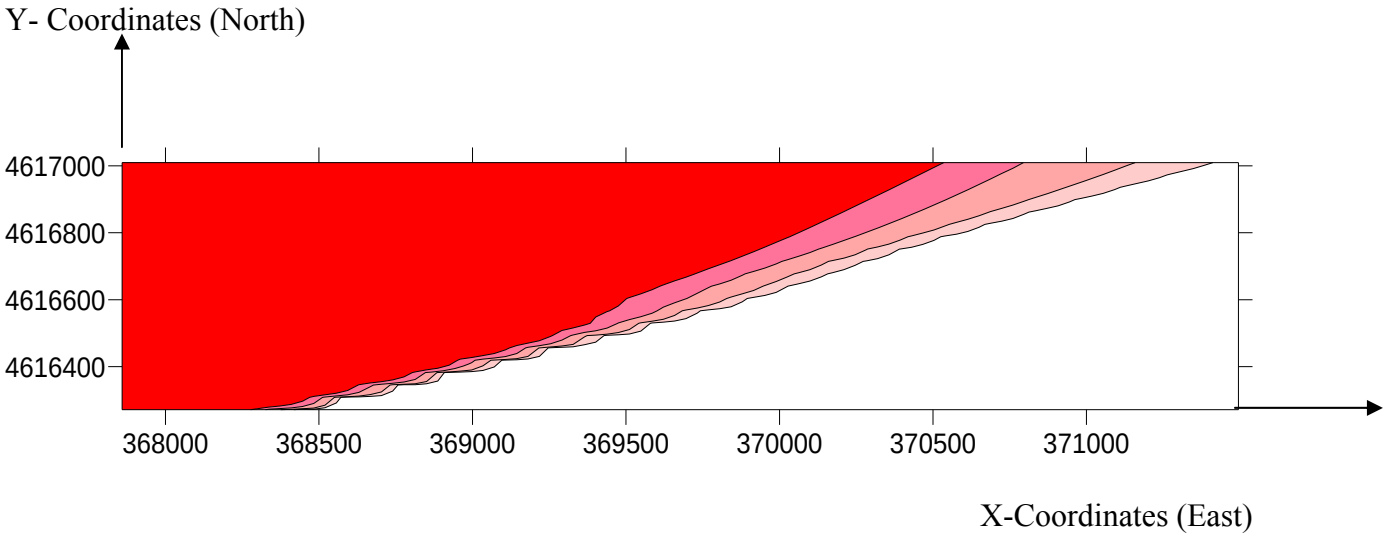


Figure 8.3a. Baseline Contour map showing the distribution of perchlorate in groundwater in 2001 at Z = 0 ft **(Model 1)**

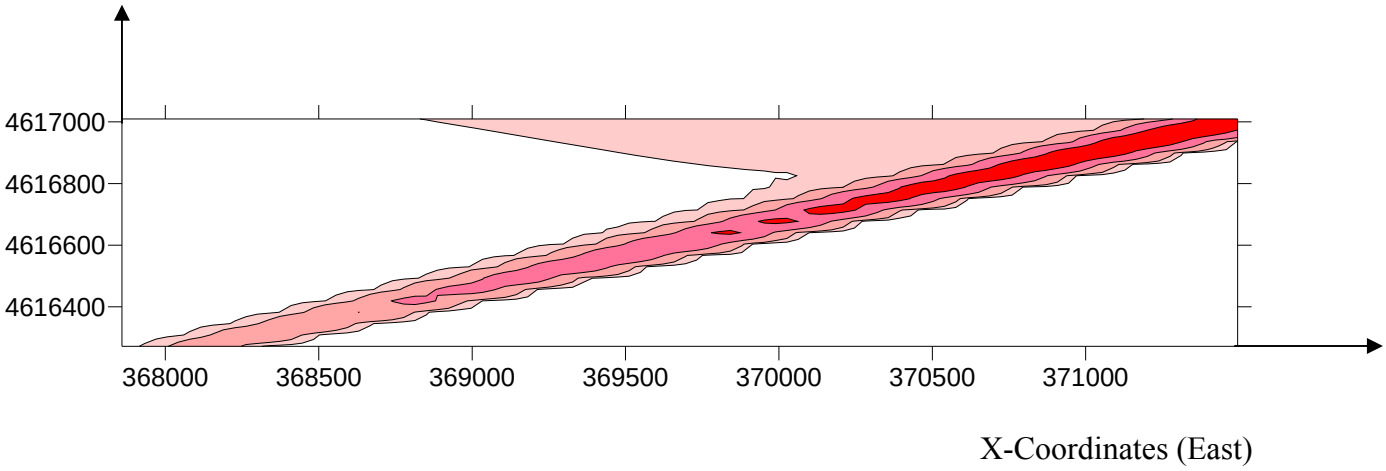


Figure 8.3b. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z =0 ft and well number 165 excluded **(Model 2)**

Y- Coordinates (North)

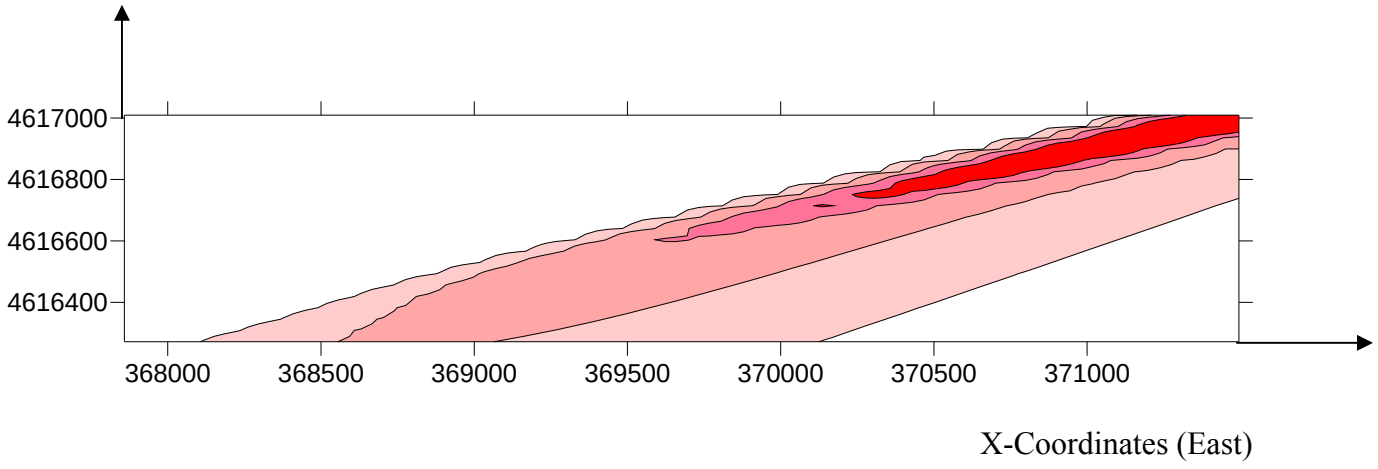


Figure 8.3c. Contour map showing the distribution of perchlorate in groundwater in 2001 at Z =0 ft and well number 32 excluded **(Model 3)**

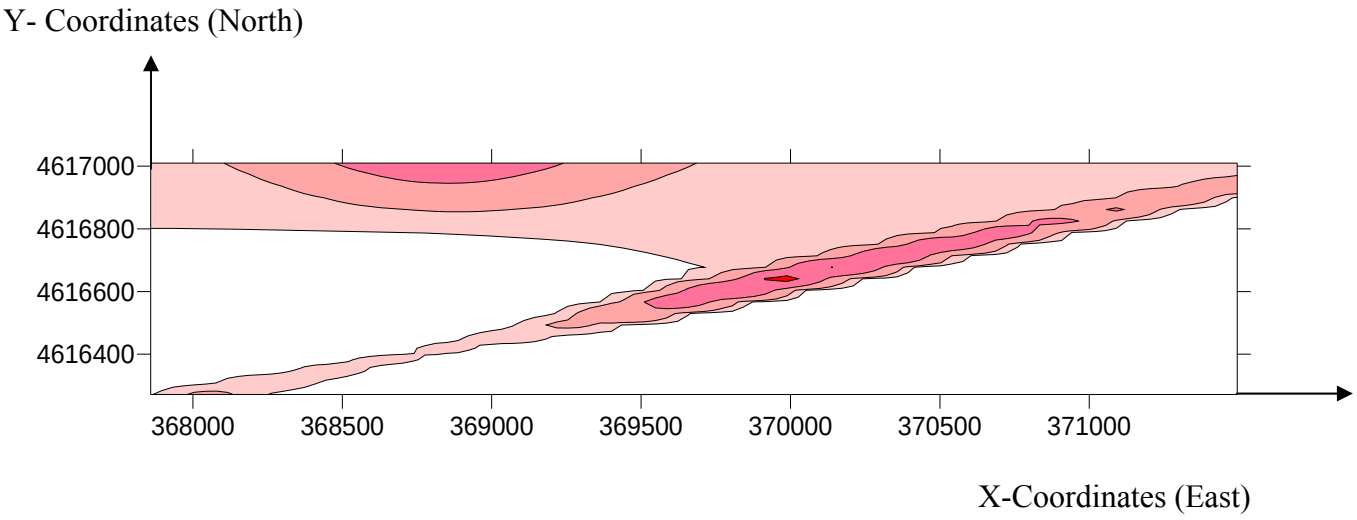


Figure 8.4a. Baseline Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 0 ft. **(Model 1)**

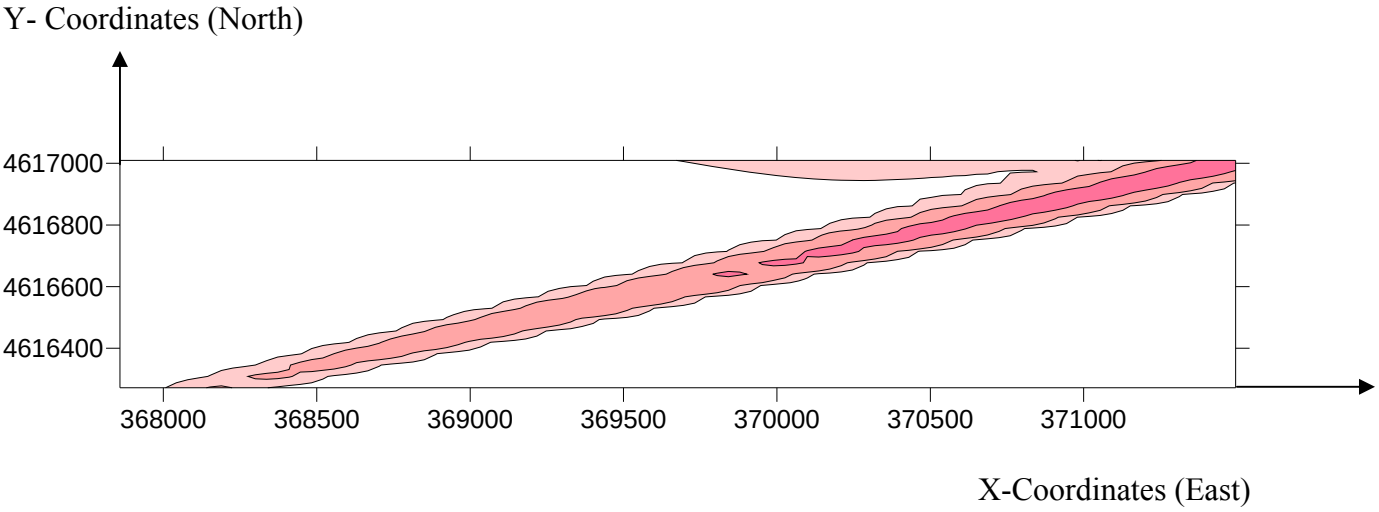


Figure 8.4b. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 0 ft and well number 165 excluded **(Model 2)**

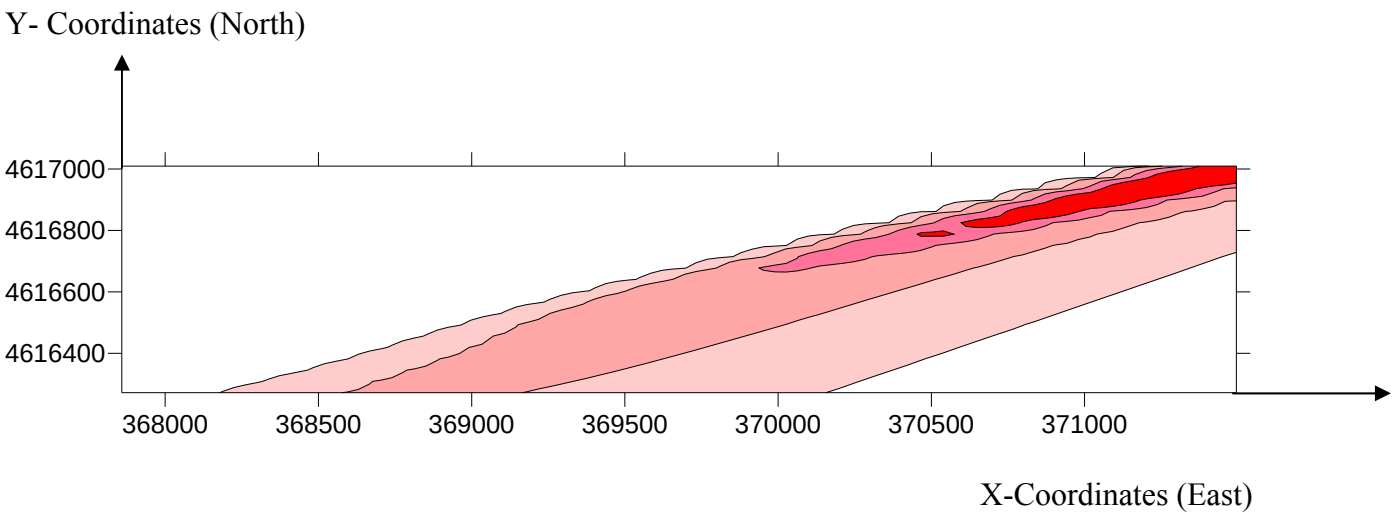


Figure 8.4c. Contour map showing the distribution of perchlorate in groundwater in 2002 at Z = 0 ft and well number 32 excluded **(Model 3)**

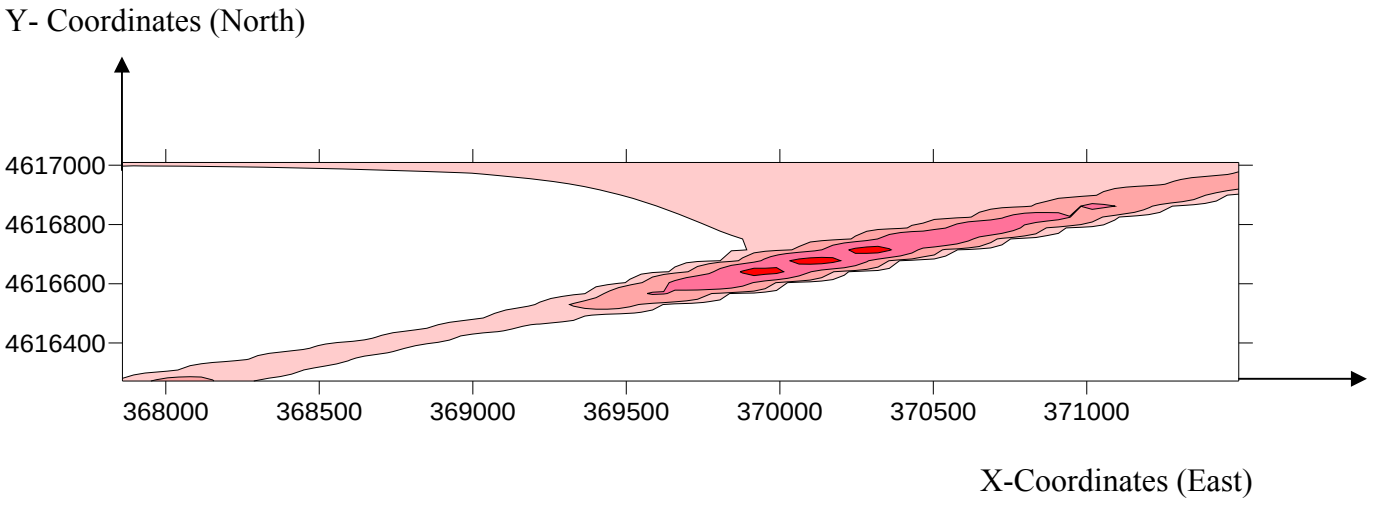


Figure 8.5a. Baseline contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 0 ft. **(Model 1)**

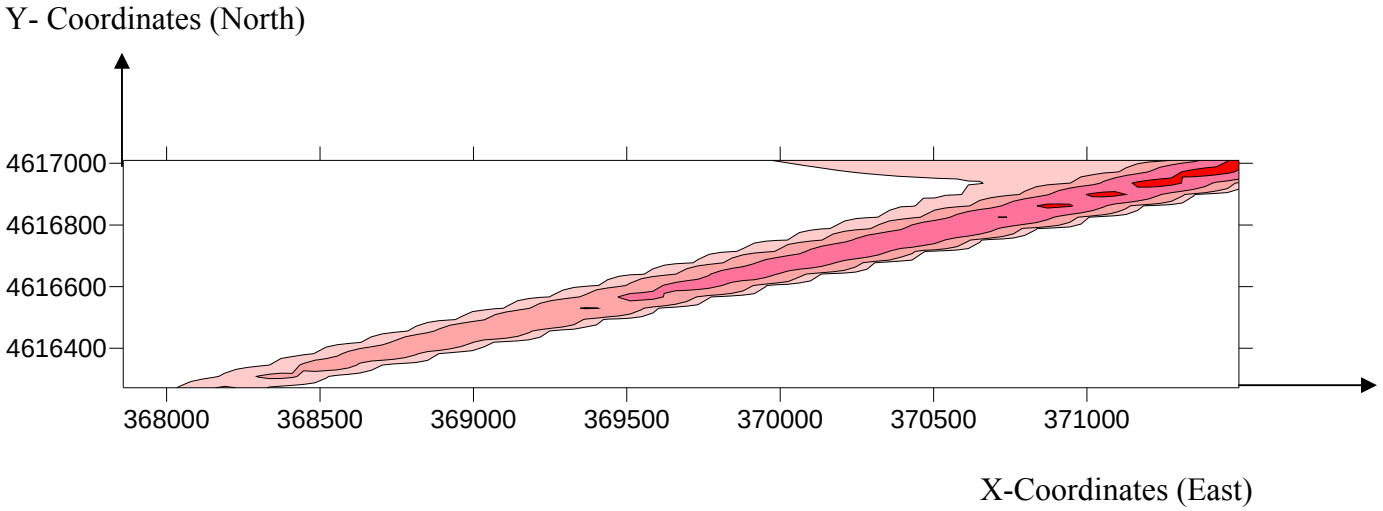


Figure 8.5b. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 0 ft and well number 165 excluded **(Model 2)**

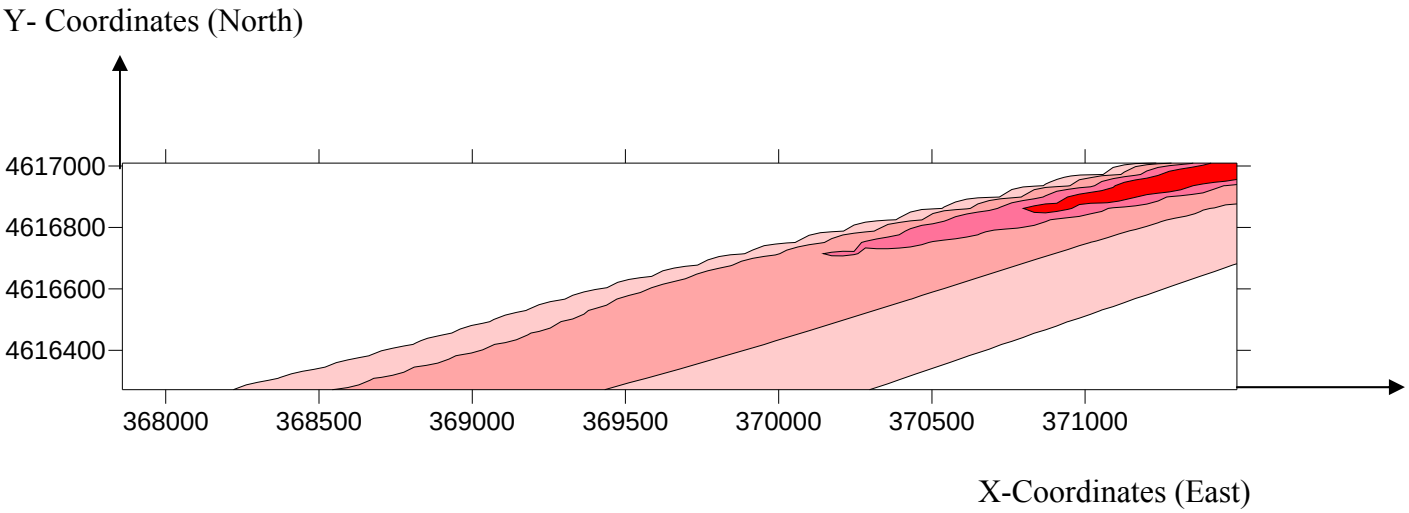


Figure 8.5c. Contour map showing the distribution of perchlorate in groundwater in 2003 at Z = 0 ft and well number 32 excluded **(Model 3)**

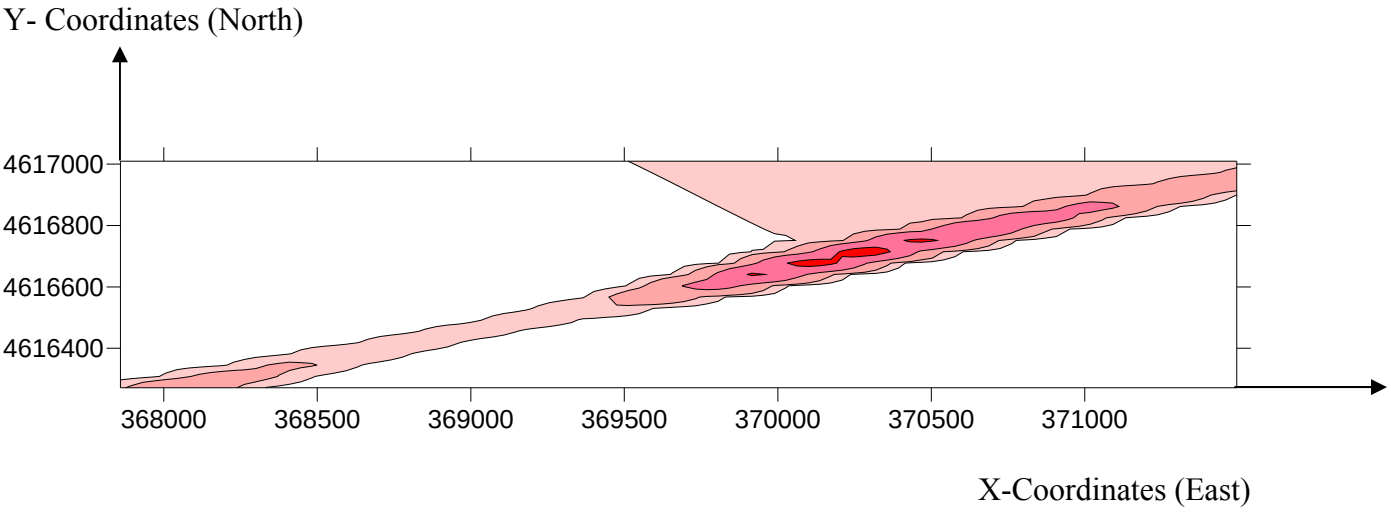


Figure 8.6a. Baseline contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 0 ft. **(Model 1)**

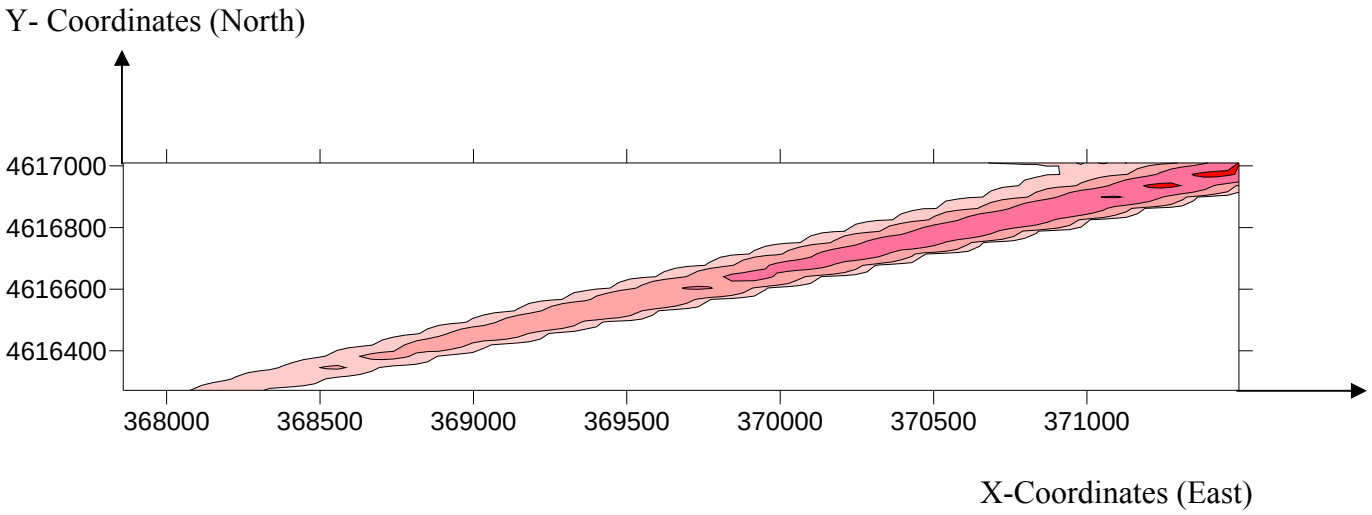


Figure 8.6b. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 0 ft and well number 165 excluded **(Model 2)**

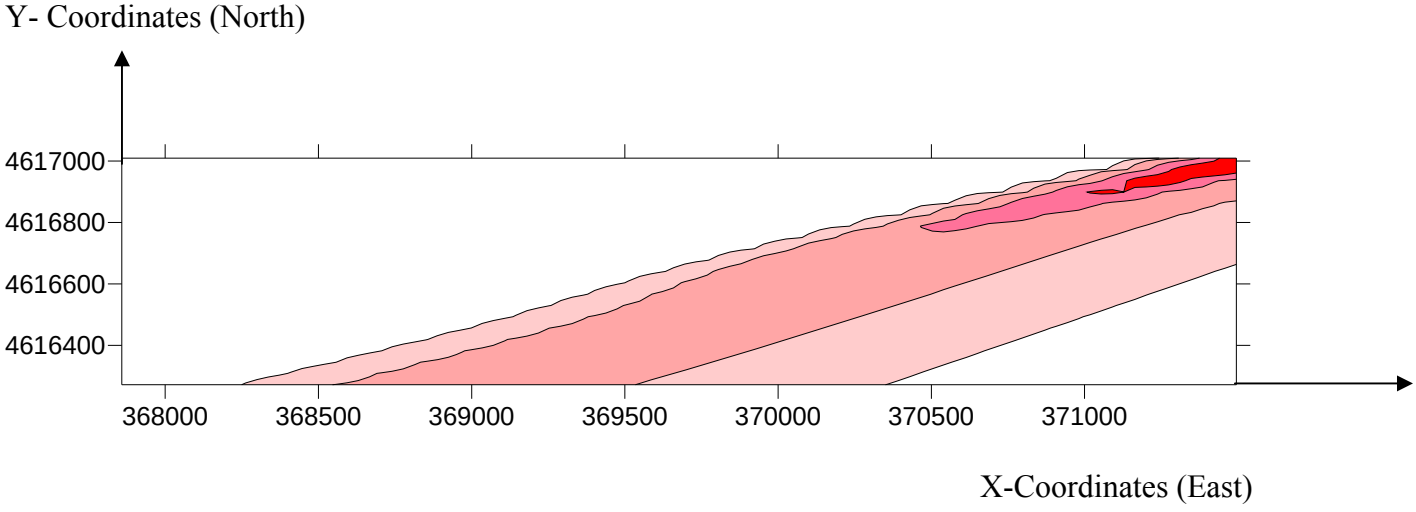


Figure 8.6c. Contour map showing the distribution of perchlorate in groundwater in 2004 at Z = 0 ft and well number 32 excluded **(Model 3)**

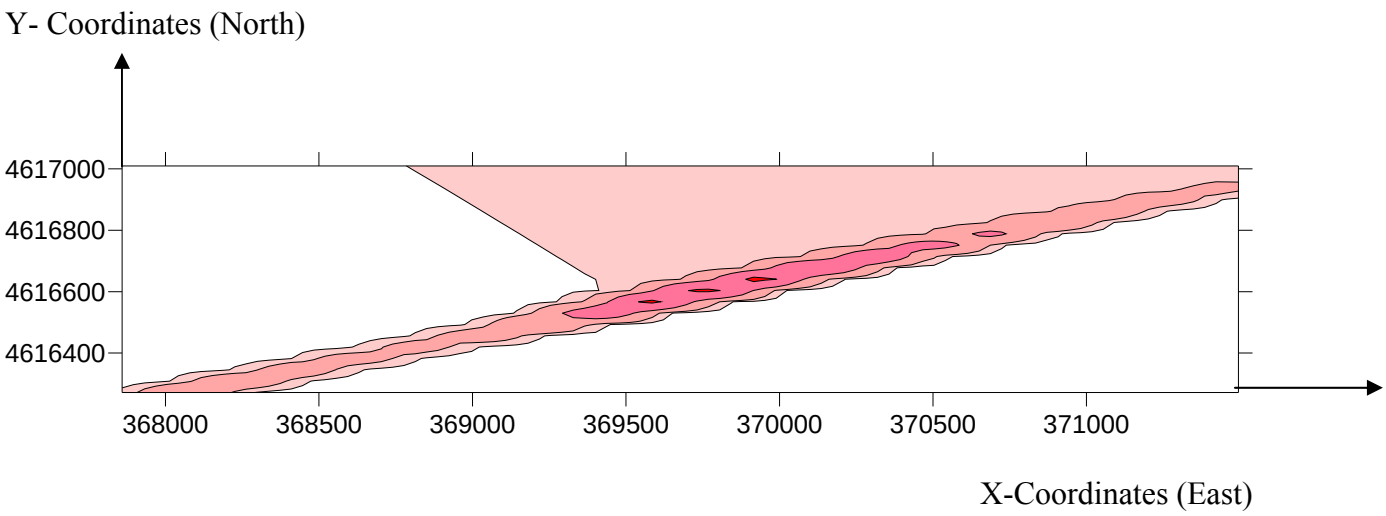


Figure 8.7a. Baseline Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 0 ft **(Model 1)**

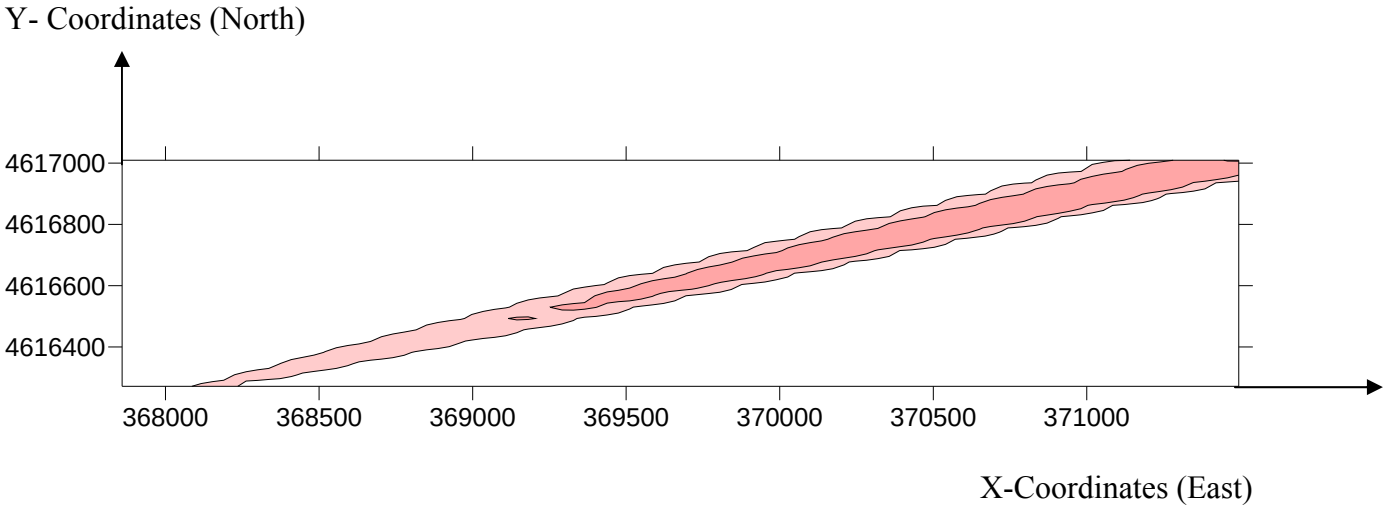


Figure 8.7b. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 0 ft and well number 165 excluded **(Model 2)**

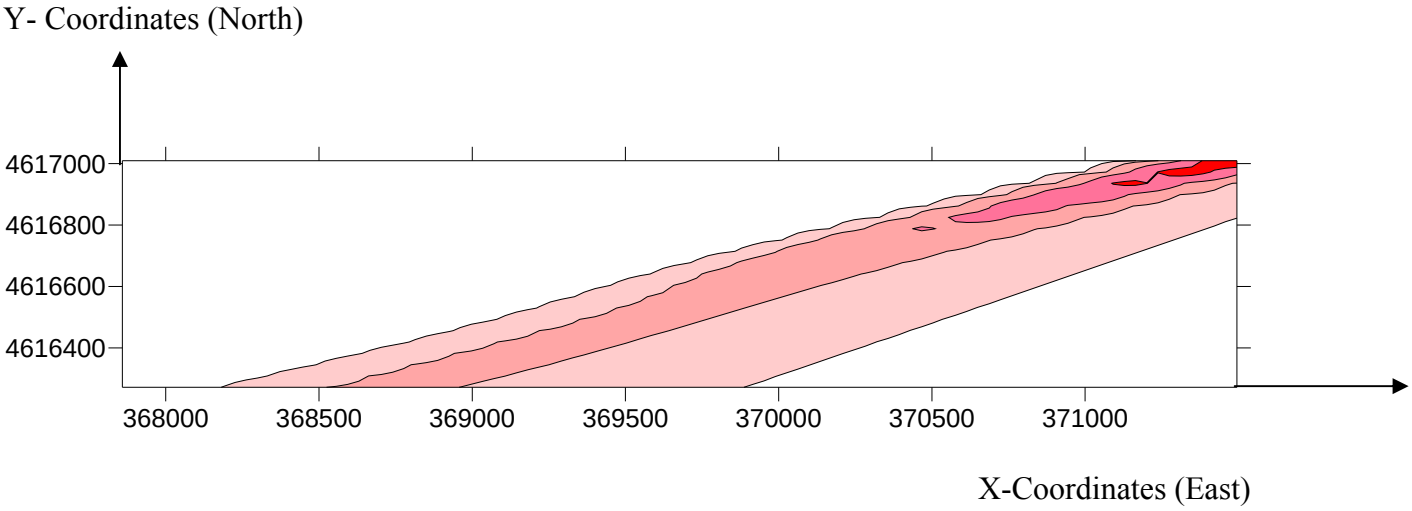


Figure 8.7c. Contour map showing the distribution of perchlorate in groundwater in 2005 at Z = 0 ft and well number 32 excluded **(Model 3)**

CHAPTER 9 - **Summary, Conclusions and Recommendations**

9.1 Summary

This study explored the potential use of ANNs for profiling and characterization of various environmental sites. It investigates the following environmental site profiling cases:

1. Two-dimensional and three-dimensional characterizations of a hypothetical data-rich site by various profiling methods
2. Two-dimensional characterizations of the inorganic materials (lead and copper) in soil and groundwater at a landfill site
3. Three-dimensional, time-related profiling of explosive-related contaminants (perchlorate) at the Massachusetts Military Reservation site

When examining the performance of various site profiling methodologies for a comparative analysis, a static ANN with back-propagation algorithm was used to model the environmental containment at a hypothetical data-rich contaminated site. The performance of the ANN profiling model was then compared to the following profiling models: Inverse Distance to a Power, Kriging, Minimum Curvature, Modified Shepard's, Nearest Neighbor, Polynomial Regression, Radial Basis Function, and Local Polynomial. The comparison showed that the ANN-based models proved to yield the lowest RMSE values in the 2-D and 3-D comparison cases (%RMSE is 19.1% for the 25-ft interval, 3.7% for the 10-ft interval, and 6.4% for the 3-D case). The ANN-based profiling models also produced the best contaminant distribution contour maps when compared to the actual maps. Along with the fact that ANN is the only profiling methodology that allows for efficient 3-D profiling, this study clearly demonstrates that ANN-based methodology, when properly used, has the potential to provide the most accurate

predictions and site profiling contour maps for a contaminated site. Its flexibility is demonstrated by its potential to accurately predict values of a certain contaminant parameter at a specific location, when only supplied with *x*, *y,* and *z* (for 3-D cases) coordinates. In addition, the ANN model provided the most reliable predictions about the location and extent of contamination for the hypothetical site.

The environmental site profiling in the comparative study proved the feasibility of ANN-based methodology to profile a contaminated site. Based on this fact, ANN with a back-propagation learning algorithm was utilized in the site characterization of contaminants at the Kansas City landfill. The use of ANN profiling models made it possible to obtain reliable predictions about the location and concentration of lead and copper contamination at the associated Kansas City landfill site. The resulting profiles can be used to determine additional sampling locations, if needed, for both groundwater and soil in any contaminated zones.

Moreover, the extent of the remediation zones can be properly assessed, reducing the associated cost of further sampling and/or remediation. As a result of this research, the site investigating team was able to capitalize on the information that the developed ANN models generated in order to determine the strategic locations for further testing. This increased efficiency reduced the need for sampling and testing in the relatively uncontaminated zones, which translated into significant cost and time savings.

ANN-based profiling models proved to be a successful profiling methodology to accurately predict the amount and distribution of environmental contaminants at the Kansas City landfill site. As a result of these findings, ANN-based models were applied to a complex contaminated site to determine their effectiveness.

Back-propagation networks were developed for the MMR Demo 1 site using the TR-SEQ1, a three-layered ANN training software package developed by Najjar (2001). The purpose of the ANN model was to predict the concentrations of perchlorate at the MMR from appropriate input parameters. To determine the most-appropriate input parameters for this model, three different cases were investigated using nine potential input parameters.

In the initial case, all nine potential input parameters (X, Y, G, Z, T, R, INJ1, INJ2 and INJ3) were used to develop the desired perchlorate prediction model. This model produced an ASE value of 0.0025 and R^2 value of 0.607. In the second case, one parameter (G) was eliminated to arrive at eight inputs (X, Y, Z, T, R, INJ1, INJ2 and INJ3). This model produced an ASE value of 0.0030 and R^2 value of 0.57. In the final case, two parameters (R and G) were eliminated to arrive at seven input parameters (X, Y, Z, T, INJ1, INJ2 and INJ3). This model produced an ASE value of 0.0032 and R^2 value of 0.503.

When comparing the three cases, it was observed that the eight inputs case produced an ASE value 20% greater than the nine inputs case while the R^2 value decreased by 6.5%. The seven inputs case produced an ASE value 28% greater than the nine inputs case while the R^2 decreased by 17%.

Although the finding for seven inputs case and eight inputs case were somewhat comparable, the nine inputs case model outperformed the seven and the eight inputs case models, trials therefore identifying it as the optimal ANN model for this study. It was determined that the optimal network model for the MMR perchlorate prediction model contained nine input parameters, nine hidden node, and one output parameter (9-9-1).

The ANN modeling used in this case demonstrates the neural network's ability to accurately predict perchlorate contamination using multiple variables. When comparing the

trends observed using the ANN-generated data and the actual trends identified in the MMR 2006 System Performance Monitoring Report, both agree that perchlorate levels are decreasing due to the use of the ETR systems. This proves that the ETR systems were both effective and necessary for the removal of perchlorate contamination at the Demo 1 site, as demonstrated in the contour maps.

Using the knowledge obtained from the MMR perchlorate prediction model, a similar ANN with a back-propagation learning algorithm was developed to model the data importance at the Massachusetts Military Reservation site. In various testing trials, twenty-eight back-propagation ANN models were developed, which excluded or included certain groundwater monitoring wells. These models were then used to investigate the minimum number of groundwater wells necessary to characterize the Demo 1 site accurately.

After comparing the ASE value and the contour maps of all of the exclusion or inclusion of certain groundwater monitoring wells models to the baseline contour map (Model 1), it is apparent that sampling errors associated with data obtained from specific wells may negatively impact the prediction accuracy which could lead to millions of dollars in wasted remediation costs. Although ANN is an accurate profiling methodology when provided with sufficient and accurate data, this study demonstrates that few of the MMR-Demo 1 groundwater monitoring wells may not be needed for an accurate site characterization assessment.

9.2 Conclusions

This research demonstrates the advantages of ANN site characterization modeling in contrast with traditional modeling. First, no complex mathematical formulations were developed to describe the behavior of the contaminants, and the ANN model was built up simply by training on the available laboratory/analytical data. Second, the trained-ANN model can simulate

new scenarios without the need for any additional laboratory analytical-based information. Third, the developed ANN model is convenient for practical usage by either acting as a standalone simulator or by being implemented into another program (Microsoft Excel spreadsheet or Surfer Program). Fourth, flexibility and generality characterized the generated ANN-based models. Once a decision is made for what networks is to represent a site, this network can be readily used to predict the contaminant values at any desired location—this demonstrates flexibility. The only parameter a trained network needs in order to provide such predictions is the input data vector such as (x, y, z) coordinates of the point at which a prediction is desired. Generality lies in ANN's power to capture the mode of change of a contaminant's parameters based on all available data.

As was noted out in this site characterization study, it is necessary to mention that the accuracy of the developed neural network depends on the accuracy of the database used. If the database contains a significant amount of erroneous data, or if the database is too small to capture the features that the neural network is aimed to predict, the neural network will generate significantly incorrect predictions.

Finally, site characterization has posed an everlasting problem in regard to the accuracy of sampling and testing. This is due mainly to the effect of sample size and the practically unattainable and unknown exact (actual) variability of the contaminants' behavior. In all cases of site profiling, the predicted distribution might be far from the actual distribution (although actual distribution may not be known). Therefore, site characterization using any model involving the neural network approach should not be considered final. In other words, the profiling obtained via such methodology (or any other methodology) should be regarded as a decision-making tool that may lead to conducting a more thorough, but focused sampling strategy.

The major conclusions obtained from the present study can be summarized as follows:

- Characterization task-related uncertainties of site contaminations were curtailed by the use of ANN-based models.
- Use of ANNs in site characterization tasks demonstrates their prevalence over currently used site characterization methods.
- ANN-based models are distinguished by simplicity and flexibility.
- Visualization of the site by means of contour maps can be achieved after employing the database generated from ANN-based models.
- The structure of ANN-based models developed in this study can be used as a guide for future investigation to train and test new models on new data.
- 3-D visualization of the contaminants helps identify contaminated zones and select additional sampling locations, if warranted.
- In order to determine the input domain, ANN can be used to study the effect of each independent input parameter on the output. This capability also could be used to optimize the network design of groundwater monitoring wells.
- The ANN-based models proved to yield the lowest RMSE values in the 2-D and 3-D comparison cases
- The ANN-based profiling models produced the best contaminant distribution contour maps and 45-degree scatter graphs for the 2-D and 3-D profiling cases
- ANN is the only profiling methodology that allows for efficient 3-D profiling
- ANN-based profiling methodology, when properly used, has the potential to provide the most accurate predictions and site profiling contour maps for a contaminated site.

- It is possible to use back-propagation ANN-modeling to accurately predict groundwater and soil contamination using limited known data
- Contour maps generated from ANN-based models can significantly help in the 3-D subsurface site visualization tasks and data importance of the MMR-Demo 1 Groundwater Monitoring Network
- ANN based modeling methodology can be used as an assistive device in the determination of the effectiveness of the existing groundwater network and the need for the addition or removal of groundwater wells.
- This study demonstrates the ease, flexibility and robustness of ANN in modeling complex contaminated sites.
- ANN back-propagation modeling can be used to accurately predict groundwater and soil contamination using limited known data. This allows for the study of not only small sites, but larger areas such as Demo 1.
- ANN methodology provides for more in-depth site characterizations at a lower cost, due to the decreased need for sampling and testing data.

9.3 Recommendations

This study is among a relatively limited number of research efforts looking specifically at the use of ANNs for predicting the behavior of environmental contaminants in site characterization. As a result, it provides an important development in the understanding of the applicability, strengths, and limitations of this modeling approach. Based on results of this study, the following basic recommendations are made for future research on the feasibility of using multi-layer feed-forward ANNs in site characterization modeling:

- Results achieved in this study can form the basis for other researchers to explore and uncover other disguised capabilities for ANN.
- Decisions to include ANN-based models in site characterization tasks prior to selecting the location of samples are crucial.
- Vectors (inputs) used in the ANN-based models generated herein could be modified to include additional parameters to examine their effect on the predicted outputs.
- Additional, similar investigations to further verify this promising approach are needed.
- Dynamic ANN profiling methods need to be explored for their viability in future characterization tasks.

References

Abrahart, R.J. & See, L. (2000). *Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments*. Hydrological Processes. 14, 2157-2172.

Adeli, H. (2001). Neural networks in civil engineering: 1989-2000. *Computer-Aided Civil and Infrastructure Engineering. 16,* 126-142.

Ahlfeld, D. P. (1990). "Two-stage groundwater remediation design." *ASCE Journal of Water Resources Planning and Management.* 116(4): 517-529.

Air Force Center For Excellence. (2004). *Cleaning up the Massachussetts Military Reservation 2004 Annual Report*. Retrieved March 19, 2008, from Installation Restoration Program web site: http://mmr.org/IRP/reports/annual/2004/MMR2004AnnualReport.pdf

Ali, H. & Najjar, Y. (1999). Neuronet-based approach for assessing the liquefaction potential of soils. *Transportation Research Record 1633,* 3–8.

AMEC. (2001). Technical Team Memorandum 01-10, Demolition Area 1 Soil Report. Massachusetts Military Reservation Cape Cod, Massachusetts. June 7, 2001. AMEC Earth and Environmental, Westford, Massachusetts.

AMEC. (2001). Technical Team Memorandum 01-6, Impact Area Groundwater Study Program (IAGWSP), Central Impact Area Groundwater Report (in preparation). AMEC Earth and Environment, Inc., Westford, MA.

Amegashie, F., Shang, J. Q., Yanful, E. K, Ding, W, Al-Martini, S. (2006). *Using complex permittivity and artificial neural networks to identify and classify copper, zinc, and lead contamination in soil.* NRC Research Press web site: http://cgj.nrc.ca.

Amegashie, S, Yanful J.Q., & E.K. D. (2006). "Using complex permittivity and artificial neural networks to identify and classify copper and zinc" *Canadian Geotechnical Journal*. 43: 100-109.

Anderson, D. & McNeil, G. (1992). Artificial Neural Networks Technology Data & Analysis Center for Software, 775 Daedalian Drive, Rome, NY 13441-4909.

Babovic, V. & Bojkov, V. H. (2001). D2K Technical Report D2K TR 0401-1. DHI, Copenhagen.

Basheer, I. (1998). "Neuromechanistic-based modeling and simulation of constitutive behavior of fine-grained soils," *Ph.D. Dissertation, Kansas State University.*

Basheer, I. A. & Najjar, Y. M. (1996). Modeling coliform mortality in waste stabilization ponds. *ASCE-Journal of Environmental Engineering*, *122*(5), 449.

Besaw, L.E.; Rizzo, D.M; Mouser, P.J. (2006). Application of an artificial neural network for analysis of subsurface contamination at the Schuyler Falls Landfill, NY. In: Voinov, A., "Landfill closures spur recycling and composting," Anonymous. *BioCycle*, Apr 1999, p. 18.

Brandhuber, P., & Clark, S. (2005). *Perchlorate Occurrence Mapping*. American Water Works Association, Washington, DC.

Bredehoeft, J. D., & Young, R. A. (1983). "Conjunctive use of groundwater and surface water for irrigated agriculture: risk aversion." *Water Resour. Res.* 19(5): 1111-1121.

Buszewski, B. & Kowalkowski, T. (2006). A new model of heavy metal transport in the soil using nonlinear artificial neural networks. *Environmental Engineering Science, 23*(4), 589–596.

CalEPA (California Environmental Protection Agency). 2002. *Public Health Goal for Perchlorate in Drinking Water*. Office of Environmental Health Hazard Assessment, Pesticide Environmental Toxicology Section.

Chin, Y. S., Szu, K. P., Fen, L. B., & Pen, S. H. (2004). *Twelve different interpolation methods: A case study of Surfer 8.0.* International Society for Photogrammetry and Remote Sensing. Istanbul.

Clausen, J., Robb, J., & Fitzpatrick, B. (2004). Range Sustainability Recommendations Based on Extensive Studies at Camp Edwards. *Battelle Range Sustainment Conference.* New Orleans, LA: Battelle Range Sustainment Conference.

Consortium for Atlantic Regional Assessment. (2005). Cape Cod Case Study. Retrieved March 18, 2008, from Groundwater Issues: http://simlab.uri.edu/cara/geology_groundwater.htm.

Cressie, N.A.C. (1990), The origins of Kriging, *Mathematical Geology*, v. 22, p. 239–252.

Dahlin, T., Garin, H. & Palm, M. (2004). Combined Resistivity Imaging and RCPT for Geotechnical Pre-investigation, *Procs. NGM 2004, Ystad, Sweden*, 18-20 May 2004, 9p.

Davis, John C. (1986). *Statistics and Data Analysis in Geology*, John Wiley and Sons, New York.

Dougherty, D. E., & Marryott, R. A. (1991). "Optimal groundwater management: 1. Simulated annealing." *Water Resour. Res.* 27(10): 2493–2508.

Dowla, F.U. & Rogers, L. (1995). Solving problems in environmental engineering and geosciences with artificial neural networks. The MIT Press.

Draper, N., & Smith, H. (1981). *Applied Regression Analysis.* Wiley-Interscience.

Fausett, L. (1994). *Fundamentals of Neural Networks: Architecture*, Algorithms and Applications, Prentice-Hall, NJ.

Felker, V. (2005). *Characterizing the Roughness of Kansas PCC and Superpave Pavements.* Doctoral dissertation, Kansas State University School of Civil Engineering.

Franke, R. (1982). Scattered Data Interpolation: Test of Some Methods. *Mathematics of Computations*, 181-200.

Franke, R., & Nielson, G. (1980), Smooth Interpolation of Large Sets of Scattered Data, *International Journal for Numerical Methods in Engineering*, v. 15, p. 1691–1704.

Gailey, R. M., Crowe, A. S., & Gorelick, S. M. (1991). Coupled process parameter estimation and prediction uncertainty using hydraulic head and concentration data. In *Advances in Water Resources* 14(5): 301–314.

Golden Software, Inc. (2007). *Surfer User's Guide*.

Grossman, E., Cifuentes, L., & Cozzarelli, I. (2002). Anaerobic methane oxidation in landfill-leachate plume. *Environmental Science & Technology*, 36.

Groundwater Foundation (2006). *Groundwater Contamination Concerns*. Retrieved July 23, 2008, from Groundwater Foundation: http://www.groundwater.org/index.html.

Hajmeer, M., Basheer, I. & Najjar, Y. (1996). The growth of *eschrichia coli 015:H7*—A back-propagation network approach. In Dagli, C. H. et. al. (Eds.), *Intelligent Engineering Systems Through Artificial Neural Networks, 6*, 635–640.

Ham, F. & Kostanic, I. (2000). Principles of Neurocomputing for Science and Engineering, Prentice Hall.

Hanson S.J. (1995). "Backpropagation: Some Comments and Variations," in "*Backpropagation: Theory, Architecture, and Applications,"* Rumelhart D.E. and Yves C. (Eds.), Lawrence Erlbaum Assoc., NJ, pp. 237-271.

Haykin, S. (1994). Neural Networks. A Comprehensive Foundation. New York: Macmillan Publishing Company.

Hegazy, Y. A., Mayne, P. W. & Rouhani, S. (1996). Geostatistical assessment of spatial variability in Piezocone tests. In Shackelford, C. D. et al. (Eds.) *Uncertainty in the Geological Environment: From Theory to Practice. ASCE Geotechnical Special Publication No. 58*: Vol. 1 (pp. 254–268).

Herman, D. & Frankenberger, W. (1999). Bacterial reduction of perchlorate and nitrate in water. *Journal of Environmental Quality, 28*(3), 1018-24.

Hertz, J., Krogh, A., & Palmer, R. G., (1991). *Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing Company, New York, pp. 130–141.

Huang, C. (2006). *On the use of the Artificial Neural Network in Geo-Engineering Applications*. Doctoral dissertation. Kansas State University School of Civil Engineering.

Huang, C., Najjar, Y., & Romanoschi, S. (2006). Characterizing the fatigue life of asphalt concrete. *Intelligent Engineering Systems Through Artificial Neural Networks, 16,* 369-374.

Huang, Y.F., Huang, G.H. & Dong, M.Z. (1992). *Development of an artificial neural network for predicting minimum miscibility pressure in CO2 flooding.* Journal of Petroleum Science and Engineering.

Hunt, J. R., Sitar, N. S., & Udell, K. S. (1988). Nonaqueous phase liquid transport and cleanup: 1. Analysis of mechanisms; 2. Experimental studies. *Water Resour. Res.* 24(8): 247–1270.

Isaaks, E. H. & Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics*, Oxford University Press, New York.

Itani, O. & Najjar, Y. (2000). "3-D modeling of spatial properties via artificial neural networks," *Transportation Research Records***,** No. 1709: 50–59.

Itani, O. & Najjar, Y. M. (2000). "3-D modeling of spatial properties via artificial neural networks," *Transportation Research Board,* No. 1709, pp. 50–59.

Jain A.K., Mao J. & Mohiuddin, K.M. (1996), "Artificial neural networks: A tutorial," *Computer, IEEE,* March 1996, pp. 31–44.

Jakeman, A.J. & Rizzoli, A.E. (eds). Proceedings of the iEMSs Third Biennial Meeting: Summit on Environmental Modeling and Software. International Environmental Modeling and Software Society, Burlington, USA. CD ROM. Internet: http://www.iemss.org/iemss2006/sessions/all.html.

Jenkins, T. F., Pennington, J. C., Ranney, T. A., Berry, T. E., Miyares, P. H., Walsh, M. E., Hewitt, A. D., Perron, N., Parker, L. V., Hayes, C. A., & Wahlgren, E. (2001). *Characterization of Explosives Contamination at Military Firing Ranges.* (ERDC Technical Report TR-01-05). U.S. Army Engineer Research and Development Center, Cold Regions Research and Engineering Laboratory, Hanover, NH.

Johnson, V. M., & Rogers, L. L. (1995). "Location analysis in ground-water remediation using neural networks." *Groundwater* 3(5): 749–758.

Jones, N. L. & Davis, R. J. (1996). Three-dimensional characterization of contaminant plumes. Meeting of the Transportation Research Board, January 7–11, 1996. Washington, D.C.

Journel, A.G. (1989). *Fundamentals of Geostatistics in Five Lessons*, American Geophysical Union, Washington, D.C.

Juang, C.H.; Lin, P; Tso, T. (1997). *Interpretation of in situ test data using artificial neural networks.* Intelligent Information Systems. IIS apos; 97. Proceedings. 8–10, 68–172.

Kansas Department of Health and Environment (KDHE), Bureau of Environmental Remediation (1996). Screening site inspection: Kansas City Armory site. Topeka, Kansas.

KDHE, Bureau of Environmental Remediation (1989). Screening site inspection for the EPA model landfill site. Topeka, Kansas.

Lee, D. T., & Schachter, B. J. (1980). Two algorithms for constructing a Delaunay triangulation, *International Journal of Computer and Information Sciences, 9*(3), 219–242.

Li, J.B., Huang, G.H., Huang Y.F, Chakma, A., & Zeng G. M. (2002). *Neural network modeling of hydrocarbon recovery at petroleum contaminated sites*. The 2002 International Technical Conference on Circuits, Systems, Computers and Communications. Phuket, Thailand, July 16-19, 2002. p. 786-789.

Logan, B., Wu, J. & Unz, R. (2001). Biological perchlorate reduction in high-salinity solutions. *Water Research*, *35*(12), 3034-3038.

Lust, L. (2004). Memorandum for See Distribution. Retrieved March 22, 2008, from U.S. Army Environmental Command web site: http://aec.army.mil/usaec/compliance-p2/perchlorate2004.pdf.

MAARNG, 2000. Immediate Response Action (IRA) Plan, Demolition Area 1 –Massachusetts Military Reservation. Impact Area Groundwater Study Office, MAARNG, Camp Edwards, MA.

Mandavilli, S., Najjar, Y. & Abu-Lebdeh, G. (2005). Modeling crash rates for the Kansas Rural Expressway Network. *Intelligent Engineering Systems through Artificial Neural Networks, 15*, 721-730.

Massachusetts Army National Guard (MAARNG) (2000). Immediate Response Action (IRA) Plan, Demolition Area 1–Massachusetts Military Reservation, DEP Release Tracking

Number 4-15031. Impact Area Groundwater Study Office, MAARNG, Camp Edwards, MA.

Masterson, J. P., & Portnoy, J. W. (2006). Potential changes in ground-water flow and their effects on the ecology and water resources of the Cape Cod National Seashore, Massachusetts. Retrieved March 23, 2008, from The U.S. Geological Survey: http://pubs.usgs.gov/gip/2005/13/pdf/GIP_13.pdf

McKinney, D.C., & Loucks, D.P. (1992). *Network Design for Predicting Groundwater Contamination*, Water Resour. Res., (28) 1, 133–147.

Motzer, W. (2001). Perchlorate**:** problems, detection, and solutions. *Environmental Forensics*, *2*(4), 301-311.

Mryyan, S.A. & Najjar, Y. M. (2005). Investigating the Environmental Impact of an Abandon Landfill. *Intelligent Engineering System Through Artificial Neural Networks, 15.* 751–761.

Mryyan, S.A. & Najjar, Y. M. (2006). *Using Neural Network to Investigate the Environmental Impact of an Abandon Landfill.* World Water and Environmental Resources Congress 2006, pp. 250–262. American Society of Civil Engineering.

Mryyan, S.A. & Najjar, Y. M. (2007). Environmental site profiling: a comparative study. *Intelligent Engineering System Through Artificial Neural Networks, 17,* 61–66.

Najjar, Y. & Basheer, I. (1996). "*A neural network approach for site characterization and uncertainty prediction, in the geological environment: from theory to practice,*" ASCE Geotechnical Special Publication No. 58, Shackelford, C. D. et al. (Editors), Vol. 1, pp. 134–148.

Najjar, Y. & Felker, V. (2003). Modeling the time-dependent roughness performance of Kansas PCC pavements. *Intelligent Engineering Systems through Artificial Neural Networks, 13,* 883–888.

Najjar, Y. (2002). *3-D Subsurface Site Profiling*. Proceedings from the ANNIE 2002 Conference. Missouri University of Science and Technology.

Najjar, Y. & Basheer, I. (1996). "Predicting dynamic response of adsorption columns with neural nets*." Journal of Computing in Civil Engineering,* ASCE. 10(1): 31–39.

Najjar, Y. M. & Basheer, I. A. (1996). "A Neural Network Approach for Site Characterization and Uncertainty Prediction," Uncertainty in the Geological Environment: From Theory to Practice, ASCE GSP No. 58, Vol. 1, pp. 134–148.

Najjar, Y. M. & Basheer, I. A. (1996). Utilizing computational neural networks for evaluating the permeability of compacted clay liners. *Geotechnical and Geological Engineering 14*(3) 193–212.

Najjar, Y. M. (1999). Quick Manual for TR-SEQ1 ANN Training Program. Department of Civil Engineering, Kansas State University, Manhattan, Kansas.

Najjar, Y. M. & Zhang, X. (2000). "Characterizing the 3D Stress-Strain Behavior of Sandy Soils: A Neuro-Mechanistic Approach," *ASCE Geotechnical Special Publication,* No. 96, pp. 43–57.

Najjar, Y. M., & Itani, O. M. (2000). 3-D modeling of spatial properties via artificial neural networks. *Transportation Research Records,* 50–59.

Najjar, Y. M., Basheer, I. A. & McReynolds, R. L. (1996). "Neural modeling of Kansas soil swelling," *Transportation Research Board,* No. 1526, pp. 14–19.

Najjar, Y., Basheer, I., Ali, H., & McRyenolds, R. (1999). Characterizing the swelling potential of soils via neuro-reliability approach. In Dagli, C. H. et al. (Eds.) *Intelligent Engineering Systems Through Artificial Neural Networks, 9,* 1201–1206.

Najjar, Y., Reddi, L. & Basheer, I., (1996). "Site characterization by neuronets: an application to the landfill sitting problem." *Ground Water Journal*, 4, 610–620.

Najjar, Y.M. & Basheer, I.A. (1996). A neural network approach for site characterization and uncertainty prediction. In Shackelford, C.D. et al. (Eds). *Uncertainty in the Geological Environment: From Theory to Practice* (ASCE Geotechnical Special Publication No. 58, Vol. 1, pp. 134–148).

Neilson Research Corporation (2007). *Groundwater Contamination in the United States*. Retrieved August 1, 2008, from http://www.nrclabs.com/PWS/GWcontamination.htm

Nerenberg, R., Rittmann, B., Kawagoshi, Y., Gillogly, T., Lehman, G. & Adham, S. (2003). *Microbial Ecology of a Perchlorate-Reducing, Hollow-Fiber Membrane Biofilm Reactor*. American Society for Microbiology Biofilms 2003 Conference. Victoria, Canada.

Newman, D. (2001). *Geologic History Of Cape Cod, Massachusetts*. Retrieved March 19, 2008, from U.S. Geological Survey web site: http://pubs.usgs.gov/gip/capecod/glacial.html.

Nielsen, D. (2005). *Practical Handbook of Environmental site characterization and Groundwater-Water Monitoring.* The Nielsen Environmental Fields School. Galena, Ohio.

Norman Endocrine Surgery Clinic (2005). *Hypothyroidism: too little thyroid hormone*. Retrieved April 22, 2008, from Endocrine Web: http://www.endocrineweb.com/hypo1.html.

Ogden (1999). Final Field Sampling Plan for Phase II (a) Drilling to Investigate RDX Exceedances for the Camp Edwards Impact Area Groundwater Quality Study.

Massachusetts Military Reservation, Cape Cod, Massachusetts. Ogden Environmental and Energy Services. Westford, MA.

Ogden (1999). Final Work Plan for Phase (Activities, Camp Edwards Impact Area Groundwater Quality Study). Massachusetts Military Reservation Cape Cod, Massachusetts. Ogden Environmental and Energy Services, Westford, Massachusetts.

Ogden (1999). Draft Final COC Identification Process, Camp Edwards Impact Area Groundwater Quality Study, Massachusetts Military Reservation Cape Cod, Massachusetts. Westford Massachusetts.

Pennington, J. C., Brannon, J. M. & Mirecki, J. E. (2002). *Distribution and fate of energetics on DOD test and training ranges.* (Technical Report *2* ERDC TR-02-8). U.S. Army Corps of Engineers Washington, DC.

Pike, J. (2006). Massachusetts Military Reservation and Otis National Guard Base Camp Edwards. Retrieved March 18, 2008, from GlobalSecurity.org: http://www.globalsecurity.org/military/facility/mmr.htm.

Powell, M.J.D. (1990), *The Theory of Radial Basis Function Approximation in 1990*, University of Cambridge Numerical Analysis Reports, DAMTP 1990/NA11.

Rizzo, D. M. & Dougherty, D. E. (1994). "Characterization of aquifer properties using artificial neural networks: neural Kriging." *Water Resources Research,* 30(2): 48–489.

Rizzo, D. M., Lillys, T. P. & Doughety, D. E. (1996). Comparisons of site characterization methods using mixed data. In Shackelford, C. D. et al. (Eds.) *Uncertainty in the Geological Environment: from Theory to Practice*, ASCE geotechnical special publication *58*(1) 167–179.

Rojas, R., *Neural Networks: A Systematic Introduction*, Springer-Verlag, Berlin, Inc.,1996.

Roote, D. S. (2001). *Perchlorate Treatment Technologies*. Pittsburgh, PA: Ground-Water Remediation Technologies Analysis Center.

Rosqvist, H., Ahlin, A., Fourie, D., Röhrs, L. & Bengtsson, M. (2003). *Mapping of leachate plumes at two landfill sites in South Africa using geoelectrical imaging techniques*. Ninth International Waste Management and Landfill Symposium, Environmental Sanitary Engineering Centre, Cagliari, Italy.

Russell, S. J., & Norvig, P., *Artificial Intelligence: A Modern Approach*, Prentice Hall, Inc., 1995.

Sarle, W.S. (1994). "Neural Networks and Statistical Models," Proceedings of the Nineteenth Annual SAS Users Group International Conference, Cary, NC: SAS Institute, pp 1538–1550. (ftp://ftp.sas.com/pub/neural/neural1.ps).

Shepard, D. (1968), A two dimensional interpolation function for irregularly spaced data, *Proc. 23rd Nat. Conf. ACM*, pp. 517–523.

Sibson, R. (1980). A vector identity for the Dirichlet Tesselation, *Math. Proc. Cambridge Phil. Soc.*, v. 87, pp. 151–155.

Sibson, R. (1981), A brief description of natural neighbor interpolation, *Interpreting Multivariate Data*, V. Barnett editor, John Wiley and Sons, New York, pp. 21–36.

Simpson, Patrick K. (1990), *Artificial Neural Systems*, Pergamon Press, New York.

Tabach, E., Lancelot, L., Shahrour, I., & Najjar, Y. (2007). Use of artificial neural network simulation metamodeling to assess groundwater contamination in a road project. *Mathematical and Computer Modeling. 45,* 766–776.

The Interstate Technology & Regulatory Council (ITRC). (2008). Remediation technologies for perchlorate contamination in water and soil. Perchlorate Team, Washington, DC.

U.S. Army Corps of Engineers (USACE), 1999. *Findings, Conclusions, and Recommendations for the Massachusetts Military Reservation*. Falmouth, Massachusetts. United States Army Corps of Engineers Rock Island District. Rock Island, Illinois.

USACE, (2006). *System performance monitoring report, rapid response action system, Demo 1 ground water operable unit*. Concord, Massachusetts.

U.S. Department of Defense (USDOD), (2007). *DoD Perchlorate Handbook*. Department of Defense Environmental Data Quality Workgroup, Washington, DC.

USDOD (2007). Report to the Congress: Perchlorate in the southwestern United States. Retrieved April 22, 2008, from Clu-In Web site: http://clu-in.org/download/contaminantfocus/perchlorate/epa2005_1746.pdf.

U.S. Environmental Protection Agency (EPA) (2005). Basic information: Municipal solid waste. http://www.epa.gov/epaoswer/non-hw/muncpl/facts.htm.

USEPA (1997). USEPA Orders Cease-Fire on Cape Cod Base. Retrieved March 18, 2008, from FedFacs: http://www.epa.gov/compliance/resources/newsletters/civil/fedfac/fedfacs4.pdf.

USEPA (2000). Administration order for response action EPA docket number SDWA-1-2000-0014, USEPA Region 1, in the matter of training range and impact area, Massachusetts Military Reservation. Washington, D.C.

USEPA (2002). Perchlorate Environmental Contamination: Toxicological Review and Risk Characterization. Washington, DC. National Center for Environmental Assessment, Office of Research & Development.

USEPA (2008). Contaminaton Focus: Perchlorate. Retrieved March 18, 2008, from Technology Innovation Program web site: http://cluin.org/contaminantfocus/default.focus/sec/perchlorate/cat/Policy_and_Guidance/

USEPA (2007). Perchlorate. Retrieved March 18, 2008, from Federal Facilities Restoration and Reuse: http://www.epa.gov/fedfac/documents/perchlorate.htm

U.S. Geological Survey (USGS), (2002). *Land and People: Cape Cod - Where do Cape Codders Get Their Water?* Retrieved March 2, 2008, from USGS Learning Web: http://interactive2.usgs.gov/learningweb/textonly/students/landpeople_s_cc_water.htm

USGS (2005). HA 730-M *Surficial Aquifer System, Cape Cod Glacial* Aquifer *Text.* Retrieved March 23, 2008, from Ground Water Atlas of the U.S.: http://capp.water.usgs.gov/gwa/ch_m/M-text2.html

U.S. Government Accountability Office (GAO) (2004). GAO-04-601, DOD Operational Ranges. *More reliable cleanup cost estimates and a proactive approach to identifying contaminaton are needed.* Retrieved March 18, 2008, from the U.S. Government Accountability Office: http://www.gao.gov/htext/d04601.html.

Urbansky, E. T. (2000). *Perchlorate in the environment.* American Chemical Society.

Urbansky, E.T., & Schock, M.R. (1999). Issues in managing the risks associated with perchlorate in drinking water. *Journal of Environmental Management*, *56*(76).

Zhang, H., Burns, M. & Logan, B. (2002). Perchlorate reduction by a novel chemolithoautotrophic, hydrogen-oxidizing bacterium. *Environmental Microbiology*, *4*(10), 570-576.

Zhou, B. & Dahlin, T (2003). Properties and Effects of Measurement Errors on 2D Resistivity Imaging Surveying, *Near Surface Geophysics*, 1(3), 105–117.

Zupan, J. & Gasteiger, J. (1993). *Neural Networks for Chemists: An Introduction*, VCH Publishers, New York, NY.

Appendix A Groundwater Analytical Results

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
19	371303.388	4616957.897	65.94	5	68.14	220	33.05	104	0	0	0
19	371303.388	4616957.897	65.44	5	68.14	342	44.16	12	0	0	0
19	371303.388	4616957.897	65.96	5	68.14	534	74.89	41	0	0	0
19	371303.388	4616957.897	66.14	5	68.14	601	85.01	8.49	0	0	0
19	371303.388	4616957.897	64.74	5	68.14	726	95.97	18.6	0	0	0
19	371303.388	4616957.897	63.4	5	68.14	879	115.28	5.2	0	0	0
19	371303.388	4616957.897	63.14	5	68.14	949	120.69	4.1	0	0	0
19	371303.388	4616957.897	65.42	5	68.14	1613	227.04	1.86	0	0	0
31	371140.0511	4616924.767	65.39	33	32.39	221	33.05	46	0	0	0
31	371140.0511	4616924.767	65.39	15.5	55.13	221	33.05	43	0	0	0
31	371140.0511	4616924.767	64.39	15.5	55.13	342	44.16	30	0	0	0
31	371140.0511	4616924.767	64.94	15.5	55.13	487	67.82	20	0	0	0
31	371140.0511	4616924.767	65.21	33	32.21	508	68.59	19	0	0	0
31	371140.0511	4616924.767	65.57	15.5	55.13	601	85.01	16.2	0	0	0
31	371140.0511	4616924.767	64.14	33	31.14	734	95.97	1.66	0	0	0
31	371140.0511	4616924.767	64.23	15.5	55.13	734	95.97	12.5	0	0	0
31	371140.0511	4616924.767	65.74	33	32.74	842	107.79	2.98	0	0	0
31	371140.0511	4616924.767	65.74	33	32.74	842	107.79	3.04	0	0	0
31	371140.0511	4616924.767	62.95	15.5	55.13	879	115.28	12	0	0	0
31	371140.0511	4616924.767	62.34	33	29.34	949	120.69	10	0	0	0
31	371140.0511	4616924.767	62.41	15.5	55.13	949	120.69	7.2	0	0	0
31	371140.0511	4616924.767	61.69	33	28.69	1049	132.67	5.2	0	0	0
31	371140.0511	4616924.767	61.7	15.5	55.13	1049	132.67	4.9	0	0	0
31	371140.0511	4616924.767	62.5	33	29.5	1182	154.00	1.8	0	0	0
31	371140.0511	4616924.767	62.41	15.5	55.13	1185	154.00	10	0	0	0
31	371140.0511	4616924.767	65.54	33	32.54	1365	188.45	2.9	0	0	0
31	371140.0511	4616924.767	65.53	15.5	55.13	1365	188.45	4.6	0	0	0
31	371140.0511	4616924.767	65.53	15.5	55.13	1365	188.45	5.3	0	0	0
31	371140.0511	4616924.767	64.93	33	31.93	1519	212.49	0.68	0	0	0
31	371140.0511	4616924.767	64.9	15.5	55.13	1519	212.49	7.7	0	0	0
31	371140.0511	4616924.767	64.99	33	31.99	1573	221.79	0.474	0	0	0
31	371140.0511	4616924.767	64.98	15.5	55.13	1592	224.94	5.02	0	0	0
31	371140.0511	4616924.767	64.3	33	31.3	1761	237.67	7.44	7.06	4.58	4.58

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
31	371140.0511	4616924.767	64.3	15.5	55.13	1761	244.92	4.7	7.06	4.58	4.58
32	370281.3369	4616896.876	61.13	80	18.41	842	107.79	0.64	0	0	0
32	370281.3369	4616896.876	61.1	70	-0.94	842	107.79	1.97	0	0	0
32	370281.3369	4616896.876	61.12	52.5	16.64	843	107.83	1.38	0	0	0
32	370281.3369	4616896.876	61.14	80	-8.41	1060	146.06	0.66	0	0	0
32	370281.3369	4616896.876	60.14	70	-0.94	1124	146.06	2.3	0	0	0
32	370281.3369	4616896.876	60.14	70	-18.41	1124	146.06	2.3	0	0	0
32	370281.3369	4616896.876	60.16	52.5	16.64	1124	146.06	2.1	0	0	0
32	370281.3369	4616896.876	60.14	80	-18.41	1185	157.01	0.44	0	0	0
32	370281.3369	4616896.876	56.02	70	-0.94	1185	157.01	2.5	0	0	0
32	370281.3369	4616896.876	60.64	52.5	16.64	1185	157.01	1.5	0	0	0
32	370281.3369	4616896.876	65.76	80	-18.41	1417	195.87	2.2	0	0	0
32	370281.3369	4616896.876	65.71	70	-0.94	1417	196.30	2.6	0	0	0
32	370281.3369	4616896.876	65.71	70	-18.41	1417	196.30	2.8	0	0	0
32	370281.3369	4616896.876	65.16	70	-0.94	1524	212.52	3.93	0	0	0
32	370281.3369	4616896.876	64.42	52.5	16.64	1524	212.52	1.69	0	0	0
32	370281.3369	4616896.876	65.35	80	-18.41	1530	213.12	2.2	0	0	0
32	370281.3369	4616896.876	65.34	80	-18.41	1572	221.79	2.35	0	0	0
32	370281.3369	4616896.876	65.26	70	-0.94	1572	221.79	4.14	0	0	0
32	370281.3369	4616896.876	65.39	52.5	16.64	1572	221.79	1.04	0	0	0
32	370281.3369	4616896.876	65.14	80	-18.41	1676	231.62	4.78	0	0	0
32	370281.3369	4616896.876	64.82	70	-0.94	1676	231.62	4.21	0	0	0
32	370281.3369	4616896.876	64.82	70	-18.41	1677	231.62	4.03	0	0	0
32	370281.3369	4616896.876	65.12	52.5	16.64	1677	231.62	1.26	0	0	0
32	370281.3369	4616896.876	64.12	80	-18.41	1810	252.71	0.71	14.11	12.32	12.32
33	370236.8868	4616850.825	62.15	87.5	-17.95	725	95.97	1.54	0	0	0
33	370236.8868	4616850.825	62.63	70	-0.42	725	95.97	1.38	0	0	0
33	370236.8868	4616850.825	60.73	87.5	-17.95	843	107.83	2.02	0	0	0
33	370236.8868	4616850.825	61.03	70	-0.42	843	107.83	1.72	0	0	0
33	370236.8868	4616850.825	61.04	52.5	17.04	843	107.83	1.72	0	0	0
33	370236.8868	4616850.825	60.4	70	-0.42	950	120.69	2.1	0	0	0
33	370236.8868	4616850.825	60.36	52.5	17.04	950	120.69	1.6	0	0	0
33	370236.8868	4616850.825	59.65	87.5	-17.95	1049	132.67	2.2	0	0	0
33	370236.8868	4616850.825	59.65	87.5	-17.95	1049	132.67	2.2	0	0	0
33	370236.8868	4616850.825	59.66	70	-0.42	1049	135.11	1.9	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
33	370236.8868	4616850.825	59.65	52.5	17.04	1052	135.11	1.6	0	0	0
33	370236.8868	4616850.825	59.8	87.5	-17.95	1132	146.92	3	0	0	0
33	370236.8868	4616850.825	59.92	70	-0.42	1132	146.92	1.7	0	0	0
33	370236.8868	4616850.825	59.84	52.5	17.04	1132	146.92	1.3	0	0	0
33	370236.8868	4616850.825	60.47	87.5	-17.95	1185	157.01	1.6	0	0	0
33	370236.8868	4616850.825	60.54	70	-0.42	1185	157.01	1.5	0	0	0
33	370236.8868	4616850.825	60.49	52.5	17.04	1185	157.01	1.3	0	0	0
33	370236.8868	4616850.825	65.79	87.5	-17.95	1432	196.30	1.1	0	0	0
33	370236.8868	4616850.825	65.82	70	-0.42	1432	197.54	1.1	0	0	0
33	370236.8868	4616850.825	65.76	52.5	17.04	1432	197.54	0.56	0	0	0
33	370236.8868	4616850.825	65.47	87.5	-17.95	1523	212.52	0.89	0	0	0
33	370236.8868	4616850.825	65.5	70	-0.42	1524	212.52	1.06	0	0	0
33	370236.8868	4616850.825	65.48	87.5	-17.95	1570	221.79	0.47	0	0	0
33	370236.8868	4616850.825	65.52	70	-0.42	1572	221.79	0.48	0	0	0
33	370236.8868	4616850.825	65.52	70	-0.42	1572	221.79	0.41	0	0	0
33	370236.8868	4616850.825	65.31	87.5	-17.95	1675	231.62	0.83	0	0	0
33	370236.8868	4616850.825	65.34	70	-0.42	1676	231.62	0.76	0	0	0
33	370236.8868	4616850.825	65.3	52.5	17.04	1676	231.62	0.442	0	0	0
33	370236.8868	4616850.825	64.3	87.5	-17.95	1803	251.66	1.1	13.10	11.22	11.22
33	370236.8868	4616850.825	65.53	70	-0.42	1938	274.50	0.64	32.54	32.55	32.55
34	370534.0266	4616835.848	62.75	58	8.55	222	33.85	56	0	0	0
34	370534.0266	4616835.848	63.39	78	-11.61	352	47.57	109	0	0	0
34	370534.0266	4616835.848	63.35	58	8.55	352	47.57	34	0	0	0
34	370534.0266	4616835.848	63.61	58	8.55	486	67.82	28	0	0	0
34	370534.0266	4616835.848	63.59	78	-11.61	490	67.83	46	0	0	0
34	370534.0266	4616835.848	64.1	58	8.55	576	79.68	16.2	0	0	0
34	370534.0266	4616835.848	64.19	78	-11.61	577	79.68	30.8	0	0	0
34	370534.0266	4616835.848	64.19	78	-11.61	577	79.68	31.4	0	0	0
34	370534.0266	4616835.848	62.99	78	-11.61	725	95.97	17.7	0	0	0
34	370534.0266	4616835.848	62.95	58	8.55	725	95.97	5.8	0	0	0
34	370534.0266	4616835.848	61.46	78	-11.61	844	107.83	7.9	0	0	0
34	370534.0266	4616835.848	61.62	58	8.55	844	107.83	19.6	0	0	0
34	370534.0266	4616835.848	60.99	78	-11.61	962	120.69	7.1	0	0	0
34	370534.0266	4616835.848	60.99	78	-11.61	962	120.69	7.3	0	0	0
34	370534.0266	4616835.848	61	58	8.55	962	120.69	17	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
34	370534.0266	4616835.848	60.3	78	-11.61	1049	132.67	8	0	0	0
34	370534.0266	4616835.848	60.27	58	8.55	1049	132.67	14	0	0	0
34	370534.0266	4616835.848	60.98	78	-11.61	1178	153.90	8	0	0	0
34	370534.0266	4616835.848	60.99	58	8.55	1178	153.90	10	0	0	0
34	370534.0266	4616835.848	62.95	78	-11.61	1411	191.43	6.9	0	0	0
34	370534.0266	4616835.848	63.9	58	8.55	1411	195.84	7.3	0	0	0
34	370534.0266	4616835.848	63.53	78	-11.61	1525	212.61	3.43	0	0	0
34	370534.0266	4616835.848	63.43	58	8.55	1525	212.61	7.02	0	0	0
34	370534.0266	4616835.848	63.46	78	-11.61	1595	224.94	5.28	0	0	0
34	370534.0266	4616835.848	63.62	58	8.55	1595	224.94	5.23	0	0	0
34	370534.0266	4616835.848	63.33	78	-11.61	1678	232.20	3.32	0	0	0
34	370534.0266	4616835.848	63.33	78	-11.61	1678	232.20	3.1	0	0	0
34	370534.0266	4616835.848	63.3	58	8.55	1678	232.20	5.87	0	0	0
34	370534.0266	4616835.848	63.44	78	-11.61	1937	274.39	3.1	32.40	32.39	32.39
34	370534.0266	4616835.848	63.43	58	8.55	1937	274.50	3.9	32.40	32.39	32.39
35	370325.7406	4616814.787	63.08	73	-6.37	489	67.82	4	0	0	0
35	370325.7406	4616814.787	63.78	73	-6.37	580	80.07	5.4	0	0	0
35	370325.7406	4616814.787	62.43	73	-6.37	720	95.17	6.34	0	0	0
35	370325.7406	4616814.787	61.13	73	-6.37	844	107.83	6.44	0	0	0
35	370325.7406	4616814.787	59.79	73	-6.37	1052	135.11	4.2	0	0	0
35	370325.7406	4616814.787	60.8	73	-6.37	1223	158.11	3.9	0	0	0
35	370325.7406	4616814.787	62.61	73	-6.37	1698	235.74	6	0	0	0
36	370552.3802	4616858.569	62.93	59	7.03	738	96.70	1.86	0	0	0
36	370552.3802	4616858.569	62.93	59	7.03	738	96.70	2.16	0	0	0
36	370552.3802	4616858.569	61.85	59	7.03	844	107.83	3.44	0	0	0
36	370552.3802	4616858.569	61.23	59	7.03	950	120.69	4	0	0	0
36	370552.3802	4616858.569	60.47	59	7.03	1052	135.11	4.2	0	0	0
36	370552.3802	4616858.569	61.11	59	7.03	1179	153.90	3.7	0	0	0
36	370552.3802	4616858.569	64.13	59	7.03	1411	195.84	4.8	0	0	0
36	370552.3802	4616858.569	63.68	59	7.03	1522	212.49	3.13	0	0	0
36	370552.3802	4616858.569	63.68	59	7.03	1523	212.52	3.09	0	0	0
36	370552.3802	4616858.569	63.68	59	7.03	1573	221.79	1.9	0	0	0
36	370552.3802	4616858.569	63.38	59	7.03	1676	231.62	2.9	0	0	0
36	370552.3802	4616858.569	63.55	59	7.03	1937	274.50	5.3	32.40	32.39	32.39
73	371281.9525	4616942.268	68.03	5	67.78	353	47.57	6	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
73	371281.9525	4616942.268	68.03	5	67.78	530	72.62	10	0	0	0
73	371281.9525	4616942.268	68.03	5	67.78	741	97.01	3.3	0	0	0
73	371281.9525	4616942.268	65.71	5	67.78	962	120.69	1.9	0	0	0
73	371281.9525	4616942.268	66.01	5	67.78	1365	188.45	3.9	0	0	0
73	371281.9525	4616942.268	65.38	5	67.78	1519	212.49	3	0	0	0
73	371281.9525	4616942.268	65.36	5	67.78	1606	225.40	2.46	0	0	0
74	370870.0287	4616969.526	64.28	36	32.13	845	109.04	0.73	0	0	0
74	370870.0287	4616969.526	62.54	36	32.13	850	109.16	0.45	0	0	0
74	370870.0287	4616969.526	61.77	81	-12.9	1179	153.90	0.49	0	0	0
74	370870.0287	4616969.526	64.56	81	-12.9	1433	197.54	0.9	0	0	0
74	370870.0287	4616969.526	64.61	36	32.13	1433	197.54	0.39	0	0	0
74	370870.0287	4616969.526	64.61	36	32.13	1433	197.54	0.41	0	0	0
74	370870.0287	4616969.526	64.4	81	-12.9	1521	212.49	0.42	0	0	0
74	370870.0287	4616969.526	64.33	36	32.13	1522	212.49	0.39	0	0	0
74	370870.0287	4616969.526	64.13	36	32.13	1676	231.62	0.56	0	0	0
74	370870.0287	4616969.526	62.84	36	32.13	1803	251.66	0.56	13.10	11.22	11.22
74	370870.0287	4616969.526	62.84	36	32.13	1803	251.66	0.55	13.104	11.218	11.218
75	370877.7201	4616923.479	65.15	39	29.36	586	80.25	6.24	0	0	0
75	370877.7201	4616923.479	63.66	39	29.36	737	96.47	4.08	0	0	0
75	370877.7201	4616923.479	62.58	64	4.26	844	107.83	0.57	0	0	0
75	370877.7201	4616923.479	62.57	39	29.36	845	109.04	4.89	0	0	0
75	370877.7201	4616923.479	61.85	39	29.36	961	120.69	2.8	0	0	0
75	370877.7201	4616923.479	61.85	39	29.36	961	120.69	3.2	0	0	0
75	370877.7201	4616923.479	61.15	39	29.36	1052	135.11	3.6	0	0	0
75	370877.7201	4616923.479	61.93	39	29.36	1180	153.90	6.8	0	0	0
75	370877.7201	4616923.479	64.61	64	4.26	1432	197.54	0.61	0	0	0
75	370877.7201	4616923.479	64.54	39	29.36	494	67.83	9	0	0	0
75	370877.7201	4616923.479	64.63	39	29.36	1433	197.54	4.2	0	0	0
75	370877.7201	4616923.479	64.3	64	4.26	1516	212.49	0.37	0	0	0
75	370877.7201	4616923.479	64.36	39	29.36	1516	212.49	3.08	0	0	0
75	370877.7201	4616923.479	64.63	39	29.36	1516	212.49	2.84	0	0	0
75	370877.7201	4616923.479	63.86	64	4.26	1558	218.44	0.5	0	0	0
75	370877.7201	4616923.479	64.24	39	29.36	1558	218.44	2.59	0	0	0
75	370877.7201	4616923.479	64.24	39	29.36	1558	218.44	2.46	0	0	0
75	370877.7201	4616923.479	64.54	39	29.36	494	67.83	9	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
75	370877.7201	4616923.479	64.21	39	29.36	1676	231.62	1.1	0	0	0
75	370877.7201	4616923.479	64.21	39	29.36	1676	231.62	1.1	0	0	0
75	370877.7201	4616923.479	63.15	64	4.26	1803	251.66	0.356	13.10	11.22	11.22
75	370877.7201	4616923.479	63.59	39	29.36	1931	274.35	1.9	31.54	31.44	31.44
76	370894.2912	4616832.476	64.48	63	5.28	492	67.83	8	0	0	0
76	370894.2912	4616832.476	64.53	43	25.36	492	67.83	17	0	0	0
76	370894.2912	4616832.476	64.53	23	45.36	492	67.83	7	0	0	0
76	370894.2912	4616832.476	65.1	23	45.36	587	81.19	13.3	0	0	0
76	370894.2912	4616832.476	65.16	63	5.28	590	84.41	16	0	0	0
76	370894.2912	4616832.476	61.07	63	5.28	1052	135.11	11	0	0	0
76	370894.2912	4616832.476	65.08	43	25.36	590	84.41	22.1	0	0	0
76	370894.2912	4616832.476	65.08	43	25.36	590	84.41	22.5	0	0	0
76	370894.2912	4616832.476	63.48	63	5.28	727	95.97	30.6	0	0	0
76	370894.2912	4616832.476	63.56	23	45.36	727	95.97	41.2	0	0	0
76	370894.2912	4616832.476	63.56	43	25.36	737	96.47	126	0	0	0
76	370894.2912	4616832.476	62.53	63	5.28	844	107.83	15.3	0	0	0
76	370894.2912	4616832.476	64.46	43	25.36	340	44.13	11	0	0	0
76	370894.2912	4616832.476	62.51	43	25.36	844	107.83	174	0	0	0
76	370894.2912	4616832.476	62.52	23	45.36	844	107.83	175	0	0	0
76	370894.2912	4616832.476	61.79	63	5.28	961	120.69	3.1	0	0	0
76	370894.2912	4616832.476	61.8	23	45.36	962	120.69	88	0	0	0
76	370894.2912	4616832.476	61.11	23	45.36	1052	135.11	26	0	0	0
76	370894.2912	4616832.476	61.87	63	5.28	1180	153.9	200	0	0	0
76	370894.2912	4616832.476	65.05	63	5.28	1365	188.45	97	0	0	0
76	370894.2912	4616832.476	65.01	23	45.36	1365	188.45	19	0	0	0
76	370894.2912	4616832.476	64.33	63	5.28	1515	212.46	16.4	0	0	0
76	370894.2912	4616832.476	64.16	43	25.36	1515	212.49	115	0	0	0
76	370894.2912	4616832.476	64.26	23	45.36	1515	212.49	19.1	0	0	0
76	370894.2912	4616832.476	64.35	63	5.28	1572	221.79	17.9	0	0	0
76	370894.2912	4616832.476	64.37	23	45.36	1572	221.79	11.3	0	0	0
76	370894.2912	4616832.476	64.36	43	25.36	1573	221.79	93.1	0	0	0
76	370894.2912	4616832.476	64.08	63	5.28	1679	232.20	47.3	0	0	0
76	370894.2912	4616832.476	64.01	43	25.36	1684	232.20	57.2	0	0	0
76	370894.2912	4616832.476	64.06	23	45.36	1684	232.20	2.11	0	0	0
76	370894.2912	4616832.476	64.05	43	25.36	1929	274.35	25	31.25	31.13	31.13

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
76	370894.2912	4616832.476	64.07	23	45.36	1929	274.35	3.2	31.25	31.13	31.13
76	370894.2912	4616832.476	64.1	63	5.28	1930	274.35	1.6	31.39	31.28	31.28
77	370884.0444	4616878.785	64.41	43	25.91	340	44.13	28	0	0	0
77	370884.0444	4616878.785	64.52	43	25.91	495	67.83	16	0	0	0
77	370884.0444	4616878.785	65.12	43	25.91	587	81.19	13.9	0	0	0
77	370884.0444	4616878.785	63.67	103	-34.13	725	95.97	0.44	0	0	0
77	370884.0444	4616878.785	63.81	43	25.91	725	95.97	12.3	0	0	0
77	370884.0444	4616878.785	62.58	43	25.91	844	107.83	8.01	0	0	0
77	370884.0444	4616878.785	61.96	43	25.91	949	120.69	7.2	0	0	0
77	370884.0444	4616878.785	61.1	43	25.91	1052	135.12	7.2	0	0	0
77	370884.0444	4616878.785	61.91	43	25.91	1180	153.90	5.4	0	0	0
77	370884.0444	4616878.785	65.06	103	-34.13	1365	188.45	0.81	0	0	0
77	370884.0444	4616878.785	65.08	43	25.91	1365	188.45	9.1	0	0	0
77	370884.0444	4616878.785	64.5	43	25.91	1503	212.38	5.32	0	0	0
77	370884.0444	4616878.785	64.23	43	25.91	1553	218.01	5.7	0	0	0
77	370884.0444	4616878.785	64.1	43	25.91	1670	231.16	5.1	0	0	0
77	370884.0444	4616878.785	64.1	43	25.91	1670	231.16	5.1	0	0	0
77	370884.0444	4616878.785	64.25	43	25.91	1936	274.39	7	32.26	32.23	32.23
78	370909.8458	4616791.326	64.37	43	25.67	340	44.13	19	0	0	0
78	370909.8458	4616791.326	64.81	43	25.67	592	84.42	11.4	0	0	0
78	370909.8458	4616791.326	63.71	63	5.91	726	95.97	0.4	0	0	0
78	370909.8458	4616791.326	63.67	43	25.67	727	95.97	4.43	0	0	0
78	370909.8458	4616791.326	57.58	63	5.91	845	109.04	2.07	0	0	0
78	370909.8458	4616791.326	62.55	43	25.67	845	109.04	4.75	0	0	0
78	370909.8458	4616791.326	61.79	63	5.91	962	120.69	4.6	0	0	0
78	370909.8458	4616791.326	61.79	63	5.91	962	120.69	3	0	0	0
78	370909.8458	4616791.326	61.77	43	25.67	962	120.69	6.3	0	0	0
78	370909.8458	4616791.326	61.07	63	5.91	1054	135.12	4.1	0	0	0
78	370909.8458	4616791.326	61.06	43	25.67	1054	135.12	8.7	0	0	0
78	370909.8458	4616791.326	61.96	63	5.91	1181	153.90	4.9	0	0	0
78	370909.8458	4616791.326	64.44	43	25.67	495	67.83	9	0	0	0
78	370909.8458	4616791.326	61.95	43	25.67	1181	154.00	4.7	0	0	0
78	370909.8458	4616791.326	64.57	63	5.91	1433	197.54	5.3	0	0	0
78	370909.8458	4616791.326	64.46	63	5.91	1514	212.38	4.83	0	0	0
78	370909.8458	4616791.326	64.34	43	25.67	1515	212.49	8.34	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
78	370909.8458	4616791.326	64.53	43	25.67	1433	197.54	11	0	0	0
78	370909.8458	4616791.326	64.34	43	25.67	1515	212.49	8.18	0	0	0
78	370909.8458	4616791.326	64.84	63	5.91	1556	218.44	4.37	0	0	0
78	370909.8458	4616791.326	64.21	43	25.67	1557	218.44	8.2	0	0	0
78	370909.8458	4616791.326	63.98	63	5.91	1684	232.20	2.84	0	0	0
78	370909.8458	4616791.326	63.96	43	25.67	1685	232.20	6.48	0	0	0
78	370909.8458	4616791.326	64.3	63	5.91	1936	274.39	2.1	32.26	32.23	32.23
78	370909.8458	4616791.326	64.29	43	25.67	1936	274.39	3.5	32.26	32.23	32.23
114	370554.3323	4616779.681	64.03	101	-35.65	362	48.01	11	0	0	0
114	370554.3323	4616779.681	68.85	101	-35.65	438	57.62	13	0	0	0
114	370554.3323	4616779.681	64.11	101	-35.65	534	74.89	10	0	0	0
114	370554.3323	4616779.681	62.15	101	-35.65	1047	132.67	11	0	0	0
114	370554.3323	4616779.681	63.05	101	-35.65	720	95.17	22.1	0	0	0
114	370554.3323	4616779.681	63.87	44	27.37	740	96.71	127	0	0	0
114	370554.3323	4616779.681	61.69	44	27.37	879	115.28	72	0	0	0
114	370554.3323	4616779.681	61.39	101	-35.65	902	119.23	12	0	0	0
114	370554.3323	4616779.681	61.12	101	-35.65	951	120.69	14	0	0	0
114	370554.3323	4616779.681	61.14	44	27.37	951	120.69	64	0	0	0
114	370554.3323	4616779.681	62.15	44	27.37	1047	132.67	71	0	0	0
114	370554.3323	4616779.681	62.86	101	-35.65	1242	166.35	9.6	0	0	0
114	370554.3323	4616779.681	62.86	44	27.37	1258	166.35	56	0	0	0
114	370554.3323	4616779.681	64.34	44	27.37	1369	188.45	52	0	0	0
114	370554.3323	4616779.681	64.25	101	-35.65	1370	188.65	7.7	0	0	0
114	370554.3323	4616779.681	63.68	101	-35.65	1500	212.37	13.4	0	0	0
114	370554.3323	4616779.681	63.68	44	27.37	1500	212.38	42.3	0	0	0
114	370554.3323	4616779.681	63.56	101	-35.65	1570	221.79	9.67	0	0	0
114	370554.3323	4616779.681	63.62	44	27.37	1570	221.79	37.7	0	0	0
114	370554.3323	4616779.681	63.4	101	-35.65	1672	231.62	4.36	0	0	0
114	370554.3323	4616779.681	63.41	44	27.37	1672	231.62	40.8	0	0	0
114	370554.3323	4616779.681	63.32	101	-35.65	1929	274.35	1.7	31.25	31.13	31.13
114	370554.3323	4616779.681	63.4	44	27.37	1929	274.35	54	31.25	31.13	31.13
129	370565.5007	4616722.093	63.18	61	5.12	367	48.61	10	0	0	0
129	370565.5007	4616722.093	62.82	61	5.12	438	57.62	9	0	0	0
129	370565.5007	4616722.093	62.79	51	14.89	438	57.62	6	0	0	0
129	370565.5007	4616722.093	64.12	61	5.12	535	74.89	6	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
129	370565.5007	4616722.093	64.14	51	14.89	536	74.89	8	0	0	0
129	370565.5007	4616722.093	62.98	61	5.12	720	95.17	5.92	0	0	0
129	370565.5007	4616722.093	63.09	51	14.89	720	95.17	6.93	0	0	0
129	370565.5007	4616722.093	64.88	61	5.12	832	107.37	4.63	0	0	0
129	370565.5007	4616722.093	64.89	51	14.89	832	107.37	0.72	0	0	0
129	370565.5007	4616722.093	61.74	31	34.85	835	107.54	0.69	0	0	0
129	370565.5007	4616722.093	61.03	61	5.12	961	120.69	1.9	0	0	0
129	370565.5007	4616722.093	61.04	51	14.89	961	120.69	13	0	0	0
129	370565.5007	4616722.093	61.13	31	34.85	961	120.69	1.5	0	0	0
129	370565.5007	4616722.093	60.38	61	5.12	1047	132.67	2.2	0	0	0
129	370565.5007	4616722.093	61.04	51	14.89	1047	132.67	16	0	0	0
129	370565.5007	4616722.093	60.43	51	14.89	1047	132.67	15	0	0	0
129	370565.5007	4616722.093	64.43	31	34.85	1047	132.67	0.7	0	0	0
129	370565.5007	4616722.093	61.12	61	5.12	1178	153.87	5.9	0	0	0
129	370565.5007	4616722.093	61.11	51	14.89	1178	153.90	14	0	0	0
129	370565.5007	4616722.093	64.29	61	5.12	1370	188.72	8.5	0	0	0
129	370565.5007	4616722.093	64.23	51	14.89	1370	188.72	6.7	0	0	0
129	370565.5007	4616722.093	64.33	31	34.85	1370	188.72	0.59	0	0	0
129	370565.5007	4616722.093	63.77	61	5.12	1501	212.38	6.62	0	0	0
129	370565.5007	4616722.093	63.89	51	14.89	1501	212.38	5.13	0	0	0
129	370565.5007	4616722.093	63.48	61	5.12	1557	218.44	6.54	0	0	0
129	370565.5007	4616722.093	63.49	51	14.89	1558	218.44	5.27	0	0	0
129	370565.5007	4616722.093	63.12	61	5.12	1679	232.20	3.68	0	0	0
129	370565.5007	4616722.093	63.44	51	14.89	1679	232.20	4.74	0	0	0
129	370565.5007	4616722.093	63.41	31	34.85	1679	232.20	0.36	0	0	0
129	370565.5007	4616722.093	62.27	31	34.85	1802	251.63	1.2	12.96	11.06	11.06
129	370565.5007	4616722.093	63.2	61	5.12	1921	274.33	1.5	30.10	29.86	29.86
129	370565.5007	4616722.093	63.17	51	14.89	1921	274.33	4.5	30.10	29.86	29.86
139	370336.7162	4616757.016	62.71	75	-10.43	363	48.01	8	0	0	0
139	370336.7162	4616757.016	63.53	75	-10.43	536	74.89	3	0	0	0
139	370336.7162	4616757.016	61.2	115	-50.52	837	107.59	1.86	0	0	0
139	370336.7162	4616757.016	60.89	75	-10.43	837	107.59	2.77	0	0	0
139	370336.7162	4616757.016	60.66	115	-50.52	951	120.69	1.6	0	0	0
139	370336.7162	4616757.016	60.62	75	-10.43	951	120.69	1.2	0	0	0
139	370336.7162	4616757.016	60.62	75	-10.43	951	120.69	1.3	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
139	370336.7162	4616757.016	59.78	115	-50.52	1038	132.67	1.4	0	0	0
139	370336.7162	4616757.016	60.63	115	-50.52	1182	154.00	0.65	0	0	0
139	370336.7162	4616757.016	62.19	75	-10.43	439	57.62	11	0	0	0
139	370336.7162	4616757.016	63.64	115	-50.52	1378	188.72	0.42	0	0	0
139	370336.7162	4616757.016	63.67	75	-10.43	1378	189.04	13	0	0	0
139	370336.7162	4616757.016	63.02	115	-50.52	1518	212.49	0.401	0	0	0
139	370336.7162	4616757.016	62.99	115	-50.52	1592	224.94	0.595	0	0	0
139	370336.7162	4616757.016	62.99	75	-10.43	1595	224.94	0.6	0	0	0
139	370336.7162	4616757.016	62.79	115	-50.52	1677	231.62	0.505	0	0	0
139	370336.7162	4616757.016	62.82	75	-10.43	1677	231.62	3.5	0	0	0
139	370336.7162	4616757.016	62.51	75	-10.43	1923	274.33	2.94	30.38	30.18	30.18
162	370681.3427	4616701.79	64.0	54.28	9.25	748	98.24	1.55	0	0	0
162	370681.3427	4616701.79	61.54	54.28	9.25	838	107.59	2.03	0	0	0
162	370681.3427	4616701.79	61.33	54.28	9.25	950	120.69	2.4	0	0	0
162	370681.3427	4616701.79	60.60	54.28	9.25	1048	132.67	1.9	0	0	0
162	370681.3427	4616701.79	61.36	54.28	9.25	1181	154.00	3.5	0	0	0
162	370681.3427	4616701.79	61.36	54.28	9.25	1181	154.00	3.4	0	0	0
162	370681.3427	4616701.79	64.43	54.28	9.25	1378	189.04	4.4	0	0	0
162	370681.3427	4616701.79	63.86	54.28	9.25	1521	212.49	3.91	0	0	0
162	370681.3427	4616701.79	63.75	54.28	9.25	1560	218.44	4.11	0	0	0
162	370681.3427	4616701.79	63.55	54.28	9.25	1670	231.16	6.2	0	0	0
162	370681.3427	4616701.79	62.4	54.28	9.25	1802	251.63	10	12.96	11.06	11.06
165	370226.6833	4616698.383	63.03	51	13.23	493	67.83	122	0	0	0
165	370226.6833	4616698.383	63.51	51	13.23	593	84.42	102	0	0	0
165	370226.6833	4616698.383	62.03	51	13.23	740	96.71	81.2	0	0	0
165	370226.6833	4616698.383	60.96	51	13.23	838	107.59	83.5	0	0	0
165	370226.6833	4616698.383	60.42	51	13.23	952	120.69	64	0	0	0
165	370226.6833	4616698.383	59.57	51	13.23	1060	135.82	78	0	0	0
165	370226.6833	4616698.383	60.47	111	-46.78	1181	154.00	4	0	0	0
165	370226.6833	4616698.383	60.38	51	13.23	1181	154.00	110	0	0	0
165	370226.6833	4616698.383	63.64	111	-46.78	1349	187.54	2.5	0	0	0
165	370226.6833	4616698.383	63.58	51	13.23	1349	187.54	57	0	0	0
165	370226.6833	4616698.383	63.58	51	13.23	1362	187.54	58	0	0	0
165	370226.6833	4616698.383	63.84	111	-46.78	1519	212.49	3.15	0	0	0
165	370226.6833	4616698.383	62.81	51	13.23	1521	212.49	50.9	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
165	370226.6833	4616698.383	62.81	51	13.23	1521	212.49	50.9	0	0	0
165	370226.6833	4616698.383	62.73	111	-46.78	1558	218.44	3.05	0	0	0
165	370226.6833	4616698.383	62.63	51	13.23	1560	218.44	39	0	0	0
165	370226.6833	4616698.383	62.62	111	-46.78	1678	232.20	3.54	0	0	0
165	370226.6833	4616698.383	61.33	51	13.23	1679	232.20	41.3	0	0	0
165	370226.6833	4616698.383	61.33	51	13.23	1802	250.36	94	12.96	11.06	11.06
165	370226.6833	4616698.383	62.28	51	13.23	1930	274.35	9.8	31.39	31.28	31.28
172	370223.4066	4616613.115	60.09	109	-46.49	537	74.89	3	0	0	0
172	370223.4066	4616613.115	62.94	109	-46.49	629	87.43	3.394	0	0	0
172	370223.4066	4616613.115	61.3	109	-46.49	769	99.63	5.45	0	0	0
172	370223.4066	4616613.115	59.18	109	-46.49	968	123.84	7.1	0	0	0
172	370223.4066	4616613.115	59.23	109	-46.49	1054	135.82	6.8	0	0	0
172	370223.4066	4616613.115	60.14	109	-46.49	1182	154.00	6.8	0	0	0
172	370223.4066	4616613.115	63.24	109	-46.49	1383	189.04	6.8	0	0	0
172	370223.4066	4616613.115	62.61	109	-46.49	1501	212.38	4.45	0	0	0
172	370223.4066	4616613.115	62.61	109	-46.49	1501	212.38	4.44	0	0	0
172	370223.4066	4616613.115	61.47	109	-46.49	1570	221.79	4.39	0	0	0
172	370223.4066	4616613.115	62.36	109	-46.49	1613	227.04	4.1	0	0	0
172	370223.4066	4616613.115	61.88	109	-46.49	1921	274.33	2.1	30.10	29.86	29.86
173	369355.6014	4616669.876	57.85	57.2	5.18	755	98.85	0.632	0	0	0
173	369355.6014	4616669.876	59.65	57.2	5.18	801	100.11	0.672	0	0	0
173	369355.6014	4616669.876	56.4	57.2	5.18	839	107.79	0.88	0	0	0
173	369355.6014	4616669.876	56.45	57.2	5.18	951	120.69	0.7	0	0	0
173	369355.6014	4616669.876	58.56	57.2	5.18	1048	132.67	0.48	0	0	0
173	369355.6014	4616669.876	58.56	57.2	5.18	1242	166.35	1.1	0	0	0
173	369355.6014	4616669.876	61.75	57.2	5.18	1413	195.84	0.65	0	0	0
173	369355.6014	4616669.876	61.75	57.2	5.18	1413	195.87	0.602	0	0	0
173	369355.6014	4616669.876	61.35	57.2	5.18	1502	212.38	0.842	0	0	0
173	369355.6014	4616669.876	52.13	57.2	5.18	1567	221.79	0.75	0	0	0
173	369355.6014	4616669.876	52.13	57.2	5.18	1570	221.79	0.532	0	0	0
173	369355.6014	4616669.876	61.25	57.2	5.18	1671	231.62	0.621	0	0	0
173	369355.6014	4616669.876	61.15	57.2	5.18	1934	274.35	0.45	31.97	31.92	31.92
210	369996.6187	4616665.601	59.82	59.69	-0.1	887	117.19	12	0	0	0
210	369996.6187	4616665.601	59.04	59.69	-0.1	1031	129.84	9.93	0	0	0
210	369996.6187	4616665.601	59.15	59.69	-0.1	1154	150.90	12	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
210	369996.6187	4616665.601	62.02	59.69	-0.1	1496	211.17	19	0	0	0
210	369996.6187	4616665.601	61.91	59.69	-0.1	1531	213.12	23	0	0	0
210	369996.6187	4616665.601	61.93	59.69	-0.1	1595	224.94	44	0	0	0
210	369996.6187	4616665.601	61.94	59.69	-0.1	1601	225.15	43	0	0	0
210	369996.6187	4616665.601	61.78	59.69	-0.1	1677	231.62	59	0	0	0
210	369996.6187	4616665.601	59.82	59.69	-0.1	887	117.19	11	0	0	0
210	369996.6187	4616665.601	60.55	59.69	-0.1	1801	250.36	56	12.82	10.90	10.90
211	369353.6537	4616562.439	56.13	34.7	-3.6	887	117.19	3	0	0	0
211	369353.6537	4616562.439	56.4	60	-3.6	991	129.84	0.51	0	0	0
211	369353.6537	4616562.439	55.96	34.7	-3.6	1031	129.84	3.02	0	0	0
211	369353.6537	4616562.439	56.22	34.7	-3.6	1153	150.90	3.5	0	0	0
211	369353.6537	4616562.439	60.15	60	-3.6	1601	225.15	11	0	0	0
211	369353.6537	4616562.439	59.94	60	-3.6	1495	211.17	5.6	0	0	0
211	369353.6537	4616562.439	60.01	60	-3.6	1525	212.61	9.8	0	0	0
211	369353.6537	4616562.439	60.18	60	-3.6	1672	231.62	13	0	0	0
211	369353.6537	4616562.439	58.44	60	-3.6	1801	244.92	33	12.82	10.90	10.90
211	369353.6537	4616562.439	60	34.7	-3.6	1801	250.36	0.72	12.82	10.90	10.90
211	369353.6537	4616562.439	58.44	34.7	-3.6	1801	250.36	0.66	12.82	10.90	10.90
211	369353.6537	4616562.439	58.44	60	-3.6	1921	252.77	25	30.10	29.86	29.86
211	369353.6537	4616562.439	58.71	34.7	-3.6	1921	274.33	3	30.10	29.86	29.86
214	370181.7339	4616541.333	59.21	83.45	-23.19	1032	129.89	0.6	0	0	0
214	370181.7339	4616541.333	59.54	83.45	-23.19	1124	146.92	0.72	0	0	0
214	370181.7339	4616541.333	62.21	83.45	-23.19	1535	213.12	0.65	0	0	0
214	370181.7339	4616541.333	62.36	83.45	-23.19	1602	225.15	0.35	0	0	0
214	370181.7339	4616541.333	62.07	83.45	-23.19	1672	231.62	0.61	0	0	0
225	368970.7769	4616538.54	53.37	31.48	21.89	948	120.69	2.9	0	0	0
225	368970.7769	4616538.54	52.55	31.48	21.89	1047	132.67	1.5	0	0	0
225	368970.7769	4616538.54	53.15	31.48	21.89	1132	150.90	0.62	0	0	0
225	368970.7769	4616538.54	55.21	31.48	21.89	1495	210.57	1.9	0	0	0
225	368970.7769	4616538.54	55.31	31.48	21.89	1531	213.12	2.5	0	0	0
225	368970.7769	4616538.54	55.39	31.48	21.89	1602	225.15	2.62	0	0	0
225	368970.7769	4616538.54	55.19	31.48	21.89	1678	232.20	2.1	0	0	0
225	368970.7769	4616538.54	54.1	31.48	21.89	1803	251.63	3.2	13.10	11.22	11.22
225	368970.7769	4616538.54	55.35	31.48	21.89	1922	274.33	7.7	30.24	30.02	30.02
231	368986.0277	4616432.219	54.23	109.15	-55.42	968	120.69	0.51	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
231	368986.0277	4616432.219	52.9	63.33	-10.26	1048	132.67	0.45	0	0	0
231	368986.0277	4616432.219	53.02	63.33	-10.26	1131	146.92	0.6	0	0	0
231	368986.0277	4616432.219	55.58	63.33	-10.26	1490	197.54	0.58	0	0	0
231	368986.0277	4616432.219	61.3	63.33	-10.26	1530	213.12	0.63	0	0	0
231	368986.0277	4616432.219	57	63.33	-10.26	1671	231.62	0.71	0	0	0
231	368986.0277	4616432.219	58.3	63.33	-10.26	1804	251.66	0.663	13.25	11.38	11.38
231	368986.0277	4616432.219	58	63.33	-10.26	1928	274.33	0.76	31.104	30.968	30.968
240	368936.5251	4616632.908	63.65	31.45	20.55	1804	251.66	0.35	13.25	11.38	11.38
252	368590.8485	4616334.163	52.35	6.63	41.36	1242	163.21	0.55	0	0	0
252	368590.8485	4616334.163	51.97	6.63	41.36	1539	213.75	0.41	0	0	0
252	368590.8485	4616334.163	52.17	65.6	-17.59	1671	231.62	0.4	0	0	0
252	368590.8485	4616334.163	52.17	6.63	41.36	1671	231.62	0.43	0	0	0
252	368590.8485	4616334.163	53.42	65.6	-17.59	1936	274.35	0.41	32.26	32.23	32.23
255	370285.5839	4616964.46	60.53	65.43	-4.9	1193	157.01	0.54	0	0	0
255	370285.5839	4616964.46	64.9	65.43	-4.9	1348	178.56	1.1	0	0	0
255	370285.5839	4616964.46	64.74	65.43	-4.9	1432	197.54	0.36	0	0	0
258	368514.469	4616575.047	47.31	44.7	2.61	1154	152.80	0.408	0	0	0
258	368514.469	4616575.047	47.31	44.7	2.61	1161	152.80	3	0	0	0
258	368514.469	4616575.047	47.28	34.75	12.53	1161	152.80	0.49	0	0	0
258	368514.469	4616575.047	47.28	34.75	12.53	1175	152.80	1.9	0	0	0
258	368514.469	4616575.047	51.99	44.7	2.61	1307	169.09	0.4	0	0	0
258	368514.469	4616575.047	51.73	69.1	-21.95	1362	188.45	0.36	0	0	0
258	368514.469	4616575.047	51.78	44.7	2.61	1365	188.45	0.51	0	0	0
258	368514.469	4616575.047	50.96	44.7	2.61	1552	215.91	0.9	0	0	0
258	368514.469	4616575.047	50.97	69.1	-21.95	1670	231.16	0.39	0	0	0
258	368514.469	4616575.047	51.13	44.7	2.61	1671	231.62	1.4	0	0	0
258	368514.469	4616575.047	51.08	34.75	12.53	1671	231.62	0.73	0	0	0
258	368514.469	4616575.047	49.74	69.1	-21.95	1804	251.66	0.456	13.25	11.38	11.38
258	368514.469	4616575.047	51.03	44.7	2.61	1804	251.66	1.62	13.25	11.38	11.38
258	368514.469	4616575.047	49.96	34.75	12.53	1804	251.66	1.01	13.25	11.38	11.38
258	368514.469	4616575.047	53.7	69.1	-15.4	1985	274.52	0.68	39.31	39.97	39.97
258	368514.469	4616575.047	53.73	44.7	2.61	1985	283.75	4	39.312	39.974	39.974
258	368514.469	4616575.047	53.74	34.75	12.53	1985	283.75	1.9	39.312	39.974	39.974
258	368514.469	4616575.047	53.73	34.75	12.53	1985	283.75	1.9	39.312	39.974	39.974
341	369354.7273	4616497.089	59.02	55.66	3.86	1691	232.20	2.95	0	0	0

Well #	X	Y	G	Depth	Z	T	R	P	INJ1	INJ	INJ 3
341	369354.7273	4616497.089	58.29	25.16	33.76	1704	236.21	14.7	0.00	0	0
341	369354.7273	4616497.089	57.15	108.16	-49	1805	251.66	0.428	13.39	11.53	11.53
341	369354.7273	4616497.089	57.83	55.66	3.86	1805	252.71	15.5	13.39	11.53	11.53
341	369354.7273	4616497.089	57.23	25.16	33.76	1805	252.71	0.442	13.39	11.53	11.53
341	369354.7273	4616497.089	58.5	108.16	-49	1934	274.35	0.56	31.97	31.92	31.92
341	369354.7273	4616497.089	59.11	55.66	3.86	1934	274.35	40	31.97	31.92	31.92