

NM WAIDS: A Produced Water Quality and Infrastructure
GIS Database for New Mexico Oil Producers
(FINAL Technical Progress Report)

Report Start Date: October 1, 2002

Report End Date: May 31, 2005

Martha Cather, PI
Robert Lee, Project Manager
Ibrahim Gundiler, Co-PI
Andrew Sung, Co-PI

Contributors:
Naomi Davidson
Ajeet Kumar Reddy
Mingzhen Wei

June 1, 2005

DE-FC26-02NT15134

New Mexico Petroleum Recovery Research Center
New Mexico Institute of Mining and Technology
Socorro, NM 87801
(505) 835-5142

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, produce, or process disclosed or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily stat or reflect those of the United States Government or any agency thereof.

ABSTRACT

The New Mexico Water and Infrastructure Data System (NM WAIDS) was created to alleviate a number of produced water-related issues in southeast New Mexico. This project entailed the design and implementation of a Geographical Information System (GIS) and integral tools to provide operators and regulators with necessary data and useful information to help them make management and regulatory decisions.

The major components of NM WAIDS are: 1) databases on produced water and groundwater quality, cultural and natural features, and corrosion information, 2) a web site capable of displaying this data in a GIS format or accessing some of the data by text-based queries, 3) a fuzzy logic-based risk assessment tool that could be used to assess the seriousness of a spill of produced water, and 4) a corrosion and scale management toolkit that provides operators with data and information on water-related corrosion and scale problems to aid them in deciding how to address such issues. The various parts of NM WAIDS have been integrated into a website with a user-friendly interface that provides access to previously difficult-to-obtain data and information.

Over the course of the project, the following activities were accomplished:

- Creation of a water quality database, with over 7000 entries for New Mexico for produced water analyses from a variety of sources. Database creation included data acquisition, cleaning, and integration with other data sources.
- Creation of a web-based data entry system for this database that allows a user to view, enter, or edit data from a web page rather than having to directly access the database.
- Creation of a semi-automated data capturing system for use with standard water quality analysis forms, which improved the accuracy and speed of water quality data entry.
- Acquisition of ground water data from the New Mexico State Engineer's office, including chloride content and TDS (Total Dissolved Solids) for over 30,000 data points in southeast New Mexico.
- Generation of approximate locations to enable mapping of this data.
- Creation of a web-based scale prediction tool, also with a web-based interface, that uses two common scaling indices to predict the likelihood of scaling. This prediction tool can either run from user input data, or the user can select samples from the water analysis database.
- Creation of depth-to-groundwater maps for the study area.
- Analysis of water quality data by formation.
- Continuation of efforts to collect produced water quality information from operators in the southeast New Mexico area.
- Qualitative assessment of produced water corrosivity from various formations.
- Efforts at corrosion education in the region through operator visits.
- Compilation of both hard copy and online corrosion toolkit material.
- Improvement of the integrated web and GIS interface for all the information collected in this effort, including data from northwest New Mexico.
- Continued development of a fuzzy logic spill risk assessment tool that was initially developed prior to this project. Improvements include addition of parameters found to be significant in determining the impact of a brine spill at a specific site.

TABLE OF CONTENTS

DISCLAIMER	ii
ABSTRACT	iii
LIST OF TABLES	v
LIST OF FIGURES	vi
INTRODUCTION	1
EXECUTIVE SUMMARY	2
EXPERIMENTAL METHODS.....	3
PROGRESS AND DISCUSSION OF RESULTS.....	3
Database Construction	3
Groundwater Database.....	3
Produced Water Quality Database.....	3
Data Standardization and Assimilation.....	6
GIS Construction	13
Corrosion Management Toolkit.....	13
Reference Book.....	14
Formation Water Analysis.....	14
Mapping Aquifer Thickness	17
Formation Water Interpretation	18
Scale Prediction Tool.....	19
NM WAIDS Web Site	22
Databases	22
Tools	22
Maps.....	22
Manual	23
Web Pages.....	23
Technology Transfer.....	24
CONCLUSIONS.....	24
REFERENCES	26
LIST OF ACRONYMS AND ABBREVIATIONS	28

LIST OF TABLES

Table 1. Sample List of Data from Two Data Sets	29
Table 2. String Sample Pairs from Two Data Sets and Matching Results Sets.....	29
Table 3. Name and Address String Comparisons	30
Table 4. Conversion Factors for Common Anions and Cations	30
Table 5. Water Injection Statistics for Southeast New Mexico	31

LIST OF FIGURES

Fig. 1. (a) Example of a good, clean water analysis form that is easy to convert to text. (b) Example of a poor quality water analysis form. Note the fuzziness of the type and the many black specks on the page. These make OCR more difficult.	32
Fig. 2. Example of a form that has been scanned and processed. The numbered boxes are the data fields that are recorded in a database, while the dashed lines surround fields that are reference fields.....	33
Fig.3. Flowchart of the new string comparison algorithm used for cleaning data to be entered in the Produced Water Quality Database.	34
Fig. 4. Water analysis form showing two reporting methods (ppm and ep _m) for the sample.	35
Fig. 5. Reported and calculated values of chloride and magnesium in ppm and mg/l vs ep _m units. These types of comparisons are very useful for quality control of scanned and hand-entered data. If the value of the number in ep _m units divided by that of ppm or mg/l units was less than half or greater than twice the known conversion factor (for instance, 1ppm = 0.0282 ep _m for chloride), the data was considered inaccurate and values were checked by hand against the scanned image. Note that the slope of the trend lines for calculated data is equal to the conversion factor given in Table 4.	36
Fig. 6. Location of three cross sections and 10 km buffer around them. Produced water samples from wells that fall within the buffer zone were projected onto the closest cross sections in a N-S or E-W line, depending on the predominant direction of the cross section. Line A-A' marks the cross section line and buffer zone for the next three figures.....	37
Fig. 7. Cross section through Eddy and Lea counties showing produced water samples by depth and reported formation.....	38
Fig. 8. Chloride concentration of produced water samples for cross section A-A'. Interestingly, some of the most saline waters are found at relatively shallow depths.	38
Fig. 9. Cross section A-A', showing interpreted origin of the produced waters. Solid points are based on Cl/SO ₄ and Ca/Mg ratios, while open points are based on Ca/SO ₄ ratios only.....	39
Fig. 10. Stratigraphic column showing various formations of New Mexico Permian Basin. Produced water sample chemistries were obtained from almost all formations.....	40
Fig. 11. Typical piper diagram of produced water samples, (Leonard) Abo shown.	41

Fig. 12. Piper diagram of San Andres and Grayburg produced water samples.....41

Fig. 13. Home page for NM WAIDS web site. Interface includes links to databases, the GIS map server, various tools for predicting corrosion and scale, and the online corrosion manual.....42

Fig. 14. Search interface to Produced Water Quality Database.....42

Fig. 15. Results for a search using T23S as search criteria. Note that some wells have more than one sample available. Clicking on the hyperlinked (underlined) numbers will bring up a screen like that seen below in Figure 13.....43

Fig. 16. Screen shot showing samples available for the Big Freddy Unit Well 002. Clicking on one of the icons will either show more information about the sample, will allow the user to do an Odde-Tomson or Stiff-Davis scale calculation, or will select the sample to use in the water mixing program available in the Tools section of the web site.....43

Fig. 17. Screen showing water chemistry data for a selected sample.....44

Fig. 18. Screen shot of sample data entered into Stiff-Davis calculation tool.....44

Fig. 19. Stiff Davis scale calculation. Scale precipitation is likely to occur at temperatures in excess of about 155°F.45

Fig. 20. Probable composition of any scale that would form from a brine of this chemistry.....45

Fig. 21. Tools available include two scale calculation tools, a mineral composition tool, a tool that will calculate the composition of a mix of two waters, and a variety of unit conversion calculators.....46

Fig. 22. Screen shot of NM WAIDS online map, zoomed in to Lea County. Produced and ground water are shown in different color ranges, with darker colors showing as higher chloride contents. This is particularly difficult to accurately demonstrate in the grayscale image.....46

Fig. 23. Closeup of a selected well in the Groundwater Database. Using the Identify tool on the map server will bring back a screen similar to that seen below, where there is some information about the selected sample(s). Clicking on one of the hyperlinks in the Sample_ID column will provide more information, seen in Fig. 22.47

Fig. 24. Information about selected sample in groundwater database. Using either the Identify or the Hotlink tool will link to this same report screen.....47

Fig. 25. Sample record and detailed information for a well selected from the Produced Water layer of the online map.....48

Fig. 26. Screen shot of home page for the corrosion manual.	49
Fig. 27. Theory page has more information about a variety of commonly seen oilfield corrosion mechanisms.	49
Fig. 28. Screen shot of the CO ₂ Corrosion page, providing more information, likely environments of formation, pictures, and mitigation options.	50
Fig. 29. Corrosion prevention section.	50
Fig. 30. Screen shot of Scale page.	51
Fig. 31. Help section.	51

INTRODUCTION

The concept for the NM WAIDS project was generated by discussions with New Mexico oil and gas producers that revealed a number of water-related problems and issues that they had to deal with on a daily basis. Some were merely small headaches, like determining depth to groundwater in an area, while others—water-related scaling problems— took a considerable toll in time, money and equipment. Producers’ needs included:

- a better understanding of where and why corrosion and scale were occurring in their production equipment;
- knowing chemistries and volumes of produced water for possible treatment and reuse options, and
- knowing more about the ground water in producing areas to understand what areas might be more vulnerable to brine spills.

Thus, the primary goal of this project was to design and implement a robust, secure, and dynamic system on the Internet to provide New Mexico producers both data warehousing services and access to a variety of water-related data sets and tools.

EXECUTIVE SUMMARY

The New Mexico Water and Infrastructure Data System (NM WAIDS) was created to alleviate a number of produced water-related issues in southeast New Mexico. This project entailed the design and implementation of a Geographical Information System (GIS) and integral tools to provide operators and regulators with necessary data and useful information to help them make management and regulatory decisions.

The major components of NM WAIDS are: 1) databases on produced water and groundwater quality, cultural and natural features, and corrosion information, 2) a web site capable of displaying this data in a GIS format or accessing some of the data by text-based queries, 3) a fuzzy logic-based risk assessment tool that could be used to assess the seriousness of a spill of produced water, and 4) a corrosion and scale management toolkit that provides operators with data and information on water-related corrosion and scale problems to aid them in deciding how to address such issues. The various parts of NM WAIDS have been integrated into a website with a user-friendly interface that provides access to previously difficult-to-obtain data and information.

Over the course of the project, the following activities were accomplished:

- Creation of a water quality database, with over 7000 entries for New Mexico for produced water analyses from a variety of sources. Database creation included data acquisition, cleaning, and integration with other data sources.
- Creation of a web-based data entry system for this database that allows a user to view, enter, or edit data from a web page rather than having to directly access the database.
- Creation of a semi-automated data capturing system for use with standard water quality analysis forms, which improved the accuracy and speed of water quality data entry.
- Acquisition of ground water data from the New Mexico State Engineer's office, including chloride content and TDS (Total Dissolved Solids) for over 30,000 data points in southeast New Mexico.
- Generation of approximate locations to enable mapping of this data.
- Creation of a web-based scale prediction tool, also with a web-based interface, that uses two common scaling indices to predict the likelihood of scaling. This prediction tool can either run from user input data, or the user can select samples from the water analysis database.
- Creation of depth-to-groundwater maps for the study area.
- Analysis of water quality data by formation.
- Continuation of efforts to collect produced water quality information from operators in the southeast New Mexico area.
- Qualitative assessment of produced water corrosivity from various formations.
- Efforts at corrosion education in the region through operator visits.
- Compilation of both hard copy and online corrosion toolkit material.
- Improvement of the integrated web and GIS interface for all the information collected in this effort, including data from northwest New Mexico.
- Continued development of a fuzzy logic spill risk assessment tool that was initially developed prior to this project. Improvements include addition of parameters found to be significant in determining the impact of a brine spill at a specific site.

EXPERIMENTAL METHODS

No experimental methods, materials, or equipment were used for this project. Data for the water chemistry databases was collected from previously existing private and public sources. Information on corrosion issues was collected by mailed survey and by personal interviews with operators and chemical company representatives in the area. ArcGIS software was used for the mapping aquifers, well locations, and water chemistry data. Microsoft SQL Server 2000 was used to store and access water chemistry databases.

PROGRESS AND DISCUSSION OF RESULTS

Database Construction

The development of dynamic web-accessible databases for storing, managing, accessing, and analyzing data to be accessed by both web-based queries and GIS-initiated queries was a key task for this project. As data files can be quite large, the system needed to be efficient and useable by persons with varying degrees of computer literacy.

Groundwater Database

The data comprising the bulk of the groundwater quality database was obtained from the Office of the State Engineer (NMSEO), Water Rights Division District 2, in Roswell, New Mexico. This data set contains over 30,000 records for southeast New Mexico and actually covers a larger region than is addressed in the NM WAIDS project. The only chemical data available were chloride content and total dissolved solids (TDS), so these were the only water quality parameters analyzed in this study. A second data set was also acquired from the NMSEO's Santa Fe office that contains no quality data, but does have basic positional information along with well depth and groundwater depths. This database covers the entire state and was used construct maps showing depths to groundwater, a parameter believed to be important when evaluating the potential severity of a chemical or brine spill.

Produced Water Quality Database

Creation of the Produced Water Quality Database (PWQD) was one of the largest and most time-consuming tasks of this entire effort. The PWQD was compiled from a large variety of source data. A number of oil and gas producers in the area were solicited for data, and many were very generous in sharing this information. Some of the data was provided in digital format, either as Microsoft Excel spreadsheets, Microsoft Access databases, or simple text files. Much, however, came from producers as paper forms supplied to them by the various companies employed to run the water analyses. Each data source had to be analyzed to determine what kind of data was available and in what format (numeric, text, semi-quantitative), so the correct fields and data definitions could be built into the database structure. Examination of the thousands of paper forms and

digital files revealed that data could be divided into four main categories: general information, general sample properties, anions, and cations. A number of tables and views were used in the database construction: primary tables are the general sample information (items such as sample name, location, formation, physical parameters), anion information (CO₃, SO₄, etc.), and cation information (Ca, Na, Mg, etc). After the initial database was constructed, digital data sets were imported into the database. As new data sets, either paper or digital, were acquired, some modifications to the PWQD were needed, but researchers found that the basic structure was able to accommodate new data types.

Web-Based Data Entry System

One component of the PWQD is a web-based data entry system. This was implemented for two reasons. First, much of the data entry for the initial database was done by student employees. It was more efficient, and also more secure, to have these employees work at their own computers and use a system of forms to enter data, rather than having them actually manipulate the database. Users can examine, add, or edit data via the web entry system; however it would be difficult for a user to inadvertently damage the entire database or delete large amounts of information. Another reason for web entry was to create a system into which producers could enter their own data, should they wish to. Although significant time was spent developing the web entry system, its use has been discontinued. Semi-automated data capture, discussed below, proved superior as a method for data entry of form data, and few operators have expressed any interest in keying in their own data.

Automatic Data Capture System

A second component of the PWQD, and one which received significant effort in the development of this project, is the Automatic Data Capturing System (ADCS). The objective of the ADCS was to speed the transfer of data from the multitude of paper forms into digital format, and to incorporate this data into the existing SQL databases.

Researchers collected over 3000 water quality analysis forms for input into the database. There was an average of 30 fields on each form from which data had to be collected, and there were many types of forms, so the data types were not always the same from form to form. Previous experience with the web-entry system showed that it took at least five minutes per form just to enter the data without any verification. Manual input proved to be impractical for the number of forms obtained, and was also prone to significant typographical errors. The new data capturing system requires less than a minute for each form if the form is well structured and clear and uses a previously defined form structure. This time includes any time for manual verification. In some difficult cases, such as forms filled in with cursive hand writing and unstructured forms, form processing may take up to three minutes depending on how bad the form is. In all but the worst cases, there was a time savings of 50-60%. Many of the documents processed were of poor quality (some date back to 1950); once processing of the older documents was completed, processing rates were higher still, leading to greater savings in time and money. An additional advantage is that now there is a digital record of each image, so if there is a question about the data, the actual form image can be examined.

An examination of currently existing processes and software was done as part of an assessment of the usefulness of ADC for our needs. Off-the-shelf technology (ABBY FormReader and ABBY FineReader) was used in our data capturing process, so most efforts were devoted to the actual data capture, quality control, and integration with previously-existing databases.

Technology behind the System

Optical Character Recognition (OCR) is the key technology by which an ADCS works. The OCR technology helps translate what is written on a piece of paper into an electronic text or number format. When a paper is scanned, it is made into a digital picture which the OCR software can interpret. The software/hardware system that recognizes characters from a registered image (part of the image from which individual characters to be recognized) can be divided into three operational steps: document analysis, character recognition, and contextual processing.

Document Analysis

Text is extracted from the document image in a process known as document analysis. Reliable character segmentation and recognition depend upon both original document quality and registered image quality. Processes that attempt to compensate for poor quality originals and/or poor quality scanning include image enhancement, underline removal, and noise removal. Image enhancement methods emphasize character versus non-character discrimination. Underline removal erases printed guidelines and other lines which may touch characters and interfere with character recognition and noise removal erases portions of the image that are not part of the characters.

Prior to character recognition it is necessary to isolate or segment individual characters from the text image of the word. Many OCR systems use connected components for this process. For those connected components that represent multiple or partial characters more sophisticated algorithms are used. In low quality or non-uniform text images these sophisticated algorithms may not correctly extract characters and thus, recognition errors may occur. Recognition of unconstrained handwritten text can be very difficult because characters cannot be reliably isolated especially when the text is cursive handwriting. This process of separation of individual character image files is called segmentation. A significant portion of the data we have acquired is handwritten. Fortunately recognition of numeric characters is fairly successful even for handwritten data.

Form Analysis

Two types of documents are typically found in ADC systems; form-based and text-based. Text-based documents have variable structure and all text images are converted to characters. This report would be an example of a text-based document. Form-based documents have a structured layout with variable and static information. A form such as an IRS 1040 form would be an example. Our project utilized form-based document processing. Form-based ADC systems

have two main requirements. For optimum speed and accuracy, all fields, graphic elements, separators, etc. must have the exact same location on all copies of the form. It is possible to automate the process if the forms fields are not at the same position on all the forms, but it adds an overhead of extra time in manually aligning the zones. Also, forms should have some reference points such as registration or rectifying marks to help in correction of any skew and removal of linear or nonlinear scanning defects.

Automatic form processing consists of three stages: batching, template creation, and form input. Batching is simply the separation of forms into batches of forms with an identical layout. Template creation tells the system where in the document a particular piece of information is found. After the template creation is done the system is ready to process the all forms that conform to that particular template. Form input can be done either as a batch, where all forms are first scanned and then processed, or it can be done individually, where each form is scanned and processing occurs immediately.

Form input is further divided into four stages 1) Scanning 2) Recognition 3) Verification and 4) Export. Scanning is performed using either an automatic document feed (ADF) or flat bed scanner. After scanning the image, any necessary preprocessing of the image, such as image rotation, page orientation detection, and noise reduction is done. After preprocessing, the image is ready for interpretation. The interpretation module first matches the template and the individual fields are recognized, and image pixels are converted into text or numbers. The program then performs a verification of field data by invoking several predefined rules. Fields which are not recognized properly are then highlighted for manual correction. This is the most time-consuming part of the ADC process, and the amount of time spent for verification increases drastically if the forms are not clear, are skewed, or if they have lot of noise. After the automatic and manual verifications, the data is ready for export either directly to an ODBC compatible database or into a CSV or spreadsheet file.

Figure 1a shows a typical good-quality document from which water quality information is extracted, while Fig. 1b shows one of the more difficult documents that was used. Both documents are treated in the way described above, but the second required more manual verification. Figure 2 shows the completed record after processing; information on the paper form has been transferred into an electronic image.

Although the ADCS greatly speeded the process of turning paper forms into digital data, it introduced several complications of its own that required considerable work to overcome. Two major issues had to be addressed. One was that of data conversion quality: how accurate was the OCR process for recognizing numerical data? The second, larger task was that of data standardization and assimilation—which will be addressed first in this Final Report.

Data Standardization and Assimilation

In order to assimilate data into a preexisting data structure that allows construction of meaningful search queries, (such as search by location or search by well name) the data must exist in a standardized format. Then, when the data is incorporated into the database, there must be an assurance that it does not duplicate data that is already in the database.

The API number is the most commonly-used unique identifier in most databases of oil and gas well data; and it is what we use identify wells in our production database. Unfortunately although much of the water quality data was obtained from producing oil and gas wells, the

forms and even electronic data often did not have an associated API recorded. Other data such as complete locations, formation names, and correct well names were often missing as well. Most forms did include a partial-to-complete well name or location; thus well location and well name became critical in our data schema. With a correct name and location, it is generally possible to link data in the produced water database to other databases maintained by both the state of New Mexico and the Petroleum Recovery Research Center that contain information such as API, location in X,Y coordinates, correct well and operator name, field, pool, and production information for most of the wells in the state. Thus, it was important for the project's own data on water to be easily cross-referenced to all this existent data.

When water quality forms are scanned, the OCR process simply turns the pixels in a particular region of the image into a text string. The operator of the system then must specify to the software program that a certain region on a particular form contains information of a particular type; location, for example. Then the program places the text string it finds in that region into the appropriate location field in a preliminary database. The problem for our example field, location, is that well locations were designated in many different ways on the scanned forms. M-11-30-7, SW/SW 11/30/7, M 11 30N 7W, 30-7-11 and M-11-30N7W are just some of the permutations that were used to describe the same parcel of land in the southwest corner of section 11, Township 30 North and Range 7 West. Similar variations exist in well names, pool names, and field names.

Additionally, for our example of location, almost all water chemistry reports were incomplete. Locations reported in unit letter, section, township, range (ULSTR) format are not precise, but only pinpoint a well down to a 40-acre parcel of land. Most samples lacked unit letter designations, coarsening the resolution down to 160 acres. In order to place the information into the GIS system that comprised one of the ultimate goals of this project, samples were required to have an actual X,Y location in either latitude/longitude or UTM coordinates. We had another database that already contained locations in X,Y coordinates for most of the wells in the state, including most wells that were sources of produced water data. Linking well information in the produced water database to that in the other database would provide location accurate to within a few hundred feet, rather than an approximation that might be several thousand feet off.

This example demonstrates only one of the reasons why data standardization and cleanup became an essential part of this project. Other useful linked information included water production volumes, well status, production histories, injection volumes and more. The tremendous amount of data cleanup and standardization would have been too time-consuming to do by hand. A survey of available literature on data cleaning showed that while there were some helpful algorithms already available, an improved comparison algorithm would be necessary for this task. No previous use of an algorithm exactly like the one we designed was found in our review, so although it does build on the work of previous researchers, we have called this a new approximate string comparison algorithm.

Creation of a New Approximate String Comparison Algorithm

Background

In the context of data cleaning, it is important to identify equivalent data in multiple sources, and more importantly, to identify data in multiple sources about the same real-world entity. The first problem is referred to as a field matching problem, and the second one as a record linkage problem. These two problems are independent of each other, yet mutually related. In order to identify equivalent data in the multiple sources, it is necessary to determine whether or not two syntactic values are alternative designations of the same semantic entity. For example, a method is needed to determine if “M-11-30-7,” “SW/SW 11/30/7,” and “M 11 30N 7W” refer to the same well. The problem can be described as a field matching problem.

The field matching problem evaluates the similarity of different values of the same attribute of an object.¹ Its solution and algorithm can be applied to solve record linkage problems. Data from heterogeneous sources often contains problems such as misspellings, typographical variations, non-standard abbreviations, or differences in the details of data format, both at field level and at record level. In database applications, string comparison algorithms for field matching or record linkage purposes are designed to detect strings that are duplicates of each other, but not necessarily textually identical. Using the name of a person as an example, “David Smith” and “D. Smith” could be the same person, but there are differences in characters between these two strings representing the name. Approximate string matching rather than strict character-by-character comparison is necessary for appropriate identification of these two strings as the same entity.

Approximate string comparison has been a subject of research in computer science for many years.² There are several approximate string comparison algorithms, each having its own characteristics. Most can be divided into one of two classes: character-by-character and word-by-word. Simply, algorithms in the character-by-character class compare strings character by character, and those in the word-by-word class compare strings word by word.

Time complexity is an important feature of any string-matching algorithm. Time complexity is defined as the way in which the number of steps required by an algorithm varies with the size of the problem it is solving.³ Time complexity is normally expressed as an order of magnitude: $O(N^2)$ means that if the size of the problem (N) doubles, then the algorithm will take four times as many steps to complete. Some string-matching algorithms may have very good results for data comparison but the time complexity can make them unusable for all but the smallest data sets.

Character-Based String Comparison Algorithms

Character-based string comparison algorithms compare two strings character by character, although they have different strategies to handle misspellings and variations in strings. Edit distance^{1,4} is a common measure of textual similarity. Given two strings, s_1 and s_2 , their edit distance, denoted $\Delta e(s_1, s_2)$, is the minimum number of edit operations, insertions, deletions, and substitutions of single characters that are needed to transform s_1 to s_2 . For example, $\Delta e(\text{“Harrison”}, \text{“Harison”}) = 1$. This algorithm compares two strings, s_1 and s_2 , character by character and its time complexity is $O(|s_1| * |s_2|)$, where $|s_1|$ is the length of string s_1 and $|s_2|$ is the

length of string s2. Further refinements of this basic concept have been introduced over the years.⁴⁻¹⁰ Refinements have included methods to allow gaps, or sequences of non-matching symbols, improvements in time efficiency, and to allow for^{2,11,12} partial agreement between two strings.

Word-Based String Comparison Algorithms

Word-based string comparison algorithms are different from character-by-character algorithms. These algorithms compare two strings word by word, rather than character by character. A literature search has revealed only two word-by-word string comparison algorithms thus far. One is a basic field matching algorithm and the other a recursive field matching algorithm.

The basic field matching algorithm splits strings into sequences of atomic strings and then sorts the atomic string sequences. An atomic string is a sequence of alphanumeric characters delimited by punctuation characters. Two atomic strings match if they are the same string or if one is the prefix of the other. The matching degree of two strings is defined as the number of common atomic strings divided by the average number of atomic strings.⁶ This basic algorithm will handle two situations; that where two atomic strings match each other exactly, and that where one atomic string is the prefix of the other. For example, the degree of match between the strings “Comput. Sci. & Eng. Dept.” and “department of computer science” is $3*2/7 = 0.86$. The algorithm processes the string matching word by word in a straightforward way, which causes the order of atomic strings in strings to have no influence on the matching degree. This is the most important advantage of this algorithm. Because string pairs commonly do have some out-of-order problems, the lack of influence of order on the matching degree makes this method superior to character-based comparison algorithms. However, the algorithm still does not cover all possible string matching situations, limiting its application.

A recursive field-matching algorithm⁶ was developed to make up for the disadvantages of the basic field matching algorithm. The recursive algorithm improves the basic algorithm by extending the string matching situations. In addition to the two string-matching situations handled by the basic algorithm, three more are given: a) one string is the combination of a prefix and a suffix of the other string, as “dept” and “department”; b) one string is the acronym of the other string, as “UCSD” and “university of California, San Diego”; and c) one string is the concatenation of prefixes from the other string, as “CalTech” and “California Institute of Technology”;

These string matching situations cover all possible string matching situations one can encounter when comparing two character strings. However, this algorithm is much more time-intensive to run, and still excludes many cases encountered in real-world string-matching applications.

From the above description, one can see that there are important differences between these two classes of string comparison algorithms. Character-by-character algorithms compare strings by character so the order of characters or words in the strings is very important; word-by-word algorithms do the same thing by word, hence only atomic string-matching matters and the order of atomic strings is ignored. Word-based matching algorithms are much more flexible than character-based matching algorithms for long multiple-word strings and character-based matching algorithms are better for spelling error checking. The word-based string comparison

algorithms increase the flexibility of string comparison by ignoring order problems and enrich its applications in various fields.

A New String Comparison Algorithm

Different string comparison algorithms fit different string comparison situations. Analysis of specific data characteristics is quite important prior to use of any algorithm. Analysis of our specific data problems showed data match issues that no single matching algorithm that we found was capable of handling in a time-efficient and correct manner. Table 1 shows data samples from two data sets that were used to create the final Produced Water Quality Database. This table demonstrates some of the differences between data that are equivalent for the same entity. The data on the left is from the scanned forms in a preliminary database, while the data on the right is the correct data according to the State of New Mexico's well database.

From Table 1, the following data problems can be detected.

- Out-of-order problems, like strings in pair 1;
- Typographical problems, like strings in pairs 2 and 3;
- Atomic string missing problem, like strings in pair 4;
- Abbreviation problems, like strings in pair 5;
- Numerical data matching problems, in all above samples;
- Mixture of numbers and characters in target strings in all above samples;
- Mixed atomic strings with characters and numbers, like strings in pair 6.

From the observation of the above strings, we can see three kinds of strings to match after removing all stop words and characters (words such as “and,” “in,” “the,” “of,” “inc,” and special characters or punctuation marks) from strings.

- Character strings, like “SAN”;
- Numerical strings, like “001”;
- Mixed strings, like “2A”;

Combining all these kinds of strings, the following string-matching cases can be derived from our analysis.

- Matching between character strings, like “SAN JUAN” and “SJ”;
- Matching between numerical strings, like “2” and “002”;
- Matching between mixed strings, like “2A” and “002A”;
- Matching between a character string and a numerical string, like “April” and “04”;

Based on the above string-matching analysis, a new approximate string comparison algorithm was developed specifically for our data problems. Figure 3 is a flowchart of this new string comparison algorithm. It is based on previous word-based string comparisons, particularly the work of Monge and others,⁶ but is greatly expanded to include the following matching problems:

For two character strings.

- Two strings exactly match each other ignoring case, as “match” and “Match”;

- One string is the prefix of the other string, as “sci” and “science”;
- One string is the combination of prefix and postfix of the other string, as “dept” and “department”;
- One string is the acronym of the other string, as “UCSD” and “University of California, San Diego”;
- One string is the concatenation of prefixes from the other string, as “CalTech” and “California Institute of Technology”;

For two numerical strings.

- Two strings match exactly if their real numerical values are same, as “012” and “12”, or “3.5” and “3.5000”

For two mixed strings.

- Two mixed strings match exactly if two strings both are in pattern “CharactersNumbersCharacters” or “NumbersCharactersNumbers” and the character parts of two strings match exactly according to the matching rules for character strings and the numerical parts of two strings match exactly according to the matching rule for numerical string, as “2A” and “002A”.

For one character string and one numerical string.

- Two strings match exactly if the character string is the exact alternative presentation of the numerical string, as “one” and “001”.

By extending the basic field matching algorithm, our new string comparison algorithm covers all possible string matching situations found after removing stop words in strings.

Since this string comparison algorithm is used to compare strings from large data sets, time efficiency becomes a very important factor. Time complexity of the basic field-matching algorithm, which is dominated by the time involved in sorting two sequences of atomic strings, is $O(n \log n)$ where n is the maximum number of atomic strings in either string. The time complexity of the new algorithm is same as that of the basic field-matching algorithm. The new algorithm employs an optimization method to mark all matching atomic strings to prevent matching atomic strings being revisited. This greatly improves the time efficiency of the algorithm.

As a test of this system, the NM WAIDS team used the new approximate string comparison algorithm to compare strings representing the entity, “names of oil producers in New Mexico,” in different data sets and to link corresponding records with semantically equivalent but not textually identical strings. There were more than 4,000 records to process in the target data set, and over 200,000 records in the source data set. Results of the new algorithm showed a very good improvement. Table 2 shows examples from two data sets in question and matching results based on the new algorithm, compared with those using the previous character-based or word-based algorithms. Character-based algorithms are represented by an edit distance algorithm and word-based algorithms by basic field matching. From this table, it is easy to see that the new algorithm improves the degree of string matching by including more string matching situations. The new algorithm also includes a new method of calculating the degree of match.

Since approximate string-matching algorithms are widely used in name and address matches, examples are taken from these areas to test the accuracy of the new algorithm. Table 3

shows the string comparison results for some sample string pairs. From this table, one can see that the new algorithm generally gets higher matching result scores, especially for strings containing abbreviations. Examination of the two tables reveals that.

- Edit-distance algorithms are best suited for comparing two one-word strings and multi-word strings without out-of-order problems;
- Edit-distance algorithms produce poor matching results for strings having out-of-order problems and abbreviations;
- Basic field matching is suitable for comparing two strings having all atomic strings being character strings, and numerical strings with the same format and scale;
- The new algorithm produces good results for strings with character strings, numerical strings and mixed strings;
- Both the basic field-matching algorithm and the new algorithm have difficulty with strings containing spelling problems, especially for one-word strings with spelling errors. The new algorithm appears to handle multi-word strings with spelling errors more accurately than the other algorithms.

Even with application of this improved string-matching algorithm, ambiguous cases were found where the computer alone could not make the link between data from one source and another. These cases were all flagged and examined by hand. In many cases, it was possible to correctly link the data. There were, however, a number of instances where there was simply not enough information associated with the water quality report to accurately link the water data to important identifying information from our other databases. Such an example would be a water report that identified the sample source as State Com #1, or Sample #23, or Tank Battery SJ 30-6. These cases were not included in the final database.

Data Quality Control

One additional phase of data quality control was that of checking the accuracy of the ADC system. Although the operators of the system tried to be conscious of checking data accuracy, it was felt that an additional check on data quality was necessary, particular for the numeric water chemistry data. One of the best ways to check data accuracy was to examine the relationships in major element chemistry. Most water sample reports have data presented in two different units (Fig. 4). This might be in parts per million (ppm), milligrams per liter (mg/l), milligram equivalents per liter (me/l) or equivalents per million (epm). Equivalents per million, a unit of measurement involving the number of ions, is often used in studies of chemistry of natural waters and in the interpretation of analyses. In waters of low salinity, the unit epm is numerically the same as the unit milligram equivalents per liter (me/l). For practical purposes, they can be considered identical. Concentrations expressed in units of weight as parts per million (ppm) are sometimes desired for a particular purpose¹³ and are a very common reporting unit. Conversion factors for some common anions and cations are included in Table 4. The equivalent weight of chloride is 35.5; thus 5 epm of chloride is the same as 177.5 ppm and 1 ppm chloride = 0.0282 epm.. These linear relationships can be used to check data quality. Figure 4 shows graphs of chloride and magnesium reported in ppm vs epm for some of the scanned data; similar checks can be run on all other data.

Figure 5 shows that most reported data points lie on or very close to a line whose slope corresponds to the conversion values determined by ppm/epm. Points that lie significantly away

from this line are considered suspect. These data points are flagged and checked against the scanned images of the data forms. In approximately half the cases, the error was found to be in the conversion of the image to text. The most common conversion error occurred in cases where the OCR program could not distinguish between a comma and a decimal point. In the other half of the cases checked, the OCR conversion was correct, and the problem lies in the actual data itself. A decision was made to keep the data in the database, but to add the notation to that particular data field that the data itself may have an error. In this way, the decision to use the data is left up to the individual database user.

The Produced Water Quality Database now contains approximately 7500 records. Of these, about 4300 are from southeast New Mexico. The remainder comes primarily from the San Juan Basin and southern Colorado. Although these data are from wells outside the study area, strong industry support compelled us to add this data into the database.

GIS Construction

Construction of the NM WAIDS Geographic Information System began in the first year of the project. Layers were created for the pilot area, including layers showing roads, metropolitan areas, rivers and streams, groundwater well locations, and produced water well locations. Once the water quality databases were created and cleaned, these too were added to the GIS. The one significant problem that remained with the construction of the initially-proposed GIS is that of infrastructure—pipelines and handling facilities. Although source data was once available on a public web server, access to this data has been restricted to registered operators and employees of various government agencies for security reasons. The initial GIS was created using ArcMAP, a program that allows users to manipulate data and make maps on their own computers. After the GIS was constructed, it was transferred to a server where it can be accessed publicly on the Internet using ArcIMS technology.

Corrosion Management Toolkit

A third key area of the NM WAIDS project is the Corrosion Management Toolkit (CMT). Although produced water in itself is not corrosive, the interaction of water with atmospheric or injected gases and other liquids can cause many corrosion and scale problems. The NM WAIDS team determined that the toolkit would consist of the following components:

1. **Reference Book**: Descriptions of common types of corrosion-related problems, including photos or drawings. This portion also includes possible mitigation approaches. The web version of this component also includes contact information and links to companies that provide these solutions in the Permian Basin area.
2. **Formation Water Analysis**: A qualitative analysis of produced waters in the region, by formation. Regional variations are noted, and if enough data is available, maps to graphically depict the variations are generated.
3. **Scale Prediction Tools**: Online tools that will allow the user to predict scaling tendencies of either a single water or a mix of waters. The tools allow users to input their own data, or they can request samples from the PWQD if they lack sample data. A simple online corrosion prediction tool was considered. However, this is a very complex and not easily-quantified problem; many chemical vendors have spent considerable time and effort

developing commercial software products to do this. For these reasons, we elected to devote our efforts to the scale prediction tool alone.

Reference Book

In the construction of the reference book, visits both to operators in the region and to the chemical vendors provided a number of examples, photograph, and drawings of corrosion and scale. Researchers were granted permission to use some of this material, both graphical and text, in the construction of the reference book. Considerable efforts were made to contact operators and chemical companies for information regarding corrosivity of specific formation waters through letters and phone calls. Responses were collected by asking those familiar with various formations to fill out a form with information about them.

Formation Water Analysis

Literature Summary of Geochemistry

Significant work was done in the analysis and interpretation of formation water chemistries of both our produced water and groundwater data. A literature search for existing studies comprised one of the first steps in understanding the formation water chemistry in the Permian basin. There are a large number of references concerning water geochemistry in the Permian basin, which can be loosely grouped into three categories: 1) older references describing the general hydrology of the upper aquifers and groundwater use in southeast New Mexico; 2) Detailed hydrologic and geochemical studies of the Waste Isolation Pilot Plant (WIPP) site, near Carlsbad, New Mexico, and 3) Geochemical studies of the Palo Duro basin, a basin with similarly aged rocks located in the Texas panhandle and separated from the Permian Basin by the Matador Arch. Remarkably little study has been done examining water chemistries of deeper formations in New Mexico.

References about groundwater and surface water in Lea, Eddy, and Chaves counties are quite numerous. However, many of these references are old, and most that concern our study area cover only the upper fresh water aquifer, not the deep brines. In addition, there is very little interpretation about the origin or history of waters in the basin. There are many geochemical studies of the WIPP site, near Carlsbad, New Mexico, that go into great detail about the Rustler and Salado Formations of the Ochoan series, including major and minor solutes and stable isotope data of the waters. Unfortunately the study area of these reports is much smaller than the NM WAIDS pilot area, encompassing only about 1600 km², an area including and surrounding the actual WIPP site.

Palo Duro Studies Aid Interpretation of NM Permian Data

Several good studies of water geochemistry in the Permian Basin in Texas, including the Palo Duro basin, located in the Texas panhandle were quite useful as interpretive aids, although they were outside of the pilot study area in southeast New Mexico. For example, Bassett and Bentley¹⁴ used drill-stem test (DST) data from petroleum wildcat wells to find pressure (head) and transmissivity of the saline deep basin aquifer, and found the hydrodynamics of the Palo

Duro basin showed a general decline of head from west to east, following the topographic dip from the Rocky Mountains. Studies by Fisher and Kreitler¹⁵ further analyzed the saline deep basin aquifer of the Palo Duro basin by using four wells drilled for the U.S. Department of Energy Salt Repository Program and two wells drilled by independent oil and gas companies. Samples from these wells were analyzed for major, minor, and trace ions, as well as isotopic compositions. Formation waters are Na–Cl brines that contain between 124-290 g/L TDS. They fitted their data to hydrodynamic models, which suggest that the basin has been completely flushed by meteoric water, and thus brine chemistry has evolved strictly through water-rock interactions. Chemical and isotopic compositions of the samples suggest two groups of waters: western Palo Duro basin samples, and eastern and central Palo Duro basin samples.

Stueber and others¹⁶ examined waters in carbonates in the San Andres-Grayburg, Wolfcamp, Pennsylvanian, and Devonian formations in the Central Basin platform in west Texas. They looked at chemical and isotopic signatures from these waters and found basically two groups of waters. The first group, loosely labeled saline meteoric water, includes samples from the San Andres Formation and Devonian limestones. This group has salinities of 26-59 g/L and $\delta D - \delta^{18}O$ values in the same range as modern precipitation and groundwater in the Ogallala aquifer. This water probably acquired its salinity from halite dissolution, shown by Na, Cl, and Br concentrations. The second group of water samples, called modified evaporitic marine brines, is from Pennsylvanian and Wolfcamp limestones. These samples are more saline, with salinities of 70-215 g/L and are apparently a mixture of two fluids, most likely highly evaporated seawater and saline meteoric water similar to the first group waters. These modified evaporitic marine brines were the dominant fluids in Paleozoic carbonates until the late Tertiary, when the tectonic uplift began 5-10 Ma, apparently causing meteoric water to flow into deeper strata. These samples were all taken from formations composed of carbonate, in the form of calcite or dolomite. The different groups are also delineated using TDS and ratios of some major ions, including Cl-/SO₄²⁻ and Ca²⁺/Mg²⁺ ratios. The two different Ca/Mg ratios may or may not have any significance, because all the samples were from carbonate formations. The Cl-/Na⁺+K⁺ ratio, in conjunction with the TDS, was also looked at. Although there is some separation between groups, it is mostly a function of the TDS; there is not enough difference between the Cl-/Na⁺+K⁺ ratios to separate the two groups. Stueber did not include HCO₃⁻ in his description, so that ion is ignored; however this seems reasonable considering the highly dynamic nature of both the pH and HCO₃⁻ concentration when a produced water sample is taken.

Analysis of Water Chemistry

NM WAIDS researchers found that it was possible to use a similar procedure for Permian Basin produced water analyses. A subset of our data, a database obtained from the USGS, was used for this work. Samples were analyzed by individual formation, and again, four distinct groups were seen. The first group corresponds with the first group described by Stueber and others,¹⁶ saline meteoric water. This group has a TDS of less than 75 g/L, a Cl/SO₄ ratio of less than 50, and a Ca/Mg ratio between 2-4. The second group also corresponds with the second group of Stueber et al., the modified evaporitic marine brines, and has a TDS of 125 g/L or above, a Cl/SO₄ ratio greater than 50, and a Ca/Mg ratio between 4 and 7. The third group is a mixture of the first two groups. It has a TDS between 75 and 125 g/L, a Cl/SO₄ ratio less than 50, and a Ca/Mg ratio between 4 and 6. The fourth group includes all the samples that do not fall into the other three. The next analytical step was to examine samples for regional variability in water chemistry, using the sample groupings derived from the above exercise.

Because the percentage of samples in the first three groups for each formation was often not high, it was decided to show groups based on the Cl/SO₄ ratio and TDS only (omitting the Ca/Mg ratio requirement) to augment the number of samples in each group. Once the data was grouped, the spatial variations were examined both in planimetric view using ArcView GIS and in three dimensions by plotting the data on to geologic cross sections.

Geologic Cross Sections

Three geologic cross sections were created based on water sample distribution patterns (Fig. 6). The data for the cross sections came from scout cards in the New Mexico Bureau of Geology Subsurface Library. These scout cards provided information such as well location, elevation of the well in feet above sea level, the depth of formation tops in feet from the surface, and some DST information regarding oil production. The location, elevation, and formation tops were used to create these cross sections. An attempt was made to get good data from one well in each section that a line of the cross section crossed (i.e., one well per mile.). Most wells did not have every formation top, so some interpolation was necessary where the data was lacking.

Once the cross sections were created, all produced water samples located ten kilometers or closer to each cross section were selected using ArcGIS, and then projected onto the closest cross sections in a N-S or E-W line, depending on the predominant direction of the cross section. The samples were plotted as depth versus latitude for the N-S cross sections, or depth versus longitude for the E-W cross section. Adobe Illustrator[®] was used to place the water quality data on top of the cross sections.

Three sets of data were projected on each cross section, so there are three figures per cross section. Only one of these sets, the east-west cross section through Lea and Eddy counties, is shown in this report. Figure 7 shows the reported formation of the produced water sample, which illustrates how well the water sample formations match the geologic cross section. Figure 8 shows the chloride concentration of the produced water samples, for comparison with the map view figures. Chloride concentration is used as a proxy for the distribution of TDS within the cross section. Figure 9 shows the interpreted origin of the produced waters.

Water Analysis Interpretation

Produced water samples were obtained from almost all formations (Fig. 10). In comparison with groundwater data analyzed earlier, produced waters typically had a much higher average chloride concentration than the groundwater samples. The piper diagrams for all formations except those in the Artesia group are similar and show halite dissolution (Figs. 11 and 12). The larger range of Artesia group samples in the piper diagrams is likely because of the samples with low chloride (higher water quality). Such samples are likely recent meteoric water and show the influence of the smaller amount of time for equilibration with local minerals.

Looking at the samples grouped by chemical ratio, high variability within formations and short sampling distances is evident. Thus, spatial trends are not as evident as they were for groundwater, even with large numbers of samples. However, it appears that there are more modified brines (group 2) with distance east, especially once over the Central Basin Platform, which is consistent with Dutton's¹⁸ findings of eastward regional flow.

A Question of Injected Water

Because of the large number of injection wells in southeast New Mexico, there was a question as to whether a study of produced water chemistry was even valid. Water injection occurs over the entire study area, (Fig. 6) in nearly all formations, with the exception of those formations with low porosities and permeabilities (“tight” formations). Two types of water injection wells are found in the area. A simple water injection well is used to keep pressure up within a formation, and usually injects water of similar composition to the water produced from that formation. A salt-water disposal well (SWD) is a deep well used to dispose of large amounts of highly saline brines, which may be of any composition. Most SWD wells inject into highly permeable formations that are thought to be well isolated from either hydrocarbon-producing zones or potential aquifers.

Water injection for pressure maintenance typically uses produced water from the same formation in order to prevent scaling within the formation from mixing incompatible waters. Nearly the same amount of water is injected that is produced (see Table 5); however, the composition of the injected water is usually somewhat different from the original produced water composition. There are several reasons for this difference in composition. The first is that there is often mixing of produced waters in a holding tank within an oil field; as producing and reinjection continues over time, the composition of water within a formation will tend to homogenize over distance. The second reason is that oil makes up a certain percentage of the produced water, and for injection into the ground, the lost volume of liquid is often made up by fresh water (often from the Ogallala Formation.) Again, this is to prevent possible scaling from mixing two incompatible brines; water from the Ogallala is fresh enough that the produced brine will only be diluted.

Although the large amounts of water that are produced and re-injected does mean that water cycling occurs (in which the water in a formation is replaced over time by injected water), the water remains compositionally similar to the original water. In addition, the study area is so large that regional trends will not be affected much; thus, it is valid to use the produced water to evaluate whether and where meteoric waters are moving in the deep basin.

Mapping Aquifer Thickness

One of the tasks for this project included the creation of a tool that could enable operators to evaluate the relative seriousness and degree of response required for a brine spill at a particular location. It has been determined that one of the most important parameters for this determination is the thickness of the aquifer at a given location (R. Hicks, pers. comm., 2004). This data has not been published in a form that is easily quantified and placed into a GIS system, so this became a necessary task for completion of the project. The primary aquifer in southeast New Mexico is the Ogallala, and possibly Dockum Formations of Quaternary and Tertiary age. The bottom of the aquifer is formed by the “redbed surface” described by Nicholson and Clebsh¹⁹ and a contour map of the redbeds in feet above sea level is found in Plate 1 of Nicholson and Clebsh.¹⁹ This map was scanned and converted to an image file, which was used for the rest of our aquifer thickness calculations. Contours were digitized using ArcGIS and the resulting digital contour plot was turned into a grid using “topogrid,” an Arc command that interpolates by a method similar to splining.

The top of the aquifer was found by mapping the water surface, which required several steps. The first step was to determine the depth to water using well data downloaded from the NMSEO website (<http://www.seo.state.nm.us/water-info/index.html>). It was necessary to

convert the data to latitude/longitude coordinates. The data was in feet below the surface, so in order to find the aquifer thickness, the land surface elevation was needed. This was obtained by use of a digital elevation map (DEM) provided by the New Mexico Bureau of Geology and Mineral Resources. Then, the depth to water was subtracted from the land surface grid and then a surface (grid) was interpolated using both kriging and inverse distance weighting (IDW) methods. Unfortunately, the resultant grids contained numerous “bull’s-eyes”—artifacts due to poor sample distribution because the wells were spaced quite far apart in the required area.

The rebed surface was subtracted from the water surface, and the resulting grid showed that a large percentage of the area had negative values, i.e., the rebed surface was above the water surface. Surfaces obtained by subtracting the water surface from the land surface (resulting in the thickness of the unsaturated zone) also produced negative values. A second data set with denser well spacing in the area delineated by the original rebed map was obtained. After this data set was cleaned by removing all “zero” data points, it was kriged to produce the depth of the unsaturated zone. When this grid was subtracted from the land surface to get the water surface, it was encouraging to see that the highest and lowest elevations of the water surface were greater than the highest and lowest elevations of the rebed surface.

Finally, the rebed surface was subtracted from the water surface to get the aquifer thickness. Unfortunately, although there was some improvement in the new grid, it was still about half negative values, implying that no aquifer was present. Looking at a grid of land surface – rebed surface (i.e., depth to rebed) shows that in areas, the rebed surface is quite near or even up to 20 feet above the ground surface in the original area of the contour map, so it makes sense that there would probably be more negative values in the aquifer thickness grid. Unfortunately, the data from the NMSEO is not sufficiently detailed in terms of defining what formation a particular well is producing water from. It is likely that in areas, the groundwater is probably coming from the rebeds or deeper formations rather than the Ogallala aquifer. Further research is necessary to determine the answers to this problem.

Formation Water Interpretation

During the later stages of this project, a tremendous amount of work was done on interpretation of both produced water and groundwater chemistry, and a M.S. thesis was generated from this work. The amount of work was far beyond the scope of this report; the thesis abstract is included here for reference purposes.

The objective of this project is to use existing groundwater and produced water databases to describe the groundwater chemistry of geologic formations, to map the geochemical distributions and trends of solutes, and to discover where, and in which formations, groundwater flushing (in which relatively fresh water moves through a formation, eventually replacing the original saline brine) is taking place. The two main databases used for this purpose include the State Engineer’s groundwater quality database, and the USGS database of produced water. Although trends are present, water quality in both databases is highly variable within formations and short sampling distances. Locations for three geologic cross sections across Lea and Eddy counties were chosen after sample distribution was mapped.

In general, shallow groundwater samples found in formations at or near the surface have low chloride concentrations, with the majority of samples having a chloride concentration of less than 1000 mg/L. Their high quality most likely reflects their origin

as meteoric water and short residence time within the aquifer. Where the number of samples is sufficient, spatial trends are usually evident, and are usually shown as decreasing chloride concentration with distance away from the Pecos River.

There are distinct chemical trends in the deep basin groundwaters, which are controlled by the flushing of meteoric water through high permeability formations. Stueber (1998, AAPG Bulletin, 82, 1652-1672) and Bein and Dutton (1993, GSA Bulletin, 105, 695-707) used stable isotopes, strontium isotopes, and Cl/Br ratios in order to differentiate connate brines from meteoric waters. Unfortunately I lack these types of measurements; however, the major ion data can be used to break out similar groups, or genetic classifications.

The work of previous investigators, together with hydrogeological data gathered in the course of this project, has indicated that several patterns of water movement should be expected. These include 1) eastward regional flow, 2) relatively high flow through highly fractured carbonates such as the reef zones in the Capitan and Abo, and regionally extensive carbonates such as the Mississippian through Ordovician, 3) more intermediate flow rates through carbonates with interbedded shales such as the Pennsylvanian and Wolfcamp, and 4) low flow rates through formations with variable lithology including carbonates, evaporites, redbeds, and shales, including the Artesia Group and Upper Leonardian Formations. The Delaware Mountain Group and Ochoan formations, composed of low permeability fine-grained sandstone, and evaporates respectively, are expected to have very low flow rates and briney waters. Although there is a large amount of variability between samples, these trends are born out in general. A major exception is the presence of brines within the Ordovician despite its carbonate composition; this is because the Ordovician is cut off from the recharge zone by a major fault zone (the Central Basin Platform) and is not vertically connected with the upper formations.

The Permian Basin in southeast New Mexico is complex both geologically and hydrologically. The basin lithology and history, combined with the interaction of groundwater as it moved through the deep basin aquifer through geologic time, has greatly influenced the chemical characteristics of waters within the basin in a reasonably consistent and predictable fashion. The uplift and eastward tilting of the area in the late Tertiary and Quaternary is likely the cause of much of the chemical distribution of the waters that we see today, although it is impossible to tell the timing of groundwater flushing without more detailed chemical analyses and modelling. (Davidson, N.J.R.: "Groundwater and Produced Water Quality of the Permian Basin, Southeast New Mexico," M.S. thesis, NMIMT. Dec. 2003.)

Scale Prediction Tool

One of the goals of the project was to help producers understand and predict when and where corrosion and scale might be problems. Corrosion prediction was beyond the scope of this study; however a useful scale prediction tool was developed as part of our work.

Methodology

All minerals are soluble in water to a limited extent and solubility calculations are used to predict scale formation. The limit of a mineral's solubility is called the solubility product, K_{sp} . Most minerals are salts and will dissociate into ions, for example $\text{NaCl} \leftrightarrow \text{Na}^+ + \text{Cl}^-$. Multiplying the concentrations of the ions together will give the ion product (IP). If the solution is saturated with that mineral and in equilibrium, the ion product will equal the solubility product K_{sp} .

The Saturation Index (SI) is the logarithmic ratio of the ion product and the solubility product, $\text{SI} = \log (\text{IP} / K_{sp})$. In other words, the SI is the log of the actual amount of mineral-forming ions over the solubility of that mineral. Thus, a saturated solution (in equilibrium) will have a SI of 0, an undersaturated solution will have a negative SI, and a supersaturated solution will have a positive SI. In addition, the saturation index has a logarithmic scale. For example, a solution with a SI of 3 is 10 times more oversaturated than a solution with a SI of 2. It is important to remember that a positive SI does not necessarily mean that scale will form, since the kinetics of scale formation may be too slow; rather, it is an indicator that formation is possible. The solubility constant for any salt will vary with pressure, ionic strength, temperature, and possibly pH. All salts increase in solubility as the pressure is increased, because when a salt is dissolved in water, there is a decrease in volume. As the ionic strength is increased, solubility increases up to a point when the water simply cannot hold any more salt, and then decreases. The effect of temperature on solubility differs with different salts.

In the oilfield, the most common scale-forming minerals are calcium carbonate (calcite), calcium sulfate, usually in the form of gypsum, and barium and/or strontium sulfates. Both barium and strontium have very low solubilities; thus, when barium or strontium and sulfate are present together in water, scale formation is very likely. At atmospheric pressure, barite is about 20 times less soluble than calcite, which is in turn almost 500 times less soluble in water than gypsum.¹⁷ However, barium and strontium are not usually major species in natural waters, so calcium carbonate and sulfate scales are much more common.

Many methods of predicting scale exist in the oil industry. Our tool uses two industry standards, the Stiff-Davis and Odco-Tomson methods.

Stiff-Davis Method

The Stiff-Davis method is one of the easiest ways to calculate calcium carbonate scaling tendencies (Calcite Saturation Index) in brines.²⁰ It is valid between temperatures from 0–100°C (32–212°F) and ionic strengths from 0–4. Inputs needed are pH, alkalinity, calcium concentration, and ionic strength. The Stiff-Davis method is very simple, but it may not be accurate if the pH is not measured immediately at the sample site. In addition, it does not take into account the total pressure or amount of dissolved or undissolved carbon dioxide gas.

Skillman/McDonald/Stiff Method

The SMS method predicts the solubility of gypsum scale (NOT the Saturation Index OR the total possible scale formed),²¹ and is valid between temperatures of 10–80°C (50–176°F) and ionic strengths from 0–6. Inputs needed are ionic strength, and sulfate and calcium ion concentrations. It does not take pressure into account, or any barium or strontium concentrations, which would most likely precipitate out barite (BaSO_4) or celestite (SrSO_4), reducing the available sulfate ion concentration. The solubility constant, K_{sp} , is calculated using the graph of K_{sp} versus temperature and ionic strength.¹⁷ The SMS method can also be used to predict the solubility of barite and celestite scale, if the K_{sp} variation with temperature and pressure are known.

Oddo-Tomson Method

The Oddo-Tomson method is another way of predicting the formation of calcium carbonate and various sulfate scales by calculating SI values.²² It is valid between temperatures of 0–200°C (32-392°F), ionic strengths of 0–4.0, and pressures of 1–1380 bar (0-20000 psig).¹⁷ It calculates different saturation indices for the different types of calcium sulfate minerals, including gypsum ($\text{CaSO}_4 \cdot 2 \text{H}_2\text{O}$), hemihydrate ($\text{CaSO}_4 \cdot \frac{1}{2} \text{H}_2\text{O}$), and anhydrite (CaSO_4). Gypsum is the most common scale former; it occurs at relatively low temperatures. Above about 100°C (212°F), anhydrite is the stable phase; however, hemihydrate can form in temperatures ranging from 90 to 120°C. The Oddo-Tomson method can also predict the formation of barium and strontium sulfate scales. User inputs needed are chemical analysis (including calcium, barium, strontium, bicarbonate, carbonate, and sulfate ions), temperature in °F, pressure in psia (psig + 14.7), and mole percentage of carbon dioxide in the gas phase or, if there is no gas phase, the amount of dissolved carbon dioxide in the water. If the amount of carbon dioxide is unknown but there is an accurate pH measurement, the method uses the pH to calculate the saturation indices.

The Oddo-Tomson method is more accurate than the Stiff-Davis method because it takes pressure as well as temperature and ionic strength into account. In addition, the method does not require a pH measurement, but calculates the pH based on the amount of carbon dioxide gas and bicarbonate in the water. This allows a greater accuracy in calculating the actual saturation index of a water sample, since pH measurements decline in accuracy very quickly after the sample is taken out of its natural environment.

Several web pages were created to make the prediction tool available on the Internet. The web pages consist of the Stiff-Davis Calculation Page, the Oddo-Tomson Calculation Page, and the Mixing Calculation Page. The Stiff page uses the Stiff-Davis method for SI calculation for calcite, and the SMS method for solubility of gypsum and barite. It also calculates the total possible scale amount for gypsum and barite using Eq. 3.16 of Patton (total calcite scale is calculated using a modified version of this equation).¹⁷ The calculations provide the user with both graphs or tables of calcite SI, and gypsum and barite scale solubility, or total possible scale vs. temperature. Each graph has different temperature points because the original solubility data was collected at different temperatures. When using this page, it must first be specified whether the user would like graphs of the solubility or total possible scale. Regardless, the first graph will be the calcite SI vs temperature.

The Oddo-Tomson page uses the Oddo-Tomson method to predict saturation indices for calcite, gypsum, barite and celestite, taking temperature, pressure, and ionic strength into account. This page can also calculate pH values, if one inputs the mole fraction of CO_2 in the gas phase or, if no gas phase is present, the concentration of dissolved CO_2 in mol/L in the water sample. If the pH value is known to be accurate, the measured pH can be used; however, using a measured pH is discouraged because the method was developed partially as a way to calculate saturation indices (and pH) without a measured pH value. This is because changes in temperature and pressure as the water sample is taken from a formation often leads to degassing, which in turn leads to inaccurate pH measurements, even if measured as soon as a water sample is taken.

Mixing of formation waters will often result in some scale formation, whether mixing different waters in a holding tank or injecting incompatible water into a formation. The mixing page will allow a user to mix waters with a known chemical analysis. Users can specify the ratio that the waters are mixed at, and the program will assume complete mixing, resulting in the maximum amount of possible scale. (A mixing ratio of 1:1 will give the maximum possible

scale.) The result will be the chemical composition of the mixed water, before any precipitation occurs. This data can then be plugged into the Oddo-Tomson Page to give an idea of what scales will form, and in what amounts.

In all the pages, users can either input their own chemical analyses, or they can first search the database for water samples that can then be used for any of the calculations. This would allow a person to select water analyses from location near their area of interest or from the same formation for use in the calculations.

NM WAIDS Web Site

A substantial amount of effort in the later stages of the project was devoted to the creation of the NM WAIDS web site, the culmination of the project's effort. The site is now fully operational; it is available to the public at <http://octane.nmt.edu/waterquality>. The home page is shown in Fig. 13. The web site is divided into four major parts: Databases, Maps, Tools, and Manual.

Databases

The Database portion of the web site allows access to the Produced Water Quality Database (PWQD) and the Groundwater Database (GD). The PWQD is searchable by section, township, range, well name, field, formation, and county (Fig. 14), while the GD allows for search by section, township, range, formation, sample date, and chloride content. The database query returns a list of samples that match the search criteria (Fig. 15), and then the user can select the sample(s) they wish to see more information for by clicking on a hyperlink. Figure 16 shows the page that is returned by this action. In this case, several samples were available for the selected well. Users can either select an individual sample and view more details about sample chemistry (Fig. 14), use in a Stiff Davis or Oddo Tomson scale calculation (Figs. 18 and 19), or select that sample to use in a water mixing calculation (Fig. 20). Stiff Davis and Oddo Tomson calculations were discussed in detail a previous report²³; these are two different methods of calculating the likelihood of scale formation from water of a given chemistry. Users can also see a summary of water volumes produced by a selected well for the past three years, and they can follow links to obtain more detailed information water production or injection that is available from other databases.

Tools

The Tools section of the NM WAIDS site is designed to provide access to several of the calculation and conversion tools (Fig 21). The calculation tools are the same ones that are used in the Data section to perform scale and mineral composition calculations; however when accessed via the Tools menu, users will need to use their own data for sample chemistry. The units conversion tools provide a number of different calculators for converting units of temperature, volume, concentration, pressure, distance, and mass.

Maps

The Maps section of the NM WAIDS web site currently displays layers for both produced and groundwater samples along with relevant cultural data such as section, township, and county boundaries. Figures 22–25 show screen shots from the online map section. Map users can select one or multiple points in either the groundwater or produced water layers, and by using either the Identify or Hotlinks tools, the databases are queried for detailed information about the selected samples. The map section also includes the fuzzy risk analysis tool, simple tools for estimating depth to groundwater and distance to water sources, and material concerning water quality for both produced and groundwater databases.

Manual

Figure 26 shows a screen shot of the entry page for the *Online Corrosion Manual*. The manual has been organized into sections on general corrosion theory, influence of gases on corrosion, equipment-specific types of corrosion, corrosion prevention, and references. The theory pages have basic information and theory on the main types of corrosion and why they occur. Subheadings include uniform, galvanic, crevice and pitting, hydrogen damage, environmentally induced, and erosion corrosion. The gases pages discuss specific gases in the oil field that can cause corrosion, including O₂, H₂S, and CO₂, as well as microbially influenced corrosion. Each section has expanded theory on the different corrosion types, occurrences, pictures of corrosion, and prevention and/or mitigation techniques. The equipment pages have information on the specific types of corrosion in oilfield equipment. Included are pages for artificial lift wells, casing and tubing, surface equipment and enhanced oil recovery operations. The prevention page has information on cathodic protection, coatings, and chemical corrosion inhibitors, with their different types and uses. A link to an interactive, ArcIMS map of New Mexico shows areas of prevalent corrosion environments such as hydrogen sulfide in large areas of southeast New Mexico, and carbon dioxide corrosion in northwest New Mexico. Finally, there are numerous references provided to help guide users to more detailed corrosion prevention information and suppliers of products to help mitigate oilfield corrosion.

Web Pages

From the entry page for the *Corrosion Manual* (Fig. 26), one can link to sections on corrosion theory, corrosion as related to the oilfield, corrosion mitigation, and links. The Theory section provides a more scientific and detailed look at corrosion, with links to pages discussing various types of corrosion including uniform, galvanic, crevice and pitting, hydrogen damage, environmentally induced cracking, and erosion corrosion (Fig. 27). There are many unique environments in the oil field where corrosion occurs. Some of the most common corrosive gases in the oil field are carbon dioxide and hydrogen sulfide, which form weak acids in water, and oxygen, which is a strong oxidizer. Microbial activity may cause corrosion alone, create more corrosive gases, and/or act to induce blockage within pipes. The Oil Field Practice page contains links to pages that discuss the effects of these factors, where they commonly occur (such as wells, tanks, separation facilities, and flow lines), and what common types of corrosion occur with these present (Fig. 28). The Prevention section discusses several types of prevention techniques in general terms, including coatings, cathodic protection, and chemical inhibitors (Fig. 29). A separate page has been included for the process of scale formation. This page (Fig. 30) covers the common scale types seen in the oil field, and several methods of scale prediction, including the Stiff Davis, Skillman McDonald Stiff, and Oddo Tomson methods. The scale

prediction and mixing tools are explained, and there are links to these pages. References are included. Finally, there is a Links section that includes two groups of links; “sponsored links” are links to companies that have let us use pictures, presentations, books, or other useful information, on the corrosion website. “Other links” are links to websites with a great deal of helpful information on corrosion or that specialize in certain aspects of corrosion control but that did not contribute to the website. The Help section (Fig. 31) contains a variety of material including some information on site organization, references used in the website, and a list of numerous books and journals that may be helpful to other people doing research about corrosion issues both in general and in the oil field industry.

Technology Transfer

Work done for this project has generated several papers and talks that were presented at international technical conferences. These papers and presentations include:

Wei, M., Sung, A., and Cather, M.: “Mining Spatially Abnormal Data in Spatial Databases,” presented at the 55th Annual Technical Meeting and 5th Canadian International Petroleum Conference, Calgary, Alberta, Canada, June 8–10, 2004.

Wei, M., Sung, A., and Cather, M.: “Data Quality Control for New Mexico Produced Water Chemistry Database,” presented at the 55th Annual Technical Meeting and 5th Canadian International Petroleum Conference, Calgary, Alberta, Canada, June 8–10, 2004.

Wei, M., Sung, A., and Cather, M.: “Detecting Spatial Outliers Using Bipartite Outlier Detection Methods,” 2004 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, June 22–24, 2004.

In addition to these major conferences, presentations about this project were made at several venues including the Southwest PTTC Workshop “Produced Water: Management and Issues” April 9, 2004 in Farmington, New Mexico, a presentation to the San Juan Basin Working Group in March, 2004 in Farmington, New Mexico, and a presentation to the Roswell Geological Society in April of 2005. A master’s thesis was also generated: (Davidson, N.J.R.: “Groundwater and Produced Water Quality of the Permian Basin, Southeast New Mexico,” M.S. thesis, NMIMT. Dec. 2003.). Several student independent studies also arose out of the work performed for this project.

Finally, a CD has been created that contains much of the same data and maps that are on the web site. The corrosion manual is also available from this CD. The only material that will not be on the CD are the scale calculation, water mixing, and units conversion tools. A simple GIS viewer that can be used with the data is available from the ESRI website, and instructions for downloading and using this program are available on the CD. To date, the greatest perceived benefit of this project has been the compilation of the tremendous amount of produced water data. Extensive interest has shown in this data, and several private companies are using the it for analyzing optimal locations to site their own produced water re-use projects.

CONCLUSIONS

In the first year of the project, a great deal of progress was achieved in all phases The

project team assembled databases of both produced water quality and groundwater quality, created a web-based data entry system that allows remote users to add data to the PWQD, and created an automatic data capturing system to allow for more accurate data collection. They obtained a great deal of produced water chemical data, contacted many Permian Basin operators on corrosion and other water-related issues, and created a web-based tool to predict scale type and probability based on two of the most commonly-used industry methods. They also began to assemble qualitative water quality information to be examined for variability by formation, region, and type of problem. Geographic coordinates were created for all sample locations with sufficient information and other pertinent information was assembled such as base maps for roads, metropolitan areas, political boundaries, geology, hydrology, and similar data. The NM WAIDS team also calculated aquifer thickness for the area, a possible key parameter in evaluation of brine spills.

Permian Basin operators showed themselves to be helpful and interested in the project. Yates Petroleum Corporation of Artesia has been particularly generous in the sharing of information, time, and source codes that were a partial basis for the scale calculation tools.

The greatest hurdles in the first year were encountered in assembling and entering data for the produced and ground water quality databases, and in cleaning up this data not only for accuracy's sake, but also to allow for integration with other data sets.

The second year of the project saw the creation and near-completion of the web site and assembly of the many GIS maps that were used for water analysis. Two different fuzzy tools that have been examined by various members of the NMOCD and the NMOGA Chlorides Working Group were deployed within the GIS (although neither of these tools will necessarily be the tool that is finally approved by the NMOCD and NMOGA). As the project entered its final period, data for produced water chemistry from outside the initial proposed study area were incorporated into the PWQD to make a more useful product. Minor adjustments and refinements needed to be made on the website

The New Mexico Water and Infrastructure Data System (NM WAIDS) is now offered to the public to help answer a number of produced water-related questions in southeast New Mexico. The website offers a Geographical Information System (GIS) and integral tools that will provide operators and regulators with necessary data and useful information to help them make management and regulatory decisions.

The major components of this system are: 1) databases on produced water quality, cultural and groundwater data, oil pipeline and infrastructure data, and corrosion information, 2) a web site capable of displaying produced water and infrastructure data in a GIS or accessing some of the data by text-based queries, 3) a fuzzy logic-based, site risk assessment tool that can be used to assess the seriousness of a spill of produced water, and 4) a corrosion management toolkit that will provide operators with data and information on produced waters that will aid them in deciding how to address corrosion issues. These elements of the NM WAIDS project are now all integrated into a website with a user-friendly interface that will provide access to previously difficult-to-obtain data and information.

REFERENCES

1. Jin, L., Li, C., and Mehrotra, S.: "Efficient Record Linkage in Large Data Sets," *Proc.*, Eighth Int. Conf.on Database Systems for Advanced Applications (DASFAA), Tokyo, Japan, (2003).
2. Porter, E.H. and Winkler, W.E.: "Approximate string comparison and its effect on an advanced record linkage system," *in* Proceeding of an International Workshop and Exposition - Record Linkage Techniques, Arlington, VA, USA, (1997).
3. FOLDOC, The Free On-line Dictionary of Computing:. URL:<http://www.foldoc.org/>, Editor Denis Howe.
4. Navarro, G.: "A Guided Tour to Approximate String Matching," *ACM Computing Surveys* (2001) **33**, No.1, 31 – 88.
5. Monge, A.E.: "Adaptive Detection of Approximately Duplicate Database Records and the Database Integration Approach to Information Discovery." PhD Dissertation, University of California, San Diego, California (1997).
6. Monge, A.E., and Elkan, C.P.: "The Field Matching Problem: Algorithm and Application," *Proc. Second International Conference on Knowledge Discovery and Data Mining*, Portland Oregon, (1996), p. 267-270, AAAI Press.
7. MacCallum, R.M.: "Computational Analysis of Protein Sequence and Structure," On-line document, URL:<http://www.sbc.su.se/~maccallr/thesis/thesis.html>, 1997.
8. Barton, G.J.: "An Efficient Algorithm to Locate All Locally Optimal Alignments Between Two Sequences Allowing for Gaps," *Computer Application in the BioSciences* (1993) **9**, No. 3, 729-734.
9. Monge, A.E.: "An Adaptive and Efficient Algorithm for Detecting Approximately Duplicate Database Records," On-line document, URL:<http://citeseer.nj.nec.com/monge00adaptive.html>, (2000).
10. Monge, A.E., and Elkan, C.P.: "An Efficient Domain-independent Algorithm for Detecting Approximately Duplicate Database Records," *SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery* (1997), Tucson, Arizona.
11. Winkler, W.E.: "String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage," *in* Proceedings of the Section on Survey Research Methods, American Statistical Assn., pp. 354-359, (1990).
12. Winkler, W.E.: "Advanced methods for record linkage." Technical report, Statistical Research Division, U.S. Census Bureau, Washington, DC, 1994.
13. Wilcox, L.V., and Magistad, C.C.: "Interpretation and Analyses of Irrigation Waters and the Relative Tolerance of Crop Plants," Bureau of Plant Industry, Soils and Agr. Engineering,

Agricultural Research Administration, U. S. Department of Agriculture, Riverside, California, (1943).

14. Bassett, R. L. and Bentley, M. E.: Geochemistry and Hydrodynamics of Deep Formation Brines in the Palo Duro and Dalhart Basins, Texas, USA. *Journal of Hydrology* (1982) 59, 331-372.

15. Fisher, R. and Kreidler, C.: Geochemistry and Hydrodynamics of Deep Basin Brines, Palo Duro Basin, Texas, USA. *Applied Geochemistry* (1987) 2, 459-476.

16. Stueber, A. M., Saller, A. H., and Ishida, H.: "Origin, Migration, and Mixing of brines in the Permian Basin: Geochemical Evidence from the Eastern Central Basin Platform, Texas. *AAPG Bulletin* (1988) 82, 9, 1652-1672.

17. Patton, C.C.: *Applied Water Technology*, Campbell Petroleum Series, Norman, Oklahoma (1986).

18. Dutton, A. R.: "Hydrogeologic and Hydrochemical Properties of Salt-Dissolution Zones, Palo Duro Basin, Texas Panhandle Preliminary Assessment," The University of Texas at Austin, Bureau of Economic Geology Geological Circular No. 87-2 (1987).

19. Nicholson, A. Jr. and Clebsch, A. Jr.: Geology and Ground-Water Conditions in Southern Lea County, New Mexico: Ground-Water Report. Technical Report 6, State Bureau of Mines and Mineral Resources, Socorro, NM (1961).

20. Stiff, H.A., and Davis, L.E.: "A Method for Predicting the Tendency of Oil Field Waters to Deposit Calcium Carbonate," *Transactions AIME* (1952) 195, 213-216.

21. Skillman, H.L., McDonald, J.P. Jr., and Stiff, H.A.: "A Simple, Accurate, Fast Method for Calculating Calcium Sulfate Solubility in Oil Field Brine," presented at the Spring Meeting of the Southwestern District, API, Lubbock, Texas, March 12-14, 1969.

22. Oddo, J.E., and Tomson, M.B.: "Why scale forms in the oil field and methods to predict it," paper SPE 21710 presented at the 1991 Production Operations Symposium, Oklahoma City.

23. Cather, M.E.: "NM WAIDS: A Produced Water Quality and Infrastructure GIS Database for New Mexico Oil Producers," annual report, Contract No. DE-FC26-02NT 15134, U.S. DOE, Washington, DC (March 2003).

LIST OF ACRONYMS AND ABBREVIATIONS

NM WAIDS – New Mexico Water and Infrastructure Data System
GIS – Geographic Information System
TDS – Total Dissolved Solids
NMSEO – New Mexico State Engineer’s Office
PWQD – Produced Water Quality Database
ADCS – Automatic Data Capture System
OCR – Optical Character Recognition
ADF – Automated Document Feed
ODBC – Open Database Connectivity
CSV – Comma Separated Variable
PLSS – Public Land Survey System
GCDB – Geographic Coordinate Database
CMT – Corrosion Management Toolkit
SMS – Skillman/McDonald/Stiff
IP – Ion Product
SI – Saturation Index
 K_{sp} – Solubility Product
DEM – Digital Elevation Model
WIPP – Waste Isolation Pilot Project
IDW – Inverse Distance Weighting
NMOCD – New Mexico Oil Conservation Division

Table 1. Sample List of Data from Two Data Sets

Pair	String From Water Quality Data Set	Matched String From State Source Data Set
1	#1 DICKINSON HEIRS	DICKINSON HEIRS 001
2	JENNING'S FEDERAL NO. 1	JENNINGS FEDERAL 001
3	FEDERAL-KEOHANE ETAL #2	KEOHANE ET AL FEDERAL 002
4	30-6 POW #3	SAN JUAN 30 6 UNIT POW 003
5	SJ 29-6 105	SAN JUAN 29 6 UNIT 105
6	GAMBLING 2A	GAMBLING 002A

Table 2. String Sample Pairs from Two Data Sets and Matching Results Sets

Pair	Strings from Water Quality Data Set	Matched Strings From Source Data Set	Edit Distance	Basic Field Matching	New Algorithm
1	#1 DICKINSON HEIRS	DICKINSON HEIRS 001	6	0.67	1.0
2	A-1 EAVES	EAVES A 001	9	0.67	1.0
3	E M HARTMAN 1	EM HARTMAN 001	3	0.57	1.0
4	C. M. FARNSWORTH B 5	FARNSWORTH B FEDERAL 005	14	0.44	0.67
5	FEDERAL-KEOHANE ETAL #2	KEOHANE ET AL FEDERAL 002	17	0.67	1.0
6	30-6 POW #3	SAN JUAN 30 6 UNIT POW 003	15	0.55	0.73
7	SJ 29-6 105	SAN JUAN 29 6 UNIT 105	12	0.8	0.9
8	GAMBLING 2A	GAMBLING 002A	2	0.5	1.0

Table 3. Name and Address String Sample Pairs and Result Comparison

Pair	Strings from Water Quality Data Set	Matched Strings From Source Data Set	Edit Distance	Basic Field Matching	New Algorithm
1	Colette Johnen	John Colette	12	1.0	1.0
2	Liu Hang Xiang	Liu H. X.	7	1.0	1.0
3	CalTech	California Institute of Technology	24	0.5	0.75
4	NM Tech	New Mexico Institute of Mining and Technology	31	0.57	0.71
5	600 113 th st. apt. 5a5	600 113 th st. ap. 585	2	0.6	0.8
6	SHACKLEFORD	SHACKELFORD	2	0.0	0.0

Table 4. Conversion Factors for Common Anions and Cations

Cation	Equivalent weight	Conversion Factor (1/equivalent weight)	Anion	Equivalent Weight	Conversion Factor (1/equivalent weight)
Calcium (Ca)	20	0.05	Carbonate (CO ₃)	30	0.0333
Magnesium (Mg)	12.2	0.08197	Bicarbonate (HCO ₃)	61	0.0164
Sodium (Na)	23	0.0435	Sulfate (SO ₄)	48	0.0208
Potassium (K)	33.1	0.0302	Chloride (Cl)	35.5	0.0282
			Nitrate (NO ₃)	62	0.0161

To convert ep_m to pp_m, multiply the concentration in ep_m by the equivalent weight. To convert pp_m to ep_m, divide the concentration in pp_m by the equivalent weight.

Table 5. Water Injection Statistics for Southeast New Mexico

Eddy				
		2001	2002	2003 (as of 7/21/03)
	Cum. Prod. H ₂ O (bbl)	172329258	166563993	64194936
	Cum. Injected H ₂ O (bbl)	152829075	146926205	63197886
	% of Cum. Prod. H ₂ O injected	89	88	98
Chaves				
		2001	2002	2003 (as of 7/21/03)
	Cum. Prod. H ₂ O (bbl)	7051815	6446153	2999047
	Cum. Injected H ₂ O (bbl)	6819553	6772747	2822281
	% of Cum. Prod. H ₂ O injected	97	105	94
Lea				
		2001	2002	2003 (as of 7/21/03)
	Cum. Prod. H ₂ O (bbl)	429351183	412605679	154455426
	Cum. Injected H ₂ O (bbl)	403614286	428160672	156906731
	% of Cum. Prod. H ₂ O injected	94	104	102
Total				
		2001	2002	2003 (as of 7/21/03)
	Cum. Prod. H ₂ O (bbl)	608732256	585615825	221649409
	Cum. Injected H ₂ O (bbl)	563262914	581859624	222926898
	% of Cum. Prod. H ₂ O injected	93	99	101

EL PASO NATURAL GAS COMPANY
SAN JUAN DIVISION
FARMINGTON, NEW MEXICO
PRODUCTION DEPARTMENT WATER ANALYSIS



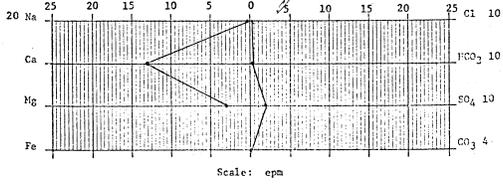
Analysis No. 1-9722 Date 10-24-79
 Operator ENPG Well Name SJ 30-6 8 A
 Location NW 31-30-7 County Rio Arriba State N.M.
 Field _____ Formation _____
 Sampled From 1417 W
 Date Sampled _____ By _____

Tbg. Press.	Csg. Press.	Surface Csg. Press.
ppm	ppm	ppm
Sodium <u>104</u>	<u>5</u>	Chloride <u>12</u>
Calcium <u>262</u>	<u>13</u>	Bicarbonate <u>127</u>
Magnesium <u>36</u>	<u>3</u>	Sulfate <u>875</u>
Iron _____	_____	Carbonate <u>0</u>
H ₂ S _____	_____	Hydroxide <u>0</u>

Total Solids Dissolved 1120
 pH 7.4
 Sp. Gr. 1.0021 at _____ °F
 Resistivity 580 ohm-cm at 77 °F

cc: D.C. Adams
 F.A. Ullrich
 E.R. Paulok
 J.N. McCarthy
 A.M. Smith
 W.B. Shropshire
 File
 80 Ft.
 C.Bo'Nan

Chemist



(a)

EL PASO NATURAL GAS COMPANY
SAN JUAN DIVISION
FARMINGTON, NEW MEXICO
PRODUCTION DEPARTMENT WATER ANALYSIS

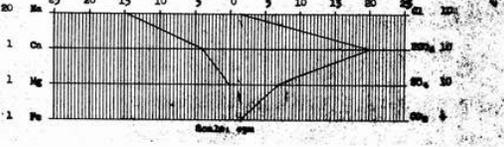
Analysis No. 1-3053 Date JULY 30, 1962
 Operator EL PASO NATURAL GAS COMPANY Well Name FERRIS NO. 2
 Location SW 11-30-7 County RIO ARriba State NEW MEXICO
 Field BLANCO Formation MESA VERDE
 Sampled From HEADHEAD
 Date Sampled JULY 5, 1962 By CONNELLY
 Tubing Pressure 930 Casing Pressure 1000 Surface Casing Pressure 246

ppm	ppm	ppm	ppm
Sodium <u>6585</u>	<u>286</u>	Chloride <u>595</u>	<u>17</u>
Calcium <u>80</u>	<u>4</u>	Bicarbonate <u>12150</u>	<u>190</u>
Magnesium <u>0</u>	<u>0</u>	Sulfate <u>3300</u>	<u>60</u>
Iron _____	_____	Carbonate <u>170</u>	<u>6</u>
H ₂ S _____	<u>Absent</u>	Hydroxide <u>0</u>	<u>0</u>

Total Solids Dissolved 12260
 pH 8.75
 Sp.Gr. 1.015 at _____ °F
 Resistivity 68 ohm-cm at _____ °F

cc: E.E. Alsup
 L.M. Parrish, Jr.
 J.E. Ashworth
 E.S. Oberly
 L.D. Galloway
 R. Fritchard (c)
 R.L. Ahrens
 A.H. Viesens
 file

Chemist



(b)

Fig. 1. (a) Example of a good, clean water analysis form that is easy to convert to text. (b) Example of a poor quality water analysis form. Note the fuzziness of the type and the many black specks on the page. These make OCR more difficult.

Editor of Template 1

Template Blocks View Image Tools Help

Close

EL PASO NATURAL GAS COMPANY
 SAN JUAN DIVISION
 FARMINGTON, NEW MEXICO
 PRODUCTION DEPARTMENT WATER ANALYSIS

Analysis No. 1 1-8572 Date 2 October 01, 1975

Operator 3 E.P.N.G. Well Name 4 Horton #6

Location 5 79 County 6 State 7

Field 9 Formation 8 M. V.

Sampled From 10 Bradenhead

Date Sampled 11 09-23-75 by Thomas F. West

Tubing Pressure 12 Casing Pressure 13 Surface casing pressure 14

	ppm	eppm		ppm	eppm
Sodium	153330	16 145	Chloride	231680	24 47
Calcium	17350	18 17	Bicarbonate	2515	26 0.2
Magnesium	1980	20 2.4	Sulfate	275615	28 117
Iron	21 Present		Carbonate	29 0	30 0

Fig. 2. Example of a form that has been scanned and processed. The numbered boxes are the data fields that are recorded in a database, while the dashed lines surround fields that are reference fields.

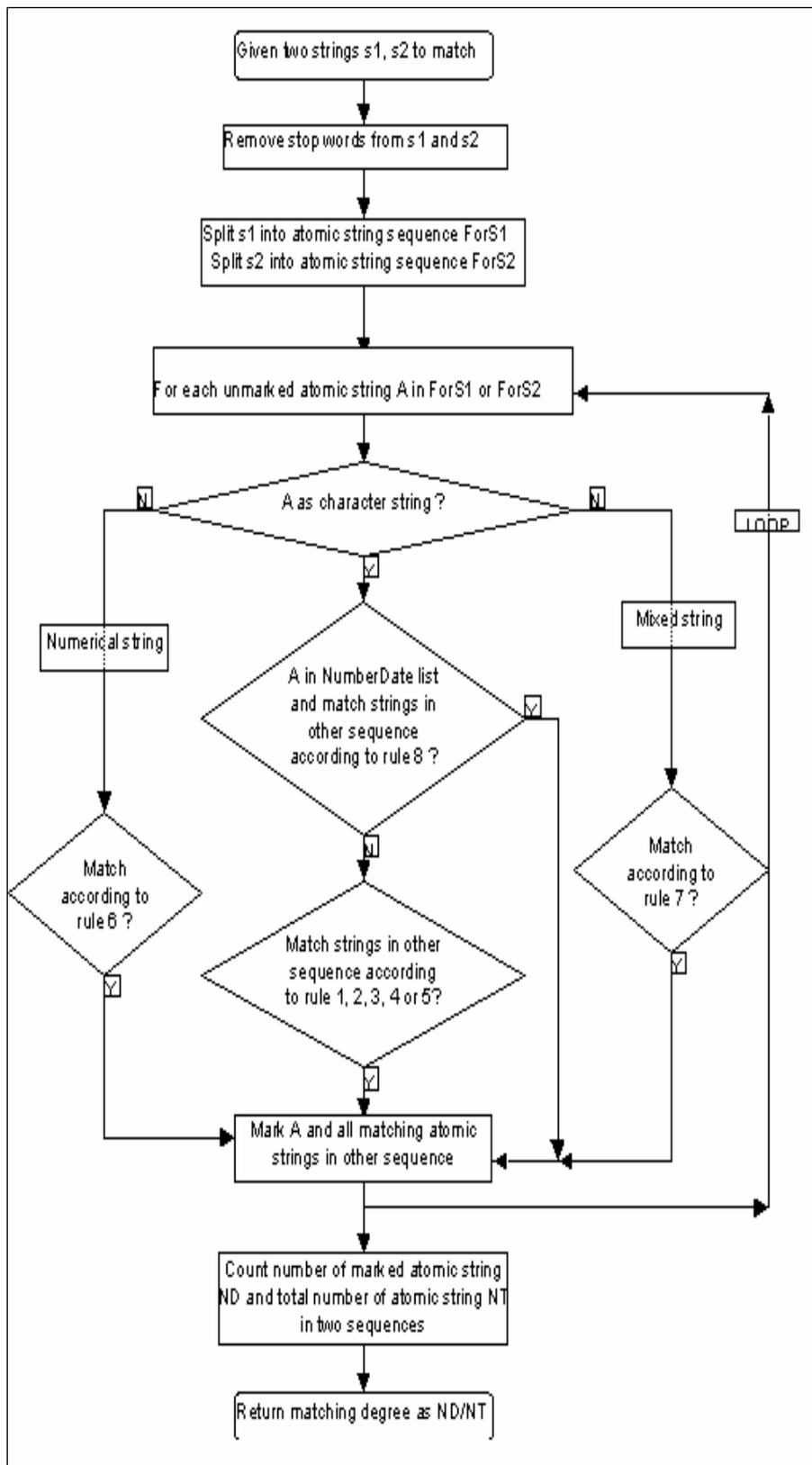


Fig. 3. Flowchart of the new string comparison algorithm used for cleaning data to be entered in the Produced Water Quality Database.

EL PASO NATURAL GAS COMPANY
 SAN JUAN DIVISION
 FARMINGTON, NEW MEXICO
 PRODUCTION DEPARTMENT WATER ANALYSIS

Analysis No. 1-8572 Date October 01, 1975

Operator E.P.N.G. Well Name Horton #6

Location 29-35 County _____ State _____

Field _____ Formation M. V.

Sampled From Bradenhead

Date Sampled 09-23-75 by Thomas F. West

Tubing Pressure _____ Casing Pressure _____ Surface casing pressure _____

	ppm	epm		ppm	epm
Sodium	<u>3330</u>	<u>145</u>	Chloride	<u>1680</u>	<u>47</u>
Calcium	<u>350</u>	<u>17</u>	Bicarbonate	<u>15</u>	<u>0.2</u>
Magnesium	<u>30</u>	<u>2.4</u>	Sulfate	<u>5615</u>	<u>117</u>
Iron	<u>Present</u>		Carbonate	<u>0</u>	<u>0</u>
H ₂ S	<u>Absent</u>		Hydroxide	<u>0</u>	<u>0</u>

Fig. 4. Water analysis form showing two reporting methods (ppm and epm) for the sample.

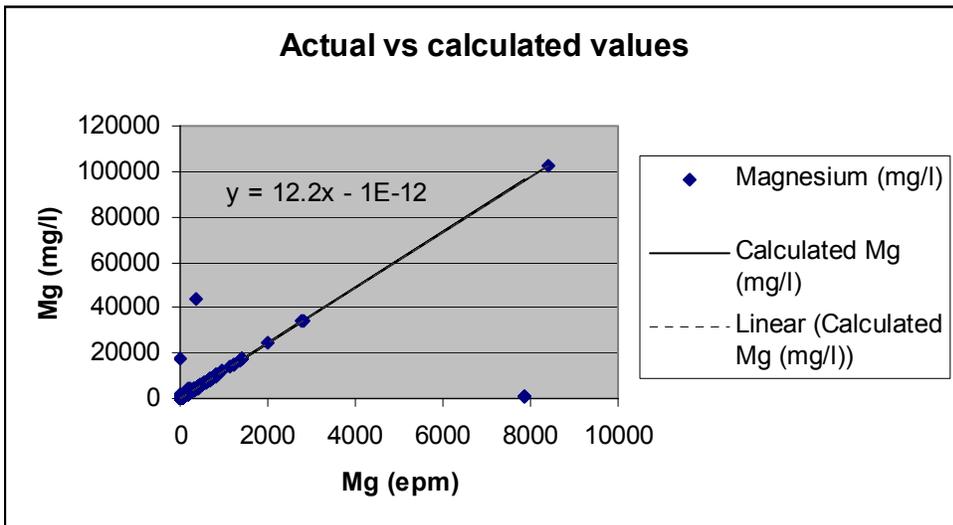
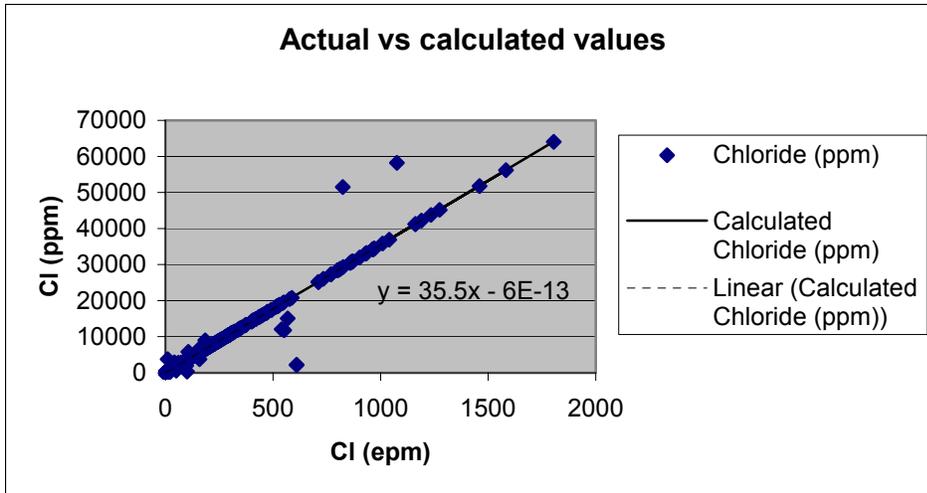


Fig. 5. Reported and calculated values of chloride and magnesium in ppm and mg/l vs epm units. These types of comparisons are very useful for quality control of scanned and hand-entered data. If the value of the number in epm units divided by that of ppm or mg/l units was less than half or greater than twice the known conversion factor (for instance, 1ppm = 0.0282 epm for chloride), the data was considered inaccurate and values were checked by hand against the scanned image. Note that the slope of the trend lines for calculated data is equal to the conversion factor given in Table 4

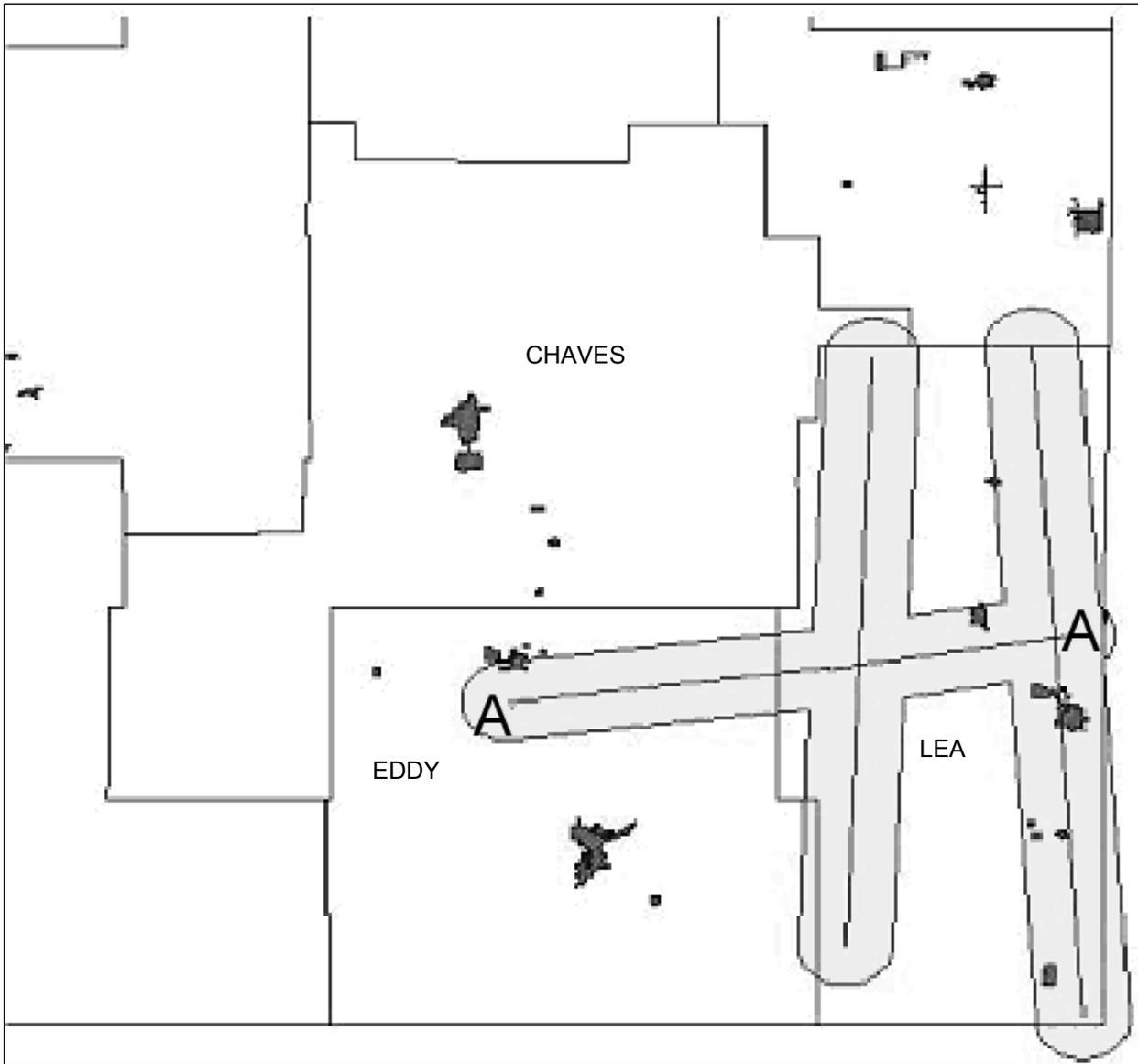


Fig. 6. Location of three cross sections and 10 km buffer around them. Produced water samples from wells that fall within the buffer zone were projected onto the closest cross sections in a N-S or E-W line, depending on the predominant direction of the cross section. Line A-A' marks the cross section line and buffer zone for the next three figures.

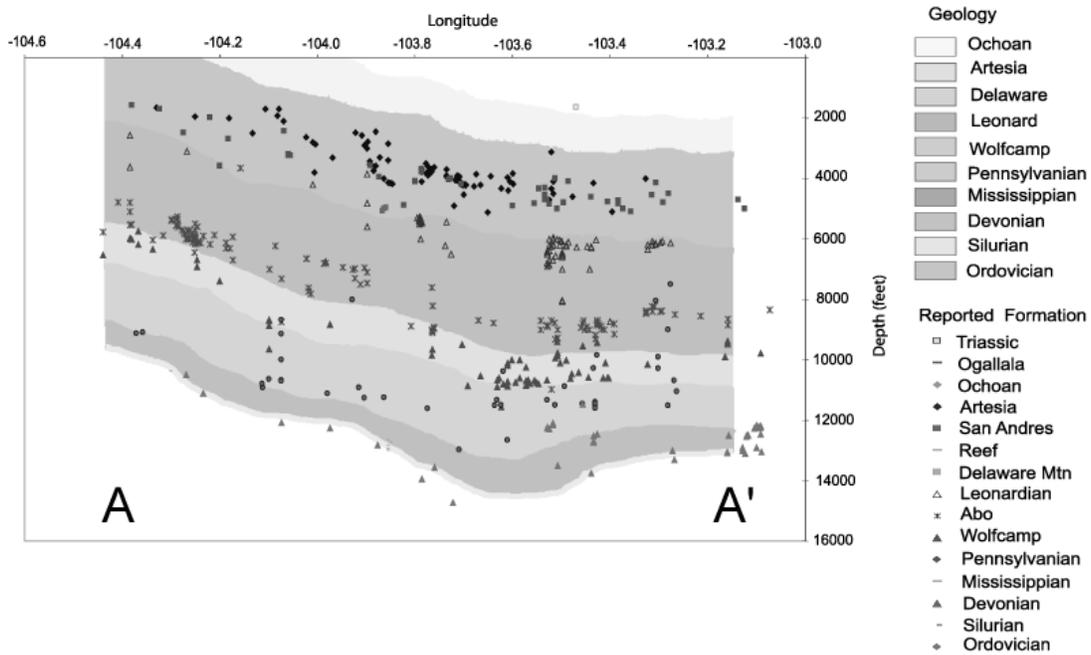


Fig. 7. Cross section through Eddy and Lea counties showing produced water samples by depth and reported formation.

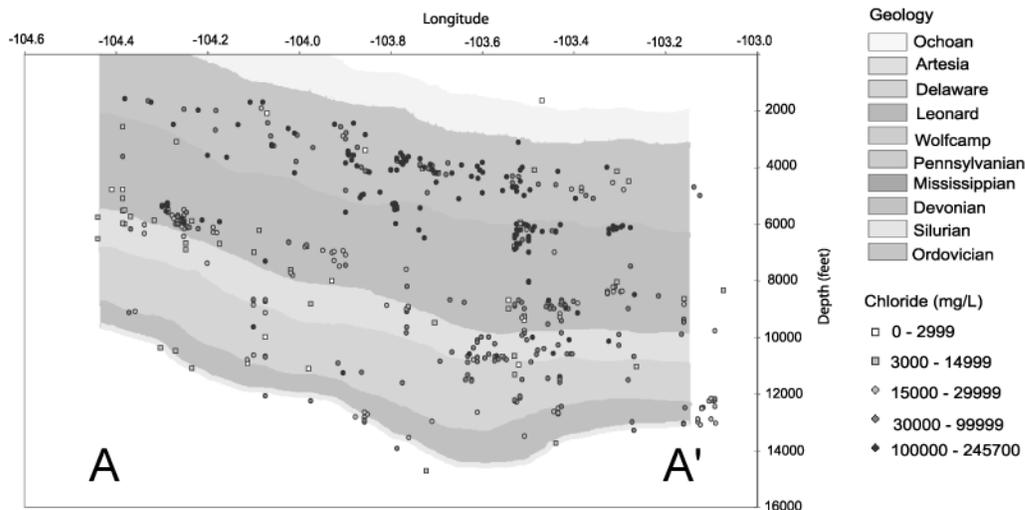


Fig. 8. Chloride concentration of produced water samples for cross section A-A'. Interestingly, some of the most saline waters are found at relatively shallow depths.

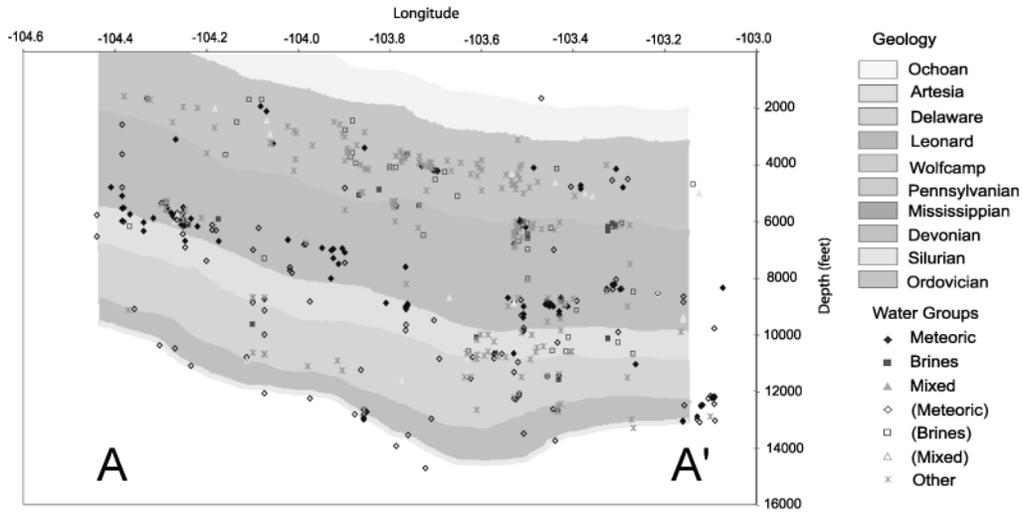


Fig. 9. Cross section A-A', showing interpreted origin of the produced waters. Solid points are based on Cl/SO₄ and Ca/Mg ratios, while open points are based on Ca/SO₄ ratios only.

Era Period Epoch or Age				Cross section divisions				
Phreatic Aquifer	Cenozoic	Tertiary	Quaternary	Playa Deposits Windblown Sand				
			Pleistocene	Alluvium Travertine Terrace Alluvium Berino Soil Mescalero Caliche				
				Galuna Formation				
				Ogallala Formation				
	Mesozoic	Triassic	Comanche Series	Dockum Formation				
				Chinle Group			Santa Rosa sandstone Dewey Lake Redbeds Formation	Triassic
	Evaporite Confining Layer	Paleozoic	Permian	Ochoan Series		Rustler Formation Forty-Niner Member Magenta Dolomite Tamarisk Member Culebra Dolomite Virginia Draw Member Upper Member McNutt Potash Member Lower Member Tessey Limestone Castile Formation Painhorse Member Anhydrite IV Member Halite III Member Anhydrite III Member Halite II Member Anhydrite II Member Halite I Member Anhydrite I Member Basal Limestone Member	Ochoan	
				Guadalupian Series		Tanhill Formation Yates Formation Seven Rivers Formation Queen Formation Shattuck Member Penrose Sand Member	Artesia	
				Artesia Group		Grayburg Formation	Ocotillo Silt Member	(San Andres)
				Reef		San Andres Formation Cherry Canyon tongue Sandstone Cutoff Shale	Reef	
Delaware Mtn. Group				Capitan Limestone Capitan Limestone Goat Seep Dolomite Bell Canyon Formation Cherry Canyon tongue Formation Brushy Canyon Formation	Forereef Member Reef Member Manzanita Limestone Member South Wells Limestone Member Getaway Limestone Member	Delaware Mtn Group		
Leonardian Series				Glorieta Sandstone Bone Spring Limestone Hueco Limestone Clear Fork Formation Yeso Formation Abo Formation	Padlock Blinobry Tubb Sand Drinkard Member	(Upper) Leonardian (Abo)		
Wolfcamp Series						Wolfcamp		
Virgil Series				Cisco Formation		Pennsylvanian		
Missourian Series				Canyon Formation				
Des Moines Series				Strawn Formation				
Atoka Series	Atoka Formation							
Morrow Series	Morrow Formation							
Deep Basin Aquifer	Paleozoic	Mississippian	Barnett Shale	Mississippian				
			Devonian	Woodford Shale (Thirtyone) Formation	Devonian			
			Silurian	(Wristen) Formation Fusseiman Dolomite	Silurian			
			Ortovician	Montoya Group	Ortovician			
				Simpson Group				
				Eilenburger Group				
	Cambrrian							

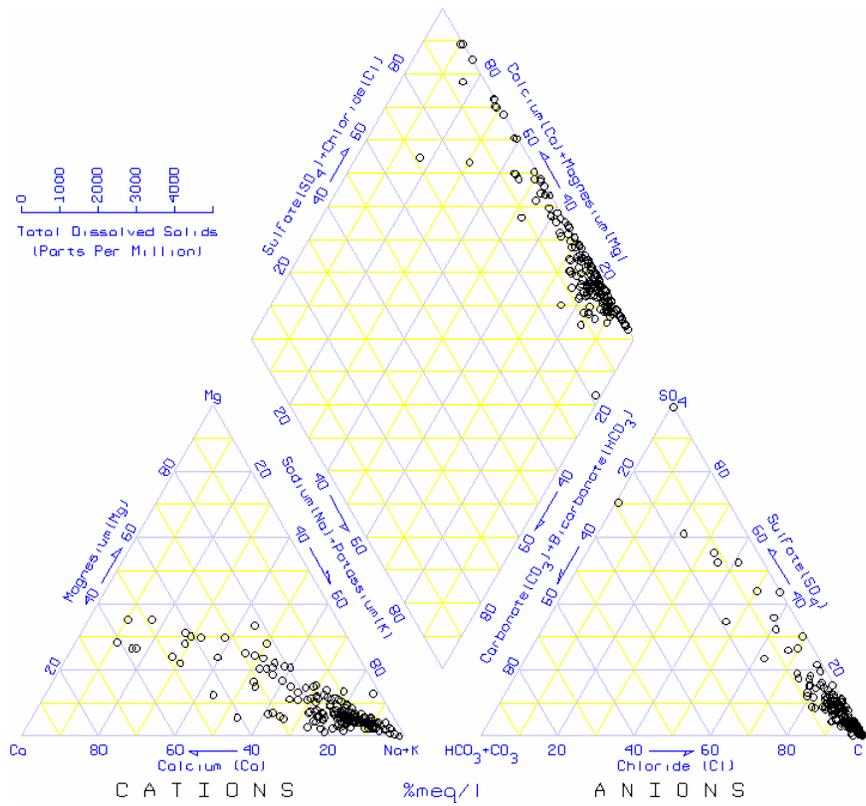
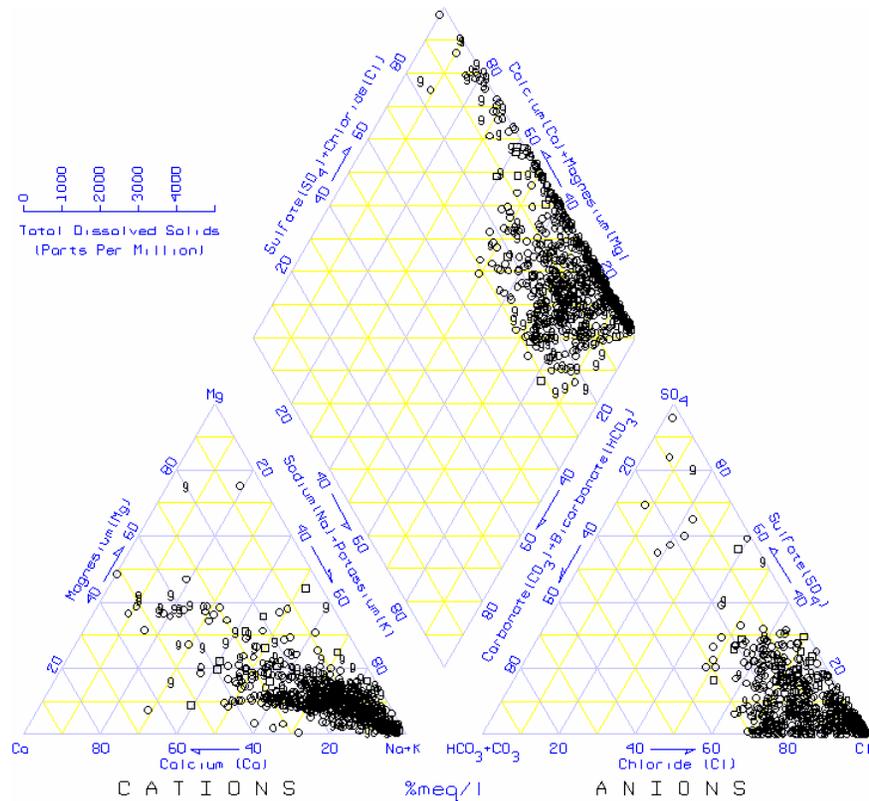


Fig. 11. Typical piper diagram of produced water samples, (Leonard) Abo shown.

Fig. 12. Piper diagram of San Andres and Grayburg produced water samples.



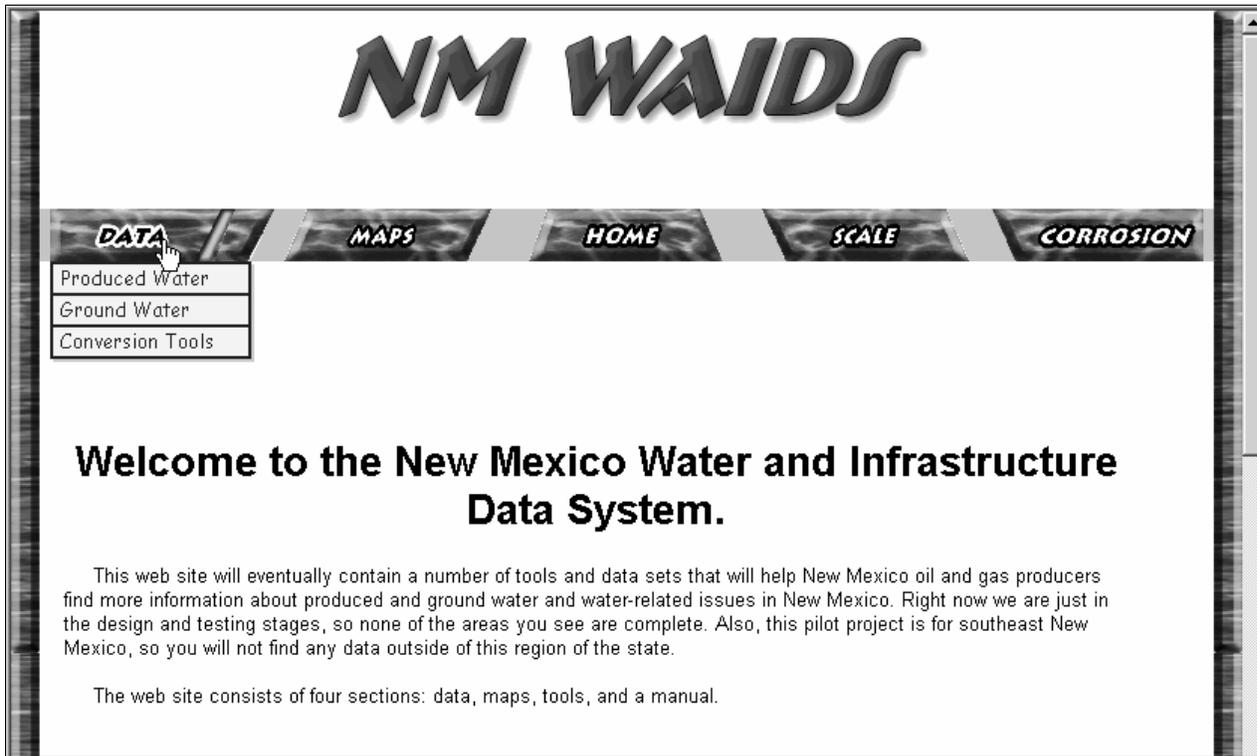


Fig. 13. Home page for NM WAIDS web site. Interface includes links to databases, the GIS map server, various tools for predicting corrosion and scale, and the online corrosion manual.

Produced Water Samples	
Water Sample Search	
API NUMBER	<input type="text"/>
SECTION	Not Specified <input type="button" value="v"/>
TOWNSHIP	Not Specified <input type="button" value="v"/> <input type="radio"/> North <input type="radio"/> South <input type="checkbox"/> Half
RANGE	Not Specified <input type="button" value="v"/> <input type="radio"/> East <input type="radio"/> West
WELL NAME	<input type="text"/>
COUNTY	Not Specified <input type="button" value="v"/>
FORMATION	Not Specified <input type="button" value="v"/>
FIELD	<input type="text"/>
<input type="button" value="Submit"/> <input type="button" value="Clear"/>	

Fig. 14. Search interface to Produced Water Quality Database.

Water Samples for Township 23 South

Instructions:
The number represents the number of water samples of certain well. Click the number if you want to download the data.

128 records are available.

#	API	WellName	S	T	R	County	Formation	Field
<u>1</u>	null	BRANTLEY GAS COM	22	23S	02E Eddy	MORRO	null	
<u>1</u>	3001500012	HUAPACHE UT	06	23S	22E Eddy	ABO	null	
<u>1</u>	3001500044	BANDANA POINT UT	13	23S	23E Eddy	DEVON	BANDA	
<u>2</u>	3001500044	BANDANA POINT UT	13	23S	23E Eddy	ELLEN	BANDA	
<u>1</u>	3001501114	PARDUE	12	23S	27E Eddy	DELAW	CASS	
<u>1</u>	3001503691	REMUDA BASIN UNIT	24	23S	29E Eddy	BEAR	REMUD	
<u>2</u>	3001503691	REMUDA BASIN UNIT	24	23S	29E Eddy	DEVON	REMUD	
<u>1</u>	3001505847	GARDNER TL	34	23S	31E Eddy	DELAW	null	

Fig. 15. Results for a search using T23S as search criteria. Note that some wells have more than one sample available. Clicking on the hyperlinked (underlined) numbers will bring up a screen like that seen below in Figure 13.

Water Samples for Well 'BIG FREDDY UNIT '
API = '3001510054'
Formation = 'MOR'
Field = 'INDIAN BASIN'

Instruction:
Click  For sample analysis.
Click  For general information about this sample.
Click  To select **this water sample** for water mixing. It will lead to the main page, and add the sample ID to mixing table.
The ions are in (mg/L) units.

SampleID	T	R	S	SO4	CL	CO3	HCO3	K	Na	Ca	Mg
6715 	23S	22E	10	25	67521	null	122	null	null	8680	443
6733 	23S	22E	10	2	72846	null	122	null	null	4880	197
6722 											

Fig. 16. Screen shot showing samples available for the Big Freddy Unit Well 002. Clicking on one of the icons will either show more information about the sample, will allow the user to do an Oddo-Tomson or Stiff-Davis scale calculation, or will select the sample to use in the water mixing program available in the Tools section of the web site.

General Information About: Sample 1296			
SOUTH LOCO HILLS UNIT006			
API	3001503448	Sample Number	
Unit/Section/ Township/Range	B / 20 / 18 S / 29 E	Field	LOCO HILLS
County	Eddy	Formation	ARTESIA
State	NM	Depth	
Lat/Long	32.73876 / -104.09520	Sample Source	well head
		Water Type	produced
Sample Date (MM/DD/YYYY)		Analysis Date (MM/DD/YYYY)	
Remarks/Description			
Cation Information (mg/L)		Anion Information (mg/L)	
Potassium (K)		Sulfate (SO)	1790
Sodium (Na)	57800	Chloride (Cl)	95000
Calcium (Ca)	950	Carbonate (CO ₃)	
Magnesium (Mg)	1740	Bicarbonate (HCO ₃)	410
Barium (Ba)		Hydroxide (OH)	
Manganese (Mn)		Hydrogen Sulfide (H ₂ S)	
Strontium (Sr)		Carbon Dioxide (CO ₂)	
Iron (Fe)		Oxygen (O)	1

Fig. 17. Screen showing water chemistry data for a selected sample.

Fig. 18. Screen shot of sample data entered into Stiff-Davis calculation tool.

Lease/Well	General Character		pH	7.8	Total hardness		Ionic Strength				
			Specific Gravity		Resistivity		Total Dissolved Solid				
Ions mg/L											
Ca ⁺⁺	950	Mg ⁺⁺	1740	Na ⁺	57800	Ba ⁺⁺	0	Fe ⁺⁺	0	Sr ⁺⁺	0
CO ₃ ⁼	0	HCO ₃ ⁻	410	SO ₄ ⁼	1790	Cl ⁻	95000	OH ⁻	0		
Stiff Davis Method											
Stiff Davis Calculate											
Graphics Choice											
<input type="radio"/> Solubility and Actual amount											
<input type="radio"/> S Index (Actual - Solubility, meq/L)											
<input type="radio"/> Total possible scale in mg/L											
<input type="radio"/> Total possible scale in PTB (A index)											
Switch to Oddo Tomson Method											
Calculate Probable Mineral Composition											
Clear											

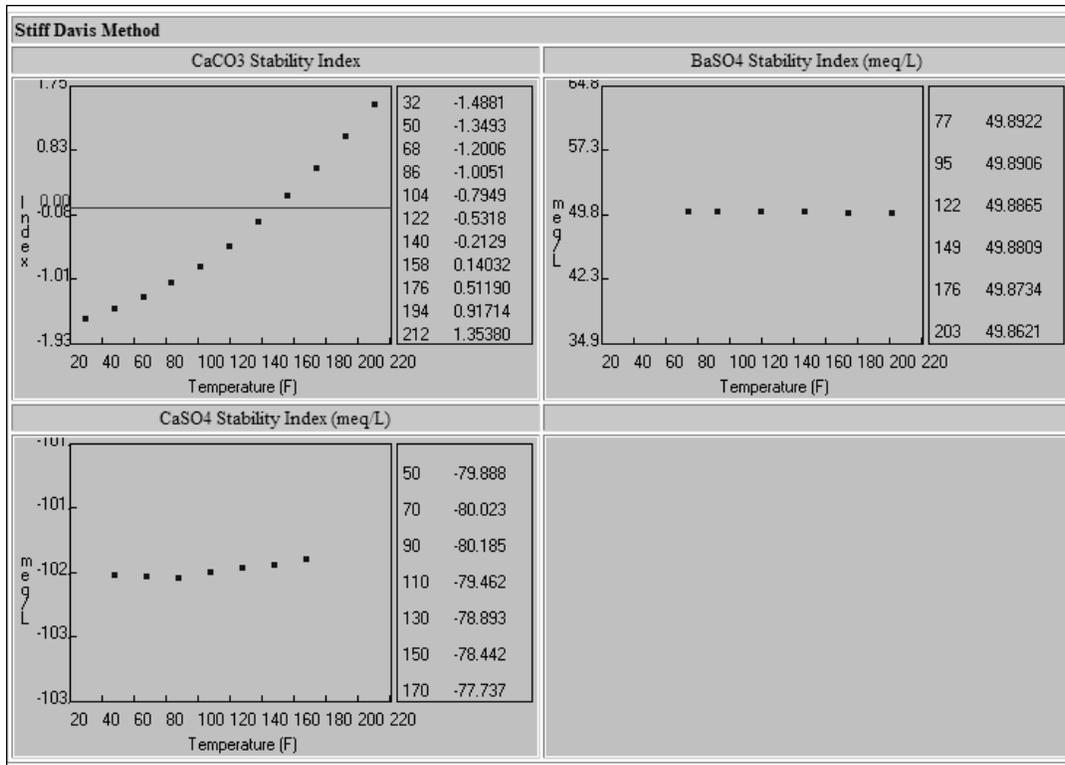


Fig. 19. Stiff Davis scale calculation. Scale precipitation is likely to occur at temperatures in excess of about 155°F.

Ions mg/L											
Ca ⁺⁺	2000	Mg ⁺⁺		Na ⁺	40000	Ba ⁺⁺		Fe ⁺⁺		Sr ⁺⁺	
CO ₃ ⁼	0	HCO ₃ ⁻	40	SO ₄ ⁼	403	Cl ⁻	89810	OH ⁻	0		
Probable Mineral Composition Meq/L											
BaCO ₃	0	CaCO ₃	0	MgCO ₃	0	Na ₂ CO ₃	0	Fe(OH) ₂	0		
Ba(HCO ₃) ₂	0	Ca(HCO ₃) ₂	0.66	Mg(HCO ₃) ₂	0	NaHCO ₃	0	FeCO ₃	0		
BaSO ₄	0	CaSO ₄	8.39	MgSO ₄	0	Na ₂ SO ₄	0	Mg(OH) ₂	0		
BaCl ₂	0	CaCl ₂	90.75	MgCl ₂	0	NaCl	1739.9	Ca(OH) ₂	0		

Figure 20. Probable composition of any scale that would form from a brine of this chemistry.

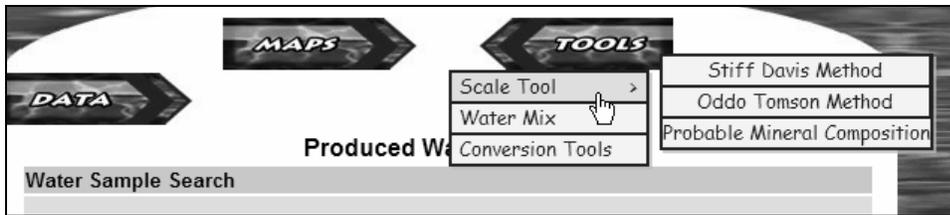


Fig. 21. Tools available include two scale calculation tools, a mineral composition tool, a tool that will calculate the composition of a mix of two waters, and a variety of unit conversion calculators.

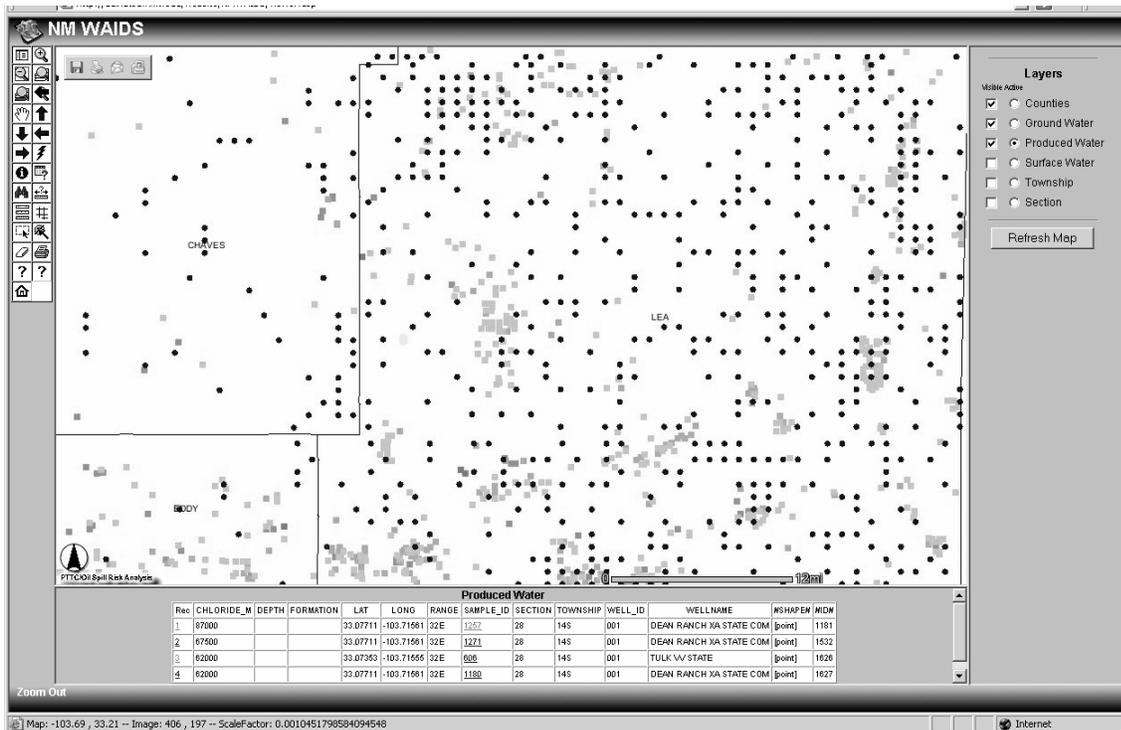
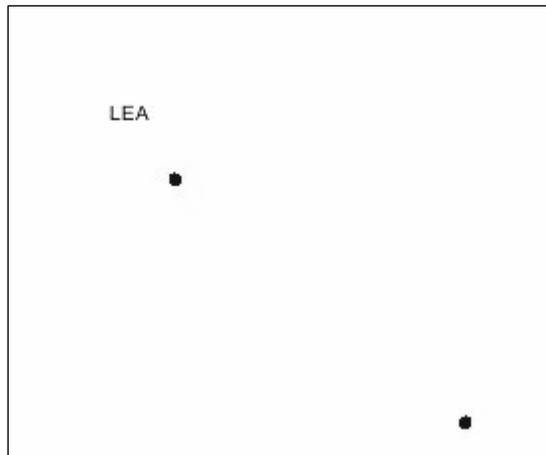


Fig. 22. Screen shot of NM WAIDS online map, zoomed in to Lea County. Produced and ground water are shown in different color ranges, with darker colors showing as higher chloride contents. This is particularly difficult to accurately demonstrate in the grayscale image.



Groundwater														
Rec	CHLORIDE_M	DATE_CLTD	DEPTH	FORMATION	LAT	LONG	PT_CLLTN	RANGE	SAMPLE_ID	SECTION	TOWNSHIP	USE	#SHAPE#	#ID#
1	112	Thu, 15 Aug 1957 00:00:00	147	OGALLALA	32.7038	-103.1358	DP	38E	7524	34	18S	MFG	[point]	22434
2	59	Mon, 13 Jul 1942 00:00:00	0	OGALLALA	32.7038	-103.1358		38E	8143	34	18S	MUN	[point]	22760
3	81	Mon, 13 Jul 1942 00:00:00	0	OGALLALA	32.7038	-103.1358		38E	7827	34	18S	MUN	[point]	22877
4	54	Wed, 9 May 1951 00:00:00	0	OGALLALA	32.7038	-103.1358		38E	7734	34	18S	MUN	[point]	22998

Fig. 23. Closeup of a selected well in the Groundwater Database. Using the Identify tool on the map server will bring back a screen similar to that seen below, where there is some information about the selected sample(s). Clicking on one of the hyperlinks in the Sample_ID column will provide more information, seen in Fig. 22.

General Information About: Sample 7734			
Section/ Township/Range	34 / 18 S / 38 E	Lat/Long	32.7038 / -103.1358
Elevation	0	Depth	0
Date Collected	5/9/1951	Chlorides	54
Collector / Point of Collection	USG /	Use	
Formation	OGALLALA		

Produced Water													
Rec	CHLORIDE_M	DEPTH	LAT	LONG	RANGE	SAMPLEID	FORMATION	SECTION	TOWNSHIP	WELL_ID	WELLNAME	#SHAPE#	#ID#
1	13210		32.70155	-103.13467	38E	4349	GBG/SA	34	18S	331	NORTH HOBBS G/SA UNIT	[point]	3190

Fig. 24. Information about selected sample in groundwater database. Using either the Identify or the Hotlink tool will link to this same report screen.

General Information About: Sample 4238			
SOUTH HOBBS G/SA UNIT 004			
API	3002507570	Sample Number	
Unit/Section/ Township/Range	K / 34 / 18 S / 38 E	Field	HOBBS
County	Lea	Formation	GBG/SA
State	NM	Depth	
Lat/Long	32.70164 / -103.13887	Sample Source	SEPARATOR
		Water Type	
Sample Date (MM/DD/YYYY)		Analysis Date (MM/DD/YYYY)	
Remarks/Description			
Cation Information (mg/L)		Anion Information (mg/L)	
Potassium (K)		Sulfate (SO)	1747
Sodium (Na)		Chloride (Cl)	6090
Calcium (Ca)		Carbonate (CO ₃)	
Magnesium (Mg)		Bicarbonate (HCO ₃)	105

Fig. 25. Sample record and detailed information for a well selected from the Produced Water layer of the online map.

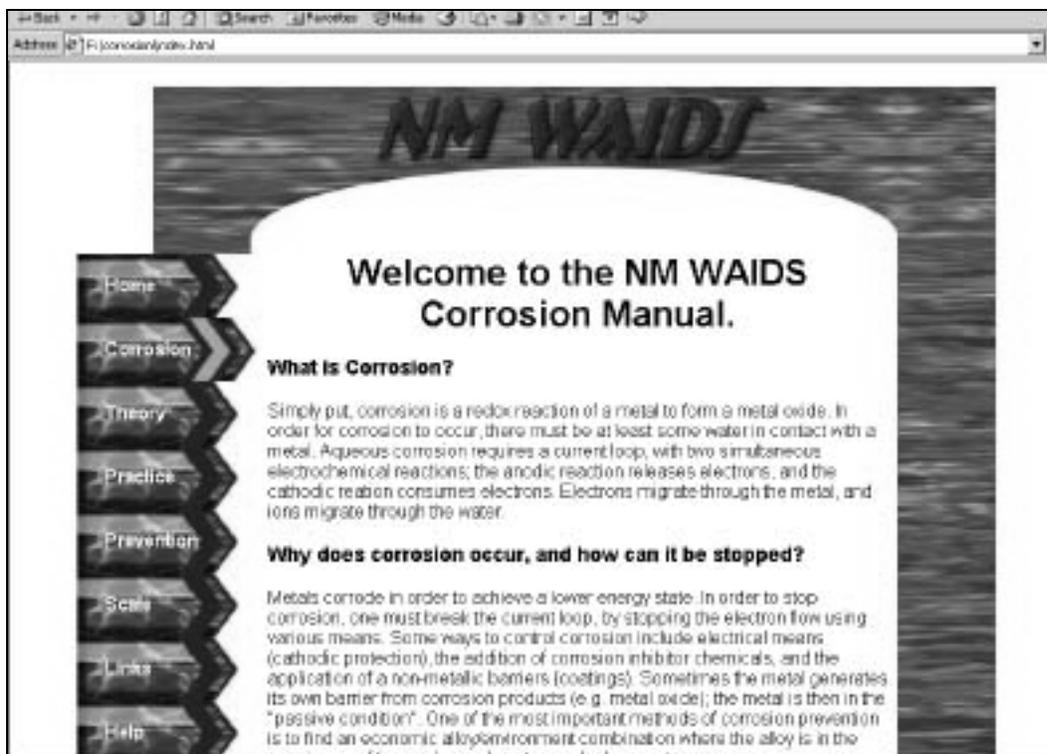


Fig. 26. Screen shot of home page for the corrosion manual.

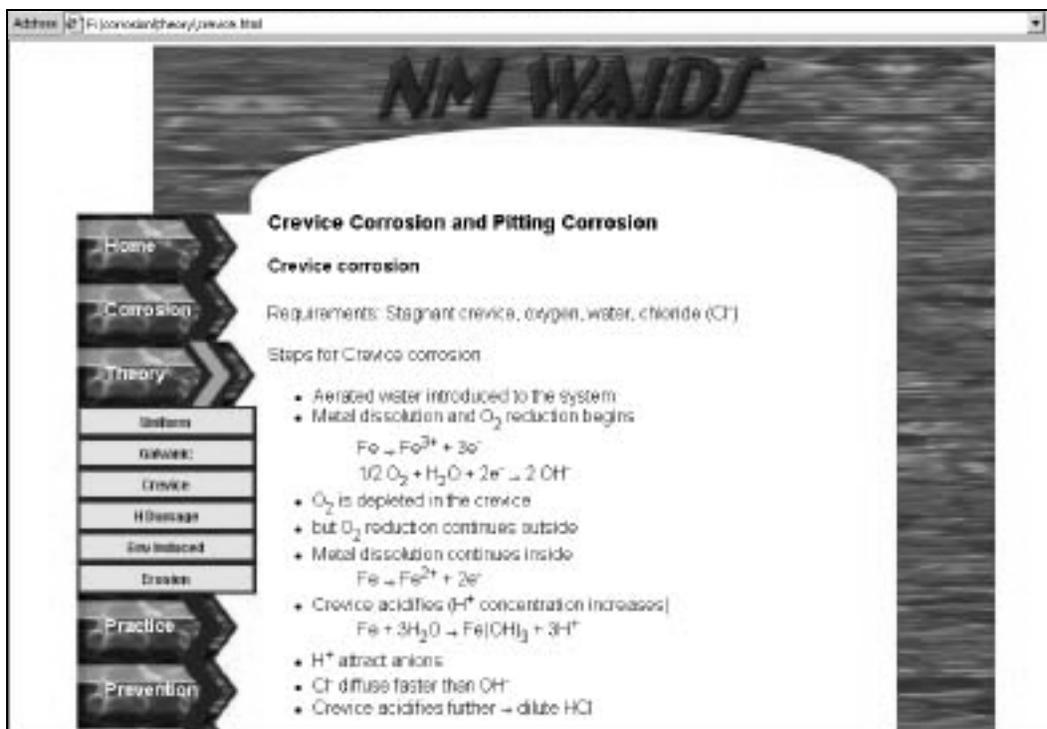


Fig. 27. Theory page has more information about a variety of commonly seen oilfield corrosion mechanisms.



Fig. 28. Screen shot of the CO₂ Corrosion page, providing more information, likely environments of formation, pictures, and mitigation options.



Fig. 29. Corrosion prevention section.

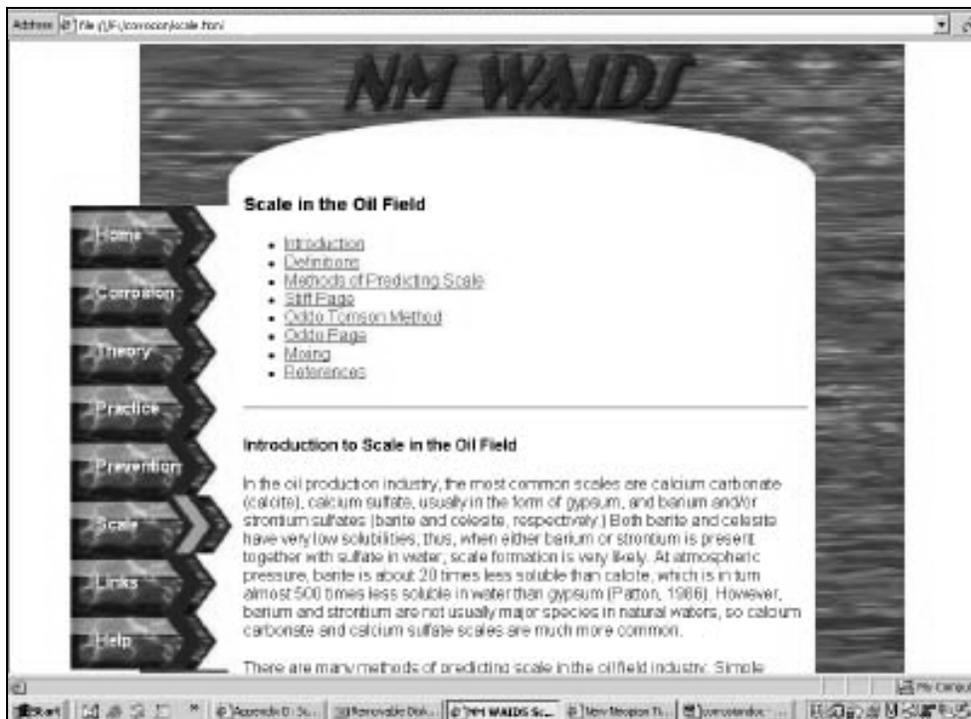


Fig. 30. Screen shot of Scale page.

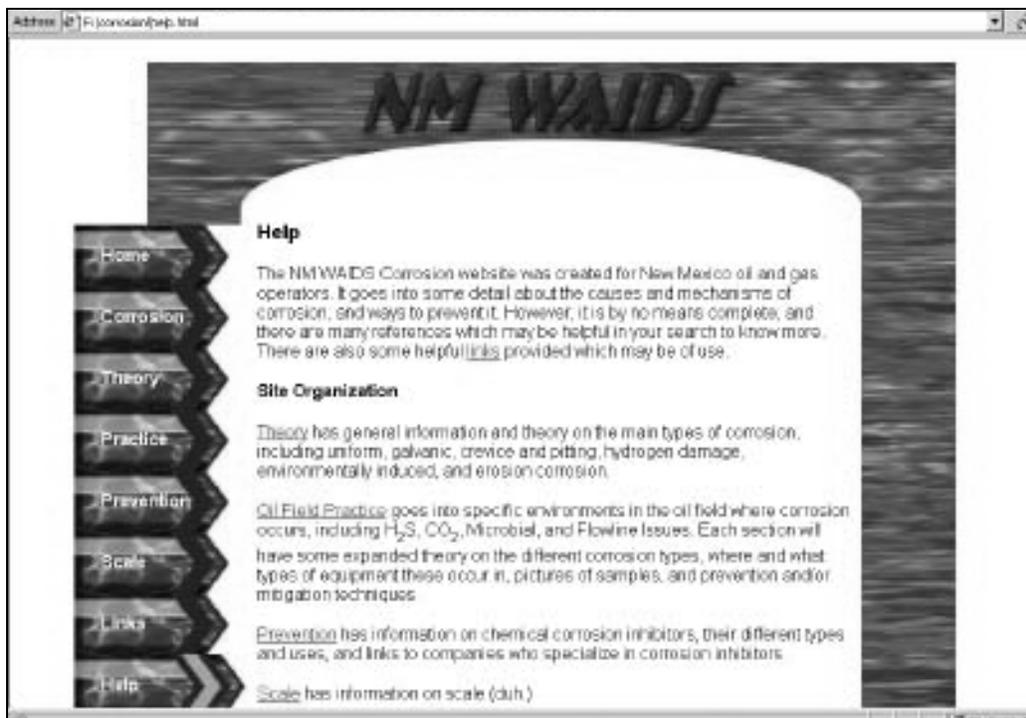


Fig. 31. Help section.