Using the root zone water quality model (RZWQM) to predict water movement through a hydrocarbon contaminated soil
by Scott Michael Anderson

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering
Montana State University

Abstract:
The quantification of hydrocarbon seepage into the groundwater is a regulatory concern for a refinery. Contaminant transport in the vadose (unsaturated soil) zone is highly dependent on the flow of soil-water. Vadose zone simulation models, such as the popular Root Zone Water Quality Model (RZWQM) can be used to predict seepage of infiltrated water to the groundwater. Basic soil properties, hydraulic parameters, hourly precipitation, and refinery groundwater data were used in the RZWQM. Soil properties were determined from soil samples obtained from five locations in a disturbed area. A representative composite soil profile, using data from the five sample locations, was used for most analyses. The Rosetta program was used to estimate the Brooks and Corey water retention and hydraulic conductivity parameters required by the RZWQM. The RZWQM predicted a nonzero quantity of seepage to groundwater for every year of simulation. Predicted mean annual seepage for the period of groundwater record (1995 thru 2002) was 11.9 cm, which increased to 21.2 cm for the comparatively wetter period of precipitation record (1949 thru 2002). Exceedence probabilities were calculated for each year based on the precipitation record. There was an 80 percent probability that 15 cm of annual seepage would be exceeded and a 20 percent probability that 26 cm of annual seepage would be exceeded. The soil profiles that had the most influence on seepage were bulk density and percent silt in the top soil layer. Monte Carlo simulations were used to determine predicted seepage confidence based on randomly generated, normally-distributed soil properties. The standard deviation of predicted seepage was relatively high (approximately 47 percent of the mean); therefore, variations in soil properties produced a large range of seepage estimates. RZWQM storage predictions were significantly over predicted, which effected annual seepage totals. Inaccuracies probably had less effect on large daily seepage events, which may be more appropriate for estimating contaminant seepage. The link between water seepage and hydrocarbon transport still needs to be evaluated.

USING THE ROOT ZONE WATER QUALITY MODEL (RZWQM)

TO PREDICT WATER MOVEMENT THROUGH A

HYDROCARBON CONTAMINATED SOIL

by

Scott Michael Anderson

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Civil Engineering

MONTANA STATE UNIVERSITY-BOZEMAN
Bozeman, Montana

August 2003

APPROVAL

of a thesis submitted by

Scott Michael Anderson

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Otto R. Stein, Ph. D. — Otto R Stein (Signature) — 8/8/03 Date

Approved for the Department of Civil Engineering

Brett W. Gunnink, Ph. D. — (Signature) — 8/11/03 Date

Approved for the College of Graduate Studies

Bruce R. McLeod, Ph. D. — Bruce R. McLeod (Signature) — 8-20-03 Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature Scott Anderson

Date 8/8/03

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

The quantification of hydrocarbon seepage into the groundwater is a regulatory concern for a refinery. Contaminant transport in the vadose (unsaturated soil) zone is highly dependent on the flow of soil-water. Vadose zone simulation models, such as the popular Root Zone Water Quality Model (RZWQM) can be used to predict seepage of infiltrated water to the groundwater. Basic soil properties, hydraulic parameters, hourly precipitation, and refinery groundwater data were used in the RZWQM. Soil properties were determined from soil samples obtained from five locations in a disturbed area. A representative composite soil profile, using data from the five sample locations, was used for most analyses. The Rosetta program was used to estimate the Brooks and Corey water retention and hydraulic conductivity parameters required by the RZWQM. The RZWQM predicted a nonzero quantity of seepage to groundwater for every year of simulation. Predicted mean annual seepage for the period of groundwater record (1995 thru 2002) was 11.9 cm, which increased to 21.2 cm for the comparatively wetter period of precipitation record (1949 thru 2002). Exceedence probabilities were calculated for each year based on the precipitation record. There was an 80 percent probability that 15 cm of annual seepage would be exceeded and a 20 percent probability that 26 cm of annual seepage would be exceeded. The soil profiles that had the most influence on seepage were bulk density and percent silt in the top soil layer. Monte Carlo simulations were used to determine predicted seepage confidence based on randomly generated, normally-distributed soil properties. The standard deviation of predicted seepage was relatively high (approximately 47 percent of the mean); therefore, variations in soil properties produced a large range of seepage estimates. RZWQM storage predictions were significantly over predicted, which effected annual seepage totals. Inaccuracies probably had less effect on large daily seepage events, which may be more appropriate for estimating contaminant seepage. The link between water seepage and hydrocarbon transport still needs to be evaluated.

INTRODUCTION

Due to unknown potential impacts to off-site locations, ConocoPhillips is concerned with the transport of petroleum hydrocarbons that exist in the vadose (unsaturated soil) zone within their Billings, Montana refinery. Contaminant transport is highly dependent on the flow of soil-water, thus a potential for a large precipitation event to transport contaminants to the relatively shallow groundwater and away from the refinery exists. A preliminary study by Kevin Germain (personal communication, 2001) used the Soil Conservation Service (SCS) curve number method in conjunction with the Thornthwaite water balance model to predict hydrologic processes at the refinery. These simplified lump-sum models are difficult to calibrate for site-specific conditions. Using average annual precipitation and crudely estimated coefficients near zero annual seepage was estimated, suggesting that hydrocarbon transport by infiltrated water was of little concern.

To better understand these hydrologic processes and enable ConocoPhillips to make more informed management decisions a robust hydrologic model, the Root Zone Water Quality Model (RZWQM), was used to model runoff, infiltration, evapotranspiration (ET), soil storage, and seepage to groundwater within the disturbed South 40 area of the Billings, Montana refinery. The RZWQM is an event-driven numerical model that utilizes site-specific soil properties and meteorological input data.

Of the hydraulic processes simulated with the RZWQM, seepage is of greatest concern here. The quantity of seepage is the result of many factors including the characteristics of the soil profile, the volume of water initially in the soil, the depth to groundwater, and the intensity, duration, and frequency of precipitation events. Accurate estimates of soil-water transport are necessary before contaminant fate and transport can be predicted.

Mean annual and daily seepage quantities were determined from measured soil properties and historical precipitation data. Variations in mean annual seepage due to uncertainty in measured soil properties were also estimated. The confidence of mean annual seepage predictions were determined with the Monte Carlo method by randomly sampling the soil input parameters based on statistical distributions determined from the measured data and then determining the output seepage distribution.

Goals and Objectives

The primary goal of this research was to provide ConocoPhillips with a best estimate of the quantity of groundwater seepage (if any) using the RZWQM. The model incorporated measured site-specific soil properties obtained from five locations (sites A, B, C, D, and E) in the South 40 area and provided estimates for a larger composite area (composite site). The most permeable soil layers were identified, the sensitivity of seepage to the specific soil properties was determined, and the model confidence was estimated. This goal was accomplished by completing the following objectives:

1. On-site data collection. Necessary field data was collected from the South 40 area of the Billings, Montana ConocoPhillips Refinery site. Infiltration rates were measured using infiltrometers. Intact soil cores were used to measure texture (percent sand, silt, and clay), bulk density, saturated hydraulic conductivity, and other important soil properties. These and other related data provided the basis for developing a comprehensive assessment of seepage in the South 40 area.

2. Seepage estimation. The RZWQM was used to predict seepage to groundwater at the sample locations and the composite site based on mean soil properties and hourly precipitation data for the period of groundwater record (8/16/1994 thru 12/31/2002). Predicted seepage values are a necessary first step in estimating contaminant seepage.

3. Seepage probability. Long-term management decisions should not be entirely based on the relatively short period of groundwater record. Seepage was predicted for the composite site with the RZWQM using the full Billings hourly precipitation record (7/3/1948 thru 12/31/2002) and a constant groundwater elevation. Exceedence probabilities are calculated for annual and daily seepage events, which can be used in risk assessment.

4. Model sensitivity. A sensitivity analysis was performed to determine which soil properties had the greatest effect on RZWQM predicted seepage. The composite site's soil properties were varied plus or minus one standard deviation for the period of groundwater record. The variability of soil properties within each of the soil layers and between sites was used to determine seepage variation.

5. Model confidence. Confidence in predicted seepage, based on measured soil properties, was assessed using the Monte Carlo method. Simulations were run for the period of groundwater record using the composite site's physical and hydraulic inputs in the RZWQM. The probability that a specific quantity of seepage will be exceeded was determined based on soil property variability.

Monte Carlo Method

Rubinstein (1981) described the Monte Carlo method as the most commonly used and powerful complex problem analysis tool. Monte Carlo is not a computer model in itself, but can be used in conjunction with any numerical model. The Monte Carlo method uses statistical distributions of input parameters to estimate the range of possible outputs. The probability of an event can be determined from the statistical distributions of model outputs. Common applications for the Monte Carlo method are maintenance operations, wind tunnel analysis, river basin modeling, and military training (Rubinstein, 1981).

If there are a large number of model inputs, a sensitivity analysis may be performed to minimize inputs used in the Monte Carlo simulations (Rubinsein, 1981). Monte Carlo simulations are often executed only on the input parameters that have the greatest effect on model output. The statistical distributions for the sample population of each input parameter simulated must be known. Common statistical distributions are: uniform, normal, lognormal, exponential, gamma, etc. The chi-square goodness-of-fit test, described by Stephens (1998), can be used to determine which statistical distribution is most appropriate for each input parameter.

Using the Monte Carlo method, independent random (technically pseudorandom) numbers are generated from the statistical distributions of each input parameter (Rubinstein, 1981). Best results are obtained by generating as many sets of random numbers as practically possible. The generated values can be input into a numerical model, such as the RZWQM.

Root Zone Water Quality Model (RZWQM)

In 1992, the United States Department of Agriculture (USDA) developed the RZWQM to simulate soil-water, nutrient, and pesticide transport in the soil vadose zone (Ahuja et al., 1999a; RZWQM, 1992). The RZWQM is capable of simulating water quality and agricultural processes such as infiltration, runoff, evapotranspiration (ET), water distribution, chemical transport, macropore flow, heat transport, plant growth, organic matter cycling, and pesticide processes (Malone et al., 2002).

The RZWQM is a mass-conservative, mixed form iterative finite-difference numerical model (Ahuja et al., 1999a). Lateral heterogeneity is not considered because the RZWQM is a point-scale model (Malone et al., 2002). Vertical heterogeneity can be accurately modeled with up to twelve distinct soil layers to the maximum depth of three meters. The effects of varying input parameters can be easily analyzed by running multiple scenarios using the built-in batch mode. A maximum of 100 years of precipitation data can be simulated with the RZWQM (Ahuja et al., 1999a).

Minimum RZQWM inputs are precipitation and meteorological data, soil horizon depths, soil texture (percentage of sand, silt, and clay), bulk density, and bottom boundary condition. A unit gradient, constant flux, or constant pressure head boundary must be set at the bottom of the profile (Ahuja et al., 1999b). A unit gradient condition is used to simulate unsaturated soil below the bottom soil layer. The constant flux condition is used with a dynamic groundwater table. The third condition models a constant pressure head at the bottom boundary using the initial matric potential.

The RZWQM calculates a water balance between precipitation, runoff, infiltration, evapotranspiration (ET), seepage, and storage using sub-hourly time intervals (Ahuja et al., 1999a). Internal to the model are three numerical grids: the first grid is divided by the input layer depths used to describe the physical properties of the soil, the second is a non-uniform grid for water redistribution, and the third is a one centimeter grid for vertical infiltration (Ahuja et al., 1999b). The bottom soil profile depth occasionally needs to be adjusted by a few centimeters to accommodate the model's numerical schemes.

The Root Zone Water Quality Model uses a Modified Penman-Montieth model to simulate potential evapotranspiration (Malone et al., 2002). ET is the combined amount of evaporation and transpiration. Evaporation is the physical process of soil or surface water entering the atmosphere, and transpiration is an evaporative process facilitated by plant biological activity (Farahani and DeCoursey, 1999). The freezing and thawing of soils are not simulated with the current version of the RZWQM. Instead, the model uses the Precipitation Runoff Modeling System (PRMS) to determine snow accumulation and melt. The PRMS model simulates the water balance and heat transport to and from the snow. When snow has accumulated on the soil surface, the surface temperature is set to zero degrees Celsius. Otherwise, the soil surface temperature is set equal to the average air temperature for the day (Flerchinger et al., 1999).

Seepage is often the most difficult hydraulic process to measure or estimate. Precipitation, infiltration, runoff, ET, and soil storage measurements must be accurate for seepage to be accurately predicted. Singh et al. (1996) determined that effective porosity, initial water content, and lateral saturated hydraulic conductivity were the most important model parameters for quantifying seepage. Unfortunately, the lateral saturated hydraulic conductivity is difficult to measure because artificial drains are usually required.

Runoff and water content predictions using the RZWQM are usually reasonable though consistent biases have been documented. Ghidey et al. (1999) used the RZWQM to predict runoff and soil-water content at an agricultural site. Measured soil texture, bulk density, and 1/3-bar water content were used. The model tended to over predict runoff during dry periods and under predict during wet periods. However, the overall predicted runoff varied by less than five percent of the measured runoff. Soil-water contents were slightly under predicted in most cases and seriously under predicted in high clay content soils. In a similar study, Malone et al. (2000) found that soil-water content predictions were generally adequate but sometimes under predicted.

A benchmark for estimating and calibrating ET is the Bowen Ratio Energy Balance (BREB). Jaynes and Miller (1999) found that ET predictions were somewhat under predicted using the RZWQM as compared to the BREB. However, the authors found that simulated annual seepage compared reasonably well to drainage from a closed subbasin. Farahani and DeCoursey (1999) found simulated seepage to be accurate when judged against measured seepage collected in metal troughs.

Ellerbroek et al. (1998) used the RZWQM and the Monte Carlo method to determine if metolachlor transport was more sensitive to irrigation rate or saturated hydraulic conductivity. The authors found that varying the water application rate had a greater affect on transport than the saturated hydraulic conductivity. In fact, conductivity was significant only at high application rates. Surprisingly, Malone et al. (2000) found that increasing the saturated hydraulic conductivity in the RZWQM decreased water and herbicide transport through relatively large cracks in the soil called macropores.

The model employs the popular Green-Ampt equation to simulate the downward flow of water through the soil profile during precipitation (Ahuja et al., 1999b). When water first enters the soil the infiltration rate is high, but gradually decreases until a constant final rate is attained (Or et al., 2002). The Green-Ampt equation is (Green and Ampt, 1911; Ahuja et al., 1999b):

$$V = K_{se} \cdot \frac{(\Psi_{wf} + H_o + Z_{wf})}{Z_{wf}}$$

where V is the infiltration rate (cm/hr), K_{se} is the effective average saturated hydraulic conductivity of the wetting zone (cm/hr), Ψ_{wf} is the matric potential (negative pressure head) at the wetting front (cm), H_o is the depth of surface ponding (cm), and Z_{wf} is the depth of the wetting front (cm). Basic assumptions of the equation are: a distinct wetting front exists in the soil where the water content is constant below and above it, the wetted soil has constant properties, and the matric potential is constant at the wetting front (Or et al., 2002).

The Richard's equation describes how soil-water is distributed in the vadose zone between precipitation events. The one-dimensional Richard's equation is built into the RZWQM (Ahuja et al., 1999b; Wu et al., 1996):

$$\frac{d\theta(\Psi)}{dt} = \frac{d}{dz}\left[K(\Psi)\left(\frac{d\Psi}{dz} - 1\right)\right]$$

where Ψ is the matric potential (cm), z is the height from a reference elevation (cm), $\theta(\Psi)$ is the soil-water content as a function of matric potential (cm^3/cm^3), and $K(\Psi)$ is the hydraulic conductivity as a function of matric potential (cm/hr). Water retention ($\theta(\Psi)$) and hydraulic conductivity ($K(\Psi)$) functions are required to determine the distribution of soil-water at any given time. Typical water retention and hydraulic conductivity curves are shown in Figures 1.1 and 1.2, respectively.

Figure 1.1. Water retention curves for three soil types (Or et al., 2002).

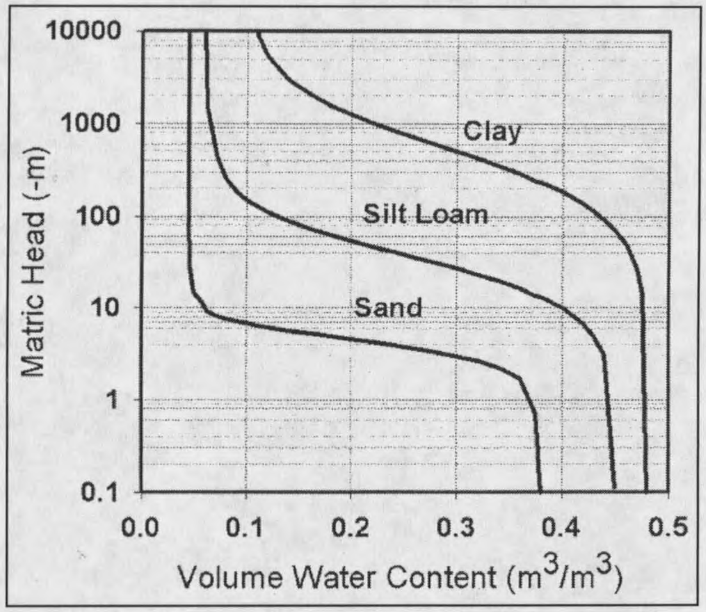
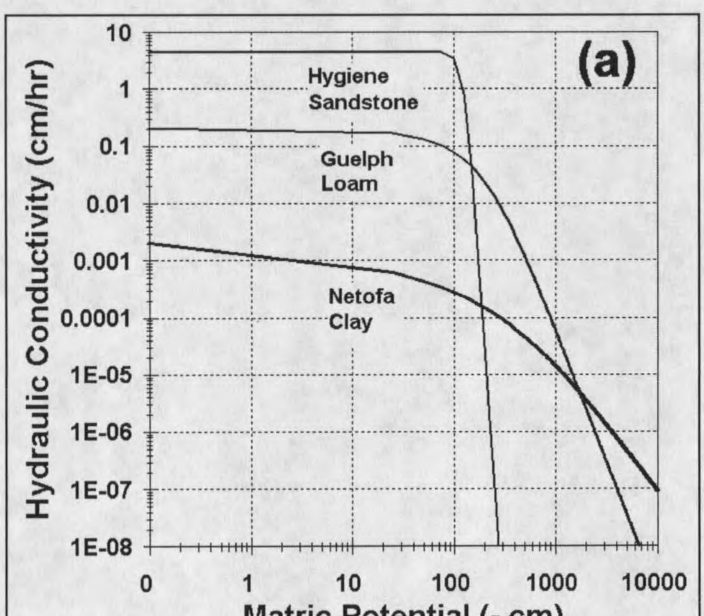


Figure 1.2. Hydraulic conductivity curves for three soil types (Or et al., 2002).



Water Retention and Hydraulic Conductivity Functions

Most numerical vadose zone models use either the Brooks and Corey (1964) or van Genuchten (1980) equations to approximate the water retention and hydraulic conductivity relationships necessary to describe the Richard's equation. The RZWQM uses the Brooks and Corey equations. An understanding of both sets of equations is useful because neither is universally accepted and there is often a need to convert between them. Figure 1.3 shows measured water retention data and calibrated Brooks and Corey and van Genuchten functions.

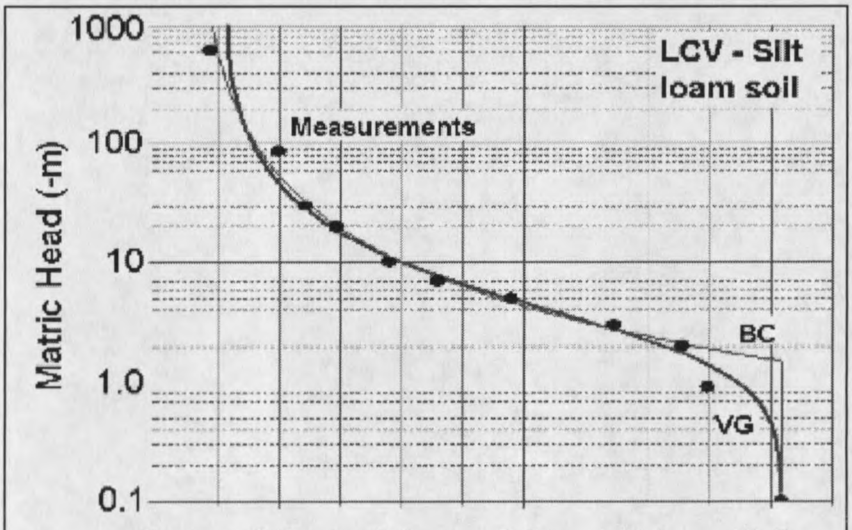
The Brooks and Corey equations are still widely used today, almost forty years after their development. Despite the popularity of these equations, a discontinuity exists at the bubbling pressure that makes them difficult to implement. Two separate matric potential ranges are required for the Brooks and Corey functions because of this discontinuity (Brooks and Corey, 1964; 1966):

For $\Psi > \Psi b$: $$\frac{\theta(\Psi) - \theta_r}{\theta_s - \theta_r} = \left(\frac{\Psi_b}{\Psi}\right)^{\lambda} \qquad K(\Psi) = K_{sat}$$

For $\Psi < \Psi b$: $$\frac{\theta(\Psi) - \theta_r}{\theta_s - \theta_r} = 1.0 \qquad K(\Psi) = K_{sat}\left(\frac{\Psi_b}{\Psi}\right)^{2+3\cdot\lambda}$$

where θ_s is the saturated water content (cm^3/cm^3), θ_r is the residual water content (cm^3/cm^3), Ψ_b is the air-entry or bubbling pressure (cm), λ is the pore-size distribution index for the soil, and K_{sat} is the saturated hydraulic conductivity (cm/hr). θ_s is often assumed equal to the soil porosity. θ_r is the amount of water in a volume of soil after the soil has thoroughly drained. Unfortunately, θ_r is difficult to measure accurately. van Genuchten (1980) suggested that a large matric potential (–15 bars) could be used as an approximation. Ψ_b is related to the largest pore diameter in a continuous network of flow pores and is approximated by the capillary pressure at which air first moves through the pores. λ is related to the size of the flow channels in the soil. Sands generally have a large λ, while clays usually have smaller values (Brooks and Corey, 1964; 1966). K_{sat} is a measure of how fast water is transported when the soil is completely saturated. It is important to note that K_{sat} is not equivalent to K_{se} in the Green-Ampt equation (Klute and Dirksen, 1986). Bouwer (1986) found that K_{se} was equal to approximately half K_{sat}.

Figure 1.3. Brooks and Corey and van Genuchten functions calibrated to measured water retention data (Or et al., 2002).



van Genuchten (1980) proposed a set of water retention and hydraulic conductivity equations that do not exhibit a discontinuity. The van Genuchten functions are:

$$\frac{\theta(\psi) - \theta_r}{\theta_s - \theta_r} = \left[1 + (\alpha \cdot \Psi)^n\right]^{-m} \qquad K(\Psi) = K_{sat} \cdot \frac{\left[1 - (\alpha \cdot \Psi)^{n-1} \cdot \left[1 + (\alpha \cdot \Psi)^n\right]^{-m}\right]^2}{\left[1 + (\alpha \cdot \Psi)^n\right]^{\frac{m}{2}}}$$

where α (1/cm), n, and m are fitting parameters and all other variables are as previously defined. As a simplification, the parameter m is usually defined equal to 1-1/n.

van Genuchten (1980) applied the proposed water retention and hydraulic conductivity functions to four types of soils: sandstone, two types of silt loams, and clay. For the sandstone and the silt loams, the model showed excellent agreement between observed and predicted hydraulic conductivities. However, the hydraulic conductivity was seriously under predicted with the clay (van Genuchten, 1980). van Genuchten et al. (1989) concluded that the unsaturated hydraulic conductivity models work reasonably well for coarse and medium-textured soils but give worse predictions for fine-textured (clay) soils.

Vereecken et al. (1989) compared five different forms of the van Genuchten water retention function ($\theta(\Psi)$) to measured water retention data. The only parameter that was varied was m. Best results were found when m was calibrated or set equal to unity. The standard form of the equation (m = 1-1/n) provided less accurate predictions for the 182 soil horizons analyzed. The parameter m was set equal to unity for a sensitivity analysis of θ_s, θ_r, α, and n. The authors found that water retention was most sensitive to the saturated water content and least sensitive to the residual water content.

Vereecken et al. (1990) compared the accuracy of four different hydraulic conductivity equations ($K(\Psi)$) using data measured from 127 soil cores. The first model was an exponential function, the second was a power function, the third had the form of the Brooks and Corey equation with the residual water content set to zero, and the fourth model was the Gardner (1958) equation, which is very similar to the van Genuchten equation. The least amount of variance was found with the Gardner equation (0.16).

Regardless of which set of $\theta(\Psi)$ and $K(\Psi)$ functions are used, the fitting parameters (Ψ_b and λ or α and n (assuming m is fixed)) must be calibrated to a specific soil or estimated from known soil properties (van Genuchten, 1980). These parameters can be calibrated to measured water retention or hydraulic conductivity data. Because of the relative ease of measurement, the $\theta(\Psi)$ function is usually calibrated and the parameters used in the $K(\Psi)$ function.

Direct Measurements

The Brooks and Corey or van Genuchten functions can be calibrated from water retention data measured directly in the field or laboratory. Water retention data usually has the shape of the water retention curves shown in Figure 1.1 (Or et al., 2002). To obtain continuous functions, the Brooks and Corey parameters Ψ_b and λ or the van Genuchten parameters α and n (with m fixed) can be calibrated to the data by minimizing the sum of the squared errors (SSE) or maximizing the squared correlation coefficient

(R^2) using spreadsheet software (Wraith and Or, 1998). The calibrated parameters can be used in both the $\theta(\Psi)$ and $K(\Psi)$ functions.

Unfortunately, the cost and time involved in direct measurements often make these methods unfeasible. For example, water retention determination in the laboratory requires desorption of an initially saturated soil to a specific matric potential (soil suction head) while simultaneously measuring the water content. This procedure must be repeated for many matric potentials to obtain a smooth curve (Ayra and Paris, 1981). A popular field method for obtaining water retention curves consists of taking numerous measurements with a tensiometer for matric potential and time domain reflectometry (TDR) probe for water content (Or et al., 2002). Hydraulic conductivity measurements are even more arduous. It can take weeks to determine the conductivities of fine-textured or highly compact soils.

Estimation Techniques

Rather than calibrating the Brooks and Corey or van Genuchten parameters from measured data, the parameters can be directly estimated from easily measured soil properties. Most often these properties are soil texture, bulk density, organic carbon content, water contents measured at specific matric potentials (i.e. 1/3 or 1/10-bar), or any combination of these. The more soil properties included the better the results (Rawls et al., 1982; Wu et al, 1996). van Genuchten et al. (1989) stated some advantages of estimation techniques over direct measurement: the governing flow equations do not need to be approximated, hydraulic conductivities can be predicted for the full range of water contents, and parameter uncertainty and model accuracy can be quantified.

The Brooks and Corey (1964) parameters must be input into the RZWQM directly or estimated by one of two methods built into the RZWQM. The first estimation method is used when only the soil texture (percent sand, silt, and clay) and bulk density are known (Ahuja et al., 1999b). The model determines the USDA Soil Conservation Service (SCS) texture class from these properties. Tabulated Brooks and Corey parameters are then assigned based on the bulk density and texture class alone.

The second RZWQM estimation method predicts the Brooks and Corey parameters using the soil scaling technique described by Warrick et al. (1977). This method can only be used if the measured saturated hydraulic conductivity and 1/3 or 1/10-bar water content are known in addition to soil texture and bulk density. The soil is first classified as it was in the previous method. Mean properties from this soil class are used as the "similar soil." The authors found that the matric potential of a soil sample could be related to the similar soil's scaled mean matric potential. The water retention curve of the similar soil is then shifted to coincide with the measured water content of the sample soil. The Brooks and Corey parameters are found by minimizing the sum of squares. Soil scaling was found to be most effective in coalescing large data sets, where it reduced the sum of squares by 34 to 90 percent (Ahuja et al., 1999b).

Cosby et al.'s (1984) method can be used to estimate the Brooks and Corey parameters directly. The authors found that soil texture was the only regression equation property of significance from a two-way analysis of variance (ANOVA). In fact, the authors found that use of only one texture parameter gave an approximation nearly as good as using two texture components (three would be redundant). Linear regression equations were developed using only percent sand, silt, or clay. Correlation for all estimated parameters to 1448 measured water retentions and hydraulic conductivities was relatively high ($0.77 < R^2 < 0.97$).

Recently, neural networks have been used to improve the accuracy of empirical regression equations (Schaap, 1999). Neural networks get their name from their resemblance to biological neural systems. Neural networks use an iterative calibration procedure to link soil inputs to hydraulic outputs. Schaap et al. (1998) defined a common neural network as a nonlinear data transformation structure consisting of input and output nodes connected to a number of hidden nodes by adaptable coefficients. In this study, the authors compared four published water retention equations and six published hydraulic conductivity functions to the neural network predictions. A total of 1209 soil samples for water retention prediction and 620 for hydraulic conductivity were used.

Schaap et al. (1998) found the root mean square residuals (RMSR) of the neural network models to be approximately 0.01 to 0.02 m^3/m^3 lower than the published regression equations. These results, however, were comparable when water content data was unavailable or when uncertainty in the data was large. Of the published water retention equations, Vereecken et al.'s (1989) predictive equation had the least amount of error (average RMSR of 0.098). The comparable neural network model had an average RMSR of 0.087. Cosby et al.'s (1984) published hydraulic conductivity equation performed best (average RMSR = 0.746 log(cm/day)). The neural network model with similar inputs (soil texture) had an average RMSR of 0.602.

Rosetta

The Rosetta program was developed by Schaap (1999) to use neural networks to predict the van Genuchten function parameters. Rosetta uses soil texture, bulk density, and up to two measured water contents to estimate α, n, and K_{sat} (m=1-1/n). Feed-forward neural networks are used along with the bootstrap method to determine parameter uncertainties. Given calculated uncertainties are plus or minus one standard deviation. The theory behind the bootstrap method is "that multiple alternative realizations of the population can be simulated from the single data set that is available (Schaap, 1999)." The predicted van Genuchten parameters must be converted into Brooks and Corey parameters in order to be input into the RZWQM (described in Chapter 3).

METHODS AND MATERIALS

Field Tests

Fieldwork was performed during November 2002 at the Billings, MT ConocoPhillips refinery. A map of the refinery is provided in Appendix A. The disturbed South 40 area is located in the Southeast ¼ of the refinery. Groundwater flow is to the Northeast. Five sample locations were chosen within the "South 40 DNAPL" shaded area in the greater South 40 area. These sites were termed A, B, C, D, and E. At each of the sites, Shelby and acetate tubes were direct-pushed into the ground with a hollow stem auger. The acetate samples from site D were taken approximately three meters (ten feet) to the east of where the Shelby tube cores were obtained because deeper sampling was possible in that area due to less gravel at the surface. Otherwise, Shelby and acetate tube samples were within one meter of each other. The capability of the drill rig, the granularity and compaction of the soil, and the strength of the tubes limited the depth that samples could be obtained in some cases. The steel Shelby tubes were 7.6-cm (3-in) in diameter and 76.2-cm (2.5-ft) long. The acetate tubes were 3.5-cm (1.4-in) in diameter and 121.5-cm (4-ft) long. A South 40 location map with the five sample locations and nearby monitoring wells is included in Figure 2.1. The location D_acetate is where site D acetate tube samples were taken.