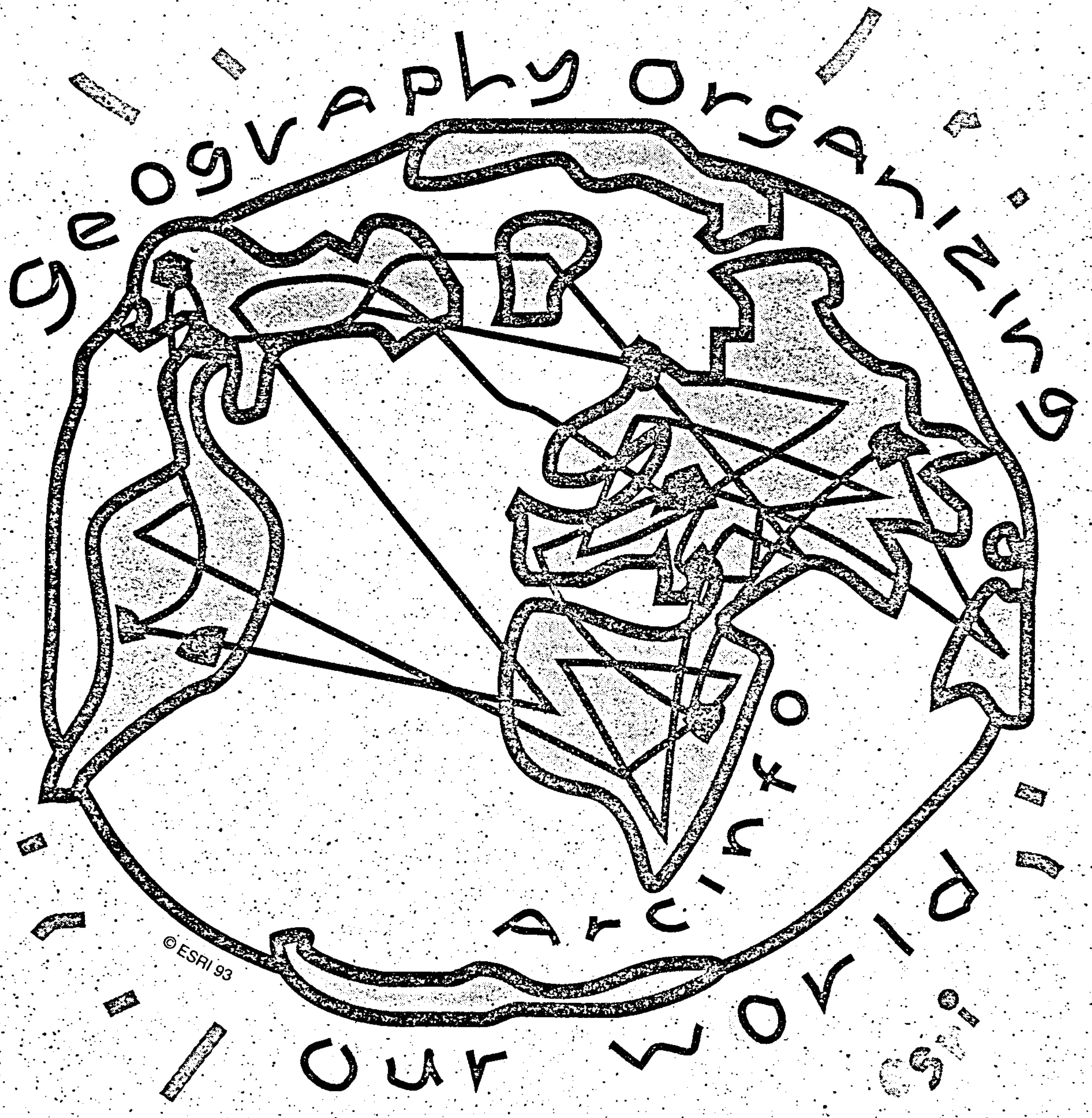
1276



# Proceedings of the Thirteenth Annual ESRI User Conference

## Volume 3

Peter J. Vaughan, Postdoctoral Research Associate
Dennis L. Corwin, Soil Scientist
Hong-Hua Wang, Graduate Research Assistant
US Salinity Laboratory, USDA/ARS
4500 Glenwood Dr.
Riverside, CA 92506
Telephone: (909) 369-4875
Fax: (909) 369-48 18

# COUPLING A CHEMICAL TRANSPORT MODEL TO A GIS DATABASE FOR ASSESSMENT OF NON-POINT SOURCE POLLUTION IN IRRIGATED AGRICULTURAL AREAS

Our purpose is the demonstration of a methodology for integrating a functional model of water and chemical transport with an ARC/INFO database. The database was created for an agricultural area of 2350 hectares in the San Joaquin Valley. Data gathered include: (1) GPS determination of the boundaries of quarter-sections; (2) chemical characteristics of soil samples obtained from 4 depths at 284 point locations; (3) soil survey maps ; (4) crop maps; (5) irrigation water delivery data (dates, amounts and chemical analyses); (6) daily evapotranspiration; and (7) crop characteristics such as maximum root depth. These parameters were organized into thematic INFO tables. The solute transport model, TETransgeo, calculates l-dimensional water and solute redistribution in vertical columns extending from the surface to the water table. An AML macro language script writes the input file for TETransgeo by drawing data from the various INFO tables, then runs TETransgeo and, finally, reads the output back into the relevant INFO tables where it becomes available for representation utilizing standard ARC/INFO techniques. Results generated by TETransgeo include the redistribution of water and solute within each soil column for a list of irrigation dates and the solute flux into the groundwater. Flux data can assess various irrigation management strategies with respect to non-point source pollution of the groundwater.

## INTRODUCTION

The development of large-scale irrigated agriculture has brought increased productivity to many areas of the world. Previously.unproductive areas such as the northernmost parts of Africa are now producing significant quantities of food. In other parts of the world existing agricultural areas have seen productivity rise substantially due to the development of irrigation along with other components of the green revolution. But this increased productivity has frequently been accompanied by environmental degradation which threatens the long term sustainability of agricultural production in addition to causing more general environmental problems such as deteriordtion of groundwater quality. This paper examines one **aspect** of irrigated agriculture that has been causing environmental degradation since ancient civilizations first developed irrigation techniques. Irrigation with saline water has resulted in increasing salinization of soils. A very significant additional problem is the deterioration of groundwater quality by increasing salinity.

It is common practice to use slightly saline water for irrigation because such water is less expensive than fresh water and, for certain soils, infiltration rates for fresh water are extremely low. Assessment of the effects of irrigation utilizing saline water requires development of a model of the vertical transport of both water and solute. Several such models of l-dimensional transport of water and solutes have been developed, some of them are reviewed by: Molz (198 1). Addiscott and Wagenet (1985) and van Genuchten and Jury (1987). Application of these models to agricultural sites has been limited due to the effort required to obtain data but one study coupled a l-dimensional finite difference code to a GIS to predict leaching of pesticides (Petach et al., 1991).

## Water flow in variably-saturated media

Above the water table soils are variably-saturated which means that both water and air are present in the pore space. The vertical movement of water through variably-saturated media is complicated because many factors influence water flow. Some of these factors are: the forces of gravity and capillary pressure; the geometry of the pore space; the wettability of the soil, and the presence of inhomogeneity in the soil caused by roots, animal burrows or variations in soil type. Any detailed mechanistic model that accounts for flow in the variably-saturated zone must take into account the relevant factors. Thus, a model of infiltration of irrigation water into the soil is quite complicated requiring depth profiles of physical properties such as the water retention curves (water content vs. soil water pressure) and the hydraulic conductivity of the soil. This kind of data is relatively easy to obtain in carefully-controlled laboratory experiments but evaluating these parameters for an agricultural field area presents a much greater challenge. For this reason, the current work utilizes a simplified functional **transport** model (Addiscott, 1977; Bond and Smiles, 1988) requiring a data set that is easier to obtain (Corwin, 1991). A mechanistic model would involve solution of the Richards equation for variably-saturated flow. The simplified functional model, called TETrans, has been described in detail elsewhere (Cot-win et al., 1991; Corwin and Waggoner, 1991).

## Water and solute transport model

The TETrans model calculates the redistribution of water and solutes in the soil profile after irrigation. The model places limits on the degree of saturation existing in a specified depth range. The lower limit is defined by the minimum water content which is the water content at the wilting point of plants. The upper limit is the field capacity. For water contents greater than field capacity, a soil will drain freely. TETrans assigns minimum water content and field capacity to each depth range (termed a compartment). Compartments are generally defined as depth ranges in which the physical properties are constant. These may be ascertained from soil survey data or by direct soil sampling. Each compartment is assigned a field capacity and a minimum water content as determined from easily measurable physical properties.

The TETrans algorithm is a mass balance calculation involving water and a single solute. When an irrigation or precipitation event occurs, the water content of the top compartment increases due to infiltration. If the field capacity is exceeded, then water drains into the next lower compartment. This process continues downwards through all compartments until all water from the irrigation/precipitation event is accounted for. Water draining from the bottom Compartment is assumed to enter the groundwater.

Water is removed from soil not only by drainage but also by evapotranspiration. The TETrans model handles evapotranspiration by defining evapotranspiration “events” which are summations of evapotranspiration over the time periods between irrigations. Evapotranspiration removes water from the soil according to a plant root water uptake distribution. The TETrans model requires that a fractional plant root water uptake distribution be defined such that each compartment will contribute some fractional amount of water to the total evapotranspiration. Several factors influence the total evapotranspiration, among them are plant type and maturity, temperature and wind speed. Evapotranspiration data is provided for agricultural areas in California by the California Irrigation Management Information System (CIMIS). The data consist of a daily baseline evapotranspiration (EtO) which is defied as evapotranspiration from a mature grass. The Et0 data are supplemented by crop coefficients which are used to correct the baseline number for a particular crop and crop maturity. For each irrigation or evapotranspiration “event”, TETrans recalculates water content and solute concentration after addition or removal of water from the soil profile.

Chemical Transport in TETrans, Over the past three decades a growing awareness of groundwater quality as an important environmental concern has led to the development of solute transport in the subsurface as an area of active scientific interest. Thus, many models have been developed for the purpose of calculating movement of solutes in the subsurface. TETrans can model the redistribution of a single solute (Corwin et al., 1991). An important concept that must be dealt with before solute transport in TETrans can be described is the concept of bypass flow. When fluids containing dissolved solids move through porous media there are several processes which cause the a variable velocity water flow. One of these is dispersion which is an actual variation in velocity (Sudicky, 1986; Deans, 1963). Consider, for example, water flowing along a circular tube. The water that is closer to the walls of the tube moves more slowly than water at the centerline. The same sort of effect occurs in each individual pore within a porous medium. The appearance of variable water velocity may also be due to the existence of immobile water locked in dead-end pores (Coats and Smith, 1964). Finally, there may be very rapid flow through large pores called macropores (Beven and Germann, 1982). TETrans lumps all of these effects together and assumes that the water fraction in a given compartment is divided into portions undergoing piston-type flow and bypass flow (Cot-win et al., 1991). Piston-type flow is flow at constant. velocity, which, for solute transport, means that an initial pulse of solute entering the soil at the top will not spread as it moves downwards. Bypass flow is that portion of water moving past a given compartment and on to the next one. The mobility coefficient, y, is the fractional amount of water involved in piston-type flow (Addiscott, 1977, 1981; van Ommen, 1985). The bypass flow fraction is 1 - y.

During movement of water through soil there may be several kinds of chemical reactions that take place. Solutes may react with each other in solution. Solutes may precipitate onto or dissolve from the surface of the solid. Adsorption/desorption of solutes on the solid phase may occur. Any or all of these processes may take place for a reactive solute. A non-reactive solute such as chloride does not participate in any of these types of reactions. The main effect of the bypass flow assumption on chemical effects is due to separation of the soil water into two chemically distinct fractions. The solutes contained in the fraction undergoing bypass flow will not react with either the solid or the fluid, these solutes are essentially nonreactive. For reactive solutes in the water fraction undergoing piston-type flow, TETrans currently treats adsorption/desorption reactions by assuming instantaneous equilibration between the solute and the adsorbed material.

Assumptions of the TETrans Model. TETrans requires certain assumptions regarding the mechanics of fluid flow in the unsaturated zone and the transport of solutes. These assumptions render the model useful from a management perspective because of diminished requirements in terms of computational resources and greatly simplified input data (Corwin et al., 1991). The most significant assumption involves dropping the time-dependence of the fluid redistribution. After an irrigation or precipitation event, the fluid is redistributed instantaneously throughout the profile. Another assumption is that water can only move downwards through the soil profile. The amount of water present in a compartment is determined by properties that are specified for that compartment, notably the field capacity and the minimum water content. The water contents of adjacent compartments are unrelated except when the field capacity is exceeded and water drains from one compartment to the compartment beneath it. Assumptions related to chemical transport include: (1) dispersion effects lumped together with macropore flow; (2) no dissolution/precipitation reactions; and (3) adsorption/desorption is instantaneous and non-hysteretic. The model also assumes that the soil profile can be characterized by a small number of homogeneous compartments.

## INCORPORATION OF TETrans INTO A GIS FRAMEWORK

The TETrans program was originally written as an interactive Fortran application that runs on both IBM-PC and hlacintosh platforms (Corwin and Waggoner, 1991). The user supplies a data set that represents conditions at a single geographic location. The data set consists of physical and chemical properties of the soil, irrigation amounts and dates, crop schedule and initial conditions for soil water content and solute concentration. Based on these data TETrans predicts soil water content and solute concentration at some future date. Other predicted quantities are the total amounts of water and solute that drain from the soil column into the groundwater. The original version of TETrans is useful as a tool for estimating the effects of certain irrigation and crop schedules for a single, homogeneous field. But, if spatial variability is significant or, if a model of a larger area is needed, then a geographically-based system would be more effective.

**The** objectives of the current work are twofold. First, a new TETrans code that can handle any number of individual locations should be written. Second, a geographically-based system for storing and accessing the data needed by the TETrans program must, be developed. These objectives were approached jointly by developing a database structure which could be implemented both in a C language program as a hierarchy of various structs and lists and in an INFO database as a set of related INFO tables (Corwin et al., 1993). The new C language program, called TETransgeo, meets the first objective by carrying out the TETrans calculations at successive locations in a selected set. The second objective was met by designing a system of related tables using the INFO module of the ARC/INFO1 GIS package.

TETransgeo is not an interactive program, it reads a text file containing all of the required input data and then writes a result file. A generalized algorithm giving the steps involved in a complete run of TETransgeo is shown in Table 1. Currently interactivity is restricted to querying the user for a few pieces of information such as dates over which the calculation should run and filenames for INFO tables. Future development will include work on a better user interface.

1. Gather data to generate a set of text files for the case.
   Example of a text tile required: Water content at several depths with fieldID and location ID numbers
2. Start the Main AML program which:
   1) prompts the user for some essential data such as dates of the simulation
   2) updates the relevant INFO tables using data contained in the text files
   3) creates a text file containing all data required by the TETransgeo program
   4) runs TETransgeo from a &sys command
   5) reads the result file created by TETransgeo into the INFO tables that store results
3. Analyze the computed results by creating maps and graphs as needed.

Table 1. Generalized algorithm for a run of TETransgeo

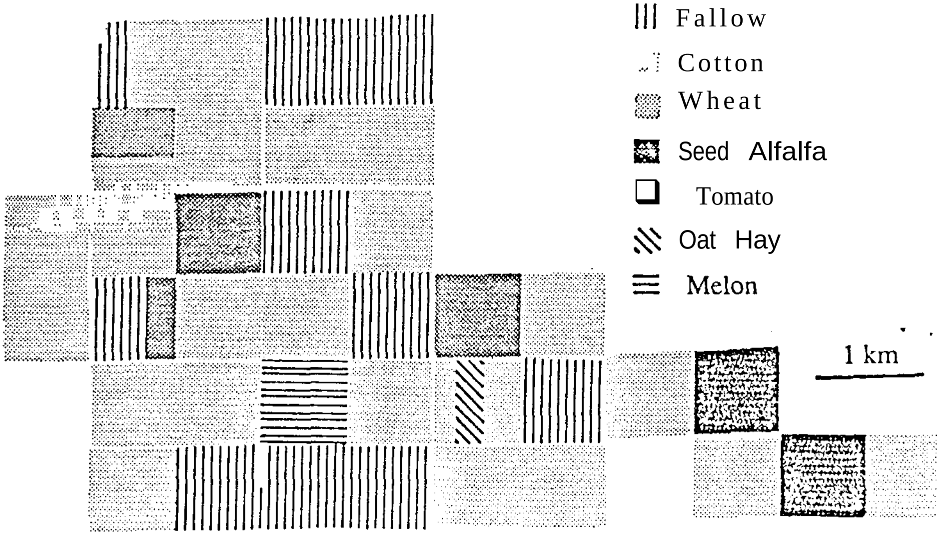


Figure 1. Crop map'of the Broadview Water District, summer 1991

## Data classification method

To provide a background for a complete description of the organization of the data and the AML macros that access the data, the following is a summary of some information relevant to the problem. The field area selected for the initial testing of TETransgeo consists of 2350 hectares in the Broadview Water District located in the San Joaquin Valley west of Fresno, California. This area was sampled at 284 sites during the spring of 1991 to determine water content and electrical conductivity of the soil water. Electrical conductivity can be used to estimate total dissolved solids by a simple empirical relation (tds = EC * 640) where tds is in mg/L and EC is in **dS/m.** The starting conditions for the TETransgeo calculation are therefore defined by these measurements. Other data can be represented in map form. For example, a map representing the crop type during the summer of 1991 is shown in Figure 1. The crop-type map follows the geographic boundaries of quarter-sections (64.8 hectares) except in those areas where a crop has been planted in an area smaller than the entire quarter-section. Other data (not shown) include the irrigation schedules which give the dates and amounts of irrigation water applied to entire quarter-sections. Periodic chemical analyses of the irrigation water also provide data on the solute concentrations of water entering the soil.

The original TETrans program has over 50 parameters (Corwin et al., 1991). Several parameters have been added in TETransgeo, notably parameters related to the newly-developed capability to calculate evapotranspiration internally. The original version of TETrans permits some parameters to vary with depth. If a depth-dependent parameter appears in an INFO table, then a separate item must be created corresponding to each compartment. Taking this into account, there are currently about 100 parameters that must be represented in various INFO tables. With this many parameters, the organization of data becomes an important factor in assuring a useful database. Some criteria utilized in database creation are: (1) maximum accessibility for the data; (2) minimum amount of data duplication and (3) logical hierarchy of parameters. These goals can be approached, in a general sense, by the establishment of a parameter classification scheme. For the current work a scheme including 4 categories was developed:

(1) Local parameters

Parameters of this type have values established for individual points. The locations of these points may be determined by statistical methods, they may lie on a grid or they may be determined by external factors such as sampling convenience.

(2) Spatial parameters

These parameters are constant within a restricted domain. The distinction between spatial data and local data depends on the sample spacing for the local data. A parameter may be known to be single-valued within a domain, but if the domain size is smaller than the point spacing for the local data representation of the parameter then representation of the parameter as local data is appropriate. Polygons represent spatial parameters in a GIS.

(3) Derived parameters

Derived parameters have no specific association with a given location and **they** generally occur in sets such as the set of parameters that describe crop properties. The selection of a particular set of values is based on the value of a discrete variable such as crop type.

(4) Regional parameters

Every environmental study must restrict itself to some geographic area. Regional parameters are those which can be considered constant over the entire area. Such parameters may be either simple constants or they may be time-dependent functions.

## Utilization of the Classification

As the first step in the construction of a database for the Broadview Water District field area, parameters that were identified as being significant to the problem of transport modeling in variably-saturated conditions were classified according to the parameter classification scheme outlined above. Of particular importance was the identification of parameters that are the derived type. Substantial savings of computer memory may be achieved by classifying as many parameters as possible into this general type. But the greatest benefit of such classification is the improvement in accessibility of the data Dividing the large data set into a set of smaller, thematic tables makes the data much easier to view and modify. An example of a set of derived parameters in the database discussed here are crop coefficients for evapotranspimtion (Figure 2).

```
typedef struct cropInfo_rec
(
        int               cropID;       /* Crop identification number */
        char              name[16];     /* Crop name */
        RootModelType     rootmdl;      /* Model of root distribution */
        double            rpmax;        /* Maximum root depth */
        double            rdc;          /* Root distribution coefficient */
        double            alpha;        /* Root water uptake distribution coefficient */
        double            kC1;          /* Evapotranspiration coefficients */
        double            kC2;
        double            kC3;
) cropInfo_rec;
```

Figure 2. C-language struct for specific crop data, cropID also appears in the location tables to identify the crop growing at a point.

The values of these coefficients are functions of crop type and can, therefore, be derived from crop type. An ID number which identifies the crop is stored in the table of location-specific data in order to idendfy the crop growing at that particular location. A separate, related table stores all the relevant data for each crop as well as the ID number.

Further classification of parameters involves the distinction between spatial and local parameters. Certain parameters, such as soil type, are classified as spatial because their values are constant within a restricted domain. The physical properties of soils needed for flow modeling are then derived parameters that may be obtained from the spatial parameter, soil type. But not all the soil properties required for the flow modeling can be obtained in this way. For other parameters, notably initial water content, measurements were made at specific locations within the field area. The initial water content is, therefore, a local parameter because it has a unique value for a single geographic location. Finally, there are regional properties such as baseline evapotranspiration which has a constant value over the entire study area for a particular day.

## construction of the Database

After classifying parameters, the database was constructed by grouping those parameters having a common theme. A diagram of the general data organization is given in Figure 3.

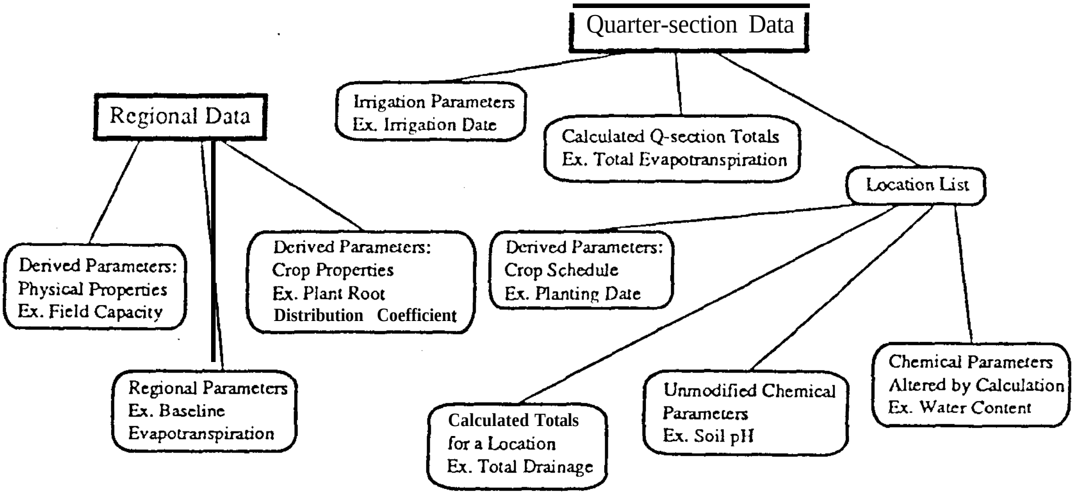


Figure 3. Data organization for both the external program and INFO database.

There are two entirely separated structures, one contains the regional parameters and derived parameters that are completely independent of location. The other structure contains all the data that are dependent, in some sense, on location within the field area. Data that are location-dependent include: soil attributes determined by soil type; irrigation data that are specific to each quarter-section; and various parameters that have been measured at point locations.

The database structure was established to provide an organization of all the data that arc of interest. Most of the data are stored in INFO tables that are exclusively reserved for user-supplied data. These tables do not store geographic information are not required by the geographic portion of the GIS. A small percentage of the data and several ID numbers are stored in tables that are associated with the geographic portion of the GIS. Examples of such tables are the point attribute tables used for representing sample locations on a map and polygon attribute tables that are used for storing spatial data such as soil maps. The ID numbers stored in these tables serve as pointers to INFO tables containing the actual data. As an example, consider a single quarter section (normally containing eight locations). In many cases a single crop grows within a single quarter-section so identical crop ID numbers are stored in the INFO record at all eight locations. If, however, the quarter-section is split into areas with different crops then the crop ID numbers will vary. The occurrence of split fields required storage of crop ID number in the point attribute tables rather than storage in the polygon attribute tables representing entire quarter-sections.

## AML Scripts and Running TETransgeo

The utility of a large database is greatly enhanced by automating the operations of storing and retrieving data. ,In ARC/INFO these operations can be carried out by AML programs. About 30 AML scripts have been written to automate most of the important data-handling operations in this application. The first operation in the generalized algorithm (Table 1) is updating the INFO tables from test files. Much of the raw data is in the form of text files so a decision was made that all INFO tables should be automatically updatable from the original text files. Text files are easy to transmit from one computer to another and can be edited either directly or stream edited using a pattern-matching programming language. Updating of the INFO tables takes place in a &DATA ARC INFO block containing a PURGE statement to remove existing records and then the ADD statement to create a new set of records.

Creation of the text input file for TETransgeo is by cursor processing. Cursors are declared at different levels in the data hierarchy and the file is written by an algorithm consisting of nested loops (Figure 4). Since several locations in a field generally have identical crop schedules,

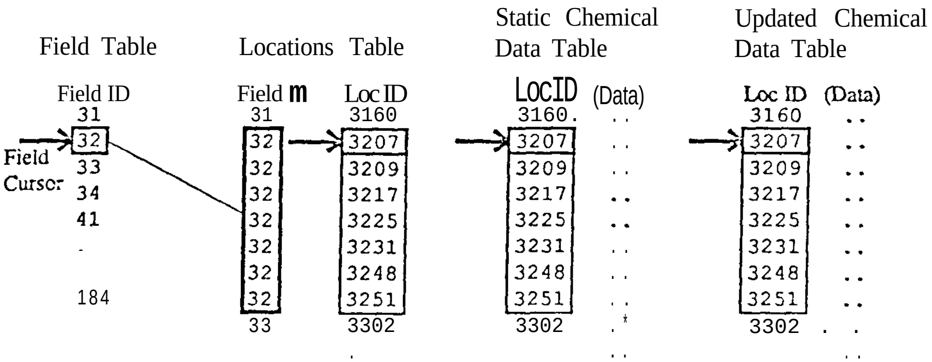


Figure 4. Cursor processing to extract data from INFO tables. Field ID number selects a set of locations from the location table which points to location-specific data in other tables.

a reverse list operation lists in the text file the location ID numbers for all locations having a particular crop schedule. This avoids duplication of crop schedules.

TETransgeo is run using an &SYS command. All the data required by TETransgeo exists in the text file so there are no command-line arguments. TETransgeo can return either a commented file which has the names of all variables and their values in a formatted output or a file which has one data value per line. The first type of file is useful for debugging and the second type can be read directly by the IMPORT command in INFO.

The last step in the generalized algorithm is the updating of the INFO tables from the results calculated in TETransgeo. This takes place through an &DATA ARC INFO block containing an IMPORT statement. The preexisting data, if any, are purged and a new set of records is generated through the ADD command. The IMPORT statement is a factor of 2-3 faster than using cursors to accomplish the same operation.

## INITIAL RESULTS

One application of TETransgeo is in the development of irrigation management strategies. Assessment of the salt-loading to groundwater is one component of the characterization of non-point source pollution. TETransgeo calculates the amount of a solute entering the groundwater at a given location and can therefore be used to estimate solute-loading over an area. To test the capability of the program we have run simulations of vertical water and solute transport at 284

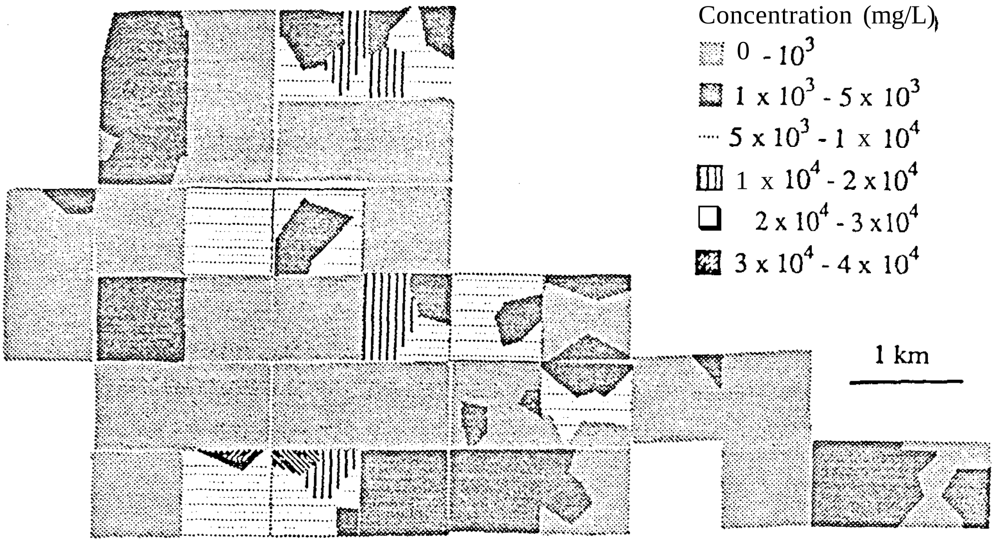


Figure 5. Compu;ted total dissolved solids in the top 0.3 m of the soil profile, 9/30/1991.

locations in the Broadview Water District. Figure 5 shows TDS concentration in the top 0.3 m at the end of the time period over which the calculation was run: May of 1991 to September 30, 1991. Several quarter-sections were fallow and therefore not irrigated during this period (Figure 1). The result was substantially decreased water content and correspondingly higher concentration

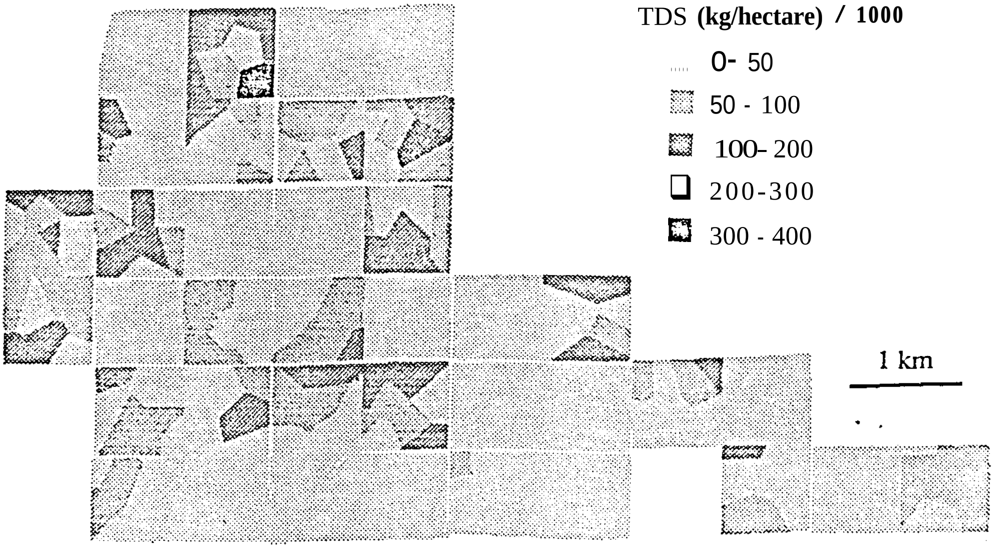


Figure 6. Computed variation in flux of total dissolved solids into groundwater (Brondview Water District, summer, 1991). Irrigation amount and TDS flux are highly correlated.

of dissolved solids near the surface. Results also indicate that solute loading to groundwater varies from zero in fallow areas to approximately 4 x $10^5$ kg/hectare (Figure 6). Quarter-sections that received large amounts of irrigation water and had high initial salt content showed a significantly greater amount of salt-loading to groundwater (Figure 6).

The operation of the TETransgeo program required about 1 minute of elapsed time on a SPARC1 2 workstation. The amount of time required to run AML's that read and write the two text files and update INFO tables is over an hour. This means that optimization of the overall calculation is best achieved by improving the efficiency of the AML scripts.

## CONCLUSION

A one-dimensional functional model (TETransgeo) that computes water and solute transport in the variably-saturated zone above the water table has been coupled with the ARC/INFO geographic information system. The reason for this coupling is utilization of the GIS to store values of the parameters required by the transport model. Arc Macro Language (AML) scripts are the essence of the coupling process. Several scripts write a data file, using data stored in several INFO files, which is read by the external TETransgeo program. Other scripts control the reading of data from a results file written by the external program. A data set collected for an agricultural field area located in the San Joaquin Valley was used as a test data set for the new program and the AMLs. Test runs on this data set indicate that solute-loading of groundwater is strongly-correlated with high initial salinity and application of greater amounts of irrigation water. The execution time for the external program, including its file manipulation operations, is about 2% of the total time required to run a full calculation. These results indicate that optimization of the reading/writing of external files by the GIS is the most significant potential area for performance enhancement.

## ENDNOTES

1. The use of brand or trade names in this report is for identification purposes only and does not constitute endorsement by the US Dept. of Agriculture.

## REFERENCES

Addiscott, T.M. "A simple computer model for leaching in structured soils." J. Soil Sci., 28 (1977): 554-563.

Addiscott, T.M. and Wagenet, R.J. "Concepts of solute leaching in soils: a review of modeling approaches." J. Soil Sci., 36 (1985): 411-424.

Beven, K. and P. Germann. "Macropores and water flow in soils." Water Resour. Res., 18 (1982.): 1311-1325.

Biggar, J.W. and D.R. Nielsen. “Spatial variability of the leaching characteristics of a field soil.” Water Resour. Res., 12 (1976.): 78-84.

Bond, W.J. and Smiles, D.E. “Predicting the average movement of reactive solutes in soils.” Soil Sci., 104 (1988): 227-233.

Coats, K.H. and B.D. Smith. “Dead end pore volume and dispersion in porous media.” Soc. Pet. Eng. J., 4 (1964): 74-84.

Cot-win, D.L. “Spatial and temporal estimates of hydraulic bypass in soil using chloride distributions.” Paper presented at Natl. Symposium on Preferential Plows, ASAE, December 1991, Pp. 202-213.

Cot-win, D.L., and Waggoner, B. L. “TETtans: A user-friendly, functional model of solute transport..” Water Sci. Tech., 24 (199 1): 57-65.

Cot-win, D.L., Waggoner, B.L., and Rhoades, J. D. “A functional model of solute transport that accounts for bypass.” J. of Environ. Qual., 20 (1991): 647-658.

Cot-win, D. L., Vaughan, P.J.. Wang, H., Rhoades, J. D., and Cone, D. G., “Coupling a solute transport model to a GIS to predict solute loading to the groundwater for a non-point source pollutant.” Paper presented at Proc. of the ASAE Applic. of Adv. Info. Tech. for Mgmt. of Nat. Res. Conf. Spokane, WA, June 17-19, 1993 (In Press).

Deans, H.H. “A mathematical model for dispersion in the direction of flow in porous media.” Soc. Pet. Eng. J., 3 (1963): 49-52.

Molz, F.J. “Models of water transport in the soil-plant system: a review.” Water Resour. Res., 17 (198 1): 1245-1260.

Petach, M.C., Wagenet, R.J., and DeGloria, S.D. “ Regional water flow and pesticide leaching using simulations with spatially distributed data.” Geoderma, 48 (1991): 245-269.

Rhoades, J.D. “Intercepting, isolating and reusing drainage waters for irrigation to conserve water and to protect water quality.” Agric. Water Mgmt., 16 (1989): 37-52.

Sudicky, E.A. “A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process.” Water Resour. Res., 22 (1986.): 2069-2082.

van Genuchten, M. Th., and Jury, W. “Progress in saturated flow and transport modeling.” Rev. Geophys. 25 (1987): 135-140.

van Ommen, H.C. “Calculating the quality of drainage water from non-homogeneous soil profiles with an extension to an unsaturated-saturated groundwater quality model including bypass flow.” Agric. Water Mgmt., IO (1985.): 293-304.