

# State-Space Mixing Cell Model of Unsteady Solute Transport in Unsaturated Soil

V J Bidwell

Lincoln Environmental, P O Box 84, Lincoln University, Canterbury, New Zealand

**Abstract** The purpose of this model is to enable implementation of the theory of linear systems control in operational management of waste and fertiliser applications onto land, so that the underlying groundwater is protected from pollution by leachate. The state-space form of the model enables use of the extensive theory and available software on stochastic linear systems. In particular, the Kalman filter is relevant to the imperfectly understood and highly variable processes of solute transport and transformation in field soils. The series of mixing cells was selected as a linear system model of one-dimensional, vertical, advective-dispersive transport, and based on cumulative soil water drainage as the index variable for application to unsteady flow in unsaturated soil. For each cell, solute transfer between mobile and immobile soil water, as well as equilibrium and nonequilibrium linear adsorption, are represented as lumped processes by two fractions linked by rate-limited transfer. The resident solute concentrations in the cell fractions are the states of the system. The complete model of solute transport and transformation for a uniform soil has four parameters, and can be described in MATLAB<sup>®</sup> with about ten lines of code. The software library can then be used to produce the discrete form of the model, which is unconditionally stable for any drainage interval, as well as to implement state estimation and control algorithms. A demonstration of the model is reported for <sup>35</sup>S-labelled sulphate leached from five replicated lysimeters (800 mm diameter, 1100 mm depth) of an undisturbed field soil (a free-draining silt loam) under pasture receiving rainfall and irrigation. The results show satisfactory one-step-ahead forecasts with the Kalman filter for the period of record, and a forecast is given of the complete response to the solute pulse application beyond the data record.

## 1. INTRODUCTION

Application of fertilisers or wastes to the land surface can degrade the quality of underlying groundwater if these substances are in excess of crop requirements or the assimilative capacity of the soil. The model presented in this paper is designed to support management of the amount and timing of substance applications to the land so that groundwater quality is protected. The proposed management process involves monitoring leachate quality by means of lysimeters and using these data for operational decision making.

Each lysimeter is an undisturbed column of soil, surrounded by a sealed vertical casing, from which freely draining soil water is collected for analysis (Cameron *et al.*, 1992). The collected leachate is a 100% sample of drainage from the horizontal area of the lysimeter and therefore substance mass balances are measured for the soil volume. A reliable estimate of water and solute mass flux in the field is obtained with appropriate replication of lysimeter units.

For reasons explained in the following section, cumulative drainage is the index variable rather than time. For practical monitoring, leachate data are collected after significant rainfall or irrigation events, and the drainage interval since the previous collection is unknown. Therefore, the leachate concentration data represent the mean values during these irregular intervals of cumulative drainage.

## 2. MODEL

### 2.1 State-Space for Monitoring and Control

Transport and transformation of substances in soil involves processes which are highly variable and have properties which are difficult to measure accurately. Management under these circumstances is facilitated by application of the theory of stochastic control (e.g., Bertsekas, 1976; Davis and Vinter, 1985) which enables formulation of rules for monitoring and control with optimal performance. This theory requires the processes of interest be described by a mathematical model with the "state-space" structure which relates the measured outputs  $Y(t)$  of a system to the internal states  $X(t)$  in response to inputs  $U(t)$  in the form:

$$\begin{aligned}\frac{dX(t)}{dt} &= AX(t) + BU(t) \\ Y(t) &= CX(t) + DU(t)\end{aligned}\quad (1)$$

The coefficient matrices  $A$ ,  $B$ ,  $C$ , and  $D$  for the present paper are assumed to be constant with respect to the index variable  $t$ . A model in the form of (1) can easily be coded in MATLAB<sup>®</sup> software for access to a library of algorithms for linear system control (Grace *et al.*, 1992).

## 2.2 Unsteady, Unsaturated Solute Transport

For one-dimensional, vertical, steady, advective-dispersive transport of solute through soil, the probability density function  $f(x,t)$  of the distance  $x$  travelled by solute particles in time  $t$  is:

$$f(x,t) = \frac{1}{\sqrt{2\pi(2Dt)}} \exp\left\{-\frac{(x-Vt)^2}{2(2Dt)}\right\} \quad (2)$$

where  $V$  is the mean pore water velocity and  $D$  is the dispersion coefficient. The distribution (2) is Gaussian, centred on the mean travel distance  $Vt$  and with variance  $2Dt$ . There is experimental and analytical evidence that this process is applicable to non-uniform water content (De Smedt and Wierenga, 1978) and transient water flow (Wierenga, 1977; Jury *et al.* 1990). This behaviour depends on dispersion being linearly related to velocity by the dispersivity  $\lambda$ . Substitution of  $D=\lambda V$  into (2) results in dispersion which is dependent only on the mean distance travelled by the solute particles. If  $I$  is the total amount of vertical drainage which has transported the solute pulse down from the ground surface, then the mean depth of travel is  $I/\theta$  where  $\theta$  is the fraction of mobile soil water in the profile after drainage has ceased. With the additional substitution  $Vt=I/\theta$  in (2) it can be shown that the distribution with cumulative drainage  $I$ , at an observation depth  $L$  is:

$$f(I,L) = \frac{L}{I} \frac{1}{\sqrt{2\pi(2\lambda I/\theta)}} \exp\left\{-\frac{(L-I/\theta)^2}{2(2\lambda I/\theta)}\right\} \quad (3)$$

$$\text{mean} = L\theta \quad \text{variance} = 2\lambda L\theta^2$$

The ratio of variance to mean for (3) is  $2\lambda\theta$ , which is independent of the travel distance. This information is used in the next section to develop a model to simulate (3) with a structure suitable for (1).

## 2.3 Mixing Cell Simulation of Advection-Dispersion

The mixing cell concept has a long history in dispersion modelling (Bear, 1969; Bajracharya and Barry, 1994) but is generally presented as an approximation to the finite difference solution of the advection-dispersion equation, with optimal results when the spatial step is equal to twice the dispersivity. Numerical dispersion and stability restrict the size of the time step. Bajracharya and Barry (1992) developed an "improved" mixing cell model which overcomes these problems but it does not have the state-space structure desired for the present application.

A mixing cell is a volume  $W$  of water (per area transverse to the flow) through which flows a water flux  $q(t)$ . Solute of concentration  $c_i(t)$  in the inflow is instantaneously mixed with the contents of the cell, and the outflow concentration  $c_o(t)$  is the same as that within the cell. Solute mass balance for the cell is:

$$W \frac{dc_o(t)}{dt} = q(t)c_i(t) - q(t)c_o(t) \quad (4)$$

By substituting incremental drainage  $dI=q(t)dt$  into (4), and omitting the time index for convenience:

$$W \frac{dc_o}{dI} + c_o = c_i \quad (5)$$

which is a first order linear differential equation. The impulse response solution to (5) gives the distribution of solute transit through the cell in terms of  $I$  as:

$$f(I) = \frac{1}{W} \exp\left\{-\frac{I}{W}\right\} \quad (6)$$

which is the exponential distribution with mean  $W$  and variance  $W^2$ . The ratio of variance to mean is  $W$ , and by putting  $W=2\lambda\theta$  the mixing cell becomes a simple mathematical concept for simulating the first two moments of (3). The number of cells required is the mean of (3),  $L\theta$ , divided by the mean for one mixing cell,  $2\lambda\theta$ , which is  $n=L/2\lambda$ . The resulting impulse response for  $n$  cells in series is the gamma distribution:

$$f(I) = \frac{I^{n-1}}{(n-1)!(2\lambda\theta)^n} \exp\left\{-\frac{I}{2\lambda\theta}\right\} \quad (7)$$

The shape of (7) becomes closer to (3) as  $n$  (or  $L/2\lambda$ ) increases, and the mixing cell series is a good practical representation of advection-dispersion. The vertical space step is restricted to a size of  $2\lambda$  but this is not usually a practical difficulty given the precision of soil profile data.

The differential equations for the series of  $n$  mixing cells can be expressed in terms of (5), with  $a=1/W=1/2\lambda\theta$ :

$$\begin{aligned} \frac{dc_1}{dI} &= -ac_1 + ac_i \\ \frac{dc_2}{dI} &= -ac_2 + ac_1 \\ \dots & \dots \dots \dots \\ \frac{dc_n}{dI} &= -ac_n + ac_{n-1} \\ c_o &= c_n \end{aligned} \quad (8)$$

which has the state-space structure of (1) when written as:

$$\frac{d}{dI} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \dots \\ c_{n-1} \\ c_n \end{bmatrix} = \begin{bmatrix} -a & 0 & \dots & \dots & \dots & 0 \\ a & -a & 0 & \dots & \dots & 0 \\ 0 & a & -a & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a & -a & 0 \\ 0 & \dots & \dots & 0 & a & -a \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \dots \\ c_{n-1} \\ c_n \end{bmatrix} + \begin{bmatrix} a \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{bmatrix} c_i \quad (9)$$

$$c_o = \begin{bmatrix} 1 \\ \dots \\ 0 \end{bmatrix} c_n + \begin{bmatrix} 0 \\ \dots \\ 1 \end{bmatrix} c_i$$

The continuous-drainage model (9) is the basis for obtaining the discrete state-space models for any drainage interval by means of MATLAB® software. These models are unconditionally stable, and can include filtering to account for leachate concentration being the mean value over the drainage interval rather than the instantaneous value.

## 2.4 Solute Transformations

With advective-dispersive transport represented by the system of mixing cells, solute transformations may be considered as lumped processes within each mixing cell subsystem (Sardin *et al.*, 1991). The following discussion considers only linear, equilibrium and nonequilibrium solute transformations. These are described from a process point of view in this section and will then be re-parameterised in terms of dynamic behaviour in section 2.5.

### 2.4.1 Mobile-Immobile Soil Water

Each mixing cell has a mobile water fraction  $\theta_m$  with solute concentration  $c$ , which is part of the advective-dispersive transport through all the cells, and an immobile fraction  $\theta_{im}$  with concentration  $s$ , which is connected only to the mobile fraction of the same cell (Figure 1). The differential equations for the  $j$ th cell are, for mass balance:

$$2\lambda\theta_m \frac{dc_j}{dI} + 2\lambda\theta_{im} \frac{ds_j}{dI} = c_{j-1} - c_j \quad (10)$$

and rate-limited diffusion:

$$\frac{ds_j}{dI} = r(c_j - s_j) \quad (11)$$

in which  $r$  is a transfer coefficient with dimension (drainage units)<sup>-1</sup> rather than (time)<sup>-1</sup>, because the nonequilibrium processes are with reference to cumulative drainage.

### 2.4.2 Linear Adsorption-Desorption

Linear equilibrium adsorption within either the mobile or immobile fraction can be quantified by a "retardation coefficient"  $R$ , specific to each fraction, which is the ratio of total resident mass in equilibrium, per volume of soil, to the solute mass in the soil water fraction. Inclusion of equilibrium adsorption in (10) gives:

$$2\lambda\theta_m R_m \frac{dc_j}{dI} + 2\lambda\theta_{im} R_{im} \frac{ds_j}{dI} = c_{j-1} - c_j \quad (12)$$

The only form of nonequilibrium adsorption-desorption considered in the present conceptual discussion is that due to the rate-limited diffusion described by (11). Therefore,

any equilibrium adsorption within the immobile fraction is observed externally as being nonequilibrium.

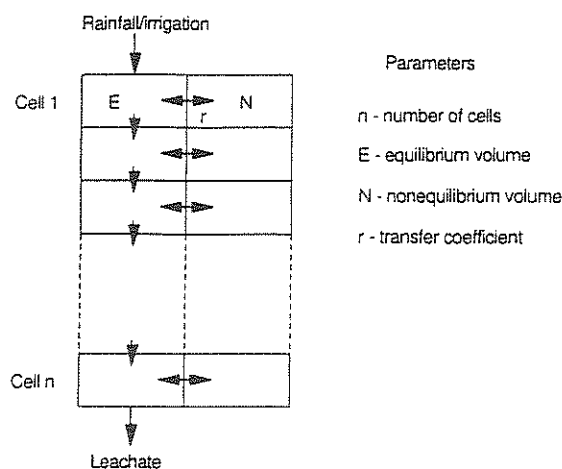


Figure 1: Conceptual structure of the mixing cell model

## 2.5 The Complete Dynamic Model

By making the substitutions:

$$E = 2\lambda\theta_m R_m \quad N = 2\lambda\theta_{im} R_{im} \quad (13)$$

in (12), and substituting (11) into (12), the differential equations for the  $j$ th cell become:

$$\begin{aligned} \frac{dc_j}{dI} &= \frac{1}{E} c_{j-1} - \frac{(1+rN)}{E} c_j + \frac{rN}{E} s_j \\ \frac{ds_j}{dI} &= rc_j - rs_j \end{aligned} \quad (14)$$

The four parameters  $E$ ,  $N$ ,  $r$ , and the number of cells  $n$ , are the maximum set that can be identified from a series of leachate concentration data. However, the equivalences (13) can be useful for suggesting initial values from other process knowledge, or for relating calibrated parameter values to physical reality. Any form of adsorption-desorption with first-order dynamics (relative to drainage) will be indistinguishable from the concepts used in the model development.

The state-space structure (9) is extended to include the additional states  $s_j$  ( $j=1, n$ ) and the terms in (14):

$$\frac{d}{dt} \begin{bmatrix} c_1 \\ \dots \\ c_n \\ s_1 \\ \dots \\ s_n \end{bmatrix} = \begin{array}{c|c} \text{CC} & \text{CS} \\ \hline \text{SC} & \text{SS} \end{array} \begin{bmatrix} c_1 \\ \dots \\ c_n \\ s_1 \\ \dots \\ s_n \end{bmatrix} + \frac{1}{E} \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix} c_i$$

$$c_o = c_n$$

$$\text{CC} = \frac{1}{E} \begin{bmatrix} -1 & 0 & \dots & \dots & \dots & 0 \\ 1 & -1 & 0 & \dots & \dots & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 & -1 & 0 \\ 0 & \dots & \dots & 0 & 1 & -1 \end{bmatrix}$$

$$\text{CS} = \frac{rN}{E} I(n) \quad \text{SC} = r I(n) \quad \text{SS} = -r I(n) \quad (15)$$

where  $I(n)$  is the identity matrix of size  $n$ .

The model (15) can be described in MATLAB<sup>®</sup> with about ten lines of code.

### 3. DEMONSTRATION

The emphasis in the following demonstration is on the role of the model in describing prior knowledge about the dynamic behaviour of solute transport, which can then be incorporated into a Bayesian approach to prediction and control with uncertain information.

#### 3.1 Experimental Data and Parameter Estimation

The data are from an experiment (Fraser *et al.*, 1994) on five replicated lysimeters (800 mm diameter, 1200 mm depth) of undisturbed soil (a free-draining silt loam) under pasture, which received rainfall and irrigation over a period of one year. Several solutes were applied as a surface pulse to the soil (at field capacity) with 4 mm of water to represent animal urine, and immediately followed by 10 mm of simulated rainfall. Leachate drainage, in this research environment, was sampled at regular intervals of 30 mm. Solute breakthrough data for <sup>35</sup>S-labelled sulphate from the same experiment (McLaren *et al.*, 1993) were selected for demonstration in the present paper.

The solute breakthrough curves (BTC) do not represent the pulse response of transport by vertical soil water movement through the soil, because of plant uptake of

solute and water in the root zone. Solute, which does arrive at the measurement depth, has been transported by vertical water movement which is larger than the measured cumulative drainage. The result is that the solute input history, in terms of measured cumulative drainage, is unknown and therefore the usual system identification techniques cannot be applied.

The parameters  $n$ ,  $E$ ,  $N$ , and  $r$  in (15) were fitted by visually comparing the impulse response of (15) with the BTC data, with both scaled to unit area for the length of record. Figure 2 shows the nature of the comparison for the sulphate data. The criterion for a satisfactory "fit" was that the steepest and flattest gradients of the impulse response should match corresponding features of the BTC, as a measure of dynamic behaviour. The parameters were fitted sequentially, because the nitrate data were described adequately by advection-dispersion ( $n$ ,  $E$ ), bromide data required the addition of diffusion ( $r$ ,  $N$ ) and the sulphate data needed adsorption in the immobile fraction (same  $n$ ,  $E$ ,  $r$ , but different value of  $N$ ). Space limitation in the present paper does not permit a complete reporting of this process.

The parameter values for the sulphate data used in this demonstration (Figure 2) are:  $n=20$ ,  $E=10.3$  mm,  $N=28$  mm,  $r=0.005$  mm<sup>-1</sup>. For the measurement depth  $L=1100$  mm, these values correspond to dispersivity  $\lambda=27.5$  mm, and mobile water fraction  $\theta_m=0.19$ . The value of  $N=4$  mm obtained for bromide (no adsorption) implies that the immobile water fraction is  $\theta_{im}=0.07$  and the retardation coefficient for sulphate in the immobile fraction is  $R_{im}=7$ .

#### 3.2 State Estimation with the Kalman Filter

Application of substances to the land surface in response to feedback information from monitoring of leachate concentration may be considered as a controlled system which tracks the consumption of solutes by plant uptake, decay, or other biochemical processes. The best information for feedback to operational action is based on estimation of the current states of the system. For the model (15) the states are the values of  $c_j$  and  $s_j$ , the resident concentrations of the cells.

Uncertainty about process knowledge and imperfect measurement can be accounted for by appending error components to (15). The relative magnitudes of the variance of these error terms for the demonstration are shown in Figure 3. These are based on the error variance of the experimental data for the measurement error, and subjective judgement about processes in the soil profile for the state errors. The most significant feature of the latter is the high relative variance ascribed to the root zone (cells 1-5), especially near the ground surface, in comparison to the lower value for the deeper part of the profile where there is insignificant uptake.

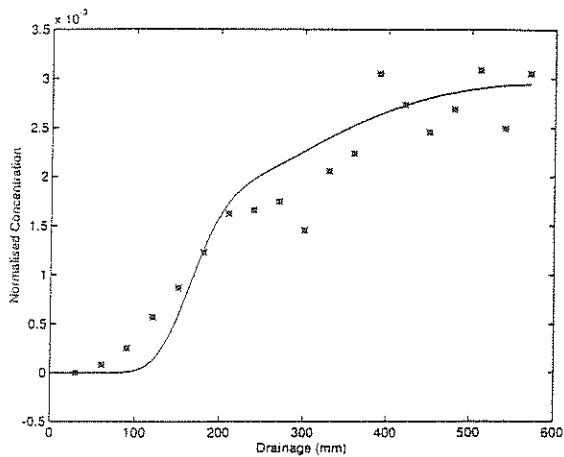


Figure 2: Impulse response of the model (-) for use with  $^{35}\text{S}$ -labelled sulphate data (\*).

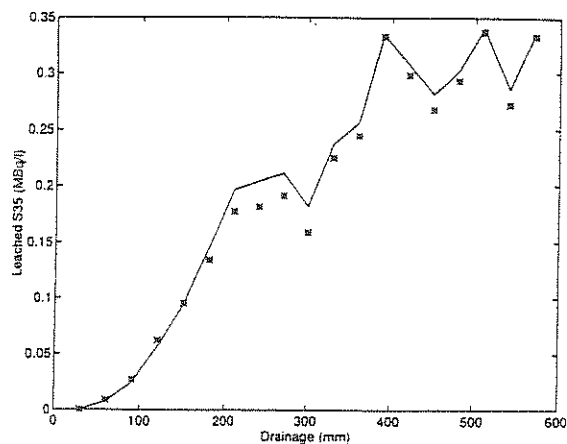


Figure 4: One-step-ahead forecast (-) of the  $^{35}\text{S}$  data (\*).

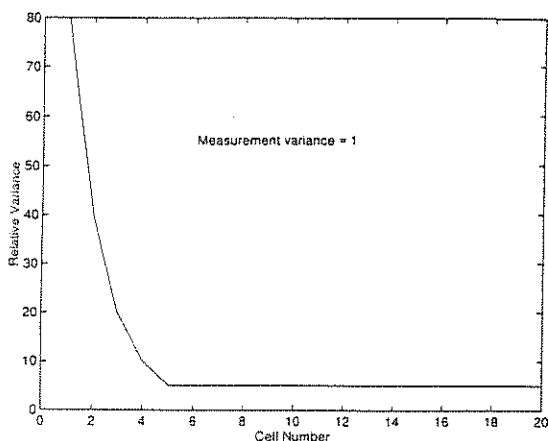


Figure 3: Uncertainty about process knowledge expressed as error variance relative to measurement error.

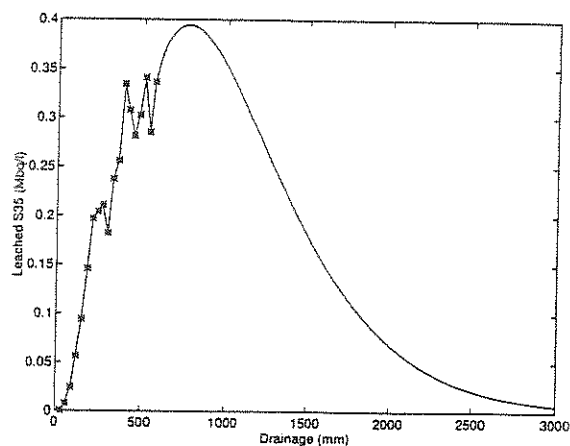


Figure 5: Forecast of the complete  $^{35}\text{S}$  breakthrough curve.

The Kalman filter (e.g., Davis and Vinter, 1985: 117-127) is a Bayesian statistical procedure for combining prior knowledge about a dynamic system together with information from measured data to produce an optimal estimate of the states of the system, given assumptions about the error properties. In the present demonstration, model (15) was combined with the sulphate BTC data and the uncertainty profile of Figure 3, by means of the MATLAB<sup>®</sup> Control System procedures "dlqe" and "destim" (Grace *et al.*, 1992), to form the state estimator which is itself in state-space form.

The state estimator was tested for its ability to forecast the system states ( $c_j$ ,  $s_j$ ), and hence the output  $c_o$ , at one drainage interval beyond each observed BTC data value, using the procedure "dlsim" with the data series as input. The initial state values at the beginning of the series were set to zero except for the top cell which was determined from the value of the solute pulse applied at the ground surface. The water depth applied with the solute pulse

(14 mm) was approximately the combined water fraction for one cell, and therefore the solute pulse dose of  $740 \text{ MBq m}^{-2}$   $^{35}\text{S}$  was allocated to an equilibrium concentration in the total water storage equivalent ( $E+N=38.3 \text{ mm}$ ) of the top cell as an approximation to initial conditions. Figure 4 shows the results of the one-step-ahead forecasts.

Forecasts of leachate concentration can, of course, be made for any number of drainage intervals beyond a measured value, but then the knowledge basis is the model (15) with initial state values as estimated from this last value and with no further correction. Figure 5 shows the full leachate response to the solute pulse as forecasted (using "dinitial") from the last measured value. Approximately 14% of the  $^{35}\text{S}$  dose had been leached at the end of the measured portion of the BTC, whereas the complete estimated response accounts for 66% at 3000 mm of leachate.

#### 4. DISCUSSION AND CONCLUSIONS

The purpose of the model development was to enable access to control theory and supporting software so that this technology could be applied to managing the impact on groundwater of applying wastes and fertilisers to the land. The model is designed for application to leachate data from undisturbed field soils where uncertainty of process parameters and measurement is high. Processes of advective-dispersive transport, mobile-immobile soil water, and linear adsorption (equilibrium and nonequilibrium) were explicitly included in the model development, but have been shown to be represented by a smaller set of dynamic characteristics. Sources and sinks, such as plant uptake, are implicitly included in the concept of this model being part of a controlled system which is tracking these assimilative processes.

Demonstration of the model was centred on its role as a component in an application of the Kalman filter, a pivotal theory in estimation and control under uncertainty. The results (Figure 4) display satisfactory dynamic tracking in response to new data but the error between forecast and measurement shows some persistence, which indicates imperfections in specification of the model and uncertainty profile (Figure 3). The forecast of the complete response, at 1100 mm depth, to a solute application at the land surface (Figure 5) demonstrates again how the model can be used to incorporate available data and subjective judgements about uncertainty into making an optimal estimate of dynamic behaviour.

#### 5. ACKNOWLEDGEMENTS

The experimental data used for the model development and demonstration were contributed by Ron McLaren (staff) and Trish Fraser (former postgraduate student), Soil Science Department, Lincoln University. The reported research is part of the Groundwater Quality Protection programme funded by the Foundation for Research, Science and Technology, New Zealand.

#### 6. REFERENCES

- Bajracharya, K. and D.A. Barry, Mixing cell models for nonlinear nonequilibrium single species adsorption and transport, *Water Resources Research*, 29(5), 1405-1413, 1992.
- Bajracharya, K. and D.A. Barry, Note on common mixing cell models, *Journal of Hydrology*, 153, 189-214, 1994.
- Bear, J., Hydrodynamic dispersion. Chapter 4 in *Flow Through Porous Media*, R.J.M. De Wiest (ed.), Academic Press, 1969.
- Bertsekas, D.P., *Dynamic Programming and Stochastic Control*, Academic Press, 1976.

- Cameron, K.C., N.P. Smith, C.D.A. McLay, P.M. Fraser, R.J. McPherson, D.F. Harrison and P. Harbottle, Lysimeters without edge flow: an improved design and sampling procedure, *Soil Science Society of America Journal*, 56, 1625-1628, 1992.
- Davis, M.H.A. and R.B. Vinter, *Stochastic Modelling and Control*, Chapman and Hall, 1985.
- De Smedt, F. and P.J. Wierenga, Solute transport through soil with nonuniform water content, *Soil Science Society of America Journal*, 42, 7-10, 1978.
- Fraser, P.M., K.C. Cameron and R.R. Sherlock, Lysimeter study of the fate of nitrogen in animal urine returns to irrigated pasture, *European Journal of Soil Science* 45, 439-447, 1994.
- Grace, A., A.J. Laub, J.N. Little and C.M. Thompson, *Control System Toolbox - Users Guide*, The Mathworks Inc., 1992.
- Jury, W.A., J.S. Dyson and G.L. Butters, Transfer function model of field-scale solute transport under transient water flow, *Soil Science Society of America Journal*, 54, 327-332, 1990.
- McLaren, R.G., K.C. Cameron and P.M. Fraser, A comparison of the effects of subsoiling on plant uptake and leaching losses of sulphur and nitrogen from a simulated urine patch, *Plant and Soil*, 155/156, 375-378, 1993.
- Sardin, M., D. Schweich, F.J. Leij and M. Th. van Genuchten, Modeling the nonequilibrium transport of linearly interacting solutes in porous media: a review, *Water Resources Research*, 27(9), 2287-2307, 1991.
- Wierenga, P.J., Solute distribution profiles computed with steady-state and transient water models, *Soil Science Society of America Journal*, 41, 1050-1055, 1977.