

## **9. MONTE CARLO SIMULATION OF SOLUTE TRANSPORT FROM SMALL SOURCES IN DEEP, VARIABLY SATURATED SOILS**

### **9.1 Introduction**

Contamination of groundwater resources has been one of the largest environmental concerns of the last decade. Most pollution sources, whether they are agricultural, domestic, or industrial, non-point or point pollution sources, are at or near the surface. Much of the pollutants must therefore travel through the unsaturated zone between the surface and the groundwater before reaching the water table. A thorough understanding of the transport processes in the unsaturated zone is essential to assess the contamination risk of groundwater resources and to predict the travel time from a pollution source to a drinking water well.

The advection-diffusion equation (ADE, eqn. 5-1) is generally used to describe the movement of solutes in the unsaturated zone. Like the study of groundwater pollution the assessment of solute and contaminant transport through the unsaturated zone is hampered by the uncertainty caused by the heterogeneous structure of the geologic material. However, the list of soil properties that determines solute displacement and plume spreading includes more than just the saturated hydraulic conductivity, since the unsaturated hydraulic conductivity depends directly on the soil water tension or the soil water content. In the previous chapter some of the effects of soil spatial variability on the soil water tension and velocity distribution in the soil were studied. Field studies have shown that the heterogeneity of the soil moisture flux leads to variable, time- and scale-dependent parameters in the ADE.

Stochastic transport models have been developed to describe average solute displacement and plume spreading in variable flux fields (e.g. Gelhar and Axness, 1983; Dagan, 1982, 1984; Neuman et al., 1987). The objective of these stochastic models is to cast solute transport in heterogeneous geologic media into an equivalent quasi-homogeneous ADE

i.e., to define the mean transport velocity and the average plume-spreading in terms of the statistical parameters that describe the spatially random velocity field. Stochastic analysis of transport in porous media such as Dagan (1984, 1988), Neuman and Zhang (1990), Rubin (1990), Zhang and Neuman (1994c) suggests that the effective dispersion or "macrodispersion" varies with time. These models have primarily been developed and applied to transport in groundwater, although they are in principle valid also for unsaturated flow.

Past efforts of modeling unsaturated solute transport have concentrated on the analysis of one-dimensional transport in steady state flow fields. A basic tenet of such models is the assumption that the soil consists of an ensemble of independent, homogeneous vertical stream tubes. For each stream-tube, the vertical flow and transport parameters are defined randomly. (Dagan and Bresler, 1979; Bresler and Dagan, 1981; Amoozegar-Fard et al., 1982; Jury, 1982; Simmons, 1982; Jury et al., 1986; Butters and Jury, 1989; Destouni and Cvetkovic, 1989, 1991). Horizontal solute displacement and lateral dispersion is neglected, an assumption that has been found valid mainly in applications to field-scale solute transport from non-point pollution sources through shallow soils. Destouni (1992) incorporated the effect of vertically variable saturated conductivity into the stream-tube model.

Field studies of multi-dimensional unsaturated transport have been documented by Ellsworth and Jury (1991) and by Wierenga et al. (1991). These studies have demonstrated the limitations of one-dimensional unsaturated transport models if the lateral extension of the plume is small relative to the depth of groundwater. Lateral solute spreading and plume contraction and expansion were found to be important mechanisms affecting the movement of the solute plume. The dispersion of solutes in a more general two-dimensional heterogeneous soil has been studied numerically by Russo (1991) and by Russo and Dagan (1991) who suggested that macrodispersion of unsaturated transport is amenable to the same stochastic transport analysis as those known for aquifer contamination. More recently, Russo (1993a) combined the three-dimensional stochastic formulation of unsaturated flow in heterogeneous media by Yeh et al. (1985a,b) with the Lagrangian transport analysis by Dagan (1982, 1984) to derive analytical

expressions for the displacement of the center of a plume and the average plume spreading. In a complementary study, Russo (1993b) derived the temporal moments of solute arrival time in a three-dimensionally heterogeneous soil based on the work by Cvetkovic et al. (1992).

Both spectral analysis of unsaturated flow and Lagrangian analysis of solute transport are limited to mildly heterogeneous media with a normally distributed velocity. The findings in the previous chapter suggest that the velocity pdf is neither normally distributed nor well described by the first order spectral perturbation analysis. It was found that spectral analysis may be much more limited with respect to predicting velocity moments than with respect to predicting the soil water tension distribution. While Russo's (1993a,b) analysis is a useful tool for the validation of numerical models and to obtain approximate estimates, it may be inappropriate for many field applications with highly heterogeneous soils.

In this study, transport through unsaturated soils is re-examined without assuming that the spatial variability is small, without limiting the study to one-dimensional transport of non-point pollution, and without depending on the assumption that transport is ergodic, which is invoked in the numerical study by Russo (1991) and which is necessary for analytical results such as those of Dagan (1984) to hold. Going beyond these restrictions allows one to analyze the behavior of plumes from sources of relatively small lateral extent occurring for example under leaky storage tanks and damaged liner systems. These sources can generally be viewed as point-sources. In contrast to solute plumes from non-point sources, the lateral solute movement plays an important role for point source plumes and is a significant source of uncertainty for transport models of heterogeneous soils. Solute transport from small sources in heterogeneous porous media is not ergodic i.e., the mean concentration derived stochastically is not identical to the actual concentration. Plumes from small sources (lateral extent on the order of or smaller than the correlation scale of the soil hydraulic conductivity) require several tens if not hundreds of correlation scales in mean travel distance before the ergodic condition is met and the concentration variance vanishes.

Therefore the Monte Carlo technique is applied to study the spatial and temporal

variability of non-ergodic solute plumes in variably saturated porous media. A variety of different hypothetical soil sites is investigated to establish a stochastic analysis of the solute concentration, the solute flux, and the solute plume spreading as a function of the various independent parameters characterizing a spatially variable soil. For efficiency, the two-dimensional steady-state head and velocity distribution in a vertical soil profile is computed using the ASIGN approach (Harter and Yeh, 1993; see chapter 7). For each realization of a velocity field the movement of a small plume through the soil is predicted as a function of time. From a set of 300 realizations within each Monte Carlo simulation the local concentration moments (mean and variance) are computed over time, the first and second moments of each individual plume are monitored over time, and solute flux is recorded and statistically analyzed as a function of location and time. The technical implementation and design of the Monte Carlo simulations is described in section 9.2. In section 9.3, the results for the concentration moments and the solute plume spreading are reported and compared to the linear macrodispersion model by Dagan (1988). In section 9.4 solute flux and travel time in various soils are investigated and compared to the Lagrangian particle travel time model by Cvetkovic et al. (1992). Both the linear macrodispersion model and the Lagrangian travel time analysis are coupled with the linearized first order perturbation analysis of unsaturated flow described in chapter 4. The chapter ends with summary and conclusions in section 9.5.

## 9.2 Implementation of the Monte Carlo Simulation and Statistical Methods

### 9.2.1 Transport Model

The same Monte Carlo procedure as described in chapter 8.2 is used for the stochastic analysis of transient transport. The procedure described in 8.2 computes steady-state, unsaturated velocity fields  $\mathbf{v}(\mathbf{x})$  given spatially correlated random realizations of  $f$  and  $a$  with means  $F$ ,  $A$ , variance  $\sigma_f^2$ ,  $\sigma_a^2$ , and an exponential spatial correlation characterized by the horizontal and vertical correlation scales  $\lambda_{fx}$  and  $\lambda_{fz}$ , respectively. For each realization of  $\mathbf{v}(\mathbf{x})$ , transient transport is simulated with an algorithm based on the modified method of characteristics (MMOC) and described in Yeh et al. (1993). Chapter 5 gives an overview of the numerical procedure and describes the particular implementation used for this study. In this chapter, only transport of non-reactive solutes is analyzed. Recall from chapter 5 that only advective transport is computed (for efficiency), while local or "pore-scale" dispersion is introduced through numerical dispersion. The effective equation solved by MMOC given a random realization  $\mathbf{v}(\mathbf{x})$  of the velocity field is:

$$\frac{\partial c(\mathbf{x},t)}{\partial t} + \mathbf{v}(\mathbf{x})\nabla c(\mathbf{x},t) = \nabla \cdot (\mathbf{D}_{num}(\mathbf{x})\nabla c(\mathbf{x},t)) \quad (9-1)$$

Although the numerical dispersion  $\mathbf{D}_{num}(\mathbf{x})$  cannot be rigorously quantified, preliminary numerical experiments showed that the resulting effective local dispersivity  $\mathbf{d}_{num} = (\mathbf{D}_{num} / V_z)$  is on the order of 1/10 to 1/100 of the element size and therefore much smaller than the correlation scale of the velocity variations (see chapter 8).  $V_z$  is the mean vertical velocity. It will be shown later that the effective local dispersivity is only of minor importance for the overall plume movement. The main contribution of the effective local dispersivity is limited to the first few time steps i.e., very early time ( $t \ll \lambda_{fz}/V_z$ ). At early time an initially small plume with uniform concentration  $c_0$  is numerically dispersed over a larger area due to the sharp concentration gradient at the boundary of the plume. Numerical dispersion is caused by the bilinear interpolation scheme employed in the modified method of characteristics to compute

the concentration at locations not identical to those of finite element nodes. At later times, the concentration gradients are much smaller, and the plume movement is dictated predominantly by advective processes. While numerical dispersion is artificial, its net effect is consistent with many field findings and with the stochastic transport theories of Gelhar and Axness (1983), and Neuman et al. (1987), which explicitly account for pore-scale dispersion. Implicit (numerical) or explicit (parametric) local dispersion is expected to affect primarily the local concentration variance (Dagan, 1982) and the asymptotic (i.e. late time) magnitude of the lateral macrodispersion (Gelhar and Axness, 1983; Neuman et al., 1987).

### **9.2.2 Moment Analysis**

In this analysis, the statistical description of three phenomena associated with solute transport is addressed: the statistical analysis of the spatial moments of the actual concentration plumes, the statistical moments of the local concentration, and the statistical moments of the solute flux at a given distance from the plume source. The first two phenomena are closely associated with the concept of macrodispersion (Dagan, 1982, 1984, 1988; Gelhar and Axness, 1983; Neuman et al., 1987) and are grouped together in the spatial analysis of solute transport. The latter analysis is often treated separately and deals primarily with the statistical analysis of particle travel times to a given distance from the source of the contamination (Cvetkovic et al., 1992; Russo, 1993b). Only recently, the spatial and temporal analysis of solute transport has been treated in a unified manner both numerically and analytically (Bellin et al., 1992; Zhang and Neuman, 1994a-d)

### 9.2.2.1 Spatial Concentration Distribution

Concentration is a non-stationary random field variable (RFV, see section 2.5.1) in space and time. Spatial plume moments and local statistical moments of the concentration distribution are therefore RFV functions of time. In this analysis the computation of the local concentration sample mean  $\langle c(\mathbf{x},t) \rangle$  and of the local sample variance  $\sigma_c^2(\mathbf{x},t)$  is limited to four discrete points in time (see below). The computation of local concentration moments is equivalent to the computation of local head sample moments (see chapter 8). No covariances and cross-covariances are computed. The spatial moments of the mass distribution in each solute plume are computed at each time-step of each realization. The zero order, first order, and second order spatial moments of an actual concentration plume are given as:

$$\begin{aligned}
 M_0(t) &= \sum_{i=1}^{NN} c(\mathbf{x}^i, t) \theta(\mathbf{x}^i) \Delta x^i \Delta z^i \\
 M_x(t) &= 1/M_0 \sum_{i=1}^{NN} c(\mathbf{x}^i, t) \theta(\mathbf{x}^i) \Delta x^i \Delta z^i x^i \\
 M_z(t) &= 1/M_0 \sum_{i=1}^{NN} c(\mathbf{x}^i, t) \theta(\mathbf{x}^i) \Delta x^i \Delta z^i z^i \\
 M_{xx}(t) &= 1/M_0 \sum_{i=1}^{NN} c(\mathbf{x}^i, t) \theta(\mathbf{x}^i) \Delta x^i \Delta z^i (x^i)^2 \\
 M_{zz}(t) &= 1/M_0 \sum_{i=1}^{NN} c(\mathbf{x}^i, t) \theta(\mathbf{x}^i) \Delta x^i \Delta z^i (z^i)^2
 \end{aligned} \tag{9-2}$$

$M_0(t)$  is the total mass in the finite element domain, where the finite element domain consists of NN nodes connecting rectangular elements of sidelength  $\Delta x$  and  $\Delta z$ .  $\theta(\mathbf{x}^i)$  is the arithmetic average of the water content in the four elements surrounding node  $i$ . In this study  $\theta$  is assumed constant throughout the domain (chapter 8).  $M_x(t)$  and  $M_z(t)$  are the horizontal and vertical position, respectively, of the center of mass of the plume  $c(\mathbf{x},t)$ .  $M_{xx}(t)$  and  $M_{zz}(t)$  are the horizontal and vertical moment of inertia, respectively. The moment of inertia is a measure

of the average plume spreading around the center of mass. The five spatial moments in (9-2) are computed for each time-step of each realization in the Monte Carlo simulation. Similar to the concentration moments, the spatial plume moments are a random process indexed on the real line representing time (see section 2.5). From the realizations of the random, time-dependent moment functions, sample means and sample variances are computed at 500 equal-distanced time intervals distributed over the total amount of simulation time (see below).

Due to the principle of mass conservation the sample mean  $\langle M_0(t) \rangle$  must be constant with time during the early part of the simulation when the solute plume is entirely confined within the boundaries of the finite element domain. Any variance in  $M_0(t) > 0$  would indicate mass balance errors due to the numerical transport solving method. Computing the variance of  $M_0(t)$  is therefore an important opportunity to assess potential mass balance errors in MMOC. The expected values of the first spatial moments,  $\langle M_x(t) \rangle$  and  $\langle M_z(t) \rangle$ , are a measure of the average plume displacement and must coincide with the center of mass of the mean plume concentration  $\langle c(\mathbf{x}) \rangle$ . The mean of the second spatial moment  $\langle M_{ii}(t) \rangle$  ( $i=x,z$ ) is a representative measure of the average spreading of the plume around its centroid.  $\langle M_{ii}(t) \rangle$  is not identical to the second moment  $X_{ii}(t)$  of the mean concentration plume (c.f. Dagan, 1990). The first and second spatial moment  $X_i$  and  $X_{ii}$  of the mean concentration plume  $\langle c(\mathbf{x},t) \rangle$  are computed as in (9-2) with  $c(\mathbf{x},t)$  replaced by  $\langle c(\mathbf{x},t) \rangle$  and  $M$  replaced by  $X$ . From statistical principles for turbulent mixing (Fischer et al., 1979), illustrated by Kitanidis (1988) and Dagan (1990) for porous media transport, it follows that  $X_{ij}$  and  $\langle M_{ij} \rangle$  are related through:

$$\langle M_{ii}(t) \rangle + \text{var}(M_i(t)) = X_{ii}(t) \quad i=x,z \quad (9-3)$$

where  $\text{var}(M_i)$  is the sample variance of the first spatial moment  $M_i$ .  $\text{var}(M_i)$  is a measure of the uncertainty regarding the actual center of a solute plume. Hence, the spatial spreading of the mean concentration plume  $X_{ii}$  is the sum of the expected moment of inertia of the actual plume plus the uncertainty about the center of mass of the plume (Figure 9.1). A general result of stochastic transport analysis is the fact that only for plumes of large initial lateral spreading or

at very late times the uncertainty regarding the center of the plume vanishes relative to the size of  $\langle M_{ii} \rangle$ . Then  $\langle M_{ii} \rangle$  and  $X_{ii}$  become interchangeable. Such a plume is called ergodic (chapter 2).

### 9.2.2.2 Solute Flux Characteristics

While much research has been devoted to define the spatial concentration distribution in terms of  $\langle c(\mathbf{x},t) \rangle$  and  $\sigma_c^2(\mathbf{x},t)$ , the stochastic analysis of the solute mass flux  $s(\mathbf{x},t)$  has only recently been investigated (Dagan et al., 1992; Cvetkovic et al., 1992, Russo, 1993; Neuman, 1993; Zhang and Neuman, 1994c). Solute mass flux is an important variable in many regulatory applications, where interest is not so much focused on the spatial distribution of a contamination plume, but on the temporal distribution of solute mass flux across a compliance boundary. It is of particular interest in the study of unsaturated transport, since a common remediation and site assessment question is: When, where, and how much solute mass will arrive at the water table? What is the uncertainty of the prediction due to variable travel times that are caused by the spatial heterogeneity of the soil?

Solute flux  $s(\mathbf{x},t)$  is defined as the mass of solute per unit area and unit time passing through a surface element of unit normal  $\eta$ . Neglecting pore-scale dispersion, it can be related to the resident concentration  $c(\mathbf{x},t)$ :

$$s(\mathbf{x},t) = \mathbf{s}(\mathbf{x},t) \cdot \boldsymbol{\eta} = c(\mathbf{x},t) \boldsymbol{\theta}_e \cdot \mathbf{v}(\mathbf{x}) \cdot \boldsymbol{\eta} \quad (9-4)$$

(9-4) is adopted specifically for use in the numerical model, such that  $s(\mathbf{x},t)$  can easily be computed from the resident concentration  $c(\mathbf{x},t)$ . The definition of (9-4) yields a flux-averaged concentration  $c_q = \int (\mathbf{s} \cdot \boldsymbol{\eta} \, dA) / \int (\boldsymbol{\theta}_e \cdot \mathbf{v} \cdot \boldsymbol{\eta} \, dA)$  equal to the resident concentration. Setting the flux averaged concentration equal to the resident concentration is justified since the advective mass flux is much larger than the dispersive mass flux (Parker and VanGenuchten, 1983). No diffusion is included in these simulations and since the effective dispersion is related to the

seepage velocity, the effective dispersive mass flux is always smaller than the advective mass flux.

In this study, a number of statistical tools are employed to analyze solute flux at a given compliance surface CS (CS reduces to a line CL in two dimensional transport). The total solute mass flux  $S(t)$  is the integrated mass flux across the compliance surface:

$$S(t) = \sum_{n=1}^{NC} s(\mathbf{x}_i, t) \Delta x_i \quad (9-5)$$

where NC is the number of finite element nodes along the horizontal CS and  $\Delta x_i$  is the average element width to the left and right of node  $i$ , in other words, the concentration is linearly weighted between nodes. The mean  $\langle S(t) \rangle$  and variance  $\sigma_s^2(t)$  of  $S(t)$  are computed from the individual realizations of the integrated mass breakthrough curves  $S(t)$ . For this study, four CLs are defined at dimensionless vertical distances  $Z' = Z/\lambda_{tz} = 5.4, 11.6, 17.8,$  and  $23.8$  from the solute source area.

In addition to the stochastic analysis of total mass breakthrough  $S(t)$ , a stochastic analysis of solute mass flux arrival time  $t_a(\mathbf{x}, s_0)$ , and solute peak flux time  $t_p(\mathbf{x})$  is undertaken. The solute mass flux arrival  $t_a(\mathbf{x}, s_0)$  is the time at which the solute mass flux  $s(\mathbf{x}, t)$  first exceeds some compliance mass flux  $s_0$  at the location  $\mathbf{x}$  on the CL. Nineteen different  $s_0$  are defined (see below) for each of which, the mean  $\langle t_a(\mathbf{x}, s_0) \rangle$ ,  $\langle t_p(\mathbf{x}) \rangle$  and the variance  $\text{var}_{t_a}(\mathbf{x}, s_0)$ ,  $\text{var}_{t_p}(\mathbf{x}, s_0)$  are computed. The peak flux time  $t_p(\mathbf{x})$  is equal to the time of highest concentration or peak solute flux, since the velocity field is at steady-state. The moments of the arrival and peak time are investigated only at one horizontal compliance surface in the center of the simulation domain. The results are normalized with respect to the constant water content  $\theta$  and the mean vertical velocity  $V_z$ . A more detailed discussion of the physical importance of these temporal moments is given in the analysis of the results.

### 9.2.3 Parameters and Model Design

The numerical simulations are implemented for a soil cross-section that is 12.8 m deep and between 7.6 m and 30 m wide depending on the expected horizontal solute spreading. The model simulates the instantaneous injection of a small source solute slug into the soil by specifying an initial concentration  $c_0 = 1$  for a horizontal rectangle of 3 by 2 nodes (concentration is specified as a nodal property in the MMOC algorithm). The total applied mass therefore depends on the chosen grid-discretization. Since all results are normalized with respect to the total mass or the initial concentration, the results are only dependent on the ratio between the initial plume size and the correlation scale of the soil texture.

For the transport simulations a subset of the different soil-types investigated in the previous chapter is selected. All soils investigated here have a hypothetical vertical correlation length  $\lambda_{fz}=50$  cm. The horizontal correlation length  $\lambda_{th}$  varies from 50 cm to 300 cm. The discretization of the domain is 10 cm in the vertical. The horizontal discretization is 10 cm for  $\lambda_{th}=50$  cm, 15 cm for  $\lambda_{th}=150$  cm, and 30 cm for  $\lambda_{th}=300$  cm. Thus the size of the initial solute slug relative to the correlation scale of the saturated hydraulic conductivity is 40% in the vertical; 60% in the horizontal for the isotropic soils, and 30% in the horizontal for the anisotropic soils. The total size of the finite element domain is 128 elements vertically and between 76 and 100 elements horizontally. The horizontal domain size was chosen such that the solute plume would not spread beyond the vertical boundaries of the finite element domain. An overview of the different hypothetical soils is given in Table 9.1 together with the actual size and discretization of the respective flow and transport models. The same soil mc3 is used as base soil site as in the flow analysis (isotropic, unit variance in  $f$ ). The variance of  $f$  varies from 0.01 to 4, the variance of  $a$  from 0.0001 to 0.04. The correlation  $\rho_{af}$  between  $f$  and  $a$  is either 1 or 0. The geometric mean of  $\alpha$  ( $=\Gamma$ ) is always  $0.01 \text{ cm}^{-1}$ . The average water tension varies from -150 cm to -3000 cm.

As for the flow analysis in chapter 8, the boundary conditions imposed on the solution of the steady state head field are of the random Dirichlet type. The random head boundaries are obtained from the first order spectral solution for the head given the particular realization of  $f$

and  $a$ . From the previous chapter it is known that the first order head boundaries significantly alter the velocity distribution in the vicinity of the boundary. These boundary effects would adversely affect transport across or near the boundary. To avoid the erroneous impact of the first order boundary, the solution of the flux is separated from the transport simulation. Steady-state flux is obtained for a finite element domain that is five to ten elements larger around each side than the finite element domain for the transport simulation. In other words, the boundaries of the transport model are located in the interior of the flow model. The size of the peripheral "cut-out" within the flow-model is determined from the results of the previous flow analysis (chapter 8, Figure 8.14). The number of rows or columns cut off of each side of the flow field is indicated in the right column of Table 9.1. For the transport model itself, no boundary conditions are necessary, since only advective transport is solved. The Monte Carlo simulations are based on 300 realizations of each soil site.

The previous chapter elucidated the dependence of the velocity field and its spatial variability upon the various parameters governing unsaturated flow. The spatial variability of the velocity field controls the uncertainty (or spatial variability) of solute flux and the solute concentration distribution through the physical dependence of concentration upon velocity expressed in the transport equation (eqn. 5-1). The velocity field is controlled by Darcy's law, and hence by the spatial variability of the unsaturated hydraulic conductivity  $y$ . The latter is the single most important parameter used to describe or predict the uncertainty in solute transport. Its spatial variability, as described in chapter 8, depends on the variance of the independent input parameters  $f$  and  $a$ , on their correlation scale and anisotropy ratio, and on the mean soil water tension  $H$ . For the evaluation of the transport simulation, the soils are therefore grouped into four categories, each of which addresses the sensitivity of stochastic transport to one specific parameter:

1. category: isotropic, wet soil: #2, #8, #3, and #9. Only the variance of  $f$  and  $a$  change. The actual average sample variance of  $f$  is 0.01, 0.11, 0.95, and 3.62, respectively (see chapter 8). The variance of  $a$  is  $10^{-4}$ ,  $10^{-2}$ ,  $10^{-2}$ , and  $4 \cdot 10^{-2}$ , respectively.

2. category: anisotropic, wet soil: #12, #29, #28, #22. Again, only the variance of  $f$  and  $a$  change. For #29, #28, and #22, the variances of  $f$  and  $a$  are 0.95, 2.15, 3.67, and  $10^{-2}$ ,  $4 \cdot 10^{-2}$ ,  $4 \cdot 10^{-2}$ , respectively. #12 and #29 have identical parameter variances, but unlike all other example soils #12 has perfectly correlated  $f$  and  $a$  random fields, which slightly reduces the unsaturated hydraulic conductivity variance relative to #29 (see chapter 8).
3. category:  $\sigma_f^2=1$ , wet soil: #3, #31, #29. Only the horizontal correlation scale of  $f$  (and  $a$ ) change: For the three soils they are 50 cm, 150 cm, and 300 cm which results in aspect ratios  $v = \lambda_{fh}/\lambda_{fz} = 1, 3, \text{ and } 6$ , respectively.
4. category: anisotropic dry soil: #15 and #21. These are both dry soils with mean soil water tensions  $H$  of -1000 cm and -3000 cm, respectively. The first soil is otherwise the same soil as #29, while the second soil is otherwise identical to #12, which has perfectly correlated  $f$  and  $a$  random fields. The two dry soils are contrasted with two wet soils of similar unsaturated hydraulic conductivity variance to investigate whether unsaturated transport can be characterized by  $\sigma_y^2$  and  $C_{yy}$  alone: #15 is contrasted with #28 ( $\sigma_y^2=1.5$  and 1.8, respectively), #21 is contrasted with #22 ( $\sigma_y^2=3.2$  for both).

### 9.3 Spatial Analysis of Solute Transport under Uncertainty

#### 9.3.1 General Characteristics of Solute Movement and of its Statistical Representation

The dynamics of the spatial distribution of solute concentration in unsaturated porous media and its direct dependence on the variability in velocity is illustrated in Figure 9.2 for three anisotropic soils with distinctly different  $\sigma_y^2$ . In mildly heterogeneous soil the streamlines are more or less vertical (Figure 9.2a). Deviation from a parallel vertical flow pattern becomes stronger as the variability of  $y$  increases (Figure 9.2b,c). In more heterogeneous soils streamlines tend to cluster in preferential flow areas, and show a large variability over short distances. If the soil is very heterogeneous, the streamlines may exhibit a horizontal

displacement of several meters (a few correlation scales, Figure 9.2c). Relative to its entry position, the maximum observed total horizontal displacement of a streamline at the bottom of the 12 m deep vertical section is 1 to 2 m for  $\sigma_y^2=0.74$  (soil #29), on the order of 5 m for  $\sigma_y^2=1.48$ , and on the order of 10 m for  $\sigma_y^2=3.20$ . While the single realizations of streamlines depicted in Figure 9.2 bear limited statistical significance, they clearly illustrate the degree of uncertainty associated with making predictions about the travel path of a small solute plume and help to understand the results of the stochastic transport analysis. These examples also illustrate that parallel streamtube models are limited in their applications when modeling solute transport from point sources, because they neglect the horizontal displacement of solutes.

The individual solute plumes (Figure 9.2d-f) show the combined effects of travel path variability, local dispersion and travel velocity variability. Since the plume center of mass is an integrated measure of the entire plume displacement, its travel path varies less than the corresponding streamlines. With increasing variance, the solute plumes assume increasingly erratic shapes, and for the same travel time, the travel distance of the plume center of mass becomes more variable. The peak concentration of the plume strongly depends on the total travel distance of the individual plume. If the solute is initially placed into an area of low unsaturated permeability, the plume will travel only slowly for an extended period of time and disperse very little i.e., the peak plume concentration will remain high. In contrast, a solute plume that is initially placed into a relatively high permeability area will travel quickly, disperse, and the peak concentrations will drop rapidly.

The mean concentration from the Monte Carlo simulation (Figure 9.2g-i) indicates the stochastically predicted location and spreading of the solute plumes at  $t'=10$  ( $t'=\lambda_{tz}/V_z$ ), the same time at which the snapshot of the actual plumes (Figure 9.2d-f) are taken. As expected, the mean plumes exhibit more spreading and lower peak concentrations than the actual plumes due to the variability  $\text{var}(M_i)$  ( $i=x,z$ ) in the travel path of the plume center (9-3). For strongly variable soil flux ( $\sigma_y^2>1$ ), Figure 9.2i qualitatively illustrates how  $\text{var}(M_i)$  contributes significantly to the total spreading  $X_{ii}$  of the mean concentration plume. In that case  $X_{ii}$  is a poor measure of the "typical"

solute distribution  $M_{ii}$  (Kitanidis, 1994).

Small irregularities can be seen in the contours due to the random nature of the sample mean (Figure 9.2g-i). The irregularities in the mean plume shape and the skewness of the mean concentration along the vertical axis are more notable at higher  $\sigma_y^2$ , since the potential error of the sample mean concentration increases with increasing velocity and concentration variance. For  $\sigma_y^2 > 1$ , the Monte Carlo simulations (which are based on 300 realizations) become potentially very sensitive to outliers. Those realizations with very small unsaturated hydraulic conductivities at and/or immediately around the source produce solute plumes which move extremely slow and therefore have peak concentrations close to the initial concentration throughout the entire simulation time (Figure 9.3). Even one realization with a very slow moving plume may significantly bias the arithmetically averaged concentration at later times, since otherwise the concentration at the source quickly drops by several orders of magnitude. In the example problems all simulations with  $\sigma_y^2 > 1.5$  show significant 'residual' mean concentrations at the source or within 1 correlation scale distance of the source. In this context, 'significant' is defined as any concentration larger than approximately 1/100th of the observed peak concentration. In these cases the concentration contours in the immediate vicinity of the source are also erratic and must therefore be attributed to outliers (see e.g. Figure 9.2h).

To investigate the accuracy of the sample mean concentrations calculated from 300 realizations, the rate of convergence in the Monte Carlo simulation is demonstrated for the most variable soil in this study (#21,  $\sigma_y^2=3.2$ ) i.e., the soil for which the sample error is the largest. Figure 9.3 shows the mean concentration at four different times computed from 300 realizations (Figure 9.3a-d) and from 100 realizations (Figure 9.3e-h). The contour map for 300 realizations is significantly smoother than the contour map for 100 realizations. Near the source and at the outermost contour level (<1% of peak mean concentration) the sample accuracy appears to be the least. Otherwise the dominant features of the mean plume are well-defined when based on 300 realizations. With only 100 realizations, the general pattern of the mean concentration distribution can already be seen, but there is also a large degree of erratic spatial variability. It

is obvious that hundreds (if not thousands) of additional realizations would be necessary to completely remove the secondary peak near the source caused by an outlier within the first 100 realizations. However, 300 realizations are considered adequate to give relatively accurate results for the spatial moments of the mean plume and the mean breakthrough curve and its variance, because they are integrated measures. They are also adequate to give an approximate description of the spatial mean concentration distribution (Figure 9.3a-d,i).

Therefore it appears that the skewness in the vertical distribution of the mean concentration, which increases with  $\sigma_y^2$ , cannot be explained as an outlier problem. Another explanation for the skewed mean concentration distribution must be sought. Recall that the longitudinal velocities have approximately a lognormal pdf (chapter 8). The theory by Neuman (1993) and the results of Zhang and Neuman (1994d) show that a lognormal velocity pdf indeed causes a skewed shape of the concentration plume at early travel time. Zhang and Neuman (1994d, Figures 1 and 2) indicate that the skewness of the mean concentration plume vanishes after only one correlation scale in travel distance. However, their examples are for mildly heterogeneous media, in which case the Monte Carlo simulations here also produce a Gaussian mean concentration distribution (Figure 9.2g,h). For highly heterogeneous soils ( $\sigma_y^2=3.2$ ), the skewness of the concentration plume along the longitudinal direction is very strong even at  $t'=10$ . Only at  $t'=20$  the skewness appears to converge towards a more Gaussian distribution (Figure 9.3c). For soil site #15 ( $\sigma_y^2=1.5$ ), the skewness is very strong at  $t'=5$  (not shown), but has almost disappeared at  $t'=10$  (Figure 9.2h).

These observations have a significant practical aspect. The relatively high 'residual' mean concentrations at  $t'=10$  for the large variance soils #9, #21, and #22 ( $\sigma_y^2>3$ ) (see e.g. Figure 9.3) suggest that several hundred correlation scales of travel distance may be required before a solute plume achieves ergodicity. The use of stochastic transport models that are based on the ergodicity assumption is therefore not justified for small plume sources and highly variable soils, since the unsaturated zone is either not that thick or it is unlikely to be characterized by a stochastically homogeneous random medium throughout its vertical extent.

The variability of the concentration from realization to realization is illustrated by the variance maps (Figure 9.2k-m). The concentration variance, however, is not a good measure of uncertainty since the mean concentration is nonstationary. The concentration variance is largest near the center of the mean concentration plume simply because the concentrations are largest at the center of the plume. Also note that the largest variances are found in the second of the three example soils, which exhibits a  $\sigma_y^2$  of less than half the largest  $\sigma_y^2$  in the three examples of Figure 9.2. Intuitively it would be expected that the uncertainty increases with  $\sigma_y^2$ . The discrepancy in the behavior of the peak variance is due to the fact that at higher  $\sigma_y^2$  the peak mean concentration decreases thus countering the overall increase in concentration variability. Another problematic aspect of the concentration variance is that it is even more sensitive to outliers than the concentration mean due to its second order nature (e.g. Figure 9.2l). A considerably better measure of uncertainty is the concentration coefficient of variation  $CV_c(\mathbf{x},t)$ :

$$CV_c = \frac{\sigma_c}{\langle c \rangle} \quad (9-6)$$

where  $\langle \rangle$  refers to sample averages and the evaluation is at  $\mathbf{x}$  and  $t$ . The  $CV_c$  measures the variability of the concentration at location  $\mathbf{x}$  and time  $t$  not in absolute terms but relative to the mean concentration at the same location and time. Unlike the concentration variance, the  $CV_c$  has a *minimum* at or near the location of the peak mean concentration and increases with distance from the plume center. This is consistent with the analysis of concentration variability in saturated transport (Dagan, 1984; Rubin, 1991b). In terms of  $CV_c$  the uncertainty about the high concentrations in the center of the plume is the lowest while uncertainty about the very low mean concentrations at the edges of the plume is highest. The added advantage of using  $CV_c$  instead of  $\sigma_c^2$  as a measure of uncertainty is its inherent ability to neutralize outliers of the type discussed above. Outliers not only increase the concentration mean but also the concentration variance near the plume source. The resulting  $CV_c$  remains very high (as expected) relative to the  $CV_c$  at the plume center. Overall the lower  $CV_c$  is much less sensitive

to sampling error than either the variance or the mean.

The general shape of the  $CV_c$  plume shows a similar sensitivity to soil heterogeneity and soil moisture as the mean concentration plume. At the plume center the  $CV_c$  increases with  $\sigma_y^2$  while it decreases at the periphery of the mean plume location (see chapter 10). With increasing aspect ratio vertical spreading decreases and horizontal spreading increases. The large  $CV_c$  ( $\gg 1$ ) in all areas except the plume center is yet another indication that the concentration RFV  $c(\mathbf{x},t)$  is not Gaussian distributed. Cushey et al. (1993) have also shown that the concentration in porous media has a non-Gaussian pdf. These findings must be taken into consideration when it comes to the practical meaning of both mean concentration and concentration coefficient of variation (or the concentration variance).

### 9.3.2 The Minimum $CV_c$ - an Empirical Stochastic Analysis

As expected, the minimum  $CV_c$  (used as an indicator of the overall uncertainty) increases with the textural variability of the soil expressed by  $\sigma_f^2$  (Figure 9.4a,b). Surprisingly, the minimum  $CV_c$  in the  $\sigma_f^2=0.01$  soil is still significant throughout the simulated time-span ( $\approx 0.2$ ). On the other end of the scale, the largest observed minimum  $CV_c$  are for anisotropic soils of high variability in  $y$  and do not exceed 1.6 (#21, Figure 9.4c). The location of the minimum  $CV_c$  coincides with the location of the maximum mean concentration only for the two plumes with  $\sigma_f^2 < 0.15$ . For all other soils, the minimum  $CV_c$  location is lower than the point of maximum mean concentration (Table 9.2). Generally the distance between the two is less than  $1\lambda_{fz}$ . But for  $t'=20$  the distance may be from  $2\lambda_{fz}$  to  $10\lambda_{fz}$ , depending on the soil variability. Non-Gaussian vertical velocity pdfs and higher variability of the soil flux causes stronger skewness of the vertical mean concentration plume and larger distances between the two locations of maximum concentration and minimum  $CV_c$ .

The results are consistent with the theoretical findings of Rubin (1991b), who estimated the concentration mean and concentration coefficient of variation in a mildly heterogeneous

saturated porous medium ( $\sigma_y^2 = 0.21$ ). Under the assumption that the velocity pdf is Gaussian, it was concluded that the highest concentration variance and lowest  $CV_c$  coincide with the center of the mean concentration plume. The results here illustrate that Rubin's conclusion does not extend to porous media of higher variability and lognormally distributed velocities, where the minimum  $CV_c$  location has travelled further than the peak mean concentration location. The results are in contrast with those of Zhang and Neuman (1994d, Figure 1). In their analysis of solute transport with lognormally distributed longitudinal velocity, they came to the same conclusion as Rubin (1991b) i.e., the lowest  $CV_c$  occurs at the location of the largest  $\langle c \rangle$ . Again it should be noted that their result is obtained for a mildly heterogeneous porous medium.

Unfortunately, no information on the  $CV_c$  is available for  $t' < 5$  or at any other times besides those at which the concentration distribution is recorded ( $t' = 5, 10, 20, 40$ ; in case #28:  $t' = 4, 8, 16, 31$ ;  $t' = \lambda_{tz}/V_z$ ). For illustration purposes the three data points of each soil site shown in Figure 9.4 are interpolated with a 3rd order polynomial on the interval  $t' = [4, 20]$ . The  $CV_c$  decreases with time or remains almost constant in the isotropic soils (category 1, Figure 9.4a), but initially increases and later remains constant or decreases again in the anisotropic soils (category 2, Figure 9.4b). The minimum  $CV_c$  at  $t' = 5$  is higher in the isotropic soils than in the anisotropic soils (category 3, Figure 9.4c). At later times ( $t' = 10, 20$ ) the observed difference in  $CV_c$  for different aspect ratios decreases, and in some instances the anisotropic soil will have a higher  $CV_c$  than the isotropic soil of comparable textural variability (compare  $\sigma_f^2 = 3.6$ , isotropic/anisotropic soils, Figure 9.4a,b). At the last output time ( $t' = 40$ ) all or most of the plume has passed through the domain, and the actual minimum  $CV_c$  of the plume cannot be observed. For soils of comparable  $\sigma_y^2$ , but different mean head (4th category), the temporal dynamic and magnitude of the minimum  $CV_c$  is very similar but not identical (Figure 9.4d).

In contrast to these findings, other analytical studies (Rubin, 1991b) and field analyses (Barry and Sposito, 1990), both for transport in the saturated zone, report a steady increase in the  $CV_c$  with time (maximum analyzed travel distance approximately  $40\lambda$ ). The argument was made that the deterministic effect of the source shape and location of the plume wears out over

time. Thus one would observe increasing prediction uncertainty. Observations here suggest that the minimum  $CV_c$  in the tested soils may reach a maximum around or before  $t'=20$ , then decrease at a slow rate. However, the record is too short to be conclusive. In addition, the time span of the initial increase in minimum  $CV_c$  seems to be strongly influenced by the aspect ratio of the anisotropy and by the variance of the unsaturated conductivity.

### 9.3.3 Spatial Spreading of the Mean Plume, Mean Spatial Spreading of Plumes, and Variability of the Plume Center of Mass

In this section the stochastic dependence of the three moments  $X_{jj}$ ,  $\text{var}(M_j)$ , and  $\langle M_{jj} \rangle$  ( $j=x,z$ ) on the heterogeneity and moisture content of the unsaturated zone is examined under fully nonlinear conditions (numerical Monte Carlo simulations). The results are used to assess the validity of what is called the linear macrodispersion theory.

Until recently, the thrust of stochastic analyses regarding transport in heterogeneous porous media has not so much dealt with the mean concentration and concentration variance itself, as it has focused on the second spatial moment or moment of inertia  $X_{jk}$  of the ensemble mean concentration plume (9-2). Under the assumption of normally distributed velocities, the spatial distribution of the mean concentration is Gaussian, and  $X_{jk}$  is then intrinsically related to the apparent Fickian macrodispersion  $D_{jk}$  through:

$$D_{jk}(t) = \frac{1}{2} \frac{dX_{jk}(t)}{dt} \quad (9-7)$$

With the aid of the time-dependent macrodispersion coefficient, the ensemble mean concentration can be estimated through a quasi-Fickian analysis i.e., by solving (5-1) or (9-1) with  $D(t) = D_{jk}(t) + D_d$  in lieu of  $D = D_d$ , where  $D_d$  is the local (small-scale) dispersion coefficient. In all linear theories of macrodispersion (Gelhar and Axness, 1983; Dagan, 1984, 1988; Neuman et al., 1987), the second moment (moment of inertia) of the mean concentration

plume is estimated analytically based on the following fundamental result of turbulent diffusion (Dagan, 1984, eqn. 3.20, 1988, eqn.3, A2; here formulated for two-dimensional, vertical, uniform flux):

$$X_{jj}(t) = \int_0^t \int_0^t C_{v_j, v_j}(V_z t', V_z t'') dt' dt'' \quad (9-8)$$

where  $j = x, z$  and  $V_z$  is the mean velocity in the vertical direction.  $C_{v_j, v_j}$  is the covariance of the velocities at two locations  $\mathbf{x}'$  and  $\mathbf{x}''$  corresponding to mean displacement at  $t'$  and  $t''$ , respectively.  $\mathbf{x}'$  and  $\mathbf{x}''$  are assumed to be located along the travel path of the mean solute plume. This latter assumption is the most important restriction to the analysis, since it neglects any deviation of the actual plume from the mean travel path. Neuman and Zhang (1990) and Neuman (1993) have overcome the assumption by using a quasi-linear analysis and a Eulerian-Lagrangian approach, respectively. Equation (9-8) also assumes both Eulerian and Lagrangian stationarity of the velocity field (Dagan, 1988).

The stationarity conditions are met in the numerical transport simulations under investigation. Also, in chapter 4 the second moment  $C_{v_j, v_j}$  of the velocity is derived to first order for soils of arbitrary covariance functions in  $f$  and  $a$ . The velocity covariance function is obtained from a spectral analysis and a numerical fast Fourier transform of the resulting spectral density function. The mean travel velocity is given to first order in (4-43, 8-16). The second spatial moments of the mean plume (9-8) can therefore be readily computed. The validity of the linear macrodispersion theory over a wide range of soil heterogeneity is analyzed by comparing the analytical with the numerical results.

Since all example soils of the Monte Carlo analysis assume the same correlation functions for the input RFVs  $f$  and  $a$  (identical correlation scales), the covariance functions  $C_{yy}$ ,  $C_{v_1, v_1}$ , and  $C_{v_2, v_2}$  and by virtue of (9-8) also the spatial moments  $X_{zz}(t)$  and  $X_{xx}(t)$  are directly proportional to the variance factor  $\sigma^2$  (see chapters 4 and 8):

$$\sigma^2 = \sigma_f^2 (1 + 2\rho H' + H'^2) \quad (9-9)$$

$H' = \zeta\Gamma H$  is a dimensionless form of the mean soil water tension. The spatial moments can therefore be normalized with respect to four of the input parameters: the correlation between  $a$  and  $f$ ,  $\rho_{af}$ , the ratio between the standard deviations of  $a$  and  $f$ ,  $\zeta$ , the mean head,  $H$ , and the input variance of  $f$ ,  $\sigma_f^2$ , and of  $a$ ,  $\sigma_a^2$ . The remaining input parameters are the geometric mean of  $\alpha$ ,  $\Gamma$ , and the vertical and horizontal correlation scales of  $f$ ,  $\lambda_{fz}$  and  $\lambda_{fx}$ .  $\Gamma$  is constant throughout all simulations. The dimensionless spatial moments  $X_{zz}' = X_{zz}/(\lambda_{fz}^2 \sigma^2)$  and  $X_{xx}' = X_{xx}/(\lambda_{fx}^2 \sigma^2)$  for the three different anisotropy ratios in Table 9.1 give the analytical solution (9-8) to all hypothetical soils investigated here (Figure 9.5a,b). The initial dimensions of the plume are negligible for all but the least variable soil ( $\sigma_f^2=0.01$ ) and are not taken into account in Figure 9.5.

As the aspect ratio  $v$  increases, the longitudinal, vertical spatial spreading  $X_{zz}'$  computed from (9-8) decreases considerably. This is consistent with the theoretical results for three-dimensional saturated porous media (Dagan, 1988). At early time, the transverse, horizontal spreading  $X_{xx}'$  is smaller in soils with larger  $v$ . But the less anisotropic the soil, the earlier it reaches its maximum transverse macrodispersivity, which is also higher. Hence, at late time  $X_{xx}'$  is larger for soils with higher transverse anisotropy. The theoretical impact of anisotropy on  $X_{zz}'$  and  $X_{xx}'$  is considerably stronger than in three dimensions (compare to Russo, 1993a, Fig. 7). From the corresponding curves for the apparent longitudinal and transverse macrodispersivity  $D_{jj}' = 0.5 X_{jj}' / t'$  (Figure 9.5c,d) it is obvious that the asymptotic apparent macrodispersivity will only be reached after 100 to several hundred correlation scales travel distance. In a three-dimensional medium, the asymptotic value is reached significantly earlier (Russo, 1993a, Figure 8). The right vertical axis in Figure 9.5c,d is scaled to give the actual macrodispersivity  $\alpha_{jj}$  for soils with  $\sigma_f^2=1$ ,  $\lambda_{fz}=50$  cm,  $H=150$ cm,  $\Gamma=0.01$  1/cm,  $\zeta=0.1$ , and  $\rho_{af}=0$  (soils #3, #31, #29). The asymptotic longitudinal macrodispersivity for these soils is 51 cm, and the maximum transverse dispersivity varies from 3 cm in the isotropic soil to 1.7 cm in

the most anisotropic of the three soils. None of these theoretical results considers local dispersion i.e.,  $D_d = 0$ .

From the linear theory, the spatial moments of the mean plume are expected to increase as the soil dries out, if  $\rho_{af}=0$  (9-9). For correlated soils ( $\rho_{af}=1$ ) the variance of  $y$  and hence the solute spreading decreases as the soil begins to dry out, reaches a zero variance at  $H' = 1$  (9-9), and then increases for very dry soils (Figure 9.5e). Note that the variance of the correlated case reaches 0 only if the correlation functions for  $a$  and  $f$  are identical. The effect of decreasing soil moisture and increasing soil water tension becomes significant for  $H' > 0.01$ . At high soil water tensions  $H' > 100$  the correlation between  $f$  and  $a$  plays an insignificant role and the scaling factor  $\sigma^2$  grows with  $(H'^2 \sigma_f^2)$ .

Figure 9.6 shows the results for the spatial spreading  $X_{zz}'$  and  $X_{xx}'$  of the mean concentration plume in the numerical simulations. The results are shown separately for the four different simulation categories. For the isotropic soils, the fully non-linear solution deviates strongly from the linearized analytical solution for all  $\sigma_f^2 \ll 1$  due to the influence of local dispersion (whose dimensionless magnitude increases with  $1/\sigma_f^2$  as  $\sigma_f^2 \rightarrow 0$ ). The smaller the variability of the soil, the larger are both the longitudinal and transverse spreading relative to the theoretical result. For  $\sigma_f^2=0.95$ , the actual longitudinal spatial spreading of the mean concentration is very close to, but slightly smaller than the analytical solution. At higher soil variability the numerical results again increase relative to theory. However, for  $\sigma_f^2=3.6$  the mean concentration plume reaches the outflow boundary very early resulting in an artificial deviation from the expected behavior in both the transverse and longitudinal spreading. For the anisotropic soils (second category), the lowest variance is  $\sigma_f^2=0.95$  and as for the isotropic soils, an increase is observed in both the horizontal and vertical spreading with increasing soil variability. In contrast to the isotropic soils, the dimensionless vertical (longitudinal) spreading is much larger relative to the linear macrodispersion theory, even for  $\sigma_f^2=0.95$ , since the decrease in longitudinal spatial spreading with increasing  $v$  is not as strong for the numerical results as for the analytical solutions (category 3). Similarly drier soils with larger  $(1 + 2\rho\zeta\Gamma H + (\zeta\Gamma H)^2)$

exhibit a stronger deviation from the linear macrodispersion theory (category 4).

In general, the numerically obtained transverse  $X_{xx}'$  curves follow the same ups and downs with varying  $\sigma^2$  as the longitudinal  $X_{zz}'$ . But the actual deviation from the linear theory is much larger in the transverse direction than in the longitudinal direction. The transverse spreading for  $\sigma_r^2=0.95$  in an isotropic soil is approximately twice as large as the analytical prediction. This confirms similar results by Rubin (1991b) in a stochastic analysis of a groundwater transport site with  $\sigma_y^2=0.29$ . Based on a nonlinear stochastic transport model, Rubin found that the transverse spreading is approximately twice the value predicted by linear macrodispersion theory (Dagan, 1984, 1988). For the anisotropic soils, the transverse, horizontal spreading of the mean concentration plume from the Monte Carlo analysis is several times larger than predicted from (9-8). The underestimation of the horizontal mean plume spreading by (9-8) is significantly greater for the anisotropic soils than for the isotropic soils (category 3).

The variance  $\text{var}'(M_z)$  of the plume center of mass in the vertical direction is also larger than the analytical results for  $X'_{zz}$  with the exception of isotropic soils. It exhibits a very similar sensitivity though not as pronounced, to changes in soil variability, soil moisture, and anisotropy as the vertical spreading of the mean concentration plume (Figure 9.7a-d); a very small decrease from the isotropic soil  $\sigma_r^2=0.01$  to the soils  $\sigma_r^2=0.95$ , and an increase in  $\text{var}'(M_z)$  for both anisotropic and isotropic soils with  $\sigma_r^2 > 1$ . In the wet, anisotropic soils with  $v=6$ ,  $\text{var}(M_z)$  is almost identical with the theoretical curve for  $X'_{zz}$ . As  $v$  decreases, the increase in  $\text{var}'(M_z)$  is not as strong as in the theoretical  $X'_{zz}$ . On the other hand, a significant increase is observed in  $\text{var}'(M_z)$  for the two dry soils relative to wet soils of similar unsaturated conductivity variance. In the horizontal direction, the variance  $\text{var}'(M_x)$  of the plume center of mass is almost identical with the theoretical  $X'_{xx}$  for the isotropic soils, but several times larger than the theoretical  $X'_{xx}$  for all anisotropic soils. The disagreement between the two results increases significantly for higher  $\sigma_r^2$  and for dry soils (Figure 9.7e-h).

The numerical results for the spatial moments of the mean concentration plume cannot

be strictly compared with the theoretical results for the ensemble mean concentration, because the theoretical results do not account for the local dispersion  $D_d$ . Unfortunately, no definite number can be associated with  $D_{num,jk} = D_{d,jk}$ , but initial empirical results have shown that  $\alpha_{num,jj} = D_{num,jj}/V_z$  is on the order of 1/10th or less of the element size in the j-th direction (see section 9.2.1). Since  $D_{num,jk}$  is constant with time, a term  $X_{num,jk} = 2tD_{num,jk}$  must be subtracted from the spatial moments of the mean concentration to obtain the actual  $X_{jk}$ . Note that the dimensionless  $X'_{num,jk} = (2\alpha_{num,jk} t')/(\lambda_{fz}\sigma_f^2)$ . For small  $\sigma_f^2$  the numerical dispersion therefore has a considerable impact on the spatial moments of the mean concentration. The initial findings on numerical dispersion are re-evaluated by assessing the observed mean spreading  $\langle M_{jj} \rangle$  around the center of mass of each plume. Taking into account the initial size of the plume  $M_{jj}(0)$  the mean apparent dispersivity of plumes,  $\alpha_{Mjj}$ , is:

$$\alpha_{M_{ij}} = \frac{(\langle M_{ij}' \rangle - M_{ij}'(0)) \lambda_{fz} \sigma^2}{2t'} \quad (9-10)$$

$(\langle M_{ij}' \rangle - M_{ij}'(0))/(2t')$  is plotted in Figure 9.8. For  $\sigma_f^2=0.01$  the effect of numerical dispersion is largest relative to the effect of spatial variability. There, the mean apparent dispersivity is approximately constant with time and is 1.08 cm in the vertical and 0.12 cm in the horizontal. Since the element size is square with side-length 10 cm, the initial assumption that local dispersivity is on the order of 1 cm or less is confirmed. In other soils,  $\alpha_{M_{ij}}$  is significantly larger and is attributed mainly to the effect of spatially variable velocity fields.

## 9.4 Temporal Analysis of Solute Transport under Uncertainty

### 9.4.1 Integrated Breakthrough at a Compliance Surface

For many regulatory purposes, the distribution of a contaminant in the unsaturated zone is not as much of interest as the breakthrough of the contaminant or solute at the groundwater table, which is at some distance  $Z$  from the source area. Cvetkovic et al. (1992) introduced a theoretical multidimensional model based on the Lagrangian analysis of solute transport (Dagan et al., 1992) to predict the aerially integrated averaged breakthrough and its variance at an arbitrary compliance surface located normal to the mean flow direction at distance  $Z$  from a solute source of initial volume  $V_0$  [a similar approach based on the theory by Neuman (1993) was used by Zhang and Neuman (1994c)]. Since the model requires stationarity in the velocity moments and a mean uniform flow field, it applies directly to the unsaturated transport experiments performed in this study. The simulations presented here are for two dimensions. Thus, the initial source is an area  $A_0$  and the compliance surface CP reduces to a linear compliance level CL. The theoretical mean total solute flux  $\langle S(t,Z) \rangle$  and the variance of the total solute flux  $\sigma_s^2(t,Z)$  across the CL at distance  $Z$ , at time  $t$  are obtained from the following fundamental relationships (Dagan et al., 1992):

$$\begin{aligned} \langle S(t, \mathbf{Z}) \rangle &= c_0 \int_{A_0} g_1(t, Z - a, \mathbf{a}) d\mathbf{a} \\ \sigma_s^2(t, \mathbf{Z}) &= c_0 \int_{A_0} \int g_2(t, Z - a', \mathbf{a}' - \mathbf{a}'') - g_1(t, Z - a', \mathbf{a}') g_1(t, Z - a'', \mathbf{a}'') d\mathbf{a}' d\mathbf{a}'' \end{aligned} \tag{9-11}$$

where  $c_0$  is the initial, uniform solute concentration in the source area  $A_0$ , and  $a$  is the vertical coordinate of the location vector  $\mathbf{a}$ .  $g_1(t, Z - a, \mathbf{a})$  is the travel time probability density function (pdf) of a particle originating at  $\mathbf{a}$  in  $A_0$  and passing through the CL at time  $t$ . Correspondingly, the joint two-particle travel time pdf  $g_2(t, Z - a', \mathbf{a}' - \mathbf{a}'' )$  is the probability that two particles originating at  $\mathbf{a}'$  and  $\mathbf{a}''$  in  $A_0$  traverse the CL at time  $t$ . The travel time pdfs depend on the statistics of the velocity. Cvetkovic et al. (1992) derive a first order approximation of the first two moments of the joint two-particle travel time pdf under the assumption that the travel path of a particle does not deviate significantly from the mean flow direction:

$$\begin{aligned} \langle \tau \rangle (Z) &= \frac{Z}{V_z} \\ \sigma_{\tau, \tau} (Z, \mathbf{a}', \mathbf{a}'') &= \frac{1}{V_z^4} \int_{a'}^Z \int_{a''}^Z C_{v_z, v_z} (z' - z'', x' - x'') dz' dz'' \end{aligned} \tag{9-12}$$

where  $(a', x')$  and  $(a'', x'')$  are the vertical and horizontal coordinates of  $\mathbf{a}'$  and  $\mathbf{a}''$ . Using these travel time moments, Cvetkovic et al. (1992) then hypothesize that  $g_1$  and  $g_2$  are lognormal pdfs, an approximation that is strictly valid only for lognormal velocity pdfs as  $Z \rightarrow 0$ . They point out, however, that lognormal pdfs for  $g_1$  and  $g_2$  are consistent with the first order approximation (9-12) at all distances from the source, if the velocity field is lognormally distributed, an assumption that holds for the numerical simulations (chapter 8). The form of the lognormal pdfs  $g_1$  and  $g_2$  can be found e.g., in Bras and Rodriguez-Iturbe (1985). Using (9-12) with the linearized velocity covariance function  $C_{v_z, v_z}$  derived in chapter 4, the double integrals in (9-11) can be obtained by numerical evaluation. Since  $C_{v_z, v_z}$  varies smoothly and is numerically evaluated with a discretization  $\Delta x = 0.1 \lambda_r$  (chapters 4, 8), a very accurate estimate

of the integrals (9-12) and (9-11) is obtained by applying an extended form of Simpson's rule first to (9-12) and subsequently to (9-11) using the same discretization as in  $C_{vz,vz}$ . In the evaluation, the vertical extent of the source area is taken into full consideration. A conceptually similar approach is taken by Russo (1993b) but based on a different derivation of  $C_{vz,vz}$  and different integration procedures. From (9-12a) and (9-11a) he evaluated the mean (but not the variance) of the unsaturated breakthrough curve from a source of negligible longitudinal extension for a three-dimensional soil with normally distributed pore-size parameter  $\alpha$ .

As in section 9.3, the travel time variance and covariance (9-12) can be obtained normalized with respect to  $\lambda_{fz}$  and the scaled input variance  $\sigma^2$ . But due to the nonlinear form of the lognormal pdfs  $g_1$  and  $g_2$ , the mean and variance of the solute flux (9-11) would not be associated with the same normalization. Since solute flux has dimensions of mass per time, the dimensionless solute flux  $S'$  is defined as:

$$S' = \frac{S\lambda_{fz}}{M_0 V_z \theta} \quad (9-13)$$

For non-reactive, mass-conservative solutes, the area under the breakthrough curve  $S'(t')$  must be unity. The numerical mean breakthrough curves (mean BTCs) can be integrated up to  $t'=40$ . Only at  $Z'=5.4$  breakthrough is completed in all but the most heterogeneous soil before  $t'=40$ . Mass balance errors in the Monte Carlo mean BTCs range from +1% to +5%. The mass-balance in the numerically evaluated analytical BTC (9-11) is accurate to within less than 0.1%. For  $Z'=11.6$ , the analytical ensemble solute flux moments (9-11) and the numerical sample solute flux moments are plotted in Figure 9.9 (as solid and dashed lines, respectively).

*Numerical simulation results.* For both isotropic and anisotropic soils, the mean BTC increases with soil heterogeneity (Figure 9.9a,b). The increased spreading is associated with lower and much earlier peak solute flux. In the isotropic soils, peak solute flux decreases from 0.25 to 0.06 as the variance of  $f$  increases from 0.11 to 3.62. In the anisotropic soils of equivalent heterogeneity solute breakthrough is less disperse with higher peak concentrations.

At  $\sigma_f^2=0.95$ , peak solute flux is 0.10 in the isotropic wet soil and 0.15 in the anisotropic ( $v=6$ ) wet soil. At  $\sigma_f^2=3.6$ , peak solute flux has reduced to 0.06 and 0.08, respectively. Very similar breakthrough is observed for the two pairs of wet and dry soils with similar  $\sigma_y^2$  (Figure 9.9d). The dry soil #15 has a slightly more peaked breakthrough than #28, owing perhaps to its 17% smaller  $\sigma_y^2$ . In contrast, dry soil #21 shows a more smearing than wet soil #22, even though their  $\sigma_y^2$  are identical. Some of the difference may be attributed to the 7% mass balance difference between #21 and #22 (#22 is the only soil with a relatively high mass balance error: +12%).

The skewness and spread of the mean BTC is reflected in the skewness and spread of the standard deviation of the solute flux (Figure 9.9e-h). The peak standard deviation does not decrease as rapidly with increasing  $\sigma_f^2$  as the mean BTC, suggesting an increase in the solute flux coefficient of variation as the soils become more heterogeneous. The shape of the standard deviation curve (STDC) is almost bimodal with a small secondary peak after the maximum standard deviation occurred. The bimodality is more pronounced in the isotropic soils than in the anisotropic soils. Also in the isotropic soils, both peaks in the STDC occur before the mean BTC reaches its maximum (except in the soil with  $\sigma_f^2=0.11$ ). In the anisotropic soils, the peak of the mean BTC falls between the two maxima of the STDC (Figure 9.9f). As for the mean BTC, little differences are observed in the STDC between the soils of the two wet/dry soil pairs (Figure 9.9h).

*Comparison to analytical results.* Surprisingly, perhaps, there is a good visual agreement between the sensitivity of the theoretically determined and the numerically obtained BTCs and STDCs to the variations of the soil hydraulic conductivity variance, change in soil water tension, and change in anisotropy (Figure 9.9). All theoretical BTCs are less dispersed than the numerical results with a peak solute flux that is generally between 15% and 30% larger. Larger differences occur in soils with more heterogeneous flow patterns. Only in the driest soil (#21) the difference exceeds 30% (approx. 70% more peak solute flux). The theoretical standard deviation is approximately twice as large as the numerical standard

deviation with a tendency to a greater difference at low  $\sigma_f^2$  and low  $v$ . Both numerical and analytical STDC are much broader than the respective mean BTCs suggesting a very high uncertainty about the prediction of the frontal breakthrough and tailing of the mean plume. The bimodality of the STDC is not observed in the analytical results. The general observations made for the BTCs and STDCs at  $Z'=11.6$  (Figure 9.9) are also made at other depths of the CL (an example is shown in Figure 9.10)

The smaller spreading of the mean theoretical BTC curve (when compared to the simulation results) is consistent with the observed difference between the longitudinal spatial moments from the analytical stochastic macrodispersion analysis and the numerically determined spreading of the mean concentration plume, the latter of which is generally found larger. As mentioned above the differences can be attributed to the first order approximations assumed in the derivation of (9-11). Note that the illustration of the numerical and analytical results of the least variable soil (wet, isotropic,  $\sigma_f^2=0.01$ ) are omitted. For that soil site the analytical results show a very sharp peak, which is approximately twice as large as the numerical result. However, in the latter local dispersion plays a considerable role in the overall plume spreading since the spatial variability of the velocity field is almost negligible.

#### **9.4.2 Local Compliance with Maximum Contamination Flux Levels**

This section is concerned with the statistics of local solute breakthrough in contrast to the integrated breakthrough investigated in the previous section. The arrival time  $t_a$  of the compliance solute flux  $s_0$  and the solute flux peak passing time  $t_p$  are used as parameters to describe the frontal part of the breakthrough curves at each location  $x$  of the CL (recall that  $x$  is defined for a line only). In fact,  $t_p$  is only a particular aspect of the  $t_a$  function. The choice of these two parameters, instead of the entire local BTCs, has several advantages. The computational effort of storing and evaluating these data in a Monte Carlo simulation is orders of magnitudes smaller than the statistical evaluation of the local BTCs( $x$ ). In the numerical

simulation,  $t_p$  is defined simply as a function of location and is therefore a very small array containing the same number of elements as a single row in the simulation domain. The arrival time  $t_a$  is also defined as a function of horizontal location on the CL, but it is parametrized with respect to the compliance solute flux.  $t_a$  is compiled for 19 levels of dimensionless solute flux  $s/s_0$ , varying logarithmically from  $10^{-9}$  to 1. Hence,  $t_p$  is essentially a function defined for the two-dimensional plane  $(x, s/s_0)$ . Computationally, these are 19 times the amount of data stored for  $t_p$ . In contrast, the information of a solute BTC at  $(x, t)$ , recorded for every time-step, takes as many data as there are time-steps for each grid-node on the simulated CL. These time-steps vary from realization to realization depending on the maximum Courant number (automatic time-stepping, chapter 5). For the assembly of the sum and sum of square arrays of these data variables (see above), the individual BTCs have to be interpolated onto a regular grid of the time-axis.

To illustrate the conceptual link between the arrival time function  $t_a(x, s/s_0)$ , the peak time function  $t_p(x)$ , and the BTC function  $s/s_0(x, t)$ , recall that all are functions in the same three-dimensional space defined by the real variables  $s/s_0$ ,  $x$ , and  $t$ . The BTC function  $s/s_0(x, t)$  has a unique solution for each  $(x, t)$  and can be visualized as a longitudinal (parallel to the  $t$  coordinate) trending "mountain" bulging above the  $(x, t)$  plane. The BTC function can be uniquely mapped by projecting the mountain onto a contour map in the  $(x, t)$  plane. In contrast, neither the function  $t(x, s/s_0)$  nor the function  $x(t, s/s_0)$  have unique solutions. For each  $x$ ,  $s/s_0$  is obtained twice (in the front and in the tail of the BTC(x) i.e. one can look at the mountain from the front or the back). And at each time,  $s/s_0$  is obtained twice along the CL (to the left and to the right of the center of the breakthrough i.e. one can look at the mountain from the left or the right side). But if the three-dimensional cube spanned by  $\mathbb{R}^3(s/s_0, x, t)$  is sliced such that  $t(x, s/s_0)$  has only a single solution (in other words by deciding that one is only interested in the shape of the mountain as it is seen either from the front or the back), either the frontal part of the solute flux "mountain" or the tail part of the solute flux "mountain" can be projected onto a contour map in the  $(x, s/s_0)$  plane of the cube. The  $t_a(x, s/s_0)$  function is the frontal projection

of the solute flux "mountain". And  $t_p(x)$  is the contour in the  $(x,t)$  plane along which the cut was made.

In the statistical analysis, there is a fundamental difference, however, between the mean BTC( $x,t$ ) function and the mean  $t_a(x,s/s_0)$  function and their associated variances. The former will average the solute flux at a particular time and location, while the latter will average the time associated with a particular location and solute flux. From the analysis of the spatial mean concentration it has become obvious that the concentration not only has a truncated, but also highly skewed pdf as evidenced by the outlier problems. The meaning of averages and variances of non-Gaussian pdfs is questionable. In contrast, the pdf of the arrival time is not truncated (a zero arrival time is impossible) but it has the difficult property that arrival time is not necessarily defined for every location  $x$  and every solute flux  $s/s_0$ . At distance  $x$  certain solute fluxes  $s/s_0$  (and any higher flux) never occur. From realization to realization, arrival time may or may not be defined for the location  $x$  and level  $s/s_0$ . In statistical terms, the empty set  $\{\emptyset\}$  must be assigned a certain probability. While this poses difficulties in approaching the problem theoretically, an approximate solution for the numerical sampling in a finite sampling space (number of realizations) can be constructed by simply ignoring those samples at  $(x,s/s_0)$  that are not defined. In the Monte Carlo simulation, a zero is added to the sum and sum of square arrays, if no solute flux ever occurred at a certain level and location during a particular transient transport realization. In addition a counting file is kept to count the number of occurrences of  $t_a(x,s/s_0)$ , which is equal to or smaller than the total number of realizations. The resulting sums and sums of squares are eventually divided not by the number of realizations but by the number of occurrences of  $t_a(x,s/s_0)$ .

Thus not only an average time and time variance is obtained, but also a histogram of the probability that  $s/s_0$  is exceeded at location  $x$ , which is in essence an inverse cumulative distribution function ("iCDF" = 1-CDF) for  $s/s_0$ . Obtaining a iCDF for solute flux is extremely important for many regulatory purposes, which deal with the likelihood that a certain contamination level is ever exceeded after the installation of a potential contamination source

or after completion of site remediation. In addition, the physical meaning of "expected time of solute flux exceedance" and the variability of the arrival time is in many instances more significant than the "expected solute flux at time  $t$ " and its potential variability. The probability of arrival time of a certain solute flux level is similar to assessing the probability that a solute particle originating from a contamination source will arrive within a certain time-frame. Many regulatory statutes require that a certain contamination level may not be exceeded at any given time after installment of a potential pollution source, or that certain contamination levels are unlikely to occur for  $X$  number of years after installment of the potential pollution source (Neuman, 1991). Such measures can be formulated as conditional probability measures of time of exceedance given a solute flux ( $\text{pdf}(t_a)|s/s_0$ ) or as conditional probability measures of solute flux levels given a time ( $\text{pdf}(s/s_0)|t_a$ ).

The evaluation of the local solute breakthrough is essentially complimentary to the spatial analysis of the concentration mean and variance. The results are expected to be analogous to the findings of the previous sections. The numerical evaluation of the arrival time of peak solute flux is shown in Figure 9.11 for the first, third, and last soil categories. All times are normalized with respect to the mean residence time  $Z/V_z$ . The average peak arrival time occurs earliest for the least heterogeneous soils (isotropic, wet,  $\sigma_f^2=0.01, 0.11$ ), since the BTC in these soils are least skewed (Figure 9.11a, see also Figure 9.9a). Not much difference in  $\langle t_p \rangle$  is seen between these two soils. Notice that the peak arrival time for the highly heterogeneous soil (isotropic, wet,  $\sigma_f^2=3.62$ ) is significantly later than for the other three soils. The peak solute flux occurs later at location on the CL further away from its center (where the center  $x=0$  is defined as being located vertically underneath the center of the plume source). The shape of the  $t_p$  contour is not very sensitive to soil heterogeneity indicating that the delay in peak arrival time at a location  $x$  (with respect to the  $\langle t_p \rangle(x=0)$ ) is the same no matter how heterogeneous the soil. With respect to the shape of the  $\langle t_p \rangle$  line, the anisotropy ratio of the soil is much more significant (Figure 9.11b). At higher anisotropy, the peak time on the sides of the CL will occur earlier than in isotropic soils due to the stronger horizontal spreading of

the solute plume. Again the differences between dry and wet soils are insignificant if the flux field has a similar variability (Figure 9.11c). For all soils the variance of the peak arrival time does not vary with horizontal location, even though the mean arrival times are larger at larger  $x$ . The variance increases mainly with  $\sigma_f^2$  and decreases only slightly with higher anisotropy (Figures 9.11d-f).

The mean arrival  $\langle t_a \rangle$  map, a front view or projection into the  $(x, s/s_0)$  plane of the solute flux "mountain" generally has the expected behavior that a given solute flux level is exceeded earlier towards the center of the CL and later as  $|x|$  increases. The higher the solute flux level, the later it will on average be exceeded (Figure 9.12). Note that the solute flux is plotted on a logarithmic scale. Except for the very highest solute flux levels (the only ones distinguishable in the BTCs of Figure 9.9, where solute flux is plotted on an arithmetic scale), the mean time of first exceedance of a solute flux level is significantly earlier than 1. The mean arrival time (relative to the average residence time  $Z/V_z$ ) for the lowest solute flux level ( $s/s_0=10^{-9}$ ) is between 0.45 and 0.55 for all investigated soils. Only in isotropic soils with  $\sigma_f^2 < 1.0$  the mean arrival time for any solute flux level never drops below 0.5. The mean arrival times increase in an almost linear fashion radially away from the  $(x=0, s/s_0=10^{-9})$  point. For higher variances and larger anisotropies the shape of the "mountain" becomes broader (but not flatter owing to the log scale on the vertical axis). With increasing anisotropy and increasing  $\sigma_f^2$ , the earliest arrival time for the low solute flux levels decreases everywhere on the CL, although  $\langle t_a \rangle$  shows very little sensitivity in the anisotropic soils at  $x=0$ , once  $\sigma_f^2 \geq 0.95$ .

The highest solute flux levels are reached in the least variable soils, while the maps for the most highly variable soils (Figure 9.12e,f) prove that outliers are observed not only on the slow travel time end (see section 9.3), but also on the fast travel time end of time scales. In a reversion of the general trend that higher solute fluxes occur later, the highest solute fluxes in the three soils with  $\sigma_f^2 > 0.9$  occur on average earlier than some lower flux levels. Since these high flux levels are associated with low likelihoods (see below), they can only be attained if the solute source is in a preferential flow area, which displaces the plume relatively fast past the CL.

Due to the varying number of samples underlying the sample mean and sample variance, the estimation errors of the sample mean and sample variance of  $t_a$  may vary significantly and are expected to be high near the margin of the map, where the likelihood of exceedance is least. The more variable the soil, the more erratic are the contour lines of both the  $\langle t_a \rangle$  and the  $CV_{ta}$  maps (Figure 9.12 and 9.13).

The coefficient of variation  $CV_{ta}$  of solute flux exceedance time shows that the highest uncertainties are about the arrival times of the lowest solute flux levels at the center of the CL (Figure 9.13). The only exception are the two, mildly heterogeneous, isotropic soils (Figure 9.13a-b), where the  $CV_{ta}$  increases first and then decreases again, with larger distance from the center of the CL. Clearly, the  $CV_{ta}$  rises with soil variability. For the isotropic, wet soil with  $\sigma_f^2=0.11$ , the  $CV_{ta}$  varies from 0.3 to 0.5 (Figure 9.13a). In the anisotropic soils with  $\sigma_y^2=3.2$ , the  $CV_{ta}$  varies from 0.5 near the edges to 1.1 in the center of the CL (Figures 9.13e-f). Anisotropy decreases the uncertainty about the mean arrival time considerably for soils of comparable variability in  $\sigma_f^2$ . The observed decrease of uncertainty towards the edges of the "mountain" must be seen in connection with the decreasing likelihood of such solute fluxes ever to be exceeded. Figure 9.14 gives the sample probability that a solute flux level is ever exceeded. Each curve presents the iCDF of solute flux at a particular location  $x$  of the left half of the CL. Like the mean  $\langle t_a \rangle$ , the iCDF is symmetric to the center of the CL. The rightmost curves represent the highest likelihood for the highest solute flux levels and are associated with  $x=0$ . Each curve further to the left in the graphs of Figure 9.14 is for an increasing distance  $x$  from the center. The interval between locations  $x$  of neighboring iCDF plots is indicated as  $\Delta x$ . The iCDF is steepest for the soils of lowest variability indicating the least variance in solute flux. At each soil, the steepest iCDF is observed for the center location, which has the highest mean (rightmost curve) and the least variance. The further away from the center of the CL, the lower the mean solute flux and the smaller the slope of the iCDF, hence the larger the variance. Note that these plots are on a logarithmic scale for  $s/s_0$ . For a relatively homogeneous soil (Figure 9.14a), the iCDF is relatively symmetric with respect to the geometric mean  $s/s_0$ , and

solute flux is lognormally distributed. At higher variances the top part of the slope of the iCDF decreases with a longer tail towards lower solute flux levels, indicating that on a logarithmic scale for  $s/s_0$ , the probability distribution of  $s/s_0$  (derivative of the iCDF) is significantly skewed. Skewness also increases with distance from the center.

## 9.5 Summary and Conclusions

In this chapter solute transport from local sources of small lateral extent in unsaturated soils with steady-state mean uniform flow conditions is analyzed. A number of numerical Monte Carlo transport simulations were implemented for a variety of soil conditions. The sensitivity of solute transport and the uncertainty of its prediction is investigated for soils of varying variability in  $\sigma_f^2$  and  $\sigma_a^2$ , for soils with anisotropy ratios ranging from  $\nu=1$  to  $\nu=6$ , for soils of different mean soil water tension  $H$ , and for soils with correlated and uncorrelated  $f$  and  $a$  parameters. Several different aspects of solute transport have been addressed: The dynamics of the mean concentration plume and of the concentration variance, the spatial inertia moments of the mean concentration plume, the mean inertia moments of individual plumes around their center of mass, and the variability of the center of mass of individual plumes. The spatial description of solute transport was contrasted with the temporal description of solute transport i.e., the solute breakthrough at a compliance surface or compliance level (CL) some distance away from the solute source. The mean and variance of the total mass flux across the CL is determined and an alternative interpretation of uncertainty with regard to the arrival time of a solute at a compliance surface is offered. Instead of predicting the uncertainty about the solute flux level at a given time, the uncertainty about the time of first exceedance of a given solute flux level is stochastically determined.

In agreement with theoretical predictions by Dagan (1986), the prediction of a solute plume of small initial extension (relative to the correlation scale of soil heterogeneity) is associated with large uncertainties. For highly heterogeneous soils ( $\sigma_f^2 > 2$ ) the mean

concentration plume is of dimensions that have little to do with the actual size of the plume. For soils with an unsaturated hydraulic conductivity variance  $\sigma_y^2 = 3.2$ , the variability in the vertical and horizontal displacement of the plume center at  $t'=10$  accounts for 71% and 55%, respectively, of the spreading observed in the mean concentration. As a consequence, the predicted average peak concentration is almost an order of magnitude lower. At  $\sigma_f^2=1$  and  $t'=10$ , the vertical and horizontal variability of the plume center displacement still accounts for 61% to 68% of the longitudinal and for 47% to 42% of the lateral mean concentration spreading, with the former limit being for isotropic soils and the latter for anisotropic soils of aspect ratio 6. Large residual mean concentration at less than 20 correlation scales from the source for  $t'=40$  is an important indication that the variability of the displacement of the plume center accounts for significant amounts of the prediction uncertainty even at travel distances larger than 40 correlation scales. This conclusion is confirmed by the significant variability of the solute mass flux across the bottom boundary (23 correlation scales from the source) at  $t'=40$ .

The numerical results are contrasted with linear theoretical models of stochastic transport in porous media. To describe the spatial moments of the solute plume, a linear macrodispersion model is developed for the two-dimensional unsaturated transport under investigation based on the analytical flow model developed in chapter 4 (see also Yeh et al., 1985a,b) and on the linear macrodispersion theory by Dagan (1984, 1988). For the solute breakthrough at a compliance surface the first order perturbation analysis of unsaturated flow is applied to the travel time analysis developed by Dagan et al. (1992) and Cvetkovic et al. (1992). Analytical predictions of the mean solute breakthrough curve and the variance of total solute flux at the unsaturated CL are obtained as a function of time.

The linear macrodispersion theory predicts that for soils with identical  $\lambda_r=\lambda_a$  and  $\Gamma$ , the second spatial moments are directly proportional to the theoretical unsaturated hydraulic conductivity variance  $\sigma_y^2$ , since the second spatial moment and  $\sigma_y^2$  both grow with  $\sigma = \sigma_f^2 (1+2\rho H' + H'^2)$ , where  $H' = \zeta\Gamma H$  is a dimensionless measure of the mean soil water tension.

From these theoretical considerations it is expected that a direct proportionality between the spatial moments of the concentration and the unsaturated hydraulic conductivity does not exist, when either the aspect ratio of the correlation scales or the geometric mean of  $\alpha$  ( $=\Gamma$ ) changes. It does also not hold, if the correlation functions for  $f$  and  $a$  are not identical. Together with the findings in chapter 8 ( $\lambda_y$  does not decrease significantly with respect to  $\lambda_r$  as the soil water tension increases) both the numerical and analytical results presented here contradict the hypothesis by Russo and Dagan (1991) that  $\lambda_{yz}\sigma_y^2 \approx \lambda_{rz}\sigma_f^2$  independent of the mean soil water tension (see also chapter 6).

The numerical simulations have shown that the longitudinal spatial moments of the mean concentration in anisotropic soils with moderate to high variability are indeed more or less proportional to the unsaturated hydraulic conductivity. In isotropic soils, however, no direct proportionality is found between  $X_{zz}$  and  $\sigma_y^2$ . Only the variance of the vertical (longitudinal) location of the center of plume position is in all soils found to be directly proportional to  $\sigma$  and hence to  $\sigma_y^2$ . Otherwise the accuracy of the linear macrodispersion theory is limited. The actual mean concentration spreading is found to be not only larger than predicted by the theory, particularly in anisotropic soils, but also in contrast with the very concept of "macrodispersion". The latter is based on the pseudo-Fickian advection-dispersion model, i.e., a Gaussian spatial concentration distribution. The numerical simulations, however, show that the spatial distribution of the average concentration is non-symmetric and highly skewed along the longitudinal axis due to the lognormally distributed vertical velocity component. For the anisotropic soils of moderate to high variability and aspect ratio 6, the actual longitudinal spreading is 60% to 80% larger than predicted.

Numerically computed transverse mean plume spreading and displacement variance of the plume centers also varies nonlinearly with  $\sigma_y^2$ . Transverse spreading of the mean plume far exceeds the predictions of the linear macrodispersion theory. Like for longitudinal spreading, this study finds that the difference between theory and simulation of transverse spreading increases as the anisotropy ratio becomes larger. In a dry soil with aspect ratio 6 and

$\sigma_y^2=3.2$ , the horizontal spreading of the actual (numerical) mean plume is almost one order of magnitude larger than predicted by the theory, and most of the spreading is due to variability in the horizontal plume center displacement. Except for the case of the least variable soil examples, the larger actual mean plume spreading cannot be explained by the effect of local dispersion, which has been neglected in the macrodispersion model. Rather it is the well-known limitation of the macrodispersion model itself, which explains the difference. The main assumption of the linear macrodispersion model is that a particle deviates only insignificantly from the mean travel path, which is obviously erroneous for highly heterogeneous soils (see Figure 9.2c).

Although not specifically addressed in this study, the results do not confirm the findings of Bellin et al. (1992) who concluded from their numerical studies of saturated flow that the erroneous effects of linearized flow and linearized transport may cancel each other. However, their study was limited to the isotropic case with  $\sigma_y^2=\sigma_f^2 \leq 1.6$ . For those conditions the results in this study also indicate a good match between the theoretical model and the numerical results. Only under anisotropic conditions large differences between theoretical and numerical solutions are found, even in mildly heterogeneous soils.

The findings with respect to the underestimation of horizontal spreading are consistent with Rubin (1990), Neuman and Zhang (1990), Tompson et al. (1990), and Zhang and Neuman (1994c). However, the same authors come to the conclusion that the linear macrodispersion model will overpredict longitudinal spreading of the mean concentration. Again, this is in contrast to the results here, where the macrodispersion model generally underestimates the longitudinal macrodispersion. The differences are attributed to the different anisotropy structure modeled here. The above mentioned papers model groundwater transport in isotropic and anisotropic porous media, where the mean flow is generally assumed to be parallel to the direction of strongest correlation. In contrast, mean flow in this study is normal to the direction of strongest correlation. Together with the results of Bellin et al. (1992) this may be an indication that the linear macrodispersion model best predicts transport in isotropic porous

media, underpredicts longitudinal spreading for  $\lambda_l/\lambda_t < 1$  and overpredicts longitudinal spreading for  $\lambda_l/\lambda_t > 1$ , where  $\lambda_l/\lambda_t$  is the ratio of the correlation scales in longitudinal and transverse direction of mean flow.

Similar differences occur in the estimation of the spreading of the mean breakthrough curve across the CL. The theoretical model by Cvetkovic et al. (1992) based on lognormal two-particle joint travel time pdfs with first order parameters, underestimates the spreading of the mean breakthrough curve observed in the numerical simulation. It is generally very accurate in predicting the front end of the mean BTC, but overestimates the peak concentration by 15% to 30% and in highly heterogeneous soils up to 70%. All numerically obtained mean breakthrough curves exhibit considerably more tailing than predicted by the theory. The theory overpredicts the variance of the solute flux by a factor 2 resulting in conservative estimates of the uncertainty (measured in terms of the coefficient of variation) about the solute flux at time  $t$ . Overall I conclude that the linear macrodispersion model as well as the travel time model adopted here for unsaturated flow conditions will give reasonable estimates of the mean concentration in space and mean solute flux across CL, if the purpose is to obtain rough estimates of the mean concentration plume and the mean BTC. For predicting uncertainties associated with extreme events (e.g., early arrival) or associated with very small concentrations, the theoretical models are insufficient.

The statistical analysis of the first time or arrival  $t_a$  of exceedance of a given solute flux  $s/s_0$  at a location  $x$  of the CL has yielded additional insights into the stochastic behavior of solute transport. Not only is it a very attractive and efficient alternative to the numerical Monte Carlo evaluation of the local BTC( $x$ ). It also offers an important practical tool to predict the mean and uncertainty about the time when a given compliance level will be exceeded, an approach that has in the past neither been addressed theoretically nor numerically. The analysis of travel time variability that is suggested here also yields probability distribution functions for the likelihood that a given solute flux level will ever be exceeded at location  $x$  of the CL. However, it does not indicate for how long the compliance level is exceeded or what the total

mass flux will be at  $x$ . The empirical sample probability of exceedance of solute flux levels has a form best described as a slightly skewed quasi-lognormal Gaussian CDF. The average  $\langle t_a(s/s_0) \rangle$ , the time of first exceedance of  $s/s_0$ , shows that the same solute flux levels are reached later as the distance from the center of the CL increases. The least uncertainty is associated with predicting the arrival time of the highest solute flux levels at a location  $x$ , since high solute flux levels can only be attained, if the travel path is fairly direct and undistorted and hence, the travel time is relatively short (physical constriction). This must not be confused with the likelihood that such high solute flux levels occur. Indeed the high flux levels are the least likely. A weakness of the numerical approach to arrival time analysis is that the number of samples decreases with larger  $x$  and larger  $s/s_0$ . The statistical sample moments of  $t_a$  and  $t_p$  have therefore increasing sample error. Note that this type of analysis is different from the numerical or Lagrangian travel time analysis of particles such as e.g. in Smith and Schwartz (1981a) and Cvetkovic et al. (1992).

In light of the high variability of the plume center displacement, the strong skewness of the mean concentration plume, the non-Gaussian distribution of the concentration RFV  $c(\mathbf{x},t)$ , and the very long travel times required for ergodicity to occur, the fundamental problem highlighted in this chapter is the high uncertainty of predicting solute transport with *any* unconditional stochastic model for unsaturated transport, if the soil flux variability becomes large ( $\sigma_y^2 \geq 1$ ). The differences between the stochastically more accurate numerical simulations and the predictions of the macrodispersion and travel time theories may seem minor compared to the differences between the mean predictions and actual values of concentrations and solute fluxes. The conditional simulation approach introduced in the next chapter offers an alternative to the unconditional stochastic models applied in this chapter. But the detailed information obtained in this chapter about the variability of solute transport from small sources - under very idealized conditions - is in agreement with the empirical results reported in numerous field studies: It is generally difficult to model or predict the actual transport behavior of a solute in a heterogeneous soil to a high degree of accuracy.

**Table 9.1**

Input parameters for the various hypothetical soil sites used in the unconditional transport analysis:  $\sigma_f^2$ : variance of  $f = \log K_s$  (log: natural logarithm),  $\sigma_a^2$ : variance of  $a = \log \alpha$ ,  $\rho_{af}$ : correlation coefficient between  $f$  and  $a$ ,  $\Gamma$ : geometric mean of  $\alpha$ ,  $\Delta x$ : horizontal discretization of finite elements,  $\Delta z$ : vertical discretization of finite elements,  $\lambda_{fx}$ : horizontal correlation length of  $f$ ,  $\lambda_{fz}$ : vertical correlation length of  $f$ .

name	$\sigma_f^2$	$\sigma_a^2$	$\rho$	$\Gamma$	H	$\Delta x$	$\lambda_x$	cut
#3	1.0	0.01	0	0.01	-150	10	50	5
#2	0.01	$10^{-4}$						2
#4			1					5
#8	0.12							5
#9	4.0	0.04						10
#12			1			30	300	5
#15					-1000	30	300	5
#21			1		-3000	30	300	10
#22	4.0	0.04				30	300	10
#28	2.25	0.04				30	300	5
#29						30	300	5
#31						15	150	5

**Table 9.2**

Dimensionless distance  $z/\lambda_{iz}$  between the location of the minimum  $CV_c$  and the location of the maximum  $\langle c \rangle$  as a function of dimensionless time  $t' = tV_z/\lambda_{iz}$ . The minimum  $CV_c$  is always located below the maximum  $\langle c \rangle$  i.e., is travelling at a faster rate. Dimensionless times in parentheses refer to soil site #28.

soil type	$\sigma_v^2$	v	$t' = 5$ (4)	$t' = 10$ (8)	$t' = 20$ (16)
#2	.01	1	0.0	0.0	0.0
#8	.10	1	0.1	0.0	0.0
#3	.85	1	0.8	0.6	0.4
#9	3.4	1	2.6	5.2	14.8
#31	.75	3	0.8	1.0	3.2
#12	.53	6	0.4	0.4	1.2
#29	.79	6	0.4	0.6	1.8
#15	1.5	6	0.8	1.6	1.4
#28	1.8	6	0.8	1.8	4.8
#22	3.2	6	2.6	9.6	5.2
#21	3.2	6	4.2	8.6	4.8

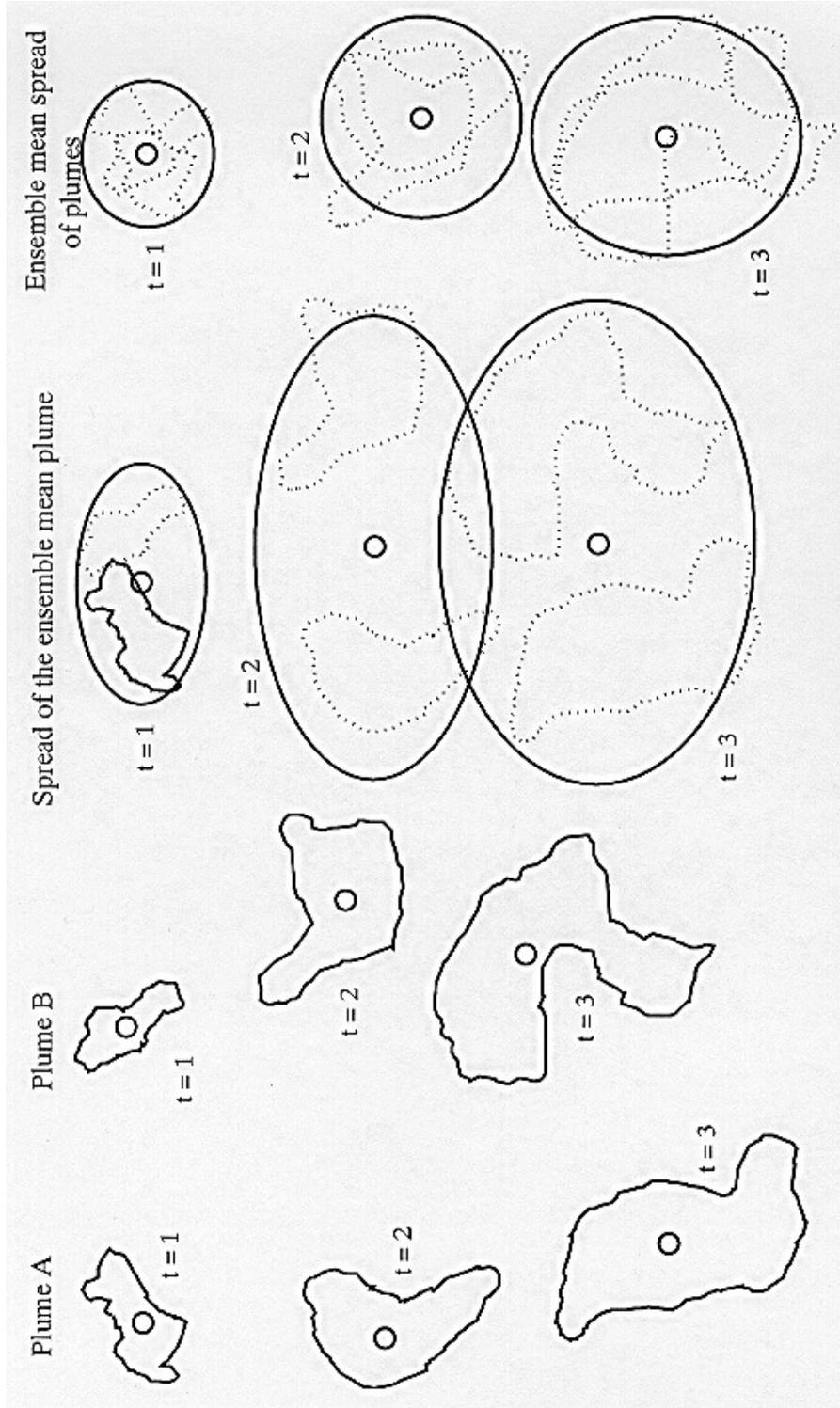


Figure 9.1: mean spreading  $\langle M \rangle$  of plumes in steady state, statistically homogenous porous medium.

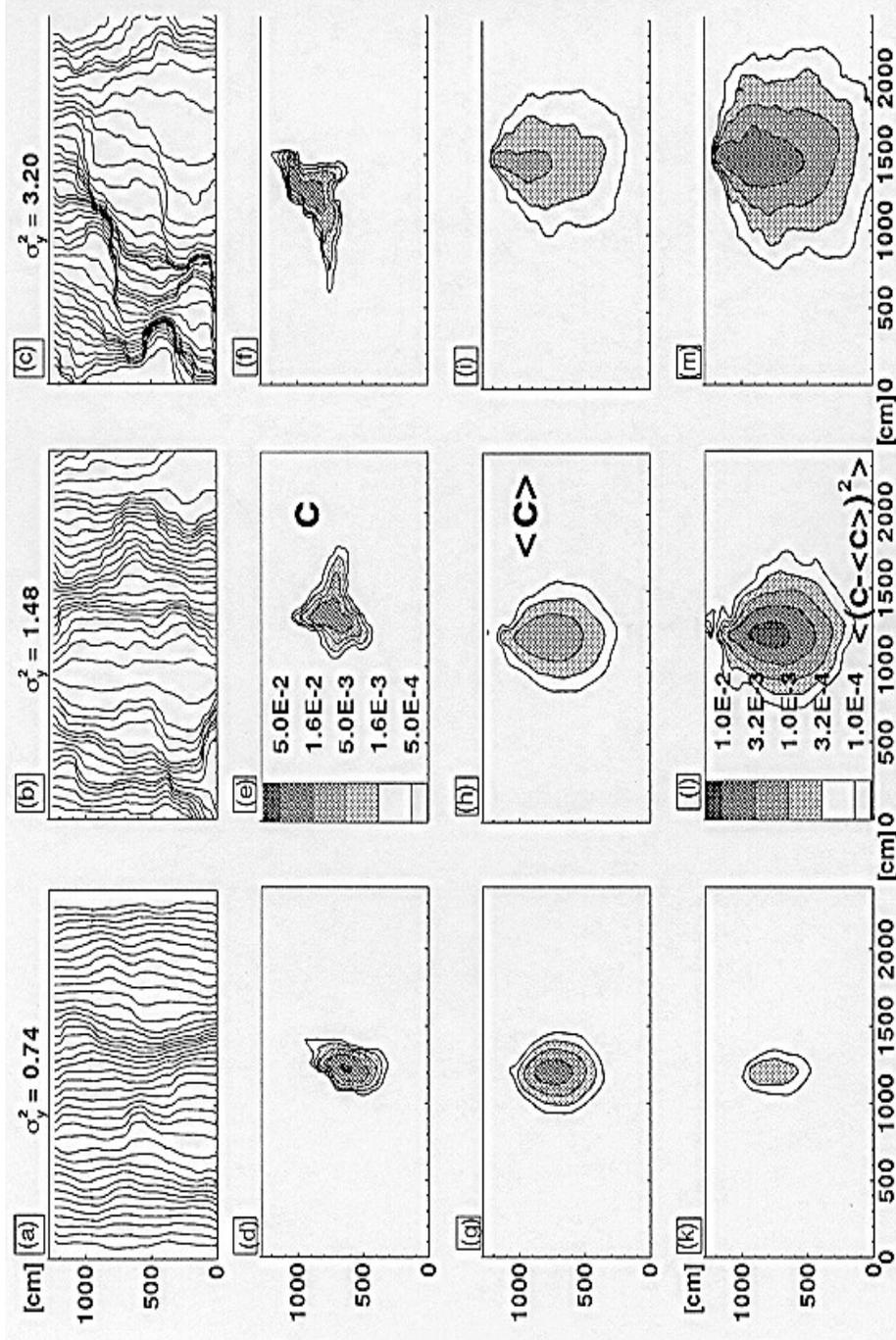


Figure 9.2: sample mean concentration (g,h,i) and sample concentration variance (k,l,m) for the three soils: #29 (left column), #15 (center column), and #22 (right column).

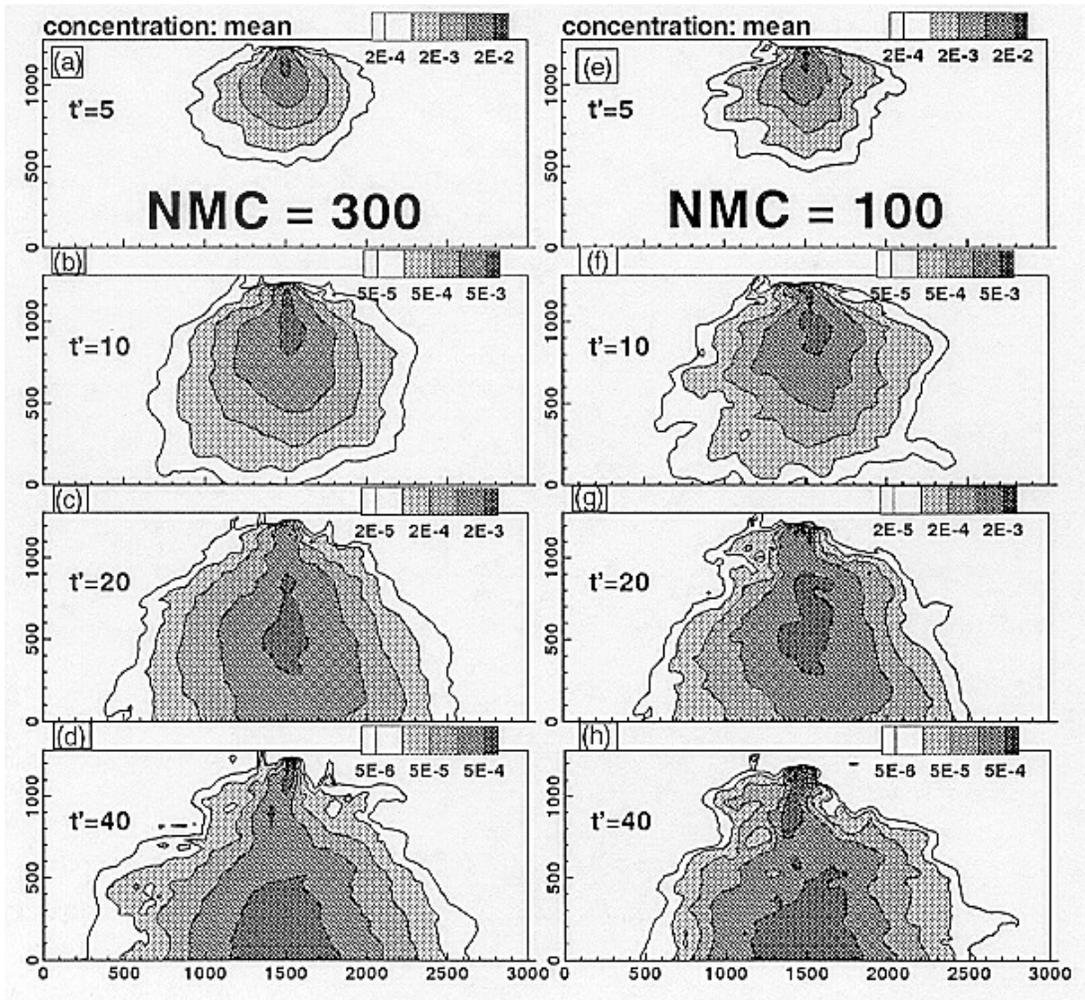


Figure 9.3: Sample mean concentration after 300 realizations (a-d, left column) and after only 100 realizations (e-h, right column) of transport in a strongly heterogeneous flow field (soil site #21,  $\sigma_y^2 = 3.2$ ).

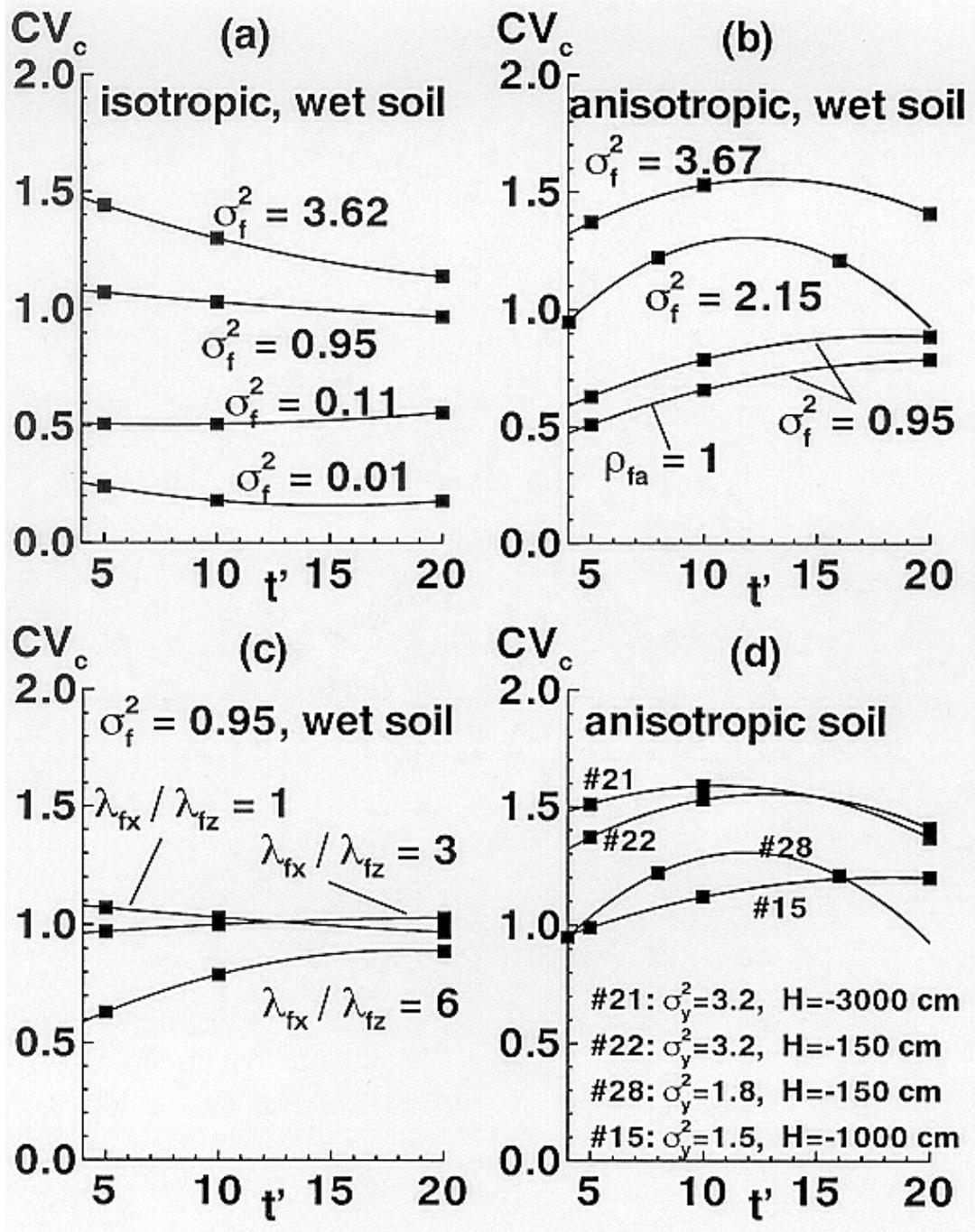


Figure 9.4: Minimum concentration coefficient of variation at or near the center of the mean plume. Except where indicated, f and a are not correlated ( $\rho_{fa} = 0$ ).

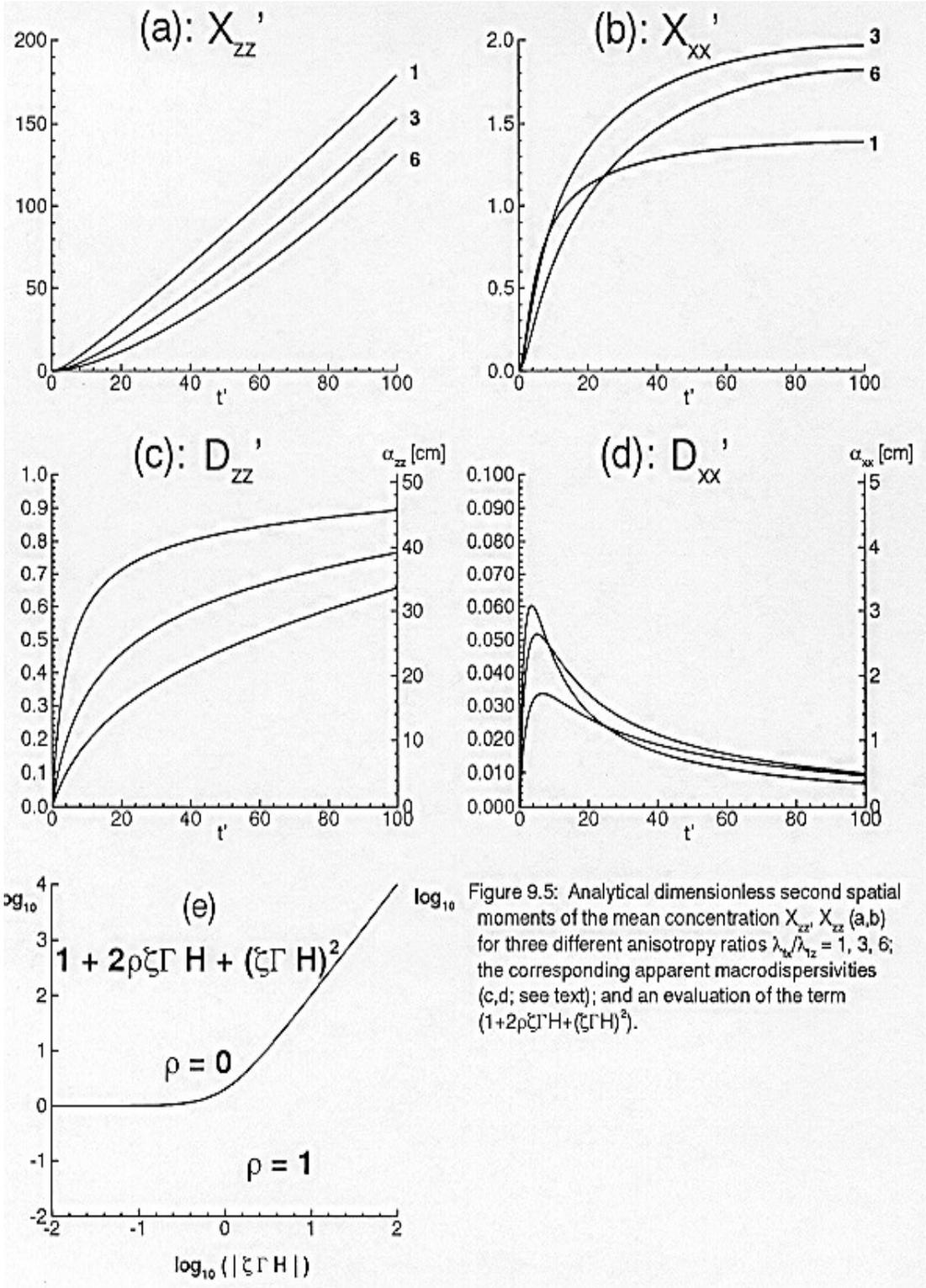


Figure 9.5: Analytical dimensionless second spatial moments of the mean concentration  $X'_{zz}$ ,  $X'_{xx}$  (a,b) for three different anisotropy ratios  $\lambda_x/\lambda_z = 1, 3, 6$ ; the corresponding apparent macrodispersivities (c,d; see text); and an evaluation of the term  $(1+2\rho\zeta\Gamma H+(\zeta\Gamma H)^2)$ .

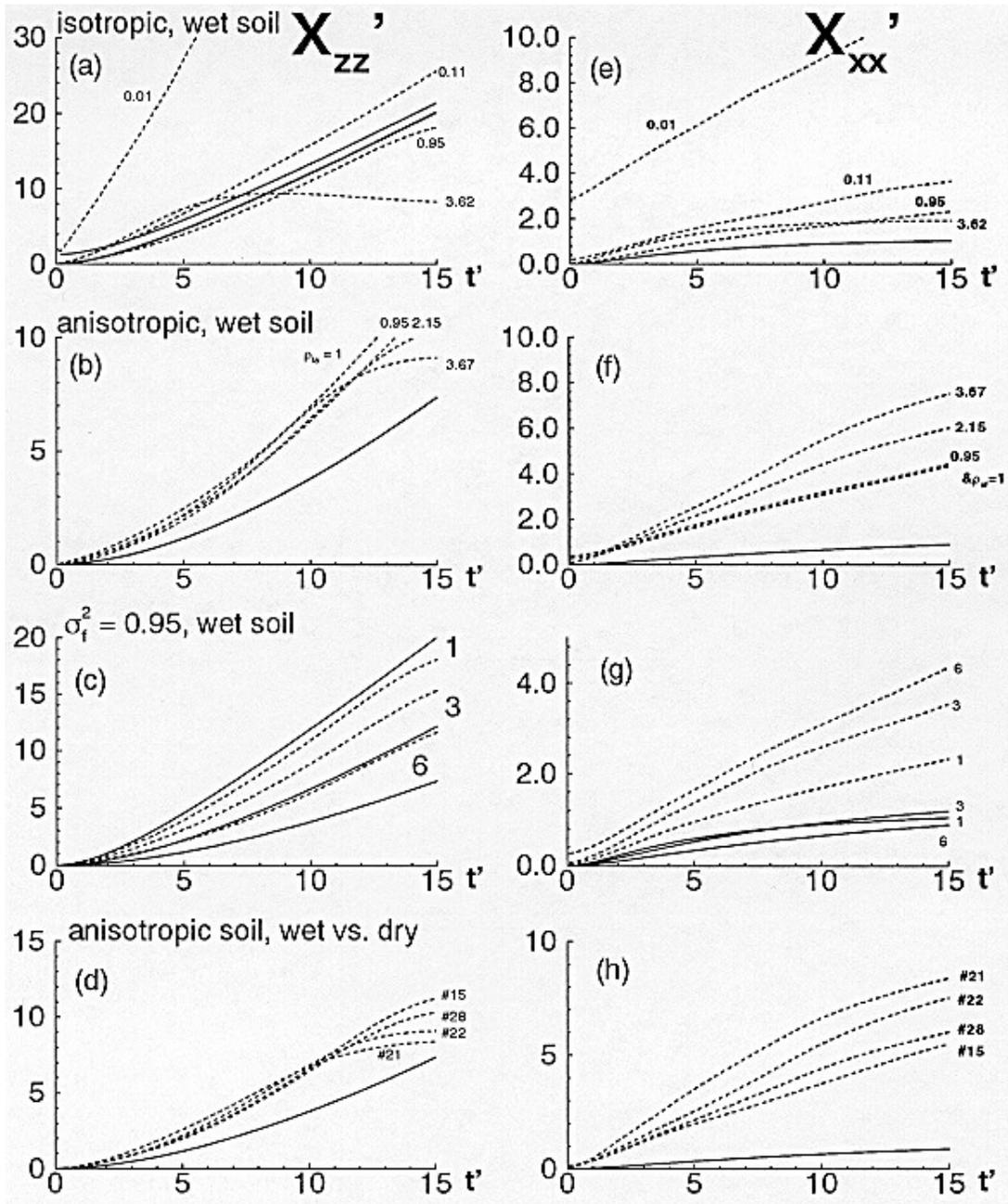


Figure 9.6: Vertical (a-d, left column) and horizontal (e-h, right column) mean concentration second spatial moment. Solid lines are analytical results, while dashed lines are from the Monte Carlo analysis. Labels are explained in Figure 9.4 and Table 9.1 (also see section 9.2.3).

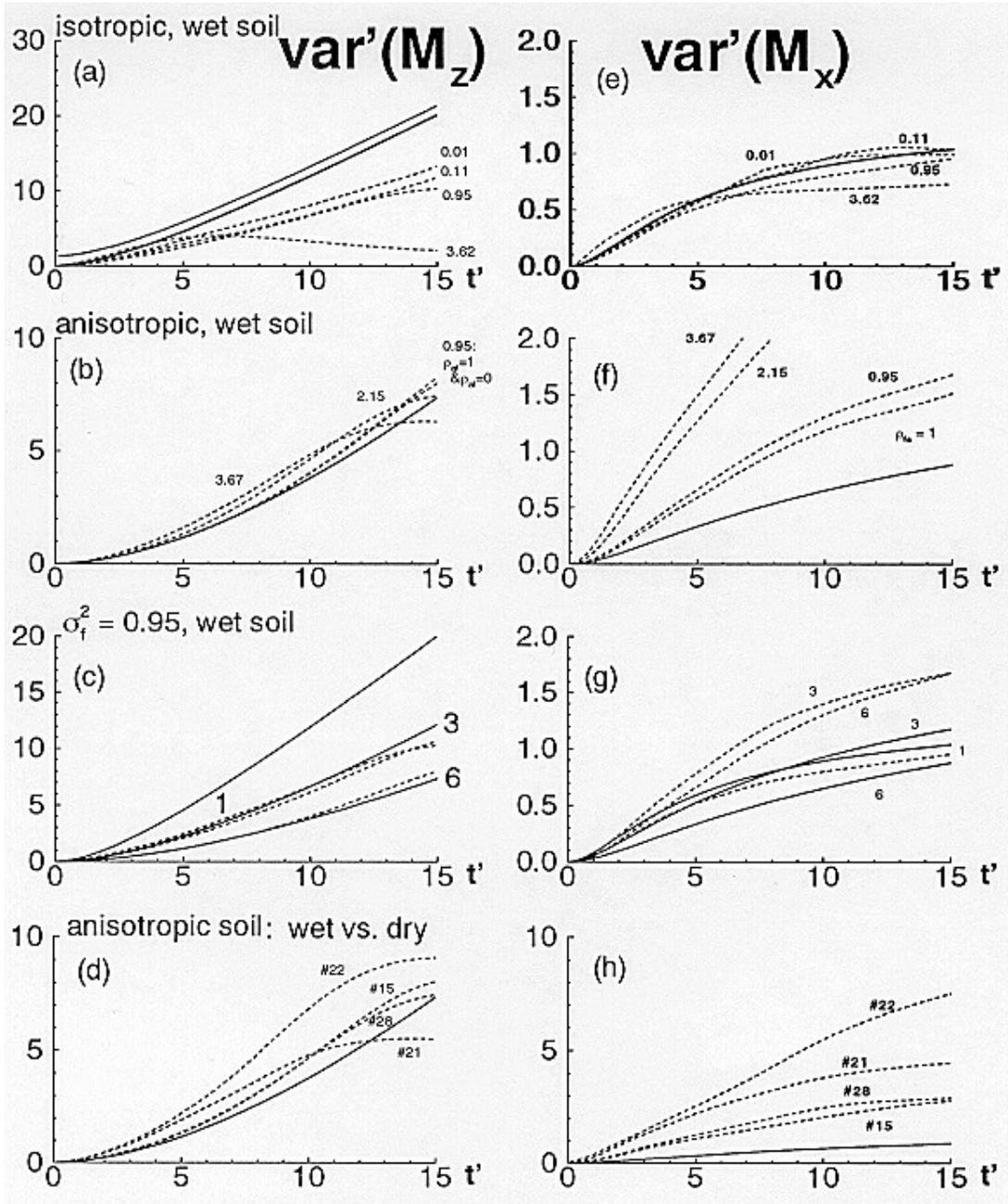


Figure 9.7: Vertical (a-d, left column) and horizontal (e-h, right column) variance of concentration center of mass. Solid lines are analytical results, while dashed lines are from the Monte Carlo analysis. Labels are explained in Figure 9.4 and Table 9.1 (also see section 9.2.3).

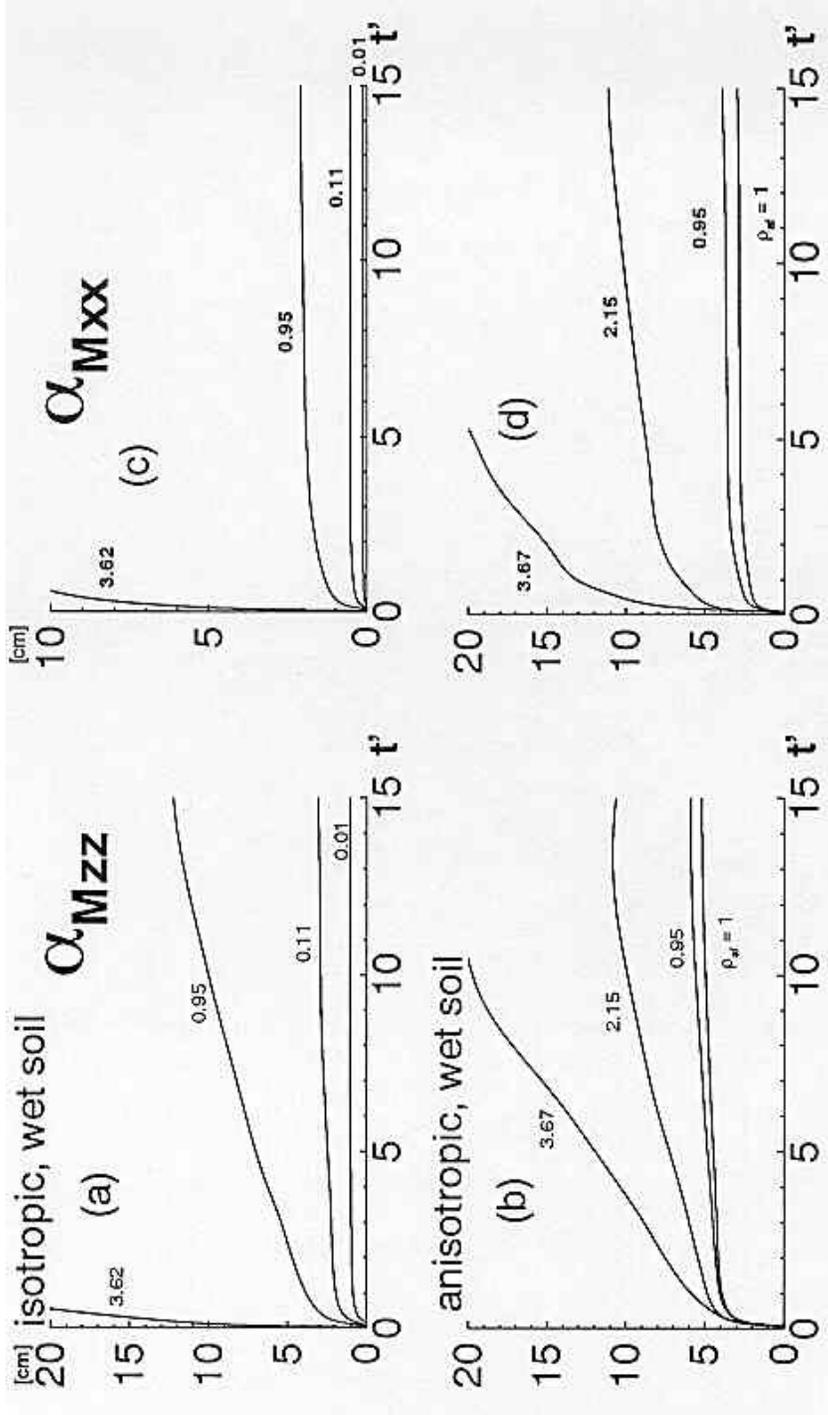


Figure 9.8: [cm] derived from the sample M Table 9.1 For computation of

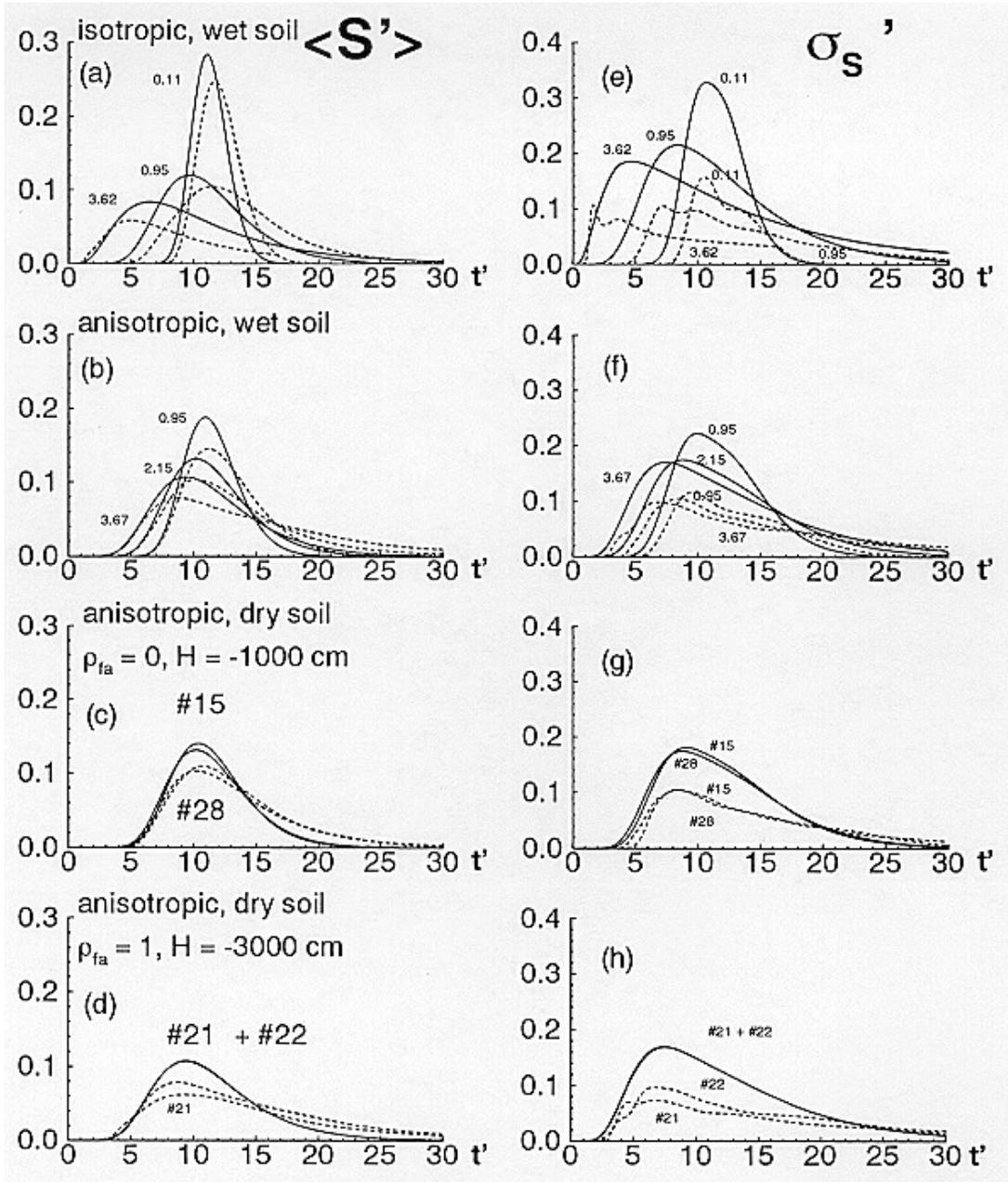


Figure 9.9: Mean solute flux (a-d; left column) and standard deviation of the solute flux (e-h; right column) at a distance  $z'=11.6$  from the center of the source. Solid lines are the analytical model (eqn. 9-11), dashed lines are from the Monte Carlo analysis. For labeling refer to Figure 9.4.

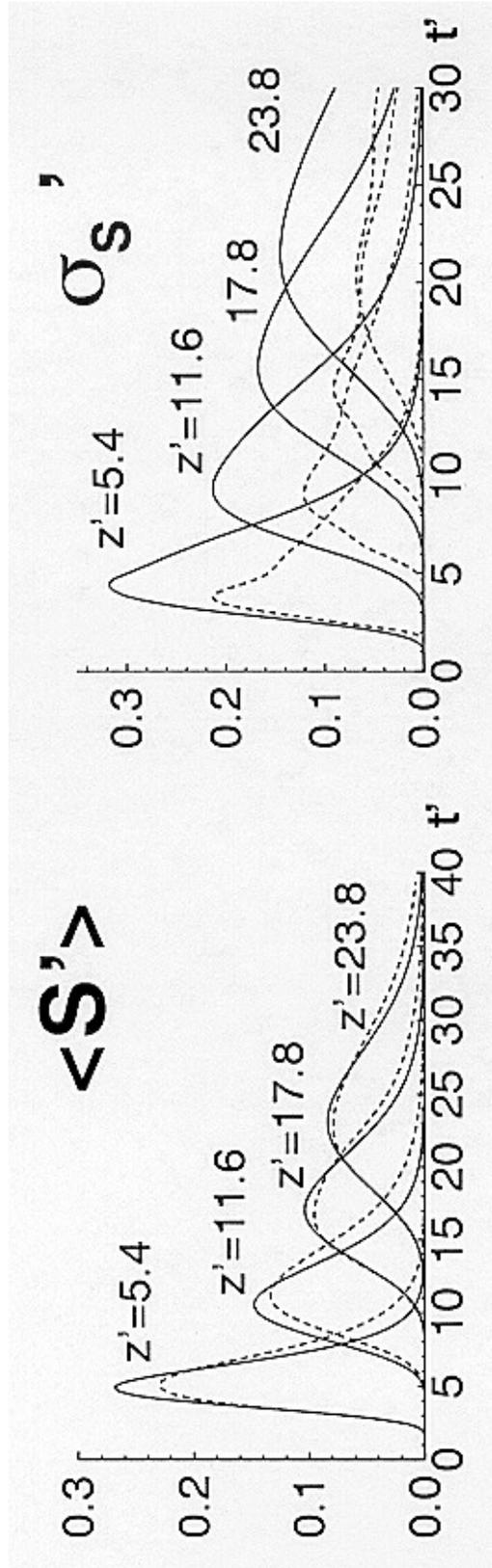


Figure 9.10: anisotropic, wet,

$$\sigma_r \approx 0.95,$$

$$\nu = 3.$$

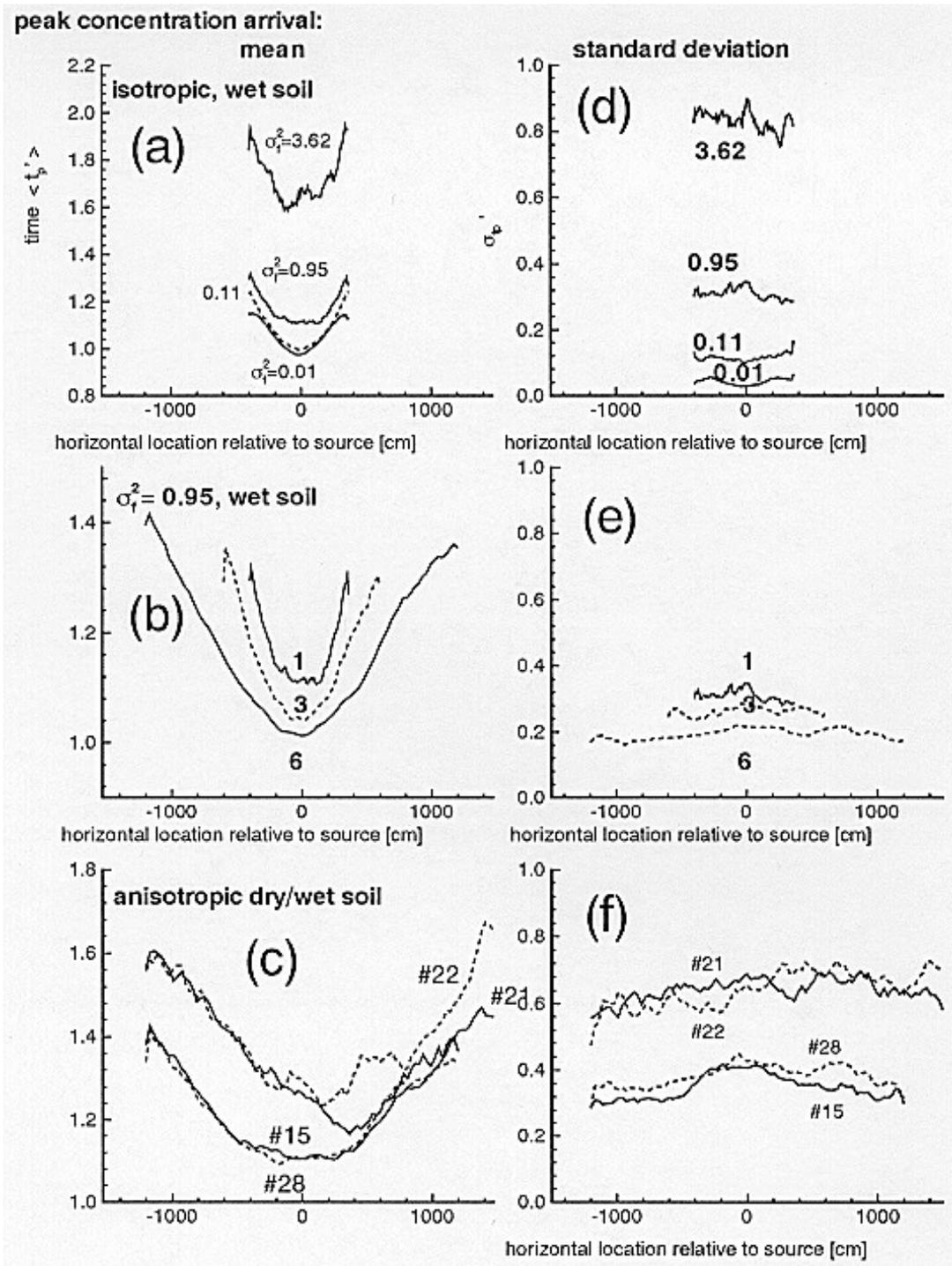


Figure 9.11: Mean (a-c; left column) and standard deviation (d-f; right column) of the arrival time  $t_p$  of peak concentration at the compliance surface. All results are from numerical analysis. Labeling identical to previous Figures.

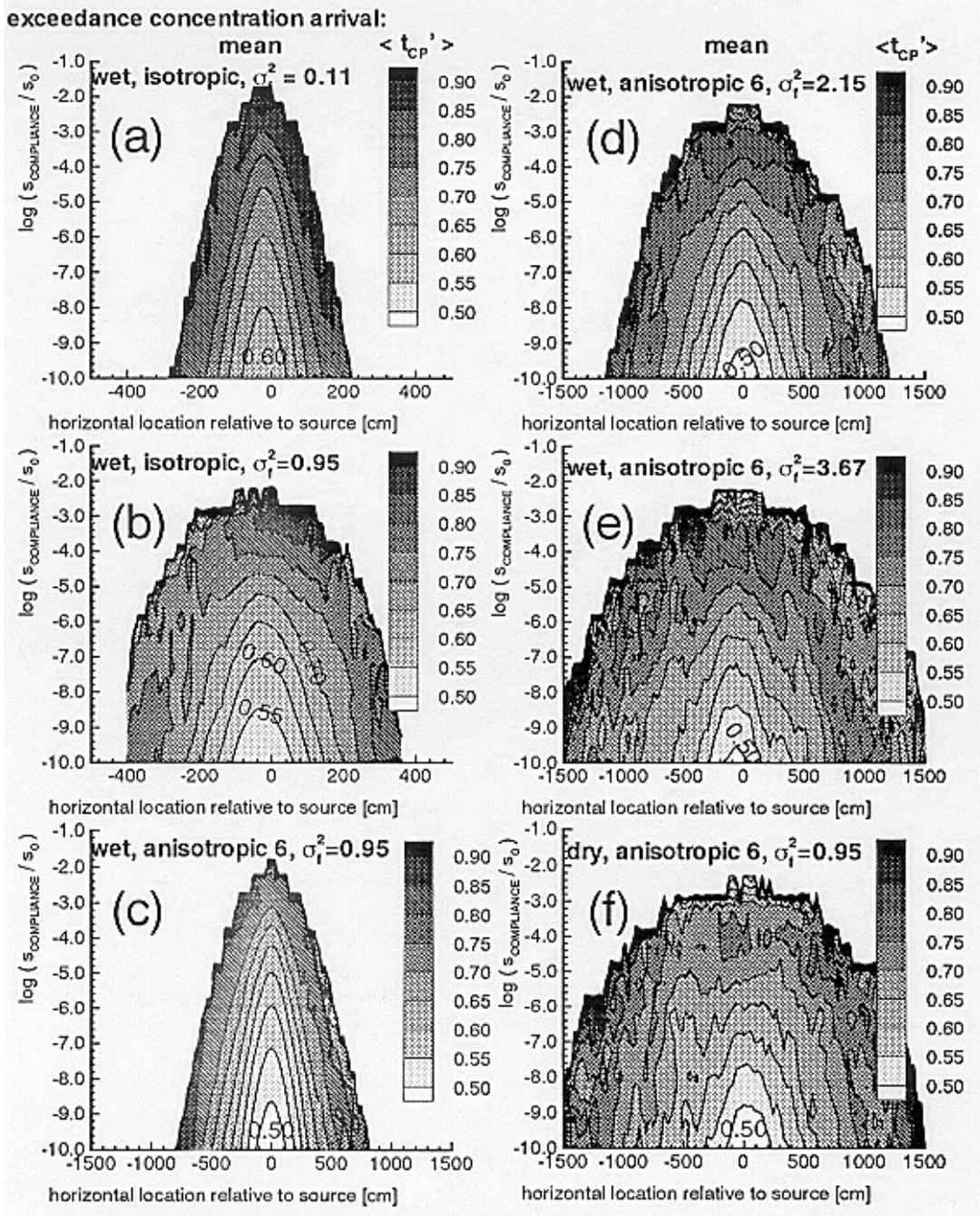


Figure 9.12: Mean arrival time  $\langle t \rangle$  of first exceedance of the compliance concentration. The arrival time is contoured as a function of compliance concentration (y-axis) and as a function of horizontal location (x-axis). The time is normalized with respect to the mean travel time to the CP surface.

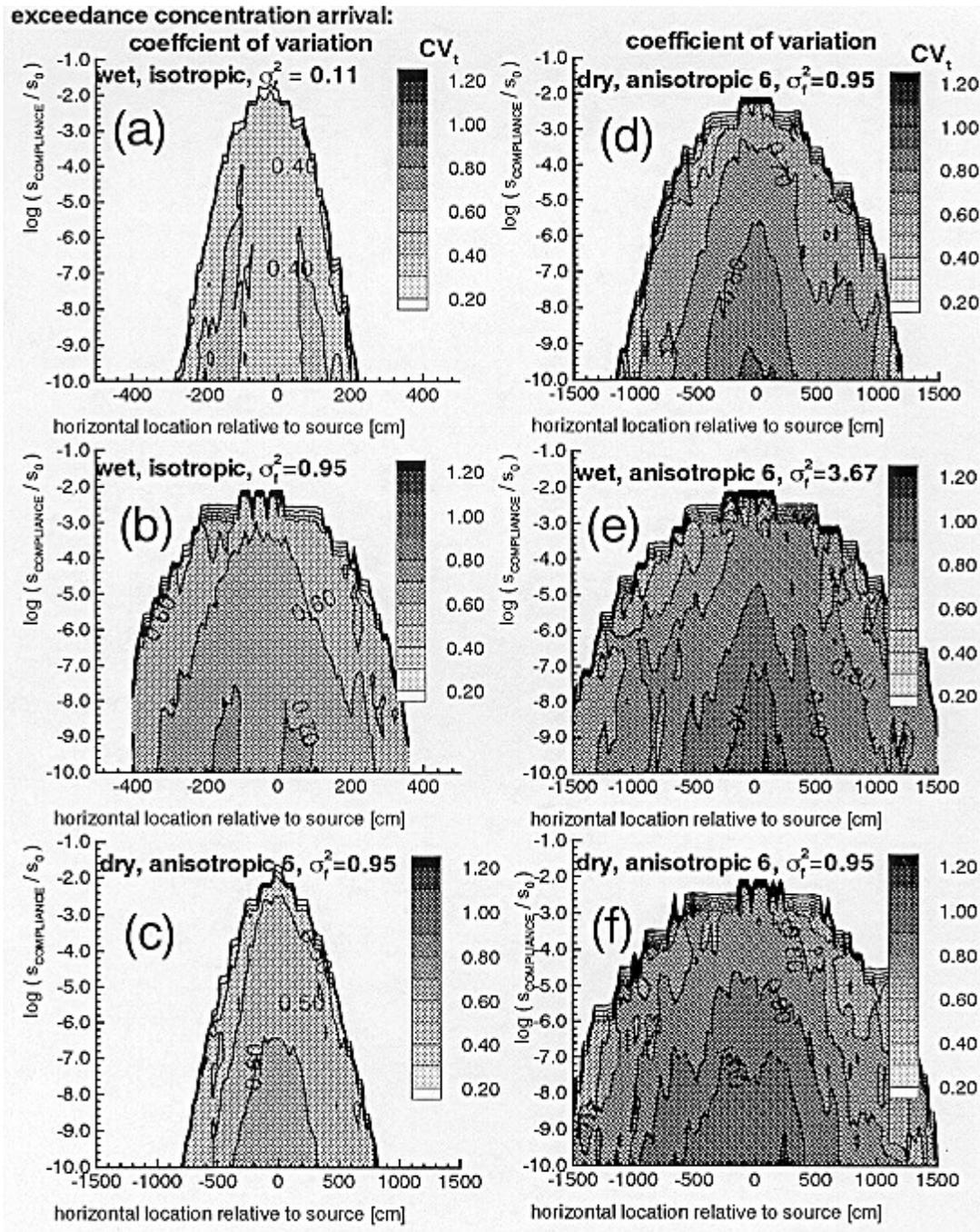
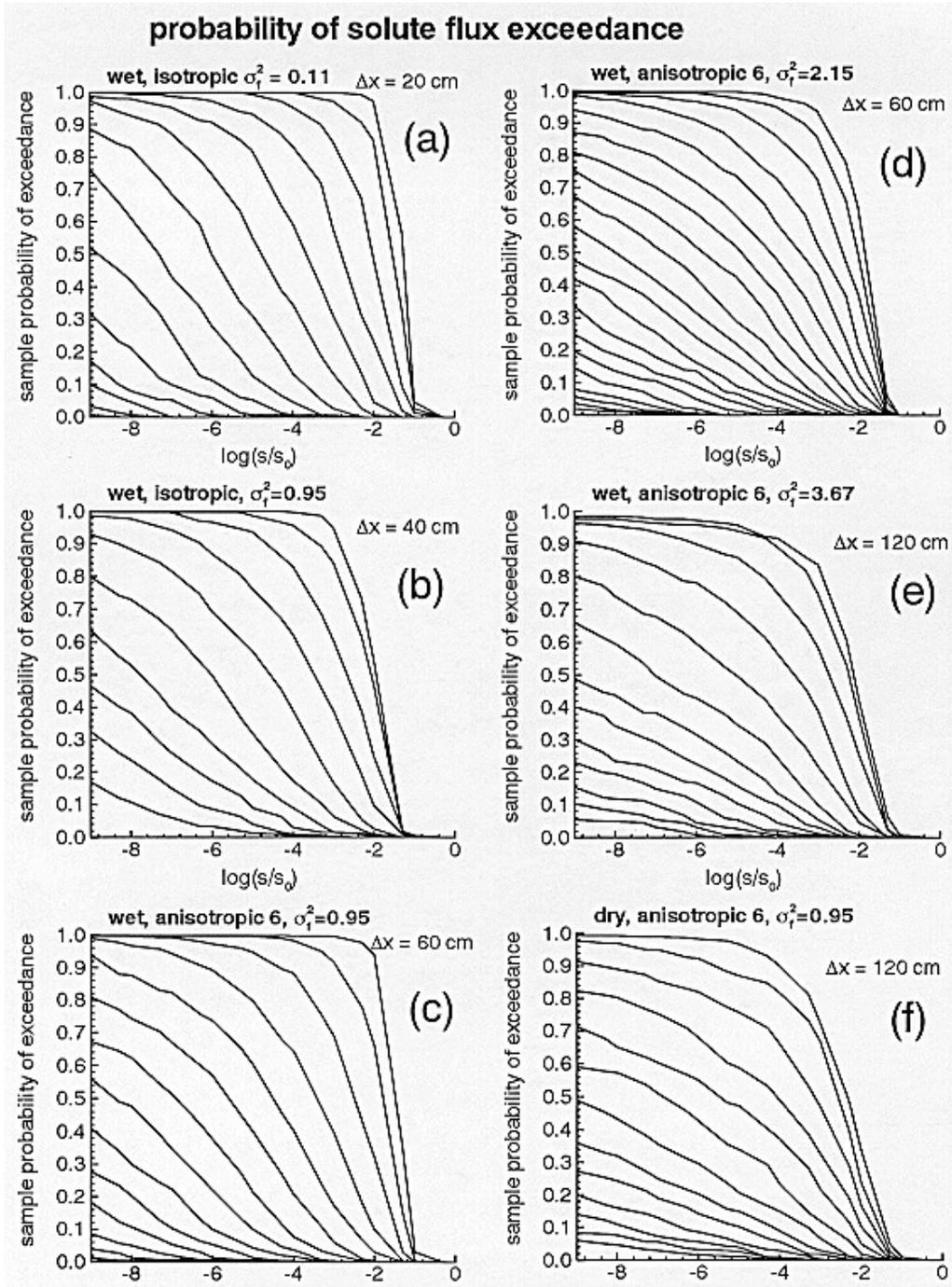


Figure 9.13: Coefficient of variation of arrival time  $CV_{t_{cp}}$  of first exceedance of the compliance concentration. The CV of arrival time is contoured as a function of compliance concentration (y-axis) and as a function of horizontal location (x-axis).



Fig

ure 9.14: Probability of exceedance of solute flux  $s/s$  at different horizontal distances  $x$  relative to the location of the source. The curve with the highest probability is at  $x = 0$ . Subsequent curves are at distance  $x$  that are  $\Delta x$  apart from the neighboring curve.