

**8. STOCHASTIC ANALYSIS OF STEADY-STATE  
FLOW IN HETEROGENEOUS UNSATURATED SOILS  
VIA INTENSIVE MONTE CARLO SIMULATION**

**8.1 Introduction**

Spatial heterogeneity of soil textural properties causes a potentially high degree of variability in the soil moisture flux. In general, the amount of field information that is available to understand or predict moisture flux in the heterogeneous unsaturated zone is limited. Consequently, the modeling of infiltration events and of soil moisture transport to the water table is associated with uncertainty. In many applications quantification of the uncertainty by stochastic analysis is necessary to assess certain hazards or risks. Statistical models enable us to quantify the average soil textural properties, their variability, and their spatial correlation (chapter 2). They are utilized as input for the stochastic analysis of the physical principles governing soil moisture movement. Stochastic analysis provides a statistical description of the variability of soil moisture movement.

Over the past two decades field studies at numerous sites have been used to determine statistical models for describing the variability of soil texture, saturated and unsaturated hydraulic conductivity, soil moisture content, and soil water tension (Ahuja et al., 1984; Anderson and Cassel, 1986; Burden and Selim, 1989; Byers and Stephens, 1983; Cameron, 1978; Ciollaro and Comegna, 1988; Field et al., 1984; Greenhotlz et al., 1988; Greminger et al., 1985; Hopmans et al., 1988; Lauren et al., 1988; Mulla, 1988; Naney et al., 1988; Nielsen et al., 1973; Russo et al., 1981; Russo, 1984; Saddiq et al., 1985; Smettem, 1987; Vieira et al., 1981; Wagenet and Addiscott, 1987; Wierenga et al., 1989). The basic tenet of these field studies has been that the soil hydraulic parameters, which relate unsaturated hydraulic conductivity to soil moisture content and soil water tension, may vary by orders of magnitude over very short distances (decameters to meters). They are best characterized by a

lognormal probability density function (White and Sully, 1992). The variability in soil moisture content and in soil water tension (here generally referred to as suction head or simply 'head') has been found to be very significant with coefficients of variation that often exceed 70%.

This chapter is concerned with the stochastic analysis of these spatially variable field properties. In the stochastic context, the spatially variable physical properties are referred to as random field variables (RFVs, see section 2.5.1) to emphasize that they are not deterministic but described by a probability density function (pdf) (see chapter 2). The stochastic analysis here focuses on characterizing the mean, variance, and covariance of the dependent RFVs  $K$  (unsaturated hydraulic conductivity), head  $h$ , and moisture flux  $v$  as a function of the two RFVs  $K_s$  (saturated hydraulic conductivity) and  $\alpha$  (soil pore size distribution parameter). The physical equations relating  $K_s$  and  $\alpha$  to  $K$ ,  $h$ , and  $v$  are the exponential unsaturated hydraulic conductivity model by Gardner (1958) (eqn. 4-8), the governing unsaturated flow equation (Richards equation, eqn. 4-1), and Darcy's law (eqn. 4-2) (see chapter 4).

In the past, several approaches have been suggested for the stochastic analysis of unsaturated flow problems. These approaches are either based on purely analytical methods or on numerical computer models. Analytical methods (Yeh et al., 1985a,b; Mantoglou et al., 1987a,b,c; Yeh, 1989; also see chapter 4) offer the advantage of providing general mathematical solutions (in form of equations) and an explicit insight into the interdependencies of the statistical parameters for the RFVs  $K_s$ ,  $\alpha$ ,  $K$ ,  $h$ , and  $v$ . Analytical solutions are limited, however, to quasi-infinite soils of mild to moderate variability ( $\sigma_y^2 < 1$ ,  $y = \log K$ ; log refers to the natural logarithm). In contrast, the numerical stochastic analysis of unsaturated flow provides almost unlimited flexibility in designing the model to match with the particular conditions at a field site or to address particular problems of fundamental interest that are difficult to address analytically (Ababou, 1988; Hopmans et al., 1988; Ünlü et al., 1990; Polmann et al., 1991; Russo, 1991). The Monte Carlo technique is particularly attractive, since no stationarity assumptions are needed (Hopmans et al., 1988). Monte Carlo simulations are common in the stochastic analysis of groundwater flow and transport problems (e.g. Freeze,

1975; Delhomme, 1979; Smith and Freeze, 1979; Smith and Schwartz, 1980, 1981a,b; Clifton and Neuman, 1982; Rubin, 1990; Rubin, 1991a,b). The flexibility, however, comes at the expense of rigor and - more importantly - at the expense of potentially enormous computational costs. Single numerical solutions of the nonlinear, heterogeneous flow problem and in particular the steady-state solution are expensive to obtain, let alone multiple solutions in a Monte Carlo simulation (chapter 6). It is therefore not surprising that the numerical analysis of unsaturated flow in heterogeneous soils has been limited both in the number of studies published and in the number of random realizations implemented for each study.

Recently, Harter and Yeh (1993) have developed an efficient combined analytical-numerical method (called ASIGNing) that reduces the cost of computing the solution to Richards equation ( $K$ ,  $h$ ) and Darcy's law ( $\mathbf{v}$ ) by two orders of magnitude, even for highly heterogeneous input random fields  $K_s$  and  $\alpha$  (chapter 7). In this chapter, ASIGNing is applied as the cornerstone to Monte Carlo simulations with a large number of realizations ( $N=1000$ ). The objective is to obtain highly accurate stochastic solutions of the dependent RFVs  $K$ ,  $h$ , and  $\mathbf{v}$  in two-dimensional, vertical, steady-state unsaturated flow-fields of moderately to strongly heterogeneous soils in order to implement a stochastic analysis with respect to the statistical input parameters describing the lognormally distributed RFVs  $K_s$  and  $\alpha$ . Accuracy here refers not only to the numerical accuracy, but also to the statistical accuracy of the sample moments obtained from the Monte Carlo simulation (compare to chapter 6). The study is intended to provide new insight to the problem of variably saturated flow in highly heterogeneous porous media and to critically assess the assumptions and the range of validity of the analytical stochastic steady-state flow model by Yeh (1985a,b), which has here been adopted for 2-D flow and the particular case of  $\alpha$  having a lognormal pdf (chapter 4). All past studies of unsaturated flow in heterogeneous porous media have been limited to soils with a variance of  $\gamma$  no larger than one ( $\sigma_\gamma^2 \leq 1$ ). Field studies have shown that the variance of the logarithms of the saturated and unsaturated hydraulic conductivity often exceed 1 and may be as large as 3, sometimes even higher (Nielsen et al., 1973; Vieira et al., 1981; Anderson and Cassel, 1986; Ciollaro and

Comegna, 1988; Lauren, 1988; Wierenga et al., 1989). In this study hypothetical isotropic and anisotropic soils are investigated with  $\sigma_y^2$  ranging from 0.01 to 3.2. The variability in  $y$  may be due to the dryness of the soil (large mean soil water tension), or due to a high variability in  $K_s$  and  $\alpha$ , or due to a combination of these.

It is expected that the results will be useful for the assessment of soil moisture movement in variably saturated soils, and also for the assessment of unsaturated transport since the second-order moments characterizing unsaturated flow are also used to estimate solute transport in heterogeneous porous medium (Russo, 1993a,b; Dagan, 1982, 1984; Rubin, 1990, 1991a, 1992) (see chapter 9). Among others, the proposed model allows accurate estimates of the head covariance function and of the cross-covariance function between head,  $K_s$ , and  $\alpha$  for arbitrary boundary conditions. These (cross-) covariances are necessary to implement conditional simulations of unsaturated flow and transport (chapter 10).

Polmann et al. (1991) have pointed out the importance of model-generated data both to validate analytical models and to improve our understanding of unsaturated flow processes: Large amounts of data are difficult to obtain in the field and problems of sampling accuracy and soil heterogeneity become intertwined. In contrast, *"a simulation experiment based on model-generated data enables us to focus on individual sources of heterogeneity while holding others fixed. If carried out systematically, this approach can identify the critical factors which control moisture movement through heterogeneous soils"* (ibid., p.1448). The following analysis is presented in this spirit.

First a rigorous definition is given for the term Monte Carlo and some simple measures are introduced to determine the sample accuracy of numerical stochastic results as a function of the number of realizations. Then an outline is given of the actual implementation of the Monte Carlo simulations. The results are presented separately for each of the dependent RFVs  $\log K$ ,  $h$ , and  $v$ . The general statistical properties of each of these RFVs are discussed, the sample error associated with the Monte Carlo simulation is determined, boundary effects are investigated, the dependency of the RFVs on the independent parameters is analyzed and

compared with the analytical stochastic model introduced in chapter 4.

## 8.2 Monte Carlo Simulation

### 8.2.1 Definition and Theoretical Sampling Accuracy

The Monte Carlo method is defined as a random sampling procedure used to numerically evaluate the integral:

$$G = \int_{\Omega_0} g(\mathbf{X})p(\mathbf{X})d\mathbf{X} \quad (8-1)$$

where  $G$  is the expected value of the random variable  $g(\mathbf{X})$  defined in the sample space  $\Omega$ .  $g(\mathbf{X})$  is an analytical function of a vector  $\mathbf{X}$  of random variables or random field variables with a joint-pdf  $p(\mathbf{X})$ . The numerical integration by Monte Carlo is performed as a *game of chance* (Kalos and Whitlock, 1986), where  $N$  sets of random (field) variables  $\mathbf{X}$  are sampled from the joint-pdf  $f(\mathbf{X})$ . In Monte Carlo sampling is equivalent to generating a random number or random field (chapter 3). The integral above is approximated by the sum:

$$G_N = \frac{1}{N} \sum_i g(\mathbf{X}_i) \quad (8-2)$$

$G_N$  is the sample mean. Through the fundamental theorem of large numbers it is guaranteed that  $\langle G_N \rangle$  converges in the mean square to  $G$  ( $\langle \rangle$  indicates expected value)

If  $\mathbf{X}_i$  is independent of  $\mathbf{X}_j$ ,  $i \neq j$ , and if it is known *a priori* that  $g(\mathbf{X})$  has a Gaussian pdf or if  $N$  is very large, the sampling error (variance)  $\epsilon_G^2$  of the normally distributed sample mean  $G_N$  is (Haan, 1977; Kalos and Whitlock, 1986):

$$\epsilon_G^2 = \frac{\sigma_g^2}{N} \quad (8-3)$$

where  $\sigma_g^2$  is the variance of  $g(\mathbf{X})$ . Note that the sample moments are assumed to have a Gaussian distribution rather than the commonly applied t-distribution, which is justified since samples of at least 40 independent measurements are subsequently evaluated. Since  $\sigma_g^2$  is not

known, it must be estimated by:

$$\text{var}_g = \frac{1}{N} \sum g^2(\mathbf{X}_N) - \left( \frac{1}{N} \sum g(\mathbf{X}_N) \right)^2 \quad (8-4)$$

The sample variance  $\text{var}_g$  itself has an associated sampling error. For the square-root  $s_g$  of  $\text{var}_g$ , the sampling error (variance)  $\epsilon_{s,g}^2$  of the sample standard deviation  $s_g$  is approximately (Yevjevich, 1972):

$$\epsilon_{s,g}^2 \approx \frac{\sigma_g^2}{2N} \quad (8-5)$$

For the sample variance  $\text{var}_g$  itself, simple heuristic considerations lead to the following expression of the expected sampling error (standard deviation)  $\epsilon_{v,g}$  given  $\epsilon_{s,g}$  :

$$\epsilon_{v,g} = \frac{(\sigma_g + \epsilon_{s,g})^2 - (\sigma_g - \epsilon_{s,g})^2}{2} \quad (8-6)$$

which simplifies with the help of (8-5) to:

$$\epsilon_{v,g} \approx \frac{2\sigma_g^2}{\sqrt{2N}} \quad (8-7)$$

(8-3) and (8-7) can be used to estimate the sampling errors of past Monte Carlo studies of unsaturated moisture movement. Ünlü et al. (1990) implemented Monte Carlo simulations with 50 realizations of the moisture redistribution process in a one-dimensional soil column. The column consisted of 100 random soil layers of varying soil properties. The study addressed the sensitivity of the head and vertical flux moments to the statistical input parameters and to the boundary conditions. Using (8-3) with  $N=50$  and the sample head variances reported by the authors, the 95% confidence interval ( $\pm 2\epsilon_c$ ) for sample mean head ranges from  $\pm 0.2$  cm to  $\pm 10$  cm ( $\pm 28\%$  of the sample standard deviation). This is a small range given that the mean suction head varies over several hundred centimeters along the vertical column, but relative to the

standard deviation it is rather significant.

The same accuracy of sample mean heads is found for Hopmans et al. (1988) who used 10 realizations of a two-dimensional, hypothetical soil cross-section consisting of 50 vertically homogeneous columns i.e., with vertically constant  $K_s$  and soil pore size distribution, but variable mean head. The analysis was used to derive the nonstationary head and flux distribution moments as a function of the distance above water-table by averaging over all 10 samples in all 50 columns in the same horizontal layer. For a conservative estimate of the associated sampling error it can be assumed that after a horizontal distance of roughly 10 soil columns the local head and flux moments are completely independent of each other. Then,  $N = 10 * 50/10 = 50$  (the number of realizations times the number of independent soil columns per realization), the same as in Ünlü et al. (1990).

The 95% confidence interval of the sample variance ( $\pm 2\epsilon_{v,g}$ ) in these two studies ranges from 60% to 140% of the ensemble standard deviation (as represented by the square-root of the sample variance). The same confidence interval applies to the sample covariance functions computed in Hopmans et al. (1988). While the results from both studies may be considered accurate for practical purposes, the sampling error of the mean and variance in head and flux is large enough to question the use of these results for comparison with analytical results.

From (8-7) it is straightforward to determine that the number of realizations necessary to reduce the 95% confidence interval of the sample variance (or covariance) to within  $\pm 10\%$  and  $\pm 5\%$  of the ensemble variance is  $N \geq 800$  and  $N \geq 3200$ , respectively. In other words, roughly 1000 realizations are necessary to estimate the local sample variance (covariance) such that in 19 of 20 Monte Carlo simulations the sample error is less than 10% of the ensemble variance. If the variance is obtained not only by averaging over the  $N$  samples at  $\mathbf{x}$ , but also by averaging spatially, less realizations are needed and the confidence interval will depend on the spatial correlation of the RFVs. Spatial averaging, however, assumes that the field is weakly stationary i.e., that the ensemble mean  $G$  and ensemble variance  $\sigma_g^2$  are identical for all  $\mathbf{x}$  in the simulation domain. In this study, the primary interest is to obtain the sample mean, variance,

and covariance separately for each  $\mathbf{x}$  so that stationarity of the dependent RFVs does not have to be assumed *a priori*.

### 8.2.2 General Computational Procedures

Applied to the unsaturated flow problem,  $g(\mathbf{X})$  is any of the random field variables  $\log K$ ,  $h$ , and  $\mathbf{v}$ .  $\mathbf{X}$  is a vector of the two RFVs  $K_s$  and  $\log \alpha$ . Note that the mean and variance of the RFV  $g(\mathbf{X})$  and the sampling errors  $\epsilon_G$  and  $\epsilon_{s,g}$  are functions of location  $\mathbf{x}$  unless stationarity is assumed. To obtain the sums (8-2) and (8-4) individual realizations  $\mathbf{X}_i$  of the Gaussian distributed RFVs  $f = \log K_s$  and  $a = \log \alpha$  must be generated. For each realization of  $f$  and  $a$ , the corresponding random field solutions of  $y$ ,  $h$ ,  $v_x$  (horizontal flux), and  $v_z$  (vertical flux) are computed by using the ASIGNing technique described in chapter 7. The results are evaluated not only to determine the first and second moment of the pdfs, (8-2) and (8-4), but also the histograms of both the independent RFVs  $f$  and  $a$  (for control) and the dependent RFVs  $y$ ,  $h$ ,  $v_x$ , and  $v_z$ . The local sample covariance field of a RFV  $g$  centered around point  $\mathbf{x}$ ,  $cov_g(\mathbf{x}, \xi)$ , is computed by:

$$cov_g(\mathbf{x}, \xi) = \frac{1}{N} \sum_i^N [g_i(\mathbf{x})g_i(\mathbf{x}+\xi) - G_N(\mathbf{x})G_N(\mathbf{x}+\xi)] \quad (8-8)$$

And similarly the local cross-covariance field for two RFVs  $g$  and  $g'$  is defined as:

$$cov_{gg'}(\mathbf{x}, \boldsymbol{\xi}) = \frac{1}{N} \sum_i^N g_i(\mathbf{x}) g_i'(\mathbf{x} + \boldsymbol{\xi}) - G_N(\mathbf{x}) G_N'(\mathbf{x} + \boldsymbol{\xi}) \quad (8-9)$$

Here,  $g$  and  $g'$  denote any of the RFVs  $f$ ,  $a$ ,  $y$ ,  $h$ ,  $v_x$ , and  $v_z$ . Note that unlike in any previous numerical studies of unsaturated flow, the sample covariance and cross-covariance is evaluated over the sample space  $N$  and does not involve any spatial averaging. During the Monte Carlo simulation, the sums in (8-2), (8-4), and (8-8), (8-9) are updated after each realization i.e., the outcomes  $g_i(\mathbf{x})$ ,  $g_i^2(\mathbf{x})$ ,  $g_i(\mathbf{x})g_i(\mathbf{x} + \boldsymbol{\xi})$ , and  $g_i(\mathbf{x})g_i'(\mathbf{x} + \boldsymbol{\xi})$  are added to the sum of their outcomes from previous realizations  $j < i$ , respectively. For efficient data management, the actual results of each realization are discarded once all sums and histograms (see below) have been updated. After the Monte Carlo simulation is completed, the sample means, variances, and (cross-)covariances (8-2), (8-4), (8-8), and (8-9) are obtained explicitly from the respective sums.

The sample mean  $G_N$  and sample variance  $var_g$  are themselves two-dimensional realizations of RFVs and summary statistics can be obtained by spatially averaging over the sample mean field and the sample variance field. For  $N=1000$ , the spatial average of the local moments (average sample mean and average sample variance) has a very narrow confidence interval i.e., it is a very accurate estimate of the true mean and variance of the dependent RFV's provided that the mean and variance fields are found to be weakly stationary. Since the statistical input parameters  $F$ ,  $A$ ,  $H$ ,  $\sigma_f^2$ , and  $\sigma_a^2$  in this study are all independent of location, the statistical moments of  $y$ ,  $h$ , and  $\mathbf{v}$  must also be weakly stationary. In the weakly stationary Gaussian case, the spatial variance of the sample mean field,  $var(G_N)$ , and the standard deviation of the sample variance field,  $std(var_g)$ , also provide an estimate of the sampling error and should be similar to (8-3) and (8-7) if the sample moments indeed converge in the mean square sense (see chapter 2). For better comparison, the spatial variance of the local sample means is normalized by the expected sampling error  $\epsilon_G^2$ , where the latter is obtained by using the average sample variance  $\langle var_g \rangle$  rather than the (unknown) ensemble variance  $\sigma_g^2$  in (8-3). Then the dimensionless actual sample error of the sample mean is defined as:

$$\epsilon_G^{/2} = N \frac{\text{var}(G_N)}{\langle \text{var}_g \rangle} \quad (8-10)$$

where  $N$  is the number of realizations in the Monte Carlo simulation. Similarly, the dimensionless actual sample error of the sample variance is defined by using (8-7):

$$\epsilon_{\text{var.g}}^{/} = \frac{\sqrt{2N \text{var}(\text{var}_g)}}{2 \langle \text{var}_g \rangle} \quad (8-11)$$

As in (8-10) and (8-11) an apostrophe ' with a statistical moment subsequently indicates that the moment has been normalized and that it is dimensionless.

The sample covariance and cross-covariance fields are computed in a window of half the side-length of the simulation domain centered around each of nine sample locations  $\mathbf{x}$  (Figure 8.1). In other words, (cross-)covariance values around  $\mathbf{x}$  are computed only for separation distances  $|\xi|$  not exceeding one-fourth of the domain-length in each principal direction. The choice of the locations  $\mathbf{x}_i$  and the size of the windows surrounding them is dictated by several objectives: to provide local sample (cross-)covariance fields that can be checked for spatial trend; to spread the locations  $\mathbf{x}_i$  as far apart as possible to minimize correlation between the sample (cross-)covariance fields; and finally to provide equally sized sample fields. The latter is necessary to obtain average (cross-)covariance fields  $C_{gg}(\xi)$  for each lag distance  $\xi$ :

$$C_{gg}(\xi) = 1/9 \sum_{i=1}^9 \text{cov}_{gg}(\mathbf{x}_i, \xi) \quad (8-12)$$

The window for the covariance and cross-covariance fields around the center point  $\mathbf{x}_{\text{center}}$  of the simulation domain is chosen to be as large as the simulation domain itself to provide additional information on  $\text{cov}_{gg}(\mathbf{x}_{\text{center}}, \xi)$  at lag distances up to one-half of the domain size in each dimension.

Finally, two classes of histograms are computed: The local histogram of the RFVs at  $\mathbf{x}_{\text{center}}$  and the total histogram of all outcomes of each RFV regardless of location. The

histograms are updated after each realization. Figure 8.2 shows a simple flow-chart of the Monte Carlo simulation procedure.

### 8.3 Simulation Parameters and Implementation

This study is geared towards the equivalent of a formal stochastic analysis and not towards a particular field application. Nevertheless, it is imperative to implement the numerical analysis such that the demands of actual model applications are addressed. An intensive study of field heterogeneity in an arid soil was implemented by Wierenga et al. (1989, 1991) near Las Cruces, New Mexico. Measurements of the in situ and laboratory saturated hydraulic conductivity, soil water content, and soil water retention function provide valuable information about the magnitude of field soil variability. The saturated hydraulic conductivity was found to be on the order of  $10^0 - 10^5$  cm/d and to have a lognormal distribution. Variances in  $\log K_s$  (natural based logarithms) range from as small as 0.1 to as large as 3 depending on the measurement method and the soil layer. The overall variability of  $\log K_s$  at the Las Cruces site is approximately 1.5. A geostatistical analysis of these  $\log K_s$  data reveals that the correlation structure can be modeled by an exponential covariance function with an integral scale of a few meters in the horizontal direction and an integral scale of a few decimeters in the vertical direction. To describe the spatial variability of pore-size distribution related parameters, Wierenga et al. (1989) fitted the VanGenuchten model (4-4) to the empirical retention curves obtained from soil cores. A statistical analysis of their data shows that the VanGenuchten  $\alpha$  is lognormally distributed with a geometric mean  $\Gamma = 0.04$  and variance  $\sigma_a^2 = 0.3$ . The VanGenuchten  $n$  is also best fitted by a lognormal distribution with a geometric mean  $n$  of 1.6 and a variance in  $\log(n)$  of 0.02. Covariance functions similar to that of  $\log K_s$  were found for the parameters characterizing the pore-size distribution (unpublished study report, Harter, 1991). This information of spatial variability of unsaturated hydraulic properties in the field provided the basis for the design of the Monte Carlo simulations.

The input parameters for all Monte Carlo simulations (MCSs) in this study are (Table 8.1): the variances of  $f$  and  $a$ ,  $\sigma_f^2$ ,  $\sigma_a^2$ , the correlation between  $a$  and  $f$ ,  $\rho_{af}$ , the geometric mean  $\Gamma$  of the soil parameter  $\alpha$  (eqn. 4-8), the horizontal and vertical grid discretization,  $\Delta x$ ,  $\Delta z$ , and the horizontal and vertical correlation scales of  $f$  (and  $a$ ),  $\lambda_{fx}$ ,  $\lambda_{fz}$ . Only steady-state gravity flow is considered. Each Monte Carlo simulation is designed to simulate one particular, hypothetical soil site. The different sites (simulations) are arbitrarily labeled as #M, where  $M \in \{2,3,\dots,31\}$ . To keep matters simple and transparent, a base soil site is defined (#3). From site to site, one or a few of the input parameters are systematically varied. Only the differences to the base site are listed in Table 8.1.

Recall from chapter 7, that the ASIGNing technique allows the use of random Dirichlet type boundaries by setting the head on the boundaries equal to the spectrally derived solution. The solution for a particular sample of random fields  $f$  and  $a$  is obtained quasi-analytically in the spectral domain (Harter and Yeh, 1993; see also chapter 7) The underlying assumption is that the RFV  $h$  is weakly stationary, normally distributed, and that boundaries are at infinity. The spectral solution for the head is defined through the spectral representations of  $f$  and  $a$ , and through the ensemble means  $\Gamma$  and  $H$ . The mean vertical flux is controlled through  $\Gamma$ ,  $H$ , and the covariance function of  $f$  and  $a$ . This type of boundary condition allows the simulation of a finite portion of a quasi-infinite domain, which is consistent with many field applications. Boundary conditions are rarely known with certainty.

Within the steady-state unsaturated flow profile the spatial variability of the soil water content  $\theta$  is neglected. For simplicity a constant  $\theta=1$  is used in the numerical simulations. Then the pore velocity is equal to the Darcian flux  $\mathbf{q}$  and is automatically normalized with respect to the (constant) soil water content. This greatly simplifies the flux analysis and - in subsequent chapters - will prove advantageous in the analysis of solute transport through unsaturated media, since the results are very general and independent of the relationship between  $\theta$  and  $h$ . Note that this assumption does not neglect the change in water content from a very wet to a very dry soil. It only postulates that the water content spatial variability at a given  $H$  is

negligible. The velocity distribution can also be normalized with respect to the mean saturated hydraulic conductivity, the only other RFV depending on units of time, since the choice of the time-units is arbitrary. Setting the mean saturated hydraulic conductivity to 1[cm/day], the actual steady-state results for the velocity distribution given a mean saturated hydraulic conductivity of  $x$  [cm/day] are obtained by multiplying the velocity with  $x$ .

The mean pressure head in the base soil is -150 cm and varies in other simulations from -100 cm to -3000 cm. The mean pressure head is chosen to avoid partial saturation of significant parts of the soil domain if the soil is very heterogeneous. Partial saturation poses no problem to the Monte Carlo simulation, but cannot be taken into account by the first order stochastic analysis to which the numerical results are compared. The base site has a unit variance of  $f$ ,  $\sigma_f^2=1$ . In other soils,  $\sigma_f^2$  is as small as 0.01 and as large as 4. The geometric mean  $\Gamma$  of  $\alpha$  is  $0.01 \text{ cm}^{-1}$  with a variance  $\sigma_a^2 = 0.01$  such that  $\zeta = \sigma_a/\sigma_f = 0.1$ . All simulations are implemented using the exponential covariance function (7-5) for  $f$  and  $a$ . The RFV  $a$  is either perfectly correlated with  $f$  ( $\rho_{af}=1$ ) or - as in the base soil site - independent of  $f$  ( $\rho_{af}=0$ ). It has a correlation scale that is always identical to that of  $f$ . In the sensitivity analysis the geometric mean of  $\alpha$  is increased to values as large as  $0.1 \text{ cm}^{-1}$  and the variance of  $a$  varies between  $10^{-4}$  and 0.6. The correlation scale is systematically varied from as little as 12.5 cm in both the horizontal and vertical direction to as much as 300 cm in the horizontal and 50 cm in the vertical. The base soil is isotropic with a correlation scale  $\lambda_f = 50 \text{ cm}$ . The discretization of the base soil site yields squared finite elements of  $(10)^2 \text{ cm}^2$  or  $1/5^{\text{th}}$  of the correlation scale in each dimension. The vertical discretization ( $\Delta z=10 \text{ cm}$ ) is chosen according to the results of the grid design analysis in chapter 6. Different finite element discretizations are also selected to test the grid-design with the Monte Carlo approach (as opposed to the single simulation technique used in chapter 6). An isotropic case is simulated with  $\Delta x = \Delta z=2.5 \text{ cm}$  and a correlation scale  $|\lambda_f| = 12.5 \text{ cm}$  in each direction (#19). The effect of horizontal discretization on the stochastic results is tested by comparing an anisotropic case ( $\lambda_{f_x}=300 \text{ cm}$ ,  $\lambda_{f_z} = 50 \text{ cm}$ ) with  $\Delta x = 30 \text{ cm}$  and  $\Delta z=10 \text{ cm}$  (#12) with the same case, but different horizontal discretization

$\Delta x = 10$  cm (#11).

#### 8.4 Random Field Generator: Evaluation

The performance of the random field generator used to generate the two-dimensional input random fields of  $K_s$  and  $\alpha$  has been discussed in chapters 3 and 6. The spectral generator produces numerically undistorted random fields with sample moments that are in excellent agreement with the specified ensemble moments. The only significant weakness of the technique described in chapter 3 is a small reduction in the variance of the random fields: The sample variance is generally 5% lower than specified, while the sample covariance reproduces the desired covariance structure at non-zero lags with very good accuracy. No consistent error is observed for the sample mean. These results from chapter 3 are confirmed by the sample moments obtained for  $f$  and  $a$  in the unsaturated flow simulations. The histograms of  $f$  and  $a$  show a smooth Gaussian-like distribution, and no consistent artificial spatial pattern is observed in the two-dimensional map of the input sample moments (Figure 8.3d,h). In the base soil #3, the local sample mean  $F_{1000}(\mathbf{x})$  of  $f$  varies from -0.1 to 0.1 with a (spatial) standard deviation of the sample mean of 0.033. The expected standard deviation (8-3) of the sample mean is  $\epsilon_F = 0.032$  ( $N=1000$ ) (Figure 8.3a). The local sample variance of  $f$  varies from 0.8 to 1.1 with a spatial average of 0.94, which is 6% below the specified ensemble variance (Figure 8.3b). The spatial standard deviation of the local sample variance is 0.043 which is very close to the expected  $\epsilon_{v,f} = 0.042$  (8-7). Similar results are found for the sample moments of  $a$ . The average sample covariance function (8-8) for  $f$  and  $a$  are isotropic (Figure 8.3c,g) and a cross-section shows that they are in excellent agreement with the specified exponential covariance function (Figure 8.6).

#### 8.5 Stochastic Analysis of the Unsaturated Hydraulic Conductivity

The stochastic analysis of the dependent variables  $y$ ,  $h$ ,  $v_x$ , and  $v_z$  is organized in the following manner: Using primarily the results from the base soil site #3, general observations are summarized regarding the structure of the random fields, the structure, stationarity, and sampling error of the sample mean and sample variance fields, and the structure of the covariance fields. The covariance fields are qualitatively compared with analytically obtained covariance functions (chapter 4). The histograms are described to draw conclusions about the empirical pdf of the dependent parameters. Then a quantitative analysis is implemented regarding the stochastic dependence of the mean, the variance, and the covariance of the dependent RFVs on the variances  $\sigma_f^2$ ,  $\sigma_a^2$ , the correlation  $\rho_{af}$  between  $a$  and  $f$ , the mean head  $H$ , the geometric mean  $\Gamma$  of  $\alpha$ , the horizontal and vertical correlation scales  $\lambda_{fx}$  and  $\lambda_{fz}$  of  $f$ , the anisotropy aspect ratio  $v = \lambda_{fx}/\lambda_{fz}$ , and the grid discretization  $\Delta x$  and  $\Delta z$ . The numerical results are directly compared with the first order analytical solutions.

### 8.5.1 General Observations

*Sample mean and sample variance field.* At site #3, the sample mean and variance fields of  $y$  have the random character of individual realizations (Figure 8.3i,k). The dimensionless errors (8-10) and (8-11) of the sample mean and sample variance are 1.21 and 0.99, respectively. Almost identical sampling errors are observed for other soils with the same or less variability in  $y$ . At the anisotropic soil sites, the dimensionless error of the sample mean reduces to 0.82 due to the smaller size of the simulation domain relative to the correlation scale of  $f$  (see Table 8.1). No significant trend or other artificial spatial features indicates a deviation from second order stationarity. Second order stationarity for  $y$  and other dependent RFVs is expected since the random head boundary conditions are weakly stationary and the mean vertical flow therefore uniform. Notice that the sample mean and variance of  $y$  on the boundary are not notably different from the interior of the simulation domain. However, for soils with larger aspect ratio  $v$  than the base soil site #3 or larger variability  $\sigma_y^2$  or both, the variance within  $0.5\lambda_{fz}$  from the

bottom boundary and within  $2\lambda_{fz}$  from the top boundary tends to be lower than in the rest of the domain by up to approximately 30%. A similar variance reduction is observed within  $0.2\lambda_{fx}$  of the horizontal boundaries. These boundary effects on the sample variance of  $y$  increase  $\epsilon'_{v,g}$  to 1.26 in the highest variance soil (#21) while reducing  $\epsilon'_G$  to 0.67. No significant boundary effects are observed for the sample mean  $Y$  of  $y$  at any soil site.

The average reduction of  $\sigma_y^2$  at and near the boundary is due to setting the head values on the boundary equal to the first order approximation of the head, given the random fields of  $f$  and  $a$ . The statistical moments of the head on the boundary are therefore not entirely consistent with those in the interior of the domain. As will be discussed below, the head and velocity variances near the boundary increase significantly for soils with high variability in moisture flux due to the approximate nature of the first order quasi-analytical head boundary conditions. It is not clear, however, why there is a reduction and not an increase in  $\sigma_y^2$  near the boundary (relative to the interior).

*Sample covariance.* For all soils the sample covariance field  $C_{yy}$  of  $y$  is very similar to the input covariance field  $C_{ff}$  reflecting the physical observation that the random fields of  $y$  have a very similar random structure as the random fields of  $f$  and  $a$  (compare e.g., Figure 8.3i,k vs. Figure 8.3a,b). At the isotropic soil sites (isotropic with respect to  $f$ ),  $C_{yy}$  has a very small, but notable anisotropy with larger vertical than horizontal correlation scale (Figure 8.6). The correlation lengths in both directions are approximately 10% smaller than  $\lambda_f$ . The horizontal covariance is of the "hole-covariance" type i.e., it becomes negative at some lag distance and then tends asymptotically to zero. The vertical covariance of  $y$  remains positive for all lag-distances. In the anisotropic soils with  $\nu > 1$ , the situation reverses to a hole-type covariance function in the vertical direction and an exponential type covariance function in the horizontal direction. Again, the correlation lengths are approximately 10% smaller than those for  $f$  (Figure 8.7). These findings are in excellent agreement with the theoretical covariance function derived in chapter 4. In Figure 8.5d a single sample covariance  $C_{yy}$  is plotted for an anisotropic wet soil with  $\nu=3$  and  $\sigma_f^2=0.95$  (#31) to illustrate the qualitative agreement between

the analytical solution and the numerical results. Similar qualitative agreement of the sample  $C_{yy}$  with the analytical  $C_{yy}$  is found at all sites. The correlation function  $\rho_{yy}$  is also in good quantitative agreement for all soil sites, even those with strong variability (compare Figure 8.6, a mildly heterogeneous, isotropic soil, with Figure 8.7, a strongly heterogeneous, anisotropic soil).

*Histogram.* The histograms of  $y$  at all soil sites indicate that  $y$  is Gaussian-like distributed i.e., the unsaturated hydraulic conductivity seems lognormally distributed (Figure 8.3m). Only at the driest soil site (#21), which is also the soil with the highest variability in  $y$ , the histogram has a slight tail towards lower  $y$ . No distribution tests were implemented.

### 8.5.2 Moment Analysis of the Unsaturated Hydraulic Conductivity

For the stochastic analysis, only the spatially averaged sample moments are considered. To eliminate the non-stationary effects near the boundary, the average sample mean and variance of each dependent RFV are obtained by averaging over the center 33 by 33 nodes ( $h, v_x, v_z$ ) or the center 32 by 32 elements ( $y$ ).  $y$  is computed for each element from Gardner's equation (4-8) by arithmetically averaging the head values on the four nodes surrounding the element.

Since the correlation functions of  $f$  and  $a$  are identical within each soil site, inspection of 4-31, 4-39, and 4-47 in chapter 4 suggests that the variances of all dependent parameters of interest -  $h, y, v_x$ , and  $v_z$  - can be normalized by the variance factor  $\sigma^2$ :

$$\sigma^2 = \sigma_f^2 (1 + \rho \zeta \Gamma H + (\zeta \Gamma H)^2) \quad (8-13)$$

Note that  $\sigma^2$  is not identical with either the saturated conductivity variance  $\sigma_f^2$  nor with the unsaturated conductivity variance  $\sigma_y^2$ . The dimensionless unsaturated hydraulic conductivity variance  $\sigma_y'^2$ :

$$\sigma_y'^2 = \frac{\sigma_y^2}{\sigma^2} \quad (8-14)$$

as well as the dimensionless variances of the other dependent RFVs are in first order independent of the mean soil water tension  $H$ , the correlation  $\rho_{af}$ , and the variances of  $f$  and  $a$ . The analytical, dimensionless variances and covariance functions of all RFVs including  $y$  are therefore only functions of the correlation scales of  $f$  and the geometric mean  $\Gamma$  of  $\alpha$ . The numerical analysis shows, however, that the actual stochastic relationship between the dependent RFV second moments and the independent RFV pdfs is more complex than suggested by the analysis in chapter 4. The following results will illustrate this for the unsaturated hydraulic conductivity variance. The stochastic analysis of other RFVs is given in subsequent sections.

*Dependence on input variance.* The average sample mean  $Y$  of the log unsaturated hydraulic conductivity changes proportional to  $H$  such that for all sites the first order approximation of  $Y$  (4-35),  $Y=F+H\Gamma$ , holds very accurately (deviations of less than 1%). Figure 8.8b shows the normalized soil variances  $\sigma_y^2$  as a function of the input variance  $\sigma_f^2$ , aspect ratio  $v$ , and vertical correlation scale  $\lambda_{fz}$ . All soils have the same mean tension head  $H = -150$  cm. The random fields of  $f$  and  $a$  are independent. The variance ratio  $\zeta = \sigma_a / \sigma_f$  is 0.1. It is obvious from Figure 8.8b that the results are not quite independent of the actual magnitude of the variances in  $f$  and  $a$ . For the three least variable soil sites ( $\sigma_f^2 = 0.01, 0.11, 0.95$ ), the actual  $\sigma_y^2$  is approximately 4% smaller than the first order results indicating (as expected) a very good agreement between the numerical and the analytical results (Table 8.2) considering that the variance of the input random fields is also approximately 4%-5% smaller than specified.

With increasing  $\sigma_f^2$ , the dimensionless  $\sigma_y^2$  increases more or less linearly. In the anisotropic soils the increase relative to the analytical solution is larger. At  $\sigma_f^2 = 3.6$ , the numerical  $\sigma_y^2$  is 4% larger than the analytical  $\sigma_y^2$  in the isotropic soil and 10% and 16% larger in the anisotropic soils with  $\lambda_{fz} = 50$  cm and 30 cm, respectively. A careful analysis of these results reveals that the differences partly stem from an increasing difference in the sample mean head (used for the normalization (8-14)) at higher variances. The difference between the actual (dimensional) and analytical  $\sigma_y^2$  does not exceed 11%, even for the most variable soils ( $\sigma_f^2 = 3.2$ ) including the dry sites that are not shown here.

The dependence on  $\sigma_a^2$  alone is demonstrated by comparing the isotropic base site ( $\sigma_a^2=0.01$ , #3) with a soil having  $\sigma_a^2=0.64$  (#26). In the latter soil the difference between the actual and analytical  $\sigma_y^2$  (dimensional) is 16%.

Overall, the MCSs indicate that the first order analytical estimate of the mean  $Y$  of the unsaturated hydraulic conductivity is very accurate even for strongly heterogeneous soils. The analytical solutions underestimate  $\sigma_y^2$  at large variances of  $f$  and  $a$ . For practical purposes, the 10%-16% error of the analytical solution in very heterogeneous flow fields (large  $\sigma_y^2$ ) is negligible.

*Dependence on soil water tension and the correlation between  $f$  and  $a$ .* Again, the first order approximation of  $Y$  gives very accurate predictions (to within 1%) of the observed  $Y$ . Apart from the differences between analytical and numerical solutions for  $\sigma_y^2$  discussed in the previous paragraph, neither the mean head, nor the correlation coefficient  $\rho_{af}$  have a remarkable effect on  $\sigma_y^2$ . Due to the particular form of the variance factor  $\sigma^2$ , the variance of all RFVs goes to 0 for  $\rho_{af}=1$  as  $H \rightarrow -1/\zeta\Gamma$ . At soil water tensions that are more negative than this limit, the variances of all RFVs increase again (see also chapter 9). It is found that these results indeed hold for the numerical simulation. In correlated soils, the first order solution for the moments of  $y$  is accurate over a larger range of soil tensions than in uncorrelated soils. Only at a very dry head ( $H=-3000$ ) with a large  $\sigma_y^2$ , the dimensionless  $\sigma_y^2$  increases relative to the analytical solution (Figure 8.9b).

*Dependence on  $\Gamma$  and the correlation scale of the soil.* The second moments of the dependent RFVs depend nonlinearly on  $\Gamma$  and  $\lambda$  as shown in Figure 8.10b and are also found to be in good agreement with the first order analytical solution, even for large  $\Gamma = 0.1 \text{ cm}^{-1}$ .

*Dependence on aspect ratio and grid discretization.* The difference between the numerical and analytical  $\sigma_y^2$  increases for larger aspect ratio  $v$  and longer vertical correlation scale  $\lambda_{vz}$ . Figure 8.11b and shows that overall the variance of  $y$  decreases as the aspect ratio increases, as expected from the first order analysis. Different horizontal element discretization (#11 vs. #12) does not influence the results for  $y$ .

## 8.6 Stochastic Analysis of the Soil Water Tension

### 8.6.1 General Observations

*Sample mean and sample variance field.* The sample mean and variance fields of the soil water tension have a very different random character compared to  $y$ ,  $f$ , or  $a$  (compare Figure 8.3i,k with Figure 8.4a,b): The visual patterns are much less erratic and significantly broader with only a few relatively large areas of randomly high and low sample values. This pattern is a reflection of the much less erratic nature of the underlying realizations of  $h$ , which exhibit a similarly smooth pattern (see Figures 7-2 through 7-4 in chapter 7).

The dimensionless error of sample mean and sample variance,  $\epsilon'_G{}^2$  and  $\epsilon'_{v.g}$ , are 0.66 and 1.09, respectively at the isotropic base soil site (#3). Recall that the dimensionless error reflects the spatial variability of the sample mean and sample variance within the simulation domain relative to the expected variability in an infinite domain. The low error of the sample mean is caused by the strong correlation of the head sample mean values within the simulated domain; the sample domain is small relative to  $\lambda_h$  (approximately  $3 \lambda_{hx}$  and  $5 \lambda_{hz}$ ). In the anisotropic soils (anisotropy in  $f$ ),  $\epsilon'_G{}^2$  becomes even smaller ranging from 0.22 in the less variable soils (e.g. #12) to 0.46 in the most variable soils (e.g. #22). This significant reduction in the sample error comes despite the fact that e.g., for the anisotropic soils with  $v=6$ , the correlation scales are  $\lambda'_{hx} \approx 2$  and  $\lambda'_{hz} \approx 3$ , which means that the relative domain size (measured in  $\lambda_h$ ) remains approximately the same as in the isotropic soils. The low  $\epsilon'_G{}^2$  indicates that the sample error associated with the *average* sample mean approaches that of the *local* sample mean. The variance sample error  $\epsilon'_{v.g}$  is approximately 1. In the anisotropic soils of moderate variability it generally is within 5% of 1, and increases to 1.1 in the isotropic soils. In the strongly variable soils  $\epsilon'_{v.g}$  increases up to 1.4.

Boundary effects are insignificant at the base soil site (Figure 8.4a,b). But in more heterogeneous soils and particularly in soils of stronger horizontal anisotropy, the variance increases by up to 30% in a boundary region that is  $\lambda_{tz}=2$  thick near the horizontal boundaries,

but only  $\lambda_{fx}=0.2$  wide near the vertical boundaries. In other words, the boundary effect is particularly dominant into the direction of mean flow. The variance increase is due to the larger variance in the first order head perturbation solution on the boundary (see discussion below). The artificial impact of the first order random head boundary in the anisotropic and strongly heterogeneous soils is - spatially - much less dominant than constant head or flux boundary conditions, which have traditionally been used in Monte Carlo simulations. For the saturated case, Rubin and Dagan (1988, 1989) estimate that the boundary effects of such non-random boundaries vanishes only at a distance of at least  $1\lambda_r$  to  $2\lambda_r$  from the boundary.

*Covariance sample field.* The "smoothness" of the realizations of  $h$  and of the pattern in the sample mean and variance fields is quantitatively captured by the covariance function, which has a much larger correlation scale than  $f$  in both the horizontal and vertical direction (Figure 8.4c). Even for the isotropic base soil #3, the head covariance  $C_{hh}$  is anisotropic with  $\lambda'_{hx} = \lambda_{hx}/\lambda_{fx} \approx 4.5$  and  $\lambda'_{hz} = \lambda_{hz}/\lambda_{fz} \approx 2.5$ . The anisotropy is reflected in the horizontally elongated pattern structure of the sample mean and variance fields in Figure 8.4a,b. The numerically obtained covariance function is well predicted by the analytical covariance function for the head  $C_{hh}$  (Figure 8.5a,b). The covariance function is similar to an anisotropic Gaussian function, particularly near the origin, which explains the smoothness of the random head fields (Figure 8.6). For larger lag distances both the analytical and numerical covariance fields deviate from the oval shape of the Gaussian covariance. In the vertical direction, the head covariance is a "hole"-type function (see discussion of  $C_{yy}$ ) regardless of the type of soil investigated here. The limitation of the domain size does not allow an assessment of the type of covariance function in the horizontal direction. The differences between numerical and analytical covariance functions are primarily due to sampling variability and due to the different variances. The normalized head correlation functions  $\rho_{hh}$  from the first order analysis and the Monte Carlo analysis are in good agreement not only for mildly heterogeneous soils ( $\sigma_y^2=0.1$ , #8, see Figure 8.6), but also for highly heterogeneous soils ( $\sigma_y^2=3.2$ , #22, Figure 8.7). For strongly heterogeneous soils of any anisotropy ratio, the analytical correlation function tends

to underestimate the vertical correlation of the Monte Carlo results. In anisotropic soils this is also true for the head horizontal correlation.

*Histograms.* The histogram for the total of sample head values is not significantly different from the histogram for the head values sampled at the center of the simulation domain. Figure 8.4d shows the total sample histogram for the base soil site. At this and most other soil sites, the histogram follows the symmetric Gaussian pdf, which confirms a basic assumption of previous analytical studies of flow in heterogeneous soils (Yeh et al., 1985a,b,c; Mantoglou et al., 1987a,b,c). Only in the most heterogeneous soils (those with the highest  $\sigma_h^2$ ) and in soils with a large  $\Gamma$  (=0.1 in #30) an almost negligible but consistently notable tail towards more negative head values develops.

### 8.6.2 Moment Analysis of the Soil Water Tension

*Dependence on input variance of  $f$  and  $a$ .* In the least variable soil (isotropic soil site #2 with  $\sigma_f^2=0.01$ ) the average mean head deviates less than 0.1% from the mean head (-150 cm) prescribed for the first order perturbation solution on the boundary. As the variance of  $f$  and  $a$  increase, the mean head drops slightly to -150.9 cm in the isotropic base soil site with  $\sigma_f^2=1$  (#3) and to -156.3 cm in one of the most variable, wet anisotropic soil sites with  $\sigma_f^2=4$  (#22). In drier soils the difference between prescribed and average measured mean head does not exceed 1%, even if the head variance is very large (e.g. #21). The actual sample mean head is not sensitive to any of the other input parameters. The rest of this section will therefore only discuss the dependence of the sample head variance on the various input statistics.

For the least variable three isotropic soil sites, the results of the head variance are shown in Table 8.2 together with the analytical head variance solution. For those three soils the largest difference between numerical and analytical head variance is observed for the least variable soil. There, the variance in the Monte Carlo is 4% higher than the analytical variance  $\sigma_h^2$ . At  $\sigma_f^2=1$  (#3) the difference is reduced to 2%. For practical purposes, the first order analytical solution

is considered accurate for isotropic soils of  $\sigma_f^2 \leq 1$ .

The normalized head variance  $\sigma_h'^2$  is:

$$\sigma_h'^2 = \frac{\sigma_h^2}{\sigma^2 \lambda_{fz}^2} \quad (8-15)$$

where the variance factor  $\sigma^2$  is defined in (8-13). Figure 8.8a shows that the numerical head variance in the isotropic soils is well modeled by the analytical head variance even at  $\sigma_f^2=3.6$ . In the anisotropic soils, the head variance is also very accurately predicted from theory for  $\sigma_f^2=1$ , but decreases linearly (relative to the first order solution) at higher variances. The decrease is the strongest for those soils with the highest aspect ratio, such that for a wet soil with  $v=6$  and  $6.7$ , and with  $\sigma_f^2=3.6$  the head variance in the Monte Carlo simulation is only 75% of the analytically obtained variance. This is consistent with the boundary effects observed particularly in the anisotropic soils.

The robustness of the first order head perturbation solution at variances up to  $\sigma_f^2=1$  has previously been discussed in the literature, but with respect to saturated groundwater flow. Dagan (1985) computed second order corrections for the head moments in an infinite aquifer and found that for  $\sigma_f^2=1$  first order head moments are within 10% of the second order head moments. Gutjahr and Gelhar (1981) concluded from their analysis that the spectral first order approximation of the head moments in saturated porous media is valid even for variances  $\sigma_f^2$  much larger than unity. As Gelhar (1986) noticed, no such evaluation has been made for the spectral analysis of unsaturated flow. While this study does not address the issue analytically, the simulations clearly indicate the general trend: First order analysis will significantly overestimate  $\sigma_h'^2$  at large  $\sigma_f^2$  and  $\sigma_a^2$ . It may be argued that the difference is due to the small simulation domain. However, the average sample variance as well as the sample variance in the center of the simulation domain change insignificantly, when the vertical and horizontal domain size is increased to 150 nodes and 100 nodes, respectively.

*Dependency on mean soil water tension.* In anisotropic soils the numerically obtained

$\sigma_h'^2$  decreases as the soil becomes drier and as  $\sigma_y^2$  increases (Figure 8.9a). For the driest soil (#21:  $H=-3000$  cm,  $v=6$ ,  $\sigma_y^2=3.2$ ) the decrease is approximately 25% i.e., almost the same as for the highly heterogeneous wet soils (#22:  $H=-250$  cm,  $v=6$ ,  $\sigma_y^2=3.2$ ). The overall effect of mean soil water tension on the actual variance of the head depends on  $\rho_{af}$  as discussed in the analysis of  $y$ . For mean head much more negative than the critical head  $H=-1/\zeta\Gamma$ , the head variance will increase in any soil but not as strong as suggested by the form of the variance factor  $\sigma^2$  (8-13).

*Dependency on  $\Gamma$  and vertical correlation scale.* A similar deviation from the first order results is not observed for varying vertical correlation scale or increased coarseness of the soil texture (larger average pore size distribution parameter  $\Gamma$ ), if  $\sigma_f^2$  does not significantly exceed unity. As Figure 8.10a shows, there is excellent agreement between first order analysis and Monte Carlo analysis.

*Dependency on aspect ratio and grid discretization.* The effects of aspect ratio are also well modelled by the first order approximation (Figure 8.11a) if  $\sigma_f^2=1$ . As the aspect ratio  $v$  increases,  $\sigma_h'^2$  decreases slightly relative to the analytical solution. A threefold increase in the horizontal element size (#12 vs. #11) increases the head variance slightly (about 2%).

## 8.7 Stochastic Analysis of the Velocity

### 8.7.1 General Observations

*Sample mean and variance fields.* The random structure of the horizontal and vertical velocity fields are very peculiar and distinctly different from those of other RFVs. In Figure 8.12, a single realization of corresponding  $v_x$  and  $v_z$  fields in the base soil site (#3) are shown. The horizontal velocity map has a distinctly symmetric pattern of diagonally trending narrow stripes with strong negative velocities (dark NE-SW trending "canyons") and counter-diagonally trending narrow stripes with strong positive velocities (white NW-SE trending "cloud streamers") in an otherwise relatively homogeneous velocity field with very small horizontal

velocities. The vertical velocity map, in contrast, is not quite unlike the map of a braided channel network in a river valley or of the preferential flow paths that have been reported to occur in soils: While most portions of the soil domain have relatively small vertical velocities (lightly colored areas), high velocities (dark areas) form a braided network of narrow channels with a predominantly vertical direction. The diagonal streaks of  $v_x$  and the vertical braided channels of  $v_z$  reappear in the sample mean and sample variance maps (Figure 8.4e,f,i,k), but in a much more vivid, livelier, more interwoven, and more erratic manner. Graphically speaking, the laziness of the landscape in the individual realizations is replaced by a vivid pattern in the sample mean and variance fields. This is in contrast to the observations for the sample mean and variance fields of  $f$ ,  $a$ ,  $y$ , and  $h$ , which are all very similar in character to individual realizations.

Individual realizations are relatively homogeneous over large areas (the laziness) with extreme values (the canyons and clouds and river channels) interrupting them at a spatial interval that is on average significantly longer than the transverse size of the channels and streaks. In the sample mean and variance fields the diagonal and cross-diagonal streakline pattern and the vertical braided channel pattern are preserved, but the transverse extent of the channels and streaks is narrower, the frequency of streaks and channels has increased, and they are much less continuous.

What is the explanation for the particular pattern of the individual realizations and the character of the sample mean and variance fields? And what is its significance? Before analyzing the statistical description of the velocity RFV, it is important to pursue these questions to better understand the physical nature of moisture movement in heterogeneous soils.

The horizontal and vertical velocity realizations depicted in Figure 8.12 must be seen as a unity since they are two components of a single vector  $\mathbf{v}$ . The areas of very large positive and negative horizontal velocity occur in those parts where the vertical velocity is also large (indicated by the channels) but where the channels are inclined relative to the vertical axis. In the isotropic soil of Figure 8.12, the horizontal component of the velocity is rarely much larger

than the vertical component, therefore the diagonal orientation of the streaks (instead of a horizontal or near horizontal orientation). Since most of the flow is vertical, the vertical velocity map can be seen as almost representative of  $\mathbf{v}$ . The velocity map in Figure 8.12b indicates that soil moisture movement is spatially variable but statistically homogeneous (chapter 2) soils tends to be along preferential flow paths i.e., the majority of soil moisture moves through only a small portion of the entire soil domain. In a large part of the soil domain, moisture flux is relatively small. The simulations show that the concentration of moisture flux into small channels increases as the heterogeneity of the soil increases or as the soil becomes drier. Similar patterns of flow channeling are shown by Moreno et al. (1989) who modeled Darcian flow in a two-dimensional, single fracture with varying aperture and high variability of fracture resistance (which is inversely related to the conductivity). Channeling has also been observed in field soils, where channeling due to soil heterogeneity and channeling due to wetting front instability (fingering) together may greatly enhance the variability of the flux field (Glass et al., 1988).

From these physical observations it is expected that both the vertical and horizontal velocity distribution have a non-Gaussian, highly skewed distribution. As discussed in more detail below, the velocity components are indeed non-Gaussian, lognormal-like distributed (Figure 8.13). Hence the usefulness of the first and second moments as measures of the pdf of  $\mathbf{v}$  is limited. Since the sample mean and variance are obtained through arithmetic averaging, the large velocity areas of individual realizations carry much weight in the sample mean and variance. The maps of the two sample moments become like a collection of the many streaks and channels of the individual realizations.

*Boundary effects.* The sample mean and sample variance fields of the velocity have boundary effects that quantitatively are very strong, even at the base soil site, where no significant boundary effects are observed for other RFVs (Figure 8.4e,f,i,k). Directly on the boundary, extremely low and high values occur in the sample mean of both velocity components. The sample variance of the velocity is much higher at the boundary than in the

interior of the domain: at the base site (#3) by up to an order of magnitude for  $\sigma_{vx}^2$  and a factor 5 for  $\sigma_{vz}^2$ . The spatial extension of the boundary effects are particularly strong for  $\sigma_{vx}^2$  into the mean flow direction: along the horizontal boundaries they are significant within  $1\lambda_{fz}$  (Figure 8.14a,b). Along the vertical boundaries and for all boundaries around the  $\sigma_{vz}^2$  map the effect is limited to  $0.5\lambda_{fz}$  and less. The boundary effect has a spatially larger extent in the anisotropic soils of equivalent heterogeneity, which increases with the variability of  $y$ . If  $\sigma_y^2 \approx 3$ , the boundary effect significantly increases  $\sigma_{vx}^2$  and  $\sigma_{vz}^2$  within almost  $3\lambda_{fz}$  from the bottom boundary and within almost  $2\lambda_{fz}$  from the top boundary (Figure 8.14e,f). For this soil the velocity variance on the boundaries is three orders of magnitude larger than in the interior of the domain. The spatial extent of the boundary effect on the velocity moments is slightly smaller than the spatial extent of constant flux boundary conditions in saturated flow (compare Figure 8.14 to Bellin et al., 1992, Figure 3).

The very strong though spatially limited boundary effects on  $\sigma_v^2$  are again caused by the approximate nature of the first order head perturbation solution used as Dirichlet boundary conditions in the numerical model. But while the statistical moments of the first order head approximation are in excellent agreement with those of the Monte Carlo simulation for  $\sigma_f^2 = 1$ , the velocities derived from the first order perturbation head distribution along the boundary are extremely erratic. A simple method to circumvent the erroneous boundary effects in transport simulations is described in chapter 9.

*Covariance fields.* The covariance fields for  $v_x$  and  $v_z$  are a reflection of the diagonal and counter-diagonal patterns, and of the braided vertical patterns, respectively, of the high velocity areas (Figure 8.4g,l). For increasing aspect ratios, the two diagonal main axes of the horizontal covariance function become flatter (Figure 8.5e,f) indicating that the diagonal flow patterns observed in the isotropic soil (Figure 8.4g) tend to become more horizontal as the aspect ratio increases. For mildly to moderately heterogeneous soils such as the base soil (#3), the analytical covariance functions obtained from the first order spectral density functions for the two velocity components are in very good qualitative, if not quantitative agreement with the

numerically derived velocity covariance functions (Figures 8.5 and 8.6).

In soils (wet or dry) with  $\sigma_y^2 > 1$ , the vertical and horizontal cross-sections of the analytical velocity correlation functions deviate significantly from the numerically determined solutions. The numerically obtained horizontal velocity correlation function is almost identical in the transverse (horizontal) and the longitudinal (vertical) direction, while first order analysis predicts a much shorter transverse correlation scale and a much larger longitudinal correlation scale. The numerical vertical velocity correlation function has also a much shorter longitudinal correlation scale than the analytical correlation function. Only the transverse hole-type correlation of  $v_z$  is very accurately predicted for all soils (compare Figures 8.6 and 8.7).

*Histograms.* The histogram for  $v_z$  is - as expected - skewed (see, for example, Figure 8.4m). The velocity statistics are obtained on the untransformed RFVs. Logarithms were not taken during the simulation due to the fact that neither the vertical nor the horizontal velocity component is restricted to either positive or negative values only. A graphical method to investigate whether the vertical velocity is indeed lognormal-like distributed consists of a plot of the histogram on a lognormal axis (Figure 8.13). For very small input variances of  $f$  ( $\sigma_f^2 = 0.01$ ) the histogram of  $v_z$  can either be interpreted as normal or as lognormal (Figure 8.15), but at higher variances, the histogram is always skewed on the arithmetic scale, even for mildly heterogeneous soils (e.g.  $\sigma_f^2 = 0.11$ , #8, see Figure 8.15). This result is in accordance with the histograms obtained from Monte Carlo simulations of saturated flow in two- and three-dimensional heterogeneous media by Bellin et al. (1992) and Levin (1994). Figure 8.16 shows the total histogram as well as the histogram of the center point of the simulation domain for  $v_x$  and  $v_z$ . The two types of histograms are generally identical. The histogram for the center point is of course based on only  $N=1000$  values, while the total histogram is based on 4 million data ( $64^2 N$ ). Even the histogram plotted on the logarithmic axis has a significant skewness if the soils are dry or very heterogeneous (Figure 8.16h,i). In the soils with very high variability, a significant amount of vertical velocities is positive (upward) (Figure 8.16h,i,k).

The horizontal velocity histogram plotted on an arithmetic scale (Figure 8.4h) seems

to be Laplacian (symmetric exponential decay, chapter 2), but plotting the histograms of  $|v_x|$  on the logarithmic scale reveals that the pdf for  $v_x$  must be differentiable for  $|v_x| \rightarrow 0$  (Figure 8.16a-e). The histogram of  $v_x$  resembles a Gaussian function only if the unsaturated hydraulic conductivity variance is very small ( $\sigma_y^2 < 0.2$ ). This is in contrast to the findings of Bellin et al. (1992) and Levin (1994) who argue that the horizontal velocity component in their simulations has a normal pdf even in very heterogeneous saturated porous media. However, visual inspection of the numerical velocity cdf and the Gaussian cdf in Figure 7d of Bellin et al. (1992) indicates that their transverse velocity pdf qualitatively tends away from the Gaussian pdf towards a similar shape shown for the unsaturated velocity pdf e.g. in Figure 8.4h!

### 8.7.2 Moment Analysis of the Velocity

*Dependence on input variance of  $f$  and  $a$ .* Due to the mean vertical, uniform flux, the mean horizontal velocity must be 0. In all simulated soils, the average sample mean horizontal velocity  $V_x$  is at least three orders of magnitude smaller than the mean vertical  $V_z$  and can therefore indeed be considered as being negligible. The first order analytical mean  $V_z$  is equal to:

$$V_z = \frac{K_m}{\theta} \quad (8-16)$$

where  $K_m = \exp(Y)$  is the geometric mean of the unsaturated hydraulic conductivity. The first order analysis, of course, assumes that both the vertical and horizontal velocities have a normal distribution. Nevertheless, the difference between analytical and average sample  $V_z$  in the isotropic soils with  $\sigma_f^2 \leq 1$  is 2% at the most (Table 8.2). For the most heterogeneous soils (#9) the Monte Carlo  $V_z$  is 10% larger than (8-16). In contrast, the average sample  $V_z$  in the anisotropic, wet soils with  $v=6.0$  and  $v=6.7$ ,  $\sigma_f^2=1$ , is more than 20% smaller than the analytical  $V_z$  and decreases to less than 50% of (8-16). The decrease in the average arithmetic sample mean velocity relative to the analytical mean velocity must be explained with the neglect of

higher order moments in (8-16) and with the lognormal distribution of  $v_z$ , which yields a preferential flow pattern as the variance increases, particularly in anisotropic soils. The numerical results show that the average steady state flux in highly heterogeneous soils depends strongly on the aspect ratio.

For a better physical explanation,  $v_z$  must be analytically evaluated to higher order than in chapter 4. Yeh et al. (1985b) used a mixed first and second order approach to determine the effective hydraulic conductivity  $K_e$  in a vertically uniform flow field. The effective conductivity is defined as  $K_e = \langle q_z \rangle / J_z$ , where  $\langle q_z \rangle$  is the mean vertical flux and  $J_z$  is the mean vertical total potential gradient. All of the simulations here preserve the mean unit gradient condition. Since the soil water content in all simulations is constant and identical to 1, the average sample mean  $V_z$  from the MCS becomes the effective hydraulic conductivity. As Figure 2 in Yeh et al. (1985) indicates,  $K_e/K_m$  is expected to be slightly larger than 1 in isotropic soils, but only about 0.5 for  $v=10$  ( $\Gamma\lambda_{fz} = 0.5$ ). Qualitatively and quantitatively, their findings are therefore confirmed by the numerical simulations.

The average sample velocity variances  $\sigma_{vx}^2$  and  $\sigma_{vz}^2$  differ by 0% and -3% from the analytical solutions for the least variable soil (#2,  $\sigma_f^2=0.01$ ). In isotropic, wet soils with  $\sigma_f^2=0.1$  and 1 (#8, #3), the differences of the numerical to the analytical solutions are of similar magnitude (Table 8.2). Again a more rigorous analysis can be performed by using the dimensionless variances of  $v_x$  and  $v_z$ , which are defined by:

$$\begin{aligned} \sigma_{vx}^2 / 2 &= \frac{\sigma_{vx}^2}{\sigma^2 K_m^2} \\ \sigma_{vz}^2 &= \frac{\sigma_{vz}^2}{\sigma^2 K_m^2} \end{aligned} \tag{8-17}$$

where the variance factor is defined in (8-13). The numerical  $\sigma_{vz}^2$  and  $\sigma_{vx}^2$  are plotted in Figure 8.8c,d for wet soils with three different anisotropy ratios. The Monte Carlo  $\sigma_{vx}^2$  increases significantly in all soils as  $\sigma_f^2$  increases. In the most heterogeneous soils, the average sample  $\sigma_{vx}^2$  exceeds the analytically predicted by a factor of 5 in the two anisotropic

soils and by a factor of 2.5 in the isotropic soil. The Monte Carlo  $\sigma'_{vz}{}^2$  in the isotropic strongly variable soil is also larger than predicted (factor of 1.5), while it decreases with the variance of  $f$  in the anisotropic soil. The results are difficult to interpret, since the RFVs are not normally distributed. But they clearly show the limitations of the perturbation approach with respect to the velocities.

*Dependence on mean soil water tension.* The stochastic dependence on the mean head found in the MCS deviates from the analytical results in a similar way, if both the previous and these results are stated not in terms of  $\sigma_f^2$  and  $H$ , but in terms of  $\sigma_y^2$ : The deviations become stronger as  $\sigma_y^2$  increases, which may be due to either a larger soil textural variability or a drier soil. The magnitude and direction of the deviations are independent of whether the higher  $\sigma_y^2$  is due to large negative  $H$  or due to high  $\sigma_f^2$  (Figure 8.9c,d). Recall that for  $\rho=1$ , the variance of the RFVs theoretically decreases to 0 at  $H=-1000$  cm.

*Dependence on  $\Gamma$  and the vertical correlation scale.* The average sample velocity variance follows a similar stochastic function as the theoretical curve but decreases not as quickly with increasing  $\Gamma\lambda_{fz}$  as predicted by 1st order analysis (Figure 8.10c,d).

*Dependence on aspect ratio and grid discretization.* The velocity variance decreases with increasing aspect ratio, just as the mean vertical velocity decreases. At  $\sigma_f^2=1$ , the influence of the horizontal correlation scale on the accuracy of the analytical solution is negligible. Grid discretization has no significant impact on the solution (Figure 8.11c,d).

## 8.8 Stochastic Analysis of the Cross-Covariance Functions

Cross-covariances are of interest for various reasons. First, many analytical stochastic models of unsaturated or general porous media flow and transport rely on first-order analytical formulations of the cross-covariances (e.g. Dagan, 1984, 1987; Yeh et al., 1985a,b; Mantoglou et al., 1987a,b; Rubin, 1990; Cvetkovic et al., 1992; Russo, 1993a,b). Second, the cross-covariance is necessary for the implementation of conditional simulation, which will be

described in chapter 10. The cross-covariances of common interest are  $C_{fh}$ ,  $C_{f,vx}$ , and  $C_{f,vz}$ . Figure 8.17 shows an anisotropic example (#31) of single sample fields (not the average covariance field!) of each of those cross-covariances. The cross-covariances have features, which are only partially reflected in their respective horizontal and vertical cross-sections. By inspection of Figure 8.17 it can be seen that the horizontal and vertical cross-sections of  $C_{f,vx}$  and  $C_{h,vx}$ , for example, would have little information content if taken horizontally or vertically through the origin. Unlike the covariance fields, the cross-covariances are neither symmetric with respect to the origin, nor symmetric with respect to the major coordinate axes. The complex structure of the cross-covariance functions will make it difficult to define such cross-covariances from field measurements, unless a large number of samples are taken throughout the area of interest.

Of practical interest is the fact that all cross-covariances except  $C_{f,vx}$  and  $C_{h,vx}$  are much stronger in the vertical direction than in the horizontal direction. Hence the information content of one variable with respect to another variable is predominant within the same vertical region but bears less predictive capacity with respect to other variables in the same horizontal region. Another important feature to be noticed is the non-zero lag-distance at which the highest absolute cross-correlation is reached. Also, the correlation can be either negative or positive. For example, head values have a positive correlation to  $f$  values that are approximately  $3\lambda_f$  further upward and a strong negative correlation to  $f$  values that are approximately  $1\lambda_f$  downward from the location of the head measurement. There is comparatively small cross-correlation into the horizontal direction. The cross-correlation between  $f$  and  $h$  at the same location is only about half of the strongest cross-correlation between  $f$  and  $h$  at the optimal distance. The knowledge of the particular structure of the cross-covariance function is helpful in the design of monitoring networks, in particular if conclusions on the state of one RFV are drawn from the state of another RFV (see chapter 10).

The relatively large differences between numerical and analytical solutions for the cross-covariances of the two least variable isotropic soils (approximately 10% to 20%, see Figure 8.6)

are due to a simplification in the computation of the numerical sample cross-covariance. The RFVs  $f$  and  $a$  are element properties while the RFVs  $h$ ,  $v_x$ , and  $v_z$  are node properties. Since  $y$  depends on both nodal and elemental properties ( $h$ ,  $f$ , and  $a$ ), nodal properties must be extrapolated to the element or vice versa. For the sampling procedure here, the head values  $h_i^k$  of the four nodes  $i$ ,  $i=1..4$ , around an element  $k$  are averaged, and  $y$  is obtained as elemental property of  $f$ ,  $a$ , and the average head  $h^k$  in a particular element. Averaging the head values introduces a small reduction in the variance of the unsaturated hydraulic conductivity. However, since the head values have a strong spatial correlation, the variance reduction can be neglected. In the sampling process for the sample cross-covariance (8-9) it is assumed - for simplicity - that the location of an element is identical to the location of the lower left node of that element. The error in the cross-covariance fields relating nodal with elemental properties stems from the discrepancy between the assumed identity of element and node location.

Note that the correlation  $\rho_{af}$  has a significant impact on the cross-covariances of  $f$ ,  $h$ , and  $v_z$ : In the correlated soils the vertical cross-covariance structure of  $C_{fh}$  and  $C_{f,vz}$  inverts itself at  $H_{\min} = -1/\zeta\Gamma$ , such that in dry soils  $C_{fh}$  has a minimum at negative lag distances and a maximum at positive (upward) lag distances. The cross-correlation between  $f$  and the (negative, downward) vertical velocity becomes positive at lower head pressure, because under dry conditions soils with high saturated hydraulic conductivity and coarse texture (large  $\alpha$ ) are assumed much less permeable than soils with low saturated hydraulic conductivity and fine texture (low  $\alpha$ ). In the uncorrelated soils the correlation between  $f$  and  $h$  and  $f$  and  $v_z$  weakens as the soil dries out, which can be seen by comparing, for example, the cross-covariance  $C_{fh}$  for the correlated soil at -2000 cm (#24) and the uncorrelated, anisotropic soil at -1000 cm (#15). Both have approximately the same absolute maximum of 30 cm, although the head variance in the uncorrelated soil is almost twice as large as in the correlated soil (Figure 8.18).

## 8.9 Summary and Conclusions

For the first time, intensive Monte Carlo simulations of unsaturated, steady-state gravity flow have been implemented for a large range of different soils. The Monte Carlo simulations take advantage of the ASIGNing technique introduced in chapter 7, which combines the efficiency of spectral perturbation analysis with the flexibility and accuracy of finite element modeling. It allows for the fast simulation of steady-state head and flux in two-dimensional vertical soils. It is applicable to a great variety of different soils and is therefore well-suited for the stochastic simulation of unsaturated flow at actual field sites. The approach is here used for a stochastic analysis of the unsaturated hydraulic conductivity, the soil water tension, and the soil water flux. The simulations are all implemented with 1000 realizations on a finite element domain of 64\*64 rectangular elements. To avoid aliasing effects in the fast Fourier transform of the random field and initial guess generator, the size of the initial random field is at least 10 by 10 correlation scales  $\lambda_f$  and in most cases exceeds 20 by 20  $\lambda_f$  as recommended by Gutjahr et al. (1989). The large amount of realizations for each Monte Carlo simulation results in a very small variability of the sample moments, which allows for both a numerical model validation and an evaluation of first order analytical solutions that were introduced almost a decade ago (Yeh et al., 1985a,b, chapter 4), but have never been rigorously tested for their validity in mildly and strongly heterogeneous soils. By comparing the spatial variability of the local sample mean and variance with the expected variability of the sample moments it was shown that the Monte Carlo simulations indeed converge and that the theoretical variability (8-3) and (8-7) of the sample mean and sample variance provide good estimates of the actual sample error. While the analysis here is limited to the case of exponential input covariance functions, the numerical and analytical methods introduced in this work are both applicable to arbitrary input covariance functions and arbitrary correlation structures between  $\log K_s$  and  $\log \alpha$ . In many field situations, these covariance and cross-covariance functions are obtained from geostatistical analysis (Isaaks and Srivastava, 1990). Both the analytical and numerical approach

can in principle also be used to analyze flow in unsaturated soils of multidirectional flow with arbitrary, spatially constant mean head gradients (see also Yeh et al., 1985b).

Beyond the analytical solutions of chapter 4, the Monte Carlo model introduced in this chapter provides not only more accuracy, but mainly the flexibility offered by the finite element model with respect to boundary conditions and the particular probability distributions of the input parameters. Moreover, the combination of spectral analysis and numerical model (ASIGNing) makes it possible to simulate quasi-infinite domains or semi-infinite domains (e.g. with random head vertical boundaries, flux boundary at the top and water table at the bottom). The analysis has shown that even for very heterogeneous flow fields (large  $\sigma_y^2=3.2$ ) the use of the initial first order perturbation solution as random head boundary adversely affects the results within no more than one or two correlation scales from the boundaries, which is similar to the effect of using constant head or constant flux boundaries. In mildly to moderately heterogeneous flux fields ( $\sigma_y^2 \leq 1$ ), the use of random head boundaries obtained by first order analysis is less biased than the use of constant head or flux boundary conditions, if deterministic boundary conditions are not truly justified. Mixed deterministic/random boundaries can also be introduced by conditioning as demonstrated in chapter 10. Arbitrary boundaries could also be specified including non-stationary boundary conditions (e.g. above a water table).

The analytical solutions derived in chapter 4 for the mean and variances of the dependent RFVs  $y$ ,  $h$ , and  $v$  are found to be - for all practical purposes - very similar to the numerical solutions if the resulting variance of  $y$  is less than 1.0. (mildly variable flow). The two-dimensional covariance and cross-covariance functions are also in good agreement with numerically sampled models. For  $\sigma_y^2 \geq 1$  (moderately to strongly variable flow) the analysis of chapter 4 provides some general insights, but the actual, fully nonlinear (numerical Monte Carlo) solutions differ in parts very significantly. The most important findings of the stochastic analysis of the dependent RFVs and the comparison with the analytical solutions of chapter 4 are summarized here:

The first and second moments of  $y$  are very robust with respect to first order

perturbation analysis. Even for very heterogeneous flow, the differences between Monte Carlo and perturbation analysis were less than 16%. The moments of the unsaturated head are generally also well estimated by the first order analysis. If  $\sigma_y^2 > 1$  i.e., if a soil is either strongly variable in the saturated hydraulic conductivity and in the soil pore parameter  $\alpha$  or if it is a dry soil, the variance of the head is significantly overestimated by the perturbation analysis (up to 30 %). The head correlation function is in very good agreement with the numerically obtained correlation field throughout the simulations. At large lag-distances, the sample  $\rho_{hh}$  in the Monte Carlo simulation is slightly higher than predicted, possibly because of boundary effects in the numerical simulation. The sample pdfs of  $h$  and  $y$  are always found to be Gaussian, except for very high  $\sigma_\alpha^2$ , where  $y$  showed a small but notable skewness.

The probability distributions of the two velocity components are skewed. The vertical velocity is best described by a lognormal pdf. At very high  $\sigma_y^2$ , however, the sample pdf (histogram) of  $v_z$  extends beyond zero velocity and shows that a significant number of nodes with upward velocities exists. Such a pdf cannot be modeled with the lognormal function. The pdf of  $v_x$  is symmetric and has an exponential decay as  $|v_x|$  increases. However, it is differentiable for  $v_x \rightarrow 0$  as shown by plotting the pdf of  $\log|v_x|$ . Since the first order perturbation analysis assumes normal RFVs, it is generally much less accurate in predicting the flux (velocity) than in predicting the head and the unsaturated hydraulic conductivity at equal variability of  $y$ . Only the decrease in mean vertical velocity is well predicted by using the mixed order effective hydraulic conductivity analysis of Yeh et al. (1985b).

The Monte Carlo sample correlation fields and the analytically determined correlation functions for the velocity are in good qualitative agreement for all the tested soils. The covariance function of the horizontal velocity is symmetric with respect to the origin, but has its major axes diagonal to the major axes of anisotropy. The vertical velocity covariance is strongly anisotropic with a larger vertical than horizontal correlation scale, even if the underlying hydraulic soil properties are isotropic. The covariance functions of the two velocity components reflect the peculiar flow structure in heterogeneous soils, which was shown to take

place in a preferential or channel type flow pattern. Even though the underlying random structure of the soil is statistically homogeneous, most of the moisture mass is transported through only a small fraction of the soil. This is in good agreement with field findings (Glass et al., 1988) and the numerical analyses of flow in single fractures with high variability in their conductivity (Moreno et al., 1989).

Similar second order moments for  $h$ ,  $y$ ,  $v_x$ , and  $v_z$  are found for wet, texturally heterogeneous soils and dry, texturally rather homogeneous soils with an equal degree of unsaturated logK heterogeneity. The similarity does not extend to the cross-covariances, which depend not only on the mean head, but also on the correlation between  $f$  and  $a$ . For  $\rho_{af}=0$  the cross-correlation between  $f$  and  $h$ , and between  $f$  and  $v_z$  weakens with increasing soil-water tension. If  $\rho=1$ , the unsaturated hydraulic conductivity parameters are completely determined by  $h$  and  $f$ . Consequently a strong correlation not only between  $f$  and  $h$ , but also between  $f$  and  $v_z$  exists even in dry soils. The use of tension measurements for the conditional simulation of  $f$  and  $a$  random fields is discussed in chapter 10.

The grid-discretization criteria developed in chapter 6 have been proven to provide accurate solutions not only in the context of single large simulations, but also for Monte Carlo simulations with a large number of realizations. A fairly coarse vertical discretization of 10 cm has been shown to provide results of accuracy equal to that of a fine discretization (2.5 cm).

In conclusion this study has shown both the applicability and limitations of the first order perturbation solutions developed in chapter 4 for two-dimensional heterogeneous soils with lognormally distributed  $\alpha$ . ASIGNing provides a flexible tool to implement Monte Carlo simulations efficiently on today's available workstations.

Table 8.1

Input parameters for the various hypothetical soil sites:  $\sigma_f^2$ : variance of  $f=\log K_s$ ,  $\sigma_a^2$ : 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	$\sigma_f^2$	$\sigma_a^2$	$\rho_{af}$	$\Gamma$	H	$\Delta x$	$\Delta z$	$\lambda_{fx}$	$\lambda_{fz}$
#3	1.0	0.01	0	0.01	-150	10	10	50	50
#2	0.01	10 <sup>-4</sup>							
#4			1						
#6					-1000				
#8	0.12								
#9	4.0	0.04							
#10			1		-3000				
#11			1					300	
#12			1			30		300	
#13						20		200	30
#15					-1000	30		300	
#19						2.5	2.5	12.5	12.5
#20							2.5		12.5
#21			1		-3000	30		300	
#22	4.0	0.04				30		300	
#23			1		-3000	20		200	30
#24			1		-2000	30		300	
#25	4.0	0.04				20		200	30
#26		0.64							
#27	20	0.09	1	10 <sup>-4</sup>	-1.8E5	30		300	
#28	2.25	0.04				30		300	
#29						30		300	
#30				0.1	-100				
#31						15		150	

Table 8.2

Comparison of the numerical and first order analytical stochastic solutions for the mean and variance of the dependent RFVs head  $h$ , unsaturated hydraulic conductivity  $y$ , horizontal velocity  $v_x$ , and vertical velocity  $v_z$ . #2, #8, and #3 are three different Monte Carlo simulations with  $\sigma_f^2 = 0.01, 0.1, \text{ and } 1.0$ , respectively. All other parameters are identical to base case #3 (Table 8.1).

	#2 numerical - analytical		#8 numerical - analytical		#3 numerical - analytical	
h:						
mean	-150.1	-150.0	-150.4	-150.0	-150.9	-150.0
variance	11.2	10.6	157	150	1079	1060
y:						
mean	-1.499	-1.500	-1.503	-1.500	-1.498	-1.500
variance	8.53E-2	8.90E-2	.121	.126	.858	.887
$v_x$ :						
mean	-3.18E-5	0.00000	-1.03E-4	0.00000	-4.35E-4	0.00000
variance	5.49E-5	5.49E-5	8.03E-4	7.78E-4	7.23E-3	5.49E-3
$v_z$ :						
mean	-.2232	-.2231	-.2230	-.2231	-.2293	-.2231
variance	1.71E-4	1.76E-4	2.44E-3	2.50E-3	1.93E-2	1.76E-2

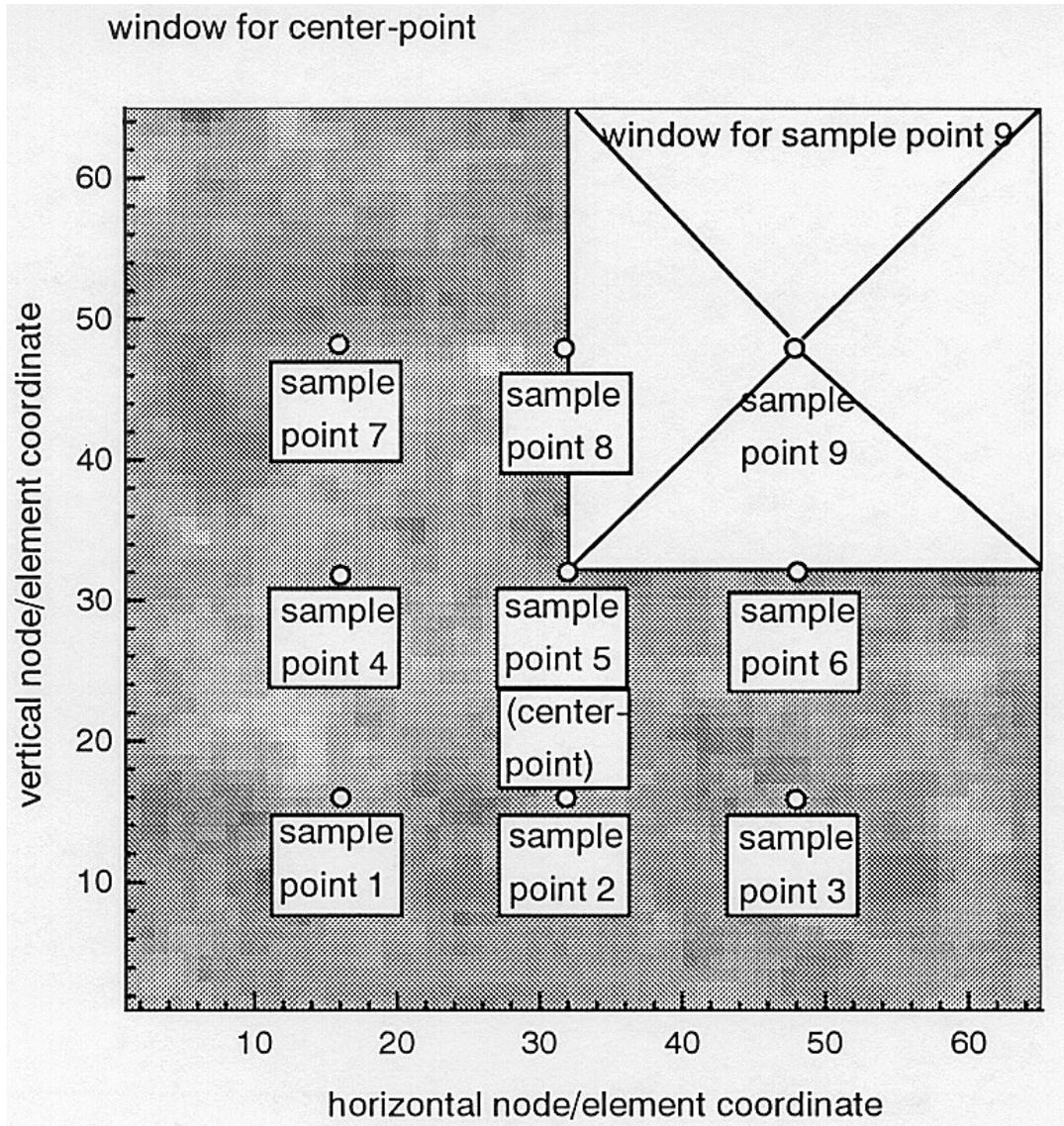


Figure 8.1: Location of the sampling points for which local covariance and cross-covariance fields are obtained. The local (cross-)covariance fields are computed in a 31 by 31 window around each sample point. The center point has the entire field as window.

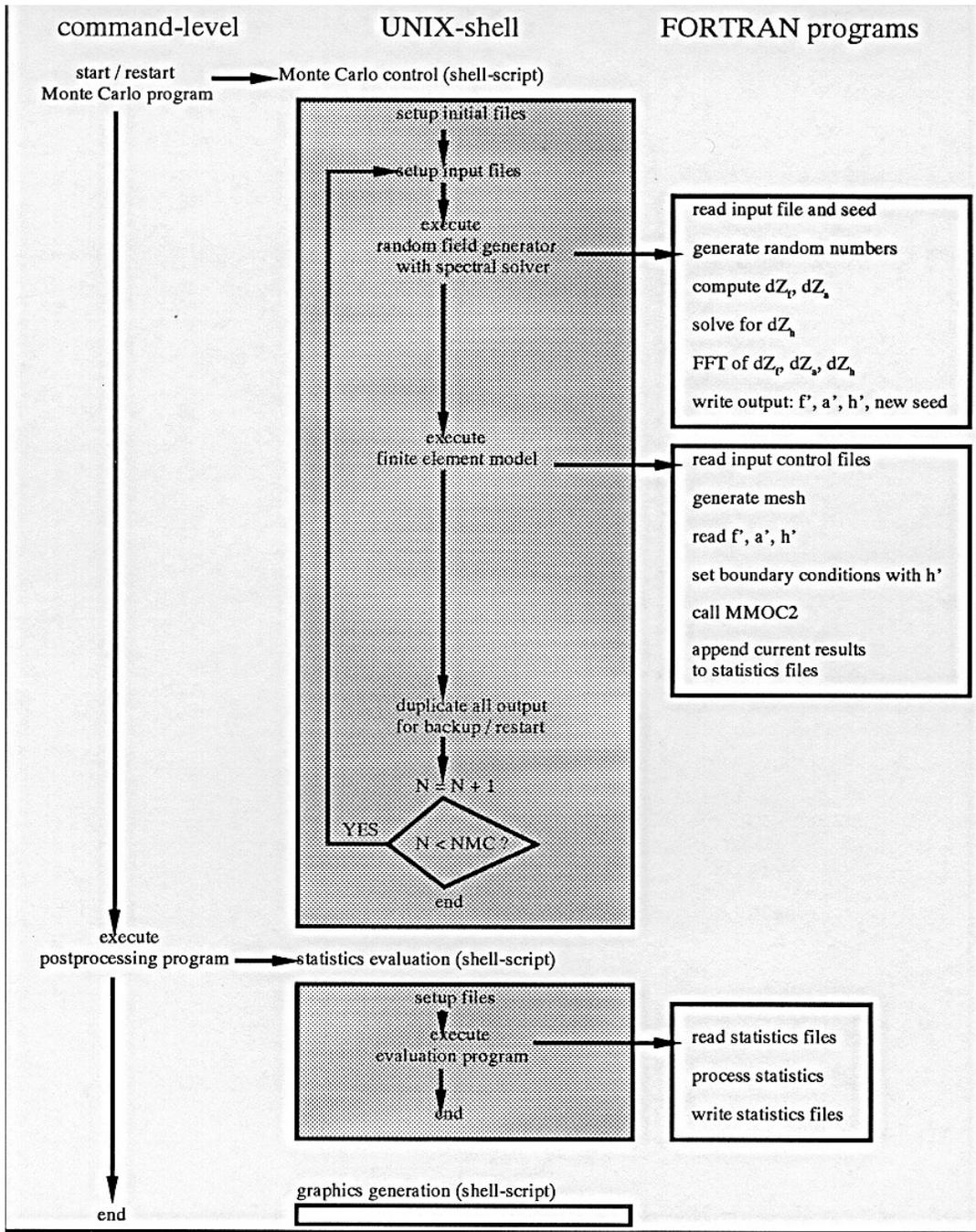


Figure 8.2: Flow chart of the Monte Carlo simulation.

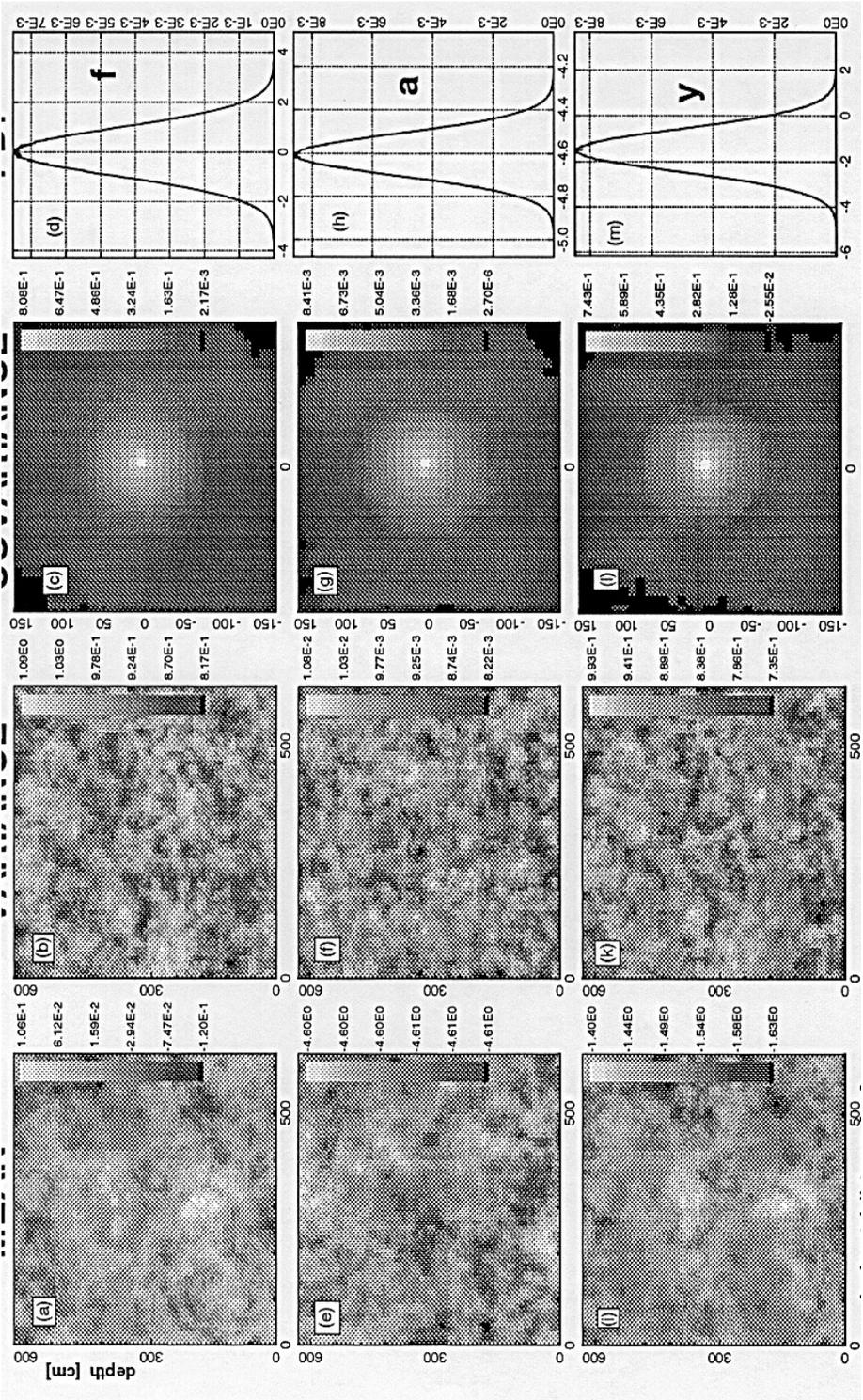


Figure 8.3: the base soil site (#3).

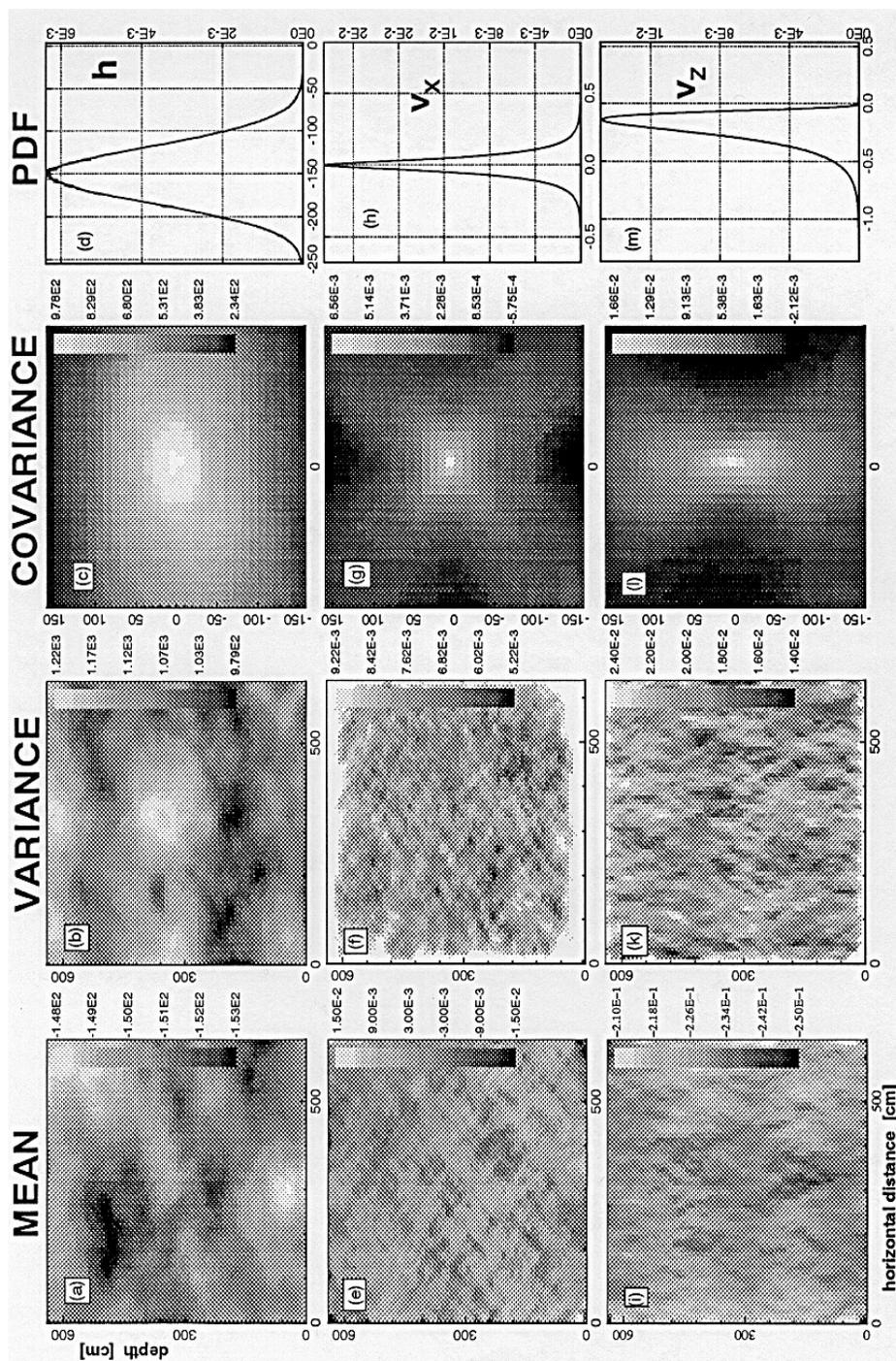


Figure 8.4:  $V_z$  (Fig. 8.4i-m) of the base soil site (#3).

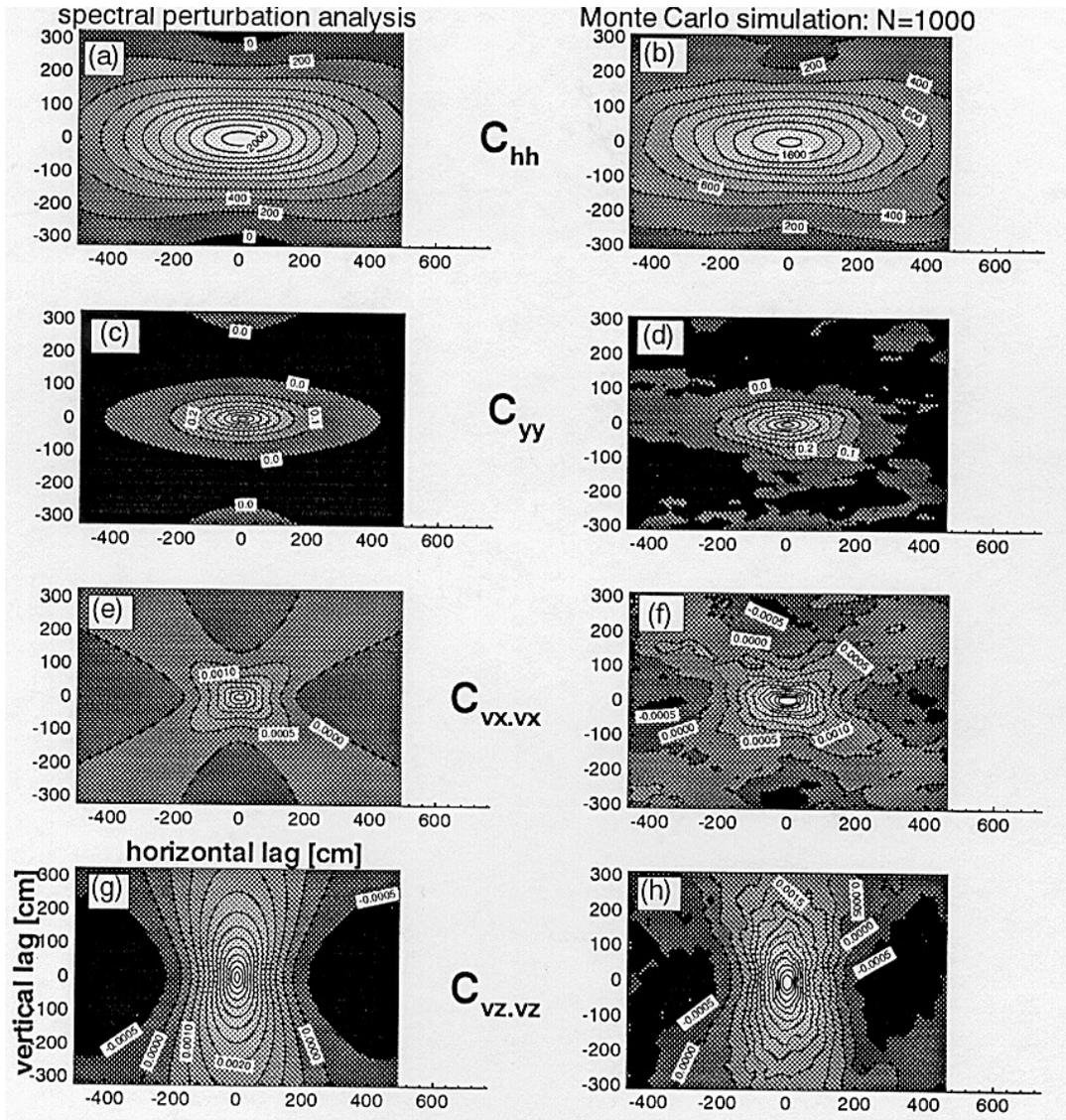


Figure 8.5: Analytical ensemble covariance fields (left) and Monte Carlo sample covariance fields (right) around the center-point of the simulation domain in an anisotropic soil site (#31:  $\lambda_x = 150$  cm,  $\sigma_1^2 = 1$ ,  $H = -150$  cm ).

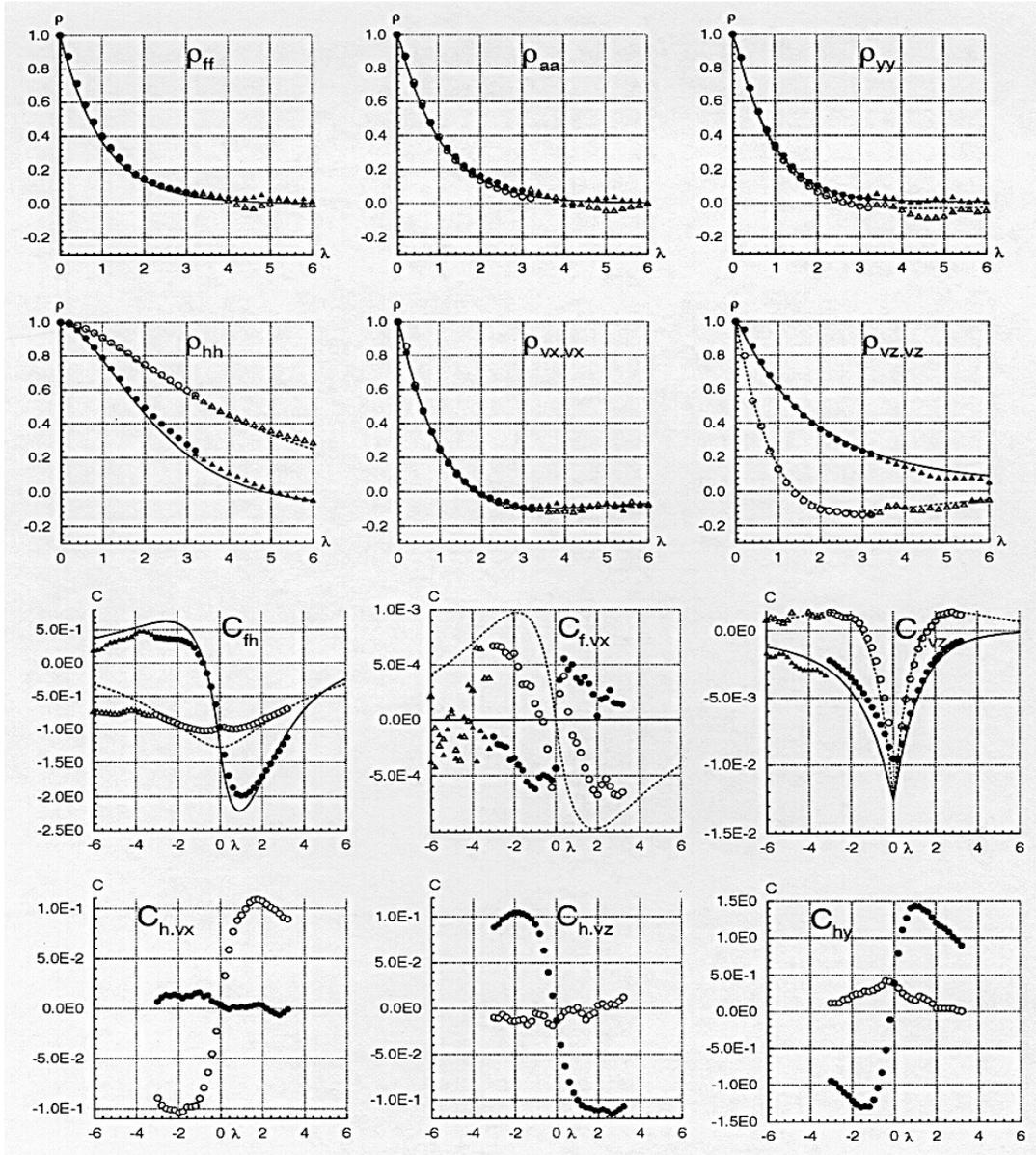


Figure 8.6: Cross-sections of the correlation functions  $\rho$  and cross-covariance functions  $C$  for the #8 soil site. Horizontal and vertical first order analytical functions are indicated by dashed and solid lines, respectively. Horizontal and vertical MC sample functions are noted with hollow and solid symbols, respectively. Circles are average sample functions, triangles are from the center-point correlation and cross-covariance samples (see text).

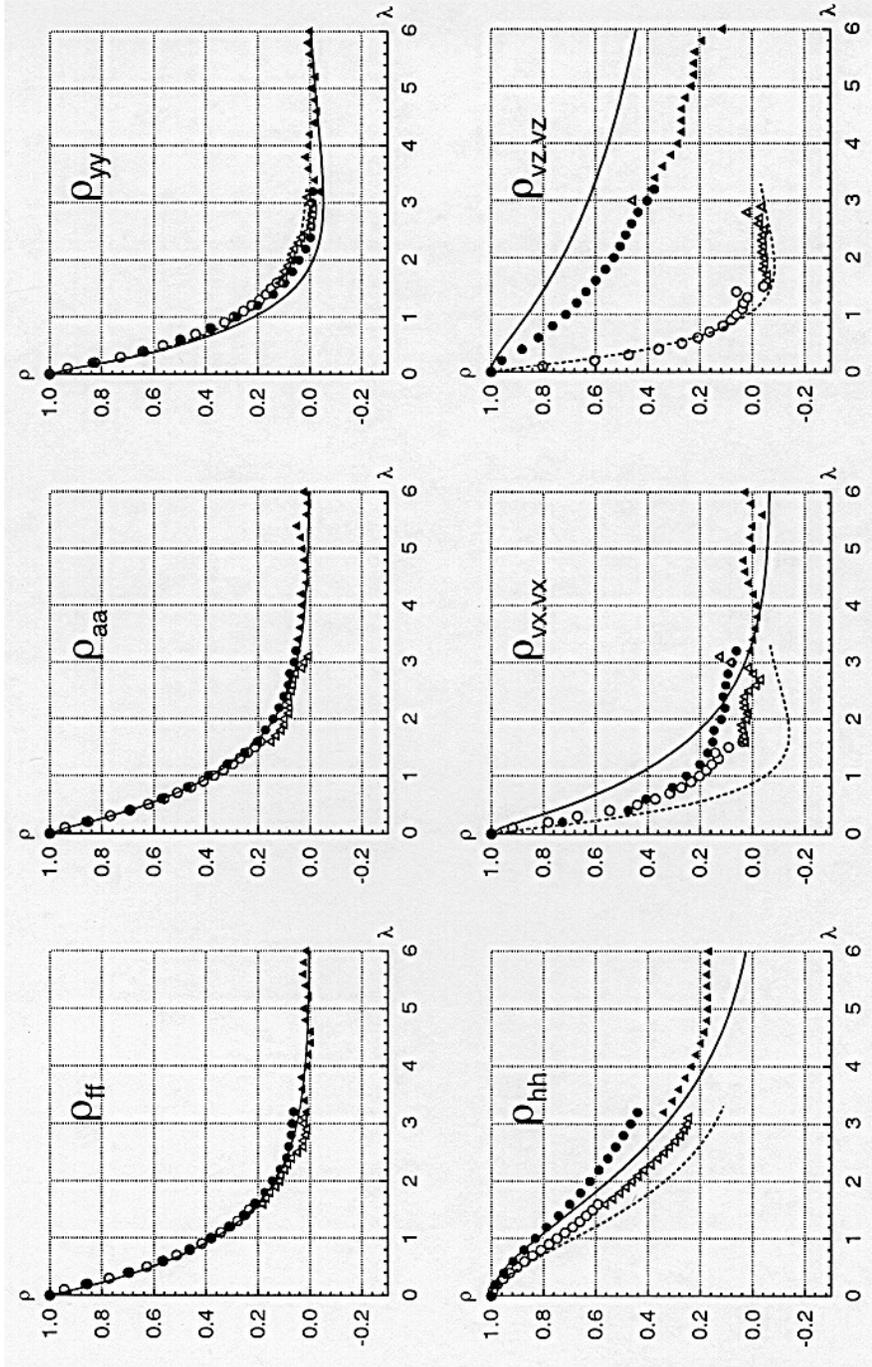


Figure 8.7: (symbols for an anisotropic, highly variable, wet soil (#22). Legend as in Figure 8.6.

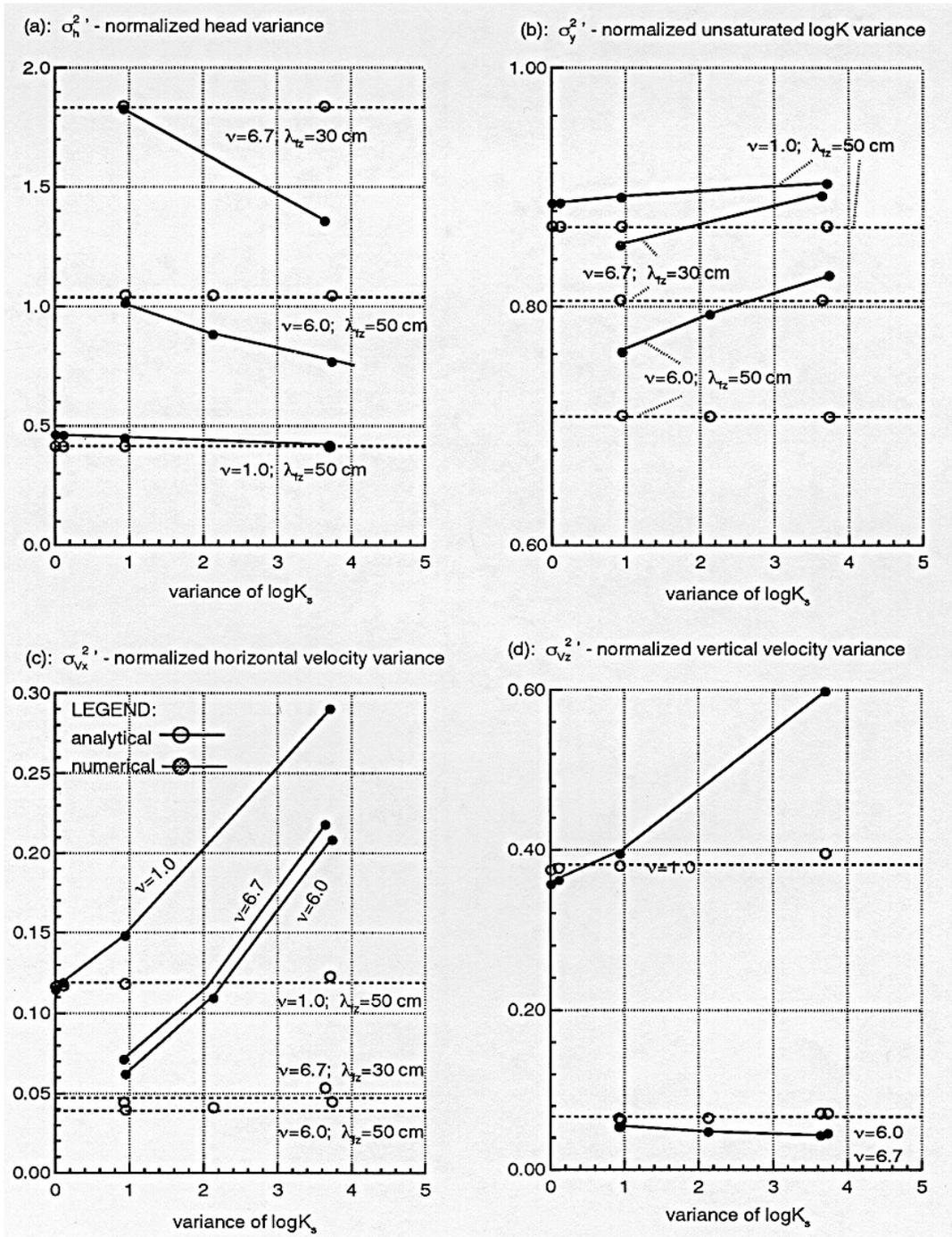


Figure 8.8: Dimensionless variances of the dependent RFVs as a function of the variance of the RFV  $f (= \log K_s)$ .

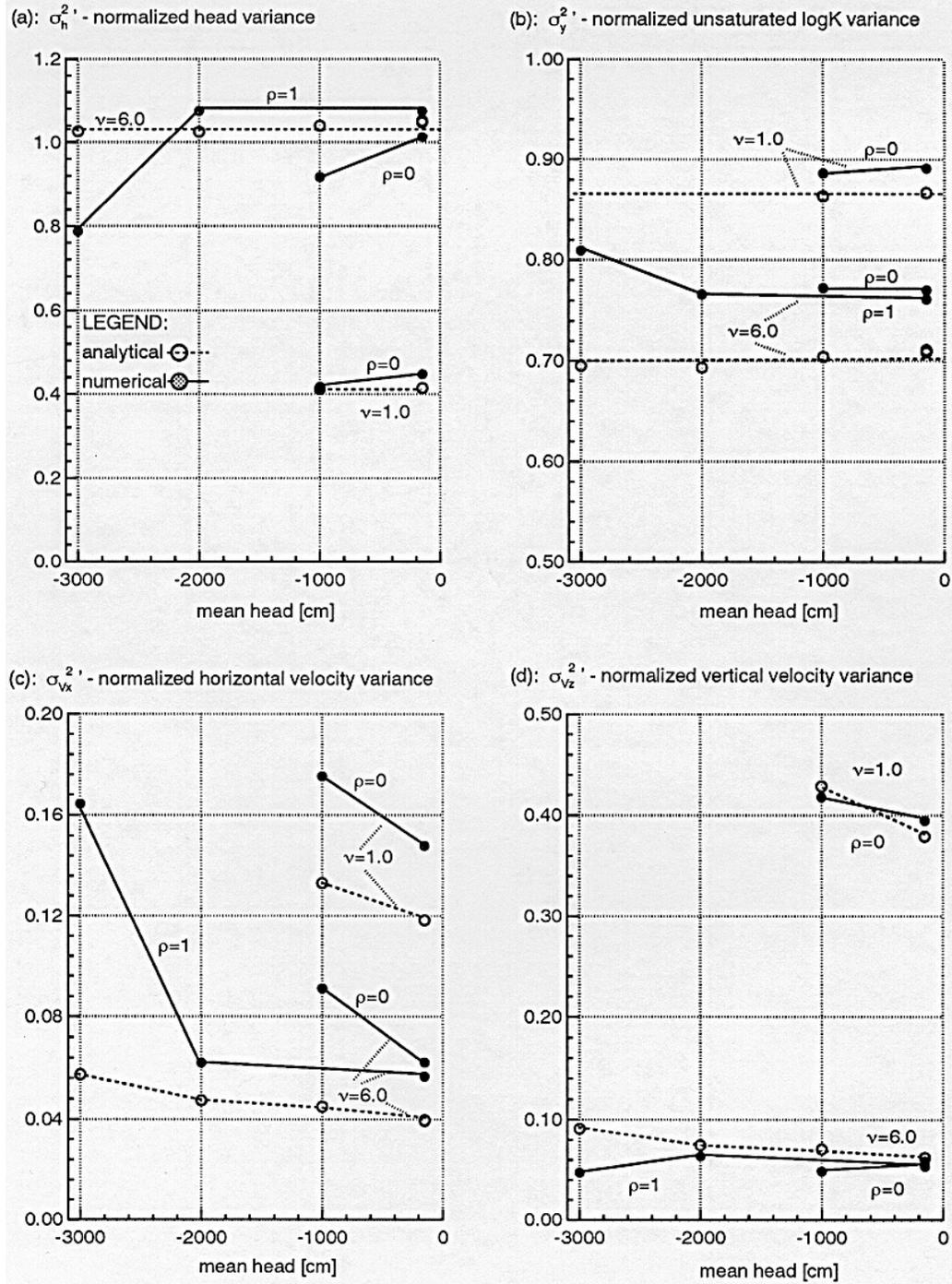


Figure 8.9: Dimensionless variances of the dependent RFVs as a function of the mean head.

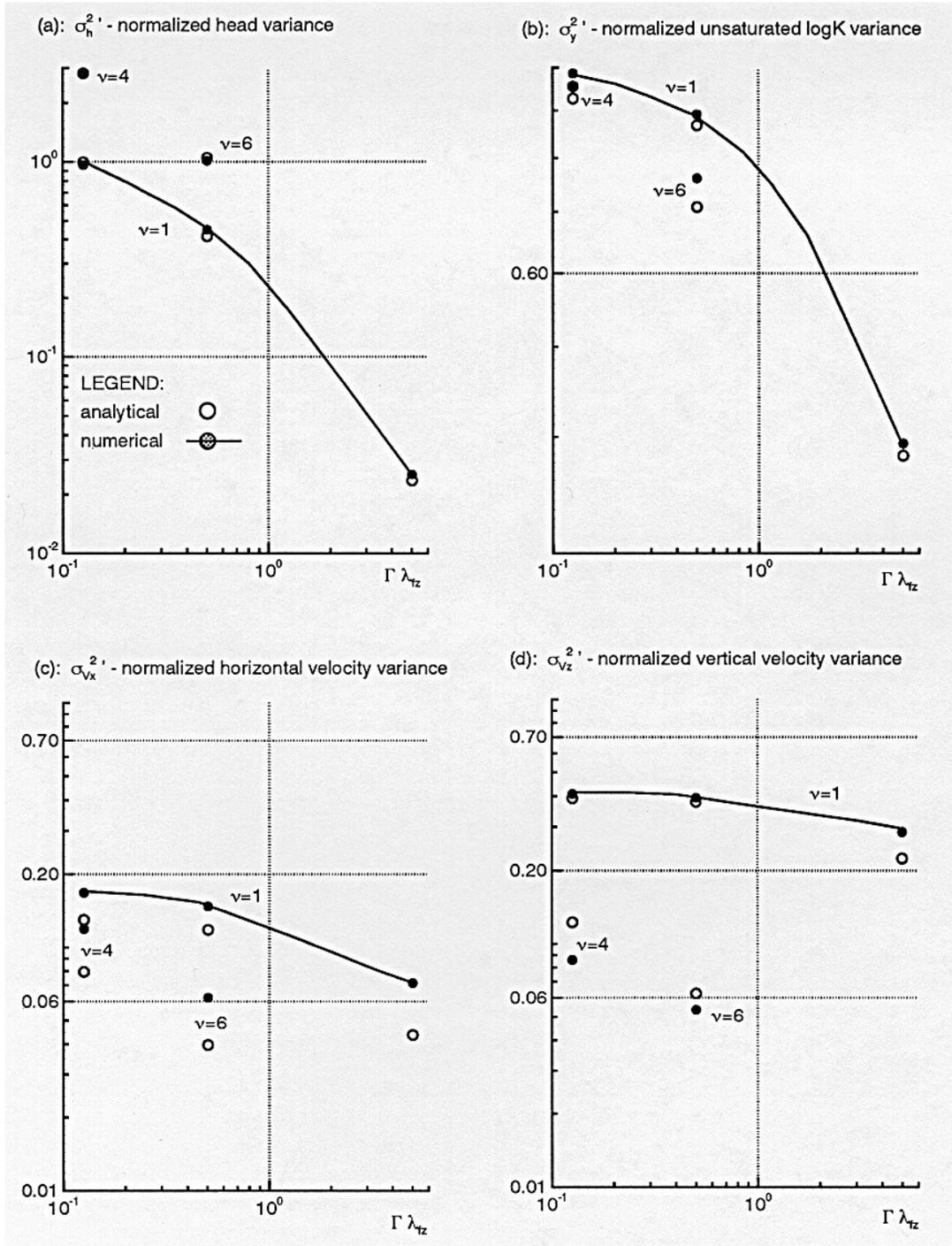


Figure 8.10: Dimensionless variances of the dependent RFVs as a function of  $\Gamma\lambda_{TZ}$ , where  $\Gamma$  is the geometric mean of  $\alpha$  and  $\lambda_{TZ}$  is the vertical correlation scale of  $f$ .

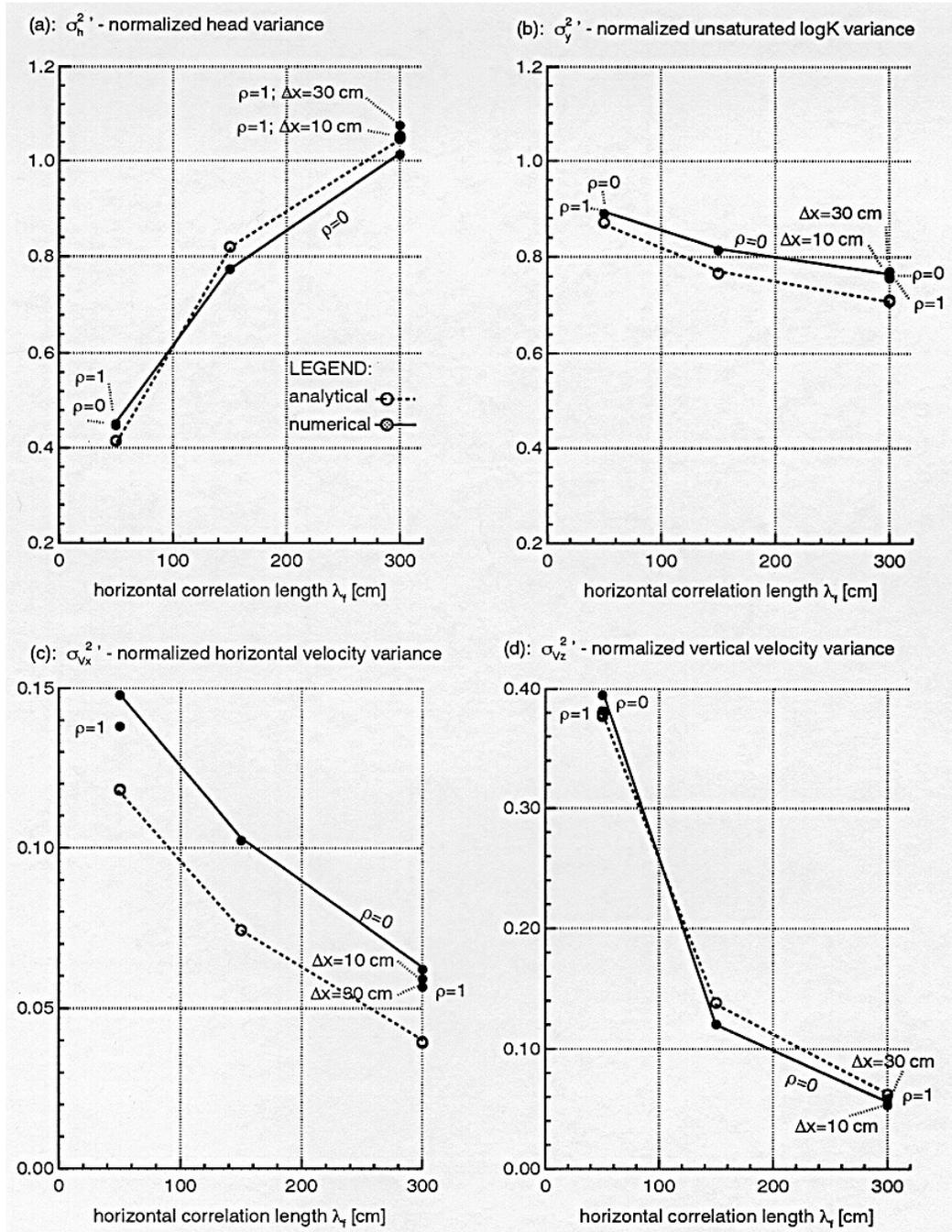


Figure 8.11: Dimensionless variances of the dependent RFVs as a function of the horizontal correlation scale of  $f$ . The vertical correlation scale of  $f$  is 50 cm.

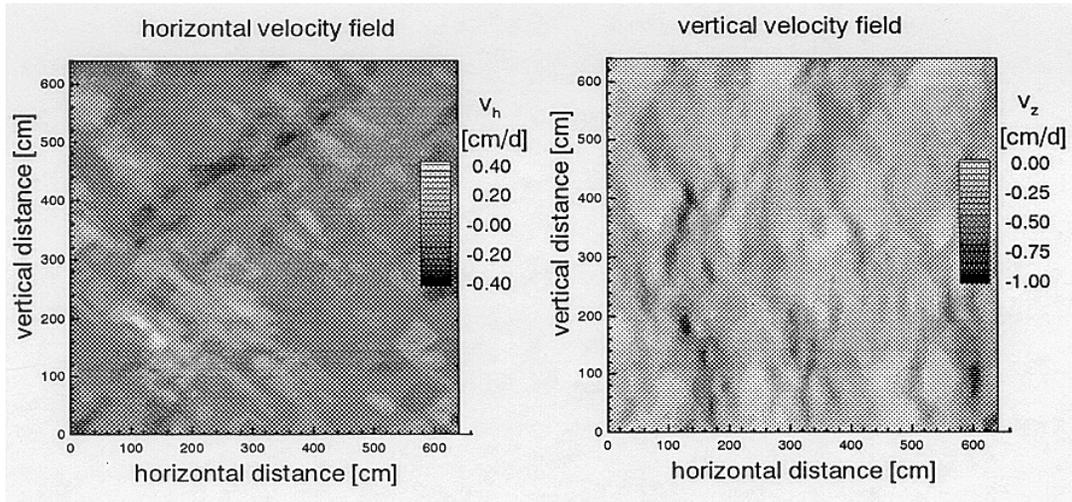


Figure 8.12: Single solution for  $v_x$  (left) and  $v_z$  (right) from a random field realization of  $f$  and  $\alpha$  for the isotropic base soil site (#3).

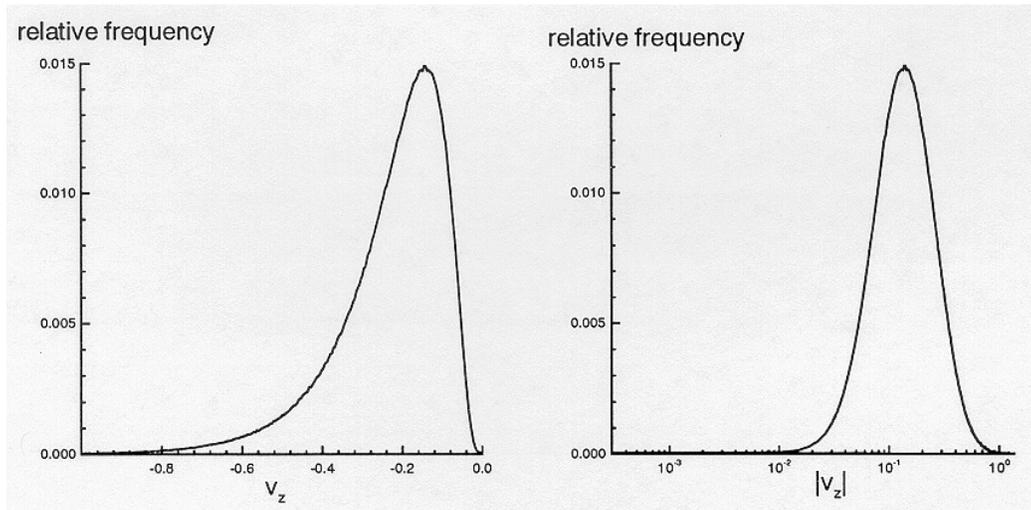


Figure 8.13: Histogram of  $v_z$  at the isotropic base soil site (#3) on an arithmetic scale (left) and on an absolute logarithmic scale (right).

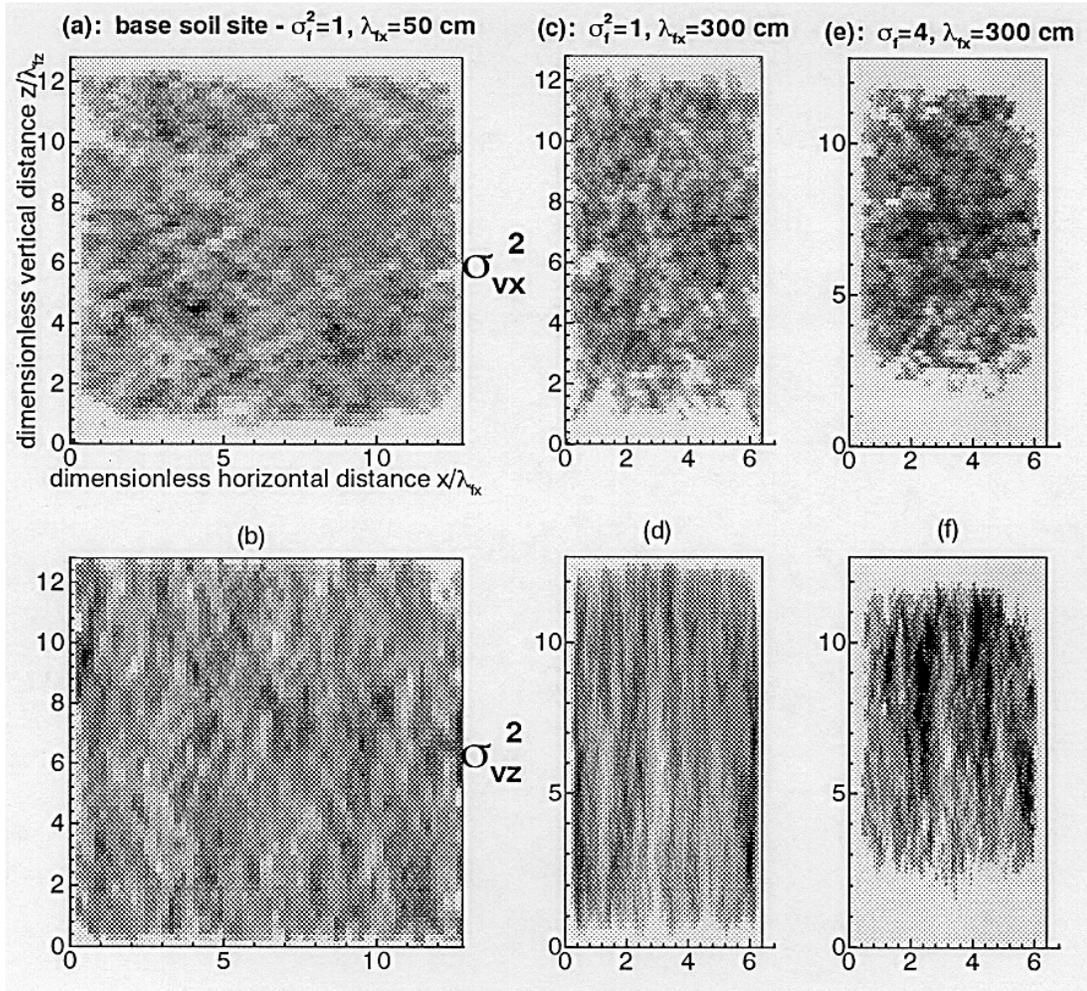


Figure 8.14: Influence of the first order perturbation head boundaries on the horizontal (top) and vertical (bottom) velocity variance for the isotropic base soil (#3, Fig. 8.14a,b), an anisotropic soil (#29, Fig. 8.14c,d), and an anisotropic high f-variance soil (#22, Fig. 8.14e,f). The boundary effect is indicated by the (white) high velocity variance areas near the boundary. Darker shades indicate lower variance.

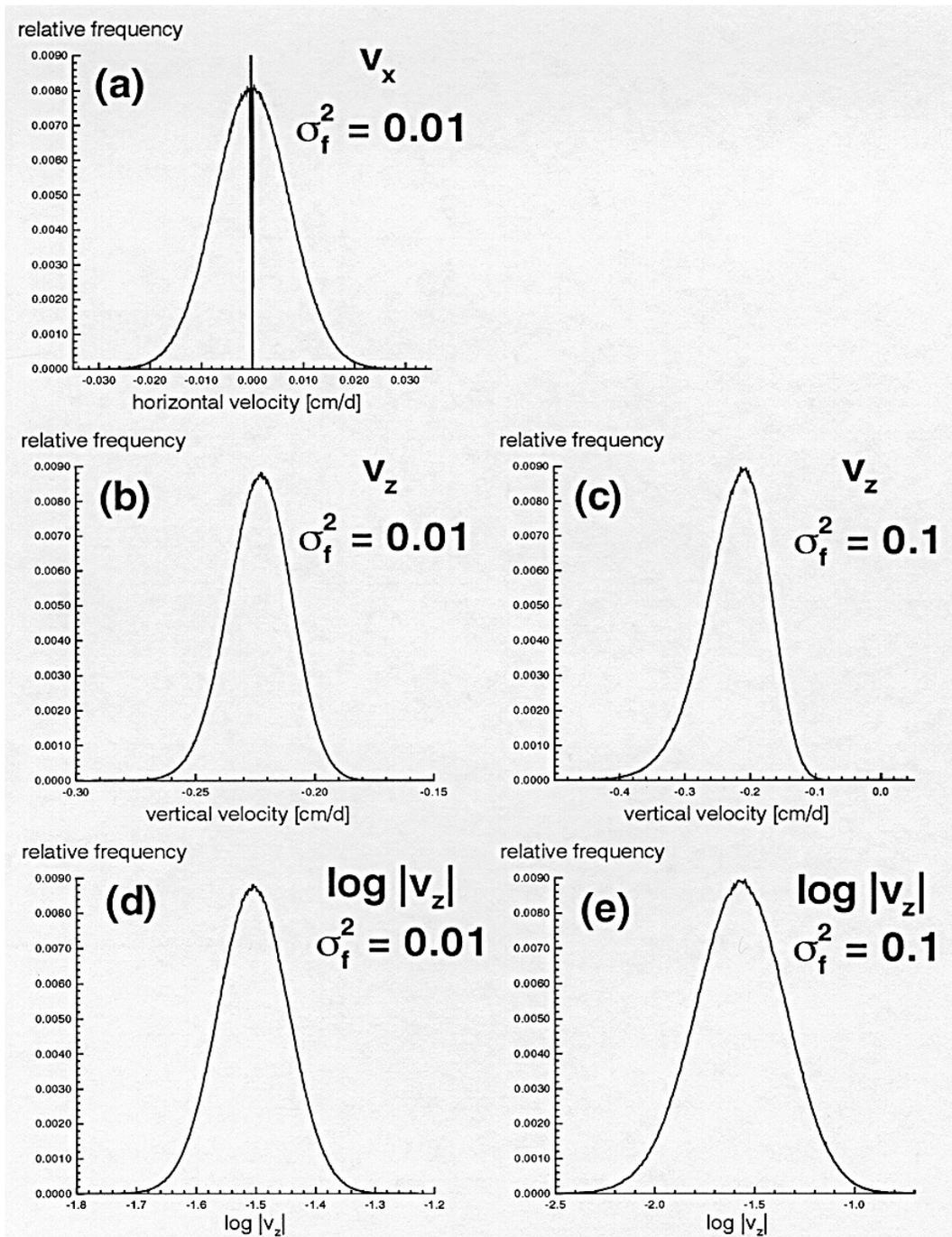


Figure 8.15: Histograms of the horizontal velocity at  $\sigma_f^2 = 0.01$  (a). Histograms for the vertical velocity are given for  $\sigma_f^2 = 0.01$  (b) and  $0.1$  (c). The latter histograms are also plotted both on a logarithmic scale (Fig. 8.15d,e).

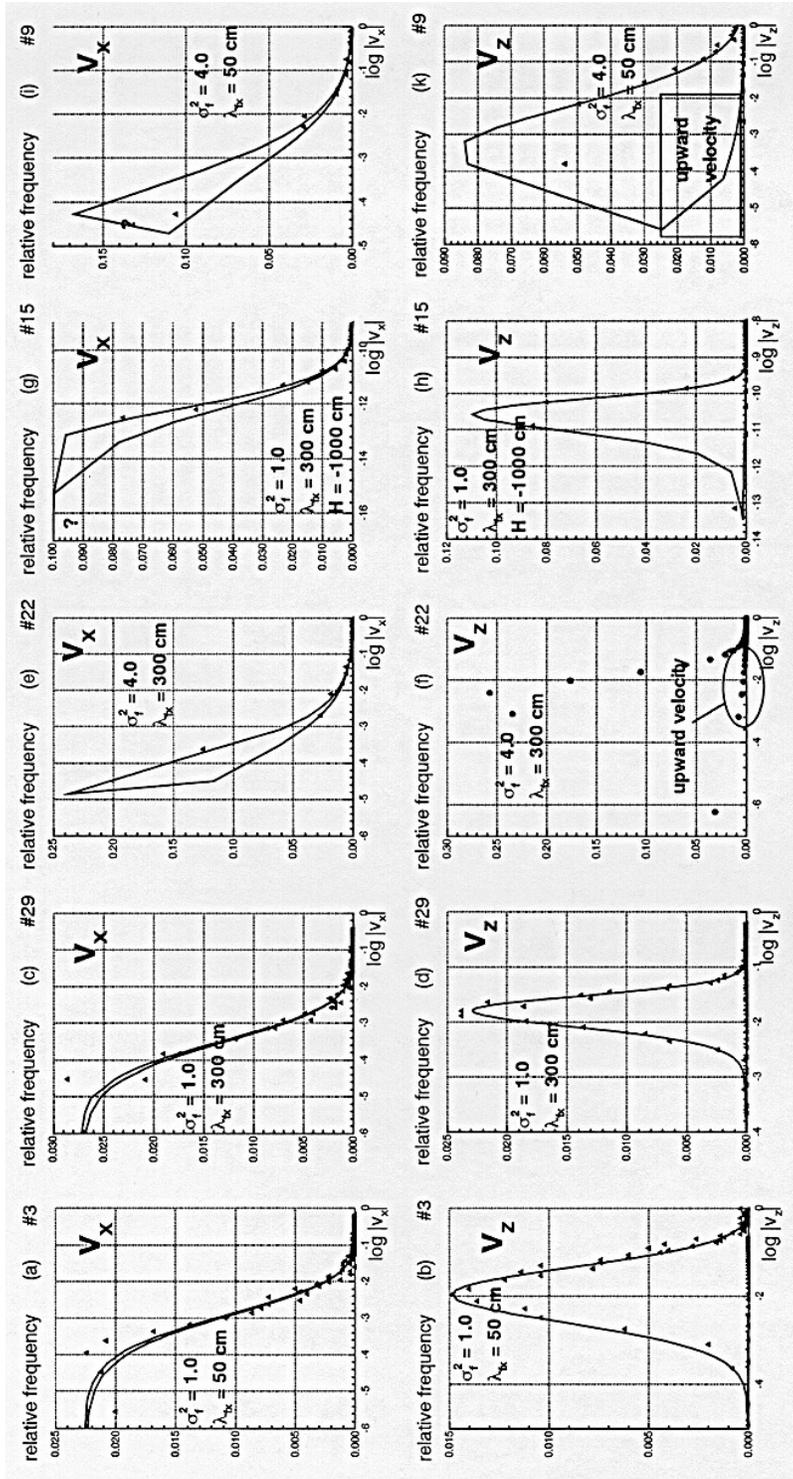
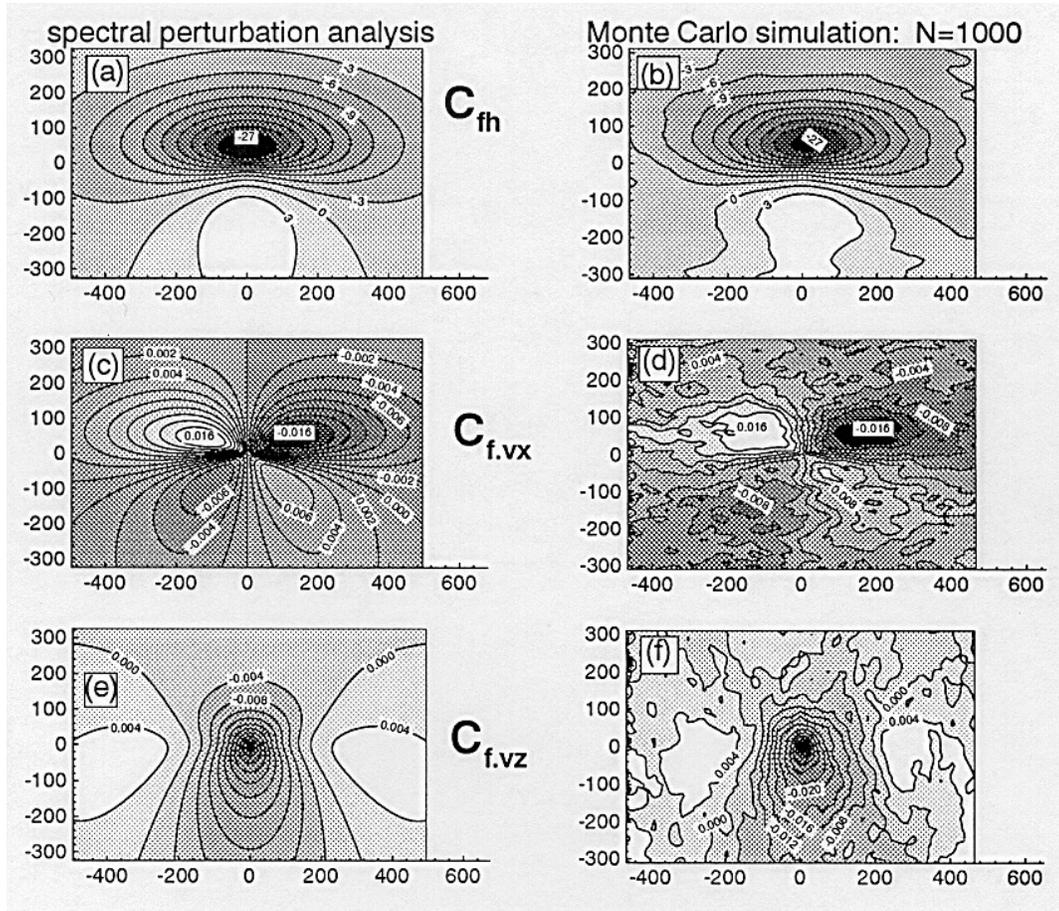


Figure 8.16:

soil), soil site #29 (anisotropic), soil site #22 (anisotropic, f-high variability), soil site #15 (anisotropic, dry), and soil site #9 (isotropic, high f-variability). In all three simulations

$\lambda_{hk}$  and  $\sigma_f^2$  are histogram parameters obtained from the histograms. The histograms are histograms for all



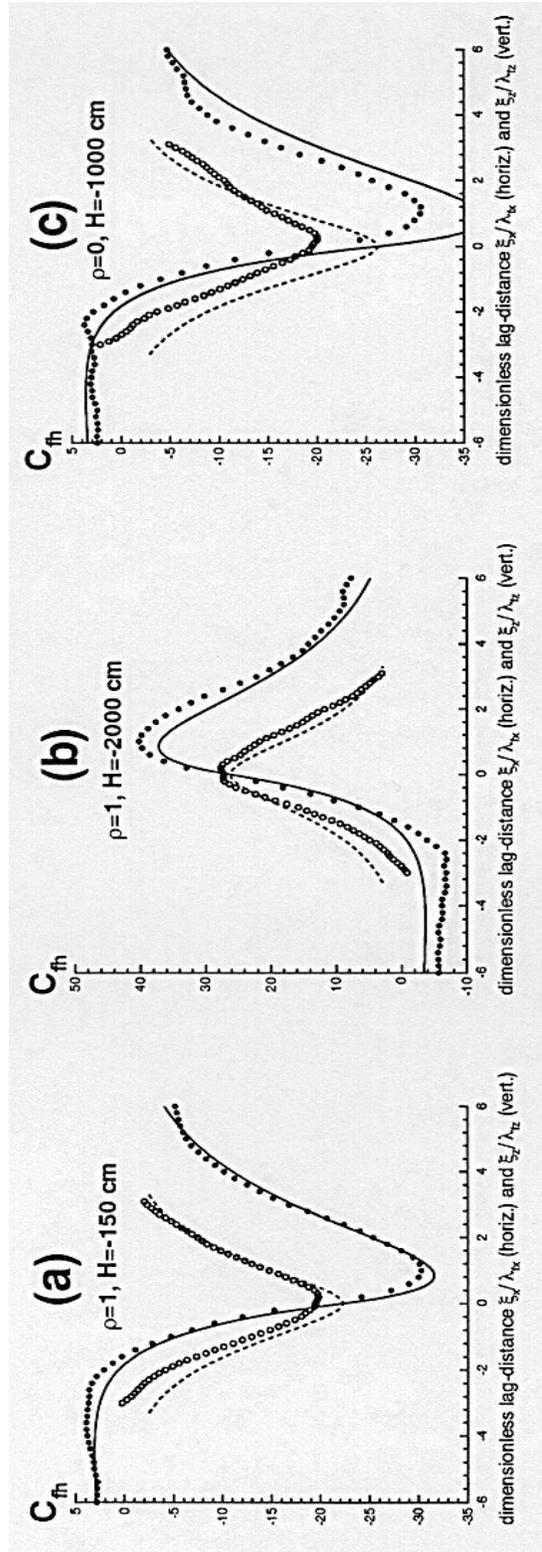


Figure 8.18: heterogeneity but different correlation between f and log left to right: #12, #24, #15). Legend as in Figure 8.6.