SOLUTION OF A PROBLEM IN TRANSIENT GROUNDWATER FILTRATION BY MONTE CARLO METHODS

V. N. Émikh

This method is used mainly for steady-state filtration [1, 2], but a transient-state problem for petroleum hydraulics has been solved in this way [3]. Here I consider a typical case of transient infiltration of groundwater with a free surface; some special features are noted.

1. General. For the region G ($x>0$, $y>0$, $t>0$ we consider the equation

$$\frac{\partial h}{\partial t}=\frac{k}{\mu}\left[\frac{\partial}{\partial x}\left(h\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial yx}\left(h\frac{\partial h}{\partial y}\right)\right] \tag{1.1}$$

subject to the boundary conditions

$$h(x, y, 0)=H_0, \qquad h(0, y, t)=h(x, 0, t)=H_1 \tag{1.2}$$

The problem of (1.1) and (1.2) is that of transient flow of groundwater in an unpressurized stratum having a horizontal impermeable layer. The infiltration region is a square bounded on two sides by mutually perpendicular channels, which coincide with the positive semiaxes of the (x, y) coordinate system. In (1.1), k and μ are the coefficients of filtration and water release of the ground, t is time, and h is the height of the free surface of the groundwater above the impermeable layer. The groundwater level is H_0 at $t=0$ throughout the infiltration region and subsequently alters when the level in the channels changes instantaneously to H_1.

This is known as Kamenskii's problem [4] and is convenient for the Monte Carlo method in that the results can be compared with those from various finite-difference schemes [4, 5].

The method has so far been developed only for linear differential equations, so we linearize (1.1) by putting $h=h^*$ in the parentheses, in which h^* is some mean ordinate of the free surface in region G. We divide G up via a net with steps of l in the spatial coordinates and τ in time in such a way that the nodes with zero values lie at the boundary of the region. We represent the derivatives in the linearized equation as the ratios of finite differences [6] to get a system of equations whose detailed form is dependent on the particular scheme.

2. Explicit Finite-Difference Scheme. In place of (1.1) we have the following system of equations at the nodes of the net:

$$H_{i,j,s}=\left(1-\frac{4a^2\tau}{l^2}\right)H_{i,j,s-1}+\frac{a^2\tau}{l^2}\left(H_{i-1,j,s-1}+H_{i+1,j,s-1}+H_{i,j-1,s-1}+H_{i,j+1,s-1}\right)$$
$$\left(a^2=\frac{kh^*}{\mu};\ i, j, s=1, 2, \ldots\right) \tag{2.1}$$

Conditions (1.2) become

$$H_{i,j,0}=H_0, \qquad H_{0,j,t}=H_{i,0,t}=H_1 \tag{2.2}$$

Novosibirsk. Translated from Zhurnal Prikladnoi Mekhaniki i Tekhnicheskoi Fiziki, Vol. 9, No. 4, pp. 163-168, July-August, 1968. Original article submitted January 12, 1968.

$H_{i,j,s}$ is determined by the Monte Carlo method [1, 2, 7] by following a wandering particle from node (i, j, s). Each step is a transition from the given node to one of the adjacent nodes related by the difference scheme. Transition to an adjacent node occurs with a probability equal to the coefficient to H at that node, and the relation between the H at the nodes should be solved for H at the starting node, as should (2.1). The random walk ceases if the particle passes outside the region, and this imposes a penalty equal to the value of the function at the exit point, while the particle starts again from node (i, j, s). Repetition of this process gives a statistical estimate of the mathematical expectation of the penalty for node (i, j, s), which has been shown [1, 7] to equal $H_{i,j,s}$, which approximates h at the corresponding point. This can be expressed via

$$H_{i,j,s} = \sum_{r=1}^{R} b^{(r)}_{i,j,s} H_r \tag{2.3}$$

in which H_r is the value of H at boundary node r (including the nodes of the initial layer), while $b^{(r)}_{i,j,s}$ is the probability of reaching boundary node r from node (i, j, s).

Steps l and τ are chosen on the basis of the stability condition for the explicit scheme, which is (see chapter 1, §11, of [6]):

$$a^2\tau / l^2 \leqslant 1/4 \tag{2.4}$$

We put (compare [4])

$$a^2\tau / l^2 = 1/4 \tag{2.5}$$

and then (2.1) becomes

$$H_{i,j,s} = 1/4 (H_{i-1,j,s-1} + H_{i+1,j,s-1} + H_{i,j-1,s-1} + H_{i,j+1,s-1}); \; i, j, s = 1, 2, \ldots \tag{2.6}$$

Then (2.6) implies that the particle at each stage passes with a probability of 1/4 to one of the four nodes of the previous time layer and after step s is in the initial layer if it has not passed before this into one of the channels. Then this scheme means that the random walk started in time layer s certainly ends not later than after s steps.

Violation of (2.4) within this scheme means that there is a negative probability of passing from a node to one placed directly under it.

The random walk is performed as follows. The section [0.1] is divided into four parts, each corresponding to a certain event (passage of the wandering particle from a given node to one of the four adjacent ones). The length of each section equals the probability of the corresponding event, this probability being 1/4 for all events in this case. The particle passes to the adjacent node set by the value of the random number η $(0 < \eta < 1)$ produced by a generator, which is called at each step. These random numbers must be uniformly distributed in the range 0-1. We used the No. 1 random-number generator of the M-20 computer (see Chapter V, §8 of [8]), and Table 1 shows that this meets the requirement well, in which Δ is the interval and n is the total number of random numbers generated.

The problem of (2.2) and (2.6) was solved with the parameters used in [4]: k=5 m/day, μ=0.06, $h^* = H_0 = 30$ mm, $H_1 = 40$ m, $l = 1000$ m, $\tau = 100$ days, with the steps in l and τ chosen in accordance with (2.5).

An important feature is the number N of steps that must be performed from node (i, j, s) for $H_{i,j,s}$ to be determined with sufficient accuracy without excessive time consumption. The optimum N in any particular case should be determined by trial, but available theoretical estimates can be used to obtain a first approximation. It is known (Chap. V, §2 of [7]) that the following is the N for which the error ε is not exceeded with a probability 0.997 in calculating the mathematical expectation $M\xi$ of an independent random variable ξ:

$$N = 9D\xi / \varepsilon^2$$

TABLE 1

Δ	$n=10$	10^2	10^3	10^4	10^5	10^6
0—0.1	1	11	98	958	10025	99849
0.1—0.2	0	8	86	981	10057	100747
0.2—0.3	0	10	128	1088	10135	99929
0.3—0.4	1	11	101	1037	9901	100196
0.4—0.5	0	8	95	978	10005	99915
0.5—0.6	2	6	86	980	10118	100109
0.6—0.7	2	8	94	996	9926	100048
0.7—0.8	1	10	94	979	9881	99739
0.8—0.9	1	13	100	1044	10036	99480
0.9—1.0	2	15	122	959	9916	99988

in which $D\xi$ is the dispersion of ξ. In this problem ξ takes the two values $\xi_1=30$ m and $\xi_2=40$ m. It is readily shown that the maximum dispersion corresponds to $M\xi=35$ m; then the events $\xi=\xi_1$ and $\xi=\xi_2$ are equally probable, $\rho(\xi_1)=\rho(\xi_2)=0.5$. We have

$$\max D\xi = (30\,\text{m} - 35\,\text{m})^2\,0.5 + (40\,\text{m} - 35\,\text{m})^2\,0.5 = 25\,\text{m}^2$$

From max $D\xi$ (which gives N with a margin of safety) and putting $\varepsilon=0.5$ m, for example, we get

$$N = \frac{9\cdot 25\,\text{m}^2}{0.25\,\text{m}^2} = 900 \tag{2.7}$$

A program was written for the M-20 to find H at the diagonal points of the grid for several instants. Table 2 gives results for t=1500 days. The last line is H from

$$H(x, y, t) = H_1 - (H_1 - H_0)\,\Phi\left(\frac{x}{2a\sqrt{t}}\right)\Phi\left(\frac{y}{2a\sqrt{t}}\right) \quad \left(\Phi(\zeta) = \frac{2}{\sqrt{\pi}}\int_0^{\zeta} e^{-u^2}\,du\right) \tag{2.8}$$

which is the exact solution to (1.1) after linearization. Table 2 gives H (in m) at the $\{i, i\}$ diagonal nodes for t=1500 days, the top 2/3 of the table (for N of 10^3 and 10^4) via the explicit scheme, while the bottom third (for $N=10^3$) via the inexplicit one (the quantities in parentheses are the deviations from the result of (2.4) as % of the latter).

First we consider the part of the table containing the results from the explicit scheme.*

The quantities in the line marked (2.4) are the H calculated directly from that equation, while (2.8) is the exact solution. The other lines in the table have this error mingled with the error from the Monte Carlo method, and the latter is revealed by comparing the result for $l=1000$ m and $\tau=100$ days with that from Kamenskii's scheme.

We see that the error of the Monte Carlo method does not exceed 0.62% for N=100, as (2.7) indicates, this error representing 0.2 m, i.e., being well within the limit set in choosing N. There is a tendency for the error at $N=10^4$ to be less than that for $N=10^3$, but the second and third nodes show fluctuations that deviate from this trend.

The more closely spaced grid obeying (2.5) produces somewhat higher accuracy, as is clear from the table, but the fluctuations tend to obscure the dependence of the error of approximation on the spacing, though this error is [6] of the order of $l^2+\tau$.

About 10 sec of machine time is needed to obtain a solution at one point for t=1500 days, $l=1000$ m, and $\tau=100$ days (the boundary). This time becomes minutes if the difference scheme is used directly, since to obtain a value for the level at a point in the region at a given instant one has to calculate for all previous time steps. This means that the Monte Carlo method can be recommended when the state of groundwaters has to be predicted for a long time ahead. The choice of l, τ, and N should be based on providing the required accuracy with reasonable economy in machine time. This means that it is undesirable to increase N or reduce the grid spacing.

*All the calculations were performed by L. U. Kolner, graduate student at Novosibirsk University.

TABLE 2

N	l м	τ days	{1.1}	{2.2}	{3.3}	{4.4}	{5.5}	{6.6}	{7.7}
10^8	1000	100	39.22 (0.023)	37.27 (0.035)	34.94 (0.198)	32.60 (0.616)	31.38 (0.045)	30.55 (0.199)	30.15 (0.294)
	500	25	39.32	37.13	34.64	32.70	31.27	—	—
	250	6.25	39.10	37.10	34.74	32.16	31.11	—	—
10^4	1000	100	39.221 (0.020)	37.165 (0.247)	34.940 (0.198)	32.750 (0.159)	31.396 (0.006)	30.583 (0.091)	30.213 (0.036)
	500	25	39.139	37.215	34.614	32.640	31.345	—	—
	250	6.25	39.151	37.191	—	—	—	—	—
10^3	1000	100	39.09	36.76	34.46	32.86	31.50	30.72	30.35
	1000	300	38.94	36.76	34.51	32.49	31.63	30.78	30.29
	1000	700	38.70	36.43	33.85	32.57	31.50	30.89	30.47
	1000	1500	38.10	36.18	33.43	32.20	31.44	30.87	30.50
(2.4)	1000	100	39.229	37.257	34.871	32.802	31.394	30.611	30.239
(2.8)			39.199	37.138	34.721	32.673	31.314	30.552	30.219

3. Inexplicit Finite-Difference Scheme. A very simple scheme [6] may be used to represent (1.1) as

$$H_{i,j,s} = b_\tau H_{i,j,s-1} + b_l (H_{i-1,j,s} + H_{i+1,j,s} + H_{i,j-1,s} + H_{i,j+1,s}) \quad (i, j, s = 1, 2, \ldots) \tag{3.1}$$

Here

$$b_\tau = \left(1 + \frac{4kh^*\tau}{\mu l^2}\right)^{-1}, \qquad b_l = b_\tau \frac{kh^*\tau}{\mu l^2} \tag{3.2}$$

The structure of this scheme produces some features in the random walk. The particle at a node in time layer s can pass to one of the four adjacent nodes in the same layer with a probability $4b_l$, while the probability of reversion to the previous time layer is b_l.

As G takes the form of a square, the walk is unrestricted in two directions, and jump to the previous time layer is only one of several possible events, so there is a danger that the walk may extend too long in particular cases.

The distribution of the random numbers is such that the relative frequency of occurrence in the part of the 0-1 range corresponding to jump to the previous layer is approximately proportional to the length of that part. Then a particle starting from a certain node in layer s is predetermined to enter the initial layer (or to escape from one of the boundaries before this) on completing γ steps, whose mathematical expectation M_γ is

$$M\gamma = s / b_\tau$$

Let t_s be the instant corresponding to layer s. Then $s = t_s/\tau$. From (3.2)

$$M\gamma = t_s (1 + a^2\tau / l^2) / \tau \tag{3.3}$$

This inexplicit scheme is absolutely stable, so there is no need to consider (2.4) in choosing the step sizes. As step τ is increased with respect to time there occurs in accordance with (3.2) a redistribution of the probabilities b_l and b_τ, during which the former increases. This does not mean that the particle reaches the boundary more rapidly as τ increases, since it may recede from the boundary during the walk. More definitely, (3.3) indicates that the possibility of escape to the initial layer increases with τ, and the reduction in b_τ is more than balanced by the reduction in the number of time layers between the given layer and the initial one. The computation time is thus reduced as τ increases, but this gain is obtained at the expense of a loss of accuracy due to increase in the error of approximation.

This feature is reflected in the H from (3.1) of Table 2. The mean computation time per node for the first seven nodes was about 30 sec for N=1000 and τ=100 days in the inexplicit scheme. This exceeds the time in the explicit scheme, which also provides higher accuracy, and so the explicit one is to be preferred.

4. Hydraulic Coupling with Adjacent Horizons, and Effects of Boreholes and Rock Inhomogeneity. The essential features of the Monte Carlo method as applied in this case have already been presented [1, 2] for steady-state flows, and they can be transferred to transient states.

If the horizon has a hydraulic coupling to an underlying pressurized layer via a band of thickness m and filtration coefficient k_1, the equation for h differs from (1.1) in the presence on the right of a term $k_1/\mu m(H_2-h)$, in which H_2 is the head in the underlying bed. The explicit scheme gives us the following system of finite-difference equations:

$$H_{i,j,s}=\left(1-\frac{4a^2\tau}{l^2}-\omega\tau\right)H_{i,j,s-1}+\frac{a^2\tau}{l^2}(H_{i-1,j,s-1}+H_{i+1,j,s-1}+ \\ +H_{i,j-1,s-1}+H_{i,j+1,s-1})+\omega\tau H_2 \\ (\omega=k_1/\mu m,\ i,j,s=1,2,\ldots) \tag{4.1}$$

Here there is a further possible event: entry to the underlying horizon (probability $\omega\tau$). A similar situation has been considered [1, 2] for finding the particle in the borehole node, at which the head is given.

The following are values of H for $\{i,i\}$ diagonal nodes with $H_2=H_0=30$ m, N=1000, $t\approx1500$ days; the other parameters and boundary conditions are as in Kamenskii's problem, while k_1/m takes the following values (day^{-1}) for the first, second, and third lines; 4×10^{-3}, 4×10^{-4}, and 4×10^{-5}.

	{1,1}	{2,2}	{3,3}	{4,4}	{5,5}	{6,6}	{7,7}	{8,8}
Hm=	30.83	30.01	30.0	30.0	30.0	30.0	30.0	30.0
Hm=	33.62	32.15	30.21	30.07	30.0	30.0	30.0	30.0
Hm=	38.06	36.37	33.17	31.64	30.89	30.47	30.15	30.02

The exact values of t differ somewhat in each form, being multiples of τ, which in order to simplify (4.1) was chosen from

$$1-4a^2\tau/l^2-\omega\tau=0$$

Consider further the case where there is a borehole with a given flow rate. We distribute this rate evenly over a square cell of area l^2 centered on the $\{i,j\}$ node nearest to the borehold. We use the explicit scheme and choose the step size in accordance with (2.5) to get for node $\{i,j\}$, which we call [1, 2] the special one, that

$$H_{i,j,s}=\tfrac{1}{4}(H_{i-1,j,s-1}+H_{i+1,j,s-1}+H_{i,j-1,s-1}+H_{i,j+1,s-1})-(\tau/\mu l^2)Q\quad(s=1,2,\ldots)$$

Equations of the form of (2.6) apply for the other nodes.

Here the random walk has the feature that $(\tau/\mu l^2)$ Q=Q/600 m is subtracted from the sum of the accumulated penalty if the particle enters the special node, and the walk continues to exit at the boundary or the initial layer.

The following are values of H for N=1000 with the borehole at the $\{3,5\}$ node and having flow rates of 1000 m^3/sec (first line) and 10,000 m^3/sec (second line):

	{1,1}	{2,2}	{3,3}	{4,4}	{5,5}	{6,6}	{7,7}	{8,8}	{9,9}
Hm=	39.132	36.982	34.356	32.220	30.846	30.343	30.077	30.034	30.028
Hm=	38.712	35.116	29.450	24.437	24.955	27.941	29.419	29.880	30.023

If the rock is inhomogeneous, the probability of passage to adjacent nodes is related to the local k around the initial node, and this varies, which complicates the calculation. In the explicit scheme, l and τ should be chosen from the maximum k in accordance with (2.4).

5. Conclusions. The Monte Carlo method has the following features as applied to transient flows.

1. The method has been developed for linear differential equations, so its use in unpressurized infiltration is dependent on the availability of a suitable linearization.

2. The method has considerable advantages over finite-difference methods as regards time needed when the requirement is for long-term prediction.

3. An adequate number of walks from node $\{i, j, s\}$ may be employed with recording of the frequency of entry to boundary node r to determine statistically the probability $b^{(r)}_{i,j,s}$ of that event, which is dependent on the grid steps, mean flow rate, and hydrogeological parameters of the stratum. The values of $b^{(r)}_{i,j,s}$ for all boundary nodes allows one to use (2.3) to calculate $H_{i,j,s}$ for various H_r if h^* does not vary when H_r is altered. Shvidler (§13 of [2]) has pointed out this possible use of (2.3) as a fundamental solution.

LITERATURE CITED

1. M. I. Shvidler, "Monte Carlo solution of planar filtration problems," Izv. AN SSSR, OTN, Mekhanika i mashinostroenie, no. 1, 1963.
2. M. I. Shvidler, Filtration Flows in Inhomogeneous Media [in Russian], Gostoptekhizdat, Moscow, 1963.
3. V. A. Isyakaev, "Monte Carlo solution of a problem in spatial infiltration," PMTF [Journal of Applied Mechanics and Technical Physics], no. 2, 1967.
4. G. N. Kamenskii, "Prediction of groundwater conditions and waterlogging," Collection: Papers on Hydrogeology [in Russian], Trudy lab. gidrogeol. probl. im. F. P. Savarenskogo, vol. 20, 1958.
5. L. M. Pleshakova and V. G. Pryazhinskaya, "Some methods of numerical solution of a problem in three-dimensional unsteady percolation," PMTF [Journal of Applied Mechanics and Technical Physics], no. 2, 1965.
6. V. K. Saul'ev, Integration of Equations of Parabolic Type by the Method of Grids [in Russian], Fizmatgiz, Moscow, 1960.
7. N. P. Buslenko, D. I. Golenko, I. M. Sobol, V. G. Stragovich, and Yu. A. Shreider, The Monte Carlo Method [in Russian], Fizmatgiz, Moscow, 1962.
8. V. F. Lyashenko, Programming for the M-20, BÉSM-3M, BÉSM-4, and M-220 Computers [in Russian], Izd. Sov. radio, Moscow, 1967.