Reprinted from Ground Water Proceedings
Int'l Symposium/Hy Div/ASCE/Nashville, TN/July 29-Aug. 2, 1991

Prediction Method of Concentration Distribution of Groundwater Pollutants Using a Field Monitoring Network System

Akira Kawamura[1], Dr. Eng.
Kenji Jinno[2], Dr. Eng.

Abstract

A method based on the extended Kalman filter (EKF) is developed to identify the parameters of a two-dimensional stochastic convective-dispersion (SCD) differential equation and to predict in real-time the spatial concentration distribution of groundwater pollutants using a field monitoring network system. To study the various characteristics of this method, it is applied to a synthetically generated concentration distribution. Results show that this method provides an effective parameter identification, thereby giving accurate predictions of the concentration distribution. The factors which effect the accuracy of the prediction are also studied.

Introduction

Recently, it has been reported that groundwater pollution caused by chlorinated hydrocarbons (Jinno et al., 1986), which are used in high-technology industries, and agricultural chemicals like pesticides are increasing not only in Japan but all over the world. In order to take countermeasures against groundwater pollution, it is indispensable to identify the source of the pollutants immediately and to predict the concentration distribution precisely. The transport of groundwater pollutants is known as a SCD phenomenon, caused by random groundwater flow. When this phenomenon is analyzed, the initial and boundary conditions are assumed known beforehand. For practical problems, however, the initial and boundary conditions are often unknown, and the parameters of the convective-dispersion equation should be treated in a

---

[1]Research Associate, [2]Associate Professor,
Department of Civil Engineering(SUIKO),Kyushu University,
6-10-1 Hakozaki, Higashi-ku, Fukuoka 812, Japan

stochastic way (Jinno et al.,1989). Numerical simulation by finite difference or finite element methods is commonly employed for the prediction of the concentration distribution of pollutants. These methods generally treat the parameters as known and need to assume the deterministic initial and boundary conditions. These assumptions are ambiguous for natural groundwater flow systems. Also, measured data of concentration of the pollutants at each observation point in the field should be effectively used for improving the prediction accuracy by a feedback system.

In order to take into account all above factors, an attempt is made to develop a method in which the identification of parameters of a two-dimensional SCD differential equation is done simultaneously with the prediction of concentration distribution using the EKF. The method involves a monitoring algorithm based on a feedback system, and the observation of the concentration in a field monitoring network system is effectively used for both prediction of the transport of pollutants and parameter identification in the equation. To study the various characteristics of the proposed method, it is applied to a synthetically generated concentration distribution. The effect of the number and the arrangement of observation stations and the number of terms in the double Fourier series expansion (DFSE) on the accuracy of the prediction is also discussed in this study.

Formulation for Prediction of Concentration

When the x-axis is taken along the direction of the mean groundwater flow, the two-dimensional, constant coefficient SCD equation is expressed as follows:

$$\frac{\partial C(x,y,t)}{\partial t} + u\frac{\partial C(x,y,t)}{\partial x} = D_x\frac{\partial^2 C(x,y,t)}{\partial x^2} + D_y\frac{\partial^2 C(x,y,t)}{\partial y^2} - \gamma C(x,y,t) + \varepsilon(x,y,t) \quad (1)$$

where C(x,y,t) is concentration, u is the mean velocity, D_x and D_y are coefficients of dispersion for the x- and y-directions respectively, γ is coefficient of attenuation, $\varepsilon(x,y,t)$ is the zero-mean Gaussian white noise, x and y are the locations and t is time. The physical parameters (PPs) u, D_x,D_y and γ are assumed to be constant in space. The last term in equation (1) is added to take into account the uncertainties which are inherent in modeling the phenomenon.

The stochastic partial differential equation (1) is transformed into an ordinary differential equation using DFSE. Specifically, C(x,y,t) and $\varepsilon(x,y,t)$ are expanded as:

$$C(x,y,t) = \frac{A_{00}}{2} + \sum_{m=1}^{M}\sum_{n=0}^{N}\left[A_{mn}(t)\cos F_1(x,y,m,n) + B_{mn}(t)\sin F_1(x,y,m,n)\right]$$

$$+\sum_{m=0}^{M}\sum_{n=1}^{N}\left[C_{mn}(t)\cos F_2(x,y,m,n)+D_{mn}(t)\sin F_2(x,y,m,n)\right] \quad (2)$$

$$\varepsilon(x,y,t)=\sum_{m=1}^{M}\sum_{n=0}^{N}\left[E_{mn}(t)\cos F_1(x,y,m,n)+F_{mn}(t)\sin F_1(x,y,m,n)\right]$$
$$+\sum_{m=0}^{M}\sum_{n=1}^{N}\left[G_{mn}(t)\cos F_2(x,y,m,n)+H_{mn}(t)\sin F_2(x,y,m,n)\right] \quad (3)$$

$$F_1(x,y,m,n)=2\pi mx/\lambda_x+2\pi my/\lambda_y,\; F_2(x,y,m,n)=2\pi mx/\lambda_x-2\pi my/\lambda_y \quad (4)$$

where A_{mn}, B_{mn}, C_{mn}, D_{mn}, E_{mn}, F_{mn}, G_{mn}, H_{mn} are Fourier coefficients (FCs), M and N are the numbers of terms in the DFSE, λ_x and λ_y are wavelengths for x- and y-directions respectively. Substitution of equation (2) and (3) into equation (1) yields a set of ordinary differential homogeneous equations with respect to wave numbers m and n as follows:

$$\begin{bmatrix} dA_{mn}(t)/dt \\ dB_{mn}(t)/dt \\ dC_{mn}(t)/dt \\ dD_{mn}(t)/dt \end{bmatrix} = \begin{bmatrix} -P_{mn} & -Q_m & 0 & 0 \\ Q_m & -P_{mn} & 0 & 0 \\ 0 & 0 & -P_{mn} & -Q_m \\ 0 & 0 & Q_m & -P_{mn} \end{bmatrix} \begin{bmatrix} A_{mn}(t) \\ B_{mn}(t) \\ C_{mn}(t) \\ D_{mn}(t) \end{bmatrix} + \begin{bmatrix} E_{mn}(t) \\ F_{mn}(t) \\ G_{mn}(t) \\ H_{mn}(t) \end{bmatrix} \quad (5)$$

where

$$P_{mn} = D_x(2\pi m/\lambda_x)^2+D_y(2\pi n/\lambda_y)^2+\gamma\,,\quad Q_m = u(2\pi m/\lambda_x) \quad (6)$$

Under unknown initial and boundary conditions, the concentration distribution is predicted using the EKF on the basis of information from the observation stations distributed in the area, i.e., the field monitoring network system. The unknown PPs and the coefficients of the DFSE are identified in the prediction process by the EKF (Athans et al., 1968; Kawamura et al., 1986). The system state vector X is to be estimated by the EKF model as follows:

$$X(t) = \left[u\ D_x\ D_y\ \gamma\ A_{00}(t),\cdots,A_{mn}(t),B_{mn}(t),C_{mn}(t),D_{mn}(t),\cdots\right]^T \quad (7)$$

where the symbol ^{T} indicates transposed matrix. The number of elements of X(t) is (4MN+2M+2N+5). Here, the dynamics of the PPs u, D_x, D_y, and γ are assumed constant in time for the EKF formulation such that:

$$du/dt = 0,\quad dD_x/dt = 0,\quad dD_y/dt = 0,\quad d\gamma/dt = 0 \quad (8)$$

However, this does not mean that the PPs do not change in time through the parameter identification process by the EKF (Kawamura et al., 1986). Combining equations (5) and (8), the system equation is obtained as follows:

$$dx/dt = f(X(t)) + v(t) \quad (9)$$

Equation (9) is a set of nonlinear functions of u, D_x, D_y, γ, A_{00}, A_{mn}, B_{mn}, C_{mn}, and D_{mn}. Here the Taylor series expansion of f(X(t)) is carried out, and more than or equal to second-order terms of the series are neglected in order to linearize so that the EKF can be applied.

The observation vector of the EKF is the concentration distribution measured at spatially arranged observation stations at some sampling time interval with observation noise w(x,y,t). The concentration distribution given by equation (2) is written as:

$$C(x,y,t)=[\ 0\ \ 0\ \ 0\ \ 0\ \ 1/2\cdots\cos F_1(x,y,m,n),\ \sin F_1(x,y,m,n),$$
$$\cos F_2(x,y,m,n)\ ,\ \sin F_2(x,y,m,n)\ ,\cdots]X(t)\ +\ w(x,y,t) \quad (10)$$

Details of the EKF are described by Kawamura et al.(1986) or Jinno et al.(1990).

Simulation Results

Figure 1 shows the contour map (solid lines) generated by equation (1) and disturbed by the random noise ε. Here 10 steps equal 1 day. The following values were used for the synthetic generation: u=1.0(m/day), D_x=3.0(m^2/day), D_y= 1.0(m^2/day), γ=0(1/day), λ_x=150(m), λ_y= 100(m), and M=N=10. The Gaussian random noise E_m and F_m were assumed to have a mean of zero and with a standard deviation equal to 0.0005(g/m^3). Equation (5) was solved numerically using a finite difference method. The discrete time interval is Δt = 0.1 day. Here, the initial concentration distribution for the simulation study was the analytical solution of equation (1) (excluding the term ε) at t = 10 days under the following initial condition:

$$C(x,y,0)=\begin{cases} C_0 & (x_1\leq x\leq x_2\ ,\ y_1\leq x\leq y_2) \\ 0 & (0\leq x<x_1\ ,\ x_2<x\leq\lambda_x\ ,\ 0\leq y<y_1\ ,\ y_2<y\leq\lambda_y) \end{cases} \quad (11)$$

where, x_1, x_2, y_1 and y_2 were set to 19.5, 20.5, 24.5, 25.5(m), respectively. When the distribution was assumed observable at some observation points, an observation noise w(x,y,t) with a mean of zero and a standard deviation equal to 0.1(g/m^3) was added as shown in equation (10). In this simulation, the 36 observation points to monitor the concentration.are shown as cross-marks in Figure 1 Observed values were obtained every 10 steps, i.e., once a day.

The following initial conditions for the applications of the EKF were used in the simulation. The initial values of the PPs and the FCs were assumed to be 50% of the true values. The diagonal elements of the system noise covariance matrix for EKF were taken as 1% of the initial state vector X(0) for PPs, 5% for FCs, and off-diagonal as zero. The broken lines of Figure 1 show the predicted 10-step ahead concentration distribution. Figure 2 shows some of the identified PPs and FCs. Figure 3 shows observed values (vertical bars) and predicted values (circle and cross marks) at two observation stations out of 36 (No.1 and No.2, respectively) whose locations are shown in Figure 1.

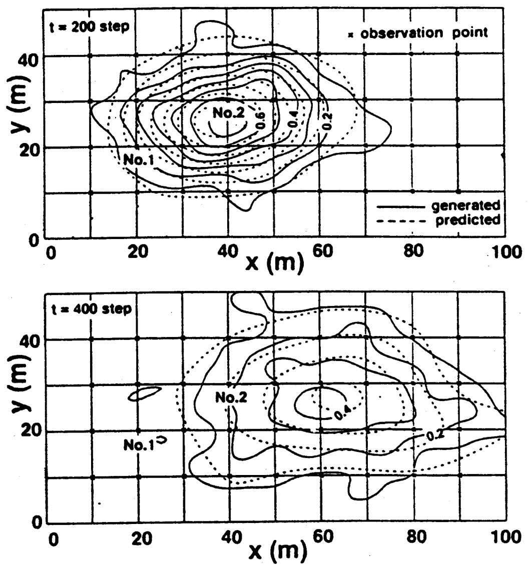

Figure 1. Generated and Predicted Concentration Distribution

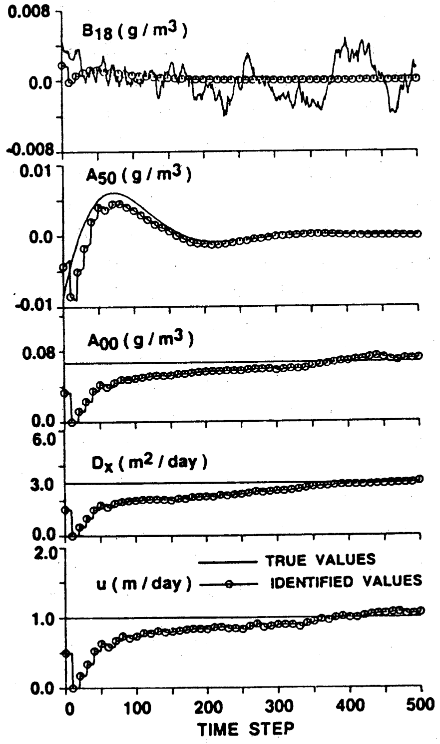

Figure 2. Identification of Physical Parameters and Fourier Coefficients

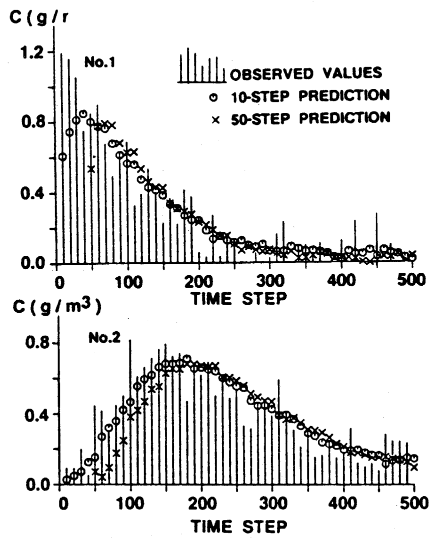

Figure 3. Observed an Predicted Values at the Observation Stations

Discussion and Summary

Figure 1 shows that the disturbance by the noise ε in equation (1) becomes dominant as time passes. The predicted 10-step ahead distributions are shown to be accurate, except for the fluctuations corresponding to the high-frequency components.

Figure 2 shows the process of identification in some PPs and FCs. They converge to true values gradually except for B_{18}. The convergence depends on the estimation covariance of the system noise for the EKF. If the covariance is set larger, the convergence tends to fluctuate. The true values of coefficient B_{18} fluctuate depending on the disturbance by the system noise ε, but the EKF identifies the main trend only, and not effects of the random noise. These characteristics are the same as for the other coefficients. FCs of wave numbers higher than five become less than those of the ε as time passes. This means that the contribution of the high wave numbers to the concentration distribution is small.

As shown in Figure 3, both 10-step and 50-step ahead predictions become accurate soon after the initial errors of PPs and FCs given as 50% of the true values. These are converged around the 50th step, even though the observed values contain Gaussian white observation noise.

Beside the above results, the following characteristics were obtained through this study. The factors which affect the accuracy of prediction and identification are firstly the initial values for PPs and FCs, and secondly the number and the arrangement of observation stations in the field. An accurate initial guess, sufficient number and arrangement of the observation stations are key factors for accurate prediction and identification. The estimation covariance of system noise for the EKF is also influential especially for the accuracy of the identification of PPs. It was also found in this study that the influence of the number of terms in the DFSE is small as long as more than five terms in each direction is used.

Acknowledgement

This research was supported by the Tojiro Ishihara Scholarship Fund awarded in 1990.

References

Athans,M., Wishner,R.P. & Bertolini,A. (1968) : Suboptimal state estimation for continuous-time nonlinear system from discrete noisy measurement. IEEE Trans. on Automatic Control AG-13 (5), 504-514.

Jinno,K., Ueda,T.& Yasuda,H. (1986) : Analytical approach to monitoring pollution by chlorinated solvents. In: Monitoring to Detect Changes in Water Quality Series (Proc. Budapest Symp., July 1986), 185-196, IAHS Publ. no.157.

Jinno,K., Kawamura,A., Ueda,T. & Yoshinaga,H. (1989) : Prediction of the concentration distribution of groundwater pollutants. In : Groundwater Management : Quantity and Quality (Proc. of Benidorm Symp., October 1989), 131-142, IAHS Publ. no.188.

Kawamura,A., Jinno,K., Ueda,T. & Medina,R.R. (1984) : Detection of abrupt changes in water quality time series by the adaptive Kalman filter, In: Monitoring to Detect Changes in Water Quality Series (Proc. Budapest Symp., July 1986), 285-296, IAHS Publ. no.157.