

# 회귀모형을 이용한 지하수문계의 매개변수 추정에 관한 연구

A Study on the Parameter Estimation of Groundwater  
Hydrologic System by means of Regression Modeling

이 원 환 \*, 조 원 철 \*\*, 정 일 문 \*\*\*

Lee, Won Hwan, Cho, Won Cheol, Jung, Il Moon

## 1. 서론

지하수 흐름 모형은 수학적 모형의 하나인데 이는 흐름계의 물리적인 모형을 수학적인 변수, 매개변수, 상수를 포함하는 표현으로 바꾼 것이다<sup>7)</sup>. 따라서 수학적모형은 물리적인 계의 단순화를 포함한다. Krumbein과 Graybill<sup>17)</sup>은 수학적 모형을 세가지 형태로 분류하였는데 그것은 결정론적 모형과 통계학적 모형, 그리고 추계학적 모형이다. 결정론적 모형은 종속변수들이 독립변수, 매개변수, 상수를 포함하는 표현으로 부터 정확하게 계산되는 것이다. 통계학적 모형은 하나 혹은 더 많은 무작위 성분이 더해지는 결정론적 모형이다. 추계학적 모형은 통계학적 모형과 유사하게 사용되며 통계학적 과정에서 추가된 무작위과정에 추계학적인 과정을 포함시킨다. 결정론적 모형과 통계학적 모형, 그리고 추계학적 모형으로 크게 대별되는 수학적 모형이 지하수 흐름 모형에 적용될 때는 위 세가지 모형의 특징들을 적절히 살리는 것이 빈약한 자료와 또한 그 자료값의 신빙성 여부를 검토함에 있어서도 보다 정확한 결과를 산출할 수 있다고 하겠다. 이러한 문제들은 대수층의 수문지질학적매개변수(수리전도율, 투수량계수, 저류도)와 생성 및 소멸량, 경계유동량에 대한 매개변수 추정치를 구하는 방법을 연구함으로써 재현되어 왔다. 현장 대수층에서의 공간적으로 변화하는 투수성의 산출은 Gutjahr 등<sup>6)</sup>, Smith 와 Freeze<sup>8)</sup>등에 의해 추계학적과정을 고려함으로써 연구되어 왔으며 Cooley는 정상 지하수 흐름에 대한 매개변수추정과 신뢰도산정방법의 이론 및 수치특성과 통계분석에 관한 연구를 발표하였다<sup>3, 4)</sup>. 본 연구는 Cooley와 Naff가 개발한 회귀모형<sup>2)</sup>을 분석하여 실제지역에 적용하여 봄으로써 그 지역의 대수층 현황을 파악하여 결정론적모형의 입력치로 주어지는 여러 매개변수를 추정함에 그 목적을 두었다.

---

\* 정회원, 연세대학교 토목공학과 교수

\*\* 정회원, 연세대학교 토목공학과 부교수

\*\*\* 정회원, 연세대학교 대학원 토목공학과 박사과정

## 2. 기본이론

모형이 매개변수내에서 비선형이지만 종속변수에 대하여 선형일 때 모형은 다음과 같은 표준 비선형 회귀형태로 쓰여진다.

$$Y = f(\xi_1, \xi_2, \dots, \xi_k; \beta_1, \beta_2, \dots, \beta_p) + \varepsilon \quad (1)$$

n개의 관측점이 있을 때 식(1)은 다음과 같은 형태로 나타난다.

$$Y = f(\xi, \beta) + \varepsilon \quad (2)$$

$Y, f, \xi, \beta, \varepsilon$  등은 모두 행렬의 형태이다. (2)식은 일반적인  $\beta$ 의 추정치  $b$  와  $\varepsilon$ 의 추정치  $e$  의 항으로 나타낼 수 있다.

$$Y = f(\xi, b) + e \quad (3)$$

선형인 경우에는 식(3)에 대한 회귀해는 가중된 잔차의 제곱합을 최소화함으로 다음과 같이 얻을 수 있다.

$$S(b) = e^T w e = (Y - f(\xi, b))^T w (Y - f(\xi, b)) \quad (4)$$

그러나 식(3)이 비선형일 경우 문제의 해는 선형의 경우와 같이 직접적이지 못하다. 따라서 초기변수들에 대한 선형근사화가 이루어진다. 방법을 유도하기 위해서 식(3)은 절단된 Taylor급수에 의해 초기 변수  $b_0$ 에 대해 확장된다.

$$f(\xi, b) \cong f(\xi, b_0) + X_0(b - b_0) \quad (5)$$

$$X_0 = \{ X_0^{ij} \} = \left\{ \frac{\partial f_i}{\partial b_j} \Big|_{b=b_0} \right\}_{(n \times p)} \quad (6)$$

$f_i$ 는  $i$ 점에서의 계산값이고  $X_0$ 는 민감도 계수행렬이다.

식 (5)를 이용하여 식(3)은 다음과 같이 나타난다.

$$Y - f(\xi, b_0) \cong X_0(b - b_0) + e \quad (7)$$

이것은 선형모형에 대한 증분된 표현이다. 모형이 비선형일때 식(7)은  $b_0$  부근에서 매개변수에 대해 선형으로 근사화된 모형이 된다.  $\beta$ 에 대한 근사화된 최선의 추정치는 다음식을 최소화함으로  $b$ 에 대해 얻을 수 있다.

$$S(b) = e^T w e \cong (Y - f(\xi, b_0) - X_0(b - b_0))^T w (Y - f(\xi, b_0) - X_0(b - b_0)) \quad (8)$$

이 과정이 선형모형인 경우와 같이 수행되고 정규방정식을 산출한다.

$$X_0^T w X_0 d_1 = X_0^T w (Y - f(\xi, b_0)) \quad (9)$$

여기서 첨자 1은 첫번째 근사해를 나타낸다. 그리고

$$d_1 = b_1 - b_0 \quad (10)$$

를 나타낸다. 비선형 회귀분석의 반올림 오차를 줄이기 위해 식(2)를 축척화 (scaling)하는 과정이 사용된다. 이 방법은 선형 회귀모형의 경우와 동일한 것으로 식(11)에서 (16)까지 나타나 있다.

$$C_0^T X_0^T w X_0 C_0 C_0^{-1} d_1 = C_0^T X_0^T w (Y - f(\xi, b_0)) \quad (11)$$

$$S_0^T w S_0 \delta_1 = S_0^T w (Y - f(\xi, b_0)) \quad (12)$$

$$S_0 = X_0 C_0 \quad (13)$$

$$\delta = C_0^{-1} d_1 \quad (14)$$

$$C_0 = \text{diag}\{(A_{11})^{0-\frac{1}{2}}, (A_{22})^{0-\frac{1}{2}}, \dots, (A_{pp})^{0-\frac{1}{2}}\} \quad (15)$$

$$A_0 = X_0^T W X_0 \quad (16)$$

$A_{ii}^0$ 는  $A_0$ 의 대각선 성분이다. 방정식 (9)가 정확하지 않으므로 방정식 (4)는 최소화된 참값이라고는 볼 수 없으며  $b_1$ 역시 최적의 집합이라고는 볼 수 없다. 따라서  $b_1$ 이  $b_0$ 에 대입되어 지고 또 다른 개선된 추정치를 산출할 때 까지 과정은 계속적으로 반복된다. 일반적인 반복계산식이 다음과 같이 표현된다.

$$S_r^T W S_r \delta_{r+1} = S_r^T W (Y - f(\xi, b_r)) \quad (17)$$

$$S_r = X_r C_r \quad (18)$$

$$\delta_{r+1} = C_r^{-1} (b_{r+1} - b_r) = C_r^{-1} d_{r+1} \quad (19)$$

$$C_r = \text{diag}\{(A_{11})^{r-\frac{1}{2}}, (A_{22})^{r-\frac{1}{2}}, \dots, (A_{pp})^{r-\frac{1}{2}}\} \quad (20)$$

여기서,  $A_{ii}^r$ 은  $A_{ii}^0$ 과 상사성을 이룬다.  $\delta_{r+1} \rightarrow 0$ 일 때 방정식 (7)은  $(Y - f(\xi, \hat{b})) = \hat{e}$  이 된다. 동시에  $\hat{b}$ 는 방정식 (4)의  $S(\hat{b})$ 를 최소화시킨다. 즉,  $S(\hat{b}) = (Y - f(\xi, \hat{b}))^T W (Y - f(\xi, \hat{b}))$ 가 비선형 회귀분석을 최소화한다. 이처럼  $S(b)$ 를 최소화 하는 과정을 Gauss-Newton 방법이라 한다. 방정식 (17)로 정의된 Gauss-Newton 방법은 언제나 수렴하는 것은 아니다. 따라서 수렴을 유도하기 위해 완충변수(Damping parameter)  $\rho$ 가 도입된다.

$$b_{r+1} = \rho d_{r+1} + b_r \quad (21)$$

여기서,  $d_{r+1} = C_r \delta_{r+1}$  이다. 만일  $0 < \rho < 1$  이면 계산된 매개변수치의 변화는  $\rho=1$ 일때 보다 작아진다.  $d_{r+1}$ 의 개선과정에서 양의 매개변수  $\mu$ 값을 계수행렬의 주 대각요소에 구성하게 되는데 이를 Marquardt 매개변수라 하며 수학적으로 다음과 같이 표현된다.

$$(S_r^T W S_r + \mu I) \delta_{r+1} = S_r^T W (Y - f(\xi, b_r)) \quad (22)$$

해를 구하는 절차를 간단히 정리하면 다음과 같다.

- 가.  $f(\xi, b_r)$ 와  $S_r$ 를 초기매개변수  $b_r$ 과 식(6)과 (18)을 결합시켜 구한다.
- 나.  $\delta_{r+1}$  를 구하기 위해 식(22)을 푼다.
- 다.  $d_{r+1}$  를 구하기 위해 식(19)를 푼다.
- 라.  $b_{r+1}$  를 구하기 위해 식(21)을 푼다.
- 마. 만일  $|d_i^{r+1}| > c|d_i^r|$  인지 검사한다. 여기서  $\epsilon$ 는 0.01등과 같은 작은 수이며  $b_i^r \neq 0$  에 대해  $c = b_i^r / b_i^r = 1$  에 대해  $c=1$  이다.
- 바. 만일  $|d_i^{r+1}| > c|d_i^r|$  이면 증가치  $r$ 는 1로 되돌아 간다. 그렇지 않으면 과정은 수렴된다.

$$D(\mathbf{h}, \xi, \beta) \mathbf{h} = \mathbf{q}(\mathbf{h}, \xi, \beta) \quad (23)$$

식(23)은 종속변수  $\mathbf{h}$ 가  $m$ 차의 해 벡터인 항을 포함한 비선형 행렬방정식이다.  $D$ 는  $m$ 차의 정칙계수행렬이며  $\mathbf{h}, \xi$ , 그리고  $\beta$ 의 함수이다. 그리고  $\mathbf{q}$ 는  $\mathbf{h}, \xi$ , 그리고  $\beta$ 의 함수인  $m$ 차의 벡터이다.  $m$ 차는 관측점수  $n$  과는 무관하지만 문제의 해의 수치근사를 좋게 하는데 필요하다. 만약 방정식 (23)이  $\mathbf{h}$ 에 대해 선형이고 따라서  $D$ 와  $\mathbf{q}$ 가  $\mathbf{h}$ 의 함수가 아니라면 방정식은  $\mathbf{h}$ 에 대해 직접적으로 풀린다. 이러한 경우에 회귀해를 구하기 위해 Gauss-Newton 방법이 사용된다. 그 절차는 다음과 같다.

$$\mathbf{h} = D^{-1} \mathbf{q} \quad (24)$$

이는 변수  $\mathbf{h}$ 에 대해 양해적이다. 여기서  $\mathbf{h}$ 와 식(2)의  $\mathbf{f}$ 는 같은 물리량의 표현이며, 단지 그 차이라고 하면  $\mathbf{h}$ 의 요소들은 수치해로부터 구해진 점에서의 종속변수 값이며,  $\mathbf{f}$ 는 관측점에서 계산 되어진 종속변수값이다.

Gauss-Newton해를 형성하기 위한 최종단계는 민감도 행렬  $X$ 를 유도하는 것이다. 이를 수행하기 위해 식(24)를 일반적인 매개변수 집합  $\mathbf{b}$ 의 항으로 쓰고 다시  $\mathbf{b}$ 에 관해 편미분한다.

$$D \frac{\partial \mathbf{h}}{\partial b_j} + \frac{\partial D}{\partial b_j} \mathbf{h} = \frac{\partial \mathbf{q}}{\partial b_j}, \quad j=1, 2, \dots, p \quad (25)$$

혹은,

$$\frac{\partial \mathbf{h}}{\partial b_j} = D^{-1} \left[ \frac{\partial \mathbf{q}}{\partial b_j} - \frac{\partial D^{-1}}{\partial b_j} \mathbf{h} \right], \quad j=1, 2, \dots, p \quad (26)$$

$\partial \mathbf{h} / \partial b_j$ 의 값은  $m$ 지점에 대한 민감도 행렬의 열을 형성한다. 민감도행렬  $X_r$ 은 (26)식의 방정식을 이용하여 첫번째  $\partial \mathbf{h} / \partial b_j$ 를 계산함으로써 밝혀지고  $\mathbf{h}$ 로부터  $\mathbf{f}$ 를 얻기 위해 위에 기술된 두 가지 절차중의 하나를 따른다.  $\mathbf{f}(\xi, \mathbf{b}_r)$ 과  $X_r$ 의 수치 추정치를 구하는 과정을 통합함으로써 Gauss-Newton방법에 대해 주어진 알고리즘은 (24)식의 회귀해를 정확하게 밝히기 위해 뒤따른다.

### 3. 지배방정식 및 적분된 유한차분모형

대부분의 지하수 흐름의 일반적인 흐름 방정식이 다음과 같이 주어진다.

$$\frac{\partial}{\partial x} \left[ T_{xx} \frac{\partial h}{\partial x} \right] + \frac{\partial}{\partial y} \left[ T_{yy} \frac{\partial h}{\partial y} \right] + R(H-h) + W + \sum_{L=1}^N \delta(x-a_L) \delta(y-b_L) Q_L = S \frac{\partial h}{\partial t} \quad (27)$$

$T_{ii}(x, y)$  =  $x, y$ 방향의 투수량계수 ( $K_{ii} \times b$ )로서  $x, y$ 의 연속함수

$K_{ii}(x, y)$  =  $i$  방향의 대수층의 수리전도도(Hydrarulic Conductivity)

$b(x, y)$  = 대수층의 두께

$R(x, y)$  = 하상 밑에 놓인 토사 혹은 대수층 위 아래의 준 대수층 (aquitard)의 두께로 나눈 수리전도계수

$W(x, y, t)$  = 면적분포된 source-sink 항

$\sum_{L=1}^N \delta(x-a_L)\delta(y-b_L)Q_L = N$ 양수정에 대한 Dirac delta, 각 양수정은  $(a_L, b_L)$  에서  $Q_L(t)$ 의 양수율로 양수

$S(x, y)$  = 저류계수

$h(x, y, t)$  = 대수층의 수리수두

$H(x, y, t)$  = 피압하상바닥이나 준 대수층(aquitard) 끝에서의 수두

$x, y$  = 직교좌표

$t$  = 시간

방정식 (27)로 부터  $\partial h/\partial t=0$ 이라고 가정함으로 분석할 문제에 대한 지하수 흐름 지배방정식은 다음과 같이 표현된다.

$$\frac{\partial}{\partial x} \left[ T_{xx} \frac{\partial h}{\partial x} \right] + \frac{\partial}{\partial y} \left[ T_{yy} \frac{\partial h}{\partial y} \right] + R(H-h) + W + \sum_{L=1}^N \delta(x-a_L)\delta(y-b_L)Q_L = 0 \quad (28)$$

본 연구에 사용된 수치모형은 방정식 (28)에 대한 해를 구하기 위해 도입 되는데 적분된 유한차분법을 사용한다. 격점들로 이루어진 직사각형 격자는 그림 1과 같이 나타난다. 각각의 격점들은 부영역에 둘러싸여 있는데, 이 부영역들은 인접한 격점간의 절반을 취한 모서리를 연결한 직사각형 영역이다. 전형적인 격점의 좌표  $(i, j)$ 는  $(x_i, y_j)$ 와 같이 주어진다. 그림 1에 나타난 것과 같이 방정식 (28)은 격점  $(i, j)$ 를 둘러싼 부영역에 대해 적분가능하다.

$$\int_{\Delta y_j} \left[ T_{xx} \frac{\partial h}{\partial x} \right]_{i+1/2} dy - \int_{\Delta y_j} \left[ T_{xx} \frac{\partial h}{\partial x} \right]_{i-1/2} dy + \int_{\Delta x_i} \left[ T_{yy} \frac{\partial h}{\partial y} \right]_{i+1/2} dx - \int_{\Delta x_i} \left[ T_{yy} \frac{\partial h}{\partial y} \right]_{i-1/2} dx + \int_{\Delta x_i} \int_{\Delta y_j} R(H-h) dx dy + \int_{\Delta x_i} \int_{\Delta y_j} W dx dy + \sum_{p=1}^{N_{ij}} Q_p = 0 \quad (29)$$

$$\left. \begin{aligned} \Delta x_i &= \frac{1}{2} (\Delta x_{i+1/2} + \Delta x_{i-1/2}) \\ \Delta y_j &= \frac{1}{2} (\Delta y_{j+1/2} + \Delta y_{j-1/2}) \end{aligned} \right\} \quad (30)$$

그리고,  $N_{ij}$ 는 부영역  $\Delta x_i \Delta y_j$  내의 양수정의 수이다.

#### 4. 모형의 적용

본 연구의 대상지역은 고창군 대산면 일대로 관측정 및 양수정이 밀집된 지역으로 대부분이 층적관정이다<sup>1)</sup>. 각 관정들의 개발내역자료로 부터 얻은 17지점의 지하수위 실제 관측값으로부터 지하수 등고선을 구하여 육안으로 선형보간한 56개 지점의 지하수위 관측치와 3종류의 영역으로 구분된 투수량계수 및 피압하상의 수리전도율, 충진율, 생성과 소멸량, 경계유동량등이 입력자료로 사용되었다. 피압하상의 수리전도율(HC)은 지질을 실트질로 가정한 값으로, 지하수 충전(Q)은 강우에 의한 주입이 대부분임을 감안하여 이 지역의 연평균 강우량(1,266.8mm)으로 부터 일평균 강우량을 산정하여 이를 다시 지표로 침투하는 양으로 변환시켜 사용하였다. 입력에 사용된 값들이 표 1에 나타나 있다.

표 1. 대수층 매개변수의 입력자료치(초기값)

영역별 대수층 매개변수의 초기치				
영역	T <sub>xx</sub> (m <sup>2</sup> /s)	T <sub>yy</sub> (m <sup>2</sup> /s)	HC (d <sup>-1</sup> )	Q (m/day)
1	177.00	177.00	0.10000E-02	0.10000E-03
2	20.000	20.000	0.10000E-02	0.50000E-03
3	60.000	60.000	0.50000E-02	0.10000E-03

경계수위는 3지점에 대해 기지값을 부여했으며 경계유동량은 관측정간의 동수경사와 영역의 투수량계수를 사용하여 초기치를 정하였다. 수치모형을 적용하기 위한 격자망은 16×7로 구성하였으며 영역분할은 3부분으로 하였다. 수치비선형 회귀분석으로 구한 매개변수 집합이 표2에 나타나 있고 이 값에 의한 최종 계산수두값을 관측수두와 비교하여 본 결과 그림 2와 같은 결과를 산출하였다.

표 2. 추정된 적정 매개변수(오차의 제곱합=120.05, 오차분산=2.61)

번호	변수명	변수의 추정치	오차의 제곱합
1	T <sub>1</sub>	188.27	78.098
2	T <sub>2</sub>	14.299	19.561
3	T <sub>3</sub>	26.533	34.714
4	R <sub>1</sub>	0.18803E-02	0.20504E-02
5	R <sub>2</sub>	0.22501E-02	0.33536E-02
6	R <sub>3</sub>	0.12425E-02	0.16474E-02
7	W <sub>1</sub>	-.92747E-03	0.12150E-02
8	W <sub>2</sub>	-.13362E-02	0.42874E-02
9	W <sub>3</sub>	-.56269E-03	0.98880E-03
10	QB1	0.44831	0.95956
11	QB2	0.81895	1.0927
12	QB3	0.14228	0.37308
13	h <sub>B1</sub>	35.750	1.0023
14	h <sub>B2</sub>	13.102	0.97354
15	h <sub>B3</sub>	10.934	1.7025

## 5. 결론

본 연구는 정상상태의 지하수 흐름을 지배하는 방정식내의 여러가지 수문 지질학적 대수층 변수, 생성 및 소멸량, 경계유동량등을 유한차분화된 2차원 영역에 대한 수치 비선형 회귀분석을 이용하여 매개변수 추정치를 찾아내는 모형을 적용한 것으로서 다음과 같이 요약될 수 있다.

1. 관측수두와 계산수두와의 상관도는 0.98로 적합되었다.
2. 1과 같은 상관성을 유지하면서 전체적인 지하수문제에 대한 변수들의 수렴을 유도해 내는 15개의 대수층 및 흐름 매개변수를 구할 수 있었다.
3. 지형의 급변으로 수두변화가 큰 지역에 대해서는 수치해의 정확도가 낮게 나타났으며 이는 격자의 크기와 관측수두산정시 발생하는 육안보간 오차의 영향 때문일 것으로 사료된다.
4. 본 모형으로 기존의 지점추출로 부터 구해지는 투수량계수의 지역적인 분포 및 여러 대수층 변수를 예측하여 지하수의 분포양상을 구할 수 있었다.

## 6. 참고문헌

1. 농어촌 진흥공사, 『고창지구 광역수리지질 시범조사 보고서』, 1991.
2. Cooley, R.L. and R.L. Naff., "Regression Modeling of Groundwater Flow", Techniques of Water Resources Investigations of U.S.G.S., Book 3, chapt B4, 1990.
3. Cooley, R.L., "A method of estimating parameters and assessing reliability for models of steady-state groundwater flow, 1-Theory and numerical properties", Water Resources Research, vol.13, No.2, pp. 318-324, 1977.
4. Cooley, R.L., "A method of estimating parameters and assessing reliability for models of steady-state groundwater flow, 2-Application of statistical analysis", Water Resources Research, vol.15, no.3, pp. 603-617, 1979.
5. Freeze, R.A. and J.A. Cherry, Groundwater, Prentice Hall, pp.14~79, 1979.
6. Gutjahr, A.L., Gelhar, L.W., Bakr, A.A., and MacMillan, J.W., "Stochastic analysis of spatial variability in subsurface flows, 2-Evaluation and application", Water Resources Research, vol.14, no.5, pp.953-959, 1978.
7. Krumbain, W.C., and Graybill, F.A., An Introduction to Statistical

7. Krumbein, W.C., and Graybill, F.A., *An Introduction to Statistical Models in Geology*, McGraw Hill, 1965.
8. Smith, L., and Freeze, R.A., "Stochastic analysis of steady-state groundwater flow in a bounded domain, 2- two dimensional simulations", *Water Resources Research*, vol.15, no.6, pp.1543-1559, 1979.

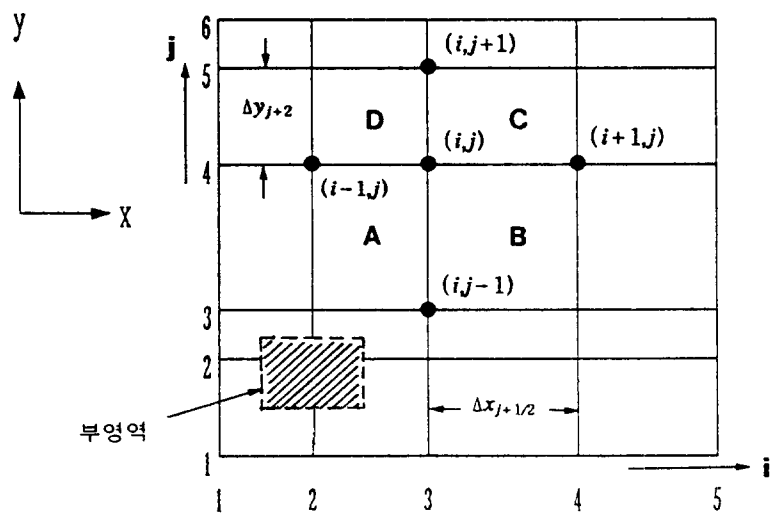


그림 1 직사각형 격자의 예

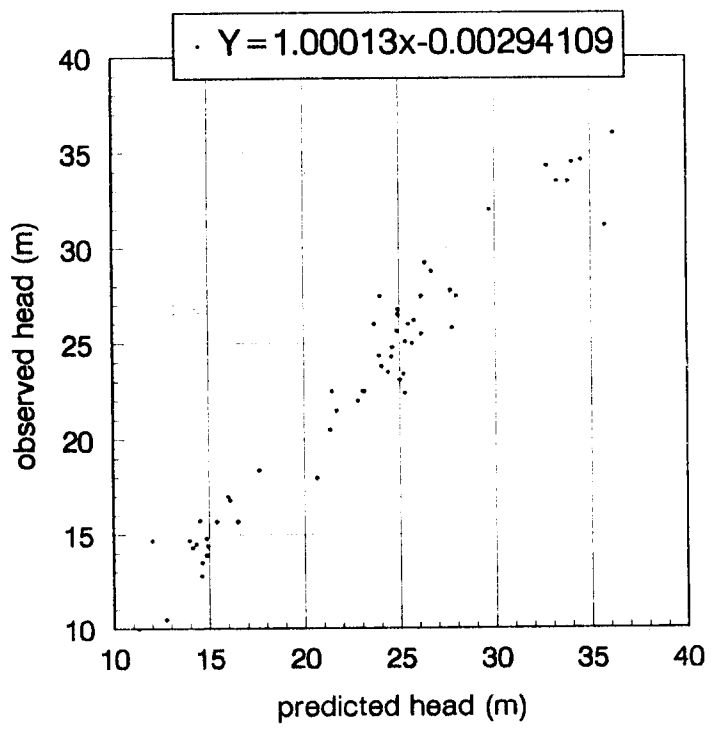


그림 2 관측수두와 계산수두와의 상관도 (상관계수 = 0.98)