2. Parameter estimation in rainfall-watertable relationship using Kalman Filter

5 authors, including:

Isaac Dontwi
Kwame Nkrumah University Of Science and Technology
61 PUBLICATIONS **397** CITATIONS

Thomas Atta-Darkwa
University of Energy and Natural Resources
36 PUBLICATIONS **410** CITATIONS

GLOBAL JOURNAL OF ENGINEERING, DESIGN & TECHNOLOGY
(Published By: Global Institute for Research & Education)
www.gifre.org

Parameter Estimation in Rainfall-Watertable Relationship Using Kalman Filter

Ofori, E., Owusu-Ansah, E., Kyei-Baffour, N., Dontwi, I. K. and Atta-Darkwa, T.
Department of Agricultural Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana.

Abstract

Employing a proper groundwater recharge estimation technique is extremely important for efficient water resource development in a groundwater basin. This paper describes the estimation of groundwater recharge in the Besease basin using the linear Kalman filter mathematical model. The physical model estimated the watertable levels and subsequently derived the infiltration parameters from rainfall inputs and groundwater levels data. The Kalman Filter method used as a recharge estimate resulted in a fit between the simulated hydraulic head and observed sub-surface water level fluctuation. The results show that the infiltration parameter varied considerably over the period of time when it was assumed as time dependent with the recharge values ranging between 0.0-1.27 % for P4 and 0.0-16.5 % for P14 of the incident rainfall. A very high infiltration factor α was obtained when considerable rain fell during June 2009, October 2009 and in June and July, 2010. However, during the periods from December 2009 to April 2010, the infiltration factor was zero which suggested that infiltrated water could not reach the water table but was retained in the unsaturated zone to replenish moisture deficit. Therefore, efficient application of irrigation water, knowledge about the moisture regime and the cropping pattern in the basin is fundamental for ensuring optimal moisture content and watertable level management.

***Keywords**:* Kalman Filter, Model, Infiltration Parameter, Watertable

Introduction

The maximum quantity of water that can be extracted from an aquifer usually depends on the recharge levels of the aquifer. One of the usual forms of recharge is by rainfall. In the case of unconfined aquifers a fraction of the rainfall reaches the watertable and the rest is either lost as evapotranspiration or runoff. This fraction which reaches the aquifer determines the safe yield of the aquifer and hence its estimation to a reasonable degree of accuracy is essential for the proper management of aquifers. The infiltration due to rainfall depends upon several factors like surface-level gradients, sand particle size in the unsaturated zone, depth of watertable level from the surface, intensity of rainfall and so on. These factors not only vary spatially but also with respect to time. The infiltration rates could generally be determined by three methods namely experimental methods, conceptual and time series models. Experimental methods often include the use of lysimeters. Conceptual models include mass-balance. For example Caro and Eagleson (1981) estimated aquifer recharge due to rainfall from an annual water balance. Time series analysis offers a black-box approach for the determination of recharge parameters given a history of rainfall events and watertable readings. Stephenson and Zuzel (1981) noted that for a precipitation of 147 mm the net groundwater recharge was 71 mm which represented 48.3% of rainfall. It was concluded from the study that rainfall in excess of 20–30 mm or higher intensity cloud bursts are major contributors to groundwater recharge.

Rennolls *et al* (1980) used a first order auto-regressive model to describe the response of the watertable level in a borehole to a series of rainfall events. The model parameters λ and α were estimated using maximum likelihood method to be 0.88 and 1.13 respectively. Viswanathan (1983) also developed a model for aquifer in order to estimate the groundwater levels from a history of rainfall observations and past groundwater levels to determine the recharge levels of unconfined aquifers. Matsumoto (1992) utilised a multiple regression analysis to eliminate not only the responses of barometric pressure and earth tide but also precipitation from the groundwater level variation. Modelling groundwater flow faces the problem of modelling an invisible asset. In the field of groundwater studies, groundwater models manage to reproduce the dynamics of the variation of the piezometric heads but they tend to be biased. To circumvent this, it is possible to include the additional information contained in the observations by using data assimilation. Kalman filtering is the most popular approach to data assimilation in hydrological modelling because of its simplicity of implementation and the development of a number of sub-optimal schemes that can be used to deal with high dimensional systems (Riechle *et al*, 2002 and Eigbe *et al*, 1998). This paper describes the model parameter estimates for the Kalman Filter method and determines the watertable levels from series of rainfall inputs and recorded daily water levels.

Study Area

Besease is a predominant farming area in the Ejisu Municipal District of the Ashanti Region in Ghana. The site lies within Latitude 1° 15¹ N and 1° 45¹ N and Longitude 6°15¹ W and 7° 00¹ W. The study area covers about 72 ha of the valley bottom lands at Besease (Figure 1). The climate of the study area is mostly related to the semi-humid type. The region is characterised with two distinct seasons, the wet season which begins from April and ends in October whiles the dry season extends from the month of November-March. The wet seasons can be categorised under two rainy seasons.

The major rainy season which ranges from mid-March to July and the minor rainy season starts from September to mid-November. The mean annual rainfall is 1420 mm; mean monthly temperature is 26.5°C, the relative humidity ranges from 64% in January to 84% in August. The average monthly maximum and minimum evapotranspiration (ET_o) for the study area are 127.5 mm and 64.7 mm respectively and has a an annual ET_o of 1230 mm. The area is drained by the Oda River which is seasonal and whose basin is about 143 km² (Kankam-Yeboah *et al*, 1997). The study area is located in the moist semi-deciduous forest zone.

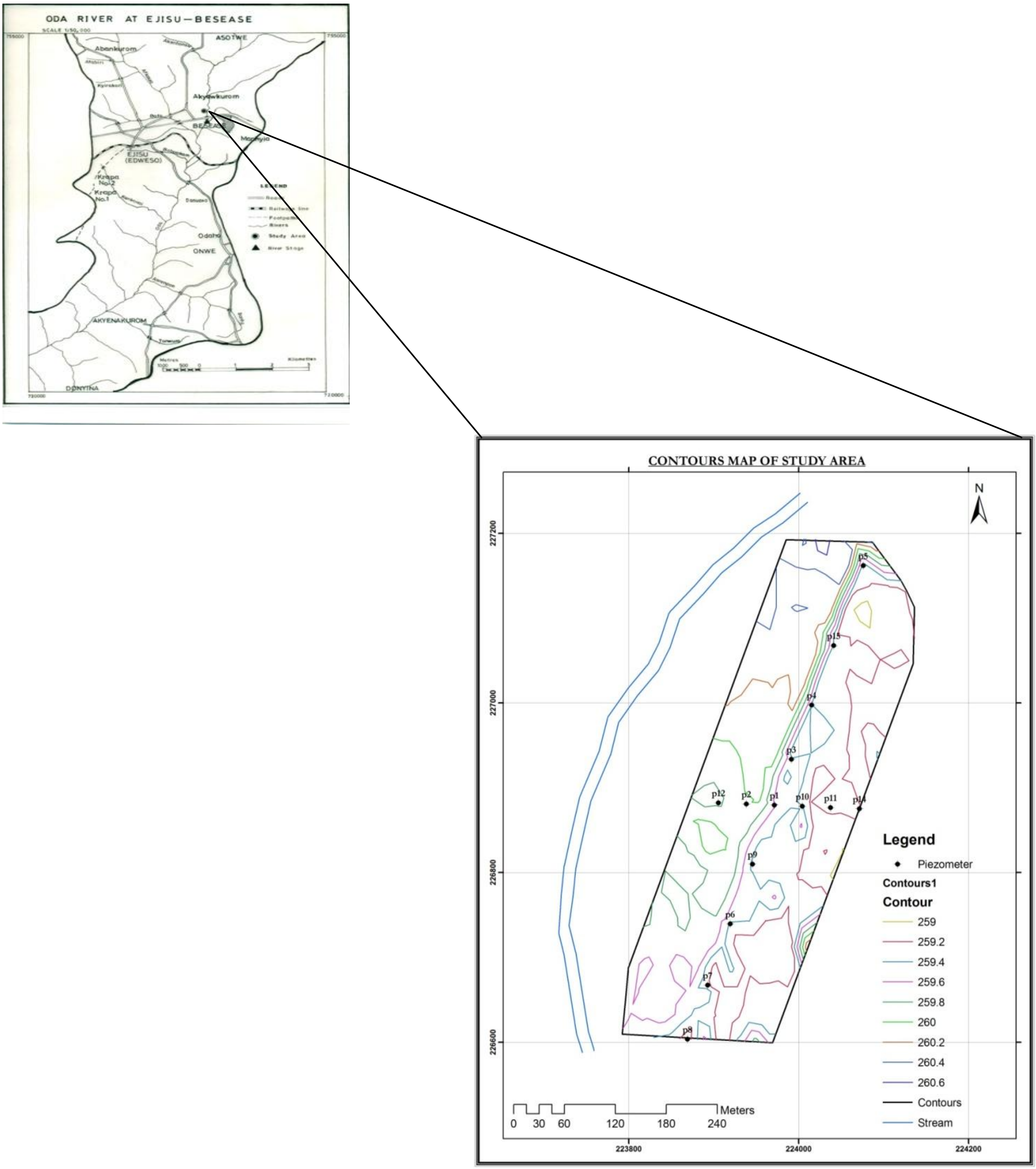


Figure.1 Map of Besease project site showing field piezometric network

Grass species prominently found in the valley bottom are *Santrocema trifolia, Chromolaeve ordorata, Imperata cylindrical, Mimosa pigra,Ceiba patendra, Centrosema pubescens and Mariscus flabelliformis*. Plant species like *Raphia hookeri (Raphia palm), Alstonia boonei, Malotus oppositifolius* and *Pseudospondias microcarpa* extends along the margins of the Oda River. Soils of the Ejisu-Besease can be found in the soil map of Kumasi area. The study area lies in the Offin soil series which are grey to light brownish grey, poorly drained alluvial sands and clays developed within nearly flat but narrow valley bottoms along streams. The series have very slow internal drainage, very slow runoff, rapid permeability and moderate water holding capacity. The geology of the watershed is relatively heterogeneous and mainly composed of Phyllites, quartzite, shale, Tarkwaian and Voltaian-sandstone and limestone. The Phyllites which underlie 59 % of the area consist of upper and lower Birimian rocks. Very few rock outcrops were encountered in the survey as the rocks are deeply weathered. The weathered phyllite is soft and easily broken, recognizable pieces and is typically found at 2-3 m below surface. Soils found within the Oda River catchment are grouped as those derived from granites, sandstones, alluvial materials, greenstone, andesite, schist and amphibolities. Specifically the soils are Orthi-ferric Acrisol, Eutric Fluvisol, Gleyic Arenosols, Eutric Gleysols and Dystri-Haplic Nitisol. The Besease aquifer is composed of heterogeneous sequence of layers which is dominated by sand, clayey sand and silts. The valley bottom is developed by small holder farmers who cultivate rice in the wet season and also grow vegetables like cabbage, lettuce, sweet pepper, cauliflower, cucumber and okra and other cereals like maize in the dry season when the watertable is low.

Materials and Methods

Groundwater Level Monitoring

Wetland groundwater level fluctuations was monitored through a network of 14 piezometers installed using a hand auger along a longitudinal and transverse transect at the Besease site as shown in Figure 1. The piezometers consisted of PVC pipes of 7.62 cm diameter screened over the bottom 20 cm with holes of 0.3 cm diameter. The depth of the pipes ranged from 1.8-3 m. Sand was packed around the screens and the rest of the annulus hole was backfilled with auger cuttings and then grout placed on the top to prevent surface water entry. The cup covering the top of the pipes were not hermetically closed to prevent build up of pressure in the piezometer during phases of groundwater rise. Depth to watertable was measured for every two days with greater frequency during rain events by inserting a measuring tape down into the piezometers and observing when it encountered the water surface. The elevations of the piezometers were surveyed to benchmarks to allow adjusting the water levels in the wells to the local datum.

Modelling of Groundwater flow

The linear relationship (Viswanathan, 1983) between watertable level and rainfall is:

$$y_k = a_{1,k} y_{k-1} + a_{2,k} R_k + a_3 + \epsilon_k \quad (1)$$

Where y_k and R_k are water level in a borehole and rainfall on day " k " and y_{k-1} is the water level in a borehole in day "k-1", a_1, and a_3 represent the natural drainage characteristics of the aquifer and a_2 represents infiltration or recharge characteristics of the aquifer due to rainfall.

To simplify notation the following representations were made:

$$(y_{k-1}\ , R_k\ , 1) = (x_{1,k}\ \ x_{2,k},\ 1) \quad (2)$$

Then

$$y_k = a_{1,k} x_{1,k} + a_{2,k} x_{2,k} + a_3 + \epsilon_k \quad (3)$$

$$= H_k X_k + \epsilon_k \quad (4)$$

Where,

$$X_{k\ k}^T = (a_{1,k}\ \ a_{2,k}\ \ a_3) \quad (5)$$

and

$$H_k = (x_{1,k}\ \ x_{2,k}\ \ 1) \quad (6)$$

For Recursive Formulation

Let

$$\Delta a_{j,k} = v_{j,k}$$

$$a_{j,k} - a_{j,k-1} = v_{j,k} \quad (7)$$

$$a_{j,k} = a_{j,k-1} + v_{j,k} \quad (8)$$

Also let $x_{j,k} = v_{j,k}$

Then

$$x_{1,k} = x_{1,k-1} + v_{1,k}$$

$$x_{2,k} = x_{2,k-1} + v_{3,k}$$

$$x_{3,k} = x_{3,k-1} + v_{3,k} \quad (9)$$

Matrix Representation:

$$\begin{bmatrix} Z_{1,k} \\ Z_{2,k} \\ Z_{3,k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Z_{1,k-1} \\ Z_{2,k-1} \\ Z_{3,k-1} \end{bmatrix} + \begin{bmatrix} v_{1,k} \\ v_{2,k} \\ v_{3,k} \end{bmatrix} \quad (10)$$

The present problem in the aquifer is to estimate the parameters as a function of time, given $R_k\ and\ y_k$, and the estimation of the parameters a_l, a_2 and a_3 is done using the Kalman Filter technique.

The Kalman Filter

The problem of estimating a set of parameters which varies according to a known parameter variation law as:

$$a_k = F a_{k-1} + \varepsilon_k \quad (11)$$

Is similar to the problem of estimating the state vector

$$X_k = (X_1 \dots \dots . . X_n)_k^t$$

Of a linear discrete time stochastic system of the form:

$$X_k = F X_{k-1} + v_k \quad (12)$$

For a p dimensional vector of noisy measurement:

$y_k = (y_1 \dots \dots . . y_p)^t_k$
Linearly related to the state by an observational equation of the form:

$$y_k = H_k X_{k-1} + \epsilon_k \quad (13)$$

Where,
F = I the linear model operator, X_k is the state vector at time k, v_k the model error. The model error is assumed to be time-uncorrelated, normally distributed, with zero mean and covariance matrix Qk (size n · n), also named the model error covariance, y_k is the observation at time k, H_k the observation matrix and ϵ_k the observation error. The observation error is assumed to be time-uncorrelated, normally distributed with zero mean and covariance matrix Rk (size q · q), also named the observation error covariance. Observation error and model error are assumed to be uncorrelated.

$$v_k \sim N(0, Q_k)$$
$$\epsilon_k \sim N(0, R_k)$$

Also F_k and H_k are respectively nxn and pxn matrices.

Kalman Filtering Scheme

Stage 1: Prediction - No knowledge from measurement
The predicted state vector is given by the deterministic model propagation

$$X_k^- = F X_{k-1}^+ \quad (14)$$

The predicted covariance matrix is propagated through the following equation:

$$P_k^- = F P_{k-1}^+ F^T + Q_k \quad (15)$$

Stage 2: Data Assimilation (measurement information used)
The innovation vector is defined as the difference between the observations and the forecast state variables:

$$d_k = y_k - H_k X_k^- \quad (16)$$

and its covariance matrix is:

$$H_k P_{k-1}^+ H_k^T + R_k \quad (17)$$

The Kalman gain is derived by requiring that X_k^+ is the minimum variance estimate of X_k given the observation y_k :

$$K_k = P_k^- H_k^T (H_k P_{k-1}^+ H_k^T + R_k)^{-1} \quad (18)$$

The analysis step is:

$$X_k^+ = X_k^- + K_k \left(y_k - H_k X_k^- \right) \quad (19)$$

and

$$P_k^+ = (I - K_k H_k) P_k^- \quad (20)$$

Where:
X_k^- = *a priori* state estimate
X_k^+ = *a posteriori* state estimate
P_k^- = covariance matrix of the predicted error (e_k - e_k^-)
P_k^+ = covariance matrix of the updated error (e_k - e_k^+)
The analysis step is in fact a linear combination of the observations and the model forecast. The Kalman gain describes how the innovations are spread over the entire state space and weights how strong the correction should be. If the model forecast is more certain than the observation, i.e. $P_k^- \ll R_k$, then the gain is close to zero and $X_k^+ \to X_k$. In case. $P_k^- \gg R_k$ then the gain is close to one, and the analysis is close to the observations.

Solution

Let the parameter vector be:

$$a_k = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

and the matrix H becomes a vector of the following form:
$H_k = (y_{k-1} \; R_k \;\; 1)^t_k$
For a simple random walk model: $F_k = I$
Where I is the identity matrix
Therefore Equations (14) and (15) become:

$$X_k = X_{k-1} \quad (21)$$
$$P_k^- = P_{k-1}^+ + Q_k \quad (22)$$

Hence Equations (21) and (22) are prediction algorithms and Equations (19) and (20) are also correction algorithms for the estimation of time dependent parameters a_1, a_2 and a_3. In Equation (22), the Q matrix is chosen to be diagonal with the diagonal elements selected to represent the expected rate of variation of the parameters between the sample intervals. Differing expected rate of change can be specified for different parameters. Any parameter that is known to be time invariant can simply be handled by setting the appropriate diagonal element to zero (Young, 1974).

If the parameters a_1, a_2 and a_3 are known in Equation 1, then the watertable level can be estimated (Fig 2, Fig 7 and Fig 8) using the equation:

$$y_k = a_{1,k} y_{k-1} + a_{2,k} R_k + a_3 + \varepsilon_k \quad (23)$$

Where y_k is the estimate of level on day k in Equation (1) and there are three components that affect the watertable level on day k. These are water level y_{k-1} in a borehole on day ''k-1'' and rainfall R_t on day "k" and unknown external influences exhibited by the parameter a_3. In estimating the parameters, three scenarios that can exist with respect to the parameters that correspond to the above three components are:

- the parameters a_1, a_2 and a_3 are all time invariant;
- the parameter a_2 was assumed to be time dependent with the value of the diagonal element that corresponds to a_2 was arbitrarily chosen as 0.01 and the parameters a_l and a_3 were treated as time invariants;
- all the parameters are time dependent with equal weightage.

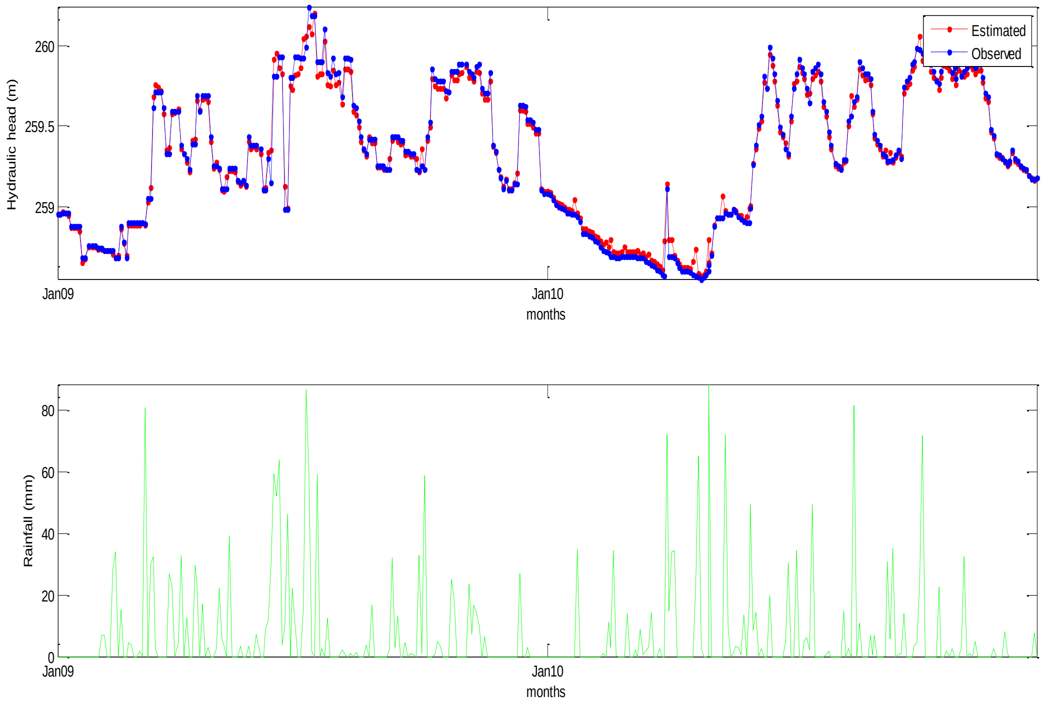
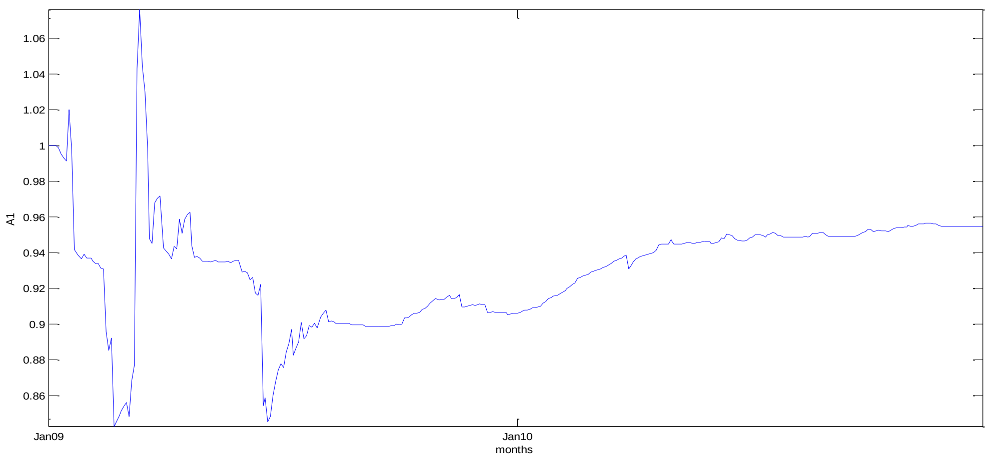


Figure 2: Rainfall and water table level during the years 2009 and 2010 for Q = 0

Results and Discussions

Parameter Estimation



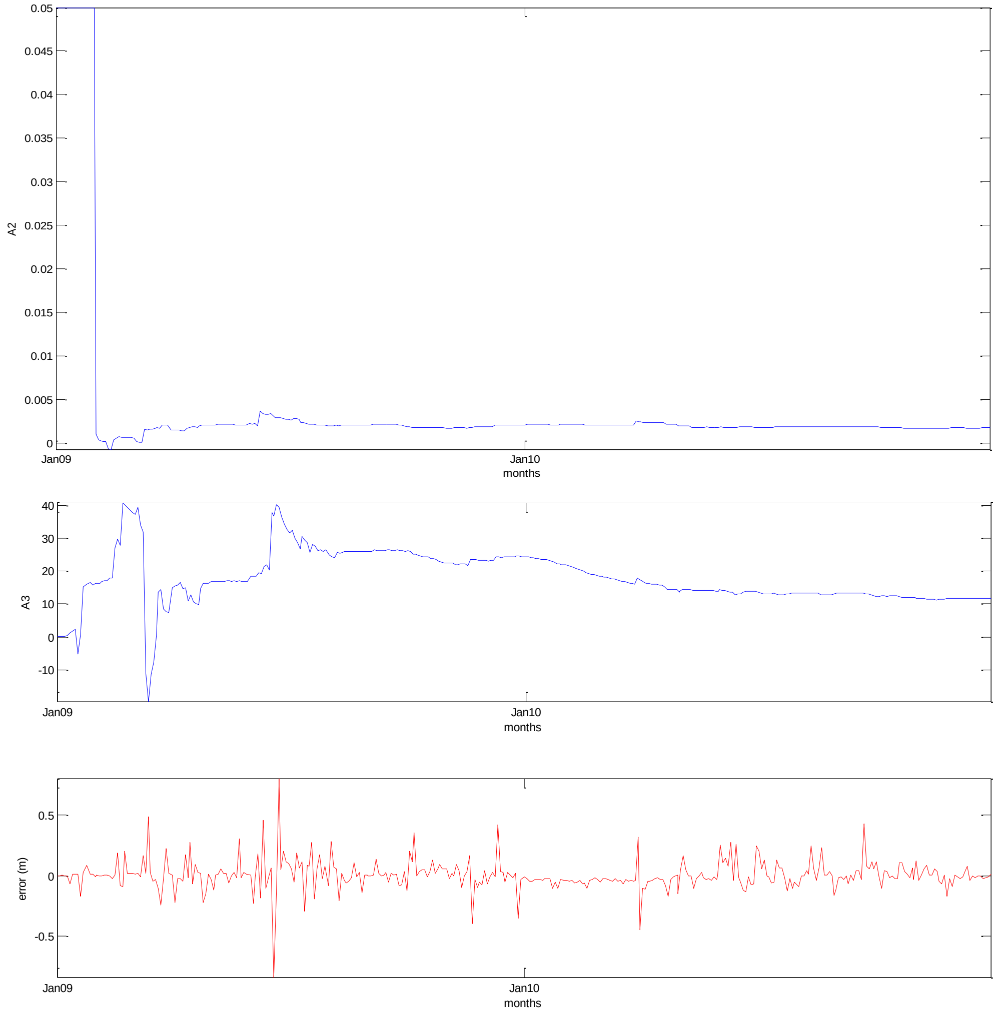


Figure 3: Parameter variation with Q = 0

Figure 3 shows the variation parameters a_1, a_2, a_3 and errors with the diagonals of the matrix **Q** in Equation (21) being taken as zero. This means that the parameters a_1, a_2 and a_3 are assumed to be time invariant. Consequently the variation in the parameters was extremely slow. The infiltration due to rainfall is given by the parameter a_2 which varied between 0.0 and 0.005. The rise in watertable level due to rainfall alone is:

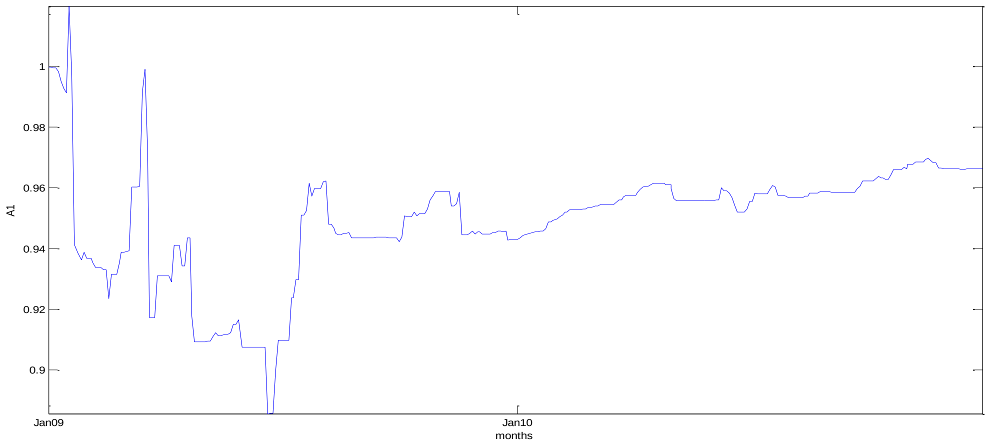
$$\Delta h = a_2 R_t \qquad (22)$$

When α is given as the fraction of rainfall that reaches the watertable, then the rise in watertable could also be expressed as:

$$\Delta h = \frac{\alpha R_t}{Sy} \qquad (23)$$

Where, Sy is the specific yield. Combining Equations (22) and (23):

$$\alpha = a_2 . Sy$$



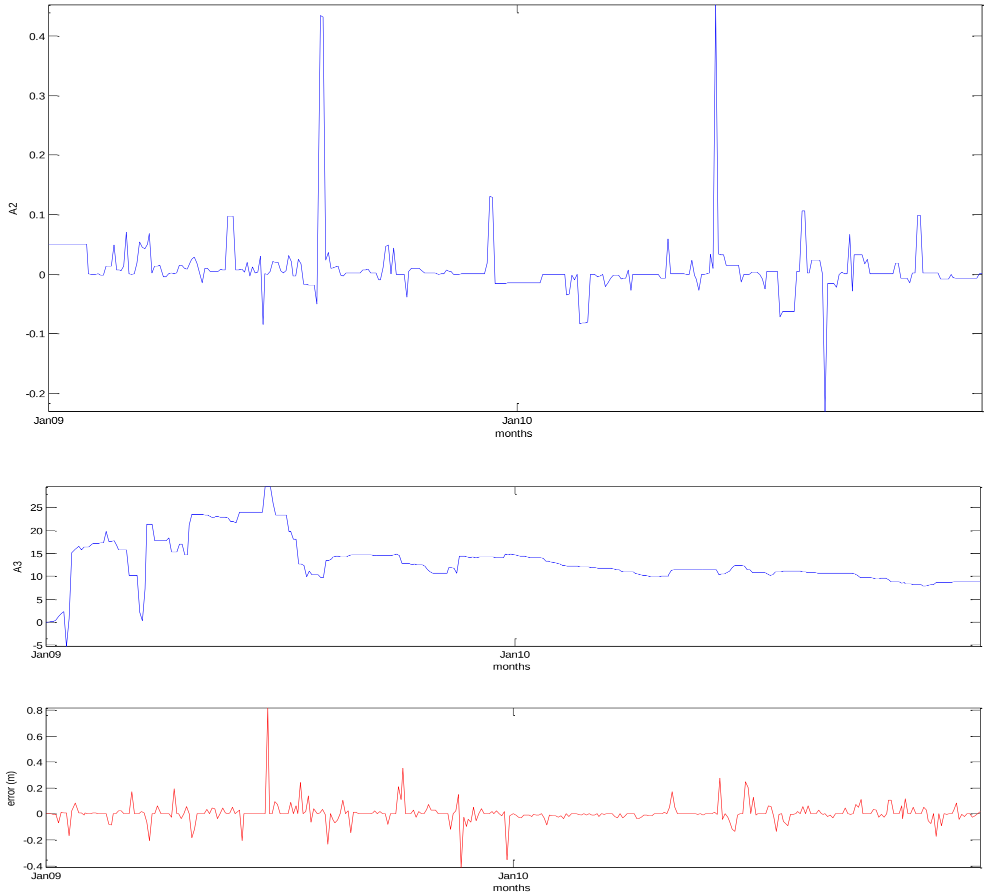
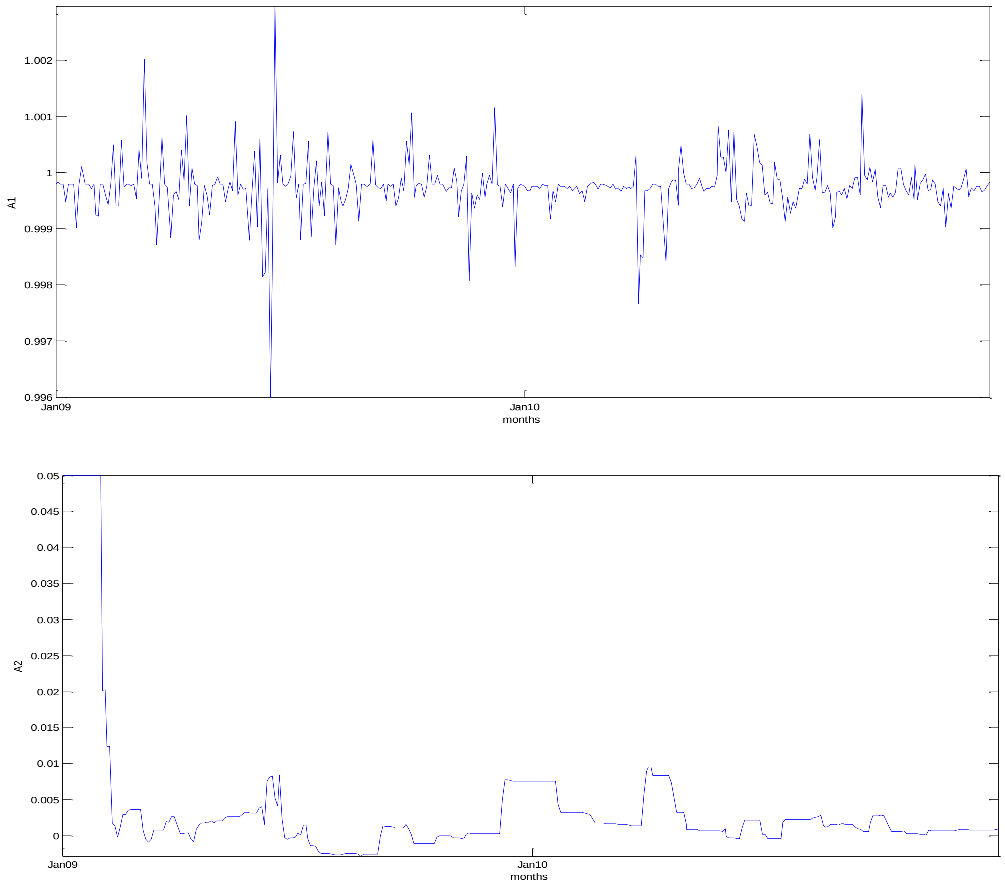


Figure 4: Parameter variation with Q = 0.01 for the diagonal element a_2



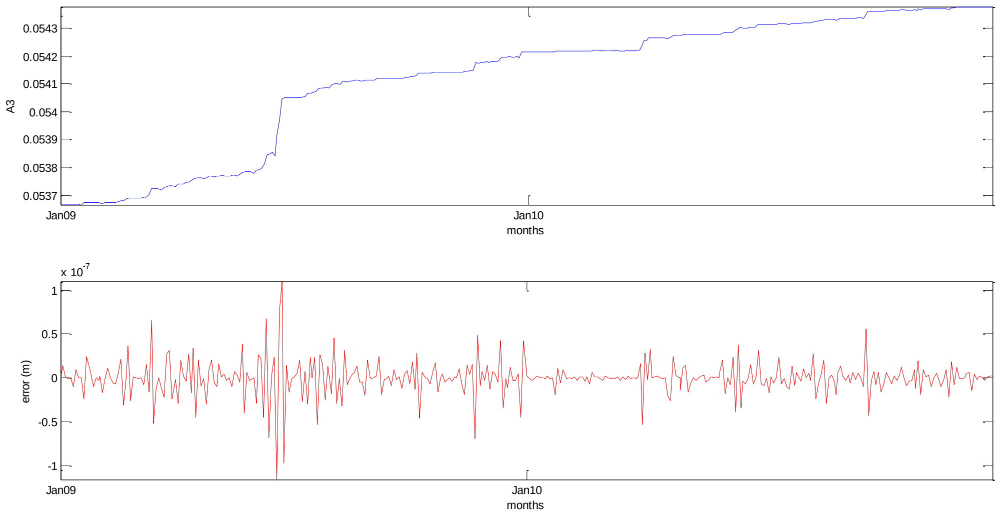
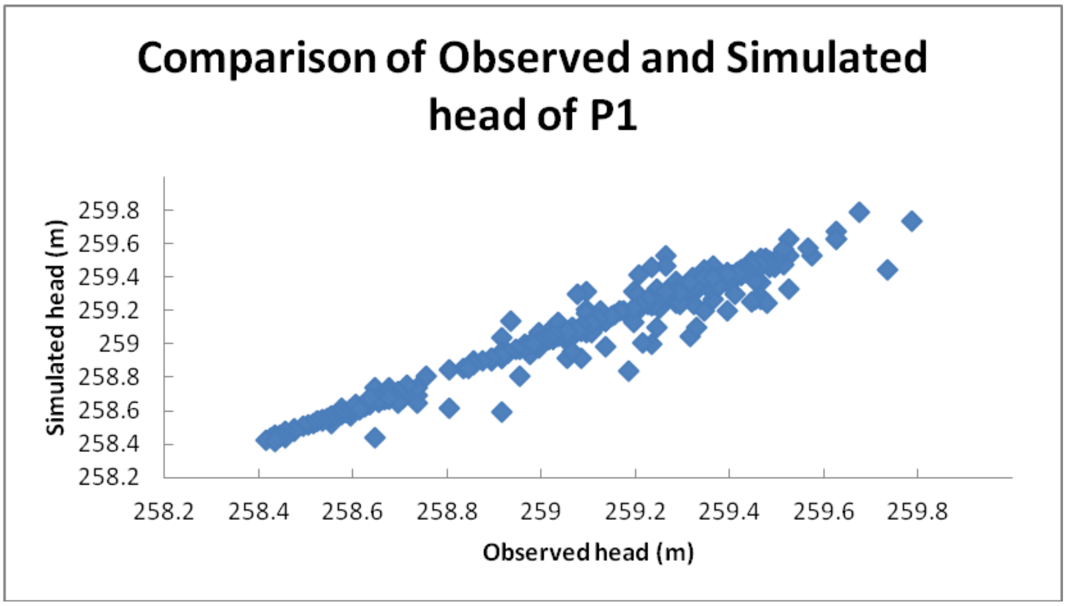


Figure 5: Parameter variation with Q = 0.01 for all the diagonal element a_1,a_2 and a_3

With the average value of specific yield (Sy) for the field being 0.15, the infiltration factor α for the period 2009 and 2010 varied between 0.0 and 0.00075. However, the error between the estimated and actual watertable levels is substantially high owing to the assumption that the parameters are time invariant. In the next scenario, the parameter corresponding to the rainfall a_2 was assumed to be time variant and the value of its diagonal element arbitrarily taken as 0.01. The value of 0.01 was chosen from trial and error, so that the error between the measured and calculated watertable levels was nearly equal to zero. The rest of the parameters a_1 and a_3 were assumed to be time independent.

From Figure 4, the results show that the parameter a_2 varied considerably over the period of time, also with the recharge values ranging between 0.0-1.27 % for P4 and 0-16.5 % for P14 of the incident rainfall. In the last analysis a constant value 0.01 is used for the diagonal elements for the matrix Q, which means that all the parameters are time variant with equal weightage. Results also show α varying between 0.0 and 0.15 % of the annual rainfall for all the piezometers. The Kalman Filter method used as a recharge estimate resulted in a fit between the simulated hydraulic head and observed sub-surface water level fluctuation (Figure 6).



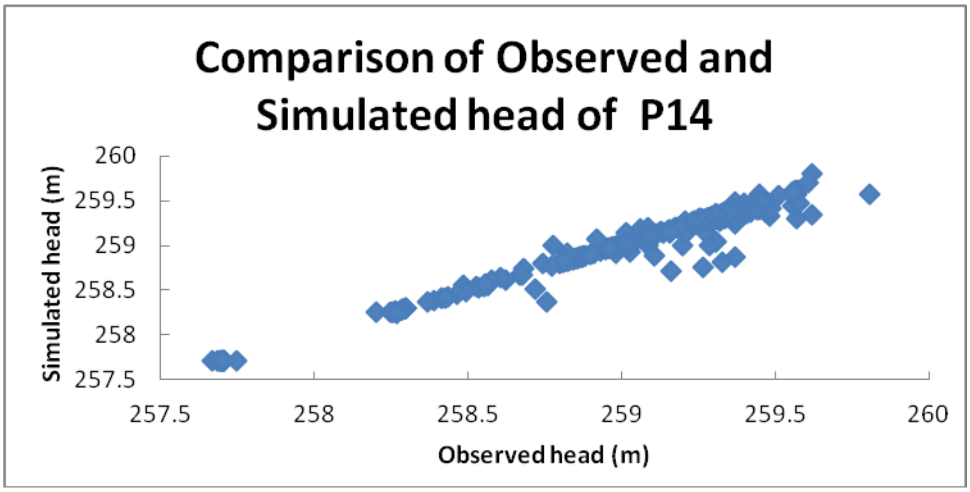


Figure 6: Comparison of simulated and observed heads in the Besease Inland valley wetland

Table 1 Recharge values based on the present day rainfall

Piezometer	Specific yield	Fraction of rainfall(α)	Recharge	% of Rainfall
P1	0.15	0-0.15	0-0.0225	0-2.25
P2	0.15	0-0.46	0-0.069	0-6.95
P3	0.15	0-0.22	0-0.033	0-3.3
P4	0.15	0-0.085	0-0.0127	0-1.27
P5	0.15	0-0.38	0-0.057	0-5.7
P6	0.15	0-0.97	0-0.1455	0-14.5
P7	0.15	0-0.98	0-0.147	0-14.7
P8	0.15	0-0.33	0-0.0495	0-4.95
P9	0.15	0-0.24	0-0.036	0-3.6
P10	0.15	0-0.12	0-0.018	0-1.8
P11	0.15	0-0.16	0-0.024	0-2.4
P12	0.15	0-0.7	0-0.105	0-10.5
P13	0.15	0-0.7	0-0.105	0-10.5
P14	0.15	0-1.1	0-0.165	0-16.5

The error between the calculated and the observed watertable level (Fig 5) is nearly equal to zero. It could be deduced from the assumption that, the time variant parameters of a_2 only was considerable in the parameter values estimated (Fig 4) is more appropriate and realistic than the time invariant assumptions and the time variant parameters with equal weightage. From Figures 4 and 5, when considerable rain fell during June 2009, October 2009, June and July 2010, the infiltration factor α was very high. However, during the periods December 2009 to April 2010, the infiltration factor was zero which indicates that infiltrated water could not reach the watertable but was retained in the unsaturated zone to replenish moisture deficit.

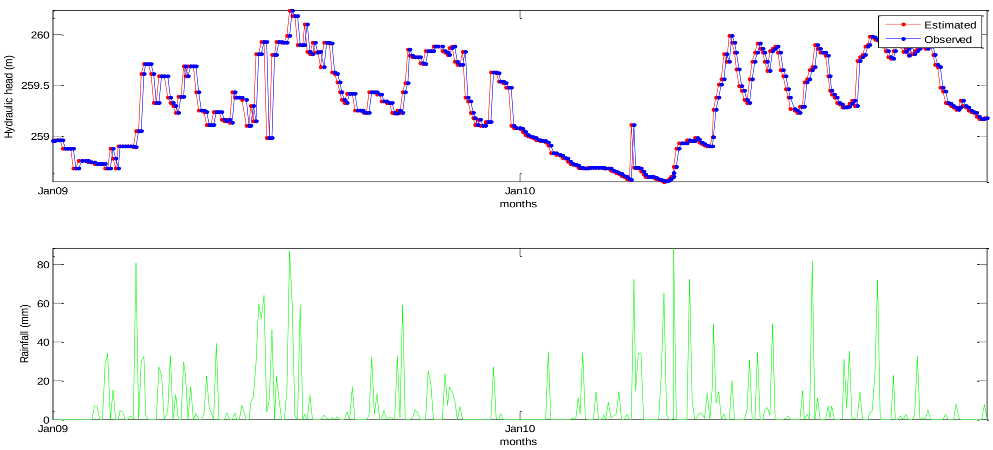


Figure 7: Rainfall and watertable level during the years 2009 and 2010 for a_2, Q= 0.01

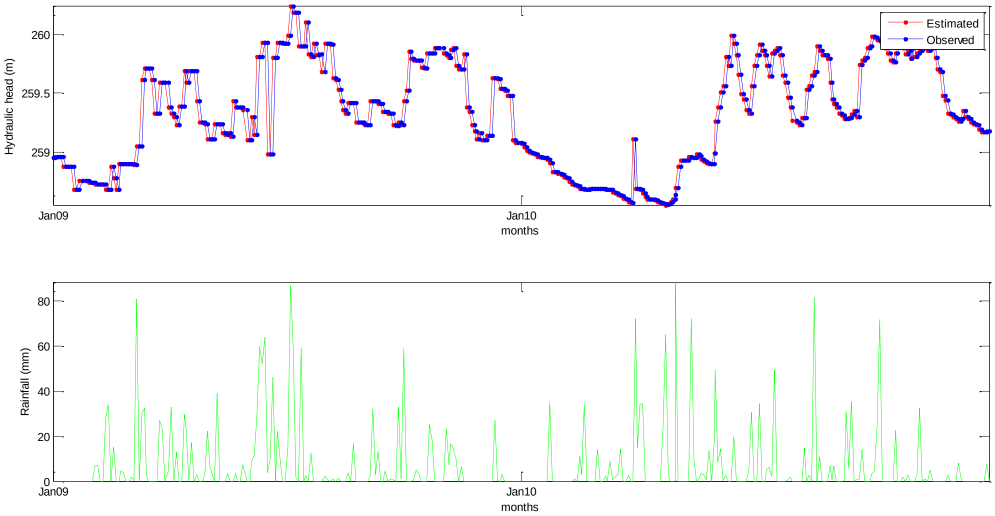


Figure 8: Rainfall and watertable level during the years 2009 and 2010 for a_1,a_2,a_3, Q = 0.01

Conclusions and Recommendations

The Kalman filter method was used to estimate groundwater recharge due to rainfall in an unconfined aquifer with a prior knowledge of the rainfall and history of watertable levels. From the field studies conducted at the inland valley bottom of Besease, it shows that the assumption that, the time variant parameter of the rainfall infiltrating factor is more appropriate and realistic than the time invariant assumptions and the time variant parameters with equal weightage, and that gave a considerable variation in the parameter values. The infiltration factor between the years 2009 and 2010 varied between 0.0 % and 16.5 % of the rainfall. The lowered estimated water levels during the dry periods suggest smaller infiltration factor and therefore must be taken into account for the estimation of safe yield from aquifers in small basins such as Besease.

References

Caro, R. and Eagleson, P.S. 1981. Estimating aquifer recharge due to rainfall. J. Hydrol., 53:185-211.

Eigbe, U., Beck, M. B., Wheater, H.S. and Hirano. F. 1998. Kalman filtering in groundwater flow modelling: problems and prospects. Stoch Hydrol Hydraul.12:15–32.

Kankam Yeboah, K., Duah, A. A., and Mensah, F. K. 1997. Design and construction of Hand Dug wells at Besease Technical report, Water Research Institute (CSIR), Accra, Ghana. 16 pp.

Matsumoto, N. 1992. Regression analysis for anomalous changes of groundwater level due to earth quakes. Geophys. Res. Lett., 19(12) 1193-1196.

Reichle, R. H., McLaughlin, D. B. and Entekhabi, D. 2002. Hydrologic data assimilation with the ensemble Kalman filter. Monthly Weather Rev ;130:103–14.

Rennolls, K., Carnell, R. and Tee, V. 1980. A Descriptive model of the relationship between fainfall and Soil water table. J. Hydrol., 47: 103-114.

Stephenson, C.R. and Zuzel, J. F., 1981. Groundwater recharge characteristics in a semi-arid environment o f southwest Idaho. Journal of Hydrology 53: 213-227.

Viswanathan, M. N., 1983. Recharge characteristics of an unconfined aquifer from the rainfall water-table relationship. J. Hydrol., 70: 233-250.

Young, P.C., 1974. Recursive approaches to time series analysis. The institute of Mathematics and its applications – Bulettin. May-June. pp 209-224.