

PREDICTING GROUNDWATER LEVEL USING FOURIER SERIES INTEGRATED WITH LEAST SQUARE ESTIMATION METHOD

Manoj K. Jha

Department of Civil, Architectural and Environmental Engineering,
North Carolina A and T State University, NC 27411, Greensboro, USA

Received 2014-02-04; Revised 2014-02-17; Accepted 2014-03-26

ABSTRACT

Groundwater level data is an important indicator of the availability and distribution of groundwater resources of the region. However, it is difficult to understand the continuous and discrete fluctuations of the groundwater level which is controlled by various factors. This study demonstrated the use of Fourier series integrated with the least square estimation method to predict the groundwater level especially in the case of seasonal-sensitive groundwater fluctuations. It was observed that the designed method was able to model the groundwater-table data, collected at the Hagan Stone Park station in Greensboro, North Carolina, with a fair degree of accuracy with a testing mean square error of 0.0735.

Keywords: Groundwater Level, Fourier Series, Modeling, Greensboro

1. INTRODUCTION

Groundwater is the principal source of drinking water for about 50% of the population in the United States (Solley *et al.*, 1998). The availability of groundwater varies and fluctuates widely due to short-term and long-term changes in climate, withdrawal and land use. The key data which provides information on groundwater distribution and availability is water-level measurement from observation wells. It also provides critical information regarding hydrologic stresses acting on aquifers and how these affect groundwater dynamics such as recharge, storage and discharge. Systematic and long-term measurements of water-level are essential for the evaluation of groundwater condition for sustainable development of the resources (design, implement and monitor effectiveness of groundwater management and protection) and for the development of groundwater models for scenario and policy analyses as well as forecast trends.

Regular monitoring of groundwater levels in multiple observation wells in a region will provide opportunity for comparative analysis of data for

identifying direction and gradients of groundwater flow. **Figure 1** illustrates a groundwater monitoring well with electric sounding devices which are frequently used to measure depth of groundwater in observation wells.

The variation of groundwater fluctuation depends on various factors including specific geographic, geological and meteorological condition. Groundwater level fluctuation is continuous in nature but changes frequently to discrete on the interruption of human activity. A continuous sequence can be simulated by discrete data; however, the accuracy will depend on the time interval of these data. The shorter time interval, the more accurate actual series can be simulated. Precipitation and evaporation have a periodic impact on groundwater level fluctuation whereas manual extractions (for irrigation and others) are rather uncertain. Some natural factors such as crustal movement and tide and human factors provide some random changes.

There have been several modeling development and application to simulate the changes in groundwater depth under both continuous and discrete conditions.



Fig. 1. Groundwater level monitoring Station (Site: Hagan Stone Park, Greensboro, North Carolina)

Groundwater flow modeling is an integral component of the watershed modeling where groundwater is simulated separately for quantifying water balance and pollutant transport (Jha *et al.*, 2010; Jha 2011; Jha *et al.*, 2013, Jha and Gassman, 2013b; Amatya and Jha, 2011; Amatya *et al.*, 2011; 2013; Secchi *et al.*, 2011; Rabotyagov *et al.*, 2010). However, watershed models simulate groundwater component as a process-based mechanism and used in mass balance approach on a temporal scale. On the other hand, groundwater is also modeled empirically using data-fitting techniques. Daliakopoulos *et al.* (2005) compared different types of artificial neural networks including feed-forward neural network, Elman or recurrent neural network and Radial basis function neural network to predict the ground water level. The first method was found to achieve the best fit for groundwater level fluctuations which are influenced mainly by precipitation and surface discharge. They also demonstrated that the neural network techniques are useful prediction tools for groundwater modeling. Similarly, Yang *et al.* (2009) compared integrated time-series method with back-propagation neural network. Both methods were found to predict well with the exception of underestimating and/or overestimating extreme values. The possible reason provided was the poor model calibration where model was not well-trained due to the lack of samples of extreme situations.

This study attempted to apply an integrated time-series method with an emphasis to the “seasonal fluctuations” component of the time series. The Fourier series (similar to the integrated time-series method) was integrated with the least square estimation method to model the groundwater level fluctuations. A well-observation study site located in Greensboro, North

Carolina was chosen to test the technique. Successful application of the technique will provide a robust tool for groundwater modeling where groundwater level fluctuation is heavily influenced on a seasonal basis.

2. MODEL DEVELOPMENT

Traditionally, seasonal components are modeled as a truncated Fourier series. The change in groundwater level can be expressed as:

$$\text{Water Table} = \text{Base} + \text{Growth Rate} * w + \text{Seasonal}_w$$

And, the seasonal components can be written as:

$$\text{seasonal}_w = \sum_{i=1}^N A(i) \cos(w_0 * i * w) + B(i) \sin(w_0 * i * w)$$

where, $w_0 = \pi/365$. It is assumed that there are 365 days in a year and base and growth rates are linear functions of the week variable (w).

The direct application of solving the normal equations for least square problems become cumbersome as the number of unknown constants in a problem grows. Generally we can express this kind of problem as:

$$y = w_0 x_0 + w_1 x_1 + \dots + w_n x_n$$

For discrete events, T , above equation can be rewritten as:

$$Y = X * W$$

And so the Fourier series matrix X is:

$$X = \begin{bmatrix} 1 & 1 & \sin \omega & \cos \omega & \dots & \sin 3\omega & \cos 3\omega \\ 2 & 4 & \sin 2\omega & \cos 2\omega & \dots & \sin 6\omega & \cos 6\omega \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ n & n^2 & \sin n\omega & \cos n\omega & \dots & \sin 3n\omega & \cos 3n\omega \end{bmatrix}$$

Above expression of the truncated Fourier series integrated with the least squares modeling technique can be approximated by the Optimized Linear Associative Memory (OLAM) technique (Haykin, 1999). The weights of the OLAM guarantee perfect retrieval of stored memories given that the columns of both the x and y fields are linearly independent (a basic assumption in this model). Both the X and Y fields are

separated into training set and test set. The forward OLAM weights are calculated as:

$$W = (X^T * X)^{-1} * X * Y$$

3. MODEL TESTING

Guilford County of the state of North Carolina in USA is experiencing heavy industrial development and population growth, which resulted in an ever-increasing demand for water. In the rural areas of the county, groundwater remains the water supply resource for domestic use (NC-DPH, 2007). Because the amount of groundwater in the bedrock aquifers available in the county for potable water is largely unknown, the availability of groundwater as a present and future resource has been a concern for the water supplies in the suburban communities. Therefore, it is necessary to establish a long-term monitoring network to monitor and determine whether the increasing new developments in the county will impact the availability of groundwater resources. The data obtained from the long-term groundwater monitoring would benefit county government for the policy making for new

development and permitting process in the county and area developers for considering groundwater resources for the new development.

The city of Greensboro in Guilford County installed a total of eight stations across the county for this monitoring network. Each site has groundwater level data collected on a daily basis since May of 2003. Here, we analyzed the time-series data for historical trend and selected the monitoring station at Hagan Stone Park due to its noticeable fluctuations in the seasonal characteristics. **Figure 2** shows the historical data on a daily basis from May 2003 to Feb 2011.

It can be observed that the water table has a significant relationship with the seasonal affections with a decreasing trend over the years. There is a consistent annual peak every May and an annual valley every November. The curve also shows an approximate sine wave shape. In southeastern part of the United States, the summer precipitation peaks are due to the rainfall season and the winter valleys are due to the dry season. The strong seasonal pattern was found suitable for the application of the truncated Fourier series in conjunction with the least square estimation method.

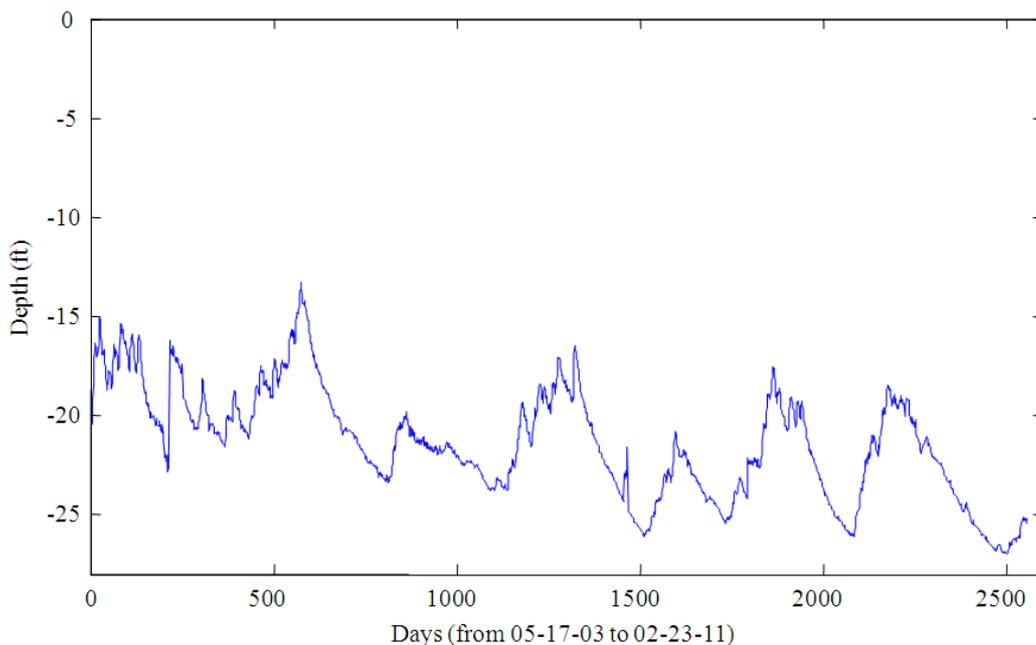


Fig. 2. Historical groundwater level monitored at Hagan Stone Park, Greensboro, North Carolina

4. RESULTS AND DISCUSSION

The raw dataset was divided into two parts. The first part was the training part containing 1,300 samples from May 2003. Using the aforementioned model, we generated an input vector (X) consisting of 1300 rows and 15 columns, with column 1 representing the bias, column 2 the growth variable, column 3 the super growth variable and columns 4-15 6 successive harmonics of the Fourier series (Table 1). The output vector Y will consist of the 1300 daily water table data. The training result is shown in Fig. 3. Statistical analysis yielded a mean square error of 0.4382. Figure 4 shows the error curve which seem to be equally distributed in both sides.

The calculated weight matrix was used to test the model by simply applying the equation $Y = X*W$. The test part is selected as the next 100 data following the data used in the training. The testing input matrix consists of 100 rows and 15 columns. The test results are plotted and shown in Fig. 5. Here, the mean square error was found to be 0.0735. Figure 6 shows the error curve.

It was observed that the mean square error of the testing section was smaller than that of the training section. The possible reason may be the larger number of data samples for the training section. From the satisfactory results of the model application during the training and testing sections, it can be concluded that the Fourier series and least squares estimation method performed very well in this specific application.

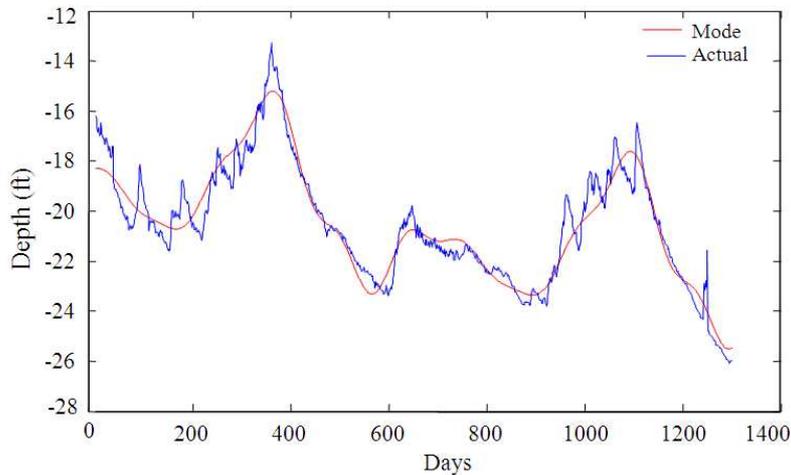


Fig. 3. Actual Vs. modeled data in the training section

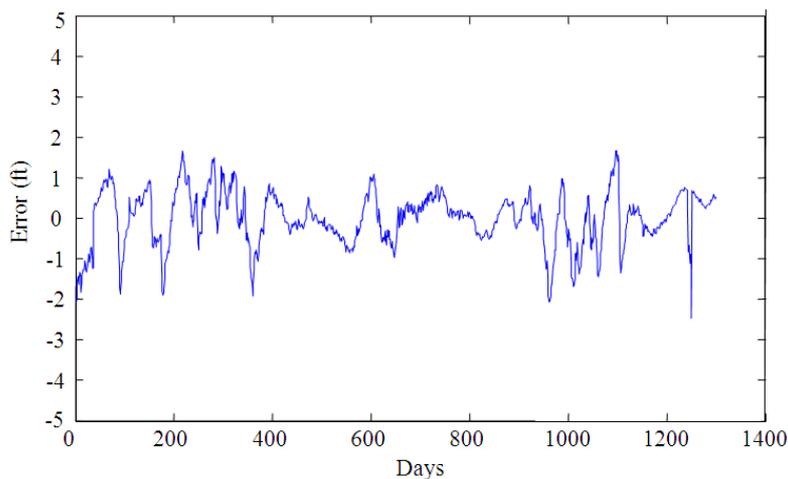


Fig. 4. Error curve of training performance

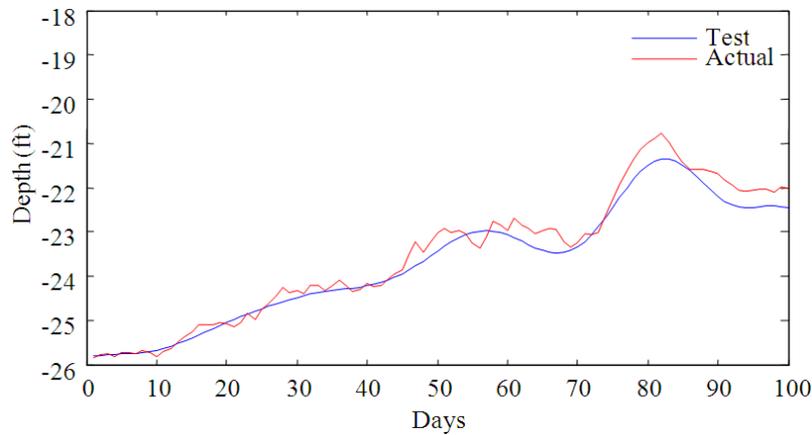


Fig. 5. Performance of the model during the testing period

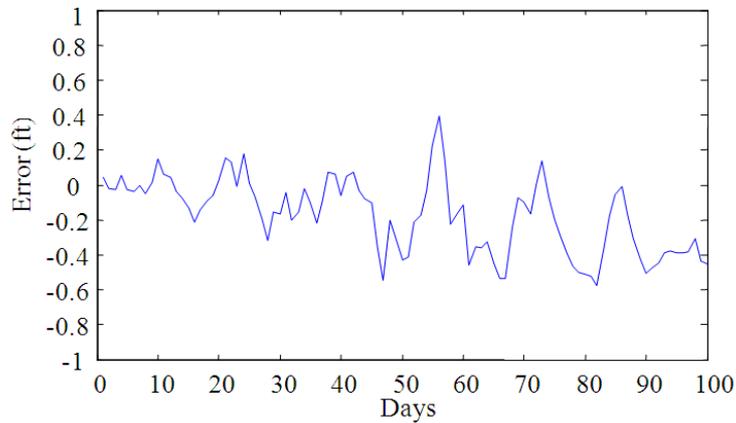


Fig. 6. Error curve during the testing performance

Table 1. Weight matrices

w1	-18.175	w9	-0.24129
w2	-0.004460	w10	-0.035270
w3	7.90E-07	w11	0.028418
w4	0.327930	w12	0.072356
w5	-1.696200	w13	-0.372070
w6	-0.527600	w14	0.241810
w7	1.883600	w15	0.275990
w8	-0.103960		

5. CONCLUSION

The Fourier series was integrated with the least square estimation method to model the groundwater level fluctuations. It was observed that the OLAM was able to model the ground water table data, collected at the Hagan Stone Park station in Greensboro, North Carolina, with a fair degree of accuracy with a testing mean square error of 0.0735.

It can be concluded that the Fourier series integrated with the least squares estimation method is an excellent modeling tool. Its ability to model multivariate problems without making complex dependency assumptions among input variables is quite useful. It attempts to draw links between sets of input data and observed outputs and so it is considered a very robust tool for time-series analysis. However, there are several issues that cannot be addressed by it alone. Some of these include: (1) Effects of special weather and abnormal groundwater flow movement cannot be properly reflected; (2) Number of harmonics cannot be well decided without help of statistical method; and (3) Performance can be significantly downgraded if the samples are not well collected.

Further consideration and enhancement of the model may be obtained by using the statistics to analyze the

residual between the model and actual data. This will provide opportunity to identify potential periodic occurrence effects on the water table.

6. REFERENCES

- Amatya, D.M. and M.K. Jha, 2011. Evaluating the SWAT model for a low-gradient forested watershed in coastal South Carolina. *Trans. ASABE*, 54: 2151-2163. DOI: 10.13031/2013.40671
- Amatya, D.M., M.K. Jha, A. Edwards, T. Williams and D. Hitchcock, 2011. SWAT-based streamflow and embayment modeling of karst affected Chapel Branch Watershed, SC. *Trans. ASABE*, 54: 1311-1323. DOI: 10.13031/2013.39033
- Amatya, D.M., M.K. Jha, A. Edwards, T. Williams and D. Hitchcock, 2013. SWAT model prediction of phosphorus loading in a South Carolina Karst watershed with a downstream embayment. *J. Environ. Protect.*, 4: 75-90. DOI: 10.4236/jep.2013.47A010
- Daliakopoulos, I.N., P. Coulibaly and I.K. Tsanis, 2005. Groundwater level forecasting using artificial neural networks. *J. Hydrol.*, 309: 229-240. DOI: 10.1016/j.jhydrol.2004.12.001
- Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation*. 2nd Edn., Prentice Hall, Upper Saddle River, ISBN-10: 0780334949, pp: 700.
- Jha, M.K., K.E. Schilling, P.W. Gassman and C.F. Wolter, 2010. Targeting land-use change for nitrate-nitrogen load reductions in an agricultural watershed. *J. Soil Water Conser.*, 65: 342-352. DOI: 10.2489/jswc.65.6.342
- Jha, M.K., 2011. Evaluating hydrologic response of an agricultural watershed for watershed analysis. *Water*, 3: 604-617. DOI: 10.3390/w3020604
- Jha, M.K., P.W. Gassman and Y. Panagopoulos, 2013. Regional changes in nitrogen loadings in the Upper Mississippi River Basin under predicted mid-century climate. *Regional Environ. Change*, DOI: 10.1007/s.10113-013-05539-y
- Jha, M.K. and P.W. Gassman, 2013. Changes in hydrology and streamflow as predicted by modeling experiment forced with climate models. *Hydrol. Process.*, 28: 2772-27181. DOI: 10.1002/hyp.9836
- NC-DPH, 2007. *Guilford County Groundwater Network Status Report*. HERA Team. Division of Environmental Health. North Carolina Department of Public Health.
- Rabotyagov, S., M.K. Jha and T. Campbell, 2010. Impact of crop rotations on optimal selection of conservation practices for water quality protection. *J. Soil Water Conser.*, 65: 369-380. DOI: 10.2489/jswc.65.6.369
- Secchi, S., P.W. Gassman, M.K. Jha, L. Kurkalova and C.L. Kling, 2011. Potential water quality changes due to corn expansion in the Upper Mississippi River Basin. *Ecol. Applic.*, 21: 1068-1084. DOI: 10.1890/09-0619.1
- Solley, W.B., R.R. Pierce and H.A. Perlman, 1998. *Estimated use of water in the United States in 1995*. U.S. Geological Survey Circular.
- Yang, Z.P., W.X. Lu, Y.Q. Long and P. Li, 2009. Application and comparison of two prediction models for groundwater levels: A case study in Western Jilin Province, China. *J. Arid Environ.*, 73: 487-492. DOI: 10.1016/j.jaridenv.2008.11.008