Computers & Geosciences 36 (2010) 936–944

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Development of genetic algorithm-based optimization module in WHAT system for hydrograph analysis and model application

Kyoung Jae Lim[a,1], Youn Shik Park[a,1], Jonggun Kim[a,1], Yong-Chul Shin[a,1], Nam Won Kim[b,2], Seong Joon Kim[c,3], Ji-Hong Jeon[d,4], Bernard A. Engel[e,*]

[a] Regional Infrastructures Engineering, Kangwon National University, Chuncheon, Kangwon, Republic of Korea
[b] Water Resources Research Department, Korea Institute of Construction Technology, Kyounggi, Republic of Korea
[c] Department of Civil & Environmental System Engineering, KonKuk University, Seoul, Republic of Korea
[d] Department of Environmental Engineering, Andong National University, Andong, KyougsangBukdo, Republic of Korea
[e] Department of Agricultural and Biological Engineering, Purdue University, 225 South University Street, West Lafayette, IN 47907-2093, USA

A R T I C L E I N F O

Article history:
Received 6 July 2009
Received in revised form
26 January 2010
Accepted 27 January 2010

Keywords:
Baseflow separation
Recession curve analysis
Water quality
L-THIA
WHAT

A B S T R A C T

Many hydrologic and water quality computer models have been developed and applied to assess hydrologic and water quality impacts of land use changes. These models are typically calibrated and validated prior to their application. The Long-Term Hydrologic Impact Assessment (L-THIA) model was applied to the Little Eagle Creek (LEC) watershed and compared with the filtered direct runoff using BFLOW and the Eckhardt digital filter (with a default BFI_{max} value of 0.80 and filter parameter value of 0.98), both available in the Web GIS-based Hydrograph Analysis Tool, called WHAT. The R^2 value and the Nash–Sutcliffe coefficient values were 0.68 and 0.64 with BFLOW, and 0.66 and 0.63 with the Eckhardt digital filter. Although these results indicate that the L-THIA model estimates direct runoff reasonably well, the filtered direct runoff values using BFLOW and Eckhardt digital filter with the default BFI_{max} and filter parameter values do not reflect hydrological and hydrogeological situations in the LEC watershed. Thus, a BFI_{max} GA-Analyzer module (BFI_{max} Genetic Algorithm-Analyzer module) was developed and integrated into the WHAT system for determination of the optimum BFI_{max} parameter and filter parameter of the Eckhardt digital filter. With the automated recession curve analysis method and BFI_{max} GA-Analyzer module of the WHAT system, the optimum BFI_{max} value of 0.491 and filter parameter value of 0.987 were determined for the LEC watershed. The comparison of L-THIA estimates with filtered direct runoff using an optimized BFI_{max} and filter parameter resulted in an R^2 value of 0.66 and the Nash–Sutcliffe coefficient value of 0.63. However, L-THIA estimates calibrated with the optimized BFI_{max} and filter parameter increased by 33% and estimated NPS pollutant loadings increased by more than 20%. This indicates L-THIA model direct runoff estimates can be incorrect by 33% and NPS pollutant loading estimation by more than 20%, if the accuracy of the baseflow separation method is not validated for the study watershed prior to model comparison. This study shows the importance of baseflow separation in hydrologic and water quality modeling using the L-THIA model.

Published by Elsevier Ltd.

1. Introduction

Significant areas have been converted into impervious areas with human induced development activities in recent years. This "urban sprawl" has been a dominant phenomenon in urbanized regions worldwide according to the U.S. EPA (2001) due to social and economic benefits from development. However, the negative impacts of "urban sprawl" on hydrology and water quality have been recognized only recently. Thus, many computer simulation models have been developed and utilized to assess the impacts of urban sprawl to assist in environment-friendly land use planning.

The Long-Term Hydrologic Assessment Tool (L-THIA) (Harbor, 1994; Bhaduri et al., 2001; Lim et al., 2001, 2006), Soil and Water Assessment Tool (SWAT) (Arnold et al., 1995), and Hydrological Simulation Program—Fortran (HSPF) (Bicknell et al., 1997) models have been frequently used for this purpose. The accuracies of these models should be validated prior to their application in land use planning. The water quality components of these models rely on their

* Corresponding author. Tel.: +1 765 494 1162; fax: +1 765 496 1115.
E-mail addresses: kjlim@kangwon.ac.kr (K.J. Lim), incrediblemorning@gmail.com (Y.S. Park), kimjg23@gmail.com (J. Kim), asterion17@lycos.co.kr (Y.-C. Shin), nwkim@kict.re.kr (N.W. Kim), kimsj@konkuk.ac.kr (S.J. Kim), jhjeon@andong.ac.kr (J.-H. Jeon), engelb@purdue.edu (B.A. Engel).
[1] Tel.: +82 33 250 6468; fax: +82 33 251 1518.
[2] Tel./fax: +82 31 910 0256.
[3] Tel./fax: +82 2 450 3749.
[4] Tel./fax: +82 54 820 5093.

0098-3004/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.cageo.2010.01.004

corresponding hydrologic components. Thus, researchers and modelers typically first calibrate and validate the hydrologic component of models. In most cases, these models simulate direct runoff and baseflow components separately; thus it would be desirable to calibrate and validate the direct runoff and baseflow modules separately.

For accurate model calibration and validation, the direct runoff and baseflow components from stream flow have to be first separated. There are numerous methods, called "hydrograph analysis" or "baseflow separation", available to separate baseflow from measured stream flow hydrographs. The traditional hydrograph analysis methods are not very efficient because these subjective techniques do not provide consistent results. Thus, the U.S. Geological Survey (USGS) has developed and distributed an automated hydrograph analysis program called HYSEP (Sloto and Crouse, 1996). Digital filtering methods have also been used in baseflow separation because they are easy to use and provide consistent results (Lyne and Hollick, 1979; Chapman, 1987; Nathan and McMahon, 1990; Arnold et al., 1995; Arnold and Allen, 1999; Eckhardt, 2005). The BFLOW (Lyne and Hollick, 1979; Arnold and Allen, 1999) and the Eckhardt (Eckhardt, 2005) digital filters are widely used for hydrograph analysis. Although a new parameter was introduced in the Eckhardt digital filter to reflect local hydrogeological situations (BFI_{max}) and representative values were proposed for various aquifers, the use of a BFI_{max} value specific to local conditions is strongly recommended instead of using the proposed representative BFI_{max} values.

Lim et al. (2005) developed the Web GIS-based Hydrograph Analysis Tool (WHAT) (https://engineering.purdue.edu/~what) to provide fully automated functions for baseflow separation. Lim et al. (2005) compared the filtered baseflow for 50 gaging stations in Indiana, USA using the Eckhardt filter with a default BFI_{max} value of 0.80 and the BFLOW filter. The Nash–Sutcliffe coefficient values were 0.91 for 50 Indiana gaging stations. However, the use of the default BFI_{max} value of 0.80 for 50 Indiana gaging stations is not recommended because the filtered baseflow data were compared with the results from the BFLOW digital filter which did not reflect the physical characteristics in the aquifer and watershed.

Lim et al. (2006) validated the L-THIA model accuracy by comparing L-THIA daily direct runoff with daily direct runoff values estimated using BFLOW and observed flow. The Nash–Sutcliffe coefficient values were 0.60 for both calibration and validation periods. Although the studies by Lim et al. (2005, 2006) provided higher statistics for comparisons, none of studies by Lim et al. (2005, 2006) was compared with the direct runoff values separated using digital filters reflecting the local hydrogeological conditions. There is a need for an automated module to separate baseflow accurately from stream flow in calibration and validation of hydrologic and water quality models, such as L-THIA and SWAT.

The objectives of this study are to: (1) develop an automated module in the WHAT system for determination of optimum Eckhardt BFI_{max} value and filter parameter values using Genetic Algorithm (GA) techniques by comparing the filtered baseflow with baseflow from recession curve analysis; (2) assess the hydrologic and water quality impacts of using optimized BFI_{max} values, rather than a default BFI_{max} value of 0.80 provided by Eckhardt (2005).

2. Literature review

2.1. Long-Term Hydrologic Impact Assessment (L-THIA) ArcView GIS

The L-THIA model was developed to estimate direct runoff using the CN method (Harbor, 1994). It utilizes daily rainfall depth, land use, and hydrologic soil group data. An ArcView GIS interface was developed and enhanced over the last several years to provide a user-friendly interface to the L-THIA model (Bhaduri et al., 2001; Lim et al., 2001), and a Web-based L-THIA was developed and is accessible from http://www.ecn.purdue.edu/runoff/. The L-THIA model estimates NPS pollutant loadings by multiplying the estimated daily direct runoff by pollutant loading coefficients, called Event Mean Concentration (EMC) values, associated with land use (Lim et al., 2001) (Fig. 1).

The L-THIA model has been used in many research efforts for land use impact assessment (Pandey et al., 2000; Bhaduri et al., 2001; Kim et al., 2002). A daily version of the L-THIA model is also available, and Lim et al. (2005) applied the daily version of L-THIA for the comparison of yearly direct runoff values. The comparison gave reasonable results with a Nash–Sutcliffe coefficient of 0.67. Lim et al. (2006) developed an automatic calibration tool to calibrate the daily version of the L-THIA model, and its comparison was reasonable with a Nash–Sutcliffe coefficient value of 0.60.

2.2. Hydrograph analysis

The hydrograph describes flow versus time, and the shape of a hydrograph varies depending on physical and meteorological conditions in the watershed (Bendient and Huber, 2002). As described in the "Introduction" section, there are many methods available to separate baseflow from stream flow hydrographs. To overcome limitations in traditional baseflow separation methods, many automated programs, such as the USGS HYSEP (Sloto and Crouse, 1996) and the BFLOW (Arnold and Allen, 1999), were developed to provide consistent baseflow separation results. The USGS HYSEP (Sloto and Crouse, 1996) provides three baseflow separation methods: fixed interval, sliding interval, and local minimum method. However, HYSEP is not a user-friendly program and requires a great deal of user intervention to prepare input data and run the program. Arnold and Allen (1999) compared filtered baseflow data using the BFLOW digital filtering technique (Lyne and Hollick, 1979; Arnold and Allen, 1999) with measured data; the R^2 value for this comparison was 0.83. However, the BFLOW digital filtering method does not consider aquifer characteristics in separating baseflow from stream flow because the digital filtering method by Lyne and Hollick (1979) was originally used in signal analysis and processing to separate high frequency signals from low frequency signals. Thus, Eckhardt (2005) proposed a general form of the digital filtering method with representative BFI_{max} (maximum value of long-term ratio of baseflow to total stream flow) parameter values for various hydrogeological situations to minimize the subjective influence of using BFI_{max} on baseflow separation (Eckhardt, 2005). Eckhardt (2005) estimated representative BFI_{max} values for different hydrological and hydrogeological situations. Eckhardt (2005) proposed the use of BFI_{max} values of 0.80 for perennial streams with porous aquifers, 0.50 for ephemeral streams with porous aquifers, and 0.25 for perennial streams with hard rock aquifers. However, use of a BFI_{max} value specific to local conditions is recommended.

2.3. Web-based Hydrograph Analysis Tool (WHAT)

The Web GIS-based Hydrograph Analysis Tool (WHAT) (https://engineerg.purdue.edu/~what) was developed (Lim et al., 2005) to provide a Web GIS interface for the 48 continental states in the USA for baseflow separation using a local minimum method, the BFLOW digital filter method, and the Eckhardt filter method. The Web Geographic Information System (GIS) version of the WHAT system accesses and uses U.S. Geological Survey (USGS) daily stream flow data from the USGS web server. To evaluate WHAT performance, the filtered results using the Eckhardt digital filter with default BFI_{max} value of 0.80 were compared with the results from the BFLOW filter method that was previously validated (Lim et al., 2005). The Nash–Sutcliffe coefficient values were 0.91, and the R^2 values were over 0.98

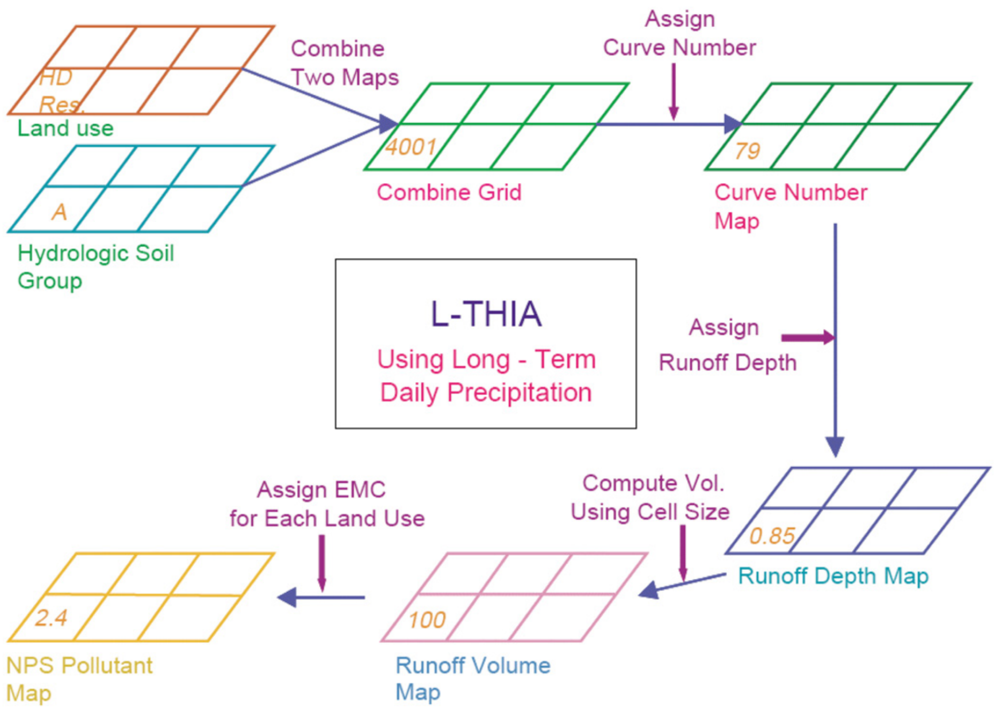

Fig. 1. Overview of L-THIA GIS system (adapted from Lim et al., 2001).

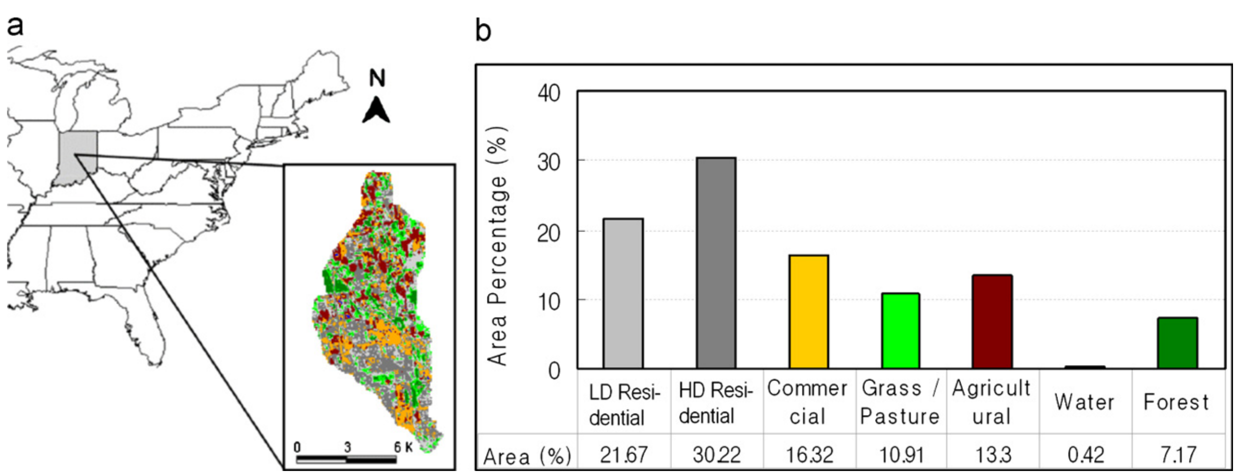

Fig. 2. Location of Little Eagle Creek (LEC) watershed (a) and its land use for 1991 (b) (adapted from Lim et al., 2006).

for 50 Indiana gaging stations. However, the use of the default BFI_{max} value of 0.80 for 50 Indiana gaging stations is not recommended because the filtered baseflow data were compared with the results from the BFLOW digital filter which did not reflect physical characteristics in the aquifer and watershed. Although the WHAT system cannot consider reservoir release and snowmelt that can affect stream hydrographs, the fully automated WHAT system can play an important role for sustainable ground water and surface water exploitation.

3. Methodology

3.1. Study area

In this study, the Little Eagle Creek (LEC) watershed near Indianapolis, Indiana was chosen for daily direct runoff comparison because it was used in the Lim et al. (2006) study for the comparison of L-THIA estimated direct runoff with the filtered direct runoff using the BFLOW digital filter. The LEC watershed is 70.5 km^2 in size (Fig. 2(a)). Fig. 2(b) shows the 1991 land uses for the LEC watershed. Urbanized land area in the LEC watershed was approximately 68% of the total land area in 1991 (Lim et al., 2005). The LEC land use data with the 1991 daily rainfall data were used in this study to simulate the effects on direct runoff estimation of using an improved BFI_{max} value for the LEC watershed rather than using the default BFI_{max} value of 0.80 provided in the WHAT system.

3.2. Baseflow separation using BFLOW and Eckhardt filters in the WHAT system

In the study by Lim et al. (2006), the L-THIA model was calibrated using BFLOW (Arnold and Allen, 1999) filtered direct runoff results. In the study described herein, the BFLOW filter and the Eckhardt filter (with a default BFI_{max} value of 0.80) were used to separate baseflow from the 1991 stream flow. The filtered direct runoff from BFLOW and the Eckhardt (with a default BFI_{max} value of 0.80) filters were compared to examine the agreement of L-THIA estimated direct runoff values with those from the BFLOW and the Eckhardt filter (with a default BFI_{max} value of 0.80).

3.3. Comparison of L-THIA estimated direct runoff using filtered direct runoff

In this study, the L-THIA model was calibrated using the BFLOW and the Eckhardt filtered direct runoff values. The L-THIA estimated direct runoff was compared with the BFLOW filtered direct runoff values as well as Eckhardt filtered direct runoff values to examine the validity of using the default BFI_{max} value of 0.80 in the Eckhardt digital filter. The BFLOW and Eckhardt filter parameters are not site specific and thus may not reflect the local hydrogeological conditions in the LEC watershed, Indiana, USA.

3.4. Development of genetic algorithm-based optimization for determination of optimum BFI_{max} and filter parameter values

One of the objectives of this study was to develop an automated module in the WHAT system using a Genetic Algorithm (Holland, 1975) technique to determine the optimum BFI_{max} value and filter parameter value to be used in the Eckhardt digital filtering method. The recession curve analysis method can be used to separate baseflow from long-term stream flow. However, this is time consuming and typically results in inconsistent results because it is not easy to identify the deflection point in the storm recession curves. Thus, a Genetic Algorithm-based optimization module to find the optimum BFI_{max} and filter parameter values (BFI_{max} Genetic Algorithm-Analyzer, BFI_{max} GA-Analyzer) was developed in this study. The BFI_{max} GA-Analyzer determines the optimum BFI_{max} value and filter parameter for the Eckhardt filter by comparing filtered baseflow with baseflow from recession curve analysis until it finds the maximum Nash–Sutcliffe coefficient value.

Genetic Algorithms (GA) were pioneered and developed by Holland (1975) at the University of Michigan in the 1960 and the 1970s (Holland, 1975). The GA is useful in finding the optimal solution to solve variable optimization problems through collective learning processes within a population of individual candidate solutions, and has been used for many researches (Chemin and Honda, 2006; Chung et al., 2009; Kollat and Reed, 2006; Matta, 2009). The GA is based on the principles of 'survival of the fittest', sets up a population of individuals to the problem, and attempts to create the individual for the 'best fitness'. There are three operators in the GA—selection, crossover, and mutation. With the selection operator, the population changes from poorer solutions to better solutions to remain. With the crossover operator, the genetic material of several pairs of solutions and some of their values are traded. It provides better fitness individuals a higher probability of being selected. Mutation alters a small percentage of individuals (one or more of their values) in the population (Holland, 1975; Georgieva and Jordanov, 2009; Wu et al., 2006).

The Eckhardt filter separates the baseflow from stream flow with the equation as shown below:

$$b_t = \frac{(1\text{-}BFI_{max})\alpha + b_{t-1}(1\text{-}\alpha)BFI_{max}Q_t}{1\text{-}\alpha BFI_{max}} \quad (1)$$

where b_t is the filtered baseflow at the t time step; b_{t-1} is the filtered baseflow at the $t-1$ time step; BFI_{max} is the maximum value of long-term ratio of baseflow to total stream flow; α is the filter parameter; and Q_t is the total stream flow at t time step.

The BFI_{max} GA-Analyzer module determines the optimum BFI_{max} value and filter parameter through a 500 generation GA analysis. The BFI_{max} GA-Analyzer module compares the filtered baseflow using the Eckhardt BFI_{max} and filter parameter values selected in each generation with the baseflow from recession curve analysis until the optimum BFI_{max} and filter parameter is determined (Fig. 3). The optimum BFI_{max} value and filter parameter values are transferred to the WHAT main interface for long-term hydrograph separation analyses.

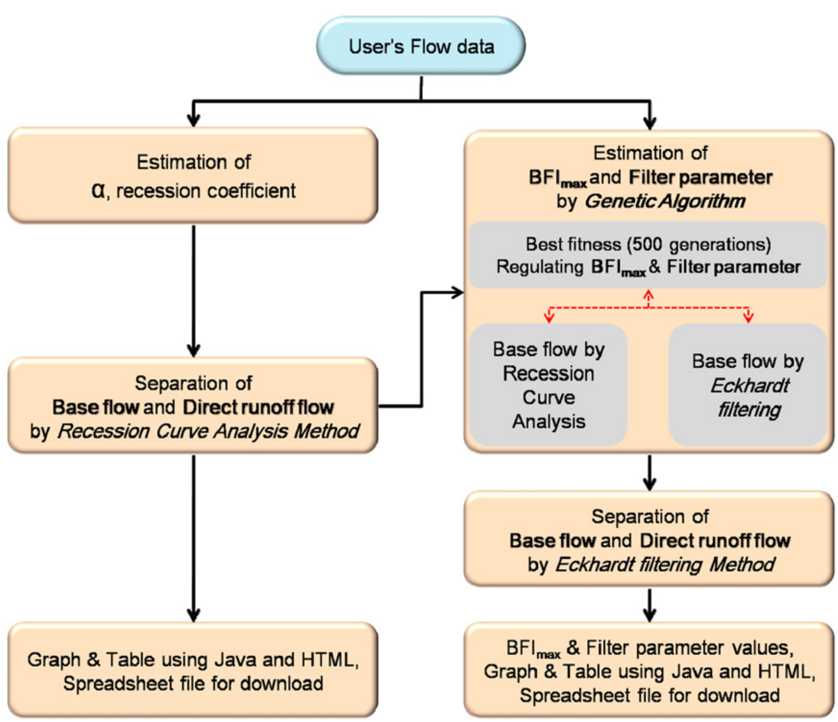

Fig. 3. WHAT BFI_{max} GA-Analyzer module developed to determine optimum BFI_{max} value and filter parameter.

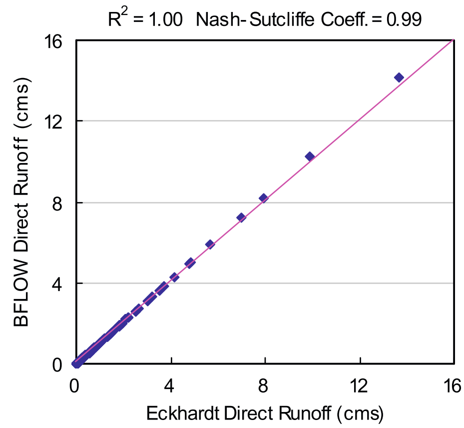

Fig. 4. Comparison of Eckhardt filtered direct runoff (with a default BFI_{max} value of 0.80 and a default filter parameter value of 0.98) and the BFLOW filtered results for Little Eagle Creek (LEC) watershed.

3.5. Hydrologic and water quality impacts of using optimum BFI_{max} parameter

The optimum BFI_{max} and filter parameter values for the LEC watershed were obtained using the automated recession curve analysis module and the BFI_{max} GA-Analyzer module developed in this study. The direct runoff with the optimized BFI_{max} and filter parameter values were used to calibrate the L-THIA model, because the optimum BFI_{max} and filter parameter values reflect baseflow recession curve characteristics of the LEC watershed well. The L-THIA results calibrated with the default BFI_{max} value of 0.80 and filter parameter value of 0.98 were compared with those obtained with the optimized BFI_{max} and filter parameter values to examine potential errors in L-THIA estimated direct runoff with use of the default BFI_{max} and filter parameter values in the Eckhardt digital filter. As with most hydrologic/water quality models, the water quality module of the L-THIA model is affected by the hydrologic

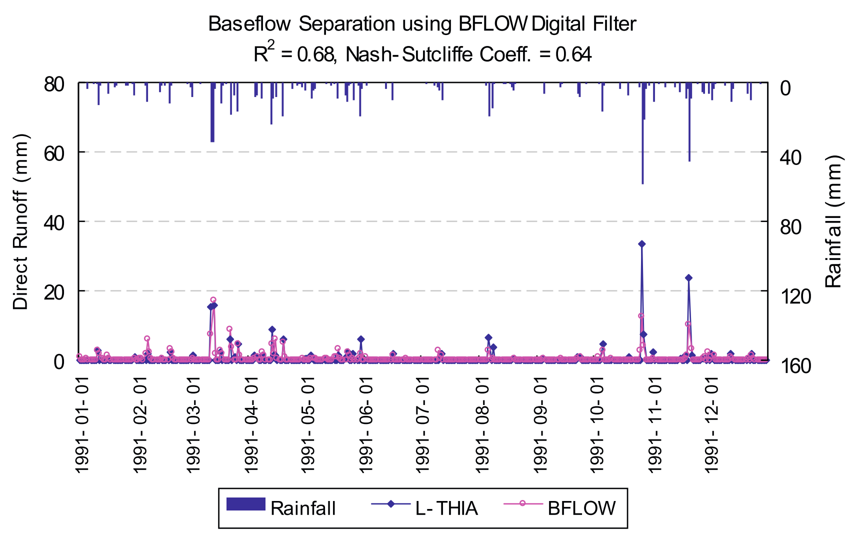

Fig. 5. Comparison of L-THIA estimated direct runoff with BFLOW filtered direct runoff.

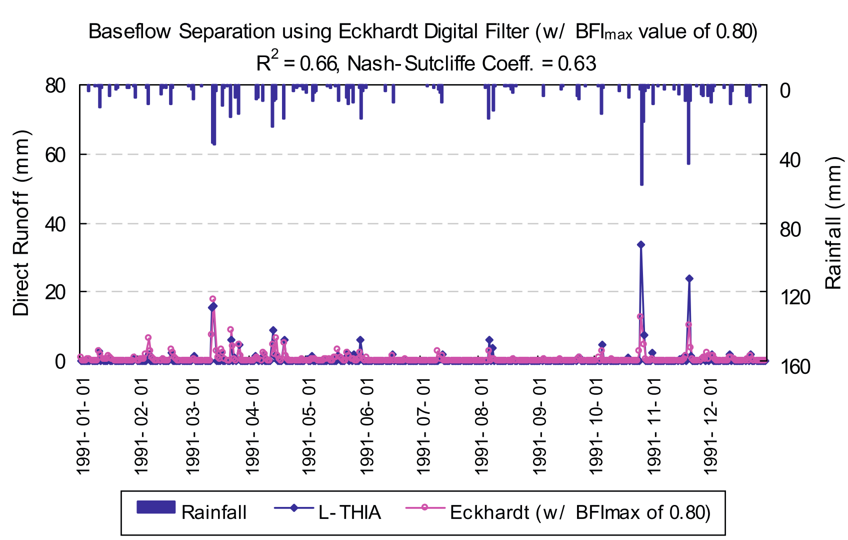

Fig. 6. Comparison of L-THIA estimated direct runoff with Eckhardt filtered direct runoff with a default BFI_{max} value of 0.80.

component of the L-THIA system. Thus, the water quality impact with the use of optimized BFI_{max} and filter parameter values was also investigated.

4. Results

4.1. Baseflow separation using BFLOW and Eckhardt filters using WHAT system

The default BFI_{max} value of 0.80 is provided by Eckhardt (2005) for perennial streams with porous aquifers. Thus, the default BFI_{max} value of 0.80 is commonly used in separating baseflow from observed flow by most users. The filtered direct runoff values using the BFLOW filter and the Eckhardt filter (with a default BFI_{max} value of 0.80 and a default filter parameter value of 0.98) were compared. The R^2 value was 1.00 and the Nash–Sutcliffe coefficient was 0.99 for the comparison of the BFLOW filtered direct runoff and the Eckhardt filtered direct runoff (with a default BFI_{max} value of 0.80) as shown in Fig. 4, indicating the calibrated CN values for the LEC watershed using the BFLOW filtered direct runoff will also work well with the Eckhardt filtered direct runoff results (with a default BFI_{max} value of 0.80 and filter parameter value of 0.98).

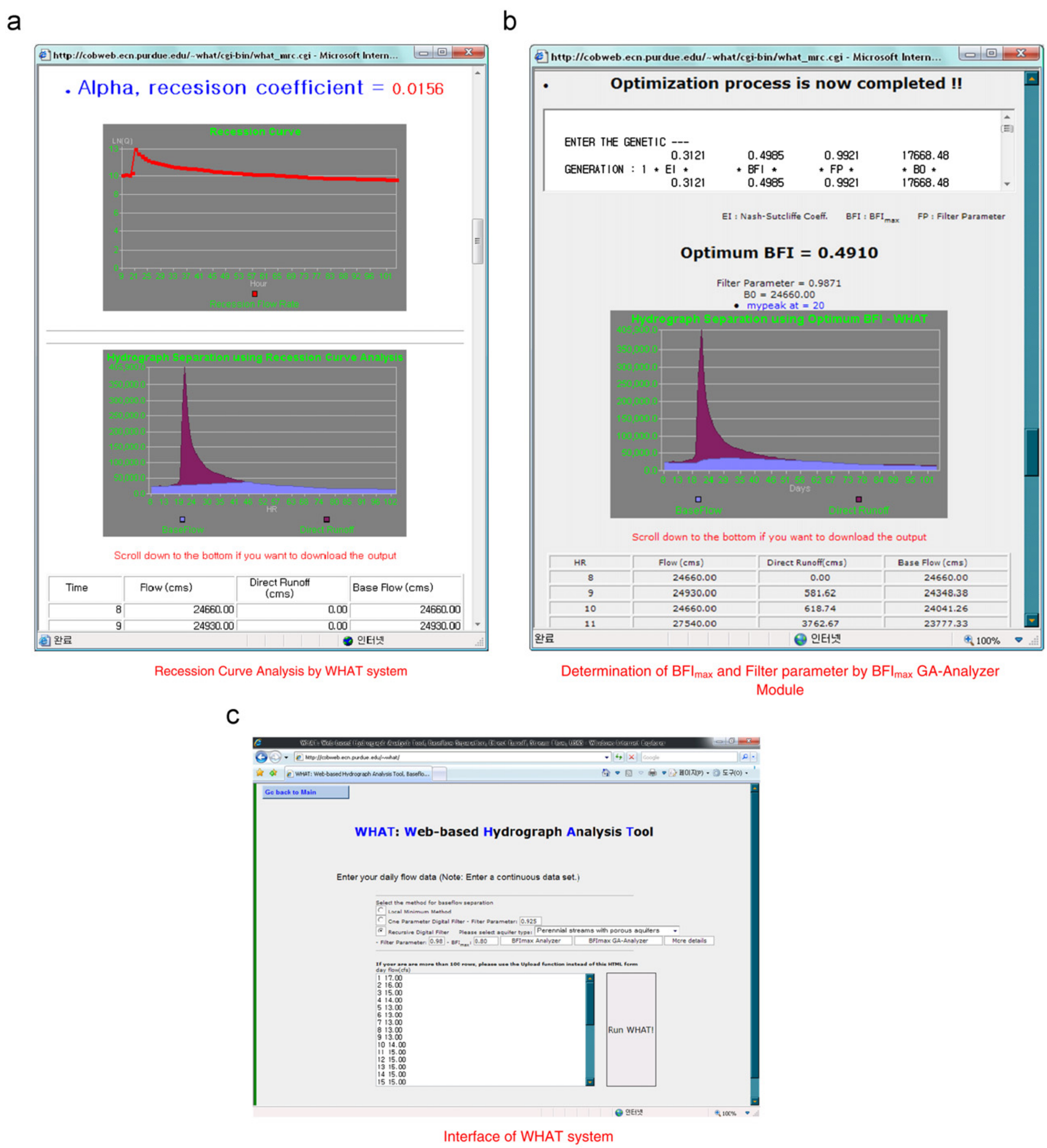

Fig. 7. Automated recession curve analysis and BFI_{max} GA-Analyzer module to determine BFI_{max} parameter and filter parameter of Eckhardt digital filter in WHAT web system (https://engineering.purdue.edu/~what).

4.2. Comparison of the L-THIA estimated direct runoff using filtered direct runoff

The CN values were adjusted for best fits between the L-THIA estimated direct runoff and the BFLOW and the Eckhardt (with a default BFI_{max} value of 0.80 and a default filter parameter of 0.98) filtered direct runoff values. Fig. 5 shows the comparison of the L-THIA estimates with the BFLOW filtered values. The R^2 value was 0.68 and the Nash–Sutcliffe coefficient was 0.64, indicating the L-THIA model is capable of simulating direct runoff for the LEC watershed reasonably well. Santhi et al. (2001) suggested a Nash–Sutcliffe value of 0.5 or higher indicates a hydrologic model performs at an acceptable level. Fig. 6 shows the comparison of L-THIA runoff estimates with the Eckhardt filtered values with a default BFI_{max} value of 0.80 and a default filter parameter value of 0.98. The R^2 value was 0.66 and the Nash–Sutcliffe coefficient was 0.63. Thus, one can infer that the L-THIA model performs well in predicting the direct runoff for the LEC watershed since the R^2 and the Nash–Sutcliffe coefficient values were higher than 0.60 in both cases. The similar R^2 and Nash–Sutcliffe coefficient values were because of similarity in filtered direct runoff values using the BFLOW filter and the Eckhardt filter with a default BFI_{max} value of 0.80 as shown in Fig. 4.

However, the BFLOW filtering method and the Eckhardt filtering method do not accurately reflect hydrogeological conditions in the LEC watershed when separating the baseflow from the stream flow. Therefore, the accuracy in filtered baseflow needs to be investigated to provide reliability in L-THIA estimated direct runoff.

4.3. Development of genetic algorithm-based optimization for determination of optimum BFI_{max} parameter

To determine the optimum BFI_{max} parameter, the recession curve method was automated and the BFI_{max} GA-Analyzer module was developed and integrated into the WHAT system as shown in Fig. 7. Users are required to provide or upload stream flow data into the WHAT system for recession curve analysis (Fig. 7(a)). The BFI_{max} GA-Analyzer module determines the optimum BFI_{max} value and filter parameter value as shown in Fig. 7(b). Then, the optimum BFI_{max} value is transferred to the main WHAT interface automatically for users' convenience. This fully automated "BFI_{max} GA-Analyzer" was integrated into the WHAT interface (Fig. 7(c)).

Fig. 8 shows the stream flow data for the LEC watershed and the filtered baseflow data from the recession curve analysis. Baseflow separated using the recession curve analysis takes 58.03% of the total stream flow. Using the automated recession curve analysis tool and the BFI_{max} GA-Analyzer module developed in this study (Fig. 7), the optimum BFI_{max} value of 0.491 and the filter parameter value of 0.98 for the LEC watershed were determined. Fig. 9 shows the filtered baseflow using the optimum BFI_{max} value of 0.491 and filter parameter value of 0.987. Baseflow using the optimum BFI_{max} value takes 58.03% of the total stream flow. The minor differences in the filtered baseflow (Figs. 8 and 9) were because the recession curve analysis determines the deflection point of the recession curve as an ending point of direct runoff, while the Eckhardt digital filter does not consider any physical aquifer characteristics. However, the BFI_{max} value of 0.491 is the optimum value for the LEC watershed compared with the filtered baseflow using the recession curve analysis.

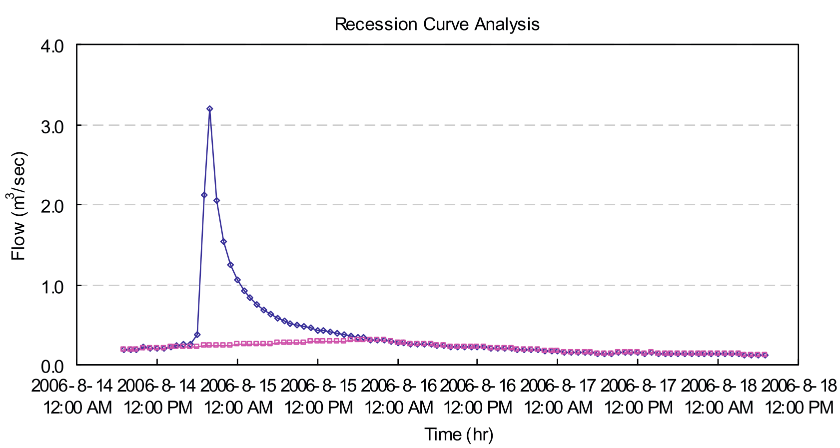

Fig. 8. Baseflow separation using recession curve analysis.

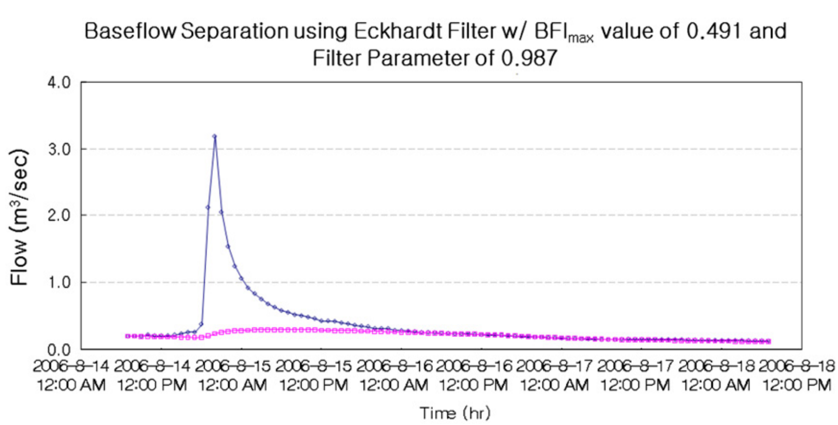

Fig. 9. Baseflow separation using Eckhardt digital filter with optimized BFI_{max} value of 0.491 and filter parameter value of 0.987.

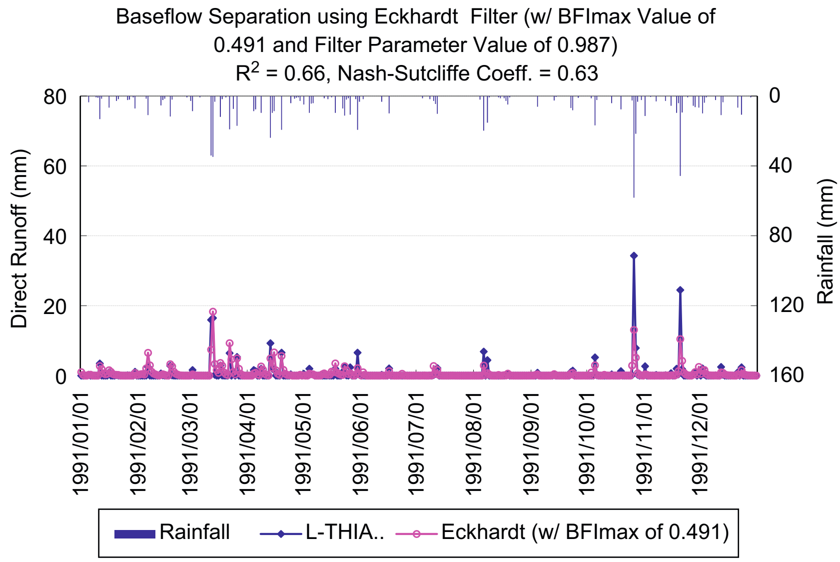

Fig. 10. Comparison of L-THIA estimated direct runoff with Eckhardt filtered direct runoff with an optimized BFI_{max} value of 0.491 and filter parameter value of 0.987.

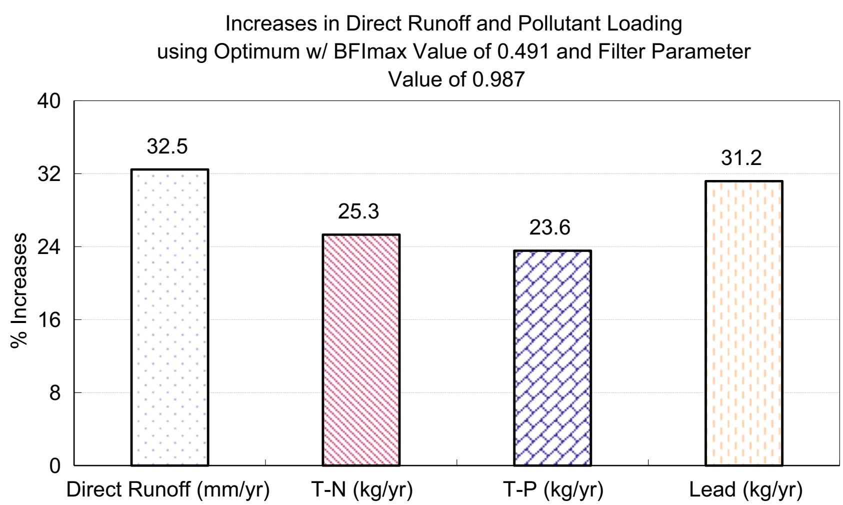

Fig. 11. Increases in direct runoff and pollutant loading using an optimum BFI_{max} value of 0.491 and filter parameter value of 0.987.

4.4. Hydrologic and water quality impacts of using optimum BFI_{max} parameter and filter parameter

The L-THIA model was calibrated with the direct runoff separated using the Eckhardt filter with the optimized BFI_{max} value of 0.491 and the optimized filter parameter value of 0.987. The R^2 value was 0.66 and the Nash–Sutcliffe coefficient value was 0.63 as shown in Fig. 10. These statistics are the same as those from the Eckhardt filter with the default BFI_{max} value of 0.80. However, the L-THIA direct runoff values calibrated using the optimized BFI_{max} value of 0.491 is greater than those with a default BFI_{max} value of 0.80 by 32.5%, although the R^2 and Nash–Sutcliffe coefficient between the L-THIA estimated direct runoff and the Eckhardt filtered direct runoff with different BFI_{max} values (default value of 0.80 and optimized value of 0.491) are the same.

This result indicates that the L-THIA estimated direct runoff values with acceptable performance statistics (Nash–Sutcliffe coefficient value of 0.60 or higher, Santhi et al., 2001) for the comparison with filtered direct runoff values may not match the "true" direct runoff values. There could be substantial differences in the L-THIA estimated direct runoff as shown in this study. This mismatch will also result in errors in estimated pollutant loads using the L-THIA model. The water quality impacts of using the optimum BFI_{max} value were computed for the LEC watershed as shown in Fig. 11. The total N loadings increased by 25%, total P by 23%, and total Lead by 31% with use of an optimum BFI_{max} value of 0.491. Thus, neglecting validation of the accuracy in baseflow separation using the non-physically based digital filtering methods can result in estimated differences of more than 20% in NPS loadings as shown in the LEC watershed.

5. Conclusions and discussion

In this study, the optimum BFI_{max} value and filter parameter value of the Eckhardt digital filter for the LEC watershed were determined with the BFI_{max} GA-Analyzer module developed in this study. Although the comparison between the L-THIA estimates and the Eckhardt filtered direct runoff with the default BFI_{max} value of 0.80 and a default filter parameter value of 0.98 provided the same accuracy (Nash–Sutcliffe coefficient values of 0.63 in both cases) compared with those with the optimized BFI_{max} value of 0.491 and filter parameter value of 0.987, there was a 33% increase in the L-THIA estimated direct runoff with use of an optimized BFI_{max} value and filter parameter values for the LEC watershed. Estimated NPS pollutant loadings were more than 20% greater with the optimum BFI_{max} value and filter parameter value for the LEC watershed. The results obtained in this study indicate that hydrologic and water quality models, such as L-THIA, should be validated with properly measured hydrologic data before their application in land use and water quality management planning as shown in this study.

The LEC watershed optimum BFI_{max} value of 0.491 and filter parameter value of 0.987 were determined with the automated recession curve analysis and the BFI_{max} GA-Analyzer module of the WHAT system. The current WHAT version developed in this study only considers single storm hydrographs. Thus, the automated recession curve analysis method should be extended for use with multiple storm hydrographs or long-term storm hydrographs to derive information on the Master Recession Curve. It is expected that this functionality will enhance the accuracy in filtered baseflow using the Eckhardt digital filter. The WHAT Web GIS system (https://engineering.purdue.edu/~what) can be used in the calibration and validation processes of hydrologic and water quality models.

Acknowledgement

This work was supported by a grant (code: 2-2-3) from Sustainable Water Resources Research Center of 21st Century Frontier Research Program.

References

Arnold, J.G., Allen, P.M., 1999. Validation of automated methods for estimating baseflow and groundwater recharge from stream flow records. Journal of American Water Resources Association 35 (2), 411–424.

Arnold, J.G., Allen, P.M., Muttiah, R., Bernhardt, G., 1995. Automated baseflow separation and recession analysis techniques. Groundwater 33 (6), 1010–1018.

Bendient, P.B., Huber, W.C., 2002. Hydrology and Floodplain Analysis third ed Prentice Hall, New Jersey.

Bhaduri, B., Minner, M., Tatalovich, S., Harbor, J., 2001. Long-term hydrologic impact of urbanization: a tale of two models. Journal of Water Resources Planning and Management 127, 13–19.

Bicknell, B.R., Imhoff, J.C., Kittle Jr., J.L., Donigian Jr., A.S., Johanson, R.C., 1997. Hydrological Simulation Program—Fortran, User's Manual for Version 11. US Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA EPA/600/R-97/080, 755 pp..

Chapman, T.G., 1987. Unit hydrograph identification using only streamflow data. Transactions of the Institution of Engineers, Australia CE29 (3), 187–191.

Chemin, Y., Honda, K., 2006. Spatiotemporal fusion of rice actual evapotranspiration with genetic algorithms and an agrohydrological model. IEEE Transactions on Geoscience and Remote Sensing 44 (11), 3462–3469.

Chung, J.W., Oha, S.M., Choi, I.C., 2009. A hybrid genetic algorithm for train sequencing in the Korean railway. Omega 37, 555–565.

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes 19 (2), 507–515.

Georgieva, A., Jordanov, I., 2009. Global optimization based on novel heuristics, low-discrepancy sequences and genetic algorithms. European Journal of Operational Research 196, 413–422.

Harbor, J., 1994. A practical method for estimating the impact of land use change on surface runoff, groundwater recharge and wetland hydrology. Journal of American Planning Association 60, 91–104.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, MI 183 pp.

Kollat, J.B., Reed, P.M., 2006. Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. Advances in Water Resources 29, 792–807.

Lim, K.J., Engel, B.A., Kim, Y., Harbor, J., 2001. Development of the long-term hydrologic impact assessment (L-THIA) WWW systems. In: Stott, D.E., Mohtar, R.H., Steinhardt, G.C. (Eds.), Sustaining the Global Farm—Selected papers from the 10th International Soil Conservation Organization Meeting, West Lafayette, IN, International Soil Conservation Organization in Cooperation with the USDA and Purdue University, West Lafayette, IN, pp. 1018–1023.

Lim, K.J., Engel, B.A., Tang, Z., Choi, J., Kim, K., Muthukrishnan, S., Tripathy, D., 2005. Automated web GIS-based hydrograph analysis tool, WHAT. Journal of the American Water Recourse Association 41 (6), 1407–1416.

Lim, K.J., Engel, B.A., Tang, Z., Muthukrishnan, S., Choi, J., Kim, K., 2006. Effects of calibration on L-THIA GIS runoff and pollutant estimation. Journal of Environmental Management 78 (1), 35–43.

Lyne, Y.D., Hollick, M., 1979. Stochastic time-variable rainfall-runoff modeling, Proceedings Hydrology and Water Resources Symposium. Institution of Engineers Australia, Perth, Australia, pp. 89–92.

Matta, M.E., 2009. A genetic algorithm for the proportionate multiprocessor open shop. Computers & Operation Research 36, 2601–2618.

Nathan, R.J., McMahon, T.A., 1990. Evaluation of automated techniques for baseflow and recession analysis. Water Resources Research 26 (7), 1465–1473.

Pandey, S., Gunn, R., Lim, K.J., Engel, B.A., Harbor, J., 2000. Developing a web-enabled tool to assess long-term hydrologic impact of land use change: information technologies issues and a case study. Urban and Regional Information Systems Journal 12 (4), 5–17.

Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R., Hauck, L.M., 2001. Validation of the SWAT model on a large river basin with point and nonpoint sources. Journal of the American Water Resources Association 37 (5), 1169–1188.

Sloto, R.A., Crouse, M.Y., 1996. HYSEP: a computer program for stream flow hydrograph separation and analysis. U.S. Geological Survey, Water-Resources Investigation Report 96-4040, Reston, VA, 46 pp.

U.S. EPA., 2001. Our built and natural environment: a technical review of the interactions between land use, transportation, and environmental quality. U.S. Environmental Protection Agency Report EPA 231-R-01-002, Washington, DC, 102 pp. Available at 〈http://www.epa.gov/dced/pdf/built.pdf〉. Accessed in September 2006.

Wu, J., Zheng, C., Chien, C.C., Zheng, L., 2006. A comparative study of Monte Carlo simple genetic algorithm and noisy genetic algorithm for cost-effective sampling network design under uncertainty. Advanced in Water Resources 29, 899–911.