Analyzing the relationship between peak runoff discharge and land-use pattern – a spatial optimization approach

I.-Y. Yeo\textsuperscript{1} and J.-M. Guldmann\textsuperscript{2}

\textsuperscript{1}Department of Geography, The University of Maryland, College Park, Maryland, USA
\textsuperscript{2}Department of City and Regional Planning, The Ohio State University, Columbus, Ohio, USA

Received: 30 March 2009 – Accepted: 1 April 2009 – Published: 28 April 2009

Correspondence to: I.-Y. Yeo (iyeo@umd.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

This paper investigates the impacts of land-use patterns on watershed hydrology and characterizes the nature of this relationship. The approach combines a spatially explicit, process-based hydrological simulation model, a land-use optimization model, the Integrated Hydrological and Land-Use Optimization (IHLUO) model, and an extensive GIS database. Numerical experiments are conducted to assess changes in the peak discharge rate under various spatial land-use arrangements, and to delineate the optimal land distribution that minimizes the peak discharge. The area of application is a catchment of the Old Woman Creek watershed in the southwestern coastal area of Lake Erie, OH. The global optimality of the delineated land pattern at a 30-m resolution is evaluated using a combinatorial statistical method. A large number of solutions has been generated from clearly different initial solutions, and these solutions turn out to be very close to each other, strongly supporting the case for a convex relationship between peak discharge and land-use pattern. The Weibull distribution is used to generate a point estimate of the global optimal value and its confidence interval. The peak discharge function is further examined in light of the underlying physics used in the simulation model.

1 Introduction

Understanding watershed hydrology processes and their linkages to land cover changes is important for controlling nonpoint source (NPS) water pollutants, which are produced by land-using activities (Novotny, 2003) and carried into waterways by stormwater runoff. Effective management of NPS pollution calls for efforts to identify pollution sources and pathways, and to minimize production of pollutants and their delivery to waterways. Complex watershed processes have been modeled mathematically or empirically, leading to computer simulation models that can provide key information in better understanding the interactions among the various physical systems in
the watershed. Such models are important to predict changes in the watershed system, due to land-use or management practice changes (Beven, 1989; Grayson et al., 1992), and have been widely utilized to develop area-wide Best Management Practices (BMP) or predict the hydrological effects of future land-use and climate changes (Quilbé et al., 2008). However, these models cannot explicitly link pollution sources and yields, and consider necessarily a limited number of scenarios, which cannot lead to the optimal selection and location of BMPs.

Some recent studies (Srivastava et al., 2002; Nicklow and Muleta, 2001; Muleta and Nicklow, 2002; Seppelt and Voinov, 2002; Kaur et al., 2004) have attempted to overcome the limitation of a scenario-based approach by integrating optimization methods into simulation models, but have failed to explicitly relate pollution sources and yields at a high level of geographical disaggregation. More recently, Yeo et al. (2004) have developed the Integrated Hydrological and Land-Use Optimization (IHLUO) model to delineate high-resolution (30 m) land-use patterns that minimize the peak discharge flow. Building upon a spatially explicit hydrological model, the IHLUO model identifies critical pollution sources and pathways, while accounting for the physical heterogeneity of and the spatial dynamics across the watershed. The IHLUO model has been applied to a catchment of the Old Woman Creek (OWC) watershed in the southwestern coastal area of Lake Erie (OH), and numerical results describing the emergent properties of the optimal land use patterns under various storm sizes have been reported. Yeo et al. (2007) have also integrated the IHLUO model into a hierarchical land-use optimization scheme, where the allocations at the higher levels (subwatersheds, catchments) for larger areas are implemented via quadratic programming, with peak runoff functions estimated using simulation-generated pseudo data. This earlier research has demonstrated the importance of the spatial configuration of land use in controlling peak storm water runoff, with the optimal land pattern reducing the peak discharge rate at the watershed outlet by more than 40% under a 1-year storm, as compared to the existing land-use pattern.

The purpose of this research is to further investigate (1) the properties of the optimal
solution, in particular to assess its global optimality, and (2) the functional relationship between the spatial distribution of land uses and the peak runoff flow to support the case for global optimality. Systems analysis approaches such as spatial optimization methods have been efficient in selecting types and placement of BMPs, but without ensuring the global optimality of the obtained solution. These problems involve the simulation of highly nonlinear dynamics across the landscape (with no assumption of equilibrium) and of physical processes that cannot be expressed with analytical functions. Therefore, the obtained solution cannot be evaluated using the standard mathematical conditions for global optimality, and the quality of a solution has been addressed by merely demonstrating algorithmic convergence of the solution. Assurance of the global optimality of the nonlinear problem solution is critical to provide confidence in decision making and to further extend integrated optimization approaches to watershed systems. This requires understanding the nature of the relationship between the watershed system and the decision variables.

This paper presents an innovative combination method that uses a numerical procedure and the Weibull distribution to assess the convergence of a solution toward the global optimum. The results obtained from the IHLUO model suggests that the functional relationship between the spatial distribution of land uses and the peak runoff flow is convex, an important implication in hydrology. Examining the underlying physics of the hydrological processes used in the simulation model provides additional support for convexity and global optimality. The hydrological feedbacks from various land patterns are analyzed at a 30-m resolution. A 1-year design storm is selected, because land management alone, without any other constructive BMP, is effective only under a small-size storm (see the pilot study in Yeo et al., 2004 to analyze optimal land allocations under different storm sizes).

The remainder of this paper is organized as follows. Section 2 describes the modeling methodology. A brief discussion of the spatially explicit hydrologic model and the IHLUO model is presented in Sects. 2.1 and 2.2, and the statistical basis for assessing the global optimality of the obtained solution using the Weibull distribution is discussed.
in Sect. 2.3. The results of a numerical application on a small catchment in the OWC watershed are reported and analyzed in light of the underlying physics of the hydrological model in Sect. 3. The findings and limitations of this research are summarized in Sect. 4.

2 Simulation-optimization methodology

2.1 A spatially explicit hydrological model

The relationship \( f(X) \) between a land-use pattern \((X)\) and the resulting peak discharge rate at the watershed outlet is explored through a spatially explicit hydrological model. Since such a relationship is very difficult to derive from field studies, a process-based computer model is developed to simulate this relationship, by modifying the SCS curve number (CN) method. The CN method is chosen, because (1) the land-water relationship is directly expressed in terms of hydrologic soil groups and land use/cover conditions (McCuen, 1982; USDA, 1986; Bingner and Theurer, 2001), and (2) it meets computing resources requirement for hydrologic simulations and optimization. Due to its simplicity and accuracy, it has been widely utilized and embedded into various watersheds for hydrology, flood analysis, and water quality modeling (Garen and Moore, 2005), including Soil and Water Assessment Tools (SWAT) (Arnold et al., 1998), AGRicultural Non-Point Source Pollution Model (AGNPS) (Bingner and Theurer, 2001; Young et al., 1989), and Erosion Productivity Impact Calculator or Environmental Policy Integrated Climate (EPIC) (Williams et al., 1984; Williams and Meinardus, 2004).

The conventional CN method yields lumped effects by taking weighted averages of the modeling parameters. To account for the impacts of the spatial variability in land-use changes, the modeling parameters are assigned to each cell (30 m) without any spatial aggregation or averaging. The volume of runoff \((Q)\) is computed as:

\[
Q = \frac{(P - 0.2S)^2}{P + 0.8S},
\]  

(1)
where $P$ is the precipitation, and $S$ is the moisture retention, estimated from the runoff curve number (CN), with:

$$S = 254 \left[ \frac{100}{CN} - 1 \right].$$

(2)

The quantities $P$, $Q$, and $S$ are measured in millimeters [mm]. Note that groundwater flows are not modeled, and the antecedent soil moisture condition is considered by using the default estimation of the SCS method (USDA, 1986), setting $I_a=0.2S$ ($I_a$=initial abstraction, in inches). Details are provided in McCuen (1982), USDA (1986), and Bingner and Theurer (2001).

Runoff flows are accumulated following the flow paths determined by topography. The flow routing direction is determined by the D-8 method (i.e., eight flow directions), which assigns the runoff on a given cell to the lowest-elevation cell among the eight surrounding cells (O’Callaghan and Mark, 1984). The runoff process is analyzed at the flow cell level, then at the flow path level, then at the catchment/watershed level. The runoff over a flow path is obtained by summing up the storm runoffs occurring at all the cells along the flow path. The total runoff volume at the watershed outlet is obtained by summing up the runoffs occurring along all the flow paths within the watershed (Olivera, 1996).

A similar approach is applied to estimate the time of concentration, $T_c$ (i.e., the longest hydraulic distance to the watershed outlet). Rather than computing $T_c$ from the predefined longest geographical distance from the watershed outlet, the simulation model calculates it by keeping track of the flow time for every pathway to better account for the spatial variability in land configuration. The travel time of a flow path to the watershed outlet is calculated by summing up the travel times for all cells along the path. The maximum travel time across all paths is selected as the time of concentration. The travel time for each cell is determined according to its flow type: overland flow, shallow concentrated flow, or channel flow (USDA, 1986). Surface roughness, channel shapes, flow patterns, and slopes are the main factors affecting time of concentration and flow time. Land use mainly affects surface roughness (i.e., Manning’s $n$ coefficient), and,
therefore, runoff flow velocity and travel time. Overland flow is estimated using Manning’s kinematic solution, and channel flow is calculated with Manning’s equation. The velocity from a shallow concentrated flow cell is calculated with the empirical and explicit formula developed by USDA (1986), and Bingner and Theurer (2001). Using field observations, the shallow concentrated flow cell velocity is expressed as a function of watercourse slope and type of channel. See TR-55 for details of this method (USDA, 1986).

After calculating the time of concentration and the total amount of runoff, the peak runoff rate is determined using the extended TR-55 procedure (Bingner and Theurer, 2001), with:

\[
Q_p = 2.78 \cdot 10^{-3} P_{24} D_a \left[ \frac{a_p + (c_p \cdot T_c) + (e_p \cdot T_c^2)}{1 + (b_p \cdot T_c) + (d_p \cdot T_c^2) + (f_p \cdot T_c^3)} \right] 
\]  

(3)

where \(Q_p\) is the peak discharge [m³/s], \(D_a\) the area of the spatial unit [ha], \(P_{24}\) the 24 h effective rainfall over the total drainage area [mm], \(T_c\) the time of concentration [h] it is our standard for hour, and the coefficients, \(a_p, b_p, c_p, d_p, e_p,\) and \(f_p,\) are determined by the ratio of initial abstraction (\(I_a\)) to 24-h precipitation (\(P_{24}\)). See Bingner and Theurer (2001) for the values of these coefficients. See Yeo et al. (2004) for further description of the hydrological model.

2.2 Integrated Hydrological and Land-Use Optimization (IHLUO) model

An optimization method is integrated with the spatially explicit hydrological model to delineate the land-use pattern that minimizes the peak storm runoff at the watershed outlet, under given land-use constraints, with:

\[
\text{Min } f(X) = \text{Peak Runoff Rate} 
\]  

(4)
Subject to:
\[ \sum_{i=1}^{n} x_{il} = T_l \quad \forall l = 1, 2, \ldots, m \]  
(5)

\[ \sum_{l=1}^{m} x_{il} = A_i \quad \forall i = 1, 2, \ldots, n \]  
(6)

where \( x_{il} \) is the amount of land use \( l \) in cell \( i \), \( T_l \) is the total amount of land use \( l \) in the watershed area, and \( A_i \) is the area of cell \( i \). \( f(X) \) is the objective function that evaluates the peak discharge at the watershed outlet resulting from the land-use pattern \( X \), is nonlinear, cannot be expressed analytically, and is numerically evaluated using the runoff simulation model. Constraint (5) guarantees the achievement of watershed land-use targets and constraint (6) the full coverage of cell \( i \) by land uses.

The nonlinear program (4–6), with an objective function that can only be evaluated through simulation, is solved using sequential linear programming, also known as the method of convex combinations (Wagner, 1976; Venkataraman 2001). A linear approximation of the nonlinear objective function is based on a first-order Taylor’s series expansion, with:

\[ f(Z) \approx f(X^k) + \sum_{j=1}^{J} \frac{\partial f(X^k)}{\partial x_j} (z_j - x_j^k) , \]  
(7)

where \( f(Z) \) is the linear approximation of the original function, \( X^k = \{x_j^k\} \) is the \( k \)-th trial point, and \( Z = \{z_j\} \) is any point in the neighborhood of \( X^k \). The goal is to minimize the second term in (7), subject to constraints (5–6), and the model becomes a linear program.

This linear approximation is not straightforward, because the objective function is not analytically closed. The partial derivatives are approximated by numerically evaluating
the changes in the objective function (storm runoff at the watershed outlet) resulting from very small changes in land uses, cell by cell and land use by land use. This procedure generates a mixture of land uses in each cell. The objective function is, each time, computed using the hydrological simulation model. The optimal step size along the gradient is determined by evaluating the objective function for steps varying in the interval [0–1], with an increment of 0.01. See Yeo et al. (2004) for further details.

2.3 Assessment of local optima for global optimality

The IHLUO model is nonlinear, and the gradient-based algorithm only guarantees locally optimal solutions. In the case of minimization, this algorithm produces the globally optimal solution only if the objective function is convex over the feasible domain, which requires that its Hessian matrix be negative definite everywhere (Intriligator, 1976). The Hessian matrix is made up of the second-order partial derivatives of the objective function. Since the IHLUO objective function cannot be expressed analytically, it is impossible to directly assess its convexity.

However, it is possible to assess how close a local optimum is to the global one by using a combinatorial statistical method, developed by Golden (1978) and Los and Lardinois (1982). The idea is to devise a systematic procedure to generate independent local optima and to obtain a point estimate for the global optimum by applying statistical extreme-value theory. Golden (1978) investigates how far a heuristic solution is from the global optimum in the case of the traveling salesman problem. Los and Lardinois (1982) apply a hill-climbing technique to solve the optimal network design problem. In this research, an initial solution (i.e., an initial land-use pattern) is generated by randomly assigning land uses to cells, while keeping the total amount of each land use constant at the watershed level, and a local optimum is then generated by the IHLUO model.

The local optima generated with various initial states are statistically analyzed, using a Weibull Distribution. This distribution has been widely used to analyze survival data, to assess stability, and to measure risk (Aitkin and Clayton, 1980). In hydrology, the
Weibull distribution has been utilized to fit hydrographs (Bhunya et al., 2006) and analyze the trend in extreme events, including annual peak discharges, annual minimum flows, and annual maximum rainfall intensities (Clarke, 2002). One of its merits is its independence from the parent distribution. It is derived as the asymptotic distribution of the smallest order statistics, known as Type III (Fisher and Tippett, 1928; Gumbel, 1958). The Weibull cumulative distribution function \( \Phi(x) \) and its density function \( \delta(x) \) are:

\[
\Phi(X) = 1 - \exp \left[ - \left( \frac{X - a}{b} \right)^c \right] \tag{8}
\]

\[
\delta(X) = \frac{c}{b} \left( \frac{X - a}{b} \right)^{c-1} \exp \left[ - \left( \frac{X - a}{b} \right)^c \right] \tag{9}
\]

\( X \geq a > 0; c > 0, b > 0. \)

The parameters \( a, b, \) and \( c \) are defined as the location, scale, and shape parameters, with \( \Phi(a+b) = 1-e^{-1}(\approx 0.63). \) The parameter \( a, \) the lower bound of the Weibull Distribution, is also the lower bound of the parent-distribution, and is considered as the global optimum (minimum).

The best local optimum is used as an estimate of the global optimum, and its reliability is statistically evaluated by analyzing the empirically fitted Weibull distribution of local optima. If \( R \) observations (i.e., sample size) are drawn from a Weibull distribution, and if \( x_{(1)}^h \), the best local optimum, is the first element (i.e., smallest number) in an ordered set, it follows that:

\[
\Pr\{x_{(1)}^h \leq a + b\} = 1 - \Pr\{x_{(1)}^h > a + b\} \nonumber \\
= 1 - \left\{1 - \Phi_{x_{(1)}^h}(a + b)\right\}\left\{1 - \Phi_{x_{(2)}^h}(a + b)\right\} \ldots \left\{1 - \Phi_{x_{(R)}^h}(a + b)\right\} = 1 - \left(1 - e^{-1}\right)^R = 1 - e^{-R}. \tag{10}
\]
The estimator $\hat{a}$ for the location parameter $a$ is used as point estimator for the global optimum ($\hat{x}^*$). Its confidence interval with a significance level of $100(1-e^{-R})\%$ is (Golden and Alt, 1979; Dergis, 1985):

$$\Pr\{x_h^{(1)} - \hat{b} \leq \hat{x}^* \leq x_h^{(1)}\} \approx 1 - e^{-R}. \quad (11)$$

However, the interval calculated by this formula is too large. It was tightened by Los and Lardinois (1982), using a real number $S$, with:

$$\Phi \left( a + b / S \right) = 1 - \exp \left[ -\left(1 / S\right)^c \right]. \quad (12)$$

Then, the confidence interval is given as:

$$\Pr\left\{ x_h^{(1)} - \hat{b} / S \leq \hat{x}^* \leq x_h^{(1)} \right\} = 1 - \exp \left( -R / S^c \right). \quad (13)$$

If the confidence level $(1-\alpha)$ is expressed as

$$1 - \exp \left( -R / S^c \right) = 1 - \alpha, \quad (14)$$

the real number $S$ is

$$S = (-R / \ln \alpha)^{1/\hat{c}}. \quad (15)$$

Estimating the global optimum value and its confidence interval can then be used to assess convergence toward the global optimum.

### 3 Numerical application

#### 3.1 Data

The methodology is applied to a small catchment in the Old Women Creek watershed (Ohio). Due to the large computing requirements of the IHLUO model and the need to generate many independent local optima for statistical assessment, the numerical...
application is performed on a small catchment, with a few land-use categories, a simple drainage network, and simple spatial distributions of land uses and soil types.

This catchment is overlaid by a grid of 1732 30-m cells, with land-use/cover classified into three categories—agriculture, conservation, and urban. The catchment is predominantly agricultural (78%), conservation land uses (grass/woods) represent 12.6% of the area, urban land use makes up 1.5%, and the remainder represents water (8%). Roads make up most of the urban land (25 cells), except for one cell of built-up structures. See Yeo et al. (2004) for further descriptions of the OWC and data sources and processing. The land-use, soil, and topography structures of the catchment are illustrated in Fig. 1.

The hydrological model is a single-event distributed system. The simulation is run with 1-year storm (2.88 mm) event, determined using daily precipitation data available on site (OWC watershed) and the Extreme Value Type I probability distribution function (Chow et al. 1989). Since the data are only available in daily steps, it is assumed that the precipitation pattern follows a SCS II rainfall time distribution (USDA, 1986). The Bulletin 17B method (IACWD, 1982) was used to calibrate the simulation model with 1-, 2-, 5- and 10-year flood frequency analysis at the 95% confidence level. See Yeo et al. (2004) for further discussions on model implementation and validation.

### 3.2 Statistical assessment of global optimality

Five hundred land-use maps have been generated by randomly assigning land-use types to the 1567 cells of this catchment, that are neither road nor water. The total amounts of land use are kept constant across these maps: 22 urban cells, 1307 agricultural cells, and 237 conservation cells. These totals correspond to the optimal land allocation obtained with the hierarchical optimization model developed by Yeo et al. (2007). Starting with population and other forecasts for the whole watershed and the year 2020 (see Yeo and Guldmann, 2006), this model allocates land use first at the subwatershed, and next at the catchment level, using quadratic programming. The resulting aggregate land allocation for this catchment is different from the current one,
with 2.7% urban, 75.5% agricultural, and 13.7% conservation land. The IHLUO model is then applied to these 500 land use allocations at the 30-m cell level, and the resulting optimal allocations that minimize peak stormwater runoff at the catchment outlet are further analyzed statistically. Nine identical local solutions were eliminated in order to satisfy the assumptions of the Fisher-Tippett theorem, which requires independence of the observations in the sample (Los and Lardinois, 1982; Dergis 1985).

In order to illustrate the range of the 491 initial solutions, the catchment is divided into three sub-regions, as illustrated in Fig. 2, and statistics for the numbers of agricultural, conservation, and urban cells allocated to each sub-region are reported in Table 1. These allocations vary significantly within each sub-region, pointing to a wide range of initial solutions. This range is mirrored by the range of the corresponding peak discharge rates, which vary from 0.25 m$^3$/s to 0.5 m$^3$/s (Fig. 3A). In contrast, the optimal allocations generated by the IHLUO model display little variability, with much smaller standard deviations and coefficients of variations (Table 1). The corresponding peak discharges vary within the very narrow range of [0.254073–0.254298] m$^3$/s (Fig. 3B).

Figure 4 displays five maps corresponding to the optimal allocations for the 1-, 25-, 50-, 75-, and 100 percentiles of the peak runoff flow, and the five initial land allocations leading to these optimal allocations. The optimal maps are very close to each other, but significantly different from the initial allocations used to generate them. At the optimum, most of the urban land is allocated to upland areas, near the upland boundary of the catchment, away from the waterways and roads, and at low density. Urban land is buffered by conservation land, in order to offset its impacts on runoff volume and traveling time to the stream. Denser conservation land is allocated near the catchment outlet along the waterways, and at the edge of the catchment, where the slope is steep, but is avoided in areas with low infiltration capacity (i.e., soil type C or D), increasing the traveling time of runoff flows, but reducing the runoff volume.

These optimal peak runoff values are obtained at convergence, that is, when the sum of the squared differences between the land allocations of two consecutive iterations in
the optimization procedure $\left( (\Delta X_k^T \Delta X_k) \right)$ is less than $\varepsilon = 10^{-8}$. It is very likely that
with a much smaller convergence criterion $\varepsilon$, the algorithm would converge to the same optimum for all the initial solutions. However, because the IHLUO model is gradient-based, the convergence becomes very slow after about 10 iterations, and reaching the exact global optimum is unfeasible timewise. Note that if Newton’s method, based on the Hessian matrix, had been used, convergence would likely have been much quicker, but the numerical estimation of the second derivatives of $f(X)$ is not feasible, as discussed in Sect. 2.2.

However, the statistical methodology presented in Sect. 2.3 can be applied to the 491 “optimal” peak discharge rates, which can be viewed as representing a random sample for $\varepsilon$-level convergence, and a three-parameter Weibull distribution is estimated, using a Maximum Likelihood Estimator (MLE). The results are presented in Table 2. The distribution of the sampled data is presented in Fig. 5.

The global optimum and its confidence interval (CI) are estimated from the $\varepsilon$-level optimal values, and the results are summarized in Table 3. The best value from the numerical experiment (i.e., $x_h^{(1)}$ or the upper bound of the CI) is 0.254073, approximately 0.01% above the point estimate of the global optimum ($\hat{x}^* = 0.254047$). The lower bound of the CI is 0.254023, about 0.02% below the best local optimum value ($x_h^{(1)}$). These results make a strong case that the best obtained solution can be considered the global optimal solution.

### 3.3 Explaining global optimality

If the local minimum generated by a hill-climbing algorithm always turns out to be the global one, the objective function is necessarily convex. The previous results therefore suggest that the peak runoff behaves like a convex function with regard to the land-use variables. However, as this function cannot be expressed analytically in a closed form, its convexity cannot be proven through an analysis of its Hessian matrix. As
discussed earlier in Sect. 2, the main components of the hydrological simulation, the calculations of the total runoff volume and traveling times, are involved in estimating the peak discharge. How the land-use variables affect these two components and the whole hydrological system is further examined in the following sections to provide additional support for the convexity of the objective function, though no formal proof.

3.3.1 Estimation of the runoff volume

As described in Eq. (1), the CN method is used to estimate the volume of runoff \( Q \) as a function of precipitation \( P \) and the moisture retention \( S \). While precipitation is exogenous to the simulation model, the parameter \( S \) is solely a function of the curve number (Eq. 2), which is endogenous, as it depends upon land cover and soil type. Let \( x_{i/l} \) be land use \( l \) in cell \( i \), and \( c_l \) the curve number for land use \( l \). Since soil types do not vary across the watershed, the curve number \( (cn_i) \) for cell \( i \) is:

\[
 cn_i = \sum_l c_l x_{i/l} \tag{16}
\]

where \( \sum_l x_{i/l} = 1 \).

Therefore, the parameter \( S_i \) and the runoff volume \( Q_i \) of cell \( i \) are functions of the vector of the land-use variables, \( X_i = (x_{i/1}, \ldots x_{i/l}, \ldots x_{i/L}) \), with:

\[
 S_i = F(X_i) \tag{17}
\]

\[
 Q_i = G(X_i) \tag{18}
\]
The runoff volume $Q_{pa}$ along the flow path $pa$ to the watershed outlet is estimated by summing the runoff volumes in all cells $i$ in the path, with:

$$Q_{pa} = \sum_{i \in pa} Q_i = \sum_{i \in pa} G(X_i) = \sum_{i \in pa} \left[ \frac{(P_i - 0.2F(X_i))^2}{P_i + 0.8F(X_i)} \right] = \sum_{i \in pa} \left[ \frac{P_i + 0.2 \left( \frac{100}{\sum c_i x_{ii}} - 1 \right)}{P_i + 0.8 \left( \frac{100}{\sum c_i x_{ii}} - 1 \right)} \right]^2. \quad (19)$$

$P_i$ is the amount of precipitation in cell $i$. The routing path is determined by the D-8 method, which allows the flow to move to the lowest point among the neighboring cells. Equation (20) is identical to Eqs. (1–2), but applies to the runoff volume along the flow path, instead of to a single cell. Although the total runoff is expressed analytically as a function of the land-use variables, it is difficult to characterize the convexity of the function presented in Eq. (20), because of the term $F(X_i)$. A numerical analysis has been conducted to better understand this function. With the given land allocation and soil distribution (Fig. 1), the CN values ($\sum c_i x_{ii}$) vary in the range of [77–82] for the study catchment. The runoff volume ($Q_{pa}$) generated for these CN values was computed. The results, presented in Fig. 6, show a monotonously increasing and slightly convex relationship between $Q_{pa}$ and CN. As CN is a linear function of the $X_i$’s, $Q_{pa}$ is a convex function of the $X_i$’s.

### 3.3.2 Estimation of the runoff travel time

The travel time of the stormwater runoff is also estimated by summing up the flow time of each cell along a path. As discussed in Sect. 2.1., the most influential parameters for flow time are land uses and topography. The overland flow is a function of Manning’s roughness coefficient, flow length, and slope, the shallowly concentrated flow is determined by slope, and the channel flow is computed with Manning’s roughness coefficient, channel length and area, and slope. Manning’s roughness coefficient, a pa-
rameter for surface friction and resistance, is a function of land-cover, and topography determines the flow direction, slopes, and drainage patterns.

Rather than estimating the time of concentration from the geographically most distant location, the hydrological model keeps track of all flow paths to the watershed outlet, and assigns a specific flow type (i.e., sheet flow, shallowly concentrated flow, and open channel flow) to each cell on each path. This is necessary to account for the impacts of site-specific land-use changes, as surface cover affects the Manning’s roughness coefficient used in flow time estimation. Then, the flow time is explicitly calculated for each cell $i$, and the total flow time over path $p_a$ is estimated by the sum of the travel times over all the consecutive flow segments along the flow path. Then, the time of concentration $T_c$ is determined by choosing the flow path with the maximum travel time. Since flow path and type are fixed by watershed topography and geography, according to the D-8 method and the TR-55 method (USDA, 1986), the time component of the hydrological model is necessarily a function of the land-use variables vector $X$:

$$T_c = H(X)$$

(20)

### 3.3.3 Estimation of the peak discharge

The peak discharge rate at the watershed outlet (Eq. 4) is estimated using the extended TR-55 method (Bingner and Theurer, 2002), which requires the following inputs: the runoff volume, the time of concentration, and the unit peak regression coefficients, $a_{-f}$. These coefficients are determined by the rainfall distribution and the ratio $I_a/P_{24}$. The initial abstraction ($I_{a,i}$) of cell $i$ is estimated as 20% of the moisture retention $S_i$ (Eq. 2), which is itself dependent upon the land uses in cell $i$ (USDA, 1986), with:

$$I_{a,i} = 0.2 \cdot S_i = 0.2 \left( \frac{1000}{\sum c_i X_{il}} - 10 \right) = y(X_i).$$

(21)
The initial abstraction for the watershed is then estimated as the average value of $I_{a,i}$:

$$I_{a}/P_{24} = \sum_{i} I_{a,i}/P_{24} = \sum_{i} y(X_i)/P_{24} = Y(X). \quad (22)$$

The constants ($a$–$f$) are derived from a look-up table, and Fig. 7 presents the changes in their values as function of the ratio $I_{a}/P_{24}$, for Type II rainfall. Except for coefficient $a$, these curves are strongly nonlinear. Once the values of parameters $a$–$f$ are determined, the peak discharge rate is computed by:

$$Q_p = 2.5 \cdot 10^{-2} P_{24} D_a \cdot \left[ \frac{a_p + (c_p \cdot T_c) + (e_p \cdot T_c^2)}{1 + (b_p \cdot T_c) + (d_p \cdot T_c^2) + (f_p \cdot T_c^3)} \right] = 2.5 \cdot 10^{-2} \sum_{1}^{N_{Pa}} Q_{Pa} L(H(X), Y(X)) \quad (23)$$

where $Y(X)$ represents the ratio $I_{a}/P_{24}$ that determines the regression coefficients $a$–$f$ (Eq. 23), $H(X)$ represents the time of concentration $T_c$ (Eq. 22), and $N_{Pa}$ is the number of all possible paths. The right-hand side of Eq. (24) is essentially identical to Eq. (4), as the total runoff at the watershed outlet ($\sum_{1}^{N_{Pa}} Q_{Pa}$) is equal to the product of the total drainage area by the effective rainfall ($P_{24} D_a$), which is the amount of precipitation that is neither retained by the land surface nor infiltrated into the soils. Equation (23) relates two components, the total runoff volume and the time of concentration, to the peak runoff. As the runoff volume ($\sum_{1}^{N_{Pa}} Q_{Pa}$) has been shown to be convex (Sect. 3.3.1), Eq. (24) is further analyzed by focusing on the component, $L(H(X), Y(X))$.

The function $L(H(X), Y(X))$ is computed with selected values for $T_c$ and the $I_{a}/P_{24}$ ratio for the study area. Numerical results are presented in Fig. 8 for $I_{a}/P_{24} = (0.00, 0.25, 0.50, 0.75)$ and $T_c$ in the range [0–2 h], which covers all possible $T_c$ values in the 500 initial land-use patterns. The relationships presented in Fig. 8 are either convex or linear. However, the convexity of Eq. (24) cannot be guaranteed, as the multiplication
of two convex functions, $L(H(X), Y(X))$ and $\sum_{1}^{NPa} Q_{Pa}$, cannot be mathematically proven to be convex.

4 Conclusions

This paper has investigated and characterized the relationship between land-use patterns and watershed hydrology. The IHLUO model has been used to delineate optimal land patterns that minimize the peak discharge rate at the watershed outlet. As the relationship between land use and peak discharge is highly nonlinear, the global optimality of a local optimum cannot be guaranteed a priori. A large number of solutions has been generated from clearly different initial solutions, and these solutions turned out to be very close to each other, strongly supporting the case for a convex relationship between peak discharge and land-use pattern. The obtained solutions can be viewed as the solutions achieved with a given convergence criterion, and the Weibull distribution has been used to generate a point estimate of the global optimum and its confidence interval.

The convexity of the objective function has been further investigated by examining the underlying physics of the hydrological model in terms of land-use variables and by performing numerical evaluations of its main components. The presented mathematical arrangements of conceptual hydrological model provide a unique way to assess the impacts of land use changes across the multiple hydrologic processes of watershed system.

The numerical results strongly support, though do not fully prove, the case for convexity. This finding allows the application of the IHLUO model to much larger watersheds, with many more decision variables, assuring that the obtained solution is reliably the global optimum. It should be noted, however, that the conclusion on the convexity of the land-water relationship may only be valid for the study area, as geographical
characteristics (e.g., topography, soil type) can affect the structure of this relationship. Despite this limitation, this research has proposed an innovative approach to characterize and understand the spatial relationship between land-use pattern and peak discharge. The proposed method can be adapted into other optimization integrated models to develop watershed BMPs and used to assess the closeness of the obtained solution to the global optimum.

References

Relationship between peak runoff and land-use pattern

I.-Y. Yeo and J.-M. Guldmann

Olivera, F.: Spatially distributed modeling of storm runoff and non-point source pollution using geographic information systems, Ph.D. Dissertation, University of Texas, Austin, 1996.
Quilbé, R., Rousseau, A. N., Moquet, J.-S., Savary, S., Ricard, S., and Garbouj, M. S.: Hydrological responses of a watershed to historical land use evolution and future land use
scenarios under climate change conditions, Hydrol. Earth Syst. Sci., 12, 101–110, 2008, 
http://www.hydrol-earth-syst-sci.net/12/101/2008/.
Seppelt, R. and Voinov, A.: Optimization Methodology for Land Use Patterns Using Spatially 
US Department of Agriculture (USDA): Urban Hydrology for Small Watersheds: TR-55, USDA, 
Wagner, H. M.: Principles of operations research: with applications to managerial decisions, 
Williams, J. R., Jones, C. A., and Dyke, P. T.: A Modeling Approach to Determining the Relation-
Yeo, I., Gordon, S. I., and Guldmann, J.-M.: Optimizing Patterns of Land Use to Reduce Peak 
Runoff Flow and Nonpoint Source Pollution with an Integrated Hydrological and Land Use, 
Table 1. Summary statistics for the initial and optimal allocations (unit: 30 m cell).

<table>
<thead>
<tr>
<th>Statistical measure</th>
<th>Land use</th>
<th>Sub-Region 1</th>
<th>Sub-Region 2</th>
<th>Sub-Region 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Initial allocation</td>
<td>Optimal allocation</td>
<td>Initial allocation</td>
</tr>
<tr>
<td>Mean</td>
<td>Urban</td>
<td>5.152</td>
<td>4.192</td>
<td>7.435</td>
</tr>
<tr>
<td></td>
<td>Agriculture</td>
<td>298.635</td>
<td>294.271</td>
<td>454.994</td>
</tr>
<tr>
<td></td>
<td>Conservation</td>
<td>51.559</td>
<td>54.646</td>
<td>83.571</td>
</tr>
<tr>
<td>Median</td>
<td>Urban</td>
<td>5.000</td>
<td>4.197</td>
<td>7.000</td>
</tr>
<tr>
<td></td>
<td>Agriculture</td>
<td>298.000</td>
<td>294.225</td>
<td>455.000</td>
</tr>
<tr>
<td></td>
<td>Conservation</td>
<td>52.000</td>
<td>54.719</td>
<td>84.000</td>
</tr>
<tr>
<td>Minimum</td>
<td>Urban</td>
<td>0.000</td>
<td>3.217</td>
<td>2.000</td>
</tr>
<tr>
<td></td>
<td>Agriculture</td>
<td>282.000</td>
<td>292.932</td>
<td>433.000</td>
</tr>
<tr>
<td></td>
<td>Conservation</td>
<td>37.000</td>
<td>53.122</td>
<td>59.000</td>
</tr>
<tr>
<td>Maximum</td>
<td>Urban</td>
<td>11.000</td>
<td>5.010</td>
<td>14.000</td>
</tr>
<tr>
<td></td>
<td>Agriculture</td>
<td>315.000</td>
<td>295.935</td>
<td>481.000</td>
</tr>
<tr>
<td></td>
<td>Conservation</td>
<td>68.000</td>
<td>55.852</td>
<td>102.000</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>Urban</td>
<td>2.037</td>
<td>0.307</td>
<td>2.359</td>
</tr>
<tr>
<td></td>
<td>Agriculture</td>
<td>6.049</td>
<td>0.569</td>
<td>7.151</td>
</tr>
<tr>
<td></td>
<td>Conservation</td>
<td>5.580</td>
<td>0.514</td>
<td>6.809</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>Urban</td>
<td>0.395</td>
<td>0.073</td>
<td>0.317</td>
</tr>
<tr>
<td></td>
<td>Agriculture</td>
<td>0.020</td>
<td>0.002</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>Conservation</td>
<td>0.108</td>
<td>0.009</td>
<td>0.082</td>
</tr>
</tbody>
</table>
Table 2. Data fitting with a Weibull distribution.

<table>
<thead>
<tr>
<th>Parameter estimation</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location (Threshold): a</td>
<td>0.254047</td>
</tr>
<tr>
<td>Scale: b</td>
<td>0.000156</td>
</tr>
<tr>
<td>Shape: c</td>
<td>4.448359</td>
</tr>
<tr>
<td>Mean of sample data</td>
<td>0.25419</td>
</tr>
<tr>
<td>Standard deviation of sample data</td>
<td>0.000034</td>
</tr>
<tr>
<td>Minimum of sample data</td>
<td>0.254073</td>
</tr>
<tr>
<td>Maximum of sample data</td>
<td>0.254298</td>
</tr>
</tbody>
</table>

Goodness-of-Fit Tests for Weibull Distribution

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 0.356</td>
<td>Pr&gt;W-Sq 0.084</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 2.593</td>
<td>Pr&gt;A-Sq 0.036</td>
</tr>
<tr>
<td>Chi-Square</td>
<td>Chi-Sq 42.572 (d.f. = 10)</td>
<td>Pr&gt;Chi-Sq &lt;0.001</td>
</tr>
</tbody>
</table>

Pr > 0.05 indicates a good fit.
**Table 3.** Estimation for the global optimum and confidence interval.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{x}^*$ = 0.254047</td>
<td>R=491</td>
</tr>
<tr>
<td>$x_{h(1)}^h$ = 0.254073</td>
<td>$\alpha = 0.05$ (95% CI)</td>
</tr>
<tr>
<td>$x_{h(n)}^h$ = 0.254298</td>
<td>S=3.1466</td>
</tr>
<tr>
<td></td>
<td>95% CI of $\hat{x}^*$ = (0.254023, 0.254073)</td>
</tr>
</tbody>
</table>

R: Sample Size without replication; $\hat{x}^*$: global optimum
Fig. 1. Characteristics of the study area.
Fig. 2. Sub-regions in the OWC catchment.

Sub-Region 1
Total no. of cells: 356
Variable+ no. of cells: 356

Sub-Region 2
Total no. of cells: 616
Variable+ no. of cells: 546

Sub-Region 3
Total no. of cells: 760
Variable+ no. of cells: 668
+ Non-water and non-road

Relationship between peak runoff and land-use pattern

I.-Y. Yeo and J.-M. Guldmann
Fig. 3. Distributions of the peak discharge rates.
Fig. 4. Initial and optimal land use maps. Note: The optimal land maps (4B) are generated using three land use categories for illustration purpose only: urban (Urb), Agriculture (Ag), and Conservation (Con). Only the dominant land use category is coded for each cell, except in the case where two land categories are in the same density group. Three density groups are used: Low=0–30%, Med (medium)=30–60%, High=60–100%.
Fig. 5. Observation vs. fitted cumulative density function.
Fig. 6. Numerical assessment of peak discharge rate with varying curve number.
Relationship between peak runoff and land-use pattern

I.-Y. Yeo and J.-M. Guldmann

Fig. 7. Regression coefficients for peak discharge (Bingner and Theurer, 2002).
Fig. 8. Numerical assessment of $L(H(X), Y(X))$ under type II rainfall distribution.