Prediction of land-use conversions for use in watershed-scale hydrological modeling: a Canadian case study

Antoun El-Khoury  
Department of Civil Engineering, University of Ottawa

Ousmane Seidou  
Department of Civil Engineering, University of Ottawa

David R. Lapen  
Agriculture and Agri-food Canada

Mark Sunohara  
Agriculture and Agri-food Canada

Que Zhenyang  
Department of Civil Engineering, University of Ottawa

Majid Mohammadian  
Department of Civil Engineering, University of Ottawa

Bahram Daneshfar  
Agriculture and Agri-food Canada

Land-use conversion models elucidate the complexities and spatial interdependencies of components of land use systems and provide insights into future land-use configurations. In this paper, the 2012–2050 future land-use patterns in the South Nation (SN) River basin, located in eastern Ontario, Canada, were generated with a modified version of the CLUE model and a 2011 reference map. The SN is an example of a basin where some water quality endpoints have dropped below acceptable limits because of a combination of intensive agriculture, urbanization, and climate change. Five historical land-use maps were used to identify the historical trends in generalized land-use classes. Seven demographic and geographic factors were used to derive the spatial distribution of land suitability to each land-use class. The methodology was first validated by simulating land-use changes from 1991 to 2011 starting from the 1991 reference map, and comparing the simulated 2011 map to the 2011 reference map. Then, the 2012–2050 land-uses were generated, assuming historical trends derived from historical reference maps will continue in the future. Environmental impacts of the projected land-use changes were discussed.

Keywords: Land-use projection, CLUE, environmental degradation

Prévision de la réaffectation des sols pour la modélisation hydrologique des bassins versants: une étude de cas canadienne

Les modèles de réaffectation des sols permettent de décortiquer les complexités et les interdépendances spatiales des éléments de base des systèmes d'occupation des sols et donnent un aperçu des formes

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

*Correspondence to/Adresse de correspondance: Ousmane Seidou, Department of Civil Engineering, University of Ottawa, 161 Louis Pasteur Office A113. Email/Courriel: oseidou@uottawa.ca
The generation of future land–use maps of the basin; (2) the development of climate change scenarios for the basin; and (3) running a spatially distributed hydrological model to assess combined impact of land–use and climate changes on water quantity and water quality (e.g., nitrogen, phosphorus loads). Only the generation of future land–use maps is presented here; the remainder of the study can be found in El-Khoury et al. (2014). The generalized land–use conversion model used in the El-Khoury et al. (2014) study was the CLUE model (Verburg et al. 2002; Verburg and Overmars 2009). The CLUE model was modified here, in order for CLUE outputs to be more efficiently utilized within the hydrological model. A modification to CLUE land–use allocation algorithms was made here to ensure a quicker convergence time. The modified CLUE model was used in this paper to generate future land–use configurations using historical trends and spatial suitability equations. Historical land–use maps of the SN River basin corresponding to various dates (1966, 1991, 2000, 2005, 2011) were used as observed data to inform land–use change trends. A number of potential drivers were derived from various sources (Pavlic et al. 2002; Statistics Canada 2006a, b), and then used to derive land–use suitability equations (i.e., equations that will link the drivers to the probability of occurrence of a given land–use in a given pixel). These equations and the trends estimated from historical reference maps were subsequently used with the land–use allocation algorithm to generate the 2011 map starting from the 1991 reference map, and the 2050 map starting from the 2011 reference map. To validate the methodology, the simulated 2011 map was compared to the 2011
Materials and Methods

Study area
The SN River basin is located in eastern Ontario, Canada (Figure 1). The SN basin covers an area of 3900 km² with an estimated population of approximately 115,000 inhabitants distributed in 15 municipalities. Lumbering and agriculture led to widespread deforestation beginning in the nineteenth century, where wetland areas dropped from ~47% to, in early 1980s, ~16% (Coyne 2001). Analyses of satellite-derived land-use maps suggest that cropland areas were on the decline between 1966 and 2000, but started to increase again after 2000. Unsurprisingly, water quality in the basin has declined during the last decades. Total nitrogen loading in surface water has increased significantly since the 1960s (Fleming and Fraser 1999); total phosphorus concentrations regularly exceed 0.03 mg/L which is the provincial water quality objective (South Nation Conservation Authority 2006). Urbanization and agricultural activities have contributed to enhanced contaminant loading in the basin.

The CLUE land-use projection process overview
The following data were gathered or generated to inform CLUE modeling:

1. The target surface for each land-use class at each time step in the simulation period. These target surfaces are called land-use demands. In this paper, the actual land-use surfaces in the reference maps were used in the validation exercise; the projected land-use surfaces were used as land-use demands for future periods.
2. Location suitability maps that represent the competitive capacity among land–use classes at a specific location. By location suitability we mean the probability that a specific land–use will occur at that location. The calculation of suitability maps was based on empirical analysis of preferences between land–use classes and potential driving factors. In the present study, two binary logistic regression equations were used to link the probability of occurrence of each land–use class to: (a) geographic and demographic factors; and (b) the type and dominance of all available land–use classes in a spatial neighbourhood.

3. Spatial policies and restrictions that represent areas where a particular type of land–use conversion cannot take place. Local policies (e.g., zoning, political boundaries, biophysical boundaries) can restrict a certain land–use at a given location even if it has a higher suitability than other land–uses.

4. Land–use specific conversion settings for all land–use classes in the study area. The conversion settings are implemented into the model in two forms: a matrix of conversions and a set of elasticity coefficients representing the resistance of a given land–use to change. The use of elasticity coefficients is explained later in the text.

In the CLUE algorithm as described in Verburg et al. (2002), once all the above information is available, land–use patterns at each step are iteratively allocated by varying the elasticity of each land–use and allocating to a given pixel, the land–use having the higher sum of location preference and elasticity. Iterations are stopped when the allocated land–use surfaces are reasonably close to the demands. There is, however, no guarantee that there is a combination of elasticity values that will result in a reasonable solution. A systematic search in the elasticity space proved to be time consuming (a problem associated with $k$ different land–uses, and attempting $n$ values for each elasticity value, leading to $k^n$ map allocation trials) and often did not give a solution within the desired precision. For some cases, increasing the value of $n$ to ten–fold results in over 100,000 trials, which did not lead to an acceptable solution. A modification to the algorithm (described later in the text) solved the convergence issue while keeping the general principles of the CLUE model valid.

Land–use observations and estimation of historical trends

Projection of land–use demand up to 2050 will be driven by complex socioeconomic processes occurring inside and outside the study area that are only partially described by existing land–use trends. In this paper, time series of historical land–use maps were built from: (1) the 1966 partial map of the area (Natural Resources Canada 2002); (2) the 1991 land–use map (AAFC 1991); (3) a 2000 land–use map (CH2M HILL 2001); and (4) a 2005 and 2011 map (AAFC 2005, 2011). The 1966 and the 2011 land–use maps are given in Figure 2.

The maps have different geographical extents so the analysis was restricted to their spatial collocation. They also use different, and possibly inconsistent, classifications, which is not surprising when data from different sources are utilized. A reclassification was conducted by reassigning original land–use class designations to one of the following generalized six classes: bare, urban, cropland, other, forest, and water. The correspondence between the original classes and the new classes for the 1966 and 2011 maps is provided in Table 1. Such reclassification introduced an additional degree of uncertainty in the reference map. The area of each class was then calculated inside the common area, multiplied by the ratio of total river basin area to the common area. The time series of land–use surface were used for trend calculations as follow:

**Urban areas**: exponential and linear trends were estimated using the 1966–2011 time series. Based on the sum of squared errors, it was found that the exponential trend (plotted on Figure 3b) best represented the evolution on urban areas.

**Cropland areas**: extracted cropland surfaces suggest that cropland areas have been continuously decreasing since 1966 but that the decrease has stopped and even was inverted since 2000. Generalized increases in crop prices support this finding. A linear trend was estimated for cropland areas using the 2000–2011 values, and that trend was used to extend the time series up to 2050 assuming that historical trends will continue in the future. Historical cropland areas between 1966 and 2011 and projected values between 2011 and 2050 are presented on Figure 3b. The interception of the projection equation was set to have continuity in land–use surfaces for year 2011.
Forested areas: there was no need to develop a trend line for this land-use as it can be deduced from the variations of urban and cropland areas (as a residual).

After the trends were calculated, all maps were resampled to a 30 m resolution and the calculations were carried in a rectangular region of 5554 (north-south) by 6619 (east-west) cells containing the river basin.

Location suitability equations
Land-use changes are driven by a large array of factors such as local economic activities, physiographic suitability for agricultural activity, proximity to water or transportation infrastructure, political boundaries, etc. For a given land-use modeling project, some of these factors are readily available while others are not. The modeller has to choose among known factors, a subset of factors that he/she believes are likely to have a significant impact on land-use conversions (expert knowledge). In this paper, we restricted our search for drivers to those that could be calculated from geographical data sets that were readily available at the beginning of the study. These data sets include the 2005 road infrastructure map (Statistics Canada 2006a); the municipal population density calculated from the 2006 census of population (Statistics Canada 2006b); and the spatial distribution of water bodies in the watershed (Pavlic et al. 2002). Several drivers...
were extracted from these maps: one demographic factor (population density) and six geographic factors, which include the distance to the closest city; the distance to the closest major city (population > 50,000); the distance to the closest small city (population < 50,000); the distance to the closest major road (primary highway or expressway); the distance to the closest paved road; and the distance to the closest water body.

The CLUE model uses logistic regression to estimate the suitability of a given pixel in a given simulation space to a given land-use $i$, which is measured by a probability $P_i$ linked to biogeophysical and socioeconomic parameters by the following regression model:

$$
\log \left( \frac{P_{loc,i}}{1 - P_{loc,i}} \right) = \sum_{j=0}^{n} \beta_{loc,j}X_{j,i}
$$

where $X_{j,i}$ is the $j^{th}$ location affecting the suitability of land-use $i$, $\beta_{loc,j}$ is a coefficient to be estimated, and $P_{loc,i}$ is the location suitability based on location factors. Since maps of biogeophysical and socioeconomic parameters are available, equation one can be used to plot a map of the suitability of each point in the study region for a given land-use. When allocating an amount $X$ of land-use $i$, the model will give priority to pixels with high suitability values. Given that the probability of occurrence of a given land-use depending on its neighbourhood, a special group of location parameters $F$ (also called enrichment factor) are calculated as follows:

$$
\log \left( \frac{P_{enr,i}}{1 - P_{enr,i}} \right) = \sum_{d \in D} \left( \beta_{enr,k,d}F_{i,k,d} \right)
$$

where $n_{k,d,i}$ is the number of cells of land-use of type $k$ in the neighbourhood with size $d$ of cell $i$, $n_{d,i}$ is the number of cells in the neighbourhood, $N_k$ is the number of cells of land-use of type $k$ in the map, and $N$ is the number of cells in map. In this paper enrichment factors $F_{i,k,d}$ were calculated for $d=100$ m, 500m, 1000 m, 5000 m and 10000 m. A second equation is derived as follow:

Table 1
Correspondence between the original and the new classifications for the 1966 and 2011 maps

| Original class                    | Hydrological model class | Original class                    | Hydrological model class |
|----------------------------------|--------------------------|----------------------------------|--------------------------|
| Productive woodland              | Forest                   | Cloud                            | Other                    |
| Improved pasture and forage crops| Cropland                 | Water                            | Water                    |
| Nonproductive woodland           | Forest                   | Exposed land                     | Bare                     |
| Swamp, marsh or bog              | Water                    | Developed                         | Urban                    |
| Unimproved pasture and range land| Cropland                 | Shrubland                         | Forest                   |
| Mines, quarries, sand and gravel pits | Bare                      | Wetland                           | Forest                   |
| Unproductive land rock           | Bare                     | Grassland                         | Cropland                 |
| Outdoor recreation               | Forest                   | Agriculture (undifferentiated)    | Cropland                 |
| Urban built up area              | Urban                    | Hay / Pasture                     | Cropland                 |
| Unproductive land sand           | Bare                     | Too wet to be seeded              | Cropland                 |
| Orchards and vineyards           | Cropland                 | Fallow                            | Cropland                 |
| Horticulture                     | Cropland                 | Cereals                           | Cropland                 |
| Unmapped areas                   | Other                    | Corn                              | Cropland                 |
| Water                            | Water                    | Soybeans                          | Cropland                 |
| Cropland                         | Cropland                 | Beans                             | Cropland                 |
|                                  |                          | Vegetables                        | Cropland                 |
|                                  |                          | Fruits                            | Cropland                 |
|                                  |                          | Herbs                             | Cropland                 |
|                                  |                          | Other crops                       | Cropland                 |
|                                  |                          | Coniferous trees                  | Forest                   |
|                                  |                          | Deciduous trees                   | Forest                   |
|                                  |                          | Mixed trees                       | Forest                   |
where $\beta_{enr,i}$ are coefficients to be estimated, $D$ is an ensemble of selected distances to calculate enrichment factors, and $P_{enr,i}$ is the neighbourhood suitability. Equation (2) is an attempt to capture the fact that new urban developments or new croplands are more likely to occur in an area where there are already pixels with similar land-uses in the neighbourhood. Given the large number of potential predictors, a forwarded stepwise procedure was used in the estimation of the coefficients of the two logistic equations and only predictors with a p-value below 0.05 were retained. (2) and (3) were fitted for all three land-uses and presented respectively in Tables 1 and 2 for urban, cropland, and forest land-uses, for years 1991, 2000, 2005, and 2011. Apart from two exceptions (urban land-use, year 2000 and 2011), all location factors were found significant and retained in the location suitability equation (Table 2). Similarly, each type of land-use was found to significantly influence the suitability of its neighbourhood for urban, cropland, and forest land for at least one neighbourhood size (Table 3). The quality of the fit was measured by plotting the Receiver Operating Characteristic (ROC) for each equation (Figures 4 and 5). ROC curves are commonly used to assess the performance of logistic regression models. In a ROC curve the true positive rate (the percentage of times the model correctly predicted the presence of a given land-use) is plotted in function of the false positive rate (the percentage of time the model predicts a given land-use while the actual land-use is different) for different cut-off points (threshold above which the output of the logistic regression model is considered a positive outcome). Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold (in this case the calculated location preference). A perfect model has a ROC curve that passes through the upper left corner of the graph. Therefore the closer the ROC curve is to the upper left corner of the graph, the higher the overall accuracy of the test (Zweig and Campbell 1993). A model is considered to have skill if the ROC curve is above the diagonal, and the area under the curve is above 0.5. Examination of Figures 4 and 5 show that location suitability equations for urban land-use are well fitted (because of the inclusion of distances to the center of cities in the predictors) while location equations for forest and cropland have lower but positive skill as measured by the area under the curve. As expected, the skills of the neighbourhood suitability equations are high for all three land-uses.

**Land-use allocation**

Once the location preferences for each land-use class are calculated, the following allocation
steps are taken to assign the suitable land-use changes:

1. Determination of all pixels that are allowed to change by excluding those which are parts of protected areas.
2. For each grid cell \( i \), the total probability \( \text{TOTALP}_{i,k} \) is calculated for each of the land-use class \( k \) according to the following equation:

\[
\text{TOTALP}_{i,k} = P_{\text{loc},i,k} + ELAS_k + ITER_k
\]

where \( P_{\text{loc},i,k} \) is the suitability of location \( i \) for land-use class \( k \) based on location factors; \( P_{\text{enr},i,k} \) is the suitability of location \( i \) for land-use class \( k \) based on enrichment factors; \( ELAS_k \) is the conversion elasticity for land-use \( k \); and \( ITER_k \) is an iteration variable that is specific to the land-use type and indicative for the relative competitive strength of the land-use type. \( ELAS_k \) is the land-use type specific elasticity to change value, only added if grid-cell \( i \) is already under land-use type \( k \) in the year considered. In this paper \( ELAS_k \) is set to infinite for water and urban land-uses so that they cannot be converted to another land-use, and zero for all other land-use classes. \( ELAS_k \) is changed by the model at each time step when attempting to fulfill land-use demands.

3. The land-use allocation procedure used in this paper is different from the one described in Verburg et al. (2002). It was found that the procedure described in Verburg et al. (2002) required very long iterations and did not always converge to the desired amounts of land-uses when applied to the study site here. In this paper, land-use changes are calculated in the following order for each year: urban, cropland, and forest. For each of these land-uses (assume it is land-use \( k \)), if the demand for the current year is larger than the area allocated to the area currently allocated to land-use \( k \), \( ITER_k \) is adjusted until the demand becomes equal to the surface occupied by land-use \( k \). This allocation process will cause a certain number of grid cells to change land-use class. Pixels for which the land-use was changed are no longer allowed to change until the next year.

The total allocated area of each land-use is therefore equal to the land-use demand areas presented in the demand file. The ultimate map of the year in question is generated in one loop, much faster than the iterative process used by CLUE in Verburg et al. (2002).

Land-use specific conversion settings were set to forbid the conversion of water and urban to other types of land-use. Bare and undefined land-uses were merged into cropland and forest respectively. Given the small percentage of bare and undefined land in the maps, this choice made the modeling easier without significantly impacting the results. Forest and cropland were allowed to freely change to any other land-use except water. No spatial policies were implemented in the model.
Table 3
Coefficients of the neighbourhood suitability equations for urban, cropland and forest land uses

| Suitability Factor | 1991 | 2000 | 2005 | 2011 | 1991 | 2000 | 2005 | 2011 | 1991 | 2000 | 2005 | 2011 |
|--------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| Constant           | 2.4E-1 | -2.0E-1 | -2.1E+0 | -2.2E-1 | -2.6E+0 | -2.7E+0 | -3.8E+0 | -2.7E+0 | -3.9E-1 | 7.2E-2 | -2.1E+0 | 1.3E-1 |
| Bare (10000 m)     | 1.4E-2 | 4.9E-01 | 7.0E-1 | 3.1E+1 | 2.3E+1 | -1.7E+1 | -5.4E+0 | -1.0E+1 |          |         |         |      |
| Bare (100 m)       | 1.2E+2 | 3.7E-1 | 3.8E-1 | -6.7E-2 | -1.3E+2 | -8.3E+1 | -1.2E+2 | -3.1E+1 | -4.2E-1 |          |         |      |
| Bare (1000 m)      | 6.6E+3 | 5.1E+3 | 2.1E-3 | -1.6E+4 | -4.1E+3 | -4.1E+3 | -5.8E+3 | -4.8E+3 | -4.8E+3 | -1.9E+3 |          |      |
| Bare (5000 m)      | -9.5E+0 | -2.9E+0 | -6.5E+0 |          |          |          |          |          |          |          |          |      |
| Bare (500 m)       | 1.9E-2 | -1.1E+2 | 1.5E+2 | 5.7E+2 | -1.0E+2 | -1.7E+2 | -2.1E+2 | 1.2E+2 | -1.4E+2 |          |         |      |
| Cropland (10000 m) | 4.8E+3 | -4.5E+1 | 2.0E+1 | -1.1E+2 | 2.2E+2 | -1.4E+2 | -3.8E+1 | -1.5E+2 | -5.9E+1 | 3.8E+1 | -2.4E+1 | 7.9E+1 |
| Cropland (1000 m)  | 1.6E+3 | 6.0E+2 | 1.8E+3 | -1.0E+3 | 4.8E+3 | -5.6E+2 | 1.0E+3 | -1.5E+3 | -1.6E+2 | 6.5E+2 |          |      |
| Cropland (100 m)   | 2.2E+5 | 1.6E+5 | 2.8E+5 | 1.2E+5 | -4.9E+5 | -1.5E+5 | -1.2E+4 | -1.1E+5 | -2.0E+5 | -1.5E+5 | -4.1E+4 | -1.2E+5 |
| Cropland (5000 m)  | -2.2E+2 | 5.8E+1 | -9.5E+1 | 2.8E+2 | -1.1E+3 | 4.9E+2 | 2.0E+2 | 2.0E+2 | -8.9E+1 | 3.4E+1 | -1.9E+2 |          |
| Cropland (500 m)   | -9.4E+3 | -4.4E+3 | -1.0E+4 | 2.1E+3 | 1.5E+4 | 4.5E+3 | 1.8E+3 | 8.6E+3 | 4.7E+3 | -1.8E+3 |          |      |
| Forest (10000 m)   | -3.7E+1 | -4.7E+1 | -6.2E+1 |          |          |          |          |          |          |          |          |      |
| Forest (100 m)     | -1.0E+3 |          |          | -1.3E+3 | 1.5E+3 |          |          |          |          |          |          |      |
| Forest (1000 m)    | -7.3E+4 | -1.1E+5 | -1.7E+4 | -5.6E+4 | -2.5E+5 | -1.3E+5 | -1.5E+4 | -9.0E+4 | 8.0E+4 | 1.2E+5 | 1.3E+5 | 6.4E+4 |
| Forest (5000 m)    | 1.0E+2 | -1.8E+1 | 1.8E+2 | -1.1E+3 | 1.5E+2 |          |          |          |          |          |          |      |
| Forest (500 m)     | 1.9E+3 | 4.0E+3 | 1.7E+3 | 4.4E+3 | 2.6E+3 | -2.1E+3 | -4.0E+3 | -4.8E+3 | -1.5E+3 |          |          |      |
| Urban (10000 m)    | 1.3E+1 | 9.2E+0 |          | -2.7E+1 | -3.9E+1 | 5.3E+0 | -5.1E+1 | 2.9E+0 |          |          |          |      |
| Urban (1000 m)     | 4.6E+1 | -4.7E+1 | 1.6E-1 | -4.3E+1 | 2.0E+2 | 1.7E+2 | 2.0E+2 | 2.4E+2 | 2.2E+1 | -1.4E+1 |          | -9.4E+1 |
| Urban (100 m)      | -2.6E+2 | -3.5E+3 | -2.2E+3 | -4.9E+3 | 3.7E+3 | 1.4E+4 | 2.2E+4 | 1.3E+4 | -4.3E+3 | -5.7E+3 | -3.0E+2 | -3.8E+3 |
| Urban (5000 m)     | -1.4E+1 | -1.0E+1 | -2.9E+0 | 1.3E-1 | 2.2E+1 | 6.3E-1 | 8.7E+1 | 2.9E+0 |          |          |          |      |
| Urban (500 m)      | -2.4E+2 | 1.8E+2 |          | 2.4E+2 | -1.1E+3 | -1.1E+3 | -1.0E+3 | 1.5E+2 | 1.7E+2 |          | 1.6E+2 |          |
| Water (10000 m)    | -4.1E+0 |          | -6.0E+0 | 3.3E+0 | 2.2E+0 | 3.1E+0 | -2.6E+0 | -5.4E+0 | 5.0E+0 |          |          |      |
| Water (1000 m)     | 6.7E+0 |          |          | 3.8E+0 | -3.0E+0 | -5.2E+1 | -1.1E+1 | -1.4E+1 | -1.5E+1 |          |          |      |
| Water (100 m)      | -4.4E+2 | -8.2E+2 | -1.3E-3 | -1.1E-3 | -2.4E+3 | -1.3E+3 | 8.3E+2 | -1.1E+3 | -4.4E+2 | -9.1E+2 | -8.6E+2 | -1.5E+2 |
| Water (5000 m)     | 1.6E+0 |          | -4.5E+0 | 3.0E+1 | 6.4E+0 | -3.9E+0 | 6.1E+0 | -2.0E+0 | 8.7E+0 | 2.4E+0 |          |
| Water (500 m)      | 1.3E+1 | 3.9E-1 | -2.8E+1 | 3.9E+1 |          |          |          |          |          |          |          |      |

Validation procedures
Two types of validation procedures were selected to compare the simulated 2011 map to the 2011 reference maps: (1) an intuitive visual validation procedure in which specific conversions in specific areas of the map were compared to expectations (expert opinion), and (2) the multiple resolution validation procedure proposed by Costanza (1989) where the similarity of the two maps at a given spatial resolution is summarized in one number between 0 and 100%. The main idea behind the multiple resolution validation procedure is that comparing the two maps at one resolution is insufficient to describe the similarity of land–use patterns especially if the resolution is very fine, as “near misses” will be given a zero weight. The solution is the following validation index $F_w$ for various window sizes $w$:

$$F_w = \frac{\sum_{s=1}^{w} \left[ 1 - \sum_{k=1}^{p} \frac{|a_{ik} - a_{kj}|}{2w^2} \right]}{t_w}$$

where $w$ is the dimension of the sampling window, $F_w$ is fit for sampling window $w$, $a_{ik}$ the number of cells of land–use class $i$ in map $k$, $p$ the number of land–use classes, and $t_w$ the number of sampling windows in the map. While $F_w$ may be small at the highest resolution (smallest $w$), if $F_w$ increases rapidly with $w$ then stabilizes quickly, it can be concluded that the patterns between the two maps are well matched.

Results and Discussion
In order to validate the land–use allocation algorithm, the 2011 land–use was simulated starting
from 1991, and simulated urban, cropland, and forest lands were compared to observed respective land-use changes, as previously discussed. Unidentified land-uses (0.3% of the 1991 reference map) were merged in the forest land-use, while bare lands (10% of the 2011 reference map) were assimilated to cropland. The observed reference and simulated maps for 2011 (Figure 6) were used in the two validation procedures.

Visual validation
Since the observed and simulated land-use changes were not dramatic, differences between the two maps was difficult to assess visually. It was more practical to generate a map of each class of land-use showing the type of conversion that took place. Figures, 7, 8, and 9 show maps of undisturbed areas (yellow), lost areas (green), and new areas (blue) for urban, forest, and cropland and forest, respectively. A few unexpected green areas can be found on the top panel of Figure 7 showing that areas which were classified as urban in the 1991 reference maps are no longer classified as such in the 2011 reference map, while urban areas are unlikely to be replaced by other land-uses in the future. These green areas are probably due to discrepancies in the classification methods between the 1991 data (AAFC 1991) and the 2011 data (AAFC 2011). The simulated map of 2011 shows no loss of urban areas since the conversion restriction described in the previous section forbids it. More "new urban" areas are visible on the reference map to compensate for lost urban areas on the same map. The new urban areas are uniformly distributed in the top 80% of the reference map, and the top 60% of the simulated map. These results suggest that the northern portion of the river basin (closest to the City of Ottawa) will be developing earlier than the southern parts of the watershed. Figure 8 shows very similar patterns of forest conversions between the simulated map and the reference map. Most of the undisturbed forest is...
concentrated in two regions north and south of the watershed, while new forested areas are uniformly distributed except in a southwest–northeast oriented strip in the center of the watershed where their density is lower. Clusters of new forested areas are visible on the reference map but not on the simulated map. Yet regulatory restrictions regarding foresting, and reforesting, are highly dynamic and moreover, currently there is considerable deforestation for the purpose of increasing arable lands for agriculture in the region. In the US Corn Belt, a recent doubling in commodity prices has created incentives for landowners to convert natural lands to corn and soybean cropping systems (Wright and Wimberly 2013), and the trend to clear land for agriculture is notable in many parts of Canada as well, including the SN River basin (e.g., AgriNews 2010). Thus it will be unclear how this deforestation trend will be contained by regulatory action responding to these trends.

Like forested areas, similar conclusions from the simulations can be drawn for cropland areas (Figure 9), where lost cropland are uniformly distributed except in a southwest–northeast oriented strip in the center of the watershed where their density is lower. Overall, given the level of uncertainty in the reference maps, the performance of the land-use allocation model seemed satisfactory.

Multiple resolution validation

In order to confirm the added value of the land-use generation algorithm, the previously described validation index was used to compare the 2011 simulated map with both the 2011 and 1991 reference maps, at resolutions of 30 m, 90 m, 300 m, 900 m, 3000 m, 6000 m, 15000 m, and 30000 m (Figure 10). The plot of the validation index shows that above a resolution of 2000 m, the 2011 simulated map has more similarity with the 2011 reference map than with the 1991 reference map (97% versus 90% accuracy at a resolution of 3000 m). It is closer to the 1991 reference map at resolutions

Figure 5
ROC curves of the suitability equations based on enrichment factors
lower than 2000 m because of the relatively large number of pixels which were never converted. These results show that the land-use allocation algorithm captures the land-use dynamics in the watershed at and above a 2000 m scale.

Simulated 2012–2050 land-uses
Starting with the 2011 land-use map and using the 2012–2050 cropland and urban land-use demands, land-use patterns were simulated up to 2050.

Undefined land-uses (0.3% of the 2011 reference map) were merged with forested areas while bare lands (2.3% of the 2011 reference map) were assimilated with cropland. Water area in 1991 is assumed exact and does not vary through simulations. Given that the trend in cropland area is assumed to be inverted starting from 2010, the generated maps show a continuous gain in cropland and a continuous loss in forest between 2012 and 2050. This is consistent with current trends in the river basin due to reasons already discussed.

Figure 6
Simulated and observed 2011 land-use maps

Figure 7
Simulated and observed changes in urban areas between 1991 and 2011
The algorithm was able to allocate each land-use with an accuracy of less than a pixel area (900 m²). The generated 2025 and 2050 maps are presented in Figures 11 and 12. As expected, the concentration of new cropland and lost forest areas follow the same pattern as described in the previous section, but in the inverse direction. The pattern of new urban areas (Figure 12, top right panel) is similar to the pattern of areas that were converted to urban between 1991 and 2011 (Figure 7, top panel), suggesting the model is able to reasonably allocate urban areas if the right demand is assigned for a given year. At the end of the process, a set of 47 land-use maps were generated that serve to fuel input to hydrological simulations, and identify highly impacted areas in the basin that should receive special attention from conservation authorities. There are a few possible improvements that could be considered for the future land-use generation process:

1. The land-use demand projection could be made more sophisticated by including the socioeconomic dynamics of the basin, or at least the

Figure 8
Simulated and observed changes in forest between 1991 and 2011

Figure 9
Simulated and observed changes in cropland areas between 1991 and 2011
development projections of the municipalities in the watershed.

2. The choice of the predictors in the logistic regression was based on a combination of expert opinion on importance, parsimonious principles, and data availability. A greater number of important predictors could have been exploited, but at the potential expense of keeping modeling parsimonious. Example of predictors that could have been used to explain land use changes are crop and lumber prices.

3. The model does not set restrictions on cropland expansion while there are several protected forests on the watershed. Local policies restricting the conversions of these protected forests to cropland in the model could have been implemented.

Despite these limitations, the generated maps were deemed realistic enough to be used for the remainder of the project.

Environmental and policy implications of the projected land-use changes

The modeling work conducted in this paper provides for a reasonable simulation of land-use change trajectories in the basin, assuming no restriction on land-use conversions is in place—a gross simplification given that land conservation authorities regulate activities in areas that are thought to be sensitive, such as drinking water source protection areas, and have forest protection policies in place. The following critical areas of land-use changes can be identified, all of which likely have significant environmental impact:

1. Relatively rapid urbanization in the areas close to the Ottawa/Gatineau metropolitan area in the north, and along highways (Figure 7). Urbanization is usually associated with increased point and non-point source pollutant loads, such as waste water treatment plant effluent, septic leakages, pesticide and fertilizer use, etc. (e.g., Qin et al. 2010; Walters et al. 2011; Wilkes et al. 2013).

2. Disappearance of small forested areas and their replacement with croplands throughout the watershed. On the simulated 2050 map, most forested areas will be concentrated in two spots located in the northern and southern part of the watershed. Forested areas are important to the health of a watershed as they have the following major environmental and protective functions (Gottle and Sène 1997): (a) protection of water resources; (b) soil protection; (c) influence on the local climate and reduction of gas emission impacts; (d) conservation of the natural habitat.
and biological diversity; and (e) recreational and other social functions of forests. All these functions will be significantly affected if the projected land-use changes occur. Furthermore, the increase of croplands areas will inevitably result in higher nutrient loading, and potentially increases in agriculturally derived pathogens (Barton and Farmer 1997; Almasri and Kaluarachchi 2004; Wilkes et al. 2011, 2013).

Given that water quality is already a concern to conservation authorities, the results presented in this paper are a call for strong policies to control land-use conversions or restrict the allowed activities in some of these land-use policies that specifically control propagation of urbanization and agriculture at the expense of forested areas. Surprisingly, according to CCEA (2013), no conservation area is currently located inside the SN river basin. The results presented in this paper suggest that some of the smallest forest areas, especially those close to the water bodies, should be protected. Wilkes et al. (2013) showed that the presence of these forested areas could reduce exposure risks of some zoonotic pathogens such as livestock and human associated Cryptosporidium. The generated 2011–
2050 maps can also assist the South Nation Conservation Authority in the prioritization of interventions aiming to protect vulnerable water sources.

Usage of the generated maps to estimate future evolution of water quantity and quality in the basin

It is broadly recognized that land-use modifications can severely alter both water quantity and quality patterns (Fohrer et al. 2001; Tong and Chen 2002). Water quality is currently a major concern in the watershed because of agricultural activities that generate potentially harmful pollutants that are delivered to the river system through surface and groundwater flow. The findings in this paper suggest that pollution from agricultural lands should increase as cultivated areas expand. Pollution from urbanized areas would increase as well, and as such, germane point sources need to be accounted for dynamically over time. Distributed or semi-
distributed hydrological models that account for land–use may be used with the maps generated in this paper to quantitatively assess how water quality will be affected by the projected changes in land–use. Expected changes in climate regimes may be integrated to have an even better picture of future hydrological changes. In El–Khoury et al. (2014), the maps generated in this paper were used in combination with climate change scenarios to estimate the combined effects of climate and land–use changes on six water quality (N and P) and quantity parameters. The simulations were carried out using the SWAT (Soil and Water Assessment Tool: Neitsch et al. 2009) to simulate the future hydrological regime of the South Nation river. They found that changes in streamflow will be driven by climate, whereas changes in water quality parameters will be driven by land–use changes—underscoring the importance of the coupling of land–use and climate–change modeling in predicting future hydrological parameters at river basin scales.

Conclusions

A modified version of the CLUE model was used to generate the 2012–2050 land–use maps of the SN river basin. Six different land–uses (bare, urban, cropland, other, forest, and water) were considered; binary logistic regression was used to estimate the location and neighbourhood suitability of each location within the watershed. The land–use allocation procedure was successfully validated using both visual assessment and a multiple resolution procedure. Historical trends in each of the six land–use classes were estimated from remote–sensing maps and used to generate land–use maps for the 2012–2050 periods. Results suggest that plausible future land–use changes could increase the water quality burden, if mitigation measures such as agricultural Best Management Practices (BMPs), forest protection, and reforestation, are not implemented.

Acknowledgements

This work was jointly funded by Agriculture and Agri–food Canada and the NSERC Discovery Grant of the main author. We would like to thank the editor and the reviewers for their time and valuable input that helped greatly improve the paper.

References

AAFC (Agriculture and Agri–food Canada). 1991. Regional environmental information system (REIS). (Site accessed 23 October 2008; no longer active.) http://reis.agr.ca/cansis/index.html.

——. 2005. Land–use map of the South Nation watershed. Unpublished data set, Obtained courtesy of AAFC. Ottawa, Ontario.

——. 2011. Land–use map of the South Nations watershed. Unpublished data set, Obtained courtesy of AAFC. Ottawa, Ontario.

AgriNews. 2010. MNR minister attends event highlighting Dundas County deforestation. http://www.agrinewsinteractive.com/archives/article–11217.htm.

Almasri, M. N., and J. J. Kaluarachchi. 2004. Assessment and management of long–term nitrate pollution of ground water in agriculture–dominated watersheds. Journal of Hydrology 295(1–4): 225–245.

Barton, D. R., and M. Farmer. 1997. The effects of conservation tillage practices on benthic invertebrate communities in headwater streams in southwestern ontario, canada. Environmental Pollution 96(2): 207–215.

CCEA. 2013. The conservation areas reporting and tracking system (carts): data current as of 31 March, 2013. http://www.ccea.org/en_carts.html.

CH2M HILL. 2001. Eastern Ontario Water Resources Management Study Final Report. CH2M HILL Canada Limited, Prescott–Russel, Ontario.

Costanza, R. 1989. Model goodness of fit: A multiple resolution procedure. Ecological Modelling 47(3–4): 199–215.

Coyne, P. J. 2001. Reflections of the South Nation watershed: a pictorial history of people and natural resources. Renfrew, ON: General Store Publishing House.

El–Khoury, A., O. Seidou, D. R. Lapen, Z. Que, M. Mohammadian, M. Sunohara, and D. Bahram. Combined impacts of future climate and land use changes on discharge, and nitrogen and phosphorus loads for a Canadian river basin. Forthcoming in Journal of Environmental Management.

Fiquepron, J., S. Garcia, and A. Stenger. 2013. Land–use impact on water quality: Valuing forest services in terms of the water supply sector. Journal of Environmental Management 126: 113–121.

Fleming, R., and H. Fraser. 1999. Nitrate and phosphorus levels in selected surface water sites in Southern Ontario–1964–1994. Research report. University of Guelph, Ridgetown College. http://www.ridgetown.uoguelph.ca/research/documents/fleming_nitphos.pdf.

Fohrer, N., S. Haverkamp, K. Eckhardt, and H. G. Frede. 2001. Hydrologic response to land use changes on the catchment scale. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere 26(7–8): 577–582.

Gottle, A., and E. Sène. 1997. Forest functions related to protection and environmental conservation. Unasylya, an International Journal of Forestry and Forest Industries 48(190–191): 30–37.

Koomen, E., J. Stillwell, A. Bakema, and H. J. Scholten. 2007. Modelling land–use change: progress and applications. Dordrecht: Springer.

Lambin, E., M. Rounsevell, and H. Geist. 2000. Are agricultural land–use models able to predict changes in land–use intensity? Agriculture, Ecosystems & Environment 82(1–3): 321–331.
Le, Q. B., S. J. Park, and P. L. G. Vlek. 2010. Landuse dynamic simulator (LUDAS): a multi–agent system model for simulating spatio–temporal dynamics of coupled human & landscape system. *Ecological Informatics* 5(3): 203–221.

Lee, S. W., S. J. Hwang, S. B. Lee, H. S. Hwang, and H. C. Sung. 2009. Landscape ecological approach to the relationships of land use patterns in watersheds to water quality characteristics. *Landscape and Urban Planning* 92(2): 80–89.

Natural Resources Canada. 2002. Canada land inventory (1:50 000) – land use (circa 1966). http://data.gc.ca/.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams. 2009. *Soil and Water Assessment Tool* theoretical document version 2009, Texas Water Resources Institute Technical Report No. 406. College Station: Texas A&M University.

Qin, H. P., S. T. Khu, and X. Y. Yu. 2010. Spatial variations of storm runoff pollution and their correlation with land-use in a rapidly urbanizing catchment in China. *Science of The Total Environment* 408(20): 4613–4623.

Pavlic, G., R. A. Fernandes, W. Chen, R. Fraser, and S. Leblanc. 2002. Methods for deriving Canada–wide geo–spatial datasets in support of environmental monitoring and modelling. In: *Proceedings of the Joint International Symposium on Geospatial Theory, Processing and Applications*. http://www.isprs.org/proceedings/XXXIV/part4/pdfpapers/384.pdf.

Sala, O. E., Chapin, F. S. III, J. J. Armesto, E. Berlow, J. Bloomfield, R. Dirzo, E. Huber–Sanwald, L. F. Huenneke, R. B. Jackson, A. Kinzig, R. Leemans, D. M. Lodge, H. A. Mooney, M. Oesterheld, N. L. Poff, M. T. Sykes, B. H. Walker, M. Walker, and D. H. Wall. 2000. Global biodiversity scenarios for the year 2100. *Science* 287(5459): 1770–1774.

Seboonruang, U. 2012. A statistical assessment of the impact of land uses on surface water quality indexes. *Journal of Environmental Management* 101: 134–142.

South Nation Conservation Authority. 2006. *State of the Nation*. Finch, Ontario: South Nation Conservation Authority.

Statistics Canada. 2006a. National road network 2006. http://www12.statcan.gc.ca/census-recenmment/2011/geo/RNF-FRR/index-eng.cfm.

Statistics Canada. 2006b. Population and dwelling counts. A portrait of the Canadian population. http://www12.statcan.gc.ca/census-recenmment/2006/rt-td/pd-pl-eng.cfm.

Tong, S. T. Y., and W. Chen. 2002. Modeling the relationship between land use and surface water quality. *Journal of Environmental Management* 66(4): 377–393.

van Delden, H., J. van Vliet, D. Rutledge, and M. Kirkby. 2011. Comparison of scale and scaling issues in integrated land–use models for policy support. *Agriculture, Ecosystems & Environment* 142(1–2): 18–28.

van Vliet, J., J. Hurkens, R. White, and H. van Delden. 2012. An activity–based cellular automaton model to simulate land–use dynamics. *Environment and Planning B: Planning and Design* 39(2): 198–212.

Verburg, P. H., and K. P. Overmars. 2009. Combining top–down and bottom–up dynamics in land use modeling: exploring the future of abandoned farmlands in Europe with the dyna–CLUE model. *Landscape Ecology* 24(9): 1167–1181.

Verburg, P. H., W. Soepboer, A. Veldkamp, R. Limpiada, V. Espaldon, and S. S. Mastura. 2002. Modeling the spatial dynamics of regional land use: The CLUE–S model. *Environmental Management* 30(3): 391–405.

Walters, S. P., A. I. Thebo, and A. B. Boehm. 2011. Impact of urbanization and agriculture on the occurrence of bacterial pathogens and stx genes in coastal waterbodies of central California. *Water Research* 45(4): 1752–1762.

White, R., and G. Engelen. 1997. Cellular automata as the basis of integrated dynamic regional modelling. *Environment and Planning B: Planning and Design* 24(2): 235–246.

Wilkes, G., N. J. Ruecker, N. F. Neumann, V. P. J. Gannon, C. Jokinen, M. Sunohara, E. Topp, K. D. M. Pintar, T. A. Edge, and D. R. Lapen. 2011. Spatiotemporal analysis of cryptosporidium species/genotypes and relationships with other zoonotic pathogens in surface water from mixed–use watersheds. *Applied and Environmental Microbiology* 79(2): 434–448.

Wilkes, G., J. Brassard, T. A. Edge, V. Gannon, C. C. Jokinen, T. H. Jones, R. Marti, N. F. Neumann, N. J. Ruecker, M. Sunohara, E. D. Topp E, and D. R. Lapen. 2013. Coherence among different microbial source tracking markers in a small agricultural stream with or without livestock exclusion practices. *Applied and Environmental Microbiology* 79(20): 6207–6219.

Wright, C. K., and M. C. Wimberly. 2013. Recent land use change in the Western Corn Belt threatens grasslands and wetlands. *Proceedings of the National Academy of Sciences* 110(10): 4134–4139.

Zweg, M. H., and G. Campbell. 1993. Receiver–operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 39(4): 561–577.