ABSTRACT

The objective of this study was to evaluate Normalized Difference Vegetation Index (NDVI) data derived from QuickBird (QB) satellite imagery to map impervious areas and open spaces for runoff curve number (CN) determination. The image was collected on April 26, 2004, and provided by the City of Sioux Falls, SD. The image was radiometrically and geometrically corrected prior to this study. The imagery has four bands (including the blue, green, red, and near-infrared) and was imaged at a 2.39-meter spatial resolution. The red channel (band 3: 630-690 nm) and the near infrared channel (band 4: 760-900 nm) were used to create the QB NDVI dataset (Band 4 - Band3 / Band 4 + Band3). This research employed the urban land cover classification scheme of the runoff curve number table in the TR-55 (NRCS 1986) publication. An unsupervised spectral classification approach was used in this study. The research hypothesis was that high-resolution NDVI could improve the efficiency and effectiveness of urban land cover data extraction. For the accuracy assessment process, a simple random sampling scheme was employed using a random number generator to select points from the classified QB NDVI imagery for comparison to the 0.6-meter spatial resolution orthoimage acquired on April 23, 2004. The overall and individual Kappa coefficients were calculated. The overall accuracy for the two QB NDVI thematic maps produced was 95% and 73%, respectively. The average CN values of homogeneous land use / land cover and soil type covering approximately 30% of the study area were determined and compared to the SCS CN in order to validate the utility of QB NDVI imagery for CN calculation. The overall average CN produced from the spatial modeling was 85 and the average SCS CN was 79.

INTRODUCTION

Remote satellite sensing application in hydrology did not start until the first launch of Earth Resources Technology Satellite (ERTS) [later renamed Landsat 1] in July of 1970 (Schultz, 1988). Remotely sensed data is rapidly becoming an important source of information in hydrologic modeling, monitoring, and research (Engman 1982; 1993; Melesse and Graham 2004). Its advantages include spectral, radiometric, spatial, and temporal resolutions including its ability to generate information and to provide a means of observing hydrological state variables over large areas or small complex basins. It offers an alternative to conventional data collection employed in the estimation of certain hydrologic model parameters. The primary benefits obtainable from using remote sensing are the reduction of manpower, time, and monetary resources required for data collecting and processing. For some situations, remote
sensing provides information that was previously impossible or impractical to obtain using conventional techniques (Thomas and McCuen, 1979). Therefore, remote sensing technology is a very important tool for successful hydrologic model analysis, prediction, and validation. The National Resource Conservation Service (NRCS), formally the Soil Conservation Service (SCS) has developed a series of hydrologic models in water resource planning and design. The SCS Curve Number (SCS-CN) method is the best-known component of the SCS hydrologic models. According to Ponce and Hawkins (1996), the origins of the CN methodology can be traced back to the thousand of infiltrometer tests carried out by SCS in the late 1930s and early 1940s. The intent was to develop basic data to evaluate the effects of watershed treatment and soil conservation measures on the rainfall-runoff process. A major catalyst for the development and implementation of the runoff CN methodology was the passage of the Watershed Protection and Flood Prevention Act (Public Law 83-566) in August 1954. The method and its use are described in the SCS National Engineering Handbook Section 4: Hydrology (NEH-4) [NRCS, 1972]. The document has since been revised in 1956, 64, 65, 71, 72, 85, and 1993. The CN method is well established in hydrologic engineering and environmental impact analyses. Its major advantage are (1) its simplicity; (2) its predictability; (3) its stability; (4) its reliance on only one parameter; and (5) its responsiveness to major runoff-producing watershed properties (soil type, land cover/treatment, surface condition, and antecedent condition) [Ponce and Hawkins, 1996]. Consequently, the methodology has been widely used in numerous applications by practicing engineers and hydrologists nationally and internationally (McCuen, 1982, 2002, 2003; Ponce and Hawkins, 1996; Golding, 1997; McCuen and Okunola, 2002; Mishra and Singh, 2003).

The CN is an index that represents the combined hydrologic effect of soil group, land cover type, agricultural land treatment class, hydrologic condition, and antecedent soil moisture (McCuen, 1982). It is the best-known component of a series of SCS hydrologic models, most noticeably the TR-55 (NRCS, 1986) procedures and the TR-20 computer model (NRCS, 1992; McCuen, 2002). The CN calculation requires use of storm rainfall and associated streamflow data for annual floods to derive the best means of establishing runoff curve numbers. A CN of 100 represents a condition of zero potential retention or 100% runoff of incident precipitation.

Determination of the CN has been traditionally a time-consuming and labor-intensive procedure. A basic problem exists in quantifying detailed spatial extent and distribution of various land cover classes. Therefore, remote sensing technology has become one of the primary methods for acquiring land cover data for providing detailed mapping information as input into the CN method (Jackson and Ragan, 1977; Engman 1982, 1993; Draper and Rao, 1986; Schultz, 1996; Dubavah et al., 2000).

In previous studies, one of the main problems of using remotely sensed data for estimation of land cover is a spatial resolution issue. Moderate spatial resolutions are insufficient to generate the detail necessary to utilize the published SCS land cover table used in the SCS procedures (Ragan and Jackson, 1980; Engman, 1982, 1993). Engman, in 1982 and 1993 suggested that one must develop a land cover table analogous that is compatible to the Landsat data. Landsat may not be acceptable in small areas with certain land cover types. Bondelid et al. (1982) stated that special care should be exercised in analyzing small watersheds that contain heterogeneous land cover causing the largest CN errors. Engineering designs utilizing the curve number are widely used but the curve number is sensitive to errors in the estimated input values used to derive it (Bondelid et al., 1982; McCuen, 2002).

Remote sensing, GIS, GPS, and image processing technological capability have improved significantly since the previous studies. Commercial, high-resolution satellite imagery creates unique opportunities to address and potentially solve old problems. The purpose of this applied research is to bridge remote sensing and GIS to produce scientific knowledge by designing technological methodologies for determining the composite of curve numbers in the urban hydrology for small watersheds technical release 55 (TR-55) table NRCS (1986) publication for practicing professionals in urban water resource planning. The theoretical foundation is that mapping a high resolution NDVI image generated using the ISODATA algorithm is an efficient and effective information extraction approach for the composite of runoff curve number calculation.

**METHODS**

In this research, newly designed technological methodologies were proposed for mapping urban land use / land cover using QuickBird NDVI imagery and developing GIS spatial modeling for generating the composite of curve numbers in the TR-55 table, NRCS (1986) publication. To achieve this goal, a sequence for digital processing and analysis was proposed and implemented (Figure 1). In order to assess the utility ofQB NDVI imagery and to validate the composite runoff curve number calculation using designed GIS spatial modeling, various procedures and
comparisons of the generated runoff coefficients were presented, reviewed, and validated by practicing professionals, including the City Drainage Engineer and the City Geographic Information Systems manager, for the City of Sioux Falls, South Dakota in August 2005.

Figure 1. Sequence for digital processing and analysis.

Study Area

The City of Sioux Falls is located in southeastern South Dakota. Sioux Falls covers 40,208 acres (16271.6 hectares), and has a population of approximately 138,000. Climatologically, Sioux Falls is in a continental climate with an average rainfall of 23.86 inches (60.6 cm) and approximately 40 inches (101.6 cm) of snow annually. The majority of the precipitation comes in spring and summer with May and June being the months of maximum precipitation for the region. The study area for this project is centered on 43° 31’ 18” north latitude and 96° 44’ 42” west longitude in the southwestern part of the city. This area encompasses 2,901.42 acres (1174.2 hectares) with various land use / land cover types and includes all or part of 26 hydrological urban sub-basins (Figure 2).

Figure 2. The 2004 images of the study area at the same scale: (a) The QuickBird multi image (4-3-2) on the left and (b) The QuickBird NDVI image on the right.
Digital Data & Pre-Processing

The 2004 remotely sensed data and vector GIS data layers were provided by the City of Sioux Falls. They include orthophoto mosaics with 2 ft. (0.6 m) and 0.5 ft. (0.15 m) resolution, a QuickBird satellite image with 7.96 ft. (2.39 m) resolution and include data layers such as parcels, city limit boundary, hydrography, and streets. The data sets were processed and merged to combine the remotely sensed data with the GIS layers for the study area. The April 23, 2004 orthophoto was used as a reference to integrate all data sets into the same map projection and unit, with the Universal Transverse Mercator (UTM) map projection, World Geodetic System (WGS) 1984, and horizontal units in feet. Software used in this project consisted of ERDAS Imagine 8.7, ArcMap 9, ArcView 3.3, and Microsoft Excel 2002.

The two digital orthophotographs were acquired on April 23, 2004 and May 20, 2002, respectively. Both orthophotos were collected and scanned to meet the combined requirements of National Map Accuracy Standards; 90% of all contours will be within ½ contour interval, and 90% of horizontal positions shall be within 1/30 of one inch at the specified map scale. The specified horizontal map scale was 1 inch = 200 ft (60.9m). The flight height above terrain was 4,000 ft (1219.2m) for a photo negative scale of 1 inch = 667 ft (203.3m). Contour interval for the elevation data was 2 ft (0.61m). The radiometric density of the digital orthophotograph was 10 bits per channel with a geometric precision of 1 micron. Both orthophotos were provided in UTM projection, WGS84 datum, and planar distance survey units in feet. Additionally, both orthophotos were subset to the study area and the 2002 orthophoto was degraded 3x3 using ERDAS Imagine 8.7. Overlaying of the subset orthophotos demonstrated good alignment between images. Therefore, an image to image registration was not required.

The QuickBird leaf-on image was collected on April 26, 2004. The image was radiometrically corrected and orthorectified prior to this study, with UTM projection, WGS84 datum, and units in meters. The imagery has four bands (including the blue, green, red, and near-infrared) and was imaged at a 2.39-meter spatial resolution in unsigned 16 bit data type. The subset QB image was created in unsigned 8 bit data type. The image was observed with no signs of atmospheric contamination. The statistical spectral information showed a normal distribution with similar mean, median, and mode values. Geometric correction was performed. The subset scene was registered to the 2004 orthophoto with 31 ground control points and the RMS was 0.6940 pixel. Figure 3 demonstrates the result. The red channel (band 3: 630-690 nm) and the near infrared channel (band 4: 760-900 nm) were processed to create the QB NDVI image (NDVI = Band 4 - Band3/ Band 4 + Band3) in unsigned 8 bit data type for the study area.

The 2004 NRCS SSURGO 1:24,000 data set of the soil survey area (SSA) named “Minnehaha County, South Dakota” and its standard metadata files were downloaded from the Soil Data Mart at http://soildatamart.nrcs.usda.gov on June 7, 2004. It is a digital soil survey and generally is the most detailed level of soil geographic data developed by the National Cooperative Soil Survey. The soil layer was in a single zip file in ArcView shapefile format, with a geographic coordinate system and using North American Datum (NAD) 1983. This layer was processed using ArcView 3.3 geoprocessing and Microsoft Excel 2002 to create a soil layer associated with hydrologic soil groups in the project coordinate system for this study area. Overlaying of the soil hydro features to the subset QB image demonstrated a good fit. Therefore, a vector to image registration was not necessary.

![Image](image_url)

Figure 3. The registered 2004 images comparison at same scale and location:
(a) The QuickBird multi image (4-3-2) on the left with the 2004 Orthoimage (1-2-3) on the right.
(b) The QuickBird NDVI image on the left with the 2004 Orthoimage (1-2-3) on the right.
Classification Approach

With the recent development of new remote sensing systems, there are an increasing number of commercially high resolution data available. There is a considerable amount of previous research that has been devoted to exploring the magnitude and impact of spatial resolution on image analysis when changing scale from coarse to fine resolutions. Due to the more heterogeneous spectral-radiometric characteristics within land use / land cover unit portrayed in high resolution images, many applications of traditional single-resolution classification approaches have not led to satisfactory results. In general, traditional single-resolution classification procedures such as supervised, unsupervised, and hybrid training approaches are inadequate for discriminating between land use / land cover classes where spectral/spatial features and spatial patterns vary as a function of spatial resolution. Consequently, there is an increasing need of researching for more efficient approaches to process and analyze these data products (Chen and Stow, 2003).

In this research, the relatively simple Normalized Difference Vegetation Index (NDVI) was introduced in order to reduce heterogeneous spectral-radiometric characteristics within land use / land cover surfaces portrayed in high resolution images. It also was used for improving the accuracy level of details for mapping impervious surfaces and open spaces with different hydrologic conditions as used in the proposed SCS runoff curve number (CN) calculation in this study. Additionally, this research presented applying the principle behind NDVI using high spatial resolution QuickBird imagery based on the unique interaction differences in land use / land cover surfaces with reflected or emitted electromagnetic radiation between QB Red band (630-690nm) and QB NIR band (760-900nm).

NDVI has been shown to correlate green leaf biomass and green leaf area index. Chlorophylls, the primary photosynthetic pigments in green plants absorb light primarily from the red and blue portions of the spectrum, while a higher proportion of infrared is reflected or scattered. As a result, vigorously growing healthy vegetation has low red-light reflectance and high near-infrared reflectance, and hence high NDVI values. Impervious surfaces (e.g. Asphalt, concrete, buildings, etc.) and bare land (e.g. bare soil, rock, dirt, etc.) have similar reflectance in the red and the near infrared, so these surfaces will have values near zero. NDVI is calculated using the formula (Near infrared – red)/(Near infrared + red). This relatively simply algorithm produces output values in the range of -1.0 to 1.0.

Figure 4 shows a visual comparison of spectral correlation between QuickBird NDVI in unsigned 8 bit data type (0-255) for impervious areas and open spaces in the urban environment. The impervious surfaces are the dark areas with low NDVI DNs and the brighter areas are vegetative covers with high NDVI DNs.

![The 2004 QuickBird NDVI](image)

(a) The 2004 QuickBird NDVI

![The QuickBird multi image (4-3-2)](image)

(b) The QuickBird multi image (4-3-2)

![The 2004 Orthophoto (1-2-3)](image)

(c) The 2004 Orthophoto (1-2-3)

Figure 4. The registered 2004 images visual comparison of impervious areas and vegetative open spaces at same scale and location in residential areas.
This research employed the urban land cover of the runoff curve number table in the TR-55, NRCS (1986) publication (Figure 5). The classification schemes defined from the TR-55 are detailed in Table 1 and Table 2. The defined classification schemes were used to generate two QB NDVI thematic maps using the ISODATA algorithm, Gaussian Maximum Likelihood Classifier, with ERDAS Imagine 8.7; assigning unsupervised spectral classes and assessing the accuracy of the maps. The detailed map information from the second thematic map was then converted as a GIS input into a spatial modeling for calculating the composite of CN’s using the composite runoff index spatial model below.

This spatial model based on the composite of the CN’s equation below can be applied for calculating the runoff index for both the SCS CN method and the rational method; the CN number calculated for the SCS CN method and a runoff coefficient is the “c” value used in the rational method recommended by the American Society of Civil Engineers and the Water Pollution Control Federation.

The Composite of Runoff Index’s Spatial Model (© 2005 Pravara Thanapura. Use with permission):

\[
\text{Runoff Index}_{\text{composite}} = \left( \frac{\text{% of area covered}}{\text{CN}} \right) \times \text{CN}
\]

Where:
- \( \text{CN}_{\text{composite}} \) is the sum of the component runoff index within an area (i) delineated by description of area;
- \( \text{% of area covered} \) is the component (j) of the area (i) divided by the total area (i). Note that the component of the area is delineated by surface characteristics of land use / land cover and/or hydrologic soil group;
- \( \text{CN}(j) \) is a runoff index of the component (j) of the area (i) determined by the surface characteristics and/or its hydrologic soil groups.

The Composite CN’s Weigh Average Equation:

\[
\text{CN}_{\text{composite}} = \left( \frac{\text{% of area covered} \times \text{CN}}{\text{CN}} \right)
\]

Where:
- \( \text{CN}_{\text{composite}} \) is the sum of composite runoff curve number within an area;
- \( \text{% of area covered} \) is a component of the area divided by the total area;
- \( \text{CN} \) is the runoff curve number of the area.

In this study, the composite of CN’s calculation requires the combined surface characteristics and hydrologic effect of soil groups. The surface characteristic was derived from the two QB NDVI thematic maps and the hydrologic soil groups were obtained from the 2004 NRCS SSURGO 1:24,000 data set. The soil information covered by hydrologic areas such as rivers or canals are not surveyed or were not available in this data set. Therefore, they were excluded from the decision rules. Swimming pools were considered as impervious areas since these small size water features appeared spectrally impervious on the QB NDVI. Additionally, they retain precipitation and act as pervious surfaces during storms, thus the decision rules would result in a conservative calculation of CN’s.

There is one heavy industrial area in the southern part of the study area. This area was an oil and fuel storage facility. The tanks and their associated grass covered containment areas approximately 0.38 % of the study areas 2,901.42 acres (1174.2 hectares). The tank structures themselves were impervious but they were spectrally ruled as open spaces, which resulted in a conservative composite CN calculation. Precipitation falling on the containment areas does not interact with the surface and groundwater, since precipitation falling in these areas evaporates naturally or is pumped out and treated as wastewater.
# Chapter 2

## Estimating Runoff

<table>
<thead>
<tr>
<th>Cover type and hydrologic condition</th>
<th>Average percent impervious area #</th>
<th>Curve numbers for hydrologic soil group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>

### Fully developed urban areas (vegetation established)

- **Open space (lawns, parks, golf courses, cemeteries, etc.):**
  - Poor condition (grass cover < 50%)
  - Fair condition (grass cover 50% to 75%)
  - Good condition (grass cover > 75%)

- **Impervious areas:**
  - Paved parking lots, roofs, driveways, etc.
  - Streets and roads:
    - Paved, curbs and storm sewers (excluding right-of-way)
    - Paved, open ditches (including right-of-way)
    - Gravel (including right-of-way)
    - Dirt (including right-of-way)
  - Western desert urban areas:
    - Natural desert landscaping (pervious areas only)
    - Artificial desert landscaping (impervious weed barrier, desert shrub with 1-to-2 inch sand or gravel mulch and basin borders)

- **Urban districts:**
  - Commercial and business
  - Industrial

- **Residential districts by average lot size:**
  - 1/8 acre or less (town houses)
  - 1/4 acre
  - 1/3 acre
  - 1/2 acre
  - 1 acre
  - 2 acres

### Developing urban areas

- **Newly graded areas:**
  - (pervious areas only, no vegetation)

- **Idle lands (CN's are determined using cover types similar to those in Table 2-2c).**

---

1. Average runoff condition, and Iₐ = 0.28.
2. The average percent impervious area shown was used to develop the composite CN's. Other assumptions are as follows: impervious areas are directly connected to the drainage system, impervious areas have a CN of 98, and pervious areas are considered equivalent to open space in good hydrologic condition. CN's for other combinations of conditions may be computed using figure 2-3 or 2-4.
3. CN's shown are equivalent to those of pasture. Composite CN's may be computed for other combinations of open space cover type.
4. Composite CN's for natural desert landscaping should be computed using figures 2-3 or 2-4 based on the impervious area percentage (CN = 98) and the pervious area CN. The pervious area CN's are assumed equivalent to desert shrub in poor hydrologic condition.
5. Composite CN's to use for the design of temporary measures during grading and construction should be computed using figure 2-3 or 2-4 based on the degree of development (impervious area percentage) and the CN's for the newly graded pervious areas.

---

*Figure 5. TR-55 runoff curve number table for urban areas (United States Department of Agriculture, 1986).*

---

*Pecora 16 “Global Priorities in Land Remote Sensing”

October 23 – 27, 2005 ** Sioux Falls, South Dakota*
Table 1. Initial land use / land cover classification scheme of QuickBird NDVI thematic map#1 (Figure 6) [© 2005 Pravara Thanapura. Use with permission].

<table>
<thead>
<tr>
<th>Labels</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land Use / Land Cover</td>
<td></td>
</tr>
<tr>
<td>Character of Surface:</td>
<td></td>
</tr>
<tr>
<td>Impervious Areas</td>
<td>If land area has &lt; 25% covered with areas characterized by</td>
</tr>
<tr>
<td></td>
<td>vegetative open spaces then Impervious Areas (1)</td>
</tr>
<tr>
<td></td>
<td>If land area &gt; or = 75% characterized by impervious surfaces</td>
</tr>
<tr>
<td></td>
<td>(e.g. Asphalt, concrete, buildings, etc.) then Impervious Areas (1)</td>
</tr>
<tr>
<td></td>
<td>If land area &gt; or = 75% covered by bare land (e.g. Bare rock,</td>
</tr>
<tr>
<td></td>
<td>gravel, silt, clay, dirt, and sand or any other earthen materials</td>
</tr>
<tr>
<td></td>
<td>then Impervious Areas (1)</td>
</tr>
<tr>
<td>Open Spaces</td>
<td>Else if land area &lt; 25% covered with areas characterized by impervious surfaces then Open Spaces (2)</td>
</tr>
<tr>
<td></td>
<td>If land area &gt; 75% covered with vegetation naturally existing or planted (e.g. grass, plants, tree (leaf-on /leaf-off), forest, shrub, and scrub, etc.) then Open Spaces (2)</td>
</tr>
<tr>
<td></td>
<td>Else Impervious Areas (1)</td>
</tr>
</tbody>
</table>

Table 2. Secondary land use / land cover classification scheme of QuickBird NDVI thematic map#2 (Figure 7) [© 2005 Pravara Thanapura. Use with permission].

<table>
<thead>
<tr>
<th>Labels</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land Use / Land Cover</td>
<td></td>
</tr>
<tr>
<td>Character of Surface:</td>
<td></td>
</tr>
<tr>
<td>Open Spaces</td>
<td>If land area = or &lt; 75% covered with vegetation naturally existing or planted (e.g. grass, plants, tree (leaf-on /leaf-off), forest, shrub, and scrub, etc.) then Open Spaces - Poor condition (1)</td>
</tr>
<tr>
<td></td>
<td>Else if land area &gt; 75% covered with vegetation naturally existing or planted (e.g. grass, plants, tree (leaf-on /leaf-off), forest, shrub, and scrub, etc.) then Open Spaces - Good to fair condition (2)</td>
</tr>
<tr>
<td></td>
<td>Else Open Space Poor condition (1)</td>
</tr>
</tbody>
</table>

**Sampling Design & Accuracy Assessment**

A simple random sampling method was chosen to ensure unbiased sample selection. To provide a statistically sound assessment of accuracy, the simplified, worst-case scenario equation was used to calculate the sample sizes to insure sufficient samples for filling in an error matrix for this project:

The Conservative Sample Size Equation (Congalton and Green, 1999):

\[ n = \frac{B}{4b^2} \]

Where:
- \( n \) = Total samples of all classes
- \( B \) = The upper \((\alpha/k) \times 100\)th percentile of the Chi-Squared distribution with 1 degree of freedom \((\alpha)\); \( k \) = number of classes
- \( b \) = Significant level = +/- 5% accuracy \((\alpha - 0.95 = 0.05)\)
Over 500 sample points were generated for each thematic map using the ISODATA algorithm to insure good distribution across the study area. Figure 6 and 7 show the accuracy assessment sites acquired for the entire classification of each QB NDVI thematic map. To accomplish this, ERDAS 8.7's random number generator function was used to create random points from the QB NDVI thematic maps for comparison to the 2004 2 ft. (0.6 m) and the 2002 0.5 ft. (0.15 m) spatial resolution orthophotos. The program automatically discarded random points not sufficiently representing relatively homogenous areas. This function minimized registration errors of the image and helped to accurately interpret a true representation of each reference point on the orthophoto. To assure visual interpretation consistency, these reference points were individually evaluated by one team member with many years of photo interpretation experience. In some cases, there were conflicts due to ‘lean’ in the orthophoto or mixed pixels around the edges. These points were re-evaluated using the 2004 orthophoto. Quality control was performed on the data set by the principal author. Finally, the error matrices were generated and the overall and individual Kappa coefficients were calculated (Table 3 and 4).

Figure 6 Classification result of NDVI QuickBird thematic map #1 and 502 accuracy sites.

Figure 7 Classification result of NDVI QuickBird thematic map #2 and the accuracy assessment sites.
GIS Spatial Modeling

A spatial modeling or analysis is a set of analytical procedures applied with a GIS. A spatial model is a set of clearly defined analytical procedures applied with a GIS. A spatial model is a set of procedures that run on a spatial database to derive new information describing a new spatial relationship. It is structured as a set of rules including a combination of logical expressions, mathematical procedures, geometric models, such as calculating the Euclidean distance between objects, generating buffers, calculating areas and perimeters, and so on; (2) coincidence models, such as polygon overlay; and (3) adjacency models (pathfinding, redistricting, and allocation). All three model categories support operations on geographic data objects such as points, lines, polygon, TINS, and grids (ESRI staff, 1995).

In GIS, there are three categories of spatial modeling functions that can be applied to geographic data objects: (1) geometric models, such as calculating the Euclidean distance between objects, generating buffers, calculating areas and perimeters, and so on; (2) coincidence models, such as polygon overlay; and (3) adjacency models (pathfinding, redistricting, and allocation). All three model categories support operations on geographic data objects such as points, lines, polygon, TINS, and grids (ESRI staff, 1995).

In this study, the geometric and coincidence modeling functions in ArcView3.3 geoprocessing were designed in a sequence of steps to build a spatial modeling based on the composite runoff index spatial model. The process of spatial

---

**Table 3. Accuracy assessment result of NDVI QuickBird thematic map#1**

<table>
<thead>
<tr>
<th>Class Name</th>
<th>Labeling Spectral Clusters</th>
<th>Areas (Acres)</th>
<th>Areas (%)</th>
<th>Reference Totals</th>
<th>Classified Totals</th>
<th>Number Correct</th>
<th>Producers Accuracy (%)</th>
<th>Users Accuracy (%)</th>
<th>Kappa Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impervious areas</td>
<td>1-40</td>
<td>1270.98</td>
<td>43.72</td>
<td>258</td>
<td>216</td>
<td>210</td>
<td>81.40</td>
<td>97.22</td>
<td>0.9429</td>
</tr>
<tr>
<td>Open spaces</td>
<td>41-100</td>
<td>1635.81</td>
<td>56.28</td>
<td>244</td>
<td>286</td>
<td>238</td>
<td>97.54</td>
<td>83.22</td>
<td>0.6734</td>
</tr>
</tbody>
</table>

**Table 4. Accuracy assessment result of NDVI QuickBird thematic map#2**

<table>
<thead>
<tr>
<th>Class Name</th>
<th>Labeling Spectral Clusters</th>
<th>Areas (Acres)</th>
<th>Areas (%)</th>
<th>Reference Totals</th>
<th>Classified Totals</th>
<th>Number Correct</th>
<th>Producers Accuracy (%)</th>
<th>Users Accuracy (%)</th>
<th>Kappa Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor</td>
<td>1-25</td>
<td>730.66</td>
<td>25.14</td>
<td>242</td>
<td>239</td>
<td>164</td>
<td>67.77</td>
<td>68.62</td>
<td>0.4547</td>
</tr>
</tbody>
</table>

**GIS Spatial Modeling**

A spatial modeling or analysis is a set of analytical procedures applied with a GIS. A spatial model is a set of clearly defined analytical procedures that run on a spatial database to derive new information describing a new spatial relationship. It is structured as a set of rules including a combination of logical expressions, mathematical procedures, and criteria built using analytical tools in GIS which are applied for the purpose of simulating the process (ESRI staff, 1995).

In GIS, there are three categories of spatial modeling functions that can be applied to geographic data objects: (1) geometric models, such as calculating the Euclidean distance between objects, generating buffers, calculating areas and perimeters, and so on; (2) coincidence models, such as polygon overlay; and (3) adjacency models (pathfinding, redistricting, and allocation). All three model categories support operations on geographic data objects such as points, lines, polygon, TINS, and grids (ESRI staff, 1995).

In this study, the geometric and coincidence modeling functions in ArcView3.3 geoprocessing were designed in a sequence of steps to build a spatial modeling based on the composite runoff index spatial model. The process of spatial
modeling created a new data set (ArcView shapefile) of geographic features. Each shapefile contained new polygons that described new spatial relationships of the data sets among surface characteristics of land use / land cover and hydrologic soil groups including its CN.

Next is an example of applying the composite runoff index spatial model to develop spatial modeling for the composite of CN’s calculation within the study area. Finally, the CN results were compared to the TR-55, NRCS (1986) publication (Figure 5) in order to validate the utility of QB NDVI imagery, as shown in Table 8.

The Composite of Runoff Index’s Spatial Model for SCS CN and SCS adjusted CN (© 2005 Pravara Thanapura. Use with permission):

\[
CN_{\text{composite}} = \left( \% \text{ of area covered}_{(j)} \times CN_{(j)} \right)
\]

Where: \(CN_{\text{composite}}\) is the sum of the component CN within a parcel boundary (i) delineated by the land activity code of the City of Sioux Falls (Table 5); 
\(\% \text{ of area covered}\) is a component (j) of the area (i) divided by the total area (i). Note that the component of the area is a subset of polygon features delineated by surface characteristics of land use / land cover (GIS layer = ArcView shapefile) derived from two QB NDVI thematic maps using the ISODATA algorithm and hydrologic soil group as described previously (Figure 8); 
\(CN_{(j)}\) is a runoff curve number (Table 6) of the component (polygon)[j] of the area (i) determined by the surface characteristics and its hydrologic soil groups.

<table>
<thead>
<tr>
<th>Activity Code and Description</th>
<th>SCS CN Cover Type(^1)</th>
<th>Rule by Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 Single family</td>
<td>Residential districts by average lot size -</td>
<td>(\geq 0.1) and (\leq 0.125) acre</td>
</tr>
<tr>
<td></td>
<td>Residential 1/8 acre (0.125 acres or 505.85 sq.m(^2))</td>
<td>(\geq 404.68) and (\leq 505.85) sq.m(^2)</td>
</tr>
<tr>
<td>11 Single family</td>
<td>Residential districts by average lot size -</td>
<td>(&gt; 0.125) and (\leq 0.25) acre</td>
</tr>
<tr>
<td></td>
<td>Residential 1/4 acre (0.25 acres or 1011.71 sq.m(^2))</td>
<td>(&gt; 505.85) and (\leq 1011.71) sq.m(^2)</td>
</tr>
<tr>
<td>11 Single family</td>
<td>Residential districts by average lot size -</td>
<td>(&gt; 0.25) and (\leq 0.333) acre</td>
</tr>
<tr>
<td></td>
<td>Residential 1/3 acre (0.333 acres or 1347.59 sq.m(^2))</td>
<td>(&gt; 1011.71) and (\leq 1347.59) sq.m(^2)</td>
</tr>
<tr>
<td>11 Single family</td>
<td>Residential districts by average lot size -</td>
<td>(&gt; 0.333) and (\leq 0.5) acre</td>
</tr>
<tr>
<td></td>
<td>Residential 1/2 acre (0.5 acres or 2023.41 sq.m(^2))</td>
<td>(&gt; 1347.59) and (\leq 2023.41) sq.m(^2)</td>
</tr>
<tr>
<td>11 Single family</td>
<td>Residential districts by average lot size -</td>
<td>(&gt; 0.5) and (\leq 1) acre</td>
</tr>
<tr>
<td></td>
<td>Residential 1 acre (4046.83 sq.m(^2))</td>
<td>(&gt; 2023.41) and (\leq 4046.83) sq.m(^2)</td>
</tr>
<tr>
<td>11 Single family</td>
<td>Residential districts by average lot size -</td>
<td>(&gt; 1) and (\leq 2) acres</td>
</tr>
<tr>
<td></td>
<td>Residential 2 acres (8093.65 sq.m(^2))</td>
<td>(&gt; 4046.83) and (\leq 8093.65) sq.m(^2)</td>
</tr>
<tr>
<td>31 Banks and Financial Institutions</td>
<td>Urban Districts - Commercial and business</td>
<td>(\geq 0.5) acres or (\leq 2023.41) sq.m(^2)</td>
</tr>
<tr>
<td>33 Other offices</td>
<td>Urban Districts - Commercial and business</td>
<td>(\geq 0.5) acres or (\leq 2023.41) sq.m(^2)</td>
</tr>
<tr>
<td>64 Warehousing, Distribution, and Wholesale</td>
<td>Urban Districts - Industrial</td>
<td>(\geq 0.5) acres or (\leq 2023.41) sq.m(^2)</td>
</tr>
</tbody>
</table>

\(^1\) The samples of activity codes are similar to the cover type of the urban land cover of the runoff curve number table in the TR-55, NRCS (1986) publication [Figure 5].
Figure 8 The runoff index GIS data sets derived from the 2004 parcel layer by activity codes delineated by NDVI QuickBird information and hydrologic soil groups.

Table 6. Runoff curve numbers in Percentage for the CN calculation defined from the TR-55, NRCS (1986) Publication (Figure 5).

<table>
<thead>
<tr>
<th>Land Use / Land Cover</th>
<th>Curve numbers for hydrologic Soil Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Character of Surface:</td>
<td>A¹  B¹  C¹  D¹</td>
</tr>
<tr>
<td>Impervious Areas</td>
<td>98  98  98  98</td>
</tr>
<tr>
<td>Open Spaces - Poor Condition</td>
<td>68  79  86  89</td>
</tr>
<tr>
<td>Open Spaces - Good to Fair Condition¹</td>
<td>49  69  79  84</td>
</tr>
</tbody>
</table>

¹ CN of fair condition is use for open spaces – good to fair condition.

¹ A: The soil characteristics are deep sand, deep loess, and aggregated silts (McCuen, 1982).
² B: The soil characteristics are shallow loess and sandy loam. (McCuen, 1982).
³ C: The soil characteristics are clay loams, shallow sandy loam, soils low in organic content, and soils usually high in clay (McCuen, 1982).
⁴ D: The soil characteristics are swell significantly when wet, heavy plastic clays, and certain saline soils (McCuen, 1982).
# Table 8. The CN Results and SCS CN comparison.

<table>
<thead>
<tr>
<th>GIS ID</th>
<th>Activity Code and Description</th>
<th>SCS CN Cover Type1</th>
<th>Rule by Area</th>
<th>Samples</th>
<th>Total Area2</th>
<th>CN Result (avg.)</th>
<th>SCS CN Hydrologic Soil Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Single family Residential districts by average lot size - Residential 1/8 acre (0.125 acres or 505.85 sq.m)</td>
<td>&gt;= 0.1 and &lt;= 0.125 acres</td>
<td>312</td>
<td>35.55 acres</td>
<td>86.88</td>
<td>85</td>
<td>99.94%</td>
</tr>
<tr>
<td>2</td>
<td>Single family Residential districts by average lot size - Residential 1/4 acre (0.25 acres 1011.71 sq.m)</td>
<td>&gt; 0.125 and &lt;= 0.25 acres</td>
<td>3284</td>
<td>576.6 acres</td>
<td>84.30</td>
<td>75</td>
<td>99.76%</td>
</tr>
<tr>
<td>3</td>
<td>Single family Residential districts by average lot size - Residential 1/3 acre (0.3330 acres 1347.59 sq.m)</td>
<td>&gt; 0.25 and &lt;= 0.3330 acres</td>
<td>357</td>
<td>100.06 acres</td>
<td>82.44</td>
<td>72</td>
<td>100.00%</td>
</tr>
<tr>
<td>4</td>
<td>Single family Residential districts by average lot size - Residential 1/2 acre (0.5 acres 2023.41 sq.m)</td>
<td>&gt; 0.3330 and &lt;= 0.5 acres</td>
<td>140</td>
<td>54.51 acres</td>
<td>82.03</td>
<td>70</td>
<td>100.00%</td>
</tr>
<tr>
<td>5</td>
<td>Single family Residential districts by average lot size - Residential 1 acre (4046.83 sq.m)</td>
<td>&gt; 0.5 and &lt;= 1 acres</td>
<td>43</td>
<td>26.1 acres</td>
<td>76.84</td>
<td>68</td>
<td>95.57%</td>
</tr>
<tr>
<td>6</td>
<td>Single family Residential districts by average lot size - Residential 2 acres (8093.65 sq.m)</td>
<td>&gt; 1 and &lt;= 2 acres</td>
<td>4</td>
<td>4.56 acres</td>
<td>85.27</td>
<td>65</td>
<td>100.00%</td>
</tr>
<tr>
<td>7</td>
<td>Banks and Financial Institutions Urban Districts - Commercial and business</td>
<td>&gt;= 0.5 acres</td>
<td>17</td>
<td>14.96 acres</td>
<td>94.19</td>
<td>92</td>
<td>100.00%</td>
</tr>
<tr>
<td>8</td>
<td>Other offices Urban Districts - Commercial and business</td>
<td>&gt;= 0.5 acres</td>
<td>56</td>
<td>63.96 acres</td>
<td>93.21</td>
<td>92</td>
<td>100.00%</td>
</tr>
<tr>
<td>9</td>
<td>Warehousing, Distribution, and Wholesale Urban Districts - Industrial</td>
<td>&gt;= 0.5 acres</td>
<td>11</td>
<td>20.64 acres</td>
<td>86.31</td>
<td>88</td>
<td>94.00%</td>
</tr>
</tbody>
</table>

**RESULTS & DISCUSSION**

The overall accuracy of the two QB NDVI thematic maps produced is presented in Table 3 and Table 4 using classification schemes from Table 1 and Table 2, respectively. Figure 6 and 7 shows the two NDVI QuickBird thematic maps with the highest overall classification accuracy used as a GIS input into the Sioux Falls spatial modeling for the composite of CNs calculations.

Table 3 shows the overall accuracy results of mapping impervious areas and open spaces obtained from five different labeling criteria when the same unsupervised image of 100 spectral clusters was used. The Map ID#13 with the combination labeling criteria of impervious areas (DNs range between 1-50) and open spaces (DNs range between 51-100) achieved the highest overall Kappa value of 0.9003, with the overall classification accuracy of 95.05%. Comparing the accuracy results to the accuracy results obtained from other labeling criteria showed slightly differing classification accuracy improvement increasing toward the mid DNs of QB NDVI. This pattern of change in classification accuracy showed the potential correlations between increasing and decreasing of the DNs and amounts of open spaces (vegetative areas) and impervious areas (non vegetative areas) [Figure 4]. The DNs range from 126 to 255 indicates increasing impervious surfaces, shown in shades of dark tones. The DNs range from 126 to 255 indicates increasing openspaces, shown in shades of light tones.
Table 4 shows the overall accuracy results of mapping two classes obtained from two different labeling criteria utilizing the same unsupervised image of 50 spectral clusters. The overall Kappa value for the Map ID# 21 and Map ID#22 produced was only 0.44 and 0.45, respectively. The results showed that the spectral classifications were not spectrally distinct among open spaces-poor and good-to-fair conditions. Map ID#21 was used in this study since its producer and user accuracy in percentages were similar which indicated higher map accuracy compare to Map ID#22.

Comparing the overall accuracy of the two thematic maps demonstrated using the labeling pixel criteria with the DNs 50/50 spectral cluster thresholds for assigning labels into two classes was the best method to use. This tended to improve classification accuracy with high Kappa values and similar percentages of producer and user accuracy.

Finally, the average CN values of different land use / land cover descriptions covering homogeneous soil types were calculated and compared to the SCS CN. The results of this comparison are shown in Table 8 in order to validate the utility of QB NDVI imagery for CN calculation. The overall average CN covering approximately 30% of the study area produced was 85 and the average SCS CN was 79. Comparing CN value within GIS ID produced from different approaches demonstrated that the new design technological methodologies produced conservative CN results similar to those found in the SCS CN table.

The finer resolution imagery used in this study allows for better discrimination in land use / land cover and the accuracy results validate the hypothesis of this study. This is reflected in the fact that overall accuracy for the two QB NDVI thematic maps produced was 95% and 73%, respectively. Additionally, the average CN values among the GIS ID samples described as single family showed that, decreasing average lot size, produced higher CN numbers or higher runoff. This indicates increasing amounts of impervious areas and decreasing amounts of green vegetation as expected.

CONCLUSIONS

This applied research demonstrated that the proposed spatial modeling utilizing the fine spatial resolution remotely sensed data and a GIS approach were successful in accurately automating and mapping land use / land cover surface characteristics and determining the runoff curve number in the urban hydrology for small watersheds. Therefore, the hypothesis of this study was accepted. Mapping fine spatial resolution QuickBird NDVI imagery generated using the traditional unsupervised classification was a more accurate, simpler, and efficient data extraction approach that increased speed and potentially reduces costs of both the analysis and mapping process. The application of this approach could be beneficial to municipal engineers involved with designing or maintaining stormwater management systems and drainage improvement projects.

ACKNOWLEDGEMENTS

The author would like to thank the following employees of the City of Sioux Falls, South Dakota, for their contribution to the success of this research project: Steve Van Aartsen, GIS supervisor and his staff, for all the data and information contributed for this project; Jeff Dunn, City Drainage Engineer, for his time and the literature provided on hydrology in Sioux Falls; and Sam Trebilcock, Transportation Planner for his assistance and interest.

The author would also like to thank the co authors for their support, input, effort, and specialized expertise in making this project a success. Additional thanks go to Mr. Kevin Dalsted, Director, Engineering Resource Center, South Dakota State University, Brookings, South Dakota, for the use of the hardware and software that made this research possible.
REFERENCES


Pecora 16 “Global Priorities in Land Remote Sensing”
October 23 – 27, 2005 * Sioux Falls, South Dakota