

## 신경망기법으로 분류한 토지피복도의 CN값 산정 적용성 검토

### A Study of Runoff Curve Number Estimation Using Land Cover Classified by Artificial Neural Networks

김 흥 태\* / 신 현 석\*\*  
Kim, Hong Tae / Shin, Hyun Suk

---

#### Abstract

The techniques of GIS and remote sensing are being applied to hydrology, geomorphology and various field of studies are performed by many researcher, related those techniques. In this paper, curve number change detection is tested according to soil map and land cover in mountain area. Neural networks method is applied for land cover classification and GIS for curve number calculation.

The first, sample area are selected and tested land cover classification, NN(84.1%) is superior to MLC(80.9%). So we selected NN with land cover classifier. The second, curve number from the land cover by neural network classifier(57) is compared with that(curve number) from the land cover by manual work(55). Two values are so similar.

The third, curve number classified by NN in sample area was applied and tested to whole study area.

As results of this study, it is shown that curve number is more exact and efficient by using NN and GIS technique than by (using) manual work.

**Keyword** : Remote sensing, Neural network, Land cover, Soil map, Curve number

---

#### 요 지

GIS기법과 원격탐사 기법은 수문학의 지형자료 구축과 응용 분야에 활발하게 이용되고 있으며 다방면에서 많은 연구가 진행 중이다. 본 연구에서는 산악지역에서 토양 특성과 토지 피복 상태에 따라 유출 특성이 어떻게 나타나는지를 CN값을 산정하여 평가 하였다. 토지 피복 분류에 신경망 기법을 사용하여 보다 적합한 분류 방법을 모색하고자 했고, CN값 산정을 위한 연산에 GIS기법을 사용하였다.

우선 샘플지역을 선정하여 토지 피복의 정확도를 평가하면, 기존의 최우도법(80.9%)과 신경망 기법(84.1%)에서 신경망 기법 분류 결과가 상대적으로 우수하므로 신경망 기법으로 토지 피복을 분류하였다. 그리고 SCS방법으로

---

\* 정희원 · 부산대학교 토목공학과 박사과정  
(E mail: htkim8@hanmail.net)  
\*\* 정희원 · 부산대학교 토목공학과 조교수  
(E mail: hsshin@pusan.ac.kr)

토양도를 이용하여 AMC-II 조건하에서 CN값을 산정하면 수작업 토지이용도는 55, 신경망 분류 토지 피복도는 57로 비슷한 결과로 나타났다. 이를 토대로 전체 유역에 대해서 신경망 기법으로 분류한 토지 피복도를 사용하여 CN값을 산정하여 적용함으로써 타당성을 증명했다. 앞으로 신경망 기법을 이용한 토지 피복 분류와 GIS기법의 적용으로 보다 정확하고 신속한 CN값 산정이 가능할 것으로 사료된다.

**핵심용어** : 원격탐사, 인공신경망, 토지피복, 토양도, CN

## 1. 서 론

GIS기법과 원격탐사기법은 수문학의 지형자료 구축과 응용분야에 활발하게 이용되고 있고, 현재에도 수문학 지형자료의 구축이 다양한 분야에서 많은 연구가 진행 중이다. 미계측유역의 유출량을 산정하려면 유출우량을 결정해야하는데, 일반적으로 SCS방법의 유출곡선지수, CN(Runoff Curve Number) 값을 산정하여 유출우량을 계산한다. 그러나 CN값을 산정하는 것은 간단한 문제가 아니며 복잡한 작업과정과 시간이 필요한 힘든 작업이다. 그래서 보통 적절한 값으로 가정하여 산정하는데 유출량이 CN값에 민감한 소유역의 경우 잘못된 결과를 초래할 가능성이 크므로 소유역에 대한 적절한 CN값이 산정되어야 한다. 기존의 CN값 산정 방법은 제작 시간이 오래된 지형도나 항공사진에서 수작업으로 추출한 토지피복도와 이를 토양도와 중첩함으로써 CN값을 수작업으로 연산해내는 복잡하고 부정확한 방법을 사용하였다. 그러나 최근에는 광범위하고 최신의 수문지형자료를 위성영상자료에서 쉽게 획득할 수 있는 연구가 수행되고 있다. Rangan 등(1980)은 Landsat 영상자료를 이용하여 토지피복을 분류하고 영상자료로부터 분류된 피복상태에 적합한 CN값을 제시한 바 있다. 임상준 등(1997)은 강우-유출 자료로부터 논외의 최대 잠재 지류량의 확률적 특성을 분석하여 선행 강우조건에 따른 CN값을 추정하였다. 박희성(1998)에 의하면 위성영상을 사용한 유출곡선지수의 추정은 기존의 방법에 비하여 증가하거나 향상된 결과를 보여주었으며, 단일의 토지피복을 가정하였을 때보다 복합적인 토지피복을 가정하였을 때 큰 차이는 아니지만 좀더 향상된 결과를 보였다. 한편 최근에는 양인태(1998)는 Landsat TM영상의 가시밴드의 밴드별 DN(Digital Number)을 결합하여 CN값과의 관계를 비선형 다변량 해석으로부터 구하여 미지점에 대한 유출계수를 자동으로 구할 수 있는 회귀함수식을 유도하였고, 조홍제 등(2001)은 위성영상자료를 이용하여 삼림분포 밀도를 정량화하고 삼림분포 밀도가 유출특성에 미치는 영향을

CN값으로부터 검토하였는데 개략토양도보다 정밀토양도의 사용이 합리적임을 제시했다.

본 연구에서는 원격탐사자료와 신경망기법, GIS기법을 이용하여 효율적이고 신속한 방법으로 보다 정확한 CN값을 산정하는 방법을 모색하고자 한다.

우선 작업지역을 지리산의 대원사 유역으로 정하여 이곳의 토양도와 토지 피복도를 구축하고, 대원사 유역에서 샘플지역을 재선정하여 CN값을 산정하여 타당성을 검토한 후 대원사 유역의 CN값을 산정하고, 검증을 위해 지리산 일대 전체 유역에 대해 CN값을 산정했다. 토양도는 농진청에서 제작한 개략토양도(1/50,000)를 스캐닝하여 벡터화 하는 방법을 사용하여 구축하였고, 토지 피복도는 인공위성영상 자료인 Landsat TM 영상을 분류기법에 따른 토지피복분류의 성능과 정확도를 비교하여 최적의 분류기법을 채택하여 분류에 사용했고, 또한 수작업으로 산정한 토지피복자료와 비교하여 적용성을 평가하였다.

본 연구에서는 토양도는 일정한 값으로 CN값 산정의 변화 요소가 아니므로 그대로 사용하고, 토지 피복도는 산정 방법과 분류기법에 따라 차이가 생길 수 있으므로 이를 고려하여 산정방법을 수작업과 위성영상분류에 의한 값으로 비교하였고, 분류기법은 기존의 통계적인 분류기법과 신경망기법을 비교 검토하였다.

## 2. 영상 분류 이론

### 2.1 최우도분류법

최근 가장 널리 사용되어온 최우도법, MLC(Maximum Likelihood Classification)은 영상 분류를 위한 통계적 기법의 일종으로 분류를 위한 학습자료(Training Data)를 필요로 한다는 점에서 감독분류의 일종이기도 하다. 이때 분류자료의 통계적인 특성은 자료가 정규분포를 가진다고 가정하여 모수분류법의 일종이다. 최우도법의 이론 및 적용 절차를 Jensen(1996)의 적용을 참고하여 정리하면 다음과 같다.

이론을 설명하기 위해 분류를 위한 학습자료 군(Set)

을,  $X$ 라 하고, 자료를  $c$ 개의 등급(Class)으로 분류한다고 가정한다. 임의 지역의 Landsat TM 영상 자료에서, 식(2.1)

$$\left( \begin{array}{l} \text{인공위성영상의 Band : } j=1, \dots, m, \\ \quad \quad \quad \quad \quad \quad \quad (m = \text{Band수}) \\ \text{분류하고자 하는 Class : } i=1, \dots, c, \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad (c = \text{Class수}) \end{array} \right) \quad (2.1)$$

이고 각 밴드가 ( $s \times t$ )의 배열을 가진다면 총 화소수  $n = s \times t$ 이다.

우선, 자료준비 단계에서, 학습 자료를  $X_j$ 라 할 때, 각 등급별로 지정된 학습 자료에서 매개변수인  $M_i$ (등급  $i$ 의 평균추정벡터( $m \times 1$ ))와  $V_i$ (등급  $i$ 의 공분산행렬( $m \times m$ ))을 계산할 수 있다.

두 번째로, 시험(Test) 자료를  $XT_j$ 라 할 때, 식(2.2)와 같고

$$XT^k = \begin{pmatrix} XT_1 \\ XT_2 \\ \vdots \\ XT_m \end{pmatrix}^k \quad (2.2)$$

여기서,  $j=1, \dots, m, k=1, \dots, n$ 으로 둘 수 있다.

세 번째로, 식(2.3)의 각 등급별 우도값을 계산하는 단계에서

$$\begin{aligned} P_i &= f(V_i, M_i, XT^k) \\ P_i &= \{-0.5 \log_e [\det(V_i)]\} \\ &\quad - [0.5(XT^k - M_i)^T V_i^{-1} (XT^k - M_i)] \quad (2.3) \\ P_c(\text{결정 Class}) &= \max(P_i) \\ &\rightarrow \text{Bayesian theory} \end{aligned}$$

각 화소마다 가장 큰  $P_c$ 값을 가지는 등급이 각 화소에 할당된다.

## 2.2 신경망 이론

신경망(Neural Networks)의 구조는 Bishop(1995)과 오창석(1996)의 문헌에서 입력층, 은닉층, 출력층으로 단순하지만 상당히 상호 연결된 처리 요소인 수많은 뉴런들로 구성된 정보처리 구조들로 설명되고 입력층은 입력 정보에 대하여 각 입력값에 가중치(Weight)를 곱

해서 은닉층으로 전달한다. 은닉층은 입력층에서 각기 다른 가중치에 의해 변환된 자료들을 모두 합하여 전달 함수를 통하여 출력하게 되고 이 값들은 다시 은닉층과 출력층 사이의 가중치와 곱하여져 출력층으로 보내진다. 출력층은 은닉층에서 각기 다른 가중치에 의해 변환된 자료들을 받으며, 각 신경단위는 합산된 자료들을 출력값의 목적에 따라 이진 값의 출력을 위한 시그넘(Signum)함수, 연속값의 출력을 위한 선형 또는 비선형 함수에 의하여 변형시켜 출력한다.

본 연구에선 경험적으로 신경망 기법으로써 성능이 가장 우수했던 RBFN(Radial Basis Function Networks)을 사용했으며, 방사형 기초함수에 기반을 둔 모형을 사용했는데 Haykin(1994)에 의하면 RBFN은 전 방향으로 연결되는 연속적인 세 개의 층으로 구성되어 있고 입력층에서 은닉층으로의 변환은 방사형 기초 함수(RBF)를 이용한 비선형이고, 반면에 은닉층에서 출력층으로의 변환은 선형이다. 은닉층은 많은 수의 절점과 가중치 벡터로 고려될 수 있는 "중심(center)"이라고 불리는 매개변수로 구성되어 있다. 중심에서 입력 벡터가 얼마나 떨어져 있는가를 측정하는데 표준 Euclidean 거리가 사용된다. 각 절점에서, 중심과 네트워크의 입력벡터 사이의 Euclidean 거리는 계산되고 은닉층 절점의 출력을 결정하는 비선형 함수에 의해서 변환된다. 출력층은 선형 방식에 의해서 은닉층의 결과를 조합한다. RBF 네트워크의 출력  $y$ 는 식(2.4)와 같이 계산된다.

$$y = f(x, d) = \sum_{i=1}^n w_i R_i(x) + w_0 \quad (2.4)$$

여기서,  $w_i$ 는 은닉 절점과 출력 절점 사이의 연결강도,  $w_0$ 는 편향치(Bias),  $x$ 는 입력 벡터이다. 함수  $R_i : R^n \Rightarrow R$ 은 식(2.5)와 같은 일반적인 형태를 가지는 방사형 기초 함수이다.

$$R_i = \phi \|x - m_i\| \quad (2.5)$$

$\phi(\cdot)$ 는 원점에서 최대값을 가지고, 무한으로 갈수록 급하게 하강곡선을 그린다. 함수  $\phi(\cdot)$ 은 입력과 중심 사이의 Euclidean 거리가 증가할 때, 0(Zero)으로 접근한다.  $\phi(\cdot)$ 에 대한 여러 가지 선택이 있으나, 방사형 기초 함수의 일반적인 사용은 식(2.6)과 같은 가우

시안(Gaussian) 함수가 사용된다.

$$R_i = -\exp\left(-\sum_{i=1}^n \frac{\|x_i - m_{ii}\|^2}{2\sigma_{ij}^2}\right) \quad (2.6)$$

여기서  $m_i^T = [x_{i1}, c_{i2}, \dots, c_{in}]$ 은 허용되는 구간에서의 중심이고,  $\sigma_{ij}$ 는 가우시안 함수의 폭이다.

은닉층 노드수는  $m_i$ 에 근접한 입력 벡터에 최대 반

응을 일으키며 각 은닉층 노드는 입력 공간에서 자체의 수용층(Receptive Field),  $R_i(x)$ 를 가지는데 그것은  $\sigma_i$ 에 비례해서 크기를 가진  $m_i$ 에 중심을 둔 영역이다.

분류를 위해 사용된 토지피복 분류를 위해 사용된 RBFN 신경망 구조는 그림 2.1과 같이 입력층, 은닉층, 출력층으로 나뉜다. 입력층은 1개의 픽셀에 해당하는  $i$ 개 밴드의 화소 값이 각각 입력되고, 전달함수에 의해 은닉층으로 전달된다. 은닉층은 최적의 수렴율과 속도

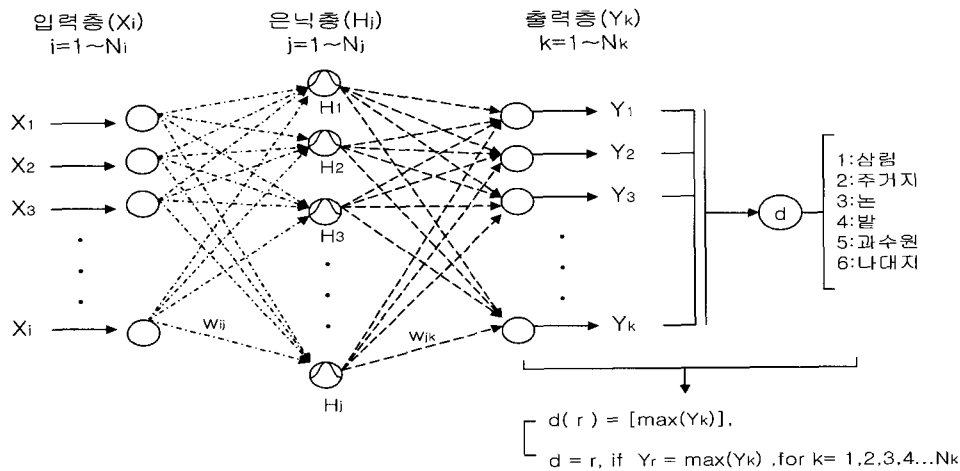


그림 2.1 토지피복분류 RBFN의 구조

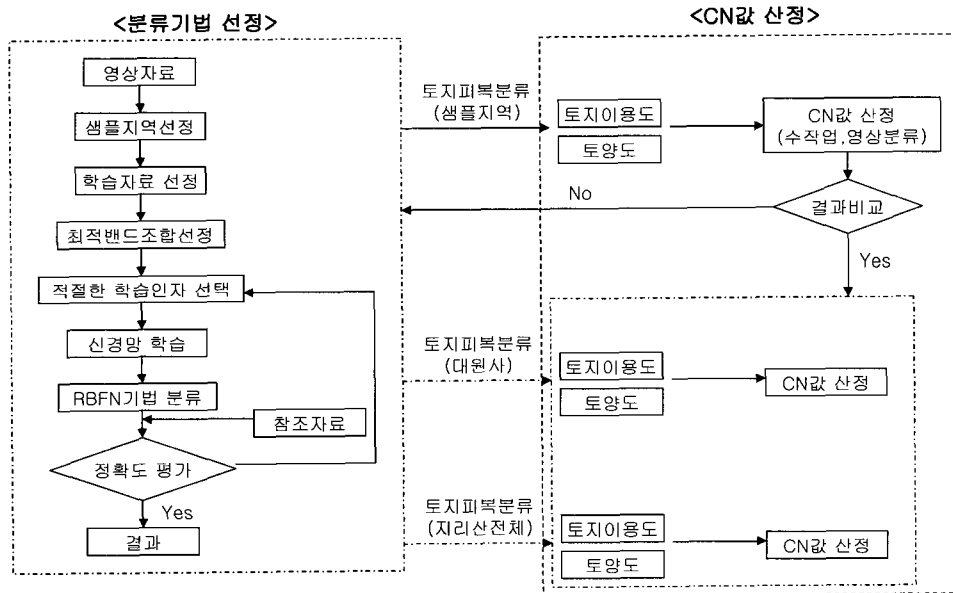


그림 3.1 실험 과정의 흐름도

를 가지도록 적절한 개수를 선택하였으며 출력층과의 전달함수에 의해 출력층으로 전달된다. 은닉층과의 가중 함수에 의해 계산된 출력값은 그대로 토지피복분류 값으로 사용될 수 없으므로 적절한 방법으로 변환해야 한다. 토지이용 분류 값의 결정은 우선, k개의 분류항목을 각 항목의 (1,k)번째 값이 1이 되고 나머지는 0(Zero)이 되도록 k개의 행렬 1×k를 만드는데 합치면 k×k의 고유 행렬이 된다. 출력층에 계산된 값이 1×k의 행렬로 나타나면 여기서 최대치만 1로 두고 나머지는 모두 0(zero)로 처리한다. 따라서 1이 위치한 행의 위치값 r을 읽으면 이것이 토지피복 분류 값 d로 환산된다. 이러한 방식으로 각 픽셀에 화소 값이 모두 d값으로 변환되면서 토지피복 분류도가 완성된다.

### 3. 적용에

#### 3.1 적용 방법

토지피복도 분류를 위해 통계적 방법으로 최우도법을, 신경망 기법으로 RBFN를 사용했는데, 전반적인 실험 과정을 정리하면 다음 그림 3.1과 같다.

#### 3.2 영상분류실험

##### 3.2.1 대상구역선정 및 적용자료

연구대상지역은 지리산 대원사 구역으로 중심의 위치는 경위도 좌표로 위도 35°21'38.15", 경도 127°45'59.65"이며, 구역의 특성은 구역면적은 34.64 km<sup>2</sup>이고, 유로연장은 15km 정도로 나타났고, 고도차는

1,708m, 유로평균경사 0.114(m/m)로 상당히 가파르며, 하천은 상류측은 서에서 동으로 진행하다 중하류에서는 남동쪽으로 흐르고 있다.

대상지역의 전처리 과정으로 기하보정을 실시했는데 중부원점을 사용했고, 영상처리 패키지인 Image Analyst 7.1을 이용하여 수행되었으며, 화소값의 재배열 방법으로는 최근치 보간법을 이용하였으며, 좌표값의 변환식은 Affine-1차 다항식이 사용되었다. 본 연구에서 사용된 영상자료는 95년 5월에 촬영된 지리산 지역 30m 해상도의 Landsat TM 영상으로 대원사구역(362×317)과 샘플지역(100×108)의 분석을 위해 연구 목적에 맞게 적절히 절취하고 재배열(Reasampling)하여 사용했는데 그림 3.2는 기하보정의 흐름도이고 그림 3.3은 실제 대상 구역의 적용결과이다.

##### 3.2.2 분류 항목의 선정

학습을 위한 화소수가 학습에 미치는 민감도를 분석하기 위해, 학습 자료군(Training Dataset)을 각각 S1, S2, S3, S4, S5 으로 세분하여 5개의 세트를 구성하였고, 이때 세트의 화소수는 임의로 무작위로 선정하였다. 여기서 초기에 선택할 수 있는 최대의 학습 자료군을 S1으로 지정하였고, 약 2000개 정도의 학습 자료(Training Data)를 추출하여 S2, 약 1000개 정도를 S3, 약 500개 정도를 S4, 약 250개 정도를 S5 으로 정했는데, 다음 표 3.1은 학습 자료의 분류항목별 화소수를 정리한 것이고, 그림 3.4는 이들 학습 자료군의 샘플지역 내에서 위치와 자료의 크기를 나타낸 것이다.

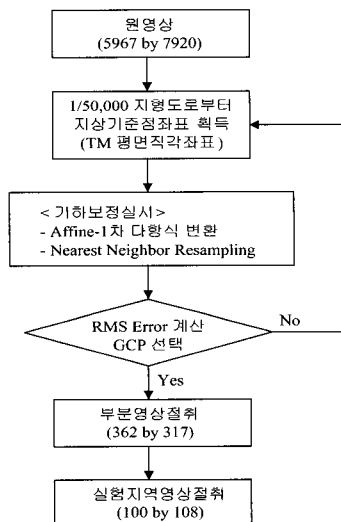


그림 3.2 기하보정의 흐름도

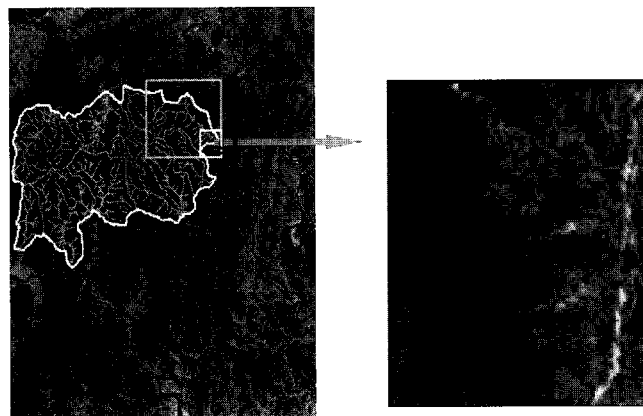
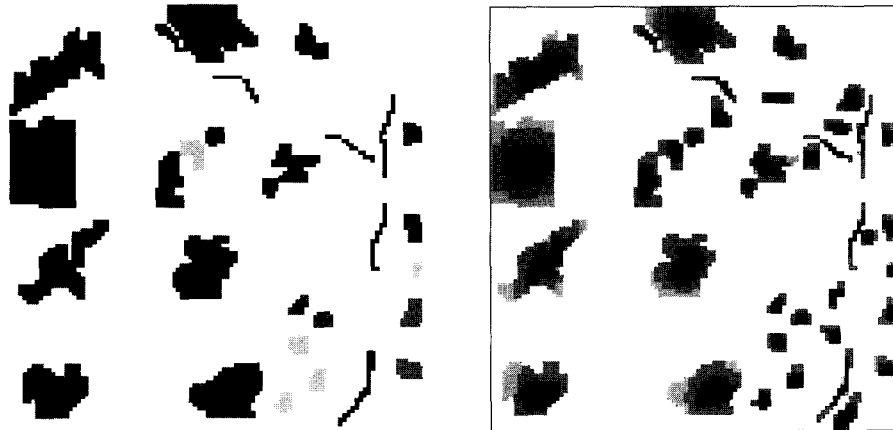


그림 3.3 지리산 지역 영상과 샘플지역 영상

표 3.1 학습자료의 분류항목별 화소수

분류 항목	S1	S2	S3	S4	S5
삼림	1706	1411	661	345	130
주거지	128	128	66	45	25
논	232	232	149	96	56
밭	112	116	74	46	30
과수원	62	62	31	24	8
나대지	130	130	82	58	39
합계(화소수)	2,371	2079	1063	614	288



(a) 학습 자료군의 위치와 항목  
 (b) 학습자료군별 자료의 크기  
 ( ■ : 삼림, ■ : 주거지, □ : 논, ■ : 밭, ■ : 과수원, ■ : 나대지)

그림 3.4 학습 자료군의 위치와 자료의 크기

표 3.2 선정된 검증 자료의 항목별 화소수

분류항목	학습 검증 자료군(화소수)	평가 검증 자료군(화소수)
삼림	13	24
주거지	7	21
논	9	24
밭	5	13
과수원	4	6
나대지	8	20
총계	46	108

### 3.2.3 분류결과 검증을 위한 검증자료의 선정

분류결과의 정확도를 평가하기 위해선 자료를 비교할 대상이 필요한데, 여기선 지형도에서 추출한 지형 정보 값을 참값으로 가정하고 검증자료를 선정했다. 학습 자료 군을 취득하는 방법과 유사한 방법으로 검증자료군(Reference Dataset)을 선정하였는데, 각 항목의 학습 자료군의 중심점을 학습 검증 자료군(Training Reference Dataset)으로 두어 각 기법의 학습 성능을 평가하는 기준으로 삼고, 각 항목의 학습 자료 군에서 지형도와 영상에서 명확하게 구분될 수 있고 학습에 사용되지 않은 지점들을 평가 검증 자료군(Test Reference Dataset)으로 구성하여 계산된 분류 정확도에 대한 평가 기준으로 정했는데 다음 표 3.2와 같다.

### 3.2.4 최적밴드조합 선택

밴드 조합수별 최적 밴드조합을 선정하기 위해 평균 변환 발산도(Average Transformed Divergence)를 계산하는데, 이것은 밴드간의 분리도 정도를 평가해주는 값으로 1900이상이면 밴드간의 간섭 없이 적당히 분리된 상태라 판단한다. 밴드 조합 중 분리도가 1900 이상인 것 중 가장 좋은 밴드 조합을 학습 자료군 별로 정리하면 다음 표 3.3과 같은데 여기서, × 표시는 분리도가 1900 이하인 자료 군으로 사용하지 않음을 나타낸다.

### 3.2.5 신경망 분류를 위한 학습인자 최적화

신경망을 빠르고 효율적으로 학습시키기 위해서는 자료 특성과 크기에 감안하여 초기 연결강도, 학습률, 은닉층의 노드수 등 여러 학습인자를 최적화 하는 과정

이 필요한데 SSE 값을 고정하고 학습자료와 검증자료의 정확도를 폭함수 0.05~0.5사이의 값으로 실험한 결과 폭함수 값은 0.2, 은닉층 노드수는 150에서 학습 자료의 정확도 95.65%, 검증자료의 정확도 86.24%로 최적의 분류 결과를 보였고, 반복횟수는 RBFN의 특성상 은닉층의 노드수와 같은 150회가 적당했다.

## 3.3 분류결과 검증

### 3.3.1 정확도 평가

분류결과의 정확도를 표현하는 가장 보편적인 방법 중 하나로 Kiefer(1994)는 오차 행렬(Error Matrix)을 사용했다. 오차 행렬은 항목 대 항목에 대해, 알고 있는 검증자료(실제 지표자료)와 자동 분류된 결과물들 사이의 일치성을 비교하여 놓은 것으로 분류정확도를 평가할 수 있는 각 항목들의 개수와 똑같은 개수의 행과 열로 이루어진 사각형태이다. 이런 매트릭스로 평가할 수 있는 정확도는 전체 정확도(Overall Accuracy), 생산자 정확도(Producer Accuracy), 사용자 정확도(User Accuracy),  $K'$  (Kappa Coefficient) 등이 있다. 본 연구에선 Jensen(1996)이 사용한  $K'$ 를 사용하여 정확도를 평가 했는데 이는 오차 행렬로부터 정확도를 평가하기 위한 수단으로 이용되며 아래의 식(3.1)과 같이 정의되어 진다. 여기에서  $x_{ii}$ 는 오차 행렬의  $i$  번째의 대각선요소를 뜻하며,  $x_{i+}$ 와  $x_{+i}$ 는 각각 행렬의 주변분포(Marginal Distribution)를 나타내기 위한 요소이고,  $N$ 은 행렬에 포함된 모든 화소의 수를 뜻한다.

표 3.3 학습자료군별 각 밴드 조합수의 최적 분리도 결과

학습자료군	최적 밴드 조합(위)과 평균 변환 발산도(아래)				
	2개	3개	4개	5개	6개
S1	×	145	1245	12457	123457
	×	1962	1988	1993	1996
S2	×	245	1245	12457	123457
	×	1971	1991	1994	1997
S3	15	245	1245	12457	123457
	1916	1990	1999	1999	2000
S4	35	235	1245	12345	123457
	1963	1997	2000	2000	2000
S5	35	257	2457	23457	123457
	2000	2000	2000	2000	2000

$$K' = \frac{N \cdot \sum x_{ii} - \sum (x_{i+} \cdot x_{+i})}{N^2 - \sum (x_{i+} \cdot x_{+i})} \quad (3.1)$$

이러한  $K'$ 는 일반적인 경우에는 널리 사용되고 있는 전체 정확도와는 다른 결과를 나타내며, 실제로 전체 정확도보다 작은 값을 보이는데, 이러한 차이는 전체 정확도와  $K'$ 가 담고있는 정보의 차이에 기인하게 된다. 전체 정확도의 경우에는 단순히 행렬의 대각선 요소들만을 이용하여 계산하게 되므로, 주변분포, 즉 생략 오

차(Omission Error)와 포함 오차(Commission Error)를 고려할 수 없으나,  $K'$ 의 경우는 행렬의 주변 분포를 이용하여 계산되어지는 값이므로 행렬의 대각선 요소가 아닌 요소들도 고려하게 되며, 따라서 이러한 생략 오차와 포함 오차라는 오차의 영향에 대한 정보를 포함하게 된다. 따라서 우연으로 인한 결과를 포함하고 있는 전체 정확도보다는  $K'$ 의 경우가 분류의 정확도를 표현하는 데에 보다 바람직하다고 판단하여 본 연구에서는  $K'$ 를 사용하여 정확도를 평가했다.

표 3.4 학습자료에서 밴드 조합수별 각 분류 기법의 학습자료수에 대한 정확도

밴드 조합수	분류기법	학습 자료군의 학습 검증 자료의 $K'$				
		S5	S4	S3	S2	S1
2개 밴드	MLC	64.6%	73.1%	70.3%	×	×
	RBFN	57.0%	66.8%	61.2%	×	×
3개 밴드	MLC	67.7%	78.5%	81.0%	83.7%	91.9%
	RBFN	64.4%	67.0%	78.1%	75.2%	72.5%
4개 밴드	MLC	73.0%	86.6%	89.2%	89.2%	89.2%
	RBFN	81.1%	89.2%	86.4%	75.3%	72.5%
5개 밴드	MLC	75.7%	89.3%	89.2%	91.9%	91.9%
	RBFN	81.0%	89.2%	83.7%	83.6%	78.0%
6개 밴드	MLC	78.4%	86.5%	86.5%	94.6%	94.6%
	RBFN	91.9%	94.6%	89.2%	75.2%	75.2%

표 3.5 평가자료에서 밴드 조합수별 각 분류 기법의 학습자료군에 대한 정확도

밴드 조합수	분류기법	학습 자료군의 평가 검증 자료의 $K'$				
		S5	S4	S3	S2	S1
2개 밴드	MLC	55.7%	53.0%	63.4%	×	×
	RBFN	44.4%	50.2%	61.2%	×	×
3개 밴드	MLC	61.7%	67.5%	68.6%	72.0%	76.5%
	RBFN	59.2%	59.3%	75.1%	77.3%	76.1%
4개 밴드	MLC	70.7%	74.2%	73.1%	73.1%	74.2%
	RBFN	62.8%	79.6%	85.3%	80.7%	76.2%
5개 밴드	MLC	71.7%	76.4%	76.5%	77.6%	78.7%
	RBFN	73.0%	80.8%	81.9%	80.7%	81.8%
6개 밴드	MLC	70.5%	77.5%	77.5%	78.7%	80.9%
	RBFN	75.1%	84.1%	84.1%	82.9%	79.5%

표 3.6 기법별 최적 학습자료군의 화소수와 밴드 조합수

분류기법	최적 학습 자료군(화소수)	최적밴드조합(밴드수)	학습정확도( $K'$ )	평가정확도( $K'$ )
MLC	S1(2370개)	123457(6개)	94.6%	80.9%
RBFN	S4(614개)	123457(6개)	94.6%	84.1%



### 3.3.2 최적 분류 조합 선정

표 3.2의 학습 검증 자료와 평가 검증 자료를 이용한 각 밴드 조합수에 따른 기법별 학습 자료수에 대한 정확도 평가 결과는 표 3.4와 표 3.5와 같고, 여기서 공통으로 80%이상의 분류 결과를 나타내는 밴드 조합수는 4개 이상이며, 분류 기법은 MLC와 RBFN, 그리고 학습 자료군은 S1, S2, S3, S4이다.

또한, 통계적기법과 신경망기법의 최적 학습 자료군과 밴드 조합수를 결정하면 다음 표 3.6과 같다.

### 3.3.4 선정된 기법별 분류 및 성능 시험

#### (1) 선정된 기법별 분류

그림 3.5는 샘플지역의 토지피복도로 각각 지형도에

서 추출한 토지피복도, MLC, RBFN기법으로 분류한 토지피복도이다.

#### (2) 최적 기법의 성능 시험

최적의 분류 정확도를 나타내는 신경망 기법(RBFN)을 이용하여 대원사 전체 유역에 대해 시험한 토지피복도 분류 결과를 나타내면 그림 3.6과 같고 지리산 전체 유역에 대해 시험한 결과는 그림 3.7과 같다.

## 3.4 대원사 유역의 수작업과 영상분류 결과의 CN값 비교

### 3.4.1 CN값 산정 이론

본 연구의 유출곡선지수(CN)의 산정은 SCS방법을

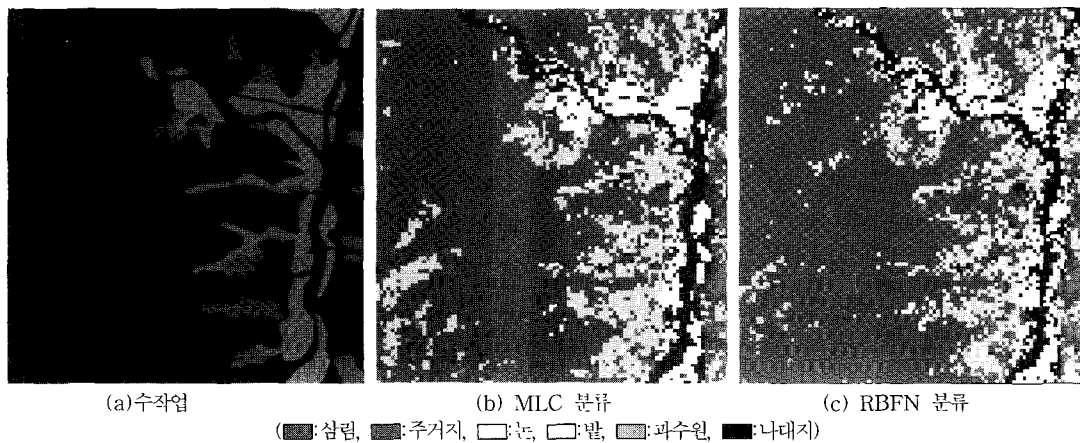


그림 3.5 샘플지역의 토지피복도

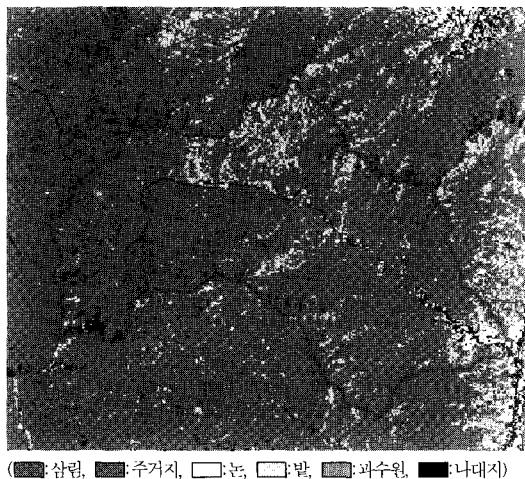


그림 3.6 대원사 유역의 토지피복도

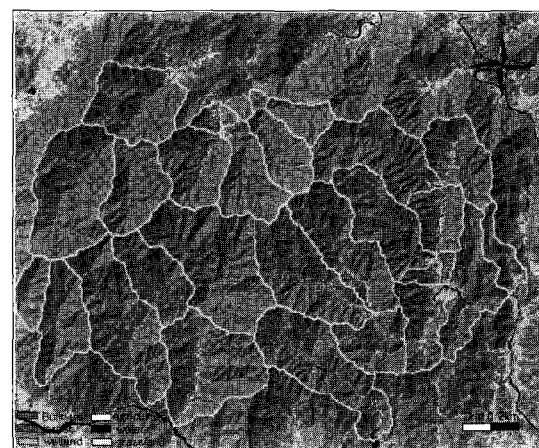


그림 3.7 지리산 전체유역의 토지피복도

사용하였는데 윤용남(2002)의 순서에 의하면 CN값은 토지피복, 수문학적 토양군, 그리고 선행함수조건에 의해 결정되어질 수 있다. 토지피복과 수문학적 토양군의 자료를 이용하여 CN값을 추정하여 면적 가중 평균법으로 대표 CN값을 도출할 수 있다. SCS의 초과우량 산정식은 아래와 같이 계산된다.

$$Q = \frac{(P-0.2S)^2}{P+0.8S} \quad (3.2)$$

(Q:누가 유출량, P:누가 강우량, S:최대 잠재 저류량)

$$\text{한편, CN값은 } CN = \frac{25400}{254+S} \quad (3.3)$$

CN값이 결정되면 식(3.3)로부터 최대 잠재저류량(S)을 구하고 식(3.2)에서 Q를 구할 수 있다. 토양분류와 지면 피복 상태를 나타내는 토지 피복을 알면 이에 대한 CN값을 토양 피복형 CN값(AMC-II, Ia=0.2S)으로부터 결정할 수 있으나 이 값은 토양의 수분상태가

건조하거나 습하지 않은 중간 값이므로 강우이전 토양의 수분상태에 따라 SCS는 선행함수조건을 설정함으로써 5일 선행강우량 P5를 기준으로 결정하는데, AMC-II조건에 대한 CN값을 결정한 다음 건조상태(AMC-I)와 습윤상태(AMC-III)에 대한 CN값은 다음 식으로부터 산정한다.

$$CN(I) = \frac{4.2CN(II)}{10-0.058CN(II)} \quad (3.4)$$

$$CN(III) = \frac{23CN(II)}{10+0.13CN(II)} \quad (3.5)$$

### 3.4.1 기초자료

표 3.7는 토지 피복도에 해당하는 CN값 이고 그림 3.8은 대원사 지역의 토양도이며, 표 3.8은 토양 구분이다.

표 3.7 토지피복도별 CN값

토지피복상태	배수상태	토양형			
		A	B	C	D
삼림	나쁨	45	66	77	88
	중간	36	60	73	79
	좋음	25	55	70	77
주거지	-	59	74	82	86
논	나쁨	70	79	84	88
	좋음	65	75	82	86
밭	나쁨	65	76	84	88
	좋음	63	75	83	87
과수원	-	43	65	76	82
나대지	-	76	85	89	91

표 3.8 대원사 지역의 토양구분

토양명	배수상태	토양의 성질
Mac	매우양호	A
Anc	약간양호	B
Ro	매우불량	D
Mab	매우양호	A
Anb	약간양호 또는 약간불량	B

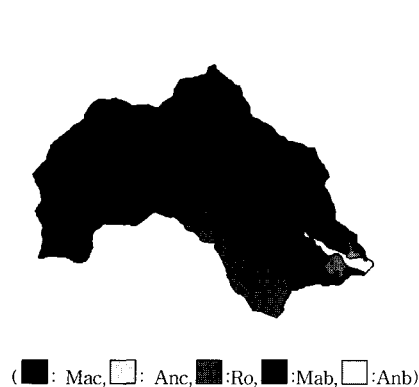


그림 3.8 대원사지역의 토양도

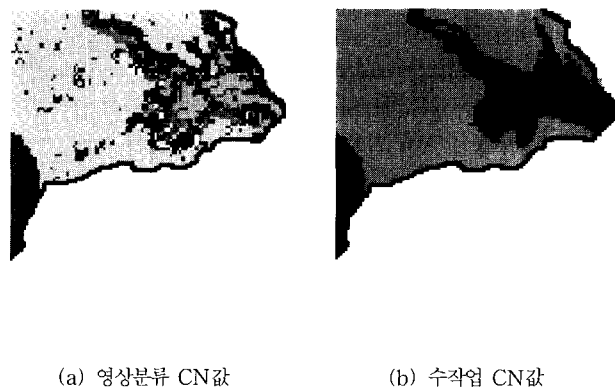


그림 3.9 영상분류시(a)와 수작업시(b)의 CN값

표 3.9 수작업시의 CN값

토지 이용 상태	토양형								총면적 (km <sup>2</sup> )	평균 CN
	A		B		C		D			
	면적 (km <sup>2</sup> )	CN	면적 (km <sup>2</sup> )	CN	면적 (km <sup>2</sup> )	CN	면적 (km <sup>2</sup> )	CN		
삼림	2.485	45	0.175	60	-	-	0.212	88	2.871	49.1
주거지	0.005	59	0.033	74	-	-	-	-	0.039	71.9
논	0.008	65	0.328	75	-	-	-	-	0.336	74.8
밭	0.015	63	0.068	75	-	-	-	-	0.083	72.8
과수원	-	-	-	-	-	-	-	-	0.000	0.0
나대지	0.077	76	0.228	85	-	-	-	-	0.305	82.7
평균치	2.591	-	0.831	-	-	-	0.212	-	3.633	55.1

표 3.10 영상 분류시의 CN값

토지 이용 상태	토양형								총면적 (km <sup>2</sup> )	평균 CN
	A		B		C		D			
	면적 (km <sup>2</sup> )	CN	면적 (km <sup>2</sup> )	CN	면적 (km <sup>2</sup> )	CN	면적 (km <sup>2</sup> )	CN		
삼림	2.083	45	0.092	60	-	-	0.203	88	2.378	49.3
주거지	0.176	59	0.119	74	-	-	0.007	86	0.302	65.5
논	0.122	65	0.278	75	-	-	-	-	0.400	72.0
밭	0.113	63	0.153	75	-	-	0.001	88	0.266	70.0
과수원	0.035	43	0.022	65	-	-	-	-	0.057	51.4
나대지	0.063	76	0.167	85	-	-	-	-	0.230	82.5
평균치	2.591	-	0.831	-	-	-	0.212	-	3.633	56.8

표 3.11 대원사 구역의 CN값

토지 이용 상태	토양형								총면적 (km <sup>2</sup> )	평균 CN
	A		B		C		D			
		CN	면적 (km <sup>2</sup> )	CN	면적 (km <sup>2</sup> )	CN	면적 (km <sup>2</sup> )	CN		
삼림		45	0.138	60	-	-	2.427	88	26.738	49.0
주거지		59	0.092	74	-	-	0.077	86	2.758	60.6
논		65	0.274	75	-	-	0.115	88	0.627	69.6
밭		63	0.150	75	-	-	0.079	88	2.138	65.6
과수원		43	0.033	65	-	-	0.007	82	1.113	44.9
나대지		76	0.175	85	-	-	0.079	91	1.315	77.9
평균치		-	0.861	-	-	-	2.786	-	34.689	52.3

3.4.2 샘플지역의 CN값 비교

그림 3.9는 샘플지역의 수작업 시와 영상분류시의 CN값 결과이고 이들을 표 3.9, 표3.10과 같이 결과를 비교하면 55.1, 56.8로 비슷한 값으로 나타났다.

3.4.3 대원사 구역의 CN값 산정

샘플지역의 결과를 바탕으로 대원사 구역에 대해서도 CN값을 산정하면, 표 3.11과 같고 이를 도시하면 그림 3.10과 같다. 이들의 결과는 AMC-II 조건하에선

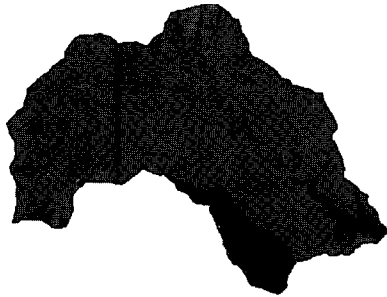


그림 3.10 대원사 유역의 CN값

52, AMC-III조건하에선 71의 값을 얻을 수 있었다.

#### 3.4.4 지리산 전체유역의 CN값 산정

위의 결과를 바탕으로 지리산 전체 유역에 대해 CN값 산정을 시험한 결과, 대부분 토양형은 A이고 토지피복 상태는 삼림으로 대부분은 56인 것으로 나타났다.

이는 일반적 자연유역의 CN값이 75내외임을 고려하면 상대적으로 작은 값이나, 대부분의 지역이 국립공원으로 지정되어 토지피복 현황이 삼림으로 일정한 데 기인하는 것으로 판단된다.

#### 3.5 비교 고찰

토지피복도 분류과정과 CN값 산정을 위한 진행과정에 따라 결과를 정리하면 다음과 같다.

1. 표 3.4와 표 3.5의 결과를 학습 자료군의 화소수의 측면에서 보면, 신경망기법(RBFN)으로 분류한 경우가 대체로, 학습 자료군의 화소수가 적을 경우에 통계적 기법(MLC) 분류 결과보다 좋게 나타났다.
2. 표 3.4와 표 3.5의 결과를 밴드 조합수의 측면에서 보면, 통계적 기법(MLC)과 신경망기법(RBFN) 모두에서 밴드 조합수가 6개일 경우에 최상의 분류 결과를 보였지만, 신경망기법(RBFN)의 경우는 적은 밴드 조합수에서도 좋은 결과를 보였다.
3. 최상의 분류 정확도를 나타내는 신경망기법(RBFN)에서 동일한 학습인자를 사용해서 대원사 유역에 성능 평가에서 표로 정리하지는 않았지만 학습 검증 자료의 경우 86.4%, 평가 검증 자료의 경우 78.5%로 샘플지역의 검증자료 표 3.6의 94.6%, 84.1%보다 낮아졌지만, 그림 3.4와 그림 3.5를 비교하면 비슷한 양상을 나타내며

로, 본 연구에서 선정된 최적의 학습 자료 군과 분류기법으로 지리산 전체 지역에 시험한 결과 지형도에서 나타난 토지이용도와 유사한 양상을 나타냈다.

4. 대원사 유역의 샘플지역에 대해 인공위성영상자료를 신경망기법(RBFN)으로 분류한 토지피복도와 수작업으로 추출한 토지피복도의 CN값 결과를 산정하여 비교하였는데 표 3.9와 표 3.10에서 알 수 있듯이 각각 55.1, 56.8로 비슷한 결과를 나타냈다.
5. 샘플지역의 CN값 산정결과를 바탕으로 대원사유역과 지리산 전체유역 각각에 대해 CN값을 산정했는데, 지리산 전체를 여러 소유역으로 분할하여 AMC-II조건하에 CN값을 산정한 결과 52~56 범위의 값으로 나타났다.

#### 4. 결 론

본 연구에서는 토지 피복 분류에 신경망기법을 적용함으로써 보다 나은 분류기법의 모색을 시도하였고 최적의 분류기법으로 인공위성영상을 분류하여 토지 이용도를 산출하였다. 또한 본 연구의 목적인 수문학적 지형인자로서의 활용을 위해 영상 분류로 얻은 토지 피복도를 이용해 CN값을 산정하고 기존의 방법과 비교 검토하여 타당함을 증명하였는데 이들 각각의 결과를 정리하면 다음과 같다.

- (1) 토지 피복 분류기법에서 결과적으로 신경망기법(RBFN)이 학습 자료수와 밴드조합수가 적을수록 최우도법(MLC) 보다 분류를 잘하고 학습 자료수와 밴드조합수가 적절히 많을 경우는 최우도법(MLC)과 신경망기법(RBFN) 모두 사용 가능하고, 또한 그림 3.5, 그림 3.6 등의 분류된 결과를 살펴보면 분류기법의 특성상 곳곳에 노이즈가 나타나고 CN값의 결과에도 그림 3.9, 그림 3.10과 같이 영향을 미친다. 그러므로 필터링 또는 참조자료를 이용한 직접 수정 등의 후처리 작업이 요구된다.
- (2) CN값 산정은 SCS방법으로 AMC-II 조건에서 수작업 토지이용도와 영상 분류 토지이용도를 샘플지역을 통해 표 3.9와 표 3.10과 같이 비교하였는데 서로 비슷한 값으로 산정되어 영상분류자료의 사용으로 CN값 산정의 타당함을 증명하였다. 또한 기존의 수작업으로 인한 주관적인 토지이용도의 사용으로 인한 오류를 어느 정도

줄일 수 있게 되었고 영상 분류자료의 사용으로 보다 신속한 CN값 산정이 기대된다.

- (3) 샘플지역의 CN값 비교와 대원사, 지리산 전체지역의 시험결과에서 알 수 있듯이 원격탐사와 신경망 기법을 이용한 토지피복도 분류 결과는 수문학적 지형자료로 적용가능하며 지리산유역 일대에서는 일정한 값을 나타내는데 이는 삼림지역의 분포가 타 분포 보다 크기 때문인 것으로 사료되고 차후 연구에서는 농경지역이나 도시지역의 적용으로 모형의 적절성을 시험할 예정이다.
- (4) 본 연구에선 신경망기법(RBFN)으로 다양한 유역의 토지 피복도를 분류하고 CN값을 산정함으로써 소유역 산악지역의 일반적인 CN값 기준을 제시하였다.

### 감사의 글

본 연구는 2001-2002년 한국과학재단 지역대학 우수과학자 지원연구비로 이루어진 것으로 본 연구를 가능케 한 당국에 감사드립니다.

### 참 고 문 헌

박희성 (1998). **위성영상을 이용한 토지피복 분류 및 유출곡선번호의 추정**. 석사학위논문, 서울대학교, pp. 30-40.

임상준, 박승우 (1997). "논의 유출곡선번호 추정." **한국수자원학회지**, 한국수자원학회, 제30권, 제 4호, pp. 379-387.

양인태 (1998). "랜드셋 자료로부터 유역의 유출계수 자동추출 기법 개발에 관한 연구." **대한토목학회지**, 대한토목학회, 제 18권, 제 III-4호, pp. 485-494.

오창석 (1996). **뉴로컴퓨터**. 지성출판사, pp. 67-92.

윤용남 (2002). **공업수문학**. 청문각, pp 145-149.

조홍제, 김광섭, 이충희 (2001). "Landsat 영상을 이용한 CN값 산정에 관한 연구." **한국수자원학회지**, 한국수자원학회, 제34권, 제 6호, pp. 735-743.

Christopher M. Bishop (1995). *Neural Networks for Pattern Recognition*. Clarendon Press, pp. 1 16-193.

John R. Jensen (1996). *Introductory Digital Image Processing; A Remote Sensing Perspective, 2nd edition*. Prentice Hall, pp. 197-231.

Lillesand, Kiefer (1994). *Remote Sensing and Image Interpretation 3rd Edition*. John Wiley & Sons Inc., pp. 585-647

Robert M. Rangan and Thomas J. Jackson (1980). "Runoff Synthesis Using Landsat and SCS Model." *Journal of the Hydraulics Division*, May, 1980. pp. 667-678.

Simon Haykin (1994). *Neural Networks; A Comprehensive Foundation*. Macmillan College Publishing Company, pp. 236-284.

(논문번호:02-97/접수:2002.12.23/심사완료:2003.06.25)