

## 氣象資料를 利用한 산불發生確率模型의 開發\*<sup>1</sup>

崔 寬<sup>2</sup> · 韓祥烈<sup>2</sup>

### Developing Forest Fire Occurrence Probability Model Using Meteorological Characteristics\*<sup>1</sup>

Kwan Choi<sup>2</sup>, Sang Yoel Han<sup>2</sup>

#### 要 約

본격적인 산림자원조성시대를 대비하기 위해서는 가장 큰 피해를 주고 있는 산불에 대한 과학적이고 합리적인 산불발생에 관한 예측모형의 개발이 필수적이다. 따라서 본 연구는 이에 대한 현실적인 방안으로서 대구·경상북도지역을 대상으로 하여 기상요인을 이용한 산불발생확률모형을 개발하고자 수행하였다. 이를 위해 먼저 산불발생일의 모든 기상자료들을 검토하여 이들 기상요인과 산화발생빈도와의 함수관계를 파악하여 의미 있는 기상요인을 규명하고, 이와 병행하여 局地的 次元의 기상자료 획득이 현실적으로 불가능하기 때문에, 각 시·군별 산불발생일의 유사정도를 적절한 통계적 기법에 이용하여 定量化하고 이를 토대로 조사대상지역 범주화하였다. 그 결과 산불발생에 영향을 미치는 기상요인으로는 상대습도, 일조시간, 강우후 경과일로 밝혀졌으며, 조사대상지역은 대구를 중심으로 한 중남부지역, 안동을 중심으로 한 북부지역, 포항을 중심으로 한 동부해안지역으로 구분되었다.

따라서 구분된 각 지역의 時系列 기상자료를 이용하여 logistic과 probit model을 기초로 한 산불발생확률모형이 개발되었다. 모형의 실효성을 검증하기 위하여 과거 기상자료를 대입한 결과 상당한 정도의 예측능력이 확인되어, 이를 이용한 효율적인 감시활동과 진화장비의 배치 등 산불예방활동의 효율성을 제고할 수 있을 것으로 기대된다.

#### ABSTRACT

Preparing the era of forest resources management requires studies on forest fire. This study attempted to develop forest fire occurrence model using meteorological characteristics for the practical purposes of forecasting forest fire danger rate. To accomplish this goal, the relationships between forest fire occurrence and meteorological characteristics are estimated. In the process, the forest fire occurrence pattern of the study region(Taegu-Kyungpook) is categorized by employing quantification IV method. The study region was divided into three areas such as, Taegu, Andong and Pohang area.

The meteorological variables emerged as affective to forest fire occurrence are relative humidity, longitude of sunshine, and duration of precipitation. To estimate the probability of forest fire danger, forest fire occurrence of three areas are regressed on the time series data of affective meteorological variables using logistic and probit model.

The effectiveness of the models estimated are tested and showed acceptable degree of goodness. Those models developed would be helpful to increase the efficiency of forest fire management such as detection of forest fire occurrence and effective disposition of forest fire fight equipments.

<sup>1</sup> 접수 1995년 6월 10일 Received on June 10, 1995.

<sup>2</sup> 경북대학교 농과대학 임학과 Department of Forestry, Kyungpook National University, Taegu 702-701, Korea.

\* 이 논문은 1993년도 교육부지원 한국학술진흥재단의 자유공모과제학술연구 조성비에 의하여 수행되었음.

Key words : forest fire, fire danger-rating, effective humidity, logistic model, probit model.

## 緒 論

육림기에 접어든 우리나라의 산림은 林內 可燃性物質의 증가에 따라 산불발생의 위험이 높아지고 있으며, 그 피해 또한 대형화되고 있는 추세이다. 산림청(1994) 보고에 의하면 1984년부터 1993년까지 10년간 전국에서 총 1,567건의 산불이 발생하였으며 피해면적은 10,141.63 ha, 피해액은 약 14억 4천만 원에 달한다. 이는 산림피해의 원인중 가장 큰 비중을 차지하는 것으로 본격적인 산림자원조성 시대를 대비하기 위해서는 합리적인 산불관리체계의 구축이 시급히 요청된다.

산불발생의 원인은 등산객, 성묘객, 무속행위자, 산림약초채취자 등 입산자의 부주의에 의한 실화가 가장 많으며 그 밖에 논·밭두렁 소각, 담뱃불, 군사훈련, 어린이 불장난 등에 의해 발생하고 있다. 산불은 일차적인 원인으로 다양한 인위적 요인에 의하여 시작되지만 이와 같은 실화가 산림에 피해를 줄 정도로 발화·연소·확산되기 위해서는 강우량, 풍속, 상대습도 및 林內 가연성물질의 습도 등 연소환경을 구성하는 기상적요인과 깊은 상관관계를 가지고 있다. 대구·경상북도지역에서 최근 12년간(1982-1993) 발생한 산불 346건중 건조한 봄철(3월-5월)에 발생한 산불이 243건으로 전체 산불발생의 70.2%를 차지하고 있다. 이는 봄철이 타계절보다 강우량이 적고 맑은 날이 많아 상대습도가 가장 낮은 시기이면서 산림에서는 새잎이 돋아나지 않고 건조한 낙엽이 많이 쌓여 林內 可燃性物質이 매우 건조하기 때문이다.

따라서 정확한 산불화재위험을 예측을 위해서는 인위적 요인과 기상적 요인을 함께 고려하는 것이 이상적이라 하겠다. 그러나 산불화재위험을 예측 모델을 설정함에 있어 인위적 요인을 대표할 수 있는 변수의 선정이 어려울 뿐 아니라, 이들 변수들의 크기를 매일 예측한다는 것은 사실상 불가능하다. 그리하여 각국에서도 기상적 요인과 산불과의 관계규명을 통하여 지역별 산불화재 위험도를 예측하는 실용적인 모델개발에 주력하여 체계적이고 합리적인 산불관리체계를 구축하여 왔다.

미국에서는 1940년대부터 林內 可燃性物質의

함수율을 중심으로 하는 모델이 개발되었다. 1972년 Deeming 등에 의하여 개발된 Fire-Danger Rating System은 산불진행에 따른 물리적 습성을 근거로 9개의 Fuel Model을 구분하고, 기본적인 몇 가지 구조를 중심으로 指數化하여 완성되었다. 1978년에는 NFDRS(National Fire Danger Rating System)에 건조상태를 나타내는 여러 지수를 개선하고 11개 이상의 Fuel Model을 추가하여 더욱 세분화하였고 인위적 요인과 자연적 요인에 의한 산불발생의 指數를 분리하여 현존하는 식생들의 수분함유상태에 대한 예측모형을 추가·개발하였다. 그 후 산불위험을 추정체계를 보완하는 많은 연구들이 수행되었는데, Haines 등(1978)은 NFDRS의 Spread Index를 보완하여 NFDRS의 유효성을 확인하였다. Boychuck(1987)는 산불화재의 鑛火를 위한 효율적인 인원 및 장비의 배치를 위해 Markov Chain Model을 개발하여, 모의시험을 통해 산불방제요원 및 장비의 기회비용을 30%까지 절감할 수 있다고 보고하였다.

일본에서는 松本由友(1951)가 기상과 산불발생과의 관계를 분석하였으며, 林野廳(1978)에서는 산불화재와 기상관계를 분석하여 지역별 발생상황도를 작성하였다.

우리나라에서는 정연하 등(1989)이 임업연구원 및 광릉지역에서 강우후 낙엽 및 연료봉을 이용하여 fuel type별 건조과정을 林內·外, 溫·濕度 변화와 함께 조사하였다. 최관, 한상열(1990)은 대구를 중심으로 산불화재예측모델을 線形確率模型으로 수식화하였는데, 산불발생에 영향을 주는 기상요인으로는 실효습도, 기온, 강우후경과일, 일조시간, 풍속 등으로 나타났다. 이시영(1990)은 산불연소 및 진행속도에 영향을 주는 주요 환경인자를 이용하여 실내 연소실험을 실시하여 각 인자별 연소온도 및 진행속도를 조사하여 산불피해 예측모델의 적용성을 제고하였다.

본 연구에서는 산불발생사례를 이용하여 지역적 산불발생의 특성을 數量化法을 통해 Grouping하여 조사지역을 구분하고, 구분된 각 特性群의 산불발생일과 비발생일의 기상자료를 분석하여 산불발생확률모형을 개발하였다. 개발된 모형의 실효성을 검증하기 위하여 과거의 산불발생사례와 기상자료를 토대로 모형의 유효성 검증도 함

게 수행되었다. 이들 모형을 이용한 산불발생의 위험을 예측은 산불의 예방, 조기발견 및 진화장비와 인원의 효율적인 배치를 통해 산불발생에 따른 시간적·경제적 비용을 절감하여 산불피해를 최소화할 수 있으며, 더 나아가 통합적인 산불관리체계의 이론적 기초로도 제공될 수 있을 것으로 기대된다.

**研究資料 및 方法**

**1. 研究資料**

최근 12년간(1982-1993) 대구시와 경상북도에 보고된 산불발생자료를 이용하여 산불발생일의 기상특성, 산불 발생빈도, 원인, 피해상황 등 일반적 사항을 파악하고, 산불발생일의 기상정보를 중앙기상대의 日氣象資料를 이용·분석하였다. 기상자료는 산불과 깊은 연관이 있으리라 짐작되는 상대습도, 기온, 풍속, 일조시간, 강우, 강우후 경과일 등의 자료들이 검토되었다.

**2. 研究方法**

과거 산불발생일의 기상정보를 수집하여 각각의 기상자료와 산불발생건수를 분석하여 그들의 관련성을 확인한 후, 여러 통계적 기법을 이용, 각각의 기상요인이 종합적으로 작용하여 산불발생에 어떠한 영향을 미치는가를 분석하여 산불발생위험예측의 정확성을 제고하고자 하였다.

조사기간의 산불발생사례를 월별, 발생시각별로 조사하여 산불발생의 季節變化, 日變化를 파악한 결과 규칙성이 발견되어 이에 영향을 줄 수 있다고 판단되는 기상인자와 산불발생건수와의 관련성을 回歸分析을 통해 규명하였다. 또한 대구·경상북도지역내의 각 시·군들간의 산불발생의 類似性(similarity)과 非類似性(dissimilarity)을 數量化法을 통해 조사지역을 구분하였다. 구분된 地域群의 대표적 시·군을 선정하여 時系列 기상자료(산불발생과 연관이 규명된 기상요인)와 산불발생의 有無를 logistic model과 probit model을 이용하여 확률모형을 개발하였다.

개발된 산불발생확률모형의 적합성과 실효성을 확인하기 위하여 모형에 포함된 기상인자와 산불발생자료들을 추정모형에 대입하여 조사기간 내일의 확률을 산출하여 실제 산불이 발생한 날과 산불이 발생하지 않은 날의 확률들을 비교·검토

하여 모형의 실효성과 적합정도를 확인하였다.

**산불과 氣象**

**1. 산불발생의 季節變化와 日變化**

대구·경상북도지역의 년도별 산불발생건수는 1982년부터 1993년까지 총 346건으로 1982년 16건에서 1993년 65건으로 점차 증가하고 있는 추세에 있다(그림 1). 12년간의 지역별 산불발생건수를 보면, 대구가 84건으로 전체 산불발생의 24.3%, 포항·영일군이 40건으로 11.6%, 달성군이 28건으로 8.1%, 안동시·군이 21건으로 6.1%의 순으로 나타났다. 월별 산불발생현황은 4월과 3월이 149건, 75건으로 전체 산불발생건수의 61.8%로 봄철에 가장 많이 발생하고 있으며(그림 2), 매 3년간의 월평균 산불발생건수를 파악해본 결과 매년 규칙적으로 4월과 3월에 산불이 집중적으로 발생하고 있다(그림 3).

이처럼 산불이 4월에 집중적으로 발생하는 원인을 규명하기 위하여, 1993년 대구, 안동, 포항, 울진 기상대의 월평균 기온, 상대습도, 누적강우량, 일조시간을 검토한 결과 월평균 기온의 경우 雨期인 6, 7, 8월을 제외하고 4, 5월의 기온이 15°C에서 20°C 범위로 비교적 따뜻한 기온

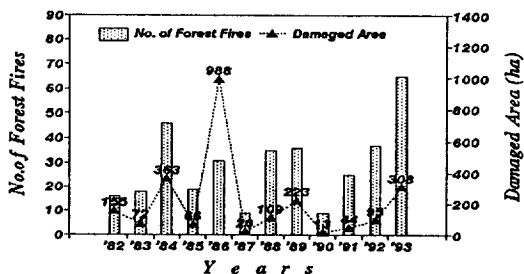


Fig. 1. No. of forest fire and damaged area in Taegu-Kyungpook.

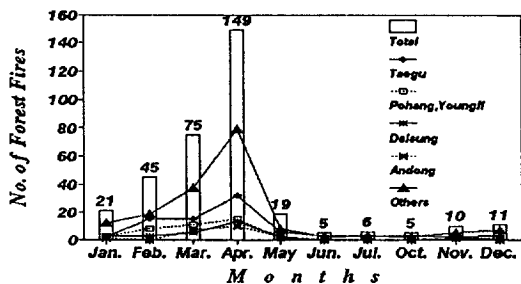


Fig. 2. No. of forest fires by month.

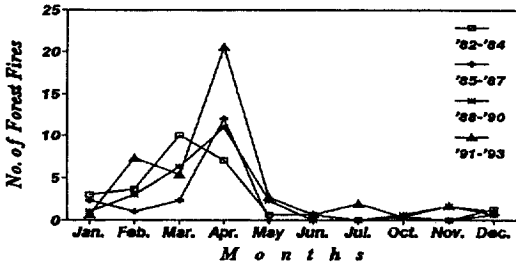


Fig. 3. No. of fires of each 3 years by month

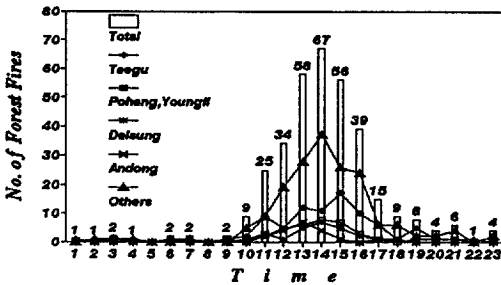


Fig. 4. No. of forest fires by occurrence time.

을 유지하는 시기이며, 상대습도와 누적강우량은 4월이 가장 낮았으며, 일조시간은 가장 길게 나타났다.

·산불발생의 日變化는 앞서 언급된 季節變化와 밀접한 관계가 있는 기상요소의 日變化가 매우 밀접한 관련을 가진다. 기상의 日變化는 오후 2시를 전후하여 기온이 가장 높이 올라가며 이와 대조적으로 습도는 가장 낮은 시간대를 이루고 있는 특징과 깊은 관련이 있다. 조사된 12년간의 지역별 산불발생의 자료를 검토하여 보면 오후 2시를 전후하여 산불이 가장 많이 발생하고 있음을 볼 수 있다. 즉 오전 12시에서 오후 4시까지의 산불발생건수는 254건으로 전체 산불발생건수의 73.4%를 점하고 있으며, 이러한 현상은 그림 4에서 보는 바와 같이 지역에 따라 큰 차이가 없는 것으로 나타났다.

2. 산불발생 원인

산불발생건수를 지역별·원인별로 살펴보면 그림 5와 같다. 대구·경상북도지역 산불의 경우 인위적 실화가 대부분인데 여기에는 등산객·성묘객·무속행위자·약초채취자 등에 의한 입산자 부주의에 의한 실화가 전체 산불의 45.4%를 차

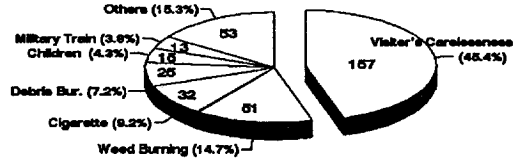


Fig. 5. Percentage of forest fire sources.

지하고, 논·밭두렁의 소각시 부주의에 의한 규반소각이 14.7%, 입산객의 담뭍불 부주의에 의한 실화가 9.2%, 폐비닐이나 농업용 쓰레기와 같은 農産廢棄物 소각시 부주의에 의한 실화가 7.2%, 어린이 불장난이 4.3%, 군훈련중 포탄이나 병사의 부주의에 의한 산불이 2.8% 발생한 것으로 나타났다. 이 밖에도 고압선 누전, 정신 이상자 방화, 인근지역 火焰에 의한 飛火 등에 의해 산불이 발생한 것으로 나타났다.

3. 산불발생빈도와 氣象과의 관계

산불발생과 기상과의 관계는 季節變化와 日變化를 통하여 그 관련성을 확인할 수 있어 산불발생에 관계하는 기상요소와 산불발생의 빈도와와의 관계를 규명하여 산불예방을 위한 기상정보의 이용 가능성을 고찰하였다.

이를 위해 대구·경상북도 전역의 산불발생지별 기상정보의 획득이 필수적이지만 각 시·군의 기상자료를 수집한다는 것은 현실적으로 불가능하므로 기상대가 설치되어 있는 대구, 포항, 안동, 울진을 중심으로 이들 지역의 산불발생일의 日氣象資料를 검토하였다.

산불이 발생한 당일의 기상과 산불발생의 빈도와의 구조적 관계를 밝히기 위하여는 먼저 적절한 함수의 형태(functional form)가 결정되어야 함으로, 본 연구에서는 從屬變數와 獨立變數사이에는 線形可能(linearizable)한 관계가 존재하는 것으로 가정하였다. 이러한 線形可能函數에는 獨立變數와 從屬變數가 非線形(nonlinear)의 형태를 취하더라도 推定母數에 대하여 線形的의 형태로 표현될 수 있는 모든 함수형태를 포함하였다.

1) 산불발생빈도와 상대습도

물질의 함수량과 燃燒의 관계는 공기에 포함되어 있는 水分의 정도를 나타내는 상대습도를 통하여 파악될 수 있다. 여기서 상대습도는 산불발생의 日變化에서 살펴본 바와 같이 하루중 상대습

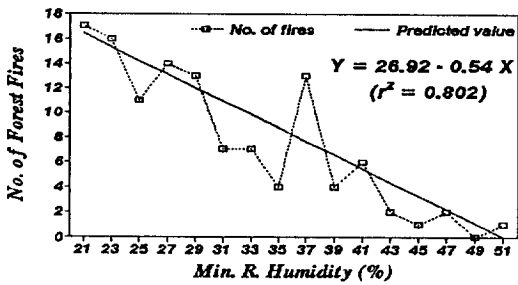


Fig. 6. The relation of no. of forest fire and min. R. Humidity.

도가 가장 낮은 시각인 오후 2시를 전후하여 주로 산불이 발생하였으므로 日日最小相對濕度를 통하여 검토하였다.

日日最小相對濕도와 산불발생빈도와의 관계를 추정한 결과 그림 6과 같이 나타났다. 日日最小相對濕도가 낮을 경우 산불발생의 빈도는 높았으며 이는 상대습도의 정도에 따라 산림내 可燃性物質의 건조정도가 결정되는 것으로 판단되어 진다.

2) 산불발생건수와 일조시간

기상요인중 습도가 可燃性燃料의 수분함유상태에 영향을 주는 요인이라고 한다면, 일조시간은 可燃性燃料의 온도를 결정하는 기상요인으로 생각할 수 있다. 일조시간이 긴 날은 산림내 可燃性物質의 표면온도를 상승시켜, 發火溫度에 보다 가까이 접근되며, 可燃性物質은 기온상승으로 인해 건조되면서 산불발생의 위험성을 상승시킨다. 이외에도 일조시간은 그 날의 기상이 맑고 흐림에 관계되는 요인으로, 맑은 날은 산림휴양객과 등산객이 흐린 날에 비해 보다 많이 증가되어 그만큼 산불발생빈도의 위험성이 높다할 수 있다. 이러한 두 가지 측면에서 일조시간과 산불발생빈도와의 관계를 그림 7에 나타내었다. 일조시간의 길어짐에 따라 산불발생건수가 증가하는 것을 확

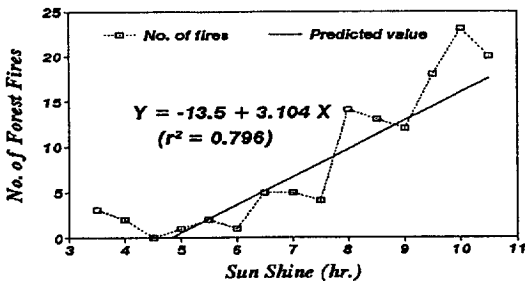


Fig. 7. The relation of no. of forest fire and sunshine.

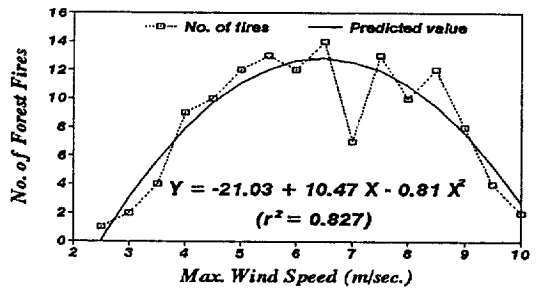


Fig. 8. The relation of no. of forest fire and max. wind speed.

인할 수 있었다.

3) 산불발생건수와 풍속

상대습도와 일조시간이 산불발생에 영향을 주는 기상요인이라 한다면, 풍속은 산불이 발생하였을 경우 산불확산에 영향을 주는 기상인자이다. 또한 풍속은 산림내 통풍에 관계하여 可燃性燃料의 水分含有와도 관계된다고 할 수 있다.

그림 8에서와 같이 日日最大風速과 산불발생빈도와의 관계에서 풍속이 증가함에 따라 산불발생빈도가 점점 증가하여 5m-8m에서 가장 높았으며 그 이상 풍속이 증가하면 산불발생의 빈도가 줄어드는 것으로 나타났다. 이는 상대적으로 풍속이 빠른 날이 적은 원인도 있겠으나, 일조시간이 그날의 날씨에 관계되는 것과 같이 풍속이 빠른 날은 야외휴양활동보다는 실내활동을 주로 하여 산림내로의 접근이 용이치 않아 그만큼 산불발생의 위험성이 적은 것으로 판단된다.

數量化 VI法을 利用한 市·郡地域의 範疇化

조사지역의 산불발생사례를 분석한 결과 산불발생은 기상요인과 매우 밀접한 관계를 가지고 있음을 확인할 수 있었다. 이러한 각 기상요인과 산불과의 관련성을 규명하고 여러 기상요소가 종합적으로 작용하여 산불발생에 미치는 영향을 평가하기 위해서는 산불발생지의 時系列 기상자료의 수집이 필요하다. 그러나 대구·경상북도지역의 산불발생지별 산불발생의 기상자료를 수집한다는 것은 현실적으로 불가능하므로 본 연구에서는 어떤 임의의 지역들에 있어 기상요소가 유사하다면 이들 지역의 산불발생의 出火日이 동일할 것이라는 전제하에, 기상요인의 類似程度를 Grouping하여 산불발생 위험 예측모형의 지역별

분류로 이용하고자 範疇化를 시도하였다.

각 시·군별로 산불발생일을 분류하여 동일한 날에 산불이 발생한 시·군을 짝지으면, 이들 사이의 산불발생에 관한 지역별 類似性을 정의할 수 있다. 이를 위해 기상조건에 대한 多次元 공간내의 축약된 표현과 그것에 의한 분류를 목표로 하는 數量化 IV法을 이용하여 분석에 적용하였다.

1. 分析方法

數量化 IV法은 類似性 또는 非類似性을 근거로 정식화하는 것이 보통이나 본 연구에서는 非類似性을 바탕으로 數量化 IV法을 구성하였다. 동일한 날 산불이 발생한 두 지역 i와 j 지역간의 類似性을 e<sub>ij</sub>, 非類似性을 d<sub>ij</sub>라고 두면, 한 지역 i에 數量化 값 x<sub>i</sub>를, 다른 지역인 j에 x<sub>j</sub>를 부여할 수 있으며(i, j=1, ..., n) 非類似性 d<sub>ij</sub>가 큰 경우에는 지역 i와 지역 j를 멀리 놓아야 한다는 것을 근거로 목적함수와 제약식을 아래와 같이 설정하였다.

목적함수: max ∑<sub>i</sub> ∑<sub>j</sub> d<sub>ij</sub> (x<sub>i</sub> - x<sub>j</sub>)<sup>2</sup>

제약식: ∑<sub>i</sub> x<sub>i</sub> = 0, ∑<sub>i</sub> x<sub>i</sub><sup>2</sup> = 1

목적함수와 제약식을 n×n 행렬 D=(d<sub>ij</sub>)와 n×1 벡터 x=(x<sub>1</sub>, ..., x<sub>n</sub>)'을 써서 표현하면 아래와 같다.

목적함수: max x' H x

제약식: 1'x=0, x'x=1,

여기서

H = diag (D1) + diag (1'D) - D - D'

따라서 최적화는 특성방정식 H x = λ x를 풀어 얻을 수 있다. 즉 λ와 x는 대칭행렬 H의 고유근과 고유벡터이며 특히 그때의 목적함수의 값이 λ가 되므로 제 1축의 數量化 값은 최대 고유근에 해당하는 고유벡터로부터 유도된다. 한편 제 1축의 數量化와 직교하면서 그 다음의 최대값을 갖는 제 2축의 數量化를 생각할 수 있는데 이것은 두번째 큰 고유근에 해당하는 고유벡터로부터 산출될 수 있다.

2. 研究資料

조사기간동안 발생한 346건의 산불발생사례중

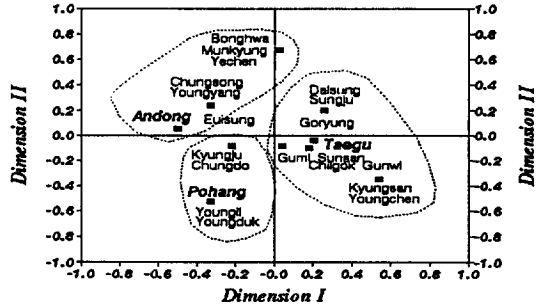


Fig. 9. The result of metric MDS by Quantification IV.

동일한 날에 발생한 각 시·군의 산불발생사례는 총 179건으로 이중 1 ha이상의 산불발생사례는 116건으로 집계되었다. 각 시·군을 隣近別로 묶어 총 10개의 시·군으로 재집계하여 동일한 날에 발생한 자료를 이용하여 산불발생의 친근성의 척도로 각 지역별 不一致個數 행렬을 이용하였다. 이때 불일치개수가 클수록 친근성이 낮은 것을 의미하기 때문에 각각의 불일치개수에 음의 부호를 붙여 불일치개수가 클수록 친근성이 높도록 재구성하여 數量化 IV法에 적용하였다.

3. 數量化 IV法의 分析結果

고유근과 고유벡터에서 고유치가 가장 큰 1, 2축을 이용하여 多次元的 정보량을 2차원으로 선택하여 각 지역별 고유벡터를 산출하였다(그림 9). 이를 기초로 하여 대구·경상북도지역을 크게 3 Group으로 나눌 수 있었다. 첫번째 Group은 대구를 중심으로 한 중남부지역이며, 두번째 Group은 안동을 중심으로 한 북부지역, 3번째 Group은 포항을 중심으로 한 동부해안지역으로 대별되었다.

따라서 다음에 논의될 산불발생확률모형의 개발은 대구·경상북도지역을 대구지역, 포항-영일지역, 안동지역으로 구분하여 각 地域群의 산불발생확률모형을 제시하였다.

氣象資料를 산불發生確率模型의 開發

1. logistic과 probit 모형의 이론적 배경

範疇화된 地域群으로 선정된 3지역의 時系列 기상자료(산불발생과 연관성이 규명된 기상요인)와 산불발생의 有無를 이용한 산불발생확률모형은 먼저 산불발생의 有無를 從屬變數로 두고 앞

에서 검증된 기상요인들을 獨立變數로 두었을 때, 회귀모형을 생각하면 식 (1)과 같다.

$$Y_j = \beta_0 + \beta_{1x}X_j + \epsilon_j \quad j=0, 1, 2, \dots, n \quad (1)$$

$Y_j=0$ (산불발생이 없는 경우) 또는  
 $1$ (산불발생이 있는 경우)

식 (1)에서  $E(\epsilon_j)=0$ 이라 하면 반응함수는 특별한 의미를 갖는다.

즉,

$$Y_j = \beta_0 + \beta_{1x}X_j \quad (2)$$

식 (2)에서 從屬變數  $Y_j$ 는 0 또는 1의 값을 가지므로 Bernoulli 確率變數가 된다. 따라서 식 (2)는 식(3)으로 표현할 수 있다.

$$E(Y_j) = 0 \cdot P[Y_j=0] + 1 \cdot P[Y_j=1] = P_j \quad (3)$$

반응함수  $E(Y_j)$ 는 獨立變數가  $X_j$ 의 값을 가질 때, 從屬變數가  $Y_j=1$ 이 되는 확률로 표현된다. 이처럼 從屬變數가 指示變數(dummy variable)로 나타내어지는 경우에 관하여 從屬變數와 獨立變數와의 관계를 線形 혹은 非線形으로 나타낼 수 있다. 이에 대한 분석방법으로 세 가지 과정이 도입되는데 선형관계를 가정한 線形確率模型(linear probability model)의 GLS(generalized least squares)와 線形關係가 아닌 非線形(non-liner)으로 나타나는 것을 가정하여 누적밀도함수가 logistic 분포로 구성된다고 하는 logistic model (LM), 정규확률누적분포에 입각한 probit model (PM)이 이용된다.

이들 방법중에서 LPM은 異分散(heteroskedasticity)의 문제와 0의 이하, 1이상의 확률이 나올 수 있으므로, logistic model(LM)과 probit model (PM)이 분석에 주로 이용된다.

logistic 모형은

$$\ln \left[ \frac{P}{1-P} \right] = Z_i$$

여기서  $P$  : 산불이 일어날 확률,  
 $1-P$  : 산불이 일어나지 않을 확률,  
 $Z_i$  :  $\beta_0 + \beta_{1x}X_j$ ,  $X_j$  는 조사당일의 기상변수

$\ln [P/(1-P)]=Z_i$ 의 양변에 자연지수(exponential)를 취하여 산불이 발생할 확률에 대하여 풀면 아래와 같다.

$$P = \frac{1}{1 + \exp(-z_i)}$$

이상의 logistic 함수의 특징은 반응함수  $E(Y_j)$ 의 인  $0 \leq E(Y_j) \leq 1$ 의 조건을 만족하며, 함수의 형태가 非線形이지만 이를 쉽게 線形으로 변환시킬 수 있는 장점이 있다.

probit 모형은 확률  $P$ 의 probit(probability unit)  $\Phi^{-1}(P)$ 를 말하는데, 이는 logistic 모형과 매우 유사하나, 누적밀도함수가 logistic 분포가 아닌 표준정규분포의 누적분포함수로 아래의 식과 같이 구성되어 진다.

$$P(Y=1 | X) = \phi(Z_i) = \int_{-\infty}^{z_i} \exp(-u^2/2) / \sqrt{2\pi} du$$

## 2. 推定結果

이상의 모형을 배경으로 logistic과 probit 모형의 從屬變數를 산불발생의 有無로 하고, 說明變數를 산불발생에 영향을 미치는 것으로 규명된 기상요인들로 설정하여 最尤推定方法(maximum likelihood estimation method)을 이용하여 분석을 실시하였다. 산불발생의 有無와 기상자료는 최근 3년간(1991-1993) 산불이 많이 발생하는 1월에서 5월까지로 설정하였으며, 분석대상지는 數量化 IV法の 결과에 따라 대구, 안동, 포항 3곳의 사례를 이용하였다.

설명변수중 습도에 관계하는 변수로는 상대습도의 누적치인 實效濕度를 이용하였는데 實效濕度는 當日과 前日들의 상대습도에 加重值를 붙여 얻어지는 평균습도로, 본 연구에서 이용한 實效濕度는 當日의 상대습도를 포함하여 4일간의 상대습도의 누적치를 이용하였으며, 實效濕度の 係數로는 0.7을 사용하였다.

즉,

$$He = \frac{H(0) + r(H1) + r^2(H2) + r^3(h3)}{1 + r + r^2 + r^3}$$

He : 實效濕度  
 H(0) : 當日の 평균 상대습도  
 H(n) : n일 전의 상대습도  
 r(=0.7) : 實效濕度 係數

규명된 기상자료들의 다양한 조합이 시도되어, 유의성이 없는 변수를 제거하여 최종모형에 實效濕度, 일조시간, 강우후 경과일의 함수로 대구, 안동, 포항의 산불발생확률모형이 개발되었다.

표 1에 나타난 대구지역의 산불발생확률모형에서는 3개의 설명변수 모두가 logistic, probit 모형에서 유의수준 10% 내에서 유의성이 있었으며, 모형의 적합도를 나타내는 pseudo R<sup>2</sup>의 값은 logistic 모형이 30.4, probit 모형이 30.3으로 두 모형이 거의 같은 것으로 나타났다. 포항-영일지역의 산불발생확률모형에서는 3개의 설명변수중 강우후 경과일이 유의수준 10% 내에서 유의성이 없어 이를 제거하여 實效濕度和 일조시간만을 이용한 최종 모형이 추정되었다. logistic과 probit 모형의 pseudo R<sup>2</sup>의 값은 각각 19.6으로 두 모형의 적합도가 거의 동일한 것으로 나타났다. 안동지역의 산불발생확률모형에서는 포항-영일지역과는 달리 일조시간이 유의수준 10% 내에서 유의성이 없어 이를 제거하여, 實效濕度和 강우후 경과일의 함수로 추정되었다. pseudo R<sup>2</sup>는 logistic 모형이 10.6, probit 모형이 10.5로 거의 유사하게 나타났다.

3지역에 대한 logistic과 probit 모형의 추정력은 거의 차이가 없는 것으로 나타났기 때문에 비교적 추정이 용이한 logistic model을 이용하여 최종 산불발생확률모형을 구축하였다.

## 結 論

앞에서 logistic 모형을 기초로 추정된 산불발생확률모형의 실효성검정을 위해서 모형에 포함된 과거 기상자료들을 대입하여 산불발생일의 발생확률과 산불이 발생하지 않은 날의 확률과의 비교·검토를 통하여 모형의 적합성 여부를 검증하여 보았다.

3년간(1991년-1993년) 산불이 많이 발생하는 1월에서 5월까지 대구, 안동, 포항지역의 實效濕度, 일조시간, 강우후 경과일을 최종모형에 대입하여 산불발생 확률을 산출하였다. 그 결과 대구지역에 있어서는 평균 산불발생확률이 0.07로 나타났으며, 실제로 산불이 발생한 33건에 대한 산불발생확률에 있어서는 평균치인 0.07이하에서 산불이 발생한 경우가 8건이며, 나머지 25건은 모두 평균산불발생확률 이상에서 산불이 발생하였다. 포항지역에서는 평균 산불발생확률이 0.037로 나타났으며, 실제로 산불이 발생한 17건에 대한 산불발생확률에 있어서는 평균치인 0.037이하에서 산불이 발생한 경우가 단 3건이며, 나머지 14건은 모두 평균산불발생확률 이상에서 산불이 발생하였다. 특히 산불발생확률의 범위가 0.27에서 0.33으로 높게 나타난 8일 가운데 실제로 산불이 4건이나 발생한 것으로 나타났다. 안동지역

**Table 1.** Forest fire occurrence probability models, parameter estimation for logistic and probit model.

Variable	Taegu		Pohang		Andong	
	logistic	probit	logistic	probit	logistic	probit
Intercept	-1.8924 (1.9623)	-1.1956 (1.0081)	-3.8724*** (2.2142)	-1.9999* (1.0284)	2.6146 (2.4552)	0.5208 (1.1500)
E. Humidity	-0.0699 <sup>b</sup> (0.0361)	-0.0335* (0.0183)	-0.1101*** (0.0369)	-0.0541** (0.0180)	-0.1663*** (0.0604)	-0.0649** (0.0268)
Sun Shine	0.2087** (0.0841)	0.1041** (0.0411)	0.5713*** (0.1914)	-0.2734** (0.0936)		
<sup>a</sup> No. of Days after Precip.	0.0520** (0.0237)	0.0290** (0.0127)	-	-	0.0526* (0.0305)	0.0207* (0.0138)
Model $\chi^2$	199.2	198.2	112.39	112.08	53.70	53.10
pseudo R <sup>2</sup>	30.4	30.3	19.6	19.6	10.6	10.5
N	455	455	457	457	454	454

Numbers in parentheses are asymptotic standard errors.

<sup>a</sup> No. of days after precipitation(0.6 mm ≤ Precip.)

<sup>b</sup> Single asterisk indicates significant at 10% level; double asterisk indicates significant at 5% level; three asterisk indicates significant at 1% level.



의 경우에는 평균 산불발생확률이 0.013으로 나타났으며, 실제로 산불이 발생한 6건에 대한 산불발생확률에 있어서는 평균치인 0.013 이상에서 모두 산불이 발생한 것으로 나타났다.

이상의 결과에서 매일 예보되는 기상정보를 이용하여 상당한 정도의 예측능력을 가진 지역별 산불발생확률의 추정이 가능하기 때문에 실용화 전망은 밝을 것으로 기대된다. 본 연구에 이용된 방법을 통하여 대구·경상북도지역 뿐만 아니라 전국적 규모의 기상정보 시스템을 이용한 산불발생예측모형으로의 확대가 가능하며, 이를 토대로 산불발생확률이 높게 나타나는 지역에 대한 강도 높은 감시활동과 진화장비의 사전 배치 등을 통하여 산불방지체제의 효율성을 제고할 수 있을 것으로 기대된다.

參 考 文 獻

1. 기상청. 1982-1993. 월별기상정보.
2. 남궁평·홍종천. 1994. 범주형자료의 통계분석. 자유아카데미.
3. 산림청. 1994. 임업통계연보(제 24호).
4. 송문섭·이영조·조신섭·김병천. 1993. SAS를 이용한 통계자료분석. 자유아카데미.
5. 이시영. 1990. 환경인자가 산불의 온도및 진행속도에 미치는 영향. 동국대학교 대학원 석사논문.
6. 임업연구원. 1993. 일본의 최근 임야화재의 경향과 대책. 월간임업정보 (30호).
7. 정연하·이시영·염육철·여운홍. 1989. 산불위험율예측에 관한 연구. 임업연구원보고 38:117-123.
8. 최 관·한상열. 1990. 산림화재 예측 모델의 개발. 한국임학회지 80(3).
9. 허명희. 1989. SAS 범주형 데이터 분석(고려대학교 통계연구소 통계분석 강의총서(5)). 자유아카데미.

10. 홍성길. 기상과 화재. 1987. 교학연구사.
11. 松本由友. 1951. 森林火災の發生と溫度の關係に就て. 日林誌 59: 179-181.
12. 田中 豊, 脇本和昌. 1992. 多變量統計解析法. 自由아카데미
13. Boychuk, D., and D.L. Martell. 1988. A Markov Chain Model for Evaluating Seasonal Fire Fighter Requirements. Forest Science 34.
14. Cunningham, A.A., and D.L. Martell. 1973. A Stochastic Model for the Occurrence of Man-Caused Forest Fires. Forest Science 29(4)
15. Cunningham, A.A., and D.L. Martell. 1976. The Use of Subjective Probability Assessments to Predicted Forest Fire Occurrence. Can. J. For. Res. 6(3).
16. Deeming, J.E., R.E. Burgan, and J.D. Cohen. 1971. The National Fire-Danger Rating System 1972. USDA Forest Serve Gen Tech Rep INT.
17. Deeming, J.E., R.E. Burgan, and J.D. Cohen. 1977. The National Fire-Danger Rating System 1978. USDA Forest Serve Gen Tech Rep INT.
18. Fried, J.S., and J.K. Gilless. 1988. Stochastic Representation of Fire Occurrence in a Wildland Fire Protection Planning Model for California. Forest Science 34.
19. Haines, D.A. 1983. Fire-Danger Rating and Wildfire Occurrence in the Northeastern United States. Forest Science 29(4).
20. Van Wager, C.E. 1974. Structure of the Canadian Forest Fire Weather Index. Can. For. Serve Pub. I 1333. Petawa Forest Exp Stn. Chalk River Ont..