

Article

기계학습 기반의 산불위험 중기예보 모델 개발

박수민 ^{1)*} · 손보경 ²⁾ · 임정호 ^{3)†} · 강유진 ²⁾ · 권춘근⁴⁾ · 김성용⁴⁾

Development of Mid-range Forecast Models of Forest Fire Risk Using Machine Learning

Sumin Park ^{1)*} · Bokyung Son ²⁾ · Jungho Im ^{3)†} ·
Yoojin Kang ²⁾ · Chungun Kwon⁴⁾ · Sungyong Kim⁴⁾

Abstract: It is crucial to provide forest fire risk forecast information to minimize forest fire-related losses. In this research, forecast models of forest fire risk at a mid-range (with lead times up to 7 days) scale were developed considering past, present and future conditions (i.e., forest fire risk, drought, and weather) through random forest machine learning over South Korea. The models were developed using weather forecast data from the Global Data Assessment and Prediction System, historical and current Fire Risk Index (FRI) information, and environmental factors (i.e., elevation, forest fire hazard index, and drought index). Three schemes were examined: scheme 1 using historical values of FRI and drought index, scheme 2 using historical values of FRI only, and scheme 3 using the temporal patterns of FRI and drought index. The models showed high accuracy (Pearson correlation coefficient >0.8, relative root mean square error <10%), regardless of the lead times, resulting in a good agreement with actual forest fire events. The use of the historical FRI itself as an input variable rather than the trend of the historical FRI produced more accurate results, regardless of the drought index used.

Key Words: Forest fire risk index, Machine learning, Mid-range forecast of forest fire risk

요약: 산불로 인한 피해를 최소화하기 위해서 산불위험 예보 정보를 제공하는 것은 필수적이다. 따라서 본 연구에서는 우리나라를 대상으로 기계학습 기반의 산불위험 중기예보(1일 후부터 7일 후까지) 모델을 개발

Received October 8, 2022; Revised October 17, 2022; Accepted October 20, 2022; Published online October 31, 2022

¹⁾ 한국항공우주연구원 위성활용부 선임연구원(Senior Researcher, Satellite Application Division, Korea Aerospace Research Institute, Daejeon, Republic of Korea)

²⁾ 울산과학기술원 도시환경공학과 석·박사과정생(Combined MS/PhD Student, Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea)

³⁾ 울산과학기술원 도시환경공학과 교수(Professor, Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea)

⁴⁾ 국립산림과학원 산림환경보전연구부 임업연구사(Researcher, Department of Forest Environment and Conservation, National Institute of Forest Science, Seoul, Republic of Korea)

† Corresponding Author: Jungho Im (ersgis@unist.ac.kr)

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

하였다. Global Data Assimilation and Prediction System (GDAPS)의 기상예보 자료와 기 개발된 산불위험지수 (Fire Risk Index, FRI)의 과거 및 현재 정보, 그리고 기타 환경요소(i.e., 고도, 산불다발지수, 가뭄지수)의 현재 정보를 반영하여 모델을 개발하였다. 본 연구에서는 실시간 학습을 통해 모델을 개발하였으며, 효율적인 모델 개발을 목적으로 과거 산불위험지수와 가뭄지수의 유무를 고려하여 세가지 경우(Scheme 1: 과거 산불위험지수 및 가뭄지수, Scheme 2: 과거 산불위험지수, Scheme 3: 과거 산불위험지수 변화 추세 및 가뭄지수)로 연구를 수행하였다. 본 연구에서 개발된 산불위험예보모델은 예보기간에 상관없이 높은 정확도(피어슨 상관계수 (Pearson correlation) >0.8, relative root mean square error <10%)를 나타냈으며, 실제 산불 발생 건에 대해서도 유의미한 결과를 보였다. 과거 산불위험지수의 추세보다는 산불위험지수 값 자체를 입력변수로 사용하는 것이 높은 정확도를 보였으며, 가뭄지수 사용과 관계없이 좋은 결과를 나타냈다.

1. 서론

산불은 주요 자연재해 중 하나로 기상, 식생, 지형환경에 따른 자연적 요인이나 입산자 실화와 같은 인위적 요인으로 인해 발생한다(Jolly *et al.*, 2019). 산불 발생은 해충이나 질병을 소멸시키고 더 나은 재생을 위한 영양소 제공 등의 긍정적인 영향도 있으나, 산불 발생의 빈도와 규모가 증가함에 따라 부정적인 영향(e.g., 호흡기 질환, 생태계 파괴, 기후변화 가속화)이 더 크게 작용할 수 있다(Abdollahi *et al.*, 2018; Farfán *et al.*, 2021; Yarragunta *et al.*, 2020). 산불로 인한 피해를 최소화하기 위해서는 산불 발생에 대한 선제적인 대응이 필요하며, 선제적인 대응을 위해서는 산불위험에 대한 예보정보가 필수적이다. 특히 산불위험 중기예보는 산불발생 시 필요한 진화 자원의 재배치와 산불 발생 후의 확산 정도, 피해 정도를 예측하는데 도움이 된다(Boychuk *et al.*, 2020).

이러한 이유로 국내 및 국외에서는 산불예보시스템을 통해 산불위험 정보를 제공하고 있다. 미국 산림청 (United States Forest Service, USFS; <https://www.fs.usda.gov/managing-land/fire>)에서는 기상인자, 가뭄지수, 연료 정보 등을 융합하여 산불위험에 대해 낮음, 중간, 높음, 매우 높음 등으로 예보하고 있다. 캐나다 산불정보시스템(Canadian Wildland Fire Information System, CWFIS; <https://cwfis.cfs.nrcan.gc.ca/home>)에서는 기상예보자료와 Initial Spread Index (ISI), Duff Moisture Code (DMC), Drought Code (DC), Buildup Index (BUI) 등을 활용하여 최대 14일까지 산불 위험 예측을 수행하고 있으며, 유럽에서는 캐나다와 동일한 방식으로 유럽 산불정보시스템(European Forest Fire Information System, EFFIS; <http://effis.jrc.ec.europa.eu>)를 통해 최대 9일 후까지의

산불 위험 예측정보를 제공하고 있다. 하지만 국내에서는 국가산불위험예보시스템(<http://forestfire.nifos.go.kr/main.action>)을 통해 기상청 동네예보에서 자료를 이용하여 48시간까지의 산불위험 정보만 제공하고 있다. 우리나라는 최근 소규모 산불 뿐만 아니라 대형 산불의 빈도 또한 증가함에 따라 산불로 인한 피해가 급증하고 있다. 산불발생에 대한 선제적 대응을 위해서는 단기 예보뿐만 아니라 중기(약 7일)예보 또한 필요하다.

산불위험 정보를 추출하기 위해 국내외적으로 많은 연구들이 수행되었다. 기상자료(e.g., 기온, 습도, 풍속)만을 이용하거나, 기상자료 외 추가적으로 위생기반의 식생지수를 이용하거나, 또는 가뭄관련 인자 및 인적 요소(e.g., 도로밀도, 인구수)를 추가적으로 이용(Gudmundsson *et al.*, 2014; Kang *et al.*, 2019; Kang *et al.*, 2020)하여 산불위험지수를 개발하였다. 개발에 사용되는 기법은 주로 일반 통계기법(e.g., 선형회귀, 로지스틱 회귀)이었으며(Preisler and Westerling, 2007; Gudmundsson *et al.*, 2014; Pham *et al.*, 2020), 최근 기계학습을 이용하여 산불위험지수를 개발하고 있다(Kang *et al.*, 2020). 이처럼 대다수의 연구가 산불위험지수를 개발하는 것으로 이루어지고 있으며, 산불위험예측에 대한 정보를 추출하는 것은 기상예보 자료를 기 개발된 산불위험지수에 입력하거나, 기상예보자료만을 이용하여 모델을 개발함으로써 산불위험예측을 수행하고 있다(Yan *et al.*, 2022). 하지만 산불위험예측에 있어 미래 기상 상태 뿐만 아니라 가까운 과거 및 현재의 환경 또한 중요한 요소이다. 따라서, 보다 정확한 산불위험예측을 위해서는 1) 기상자료 외 기타 환경인자를 사용하고, 2) 과거 및 현재 그리고 미래상태를 동시에 고려해야한다.

우리나라는 최근 산불 빈도가 증가하는 추세로, 산불

로 인한 피해 또한 증가하고 있다. 국내 산불위험예보 시스템은 기상자료 기반의 산불위험지수를 이용하여 미래 기상자료만을 사용하여 산불위험의 단계예측 (~48시간)에 대한 정보를 제공하고 있다. 우리나라는 입산자 실화와 같은 인적발화가 주 원인이므로 인적요소(e.g, 도로 밀도)를 고려한 산불위험지수를 사용해야 하며(Kim *et al.*, 2019; Kang *et al.*, 2020), 선제적 대응을 위한 예측 기간을 연장할 필요가 있다. 따라서 본 연구에서는 우리나라를 대상으로 Kang *et al.* (2020)이 개발한 산불위험지수의 시계열 정보와, 가뭄영향, 미래 기상정보 등을 함께 고려하여 산불위험 중기예측 모델을 개발하였다. 예측 모델은 1일에서 7일까지 일별 기준으로 각각 개발되었으며, 독립변수와 종속변수의 비선형적 관계를 고려하기 위해 기계학습 기법 중 하나인 랜덤포레스트(random forest)를 이용하였다.

2. 연구지역 및 자료

1) 연구지역

본 연구에서는 산림청(www.forest.go.kr)에서 제공하는 임상도를 기준으로 우리나라의 전 산림 지역을 연구 지역으로 선정하였다(Fig. 1). 우리나라에서는 최근 10년간 연평균 481건의 산불이 발생하여 연평균 1,087 ha가 소실되었으며 주원인은 입산자 실화(38%)이다. 계절별로는 봄철(45%)과 겨울철(43%)에 주로 산불이 발생하였고, 지역적으로는 경기도에서 산불 발생 빈도 수가 높게 나타났으며, 강원도 및 경상북도에서 산불로 인한 피해면적이 큰 것으로 나타났다.

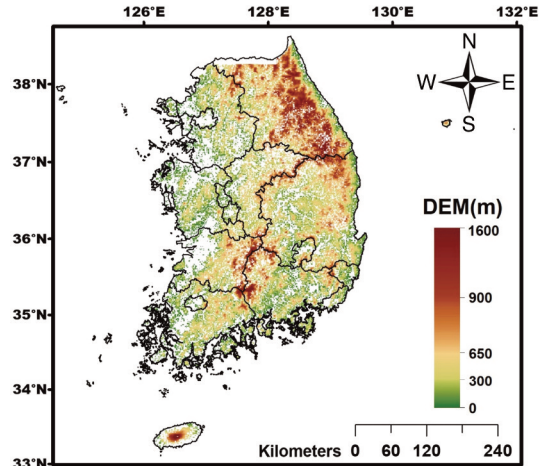


Fig. 1. The study area of this research. The Shuttle Radar Topography Mission (SRTM) 30 m digital elevation map (DEM; downloaded from <http://eros.usgs.gov/elevation-products>) is used as a background image in forest areas.

2) 산불위험지수

본 연구에서는 Kang *et al.* (2020)이 우리나라 지역을 기준으로 개발한 산불위험지수인 Fire Risk Index (FRI)를 이용하여 산불위험중기예보 모델을 개발하였다 (Table 1). FRI는 고위험 지역을 추출하기 위한 통합모델과, 기상요소만을 반영하는 기상 모델의 앙상블을 통해 산출된다. 통합모델은 시기를 나타내는 day of year (DOY), 임상도, 밀도, 고도, 도로밀도, 인구밀도, 풍속, 상대습도, 온도 등을, 기상모델은 풍속, 상대습도, 온도 등을 이용하여 인공지능 기법 중 하나인 Catboost를 통해 각각 산출된다. 각각 만들어진 통합모델과 기상모델을 8:2로 가중평균화하여 최종적으로 산불위험지수를

Table 1. Data used for developing mid-range forecast models of forest fire risk in this study

Data reference		Spatial resolution	Time resolution
FRI (Kang <i>et al.</i> , 2020)		1 km	hourly
SRTM DEM (http://eros.usgs.gov/elevation-products)		90 m	-
Forest fire hazard index (Source: National Institute of Forest Science)		-	-
Drought index (Park <i>et al.</i> , 2019)		1 km	daily
GDAPS (Source: Korea Meteorological Administration)	precipitation	10 km	daily
	land surface temperature		
	air temperature		
	relative humidity		
	wind speed		

산정한다. 개발된 산불위험지수는 낮음(0-0.5), 보통(0.51-0.65), 높음(0.66-0.85), 매우 높음(0.86-1)의 등급으로 산불 위험도를 나타낸다. 본 연구에서는 Kang *et al.* (2020)과 동일한 자료를 이용하여 5 km의 해상도의 산불위험지수를 산출하였으며, 산불위험중기예보 모델에서 입력변수 및 종속변수로 사용하였다.

3) 기타 환경인자

산불위험중기예보 모델 개발을 위해 고도, 산불 다발지수, 가뭄 지수 등을 입력변수로 이용하였다(Table 1). 고도는 Shuttle Radar Topography Mission (SRTM)의 Digital Elevation Model (DEM)을 5 km로 변경 후 이용하였다. 국립산림과학원에서 제공받은 산불다발지수(forest fire hazard index)는 1991년부터 2015년 기간에 발생한 산불을 수치 지도화한 것으로 산불이 빈번하게 발생할 수 있는 지역에 대한 정보를 나타낸다. 본 연구에서는 과거에 산불이 빈번히 발생했다는 것은 현재에도 발생할 가능성이 높다고 가정하고 5 km 격자로 재가공하여 입력변수로 사용하였다.

본 연구에서는 가뭄의 심도가 산불 발생에 영향을 끼칠 수 있다는 기존 연구와 현재 운영되고 있는 국외 산불위험예보시스템을 참고하여 Park *et al.* (2019)이 제안한 가뭄지수를 산불위험 중기예보 모델에 입력자료로 이용하였다. 가뭄지수는 우리나라를 대상으로 위성기반의 토양수분, Normalized Difference Water Index (NDWI), Precipitation Condition Index (PCI)를 결합한 것으로 실제 산불 발생과 관련성이 높다(식(1)). 가뭄지수 산정 시 사용한 자료는 기존 연구와 동일하게 적용하였으며, 최종적으로 산불위험지수와 같은 5 km 격자의 일별 형태로 변환하여 사용하였다.

$$drought\ index = 0.4 \times Soil\ moisture + 0.2 \times NDWI + 0.4 \times PCI \quad (1)$$

4) 수치예보모델 산출물

산불위험은 기상상황에 따라 변할 수 있으므로, 산불위험예보 모델 개발에 있어서 기상예보자료의 사용은 필수적이다. 본 연구에서는 산불위험 중기예보에 적합한 예측기간을 가지는 Global Data Assimilation and Prediction System (GDAPS) 산출물을 사용하였다(Table 1). 기상청에서는 2016년 7월 1일부터 약 10 km의 해상도,

한시간 간격으로 최대 10일까지의 예보장을 제공하고 있다. 본 연구에서는 강수량, 지표면 온도, 대기온도, 풍속, 상대습도를 이용하였으며, 예측 기간에 맞춰 강수량은 누적 강수량으로, 그 외 산출물은 평균값으로 산정하였다. 예를 들어, 3일 후의 산불위험 정도를 예측한다고 가정하면, 강수량은 당일로부터 1일 후부터 3일 후까지의 누적량을, 그 외 인자는 평균값을 사용하였다. GDAPS의 모든 산출물은 다른 입력변수와 동일하게 5 km의 격자로 변환되었다.

3. 연구방법

1) 연구방법

2016년 7월부터 2018년 12월까지의 산불위험지수, 고도, 산불다발지수, 가뭄지수, GDAPS 기상예보자료의 일별, 5 km 격자로 전처리 후 기계학습 기법 중 하나인 랜덤포레스트를 통해 산불위험 중기예보 모델을 개발하였다(Fig. 2). 본 연구에서는 실시간 훈련을 기반으로 모델이 매일 업데이트되는 방식으로 연구를 수행하였다. 이는 1) GDAPS가 주기적으로 업데이트되는 것을 고려하고, 2) 산불위험에 대한 최신 경향을 훈련자료로 사용하기 위함이다. 따라서, 당일로부터 30일 전까지의 입력자료를 훈련자료로 하여 당일로부터 1일 후부터 7일 후까지 예측하는 것으로 연구를 진행하였다(Fig. 3).

산불위험 중기예보 모델은 산불위험지수의 현재 및 과거정보(7일)와 기존 연구에서 언급되었던 가뭄의 영향을 반영하는 것을 기본으로 개발되었다(Scheme 1, Table 2). 하지만, 많은 입력변수를 사용하여 미미한 정확도 개선은 모델의 실시간 훈련에 비효율적으로 작용할 수 있다. 따라서 본 연구에서는 가뭄지수를 제외했을 경우(Scheme 2)와 산불위험지수의 현재부터 과거 7일까지의 경향(Scheme 3)으로 입력변수를 줄였을 경우에 대해 추가적으로 모델을 각각 개발하였다(Table 2).

랜덤포레스트는 여러 개의 의사결정나무를 앙상블하여 모델을 생성하는 기법으로, 각각의 의사결정나무는 동일한 확률 분포를 가지면서 동시에 독립적으로 구성되어있으며, 최종적으로 평균/표결(average/voting)을 통해 최종 결과물을 산출한다(Breiman, 2001; Sothe *et al.*, 2020). 학습자료를 임의로 선택하여 입력 변수의 조합을

Table 2. Three schemes for random forest models

Scheme	Input variables in random forest model		The number of input variable
	Common variable	Additional variable	
1	forest fire frequency index DEM	FRI (t, t-1, t-2, ..., t-6) drought index	15
2	GDAPS precipitation GDAPS land surface temperature GDAPS air temperature, GDAPS relative humidity	FRI (t, t-1, t-2, ..., t-6)	14
3	GDAPS wind-speed	Slope of FRI drought index	9

There are seven common variables and additional variables that vary by scheme. There are seven common variables in three schemes. GDAPS products were accumulated or averaged by forecasting period.

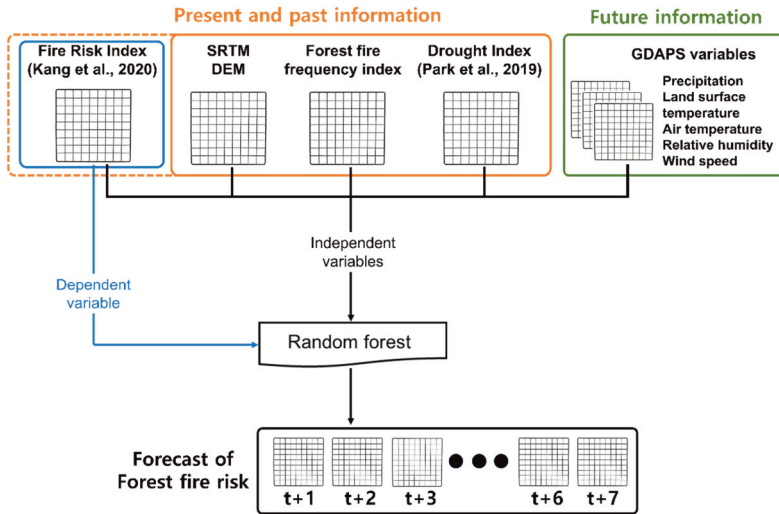


Fig. 2. The overall flowchart of developing mid-range forecast models of forest fire risk. The orange box represents present and past information and the green box indicates future information in input variables. "t" means the present day.

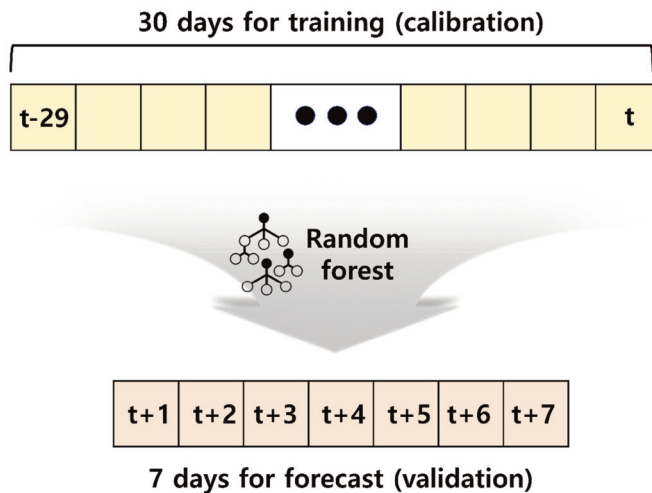


Fig. 3. Calibration and validation in random forest models. The yellow box means training data and the pink box indicates target days of forecasting with lead times. "t" means the present day.

임의 구성함으로써 이상치(outlier)의 영향을 최소화한다 (Sahoo *et al.*, 2020). 본 연구에서는 다른 인공지능 기법 (e.g, 서포트벡터머신, 인공신경망)에 비해 모델 파라미터의 개수가 적고, 정확도 및 시간으로 모두 고려했을 때 실시간 학습에 효율적이기 때문이다. 랜덤포레스트에서 가장 중요한 파라미터인 트리 개수를 최적화하기 위해 본 연구에서는 2016년 7월부터 2016년 12월까지의 자료를 통해 트리개수를 100개로 정하였다. 모델의 결과는 피어슨 상관계수(Pearson correlation)와, relative root mean square error (rRMSE)를 이용하여 정확도를 산정하였다.

4. 연구결과 및 토의

1) 모델 정확도 분석

각 Scheme 별로 개발된 모델의 정확도는 Fig. 4, 5와

같다. Fig. 4는 예보 기간에 따른 피어슨 상관계수를 나타낸 것으로 1에 가까울수록 모델의 성능이 우수함을 나타낸다. 예보 기간에 상관없이 대체로 0.8 이상의 값을 나타냈다. 예보 기간이 길어질수록 예보능력이 감소하는 것으로 나타났으며, 값의 범위가 점점 커지는 경향을 보였다(예보 기간이 길어질수록 상자 크기가 커지며, 범위 밖의 값 개수가 증가함). 이렇게 정확도가 감소하는 이유는 1) 모델 자체의 오류, 2) GDAPS의 예측능력이 예보 기간이 길어질수록 감소하는 것 때문으로 예상된다. 특히 강수량은 2일 혹은 3일이 지나면 현장관측 자료에 비해 크게 정확도가 감소하는 경향을 보였다. Scheme 1과 Scheme 2의 차이는 가뭄지수의 사용 여부에 따른 것으로 사용했을 때와 사용하지 않았을 때 Pearson 상관계수 값에 있어 큰 차이를 보이지 않는 것으로 나타났다. 이는 과거 기상상태가 산불위험지수에 포함되어 있어 가뭄지수의 영향력이 크게 나타나지 않

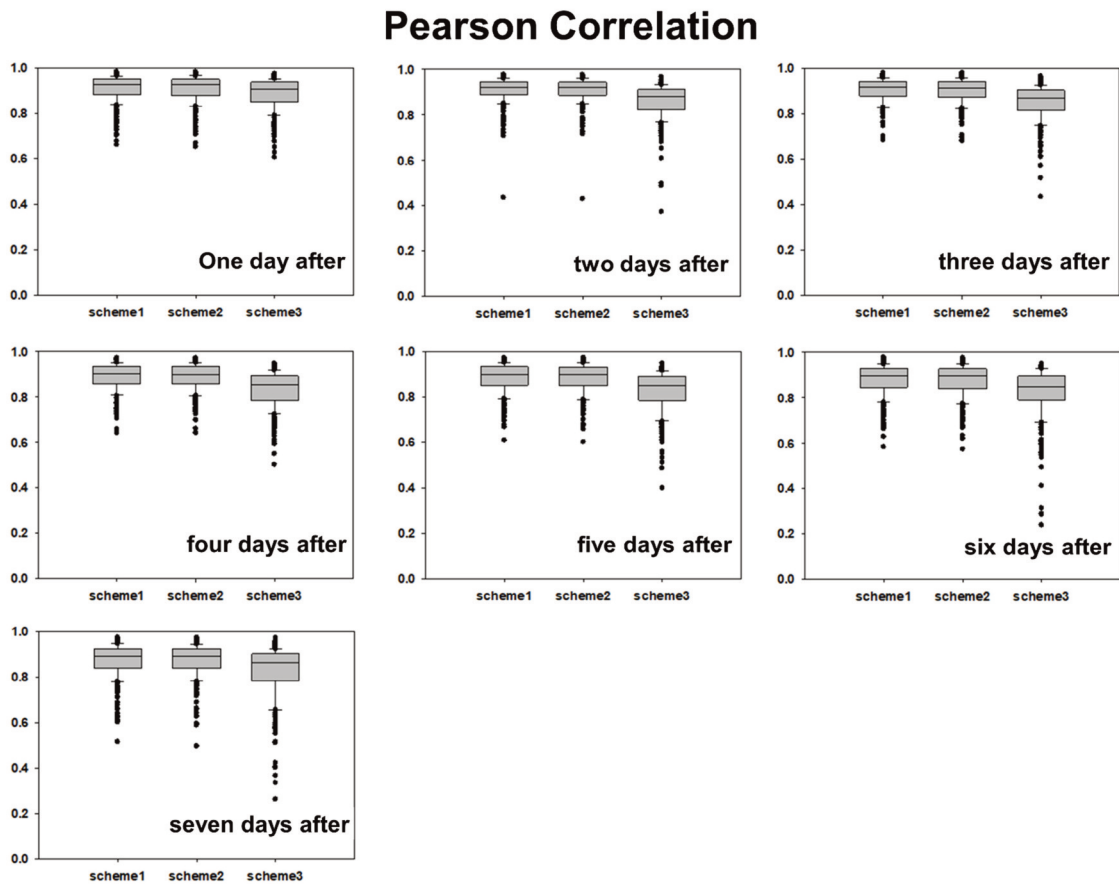


Fig. 4. Box plot of Pearson correlation coefficient values for each scheme. Refer to Table 2 for scheme information.

Relative root mean square error (rRMSE, %)

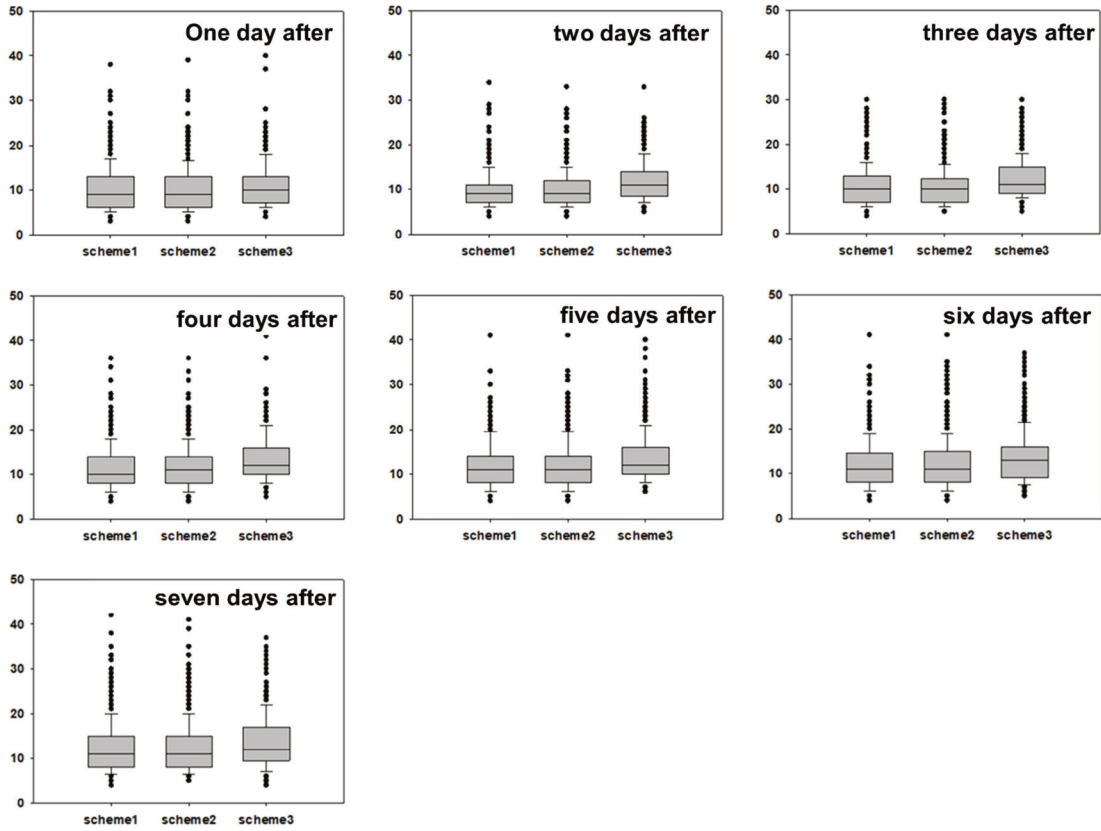


Fig. 5. Box plot of relative root mean square error (rRMSE) for each scheme. Refer to Table 2 for scheme information.

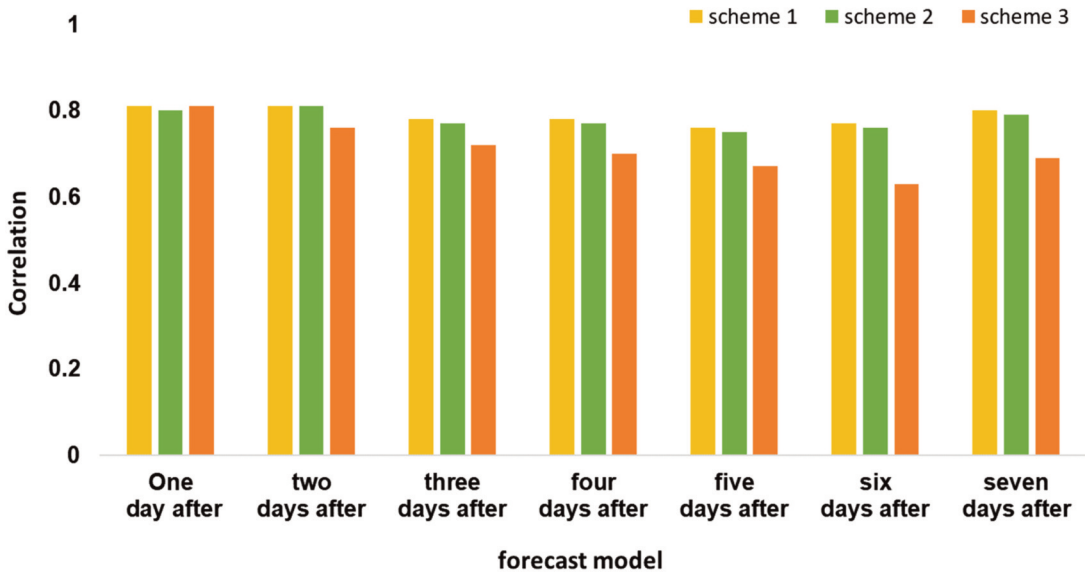


Fig. 6. The Pearson correlation for each scheme for forest fire events occurred in 2017. Yellow, green and orange bars represent schemes 1, 2, and 3, respectively.

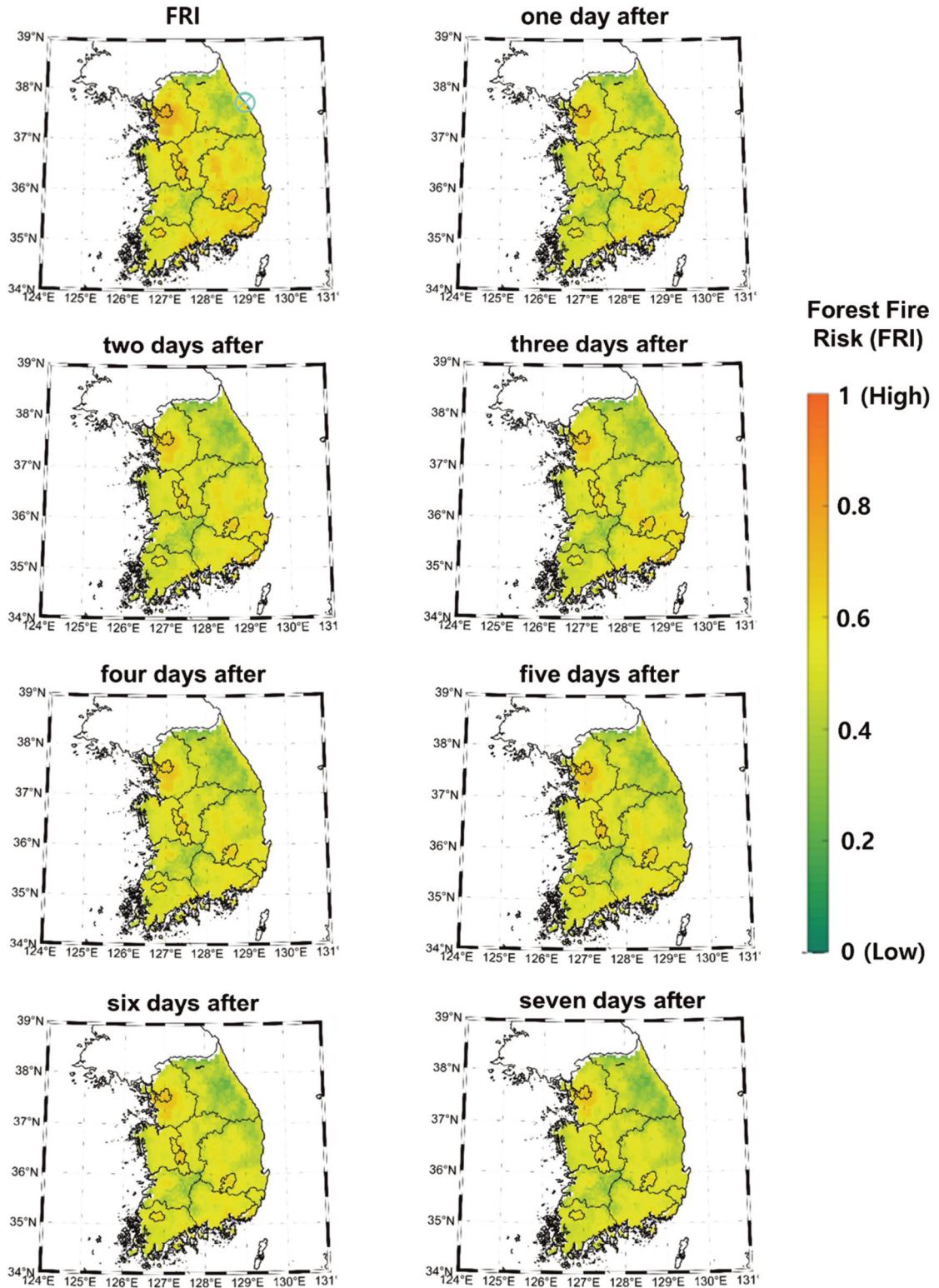


Fig. 7. Results of scheme 2 targeting March 10, 2017. A light blue circle on the left top image indicates the location of the forest fire event occurred on that day. The green to red color scheme represents low to high forest fire risk.

는 것으로 추측할 수 있다. Fig. 5는 rRMSE를 나타낸 것으로 Fig. 4와 비슷한 분포를 나타냈다. rRMSE는 평균적으로 10% 정도를 나타내었으며 크게는 50%의 값을 나타내기도 했다. 예보 기간이 길어질수록 rRMSE의 값이 커졌으며 Scheme 3의 경우 예보기간에 상관없이 가장 큰 rRMSE를 나타냈다. 이는 예측하고자 하는 날의 산불위험 정도가 과거 7일 동안의 산불위험지수 추세와 상이함에 따른 결과라고 말할 수 있다.

2) 산불 현장자료와의 비교

2017년 실제 산불 발생 302건에 대해 산불위험지수의 참값과 모델을 통해 예측된 값을 비교해보았다(Fig. 6). Scheme 1과 Scheme 2는 0.75-0.81의 상관계수 값을 나타냈으며 Scheme 3의 경우는 조금 낮은 0.63-0.81의 범위를 나타냈다. 앞선 결과와 마찬가지로 가뭄지수 유무에 따른 결과의 차이는 크지 않은 것으로 나타났다. 본 연구에서 사용한 가뭄지수는 산불 발생 건과 상관성이 높은 것으로 기존 연구에서 소개되었으나(Park *et al.*, 2019), 과거 산불위험지수 값으로 인해 모델에서 큰 영향력이 없는 것으로 추측할 수 있다. 모델의 정확도 및 효율성을 고려할 때, Scheme 2가 전체적으로 우수한 것으로 나타났으나, 실제 산불 발생 건에 대해 Scheme 1 또는 Scheme 3가 좋은 결과를 나타내기도 했다.

2017년 3월 9일 일부 지역의 산불위험이 “높음”으로 나타났으며, 강원도에서는 오전부터 10일 저녁까지 산불이 발생하여 160 ha의 면적을 소실하였다(Fig. 7). 각 Scheme 별로 rRMSE를 산정한 결과 Scheme 1에 대해서는 예측기간(1~7일) 순서대로 7, 9, 10, 10, 10, 12%를, Scheme 2에 대해서는 7, 10, 10, 10, 11, 11, 13%를(Fig. 7), Scheme 3에 대해서는 6, 9, 8, 6, 10, 7, 9%를 나타냈다. 단기 예측에서는 실제 산불위험지수와 비슷한 수치를 나타내었으나, 예측기간이 길어짐에 따라 값의 차이가 증가하는 경향을 보였다. 이는 앞서 언급했듯이, 예보기간이 길어짐에 따라 기상예측 정확도 감소와 모델 자체의 오류로 인해 발생하는 문제이다. 하지만 전체 공간적인 분포는 참값과 일치하는 경향을 보였다. 산불이 발생한 해당 지역의 실제 산불위험지수는 0.66으로 위험정도가 “높음”이었으며, 세 모델(*i.e.*, Scheme 1, Scheme 2, Scheme 3)에서 모두 0.6 정도의 값으로 나타났다. 앞서 전체적인 정확도 비교에서는 가뭄지수를 제외한 Scheme 2

가 상대적으로 우수한 결과를 나타냈으나 해당일에는 Scheme 3이 더 좋은 결과를 나타냈다. 이러한 결과는 3월에 증가하는 산불위험 추세가 반영된 이유라고 추측할 수 있다. 추후, 각 Scheme의 결과를 분석하여 통합된 모델을 개발할 필요가 있다.

5. 결론

본 연구에서는 산불위험지수의 과거 및 현재 정보와 미래 기상정보를 융합하여 우리나라를 대상으로 산불위험 중기예보 모델을 개발하였다. 인적발화가 주를 이루는 우리나라의 산불발생 특징을 이용하여 인적요소가 고려된 산불위험지수를 이용하였으며, 기타 환경요소(*i.e.*, 가뭄지수, 고도, 산불다발지수)도 이용하여 랜덤포레스트 기반의 모델을 개발하였다. 효율적인 모델 개발을 위해 가뭄지수 및 과거 산불위험지수를 변형하며 세 가지 경우(Scheme 1: 과거 산불위험지수 및 가뭄지수, Scheme 2: 과거 산불위험지수, Scheme 3: 과거 산불위험지수 변화 추세 및 가뭄지수)로 모델을 개발하였다. 상관계수 및 rRMSE 기반으로 개발된 모델은 모두 유의미한 정확도를 나타냈으며, 실제 산불 사건에 대해서도 좋은 결과를 나타냈다.

하지만, 본 연구에서 개발된 중기예보 모델은 일부 개선점이 필요하다. 첫째, 실제 산불 발생 건과 세 가지 Scheme에 따른 결과를 비교했을 때, 우수한 능력을 나타내는 Scheme이 상이하였다. 이는 과거 산불위험지수의 시계열 패턴과 미래 기상자료의 정확도에 따라 생기는 차이일 것으로 예상되며, 추후 Scheme의 앙상블 적용 혹은 입력변수의 조정 등을 통해 산불위험예보에 적합한 모델을 최종 선정할 수 있을 것으로 기대된다. 둘째, 예측된 산불위험지수의 낮은 해상도이다. 본 연구에서는 GDAPS의 낮은 공간해상도를 고려하여 5 km 해상도로 산불위험 예보 결과를 산출하였다. 하지만 우리나라는 10 ha 이하의 소규모 산불도 많이 발생하기 때문에 고해상도의 산불위험예보가 필요하다. 고해상도의 기상예보자료가 제공된다면 고해상도의 산불위험예보가 가능할 것으로 예상된다.

사사

이 연구는 본 과제는 행정안전부 재난안전 부처협력 기술개발사업(20009742), 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업(IITP-2022-2018-0-01424), 국립산림과학원의 ‘기상 빅데이터를 활용한 산불위험 통합 예보체계 구축(과제번호: FE 0500-2018-01-2022)’의 지원을 받았으며, 이에 감사드립니다.

References

Abdollahi, M., T. Islam, A. Gupta, and Q. K. Hassan, 2018. An advanced forest fire danger forecasting system: Integration of remote sensing and historical sources of ignition data, *Remote Sensing*, 10(6): 923. <https://doi.org/10.3390/rs10060923>

Breiman, L., 2001. Random forests, *Machine Learning*, 45(1): 5-32. <https://doi.org/10.1023/A:1010933404324>

Boyчук, D., C.B. McFayden, J. Evens, J. Shields, A. Stacey, D.G. Woolford, M. Wotton, D. Johnston, D. Leonard, and D. McLarty, 2020. Assembling and customizing multiple fire weather forecasts for burn probability and other fire management applications in Ontario, Canada, *Fire*, 3(2): 16. <https://doi.org/10.3390/fire3020016>

Farfán, M., C. Dominguez, A. Espinoza, A. Jaramillo, C. Alcántara, V. Maldonado, I. Tovar, and A. Flamenco, 2021. Forest fire probability under ENSO conditions in a semi-arid region: a case study in Guanajuato, *Environmental Monitoring and Assessment*, 193(10): 1-14. <https://doi.org/10.3390/fire3020016>

Gudmundsson, L., F.C. Rego, M. Rocha, and S. I. Seneviratne, 2014. Predicting above normal wildfire activity in southern Europe as a function of meteorological drought, *Environmental Research Letters*, 9(8): 084008. <http://dx.doi.org/10.1088/1748-9326/9/8/084008>

1748-9326/9/8/084008

Jolly, W.M., P.H. Freeborn, W.G. Page, and B.W. Butler, 2019. Severe fire danger index: A forecastable metric to inform firefighter and community wildfire risk management, *Fire*, 2(3): 47. <https://doi.org/10.3390/fire2030047>

Kang, Y.J., S.M. Park, E.N. Jang, J.H. Im, C.G. Kwon, and S.J. Lee, 2019. Spatio-temporal enhancement of forest fire risk index using weather forecast and satellite data in South Korea, *Journal of the Korean Association of Geographic Information Studies*, 22(4): 116-130 (in Korean with English abstract). <https://doi.org/10.11108/kagis.2019.22.4.116>

Kang, Y., E. Jang, J. Im, C. Kwon, and S. Kim, 2020. Developing a new hourly forest fire risk index based on catboost in South Korea, *Applied Sciences*, 10(22): 8213. <https://doi.org/10.3390/app10228213>

Kim, S.J., C.H. Lim, G.S. Kim, J. Lee, T. Geiger, O. Rahmati, Y. Son, and W. K. Lee, 2019. Multi-temporal analysis of forest fire probability using socio-economic and environmental variables, *Remote Sensing*, 11(1): 86. <https://doi.org/10.3390/rs11010086>

Park, S., B. Son, J. Im, J. Lee, B. Lee, and C. Kwon, 2019. Development of satellite-based drought indices for assessing wildfire risk, *Korean Journal of Remote Sensing*, 35(6-3): 1285-1298 (in Korean with English abstract). <https://doi.org/10.7780/kjrs.2019.35.6.3.11>

Pham, B.T., A. Jaafari, M. Avand, N. Al-Ansari, T. Dinh Du, H.P.H. Yen, T.V. Phong, D.H. Nguyen, D. Mafi-Gholami, I. Prakash, H. Thi Thuy, and T.T. Tuyen, 2020. Performance evaluation of machine learning methods for forest fire modeling and prediction, *Symmetry*, 12(6): 1022. <http://dx.doi.org/10.3390/sym12061022>

Preisler, H.K. and A.L. Westerling, 2007. Statistical model for forecasting monthly large wildfire events

- in western United States, *Journal of Applied Meteorology and Climatology*, 46(7): 1020-1030. <https://doi.org/10.1175/JAM2513.1>
- Sahoo, S., P.K. Gupta, and S.K. Srivastav, 2020. Inter-calibration of DMSP-OLS and SNPP-VIIRS-DNB annual nighttime light composites using machine learning, *GIScience & Remote Sensing*, 57(8): 1144-1165. <https://doi.org/10.1080/15481603.2020.1848323>
- Seong, N., M. Seo, K.-S. Lee, C. Lee, H. Kim, S. Choi, and K.-S. Han, 2015. A water stress evaluation over forest canopy using NDWI in Korean peninsula, *Korean Journal of Remote Sensing*, 31(2): 77-83 (in Korean with English abstract). <https://doi.org/10.7780/kjrs.2015.31.2.3>
- Sothe, C., C.M. De Almeida, M.B. Schimalski, L.E.C. La Rosa, J.D.B. Castro, R.Q. Feitosa, M. Dalponte, C.L. Lima, V. Liesenberg, G.T. Miyoshi, and A.M.G. Tommaselli, 2020. Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data, *GIScience & Remote Sensing*, 57(3): 369-394. <https://doi.org/10.1080/15481603.2020.1712102>
- Yan, W., J. Ren, J. Feng, Y. Duan, and C. Wei, 2022. A New Forest Fire Risk Rating Forecast Model Based on XGBoost, *Proc. of 2022 International Seminar on Computer Science and Engineering Technology (SCSET)*, Indianapolis, IN, USA, Jan. 8-9, pp. 227-230. <https://doi.org/10.1109/SCSET55041.2022.00060>
- Yarragunta, Y., S. Srivastava, D. Mitra, and H.C. Chandola, 2020. Influence of forest fire episodes on the distribution of gaseous air pollutants over Uttarakhand, India, *GIScience & Remote Sensing*, 57(2): 190-206. <https://doi.org/10.1080/15481603.2020.1712100>