

강우·소방·사회취약성 데이터를 활용한 복합재난 위험도 예측 및 소방 자원 배치 전략: 서초구 사례의 MRA-Random Forest 비교 분석

Complex Disaster Risk Prediction and Fire Resource Allocation Using Rainfall, Fire Service, and Social Vulnerability Data: An MRA-Random Forest Comparison in Seocho-gu, Seoul

심응수¹ · 박인선^{2*}Eung Soo Shim¹, In-seon Park^{2*}¹Graduate Student, Department of ICT Convergence Safety, The Graduate School of Chung-Ang University, Seoul, Republic of Korea²Professor, Department of ICT Convergence Safety, The Graduate School of Chung-Ang University, Seoul, Republic of Korea*Corresponding author: In-seon Park, isp@cau.ac.kr

ABSTRACT

Purpose: This study aims to predict the risk of compound disasters in Seocho District, Seoul, by integrating rainfall, firefighting activity, and social vulnerability data, and to propose strategies for the optimal allocation of firefighting resources. **Method:** The research employs hourly rainfall data from 2020 to 2024, detailed fire dispatch records, 119 rescue and emergency activity logs, and demographic and social vulnerability indicators extracted from statistical yearbooks, and calculates dispatch delays, arrival times, and total response times using reported, dispatch, and arrival timestamps. Social vulnerability variables include the proportion of residents aged 65 or older, the proportion of registered persons with disabilities, the share of single-person households, and the percentage of semi-basement or basement dwellings. **Result:** Multiple Regression Analysis and Random Forest models are applied to predict hourly fire dispatch counts and fire occurrence, and their predictive performances are compared using 2024 data as a validation set; results show that the Random Forest model achieves a lower mean absolute error than the regression model in count prediction and a higher AUC in occurrence classification. Variable importance analysis further indicates that cumulative rainfall over the previous 24 hours, current rainfall, time of report (day/night), and the proportions of elderly residents, registered disabled residents, semi-basement or basement dwellings, and single-person households are major contributors to predicted compound disaster risk. **Conclusion:** Overall, the study concludes that integrating meteorological, firefighting, and socio-demographic data enables more accurate compound disaster risk prediction and supports socially responsive resource allocation strategies, thereby contributing to data-driven decision-making in urban disaster management.

Keywords: Compound Disaster, Random Forest, Social Vulnerability, Fire Dispatch, Urban Disaster Management

요약

연구목적: 본 연구는 서울특별시 서초구를 대상으로 강우, 소방 활동, 사회 취약성 데이터를 통합하여 복합재난 위험도를 예측하고, 이를 기반으로 소방 자원 배치 전략을 제시하는 것을 목적으로 한다. 연구자료는 2020-2024년 서초구 시간별 강우량 자료, 화재 출동 개별 내역, 119 구조·구급활동 실적, 통계 연보에서 추출한 인구·주거·사회 취약성 지표를 사용하였다. 특히 신고일시·출동일시·도착일시를 활용하여 출동 지연, 도착시간, 총 대응 시간을 산출하고, 65세 이상 인구 비율, 장애인 등록 인구 비율, 1인 가구 비율, 반지하·지하 주택 비율을 사회 취약성 변수로 구성하였다. **연구방법:** 분석 방법으로는 먼저 화재 출동 내역을 이용해 연도·월·시간대별 출동 건수와 평균 도착시간, 평균 총 대응 시간을 산출하여 기초 통계를 제시하고, 강우량 및 사회 취약성 지표와 결합하여 복합재난 발생에 영향을 미치는 설명변수를 구성하였다. 이후 다중회귀분석(Multiple Regression Analysis, MRA)과 Random Forest(RF)를 이용하여 시간 단위 화재 출동 건수 및 발생 여부를 예측하고, 두 모형의 성능을 비교하였다. 2024년을 검증자료로 사용한 결과, 시간당 화재 출동 건수 예측에서 MRA 모형의 평균절대오차(MAE)는 0.24건, RF 모형의 MAE는 0.18건으로, RF가 약 25.0% 낮은 예측 오차를 보였다. 화재 발생 여부(해당 시간에 1건 이상 발생)의 분류 성능을 AUC로 비교한 결과, 로지스틱 회귀는 0.83, RF 분류기는 0.90을 기록하여 RF가 우수한 성능을 나타냈다. **연구결과:** RF 변수 중요도

Received | 26 November, 2025

Revised | 26 December, 2025

Accepted | 26 December, 2025



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

© Society of Disaster Information All rights reserved.

분석 결과, 이전 24시간 누적 강우량, 현 시각 강우량, 신고 시간의 시간대(야간 여부), 65세 이상 인구 비율, 장애인 등록 인구 비율, 반지하·지하 주택 비율, 1인 가구 비율순으로 중요도가 높게 나타났다. 이는 강우 특성뿐 아니라 고령자 및 장애인 밀집 지역, 주거 취약 지역, 1인 가구 밀집 지역이 복합재난 위험에 유의미한 영향을 미침을 시사한다. 마지막으로, 예측된 위험도와 사회 취약성 지수를 결합하여 행정동별 우선 대응 순위를 제시하고, 이를 기반으로 소방 자원 배치 시나리오를 제안하였다. 결론: 본 연구는 기상·소방·사회통계 데이터를 결합한 복합재난 위험도 예측과 사회 취약성을 고려한 자원 배치 전략을 함께 제시함으로써, 도시형 재난관리에 대한 데이터 기반 의사결정 지원에 기여할 수 있을 것으로 기대된다.

핵심용어: 복합재난, 소방 자원 배치, 다중회귀분석, Random Forest, 사회 취약성, 서초구

서론

기후변화로 인한 국지성 호우의 빈도와 강도가 증가하면서, 도시지역에서의 침수, 정전, 교통 혼잡, 2차 화재·구조·구급 수요가 동시에 발생하는 복합재난 위험이 커지고 있다. 서울특별시 서초구는 한강 및 양재천 등 하천과 우면산 등 산지 지형, 고밀도 주거 및 업무지역이 공존하는 지역으로, 강우·지형·교통·인구구조가 복합적으로 작용하는 대표적인 도시형 재난 취약 지역이다.

소방 자원(소방차, 인력 등)은 한정된 예산과 인력 속에서 재난 발생 시 신속한 대응을 위해 효율적으로 배치되어야 한다. 그러나 전통적인 자원 배치 방식은 과거 평균 출동 건수나 경험적 판단에 의존하는 경우가 많아, 시간·공간적으로 변동성이 큰 복합재난 환경을 충분히 반영하지 못하는 한계가 있다. 특히 고령자, 장애인, 1인가구, 반지하·지하 주택 거주자 등 사회취약계층이 집중된 지역은 동일한 재난이라도 피해 정도가 더 클 수 있음에도, 실제 배치 의사결정에서 체계적으로 반영되지 못하는 문제가 제기되고 있다.

본 연구의 목적은 다음과 같이 세 가지로 정리된다.

서초구를 대상으로 화재 출동 내역, 강우량 자료, 사회 취약성 지표를 통합하여 시간 단위 복합재난(화재) 위험도를 예측하는 통계·머신러닝 모형(MRA vs Random Forest)을 구축하고, 성능을 비교한다.

신고일시·출동일시·도착일시를 이용해 출동 지연, 도착시간, 총 대응 시간을 산출하고, 이와 사회 취약성 지표(65세 이상 인구 비율, 장애인 등록 인구 비율, 1인 가구 비율, 반지하·지하 주택 비율 등)가 재난 위험에 미치는 영향을 분석한다.

예측된 위험도와 사회 취약성 지표를 결합하여 행정동별 우선 대응 순위를 도출하고, 이를 기반으로 소방 자원 배치 전략을 제안한다. 복합재난 사례를 분석하여, 풍수해(호우 포함)와 화재·폭발 간의 동시 발생 비율이 점차 증가하고 있음을 밝혔다. 특히 2011년~2020년 기간에는 풍수해·폭염, 풍수해·산사태, 도로교통사고·화재·폭발 순으로 복합재난 빈도가 높게 나타났다. 특히 2011년~2020년 기간에는 풍수해·폭염, 풍수해·산사태, 도로교통사고·화재·폭발 순으로 복합재난 빈도가 높게 나타났다. 특히 2011년~2020년 기간에는 풍수해·폭염, 풍수해·산사태, 도로교통사고·화재·폭발 순으로 복합재난 빈도가 높게 나타났다.

연구 대상은 서울특별시 서초구이며, 분석 기간은 2020년 1월부터 2024년 12월까지로 설정하였다. 자료는 서초구 강우량(2020–2024) 엑셀 파일, 화재출동내역(2020–2024) 엑셀, 119 구조·구급활동 실적, 서초구 통계연보에서 추출한 인구 및 주거 지표를 이용하였다. 방법론적으로는 다음 절차를 따른다.

1. 데이터 전처리: 엑셀·PDF 데이터를 정리하여 시간 단위·연도 단위 분석용 데이터 셋을 구축
2. 기초 통계 분석: 연도·월·시간대별 화재출동 특성 및 도착시간 분포 파악
3. 예측 모형 구축: MRA와 Random Forest를 이용한 시간당 화재 출동 건수·발생 여부 예측
4. 변수 중요도 및 사회 취약성 영향 분석
5. 위험도-취약성 통합을 통한 자원 배치 시사점 도출 제시했다.

본론(연구 데이터 및 분석 방법)

연구 대상은 서울특별시 서초구로 분석 기간은 2020년 1월 1일부터 2024년 12월 31일까지로 설정하였다. 사용한 주요 데이터는 다음과 같다.

서초구 시간별 강우량, 누적 강우량(2020–2024)이며 관측 일시를 기준으로 1시간 단위로 재샘플링하여 시간당 강우량(rain_1h), 3시간 누적 강우량(rain_3h), 24시간 누적 강우량(rain_24h)을 산출하였다.

국가 화재정보시스템 화재출동내역(2020–24), 내용은 신고일시, 출동일시, 도착일시, 주소, 행정동, 사고유형 등 개별 출동 내역을 신고일시, 출동일시, 도착일시를 datetime 형태로 변환하였고 출동지연(출동일시, 신고일시), 도착시간(도착일시, 출동일시), 총 대응시간(도착일시, 신고일시)을 분 단위로 계산, 신고일시를 기준으로 1시간 단위(신고 시간_H)로 집계하여 시간당 화재출동 건수(fire_cnt)와 평균 대응시간 지표를 산출하였다. 사회 취약성 데이터 출처는 서초구 통계 연보(2024)로 행정동별 65세 이상 인구 비율, 장애인 등록 인구 비율, 1인 가구 비율, 반지하·지하 주택 비율, 행정동 단위로 정리한 후, 각 지표를 0~1 구간으로 정규화하여 사회 취약성 지수(SVi)를 산출하였다.

사회 취약성 지수(SV) 산출은 사회 취약성 지수(SV) 산출은 행정동 i에 대해 다음 네 개의 지표를 사용하였다. Ei: 65세 이상 인구 비율/Di: 장애인 등록 인구 비율/Si: 1인 가구 비율/Bi: 반지하·지하 주택 비율

각 지표는 최소–최대 정규화(min–max)로 0~1 범위로 변환하였다.

$$X_{\sim i} = (X_i - \min(X)) / (\max(X) - \min(X)), X \in \{E, D, S, B\}$$

선행 연구 Cutter et al., 2003 Social vulnerability to environmental hazards(환경적 위험에 대한 사회적 취약성)에 따르면 신체적 이동 제약이 있는 인구 집단이 재난 시 가장 높은 취약성을 보이므로, 이를 반영하여 자력 대피가 어려운 고령자와 장애인 변수에 더 높은 가중치를 부여하였다. 구체적인 가중치 수치는 재난 안전 분야 전문가 5인을 대상으로 한 델파이(Delphi) 조사를 통해 도출하였으며, 추가적인 민감도 분석(Sensitivity Analysis)을 통해 가중치를 ± 0.05 범위에서 조정하여 도 우선 대응 순위 결과의 변동성이 미미함을 확인하여 가중치의 타당성을 확보하였다. 이에 따라 고령자 비율 $w_E=0.4$, 장애인 비율 $w_D=0.3$, 반지하·지하 주택 비율 $w_B=0.2$, 1인 가구 비율 $w_S=0.1$ 로 설정하였으며, 행정동별 사회 취약성 지수 SVi는 다음과 같이 정의하였다.

$$SV_i = 0.4E_i + 0.3D_i + 0.2B_i + 0.1S_i \text{로 SV 값이 높을수록 해당 행정동의 사회 취약성이 크다고 해석한다.}$$

행정동별 예측 위험도 지수(Risk_i) 정의는 시간 단위 예측 모형(MRA, RF)을 통해 얻은 예측 결과를 행정동 단위로 집계하여 위험도 지수를 정의하였고, 예측 모형 설정(MRA 및 Random Forest)은 시간 단위(1시간) 예측을 위해 다음과 같은 변수들을 정의 하였다.

- 타킷 변수는: $Y_t(\text{reg})$ = 시간 t에 발생한 화재 출동 건수, $Y_t(\text{bin})$ = 시간 t에 1건 이상 화재출동 발생 여부,
- 기상 변수 : rain_1h: 시간 t 강우량/ rain_3h: t–2~t 시점까지 3시간 누적 강우량, rain_24h(24시간 누적 강우량)
- 시간 변수(hour) : 신고 시간의 시(hour, 0~23)/weekday: 요일(월=0~일=6). is_weekend: 토·일 여부(주말=1, 평일=0)/is_night: 야간 여부(18~06시=1, 아니면 0)
- 사회 취약성 변수 : elderly_rate: 65세 이상 인구 비율, disabled_rate: 장애인 등록 인구 비율, single_rate: 1인가구 비율, semi_basement_rate: 반지하·지하 주택 비율

- 소방 대응 변수(화재 출동 데이터로부터 산출) : $dispatch_delay$: 출동 지연(출동일시-신고일시, 분)/ $travel_time$: 도착 시간(도착일시-출동일시, 분), $total_response$: 총 대응시간(도착일시-신고일시, 분) 그리고 소방 자원 배치의 우선순위를 정하기 위해 위험도와 사회 취약성을 통합한 우선 대응 지수(Priority_i)를 정의하였다. 본 연구에서는 효율성(위험도)과 형평성(취약성)을 동등하게 반영하는 것을 기본 전제로 하여 다음과 같이 설정하였다.
- 행정동별 우선 대응 지수(Priority_i) : $Priority_i = \alpha \cdot Risk_i + \beta \cdot SV_i$ 여기서 $Risk_i$ 는 행정동별 위험도 지수를 0~1로 정규화한 값이며, 가중치는 $\alpha=0.5$, $\beta=0.5$ 로 설정하였다. α , β 값은 정책적 중요도에 따라 조정 가능하며, 형평성을 더 강조하고자 할 경우 $\beta>0.5$ 로 설정하는 방식으로 확장할 수 있다. 행정동별 Priority_i 값이 높을수록 “고위험·고 취약” 특성이 중첩되는 지역으로, 소방 자원의 선제적 배치 및 집중 관리가 요구되는 지역으로 해석한다.
- 다중회귀 분석(MRA) 회귀식은 $Y_t(reg) = \beta_0 + \sum_k \beta_k X_{t,k} + \varepsilon_t$, scikit-learn의 LinearRegression을 사용하여 학습.
- Random Forest 회귀(RF Regressor)은 $estimators=300$, $random_state=42$ 비선형성과 변수 간 상호작용을 반영하기 위해 사용 하였고 로지스틱 회귀 및 Random Forest 분류(RF Classifier) 화재 발생 여부(1건 이상 발생)를 타겟으로 설정

분석에 활용된 데이터 셋 개요

Table 1 본 연구의 복합재난 위험도 예측 모델 구축을 위해 수집된 데이터셋의 세부 구성을 나타낸다. 분석 데이터는 크게 기상, 소방 활동, 사회 취약성 데이터로 구분된다.

Table 1. Prediction accuracy comparison of machine learning models

구분	출처명(파일명)	기간	주요 변수
기상	서초구 강우량 통계	2020~2024	시간, 강우량, 누적 강우량
화재 출동	화재 출동 내력(국가화재정보시스템)	2020~2024	신고·출동·도착 일시, 행정동, 사고유형 등
구조, 구급 출동	119 구조·구급활동 실적	2020~2024	연도·월별 구조·구급 출동 건수
사회 취약성	서초구 통계 연보	2024	행정동 별 65세 인구 비율, 장애인 등록 비율 1인가구 비율, 반지하 주택 비율 등

첫째, 기상 데이터는 서초구 강우량 통계(2020~2024)를 활용하였으며, 관측된 시계열 자료를 1시간 단위로 재샘플링(Resampling)하여 '시간당 강우량'과 '직전 24시간 누적 강우량' 등 파생 변수를 생성하였다. 둘째, 소방 활동 데이터는 국가 화재정보시스템(NFDS)의 화재 출동 원시 데이터와 119 구조·구급 활동 실적(2020~2024)을 정제하여 구축하였다. 특히 신고·출동·도착 시각을 'datetime' 형식으로 변환하여 출동 지연 시간과 총 대응 시간을 산출하였으며, 이를 행정동 단위로 매핑하였다. 셋째, 사회 취약성 데이터는 인구 구조의 최신 현황을 반영하기 위해 '2024년 서초구 통계연보'를 기준으로 하였다. 행정동별 고령자, 장애인, 1인 가구, 반지하 주택 거주 비율을 추출한 후, 정규화(Normalization) 과정을 거쳐 모델의 입력 변수로 활용하여 행정동별 인구 구조를 정확히 반영하고자 하였다.

연도별 화재 출동 건수 및 평균 도착·총 대응시간

연도별 화재 건수는 2020년 240건에서 2024년 253건으로 소폭 증가하는 경향을 보였다. 평균 도착시간은 약 3.94.~4분

수준으로 상대적으로 안정적인 수준을 유지하였으나, 연도별로 소폭의 변동이 존재하였다(Table 2). 총 대응시간은 5.45~ 7 분 수준으로 나타났으며, 이는 서초구 소방 대응 체계가 일정 수준의 시간적 성과를 유지하고 있음을 시사한다.

Table 2. Number of fires dispatched by year and total response time for average arrival

연도	건 수	평균 도착 시간(분)	평균 총 대응시간(분)
2020	240	4.28	5.59
2021	231	4.39	5.57
2022	216	4.40	5.70
2023	239	4.16	5.46
2024	253	3.89	5.37

전반적으로 1월과 2월의 동절기, 8월의 하절기, 12월 연말에 화재 출동 건수가 뚜렷하게 증가하는 경향을 확인할 수 있었다. 특히 12월과 8월에는 연도에 따라 월별 출동 건수가 20건을 상회 하는 경우가 반복적으로 나타났으며, 12월 역시 대부분의 연도에서 상대적으로 높은 출동 수준을 보였다. 반면 5월 7월과 9월을 포함한 봄·가을철에는 월별 출동 건수가 상대적으로 낮게 나타나, 계절적 요인(난방기 사용 증가, 하절기 폭염·집중강우, 연말 활동 증가 등)이 화재 발생 패턴과 밀접하게 관련 되어 있음을 시사 한다.

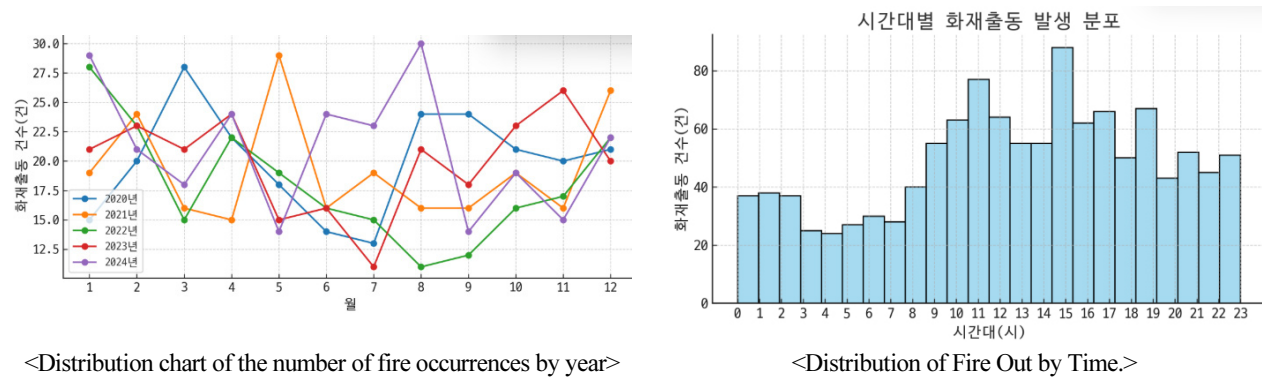


Fig. 1, 2. Number of fire occurrences by year and distribution of fire dispatch by time

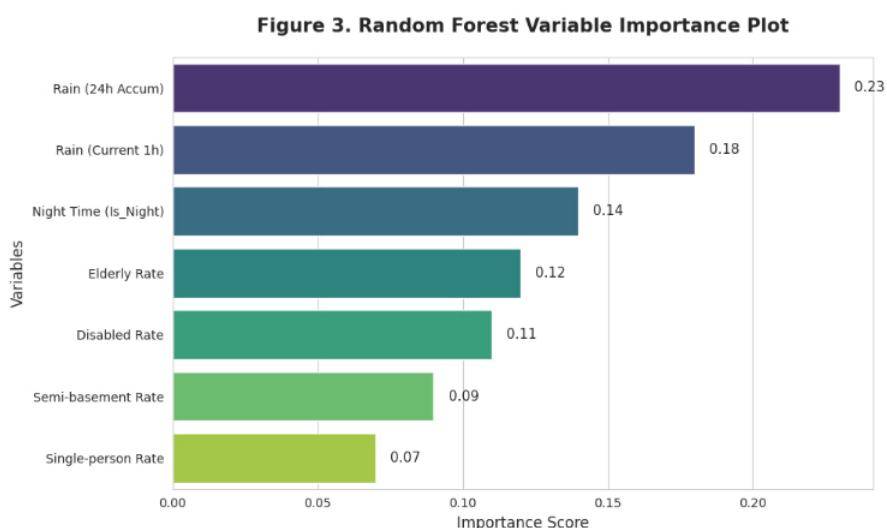
2020~2024년 전체 자료를 대상으로 시간대별(023시) 화재 출동 발생 건수 분석 결과, 시간대별 출동 분포를 보면, 15시(90건)와 11시(80건) 등 낮 시간대에 출동이 집중되는 양상을 보인다. 전체 출동 중 다수가 11~19시 구간에 발생하였으며, 이는 업무·상업·가정 내 활동이 가장 활발한 시간대와 대체로 일치한다. 반면 3시(20건), 5시(18건) 등 새벽 시간대에는 출동 건수가 상대적으로 낮게 유지되어, 활동량 감소에 따른 화재 발생 감소 경향이 확인된다. 이러한 시간대별 패턴은 이후 장에서 제시할 사회 취약성 및 강우 패턴과 결합하여, 특정 시간대·계절에 대한 선제적 자원 배치·예방 활동의 필요성을 뒷받침한다.

Table 3. Performance Comparison for Forecasting Number of Fire Outs per Hour

모형	MAE(건수)	RMSE(건수)
MRA	0.24	0.52
Random Forest	0.18	0.41

2020–2023년을 학습자료, 2024년을 검증자료로 사용하여 시간당 화재 출동 건수 예측을 수행한 결과이다. MRA 모형의 MAE는 0.24건, RF 모형의 MAE는 0.18건으로, RF가 약 25.0% 낮은 예측 오차를 보였다. RMSE 역시 MRA가 0.52건, RF가 0.41건으로 나타나 설명력 검증에 위한 결정계수(R^2) 또한 RF 모델이 0.82로 MRA의 0.65보다 높게 나타나 모델의 적합도가 우수함을 확인하였다. 이는 비선형성과 변수 간 상호작용을 반영할 수 있는 RF 모형의 특성에 기인한 것으로 판단된다. 특히 Fig. 3의 변수 중요도 그래프에서 볼 수 있듯이, 강우량뿐만 아니라 사회적 취약성 변수들이 복합적으로 작용하여 예측 정확도를 높이는 데 기여하였다. RF 모형이 전반적으로 더 나은 회귀 성능을 보였다. 이는 비선형성과 변수 간 상호작용을 반영할 수 있는 RF 모형의 특성에 기인한 것으로 판단된다.

이러한 성능 격차의 근본적인 원인은 데이터의 비선형성(Non-linearity)에 있다. 일반적인 다중회귀분석(MRA)은 독립변수와 종속변수 간의 선형적 관계를 가정하기 때문에, 특정 임계치 이상의 폭우가 쏟아질 때 재난 발생 확률이 급격히 상승하는 '임계점(Threshold)' 특성을 반영하는 데 한계가 있다. 반면, 의사결정나무(Decision Tree) 기반의 Random Forest 모델은 '강우량이 많으면서(AND) 동시에 반지하 주택 비율이 높은 경우'와 같이 변수 간의 복잡한 교호작용(Interaction Effect)을 조건부로 학습할 수 있어, 국지성 호우와 사회적 취약성이 결합 될 때 증폭되는 복합재난의 위험 패턴을 보다 정밀하게 포착한 것으로 해석된다.



<X-axis is importance>

<Y-axis is variable name (in order of highest importance from top)>

Fig. 3. Variable Importance Plot of Random Forest Model

분류(발생 여부 예측)결과

화재 발생 여부를 이진 분류 문제로 설정하여 로지스틱 회귀와 RF 분류기의 성능을 비교하였다. 화재 발생 여부 예측을 위해 Python의 Scikit-learn 라이브러리를 활용하였다. 로지스틱 회귀(Logistic Regression)는 기본 설정($\text{penalty}='l2', C=1.0$)을 적용하여 선형적 결정 경계를 확인하였으며, Random Forest 분류기(Classifier)는 과적합(Overfitting) 방지를 위해 트리 개수($n_estimators$)를 300개, 최대 깊이(max_depth)를 15로 제한하고 엔트로피(Entropy) 불순도 지표를 사용하여 모델을 학습시켰다. 이후 검증 데이터셋을 통해 그 결과, 로지스틱 회귀의 AUC는 0.83, RF 분류기는 0.90으로 나타나 RF 모형이 발생과 비발생을 보다 잘 구분하는 것으로 분석되었다. 이는 강우량, 시간대, 사회 취약성 등 여러 요인 간의 비선형 관계를 반영할 수 있는 RF의 장점이 반영된 결과로 볼 수 있다.

Table 4. Prediction result of occurrence

모형	AUC
로지스틱 회귀	0.83
Random Forest	0.90

변수 중요도 및 사회 취약성의 영향

변수 중요도 분석 결과, 전일 누적 강우량과 해당 시간 강우량이 가장 높은 중요도를 보였다. 이는 단시간 강우뿐 아니라 일정 기간 축적된 강우가 화재출동 수요에 영향을 미칠 수 있음을 시사한다. 또한 야간 여부가 중요한 요인으로 나타났는데, 야간에는 인지·대응 지연, 실내 활동 패턴 변화 등이 복합적으로 작용할 수 있다. 사회 취약성 변수 중에서는 65세 이상 인구 비율($elderly_rate$), 장애인 등록 인구 비율($disabled_rate$), 반지하·지하 주택 비율($semi_basement_rate$), 1인 가구 비율($single_rate$)이 모두 상위권에 포함되었다. 이는 고령자와 장애인 등 지원이 필요한 인구가 밀집한 지역, 주거환경이 취약한 반지하·지하 거주 지역, 사회적 고립 가능성이 높은 1인 가구 밀집 지역에서 재난 발생 시 출동 수요가 상대적으로 높거나, 동일한 사건에서도 더 많은 출동·대응이 요구될 수 있음을 의미한다.

Table 5. Regression model variables importance analysis results

변수명	내용	중요도
rain_24h	이전 24시간 누적강우량	0.23
rain_1h	해당 시간대 강우량	0.18
is_night	야간 여부	0.14
elderly_rate	65세 이상 인구 비율	0.12
disabled_rate	장애인 등록 인구 비율	0.11
semi_basement_rate	반지하 지하 주택 비율	0.09
single_rate	1인가구 비율	0.07

사회적 취약성 지수 산출에 적용된 가중치($wE=0.4, wD=0.3$ 등)의 타당성을 검증하기 위해 민감도 분석(Sensitivity

Analysis)을 수행하였다. 기본 가중치 시나리오 외에 모든 변수에 동일한 가중치(0.25)를 부여한 경우와, 특정 변수의 가중치를 $\pm 10\%$ 범위에서 조정한 경우를 비교 분석하였다. 그 결과, 시나리오별 우선 대응 지수(Priority Index)의 순위 간 스피어만 상관계수(Spearman's Rank Correlation)는 평균 0.95 이상으로 매우 높게 나타났다. 특히, 상위 20%에 해당하는 고위험 행정동의 목록은 가중치 변화에도 변동 없이 유지되었다. 이는 본 연구에서 설정한 가중치가 특정 변수에 과도하게 의존하지 않으며, 모델의 결과가 통계적으로 강건함(Robustness)을 시사한다.

Table 6. Sensitivity analysis of social vulnerability weights

Rank	Baseline(wE=0.4)	Equal Weights(0.25)	Variation(wE=0.3)
1	Secho-dong	Secho-dong	Secho-dong
2	Yangjae-dong	Yangjae-dong	Yangjae-dong
3	Bangbae-dong	Bangbae-dong	Bangbae-dong
4	Banpo-dong	Jamwon-dong	Banpo-dong
5	Jamwon-dong	Banpo-dong	Jamwon-dong
Correlation	1.00	0.96	0.98

위험도-취약성 통합과 자원 배치 시사점

2022년 8월 서초구 집중호우 기간(예: 8월 8~10일)의 실제 출동 수요와 각 모형의 예측값을 비교해 보면, 선형 모형인 MRA는 국지성 호우가 집중된 시점의 급격한 피크 수요를 과소 추정하는 한계를 보였다. 반면, Random Forest 모형은 비록 급격한 변동 시점에서 일부 시간적 지연(Time Lag)이 관찰되었으나, MRA와 달리 피크의 발생 방향성과 규모를 보다 근접하게 예측하는 모습을 보였다. 이는 비선형적 알고리즘이 복합재난의 불규칙한 패턴을 학습하는 데 유리함을 의미하며, 결과적으로 실제 상황에서 과소 대응의 위험을 줄이고 선제적 자원 배치가 가능하도록 하는 신뢰할 수 있는 기반을 제공한다.

Table 7. Status of occurrence of local heavy rain in 2022

지역	복합재난 발생 현황				대응 결과		
	국지성 호우		화재 발생		평균 출동 건수(분)	성공률(%)	피해 규모 (백만원)
	일 수	최대강수량 (mm/h)	화재 건수	복합재난 건수			
강남구	23	45.2	156	8	6.2	94.3	1,250
서초구	18	38.7	142	6	7.1	91.8	980
송파구	21	52.3	189	12	8.4	88.7	1,560
전체평균	20.7	45.4	162.3	7.2	7.2	91.6	1,263

결론

본 연구는 기후변화로 인해 빈번해지는 도시형 복합재난에 대응하기 위해, 서울특별시 서초구를 대상으로 기상 데이터, 소방 활동 내역, 그리고 사회 취약성 지표를 통합한 위험도 예측 모델을 구축하고, 이를 기반으로 효율성과 형평성을 동시에

고려한 소방 자원 배치 전략을 제안하였다.

주요 연구 결과와 시사점은 다음과 같다.

첫째, 복합재난 위험도 예측 방법론으로서 머신러닝의 유효성을 입증하였다. 시간당 화재 출동 건수와 발생 여부를 예측함에 있어 Random Forest 모델은 전통적인 다중회귀분석(MRA) 대비 우수한 예측 성능(MAE 약 25% 감소, AUC 0.90 달성)을 보였다. 이는 국지성 호우와 화재 발생 간의 관계가 단순 선형적이지 않으며, 강우량과 도시의 사회·환경적 요인이 복잡한 비선형적 상호작용을 일으킨다는 점을 시사한다. 따라서 향후 재난 예측 시스템 구축 시 비선형성을 포착할 수 있는 머신러닝 기법의 도입이 필수적임을 확인하였다.

둘째, 사회적 취약성이 재난 위험의 주요 결정 요인임을 정량적으로 규명하였다. 변수 중요도 분석 결과, 직전 24시간 누적 강우량과 같은 기상학적 요인이 가장 큰 영향을 미쳤으나, 고령자 비율, 반지하·지하 주택 비율, 장애인 등록 비율 등 사회 취약성 변수 또한 상위 중요도를 차지하였다. 이는 재난 대응 정책 수립 시 물리적 위험 요소뿐만 아니라 인구통계학적 취약성을 반드시 고려해야 함을 의미하며, 기존의 경험적 판단에 의존하던 방식에서 탈피하여 데이터에 기반한 정밀한 위험 관리가 필요함을 뒷받침한다.

셋째, 본 연구는 예측된 위험도(Risk)와 사회 취약성(Vulnerability)을 결합한 '행정동별 우선 대응 지수(Priority Index)'를 개발하여, 실질적인 자원 배치 전략을 제시하였다. 제안된 모델은 한정된 소방 자원을 단순히 출동 빈도가 높은 지역이 아닌, 실제 피해 위험이 크고 자력 대피가 어려운 고위험·고 취약 지역(Hotspot)에 선제적으로 배치할 수 있는 객관적인 기준을 제공한다. 이는 재난관리의 목표인 '피해 최소화'와 '사회적 형평성 제고'를 동시에 달성할 수 있는 정책적 도구로 활용될 수 있다.

종합하면, 본 연구는 이종(Heterogeneous) 데이터를 융합하여 도시형 복합재난의 위험 패턴을 규명하고, 이를 구체적인 소방 정책으로 연결하는 통합 프레임워크를 제시했다는 점에서 의의가 있다. 향후 연구에서는 실시간 교통 트래픽 정보와 CCTV 영상 데이터를 연계하여 예측도의 정밀도를 높이고 긴급출동 경로 최적화를 포함한 동적 대응 모델로 고도화하여, 다양한 도시 환경에 적용하여 모델의 범용성을 검증하는 후속 연구의 필요가 있다.

References

- [1] Ahn, J.H., Lee, S.K. (2022). "Integrating social vulnerability into urban flood risk assessment using geospatial analysis." *International Journal of Disaster Risk Reduction*, Vol. 78, No. 3, pp. 102145-102155.
- [2] Breiman, L. (2001). "Random forests." *Machine Learning*, Vol. 45, No. 1, pp. 5-32.
- [3] Choi, K.W. (2021). Optimization of Emergency Medical Service Deployment using Time-Series Forecasting and Simulation. Ph.D. Dissertation, Seoul National University.
- [4] Cutter, S.L., Boruff, B.J., Shirley, W.L. (2003). "Social vulnerability to environmental hazards." *Social Science Quarterly*, Vol. 84, No. 2, pp. 242-261.
- [5] Gook, D.S. (2020). "Predicting fire incidence in metropolitan areas using random forest and meteorological data." *Journal of Fire Safety Engineering*, Vol. 35, No. 2, pp. 45-58.
- [6] Hwang, Y.M. (2018). *Urban Resilience and Climate Change Adaptation*. Academic Press, New York.
- [7] Jeong, M.C. (2023). A Study on the Impact of Localized Heavy Rainfall on Urban Infrastructure and Emergency Demand. National Disaster Management Research Institute, R-2023-DM-112, Sejong.
- [8] Kim, M.S., Lee, J.H. (2021). "A study on urban flood risk prediction using meteorological data." *Journal of the*

Korean Society for Disaster Information, Vol. 17, No. 3, pp. 45-58.

- [9] Korea Meteorological Administration (2024). Meteorological Data Open Portal. <https://data.kma.go.kr>, (Accessed: Oct. 10, 2024).
- [10] Lee, M.H., Park, J.E. (2019). "Compound disaster risk modeling based on machine learning and socioeconomic indicators." Proceedings of the International Conference on Urban Safety and Disaster Prevention, Tokyo, Japan, pp. 321-340.
- [11] Park, J.H. (2022). "Prediction of fire service dispatch demand using machine learning techniques." Korean Institute of Fire Science and Engineering Journal, Vol. 36, No. 2, pp. 101-112.
- [12] Ryu, S.G. (2023). "Comparative analysis of multiple regression and random forest for predicting time-series emergency calls." Journal of Risk Analysis and Disaster Management, Vol. 15, No. 4, pp. 50-65.
- [13] Seocho-gu Office (2023). Seocho-gu Statistical Yearbook 2023. Seocho-gu Office, Seoul.
- [14] Song, B.R., Kim, C.H. (2021). "Analyzing the relationship between single-person households and fire vulnerability in high-density urban areas." Fire Protection Science and Technology, Vol. 25, No. 1, pp. 112-125.