Machine Learning–Based Analysis and Prediction of 10-h Dead Fuel Moisture Content Using Automated Weather Observations in Gangwon Province, South Korea

SANGHEE CHAE,[a] KYU RANG KIM,[a] JUNG HYUK KANG,[a] HYEONG-SE JEONG,[a] AND SEUNGBUM KIM[a]

[a] *Research Applications Department, National Institute of Meteorological Sciences, Seogwipo, South Korea*

(Manuscript received 6 May 2025, in final form 20 August 2025, accepted 22 October 2025)

ABSTRACT: This study analyzed 10-h dead fuel moisture content (DFMC) data collected throughout the year 2024 at hourly intervals from 11 observation stations in Gangwon Province, South Korea, and developed a random forest–based prediction model. This model was chosen for its robust ability to capture the nonlinear relationships between diverse environmental factors and fuel moisture content, offering high predictive accuracy and interpretability. The 10-h DFMC exhibited distinct seasonal patterns, with notably low values in spring (March–May) and December, corresponding to periods of high forest fire occurrence on the Korean Peninsula. Topographical influences were evident; the highest-elevation station (DG) maintained high moisture content in February–March due to snow effects, while coastal stations (BG, GN, YG) showed greater variability in February compared to inland stations (BC, CS), likely due to frequent humidity fluctuations from sea fog and sea/land breezes, as well as higher precipitation frequency, highlighting the importance of maritime influences. The random forest model demonstrated good performance with R^2 = 0.80, root-mean-square error (RMSE) = 2.73%, and mean absolute error (MAE) = 1.93% in the test dataset. Variable importance analysis revealed that relative humidity was the most influential factor (0.7562), followed by hour (0.0948), temperature (0.0506), and Julian day (0.0499), while precipitation showed unexpectedly low importance (0.0042). Model performance varied by location, with the best results at BP (R^2 = 0.86, RMSE = 2.10) and MO (R^2 = 0.85, RMSE = 2.17) stations, and the poorest at the high-elevation DG station (R^2 = 0.76, RMSE = 3.18). The developed model provides a foundation for effectively predicting fuel moisture content in areas without real-time 10-h DFMC monitoring by using only meteorological data, though future improvements should incorporate additional topographical and environmental variables to enhance prediction accuracy.

KEYWORDS: Annual variations; Forest fires; Machine learning

1. Introduction

Wildfires are one of the natural disasters that cause serious damage to forest ecosystems and human society worldwide. The Gangwon region of South Korea is known for having the highest frequency and scale of wildfire damage due to its topographical and climatic characteristics. According to the Korea Forest Service's wildfire damage records (Korea Forest Service 2024), wildfires in the Gangwon region are showing a trend of increasingly extensive damage, as seen in cases such as the large-scale East Coast wildfire in 2000 and the Goseong and Sokcho wildfires in 2019. The increase in dry periods and strong wind frequency due to climate change is further exacerbating these risks.

One of the key factors influencing wildfire ignition and spread is the moisture content of dead fuels in forests (Carlson and Burgan 2003; Cawson et al. 2020; Chae 2014). Dead fuels refer to materials whose moisture content depends solely on surrounding meteorological conditions, including leaf litter on the soil surface and dead branches (Aguado et al. 2007; Simard 1968; Chae 2014). Dead fuel moisture content (DFMC) is a critical factor determining the ignition probability and rate of spread of wildfires, and accurately predicting it is essential for establishing effective wildfire management and prevention systems (Brown et al. 2022).

Time lag refers to the rate at which dead fuels respond to changes in atmospheric moisture. Based on this response time, fuels are categorized into different classes (e.g., 1, 10, 100 h) that correspond to materials ranging from fine grasses to large logs (Bradshaw et al. 1984). This study focuses specifically on 10-h DFMC (10-h DFMC). This class represents the moisture in medium-sized twigs and small branches (roughly 0.6–2.5 cm in diameter). As the name implies, these fuels take approximately 10 h to reach equilibrium with new environmental moisture conditions (Lee et al. 2019).

Traditional methods of measuring dead fuel moisture content involve collecting fuel samples and calculating the ratio by measuring their weight before and after drying (Viegas et al. 1992; Aguado et al. 2007; Desbois et al. 1997; Zormpas et al. 2017; Chae 2014). Recently, real-time in situ monitoring has been implemented using fuel moisture sensors. Additionally, research has been conducted to estimate spatial distribution using remote sensing data such as satellite observations (Chuvieco et al. 2002; Danson and Bowyer 2004; Yebra et al. 2013). Furthermore, studies are being carried out to estimate fuel moisture content through empirical models based on meteorological elements or machine learning methods (Shmuel et al. 2022; Rodrigues et al. 2024).

Recent advancements in machine learning techniques have led to active research in modeling complex natural phenomena by integrating various environmental variables. In particular, the random forest algorithm provides high predictive

Denotes content that is immediately available upon publication as open access.

Corresponding author: Kyu Rang Kim, krk9@korea.kr

DOI: 10.1175/JAMC-D-25-0068.1

TABLE 1. Information on 10-h DFMC observation stations in Gangwon Province installed by NIMS.

Station full name (acronym)	Station No.	Latitude (°)	Longitude (°)	Elevation (m)	Characteristics
Daegwallyeong (DG)	100	37.67713	128.71834	772.43	Mountainous
Daehwa (DH)	597	37.54548	128.44108	561.85	
Bongpyeong (BP)	525	37.61234	128.37725	552.11	
Jinbu (JB)	560	37.64794	128.56447	540.54	
Myeon-on (MO)	660	37.56197	128.37762	533.47	
Pyeongchang (PC)	526	37.37748	128.39469	303.18	Low-mountainous
Cuncheon-sinbuk (CS)	978	37.95461	127.77626	140.37	Inland plain
Bukcuncheon (BC)	93	37.94738	127.75443	95.78	
Bukgangneung (BG)	104	37.80456	128.85535	75.24	Coastal
Gangneung (GN)	105	37.75147	128.89099	27.12	
Yeongok (YG)	566	37.85155	128.81924	8.36	

accuracy and interpretability through an ensemble of multiple decision trees and is known to be suitable for estimating variables influenced by diverse environmental factors, such as dead fuel moisture content, due to its excellent ability to capture nonlinear relationships (Breiman 2001; Wright and Ziegler 2017; Kim et al. 2021; Lee et al. 2020).

The Gangwon region of South Korea has unique climatic characteristics, featuring complex mountainous terrain centered around the Baekdudaegan mountain range, with coexisting maritime climate influences from the East Sea and inland effects. Development of a region-specific dead fuel moisture content estimation model is necessary because of the complex topographical and climatic characteristics of Gangwon Province—particularly the coexistence of maritime influence. Unlike previous domestic and international studies that primarily focused on limited environmental conditions (utilizing very few points, forests, etc.), this study attempted to develop a more universal model encompassing the diverse environmental conditions.

This study has two primary objectives. First, we analyze the monthly and seasonal variability of 10-h DFMC using hourly in situ observations from 11 representative stations across diverse environments in Gangwon Province during 2024. Second, we develop and validate a random forest–based prediction model for 10-h DFMC. The model is designed to use a minimal set of inputs: standard meteorological data (temperature, relative humidity, wind speed, and precipitation), complemented by time of day and Julian day to account for the diurnal and seasonal patterns of solar radiation. The ultimate goal of this research is to create a universally applicable model that can reliably estimate 10-h DFMC from standard meteorological data anywhere in the world, thereby supporting improved forest fire risk assessment and management strategies.

Our contributions are fourfold: 1) We place station-resolved distributions alongside all-station aggregated distributions to compare observational behavior and model fidelity within a unified framework; 2) we provide explicit spatial skill visualization that links errors to physiography; 3) we perform event-based diagnostics contrasting snow-covered versus snow-free weeks to identify missing snow-state information as the primary source of winter bias; and 4) we contextualize 2024 relative to 1991–2020 normals to discuss generalizability under an anomalously dry year.

2. Research methods

a. Data

The National Institute of Meteorological Sciences (NIMS) has been conducting real-time in situ observations of 10-h DFMC at 11 stations established in the Gangwon region in late October 2023 to study the effects of cloud seeding and wildfire-related research. Detailed information about the observation stations is shown in Table 1. Gangwon Province has coastal areas to the east along the East Sea and mountainous terrain to the west along the Taebaek Mountains (Fig. 1). Observation stations were established in coastal regions (YG, GN, and BG), mountainous regions with elevations above 300 m (DG, DH, BP, JB, MO, and PC), and inland plain regions (BC and CS) to represent observations based on elevation and regional characteristics (Table 1). Additionally, the automatic weather system (AWS) stations used in this study for meteorological observations were located at the exact

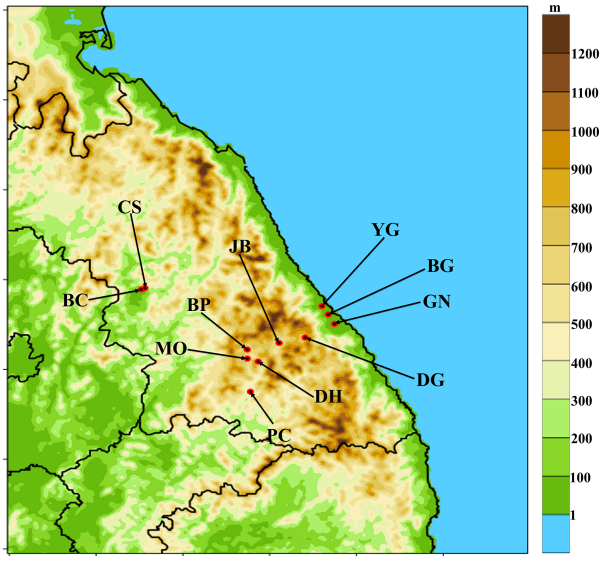

FIG. 1. Distribution of the 10-h DFMC observation stations in Gangwon Province. Observations at these stations began in late October 2023, and the data analyzed in this study cover the period from 1 Jan to 31 Dec 2024. Meteorological variables were simultaneously observed by the Korea Meteorological Administration's AWS located at the same 10-h DFMC sites.

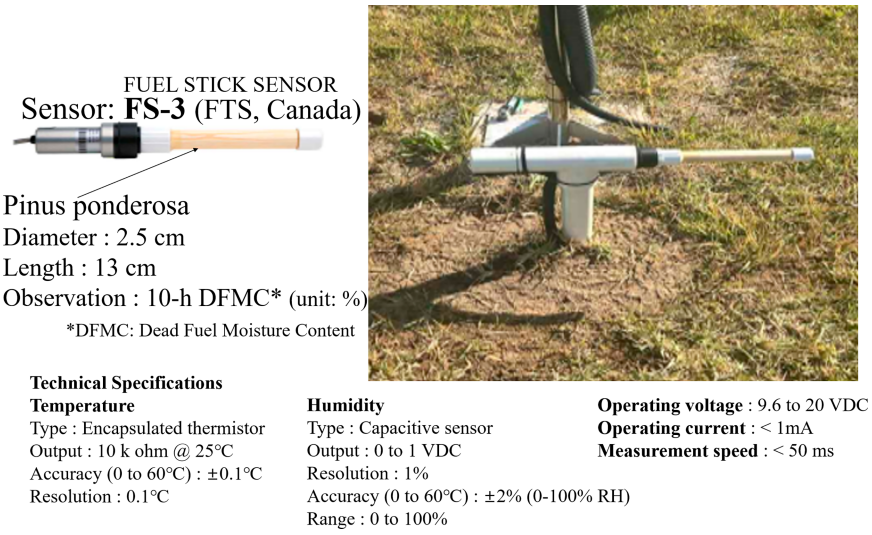

FIG. 2. Photographs of sensors installed at the observation station and their technical specifications.

same sites as the DFMC observation stations listed in Table 1. Therefore, both fuel moisture and meteorological variables were simultaneously observed at all 11 locations presented in Table 1.

The sensor installed for fuel moisture observation is the FS-3 sensor from FTS Inc. (Fig. 2). This sensor contains temperature and humidity sensors embedded in ponderosa pine material, with a diameter (D) of 2.5 cm and a length of 13 cm. It observes the 10-h time-lag DFMC (10-h DFMC). The sensors were installed facing north to minimize the effect of direct sunlight, which could otherwise elevate the sensor temperature and, consequently, result in inaccurately low fuel moisture readings. Additionally, the sensors were positioned approximately 20–30 cm above the ground to avoid direct influence from the soil surface. The sensors record observations at 1-min intervals, and the data used for analysis were hourly averages from 0000 Korea Standard Time (KST) 1 January 2024 to 2300 KST 31 December 2024.

Korea Meteorological Administration's AWS observation data were used to estimate fuel moisture content. AWS data were collected from the same site where fuel moisture observations were conducted, with the station numbers in Table 1 corresponding to the KMA AWS station numbers. The AWS data period used was identical to the fuel moisture data period, utilizing hourly average data from 0000 KST 1 January 2024 to 2300 KST 31 December 2024. The selected meteorological elements were 2-m air temperature (°C), 2-m air humidity (%), 10-m wind speed (m s^{-1}), and 1-h accumulated precipitation (mm). Additionally, since there were no solar radiation data at the observation stations, Julian day and hour information were included in the dataset to reflect solar radiation elements. Because solar radiation was not measured at the DFMC sites, we used hour and Julian day to partially represent diurnal/seasonal radiation forcing. For portability across sites lacking specialized data, we intentionally limited predictors to AWS variables; thus, static physiographic (elevation, slope/aspect, coastal proximity), vegetation structure, and snow information were not used in the present model.

To ensure data quality, we applied comprehensive validation procedures that including automatic detection and removal of abnormal values (such as sensor startup errors, persistent identical readings, and transmission dropouts), outlier identification and exclusion of unrealistic values (e.g., persistent 0% or negative moisture content), and complete case analysis, where any record with missing values in one or more variables was excluded from model training and evaluation datasets.

b. *Development of machine-learning-based 10-h DFMC estimation model*

Random forest technique was applied to estimate dead fuel moisture content ($\hat{y}$) using the RandomForestRegressor model from the Python-based scikit-learn library (Pedregosa et al. 2011). This model is an ensemble composed of multiple decision trees, where each tree performs an individual estimation and the average of these values is calculated as the final estimate (Fig. 3, $\hat{y}$):

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(\mathbf{x}).$$

In this equation, $\mathbf{x}$ is the input feature vector, which includes 2-m temperature, 10-m wind speed, 1-h accumulated precipitation, 2-m relative humidity, Julian day, and time. The T is the total number of decision trees in the random forest algorithm. The $f_t(\mathbf{x})$ is the fuel moisture value estimated by the decision tree for the input $\mathbf{x}$. In other words, the final fuel moisture prediction is calculated by averaging the predicted values from each decision tree for a given input. This approach reduces the variance of predictions compared to a single decision tree and provides more stable results. According to previous research, the random forest approach prevents model overfitting and quickly processes substantial amounts of data to produce optimal results (Kim et al. 2021).

To assess potential overfitting, the original dataset was divided into training and test sets at a 7:3 ratio, with the model

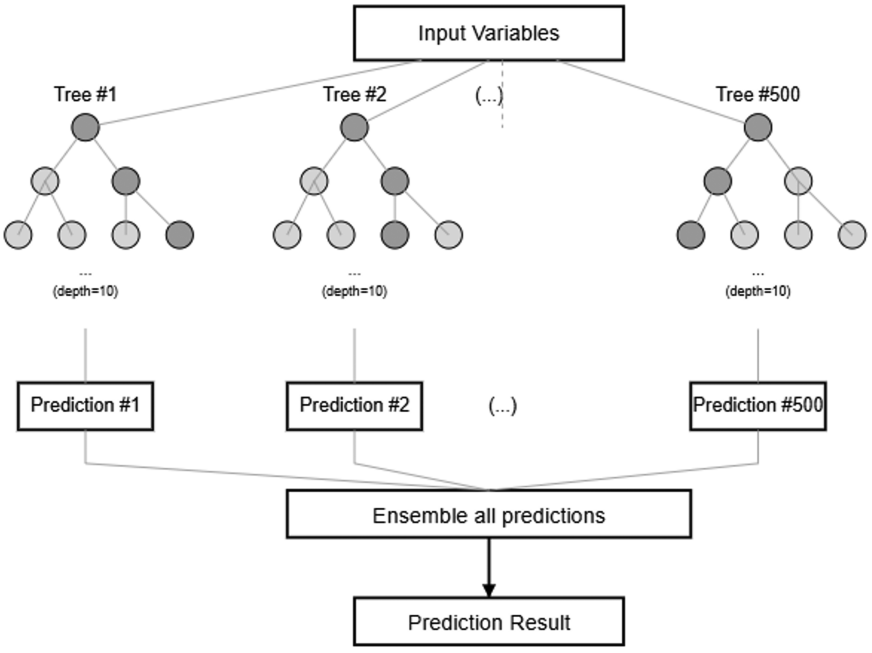

FIG. 3. Schematic diagram of the prediction algorithm of the RF model (Kim et al. 2021).

trained on the training set and evaluated on the test set. When extracting samples for each tree from the training data, random_state was set to 42 to ensure the same samples were selected each time. By varying the number of trees from 100 to 1000 and tree depths from 1 to 20 while comparing the coefficient of determination, we optimized the model by adjusting these parameters.

In addition, to verify the suitability of the random forest (RF) model, we conducted comparative experiments with linear regression, a kernel-based nonlinear model [support vector regression (SVR)], and an advanced ensemble method [gradient boosting regressor (GBR)]. Detailed performance metrics are summarized in Table A1 in appendix A.

c. *Validation of 10-h DFMC estimates*

To ensure robust evaluation of model performance and spatiotemporal generalizability, we implemented a comprehensive validation framework consisting of three complementary approaches. Initially, we performed a conventional random split validation using 70% of the data for training and 30% for testing. However, recognizing that this approach may not adequately assess the model's ability to generalize across different temporal and spatial conditions, we additionally conducted 5-fold cross validation and seasonal cross validation.

The 5-fold cross validation was implemented to evaluate model stability and reduce the potential bias associated with a single train–test split. Data were randomly partitioned into five equal subsets, with each subset serving as the validation set once while the remaining four subsets were used for training. This process was repeated five times, and performance metrics were calculated as the mean and standard deviation across all folds.

Seasonal cross validation was conducted to assess the model's performance across different temporal conditions, which is particularly important given the seasonal variability in fuel moisture content. For this validation approach, we divided the dataset by seasons (winter, spring, summer, and autumn) and trained the model on three seasons while testing on the remaining season. This process was repeated for all possible combinations, allowing us to evaluate how well the model generalizes to unseen seasonal conditions.

Performance was evaluated using three metrics: the coefficient of determination (R^2), the root-mean-square error (RMSE), and the mean absolute error (MAE). These metrics provide complementary insights into model accuracy, with R^2 indicating the proportion of variance explained by the model, RMSE penalizing larger errors more heavily, and MAE providing a measure of average prediction error:

$$R^2 = \left(\frac{\sum_{i=1}^{n} (\hat{y}_i - \overline{\hat{y}})(o_i - \overline{o})}{\sqrt{\sum_{i=1}^{n} (\hat{y}_i - \overline{\hat{y}})^2 \sum_{i=1}^{n} (o_i - \overline{o})^2}} \right)^2,$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - o_i)^2},$$

$$\mathrm{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - o_i|.$$

In these equations, $\hat{y}$ represents the estimated values, and o represents the observed values. In this study, the model's performance was comprehensively evaluated considering the aforementioned indices.

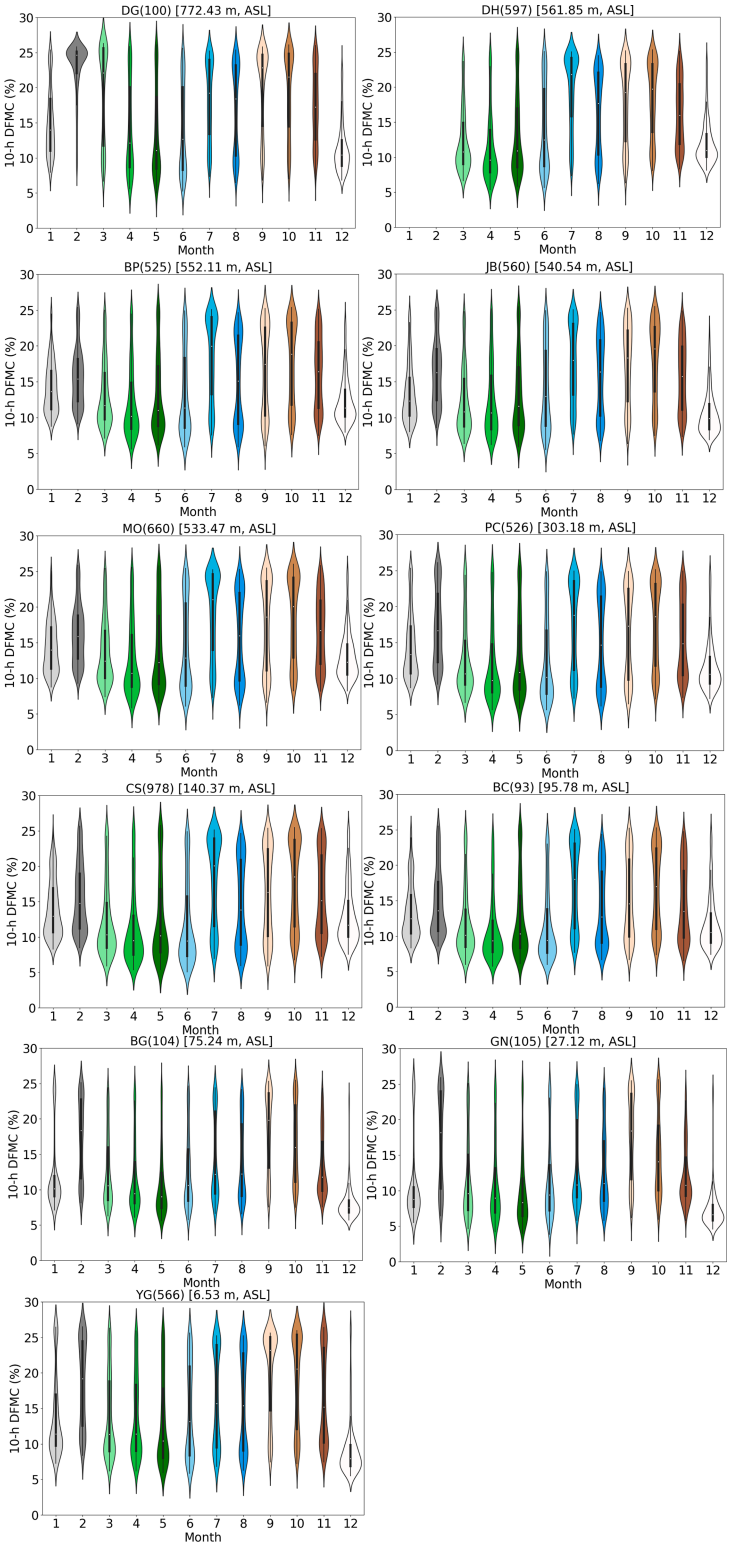

FIG. 4. Monthly variability of 10-h DFMC (%) at 11 observation stations, based on hourly data from 2024. For each violin, the white circle marks the median, the central black bar shows the interquartile range, and the shape represents the data's probability density. Colors correspond to the seasons: green tones for spring (March–May), blue tones for summer (June–August), brown tones for autumn (September–November), and white to gray for winter (December–February).

TABLE 2. Statistical values of monthly observed 10-h DFMC (%) measurements for each station (DG, DH, BP, JB, PC, and CS). The values were expressed to two decimal places.

	January	February	March	April	May	June	July	August	September	October	November	December
							DG					
Mean	15.10	22.79	19.18	14.17	13.62	14.15	18.34	17.03	19.86	19.38	17.32	11.33
Std	4.91	3.99	6.82	6.47	6.37	6.32	5.60	6.28	5.94	5.95	5.30	3.57
Min	7.93	8.21	6.67	5.59	5.00	5.27	7.36	6.49	6.80	7.01	6.77	6.84
Median	13.95	24.66	22.10	12.10	11.04	12.64	19.23	18.42	22.82	21.53	17.25	10.43
Max	25.39	25.79	26.32	25.85	25.85	26.11	25.64	25.22	25.37	25.79	26.09	25.99
							DH					
Mean	—	—	12.38	11.47	13.09	14.14	19.57	16.50	17.81	18.44	16.20	12.07
Std	—	—	4.58	5.08	5.66	6.09	5.31	5.80	5.81	5.33	4.74	3.24
Min	—	—	6.63	5.36	5.84	5.66	7.35	6.23	6.31	8.08	8.32	8.13
Median	—	—	10.74	9.57	10.92	12.45	21.80	17.66	19.21	19.72	15.99	10.98
Max	—	—	24.74	25.02	25.53	25.06	25.01	24.49	25.20	25.32	25.15	24.63
							BP					
Mean	14.19	15.68	13.30	12.10	13.26	13.44	18.45	15.17	16.70	17.68	16.21	12.36
Std	3.56	4.04	4.65	5.03	5.57	5.83	5.77	6.02	6.09	5.80	5.01	3.30
Min	8.78	9.10	7.12	5.59	6.17	5.90	7.53	5.90	6.03	7.59	7.52	7.92
Median	13.63	15.38	11.71	10.19	10.99	11.34	19.95	15.08	17.45	18.84	16.41	11.35
Max	24.91	25.35	24.97	25.14	25.67	24.88	25.08	24.66	25.21	25.38	25.13	24.35
							JB					
Mean	13.48	16.36	12.43	12.44	13.29	14.13	17.69	15.90	17.37	18.15	15.68	10.62
Std	4.15	4.51	4.67	5.08	5.47	5.79	5.33	5.47	5.51	5.36	4.85	3.07
Min	8.03	8.56	6.37	6.21	5.93	5.56	7.70	6.79	6.35	7.48	7.37	6.90
Median	12.32	16.27	10.81	10.65	11.58	12.94	17.99	16.37	18.33	19.56	15.72	9.87
Max	24.74	25.36	24.76	25.15	25.44	24.91	25.03	24.62	25.24	25.46	24.82	22.53
							MO					
Mean	14.44	16.22	13.79	12.74	14.06	14.59	19.11	15.87	17.54	18.57	16.63	13.07
Std	3.64	4.14	4.60	5.17	5.73	6.17	5.75	6.03	6.12	5.84	4.96	3.37
Min	8.76	9.30	7.39	5.85	6.53	6.08	7.32	6.21	6.60	8.22	8.16	8.25
Median	13.97	15.88	12.44	10.73	12.20	12.84	21.06	15.96	18.58	20.06	16.70	12.29
Max	25.09	25.84	25.41	25.55	25.78	25.44	25.22	25.28	25.49	25.99	25.80	25.37
							PC					
Mean	14.62	17.21	12.80	12.12	13.14	12.54	17.52	15.04	16.39	17.53	15.53	11.79
Std	4.58	5.24	5.12	5.55	5.88	5.79	6.02	6.15	6.18	5.76	5.24	3.74
Min	8.73	9.21	6.67	5.62	5.82	5.62	7.26	6.20	6.47	7.16	7.37	7.21
Median	13.35	16.67	10.68	9.72	10.85	10.16	18.77	14.66	17.28	18.62	14.86	10.60
Max	25.31	26.11	25.83	25.65	25.67	24.86	24.93	24.86	24.89	25.15	24.94	25.10
							CS					
Mean	14.08	15.53	12.28	11.29	12.68	11.90	17.97	14.65	16.24	17.59	15.97	13.08
Std	4.07	4.80	5.22	5.27	6.15	5.94	6.16	6.08	6.16	6.05	5.65	4.22
Min	8.27	8.49	5.89	5.34	5.40	5.07	6.95	5.95	5.89	6.77	7.25	7.56
Median	12.95	14.74	10.33	9.51	10.17	9.47	20.06	13.83	16.32	18.50	15.14	11.69
Max	25.14	25.68	25.85	25.61	25.78	24.91	25.11	24.65	25.38	25.46	25.33	25.55

In addition to the standard metrics, we conducted an event-based validation to diagnose error sources. At DG, we paired a snow-covered week (20–27 February 2024) with a snow-free summer week (1–7 August 2024) to examine potential sensor malfunction, data anomalies, and feature insufficiency. DFMC was analyzed together with concurrent AWS variables (temperature, relative humidity, wind speed, precipitation) and snow depth/cover.

3. Results

a. Analysis of observed 10-h DFMC data

The monthly variability of 10-h DFMC was investigated using data from 11 stations in the Gangwon region. Figure 4 complements the aggregated analysis by presenting station-resolved violin plots that reveal spatial heterogeneity and local distributional characteristics, including multimodality patterns that vary across sites due to topographic and microclimatic influences. At most observation stations, monthly 10-h DFMC showed lower trends during spring (March–May) and December of winter (Tables 2 and 3). This suggests that spring (March–May) and December of 2024 were vulnerable periods for forest fires in the Gangwon region.

The DG station, located at the highest elevation in the mountainous area, showed the highest 10-h DFMC distribution in February. At this station, heavy snowfall in February covered the sensors, and, due to low winter temperatures, the snow did not melt but remained covering the sensors

TABLE 3. Statistical values of monthly, the entire period, and seasonally categorized observed 10-h DFMC (%) measurements for each station (BC, BG, GN, and YG) and all stations. The values were expressed to two decimal places.

	January	February	March	April	May	June	July	August	September	October	November	December
							BC					
Mean	13.46	14.78	11.87	10.89	12.58	11.48	17.18	13.98	15.31	16.69	14.56	11.87
Std	3.96	4.63	4.85	4.60	5.58	5.17	5.87	5.35	5.62	5.70	5.19	3.92
Min	8.35	8.58	6.00	5.73	6.14	6.00	7.69	6.67	6.67	7.37	6.54	7.38
Median	12.47	13.60	10.11	9.34	10.31	9.39	18.03	12.73	14.57	16.99	13.48	10.53
Max	24.83	25.51	25.15	24.81	25.75	24.55	25.02	24.33	25.23	25.30	24.72	25.14
							BG					
Mean	12.20	17.32	12.56	11.62	10.77	12.61	14.58	14.08	18.30	16.43	13.43	8.10
Std	5.22	5.64	5.11	5.10	4.84	5.33	5.87	5.54	5.63	5.60	4.53	2.63
Min	7.09	8.12	6.26	6.08	5.66	5.73	6.62	6.59	7.52	7.47	7.55	5.69
Median	10.11	18.36	10.55	9.42	8.99	10.59	12.15	12.21	19.81	15.98	11.79	7.46
Max	24.86	25.15	24.38	24.74	25.35	24.66	24.42	24.15	25.34	25.48	24.37	23.72
							GN					
Mean	11.02	17.17	11.59	10.94	9.99	11.38	13.97	12.94	17.51	15.07	12.24	7.40
Std	5.73	6.75	5.74	5.61	5.12	5.71	6.09	5.41	5.99	5.79	4.38	3.01
Min	5.52	6.44	4.63	4.17	3.95	3.83	5.77	5.79	6.17	5.93	5.77	4.59
Median	8.78	18.16	9.57	8.97	8.31	9.42	10.81	11.03	18.39	14.09	10.78	6.61
Max	25.53	25.93	25.08	25.41	25.57	25.04	24.93	24.47	25.45	25.60	24.63	24.70
							YG					
Mean	14.01	18.55	13.94	13.79	13.08	14.58	16.34	15.83	19.89	18.79	16.59	9.24
Std	5.91	6.06	5.96	5.86	6.13	6.51	6.71	6.64	6.11	6.47	6.50	4.11
Min	7.23	8.40	6.16	6.16	5.96	5.86	6.82	6.27	7.43	7.21	7.42	5.48
Median	11.38	19.18	11.34	11.41	10.44	13.16	15.68	15.37	23.17	20.55	15.17	7.94
Max	26.45	26.47	26.31	25.93	25.94	25.62	25.25	25.23	25.63	25.96	26.41	26.57
							All stations					
Mean	13.66	17.15	13.29	12.14	12.69	13.17	17.34	15.18	17.54	17.67	15.49	11.29
Std	4.78	5.48	5.62	5.46	5.81	6.00	6.11	6.01	6.08	5.91	5.35	4.18
Min	5.52	6.44	4.63	4.17	3.95	3.83	5.77	5.79	5.89	5.93	5.77	4.59
Median	12.24	16.81	11.11	9.98	10.40	10.90	18.20	14.66	18.75	18.45	14.74	10.30
Max	26.45	26.47	26.32	25.93	26.11	25.64	25.25	25.37	25.79	26.09	26.41	26.57

All stations for period	All period	Spring (March–May)	Summer (June–August)	Fall (September–November)	Winter (December–February)
Mean	14.77	12.71	15.26	16.91	14.13
Std	6.03	5.65	6.27	5.87	5.41
Min	3.83	3.95	3.83	5.77	4.59
Median	13.24	10.51	14.41	17.13	12.49
Max	26.57	26.32	25.64	26.41	26.57

for an extended period, resulting in a tendency to maintain high fuel moisture distribution. In this context, "high DFMC distribution" means the distribution is shifted toward the upper quantiles relative to the same period/group, with higher median/upper quantiles and fewer low-DFMC observations.

The DH, BP, JB, and MO stations, located at 500-m elevation, showed similar monthly plots. For the DH station, valid (nonmissing) observations began from 1600 UTC 29 February, so February was excluded from the plot calculations. Excluding the DG station, most locations showed the smallest variation in 10-h DFMC during spring (March–May), indicating that consistently dry conditions were maintained during this period. In contrast, summer (June–August) and autumn (September–November) showed larger variations, which appeared to be caused by precipitation events or instability in weather conditions.

Analysis of DFMC data from all 11 observation stations through monthly boxplots in this study revealed that extreme values near the maximum were consistently detected at all stations particularly in December. This phenomenon appears to be closely related to the lack of precipitation during winter. During the winterly dry periods, precipitation decreases significantly, which lowers atmospheric humidity and promotes moisture evaporation from fuels. December, in particular, is a seasonal transition period when temporary weather changes or micromoisture inputs such as dew or frost may occur on certain dates. These can cause rapid changes (increases) in fuel moisture content in an otherwise dry environment, resulting in extreme values near the maximum.

At coastal stations BG, GN, and YG, located at lower elevations, the 10-h DFMC showed large variations in February, whereas at BC and CS stations, located in inland areas at lower elevations, the 10-h DFMC variations in February were smaller. Coastal areas can experience greater humidity fluctuations due to direct influence from the sea. The daily cycle of sea and land breezes, the formation and dissipation of fog (including sea fog from marine moisture and radiation fog over

land surfaces), and other coastal phenomena can cause frequent changes in atmospheric humidity, potentially resulting in high variability in fuel moisture content. Additionally, coastal areas may experience a higher precipitation frequency in winter compared to inland areas, which can cause frequent changes in fuel moisture content. While coastal areas generally have higher relative humidity than inland areas due to moisture from the sea, they also experience greater variability depending on wind and weather conditions. These humidity variations likely had a direct impact on the fluctuations in fuel moisture content.

At all observation stations, the December distribution (white violin) consistently shows a narrow upper portion and a wide lower portion. This asymmetric distribution pattern clearly suggests that 10-h DFMC values in December are predominantly concentrated in the lower range. In such cases, the mean value tends to be greater than the median value. In contrast, from July to October (blue to brown series violins), most stations show wide distributions with multimodal characteristics, confirming that fuel moisture content variability is greater during summer and autumn.

In addition to coastal sea–land breeze effects, terrain-induced circulations likely contribute to the station-to-station contrasts in DFMC. In mountainous areas, daytime upslope/valley winds and nocturnal downslope drainage flows/cold pools can modulate moisture transport and thus fuel drying–wetting cycles. These meso- to local-scale thermally driven processes—varying with ridges, slopes, and valleys—can produce nocturnal wetting/fog or enhance daytime drying on sun-exposed slopes, providing context for the site-to-site differences observed here.

Analysis of the violin plots revealed that the DG station showed unique temporal variability that differentiated it from other observation stations. A particularly noteworthy feature was the DFMC distribution in February–March, when fuel moisture content remained very high at an average of over 25% with the distribution concentrated in the upper range. This pattern differs from the typical climate characteristics of the Korean Peninsula, where spring is generally known as a dry period, reflecting the special microclimate conditions of the high mountain area. Specifically, snowmelt in high elevations or increased fog frequency during this period likely contributed to the increase in fuel moisture content. In contrast, from April to June, DFMC values decreased significantly and the distribution range expanded considerably to approximately 5%–25%. This wide distribution range suggests that the variation in fuel moisture content increased as temporary precipitation events and dry periods alternated during a period of high variability in spring weather conditions. In May–June, particularly, the violin plots showed long and slender shapes, indicating that fuel moisture content values were relatively evenly distributed across the entire range. During summer and autumn (July–November), DFMC values rose again, forming the center of the distribution at approximately 20%–25%, with violin plot patterns showing thicker upper portions. This indicates that relatively high DFMC values were frequently observed during this period, likely influenced by summer monsoons, concentrated rainfall, and autumn dew formation. In winter (December–January), DFMC values decreased again to approximately 10%–15%, with distribution shapes showing simple and narrow patterns. This indicates reduced variability in fuel moisture content due to the influence of dry winter climate. The DG station's unique annual DFMC variation pattern suggests that fuel moisture dynamics in high mountain areas may be regulated by different mechanisms than in lower elevations. In particular, the phenomenon of high DFMC values during the generally dry spring period implies that a differentiated approach by elevation is necessary in developing forest fire risk assessment and prediction models.

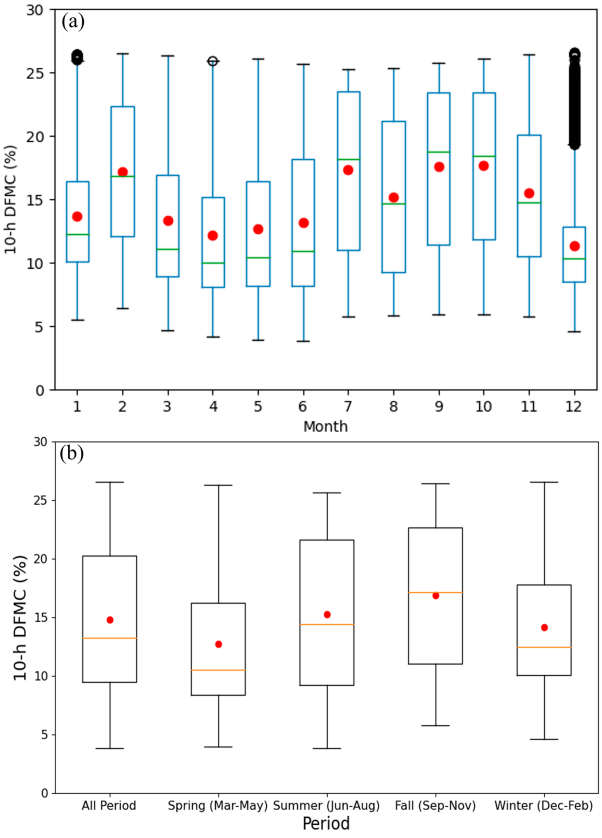

FIG. 5. (a) Monthly boxplots showing data from all observation stations. (b) Seasonal boxplots displaying data from all stations across the entire study period. The red dots indicate mean values.

Overall, most stations showed a trend of higher DFMC values from June to September (light sky blue to light brown violins) and lower values from March to May (green series violins), though the magnitude of this seasonal variation differed among stations. These station-specific and seasonal variation patterns suggest that fuel moisture content in the Korean Peninsula is complexly influenced by various local factors such as topography and elevation, in addition to climatic elements. Particularly, the low DFMC distribution pattern in spring corresponds with the increased forest fire risk during this period.

To establish the overall distributional context, Fig. 5 shows pooled 10-h DFMC observations from all stations, summarizing temporal patterns across monthly, seasonal, and full-period scales. Analysis of temporal variability in 10-h DFMC revealed

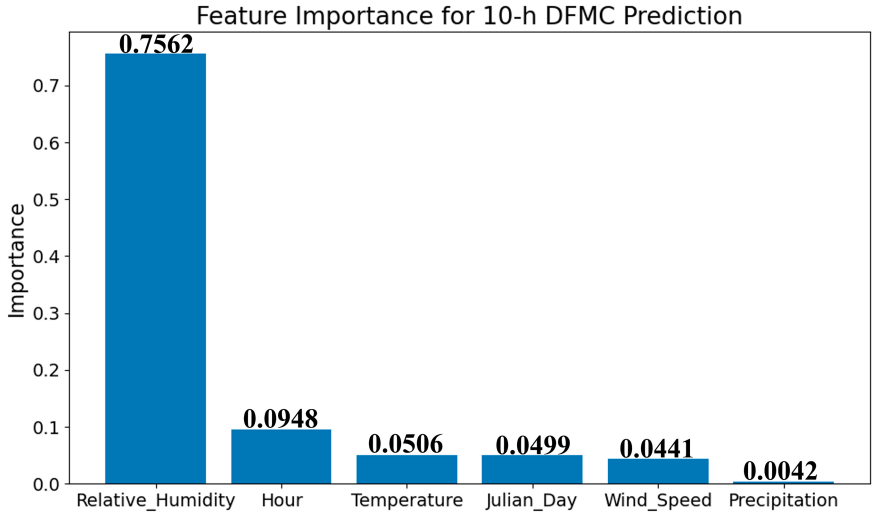

FIG. 6. Importance of meteorological variable inputs in the RF model.

distinct seasonal patterns. In the monthly analysis, maximum values remained relatively constant at approximately 25%–26% throughout the entire period, while median and mean values showed considerable variation. Particularly high average DFMC values (approximately 17.15%–17.67%) were observed in February, July, September, and October, whereas December showed the lowest average value (approximately 11.29%). In the seasonal analysis, autumn (September–October) recorded the highest median (approximately 18.75%) and mean values (approximately 17.67%), and also showed the greatest variability. In contrast, spring (March–May) exhibited the lowest median (approximately 10%) and mean values (approximately 13%). These seasonal variations are associated with the typical climate patterns of the Korean Peninsula, reflecting characteristics of the spring dry period and the relatively wet autumn period.

Notably, the meteorological background of 2024 itself was anomalous. At the high-elevation DG station, annual precipitation was only 69% of normal despite higher relative humidity (+3.4%), while at GN, precipitation was near normal overall but extremely imbalanced across seasons, with summer rainfall at only ~36% of normal and autumn exceeding 150% (Tables B1 and B2 in appendix B). Both sites recorded +1.8°C above normal annual temperature. These anomalies provide important context for the seasonal DFMC variability observed in this study.

b. *Random forest model*

The model configuration of 500 trees and a tree depth of 10 was selected as optimal based on achieving the highest R^2 values with minimal performance gap between training (70%) and test (30%) datasets, indicating good generalization without overfitting. Each tree performs splitting using the features and calculates the importance of these features during this process. Importance values are expressed as real numbers between 0 and 1, representing the relative importance among features. Analysis of variable importance in the machine learning model for 10-h DFMC prediction revealed that relative humidity had a significantly higher influence with an importance value of 0.7562 compared to all other variables combined (Fig. 6). This suggests that most of the variability in 10-h DFMC can be explained by atmospheric relative humidity. The second most important variable was time (hour, 0.0948), indicating that diurnal variability is a significant factor in 10-h DFMC prediction. The 2-m air temperature (0.0506) and Julian day (0.0499) showed similar levels of importance, suggesting that direct temperature effects and seasonal variability influence fuel moisture content to a similar degree. Wind speed (0.0441) showed moderate influence, while precipitation (0.0042) showed surprisingly low importance contrary to expectations.

These results suggest that 10-h DFMC responds more significantly to sustained atmospheric humidity conditions rather than direct precipitation events, and the overwhelming importance of relative humidity implies that special attention should be paid to this variable when developing 10-h DFMC prediction models. Furthermore, these findings align with the research results of Resco de Dios et al. (2015), which reported that fine fuels respond to daily temperature and relative humidity changes due to their short lag times.

To further quantify the contribution of each meteorological variable to the prediction model, we conducted Shapley additive explanations (SHAP) and partial dependence plot (PDP) analyses (Fig. 7). The SHAP summary plot revealed that relative humidity exerted the strongest influence on 10-h DFMC predictions, consistent with the random forest Gini importance results. In contrast, precipitation exhibited SHAP values clustered around zero, indicating a minimal direct contribution to the model output. This pattern is likely attributable to the hourly temporal resolution of our dataset and the 10-h DFMC's lag-response characteristics. Specifically, the sustained high atmospheric relative humidity following rainfall events, rather than the instantaneous precipitation amount itself, appears to be the primary driver of DFMC

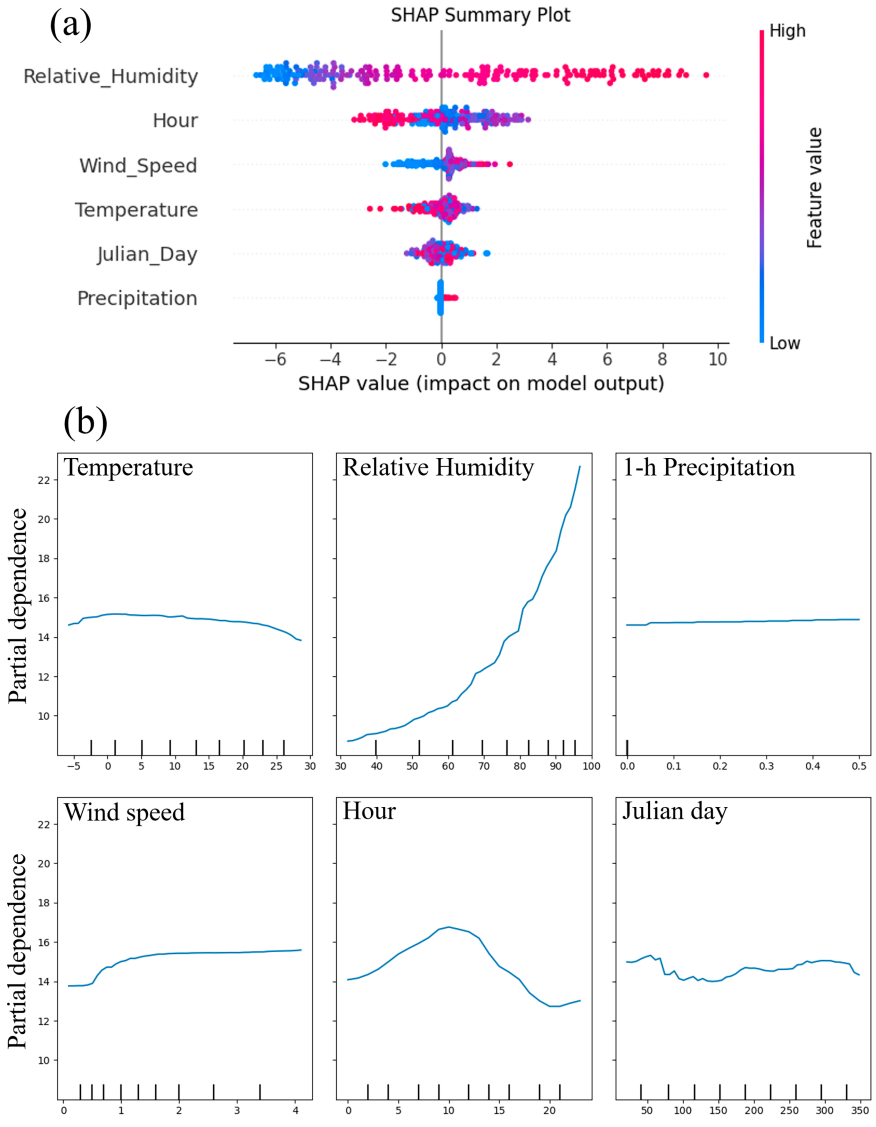

FIG. 7. Model interpretability analysis using SHAP values and PDPs for the 10-h DFMC prediction model. (a) SHAP summary plot showing the impact magnitude and direction of each meteorological variable on model predictions. Each dot represents a single prediction, with colors indicating feature values (blue: low, red: high), and the horizontal position showing SHAP values (negative: decreasing fuel moisture, positive: increasing fuel moisture). (b) PDPs illustrating the average marginal effect of each feature on the predicted fuel moisture while holding other variables constant.

variation, leading to precipitation's influence being indirectly captured through the humidity variable.

The PDP results further supported this interpretation. Predicted DFMC values remained nearly constant across the observed precipitation range, which can be explained by the predominance of zero-rainfall hours and the absorption of postrainfall effects into other meteorological variables, particularly relative humidity and temperature. In contrast, relative humidity showed a clear nonlinear positive relationship with DFMC, with predictions increasing sharply at higher humidity levels. These findings suggest that the low precipitation importance in our model should not be interpreted as a lack of relevance, but rather as a consequence of variable design, temporal resolution, and the absence of lagged interaction terms. Future work will incorporate cumulative precipitation, time since rainfall, and explicit interaction terms to better isolate and quantify precipitation effects.

c. *Validation of the 10-h DFMC estimation model*

We compared the predictions from the random forest–based estimation model with observations using automatic weather observation data from 11 stations in the Gangwon

region. Figure 8 shows scatterplots for both training and test sets. The training set contained 65 657 samples, with R^2 of 0.82, RMSE of 2.53%, and MAE of 1.80. The test set, which represents the actual model performance, contained 28 139 samples, with R^2 of 0.80, RMSE of 2.73%, and MAE of 1.93. These index values are rounded to the third decimal place. The performance difference between the training and test sets was not significant, indicating no overfitting problems and showing good model performance. Examining the scatterplot, the model estimated relatively well in the center for lower 10-h DFMC ranges but tended to underestimate compared to observations in the higher 10-h DFMC ranges above 20%.

In addition, we compared the performance of the RF model with other machine learning techniques, including a statistical model (linear regression), a kernel-based nonlinear model (SVR), and an advanced ensemble method (GBR). The results showed that the RF model substantially outperformed linear regression (R^2 = 0.56, RMSE = 4.01%, MAE = 3.31%) and SVR (R^2 = 0.65, RMSE = 3.57%, MAE = 2.71%), while achieving nearly identical performance to GBR (R^2 = 0.81, RMSE = 2.61%, MAE = 1.86%). This confirms that RF not only captures complex nonlinear relationships more effectively than simpler approaches but also achieves predictive skill comparable to state-of-the-art ensemble methods. Detailed results are summarized in Table A1 in appendix A.

The comprehensive validation framework revealed distinct patterns in model performance across different evaluation approaches. The conventional random split validation provided initial performance benchmarks, while the additional cross-validation methods offered deeper insights into model robustness and generalization capabilities. The 5-fold cross-validation results demonstrated consistent model performance with relatively low variability across folds. The model achieved an R^2 of 0.7901 ± 0.0035, indicating that approximately 79% of the variance in fuel moisture content was explained by the model with minimal variation between folds. The RMSE was 2.7562 ± 0.0316, and the MAE was 1.9472 ± 0.0134, both showing small standard deviations that suggest stable model performance regardless of the specific training–validation split used.

The seasonal cross-validation results revealed more substantial variability in model performance, highlighting the influence of temporal factors on prediction accuracy. The overall seasonal cross validation yielded an R^2 of 0.6093 ± 0.1583, with an RMSE of 3.4906 ± 0.5171 and an MAE of 2.6256 ± 0.4557. The larger standard deviations in these metrics reflect the significant seasonal differences in model performance. Examining individual seasonal performance, the model demonstrated the highest accuracy during summer months (R^2 = 0.7866, RMSE = 2.9003, MAE = 2.1098) and autumn months (R^2 = 0.7246, RMSE = 3.0812, MAE = 2.2800). In contrast, winter showed the poorest performance (R^2 = 0.3835, RMSE = 4.1545, MAE = 3.2490), followed by spring (R^2 = 0.5426, RMSE = 3.8265, MAE = 2.8636).

This seasonal variation in model performance can be attributed to the inherent characteristics of fuel moisture dynamics throughout the year and region-specific environmental factors. During summer and autumn, fuel moisture patterns tend to be more stable and predictable, likely due to more consistent

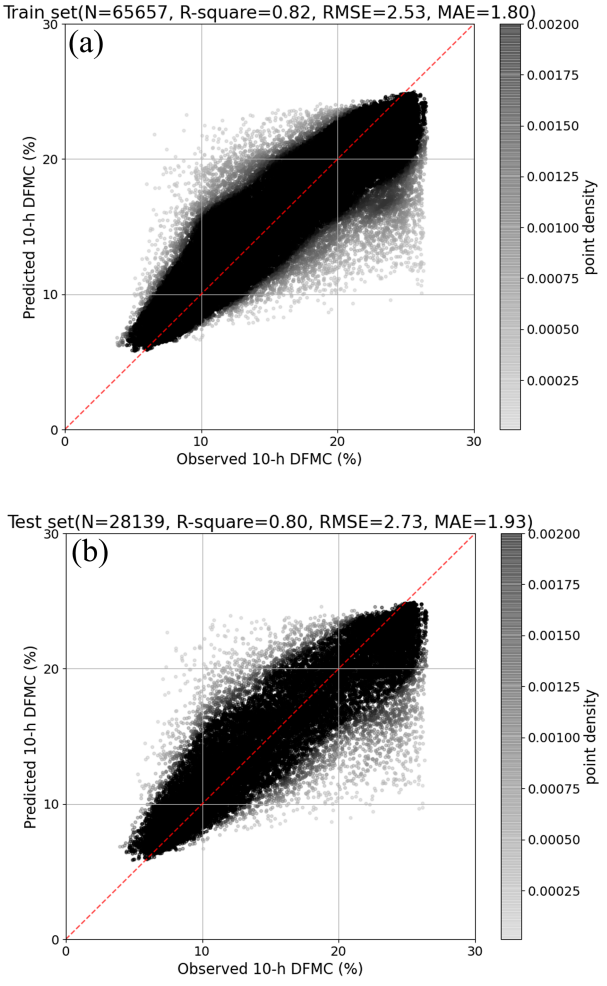

FIG. 8. Scatterplots of observed vs predicted 10-h DFMC for (a) the training set and (b) the test set. The red lines represent the 1:1 line.

weather patterns and vegetation phenology. Conversely, the poor performance during winter and spring periods can be primarily attributed to the prolonged snow retention characteristic of the Gangwon region, where snow cover persists well into the spring, particularly in mountainous areas. This performance limitation appears to be related to the model's inability to properly predict the phenomenon of fuel moisture sensor readings, which remains consistently high during winter and early spring—periods characterized by heavy and prolonged snow cover. The inadequate representation of snow-related conditions in the training dataset likely contributed to this limitation, as the model lacked sufficient exposure to the complex interactions between snow cover, soil moisture, and fuel moisture dynamics that are prevalent in this region.

Figure 9 is a time series graph for observations (blue line) and estimates (red line with 50% transparency) for each of the 11 stations. Overall, the estimates were shown to simulate the time series of observations well. Table 4 summarizes the performance evaluation indices by station. Stations with good performance were BP (R^2 = 0.86, RMSE = 2.10, MAE = 1.48),

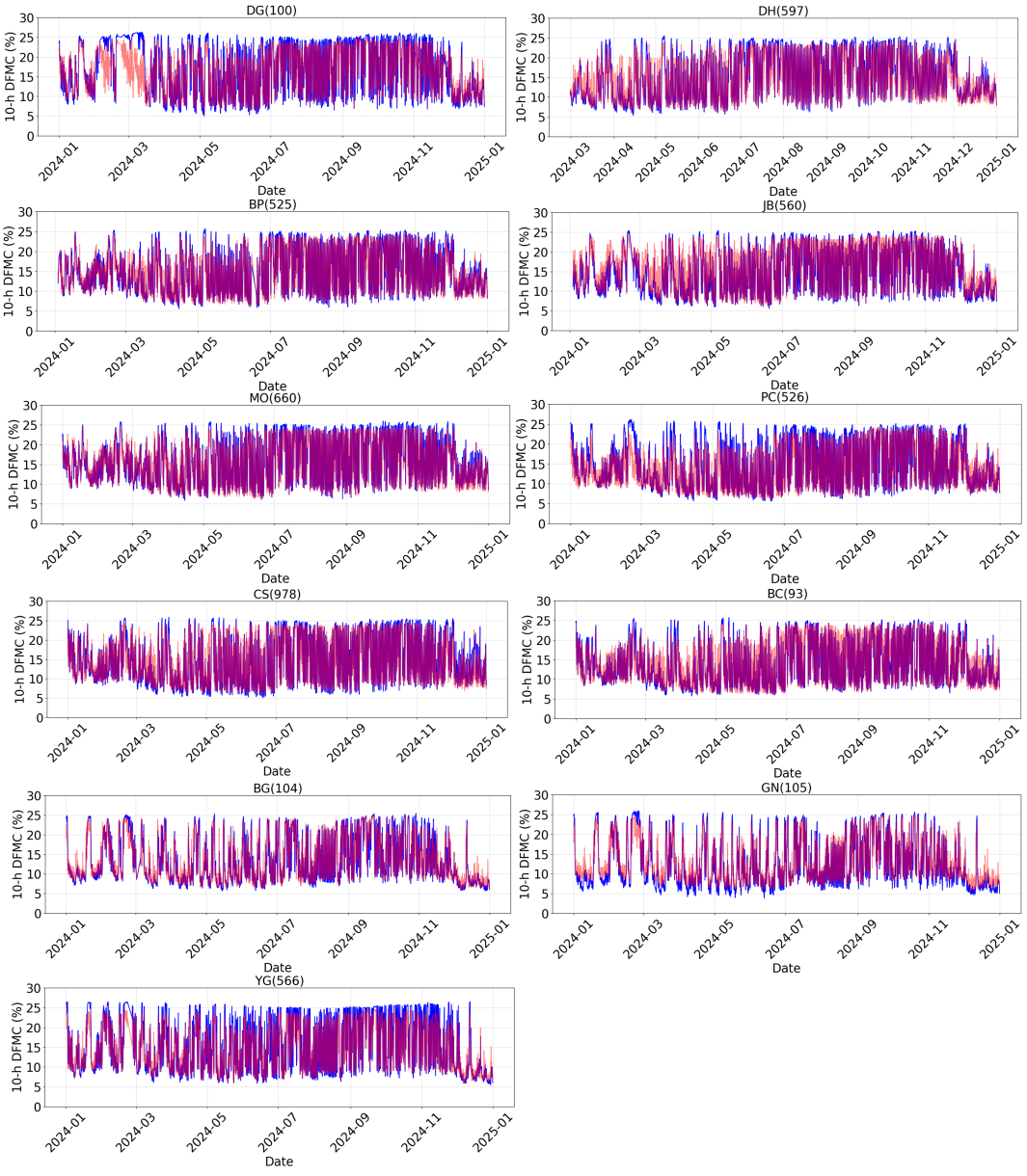

FIG. 9. Time series comparison for 10-h DFMC at 11 sites throughout the year 2024. The blue line represents the observed values, and the red line with 50% transparency represents the predicted values. The purple sections indicate where the observed and predicted values overlap, showing periods of agreement. Each subplot represents a different observation site, identified by its unique code (e.g., DG, DH).

MO (R^2 = 0.85, RMSE = 2.17, MAE = 1.60), BC (R^2 = 0.84, RMSE = 2.16, MAE = 1.53), and CS (R^2 = 0.84, RMSE = 2.40, MAE = 1.66), in that order. The station with the poorest estimation performance was DG (R^2 = 0.76, RMSE = 3.18, MAE = 2.25). As explained earlier, in the case of DG, heavy snowfall during winter and early spring covered the sensors for a long time, resulting in high observed fuel moisture; however, the automatic weather observation sensors record atmospheric relative humidity, temperature, and wind speed outside of the snow cover, which likely caused the estimates to be lower than observations. Examining each station, the ranges of each validation index were as follows: R^2 ranged from 0.76 to 0.86, RMSE from 2.10% to 3.18%, and MAE from 1.48 to 2.24.

Among coastal observation stations, it was uniquely observed at the GN station that the model tended to predict fuel moisture higher than observed values in the low 10-h DFMC range. It is noteworthy that this bias did not appear at the other coastal stations (YG, BG). This is likely due to unique topographical characteristics or local microclimate conditions around the GN station. It appears these characteristics were not sufficiently captured during the model training process.

TABLE 4. Performance evaluation results of 10-h DFMC estimation by each station. The values were rounded at the third decimal place and presented to two decimal places.

Acronym of station	R^2	RMSE	MAE
DG	0.76	3.18	2.25
DH	0.79	2.74	1.95
BP	0.86	2.10	1.48
JB	0.74	2.84	2.13
MO	0.85	2.17	1.60
PC	0.79	2.72	1.82
CS	0.84	2.40	1.66
BC	0.84	2.16	1.53
BG	0.80	2.63	1.81
GN	0.83	2.53	1.91
YG	0.82	2.85	2.15

Such station-specific peculiarities suggest the need for regional adaptation of fuel moisture prediction models.

To reveal the spatial distribution of stationwise skill, we mapped RMSE as bubble size on a georeferenced basemap (Fig. 10). Mountainous sites (e.g., DG, JB) tend to exhibit larger RMSE, whereas low-mountainous/inland-plain sites (e.g., PC, BC, CS) show smaller errors. Coastal stations (BG, GN, YG) are generally lower than mountainous sites, although YG stands out with a comparatively higher error. Exact values and additional metrics (R^2, MAE) are provided in Table 4.

d. Event-based validation and diagnosis at DG

Contrasting a snow-covered, subfreezing winter week (20–27 February 2024) with a snow-free summer week (1–7 August 2024) reveals that the observations form a high plateau (~25%–26%) in winter, while the model systematically underestimates (Fig. 11a). Because the model is driven solely by ambient AWS variables, it cannot represent the very moist microenvironment of an FS-3 probe that is partially/fully embedded in snow. In summer, with no snow cover and strong diurnal forcing, observations and predictions closely track, and the winter bias largely vanishes (Fig. 11b). This contrast indicates that DG underperformance primarily arises from a process mismatch due to missing snow-state information (snow cover/depth) rather than a malfunctioning sensor. Given DG's high-elevation setting and the FS-3 mounting height (20–30 cm above ground), burial following snowfall is physically plausible. Auxiliary variables (snow depth, temperature, wind, precipitation) remain temporally consistent with DFMC behavior, and we found no signatures of failure (e.g., limit flatlining or unrealistic step changes). This aligns with the winter period (February and March) underestimation, seen in the DG time series in Fig. 9.

4. Conclusions and discussion

In this study, we analyzed hourly 10-h DFMC data observed at 11 stations in the Gangwon region throughout 2024 and developed a random forest-based fuel moisture estimation model using automatic weather observation data. The main conclusions and academic contributions of this study are as follows.

The 10-h DFMC at observation stations in the Gangwon region showed distinct seasonal variation patterns. Particularly in spring (March–May) and December, fuel moisture appeared low, which coincides with periods of high forest fire frequency on the Korean Peninsula. Differences according to topographical characteristics were also observed; the DG station, at the highest elevation, maintained higher fuel moisture content in February–March due to the influence of snow. Additionally, coastal areas (BG, GN, YG) showed greater variation in 10-h DFMC in February compared to inland areas (BC, CS), confirming that maritime influence plays an important role in the variability of fuel moisture content.

The 10-h DFMC estimation model based on random forest showed good performance on the test set with R^2 = 0.80, RMSE = 2.73%, and MAE = 1.93. When examined by observation station, R^2 ranged from 0.76 to 0.86, RMSE from 2.10% to 3.18%, and MAE from 1.48 to 2.24. The performance metrics

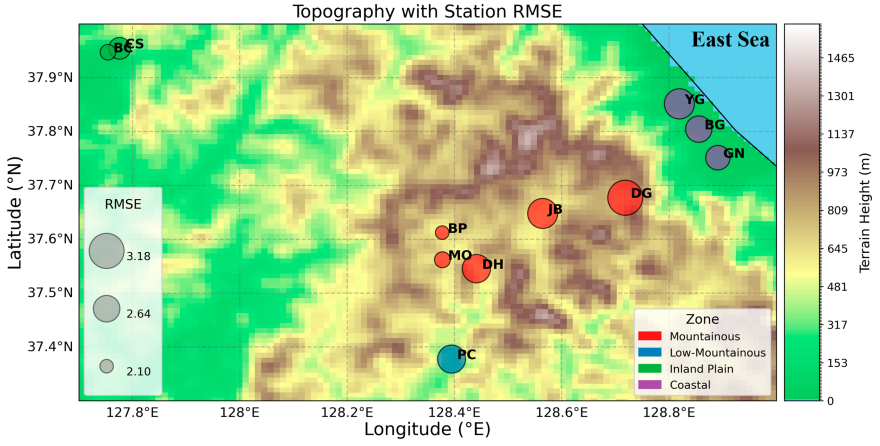

FIG. 10. Bubble map of station-level RMSE for the 10-h DFMC prediction model. Bubble colors indicate terrain classes (mountainous, low-mountainous, inland plain, and coastal), and bubble sizes represent the RMSE magnitude.

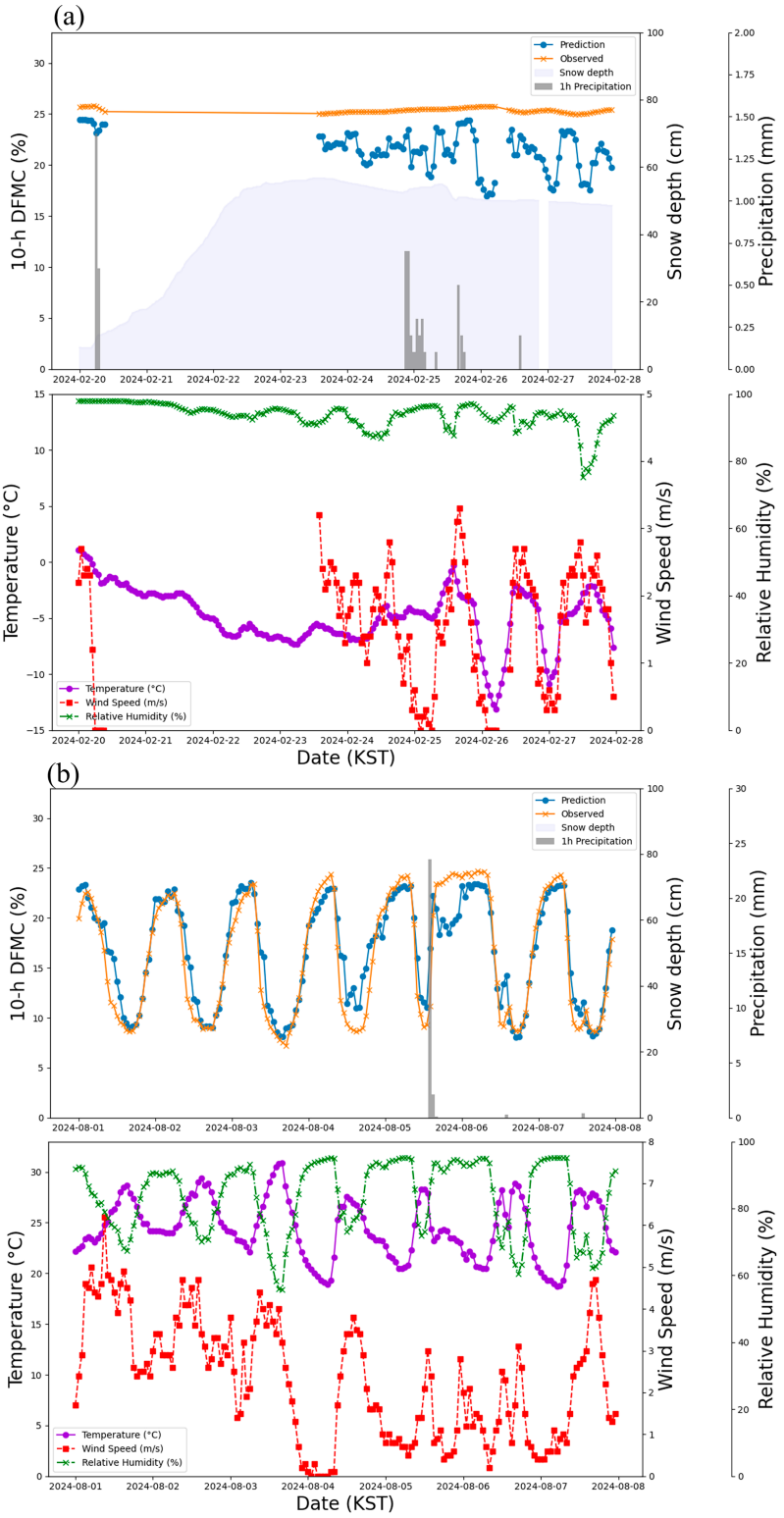

FIG. 11. Event-based validation at DG. (a) Winter case (20–27 Feb 2024): (top) 10-h DFMC predictions (blue) and observations (orange), hourly precipitation (gray bars), and snow depth (blue shading). (bottom) Air temperature (purple), wind speed (red), and relative humidity (green). Under persistent snow and sub-freezing temperatures, observations plateau around 25%–26%, and predictions underestimate. (b) Summer case (1–7 Aug 2024): With no snow, observations and predictions closely track each other, and the winter bias disappears.

of our random forest model (R^2 = 0.76–0.86, RMSE = 2.10–3.18%) are comparable to those reported by Lee et al. (2020) (R^2 = 0.77–0.82, RMSE = 2.05–2.84%), albeit with slightly broader ranges. In addition, our comparative analysis with linear regression, SVR, and GBR confirmed that the RF model clearly outperformed the simpler models and performed nearly indistinguishably from GBR (Table A1 in appendix A). These results demonstrate that machine learning approaches based on random forest algorithms can generate reliable predictions of 10-h dead fuel moisture content specifically calibrated for the forest ecosystems of the Korean Peninsula, with particular applicability to the diverse topographic and climatic conditions of Gangwon Province. Variable importance analysis revealed that relative humidity (0.7562) showed overwhelmingly high influence, followed by time (0.0948), temperature (0.0506), and Julian date (0.0499). The notably low importance of precipitation (0.0042) suggests that 10-h DFMC responds more significantly to sustained atmospheric humidity conditions rather than direct precipitation events.

This study makes the following academic and practical contributions:

First, it provides the first comprehensive analysis of fuel moisture content variability in the unique geographical and climatic environment of the Gangwon region on the Korean Peninsula. This aligns with previous studies that emphasized the importance of developing regionally specialized forest fire risk prediction models in the context of global climate change.

Second, the 10-h DFMC prediction model using the random forest algorithm achieved high accuracy with minimal meteorological variables, demonstrating that effective predictions are possible even without automated DFMC sensor networks in various forest environments. This suggests that forest fire risk prediction systems can be established even in resource-constrained environments.

Third, the analysis of model performance by station clearly shows the complex effects of topography and microclimate on fuel moisture, specifically presenting the need for model adjustments in high mountain areas and coastal regions. This empirically supports previous research suggestions emphasizing the importance of "testing models under various canopy conditions" and "approaches reflecting topographical characteristics."

Fourth, this study provides a foundation for fuel moisture prediction models that consider the unique seasonality and topographical diversity of the Korean Peninsula, establishing scientific evidence directly applicable to forest fire risk forecasting and prevention management.

To maximize operational portability, we restricted the predictors to routinely available AWS variables (temperature, relative humidity, wind speed, and precipitation) and calendar terms (hour of day and Julian day). Solar radiation data were unavailable at our study sites. Station-specific error patterns—including the highest RMSE at the snow-prone, high-elevation DG site and underestimation of low DFMC at GN—suggest that static physiographic and surface-state factors were missing from the current feature set. The event-based validation at DG further confirmed that its winter bias stemmed from snow burial of the probe, not sensor malfunction. Future model improvements should therefore incorporate snow-state indicators, freezing flags, and terrain/vegetation covariates to enhance robustness. Furthermore, based on the model developed in this study, it will be possible to develop an advanced prediction system and establish a real-time forest fire risk assessment system (Lyell et al. 2024).

Although this study focused on the Gangwon Province as a case study, the proposed machine learning model was developed using general meteorological variables that are widely available across regions. As such, the model structure is inherently transferable and could be applied to other regions with appropriate local training data. Future studies should evaluate the model's portability and performance across diverse climatic and ecological zones to assess its broader applicability. This is ultimately expected to contribute to effective forest fire management and damage reduction in response to intensifying wildfires worldwide.

Acknowledgments. This research was supported by the Korea Meteorological Administration Research and Development Program, "Advanced Research on Bio- and Agricultural Meteorology" (Grant KMA2018-00626). We would like to express our gratitude to Researcher Hae Jeong Koo from the same department for her efforts in installing and implementing the in situ 10-h DFMC sensor network in the Gangwon Province, South Korea. We also thank the anonymous reviewers for their constructive comments and suggestions that improved the quality of this manuscript.

Data availability statement. The data that support the findings of this study are available from the corresponding author from NIMS, KMA, Republic of Korea.

APPENDIX A

Comparison of Machine Learning Models

To validate the performance of the random forest (RF) model selected in this study, we conducted a comparative analysis against three representative machine learning approaches: a statistics-based linear regression model, a kernel-based nonlinear model, support vector regression (SVR), and an ensemble-based model, gradient boosting regressor (GBR), which is a representative boosting technique similar to XGBoost.

The results confirmed that the RF model achieved superior or equivalent performance compared to the other methods (Table A1). In particular, the RF model showed an R^2 of 0.80, RMSE of 2.73%, and MAE of 1.93% on the test dataset, substantially outperforming both linear regression (R^2 = 0.56, RMSE = 4.01%, MAE = 3.31%) and SVR (R^2 = 0.65, RMSE = 3.57%, MAE = 2.71%). When compared to the GBR model (R^2 = 0.81, RMSE = 2.61%, MAE = 1.86%), the RF model demonstrated nearly identical performance, indicating that RF is on par with advanced ensemble methods.

TABLE A1. Comparison of the RF model's performance with that of other machine learning models.

Model	Best hyperparameters	Training set			Test set		
		R^2	RMSE	MAE	R^2	RMSE	MAE
Linear regression	fit_intercept: True, normalize: True	0.55	4.02	3.30	0.56	4.01	3.31
SVR	C:1, epsilon: 0.01, gamma: scale, kernel: rbf	0.65	3.57	2.69	0.65	3.57	2.71
GBR	learning_rate: 0.1, max_depth: 5, max_features: None, n_estimators: 100, subsample: 0.7	0.82	2.56	1.82	0.81	2.61	1.85
RF	n_estimators: 500, max_depth = 10, random_state = 42	0.82	2.53	1.80	0.80	2.73	1.93

APPENDIX B

Climatological Context of 2024 (Relative to 1991–2020 Normal Values)

To evaluate representativeness and provide a climatological backdrop for 2024, we compared 2024 conditions with 1991–2020 normals at two long-record, regime-representative sites: Daegwallyeong (DG, mountainous) and Gangneung (GN, coastal). We computed monthly cumulative precipitation and monthly means of temperature, relative humidity (RH), and wind speed (Table B1). Seasonal means follow MAM/JJA/SON/DJF; because the analysis window spans January–December 2024, "winter 2024" aggregates January–February and December 2024, and normal values were aggregated over the same calendar months for comparability.

The climatological assessment for 2024, using two representative stations with long-term data—DG for the mountainous region and GN for the coastal region—is as follows: DG station (mountainous region): 2024 was a significantly drier year than normal, with summer (June–August) precipitation totals being only 48.7% of their respective normals (Table B2). While the annual mean temperature was warmer than normal, precipitation in the fall and winter was either near normal (winter: 99.1%) or above normal (fall: 104.1%), showing distinct seasonal variability. GN station (coastal region): At the GN station, summer precipitation was only 35.7% of normal, indicating an exceptionally dry summer. However, this was followed by anomalously wet conditions in the fall (157.1%) and winter (175.1%).

This analysis clearly demonstrates that 2024 was not a "climatologically normal" year but rather one characterized by distinct, regional anomalies. Particularly, the dry spring and summer conditions provide a crucial context for our DFMC findings, as they validate our model's performance

TABLE B1. Monthly meteorological variables for 2024 compared to long-term climatological normals (1991–2020) at DG (Daegwallyeong) and GN (Gangneung) stations. The variables presented are cumulative precipitation (mm), mean temperature (°C), mean relative humidity (%), and mean wind speed (m s^{-1}), which provide context for the climatological representativeness of the study period.

		Precipitation (mm)		Temperature (°C)		Relative humidity (%)		Wind speed (m s^{-1})	
		DG	GN	DG	GN	DG	GN	DG	GN
January	2024	16.8	46.8	−5.7	2.2	72.7	54.1	3.7	2.6
	Normal	51.7	47.2	−7.0	0.9	66.4	46.7	4.8	3.4
February	2024	119.6	175.3	−2.9	3.6	85.8	75.2	2.6	1.8
	Normal	52.4	51.5	−4.6	2.7	65.7	49.3	4.4	3.0
March	2024	52.7	66.0	0.4	7.0	64.9	58.2	3.5	2.3
	Normal	73.8	66.4	0.4	7.1	65.8	52.8	4.2	2.9
April	2024	53.9	41.9	9.6	15.8	69.6	59.7	2.4	1.9
	Normal	90.0	78.5	6.9	13.0	61.9	52.3	4.3	2.8
May	2024	64.9	76.8	13.2	19.5	66.3	55.8	3.1	2.4
	Normal	113.1	82.8	12.5	18.0	67.5	59.2	3.9	2.5
June	2024	110.9	93.8	18.6	24.3	72.2	64.2	2.2	1.8
	Normal	161.8	116.4	16.2	21.3	79.4	69.3	2.9	1.9
July	2024	211.6	115.7	22.0	27.9	85.7	71.5	3.9	2.4
	Normal	338.2	252.8	19.6	24.7	86.2	74.8	3.3	1.9
August	2024	95.2	87.2	22.8	28.7	84.8	73.4	2.1	1.9
	Normal	357.1	281.5	19.7	25.0	87.4	76.3	2.7	1.9
September	2024	249	479.5	18.5	23.1	89.9	82.8	1.8	1.7
	Normal	254.9	236.8	14.6	20.5	85.5	73.0	2.4	2.1
October	2024	163.6	167.4	10.2	16.7	87.2	74.0	1.7	1.9
	Normal	100.9	116.5	8.7	15.6	76.9	61.6	3.2	2.5
November	2024	29.3	32.9	4.1	11.5	81.8	64.1	2.6	2.2
	Normal	68.7	79.5	2.3	9.5	70.2	52.6	4.3	2.9
December	2024	2.2	16.9	−4.3	3.8	60.4	41.8	4.5	2.8
	Normal	35.8	37.8	−4.5	3.3	66.6	45.6	4.9	3.3

TABLE B2. Annual and seasonal meteorological summary for Daegwallyeong and Gangneung in 2024. For precipitation, the value in parentheses indicates the percentage of the normal (%), while for other meteorological elements, it represents the deviation from the normal.

Station name	Period	Precipitation (mm)	Temperature (°C)	Relative humidity (%)	Wind speed (m s^{-1})
DG	Annual	1169.7 (69%)	8.9 (+1.8)	76.7 (+3.4)	2.8 (−1.0)
	Spring	171.5 (61%)	7.7 (+1.1)	66.0.9 (+1.8)	3.9 (−0.7)
	Summer	417.7 (48.7%)	21.3 (−1.3)	81.6 (−0.3)	2.3 (−0.5)
	Fall	441.9 (104.1%)	9.6 (+1.1)	86.3 (+8.8)	2.0 (−1.3)
	Winter	138.6 (99.1%)	−4.3 (+1.1)	73.0 (+6.8)	3.6 (−0.6)
GN	Annual	1400.2 (97%)	15.3 (+1.8)	64.5 (+5.0)	2.6 (+0.5)
	Spring	184.7 (84.4%)	13.2 (+0.5)	57.9 (+3.1)	2.2 (−0.5)
	Summer	232.3 (35.7%)	25.0 (+0.3)	69.7 (−3.8)	2.6 (+0.6)
	Fall	679.8 (157.1%)	10.9 (+1.4)	73.6 (+11.2)	1.9 (−0.6)
	Winter	239.0 (175.1%)	−0.4 (+4.0)	57.0 (+9.8)	2.3 (−0.9)

under anomalous, high-risk drought conditions. This allows us to both acknowledge the limitation of a single-year dataset and demonstrate the robustness of our model.

REFERENCES

Aguado, I., E. Chuvieco, R. Borén, and H. Nieto, 2007: Estimation of dead fuel moisture content from meteorological data in Mediterranean areas. Applications in fire danger assessment. *Int. J. Wildland Fire*, **16**, 390–397, https://doi.org/10.1071/WF06136.

Bradshaw, L. S., J. E. Deeming, R. E. Burgan, and J. D. Cohen, 1984: The 1978 National Fire-Danger Rating System: Technical documentation. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station General Tech. Rep. INT-16, 49 pp, https://doi.org/10.2737/INT-GTR-169.

Breiman, L., 2001: Random forests. *Mach. Learn.*, **45**, 5–32, https://doi.org/10.1023/A:1010933404324.

Brown, T. P., A. Inbar, T. J. Duff, P. N. J. Lane, and G. J. Sheridan, 2022: The sensitivity of fuel moisture to forest structure effects on microclimate. *Agric. For. Meteor.*, **316**, 108857, https://doi.org/10.1016/j.agrformet.2022.108857.

Carlson, J. D., and R. E. Burgan, 2003: Review of users' needs in operational fire danger estimation: The Oklahoma example. *J. Remote Sens.*, **24**, 1601–1620, https://doi.org/10.1080/01431160210144651.

Cawson, J. G., and Coauthors, 2020: Exploring the key drivers of forest flammability in wet eucalypt forests using expert-derived conceptual models. *Landscape Ecol.*, **35**, 1775–1798, https://doi.org/10.1007/s10980-020-01055-z.

Chae, H. M., 2014: The analysis on moisture contents of forest fuel using weather factors in forest land and weather station. *J. Korean Soc. Hazard Mitigation*, **14**, 205–212, https://doi.org/10.9798/KOSHAM.2014.14.5.205.

Chuvieco, E., D. Biano, I. Aguado, and D. Cocero, 2002: Estimation of fuel moisture content from multitemporal analysis of Landsat Thematic Mapper reflectance data: Applications in fire danger assessment. *Int. J. Remote Sens.*, **23**, 2145–2162, https://doi.org/10.1080/01431160110069818.

Danson, F. M., and P. Bowyer, 2004: Estimating live fuel moisture content from remotely sensed reflectance. *Remote Sens. Environ.*, **92**, 309–321, https://doi.org/10.1016/j.rse.2004.03.017.

Desbois, N., M. Deshayes, and A. Beudoin, 1997: Protocol for fuel moisture contents measurements. "A review of remote sensing methods for the study a large wildland fires", E. Chuvieco, Ed., Universidae de Alacala Rep., 61–72, https://www.researchgate.net/profile/EmilioChuvieco/publication/313057542_Remote_sensing_of_burned_areas_A_review_A_review_of_remote_sensing_methods_for_the_study_of_large_wildland_fires/links/5a09c85ba6fdcc1b976cc7fb/Remote-sensing-of-burned-areas-A-review-A-review-of-remote-sensing-methods-for-the-study-of-large-wildland-fires.pdf#page=65.

Kim, B.-Y., J. W. Cha, K.-H. Chang, and C. Lee, 2021: Visibility prediction over South Korea based on random forest. *Atmosphere*, **12**, 552, https://doi.org/10.3390/atmos12050552.

Korea Forest Service, 2024: Statistical year book of forest fire. Accessed 16 April 2025, https://www.forest.go.kr/kfsweb/kfi/kfs/frfr/selectFrfrStatsNow.do.

Lee, H., M.-S. Won, S.-H. Yoon, and K.-C. Jang, 2019: Modeling and mapping fuel moisture content using equilibrium moisture content computed from weather data of the automatic mountain meteorology observation system (AMOS). *J. Korean Assoc. Geogr. Inf. Stud.*, **22**, 21–36, https://doi.org/10.11108/kagis.2019.22.3.021.

——, M. Won, S. Yoon, and K. Jang, 2020: Estimation of 10-hour fuel moisture content using meteorological data: A model inter-comparison study. *Forests*, **11**, 982, https://doi.org/10.3390/f11090982.

Lyell, C. S., and Coauthors, 2024: Forecasting dead fuel moisture content below forest canopies—A seven-day forecasting system. *Agric. For. Meteor.*, **358**, 110217, https://doi.org/10.1016/j.agrformet.2024.110217.

Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, **12**, 2825–2830.

Resco de Dios, V., A. W. Fellows, R. H. Nolan, M. M. Boer, R. A. Bradstock, F. Domingo, and M. L. Goulden, 2015: A semi-mechanistic model for predicting the moisture content of fine litter. *Agric. For. Meteor.*, **203**, 64–73, https://doi.org/10.1016/j.agrformet.2015.01.002.

Rodrigues, M., V. Resco de Dios, Â. Sil, Á. C. Camprubí, and P. M. Fernandes, 2024: VPD-based models of dead fine fuel moisture provide best estimates in a global dataset. *Agric. For. Meteor.*, **346**, 109868, https://doi.org/10.1016/j.agrformet.2023.109868.

Shmuel, A., Y. Ziv, and E. Heifetz, 2022: Machine-learning-based evaluation of the time-lagged effect of meteorological factors on 10-hour dead fuel moisture content. *For. Ecol. Manage.*, **505**, 119897, https://doi.org/10.1016/j.foreco.2021.119897.

Simard, A. J., 1968: The moisture content of forest fuels. I. A review of the basic concepts. Forest Fire Research Institute

Information Rep., 47 pp., https://scholar.google.com/scholar_lookup?&title=The%20moisture%20content%20of%20forest%20fuels-1.%20A%20review%20of%20the%20basic%20concepts&publication_year=1968&author=Simard%2CAJ.

Viegas, D. X., M. T. S. P. Viegas, and A. D. Ferreira, 1992: Moisture content of fine forest fuels and fire occurrence in central Portugal. *Int. J. Wildland Fire*, **2**, 69–86, https://doi.org/10.1071/WF9920069.

Wright, M. N., and A. R. Ziegler, 2017: Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Software*, **77** (1), 1–17, https://doi.org/10.18637/jss.v077.i01.

Yebra, M., and Coauthors, 2013: A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. *Remote Sens. Environ.*, **136**, 455–468, https://doi.org/10.1016/j.rse.2013.05.029.

Zormpas, K., C. Vasilakos, N. Athanasis, N. Soulakellis, and K. Kalabokidis, 2017: Dead fuel moisture content estimation using remote sensing. *Eur. J. Geogr.*, **8**, 17–32.