Factor Analysis of Weather Conditions Impact on Firefighter Interventions

Naoufal Sirri[1[0000−0002−9900−337X]] and Christophe Guyeux[1[0000−0003−0195−4378]]

FEMTO-ST Institute, UMR 6174 CNRS, University of Franche-Comté, Belfort, France, {naoufal.sirri,christophe.guyeux}@univ-fcomte.fr

Abstract. This research builds upon a series of studies investigating various categories of variables and their implications for firefighting. It is motivated by the recognition that, while weather conditions impact the environment and health, they are a critical factor in firefighting interventions. Firefighters respond primarily to emergencies, prompting an examination of how the weather might influence their operations. Our study aims to thoroughly evaluate the impact of weather conditions on firefighting interventions, with a particular focus on their frequency. The importance of this study is highlighted by the significant effects of flood and accident risks. In France, extreme temperatures have been shown to increase mortality rates. Climate change has caused a rise in the frequency of extreme heat days, doubling the associated mortality, as noted in a 2022 public health survey. This underscores the urgency of our research. Spanning nine years from 2015 to 2024, our methodology involves data preparation, detailed analysis, and the application of the XGBoost predictive model, renowned for its speed and resilience to outliers. The iterative training process selects features that enhance the RMSE score over 24 hours, highlighting the critical importance of weather-related variables across all time frames. Our primary finding is that these variables have an immediate and sustained impact, with notable transparency in mid-term horizons due to the consistent influence of certain variables on interventions. This indicates their increased relevance for predicting outcomes over the entire duration. This refined understanding of the models linked to incident occurrence provides a practical approach for anticipating resource management, enhancing firefighters' response times, and saving lives by reducing intervention failures during major incidents.

Keywords: Firefighters intervention · Climate impact · Feature selection · Data mining · Intervention causes.

1 Introduction

The impact of climatic conditions on the frequency and nature of firefighting interventions, especially in critical areas such as public health and the environment, is substantial. In France, extreme temperatures increase mortality rates. Although the risks have diminished over time, climate change has resulted in more frequent high-temperature days, doubling the mortality associated with such temperatures [1]. However, leveraging weather surveillance data for detecting specific events is challenging due to the intricacies of spatial and temporal variability and the need for precise detection thresholds. This study aims to investigate the potential correlation between climate-related variables and firefighting interventions, contributing to the development of more effective monitoring strategies.

Recent years have seen notable strides in climate prediction and firefighting interventions leveraging artificial intelligence (AI) and its applications. The study [2] utilized AI techniques such as Random Forest (RF), Artificial Neural Network (ANN), and Support Vector Regression (SVR) to predict monthly reservoir inflows, highlighting RF's superior performance and interpretative capabilities. Climate indices further enhanced forecast accuracy by revealing correlations between climate conditions and local hydrology over multiple months. On the other hand, the study [3] implemented a deep learning model for forecasting severe convective weather, utilizing numerical weather prediction data, surpassing traditional methods with enhanced accuracy, and is now operational at the National Meteorological Center of China for guiding severe weather forecasts. However, [4] emphasized the urgent need to integrate climate risk assessment into business operations, highlighting the potential of artificial intelligence, particularly deep learning algorithms, in predicting and managing climate extremes to promote climate-resilient business ecosystems. Furthermore, [5] developed recurrent and convolutional neural networks to extract and predict features of rare and impactful weather events. However, studies on predicting firefighter interventions remain scarce. [6] applied natural language processing techniques to extract features from weather bulletin texts to predict peak intervention periods caused by rare events. Similarly, [7] illustrated the feasibility of maintaining an updated and refined database by employing specialized feature selection tools tailored for XGBoost. Finally, within the same context as our paper, [8] examined the reliability of predictions related to firefighting interventions to optimize resources and response time. They also conducted three separate studies analyzing the impact of specific factors of air quality, solar activity, and river height on firefighting interventions, as detailed in [9], [10], and [11], respectively.

Our research group addresses this issue by adopting an innovative approach to investigate the influence of climate on firefighting interventions. Utilizing extensive historical datasets that include weather information and firefighting activities, we examine this relationship. The primary objective of this paper is to assess the impact of climatic conditions on firefighting interventions, with a particular focus on their frequency. To achieve this, we formulated a hypothesis and its prediction within the framework of our study:

- Hypothesis: Climatic conditions, specifically extreme temperatures and precipitation levels, significantly impact the frequency and nature of firefighting interventions.
  Prediction: An increase in extreme temperatures (both hot and cold) will correlate with a higher frequency of firefighting interventions due to heat-related incidents such as wildfires during hot weather and heating-related fires during cold weather.

To ensure a systematic and organized approach to our research, we meticulously developed and followed a structured methodology. Section 2 provides a comprehensive overview of the methods and materials utilized in our experiments, outlining our experimental strategy. Section 3 details the research process and presents the results obtained. Section 4 offers a thorough analysis and critical discussion of the findings, addressing the initial research question and emphasizing relevant implications and interpretations. Finally, Section 5 concludes the study by summarizing the key findings, highlighting significant contributions, and suggesting directions for future research. This methodological framework enabled a systematic exploration of the relationship between climatic conditions and firefighting interventions, while maintaining a cohesive and comprehensive approach throughout our investigation.

2 Material and Method

2.1 Data Preparation

Data Acquisition This study employs a comprehensive dataset from the Doubs Fire and Rescue Service (SDIS 25), France, which includes 323,225 documented interventions spanning from January 1, 2015, to June 30, 2024. Each intervention entry is meticulously recorded, providing an identification code, precise start and end timestamps, geographical coordinates, intervention type, and response durations. A variety of contextual features have been integrated to thoroughly characterize the circumstances of each incident. These features encompass meteorological conditions, solar activity metrics, river levels in the Doubs department (reflecting the responsibility of French firefighters for various rescues, including floods), air quality indices, epidemiological statistics, holiday periods, lunar phases, and other pertinent factors. To predict future interventions, a comprehensive dataset was constructed by merging the firefighter intervention records with supplementary information from diverse sources. The detailed methodology is outlined below:

- Initially, our meteorological data source was Météo France [22] (the French public weather service). However, challenges arose due to restricted access to remote primary stations and a three-hour sampling interval, which impacted both geographical and temporal precision. Despite Météo France's comprehensive weather risk bulletins, the addition of MeteoStat [23] was essential to overcome these limitations. MeteoStat provides future forecasts, thereby extending the capabilities offered by Météo France.

  - The data extracted from three meteorological stations in the Doubs department includes a comprehensive range of atmospheric measurements. These encompass cloud cover, temperature, barometric trends, atmospheric pressure, humidity levels, dew point, precipitation in the last hour, precipitation in the last three hours, average wind speed recorded every 10 minutes, gust speed over a specified interval, average wind direction tracked every 10 minutes, horizontal visibility, and prevailing weather conditions, all sourced from Météo France.
  - Collection of weather advisories from MétéoFrance encompasses various meteorological risks such as floods, storms, snowfall, wind, heatwaves, severe cold, precipitation, and freezing rain. Each advisory is categorized by color codes (green, orange, red, yellow), enhancing the significance of the meteorological dataset [24].
  - The MeteoStat API was utilized to access climatic variables such as precipitation, wind speed and direction, temperature, snowfall, dew point, pressure, and humidity from freely available meteorological and climatic datasets. Temperature data was extracted from a wide network of 11x11 grids covering the entire department.
- To assess the influence of air quality, we utilized features such as particulate matter (PM2.5, PM10), ozone (O_3), and nitrogen oxides (NO_2) from various nearby air quality monitoring stations, as detailed in [14].
- To assess the influence of solar activity, features such as the X-ray emissions, the 10 cm radio flux, sunspot area, and sunspot number were incorporated, as detailed in [13].
- Continuous data from NASA's VIIRS and MODIS satellites were collected, capturing Earth images with varying wavelengths and resolutions to analyze fire spread in specific areas, as shown in [15] and [16].
- The [17] libraries were used to calculate spatial separations between Earth, Moon, and Sun. Astral [18] was employed to analyze various parameters related to the Sun and Moon, including lunar phases and moonrise, to assess their influence on natural disasters. Additionally, sunrise and sunset data were utilized to establish a boolean variable signifying "night" or "day".
- Data on water levels from the initial forty rivers of the Doubs were sourced from the government service "Hydroreel" and integrated into the study. The process involved populating a dictionary with the closest average measurements for each hourly interval, as detailed in [12].
- Incorporation of weekly epidemiological data from the Sentinelles network, covering conditions such as acute diarrhea, influenza, and chickenpox, as documented in [19].
- Integration of temporal details including year, month, day, hour, day of the week, and day of the year, alongside information about holidays, academic breaks sourced from [21], and events such as Ramadan observances, closures, and curfews.
- Integration of variables related to French league football and Champions League matches, identified as potential factors influencing interventions, as documented in [20].

The chosen variables, identified through an analysis of firefighter responses, are intended to gauge their potential impact. This comprehensive approach allows for the inclusion of possibly non-significant variables while facilitating the assessment of relationships among factors linked to meteorological conditions and other parameters. These variables encompass a range of risks including accidents, fires, and floods, central to the dataset on firefighter interventions.

Data Pre-processing During the data processing phase, linear interpolation was applied to address missing values in specific meteorological datasets. To refine our learning model, we employed two techniques from the Scikit-learn library [25]. The "StandardScaler" method was used to normalize numerical features, including variables such as wind speed, wind direction, gusts, humidity, cloud cover, visibility, temperature, precipitation, dew point, acute diarrhea, influenza data, chickenpox, year, hour, lunar distance, and river levels. This approach adjusts the value distribution to achieve a mean of zero and a unit variance.

Furthermore, we utilized the "TargetEncoder" technique [26] to encode categorical attributes such as year, month, day of the week, barometric trend, events, and holidays. This method involves transforming these variables by replacing each category with the average of the corresponding target variable. We preserved the original target values (number of interventions) as discrete entities, as they provide a more accurate representation of intervention distribution.

Data Mining A comprehensive review of the dataset was essential for extracting relevant information for our study. On average, approximately 30,000 interventions were recorded annually, with a notable upward trend in intervention frequency over time. Table 1 presents the key statistical measures for various parameters within the climate variable family. This table provides a concise summary of central tendencies, variability, and distribution patterns observed across different categories of climate-related variables. All these variables are represented as continuous entities, primarily exhibiting a right-skewed distribution.

Table 1: Data analysis of climate variables

Variable	**Mean**	**Std**	**IQR**	**Max**
humidity (g/m^3)	9.4	2.6	3.1	15.1
dewPoint (°C)	10.2	3.1	3.8	16.6
lastHourRainfall (mm)	2.6	1.4	1.8	5.2
last3hHourRainfall (mm)	5.4	2.2	3.0	10.2
temperature (°C)	15.1	5.1	7.2	27.3
snow (cm)	16.2	6.5	8.2	22.3
windDirection (°)	177.6	42.8	61.9	264.5
precipitations (mm)	50.3	20.1	24.3	94.2

Furthermore, examining these variables through time series analysis reveals a significant degree of similarity. To support this observation, we conducted a correlation analysis among these variables, the results of which indicate substantial correlations. These correlations range from 0.65 to 0.88, suggesting that these variables provide redundant information for prediction and indicating the existence of a common influencing factor.

To conclude our analysis, we assessed the impact of weather conditions on the overall target variable through correlation analysis. The results showed a weak and statistically insignificant effect on the overall target variable. These findings offer an initial glimpse into the limited role of weather conditions in shaping intervention patterns. Although no significant correlations were identified, the possibility of a more complex relationship underscores the nuanced interaction between climate variables and incidents requiring emergency services.

2.2 Feature and Model Selection

Previously, it was common to include all available features in the training dataset based on the assumption that more information would lead to an optimal model. However, there are two primary reasons to limit the number of features. First, some variables may have strong interdependencies, while others may contribute minimally to predictive power, potentially decreasing model generalization or introducing redundant information. Second, incorporating numerous features can significantly increase computational complexity without proportionately enhancing model performance. Thus, utilizing a more concise set of features is promising for achieving more efficient outcomes. In this study, we used the "feature importance" method for feature selection, assigning scores to each variable in the dataset, with higher scores indicating greater relevance [28]. A threshold was established to retain the 400 most pertinent features. Various selection techniques were applied, including:

1. High Variance: retaining features with variances greater than 0.5.
2. Pearson and Spearman correlation coefficients: excluding correlations with the target variable where the absolute value is 0.4 or lower.
3. Chi-Square Selector: used the chi-square test to assess the association between each feature and the target variable after normalizing the features with the 'Min-Max Scaler' function.
4. Extreme Gradient Boosting (XGBoost) [29]: utilized preset hyperparameters (maximum depth = 7, number of estimators = 100000, early stopping after 10 rounds), performed training, and then calculated or derived feature importance.
5. Light Gradient Boosting Model (LightGBM) [30]: applied defined hyperparameters (learning rate = 0.1, objective function = regression, metric = RMSE, number of leaves = 2^7, maximum depth = 7, number of estimators = 100000, early stopping after 10 rounds), and obtained feature importance as previously described.

Following this selection process, a refined list of features was derived, emphasizing those consistently identified by multiple techniques. Out of the initial 3,912 features, approximately 10% were retained, resulting in a subset of 400 features utilized for model training.

In our model selection process, we have chosen Extreme Gradient Boosting (XGBoost) [29] for its renowned scalability, efficiency, flexibility, and speed. XGBoost excels in handling complex datasets by integrating gradient boosting principles, continuously improving the performance of weaker learners. Its widespread adoption spans various domains, including finance and healthcare, highlighting its versatility in tackling regression, classification, and ranking tasks with exceptional precision. Notably, XGBoost's capabilities extend to managing missing data, conducting feature selection, and adapting seamlessly to diverse data structures and characteristics. Its robustness and effectiveness in predictive modeling make it a preferred choice for advanced machine learning applications.

2.3 Strategies Employed for The Prediction Tool

To translate our theoretical analysis into practical implementation, we established a robust training and optimization pipeline (see Figure 1).

1. We conducted multiple training iterations across different timeframes, spanning from 1 to 24 hours, to assess the climatic impact up to 24 hours ahead.
2. Following that, we conducted feature extraction as detailed in the "Feature Selection" section.
3. In each iteration, we integrated every selected variable from the previous phase into the XGBoost model for training purposes.
4. In each iteration, we divided the dataset into training, testing, and validation subsets. Initially, 20% of the data was reserved for testing, with the remaining 80% used for training and validation. Later adjustments were made to allocate 80% for training and 20% for validation.
5. We trained the XGBoost model using predefined hyperparameters (refer to Table 2) and implemented early stopping with "early_stopping_rounds=15". This approach allowed us to stop training if there was no improvement in performance on the validation set for 15 consecutive rounds, influenced by the fixed 100,000 estimators in our model.
6. We computed prediction scores using the RMSE metric (Equation 1) and compared them with previous results. Variables that demonstrated improvement were retained in the input list through sequential feature selection, and adjustments were made to the preceding scores accordingly.

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

7. Ultimately, we pursued fine-tuning of prediction outcomes by experimenting with different combinations of hyperparameters (refer to Table 3). This

effort involved using "early_stopping_rounds=25" in conjunction with a fixed parameter set of 100,000 estimators. Through our iterative approach, we identified optimal hyperparameter configurations that significantly enhanced performance.

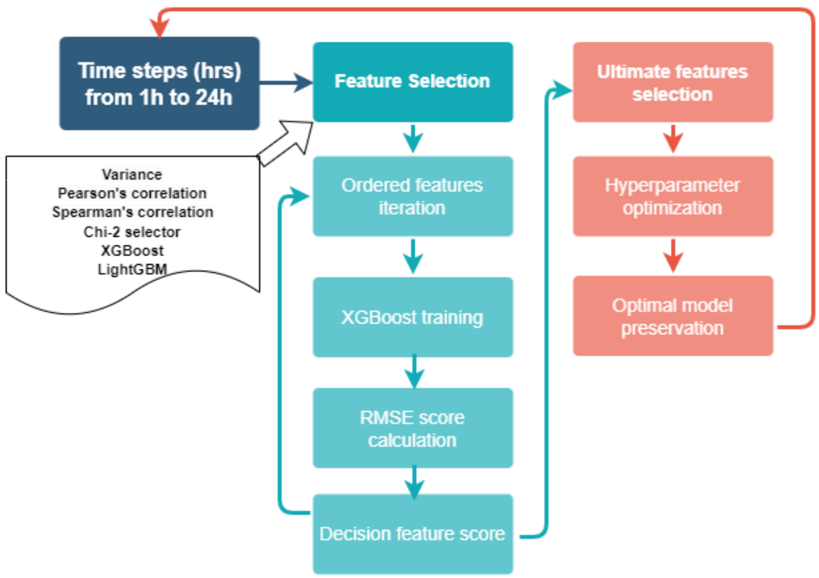


Fig. 1: Training pipeline representation model

Table 2: XGBoost default hyperparameters

Hyperparameter	**Values**
max_depth	7
min_child_weight	1
gamma	0
subsample	0.8
colsample_bytree	0.8
learning_rate	0.1

Table 3: Settings XGBoost hyperparameters

Hyperparameter	**XGBoost**
`max_depth`	[2, 14]
`min_child_weight`	[0, 14]
`gamma`	[0.0, 0.4]
`subsample`	[0.6, 0.9]
`colsample_bytree`	[0.6, 0.9]
`learning_rate`	[0.01, 0.009]

3 Results

Extensive efforts have been made to compile, process, and consolidate a comprehensive dataset, emphasizing the integration of information from various sources regarding past interventions. Special attention was given to evaluating the significance of each variable within this context. As outlined earlier, the core of this study is to uncover pertinent insights about the impact of climatic conditions on the prediction of interventions. The previous section detailed a methodology for identifying key features. Following the training of our model over 24 horizons, we present the outcomes of the feature selection process, emphasizing the weather-related variables that contributed to improved prediction accuracy (see Table 4).

Furthermore, the classification results generated by our model are presented and ready for interpretation in the subsequent section. Table 5 details the predictive accuracy of the XGBoost model regarding the number of interventions across various temporal horizons, incorporating weather variables, with the most significant performances highlighted in bold. It also shows the RMSE values before and after the inclusion of weather variables, along with the percentage improvement. Figure 2 displays the outcomes from 300 samples aimed at predicting atypical intervention counts for the fourteenth hour. A notable RMSE improvement of 6.53% is achieved by utilizing a single weather variable in the midst of horizons, compared to incorporating alternative variables in preceding horizons. Figures 3(a) and 3(b) provide a visual representation of prediction accuracy, illustrating a maximum deviation range of 0 to 17 errors for the second hour before and after integrating the weather variable, respectively. As the XGBoost model produces decimal predictions (e.g., 7.16 interventions), the results have been rounded to the nearest whole number (in this case, 7 interventions) to ensure practical relevance.

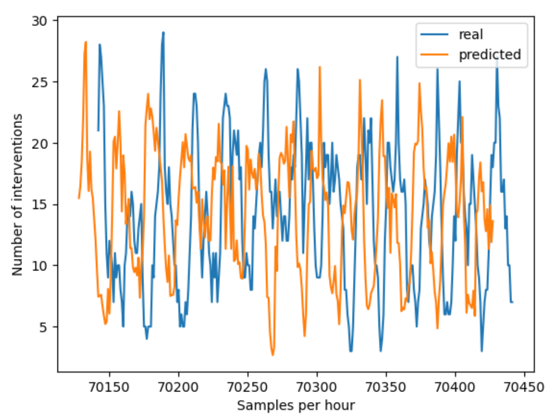


Fig. 2: Prediction for the 14th hour

Significantly, XGBoost achieved its best RMSE score of 2.2046 at the fourteenth hour, notably influenced by the inclusion of weather variables, particularly temperature. The improvement in scores began early in the time horizons and persisted due to the recurrent selection of this variable. Weather-related variables such as precipitations and dew point consistently contributed to score improvements every two hours, with

Table 4: Exploring the impact of Feature Selection on the analysis of climate

Time horizon	**Height of rivers**	**Feature Selection Technique**	**Rank**
1h	humidity	Variance and XGBoost	203
	precipitations	Variance,Chi-Square and LightGBM	188
	dewPoint	Variance, XGBoost and LightGBM	165
3h	temperature	Variance, Spearman and LightGBM	209
	precipitations	Variance,Chi-Square and LightGBM	172
	dewPoint	Variance, XGBoost and LightGBM	160
	lastHourRainfall	Variance, Pearson and LightGBM	142
5h	temperature	Variance, Spearman and LightGBM	205
	precipitations	Variance,Chi-Square and LightGBM	154
	dewPoint	Variance, XGBoost and LightGBM	146
	lastHourRainfall	Variance, Pearson and LightGBM	129
	last3hHourRainfall	Variance, XGBoost and LightGBM	120
6h	windDirection	Variance and LightGBM	182
	precipitations	Variance,Chi-Square and LightGBM	142
	lastHourRainfall	Variance, Pearson and LightGBM	121
	last3hHourRainfall	Variance, XGBoost and LightGBM	112
8h	humidity	Variance and XGBoost	172
	windDirection	Variance and LightGBM	166
	temperature	Variance, Spearman and LightGBM	152
	last3hHourRainfall	Variance, XGBoost and LightGBM	99
10h	humidity	Variance and XGBoost	162
	dewPoint	Variance, XGBoost and LightGBM	110
	last3hHourRainfall	Variance, XGBoost and LightGBM	91
14h	temperature	Variance, Spearman and LightGBM	102
17h	temperature	Variance, Spearman and LightGBM	96
	dewPoint	Variance, XGBoost and LightGBM	81
	precipitations	Variance,Chi-Square and LightGBM	77
19h	temperature	Variance, Spearman and LightGBM	91
	dewPoint	Variance, XGBoost and LightGBM	80
	precipitations	Variance,Chi-Square and LightGBM	73
21h	humidity	Variance and XGBoost	82
	precipitations	Variance,Chi-Square and LightGBM	76
	dewPoint	Variance, XGBoost and LightGBM	63
	temperature	Variance, Spearman and LightGBM	52
	windDirection	Variance and LightGBM	43
23h	humidity	Variance and XGBoost	75
	dewPoint	Variance, XGBoost and LightGBM	63
	temperature	Variance, Spearman and LightGBM	42

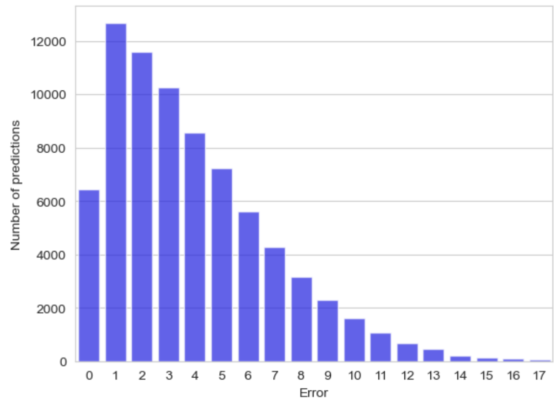
a significant enhancement observed in half of the horizons. Following this, XGBoost was comprehensively trained across all horizons using a detailed grid search procedure for parameter tuning, systematically exploring various combinations to determine the optimal solution. Table 6 provides a comprehensive overview of the hyperparameters identified and employed for achieving the best RMSE score and the highest percentage improvement at the fourteenth hour, ensuring the model's peak performance.

Table 5: Insights into classification outcomes

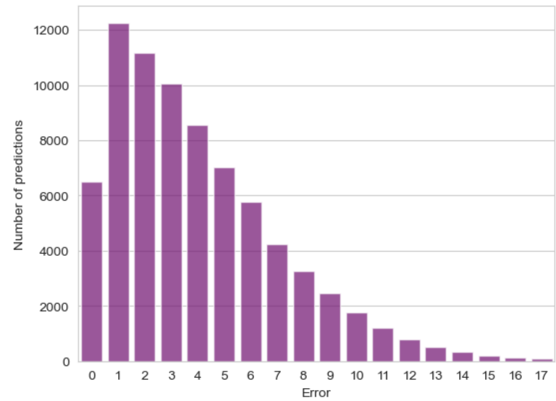
Time horizon	Weather	RMSE pre-sel	RMSE post-sel	Improvement
1h	humidity	3.1501	3.1346	0.49%
	precipitations	3.0708	3.0303	1.31%
	dewPoint	2.9708	2.9538	0.57%
3h	temperature	3.6304	3.6139	0.45%
	precipitations	3.5028	3.4856	0.49%
	dewPoint	3.4856	3.4556	0.86%
	lastHourRainfall	3.3702	3.3504	0.59%
5h	temperature	3.8566	3.8355	0.55%
	precipitations	3.7833	3.7634	0.52%
	dewPoint	3.6407	3.6062	0.95%
	lastHourRainfall	3.4010	3.3795	0.63%
	last3hHourRainfall	2.3840	2.3677	0.68%
6h	windDirection	3.7445	3.7182	0.70%
	precipitations	2.7236	2.6401	3.16%
	lastHourRainfall	2.6022	2.5492	2.07%
	last3hHourRainfall	2.5324	2.4950	1.49%
8h	humidity	3.8538	3.7571	2.57%
	windDirection	3.5660	3.5468	0.54%
	temperature	2.7296	2.6501	3.00%
	last3hHourRainfall	2.4555	2.4250	1.25%
10h	humidity	3.8945	3.8366	1.50%
	dewPoint	2.5761	2.4289	6.06%
	last3hHourRainfall	2.3588	2.3383	0.87%
14h	temperature	2.3486	**2.2046**	**6.53%**
17h	temperature	3.9290	3.8945	0.88%
	dewPoint	3.6809	3.6506	0.83%
	precipitations	3.6506	3.6388	0.32%
19h	temperature	3.8838	3.8466	0.96%
	dewPoint	3.6575	3.6242	0.91%
	precipitations	3.7825	3.7445	1.01%
21h	humidity	4.1144	4.0427	1.77%
	precipitations	4.0427	4.0060	0.91%
	dewPoint	3.7203	3.6924	0.75%
	temperature	3.6117	3.5904	0.59%
	windDirection	3.5904	3.5786	0.32%
23h	humidity	4.0203	4.0034	0.42%
	dewPoint	3.9143	3.8717	1.10%
	temperature	3.6878	3.6392	1.33%

4 Discussion

This study aimed to evaluate the impact of changing climatic conditions on the prediction of firefighter interventions over nine years, from 2015 to 2024. Our methodological framework was carefully designed, involving detailed data preparation and comprehensive analysis to guide our experimental decisions. We selected features through various



(a) Prediction's error at the 14th hour pre-selection



(b) Prediction's error at the 14th hour post-selection

Fig. 3: Insights from predictive modeling

Table 6: Best hyperparameter for the 14th hour

Hyperparameter	**Values**
max_depth	5
min_child_weight	1
gamma	0.2
subsample	0.9
colsample_bytree	0.6
learning_rate	0.1

statistical techniques and machine learning methods, with XGBoost chosen as the predictive model for its robustness in handling outliers, which is particularly beneficial for continuous weather variables (see Table 1). Our approach included an iterative learning pipeline, where features that enhanced the RMSE score over a 24-hour horizon were systematically selected (see Figure 1). The findings, detailed in Table 5, underscore the importance of weather-related variables, especially from the initial stages and throughout extended periods. Notably, at the 14th hour, the inclusion of the "temperature" variable significantly contributed to an RMSE score of 2.2046 and the highest improvement of 6.53%, impacting predictions from the 3rd hour to the final hours. Additionally, variables such as "dew point" and "precipitations" consistently improved the results every two hours during prolonged periods.

This observation indicates a consistent and gradual impact of climatic variables throughout all time periods, underscoring their enhanced usefulness for both short-term and long-term forecasts. Our results demonstrate that certain weather-related variables significantly influence predictions across all horizons, owing to diverse temporal and atmospheric factors. From these results, it can be inferred that the presence of variables such as "precipitation," "dew point," and "humidity" is associative. Precipitation can affect the probability and severity of fires, with heavy rainfall potentially reducing fire risk by increasing soil and vegetation moisture, whereas insufficient rainfall can heighten fire risk by drying out fuels. The dew point, a measure of air moisture, signifies high humidity when elevated, which can facilitate fire suppression by slowing fire spread in humid air. However, it can also increase the likelihood of fog, impacting visibility and potentially leading to accidents requiring firefighter intervention. Additionally, rainfall can have a mid-horizon impact if it is intense. This effect is evident in our results, as rainfall improved outcomes from the outset until its impact diminished midway, where we observed the best improvement. This conclusion can aid firefighters in preparing and optimizing resources at critical times for flood or landslide issues. In summary, these factors lead us to conclude that climatic conditions significantly impact the environment, often resulting in dangerous consequences that systematically necessitate firefighter interventions.

The investigation aims to understand the scarcity of selected variables from the weather conditions category that could improve prediction accuracy. For instance, considering the prediction improvement to 2.2046 after selecting the "Temperature" variable at the fourteenth hour, it becomes evident that other features from the same category provide redundant information and thus cannot be selected as they do not enhance the prediction outcome. Furthermore, as shown in Figure 4, calculating the correlation of variables from different categories with the climatic conditions feature reveals a high correlation among variables within the same family, placing them at the forefront. A further detailed analysis of prediction improvement after selecting the most relevant variables from the weather conditions category, as illustrated in Figure 5, shows a consistently higher average prediction improvement during winter. This strongly supports the inclusion of variables from this category during this season to achieve promising forecasting results.

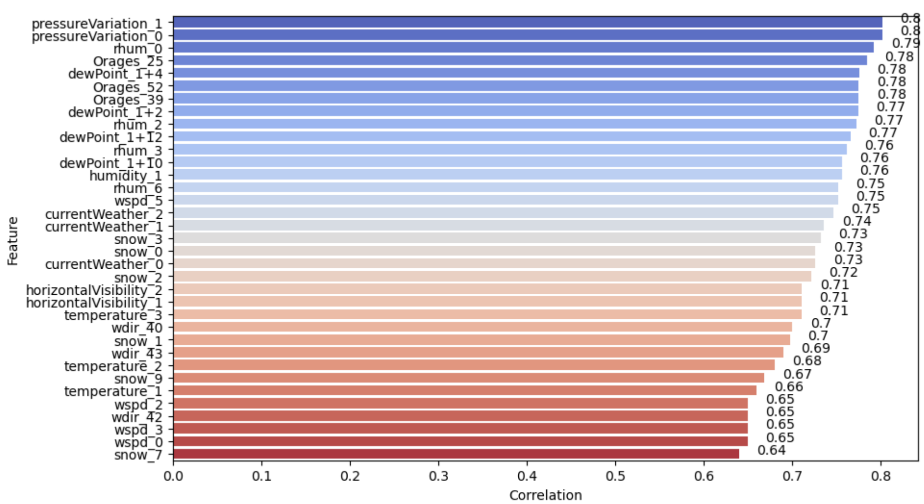


Fig. 4: Correlation with the "Temperature" at the 14th hour

A detailed analysis, as shown in Figure 3, confirms a reduction in prediction errors, validating our findings. Specifically, the figure illustrates a decrease in errors, with an increase in zero errors and a reduction in one and two errors after incorporating

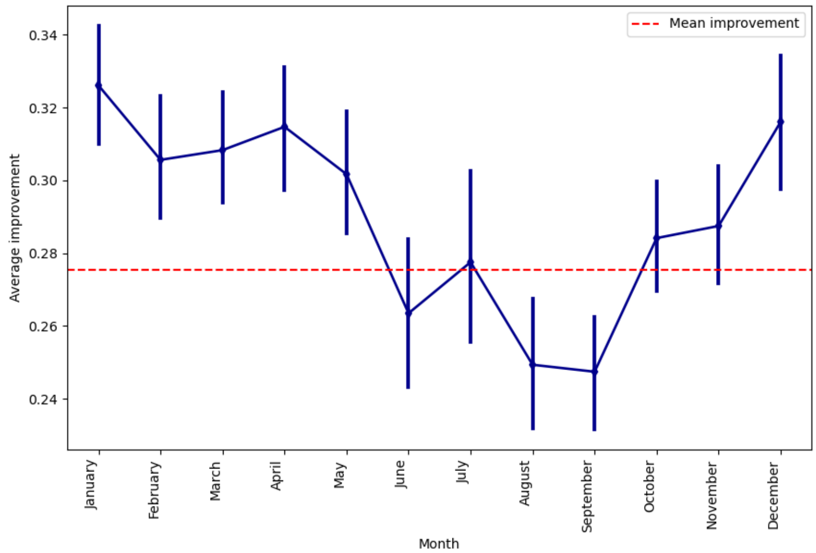


Fig. 5: Improvement statistics per month

the weather variable "Temperature" compared to the errors before its inclusion. This study deepens our understanding of the impact of climatic conditions on firefighter interventions, offering new insights. However, our results are tempered by limitations such as the lack of comparison with other boosting models or machine learning methods and the exclusive use of grid search for hyperparameter optimization. Future work will explore additional climatic variables, employ alternative models, and adopt more advanced sequential optimization methods, such as Bayesian optimization.

5 Conclusion

This study extends a series of articles that investigate various categories of variables and their implications for firefighting interventions. By focusing on predicting the number of future interventions, a crucial aspect for fire services globally, this paper examines the impact of weather conditions on emergency service forecasts. Using a comprehensive nine-year dataset from the Fire and Rescue Service (SDIS 25) in the Doubs region of France, the research explores how the predictive model identifies significant trends related to weather variables at different times of the day. The results indicate that these variables can have a sustained impact from the beginning and over an extended period. Notably, the variables "precipitation", "dew point" and "humidity" were associated from the start, leading to significant improvements in mid-term forecasts and continued modest improvements over longer intervals. Although this approach is valuable for resource management and optimizing response times, continuous improvements are necessary for seamless integration into firefighting decision-making processes.

Continuing our research efforts, we will explore the nuanced impact of additional variables within the same category of weather conditions on forecasts across various time horizons. Our goal is to extract valuable insights from these areas and identify the most influential variables based on the types of interventions and forecasting periods. Leveraging this information, we aim to refine our forecasts, both short- and long-term, by selecting models and hyperparameters that align with the increased relevance of specific variables. Additionally, we will focus on integrating large language models (LLMs) to enhance predictions based on query submissions and the use of different agents to obtain adaptive responses. Our primary objective is to improve our predictive capabilities, enabling the implementation of operational strategies that proactively respond to the demand for firefighting services.

References

1. Public Health France. https://www.santepubliquefrance.fr/les-actualites/2022 /climat-et-sante-l-evolution-des-temperatures-a-t-elle-un-impact-sur-la-mortalite-en-france. Accessed on 10 July 2024.
2. Yang TianTian, Yang TianTian and Asanjan, AA and Welles, E and Gao XiaoGang, Gao XiaoGang and Sorooshian, S and Liu XiaoMang, Liu XiaoMang, *Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information*, *Advancing earth and space sciences*, vol. 53, pp. 2786-2812, 2017.
3. Zhou, Kanghui and Zheng, Yongguang and Li, Bo and Dong, Wansheng and Zhang, Xiaoling, *FJournal of Meteorological Research*, vol. 33, pp. 797–809, 2019, Publisher: Springer.
4. Singh, Shivam and Goyal, Manish Kumar, *Enhancing climate resilience in businesses: The role of artificial intelligence*, *Journal of Cleaner Production*, vol. 418, pp. 138228, 2023, Publisher: Elsevier.
5. Fang, Wei and Xue, Qiongying and Shen, Liang and Sheng, Victor S, *Survey on the application of deep learning in extreme weather prediction*, *Atmosphere*, vol. 12, pp. 661, 2021, Publisher: MDPI.
6. Cerna, Selene and Guyeux, Christophe and Laiymani, David. *The usefulness of NLP techniques for predicting peaks in firefighter interventions due to rare events.* In: *Neural Computing and Applications*, vol. 34, no. 12, pp. 10117–10132, 2022, Publisher: Springer.
7. Guyeux, Christophe and Makhoul, Abdallah and Bahi, Jacques M. *How to build an optimal and operational knowledge base to predict firefighters' interventions.* In: *Proceedings of SAI Intelligent Systems Conference*, Springer, 2022, pp. 558–572.
8. Sirri, Guyeux, *Firefighter Intervention Predictive Modeling: Reliability Assessment.* In: *International Conference on Circuit, Systems and Communication*, IEEE , 2024. Accepted for presentation on 28 June 2024.
9. Sirri, Guyeux, *Air Quality Impact on Firefighter Interventions: Factors Analysis.* In: *The 7th International Conference on Big Data and Internet of Things*, Springer, 2024. Accepted for presentation on 17 April 2024.
10. Sirri, Guyeux, *Solar activity Impact on Firefighter Interventions: Factors Analysis.* In: *The 5th International Conference on Deep Learning and Applications*, Springer, 2024. Accepted for presentation on 11 July 2024.
11. Sirri, Guyeux, *River Levels Affecting Firefighter Interventions: Factor Analysis.* In: *The 28th International Database Engineered Applications Symposium*, Springer, 2024. Proposed for presentation on 28 August 2024.
12. Ministry of Ecological Transition. http://www.hydro.eaufrance.fr/. Accessed on 10 July 2024.

13. NASA. https://www.swpc.noaa.gov/. Accessed on 10 July 2024.
14. ATMO-BFC. https://www.atmo-bfc.org/accueil. Accessed on 10 July 2024.
15. NASA. https://lance.modaps.eosdis.nasa.gov/viirs/. Accessed on 10 July 2024.
16. MODIS. https://lance.modaps.eosdis.nasa.gov/modis/. Accessed on 10 July 2024.
17. Skyfield. https://github.com/skyfielders/python-skyfield. Accessed on 10 July 2024.
18. Astral. https://pypi.org/project/astral/0.5/. Accessed on 10 July 2024.
19. The Sentinel Network. https://www.sentiweb.fr/?page=table. Accessed on 10 July 2024.
20. Soccer. https://www.footendirect.com/. Accessed on 10 July 2024.
21. Ministry of National Education. http://www.education.gouv.fr/pid25058/le-calendrier-scolaire.html. Accessed on 10 July 2024.
22. Météo-France. https://www.ecologie.gouv.fr/. Accessed on 10 July 2024.
23. Meteo-Stat. https://pypi.org/project/meteostat/. Accessed on 10 July 2024.
24. Vigilance-France. https://vigilance.meteofrance.fr/fr. Accessed on 10 July 2024.
25. Pedregosa, Fabian and Varoquaux, Gael and Gramfort, Alexandre and Michel, Vincent and Thirion, Bertrand and Grisel, Olivier and Blondel, Mathieu and Prettenhofer, Peter and Weiss, Ron and Dubourg, Vincent and others. *Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research*, 12, 2011, pp. 2825–2830.
26. Target Encoder. https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.TargetEncoder.html. Accessed on 10 July 2024.
27. Garreta, Raúl and Moncecchi, Guillermo. *Learning scikit-learn: machine learning in Python: experience the benefits of machine learning techniques by applying them to real-world problems using Python and the open source scikit-learn library.* (No Title), 2013.
28. Zien, Alexander and Krämer, Nicole and Sonnenburg, Sören and Rätsch, Gunnar. *The feature importance ranking measure.* In: *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2009, Bled, Slovenia, September 7-11, 2009, Proceedings, Part II*, pp. 694–709, 2009, Springer.
29. Chen, Tianqi and Guestrin, Carlos. *Xgboost: A scalable tree boosting system.* In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, 2016.
30. Ke, Guolin and Meng, Qi and Finley, Thomas and Wang, Taifeng and Chen, Wei and Ma, Weidong and Ye, Qiwei and Liu, Tie-Yan. *Lightgbm: A highly efficient gradient boosting decision tree.* In: *Advances in neural information processing systems*, volume 30, 2017.