Medium-Term Prediction for Ambulance Demand of Heat Stroke using Weekly Weather Forecast

Tetsuya Nakai [1], Sachio Saiki [2], Masahide Nakamura [1,3]
[1]Graduate School of System Informatics, Kobe University, 1-1 Rokkodai-cho, Nada, Kobe, 657-0011, Japan
[2]Kochi University of Technology 185 Miyanokuchi, Tosayamada Kami City, Kochi, 782-8502, Japan
[3]Riken AIP, 1–4–1 Nihon-bashi, Chuo-ku, Tokyo 103–0027
Email: [1]manda@ws.cs.kobe-u.ac.jp, [2]saiki.sachio@kochi-tech.ac.jp, [3]masa-n@cs.kobe-u.ac.jp

***Abstract*—In our joint research with the Kobe City Fire Department, we have been studying the effective use of emergency resources. In our previous research, we developed a prediction model for the number of heat stroke victims in Kobe City by applying machine learning to past weather observation data and emergency dispatch records. It takes a certain period of time to prepare the necessary emergency measures in the field. Therefore, it is necessary to make medium-term forecasts for the next week or so, and the accuracy of such forecasts is not known. In addition, conventional models refer to items of past weather data, which are difficult to obtain from current weather forecasts. Therefore, in this paper, we investigate a new method for predicting the number of heat stroke victims in the medium term by using past weather forecast data. From the proposed method, we know that it is possible to predict the number of heat stroke victims in 7 days by utilizing the actual weather forecast data in Kobe City.**

***Index Terms*—heat stroke, ambulance, smart city, demand prediction, machine learning**

I. Introduction

With the development of the ICT systems and IoT technologies, many companies and local governments are acquiring and accumulating a wide variety of data, and utilization for town development is increasing. Some data are open to the general public as open data and people can use it for innovation. By utilizing the data, aiming for a more efficient and sustainable city, which is so-called *smart city* [1], is now a global trend.

As part of its efforts, the Kobe City Fire Department records emergency dispatch logs using an ICT system. The emergency dispatch log contains the data of the dispatching team and vehicle, the time of the call and the rush to the scene, the age of the injured person, the data on the injury or illness, and the data on the hospital to which the injured person was transported. These huge amounts of accumulated big data are called "emergency big data". By analyzing and utilizing emergency big data, it is possible to improve emergency services. For example, reallocating emergency resources to meet emergency demand, or developing dispatch strategies in anticipation of emergency demand.

In recent years, the number of emergency medical services has been increasing, and emergency medical care is under pressure. In particular, it is difficult to prepare for the constant acceptance of injuries and illnesses that increase rapidly only under certain conditions.

In this study, we focus on heat stroke, which has been increasing in incidence in recent years, as a factor that causes a temporary shortage of emergency medical services. In previous studies, we have shown that heat stroke causes a pressure on emergency medical services. In addition, we have developed a prediction model that calculates the number of people transported due to heat stroke based on past weather data such as temperature, humidity, and weather.

In order to make forecasts that can be used in practice, it is necessary to use future weather information such as weather forecasts. However, traditional forecast models refer to items of past weather data. This raises the issue that they cannot be adapted to current weather forecast data. In addition, it takes at least four to seven days to prepare for emergency measures such as an increase in the number of temporary emergency teams. Therefore, it is necessary to make medium-term (for next week) predictions, the accuracy of which is not clear.

In this paper, we define the following research questions R1 and R2 to solve the problems of the previous model.

- RQ1: Is it possible to predict the number of people who will be transported for heat stroke in the medium term?
- RQ2: Is the method proposed in RQ1 effective in Kobe City?

As a key-idea for RQ1, we will build a prediction model by using only the items included in the weekly weather forecast. The available items are the maximum temperature, the minimum temperature, and the weather. Furthermore, by considering the threshold of temperature at which heat stroke can occur, we reflect regional characteristics in our predictions.

In this study, we construct a multiple regression prediction model using the Partial Least Squares Regression (PLS) algorithm to better explain the prediction results. In addition, we apply a heuristic that multiplies the predictions by 0.85 on rainy days in order to reflect the empirical knowledge provided by the paramedics.

Based on the proposed method, we construct a model using the data of Kobe City as an example. From the Japan Meteorological Agency (JMA), we will obtain and utilize

TABLE I
EXAMPLES OF EMERGENCY BIG DATA

No	Unit	Emergency call at	Arrival on site at	⋯	Arrival location	age	gender	injuries	degree
1	Chuou	2015-01-01 00:01:30	2015-01-01 00:07:30		chuo-ku	70	male	acute intoxication	mild
2	Suma	2016-08-16 14:20:10	2016-08-16 14:30:00		shirakawa. suma-ku	22	female	heat stroke	mild
3	Nada	2017-06-15 19:20:05	2017-06-15 19:27:00		bingo-cho, nada-ku	45	male	cold syndrome	moderate

TABLE II
METEOROLOGICAL DATA IN AUGUST 2020

day	precipitation(mm)	temperature(degree celsius)			humidity(%)		wind speed(m/s)		weather(6 : 00 to 18:00)
		average	max	min	average	min	average	max	
1	—	28.4	31.7	24.6	68	53	2.4	4.9	sunny then shortly cloudy
2	0.0	28.6	32.1	26.3	73	57	3.1	6.5	cloudy and shortly rainy and then sunny, with some thunde

TABLE III
EXAMPLE OF A WEEKLY WEATHER FORECAST BY JMA

Announced at 17:00 on July 23, 2021							
date		23(Fri)	24(Sat)	25(Sun)	⋯	29(Thu)	30(Fri)
weather		sunny and clondy	sunny and clondy	sunny then clondy	⋯	clondy and sunny	clondy and sunny
chance of precipitation(%)		-/-/-/20	0/0/10/10	10	⋯	20	20
Kobe temperature (degree celsius)	max	-	33	33 (31～34)	⋯	32 (29～34)	33 (30～35)
	min	-	26	26 (25～27)	⋯	26 (24～27)	26 (24～27)

historical weather data for each day (854 days in total) from June to September of each year from 2013 to 2019 in Kobe City. In addition, the number of people transported with heat stroke on each day is obtained from the emergency big data. Using these data, we developed a prediction model.

We input weather data for each day from June to September 2020 in Kobe City to the constructed model, and predicted the number of people transported with heat stroke on the same day. Applying the heuristic, the model made predictions with a mean error of 2.45 people.

In RQ2, we feed the weather forecast data into the prediction model constructed in RQ1. Then, we confirm whether it is possible to predict heat stroke in Kobe City in the medium term. The Kobe District Meteorological Observatory provided us with the information stored by the Japan Meteorological Agency.

For each day from June to September 2020, the number of people transported for heat stroke was predicted using the weekly weather forecast announced 7 days in advance. Applying the heuristic, the model made predictions with a mean error of 2.79 people. We also predicted the value range of the number of people using the expected range of temperature. As a result, the probability that the measured value would be included in the value range was 93.4%, and the average size of the value range was 10.03.

II. PRELIMINARY

A. Heat Stroke

Heat stroke is defined as “It’s when heat balance is lost due to temperature rise, causing heat accumulation and affecting organ function” [2]. In most cases, heat stroke occurs when the body’s heat production exceeds heat dissipation and the body temperature rises. Heat stroke can be caused by environmental conditions such as temperature and humidity, physical conditions such as old age and bad health, and behavioral conditions such as strenuous exercise and prolonged outdoor work. For example, in terms of environmental conditions, heat stroke is more likely to occur on days with high temperature and humidity, strong sunshine, weak wind, and sudden heat. In addition, heat stroke risk increases with the duration of hot weather [3].

The main symptoms of heat stroke can be mild, such as dizziness or lightheadedness, or severe, such as convulsions or loss of consciousness. In severe heat stroke, the mortality rate is high, and sequelae are not uncommon. In recent years, as a precaution against COVID-19, “new lifestyle” [4] has been implemented in which people are encouraged to keep a physical distance and wear masks. However, wearing masks can cause heat buildup, and some cases of heat stroke have occurred.

B. Emergency big data

As part of its efforts, the Kobe City Fire Department records emergency dispatch logs using an ICT system. The emergency dispatch log contains the data of the dispatching team and vehicle, the time of the call and the rush to the scene, the age of the injured person, the data on the injury or illness, and the data on the hospital to which the injured person was transported. Table I shows examples emergency big data. Note that the contents of this data are tentative and some data have been omitted.

C. Meteorological data

The Japan Meteorological Agency (JMA) has been recording and accumulating historical *meteorological data* [5] since 1872. Meteorological data refers to temperature, humidity, precipitation, weather, etc. for a specific day or time period. In recent years, not only daily data, but also hourly and ten-minute data have been recorded. Table II shows a part of the daily meteorological data of Kobe City in August 2020. Some of the meteorological data, such as temperature and humidity, include meteorological statistics such as the highest and lowest values for the day. These data are released to the public as *open data*, and are expected to be used as an opportunity for innovation.

In addition, a *weekly weather forecast* [6] is issued that predicts the weather and temperature for the next seven days from the date of the forecast. Table III shows the weekly weather forecast for Hyogo Prefecture, Japan, released at 11:00 a.m. on July 23, 2021.The weather $W(d)$, the probability of precipitation $RP(d)$, the maximum temperature $T_{max}(d)$, and the minimum temperature $T_{min}(d)$ for a particular date d are shown in each column. The temperature row after two days shows after the predicted range of temperatures. In this paper, the lower end of the forecast range of the maximum temperature for date d is $LT_{max}(d)$, the upper end of the forecast range of the maximum temperature is $UT_{max}(d)$, the lower end of the forecast range of the minimum temperature is $LT_{min}(d)$, and the upper end of the forecast range of the minimum temperature is $UT_{min}(d)$.

D. Predicting heat stroke using emergency big data

In our previous research, we constructed a prediction model for the number of people transported for heat stroke by using machine learning based on past weather data and emergency dispatch logs [7]. In this model, we input the maximum temperature (degree celsius), the minimum humidity (%), the weather (sunny, cloudy, rainy), and the maximum temperature (degree celsius) of the day before the forecast day. After that, we can predict the number of people transported for heat stroke in Kobe City in one day. By utilizing this model, it is hoped that more concrete and effective countermeasures against heat stroke can be developed, such as the formulation of mobilization strategies in anticipation of future emergency demand.

However, we found that there are two problems in utilizing the conventional model for emergency strategies.

1) It cannot be adapted to the current weather forecast because it uses items that exist only in historical meteorological data.
2) We need a medium-term forecast of the next 4-7 days for the preparation of emergency strategies.

Regarding (1), referring to Tables II and III, we can find that there are some items that are included in the historical weather data but not in the weekly weather forecast. For example, average humidity, minimum humidity, precipitation, wind speed, and so on. “Minimum humidity” is an item used in conventional forecast model. Therefore, the conventional model cannot make predictions based only on the weekly weather forecast information.

Regarding (2), the Kobe City Fire Department prepares a *temporary EMS* team for days when the risk of heat stroke is high. In order to prepare for this temporary ambulance team, the Fire Department ensures the availability of vehicles and staff, and adjusts the waiting area for the team. This requires a 4-7 day preparation period at least. Therefore, it is necessary to predict the number of people transported for heat stroke in the medium term after one week. However, the accuracy of the medium term prediction is unclear, and we are not sure whether it can be applied to actual operations.

E. Research Questions

In this paper, we define the following research questions R1 and R2 to solve the problems of the previous model.

- RQ1: Is it possible to predict the number of people who will be transported for heat stroke in the medium term?
- RQ2: Is the method proposed in RQ1 effective in Kobe City?

If we can predict the number of people transported for heat stroke in the medium term based on the weekly weather forecast, we can take emergency measures in accordance with the prediction. If we can predict the number of people transported for heat stroke in the medium term based on the weekly weather forecast, we can take emergency measures in accordance with the prediction.

III. PREDICTING THE NUMBER OF PEOPLE TRANSPORTED FOR HEAT STROKE IN THE MEDIUM TERM

A. Key Idea

In order to forecast the number of people transported for heat stroke in the medium term, we use the weekly weather forecast published by the Japan Meteorological Agency. There are two points where this differs from conventional prediction model.

- C1: There are some items that are present in the historical weather data but are not present in the weekly weather forecast.
- C2: Predictions based on weather forecasts do not guarantee the accuracy of the values.

Regarding C1, there is no “minimum humidity” item in the weekly weather forecast, so it cannot be applied to the conventional prediction model. In order to solve C1, we propose the following key idea.

- A1: Construct a prediction model using only items that exist in both historical weather data and weekly weather forecasts.

This reduces the number of available items, but allows us to make predictions using weekly weather forecast.

Regarding C2, since the weekly weather forecast is a predicted value, it may be significantly different from the actual observed value. In addition, the more future weather forecasts are, the less accurate they are. These factors are also likely to affect the prediction. In order to solve C2, we propose the following key idea.

- A2-1: To make the prediction model as interpretable as possible.
- A2-2: To predict not only the predicted value, but also its value range

In A2-1, the prediction process can be interpreted by deriving the predicted values from the multiple regression equation. This allows for a clear explanation of the basis for the predictions and an investigation of the causes when the predictions are wrong.

In A2-2, we utilize the range of predicted temperatures in the weekly weather forecast. This allows us to consider the case where it is hotter than forecast, and the case where the district is cooler than forecast.

B. A1. Feature Engineering

In order to predict the number of heat stroke victims, we need *explanatory variables* that strongly influence the prediction. In this paper, we choose items existing in both the historical weather data and the weekly weather forecast that are strongly attributable to the occurrence of heat stroke. From section II-C, the available items as explanatory variables are "weather (weather summary), maximum temperature, and minimum temperature".

In order to make the temperature feature more prominent, we define a new feature.

$$SqT_{max}(d) = K_\alpha \times \{T_{max}(d) - \alpha\}^2$$
$$SqT_{min}(d) = K_\beta \times \{T_{min}(d) - \beta\}^2$$

$$K_\alpha = \begin{cases} 1 & (T_{max}(d) \geq \alpha) \\ -1 & (T_{max}(d) < \alpha) \end{cases}$$

$$K_\beta = \begin{cases} 1 & (T_{min}(d) \geq \beta) \\ -1 & (T_{min}(d) < \beta) \end{cases}$$

$$\alpha : T_{max}(d) threshold for heat stroke to occur$$
$$\beta : T_{min}(d) threshold for heat stroke to occur$$

The α and β in the above formulas can be freely determined based on the historical records of the regions targeted by the prediction model. For example, we determine α and β for Kobe City.

Figure 1 shows the relationship between min/max temperature and the average number of people transported for heat stroke per day in Kobe City. It can be seen that in Kobe City, heat stroke gradually occurs when the minimum temperature exceeds 22 degree and the maximum temperature exceeds 25 degree Therefore, we define $\alpha = 25, \beta = 22$. It can also be seen that the number of people transported for heat stroke increases rapidly when the maximum temperature exceeds α or the minimum temperature exceeds β. In order to strongly reflect this impact, we adopt the squared equation for $SqT_{max}(d), SqT_{min}(d)$. On the other hand, when the maximum temperature is below α or the minimum temperature is below β, it is considered that heat stroke rarely occurs. In order to reflect this impact, we adopt the condition that K_α, K_β can be negative.

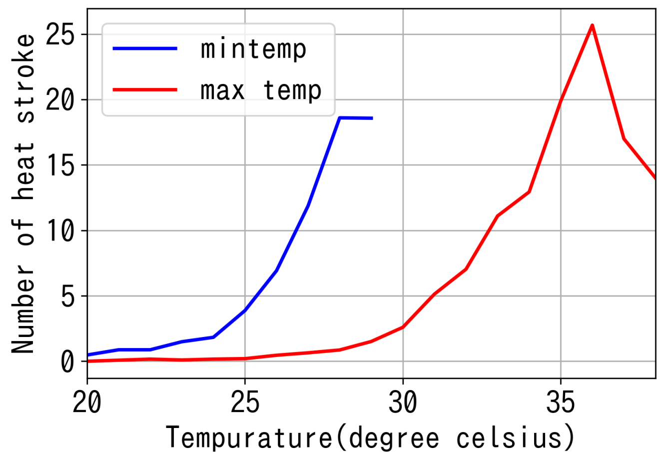

Fig. 1. Average number of persons transported for heat stroke per day at specific temperatures

Next, we will consider using the weather features as variables. For the weather summary, as shown in Table II, there are more than 20 types of complex expressions, such as " sunny and shortly cloudy" and "cloudy and shortly rainy and then sunny, with some thunder". These expressions are polite. However, they may negatively affect the forecast accuracy and are difficult to collate with the weekly weather foHowever, they may negatively affect the prediction accuracy and are difficult to collate with the weekly weather forecast. Therefore, it is necessary to convert the weather into a simple expression. Because the intensity of sunlight affects heat stroke, we focus only on "whether it is sunny or not". Then, we define the features as follows.

$$IS_{am}(d) = \begin{cases} 1 & (The weather in the AM is sunny) \\ 0 & (The weather in the AM is not sunny) \end{cases}$$

$$IS_{pm}(d) = \begin{cases} 1 & (The weather in the PM is sunny) \\ 0 & (The weather in the PM is not sunny) \end{cases}$$

We will explain how to convert a weather summary into $IS_{am}(d), IS_{pm}(d)$. The weather in the AM is the first word in the weather summary. Next, we check if the word "then" exists in the weather summary. The word "then" means that the weather changes with time. If included, the weather in the PM is the next word after "then" in the weather summary. If not included, the weather in the PM is the same as the weather in the AM.

As an example, we convert "cloudy and shortly rainy and then sunny, with some thunder", the weather of August 1, 2020, into $IS_{am}(d), IS_{pm}(d)$. The first word is "cloudy", so theweather in the AM is "cloudy". Next, since the word "then" is included, the weather in the PM is "sunny" from the next

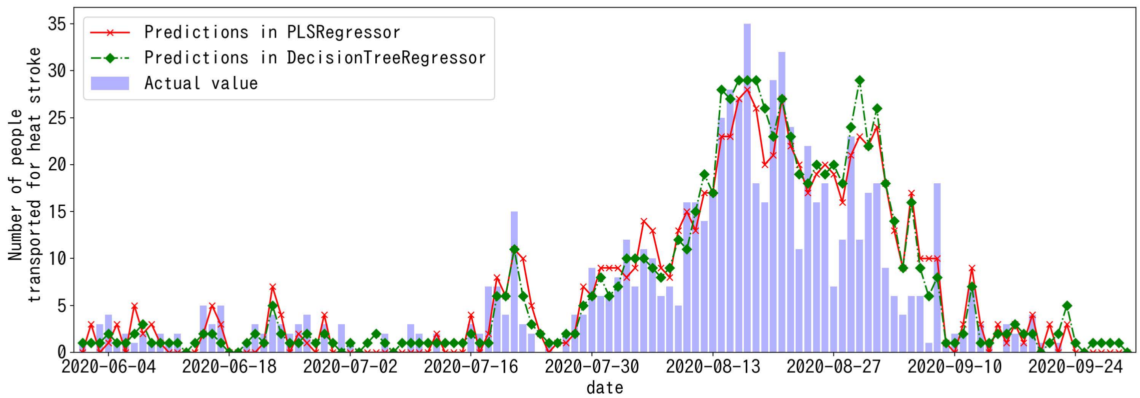

Fig. 2. Comparison of predictions and actual numbers of people transported for heat stroke

word. Applying these conditions to the variables, we can get $IS_{am}(2020/08/01) = 0, IS_{pm}(2020/08/01) = 1$.

From the above features, the following six variables are selected as explanatory variables. These were examined based on the factors that caused the occurrence of heat stroke in section II-A. The date of the predicted date of the explanatory variables is d, and the date n days before the predicted date is $d - n$.

- $SqT_{max}(d)$
- $SqT_{min}(d)$
- $SqT_{max}(d-1)$
- $SqT_{max}(d-2)$
- $IS_{am}(d)$
- $IS_{pm}(d)$

C. A2-1. Construction of prediction models using multiple regression models

In order to construct a prediction model, we require not only explanatory variables but also *target variables*, which is the value we want to predict. The aim of this paper is to predict the number of people transported for heat stroke. In this chapter, we predict the number of people transported by heat stroke in Kobe City on a day, using data from Kobe City as an example.

To build explanatory variables, we obtain the meteorological data for each day from June to September of each year from 2013 to 2019 (854 days in total) in the JMA website as a csv file. Then, we transform the data into the variables described in section III-B by Python. We also normalize the obtained variables and adjust their values to follow a standard normal distribution. As the objective variable, we obtain the number of people transported with heat stroke on the same day from the emergency big data of Kobe City.

In order to create the prediction model, we utilize *scikit-learn*, an open source machine learning library in Python. As a machine learning algorithm, we adopt Partial Least Squares Regression (PLS). PLS is a type of multiple regression analysis. It is characterized by its capability to derive regression equations even when the correlation between the explanatory variables is high. For example, weather and maximum temperature are highly correlated and the regression coefficients are often unstable. By using PLS, this problem can be solved.

From the above, we input the explanatory and objective variables for all 854 days obtained, and constructed a multiple regression prediction model by PLS. Also, we construct a regression tree model proposed as a conventional model using the same variables. For each model, we calculate the mean absolute error (MAE) and root mean square error (RMSE), and rounded them to the third decimal place. We then compare the accuracy of the models based on these values.

D. Evaluation of prediction accuracy

For the constructed two models, we compare the predictions with the actual measured values obtained from the emergency big data. In the evaluation, we use data that does not overlap with the data for model building. For this purpose, we will use meteorological data and emergency big data for each day from June to September 2020 (122 days in total). By creating explanatory variables and applying them to each of the models, we obtained the predictions for the number of people transported for heat stroke. The predictions output by the multiple linear regression models are sometimes to the right of the decimal point and sometimes to a negative number. However, the actual number of people can only be expressed as an integer greater than or equal to zero. Therefore, we perform two processes on the output predictions: “rounding off the decimal point” and “setting the prediction value to 0 if a negative value is output”.

Figure 2 shows the predictions and actual numbers of people transported for heat stroke from June to September 2020. The horizontal axis represents each day in 2020, and the vertical axis represents the number of people transported for heat stroke. Line graphs show the predictions from each model, and bar graphs show the actual values. From this graph, we consider that there is no significant difference in the values output by each model.

TABLE IV
ACCURACY EVALUATION INDICATORS

model	MAE	RMSE
PLS	2.58	3.65
RegressionTree	2.50	3.92
Conventional model	2.18	3.44

Table IV shows the indicators for evaluating the error of each prediction model, and as a comparison, the metrics for previous studies.

Comparing PLS and RegressionTree, PLS has a slightly larger MAE but a smaller RMSE. Therefore, it shows that the large error between prediction and actual values is small. However, comparing PLS with the conventional model, PLS could not outperform the conventional model in terms of accuracy.

E. Heuristics for improving accuracy

The PLS model constructed in section III-C only considers whether the weather is "sunny" or not. Therefore, it excludes the factor "rain", which is likely to reduce the risk of heat stroke. Due to this, we consider that the accuracy of the PLS model was not good.

In order to improve the accuracy of the PLS model, we apply "heuristics". A heuristic is to derive an answer from empirical rules or intuition. In this paper, we apply a heuristic that "multiplies the output prediction by 0.85 on rainy days", based on interviews with fire department staff and trial-and-error parameter tuning. Whether it has rained or not is determined by whether the word "rain"is included in the weather summary. We think that this will reflect the effect of the decrease in the risk of heat stroke due to sudden rainfall.

TABLE V
ACCURACY EVALUATION INDICATORS WITH HEURISTICS

model	MAE	RMSE
PLS	2.45	3.41
RegressionTree	2.43	3.80
Conventional model	2.18	3.44

Table V shows the indices when the heuristic is applied. Note that the heuristic is not adapted to the conventional model because it uses the information of rain. The results show that the accuracy of the proposed model has been improved. It can also be seen that the PLS has a smaller RMSE than the conventional model.

F. A2-2. Reasoning with Models

In order to represent the range of predicted values for the number of people transported for heat stroke, we need explanatory variables for cases where the predicted value is greater or less than the original value. We cannot change the type of explanatory variables for the model, so we will consider changing the values of the variables.

To represent the upper part of the value range of the prediction, we assume that the temperature has changed to the maximum of the forecast range and that the weather has changed to "sunny", and change the variables. Conversely, to represent the lower part of the value range, assume that the temperature has changed to the minimum of the forecast range and change the variable.

IV. EVALUATING THE PROPOSE METHOD IN KOBE CITY

In this chapter, we check whether it is possible to predict based on the actual weather forecast by using the constructed model. In addition, we evaluate how much the error of the weather forecast affects the prediction accuracy.

As weather forecast data, we use the weekly weather forecast released at 5:00 p.m. each day from May 1 to September 30, 2020. The Kobe District Meteorological Observatory provided us with the information stored by the Japan Meteorological Agency.

First, we evaluate the reliability of the weekly weather forecast by comparing it with actual meteorological data. In this way, we evaluate the input accuracy of the medium-term prediction. Second, we input the weekly weather forecast into the prediction model and get predictions of the number of people transported for heat stroke. Finally, we evaluate the error between the predictions and the actual number of people using MAE and RMSE.

A. weekly weather forecast

TABLE VI
TEMPERATURE ERROR IN WEEKLY WEATHER FORECAST

Forecast for the next X days	$T_{max}(d)$		$T_{min}(d)$	
	MAE	RMSE	MAE	RMSE
2 days	1.20	2.17	0.80	1.03
3 days	1.32	2.69	0.90	1.24
4 days	1.35	2.83	0.98	1.46
5 days	1.43	3.04	1.02	1.69
6 days	1.42	3.08	1.02	1.74
7 days	1.48	3.24	1.10	1.88

TABLE VII
TEMPERATURE FORECAST RANGE ERROR IN WEEKLY WEATHER FORECAST

Forecast for the next X days	Tmax		Tmin	
	Prediction accuracy(%)	Average range	Prediction accuracy(%)	Average range
2 days	77.0	3.30	77.0	2.49
3 days	81.1	3.49	77.0	2.70
4 days	74.6	3.58	77.0	2.81
5 days	79.5	3.94	79.5	3.01
6 days	77.0	3.93	77.0	3.26
7 days	76.2	4.25	82.8	3.57

The weekly weather forecast includes the weather, maximum temperature, minimum temperature, and range of temperature for two to seven days after forecast date. From those data, we extract only the forecast data for 2-7 days after the forecast date and from June 1 to September 30, 2020. Then, compare with the actual weather data.

Table VI shows the indices to evaluate the temperature forecast error for the next two to seven days. We can see

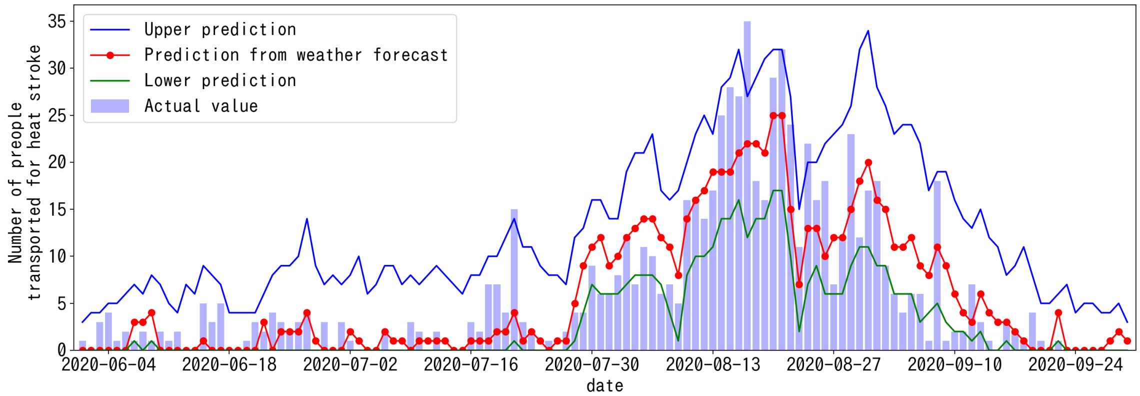

Fig. 3. Prediction of the number of heat stroke and its range in 2020 using weekly weather forecast

that the farther away from the forecast date, the more error is likely to occur in the temperature forecast. Also, we can see that even if the forecast is close to the forecast date, it is not always true.

In addition, Table VII represents an indicator that evaluates the temperature forecast range for the next two to seven days. Prediction accuracy represents the percentage of days where the actual temperature values are included in the predicted range of the weather forecast. From this table, we can see that the accuracy stays at about 78%. On the other hand, the farther days from the forecast date, the wider the forecast range is. This means that the more distant the day is from the forecast date, the less accurate the temperature prediction is.

B. Medium-term predictions using weather forecasts

In this section, we predict heat stroke after 4 following days in order to evaluate the prediction in the medium term. Next, we evaluate the accuracy of the predictions by comparing them with the actual number of heat stroke victims transported. The weekly weather forecast data obtained in this paper exists for up to seven days after the forecast publication date. Therefore, we will make predictions from 4 to 7 days after the forecast date.

TABLE VIII
EVALUATION OF PREDICTION ACCURACY USING WEEKLY WEATHER FORECAST

Forecast for the next X days	MAE	RMSE
4 days	2.89	4.23
5 days	2.93	4.30
6 days	2.98	4.29
7 days	2.79	3.87

The data for evaluation is the weather forecast data for 122 days from June 1, 2020 to September 30, 2020 in Kobe City. Based on the data for each day, we predict the value and compare it with the number of people transported by heat stroke in Kobe City on the same day.

Table VIII shows the metrics for evaluating the error in predictions after 4-7 days. According to the predictions from 4 to 6 days later, we can see that the more distant the day, the less accurate the prediction. On the other hand, the accuracy is higher for 7 days later than for other days.

C. Prediction of value ranges using temperature forecast ranges.

In this section, we predict the value range of the predictions proposed in section III-F from the actual weekly weather forecast. Then, we evaluate the accuracy of the prediction by whether the actual number of heat stroke is included in the value range of the prediction. The time period and data are the same as in section IV-B. We will analyze two items to evaluate the accuracy. The first is the percentage of days in which the actual number of heat stroke is within the range of the predicted value. The second is the mean size of the value range of the prediction. The size is the difference between the upper and lower ends of the value range of the prediction.

TABLE IX
PREDICTION ACCURACY AND MEAN RANGE USING TEMPERATURE FORECAST RANGES.

Forecast for the next X days	Prediction accuracy(%)	Average range
4 days	83.6	8.88
5 days	82.0	8.96
6 days	83.6	9.22
7 days	93.4	10.03

Table IX shows an index of the accuracy of the range of values for the prediction after 4-7 days. It can be seen that the accuracy of the prediction after 7 days is higher than the prediction after 4 to 6 days. However, the size of the range of predicted values becomes larger the further away the data is, and we cannot say that the accuracy after 7 days is good.

Figure 3 shows the range of predictions and actual values of heat stroke cases using the weather forecast for seven days later. It can be seen that some days the actual value are above the prediction and some days are below the prediction.

V. CONSIDERATION

In RQ1, we constructed a new heat stroke prediction model using only the items that exist in the weekly weather forecast. In the process, we developed features that can reflect regional characteristics. This makes it possible to construct a prediction model for regions other than Kobe City. Also, by predicting the range of prediction, we can expect countermeasures to be taken in cases where cases of heat stroke are more likely to occur. In addition, the model constructed by PLS facilitates the interpretation of the prediction process. On the other hand, the model constructed in this paper has a larger average error than the conventional model. We consider that this is because the minimum humidity, a explanatory variable used in the conventional model, has a strong influence on heat stroke. In order to improve the accuracy of the model, it is necessary to examine alternative explanatory variables to replace the minimum humidity.

In RQ2 we first analyzed the accuracy of weather forecasts. It was found that the accuracy of temperature forecasts was worse on more distant days, and that even the most recent forecasts were not always correct.

Next, we predicted the number of people transported for heat stroke based on the weekly weather forecast. It was found that predictions based on weather forecasts were less accurate than those based on actual weather measurements. It was also found that the more distant the date, the lower the accuracy. On the other hand, the predictions for the next 7 days were more accurate than the predictions for other days. Looking into the details, it was found that the accuracy of predictions for the next 7 days was better than the predictions for other days on August 9 and August 27.

As for August 9, we found a large error in the prediction using past meteorological data. This day was the second day of a three-day weekend, so it is highly likely that people went out of Kobe City for vacation or stayed home due to COVID-19. Therefore, this may have influenced the decrease in the number of heat stroke cases in Kobe City. The weather forecast for the next seven days was “cloudy”, although it was actually sunny. Thus, it is considered that the predicted values were closer to the actual values.

These results indicate that forecast errors can have both positive and negative effects on the prediction.

Finally, we predicted the value ranges using temperature forecast ranges. The results showed that the range of predictions became larger for the mid-term predictions. The results also showed that the accuracy of the predictions did not change much. Figure 3 shows that the measured values on July 21 were above the predicted range, while the measured values on September 9 were below the predicted range.

For July 21, the actual temperature was within the forecast range of the weekly weather forecast. However, it was the first day in 2020 that the temperature exceeded 30 degree, and the temperature rose significantly compared to the previous day. Therefore, it is considered that more heat stroke than predicted occurred. As for September 9, different from the forecast, it rained. This resulted in the daytime temperature not rising much. As a result, the temperature was lower than forecasted, and heat stroke did not occur to a small number.

In summary, it is possible to predict the number of people transported for heat stroke in the medium term using weekly weather forecasts. On the other hand, the influence of forecast error is strong, and the prediction range becomes large.

In future research, we will ask people to use this prediction in the emergency medical services. In the process, we will examine whether the predictions can be used as emergency measures.

VI. CONCLUSION

In this paper, we constructed a prediction model to predict the number of people transported for heat stroke in the medium term from 4 to 7 days. In contrast to conventional studies, we used only the items included in the weekly weather forecast as explanatory variables. These explanatory variables were processed into features that can reflect regional characteristics. In addition, the prediction process was made interpretable.

After the construction, we predicted the number of heat stroke in the medium term using the actual weekly weather forecast of Kobe City. Based on the forecast for the next 7 days, we predicted the number of heat stroke with a mean error of 2.80.

In the future, we will explore new explanatory variables that are effective for medium-term predictions. And we will ask people to use this prediction in the emergency medical services. In the process, we will examine whether the predictions can be used as emergency measures.

VII. ACKNOWLEDGEMENTS

This research was partially supported by JSPS KAKENHI Grant Numbers JP19H01138, JP18H03242, JP18H03342, JP19H04154, JP19K02973, JP20K11059, JP20H04014, JP20H05706 and Tateishi Science and Technology Foundation (C) (No.2207004).

This research is conducted in collaboration with the Kobe City Fire Department.

REFERENCES

[1] M. Deakin and H. Al Waer, “From intelligent to smart cities,” *Intelligent Buildings International*, vol. 3, no. 3, pp. 140–152, 2011.

[2] A. Bouchama and J. P. Knochel, “Heat stroke,” *New England journal of medicine*, vol. 346, no. 25, pp. 1978–1988, 2002.

[3] S. Kodera, T. Nishimura, E. A. Rashed, K. Hasegawa, I. Takeuchi, R. Egawa, and A. Hirata, “Estimation of heat-related morbidity from weather data: A computational study in three prefectures of japan over 2013–2018,” *Environment international*, vol. 130, pp. 1–9, 2019.

[4] L. Ministry of Health and Welfare, “Example of practicing new lifestyle,” https://www.mhlw.go.jp/stf/covid-19/kenkou-iryousoudan_00006.html, (Accessed on 09/15/2021).

[5] “Japan meteorological agency,” http://www.data.jma.go.jp/obd/stats/etrn/index.php, (Accessed on 09/15/2021, Japanese text only).

[6] “Jma one-week forecasts,” https://www.jma.go.jp/bosai/map.html#contents=forecast&lang=en, (Accessed on 09/15/2021).

[7] T. Nakai, S. Saiki, and M. Nakamura, “Analyzing and predicting ambulance demand for heat stroke using fine-grained dispatch log,” *IEICE Technical Report*, vol. 120, no. 232, pp. 13–19, 2020.