

## 気象予報を活用した熱中症救急需要の中期予測手法の検討

中井 哲也<sup>†</sup> 中村 匡秀<sup>†,‡</sup> 佐伯 幸郎<sup>†,‡,‡‡</sup>

<sup>†</sup> 神戸大学 〒657-8501 神戸市灘区六甲台町 1-1

<sup>‡</sup> 理化学研究所・革新知能統合研究センター 〒103-0027 東京都中央区日本橋 1-4-1

<sup>‡‡</sup> 高知工科大学 〒782-8502 高知県香美市土佐山田町宮ノ口 185

E-mail: <sup>†</sup>manda@ws.cs.kobe-u.ac.jp, <sup>‡</sup>masa-n@cs.kobe-u.ac.jp, <sup>‡‡</sup>saiki.sachio@kochi-tech.ac.jp

あらまし 我々はこれまでの研究で、消防局の救急出動記録と気象観測データを機械学習にかけ、神戸市の熱中症搬送者数の予測モデルを構築し、直近未来の熱中症救急需要予測が可能であることを示した。しかしながら、現場において必要な救急施策を準備するためには、向こう1週間程度の中期的な予測を行う必要があり、その予測精度は分かっていない。そこで本稿では、4-7日程度の中期的な熱中症搬送者数を予測する新たな手法を検討する。具体的には、週間天気予報に含まれる項目のみを説明変数とする予測モデルの構築を行う。また、予測したい地域における、熱中症が発生しうる気温の閾値を加え、その地域の特徴として反映させる。予測結果を現場に説明しやすくするため、部分的最小二乗回帰 (Partial Least Squares Regression, PLS) でモデル構築を行う。さらに雨が降った日の予測は0.85倍するヒューリスティックを適用する。提案手法で構築したモデルに対し、過去の週間天気予報のデータを投入して評価を行った。神戸市における2020年6月から9月までの各日において、7日前に発表された週間天気予報を用いて、熱中症搬送者数を予測したところ、平均誤差2.79人の予測が可能であることが分かった。

キーワード 熱中症, 救急医療, スマートシティ, 需要予測, 機械学習

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

Tetsuya NAKAI<sup>†</sup>, Masahide NAKAMURA<sup>†,‡</sup>, and Sachio SAIKI<sup>†,‡,‡‡</sup>

<sup>†</sup> Kobe University Rokkodai-cho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan

<sup>‡</sup> Riken AIP, 1-4-1 Nihon-bashi, Chuo-ku, Tokyo 103-0027

<sup>‡‡</sup> Kochi University of Technology 185 Miyanokuchi, Tosayamada, Kami City, Kochi, 782-8502, Japan

E-mail: <sup>†</sup>manda@ws.cs.kobe-u.ac.jp, <sup>‡</sup>masa-n@cs.kobe-u.ac.jp, <sup>‡‡</sup>saiki.sachio@kochi-tech.ac.jp

**Abstract** In our previous study, we have developed a prediction model that forecasts the number of heatstroke cases within the immediate future, by applying machine learning to observational weather data and ambulance dispatch log of Kobe Fire department. However, the practical fields require mid-term forecasting to prepare for necessary measures, which is not assumed in the previous model. In this paper, we propose a new method that forecasts the number of heatstroke cases for mid-term future from 4 to 7 days. Specifically, we first construct a model with only attributes in 7-day weather forecast. We then find a temperature threshold that triggers many heatstroke cases in the target area, and use the threshold for feature engineering. To construct an interpretable model, we use the partial least squares regression (PLS) method. We also apply a heuristic that multiplies 0.85 to the predicted demand in a rainy day. We have evaluated the proposed method using data archive of 7-day weather forecast in Kobe. For each day from June 1st to September 30th of 2020, we predicted the number of heatstroke cases using weather forecast published 7 day before. As a result, the proposed model well predicts the number within the mean absolute error of 2.79.

**Key words** heat stroke, ambulance, smartcity, demand prediction, machine learning

## 1. はじめに

ICTシステムやIoT技術の急速な発展により、企業のみならず政府や自治体が多種多様なデータを取得・蓄積し、施策や街づくりに役立てる動きが活発になっている。こうした様々なデータを利活用することで、より効率的で持続的な都市、スマートシティ [1] を目指すことは、今や世界的な潮流となっている。

神戸市消防局では、この様な取り組みの一環として、救急出動に関するデータをICTシステムにより収集・記録している。救急出動データには、各救急出動に対して、出動隊や車両のデータ、入電や現場への駆けつけなどの時刻データ、傷病者の年齢や傷病に関するデータ、搬送先の病院のデータ等が記録されている。これらの膨大なデータが蓄積されたビッグデータ（以降、救急ビッグデータと呼ぶ）を分析・活用することで、これまで定量的な評価が困難であった、救急需要を満たす救急リソースの再配置、救急需要を見越した出動戦略の策定、などの救急事業の改善が実現可能である。

近年、救急搬送数が増加傾向にあり、救急医療のひっ迫が指摘されている。特に、慢性的なひっ迫ではなく、特定の条件下においてのみ急激に増加するような傷病への対応は、常に受け入れ態勢を整えることは難しく、救急施策として考慮することは困難である。

本研究では、このような一時的な救急医療をひっ迫する要因として、近年その発生増加が大きな問題となっている熱中症に着目する。これまでの研究において、熱中症が救急医療のひっ迫に影響を与えることを明らかにした。加えて、気温・湿度・天気などの過去の気象データから、任意の日の熱中症搬送者数を求める予測モデルを構築した。この予測モデルを活用し、未来の熱中症搬送者数を具体的な数値を提示することで、より説得力のある救急施策に活用することが期待される。

実際に活用可能な予測を行うためには、天気予報などの未来の気象情報を活用し、予測を行う必要がある。しかし、従来の予測モデルでは、過去の気象データの項目を参考にしていないため、現在の天気予報データに適応できないという課題が挙げられる。加えて、臨時隊の増隊とあって救急施策を準備するためには、少なくとも4~7日の準備期間を要する。そのため、向こう1週間程度の中期的な予測を行う必要があり、その予測精度は分かっていない。

そこで本稿では、従来モデルの課題を解決するため、以下のR1, R2のリサーチ・クエスチョンを設定する。

- RQ1: 中期の熱中症搬送者数の予測は可能か？
- RQ2: 神戸市において、RQ1での提案手法は有効か？

RQ1へのキーアイデアとして、週間天気予報に含まれる項目のみを活用し、予測モデルの構築を行う。利用可能な項目として、予測日の最高気温、最低気温、天気がある。これに加え、予測したい地域における、熱中症が発生しうる気温の閾値を加えることで、その地域性を反映させる。

本研究では、予測結果を説明しやすくするため、部分的最小二乗回帰 (Partial Least Squares Regression, PLS) という手

法を用いて、重回帰予測モデルを構築する。加えて、予測値に対して、雨が降った日の予測は0.85倍するヒューリスティックを適用し、救急隊員による経験則を反映させる。

提案手法を元に、神戸市のデータを例としてモデルを構築する。利用するデータとして、気象庁より、神戸市における、2013年から2019年の各年6月から9月の各日(全854日)の過去の気象データを取得する。加えて、神戸市消防局が持つ同期間の救急ビッグデータより、各日の熱中症搬送者数を取得する。以上のデータを用いて、予測モデルを構築した。

構築したモデルに対して、神戸市における2020年6月から9月までの各日の気象データを入力し、同日の熱中症に対する救急出動回数を予測しヒューリスティックを適用したところ、平均誤差2.45人の予測が可能であることが分かった。

RQ2では、RQ1で構築したモデルに、実際の天気予報データを投入することで、神戸市における中期の熱中症予測が可能か確認する。本稿では、天気予報データとして、2020年5月1日から9月30日各17時に発表された週間天気予報を活用する。このデータは、神戸地方気象台より、気象庁保有の情報を提供して頂いた。

2020年6月から9月までの各日において、7日前に発表された週間天気予報を用いて、熱中症搬送者数の予測しヒューリスティックの適用したところ、平均誤差2.79人の予測が可能であることが分かった。また、週間天気予報に含まれる気温の予想範囲を用いることで、熱中症予測値域を出力した。その結果、実測値が予測の値域に含まれる確率が93.4%、値域の平均の大きさが10.03であった。

## 2. 準備

### 2.1 熱中症

近年、救急搬送数が増加傾向にあり、救急医療のひっ迫が指摘されている。特に慢性的なひっ迫ではなく、特定の条件下においてのみ急激に増加するような傷病への対応は、常に受け入れ態勢を整えることは難しく、救急施策として考慮することは困難である。このような一時的な救急医療をひっ迫する要因として、近年その発生増加が問題となっている熱中症に着目する。

熱中症は「体温が上昇して重要な臓器が高温にさらされたりすることにより発症する障害の総称」と定義されている [2]。多くの場合、体内の熱産生が熱放散を上回ることによって、体の体温が上昇し熱中症が発症する。熱中症を引き起こす要因として、気温や湿度といった環境条件、高齢や体調不良といった身体条件、激しい運動や長時間の屋外作業といった行動条件がある。例えば環境条件に着目すると、高温多湿の日、日差しが強い日、風が弱い日、急に暑くなった日は、熱中症が発生しやすいことが分かっている。加えて、過去の気象が熱中症発生に影響しており、暑い日が続くことで熱中症発生リスクが向上することが分かっている [3]。

熱中症の主な症状として、めまいや立ちくらみ等の軽症から、けいれんや意識障害等の重症まで多岐にわたる。重度の熱中症（熱射病と呼ばれる）では、死亡率が高く、後遺症が残ることも少なくない。近年では、COVID-19(新型コロナウイルス)対策

表 1 救急ビッグデータの例

No	駆けつけ部隊	通報日時	現場到着日時	...	傷病者発生場所	年齢	性別	傷病名	傷病程度
1	中央救助隊	2015-01-01 00:01:30	2015-01-01 00:07:30		神戸市中央区	70	男	急性アルコール中毒	軽傷
2	須磨救急隊	2016-08-16 14:20:10	2016-08-16 14:30:00		須磨区白川	22	女	熱射病・熱中症	軽傷
3	灘救急隊	2017-06-15 19:20:05	2017-06-15 19:27:00		灘区備後町	45	男	かぜ症候群	中等症

表 2 2020 年 8 月における、過去の気象データの例

日	降水量 (mm)	気温 (°C)			湿度 (%)		風速 (m/s)		天気概要	
		平均	最高	最低	平均	最小	平均	最大	昼 (6~18 時)	夜 (18 時~6 時)
1	—	28.4	31.7	24.6	68	53	2.4	4.9	晴後一時曇	曇時々晴
2	0.0	28.6	32.1	26.3	73	57	3.1	6.5	曇一時雨後晴, 雷を伴う	晴時々曇

表 3 気象庁による、週間天気予報の例

2021 年 07 月 23 日 17 時 神戸地方気象台 発表									
日付	23 日 (金)	24 日 (土)	25 日 (日)	26 日 (月)	27 日 (火)	28 日 (水)	29 日 (木)	30 日 (金)	
兵庫県天気	晴時々曇	晴後時々曇	晴時々曇	晴時々曇	晴時々曇	晴時々曇	曇時々晴	曇時々晴	
降水確率 (%)	-/-/-/20	0/0/10/10	10	10	20	10	20	20	
神戸 気温 (°C)	最高	-	33 (31~34)	33 (31~35)	32 (31~34)	32 (30~34)	32 (29~34)	33 (30~35)	
	最低	-	26 (25~27)	26 (24~27)	26 (24~27)	26 (23~27)	26 (24~27)	26 (24~27)	

として、身体的距離を保ち、マスクの着用が推奨される「新しい生活様式」が実施されている。しかし、マスクを着用することで熱がこもり、体温調節が難しくなることで、熱中症を発症する事案が発生している。

## 2.2 消防局が持つ救急ビッグデータ

我々が共同研究を行っている神戸市消防局では、スマートシティを目指す取り組みの一環として、救急出動に関するデータを ICT システムにより収集・記録している。

救急出動に関するデータには、各救急出動に対して、出動隊や車両のデータ、入電や現場への駆けつけなどの時刻データ、傷病者の年齢や傷病に関するデータ、搬送先の病院のデータ等が記録されている。表 1 は、救急ビッグデータの例を表している。なお、提示しているデータの内容は仮のものであり、一部のデータを削除及び省略していることに注意されたい。これらの膨大なデータが蓄積されたビッグデータ（以降、救急ビッグデータと呼ぶ）を分析・活用することで、これまで定量的な評価が困難であった、救急需要を満たす救急リソースの再配置、救急需要を見越した出動戦略の策定、などの救急事業の改善が実現可能である。

## 2.3 気象データ

気象庁では、1872 年から現在に至るまでの過去の気象データ [4] を記録、蓄積している。気象データとは、特定の日や時間帯における、気温、湿度、降水量、天気などを指す。近年では、1 日の気象データだけでなく、1 時間ごと、10 分ごとの気象データも記録、蓄積している。表 2 は、2020 年 8 月における日ごとの神戸市の気象データの一部を表している。気温や湿度などの気象データの一部には、その日における最高値や最低値という気象統計値が表される。このようなデータはオープンデータとして一般に公開され、イノベーションの契機に活用することが

期待されている。

また、発表日翌日から 7 日先までの天気、気温などを予報する週間天気予報 [5] を発表している。表 3 は、2021 年 07 月 23 日 11 時に発表された、兵庫県の週間天気予報を表している。特定の日付  $d$  における天気  $W(d)$ 、降水確率  $RP(d)$ 、最高気温  $T_{max}(d)$ 、最低気温  $T_{min}(d)$  が、列ごとに表される。気温の行に注目すると、予報日の 2 日後以降の気温には、気温の予測範囲が表示されている。本稿では、日付  $d$  の最高気温の予測範囲の下端を  $LT_{max}(d)$ 、最高気温の予測範囲の上端を  $UT_{max}(d)$ 、最低気温の予測範囲を下端  $LT_{min}(d)$ 、最低気温の予測範囲の上端を  $UT_{min}(d)$  とする。

## 2.4 救急ビッグデータを活用した熱中症予測

これまでの研究において、過去の気象データと救急出動記録を機械学習にかけることで、熱中症搬送者数の予測モデルを構築した [6]。これは、予測日の最高気温 (°C)、最低湿度 (%), 天気 (晴れ・曇り・雨) 及び予測日前日の最高気温 (°C) を入力することで、神戸市における 1 日の熱中症搬送者数 (人) の予測値を求めるものである。熱中症予測モデルを活用することで、将来の救急需要を見越した出動戦略の策定といった、より具体的に効果的な熱中症対策を立てることが期待される。

しかし一方で、従来モデルを救急戦略に組み込むためには、次に挙げる 2 点の問題があることが分かった。

- (1) 過去の気象データに存在する項目を利用しているため、現在の天気予報に適応出来ない
- (2) 救急戦略の準備のため、4~7 日後の中期的な予測が必要である

(1) に関して、表 2 と表 3 を参照すると、過去の気象データには存在するが、週間天気予報には存在しない項目があることが分かる。例えば、平均湿度、最小湿度、降水量、風速などで

ある。この中の「最小湿度」は、従来の予測モデルに使われる要素であることから、週間天気予報の情報のみでは、従来のモデルをそのまま活用することができない。

(2)に関して、現在、熱中症の発生リスクの高い日に対して臨時救急隊を準備することが出動戦略として挙げられている。この臨時救急隊の準備には、車両や人員の確保、臨時救急隊待機場所の調整などがあり、少なくとも4~7日の準備期間を要する。そのため、向こう1週間程度の中期的な熱中症搬送者数の予測が必要となる。しかしながら、中期的な予測の精度は不明であり、実際の業務に適応可能かどうかは分からない。

## 2.5 リサーチ・クエスチョン

本稿では、従来モデルの課題を解決するため、以下のR1, R2のリサーチ・クエスチョンを設定する。

- RQ1: 中期の熱中症搬送者数の予測は可能か？
- RQ2: 神戸市において、RQ1での提案手法は有効か？

週間天気予報から中期的な熱中症搬送者数の予測を行うことが出来れば、予測に合わせた救急施策に繋げることが出来る。加えて、具体的な数値を提示することで、施策に対する妥当性や効果を説明することが出来る。

## 3. RQ1: 中期の熱中症搬送者数の予測は可能か？

### 3.1 キーアイデア

中期の熱中症搬送者数の予測を行うために、気象庁が発表する週間天気予報を活用する。これにより、従来の予測モデルと異なる点が2点挙げられる。

- C1: 過去の気象データに存在するが、週間天気予報に存在しないデータ項目がある
- C2: 気象データの予報を元に熱中症搬送者数の予測を行うため、値の正確性が保障されない

C1について、週間天気予報に「最低湿度」の項目がないため、そのままでは従来の予測モデルに適用できない。C1を乗り越える為に、次のキーアイデアを提案する。

- A1: 過去の気象データ、週間天気予報の両方に存在する項目のみで、予測モデルを構築する

これにより、利用可能な項目は減るが、天気予報を活用した中期の予測に適用出来るようになる。

C2について、週間天気予報は予測値であるため、実際の観測値と大きく異なる可能性がある。また、中期の天気予報になるほど、その精度が落ちると考えられ、熱中症搬送者数の予測にも影響する可能性が高い。C2を乗り越える為に、次のキーアイデアを提案する。

- A2-1: 構築する予測モデルを、可能な限り解釈可能にする
- A2-2: 予測値に加え、その値域も予測モデルより出力する

A2-1では、予測値を重回帰式から導き、予測の仮定を解釈可

能にする。これにより、予測値の根拠となる説明や、予測が外れた場合の原因調査を明確に行うことが出来るようになる。

A2-2では、週間天気予報の気温の予測範囲を活用する。気温の予測範囲を活用することによって、予報よりも暑くなり熱中症リスクが上がる場合、逆に予報よりも涼しくなり熱中症リスクが下がる場合を検討することが出来る。

### 3.2 A1. 特徴開発

熱中症搬送者数を予測するために、その数に強く影響を与える要因、すなわち説明変数が必要である。本稿では、過去の気象データ及び週間天気予報の両方に存在する項目の内、熱中症の発生に強く起因する要素を説明変数とする。2.3節より、説明変数として利用可能な項目は、「天気(天気概要)・最高気温・最低気温」である。

ここで、気温の特徴をより際立たせるため、新たな特徴量を定義する。

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

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

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

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

$\alpha$ : 熱中症が発生する  $T_{max}(d)$  の閾値

$\beta$ : 熱中症が発生する  $T_{min}(d)$  の閾値

上式における  $\alpha$  及び  $\beta$  は、予測モデルの対象となる地域の実績に基づいて自由に決定出来る。ここでは、神戸市を例に  $\alpha$  及び  $\beta$  を決定する。

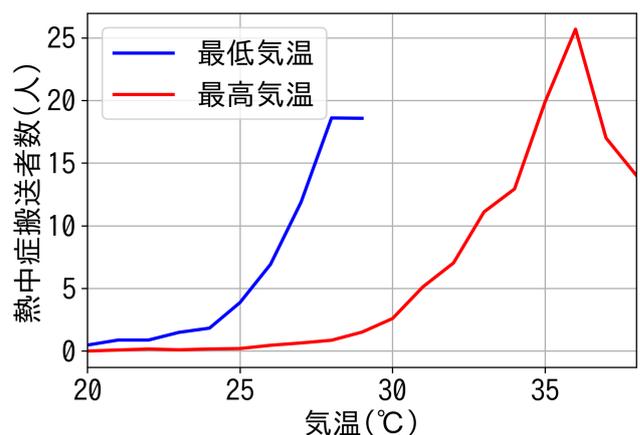


図1 特定の気温における1日の平均熱中症搬送者数

図1は、神戸市における、最低気温及び最高気温と、1日の平均熱中症搬送者数の関係を表している。これより、神戸市では、最低気温が22℃、最高気温が25℃を超えると、徐々に熱中症が発生すると思われる。よって、 $\alpha = 25, \beta = 22$ と設定

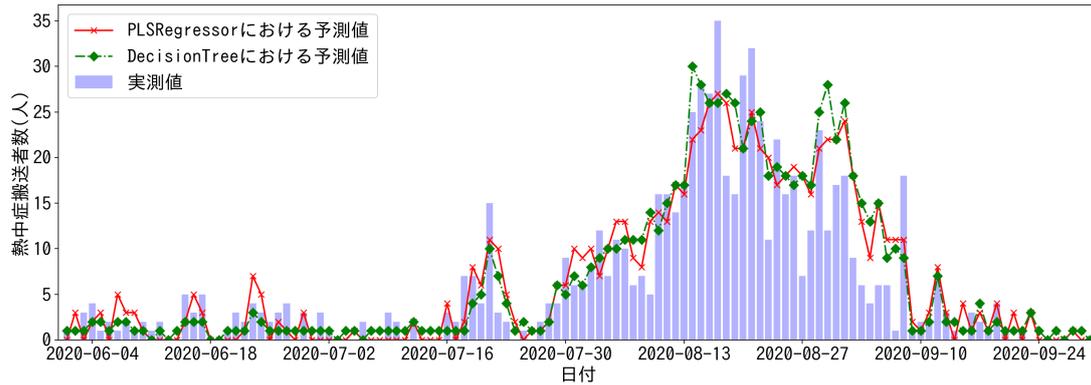


図 2 熱中症搬送者数の予測値と実測値のプロット

する。また、最高気温及び最低気温がそれぞれ  $\alpha$ ,  $\beta$  の値を超えると、熱中症搬送者数が急激に増加することが分かる。この影響を強く反映させるため、 $SqT_{max}(d), SqT_{min}(d)$  に 2 乗式を採用している。一方、最高気温及び最低気温がそれぞれ  $\alpha$ ,  $\beta$  の値を下回った場合、熱中症搬送者が殆ど出ないと考えられる。そのため、 $K_\alpha, K_\beta$  が負の値になりうる条件を与え、その影響を反映させている。

次に、天気の特徴量を変数にすることを考える。天気概要について、表 2 にある通り、「晴後一時曇」や「曇一時雨後晴、雷を伴う」といった複雑な表現が混じり、その種類は 20 以上に及ぶ。これらの表現は正確であるものの、予測精度に悪影響を及ぼす可能性や、週間天気予報との突合が難しいという問題がある。そのため、天気を単純な表現に変換する必要があると考える。また、熱中症には日差しの強さが影響していることから、「晴かどうか」にのみ着目し、以下のように特徴量を定義する。

$$IS_{am}(d) = \begin{cases} 1 & \text{(午前の天気が晴である)} \\ 0 & \text{(午前の天気が晴でない)} \end{cases}$$

$$IS_{pm}(d) = \begin{cases} 1 & \text{(午後の天気が晴である)} \\ 0 & \text{(午後の天気が晴でない)} \end{cases}$$

ここで、天気概要から  $IS_{am}(d), IS_{pm}(d)$  に変換する方法を説明する。天気概要の 1 文字目に着目し、これを午前の天気とする。次に、天気概要に「後」という言葉が存在するか確認する。この「後」という言葉は、天気が時間によって変化するという意味を持っている。「後」が含まれていた場合、その次の文字を午後の天気とする。含まれていない場合は、午前の天気と午後の天気を同じとして扱う。

例として、2020 年 8 月 1 日の天気「曇一時雨後晴、雷を伴う」を変数に変換する。1 文字目の「曇」より、午前の天気は「曇」とする。「後」が含まれているため、その次の文字「晴」より、午後の天気は「晴」とする。これを変数の条件に当てはめると、 $IS_{am}(2020/08/01) = 0, IS_{pm}(2020/08/01) = 1$  となる。

これらの特徴量から、2.1 節の熱中症の発生要因を元に、以下の 6 変数を説明変数として選択する。なお、説明変数の予測日の日付を  $d$ 、予測日より  $n$  日前の日付を  $d - n$  とする。

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

### 3.3 A2-1. 重回帰モデルによる予測モデル構築

機械学習を用いたモデル構築では、予測の元となる説明変数のほかに、予測したい値である目的変数が必要となる。本稿の目的は、熱中症搬送者数を予測することであり、目的変数は任意の日の熱中症搬送者数となる。本章では、神戸市のデータを例として、神戸市における 1 日の熱中症搬送者数を予測する。

説明変数は、気象庁ホームページにおいて 2013 年から 2019 年の各年 6 月から 9 月における各日 (全 854 日) の気象データを csv ファイルとして取得し、Python で 3.2 節の特徴量に加工する。また、得られた特徴量を正規化し、標準正規分布に従うよう、値を変換する。目的変数である各日の熱中症搬送者数は、データベースに保存されている神戸市消防局が持つ救急ビッグデータから、同日の熱中症に対する救急出動回数を取得する。

予測モデルを作成するために、Python のオープンソース機械学習ライブラリである *scikit-learn* を利用する。学習のためのアルゴリズムとして、部分的最小二乗回帰 (Partial Least Squares Regression, PLS) を用いる。PLS は重回帰分析の 1 つであり、説明変数の間の相関が高い場合においても、回帰式を導くことが出来るという特徴がある。今回のように、天気と最高気温のといった、相関の強い説明変数を使うと、回帰係数が不安定になる場合が多い。PLS を用いることで、この現象を解消することが出来る。

以上から、取得した全 854 日の説明変数及び目的変数を入力し、PLS による重回帰予測モデルの構築を行った。また、同様の変数を用いて、従来のモデルとして提案した回帰木 (RegressionTree) モデルを構築する。そして、それぞれのモデルに対して、平均絶対誤差 (MAE) と二乗平均平方根誤差 (RMSE) を計算し、小数点第 3 位を四捨五入することで、モデルの精度を比較する。

### 3.4 予測精度の評価

構築した2つのモデルに対して、モデルから出力された予測値と、実際の救急ビッグデータから得られた実測値を比較することで、モデルの予測精度の評価を行う。評価に利用するデータは、モデル作成のデータと重複しないよう、2020年6月から9月までの全122日の気象データ及び救急ビッグデータとする。3.3節と同様に説明変数を作成し、モデルそれぞれに説明変数を適用することで、熱中症搬送数の予測値を取得した。重回帰モデルが出力する予測値は、小数点以下のもの、負の数になるものがある。本予測対象である熱中症搬送者数は、0以上の整数でのみ表現されることから、出力された予測値に対して、「小数点以下を四捨五入する」「負の値が出力された場合、予測値を0とする」の2つの処理を行う。

図2は、2020年6月から9月までの、熱中症搬送者数の予測値と実測値を表している。横軸は2020年の各日を、縦軸は熱中症搬送者数である。また、各予測モデルから得られた予測値を折れ線グラフで、実際に出勤した実測値を棒グラフで表している。このグラフから、各モデルが出力する値に大きな差異は見られなかった。

表4 回帰モデルを評価する指標

モデル	MAE	RMSE
PLS	2.58	3.65
RegressionTree	2.50	3.92
従来モデル	2.18	3.44

表4は、各予測モデルの誤差を評価する指標を表している。また比較として、従来のモデル及び説明変数での指標を追記している。

PLSとRegressionTreeを比較すると、PLSはMAEが少し大きいものの、RMSEが小さく、予測と実測の大きな誤差が少ないことが分かる。また、PLSと従来モデルを比較すると、精度という点では従来モデルを上回ることが出来なかった。

### 3.5 精度向上の為のヒューリスティック

3.3節で構築したPLSモデルでは、天気「晴」かどうか見えていない。そのため、熱中症リスクを下げる可能性が高い「雨」という要素を排除している。これにより、従来モデルより精度が落ちたと考えられる。

PLSモデルの精度向上を目指し、ヒューリスティックを導入する。ヒューリスティックとは、経験則や直感から答えを導き出すことである。本稿では、消防局員への聞き取りと、試行錯誤的なパラメータ調整から、「雨の日は、出力される予測値を0.85倍する」というヒューリスティックを導入する。また、雨が降ったかどうかは、天気概要において、「雨」という文字が含まれるかで判定する。これにより、にわか雨などによる熱中症発症リスク低下の影響を反映できると考えている。

ヒューリスティックを導入した場合の指標を、表5に示す。ただし、従来モデルでは雨という情報を利用していることから、ヒューリスティックの適応はしていない。これより、提案手法におけるモデルの精度が改善したことが分かる。また、PLSが

表5 ヒューリスティックを導入した場合の指標

モデル	MAE	RMSE
PLS	2.45	3.41
RegressionTree	2.43	3.80
従来モデル	2.18	3.44

従来モデルより、最小二乗誤差が小さいことも分かる。

### 3.6 A2-2. モデルを使った推論

熱中症搬送者数の予測値の値域を表すには、本来の予測値より大きく出る、少なく出る場合の説明変数が必要となる。モデルに使う説明変数の種類は変更できないため、その中身を変化させることを検討する。

予測値の値域の上部を表すには、気温の予測範囲の最大値と、天気が「晴」に変化したと仮定して、説明変数を変更する。逆に値域の下部を表すには、気温の予測範囲の最小値に変化したと仮定し、説明変数を変更する。

表6 予測値の値域の上限を予測する説明変数の例

日付	25日	26日	27日	28日
天気(上端)	晴	晴	晴	晴
最高気温(上端)	34	35	34	34
最低気温(上端)	27	27	27	27
天気(下端)	晴時々曇	晴時々曇	晴時々曇	晴時々曇
最高気温(下端)	31	31	31	30
最低気温(下端)	25	24	24	23

説明変数を変更する例を、表6に表す。これは、表2の気温の予測範囲を参考にした。このようにして、予期される熱中症リスクがより上がった場合、より下がった場合を表現出来る。

## 4. RQ2: 神戸市において、RQ1での提案手法は有効か?

本章では、提案手法を用いて構築したモデルを活用し、実際の天気予報を元にした予測が可能か確認する。また、天気予報の誤差がどの程度予測精度に影響するか、指標を用いて評価する。

本稿では、神戸市における2013年から2019年の各年6月から9月の救急ビッグデータ及び過去の気象データを用いて、モデル構築を行う。天気予報のデータとして、2020年5月1日から9月30日各日17時に発表された週間天気予報を活用する。このデータは、神戸地方気象台より、気象庁保有の情報を提供して頂いた。

初めに、週間天気予報と実際の気象データを比較し、天気予報の信頼性を評価する。これにより、中長期予測の入力精度を評価する。次に、実際に構築したモデルに週間天気予報を入力し、熱中症搬送者数の予測値を出す。最後に、予測値と実測値の誤差を、MAE及びRMSEを用いて評価する。

### 4.1 取得したデータ

天気予報のデータには、各予報発表日の2日後から7日後の天気、最高気温、最低気温、気温の予測範囲が含まれる。その中から、特定の日数後における、2020年6月1日から9月30日

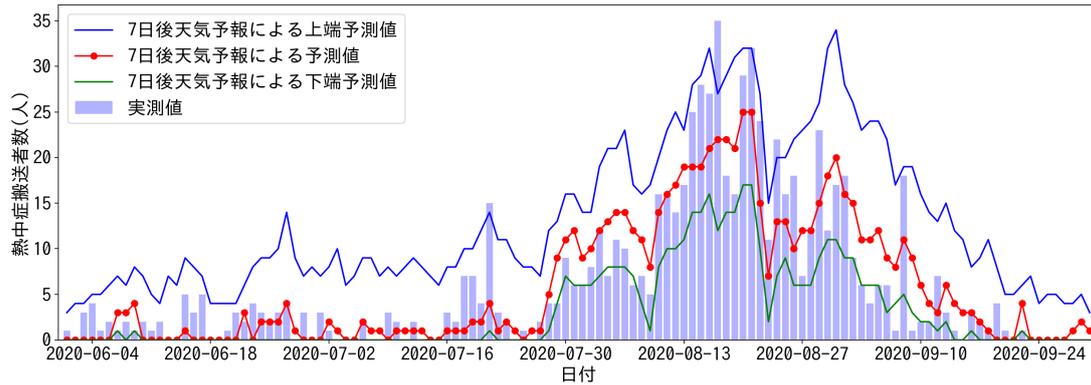


図3 週間天気予報を用いた、2020年における熱中症搬送者数の予測値とその値域

に対する122日分の予報データのみ抽出し、実際の気象データと比較する。例えば、2日後の天気予報データを抽出する場合、5月30日に発表した2日後(6月1日)の予報、5月31日に発表した2日後(6月2日)の予報と抽出し、最後に9月28日に発表した2日後(9月30日)の予報を取得し、122日分のデータを分析する。加えて、気温の予測範囲内に実際の気象データがあるかどうか、その的中率と予測範囲の平均値を分析する。

表7 週間天気予報の気温の誤差

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

表7は、2日後から7日後の気温の予報誤差を評価する指標を表している。これより、予報日より遠い日ほど、気温の予報に誤差が生じやすいことが分かる。また、予報日に近い予報であっても、必ず当たるというわけではないことが分かる。

表8 週間天気予報の気温の誤差

日付	Tmax		Tmin	
	的中率 (%)	平均範囲	的中率 (%)	平均範囲
2日後	77.0	3.30	77.0	2.49
3日後	81.1	3.49	77.0	2.70
4日後	74.6	3.58	77.0	2.81
5日後	79.5	3.94	79.5	3.01
6日後	77.0	3.93	77.0	3.26
7日後	76.2	4.25	82.8	3.57

また、表8は、2日後から7日後の気温の予測範囲を評価する指標を表している。これより、気温の予測範囲の的中率は、約78%で推移している。しかし一方で、予報日より遠い日ほどその予測範囲が広がっており、気温の予測精度が低いことが分かる。

#### 4.2 天気予報を活用した中長期予測

本節では、中長期の予測評価のため、4日後以降の熱中症予測を行い、実際の熱中症搬送者数と比較する、その精度を評価する。本稿で取得した週間天気予報データは、予報発表日の7日後までであることから、4日後から7日後の予測を行う。

評価に利用するデータは、神戸市における2020年6月1日から9月30日に対する122日分の天気予報データとする。各日のデータを元に、予測値の出力及びヒューリスティックを適用した結果を、神戸市の同日の熱中症に対する救急出動回数と比較する。

表9 週間天気予報を用いた予測精度

予測日	MAE	RMSE
4日後	2.89	4.23
5日後	2.93	4.30
6日後	2.98	4.29
7日後	2.79	3.87

表9は、特定日後の予測の誤差を評価する指標を表している。これから、4日後から6日後の予測では、遠い日ほど精度が落ちることが分かる。しかし一方で、7日後は他の日より精度が高いことが分かる。

#### 4.3 予測幅を活用した予測の的中率

本節では、3.6節で提案した予測値の値域を、実際の週間天気予報から予測し、実際の熱中症搬送者数が値域に含まれているか確認することで、その精度を評価する。期間やデータは4.2節と同様、4日後から7日後の予測を行う。精度を評価する項目として、熱中症搬送者数の実測値が予測値の値域に含まれている割合を分析する。また、予測値の値域の上端と下端の差を予測値の値域の大きさとし、その平均値を分析する。

表10 予測幅を活用した予測の的中率とその大きさ

予測日	的中率 (%)	平均範囲
4日後	83.6	8.88
5日後	82.0	8.96
6日後	83.6	9.22
7日後	93.4	10.03

表 10 は、特定日後の値域に関する指標を表している。これより、4 日後から 6 日後のデータに比べ、7 日後のデータ的中率が高いと分かる。しかし、予測値の値域の大きさは、遠い日ほど大きくなり、7 日後の精度が良いとは言えない。

また、図 3 は、7 日後の天気予報を利用した、熱中症搬送者数の予測値の値域と実測値を表している。これより、実測値が予測値の上端を上回る日、予測値の下端を下回る日が見られた。

## 5. 考 察

RQ1 では、週間天気予報に存在する要素のみを用いて、新たな熱中症予測モデルを構築した。地域性を反映させられる特徴量の開発により、神戸市以外の地域でも活用出来ると考えている。また、予測範囲を示すことで、悲観的な事象を想定した対策が期待出来る。加えて、PLS による説明可能なモデル構築により、従来モデルでは難しかった予測過程の解釈が容易になった。一方で、従来モデルと比較して、本稿で構築したモデルは平均誤差が大きい。これは、従来モデルで説明変数として利用していた「1 日の最低湿度」が、熱中症に強く影響する要素であるからと考えられる。より精度を上げるため、天気と気温を組合わせた新たな特徴量を提案し、最低湿度の代わりとなる説明変数を検討することが、今後の課題として挙げられる。

RQ2 では、初めに天気予報の精度について分析した。これより、気温の予測精度は、中期的になるほど悪くなること、短期の予測においても必ず当たるわけではないことが分かった。

次に、本稿で作成した予測モデルを利用し、実際の天気予報を元にした熱中症予測を行った。これより、天気予報を元にした予測は、気象の実測値を元にした予測より精度が落ちること、より中期の予測ほど精度が落ちる事が分かった。また、7 日後の予測が、他の日の予測よりも精度が良い結果が得られた。他の予測と比較すると、8 月 9 日、8 月 27 日などが、7 日後予測が他の日の予測より精度がいいと分かった。

8 月 9 日に関して、過去の気象を用いた予測においても誤差が大きいことが分かっている。この日は 3 連休の 2 日目であることから、神戸市外への外出や、コロナ禍による在宅が、神戸市における熱中症数の減少に影響していると考えられる。7 日後予測が実測値に近い値を出しているのは、天気予報において、実際は晴れであったが、「くもり」と予測していたことから、予測値が下がったと考えられる。

8 月 27 日に関しても同様に、過去の気象を用いた予測においても誤差が大きいことが分かっている。該当日の気温の実測値が高く、経験則からも熱中症発症のリスクが高いと考えられるが、実際熱中症者数は少なかった。この原因として、昼過ぎから雷が多く落ちたことから、外出を控えた可能性が考えられる。加えて、暑熱順化という、人間が暑さに慣れていく現象により、他の日より熱中症数が減少したと考えられる。7 日後予測が実測値に近い値を出しているのは、実測値より約 2℃ほど低く予報していたためであると考えられる。以上のことから、天気予報の誤差が、予測の精度に良い影響も、悪い影響も与えることが分かった。

最後に、気温の予測範囲を活用した、値の範囲予測を行った。

これより、中期の予測になるほど、予測範囲が大きくなるが、その的中率に変化がないことが分かった。図 3 の予測を見ると、7 月 21 日の実測値は予測範囲を上回り、9 月 9 日の実測値は予測範囲を下回っている。

7 月 21 日に関して、実際の気温は予測範囲内に収まっていたが、2020 年で初めて 30℃を超える日であり、前日から一気に気温が上がったこともあり、予測範囲以上の熱中症が発生したと考えられる。9 月 9 日に関して、予報には存在しない雨が降ったことで、昼間の気温があまり上がらず、予報よりも気温が低くなったことが要因として上げられる。

以上より、週間天気予報による熱中症搬送者数の中期予測は可能であったが、予報の誤差による影響も強く、予測範囲が大きくなってしまふことが分かった。

今後、この予測が実際の救急医療現場に対して有効かどうか、救急施策として活用可能かどうかを検討していく。

## 6. おわりに

本稿では、4~7 日後の中期的な熱中症搬送者数を予測するために、週間天気予報に含まれる項目のみを説明変数とした、説明可能な予測モデルを構築した。天気予報から、地域性を反映出来る特徴量を開発することで、特定地域によらないモデルが開発可能である。

加えて、実際の神戸市の週間天気予報を利用し、中期的な予測を行った。7 日後の予報を元に予測を行った結果、実際の熱中症者数との平均誤差が 2.80 であった。

今後の展望として、中期的な予測に有効な、新たな説明変数を模索する。同時に、構築したモデルを用いた予測が救急医療現場でどのように活用出来るか、実際に利用してもらい評価してもらおう。

謝辞 本研究の一部は JSPS 科研費 JP19H01138, JP18H03242, JP18H03342, JP19H04154, JP19K02973, JP20K11059, JP20H04014, JP20H05706, 立石科学技術振興財団の研究助成 及び、救急振興財団の救急に関する調査研究事業助成を受けて行われている。

本研究は神戸市消防局との共同研究の一環で行われている。

### 文 献

- [1] M. Deakin and H. Al Waer, "From intelligent to smart cities," Intelligent Buildings International, vol.3, no.3, pp.140-152, 2011.
- [2] 環境省, "熱中症環境保健マニュアル 2018," [https://www.wbgt.env.go.jp/heatillness\\_manual.php](https://www.wbgt.env.go.jp/heatillness_manual.php). (Accessed on 10/12/2020).
- [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] "気象庁 過去の気象データ検索," <http://www.data.jma.go.jp/obd/stats/etrn/index.php>. (Accessed on 10/12/2020).
- [5] "気象庁 週間天気予報," <http://www.jma.go.jp/jma/kishou/known/kurashi/shukan.html>. (Accessed on 07/23/2021).
- [6] 中井哲也, 佐伯幸郎, 中村匡秀, "救急出動記録を活用した熱中症に関連する救急需要の分析と予測," 電子情報通信学会技術報告書, vol.120, no.232, pp.13-19, 2020.