

令和3年7月27日

報道各位

国立循環器病研究センター
関西大学
国立環境研究所

AIと天気情報等の活用による熱中症発症数の高精度予測

－ 熱中症発症数 AI 予測モデル開発の成功

国立循環器病研究センター（大阪府吹田市、理事長：大津欣也、略称：国循）の予防医学・疫学情報部の尾形宗士郎（上級研究員）、西村邦宏（部長）らと、関西大学（大阪府吹田市、学長：前田裕）の環境都市工学部の尾崎平（教授）、北詰恵一（教授）らと、国立環境研究所（茨城県つくば市、理事長：木本昌秀）の山崎新（エコチル調査コアセンター長）、山形与志樹（地球環境研究センター客員研究員）らの研究グループは、人工知能（AI）技術の機械学習を用いて、気象データ等から熱中症発症数を高精度に予測するAIモデルを世界で初めて作成し、Nature Communications 誌に2021年7月28日付で掲載されました。

当AIモデルの特徴は下記通りです。

- ・熱中症による救急搬送の全症例件数と中等症以上症例件数（入院診療、長期入院、死亡例）のそれぞれを、12時間毎に市町村単位で予測することができます。
- ・熱中症発症数が急上昇する日（ピークの日）も高精度に予測することができます。
- ・当AIモデルでは、天気情報と暦情報と市町村の統計情報（人口、高齢化率等）といったルーティンで収集されている情報をAIで扱うことにより、熱中症発症数を予測することができます。当AIモデルの実装に求められる情報は多くの地域で収集しているため、実装は比較的容易と考えます。

■概要

背景

地球温暖化にともない猛暑日が増加し、熱中症発症リスクが増加すると想定されます。近年、世界各地において熱波が発生し多数の熱中症が報告されています。また、気温と湿度と日射量を複合したWBGT（湿球黒球温度）は、気候条件から熱中症リスクを示す指標として広く使用されています。しかし、WBGTで熱中症リスクを評価すると、日本の7月—8月の多くの日がハイリスクとなり、熱中症リスクの詳細な評価が困難でした。そこで、本研究は市町村の消防署より提供を受けた匿名化済みの熱中症搬送情報、The Weather Company社による高解像度気象データ、AI技術の機械学習を用いて、12時間毎の市町村単位の熱中症発症者数を高精度に予測するAIモデルを作成しました。

方法

近畿地方 16 市町村の 2015、2016、2017 年の 6 月から 9 月のデータを予測モデル作成のための訓練データセットとし（熱中症による救急搬送の全症例件数=11349 件）、2018 年の 6 月から 9 月のデータを予測モデルの精度を検証するためのデータセットとしました（熱中症による救急搬送の全症例件数=7513 件）。天気情報と暦情報と公開されている市町村の統計情報（人口、人口男女比、高齢化割合、緑地面積等）を特徴量とし、熱中症による救急搬送の全症例件数と中等症以上症例件数（入院診療、長期入院、死亡例）を教師データとし、複数の機械学習アルゴリズムで（一般化線形モデル、一般化加法モデル、ランダムフォレスト、勾配ブースティング）、AI 予測モデルを作成しました。加えて、比較のために、古典的な統計モデルである一般化線形モデルと WBGT と市町村人口のみで作成した古典予測モデルも作成しました。作成した予測モデルと 2018 年の天気情報と暦情報と市町村の統計情報によって、2018 年における 12 時間毎の市町村単位の熱中症発症数の予測値を求め、それに対応する実測数との差をみることで、作成した予測モデルの精度を評価しました。市町村単位の 12 時間毎の予測精度は二乗平均平方根誤差（RMSE）で評価し、熱中症発症数が多い日の全市町村の 1 日単位の予測精度は平均絶対パーセント誤差（MAPE）で評価しました。なお、RMSE、MAPE とともに低いほど予測精度が高いことを示します。

結果

熱中症による救急搬送の全症例件数に対して、古典的な予測モデルでは、RMSE が 3.73（件数）で MAPE が 43.0%でした。一方で、我々が作成した最良の AI モデルでは、RMSE が 2.97（件数）で MAPE が 14.8%でした。その結果をわかりやすく示したものが図 1 になります。

図 1 は、熱中症による救急搬送の全症例件数の観測件数（黒線）と予測件数（赤線）の比較を示し（16 市町村の一日単位の合算値）、黒色と赤色の線が近いほど予測精度が良いことを示します。2015 年から 2017 年は予測モデル作成をした訓練データセットでの結果となっており、一般的に訓練データセットでの予測精度は高いものが得られます。着目すべき箇所は、2018 年のモデル精度検証のための検証データセットの結果になります。図 1 の上部は古典的な予測モデルの予測結果を示し、熱中症発症数ピーク時の予測精度が低いです。しかし、図 1 の下部の AI 予測モデルでは、ピーク時も含めて観測件数と予測件数が非常に類似していることが示され、予測精度が高いことが示されています。

また、熱中症による救急搬送の中等症以上症例件数（入院診療、長期入院、死亡例）に対して、古典的な予測モデルでは、RMSE が 1.14（件数）で MAPE が 37.7%でした。一方で、我々が作成した最良の AI モデルでは、RMSE が 0.83（件数）で MAPE が 10.6%でした。

結論と意義

本研究において、天気情報と暦情報と市町村の統計情報を用いた熱中症発症数予測モデルは、熱中症による救急搬送の全症例件数と中等症以上症例件数（入院診療、長期入院、死亡例）を高精度に予測することに成功しました。気象条件と熱中症発症数が関連することは、既に多くの論文で報告されていますが、本研究のように熱中症発症数・重症度・発症数ピークを高精度に予測する AI モデルは世界初の研究結果となります。

本 AI モデルは天気情報と暦情報と市町村の統計情報といったルーティンで収集されている情報を用いて熱中症発症数を予測することができるので、本 AI モデルは比較的容易に社会実装できると考えています。将来的に熱中症アラートを高精度に発信することで、多くの方の熱中症予防につながることを期待しています。

■ 発表論文情報

著者 : Soshiro Ogata, Misa Takegami, Taira Ozaki, Takahiro Nakashima, Daisuke Onozuka, Shunsuke Murata, Yuriko Nakaoku, Koyu Suzuki, Akihito Hagihara, Teruo Noguchi, Koji Iihara, Keiichi Kitazume, Tohru Morioka, Shin Yamazaki, Takahiro Yoshida, Yoshiki Yamagata, Kunihiro Nishimura
題名 : Heatstroke predictions by machine learning, weather information, and an all-population registry for 12-hour heatstroke alerts
掲載誌 : Nature Communications

■ 謝辞

本研究は、下記機関より資金的支援を受け実施されました。

- ・ 独立行政法人環境再生保全機構の環境研究総合推進費（JPMEERF20191005）
- ・ 国立研究開発法人国立循環器病研究センター 循環器病研究開発費（30-6-15）

図 1

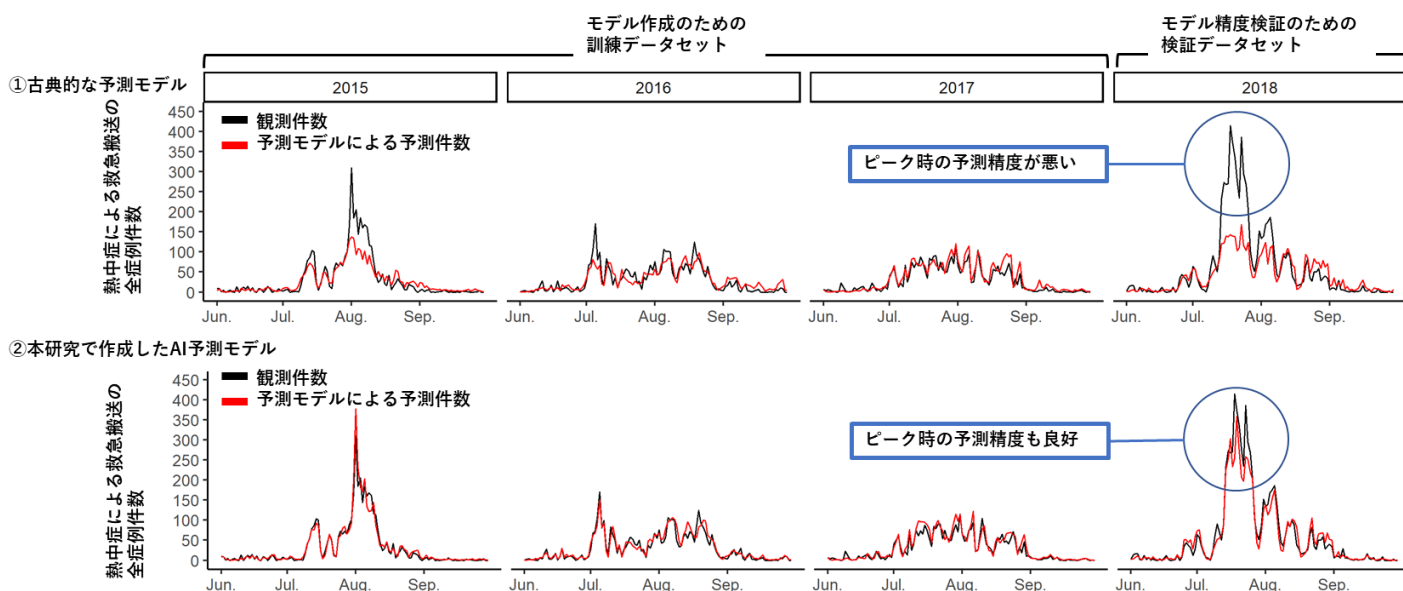


図 1. 熱中症による救急搬送の全症例件数の観測件数と予測件数の比較（16市町村の一日単位の合算値）。

図の解説：黒色と赤色の線が近いほど予測精度が良いことを示します。2015年から2017年は予測モデル作成をした訓練データセットでの結果となっており、一般的に訓練データセットでの予測精度は高いものが得られます。着目すべき箇所は、2018年のモデル精度検証のための検証データセットの結果になります。図 1 の上部は古典的な予測モデルの予測結果を示し、熱中症発症数ピーク時の予測精度が低いです。しかし、図 1 の下部の AI 予測モデルでは、ピーク時も合わせて観測件数と予測件数が非常に類似していることが示され、予測精度が高いことが示されています。

Nature Communicationsに掲載の図を改変し掲載しています。

※この報道資料は、大阪科学・大学記者クラブ、環境省記者クラブ、環境記者会、筑波研究学園都市記者会の皆様にお届けしています。

【報道機関からの問い合わせ先】

国立循環器病研究センター（代表）06-6170-1070
総務課広報係（内線 31120）大谷・樹崎