

ディープラーニングを用いた 移動測定局におけるPM_{2.5}濃度の予測

Prediction of PM_{2.5} concentrations at mobile air quality monitoring stations using deep learning

太田 耕右*¹ 大熊 一也 天野 直哉 佐久間 隆 三沢 松子

Kohsuke OHTA, Kazunari OOKUMA, Naoya AMANO, Takashi SAKUMA, Matsuko MISAWA

微小粒子状物質（以下、「PM_{2.5}」）による呼吸器・循環器系への影響に関する懸念が近年高まっている。県民への早急な注意喚起の発令のため、濃度予測手法の確立が必要であると考えられるが、大気環境移動測定車による測定地点（多賀城及び七ヶ浜）においては測定データが少なく、予測を行なうためのデータ量が不足している。本研究では、AI（人工知能）技術の一種であるディープラーニングを用いて、両地点の周辺に位置する常時監視測定局のデータを基に作成した拡張データを学習させ、両地点における7日後のPM_{2.5}濃度（1時間値）を予測した、その結果、AIに学習させるデータ（以下、「教師データ」）の種類によって異なる予測精度が得られた。特に、両地点の周辺に位置する測定局の過去のPM_{2.5}若しくは浮遊粒子状物質（以下、「SPM」）の濃度データを学習させた場合により高い精度が得られ、両地点において、実測値と予測値の相関係数は約0.7となった。

キーワード：大気汚染常時監視；PM_{2.5}予測；ディープラーニング；再帰型ニューラルネットワーク；拡張データ
Key words：Air pollution monitoring；Predictoin of PM_{2.5}；Deep learning；Recurrent neural network；Data augmentation

1 はじめに

PM_{2.5}は代表的な大気汚染物質であり、呼吸器・循環器系疾患の発生率や死亡率を有意に増加させることが明らかになっている。そこで、県民への早急な注意喚起の発令や大気汚染の未然防止に向けて有用な情報を得るためにも、PM_{2.5}濃度の予測は重要であると考えられる。

また近年、AI技術が環境分野においても活用されるようになってきている^{1,2)}。環境分野におけるAIの活用は数十年前から試みられてきたが、情報工学分野以外の研究者が容易にAIを利用できる環境ではなく、また予測精度も十分ではなかったため、実用化は進まなかった。しかし最近では、家庭用の一般的なコンピュータに無料で高性能なAIを搭載できる環境が整ってきており、AIを用いて大気汚染物質の濃度予測を行う試みが報告されるようになった。

著者らはこれまで、ディープラーニングの一種である再帰型ニューラルネットワーク（以下、「RNN」）を用いて、塩釜保健所岩沼支所屋上にある常時監視測定局（岩沼局）の測定結果を教師データとして、同地点における1週間後の光化学オキシダント（以下、「Ox」）及びPM_{2.5}濃度の予測に取り組んできた³⁻⁵⁾。

これらの研究においては、ある程度の予測精度（実測値と予測値との相関係数 r が0.70以上）を確保するために、数年分の1時間値データが教師データとして必要であった。しかしながら、固定の測定局がない地域や、数

年分の教師データがない地点においても、濃度予測の必要性は生じてくる。



図1 大気環境移動測定車による測定イメージ

（2020年度撮影。なお、2017～2019年度は車種・測定項目が異なる）



図2 移動測定局（多賀城・七ヶ浜）及び周辺局（福室、塩釜）の位置関係

*1 現 東部保健福祉事務所登米地域事務所

宮城県多賀城市及び七ヶ浜町においては、大気環境移動測定車(図1)を用いて、2017年度から図2に示す2地点(多賀城・七ヶ浜)において測定しているが、2017年度から2019年度においては測定期間は1年のうち4週間(1週間×4回)であった。このような測定データが少ない地点についても、濃度の予測が可能となれば、大気汚染常時監視体制の強化につながるため、予測手法の開発が必要と考えられる。

以上のことから、本研究では、取得可能なデータ量が限られる地点における大気汚染物質濃度の予測を目的とし、ディープラーニングを用いて多賀城市及び七ヶ浜町における7日後のPM_{2.5}濃度(1時間値)の予測を試みた。

2 予測方法

一般的なAI技術の区分を図3に示すが、AI技術の一種である機械学習のうち、多層化した(複雑な)ニューラルネットワークを使用したものが「ディープラーニング」と呼ばれている。本研究では、教師データによる機械学習モデルの構築には、ディープラーニングのひとつであるRNNを用いた。RNNは、時系列データの予測を得意とする手法であり、気象、株価、顧客行動等の予測に幅広く活用されている。

図4にRNNの模式図を示す。通常のニューラルネットワークでは、ある層の出力は、次の層の入力に利用されるのみである。一方で、RNNでは出力の一部が「再帰的に」同じセルに戻る(図4中、隠れ層における赤色の矢印)。この結果、ある時点での入力が、それ以降の出力に影響を及ぼし、結果として時系列データの処理を行うことができる。

開発環境についてはJupyter notebook (Python3.5)を使用し、RNNの実装においてはTensorFlow 1.2.1を使用した。PC構成はプロセッサにIntel Core i5、CPUに3.30GHz×2、メモリ4GBを搭載したものをを用いた。

予測地点は、図2における2地点(多賀城・七ヶ浜)とし、教師データとして2017年5月30日1時~2019年6月4日24時を使用し、Googleが提供するRNNを使用してディープラーニングを行った。

予測精度の評価には、大気環境移動測定車による測定日(表1)のうち2019年度分のデータを使用した。予測対象は1週間(168時間)後の多賀城又は七ヶ浜におけるPM_{2.5}濃度(1時間値)とし、168時間前~現在までの連続データを用いて予測した。

教師データの種類については、PM_{2.5}二次粒子の生成機構⁶⁾(図5)等を参考に、周辺の常時監視測定局(福室局及び塩釜局、図2参照)における長期欠測のないデータを選定し、以下のとおりとした。

- ①PM_{2.5} [μg/m³] (多賀城又は七ヶ浜)
- ②Ox [ppb] (福室局)
- ③NOx [ppb] (福室局)
- ④PM_{2.5} [μg/m³] (福室局)

- ⑤Ox [ppb] (塩釜局)
- ⑥NOx [ppb] (塩釜局)
- ⑦SPM [μg/m³] (塩釜局)
- ⑧遷移状態が生じない場合における気相中の分子同士の衝突による化学反応論⁷⁾における、反応速度定数の温度関数部分である $T^{0.5} \times \exp(-E/RT)$ (以下「K」という。)

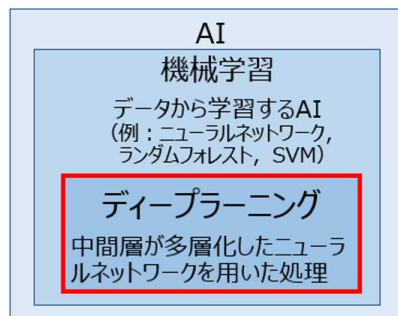
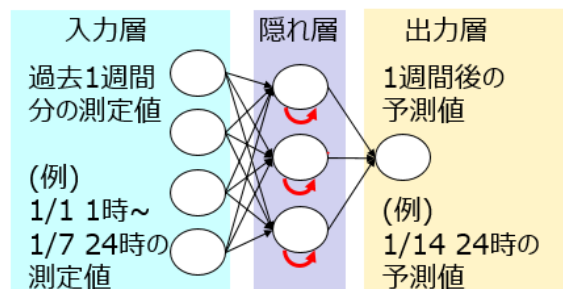


図3 AI技術の区分



※データの入出力はいずれもExcel形式

図4 RNNの模式図

表1 大気環境移動測定車による測定期間

	2017 (H29)年度	2018 (H30)年度	2019 (R元)年度
多賀城	6/6~6/12	6/19~6/25	6/20~6/26
	7/21~7/27	8/14~8/20	9/13~9/19
	10/20~10/26	10/11~10/17	11/12~11/18
七ヶ浜	1/26~2/1	2/13~2/19	2/11~2/17
	6/30~7/6	5/25~5/31	6/12~6/18
	7/13~7/19	8/23~8/29	9/21~9/27
	10/12~10/18	10/19~10/25	11/20~11/26
	2/15~2/21	2/21~2/27	2/19~2/25

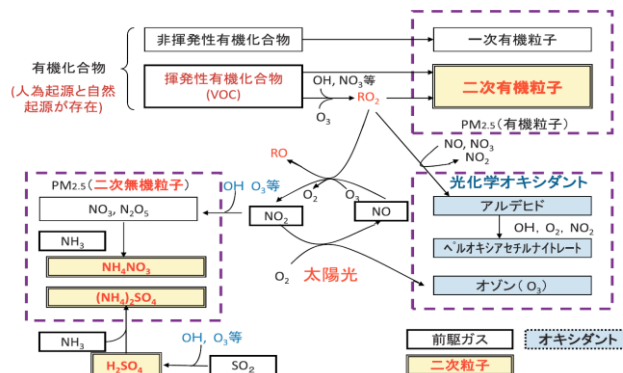


図5 PM_{2.5}二次粒子の生成経路⁶⁾

⑧の補足として、衝突反応の模式図及び反応速度定数の式を図6に示す。ここで、Tの値は仙台管区気象台における気温[K], Rは気体定数[J/(K・mol)], Eは活性化エネルギー[J/mol]を表す。なお、PM_{2.5}の生成反応に係るEの値については一般的に知られていないため、本研究ではNO_xの酸化還元反応に係るEの値⁸⁾を参考に、E=200[kJ/mol]と仮定して各時間におけるTの値からKを計算した。



$$\text{反応速度定数} = \frac{\sigma(8k/\pi\mu)^{1/2}}{\text{定数}} \times (T)^{1/2} \exp(-E/RT)$$

Tの関数：以下Kと定義

σ: 反応断面積 k: ボルツマン定数
 μ: 実効質量
 T: 気温 R: 気体定数
 E: 反応の活性化エネルギー
 (ここでは200[kJ/mol]と仮定)

以上のうち少なくとも①PM_{2.5}[μg/m³] (多賀城又は七ヶ浜)は教師データとして採用し、①に加えて②～⑧のいずれか又はすべてを教師データとした場合における、予測精度である実測値と予測値との相関係数(以下、「r」)の変化を調べた。

なお、教師データ(2017年5月30日1時～2018年12月31日24時)のうち大気環境移動測定車による測定が行われなかった期間における①PM_{2.5}[μg/m³] (多賀城又は七ヶ浜)については、以下の方法で拡張データを作成し、教師データとした。

図6 気相分子同士の衝突反応の模式図⁷⁾

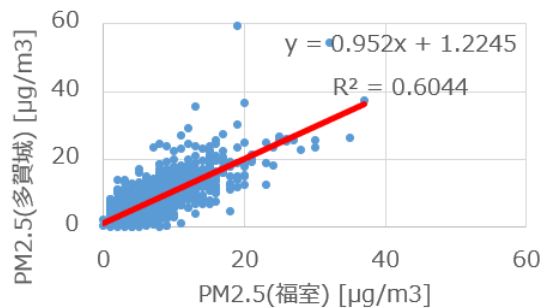


図7 2017～2018年度の測定期間における多賀城PM_{2.5}と福室PM_{2.5}の関係

- 1) 大気環境移動測定車による測定期間(表1)のうち、教師データに使用されている2017及び2018年度の測定期間において、周辺局のうち最も欠測期間の少ない福室局のPM_{2.5}濃度と、多賀城及び七ヶ浜におけるPM_{2.5}濃度との回帰式(x: 福室局のPM_{2.5}濃度, y: 多賀城又は七ヶ浜におけるPM_{2.5}濃度, としたとき, y=ax+bにおける定数a,b)を求めた。なお、求められた回帰式は以下のとおりであった(図7, 8)。

$$\text{多賀城 PM}_{2.5} = 0.952 \times \text{福室 PM}_{2.5} + 1.2245$$

$$\text{七ヶ浜 PM}_{2.5} = 0.7332 \times \text{福室 PM}_{2.5} + 2.6808$$

- 2) 多賀城及び七ヶ浜のデータを用いてそれぞれ得られた回帰式を用いて、大気環境測定車による測定が行われなかった期間における多賀城及び七ヶ浜のPM_{2.5}濃度を計算し、拡張データとした。

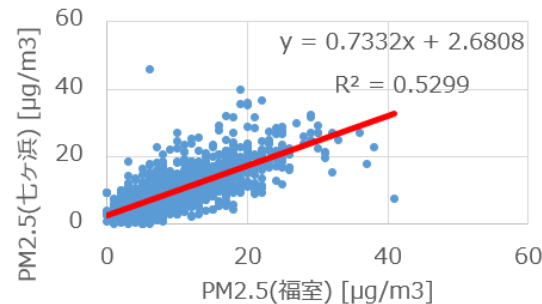


図8 2017～2018年度の測定期間における七ヶ浜PM_{2.5}と福室PM_{2.5}の関係

3 結果・考察

予測精度の指標であるrと、教師データの種類との関係を表2に示す。表2より、多賀城においては福室局のPM_{2.5}を、七ヶ浜においては塩釜局のSPMを学習させた場合にrが最も高い値となり、実測値をある程度再現できたことが確認された(図9, 10)。

表2 教師データの種類とrとの関係

教師データの種類※	実測値と予測値との相関係数	
	多賀城	七ヶ浜
福室Ox	0.39	0.47
福室NOx	0.68	0.62
福室SO ₂	0.66	0.67
福室PM_{2.5}	0.72	0.26
塩釜Ox	0.46	0.48
塩釜NOx	0.67	0.60
塩釜SPM	0.65	0.70
K	0.71	0.66
すべて	0.22	0.13

この理由としては、2017～2019年度において多賀城では西寄りの風、七ヶ浜では北寄りの風の頻度が高かったことから、風によるPM_{2.5}及びSPMを含む大気の影響が影響すると推察された(図11)。

また、Kを教師データとして用いた場合にも比較的高いrの値となったため、PM_{2.5}の生成機構には衝突反応が関連していると考えられる。

さらに、用意したデータ(2の①～⑧)をすべて教師データとした場合、rの値が最も小さくなった。このことから、ディープラーニングでは教師データの種類が多

※このほか、多賀城もしくは七ヶ浜のPM_{2.5}(拡張データと実測データ)も合わせて教師データとして使用

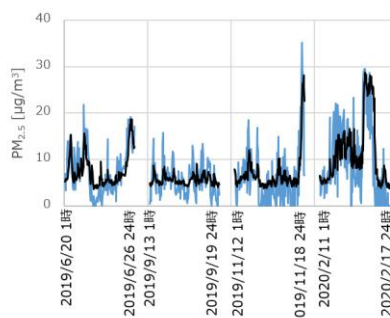
いほど精度が向上するというわけではなく、関連度の高い物質を選定する必要があることが示唆された。

5 まとめ

本研究では、RNN を用いて移動測定局における 1 週間後の PM_{2.5} 濃度の 1 時間値を予測した。結果、教師データが少ない場合でも、周辺局における測定データをもとに拡張データを作成しデータを補完すれば予測は可能であった。またこの方法は、光化学オキシダントといった他の物質の濃度予測にも適用できる可能性がある。

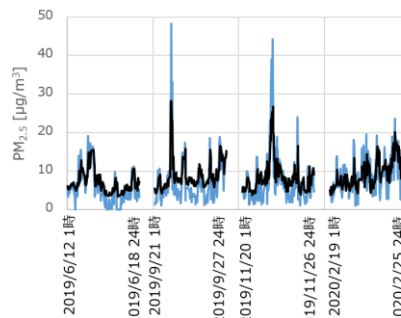
参考文献

- 1) 井上ら, 第 60 回大気環境学会年, 2019, P-144#.
- 2) 荒木ら, 第 60 回大気環境学会年会, 2019, 2D0900.
- 3) 太田ら, 第 61 回大気環境学会年会, 2020, O-L-009.
- 4) 太田耕右, 宮城県保健環境センター年報, 2020(38), p.43-46.
- 5) 太田ら, 第 46 回全国環境研協議会北海道・東北支部研究発表会, 2021.
- 6) 坂本和彦, 表面科学, 2015(36), p.141-143.
- 7) P. Atkins, アトキンス物理化学, 2009, p.935.
- 8) J. Nicholas, “Chemical kinetics”, 1976.



—実測値(多賀城) —予測値(多賀城)

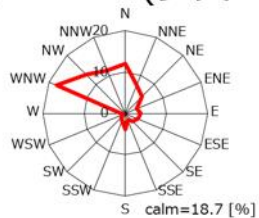
図 9 多賀城 PM_{2.5} と福室 PM_{2.5} を用いた場合の予測値と実測値



—実測値(七ヶ浜) —予測値(七ヶ浜)

図 10 七ヶ浜 PM_{2.5} と塩釜 SPM を用いた場合の予測値と実測値

風向頻度(多賀城)



風向頻度(七ヶ浜)

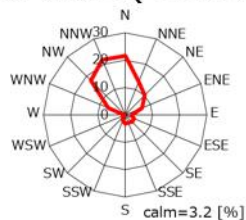


図 11 多賀城と七ヶ浜における風向頻度と位置関係