明星大学　理工学部　総合理工学科　環境科学系

令和　6年度卒業論文

**単一地点におけるDNNを用いたOx短期予測の**

**最適パラメータ探索**

学籍番号　21T7-008

氏名　今給黎　樹

研究室名　大気科学研究室

指導教官　櫻井　達也

要旨

目次

1. 研究の背景と目的
2. 機械学習とは
   1. 機械学習
   2. ニューラルネットワーク
   3. DNN
3. 研究手法
   1. 対象地点
   2. 使用データ
   3. モデル詳細
   4. 評価方法
      1. モデル評価
      2. 特徴量評価
   5. 特徴量の探索
      1. 全データによる学習
      2. 高濃度域の上位10個による学習
      3. 高濃度域の上位20個による学習
      4. 高濃度域の上位30個による学習
      5. 高濃度域と低濃度域の上位20個による学習
4. 結果
   1. 全データによる学習
   2. 高濃度域の上位10個による学習
   3. 高濃度域の上位20個による学習
   4. 高濃度域の上位30個による学習
   5. 高濃度域と低濃度域の上位20個による学習
5. 考察
6. まとめ
7. 謝辞
8. 参考文献
9. 研究の背景と目的

光化学オキシダント(Ox)は大気汚染物質の一つである。これが大気中に滞留すると上空がもやがかかったようになり、視界が悪くなる状態である「光化学スモッグ」を引き起こし、呼吸器系や循環器系、代謝系、神経系などといった器官に悪影響を及ぼす。そのため環境省では光化学スモッグによる健康被害を抑えるために、大気汚染防止法に基づき光化学オキシダント注意報を定めている。これは大気汚染防止法に基づきOx濃度の１時間値が0.12ppm以上になり、気象条件からみてその状態が継続すると認められる場合に都道府県知事等が発令を行うものである1）。令和5年のOx注意報等の発令状況は、発令都道府県数が17都府県、発令延日数が45日であり、令和4年(12都府県、41日)と比べ増加している1）。また、被害届出状況として令和５年の光化学大気汚染によると思われる被害の届出人数は２人(１県）であり、令和４年(０人）と比較して増加している1）。

注意報の発令は各自治体が常時監視局の実測値などをもとにこれから先のOx濃度を推測し、注意報発令の判断行う。これには専門的知識や経験が必要になるため、的確な判断を行うことは容易ではない。このような背景のもと、AI(人工知能)技術の一種である機械学習(ディープラーニング)を用いて、短期的な高濃度予測を行い判断の補助を行おうという研究がなされている。細越（2022）2）は、Ox高濃度が継続する可能性を事前に予測できれば健康被害の抑制、そして自治体による早期対策の検討を可能にすることに繋がるという考えのもと、Deep Neural Network(DNN)を用いて、常時観測値からリアルタイムで大気汚染(Ox濃度）の動向を予測することの有用性を示した。また現在、細越（2024）では、複数地点のデータを取り込むことによる予測精度向上を目的とした研究が行われている。これは、関東地域特有の海陸風循環を表現する気象条件のデータ、更には化学物質の輸送を表現するための他地点のデータなどを学習させ、主に高濃度域の予測精度を上げるといった試みである。具体的には、選定された代表地点における限定データの学習から、それらの周辺地域まで含めた広範囲における濃度予測を目指している

機械学習において、目的変数（予測対象）に対してどのような説明変数（特徴量）を学習させればよいかを検討することは予測精度に影響するため重要である。細越（2022）の研究において選定させれた特徴量は、オキシダントの生成に関与するNOx、Ox、NMHC、TEMPであった。対象時間から何時間前までのデータを取り込むのかの検討は行われていたが、これらがOxを予測するのに最適な特徴量であるかの検討は行われていない。また、Oxを予測するために最適な特徴慮に関する知見は限られている。このことから、Oxを予測するために最適な特徴量を探索することによって、予測精度の向上が期待できると考えた。自治体等の現場において重要視されるのは、注意報を発令すべきか否かという点である。そのため、高濃度に関して有意予測できることが求められる。そこで、発令の基準である120ppbからバッファを持たせ、本研究では80ppb以上を高濃度と定義した。作成したAIモデルが高濃度を有意に予測できるかをOx生成に関与するNOx、Ox、NMHC、TEMPを基準とし、単一の常時監視局を対象としたベンチマーク試験を実施することで、Ox予測における最適な特徴量の選定を試みることとした。

1. 機械学習とは

2.1　機械学習

近年、情報化が進み多岐の分野にわたり大量のデータを蓄積している。機械学習はそのデータをコンピューターに学習させることによって解析や予測を行わせる。これは、様々な分野で活用されており、物品販売量と顧客の購買情報に関するデータを収集・分析するマーケティングや去の販売データ・購入履歴を基に、自動でその顧客に合ったおすすめ商品やサービスが表示されるレコメンデーション、不良品を検知するための画像認識などと実用化が進んでいる。機械学習を行うにあたっては、目的変数と説明変数が必要になる。目的変数は、解析もしくは予測対象であり、説明変数はこれを用いて解析もしくは予測対象を表現するために用いられるもので、特徴量や独立変数などと呼ばれるものである。これらを与えることによって学習を行うことでモデルを作成する。このようにできたモデルで誰でも予測が簡単に行え、説明が難しい事象に対して特徴を見出して解析もしくは予測を行うことができるのが機械学習の優れている点である。ただ欠点として、コンピューターによって見出される特徴はブラックボックス化されており、そのモデル内での特徴量がどのように影響を及ぼしているかはわかりづらい。そのため、目的変数を適切に表現するための特徴量は目的変数毎に適切なものを選択しないと行けず、そこに定跡はなく基本は手作業で探索することが求められる。

2.2　ニューラルネットワーク

機械学習に用いられるニューラルネットワークは、人間の脳のニューロンの構造や働きをもとに考案されたモデルである。ニューラルネットワークとは、入力された特徴量の値に対して重みを掛け、バイアスを加え、活性化関数に通すことで出力の値を表現する非線形の数式であり、全結合層とも称される。活性化関数は、各層のニューロンから得た値を通すことで出力の値に整える働きをする様々な式の総称である。この活性化関数の働きにより、ニューラルネットワークを用いた様々な非線形的特徴量と出力の関係に対しても近似を行うことが可能となった。ここでの層とは重みとバイアス、活性化関数を用いた処理を1回行う単位のことであり、通常のニューラルネットワークではこの層を重ねて構成している。

ニューラルネットワークにおいて、外部から入力を受け取る層を入力層、外部に値を出力する層を「出力層」といい、前の層の出力を入力として処理を行い次の層の入力に対して値を出力する層を「隠れ層」という。図1におけるy\_predが「出力層」、x\_inputとy\_predの間のものが「隠れ層」のイメージとなる。また、入力層から隠れ層、出力層へ値を前に進め出力を求める処理を伝播という。ニューラルネットワークを用いた学習では、入力に対して得られた出力をより正確にするために重みとバイアスを更新する誤差逆伝播法を用いる。誤差逆伝播法では伝播の処理とは逆向きに、出力の値と正解の値の誤差から重みとバイアスを調整する4）。

ダイアグラム

自動的に生成された説明

図1 ニューラルネットワークのイメージ

2.3　DNN（Deep Neural Network）

本研究で用いるDNNは機械学習の一種であり、ニューラルネットワークの中華層が多層になった、多層ニューラルネットワークを用いた機械学習手法の総称である。従来は人が設計した特徴量を用いて機械学習を行っていたが、ディープラーニングは機械が自ら特徴を設計、抽出し、学習を行う。その結果ディープラーニングを用いた手法は従来手法よりも高い識別率を示している。DNNは機械学習の技術として幅広い分野での利用が進められており、画像認識、音声認識、翻訳等といった場面で使用されている。

3.　研究手法

本研究において、高濃度を有意に予測できるモデルを作成できることが目的である。単一点において、この目的を満たす特徴量の探索を行った。

3.1 対象地点

本実験において先行研究3)との比較を行うため先行研究でテスト地点として選定されていた地点を対象とした。なお、その地点は図2に示した東秩父、鴻巣、幸手、東青梅、所沢、草加、多摩市愛宕、世田谷区、南葛西の計9地点である。

マップ

自動的に生成された説明

図2　対象地点

3.2　使用データ

使用データには、国立環境研究所の測定物質全ての時間値データを用いた。データの前処理として、欠損値に関しては前後の値から線形補完を行い、正規化や標準化を行わずそのままのデータを投入した。また、データテーブルは下記の表1のようになっている。ここでtはある日時のことを指し、目的変数として1~3時間後のOx濃度があり、それに対する特徴量としてある日時tから24時間前までのデータを取り込んだ。

対象期間は以下の通りとし、この期間の予測を行わせた。

学習データ ：2018年　4月1日～3月31日（データ数n:8505）

テストデータ ：2019年　4月1日～3月31日（データ数n:8550）

表1 Ox濃度予測モデルに使用するデータセット構造

テーブル

自動的に生成された説明

また、特徴量においてその地点で測定してるすべてのデータを候補にすること、地点ごとに測定している物質が異なるため対象地点において測定している物質を表2にまとめた。

表2　対象地点における測定物質

テーブル

自動的に生成された説明

3.3　モデル詳細

　学習時のユーザーが設定するハイパラメータを表3にまとめた。

表3　モデルのハイパラメータ

文字と写真のスクリーンショット

自動的に生成された説明

Pythonのバージョンは3.12であり、学習にはFacebook社の提供するオープンソースフレームワークであるPyTorchを用いた。計算実験に使用したハードウェアは以下の通りである。

OS Windows 11 HOME バージョン23H2

CPU AMD Ryzen 7 5700X @4.6GHz 8コア

GPU NVIDIA GeForce RTX 3070

3.4　評価方法

3.4.1　モデル評価

　本実験において優れたモデルは注意報等を発令する補助となるものであり、高濃度を有意に予測することが求められる。そのため、評価方法として時間値データの (i) 予測値と実測とのRMSE(平均二乗平方根誤差)値が低いこと、(ii) 高濃度時の追跡率が高いこと、(iii) 高濃度時のRMSEが低いこと、といった3つの観点をもとに優れたモデルを作成できる最適な特徴量を探索することとした。(i) はモデル全体の精度を確認するものであり、一般的にRMSE値が低いと実測と予測との誤差がすくないと解釈できる。ただ、これだけであると低濃度帯での予測精度はよいが高濃度帯での予測精度は悪いといった際にもRMSE値が低いことがありうる。これだけでは本実験での目的を満たすことはできない。そこで、(ii)～(iii)のような高濃度域に対する評価が必要である。(ii) では、実測値が高濃度を超えた時に対する予測が高濃度を超えている割合であり、そのモデルで高濃度を予測できているかの可否を評価している。そして、(iii) の実測値が高濃度を超えた時のみに絞りRMSEを算出することによって、どれほどの精度で高濃度を予測しているかを評価した。

また、再現率、適合率、調和平均を算出し評価を行った。実測で高濃度を観測した日に予測で高濃度を超えた日の割合である再現率は式に表される。予測で高濃度を観測した日に実測で高濃度を観測した日の割合である適合率は式で表される。これら二つの割合の平均を取るために調和平均(F値)である式(1)で表される。ここにおける調和平均は高濃度域に関してどれほど有意に予測できるかを示した値となる。そのためこの値が一番高いものが目的を満たせるモデル、特徴量だと判断できる。これらの評価をもとに各モデルを評価した。

最終的な評価は調和平均、高濃度RMSE、特徴量の個数の順で優先度を設定し評価を最も優れていたものを抽出する。

モデルを作成する際、同じデータを与えても違う精度のモデルができる。そのため、評価する際はモデルを10回作成しそれぞれの値の平均をとり評価を行った。

3.4.2　特徴量評価

　特徴量の評価を行うためオープンソースであるSHAP（SHapley Additive exPlanations）を用いて特徴量の需要度を計算させ、モデルの評価同様10回算出した平均を用いて特徴量需要度として特徴量の評価を行った。SHAPではシャープレイ値と呼ばれる目的変数に対して、ある特徴量がどれほど貢献しているか（目的変数を説明するために重要か）を表す。これは、平均的な予測値をベースに、ある特徴量が加わったときの予測値の変動量を、その特徴量の限界貢献度として扱い、全ての順序の組み合わせで限界貢献度を求め、その平均をとることで求まる5)。すなわち、値が大きい時予測を行うにあたって需要度が高い。このようにして求めた値を用いて各特徴量がOxの予測に対してどれほどの需要なのかを評価した。この時、高濃度域と低濃度域（高濃度域以外）に対する特徴量需要度を算出させた。

3.5　特徴量の探索

探索を行うにあたって、高濃度域で特徴量需要度が高いものが最適な特徴量になると仮説を立てた。それぞれの地点において測定している物質の時間値データを全て学習させる。これにより、どの物質の何時間前のデータが需要度として高いかを確認する。特徴量需要度は高濃度以上の時と、高濃度以外の時で特徴量需要度を分けて算出した。

ここで算出した特徴量需要度を降順に並び替えた。高濃度以上の上位10個、20個、30個、高濃度以外の上位20個、高濃度以上の上位20個を合わせたものを学習させた。この時の特徴量の有意性を評価させた。

1. 結果

以下に地点ごとの結果を示した。また、下記における「lag = n」はn時間後予測のことを指す。

* 1. 幸手

表4、5に幸手の時間値評価、日別評価を示した。

1時間後予測の時間値データでは、調和平均が上位10個において最も高い76.852%であった。また、高濃度RMSEに関してはベンチマークが最も精度がよく8.654であり、上位10個の高濃度RMSEで8.701であった。モデル全体の評価であるRMSEを見ると、ベンチマークよりも上位10個の精度がよく特徴量の個数も上位10個の方が明らかに少ない。日別評価では上位30個の調和平均が80.000%と一番高く時点で上位10個の78.125%であった。日別評価で見た時上位30個がよく見えるが時間値で見た時の高濃度RMSEが11.159と上位10個と比べた時に精度が悪い。そのため、1時間後予測に有意なのは上位10個だと考えられる。

2時間後予測の時間値データでは、調和平均が上位10個において抜きんでて高い62.745%であった。日別評価に関しても同様に上位10個において抜きんでて高い66.667%であった。このことから、2時間後予測に有意なのは上位10個であると考える。

3時間後予測では、調和平均が上位20個において最も高い47.619%であり、時点で全データの43.210%であった。高濃度RMSEは、全データが19.077で、上位20個が19.369であり、全データの方が精度はよい。ただ、特徴量の個数は、上位20個の方が明らかにデータ数が少ない。また、日別評価に関して上位20個が抜きんでて高い46.154%であった。このことから、3時間後予測に有意なのは上位20個であると考える。

表4　幸手における時間値評価

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 再現率 | 適合率 | 調和平均 | RMSE | 高濃度RMSE |
| ベンチマーク\_lag = 1 | 83.495 | 69.919 | 76.106 | 4.818 | 8.654 |
| 全データ\_lag = 1 | 34.951 | 94.737 | 51.064 | 4.527 | 14.481 |
| 上位10個\_lag = 1 | 80.583 | 73.451 | 76.852 | 4.100 | 8.701 |
| 上位20個\_lag = 1 | 67.961 | 81.395 | 74.074 | 4.381 | 9.331 |
| 上位30個\_lag = 1 | 60.194 | 89.855 | 72.093 | 4.085 | 11.159 |
| 高低上位20個\_lag = 1 | 60.194 | 84.932 | 70.455 | 3.849 | 10.536 |
| ベンチマーク\_lag = 2 | 47.573 | 63.636 | 54.444 | 6.966 | 14.153 |
| 全データ\_lag = 2 | 28.155 | 70.732 | 40.278 | 6.912 | 18.899 |
| 上位10個\_lag = 2 | 62.136 | 63.366 | 62.745 | 6.715 | 13.273 |
| 上位20個\_lag = 2 | 37.864 | 75.000 | 50.323 | 6.661 | 17.040 |
| 上位30個\_lag = 2 | 37.864 | 76.471 | 50.649 | 6.695 | 16.631 |
| 高低上位20個\_lag = 2 | 20.388 | 80.769 | 32.558 | 7.586 | 21.279 |
| ベンチマーク\_lag = 3 | 25.243 | 63.415 | 36.111 | 9.098 | 21.029 |
| 全データ\_lag = 3 | 33.981 | 59.322 | 43.210 | 8.708 | 19.077 |
| 上位10個\_lag = 3 | 24.272 | 65.789 | 35.461 | 8.716 | 23.174 |
| 上位20個\_lag = 3 | 38.835 | 61.538 | 47.619 | 9.093 | 19.369 |
| 上位30個\_lag = 3 | 31.068 | 69.565 | 42.953 | 8.707 | 20.338 |
| 高低上位20個\_lag = 3 | 27.184 | 70.000 | 39.161 | 8.528 | 21.118 |

表5　幸手における日別評価

|  |  |  |  |
| --- | --- | --- | --- |
|  | 再現率 | 適合率 | 調和平均 |
| ベンチマーク\_lag = 1 | 80.645 | 73.529 | 76.923 |
| 全データ\_lag = 1 | 41.935 | 86.667 | 56.522 |
| 上位10個\_lag = 1 | 80.645 | 75.758 | 78.125 |
| 上位20個\_lag = 1 | 70.968 | 84.615 | 77.193 |
| 上位30個\_lag = 1 | 70.968 | 91.667 | 80.000 |
| 高低上位20個\_lag = 1 | 61.290 | 79.167 | 69.091 |
| ベンチマーク\_lag = 2 | 48.387 | 60.000 | 53.571 |
| 全データ\_lag = 2 | 22.581 | 63.636 | 33.333 |
| 上位10個\_lag = 2 | 67.742 | 65.625 | 66.667 |
| 上位20個\_lag = 2 | 41.935 | 65.000 | 50.980 |
| 上位30個\_lag = 2 | 38.710 | 80.000 | 52.174 |
| 高低上位20個\_lag = 2 | 12.903 | 80.000 | 22.222 |
| ベンチマーク\_lag = 3 | 19.355 | 66.667 | 30.000 |
| 全データ\_lag = 3 | 29.032 | 56.250 | 38.298 |
| 上位10個\_lag = 3 | 19.355 | 60.000 | 29.268 |
| 上位20個\_lag = 3 | 38.710 | 57.143 | 46.154 |
| 上位30個\_lag = 3 | 29.032 | 64.286 | 40.000 |
| 高低上位20個\_lag = 3 | 22.581 | 63.636 | 33.333 |

* 1. 鴻巣

表6、7に鴻巣の時間値評価、日別評価を示した。

1時間後予測の時間値データでは、調和平均が上位20個において最も高い76.852%であった。モデル全体の評価であるRMSEを見るとベンチマークよりも精度が悪い。ただ、高濃度RMSEに関して、最も精度がよく8.995であった。日別評価で見た時、調和平均は上位20個が最も高い85.714%であった。このことから1時間後予測に有意なのは上位20個だと考えられる。

2時間後予測の時間値データでは、調和平均が上位10個において抜きんでて高く66.055%であった。高濃度RMSE、RMSEどちらにおいても上位10個が最も精度がよかった。日別評価で見ても、上位10個の調和平均が抜きんでて高い。このことから2時間後予測に有意なのは上位10個だと考えられる。

3時間後予測では、調和平均が上位10個において最も高い56.652%であり、時点で上位20個の56.502%であった。高濃度RMSEだと上位20個が19.062、上位10個は20.845であり、上位20個の方が予測精度がよい。RMSEに関しても上位20個の方が精度

表6　鴻巣における時間値評価

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 再現率 | 適合率 | 調和平均 | RMSE | 高濃度RMSE |
| ベンチマーク\_lag = 1 | 70.635 | 83.178 | 76.395 | 4.125 | 9.955 |
| 全データ\_lag = 1 | 58.730 | 88.095 | 70.476 | 4.429 | 11.559 |
| 上位10個\_lag = 1 | 50.000 | 90.000 | 64.286 | 4.880 | 12.916 |
| 上位20個\_lag = 1 | 87.302 | 74.324 | 80.292 | 4.289 | 8.995 |
| 上位30個\_lag = 1 | 76.984 | 80.833 | 78.862 | 4.164 | 9.050 |
| 高低上位20個\_lag = 1 | 69.048 | 89.691 | 78.027 | 3.817 | 9.941 |
| ベンチマーク\_lag = 2 | 42.857 | 75.000 | 54.545 | 6.925 | 15.994 |
| 全データ\_lag = 2 | 48.413 | 68.539 | 56.744 | 6.694 | 15.686 |
| 上位10個\_lag = 2 | 57.143 | 78.261 | 66.055 | 6.350 | 14.576 |
| 上位20個\_lag = 2 | 43.651 | 85.938 | 57.895 | 6.696 | 16.245 |
| 上位30個\_lag = 2 | 26.984 | 85.000 | 40.964 | 6.999 | 19.214 |
| 高低上位20個\_lag = 2 | 30.952 | 88.636 | 45.882 | 6.841 | 19.894 |
| ベンチマーク\_lag = 3 | 40.476 | 69.863 | 51.256 | 8.885 | 19.222 |
| 全データ\_lag = 3 | 32.540 | 68.333 | 44.086 | 8.576 | 21.492 |
| 上位10個\_lag = 3 | 52.381 | 61.682 | 56.652 | 8.808 | 20.845 |
| 上位20個\_lag = 3 | 50.000 | 64.948 | 56.502 | 8.637 | 19.062 |
| 上位30個\_lag = 3 | 26.190 | 89.189 | 40.491 | 9.071 | 23.733 |
| 高低上位20個\_lag = 3 | 50.794 | 60.377 | 55.172 | 8.893 | 17.052 |

表7　鴻巣における日別評価

|  |  |  |  |
| --- | --- | --- | --- |
|  | 再現率 | 適合率 | 調和平均 |
| ベンチマーク\_鴻巣\_lag = 1 | 75.758 | 83.333 | 79.365 |
| 全データ\_鴻巣\_lag = 1 | 66.667 | 88.000 | 75.862 |
| 上位10個\_鴻巣\_lag = 1 | 57.576 | 86.364 | 69.091 |
| 上位20個\_鴻巣\_lag = 1 | 90.909 | 81.081 | 85.714 |
| 上位30個\_鴻巣\_lag = 1 | 78.788 | 83.871 | 81.250 |
| 高低上位20個\_鴻巣\_lag = 1 | 69.697 | 92.000 | 79.310 |
| ベンチマーク\_鴻巣\_lag = 2 | 54.545 | 78.261 | 64.286 |
| 全データ\_鴻巣\_lag = 2 | 45.455 | 62.500 | 52.632 |
| 上位10個\_鴻巣\_lag = 2 | 69.697 | 82.143 | 75.410 |
| 上位20個\_鴻巣\_lag = 2 | 48.485 | 88.889 | 62.745 |
| 上位30個\_鴻巣\_lag = 2 | 30.303 | 71.429 | 42.553 |
| 高低上位20個\_鴻巣\_lag = 2 | 30.303 | 83.333 | 44.444 |
| ベンチマーク\_鴻巣\_lag = 3 | 42.424 | 66.667 | 51.852 |
| 全データ\_鴻巣\_lag = 3 | 33.333 | 68.750 | 44.898 |
| 上位10個\_鴻巣\_lag = 3 | 63.636 | 67.742 | 65.625 |
| 上位20個\_鴻巣\_lag = 3 | 48.485 | 57.143 | 52.459 |
| 上位30個\_鴻巣\_lag = 3 | 21.212 | 77.778 | 33.333 |
| 高低上位20個\_鴻巣\_lag = 3 | 54.545 | 56.250 | 55.385 |

* 1. 東秩父
  2. 高濃度域の上位20個による学習
  3. 高濃度域の上位30個による学習
  4. 高濃度域と低濃度域それぞれの上位20個による学習

5.　考察

6.　まとめ

7.　謝辞

8. 参考文献

1)・・・　環境省（令和5年）: “令和５年光化学大気汚染の概要－注意報等発令状況、被害届出状況－”, 環境省

2)・・・　細越英彰(2022) : “ディープラーニングを用いた光化学オキシダント濃度の短期予測”, 明星大学　理工学部　総合理工学科　環境科学系, 令和　4年度卒業論文

3)・・・　細越英彰(2024) : 細越さんの修論掲載予定

4)・・・　星野 智恵子,吉野 裕紀 : 計測データの欠損を考慮した重機の経路最適化

5)・・・　川越雄介(2021/4/14) : SHAP を用いて機械学習モデルを説明する　Data Robot