

調査研究課題名：地方公共団体におけるシミュレーションモデルを活用した
光化学オキシダント対策の検討に関する調査研究（課題 2）

委託業務名：地方公共団体におけるシミュレーションモデルを活用した
光化学オキシダント対策の検討に関する調査研究

研究代表者：山村 由貴

所 属 機 関：福岡県保健環境研究所

【研究概要】

本研究では、福岡県を対象とした光化学オキシダント濃度予報システムを構築した。この予報システムでは、汎用性のあるモデルによる大気汚染シミュレーションを用い、数日先までの福岡県内のオキシダント濃度を地域別に予測する。加えて、排出量を変化させることにより、越境汚染と地域汚染の寄与判別や県内 NOx、VOC の排出量削減効果も同時に予測が可能である。

さらにこの予測の精度向上のため、機械学習を用いた誤差補正機構も開発した。誤差補正には、オキシダント生成に関係すると考えられる要素である気象、地理、大気汚染物質等の計算値を説明変数、オキシダント濃度の観測値と計算値の差を目的変数とした機械学習モデル、すなわち、説明変数から、発生し得る誤差濃度を予測する機械学習モデルを作成した。過去 3 年分の計算値・観測値を学習データとして機械学習を行った後、別の 1 年分をテストデータに設定してこの機械学習モデルを検証した結果、誤差補正による予測精度の向上が確認できた。また、健康被害低減の観点から、高濃度の予測誤差を最小化するため、オキシダント濃度の観測値が 80ppb 以上でかつ計算値が観測値よりも 3 割以上小さいデータに重みをかけて学習させることで、高濃度の見逃し低減に特化した学習モデルも作成可能であることが分かった。予報システムの予測結果から説明変数を抽出して学習モデルに読み込ませ、誤差を予測させて補正するという一連の作業を自動化し、誤差補正機構として予報システムに搭載することで、より精度の高い予報が可能となった。

予報システムは既に県での運用を開始しており、予測結果は福岡県のほか、北九州市、福岡市、大牟田市、久留米市の気象担当職員とも共有し、オキシダント高濃度時における行政対応に活用している。また、予報システムで高濃度が予測された際は、県民公式ラインで情報を配信しており、健康被害の低減に貢献している。

1 研究従事者（○印は研究代表者）

○氏名（所属）： 山村 由貴（福岡県保健環境研究所環境科学部大気課）

氏名（所属）： 山本 重一（福岡県保健環境研究所環境科学部大気課）

氏名（所属）： 熊谷 博史（福岡県保健環境研究所管理部企画情報管理課）

氏名（所属）： 城山 宗一郎（福岡県環境部環境保全課大気係）

氏名（所属）： 丸林 啓太（福岡県環境部環境保全課大気係）

氏名（所属）： 菅田 誠治（国立環境研究所）

2 研究目的

本研究の目的は、「越境汚染の寄与割合と排出量削減要請効果の試算を含む、光化学オキシダ

ントの濃度予報と自動配信が可能なシステム（以下、予報システム）を構築し、県民の健康被害を低減すること」であり、そのために「機械学習による誤差補正機能をシステム内に組み込み、予報精度を高めること」である。

3 研究対象及び方法

3.1 予報システムの構築

福岡県内の光化学オキシダントを対象とし、気象モデル WRF と大気質モデル CMAQ を用いて、3 日先までの濃度予測計算を自動で行うシステムを構築した。詳細な設定等は、3.1.1 に示す通りである。

3.1.1 計算条件

計算に使用したモデル及び計算設定を表 3.1、排出量データを表 3.2 に示す。排出量データのうち、REAS version 2 は 2008 年の排出量が最終年であることから、近年の濃度変化を反映させるため、先行研究に基づいて中国からの SO₂ 排出量を 60%、NO_x 排出量を 40%、韓国からの SO₂ 排出量を 40%削減して使用した。なお、今後、排出量データの更新に対応できるよう、予報システムのプログラムは、使用する排出量データを容易に変更可能な設計にしている。

計算領域を図 3.1 に示す。水平解像度及び計算領域は、国立環境研究所が開発した大気汚染予報システム（VENUS）を参考に、福岡県を高解像度で計算すること、越境汚染の影響評価に有効な五島・壱岐・対馬を D03 に入れること、全ての計算が翌朝 9 時までに完了できること等を考慮して設定した。

計算は、1 回の予測において、全排出量を含む通常計算（以下、通常計算）、国外人為起源排出量を含まない計算（以下、国外ゼロ計算）、福岡県内の人為起源 NO_x、VOC のみを 20%削減した計算（以下、排出量削減効果の試算）の 3 ケースを行うこととした。通常計算、国外ゼロ計算は D01～D03、排出量削減効果の試算は D03 のみ実行する。

国外ゼロ計算は、越境汚染の寄与濃度（以下、越境寄与濃度）を概算する（越境寄与濃度＝通常計算－国外ゼロ計算）ため実行するものである。また、排出量削減効果の試算の削減量 20% は、福岡県の「光化学オキシダント（光化学スモッグ）に係る緊急時対策基本要綱」に基づいて設定した。なお、排出量削減は当日要請を受けてから実施されるため、排出量削減効果の試算について、初期条件（計算前日の濃度）は通常計算の結果を使用し、計算期間（当日～3 日先）のみ、NO_x、VOC を削減した排出量データを用いて計算する設定とした。

表 3.1 計算設定

モデル	WRF Version 3.9.1.1（気象） CMAQ Version 5.0.2（大気質）
水平解像度	45 km×45 km (D01:東アジア) 15 km×15 km (D02:西日本) 5 km ×5 km (D03:北部九州)
鉛直層数	39 (~100 hPa)
地理データ	USGS GTOPO30 (Altitude) USGS 25 Category (Land Use Category)
気象データ	NCEP GFS (0.25×0.25°)

表 3.2 排出量データ

東アジア人為起源(日本を除く)	Regional Emission inventory in Asia (REAS) version 2
国内人為起源 (自動車・船舶を除く)	EAGrid2010-Japan
国内船舶	OPRF's Ship Emission inventory
国内自動車	JATOP Emission Inventory -Data Base 2011 Automobile Source (JEI-DB2011-AS)
植物起源 VOC	Model of Emissions of Gases and Aerosols from Nature (MEGAN) version2.04
火山	気象庁火山活動報告資料を参考に、九州に与える影響の大きい阿蘇・桜島・諏訪之瀬島・西之島の SO ₂ 噴出量を考慮

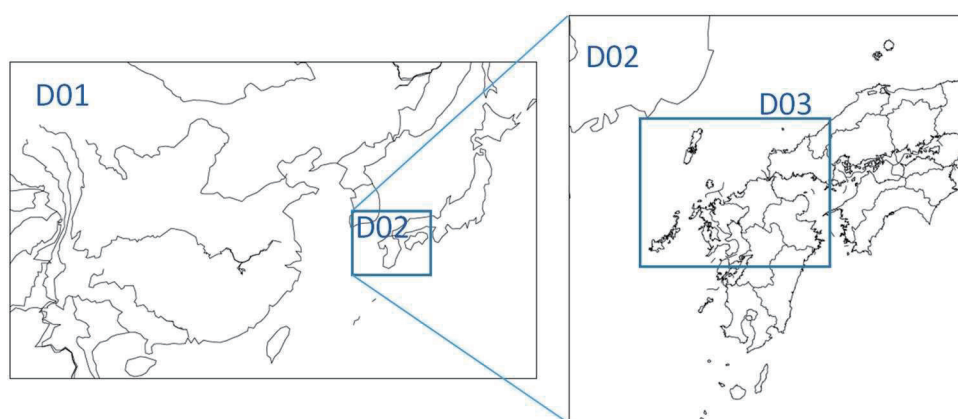


図 3.1 計算領域

3. 1. 2. 計算結果の共有

計算の結果は、オキシダントを含む数十種類の大気汚染物質の 1 時間毎の濃度情報を含んだファイルとして出力される。本システムでは、このファイルを図 3.2 に示すようなオキシダント濃度のコンターマップに加工し、pdf ファイルとして保健環境研究所と県庁担当部署（環境保全課）が共有可能なフォルダへ保存する。

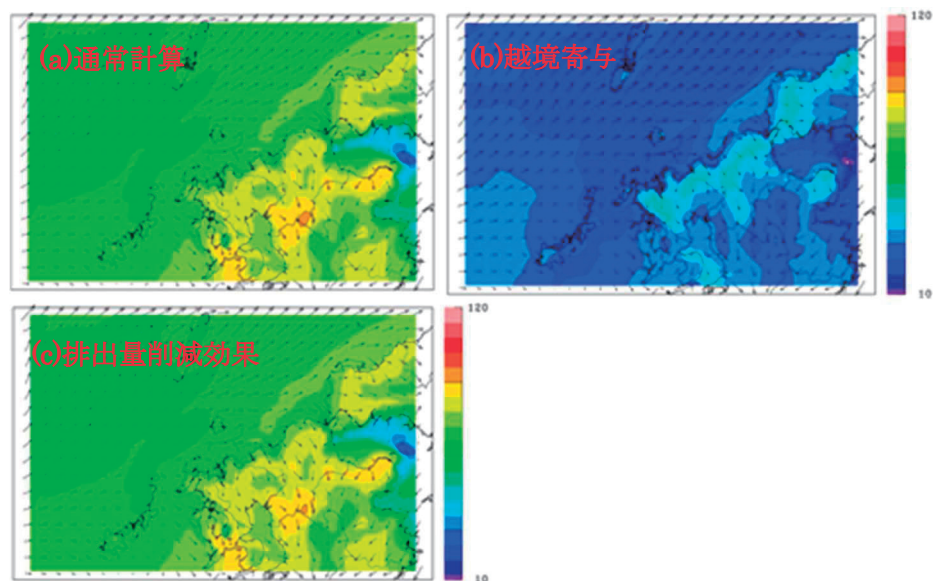


図 3.2 コンターマップ（例）

(a)通常計算 (b)越境寄与濃度 (c) 排出量削減効果の試算

通常計算による濃度予測結果については、福岡県のオキシダント注意報発令区域及び保健福祉環境事務所の管轄エリアを考慮し、北九州・福岡・筑豊・筑後の4地域に分け（図 3.3(a)）、0～6、6～12、12～18、18～24 時の4つの時間区分に対して、各地域の最高濃度（1 時間値）を算出する。この計算結果はエクセルファイルとして出力すると共に、担当職員のアドレスへ自動で送信される（図 3.3(b)）。なお、図 3.3(b)中の”地域名_Warn”の列は、注意報発令基準となる濃度 120ppb を超えると予測された場合に○が記載される。

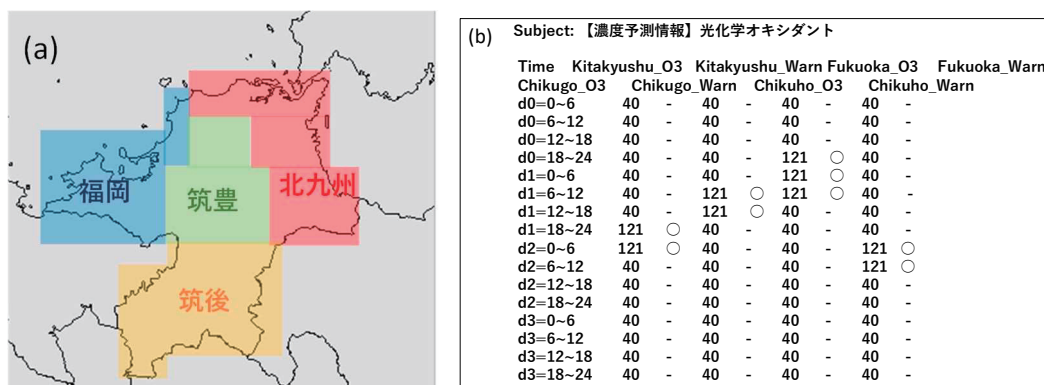


図 3.3 (a)福岡県内の地域区分概要 (b)メール送信内容（例）

3. 1. 3. 計算の流れ

主な計算の流れは以下のとおりである。なお、以下の過程はシステム上で全て自動化されており、予測日当日の朝に完了し、配信するよう毎日自動で実行されるスケジュールとなっている。

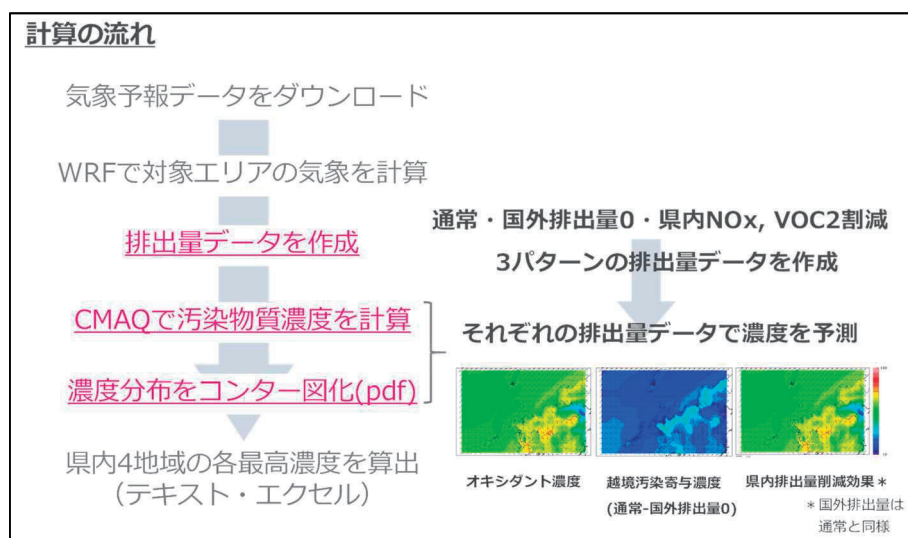


図 3.4 計算の流れ

3. 2. 誤差補正

本研究では、オキシダント濃度に関係すると考えられる要素を説明変数、オキシダント濃度計算値の誤差を目的変数とし、過去のデータを用いた機械学習によって、オキシダント濃度計算値の誤差を予測する学習モデルを作成した。作成した学習モデルの予報システムへの組み込みについては、5.4 で述べる。

3. 2. 1. 学習データ

学習データには、大気モデルによる計算値及び福岡県内でオキシダントを測定している大気汚染常時監視局（2018、2019 年は 45 局、2021 年以降は 43 局）のオキシダント濃度（以下、観測値）を用いた。データの期間は 2018、2019、2021、2022 年、時間分解能は 1 時間である。なお、2020 年は新型コロナウイルスの影響に伴う活動自粛等によって、排出量が例年と異なる挙動をしていた可能性があるため除外した。計算値は、各格子の緯度経度情報から、各大気汚染常時監視局に最も近い格子の値を用いた。また、計算値にオキシダントは無いため、オゾンをおキシダントと等価と仮定している。

モデル構築に用いる学習データ及びモデル評価に使用するテストデータは、学習データの期間が補正精度に及ぼす影響を検証するため、以下の 2 ケースを設定した。ケース 1 は 2021 年 1 年間のデータを Hold out 法で 7:3 に分割し、7 割を学習データ、3 割をテストデータとしている。ケース 2 は 3 年間を学習データ、別の 1 年間をテストデータとしている。なお、データ数=（測定局数）×（測定時間数）である。この 2 ケースの比較結果は、3.1 で述べる。

ケース 1

学習データ : 2021 年 1 月 1 日～12 月 31 日の 1 年分のデータのうち 7 割
(データ数 260526)

テストデータ 1 : 学習データで使用していない残りの 3 割
(データ数 111653)

テストデータ 2 : 2022 年 1 月 1 日～12 月 31 日
(データ数 362929)

ケース 2

学習データ : 2018 年 1 月 1 日～12 月 31 日
 2019 年 1 月 1 日～12 月 31 日
 2021 年 1 月 1 日～12 月 31 日
 (データ数 (3 年合計 1191360))
 テストデータ : 2022 年 1 月 1 日～12 月 31 日
 (データ数 362929)

説明変数は、計算値に含まれ、オキシダントの生成に関係すると考えられる 16 要素 (表 3.3) とし、目的変数はオキシダント濃度の誤差 (計算値 - 観測値[ppb]) とした。月と時刻は、そのまま説明変数に用いた場合、数字の大きさの違いのみが考慮され、1 月と 12 月、時刻の 1 時と 24 時は、離れた関係となる。しかし、実際は 12 月と 1 月は共に冬季、1 時、24 時は共に夜間であり、類似した点が多いと考えられる。そこで、三角関数 (sin, cos) を使用して周期関数に変換し、円状に配置することで周期性を反映させた。時刻については、式(3.1)、(3.2)で変換し、月については、式(3.3)、(3.4)に示すとおり、月に日付を組込んで周期化を行った。周期化の結果を図 3.5 に示す。

$$\text{mod_time_x} = \sin(t / t_max \times 2\pi) \quad \dots \text{式(3.1)}$$

$$\text{mod_time_y} = \cos(t / t_max \times 2\pi) \quad \dots \text{式(3.2)}$$

t : 該当時刻、t_max : 時刻の最大数

$$\text{mod_month_x} = \sin(m / m_max + d/d_max) \times 2\pi \quad \dots \text{式(3.3)}$$

$$\text{mod_month_y} = \cos(m / m_max + d/d_max) \times 2\pi \quad \dots \text{式(3.4)}$$

m : 該当月、d : 該当日、m_max : 月の最大値、d_max : ひと月の日数

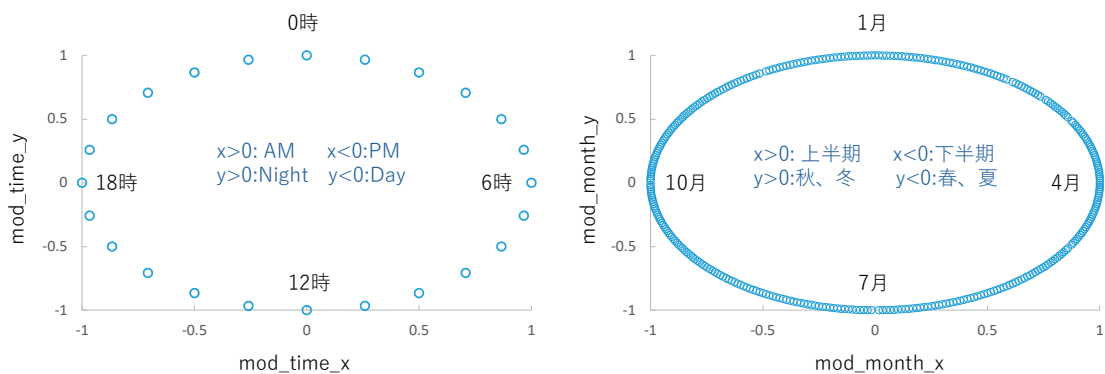


図 3.5 周期関数による変換

大気汚染物質については、オキシダントの生成に関係する NO、NO₂、VOC を説明変数とした。なお、VOC には、人為起源、植物起源の中で最も排出量が多く、かつオゾン生成能の高いトルエンとイソプレンを選択した。

表 3.3 説明変数

属性	変数
気象	気温、気圧、風向、風速、降水量、水蒸気量、日射量
月・時刻・位置	月、時刻、緯度、経度
大気汚染物質	O ₃ 、NO、NO ₂ 、トルエン、イソブレン

3. 2. 2. 機械学習モデル

機械学習のアルゴリズムは、アンサンブル回帰とニューラルネットワーク回帰とし、2つのモデルの精度を比較した。また、線形回帰モデルも作成し、機械学習モデルとの精度比較を行った。モデルの構築には、MathWorks 社の MATLAB 2023b の Statistics and Machine Learning Toolbox の関数、fitrensemble、fitnet、fitlm を用いた。

fitrensemble はアンサンブル回帰学習アルゴリズムである。本研究では、アンサンブル学習サイクル数 297、最適化回数 30 とし、MATLAB のオプション‘OptimizeHyperparameters’を使用して、ベイズ最適化によって、5 分割交差検証損失を最小化するハイパーパラメーターを求めた。得られた主なハイパーパラメーターを表 3.4 に示す。このハイパーパラメーターでの 5 分割交差検証損失は、3.9 であった。

表 3.4 アンサンブル回帰のハイパーパラメーター

アンサンブル集約法	バギング
決定木の分割条件	平均二乗誤差
木の数	297
決定木を構成する最大特徴数	6
決定木の最大深さ	学習データ数-1
ノードの最小サンプル数	1

fitnet はフィードフォワードニューラルネットワーク（以下、FNN）回帰学習アルゴリズムである。FNN は入力層、1 つ以上の隠れ層、出力層から構成され、情報は入力から出力までの一方向にのみ移動する。本研究では、fitrensemble と同様に、オプション‘OptimizeHyperparameters’を使用してハイパーパラメーターの最適化を行った。得られた主なハイパーパラメーターを表 3.5 に示す。このハイパーパラメーターでの 5 分割交差検証損失は、4.6 であった。

表 3.5 FNN 回帰のハイパーパラメーター

全結合層の活性化関数	双曲線正接関数
隠れ層の層数	3
各隠れ層のノード数	297 146 11
正則化	0.076

fitlm は、入力したデータに対して、説明変数から構成される線形式（以下、線形回帰モデル）を作成する関数である。各モデルの比較結果は、5.2 で述べる。

3. 2. 3. 重みを用いた学習

MathWorks 社の MATLAB 2023b の Deep Learning Toolbox では、ニューラルネットワークの学習時に、任意の条件を持つデータに重みを付けて学習させることが可能である。本研究ではこの機能を、高濃度の見逃しを減らすことを目的として用いた。

学習アルゴリズムには、浅層の多層ニューラルネットワーク（以下、SMNN）回帰である train を用いた。Deep Learning Toolbox には‘OptimizeHyperparameters’オプションが無いため、ハイパーパラメーターは手動で設定する必要がある。エポック数 4000、隠れ層・出力層の伝達関数は MATLAB のデフォルト値（隠れ層の伝達計数：双曲線正接シグモイド伝達関数、出力層の伝達関数：線形伝達関数）とし、隠れ層数とノード数を表 3.6 に示す 2 ケースで検討した。重みは、オキシダント濃度の観測値が 80ppb 以上でかつ計算値が観測値よりも 3 割以上小さいデータ、すなわち計算値が大きく過小評価となったデータに対して設定した。重みの検証は、前述のケースで最も精度の高かった隠れ層数とノード数設定で行った。重みは 10 と 50 の 2 ケースを設定し、重みなしの場合と比較した。これら重みの比較結果は、3.3 で述べる。

表 3.6 SMNN 回帰のノード数と隠れ層数の検証条件

	ケース 1	ケース 2
隠れ層の層数	1	1、2、3 層で比較
各隠れ層のノード数	100 と 300 で比較	100

4 研究目標（計画）

【令和 4 年度】

令和 4 年度は、過去データを使って、機械学習による計算誤差の補正手法を確立する。本研究で購入するワークステーションに大気モデルとともに機械学習用ソフトを導入し、計算環境を整備する。計算領域や解像度等の計算条件は、情報公開を担当する県の環境行政担当部署と協議し、設定する。その条件で過去数年間分の時間値濃度を計算し、計算値と観測値を比較する。比較の結果、有意な誤差が起こる条件の法則性を見出し、計算値と観測値の誤差が最小となるための補正式を算出する。補正式の算出・適用は、NO_x、VOC 排出量や越境汚染を削減した計算に対しても行う。削減計算に適する補正式は、過去の様々な濃度や気象条件等におけるデータ群を解析・検証し、算出する。なお、誤差の法則性や補正式算出過程は、条件が異なる他県でも適用できるように、編集可能なプログラムとする。

【令和 5 年度】

予報専用のワークステーションを導入し、令和 4 年度に作成した誤差のパターンを使用して、予報結果に対し、機械学習を用いて誤差の発生要因を自動判断し、対応する補正式を適用する機構をシステムに組み込む。さらに、大気モデルによる予報と配信を自動で行うシステムを構築する。補正後の予報結果及び NO_x・VOC 排出量削減パターンの計算結果については、県行政担当課と情報共有し、研究者、行政担当者の両方で活用を検討する。なお、構築したシステムや得られた成果について、国立環境研と地方環境研とのⅡ型共同研究「光化学オキシダント等の変動要因解析を通じた地域大気汚染対策提言の試み」（2023-2025 年度実施予定）を通して、他の自治体と情報共有する。

5 研究成果

【令和4年度】

令和4年度は「光化学オキシダントの濃度予報と自動配信が可能なシステム」の構築を先行して行った。この予報システムにより、福岡県内の地域ごとのオキシダント高濃度を事前に予測することが可能になった。予報システムの運用については、行政担当者と協議し、休日の監視体制（職員配備）の判断基準として使用している。オキシダント濃度が120ppbを超えると予測された場合は、該当地域の県民に対して福岡県 LINE で高濃度情報を配信することとした。また、福岡県では注意報の発令時に、光化学オキシダント濃度上昇の主要因が大陸からの大気汚染物質の移流と判断される場合には、主要ばい煙及び VOC 排出者への排出削減要請を行わないこととしている。この判断の資料の1つとして、予報システムから得られる越境寄与を活用することとした。

CMAQ では、オキシダントだけでなく、 SO_2 や NO_2 、 SO_4^{2-} 粒子等、様々な化学物質の濃度も同時に計算・出力される。①2022年12月に福岡市及びその周辺で発生した異臭発生事例では SO_2 、②2023年1月に発生した高濃度PM2.5事例ではPM2.5（微小粒子状 SO_4^{2-} 、 NO_3^- 、 NH_4^+ 、EC、OCの和をPM2.5として概算した）の予測結果を解析し、濃度予測や要因解明を行った。詳細を以下に示す。なお、得られた解析結果は、知事定例記者会見の説明資料として提供し、国立環境研と地方環境研とのⅡ型共同研究「光化学オキシダント等の変動要因解析を通じた地域大気汚染対策提言の試み」（2023-2025年度実施）メーリングリストを通じて、他自治体と共有した。

①2022年12月21日午前7時頃、福岡市内で「硫黄のような臭いがする」との通報が相次いだ。この時、福岡市内にある測定局の SO_2 時間値濃度は、最大で0.023ppmまで上昇していた。 SO_2 濃度が上昇する要因としては、もっとも典型的な例として火山が考えられる。ただし、 SO_2 の刺激臭が感知されるのは0.3-1.4ppmであることから、臭いの原因は SO_2 ではなく、 SO_2 と同様に火山から噴出された H_2S の可能性が高いと考えられた。CMAQには H_2S 等が含まれていないため、 SO_2 と H_2S 等が類似する挙動をしたと仮定し、 SO_2 の発生源解析を行った。CMAQによる SO_2 計算結果の水平分布図を図5.1(a1)(a2)(a3)、桜島付近と福岡市付近を含む同一経度（図5.1(a1)の赤線）に沿った断面図を図5.1(b1)(b2)(b3)に示す。図5.1(b1)(b2)(b3)から、桜島から噴出した SO_2 が北上しながら高度を下げ、異臭通報が多く寄せられた21日午前に、福岡市付近で地上に到達していた可能性が示唆された。

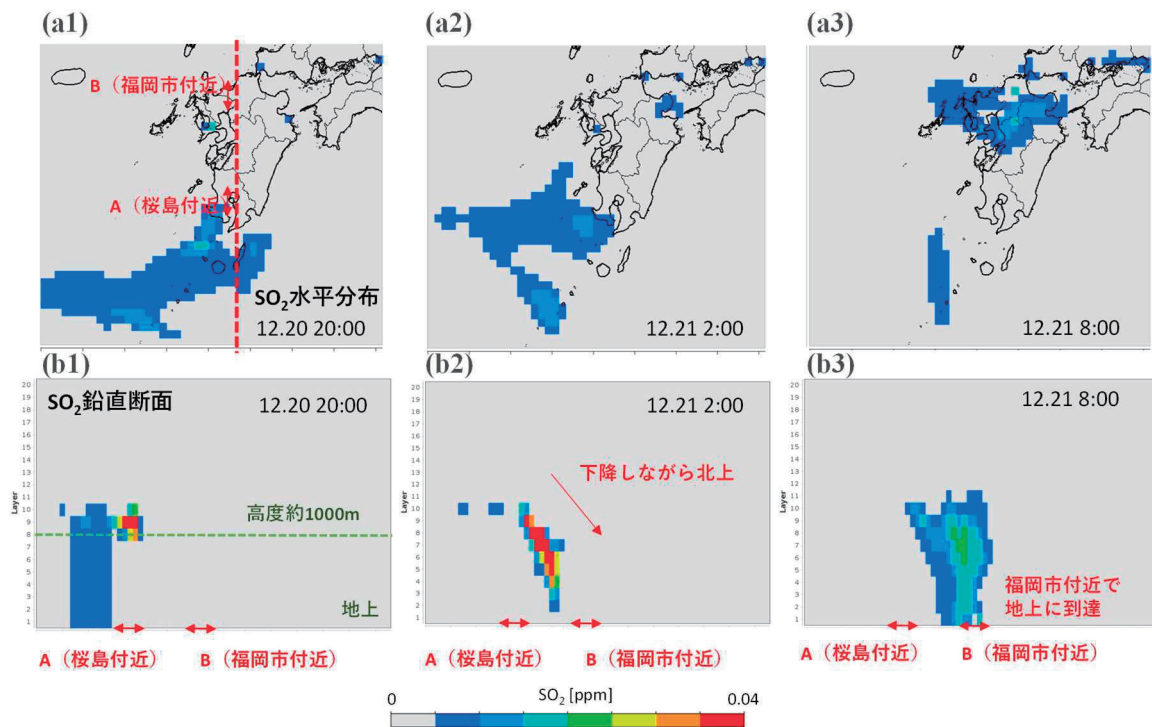


図 5.1 SO₂ 濃度分布
(a1)(a2)(a3)水平分布 (b1)(b2)(b3)鉛直断面

②2023 年 1 月 5 日～7 日にかけて、九州・山口各県で $50\mu\text{g}/\text{m}^3$ を超える高濃度の PM_{2.5} が観測され、福岡県内では霧のように霞んだ状態となっていた。この高濃度の PM_{2.5} を事前に予測し、オキシダントの場合と同様に、福岡県 LINE で高濃度情報を配信した。この時の予測結果を図 5.2(a)に示す。さらに、越境寄与濃度の予測結果(図 5.2(b)) から、高濃度の主要要因は越境汚染であったことが示唆された。

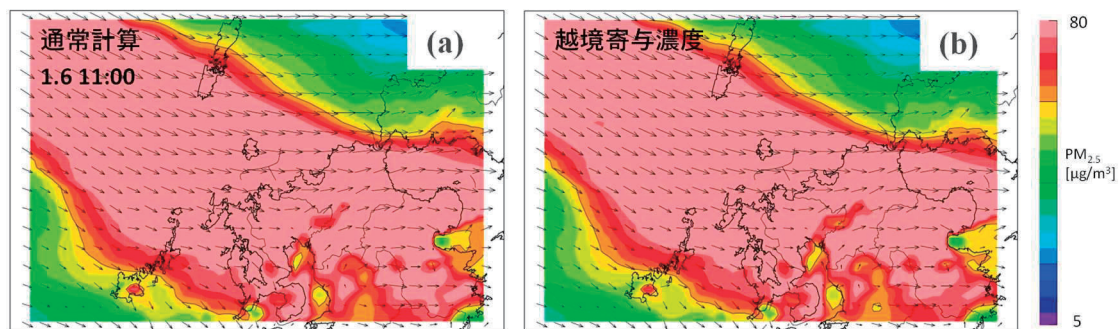


図 5.2 PM_{2.5} 濃度予測結果 (1 月 6 日 11 時)
(a)通常計算 (b)越境寄与濃度

【令和 5 年度】

令和 5 年度は、令和 4 年度に作成したシステムに追加する「機械学習による誤差補正機能」構築のため、誤差補正を行う機械学習モデル及び誤差補正を自動で行うプログラムの作成を行った。なお、計画では誤差をパターン別に分けて補正を行うモデル(分類モデル)を想定していたが、予報データの説明変数から直接誤差の値を予測するモデル(回帰モデル)で十分な精度を得られることが分かったことから、回帰モデルに変更した。

5. 1. 学習データの検証結果

3.2.1 で示したケース 1、2 について、FNN 回帰学習アルゴリズム `fitrnet` で検証を行った。

図 5.3(a)は、ケース 1 のテストデータ 1 のオキシダント濃度を、観測値と計算値で比較したものである。このテストデータ 1 に対して、ケース 1 の学習データで学習したアンサンブル回帰モデル（以下、ケース 1 学習モデル）を用いて誤差補正した結果を、図 5.3(b)に示す。誤差補正は、説明変数をケース 1 学習モデルに入力し、得られた目的変数 [オキシダント濃度の誤差（計算値 - 観測値[ppb]）] を、もとの計算値から引くことで行った。それぞれの二乗平方根誤差 RMSE、正規化平均バイアス NMB、ピアソン相関係数 PCC を表 5.1 に示す。RMSE、NMB、PCC は、式(5.1)~(5.3)に示すとおりである。補正前が RMSE19.8、NMB-0.23、PCC0.22 であるのに対し、補正後は RMSE6.8、NMB-0.23、PCC0.91 になっており、テストデータ 1 に対しては、大きく精度が向上していることが分かる。

一方、ケース 1 のテストデータ 2 のオキシダント濃度を、観測値と計算値で比較した結果を図 5.3(c)に示す。また、テストデータ 2 に対して、ケース 1 学習モデルを用いて誤差補正した結果を図 5.3(d)に示す。それぞれの RMSE、NMB、PCC を表 5.1 に示す。図 5.3(c)(d)は、図 5.3(a)(b)と異なり、補正前は RMSE16.6、NMB-0.18、PCC0.53 であるのに対し、補正後は RMSE20.8、NMB-0.20、PCC0.12 であり、補正によって精度が下がっていることが分かる。これは、ケース 1 学習モデルが 2021 年のデータに対して過学習を起こし、2022 年のデータには適合できなかったためと考えられる。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Mi - Oi)^2} \quad \dots \text{式(5.1)}$$

$$NMB = \sum_{i=1}^N (Mi - Oi) / \sum_{i=1}^N Oi \quad \dots \text{式(5.2)}$$

$$PCC = \sum_{i=1}^N [(Mi - \overline{Mi})(Oi - \overline{Oi})] / \sqrt{\sum_{i=1}^N (Mi - \overline{Mi})^2 \sum_{i=1}^N (Oi - \overline{Oi})^2} \quad \dots \text{式(5.3)}$$

Mi : 計算値、 Oi : 観測値、 $\overline{Mi}$: 平均値、 N : データ数

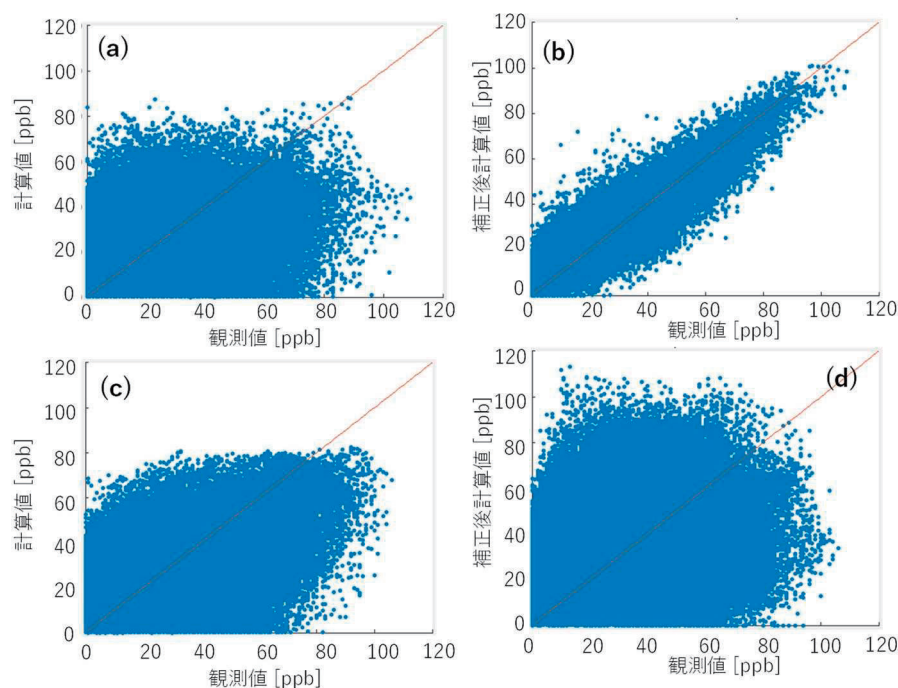


図 5.3 ケース 1 の誤差補正精度

((a)補正前テストデータ 1、(b)テストデータ 1 をケース 1 学習モデルで補正、
(c)補正前テストデータ 2、(d)テストデータ 2 をケース 2 学習モデルで補正)

表 5.1 ケース 1 の RMSE、NMB、PCC

	RMSE	NMB	PCC
補正前テストデータ 1	19.8	-0.23	0.22
テストデータ 1 をケース 1 学習モデルで補正	6.8	-0.23	0.91
補正前テストデータ 2	16.6	-0.18	0.53
テストデータ 2 をケース 2 学習モデルで補正	20.8	-0.20	0.12

ケース 2 のテストデータのオキシダント濃度を、観測値と計算値で比較した結果を図 5.4(a)に示す。このテストデータに対して、ケース 2 の学習データで学習したアンサンブル回帰モデル（以下、ケース 2 学習モデル）を用いて誤差補正した結果を、図 5.4(b)に示す。また、それぞれの RMSE、NMB、PCC を表 5.2 に示す。補正前が RMSE16.6、NMB-0.18、PCC0.53 であるのに対し、補正後は RMSE10.7、NMB-0.16、PCC0.76 になっており、補正による精度の向上が見られた。ケース 2 の場合、テストデータは学習データとは異なる年のデータであるが、ケース 1 学習モデルの補正（図 5.3(c)）と異なり、ケース 2 学習モデルは適切な補正を行うことができる。

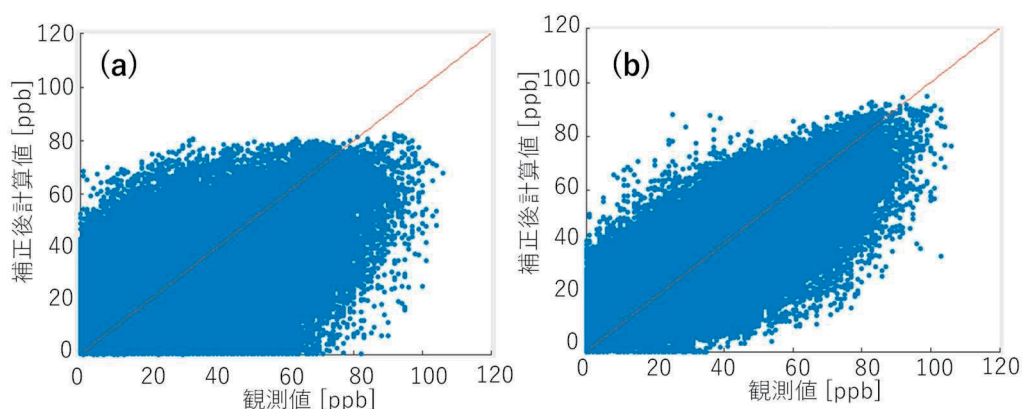


図 5.4 ケース 2 の誤差補正精度
((a)補正前テストデータ、(b)テストデータをケース 2 学習モデルで補正)

表 5.2 ケース 2 の RMSE、NMB、PCC

	RMSE	NMB	PCC
補正前テストデータ	16.6	-0.18	0.53
テストデータをケース 2 学習モデルで補正	10.7	-0.16	0.76

図 5.3(c)(d)と図 5.4(a)(b)で示された結果からも、未知のデータに対して適切に補正を行うことのできる学習モデルを作成するためには、単年のデータではなく、複数年のデータを学習に用いる必要があることが分かった。以降の節で行う検証に用いる学習・テストデータは、ケース 2 のデータを用いた。

5. 2. 学習モデルの比較結果

3.2.1 のケース 2 データを用い、3.2.2 で示したアンサンブル回帰、FNN、線形回帰による誤差補正を行い、精度を比較した。

補正前のテストデータの観測値と計算値を図 5.5(a)に示す。ケース 2 の学習データをアンサンブル回帰、FNN に学習させ、これら学習済モデルを用いて、テストデータの誤差補正を行った結果を図 5.5(b)(c)に示す。補正方法は 5.1 と同様である。また、学習データから線形回帰モデルを作成し、このモデルを用いてテストデータの誤差補正を行った結果を図 5.5(d)に示す。それぞれの RMSE、NMB、PCC を表 5.3 にまとめる。なお、図 5.5(a)は図 5.4(a)、図 5.5(b)は図 5.4(b)と同じものであるが、他のモデルと並べて比較するため、再掲している。表 5.3 から、RMSE、NMB、PCC はいずれのモデルで補正した場合も、補正前に比べて向上していた。しかし図 5.5(d)から分かるように、線形回帰では観測値が 80ppb 以上のデータに対して、計算値を過小に補正する傾向がみられた。一方、図 5.5(b)(c)から、アンサンブル回帰と FNN 回帰は線形回帰のような傾向はみられず、いずれの観測値濃度レベルに対しても、補正前に比べてデータの分散が抑えられていた。このことから、誤差補正に使用するモデルは、アンサンブル回帰と FNN 回帰が有効であることが分かった。

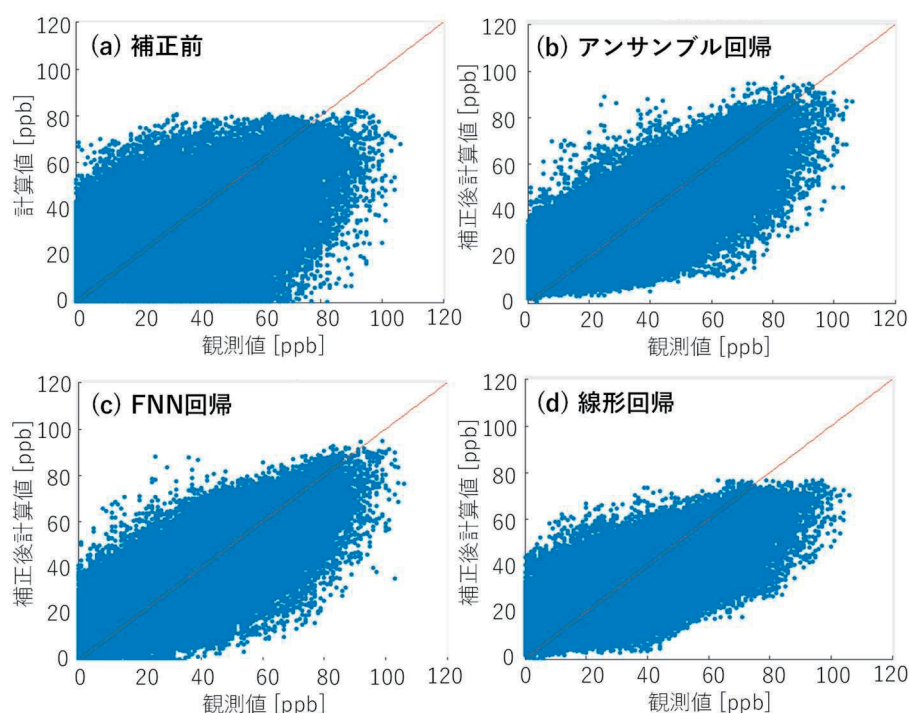


図 5.5 アンサンブル回帰モデル、FNN モデル、線形回帰モデルの誤差補正精度
 ((a)補正前テストデータ、(b) アンサンブル回帰、(c) FNN 回帰、(d) 線形回帰)

表 5.3 アンサンブル回帰モデル、FNN 回帰モデル、線形回帰モデルの RMSE、NMB、PCC

	RMSE	NMB	PCC
補正前テストデータ	16.6	-0.18	0.53
アンサンブル回帰	10.2	-0.16	0.78
FNN 回帰	10.7	-0.16	0.76
線形回帰	11.5	-0.16	0.71

5. 3. 重みを用いた学習の検証結果

3.2 で作成したモデルに比べ、高濃度の見逃しを減らすことに特化したモデルを、3.2.3 で示した Deep Learning Toolbox の SMNN 回帰アルゴリズム `train` を用いて作成した。

最適なノード数と隠れ層数の検証 (3.2.3 表 3.6) 結果を図 5.6 に示す。図 5.6 (a)(b)は層数を 1、ノード数をそれぞれ 100、300 としたモデルの補正結果、(c)(d)はノード数を 100、層数をそれぞれ 2、3 とした検証結果である。各結果の RMSE、NMB、PCC を表 5.4 に示す。表 5.4 と(a)(b)から、ノード数が増加すると精度が下がっていることが分かる。一方、(b)(c)(d)から、層数が増加すると精度が上がっていることが分かる。この結果から、重みの検証は、最も精度の高かったノード数 100、隠れ層数 3 のモデルで行うこととした。

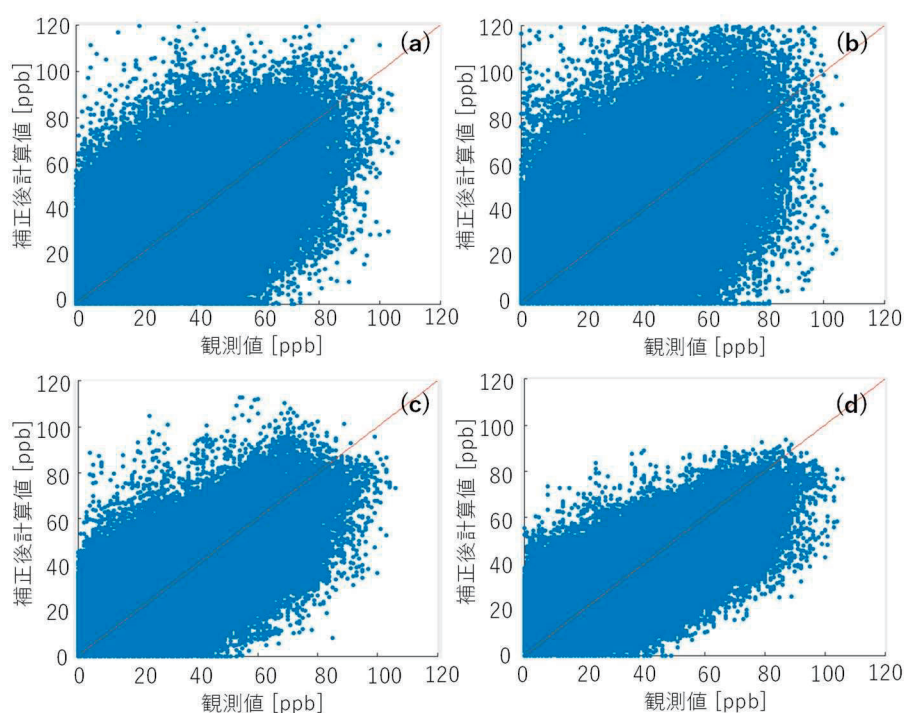


図 5.6 ノード数、隠れ層数の検証

((a) 層数 1 ノード数 100、(b) 層数 1 ノード数 300、
(c) 層数 2 ノード数 100、(d) 層数 3 ノード数 100)

表 5.4 ノード数、隠れ層数の検証 (RMSE、NMB、PCC)

		RMSE	NMB	PCC
ケース 1 (層数 1)	ノード数 100	17.1	0.087	0.50
	ノード数 300	18.9	0.020	0.46
ケース 2 (ノード数 100)	層数 2	13.0	0.00	0.65
	層数 3	11.9	-0.0081	0.69

学習データのうち、オキシダント濃度の観測値が 80ppb 以上でかつ計算値が観測値よりも 3 割以上小さいデータの重みを 10、50 に設定して学習させたモデルで補正した結果を、それぞれ図 5.7(c)(d)に示す。比較のため、補正前のテストデータの観測値と計算値を図 5.7(a)、重みをかけずに学習したモデルで計算値を補正した結果を図 5.7(b)に示している。補正方法は 3.2 と同様である。また、RMSE、NMB、PCC を表 5.5 に示す。図 5.7 と表 5.5 から、補正前は NMB が-0.18、観測値のオキシダント濃度が 80ppb 以上のデータに限定すると-0.36 と、計算値は過小評価傾向があり、オキシダント濃度が高いほど過小傾向が大きくなっていることが分かる。重みなしで補正をかけると、全データ、80ppb 以上のデータともに、RMSE、NMB、PCC いずれの値も改善しているが、80ppb 以上のデータの NMB は-0.26 であり、まだ過小評価であることが分かる。重み 10 モデルで補正すると、全データに対する RMSE、NMB、PCC は、重みなしモデルに比べると精度が下がっているが、補正前に比べると精度が上がっている。一方、80ppb 以上のデータに対しては、RMSE19.4、NMB-0.14 と、重みなしモデルよりもさらに過小評価が改善されていることが分かる。重み 50 モデルで補正すると、80ppb 以上のデータに対しては、

RMSE17.4、NMB-0.072 であり、さらに過小評価は改善されているが、全データに対しては、RMSE、NMB の精度が補正前のデータよりも下がっていた。

以上のことから、高濃度の見逃しを減らし、かつ全データに対しても補正前より精度を上げたい場合は、重み 10 モデルが適していると考えられた。一方、高濃度の見逃しを減らすことに特化させ、例えば、全データに対してではなく、補正後の計算値が 80ppb を超えると予測された場合のみに補正を適用するといった運用方法で使用する場合は、重み 50 モデルが適していると考えられた。

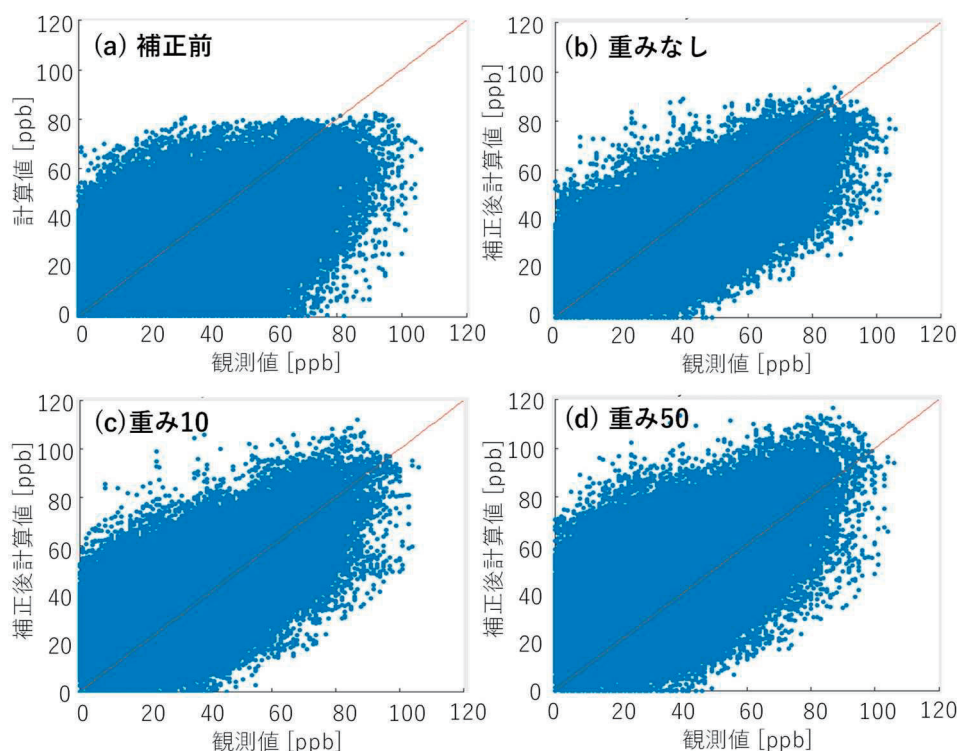


図 5.7 重みごとの誤差補正精度
((a)補正前テストデータ、(b)重みなし、(c)重み 10、(d)重み 50)

表 5.5 重みごとの RMSE、NMB、PCC

		RMSE	NMB	PCC
補正前テストデータ	全データ	16.6	-0.18	0.53
	観測 80ppb 以上	34.1	-0.36	-
重みなし	全データ	11.9	-0.0081	0.69
	観測 80ppb 以上	25.8	-0.26	-
重み 10	全データ	13.9	0.16	0.66
	観測 80ppb 以上	19.4	-0.14	-
重み 50	全データ	17.6	0.32	0.63
	観測 80ppb 以上	17.5	-0.072	-

5. 4. 予報システムへの組み込み

5.2、5.3 で作成したいずれかの学習モデルを予報システムに組み込むことで、毎日の予報結果に対し、自動で誤差補正をすることが可能になる。具体的には、以下の手順を自動で行うプログラム（以下、誤差補正プログラム）を作成した。

- ① 予報システムの計算結果ファイルから、各計算格子の説明変数を抽出する
- ② 学習済モデルを指定して説明変数を読み込ませ、目的変数（オキシダント濃度誤差）を予測させる
- ③ 得られた濃度誤差で、予報システムの計算結果を補正する
- ④ 補正した予報結果を、新たなファイルとして保存する

6 考察

機械学習の判断根拠はブラックボックスであり、本研究においても、3.2、3.3 で作成した学習モデルがどのような根拠に基づいてオキシダント誤差を予測したのかは不明である。しかし、誤差の傾向を考察することは、学習モデルの判断根拠を推測する一助になると考えられる。

表 6.1 に、ケース 2 の学習データのオキシダント誤差（計算値 - 観測値）と各説明変数との相関を示す。また、図 6.1(a)に、学習データのうち誤差が正（計算値が過大評価、図中橙色）になったデータ件数と、負（計算値が過小評価、図中青色）になったデータ件数を月別に示す。図 6.1(b)に、(a)と同様に誤差が正になったデータ件数と負になったデータ件数を時間別に示す。なお、図 6.1 は負（図中青色）の上に正（図中橙色）を重ねて描いており、例えば、青いバーが橙のバーよりも大きい時は、負の方が正よりも件数が多かったことを表す。

表 6.1 オキシダント誤差と各説明変数との相関

month_x	month_y	time_x	time_y	緯度	経度	気温	降水量	風速	風向	気圧	水蒸気量	日射	オキシダント	NO	NO ₂	トルエン	イソブレン
-0.030	-0.084	0.29	0.057	-0.079	-0.0050	-0.49	-0.076	-0.065	0.038	0.28	-0.47	-0.17	0.49	-0.19	-0.14	-0.045	-0.035

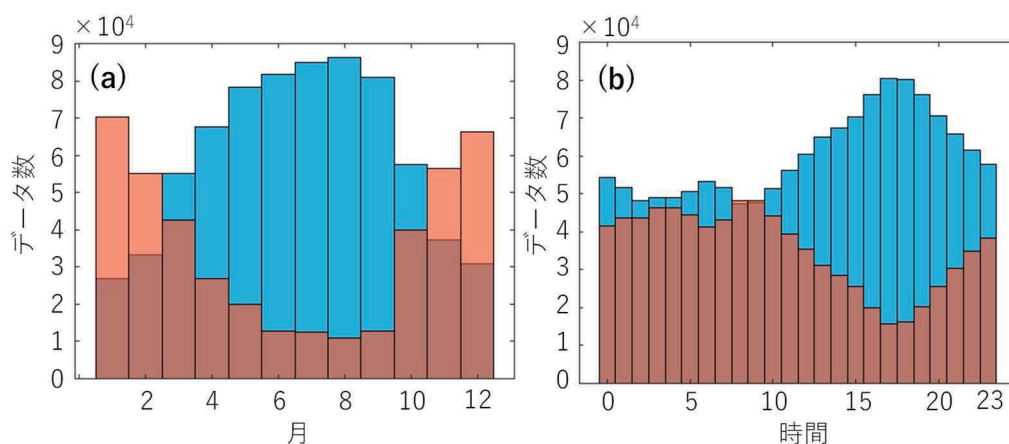


図 6.1 誤差種類別の月別・時間別データ件数

図 6.1(a)から、3~10 月は負の誤差（計算が過小）、11~2 月は正の誤差（計算が過大）の人が多いことが分かる。気温が高く水蒸気量の多い夏季を含む 3~10 月に誤差が負、気温が低く水蒸気量が少ない冬季を含む 11~2 月に誤差が正となったことが、表 6.1 の相関係数にお

いて、気温と水蒸気量がオキシダント誤差に対してやや強い負の相関を持った原因と考えられる。また、図 6.1(b)から、特に 17 時前後において、負の誤差が増えて正の誤差が減っていることが分かる。17 時前後は、地域で生成されたオキシダントが日最大値となりやすい時間帯である。一方で、移動性高気圧の停滞等によって越境汚染の影響が強くなった場合は、夜間も高濃度が続くケースが多くみられる。このことから、計算の過小は、越境汚染の濃度よりも、地域で生成するオキシダントが十分に再現できず過小になったことに起因するケースが多かったと推測された。

7 総括

- 本研究により、福岡県を対象とした詳細な光化学オキシダントの濃度予報システムが構築された。この予報システムにより、数日先までの福岡県内の地域別のオキシダント濃度、越境汚染の寄与、NO_x、VOC 排出量削減効果が予測可能になった。この予報システムは行政担当部署と連携して運用し、福岡県の環境行政において多大な貢献となっている。
- オキシダント生成に関係する 16 要素（気象、地理、大気汚染物質等の計算値）を説明変数、オキシダント濃度の計算値と観測値との誤差を目的変数とし、過去 3 年間分の観測・計算値を用いて、説明変数から誤差を予測する機械学習モデルを作成した。別の 1 年間分のデータをテストデータとして数種類のモデルの検証を行った結果、アンサンブル回帰、FNN 回帰、SMNN 回帰で、誤差補正による精度の向上がみられた。
- SMNN 回帰では、オキシダント濃度の観測値が 80ppb 以上でかつ計算値が観測値よりも 3 割以上小さいデータに重みをかけて学習させることで、高濃度の見逃し低減に特化した学習モデルが作成可能であることが分かった。
- 予報システムが出力する毎日のオキシダント濃度予測結果に対して、説明変数を抽出し、作成した学習モデルに読み込ませて誤差を予測させ、補正するという一連の作業を自動で行うプログラムを構築した。予報システムにこの誤差補正プログラムを搭載することで、より精度の高い予報が可能となった。

8 期待される成果の活用の方角性

- 予報システムは既に県での運用を開始しており、予測結果は福岡県の光化学オキシダント監視体制及び緊急時対応で行動の判断材料として活用されている。予測結果は福岡県のほか、北九州市、福岡市、大牟田市、久留米市の大気担当職員とも共有している。
- 高濃度が予測された場合には、福岡県公式 LINE により県民へ情報を配信することで、健康被害抑制に貢献している。なお、2024 年度には県のホームページ・SNS(X)で、予測結果を毎日公開することを検討している。
- 2022 年 12 月の異臭発生事例や 2023 年 1 月の PM_{2.5} 高濃度事例で示されたように、本システムはオキシダント以外の大気汚染物質に対しても、濃度予測や原因解明等に活用することが可能である。
- 本研究で構築した誤差補正プログラムを予報システムに組み込むことで、オキシダント濃度の予測精度をさらに高めることが可能となる。誤差補正プログラムに含まれる学習モデルの種類は変更可能であり、例えば、高濃度の見逃しを減らす用途としては重みをかけた学習モデルを使用するなど、目的に応じた使い分けが可能である。

- 各学習モデル自体も、学習データの変更に伴う再学習が容易に可能であるため、学習期間や排出量データの更新が行われた場合にも、柔軟に対応が可能である。また、他の自治体でこのシステムを導入する等のフレキシブルな運用が可能である。
- 誤差補正は通常計算に対してのみ行うが、越境寄与濃度、排出量削減効果の試算に対しても、天気予報の信頼度のような情報を付与することが可能である。例えば、補正する誤差が大きい場合は、通常計算の信頼度が低いため、通常計算に基づいて算出する越境寄与濃度、排出量削減効果の試算の信頼度も低くなる。そのため、これらの結果は”信頼度 低”として行政と共有し、観測データ等、シミュレーション以外の情報と合わせて解析をするなど、注意して扱う必要がある。一方、補正する誤差が小さい場合は、通常計算の信頼度が高く、越境汚染寄与計算、NO_x、VOC 削減計算の信頼度も高いと考えられるため、”信頼度 高”として共有する。
- ユーザーマニュアルの整備や、国立環境研究所と地方環境研究所（41 機関が参画）との共同研究で勉強会等を行うことで、他自治体への予報システム導入を促進することが期待できる。

【学会発表・論文】

- ・福岡県における大気シミュレーションの活用について，日韓海峡沿岸環境技術交流協議会第 2 回実務者会議, 2022.7.29
- ・大気シミュレーションモデルを用いた PM2.5、光化学オキシダントの解析事例および高濃度予測，九州衛生環境技術協議会, 2022.10.13
- ・大気シミュレーションモデルによる県独自予測情報の配信等，令和 4 年度環境白書，p52, 2022.12
- ・CMAQ 使用事例，全環研研究集会及びⅡ型共同研究による合同勉強会, 2023.3.8
- ・福岡県におけるシミュレーションによる大気汚染予測とその活用について，全国大気汚染防止連絡協議会第 68 回全国大会, 2023.11.17
- ・福岡県における火山に起因する大気汚染物質濃度の上昇について，第 50 回環境保全・公害防止研究発表会, 2023.11.17
- ・福岡県での大気シミュレーション活用事例と今後の展望，全国環境研会誌 48 巻，第 4 号，p13-18, 2023