

【5RMF-2201】 【体系番号JPMEERF20205R02】
環境研究総合推進費（環境問題対応型研究（一般課題））（令和4～5年）

機械学習によって 観測データと統合された 新しい大気汚染予測システムの開発と実働実験

令和6年6月28日 事後評価ヒアリング

課題代表者 弓本 桂也（九州大学 応用力学研究所）

サブ2代表 菅田 誠治（国立環境研究所）

研究分担者 板橋 秀一（電力中央研究所）

重点課題⑩大気・水・土壌等の環境管理・改善のための対策技術の高度化及び評価・解明に関する研究
④環境問題の解決に資する新たな技術シーズの開発・活用

行政ニーズ 非該当

研究背景

大気汚染予測システムの
予測精度は定量的に十分であるとは言えない

大気汚染予測システムは重要な社会インフラに



国立環境研究所が運用する大気汚染予測システム（VENUS）。

- ・ PM2.5や光化学オキシダントについての早期警戒
- ・ 注意報発令や解除の参考

➡ **社会の重要なインフラの一部となりつつある。**

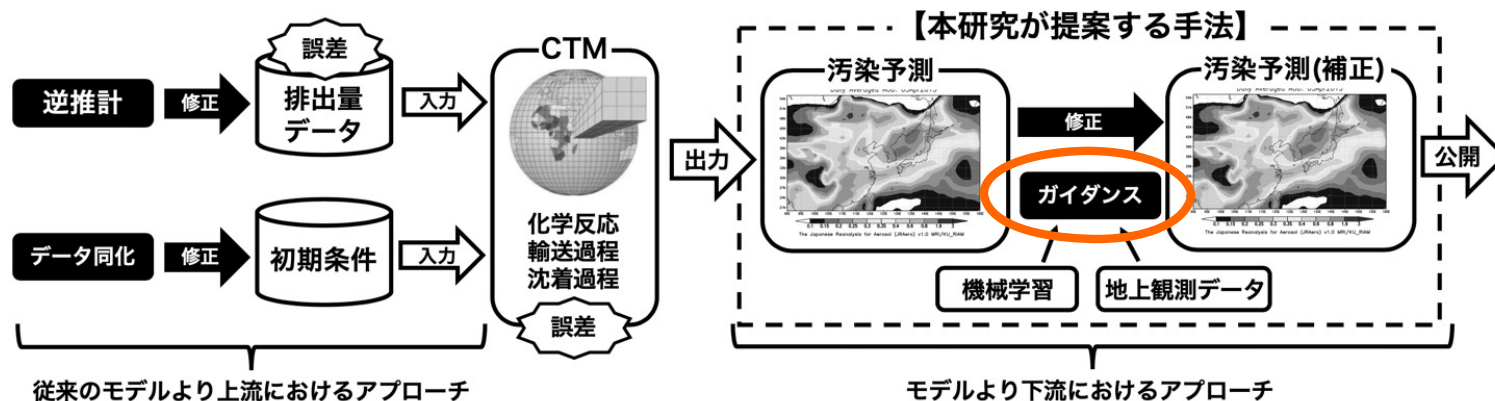
一方、光化学オキシダントの過大バイアス、高濃度PM2.5イベントの過小評価など

➡ **依然として予測精度に改良の余地があり、定量的な予測が出来ずにいる**

ガイダンスによる予測精度向上

（ガイダンス：数値天気予報で用いられる事後処理の1つで、予報の補正・翻訳と呼ばれる。）

先行課題である推進費5RF-2002（2020-2021）の中で、機械学習を応用した**大気汚染予測システム版ガイダンス**を開発、PM2.5およびオゾンの濃度予測の精度向上に非常に有効であることが確認された。



研究開発目的

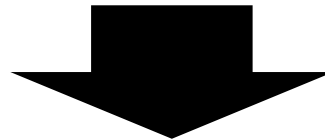
ガイドンス導入版VENUSの開発
実働実験を通じた現業運用の検討

先行課題（開発と技術検証）

【5RF-2002】

機械学習を用いた大気汚染予測システムへのガイダンス手法の開発と予測精度向上

- ① 機械学習の技術を応用し
 - ② 大気汚染広域監視システムで蓄積された観測を学習データとすることで
 - ③ 大気汚染予測版のガイダンスを構築することで
- PM2.5および光化学オキシダント予測の定量的な精度向上を実現する。



本課題（実働実験と現業運用）

- ・ ガイダンスの開発・改良
- ・ ガイダンス導入版VENUSの開発
- ・ ガイダンス導入版VENUSを用いた実働実験と現業運用の検討

[サブ1：開発班]

- ・ ガイダンスの開発・改良
- ・ ガイダンス導入版VENUSの開発
- ・ 実働実験結果の検証



[サブ2：VENUS班]

- ・ ガイダンス導入版VENUSの開発
- ・ 実働実験の実施
- ・ 現業運用にうめけた検討

研究目標および開発内容

① VENUSを想定したガイダンスの開発

- 機械学習の技術を応用し、大気汚染物質広域監視システム（そらまめ君）の観測データを教師データに、PM2.5とオゾン濃度予測精度向上を目的としたガイダンスを開発する。

② ガイダンスの改良と精度検証

[数値目標]

PM2.5: 相関係数 >0.7 、平均二乗誤差の30%改善 **オゾン:** 相関係数 >0.8 、平均二乗誤差と平均バイアスの30%改善

- 濃度区分予測を想定したガイダンスの作成と予測信頼度の算出
- 現業運用を想定したガイダンスの全国展開と精度検証

③ ガイダンス導入版VENUSの開発と現業運用を想定した実働実験の実施

- 運用中のVENUSを基礎としたシステム開発と実働実験の実施
- 現業運用にむけた課題の洗い出し

本研究の特色

- ① 機械学習手法およびガイダンスの概念の大気汚染予測への導入。
- ② 大気汚染広域監視システムで蓄積された観測値（ビッグデータ）の活用。
- ③ 機械学習でモデルと観測を統合した新しい大気汚染予測システムの開発。

研究開発内容① – 機械学習によるガイダンス構築

【学習によるガイダンス作成】

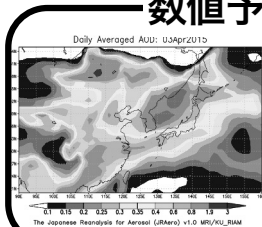
化学輸送モデル



※国立環境研究所が運用する
VENUSと同じモデル、設定を用いる

出力

数値予測



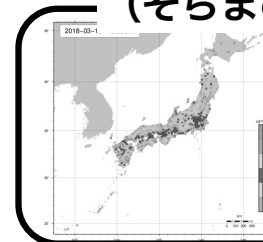
- ・PM2.5
- ・オゾン濃度
- ・気温
- ・相対湿度
- ・風向風速

入力データ

教師データ

大気汚染物質広域監視システム

(そらまめくん)



- ・PM2.5
- ・オゾン濃度
- ・地上770点

※約200万におよぶビッグデータ

③ニューラルネットワーク (機械 (深層) 学習)

【精度検証】

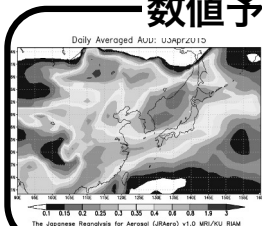
※教師データとは別の年を対象に予測実験を行い構築したガイダンスの検証を行う

化学輸送モデル



出力

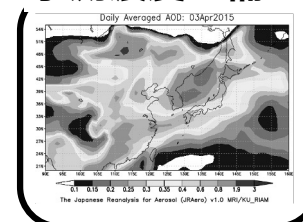
数値予測



- ・PM2.5
- ・オゾン濃度
- ・気温
- ・相対湿度
- ・風向風速

ガイダンス

予測濃度の補正



検証

移植

【ガイダンス導入版VENUSの開発と実働実験】

ガイダンス導入版VENUS



実働実験
現業化の検討

機械学習によって観測データと統合された新しい大気汚染予測システムの開発と実働

(ロング資料)

研究開発内容② – 濃度区分予測と予測信頼度

VENUSでは各地域の濃度区分予測をアイコンで提供 → ガイダンスによる濃度区分予測の精度向上

アイコン	意味
	とても少ない < 10 $\mu\text{g}/\text{m}^3$
	少ない < 15 $\mu\text{g}/\text{m}^3$
	やや多い < 35 $\mu\text{g}/\text{m}^3$
	多い < 50 $\mu\text{g}/\text{m}^3$
	とても多い < 70 $\mu\text{g}/\text{m}^3$
	非常に多い > 70 $\mu\text{g}/\text{m}^3$

福岡(PM2.5)	観測値	4月1日	4月2日	4月3日	4月4日	4月5日	4月6日	4月7日
	ガイダンスなし							
	ガイダンスあり							
	予測の信頼度	A A A A A A A A B C A A A A A A A A A B B A						
	観測値	4月8日	4月9日	4月10日	4月11日	4月12日	4月13日	4月14日
	ガイダンスなし							
	ガイダンスあり							
	予測の信頼度	A B A A B A A A B B A B A A C A A A A A A A B B						
	観測値	4月15日	4月16日	4月17日	4月18日	4月19日	4月20日	4月21日
	ガイダンスなし							
	ガイダンスあり							
	予測の信頼度	C B A A B A A A A A A A C D D B A B B A A A A B B						
	観測値	4月22日	4月23日	4月24日	4月25日	4月26日	4月27日	4月28日
	ガイダンスなし							
	ガイダンスあり							
	予測の信頼度	A A A A A A A A B C A A A A A A A A A B A A A B B A						

福岡におけるPM2.5濃度区分予測のカレンダー表記。色は濃度区分を示し暖色の濃度が高い。上から観測、ガイダンスなし予測、連続型ガイダンスを作用させた結果である。4段目は予測の信頼度（Aが最も信頼度が高く、Dがもっとも低い）。

ガイダンスなし予測に見られた過大評価（12日や19日）、過小評価（5-6日や26-27日）の修正に成功。

研究開発内容② – 濃度区分予測と予測信頼度

VENUSでは各地域の濃度区分予測をアイコンで提供 → ガイダンスによる濃度区分予測の精度向上

的中率による評価

濃度区分予測の的中率（＝濃度区分が一致した時刻／総予測時刻）

[PM2.5]

	ガイダンスなし	ガイダンスあり
福岡	47.5	68.6
北九州	50.2	68.9
大阪	55.3	66.6
名古屋	51.6	62.4
東京	47.9	65.1
仙台	28.6	50.8

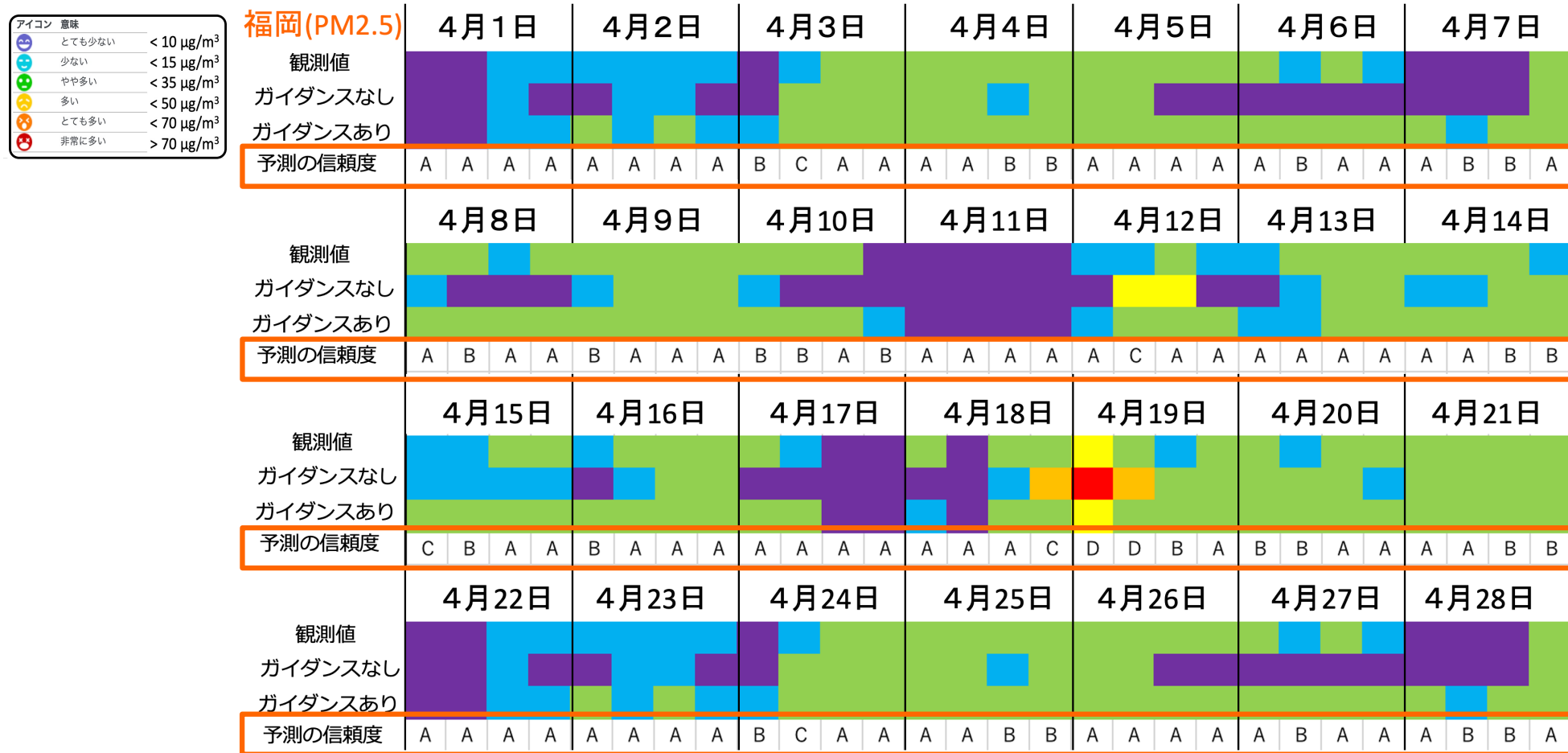
[オゾン]

	ガイダンスなし	ガイダンスあり
福岡	38.6	71.1
北九州	50.5	71.0
大阪	46.6	74.1
名古屋	42.2	75.8
東京	52.4	74.7
仙台	57.8	65.1

濃度区分予測の的中率はPM2.5濃度では10.8–21.1ポイント、オゾン濃度では7.3–33.6ポイント改善した。

研究開発内容② – 濃度区分予測と予測信頼度

気象庁の天気予報のような予測の信頼度算出 = ガイドンスの修正量を統計的に処理して算出



福岡におけるPM2.5濃度区分予測のカレンダー表記。色は濃度区分を示し暖色の濃度が高い。上から観測、ガイドンスなし予測、連続型ガイドンスを作用させた結果である。4段目は予測の信頼度（Aが最も信頼度が高く、Dがもっとも低い）。

PM2.5 : 信頼度と的中率の間に明確な相関関係 = 適切な信頼度

オゾン : 信頼度と的中率の間に相関関係なし = 予測の過大バイアスが原因か

研究開発内容③ – ガイダンスの全国展開と精度検証

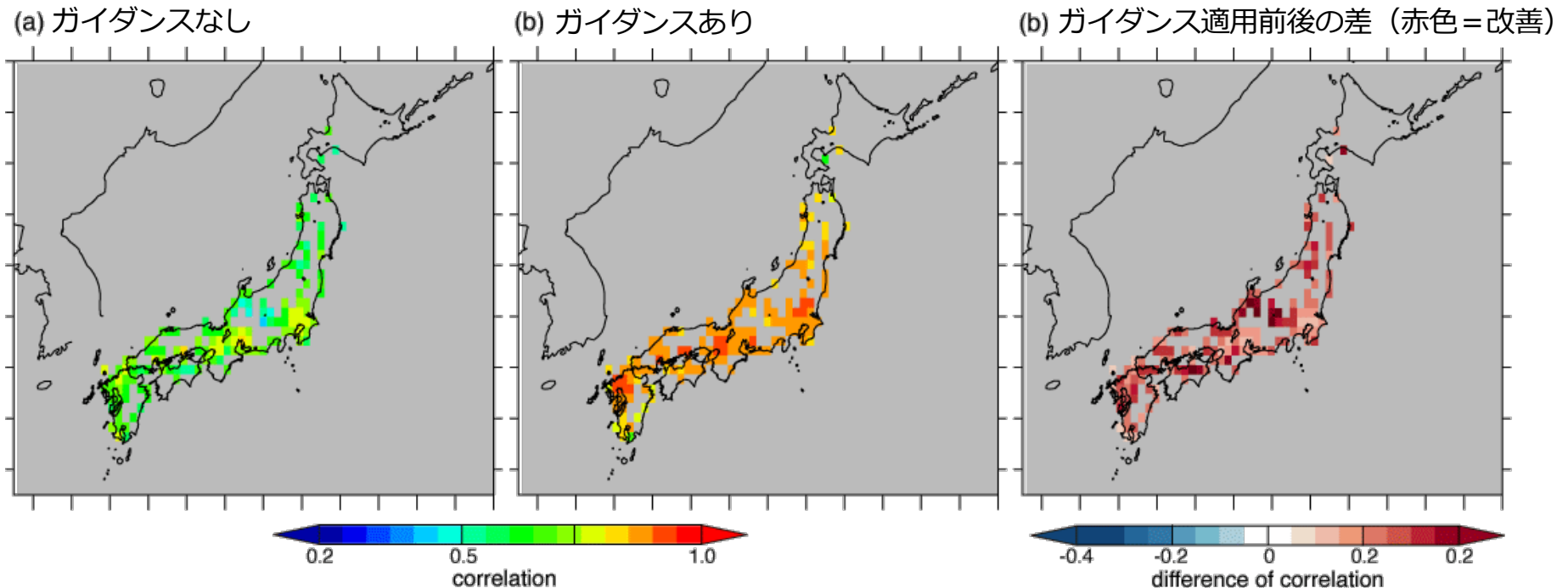
先行課題：特定の都市や地域を対象にガイダンスの開発



ガイダンス導入版の大気汚染予測を行うにあたり……

本課題：日本全国を144の地域にわけ、それぞれのグリッドでガイダンスを作成
ガイダンス適用後の濃度予測の分布表示が可能に

オゾン濃度予測におけるガイダンス適用前後の相関係数 (R) の比較



[Yumimoto et al.: Development of post processing for air quality forecasting with machine learning technique. (in prep.)]

- ・ ガイダンスの適用によって、すべてのグリッドで相関係数が改善
- ・ 欠損の地域 = 観測がないためガイダンスが作成できない地域

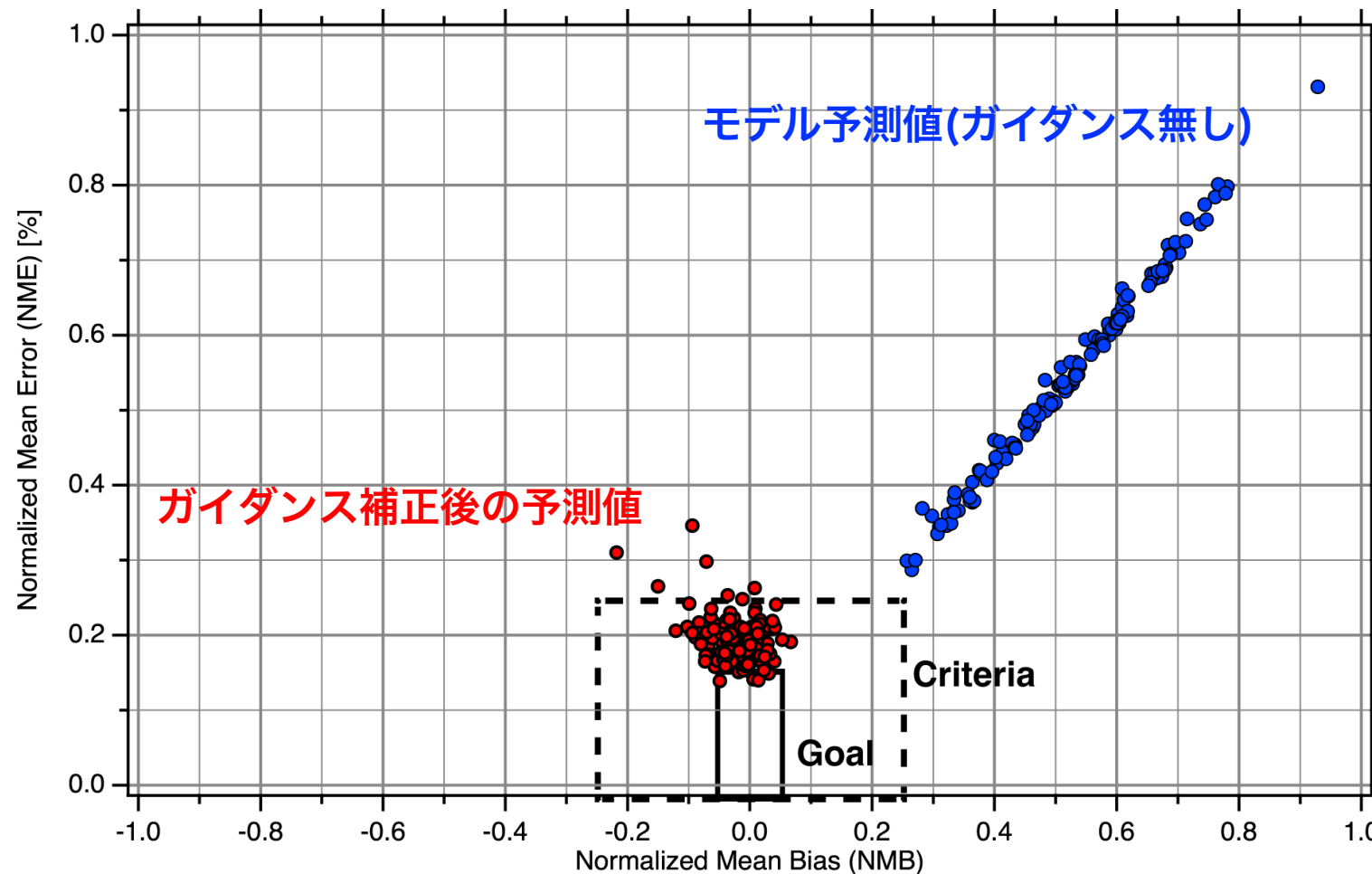
← 座標変換や画像補完技術で対応

研究開発内容③ – ガイダンスの全国展開と精度検証

Emery et al., (2017)によるオゾン濃度の数値予測に対する予測再現性の評価

($NME \leq 25\%$ かつ $NMB \leq \pm 15\%$) で許容範囲 (criteria) を達成

($NME \leq 15\%$ かつ $NMB \leq \pm 5\%$) で目標範囲 (goal) を達成



[Yumimoto et al.: Development of post processing for air quality forecasting with machine learning technique. (in prep.)]

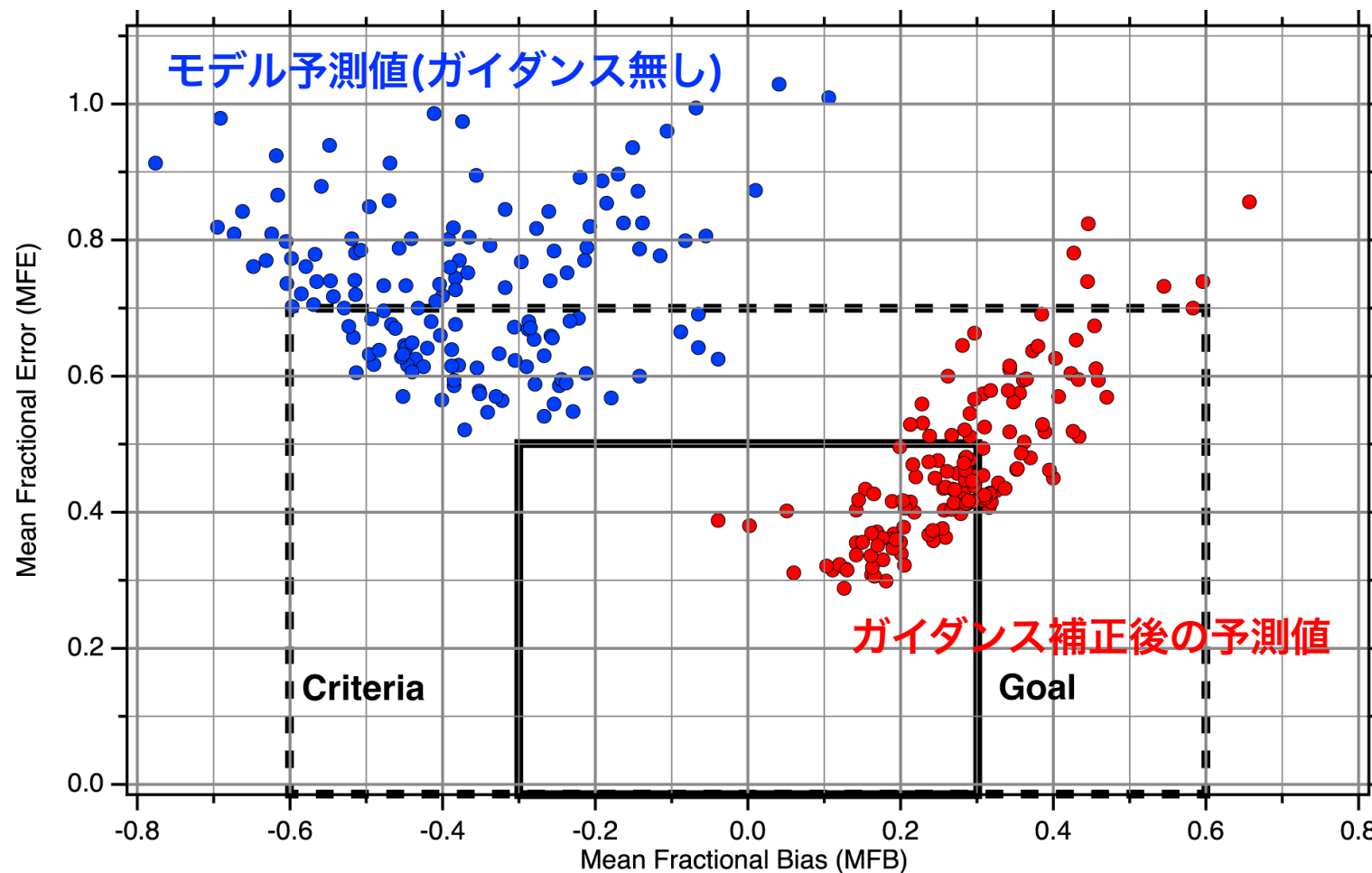
- ・ ガイダンスなし：すべてのグリッドで許容・目標範囲を達成できず。
- ・ ガイダンスあり：138/144のグリッドで許容範囲を、2/144のグリッドで目標範囲を達成。

機械学習によって観測データと統合された新しい大気汚染予測システムの開発と実働実験（事後ヒアリング資料）

研究開発内容③ – ガイダンスの全国展開と精度検証

Boylan and Russell (2006)によるPM2.5濃度の数値予測に対する予測再現性の評価

(MFE $\leq 70\%$ かつMFB $\leq \pm 60\%$) で許容範囲 (criteria) を達成
(MFE $\leq 50\%$ かつMFB $\leq \pm 30\%$) で目標範囲 (goal) を達成



[Yumimoto et al.: Development of post processing for air quality forecasting with machine learning technique. (in prep.)]

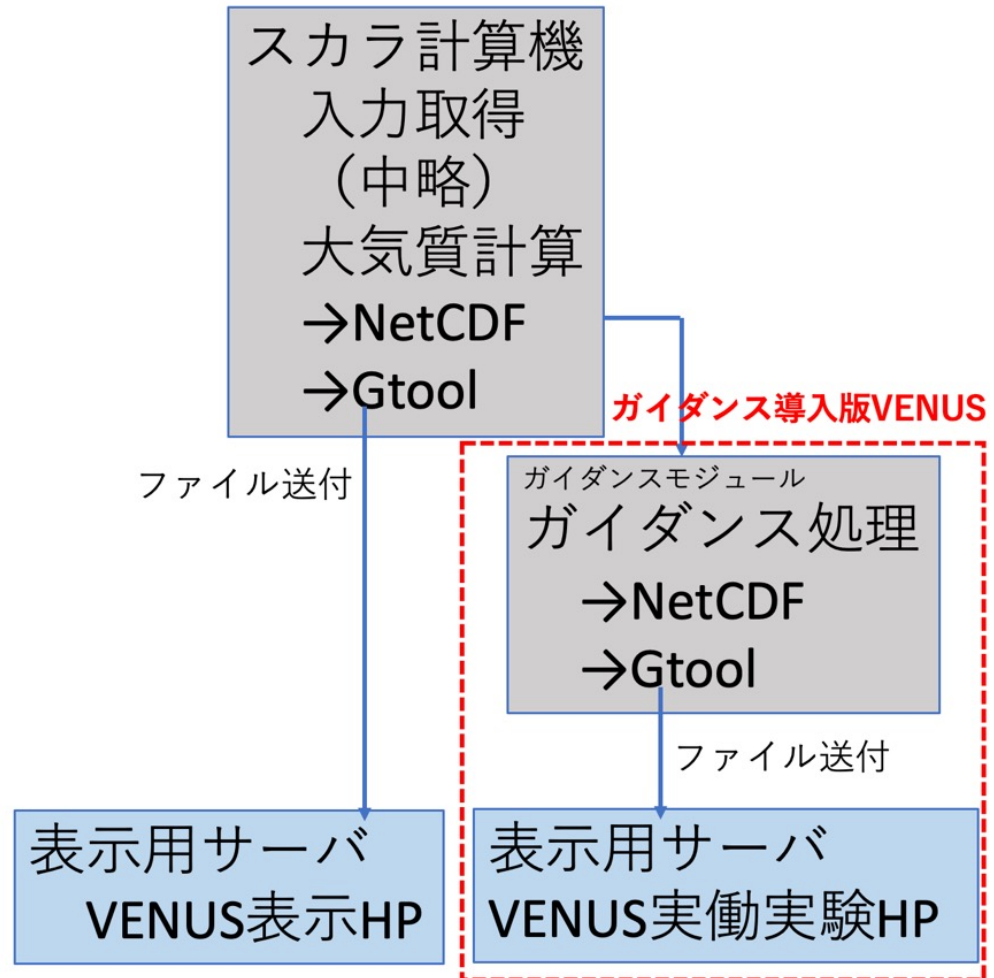
- ・ ガイダンスなし : 66/144のグリッドで許容範囲を達成。目標範囲はすべてのグリッドで達成できず。
- ・ ガイダンスあり : 138/144のグリッドで許容範囲を、79/144のグリッドで目標範囲を達成。

研究開発内容④ – ガイダンス導入版VENUSの開発

現行VENUS

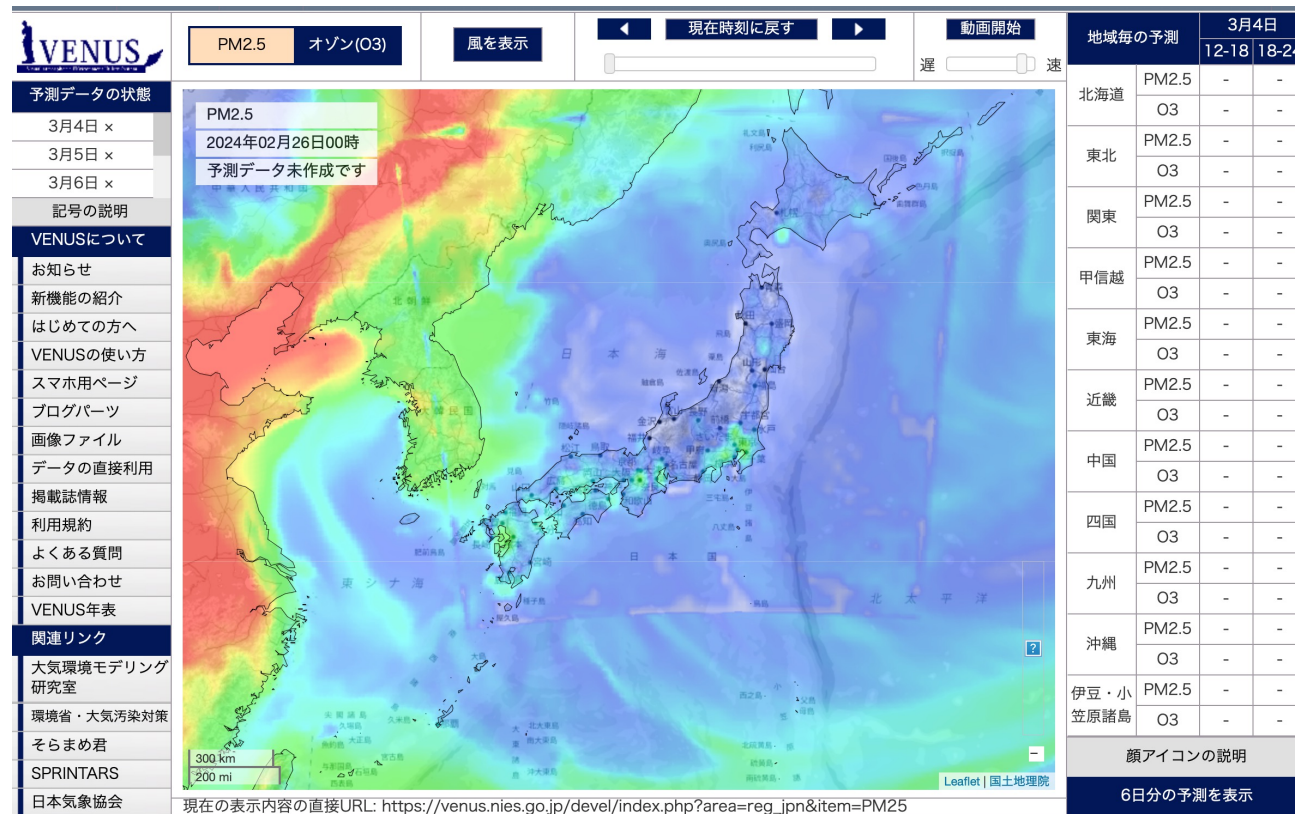


実働実験



- 既存のVENUSのシステムを基礎とし、作成したガイダンスを適用させるガイダンス導入版VENUSの開発を行った。
(データ同化と比べ、比較的用意に導入でき、計算コストも抑えられる)
- ガイダンス導入版VENUSを用いた実働実験を実施した。

研究開発内容④ – ガイダンス導入版VENUSの開発



(<https://venus.nies.go.jp/devel/>, 接続方法は報告書に記載)

ガイダンス導入版VENUSの実働実験HP。ガイダンスの適用によって都市部の局所的な高濃度が再現されるようになった。

- ・ **ガイダンス導入版VENUSは現業運用直前の段階まで準備することができた。**
- ・ 実働実験の結果、以下のような改善点を洗い出すことができた。
 - PM2.5濃度予測においてガイダンス適用後に濃度が過大になる傾向。過去のデータで学習させているため、最近のPM2.5濃度の減少傾向が反映されていないためか？
 - VENUS本体の改良とバージョンアップの影響は？

現業運用の前に、近年の観測データを含めた再学習を行う必要がある。

機械学習によって観測データと統合された新しい大気汚染予測システムの開発と実働実験（事後ヒアリング資料）

研究目標の達成状況

□ 「目標を上回る成果をあげた」と評価できる。

- 機械学習（AI）の技術を応用し、大気汚染物質広域監視システムで蓄積された観測値を教師データとして活用することで、VENUSを対象とした**大気汚染予測に特化したガイダンスの構築に成功**した。

＜PM2.5濃度予測目標＞ 相関係数0.7以上、平均二乗誤差の30%減少

- 相関係数はすべてのグリッドで改善し、67グリッド(46.5%)で0.7を超えた（ガイダンス適用前は1グリッド）。
- 平均二乗誤差はすべてのグリッドで改善し、137グリッド(全体の95.1%)で30%以上の改善を達成した。

＜オゾン濃度予測目標＞ 相関係数0.8以上、平均二乗誤差の30%減少、平均バイアスの30%減少

- 相関係数はすべてのグリッドで改善し、136グリッド(94.4%)で相関係数が0.8を超えた（ガイダンス適用前は2グリッド）。
- 平均二乗誤差はすべてのグリッドで改善し、すべてのグリッドで30%以上の改善を達成した。平均バイアスは平均で93.2%の改善を達成した。

- 運用中のVENUSを基礎とし、**ガイダンス導入版ガイダンスを開発**した。
- ガイダンス導入版VENUSの**実働実験やweb表示実験**を実施、ガイダンスの改良ポイントの洗い出しを行った。
- ガイダンス導入版VENUSの**現業運用直前の段階までの準備**を行うことができた。

環境政策への貢献 <見込まれる成果>

- 環境省と国立環境研究所が構築・運用してきた大気汚染予測システム（VENUS）にガイダンスを導入した。
- ガイダンスの導入によって、予測精度の大幅な改善を達成することができた。
- 実働実験やweb表示実験を実施。ガイダンス導入版VENUSの近年中の運用開始の見通しを立てることができた。



□ガイダンス導入版VENUSの現業運用によって

- ・ 大気汚染予測の精度向上
- ・ 越境大気汚染を含むPM2.5およびオゾンによる大気汚染現象の解明
- ・ 監視および早期警戒を通じた健康被害の軽減
- ・ 注意喚起および注意報発令の精度向上および発表・発令時刻の迅速化
- ・ これらを通じて国および地方自治体の環境政策への貢献

研究成果の発表状況等

査読付き論文に準ずる成果発表	0 件
その他誌上発表（査読なし）	0 件
口頭発表（学会等）	8 件
「国民との科学・技術対話」の実施	3 件
マスコミ等への公表・報道等	0 件
本研究に関連する受賞	0 件

<口頭発表>

1. Yumimoto et al.: Development of aerosol data assimilation system integrated with satellite retrievals, 2023 International Conference on CMAS-Asia-Pacific, Saitama, Japan, 12-21 July, 2023 (Oral)
2. Yumimoto et al.: Integration of aerosol retrieval and assimilation: current status and future direction, AERSS Annual Meeting 2023, Wuhan China, 17-19 September 2023 (Oral, invited)
3. Yumimoto et al.: Integration of aerosol retrieval and assimilation: current status and future direction, AGU Fall Meeting 2023, San Francisco, USA, 11-15 December, 2023 (Oral, invited)
4. 弓本桂也、第64回大気環境学会年会、(2023)「大気環境予測における観測データの役割」
5. 弓本桂也、菅田誠治、板橋秀一、第64回大気環境学会年会、(2023)「機械学習によって観測データと統合された新しい大気汚染予測システムの開発と実働実験」
6. 弓本桂也、第64回大気環境学会年会、(2023)「大気環境予測における観測データの役割」
7. 弓本桂也、菅田誠治、板橋秀一、第64回大気環境学会年会、(2023)「機械学習によって観測データと統合された新しい大気汚染予測システムの開発と実働実験」（招待講演）
8. 弓本桂也、鉄鋼環境基金 助成研究2023年度成果報告会、「機械学習を用いた大気汚染予測システムの予測精度向上」（2024年3月6日）

<国民との科学・技術対話>

1. 弓本桂也、九州大学総合理工学府オープンキャンパス、「機械学習について」（2022年5月21日、オープンキャンパスに出店。機械学習の基礎と大気環境予測への応用例について発表）
2. 弓本桂也、九州大学総合理工学府オープンキャンパス、「機械学習について」（2023年5月27日、オープンキャンパスに出店。機械学習の基礎と大気環境予測への応用例について発表）
3. 弓本桂也、気象学会九州支部「第24回気象教室」、「大気環境予測におけるデータ同化」（2023年11月26日）