

2025年12月19日

水・大気環境連携セミナー2025
－データで切り拓く環境研究の未来－
@自動車会館

機械学習と環境データを用いた大気質予測： 手法と適用事例

荒木 真

大阪大学大学院工学研究科

本日の内容

1. 機械学習による大気質予測手法の概要
2. 機械学習による大気質予測の事例
 - 過去30年間の日本全国PM2.5濃度推計
 - 関西地方のPM_{2.5}成分濃度および質量濃度の推計
 - ベンゾ[a]ピレンの長期濃度推計
3. まとめ

大気質予測の必要性

- ・ 大気質濃度は測定地点における点データであり、空間的に連續なデータとしては得られない。
- ・ 測定地点以外の大気質は不明 → 予測が必要



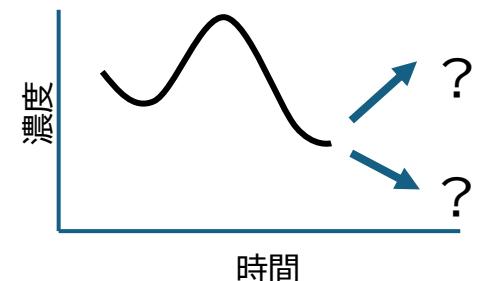
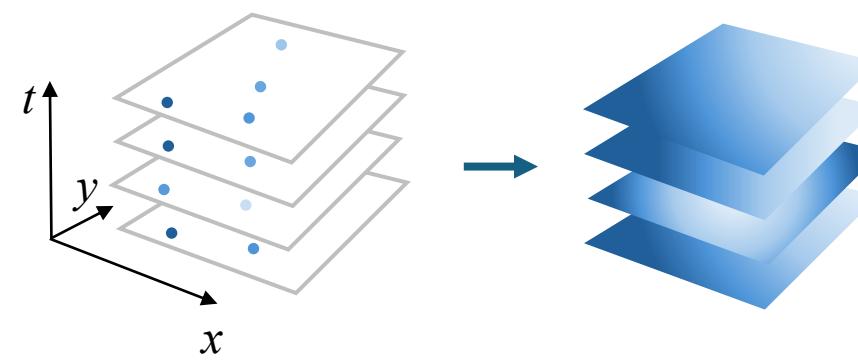
大気質予測モデル

□プロセス(物理)モデル

- ・大気質の物理・化学プロセスを数式で計算し、大気中濃度の時空間変動を得る
- ・観測値が不要(ただし、評価には必要)
- ・計算負荷が高い
- ・大気質モデル、化学輸送モデルと呼ばれるもの

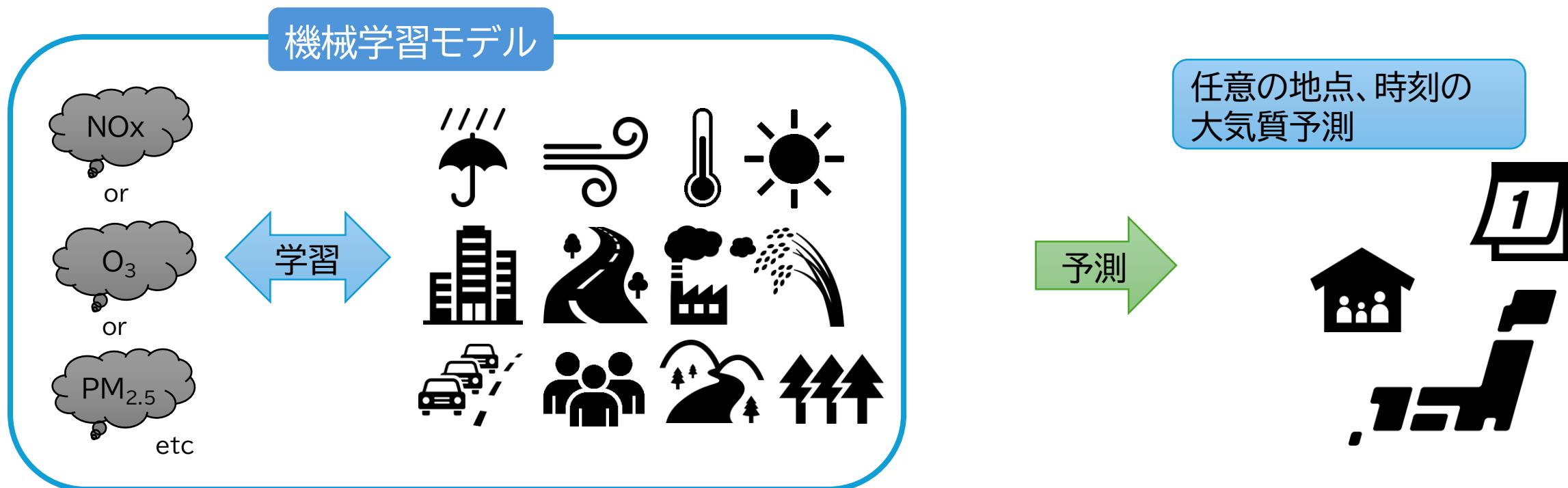
□機械学習(データ駆動型)モデル

- ・大気質濃度の(時間・空間)変動をモデルが学習・予測する
- ・観測値が必須
- ・計算負荷は高くない → 長期、広域、高解像度の予測に有利
- ・空間変動予測
- ・時系列予測

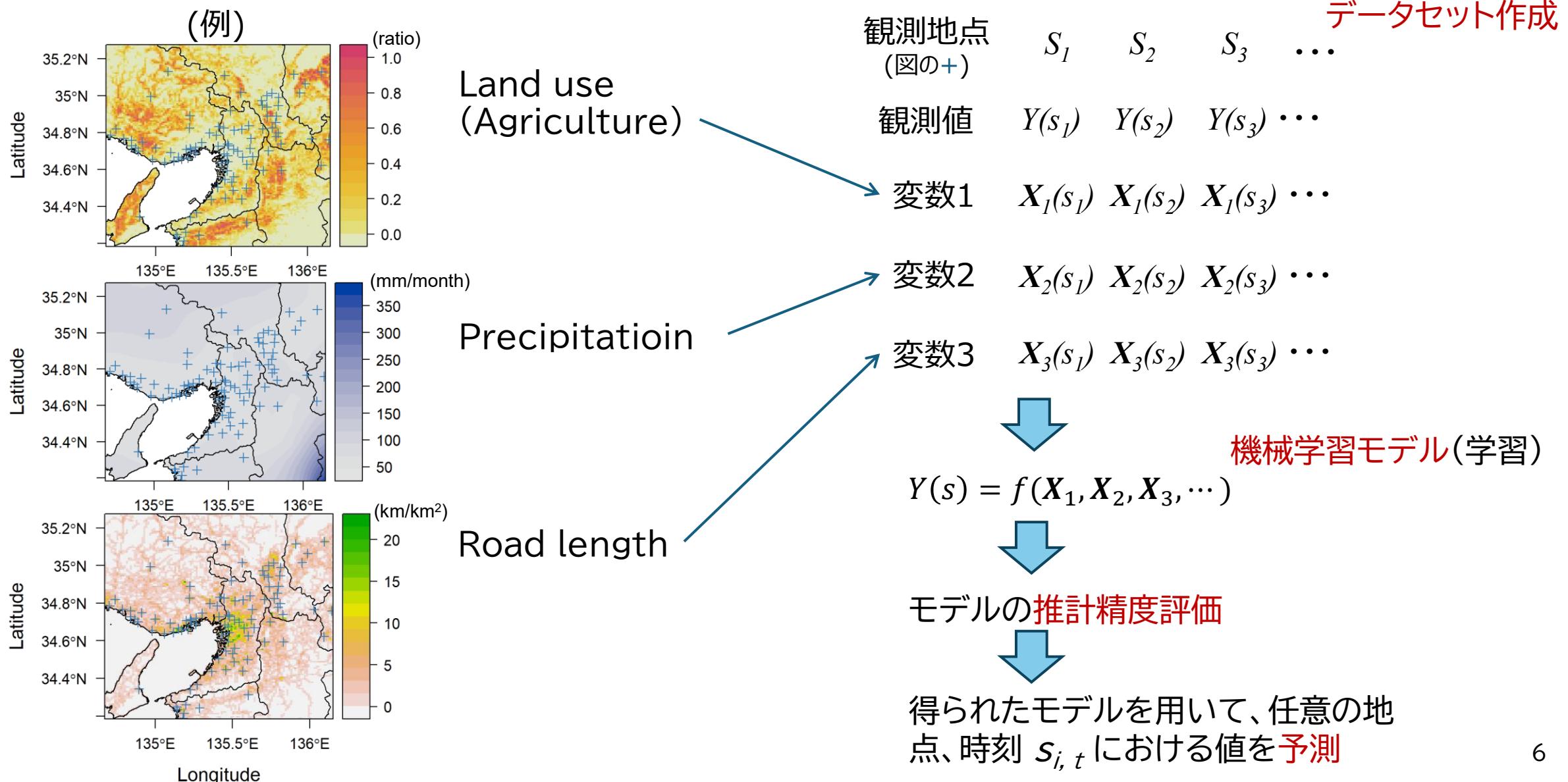


機械学習モデルによる大気質予測

- ある地点、ある時刻(日)における大気質と、対応する変数(下記に例)
- これらの(表面的な)関係を学習する
- 因果関係の解明には不向き



機械学習モデルによる大気質予測

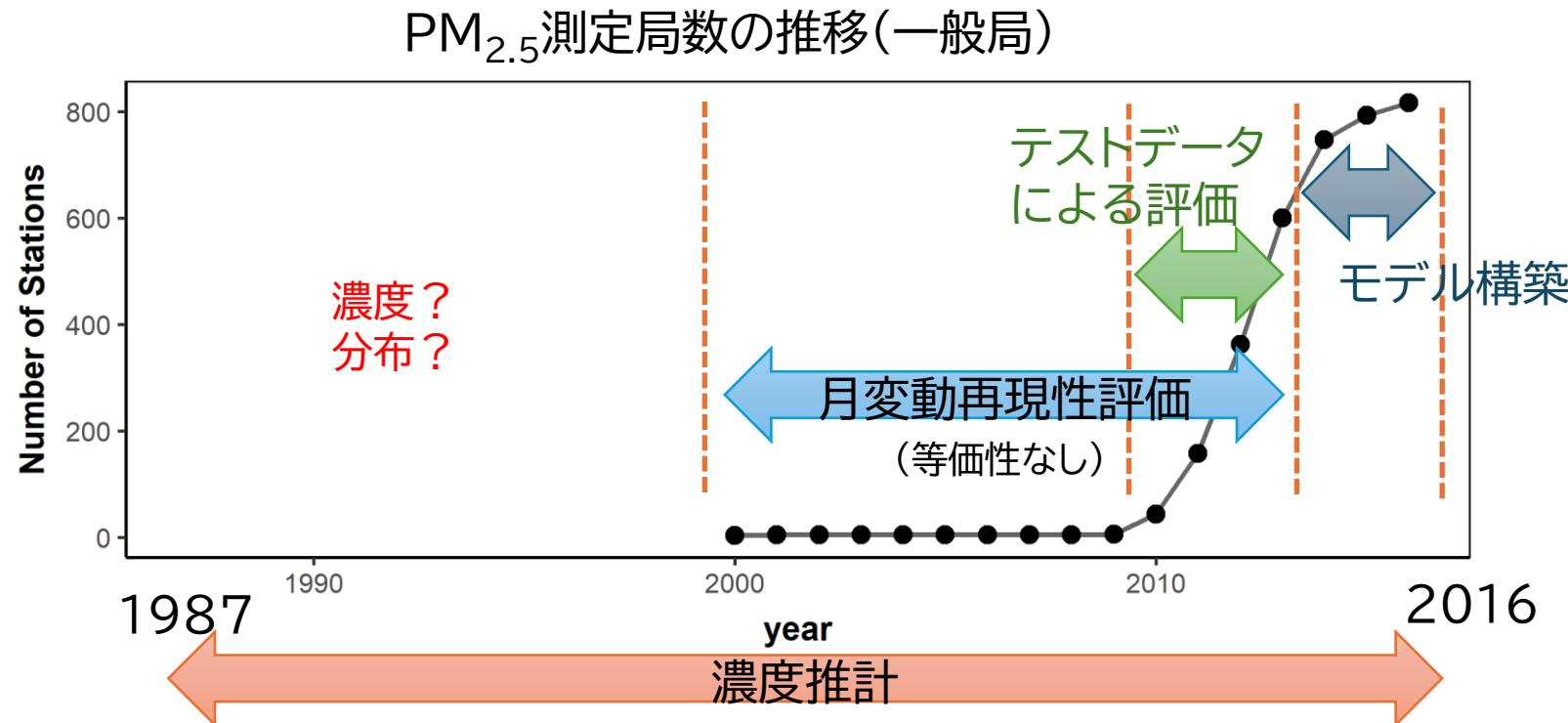


事例1

過去30年間の日本全国PM_{2.5}濃度推計

過去30年間の日本全国PM_{2.5}濃度推計

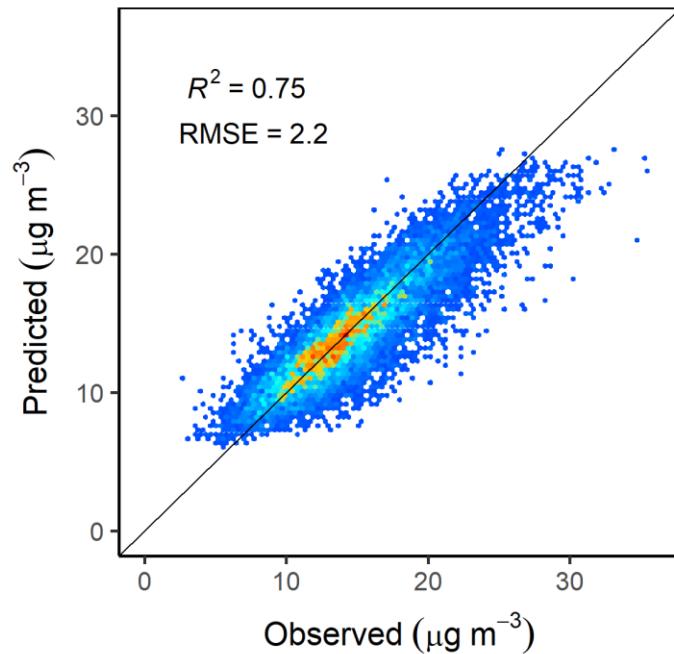
- 観測体制は2010年以降に整備(それ以前の状況はほぼ不明)
- 日本全国を対象に、PM_{2.5}濃度を月単位で1×1kmで推計
- 推計期間:1987 – 2016 年(30年)
- 時間的に外挿するため、詳細な精度評価を実施
- 单層ニューラルネットワーク + model averaging (50 models)



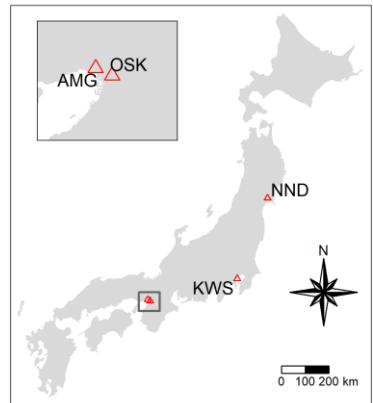
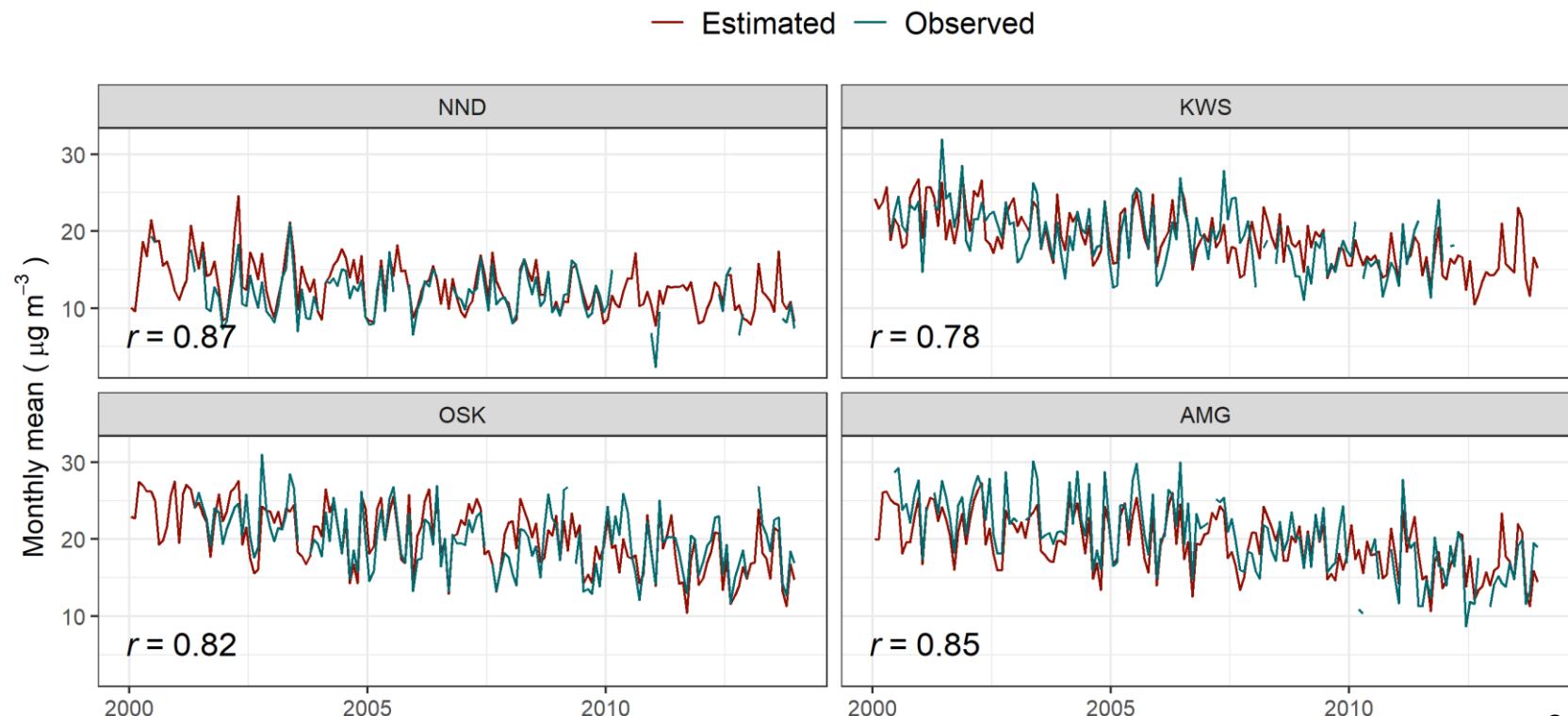
過去30年間の日本全国PM_{2.5}濃度推計：精度評価

- テストデータ(2010-2013)で良好な精度を確認($R^2=0.75$)
- 外部データ(国設局、2000-2013)で良好な月変動再現性を確認

テストデータによる精度評価結果

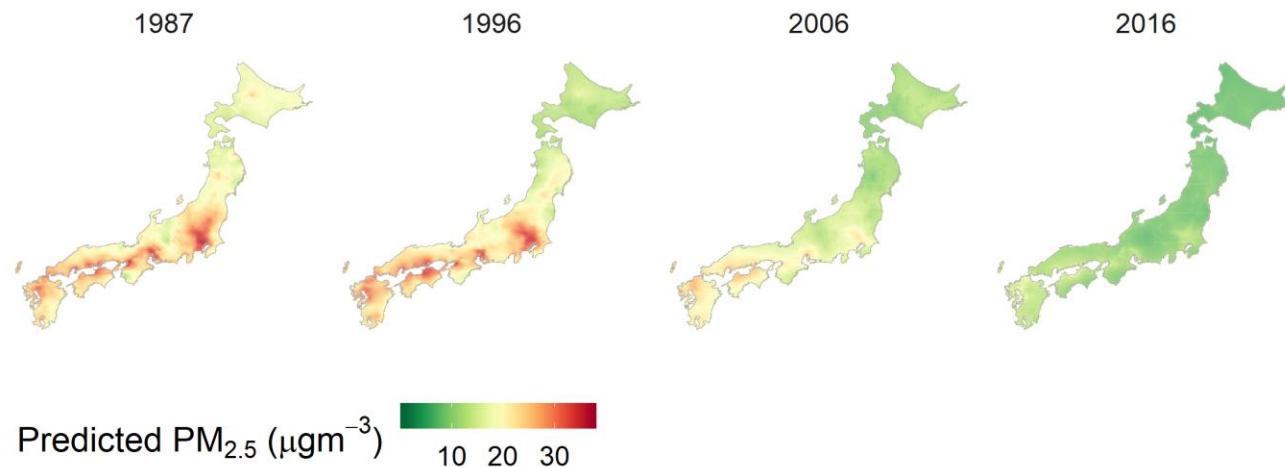


外部データによる月変動再現性評価結果

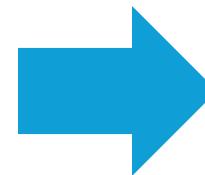
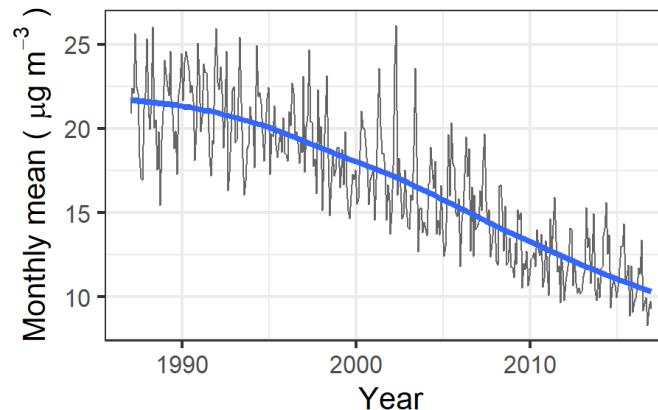


過去30年間の日本全国PM_{2.5}濃度推計:結果と健康影響評価

空間分布(月平均→年平均)

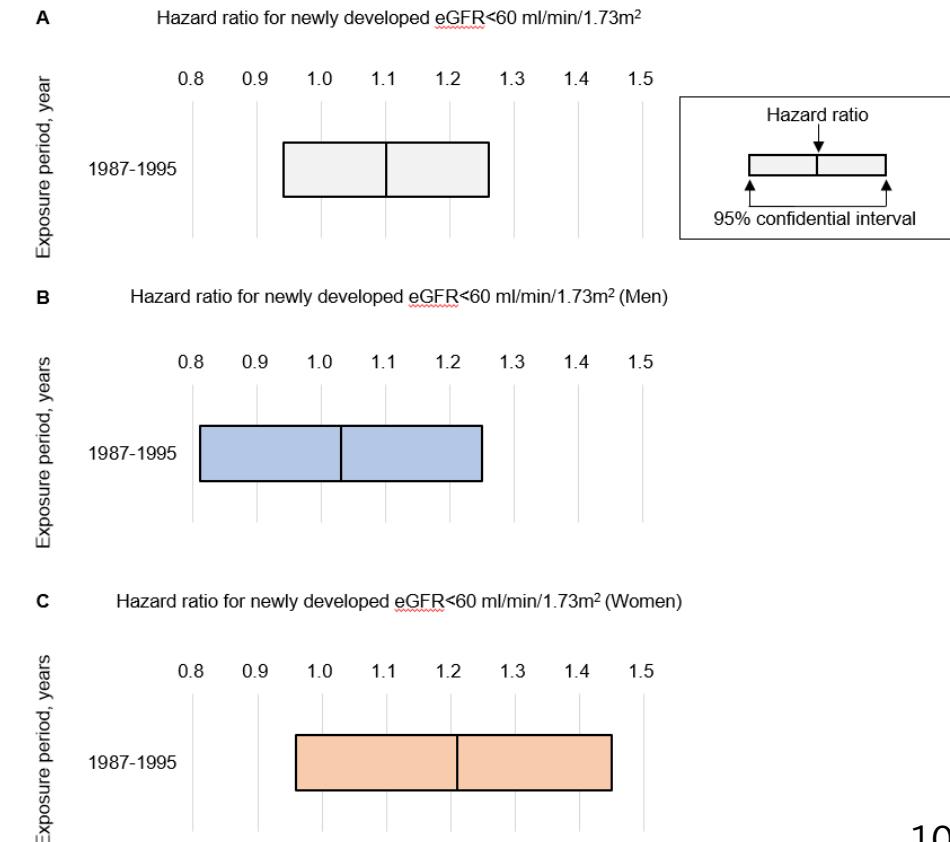


経年変化(月平均→域内平均)



- ・茨城県の市町村単位の健康診断個人データと紐づけ
- ・PM_{2.5}と慢性腎臓病との関連を評価
→ PM_{2.5}は有意な危険因子ではない

Nagai et al. JMA J. 2024.

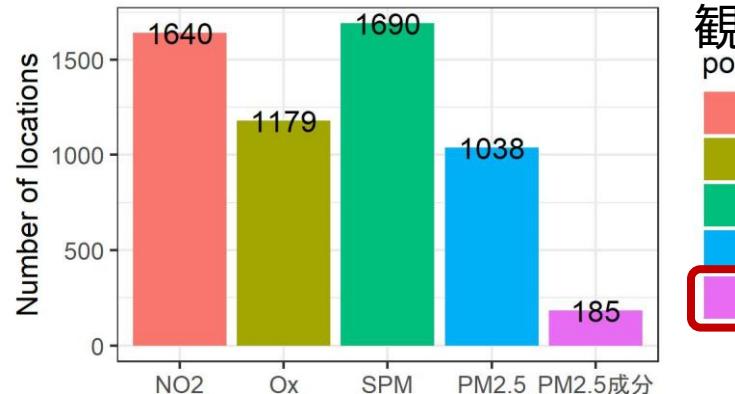


事例2

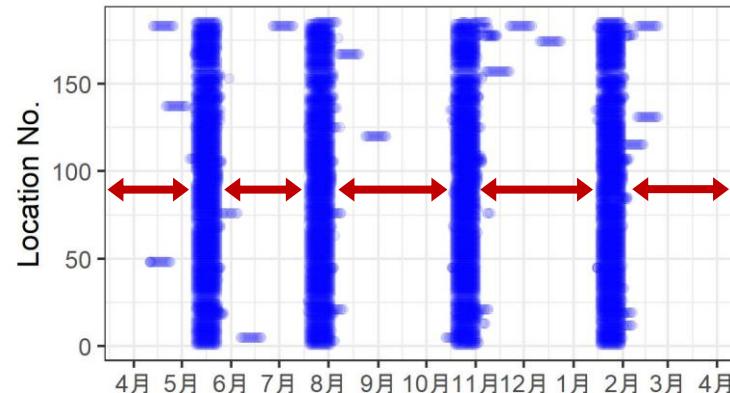
関西地方のPM_{2.5}成分濃度および質量濃度の推計

関西地方のPM_{2.5}成分および質量濃度推計

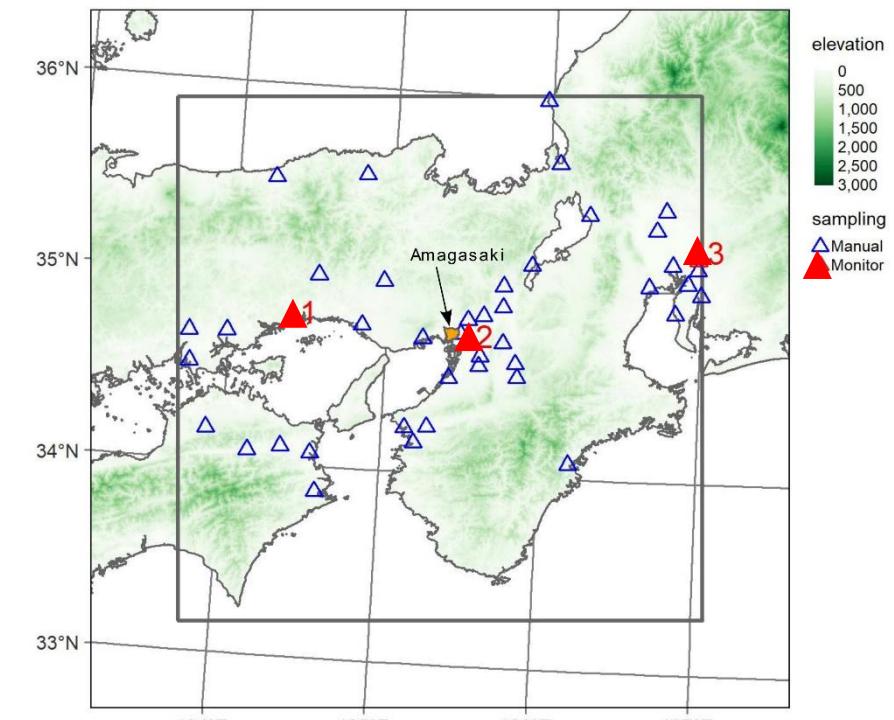
- PM_{2.5}中の硫酸イオン(SO₄²⁻)、硝酸イオン(NO₃⁻)、アンモニウムイオン(NH₄⁺)、元素状炭素(EC)、有機炭素(OC)成分濃度およびPM_{2.5}質量濃度を日単位、1km格子で推計(2010 - 2019年) → 健康影響評価
- 成分観測データは空間・時間ともに限られる(課題)
- 日変動を表す説明変数が必須 → 機械学習モデルとプロセスモデルとの融合(説明変数として導入)



観測地点数(2017年度)

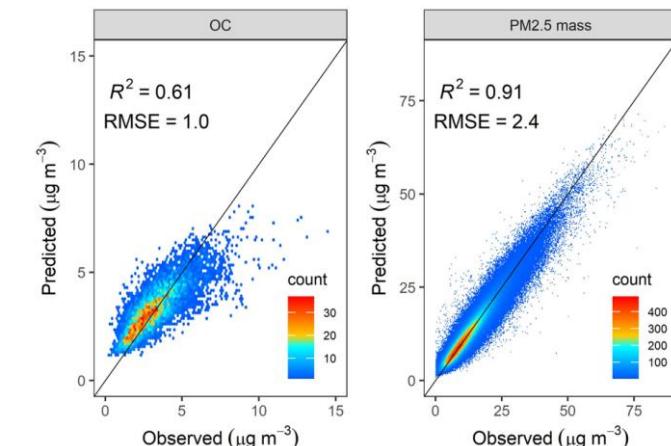
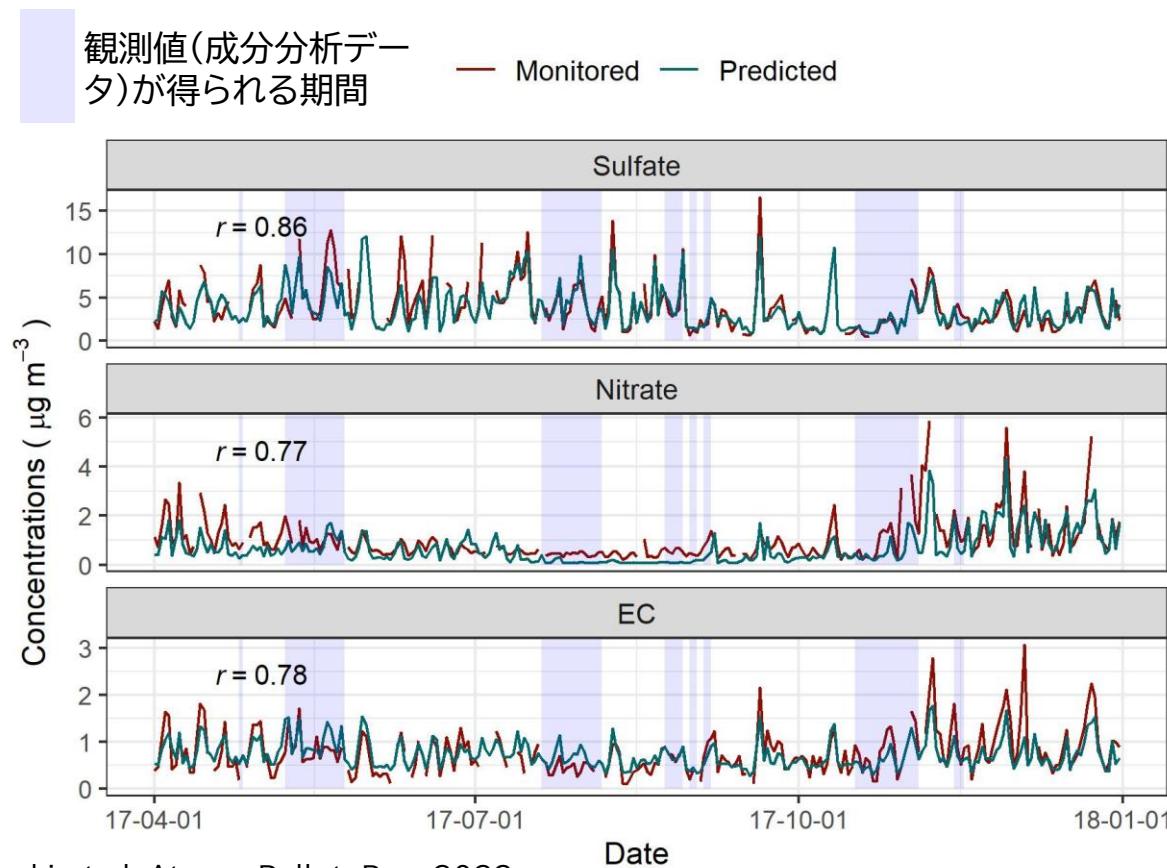


地点ごとの成分測定実施日
(2017年度)

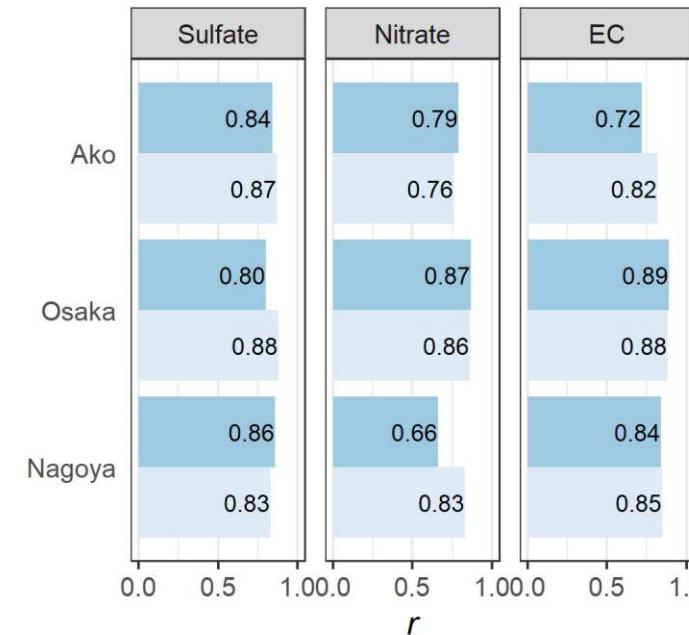


関西地方のPM_{2.5}成分および質量濃度推計:精度評価

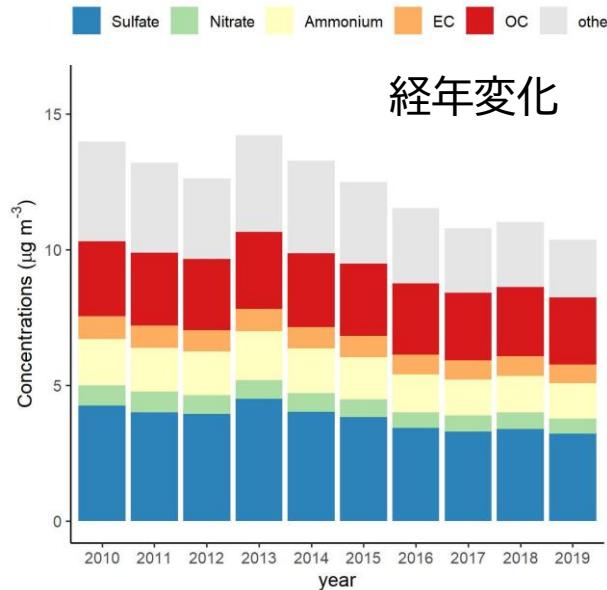
- ◆ 概ね良好な予測精度(右上図: $R^2=0.61 - 0.91$)
- ◆ 日変動を良好に再現することを確認(左図:成分自動測定器との比較)
- ◆ 観測データが得られない期間でも予測精度は低下しない
(右下図:観測データの有無による時間変動再現性)



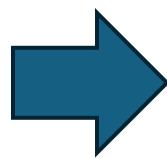
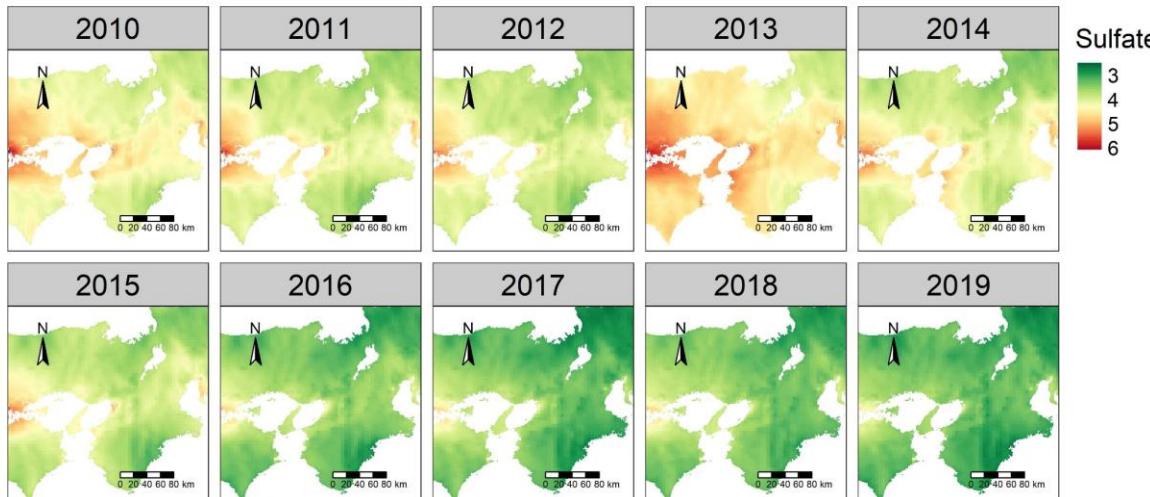
With model building data No model building data



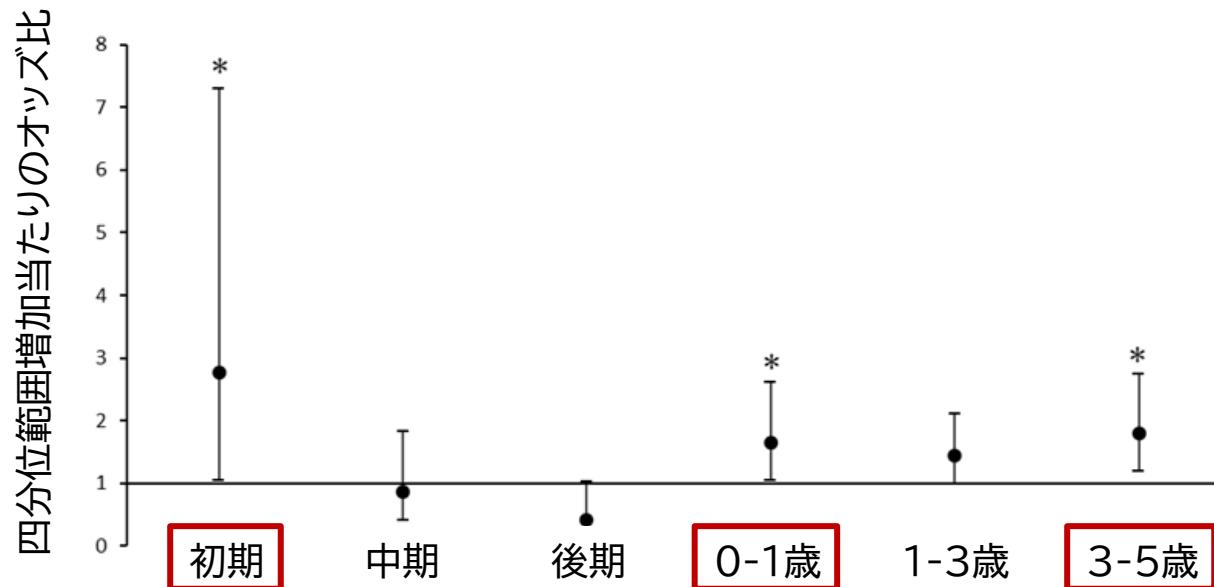
関西地方のPM_{2.5}成分および質量濃度推計: 健康影響評価



空間分布

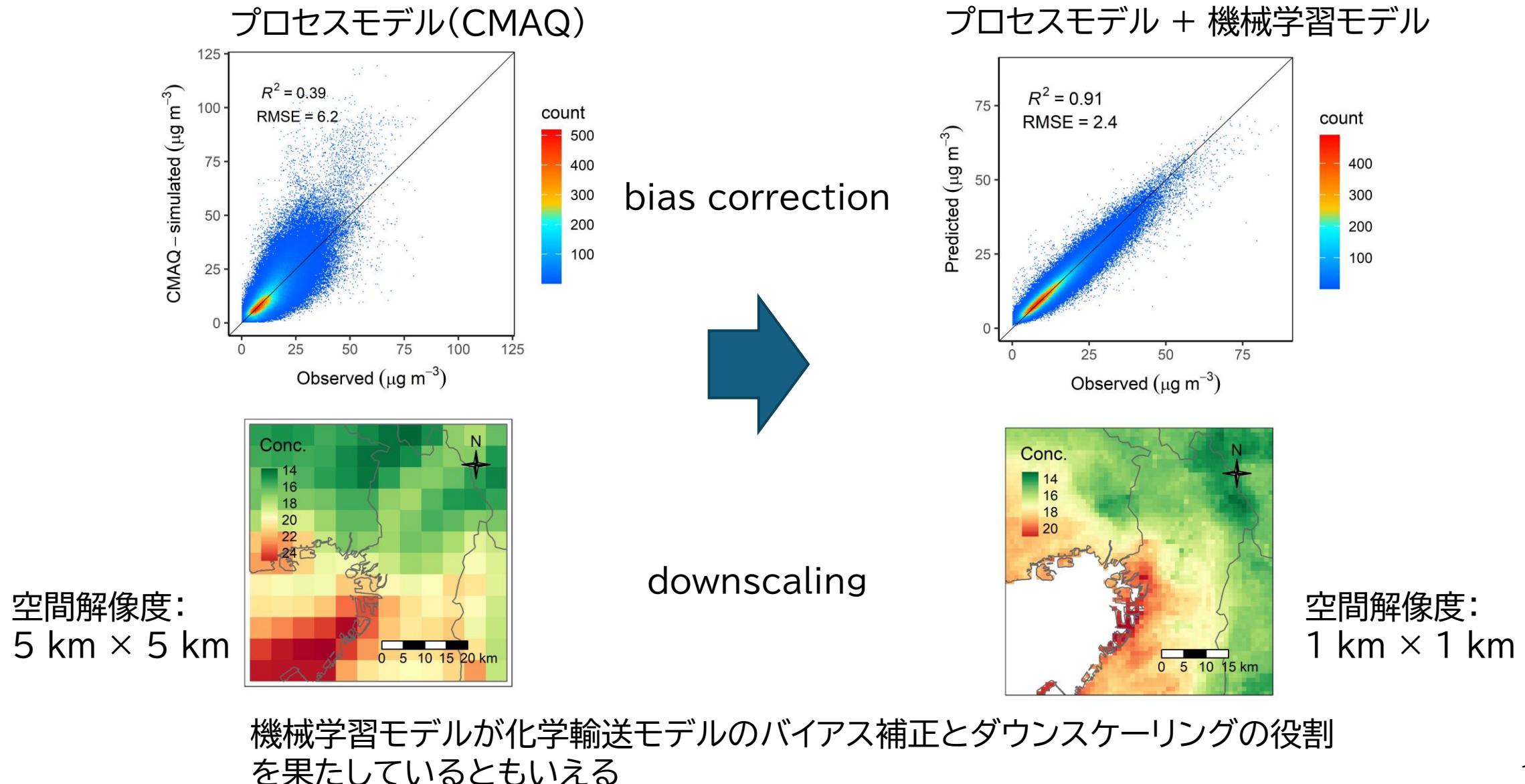


- 対象者の住所情報から曝露濃度算定し健康影響評価
- 妊娠第1期及び出生後0-1年、3-5年の屋外PM_{2.5}へのばく露量推計値が高いことと**子どもの外向的な問題行動**の多さが関連(下図)
- 妊娠中および幼児期のNO₃⁻、EC、OCへの曝露は**喘鳴**と関連



屋外PM_{2.5}ばく露量推計値と子どもが6歳時点での外向的な問題行動との関連

プロセスモデルから見た機械学習モデルの役割

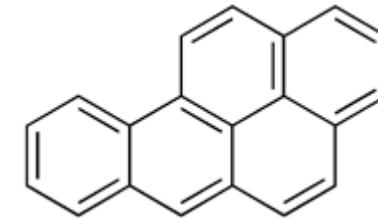


事例3

ベンゾ[a]ピレンの長期濃度推計

ベンゾ[a]ピレンの長期濃度推計

- ベンゾ[a]ピレン(BaP)は5つのベンゼン環が結合した多環芳香族炭化水素(PAHs)
- 化石燃料やバイオマスの燃焼によって非意図的に大気中に排出
- ヒトに対して発がん性があるとされている
- 有害大気汚染物質のうち、優先取組物質



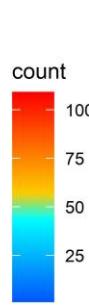
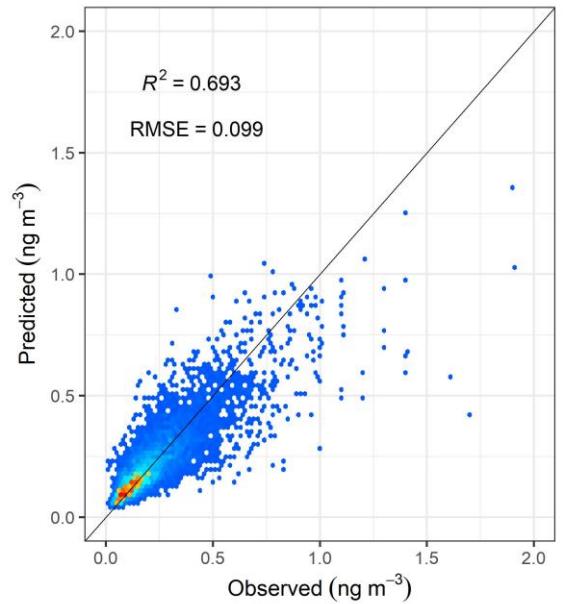
- 健康影響は労働環境における曝露によって評価されているが、その濃度は一般環境中の濃度よりも数桁高い
- 一般環境の濃度レベルでの健康影響評価には、長期間のBaP空間変動の推計が必要
- 大気常時監視として全国の自治体において測定



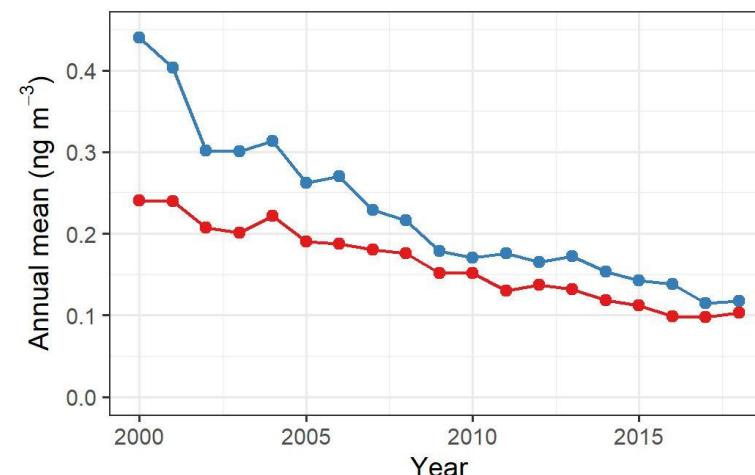
日本全国の2000 – 2018年のBaP濃度を年ごと、1 kmごとに機械学習モデルで推計、曝露濃度等の変動を得る

ベンゾ[a]ピレンの長期濃度推計: 精度評価と予測結果

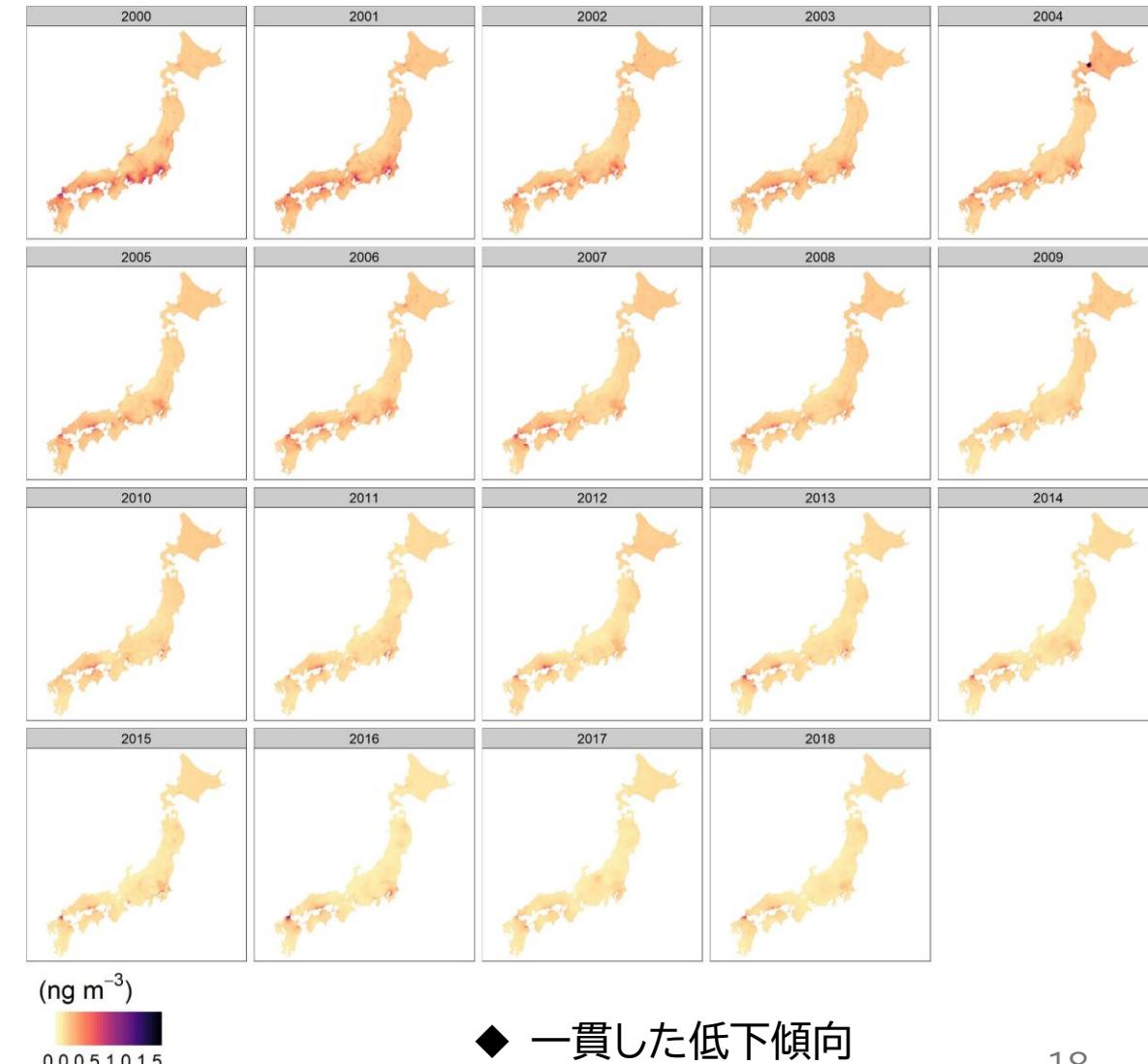
精度評価結果



BaP濃度の経年変化
● Area-weighted ● Population-weighted

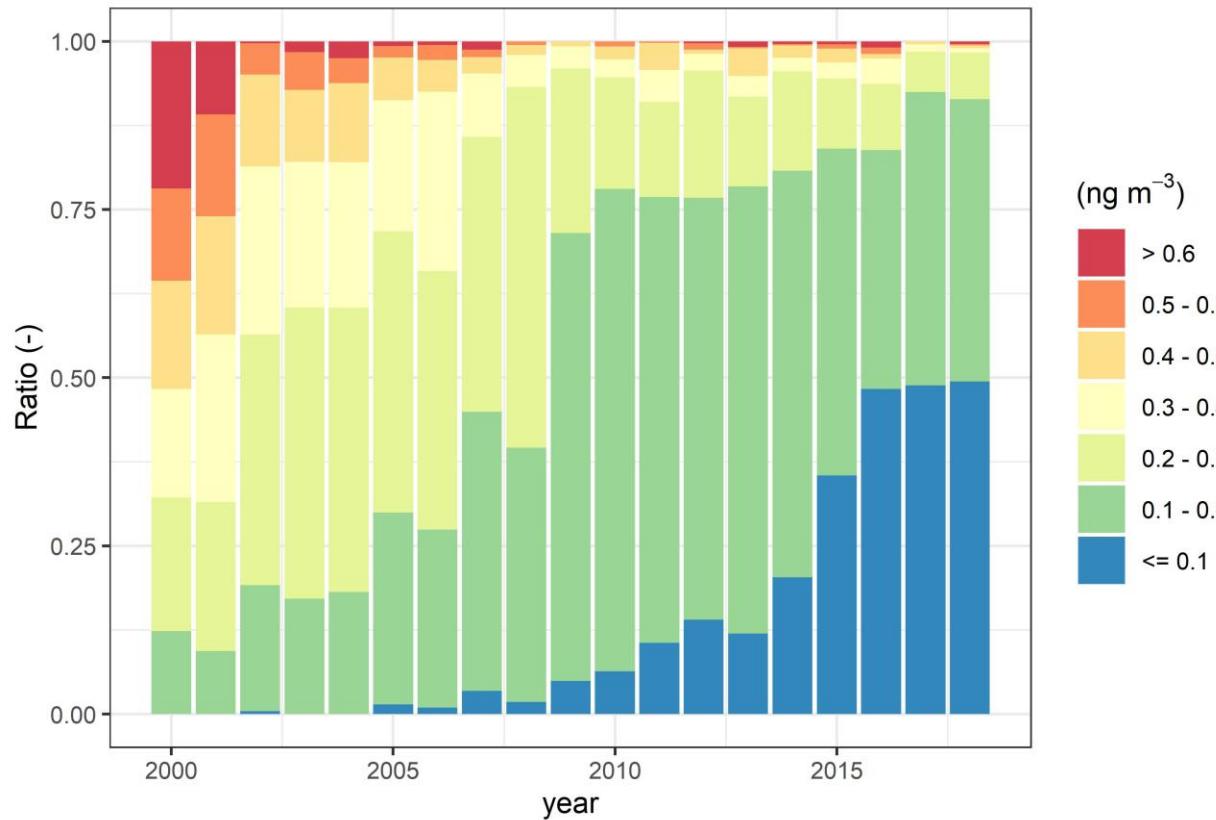


BaP濃度空間分布(2000 - 2018)

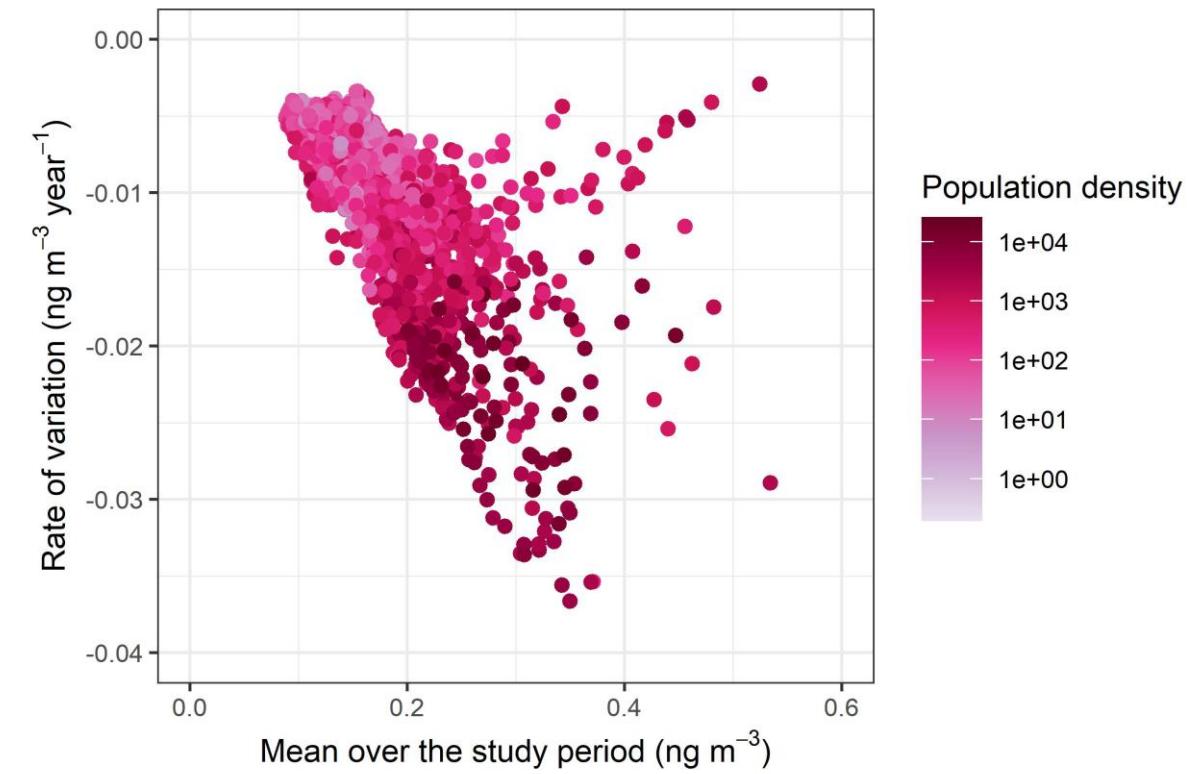


ベンゾ[a]ピレンの長期濃度推計: 予測結果

BaP曝露濃度の経年変化



BaP平均濃度と濃度減少率の関係



- ◆ 一貫した低下傾向
- ◆ 都市部で顕著に低下
- ◆ 大気汚染対策の結果

まとめ

- 機械学習モデルによる大気質予測
 - 基本的なアプローチ
 - 使用する環境データ
- 適用事例の紹介
 - 観測開始前のPM2.5濃度の推計(月単位)
 - 時間的・空間的に観測データが疎なPM2.5成分濃度の推計(日単位)
 - BaPの長期濃度推計(年単位)
- この手法は、目的によって、様々な時間・空間スケールで適用可能
- 適切なデータの利用(とデータの質)が重要

ご清聴ありがとうございました

ご質問等は、下記までお願いします。
araki@ea.see.eng.osaka-u.ac.jp

本研究はJSPS 科研費JP18H03060、JP19K12370、JP21H03205、JP22H03757、
環境研究総合推進費JPMEERF20185002、JPMEERF20195055、
JPMEERF20215005、JPMEERF20225M03、および住友財団環境研究助成の助成を受
けて実施されました。