

B-23 周辺測定局の速報値を利用した PM2.5の統計的予測システムの構築

○神野 有生^{1*}・田中 陽二²・関根 雅彦¹
・今井 剛¹・樋口 隆哉¹・山本 浩一¹

¹山口大学工学部社会建設工学科（〒755-8611 山口県宇部市常盤台2-16-1）

²東京都市大学工学部都市工学科（〒158-8557 東京都世田谷区玉堤1-28-1）

* E-mail: kanno@yamaguchi-u.ac.jp

1. はじめに

PM2.5（微小粒子状物質）は近年、疫学的知見の蓄積、環境基準の制定、測定体制の整備、メディアの報道を背景に、国民の強い関心を集めている¹⁾。現在、700以上の大気汚染常時監視測定局における1時間値（毎正時の前1時間の平均値）が、速報値として毎時WEB配信されている。一方、PM2.5が多い時に、自治体等が推奨する対策（例：屋外での長時間の激しい運動や外出をできるだけ減らす）を講じるためには、事前の予測も重要である。

PM2.5の予測として、九州大学²⁾・国立環境研究所・日本気象協会が、流体力学方程式・化学反応式などに基づく数値予測を毎日更新し、WEB配信している。しかし現状では、初期値の更新や時々刻々の値の計算にPM2.5の速報値が利用されておらず、定量的な予測は困難であるとされている³⁾。また、多くの自治体では1日2回、管内における前数時間の速報値を用いた日平均値の経験的予測・注意喚起を行っているが、これは原理的に、偏西風等によるPM2.5の移流を考慮できない。

PM2.5の移流を考慮しつつ、最新の速報値を反映できる予測として、周辺の広い地域におけるPM2.5・風速等の最新の速報値を利用した、経験式による予測が考えられる。そこで本研究では、山口県宇部市の「宇部総合庁舎」測定局を例にこのような予測式を構築し、毎時更新の予測システムとして試験運用を行った。

2. 使用データ

データ項目について、予測対象の変数（目的変数）は、直近の正時から3, 6, 9, 12, 15, 18, 24時間後までのPM2.5の1時間値の変化量である。予測に用いる変数（説明変数）

の候補は、山口・福岡・長崎県内の5測定局のPM2.5、（PM2.5の測定期間が不十分な）9測定局のSPM（粒子状物質）、4測定局の風速とその東方・北方成分、1測定局の雨量、および時刻である。全て、毎正時に定義される変数であり、PM2.5, SPMに関しては1時間値である。

データの期間は、2011 - 2012年度と2013年度の約半分である。予測対象時刻別に、目的変数の異常値の除去、全変数の標準化（平均を引いて標準偏差で割る操作）を行い、2011 - 2012年度のデータを予測式の構築用、2013年度のデータを検証用とした。さらに検証用データは、経験分布関数が構築用データとほぼ等しくなるようにリサンプリングしてから用いた。

3. 予測式の構築方法

予測式は予測対象時刻別に、構築用データを用いて、遺伝的プログラミング(GP)により構築した。GPを選んだ理由は、数式表現可能な予測式を、式形を事前に定めず柔軟に構築するためである。目的関数は一般的な残差平方和ではなく、 $5\mu\text{g}/\text{m}^3$ 以下の残差を無視した残差平方和に、式の大きさと説明変数の数に関する罰則項を加えたものとした。細かい残差の無視は、大きな残差を重点的に抑えるため、罰則項の追加は、単純で構造・挙動が把握しやすく、汎化性能に優れた式を得るためである。

4. 予測式の構築結果

(1) 予測に用いられた変数

いずれかの予測対象時刻の予測式に、説明変数として採用された変数を図-1に示す。宇部総合庁舎から約30km

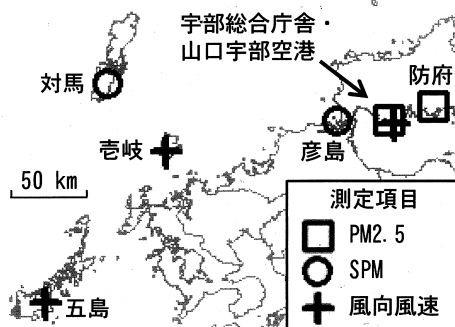


図-1. 予測に用いられる測定局の位置と測定項目

と近い防府・彦島のPM2.5, SPMは、9時間後までの全ての予測式の中に含まれていた。また、宇部総合庁舎から西に約150km離れた壱岐の風速東方成分と、約260km離れた五島の風速は、12時間後以降の全ての予測式の中に含まれていた。つまり、予測対象時刻が遠いほど、より遠くの測定局の情報が予測に使われる傾向があり、予測式がPM2.5の移流を捉えていることが示唆される。

一方、宇部総合庁舎自体のPM2.5と、約4kmしか離れていない山口宇部空港の風速情報（風速または風速東方成分）、および西に約180km離れた対馬のSPMは、全ての予測対象時刻の予測式に含まれていた。この理由は不明であるが、宇部総合庁舎のPM2.5については、「頭打ち」の効果（現在の値が大きければ今後減少しやすい；裏も真）が考えられる。実際、全ての予測対象時刻の構築用データについて、宇部総合庁舎のPM2.5は、図-1に示した他の変数と比べ、目的変数の間の相関が強かった（相関係数0.54～0.28）。

(2) 予測式の構造

得られた予測式の例として、12時間後の予測式を示す。

$$\begin{aligned} & \text{PM2.5}_{\text{宇部総合庁舎 (12時間後)}} - \text{PM2.5}_{\text{宇部総合庁舎}} \\ &= 0.145 \times \{ \text{風速東方成分}_{\text{壱岐}} - \text{風速}_{\text{山口宇部空港}} - \text{風速}_{\text{五島}} \\ &+ (0.307 - \text{SPM}_{\text{対馬}}) \times (\text{風速}_{\text{五島}} < \text{風速}_{\text{山口宇部空港}}) \\ &+ \text{PM2.5}_{\text{宇部総合庁舎}} \times \min(0.772, 0.517 \times \text{SPM}_{\text{対馬}} - \text{風速}_{\text{五島}}) \\ &+ \text{SPM}_{\text{対馬}} - 4.43 \times \text{PM2.5}_{\text{宇部総合庁舎}} \} \quad \dots \textcircled{1} \end{aligned}$$

ここに、例えばPM2.5_{宇部総合庁舎}は、宇部総合庁舎におけるPM2.5の速報値を示す。不等号<は、不等式が真のとき1を、偽のとき0を返す演算子である。またminは、括弧内の2つの数値のうち大きくない方を返す関数である。この予測式は、全変数を標準化した状態での予測式である。

式①によれば、西に位置する壱岐の風速東方成分が増加するほど、先12時間に宇部総合庁舎のPM2.5が増加する。構築用データについて、壱岐のSPMは宇部総合庁舎のSPMより、壱岐の風速東方成分に依らず大きい傾向が

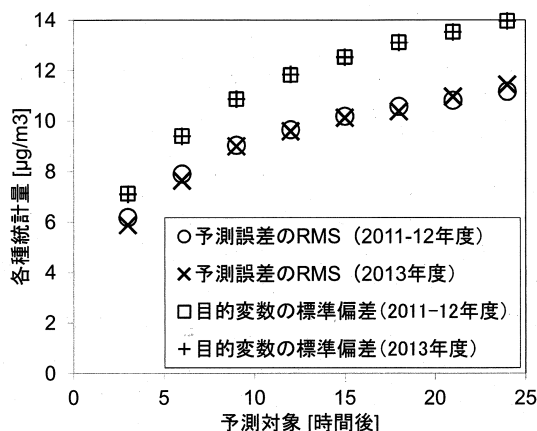


図-2. 目的変数（予測対象時刻までのPM2.5の1時間値の変化量）に関する、標準偏差と、予測誤差のRMS。標準偏差は平均値で予測した時の予測誤差のRMSである。

あった（平均値の差は5.3[μg/m³]）ため、これは偏西風による輸送の影響を表していると考えられる。式①で対馬のSPMが大きいくほど目的変数が増加することも、同様に解釈できる。さらに、式①において、宇部総合庁舎におけるPM2.5が増加するほど目的変数が減少することは、前述の「頭打ち」の効果を表していると考えられる。

このように、予測式に含まれた説明変数の一部については、式中の役割の気象学的または統計学的な解釈が可能であった。一方、式①における五島や山口宇部空港の風速のように、目的変数との関係が複雑で、役割の解釈が困難な説明変数もあった。

(2) 予測誤差のRMS

本研究の予測式構築では特殊な目的関数を用いたが、以下、解りやすさのため、一般的な誤差統計量を用いる。

図-2に示すように、予測誤差（予測値と実測値の差）のRMS（2乗平均平方根）は、構築用・検証用のいずれのデータについても、予測対象時刻が遠くなるほど増加した。一方、目的変数の標準偏差（最適な定数つまり平均値で予測した場合の予測誤差のRMS）も同様に増加し、両者の比はむしろ減少傾向を示した。しかし、いずれのデータ・予測対象時刻についても、この比が0.79を下回ることがなかった。すなわち、予測誤差のRMSについて、得られた予測式と、平均値で予測する場合との差は、高々21%である。

なお、図-2より、目的変数の予測誤差のRMSは、予測式の構築用・検証用の両データについて、同等である。これは、モデルの構築において過適合が生じておらず、予測式の汎化性能には問題がないことを示唆している。

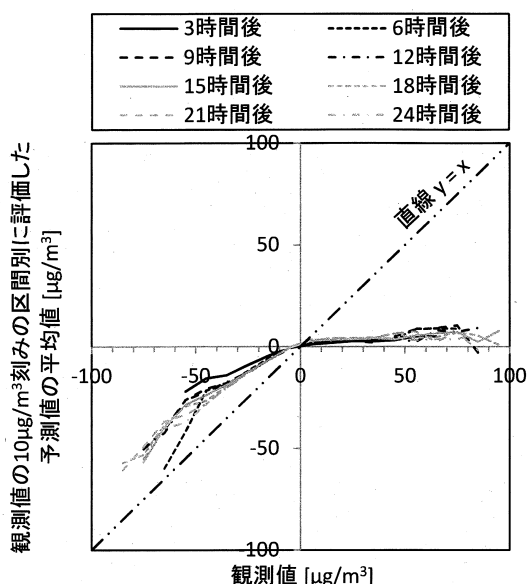


図3. 各予測対象時刻について、測定値を $10\mu\text{g}/\text{m}^3$ 刻みに分けて評価した、構築用データに関する予測値（あてはめ値）の平均値。

(3) 予測のバイアス

予測のバイアスに関する統計量として、予測対象時刻別、構築用・検証用の別に評価した予測誤差の平均値は、いずれも $0.84\mu\text{g}/\text{m}^3$ を超えず、測定の最小単位 $1\mu\text{g}/\text{m}^3$ を下回った。しかし図3のように、測定値を $10\mu\text{g}/\text{m}^3$ 刻みの区間に分けて考えると、いずれの予測対象時刻についても、測定値が負の区間で過大評価、正の区間で過小評価のバイアスが検出された。特に正の区間では、測定値が大きいほどバイアスも大きくなった。つまり、得られた予測式は、PM2.5の増加を十分に予測出来ず、増加量が大きいほど深刻な過小評価をする傾向がある。図4に例示した予測値と測定値の散布図でも、この傾向が確認できる。

5. 予測システムの実装と試験運用

得られた予測式を用いた、宇部総合庁舎におけるPM2.5の予測システムを、山口大学のサーバーに実装し、2014年6月10日に試験運用を開始した。このシステムは1時間に1回、説明変数の最新の速報値を各自治体のWEBサイトからダウンロードし、3, 6, 9, 12, 15, 18, 24時間後の予測式に代入する。さらに、宇部総合庁舎におけるPM2.5の最新の速報値を加えることで、各予測対象時刻の1時間値を予測し、表とグラフの形式でWEB配信する。

予測値は、2014年9月12日現在まで、停電・欠測などによる一時的な中断を除いて継続的に更新されており、このようなシステムが運用可能であることが示された。

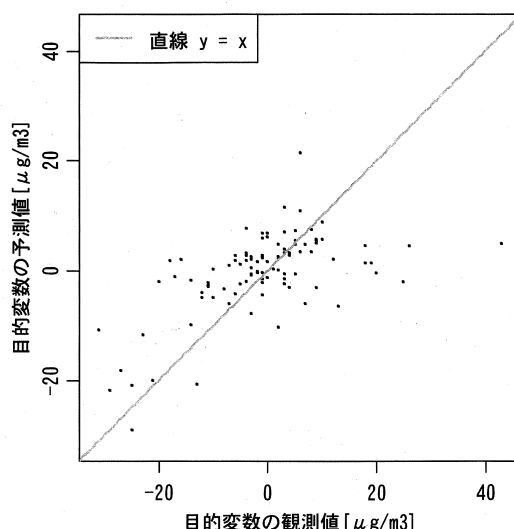


図4. 12時間後の予測式の目的変数に関する予測値と測定値の散布図。見やすさのため、単純無作為抽出した100点のみを示す。

6. おわりに

本予測システムは、PM2.5の移流を考慮しつつ、最新情報を反映した高頻度の更新が出来るという、従来にない特徴を持つ。しかし、予測の確度・精度については、予測誤差のRMSが大きいのみならず、PM2.5の急上昇を過小評価するバイアスがあり、実用的なレベルにはない。今後の改善の余地として、次の点が挙げられる。

1. 本研究では、PM2.5の測定歴が浅い測定局で代わりにSPMを用いた。データの蓄積を待てば、PM2.5のみを用いたより直接的な予測が可能となる。
2. 本研究では、予測式を構造・挙動が把握しやすい数式として得るためにGPを用いた。より柔軟な回帰モデルを使えば、確度・精度は改善しうる。

謝辞

使用データの一部は、山口県・福岡市・長崎県から提供頂いた。また一部は、環境省および気象庁のWEBサイトから入手した。使用データは未確定の速報値を含むこと、提供機関は本研究に関与しておらず一切の責任がないことを明記する。

参考文献

- 1) 微小粒子状物質 (PM2.5) に関する専門家会合：最近の微小粒子状物質 (PM2.5) による大気汚染への対応、環境省、2013。
- 2) 竹村俊彦：大気エアロゾル予測システムの開発、天気、56, pp. 455-461, 2009。