

3) Regression Kriging 法を用いた九州地域の PM_{2.5} の空間濃度分布推定

小原大翼 古澤尚英*1 豊永悟史*2 荒木真*3 山本裕典*4 山崎文雅 矢野弘道*5

要 旨

九州地域で 2014 年度から 2017 年度に観測された PM_{2.5} 日平均値に空間濃度分布予測手法である Regression Kriging 法（以下、「RK 法」）を適用した。一般的に用いられる交差検証法を用いて予測精度を評価した結果、構築した RK 法の予測精度は良好であることが示された。

キーワード：PM_{2.5}、Regression Kriging 法、空間濃度分布推定、予測精度評価

はじめに

大気汚染物質の常時監視は、大気汚染防止法に基づいて全国の自治体を実施している。微小粒子状物質 (PM_{2.5}) については 2009 年 9 月に環境基準が制定され、2010 年度以降に全国的な監視体制の整備が進められた結果、2017 年度時点では 800 以上の測定局が配置されている¹⁾。これは、PM_{2.5} の社会問題化への対応として進められた側面もあり、実際に熊本県では、国が示す事務の処理基準に基づく必要局数よりも 5 局多い 19 の一般大気環境測定局を配備し監視を行っている。

一方、PM_{2.5} 濃度は近年減少傾向にあり、全国の一般大気環境測定局の環境基準達成率は、16% (2013 年度) から 90% (2017 年度) と大きく改善している^{1)・2)}。熊本県内でも 9% (2013 年度) から 72% (2017 年度) と改善している^{3)・4)}。このような大気汚染状況の変化を鑑みれば、測定局の配置の適正化について検討を行うことが今後の課題になると考えられる。

このためには、大気汚染物質の空間濃度分布を適切に把握することが必要である。大気汚染物質の空間濃度分布予測手法の中で近年用いられている RK 法は、Land Use Regression (以下、「LUR」) 法と呼ばれる線形回帰モデルと観測値間の残差に空間補間法である Ordinary Kriging (以下、「OK」) 法を適用するものである⁵⁾。RK 法は、原理的に観測データに限られる場合も高精度での予測が可能であり、数 km 以下の空間解像度を得ることができる。また、RK 法と LUR 法の予測値を比較した研究では、前者の予測精度が良好であると報告されている⁶⁻⁸⁾。RK 法の予測は説明変数の選択などの経験則的

な要素が含まれるため、予測する領域に応じた条件の設定が重要となる。これまで日本国内では、全国スケールでの予測が行われた事例⁹⁾はあるものの、九州地域を対象とした事例は確認できない。

以上のことから、本研究では、PM_{2.5} 日平均値を用いた RK 法により九州地域の空間濃度分布を予測し、その精度を検証した。予測精度の評価は、通常用いられる交差検証法を用いて行った。

方 法

1 空間濃度分布の予測方法

RK 法では、PM_{2.5} 濃度は以下に示す 2 つの成分の和として表現される (式 1)。

$$C = C_1 + C_2 \quad (\text{式 1})$$

式 1 において、 C は、PM_{2.5} 濃度の観測値、 C_1 は、観測値において環境変数によって決定される決定論的成分、 C_2 は決定論的成分では説明できない確率論的成分を示す。さらに、式 1 の決定論的成分 C_1 は、LUR 法によって以下のとおり表すことができる。

$$C_1 = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \dots + \beta_n x_n + \beta_0 \quad (\text{式 2})$$

式 2 において、 x は土地利用や気象条件などの各説明変数の値、 $\beta_1 \sim \beta_n$ は各説明変数に対応した係数、 β_0 は切片を示す。式 1 により、確率論的成分 C_2 は、 C と C_1 の差 (以下、「残差」) として算出される。

RK 法では、観測地点ごとに得られた残差に対して OK 法を適用し、その空間濃度分布を予測する。OK 法では、最初に全ての観測地点の残差の対からその非類似度 γ^* を求める。例えば、観測地点 i 及び j における γ^* は以

*1 現球磨地域振興局保健福祉環境部 *2 現阿蘇地域振興局保健福祉環境部 *3 大阪大学工学研究科
*4 現環境センター *5 現環境生活部環境局環境保全課

下のとおり表される¹⁰⁾。

$$\gamma^* = \frac{(Z_i - Z_j)^2}{2} \quad (\text{式 3})$$

ここで、 Z_i 及び Z_j はそれぞれ観測地点 i 及び j の残差を示す。この計算を全ての観測地点のペアについて行い、全ての γ^* と観測地点間の相対距離 h との関係を示したものをバリオグラム雲という。次に、このバリオグラム雲に対して、相対距離 h に対して階級区分 \tilde{h} ごとに細分化し、 γ^* の平均値 (標本バリオグラム) を求める¹¹⁾。

$$\gamma^*(\tilde{h}) = \frac{1}{2N_h} \sum_{k=1}^{N_h} \gamma_k^*(h) \quad (\text{式 4})$$

ここで、 $\gamma^*(\tilde{h})$ が階級区分 \tilde{h} ごとの標本バリオグラム、 N_h は相対距離が h となる組み合わせの個数、 $\gamma_k^*(h)$ が相対距離 h に対する γ^* を示す。標本バリオグラムは、階級区部 \tilde{h} ごとの平均化された γ により表されるが、標本バリオグラムに対して最小二乗法により最も当てはまりが良いバリオグラムモデル¹¹⁾ を選択したものを理論バリオグラムという。なお、標本バリオグラムは、理論バリオグラムの当てはめに用いる最大距離 h_{max} (カットオフ) が規定されており¹²⁾、理論バリオグラムには、バリオグラムが定常状態になる h 、すなわち空間的自己相関がなくなる距離を示すレンジ (range) が規定されている¹¹⁾。この理論バリオグラムを用いることで、観測地点以外の地点における確率論的成分 (C_2)、すなわち残差を予測することができる。グリッドごとに求めた OK 法の出力値と、対応する LUR 法の出力値の和が RK 法の予測値となる。LUR 法で用いる説明変数は、2 節で示すとおりである。OK 法では、予備解析の結果に基づき、カットオフは 100km とし、バリオグラムモデルは、Exp(指数型)、Mat (Matern 型)、Sph (球形型) の 3 種類から最小二乗法により当てはまりの良いものを選択する形とした。本研究では、以上の操作を 1 日ごとに行い、対象期間の 4 年間で $N=1,461$ 日分の予測値が得られた。

本研究のデータ整形及び解析は、統計解析環境 R ver 3.5.1¹³⁾ 及びそのパッケージである sp¹⁴⁾、¹⁵⁾、gstat¹⁶⁾、rgdal¹⁷⁾、rgeos¹⁸⁾、mapproj¹⁹⁾、MASS²⁰⁾ を使用した。

2 Regression Kriging 法の使用データ

計算には、離島に位置する測定局 (長崎県の壱岐測定局、対馬測定局及び五島測定局) を除いた九州地域と山口県の一部を含めた北緯 31.0° から 34.2° 及び東経 128.0° から 132.4° の範囲 (以下、「対象範囲」) に存在する測定局における観測データを使用した (図 1)。各年度の解析対象の測定局数は、2014 年度は 148 局、2015 年度は 149 局、2016 年度は 151 局、2017 年度は 140 局であり、図 1 は対象範囲及び 2017 年度の測定局の位置を示してい

る。

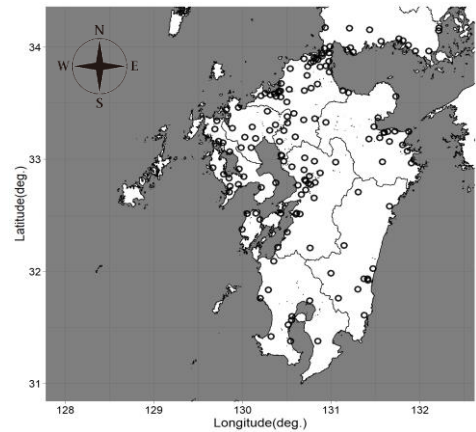


図 1 対象範囲及び測定局の位置

※) 図中○は 2017 年度の測定局地点を示す。

表 1 に、本報で使用したデータを示す。本研究では、対象領域に $1 \times 1 \text{ km}$ のグリッドを作成し予測グリッドとしており、使用データは以降に述べる方法によって予測グリッドに合わせて整形を行った。PM_{2.5} 濃度の観測データは、2014 年度から 2016 年度については、環境数値データベース²¹⁾ から 1 時間値 (確定値) を入手した。2017 年度については、対象範囲に含まれる自治体 (山口県、下関市、福岡県、福岡市、北九州市、佐賀県、長崎県、佐世保市、大分県、宮崎県、鹿児島県) から 1 時間値 (速報値) の提供を受けた。入手した PM_{2.5} 濃度の 1 時間値から、各測定局の日平均値を算出した²²⁾。このうち、日平均値が $200 \mu\text{g}/\text{m}^3$ 以上または $0 \mu\text{g}/\text{m}^3$ 以下となった日 ($N=296$) は異常値として除外したものを解析に使用した。

気象データの作成には、Weather Research and Forecasting (WRF) model ver. 3.8.1²³⁾ の前処理プログラムである WRF Preprocessing System (WPS) を利用した。位置情報の作成を行う geogrid では、Lambert 投影法による $1\text{km} \times 1\text{km}$ の水平解像度を設定した。土地利用には WRF に付随する United States Geological Survey (USGS) のデータを利用した。気象データの取り込みを行う ungrib では、気象データに National Centers for Environmental Prediction (NCEP) の水平解像度 1 度、6 時間ごとの客観解析データ (FNL)²⁴⁾ と気象庁の数値予報 GPV メソ気象モデル (MSM) 再解析データ²⁵⁾ を利用した。このようにして作成した WRF の初期条件及び境界条件から、計算対象領域の緯度 (Deg.)、経度 (Deg.)、地上付近気温 (°C)、地上付近相対湿度 (%), 地上付近気圧 (kPa) 及び風速 (m/s) を取り出して気象データとした。

人口データについては、e-Stat²⁶⁾ における 2015 年の国

勢調査 3 次メッシュ (1km×1km) データから予測グリッドの周囲 1.5 km の平均値を算出し、予測に使用した。

道路延長データ及び標高データについては、国土数値情報²⁷⁾における 2010 年の道路密度・道路延長メッシュ (1 km×1 km) データ及び 2009 年の標高・傾斜度 5 次メッシュ (250 m×250 m) データから予測グリッドの周囲 1.5 km の平均値を算出し、予測に使用した。

土地利用データについては、国土数値情報における 2016 年の土地利用細分メッシュ (100 m×100 m) データから予測グリッド内における土地利用区分の割合 (%) を算出し、さらに PM_{2.5} に対して類似の影響を与える発生源として見なされる変数は合算した。結果として、6 種類の面積比 (農用地、水域、建物用地、自然地、鉄道、道路) データを予測に使用した。

降水量データについては、過去の気象データ・ダウンロード²⁸⁾ から、対象範囲内の AMeDAS で観測された日合計降水量を入手し、OK 法を用いて空間補間を行った。なお、OK 法では、3 種類のバリオグラムモデル (Exp, Mat, Sph) より最もフィッティングが良いモデルを自動選択し、カットオフ=300km に設定した。

LUR 法で使用する説明変数は、以上の入手データから、排出、輸送、変質、沈着といった大気汚染物質の濃度を決定する重要な要素との関連性を考慮して設定した (表 2)。LUR 法の予測は、準備した説明変数から Backward stepwise 法により優位な変数が抽出され²⁹⁾、さらに予め設定した係数の方向と説明変数の係数の正負が一致したものが使用されるように設定した³⁰⁾。

表 1 使用データ

種類	入手先	期間
PM _{2.5} 質量濃度	国立環境研究所 環境数値データベース 時間値データ	2014-2017
気象	Weather Research and Forecasting model ver. 3.8.1 USGS, National Centers for Environmental Prediction FNL, 気象庁 MSM	2014-2017
人口	e-Stat (国勢調査3次メッシュ)	2015
道路延長	国土交通省 国土数値情報 道路密度・道路延長メッシュデータ	2010
標高	国土交通省 国土数値情報 標高・傾斜度5次メッシュデータ	2009
土地利用	国土交通省 国土数値情報 土地利用細分メッシュデータ	2016
降水量	気象庁HP 過去の気象データ・ダウンロード AMeDAS降水量	2014-2017

表 2 使用した説明変数及び説明変数の係数の方向

分類	説明変数	係数の方向	単位
位置	緯度 (Latitude)	±	degree
	経度 (Longitude)	-	degree
排出	道路長	+	km/km ²
	人口	+	千人/km ²
	高度	-	m
	農用地 (その他農用地, その他用地, ゴルフ) 面積比	-	km ² /km ² (%)
	水域 (河川, 海水, 海浜) 面積比	-	km ² /km ² (%)
	建物用地 面積比	+	km ² /km ² (%)
輸送	自然地 (荒地, 森林, 田) 面積比	-	km ² /km ² (%)
	鉄道 面積比	+	km ² /km ² (%)
	道路 面積比	+	km ² /km ² (%)
	風速	±	m/sec
変質	湿度	±	None(%)
	気圧	+	kPa
沈着	降水量	-	mm/h

3 予測値の精度評価

予測値の精度評価は、多くの研究でモデルの評価に用いられている一個抜き交差検証法を使用した。この手法は、1 地点の観測データのみを除外したデータセットを用いて RK 法による予測値を求める操作を全地点について繰り返し、除外した地点の予測値と観測値を比較する手法である¹⁵⁾。比較した結果は、決定係数 (R²) 及び予測値と測定値の二乗平均誤差 (RMSE) の 2 つの指標値によって評価した。これらは以下のように定義される。

$$R^2 = 1 - \frac{\sum_{l=1}^n (z_l - z_m)^2}{\sum_{l=1}^n (z_l - \hat{z}_l)^2}$$

$$RMSE = \sqrt{\frac{1}{n} \sum (z_l - z_m)^2}$$

ここで、z_lは測定値、z_mは測定値に対応する予測値、 \hat{z}_l は測定値の平均値、n はデータ数である。

4 結果と考察

図 2 に、PM_{2.5} 日平均値と対応する交差検証法により算出された予測値の散布図を示し、図中には同法による指標値の結果を示した。カナダのモントリオールで観測された PM_{2.5} の日平均値を用いた事例では、R²:0.46-0.89,

RMSE:2.36-3.34と報告されており³¹⁾、本報の指標値(R^2 : 0.90, RMSE: 2.49)はこれと同程度と良好な予測精度であった。このことから、RK法はPM_{2.5}日平均値を使用した空間濃度分布予測手法として活用可能であることが示唆された。

今後、PM_{2.5}日平均値に加えて年平均値及び高濃度平均値(高濃度日(日平均値35 $\mu\text{g}/\text{m}^3$ を超える県内の測定局が存在する日)の年間の平均値)についても同法を適用し、予測精度を検証する予定である。

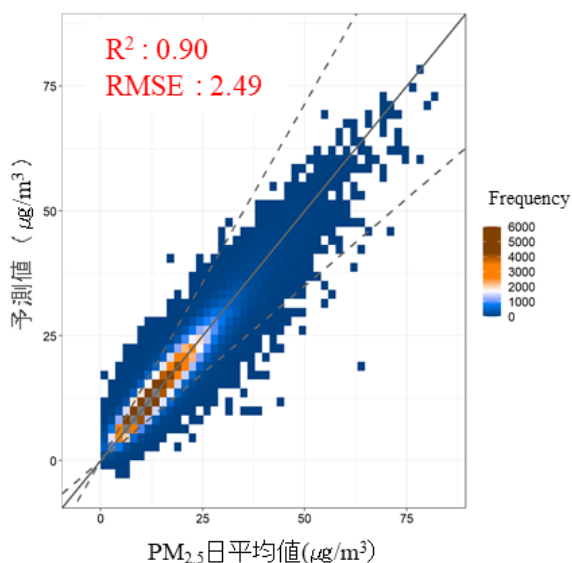


図2 PM_{2.5}日平均値と対応する交差検証法により算出された予測値の散布図

※) 実線は1:1の関係を示し、点線は0.7:1及び1:0.7の関係を示す。

まとめ

九州地域を対象に2014年度から2017年度におけるPM_{2.5}の空間濃度分布について、PM_{2.5}日平均値を用いたRK法により予測した。

交差検証法を用いた精度評価結果では、指標値が先行研究と同程度となり、良好な予測精度を示すことが確認された。

謝辞

本研究を行うにあたり、九州大学応用力学研究所教授の鶴野伊津志氏、熊本県立大学教授の張代洲氏、高崎経済大学教授の飯島明宏氏及び熊本地方気象台の西辻和也氏と篠崎覚氏には、本研究に関する有益な助言を多くいただきました。深く感謝いたします。また、福岡県環境保全課、佐賀県県民環境部環境課、長崎県環境部地域環境課、大分県生活環境部環境保全課、宮崎県環境森林

部環境管理課、鹿児島県環境林務部環境保全課、山口県環境生活部環境政策課、北九州市環境局環境監視部環境監視課、福岡市環境局環境監理部環境保全課、下関市環境部環境政策課、佐世保市環境部環境保全課には、大気汚染常時監視データの提供をいただきました。深く感謝いたします。

文献

- 1) 環境省:平成29年度大気汚染の状況(有害大気汚染物質を除く)(2019), https://www.env.go.jp/air/ref_h29.pdf.
- 2) 環境省:平成25年度大気汚染の状況(有害大気汚染物質を除く)(2015), https://www.env.go.jp/air/osen/jokyo_h25/Full.pdf.
- 3) 熊本県:大気・化学物質・騒音等環境調査報告書(第49報:平成25年度版)(2014), https://www.pref.kumamoto.jp/kiji_6089.html.
- 4) 熊本県:大気・化学物質・騒音等環境調査報告書(第53報:平成29年度版)(2018), https://www.pref.kumamoto.jp/kiji_25194.html.
- 5) 荒木 真:大気環境モニタリングネットワークの最適配置に関する研究,大阪大学大学院工学研究科博士学位論文(2015).
- 6) Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., Briggs, D. J.: Mapping of background air pollution at a fine spatial scale across the european union. *Science of the Total Environment*, 407 (6), 1852-1867 (2009).
- 7) Pearce, J. L., Rathbun, S. L., Aguilar-Villalobos, M., Naeher, L. P.: Characterizing the spatiotemporal variability of PM_{2.5} in Cusco, Peru using kriging with external drift, *Atmospheric Environment*, 43 (12), 2060-2069 (2009).
- 8) Mercer, L. D., Szpiro, A. A., Sheppard, L., Lindström, J., Adar, S. D., Allen, R. W., Avol, E. L., Oron, A. P., Laeson, T., Liu, L-J. S., Kaufman, J. D.: Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NO_x) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air), *Atmospheric Environment*, 45 (26), 4412-4420 (2011).
- 9) Araki, S., Yamamoto, K., Kondo, A.: Application of regression kriging to air pollutant concentrations in Japan with high spatial resolution, *Aerosol and Air Quality Research*, 15, 234-241 (2015).
- 10) Araki, S., Sasaki, T., Ueda, S., Yamamoto, K., Tohno, S.: A study of the prediction method of the spatial distribution

- of ozone by kriging in Otsu city and its surrounding area, *Journal of Japan Society for Atmospheric Environment*, 46, 241-250 (2011) [in Japanese].
- 11) 古谷 知之: Rによる空間データの統計分析, 朝倉書店 (2011) .
 - 12) Kubota, T., Tarumi, T.: GoogleMaps を用いた地理統計データの可視化, *Proceedings of the Institute of Statistical Mathematics*, 55 (1), 101-112 (2007) [in Japanese].
 - 13) R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austraria, <https://www.R-project.org> (2018) .
 - 14) Pebesma, E. J. and Bivand, R.S.: Classes and methods for spatial data in R, *R News*, 5 (2) , <http://cran.r-project.org/doc/Rnews/> (2005) .
 - 15) Bivand, R. S., Pebesma, E. J. and Gomez-Rubio, V.: *Applied spatial data analysis with R*, Springer, New York, <http://www.asdar-book.org/> (2008) .
 - 16) Pebesma, E.J.: Multivariable geostatistics in S: the gstat package, *Computers & Geosciences*, 30, 683-691 (2004) .
 - 17) Keitt, T. H., Bivand, R. S., Pebesma, E. J. and Rowlingson, B.: rgdal: Bindings for the Geospatial Data Abstraction Library, R package version 0.6-28, <http://cran.r-project.org/package=rgdal> (2010) .
 - 18) Bivand, R. S., Rundel, C., Pebesma, E.J., Stuetz, R., Hufthammer, K. O., Giraudoux, P., Davis, M., Santilli, S.: rgeos: Interface to Geometry Engine - Open Source ('GEOS') , <https://CRAN.R-project.org/package=rgeos> (2019) .
 - 19) Lewin-Koh, N. J., Bivand, R. S., contributions by Pebesma, E. J., Archer, E., Baddeley, A., Bibiko, H., Dray, S., Forrest, D., Friendly, M., Giraudoux, P., Golicher, D., Gomez-Rubio, V., Hausmann, P., Jagger, T., Luque, S. P., MacQueen, D., Niccolai, A., Short, T. and Stabler, B.: mapproj: Tools for reading and handling spatial objects, R package version 0.7-38, <http://cran.r-project.org/package=mapproj> (2010) .
 - 20) Venables, W. N., Ripley, B. D.: *Modern Applied Statistics with S*. Fourth Edition, Springer, ISBN 0-387-95457-0 (2002) .
 - 21) 国立環境研究所: 環境数値データベース (2019) , <https://www.nies.go.jp/igreen/>.
 - 22) 環境省: 環境大気常時監視マニュアル 第6版 (2010) , https://www.env.go.jp/air/osen/manual_6th/.
 - 23) Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. D., Barker, D. M., Duda, M. G., Huang, X-Y., Wang, W., Powers, J. G.: A Description of the Advanced Research WRF Version 3, NCAR Tech. Note, NCAR/TN-475+STR, 113 pp., National Center for Atmospheric Research, Boulder, Colorado, USA (2008) .
 - 24) National Centers for Environmental Prediction (NCEP) : CISL Research Data Archive (2017) , <http://rda.ucar.edu/datasets/ds083.2/>.
 - 25) 京都大学生存圏研究所: 生存圏データベース グローバル大気観測データ (2019) , <http://database.rish.kyoto-u.ac.jp/arch/glob-atmos/>.
 - 26) 総務省統計局 : e-Stat 国勢調査 3次メッシュ (2018) , <https://www.e-stat.go.jp/gis/statmap-search?page=1&type=2&aggregateUnitForBoundary=S>.
 - 27) 国土交通省: 国土数値情報 土地利用細分メッシュデータ (2018) , <https://nlftp.mlit.go.jp/ksj/index.html>.
 - 28) 気象庁: 過去の気象データ・ダウンロード (2018) , <https://www.data.jma.go.jp/gmd/risk/obsdl/index.php>.
 - 29) Hengl, T.: A practical guide to geostatistical mapping of environmental variables (2007) .
 - 30) Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D.: A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, 42 (33) , 7561-7578 (2008) .
 - 31) Ramos, Y., St-Onge, B., Blanchet, J., Smargiassi, A.: Spatio-temporal models to estimate daily concentrations of fine particulate matter in Montreal: Kriging with external drift and inverse distance-weighted approaches, *Journal of Exposure Science and Environmental Epidemiology*, 26, 405-414 (2016) .