

ベイズ統計を用いた PM_{2.5} 重量濃度に与える風向の影響の解析

久恒 邦裕

Bayesian Approach for the Effect of the Direction of the Wind for PM_{2.5}

Kunihiro Hisatsune

2011 年度および 2012 年度の名古屋市内 2 地点（元塩公園および八幡中学校）における自動測定器による PM_{2.5} 測定結果（時間値）について、風向がどのような影響を与えているかを評価するために CPF(Conditional Probability Function) とベイズ統計による解析を行った。その結果、いずれの解析においても元塩公園では北を、八幡中学校では南西を中心とした方向からの風が PM_{2.5} の高濃度化に影響する結果が示唆された。ベイズ統計の結果ではさらに濃度上昇への影響量や結果の信頼性とサンプル数の関係などを、定量的に評価することができた。

はじめに

微小粒子状物質 (PM_{2.5}) の環境基準が 2009 年 9 月に、“年平均値が 15 $\mu\text{g}/\text{m}^3$ 以下であり、かつ日平均値が 35 $\mu\text{g}/\text{m}^3$ 以下”と告示された。名古屋市では 2011 年度より市内 2 地点の元塩公園（自排局、以降“元塩”）と八幡中学校（一般局、以降“八幡”）において常時監視を開始し 1 時間値の自動採取を行ってきた（以降、“1 時間値”は全て自動測定機による測定値とする）。また、2012 年度からは自排局として港陽、上下水道局北営業所、千竈を、一般局として滝川小学校、白水小学校、富田支所を追加して計 8 地点での常時監視体制を整えている。

2012 年度になり、中国における PM_{2.5} 濃度の上昇とその日本への影響の可能性について大きく報道され、PM_{2.5} の注目度は一気に高まった。そのため、インターネット上で測定値を公開されているウェブサイト（“そらまめ君”¹⁾ など）については接続が困難になる状況が続いた。名古屋市においても要望の高まりを受けて、2013 年 4 月より市の HP にて 1 時間値（速報値）の逐次公開を行っている。しかし、前述の通り定められた環境基準は年平均値もしくは日平均値のみについてであり、また採用された自動測定機についても標準測定法であるフィルター捕集 - 質量法との等価性は 24 時間捕集の条件で確認されているが、1 時間値については等価性が確認されておらず参考値とされている。しかし、実際には人々の高い関心はリアルタイムでの環境情報に相当する 1 時間値にあると考えられる。

また環境省から PM_{2.5} の注意喚起の暫定指針²⁾ が示された際には、その基準は“日平均値 70 $\mu\text{g}/\text{m}^3$ を超えると予想

される場合”と定められており、その判断には早朝から午前中の 1 時間値の使用が挙げられている。そのため、1 時間値の変動について、一定の知見をもつことは重要なことであると考ええる。

PM_{2.5} は様々な成分の混合物であり、発生場所、発生源、および環境中での変化などはそれぞれの成分ごとに異なっているので、その正確な挙動を解明するためには成分分析による研究が欠かせない。しかし現在名古屋市で採用している自動測定機は 1 時間ごとの試料採取で成分分析を行う仕様にはなっておらず、技術的に困難を極める。そのため、PM_{2.5} の 1 時間値については質量濃度に対して解析を行い値の変動に対する影響を調べるのが現実的な手法である（但し、前述の通り 1 時間値の精度に関しては不十分な部分があるが、本報告では議論しない）。そこで今回は、1 時間値そのものを対象として、影響を与える要因の一つとして風向に着目しどのように影響を与えるかについて解析を行った。

調査方法

1. 調査データ

調査地点は、2011 年度と 2012 年度の 2 年間継続でデータが採取されている元塩および八幡とし、地点の差や年ごとの違いについて検討した。地点の地図は図 1 に示し、主要な国道についても記載した。PM_{2.5} の 1 時間値は、名古屋市内で測定されている常時監視データを用いた。データは 1 時間ごとに計測され、また両地点では風向・風速のデータも同時に測定されている。風向のデータは 16 方位で示されており、風速が 0.4m/s 以下のときには“無風”として

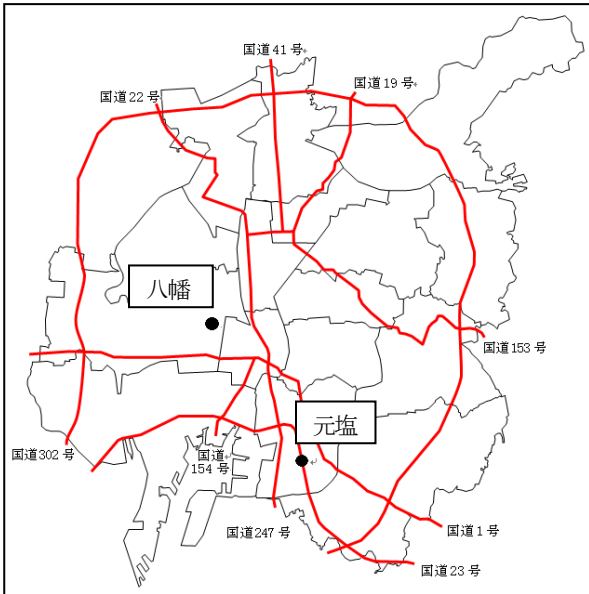


図1 調査地点概要図

記録される。今回の検討で、欠側等を除いた1時間値および風向・風速の情報が得られたデータ数は、元塩で2011年度は8784セット、2012年度は8520セット、八幡ではそれぞれ8564セット、8120セットであった。調査ごとの無風の割合は、元塩で2011年度が22.1%、2012年度が23.5%、八幡でそれぞれ3.6%と3.4%であった。なお、本検討では風向のみを検討の対象とし、風速の影響については無視した。

2. CPF 解析

風向との影響を調べるための手法の一つとして、CPF(Conditional Probability Function)解析⁵⁾を用いた。通常、CPF解析ではレセプターモデル(例えば、CMB(Chemical Mass Balance)法やPMF(Positive Matrix Factorization)法⁶⁾など)によって発生源因子を推定し、寄与濃度の高い場合の風向の頻度分布から観測地点近傍の発生源位置情報を得る手法である。CPFは式1のように定義される。

$$CPF = \frac{m_{\Delta\theta}}{n_{\Delta\theta}} \quad \text{式1}$$

ここで $\Delta\theta$ はそれぞれの風向を表し、 $m_{\Delta\theta}$ は注目している発生源因子が高寄与時に吹いた $\Delta\theta$ 方向からの風の数、 $n_{\Delta\theta}$ は調査期間に $\Delta\theta$ 方向から吹いた風の数全てを表す。CPF解析の際には、無風時の風は解析から外してCPF値を求め

た。本検討では $PM_{2.5}$ の質量濃度を解析対象としているため、発生源因子ではなく1時間値の高濃度時を高寄与時としてとらえた。高濃度時としては、全体の25パーセントイル値以上の時と定めた。

CPF解析では式1のように風向ごとに対して高濃度時の風数を全ての風数で除するが、分母となる全ての風数が風向ごとに大きく異なる場合がある。そのため、同じCPF値が得られたとしても、データの信頼性が異なる可能性があり同じ尺度での評価に適さない場合も考えられる。そこで、そのような風向ごとの風数の差を考慮する方法として、今回は次項で述べるベイズ統計を用いた解析を採用し、風向と1時間値の関係性について検討した。

3. ベイズ統計

ベイズ統計学は推測統計学のうち従来の統計学(以降、頻度主義的統計学)とは異なる理論体系をもった統計学である。ベイズ統計学の基礎となっているのは式2に示したベイズの定理であり、トーマス・ベイズ(Thomas Bayes)によって18世紀ごろに提唱された。

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad \text{式2}$$

ここで $P(A|B)$ は“事象Bが発生した条件で事象Aが起こる確率”、 $P(B|A)$ は“事象Aが発生した条件で事象Bが起こる確率”を、 $P(A)$ は“事象Aの起こる確率”、 $P(B)$ は“事象Bの起こる確率”を表す。事象Aを“任意の統計パラメータの値”とし、事象Bを“得られたデータ”だとすると、式2によって得られたデータを元に統計パラメータについて推定を行うことができる。ここで、データが得られた状態というのは起こった現実なので $P(B)$ は定数となるので、さらに式2は式3へと一般化できる。

$$P(A_i|B_j) \propto P(B_j|A_i)P(A_i) \quad \text{式3}$$

“ \propto ”は比例関係を表す。ここで $P(A_i|B_j)$ は事後確率(連続データでは事後分布)と呼ばれ、これが $P(B_j|A_i)$ で表される尤度と $P(A_i)$ で表される事前確率(連続データでは事前分布)の積に比例することが示されている。事前確率と尤度によって計算された事後確率は、また事前確率となり次の観測データの情報が尤度として計算されると、その積

によって新たな事後確率が導かれる。実際には、このようにして、得られたデータを連続して用いることで統計パラメータを更新していき解析を行う。このように、データが得られる毎にパラメータが更新されることをベイズ更新と呼び、ベイズ統計の特徴の一つとなる。近年の計算機の性能向上に伴い多くの分野で応用されており、スパムメール判定や顧客動向分析、音声認識などに使われており、また環境分野でもオゾン濃度の変動について日内変動と季節変動の分離評価³⁾や、PM_{2.5}の暴露量と循環器系疾患による死亡率との関係の評価⁴⁾にも用いられている。

今回は、ベイズ統計による解析を行うために統計用のソフトである R(version 3.0.2 64bit 版)と WinBUGS(version 1.4.3)を用いた。WinBUGS はマルコフ連鎖モンテカルロ (Markov Chain Monte Carlo : 以降, MCMC) 法の手法の一つである Gibbs sampling を行うことができるので、このソフトで事後確率からのサンプリングを行う。また、R と WinBUGS のデータのやり取りは R のパッケージの一つである R2WinBUGS を用いたが、さらにパラメータの設定を簡便にするために、北海道大学大学院の久保拓弥博士が作成した R2WBwrapper を使用した。

ベイズ統計による解析を行うためには、その現象と知りたい要因 (もしくは未知の要因) の間にある関係式 (モデル) を自ら構築して、仮定する必要がある。1 時間値に対する風向ごとの風の影響を調べるため、本研究では質量濃度を < 風の影響を受けない月ごとの平均値 (Mon) > と < 風向による上昇値 (Wid) > の合計であるとするモデルを仮定した。このことにより、任意の方向からの風が吹いたときの濃度上昇を調べることで、風向ごとの影響を評価することができる。また、観測された値 (OPM) は測定誤差や同一月の日ごとの差の影響を含んでいると考えられるので、誤差の影響のない値 (PM) を平均とし、標準偏差 σ である正規分布に従うと仮定した。以上の仮定を、以下の式 4 および式 5 として表した。

$$PM_{ab} = Mon_a + Wid_b \quad a = 1 \sim 12, b = 1 \sim 17 \quad \text{式 4}$$

$$OPM_{ab} \sim \mathcal{N}(PM_{ab}, \sigma) \quad \text{式 5}$$

ここで a は 1 月から 12 月までの 12 カ月を示す添え字であり、b は風向を北が 1、北北東が 2・・・と時計回りに 16 方位を振り分け、無風状態を 17 とし風向を示す数字である。それぞれの条件に従って PM_{ab} が決められ、式 5 はこの PM_{ab} を平均として標準偏差 σ の正規分布に OPM_{ab}

$$p(\{Mon_a\}, \{Wid_b\}, \sigma | D_{1 \sim N}) =$$

$$\frac{\prod_{i=1}^N p(D_i | \{Mon_a\} + \{Wid_b\}) p(\{Mon_a\}) p(\{Wid_b\}) p(\sigma)}{\int \int \int \cdots \int (\prod_{i=1}^N p(D_i | \{Mon_a\} + \{Wid_b\}) p(\{Mon_a\}) p(\{Wid_b\}) p(\sigma)) d\{Mon_a\} d\{Wid_b\} d\sigma} \quad \text{式 6}$$

a = 1~12, b = 1~17

が従うことを意味する。また、任意の地点・年度で得られたデータの総数を N としたときの i 番目のデータを D_i と示す。このように求めたい値 (OPM_{ab}) に対して、未知のパラメータ (Mon_a や Wid_b など) を設定するのを階層ベイズモデルといい、全体を定式化したものは以下の式 6 の通りとなる。式 6 のうち、前述の通り分母は定数となるので実際に計算をするのは分子の項のみになる。

$p(\{Mon_a\})p(\{Wid_b\})p(\sigma)$ は事前確率であり、それぞれ初期条件を設定する必要がある。通常、事前分布や初期条件はそれまでの知見などを応用する場合もあるが、一方で事前分布が結果に影響を与えることもあるため、十分ではない情報から間違った事前分布などを与えることで良い結果が得られないこともある。そのリスクを回避するため、今回は事前分布として無情報事前分布を与えた。Mon_a については得られたデータを平均 0、標準偏差 1 に規格化してから計算したので、事前分布としては範囲 - 1000 から 1000 までの均一分布を与えた。モデル設計の前提から Wid_b と σ については正の値をとる必要があるため 0 から 1000 までの均一分布とした。初期条件としては、それぞれ前述の事前分布に矛盾しない値をランダムに与えるプログラムとした。なお、Wid₁₇ は無風状態の影響なので 0 を指定した。MCMC 計算は計算回数を 50000 回として、初期条件の影響をなくすために初めの 2000 回分を廃棄した。また、計算結果は 10 ステップごとを採用し、この計算を独立で 3 回試行させた。計算結果は、トレースプロット、事後分布のグラフ、自己相関のグラフ、R-hat 値などで収束を確認した。いずれのパラメータもおおむね良好に収束していたが、Wid_b は 0 付近の値をとるものが多い一方で、正の値の制限があったためトレースプロットにやや偏りが出たが、今回はそのまま報告を行う。なお、その後、確認のため Mon_a については得られたデータを規格化せずに正の値の範囲を事前分布としたモデルでも計算を行ったが、得られた結果はほぼ同じであった。

結果と考察

初めに、それぞれの調査地点における環境基準達成状況を表 1 に示す。いずれの地点も年平均値および日平均値の 98 パーセンタイル値共に環境基準を上回っており、高い濃度で推移していることが解る。2011 年度について、4 季節各 2 週間の測定に行っている成分分析の結果⁷⁾をみると、

元塩では八幡と比較して自動車排ガスの影響が大きいことが示唆されている。

図2の上段に、各地点および年度の風配図を示した。風配

表1 PM_{2.5}測定結果

	2011年度		2012年度	
	年平均値	日平均値の98パーセントイル値	年平均値	日平均値の98パーセントイル値
元塩	17.2	41.7	18.9	50.0
八幡	18.0	44.8	16.1	39.9

図は、いずれも北西からの風が卓越していることを示しており、風向の偏りが大きいことが明らかとなった。なお、2011年度の元塩については、南風が吹かなかったのので以降において当該条件については議論の対象としない。

気象庁は気象観測所(名古屋市千種区)の風配図データを公開しており、1981年から2010年までの累積の風配図をみると、北北西の風が最も卓越して、北西、北の風がそれに続く結果となっていた。以上のことから図2の風配図の結果は、名古屋における風配図として一般的なものであると判断できる。図2の下段にはCPF解析の結果を示した。地点ごとに特徴的な形になったが、それぞれの年度ごとの大きな差はなく類似した傾向を示した。元塩公園では北を中心とした風向で高濃度になる傾向が示された。また八幡公園では、南西の方向に凸になる傾向が示され、それぞれCPFが凸になった方向の近傍において発生源がある可能性が示唆され、風向により濃度変動が起こることが明

らかとなった。次に、ベイズ統計による解析結果を示す。

まず、CPF解析の結果と比較するために、風向の影響について述べる。図3に、ベイズ統計で計算された風向ごとの影響値[Wid]とCPF解析の結果を示した。ベイズ統計の結果は、任意の風向の風が吹いた際にどの程度濃度が上昇するかを示しており、ひし形が影響の50パーセントイル値を、上下のバーが95%確信区間を示している。比較のために丸印でCPF解析の結果も示した。グラフの縦軸の単位が異なるので単純な比較はできないが、[Wid]とCPF値の風向ごとの変動は、おおむね類似したパターンを示した。元塩では北を中心とした風向で影響が大きいことが示され、八幡では南西の風向で影響が大きいと示唆された。ただし、変動のパターンが一致しない条件(例えば、2011年元塩のSSW、SWやWSWや2012年元塩のNNE、NE、SSE、S、SSW、SWなど)はその確信区間が大きくなっていることが多く、それらの条件では結果の評価に注意が必要であると考えられる。つまり、2011年の元塩においてSSWではCPF値が高い結果となったが、[Wid]の確信区間が広いことを考慮するとその結果は偶発的に得られたものである可能性が高い。一方、2012年の八幡でSWの方向からの風のときには高濃度が観測されやすいという現象は、高い確度で起こりえることだと判断できる。また、風向による濃度増加の影響は元塩で最大約25 $\mu\text{g}/\text{m}^3$ と大きな値を示した一方、八幡では約10 $\mu\text{g}/\text{m}^3$ と影響の大きさに違いがあることが示された。元塩では風向の影響で上昇するPM_{2.5}の濃度が高く、それは近傍の発生源が高い濃度でPM_{2.5}を排出している可能性があることを示している。

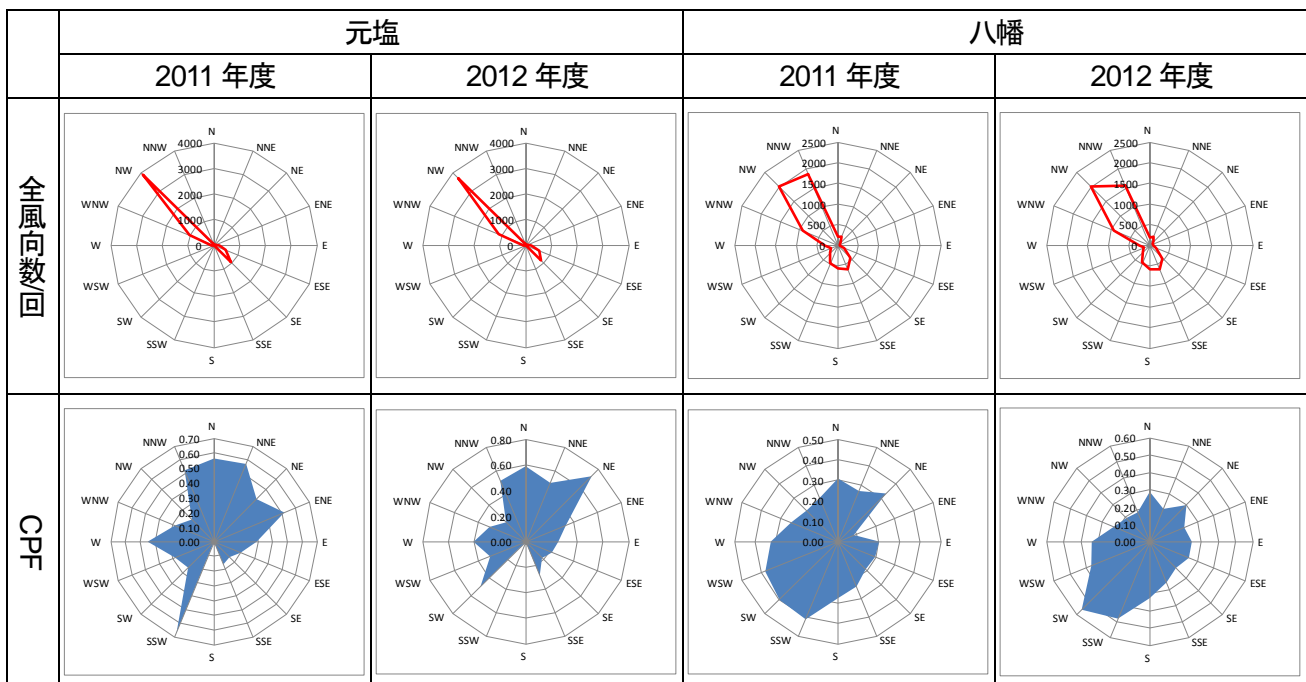


図2 風配図およびCPF解析結果

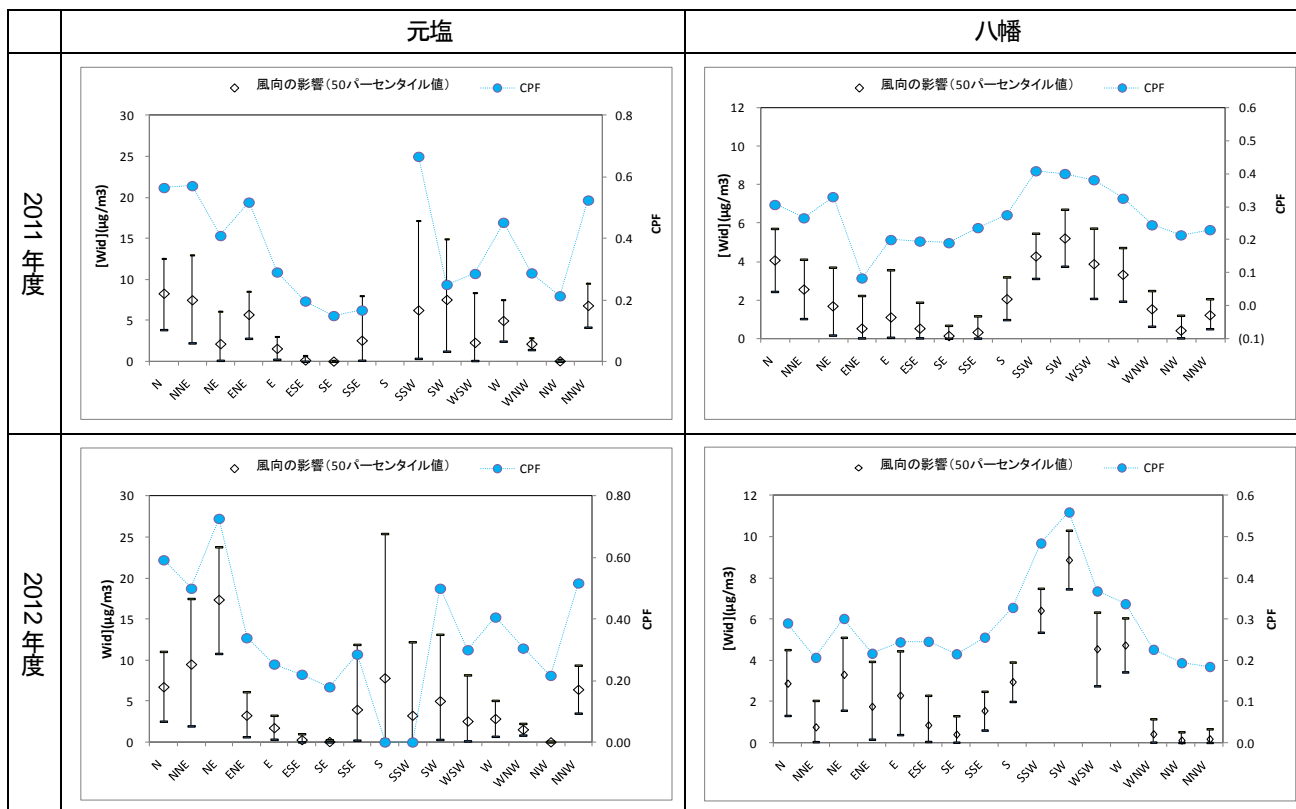


図3 PM_{2.5}質量濃度に対する風向の影響

前述の通り、元塩では自動車排ガスの影響が大きいと考えられ、風向による影響もそれを反映していると考えられる。ただし、元塩の隣接した国道23号(12時間交通量59,302台、大型車混入率22.7%)⁸⁾は測定局の西側に位置しており、風向の影響は直接的に発生源からではなく地形的な要因などで回り込む可能性があることも考慮しなければならない。八幡は、南西方向からの風で高濃度となる傾向があるが、その方向に明確な大規模発生源はみられない。南部の名古屋港近辺の工場地帯の影響が考えられるが、さらなる成分分析の結果などとの比較により高濃度要因を検証する必要がある。

次に、風向の影響を受けない月ごとの濃度[Mon]の結果を図4に示す。[Mon]については50パーセンタイル値をひし形で示し、参考として実測値の50パーセンタイル値を丸印で、95パーセンタイル区間をバーにて示した。両者の変動はおおむね一致しており、観測されるデータについては、おもに[Mon]の影響が支配的で[Wid]の影響は限定的であることが解る。また、モデルの前提から[Mon]は実測値より低い値を示すはずであるが、殆どで同等か若しくは観測値よりも高い値が示された。観測値の95パーセンタイル区間のうち50パーセンタイル値は中央よりやや低い

ところに位置しており、1時間値の中央値は平均値よりも低いと考えられ、その影響が[Mon]の値を高めにした可能性がある。また、[Mon]については、一日のなかでの時間変動やその他気象要因などの影響がまだ考慮されておらず、それらの影響を考慮することでより精度の高い統計モデルとなると考えられる。

各地点、各年度で計算された[Mon]について、月ごとの変化をグラフにまとめたものが図5になる。元塩は三角印で、八幡は線分のみで示し、2011年度データを実線、2012年度データを破線で示した。2011年度をみると、元塩と八幡の値はほぼ一致している一方、2012年度では月ごとの変動はおおむね一致しているが、元塩の方が高い傾向が示された。この差については前述の通り、モデルの精度などの問題が考えられ今後の更なる検討課題の一つである。

図3では、計算結果による95パーセント確信区間を示したが、その幅は風向ごとに大きく異なった。その風向ごとの幅に影響する要因を調べるため、風向ごとの風の数と、確信区間をグラフにして図6に示した。図6では元塩を丸印、八幡をひし形で、2011年データは中抜きで2012年データは中塗り示し、横軸(対数軸)に風向ごとの風の数を、縦軸に95%確信区間をとった。ベイズ統計による計

算は年度ごと、地点ごとに実施しておりそれぞれは独立した計算過程を経ているが、サンプル数と確信区間には一定の相関が見られた。またその傾向は年度や地点に依らず一

定の傾向を示し、例えば確信区間が $5\mu\text{g}/\text{m}^3$ 以下でデータを得たい場合には風向の数が 100 以上必要であると見積もることができる。

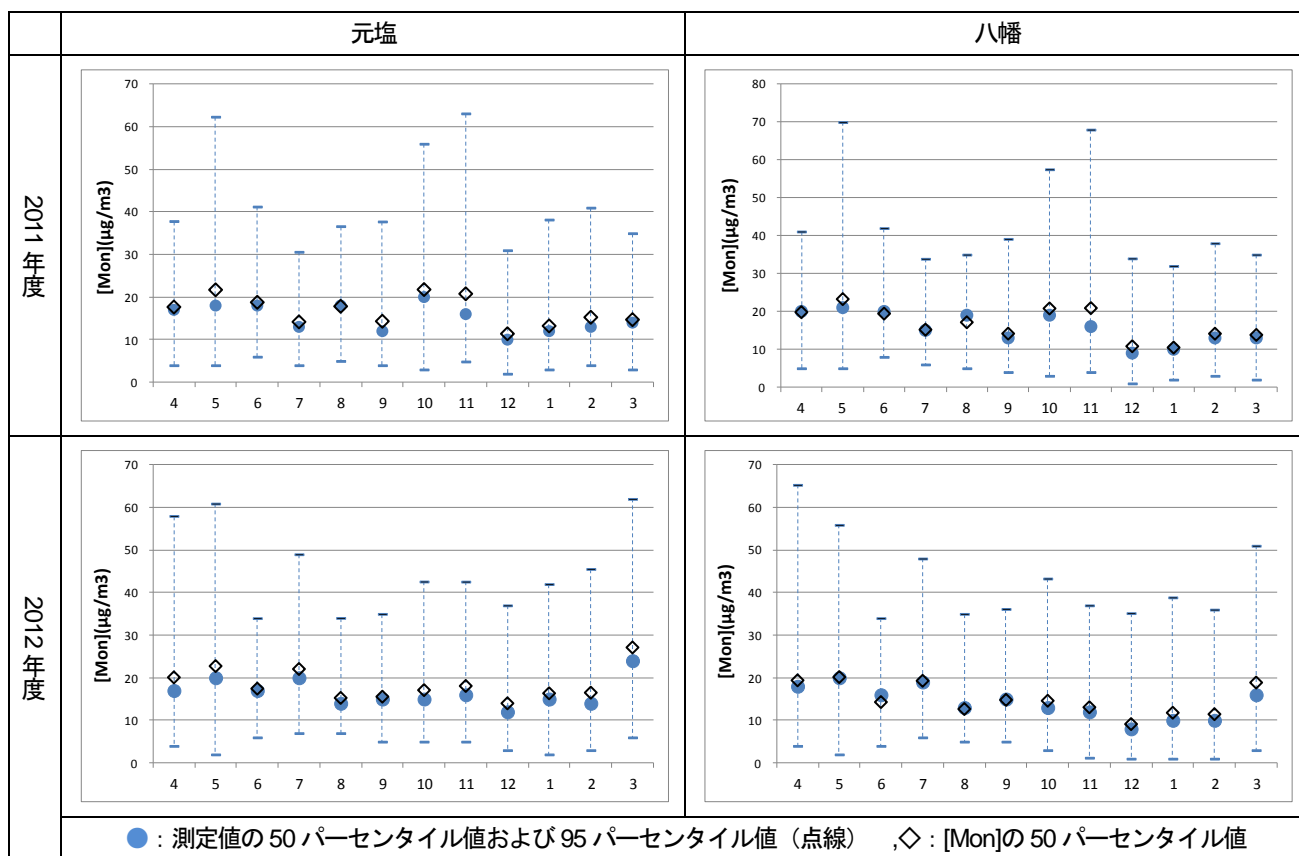


図4 . 風向の影響を受けない月ごとの濃度と観測された $\text{PM}_{2.5}$ 濃度

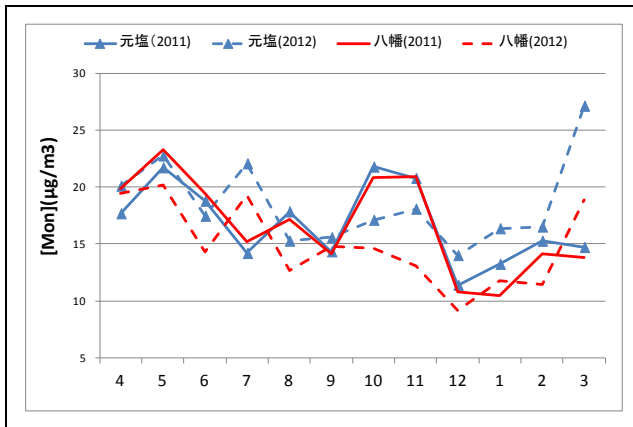


図5 風向の影響を受けない月ごとの濃度（計算値）の月変動

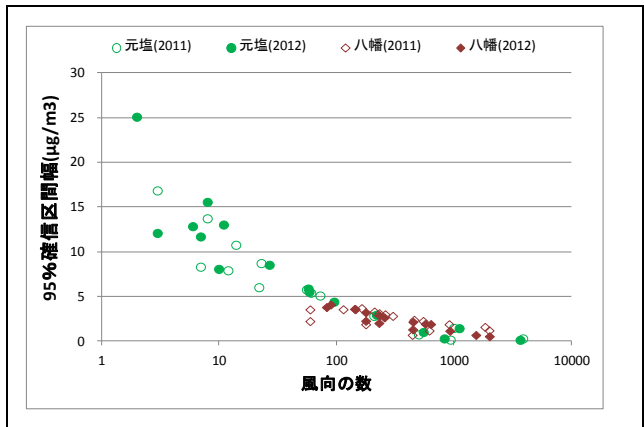


図6 95%確信区間幅と風向の数の関係

まとめ

PM_{2.5}の時間値に対する風の影響を、CPF およびベイズ統計によって解析した。いずれの結果においても、元塩では北からの風が吹くときに高濃度になる可能性が高く、八幡では南西の風の時に高濃度になる傾向があった。CPF解析では、高濃度となる風向のみを示したが、ベイズ統計による計算ではさらに、風向ごとに上昇する濃度やその信頼性も示すことができた。元塩では八幡に比べて風向によって増加する濃度が高いことが分かり、近傍の発生源の影響をより受けやすく、それは他の報告⁷⁾での結果より自動車の影響である可能性が高い。また、風向ごとの影響についてはそのサンプルの数が多くなるに従って確信区間が小さくなりその程度も定量的に評価することができた。ただし、PM_{2.5}の質量濃度については多くの発生源や生成機構があり、今回の風向の影響はその一つにすぎない。今後は、さらなる知見のもとでモデルを改良することで、質量濃度変化に与える影響をより詳細に探っていく。

文献

- 1) <http://soramame.taiki.go.jp/>
- 2) 微小粒子状物質 (PM_{2.5}) に関する専門家会合(第5回), 平成25年11月
- 3) 板野泰之, 高倉耕一: ベイズ統計手法による都市大気オゾンの日内変動と季節変動の分離評価, 大気環境学会, 46, 179-186(2011)

- 4) C. H. Holloman, et. al.: A Bayesian Hierarchical Approach for Relating PM_{2.5} Exposure to Cardiovascular Mortality in North Carolina, Environmental Health Perspectives, 112, 1282-1288(2004)
- 5) Kim, E.; Hopke P.K. (2004) Source Apportionment of Fine Particles in Washington, DC, Utilizing Temperature-Resolved Carbon Fraction. Journal of Air & Waste Management Association 54, 773-785
- 6) Patero, P.; Tapper, U. (1994) Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111-126
- 7) 山神真紀子, 久恒邦裕, 池盛文数: 微小粒子状物質 (PM_{2.5})の発生源寄与率の推定, 名古屋市環境科学調査センター年報, 1, 20-25(2012)
- 8) 平成22年度 名古屋市一般交通量概況 全国道路・街路交通情勢調査報告書 (道路交通センサス) 名古屋市