

微小粒子状物質変動の地域別トレンド分析について

(椿委員提供資料)

1. 分析対象地域と期間

大気汚染物質の環境濃度は、発生源条件、気象条件、地理的条件等の影響を受け、一定の地域内でも時間的に変化する。こうした環境濃度の時間的変化を把握するため、異なる地域の測定データによるトレンド分析を行う必要がある。

平成 13 年度から実施している微小粒子状物質曝露影響調査等の結果を用いて、東京都板橋区、大阪府守口市、福岡県福岡市の平成 13 年 4 月 1 日から平成 20 年 3 月 31 日までの 7 年間にわたり毎時測定した $PM_{2.5}$ 濃度データから、その期待値、分散の日内変動、週内変動、年内変動について、一般化加法モデルを用いて抽出した。ただし、福岡については、平成 14 年 4 月 1 日からの 6 年間のデータとなっている。

2. 前処理

上記 3 地域の 7 年間(福岡は6年間)にわたる $PM_{2.5}$ 濃度のヒストグラムを図1に示す。

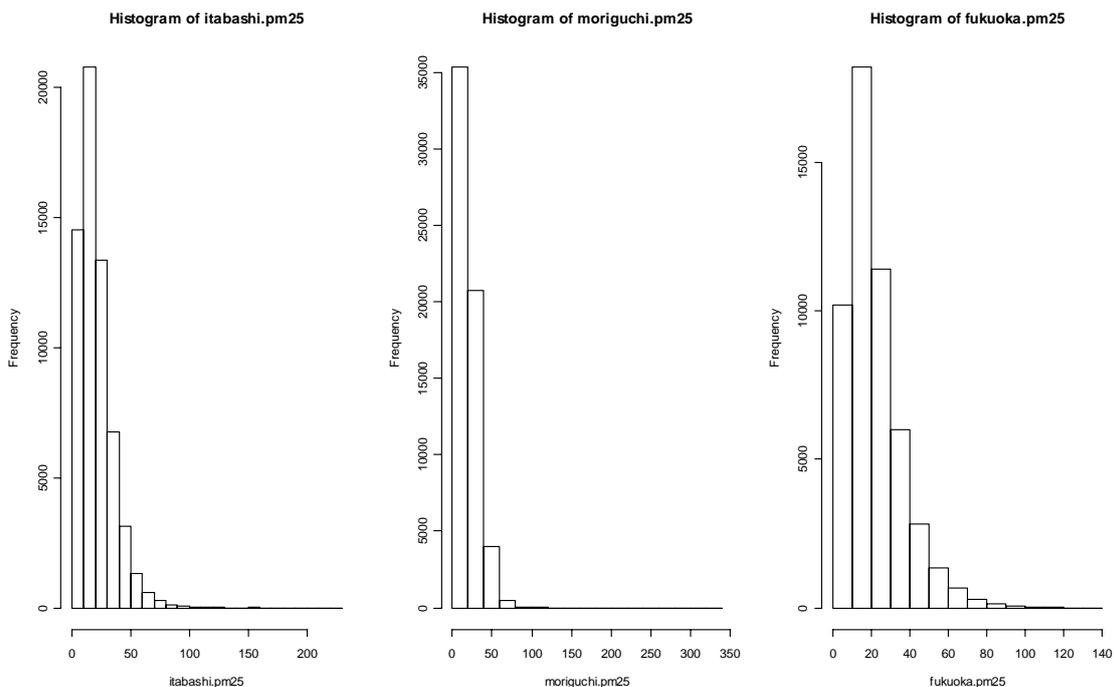


図1 3 地域の $PM_{2.5}$ 濃度のヒストグラム(左:板橋、中:守口、右:福岡)

これをボックスプロットで表示したのが図2である。

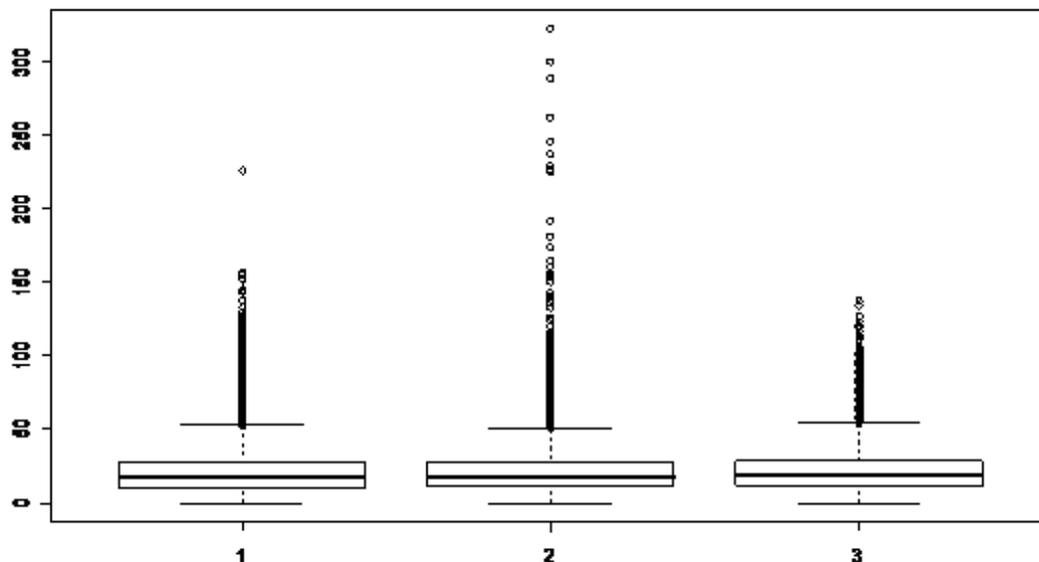


図2 図1に相当するボックスプロット(左:板橋、中:守口、右:福岡)

その5数要約は下記のとおりであり。4分位については、それほど大きな差がなく、むしろ外れ値の変動に大きなさがあることがわかる。

表1 板橋、守口、福岡のPM2.5濃度の5数要約(上段:板橋、中段:守口、下段:福岡)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00	10.40	17.40	20.55	27.30	225.50
0.00	10.90	17.40	20.22	26.80	322.10
0.00	11.40	18.30	21.67	28.40	137.00

さらに%点を詳細に調べると次のようになる。

表2 板橋、守口、福岡のPM_{2.5}濃度の%点(上段:板橋、中段:守口、下段:福岡)

90%	95%	98%	99%
39.0	47.5	59.3	68.6
37.2	44.3	52.9	60.3
40.7	50.0	61.8	70.2

これより、%点の範囲は、次のようになることがわかる。

表3 板橋、守口、福岡のPM_{2.5}濃度の%点の範囲

50%	75%	90%	95%	98%	99%
0.90	1.60	3.50	5.70	7.90	9.90

PM_{2.5} 濃度はいずれの地域でも立方根変換すると図3に示すように対称分布に近くなる。

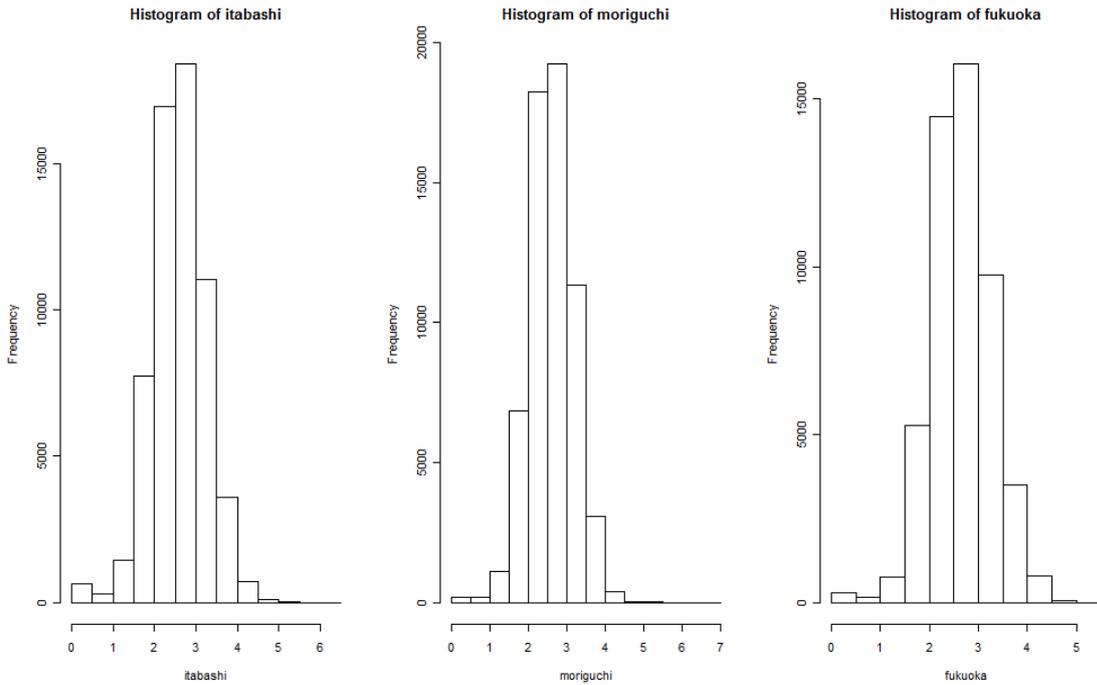


図3 3地域のPM_{2.5}濃度の立方根ヒストグラム(左:板橋、中:守口、右:福岡)

以下では、この立方根変換したPM_{2.5}濃度データ(近似的には粒径分布とみなせる)のトレンドを分析している。

3. 3地域のPM_{2.5}平均値のトレンド分析

3地域のPM_{2.5}の立方根変換データについて、日内変動成分(24時間周期トレンド)、週内変動成分(7日間周期トレンド)、年内変動成分(365.25日周期トレンド)、長期トレンドを一般化加法モデルで抽出した。これら4系統変動成分はすべて統計的に有意であった。

3地域の日内変動成分を立方根値からPM_{2.5}濃度に再換算して表示すると図4のようになる。ここで、時間原点は午前0時を意味する。福岡では、深夜期間において、それほど濃度が低下しないことが特徴的である。

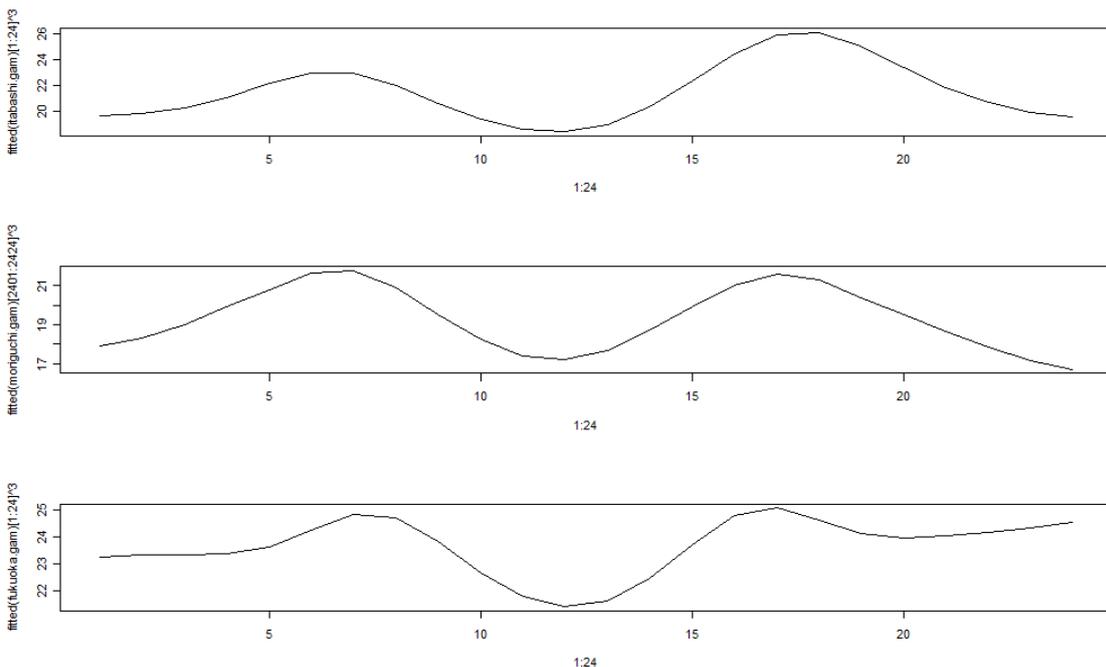


図4 PM_{2.5}日内変動成分(上段:板橋、中段:守口、下段:福岡)

次に、PM_{2.5}の週内変動成分を図示したのが図5である。ここで、時間原点は日曜午前0時を意味する。

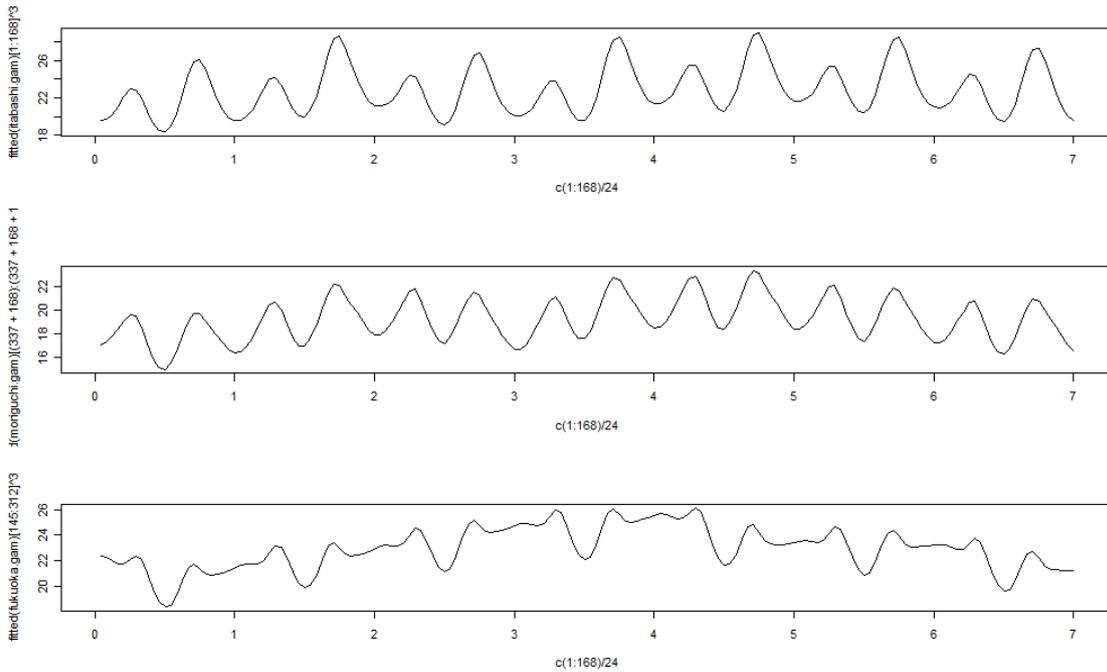


図5 PM_{2.5}週内変動成分(上段:板橋、中段:守口、下段:福岡)

次に、PM_{2.5}の年内変動成分を図示したのが図6である。ここで、時間原点は4月1日午前0時を意味する。各地区程度の差はあるが、4月から5月が減少6月に向けて上昇の傾向がみられる。12月頃から低め安定しながら徐々に4月に向けて上昇傾向になる。

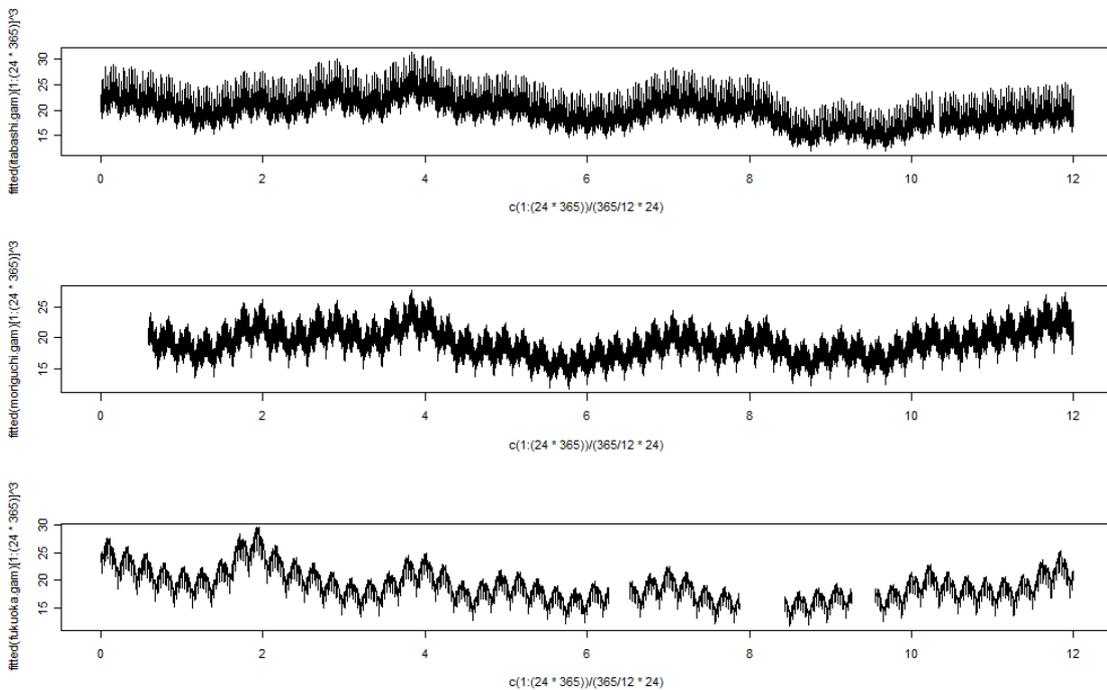


図6 PM_{2.5}年内変動成分(上段:板橋、中段:守口、下段:福岡)

図7に、今回のデータ全区間での予測値を示す。板橋に比べると福岡には顕著な減少傾向はみられない。

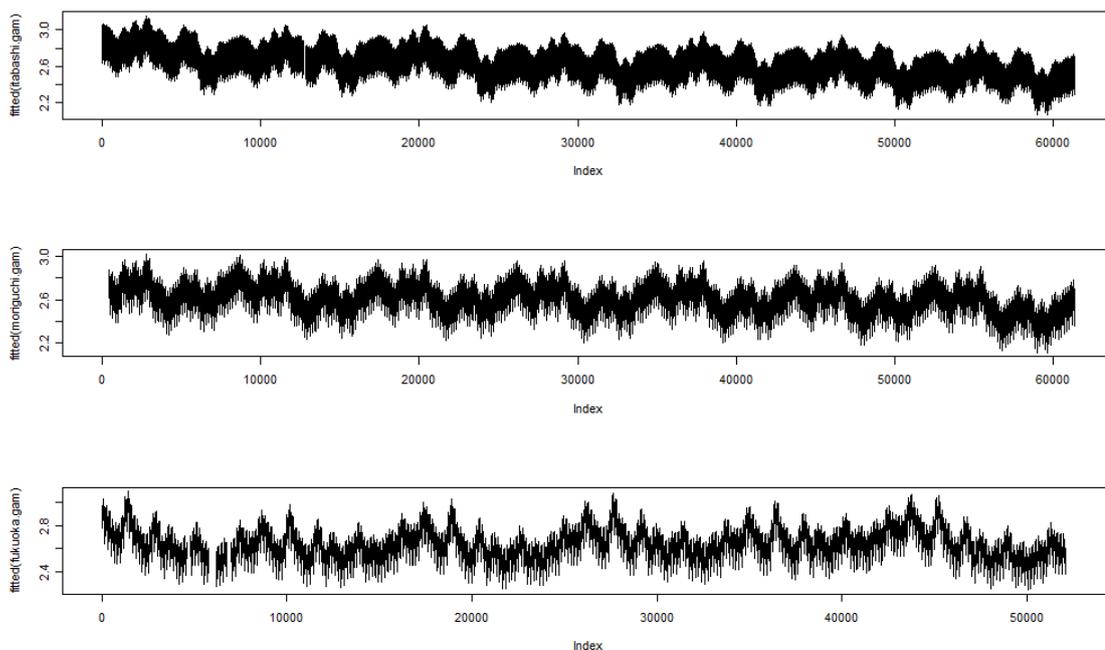


図7 PM_{2.5}長期トレンド(上段:板橋、中段:守口、下段:福岡)

4. 3 地域の PM_{2.5} 濃度立方根の分散の変動

3 地域の PM_{2.5} 濃度立方根データから図7で抽出したトレンド成分を除去した残差を抽出し、それを平均0、分散1に標準化したデータの時系列プロットを算出したのが図8である。±3標準偏差外については、いずれの地域ともに下側にずれているデータがかなりあることが特徴的である。また、ばらつきについて、変動が大きくなる時期と小さくなる時期といった残差変動にもトレンドがあることが認められる。

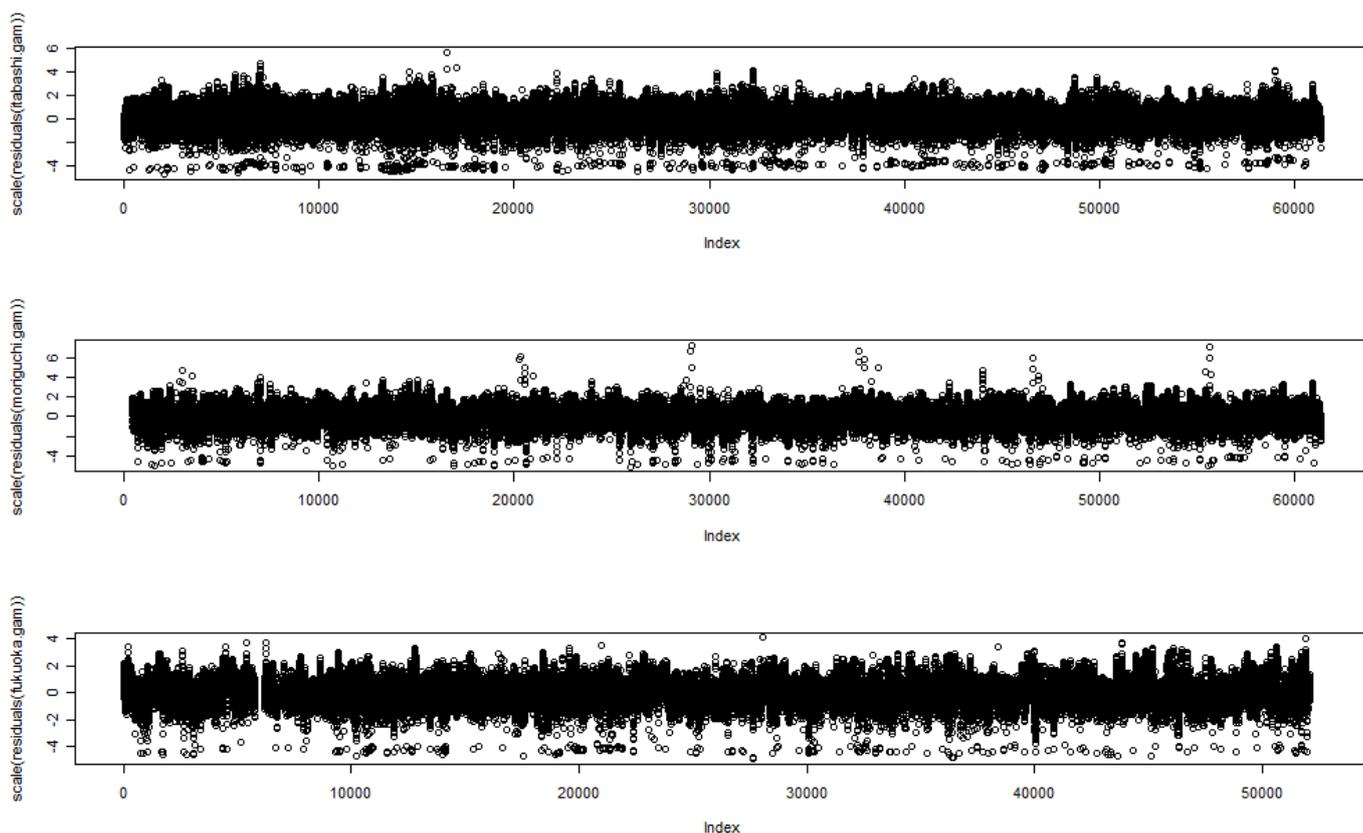


図8 残差を標準化したデータの時系列プロット(上段:板橋、中段:守口、下段:福岡)

そこで、残差二乗の系統変動を一般化加法モデル(分布はガンマ分布、リンク関数が対数関数)で分析した。時間原点の取り方については、2と同様である。

標準偏差の日内変動を図 9 に示す。PM_{2.5} の平均濃度が減少している正午頃がいずれの地域もばらつきは大きくなっており、濃度が第一のピークを迎える午前 6 時ごろが、ばらつきが小さくなっていることが分かる。

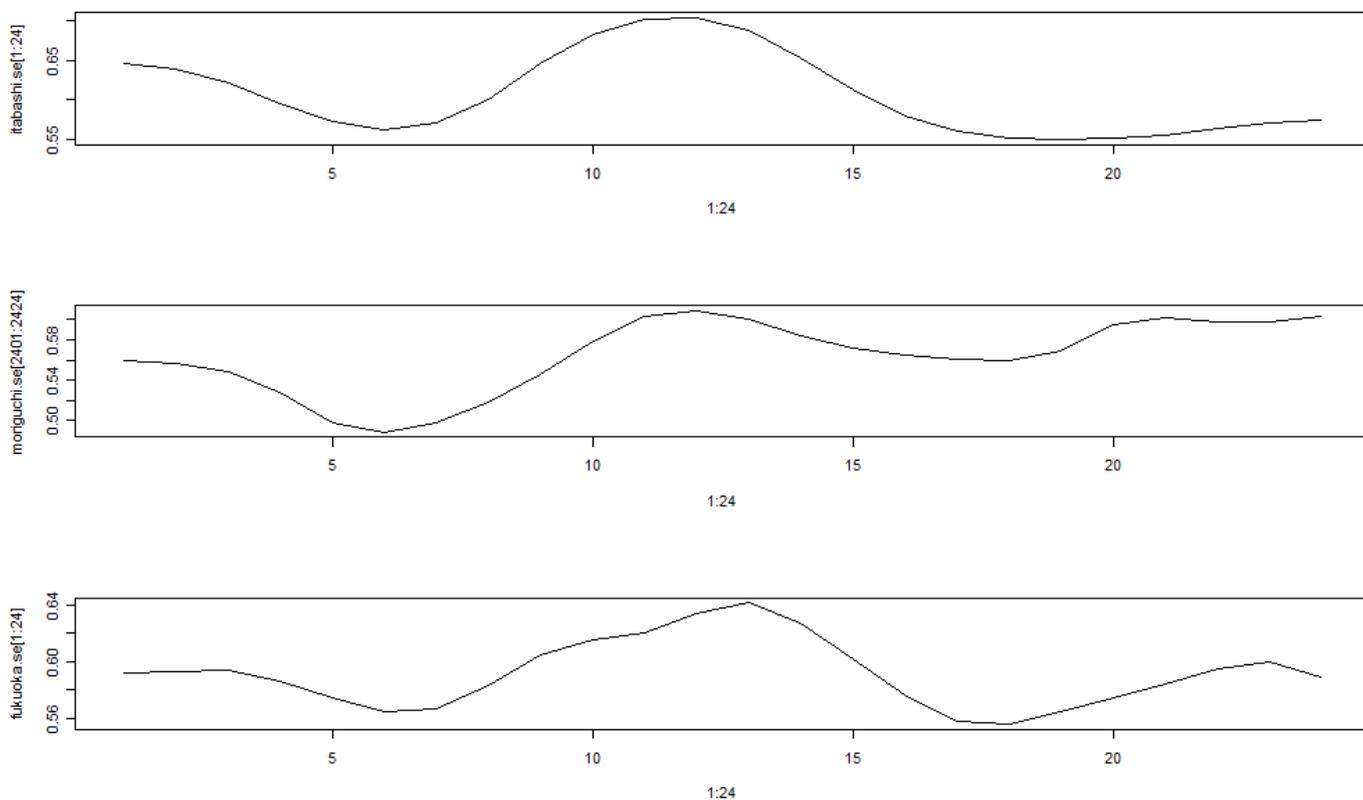


図9 変動の日内トレンド(上段:板橋、中段:守口、下段:福岡)

図10に変動の週内トレンドを示す。板橋、守口は休日、土曜のばらつきが大きい、福岡は必ずしもそのようにはなっていない。

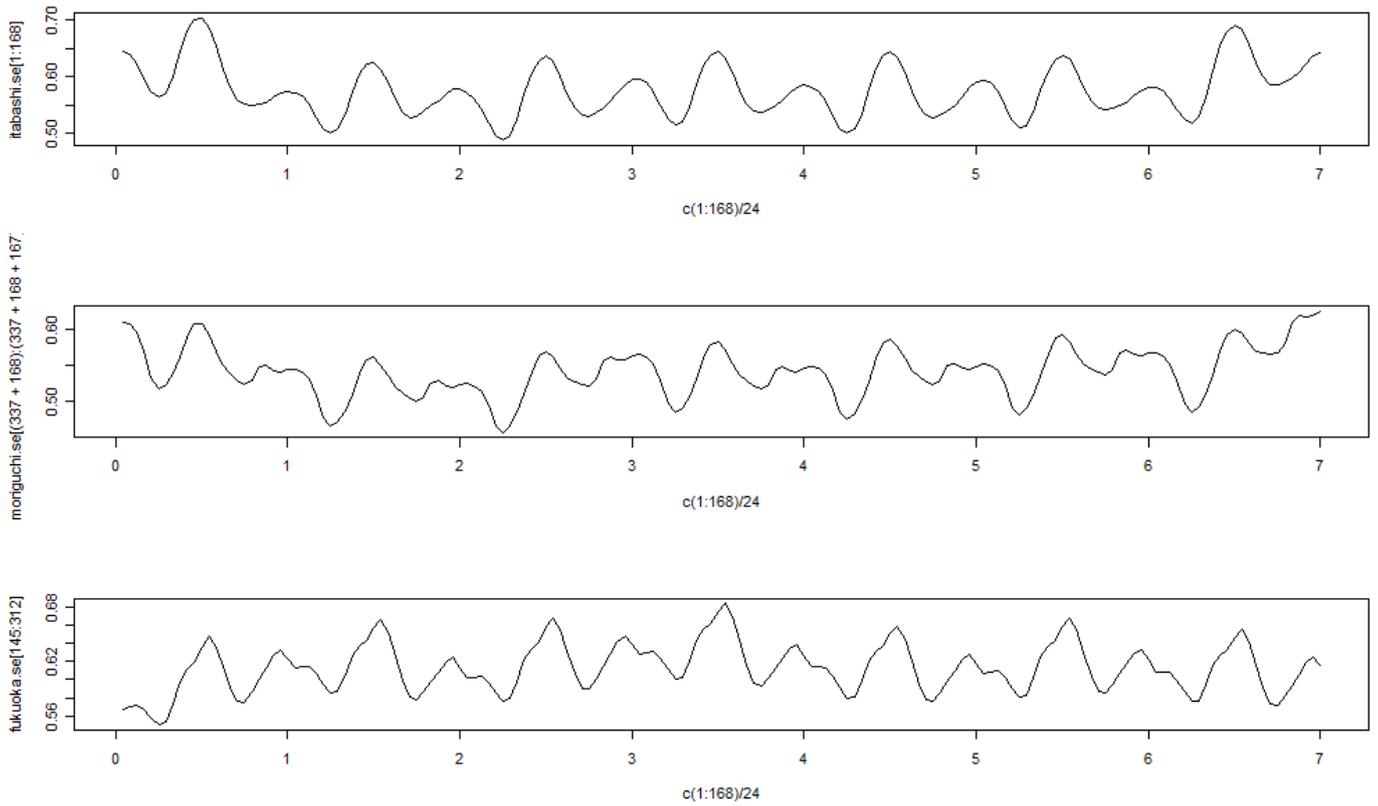


図10 変動の週内トレンド(上段:板橋、中段:守口、下段:福岡)

変動の年内トレンドについては、板橋において12月がもっとも不安定であるのに対し、福岡はむしろ4-6月が高い。守口はかなり変動が複雑であるが、傾向は福岡により近い。

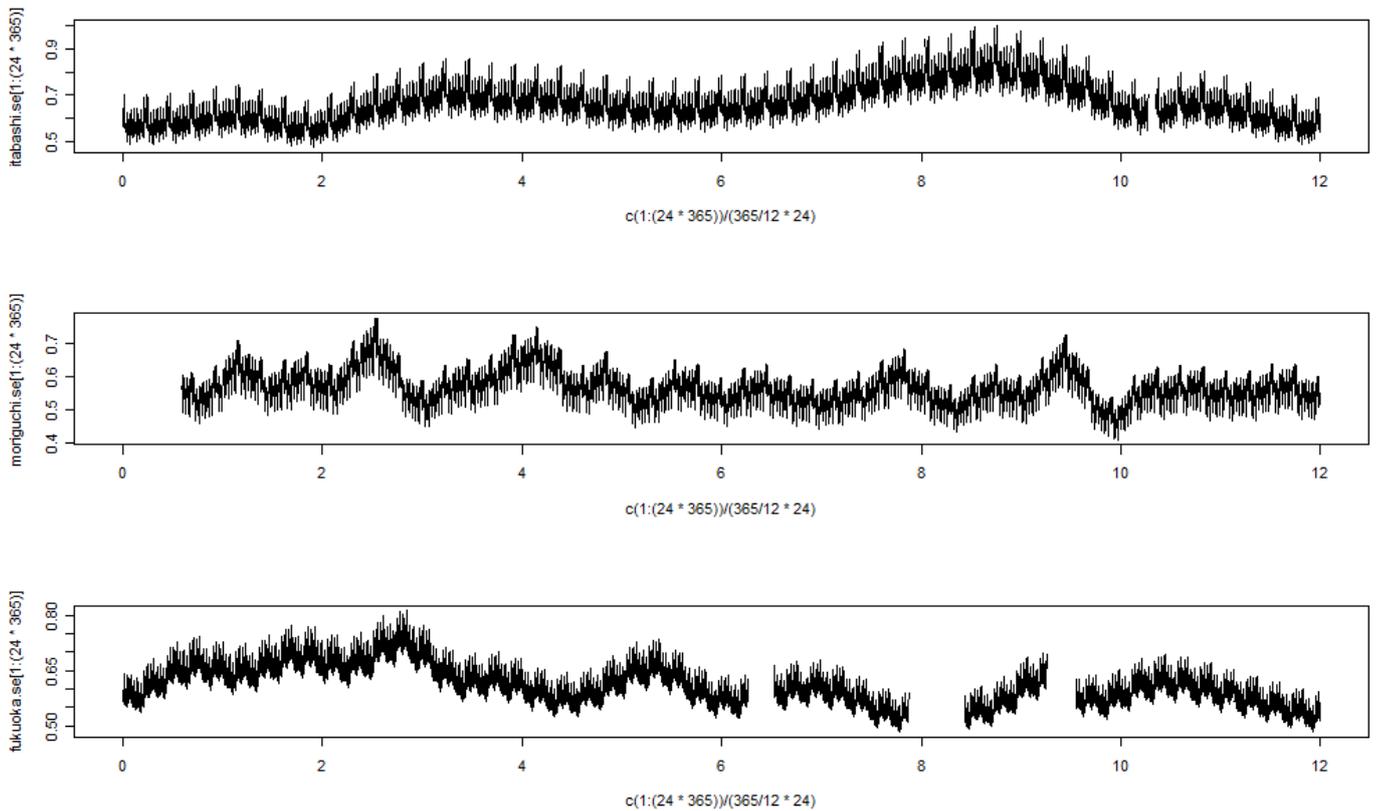


図11 変動の年間トレンド(上段:板橋、中段:守口、下段:福岡)

この $PM_{2.5}$ の立方根の変動の残差標準偏差の要約統計量を参考までに記載したのが表 5 である。守口の変動が比較的安定的であることがわかる。

表 5 $PM_{2.5}$ のトレンドからの残差標準偏差(上段:板橋、中段:守口、下段:福岡)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.3986	0.5677	0.6256	0.6362	0.6938	1.0730
0.3875	0.5265	0.5596	0.5631	0.5959	0.7905
0.4422	0.5656	0.6021	0.6056	0.6417	0.8380

ここで求めた各時点の標準偏差と各時点での期待値関数を用いて、原系列を標準化、すなわちデータから期待値を引き、標準偏差で除した系列が図 12 である。守口は安定的であるが、極端な外れ値も存在する。

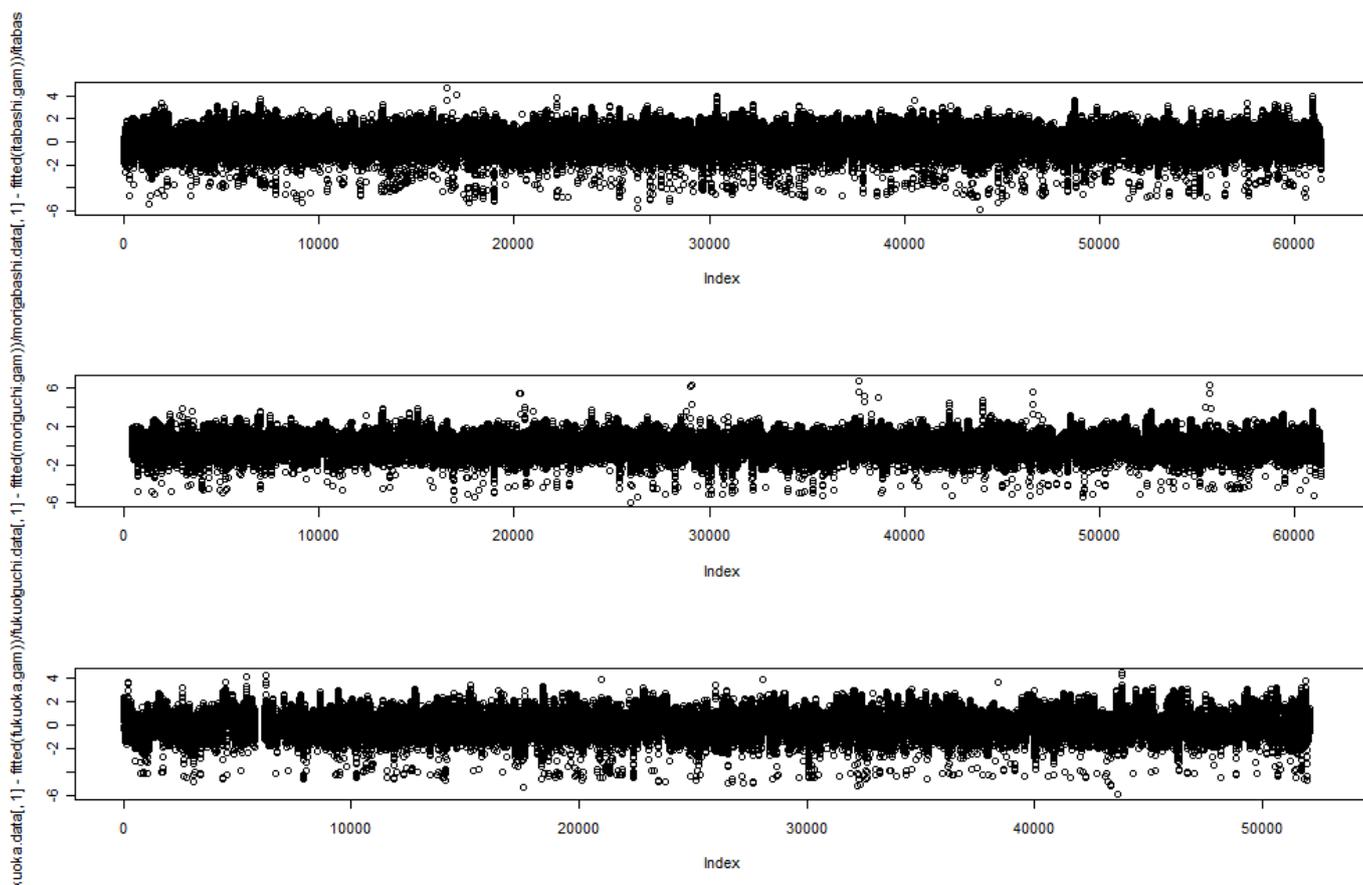


図 1 2 各時点のデータを各時点での期待値関数、標準偏差関数で標準化した系列
(上段:板橋、中段:守口、下段:福岡)

5. 期待値 + k 標準偏差と $PM_{2.5}$ 原系列の関係

3 地域の $PM_{2.5}$ 原系列に、そのトレンド推定からの予測値(赤)、予測値+2標準偏差(緑)、全データの98%点[青]を記載したのが、図 13a, b, cである。概ね98%基準は、トレンド+2標準偏差基準の代替として有効であることがわかる。なお、図 13 の横軸は、時間原点からの経過時間を意味する。

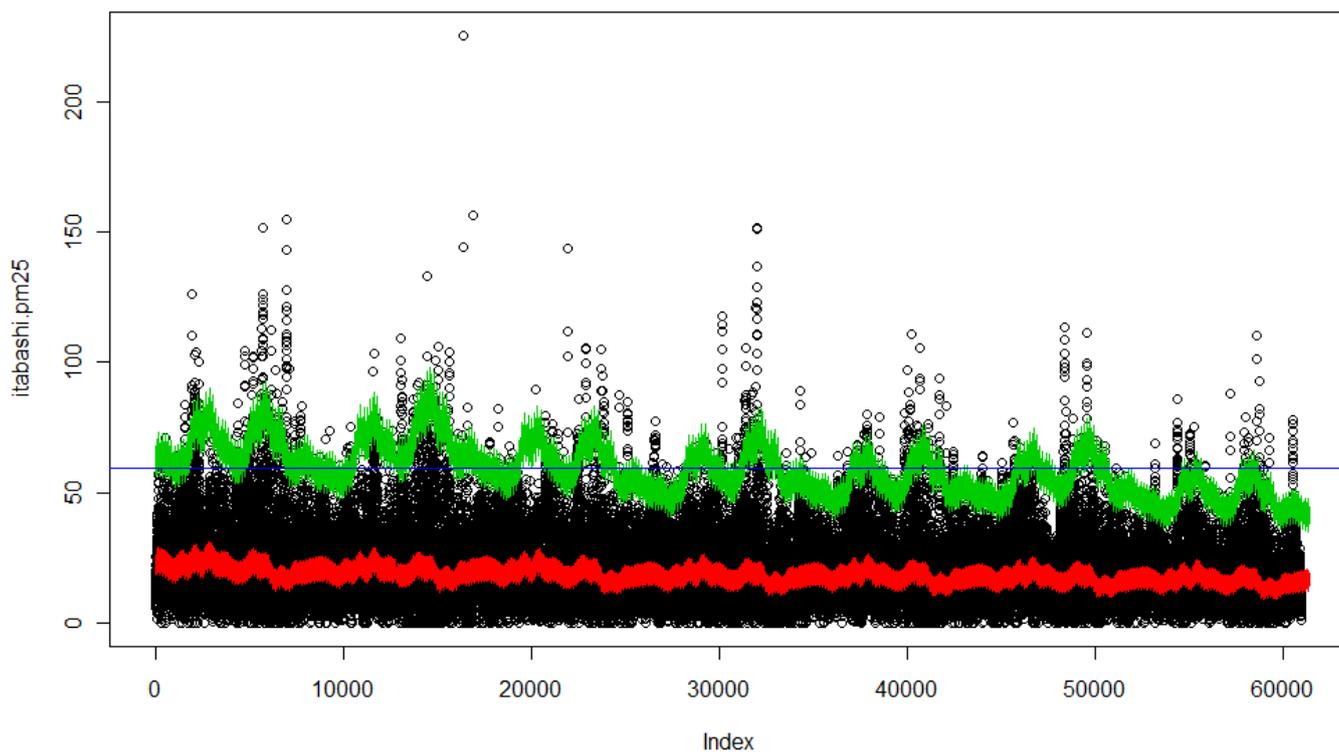


図 1 3 a 板橋の PM_{2.5} の変動期待値と上限値

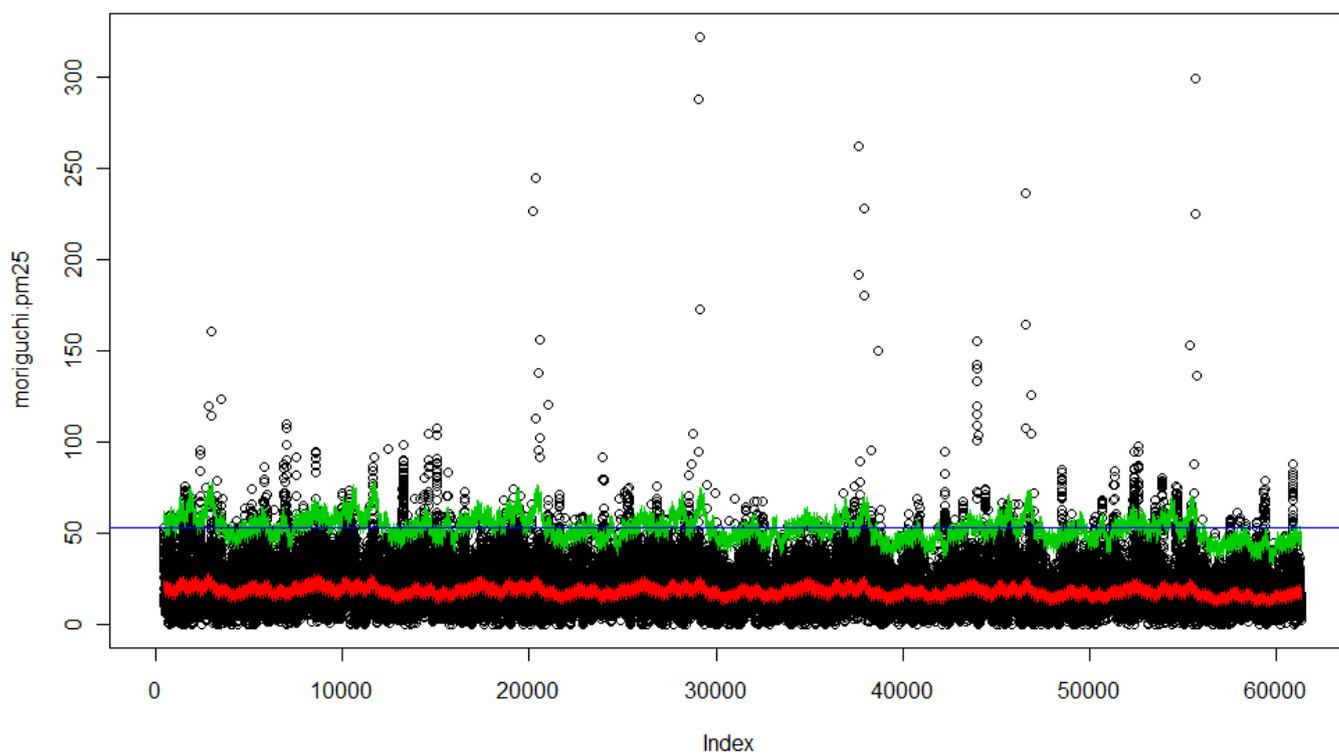


図 1 3 b 守口の PM_{2.5} の変動期待値と上限値

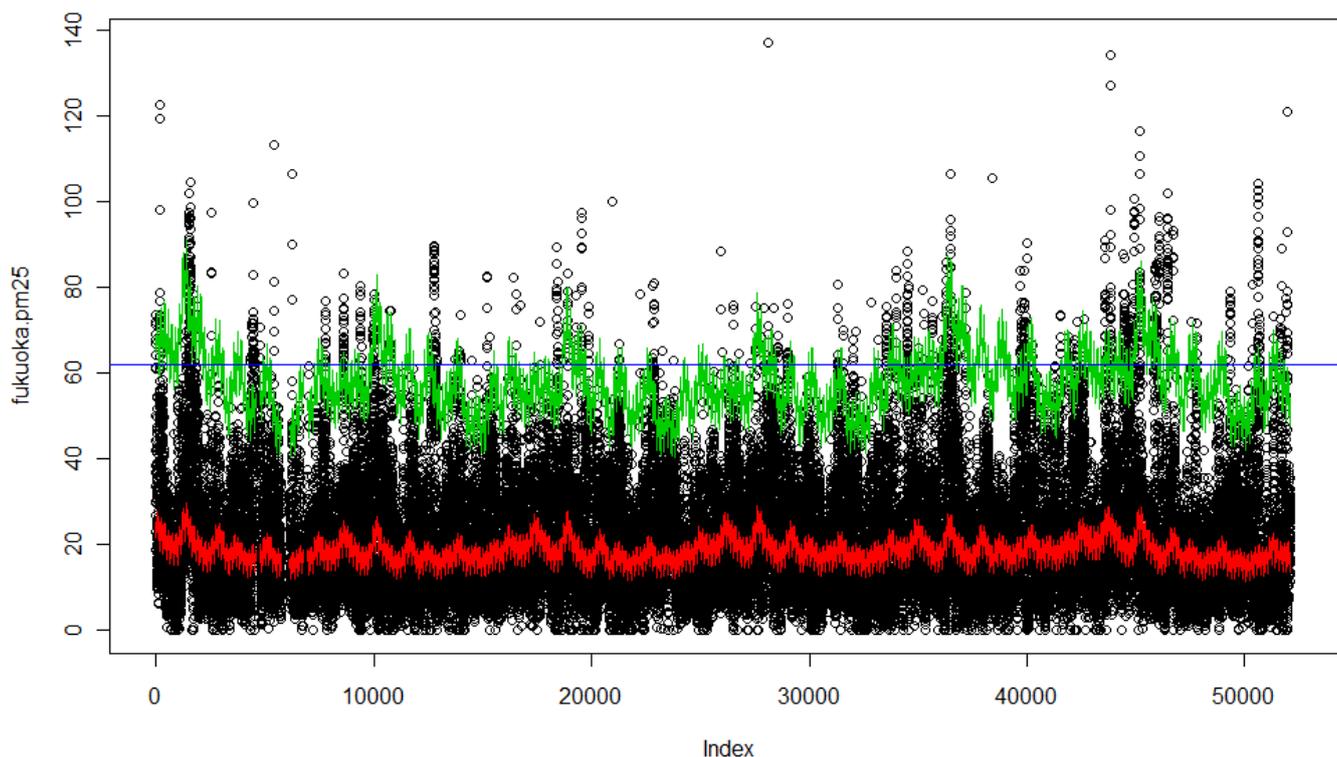


図 1 3 c 福岡の PM_{2.5} の変動期待値と上限値

6. 期待値の変動に関する日平均表示

5節の解析結果表示を基準値設定に関わる日平均値（24時間平均）で、表示する。この場合、PM_{2.5}実測値及びその期待値については、単純に24時間平均を算出すればよい。一方、期待値（予測値）の24時間平均の変動（分散）評価に当たっては、今回の分析に用いているPM_{2.5}の立方根変換値が、図14に示すように24時間変動には大きな自己相関を有しているため一定の調整が必要となる。

この自己相関から24時間平均値の実効標本数を計算すると、板橋：1.91、守口：1.65、福岡：1.61と算出される。実態として、24時間平均は24データの平均というよりは3データの平均程度しか分散減少を見込めないということである。

以上の自己相関に基づき、平均値の分散に関する評価の調整を行い、図13に対応するPM_{2.5}の24時間平均値系列にそのトレンドの24時間平均値系列（予測値系列）ならびに、予測値+2標準偏差を、それぞれ示したのが、図15a, b, cである（予測値（赤）、予測値+2標準偏差（緑）、全データの98%点[青]）。なお、図15の横軸は時間原点からの経過日数である。図13同様、いずれの地域も24時間平均値を用いる場合、98%基準は、トレンド+2標準偏差基準のほぼ同じ値となっていることがわかる。

このことから、24時間平均値の98%点以上の値をトリムした平均すれば、周期トレンドなどを加味した上でも、概ね平均+2標準偏差以上の分布については、トリムが行われることを保証する基準となっている。

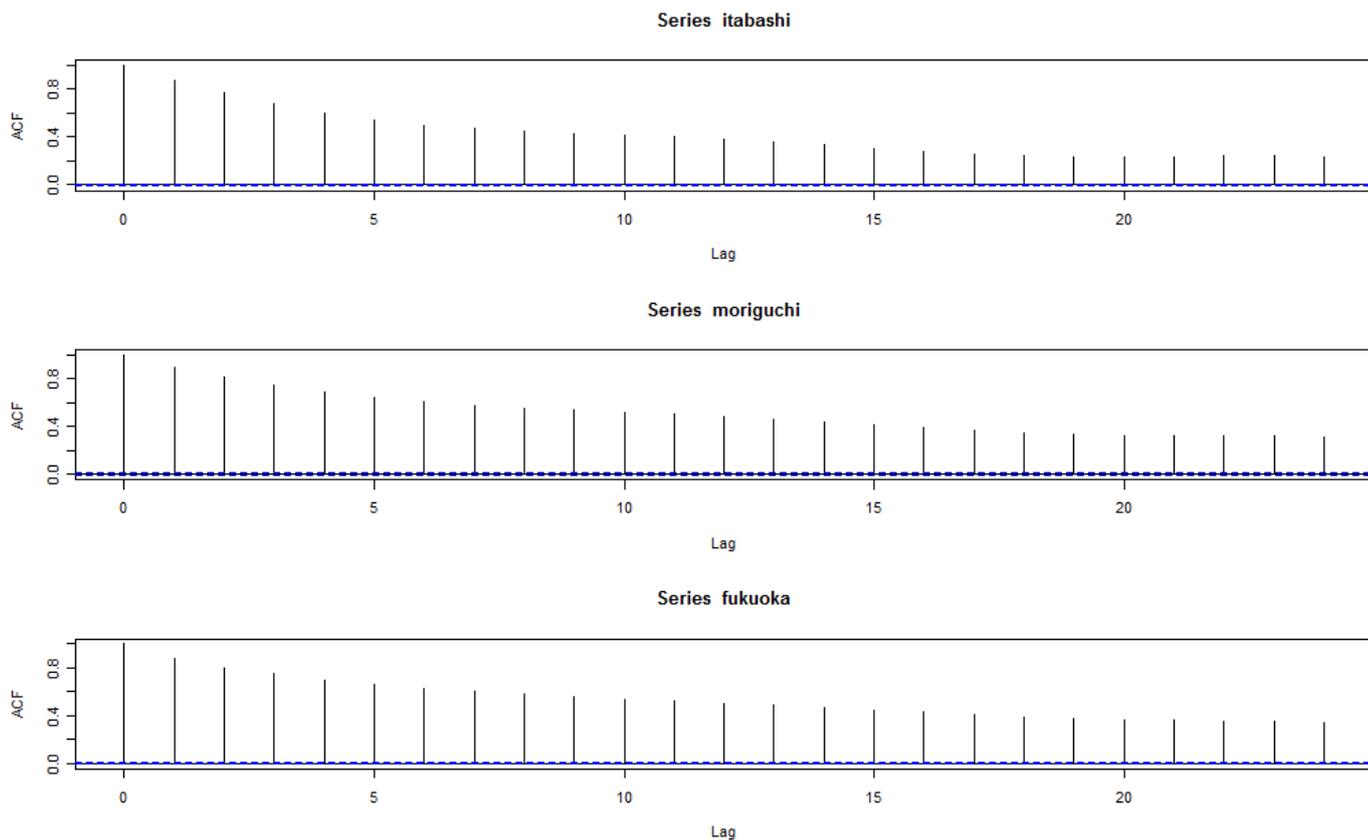


図 1 4 $PM_{2.5}$ の立方根変換値の 24 時間変動の自己相関 (上段：板橋、中段：守口、下段：福岡)

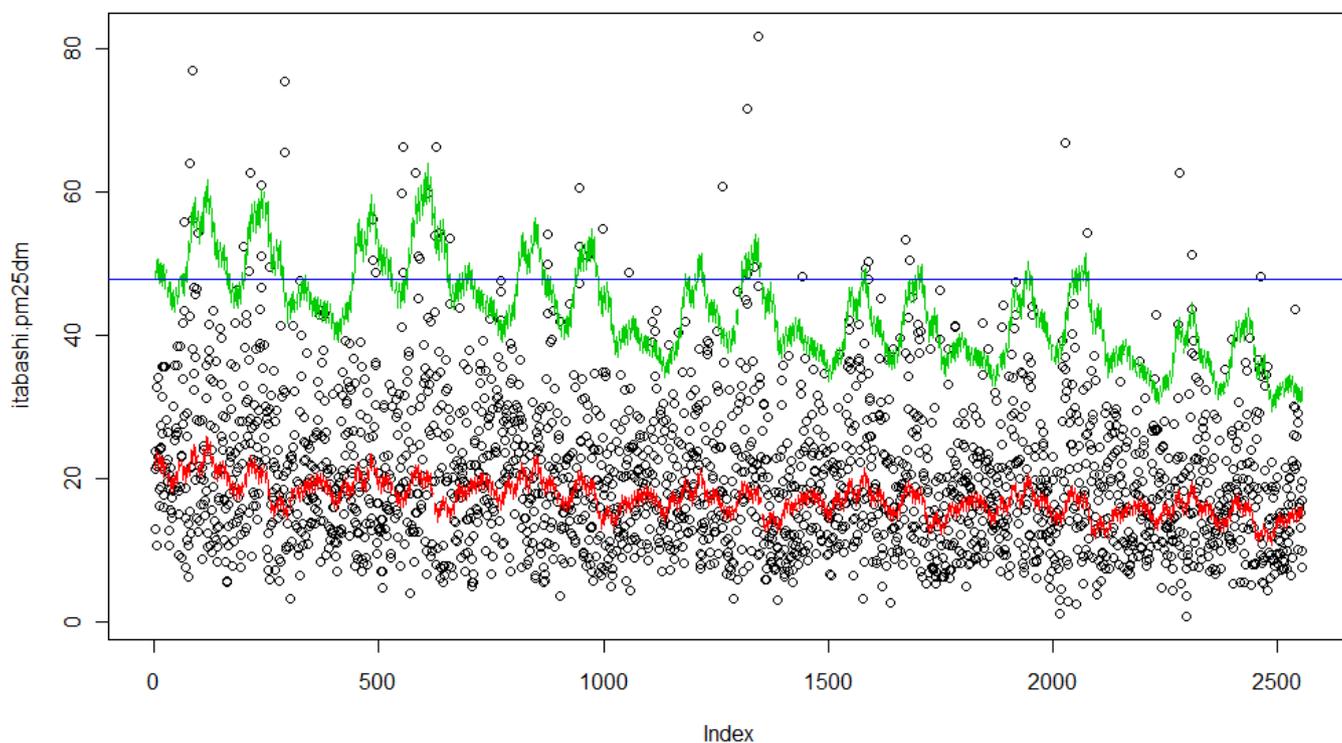


図 1 5 a 板橋の $PM_{2.5}$ の 24 時間平均値の変動期待値と上限値

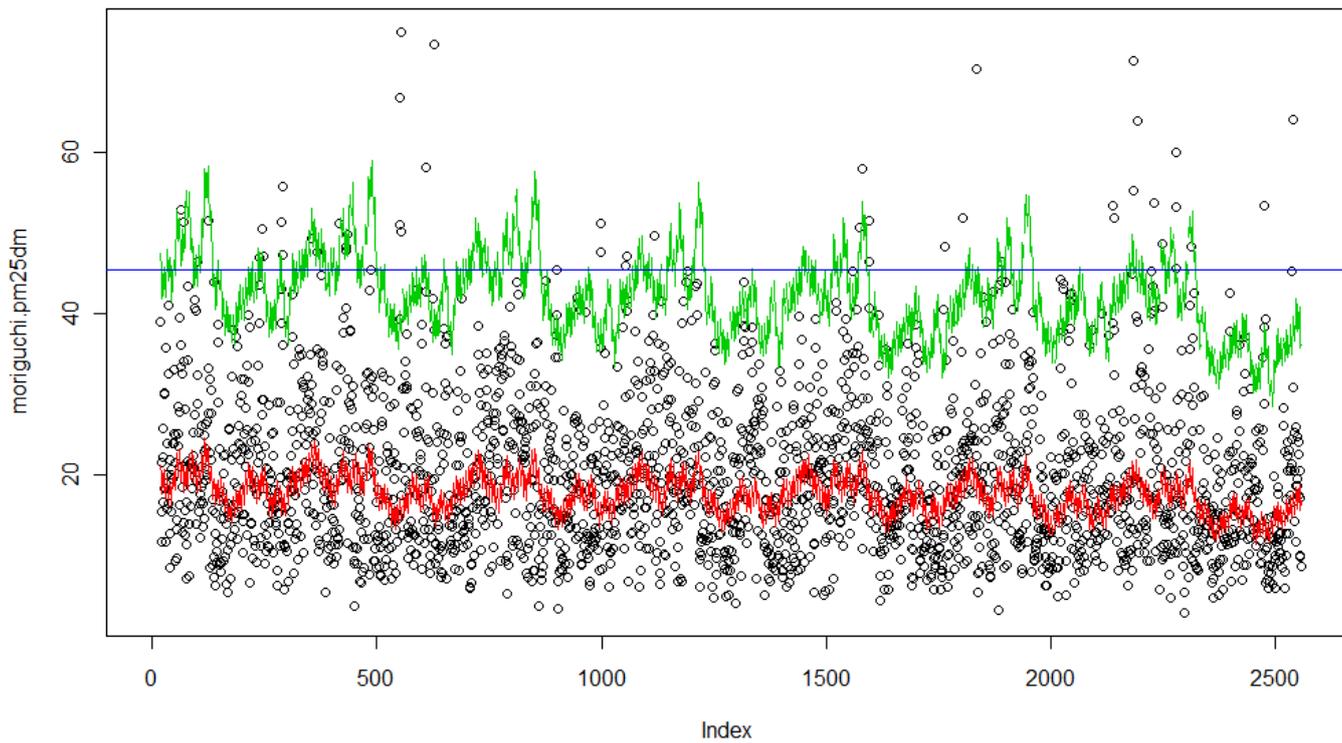


図 1 5 b 守口のPM_{2.5}の24時間平均値の変動期待値と上限値

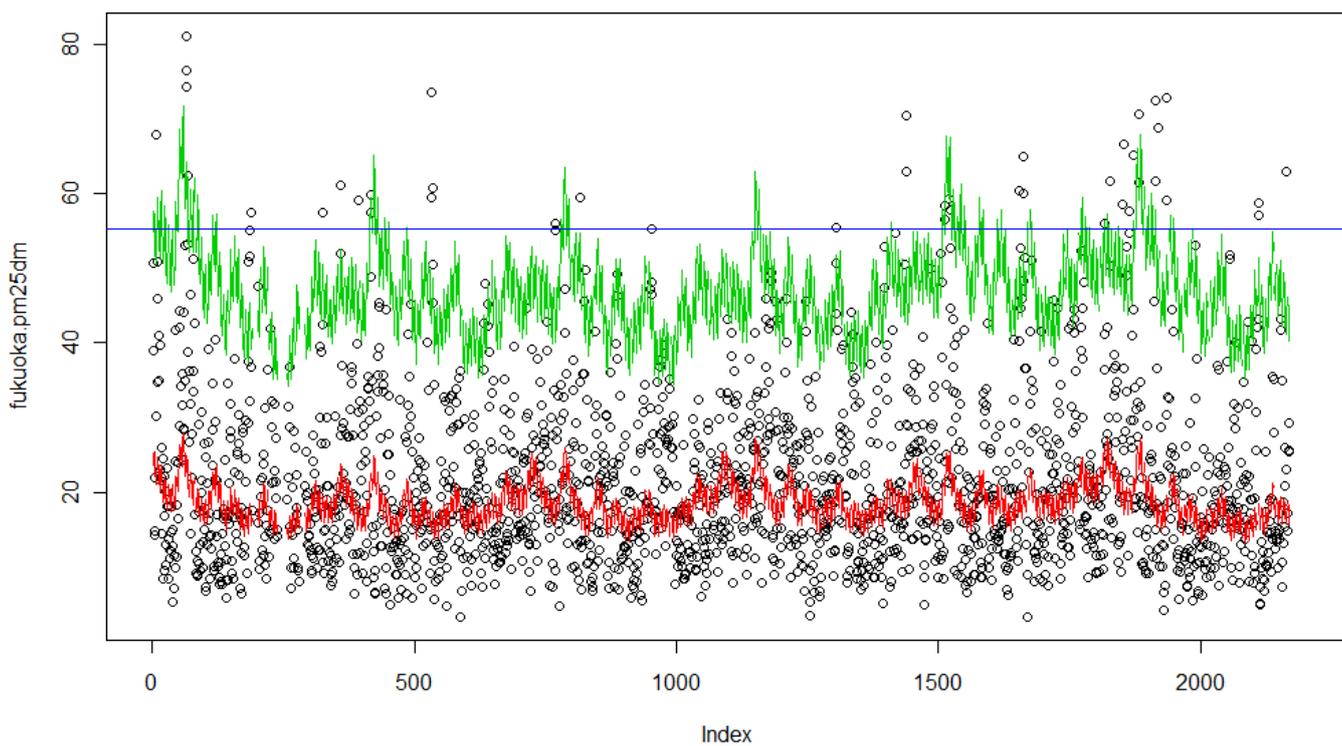


図 1 5 c 福岡のPM_{2.5}の24時間平均値の変動期待値と上限値

7. 理論的補足

今回行った方法は、筑波大学大学院ビジネス科学研究科経営システム科学専攻において若浦雅融氏が修士論文作成に際し、気温時系列の期待値と分散について長期並びに周期トレンドを抽出するために開発した2段階一般化加法モデル当てはめに基づいている(若浦, 2004)。この方法は、一段階の方法としてみれば、統計数理研究所が開発したDecompやClevelandらのSTLとほぼ同一の考え方である。

モデル的には次のようなノンパラメトリック回帰モデルを $PM_{2.5}$ 立方根データにあてはめたことになる。

第一段階での加法モデル当てはめ:

$PM_{2.5}$ 立方根 = 一般平均 + 長期トレンド成分 + 24時間周期成分 + 1週間周期成分 + 1年間周期成分 + 誤差
ただし、誤差は平均0、分散 σ^2 の正規分布に従う。

この第一段階での一般化加法モデル当てはめに基づいて、 $PM_{2.5}$ 立方根値の予測値を算出し、予測残差 e を計算している。すなわち、 $e = PM_{2.5}$ 立方根値 - モデルに基づく $PM_{2.5}$ 立方根値の予測値である。

この残差 e を用いて第2段階の一般化加法モデルが当てはめられている。

第二段階での一般化加法モデル当てはめ

$\log(E[e^2]) =$ 一般平均 + 長期トレンド成分 + 24時間周期成分 + 1週間周期成分 + 1年間周期成分
ただし、 e^2 は、変動係数 ϕ のガンマ分布(カイ二乗分布)に従う。

この2段階のノンパラメトリック回帰モデル当てはめによって、

時点 t の $PM_{2.5}$ 立方根値の期待値の推定値 m_t と予測残差二乗の期待値 s_t^2 が推定される。従って、 $PM_{2.5}$ 立方根値の近似的95%予測区間は、 $m_t \pm 2s_t$ となり、更に $PM_{2.5}$ の95%予測区間は $(m_t \pm 2s_t)^2$ となる。これが、図13に示された予測上限の求め方である。

この中で、予測残差二乗の期待値自体の評価が残差二乗時系列の平滑化というノンパラメトリックな方法で求められていることには注意が必要である。

参考文献

若浦雅嗣(2004)一般化加法モデルを用いた気温過程の平均・分散構造解析, 応用統計学 Vol. 33, pp. 181-200.