

小売販売データを用いた ARMAX モデルによる福岡市の一般廃棄物量予測

九州大学工学部 学生会員 瀨地雅也
九州大学大学院 フェロー会員 島岡隆行 正会員 杉崎康弘

1.はじめに

現在、地球は資源枯渇、地球温暖化などの問題に直面している。これらを解決する一つの方法として製品の製造、消費、廃棄物の排出・収集、中間処理、再資源化、最終処分に至る一連の流れを、プラットフォームにて一元管理、活用し、廃棄物の処理及び資源循環の最適化を行うことが挙げられる。本研究はこのプラットフォームの動脈にあたる小売の販売量と静脈にあたる一般廃棄物の量の関係性を明らかにし、AIによってその量を予測するものである。

2.データについて

本研究で使うデータは福岡市の一般廃棄物量の日ごとのデータと九州での小売業における 100 店舗当たり、1 日当たりの食品の販売個数(菓子類、加工食品、生鮮食品)である。各データの時系列グラフを図 1 に示す。

2-1 データの関連性

まず本研究で予測していく廃棄物量の自己相関及び偏自己相関を図 2 に示す。特に 7 日単位で強い相関があることが分かる。次に廃棄物量と販売個数の関係を確認するため相互相関を図 3 に示す。どの商品も -1 のところで 0.4~0.6 ほどの正の相関を持っていることが分かる。つまり商品を買ってそれがゴミになるまで一日のラグがあることが分かる。

3 予測手法

3-1 外生変数を加えた自己回帰移動平均モデル (ARMAX モデル)

本研究で予測する廃棄物量は自己相関、偏自己相関ともに相関があったので自己回帰モデル(AR モデル)と移動平均モデル(MA モデル)を組み合わせる。また販売個数をモデルに組み込むために外生変数 X を取り入れた ARMAX モデルを採用した。また日曜日と 1/1~4 は廃棄物量が 0 のためその影響もダミー変数としてモデルに組み込んだ。

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{k=1}^2 \beta_k x_{k-t} + c + \varepsilon_t$$

式 1ARMAX(p,q)モデル



図 1 各データの時系列グラフ(標準化)
(上から菓子類、加工食品、生鮮食品、廃棄物量のデータ)

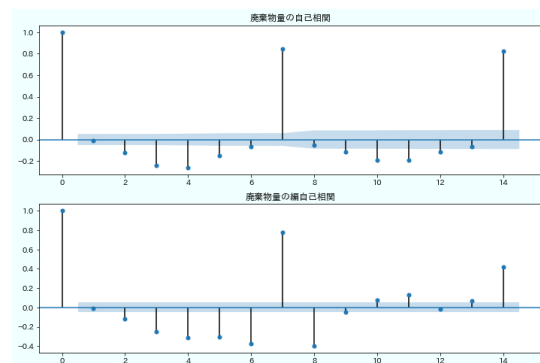


図 2 廃棄物量の自己相関(上から自己相関、偏自己相関)

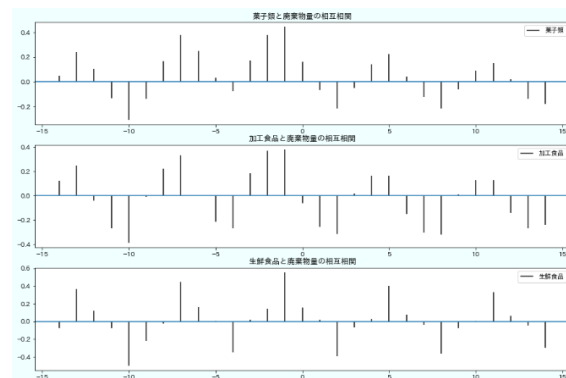


図 3 廃棄物と販売個数の相互相関(上から菓子類、加工食品、生鮮食品)

左側の項から AR 過程、MA 過程、外生変数、定数項、

誤差項となっている。また外生変数は販売個数とダミー変数の2つとなっている。p,q というのはどこまで遡って値を考慮するかというパラメータである。

3-2 パラメータ決定

パラメータ決定は AIC(赤池情報量基準)を用いた。
 $AIC = -2\log L + 2(k+1)$

式 2 AIC(赤池情報量基準)

L はモデルの最大対数尤度、k は説明変数の数であり、AIC が小さければ平均的に予測能力が高いモデルということを表す。表 1 を見ると小売データなし、ダミー変数ありのモデルの AIC が一番小さくなっている。つまりこのモデルが一番平均的な予測能力が高いということになる。このモデルのパラメータ p=9,

	ダミーなし	ダミーあり
菓子類	293.164	-203.901
加工食品	298.500	-202.484
生鮮食品	288.800	-57.632
販売個数なし	288.841	-208.692

表 1 各モデルの最小 AIC

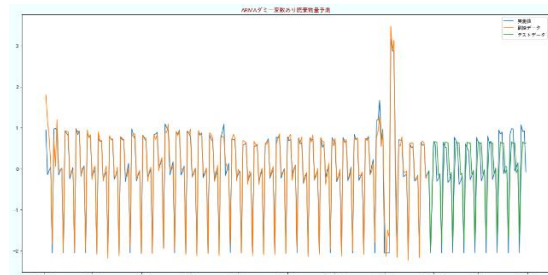


図 4 小売データなしダミー変数ありの廃棄物量予測

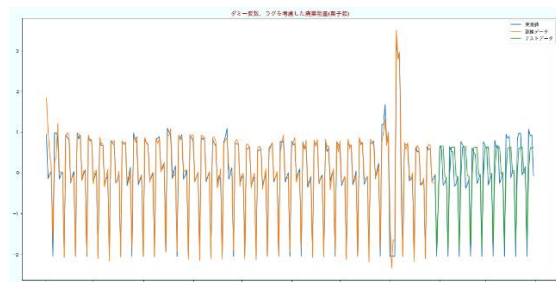


図 5 菓子類ダミー変数ありの廃棄物量予測

	ダミーなし	ダミーあり
菓子類	1.31705	0.48293
加工食品	1.35530	0.49057
生鮮食品	1.32546	0.89295
販売個数なし	1.33001	0.49700

表 2 各モデルの予測誤差

q=7であった。またデータのうち8割を訓練データ、

2割をテストデータとして使用した。

4 予測結果と考察

4-1 評価方法

予測結果の評価手法としては二乗平均平方根誤差 (RMSE)を利用した。この値が小さければ真値と予測値が近いということになる。

$$\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

式 3 二乗平均平方根誤差(RMSE)

4-2 評価結果

予測結果を図 4,図 5 及び表 2 に示す。この 8 種類のモデルの中で予測精度が高いのは RMSE=0.48293 で菓子類、ダミー変数ありのモデルとなった。

4-3 結論及び考察

- 1) 小売の販売量と廃棄物量には 0.4-0.6 の正の相関があり、商品を買って約 1 日で廃棄物になっている。
- 2) 生鮮食品と 1 日前の廃棄物量のデータ間には比較的大きい 0.6 の相関が見られた。一方、ARMAX モデルを用いた予測結果は他の外生変数のモデルと比較して誤差が大きかった。この原因は生鮮食品と AR(1)過程、もしくは MA(1)過程との間で多重共線性が起こり、予測能力が下がったと考えられる。
- 3) AIC の値が小さいモデルが必ずしも予測精度が良くはなかった。この理由は、AIC は平均的な予測能力を示しているので、データによっては予測精度と AIC の結果が一致しない時があるためと考えられる。

今後の課題として天気やイベントなど廃棄物量を変動させる要因をモデルに組み込み、予測精度を上げることが挙げられる。その場合には静脈における資源循環の最適化を目指す。

5 謝辞

本研究を遂行にするにあたり、福岡市より一般廃棄物(家庭から出るごみ)の重量データ及び株式会社マーチャング・オン社より RDS 市場データを利用させていただきました。ここに記して、謝意を表します。