

## PS-21 船舶寄港情報によるコンテナ船の積載貨物量の推計

知識・データシステム系 \*小坂 浩之

### 1. はじめに

近年、船舶自動識別装置 (Automatic Identification System, 以後 AIS とする) から取得されるデータは、海運や造船のサービスを向上させることを目的として活発な利用が進められている。活用事例の一つとして、AIS データに基づき貨物流動の把握をすることが挙げられる<sup>1)2)</sup>。AIS データに基づく船舶の積載貨物量の把握は、タンカー、バルクキャリアを対象にした試み<sup>3)4)5)6)7)</sup>が存在するが、コンテナ船の分析は見当たらない。本研究は、統計的手法や機械学習の分野で発展が著しい欠損値補完の枠組みにより、コンテナ船の積載貨物量の推計を行う。

### 2. データセット

#### 2.1 データソース

カナダの Descartes 社は、北米の港に寄港する船舶の船荷証券等を収集して、詳細な貨物データベースを作成している。本研究では、Descartes データから得られた各コンテナ船の実際の米国輸入積載貨物量を使用する。Descartes データには、寄港した船舶の仕向国、到着日、出発日、貨物サイズ、相手国等が含まれているが、寄港時の到着時刻、出発時刻、喫水は含まれていない。そこで、本研究では、AIS データに基づく Llyod's List Intelligence (LLI) の船舶動静データから、コンテナ船の到着時刻、出発時刻を、米国の The Navigation and Civil Works Decision Support Center (NDC) の船舶寄港データから、コンテナ船の喫水を確認することで、データセットを作成した。本研究では、アジア 17 各国・地域から出港し、2015 年に米国港湾に入港したコンテナ船の寄港記録に関して、2,087 件のデータセットを作成した。

#### 2.2 AIS データ

AIS データに基づいて積載貨物量を推定し、貨物の流れの構造を理解することには利点を述べる。第 1 の利点は、即時に情報を得られることである。AIS データをリアルタイムで提供するサービスもあり、月次値が数ヶ月後に発表される一般的な貿易統計よりも早く貨物情報を入手することが可能である。第 2 の利点は、AIS では、貨物の流れと個々の船舶の運航特性 (時間、距離、速度など) や船舶の仕様 (サイズ、エンジン出力など) を関連付けて船舶を特定できることである。例えば、船舶の正確な輸送距離に基づいて、貨物トンマイルを算出することができる。第 3 の利点は、世界中で収集された AIS データにより、限られた国や地域でしか公表されていない船舶の積載貨物量の情報を補完し、グローバルなデータへと発展させることができる。以上のような認識から、AIS データを用いて船舶の積載貨物量を推定することが必要

と考えられる。

#### 2.3 データの特性

データセットに含まれる各変数と積載貨物量の関係を示すために、積載貨物量を目的変数、その他の変数を説明変数として、回帰分析によりパラメータを推定した。各種の説明変数を検討した結果、船舶の Dwt、船舶の設計喫水と米国入港時の最大喫水の比である Draft rate、米国港湾での停泊時間である Port time が有効な説明変数となった。説明変数を変えた 2 つのモデルの推定結果を表 1 に示す。船舶の Dwt が、積載貨物量の説明変数としての寄与が大きいことがわかる。このデータセットを用いた回帰分析では、決定係数 (R<sup>2</sup>) が 0.25 程度の説明力である。説明変数の加工や追加により、回帰分析の推定能力が向上すると考えられるが、本研究は上記の説明変数で積載貨物量の推計を行う。

表 1 回帰分析のパラメータ推計結果

Variables	Model	
	1	2
Dwt [ton]	0.3975 (25.8)	0.3982 (24.5)
Draft rate		9650 (2.4)
Port time [day]		241.9 (3.0)
Constant	-7069 (-5.3)	-15540 (-4.3)
R <sup>2</sup>	0.24	0.25
Mean VIF	-	1.1

( ): t 値

### 3. コンテナ船の積載貨物量の推計

#### 3.1 推計方法

本研究は、欠損値補完法<sup>8)</sup>の枠組みを利用して、コンテナ船の積載貨物量の推計を行う。具体的には、2,087 件のデータセットのうち、無作為に選ばれた 1,000 件のデータは、欠損のない完全なデータと仮定する。残りの 1,087 件のデータについては、積載貨物量のデータを削除し、それを欠損データとする。本研究では、完全データにおいて、積載貨物量を目的変数とし、説明変数から目的変数を推定するモデルを作成する。説明変数は、比較を容易にするために、上述の回帰分析で用いた変数とする。モデルと欠損データの説明変数を使用して、欠損値である積載貨物量を推計する。また、推計された積載貨物量を実際の積載貨物量と比較することで、モデルの妥当性を検証する。欠損値補完法としては、回帰分析 (Regression) や多重代入法 (MI) などの統計的手法と、機

機械学習に位置づけられる人工ニューラルネットワーク (ANN) と k 近傍法 (KNN) を適用する。機械学習では、完全データが学習データに相当し、欠損データがテストデータに相当する。ANN については、隠れ層のノードが 2 の場合 (ANN1) と、隠れ層のノードが 5 の場合 (ANN2) を考える。k 近傍法<sup>9)</sup>は、モデルを作成することはなく、学習データをベクトル空間上にプロットし、説明変数が与えられた時に、学習データから距離が近い順に k 個を見つけて、多数決で目的変数の値を決定するものである。本研究では、与えられた説明変数に対して、距離が近い 3 個の学習データにおける積載貨物量の平均値を k 近傍法の推計値とした。

### 3. 2 推計精度の評価方法

推計精度の評価は、データセットを構成する船舶 ( $v$ ) の集合 ( $A$ ) において、積載貨物量の実測値 ( $y_v$ ) と推定値 ( $\hat{y}_v$ ) から、以下に示す平均平方二乗誤差 (RMSE) を計算することで行う。

$$RMSE_A = \sqrt{\frac{1}{|A|} \sum_{v \in A} (y_v - \hat{y}_v)^2} \quad (1)$$

本研究では、上述の無作為抽出を 100 回行い、各回のモデルで算出される 100 個の RMSE から平均値を算出し、無作為抽出の不確実性を減少させ、推計精度の評価を行う。

### 4. コンテナ船の積載貨物量の推計結果

図-1 は、各モデルの推定値と実測値から RMSE の平均値を算出した結果である。回帰分析の RMSE は、0.947 である。統計的手法の MI は 0.875 である。機械学習の ANN1 が 0.861、ANN2 が 0.856、k 近傍法が 0.506 である。ベンチマークとして考えることができる回帰分析と比較すると、特に k 近傍法の RMSE は小さく、積載貨物量の実測値と推定値の間で、相違が少ないことがわかる。図-2 は、k 近傍法により、欠損値を推計した結果について、積載貨物量の実績値と推計値の比較を示している。実績値と推計値の R2 値は、0.93 であり、積載貨物量の傾向が捉えられていると考えられる。今後、採用されなかった変数等を詳細に検討することにより、推計精度の向上が期待できる。

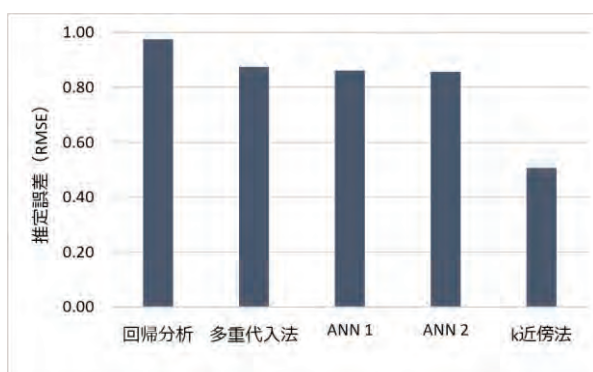


図-1 各推計方法における推計誤差 (RMSE)

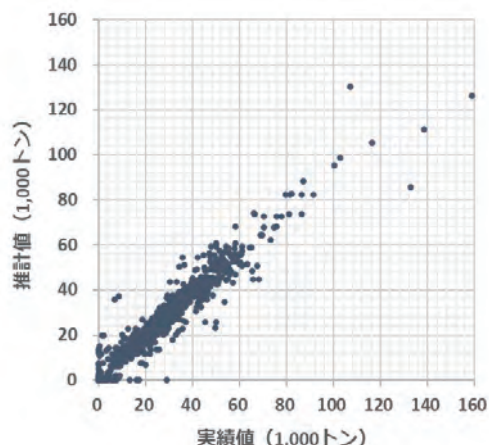


図-2 k 近傍法による実績値と推計値の比較

### 5. おわりに

本研究は、コンテナ船の積載貨物量の推計を欠損値補完の問題として検討を行った。一般的な手法である回帰式を含めた複数の推計手法の内、k 近傍法が有効となった。推計精度の向上のために、変数の詳細な検討を今後進める。また、欠損値補完は、データ特性によって、有効な手法が異なるため、他の手法の検討も必要である。今後は、有効なモデルを開発した後に、十分なデータが得られないアジア地域を対象にして、積載貨物量の推定を行う予定である。

### 参考文献

- 1) Svanberg, M., Santén, V., Hörteborn, A., Holm, H., Finnsgård, C., 2019, AIS in Maritime Research, Marine Policy, 106, Article 103520.
- 2) Yang, D., Wu, L., Wang, S., Jia, H., Li, K.X., 2019, How Big Data Enriches Maritime Research-A Critical Review of Automatic Identification System (AIS) Data Applications, Transport Reviews, 39(6), 755-773.
- 3) Adland, R., Jia, H., Strandenes, P. S., 2017, Are AIS-based Trade Volume Estimates Reliable? The Case of Crude Oil Exports, Maritime Policy & Management, 44(5), 657-665.
- 4) Arslanalp, S., Marini, M., Tumbarello, P., 2019, Big Data on Vessel Traffic: Nowcasting Trade Flows in Real Time, IMF Working Paper, WP/19/275.
- 5) Jia, H., Prakash, V., Smith, T., 2019, Estimating Vessel Payloads in Bulk Shipping Using AIS Data, International Journal of Shipping and Transport Logistics, 11(1), 25-40.
- 6) Arifin, M. D., Hamada, K., Hirata, N., Ihara, K., Koide, Y., 2018, Development of Ship Allocation Models Using Marine Logistics Data and Its Application to Bulk Carrier Demand Forecasting and Basic Planning Support, Journal of the Japan Society of Naval Architects and Ocean Engineers, 27, 139-148.
- 7) Zhou, X., Hu, Q., 2019, Estimation of Shipment Size in Seaborne Iron Ore Trade, TransNav: International Journal on Marine Navigation and Safety of Sea Transportation, 13(4), 791-796.
- 8) Enders, C. K., 2010, Applied Missing Data Analysis, Guilford Press.
- 9) Hastie, T., Tibshirani, R., Friedman, J., 2009, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. Springer.