Spatial temporal analysis of major seaport freight flows in India

Prasanta K. Sahu, Ph.D.
Corresponding Author
Assistant Professor, Department of Civil Engineering
Birla Institute of Technology and Science Pilani
Pilani, Rajasthan, India – 333031
Phone: (91) 962345 1628
Email: prasantsahu222@gmail.com

Abhishek Padhi, B.E.
Former Under Graduate Student, Department of Civil Engineering
Birla Institute of Technology and Science Pilani
Pilani, Rajasthan, India – 333031
Phone: (+91) 97722 2341
Email: abkppb@gmail.com

Gopal R. Patil, Ph.D.
Associate Professor, Department of Civil Engineering
Indian Institute of Technology Bombay
Powai, Mumbai, India – 400076
Phone: (+91) 961958 8606
Email: gpatil@iitb.ac.in

Gangadhar Mahesh, Ph.D.
Assistant Professor, Department of Civil Engineering
National Institute of Technology Karnataka
Surathkal, Karnataka, India – 575014
Phone: (+91) 948312 1039
Email: gangadhar.mahesh@gmail.com

Ashoke K. Sarkar, Ph.D.
Senior Professor, Department of Civil Engineering,
Birla Institute of Technology and Science Pilani
Pilani, Rajasthan, India – 333031
Phone: (+91) 900129 6698
Email: aksarkar@pilani.bits-pilani.ac.in

Number of Words: 5990 words + 5 Tables + 1 Figures = 7490 words

Revision Submitted to Transportation Research Board 96th Annual Meeting and Possible Publication in Transportation Research Record

Date of Submission: November 14, 2016
Abstract

This paper discusses the space time interactions among the freight flows through major ports in India by applying a spatial temporal model to estimate the freight flows at the study ports. Freight flow data at regular intervals in the form of spatial time series are collected for the twelve major ports located along the east and west coast of India. The system of freight flows is modeled thorough interactions both in time and space dimensions as a multivariate stochastic process. K-nearest neighbor method is used to find the nearest influential port for each port. The effect of the neighbor port freight on a subject port freight was analyzed with two proposed models known as space time autoregressive (STAR) model and space time moving average model. The model performances are checked through the measures of effectiveness such as mean absolute error (MAE) and root mean square error (RMSE). Lower MAE and RMSE values from STAR (1,1) model suggested better performance. This model was further utilized for demand elasticity analysis. The demand elasticities were used to understand the degree of dependency among the competing/non-competing ports. The elasticity analysis revealed that all the ports are more sensitive to changes in their own demand over time than the corresponding spatial changes indicating bulk of the demand dependency on hinterland economic activity. The crossed elasticity values are positive for all the ports except Mormugao and Mangalore. The negative demand elasticity between these ports suggested that these two ports are competing with each other as they share common hinterland in two coastal states: Goa and Karnataka. The proposed models can be used for short-term forecasting of freight flows through the major ports and assessing the impact of freight flow changes from one port to the nearest neighboring port. It is expected that this study will help port authorities and policy makers for holistic development of port system by making right investments in required locations to promote balanced development. It has also implications towards formulating policies on port development considering the importance of PPP mode of infrastructure development.

1. Introduction

Seaports are the most important hubs for transporting domestic and international freight in any nation across the world. There are 13 major ports, and 200 non-major ports along the 7517 kilometers’ coast line in India. Most of the overseas commercial trade (95% in terms of value) takes place through the Indian port system (I). In fiscal year (FY)2014–15, the total traffic volume handled by the Indian ports was 1052.1 million tons, out of which 55.25% (581.3 million tons) of total traffic was handled by major ports and the remaining 44.75% (470.9 million tons) was handled by non-major ports. The overall annual growth rate of freight traffic was 7.07% between FY2005–06 and FY2014–15, with traffic at major and non-major ports grew at a rate of 3.58% and 13.94%, respectively (2). The foreign trade exchanges have been increasing through the port system in India because of the direct implications from the nation’s foreign policy, in particular large scale investment through foreign direct investment (FDI). Bilateral agreement like Indo-EU Maritime Agreement and recent trilateral agreement between India, Iran and Afghanistan, etc. have been playing an important role in continuously increasing international trade. The ‘Make in India’ initiative aimed to accelerate the manufacturing sector in India, demands better infrastructure for the export of goods manufactured to various parts within the country and outside India. It is expected that by 2025, the ports will be required to handle a cargo of 2500 million tons per annum (MPTA) while the current port capacity in India is 1500 MPTA
(2). With the increase in the international and domestic trade volume, it is expected that the freight handled by the infrastructure sector, especially the maritime sector, will increase in a significant manner.

The increase in international trade and commercialization has put considerable pressure on the ports, operating environment, infrastructure and authorities. It has enhanced competition among the Indian ports to increase their share in the nation’s world trade. For example, acquiring higher capacity vessels with state of the art technology or sophisticated container ships will ensure the port for optimized freight transport operation. The addition of new ports, augmentation of capacity in existing ports, expansion of port associated infrastructure, adaptation of novel, reliable and improved technology, etc., requires huge sunk cost either from the port authorities or government. Therefore, it is crucial to carry out several analyses from various perspectives such as identifying probable interactions between freight flows among neighboring ports, operational-, economic-, and environmental impact analysis. The maiden step in understanding the multiple impacts is to identify the interaction among in port freight flows. The identification of such an interaction will help in estimating the impact of operational changes in the system of ports such as capacity augmentation in one or more ports. In other words, the interaction measured in terms of crossed elasticities can be used to estimate how changes in freight-flow patterns in some specific port locations are affected the freight demand at other ports in the system. Additionally, this study will be helpful for short term forecasting of freight demand.

A space time model based on space–time autoregressive moving average (STARMA) processes is used to derive the interaction between freight flows at neighboring ports. The proposed spatial-temporal model is further used for elasticity analysis to study the interaction between the Indian ports. The study ports are limited to major port systems due to limitation in data availability. This paper has six sections out of which this is first section: ‘Introduction.’ Section 2 contains the literature review where the past studies in various fields using STARMA model class are briefly discussed. The conceptual theory of the methodology adopted in this study is explained in Section 3. The spatial-temporal model development process is thoroughly discussed in section 4. Section 5 discusses about the elasticity analysis results obtained from the calibrated model. Section 6 concludes the paper.

2. Literature Review

The space–time autoregressive moving average (STARMA) model class was introduced by Pfeifer and Deutsch (3,4,5) in the early 1980s. This model class in an extension of autoregressive moving average (ARMA) model (6). Later space-time autoregressive integrated moving average (STARIMA) models were developed and used by researchers in their studies for forecasting as well as analysis purposes. These models are used in several research areas such as traffic engineering, economics, agricultural systems, biological system, ecology, neurological science, etc. STARIMA models are used to describe, analyze, and forecast a set of N observable time series. Space-time models best represent interaction among the neighboring regions of a system by considering the spatial correlation among the regions. Such models attempt to describe the dependencies between the regions of a systems across space and time.

Studying the relation and influence of neighbors with each other in a defined system has been an important research area for the last two decades. Several studies have been conducted in the past based on space-time modeling of time series data. Pfeifer and Deutsch (4) used the STARMA model to demonstrate prediction efficiency. Kamarianakis and Prastacos (7) developed
STARMA models using traffic data obtained at different places in Athens (Greece) to predict the space-time stationary traffic flow and assess the impact on other regions of the network. Lin et al., (8) conducted similar studies for predicting the short-time interval traffic volume prediction in Shanghai. The study showed that the model can be used to estimate a non-linear function. Min et al., (9) forecasted short-time traffic flow on urban intersections in Beijing, China with a combinations of STARIMA model and dynamic turn ratio prediction model. This combination of models helped in enhancing the efficiency and the forecasting performance. In a recent study, Khan et al., (10) studied the performance of STARIMA models in predicting the travel times for dense and highly varying road traffic networks in Sydney, Australia. The performance of the developed STARMA model was later analyzed for six Levels of Services (LOS) of the roads.

Reynolds and Madden (11) summarized the theory of spatial-temporal correlation analysis on a spatial-temporal dataset. The dataset consisted of temporal data of disease spread over a spatial lattice structure, which was used for description of epidemics development. Madden et al., (12) used STARIMA process to model the tobacco virus epidemics across six different fields in Kentucky, USA. Epperson (13) used ecological data to characterize the genetic population and estimate migration rates based on space time correlations. Soni et al., (14) used intervention analysis and STARMA modeling to develop a method to remove the interfering signals present in fetal bio-magnetic signals.

STARMA modeling process is also used for water resources planning and management. Dalezios and Adamowksi(15) developed a spatio–temporal moving average (STMA) precipitation model to analyze the regional meteorology and hydrology in rural watersheds. The STMA model was identified as the best fit to the hydrology dataset. Szummer and Picard (16) used spatial temporal autoregressive (STAR) model to model the image sequences of temporal textures (i.e., textures with motion). It is used in recognition and synthesis since a pixel can be generated as a combination of its neighbor pixels lagged in space and time.

Zhou and Buongiorno(17) developed a space-time econometric model for pine saw timber price predictions for 21 neighboring regions in southern USA. The researchers also performed an impulse response analysis to see if the prices are globally integrated. Niu and Tiao(18) modeled the satellite ozone data using STAR model. The satellite data was analyzed on a fixed latitude, which takes into account the temporal and longitudinal dependence of the observations. The space-time regression models are applied for trend assessment of monthly column ozone for each of the subdivided geographical blocks. Giacinto(19) developed a generalized space-time model to improve the model’s ability to cope with the spatial features of data. It also includes an application to regional unemployment analysis in Italy aimed at evaluating how a shock in a region propagates spatially extent in relation to unemployment and finding the degree of spatial heterogeneity in the process parameters. A summary from some of the above discussed studies is presented in Table 1.

Application of spatial temporal modeling mechanism on freight flow is limited to a study conducted by Garrido(20). Garrido analyzed space time interaction between truck flows through eight bridges located on Texas-Mexico Border. Also recent studies related to Indian major port systems are limited to Sahu and Patil(21), Sahu et al.,(22), and Patil and Sahu(1.23). Their study mainly focused on estimating freight demand for Indian ports using regression and time series modeling techniques. Their study models include ARIMA, SARIMA, VAR, and dynamic regression models. However, their models did not consider the effect of neighboring port freight
demand while estimating the demand for a subject port. In other words, their research did not analyze the spatial interaction between freight flows at the study ports. In the present research, the spatial temporal variation in freight flow at a neighbor port is considered while modeling the freight at a specific port.

3. Study Site Locations

The twelve study ports are strategically located along the east and west coast to facilitate international trade and commerce of the nation. Along east coast Kolkata, Paradip, Vizag, Chennai, Tuticorin and Ennore are located; whereas Mangalore, Cochin, Mormugao, Kandla, Mumbai and JNPT are placed along west coast. The geographical positions of the ports and their rankings during 2015-16 are presented in Figure 1.

Figure 1: Geographic location of study ports
Table 1: Key summary from previous studies

<table>
<thead>
<tr>
<th>Author(s) and research focus</th>
<th>Data set and source</th>
<th>Model</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pfeifer and Deutsch, 1980 (4); Regional Forecasting</td>
<td>Annual data of percentage of farms with tractors of 25 states in USA</td>
<td>STARIMA, Casetti and Semple-, Cliff and Ord model</td>
<td>STARIMA showed better results than the other two models</td>
</tr>
<tr>
<td>Kamarianakis and Prastacos, 2004 (7); Traffic Flow Forecasting</td>
<td>Traffic flow averaged for 7.5 minutes from 25 loop-detectors. Ministry of Environment and Public Works, Athens, Greece</td>
<td>STARIMA</td>
<td>Entire traffic network was modeled and predicted using a single model</td>
</tr>
<tr>
<td>Min et al., 2009 (9); Traffic Flow Forecasting</td>
<td>Traffic flow data from 50 sensors Beijing Traffic Management Bureau, China</td>
<td>STARIMA + Dynamic turn ratio prediction model</td>
<td>Combined model enhanced forecasting efficiency and performance</td>
</tr>
<tr>
<td>Lin et al., 2009 (8); Traffic Flow Forecasting</td>
<td>Traffic flow data from 78 loop detectors. Public Works in Shanghai, China</td>
<td>STARIMA</td>
<td>Better performance for short term traffic flow forecast</td>
</tr>
<tr>
<td>Khan et al., 2012 (10); Travel Time Prediction</td>
<td>Data from urban network using Loop detectors. Roads and Traffic Authority (RTA), Sydney Australia.</td>
<td>STARIMA</td>
<td>Model was well fit for steady free flow settings but was inefficient for urban settings when compared over different levels of service</td>
</tr>
<tr>
<td>Garrido, 1998 (20); Truck Flow Modeling</td>
<td>Monthly truck traffic count from eight bridges. Texas A&amp;M University and US Customs Authority, USA</td>
<td>STARMA</td>
<td>Critical bridges were identified</td>
</tr>
<tr>
<td>Szummer and Picard, 1996 (16); Temporal Texture Modelling</td>
<td>Image data</td>
<td>STAR</td>
<td>Model represented several temporal textures, and enabled good synthesis and compression</td>
</tr>
<tr>
<td>Dalezios and Adamowski, 2009 (26); Precipitation Modelling</td>
<td>Precipitation data from eleven rain gauge stations</td>
<td>STMA</td>
<td>Better prediction results</td>
</tr>
<tr>
<td>Zhou and Buongiorno, 2006 (17); Space-Time Timber Price Prediction</td>
<td>Quarterly data on timber prices Timber Mart South, Norris Foundation, USA</td>
<td>STARMA</td>
<td>Results showed that the nearest neighbor were the highest effected due to sudden change in price</td>
</tr>
</tbody>
</table>
4. Methodology

The freight flows through the ports may be realized as a system with multivariate time series and spatial dependence on the neighbor ports. This kind of system can be best represented using a space–time dependent model (3,4,5). The space-time model will help in understanding the interaction between the freight flows. The methodology for the present study is of several steps:
1) Determination of nearest neighbors to a subject port; 2) Modeling framework which consist of three stage model building procedure: model-, identification, estimation, and diagnostic checking; 3) Elasticity analysis to understand the possible interaction among port freight flows.

The nearest neighbors to a specific port is determined by adopting a very popular data mining method known as k-Nearest Neighbors (k-NN) method. k-NN method is used to find the k-Nearest Neighbors of a point. The input given to the model is an input data set with N training data points in the feature space. When a vector \( q \) is given to the k-NN model, it finds the k nearest neighbors of the given data point based on some distance measures such as Euclidean, Manhattan, Minkowski, etc. In this study, the nearest neighbor was found using the 'Great Circle Distance' concept. The Great Circle Distance concept is used to find the shortest distance between two points in a sphere along the surface of the sphere. Since we used longitude and latitude of port locations to find out the nearest neighbors of a port, this concept was the best choice. It is suggested the readers to refer Statsoft and Hardle(24,25) for more details about this method as this is not the main focus of the present research.

4.1 Modeling Framework

The STARMA model represents the time series models which are linearly dependent and lagged in both space and time. Let \( z_{it} \) represents the observations of random variable \( Z_{it} \) at \( N \) different locations in space (\( i = 1, 2, 3, \ldots, N \)) over \( T \) (\( t = 1, 2, 3, \ldots, T \)) time periods. The model is formulated with the concept of spatial lag operator. Let \( L^{(l)} \), the spatial lag operator of order \( l \), which satisfies the following two conditions as shown in Eq.(1) and (2):

\[
L^{(0)} z_{it} = z_{jt} \tag{1}
\]

\[
L^{(l)} z_{it} = \sum_{j=1}^{N} w^{(l)}_{ij} z_{jt} \tag{2}
\]

Where, \( \left\{ w^{(l)}_{ij} \right\} = \text{A set of weights} \)

The weight \( w^{(l)}_{ij} \) specifications are left to modeler’s decision. The present study assumes correlation coefficients between freight flows as weights. Euclidian distance between the ports decides the spatial order between them. The nearest ports to a port of interest are the first order neighbors. Second order neighbors are placed beyond first order neighbors, however nearer than third order neighbors. There are cases where we can have more than one neighbor in the same order. In such cases, the ideal weight is taken to be \( 1/n \), where \( n \) is the number of neighbors of same order. \( z_{jt} \) presents the linear combination of past observations and errors. In addition to the past observation at the same site, \( z_{jt} \) is also dependent on the past observation at the neighboring locations of definite spatial order. In the present study, locations refer to major port locations.
with respective latitude and longitude. With these definitions, the STARMA \((3,7)\) model is presented in Eq.\((3)\).

\[
Z_{it} = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl} \sum_{j=1}^{N} w_{ij}^{(l)} Z_{jt-k} + \sum_{k=1}^{q} \sum_{l=0}^{m_k} \theta_{kl} \sum_{j=1}^{N} w_{ij}^{(l)} a_{jt-k} + a_{it}
\]

Where \(Z_{it}\) is the freight flow through the port \(i\) during quarter \(t\), \(a_{it}\) is the random normal error associated with port \(i\) during quarter \(t\) such that the following condition is satisfied as shown in Eq.\((4)\).

\[
E[a_{it}a_{jt+s}] = \begin{cases} \sigma^2 & \text{if } i = j, s = 0 \\ 0 & \text{Otherwise} \end{cases}
\]

Where, \(E[.\] = Expected value

\(p=\) autoregression order

\(q=\) moving average order

\(\lambda_k=\) spatial order of the \(k^{th}\) autoregressive term

\(m_k=\) spatial order of the \(k^{th}\) moving average term

\(N=\) number of spatial units (ports)

\(\phi_{kl}\) and \(\theta_{kl}\) are the parameters

\(w_{ij}^{(l)}=\) level of interaction between the ports \(i\) and \(j\)

This model is referred as STARMA \((p_{\lambda_1,\lambda_2,...,\lambda_p}, q_{m_1,m_2,...,m_q})\). Two special cases of STARMA exists when either \(p=0\) or \(q=0\). The model becomes a STAR model if \(q=0\).

\[
Z_{it} = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl} \sum_{j=1}^{N} w_{ij}^{(l)} Z_{jt-k} + a_{it}
\]

The model shown is Eq.\((5)\) is referred to as STAR\((p_{\lambda_1,\lambda_2,...,\lambda_p})\) model. Similarly, for \(p=0\), only moving average terms remain and the model becomes STMA\((q_{m_1,m_2,...,m_q})\) model as presented through Eq.\((6)\).

\[
Z_{it} = a_{it} + \sum_{k=1}^{q} \sum_{l=0}^{m_k} \theta_{kl} \sum_{j=1}^{N} w_{ij}^{(l)} a_{jt-k}
\]

### 4.1.1 Model Identification

The first stage of the three-stage model building procedure is the model identification from the dataset. In the case of univariate time series models, we identify the best suited model from autocorrelation and partial-autocorrelation functions. However, in the case of space-time models the \(N^2\) possible cross-variances between all possible pairs of locations are combined to obtain proposed models. Analogous to the univariate time series models, the space-time models are identified using two-dimensional space-time autocorrelation function (STACF) and space-time
partial-autocorrelation function (STPACF). Possible model orders (p,q) for the three subclasses of space-time models (STAR, STMA, STARMA) are determined by checking the pattern of STACF and STPACF.

### 4.1.2 Model Estimation

Once the model is identified, then the optimal estimates for \( AR(\phi) \) and \( MA(\theta) \) is done using maximum likelihood estimation procedure (4). Unlike STMA and STARMA models, the STAR model parameters are estimated on the principle of linear regression theory (i.e., using conditional maximum likelihood estimation to find the least residual sum of squares). However, STMA and STARMA models are non-linear in nature. Hence, a 'gradient method' is used which uses Taylor’s Expansion to linearize the non-linear model.

### 4.1.3 Diagnostic Checking

Once a suitable model is selected and parameters are estimated, diagnostic checks are performed on the model to check if it exemplifies the data. Normally, a model may fail in two ways. Firstly, it has to be checked whether the residual of the fitted model has any significant correlation. Any observed correlation in the residuals is not desired. If the fitted model adequately represents the data should be distributed normally with mean zero, the variance-covariance matrix should be spherical. Secondly, the model can become complex. Hence, it is important to check if the parameters are statistically significant.

### 4.2 Elasticity Analysis

After the model is estimated and checked for statistical significance, the freight flows at different ports can be examined through elasticity analysis. In other words, the interaction among ports can be analyzed by determining the cross elasticities. The cross elasticities are used to understand the interaction between a particular port and its’ neighbors. It explains how the freight flow at a port is affected by the changes in freight flows in its neighbor ports. This, in turn, gives us an insight into whether the two nearest ports are competing or complementing. If the cross elasticity between two ports is negative, it means that the two ports are competing with each other to attract freight. If it is positive, it means that the freight volume of the target port increases/decreases with the increase/decrease in the volume of the neighbor port (i.e., they complement each other). Generally, the elasticity function is defined as follows.

\[
\epsilon_t(A/B) = (\partial A/\partial B) \ast (B/A); \quad \text{Where, } A, B \text{ are two continuous and differential variables.}
\]

The change in the freight flow of a port \( i \) in quarter \( t \) when the freight flow at port \( j \) is changed in quarter \( t-q \), is given by Eq. (7)

\[
\epsilon_t \left( \frac{z_{it}}{z_{jt-q}} \right) = \frac{\partial z_{it}}{\partial z_{jt-q}} \cdot \frac{z_{jt-q}}{z_{it}}
\]  

Case 1: Direct elasticity: In case \( i=j \), the Eq. (7) referred as direct elasticity. Taking spatial order is taken as 1 (\( l=1 \)) and applying Eq. (7) is to (3), the Eq.(7) gets converted to Eq.(8):

\[
\epsilon_t \left( \frac{z_{it}}{z_{it-q}} \right) = \alpha_s \frac{z_{it-q}}{z_{it}}
\]
Case 2: Cross elasticity: In case $i \neq j$, and applying Eq. (7) to (3) with spatial order 1, the Eq. (7) becomes Eq. (9):

$$
\varepsilon_l \left( \frac{z_{it}}{z_{jt-q}} \right) = \alpha_{s} w_{ij} \frac{z_{jt-q}}{z_{it}} \quad \text{for } i \neq j
$$

(9)

4.3 Data

This dataset consists of quarterly freight flows for twelve major ports of India between the periods of Quarter – I, 2001-2002 to Quarter – III, 2015-2016. The freight flows considered in this study consist of both inward and outward freight flows. Some key statistics related to port freight flow series are presented in Table 2. The port which handles the highest volume of freight in eastern and western coasts are Vizag and Kandla respectively. Mormugao port has a very high coefficient of variation (around 59%) because of the seasonal variation in the freight handled by the port. The ports in the western coast and eastern coast handle almost equal quarterly freight by volume, accounting for 50.1% (60.067 million tons) and 49.9% (59.817 million tons) respectively. The variability of freight volume for five of the six ports in the western coast lies between 19% and 32%, except for Mormugao whose variability has already been explained. Similarly, the variability for ports in the eastern coast lies between 17% and 34%, except for Ennore which has a high variability of 51.5%. The trend in the freight volume for most of the ports show an increasing trend, though the gradient for each of them varies.

### Table 2: Basic statistics for the freight flows through major ports in India

<table>
<thead>
<tr>
<th>Port</th>
<th>Average 1 (tons)</th>
<th>Std. Dev.</th>
<th>CV 2</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kandla (P1)</td>
<td>17.321</td>
<td>5.399</td>
<td>0.312</td>
<td>9.044</td>
<td>25.767</td>
</tr>
<tr>
<td>Mumbai (P2)</td>
<td>10.139</td>
<td>3.151</td>
<td>0.311</td>
<td>5.883</td>
<td>16.335</td>
</tr>
<tr>
<td>JNPT (P3)</td>
<td>12.701</td>
<td>3.736</td>
<td>0.294</td>
<td>5.444</td>
<td>16.997</td>
</tr>
<tr>
<td>Mormugao (P4)</td>
<td>7.529</td>
<td>4.416</td>
<td>0.587</td>
<td>1.353</td>
<td>19.208</td>
</tr>
<tr>
<td>Mangalore (P5)</td>
<td>8.088</td>
<td>1.631</td>
<td>0.202</td>
<td>3.788</td>
<td>10.289</td>
</tr>
<tr>
<td>Cochin (P6)</td>
<td>4.289</td>
<td>0.829</td>
<td>0.193</td>
<td>2.547</td>
<td>5.790</td>
</tr>
<tr>
<td>Tuticorin (P7)</td>
<td>5.750</td>
<td>1.701</td>
<td>0.296</td>
<td>3.048</td>
<td>9.466</td>
</tr>
<tr>
<td>Chennai (P8)</td>
<td>12.515</td>
<td>2.334</td>
<td>0.186</td>
<td>7.846</td>
<td>16.752</td>
</tr>
<tr>
<td>Ennore (P9)</td>
<td>3.799</td>
<td>1.955</td>
<td>0.515</td>
<td>1.418</td>
<td>8.25</td>
</tr>
<tr>
<td>Vizag (P10)</td>
<td>14.368</td>
<td>1.992</td>
<td>1.992</td>
<td>9.910</td>
<td>17.643</td>
</tr>
<tr>
<td>Paradip (P11)</td>
<td>11.935</td>
<td>4.048</td>
<td>0.339</td>
<td>5.419</td>
<td>19.604</td>
</tr>
<tr>
<td>Kolkata (P12)</td>
<td>11.450</td>
<td>1.946</td>
<td>0.170</td>
<td>7.320</td>
<td>15.100</td>
</tr>
</tbody>
</table>

1 Average quarterly freight flow (million tons)
2 Coefficient of variation (standard deviation/average)

Number of observations: 59 (Quarter – I, 2001-2002 to Quarter – III, 2015-2016)

5. Model Development and Estimation Process

The spatial-temporal model development and estimation was carried out in several steps using statistical computing language R3.2.3. The model orders were identified using STACF and
STPACF plots. The STACF plots showed tails cutoff and STPACF appeared to be cut off after 1 time lag at 1 spatial lag. Therefore, the model is identified as STARMA (1,0,1) or STAR (1,1). An algorithm was developed to estimate the model using ‘starma’ packages available with R software. The algorithm is explained in the subsequent subsection.

4.1 Modeling Algorithm

The STARMA model development follows a three-stage iterative model building process consisting of model identification, estimation and diagnostic checking of the selected model. The steps involved in achieving the best fitted model is explained in the following sub-sections.

4.2.1 Steps of the Algorithm

The model was developed using ‘starma’ and ‘spdep’ package of R program. The ‘starma’ package, developed by Felix Cheysson, helps in identifying, estimating and diagnosing space-time dependent STARMA models. ‘spdep’ package creates spatial weights matrix objects. The several steps in model development are as follows.

Step 1: Develop a coordinate matrix:

A matrix is developed with each row as a port location consisting of longitude and latitude as columns. Then the point coordinates are plotted. The coordinates for all the major ports have been extracted with the help of Google Map. The longitude and latitude for each port is presented below.

<table>
<thead>
<tr>
<th>Port</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Port</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Port</th>
<th>Longitude</th>
<th>Latitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kandla</td>
<td>70.22°E</td>
<td>23.03°N</td>
<td>Mangalore</td>
<td>74.88°E</td>
<td>12.87°N</td>
<td>Ennore</td>
<td>80.33°E</td>
<td>13.25°N</td>
</tr>
<tr>
<td>Mumbai</td>
<td>72.77°E</td>
<td>18.94°N</td>
<td>Cochin</td>
<td>76.14°E</td>
<td>9.58°N</td>
<td>Vizag</td>
<td>83.28°E</td>
<td>17.70°N</td>
</tr>
<tr>
<td>JNPT</td>
<td>72.95°E</td>
<td>18.95°N</td>
<td>Tuticorin</td>
<td>78.12°E</td>
<td>8.47°N</td>
<td>Paradip</td>
<td>86.68°E</td>
<td>20.27°N</td>
</tr>
<tr>
<td>Mormugao</td>
<td>73.80°E</td>
<td>15.41°N</td>
<td>Chennai</td>
<td>80.29°E</td>
<td>13.08°N</td>
<td>Kolkata</td>
<td>88.30°E</td>
<td>22.55°N</td>
</tr>
</tbody>
</table>

Step 2: Identify the nearest neighbor:

The nearest neighbor for each of the ports is determined using k-nearest neighbor method, where k=1 in this case. The nearest neighbor is then used to give spatial weight in the weight matrix. A matrix showing each of the ports and their respective neighbors is obtained using “kneigh” function of the ‘spdep’ package. The list of all the ports and their respective nearest neighbor port are shown below.

<table>
<thead>
<tr>
<th>Port</th>
<th>Nearest port</th>
<th>Port</th>
<th>Nearest port</th>
<th>Port</th>
<th>Nearest port</th>
<th>Port</th>
<th>Nearest port</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kandla</td>
<td>Mumbai</td>
<td>Mangalore</td>
<td>Mormugao</td>
<td>Ennore</td>
<td>Chennai</td>
<td>JNPT</td>
<td>Mumbai</td>
</tr>
<tr>
<td>Mumbai</td>
<td>JNPT</td>
<td>Cochin</td>
<td>Tuticorin</td>
<td>Vizag</td>
<td>Paradip</td>
<td>Mormugao</td>
<td>Mangalore</td>
</tr>
<tr>
<td>JNPT</td>
<td>Mumbai</td>
<td>Tuticorin</td>
<td>Cochin</td>
<td>Paradip</td>
<td>Kolkata</td>
<td>Ennore</td>
<td>Kolkata</td>
</tr>
<tr>
<td>Mormugao</td>
<td>Mangalor</td>
<td>Chennai</td>
<td>Ennore</td>
<td>Kolkata</td>
<td>Paradip</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 3: Determine correlation coefficient for nearest neighbor ports:

Correlation between each of the pair of nearest neighbor ports is obtained using ‘cor’ function of R package. The Pearson correlation coefficients were calculated using the port freight data for all the paired ports.
Step 4: Create weight matrices:

Zeroth order weight matrix is the matrix which shows how a given port is influenced by its zeroth order neighbor (i.e., itself). A diagonal matrix with same size as the number of ports is created. Table 3 shows the first order weight matrix which shows how a given port is influenced by its nearest neighbor. Higher order matrices were also attempted to consider in the modeling process, however, the model statistics were not found to be significant. Therefore, only first order matrix is presented here. The matrix is created with each row consisting of all values as zero except the nearest neighbor column which is filled with correlation coefficient value. Since the spatial order \((l)\) is taken as 1 in our case, only zeroth and first order weight matrix are created. Such weight matrices represent the interaction between the ports according to the modeler’s choice. Correlation matrix was taken to be the first order weight matrix since it gives the linear dependence between the nearest neighbor ports. The spatial lag operator is taken in this study as follows:

\[
w_{ij} = \begin{cases} 
\rho_{ij} & \forall i \neq j \\
0 & i = j 
\end{cases}
\]

Where, \(\rho_{ij}\) is the Pearson’s correlation coefficient of freight flows at two nearest neighbor ports \(i\) and \(j\).

Table 3: First order weight matrix (\(W^{(1)}\))

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>P9</th>
<th>P10</th>
<th>P11</th>
<th>P12</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0</td>
<td>0.799</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P2</td>
<td>0</td>
<td>0</td>
<td>0.731</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P3</td>
<td>0</td>
<td>0.731</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.012</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.899</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.899</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.209</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.631</td>
<td>0</td>
</tr>
<tr>
<td>P10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.289</td>
</tr>
<tr>
<td>P11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.289</td>
</tr>
</tbody>
</table>

Step 5: Center and scale the dataset:

The data is centered using “stcenter” method of the STARMA package in R. The data is centered and scaled using this method to avoid any intercept tem in the model. In “stcenter” function, the data is centered globally since it is considered all points come from a single process.
Step 6: Develop the model:
The space time model is created using the ‘starma’ package of R program, where the standardized dataset, weight matrices, and AR and MA orders (p, q) are given as input. The best model is obtained by trying with various combinations of p and q values and finally selecting the one with statistically significant parameters.

Step 7: Model selection:
The best model is chosen based on lower root mean square error (RMSE) and mean absolute error (MAE) using ‘caret’ function in 'starma' package.

Step 8: Elasticity Analysis:
The final model was used to carry out the elasticity analysis to understand the interaction among ports.

4.2.3 Model Estimation Results
As discussed previously, the identified model STAR (1,1) was estimated and found to be statistically significant at 90% confidence level. Additionally, we estimated STRMA (0,1,1) or STMA (1,1) and it was found to be statistically significant. We estimated STMA (1,1) as the STPACF plot was not distinctly showing cuts off. STARMA (1,1,1) model estimation results were not significant. STAR (1,1) and STMA (1,1) are presented through model M1 and M2, respectively. The model estimation results were shown below.

Model M1: STARMA (1,0,1) or STAR (1,1)
\[ z_{it} = 0.937 \times z_{it-1} + 0.049 \times w_{ij} \times z_{jt-1} \]

Where, \( z_i \) and \( z_{it-1} \) are freight flow at i for quarter t and t-1 respectively; \( z_{jt-1} \) is freight flow at j and \( w_{ij} \) is the correlation between the ports i and j.

Model M2: STARMA(0,1,1) or STMA(1,1)
\[ z_{it} = 1.017 \times a_{it-1} + 0.259 \times w_{ij} \times a_{jt-1} \]

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>STARMA (1,0,1) or STAR (1,1)</td>
<td>( \phi_{10} )</td>
<td>0.937</td>
<td>0.015</td>
<td>62.665</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>( \phi_{11} )</td>
<td>0.049</td>
<td>0.029</td>
<td>1.688</td>
<td>0.091</td>
</tr>
<tr>
<td>STARMA (0,1,1) or STMA(1,1)</td>
<td>( \theta_{10} )</td>
<td>1.017</td>
<td>0.042</td>
<td>24.401</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>( \theta_{11} )</td>
<td>0.259</td>
<td>0.088</td>
<td>2.938</td>
<td>0.003</td>
</tr>
</tbody>
</table>

It can be seen from above that AR and MA parameters are statistically significant for both the models. To choose the best model, the models were validated based on RMSE and MAE values.
1 **Model Performance**

The models performances were validated through RMSE and MAE values. They are defined in Eq.(10) and Eq.(11).

\[
MAE = \frac{1}{n} \sum_{t=1}^{n} |Y_{\text{obs},t} - Y_{\text{model},t}|
\]  

\[
RMSE = \sqrt{\frac{\sum_{t=1}^{n} (Y_{\text{obs},t} - Y_{\text{model},t})^2}{n}}
\]

Where, \(n\) number of quarters, \(Y_{\text{obs},k}\) and \(Y_{\text{model},k}\) are the observed and modeled freight flow for the \(t^{th}\) quarter, respectively.

The MAE value for STAR model was found to be much less than the counterpart STMA model. The maximum value of MAE is 0.699 from STAR (1,1) model for the case of Mormugao port whereas the minimum value of MAE is 0.870 from STMA (1,1) model in case of Mangalore port. The overall average of MAE values is 0.204 (see Table 4) from STAR (1,1) model considering all port freight flow predictions. Similar to this observation, RMSE values from STAR (1,1) model are considerably lower than that of STMA (1,1) model for all the ports. The overall average of RMSE values is 0.286 using STAR (1,1) model. Overall, the proposed model STAR (1,1) better represented east coast ports (MAE=0.165) than the west coast ports (MAE=0.204). Considering lower values for MAE and RSME from STAR (1,1) model, we utilized this model further for elasticity analysis.

**Table 4: Error indicators using STARMA (1,0,1) or STAR (1,1) model**

<table>
<thead>
<tr>
<th>Port (West coast)</th>
<th>RMSE</th>
<th>MAE</th>
<th>Port (East coast)</th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kandla (P1)</td>
<td>0.285</td>
<td>0.216</td>
<td>Tuticorin (P7)</td>
<td>0.218</td>
<td>0.101</td>
</tr>
<tr>
<td>Mumbai (P2)</td>
<td>0.222</td>
<td>0.156</td>
<td>Chennai (P8)</td>
<td>0.242</td>
<td>0.179</td>
</tr>
<tr>
<td>JNPT (P3)</td>
<td>0.200</td>
<td>0.136</td>
<td>Ennore (P9)</td>
<td>0.282</td>
<td>0.148</td>
</tr>
<tr>
<td>Mormugao (P4)</td>
<td>0.901</td>
<td>0.699</td>
<td>Vizag (P10)</td>
<td>0.194</td>
<td>0.164</td>
</tr>
<tr>
<td>Mangalore (P5)</td>
<td>0.270</td>
<td>0.173</td>
<td>Paradip (P11)</td>
<td>0.260</td>
<td>0.197</td>
</tr>
<tr>
<td>Cochin (P6)</td>
<td>0.197</td>
<td>0.082</td>
<td>Kolkata (P12)</td>
<td>0.157</td>
<td>0.198</td>
</tr>
<tr>
<td>Average</td>
<td>0.346</td>
<td>0.244</td>
<td>Average</td>
<td>0.225</td>
<td>0.165</td>
</tr>
<tr>
<td>Overall Average</td>
<td>0.286</td>
<td>0.204</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. **Elasticity Analysis Results and Discussion**

Direct and cross elasticities were determined among ports using Eq.(8) and Eq.(9) for STAR(1,1) model. Elasticities represent how the demand for a particular port for the next interval changes with the increase in demand for its neighbor port or itself in this interval (time lag taken as 1). The elasticities are reported in Table 5. The diagonal elements in the table gives us the direct elasticities for each of the ports. Direct elasticity here refers to how the flow changes in the
following quarter when the flow for this quarter increases for that port itself. Close study of all
the diagonal elements reveal that all the ports except one has elasticity less than 1 and hence are
inelastic, indicating demand stability and lesser volatility due to fluctuations in economic
environment. Direct elasticity for Mormugao port show elastic behavior since demand elasticity
is greater than 1 (1.169). This means that a small change in the last quarter’s freight flow will
have a substantial effect in the present quarter’s flow indicating demand volatility due to external
variables.

The cross elasticity is used to determine the dependency with neighbor ports. It can be observed
from Table 5 that a 10% increase in freight flow during the previous quarter for Mumbai port
(P2) would result in an increase of 0.22% in the freight flow of Kandla port (P1) for the present
quarter; whereas, a 10% increase in freight flow during the previous quarter for Kandla port
would increase the port's freight flow by 9.3% during the present quarter. Similarly, an increase
in 10% during the previous quarter of JNPT port (P3) freight flow would increase Mumbai port's
freight flow by 0.47% during the present quarter. This shows JNPT port has more influence on
Mumbai port than the influence of Mumbai port on Kandla port. The cross elasticities between
Mormugao and Mangalore is negative (-0.001). Although the value is low, the negative demand
elasticity between these ports suggested that these two ports are competing with each other as
they share common hinterland in two coastal states: Goa and Karnataka. Both the states are
major producer of iron ore. Remaining port cross elasticities are between 0.4 to 6.3%; no
competition was found among these ports.

This lack of competition can be attributed to lack of viable port options considering the less than
desirable performance levels of Indian ports (27) and the direct inelasticity found in this study.
Proposed plans for development of other ports and improvement programs for existing ports by
way of providing additional options may change both the direct and cross elasticity dynamics
which makes it imperative to understand the implications of proposed development plans before
committing investments.

Table 5: Demand elasticity analysis results

<table>
<thead>
<tr>
<th>Port</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>P9</th>
<th>P10</th>
<th>P11</th>
<th>P12</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0.933</td>
<td>0.022</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P2</td>
<td>-</td>
<td>0.924</td>
<td>0.047</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P3</td>
<td>-</td>
<td>0.029</td>
<td>0.938</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.169</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.001</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.936</td>
<td>0.063</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.031</td>
<td>0.922</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.951</td>
<td>0.004</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.031</td>
<td>0.911</td>
<td>-</td>
</tr>
<tr>
<td>P10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.941</td>
<td>0.031</td>
</tr>
<tr>
<td>P11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.010</td>
</tr>
<tr>
<td>P12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
6. Conclusion

The correlation between the freight flows and interaction between the ports was expected to be best represented by a space-time model. The quarterly time-series data obtained from the 12 major ports located along the east and west coast of India were used for analysis. The data were analyzed through a newly proposed space-time model STAR \((1,1)\) with due consideration towards understanding the influence of neighboring ports on a subject port.

Elasticity analysis of such a model revealed some interesting insights regarding the space-time dependence of a port freight flow on nearest neighbor port. Direct elasticity analysis suggested how the freight flow of a particular port in the present quarter is affected by the change in freight flow of that port in the previous quarter. The direct elasticity values for all the major ports ranged between 0.911 (Ennore) and 1.169 (Mormugao). This shows that the freight flows during the current quarter are highly dependent on the volume of its earlier quarter period. Higher direct elasticity value for Mormugao (1.169) indicated that the freight handled by the port in the present quarter is significantly affected by the change in the freight handled by the port in the previous quarter. Cross elasticity values on the other hand ranged between -0.001 and 0.063. The variation between the two extremes is almost different by a factor of 6, which is significant. It is quite high when compared to the direct elasticities because cross elasticity, unlike direct elasticity, is influenced by both space and time. Negative values of cross elasticities between Mormugao and Mangalore indicate that these two ports compete with each other for attracting freight from their common hinterland existence in two coastal states: Goa and Karnataka.

The freight flow forecasting of ports has been done till now using its past freight flow data without considering the effects of its neighbor ports. When forecasting the freight flows, it is therefore important to know if the past and current freight flow in neighbor ports can help make the forecasting even better. The proposed models can be used for short-term forecasting of freight flows through the major ports and assessing the impact of freight flow changes from one port to the nearest neighboring ports. It is expected that this study will help port authorities and policy makers for holistic development of port system by making right investments in required locations to promote balanced development. It has also implications towards formulating policies on port development considering Government of India's preferred mode of choice for infrastructure development is PPP, and policy formulation for this mode of development is required to address competition concerns considering the high sunk cost associated with ports development.

Acknowledgement

Thank to Ministry of Shipping, Mumbai Port Trust and Paradip Port Trust for the data used in this study.

References