Air Cargo Demand Modeling and Prediction

Raghavendra Totamane, Amit Dasgupta, and Shrisra Rao, Member, IEEE

Abstract—The air cargo transportation system is a large and complex service system, in which demand forecasting is a key element in the master planning process. Demand forecasting is essential for analyzing existing cargo flight schedules and identifying future facility requirements of air cargo companies. We use the Potluck Problem approach to propose a multiproducer/multiconsumer solution for predicting the cargo demand of a specific airline in a given route, and the cargo load factor for a given flight schedule on that route. This solution considers each airline as a producer and the users of air cargo services as consumers, with a producer having no explicit communication with other producers/airlines. The model analyzes the existing cargo capacity plan, highlights drawbacks, and proposes a new capacity plan to demonstrate the effectiveness of using the solution. Examples are provided to illustrate the efficacy of the approach.

Index Terms—Air cargo, cargo capacity, cargo load factor, demand prediction, multiagent system, Potluck Problem, weighted majority algorithm.

I. INTRODUCTION

A IR CARGO transportation is primarily done through larger aircraft, such as passenger airplanes, cargo airplanes, and charter airplanes. In the case of passenger airplanes, cargo supply and delivery are closely tied to passenger flight schedules, which are normally designed to facilitate passenger transport; therefore, the demand cannot be met by passenger airplanes alone, and cargo airplanes play an important role in cargo transportation. Forecasts suggest that the air cargo industry will continue to grow [1] and competition in the air cargo industry will be strong. Many passenger airlines find themselves being squeezed and see their yields and profits dropping constantly. In contrast, integrated carriers such as UPS and FedEx have been able to sustain and make good revenues and profits even during economic slowdown [2], [3]. This suggests that cargo customers are willing to pay extra price for very reliable, value added, and tailor-made services. Cargo management is gaining importance as it generates substantial revenues for airlines, more so than the saturated and stagnating passenger travel markets. A forecast report published by Boeing [4] says that air cargo traffic worldwide will continue to increase in spite of increases in jet fuel prices. A forecast report published by Boeing [4] says that air cargo traffic worldwide will continue to increase in spite of increases in jet fuel prices.

The amount of cargo supply that can be handled varies because of the varying demand for cargo every day. Irrespective of the variation in demand, cargo airlines have fixed flight schedules (e.g., five flights daily from Los Angeles to New York) for most days in a year. Even though cargo airlines use various algorithms and strategies for flight scheduling and resource planning, most of them operate at variable cargo load factors (LFs) of 50%–70%. The cargo load factor is the ratio of revenue ton-miles to the maximum possible ton-miles (assuming every flight carries its maximum cargo all the time) or the fraction of the utilized capacity over the maximum capacity [14], [15]. Potential analysis is also useful, particularly for forecasting markets in their early stages of development. Each technique has its special advantages and limitations; therefore, special care must be taken to select one or multiple techniques based on the application [5]. In particular, it is quite important to efficiently forecast and manage cargo bookings [10], [11]. Our air cargo demand forecasting is model driven and based on inductive reasoning [12], [13], led by observing the total supply of air cargo on many instances/days over the years and the average market share of airlines operating on specific regions/routes. Multiple airlines (agents) individually strategize the quantity of cargo to be handled, without any prior coordination among themselves, in a self-interested manner.

We use the Potluck Problem approach to propose a multiproducer/multiconsumer solution for predicting the cargo demand of a specific airline in a given route, and the cargo load factor for a given flight schedule on that route. This solution considers each airline as a producer and any user of air cargo services as the consumer, with a producer having no explicit communication with other producers/airlines. The model analyzes the existing cargo capacity plan, highlights drawbacks, and proposes a new capacity plan to demonstrate the effectiveness of using the solution. Examples are provided to illustrate the efficacy of the approach.

Manuscript received February 11, 2012; revised August 31, 2012; accepted September 4, 2012.

R. Totamane is an independent IT Consultant, Bangalore, India (e-mail: raghtotamane@gmail.com).
A. Dasgupta is with Unisys Global Services India, Bangalore 560 025, India (e-mail: amit.dasgupta@in.unisys.com).
S. Rao is with the International Institute of Information Technology Bangalore, Bangalore 560 100, India (e-mail: shrao@iiitb.org).

Digital Object Identifier 10.1109/JSYST.2012.2218511

1 Cargo LF is the ratio of revenue ton-miles to the maximum possible ton-miles (assuming every flight carries its maximum cargo all the time) or the fraction of the utilized capacity over the maximum capacity [14], [15].
thus can use better prediction models and improve their cargo capacity planning and flight scheduling.

Our proposed model, based on the Potluck Problem [16], approach, aims at suggesting a flexible flight schedule and cargo capacity plan based on demand prediction, which can maximize the cargo LF while maintaining existing delivery success rate. A major airline often has 100 million pounds of weekly cargo lift capacity, amounting to several hundred millions of dollars in revenue. With this volume of cargo, even a slight improvement in the forecasting technique and cargo LF is apt to have a major impact in overall savings, performance, and efficiency. Our model uses the weighted majority learning algorithm [17] with various predictors for predicting the future demand. Based on the predicted demand, available cargo capacity, and global/regional economic conditions (GDP, inflation, industry growth, etc.), and by applying various strategies, a new cargo capacity plan is suggested, thereby improving the cargo LF as well as the financial bottom line.

For our simulation and analyses, we have primarily collected the data from Unisys Logistics Management System (LMS), which is used by many leading cargo airlines to handle about 35% of the world’s air freight. In our learning algorithm, we used past years’ data (specifically from the years 2005 and 2006) of three different cargo airlines operating on different routes in the North American region as training data. Then the results were compared with the airlines’ actual performances in the year 2007 (i.e., we saw how our suggested schedules would have done compared to actual results, given the airlines’ 2007 cargo data). We did this with multiple routes and airlines, and discovered that our algorithm achieved an improvement in the cargo LF of over 9% (which translates into tens of millions of dollars) in all cases.

The remainder of this paper is organized as follows. Section II describes some of the prior work and discusses the problem in more detail. Section III describes the air cargo demand prediction model, the importance of cargo capacity planning, and how the weighted majority learning algorithm is applied to predict the cargo demand. Section IV provides the pseudocode for learning and application model. Section V discusses the results of simulations using our model. Section VI provides conclusions and suggestions for further work.

II. PROBLEM FORMULATION

A. Motivation for Air Cargo Demand Prediction

Cargo, in general, is of course an aspect of human society and trade, and its problems are as old as recorded history. There has been considerable work on cargo problems, e.g., in the container shipping industry [18]. The air cargo space problem has received attention for decades, starting with seminal works of Fetter and Steorts [19], Whitehead [20], and Marsten and Muller [21]. More recent work includes simulation and other approaches for fleet routing and planning [1], [22], [23], revenue management [24], [25], and general studies of feasibility and other issues [26], [27].

Most major airlines are interested in new strategies for improved achievement of business goals in the air cargo industry. Airlines use multiple strategies to improve their overall efficiencies of operation. Mongeau and Bès [28] address the importance of optimization of aircraft container loading and maximizing the mass of goods loaded onto the aircraft. They also address the problem of loading as much freight as possible on an aircraft while balancing the load in order to minimize fuel consumption and to satisfy stability/safety requirements.

An increase in the cargo LF directly impacts business performance. Many factors affect the cargo LF, such as efficient use of space and load balancing, center of gravity, and so on. Fok et al. [29] investigated how cargo and mail LFs can be improved through the use of mathematical optimization, and developed a Web-based application that performs long-term forecasting based on analysis of historical data, and then helps with operational load planning with mathematical optimization.

Along with the continual expansion of the air cargo market, more and more cargo airlines have been established. Facing challenges from domestic and foreign competitors, airlines need to understand customer demands and analyze their values properly in order to compete successfully in the field [30]. For air cargo and air logistics companies, data mining is greatly helpful in analyzing known data and forecasting the future developing trends of their business, as well as in figuring out the key factors to complete current tasks. These help the companies lower costs and increase profits, thus helping to place them in a more competitive position in the business world. The main purpose is to combine air cargo planning with data mining to promote the data mining technology as appropriate for use with critical functions in the air cargo business [6], [31].

In China, economic growth has been the primary drive for air cargo demand, and the aggregate forecast of future air cargo in China is projected based on the relationship between air cargo demand and economic development. Jiang et al. [32] obtained GDP projections by trend analysis and projections from the Chinese Government and recognized institutions, and developed an econometric method to determine the GDP/air-cargo relationship. Kasarda and Green [33] noted that air cargo plays an important role in overall economic development, a view that is also seen elsewhere [34].

Most airlines overbook their actual capacity (for both passengers and cargo) because part of the booked demand often does not show up at the loading gate by the time of flight departure [35]. A key element of overbooking is a model that accurately predicts the show-up rate for current bookings. Given the increasing importance of cargo, most major airlines now scrutinize estimates of show-up rates for cargo bookings. Popescu et al. [36] showed that improved estimation of the show-up rate can improve profits and customer service. Kaslingam [37] considered the problem of overbooking and presented a model given uncertain available capacity. Patel et al. [38] presented various deterministic models for determining optimal pickup times for air cargo from an airport and delivering it to a local distribution center for a global manufacturer. The solutions of these models serve as initial starting points to solve the stochastic problem with random arrival times of the flights and random custom clearance times.
and travel times. Huang and Hsu [39] presented a study on revenue management for air cargo with supply uncertainty, and their numerical experiments were based on actual operational data of a Taiwanese airline. Their results showed that accurate forecast and control of cargo space supply are critical for increasing revenue.

The air cargo supply chain is composed of shippers, freight forwarders, and airlines. The shippers send their shipments to freight forwarders, who are then responsible for contacting the airlines and procuring space to ship the cargo according to the shippers’ needs. Poppescu [40] proposed a structured method for improving the airlines’ and freight forwarders’ actions when confronted with accepting demand and acquiring capacity, respectively. Several coordinated models combining airport selection, fleet routing, and timetable setting are developed on real operating data of Taiwan Airlines [41], and preliminary results show how these models can be beneficial to airline alliances.

NASA’s scenario-based strategic planning process [42] identified global air cargo as one of several potential high-payoff vehicle classes for the year 2020. Within this vehicle class, range/payload and cost goals were established to provide a 10-fold reduction in the cost per ton-mile for air cargo shipments. The study explains technologies and configurations required to determine the opportunities and the potential to meet them. The Boeing Company issues the biennial World Air Cargo Forecast [43] to provide a comprehensive up-to-date overview of the air cargo industry. The forecast summarizes the world’s major air trade markets, identifies major trends, and presents forecasts for the future performance and development of markets and the world’s freighter airplane fleet. After strong growth of 3.9% and 3.7% in 2006 and 2007 (the most recent years for which global data is publicly available), respectively, world economic activity, as measured by GDP, is predicted to grow at nearly double the GDP growth rate. Although economic activity is the primary influence on world air cargo development, other factors must also be considered.

III. SYSTEM MODEL

A. Problem Description

Generally, in a given calendar year, most major airlines have fixed numbers of different aircraft (Boeing 747, 767, DC8, DC10, and so on), with cargo capacities ranging from 6 to 120 tons. Every airline has a fixed cargo capacity over specific regions/routes (e.g., 1000 tons/week within the North American region, 500 tons/week between North America and the European Union, and so on). Every region has a certain number of segments (e.g., North America has segments like Chicago–Boston, Los Angeles–New York, and so on), and a fixed number of airlines operate in each segment. Each airline knows about the total cargo supply per month in any segment from the freight statistics report published by the country/region [44]. Every airline does some demand forecasting and plans its flight schedule for routes/region well in advance for weeks/months and then publishes the flight schedule to its freight forwarders, partners, customers, and others. An airline is able to maximize the utilization of its cargo capacity and provide better customer satisfaction if it can predict, to a reasonable accuracy, the exact demand on a given day, and plan and schedule its flights accordingly. Cargo supply in any specific segment varies based on many factors, such as the day of the week, the month of the year, the festival season, holidays, and economic factors. An airline decides how much supply it can provide based on the demand prediction and the available cargo capacity.

Specifically, as described in Section I, we have a supply-side problem, where the available demand for a resource varies, and supply is to be adjusted accordingly. In a game-theoretical fashion, the air cargo capacity optimization problem is formulated by us as a Potluck Problem [16].

B. Potluck Problem and Demand–Supply Parity

The Potluck Problem [16] is an abstract game-theoretical paradigm of a repeated noncooperative game in a multiagent system. There are numerous producers and consumers who produce a resource, and consumption happens at specific “dinner instances” where it is desired to avoid either “starvation,” a situation where some demand remains unfulfilled, or “excess,” a situation where there is more supply of the resource by the producer agents than demanded by the consumers. Agents (both producers and consumers) do not cooperate or communicate their intentions ahead of time, but all have access to historical data regarding production and consumption at past dinner instances, and every agent acts in its self-interest.

The Potluck Problem may thus be thought of as a generalized version of the scenario where each individual provider of resources may have access to the historical demand data and must determine what resources to provide in a future time period based on that historical demand data. Thus, it is an archetype of a problem where both demand and supply vary. It is a generalization of the well-known Santa Fe Bar Problem [45] or El Farol Problem, where the supply is kept fixed but the demand may vary. However, as with the Santa Fe Bar Problem, a game of the Potluck Problem does not generally converge on its own and has oscillatory behavior in the classical sense of Cournot [46]. This is because if, in a specific dinner instance, there is starvation and the producer agents produce more for the next dinner instance, then there can be excess at the latter instance. On the other hand, if there is excess at a dinner instance and the producers reduce their production, there can be starvation at a subsequent instance. Such oscillations can continue indefinitely.

In order to take on the Potluck Problem, various “predators” and nonrational learning are used by the producers in order to predict the demand at a future dinner instance. As a multitude of predictions generally results when a predator uses a set of predictors, and as a single predictor is unlikely to be always correct, the weighted majority algorithm (WMA) [17] is used to determine the overall demand prediction. The WMA is a meta-learning algorithm for predicting a sequence of values using predictions from a finite number of predictors. It can also be considered a way to determine an overall prediction, giving more weight to the predictions of
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

reliable predictors than to those less reliable. The algorithm assumes that the predictors in the pool have no knowledge about the accuracy of their predictions, but one or more predictors may perform better than others. Our net prediction is thus based on the simple and effective method of weighted voting. The WMA is quite robust against bad predictors and is also known to perform at least as well when additional predictors are added. This is an important quality, considering that individual algorithms can have biases and other such drawbacks [47]. Other possible replacements for the WMA are not suitable in this context because they make assumptions about the data that are not likely to be always true in this domain (e.g., the Kalman filter assumes that the probability density functions of the data sources are Gaussian).

The Potluck Problem theoretical model has also been used in other domains where parity of supply and demand is required, such as in the design of solar-powered electrical microgrids [48], and in reducing the power consumption of large IT data centers [49]–[51]. In such cases, the resource produced and consumed is variously taken as electrical power, computational capacity, etc. The “dinner instances” are considered to occur hourly, daily, etc., as appropriate for each given domain.

C. Application to Air Cargo Modeling

In our formulation, an airline is a producer of the resource (cargo transfer capacity) while consumers demand the same resource at varying rates. As in the model, producers and consumers do not collude or make their intentions known in advance, but act in individual self-interest.

We aim at obtaining a near-equilibrium state where the aggregate supply of goods is nearly equal to the aggregate demand. In order to reach an equilibrium or near-equilibrium state between supply and demand at an instance $i$, each supplier agent, using the solution strategy of the Potluck Problem, predicts the demand for resources at the 10th dinner instance given past data, and based on this prediction, decides on the quantities to supply at that dinner instance. The higher the accuracy of these demand predictions by the supplier agents, the closer we get to the equilibrium between aggregate demand and supply of air cargo transport resources.

Say there are $n$ agents (airlines) who are players in the game. Consider one instance of the game (cargo supply on one day), $i$. For a player $j$, the strategy set is $0 \leq Q_j \leq \text{Max}_j$, where $Q_j$ is the amount of cargo transported by agent $j$ and $\text{Max}_j$ corresponds to the maximum amount cargo that agent $i$ can transport. Let $M_j$ denote the set of probability distributions over $Q_j$ that defines the mixed strategy (combining multiple strategies as one strategy) for agent $j$. Agent $i$ has multiple types of aircraft each having specific cargo capacities $C_q$. There can be multiple cargo flights with the same cargo capacity $C_q$ operating on a specific route. It is also possible that the same aircraft with cargo capacity $C_q$ may make multiple trips on a specific route for a given day $S_i$. The sum of cargo capacities of cargo flights operated by agent $i$ on a specific route, is given by $S_i = \sum_{q=1}^{n} C_q$. Now $S_i \in M_i$ indicates the mixed strategy of the player $i$ on a given day $i$. Then the total cargo transported by all the airlines on day $i$ is given by $S_i = \sum_{j=1}^{n} S_j$. The agent $i$ decides on the mixed strategy $S_i$ by predicting its demand for day $i$, which is denoted by $D_i$. The agents’ demand for the cargo also varies over different days. Ideally, an agent can maximize the cargo transfer capacity while consumers demand the same.

### D. Predictors and Predictions

As noted above, the cargo capacity planning approach we propose is similar to the Potluck Problem [16], [52], a generalization of the Santa Fe Bar Problem [45]. We have extended the algorithm used in the Potluck Problem [16], and this nonrational algorithm does well in maintaining parity of supply and demand in varying-demand situations, yielding better results.

For a given region or a route, cargo capacity planning is a decision-making process on scheduling of flights based on the recommendations of various putative experts (predictors), of which some are probably better than others. It makes sense that we would lean toward the recommendations of our better experts and ignore those who have performed poorly in the past. A deterministic way of implementing this is to assign a weight to each of the predictors, which represents the level of trust we currently place in them. When they predict accurately, we maintain or increase the weight given to them, and when they are incorrect, we reduce their weight in some way (e.g., by dividing by 2). Then, we compute a weighted average of the predictions to decide on the expected cargo demand.

A predictor makes use of previous cargo supply data available and makes a prediction for the demand at the current time. It is a function that uses $S_{-1}, S_{-2}, \ldots, S_{-t}$ (past data) and $D_{i-1}, D_{i-2}, \ldots, D_{i-t}$ and predicts the demand for the forthcoming day.

The predictors used here are:

1. basic statistical functions (e.g., average, minimum, maximum cargo supply over specific days of the week such as Sundays, Saturdays, other weekdays, and others);
2. cargo market share of the airline in given route compared to other airlines operating in the same route;
3. economic and quarterly GDP growth in the region (expected growth in demand);
4. monthly industry growth rate in the region;
5. average monthly inflation rate in the region;
6. monthly exports/imports growth rate in the region;
7. average cargo supply during special occasions like Valentine’s Day, Mother’s Day, and others;
8. average cargo supply during the first two quarters of the year;
9. average cargo supply during last two quarters of the year;
10. average demand over the last $j$ days (where $j$ is a tunable parameter, e.g., 15).
11) the rational predictor (presume that cargo supply on the day will be same as on the previous day); 12) maximum and minimum cargo supplied over last week; 13) cargo supply during the first week of every quarter.

Of the various predictors available, each airline chooses some $k$ predictors. To begin with, these $k$ can be all the predictors that are available, but it is reasonable to surmise that a producer would hand-pick certain specific predictors based on computation efficiency, availability, or other such factors. Then, each agent has $k$ predictions about the demand on the coming day, each of which is denoted by $O_{i,p,t}$, representing the predictor made by the $p$th predictor of airline $i$ for day $t$. The airline $i$ decides the supply $S_{i,t}$ to be made available on day $t$, based on the forecasts of the $k$ predictors, using a weighted majority approach. Each agent $i$ maintains a weight $W_{i,p,t}$ for each predictor $p$, $1 \leq p \leq k$, at time $t$, and updates it after each iteration or day, with the weight of the accurate predictors increasing and that of the inaccurate ones decreasing. The initial weights of all predictors may be equal or may be set to some predetermined or random nonzero values.

The WMA is robust against prediction errors by individual predictors, and whatever be the initial nonzero weights assigned to the predictors, their stable values always converge in the same way depending on their performances. Thus, more and new predictors can always be added to the mix, and doing this does not hurt the performance even if the new predictors turn out to be not so good.

The iterative update and learning algorithm used by the players (or agents) can be summarized in the following steps. Each airline $i$, on a day $t$, does the following:

1) predicts the demands using all the individual predictors, i.e., finds $O_{i,p,t}$, $\forall p = 1, 2, \ldots, k$;
2) predicts the demand $P_{i,t}$ using all the predictors, using a WMA;
3) decides on the supply $S_{i,t}$ that can be handled;
4) updates the weights of all predictors based on the actual demand and supply seen on the day.

The prediction of total cargo demand for a particular day of the year is calculated by taking the weighted majority [17] of the predictions made by the all predictors of the agent, that is,

$$P_{i,t} = \frac{\sum_{p=1}^{k} (W_{i,p,t} \times O_{i,p,t})}{\sum_{p=1}^{k} (W_{i,p,t})}$$

After each day, an agent updates the weights of all its predictors based on how close they were in the previous instance in predicting the actual demand. The update equation of the weight of a predictor $p$ at an instance $t$ is $W_{i,p,t+1} = W_{i,p,t} \times F$, where $F = \beta^t$, with $\beta$ being a fixed parameter chosen such that $0 < \beta < 1$. This $\beta$ measures how drastic predictions change over iterations. The smaller the $\beta$, the more drastic the changes. If the WMA is applied to a pool of functions with $\beta = 0$ with their initial weights being equal, then it is identical to the Halving Algorithm [53]. If $\beta > 0$, then the WMA gradually decreases the influence of those predictors that make a large number of mistakes and gives the predictors that make few mistakes and high relative weights. Suppose that the WMA is applied to a pool $F$ of predictors and that sequence of trials has $m$ anomalies with respect to $F$. In this case, the value of $\beta$ is slowly increased over the trials and kept constant when WMA makes no more than a constant ($\theta$) times $\log(|F|) + m$ mistakes.

If $O_{i,p,t}/D_{i,t} > 1$, then $\theta$ is set to $O_{i,p,t}/D_{i,t}$, and if $O_{i,p,t}/D_{i,t} \leq 1$, then $\theta \leftarrow D_{i,t}/O_{i,p,t}$. After updating all the weights, they are normalized to be between 0 and 1 using

$$W_{i,p,t+1} = \frac{W_{i,p,t}}{\sum_{p=1}^{k} W_{i,p,t+1}}$$

Once we predict the cargo demand $P_{i,t}$, the agent has to decide on the supply $S_{i,t}$. As explained earlier, $S_{i,t}$ is restricted to discrete values that are combinations of different cargo capacities of aircraft ($C_q$) and the numbers of trips these aircraft operated on a given route and day. Even though various strategies can be used to decide $S_{i,t}$, to begin with, the airline’s previous flight schedule can be used.

We can calculate the predicted cargo LF $LF_{i,t}$ using

$$LF_{i,t} = \frac{P_{i,t}}{S_{i,t}}$$

The notations used in the algorithm are summarized in Table I.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi$</td>
<td>Number of airlines (agents).</td>
</tr>
<tr>
<td>$t$</td>
<td>Specific day.</td>
</tr>
<tr>
<td>$k$</td>
<td>Maximum number of predictors.</td>
</tr>
<tr>
<td>$p$</td>
<td>Specific predictor $1 \leq p \leq k$.</td>
</tr>
<tr>
<td>$s$</td>
<td>Number of days.</td>
</tr>
<tr>
<td>$F$</td>
<td>Pool of experts or predictors or functions.</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Fixed parameter, $0 &lt; \beta &lt; 1$.</td>
</tr>
<tr>
<td>$m$</td>
<td>Number of errors or mistakes.</td>
</tr>
<tr>
<td>$C_q$</td>
<td>Cargo capacity of a specific aircraft type $q$.</td>
</tr>
<tr>
<td>$S_{i,t}$</td>
<td>Sum of cargo capacities ($C_q$) of all cargo flights operated by an agent $i$ on a given day $t$ for a specific route.</td>
</tr>
<tr>
<td>$D_{i,t}$</td>
<td>Cargo demand for an agent $i$ on day $t$.</td>
</tr>
<tr>
<td>$L_{i,t}$</td>
<td>Total cargo demand on a specific route for a day $t$.</td>
</tr>
<tr>
<td>$O_{i,p,t}$</td>
<td>Predicted cargo demand of an agent $i$ on a day $t$.</td>
</tr>
<tr>
<td>$P_{i,t}$</td>
<td>Predicted cargo demand of a predictor $p$ for an agent $i$ on a day $t$.</td>
</tr>
<tr>
<td>$W_{i,p,t}$</td>
<td>Weight of a predictor $p$ for an agent $i$ on a day $t$.</td>
</tr>
<tr>
<td>$LF_{i,t}$</td>
<td>Predicted LF for an airline $i$ for a day $t$.</td>
</tr>
</tbody>
</table>

IV. PROPOSED MECHANISMS

A. Learning Mechanism

The WMA requires past data for a learning model. We have designed the learning model and demonstrate it using two years or 730 days worth of past data. However, like any
Algorithm 1 Learning

Input: Previous cargo supply data of an airline \( i \) for a specific route \((S_{i,1}, S_{i,2}, \ldots, S_{i,t})\) Previous cargo demand data of an airline \( i \) for a specific route \((D_{i,1}, D_{i,2}, \ldots, D_{i,t})\) Previous total cargo demand data for a specific route \((D_{1}, D_{2}, \ldots, D_{n})\)

Output: Weights of the predictors \((W_{1,i}, W_{2,i}, \ldots, W_{t,i})\) Predicted cargo demand data of an airline \( i \) for a specific route \((P_{1,i}, P_{2,i}, \ldots, P_{t,i})\)

// Calculate and assign values to \( k \) predictors for \( e.g. \)
\( O_{i,p,r} \leftarrow 120.0; \)

// Initialize weights of the predictors to a random value
\( W[p] \leftarrow 0.5; \)

// Initialize \( \beta \) to a value
\( \beta \leftarrow 0.5; \)

for \( i = 1 \) to \( x \) do
    // Select \( k \) predictors for day \( t \)
    \( x1 \leftarrow 0.0; \)
    \( x2 \leftarrow 0.0; \)
    for \( p = 1 \) to \( k \) do
        \( x1 \leftarrow x1 + (W_{p,i} \times O_{i,p,r}); \)
        \( x2 \leftarrow x2 + W_{p,i} \); \n        if \( O_{i,p,r} > 1 \) then \( \theta \leftarrow O_{i,p,r}; \)
        else \( \theta \leftarrow \frac{O_{i,p,r}}{2}; \)
        \( W_{p,i,t} \leftarrow W_{p,i} \times \theta^{2}; \)
        // normalize weights between 0 and 1
        \( W_{p,i,t} = \frac{W_{p,i,t}}{\max_{k} W_{p,i,t}}; \)
    end
end

Algorithm 2 Air cargo demand prediction and application

Input: Weights of the predictors \((W_{1,i}, W_{2,i}, \ldots, W_{t,i})\)
Cargo demand prediction values predictors
\((O_{1,i}, O_{2,i}, \ldots, O_{t,i})\)
Future cargo supply data of an airline \( i \) for a specific route
\((S_{i,1}, S_{i,2}, \ldots, S_{i,t})\)

Output: Predicted cargo demand data of an airline \( i \) for a specific route
\((P_{1,i}, P_{2,i}, \ldots, P_{t,i})\)
Future LF of an airline \( i \) for a specific route
\((L_{F,i}, L_{F,i+1}, \ldots, L_{F,i+n})\)

for \( t = 1 \) to \( x \) do
    // Select \( k \) predictors for day \( t \)
    \( x1 \leftarrow 0.0; \)
    \( x2 \leftarrow 0.0; \)
    for \( p = 1 \) to \( k \) do
        \( x1 \leftarrow x1 + (W_{p,i} \times O_{i,p,r}); \)
        \( x2 \leftarrow x2 + W_{p,i} \); \n        \( P_{i,t} = \frac{O_{i,t}}{2}; \)
        \( L_{F,i} \leftarrow \frac{O_{i,t}}{5}; \)
    end
end

other learning models, the more the past data used, the better the accuracy and precision of prediction. Let us assume there are multiple cargo airlines operating on a specific cargo route and that we have the data for the total amount of cargo transported or actual cargo demand data per day \((D_{i})\) for \( x \) days. Now consider one of the operating airlines \( i \) for which we also have past data for cargo supply \((S_{i,1})\) and cargo demand \((D_{i,1})\) for \( x \) number of past days.

Next, we select \( k \) predictors and assign a weight to each of them. The prediction by a predictor could be based on historical data, statistical functions, economic data, or random values. For example, a predictor “average cargo supply on Saturdays” gives the average value of cargo supply of an airline \( i \) for all Saturdays over a period of \( x \) days. The remaining steps of the learning model are self-explanatory and are presented in Algorithm 1.

Even though the value of \( \beta \) is constant, choosing a suitable value plays an important role in demand prediction. The learning model in Algorithm 1 is iterated for different values of \( \beta \) starting from \( \beta = 0.1 \) to \( \beta = 0.99 \). The number of prediction errors \( m \) with a specific tolerance level is calculated for each iteration. The value of \( \beta \) for an iteration, which results in the least number of errors is taken as a constant value for the model. In our paper, we used a tolerance level of \( \pm 25\% \), leading to a \( \beta \) value of 0.87.

B. Cargo Demand Prediction and Application Mechanism

Once we have values for the weights for the predictors from the learning model, we can predict the air cargo demand of an airline \( i \), which is operating in that specific route, for any specific future day \( t \) by selecting \( k \) appropriate predictors from all possible predictors. The testing and application model are presented in Algorithm 2. Even though the learning algorithm is generic, the algorithm has to be applied first on the historical cargo data of any airline \( i \) operating in a given route, to predict the future cargo demand.

By observing the trend of the cargo LF \( L_{F,i} \) and air cargo flight schedule over a period of time, various cargo capacity planning strategies, e.g., addition or removal of flights, can be applied to optimize \( L_{F,i} \), which, in turn, affects the LF. Some such examples are illustrated further in Section V.

Strategies that increase the LF are generally better. However, maximizing LF to the highest possible value would not be able to address sudden unanticipated demand as well as some of the inherent issues in the system, such as bad weather, schedule disruption, technical snag, and crew availability. Successful air cargo carriers keep cargo LFs at not much more than 70%, indicating that having some margin on the LF enables the cargo airlines to provide reliable and on-time services in the face of uncertainties.

V. RESULTS

We note that econometric forecasts depend on a number of different factors, assumptions, and judgments, which subject
them to some degree of uncertainty. As described earlier, our algorithm uses multiple predictors, and, in fact, any other forecasting algorithm can also be used as one of the predictors. Hence, the proposed method is no less accurate on a long-term basis than any other forecasting method.

Even though airlines schedule the flights in a network of routes to meet the cargo demand of a region, as well as other key success factors such as on-time delivery, it is important to analyze cargo demand planning at a micro level or at a specific route level to ensure improved LF and, in turn, better profitability. Hence, we chose to predict the air cargo demand of some leading North American airlines for a few North American routes that have varying cargo demands. The results aim to answer the following questions about our proposed approach for air cargo demand prediction.

Q1) How correctly and closely can we predict the cargo demand?
Q2) How do economic factors and other factors influence the cargo demand?
Q3) How can our approach help improve the cargo LF?

A. Q1

To address the first question, we chose four different routes and did four simulations. The total cargo demand in each route was met by two to four airlines. In every simulation, we chose a specific route and a specific airline operating on that route. The actual total cargo supply data of a route and actual cargo supply data of the chosen airline for the years 2005 and 2006 were used as training data to obtain weights of predictors in the WMA. These predictors were then applied to predict the cargo demand of the chosen airline for the year 2007. In all four simulations, the algorithm’s weighted predictions were accurate nearly 75%–81% of the time (275–295 days in a year) with a tolerance level of ±25% of the airline’s actual cargo demand.

In the first simulation, the algorithm’s weighted predictions were accurate nearly 81% of the time (295 days in a year), with a tolerance level of ±25% cargo demand. Fig. 1 depicts demand prediction for the fourth quarter of the year 2007 of a specific route (A–B) operated by a major airline X.

B. Q2

To address the second question, the final weights of all predictors from the learning algorithm were plotted as a bar graph and compared. Fig. 2 shows the performances or weights of the predictors for a specific route (A–B) operated by a major airline X.

Notes for Fig. 2
1) Average time varying function*: This corresponds to the weight value the predictor: “average demand on Saturdays.”
2) Maximum time varying function**: This corresponds to the weight value of the predictor: “maximum demand on Sundays.”

We have used four different predictors for average statistical function: average demand on Saturdays, average demand on Sundays, average demand on Mondays, and average demand on Wednesdays. The ranges of weight values for these predictors are shown in Fig. 3.

As shown in Figs. 2 and 3, the average time-varying function predictor values ranging from 0.4 to 0.7, indicating that such predictors play major roles in determining air cargo demand trends (set for regular frequency such as daily or
weekly). The performance of an average function is influenced by large volumes of past data. Similarly, the holiday predictor value ranging from 0.4 to 0.65, playing a significant role in the airlines deciding on the number of cargo aircraft that they will need to operate to meet the demand during the holiday season.

We have also plotted the weight values and analyzed other economic predictors, such as the GDP, inflation, industrial growth, and import/export commercial activity, as shown in Fig. 2. It was observed that although the GDP and inflation have an influence in determining the future quarter or annual demand, the corresponding predictors have a relatively lower impact on the daily or weekly demand trends—this agrees well with our intuition in the matter. The air cargo demand arising from imports and exports of a nation/state is predicted to be low, as we have used regional routes for the analysis.

C. Q3

Here we answer the third question in detail.

As stated in the problem description, all airlines operate a certain number of flights (e.g., two flights daily) on specific routes, throughout the quarter/year. Each airline prepares a flight schedule based on its own prediction model(s) and/or according to the prediction provided by the freight forwarders, or predictions from other sources. On a given day, an airline operates different flight types with different capacities. Airlines also operate extra flights for holidays, such as Christmas, New Year’s Day, and so on, to handle the episodic extra demand. Once the season ends, the airlines take out the extra flight(s) and continue with an earlier off-peak schedule. Table II shows the flight schedule for route A–B of airline X.

The above flight schedule is for the period of January 8, 2007 to November 23, 2007, except during the first week of every quarter. During the first week of every quarter and from November 24, 2007 to January 8, 2008, airline X operated one additional flight on all days of the week.

Notes for Fig. 4

1) cp-1: This was the cargo capacity for airline X for route A–B from January 8, 2007 to November 23, 2007.
2) cp-2: This was the cargo capacity for airline X from November 24, 2007 to January 8, 2008.

Here we try to establish a relationship with the results of Stage 1 and Stage 2, i.e., we try to predict the cargo LF based on the predicted demand and the actual cargo capacity.

The predicted cargo LF $LF_{i,t} = \frac{P_{i,t}}{S_{i,t}}$ divided by the actual cargo capacity $S_{i,t}$.

Fig. 5 depicts the predicted cargo LF for route A–B of airline X for the fourth quarter of the year 2007.

For the data indicated in Fig. 5, it was observed that the average predicted cargo LF was 62.5% and it was also observed that actual cargo capacity is underutilized (less than 70%, which is considered as the optimum cargo LF) on most of the Thursdays, Fridays, Saturdays, and Sundays, except during the first week of every quarter and also days from November 24, 2007 to December 31, 2007. An unutilized cargo capacity of 25%–30% is almost equivalent to a wasted cargo capacity of one flight on most days (78% of the days). Hence, we predict that if airline X operates one flight less on these days, the average cargo LF would increase to 66.98%, which is a 7.16% improvement over the previously predicted cargo LF, and an effective improvement of 9.82% over the actual cargo LF. Fig. 6 shows the suggested cargo capacity for route A–B on most days in the year 2007 except during the first week of every quarter, and the days from November 24, 2007 to December 31, 2007.

As mentioned earlier, airlines can, in theory, use various strategies to improve the LF. However, when an airline has a defined flight schedule in place, the “flight deletion” strategy is...
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TTOTAMANE et al.: AIR CARGO DEMAND MODELING AND PREDICTION 9

Fig. 6. Suggested cargo capacity.

TABLE III
CARGO CAPACITY OPTIMIZATION SIMULATION RESULTS

<table>
<thead>
<tr>
<th>Airline</th>
<th>Route</th>
<th>Current Load Factor (%)</th>
<th>Load Factor After Applying Our Model (%)</th>
<th>Effective Improvement (%)</th>
<th>No. of Days Demand Overshoots the Cargo Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>A–B</td>
<td>63.39</td>
<td>69.62</td>
<td>9.82</td>
<td>5</td>
</tr>
<tr>
<td>Y</td>
<td>G–H</td>
<td>66.97</td>
<td>73.13</td>
<td>9.19</td>
<td>2</td>
</tr>
<tr>
<td>Y</td>
<td>E–F</td>
<td>64.31</td>
<td>71.11</td>
<td>10.57</td>
<td>6</td>
</tr>
<tr>
<td>X</td>
<td>C–D</td>
<td>66.23</td>
<td>72.52</td>
<td>9.49</td>
<td>7</td>
</tr>
</tbody>
</table>

the more realistic and feasible among them. Any other strategy would involve adding or rerouting a flight at short notice, and would be difficult to make happen even when suggested in theory, as it would be subject to real-world constraints such as legal and statutory approvals of flight plans, financial and logistical investments on the ground at points of origin and destination, crew availability, and so on.

D. Summary of Results

The sample air cargo demand prediction and cargo capacity planning given above are relevant to the specific instance of the airline X for a route A–B. However, the procedure we demonstrated in a particular case can be applied in a straightforward manner to another commercial carrier and with other operating routes. We have simulated the same procedure for four instances, covering two different airlines and four different routes. The results are shown in Table III.

The first simulation was done on airline X for route A–B, where the actual average cargo LF was 63.39% for the year 2007. If airline X had used our prediction model and followed suggested cargo capacity, the average cargo LF would have been 69.62%, which is an effective improvement of 9.82%. However, the actual demand overshoots the cargo capacity on five occasions or five days, which is merely 1.36% of the year.

The second simulation was done on airline Y for route G–H. The weighted majority approach with the same set of predictors used in the previous simulation was able to predict the cargo demand 76.71% of the time (280 days in a year), with a tolerance level of ±25%. Airline Y operates a variety of flights (Boeing 757-200, Boeing 762, and Boeing 763) with capacities ranging from 27 to 76 tons, and daily cargo capacity ranging from 209.5 to 393 tons over the weekdays, with 209.5 tons of cargo capacity being the lowest on Mondays and Fridays, and 393 tons on Saturdays. The predicted average cargo LF was 66.97%, and the actual cargo capacity was underutilized (with a cargo LF less than 70%) on most days in the year 2007, except during the last quarter of that year. By removing certain flights on certain days of every quarter, the actual cargo LF could have been increased from 66.97% to 73.13%, which is an effective improvement of 9.19%. Fig. 7 shows the actual cargo capacity of the route and Fig. 8 shows the suggested cargo capacity for the route. The legends Q1–Q4 represent quarters of the year 2007.

The third simulation was done on airline Y for route E–F. The weighted majority approach with the same set of predictors used in the previous simulation was able to predict the cargo demand 75.35% of the time (275 days in a year) with tolerance level of ±25%. Airline Y operates a variety of flights (Boeing 757-200, Airbus 300-600R, and Boeing 767-300ER) with capacities ranging from 44.45 to 66.6 tons, and daily cargo capacity ranging from 414.02 to 754.51 tons over the weekdays, with 414.02 tons of cargo capacity being the lowest on Mondays and Fridays, and 754.51 tons on Saturdays. As discussed previously, the predicted average cargo LF was...
VI. CONCLUSION

The simulations carried out on multiple airlines for different segments showed that the algorithm used was able to predict cargo demand with reasonably high accuracy. In this paper, we modeled the problem using some of the generic predictors, such as time-varying functions, holidays, weekends, and so on. Additional predictors, such as industrial output, industrial growth, GDP, and so on, were considered but not found to make a significant overall difference. The predictors used played a significant role in improving the efficiency of our air cargo demand predictions. The air freighter fleet forecast analysis may then proceed to consider many other factors—such as airplane capability, performance, and availability, the regional domicile level for each airline, airport slot availability, accounting for variables such as individual flight type and age, airplane size, cargo volume, retentions, utilization, strategies, and so on.

Accurate predictions of cargo demand in our simulations helped us to apply the flight deletion strategy and thereby improve airlines’ average cargo LFs by 9%–12%. However, in real-life scheduling of freight aircrafts, a thorough analysis and understanding of top-down air cargo flows and freighter capabilities can be combined with detailed bottom-up information on specific regional and operator trends and strategies. The algorithm used here can also be used in air cargo logistics management solutions, such as Unisys LMS, and the overall benefits or improvements can be benchmarked with any other enterprise cargo logistics management solution.

In the future, this work can be used to evaluate various types of predictors, which, in turn, would lead to a better understanding of the factors that influence air cargo demand. Economic models of industry and GDP growth are useful in a certain measure, but do not usually offer a concrete means of comparison of such factors with others. Using our approach, it is possible to add such factors, (e.g., predictors) in the WMA, and the stable weights (see Fig. 2) attained by the predictors can then give us a clear idea about the degree of influence that each such factor has on the matter at hand.

VII. ACKNOWLEDGMENT

The authors would like to thank H. Gupta, V. Sanjeeva, G. Rajan, and L. N. V. Samy, who facilitated this research by arranging access to much needed data. The authors would also like to thank the reviewers for their many insightful comments, which helped to improve this paper, and the participants and reviewers of the Third Annual IEEE International Systems Conference, held in March 2009 in Vancouver, BC, Canada, for their detailed and insightful comments and suggestions on a preliminary extended abstract [54] as a part of this paper. The authors also acknowledge the help of R. Mulukutla in the earlier paper.

REFERENCES

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TOTAMANE et al.: AIR CARGO DEMAND MODELING AND PREDICTION


Raghavendra Totamane received the B.E. degree from Bangalore University, Bangalore, India, in 2001. He has over 11 years of experience in the information technology (IT) industry and was with Unisys Global Services India, Bangalore, and Orange Business Services, Bangalore. He is currently an independent IT Consultant in Bangalore, India. His current research interests include supply chain management, service-oriented architecture, IT infrastructure, cloud computing, and transportation systems.

Amrit Dasgupta received the B.Sc. degree in computer science from Bangalore University, Bangalore, India, in 2002, and the M.C.A. degree from the Biju Patnaik University of Technology, Orissa, India, in 2005. He has over six years of experience in the information technology (IT) industry. Since 2006, he has been with Unisys Global Services India, Bangalore. He has worked in different domains, such as disaster recovery, logistics management, and enterprise applications. His current research interests include IT infrastructure, cloud computing, and transportation systems.

Shekhar Rao (M’08) received the M.S. degree in logic and computation from Carnegie Mellon Univer- sity, Pittsburgh, PA, and the Ph.D. degree in computer science from the University of Iowa, Iowa City. He is currently an Associate Professor with the International Institute of Information Technology Ban- galore, Bangalore, India. His current research in- terests include distributed computing, especially algorithms and formal methods for concurrent and distributed systems, solar energy and microgrids, cloud computing, energy-aware computing ("green IT"), distributed algo- rithms in sensor networks and other domains, algebraic theories of fault tolerance and systems safety, safety-critical systems, and demand-side resource management. Dr. Rao is a member of the ACM, the American Mathematical Society, and the Computer Society of India.