MODELING CARGO DEMAND FOR REGIONAL AIR TRANSPORT NETWORKS IN CANADA AND THE UNITED STATES

Martin M. Alexiev[1], Ruben E. Perez[2], Peter W. Jansen[3] & Jeremy C.-H. Wang[4]

[1]Research Assistant, Department of Mechanical and Aerospace Engineering, Royal Military College of Canada, PO Box 17000, Station Forces, Kingston, ON, K7K 7B4, Canada
[2]Professor, Department of Mechanical and Aerospace Engineering, Royal Military College of Canada, PO Box 17000, Station Forces, Kingston, ON, K7K 7B4, Canada
[3]Assistant Professor, Department of Mechanical and Aerospace Engineering, Royal Military College of Canada, PO Box 17000, Station Forces, Kingston, ON, K7K 7B4
[4]Chief Operating Officer, Ribbit Inc, 100 King Street West, Suite 6200, Toronto, ON, M5X 1B8, Canada

Abstract

Air cargo networks are an important part of aircraft operations, used for the on-time delivery of time-sensitive cargo or servicing isolated communities with no or limited road access. This research proposes a demand model for air cargo, focusing on regional feeder networks. To develop the model, multi-year historical Automatic Dependent Surveillance-Broadcast (ADS-B) data is processed and aggregated to estimate demand with a-priori knowledge of cargo network operations. The demand model then builds upon the gravity model typically used for passenger demand, adjusting it to account for the geographic and socio-economic parameters that affect the demand for air cargo. Canada and the United States were chosen as the test cases to assess the model's performance in two culturally similar countries, with possible large and sparsely populated regions for which air cargo operations are essential. The model is first applied to the lower-population markets of Canada before scaling the model to a larger market such as the United States. The impact of different factors on the test cases' demand is explored. In the Canadian context, the developed model achieves a high level of prediction accuracy with a coefficient of determination larger than 0.88. For the United States test case, a lower prediction accuracy is achieved where regional factors play a more significant role.

Keywords: Demand Model, Air Cargo, Gravity Model, System of Systems, Regional Aircraft

Nomenclature

$ADS-B$ Automatic Dependent Surveillance-Broadcast

AI Artificial Intelligence

B Distance to Nearest Body of Water (Great Lake or Ocean [nmi])

$B1900D$ Beechcraft 1900D

C Catchment Area Population Size

$C208B$ Cessna 208B Cargomaster

D Greater Circle Distance [nmi]

F Sate FRASE Index

$FRASE$ Federal Regulations and State Enterprise Index

H Distance to Nearest Highway [nmi]

I	Average After-Tax Income [CAD or USD]
$IATA$	International Air Transport Association
$ICAO$	International Civil Aviation Organization
KM	K-Means Clustering Algorithm
$No.$	Number
P	Population Size
R^2	Coefficient of determination
S	State Sales Tax
T	Average Travel Time [Hrs]
US	United States
V	Cargo Demand [lbs]
a	attractive factors of demand
ac	Aircraft
f	Cargo Density factor
g	generative factors demand
i	Origin Airport
j	Destination Airport
k, α, β, γ	weighting propotional coefficients for factors of demand
r	resistance factors of demand
v	Aircraft Max Payload

1. Introduction

Air transportation systems comprise both cargo and passenger demand. Previous studies have primarily concentrated on modeling the allocation of passenger demand between origin-destination pairs. However, the aging cargo fleet in North America presents new challenges, particularly at the regional level. By 2030, most of North America's regional aircraft cargo fleet is projected to phase out [1, 2, 3, 4, 5, 6]. The emergence of pilotless flight technologies offers an opportunity to replace the aging North American regional cargo fleet, such as reducing operational and personnel costs, expanding flight schedule flexibility, and exploring new regional flight routes.

Studies related to air transportation systems typically involve modeling demand for passengers or cargo for the allocation of aircraft or routes [7, 1, 8, 9, 4, 2, 3, 10, 11, 12]. A demand model predicts the need for a particular service based on various factors, which can be modeled using numerical or artificial intelligence (AI) methods [13]. Using numerical forecasting methods, geographical and socio-economic factors that affect passenger and cargo demand are typically used to predict demand between origin-destination pairs [14]. A gravity model is often used to assess the impact of these factors on demand, providing a straightforward method to represent a demand model while assessing the direct effects of each input parameter [7, 1, 8, 9, 4, 2, 3, 10, 11, 12].

Previous studies on air cargo operations have primarily focused on examining a single hub-and-spoke cargo network's on global or national route network structures [15, 16, 17]. For instance, *Hwang and Shiao* [15] conducted research on international air cargo traffic through Taiwan Taoyuan International Airport, where they used various factors such as total weight, Gross domestic products per capita, the product of populations, distance, annual flight frequency, and freight rate per unit of weight of each origin-destination pair to determine the coefficient of determination found to be 0.995. Although these studies have demonstrated a strong correlation between socioeconomic factors and demand prediction for a sole hub airport, they do not consider how other factors impact cargo demand in a network consisting of multi-hub airports.

This research aims to modify the commonly used gravity model methodology for passenger demand modeling of multi-hub networks and adapt the model for air cargo network demand, using the regional feeder networks in Canada and the United States (US) as test cases. The model uses multi-year historical Automatic Dependent Surveillance-Broadcast (ADS-B) data in combination with prior knowledge of cargo networks to determine the current state of each country's regional feeder network and estimate demand. The study aims to identify the impact of new and currently used socio-economic and geographical factors on demand to create an accurate demand model that can be applied to various regional markets. The data acquisition process for the route network and the gravity demand model formulation is detailed in Section 2. Section 3. applies the methodology to test cases of Canada's and the United States air cargo regional networks. The results of applying the demand model, developed in the methodology, to both markets are presented in Section 4. and the drawn conclusions and future works drawn from the results are discussed in Section 5.

2. Methodology

The methodology delineates the approach for collecting air flight data and inputting the data into a generated predictive demand model to quantify the allocation of air cargo. Canada and the United States were chosen as the test cases for the methodology to assess the performance of the demand model in a comparable, lower-population market, like Canada, before scaling the model for a larger market such as the United States.

No open-source database provides all the required information to picture each country's route networks. The data gathering challenge of air freight networks is subjected to the following challenges. The first challenge is determining who the air freight carriers within the region of interest are. Once the air freight carriers of a region are determined, the challenge becomes identifying their fleet composition. After identifying the fleet's composition, the next level of the challenge becomes determining the fleet's hub airports and flight routes. Finally, the task is to determine the quantity of cargo, measured in weight, handled on each route.

2.1 Regional Air Cargo Carriers

Regional air cargo networks predominately consist of smaller operators contracted by large freight carriers. Current Canadian airfreight liners were established using prior knowledge of cargo network operations. Air Freightliners in the United States were specified from *Jordan, Nadezhda, and Husni's* research on *Preliminary Characterization of Unmanned Air Cargo Routes Using Current Cargo Operations Survey* [5]. Primary operators of regional air cargo networks in both countries are included in Table 1, and Table 2.

The fleet of each Air Cargo Carrier was derived from FlightRadar24's open-source database of airlines, including their respective fleet, aircraft types, and tail numbers. Using prior knowledge of air cargo operations, regional routes are typically flown using small, turboprop utility aircraft. The two leading aircraft that regional cargo networks are predominately represented by are the Cessna 208B Cargomaster (C208B) and Beechcraft 1900D (B1900D). As such, only aircraft of similar size to a C208B or B1900D were considered to be a part of the regional network. Concerning the test cases of Canada and the United States, an aircraft considered part of each country's regional air cargo network must have flown only within Canada or the United States. Table 3 notes all the aircraft types considered for the test cases' regional air cargo networks.

Table 1 – Regional Canadian air cargo carriers included in this research.

Air Cargo Carriers	ICAO Designator	No. of Aircraft Tracked
Fedex	FDX	8
RiseAir	WEW	23
Skylink Express	SLQ	5

Table 2 – Regional American air cargo carriers included in this research.

Air Cargo Carriers	ICAO Designator	No. of Aircraft Tracked
Alpine Air Express	AIP	61
Ameriflight	AMF	112
Castle Aviation	CSJ	21
Corporate Air	CPT	3
FedEx	FDX	243
Freight Runners Express	FRG	21
IFL Group	IFL	5
Key Lime Air	LYM	20
Martinaire	MRA	26
Wiggins Airways	WIG	18

Table 3 – Regional air cargo aircraft included in this research.

Aircraft Name	Aircraft Abbreviation
Cessna 208 Cargomaster	C208
Dassault Falcon 20	FA20
Piper PA-31 Navajo	PA31
Saab 340	SF34
Swearingen Merlin	SW(3)(4)
Beechcraft 1900	B190
Beechcraft Super King Air 200	B200
Beechcraft Model 99	B99
Cessna 560X Citation Excel	C56X
Cessna 750X Citation 10	C750
Gulfsteam Aerospace G650	GLF6
Embraer Bandeirante 110	E110
Embraer Brasilia 120	E120
Embraer ERJ-135	E135
airchild Dornier 328JET	J328
Aerostar (1) 600	AEST
Beech 400 beechjet	BE40
Beechcraft T-34 Mentor	T34P
Cessna 408 SkyCourier	C408
ATR-42-300/320	AT43

2.2 Initial Cargo Estimation

Due to the lack of publicly available information concerning the cargo weight an aircraft holds during a given flight, cargo is treated as a resource allocation problem such that it is allocated to maximize utility. The cargo demand on each route is assumed to be limited by volume such that Air Cargo Carrier companies estimate that the payload of each route has a cargo density of 100 kg/m^3. The demand is also considered asymmetric between origin-destination city pairs such that returning and connecting flights are assumed to have a cargo density of 50 kg/m^3. Weekly cargo-by-weight [lbs] (V), between an origin (i) - destination (j) pair (i, j), is estimated as the sum of the products of cargo density factor (f) and max payload of a given aircraft flying the given route (v) as shown in Equation 1. (n) denotes a particular aircraft, where (ac) is the total number of aircraft in the route network.

$$V_{i,j} = \sum_{n=1}^{n_{ac}} f_{i,j,n} * v_{i,j,n} \tag{1}$$

2.3 Demand Model

A gravity demand model was chosen to model cargo-by-weight demand as it provides a simple numerical method to quantify the effects of each geographic and socio-economic factor. A gravity model follows Newton's Law of gravitational force such that demand is a function of the generative factors (g) of the origin airport (i), the attractive factors (a) of the destination airport (j), obstructed by any resistance factors (r) between the origin-destination pair (i, j) weighted by proportional coefficients $(k, \alpha, \beta, \gamma)$, shown in Equation 2.

$$V = k \frac{g_i^{\alpha} a_j^{\beta}}{r_{ij}^{\gamma}} \tag{2}$$

The first gravity model is a baseline consisting of commonly used geographical and socio-economic factors in passenger models cited in *Tobias, Franz, Armin's* research on *Gravity Models for Airline Passenger Volume Estimation* [14], as shown in Equation 3.

$$V_{i,j} = P_{i,j}^{k_1} C_{i,j}^{k_2} D_{i,j}^{k_3} T_{i,j}^{k_4} \tag{3}$$

Each parameter in Equation 3 above is specified in Table 4.

Table 4 – Variable definitions for Baseline Gravity Demand Model.

Variable	Description
$P_{i,j} = P_i P_j$	Population Size
$C_{i,j} = C_i / C_j$	Population Catchment Area
$D_{i,j}$	Greater Circle Distance [nmi]
$T_{i,j}$	Travel Time [Hrs]

The gravity model presented in this research builds upon the baseline model, introducing additional unique parameters which enable better prediction to the modelling of cargo demand, shown in Equation 4. Distance to the nearest highway is included in the proposed model, as cargo demand is influenced by road networks that freight trucks use for last-mile delivery. Lastly, median income was chosen as an additional parameter to develop the baseline model, as an individual's purchasing power to buy goods is influenced by their income, thereby influencing cargo demand.

$$V_{i,j} = P_{i,j}^{k_1} C_{i,j}^{k_2} D_{i,j}^{k_3} T_{i,j}^{k_4} H_{i,j}^{k_5} I_{i,j}^{k_6} \tag{4}$$

Each parameter in Equation 4 above is specified in Table 5.

Table 5 – Variable definitions new proposed factors.

Variable	Description
$H_{i,j} = H_i + H_j$	Distance to Nearest Highway [nmi]
$I_{i,j} = I_i I_j$	Median Income [CAD or USD]

3. Regional Air Cargo Flight Networks of Canada and The United States

This section presents the current Canadian and United States Regional air cargo networks used to evaluate the cargo demand model. Two years of historical ADS-B data of Origin-Destination flight paths of each airline's regional fleet were allocated between October 14, 2021, and October 14, 2023, to establish each country's current state of regional air cargo network. Dates a year after 2020 were chosen to mitigate effects that increased cargo demand during the COVID-19 pandemic might have had on the air cargo industry. The air cargo network for each market was simplified to a weekly basis to remove any one-off routes. As such, for a route to be considered a part of the network, it must have been operated at least once a week.

3.1 Canadian Network

During the above-mentioned period, the Canadian regional air cargo network had 36 planes that flew 28783 flights to 28 cities on 39 weekly routes. Every half-hour, a regional air cargo flight occurs in Canada. A view of Canada's regional air cargo flight network is shown in Figure 1, with a close-up of different regions shown in Figure 2.

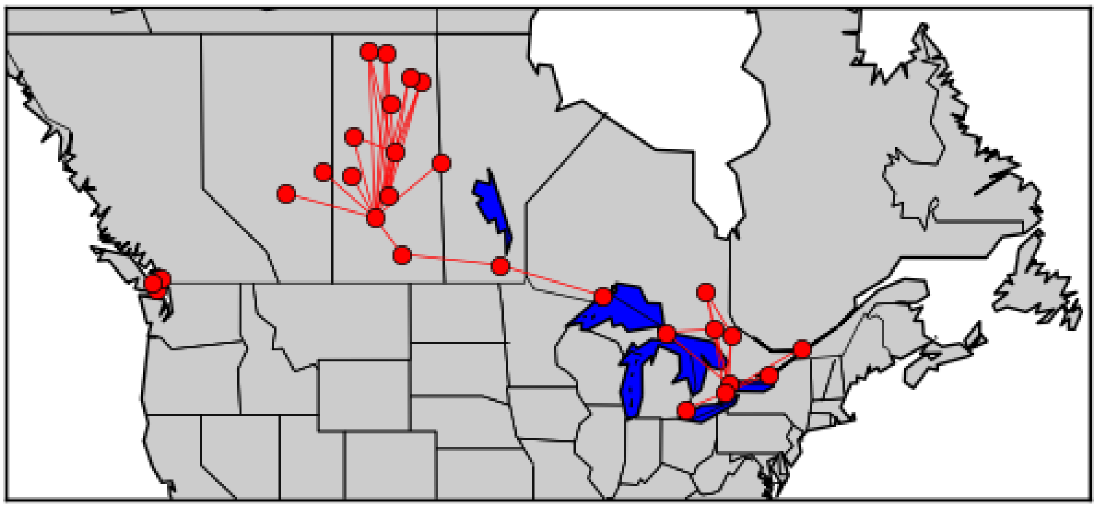
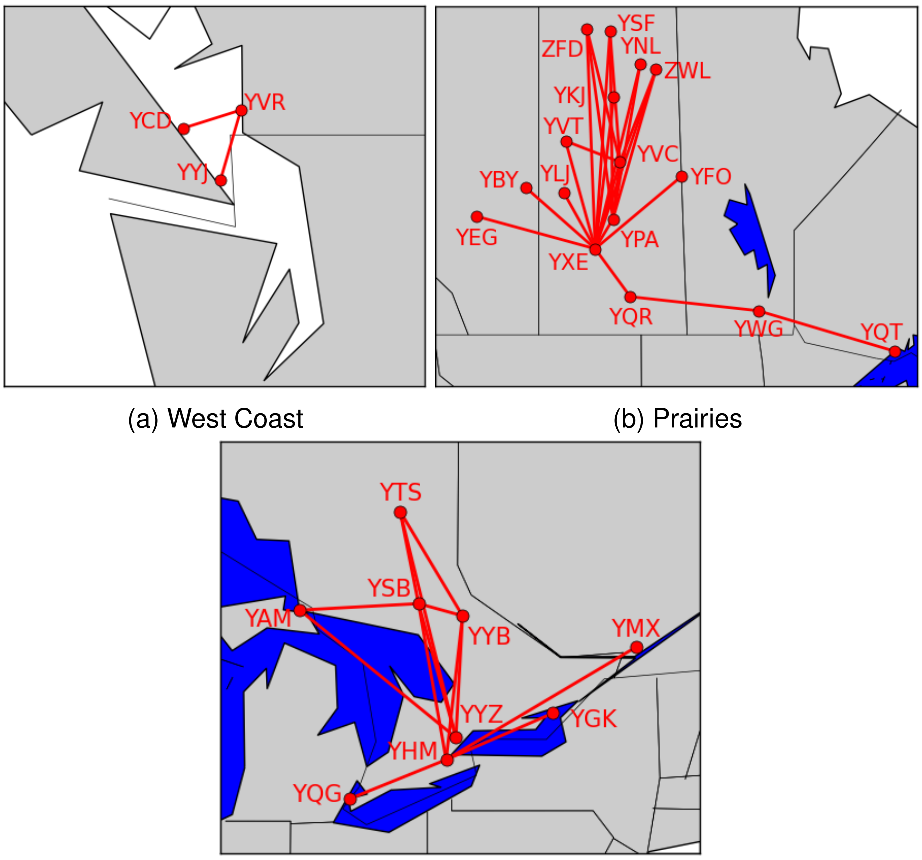


Figure 1 – Canadian regional air cargo network from October 14, 2021, to October 14, 2023.



(a) West Coast

(b) Prairies

(c) Central

Figure 2 – Close-up of Canada's regional air cargo network in different regions.

Canada generally follows the traditional Hub-Spoke route networks with minor changes to its structure in different regions. The West Coast only flies short-distance to Vancouver Island, which originates from Vancouver. In the Prairies, most routes originate from Saskatoon and spoke northwards on long-distance flights to remote communities. The Prairies hub and spoke network has routes that connect from Saskatoon, heading east to Winnipeg in Manitoba, then Thunder Bay in Northern Ontario. Central Canada Consists of two nearby hub cities, Hamilton and Toronto. Both these hub cities spoke out to span across the Great Lakes into Montreal in Quebec and as north as Timmins in Ontario, predominately using connecting routes. Most hub airports in the Canadian network are near the United States border, suggesting that road networks distribute time-sensitive trade with the United States by utilizing air cargo networks to span into Canada.

3.2 United States Network

Throughout the two years mentioned above, the United States Regional Air cargo network had 530 aircraft that flew 449652 times to 329 cities, creating 521 weekly routes. In this network, 13 flights occur every half-hour. That is 13 times more flights every half-hour compared to the Canadian market. The United States regional air cargo network is depicted in Figure 3, filtered to different regions based on the origin airport shown in Figure 4.

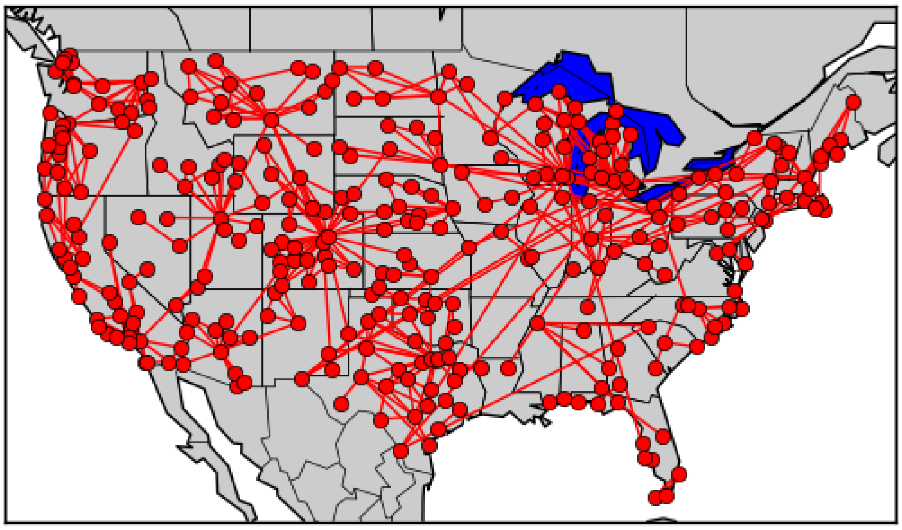
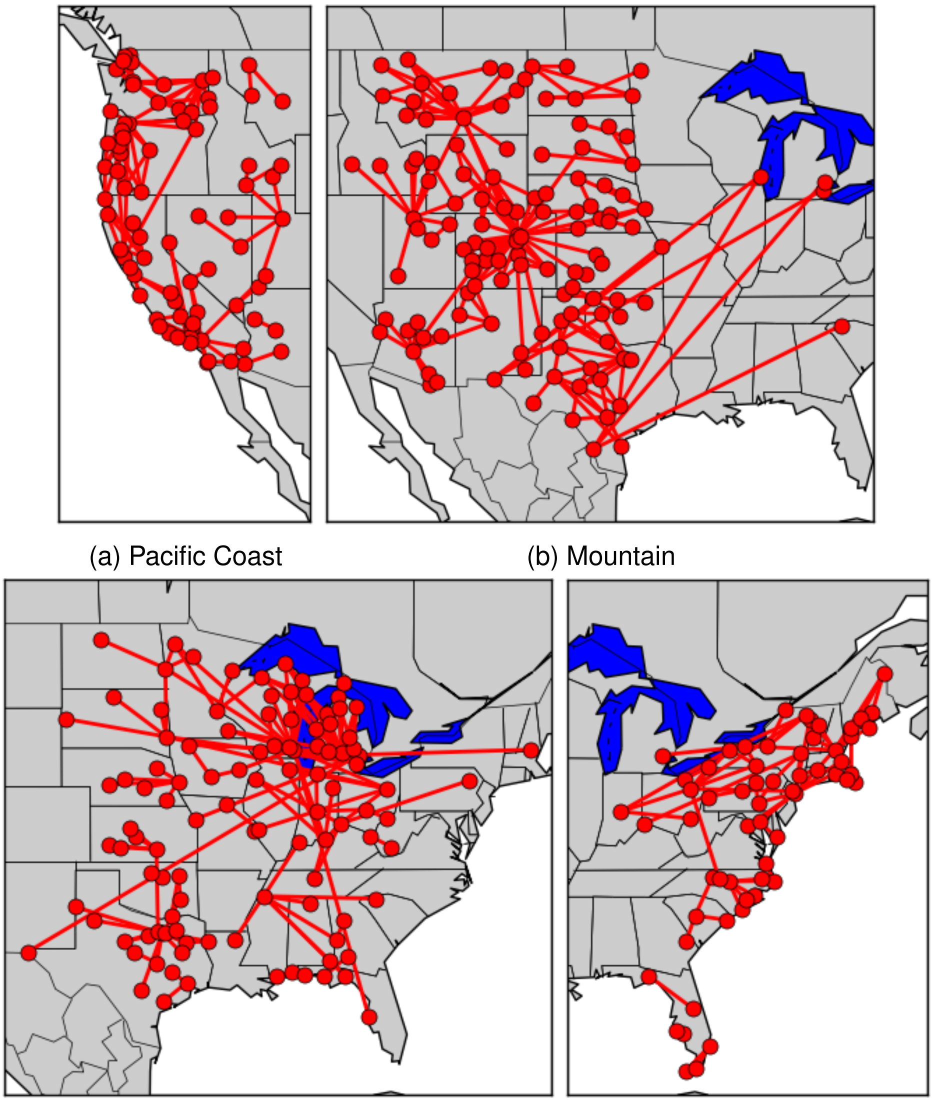


Figure 3 – US regional air cargo network from October 14, 2021 to October 14, 2023

Route patterns in the United States can be categorized into two main behaviours: coastal and landlocked. States that line the coasts or the Great Lakes have short-distanced routes concentrated along the coastline while also having routes that spoke out from the coastline to nearby states. Concentrated routes near the coastlines can be attributed to the time-sensitive nature of air cargo. The coastal behaviour is noted in Figure 4 a and d. Intra-state versions of a typical Hub-Spoke route network are observed for landlocked states, shown in Figure 4 b and c. The intra-state hub-spoke networks in landlocked states typically have a hub airport that borders another state and spokes towards the other side of the state longitudinally and laterally. Feeding into these intra-state hub-spoke routes are tails that move towards neighbouring states, closer to the nearest coastline. A hypothesis for the behaviours seen in landlocked states is that time-sensitive cargo arriving at ports on the United States coast uses air cargo route networks for on-time regional demand rather than less timely road networks.



(a) Pacific Coast
(b) Mountain
(c) Central
(d) East Coast

Figure 4 – Close-up of different regions in US regional air cargo network.

4. Results and Discussion

This section first applies the baseline and proposed demand model to the Canadian test case and further expands on understanding and improving its performance. The findings from the Canadian test case were then applied to the United States test case to validate the demand model performance with a different market.

4.1 Determination of Hub Airports

By generating a contour density agglomeration of initial weekly cargo-by-weight demand plotted against all theoretical origin-destination pairs, hub airports of a given network can be observed. Figure 5 below describes a method for identifying hub airports within a route network based on the number of route pairs with other airports applied to the Canadian network.

Dots that create vertical lines in Figure 5 denote the hub airport as the index with its spoke airports. Horizontal lines on the graph represent the same hub airports as the vertical lines. This artifact occurs because the contour densities are reflected about the diagonal due to plotting airport indices against themselves. The specks that form zigzags on the contour are due to the connecting flight loops from a hub city in Central Canada, visualized on a map in Figure 2. Hub airports of Canada's regional air cargo network, determined from Figure 5, are shown in Table 6.

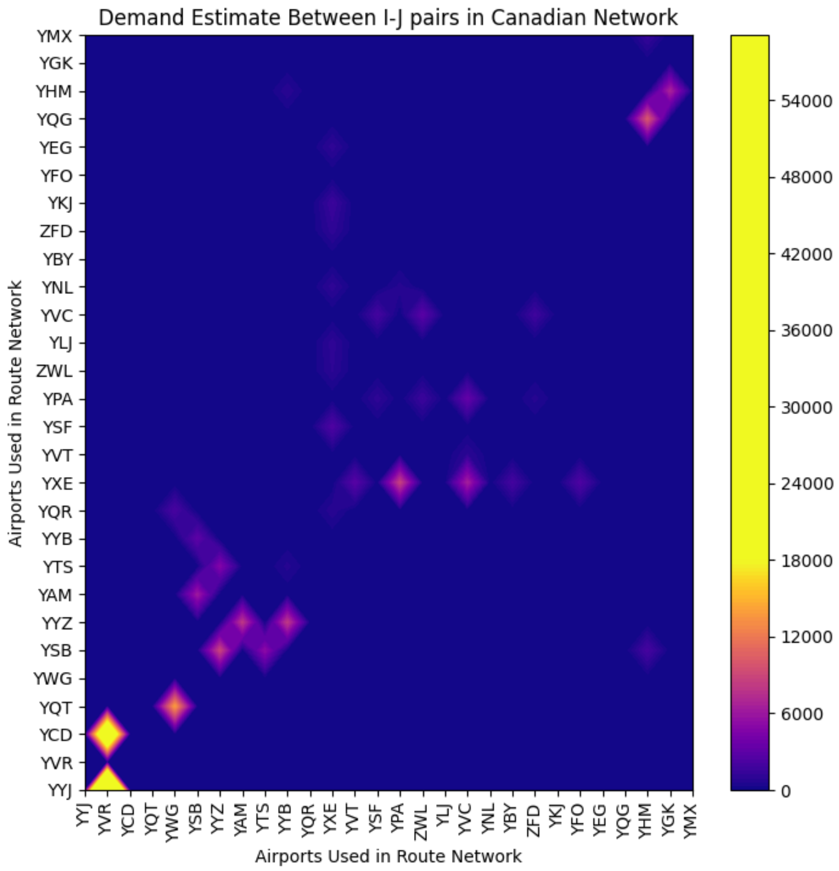
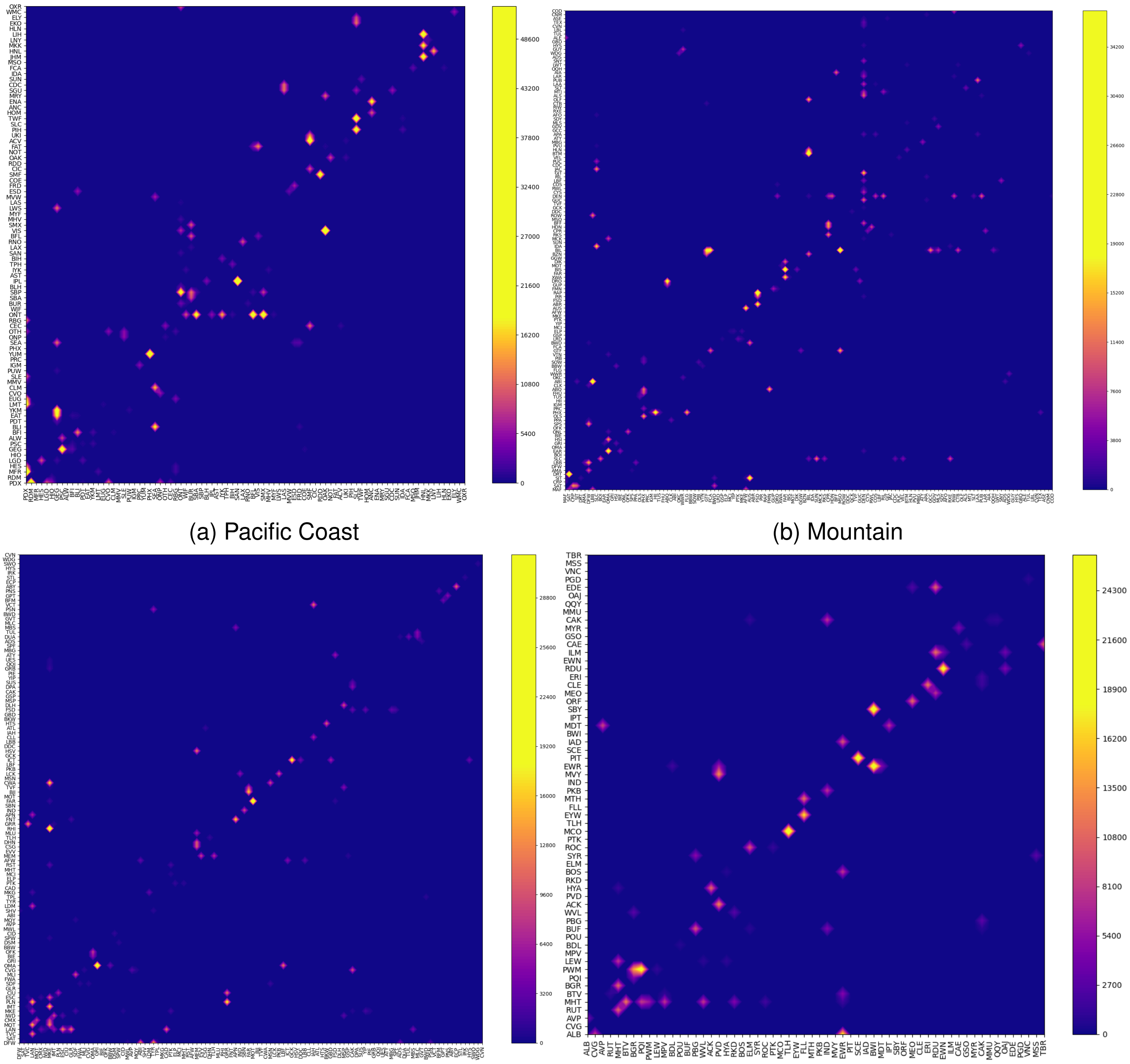


Figure 5 – Initial weekly estimated demand contoured between origin-destination pairs in Canada's network.

Table 6 – Hub airports within Canada's regional air cargo network.

Airport IATA Code	Airport City
YVR	Vancouver
YXE	Saskatoon
YYZ	Toronto
YHM	Hamilton

Figure 6 applies the same concept shown in the Canadian market, shown in Figure 5 to the United States market, but sectioned using the same sections shown in Figure 4 for visualization purposes. As the total US routes were divided into sections based on the origin airport, some asymmetry appeared across the diagonals of the plots. As such, hub airports were determined by vertical and horizontal lines. Table 7 denotes the hub airports in the United States market.



(a) Pacific Coast

(b) Mountain

(c) Central

(d) East Coast

Figure 6 – Initial weekly estimated demand contoured between origin-destination pairs in the US network.

Table 7 – Hub airports within the US regional air cargo network.

Airport IATA Code	Airport City	Airport IATA Code	Airport City
IAH	Houston	MLS	Miles City
CAK	Akron-Canton	DEN	Denver
MSP	Minneapolis–Saint Paul	MSO	Missoula
PKB	Williamstown	CPR	Casper
MOT	Minot	BIL	Billings
FAR	Fargo	AUS	Austin–Bergstrom
APN	Alpena	GTF	Great Falls
GRR	Grand Rapids	WWR	Woodward
MEM	Memphis	PHX	Phoenix
ESC	Escanaba	SLC	Salt Lake City
IMT	Iron Mountain	LBB	Lubbock
MKE	Milwaukee	MAF	Midland
LAN	Lansing	SAT	San Antonio
DFW	Dallas Fort Worth	SJT	San Angelo
PLN	Pellston	EAR	Kearney
CIU	Kincheloe	ABQ	Albuquerque
OMA	Omaha	ABI	Abilene
AFW	Fort Worth	DRO	Durango
IND	Indianapolis	MTJ	Montrose
CWA	Mosinee	WDG	Enid
ICT	Wichita	HNL	Honolulu
FSD	Sioux Falls	ANC	Anchorage
DUA	Durant	OAK	Oakland
SDF	Louisville	SMF	Sacramento
RDU	Raleigh-Durham	LAS	Las Vegas
MEO	Manteo	SMX	Santa Maria
BWI	Baltimore	VIS	Visalia
EWR	Newark	BFL	Bakersfield
FLL	Fort Lauderdale	IYK	Inyokern
PVD	Warwick	BUR	Burbank
MHT	Manchester-Boston	ONT	Ontario
ALB	Albany	SEA	Seattle–Tacoma
PWM	Portland	PSC	Pasco
BUF	Buffalo	GEG	Spokane
WVL	Waterville	PDX	Portland
SYR	Syracuse	LGD	La Grande
ROC	Rochester	EUG	Eugene
MDT	Middletown	OTH	North Bend
ILM	Wilmington	CEC	Crescent City
CAE	West Columbia	SBP	San Luis Obispo
EDE	Edenton	IPL	Imperial
LAA	Lamar	ESD	Eastsound
FAT	Fresno	SGU	St. George

4.2 Demand Model Results

The objective of the proposed demand model is to generate a model that can be applied to complete air cargo regional markets. Concerning the test cases of Canada and the US, the objective is that the demand model improves upon the baseline model, and the model can explain a majority of variation within the route network. In other words, the demand model can quantitatively explain the variation by a coefficient of determination (R^2) greater than 0.5. A summary of the results for the baseline and proposed model on Canada and the US regional air cargo networks is shown in Table 8 and Table 9, respectively.

Table 8 – Summarized results of baseline and proposed demand model applied to the Canadian regional air cargo network.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	0.682	0.396	-2.002	-1.169	-	-	0.857
+ $H_{i,j}$	0.575	0.439	-1.368	-1.609	-0.538	-	0.874
+ $I_{i,j}$	0.851	0.613	-22.952	-18.986	-	4.763	0.865
$H_{i,j}$ + $I_{i,j}$	0.400	0.324	13.971	-16.464	-0.761	-3.427	0.882

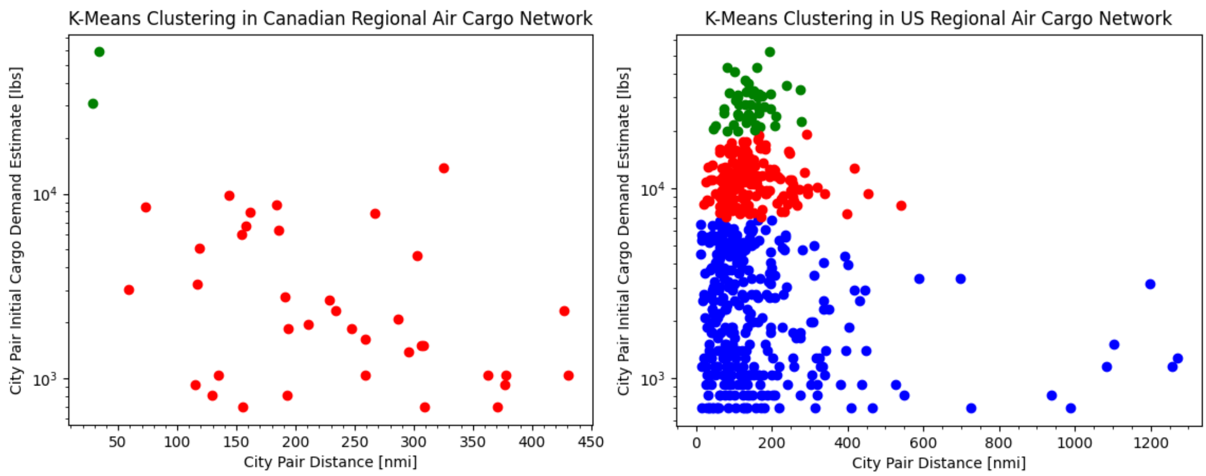
Table 9 – Summarized results of baseline and proposed demand model applied to the US regional air cargo network.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	0.070	-0.028	1.443	-1.452	-	-	0.034
+ $H_{i,j}$	0.070	-0.026	1.430	-1.434	0.075	-	0.036
+ $I_{i,j}$	0.065	-0.025	0.233	-0.212	-	0.283	0.049
$H_{i,j}$ + $I_{i,j}$	0.065	-0.024	0.258	-0.237	-0.030	0.275	0.049

Evaluating the baseline demand model for the Canadian regional air cargo network, shown in Table 8, correlates well with an R^2 value of 0.857. Changing the demand model to the proposed new cargo model increases R^2 by 3% to 0.882. Comparing the demand model results to those of the US market, shown in Table 9, the baseline model performance is 0.034, an increase of 44% to 0.049 when using the proposed model. When only including median income on top of the baseline model in the US, the R^2 increases by the same percentage, suggesting that the route's accessibility to a highway is not as important in the US route network relative to the Canadian network. The Canadian network baseline and proposed model generally perform well and surpass the target R^2 of 0.5 by 76%, while both US models shortfall the criteria by 90%, which signifies that different factors drive the US air cargo regional network.

4.3 Demand Model Improvements

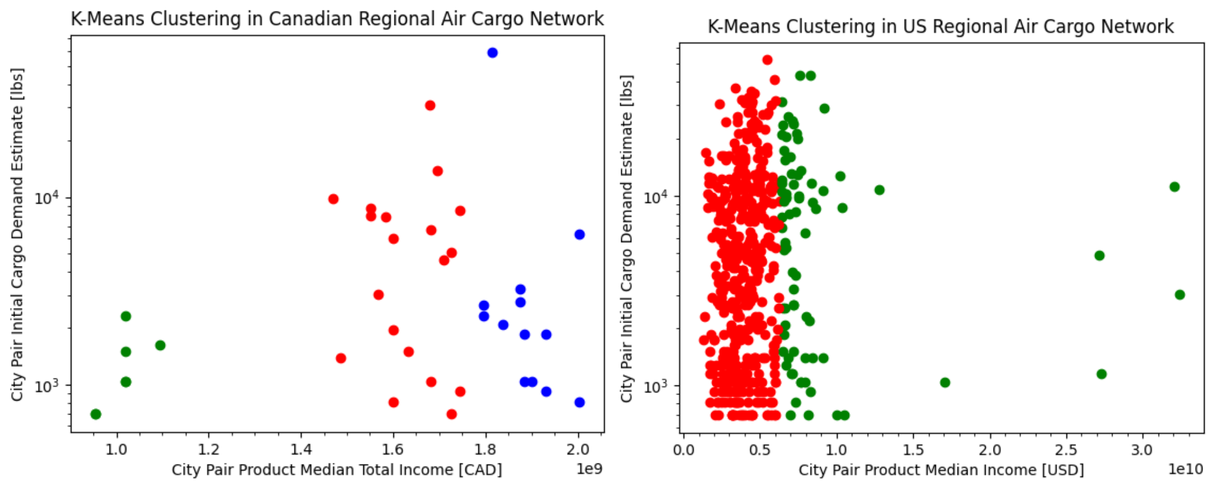
Although the initial results of the proposed demand model on both networks suggest that the Canadian and US networks rely on different variables, it does not signify that improvements can be made to both models by filtering routes by the same parameter. Passenger demand models are typically filtered by distance into a set of ranges [12]. Figure 7 applies the method of filtering by distance to both test cases using a K-Means Clustering Algorithm (KM). A K-Means Clustering Algorithm is an unsupervised machine learning algorithm that divides the data points of specified dimensions into several clusters to minimize the sum of squares within a cluster [18].



(a) Canada (b) US

Figure 7 – K-Means clustering results with respect to distance in both test cases.

Filtering initial demand estimates by city pair distance using a K-means clustering algorithm in both test cases of the Canadian and US air cargo regional networks does not create ranges of distance based on demand; instead, it creates ranges of demand based on distance. Hence, city pair distance cannot be used for a systematic approach to creating filtering ranges to apply the proposed cargo demand model. However, using a K-Means clustering algorithm on the median income in both networks provides explicit ranges to which the proposed demand model can be applied, as shown in Figure 8. Results of using the plots in Figure 8 as a reference to filter the Canadian and US markets by median income in their respective currencies are found in Table 10 and Table 11, respectively.



(a) Canada (b) US

Figure 8 – K-Means clustering results with respect to product median income between city pairs in both test cases.

Table 10 – Proposed demand model applied to product median income filtered Canadian network.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	0.682	0.396	-2.002	-1.169	-	-	0.857
$H_{i,j}$ + $I_{i,j}$	0.400	0.324	13.971	-16.464	-0.761	-3.427	0.882
$I_{i,j} < 1.4 \times 10^9$	0.024	0.034	-31.320	31.214	-0.063	8.001	0.997
$1.4 \times 10^9 < I_{i,j} < 1.75 \times 10^9$	0.147	-0.060	-5.670	5.282	-0.345	1.617	0.826
$I_{i,j} > 1.75 \times 10^9$	1.020	1.146	-18.863	10.954	6.822	3.238	0.993

Table 11 – Proposed demand model applied to product median income filtered US network.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	0.070	-0.028	1.443	-1.452	-	-	0.034
$H_{i,j}$ + $I_{i,j}$	0.065	-0.024	0.258	-0.237	-0.030	0.275	0.049
$I_{i,j} < 0.65 \times 10^{10}$	0.063	0.034	-1.878	1.927	0.095	0.768	0.068
$I_{i,j} > 0.65 \times 10^{10}$	0.046	-0.019	2.008	-1.967	-0.114	-0.090	0.020

Filtering initial demand estimate by median income for the Canadian market shows a significant improvement to the proposed model R^2 of 13% and 12.6% to 0.997 and 0.993 of the lower and upper median income brackets, respectively. In contrast, Canada's middle median income group saw a decrease of -6% to R^2 when the model was applied. Negatively proportional factors within the middle median income bracket in the Canadian market were city pair distance and highway distance, inferring that these factors either deter or are not considered a factor for demand. The US network saw an R^2 increase of 39% to its lower median income bracket but still under the criteria of 0.5 R^2, while its upper median income bracket denotes no correlation to the model. The discrepancies between the R^2 in both test cases further suggest different factors further affect the US network.

4.4 Improving Proposed Demand Model for US Test Case

Observing the varying network behaviour depicted in Figure 4 and described in Section 3.2the sparsity of the network hints that intra-state and nearby inter-state route behaviour is independent of the overall US network. Thus, breaking down the US network from coast to coast, nearby inter-state and intra-state routes could improve the model's performance.

4.4.1 Coast to Coast Breakdown

Results of breaking down the US network into four sections based on time zones, shown in Figure 4 starting from the East Coast to the Pacific Coast, and applying the proposed demand model along with filtering using median income are shown in Table 12, 13, 14 and 15.

Table 12 – Proposed demand model, with filtering by median income applied to East Coast routes of the US.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	0.050	-0.019	1.529	-1.511	-	-	0.035
$H_{i,j}$ + $I_{i,j}$	0.074	0.017	0.964	-0.853	0.250	0.098	0.078
$I_{i,j} < 0.4 \times 10^{10}$	0.065	0.231	-5.854	6. 319	0.460	1.680	0.365
$0.4 \times 10^{10} < I_{i,j} < 0.7 \times 10^{10}$	0.076	-0.049	2.422	-2.696	-0.019	-0.228	0.104
$I_{i,j} > 0.7 \times 10^{10}$	14.982	-2.578	-30.220	-28.329	17.353	-11.855	0.986

Table 13 – Proposed demand model, with filtering by median income applied to the Central routes of the US.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	-0.003	-0.034	1.727	-1.772	-	-	0.010
$H_{i,j}$ + $I_{i,j}$	2.850×10^{-5}	-0.035	2.544	-2.605	-0.183	-0.189	0.023
$I_{i,j} < 2.7 \times 10^9$	-0.052	-0.070	-2.860	-2.825	-0.375	-0.206	0.150
$2.7 \times 10^9 < I_{i,j} < 3.8 \times 10^9$	0.053	-0.070	2.158	-2.240	-0.423	-0.164	0.076
$I_{i,j} > 3.8 \times 10^9$	-0.097	0.171	-7.595	8.295	2.811	2.101	0.201

Table 14 – Proposed demand model, with filtering by median income applied to the Mountain routes of the US.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	0.103	0.006	1.299	-1.316	-	-	0.057
$H_{i,j}$ + $I_{i,j}$	0.116	0.005	-0.291	0.267	-0.249	0.362	0.079
$I_{i,j} < 4.0 \times 10^9$	0.117	-0.042	-6.094	6.018	-0.288	1.697	0.145
$4.0 \times 10^9 < I_{i,j} < 5.45 \times 10^9$	0.080	0.044	-4.318	4.554	-0.139	1.328	0.097
$I_{i,j} > 5.45 \times 10^9$	0.150	0.036	-5.212	5.189	-0.780	1.443	0.381

Table 15 – Proposed demand model, with filtering by median income applied to the Pacific Coast routes of the US.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	0.098	-0.062	1.378	-1.146	-	-	0.110
$H_{i,j}$ + $I_{i,j}$	0.101	-0.067	0.258	1.173	-0.887	0.023	0.147
$I_{i,j} < 0.65 \times 10^{10}$	0.100	-0.018	-1.372	1.681	0.413	0.610	0.178
$0.65 \times 10^{10} < I_{i,j} < 0.9 \times 10^{10}$	0.183	-0.167	-6.815	7.363	1.187	1.699	0.277
$I_{i,j} > 0.9 \times 10^{10}$	-0.706	-0.266	34.869	-31.152	6.993	-6.752	0.885

Applying the demand model when breaking the US market down from coast to coast generally improves the model's performance relative to the entire US network shown in Table 11. The improvement in the model performance further hints towards the hypothesis that different regions and states within the US market are affected by different sets of variables.

Moreover, the model performance improves further when the median income is used to filter the data. The greatest improvement seen in the model performance was in the upper median income ranges of both the East and Pacific coasts, with an R^2 of 0.986 and 0.885, respectively, passing the criteria of an R^2 value greater than 0.5. In contrast, the landlocked Central and Mountain regions behaved similarly in each median income range, with no model version passing the target R^2. Observing that the model behaves similarly in coastal and landlock regions further confirms that bodies of water play a factor in demand mentioned in Section 3.2

4.4.2 Inter-State Level Breakdown

Results from Section 4.4.1 suggest that rather than the whole US market working together, behaving the same, different states have different relations in the market. Two test cases within the US network of neighbouring states determine whether inter-state relations behave differently in different regions. Figure 9 presents routes from California, Nevada, Oregon and Washington, and Figure 10 displays routes from Michigan and Wisconsin. Their respective results are in Table 16 and Table 17.

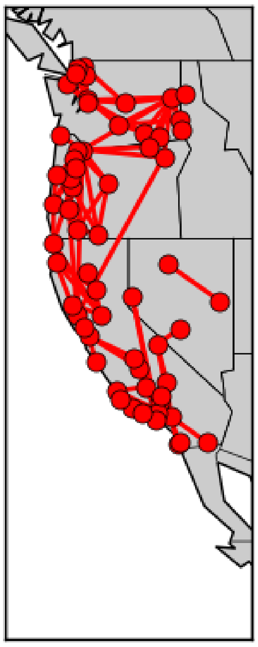


Figure 9 – Regional air cargo routes originating from California, Nevada, Oregon and Washington.

Table 16 – Proposed demand model, with filtering by median income applied to California, Nevada, Oregon and Washington routes.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	0.083	-0.033	1.448	-1.186	-	-	0.088
$H_{i,j} + I_{i,j}$	0.082	-0.035	1.210	-0.931	0.259	0.0461	0.097
$I_{i,j} < 0.65 \times 10^{10}$	0.088	0.022	-4.195	4.457	0.048	1.277	0.177
$0.65 \times 10^{10} < I_{i,j} < 0.9 \times 10^{10}$	0.017	-0.106	0.826	-0.327	0.512	0.195	0.151
$I_{i,j} > 0.9 \times 10^{10}$	-0.706	-0.266	34.869	-31.152	6.993	-6.752	0.885

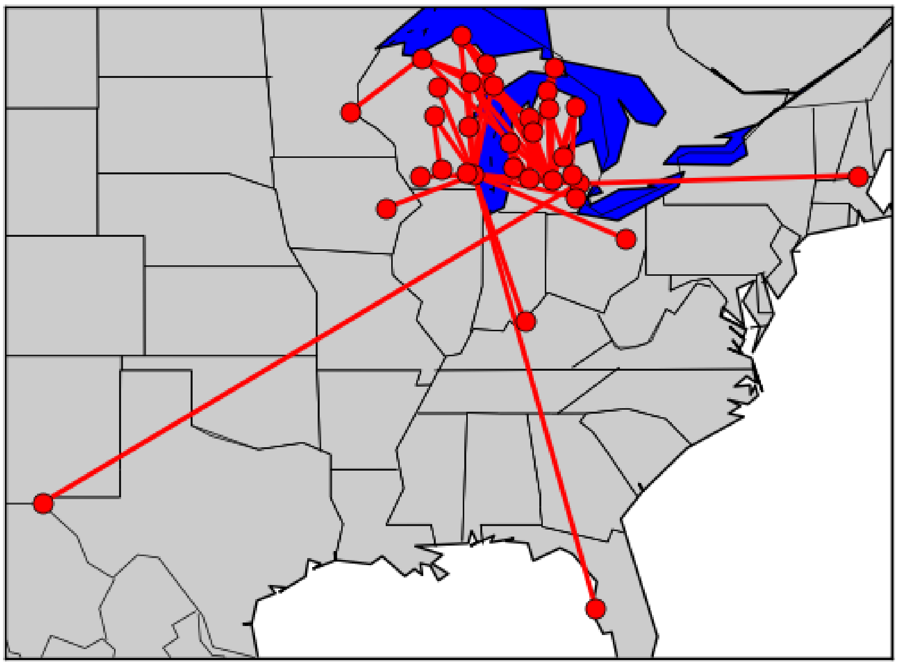


Figure 10 – Regional air cargo routes originating from Wisconsin and Michigan.

Table 17 – Proposed demand model, with filtering by median income applied to Wisconsin and Michigan routes.

	K^1	K^2	K^3	K^4	K^5	K^6	R^2
Baseline	-0.029	-0.123	1.849	-1.731	-	-	0.123
$H_{i,j} + I_{i,j}$	-0.084	-0.079	1.750	-1.520	-0.502	0.087	0.237
$I_{i,j} < 2.8 \times 10^9$	-0.081	-0.077	2.867	-2.594	-0.486	-0.180	0.195
$2.8 \times 10^9 < I_{i,j} < 3.6 \times 10^9$	-0.296	-0.316	-135.810	138.847	-1.284	32.114	0.943
$I_{i,j} > 3.6 \times 10^9$	-0.274	0.108	-117.909	120.886	5.020	2.747	0.989

Table 16 shows that the improved model does not achieve the acceptance criteria of an R^2 greater than 0.5 in all but high median income. Model performance in Wisconsin and Michigan, shown in Table 17, display results better than in California, Nevada, Oregon and Washington, with middle and high median income achieving the test criteria with R^2 values of 0.943 and 0.989, respectively. Results of the two inter-state test cases indicate that although improvements are made to the model's performance when observing routes originating from neighbouring states, the variability in prediction levels demonstrates that additional factors are not considered in the proposed model that are key contributors to the US route network.

4.4.3 Additional Factors

As mentioned in Section 4.3 implementing factors inferred from US network observations could potentially increase the proposed demand model performance within the market. Coastal routes in the US span close to the coastal line and are short distances compared to routes near the centre of the US expanded upon in Section 3.2 Moreover, hub airports near the coast cannot span radially like landlocked hubs. As in passenger demand, incorporating a factor considering an airport's distance to a body of water (W) into the proposed cargo demand model could show performance improvements. Moreover, due to varying sales taxes within each state, an individual's purchase and operation cost motivation could influence route patterns between states. A sales tax factor (S) could increase the model's fidelity. Finally, the impact of regulations from state to state varies. With 50 states in the US, the effect of these regulations can influence the ability of air cargo to flow through different states. The Federal Regulations and State Enterprise Index (FRASE, F) measures the impact of federal regulation on each state [19]. The relation between different state FRASE indices could present an opportunity for inter-country politics to have a role in route network distributions within the US market. Table 18 below summarizes the additional proposed factors to the model, and Table 19 summarizes the respective results on the US route network.

Table 18 – Proposed variables for US network.

Variable	Description
$S_{i,j} = S_i + S_j$	State Sales Tax
$B_{i,j} = B_i + B_j$	Distance to Nearest Body of Water (Geat Lake or Ocean) [nmi]
$F_{i,j} = F_i * F_j$	State FRASE Index

The revised proposed model for the US is shown in Equation 5 below.

$$V_{i,j} = P_{i,j}^{k_1} C_{i,j}^{k_2} D_{i,j}^{k_3} T_{i,j}^{k_4} H_{i,j}^{k_5} I_{i,j}^{k_6} S_{i,j}^{k_7} B_{i,j}^{k_8} F_{i,j}^{k_9} \tag{5}$$

Table 19 – Summarized results of revised proposed demand model for the US regional air cargo network.

	K^1	K^2	K^3	K^4	K^5	K^6	K^7	K^8	K^9	R^2
Baseline	0.070	-0.028	1.443	-1.452	-	-	-	-	-	0.034
$H_{i,j}$ + $I_{i,j}$	0.065	-0.024	0.258	-0.237	-0.030	0.275	-	-	-	0.049
+ $S_{i,j}$	0.073	-0.026	0.206	-0.153	0.064	0.288	-0.102	-	-	0.064
+ $B_{i,j}$	0.063	-0.024	0.446	-0.407	0.074	4.763	-	0.241	-	0.052
+ $F_{i,j}$	0.065	-0.024	0.159	-0.143	0.038	0.297	-	-	0.111	0.050
$S_{i,j}$ + $B_{i,j}$ + $F_{i,j}$	0.078	-0.029	0.209	-0.143	0.092	0.295	-0.114	-0.057	0.260	0.073
$I_{i,j} < 0.65 \times 10^{10}$	0.081	-0.019	-1.103	1.178	0.099	0.591	-0.112	-0.045	0.204	0.092
$I_{i,j} > 0.65 \times 10^{10}$	0.016	-0.033	4.385	-4.155	0.081	-0.581	0.122	-0.152	0.213	0.103

Table 20 – Revised demand model, with filtering by median income applied to East Coast routes of the US.

	K^1	K^2	K^3	K^4	K^5	K^6	K^7	K^8	K^9	R^2
Baseline	0.050	-0.019	1.529	-1.511	-	-	-	-	-	0.035
$H_{i,j}$ + $I_{i,j}$	0.074	0.017	0.964	-0.853	0.250	0.09	-	-	-	0.078
$S_{i,j}$ + $B_{i,j}$ + $F_{i,j}$	0.091	0.035	2.647	-2.357	0.398	-0.212	-0.848	-0.026	-0.589	0.124
$I_{i,j} < 0.4 \times 10^{10}$	0.131	0.109	-1.361	1.776	0.461	0.639	-0.896	0.157	-1.585	0.469
$0.4 \times 10^{10} < I_{i,j} < 0.6 \times 10^{10}$	-0.040	-0.140	27.389	-27.792	-0.161	-5.378	-3.968	0.310	0.448	0.492
$I_{i,j} > 0.6 \times 10^{10}$	0.559	0.186	60.155	-64.219	3.947	-14.260	5.763	1.581	27.350	0.475

Table 21 – Revised demand model, with filtering by median income applied to the Central routes of the US.

	K^1	K^2	K^3	K^4	K^5	K^6	K^7	K^8	K^9	R^2
Baseline	-0.003	-0.034	1.727	-1.772	-	-	-	-	-	0.010
$H_{i,j}$ + $I_{i,j}$	2.850×10^{-5}	-0.035	2.544	-2.605	-0.183	-0.189	-	-	-	0.023
$S_{i,j}$ + $B_{i,j}$ + $F_{i,j}$	0.009	-0.041	3.247	-3.359	-0.106	-0.341	0.081	-0.057	-0.252	0.027
$I_{i,j} < 2.7\times10^9$	[-0.015	-0.070	3.918	-4. 006	-0.304	-0.406	-0.855	0.076	-0.157	0.175
$2.7\times10^9 < I_{i,j} < 3.8\times10^9$	-0.187	-0.755	103.696	-108. 990	-3.376	-24.123	-1.926	3.862	-27.785	0.256
$I_{i,j} > 3.8\times10^9$	-0.096	0.0545	-9.537	10.050	1.658	2.685	-0.148	-0.376	3.532	0.325

Table 22 – Revised demand model, with filtering by median income applied to the Mountain routes of the US.

	K^1	K^2	K^3	K^4	K^5	K^6	K^7	K^8	K^9	R^2
Baseline	0.103	0.006	1.299	-1.316	-	-	0.057-	-	-	
$H_{i,j}$ + $I_{i,j}$	0.116	0.005	-0.291	0.267	-0.249	0.362	-	-	-	0.079
$S_{i,j}$ + $B_{i,j}$ + $F_{i,j}$	0.167	-0.012	-0.933	0.851	-0.137	0.506	-0.158	-0.158	0.593	0.171
$I_{i,j} < 4.0\times10^9$	0.179	-0.029	-8.065	7.872	-0.408	2.131	0.100	-0.228	1.067	0.210
$4.0\times10^9 < I_{i,j} < 5.45\times10^9$	0.174	0.0319	-2.290	2.453	-0.215	0.774	-0.162	0.025	0.413	0.223
$I_{i,j} > 5.45\times10^9$	0.172	0.028	-3.590	3.450	-0.826	1.059	-0.236	0.065	0.527	0.526

Table 23 – Revised demand model, with filtering by median income applied to the Pacific Coast routes of the US.

	K^1	K^2	K^3	K^4	K^5	K^6	K^7	K^8	K^9	R^2
Baseline	0.098	-0.062	1.378	-1.146	-	-	-	-	-	0.110
$H_{i,j}$ + $I_{i,j}$	0.101	-0.067	0.258	1.173	-0.887	0.023	-	-	-	0.147
$S_{i,j}$ + $B_{i,j}$ + $F_{i,j}$	0.123	-0.064	2.152	-1.724	0.507	-0.182	-0.035	-0.152	0.142	0.185
$I_{i,j} < 0.65\times10^{10}$	0.126	-0.018	-0.975	1.483	0.557	0.552	-0.109	-0.192	1.034	0.224
$0.65\times10^{10} < I_{i,j} < 0.9\times10^{10}$	0.318	-0.158	-0.282	1.197	1.331	0.164	-0.039	-0.369	-0.786	0.344
$I_{i,j} > 0.9\times10^{10}$	-0.658	-0.183	28. 098	-26.433	9.393	-5.507	0.823	0.0318	5.140	1.000

Table 24 – Revised demand model, with filtering by median income applied to California, Nevada, Oregon and Washington routes.

	K^1	K^2	K^3	K^4	K^5	K^6	K^7	K^8	K^9	R^2
Baseline	0.083	-0.033	1.448	-1.186	-	-	-	-	-	0.088
$H_{i,j}$ + $I_{i,j}$	0.082	-0.035	1.210	-0.931	0.259	0.0461	-	-	-	0.097
$S_{i,j}$ + $B_{i,j}$ + $F_{i,j}$	0.143	-0.022	1.974	-1.484	0.285	-0.141	-0.079	-0.200	0.215	0.176
$I_{i,j} < 0.65\times10^{10}$	0.161	0.038	-4.870	5.500	0.155	1.452	-0.348	-0.127	3.786	0.306
$0.65\times10^{10} < I_{i,j} < 0.9\times10^{10}$	0.185	-0.120	6.017	-5.123	0.722	-1.077	-0.070	-0.367	-0.525	0.326
$I_{i,j} > 0.9\times10^{10}$	-0.658	-0.183	28.098	-26.433	9.393	-5.507	0.823	0.032	5.140	1.000

Table 25 – Revised demand model, with filtering by median income applied to Wisconsin and Michigan routes.

	K^1	K^2	K^3	K^4	K^5	K^6	K^7	K^8	K^9	R^2
Baseline	-0.029	-0.123	1.849	-1.731	-	-	0.123			
$H_{i,j}$ + $I_{i,j}$	-0.084	-0.079	1.750	-1.520	-0.502	0.087	-	-	-	0.237
$S_{i,j}$ + $B_{i,j}$ + $F_{i,j}$	-0.137	-0.078	1.374	-1.277	-0.535	0.048	1.225	0.216	-0.139	0.313
$I_{i,j} < 2.8\times10^9$	-0.126	-0.157	8.461	-8.580	-0.576	-1.180	-2.405	0.225	7.790	0.313
$2.8\times10^9 < I_{i,j} < 3.5\times10^9$	-0.194	-0.456	-54.740	55.266	-0.595	12.590	4.098	0.862	-4.271	1.000
$I_{i,j} > 3.5\times10^9$	-0.380	0.231	-23.854	25.267	2.046	6.176	-4.743	-0.939	10.479	1.000

Incorporating distance to the nearest body of water, sales tax factor, and FRASE index in the proposed demand model increased R^2 49% to 0.073. Although the additional factors improved the correlation with the US market, it does not surpass the threshold of an R^2 greater than 0.5. Increasing correlation trends continue when filtering the network by median income and applying the model to certain areas of the US, but only the upper median income range of Central US routes is improved to surpass the threshold R^2 relative to the initially proposed demand model. While the factors initially proposed and discussed in this section capture a portion of the US regional route network behaviour, commercial and industrial output in the US might be equally important compared to consumer-driven air cargo transport, in contrast to the Canadian market. Considering commercial and industrial output within the proposed demand model requires additional analysis of factors such as local GDP and labour productivity rates.

5. Conclusion

This paper presents a methodology for determining air cargo networks from air flight data. Furthermore, this work proposes a model capable of predicting air cargo demand for a given route network. The regional air cargo route networks of Canada and the United States were chosen to evaluate the model's performance in a comparable, lower-population market, like Canada, before applying the model to a larger-scale market like the United States. The model's objective was to generate a demand model capable of being applied to multi-hub regional air cargo networks. The proposed model is built upon the gravity model typically used in passenger demand, adjusting for geographic and socio-economic parameters affecting cargo demand. These parameters in the proposed model include a city pair's summation of distances to the nearest highway and product median income. Applying the developed model to the Canadian market achieved a coefficient of determination larger than 0.88, relating to a high level of prediction accuracy. The United States test case achieved a lower prediction success where local variables contributed to playing more significant roles. Additional geographical and consumer-economic factors, which include local sales tax, distance to a body of water, and federal regulations and enterprise index, improved the proposed model's predictability in the United States market but mainly did not achieve an adequate coefficient of determination, excluding higher median income ranges. Future work building upon the proposed demand model will examine additional commercial economic and productivity indicators to improve the correlation with the United States regional air cargo network.

6. Contact Author(s) Email Address(es)

mailto: martin.alexiev@rmc.ca, ruben.perez@rmc.ca, peter.jansen@rmc.ca, jeremy.wang@flyribbit.com

7. Acknowledgements

The authors wish to acknowledge and thank Grace Ryde, Harnoor Singh Ahluwalia, Caleb Graham, and Elizabeth Shannon from the Advanced Aircraft Design Lab at the Royal Military College of Canada for their support in gathering and aggregating the flight and airport data used in this paper.

8. Copyright Statement

References

[1] Davendralingam, N. and Crossley, W., "Concurrent Aircraft Design and Airline Network Design Incorporating Passenger Demand Models," *9th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA, Hilton Head, South Carolina, September 2009.

[2] Davendralingam, N., Crossley, W., and Nusawardhana, A., "An Approximate Dynamic Programming Approach to Aircraft Allocation Incorporating Passenger Demand Models," *10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, AIAA, Fort Worth, Texas, September 2010.

[3] Peter W. Jansen, R. E. P., "Coupled Optimization of Aircraft Design and Fleet Allocation with Uncertain Passenger Demand," *2013 Aviation Technology, Integration, and Operations Conference*, AIAA, Los Angeles, California, August 2013.

[4] Mane, M. and Crossley, W., "Allocation and Design of Aircraft for On-Demand Air Transportation with Uncertain Operations," *Journal of Aircraft*, Vol. 49, No. 1, January–February 2012, pp. 141 – 150.

[5] Jordan Sakakeeny, Nadazhda Dimitrova, H. I., "Preliminary Characterization of Unmanned Air Cargo Routes Using Current Cargo Operations Survey," *AIAA AVIATION 2022 Forum*, AIAA, Chicago, Illinois, June 2022.

[6] (LMI), L. M. I., "LMI Automated Air Cargo Operations Market Research and Forecast," Tech. rep., NASA, April 2021.

[7] Taylor, C. and de Weck, O., "Coupled Vehicle Design and Network Flow Optimization for Air Transportation Systems," *Journal of Aircraft*, Vol. 44, No. 5, September–October 2007, pp. 1478–1486.

[8] Crossley, W. and Mane, M., "System of Systems Inspired Aircraft Sizing Applied to Commercial Aircraft / Airline Problems," *5th Aviation, Technology, Integration, and Operations Conference (ATIO)*, AIAA, Arlington, Virginia, September 2005.

[9] Mane, M. and Willcox, K., "Importance of Aircraft Type and Operational Factors for Air Taxi Cost Feasibility," *Journal of Aircraft*, Vol. 46, No. 4, July–August 2009, pp. 1222–1230.

[10] Peter W. Jansen, R. E. P., "Robust Coupled Optimization of Aircraft Design and Fleet Allocation for Multiple Markets," *AIAA/3AF Aircraft Noise and Emissions Reduction Symposium*, AIAA, Atlanta, Georgia, June 2014.

[11] Peter W. Jansen, R. E. P., "Integrated Design and Optimization of Aircraft Families and Air Transport Network," *16th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA, Washington, D.C., June 2016.

[12] Peter W. Jansen, R. E. P., "Effect of Passenger Preferences on the Integrated Design and Optimization of Aircraft Families and Air Transport Network," *17th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA, Denver, Colorado, June 2017.

[13] Andreas M. Tillmann, Imke Joormann, S. C. A., "Reproducible air passenger demand estimation," *Journal of Air Transport Management*, Vol. 112, September 2023.

[14] Tobias Grosche, Franz Rothlauf, A. H., "Gravity models for airline passenger volume estimation," *Journal of Air Transport Management*, Vol. 13, No. 4, July 2007, pp. 175 – 183.

[15] Cherng-Chwan Hwang, G.-C. S., "Analyzing air cargo flows of international routes: an empirical study of Taiwan Taoyuan International Airport," *Journal of Transport Geography*, Vol. 19, No. 4, July 2011, pp. 738 – 744.

[16] Alessandro Bombelli, Bruno M. Santos, L. T., "Analysis of the air cargo transport network using a complex network theory perspective," *Transportation Research Part E: Logistics and Transportation Review*, 2020.

[17] Turner, S. A., "An Application of a Gravity Model to Air Cargo at Vancouver International Airport," Tech. rep., The University of British Columbia, 2001.

[18] J. A. Hartigan, M. A. W., "Algorithm AS 136: A K-Means Clustering Algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 28, No. 1, 1979, pp. 100–108.

[19] Patrick A. McLaughlin, O. S., "The Impact of Federal Regulation on the 50 States," Tech. rep., Mercatus Center, George Mason University, 2016.