# Using machine learning in order to estimate the traffic mix in a building from the stops data

Lutfi Al-Sharif<sup>1,2</sup>, Richard Peters<sup>2</sup>, Matthew Appleby<sup>2</sup>, Tahani Ghaben<sup>1</sup>

<sup>1</sup>Al Hussein Technical University, Amman, Jordan <sup>2</sup>Peters Research Ltd., United Kingdom

**Keywords**: The Round Trip; elevator; lift; incoming traffic; outgoing traffic; interfloor traffic; occupant floor; entrance floor, machine learning.

**Abstract.** Previous work has established that the average number of up stops and down stops in a building during a round trip, as well as the ratio between them, could be used to estimate the mix of traffic prevailing in the building and its intensity.

Further work has used basic correlation methods to derive the mix of traffic in the building, finding the ratios of incoming traffic, outgoing traffic, and inter-floor traffic. These studies have assumed that inter-entrance traffic is zero.

This paper builds on the methodologies developed in the earlier work by introducing machine learning techniques to model the relationship between stop types and their locations within a building. The methodology requires knowledge of the types of floors in the building (occupant floors or entrance/exit floors).

The data required for machine learning will be generated in larger amounts in a reasonable time and with modest processing power, whereby the data is representative of a specific building.

#### 1 INTRODUCTION

This paper is a continuation of a series of pieces of research that attempt to infer the type of traffic in a building based on the number and type of stops that the lift makes. The following is a historical summary:

- As a simple example, the ratio of the up stops to the down stops can be used to detect the presence of an up-peak condition [1].
- A follow-up piece of work used artificial neural networks to assess the type of prevailing traffic in the building, such as up peak, down peak, two-way traffic or off peak [2].
- The concept of attempting to infer traffic data based on the number of stops was first suggested in 1992 [3] in what was then called the I-S-P method.
- Further work on this concept was carried out in 2022 [4], whereby the tables relating the expected number of stops in the up direction and the down direction were generated based on large amounts of data generated using the Monte Carlo Simulation (MCS) method.
- Basic single-dimensional regression was applied in special limited traffic cases (e.g., where incoming traffic and outgoing traffic are equal) to try to determine simple ways that can be used to find the mix of traffic based on the number of stops [5].

This paper relies heavily on the definition of a floor in a building as either an entrance/exit floor (to be referred to later as an entrance floor) or an occupant floor [6]. It further extends the work in [5] by systematically using two-dimensional regression to find specific equations that allow the user to find the incoming traffic and the outgoing traffic percentages from the four types of stops in a round trip.

Section 2 generates the data for the number of stops in a round trip based on the mix of traffic. This results in a triangular data table, ready to be used in the machine learning process. Section 3 plots

four surfaces that allow the reader to visualise the relationship that is contained in the triangular table, against the x and y axes (incoming traffic percentage and outgoing traffic percentage, respectively). Section 4 carries out the two-dimensional regression and presents the two resultant equations. Section 5 draws conclusions and discusses the next important step that will allow this methodology to become universal and allow it to be used for any size and arrangement of building.

## 2 GENERATING THE TRIANGULAR STOPS DATA

In order to implement the machine learning process, it is first necessary to generate sufficient data. The data will link the different scenarios of the traffic mix with the average number of stops in any round trip.

The ideal method for generating such data is to employ the Monte Carlo simulation (MCS) method.

In the previous piece of research [5], the approach of generating data points at different mixes of traffic was selective, rather than systematic. To understand the data and be able to come up with a suitable approach, two sets of scenarios were generated:

- The first set of scenarios attempts to keep the ratio between the incoming traffic and the outgoing traffic constant, while varying the percentage of the inter-floor traffic.
- The second set of scenarios keeps the percentage of the inter-floor traffic constant at 40%, while varying the ratio of the incoming traffic to the outgoing traffic.

In all cases, the inter-entrance traffic is kept at 0%.

Such an approach was very limited and only concentrated on a specific number of cases. This paper adopts a more systematic approach by continuously varying the incoming traffic and the outgoing traffic in increments of 10% (or 0.1) from 0% to 100% (i.e., from 0 to 1.0). When the incoming traffic attains values of 60% or more, the value of the outgoing traffic must be restricted (to prevent the total traffic values exceeding 100%). Such a restriction means that a table that has column heading of outgoing traffic ranging from 0% to 100%, and row headings of incoming traffic from 0% to 100%, can only be a "triangular" table, as the lower part of the triangular would have values the sum of which exceeds 100%, and thus must be disallowed. A resolution smaller than the value of 10% is not necessary for this exercise, as the resultant surface (which will be seen in the next section) is very smooth.

An overview of the resultant data table is shown in Figure 1 below. The figure aims to show the shape of the table (as the numbers in it are too small to read!).

incomi																outgoing tra	affic %																
mg		0			10			20			30			40			50			60			70			80			90			100	
traffic	interfloor %			interfloor %			interfloor%			interfloor%			interfloor %			interfloor%			interfloor %			interfloor%			interfloor%			interfloor%			interfloor %		
	100	up stops	down stops	90	up stops	down stops	80	up stops	down stops	70	up stops	down stops	60	up stops	down stops	50	up stops	down stops	40	up stops	down stops	30	up stops	down stops	20	up stops	down stops	10	up stops	down stops	0	up stops d	own stops
0	entrance floors	0.000	0.000		0.0000	0.9367		0.0000	1.4119		0.0000	1.6601		0.0000	1.7965		0.0000	1.8756		0.0000	1.9235		0	1.9534		0	1.9717		0	1.9835		0	1.9904
	occupant floors	6.590	6.591		6.3068	6.5885		5.9684	6.5891		5.5678	6.5886		5.0940	6.5909		4.5438	6.5876		3.8915	6.5895		3.1349	6.5899		2.2423	6.5897		1.2083	6.5895		0	6.5896
	90			80			70			60			50			40			30			20			10			0					
10	entrance floors		0.0000		0.9378	0.9373		0.9371	1.4126		0.9386	1.6603		0.9371			0.9368	1.8755		0.9373			0.9377			0.9377	1.9714		0.9377	1.9834			
	occupant floors	6.5909	6.3043		6.3059	6.3025		5.9653	6,3036		5.5669	6.3052		5.096	6.3031		4.5433	6,3048		3.8939	6.3054		3.1337	6,3055		2.2447	6.3027		1.2065	6.3058			
	80			70			60			50			40			30			20			10			0								
20	entrance floors				1.4116	0.9376		1.4107	1.4127		1.4109	1.6602		1.4115			1.4121	1.8757		1.4114			1.4126			1.4108							200
	occupant floors	6.5888	5.9663		6.3044	5.9675		5.9673	5.9657		5.5665	5.9665		5.0975	5.9673		4.544	5.9673		3.8907	5.9671		3.1337	5.9671		2.2429	5.9666					1000	
	70			60			50			40			30			20			10			0											
30	entrance floors				1.6602	0.9371		1.6605	1.4111			1.65929		1.6602			1.6594	1.8754		1.6598			1.6596								0.000		0.0.0.0
	occupant floors	6.5902	5.5662		6.3067	5.5647		5.9664	5.567		5.56582	5.56462		5.0976	5.5648		4.5411	5.5691		3.8952	5.5673		3.1313	5.5667									
	60			50			40			30			20			10			0			2.000				2000							
40	entrance floors		0.0000		1.7955	0.9363		1.7967	1.4117		1.7966	1.6592		1.79595	1.79553		1.7964	1.8755		1.796			101111111			1000			0.01			1000	
	occupant floors	6.5899	5.0973		6.3056	5.0955		5.9664	5.0972		5.5671	5.098		5.09899	5.09513		4.5447	5.0958		3.8915	5.0966												
	50			40			30			20			10			0																	
50	entrance floors		0.0000		1.8757	0.9383		1.8756	1.4127		1.8755	1.66		1.875	1.7959		1.8754	1.8760						1000									
	occupant floors	6.5886	4.5415		6.3039	4.5424		5.9673	4.5417		5.5656	4.5444		5.0956	4.5414		4.5414	4.5429															
	40			30			20			10			0			0000				11.00	10000					300			2 2 3			100	
60	entrance floors		0.0000		1.9239	0.9373		1.9237	1.4112		1.9237	1.6595		1.9229																			
	occupant floors	6.5903	3.8939		6.3027	3.8945	10	5.9657	3.8937		5.5659	3.8933		5.096	3.8935																		
70	entrance floors		0.0000	20	1.9533	0.9373	10	1.9531	1.4116	0	1.9534	1.6597																					
	occupant floors			_	6.3041	3.1321		5.9677	3.1321		5.5683	3.1329								-		_											
	occupant moors 20	6.5899	3.1311	10	6.3041	3.1321	0	5.9677	3.1321		0.0683	3.1325																					
80	entrance floors	1.0710	0.0000	10	1.9723	0.9370	- 0	1.9719	1.4118											-													
ou	occupant floors				6.3058	2.2429		5.9660	2.2428																								
	10	9.5090	6.6423	0	0.0000	6.6923		5.9660	4.2420					-						-													-
90	entrance floors	1 9832	0.0000	J	1.9832	0.9385																											
90	occupant floors				6.3042	1.2075														-													
	O Companie 10075	0.0093	4.4400		0.0092	4.60/3																											
100	entrance floors	1 9904	0.0000																														
	occupant floors		0.0000																														
	occupant moors	9.0887	0.0000																		THE RESERVE												-

Figure 1 The triangular data table that contains the stops data against the values of incoming traffic and outgoing traffic.

Figure 2 shows an enlarged view of the table. As can be seen, a highlighted group of cells contains the values of four types of stops (up on entrance floors; up stops on occupant floors; down stops on entrance floors; down stops on occupant floors) under the values of 10% outgoing traffic, 10% incoming traffic, and 80% inter-floor traffic. The values of the stops are absolute here, but they will be normalised later in the process by dividing by the number of entrance floors and the number of occupant floors as appropriate. The values of the stops under this scenario of traffic (i.e., 10%; 10%; 80%; 0%) are: At the entrance floors, 0.9378 up stops per round trip and 0.9373 down stops per round trip; at the occupant floors, 6.3059 up stops per round trip and 6.3025 down stops per round.

entrance floo	Occupant floors									
2	8									
incoming traffic %		0		outgoing traffic % 10						
traffic %	interfloor %			interfloor %						
	100	up stops	down stops	90	up stops	down stops				
0	entrance floors	0.000	0.000		0.0000	0.9367				
	occupant floors	6.590	6.591		6.3068	6.5885				
	90			80						
10	entrance floors	0.9384	0.0000		0.9378	0.9373				
	occupant floors	6.5909	6.3043		6.3059	6.3025				

Figure 2 An enlarged view of the data table, showing stops data in detail against the value of traffic components.

Figure 3 shows part of the reorganised table of data that has been prepared for the machine learning process, with three input variables (two of which are independent) and the four output variables.

Α	В	С	D	E	F	G
I/c	o/g	interfloor %	u/s ent	d/s ent	u/s occ	d/socc
0	0	100	0	0	6.59	6.591
10	0	90	0.9384	0	6.5909	6.3043
20	0	80	1.4107	0	6.5888	5.9663
30	0	70	1.6596	0	6.5902	5.5662
40	0	60	1.7967	0	6.5899	5.0973
50	0	50	1.8756	0	6.5886	4.5415
60	0	40	1.9242	0	6.5903	3.8939
70	0	30	1.9531	0	6.5899	3.1311
80	0	20	1.9719	0	6.5898	2.2425
90	0	10	1.9832	0	6.5893	1.2055
100	0	0	1.9904	0	6.5887	0
0	10	90	0	0.9367	6.3068	6.5885
10	10	80	0.9378	0.9373	6.3059	6.3025
30	10	60	1.6602	0.9371	6.3067	5.5647
40	10	50	1.7955	0.9363	6.3056	5.0955
50	10	40	1.8757	0.9383	6.3039	4.5424
60	10	30	1.9239	0.9373	6.3027	3.8945

Figure 3 Part of the data in the table following reorganisation to prepare the data for machine learning (the matching data shown in Figure 2 is highlighted here in yellow).

As was mentioned earlier, the raw data was generated using the Monte Carlo Simulation method. The MCS method can be used to faithfully generate a large amount of round-trip data [7, 8, 9, 10]. The MCS employs random sampling methods that ensure that the resultant data is a faithful representation of the parameters of the building. The MCS tool was used to generate round trip stops based on the following (similar to what is used in [6]):

- The specific number of passengers per round trip (P=13 passengers in this case).
- A predefined mix of traffic showing the relative strength of the incoming traffic, outgoing traffic, inter-floor traffic and inter-entrance floor.
- The number of occupant floors and the number of entrance/exit floors (8 and 2, respectively, in this case)
- The relative populations of the occupant floors.
- The relative arrivals at the entrance/exit floors.

The MCS methodology does not need to generate the kinematics of the lift system or produce the speed-time profiles. Moreover, it does not need to carry out any allocation of landing calls to the cars. This makes the method independent of the dispatching algorithm. But more importantly, this is an excellent example of the abstraction capability of the MCS method, where it only generates the data that is needed for the case in question, rather than spending time generating kinematic curves or allocation calls. This approach speeds up the generation of the raw data.

### 3 VISUALISING THE FOUR SURFACES

Prior to moving to try to find the best fit for the four surfaces, it is very insightful to better understand the nature of the relationships in a visual manner.

There are three independent variables: the percentage of incoming traffic, the percentage of outgoing traffic, and the percentage of inter-floor traffic. However, only two of these three are independent, as

the sum of all three values must add up to 100% (or to 1.0). Thus, in reality, there are only two independent variables (e.g., incoming traffic percentage and outgoing traffic percentage), as the third value will be dependent on the values of the other two. This is assuming that the inter-entrance traffic percentage is always zero.

In addition, there are four dependent variables. These are the average number of stops in a round trip. These four variables are:

- Stops in the up direction on entrance floors.
- Stops in the down direction on the entrance floors.
- Stops in the up direction on occupant floors.
- Stops in the down direction on occupant floors.

As there are two dependent variables, it is very convenient to plot each of the four variables as a "surface" against the two dependent variables. This results in four "surfaces" for each of the four dependent variables that have been plotted against the two dependent variables. In this case, it has been assumed that the two independent variables are the incoming traffic percentage and the outgoing traffic percentage. So, in the plots that have been shown below, the x-axis and the y-axis are the incoming traffic and the outgoing traffic percentages, respectively.

As can be seen in the four figures (Figure 4, Figure 5, Figure 6, and Figure 7), they all show a "triangular" surface that has been "curved". In two cases, the base of the triangle is at zero, either parallel to the x-axis or parallel to the y-axis, and the upper vertex of the triangle ends up at the other side. In the two other cases, the base of the triangle is at the top, either parallel to the x-axis or parallel to the y-axis, with the vertex near the zero. The limiting number of stops is the car size and the number of occupied floors. Here, the number of stops saturates at about 6.3-6.6 stops per round trip, which is about 78.8%-82.5% of the number of occupied floors.

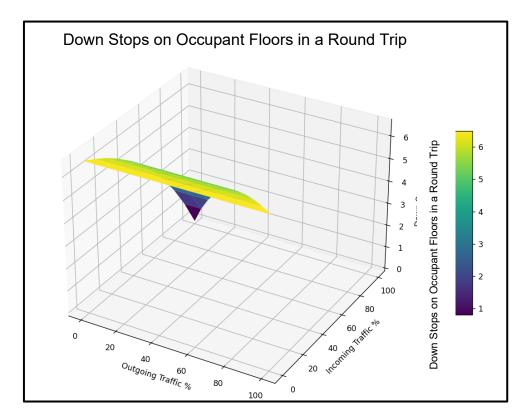


Figure 4 The plot of the surface that shows the number of down stops on the occupant floors during a round trip, plotted against the incoming traffic percentage and the outgoing traffic percentage.

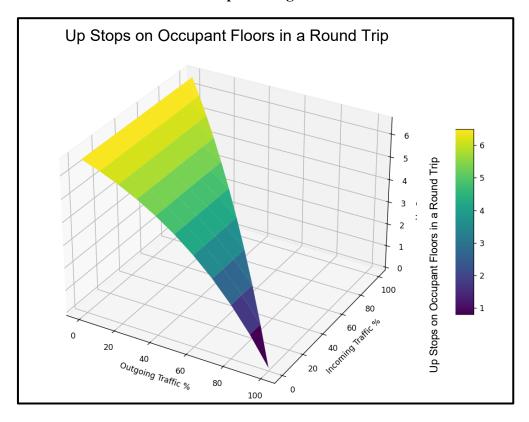


Figure 5 The plot of the surface that shows the number of up stops on the occupant floors during a round trip, plotted against the incoming traffic percentage and the outgoing traffic percentage.

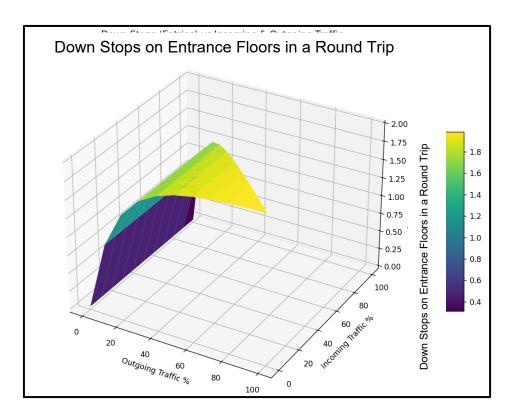


Figure 6 The plot of the surface that shows the number of down stops on the entrance floors during a round trip, plotted against the incoming traffic percentage and the outgoing traffic percentage.

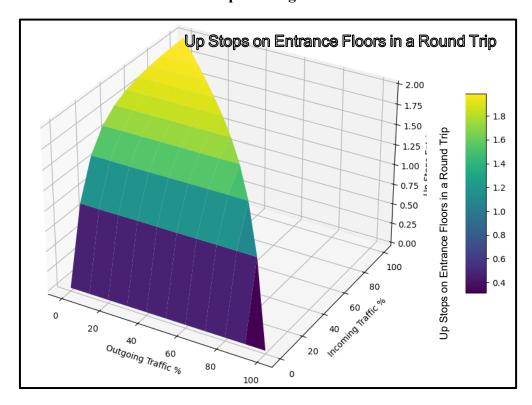


Figure 7 The plot of the surface that shows the number of up stops on the entrance floors during a round trip, plotted against the incoming traffic percentage and the outgoing traffic percentage.

It is also worth noting that the z-axis in all four cases (i.e., the average number of stops in a round trip) has not been normalised and has been shown in absolute number of stops. Using further processing (as will be seen in the next section), the number of stops can be normalised by dividing the maximum possible number of stops in the entrance floors (in this case, 2) and by dividing by the maximum number of stops in the occupant floors (in this case, 8). This will produce a "cube" that is normalised and in which each of the three sides is equal to 1 (or 100%).

### 4 TWO-DIMENSIONAL REGRESSION

Using MATLAB, two-dimensional regression tools were used to find the best fit equation that allows the user to find the percentage of the incoming traffic (i/c%), the percentage of outgoing traffic (o/g%), and the percentage of the inter-floor traffic (i/f%) based on the four types of stops that take place on average during a round trip. The equations below were developed using two-dimensional regression tools. It is worth noting that these equations use the normalised values of the stops data, by dividing the average number of stops in a round trip by the maximum number of possible stops. The following is a list of the variables used in the equations below:

- i/c: incoming traffic (number between 0 and 1).
- o/g: outgoing traffic (number between 0 and 1).
- i/f: inter-floor traffic (number between 0 and 1).
- u/s (ent): The average number of up stops on the entrance floors in a round trip, normalised by dividing it by the number of entrance floors in the building.
- u/s (occ): The average number of up stops on the occupant floors in a round trip, normalised by dividing it by the number of occupant floors in the building.
- d/s (ent): The average number of down stops on the entrance floors in a round trip, normalised by dividing it by the number of entrance floors in the building.
- d/s (occ): The average number of down stops on the occupant floors in a round trip, normalised by dividing it by the number of occupant floors in the building.

Having done the regression, these are the following two equations that resulted. The first is the resultant regression equation to find the value of the incoming traffic:

$$i/c = 85.1098 + (22.0421 * u/s ent (norm)) + (1.6454 * d/s ent (norm)) + (3.4023 * u/s occ (norm)) + (-109.0788 * d/s occ (norm))$$
 (1)

The second equation is the resultant regression equation to find the value of the outgoing traffic:

$$o/g = 85.0430 + (1.8327 * u/s ent (norm)) + (21.8450 * d/s ent (norm)) + (-109.2628 * u/s occ (norm)) + (3.7193 * d/s occ (norm))$$
 (2)

The correlation coefficient from the two regression analysis results is very high and nearly close to one, showing an excellent correlation between the number of stops and the mix of traffic parameters:

i/c equation: R<sup>2</sup> Score: 0.9908616713660177 o/g equation: R<sup>2</sup> Score: 0.9908560432436128

As the inter-entrance traffic has been assumed to be equal to 0, then the inter-floor traffic can be found by subtracting the incoming traffic and the outgoing traffic from 1, as shown below:

$$i/f = 1 - i/c - o/g$$
 (3)

For example, if we calculate that incoming traffic is equal to 10% and outgoing traffic is equal to 30%, then the interfloor traffic would be equal to 1-0.1-0.3 = 0.6 = 60%.

It is worth noting that this methodology has assumed the inter-entrance traffic is always zero for simplicity. Future work will remove this constraint and allow the inter-entrance traffic to take on non-zero values. However, in such a case, the number of independent variables will rise to three (e.g., incoming traffic, outgoing traffic, and inter-floor traffic that do not add to one, as the inter-entrance traffic will not be equal to zero in this case). As such, it will not be possible to visualise the resultant "hyper-surface", because it will have four dimensions. Nevertheless, the fact that it cannot be visualised does not prevent finding the equation that links the mix of traffic to the value of the four types of stops. In such a case, three equations will be found for the incoming traffic, the outgoing traffic and the inter-floor traffic. The inter-entrance traffic will be found by simply subtracting the three resultant values of traffic from 1.

### 5 CONCLUSIONS AND FURTHER WORK

This paper has presented a systematic methodology to attempt to find a deterministic relationship between the mix of traffic in a building (i.e., incoming traffic, outgoing traffic, and inter-floor traffic), and the number of stops on entrance/exit floors in the up direction, the number of stops on entrance/exit floors in the down direction, the number of stops on occupant floors in the up direction, and the number of stops on occupant floors in the down direction. Before processing, the number of stops in each round trip was normalised by dividing the number of stops in a round trip by the maximum possible number of stops. The maximum possible number of stops is the number of occupants, floors and/or the number of entrance floors.

Data was first generated using MCS for a specific building. For each of the 66 different combinations of mixes of traffic, one million scenarios were generated, resulting in 66 million scenarios. Each scenario represents one round trip. In each case, the number of stops in the up direction and the down direction, on both the occupant floors and the entrance floors, was recorded. The average number of stops for each type of stop was found by taking the average of all one million scenarios.

This resulted in a "triangular" data table. This table was then used to perform a machine learning exercise, to find the two equations that allow the user to find the values of incoming traffic and outgoing traffic from the four types of stops data. A very high value of correlation coefficients was obtained in both cases, giving a high level of confidence in the accuracy of the resultant equations.

The normalisation of the stops data was critical in attempting to find a universal solution. The most important piece of research that must follow this research is to attempt to carry out the same exercise that has been done in this paper on other types of buildings. If the resultant equations for the incoming traffic and outgoing traffic are identical to those found in this paper, this leads to the important conclusion that these equations are universal. This would mean that these equations can be used for any building without the need to do prior analysis on it. This would be an important breakthrough in lift traffic engineering.

## **REFERENCES**

- [1] J.R. Beebe. *Lift Management*. Ph.D. Thesis, The University of Manchester, 1980, pp 114-116.
- [2] ATP So, JR Beebe, WL Chan, SK Liu. An Artificial Neural-Network based Traffic Patterns Recognition System. *The International Journal of Elevator Engineering*, volume 1, October 1996.
- [3] Al-Sharif, L. *Predictive Methods in Lift Traffic Analysis*. The University of Manchester, 1992.
- [4] L.R. Al-Sharif, R.D.Peters, M.Appleby. Enhancing the ISP Method (Inverse Stops-Passengers) Using the Monte Carlo Simulation Method. 13th Symposium on Lift & Escalator Technologies 2022. 13 (1), 7-14
- [5] Al-Sharif L., Peters R., Appleby M., Ghaben T. Estimating the Traffic Mix in a Building by Analysing the Stops Data per Round Trip. Proceedings of the 15<sup>th</sup> Symposium on Lift and Escalator Technologies. 2024. 15: 7-14.
- [6] Al-Sharif L, Abu Alqumsan AM. An integrated framework for elevator traffic design under general traffic conditions using origin destination matrices, virtual interval, and the Monte Carlo simulation method. Building Services Engineering Research and Technology. 2015;36(6):728-750. doi:10.1177/0143624415595521.
- [7] Appleby M and Peters R D. The Round-Trip Time Simulation: Monte Carlo Implementation and Consistency with Other Techniques. *Proceedings of the 11<sup>th</sup> Symposium on Lift and Escalator Technologies (online)*. 2020.
- [8] Al-Sharif L. A Universal Methodology for Generating Elevator Passenger Origin-Destination Pairs for Calculation and Simulation. In the Book (chapter 6): Smart Cities-Their Framework and Applications, published by IntechOpen, 25th September 2020.
- [9] So A, Al-Sharif L, Hammoudeh A. Concept design and derivation of the round trip time for a general two-dimensional elevator traffic system. Journal of Building Engineering. 5: 165-177.
- [10] So A, Al-Sharif L, Hammoudeh A. Analysis of Possible Two Dimensional Elevator Traffic Systems in Large Buildings. 2014, Elevcon, Paris, France. 20: 51-61. The International Association of Elevator Engineers.

#### **BIOGRAPHICAL DETAILS**



Lutfi Al-Sharif is currently the Vice President at Al-Hussein Technical University in Amman/Jordan, and jointly Professor of Building Transportation Systems at the Department of Mechatronics Engineering, The University of Jordan. He received his Ph.D. in elevator traffic analysis in 1992 from the University of Manchester, U.K. He worked for 10 years for the London Underground, London, United Kingdom, in lifts and escalators. He has over 50 papers published in peer-reviewed journals and conferences in vertical transportation systems, is co-inventor of four patents and co-author of the 2<sup>nd</sup> edition of the Elevator Traffic Handbook, and author of the "indoor transportation" chapter in the Elsevier Encyclopedia of Transportation.



Richard Peters has a degree in Electrical Engineering and a Doctorate for research in Vertical Transportation. He is a director of Peters Research Ltd and a Visiting Professor at the University of Northampton. He has been awarded Fellowship of the Institution of Engineering and Technology and the Chartered Institution of Building Services Engineers. Dr Peters is the principal author of Elevate, elevator traffic analysis and simulation software.



Matthew Appleby is a software engineer at Peters Research Ltd. He is involved in project managing the new Elevate Online and leads frontend development. He joined Peters Research in 2019 and graduated with a first-class (Hons) degree in Software Development in 2024.



Eng. Tahani Ghobon has been working as a teaching assistant at the Electrical Engineering Department at Al Hussein Technical University (HTU) for the last 6 years. She completed her undergraduate degree in electrical engineering (Electrical Power) from the Jordan University of Science and Technology (JUST). Eng. Tahani has vast experience in the area of Machine Learning and its applications in electrical engineering applications.