ISSN 2522-3488 Vol. 7, Issue 7: 53-60, DOI: 10.25236/IJNDES.2023.070708

Urban Rail Transit Planning and Design Based on ARIMA Model and AnyLogic Software Station Pedestrian Flow Simulation

Guang Yang, Jianye Chen*

College of Water Resource & Hydropower, Sichuan University, Chengdu, 610065, China
**Corresponding author*

Abstract: *Accompanied by the integration of the world economy and high-speed development, the continuous promotion of the construction of towns and cities with the gradual saturation of urban road traffic, subway, light rail, and other rail transportation has become one of the main ways for residents of large cities to travel. This paper borrows from the network information big data and location conditions analysis, through the ARIMA time series combined with the passenger flow conveyance prediction and combined with the corresponding AnyLogic software to realize the subway entrance and exit people simulation. Accordingly, it realizes the construction planning and station design of the future rail transit trunk line in Haikou City, the capital of Hainan Province.*

Keywords: *ARIMA Model, AnyLogic Simulation Model, Urban Rail Transit Planning and Station Design, Footfall Prediction Models, Analysis of Key Urban Locations*

1. Introduction

1.1 Design Example, Haikou City Backgrounder

Haikou, also known as "Coconut City", a prefecture-level city under the jurisdiction of Hainan Province, the provincial capital, China's "One Belt, One Road" strategic pivot city, the core city of the Hainan Free Trade Port, the Beibu Gulf City Cluster important node city, is located in the northern part of Hainan Island.

Haikou City has 4 districts (county level), namely Xiuying, Longhua, Qiongshan, and Meilan; it has 21 streets, 22 towns, 211 communities, and 248 administrative villages. The main population distribution and urban construction are concentrated in the area near the sea and inland locations.

The development of urban transportation in Haikou City is mainly based on the combination of sea, land, and air amphibious development, which mainly serves the purposes of domestic and international transportation and logistics by land and sea, the convenience of transportation for tourists and the national military arrangement and strategic deployment.

1.2 Reasons for design case selection

Haikou is the capital city of Hainan Province, with good natural environmental resources, and an internationally renowned tourism industry. It has good geological conditions, but no subway has been built.

(1) In terms of strategic economic development status, Hainan Province itself relies on the tertiary industry to achieve economic growth, with great demand for people's transportation. The construction of rail transit is conducive to easing the pressure of urban traffic and improving the efficiency of economic communication.

(2) From the national strategic development plan, underground rail transit is the backbone and basis for the future development of underground space in the city. Haikou City, the modernization and development of Hainan Province is conducive to the full use of resources in the South China Sea area, driving the development and construction of the surrounding areas, providing a solid economic foundation for the defense of the South China Sea and the development of the South China Sea.

ISSN 2522-3488 Vol. 7, Issue 7: 53-60, DOI: 10.25236/IJNDES.2023.070708

Therefore, the construction of rail transit in Haikou City is a must-issue. The local government and relevant departments are actively discussing, planning, and bidding.

2. ARIMA model for the prediction of passenger conveyance

2.1 Regional Classification and Delineation of Traffic Circle Subdivisions

The Haikou Municipal Bureau of Statistics official website published the *"Haikou City, Haikou City, the end of 2021, the main data of the resident population"*, it can obtain: to the end of 2021, Haikou City, the overall resident population of 2.980 million people, the Xiuying District of 573,900 people, Longhua District of 807,300 people, the Qiongshan District of 663,500 people, Meilan District of 863,300 people. The data for each street in each district was extracted from a third-party website: *wap.tcmap.com.cn*, but the resident population of each street and county was extracted as of 2011.

Since it is a resident population, theoretically the population change satisfies the following mathematical equation in the same yearly difference:

$$N_{s2021} = N_{s2011} \cdot \prod_{i=2012}^{2021} n_i \tag{1}$$

Where N_{si} is the total resident population of the region in the year i, and n_i is the overall population growth rate of the region in the year i relative to year i-1 (including the natural growth rate and the rate of in-migration).

And in the last decade with the modernization of Haikou City, the process of rural urbanization is accelerating (as of 2021, Haikou City, has an integrated urbanization rate of 82.61%). The gap between the population growth rate of the towns and the rural areas gradually narrowed and then flat with each other [1]. It can be roughly simplified to think that the natural population growth rate of the annual population of the various streets and towns within a district is equal. Then according to the original 2011 resident population data as a ratio, we projected the population to the end of 2021. Then we estimated the average distance from the administrative area of each district to the center of the main urban area. As a result, the following planning data were obtained (only a portion of the resident population data is shown here due to space limitations), as shown in Table 1.

Table 1: Haikou City permanent population data by district

	Resident population N (10 thousand)	Footprint S (km^2)	Population density D (people/km^2)	km
Meilan district				
Haifu Street	13.01	4.3	30261.91	7.2
Lantian Street	5.83	5	11653.09	8
Boai Street	0.97	0.3	32369.70	5.5
Haidian Street	9.13	6.72	13583.71	9.6
Renmin Road Street	12.62	8	15780.23	4.3
Bailong Street	7.77	7	11098.18	8.1
Heping South Street	5.83	2.5	23306.19	6.9
Baisha Street	4.95	3.21	15428.55	6.7
Xinfu Street	1.94	8.3	2339.98	7.2
Lingshan Town	8.93	98.6	906.09	14
Yanfeng Town	5.05	119.78	421.58	33
Sanjiang Town	4.80	70.9	676.61	40
Dazhipo town	5.50	114	482.14	52
Total population/ Total area	86.33	448.61	1924.39	

The corresponding visual distribution map of the resident population is made.

ISSN 2522-3488 Vol. 7, Issue 7: 53-60, DOI: 10.25236/IJNDES.2023.070708

2.2 Activity Configuration and Demand Solving

Resident population:

In the same time series, the amount of activity generated for the resident population of each area should be limited and vary regularly over time. It is useful to define the combined mean-time activity (total distance traveled during this period) per square kilometer of an area for a small period as a function of time, population density, and the cost of economic exchange demand (here characterized by the distance of the area from the economic center of the city):

$$k = f(t, D, d) \tag{2}$$

Then the total activity K_s of the resident population from period t_1 to t_2 can be expressed as:

$$K_s = \int_{t_1}^{t_2} k \cdot S \cdot dt = \int_{t_1}^{t_2} f(t, D, d) \cdot S \cdot dt \tag{3}$$

In this paper, we use the WeChat app *"Yi Travel"* [2] to query the activity volume k in the street area of Fucheng, Haikou City, in one month. And the rest of the street activity volume is extrapolated according to the functional assumption that it is proportional to the population density and inversely proportional to the distance from the economic center of the city. To simplify the calculation, the period calculation unit is set as 24 hours a day, and temporarily ignores the influence of some irregular holidays on the activity level. Then based on the seasonal differences in the activities of certain jobs, such as teachers, students, etc., the data characteristics can be simplified to the smaller differences in the daily activity level in the same season. After that, the ARIMA algorithm is used for validation and prediction.[3]

Mobile population:

Due to the geographic conditions and economic development plans of Hainan Province, the vast majority of the floating population in Haikou City mainly comes from tourism industry sources. Therefore, the development of the tourism industry in Hainan Province will influence the floating population. At the same time, because of Hainan's tropical climate conditions, resulting in Haikou City in the fall and winter seasons of the foreign population is far more than the spring and summer which causes the formation of seasonal differences in mobile population data characteristics. In this paper, the winter sports heat map of Haikou City will be used to determine the winter sports heat map of Haikou City, while giving the corresponding proportion based on the resident population. The seasonal change characteristics and annual change trend of foreign population will be equally applicable to the prediction of ARIMA time series algorithm model.

2.3 Principal application and solution of ARIMA models

ARIMA (autoregressive moving average model), the algorithm mainly uses difference operation to process the data and obtain the trend of its function, transform the non-stationary time series into a stationary series, and then regress the lagged values of the dependent variable as well as the present and lagged values of the stochastic error term [4]. Difference operations can smooth a class of non-smooth series, if the first order difference does not smooth the time series, then second order difference, third order difference, until the number d order difference. Finally, the series is smoothed.

1st differentials:

$$\nabla X_t = X_t - X_{t-1} = (1-B)X_t$$

2nd differentials:

$$\nabla^2 X_t = X_t - 2X_{t-1} + X_{t-2} = (1-B)^2 X_t$$

In general, the j-order difference is divided into:

$$\nabla^j X_t = (1-B)^j X_t$$

∇^j becomes a difference operator of order j, then

$$\nabla^j = (1-B)^j = 1 - \binom{j}{1} B + \binom{j}{2} B^2 + L + (-1)^{j-1} + (-1)^j B^j$$

ISSN 2522-3488 Vol. 7, Issue 7: 53-60, DOI: 10.25236/IJNDES.2023.070708

Suppose $\{X_t, t=0,\pm 1,\pm 2, L\}$ is a non-smooth sequence, if there exists a positive integer j such that:

$$\nabla^j X_t = W_t$$

And $\{X_t, t=0,\pm 1,\pm 2, L\}$ is an *ARIMA(p,q)* sequence is said to be an *ARIMA(p,j,q)* sequence, then X_t is satisfied:

$$\varphi(B)\nabla^j X_t = \theta(B)\varepsilon_t$$

If $\nabla^j X_t$ is a smooth sequence but with mean $\mu \neq 0$, then $\nabla^j X_t - \mu$ is a smooth zero-mean sequence that satisfies:

$$\varphi(B)\nabla^j X_t = \theta(B)\varepsilon_t, t > j$$

Then X_t is said to be $ARIMA(p, j, q)$. If the mean value μ of the sequence is unknown, the mean $\bar{x}$ of $\nabla^j X_t$ is used for estimation.

And if the initial value X_1, X_2, L, X_j is known, then:

$$W_t = \nabla^j X_t, t = j+1, L, n$$

It is possible to recover X_t

In particular, for the seasonal model, the sequence with period s can be first differenced, i.e.:

$$\begin{aligned} \nabla_s X_t &= (1-B^s)X_t \\ \nabla_s^j X_t &= (1-B^s)^j X_t \end{aligned} \tag{4}$$

Then we can do modeling and arithmetic.

Based on the one-month time series data obtained from the program "YI Travel", ARIMA time series model was established, and it was found that the differential series had good smoothness based on the cases of j=1 and 2. Using the November data to make the difference, and then trying to calculate that *q=3,p=2* [5] fit better, the forecast chooses *(q,j,p)=(3,1,2)*or*(3,2,2)*, to make the following differential prediction data , as shown in Figure 1 and Figure 2.

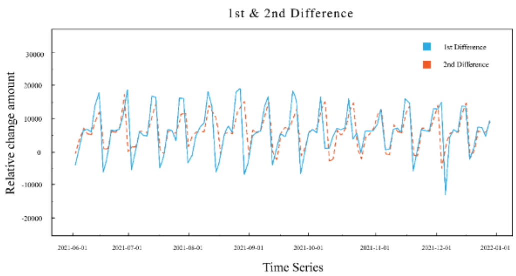

Figure 1: 1st & 2nd prediction of resident population

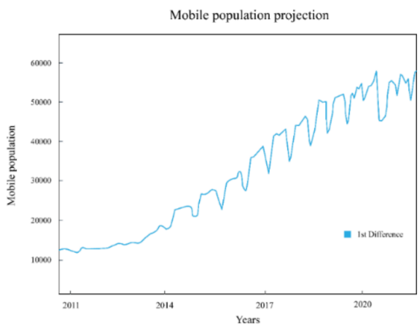

Figure 2: 1st prediction of mobile population

It can be observed that for the time series data of the activity level of the resident population, the approximate equal treatment in quarterly units is within the error tolerance. If the data prediction is according to the first-order difference, it can be observed that the activity level of the people has a large decline in the winter months of November to December, but the decline range is still within the tolerance of the regular change so that the activity level of the whole year can be averaged out and processed.

The 1st order difference is chosen because the data have a strong correlation with the time year. From the differential prediction model, the mobile population is the overall trend of increasing year by year, reflecting the data characteristics caused by the continuous development of Hainan Province's tourism industry and the increasing publicity efforts.

ISSN 2522-3488 Vol. 7, Issue 7: 53-60, DOI: 10.25236/IJNDES.2023.070708

2.4 Ridership projection

As a result of the projection and validation of the activity volume in the above-mentioned regions, the activity volume of the resident population in each region was calculated and determined. Due to the temporal regularity of the activity volume of the resident population, the quarterly data are represented by the average daily unit after validation.

There is a certain proportional relationship between the mobile population and the resident population [6], which is due to the influence of the region's unique tourist attraction location or the degree of regional economic development.

$$N_{total-population} = \varepsilon_n N_{resident-population} \tag{5}$$

Subsequently, the total population conversion ratio ε was determined for each region based on the campaign heat map relatively.

The regional activity volume is converted to the number of riders, where the conversion ratio α is related to the regional population ridership/transfer rate and the average land activity volume , as shown in Table 2.

Table 2: Average activity and ridership projections for the region in a month

Regions	Volume of activity(km/km)	Projected ridership A_s (resident population)	Projected ridership A_c (total & winter)
Meilan district			
Haifu street	174720.49	113568	306634
Lantian street	60552.39	39359	106269
Boai street	244656.11	159026	429371
Haidian street	58820.32	38233	103229
Renmin Road street	152554.46	99160	267733
Bailong street	56956.98	37022	85150
Heping South street	140411.33	91267	209915
Baisha street	95726.10	62221	143110
Xinfu street	13510.13	8781	20197
Lingshan town	2690.44	1748	2063
Yanfeng town	531.06	345	407
Sanjiang town	703.17	457	539
Dazhipo town	385.43	250	295

3. AnyLogic Station Simulation Model

3.1 Modeling of subway stations

For 3D subway station entrances and exits simulation, we directly simulate the intermediate station ordinary 4-exit subway station for simulation to explore the feasibility of its station scale.

Simulating the pedestrian flow using the pedestrian library of AnyLogic software and then model simulation using the theoretical pedestrian density statistics H (total population density D × ride conversion ratio α) derived above has the advantage of ensuring acceptable performance of service points with assumed loads, estimating the dwell time in characteristic areas, detecting potential problems in evacuating the escape from the internal geometric space, adding the effect of obstacles and many other Applications. Assuming a 16-hour operating day for commonly used entry and exit points, the predicted daily passenger flow is converted to hourly passenger flow for real-time prediction. For example, Fucheng street in Qiongshan District has a one-month projected ridership of 279097, converted in $\frac{279097}{30\times16} = 581.45\,people/hour$. In this paper, the peak passenger flow of a day is considered, and *600 people/hour* are taken for simulation.[7]

ISSN 2522-3488 Vol. 7, Issue 7: 53-60, DOI: 10.25236/IJNDES.2023.070708

Analogous to a typical two-story common site, the modeling is completed, as shown in Figure 3.

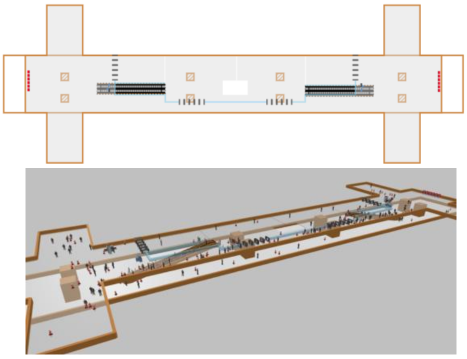

Figure 3: Physical simulation model of a classic station

3.2 Traffic generation and corresponding behavioral logic settings

In this paper, we consider the crowd generation and behavioral logic of inbound and outbound gates based on the modeling, as shown in Figure 4.

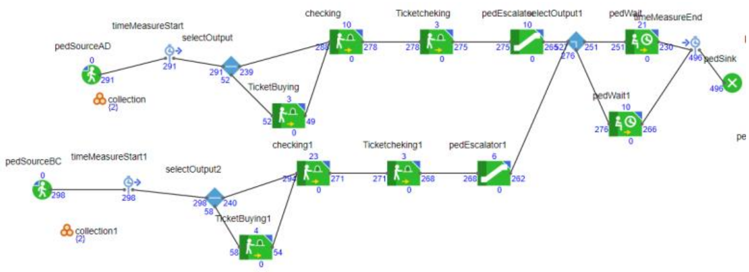

Figure 4: Inbound pedestrian behavior logic settings

Respectively simulate the subway station entrance to the left side of the AD port and the right side of the BC port of the pedestrian flow, the input data for the Haifu Street business 16 hours a day the average number of people per hour (about 600 people/hour), pedestrians buy tickets probability is set to 0.2, after checking the ticket close to the escalator to enter the B2 level. After entering B2 and waiting for waiting in the direction of both sides of the train, the probability of pedestrian riding direction is set to 0.5, and the probability of coming and going is equal. In the intelligent body collection, each person at the station began until he boarded the train, respectively, set the timeMeasureStart/End module, easy to analyze the average length of the ride.

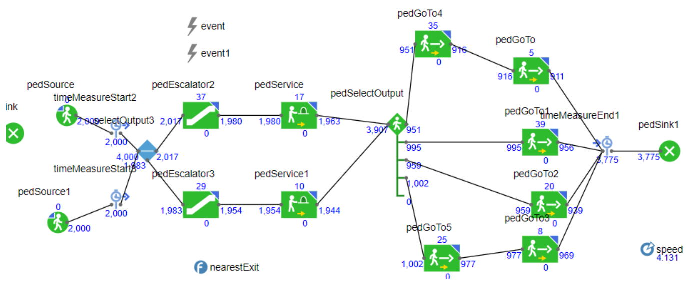

Figure 5: Outbound pedestrian behavior logic settings

In the exit simulation, the probability of an exiting passenger entering the escalator to the B1 level is set to 0.5 equally, after which the ticket is checked and exited in the nearest station. In particular, the exit choice after the exit is random, and each probability is set to 0.25, due to the model underlying logic of the shortest path, passengers will appear after the ticket checking reversed retrograde from the inbound ticket gate to the exit, to avoid this behavior does not match the reality, especially added path guidance. And the parameters event1 and event2 are set to the train schedule function, to the center of the urban area 3min a class schedule output out of the station passengers and empty is waiting for the waiting passengers.

From the figure 5, we can visualize that the degree of congestion in the subway station is relatively

ISSN 2522-3488 Vol. 7, Issue 7: 53-60, DOI: 10.25236/IJNDES.2023.070708

moderate, according to the morning and evening peak hours of the realistic daily routine, the number of people entering and exiting the station will reach 1.5 to 2.5 times the average flow of people, and off-peak hours will drop to 0.6 to 0.8 times the average flow of people. [8-9] Theoretically, the subway station can operate normally and ensure the function of people entering and leaving the station.

3.3 Corresponding statistics output

Once the model has been run, it can be intercepted to obtain the average time spent per person entering the station for a ride and the average time spent exiting the station for a departure at that subway station, as shown in Figure 6.

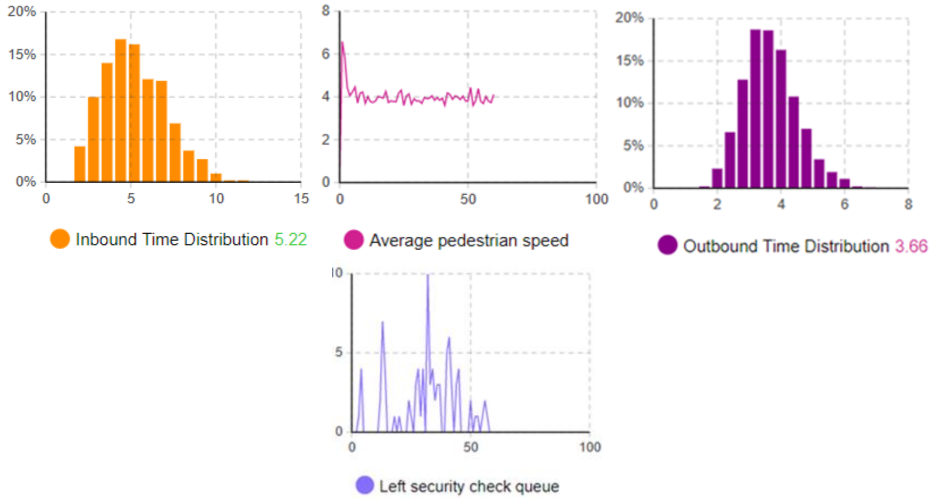

Figure 6: Relevant Data Output Statistical Charts

After the model is run, the histogram of the distribution of the inbound and outbound time of the station, the average speed of the pedestrians after the model is stabilized, and the number of people waiting in the queue at the left security checkpoint are obtained, as shown in Figure 7

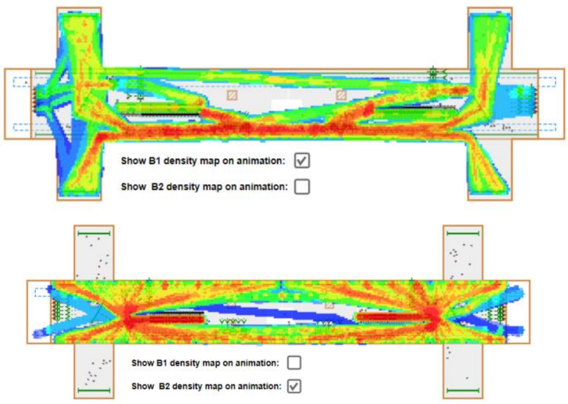

Figure 7: Spatial density map of pedestrian flow on floors B1 and B2

3.4 Summary and Evaluation of the Role of the Model and simulation

The following two measures for attention or improvement can be summarized from the experience of operating the model:

(1) To maintain an orderly and efficient operation during the security check, more security checkpoints should be opened or the efficiency of the security check should be improved.

(2) When setting up the space in the station, it should be avoided that the space at the gates and other places where the flow of people is concentrated is narrow and overcrowded so that it is easy to block the flow of people and make the overall transportation efficiency of the station drop significantly.

At the same time, the passenger flow data predicted by the ARIMA model is also closer to the actual

ISSN 2522-3488 Vol. 7, Issue 7: 53-60, DOI: 10.25236/IJNDES.2023.070708

situation and has higher credibility.[10]

4. Conclusions & Preliminary design of subway lines and stations

In recent years, China's urbanization and subway construction is in full swing. The ARIMA time series prediction of passenger flow used in the article combined with Anylogic simulation software to verify the method combined with the key location characteristics of the region can play a preliminary role in the construction of the subway, and at the same time, this design method is relatively faster and more efficient. Concerning the data and traffic circles analyzed in the previous section, the central city is identified. By the planning principles, key areas with a larger scope and wider impact are selected for the whole of Haikou City. At the same time, the existing suburban trains should be well-connected and supplemented. The rail design is shown below.[11]

Through the prediction simulation and other methods mentioned above, the subway line plan made in this paper is roughly the same as the line made in the existing plan of the city government, and the plan is more successful. Compared with the government plan, the plan in this article also focuses on the economic exchange of workers in industrial parks and the promotion of the future development of the southern countryside.

References

[1] Xinwei Gao. A Comprehensive Study on the Development of Land Surrounding Urban Rail Transit Stations in Groups Taking the Case of Hainan's Transitized Tourist Train as an Example [J]. China Construction, 2022, No. 281(04):104-105.

[2] Yuan Fang, Feng Zhan, Duochun Ma. Design of a small program based on WeChat platform for residents' travel survey [J]. Modern Information Technology, 2022, 6(14):126-128+133. DOI:10. 19850/j. cnki. 2096-4706. 2022. 014. 029.

[3] Zhiyi Guo, Hongliang Liu. Application and Construction of Multidimensional Adaptive Computing Models in Estimating the Size of Migrating Population--A Study Based on Census Data in Gansu Province [J]. Northwest Population Journal, 2018, 39(04):105-111. DOI:10. 15884/j. cnki. issn. 1007-0672. 2018. 04. 014.

[4] Hao Teng, Wei Guo. Settlement prediction based on ARIMA model with RBF neural network[J]. Geomatics & Spatial Information Technology, 2022, 45(S1):275-278.

[5] Jie Bai, Hongbo Hu. Values of regression model parameters and their uncertainty assessment in measurement [J/OL]. Acta Metrologica Sinica, 2022(12):1-6[2023-01-07]. http://kns. cnki. net/kcms/ detail/ 11. 1864. TB. 20221220. 1358. 015. html.

[6] Zhiping L .Simulation Analysis of Emergency Evacuation at Urban Rail Transit Station[J].China Transportation Review, 2018.

[7] Antonova, V. M., Grechishkina, N. A. & Kuznetsov, N. A. Analysis of the Modeling Results for Passenger Traffic at an Underground Station Using AnyLogic. J. Commun. Technol. Electron. 65, 712–715 (2020). https://doi. org/10. 1134/S1064226920060029

[8] Tianyi Ma. Analysis of differences in the value of morning and evening peak travel time for commuters [D]. Southwest Jiaotong University, 2019. DOI:10. 27414/d. cnki. gxnju. 2019. 001647.

[9] Binbin Tan. Simulation and Optimization of Urban Rail Transit Station Facility Layout Based on AnyLogic [J]. Auto Time, 2023, No. 405(09):193-195.

[10] Andrés Fielbaum, Sergio Jara-Diaz, Antonio Gschwender, Lines spacing and scale economies in the strategic design of transit systems in a parametric city. , Research in Transportation Economics, Volume 90, 2021, 100991, ISSN 0739-8859, https://doi. org/10. 1016/j. retrec. 2020. 100991.

[11] Oded Cats, Alex Vermeulen, Martijn Warnier, Hans van Lint, Modelling growth principles of metropolitan public transport networks, Journal of Transport Geography, Volume 82, 2020, 102567, ISSN 0966-6923, https://doi. org/10. 1016/j. jtrangeo. 2019. 102567.