

OPTIMIZATION SYSTEM OF RESERVED/NON-RESERVED SEATING PLANS FOR IMPROVING CONVENIENCE AND REVENUE OF INTERCITY EXPRESS TRAINS

S. NAKAGAWA, M. SHIBATA, & N. FUKASAWA

Signalling and Transport Information Technology Division, Railway Technical Research Institute, Japan.

ABSTRACT

In Japan, seating plans of reserved/non-reserved seats in intercity express trains do not necessarily correspond to passenger demands which may vary depending on many factors such as calendar days and operating sections. This mismatch possibly causes passengers' inconvenience due to passenger congestions on board, and railway operators can lose revenue. Optimizing seating plans is expected to be an effective solution for these problems, where predicting both potential passenger demands and reflections to the potential demand caused by the passenger congestions is necessary.

For establishing the prediction methods, we conducted some demand analyses and some questionnaire surveys. According to these results, we establish demand estimation models of reserved/non-reserved seats which can consider potential demands and behaviour changes if the passengers face congestion. We developed a prototype of seating plans optimization system applying these models, simulation of passengers' behaviour and genetic algorithm. This paper describes these models, the optimization system and the result of applying this system to some existing intercity express trains as a case study. We got the optimized seating plan which is expected to improve both of the passengers' convenience and the railway operators' revenue.

Keywords: congestion, genetic algorithm, intercity trains, potential demands estimation, revenue management, seat reservation, second best choice behaviour.

1 INTRODUCTION

In Japan, seating plans of reserved/non-reserved seats in intercity express trains do not necessarily correspond to passenger demands, because the demand is more flexible than the seating plans (Minami et al. [1]). This problem possibly causes passengers' inconvenience, and railway operators can lose revenue. We suppose optimizing seating plans is expected to be an effective solution, as described in Shibata et al. [2, 3].

This paper describes the seating plans optimization system we have developed and the models involved in the system, which estimate potential demands and passengers' behaviour. By conducting a case study, we can show that the optimized seating plan improves both the passengers' convenience and the railway operators' revenue.

1.1 Abbreviations and terminologies

In this article, we use the following abbreviations:

RS(s) – reserved seat(s).

NRS(s) – non-reserved seat(s).

ICT(s) – intercity express train(s).

LF – load factor.

We also use the terminology ‘OD’. This is the combination of ‘origin’ and ‘destination’ stations of passengers’ travel.

2 ESTABLISHING POTENTIAL PASSENGER DEMANDS ESTIMATION MODEL

2.1 Existing revealed demand data

To estimate the potential demand, we have to know the revealed demands of the target ICTs. We can get three kinds of the revealed demand data as described in Table 1. We use these data to calibrate the potential passenger demands estimation model.

(1) Sales results accumulated in seat reservation system

All RSs are sold through the seat reservation system. We can get sales results of RSs for all days and trains from the system. By using these data, we can observe all passengers’ OD. We have to consider that the sales results do not necessarily correspond to the revealed demands; the tickets may not necessarily be used in accordance with their validity. We also have to consider that we cannot get the sales results of NRSs from these data.

(2) OD survey for ICTs

For some ICTs, OD survey has been conducted to gain the OD data. Usually this survey is carried out in some weekdays per year, by collecting used tickets.

From this survey, we can observe OD data of not only RSs passengers but also NRSs passengers with relatively higher accuracy. On the other hand, we cannot grasp short-term change (e.g. seasonal change) of the demand from this survey, because this survey is annual.

(3) Visual count of ridership

The train conductors of ICTs have to report the number of ridership in predetermined target sections. They walk through the train and generally count the passengers in each seat class. If the target section consists of some station-to-station sections and some passengers get on board or drop off within the target section, the record should be the maximum number of ridership through the target section. The most important feature of these data is that they are recorded through all days, all trains and all of the seat classes. By analysing them, therefore, we can recognize the changes of the demands depending on the season, the day of the week, etc. On the other hand, we have to consider some defects of them – their accuracy is relatively low compared to other data and they are not OD data.

2.2 An online survey

To establish the potential passenger demands estimation model for NRS, we conducted an online survey. In this survey, we asked how passengers behave when they want to board NRS

Table 1: Existing real demand data.

	(1) Sales results recorded in seat reservation system	(2) OD survey for ICTs	(3) Visual count of ridership
Method to get the data	From the seat reservation system	Collecting used tickets	Visual count by the train conductors
Type of the data	OD	OD	Ridership on target sections
Date	All days	Target date of the survey	All days
Seat type	RS and first-class seats	All types	All types
Features	OD data of RS are available	OD data of NRS are available	Demand data on all days and of all seat types are available
Demerits	Do not always correspond with revealed demands	Cannot grasp short-term change of the demand from this data	Less accuracy than other data

and face passenger congestion. The outline of the survey is described in Table 2. We also asked how passengers choose their seat class to board and the impact of the price difference between RS and NRS (Nakagawa et al. [4]).

The results are shown in Table 3. Most of the passengers want to get a seat in NRS, although there are only few seats available in NRS. On the other hand, if there is no seat available (i.e. they have to be standing in deck or aisle), 35% of them avoid boarding the NRS.

Table 2: Outline of the online survey.

Target of the survey	Persons who: <ul style="list-style-type: none"> – belong to the age group between the age of a high school student and a 69 years old – live in Japan (excluding Okinawa Pref.) – have experiences to get on the ICT (not including Shinkansen super express) operated by JR (Japan Railways) in last one year
Period of the survey	From 30 January to 13 February 2014
Number of answerer	Target: 8,130 Valid response: 6,036 (response rate: 74.2%)

Table 3: Answers for ‘How do you behave when you want to board NRS and face passenger congestion?’

	Do not mind	Mind a little but not feeling uncomfortable	Board the ICT although feeling uncomfortable	Feeling uncomfortable and do not board the ICT
(1)	3,758	1,616	562	100
(2)	3,357	1,854	628	197
(3)	3,134	1,989	743	170

(continued)

Table 3: continued

	Do not mind	Mind a little but not feeling uncomfortable	Board the ICT although feeling uncomfortable	Feeling uncomfortable and do not board the ICT
(4)	832	714	2,388	2,102
(5)	501	350	1,989	3,196

- (1) Both the window seat and the aisle seats are available (for all empty seats, and adjacent, front and rear seats are occupied)
- (2) There are some window seats available but the aisle seats are fully occupied (ditto)
- (3) There are some aisle seats available but the window seats are fully occupied (ditto)
- (4) Cannot be seated due to full occupancy (but no one stands in the deck)
- (5) Cannot be seated due to full occupancy and the deck is crowded

2.3 THE POTENTIAL PASSENGER DEMANDS ESTIMATION MODEL

We established the potential passenger demands estimation model from the data described above. The outline of the model is shown in Fig. 1. This model estimates the potential OD data of some ICTs on a particular date from the revealed demand data of the ICTs. Because there are few revealed OD data for NRS, we tried to estimate the potential OD data indirectly; i.e. this model estimates the ridership on some sections and converts them into OD data.

- (1) Calculate the average of the counts of ridership reported on the same date over the past years, where we consider the structure of calendar date. This is treated as the revealed ridership on the section on the target date and train.
- (2) From the estimated ridership calculated in (1), we estimate the potential ridership, by using the equation describing the relationship between the potential demands and the revealed demands.
- (3) Convert the potential ridership calculated in (2) into the potential OD data of the target date and train, by applying the result of OD survey for target ICT as a OD pattern data.

The detail of each method is as follows.

2.3.1 Relationship between the potential and revealed demands (for NRS)

We tried to formulate the relationship between the revealed ridership (actual demands) and the number of the passengers who avoid to board the ICTs due to passenger congestion (lost

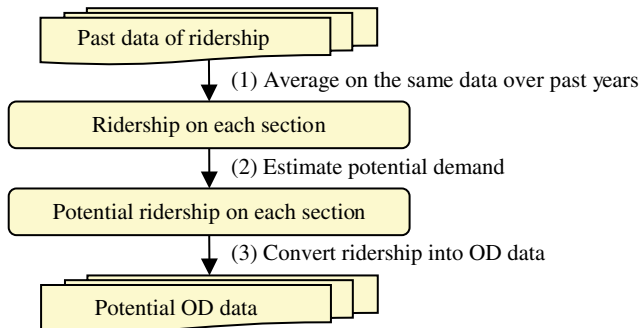


Figure 1: Outline of the potential passenger demands estimation model.

demands), by utilizing the result of the online survey described in Section 2.2. First, we assume that the LF is 75% if there are some vacant seats – either window seat or aisle seat. This is because usually half of the seats in the ICTs are window seats and the rest of the half is aisle seats. Then, we assume that when the LF is between 75% and 100%, behaviour of ones who intend to get on is average of the second and third rows in Table 3. Under this assumption, the ratio between passengers becoming the revealed demands (i.e. who take the ICT) and passengers becoming the lost demands (i.e. who abort taking the ICT) is calculated as below:

$$(3357+1854+628+3134+1989+743) : (197+170) = 11705 : 367$$

Hence, when the LF is between 75% and 100%, the lost demands are estimated to be 367/11705 of the LF minus 75%.

In the same way, when the LF is over 100%, the ratio is as below:

$$(832+714+2388) : 2102 = 3934 : 2102$$

Therefore, when the LF is over 100%, the lost demands are estimated to be the summation of 2102/3934 of the LF minus 100% and the lost demands are generated if the LF is between 75% and 100%.

By the estimation described above, we can formulate the relationship between the potential LF and revealed LF of NRS. The result is shown in Fig. 2. Note that we consider that the lost demands are not generated when the LF is less than 75%.

2.3.2 Relationship between the potential and revealed demands (for RS)

To estimate the potential demands of RS is more difficult than that of NRS, for the reasons described below.

- (1) The LF of RS, the number of seats reserved divided by the number of RS, never exceeds 100%. For this reason, we cannot conclude that ‘the ICT whose revealed LF is high has high potential demands’ with absolute certainty.
- (2) Sometimes the seat reservation request is rejected though the LF is less than 100%. This is because the request is accepted only when there is a seat available through the requested OD. Therefore, we cannot understand whether the rejection of the seat reservation request occurred or not from the LF.

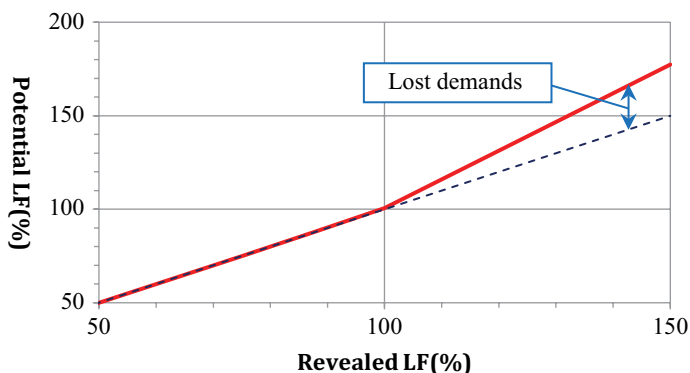


Figure 2: The relationship between the potential and revealed LF of NRS.

- (3) Most of the rejections for seat reservation requests are not recorded in the seat reservation system due to its specifications. In addition, the records of the rejections do not necessarily represent the passengers' initial choice. For these reasons, we cannot observe how many rejections of the seat reservation request have truly occurred from the seat reservation system.

We thought about a method to estimate the potential demands of RS from the revealed demand data considering these problems. First, while there is the problem (1) shown above, we decided to formulate the relationship between the potential and revealed LF of RS, which is similar to that of NRS. It is because we have no other data from which we can estimate the volume of potential demands. Next, we assumed that the rejection for the seat reservation occurred only when the LF is over 85%, as a result of an interview survey for a railway operator. Furthermore, we analysed some sales results of ICT recorded in the seat reservation system and noticed that most of the RSs are reserved on the day of operation and the rate of reservation on the previous days of operation rarely exceeds 85%. From these situations and knowledge, we supposed some assumptions described below can be applied.

- (a) When the revealed LF of RS is under 85%, no rejection for reservation request has occurred. When the LF is over 85%, some rejection has occurred at a constant rate for the excess of 85%. In other words, the equations described below are satisfied, where y is the potential LF and x is the revealed LF:

$$y = x \quad (\text{if } x \leq 0.85)$$

$$y = x + p(x - 0.85) \quad (\text{if } x > 0.85; p \text{ is a coefficient})$$

- (b) The rejection for reservation occurs only on the day of operation.
 (c) There is linear relationship between the rate of reservation on the previous day of operation and the potential LF.

With these assumptions, we conducted a linear regression analysis between the rate of reservation on the previous day of the operation and the revealed LF (from assumption (a), this is equal to the potential LF) for every ICT whose LF is under 85%. Then we calculated the coefficient p to fit the regression line with the ICTs whose LF is over 85%. Using the obtained value of p , relational expression between the potential LF and revealed LF has been calculated.

2.3.3 A method to get the OD data from the ridership on each section

As described in Section 2.1, there are few actual measured OD data for NRS. For this reason, we established a method to get the potential OD data by adjusting the past real OD data to correspond with the estimated potential ridership. The outline is as below.

- (1) Consider the previous result of OD survey (or the result of an extra survey, if the target date is a holiday) of the target ICT as the basic OD data. We note that we describe the number of passengers who take the ICT from station i to station j in this basic OD data as t_{ij} .
- (2) For every station appeared in the basic OD data, we define the coefficient a_i . We assume that the number of passengers from station i to station j in the OD data is calculated by $t_{ij} * a_i$ (as shown in Fig. 3).

Basic OD data		Get off at					
		Sta. 1	Sta. 2	Sta. 3	Sta. 4	Sta. 5	Sta. 6
Get on at	Sta. 1		10	20	50	20	30
	Sta. 2			10	40	10	20
	Sta. 3				30	10	20
	Sta. 4					30	20
	Sta. 5						10
	Sta. 6						

↓

After multiplying the coefficients		Get off at					
		Sta. 1	Sta. 2	Sta. 3	Sta. 4	Sta. 5	Sta. 6
Get on at	Sta. 1		$10a_1$	$20a_1$	$50a_1$	$20a_1$	$30a_1$
	Sta. 2			$10a_2$	$40a_2$	$10a_2$	$20a_2$
	Sta. 3				$30a_3$	$10a_3$	$20a_3$
	Sta. 4					$30a_4$	$20a_4$
	Sta. 5						$10a_5$
	Sta. 6						

Figure 3: An example of applying the converting method.

- (3) Convert the OD data assumed in (2) into the ridership on every section. Then calculate the coefficient a_i which minimizes the square error between the converted ridership and the potential ridership. The potential OD data are calculated by using this a_i . For example, when the OD data assumed in (2) are as given in the table in Fig. 3 and the potential ridership in the section between station 3 and station 4 is 200, this method calculates a_i to fit $100a_1 + 70a_2 + 60a_3$ to 200.

In other words, this method stands on the assumption that the ratio of the passengers who get off at each station to who get on a certain station is constant. Under this assumption, this method calculates the potential OD data by adjusting the basic OD data to suit the estimated potential ridership which is the input of the method.

3 ESTABLISHING SECOND BEST CHOICE BEHAVIOUR MODEL

To utilize the potential demands to determine how the transport service is, we have to consider how the customers think when they faced congestion. In other words, we have to understand how the potential demands change to revealed demands. To measure this, we carried out a questionnaire survey for the passengers of some ICTs. In this survey, we observed their choice preference if they could not take a seat/they could not get a reservation. For RS passengers, the question was ‘How did you behave if you couldn’t buy a ticket for RS due to full occupancy?’ For NRS passengers, the question was ‘How did you behave if you faced passenger congestion of NRS when you got on board?’ The outline of the survey is described in Table 4.

The result is shown in Table 5. As for RS passengers, about half of them would try to reserve RS of other trains. About 40% would decide to board on NRS. Only 4% answered that they would avoid boarding the ICT.

Though they had to stand in the deck or the aisle, 63% of the NRS passengers would choose to board the NRS. In other words, 37% of them would avoid boarding the NRS. This corresponds well to the result of the other online survey described in Section 2.2. 25% of the passengers would wait for the next train: 8% would try to reserve RS, and 17% would wait in line for NRS.

Table 4: Outline of the onboard survey.

Target of the survey	Persons who: – board the target ICT on the sections where the visual count of ridership is recorded – seat in RS or NRS – are not young children
Period of the survey	17 and 18 January 2015
Target ICT	10 trains each day
Number of answerer	Distributed: 4,194 Valid response: 3905 (response rate: 93.1%)

Table 5: The passengers’ choice behaviour if they could not take the seat they were seating.

	Passengers of RS	Passengers of NRS
Reserve RS of earlier service	261	–
Reserve RS of latter service	291	209
Board NRS of earlier service	6	–
Board NRS of this train	447	1,748
Board NRS of latter service	15	470
Abort taking ICT	41	93
Others/invalid answers	83	241

In addition, we conducted chi-squared tests for these results to check whether the attributions of the passengers and trains significantly affect the choice preference or not. The results of these tests imply that the choice preference is significantly different by some factors; for RS passengers, their age and number of the accompanying persons; and for NRS passengers, their age and OD.

From these tests, we decided to establish various second best choice behaviour models depending on these factors. As an example, in Fig. 4, we show the model for the NRSs passengers who are in their thirties travelling long-distance trip. NRS passengers have four alternatives as a second best choice as shown in Fig. 4. RS passengers have six alternatives, including ‘switch to the RS of earlier service’ and ‘switch to the NRS of earlier service’.

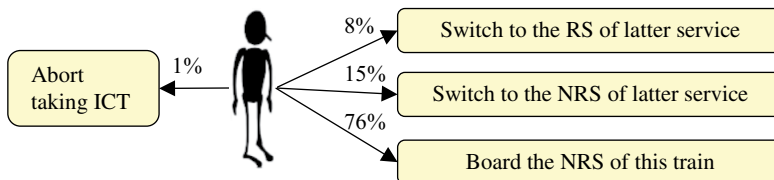


Figure 4: An example of second best choice behaviour model.

4 THE DEVELOPED SYSTEM AND THE OPTIMIZATION ALGORITHM

We developed the seating plans optimization system by using the models described above. The outline of the algorithm is shown in Fig. 5. This system is based on the genetic algorithm, whose solutions are combinations of seating plans of each target ICT and sections on the target date. In this system, to evaluate the solution, we have to consider the relationship between the potential demands and the solution, especially the effects on the demands caused by a certain seating plan. Therefore, this system involves the simulation of selling process of RS before departure and seating behaviour of NRS when a certain ICT arrives, which is called ‘S&S simulation’ hereafter.

The calculation flow of the system is described as below.

- (1) Estimate the potential OD of the target date and trains from the real demand data in the past, applying the potential passenger demands estimation models described in Section 2.
- (2) Generate various seating plans as initial solutions of the genetic algorithm.
- (3) For each of the solutions generated in (2), conduct S&S simulation using the potential OD estimated in (1). This simulation represents the selling process of RS from the launch day. The outline is as follows.
 - (3–1) We assume that the number of passengers who wanted to reserve RS every day follows the sale curve estimated by the revealed demand data. Based on this assumption, the system implements the simulation, where RSs are reserved gradually on a first-come-first-served basis. When a passenger cannot reserve the RS due to full occupancy in this simulation, he/she changes his/her choice in accordance with the second best choice behaviour model. We note that if he/she chooses to reserve another train’s RS but again faces full occupancy, he/she again changes his/her choice.
 - (3–2) When the train’s departure time comes in the simulation, the passengers of NRS, including those who decide to take NRS in (3–1), begin to try to take a seat. We assume that they make a queue on the platform and take a seat in random order if there is a NRS available. When the passenger cannot take a seat in the NRS due to full occupancy, he/she changes his/her choice in accordance with the second best choice behaviour model.

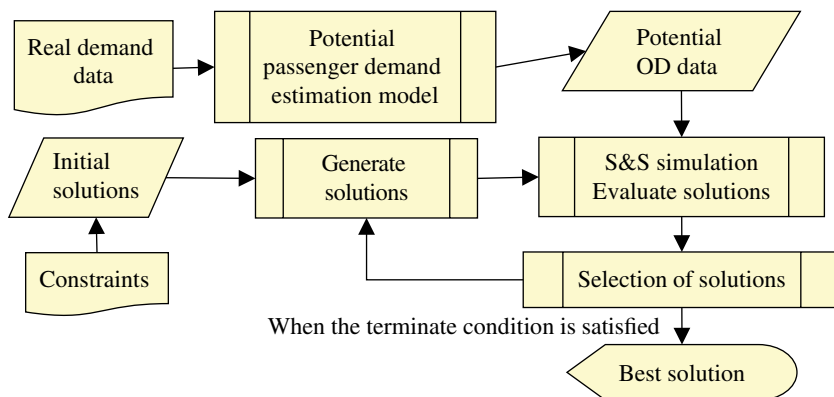


Figure 5: The outline of the seating plan optimization algorithm.

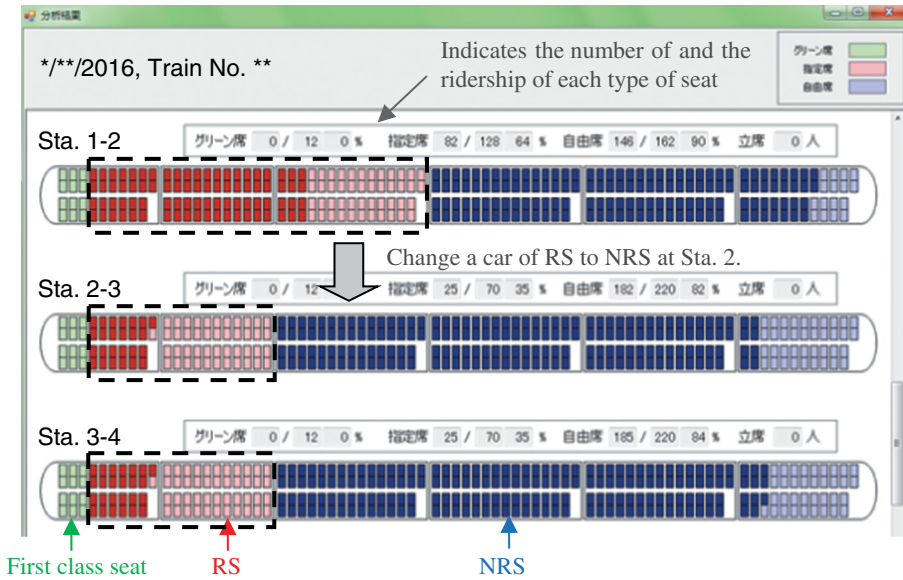


Figure 6: An example of seat map representing optimized seating plan.

- (4) Evaluate the solution considering the OD data which are the result of S&S simulation derived using procedure (3) and calculate the evaluated indices of each solution as an objective function. Based on the genetic algorithm, the system generates and selects the solutions, using a certain index which is set in advance. We can choose a certain objective function among four types of evaluation indices: the number of passengers per kilometre who take ICT but cannot seat (minimize), the difference between the LF of RS and NRS (minimize), the LF (maximize), and the total revenue (maximize).
- (5) When the condition for termination defined in advance (e.g. the limit of calculation time) is satisfied, this system terminates the continuous calculations and outputs the best solution.

The solution is displayed in the form of seat map, as shown in Fig. 6.

5 RESULT OF VALIDATION OF THE MODELS

To validate the accuracy of re-productivity of the models and S&S simulation, we estimated the OD data of a series of ICTs operated in two weeks by using the system and compared them with real demand data.

The results are shown in Figs 7 and 8. We converted the estimated OD data into the ridership on each section and compared them with the real visual count of ridership. The coefficient of correlation R is 0.787 for RS and 0.800 for NRS. These imply that the models and S&S simulation have an enough accuracy of re-productivity and these seem to be suitable for practical use.

6 RESULT OF A CASE STUDY

To evaluate the effect of this kind of flexible seating plans, we applied the seating plans optimization system to a series of ICTs operated in a certain line, as a case study. We treated two types of evaluation indices, as described in Table 6, and conducted the calculation for

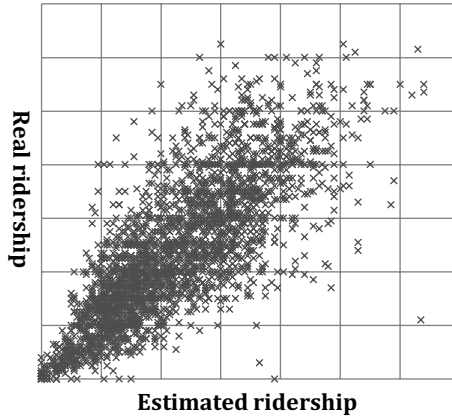


Figure 7: Validation result of accuracy (for RS).

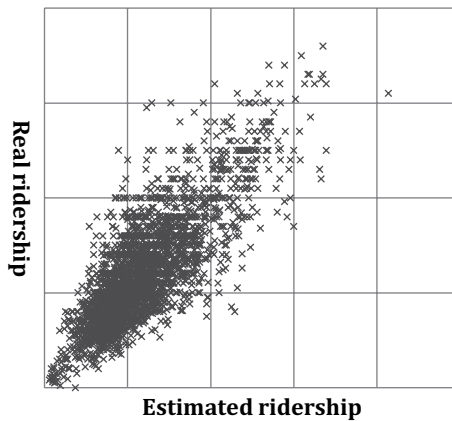


Figure 8: Validation result of accuracy (for NRS).

two days (a certain weekday and a holiday.) We also calculated the evaluation indices of the current seating plan and compared them with the indices of optimized seating plans. The results are shown in Table 6. These suggest that the optimization result possibly decreases the passengers per –kilometre (P-K) who take an ICT without taking a seat and increases the revenue.

Table 6: The effect of optimizing seating plans.

Date	Minimize P-K*		Maximize total revenue	
	P-K	Total revenue	P-K	Total revenue
A weekday	-74.1%**	+0.1%	-40.0%	+1.3%
A holiday	-60.1%	+0.1%	-9.2%	+1.1%

* P-K means 'passengers per kilometre who take an ICT but cannot take a seat'.

** Percentage in this table shows the results of comparison with the evaluation indices of current seating plan.

7 CONCLUSION

We established the potential passenger demands estimation model and the second best choice behaviour model and developed a seating plans optimization system for ICTs. This system can generate solutions which can improve both the passengers' convenience and the railway operators' revenue.

This system can be effectively utilized when the railway operators consider their ICTs' seating plan. While there are some further studies to be conducted about the train operation practically, e.g. how to show whether the seat is RS or NRS to the passengers, we aim to carry out this effective flexible seating plan for ICT operations with the cooperation of ICT operators.

REFERENCES

- [1] Minami, K., Terabe, S., Ieda, H. & Mizuguchi, M., Optimization of seat allocation on inter-city transport. *Proceedings of Infrastructure Planning*, Vol. 27, CD-ROM, 2003. (in Japanese)
- [2] Shibata, M., Terabe, S. & Uchiyama, H., An effectiveness of flexible seat class assignment on intercity express trains. *Transport Policy Studies' Review*, **13**(1), pp. 2–13, 2010. (in Japanese)
- [3] Shibata, M., Terabe, S. & Uchiyama, H., Optimization system for train set planning of inter-city rapid trains with the application of the flexible seat class assignment approach. *Proceedings of the 9th World Conference on Railway Research*, Paris, 2011.
- [4] Nakagawa, S., Shibata, M., Ozaki, N., Fukasawa, N. & Suzuki, T., A study on demand of limited express trains to optimize allocation of reserved/non-reserved seats. *Journal of Railway Engineering, Japan Society of Civil Engineers*, **18**, pp. 179–186, 2014. (in Japanese)