THE UNIVERSITY
OF BIRMINGHAM

METRO TRAFFIC OPTIMISATION ACCOUNTING FOR THE DISBENEFIT OF HALTING BETWEEN STATIONS

by

NAOKI UEDA

A thesis submitted to
The University of Birmingham
For the degree of
DOCTOR OF PHILOSOPHY

Department of Electronic, Electrical and Computer Engineering
School of Engineering
The University of Birmingham
January 2005

Synopsis

Computerised regulation for disturbed traffic in metro-type railways is proposed. Previous work has used optimisation techniques to minimise disbenefits to passengers, such as waiting time and journey time, in the objective function. The particular disbenefit of trains being halted between stations is introduced in this thesis, in combination with those already mentioned.

An effective method in real operations for preventing trains being halted between stations is to hold trains already at stations and to allow running trains to reach the next station when a particular train departure is delayed. The proposed algorithm uses this 'stop-all-trains-at-once' philosophy combined with optimisation ideas, in a sequentially structured approach.

A further consideration from real operations is the fact that it is not possible to predict precisely when the delayed train will re-start. Estimates of the re-starting time will improve as the delay increases, and the proposed scheme takes this into account.

Numerous simulations were undertaken to investigate the performance of the regulation algorithm. It is shown that the proposed regulation algorithm is effective in reducing the disbenefit to passengers from disturbed traffic for various characteristic metros with different passenger flows.

Acknowledgments

First of all, the author thanks Dr. Colin J. Goodman, who is the supervisor of this research. Dr. Goodman has given a lot of advice from a wide range of view, and helped the author get back on the right track when he was about to lose his way. The author's first language is not English, and thus Dr. Goodman had a lot of difficulties in supervising. The author is grateful particularly for this point.

Prof. Satoru Sone (Kogakuin University in Tokyo, Japan) has been a local supervisor in Japan. He has given a new idea in the research from his long experience of railway traffic management studies, and helped the author break through the walls in the research activity.

Dr. Ryo Takagi (research fellow at the University of Birmingham) has given a lot of suggestions, particularly on optimisation. He has also given a lot of advice on life in the U.K. and encouraged the author when he could not make progress.

Dr. Takafumi Koseki (associate professor of the University of Tokyo) has shown the latest achievements in his laboratory, and given many suggestions on railway traffic management studies through significant discussions.

The author thanks the above people, as well as all the other, including the colleagues of Central Japan Railway Company, where the author works, for his frequent absence from the office.

Table of contents

Not available in the digital copy of this thesis

List of figures and tables

[List of figures]

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 11

[List of tables]

List of abbreviations and symbols

List of abbreviations

[OD] Origin-destination pairs. This is used to express various kinds of passenger flow in a model line.

[OOMTS] The Object-Oriented Multi Train Simulator. Simulations performed in this research are based on this simulation program.

[Murata regulator] Simulation program developed by Murata, based on OOMTS, calculating passenger flow and discomfort values, with three kinds of discomfort criteria. Also referred to as "the Original Murata regulator".

[Murata plus halt regulator] The Murata regulator with four kinds of discomfort criteria including halt discomfort, with the same regulation algorithms as the Original Murata regulator.

[Enhanced Murata plus halt regulator] Regulation algorithms based on Murata plus halt regulator, including a crude calculation of the partial derivatives of halt discomfort and a prediction of resumption, with four kinds of discomfort criteria, including halt discomfort.

[Flat OD] OD type which has the same OD pair value for any combination of origin-destination stations.

[Centre OD] In this OD type, the flow of passengers from and to Station 14, which is regarded as the station in the city centre, is much larger than any other OD pair value.

[Sub OD] The opposite case to Centre OD. The flow of passengers from and to Station 1 is larger than any other OD pair values. Station 1 is regarded as the station in the city centre.

[Mid_1 OD] In this OD type, the flow of passengers from and to Station 8 is larger than other OD pair values.

[Mid_2 OD] This is similar to Mid_1 OD. The difference from Mid_1 is that only the flow of passengers to Station 8 is larger.

List of symbols

[D_w] Waiting discomfort

[X_w] Normalising point for waiting time (secs)

[E_w] Expected waiting time (secs)

[A_w] Actual waiting time (secs)

[D_t] Travelling discomfort

[X_t] Normalising point for travelling time (secs)

[E_t] Expected travelling time (secs)

[A_t] Actual travelling time (secs)

[D_c] Congestion discomfort

[X_c] Maximum congestion (passengers/train)

[N_c] Nominal congestion (passengers/train)

[A_c] Actual congestion (passengers/train)

[$T^i_{a_k}$] Arrival time of train i at station k

[$T^i_{d_k}$] Departure time of train i from station k

[B^i_k] The set of passengers who board train i at station k

[A^i_k] The set of passengers who alight from train i at station k

[C^i_k] The set of passengers who can board train i at station k but originally could not.

[i_b] The first train that departs from the station where the initial disturbance happens on its original schedule after the initially disturbed train.

[i_d] The initially disturbed train

[k_d] The station where the initial disturbance happens. Sometimes, this is called the disturbance site.

[$d_{i,k}$] Modified departure time of Train i from Station k (secs)

[$d_{i_d,k}$] Regulated departure time of delayed Train i_d from Station k (secs)

[f_k] Minimum run time from Station k to Station $k+1$ (secs)

[$S_j(\min)$] Minimum dwell time at Station j (secs)

[t_s] Service headway (secs)

[t_m] Minimum headway of the line (secs)

[N_{tr}] The number of trains that are running on the line (or on a considered section) at a certain instant

[N_{pl}] The number of platforms

[D_h] Halt discomfort

[X_h] Normalising point for halt time

[E_h] Expected halt time

[A_h] Actual halt time

[P(dist)] The number of passengers on a train running towards the station where disturbance in a normal operation (without any disturbance)

[P(max)] The maximum number of passengers on a train in a normal operation (without any disturbance)

1 Introduction

1.1 The nature of metros

A typical example of a "metro-type" railway is an underground railway in a big city such as the London Underground, the New York City Subway, and Tokyo Metro. The main characteristics of a "metro-type" railway are as follows.

(1) All trains stop at all stations. (With a few exceptions.)

(2) Inter-station distances are relatively short. A typical value is 1 to 2 stations per km.

(3) A "metro-type" railway generally has to handle high demand with limited facilities. Thus, it tends to have the problem of congestion, which becomes more serious when traffic is disturbed, as will be explained later.

This research deals with a "metro-type" railway. In other words, suburban railways and high-speed inter-city rail services are not the target of this research.

1.2 Need for traffic regulation and recovery from disturbance

Railways play an important role as part of public transport in the world, especially in developed countries. The railway system has some advantages compared with other forms of public transport, such as larger transport capacity, higher speed, and less emission to the environment.

On the other hand, there are a few disadvantages in the railway system, one of which is inflexibility of the facilities. Railway facilities are generally expensive to construct, as well as hard to change once they have been constructed, and the trains can only run where rails have been laid. These features lead to the fact that, if railways are operated exactly as scheduled, there are no problems; however, it is difficult to handle certain kinds of situations (e.g., disturbances) other than by original design of the infrastructure.

Thus, in order to handle disturbed traffic in railways successfully, it is important to improve the utilisation of fixed and limited facilities, in other words, to improve operational decisions. In railways, operational decisions include holding trains at stations, making trains depart from stations earlier, and so on. In this thesis, these methods of operations are called "regulation". In short, traffic regulation is necessary and important to realise flexible railways, able to respond to traffic disturbances.

Especially in metro-type railways like the London Underground, quick regulation is crucially important, since they generally have to handle high demand with limited facilities, in other words, they do not have much operating margin, even compared with other kinds of railways such as suburban lines. In this situation, disturbances may be magnified; for example, a certain amount of delay of a certain train at a certain station may cause accumulation of passengers at stations beyond this station, thus this train cannot handle all of these passengers, which may lead to more passenger load on the following trains and more delays of these trains, and so on. Under such conditions, railway systems, especially metro-type ones, may well be unstable. As Van Breusegem, et al have shown[40], metro-type railways have an intrinsic instability.

Due to this nature of metro-type railways and for other reasons, most metro-type railways are controlled by a centralised operation system, thus human operators (or computers in some cases) can direct each train to depart a station earlier or later, and to run faster or slower. In Barcelona, Spain, for example, a regulation system has been adopted in the control centre[18][19]. In metro-type railways, the instructions that operators (or computers) give to each train have a large effect on the railway traffic conditions.

In this research, the regulation method applied is limited to changing departure and arrival times of trains from/at stations. Other methods, such as changing destinations of trains, closing a part of the line and injecting extra trains are not allowed in this research. These methods are usually applied in the cases of large scale disturbances; e.g., several-hour stoppages of the service due to accidents; whereas the target of this research is a smaller disturbance, as will be stated in Chapter 4. Even so, significant discomfort for passengers will emerge if no regulation is performed. Only changing departure and arrival times will be revealed to be effective in avoiding passenger discomfort.

1.3 Use of optimisation ideas

Traditionally, railway traffic regulation is performed by experienced human operators, since adequate knowledge of the railway system and much experience of regulation is needed. This is because there are many parameters in railway traffic regulation, such as at which station a train is stuck, when and how long it is stuck, which train should be held at which station, where the depot is located, and so on. Due to this background, railway traffic regulation is difficult to computerise.

Nowadays, some traffic control systems have been realised and implemented on real railway systems. However, they are still based on the knowledge of human operators, in other words, they are heuristic. Alternatively, they only present an apparently "good" solution to the human operator, thus it is still a human operator's responsibility to choose whether the solution should be adopted.

Here, it should be noted that what is a "good" solution to a disturbed traffic situation is an important issue. In most cases, the criteria for a good solution (or a good traffic status) are determined from the railway operator's point of view. For example, the shorter the delays of trains are, the better the traffic is. That is to say, the smaller the deviations from the scheduled timetable are, the better the traffic is. Another example is that the more regular the intervals of trains, the better the traffic.

Though it may seem that these objectives (shorter delays, smaller deviations, more regular intervals, and so on) are good for passengers as well as operators, they are not necessarily so in metro-type railways, where the headway of trains is fairly short (typically 2 to 5 minutes). If the headway is short enough, passengers come to stations randomly (without checking timetables beforehand). That means, even if every train is delayed 3 minutes with respect to the schedule, it does not matter greatly for passengers, as long as trains arrive regularly. Thus, a delay affecting all trains is not a problem from the passenger perspective in a metro-type railway.

Delays of trains may matter in cases such as on a regional or a rural line, and at interchanges between lines. Even so, it is likely that, at an interchange of fairly frequent lines (London Underground interchanges, for example), delays of trains on one line do not seriously affect the others. This is again due to the reason that passengers would not check a timetable before they come to a station, thus it does not matter if they cannot take exactly the same "physical" train that they would have caught without delays.

In extreme cases, even if a train is stuck at a station, a human operator would tend to leave the trains ahead of the disturbed train to run according to the original schedule, since he/she would think that the more trains that run according to the original schedule, the better the traffic. As a result, the delayed train will have to absorb more and more passengers as it proceeds, since the preceding train will have passed through the stations a longer time ahead than usual. In this situation, a disturbance may not settle, but indeed grow, as time goes on.

As will emerge later in this thesis, it is one of the characteristics of this research that it emphasises the passenger perspective rather than the operator's point of view.

Once criteria for the "goodness" of railway traffic have been carefully determined, a theory of optimisation can be applied. Recent progress in both theory and hardware (that is, novel techniques of optimisation such as artificial intelligence, and computers with higher performance) has enabled the implementation of calculation algorithms for railway traffic regulation that were previously impossible.

This approach, using optimisation ideas, rather than depending on solely heuristic rules, is the other characteristic of this research.

1.4 Aim of this research

This research project aims at developing metro-type railway traffic regulation for disturbed conditions, thus contributing to the realisation of metro-type railways which are robust to disturbances.

There are two important points to bear in mind. First, the regulation method should include an element of optimisation. In other words, the regulation method does not depend solely on heuristic rules. Not only classical methods, but also some recently developed optimisation ideas, such as artificial intelligence techniques, will be examined. The latter ideas are sometimes called "meta-heuristic", since they exploit some heuristic ideas in order to reach the optimum quickly (not exactly the optimum, but sufficiently near the optimum). These ideas should be a candidate for regulation methods, since, due to the nature of metro-type railway traffic, a regulator has to work in real time.

Methods of utilising off-line databases will also be reviewed. "Off-line database" means a database created by off-line calculation of optimal results under selected conditions. This may contribute to reducing computation effort, since on-line calculation of optimisation can be avoided by utilising results from the off-line database. Methods for extracting the necessary information from the database quickly and efficiently must be developed.

Second, the regulation proposed in this thesis is performed from the passenger perspective, rather than the railway operators' point of view, as already stated in Section 1.2. This is realised by incorporating elements related to the passenger perspective in the objective criteria of optimisation.

In this thesis, the following four categories of passenger discomfort are used, which are to be explained fully later in this thesis. They are waiting discomfort (penalty for increase of waiting time for passengers at stations), travelling discomfort (penalty for increase of travelling time for passengers to their destinations), congestion discomfort (penalty for train congestion above a threshold) and halt discomfort (penalty for passengers halted between stations). The first three kinds of discomfort have already been introduced in previous work, while the last one is introduced within this research project.

Halt discomfort is important in metro-type railways, since passengers who are stuck in a train between stations naturally feel frustrated. Also, quite a few metro-type railways are underground lines, which means that the halted passengers would feel much more frustrated and even alarmed in a tunnel, as can be seen in some incidents on London Underground. In spite of these considerations, halt discomfort has not yet been discussed in previous research work.

In one sentence, the aim of this research is to develop a real time regulator for metro-type railways, which minimises an objective function based on the passenger perspective, including halt discomfort.

1.5 Structure of this thesis

The structure of this thesis is as follows.

Chapter 2 explains the simulation methods used in this research. All of the simulations are carried out based on the simulation methods stated in this chapter, thus it is important to explain the simulation methods just after the introduction.

Chapter 3 is devoted to giving a review of existing work. In this chapter, "the Original Murata" regulation, which plays a basic role in establishing a new regulation method in this research, is explained.

In performing a simulation, it is important to choose appropriate conditions that satisfactorily describe real circumstances. In Chapter 4, simulation conditions are explained with a view to choosing suitable parameters.

After choosing appropriate parameters, Original Murata regulation is simulated under these conditions. Results of the simulation are explained in Chapter 5. In this chapter, problems with the Original Murata regulation are revealed, and the causes of the problems are examined.

Following the discussion in Chapter 5, a new discomfort criterion called "halt discomfort" is introduced in Chapter 6. A re-formulation of the Murata regulator (later referred to as "Murata plus halt") is performed.

Chapter 7 gives an observation from real railway operation, that is, railway operators tend to stop all trains at the nearest station immediately after traffic is disturbed. This method is revealed to be fairly effective in preventing "halt discomfort" from increasing suddenly, but not effective in reducing the other three discomfort criteria. Nevertheless, a combination of the Murata regulation and this idea seems to provide a much better regulation.

After the discussion in Chapter 5 and Chapter 7, a new "phase-by-phase" regulation is introduced in Chapter 8. Prediction of the restarting of the disturbed train plays an important role in the new form of regulation.

Chapters 9-11 explain results of the new regulation from various perspectives, which are the relationship between the numbers of trains and stations in a line, where the disturbance happens, and the variation of the restarting prediction over time.

Chapter 12 summarises the results of this research and gives suggestions for further study.

2 Simulation methods

2.1 Use of OOMTS – Object-oriented Multi-train Simulator

In order to test traffic regulators, a model of a real railway is necessary. In this research, a simulation program written in C++, called "Object-oriented Multi-train Simulator" (OOMTS) has been used to simulate the traffic status under various conditions.

OOMTS was established by Siu[36], who performed his research at the University of Birmingham. As can be seen from its name, it can handle multiple trains in one simulation model, however, the original OOMTS simulation programs did not have any functions that represented passenger flow in the railway model. These functions have been developed and added to the simulation programs later by Goodman and Murata in order to evaluate traffic status from the passenger perspective and to establish regulation methods[23].

In this chapter, the abstract modelling of the line is described, rather than the detail of the simulation programs.

2.2 Model of a real railway system

OOMTS can handle various types of railways, since it reads the data files describing the topology of the line, and then constructs a model. In this thesis, the model is based on the Hong Kong Island Line.

2.2.1 Line topology

The Hong Kong Island Line has 14 stations, all of which are equipped with two passenger platforms. The layout of the line is shown in Figure 2.1. At one terminus, expressed as Stn. 1 in the figure, there is a head-shunt; at the other terminus, expressed as Stn. 14 in the figure, trains turn around by crossings.

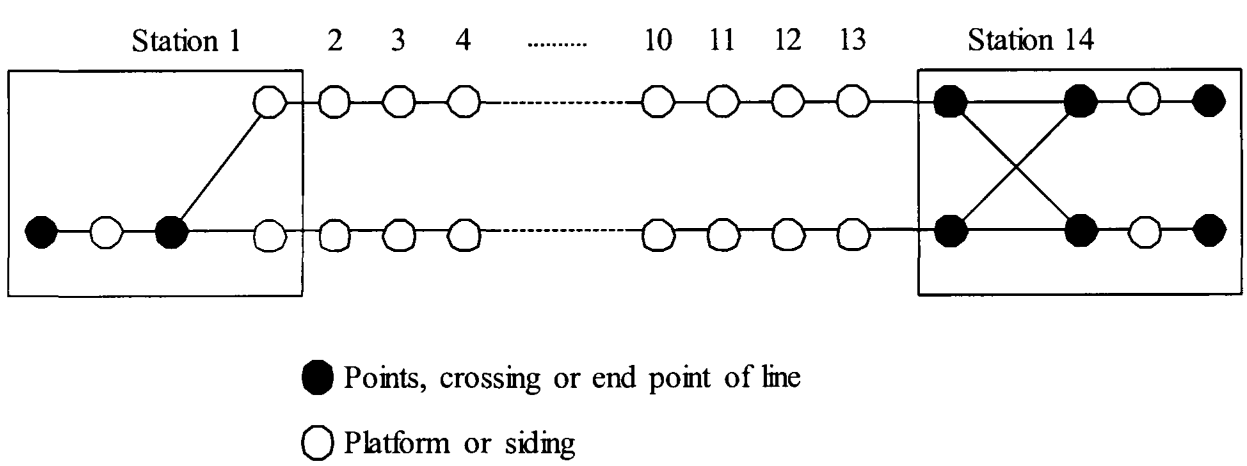

Figure 2.1: Line topology

2.2.2 Signalling system

In this thesis, the simulation model adopts a moving-block signalling system, due to its ease of simulation, although the real railway uses a fixed-block signalling system.

2.2.3 Rolling stock

The main features of the rolling stock are:

Length: 181 m,

Tare weight (vehicle weight when empty): 296×10^3 kg,

Nominal passenger load: 1600 persons,

Maximum passenger load: 2000 persons.

2.2.4 Service

Trains run at a regular interval on the line. A typical interval is 90 s. All of the trains stop at all stations, which means that, with a 90 s interval, there are 31 trains serving four round trips from station 1 to 14 and back. Travelling time from station 1 to 14 is about 22 minutes, thus a simulated time for all trains to finish their services is about 4 hours.

2.2.5 Passenger flow

The passenger flow from one station to another is defined as an origin-destination pair value, or OD for short. For example, 18 passengers per minute travel from station 1 to 7, while 3 passengers per minute travel from station 7 to 14, etc. Changing the set of OD pair values enables the simulation of various kinds of railways, such as underground lines in a city centre, city-to-suburbs lines, etc.

The passenger exchange rate at a station is defined as shown in Figure 2.2. This implies that, if there are more passengers on a train than the nominal value, the more the number of passengers exceeds this value, the slower the passengers alight from and board the train. This reflects congestion and conflict in the door openings of the train.

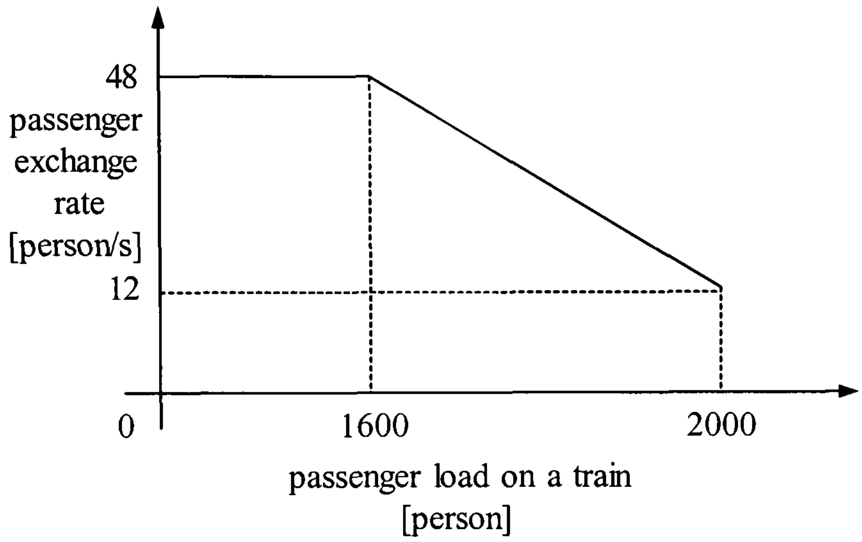

Figure 2.2: Passenger exchange rate variation with load on trains

2.3 Implementation of regulators

In the simulator, trains are operated according to the timetable stored in the simulator. The commands (arrival or departure time of trains) are issued on the basis of the information in the timetable. The regulator modifies the commands, obtains state information from the line model, calculates adjustment values that should be added to the command, and issues the commands to the trains. Figure 2.3 illustrates how it works.

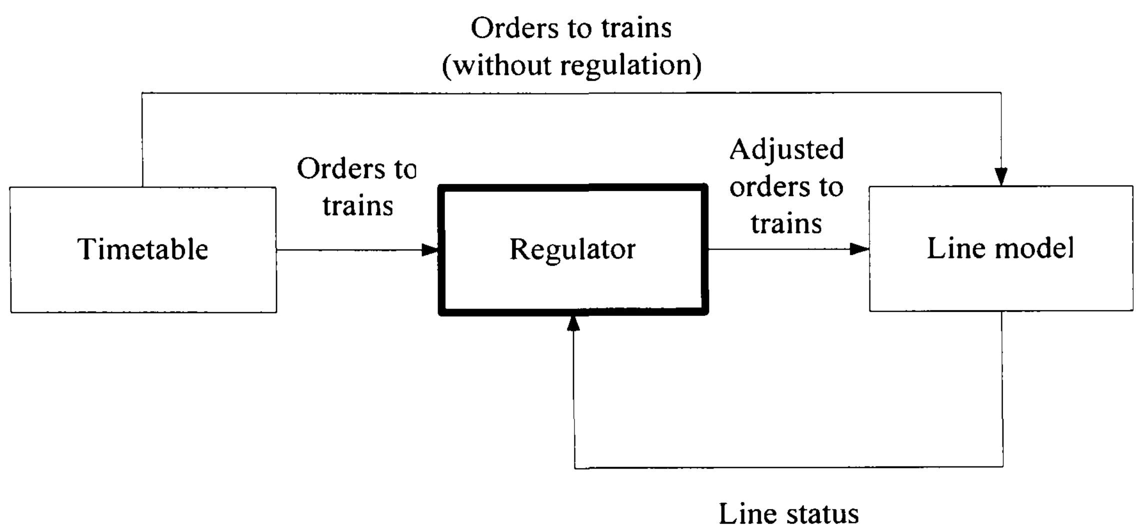

Figure 2.3: Regulators intervention

3 Review of existing work

3.1 Historical summary

Most previous work on regulators has looked at recovery to a pre-planned schedule or to a regular headway. The majority of regulation schemes has used classical optimisation techniques to find the best recovery strategy or simple heuristic control algorithms (for example, Araya, S. and Sone, S.[1]). Thia[38] has used a fuzzy expert system approach to construct an on-line rescheduling system for Singapore Mass Transit System (MRT), though it is still heuristic.

More recently, work in the research group at Birmingham has focused on the passenger experience during delays[23]. Factors, such as waiting time, travelling time and congestion, are built in to an overall evaluation function. The optimisation itself has still been achieved essentially by classical gradient projection methods[17]. This method will be explained in section 3.4.

3.2 Classical optimisation and a heuristic method

The objective functions and the methods of regulation are explained in the following sections. First, the difference between two kinds of approach to railway traffic regulation is examined. These are classical optimisation and heuristic methods. The former one needs, by its nature, an objective function on which the regulation is based, thus, an objective function is introduced on the basis of the discussion of passengers' discomfort described later in section 3.3. Finally, examples of each approach are described. These are the regulation methods invented by Murata[23] and Chua[15] respectively, both of whom were researchers in the research group at Birmingham. In this thesis, these regulators are called "the Original Murata regulator" (or, simply "Original Murata") and "Chua regulator" respectively.

3.2.1 Classical optimisation

Classical optimisation is based on an explicit objective function. With gradient-based techniques, this function is differentiated with respect to the parameters so that a set of parameters can be found which maximise (or minimise) the objective function. One advantage of the gradient method is that it is mathematically rigorous, but it does have some disadvantages, one of which is that it is sometimes unusable due to the nature of the objective function. For example, if the objective function is discontinuous, it cannot be differentiated at the point of discontinuity, thus this method is not applicable here. In this way, when classical optimisation is implemented, constraints arising from mathematical strictness should be carefully examined.

Another disadvantage of classical optimisation is that it may need such long computation times that it cannot be applied to a real time regulator. The reasons for long computation times include the fact that, when differentiating the objective function, a predictive simulation is necessary, since the phenomenon of railway traffic is time varying. A prediction of railway traffic tends to take a long time, since railway traffic has many parameters and a complicated interactive structure. This is true even for a metro because of interactions between trains.

3.2.2 Heuristic methods

Heuristic methods are another possible approach to regulation, in which a regulator issues orders according to heuristic rules implemented in the regulation algorithm. Traditionally, a human operator, sitting in front of a panel installed in a control centre, does the regulation. He or she keeps an eye on the panel displaying the status of trains running on the line and, when a disturbance happens, he or she judges the situation and decides the instructions to the trains so that the traffic disturbance will disappear as soon as possible.

A heuristic method is based on rules on which a human operator would act, for example, to make the intervals between trains equal. In this sense, a heuristic method is, in other words, part of an expert system.

The advantage of this method is its simplicity, since it is based on a human operator's experience and/or instinct and is thus not complicated. On the other hand, a heuristic method has the disadvantage that it does not guarantee optimality.

3.3 Objective functions

Before an explanation of the examples of classical optimisation, it is necessary to define the objective functions to be optimised (maximised or minimised). As is already stated in section 3.1, previous work has first focused on three kinds of discomfort that passengers may feel when traffic is disturbed. They are waiting discomfort, travelling discomfort, and congestion discomfort.

3.3.1 Waiting discomfort

Waiting discomfort D_w is defined by equation (3.1),

$$D_w = \frac{1}{(X_w - E_w)^2}(\max[0, A_w - E_w])^2 \qquad (3.1)$$

where

X_w : normalising point for waiting time (secs),

E_w : expected waiting time (secs),

A_w : actual waiting time (secs).

X_w is a linear function of E_w, which means an acceptable waiting time for all passengers, and plays a role as a weighting factor for D_w. The results shown in this thesis are calculated using the relationship between X_w and E_w described in Figure 3.1.

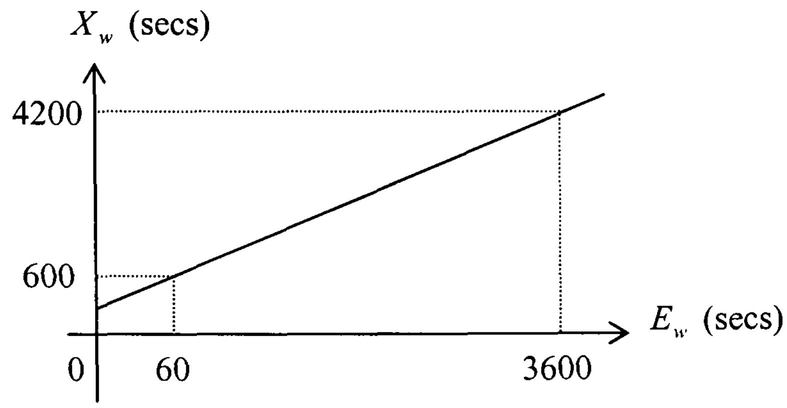

Figure 3.1: Normalising point and expected value of waiting time

Here, the expected waiting time E_w is not defined in a statistical sense. Strictly, it is the expected maximum waiting time, which may reasonably be set equal to the headway of the line in a metro-type railway, where there are frequent train services, since passengers would come to stations without checking a timetable. Thus, all the cases in this thesis are calculated under this condition.

Note that $D_w = 0$ when $A_w \leq E_w$. This means that passengers do not feel discomfort when the actual waiting time is equal to or less than the expected value.

3.3.2 Travelling discomfort

Travelling discomfort D_t is defined by equation (3.2),

$$D_t = \frac{1}{(X_t - E_t)^2}(\max[0, A_t - E_t])^2 \tag{3.2}$$

where

X_t: normalising point for travelling time (secs),

E_t: expected travelling time (secs),

A_t: actual travelling time (secs).

The form of the equation is the same as equation (3.1), with the only difference being the change in the suffix of the parameters from w to t. X_t is again a linear function of E_t, which defines an acceptable travelling time for general passengers. The results shown in this thesis are calculated based on the relationship of X_t and E_t described in Figure 3.2.

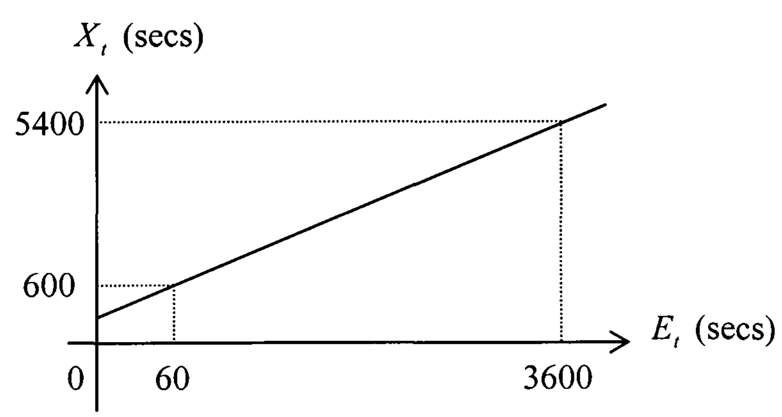

Figure 3.2: Normalising point and expected value of travelling time

Note again that $D_t = 0$ when $A_t \le E_t$. This means that passengers do not feel discomfort when the actual travelling time is equal to or less than the expected value.

3.3.3 Congestion discomfort

Congestion discomfort D_c is defined as equation (3.3),

$$D_c = \frac{1}{(X_c - N_c)^2}(\max[0, A_c - N_c])^2 \tag{3.3}$$

where

X_c : maximum congestion (passengers/train) (in other words, crush load of train),

N_c : nominal congestion (passengers/train) (in other words, discomfort threshold),

A_c : actual congestion (passengers/train).

Again, the form of the equation is the same as equations (3.1) and (3.2), though in this case, X_c and N_c are independent, (needless to say, X_c must be larger than N_c.). As is easily understood from equation (3.3), a nominal congestion N_c represents a threshold beyond which passengers begin to feel discomfort by congestion in a train. The results shown in this thesis are calculated at $X_c = 2000$ (passengers/train) and $N_c = 1600$ (passengers/train).

3.3.4 Integration of discomfort values

All the discomfort terms stated above are summed up as shown in equation (3.4).

$$F_{obj} = D_w + D_t + D_c \tag{3.4}$$

The Original Murata regulator, which is explained in the next subsection as an example of classical optimisation, uses F_{obj} as an objective function, which is to be minimised, not maximised, due to its nature.

3.4 The Original Murata regulator – an example of classical optimisation

The Original Murata regulator uses F_{obj} in equation (3.4) as an objective function, which needs to be minimised by changing the decision variables, which are, in this problem, all arrival/departure times of trains at/from stations. First, the regulator predicts the train movements and passenger flows when arrival and departure times of trains change slightly. Secondly, it calculates the partial derivatives of F_{obj}. Thirdly, it changes the arrival and departure times of the trains bases on the results of the partial derivatives of F_{obj}. Finally, it checks the constraints of arrival and departure times of trains, since there are several operational constraints, such as minimum run time between adjacent stations, minimum headway at every station, and minimum dwell time at stations to allow all passengers to exchange (or passengers to exchange until a train is full).

In the Original Murata regulator, the following four techniques have been adopted to reduce computational effort.

(1) Gradient calculation

This is adopted to reduce the computation time for the partial derivatives. The gradient calculation can be simplified by decomposing the partial derivatives of F_{obj} into the sum of the partial derivatives of the passenger discomfort functions. It can be supposed that calculating one discomfort function requires less time than calculating the all partial derivatives, if enough data is already prepared. Thus, the gradient calculation has two stages. In the first stage, a simulation is done to prepare data. In the second stage, each partial derivative is calculated.

It is not necessary to calculate partial derivatives with respect to the variables that do not affect the partial derivatives. Consider the situation in a railway when an arrival or a departure time change would contribute to identifying which calculation is necessary. When an arrival time is delayed, the waiting time of passengers at that station increases, the travelling time of passengers alighting at that station increases, and the duration of congestion increases. When a departure time is delayed, the waiting time changes for passengers who can board that train but originally could not. The travelling time of passengers and the duration of congestion are not determined at the departure of the train. Thus, the partial derivatives can be expressed as follows:

$$\frac{\partial F_{obj}}{\partial T_{a_k}{}^{i}} = \sum_{j \in B_k{}^{i}} \frac{\partial D_{w_j}}{\partial T_{a_k}{}^{i}} + \sum_{j \in A_k{}^{i}} \frac{\partial D_{t_j}}{\partial T_{a_k}{}^{i}} + \sum_{j \in A_k{}^{i}} \frac{\partial D_{c_j}}{\partial T_{a_k}{}^{i}} \tag{3.5}$$

$$\frac{\partial F_{obj}}{\partial T_{d_k}{}^{i}} = \sum_{j \in C_k{}^{i}} \frac{\partial D_{w_j}}{\partial T_{d_k}{}^{i}} \tag{3.6}$$

where

$T_{a_k}^{i}$: arrival time of train i at station k

$T_{d_k}^i$: departure time of train i from station k

B_k^i : the set of passengers who board train i at station k

A_k^i : the set of passengers who alight from train i at station k

C_k^i : the set of passengers who can board train i at station k but originally could not.

(2) Gradient projection

The technique of gradient projection is useful to reduce computation time in this highly constrained problem. Consider the example of two successive trains at a station. Figure 3.3 shows the distance-time diagram, in which partial derivatives are shown in heavy arrows. In this example, the left heavy arrow is longer, which means that the value of the partial derivative for delaying the arrival of Train 1 is larger than that for advancing the departure of Train 2. Here, if a time interval between the departure time of Train 1 and the arrival time of Train 2 is a minimum value, then the solution in the diagram which involves moving each event towards the other is not feasible, since it violates the constraint. In this case, the sum of the two partial derivatives is calculated and used to decide which direction the two events should be moved, as is shown in Figure 3.4.

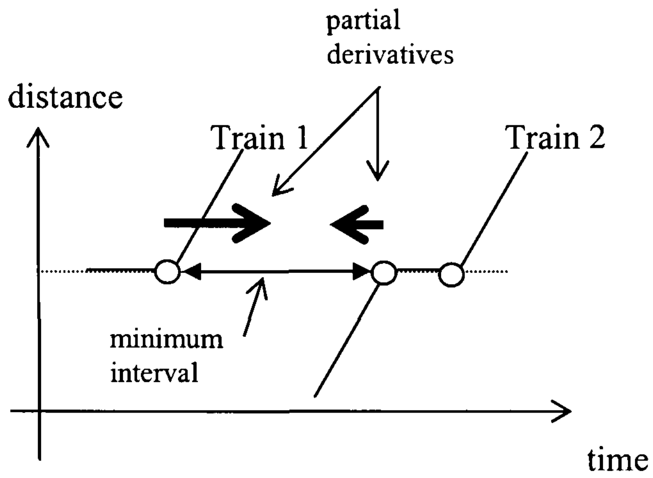

Figure 3.3: Conflict of partial derivatives with a constraint

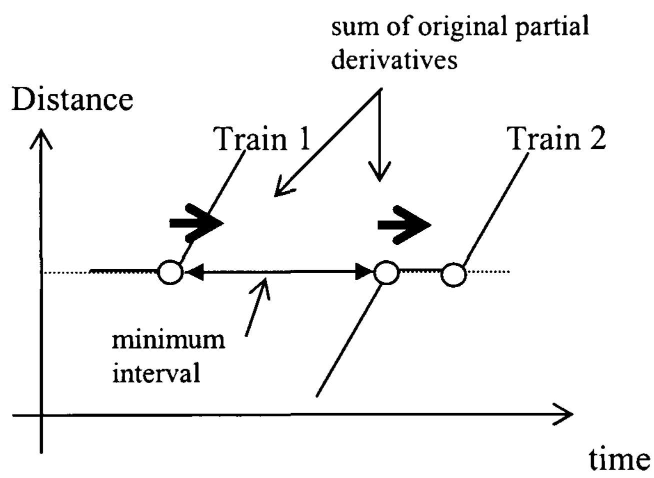

Figure 3.4: Resolution of partial derivatives conflict

By using the sum of partial derivatives as a partial derivative for each event, the gradient projection is obtained. Taking into consideration the characteristic of the problem that constraints have time-ordered and thus causal relationships, the following two-step algorithm can be obtained. First, the partial derivative is accumulated from the latest event to the earliest one.

Secondly, the accumulated value is propagated from the earliest event to the latest one.

(3) Sliding window

Railway traffic disturbances happen irregularly and generally settle out over a period of one hour or so. Thus, it is possible to calculate the optimal solution for a fixed time into the future of an hour or so, but repeat the calculation at much shorter intervals. This technique is referred to as a fixed-length optimisation window with receding horizon, or sliding window for short. It is also effective in reducing the effect of disturbances which arrive after the optimal solution has been calculated, and thus reducing the effect of modelling inaccuracies.

(4) Recursive optimisation

This means using the results of the previous optimisation cycle as the initial condition for the next optimisation cycle. This is based on the assumption that the previous solution is nearer to the solution being sought, an assumption which is likely to be true since disturbances in a railway evolve slowly.

3.5 Chua regulator – an example of a heuristic approach

Another method previously researched at Birmingham is described here. This is called the Chua regulator, and is an example of a heuristic approach. This is called heuristic because the rules adopted for this regulation are based on the experience and/or instinct of human operators[15].

The assumptions of this regulation method are:

(1) The disturbed (delayed) train runs as fast as possible and stops at stations for as short a period as possible, until it recovers to the original timetable.

(2) The trains following the disturbed (delayed) train stop at the station where the disturbance happened for the minimum period, and depart from that station at the minimum headway interval.

Other trains are also regulated in the following way, according to their position. Figure 3.5 is a simplified diagram, in which normal station stops are not shown.

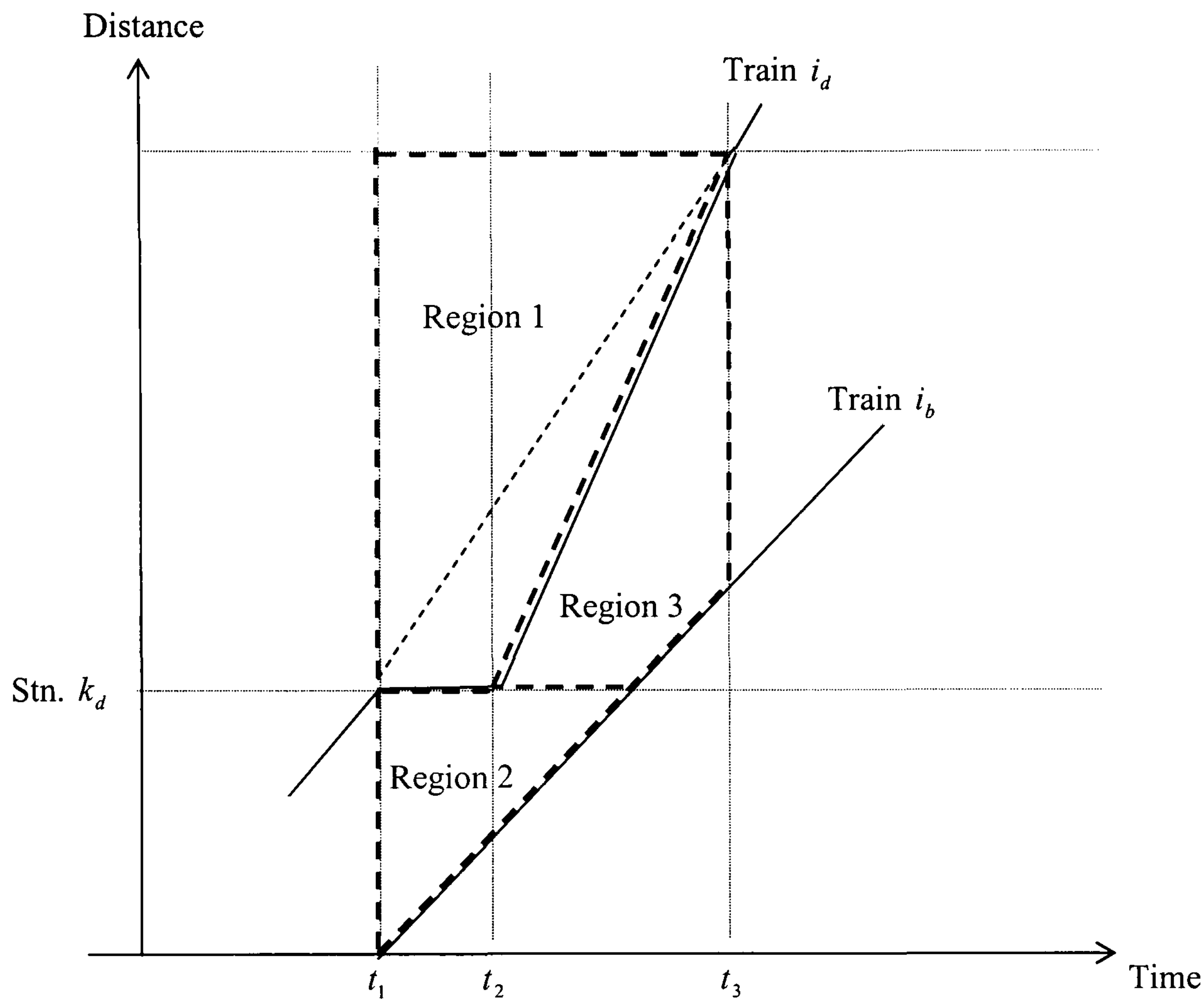

Figure 3.5: Regulation diagram

Train i_d is delayed at Station k_d from t_1 to t_2, and returns to its original schedule at t_3. Train i_b is the first train after Train i_d that departs from Station k_d on its original schedule.

Generally, it cannot be predicted how long the disturbance will last, thus t_2 in Figure 5 is only known immediately before Train i_d actually departs from Station k_d. In this regulator, Train i_d is assumed to depart from Station k_d at the start of the regulation activity that is performed every ten seconds.

Region 1:

In this region, delays are deliberately imposed on selected trains running ahead of the delayed trains while they are at stations. The criteria to select the trains to be delayed are:

(a) the train is in a station at the time of delay,

(b) the train had arrived at the station according to the original or previously re-scheduled time, and

(c) if the train left the station at the previously re-scheduled time, there would be an excessive time before the next train arrived.

The modified departure time of Train i from Station k is:

$$d_{i,k} = d_{i_d,k} - (i_d - i)t_s \tag{3.7}$$

where

$d_{i,k}$: modified departure time of Train i from Station k (secs),

$d_{i_d,k}$: regulated departure time of delayed Train i_d from Station k (secs), i.e.

$$d_{i_d,k} = t_2 + \sum_{j=k_d+1}^{k} (f_{j-1} + S_j(\min))$$

where

f_k: minimum run time from Station k to Station $k+1$ (secs),

$S_j(\min)$: minimum dwell time at Station j (secs),

i_d : delayed train number,

t_s : service headway (secs).

Region 2:

In this region, delays are also deliberately imposed on selected following trains at stations. The criteria to select the trains to be delayed are:

(a) the train is in station at the time of delay,

(b) the train had arrived at the station according to the previously re-scheduled time, and

(c) if the train left that station at the previously re-scheduled time, it would have to stop before arriving at the next station, due to the signalling constraints.

The modified departure time of Train i from Station k is:

$$d_{i,k} = t_2 + (i - i_d)t_m - S_{k_d}(\min) - f_k - \sum_{j=k+1}^{k_d - 1}(f_j + S_j(\min)) \qquad (3.8)$$

where

t_m : minimum headway of the line (secs).

Region 3:

The purpose of the regulation in this region is to enable trains to return to their nominal schedule. The modified departure time of Train i from Station k is:

$$d_{i,k} = t_2 + (i - i_d)t_m + \sum_{j=k_d+1}^{k}(f_{j-1} + S_j(\min)) \qquad (3.9)$$

4 Simulation conditions --- suitable choices of parameters

4.1 Introduction

This chapter is devoted to an explanation of the conditions under which the simulations performed in this research are based. In order to perform simulations that relate well to real circumstances and that contribute to a good understanding of the target of the research, it is important to set appropriate parameters in the simulations so that the simulation conditions agree broadly with those found in the real world.

The parameters discussed here are the ratio of the number of trains and the number of stations (or, precisely, platforms), OD (origin-destination) types, the disturbance site, and the duration of the disturbance.

4.2 Ratio of the number of trains to the number of stations

The first parameter is the ratio of the number of trains to the number of stations. In other words, this parameter explains the density of the line in terms of trains running on it.

This research introduces a new criterion that relates to passengers' discomfort when they are halted in a train that stops between stations. This means that the density of trains on a line can affect this discomfort, since, where there are significantly less trains than available platforms, trains can dwell in a platform when a disturbance happens. By this simple method, the discomfort can be drastically reduced. On the other hand, if there are significantly more trains than platforms, it is unavoidable that some trains cannot dwell in platforms and thus have to stop between stations. In this situation, it is obvious that any regulation is unable to reduce the stopped passengers' discomfort, at least initially.

According to the objective of this research, that is, to establish a regulation method for busy traffic in a metro-type railway when the operation is disturbed, the density of trains in a line should be within the typical range found for the peak time operation of real metro-type railways. When N_{tr} is the number of trains that are running over the line as a whole (or over a considered section) at a certain instance, and N_{pl} is the number of platforms associated with the line (or a considered section), the ratio N_{tr} / N_{pl} represents the density of trains, in this particular sense. Table 4.1 shows the values of N_{tr} / N_{pl} for the peak time operation of some real railways.

Table 4.1: N_{tr} / N_{pl} for the peak time operation of some real railways

Country	City	Line	N_{tr} / N_{pl}	Note
U.K.	London	Central Line	0.73	
		Piccadilly Line	0.73	
		Victoria Line	1.15	
France	Paris	M1	0.80	
		RER-A	1.00	La Defense - Nation
Japan	Tokyo	Ginza Line	0.97	
		Chuo Rapid Line	1.28	Nakano - Tokyo
China	Hong Kong	Island Line	1.06	

As can be observed in Table 4.1, on some lines, there are fewer trains running simultaneously than platforms, even in peak hours. On the other hand, an express line like the Chuo Rapid Line in Tokyo exhibits a ratio of $N_{tr} / N_{pl} > 1$, since the typical between-station distance of a rapid line is longer than that of a slow line, while the frequency of train services on an express line is the same as that in a slow line.

In the simulation tool used in this research, the headway of the simulated line can be easily changed by editing some files that determine the simulation environment. Thus, in the simulations performed in this research, the headway is changed, instead of the number of stations, in order to change the ratio N_{tr} / N_{pl}. In the simulations, three values of headway are adopted, namely, 75 s, 90 s, and 120 s. Table 4.2 shows the correspondence between the headway and the ratio N_{tr} / N_{pl}.

Table 4.2: Correspondence between headway and N_{tr} / N_{pl}.

Headway	N_{tr} / N_{pl}
75 s	1.27
90 s	1.06
120 s	0.82

As evident from Table 4.2, the variation of the headway from 75 s to 120 s covers most real railway conditions in peak hours in terms of the ratio of the number of trains and the number of platforms N_{tr} / N_{pl}.

In some cases, like Central Line and Piccadilly Line in London, the ratio N_{tr} / N_{pl} (0.73) is still smaller than the smallest simulated value (0.82 when the headway is 120s). However, this does not reduce the importance of the simulation performed, since, as has been stated at the beginning of this section, when there are fewer trains than there are available platforms, trains can dwell in a platform when a disturbance happens. This research is more concerned with the situation where trains are dense and the problem of stopped trains (and thus halted passengers) becomes paramount.

4.3 Origin-destination (OD) types

As has been stated in Chapter 2, the passenger flow from one station to another is defined as OD, or an origin-destination pair value.

The simulation tool used in this research reads the file that states all the OD values at the beginning of a simulation, which means that it has the ability to simulate various kinds of line characteristics in terms of passenger flow by changing the OD values.

Though the target of this research is set as metro-type railways, the line characteristics in terms of passenger flow are not the same for different metro-type railways. For example, in the case of the Circle Line in London, all the stations are in Zone 1, which may mean that each station basically has the same traffic volume; while on the Northern Line in London, which runs as far as Zone 5, the terminal in Zone 5 (High Barnet) is obviously less frequent than Leicester Square in terms of passenger flow.

Furthermore, even within one line, the characteristics in terms of passenger flow are not the same if the line is divided into several sections. For example, in the case of the Northern Line, the stations in Zone 1 may have basically the same passenger flow, thus the load on trains shows the same characteristics as those of the line that runs within the city centre, even though the Northern Line runs from one side of the suburbs of London to the other side.

Further still, even on the same section of a line, the characteristics in terms of passenger flow are not the same at different times. A clear example is the behaviour of commuters, that is, a lot of people go to work from their home to their office in the morning, while the opposite flow is observed in the evening.

Let us call those characteristics of a line in terms of passenger flow an "OD type". Now, each line has its own type of OD, since no two different metro-type railways in the world share all their station characteristics. Even the same line can have several types of OD, as has been discussed above. Thus, in order to investigate whether a traffic regulation is effective for this line or for that line, simulation has to be performed in every line and/or in every period of a day, which is virtually impossible, considering the number of simulation cases.

Instead of performing a large number of simulations under a variety of conditions, some typical OD types should be examined. In this thesis, five types of OD are represented. They are chosen so that they are as typical of metro-type railways as possible, and as independent of each other as possible, and have a distinct and significant role to play in the regulation behaviour.

The five types introduced in this research are named as "Flat", "Centre", "Sub", "Mid_1", and "Mid_2". Each type is explained below.

4.3.1 Flat OD

This OD has the same OD pair value for any combination of origin-destination stations. For example, Table 4.3 shows the Flat OD type in which 18 passengers per minute travel from any station to any other station. (Let us call this “Flat 18”. If the OD pair value is 24 passengers/min, then it is called “Flat 24”.)

Table 4.3: Origin-destination passenger flow of Flat 18

		Origin station number													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Destination station number	1	-	18	18	18	18	18	18	18	18	18	18	18	18	18
	2	18	-	18	18	18	18	18	18	18	18	18	18	18	18
	3	18	18	-	18	18	18	18	18	18	18	18	18	18	18
	4	18	18	18	-	18	18	18	18	18	18	18	18	18	18
	5	18	18	18	18	-	18	18	18	18	18	18	18	18	18
	6	18	18	18	18	18	-	18	18	18	18	18	18	18	18
	7	18	18	18	18	18	18	-	18	18	18	18	18	18	18
	8	18	18	18	18	18	18	18	-	18	18	18	18	18	18
	9	18	18	18	18	18	18	18	18	-	18	18	18	18	18
	10	18	18	18	18	18	18	18	18	18	-	18	18	18	18
	11	18	18	18	18	18	18	18	18	18	18	-	18	18	18
	12	18	18	18	18	18	18	18	18	18	18	18	-	18	18
	13	18	18	18	18	18	18	18	18	18	18	18	18	-	18
	14	18	18	18	18	18	18	18	18	18	18	18	18	18	-

Unit: passenger/min

Flat OD represents a situation such as when the stations are mostly in the same traffic volume, like a line that runs only within the city centre, without a large central station. Figure 4.1 shows the passenger flow at each station and on a train in the case of Flat 18. This is a theoretical example, but there are situations with short hops journey. RATP in Paris can exhibit this behaviour.

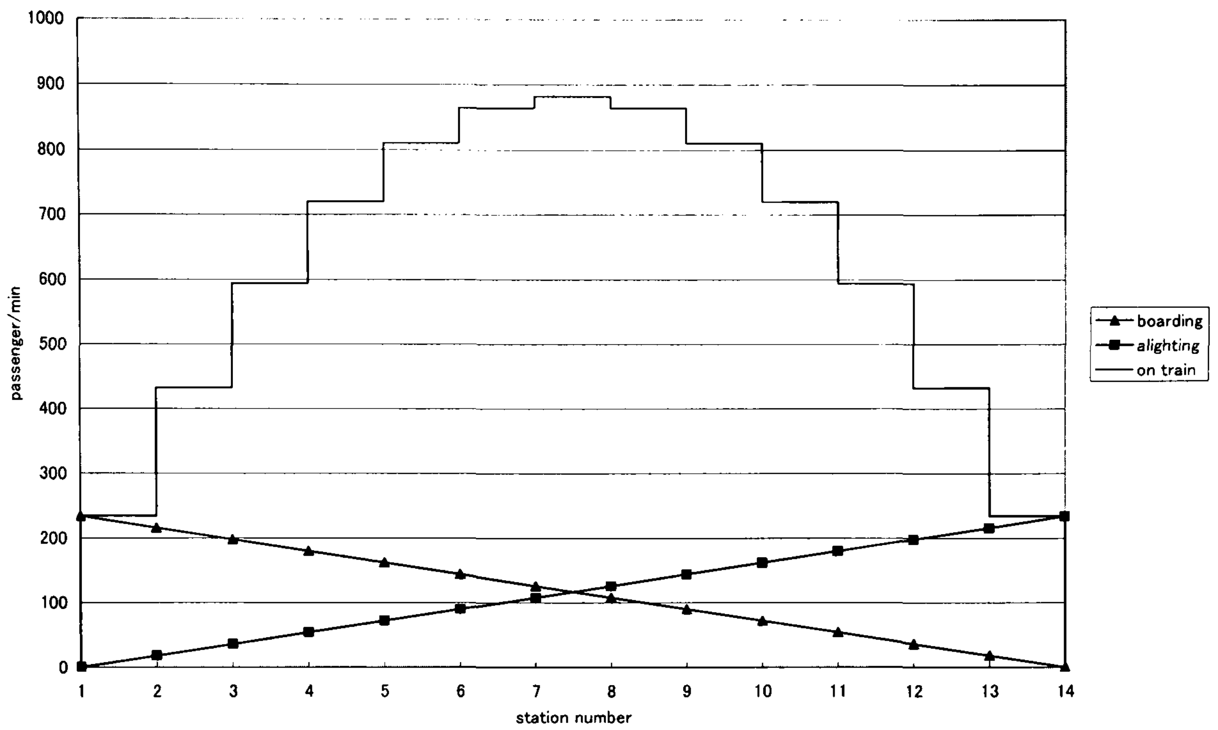

Figure 4.1: Passenger flow at each station and on a train (passenger/min, Flat 18)

The characteristics of this type are:

(1) The exchange of passengers (i.e., the sum of boarding and alighting passengers) is the same at all stations. At Station 1, passengers going to Stations 2 to 14 board a train, while no passengers alight here, thus the exchange of passengers at Station 1 is 234 passengers/min ($= 18 \times 13$) in Flat 18. At Station 2, passengers going to Stations 3 to 14 board a train, while passengers from Station 1 alight here, thus the exchange of passengers at Station 2 is again 234 passengers/min ($= 18 \times 12 + 18 \times 1$). At all of the other stations, the exchange of passengers is 234 passengers/min.

(2) The number of passengers on a train increases as it runs from Station 1 to 7, then shows the maximum value from Station 7 to 8. After leaving Station 8, the number of passengers decreases. In other words, a train is more congested in the centre of the line, while it is less congested at the edges of the line.

4.3.2 Centre OD

In this OD type, the flow of passengers from and to Station 14 is much larger than any other OD pair value. This is called "Centre OD", since Station 14 is regarded as the station in the city centre, thus trains running from Station 1 to Station 14 are going to the city centre.

Table 4.4 shows the origin-destination passenger flow per minute when the flow of passengers from and to Station 14 is 60 passengers per minute, while other OD pair values are 6 passengers per minute. (Let us call this "Centre 60".)

Table 4.4: Origin-destination passenger flow of Centre 60

		Origin station number													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Destination station number	1	–	6	6	6	6	6	6	6	6	6	6	6	6	60
	2	6	–	6	6	6	6	6	6	6	6	6	6	6	60
	3	6	6	–	6	6	6	6	6	6	6	6	6	6	60
	4	6	6	6	–	6	6	6	6	6	6	6	6	6	60
	5	6	6	6	6	–	6	6	6	6	6	6	6	6	60
	6	6	6	6	6	6	–	6	6	6	6	6	6	6	60
	7	6	6	6	6	6	6	–	6	6	6	6	6	6	60
	8	6	6	6	6	6	6	6	–	6	6	6	6	6	60
	9	6	6	6	6	6	6	6	6	–	6	6	6	6	60
	10	6	6	6	6	6	6	6	6	6	–	6	6	6	60
	11	6	6	6	6	6	6	6	6	6	6	–	6	6	60
	12	6	6	6	6	6	6	6	6	6	6	6	–	6	60
	13	6	6	6	6	6	6	6	6	6	6	6	6	–	60
	14	60	60	60	60	60	60	60	60	60	60	60	60	60	–

Unit: passenger/min

Centre OD represents a situation such as the morning scenes on a line that links the city centre and stations in the suburbs. Figure 4.2 shows the passenger flow at each station and on a train in the case of Centre 60. Obviously, the passenger load on trains near the city centre is heavier, as is observed in a commuter line in big cities.

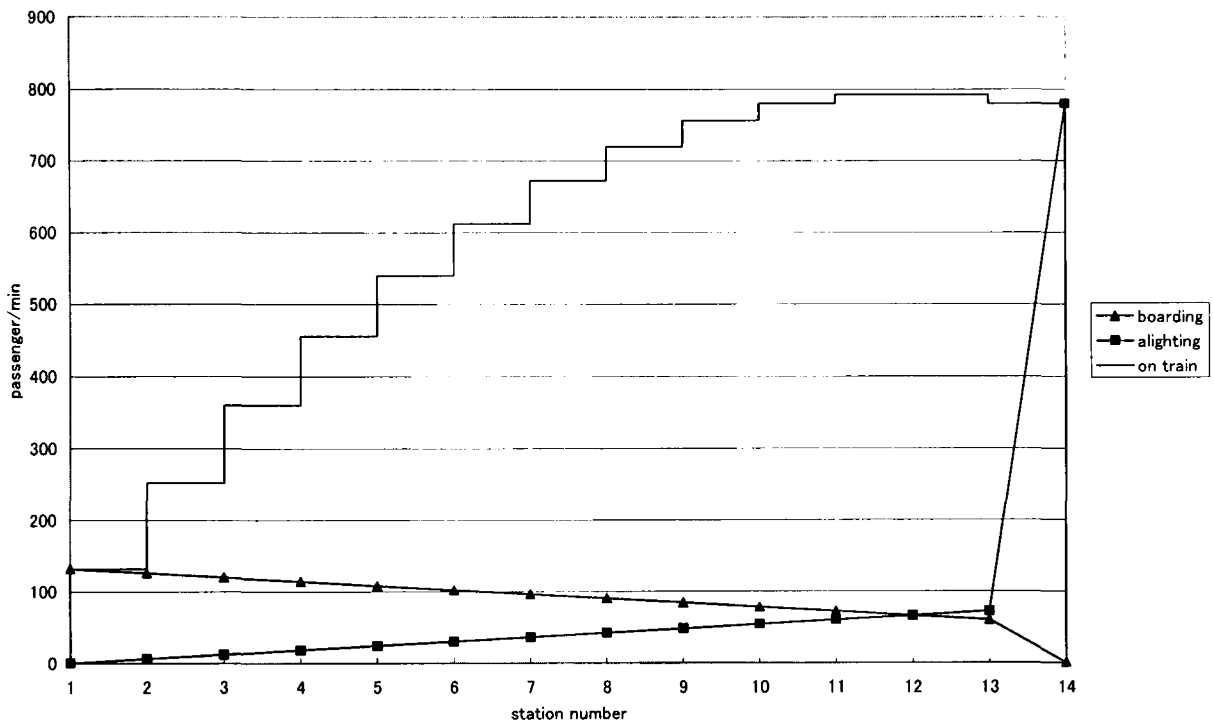

Figure 4.2: Passenger flow at each station and on a train (passenger/min, Centre 60)

4.3.3 Sub OD

This is the opposite case to Centre OD. The flow of passengers from and to Station 1 is larger than any other OD pair values. Station 1 is regarded as the station in the city centre, which means that trains running from Station 1 to Station 14 are going to the suburbs. Thus, let us call this type of OD "Suburbs OD" or simply "Sub OD".

As each OD is named in the previously introduced OD types, when 50 passengers per minute travel to and from Station 1, while for other stations the flow is 6 passengers per minute, let us call this type "Sub 50". Table 4.5 shows the origin-destination passenger flow per minute of Sub 50.

Table 4.5: Origin-destination passenger flow of Sub 50

		Origin station number													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Destination station number	1	–	50	50	50	50	50	50	50	50	50	50	50	50	50
	2	50	–	6	6	6	6	6	6	6	6	6	6	6	6
	3	50	6	–	6	6	6	6	6	6	6	6	6	6	6
	4	50	6	6	–	6	6	6	6	6	6	6	6	6	6
	5	50	6	6	6	–	6	6	6	6	6	6	6	6	6
	6	50	6	6	6	6	–	6	6	6	6	6	6	6	6
	7	50	6	6	6	6	6	–	6	6	6	6	6	6	6
	8	50	6	6	6	6	6	6	–	6	6	6	6	6	6
	9	50	6	6	6	6	6	6	6	–	6	6	6	6	6
	10	50	6	6	6	6	6	6	6	6	–	6	6	6	6
	11	50	6	6	6	6	6	6	6	6	6	–	6	6	6
	12	50	6	6	6	6	6	6	6	6	6	6	–	6	6
	13	50	6	6	6	6	6	6	6	6	6	6	6	–	6
	14	50	6	6	6	6	6	6	6	6	6	6	6	6	–

Unit: passenger/min

Sub OD represents a situation such as the evening situation on a line that links the city centre and the stations in the suburbs. Figure 4.3 shows the passenger flow at each station and on a train in the case of Sub 50. Again, obviously, the passenger load on trains near the city centre is heavier, as is observed on a commuter line in big cities.

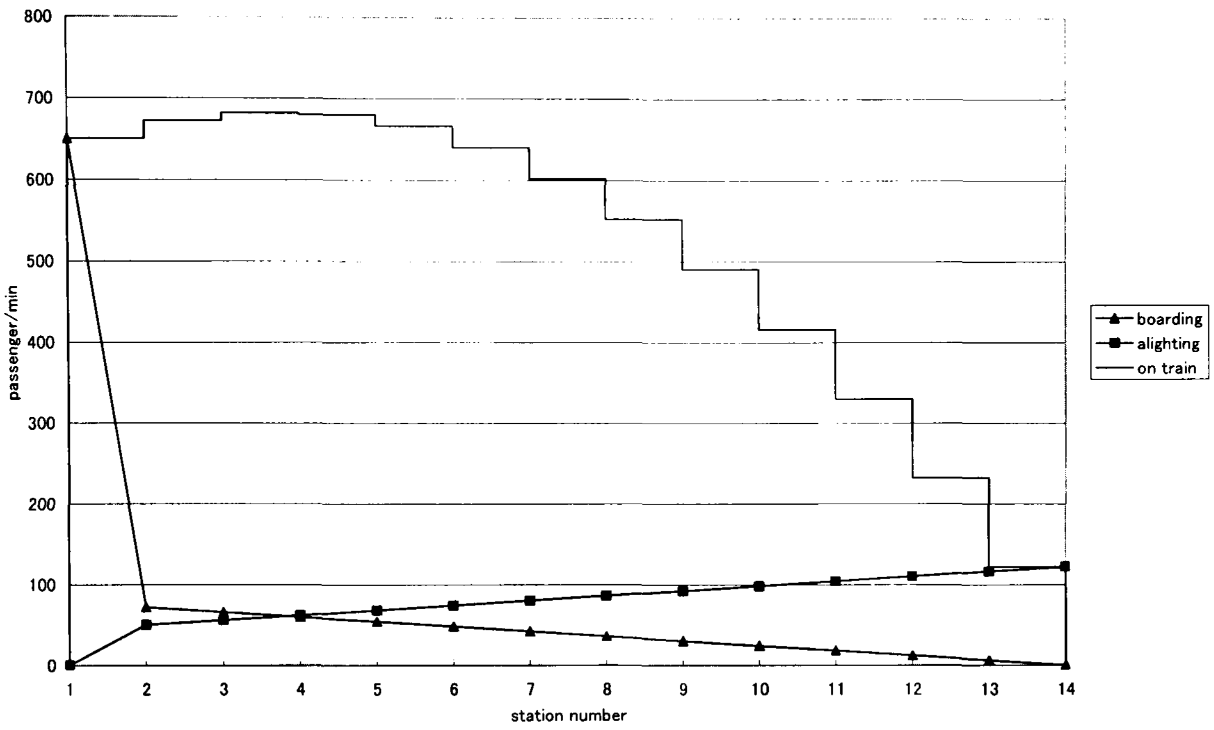

Figure 4.3: Passenger flow at each station and on a train (passenger/min, Sub 50)

4.3.4 Mid_1 OD

This OD is one where the flow of passengers from and to Station 8 is larger than other OD pair values. This is named as "Mid_1 OD", since there is a large central station in the middle of the line.

Table 4.6 shows the origin-destination passenger flow per minute in which the flow of passengers from and to Station 8 is 70 passengers per minute, while for other stations the flow is 6 passengers per minute. This is called "Mid_1 70", in the same manner as the previous cases.

Table 4.6: Origin-destination passenger flow of Mid_1 70

		Origin station number													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Destination station number	1	–	6	6	6	6	6	6	70	6	6	6	6	6	6
	2	6	–	6	6	6	6	6	70	6	6	6	6	6	6
	3	6	6	–	6	6	6	6	70	6	6	6	6	6	6
	4	6	6	6	–	6	6	6	70	6	6	6	6	6	6
	5	6	6	6	6	–	6	6	70	6	6	6	6	6	6
	6	6	6	6	6	6	–	6	70	6	6	6	6	6	6
	7	6	6	6	6	6	6	–	70	6	6	6	6	6	6
	8	70	70	70	70	70	70	70	–	70	70	70	70	70	70
	9	6	6	6	6	6	6	6	70	–	6	6	6	6	6
	10	6	6	6	6	6	6	6	70	6	–	6	6	6	6
	11	6	6	6	6	6	6	6	70	6	6	–	6	6	6
	12	6	6	6	6	6	6	6	70	6	6	6	–	6	6
	13	6	6	6	6	6	6	6	70	6	6	6	6	–	6
	14	6	6	6	6	6	6	6	70	6	6	6	6	6	–

Unit: passenger/min

Mid_1 OD represents the line that runs in the city centre, and has a large central station. The central station has both a large demand and a large supply of passengers. Figure 4.4 shows the passenger flow at each station and on a train in the case of Mid_1 70. A steep peak in the passenger load on a train in the middle of the line is observed, compared with the figure for Flat OD.

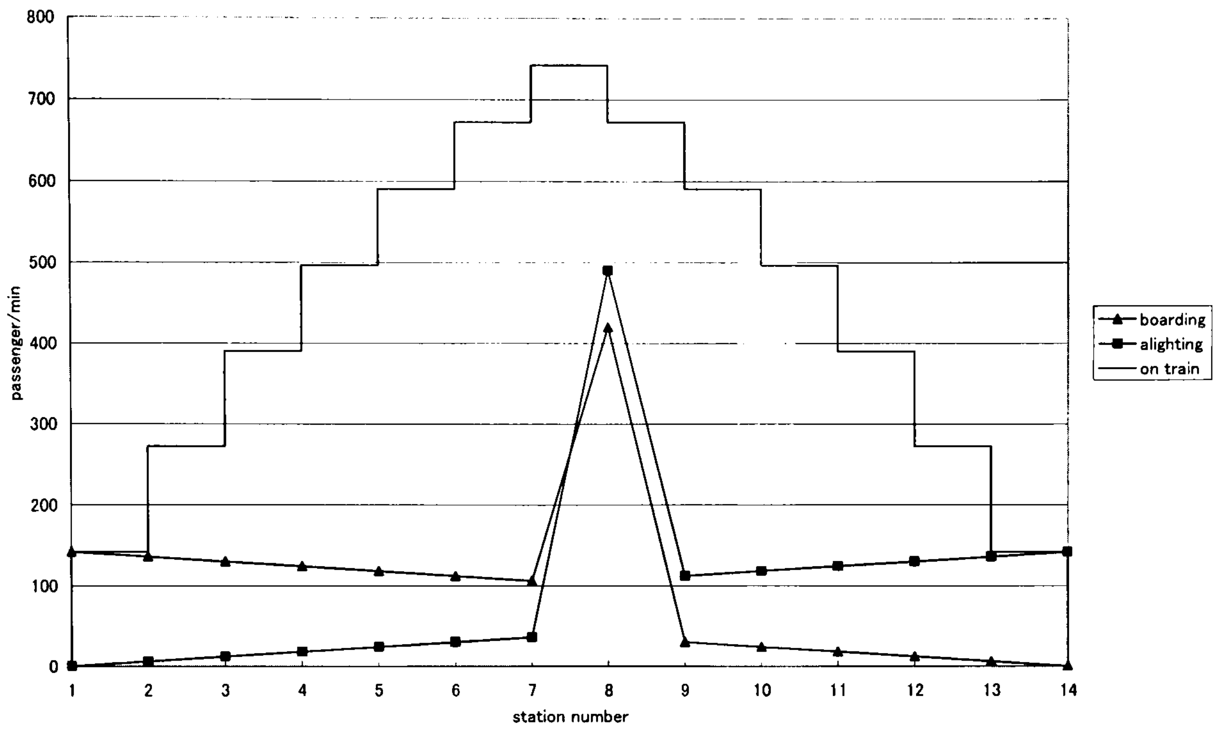

Figure 4.4: Passenger flow at each station and on a train (passenger/min, Mid_1 70)

4.3.5 Mid_2 OD

This OD also simulates the situation where there is a large station in the middle of the line. The difference from Mid_1 is only that the flow of passengers to Station 8 is larger. Table 4.7 shows the origin-destination passenger flow per minute where the flow of passengers to Station 8 is 90 passengers per minute, while for other stations the flow is 6 passengers per minute. (Let this be called "Mid_2 90".)

Table 4.7: Origin-destination passenger flow of Mid_2 90

		Origin station number													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Destination station number	1	–	6	6	6	6	6	6	6	6	6	6	6	6	6
	2	6	–	6	6	6	6	6	6	6	6	6	6	6	6
	3	6	6	–	6	6	6	6	6	6	6	6	6	6	6
	4	6	6	6	–	6	6	6	6	6	6	6	6	6	6
	5	6	6	6	6	–	6	6	6	6	6	6	6	6	6
	6	6	6	6	6	6	–	6	6	6	6	6	6	6	6
	7	6	6	6	6	6	6	–	6	6	6	6	6	6	6
	8	90	90	90	90	90	90	90	–	90	90	90	90	90	90
	9	6	6	6	6	6	6	6	6	–	6	6	6	6	6
	10	6	6	6	6	6	6	6	6	6	–	6	6	6	6
	11	6	6	6	6	6	6	6	6	6	6	–	6	6	6
	12	6	6	6	6	6	6	6	6	6	6	6	–	6	6
	13	6	6	6	6	6	6	6	6	6	6	6	6	–	6
	14	6	6	6	6	6	6	6	6	6	6	6	6	6	–

Unit: passenger/min

Mid_2 OD represents a temporary situation such as when a certain event (e.g. a football match) is held near Station 8. Figure 4.5 shows the passenger flow at each station and on a train in the case of Mid_2 90.

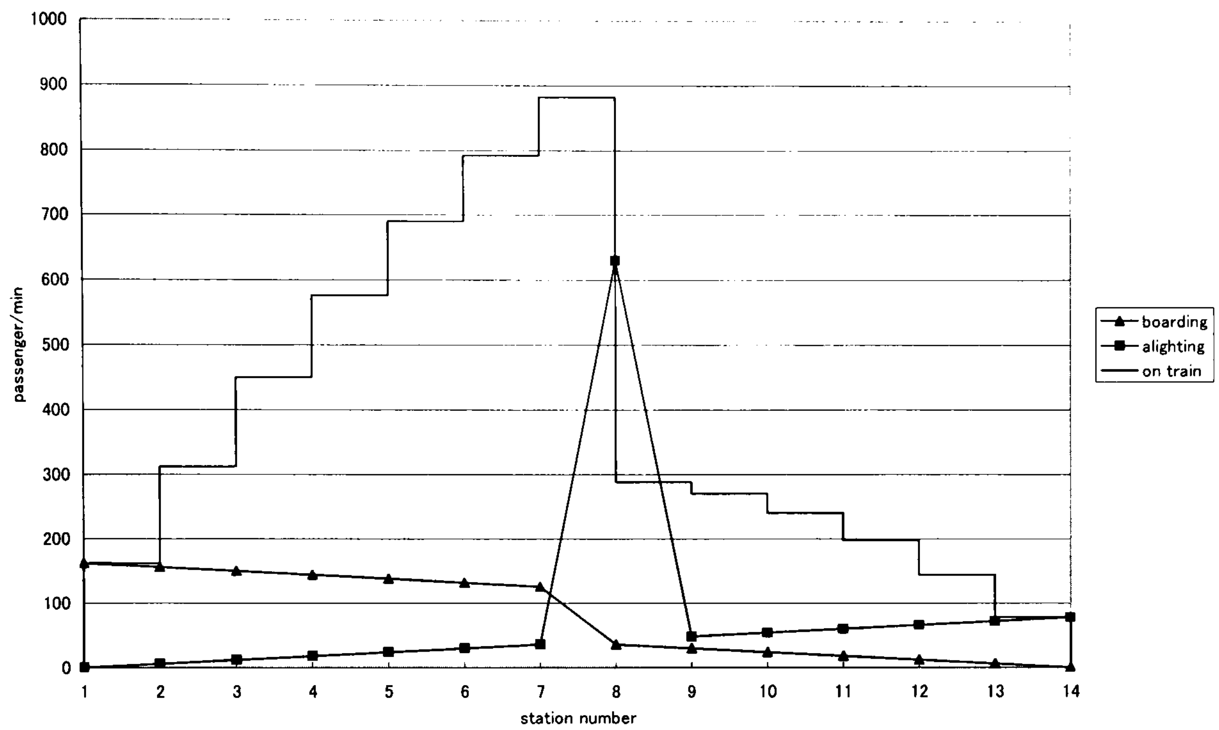

Figure 4.5: Passenger flow at each station and on a train (passenger/min, Mid_2 90)

4.3.6 Summary of OD types

Figure 4.6 summarises the characteristics of the five OD types.

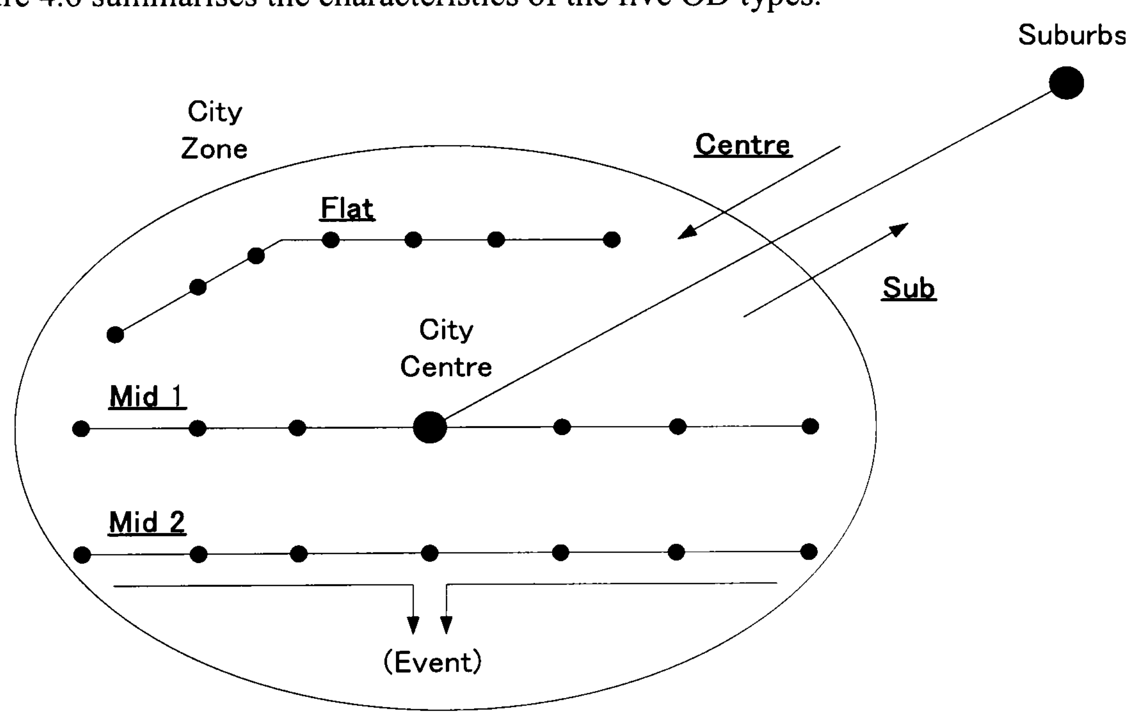

Figure 4.6: Images of each OD type

There is one point to be noted here. In the discussion above, no explanation is given as to the scale of each OD, such as Flat "18", Centre "60", Sub "50", Mid_1 "70", and Mid_2 "90". These numbers are chosen so that the maximum load of passengers on a train is mostly the same (when trains run at an equal headway, and when there is no disturbance). This is due to the following two considerations. For one thing, it is impossible to compare the results of different OD types if the load at ordinary times is very different in each case, since it means that the margins of each case are different. For the other, a real train operator would make a train service plan in which the maximum load of passengers on a train at ordinary times is below a certain rate of congestion. If the rate of congestion is too high, it is not acceptable for passengers, while, if the rate of congestion is too low, it means that the cost of the service is high and the business is not sustainable.

4.4 Disturbance site

The simulation tool used in this research has the capability to put an initial disturbance into the line by setting the disturbance site (which must be a platform) and the duration of disturbance[23]. There are 14 stations in the model used in this research, which means that there are 28 platforms in the line (*cf.* Figure 2.1).

Out of these platforms, the platform of Station 3 for a train running from Station 1 to 14 is chosen as the disturbance site (Figure 4.7).

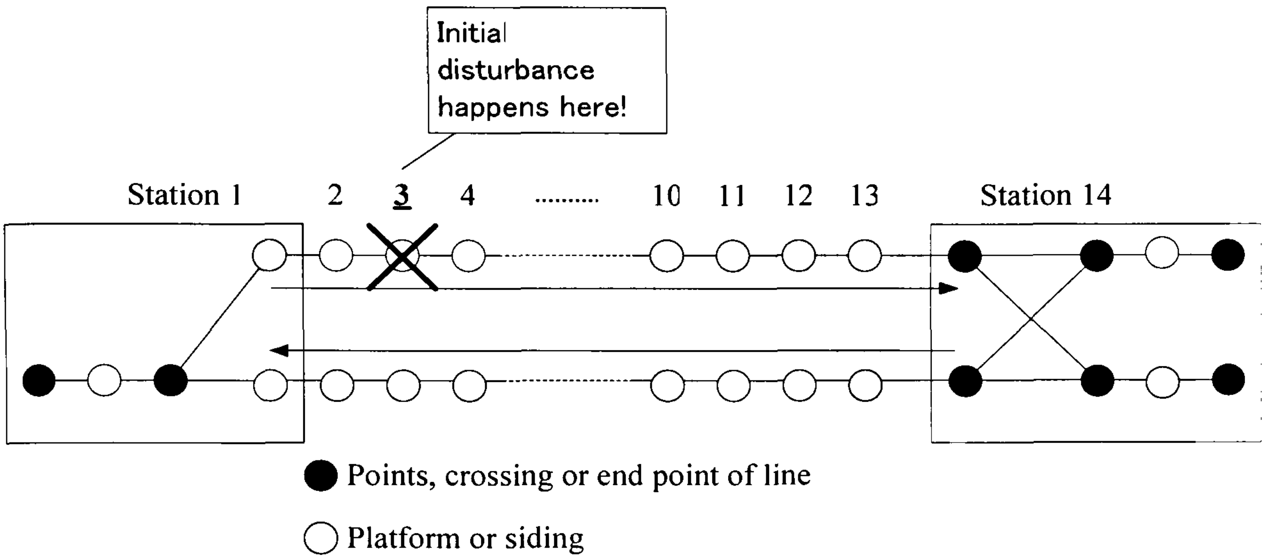

Figure 4.7: Disturbance site

The position of this disturbance site gives stringent conditions for traffic regulation. If a train is stuck here, following ones have to stop behind it, which means that the effect of the existence of a terminal cannot be ignored when disturbance lasts for long enough. Usually, a city centre terminal has a heavier passenger demand than other stations, thus there may be a lot of passengers gathering in the terminal (Station 1 in this case). This leads to a severe discomfort for passengers.

Another reason that the platform of Station 3 for a train running from Station 1 to 14 is chosen as the disturbance site is that the characteristics of the passenger load on a train show a significant difference between Centre OD and Sub OD. In Centre OD, the load of passengers on a train running from Station 1 to 14 will grow more and more as trains leave from Station 3; while in Sub OD, the load of passengers on a train of the same direction shows the maximum value from Station 3 to 4. This fact provides interestingly different cases to compare the results of regulation.

4.5 Duration of disturbance

The final parameter discussed in this chapter is the duration of the initial disturbance, which is entered into the simulator with the information of the disturbance site, as has been stated at the beginning of 4.4.

In this research, the duration of disturbance varies through 300s, 600s, 900s and 1200s. These values are set so that the duration of disturbance is three to ten times the operational headway (section 4.2).

Suppose that the line has a headway time of 3 min. For this line, the equivalent duration of disturbance is 9 to 30 min. If the disturbance is short, then regulation is not crucially important, since a crude method such as adjusting the intervals between trains is effective enough to settle traffic. On the other hand, if the disturbance is long, then the regulation that is considered in this research is not enough. As has been stated in Chapter 1, the permitted regulation in this research is to change departure/arrival times of trains from/at stations. If the disturbance is longer than about 30 min, other regulation methods are necessary such as suspending a train's operation halfway (before its destination), changing a train destination, preparing extra trains to rescue gathered passengers, closing some stations to avoid panic, etc.

4.6 Conclusion

This chapter explains the conditions of the simulations performed in this research. The parameters discussed here are the ratio of the number of trains to the number of platforms, OD types, the disturbance site, and the duration of disturbance. It is revealed that these conditions not only describe the real world well, but also provide stringent conditions for traffic regulation. Consequently, it is concluded that these conditions can be relied on in performing simulations and examining the performance of regulation algorithms.

5 Results and discussions of the Original Murata regulator

In this chapter, the results of the Original Murata regulator are presented. The regulator shows good performance in reducing the three kinds of passenger discomfort (waiting discomfort, travelling discomfort, and congestion discomfort). However, when the outputs of the regulator, train movements in particular, are closely examined, a significant difference from real railway operations is observed. A discussion follows this observation.

5.1 Simulation conditions

5.1.1 Headway

The results shown in this chapter have been calculated for a 90-second headway.

5.1.2 Passenger flow

As has been stated in Chapters 2 and 4, the passenger flow from one station to another is defined as OD, or origin-destination pair value. The results shown in this chapter are calculated for five OD types, which are referred to as "Flat 18", "Centre 60", "Sub 50", "Mid_1 70" and "Mid_2 90".

5.1.3 Disturbance of traffic

The initial disturbance is injected into the simulation by setting which train is halted at which station for how many seconds on which round trip. In this chapter, the number of the train, the number of the station, and the number of the round trip are fixed, that is, Train 11 is delayed at Station 3 on the second round trip. The duration of the disturbance varies through 300s, 600s, 900s and 1200s. Figure 5.1 gives the example of a train movement diagram under disturbed conditions without regulation, in which Train 11, the thick line in Figure 5.1, is held at Station 3 for 600s.

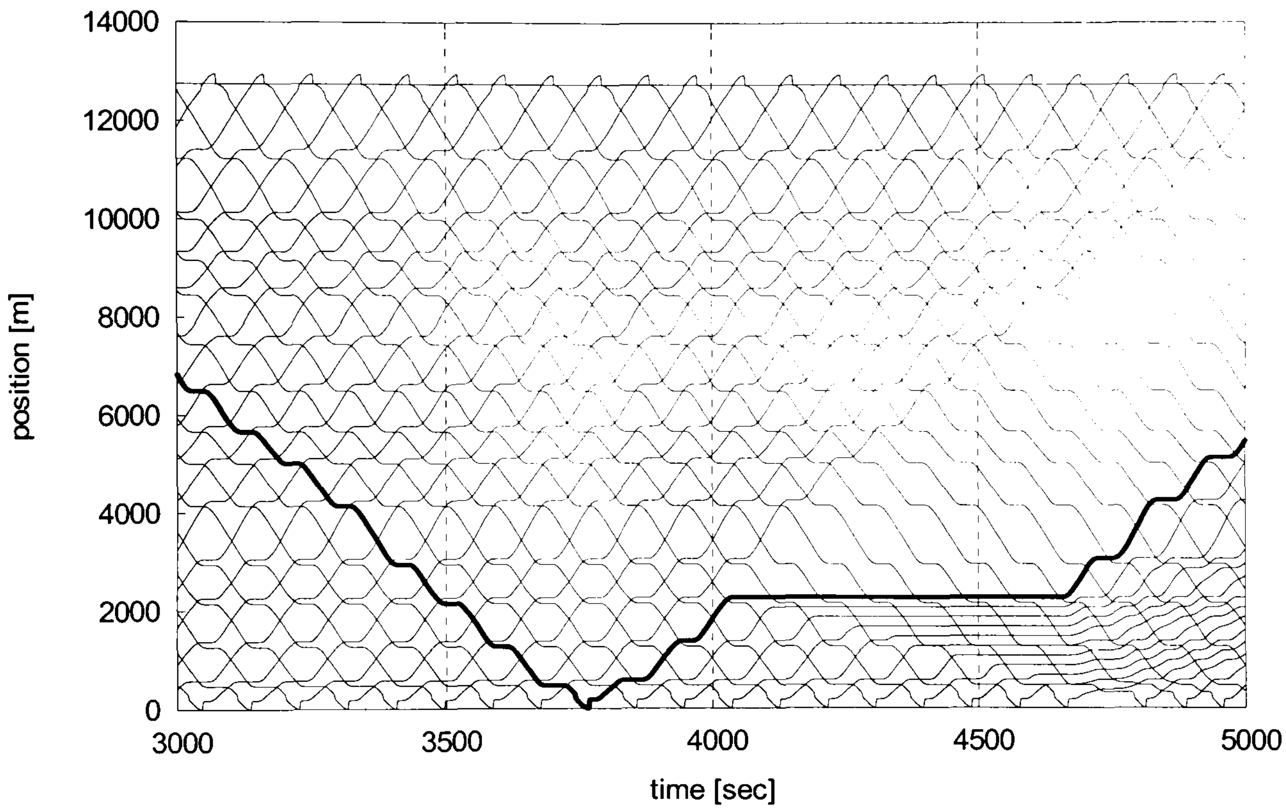

Figure 5.1: Train movement diagram under disturbed conditions without regulation

5.2 Simulation results

The results are shown in Figures 5.2-5.6, which are Flat 18, Centre 60, Sub 50, Mid_1 70 and Mid_2 90, respectively. The discomfort value is accumulated over the 10000s that has elapsed since Train 1 left Station 1.

Overall, the results show a good performance for the Original Murata regulator, which reduces total discomfort values in each case significantly. However, there is a problem with these solutions, which is to be described in the next section.

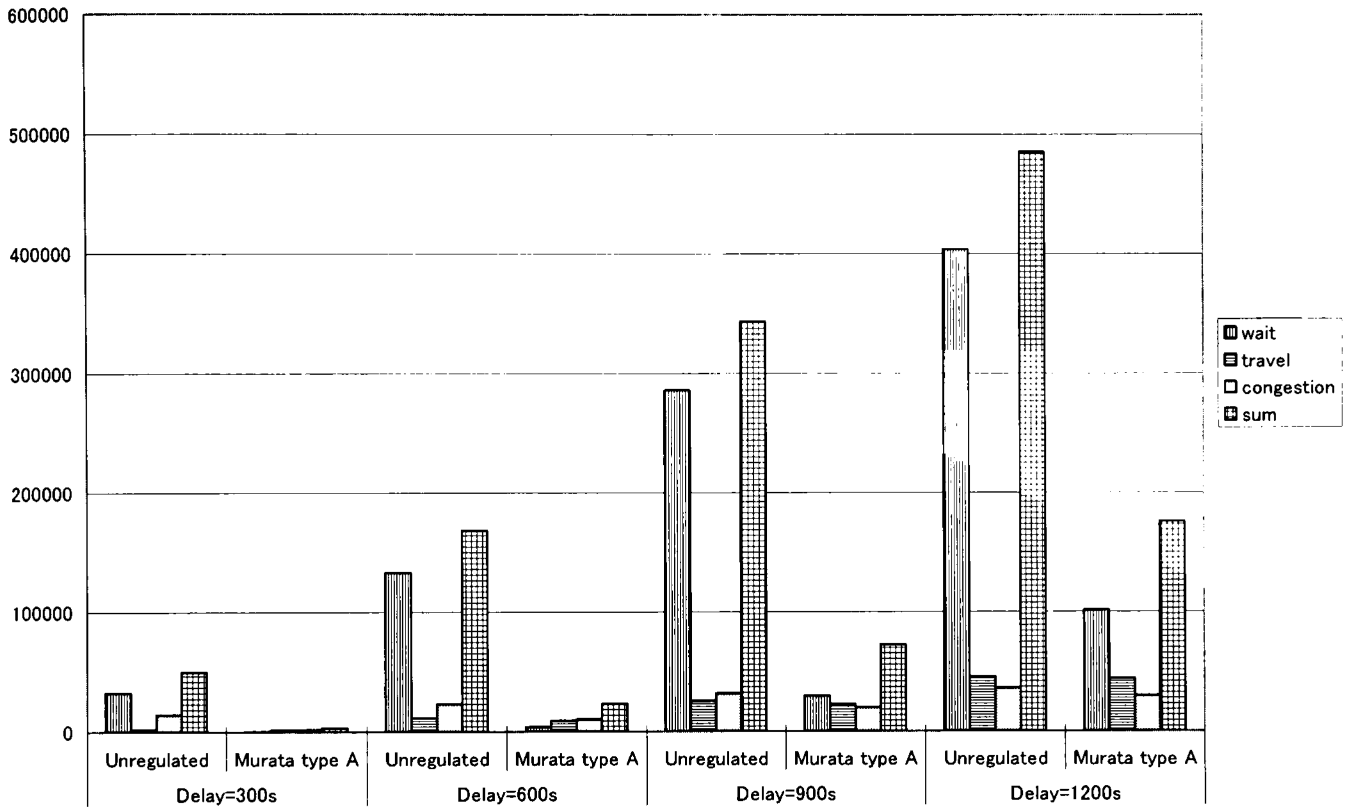

Figure 5.2: Discomfort comparison, Flat 18

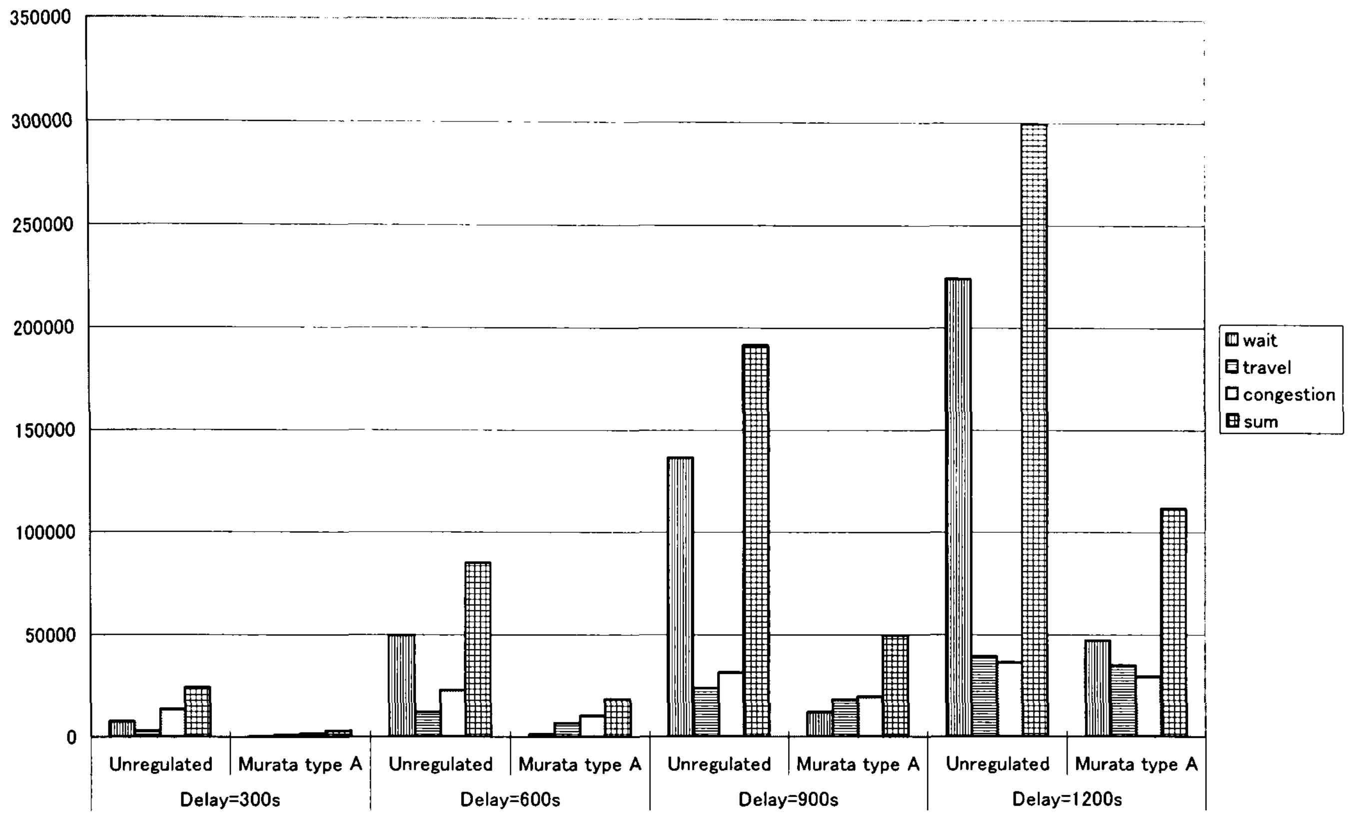

Figure 5.3: Discomfort comparison, Centre 60

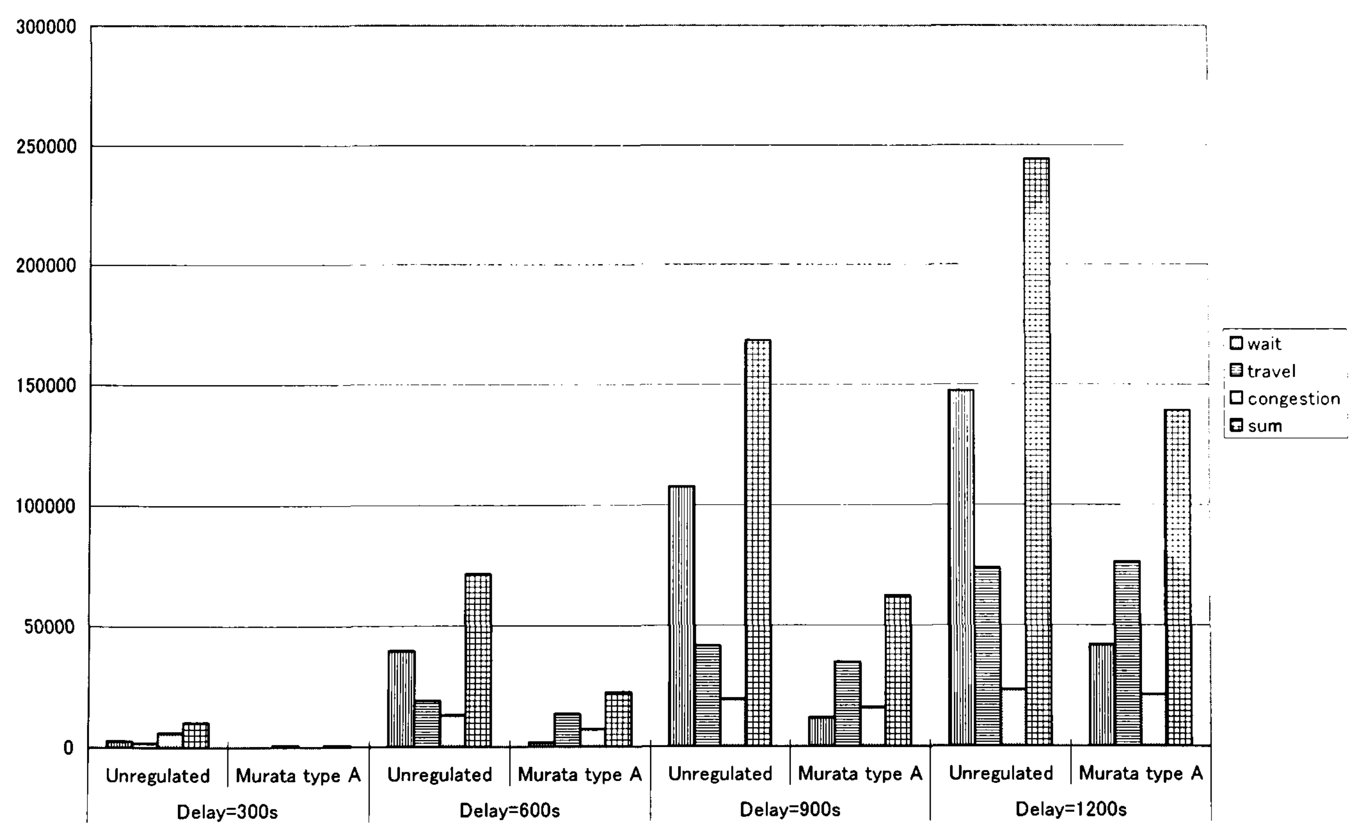

Figure 5.4: Discomfort comparison, Sub 50

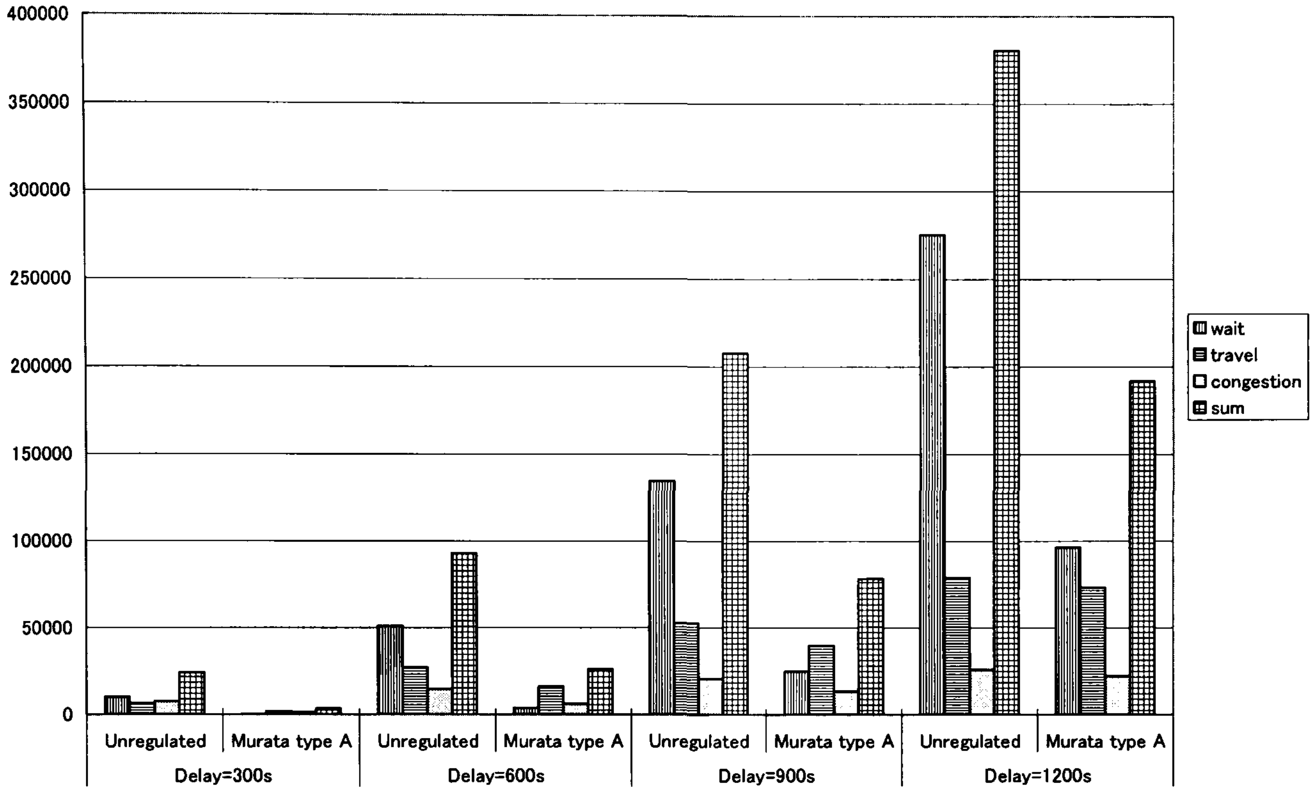

Figure 5.5: Discomfort comparison, Mid1_70

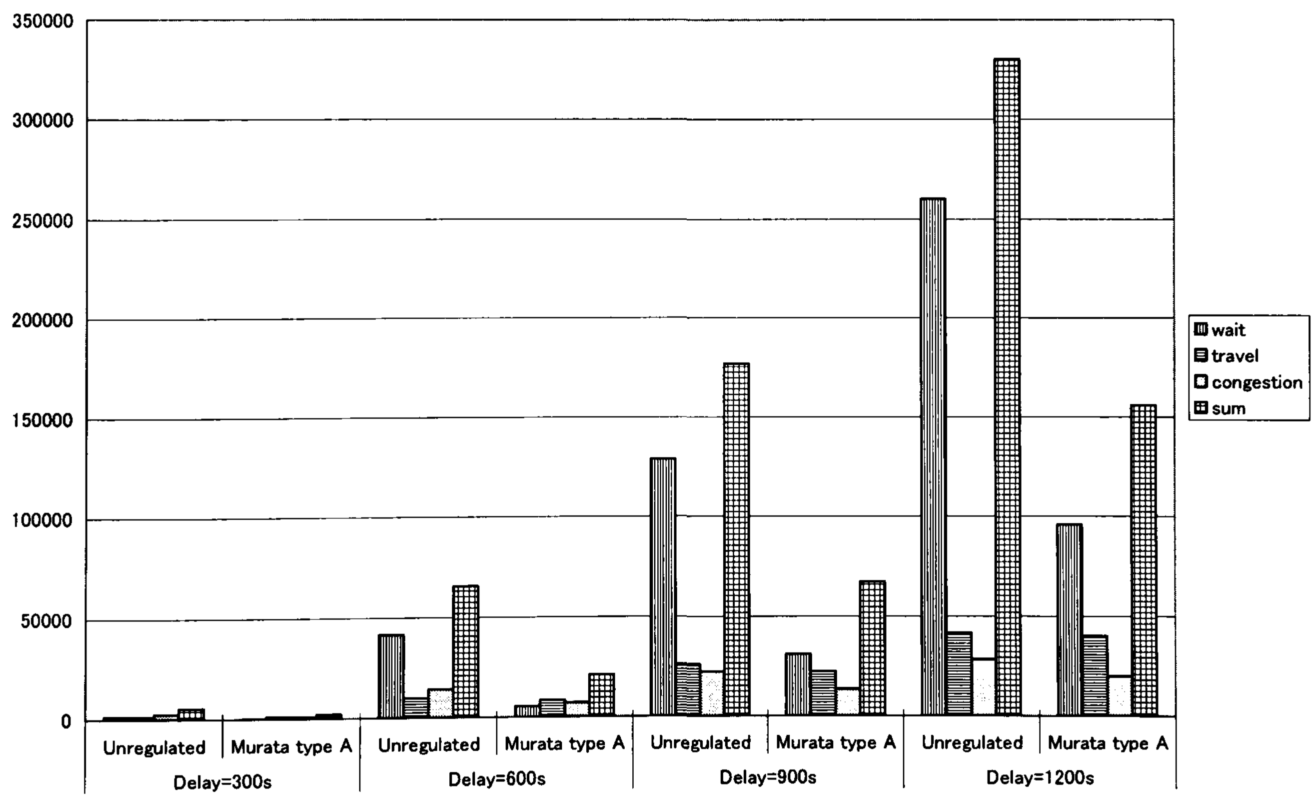

Figure 5.6: Discomfort comparison, Mid2_90

5.3 Discussion

From these results, it would appear that the Original Murata regulator shows a good performance in settling disturbed traffic. However, there is a problem in the train movements. Figure 5.7 is an example of the regulated train movements under disturbed conditions. In this figure, some of the trains running before the disturbed train (the thick line in the figure) are forced to stop at stations longer than usual, which means that they have obeyed the outputs of the regulator so that these trains can share the passengers that would originally have got on to the disturbed train.

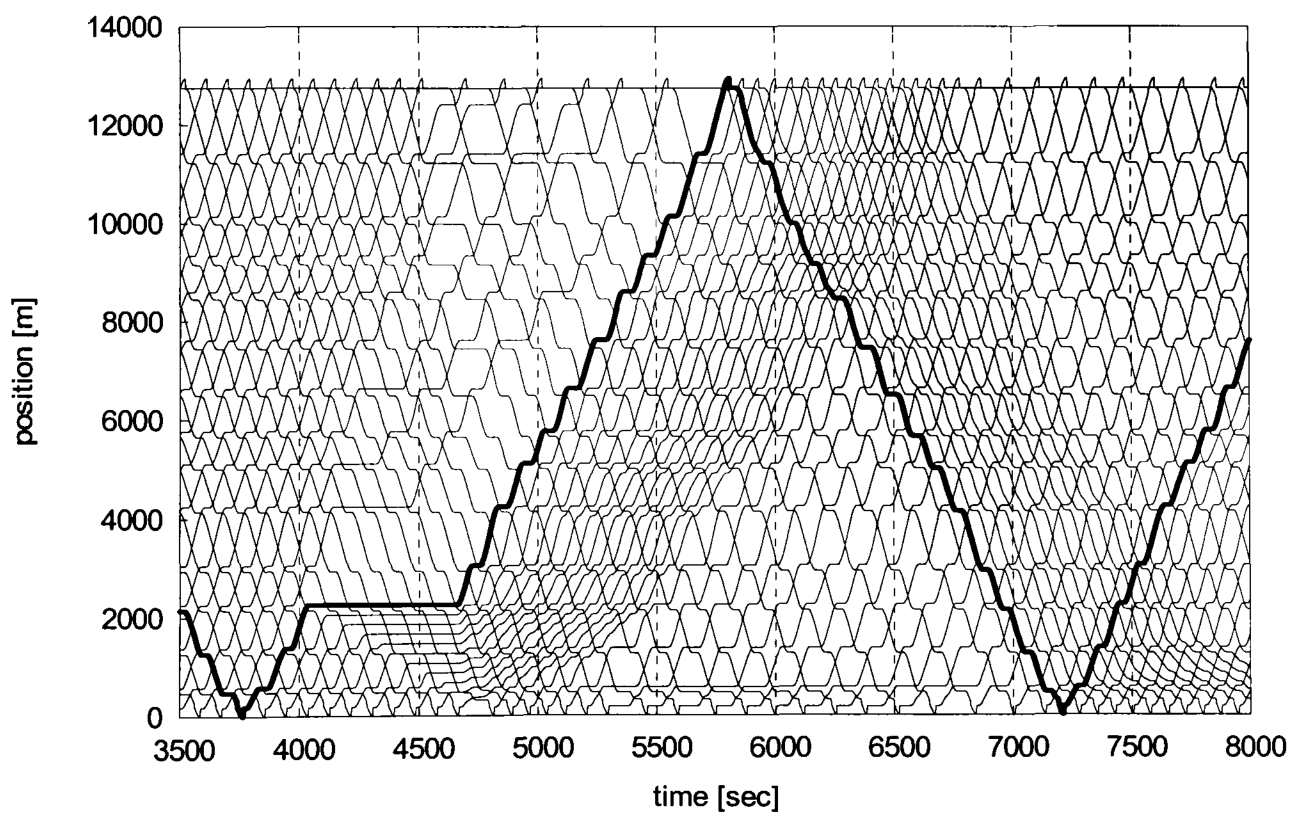

Figure 5.7: Regulated train movements under disturbed conditions

In Figure 5.7, it can be seen that many of the trains running after the disturbed train have to stop before reaching Station 3, where the initial delay happens. This situation is not desirable in real railway operations from the following two viewpoints.

First, the passengers that are stuck in the trains would feel frustrated more than those who are forced to wait at stations, and than those who have to stay even on slowly running trains. Passengers at stations can choose to abandon their journey and go back, or change a means of transport to buses, taxis, etc., while those on trains stopped behind stations cannot. Even if passengers have to stay for longer than they expected, they would feel confident that they are moving forward so long as the trains are actually moving, while passengers on completely halted trains would not. Where trains have to stop in tunnels, which is not a rare case in metro-type railways, the frustration of the stuck passengers will be even larger.

Second, trains cannot usually move backwards for safety reasons; in other words, the signalling facilities of the railway systems do not usually allow trains to move backwards. Because of this, if it eventually turns out that the stopped trains cannot move forward any more, passengers have to escape from the trains and thus they have to walk along the track, which is sometimes very dangerous, especially if the line is powered by conductor rails like the London Underground.

Although the situation where trains are halted before reaching the next station is not desirable for the two reasons given, the Original Murata regulator generates output commands that lead to this situation. As will be stated in Chapter 7, a real railway operator would probably stop all trains at once when the disturbance happens, which seems quite a crude regulation technique, but is effective in preventing trains from being stuck between stations.

Here, the following two points should be noted as to why the Original Murata regulator generates the outputs referred to above.

(1) Objective function

The objective function of the Original Murata regulator is the sum of the three kinds of passenger discomfort, which are waiting discomfort, travelling discomfort, and congestion discomfort. None of these takes account of passenger discomfort when a train is stuck and stops between stations.

This kind of discomfort is particularly important in metro-type railways, since quite a few of them are underground lines and thus, being stuck between stations may mean being stuck in a tunnel in many cases. It can easily be understood that passengers may feel severe discomfort when they are halted in a tunnel; however, the existing objective functions do not include this kind of discomfort.

In an attempt to deal with this deficiency, a new objective criterion is to be introduced in this research. This will be stated in Chapter 6.

(2) Assumption within the regulator

Intrinsic to railway traffic regulation, there is the fundamental problem that it is difficult to predict when the disturbed train will resume its journey after the disturbance happens. The Original Murata regulator avoids this problem by assuming that the disturbed train can leave the station only at the start of each regulating calculation period, which occurs every 10s.

This assumption has a significant effect on the problem of halted trains following the disturbed train, since, once the trains following the disturbed train leave the preceding stations, it is essentially impossible to get them back to the previous stations, thus they inevitably end up stopped behind the disturbed train.

5.4 Real regulation

In order to examine the problem of the Original Murata regulator, it would appear to be beneficial to examine the regulation practiced in a real railway operation when a disturbance happens.

Modern metro-type railways are generally equipped with a centralised traffic control system, as well as a communication system between the trains and the centralised traffic control system. A human operator sitting in front of the display that shows the traffic status on the line, has the ability to know when a disturbance happens, at which position on the line and on which train. He/she may know that a disturbance is occurring when he/she surveys the traffic status and finds a train, which, for example, is staying at a station for longer than scheduled. He/she can ask the crew of the train why the train cannot leave, and the crew replies that there has been a malfunction in, for example, a door control system and the doors cannot be locked. Alternatively, the crew may report a fault on the train before the operator knows it.

When the operator knows that there is a disturbance on the line, typically the first procedure to be taken is to hold all trains at stations. If there is a train that is running, then it is allowed to run to the next station and then is held there. If a train cannot enter the next station since all of the platforms of the next station are occupied, then one of the trains that occupy the station is allowed to leave the station to run to the next station if there is an available platform in the next station.

In this thesis, this procedure is called the "stop-all-trains-at-once" strategy, since all trains will stop consequently, waiting for their next orders from the centralised traffic control system.

The "stop-all-trains-at-once" strategy clearly aims at preventing trains from being stuck between stations, thus reducing the passenger discomfort arising from being stuck in a halted train. When there are more trains on the line than there are available platforms, then obviously some trains cannot enter stations and have to stop between stations. Still, this strategy seems more effective than the Original Murata regulator in reducing passenger discomfort resulting from being stuck in a halted train, since the Original Murata regulator does not rescue any halted train.

Though this strategy may seem conservative or even crude, it is apparently effective in preventing trains from being stuck between stations, thus reducing passenger discomfort when a train stops between stations. On the other hand, this style regulation may worsen performance with respect to the existing objective criteria (waiting discomfort, travelling discomfort, and congestion discomfort), since trains may have to stop at stations for a longer time if this strategy is applied, thus passengers may have to wait for trains for a longer time, may have to take a longer time to get to their destination, and may suffer from more severe congestion since the trains have to carry more passengers. Thus, there may be a trade-off relationship between this new aspect of passenger discomfort now being discussed and the existing objective criteria.

5.5 Conclusion

The Original Murata regulator has a good performance in reducing the existing three kinds of discomfort criteria (waiting discomfort, travelling discomfort, and congestion discomfort), however, it makes no attempt to prevent the trains following the disturbed train from stopping between stations behind the disturbed train. A real railway operator would try to avoid such a situation, since passengers may feel severe discomfort when their train is stuck between stations, which, in many cases, is inside a tunnel.

In order to remove this problem, a new objective criterion is necessary to describe the type of passenger discomfort that is stated above. Also, the assumption in the Original Murata regulator that the disturbed train leaves the station at the start of each regulating calculation needs to be reconsidered, since this assumption leads to the situation that the following trains leave the preceding stations without knowing that there is a disturbed train in front of them.

From an observation of real railway operations, the “stop-all-trains-at-once” strategy seems effective in removing the problem stated above. This will be stated in Chapter 7 in more detail.

6 Limits of existing objective criteria --- necessity to introduce "halt discomfort"

6.1 Limits of existing objective criteria

In Chapter 5, it has been shown that the Original Murata regulator has the problem of halting trains between stations. One of the reasons for this problem is that there is no component for the discomfort of halted passengers in the objective function. Chapter 5 has thus concluded that a new objective criterion for this discomfort is necessary.

However, the addition of such an objective criterion is not necessary if changes in the weightings for the existing discomfort criteria lead to regulation outputs that ease the halt discomfort problem. Being stuck in a halted train will generally increase the travel time discomfort, since the travelling time of halted passengers will increase if their train is halted before reaching their destinations.

In this section, the results of a 'sensitivity analysis' are described. Here, 'sensitivity analysis' means analysing the behaviour of the Original Murata regulator with respect to changes in the weighting factors of the objective function. By modifying equation (3.4), the following function is obtained.

$$F_{obj} = \alpha D_w + \beta D_t + \gamma D_c \qquad (6.1)$$

Influence of each factor is analysed independently by changing the weighting factors (α, β, γ) in equation (6.1). The cases investigated here are $(\alpha, \beta, \gamma) = (1,0,0), (0,1,0), (0,0,1)$. These describe the cases in which only one of the three terms (waiting, travelling and congestion, respectively) is involved in the regulation.

The simulation conditions are as follows. The disturbance is injected in such a way that Train 11 is halted at Station 3 for 600 s in the second round-trip. The OD type is Flat 18.

Figures 6.1 through 6.3 are the train movements when the weighting factors are $(\alpha, \beta, \gamma) = (1,0,0), (0,1,0), (0,0,1)$, respectively. In these figures, the thick line is the initially disturbed train.

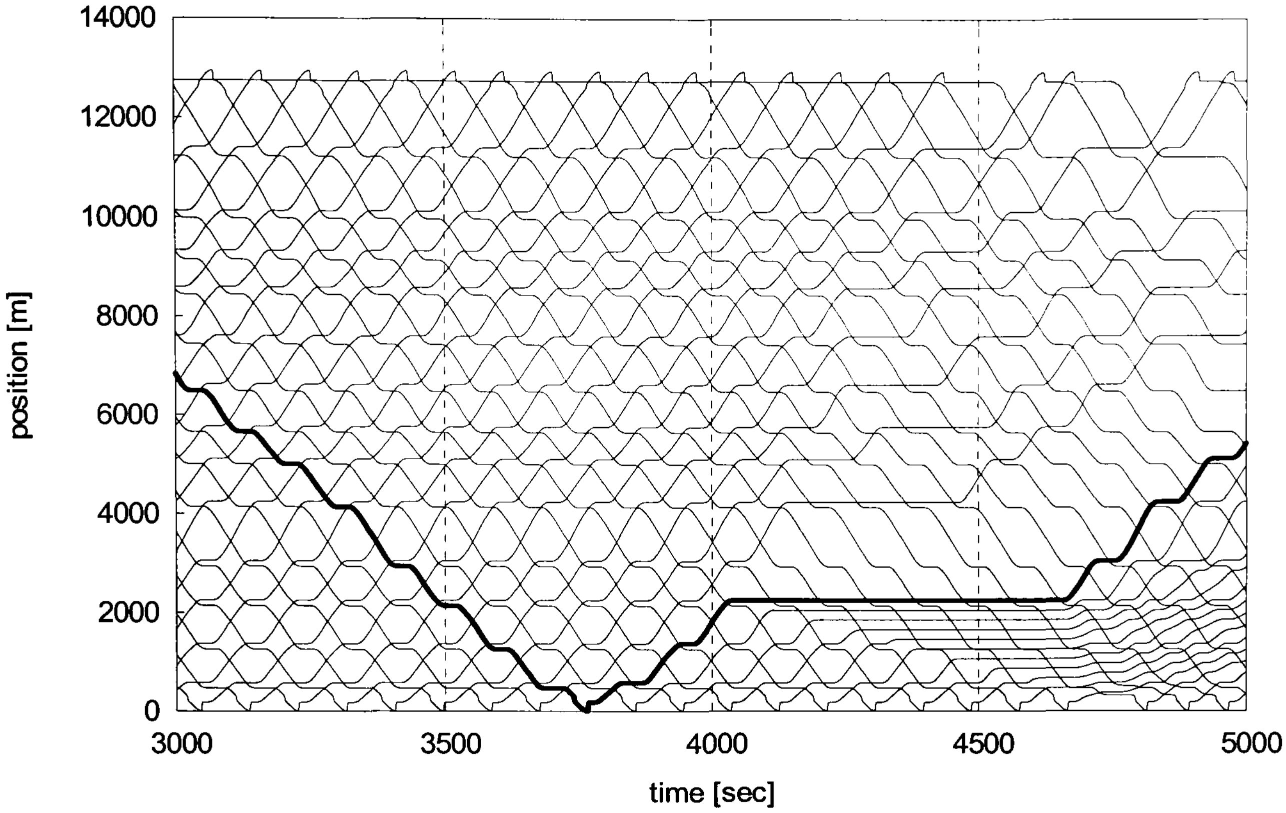

Figure 6.1: Train movements (with waiting discomfort only)

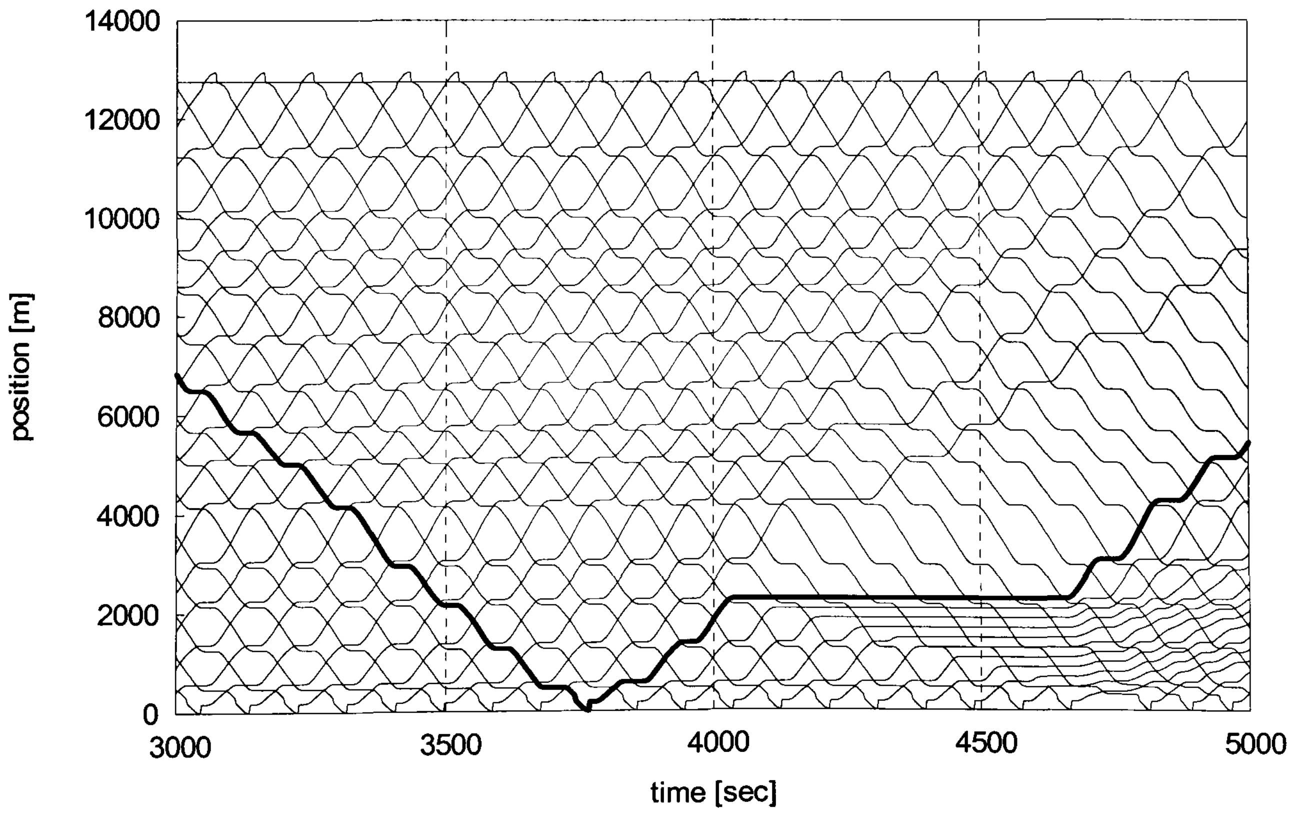

Figure 6.2: Train movements (with travelling discomfort only)

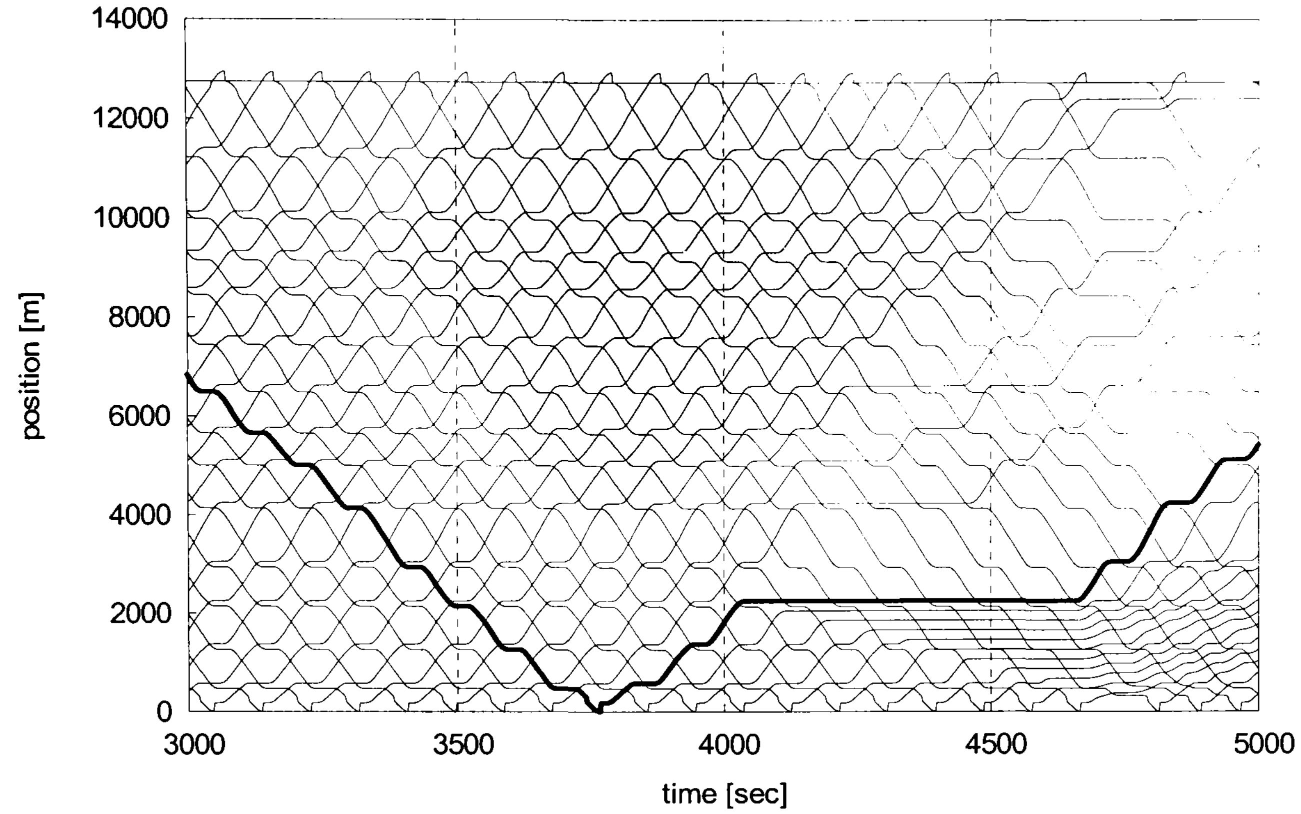

Figure 6.3: Train movements (with congestion discomfort only)

From these figures, it can be observed that trains preceding the initially disturbed train are indeed regulated in such a way that the selected objective criterion component decreases. When only waiting discomfort is sensitive to regulation (Figure 6.1), the preceding trains are held at stations for longer time than in the other cases. This is because the regulator tries to reduce waiting discomfort of the passengers at the stations ahead of the disturbance site by holding these trains there. On the contrary, when only travelling discomfort is sensitive to regulation (Figure 6.2), the preceding trains are held at stations for shorter times, since holding trains for longer time means that travelling discomfort of the passengers on these trains grows larger. When only congestion discomfort is sensitive to regulation (Figure 6.3), trains preceding the initially disturbed train are held for longer times as far as Station 8, while they leave the stations beyond

Station 8 basically on time. This is because the passenger load increases on trains bound for Station 14 prior to leaving Station 8, thus the trains have to share passengers until they leave Station 8 such that the passenger load on each train is smaller than the specified threshold.

The movements of the trains preceding the initially disturbed train change according to the sensitivity settings; however, those of the trains behind the initially disturbed train do not seem to change between the cases. Those trains leave the earlier stations one after another, despite the disturbed train being halted at Station 3. This means that the changes in the weighting for the existing three kinds of discomfort criteria do not have any effect on the movements of following trains.

6.2 Halted passengers' discomfort

In previous chapters, it has been repeatedly stated that the discomfort of passengers in a train stuck between stations is important and not negligible, but that it does not seem to have been considered in past research. The Murata regulator is not an exception; it shows a good performance in reducing the existing three kinds of passenger discomfort (waiting discomfort, travelling discomfort, and congestion discomfort). However, it regulates the disturbed traffic in such a way that the trains following the disturbed train one after another leave the station in rear of the station where the disturbance has happened, thus they have to stop behind the disturbed train. Such a situation is not acceptable in a real railway operation.

It is obvious that the outputs of the Original Murata regulator would be undesirable in a real railway operation; however, the objective function used does not include the discomfort of passengers who are stuck in a train that is halted between stations and, therefore, the regulator could not be expected to try and eliminate or reduce the time spent by trains stationary in a tunnel. In order to evaluate the performance of regulation algorithms accounting for passenger discomfort, as explained above, it is necessary to define and quantify this kind of discomfort in a manner similar to the other three kinds of discomfort.

It is also convenient to name the above stated discomfort similar to the other three kinds of discomfort. As has already been seen in the title of this chapter, this discomfort is called "halt discomfort" in this thesis. "Halt discomfort" is the discomfort felt by passengers who are stuck in a train that is halted between stations. A detailed definition is given in the next section.

6.3 Definition of halt discomfort

In order to capture the discomfort felt by passengers who are stuck in a train that is halted between stations, the halt discomfort D_h is defined as shown in function (6.2).

$$D_h = \left(\frac{\max[0, A_h - E_h]}{X_h} \right)^2 \tag{6.2}$$

where

X_h: normalising point for halt time, (see section 6.4)

E_h : expected halt time, (see section 6.4)

A_h : actual halt time.

In this definition, "halt time" is the amount of time for which a train stops completely between stations, that is to say, not at a platform. Thus, function (6.2) means that, if a train stops between stations for a longer time than the expected time (or the acceptable time before the journey resumes, in other words, a threshold time) E_h, then passengers on this train feel discomfort. The degree of this discomfort rises sharply (quadratically) if the halt time is greater than the threshold and the excess time ($A_h - E_h$) is greater than the normalising point for halt time X_h.

Here, there is one point to be noted about the normalising point. A normalising point has also been used in the definitions of waiting discomfort and travelling discomfort. In the cases of these kinds of discomfort, the normalising point is a linear function of the expected value of each parameter, that is, the expected waiting time for waiting discomfort, and the expected travelling time for travelling discomfort.

This reflects the idea that if passengers expect a longer waiting time, then a larger excess of waiting time is acceptable to them. For example, consider passengers using a line whose operational headway is 90 s and those using a line where the operational headway is 300 s. If a train arrives 600 s behind schedule, it is interesting to examine which group of passengers will get more irritated by this delay. It would be generally accepted that the first group of passengers (on a 90 s headway line) get more irritated. The same idea can be applied to travelling time.

On the other hand, the normalising point for halt time is a constant; it is not a function of the expected halt time. This is because passengers usually do not expect their train to stop at all except at a station. Even if a train stops before it reaches the next station, passengers may think that this is due to a temporary delay of the former train, thus their train will resume its journey soon. The expected halt time thus means how many seconds later a train is expected by passengers to resume its journey.

6.4 Determining constants

By observation of the behaviour of passengers in real railway systems, the author feels that it is reasonable for the experiments to set the constants in the definition of halt discomfort as follows:

$$E_h = 60 \tag{6.3}$$

$$X_h = 120 \tag{6.4}$$

This means that passengers begin to feel discomfort if their train stops between stations for longer than 60 s, and their discomfort rises sharply if they are forced to wait in a halted train for longer than 180 s.

Each objective criterion should be set so that the four kinds of discomfort can be reasonably compared with each other. The comparison of each type of discomfort is shown in Table 6.1, in which each row gives the set of conditions that results in the same value for each type of discomfort. For example, looking at the row in which each discomfort = 1, it can be seen that D_w=1 when A_w=630 s, D_t=1 when A_t=930 s, D_c=1 when A_c=2000, and D_h=1 when A_h=180 s. It can be seen that halt discomfort is set to give the sharpest increase in discomfort. This means that being halted between adjacent stations (usually in the tunnel in metro-type railways) gives passengers the largest increase in discomfort under disrupted conditions.

Table 6.1: Discomfort comparison

Discomfort	D_w ($E_w=90$)	D_t ($E_t=300$)	D_c (max=2000)	D_h ($E_h=60$, $X_h=120$)
0	$A_w \le 90$ s	$A_t \le 300$ s	$A_c \le 1600$	$A_h \le 60$ s
0.5	A_w=472 s	A_t=741 s	A_c=1883	A_h=145 s
1	A_w=630 s	A_t=930 s	A_c=2000	A_h=180 s
2	A_w=854 s	A_t=1183 s	(A_c=2166)	A_h=230 s

6.5 Evaluation with halt discomfort

In this section, the outputs of the Original Murata regulator are to be re-examined with the inclusion of halt discomfort. For the remainder of this thesis, the Original Murata regulator with the four kinds of discomfort criteria including halt discomfort, but using the same regulation algorithms as the Original Murata regulator, is called the "Murata plus halt" regulator.

Simulation conditions are the same as those in Chapter 5. Figures 6.4-6.8 are the results of the simulation.

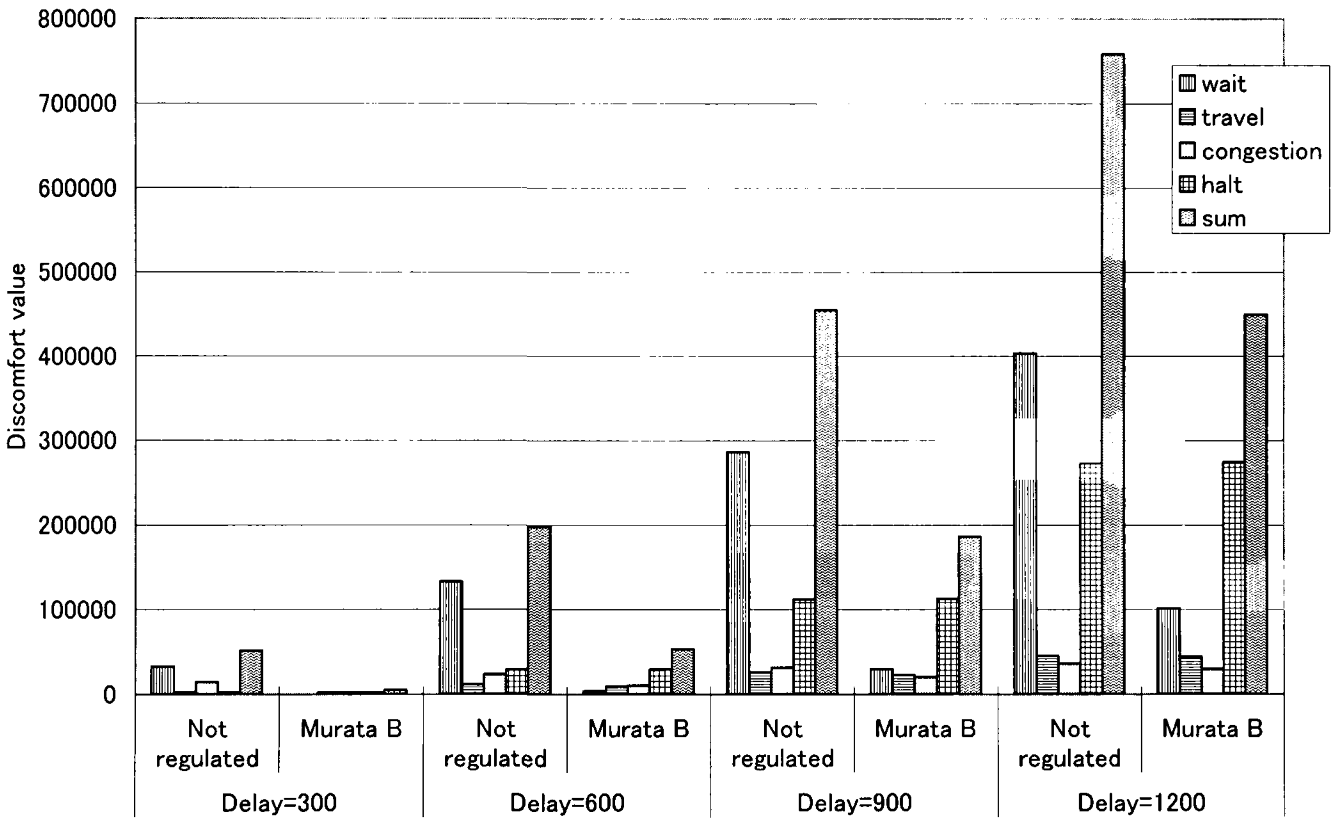

Figure 6.4: Discomfort comparison, Flat 18

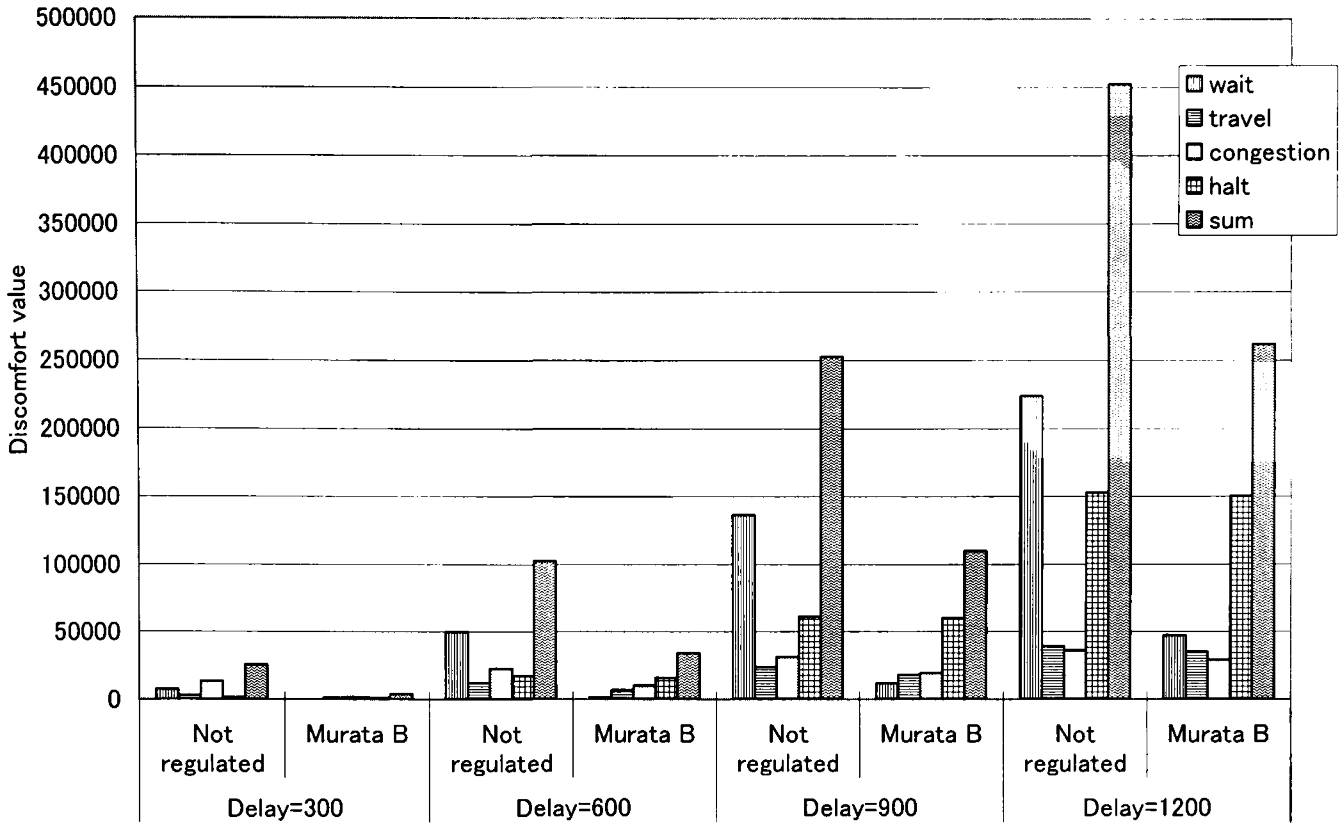

Figure 6.5: Discomfort comparison, Centre 60

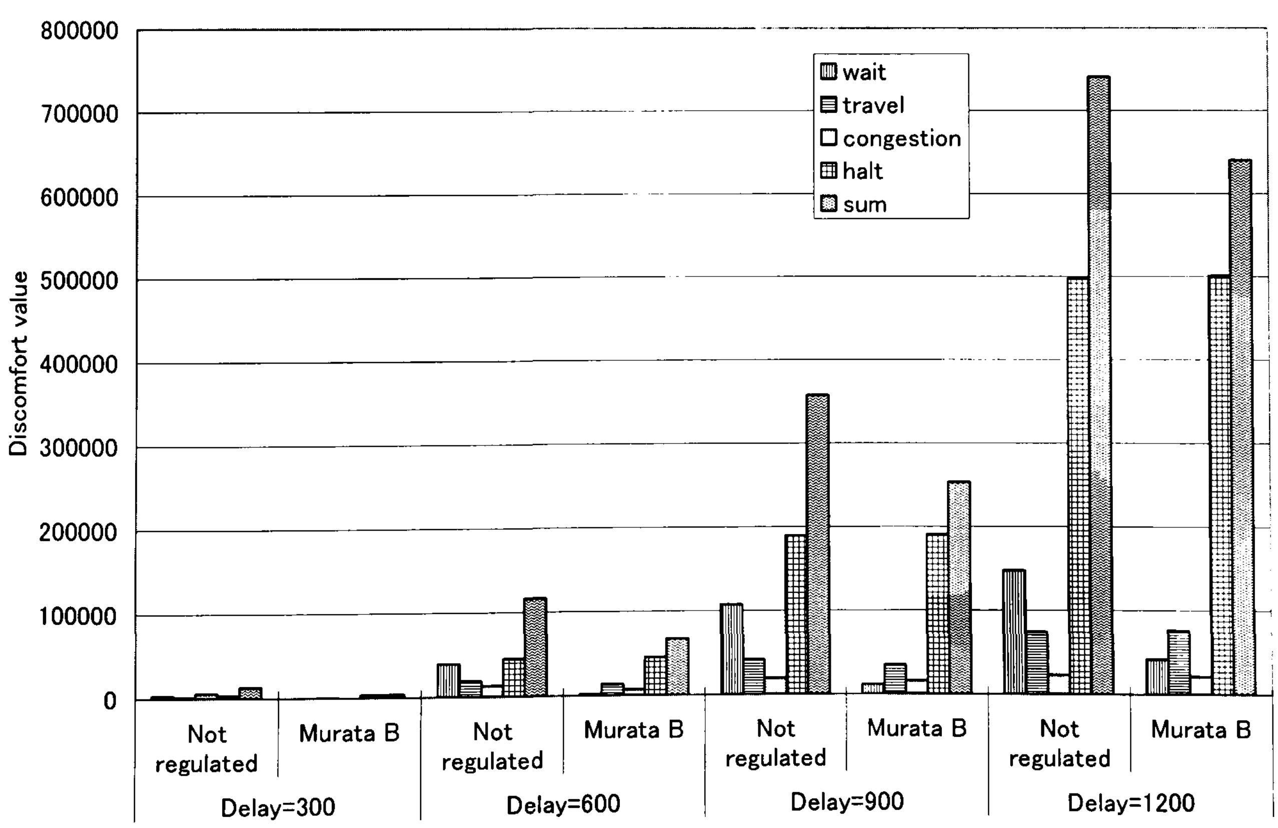

Figure 6.6: Discomfort comparison, Sub 50

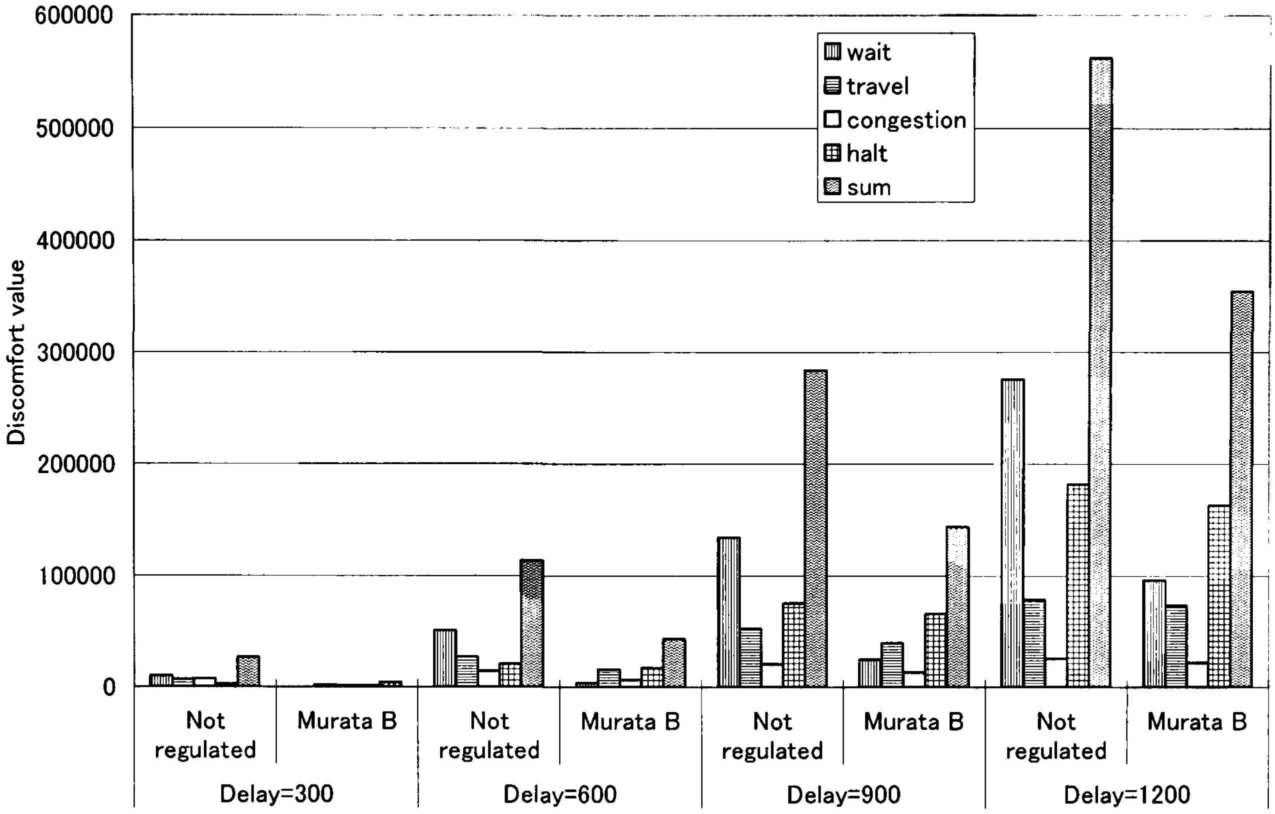

Figure 6.7: Discomfort comparison, Mid_1 70

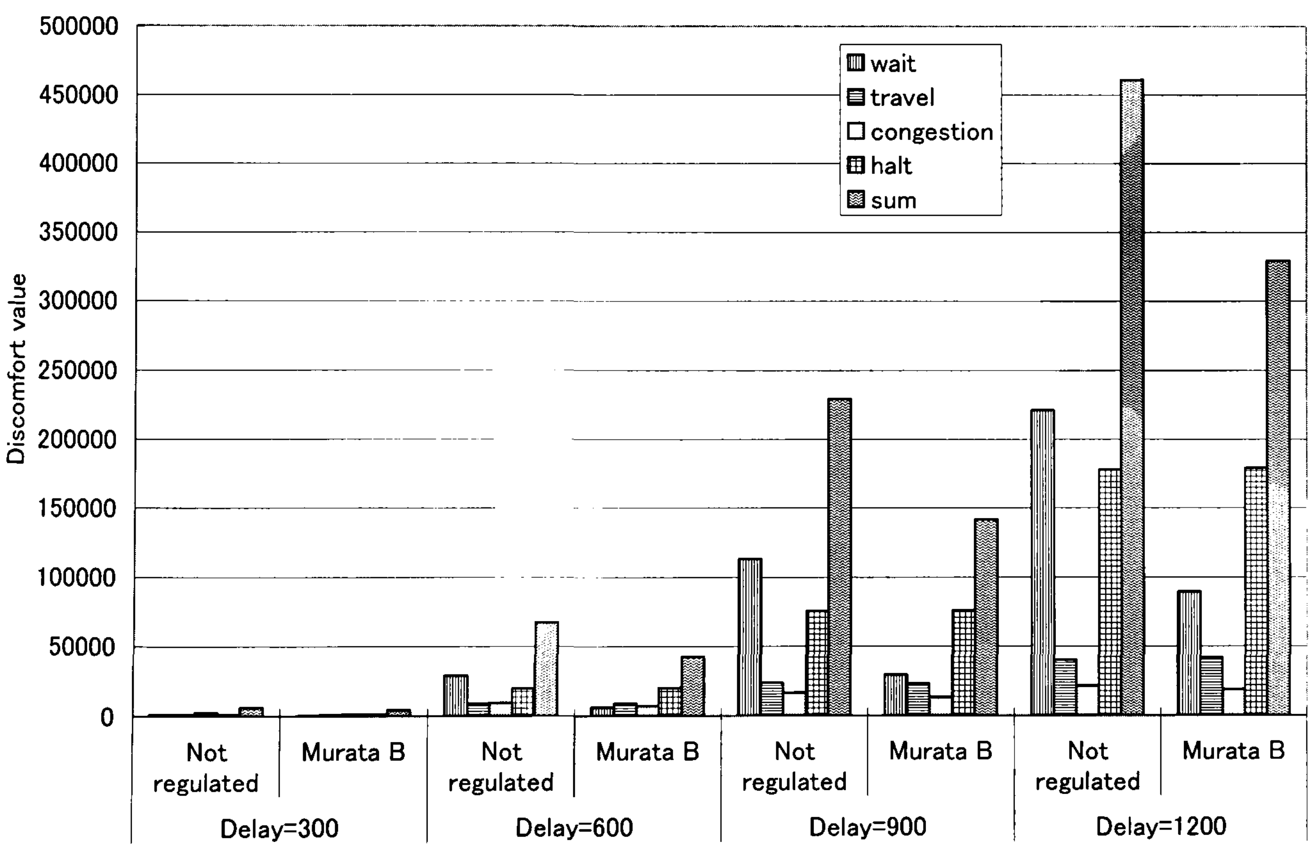

Figure 6.8: Discomfort comparison, Mid_2 90

From these results, a problem with the Murata plus halt regulator can be seen when the halt discomfort is calculated. For all OD types, the halt discomfort is almost as large as the waiting discomfort when regulation is not performed. When the Murata plus halt regulator is used, the existing three kinds of discomfort values decrease, but the halt discomfort does not.

The problem of the halted trains behind the disturbed one should not be neglected in disturbed traffic, but the Murata plus halt regulator does not act to reduce the halt discomfort. Thus, a new regulation algorithm is necessary in order to deal with this problem. Before introducing a new algorithm, the next chapter describes techniques used by real operators in an attempt to attenuate the halt discomfort problem.

7 The "stop-all-trains-at-once" strategy

7.1 Problem of the Original Murata and the Murata plus halt regulators

As has been stated in Chapter 5, the Original Murata regulator shows a good performance in reducing three kinds of passenger discomfort (waiting discomfort, travelling discomfort, and congestion discomfort). However, there is a problem in the train movements, that is the situation where trains following the disturbed train one after another leave the previous station for the station where the disturbance happens, thus having to stop behind the disturbed train.

Considering the situation in a real railway operation, it is hard to accept that such a situation should arise as a result of the regulation. It must be one of the top priorities for railway operators to prevent trains from being stuck between stations, since passengers will feel severe discomfort when their train is stuck between stations, particularly in the case of underground railways. Quite a few metro-type railways are underground lines, which means that the halted passengers would feel much more frustrated and even alarmed in a tunnel.

Two underlying reasons for this problem have already been stated in 5.3. One is that the objective criteria of the Original Murata regulator do not include the discomfort of passengers on a train that is halted between stations. Chapter 6 has been devoted to introducing a new objective criterion that describes this discomfort, as well as introducing the Murata plus halt regulator and re-examining the performance of the idea of Murata, which is inevitably poor in reducing the discomfort of passengers on halted trains.

The other reason is that it is difficult to predict when the disturbed train will resume its journey, particularly just after the disturbance has happened. In order to avoid this problem, the Original Murata regulator assumes that the disturbed train resumes immediately. This is obviously undesirable, considering that it will lead to following trains having to stop behind the disturbed train.

7.2 Evaluation with halt discomfort

In this section, the outputs of the "stop-all-trains-at-once" strategy are examined. The procedures constituting this strategy are as follows:

(1) All trains are suspended in the nearest station when the disturbance happens. If a train cannot enter the next station since all the platforms are occupied, then this train inevitably stops outside the station.

(2) The trains are restarted later when the disturbed train resumes travelling.

In order to implement the "stop-all-trains-at-once" procedures in the simulator (OOMTS, *cf.* Figure 2.3), a new regulator module has been developed, incorporating the above-mentioned procedures. The new regulator obtains the line status from the line model, and, if a disturbance is reported from the line model, the regulator module changes the departure times of trains on the line so that all the trains stop at stations (if possible) until further notice is reported from the line model that the disturbance has been removed. A brief flowchart is shown in Figure 7.1.

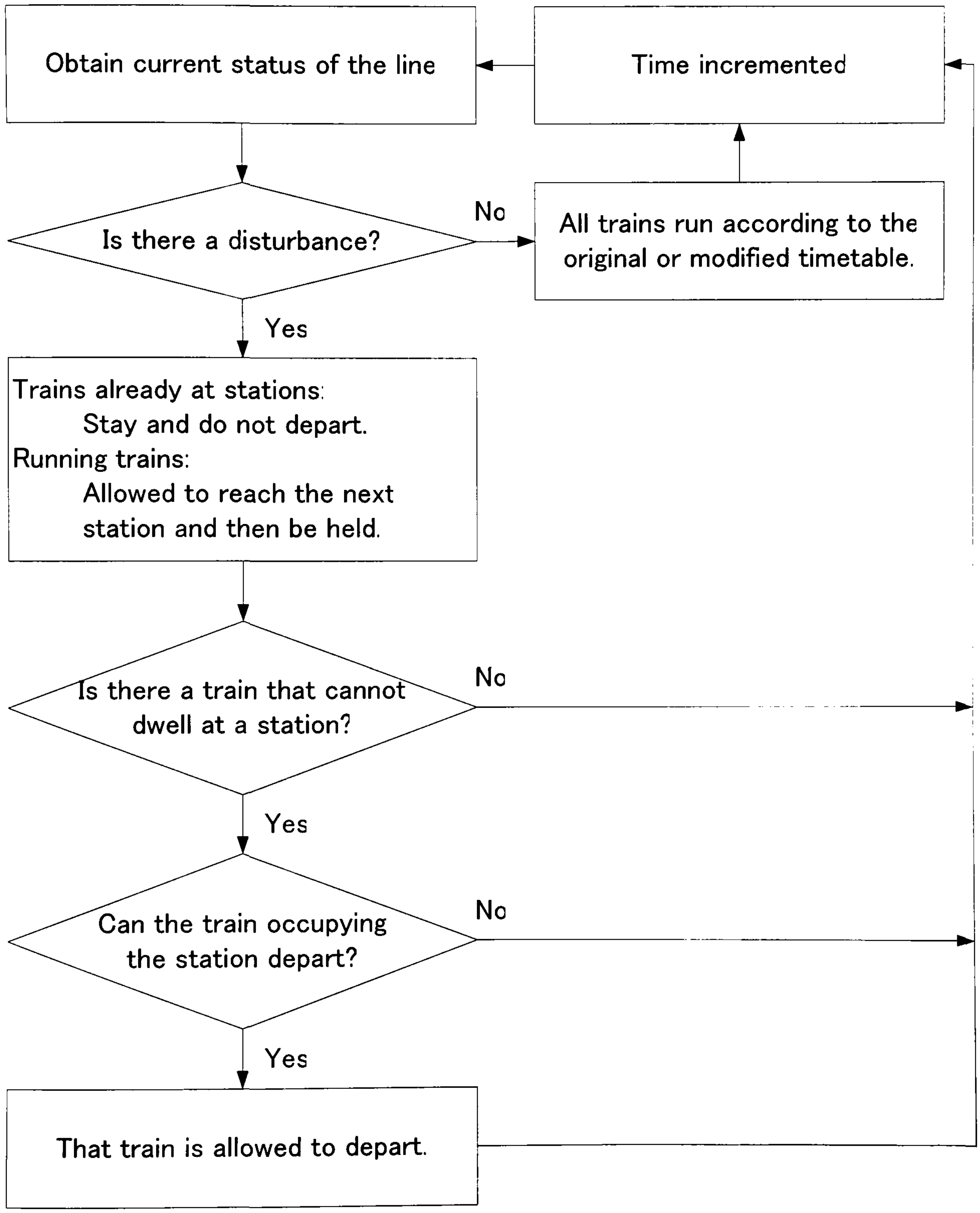

Figure 7.1: Flowchart of the "stop-all-trains-at-once" procedure

Simulation conditions are the same as those in Chapter 4. Figures 7.2-7.6 are the results of the simulation.

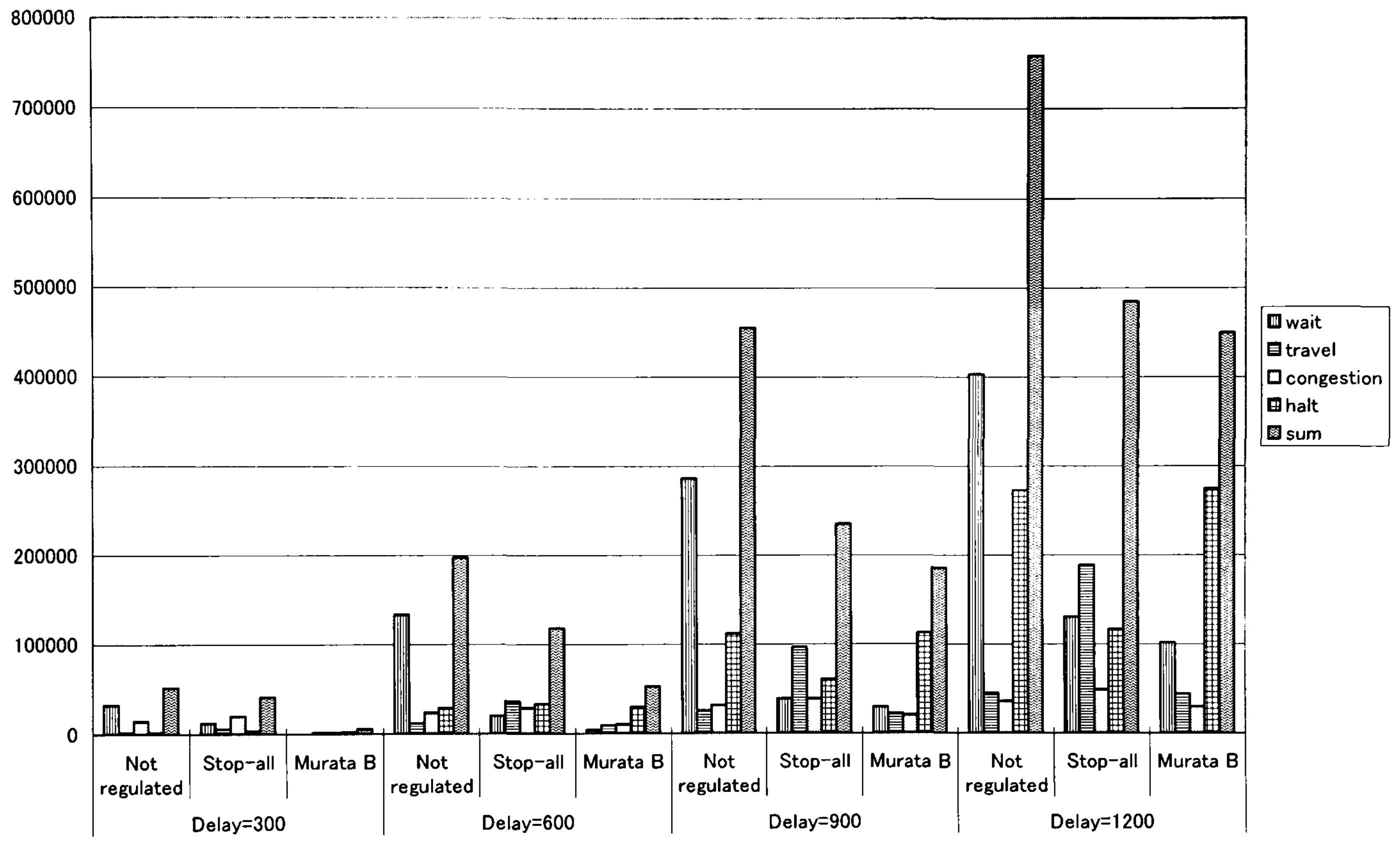

Figure 7.2: Discomfort comparison, Flat 18

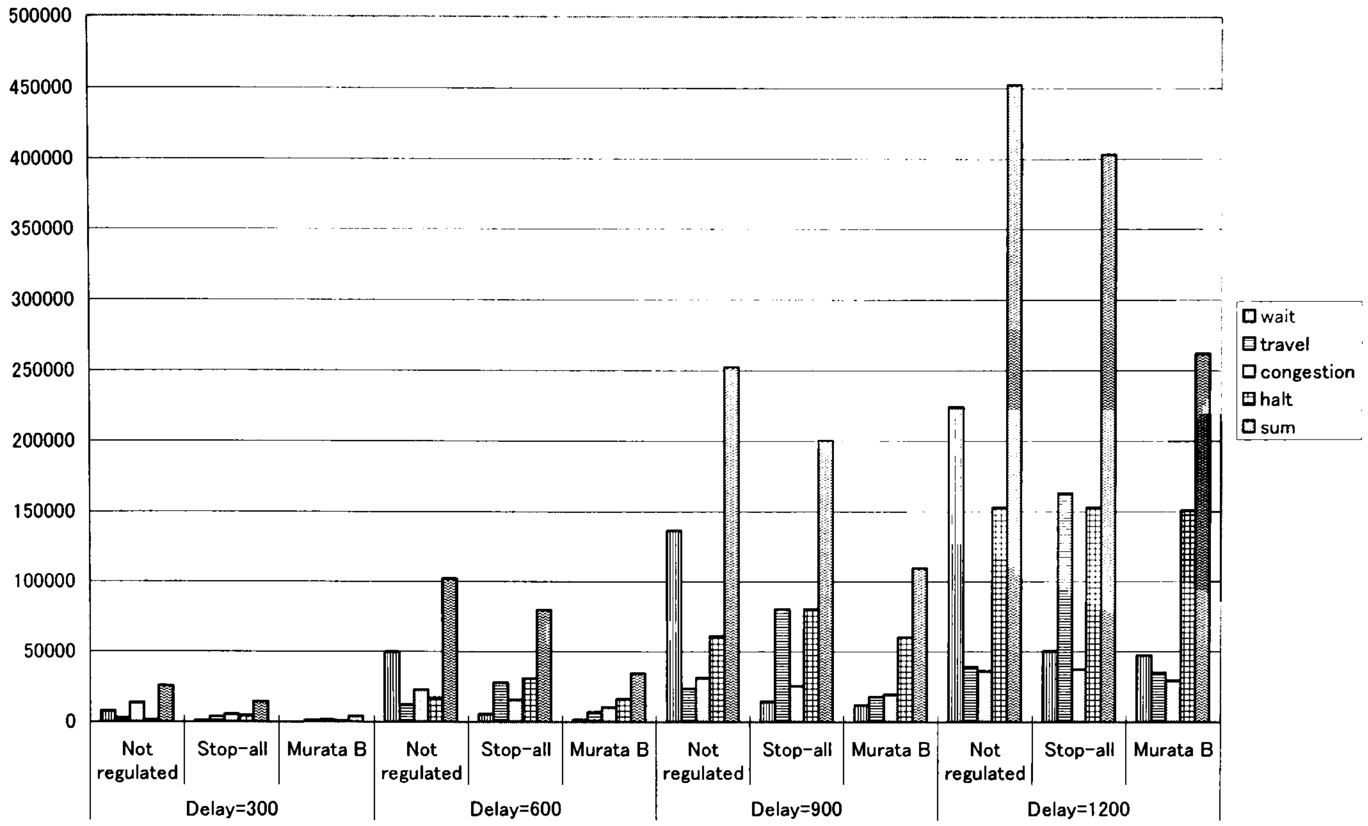

Figure 7.3: Discomfort comparison, Centre 60

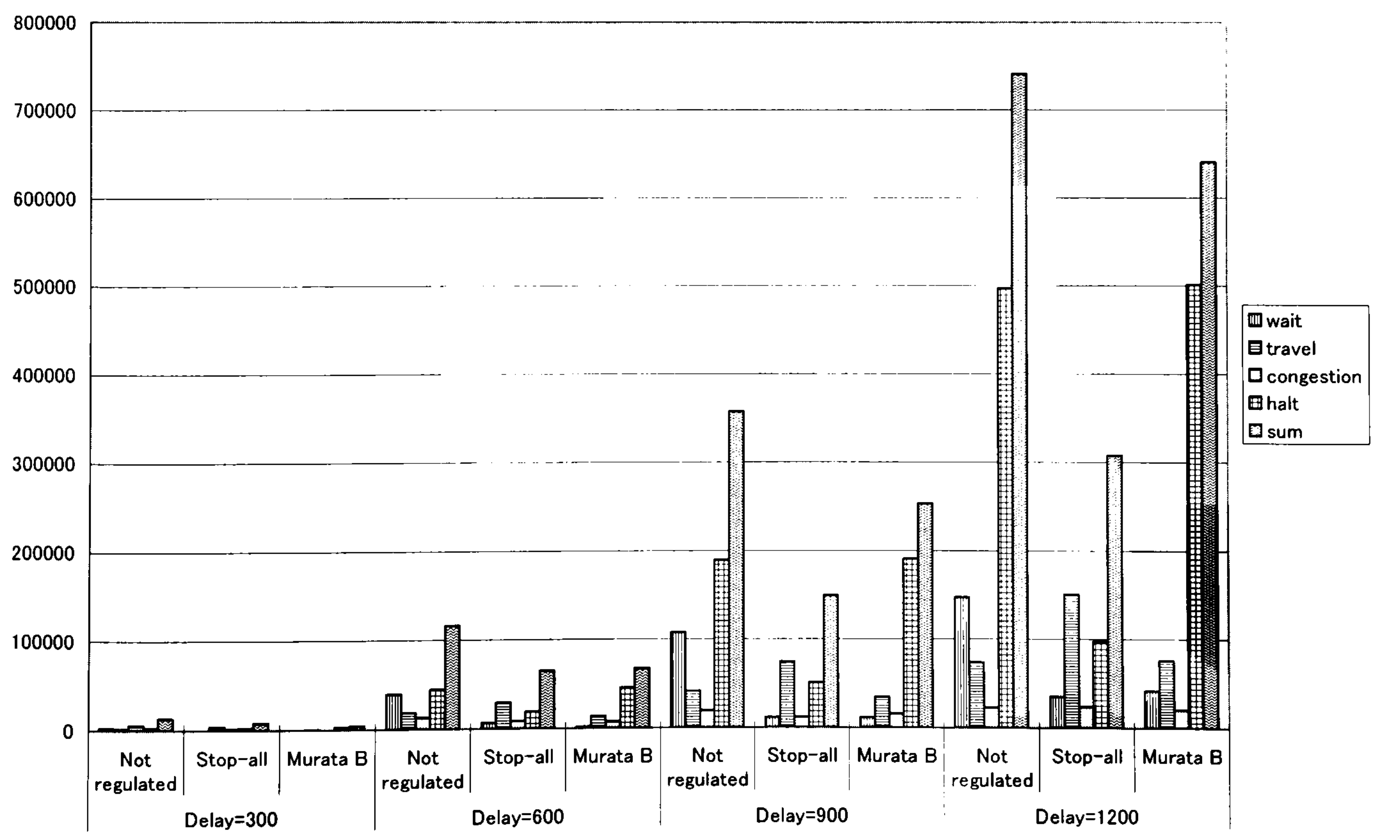

Figure 7.4: Discomfort comparison, Sub 50

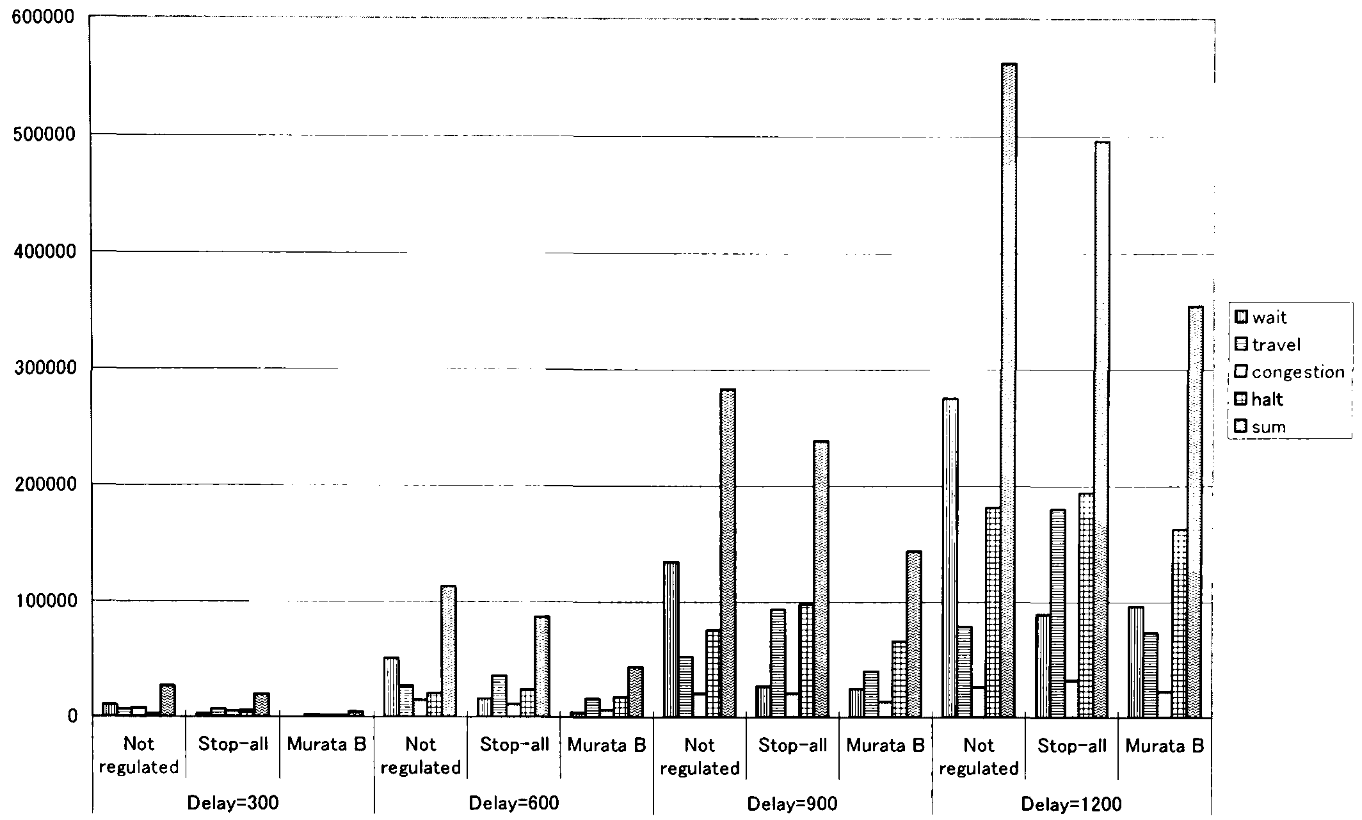

Figure 7.5: Discomfort comparison, Mid_1 70

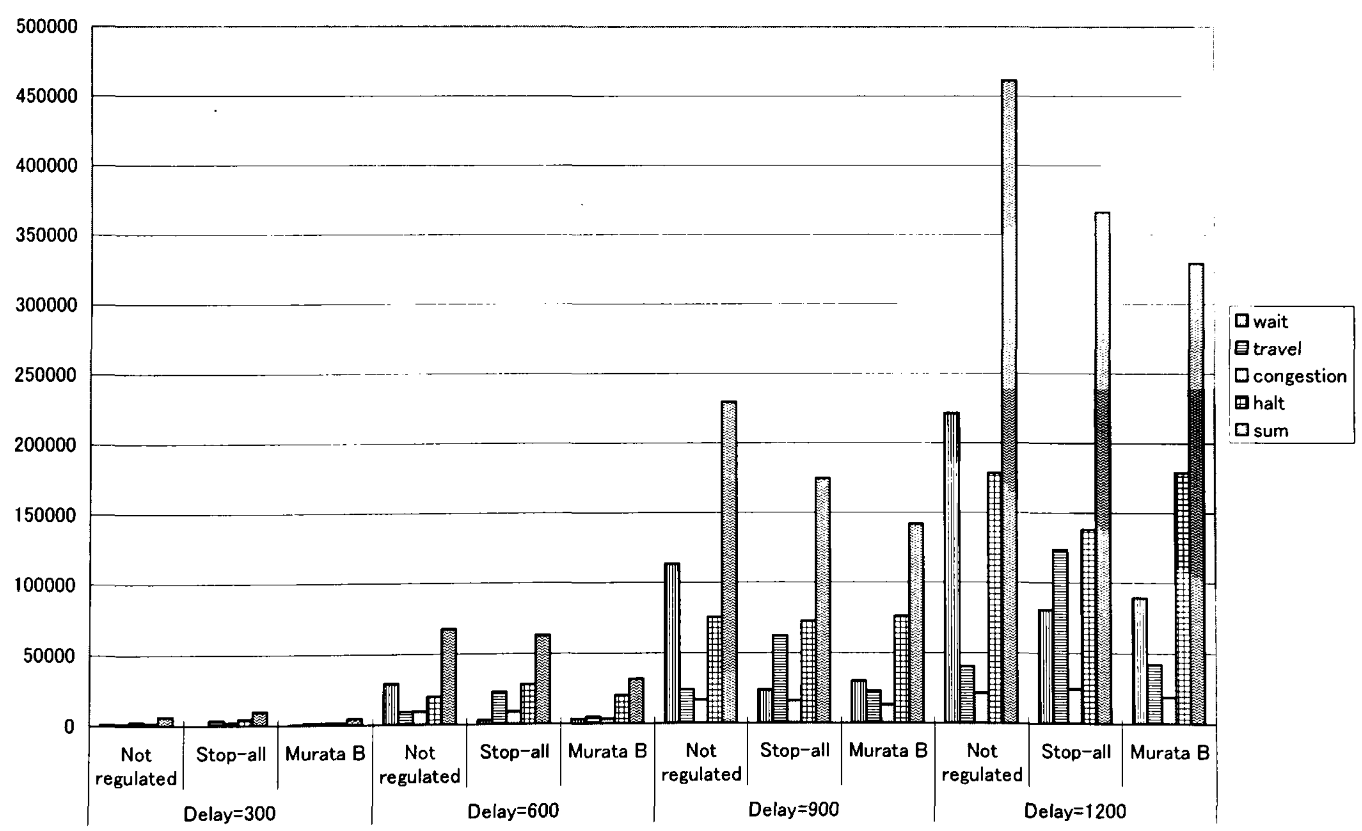

Figure 7.6: Discomfort comparison, Mid_2 90

From these results, it can be seen that the Murata plus halt regulator does not improve halt discomfort (as has been shown in Chapter 6), while the "stop-all-trains-at-once" strategy reduces halt discomfort in some cases. In total, the Murata plus halt regulator shows a better performance than the "stop-all-trains-at-once" strategy, except in the case of "Sub 50" OD, where the "stop-all-trains-at-once" idea reduces halt discomfort greatly, thus the total value of discomfort is much smaller than that for the Murata plus halt regulator. This is basically due to the fact that the number of passengers on the halted trains following the disturbed train is larger in the case of "Sub 50" OD, thus avoiding trains trailing the disturbed train from being stuck between stations contributes to reducing halt discomfort greatly.

A more detailed discussion is presented in Chapters 9-11, with the "phase-by-phase" regulation mentioned in section 7.3 being fully described in Chapter 8. The results of the regulation methods including the "phase-by-phase" regulation are discussed in Chapters 9-11.

What can be concluded here is that the Murata plus halt regulator basically has no effect on halt discomfort, while the "stop-all-trains-at-once" strategy is, although it seems crude, expected to be effective in reducing halt discomfort.

However, it should be noted that the "stop-all-trains-at-once" strategy is not always successful; the performance of this technique depends on parameters such as the OD types, the disturbance position, and so on. Also, the "stop-all-trains-at-once" strategy worsens the existing discomfort values in most cases, as presented in this Chapter. This may be due to the fact that this idea simply stops all trains, without regard to the passenger discomfort when they have to wait, even if the disturbance happens far away from their own position. In this sense, the "stop-all-trains-at-once" strategy is too crude.

Still, the "stop-all-trains-at-once" strategy does reduce halt discomfort greatly, thus it seems effective to introduce this technique into the new regulation method, which aims at reducing halt discomfort significantly, while not deteriorating the other kinds of discomfort too much.

7.3 Necessity of phase-by-phase regulation

One of the reasons why a human operator would use such a conservative method as the "stop-all-trains-at-once" strategy is that, just after the disturbance happens, it is difficult to predict exactly how long it will last. If he/she mistakenly predicts that the disturbance will be removed soon and the trains can resume, and he/she gives permission for trains to leave stations, but the disturbance still exists and the disturbed train cannot move, then trains that have set off will consequently be stuck. (It is also essentially impossible to bring trains back to their previous stations for safety reasons.)

As time goes by, however, a human operator will be receiving more and more information about the situation of the disturbance. This means that he/she may be able to predict the duration more and more accurately when a longer time has passed since the disturbance was recognised. He/she can determine how to resume operations depending on the situation; if the disturbed train is expected to resume soon, then the following train is allowed to leave immediately, since the following train will be able to enter the next station. If the disturbed train is not expected to restart soon but a certain amount of time later, then the next train cannot restart but the trains further backward may resume, and so on.

From the above considerations, it seems a good way of handling disturbances would be to combine the "stop-all-trains-at-once" strategy and the existing regulation method. In the first phase, just after the disturbance happens, the "stop-all-trains-at-once" strategy seems appropriate so as not to increase the discomfort of the passengers as a result of being stuck in halted trains. Then, as the circumstances of the disturbance gradually become clearer, and a human operator could predict the time of restarting of the disturbed train, a second phase of regulation could be entered. In this phase, the ideas of Murata may work well, since the difficulty of the prediction has now receded, thus the optimisation algorithm will work. There is still a problem of the accuracy of the prediction, particularly in the early stages, since the information regarding the disturbance still has uncertainty. Thus the prediction may change later when more detailed information is obtained.

In order to implement this overall strategy for the traffic control system, it is important to clearly define the "phases" and the "procedures" to be taken in each phase. With such detailed consideration, the idea of a "phase-by-phase" approach would appear to offer improved functionality. This will be explained in Chapter 8.

7.4 Conclusion

This chapter described the problems associated with the Original Murata and the Murata plus halt regulators, then introduces the "stop-all-trains-at-once" strategy from the observation of the real railway operations, and measures the performance of these regulations, including halt discomfort. The combination of the idea of Murata and the "stop-all-trains-at-once" strategy would appear to be effective in reducing halt discomfort, while not deteriorating the other kinds of discomfort so much.

8 New "phase-by-phase" regulation

8.1 Necessity of prediction

A basic problem in the management of a disturbed railway is that it is difficult to predict exactly how long the disturbance will last, particularly just after the disturbance has happened. The Original Murata and the Murata plus halt regulators avoid this problem by assuming that the disturbed train will resume immediately; however, this assumption is one of the underlying reasons that cause the regulator to generate outputs requiring trains to stop behind the disturbed train. The "stop-all-trains-at-once" technique does not require prediction of the resumption of the disturbed train, thus it does not suffer from this problem. However, this technique forces all trains to suspend their journeys, no matter how far away from the disturbance site a train is. In this sense, the technique is too crude.

The solution to this problem is, though it seems contradictory, to try to predict the length of the disturbance as accurately as possible. Though it is impossible to predict exactly when the disturbed train will resume immediately after the disturbance happens, the situation of the disturbance becomes clearer and clearer as time goes by. Thereby, a human operator in the control centre obtains more detailed information about the disturbance, and then can predict the time of resumption more accurately. Thus, the regulation can be performed on the basis of the latest information (or, the latest prediction).

This is one of the reasons for inventing a new "phase-by-phase" regulation, which will be explained in 8.4.

8.2 Application of the "stop-all-trains-at-once" idea

In order to avoid a sudden increase in halt discomfort, it is important to make a quick decision to suspend the trains following the disturbed train immediately after the disturbance has started; however, this is intrinsically difficult as long as the regulation depends on prediction, since it is difficult to predict exactly how long the disturbance will last, particularly straight after the disturbance has happened. To avoid this problem, a regulation method that does not need prediction is necessary.

The "stop-all-trains-at-once" idea, which has been introduced in Chapter 7, gives a pointer as to how to solve the problem, since it does not need prediction, while it shows a good performance in reducing the halt discomfort greatly. By applying this technique to the situation just after the disturbance has happened, when prediction is not available, it is expected that halt discomfort will decrease greatly.

Even though the technique is effective in reducing halt discomfort, it is not a good strategy to prolong its application, since it forces any train to stop no matter where it is, and may thus worsen the other kinds of discomfort. What should be a good strategy is to apply this idea immediately after the disturbance happens, then to withdraw it after a certain period.

This is another reason for introducing the new "phase-by-phase" regulation.

8.3 Enhanced Murata plus halt regulator

The newly developed "phase-by-phase" regulation incorporates the basic idea of Murata, though the Original Murata regulator has to be modified to incorporate a term for halt discomfort in the objective criteria. This modified regulator is called the "Enhanced Murata plus halt regulator".

The Enhanced Murata plus halt regulator includes an approximate calculation of the partial derivatives of halt discomfort. The procedure is simple; the change in halt discomfort is calculated by assuming that the increase in the running time between a pair of adjacent stations is the time for which a train stops in this section, in other words, the halt time of this train. In reality, the increase in the running time is not necessarily the time for which a train stops, thus the Enhanced Murata plus halt regulator overestimates the effect of halt discomfort. However, it still generates outputs that lead to trains bunching behind the disturbed train, since it still depends on the assumption that the disturbed train will restart immediately. When the new "phase-by-phase" regulation is introduced, this assumption is no longer made.

8.4 New "phase-by-phase" regulation

From the considerations discussed above, the following new regulation algorithm is proposed. This algorithm defines the procedure to be taken in each of several "phases", which describe the evolving situation after a disturbance happens, thus it is called "phase-by-phase" regulation. The sequence is shown in Fig. 8.1, and explained below.

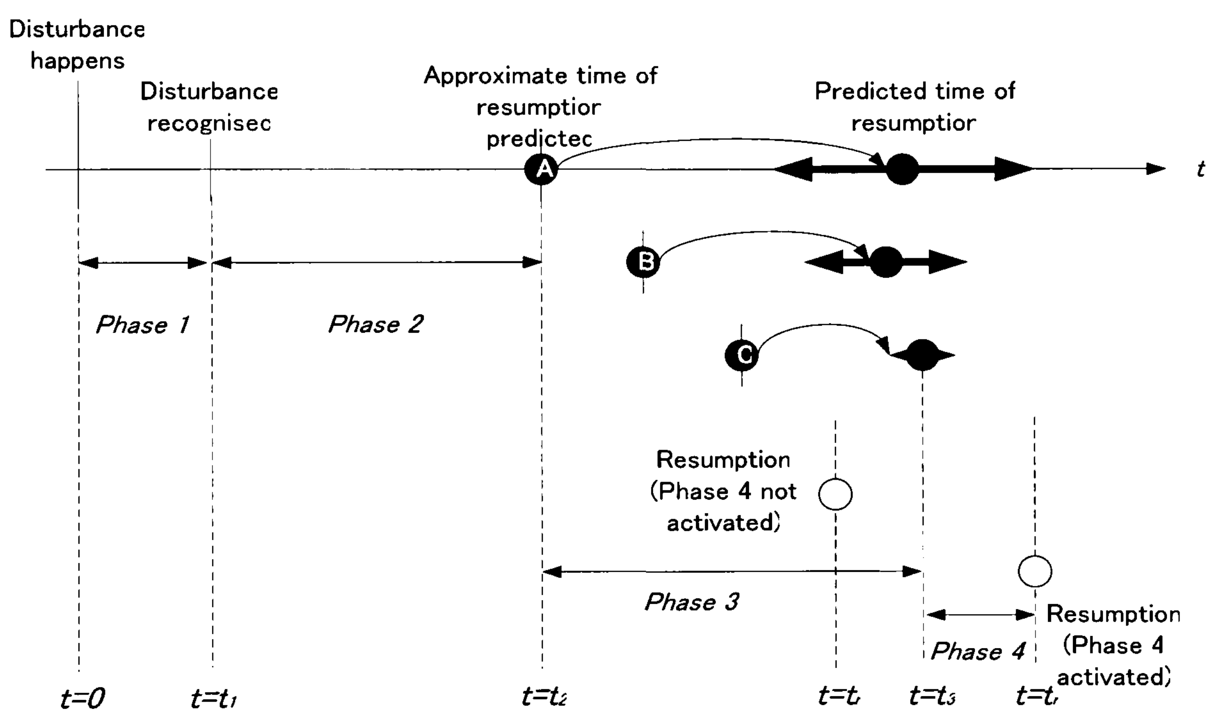

Figure 8.1: Time chart of the phase-by-phase regulation

8.4.1 Phase 1

This phase is from the occurrence of a disturbance to the recognition of it by the traffic control system.

During this phase, the Enhanced Murata plus halt regulator operates. Since the disturbance is not recognised by the traffic control system, the regulator operates on the assumption that trains standing at platforms after the scheduled departure time has passed leave the station immediately. This means that there are no differences between the decisions made by the Enhanced Murata plus halt regulator and the Original Murata regulator in this phase, since there is no prediction of the length of disturbance, and thus there is no predicted increase in the running time in any sections in the line.

8.4.2 Phase 2

This phase is from the recognition of a disturbance by the traffic control system to the time when an approximate time for the resumption of the delayed train's journey is predicted.

In this phase, it is assumed that the initially disturbed train cannot depart for as long as the maximum running time from any one station to the next station. On this assumption, a train is not allowed to depart from a station if, by so doing, it would have to stop before reaching the next station as this would still be occupied by the preceding train. Otherwise, a train obeys the output of the Enhanced Murata plus halt regulator.

This assumption has been made to prevent trains having to stop before reaching the next stations. In other words, this phase adopts the "stop-all-trains-at-once" technique, aimed at preventing halt discomfort undergoing a sudden increase after a disturbance happens. It should be noted that, if Phase 1 is long, then the "stop-all-trains-at-once" idea becomes less effective, since the order to stop all trains is issued late and some trains will have already left stations, which means that those particular trains contribute to increasing halt discomfort. This explains the importance of an early detection of the disturbance.

The duration of Phase 2 is assumed to be much longer than the maximum running time from any one station to the next station.

8.4.3 Phase 3

This phase is from the time when the first time of resumption is predicted to the final predicted time of resumption.

The Enhanced Murata plus halt regulator is applied during this phase. The following two points should be noted here.

(1) Generally, at the early stages, the prediction of the delay to resumption is subject to errors. For example, suppose the train is predicted to resume at 300 ± 100 s after the regulation enters Phase 3. However, the centre value (in this case, 300 s after the regulation enters Phase 3) has to be used for the regulation. In most circumstances, the prediction steadily becomes more and more accurate as time proceeds, moving to 300 ± 50 s, and finally, 300 s, for example. Likewise, the centre value is subject to systematic shift, i.e., 300 ± 100 s $\rightarrow$ 320 ± 50 s $\rightarrow$ 310 s, for example.

(2) If the disturbed train resumes at the predicted time of resumption or before that, the Enhanced Murata plus halt regulator continues to operate. On the

other hand, if the disturbed train does not resume until the predicted time of resumption, then Phase 4 is activated.

8.4.4 Phase 4

This phase is from the predicted time of resumption to the actual resumption.

In this phase, the train is expected to resume but does not. This implies that the train is expected to resume soon, which is the same assumption as that of the Enhanced Murata plus halt regulator. Thus, the Enhanced Murata plus halt regulator can be applied here, as it assumes that the disturbed train will resume straight away.

If a train cannot depart because another cause of disturbance is found, then this situation should be regarded as a new disturbance, and the regulation goes back to Phase 2.

8.4.5 Error in prediction

In Figure 8.1, predicted times of resumption are shown as black circles with two arrows pointing in opposite directions, to represent error bars. This reflects the idea that the prediction always has a certain range of ambiguity. When the first prediction is obtained (shown as the black circle "A" in Figure 8.1), the length of the arrows is the longest. Then, when more accurate predictions are obtained (shown as the black circles "B" and "C" in Figure 8.1), the length of the arrows becomes shorter.

It should be noted here that, even though there is a certain range of ambiguity in prediction, this ambiguity is not explicitly used in the regulation, since only the predicted time value is used. What should be emphasised is that the latest prediction is incorporated in the regulation calculations.

There may be an argument that the information regarding the error in the prediction should be used in the regulation. If a sufficient amount of information is stored, for example, there is a database of past accidents and the time to resume, it may be possible to calculate a better predicted time of resumption in a statistical sense. For example, suppose that there is a disorder in the signalling facilities of a certain station. From the past records of similar cases, it is revealed that the average repair time of this kind of failure is 600 s, with a standard deviation of 100 s. In this case, if the predicted time is set in the range of 600 ± 100 s, then the actual time of resumption is within this range with a probability of 68 per cent. Alternatively, if the predicted time is set in the range of 600 ± 200 s, then the actual time of resumption is within this range with a probability of 95 per cent.

Considering real railway operations, this kind of statistical approach may be one way of establishing a better regulation system; however, this approach is not applied in this research; instead, just a spot prediction is applied. One reason is that there are still several uncertainties in the above stated approach; for example, when the average repair time and the standard deviation are given, should the average value be applied in the regulator, or the most future value (i.e., the average repair time + the standard deviation) be applied? This needs further discussion to resolve these issues, and should be the subject of further study.

8.4.6 Implementing in simulator

The simulator used in this research (OOMTS, *cf.* Figure 2.3) was further modified to implement the new "phase-by-phase" regulation procedures. A modified regulator module was developed, incorporating the above-stated procedures to replace the regulator module referred to in 7.2. Instead of considering disturbances, the regulator module performs the optimising calculations, on the basis of the information from the line model such as the passenger flow, the passenger load on trains, etc.

8.4.7 Example of the output of the new regulator

Figure 8.2 is an example of the output of the new regulator in the form of a time-distance graph. The OD type used is "Flat 18", and the initial disturbance is 600 s, occurring at 4000 s.

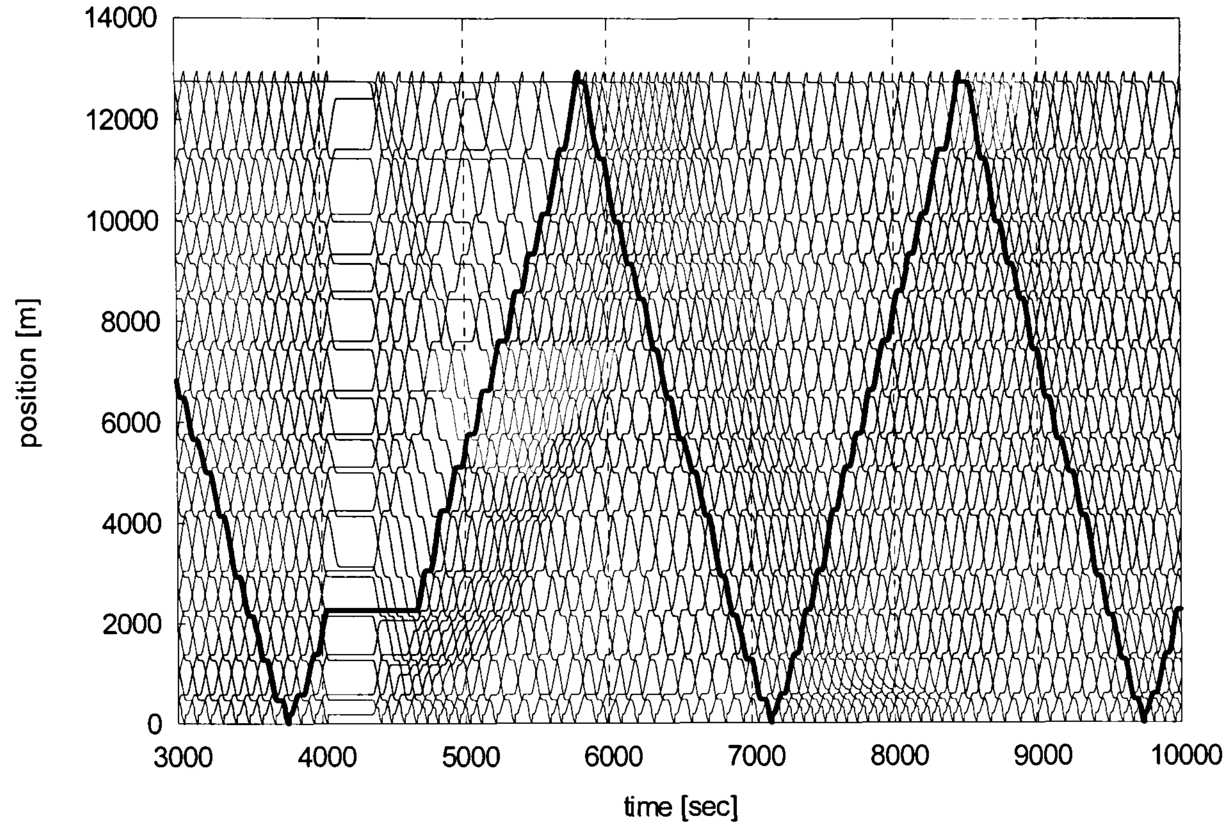

Figure 8.2: Example of the output of the phase-by-phase regulation

In this figure, the x-axis is the time elapsed from the departure of the first train from the first station, and the y-axis is the distance from the first station. The heavy line shows the initially disturbed train.

It can be seen in this figure that most trains stop at the nearest station for about 300 s from the time of about 4000 s due to the regulation in Phase 2. Some trains cannot dwell at stations due to the fact that the next station is occupied by another train. In turn, this train cannot depart for the next station because the next station is occupied by yet another train, and so on. Despite this situation, it can be confirmed that the bunching problem of the Murata regulation does not appear behind the initially disturbed train, compared with the diagram shown in Figure 5.7.

Due to the observation above, it is expected that the halt discomfort decreases much, while the other discomfort criteria do not increase much (or even decrease). The next three chapters are devoted to the discussion of the simulation results under several conditions.

8.5 Conclusion

This chapter has discussed the necessity of prediction for the regulation of disturbed railway traffic and the possibility of application of the "stop-all-trains-at-once" idea. Then the Enhanced Murata plus halt regulator has been defined so that the idea of Murata includes the calculation of the partial derivatives of halt discomfort. Though the Enhanced Murata plus halt regulator overestimates the effect of halt discomfort, it is not yet enough to output orders to prevent trains bunching, and thus a new phase-by-phase regulation, with a mixture of the "stop-all-trains-at-once" idea and the Enhanced Murata plus halt, has been introduced and explained in detail.

Though further study seems necessary regarding the statistical approach to the estimation of errors in prediction, the new phase-by-phase regulation is expected to show a better performance than the existing regulation methods. This will appear in the following chapters.

9 Results of new regulation (1) --- numbers of trains and stations

9.1 Introduction

Chapters 9, 10 and 11 present the results of the new phase-by-phase regulation, as well as results without regulation, results for the stop-all-trains-at-once technique, and results from the Murata plus halt regulator.

Each chapter is devoted to the explanation of the results from a different perspective. In this chapter, the results are examined with different relationships between the number of trains running on the line and the number of platforms. Chapter 10 describes the differences in the outputs for various OD types. Finally, Chapter 11 is devoted to the discussion assuming various levels of accuracy concerning the prediction of when a disturbed train will resume.

Before proceeding to the explanation of each case, it may be useful to explain why these perspectives are adopted.

The first factor to be considered is the relationship between the number of trains running on the line and the number of platforms. This factor is expected to have a significant effect on the performance of the stop-all-trains-at-once technique. When there are more trains on the line than available platforms, some of the trains that are ordered to stop cannot stop at platforms but have to stop between stations, which will lead to a greater halt discomfort. On the other hand, when there are less trains on the line than available platforms, and trains are ordered to stop at the nearest station, there must be some stations that are not served, thus passengers waiting at those stations have to wait longer, which will lead to a greater waiting discomfort.

The phase-by-phase regulation incorporates the method of stopping-all-trains-at-once in Phase 2, thus it is necessary to examine whether there is any effect of this factor on the performance of the new phase-by-phase regulation.

It should be noted that the new regulation does not simply order the trains on the line to serve any unoccupied stations. It runs the optimising calculation based on the original concepts of Murata in an attempt to minimise the discomfort criteria, even within the limited freedom of train movements. Thus, the new regulation is expected to show a better performance than a simple method like the stop-all-trains-at-once idea, as well as a better performance than the Murata plus halt regulation.

The second factor is the differences in the types of OD (Origin-Destination). From the results discussed in Chapter 5, no significant difference is observed in the performance of the Original Murata regulator for different OD types; however, after halt discomfort has been included in the objective criteria, and the stop-all-trains-at-once technique has been adopted, some differences are observed. For example, the Murata plus halt regulator shows a better performance than the stop-all-trains-at-once technique with the

Centre 60 OD (Figure 5.3), while the relationship between them is reversed with the Sub 50 OD (Figure 5.4). The two regulation methods have nearly the same performance with the Flat 18 OD (Figure 5.2).

The new phase-by-phase regulation is the mixture of the ideas of Murata and the stop-all-trains-at-once, thus the differences in the OD types may have an effect on the performance of the new phase-by-phase regulation.

The third and last factor relates to the accuracy of prediction concerning when a disturbed train will resume its journey i.e. restart. The new phase-by-phase regulation uses a predicted time of resumption of the disturbed train, thus the accuracy of prediction is expected to have an effect on the performance of the regulation.

When the "accuracy" of prediction is considered, not only the accuracy of the time of resumption itself but also the process of prediction may be important. As has been stated in 8.4.3, a predicted time is usually expected to have a certain range of error, like "the train will resume 300 ± 100 s later". The new phase-by-phase regulation uses the centre value (300 s in this example). It has not proved possible to explore the effects of the error limits on the optimisation results within the scope of the present study, since it needs a good amount of information about real incidents and/or accidents happened in real railway operations and recovery time from these events, which is not available in this research.

The centre value of prediction is subject to change as time proceeds after the disturbance has begun. It is expected to be important how the centre value changes during the process of regulation. For example, suppose that the disturbed train is predicted to resume 300 s after the regulation enters Phase 3, when the first prediction is obtained. Following this, a further prediction is obtained that the disturbed train is expected to resume 400 s after the regulation has entered Phase 3. In this case, it would be expected that some of the trains behind the disturbed train would depart from stations, and thus will have to stop behind the disturbed train, since it remains in the station where the disturbance occurs. This may lead to a larger halt discomfort.

On the other hand, suppose that the further prediction suggests that the disturbed train is expected to resume 200 s after the first prediction has been made. In this case, some trains may have been forced to stop outside stations or remain in stations longer than necessary, which will lead to a larger waiting and/or travelling discomfort.

The results obtained by the regulation methods are examined in terms of the above three factors. The remainder of this chapter focuses on the influence of the relationship between the number of trains on the line and the number of platforms.

9.2 Simulation conditions

The following are the simulation conditions:

(1) Disturbed train: Train 11.

(2) Station where disturbance happens: Station 3.

(3) Direction of the disturbed train: From Station 2 to Station 4.

(4) Timing of disturbance: In the second roundtrip of Train 11. This means that Train 11 runs from Station 1 to 14, then returns from Station 14 to 1, and then goes on running from Station 1 to 3, and the disturbance happens here.

(5) Time at which the discomfort is evaluated: 10000 s after the first train (Train 1) enters the line.

(6) Phase timings and accuracy of prediction: Phase 1 lasts 60 s, then Phase 2 ends 300 s after the disturbance occurs. After that, Phase 3 is activated until the disturbed train resumes. Phase 4 is not activated in this simulation. The predicted resumption time is assumed to be accurate. Hence, in Figure 8.1, $t_1 = 60\,\text{s}$, $t_2 = 300\,\text{s}$, and $t_3 = t_r =$ resumption time.

The last factor (phase timings and accuracy of prediction) can be explained further by way of an example. Consider a door failure and examine what might actually happen during the process of re-starting.

A train is dwelling at a station. It is time to depart, and the train driver operates the door switch to close the doors. However, some of the doors do not close. The driver at first thinks that a passenger or baggage is caught in the doors, and operates the door switch again and again to re-open and close the doors. The driver or the station staff now begin to suspect a fault and examine the doors systematically. They eventually realise that a door is not working properly. The situation is then reported to the control centre. These activities will take approximately one minute, thus Phase 1 is assumed to last 60 s.

The driver would continue to try the door switch in the hope that the door trouble will disappear naturally and, when he realises that it will not recover naturally, he will begin to examine the cause of the trouble, and report details to the control centre. The operator in the control centre and/or the crew will consider how long it ought to take to clear the fault, and predict the time for restarting. The length of time that is needed for them to predict the time to resumption may depend on how experienced they are, and other conditions. In the simulation in this thesis, it has been assumed to be 300 s, which may seem fairly short, but is reasonable considering that typical train defects and failures are recorded so that there is a great deal of accumulated of knowledge of incidents and the time it takes to recover from them.

9.3 Parameters

To examine the effect of the relationship between the number of trains running on the line and the number of platforms, the service headway is varied between 75 s, 90 s, and 120 s.

When the headway is 75 s, the number of trains is 37, while the number of platforms is 29. Effectively, each station has two platforms, which makes 28. Besides these, there is a siding outside Station 1. Strictly, a siding is not a platform; nevertheless, it can be considered as a platform where there are neither originating nor terminating passengers. Thus, the total number of platforms is effectively 29. Consequently, this case has $N_{tr} > N_{pl}$, where N_{tr} is the number of trains, and N_{pl} is the number of platforms.

When the headway is 90 s, the number of trains is 31, thus this case is $N_{tr} \approx N_{pl}$ (not exactly the same, but nearly so). When the headway is 120 s, the number of trains is 24, thus this case has $N_{tr} < N_{pl}$.

The OD type is "Flat" in all cases; however, the OD value varies according to the headway, so that each train carries nearly the same number of passengers in a normal situation in each case. The OD value is 21 passengers/min for 75 s headway, 18 passengers/min for 90 s headway, and 15 passengers/min for 120 s headway.

The initial delay of the disturbed train varies between 600 s, 900 s, and 1200 s.

9.4 Results and discussions

Figures 9.1 to 9.3 show the results of the simulation.

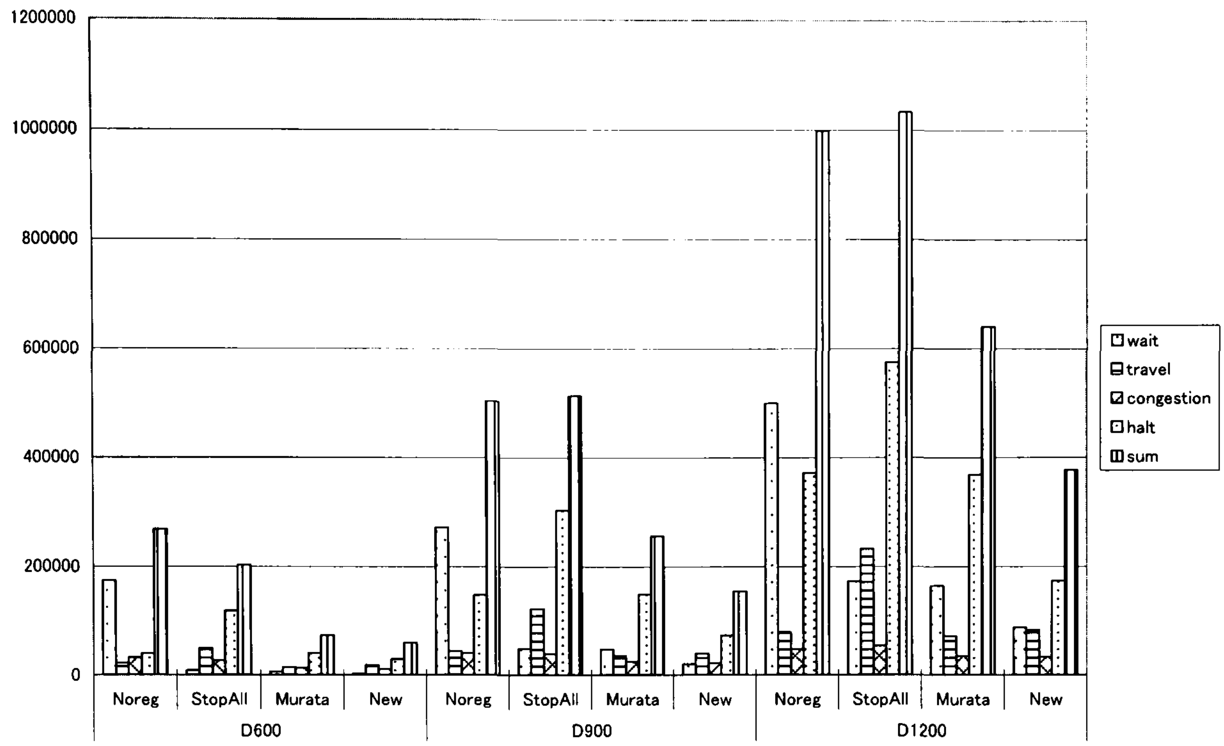

Figure 9.1: Discomfort comparison ($N_{tr} > N_{pl}$, Headway 75s, Flat 21 OD)

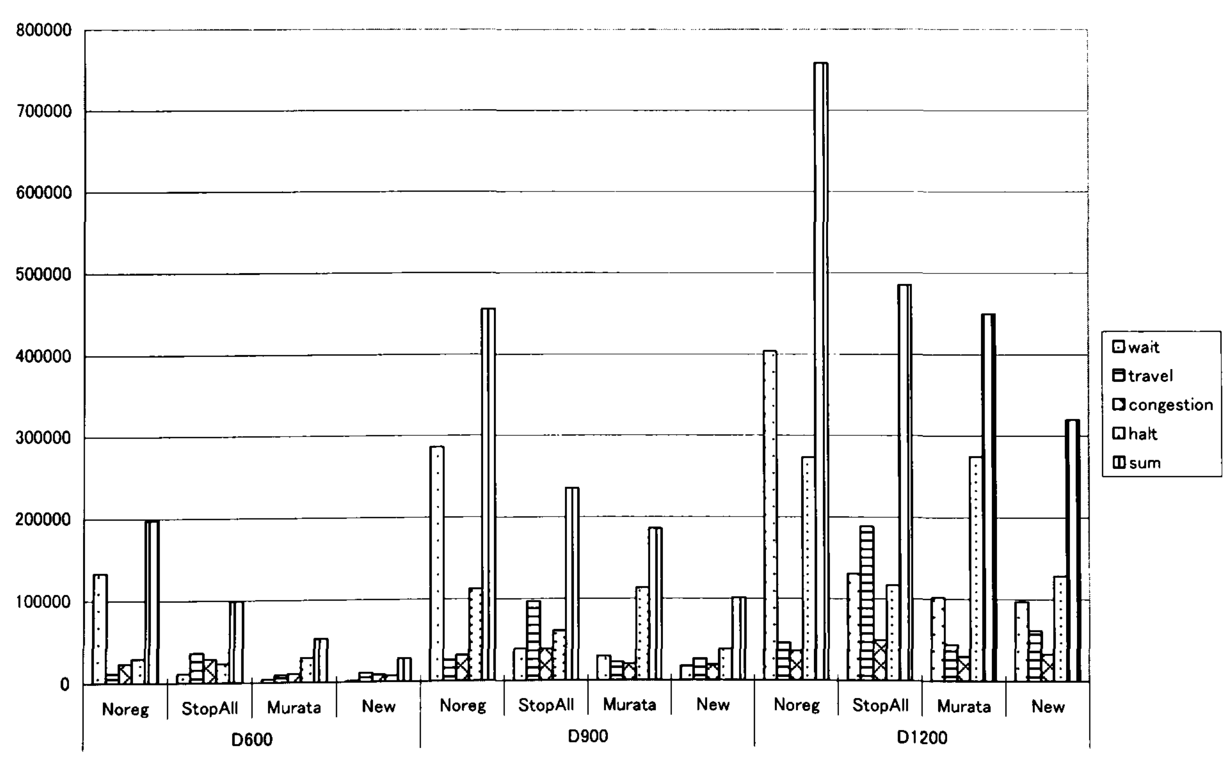

Figure 9.2: Discomfort comparison ($N_{tr} \approx N_{pl}$, Headway 90s, Flat 18 OD)

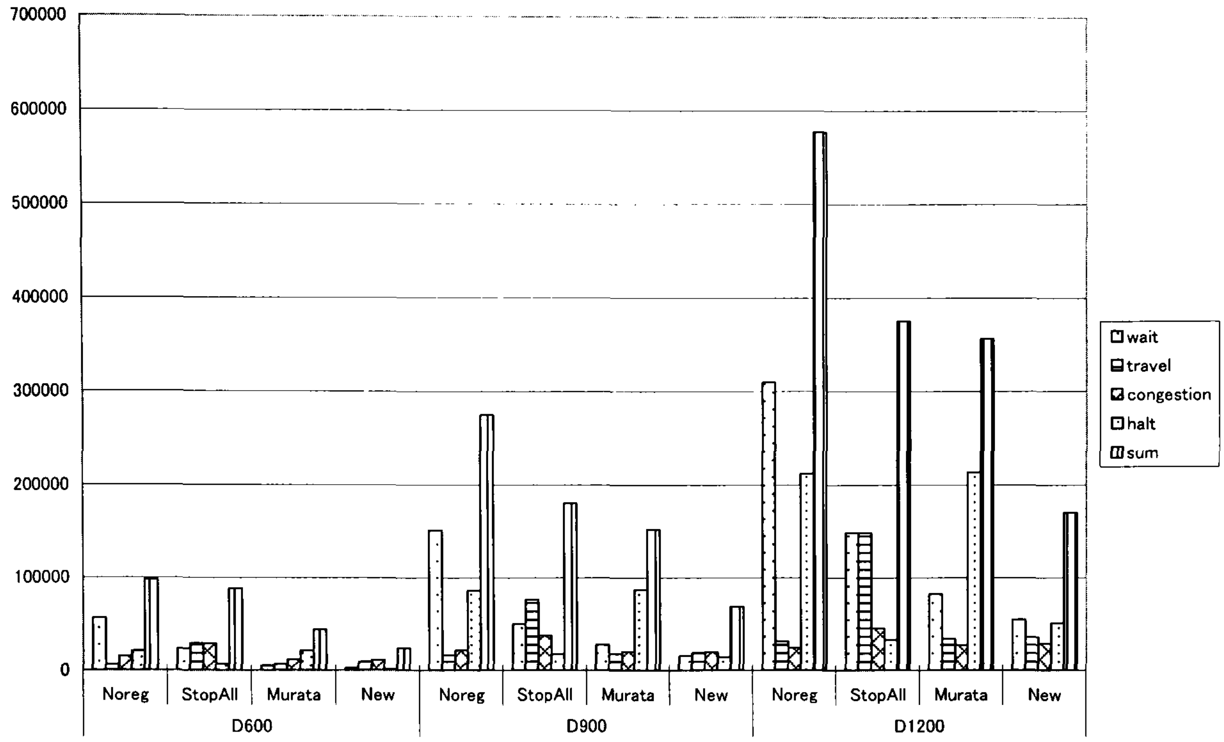

Figure 9.3: Discomfort comparison ($N_{tr} < N_{pl}$, Headway 120s, Flat 15 OD)

A significant difference is observed in the results for the "stop-all-trains-at-once" technique. When $N_{tr} < N_{pl}$, the waiting discomfort is larger than the halt discomfort, while when $N_{tr} > N_{pl}$, the relationship is reversed. When $N_{tr} \approx N_{pl}$, an intermediate result is obtained. For example, looking at the graphs when the initial delay is 900 s (which are shown as "D900" in Figures 9.1 to 9.3), the results for the unregulated simulation ("Noreg" in the figures), Murata plus halt regulation ("Murata" in the figures), and the new phase-by-phase regulation ("New" in the figures) are basically the same. Meanwhile, the results for the stop-all-trains-at-once technique ("StopAll" in the figures) show a significant difference in terms of the scale of the waiting discomfort and the halt discomfort as stated above.

This characteristic can be explained by the relationship between the number of trains and the number of platforms. When $N_{tr} < N_{pl}$, all trains stop at stations, while some platforms are not served; in other words, no train dwells at these platforms. In this situation, halt discomfort does not increase, instead, waiting discomfort for the passengers on the unoccupied platforms increases. On the other hand, when $N_{tr} > N_{pl}$, all the platforms are served, while some trains have to stop between stations, thus halt discomfort for the passengers on such trains increases.

In order to examine the relationships between the ratio of trains to platforms and the performance of the regulation methods, two additional simulation cases have been calculated. They are (1) headway = 84 s, and (2) headway = 100 s. The cases examined are shown in Table 9.1. In all cases, Flat 18 OD is adopted, and the initial disturbance duration is 600 s.

Table 9.1: Correspondence between headway and N_{tr} / N_{pl}.

Headway	N_{tr} / N_{pl}
75 s	1.27
84 s	1.13
90 s	1.06
100 s	0.96
120 s	0.82

Figure 9.4 shows the results of the simulations, in which the x-axis is the ratio of trains to platforms (N_{tr} / N_{pl}) and the y-axis is the ratio of the halt discomfort and the total discomfort values for the stop-all-trains-at-once regulation and the new phase-by-phase regulation against those for the Murata plus halt regulation.

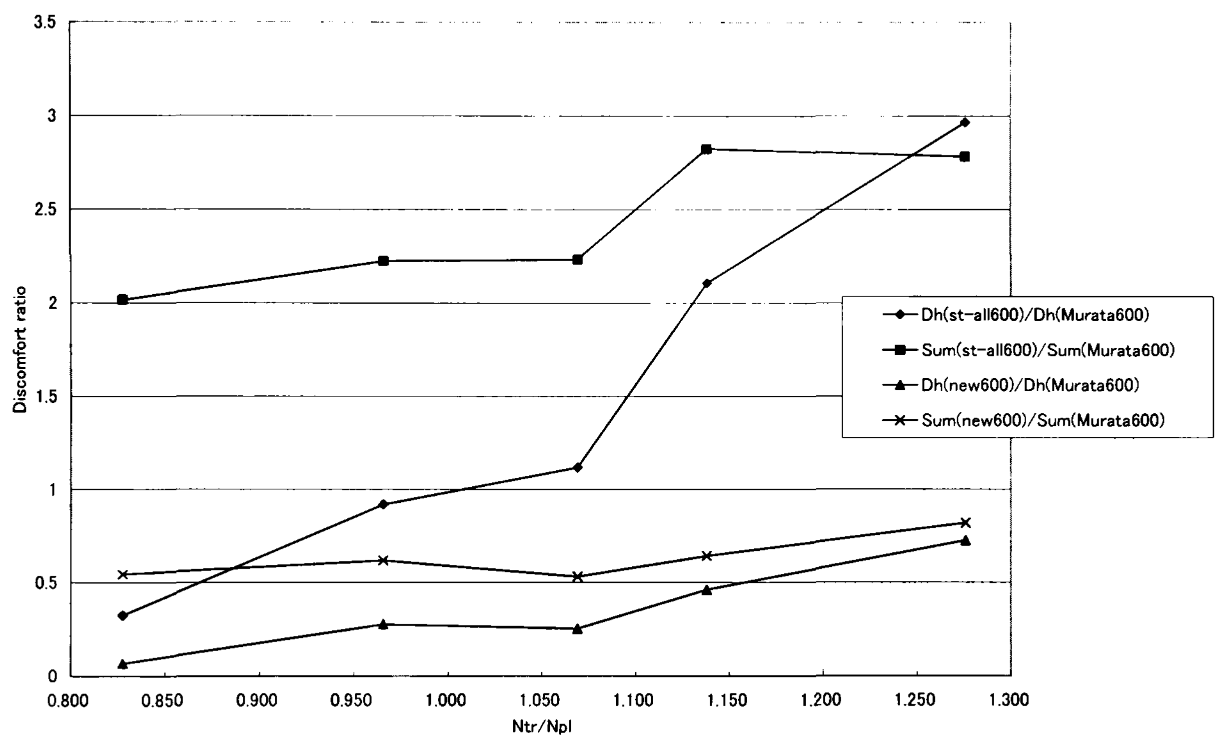

Figure 9.4: Discomfort comparison among various ratios (N_{tr} / N_{pl})

It can be observed in Figure 9.4 that the halt discomfort value for the stop-all-trains-at-once regulation is worse than that for the Murata plus halt regulation when N_{tr} / N_{pl} is 1.13 and 1.27. The total discomfort value for the stop-all-trains-at-once regulation is worse than that for the Murata plus halt regulation in all cases. This means that, when N_{tr} / N_{pl} is small, i.e., there are fairly few trains, the stop-all-trains-at-once regulation succeeds in reducing halt discomfort but not the other discomforts, thus the total discomfort value deteriorates. On the other hand, when N_{tr} / N_{pl} is large, i.e., there are many trains, the stop-all-trains-at-once regulation increases halt discomfort since some trains have to stop between stations, thus the total discomfort value deteriorates again.

From Figures 9.1 to 9.3, it can be concluded that, in all cases, the new phase-by-phase regulation shows the best performance. This result is attributed to the nature of the regulation, which is a mixture of the ideas from Murata and the stop-all-trains-at-once idea. The new phase-by-phase regulation succeeds in avoiding an increase of halt discomfort by suspending trains just after the disturbance happens. Then, across a prediction of the resumption time has been obtained, the regulator functions according to the ideas of Murata, which leads to the fact that the new regulator shows essentially the same performance as the Murata regulator in terms of discomfort values other than halt discomfort. If there is a difference, the new regulator is slightly better.

Even when $N_{tr} > N_{pl}$, the new regulator has the best performance, though it is expected that the halt discomfort will increase since some trains have to stop between stations during Phase 2 of the new regulation. This can be explained as follows. After the prediction of the restarting time is obtained (in other words, when the regulation enters Phase 3), trains can move according to commands obtained from the calculation of the partial derivatives of the objective function. Thus, trains a long way from the disturbance site are likely to restart, whereas this is not allowed in the stop-all-trains-at-once technique, where all trains have to stop until the disturbance is removed, no matter how far the train is from the disturbance site.

The regions where N_{tr} / N_{pl} is smaller than 0.8 and larger than 1.3 are not examined, since the results here can be readily anticipated. When trains are scarce, it is less and less likely that any train will have to stop between stations, thus the new phase-by-phase regulation will inevitably be more effective; conversely, when trains are densely packed, it is more probable that a number of trains will have to stop between stations, which leads to the result that even the new phase-by-phase regulation cannot reduce halt discomfort and, consequently, the total discomfort.

9.5 Conclusions

The new phase-by-phase regulation shows the best performance for all relationships between the number of trains and the number of platforms. This is attributed to the nature of the regulation, which is a mixture of the ideas of Murata (driven by a classical optimisation algorithm) and the stop-all-trains-at-once idea (somewhat heuristic and fairly crude, but often adopted in a real railway operation).

The results show that the new phase-by-phase regulation would be suitable for application to a wide range of real metro-type railways in, for example, UK, France, and Japan, see Table 4.1, where the density of trains is such that the new regulator shows clear benefits.

10 Results of new regulation (2) --- types of OD

10.1 Introduction

This chapter describes the differences in the outputs from the regulator under various OD types. The OD types that are simulated and examined here are called "Flat 18", "Centre 60", "Sub 50", "Mid_1 70", and "Mid_2 90", and have already been described in Chapter 4.

The following are the simulation conditions:

(1) Disturbed train: Train 11.

(2) Station where disturbance happens: Station 3.

(3) Direction of the disturbed train: From Station 2 to Station 4.

(4) Timing of disturbance: In the second roundtrip of Train 11. This means that Train 11 runs from Station 1 to 14, then returns from Station 14 to 1, and then goes on running from Station 1 to 3, and the disturbance happens here.

(5) Time at which the discomfort is evaluated: 10000 s after the first train (Train 1) enters the line.

(6) Phase timings and accuracy of prediction: Phase 1 lasts 60 s, then Phase 2 ends 300 s after the disturbance occurs. After that, Phase 3 is activated until the disturbed train resumes. Phase 4 is not activated in this simulation. The predicted resumption

time is assumed to be accurate. Hence, in Figure 8.1, $t_1 = 60\,\text{s}$, $t_2 = 300\,\text{s}$, and $t_3 = t_r$ = resumption time.

The initial delay of the disturbed train varies between 600 s, 900 s, and 1200 s.

10.2 Results and discussions

Figures 10.1 to 10.5 show the results of the simulation.

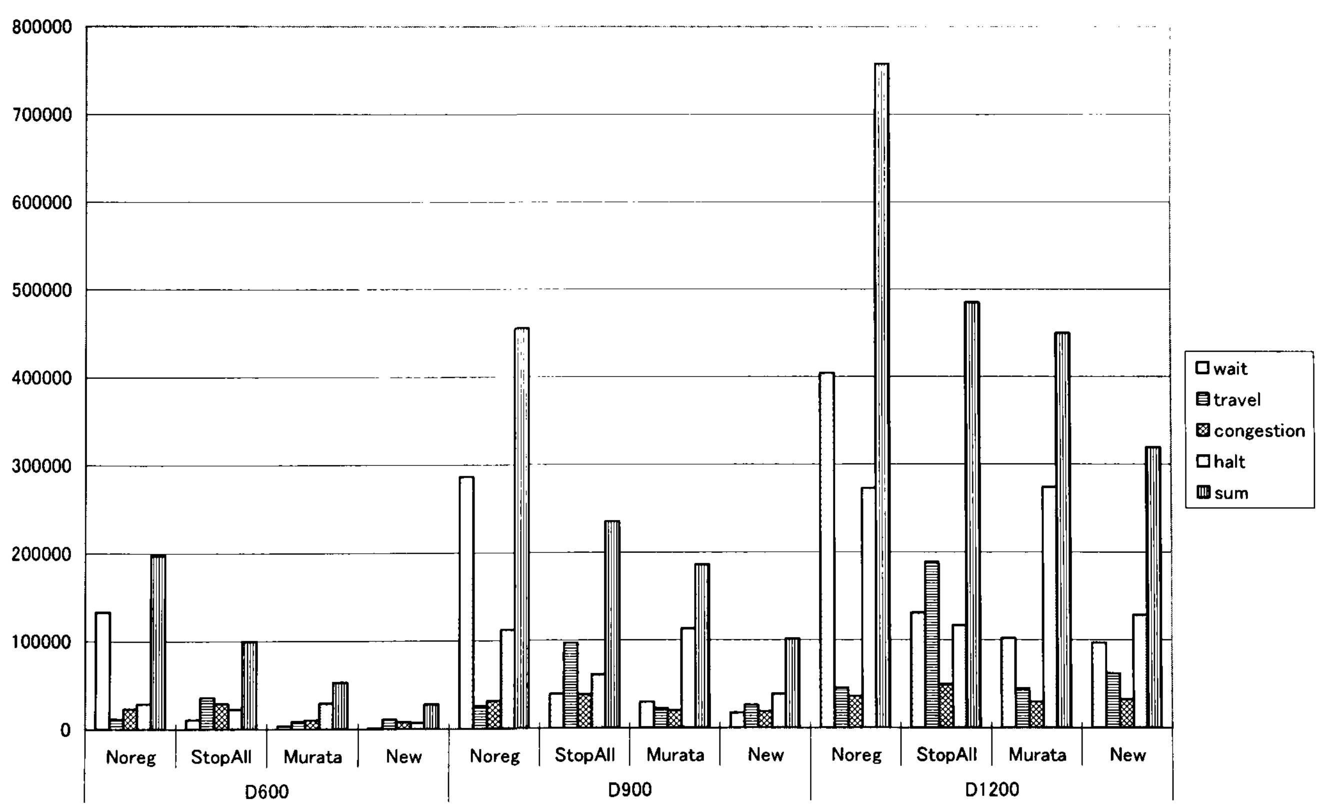

Figure 10.1: Discomfort comparison (Flat 18 OD)

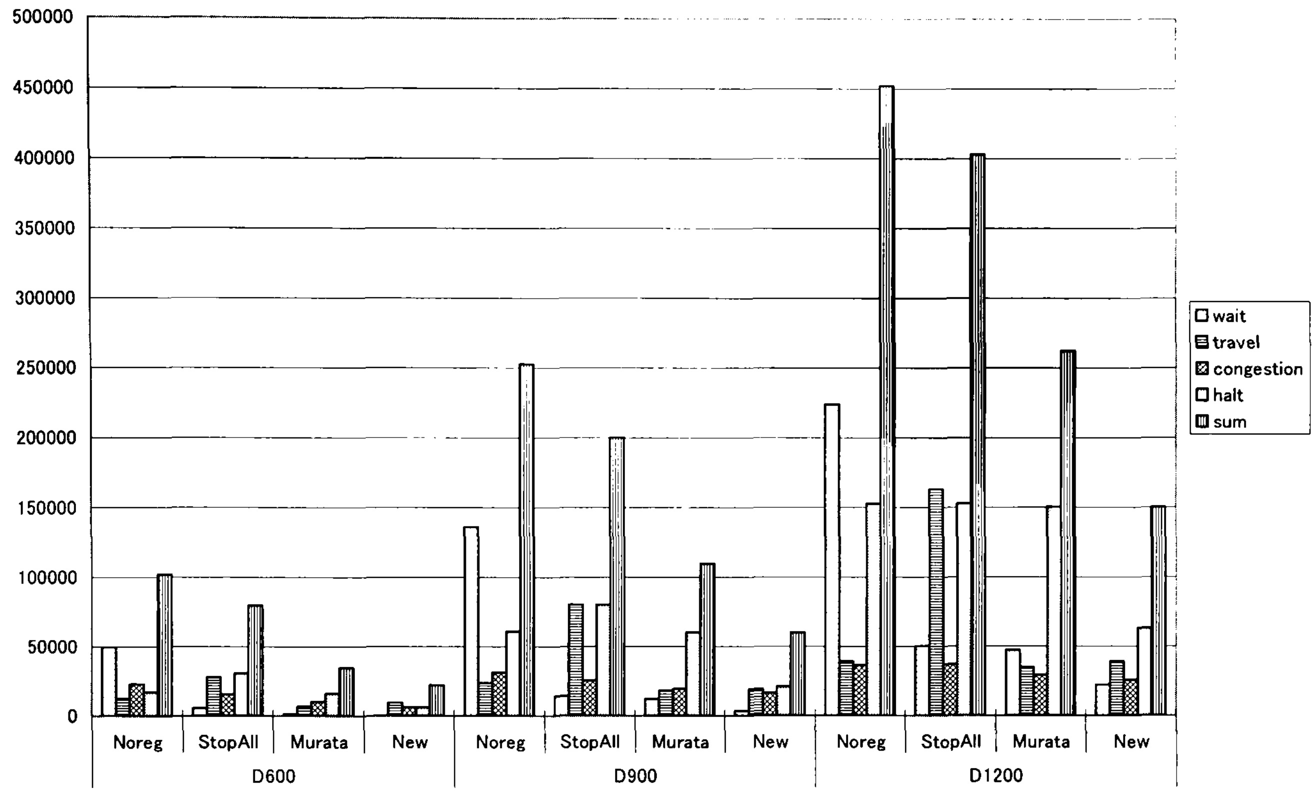

Figure 10.2: Discomfort comparison (Centre 60 OD)

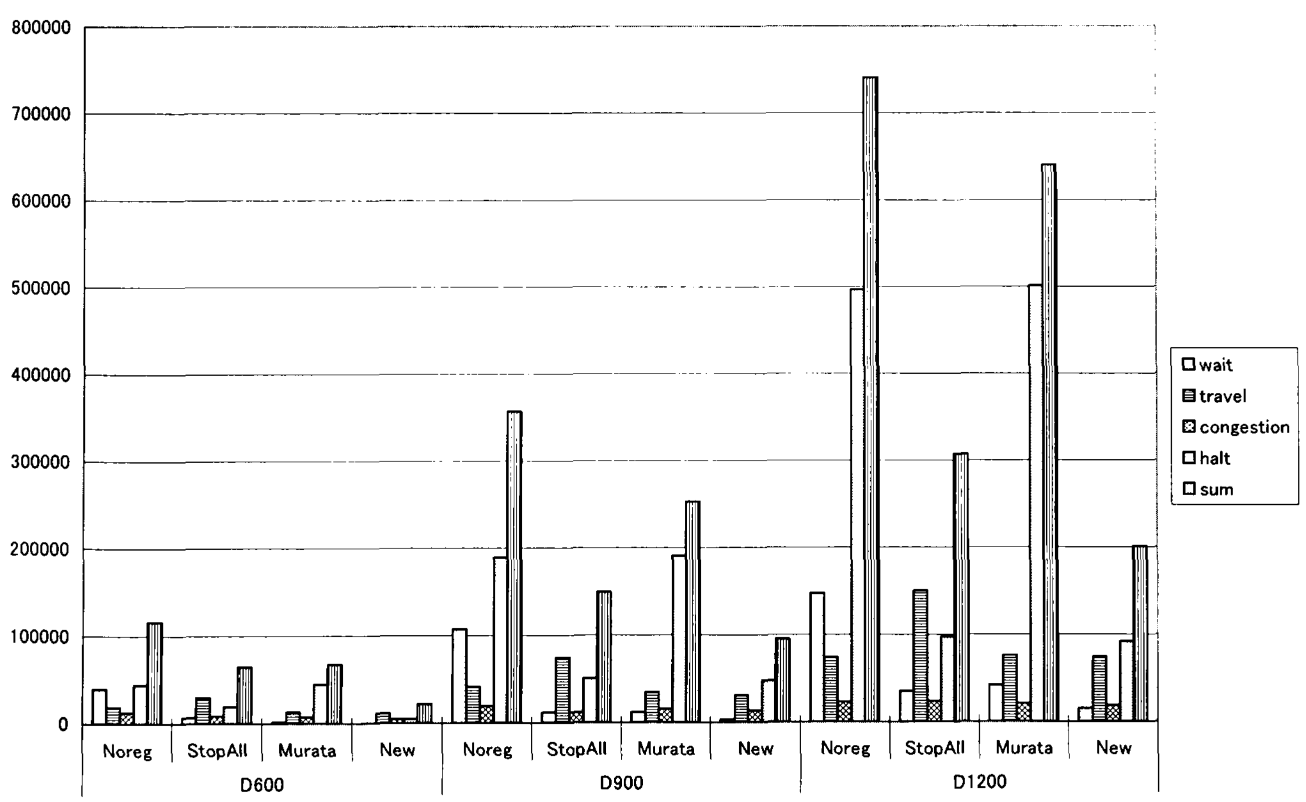

Figure 10.3: Discomfort comparison (Sub 50 OD)

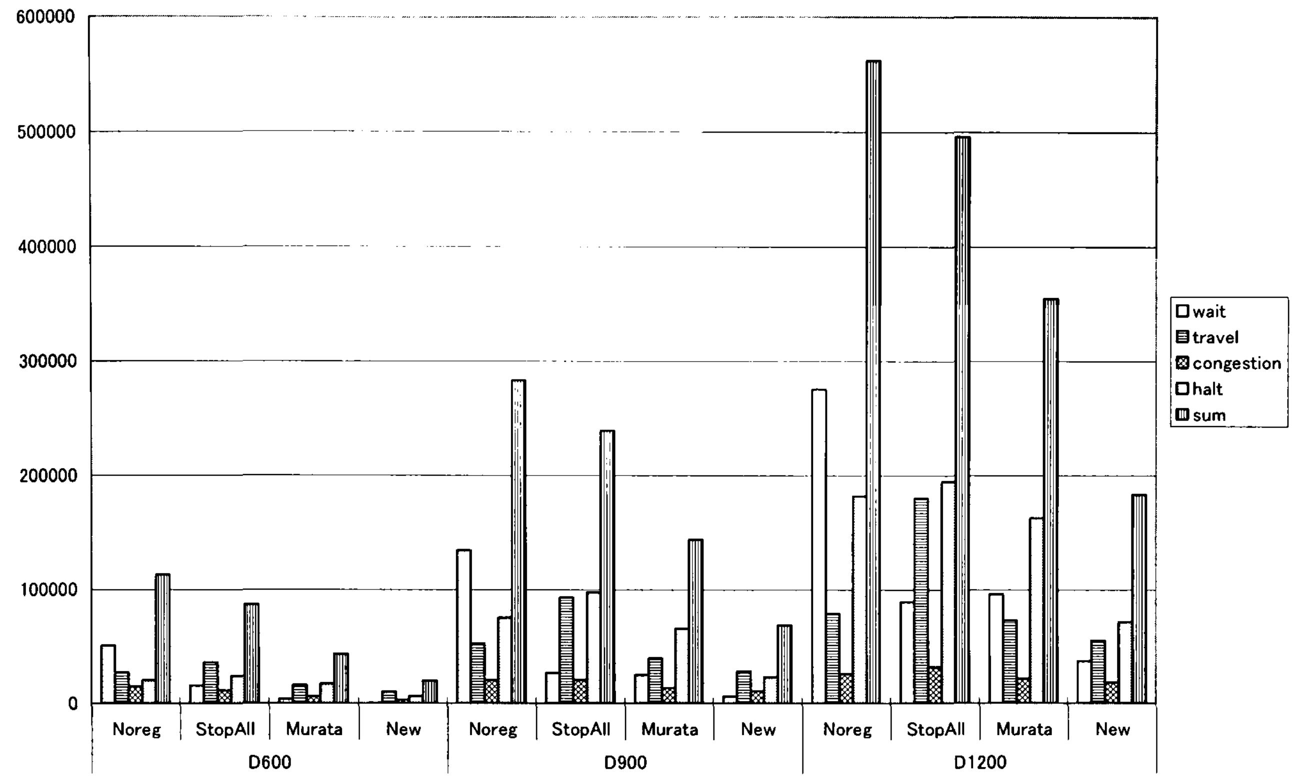

Figure 10.4: Discomfort comparison (Mid1_70 OD)

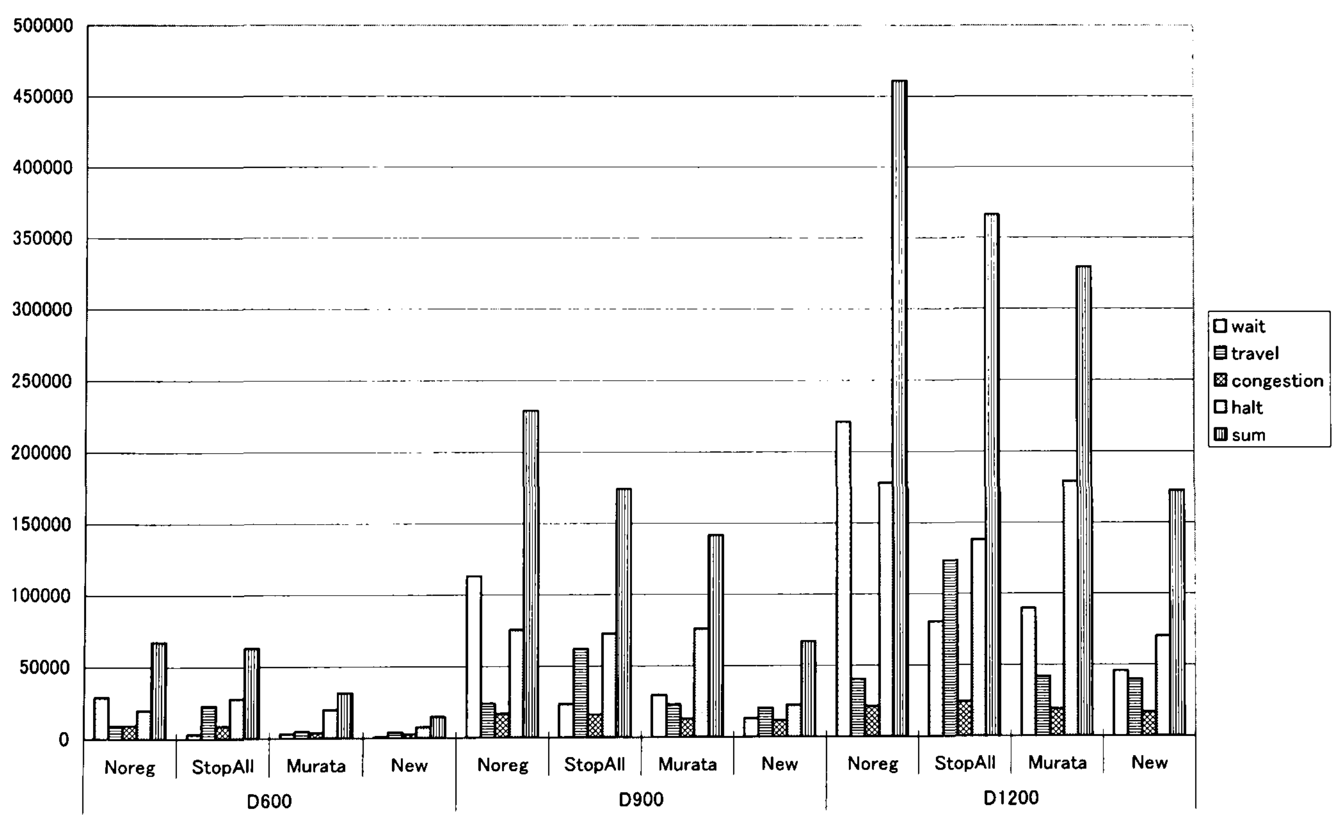

Figure 10.5: Discomfort comparison (Mid2_90 OD)

In comparing Sub OD and the other OD types, a significant difference can be observed in the performance of the stop-all-trains-at-once technique and the Murata plus halt regulator. For Sub OD, the stop-all-trains-at-once technique shows a better performance than the Murata plus halt regulator, while in the other OD types, the Murata plus halt regulator shows a better performance than the stop-all-trains-at-once idea.

This is a result of the position where the disturbance is assumed to happen. In this simulation, the disturbance happens at Station 3; Train 11 cannot leave Station 3 for between 600 s and 1200 s. The important factor is how heavy is the traffic around Station 3. In Sub OD, the traffic becomes lighter as trains go from the city centre to the suburbs, i.e., from Station 1 to 14. At Station 3, the traffic is still heavy. Thus, when a train is disturbed at Station 3 and the Murata plus halt regulator is activated, which does not include halt discomfort in its objective criteria, the halt discomfort increases greatly, since trains carrying a large number of passengers have to stop between stations. In the other OD types, this effect is not observed.

For Flat OD, the stop-all-trains-at-once technique nearly halves the halt discomfort in comparison with the unregulated case and the Murata plus halt regulator, while it increases the other kinds of discomfort, particularly travelling discomfort. By stopping all trains when the disturbance happens, a sudden increase in halt discomfort is avoided immediately after the disturbance occurs, but passengers have to wait in the trains (standing at platforms), which results in the travelling time for those passengers becoming longer. This leads to a larger travelling discomfort.

The new phase-by-phase regulation shows the best performance for the Flat OD. The results for this regulation show that the halt discomfort is as small as that for the stop-all-trains-at-once technique, while the other kinds of discomfort are as small as those using the Murata plus halt regulator. This suggests that the new regulation inherits both the merits of the stop-all-trains-at-once technique and the Murata technique, which is a primary aim of this research.

For Centre OD, the stop-all-trains-at-once technique shows much the same halt discomfort as the Murata plus halt regulator and the unregulated case, while it increases the other kinds of discomfort, particularly travelling discomfort. This is in sharp contrast to the result for Sub OD, in which the stop-all-trains-at-once idea decreases halt discomfort greatly. This is again due to the position where the disturbance occurs, as has already been discussed above. The results for Centre OD suggest that, when the disturbance happens in a less congested area, the stop-all-trains-at-once technique should not be applied, since unnecessary suspension of trains increases the passenger discomfort greatly without reducing the halt discomfort at all.

The new phase-by-phase regulation again shows the best performance for Centre OD and Sub OD. In both cases, not only does the halt discomfort decrease greatly, but the other kinds of discomfort also decrease slightly compared with the Murata plus halt regulator. This suggests that, by regulating trains immediately after the disturbance happens, the new regulation contributes to reducing the discomfort in the subsequent operation after the disturbance is cleared.

The results for Mid_1 OD and Mid_2 OD show nearly the same characteristics as those of Centre OD. This is due to the fact that the disturbance position in these cases is where the passenger load on the trains is getting heavier and heavier.

The new phase-by-phase regulation again shows the best performance in Mid_1 OD and Mid_2 OD.

10.3 Conditions at the disturbance site

As has been stated above, five types of OD have been simulated and discussed, with the disturbance site fixed as Station 3. In this section, the simulation results are explained from the viewpoint of the characteristics of the disturbance site.

The reason that the characteristics of the disturbance site are focused on is as follows. The problem in this research is inevitably multi-dimensional, with the parameters such as OD types, headway, a disturbance site, and duration of an initial disturbance. If all the parameters are closely examined, the number of simulated cases will explode; it has to be said that such a simulation is not beneficial for practical use on real railways.

In order to investigate effectively the characteristics of disturbed railway traffic and the performance of regulation, the influential parameters should be taken in the simulations. Then the problem has as low a dimension as possible, and thus the regulation method established in this research can be used in real railways without much validation.

The point of view from which the characteristics of the disturbance site are discussed here is related to the distribution of passengers in the line in each OD type. The parameter describing this point is introduced and defined as follows.

Let $P(\text{dist})$ be the number of passengers on a train running towards the station where disturbance happens (which means that the train is running from Station 2 to Station 3 in the simulations) and $P(\text{max})$ be the maximum number of passengers on a train in each OD type, both in normal operation (without any disturbance). For example, under Flat 18 OD, the passenger flow is 432 passengers per minute from Station 2 to Station 3 when there is no disturbance, while the maximum number of passenger flow from Station 7 to Station 8 is 882 passengers per minute. Thus, the ratio of the number of passengers on a train running towards the disturbance site and the maximum number of passengers on a train is:

$$P(\text{dist})/\,P(\text{max}) = 432/882 = 0.490 \qquad (10.1)$$

If the other cases are calculated in the same manner, Table 10.1 is obtained.

Table 10.1: *P*(dist)/ *P*(max) for different OD types

OD type	*P*(dist)/ *P*(max)
Flat 18	0.490
Centre 60	0.318
Sub 50	0.985
Mid_1 70	0.367
Mid_2 90	0.354
Flat 18, Station = 4	0.673

In the last row of Table 10.1 another case is described, in which the OD type is Flat 18, but the disturbance site is Station 4, not Station 3. This case has been added so that the values of the parameter *P*(dist)/ *P*(max) are evenly distributed among the simulated cases.

Figure 10.6 shows a discomfort comparison expressed as a ratio for various ratios *P*(dist)/ *P*(max), compared with the discomfort values for the Murata plus halt regulator. In this figure, "Dh" and "Sum" mean the halt discomfort and the total discomfort respectively, and "StopAll600", "Murata600", and "new600" mean the simulation results for the stop-all-trains-at-once technique, the Murata plus halt regulator, and the new phase-by-phase regulation respectively.

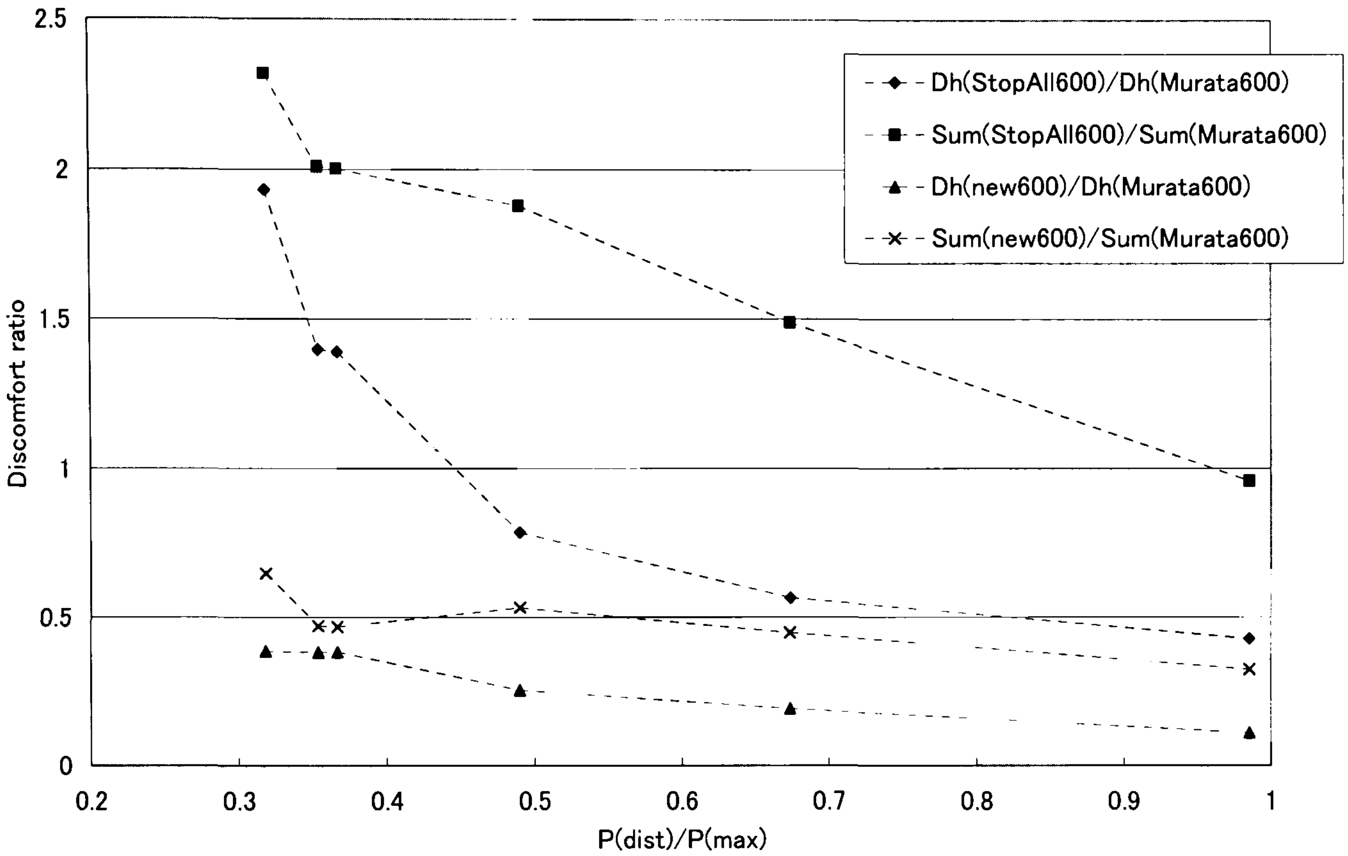

Figure 10.6: Discomfort comparison among various ratios *P*(dist)/ *P*(max)

It can be seen in Figure 10.6, when *P*(dist)/ *P*(max) becomes higher, the stop-all-trains-at-once technique is more effective in reducing halt discomfort. This demonstrates that the stop-all-trains-at-once technique is more useful when the disturbance happens at a busy station, since it can prevent heavily congested trains from stopping between stations. On the other hand, when *P*(dist)/ *P*(max) is smaller, the stop-all-trains-at-once idea is less effective in reducing halt discomfort, since it stops all trains, including congested trains running far from the disturbance site, and these trains then only contribute to increasing passenger discomfort. Without the stop-all-trains-at-once technique being used, these trains would continue to move, and would therefore not increase the halt discomfort.

The results show the new phase-by-phase regulation to be more effective than the Murata plus halt regulation in all cases and also to be better than the stop-all-trains-at-once technique applied alone. This reflects the performance of a blending of the Murata approach and the stop-all-trains-at-once technique. The new regulation successfully avoids an increase in halt discomfort by stopping all trains after the disturbance happens, keeping them stopping for a short time, and then moves the trains using the idea of Murata on the basis of prediction as to when the disturbance will be removed.

10.4 Conclusions

The new phase-by-phase regulation shows the best performance in all cases of different OD types. This confirms that a primary aim of this research has been achieved. The new phase-by-phase regulation, which is the mixture of the ideas of Murata and the stop-all-trains-at-once technique, inherits the merits of each regulation method.

The parameter introduced in 10.3, P(dist)/ P(max), explains the differences among the OD types well. It can be concluded here that, despite a lot of OD types and other parameters describing the characteristics of the model line, the performance of the regulation can be predicted by using P(dist)/ P(max), which is useful for a practical use. Taking Station 3 as a disturbance site is consequently an effective strategy to find influential parameters in the problem.

11 Results of new regulation (3) --- accuracy of prediction

11.1 Introduction

The simulation results explained in Chapters 9 and 10 are calculated on the assumption that the trains restart their operations exactly at the predicted time; in other words, the prediction is assumed to be perfectly accurate. This assumption is re-evaluated in this chapter, which examines how the accuracy of prediction of the time when the disturbed train will restart affects the performance of the new phase-by-phase regulation. The new phase-by-phase regulation calculation is based on a predicted time of resumption for the disturbed train, thus the accuracy of prediction will inevitably have an effect on the performance of the regulator.

There are two kinds of the accuracy of prediction to be discussed here. One is the variation of prediction over time. As has been stated in Chapter 8, when the first prediction is made a certain amount of time after the disturbance has happened, the prediction will be subject to error. Then, as time proceeds, the prediction should become more and more accurate, as, for example, the situation becomes clearer to the human operator in the control centre.

During the process of the regulation, the prediction may change to a later time, or to an earlier time, or it may not change at all. How the prediction changes until the trains actually restart their operations would be expected to have an effect on the performance of the regulator, since the Enhanced Murata plus halt regulator, which is incorporated into the new phase-by-phase regulation, calculates the instructions for each train according to the predicted time of restarting operation.

In terms of the variation of the prediction over time, five types of variation have been considered. They are referred to as Type 1 through 5. These five types are introduced and explained in section 11.2.

The other point to be discussed concerns the accuracy of prediction of the difference between the predicted time of resumption and the actual time of resumption. Suppose that a certain time of prediction is obtained after some revisions in the process of the regulation. It could still happen that train operation does not restart at the predicted time for reasons such as extended delay to the repair of broken doors. As has been stated in Chapter 8, if a second disturbance occurs and is expected to last for a significant time, the regulation should go back to Phase 2. Nevertheless, a small difference between the predicted time and the actual time of resumption may exist, even if another disturbance does not happen. Thus the effect of this difference on the performance of the regulation needs to be examined.

For each type of the variation of the prediction over time, three cases of the difference between the predicted time and the actual time of resumption are calculated, as described in the next section.

11.2 Accuracy of prediction

In order to examine effects of the accuracy of prediction, five types of variation in the prediction are considered. They are Type 1 through 5.

11.2.1 Type 1

Type 1 represents the situation where the first prediction is the best available and does not change during the process of regulation. However, as has been stated in 11.1, the disturbed train still may not resume at this predicted time, thus three cases are examined for each initial delay. That is, when the initial delay is 600s, the disturbed train restarts its operation at the predicted time, 60s later and 60s earlier than the predicted time. When the initial delay is 900s and 1200s, the deviation is ± 120s. Figure 11.1 shows the evolution of the prediction in Type 1 when the initial disturbance is 600s.

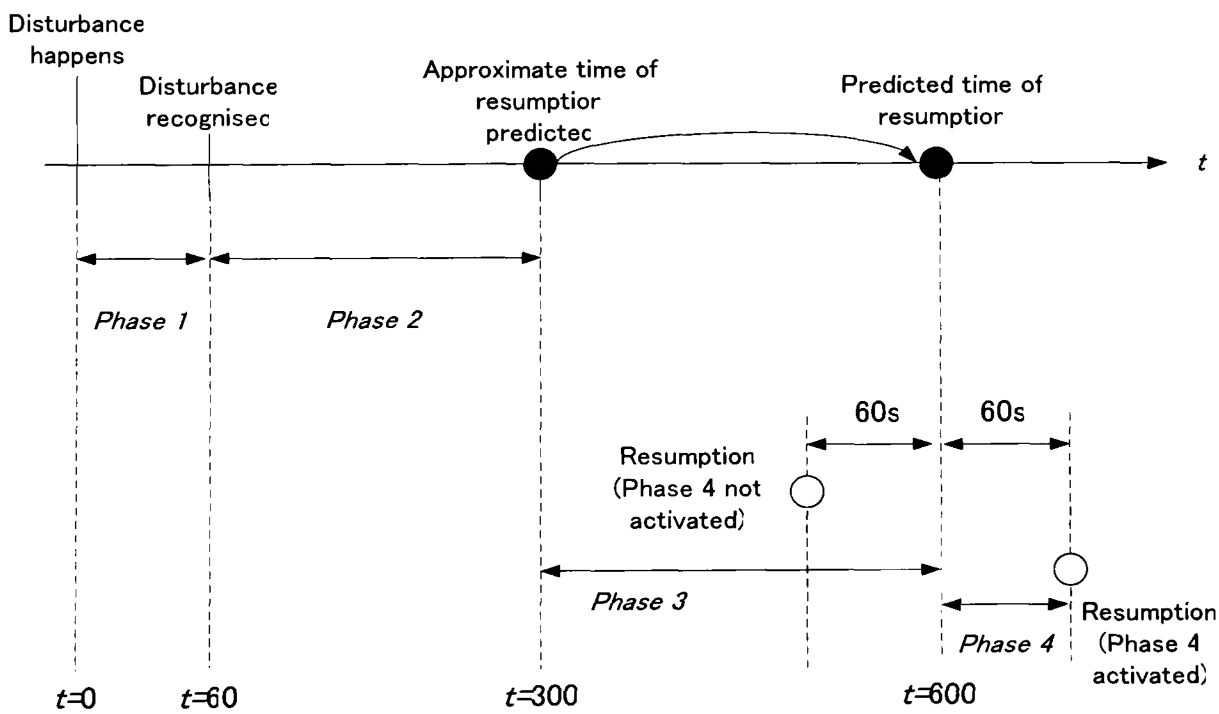

Figure 11.1: Evolution of prediction in Type 1 (Initial disturbance = 600s)

11.2.2 Type 2

In Type 2, the predicted time of resumption advances as time passes. This represents the situation where the first prediction is made "conservatively", since the cause of the disturbance situation is not clear. Then, as the situation becomes clearer, it becomes apparent that the disturbed train will resume earlier than had been previously predicted.

When the initial delay is 600s, the prediction changes as follows. Note that all the time values are measured after the regulation enters Phase 3, which starts 300s after the disturbance occurs.

(1) 0-100s ... predicted time is 500s

(2) 100-200s ...predicted time is 400s

(3) 200-300s ...predicted time is 300s

Again, the disturbed train restarts its operation at the predicted time, 60s later and 60s earlier than the predicted time. Figure 11.2 shows the evolution of the prediction in Type 2 when the initial disturbance is 600 s.

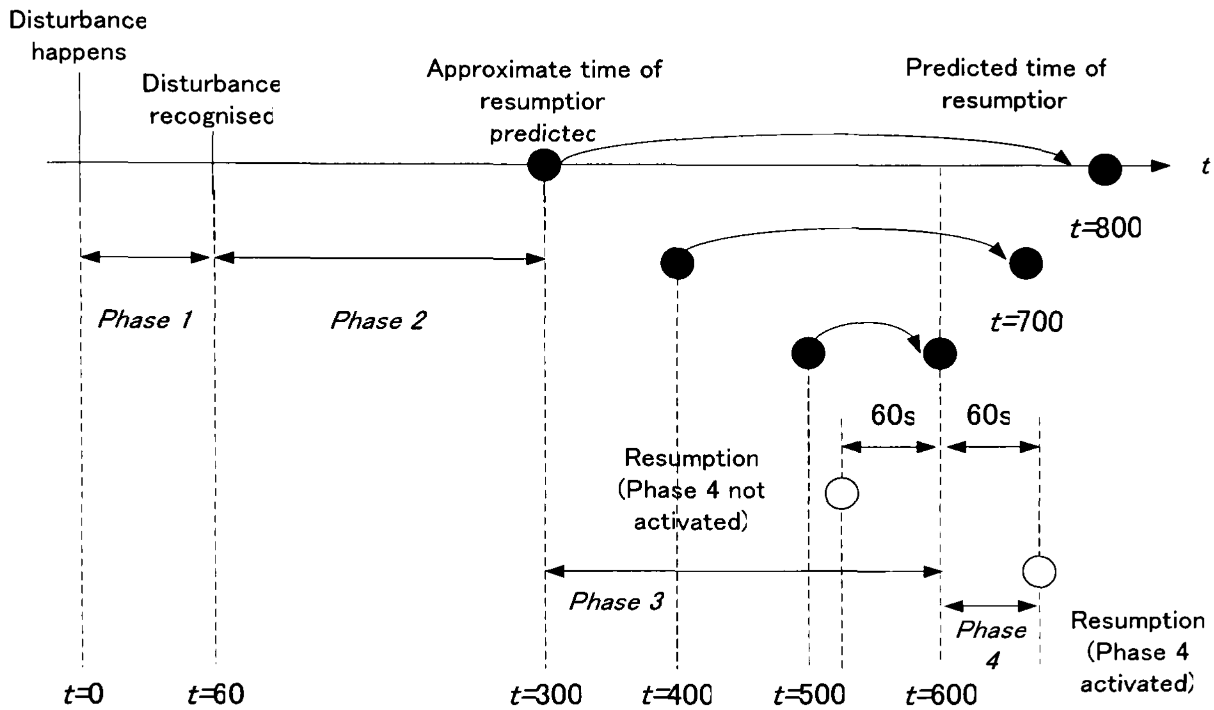

Figure 11.2: Evolution of prediction in Type 2 (Initial disturbance = 600s)

When the initial delay is 900s, the prediction changes as follows (after the regulation enters Phase 3),

(1) 0-200s... predicted time is 900s

(2) 200-400s ...predicted time is 750s

(3) 400-600s ...predicted time is 600s

The disturbed train restarts the operation at the predicted time, 120s later and 120s earlier than the predicted time.

When the initial delay is 1200s, the prediction changes as follows (after the regulation enters Phase 3),

(1) 0-300s ... predicted time is 1300s

(2) 300-600s ...predicted time is 1100s

(3) 600-900s ...predicted time is 900s

The disturbed train restarts the operation at the predicted time, 120s later and 120s earlier than the predicted time.

11.2.3 Type 3

In Type 3, the predicted time of resumption is further delayed as time passes. This represents the situation where the first prediction is made "impetuously". Then, as the situation becomes clearer, it is realised that the disturbed train will resume later than had been predicted previously. Figure 11.3 shows the evolution of the prediction in Type 3 when the initial disturbance is 600 s.

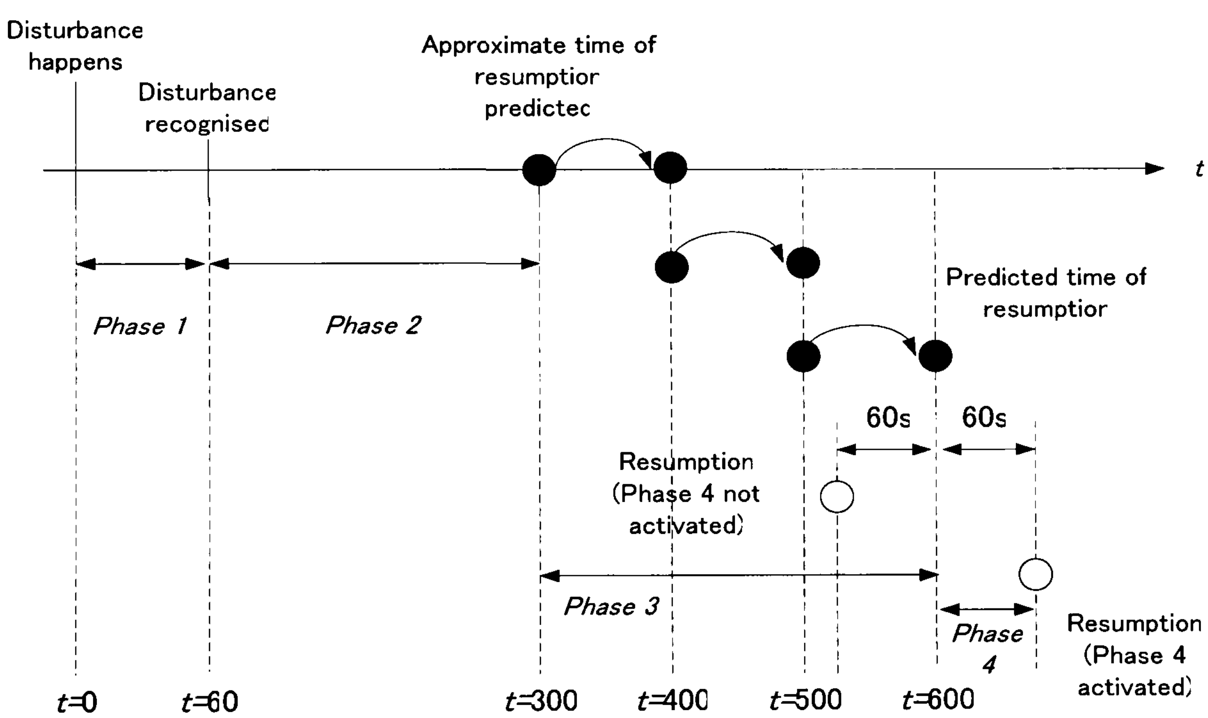

Figure 11.3: Evolution of prediction in Type 3 (Initial disturbance = 600s)

The evolution of prediction over time in all the types will be summarised at the end of 11.2.

11.2.4 Type 4

Sometimes, a prediction may vary both forwards and backwards as time proceeds. Types 4 and 5 represent this type of situation.

Type 4 represents the situation where the first prediction is made "conservatively", but later the prediction is modified to an earlier time, then finally it is again modified to a later time (but still earlier than the first prediction). Figure 11.4 shows the evolution of the prediction in Type 4 when the initial disturbance is 600 s.

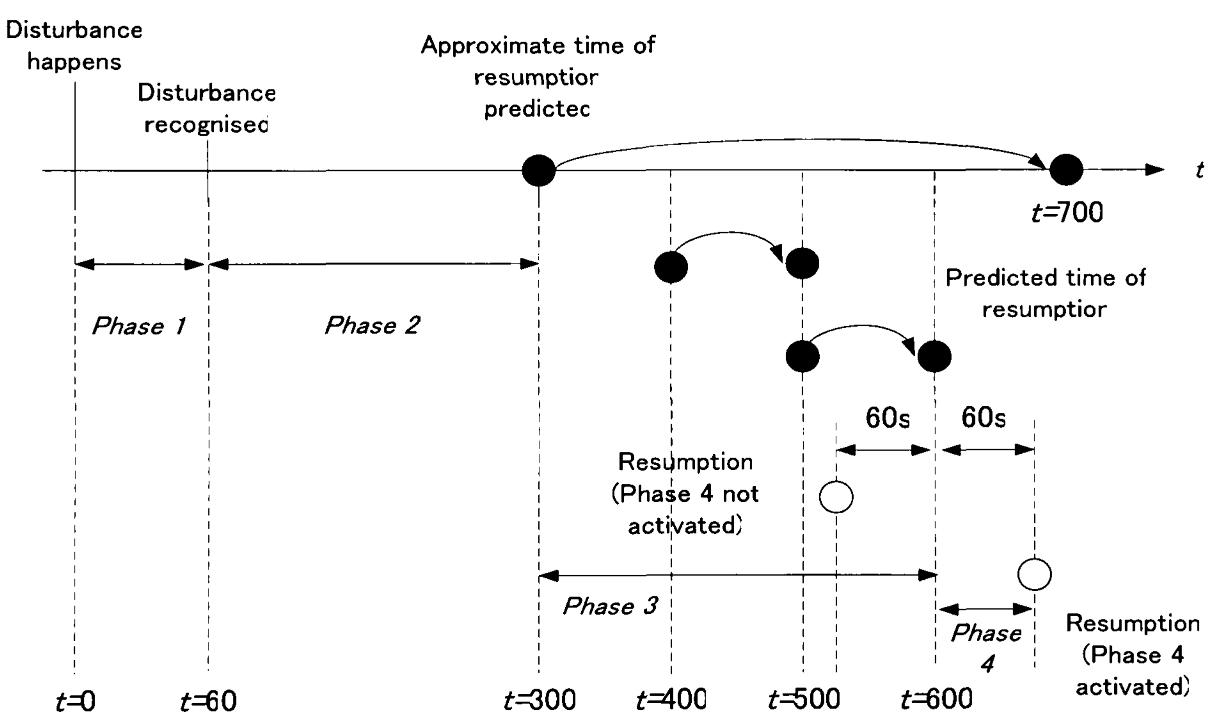

Figure 11.4: Evolution of prediction in Type 4 (Initial disturbance = 600s)

11.2.5 Type 5

Type 5 represents the situation where the first prediction is made "impetuously", but later the prediction is modified to a later time, then finally it is again modified to an earlier time (but still later than the first prediction). Figure 11.5 shows the evolution of the prediction in Type 5 when the initial disturbance is 600 s.

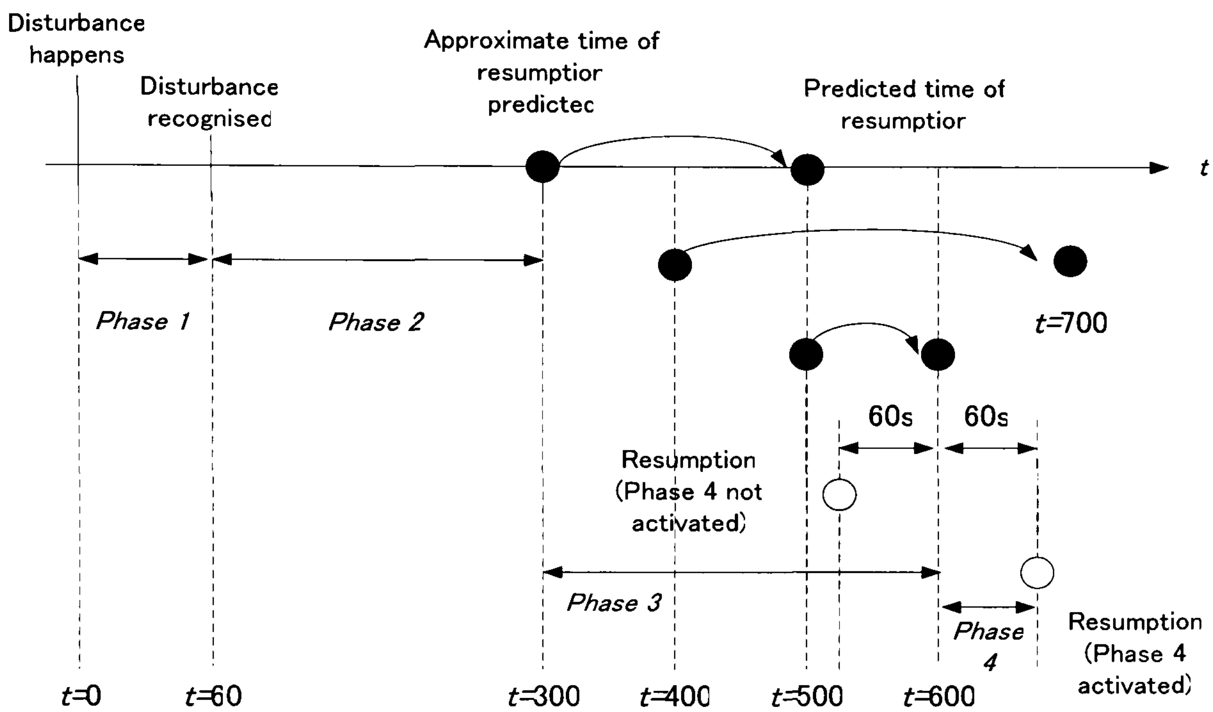

Figure 11.5: Evolution of prediction in Type 5 (Initial disturbance = 600s)

Table 11.1 through 11.3 describe the evolution of prediction over time in all the types when the initial disturbance is 600s, 900s, and 1200s respectively.

Table 11.1: Evolution of prediction in all the types (Initial disturbance = 600s)

	Type 1	Type 2	Type 3	Type 4	Type 5
0-100s	300s	500s	100s	400s	200s
100-200s	300s	400s	200s	200s	400s
200-300s	300s	300s	300s	300s	300s

Table 11.2: Evolution of prediction in all the types (Initial disturbance = 900s)

	Type 1	Type 2	Type 3	Type 4	Type 5
0-200s	600s	900s	300s	750s	450s
200-400s	600s	750s	450s	450s	750s
400-600s	600s	600s	600s	600s	600s

Table 11.3: Evolution of prediction in all the types (Initial disturbance = 1200s)

	Type 1	Type 2	Type 3	Type 4	Type 5
0-300s	900s	1300s	500s	1100s	700s
300-600s	900s	1100s	700s	700s	1100s
600-900s	900s	900s	900s	900s	900s

11.3 Parameters

As has already been explained, the initial delay of the disturbed train varies in 600s, 900s, and 1200s.

11.4 Results and discussions

Figures 11.6 to 11.8 show the results of the simulations for all regulation methods when the initial delay is 600s, 900s and 1200s respectively.

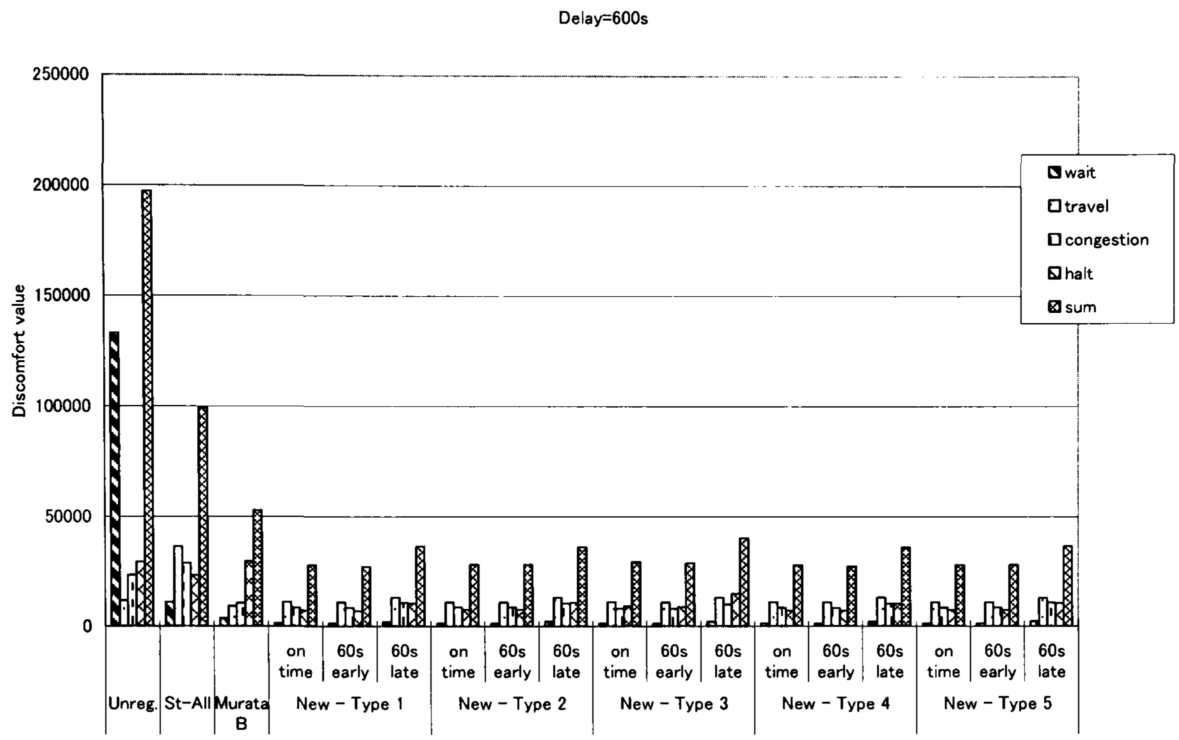

Figure 11.6: Discomfort comparison among various Types of prediction evolution

(Initial disturbance = 600s)

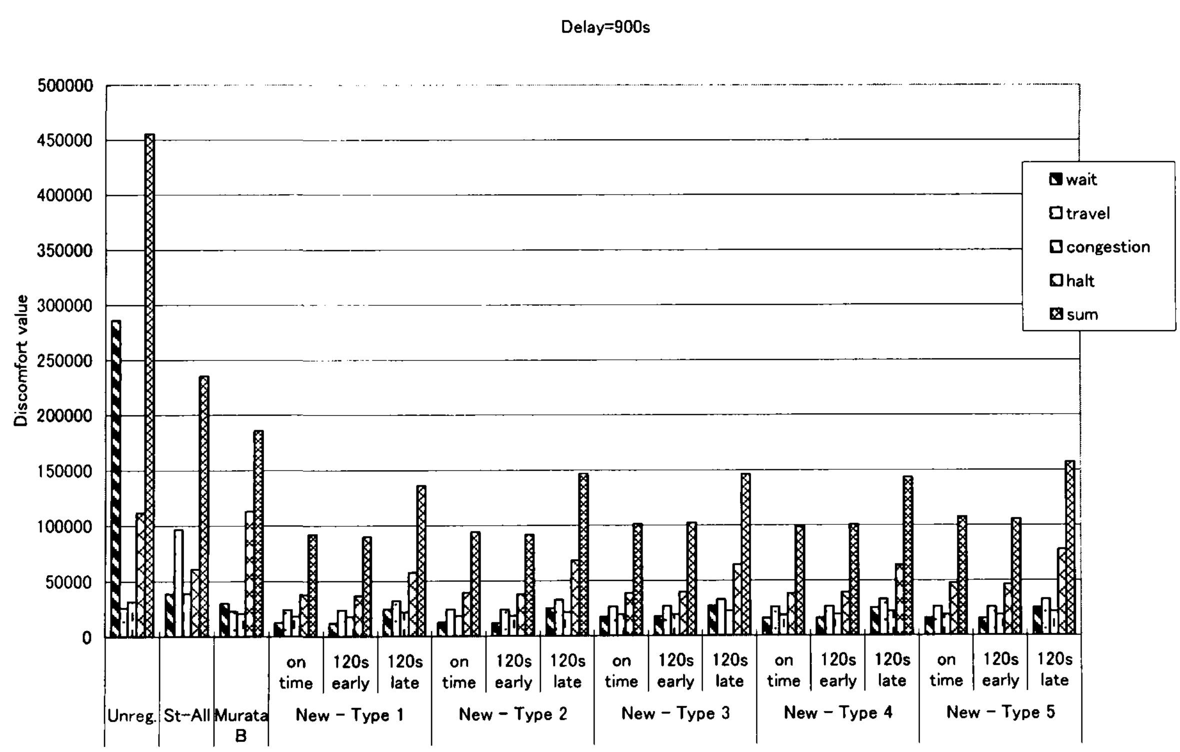

Figure 11.7: Discomfort comparison among various Types of prediction evolution

(Initial disturbance = 900s)

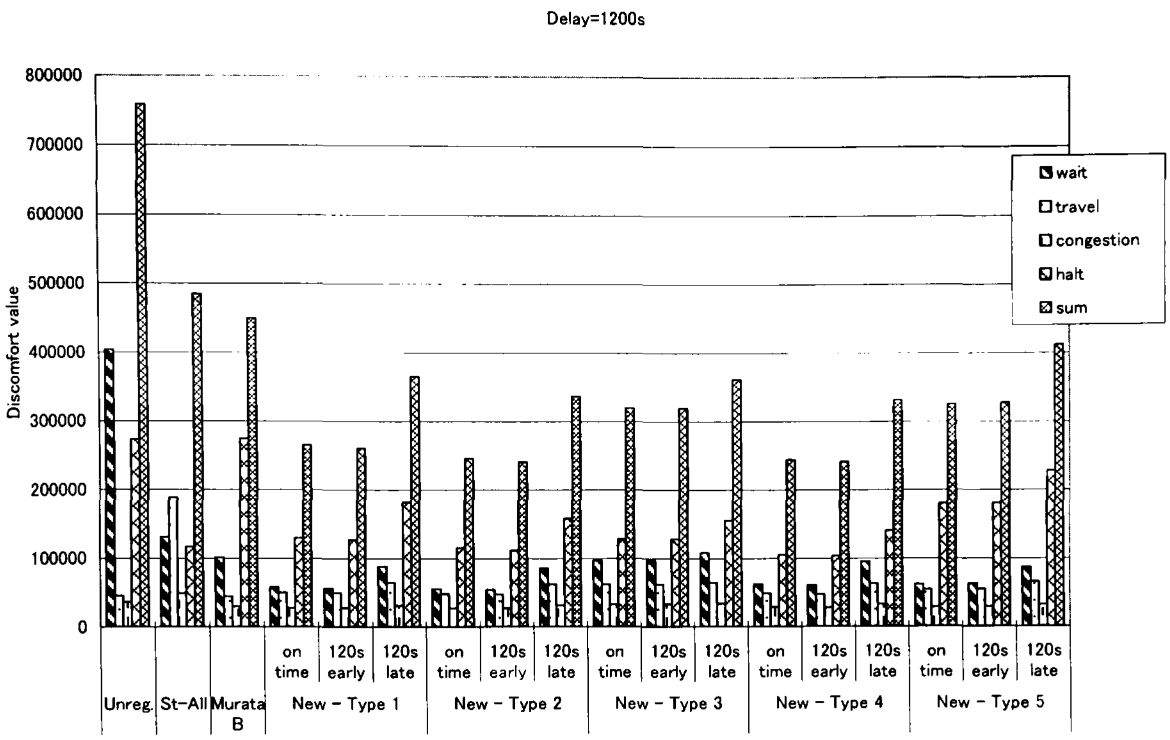

Figure 11.8: Discomfort comparison among various Types of prediction evolution (Initial disturbance = 1200s)

It can be seen in Figures 11.6 to 11.8 that Types 1 and 2 show the best performance in all cases. When the initial delay is 600s and 900s, they have almost the same performance, while when the initial delay is 1200s, the performance of Type 2 is visibly better than that of Type1. On the other hand, Type 3 is revealed to have a worse performance than that of Types 1 and 2 in all cases. This implies that the prediction of the restarting time for operation should be made "conservatively", in other words, it is better to predict that the operation will restart late, not early, in order not to increase halt discomfort.

By observing the small difference between the results for Types 2 and 4, and that between Types 3 and 5, it can be concluded that the value of the prediction in the early stages has a significant effect on the performance of the regulation, while the prediction at later times does not. This result suggests that the first prediction should be made "conservatively"; alternative "impetuous" predictions lead to changes in prediction at later times and do not improve the performance of the regulation.

Comparing the results of the cases for each Type, it can be clearly observed that, when the actual restarting time is later than the predicted time, the performance of the regulation deteriorates visibly. This result also suggests that the prediction should be made "conservatively".

11.5 Conclusions

The new phase-by-phase regulation shows the best performance for all types of variation of the prediction over time; nevertheless, it is clear that the variation of prediction over time has a significant effect on the performance of the regulation.

The results clearly demonstrates that the prediction should be made "conservatively", in other words, the operation should be predicted to restart late, not early, in order to avoid increasing the halt discomfort.

Delay of the restarting time beyond the predicted time is also the cause of a worse performance of the regulation. In the simulation, the deviation either 60s or 120s, which is not regarded as a long period compared with the initial delay (600s to 1200s).

Nevertheless, when the actual restarting time is 60s to 120s later than the predicted time, the performance of the regulation becomes visibly worse.

12 Conclusions and further study

12.1 Restatement of the outline of this research

This research aimed at developing "metro-type" railway traffic regulation, that is, a set of methods for handling disturbed traffic on metros. The target of this research has been limited to metro-type railways, since they generally have to handle high demand with limited facilities. In such a situation, railway systems, especially metro-type ones, may well be unstable, thus traffic regulation is more important than for other kinds of railways such as inter-city lines, rural lines, and so on.

The characteristics of this research are (1) to emphasise the passenger perspective rather than the operator's point of view, and (2) to use optimisation ideas, rather than depend on solely heuristic rules. Regarding (1), this research has used four kinds of passenger discomfort to evaluate the performance of regulation from the passenger perspective. They are waiting discomfort, travelling discomfort, congestion discomfort, and halt discomfort. The first three kinds of discomfort have already been introduced in previous work, while the last one is introduced within this research project. Regarding (2), the regulation method includes an element of optimisation, in this case, to minimise the total value of the four kinds of discomfort.

In this research, the regulation method applied is limited to changing departure and arrival times of trains from/at stations. Other methods such as changing destinations of trains, closing a part of the line, and preparing extra trains are not allowed in this research. Therefore, the problem that this research has tried to solve can be restated as finding the optimum set of departure and arrival times of trains from/at stations to minimise the passenger discomfort in disturbed traffic.

12.2 Conclusions of this research

The conclusions of this research are explained here, chapter-by-chapter, excluding Chapter 2, which included only the simulation methods used in this research, and does not of itself lead to conclusions as such.

Chapter 3 has been devoted to giving a review of existing work. In this chapter, the Original Murata regulation has been explained in detail, since it plays a basic foundation for establishing the new regulation method in this research. The definitions of the passenger discomfort have also been given.

The Original Murata procedure calculates the optimum solution by classical optimisation methods with four techniques adopted to reduce computation time. They are gradient calculation, gradient projection, sliding window and recursive optimisation. The objective function is the total discomfort value of three, not four, kinds of discomfort criteria, that is, the Original Murata regulator does not include halt discomfort in its objective function. The variables are departure and arrival times of trains from/at stations.

The basic idea of optimising railway traffic from the passenger perspective has been confirmed in Chapter 3.

In Chapter 4, simulation conditions have been explained concerning how to choose suitable parameters, since, in performing a simulation, it is important to choose appropriate conditions that reflect real circumstances reasonably well. The parameters discussed in this chapter are the ratio of the number of trains to the number of platforms, origin-destination types (or, OD types), the disturbance site, and the duration of disturbance.

For example, based on the model line used in this research, which is the Hong Kong Island Line, when the headway is 90s, the ratio of the number of trains and the number of platforms is 1.06. If the headways are 75s and 120s, the ratios are 1.27 and 0.82, respectively. These values cover the range of the ratios in the real life metro-type railways such as the London Underground, Paris Metro, and Tokyo Metro.

It has been shown that these conditions not only describe the real world satisfactorily, but also provide harsh conditions for traffic regulation. Consequently, it has been concluded that these conditions are a suitable basis for performing simulations and examining the performance of regulation algorithms.

After choosing appropriate parameters, the Original Murata regulation has been simulated under these conditions. Results of the simulation have been given in Chapter 5.

The Original Murata regulator has been shown to have a good performance in reducing the existing three kinds of discomfort criteria (waiting discomfort, travelling discomfort, and congestion discomfort), however, it makes no attempt to prevent the trains following the disturbed train from stopping behind the disturbed train. A real railway operator would try to avoid such a situation, since passengers may feel severe discomfort when their train is stuck between stations, which, in many cases, is inside a tunnel.

In order to remove this problem, a new objective criterion is necessary to describe the extent of passenger discomfort that is stated above. Also, the assumption in the Original Murata regulator that the disturbed train leaves the station at the start of each regulating calculation needs to be reconsidered, since this assumption leads to the situation that the following trains leave the preceding stations without knowing that there is a disturbed train in front of them.

Following the above discussion, a new discomfort criterion called "halt discomfort" has been introduced in Chapter 6. Re-formulation of the Murata regulator (referred to as "Murata plus halt") has been performed. A simple sensitivity analysis has shown the limit of the existing three kinds of discomfort criteria, which also explains the necessity of introducing the new discomfort criterion. A definition of halt discomfort has been given.

After halt discomfort has been defined, the Murata plus halt regulation has been simulated. The results have clarified the limitations of the Murata approach, and the study has thus identified the need for new techniques to be introduced into the regulation.

Chapter 7 has discussed an observation from real railway operation, that is, railway operators tend to stop all trains at the nearest station immediately after traffic is disturbed. This method has been named as the "stop-all-trains-at-once" strategy.

The performance of the regulation methods that have appeared until this chapter have been measured including halt discomfort. The stop-all-trains-at-once strategy has been shown to be fairly effective in preventing halt discomfort from rising suddenly, but is not effective in reducing the other three kinds of discomfort criteria. Nevertheless, the combination of the ideas of Murata and the stop-all-trains-at-once strategy would appear to be effective in reducing halt discomfort, while not making the other kinds of discomfort much worse.

After the discussion in Chapter 5 and Chapter 7, the new "phase-by-phase" regulation strategy has been introduced in Chapter 8. Thereafter, Chapters 9-11 have reviewed the results of the new regulation from various perspectives. There are the relationship of the number of trains to the number of stations on the line where the disturbance happens and variations of the restarting prediction over time.

In Chapter 9, the simulation results have been analysed for the effect of the relationship between the number of trains running on the line and the number of platforms. With two additional simulations performed in order to cover the wide range of the parameters, it has been revealed that, when the ratio of the number of trains to the number of platforms is high, the stop-all-trains-at-once strategy is not effective in reducing halt discomfort, since there will then be some trains that have to stop outside the stations.

The new phase-by-phase regulation is effective in all the cases simulated. This means that the mixture of the Murata approach and the stop-all-trains-at-once strategy works as predicted.

Chapter 10 has been devoted to examining the results of the simulation as they are affected by the position of the disturbance site. In order to avoid a large number of simulation cases, five types of OD have been chosen as the most independent examples with respect to each other. The results have shown the clear differences between the OD types. In order to explain the differences in more detail, the ratio of the number of passengers on a train running towards the disturbance site and the maximum number of passengers on a train has been introduced. The combination of this parameter and the disturbance site fixed at Station 3 has provided a clear view of the characteristics of each regulation method.

From this viewpoint, it has been demonstrated that the performance of the stop-all-trains-at-once strategy is more effective in reducing halt discomfort when the disturbance happens on a busy section of the line. On the other hand, this strategy is not effective when the disturbance happens on a less crowded section of the line, since it forces any train to stop no matter how far from the disturbance site it is running. Again, the new phase-by-phase regulation is effective in all the cases simulated.

The effect of the variation of the restarting prediction with time has been explained in Chapter 11. The cases where the prediction is put forward and backward as time goes by, and the case where the prediction is accurate from the first attempt, and thus does not vary over time, have been examined.

The new phase-by-phase regulation has again shown the best performance for all types of the variation of prediction over time; however, it is clear that the variation of prediction over time has a significant effect on the performance of the regulation.

Overall, the prediction should be made "conservatively", in other words, the operation should be predicted to restart late, rather than early, in order to avoid an increase in halt discomfort.

Further delay of the restarting time behind the predicted time has also been shown to be the cause of a worse performance of the regulation. When the actual restarting time is 60s to 120s later than the predicted time, the performance of the regulation becomes noticeably worse.

12.2 Further study

Some subjects to be studied in the future are outlined here.

12.2.1 Regulation methods

In this research, regulation methods are limited to changing departure and arrival times of trains from/at stations. There are some other strategies that can be adopted in real railway operations when traffic is disturbed, such as changing destinations of trains, closing a part of the line, and preparing extra trains. These have not been considered in this research. They have been excluded from the present study, since they are usually applied in the case of large scale disturbances, e.g., several-hour stoppages of the service due to accidents, whereas the target of this research is smaller disturbances.

As has been stated in this thesis, a significant discomfort of passengers will emerge if no regulation is performed when a small disturbance happens. Only changing departure and arrival times appears to be effective in avoiding such passenger discomfort.

When a much larger disturbance happens, solely changing departure and arrival times is insufficient to remove the disturbance and get back to a stable state of traffic. Some other methods, such as those mentioned above, must be taken. Dealing with a much larger disturbance in such a way is, however, also one of the most difficult items to be fully computerised in the railway industry.

If the targets for this research are to be widened, the regulation methods may have to be reconsidered, as will be discussed next.

12.2.2 Type of a targeted line

This research has targeted a metro-type railway, since the problem of the instability under disturbed conditions may be paramount in metro-type railways.

As has been stated in the chapter of introduction, a typical example of a metro-type railway is an underground railway in a big city such as the London Underground, the New York City Subway, or Tokyo Metro. The main characteristics of a metro-type railway are as follows.

(1) All trains stop at all stations. (With a few exceptions.)

(2) Inter-station distances are relatively short. A typical value is 1 to 2 stations per km.

(3) A metro-type railway generally has to handle high demand with limited facilities. Thus, it tends to have the problem of congestion, which becomes more serious in disturbed traffic.

When other kinds of railways are considered, the regulation methods may change according to the characteristics of the targeted line. For example, looking at a suburban railway that has both an express service and a local service (which may stop at all stations), the regulation methods may have to be reconsidered, since the methods so far explored are limited only to changing departure and arrival times. If other methods are not allowed, an express service can only be regulated at its pre-determined calling stations, which means that an express service has only a few chances to be regulated. It would thus seem appropriate to extend the regulation methods so that express trains may have to call at stations which were not originally on their schedule.

12.2.3 Accuracy of prediction

As has been explained when the new phase-by-phase regulation was introduced, the accuracy of prediction of the restarting time is expected to have a significant effect on the performance of the regulation. Thus, the effect of the accuracy of prediction has been examined by simulating the cases where the prediction varied over time. Furthermore, the cases where a disturbed train restarts its operation some time later or earlier than the predicted time or exactly at the predicted time have also been examined.

The simulation results have shown that the accuracy of prediction does have an effect on the performance of the regulation.

Therefore, methods for obtaining as accurate a prediction as possible when a disturbance happens need to be explored. This may include how to establish a database of past accidents and incidents, how to allocate field engineers along the line, how to build the communication tools between the trains and the control centre, and so on.

References

[1] Araya, S. and Sone, S., "Traffic Dynamics of Automated Transit Systems with Pre-established Schedules", IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-14, No.4, pp.677-687, July/August 1984

[2] Bellman, R., "Dynamic programming", Princeton University Press, New Jersey, 1957

[3] Bellman, R. E. and Dreyfus, S. E., "Applied Dynamic Programming", Princeton University Press, New Jersey, 1962

[4] Chambers, L., "The Practical Handbook of Genetic Algorithms Applications", Chapman & Hall/CRC, Florida, 2001

[5] Chang, C. S. and Du, D., "Improved optimisation method using genetic algorithms for mass transit signalling block-layout design", IEE Proceedings on Electric Power Applications, Vol. 145, No. 3, May 1998

[6] Chang, C. S. and Du, D., "Further improvement of optimisation method for mass transit signalling block-layout design using differential evolution", IEE Proceedings on Electric Power Applications, Vol. 146, No. 5, September 1999

[7] Chang, C. S. and Sim, S. S., "Optimising train movements through coast control using genetic algorithms", IEE Proceedings on Electric Power Applications, Vol. 144, No. 1, January 1997

[8] Chang, C. S. and Tian, L. F., "Worst-case identification of touch voltage and stray current of DC railway system using genetic algorithm", IEE Proceedings on Electric Power Applications, Vol. 146, No. 5, September 1999

[9] Chang, C. S. and Xu, D. Y., "Differential evolution based tuning of fuzzy automatic train operation for mass rapid transit system", IEE Proceedings on Electric Power Applications, Vol. 147, No. 3, May 2000

[10] Chang, C. S., Jiang, W. Q. and Elangovan, S., "Application of genetic algorithms to determine worst-case switching overvoltage of MRT systems", IEE Proceedings on Electric Power Applications, Vol. 146, No. 1, January 1999

[11] Chang, C. S., Low, J. S. and Srinivasan, D., "Application of tabu search in optimal system design and operation of MRT power supply systems", IEE Proceedings on Electric Power Applications, Vol. 146, No. 1, January 1999

[12] Chang, C. S., Wang, F. and Lock, K. S., "Harmonic worst-case identification and filter optimal design of MRT systems using genetic algorithms", IEE Proceedings on Electric Power Applications, Vol. 144, No. 5, September 1997

[13] Chang, C. S., Wang, W., Liew, A. C. and Wen, F. S., "Bicriterion optimisation for traction substations in mass rapid transit systems using genetic algorithm", IEE Proceedings on Electric Power Applications, Vol. 145, No. 1, January 1998

[14] Chang, C. S., Xu, D. Y. and Quek, H. B., "Pareto-optimal set based multiobjective tuning of fuzzy automatic train operation for mass transit system", IEE Proceedings on Electric Power Applications, Vol. 146, No. 5, September 1999

[15] Chua, C. K., "Disturbed running on metro systems: a simulation study", MPhil thesis, University of Birmingham, 1986

[16] Croall, I., F. and Mason, J., P., "Industrial Applications of Neural Networks", Springer-Verlag, Berlin, 1992

[17] Dixon, L.C.W., "Nonlinear Optimisation", The English Universities Press Limited, London, 1972

[18] Fernandez, A., de Cuadra, F. and Garcia, A., "SIRO: an optimal regulation system in an integrated control centre for metro lines", Computer in Railways V, Vol. 2, pp.299-308, Computational Mechanics Publications, Southampton, 1996

[19] Fernandez, A., de Cuadra, F. and Montes, F., "Traffic regulation and simulation – a predictive adaptive control system", Computer in Railways IV, Vol. 1, pp.419-427, Computational Mechanics Publications, Southampton, 1994

[20] Gen, M. and Cheng, R., "Genetic Algorithms & Engineering Optimization", John Wiley & Sons, Inc., New York, 2000

[21] Gill, P. E., Murray, W. and Wright, M. H., "Practical Optimization", Academic Press, London, 1981

[22] Goldberg, D. E., "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison-Wesley Publishing, 1989

[23] Goodman, C. J. and Murata, S., "Metro traffic regulation from the passenger perspective", IMechE Vol. 215, Part F, pp.137-147, 2001

[24] Grefenstette, J. J., "Genetic Algorithms for Machine Learning", Kluwer Academic Publishers, Massachusetts, 1994

[25] Holland, J. H., "Adaptation in Natural and Artificial Systems", The MIT Press, Massachusetts, 1992

[26] Maciejowski, J. M., "Predictive Control with constraints", Prentice Hall, Harlow, England, 2002

[27] Masters, T., "Practical Neural Network Recipes in C++", Academic Press, Inc., San Diego, California, 1993

[28] Michalewicz, Z., "Genetic Algorithms + Data Structures = Evolution Programs, Third, Revised and Extended Edition", Springer-Verlag, Berlin, 1996

[29] Mitchell, M., "An Introduction to Genetic Algorithms", The MIT Press, Massachusetts, 1996

[30] Negnevitsky, M., "Artificial Intelligence", Addison Wesley, Essex, 2002

[31] Page, G. F., Gomm, J. B. and Williams, D., "Application of Neural Networks to Modelling and Control", Chapman & Hall, London, 1993

[32] Pham, D. T., "Neural Networks for Identification, Prediction and Control", Springer-Verlag London Limited, 1995

[33] Pham, D. T. and Karaboga, D., "Intelligent Optimization Techniques", Springer-Verlag London Limited, 2000

[34] Prime, H. A., "Modern concepts in control theory", McGraw-Hill, London, 1969

[35] Rao, S. S., "Optimization theory and applications" second edition, Wiley Eastern Limited, New Delhi, 1984

[36] Siu, L. K., "An Object-oriented Railway System and Power Network Simulator", PhD thesis, University of Birmingham, 1995

[37] Swingler, K., "Applying Neural Networks --- A Practical Guide", Academic Press, London, 1996

[38] Thia, B. S., "On-line Rescheduling of Rapid Transit Systems – a fuzzy expert system approach", Master thesis, National University of Singapore, 1995

[39] Tou, J. T., "Modern Control Theory", McGraw-Hill, London, 1964

[40] Van Breusegem, V., Campion, G. and Bastin, G., "Traffic Modeling and State Feedback Control for Metro Lines", IEEE Trans. on Automatic Control, Vol. 36, No. 7, pp.770-784, July 1991

[41] Warwick, K., Irwin, G. W. and Hunt, K. J., "Neural Networks for control and systems", Peter Peregrinus Ltd., London, 1992

[42] Winter, G., Periaux, J., Galan, M. and Cuesta, P., “Genetic Algorithms in Engineering and Computer Science”, John Wiley & Sons Ltd, Chichester, 1995

[43] Yen, J., Langari, R. and Zadeh, L. A., “Industrial Applications of Fuzzy Logic and Intelligent Systems”, IEEE Press, New York, 1995

Appendix: Publication

1. Ueda, N., Goodman, C. J. & Sone, S.: "Introduction of halt discomfort in the objective criteria of regulation for metro type railways", Computers in Railways IX (Proceedings of the Ninth International Conference on Computer Aided Design, Manufacture and Operation in the Railway and Other Mass Transit Systems (COMPRAIL 2004)), pp. 753-762, Dresden, Germany, May 2004.

Not available in the digital copy of this thesis