Analysis of Distributions of Train Speed, Headway and Buffer Time: The Hague Case

Nie Lei（聂 磊）著

列车运行速度、间隔时间及缓冲时间的分析：海牙案例

北京交通大学出版社

Analysis of Distributions of Train Speed, Headway and Buffer Time: The Hague Case

列车运行速度、间隔时间及缓冲时间的分析：海牙案例

Nie Lei（聂　磊）著

北京交通大学出版社

·北京·

内 容 简 介

本书主要介绍了荷兰铁路列车运行数据分析理论与方法，以及采用的计算机分析系统和数据分析过程。本书以荷兰海牙地区铁路运营条件和实绩运行数据为例，详细地分析了列车运行图、基础设施布局、机车车辆性能等既定条件下，列车实际运行速度、实际间隔时间和实际缓冲时间的分布情况，指出了列车运行图及其他运输组织方案存在的问题，并根据分析结果提出相应的改进措施。

本书可用于相关专业技术人员的科研参考书，亦可用于铁路运输专业研究生的教学参考书。

图书在版编目（CIP）数据

列车运行速度、间隔时间及缓冲时间的分析：海牙案例 = Analysis of Distributions of Train Speed, Headway and Buffer Time: The Hague Case：英文／聂磊著．—北京：北京交通大学出版社，2008.3

ISBN 978-7-81123-239-4

Ⅰ.列… Ⅱ.聂… Ⅲ.列车-运行-研究-海牙-英文 Ⅳ.U292.4

中国版本图书馆 CIP 数据核字（2008）第 023986 号

责任编辑：张利军

出版发行：北京交通大学出版社　　电话：010-51686414

北京市海淀区高梁桥斜街 44 号　　邮编：100044

印 刷 者：北京市梦宇印务有限公司

经　　销：全国新华书店

开　　本：170×235　印张：10　字数：206 千字

版　　次：2008 年 4 月第 1 版　2008 年 4 月第 1 次印刷

书　　号：ISBN 978-7-81123-239-4/U·20

印　　数：1～1 000 册　定价：25.00 元

本书如有质量问题，请向北京交通大学出版社质监组反映。

投诉电话：010-51686043，51686008；传真：010-62225406；E-mail：press@bjtu.edu.cn。

Preface

前　言

国外铁路特别是高速铁路具有非常高的正点率，这在很大程度上得益于研究人员对提高正点率的理论与方法体系进行了深入细致的研究。根据列车运行实绩反馈指导列车运行图及基础设施和机车车辆运用方案的编制是其中的一个重要研究方向。由于我国前些年没有采集列车运行实绩的条件，而且也缺乏相应的数据分析计算机系统，到目前为止相关的研究很少。

本书是作者在欧洲著名理工大学代尔夫特理工大学（Delft University of Technology）访问工作期间，利用该大学的计算机系统和荷兰铁路的列车运行数据（精确到秒）所做的研究。本书数据虽然来自荷兰，但其理论与方法值得我们借鉴。随着我国客运专线的大量建设，旅客对列车运行正点率的要求日益提高，学习这些精确的数据分析方法有助于改善我国铁路运输组织方案，提高列车运行质量。

本书以完整的案例，系统地介绍了列车运行数据分析的理论与方法，以及采用的计算机分析系统和数据分析过程。本书的主要内容已用于博士研究生学位课程“列车运行过程组织理论与方法”等的教学。

本书在编写过程中参考了大量的文献资料，在此对这些资料的编写者表示衷心的感谢，同时也感谢代尔夫特理工大学教授 I. A. Hansen 为作者提供从事该项研究的机会。

本书的出版得到北京交通大学“轨道交通控制与安全国家重点实验室”和北京交通大学学术专著出版基金的支持，也得到了铁道部科技项目（2005F022）和校基会项目（2005KZ006）的支持，在此一并表示感谢。

由于编写时间紧促，错误和疏漏在所难免，真诚欢迎各位专家和同行批评指正，以便修改完善。

作　者

2008 年 4 月

Contents

目　　录

Chapter 1

Introduction

1.1 Outline of the Research

Train delays are one of the most important performance indicators of railway operation, which is also the result of many related factors, such as the layout of the infrastructure, the occupancy of the infrastructure, the variation of running times, arrivals and departures, and the traffic control etc. At first, the infrastructure must provide enough capacity and smooth routes. The timetable is the basic plan for train operation. On the one hand, the timetable should ensure the efficient and balanced use of the infrastructure. On the other hand, it be flexible enough to cope with stochastic disturbances during operations. For example, proper buffer times in a timetable are very important to reduce train delays and delay propagation. Stochastic disturbance of operations can not be eliminated and will always occur, but its statistical distribution can be grasped by analyzing historic and actual train operation data. To some extent, this will help to forecast and prevent train delays. Traffic control is the final means to keep train punctuality. It is always aimed to improve the efficiency of traffic control, e.g. by optimizing the actual schedule, in the development of railway science and technology. The interdependencies among the related factors are strengthened with the increase of traffic density and train speed, not only, which requires a further enhancement of the precision of time calibration, a more comprehensive and accurate analysis of train operations.

By means of the developed tools TNV-Prepare and TNV-Filter (Goverde, et al, 2000a and 2000b), train detection data and train delays at stations can be analyzed with a high precision of about a second. Earlier publications (Goverde, et al, 2001a and 2001b, Yuan 2001) deal with the distribution of train delays in some stations of the Dutch Railways. Tromp (2001) analyzes in detail one of the crossings East of Eindhoven station and reveals a significant increase of the blocking times of the hindered trains resulting in much higher occupancy of the critical track sections and a reduction of capacity. Further, Goverde (2005) performs a detailed punctuality analysis with the emphasis on fitting probability distributions of characteristic operation performance indicators and quantifying disruptive train dependencies, and develops an analytical approach to evaluate and quantify critical network dependencies on capacity utilization and timetable stability. By use of the database obtained from TNV-systems, Yuan (2006) gets more insight into the stochastic characteristics of train delay, delay propagation and real capacity utilization, and develops a probability model to predict the knock-on delays caused by route conflicts and later transfer connections in stations.

Other publications analyze the stability and robustness of railway timetables using analytical models, mathematical programming or micro-simulation. Wendler (1999) presents a stochastic model for the estimation of scheduled waiting time at station tracks for train-triples and gives a closed solution for the problem of non-utilizable time gaps at branching-off points. Huisman & Boucherie (2001) provide a queuing model reflecting the approximate dependence between free running times and scheduled headway times. The forecasted running time distribution for each train service is obtained by solving a system of linear differential equations based on Markov chains.

Zwaneveld et al. (1996) and Kroon et al. (1997) propose a Linear Programming model for the determination of optimal routes and train sequences through stations depending on scheduled train running times and fixed minimal headways. They present a branch-and-cut solution for the Weighted Node Packing Problem and applied it to a number of Dutch railway stations. Powell & Wong (2000) determine the maximum cycle performance for particular terminal station layout with up to six platforms by means of Integer Programming solved with a branch-and-bound algorithm. Billionet (2003) uses a standard commercially available integer programming software like MLP or AMPL for solving the train-platforming problem. When applying an objective function, e.g. to maximize the use of a certain track, the computing time, however, became too important. Kroon & Peters (2003) develop an optimization model for constructing cyclic timetables with variable trip times using periodic time window constraints in order to improve the robustness of a network timetable. For large timetabling problems, a cycle fixation heuristic is required to solve the algorithm in a reasonable amount of time.

Malavasi & Ricci (2001) analyze the stochastic elements in railway system using a neural model. They develop and test a self-learning simulation model capable to reproduce the impact of track occupations by delayed trains on the performance of other trains in a regional railway network of the Italian Railway. Zhu develops a simulation model based on stochastic Petri nets in order to assess the impact of incidents on the quality of operations. With a simple sensitivity analysis, the most critical elements of the network could be identified. Kaminsky (2001) introduces a so called 'buffer train' to the blocking times of each train path compensating up to 80% of the cumulative primary delays according to an assumed negative-exponential distribution. After establishing a conflict-free timetable for a large network in a part of the German Railway, the corresponding distribution of buffer times by time and space is evaluated by means of the simulation tool *Railsys*.

Rodriguez et al. (2002) develop a constraint programming model for the routing of trains in saturated corridors with explicit representation of the capacity constraints by block signals and of the temporal constraints by occupation and release of the track circuits. They apply a greedy algorithm, while the set of feasible routes is restricted to one route. Carey & Carville (2000, 2003), after having investigated the possibility for solving the train scheduling at complex stations by integer programming methods, finally present a heuristic approach for resolving conflicts between train paths and routes, while satisfying track infrastructure and headway constraints. They test the reliability and robustness of timetable options by means of simulating exogenous random delays in order to compare the costs and penalties for preferred train times and platforms including knock-on delays.

From the comprehensive literature review it becomes clear that there exist nearly no publications with regard to the variation of train speed, blocking and buffer times supported by real-world operations data. So far, the research is limited to experiments by means of simulation or mathematical analysis applying assumed distributions of primary delays and deterministic running and minimum headway times of trains.

In this book, the possible conflicting points of two main stations in The Hague area and the route allocation are discussed. Considering the train service in practice, the distribution of speed, blocking and buffer times between two successive trains at the critical crossing are computed, and the occupancies of each station track and main points are estimated. Furthermore, based on real-world train detection data, the distribution of the blocking and buffer times observed at conflicting points are determined, which explain the level of train delays.

1.2 Train Detection Data and TNV-tools

1.2.1 Train Detection Data

TNV-systems are the Dutch train describer systems that keep track of the progress of trains based on train numbers and infrastructure messages (Goverde, 2005).

In the Netherlands, train running data is generally obtained from track-side measuring equipment. In that, track sections play a key role. These sections are the elementary track segments (a pair of points, a platform track etc.) which occupation and release are determined by track circuits or axle counters. The occupancy information is primarily recorded by TNV-systems (Stam, 2007).

1.2.2 TNV-Prepare

The TNV-systems produce log files which were initially meant for incident reconstruction, containing chronological records of train number events, track section occupancy, and points and signal status with a precision of one second. The files are not directly suitable for operational analysis (Stam, 2007).

TNV-Prepare was developed during 1998/2000 in TU Delft. It is an empirical analysis tool that converts TNV-logfiles into tables, so called TNV-tables, of successive events on a route of a train line, including TNV steps and state changes of sections, signals, and switches. TNV-Prepare generates the entire trajectory at section level of incoming and outgoing routes.

1.2.3 TNV-Filter

The actual arrival and departure times at a platform stop are in general not measured and hence not recorded automatically. The closest measured event to an arrival is the platform track section occupancy, or even more close, the clearance of the last section before the platform track section. Likewise, a departure can be approximated by the measured occupancy of the track section just after the platform track section. The running time before and after the stop on the platform track section may nevertheless be considerable depending on the local station layout (platform track section length) and stop location at the platform track. TNV-Filter (Goverde, 2005)has been developed to estimate the accurate speed of approaching and departing trains, resulting in reliable and accurate arrival and departure time estimates.

As for TNV-tools and databases lack either detail or convenience. Recently, a user friendly software VTL-Tool (Stam, 2007) has been developed at NS Transport Control to generate the accurate train running data with a new data format. Additional analysis functionalities have been implemented, including calculation of departure and arrival times, identification of route conflicts, and flexible visualisation of performance statistics. The VTL-Tool has so far been applied in several operations quality projects in the Dutch railway industry.

1.3 Statistical Preliminaries

In this section we briefly introduces the normal, exponential probability distributions, outliers, and the area of statistical hypothesis testing, which are used throughout the book (Goverde, 2001a; Yuan, 2006).

1.3.1 Probability Distributions

The normal distribution naturally arises from the application of the central limit theorem which can roughly be stated as "the sum of many independent random variables with finite mean and variance is approximately normally distributed". A continuous random variable X has a normal distribution, or Gaussian distribution, if it has the probability density function

$$f(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},$$

for $-\infty<x<\infty, -\infty<\mu<\infty$, and $\sigma>0$. The two parameters are the mean and standard deviation, respectively. The normal density is symmetrical around the mean value $x=\mu$ where the density has a maximum, and the tails diminish exponentially with the square of x and a rate determined by σ.

The exponential distribution is widely used to describe phenomena involving events that occur randomly in time or space but at a constant rate. A continuous random variable X has a one-parameter exponential distribution, or negative exponential distribution, if it has the probability density function

$$f(x;\lambda)=\lambda e^{-\lambda x},$$

for $x \geqslant 0$ and $\lambda>0$. The parameter λ is the reciprocal of the mean, i.e., $E(X)=\frac{1}{\lambda}$. The exponential distribution is monotonously decreasing with maximum value λ at $x=0$.

1.3.2 Outliers

An outlier is a data point that deviates from the bulk of the data. For example, the mean of a sample is highly influenced by one extremely large (or small) value. The computation of statistics and the estimation of distribution parameters can be influenced by outliers. An outlier may be the result of a measurement or calculation error. However, it is also very well possible that it is a real observed. Assume for example that we have a sample of train departure delays including one train that departed 50 minutes late due to all unexpected rare event. Then this late train is not a representative case of the departure delay sample, and should be identified as an outlier. On the other hand, if a random variable X is exponentially distributed with mean 2 then there is still a (small) probability of a long delay time, say $x = 50$.

The impact of an outlier depends on the sample size and dispersion (spread) of the data, and one should always be aware of their presence when analysing data. However, there is no (exact) standard procedure to detect outliers. Possible outliers can be identified by using robust statistics and also graphical presentation is an effective means of detection.

For a symmetric (normal) distribution the median (middle value) is a quite robust measure of location of the data, since moving the extreme observations does not effect the sample median. If we expect a random variable to be normally distributed the sample median and sample mean should be close to each other, since they both measure the centre of symmetry. So a possible outlier may be identified by comparing the sample mean and the sample median (of normal data).

The standard deviation is also sensitive to outlying observations. A robust measure of dispersion is the interquartile range (IQR) defined as the difference between the two sample quartiles (the 0.25-quantile and the 0.75-quantile). For a normal distribution the standard deviation is equal to the IQR divided by 1.35. Therefore, for normal data possible outliers may be detected by comparing the sample standard deviation and IQR/1.35.

A useful graphical display is the boxplot that shows a measure of location (the median), a measure of dispersion (the IQR), the presence of possible outliers, and gives an indication of the symmetry and skewness of the distribution.

1.3.3 Statistical Tests

A statistical test is a probability-based method to determine whether a given parameter or model (the null hypothesis) correctly characterizes the observations. Recall that a statistic is a function of data and is thus a random variable that depends on a particular sample. Some simple examples of statistics are the mean or variation.

A statistical test is based on the value of a test statistic T that has the property that its probability distribution is (approximately) known if the null hypothesis H_0 is true. The probability that the test statistic exceeds its evaluated value t assuming that the null hypothesis is true, is called the p-value p^*,

$$p^* = P\{T \geqslant t \mid H_0 \text{ is true}\}.$$

The decision to accept the null hypothesis or reject the null hypothesis in favor of the alternative hypothesis now depends on the choice of a significance level α, which is the probability of rejecting the null hypothesis when it is in fact true. If the p-value is smaller than the significance level then the null hypothesis is rejected and otherwise the null hypothesis is accepted. In the sequel we will use the significance level $\alpha = 0.05$. The null hypothesis of a test will thus be accepted, if

$$p^* = P\{T \geqslant t \mid H_0 \text{ is true}\} \geqslant 0.05.$$

A statistical test is called a one-sample test if a set of observations is tested against a theoretical model. A two-sample test, on the other hand, is a test that compares a statistic of two sets of observations. Below we consider some statistical tests that have been used throughout the book.

1. One-sample t-test

This test is used to test whether the mean for a variable has a particular value. The main assumption in a t-test is that the data comes from a normal distribution. If this is not the case, then a nonparametric test, such as the Wilcoxon rank-sum test, may be a more appropriate test of location.

2. Two-sample t-test

This test is used to test whether two samples come from distributions with the same means. The samples are assumed to come from normal distributions. If this is not the case, then a nonparametric test, such as the Wilcoxon rank sum test, may be a more appropriate test of location.

3. Wilcoxon rank sum test

This two-sample test is used to test whether two sets of observations come from the same distribution. The alternative hypothesis is that the observations come from distributions with identical shape but different locations. Unlike the two-sample t-test, this test does not assume that the observations come from normal distributions. This test is equivalent to the Mann-Whitney test.

4. F-test

This two-sample test is used to test whether two sets of observations have equal variance. It is assumed that both data sets are drawn from normal populations. Outliers in the data may have a significant effect on the results through their relatively strong influence on the variance estimates.

5. One-sample Kolmogorov-Smirnov goodness of fit test

This goodness of fit test is used to test whether the empirical distribution of a set of observations is consistent with a random sample drawn from a specific theoretical distribution. The Kolmogorov-Smirnov goodness of fit test is generally more powerful than the chi-squared goodness of fit test for continuous variables. For discrete variables, the chi-squared test is generally preferable.

6. Two-sample Kolmogorov-Smirnov goodness of fit test

This goodness of fit test is used to test whether two sets of observations could reasonably have come from the same distribution. This test assumes that tile samples are random samples, the two samples are mutually independent, and the data are measured on at least an ordinal scale. In addition, tile test gives exact results only if the underlying distributions are continuous.