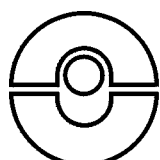


Information requirements of operating departments concerning the transport process, delays and other disruptions; criteria for assessing operating efficiency and measures for speeding up the operating process

Besoins d'informations du mouvement sur le déroulement du transport, les retards et autres perturbations ; critères permettant d'apprécier la qualité d'exploitation et mesures visant à améliorer l'exploitation

Bedarf der Betriebsleitung an Informationen über die Transportabwicklung, Verspätungen und sonstige Störungen; Kriterien zur Bewertung der Betriebslage und Maßnahmen zur Verbesserung der Betriebsführung



*Union Internationale des Chemins de fer
Internationaler Eisenbahnverband
International Union of Railways*



Document Sharing Hub



Leaflet to be classified in Section :

IV - Operating

Application :

With effect from 1st January 1983

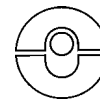
All members of the International Union of Railways

Record of update :

1st edition, January 1983 First edition

2nd edition, December 2000 Retyped in FrameMaker

The person responsible for this leaflet is named in the UIC Code



Warning

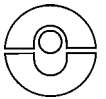
No part of this publication may be copied, reproduced or distributed by any means whatsoever, including electronic, except for private and individual use, without the express permission of the International Union of Railways (UIC). The same applies for translation, adaptation or transformation, arrangement or reproduction by any method or procedure whatsoever. The sole exceptions - noting the author's name and the source - are "analyses and brief quotations justified by the critical, argumentative, educational, scientific or informative nature of the publication into which they are incorporated"

(Articles L 122-4 and L122-5 of the French Intellectual Property Code).

© International Union of Railways (UIC) - Paris, 2000

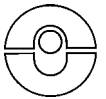
Printed by the International Union of Railways (UIC)
16, rue Jean Rey 75015 Paris - France, December 2000
Dépôt Légal December 2000

ISBN 2-7461-0230-7 (French version)
ISBN 2-7461-0231-5 (German version)
ISBN 2-7461-0232-3 (English version)

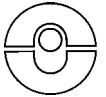


Contents

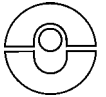
Summary	1
1 - Capture of source data	2
1.1 - General	2
1.2 - Forms of source-data capture.....	2
1.3 - Techniques for source data capture	3
2 - Processing of source data	4
2.1 - General	4
2.2 - Manual processing.....	4
2.3 - Computer processing.....	4
2.3.1 - General	4
2.3.2 - Daily processing.....	5
2.3.3 - Monthly and annual processing	5
2.4 - Evaluation of reference values	5
2.4.1 - General	5
2.4.2 - Simple reference values	5
2.4.3 - Compound reference values	6
2.5 - Presentation of results	6
2.5.1 - General	6
2.5.2 - Presentation in tables	6
2.5.3 - Presentation in graphs	7
2.5.4 - VDU and microfiche presentation	7
3 - Use of results	8
3.1 - General	8
3.2 - Analysis of time series	8
3.2.1 - General	8
3.2.2 - Determination of trends.....	9
3.2.3 - Determination of seasonal influences	9
3.3 - Forecasting of expected values	9
3.4 - Relationships between reference values	10



3.5 - Influencing of reference values.....	11
Appendix A - Plan showing data-capture and processing techniques	13
Appendix B - List of source data.....	14
Appendix C - Specimen table of daily values and aggregates compared with expected values and figures for the previous month	17
Appendix D - List of simple reference values	18
D.1 -Transport process	18
D.1.1 - Performance of tractive unit	18
D.1.2 - Performance of train	18
D.1.3 - Local services (aggregates per specified station)	19
D.1.4 - Line occupancy (aggregates per line)	19
D.2 -Disruptions	19
D.3 -Delays	20
Appendix E - List of compound reference values	21
Appendix F - Specimen histogram showing gross tonne-km for a one-year period with traction-mode percentages	23
Appendix G - Specimen of a frequency polygon showing comparisons of the number of trains cancelled and number of special trains for a reference period of one month	24
Appendix H - Specimen of a circular diagram with segments showing the percentage for splitting of trains (43%), handling and exchange of wagons (4%), new formation (46%) and locomotive change (7%) in relation to the total number of wagons handled.....	25
Appendix I - Examples of calculations for analysis of time series and forecasting of basic values.....	26
I.1 -Analysis of a time series with trend and seasonal influence	26
I.1.1 - Determination of trend according to the method of moving averages for a three-month period	26
I.1.2 - Determination of trend according to the method of orthogonal functions	28
I.1.3 - Determination of seasonal variations according to the chain index method	30
I.2 -Determination of seasonal variations in stationary time series by phase-average method	33
I.3 -Forecasts of basic values using the exponential smoothing method	35
I.3.1 - General	35
I.3.2 - Determination of initial values	37
I.3.3 - Choice of smoothing constants.....	38
I.3.4 - Calculation of forecast values	38
I.4 -Comparison of the original basic values y_i with moving average values $\bar{y}_i$	43
I.5 -Comparison of values in the time series with first and second degree trends determined by means of orthogonal functions.....	44
I.6 -Seasonal figures indexed according to the chain index method (1978...1980)	45



I.7 -Relative differences in the form of seasonal index figures determined according to the phase average method (Sunday...Saturday over 13 weeks) ...	46
I.8 -Comparison of real values y_i with forecast values v_i for 1980.....	47
I.9 -Comparison of the three forecast values v_i , v_{i+1} , v_{i+2} for 1980	48
Bibliography	49



Summary

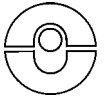
This leaflet stipulates the minimum volume of data to be captured on the transport process (services provided by trains and tractive units, local services, line occupancy) and on disruptions and delays as a basis for assessing the quality of the operating process from the point of view of planning and implementation, monitoring it at regular intervals and determining appropriate measures for its improvement.

The leaflet thus attempts to reduce the vast amount of available data to a reasonable and representative volume, and to create suitable conditions for bilateral and multilateral exchanges of information.

It contains suggestions and recommendations for data capture, processing and use.

A number of techniques that can be applied in various ways for data capture and processing are described in Appendix [A - page 13](#).

Neither information on commercial performance (e.g. loads carried), nor that on infrastructure and wagon fleets (e.g. number of wagons) has been included in the leaflet; the same applies to causes, circumstances and repercussions of serious operating incidents (especially rail accidents). The volume of such data and methods of capture, especially where research into causes is concerned, vary from one railway to another and in many cases are not the responsibility of the operating department.



1 - Capture of source data

1.1 - General

Source data

- is derived from the practical operating process and
- is captured directly, on the spot and at the time of occurrence (manual or automated capture).

As far as the work of operating departments is concerned, it is recommended that the source data listed in Appendix B - page 14 be captured.

This source data relates to

- the transport process, classified according to
 - tractive unit performance
 - train performance
 - local services
 - line occupancy
- disruptions and
- delays.

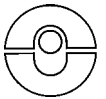
1.2 - Forms of source-data capture

Source data (see Appendix B - page 14) is captured mainly on an overall basis, though items may be classified into groups of items, e.g.

- tractive-unit performance as a whole and/or classified according to type of traction, or
- train performance as a whole and/or classified into train categories.

In certain cases a deliberate choice should be made to capture only part of the source data, for example to determine:

- wagon flows for a representative part of the timetable period (twice yearly, every two weeks);
- line occupancy, on
 - specific days of the week, e.g. Monday, Wednesday (representative for Tuesday to Thursday), Friday, Saturday, Sunday, or
 - one or more calendar days;
- the occupancy rate for major
 - marshalling yards or
 - lines only;



- train delays beyond a given threshold

and other information.

Railways are advised to make use of random selection of data (i.e. random sampling) as a form of partial capture, for example for "Statistical determination of wagon turnround times" (see *UIC Leaflet 403*), taking two points into account:

- items must be selected on a random basis;
- the degree of sampling error must be calculated.

The amount of data captured is significantly reduced by the use of random selection.

1.3 - Techniques for source data capture

As a general rule, source data is obtained visually (see Appendix [A - page 13](#)).

Railways without computer systems should capture data manually at the time of its production:

- directly on the spot (during train formation, checking of consist, etc.) or
- after transmission of a verbal radio message,

and record it in documents (train documents, local station documents, etc.).

For Railways with computer systems these documents should be adapted to the requirements of computer processing.

Data should be inputted

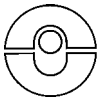
- when produced, or at the latest at the end of the working period, in the case of distributed computer terminals or processing systems, and/or
- directly:
 - when produced
 - or after the period concerned

in the case of a centralised data processing system.

If a terminal or distributed processing system is available, this presents an opportunity for data received by radio or produced locally to be keyed-in directly.

To reduce the data-capture workload, it is advisable to increase the percentage of data captured automatically (e.g. train number, length in metres, number of wagons).

Direct spoken input into the computer may be possible in the future.



2 - Processing of source data

2.1 - General

Source data captured should be

- used as a basis for operational work at station level and converted into local statistics, and
- concentrated manually or by computer at the level of intermediate management and/or aggregated on a selective basis and used to calculate reference values.

Daily values should be aggregated to make monthly values (where appropriate also weekly or ten-daily values, etc.).

Monthly values should be aggregated to make annual values, and where appropriate quarterly, half-yearly values, etc.).

Reference values are values formed by linking one item of primary information with another (e.g. train-kilometres, net tonne-kilometres) and/or by relating an item of primary information to a unit of time (e.g. number of trains per day, number of disruptions per month).

Reference values are extremely simple to calculate (basic calculations, calculation of percentages and averages) and may be worked out manually, though because of the large number of calculations required, it is preferable to use a computer so as to cut down processing work.

2.2 - Manual processing

Data in specific documents must be recorded totally or in part in new documents at local level or by regional departments, where appropriate after the following values have been calculated:

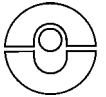
- reference values,
- absolute deviation of a value from a threshold or forecast value,
- averages.

2.3 - Computer processing

2.3.1 - General

Basic input data should be transferred unmodified by non-intelligent terminals to the mainframe computer.

Intelligent terminals or distributed computer systems should preprocess the source data (e.g. determine totals per shift, etc.) to reduce the amount of transmission work. A plausibility check should be carried out at the same time. Final processing of the data transmitted and/or directly inputted should take place in the mainframe computer.



2.3.2 - Daily processing

Values per shift, or at least per day, should also be aggregated over the month. Weekly or ten-daily aggregates may also be made if required.

To increase the significance of data, processing may be accompanied by comparisons with

- the reference threshold or expected value;
- the value for the previous day, week or ten-day period;
- the corresponding aggregates for the previous month (see Appendix C - page 17).

2.3.3 - Monthly and annual processing

Monthly processing should be carried out:

- as indicated in point 2.3.2 - page 5 for monthly data entered directly, and/or,
- by taking monthly aggregates of data processed daily.

Annual processing should be effected using monthly values with intermediate quarterly, half-yearly and nine-monthly aggregates, together with the sum of the annual values in the form of a time series. The latter must be kept continually up-to-date.

2.4 - Evaluation of reference values

2.4.1 - General

Reference values should be determined using the source data captured for:

- selected times and/or
- specific typical periods (calendar or timetable units)
- selected types of source data.

2.4.2 - Simple reference values

Simple reference values (see Appendix D - page 18) are obtained by aggregating source data for the selected calendar unit of time such as:

- number of wagons/axles of all freight trains;
- length of trains in metres;
- number of seats in all express trains;
- wagon loads;

- number of disruptions of a technical nature;
- aggregate delays at terminus for all fast freight trains on international lines (TEEM trains, etc.).

2.4.3 - Compound reference values

Compound reference values (see Appendix E - page 21) should be determined over the period selected, for example:

- multiplication of values per train and addition of these values for all trains e.g.

wagon axle-kilometres:

$$\sum_{i=1}^n (\text{number of axles} \times \text{kilometres run})_{\text{train } i}$$

where i = train 1 n,

- calculation of averages, for example:

average duration of a given type of disruption

$$\frac{\text{duration of a given type of disruption}}{\text{number of disruptions of that type}}$$

or

- calculation of percentages, for example:

degree of punctuality

$$\frac{\text{number of trains that arrived on time} \times 100\%}{\text{trains that ran}}$$

2.5 - Presentation of results

2.5.1 - General

Whatever the type of processing (manual or computer), results should be presented in the form of:

- tables,
- and graphs.

2.5.2 - Presentation in tables

It is advisable to use as

- column headings
 - the results obtained during the reference period and/or the aggregate results
 - comparisons with similar periods and/or expected values or thresholds
- and as

- side headings: the reference values

for tables compiled manually or by computer (see Appendix [C - page 17](#)).

2.5.3 - Presentation in graphs

Graphs serve to show and compare results. Visual comparison is not easy to achieve unless graphs have a standard structure and are of the same type.

- Histograms (see Appendix [F - page 23](#)) or
- frequency polygons (see Appendix [G - page 24](#))

should be used for relative and absolute values, to represent results for

- a given period, or
- as a time series (statistical series classified in terms of days, weeks, ten-day periods, months, quarters, etc.). As a general rule, linear scales should be used. A logarithmic scale may be used to present aggregates.

The reference value should be written along the Y-axis, and periods should be indicated along the X-axis.

To present relative results as percentages, a circular diagram with segments should be used (see Appendix [H - page 25](#)). The area of the circle equals 100%.

2.5.4 - VDU and microfiche presentation

Results processed by computer should as far as possible be presented in the form of tables and/or graphs, using

- visual display units for immediate operational information (short periods), and
- microfiches, for information collected over long periods.

Various methods may be used for storing data, i.e. microfiche, magnetic tape, magnetic discs, etc.

3 - Use of results

3.1 - General

The processing of results as required:

- possible forecasts of expected results, and
- relations between partial results

are set out below, so that reference values may be fully exploited.

Mathematical and statistical methods should be used for processing results and calculating forecast values. Likely short, medium and long-term repercussions on operating efficiency should be presented in a general manner. Decisions made in respect of specific situations must always take the individual interests of each Railway into account.

3.2 - Analysis of time series

3.2.1 - General

Reference values classified chronologically according to specific periods (days, weeks, ten-day periods, months, quarters, years, etc.), form a time series. Such series are determined on the basis of four random values:

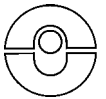
- the basic direction of the trend;
- periodic variations (seasonal influences);
- cyclic variations;
- irregular variations

which do not necessarily all occur simultaneously in a time series.

Analysis of a time series reveals values specific to the railway trends and seasonal influences, without indicating links between them.

The trend determined shows the general development excluding seasonal variations, and the seasonal variation shows recurring peaks and troughs. When results are integrated, factors such as the differing number of days in a month, bank holidays, different number of weekends etc. must be taken into account. Results obtained by analysing a time series serve mainly as a basis for determining medium and long-term measures for improvement of the operating process.

Railways with the facilities for computerised determination of trends and/or seasonal influences in time series are advised to make use of these. A large number of computer programs are available for the purpose.



3.2.2 - Determination of trends

To determine trends on a graph, results should be entered in relation to the time values using a system of coordinates, and the curve plotted freehand between the points in such a way that the areas above and below the trend line linking the empirical points are more or less equal.

Manual determination of the trend by the moving-averages method is simple and not time-consuming. The degree of smoothing of the time series is determined by the number of periods chosen and used to calculate the average.

Determination of the trend by orthogonal functions requires little effort since most of the calculation operations are given in tables. Furthermore, general trend coefficients no longer differ when variations occur in the degree of trend.

Examples of application of the last two methods with calculations are given in Appendix I - page 26.

3.2.3 - Determination of seasonal influences

Seasonal influence in a time series is a reflection of typical and regular variations (daily, weekly, monthly, according to a time-table section during the year, yearly) in rail traffic, for example variations in respect of:

- occupancy rates of express and commuter trains as a consequence of increases in holiday and leisure travel and of the drop in commuter traffic during summer months;
- or
- marshalling-yard activities, where there is an increase during the last month of the quarter and the last quarter of the year.

It is advisable to use the simple phase-average method for manual determination of seasonal influence in stationary time series (i.e. where there is little or no trend influence). Here, the phase represents the unit of time (e.g. month) within the period under consideration (e.g. year).

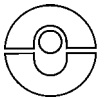
The chain-index method is recommended for manual determination of seasonal influences in time series with a strong seasonal influence. Although this method does not totally eliminate the influence of the trend, it is recommended nonetheless, since it can be applied using a simple pocket calculator.

Appendix I provides examples of corresponding calculations for the two methods.

3.3 - Forecasting of expected values

The time-series forecast enables values for a forthcoming period to be predicted:

- on the basis of the pattern shown by the time series in the past;
- on the basis of the development which the time series is expected to show (parallel, linear or non-linear development with or without trend influence).



It is very useful to have a set of forecasts available covering one or several months. Specific events occurring during the months covered by the forecast, and which strongly affect the forecast reference value, may of course produce variations.

It is advisable to apply the method of exponential smoothing for determining expected values. Appendix I, point 1.3 - page 35 gives an example of the calculation method used.

Initial values for updating average and seasonal coefficients together with the forecasts themselves, can be calculated simply and quickly. Calculation of the most appropriate smoothing constants requires time, but using a pocket or desk calculator helps considerably. Several different forecasts for events occurring in sequence can be made at one and the same time giving a reasonably accurate result.

As far as computerised forecasting of time series is concerned there exists a whole range of application-oriented methods.

Given the practical usefulness of time series in a railway context, it is advisable to apply iterative methods. The programme should be implemented in three stages

- determination of the initial values of the time series;
- choice of the method to be applied and the corresponding parameters, by means of variation on either side of the minimum error deviation.
- forecasting, by means of the method chosen , where applicable, in accordance with specified periods, with monitoring of the second stage of implementation.

3.4 - Relationships between reference values

Reference values are interdependent because of the way in which the various facets of the technical railway-operating process are themselves connected.

E.g. the ratio of the compound reference values

$$\frac{\text{net tonne-km (ntkm)}}{\text{gross tonne-km (gtkm)}}$$

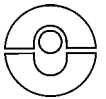
shows the volume of work (wagon tare and empty runs) required to convey the amount of freight to be forwarded. The value of the ratio indicates the efficiency of the transport process and the productive use made of wagon capacity.

In accordance with the rule for mathematical determination:

$$\frac{\text{weight of freight} \times \text{kilometres run}}{(\text{weight of wagon} + \text{weight of freight}) \times \text{kilometres run}}$$

the aim is to arrive at a relative reduction in performance expressed in gross tonne-kilometres, by

- using wagons with a $\frac{\text{payload}}{\text{tare}}$ ratio as high as possible (high capacity wagons, reduction in wagon tare);



- eliminating unnecessary empty runs, e.g. by making empty wagons from a specific unloading station or area available at that same place for reloading;
- increasing the degree to which the capacity of loaded wagons is used.

Apart from these relationships between the reference value "gross tonne-kilometres" and the reference values "net tonne-kilometres", "wagon axle-kilometres" (divided into loaded and empty) and use of wagon capacity, other relationships also exist arising from technical considerations, linked with the reference values, number of kilometres run by the tractive unit, number of kilometres run by the train, wagons per train, occupancy rate of marshalling yards, occupancy rate of lines, and others.

Moreover, the efficiency of the transport process may be further increased by measures which result in the real number of kilometres run (the multiplier in both numerator and denominator of the ntkm/gtkm ratio) being made to approach more closely the actual distance, for instance by

- optimising regulations within the computer system concerning forwarding lines of route, and
- with regard to practical implementation of the transport process, observing efficiently the regulations concerning forwarding lines of route (i.e. avoidance of detours, etc.).

3.5 - Influencing of reference values

Reference values for the transport process depend on external factors some of which are impossible to influence, on the quality of the technology used, and on the use itself to which it is put.

For example, the share represented by light running locomotives in the reference values "duration of tractive unit utilisation" and "number of kilometres run by the tractive unit" depends on whether:

- tractive unit turnaround has been optimised;
- the theoretical transport process has been implemented in practice.

Certain reference values for the transport process, e.g. use of wagon capacity, are influenced by both the railway and the customer.

The reference values:

- **Disruptions**

are influenced by technical equipment, the work of railway staff and external factors. They are caused by deviations from the planned service.

Influence is exerted on disruptions to

- avoid them by taking the necessary steps at acceptable financial outlay;
- minimize their repercussions.

- **Delays**

especially if they are the consequence of disruptions, constitute deviations from the planned service and may give rise to further deviations. Influence is only exerted on delays above a certain threshold, fixed by the Railways, and which is lower for example for passenger than for freight trains.