

## RESULTS OF THE 2005 E-RAIL TRAFFIC CENSUS

### Presentation of the results of the 2005 E-Rail Traffic Census

Transmitted by Eurostat

### **Eurostat's rail transport statistics – Annex G of Regulation 91/2003**

#### **I. INTRODUCTION**

Eurostat's rail transport statistics are predominantly<sup>1</sup> based on **Regulation (EC) No 91/2003** of the European Parliament and of the Council of 16 December 2002 which repealed Directive 80/1177/EC. Regulation (EC) No 91/2003 covers the transport of goods and passengers by rail, as well as the collection of statistics on rail safety. It takes account of the new structures in the liberalised rail transport market by simplifying the collection of data on freight transport and by including statistics on passenger transport and rail safety.

Some transmission of data under the new Regulation began in 2003. However, starting from the collection of rail freight statistics only in 2003, the number of datasets grew between 2003 and 2007. By 2007, all datasets had to be delivered to Eurostat, including the first collection of five-yearly statistics for Annexes F and G. This development was mainly due to the fact that the provisions of the Annexes were introduced gradually, as follows:

- ❖ 2003 reference year – Tables A1, A2, A3 (first set of freight transport data covering large undertakings only) and Annex I (list of undertakings) were required for the first time;
- ❖ 2004 reference year – Tables A4, A5, A6, A7, A8 and A9 (second set of freight transport data covering large undertakings only), Annex B (freight transport data covering small undertakings), Annex C (passenger transport data covering large undertakings), Annex D (passenger transport data covering small undertakings), Annex E (quarterly data covering large undertakings only) and Annex H (rail accident data covering all undertakings) were required for the first time;
- ❖ **2005 reference year** – Annex F (regional transport data at NUTS 2 level) and Annex G (statistics on traffic flows on the railway network) were required for the first time.

---

<sup>1</sup> Some additional information (mostly on equipment, network and employment) is available from the Common Questionnaire. Very few variables (length of railway network at regional level) are available from the regional questionnaire (collected via the REGWeb application).

## II. STRUCTURE OF ANNEX G DATA, DATA PROVISIONS AND TIMELINESS

The general structure of the data on traffic on the rail network is defined in Annex G to the Regulation (appendix 1), supported by the methodological Guidelines (appendix 2). Moreover, Eurostat (assisted by Denmark)<sup>2</sup> made the examples of Annex G data, plus a technical description and geographical coordinates, available to all countries (via the common Communication & Information Resource Centre Administrator, Circa) for reference purposes.

The data are contained in three tables setting out the number of freight, passenger and other trains (optional) which have to be provided. At least the TEN transport segments should be covered. Moreover, the tables should be supplemented by the list of the network segments, together with their technical characteristics and geographical coordinates.

The first collection of data on traffic flows on the rail network covered the 2005 reference year. The deadline for transmission of the data is 18 months after the end of the reference period, namely end of June 2007.

According to the *Recommendations to Governments on the E-Rail Traffic Census in Europe in 2005 jointly undertaken by the United Nations Economic Commission for Europe and Eurostat (2005 E-Rail Census)*,<sup>3</sup> the Member States of the European Union were supposed to submit Annex G data to Eurostat, supplemented by the technical description and geographical coordinates of their network segments.<sup>4</sup> Eurostat would then pass on these data to UNECE.

At the Rail Working Group in Eurostat in 2006 (where the details of Annex G data provision were discussed), some countries already gave prior warning of possible problems with supplying Annex G data to Eurostat in accordance with the deadlines specified in the Regulation.

The first data sets were provided to Eurostat in May/June 2007. Many countries which had provided the data by the deadline were obliged to revise the data or provide missing technical descriptions and geographical coordinates.<sup>5</sup> Some countries provided the data when they were reminded to do so by Eurostat. Very few took longer than 6 months to provide data or failed to provide the data.

Despite the supporting information (examples of the data, technical description and geographical coordinates) that were available well in advance, the structure and the content of the data sets were not fully satisfactory in many cases. Some countries were asked to make revisions and/or supply missing information.

Eurostat started transmitting available Annex G data to UNECE at the beginning of July 2007. At that time, Annex G data (as well as revisions or/and supporting information) were provided to UNECE as soon as they had been delivered by a country. All the data provided by Eurostat to UNECE were raw data, as Eurostat has not received all the data in the format

---

<sup>2</sup> Denmark provided Eurostat with examples of Annex G data sets as well as with a template for geographical coordinates and technical characteristics of railway segments.

<sup>3</sup> <<http://www.unece.org/trans/doc/2003/wp6/TRANS-WP6-AC4-02a1e.doc>>

<sup>4</sup> EFTA countries – Norway and Liechtenstein are subject to the same data provisions as the EU countries, under separate agreements between the EU and EFTA. Switzerland started providing rail transport data from the 2008 reference year onwards.

<sup>5</sup> Technical descriptions of segments as well as their geographical coordinates are an integral part of Annex G data; no Annex G data can be disseminated or validated without this information.

required and as the complete validation procedures have not yet been carried out. Detailed information on Annex G data provisions by country are presented in Table 1.

### **III. PROCESSING AND DISSEMINATION OF ANNEX G DATA BY EUROSTAT**

Eurostat started preparatory work on the processing of Annex G data in 2007. The work continued in 2008. The following actions have been undertaken:

1. initial brief assessment of the data supplied in order to detect missing information and most common errors in data encoding;
2. adaptation of the rail production database to incorporate Annex G data;
3. preparation of the dissemination tables and procedures for extraction from the database
4. preparation of a network segments dictionary based on the list of segments provided by the countries.

The following additional circumstances slowed down the validation and dissemination process:

1. Delays in Annex G data provision by the countries in turn delayed the creation of the railway network segments dictionary;
2. It was necessary to revise and clarify data and to supply missing information for most of the countries;
3. The length of the segment code had to be changed (from 7 to 8 characters, as the number of segments selected by some of the countries exceeded 1 000); Eurostat had to introduce an additional procedure for re-codification of all rail segments provided;
4. The specifications of two new dissemination tables were prepared in 2008 using the process for the dissemination of rail transport statistics;
5. The preparation and development of the extraction procedures followed the preparation time of the above specifications and were implemented under an IT contract in 2008;
6. The new validation rules could not be set up until after the structure of the database tables had been finalised.

Because of the above issues, Eurostat had to reschedule the process of disseminating Annex G data from mid-2008 to the beginning of 2009. The final Annex G validation and dissemination processes will basically depend on the quality of the data already provided, cooperation with the countries in the event of revisions of the data, and the quality of the IT programme used for implementing the new tables and validation rules.

**Table 1.** Annex G data delivery status (as of 08/2008)

Country	Annex G tables <sup>1)</sup>	Date of data provisions to Eurostat <sup>2)</sup>	Supportive data provisions <sup>3)</sup>	Remarks
BE	G1, G2	28/06/2008 (revisions provided 10/2008)	GEO provided TECH provided with delay (10/2008)	
BG	Not applicable (being a Candidate Country in 2005, BG was not subject to Annex G data provisions)			
CZ	G1, G2, G3	19/06/2007	GEO and TECH provided	Additional maps provided
DK	G1, G2	07/11/2007	GEO and TECH provided	Examples of Annex G data provided in advance to be used by other countries as a reference
DE	G1, G2, G3	28/06/2007	GEO – provided TECH – missing	
EE	G1, G2	22/06/2007	GEO and TECH provided with delay (10/2008)	
IE	G1, G2	20/03/2007	GEO and TECH provided	
EL	G1, G2	22/06/2007	GEO and TECH provided	
ES <sup>4)</sup>	G1, G2, G3	13/05/2007 (revisions still needed)	GEO and TECH missing	Annex G data incomplete; direction is missing;
FR	G1, G2	04/09/2007	GEO and TECH provided	
IT	G1, G2	18/07/2007	GEO and TECH provided but with delay (07/2008)	
LV	G1, G2	30/07/2007 and 07/08/2007 (revisions:15/08/2007)	GEO and TECH provided	
LT	G1, G2	05/07/2007	GEO – missing TECH – provided but with delay (01/2008)	
LU	G1, G2, G3	05/09/2007 (revisions 14/03/2008)	GEO and TECH provided with delay (03/2008)	

Country	Annex G tables <sup>1)</sup>	Date of data provisions to Eurostat <sup>2)</sup>	Supportive data provisions <sup>3)</sup>	Remarks
HU	G1, G2	03/09/2007	GEO and TECH provided with delay (02/2008)	
NL <sup>4)</sup>	G1, G2, G3	29/04/2008	GEO and TECH missing	
AT	G1, G2, G3	02/06/2007	GEO and TECH provided	
PL	G1, G2	02/05/2007	GEO and TECH provided	
PT <sup>4)</sup>	G1, G2, G3	04/07/2008 (revisions still needed)	TECH missing; GEO provided	
RO	Not applicable (being a Candidate Country in 2005, RO was not subject to Annex G data provisions)			
SI	G1, G2	14/06/2007	GEO and TECH provided	
SK	G1 G2 G3	22/06/2007	GEO and TECH provided	Maps provided
FI	G1, G2	09/07/2007	GEO and TECH provided	
SE	G1, G2, G3	27/08/2007 01/2008 revisions	GEO and TECH provided with delay (09/2007)	
UK <sup>4)</sup>	G1, G2	Data not provided yet	Not provided	
LI	G1, G2, G3	02/08/2007	GEO and TECH provided	Data provided by AT on behalf of LI
NO	G1, G2	29/06/2007	GEO and TECH provided	
HR	Not applicable (Candidate Countries are not subject to Annex G data provisions)			
TR	G1, G2, G3	07/07/2006	GEO and TECH provided, some information missing	TR provided Annex G data on voluntary basis; Eurostat hasn't received TR agreement on data to be provided to UNECE yet

\* **G1** – goods transport by network segment; **G2** – passenger transport by network segment; **G3** – other (service trains, etc.) optional dataset;

\*\* shaded area – country delayed with data provisions by at least a month;

\*\*\* **GEO** – supportive information on geographical co-ordinates of the segments;

\*\*\* **TECH** – supportive information on technical characteristics of the segments;

\*\*\*\* country notified with compliance assessment letter from the Director-General of Eurostat

#### **IV. CONCLUSIONS AND LESSONS LEARNED FROM THE 2005 ANNEX G DATA COLLECTION**

##### **A. Countries' perspective:**

Most of the countries to which Annex G data provisions applied were late with such provisions or at least were late in providing additional information. These delays can mainly be explained by the following circumstances:

– 2005 was the first reference year of the Annex G data collection – the data were collected and provided for the first time pursuant to the new Regulation 91/2003 and covered all undertakings; National Statistical Institutes should have been better prepared for setting up a system to collect and validate these data first;

– Traffic on railway network segments – sources of the data – railway undertakings proved to be an inadequate source of data (especially where the segments were not properly defined in advance and supplied to the undertakings); infrastructure managers, being responsible for railway traffic on the entire network, proved to be a better reference (for Annex G data) than railway undertakings; this makes the collection of Annex G data more complicated for countries to organise (as the other annexes to Regulation 91/2003 regard railway undertakings as a primary source of data);

– Procedure for selecting railway network segments – the survey was supposed to cover at least the TEN-T network; the selection of the network segments was left to the individual countries. Some of them were unable to establish their segments properly, or the segments selected failed to meet the requirements of the survey (e.g. to be homogeneous in terms of technical description and traffic performance).

##### **B. Eurostat's perspective:**

– As a result, delays in providing Annex G data delayed the start of Eurostat's work;

– Preparations for the validation of Annex G data were reliant on data delivery and took longer than expected, as a set of new validation rules had to be drawn up to cater for the most common errors detected;

– Preparation of the production of rail database tables, dissemination tables and IT implementation of the new validation rules were dependent on data delivery and the terms of external contracts;

– Unforeseen circumstances, such as a change in the length of the segment code (which had to be agreed beforehand with the countries, and subsequently the changes to the segment codes had to be incorporated in the Network Segments Dictionary and into the data already provided by Member States);

– Because of the delays and the lack of additional validation rules, Eurostat forwarded all raw data received direct to UNECE, in order not to delay UNECE's work;

– Eurostat has received the revisions of Annex G data and supporting information on a continuous basis; it is expecting additional revisions in the near future when the full validation process is launched – at the beginning of 2009.

### **2010 Annex G data collection – outlook**

Based on the assumption that the structure and the data coverage of Annex G data will not change:

- A greater number of countries are more likely to meet the deadline for providing 2010 Annex G data (end of June 2012) because of their experience of the 2005 data collection; the procedure for segment selection has been worked out; in most cases the list of segments will need to be updated with new segments only; the data sources are better recognized after first data collection; the structure of the data is likely to be more consistent between the countries;
- Validation of Annex G data will proceed much more quickly because of consolidated validation rules at Eurostat;
- Dissemination of Annex G data will be based on the same dissemination table structure and, as a result, the data should be disseminated no later than ~1 year after being received.

## APPENDIX 1. Annex G of Regulation 91/2003

L 14/10

EN

Official Journal of the European Communities

21.1.2003

### ANNEX G

#### STATISTICS ON TRAFFIC FLOWS ON THE RAIL NETWORK

List of variables and units of measurement	Goods transport: — number of trains Passenger transport: — number of trains Other (service trains, etc.) (optional): — number of trains
Reference period	One year
Frequency	Every five years
List of tables with the breakdown for each table	Table G1: goods transport, by network segment Table G2: passenger transport, by network segment Table G3: other (service trains, etc.), by network segment (optional)
Deadline for transmission of data	18 months after end of reference period
First reference period	2005
Notes	<ol style="list-style-type: none"><li>1. Member States shall define a set of network segments to include at least the rail TEN on their national territory. They shall communicate to Eurostat:<ul style="list-style-type: none"><li>— the geographical coordinates and other data needed to identify and map each network segment as well as the links between segments,</li><li>— information on the characteristics (including the capacity) of the trains using each network segment.</li></ul></li><li>2. Each network segment which is part of the rail trans-European network (TEN) shall be identified by means of an additional attribute in the data record, in order to enable traffic on the rail TEN to be quantified.</li></ol>

## Appendix 2. An extract of Annex G data from the Guidelines for the implementation of Regulation 91/2003 (draft v. 5.1)

### Chapter 8: Description of Annex G

#### 1.0 Overview of Annex G

- **Coverage of rail transport statistics**

Annex G of Regulation 91/2003 covers statistics on traffic flows on the rail network. Data are broken-down into three tables:

Table G1: goods transport, by network segment

Table G2: passenger transport, by network segment

Table G3: other (service trains, etc.), by network segment (optional)

Figures have to be provided in one dataset (i.e. in one file) respecting the structure and rules described below.

- **Dataset for Annex G: Statistics on traffic flows on the rail network**

Field number	Variable	Description	Coding	Table
1	RCount	Reporting country	ISO-3166-alpha2 except "UK" for United-Kingdom	G1 to G3
2	DsetID	Table identification	G1 to G3	G1 to G3
3	Year	Year of the data set	On 4 positions	G1 to G3
4	Period	Reference period	On 2 positions (A0)	G1 to G3
5	NetSeg	Network segment	On 7 positions ISO-3166-alpha2 except "UK" for United Kingdom + 'S' + <i>indication of direction (1 or 2)</i> + number on 3 positions	G1 to G3
6	FlagTen	Rail TEN segment	0: NO 1: YES	G1 to G3
7	Train	Number of trains	Numeric	G1 to G3

- **Frequency – every five years**

- **Examples of records**

ES;G1;2005;A0;ESS1001;1;272800

ES;G2;2005;A0;ESS2012;0;1625000

ES;G3;2005;A0;ESS1013;1;5600500

## Definitions

### Network Segment

A set of network segments shall be defined by the Member State and included at least the rail TEN on their national territory. It shall be defined as follows:

- the geographical coordinates and other data needed to identify and map each network segment as well as the links between segments,
- information on the characteristics (including the capacity) of the trains using each network segment.

Each network segment which is part of the rail Trans-European network (TEN) shall be identified by means of an additional attribute in the data record, in order to enable traffic on the rail TEN to be quantified.

**Geographical coordinates** means spherical co-ordinates measured in latitude and longitude. If the earth is considered to be a sphere, latitude and longitude are angles measured from the earth's centre to a point on the earth surface. Latitude and longitude are measured in degrees, minutes and seconds. The equator has latitude 0°, the North Pole 90°, and the South Pole -90°. The Prime Meridian, indicating a longitude of 0°, starts at the North Pole, passes through Greenwich, England and ends at the South Pole (Working Group document RAIL/2006/10, article 1-26).

### Rail transport European Network

**TEN (Trans-European Network) railway line** comprise of the high-speed rail lines and conventional rail lines as defined in the Decision No 1692/96/EC on the Community guidelines for the development of Trans-European transport network (Article 10) and its amendments;

### Train (Number of)

**Train** means one or more railway vehicles hauled by one or more locomotives or railcars, or one railcar travelling alone, running under a given number or specific designation from an initial fixed point to a terminal fixed point. A light engine, i.e. a locomotive travelling on its own, is not considered to be a train, except for provision of data for Annex H.

## 1.1 Description of table G1

- **Coverage of Rail transport statistics in table G1**

Table G1 covers goods transport, by network segment

- **Structure of table G1 – goods transport, by network segment**

Field number	Variable	Description	Coding
1	RCount	Reporting country	ISO-3166-alpha2 except "UK" for United-Kingdom
2	DsetID	Table identification	G1
3	Year	Year of the data set	On 4 positions
4	Period	Reference period	On 2 positions (A0)
5	NetSeg	Network segment	On 7 positions ISO-3166-alpha2 except "UK" for United Kingdom + 'S' + <i>indication of direction (1 or 2)</i> + number on 3 positions
6	FlagTen	Rail TEN segment	0: NO 1 : YES
7	Train	Number of trains	Numeric

- **Example of record :**

ES;G1;2005;A0;ESS1001;1;272800

- **Deadline for the transmission of data – 18 months after the end of the reference period**

- **Validation of data**

N°	Description of the test	Type of test	Type of error*
1	RCount must be present and belong to the ISO-3166-alpha 2 nomenclature (except for « UK »)	Existence of a field and check of value	E
2	DsetID must be present and equal to G1	Existence of a field and check of value	E
3	Year must be present and belong to the interval [2005, every five years]	Existence of a field and check of value	E
4	Period must be present and equal to [A0]	Existence of a field and check of value	E
5	NetSeg must be present	Existence of a field	E
6	FlagTen must be present and equal to 0 or 1	Existence of a field and check of value	E
7	Number of trains must be present	Existence of a field	E
8	For each segment, both directions need to be reported	Internal consistency checks	W
9	The number of trains (declared in annex G table G1) multiplied by the length of the network segment may not to be higher than the train-km (declared in annex A table 9)	External consistency checks	W

\*E: Error – W: Warning

## 1.2 Description of table G2

- **Coverage of Rail transport statistics in table G2**

Table G2 covers passengers transport, by network segment

- **Structure of table G2 – passengers transport, by network segment**

Field number	Variable	Description	Coding
1	RCount	Reporting country	ISO-3166-alpha2 except "UK" for United-Kingdom
2	DsetID	Table identification	G2
3	Year	Year of the data set	On 4 positions
4	Period	Reference period	On 2 positions (A0)
5	NetSeg	Network segment	On 7 positions ISO-3166-alpha2 except "UK" for United Kingdom + 'S' + indication of direction (1 or 2) + number on 3 positions
6	FlagTen	Rail TEN segment	0: NO 1 : YES
7	Train	Number of trains	Numeric

- **Example of record :**

ES;G2;2005;A0;ESS2012;0;1625000

- **Deadline for the transmission of data – 18 months after the end of the reference period**

- **Validation of data**

N°	Description of the test	Type of test	Type of error*
1	RCount must be present and belong to the ISO-3166-alpha 2 nomenclature (except for « UK »)	Existence of a field and check of value	E
2	DsetID must be present and equal to G2	Existence of a field and check of value	E
3	Year must be present and belong to the interval [2005, every five years]	Existence of a field and check of value	E
4	Period must be present and equal to [A0]	Existence of a field and check of value	E
5	NetSeg must be present	Existence of a field	E
6	FlagTen must be present and equal to 0 or 1	Existence of a field and check of value	E
7	Number of trains must be present	Existence of a field	E
8	For each segment, both directions need to be reported	Internal consistency checks	W
9	The number of trains (declared in annex G table G2) multiplied by the length of the network has not to be higher than the train-km (declared in annex C table C5).	External consistency checks	W

\*E: Error – W: Warning

### 1.3 Description of table G3

- **Coverage of rail transport statistics in Table G3**

Table G3 covers other traffic (service trains, etc.), by network segment (optional)

- **Structure of Table G3 – goods transport, by network segment**

Field number	Variable	Description	Coding
1	RCount	Reporting country	ISO-3166-alpha2 except “UK” for United-Kingdom
2	DsetID	Table identification	G3
3	Year	Year of the data set	On 4 positions
4	Period	Reference period	On 2 positions (A0)
5	NetSeg	Network segment	On 7 positions ISO-3166-alpha2 except “UK” for United Kingdom + ‘S’ + indication of direction (1 or 2) + number on 3 positions
6	FlagTen	Rail TEN segment	0: NO 1 : YES
7	Train	Number of trains	Numeric (optional)

- **Example of record :**

ES;G3;2005;A0;ESS1013;1;5600500

- **Deadline for the transmission of data – 18 months after the end of the reference period**

• **Validation of data**

N°	Description of the test	Type of test	Type of error*
1	RCount must be present and belong to the ISO-3166-alpha 2 nomenclature (except for « UK »)	Existence of a field and check of value	E
2	DsetID must be present and equal to G3	Existence of a field and check of value	E
3	Year must be present and belong to the interval [2005, every five years]	Existence of a field and check of value	E
4	Period must be present and equal to [A0]	Existence of a field and check of value	E
5	NetSeg must be present	Existence of a field	E
6	FlagTen must be present and equal to 0 or 1	Existence of a field and check of value	E
7	Number of trains must be present	Existence of a field	E
8	For each segment, both directions need to be reported	Internal consistency checks	W

\*E: Error – W: Warning

-----