Applying Eurostat’s ESS handbook for quality reports on Railway Maintenance Data

C. Stenström  
*Operation and Maintenance Engineering, Luleå University of Technology, Sweden*

P. Söderholm  
*Quality Technology and Management, Luleå University of Technology, Sweden*

**ABSTRACT:** The importance of data quality has become more evident with the digitalization trend and development of new asset management frameworks. Digitalization has changed maintenance work by an increasing share of condition monitoring and digitalized work order processes, which for rail infrastructure and rolling stock give rise to data sets qualifying as big data. Asset management in turn, has progressed significantly the last decades as a response to digitalization, as well as due to a changing organisational culture. ISO 55000, perhaps the best known asset management guidelines, has been adapted to railways by UIC (International Union of Railways), and the EU-projects In2Rail and In2Smart. However, the quality of the data collected has become a growing concern that has not been adequately addressed in asset management. In this study, Eurostat’s ESS (European Statistical System) handbook for quality reports has been adapted and applied to railway maintenance data. The results include a case study on data quality reporting and performance indicator specification. Practical implications are believed to be that the study will support a more structured process towards data quality management, which in turn can aid decision-making, for example by more accurate cost-benefit analysis of preventive maintenance.

Keywords: data quality, quality reporting, quality assurance framework, maintenance, asset management, European Statistical System (ESS), Eurostat, railway

1 INTRODUCTION

The European Statistical System (ESS) is the partnership between Eurostat (European Commission) and the national statistical institutes and other national authorities responsible in each member state for the development, production and dissemination of European statistics; Regulation No 223/2009 on European statistics (EC, 2009). Besides regulation No 223/2009, central documents of European statistics include ‘The European Statistics Code of Practice’, or CoP, (Eurostat, 2011), the ‘Quality Assurance Framework of the European Statistical System’, or QAF (ESSC, 2015) and the ‘ESS handbook for quality reports’ (Eurostat, 2014). The CoP provides 15 principles of European Statistics, the QAF provides recommendation for implementation of the CoP, and the ESS handbook provides guidelines for data quality reporting.

The importance of assuring quality in data collecting, storing and processing has become more evident with the increased digitalization and recent developments in asset management frameworks. Driving factors behind the digitalization trend include: cheaper, miniaturized and more efficient electronics; improved sensor technology due to microelectromechanical systems (MEMS); and more efficient mobile network standards. Some important factors behind the trend in asset management, includes, of course, the development in digitalization, but also a growing focus on quantitative indicators for the follow-up of business objectives, as well as outsourcing of support functions outside the core business. A well known series of standards in the subject is the ISO 55000 (Asset Management, ISO 2014), standards.

Follow up of reliability, from a system to component level, enables effective decision making regarding maintenance investments. Nevertheless, proper decisions require trustworthy data, i.e. the quality of the data needs to be known, which often is not the case in practice. Therefore, strategic planning needs a sub-process for assessing the data quality for following up set objectives.

Data quality frameworks of varying maturity and focus are available in the literature. The motivation for adapting Eurostat’s ESS handbook for quality reports (2014) to infrastructure managers includes:
• It is used by the member states of the European Union and their national statistical institutes, and thus, it is well developed and well known.

• It is in line with international and national standards and regulations, e.g., see EC (2009).

• The underlying QAF is accepted as a tool for quality assurance for developing composite/world indicators (OECD & JRC–EC 2008).

• Rail infrastructure managers are normally authorities with responsibility for official statistics.

OECD & JRC–EC (2008) notes that Eurostat’s quality framework is quite similar to the International Monetary Fund’s (IMF) ‘Data Quality Framework’, in the sense that both frameworks provide a comprehensive approach to quality.

In this study, Eurostat’s ESS handbook for quality reports (2014) has been adapted and applied to railway maintenance data. The aim is to improve data collection, processing and indicators for follow up of objectives and for decision support.

Reviews on developed generic data quality frameworks are available in Knight & Burn (2005) and Haug et al. (2011). Haug et al. (2011) also reviewed the impact, or cost, of poor data quality and noted that there is a lack of case studies in academic journal articles. Reviews, as well as tools for data quality, are also found in books of Batini & Scannapieco (2006) and Sadiq (2013).

One of the first and well-known frameworks are by Wang & Strong (1996). Wang & Strong (1996) performed a survey among data users, or consumers, and developed a data quality framework. The originality of the study is that the framework was developed based on data users and not only based on the researchers view. The developed framework includes 15 data quality attributes, or dimensions, categorized into four groups: intrinsic, contextual, representational and accessibility. In a later study, the research group developed a method and metrics to assess how well organizations develop information products and deliver information services to consumers (Kahn et al., 2002; Pipino et al., 2002).

A number of adaptations of data quality frameworks to maintenance exists, but are often conference contributions with limited details or contribution. One exception is Madhikermi et al. (2016) that applied a framework of Krogstie et al. (1995) to an equipment manufacturer.

Some contributions on data quality research in railways are available. Fu & Easton (2018) discussed why data quality matters in a rail context, with the aim to do future studies. King & Crowley-Sweet (2014) discussed ISO 8000 (Data quality), 9000 (Quality management) and 55000 (Asset management), and shared some experiences from implementation at Network Rail. By deploying ISO 8000, Network Rail has achieved several benefits. In particular, the organisation has been able to address concerns from the Office of the Rail Regulator with respect to data quality within Network Rail.

Kudla & Majumdar (2013) developed a data quality model with six attributes in two categories for railways based on the above-mentioned work by Pipino and Wang. The model was also applied to a dataset of workforce assaults, jointly with safety managers’ perception of data quality.

A related branch of data quality is record linkage/matching, editing and cleaning (Herzog et al., 2007; Winkler, 2004). Research studies on Eurostat’s ESS handbook for quality reports, for adaptation to industry or railways are lacking. See, for example, Scopus database. One reason can be that Eurostat’s handbook is in first hand made for national statistical institutes and for official statistics.

2 EUROSTAT’S ESS HANDBOOK FOR QUALITY REPORTS

European statistics quality criteria are (EC, 2009):

a) ‘relevance’, which refers to the degree to which statistics meet current and potential needs of the users;

b) ‘accuracy’, which refers to the closeness of estimates to the unknown true values;

c) ‘timeliness’, which refers to the period between the availability of the information and the event or phenomenon it describes;

d) ‘punctuality’, which refers to the delay between the date of the release of the data and the target date (the date by which the data should have been delivered);

e) ‘accessibility’ and ‘clarity’, which refer to the conditions and modalities by which users can obtain, use and interpret data;

f) ‘comparability’, which refers to the measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas, sectoral domains or over time;

g) ‘coherence’, which refers to the adequacy of the data to be reliably combined in different ways and for various uses.

Those criteria can be recognised from the brief literature review of the previous section.

For the statistical principles of the European statistics, EC (2009) refers to the Code of Practice (Eurostat, 2011; ESSC, 2015). There are 15 principles of which forms the foundation of European statistics:
1. Professional independence
2. Mandate for data collection
3. Adequacy of resources
4. Commitment to quality
5. Statistical confidentiality
6. Impartiality and objectivity
7. Sound methodology
8. Appropriate statistical procedures
9. Non-excessive burden on respondents
10. Cost effectiveness
11. Relevance
12. Accuracy and reliability
13. Timeliness and punctuality
14. Coherence and comparability
15. Accessibility and clarity

The meaning of reliability differs between data quality and engineering. In EC (2009), reliability is that:

"statistics must measure as faithfully, accurately and consistently as possible the reality that they are designed to represent and implying that scientific criteria are used for the selection of sources, methods and procedures”

In EN 13306 (CEN, 2017), reliability stands for the:

“ability of an item to perform a required function under given conditions for a given time interval”

This difference reflects the different backgrounds of the two areas that are combined in the present study, i.e. the ‘subject area’ of data quality and the ‘application area’ of asset management in railway. Another example is ‘punctuality’, which has a similar, but not the same, meaning regarding data quality and as one of the most important quality attributes of railway traffic respectively. Hence, it is extra important to use a stringent vocabulary and clearly indicate what the meaning of words are in the context where they are used in order to avoid misinterpretation. Note that EC’s Official Journal of the European Communities are published in multiple languages, as well as the Code of Practice.

Guidelines for preparing quality reports are given in ESS’s handbook for quality reports (Eurostat, 2014). The guidelines are referred to as a description of the statistical process.

In this work, we have simplified those guidelines to fit data that are not official statistics, which are in first hand for internal benchmarking and are used by an organisation, where statistics is not their key business function. The aim of this study is to apply the process to railway maintenance records, for data quality assurance and decision support. However, we start with a simplified process for statistics production; see Fig. 1. ‘Target quantities’ refer to the desired quantity/indicator to be obtained. In the end, it is the available data that decides how well we reach the intended quantity. Following the simplified process, we move on to the process based on Eurostat’s handbook; see Figs 2–3. The five predefined processes corresponds to chapters 2–6, part II, of the handbook (Eurostat, 2014).

---

**Figure 1.** A general statistics production process.

**Figure 2.** Process for preparing a quality report based on ESS’s handbook for quality reports.
Figure 3. The predefined processes for preparing a quality report.
3 RESULTS

As a case study, the quality reporting process is applied to repair times of heater systems in railway S&C (Switches & Crossings). The main function of the heater system is to melt snow and ice in order to enable the S&C to move and redirect passing trains. The data has been collected by Trafikverket (Swedish Transport Administration) and consist of corrective maintenance data from the fault reporting system (0Felia) and related delay data (from the Basun system). Both data sets were extracted via Trafikverket’s system for business intelligence within maintenance (LUPP). The data set is further described in the below quality report. Note that the report has been condensed to fit the limited frame of a research article. The processed data is presented in figures following the quality report.

See Eurostat’s handbook for details of the quality reporting steps.

1. Relevance

1.1 Purpose and information needs

The users’ need and purpose of the statistical product is to measure averages and spreads in active repair times of heater sub-systems, for relative comparison. Specifically, the statistical product consists of the three indicators: median repair time, standard deviation of repair times and boxplot of repair times. In this case study, we limit the analysis to the average and spread at the system level of the heaters.

1.2 Content of the statistics

1.2.1. Object and population. The objects are the population of all 7 200 S&C heater systems in Sweden. During the seven years period, those systems caused 8 774 failures out of the total 132 910 failures of the superstructure (7 %). Out of the 8 774, train delaying failures where 637 (7 %).

1.2.2 Variables. The number of train delaying failures and the corresponding active repair times are the target variables.

1.2.3 Estimators. The calculation rules applied to the repair times are: removal of outliers as per percentile and limits; calculation of the lognormal mean, median, variance and standard deviation; fitting of a lognormal distribution; and calculation of boxplot.

The percentile applied is the 90th percentile. The time limits are minimum 10 minutes and maximum 600 minutes repair time. Those rules are subjective and based on what repair times are considered as ordinary failures. Repair times outside those limits are considered as outliers, or extraordinary events.

The lognormal mean \( m \), variance \( \nu \), probability density function \( f(x) \) (or pdf), cumulative density function \( F(x) \) (cdf) and reliability function \( R(x) \) are given by:

\[
m = e^{\mu + \sigma^2/2}
\]

\[
\nu = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}
\]

\[
f(x) = \frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)
\]

\[
F(x) = \Phi\frac{\ln x - \mu}{\sigma}
\]

\[
R(x) = 1 - F(x)
\]

where \( \mu \) and \( \nu \) are the parameters of the distribution, and thus, the mean and standard deviation of the normally distributed logarithm of the variable \( X \). In this case study, the variable \( X \) is repair time of values \( x \), or \( t \). \( \Phi \) is the cumulative distribution function of the standard normal distribution \( N(\mu, \sigma) \). The box plot is given with whiskers of 1.5 IQR (interquartile range).

1.2.4-5 Statistical groups and reference times. No division of objections into groups is performed. The repair times pertain to 2010-01-01 – 2016-12-31.

2. Accuracy and reliability

2.1 Overall reliability

The reported indicators: median repair time, standard deviation of repair times and box plot, give an indication of the active repair time of the heater sub-systems. Indication means that the underlying data should be studied before taking actions, such as redesign of components.

2.2 Sources of uncertainty

Uncertainties includes:

- The failure records report repair time does not measure solely the active repair time (e.g. influenced by access to the infrastructure due to traffic).
- Only including the train delaying failure records is not a guarantee that the repair has been carried out efficiently.
- Repair times are manually logged by maintenance staff and train dispatchers (not computers).
- Specification of the component in question in the failure records are missing in 28 % of the cases.
- Subsystems are not guaranteed to consist of homogeneous objects, i.e. the same models.
- Maintenance contracts and the climate vary over Sweden, even though the regulation is national.

2.2.1-2 Selection and sample size. The sample consists of the whole population of S&C heaters in Sweden, which affects the reliability due to geographically spread out assets. The effects come from different local maintenance contracts and varying winter climate. However, the size of the sample is sufficient for statistical analysis.

2.2.3 Measurement. Repair times are manually logged by maintenance staff and train dispatchers (not computers), which affects the reliability as local maintenance contracts and contractors vary geographically.
2.2.4 Dropouts and outliers. Failure records with repair times shorter than 10 minutes or longer than 600 minutes are excluded as they are considered to be non-ordinary failures based solely on the repair time. Those limits improves the reliability. At a sub-system level, 28% of the failures are missing this data.

2.2.5 Processing. No loss of data has been reported from the data collection to the results.

2.2.6 Model assumptions. The repair times are assumed to follow a lognormal distribution. However, the reported indicator median repair time does not pertain to any distribution, and it is also a statistically robust indicator. The other indicator, standard deviation does depend on the distribution, but the box plot gives a complementary indicator of spread.

3. Timeliness and punctuality
Timeliness and punctuality are not applicable as the quality report concerns only a case study.

4. Coherence and comparability
4.1 Comparability over time. Relative comparison of values over time should note changes in the number of trains (capacity consumption) and changes of maintenance contractors.

4.2 Comparability between groups. Relative comparison of the repair times of the underlying subsystems is the goal of the indicators in the present case study. In case of comparison over geographically regions would instead give a measure of geographical effects and not how hard it is to do maintenance on the individual subsystem. At the same time, the uncertainty due to geographically spread objects need to be kept in mind.

5. Dissemination, accessibility and clarity
5.1 Access to the statistics. The statistics may be distributed through the maintenance/asset management information system. The statistics could be automatically generated on request.

5.2 Possibility of additional statistics. In general, enterprise resource planning system always allow easy access to the raw data in spread sheet formats.

5.3 Presentation. Graphically, the median repair times and spread are suggested to be presented as box plots. In tabular form, the spread is suggested to be presented in terms of standard deviation, alongside the median. Both boxplots and tabular form works well with sorting of sub-systems (statistical groups).

5.4 Documentation. Quality reports may be stored with other performance measurement documentation, such as a KPI (Key Performance Indicator) registry and KPI definition handbook.

The processed data is presented in Figs 4-6. The median repair time of Fig. 4 is 84 minutes, and of Fig. 5, the median is 75 minutes. In Fig. 6, we can see that there is in general a probability of 50% to complete a repair in about 80 minutes.
we need to note that the ‘general description’ field, which is used the most, is missing from the data set.

- The failure records repair time does not measure solely the active repair time.
- Only including the train delaying failure records is not a guarantee that the repair has been carried out efficiently.
- Repair times are manually logged by maintenance staff and train dispatchers (not computers).
- Specification of the component in question in the failure records are missing in 28% of the cases.
- Subsystems are not guaranteed to consist of homogeneous objects, i.e. the same models.
- Maintenance contracts and the climate vary over Sweden.

Factors affecting the reliability positively were found to be application of limits on repair times and use of the median and box plot as they are statistically robust.

5 ACKNOWLEDGMENTS

The work has been supported by funds from Trafikverket (Swedish Transport Administration) under contract LTU-2443-2016.

The authors would like to thank Clas-Göran Rydén and Anders Gustafsson, Trafikverket, and Per Norbin, Sweco, for their support and valuable comments.

6 REFERENCES

Eurostat. 2014. ESS handbook for quality reports, Eurostat (European Commission). DOI: 10.2785/983454


