TECHNICAL SPECIFICATION

ISO/TS 17444-2

Second edition
2017-09

Electronic fee collection — Charging performance —

Part 2: Examination framework

Perception du télépéage — Performance d'imputation —

Partie 2: Cadre d'examen

Reference number
ISO/TS 17444-2:2017(E)

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 204, *Intelligent transport systems*.

This second edition cancels and replaces the first edition (ISO/TS 17444-2:2013), which has been revised with the following changes:

— editorial and formal corrections, as well as changes, to improve readability;

— updated terminology.

A list of all parts in the ISO/TS 17444 series can be found on the ISO website.

Introduction

Electronic tolling systems are complex distributed systems involving critical technology such as dedicated short-range communication (DSRC) and global navigation satellite systems (GNSS), both subject to a certain random behaviour that may affect the computation of the charges. Thus, in order to protect the interests of the different involved stakeholders, in particular Service Users and Toll Chargers, it is essential to define metrics that measure the performance of the system as far as computation of charges is concerned and ensure that the potential resulting errors in terms of size and probability are acceptable. These metrics will be an essential tool when establishing requirements for the systems and also for examination of the system capabilities both during acceptance and during the operational life of the system.

In addition, in order to ensure the interoperability of different systems, it will be necessary to agree on common metrics to be used and on the actual values that define the required acceptable performances although this is not covered in this document.

This document is defined as a toolbox standard of examination tests plus a method for defining and documenting Specific Examination Frameworks to meet specific needs. The detailed choice of the set of examination tests within an Examination Framework depends on the application and the respective context. Compliance with this specification is understood as using the definitions and prescriptions laid out in this document whenever the respective system aspects are subjected to performance measurements, rather than using other definitions and examination methods than the ones specified in this document.

ISO/TS 17444-1 defines a set of charging performance metrics with appropriate definitions, principles and formulations, which together make up a reference framework for the establishment of requirements for EFC systems and their later examination of the charging performance.

These charging performance metrics are intended for use with any toll scheme, regardless of its technical underpinnings, system architecture, tariff structure, geographical coverage, or organizational model. They are defined to treat technical details that may be different among technologies as a “black box”. They focus solely on the outcome of the charging process, i.e. the amount charged in relation to a pre-measured or theoretically correct amount, rather than intermediate variables from various components as sensors, such as positioning accuracy, signal range, or optical resolution. This approach ensures comparable results for each metric in all relevant situations.

The metrics are designed to cover the information exchanged on the front-end interface and the interoperability interfaces between Toll Service Providers and Toll Chargers, as well as information on the end-to-end level.

Metrics for the following information exchanges are defined:

— Charge Reports;

— Toll Declarations;

— Billing Details and associated event data;

— Payment Claims on the level of user accounts;

— End-to-End Metrics which assess the overall performance of the charging process.

The proposed metrics are specifically addressed to protect the interests of the actors in a toll system, such as Toll Service Providers, Toll Chargers and Service Users. The metrics can be used to define requirements (e.g. for requests for proposals) and for performance assessment.

Toll schemes take on various forms as identified in the ISO 17575 series and ISO 14906. In order to create a uniform performance metric specification, toll schemes are grouped into two classes based on the character of their primary charging variable:

— charging based on discrete events (charges associated to the fact that a vehicle is crossing or standing within a certain zone);

— those based on a continuous measurement (duration or distance).

In all these toll schemes, tolls may additionally vary as a function of vehicle class characteristics such as trailer presence, number of axles, taxation class, operating function, and depending on time of day or day of week, such that, for example, tariffs are higher in rush hour and lower on the weekends.

With this degree of complexity, it is not surprising to find that the attempts to evaluate and compare technical solutions for Service User charging have been made uniquely each time a procurement or study is initiated, and with only limited ability to reuse prior comparisons made by other testing entities.

Examination Framework

The Examination Framework that is defined in this document is designed for measuring the metrics defined in ISO/TS 17444-1. The general aim is to achieve a maximum comparability and reproducibility of the results without restricting the technological choices in system design. Specific Examination Frameworks may be defined for the Evaluation and Monitoring Phases of a project due to the differences in the availability of equipped vehicles.

Evaluation Phase

This phase encompasses system evaluation and selection, as well as commissioning and ramp up during implementation. Important aspects of this phase are

— relatively small sample sizes, and

— well controlled behaviour of test vehicles.

Monitoring Phase

After the system has gone into operation, its behaviour needs to be monitored for several reasons, such as fine-tuning of the system performance, monitoring of SLAs between contractual partners (supplier, Toll Charger, Toll Service Provider, etc.). In this phase, the following system aspects can be expected:

— very large sample sizes possible, but with unknown behaviour of the vehicles;

— in principle all measurements from implementation phase possible, too.

Readers Guide

To understand the content of this document, the reader should be aware of the methodology and assumptions used to develop the Examination Framework and associated examination tests; therefore, a suggested reading order is given below.

a) Annex B provides details of the underlying considerations for developing the Examination Framework.

b) Annex C provides background statistical information which will enable the reader to determine sample sizes and confidence limits based on the defined performance requirements.

c) Clause 5 provides the definition of the Examination Framework for the evaluation of Charging Performance.

d) Clause 6 contains the toolbox of Examination Tests for the evaluation of charging performance for the identified scheme types.

e) Annex D contains methods which can be used to reduce the required sample sizes for metrics with high/low probabilities during the evaluation phase.

f) Annex E provides examples of Specific Examination Frameworks which have been developed in accordance with the methodology in 5.2.

Electronic fee collection — Charging performance —

Part 2: Examination framework

1 Scope

This document defines the Examination Framework for the measurement of charging performance metrics defined in ISO/TS 17444-1 to be used during Evaluation and/or on-going Monitoring.

It specifies a method for the specification and documentation of a Specific Examination Framework which can be used by the responsible entity to evaluate charging performance for a particular information exchange interface or for overall charging performance within a Toll Scheme.

It provides a toolbox of Examination Tests for the roles of Toll Charger and Toll Service Provider for the following Scheme types:

a) DSRC Discrete;

b) Autonomous Discrete;

c) Autonomous Continuous.

The detailed choice of the set of examination tests to be used depends on the application and the respective context. Compliance with this specification is understood as using the definitions and prescriptions laid out in this document whenever the respective system aspects are subjected to performance measurements, rather than using other definitions and examination methods than the ones specified in this document.

The following aspects are outside the scope of this document.

— This document does not propose specific numeric performance bounds, or average or worst-case error bounds in percentage or monetary units. Those decisions are left to the Toll Charger (or to agreements between Toll Charger and Service Provider).This document does not consider the evaluation of the expected performance of a system based on modelling and measured data from trial at another place.

— This document does not consider the specification of a common reference system which would be required for comparison of performance between systems.

— This document defines measurements only on standardized interfaces. Proprietary interfaces are excluded, because it is not possible to define standardized metrics on such system properties. These excluded interfaces are among others the link between Toll Charger RSE and central systems in DSRC systems, and the additional sensor input of GNSS modules (inertial sensors, CAN-bus for wheel ticks, etc.).

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at http://www.iso.org/obp

3.1
absolute charging error
difference between the measured charge (toll) value and the actual value as measured by a reference system where a positive error means that the measurement exceeds the actual value

[SOURCE: ISO/TS 17444-1:2017, 3.1]

3.2
accepted charging error interval
interval of the relative charging error that the toll charger considers as acceptable, i.e. as correct charging

[SOURCE: ISO/TS 17444-1:2017, 3.2]

3.3
average relative charging error
ratio between the sum of computed charges (measurement) associated to a set of vehicles during a certain period of time and the actual charge due (reference) minus 1

[SOURCE: ISO/TS 17444-1:2017, 3.3]

3.4
billing detail
information needed to determine or verify the amount due for the usage of a given service

Note 1 to entry: If the data is accepted by both the Toll Charger and the Toll Service Provider, then it is called a concluded Billing Detail which can be used to issue a Payment Claim.

Note 2 to entry: For a given Transport Service, the Billing Detail is referring to one or several valid Toll Declaration(s). A valid Billing Detail should fulfil formal requirements, including security requirements, agreed between the Toll Service Provider and the Toll Charger.

[SOURCE: ISO 12855:2015, 3.1]

3.5
charge object detection
functionality of the system responsible for detecting chargeable events associated with a *charge object* (3.6)

Note 1 to entry: This event refers to the use of a certain object and not to the mechanisms by which detection is produced.

[SOURCE: ISO/TS 17444-1:2017, 3.5]

3.6
charge object
geographic or road related object for the use of which a charge is applied

[SOURCE: ISO 17575-1:2016, 3.6]

3.8
charge parameter change
event occurring within a tolling system, that is relevant for charge calculation, such as change of vehicle category, but not for the detection of a *charge object* (3.6) itself

Note 1 to entry: Examples of this type of event are changes in vehicle category or time zone.

[SOURCE: ISO/TS 17444-1:2017, 3.7]

3.9
charge report
information containing road usage and related information originated at the Front End

[SOURCE: ISO 17575-1:2016, 3.6]

Note 1 to entry: In the 2009/750/EC, Charge Report is referred to as "Toll Declaration".

3.10
discrete toll scheme
toll scheme where the charge is calculated based on distinct events associated with the identification of *charge objects* (3.6) such as crossing a cordon, passing a bridge, being present in an area, etc.

Note 1 to entry: Each event is associated with a certain charge.

[SOURCE: ISO/TS 17444-1:2017, 3.11]

3.11
continuous toll scheme
toll scheme where the charge is calculated based on the accumulation of continuously measured parameter(s), such as, distance, time, etc.

[SOURCE: ISO/TS 17444-1:2017, 3.10]

3.13
evaluation
systematic process of determining how individuals, procedures, systems or programs have met formally agreed objectives and requirements

[SOURCE: ISO 10795:2011, 1.90]

3.14
Front End
part of a tolling system consisting of an OBE and possibly a proxy where road tolling information and usage data are collected and processed for delivery to the Back End

Note 1 to entry: The Front End comprises the on-board equipment and an optional proxy.

[SOURCE: ISO/TS 19299:2015, 3.17]

3.15
false positive event
event that was erroneously detected but did not take place

[SOURCE: ISO/TS 17444-1:2017, 3.13]

3.16
missed recognition event
usage of a *charge object* (3.6) that is not recorded by the system

[SOURCE: ISO/TS 17444-1:2017, 3.16]

3.17
monitoring
collection and assessment of status data for a process or a system

Note 1 to entry: This can be used to observe metrics during operation.

[SOURCE: ISO/TS 17444-1:2017, 3.16]

3.18
overcharging
situation when the calculated charge is above the *accepted charging error interval* (3.2)

[SOURCE: ISO/TS 17444-1:2017, 3.17]

3.19
payment claim
recurring statement referring to concluded *billing details* (3.4) made available to the payer by the payee indicating and justifying the amount due

Note 1 to entry: The payment claim is used by the Toll Service Provider to issue financial objects to its customers (e.g. invoices on behalf of the Toll Charger). A given toll payment claim refers to billing details and takes into account any specific commercial conditions applicable to a vehicle, a fleet of vehicles, a customer of a Toll Service Provider and/or a Toll Service Provider. A valid payment claim should fulfil formal requirements, including security requirements, agreed between the Toll Service Provider and the Toll Charger.

[SOURCE: ISO 12855:2015, 3.10]

3.20
charging performance metrics
specific calculations used to describe the charging performance of a system

Note 1 to entry: These calculations are technology and schema-independent.

[SOURCE: ISO/TS 17444-1:2017, 3.8]

3.21
population
totality of items under consideration

[SOURCE: ISO 3534-1:2006]

3.22
relative charging error
ratio between the *absolute charging error* (3.1) and the reference value

[SOURCE: ISO/TS 17444-1:2017, 3.19]

3.23
representative trips
trips that are of a distance larger than a defined threshold and so should be considered by the related metrics

Note 1 to entry: Only trips which exceed the threshold and cover the specific types of roads of the Toll Regime should be considered.

Note 2 to entry: The threshold may be defined as zero.

[SOURCE: ISO/TS 17444-1:2017, 3.20]

3.24
sample
subset of a population made up of one or more of its individual parts

[SOURCE: ISO 3534-1:2006, modified]

3.25
toll service user
customer of a toll service provider, i.e. one liable for toll, owner of the vehicle, fleet operator or driver depending on the context

[SOURCE: ISO 17444:2017, 3.23]

3.26
specific examination framework
particular instance of a set of examination tests defined by an entity to determine the performance of specific selected charging metrics during either evaluation and or monitoring

3.27
successful charging
situation where the user has been correctly charged according to the rules of the system

Note 1 to entry: For discrete Toll Schemes, this means that for a given chargeable journey the charge object detections have been correctly identified and for continuous schemes that the Charge determined is within the Accepted Charging Error Interval.

[SOURCE: ISO/TS 17444-1:2017, 3.21]

3.28
toll charger
legal entity charging toll for vehicles in a toll domain

Note 1 to entry: In other documents, the terms operator or toll operator can be used.

[SOURCE: ISO 17573:2015, 3.16]

3.29
toll service provider
legal entity providing customer toll services on one or more toll domains for one or more classes of vehicle

Note 1 to entry: In other documents, the terms issuer or contract issuer can be used.

Note 2 to entry: The Toll Service Provider can provide the OBE or can provide only a magnetic card or a smart card to be used with OBE provided by a third party (just as a mobile telephone and a SIM card can be obtained from different parties).

Note 3 to entry: The Toll Service Provider is responsible for the operation (functioning) of the OBE with respect to tolling.

[SOURCE: ISO 17573:2010, 3.23, modified]

3.30
toll declaration
statement to a toll charger that confirms the presence of a vehicle in a toll domain in a format agreed between the Toll Service Provider and the Toll Charger

Note 1 to entry: A valid Toll Declaration should fulfil formal requirements, including security requirements, agreed between the Toll Service Provider and the Toll Charger.

[SOURCE: ISO/TS 19299:2015, 3.44]

3.31
trip
part of space-time trajectory of a particular vehicle within a toll domain

Note 1 to entry: The exact definition of the start and end of trip is dependent on the Toll Regime and technology approach.

[SOURCE: ISO/TS 17444-1:2017, 3.27]

3.32
undercharging
situation where the calculated charge is below the *accepted charging error interval* (3.2)

[SOURCE: ISO/TS 17444-1:2017, 3.28]

3.33
user account
centrally or on-board stored transport related service rights of the user in his relationship to a service provider

[SOURCE: ISO/TS 17444-1:2017, 3.291]

3.34
user complaint
complaints from users related to a specific service provision

[SOURCE: ISO/TS 17444-1:2017, 3.30]

4 Symbols and abbreviated terms

ARCE	Average Relative Charging Error
BD	Billing Details
CCR	Continuous Charge Report
CCTV	Closed Circuit Television (ISO/TS 17444-1)
CELB	Charging Error Interval Lower Bound
CEUB	Charging Error Interval Upper Bound
CI	Charging Input
CM	Charging Metric
CR	Charge Report
CTD	Continuous Toll Declaration
DCR	Discrete Charge Report
DO	Dedicated OBE Testing
DSRC	Dedicated Short Range Communications (ISO 14906)
DTD	Discrete Toll Declaration
E2E	End to End (ISO/TS 17444-1)
EETS	European Electronic Toll Service (ISO 17573)

EFC	Electronic Fee Collection (ISO 17573)
ESA	Enforcement System ANPR
ESD	Enforcement System DSRC
FE	Front End (ISO 17575-1)
GNSS	Global Navigation Satellite System (ISO/TS 17444-1)
GPP	GNSS path post processing
ICT	Information and Communications Technology
IS	Independent Reference System
ITS	Intelligent Transport Systems (ISO/TS 17444-1)
KPI	Key Performance Indicator
MBDD	Maximum Billing Details Delay
MPCD	Maximum Payment Claim Delay
MTDD	Maximum Toll Declaration Delay
MUSD	Maximum User Statement Delay
OBE	On-Board Equipment (ISO 17573)
PC	Payment Claim
SLA	Service Level Agreement (ISO/IEC 20000-1)
SO	Simulated OBE/FE
TC	Toll Charger (ISO 17573)
TC-BO	Toll Charger back office
TD	Toll Declaration
TSP	Toll Service Provider (ISO 17573)
TSP-BO	Toll Service Provider back office
UA	User Account

5 Examination Framework

5.1 General

A toolbox of Examination Tests for the following scheme types are listed in 6.1 to 6.4:

a) DSRC Discrete (6.1 and optionally, 6.2);

b) Autonomous Discrete (6.3);

c) Autonomous Continuous (6.4).

NOTE These Scheme Types are defined in ISO/TS 17444-1.

Each of these subclauses contains the specific examination tests for the applicable Charging Metrics identified in ISO/TS 17444-1:2017:

— 5.2 defines the process that should be followed to define a specific Examination Framework for a particular purpose;

— 5.3 provides a definition of the sources of data that can be used by the Examination Tests to calculate the Charging Metrics;

— 5.4 provides the definitions of the methods of generating Charging Input referenced in the Examination Tests defined in 5.5;

— 5.5 defines the applicability of the defined Charging Metrics for the Roles of Toll Charger and Toll Service Provider for the three identified scheme types;

— 5.6 provides tables for the selection of Charging Metrics and associated Examination Tests for the Roles of Toll Charger and Toll Service Provider for each identified Scheme Type.

5.2 Method for defining a Specific Examination Framework

5.2.1 General

Figure 1 provides an overview of the process that should be followed to define a specific instance of an Examination Framework for the evaluation of Charging Metrics for the roles of Toll Service Provider and/or Toll Charger, in a particular, Toll Scheme. Further details are provided in 5.2.2 to 5.2.7.

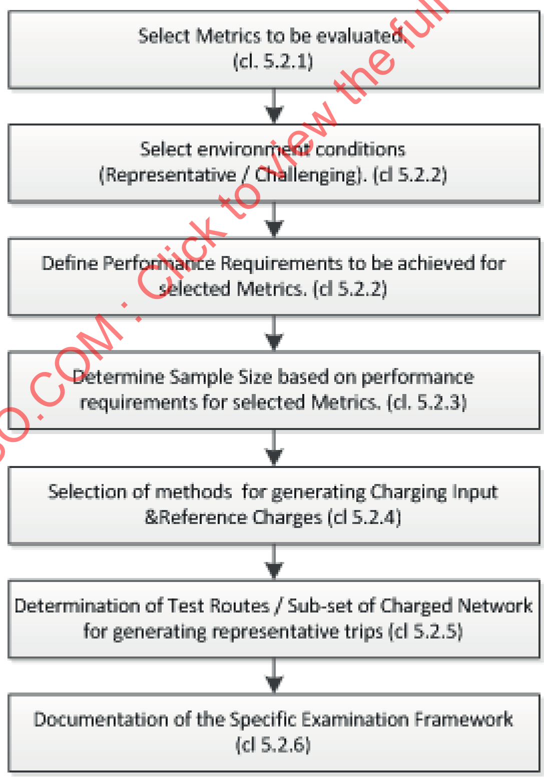


Figure 1 — Method for defining a Specific Examination Framework

5.2.2 Selection of metrics to be evaluated

The entity responsible for the definition of the specific Examination Framework shall determine the metrics to be measured in the phases of Evaluation and Monitoring for the roles of Toll Service Provider and/or Toll Charger using the appropriate tables in 5.6.

a) DSRC Discrete

— Table 5 — DSRC Discrete — Metric Selection Table

— Table 6 — DSRC Discrete — Optional DSRC Toll Declaration Metric Selection Table

b) Autonomous Discrete

— Table 7 — Autonomous Discrete — Metric Selection Table

c) Autonomous Continuous

— Table 8 — Autonomous Continuous — Metric Selection Table

5.2.3 Definition of environmental conditions and associated performance requirements

The entity responsible for the definition of the specific Examination Framework shall determine the environmental conditions (Representative/Challenging) and associated performance requirements to be met for each metric selected in 5.2.2.

NOTE 1 Assessment of Charging Metrics in a Representative Environment allows performance in the operational environment to be assessed. However, care is to be taken to ensure that the Charging Data Input/selection of representative trips is comparable to that expected for the operational environment.

The choice of representative environmental conditions will, in practice, result in a multidimensional parameter space (e.g. air moisture, topography, electromagnetic environment, etc.). It is important to choose these parameters and their values with care to ensure that tests are performed in all realistic sets of conditions (or at least the most probable ones) while keeping the number of necessary tests to a minimum.

NOTE 2 Assessments of Charging Metrics in a Challenging Environment are typically used to determine behaviour for worst case scenarios in the operational environment. Due to the nonlinear dependence of system performance on the environmental conditions, it is difficult to transpose measured performance levels to those in operational systems.

The environmental conditions and associated performance requirements to be met for each metric selected should be documented in each Examination Test within the specific Examination Framework.

NOTE 3 In cases where comparative testing is chosen (e.g. a new population of OBE is introduced into an existing tolling system), the influence of the environmental conditions on the comparison results could be reduced if the tests were performed in parallel. In this case, both populations are exposed to the same conditions. Nonetheless, it is still necessary to perform the step described in 5.2.3. This is important to ensure that the comparative test is performed under all relevant conditions; it also helps to pinpoint dependencies of performance differences to issues with robustness to certain environmental conditions, i.e. one population of equipment being more sensitive to certain environmental conditions than the other.

5.2.4 Determination of Required Sample Sizes

Based on the performance requirements set for each metric selected in 5.2.3, the entity responsible for the definition of the specific Examination Framework shall determine the sample sizes required to provide statistically significant measurements based on the respective formulas for discrete and continuous systems in Annex C. The required sample sizes shall be documented for each Examination Test in the specific Examination Framework.

5.2.5 Selection of methods for generating Charging Input and Reference Data

For each Examination Test in the Specific Examination Framework that requires specific Charging Input to be generated, the responsible entity shall document which identified option for the generation of the Charging Input shall be used (see 5.4). Where reference data are required for the calculation of the metric within a specific Examination Test, the method of generating the reference data shall be documented in the Examination Test.

As a reference for the definition of the methods for generating input data, an analysis of different data sources that can be used for that generation are identified in 5.3.

5.2.6 Determination of Test Routes/Subset of Charged Network for generating representative trips

For each selected Examination Test in the Specific Examination Framework that requires the generation of specific Charging Input/Representative Trips, the responsible entity shall determine and document the Test Route of the Charged Network which shall be used to generate the representative trips.

NOTE 1 As indicated in 5.2.3, particular care is to be taken to ensure that the selected test routes of charged network are selected to meet the required test environment conditions.

NOTE 2 In cases where the test vehicles are not under control of the test, this step is still useful and necessary: It helps in the selection of suitable test vehicles, which should be expected to drive as much as possible on the subset of the charged network. Additionally, it might be necessary to prepare the OBE for the test, e.g. generating geo-data defining the subset of the charged network and loading those data into the Front End.

5.2.7 Documentation of the Specific Examination Framework

By following the process defined in 5.2.2 to 5.2.6, the entity responsible for the definition of the Specific Evaluation Framework will have fully defined and documented the selected Examination Tests within its Specific Examination Framework. A template for the documentation of Examination Tests is provided in Annex A and examples can be found in Annex E.

5.3 Sources of data

The selection of the methods for generating input data is very much constrained by the availability of the different data sources that can be of very different nature depending on:

— the different phases (evaluation/monitoring);

— the type of system (discrete/continuous);

— the available technologies (e.g. DSRC, GNSS, ANPR).

Because the definition of metrics involves, in general, the comparison of measured system values to expected values, their computation requires both the outputs of the charging system under test and reference data that represent those expected values. The establishment of expected values is, obviously, the most complex task as it requires some independent system that has to provide expected values (which are an estimation of the truth with much higher accuracy and reliability than the operational charging system to be analysed.) or a concise statistical analysis applied to a sufficiently large data set, or in most cases both.

Data from the system under test need to be acquired at the different interfaces for which metrics are established including:

— outputs from Front-End (Charge Reports);

— outputs from Toll Service Provider back office (Toll Declarations and User Account);

— outputs from Toll Charger back office (Billing Details and Payment Claims).

The following sources of data are identified.

a) **The Operational Charging** System for the provision of the charges computed by the system that needs to be compared with the reference. These references can be obtained by any of the means described in the following items.

b) **The Operational Enforcement** System.

c) **Independent detection systems** (e.g. road side cameras, a second toll charging system or other records) that allow the determination that a vehicle has or has not passed through a particular road segment.

d) **Simulated OBEs and Front-Ends** (to feed the back office) that allow analysing the capabilities of that back office having as reference the known (simulated) data generated by those OBEs.

e) **Reference systems installed in vehicles** (e.g. highly accurate positioning system based on GPS + inertial measurement unit, Odometer, etc.).

f) **Dedicated probe vehicles**. Depending on the type of metric to be computed (in particular if it refers to discrete or continuous systems) two types of approaches can be used to know the detailed trajectory (and hence the charge due) of these probe vehicles:

   1) Known routes: the basis for the reference data is known by the *a priori* definition of those routes. This is only applicable to discrete systems since continuous systems require a detailed knowledge of the actually followed trajectory;

   2) Use of reference systems as described in item E.

g) **Data from OBE testing** based for instance in the use of GNSS signal simulators or DSRC simulators.

The potential use of these data for the different cases considered is summarized in Table 1.

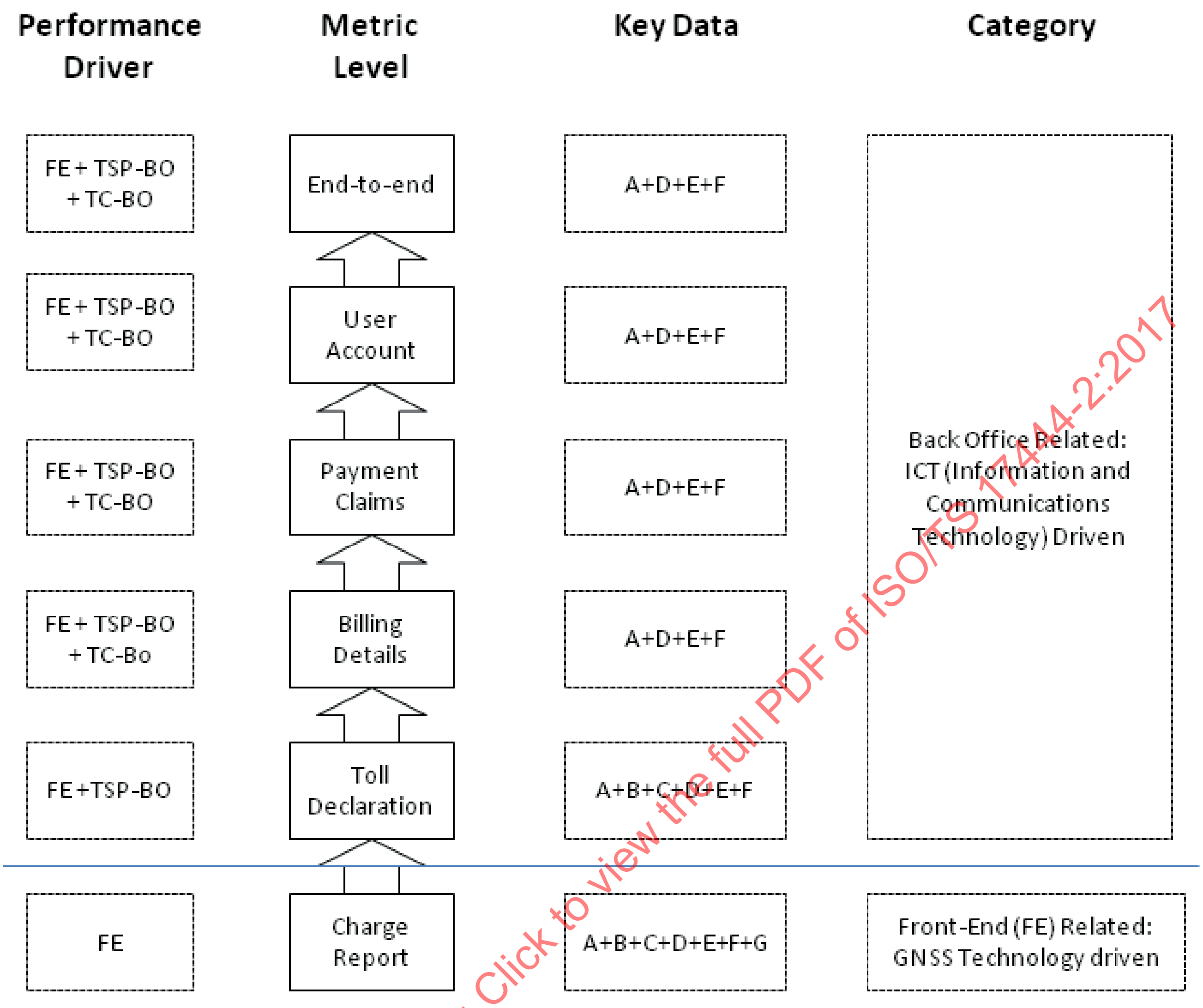
Table 1 — Identification of where different sources of data can be used (valid for both DSRC and autonomous systems)

Identifier	Source of data	Valid for			
		Phase		Scheme type	
		Evaluation	Monitoring	Discrete	Continuous
A	Operational charging system	As far as available	Yes	Yes	Yes
B	Operational Enforcement	As far as available	Yes	Yes	Yes
C	Independent detection systems	Yes	Yes	Yes	Yes
D	Simulated OBEs and FE	Yes	Yes	Yes	Yes
E	Reference systems installed in vehicles	Yes	Yes	Yes	Yes
F	Dedicated probe vehicles	Yes	Yes	Yes	Yes
G	Data from OBE Testing	Yes	Yes	Yes	Yes

Figure 2 and Figure 3 analyse for each of the levels which metrics are defined.

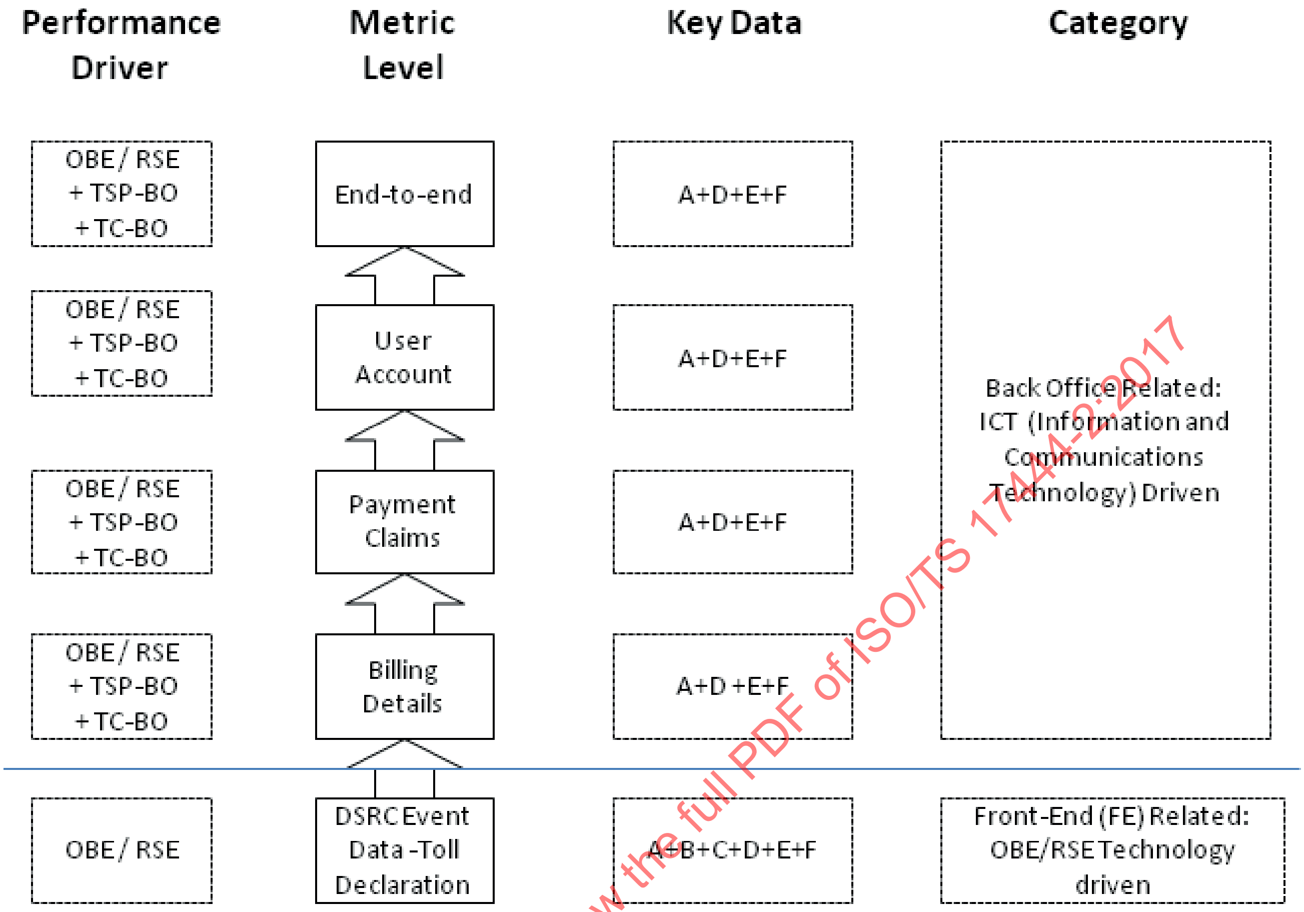
— What are the performance drivers, i.e. the key elements of the system that affects the performances?

— What is the key data (identifier from Table 1) required for the metrics computation?



NOTE Key Data refers to an identifier from Table 1.

Figure 2 — Applicability of different sources of data to different charging metrics for autonomous systems, both discrete and continuous



NOTE Key Data refers to an identifier from Table 1.

Figure 3 — Applicability of different sources of data to different charging metrics for DSRC-based systems

As a result of this analysis, it can be easily derived that two groups of metrics require data and methodology that are of a very different nature.

— Charge Reports and/or Toll Declaration metrics require specific data to determine a reference, which makes the process complex.

— Billing Details, Payment Claims, User Account and E2E metrics can be derived from the Charge Report/Toll Declaration metrics and including additional data that can be easily simulated.

This implies that a major complexity is required for the definition of the Examination Framework at Charge Report/Toll Declaration level.

5.4 Methods of generating charging input

5.4.1 General

The decisions for one or a combination of methods for generating charging input shall take into account the advantages and disadvantages of the respective examination methods. Particular attention has to be paid to:

— number of OBE tested (and what kind of variation of OBEs is necessary);

— number of toll objects versus number of different toll objects examined;

— cost vs. benefit;

— statistics.

Table 2 gives an overview of the methods of generating charging input within the Toll Scheme. In Table 3, a subset of those methods are ordered according to whether the vehicle generating the input is a Service User vehicle (“UVR”), the driver of which might not even be aware that they are participating in performance measurements, or a possibly dedicated probe vehicle (“PV”). The meaning of “predefined routes” and “reference system” are described in the following clauses.

Table 2 — Overview of methods for generating Charging Input

Name	Description	Source of data (5.3)	Identifier for later reference	Subclause
Probe Vehicles (PV)	Controlled/dedicated vehicles on either predefined routes (PVP) or in relation to a reference system (PVR)	A, B, C, E, F	PVP (for predefined routes) PVR (for reference system)	5.4.2 and 5.4.3
Service User vehicles (UVR)	Service Users vehicles in relation to a reference system	A, B, C, E	UVR	5.4.3
Compliance Checking (CC)	Controlled/dedicated vehicles or Service User vehicles at known spots where compliance checking takes places	A, B	CCD (for DSRC system information) CCA (for ANPR system information)	5.4.3.2
Simulated OBE/FE	Simulators used to generate charge object detections	A, D	SO	5.4.3
Dedicated OBE Testing	White box OBE testing	G	DO	5.4.5

Table 3 — Overview of methods for generating Charging Input by vehicles using the Toll Scheme

	Type of vehicle used to generate data	
	Probe vehicle	**Service User vehicle**
Predefined routes	PVP	not applicable
Reference system (one or more of CCD, CCA, PR, IS, GPP); see 5.4.3	PVR	UVR

5.4.2 Predefined routes (identifier: “PVP”)

Testing charging performance using predefined routes is widely used in many toll schemes. Test routes are set up to sample realistic and challenging road conditions, in order to evaluate the charging performances of the whole system which includes the operation of OBEs, roadside equipment, back offices, etc. Owing to the nature of this set-up, predefined routes are only applied for vehicles under the control of the organization responsible for the performance measurements (“probe vehicles”).

The routes shall be designed to be representative of the real Toll Domain. It will sample a driver’s overall driving behaviour in combination with occasional complex geographical features and conditions expected to challenge the technologies. The probe vehicles, which could consist of dedicated vehicles and volunteer vehicles, will undertake a variety of specific manoeuvres/test routes designed to test different aspects of the system. The distance of the routes shall be long enough and the vehicle may drive repeatedly a number of times to ensure that the test results are statistically significant. This may require a significant time (over a period of weeks or months).

NOTE If performances of OBE of autonomous systems (the proposed method is not valid for continuous measurements) are to be compared, all those OBEs can be installed in only one test vehicle; thus, reducing variations caused by external conditions and driving patterns (the OBEs also could be fitted in several vehicles to test). When comparing the performance of DSRC OBEs in a "real" environment, only one OBE should be installed per test vehicle.

Although the paths of the test vehicles are known by definition, it shall be supported by a GNSS reference to account for the fact that even for predefined routes, deviations can occur, that need to be verified later on.

5.4.3 Reference System (used in combination with identifiers: “PVR” and “UVR”)

5.4.3.1 General

A reference system is defined as a setup where a reliable reference is generated with regard to the toll due of vehicles in question. The vehicles can either be ones under the control of the organization responsible for the performance measurements (PVR) or vehicles of Service Users (UVR).

NOTE 1 “UVR” also refers to cases where the tolling service is not operational yet, but data can already be obtained from customers of a Toll Service Provider.

NOTE 2 This reference can be obtained in a two-step process where first, a reference position or track is generated and then, based on this position or track, the toll reference data necessary for the respective test.

In the following subclauses, methods for establishing a reliable reference are described.

5.4.3.2 Reference System: Comparison with a compliance checking system (identifiers: “CCD” and “CCA”)

In order to evaluate the charging accuracy of a toll scheme, DSRC transactions (CCD) or ANPR detections (CCA) from compliance checking infrastructure can be used as reference for comparison of the performance of the OBE. It assumes that such a comparison between an OBE passing a compliance checking gantry and the same OBE passing a “nearby” Charge Object can be made in a meaningful way. This procedure compares the detected Charge Object with the actual Charge Object. It enables a significant number of spot checks for all passing vehicles at all times.

The measurement basis of this method is the comparison of events (from DSRC compliance checks or ANPR compliance checks, or both) detected from gantries with (one or more) closely positioned Charge Objects and fulfilling certain other criteria. A corresponding Charging Report is searched in a certain time window around every enforcement event contained in the sampled Service User vehicle.

CCD and CCA will not have a detection accuracy of 100 %, so the reference of passing Service Users will not be complete. However this does not limit the use of this method. The subset of detected vehicles can be used as the population for which charging performance is determined.

NOTE Real-time comparison of DSRC transactions in the above mentioned way makes it possible to continuously and real-time monitor the toll scheme performance at all times. It is also low cost in terms of personnel required.

5.4.3.3 Reference System: Positioning reference system (identifier: “PR”)

Vehicles (either probe vehicles or Service User vehicles) shall be equipped with a positioning reference system. This can either be a commercial positioning system or a high-accuracy positioning system. For

UVR, Service User vehicles have to be acquired and consented, depending on data protection laws and special contracts signed. An incentive scheme might be necessary.

NOTE As an example, the vehicles could be equipped with high-performance GNSS, such as differential or kinematic GPS for positioning, supplemented by an inertial measurement unit (IMU). This would be to meet the highest possible accurate measurement of distance. Front ends under evaluation will be installed by a Toll Service Provider (TSP).

The vehicles can either run repeatedly on predefined routes or at random, in such a way that is typical for tolled vehicles. The data generated by these vehicles will be reported to a central back office (possibly through the whole computations and interfaces to the billing information to mirror the end-to-end process), analysed and compared with data generated by the reference position and/or distance as measured by the equipment in the vehicle.

Charge events are designed for the toll schemes. Reliable detection of the occurrence of certain events can be assessed as one of the key performance indicators (KPIs) for the system. The accuracy of continuous charging (time, distance) can be assessed between discrete charging events.

5.4.3.4 Reference System: Independent reference system (identifier: “IS”)

It is possible to compare events from an independent system against the output from the system under test. Examples of such independent systems are manual analysis of passing traffic, a second toll charging system (which might be already in place) or other records.

A second toll charging system is especially relevant for tests of new (discrete) interoperable toll systems or the introduction of new technology, which is then compared with an established system. The key issue with this method is the identification of the charge liable vehicles from the passing traffic or the recording of charges liable with a second system. The reliability of the independent reference system has to be established, as well as a procedure for reconciliation of charging/non-charging events when there is a difference between the IS and the system under test.

The method is similar to the comparison with a compliance checking system, but may use alternative traffic monitoring infrastructure not installed specifically for the Toll Scheme.

5.4.3.5 Reference System: GNSS path post processing (identifier: “GPP”)

GNSS path post processing is defined as a setup where a reference is generated by route analysis of the vehicle in question. The vehicles can either be ones under the control of the organization responsible for the performance measurements (PVR) or vehicles of Service Users (UVR).

The method is based on the analysis of GNSS-based vehicle tracks, which allows positioning and tracking of a single or multiple vehicles, mapping their location onto charging objects along with the routes travelled and its detailed historical analysis. This mapping of GNSS tracks onto Charge Objects will be performed by a system independent from the EFC OBE.

NOTE The GNSS tracks can be obtained from the EFC OBE itself.

5.4.4 Simulated OBE/FE (identifier: “SO”)

Simulators/Emulators that can be used to generate simulated charge object detections from the Front-End of the Toll Service Provider (Charge Reports) or Toll Charger (DSRC Transaction Reports). This method can be used in one of two ways:

a) evaluation of Charge Report Generation for given GNSS path;

b) generation of reference Charging Input for the evaluation of metrics for Toll Declaration, Payment Claims, User Account and End-2-End.

Depending on the sophistication of the simulators, it may be possible to simulate operational charging performance under certain defined conditions.

NOTE An example of the possible application of such method could be the suitability for use tests in EETS.

5.4.5 Dedicated OBE Testing (identifier: "DO")

While "black box testing" (i.e. testing based only in the information available in the "public" interfaces) is intended as the main mechanism for the definition of the examination tests described in Clause 6, it is anticipated that some metrics, especially those involving very high or very low probabilities would require a huge sample size that can make "black box testing" unfeasible due to the high cost and long schedule required.

In those cases, alternative testing methods are required and dedicated OBE testing ("white box testing) is one of those alternative methods. In this method, either simulated inputs are injected to the OBE and/or internal data analysed. This sort of analysis is potentially technology-dependent and, therefore, may require knowledge of the OBE design. As such, they cannot be defined *a priori* and therefore, they require special procedure to be agreed between the Toll Service Provider and the OBE supplier.

Dedicated OBE testing may involve the use of GNSS signal simulators (if no other position technology is hybridized within the OBE) to simulate demanding environment conditions, as well as the access to intermediate data of the OBE from which metrics can be derived.

A more detailed description of alternative methods to the one described in Clause 6, including dedicated OBE testing, are proposed in Annex D.

While the departure from the black box testing approach makes it all but impossible to provide input for reliable comparison, methods like that are widely used and expected to gain even more importance in the future. If used with care, dedicated OBE testing could provide valuable insight hard to obtain otherwise. In evaluating the result, the nonlinear system behaviour shall be taken into account (e.g. barely or just not quite receiving the signal of the fourth GNSS satellite, resulting in good or no position data).

5.5 Applicability of metrics scheme types

Table 4 defines the applicability of the defined Charging Metrics for the following types of Toll Schemes:

a) DSRC Toll Scheme (DD) — Toll Charger with one or more Toll Service Providers;

b) Autonomous Discrete Toll Scheme (AD) — Toll Charger with one or more Toll Service Providers;

c) Autonomous Continuous Toll Scheme (AC) — Toll Charger with one or more Toll Service Providers.

The entries in Table 4 Scheme Type have the following meaning when considering their inclusion in the definition of a specific Examination Framework:

— "Y" indicates that the metric is applicable;

— "O" indicates that this is an optional metric;

— " " indicates that the metric is not applicable.

In addition for each metric, the following pieces of information are provided.

a) Key Data Requirements- Data required to calculate the metric:

  1) RD — Reference Data;

  2) UA — User Account;

  3) PC — Payment Claims;

  4) BD — Billing Details;

5) TD — Toll Declarations;

6) CR — Charge Reports;

7) CCR — Compliance Check Records.

b) Evaluation — Charging Input Method Options– the applicable methods that could be used (see 5.4), entries of PVR and UVR imply the use of one or more of the Reference Identifiers: CCD, CCA, PR, IS, GPP.

c) Monitoring — Charging Input Method Options — the applicable methods that could be used (see 5.4), entries of PVR and UVR imply the use of one or more of the Reference Identifiers: CCD, CCA, PR, IS, GPP.

Table 4 — Applicability of Charging Metrics for Scheme Types and Roles

	Metric ID	Metric name	Scheme type			Key data requirements	Evaluation Charging Input Method Options	Monitoring Charging Input Method Options
			DD	AD	AC			
End-to-End	CM-E2E-1	E2E — Correct Charging Rate	Y	Y	Y	RD, UA	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-E2E-2	E2E — Overcharging Rate	Y	Y	Y	RD, UA	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-E2E-3	E2E — Undercharging Rate	Y	Y	Y	RD, UA	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-E2E-4	E2E — Late Charging Rate	Y	Y	Y	RD, UA	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
User Account	CM-UA-1	UA — Correct Charging Rate	Y	Y	Y	RD, UA	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-UA-2	UA — Overcharging Rate	Y	Y	Y	RD, UA	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-UA-3	UA — Undercharging Rate	Y	Y	Y	RD, UA	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-UA-4	UA — Accurate application of Payments and Re-funds	Y	Y	Y	UA	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-UA-5	UA — Accurate Personalisation of OBUs	Y	Y	Y	CCR	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
Payment Claims	CM-PC-1	PC — Correct Charging Rate	Y	Y	Y	RD, PC	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-PC-2	PC — Overcharging Rate	Y	Y	Y	RD, PC	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-PC-3	PC — Undercharging Rate	Y	Y	Y	RD, PC	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-PC-4	PC — Latency — TC	Y	Y	Y	PC	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-PC-5	PC — Late Payment Claims	Y	Y	Y	PC	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
	CM-PC-6	PC — Rejected Payment Claim Rate	Y	Y	Y	RD, PC	PVP,PVR,UVR,SO	PVP,PVR,UVR,SO
Billing Details	CM-BD-1	BD — Correct Charging Rate	Y	Y	Y	RD, BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-BD-2	BD — Overcharging Rate	Y	Y	Y	RD, BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-BD-3	BD — Undercharging Rate	Y	Y	Y	RD, BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-BD-4	BD — Incorrect Charging Rate	Y	Y	Y	RD, BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-BD-5	BD — Latency — TC	Y	Y	Y	BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-BD-6	BD — Late Billing Details Rate	Y	Y	Y	BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-BD-7	BD — Rejected Billing Details Rate	Y	Y	Y	RD, BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-BD-8	BD — Incorrect rejected Billing Details Rate	Y	Y	Y	RD, BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-BD-9	BD — Inferred Billing Details Rate	Y	Y	Y	RD, BD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO

Table 4 *(continued)*

	Metric ID	Metric name	Scheme type			Key data requirements	Evaluation Charging Input Method Options	Monitoring Charging Input Method Options
			DD	AD	AC			
Toll Declaration	CM-TD-1	TD — Correct Toll Declaration Generation Rate	O	Y	Y	RD, TD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-TD-2	TD — Incorrect Toll Declaration Generation Rate	O	Y	Y	RD, TD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-TD-3	TD — Late Toll Declarations Rate	O	Y	Y	RD, TD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-TD-4	TD — TSP Charge Parameter Change Rate	O	Y	Y	RD, TD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-TD-5	TD — TSP False Positive Rate	O	Y	Y	RD, TD	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-DTD-1	DTD — Correct Charging Rate (charge object detections)		Y		RD, TD	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-DTD-2	DTD — Incorrect Charge Event recognition		Y		RD, TD	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-DTD-3	DTD — Missed Charge Object Detection Rate		Y		RD, TD	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-DTD-4	DTD Overcharging Rate		Y		RD, TD	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-CTD-1	CTD Correct Charging Rate			Y	RD, TD	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-CTD-2	CTD Overcharging Rate			Y	RD, TD	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-CTD-3	CTD Accuracy of Distance/Time Measurement			Y	RD, TD	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO

Table 4 *(continued)*

	Metric ID	Metric name	Scheme type			Key data requirements	Evaluation Charging Input Method Options	Monitoring Charging Input Method Options
			DD	AD	AC			
Charge Report	CM-CR-1	CR — Correct Charge Report Generation		Y	Y	RD, CR	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-CR-2	CR — Incorrect Charge Report Generation Rate		Y	Y	RD, CR	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-CR-3	CR — Charge Report Latency		Y	Y	RD, CR	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-CR-4	CR — TSP Front End Charge Parameter Change Rate		Y	Y	RD, CR	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-CR-5	CR — TSP Front End False Positive Rate		Y	Y	RD, CR	PVP,PVR,UVR,IS,SO	PVP,PVR,UVR,IS,SO
	CM-DCR-1	DCR — Correct Charging Rate (charge object detections)		Y		RD, CR	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-DCR-2	DCR — Incorrect Charge Event recognition		Y		RD, CR	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-DCR-3	DCR — Missed Charge Object Detection Rate		Y		RD, CR	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-DCR-4	DCR — Overcharging rate (Incorrect false positive Charge Event Recognition)		Y		RD, CR	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-CCR-1	CCR — Correct Charging Rate			Y	RD, CR	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-CCR-2	CCR — Overcharging Rate			Y	RD, CR	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO
	CM-CCR-3	CCR — Accuracy of Distance/Time Measurement			Y	RD, CR	PVP,PVR,UVR,IS,SO,-DO	PVP,PVR,U-VR,IS,SO,DO

5.6 Charging Metric Selection Tables

5.6.1 General

The entity responsible for the definition of the specific Examination Framework shall use the appropriate table(s) from 5.6.2 to 5.6.4 to define, for each selected Metric to be measured during an Evaluation and/or Monitoring, the:

— Charging Input (CI) Method to be used;

— Target Value to be achieved.

The completed table(s) shall form part of the documentation of the specific Examination Framework.

5.6.2 DSRC Discrete

Table 5 should be used to define the examination tests to be performed for a specific Examination Framework in a DSRC Discrete Scheme.

Table 6 contains additional tests which might be helpful for DSRC systems, but which cannot be prescribed in a normative way, because the respective interfaces are proprietary.

Table 5 — DSRC Discrete — Metric Selection Table

Metric ID	Examination test (subclause)	Evaluation		Monitoring	
		CI method	Target value	CI method	Target value
CM-E2E-1 E2E — Correct Charging Rate	6.1.2				
CM-E2E-2 E2E — Overcharging Rate	6.1.3				
CM-E2E-3 E2E — Undercharging Rate	6.1.4				
CM-E2E-4 E2E — Late Charging Rate	6.1.5				
CM-UA-1 UA — Correct Charging Rate	6.1.6				
CM-UA-2 UA — Overcharging Rate	6.1.7				
CM-UA-3 UA — Undercharging Rate	6.1.8				
CM-UA-4 UA — Accurate application of Payments and Refunds	6.1.9				
CM-UA-5 UA — Accurate Personalisation of OBUs	6.1.10				
CM-PC-1 PC — Correct Charging Rate	6.1.11				
CM-PC-2 PC — Overcharging Rate	6.1.12				
CM-PC-3 PC — Undercharging Rate	6.1.13				
CM-PC-4 PC — Latency — TC	6.1.14				
CM-PC-5 PC — Late Payment Claims	6.1.15				
CM-PC-6 PC — Rejected Payment Claim Rate	6.1.16				
CM-BD-1 BD — Correct Charging Rate	6.1.17				
CM-BD-2 BD — Overcharging Rate	6.1.18				
CM-BD-3 BD — Undercharging Rate	6.1.19				
CM-BD-4 BD — Incorrect Charging Rate	6.1.20				
CM-BD-5 BD — Latency — TC	6.1.21				
CM-BD-6 BD — Late Billing Details Rate	6.1.22				

Table 5 *(continued)*

Metric ID	Examination test (subclause)	Evaluation		Monitoring	
		CI method	Target value	CI method	Target value
CM-BD-7 BD — Rejected Billing Details Rate	6.1.23				
CM-BD-8 BD — Incorrect rejected Billing Details Rate	6.1.24				
CM-BD-9 BD — Inferred Billing Details Rate	6.1.25				

Table 6 — DSRC Discrete — Optional DSRC Toll Declaration Metric Selection Table

Metric ID	Examination test (subclause)	Evaluation		Monitoring	
		CI method	Target value	CI method	Target value
CM-TD-1 TD — Correct Toll Declaration Generation Rate	6.2.2				
CM-TD-2 TD — Incorrect Toll Declaration Generation Rate	6.2.3				
CM-TD-3 TD — Late Toll Declarations Rate	6.2.4				
CM-TD-4 TD — TSP Charge Parameter Change Rate	6.2.5				
CM-TD-5 TD — TSP False Positive Rate	6.2.6				

5.6.3 Autonomous Discrete

Table 7 should be used to define the examination tests to be performed for a specific Examination Framework an Autonomous Discrete Scheme.

Table 7 — Autonomous Discrete — Metric Selection Table

Metric ID	Examination test (subclause)	Evaluation		Monitoring	
		CI method	Target value	CI method	Target value
CM-E2E-1 E2E — Correct Charging Rate	6.1.2				
CM-E2E-2 E2E — Overcharging Rate	6.1.3				
CM-E2E-3 E2E — Undercharging Rate	6.1.4				
CM-E2E-4 E2E — Late Charging Rate	6.1.5				
CM-UA-1 UA — Correct Charging Rate	6.1.6				
CM-UA-2 UA — Overcharging Rate	6.1.7				
CM-UA-3 UA — Undercharging Rate	6.1.8				
CM-UA-4 UA — Accurate application of Payments and Refunds	6.1.9				
CM-UA-5 UA — Accurate Personalisation of OBUs	6.1.10				
CM-PC-1 PC — Correct Charging Rate	6.1.11				
CM-PC-2 PC — Overcharging Rate	6.1.12				
CM-PC-3 PC — Undercharging Rate	6.1.13				
CM-PC-4 PC — Latency — TC	6.1.14				
CM-PC-5 PC — Late Payment Claims	6.1.15				
CM-PC-6 PC — Rejected Payment Claim Rate	6.1.16				
CM-BD-1 BD — Correct Charging Rate	6.1.17				

Table 7 *(continued)*

Metric ID	Examination test (subclause)	Evaluation		Monitoring	
		CI method	Target value	CI method	Target value
CM-BD-2 BD — Overcharging Rate	6.1.18				
CM-BD-3 BD — Undercharging Rate	6.1.19				
CM-BD-4 BD — Incorrect Charging Rate	6.1.20				
CM-BD-5 BD — Latency — TC	6.1.21				
CM-BD-6 BD — Late Billing Details Rate	6.1.22				
CM-BD-7 BD — Rejected Billing Details Rate	6.1.23				
CM-BD-8 BD — Incorrect rejected Billing Details Rate	6.1.24				
CM-BD-9 BD — Inferred Billing Details Rate	6.1.25				
CM-TD-1 TD — Correct Toll Declaration Generation Rate	6.3.2				
CM-TD-2 TD — Incorrect Toll Declaration Generation Rate	6.3.3				
CM-TD-3 TD — Late Toll Declarations Rate	6.3.4				
CM-TD-4 TD — TSP Charge Parameter Change Rate	6.3.5				
CM-TD-5 TD — TSP False Positive Rate	6.3.6				
CM-DTD-1 DTD — Correct Charging Rate (charge object detections)	6.3.7				
CM-DTD-2 DTD — Incorrect Charge Event recognition	6.3.8				
CM-DTD-3 DTD — Missed Charge Object Detection Rate	6.3.9				
CM-DTD-4 DTD Overcharging Rate	6.3.10				
CM-CR-1 CR — Correct Charge Report Generation	6.3.11				
CM-CR-2 CR — Incorrect Charge Report Generation Rate	6.3.12				
CM-CR-3 CR — Charge Report Latency	6.3.13				
CM-CR-4 CR — TSP Front End Charge Parameter Change Rate	6.3.14				
CM-CR-5 CR — TSP Front End False Positive Rate	6.3.15				
CM-DCR-1 DCR — Correct Charging Rate (charge object detections)	6.3.16				
CM-DCR-2 DCR — Incorrect Charge Event recognition	6.3.17				
CM-DCR-3 DCR — Missed Charge Object Detection Rate	6.3.18				
CM-DCR-4 DCR — Overcharging rate (Incorrect false positive Charge Event Recognition)	6.3.19				

5.6.4 Autonomous Continuous

Table 8 should be used to define the examination tests to be performed for a specific Examination Framework in an Autonomous Continuous Scheme.

Table 8 — Autonomous Continuous — Metric Selection Table

Metric ID	Examination test (subclause)	Evaluation		Monitoring	
		CI method	Target value	CI method	Target value
CM-E2E-1 E2E — Correct Charging Rate	6.1.2				
CM-E2E-2 E2E — Overcharging Rate	6.1.3				
CM-E2E-3 E2E — Undercharging Rate	6.1.4				
CM-E2E-4 E2E — Late Charging Rate	6.1.5				
CM-UA-1 UA — Correct Charging Rate	6.1.6				
CM-UA-2 UA — Overcharging Rate	6.1.7				
CM-UA-3 UA — Undercharging Rate	6.1.8				
CM-UA-4 UA — Accurate application of Payments and Refunds	6.1.9				
CM-UA-5 UA — Accurate Personalisation of OBUs	6.1.10				
CM-PC-1 PC — Correct Charging Rate	6.1.11				
CM-PC-2 PC — Overcharging Rate	6.1.12				
CM-PC-3 PC — Undercharging Rate	6.1.13				
CM-PC-4 PC — Latency — TC	6.1.14				
CM-PC-5 PC — Late Payment Claims	6.1.15				
CM-PC-6 PC — Rejected Payment Claim Rate	6.1.16				
CM-BD-1 BD — Correct Charging Rate	6.1.17				
CM-BD-2 BD — Overcharging Rate	6.1.18				
CM-BD-3 BD — Undercharging Rate	6.1.19				
CM-BD-4 BD — Incorrect Charging Rate	6.1.20				
CM-BD-5 BD — Latency — TC	6.1.21				
CM-BD-6 BD — Late Billing Details Rate	6.1.22				
CM-BD-7 BD — Rejected Billing Details Rate	6.1.23				
CM-BD-8 BD — Incorrect rejected Billing Details Rate	6.1.24				
CM-BD-9 BD — Inferred Billing Details Rate	6.1.25				
CM-TD-1 TD — Correct Toll Declaration Generation Rate	6.4.2				
CM-TD-2 TD — Incorrect Toll Declaration Generation Rate	6.4.3				
CM-TD-3 TD — Late Toll Declarations Rate	6.4.4				
CM-TD-4 TD — TSP Charge Parameter Change Rate	6.4.5				
CM-TD-5 TD — TSP False Positive Rate	6.4.6				
CM-CTD-1 CTD Correct Charging Rate	6.4.7				
CM-CTD-2 CTD Overcharging Rate	6.4.8				
CM-CTD-3 CTD Accuracy of Distance/Time Measurement	6.4.9				
CM-CR-1 CR — Correct Charge Report Generation	6.4.10				
CM-CR-2 CR — Incorrect Charge Report Generation Rate	6.4.11				
CM-CR-3 CR — Charge Report Latency	6.4.12				

Table 8 *(continued)*

Metric ID	Examination test (subclause)	Evaluation		Monitoring	
		CI method	Target value	CI method	Target value
CM-CR-4 CR — TSP Front End Charge Parameter Change Rate	6.4.13				
CM-CR-5 CR — TSP Front End False Positive Rate	6.4.14				
CM-CCR-1 CCR — Correct Charging Rate	6.4.15				
CM-CCR-2 CCR — Overcharging Rate	6.4.16				
CM-CCR-3 CCR — Accuracy of Distance/Time Measurement	6.4.17				

6 Examination Tests

6.1 Common (and DSRC Discrete) Examination Tests

6.1.1 General

The Examination Tests defined in 6.1.2 to 6.1.25 are applicable to all Scheme Types. They may be used in combination with Evaluation Tests defined in 6.2 to 6.4, depending on the Scheme Type.

The Reference Data are collected using the methods and prescriptions in 5.4.

6.1.2 ET-CM-E2E-1 E2E — Correct Charging Rate

6.1.2.1 Metric definition

This metric defines the probability that for a set of representative trips travelled by a set of Users during a time span, Δ*t*, the Average Relative Charging Error is within the Accepted Charging Error Interval.

6.1.2.2 Intended use

Typically, this is the view of the Road Operator who receives the revenue from the Toll Collection System. This metric measures the overall correct charging performance across all users.

Traditionally, metrics like these were used to specify overall charging performance where the roles of Toll Charger and Toll Service Provider were performed by the same organization.

6.1.2.3 Metric calculation method

$$\text{E2E Correct Charging Rate} = A/B \qquad (1)$$

where

A is the number of representative trips where the Relative Charging Error (RCE) is within the Accepted Charging Error Interval;

B is the total number of representative trips within the time span;

and, the Relative Charging Error (RCE) = $(C/D) - 1$

where

C is the Actual Charge for representative trip;

D is the Expected Charge for representative trip;

and

the Accepted Charging Error Interval Lower bound (CELB) = −x %;

the Accepted Charging Error Interval Upper Bound (CEUB) = +y %.

6.1.3 ET-CM-E2E-2 E2E — Overcharging Rate

6.1.3.1 Metric definition

This metric defines the probability that for any set of *representative trips* travelled by a set of Users during a time span, Δt, the Average Relative Charging Error is above the Accepted Charging Error Interval.

6.1.3.2 Intended use

Typically, this is the view of the Road Operator who receives the revenue from the Toll Collection System. This metric measures the overall overcharging across all users.

Traditionally, metrics like these were used to specify overall charging performance where the roles of Toll Charger and Toll Service Provider were performed by the same organization.

6.1.3.3 Metric calculation method

$$\text{E2E Overcharging Rate} = A/B \quad (2)$$

where

A is the number of representative trips where the Relative Charging Error (RCE) is greater than the Upper Bound of the Accepted Charging Error Interval;

B is the total number of representative trips within the time span.

and, the Relative Charging Error (RCE) = (C/D) − 1

where

C is the Actual Charge for representative trip;

D is the Expected Charge for representative trip;

and

the Accepted Charging Error Interval Lower bound (CELB) = −x %.

6.1.4 ET-CM-E2E-3 E2E — Undercharging Rate

6.1.4.1 Metric definition

This metric defines the probability that for a set of *representative trips* travelled by a set of Users during a time span, Δt, the Average Relative Charging Error is below the Accepted Charging Error Interval.

6.1.4.2 Intended use

Typically, this is the view of the Road Operator who receives the revenue from the Toll Collection System. This metric measures the overall undercharging across all users.

Traditionally, metrics like these were used to specify overall charging performance where the roles of Toll Charger and Toll Service Provider were performed by the same organization.

6.1.4.3 Metric calculation method

$$\text{E2E Undercharging Rate} = A/B \quad (3)$$

where

A is the number of representative trips where the Relative Charging Error (RCE) is less than the Lower bound of the Accepted Charging Error Interval;

B is the total number of representative trips within the time span;

and, the Relative Charging Error (RCE) = $(C/D) - 1$

where

C is the Actual Charge for representative trip;

D is the Expected Charge for representative trip;

and

the Accepted Charging Error Interval Lower bound (CELB) = $-x$ %.

6.1.5 ET-CM-E2E-4 E2E — Late Charging Rate

6.1.5.1 Metric definition

This metric defines the probability that for any set of *representative trips* travelled by a set of Users during a time span, Δt, the Charge Events appear on the User Statement later than the defined period for the Charging Scheme.

6.1.5.2 Intended use

Typically, this is the view of the Road Operator who receives the revenue from the Toll Collection System. This metric measures the overall late charging across all users.

Traditionally, metrics like these were used to specify overall charging performance where the roles of Toll Charger and Toll Service Provider were performed by the same organization.

6.1.5.3 Metric calculation method

$$\text{E2E Late Charging Rate} = A/B \quad (4)$$

where

A is the number of representative trips where the Charge Events appear on the User Statement later than the defined Maximum User Statement Delay (MUSD);

B is the total number of representative trips within the time span;

where, MUSD = x units of time.

6.1.6 ET-CM-UA-1 UA — Correct Charging Rate

6.1.6.1 Metric definition

This metric defines the probability that for any set of *representative trips* travelled by a given User during the invoicing period, the Average Relative Charging Error is within the Accepted Charging Error Interval.

6.1.6.2 Intended use

This metric is defined at the User Account level which defines the charging performance at the level of the individual Service Users.

This metric should be used when there is a defined requirement on the average Correct Charging for individual Service Users. It can be measured by the Toll Service Provider.

6.1.6.3 Metric calculation method

$$\text{UA Correct Charging Rate} = A/B \qquad (5)$$

where

- A is the number of Users where the Average Relative Charging Error (ARCE) is within the Accepted Charging Error Interval during the invoicing period;
- B is the total number of Users;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

- C is the sum of Actual Charges for representative trips for a User;
- D is the sum of Expected Charge for representative trips for a User;

and

the Accepted Charging Error Interval Lower bound (CELB) = $-x$ %;

the Accepted Charging Error Interval Upper Bound (CEUB) = $+y$ %.

6.1.7 ET-CM-UA-2 UA — Overcharging Rate

6.1.7.1 Metric definition

This metric defines the probability that for any set of *representative trips* travelled by a given User during the invoicing period, the Average Relative Charging Error is above the Accepted Charging Error Interval.

6.1.7.2 Intended use

This metric is defined at the User Account level which defines the charging performance at the level of the individual Service Users.

This metric should be used when there is a defined requirement on the average Over Charging for individual Service Users. It can be measured by the Toll Service Provider.

6.1.7.3 Metric calculation method

UA Overcharging Rate = A/B (6)

where

A is the number of Users where the Average Relative Charging Error (ARCE) is above the Accepted Charging Error Interval Upper Bound during the invoicing period;

B is the total number of Users;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

C is the sum of Actual Charges for representative trips for a User;

D is the sum of Expected Charge for representative trips for a User;

and

the Accepted Charging Error Interval Upper Bound (CEUB) = $+y$ %.

6.1.8 ET-CM-UA-3 UA — Undercharging Rate

6.1.8.1 Metric definition

This metric defines the probability that for any set of *representative trips* travelled by a given User during the invoicing period, the Average Relative Charging Error is below the Accepted Charging Error Interval.

6.1.8.2 Intended use

This metric is defined at the User Account level which defines the charging performance at the level of the individual Service Users.

This metric should be used when there is a defined requirement on the average Under Charging for individual Service Users. It can be measured by the Toll Service Provider.

6.1.8.3 Metric calculation method

UA Undercharging Rate = A/B (7)

where

A is the number of Users where the Average Relative Charging Error (ARCE) is below the Accepted Charging Error Lower bound during the invoicing period;

B is the total number of Users;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

C is the sum of Actual Charges for representative trips for a User;

D is the sum of Expected Charge for representative trips for a User;

and

the Accepted Charging Error Interval Lower bound (CELB) = $-x$ %.

6.1.9 ET-CM-UA-4 UA — Accurate application of Payments and Refunds

6.1.9.1 Metric definition

This metric defines the probability that payment transactions associated to a User Account are correct.

6.1.9.2 Intended use

This metric focuses on the ability of the TSP to correctly apply received payments and credits to User Accounts.

6.1.9.3 Metric calculation method

$$\text{UA Accurate application of Payments and Refunds} = A/B \tag{8}$$

where

A is the total number of Correctly Applied Payments and Refunds on User Accounts in a Measurement Period;

B is the total number of Applied Payments and Refunds on User Accounts in Measurement Period.

6.1.10 ET-CM-UA-5 UA — Accurate Personalisation of OBUs

6.1.10.1 Metric definition

This metric defines the probability that the OBU personalisation for any set of Users during a time span, Δt, is correct. The OBU personalisation is a given set of parameters with defined values stored in the OBU. The personalisation is correct when all parameters are readable and the read values are equal to the defined ones.

6.1.10.2 Intended use

This metric can be measured by Toll Chargers using their existing compliance checking infrastructure. It is in the Toll Charger's interest to ensure that this is maximised, as incorrectly personalised OBEs will potentially lead to detected non-compliance.

6.1.10.3 Metric calculation method

$$\text{UA Accurate personalisation of OBUs} = A/B \tag{9}$$

where

A is the number of OBEs where the agreed subset of OBE Parameters have been verified as correct by the Toll Charger Compliance Equipment in the selected time period;

B is the total number of OBEs checked by the Toll Charger Compliance Equipment in the selected time period.

6.1.11 ET-CM-PC-1 PC — Correct Charging Rate

6.1.11.1 Metric definition

This metric defines the probability that for any given Payment Claim, the Average Relative Charging Error is within the Accepted Charging Error Interval.

This metric measures the probability that the relative error in the Payment Claim used for invoicing is within defined limits to protect the interest of both the Toll Charger and the Service User.

6.1.11.2 Intended use

This metric can be used to measure the average level of correct charging at the level of Payment Claims on the User Account Statement.

6.1.11.3 Metric calculation method

$$\text{PC Correct Charging Rate} = A/B \tag{10}$$

where

A is the number of Payment Claims where the Average Relative Charging Error (ARCE) is within the Accepted Charging Error Interval;

B is the total number of Payment Claims;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

C is the sum of the Actual Charges for the Payment Claim;

D is the sum of the Expected Charge for the Payment Claim;

and

the Accepted Charging Error Interval Lower bound (CELB) = $-x$ %;

the Accepted Charging Error Interval Upper Bound (CEUB) = $+y$ %.

6.1.12 ET-CM-PC-2 PC — Overcharging Rate

6.1.12.1 Metric definition

This metric defines the probability that for any given Payment Claim, the Average Relative Charging Error is above the Accepted Charging Error Interval Upper Bound (CEUB).

This metric measures the probability that the relative error in the Payment Claim used for invoicing is above a defined limit. Protecting the interest of the Service User (i.e. avoiding excessive overcharging) requires that this probability is below a very small value.

6.1.12.2 Intended use

This metric can be used to measure the average level of overcharging at the level of Payment Claims on the User Account Statement.

6.1.12.3 Metric calculation method

$$\text{PC Overcharging Rate} = A/B \tag{11}$$

where

- A is the number of Payment Claims where the Average Relative Charging Error (ARCE) is above the Accepted Charging Error Interval;
- B is the total number of Payment Claims;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

- C is the sum of the Actual Charges for the Payment Claim;
- D is the sum of the Expected Charge for the Payment Claim;

and

the Accepted Charging Error Interval Upper Bound (CEUB) = $+y$ %.

6.1.13 ET-CM-PC-3 PC — Undercharging Rate

6.1.13.1 Metric definition

This metric defines the probability that for any given Payment Claim, the Average Relative Charging Error is below the Accepted Charging Error Interval.

6.1.13.2 Intended use

This metric can be used to measure the average level of undercharging at the level of Payment Claims on the User Account Statement.

6.1.13.3 Metric calculation method

$$\text{PC Undercharging Rate} = A/B \tag{12}$$

where

- A is the number of Payment Claims where the Average Relative Charging Error (ARCE) is below the Accepted Charging Error Lower Bound (CELB);
- B is the total number of Payment Claims;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

C is the sum of the Actual Charges for the Payment Claim;

D is the sum of the Expected Charge for the Payment Claim;

and

the Accepted Charging Error Interval Lower bound (CELB) = $-x$ %.

6.1.14 ET-CM-PC-4 PC — Latency — TC

6.1.14.1 Metric definition

This metric defines the Average Time it takes between the approval for a Billing Detail being received by the Toll Charger and the time the associated Payment Claim is created/sent by the Toll Charger.

6.1.14.2 Intended use

This metric provides an indication of the average processing time for Toll Chargers to create Payment Claims following the receipt of approved Billing Details from the Toll Service Provider

6.1.14.3 Metric calculation method

$$\text{PC Latency TC} = A/B \tag{13}$$

where

A is the sum of the total Payment Claim Delay in the measurement period;

B is the number of Payment Claims sent in the measurement period;

and, the Payment Claim Delay = $C - D$

where

C is the time the Payment Claim was sent;

D is the time of receipt of the first associated approved Billing Detail.

6.1.15 ET-CM-PC-5 PC — Late Payment Claims Rate

6.1.15.1 Metric definition

This metric defines the proportion of Payment Claims received by the TSP in a defined period where the time between the charge object detection and the receipt of the associated Payment Claim is greater than the defined period for the charging scheme.

6.1.15.2 Intended use

This metric can be used where there is a requirement for the timeliness of providing Payment Claims to the TSP. It may be used in conduction with similar metrics defined for Toll Declarations and Billing Details.

6.1.15.3 Metric calculation method

$$\text{PC Late Payment Claims Rate} = A/B \quad (14)$$

where

- A is the number of Late Payment Claims in the measurement period;
- B is the number of Payment Claims received in the measurement period;

and, Late Payment Claim is where the Payment Claim Delay is greater than the Maximum Payment Claim Delay (MPCD).

6.1.16 ET-CM-PC-6 PC — Rejected Payment Claim Rate

6.1.16.1 Metric definition

This metric defines the ratio of correctly rejected Payment Claims in relation to the total number of Payment Claims received in the measurement period.

6.1.16.2 Intended use

This metric measures the rate at which the Toll Charger generates incorrect Payment Claims that are detected by the Toll Service Provider.

6.1.16.3 Metric calculation method

$$\text{Rejected Payment Claims Rate} = A/B \quad (15)$$

where

- A is the number of Payment Claims Correctly Rejected by the Toll Service Provider in the measurement period;
- B is the total number of Payment Claims sent by the Toll Charger in the measurement period.

6.1.17 ET-CM-BD-1 BD — Correct Charging Rate

6.1.17.1 Metric definition

This metric defines the probability that for any given Billing Detail, the Average Relative Charging Error is within the Accepted Charging Error Interval.

This metric measures the probability that the relative error in the billing details used for invoicing is within a defined limit to protect the interest of both the Toll Charger and the Service User.

6.1.17.2 Intended use

This metric can be used to measure the average level of correct charging at the level of trips or line items on the User Account Statement.

6.1.17.3 Metric calculation method

$$\text{BD Correct Charging Rate} = A/B \quad (16)$$

where

A is the number of Billing Details where the Average Relative Charging Error (ARCE) is within the Accepted Charging Error Interval;

B is the total number of Billing Details;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

C is the sum of the Actual Charges for the Billing Detail;

D is the sum of the Expected Charge for the Billing Detail;

and

the Accepted Charging Error Interval Lower bound (CELB) = $-x$ %;

the Accepted Charging Error Interval Upper Bound (CEUB) = $+y$ %.

6.1.18 ET-CM-BD-2 BD — Overcharging Rate

6.1.18.1 Metric definition

This metric defines the probability that for any given Billing Detail, the Average Relative Charging Error is above the Accepted Charging Error Interval

This metric measures the probability that the relative error in the billing details ultimately used for invoicing is above a defined limit. Protecting the interest of the Service User (i.e. avoiding excessive overcharging) requires that this probability is below a very small value.

6.1.18.2 Intended use

This metric can be used to measure the average level of overcharging at the level of trips or line items on the User Account Statement.

6.1.18.3 Metric calculation method

$$\text{BD Overcharging Rate} = A/B \qquad (17)$$

where

A is the number of Billing Details where the Average Relative Charging Error (ARCE) is above the Accepted Charging Error Interval;

B is the total number of Billing Details;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

C is the sum of the Actual Charges for the Billing Detail;

D is the sum of the Expected Charge for the Billing Detail;

and

the Accepted Charging Error Interval Upper Bound (CEUB) = $+y$ %.

6.1.19 ET-CM-BD-3 BD — Undercharging Rate

6.1.19.1 Metric definition

This metric defines the probability that for any given Billing Detail, the Average Relative Charging Error is below the Accepted Charging Error Interval.

6.1.19.2 Intended use

This metric can be used to measure the average level of undercharging at the level of trips or line items on the User Account Statement.

6.1.19.3 Metric calculation method

$$\text{BD Undercharging Rate} = A/B \quad (18)$$

where

- A is the number of Billing Details where the Average Relative Charging Error (ARCE) is below the Accepted Charging Error Lower Bound (CELB);
- B is the total number of Billing Details;

and, the Average Relative Charging Error (ARCE) = $(C/D) - 1$

where

- C is the sum of the Actual Charges for the Billing Detail;
- D is the sum of the Expected Charge for the Billing Detail;

and

the Accepted Charging Error Interval Lower bound (CELB) = $-x$ %.

6.1.20 ET-CM-BD-4 BD — Incorrect Charging Rate

6.1.20.1 Metric definition

This metric defines the probability that for any predefined charge object detection that is recorded, a respective Billing Detail is incorrectly generated (the incorrect data is not detected).

"Predefined" may be defined by random measurements of determined Charge Events.

6.1.20.2 Intended use

This metric can be used to measure the proportion of Billing Details which contain one or more errors.

6.1.20.3 Metric calculation method

$$\text{BD Incorrect Charging Rate} = A/B \quad (19)$$

where

- A is the number of charge object detections incorrectly associated to Billing Details, i.e. total number of charge object detections assigned to Billing Details in the measurement period which are not correctly representing Reference Data;
- B is the total number of charge object detections, i.e. total number of charge object detections resulting from the Reference Data in the measurement period.

6.1.21 ET-CM-BD-5 BD — Latency — TC

6.1.21.1 Metric definition

This metric defines the Average Time it takes between a charge object detection occurring and the time the associated Billing Detail is created/sent by the Toll Charger.

6.1.21.2 Intended use

This metric can be used where there is a requirement for the timeliness of providing Charging information to the Service User, it may be used in conduction with similar metrics defined for Toll Declarations and Payment claims, to identify the average processing times for each step of the charging process.

6.1.21.3 Metric calculation method

$$\text{BD Latency TC} = A/B \qquad (20)$$

where

- A is the sum of the total Billing Detail Delay in the measurement period;
- B is the number of Billing Details sent in the measurement period;

and, Billing Detail Delay = $C - D$

where

- C is time the Billing Detail was sent;
- D is time of the first associated charge object detection;

6.1.22 ET-CM-BD-6 BD — Late Billing Details Rate

6.1.22.1 Metric definition

This metric defines the proportion of Billing Details received by the TSP in a defined period where the time between the charge object detection and the receipt of the associated Billing Detail is greater than the defined period for the charging scheme.

6.1.22.2 Intended use

This metric can be used where there is a requirement for the timeliness of providing Charging information to the Service User, it may be used in conduction with similar metrics defined for Toll Declarations and Payment claims.

6.1.22.3 Metric calculation method

$$\text{BD Late Billing Details Rate} = A/B \quad (21)$$

where

- A is the number of Late Billing Details in the measurement period;
- B is the number of Billing Details received in the measurement period;

and, Late Billing Detail is where the Billing Detail Delay is greater than the Maximum Billing Detail Delay (MBDD).

6.1.23 ET-CM-BD-7 BD — Rejected Billing Details Rate

6.1.23.1 Metric definition

This metric defines the ratio of correctly rejected Billing Details in relation to the total number of Billing Details received in the measurement period.

6.1.23.2 Intended use

This metric can be used to measure the performance of the Toll Charger in its ability to correctly generate Billing Details. The expectation is that this should be a small value.

6.1.23.3 Metric calculation method

$$\text{Rejected Billing Details Rate} = A/B \quad (22)$$

where

- A is the number of Billing Details Correctly Rejected by the Toll Service Provider in the measurement period;
- B is the total number of Billing Details sent by the Toll Charger in the measurement period.

6.1.24 ET-CM-BD-8 BD — Incorrectly rejected Billing Details Rate

6.1.24.1 Metric definition

This metric defines the ratio of the incorrectly rejected Billing Details in relation to the total number of rejected Billing Details in the measurement period.

6.1.24.2 Intended use

This metric can be used to measure the Toll Service Provider's performance in terms of incorrectly rejected Billing Details, i.e. those which contain no errors and, in case of Autonomous Schemes, are based on Toll Service Provider Toll Declarations.

6.1.24.3 Metric calculation method

$$\text{Incorrectly Rejected Billing Details Rate} = A/B \quad (23)$$

where

- *A* is the number of Billing Details Incorrectly Rejected by the Toll Service Provider in the measurement period;
- *B* is the total number of Billing Details sent by the Toll Charger in the measurement period.

6.1.25 ET-CM-BD-9 BD — Inferred Billing Details Rate

6.1.25.1 Metric definition

This metric defines the ratio of inferred Billing Details in relation to the total number of Billing Details in the measurement period.

6.1.25.2 Intended use

This metric is most applicable to discrete schemes where based on the information in received Toll Declarations, the Toll Charger is able to identify missing charges and generate inferred Billing Details.

In a DSRC discrete system, this can be used as a measure of DSRC detection performance for different OBE populations.

6.1.25.3 Metric calculation method

$$\text{Inferred Billing Details Rate} = A/B \tag{24}$$

where

- *A* is the number of Billing Details Generated by the Toll Charger without an associated Toll Declaration in the measurement period;
- *B* is the total number of Billing Details sent by the Toll Charger in the measurement period.

6.2 DSRC Discrete — Optional DSRC Toll Declaration Metrics

6.2.1 General

For DSRC, Discrete Schemes Evaluation Tests may be selected from 6.1. In addition, the metrics in 6.2.2 to 6.2.6 may be selected to measure the internal performance of the DSRC Toll Charger.

NOTE Formally, these metrics can only be specified for autonomous systems, where the interfaces for Charge Reports and Toll Declarations are standardized. On the other hand, it could also be used in DSRC systems. In this case, the DSRC transactions on the air interface OBE-RSE replace the ISO 17575-1 Charge Reports. The Toll Declarations of proprietary format on the interface RSE-TC back office replace the Toll Declarations defined in ISO 12855.

6.2.2 ET-CM-TD-1 TD — Correct Toll Declaration Generation Rate

6.2.2.1 Metric definition

This metric defines the probability that a DSRC Toll Declaration is correctly generated.

6.2.2.2 Intended use

This metric measures the performance of a TC system in generating correct DSRC Toll Declarations.

6.2.2.3 Metric calculation method

$$\text{TD Correct Toll Declaration Generation Rate} = A/B \quad (25)$$

where

A is the number of correctly generated Toll Declarations, i.e. number of Toll Declarations generated during the measurement period by the TC Front End which are consistent with the respective total number of charge object detections resulting from the Reference Data correctly assigned to Toll Declarations in the measurement period. Consistency requires correct representation of all charge object detections;

B is the total number of Toll Declarations, i.e. number of Toll Declarations generated by the TC Front End in the measurement period.

6.2.3 ET-CM-TD-2 TD — Incorrect Toll Declaration Generation Rate

6.2.3.1 Metric definition

This metric defines the probability that a DSRC Toll Declaration is incorrectly generated.

6.2.3.2 Intended use

This metric measures the performance of a TC system in generating incorrect Toll Declarations.

6.2.3.3 Metric calculation method

$$\text{TD Incorrect Toll Declaration Generation Rate} = A/B \quad (26)$$

where

A is the number of incorrectly generated Toll Declarations, i.e. the number of Toll Declarations generated during the measurement period by the TC Front End which are inconsistent with the respective total number of charge object detections resulting from the Reference Data correctly assigned to Toll Declarations in the measurement period. Consistency requires correct representation of all charge object detections;

B is the total number of Toll Declarations, i.e. the number of Toll Declarations generated by the TC Front End in the measurement period.

6.2.4 ET-CM-TD-3 TD — Late Toll Declarations Rate

6.2.4.1 Metric definition

This metric defines the proportion of Toll Declarations generated by the Toll Charger in a defined period where the time between the charge object detection and the generation of the associated Toll Declaration is greater than the defined period for the charging scheme.

6.2.4.2 Intended use

This metric measures the performance of the TC system in terms of the delays occurring in the process of generating Toll Declarations. Late Toll Declarations are most likely to occur when there is a communication outage between the RSI and the TC Back end.

6.2.4.3 Metric calculation method

$$\text{TD Late Toll Declarations Rate} = A/B \tag{27}$$

where

- A is the number of late Toll Declarations, i.e. the number of toll declarations generated during the measurement period by the TC Front End which are delayed by a time longer than Maximum Toll Declaration Delay (MTDD) after the actual occurrence of the respective charging events as determined from the Reference Data;
- B is the total number of Toll Declarations, i.e. the number of Toll Declarations generated by the TC Front End in the measurement period.

The Maximum Toll Declaration Delay (MTDD), i.e. maximum acceptable delay, shall be chosen before performing the test and agreed upon by the parties involved, e.g. with an SLA.

6.2.5 ET-CM-TD-4 TD — TSP Charge Parameter Change Rate

6.2.5.1 Metric definition

This metric defines the probability that for any predefined charge parameter change that takes place, the TC properly detects it.

6.2.5.2 Intended use

This metric measures the reliability of the detection and representation in Toll Declarations of DSRC charge parameter changes in the TC system.

NOTE Be aware of the distinction between charge parameter changes and Charging Events.

6.2.5.3 Metric calculation method

$$\text{TD Event Detection Rate} = A/B \tag{28}$$

where

- A is the number of correct Toll Declarations, i.e. the number of Toll Declarations containing the DSRC charge parameter changes expected based on the Reference Data during the measurement period;
- B is the total number of Toll Declarations generated by the TC in the measurement period.

6.2.6 ET-CM-TD-5 TD — TSP False Positive Rate

6.2.6.1 Metric definition

For vehicles not using the infrastructure, this metric defines the probability that for any defined charge object detection, the TC improperly detects it during the creation of Toll Declarations.

6.2.6.2 Intended use

The rate of False Positives is a critical parameter of system performance because it is directly related to customer satisfaction, the number of user complaints and the public perception of a system. This metric measures the rate of occurrence of DSRC Toll Declarations containing false positives in the TC system.

6.2.6.3 Metric calculation method

$$\text{TD False Positives Rate} = A/B \tag{29}$$

where

- A is the number of Toll Declarations containing False Positives which occur during the measurement period;
- B is the total number of Toll Declarations generated by the TC in the measurement period.

6.3 Autonomous Discrete Specific Examination Tests

6.3.1 General

For autonomous discrete systems, all tests defined in 6.1 are applicable. In addition, the measurements defined in 6.3 can be used.

The Toll Declarations are sent by the TSP Back End on the respective interface to the TC. This interface conforms to ISO 12855:2015, 6.11.

The Charge Reports are sent by the TSP Front End to the TSP Back End. This interface is designed according to ISO 17575-1.

The Reference Data is collected using the methods and prescriptions of 5.4.

6.3.2 ET-CM-TD-1 TD — Correct Toll Declaration Generation Rate

6.3.2.1 Metric definition

This metric defines the probability that a DSRC Toll Declaration is correctly generated.

6.3.2.2 Intended use

This metric measures the performance of a TSP system in generating correct Toll Declarations.

6.3.2.3 Metric calculation method

$$\text{TD Correct Toll Declaration Generation Rate} = A/B \tag{30}$$

where

- A is the number of correctly generated Toll Declarations, i.e. number of Toll Declarations generated during the measurement period by the TSP Back End which are consistent with the respective total number of charge object detections resulting from the Reference Data correctly assigned to Toll Declarations in the measurement period; consistency requires correct representation of all charge object detections;
- B is the total number of Toll Declarations, i.e. number of Toll Declarations generated by the TSP Back End in the measurement period.

6.3.3 ET-CM-TD-2 TD — Incorrect Toll Declaration Generation Rate

6.3.3.1 Metric definition

This metric defines the probability that a Toll Declaration is incorrectly generated.

6.3.3.2 Intended use

This metric measures the performance of a TSP system in generating incorrect Toll Declarations.

6.3.3.3 Metric calculation method

$$\text{TD Incorrect Toll Declaration Generation Rate} = A/B \quad (31)$$

where

- A is the number of incorrectly generated Toll Declarations, i.e. the number of Toll Declarations generated during the measurement period by the TSP Back End which are inconsistent with the respective total number of charge object detections resulting from the Reference Data correctly assigned to Toll Declarations in the measurement period. Consistency requires correct representation of all charge object detections;
- B is the total number of Toll Declarations, i.e. the number of Toll Declarations generated by the TSP Back End in the measurement period.

6.3.4 ET-CM-TD-3 TD — Late Toll Declarations Rate

6.3.4.1 Metric definition

This metric defines the proportion of Toll Declarations received by the Toll Charger in a defined period where the time between the charge object detection and the receipt of the associated Toll Declaration is greater than the defined period for the charging scheme.

6.3.4.2 Intended use

This metric measures the performance of the TSP system in terms of the delays occurring in the process of generating Toll Declarations.

6.3.4.3 Metric calculation method

$$\text{TD Late Toll Declarations Rate} = A/B \quad (32)$$

where

- A is the number of late Toll Declarations, i.e. the number of toll declarations generated during the measurement period by the TSP Back End which are delayed by a time longer than Maximum Toll Declaration Delay (MTDD) after the actual occurrence of the last of the respective charging events as determined from the Reference Data;
- B is the total number of Toll Declarations, i.e. the number of Toll Declarations generated by the TSP Back End in the measurement period.

The Maximum Toll Declaration Delay (MTDD), i.e. maximum acceptable delay, shall be chosen before performing the test and agreed upon by the parties involved, e.g. with an SLA.

6.3.5 ET-CM-TD-4 TD — TSP Charge Parameter Change Rate

6.3.5.1 Metric definition

This metric defines the probability that for any predefined charge parameter change that takes place the TSP properly detects it.

6.3.5.2 Intended use

This metric measures the reliability of the detection and representation in Toll Declarations of charge parameter changes in the TSP system.

NOTE Be aware of the distinction between charge parameter changes and Charging Events.

6.3.5.3 Metric calculation method

$$\text{TD Event Detection Rate} = A/B \tag{33}$$

where

A is the number of correct Toll Declarations, i.e. the number of Toll Declarations containing the charge parameter changes expected based on the Reference Data during the measurement period;

B is the total number of Toll Declarations sent by the Toll Service Provider in the measurement period.

6.3.6 ET-CM-TD-5 TD — TSP False Positive Rate

6.3.6.1 Metric definition

For vehicles not using the infrastructure, this metric defines the probability that for any defined charge object detection, the TSP improperly detects it during the creation of Toll Declarations.

6.3.6.2 Intended use

The rate of false positives is a critical parameter of system performance because it is directly related to customer satisfaction, the number of user complaints and the public perception of a system. This metric measures the rate of occurrence of Toll Declarations containing false positives in the TSP system.

6.3.6.3 Metric calculation method

$$\text{TD False Positive Rate} = A/B \tag{34}$$

where

A is the number of Toll Declarations containing False Positives which occur during the measurement period;

B is the total number of Toll Declarations sent by the Toll Service Provider in the measurement period.

6.3.7 ET-CM-DTD-1 DTD — Correct Charging Rate (charge object detections)

6.3.7.1 Metric definition

This metric defines the probability that for any predefined charge object detection that is recorded the corresponding Toll Declaration is correctly generated.

“Predefined” may be defined by random measurements of determined Charge Events.

6.3.7.2 Intended use

This metric tests if each single charge object detection in the Reference Data is represented accordingly in the respective Toll Declaration, measuring the performance of the TSP system for single charge object detections.

6.3.7.3 Metric calculation method

$$\text{DTD Correct Charging Rate} = A/B \tag{35}$$

where

A is the number of charge object detections associated to Toll Declarations, i.e. total number of charge object detections resulting from the Reference Data correctly assigned to Toll Declarations in the measurement period;

B is the total number of charge object detections, i.e. Total number of charge object detections resulting from the Reference Data in the measurement period.

6.3.8 ET-CM-DTD-2 DTD — Incorrect Charge Event Recognition Rate

6.3.8.1 Metric definition

This metric defines the probability that for any predefined charge object detection that is recorded, a respective Toll Declaration is incorrectly generated (the incorrect data is not detected).

"Predefined" may be defined by random measurements of determined Charge Events.

6.3.8.2 Intended use

This metric measures the rate of charge object detections incorrectly detected by the TSP system, i.e. which contain incorrect information (e.g. vehicle category, time, road category, etc.).

6.3.8.3 Metric calculation method

$$\text{DTD Incorrect Charge Event Recognition Rate} = A/B \tag{36}$$

where

A is the number of charge object detections associated to Toll Declarations incorrectly, i.e. total number of charge object detections assigned to Toll Declarations in the measurement period which are not correctly representing Reference Data;

B is the total number of charge object detections, i.e. Total number of charge object detections resulting from the Reference Data in the measurement period.

6.3.9 ET-CM-DTD-3 DTD — Missed Charge Object Detection Rate

6.3.9.1 Metric definition

This metric defines the probability that for any predefined charge object detection, an entry in the respective Toll Declaration is not generated.

6.3.9.2 Intended use

This test measures the rate of missed recognition of Charge Events, i.e. the rate of Charge Events missing in the Toll Declarations.

6.3.9.3 Metric calculation method

DTD Missed Charge Event Recognition Rate = A/B (37)

where

- A is the number of charge object detections not associated to Toll Declarations, i.e. total number of charge object detections resulting from the Reference Data in the measurement period not assigned to Toll Declarations;
- B is the total number of charge object detections, i.e. the total number of charge object detections resulting from the Reference Data in the measurement period.

6.3.10 ET-CM-DTD-4 DTD Overcharging Rate

6.3.10.1 Metric definition

For vehicles not using the infrastructure, this metric defines the probability that for any predefined charge object detection, an additional entry in the respective Toll Declaration is generated (False Positive).

6.3.10.2 Intended use

While the test ET-CM-TD-5 TD — TSP False Positive Rate counts the rate of Toll Declarations containing False Positives, this test measures the rate of single False Positives compared to the overall number of Charge Events.

6.3.10.3 Metric calculation method

DTD Overcharging Rate = A/B (38)

where

- A is the number of Charge Events associated to Toll Declarations which do not result from the Reference Data, i.e. number of Charge Events in Toll Declarations which do not represent actual usage of charged infrastructure;
- B is the total number of charge object detections, i.e. the total number of charge object detections resulting from the Reference Data in the measurement period.

6.3.11 ET-CM-CR-1 CR — Correct Charge Report Generation Rate

6.3.11.1 Metric definition

This metric defines the probability that a Charge Report is correctly generated.

6.3.11.2 Intended use

This test measures the rate of Charge Reports not containing errors.

6.3.11.3 Metric calculation method

CR Correct Charge Report Generation Rate = A/B (39)

where

A is the number of correctly generated Charge Reports, i.e. the number of Charge Reports which contain only Charge Events and charge parameter changes consistent with the Reference Data;

B is the total number of Charge Reports, i.e. the total number of Charge Reports generated by the Toll Service Provider Front-End in the measurement period.

6.3.12 ET-CM-CR-2 CR — Incorrect Charge Report Generation Rate

6.3.12.1 Metric definition

This metric defines the probability that a Charge Report is incorrectly generated.

6.3.12.2 Intended use

This test measures the rate of Charge Reports containing errors.

6.3.12.3 Metric calculation method

$$\text{CR Incorrect Charge Report Generation Rate} = A/B \tag{40}$$

where

A is the number of incorrectly generated Charge Reports, i.e. the number of Charge Reports which contain Charge Events that are not consistent with the Reference Data;

B is the total number of Charge Reports, i.e. the total number of Charge Reports generated by the Toll Service Provider Front-End in the measurement period.

6.3.13 ET-CM-CR-3 CR — Charge Report Latency

6.3.13.1 Metric definition

This metric defines the average time it takes between a Charge Event and the time the Charge Report is created/received by the Service Provider.

6.3.13.2 Intended use

This test gives the average latency of Charge Events until they are received from the Front End. Therefore, it gives information about the average Front End performance.

6.3.13.3 Metric calculation method

$$\text{Charge Report Latency} = \frac{1}{n}\sum_{i=1}^{n} A_i \tag{41}$$

where

A_i is the time span between the occurrence of charge object detection with index *i* as determined from the Reference Data and the reception of the respective Charge Report from the Front End;

n is the number of charge object detections in the measurement period.

NOTE 1 If the measurement period was chosen accordingly, this test could be used to determine the average delay for the data in a single Charge Report. If a longer period was chosen, it could also give the average delay over multiple Charge Reports.

Be aware that the measurement period influences this metric. If e.g. 24h is chosen, the average resulting latency shall be greater than 12 h for evenly spaced charge object detections.

6.3.14 ET-CM-CR-4 CR — TSP Front End Charge Parameter Change Rate

6.3.14.1 Metric definition

This metric defines the probability that the front-end properly detects any defined charge parameter change that takes place.

6.3.14.2 Intended use

This test measures the percentage of charge parameter changes which are correctly reflected in Charge Reports generated by the TSP Front End.

NOTE Be aware of the distinction between charge parameter changes and Charging Events.

6.3.14.3 Metric calculation method

$$\text{CR Event Detection Rate} = A/B \tag{42}$$

where

- A is the number of charge parameter changes which are correctly represented in Charge Reports received from the TSP Front End;
- B is the total number of charge parameter changes as determined from the Reference Data in the measurement period.

6.3.15 ET-CM-CR-5 CR — TSP Front End False Positive Rate

6.3.15.1 Metric definition

For vehicles not using the infrastructure, this metric defines the probability that for any predefined charge object detection, the front-end improperly detects it.

6.3.15.2 Intended use

The rate of False Positives is a critical parameter of system performance because it is directly related to customer satisfaction, the number of user complaints and the public perception of a system. This metric measures the rate of occurrence of Charge Reports in the TSP system containing false positives.

6.3.15.3 Metric calculation method

$$\text{CR Front End False Positive Rate} = A/B \tag{43}$$

where

- A is the number of Charge Reports containing False Positives, i.e. containing Charge Events relevant for the measurement period which do not correspond to actual usage of charged infrastructure
- B is the total number of Charge Reports, i.e. the total number of Charge Reports generated by the Toll Service Provider Front-End for the measurement period.

6.3.16 ET-CM-DCR-1 DCR — Correct Charging Rate (charge object detections)

6.3.16.1 Metric definition

This metric defines the probability that for any predefined *charge object detection* that takes place, the corresponding entry in the respective Charge Report is correctly generated.

“Predefined” may be defined by random measurements of determined charge object detections.

6.3.16.2 Intended use

This test measures the recognition rate of the TSP Front End, i.e. the percentage of road usage that is correctly detected by the Front End.

6.3.16.3 Metric calculation method

$$\text{DCR Correct Charging Rate} = A/B \quad (44)$$

where

A is the number of correctly detected charge object detections in all Charge Reports relevant for the measurement period;

B is the total number of charge object detections, i.e. the total number of charge object detections as determined from the Reference Data in the measurement period.

6.3.17 ET-CM-DCR-2 DCR — Incorrect Charge Event Recognition Rate

6.3.17.1 Metric definition

This metric defines the probability that for any predefined *charge object detection* that takes place, an entry in the respective Charge Report is incorrectly generated.

“Predefined” may be defined by random measurements of determined charge object detections.

6.3.17.2 Intended use

This test measures the percentage of Charge Events in the Charge Reports containing errors, but which correspond to actual charge object detections.

6.3.17.3 Metric calculation method

$$\text{DCR Incorrect Charge Event Recognition Rate} = A/B \quad (45)$$

where

A is the number of Charge Events in all Charge Reports containing wrong information (not detected correctly) relevant for the measurement period;

B is the total number of Charge Events, i.e. the total number of Charge Events as determined from the Charge Reports relevant for the measurement period.

6.3.18 ET-CM-DCR-3 DCR — Missed Charge Object Detection Rate

6.3.18.1 Metric definition

This metric defines the probability that for any predefined *charge object detection*, an entry in the respective Charge Report is not generated.

6.3.18.2 Intended use

This test determines the percentage of charge object detections which are missed by the TSP Front End, resulting in undercharging.

6.3.18.3 Metric calculation method

$$\text{DCR Missed Charge Event Recognition Rate} = A/B \tag{46}$$

where

- A is the number of charge object detections not represented in any Charge Report relevant for the measurement period;
- B is the total number of charge object detections, i.e. the total number of charge object detections as determined from the Reference Data in the measurement period.

6.3.19 ET-CM-DCR-4 DCR — Overcharging rate (Incorrect false positive Charge Event Recognition)

6.3.19.1 Metric definition

For vehicles not using the infrastructure, this metric defines the probability that for any predefined charge object detection, an additional entry in the respective Charge Report is generated ("False Positive").

6.3.19.2 Intended use

The rate of False Positives is a critical parameter of system performance because it is directly related to customer satisfaction, the number of user complaints and the public perception of a system. This metric measures the rate of occurrence of Charge Events representing False Positives in the TSP system.

6.3.19.3 Metric calculation method

$$\text{DCR Overcharging Rate} = A/B \tag{47}$$

where

- A is the number of Charge Events in the Charge Reports which do not correspond to actual infrastructure usage as determined from the Reference Data;
- B is the total number of charge object detections, i.e. the total number of charge object detections as determined from the Reference Data in the measurement period.

6.4 Autonomous Continuous Specific Examination Tests

6.4.1 General

For autonomous continuous systems, all tests defined in 6.1 are applicable. In addition, the measurements defined in 6.4 can be used.

The Toll declarations are sent by the TSP Back End on the respective interface to the TC. This interface conforms to ISO 12855:2014, 6.11.

The Charge Reports are sent by the TSP Front End to the TSP Back End. This interface is designed according to ISO 17575-1.

The Reference Data are collected using the methods and prescriptions defined in 5.4.

6.4.2 ET-CM-TD-1 TD — Correct Toll Declaration Generation Rate

6.4.2.1 Metric definition

This metric defines the probability that a Toll Declaration (based on a GNSS OBE) is correctly generated.

6.4.2.2 Intended use

With this metric, a TSP can measure the overall performance of the system, including its own TSP Back End (in generating Toll Declarations) and its own TSP Front End (in generating Charge Reports).

6.4.2.3 Metric calculation method

$$\text{TD Correct Toll Declaration Generation Rate} = A/B \qquad (48)$$

where

- A is the number of correctly generated Toll Declarations during the measurement period by the TSP Back End;
- B is the total number of Toll Declarations, i.e. number of Toll Declarations generated by the TSP Back End in the measurement period.

6.4.3 ET-CM-TD-2 TD — Incorrect Toll Declaration Generation Rate

6.4.3.1 Metric definition

This metric defines the probability that a Toll Declaration (based on a GNSS OBE) is incorrectly generated.

6.4.3.2 Intended use

With this metric, a TSP can measure the overall performance of the system, including its own TSP Back End (in generating Toll Declarations) and its own TSP Front End (in generating Charge Reports).

6.4.3.3 Metric calculation method

$$\text{TD Incorrect Toll Declaration Generation Rate} = A/B \qquad (49)$$

where

- *A* is the number of incorrectly generated Toll Declarations, i.e. the number of incorrect Toll Declarations generated during the measurement period by the TSP Back End;
- *B* is the total number of Toll Declarations, i.e. the number of Toll Declarations generated by the TSP Back End in the measurement period.

6.4.4 ET-CM-TD-3 TD — Late Toll Declarations Rate

6.4.4.1 Metric definition

This metric defines the proportion of Toll Declarations received by the Toll Charger in a defined period where the time between the charge object detection and the receipt of the associated Toll Declaration is greater than the defined period for the charging scheme.

6.4.4.2 Intended use

This metric measures the performance of the TSP system in terms of the delays occurring in the process of generating Toll Declarations.

6.4.4.3 Metric calculation method

$$\text{TD Late Toll Declarations Rate} = A/B \qquad (50)$$

where

- *A* is the number of late Toll Declarations, i.e. the number of toll declarations generated during the measurement period by the TSP Back End which are delayed by a time longer than Maximum Toll Declaration Delay (MTDD) after the actual occurrence of the last of the respective charging events;
- *B* is the total number of Toll Declarations, i.e. the number of Toll Declarations generated by the TSP Back End in the measurement period.

6.4.5 ET-CM-TD-4 TD — TSP Charge Parameter Change Rate

6.4.5.1 Metric definition

This metric defines the probability that for any predefined charge parameter change that takes place, the TSP properly detects it.

6.4.5.2 Intended use

This metric measures the reliability of the detection of charge parameter changes in the TSP system.

NOTE Be aware of the distinction between Charge-relevant Events and Charging Events.

6.4.5.3 Metric calculation method

$$\text{TSP Event Detection Rate} = A/B \qquad (51)$$

where

A is the number of correctly detected charge parameter changes in the Toll Declarations, during the measurement period;

B is the total number of detected charge parameter changes in the Toll Declarations sent by the Toll Service Provider in the measurement period.

6.4.6 ET-CM-TD-5 TD — TSP False Positive Rate

6.4.6.1 Metric definition

For vehicles not using the infrastructure, this metric defines the probability that for any defined charge object detection, the TSP improperly detects it during the creation of Toll Declarations.

6.4.6.2 Intended use

The rate of False Positives is a critical parameter of system performance because it is directly related to customer satisfaction, the number of user complaints and the public perception of a system. This metric measures the rate of occurrence of false positives in the TSP system.

NOTE In continuous systems, the false positives can imply the application of a wrong tariff to the distance driven within a given infrastructure, e.g. being charged for distance driven within a congestion charging zone while actually only having passed close by.

6.4.6.3 Metric calculation method

$$\text{TD False Positive Rate} = A/B \qquad (52)$$

where

A is the number of False Positives identified in all Toll Declarations which are generated during the measurement period;

B is the total number of passes of vehicles in the proximity (but outside) the charging objects in the measurement period.

6.4.7 ET-CM-CTD-1 CTD Correct Charging Rate

6.4.7.1 Metric definition

This metric defines the probability that for any set of *representative trips* travelled by a vehicle and during a certain period of time, the Average Relative Charging Error is within the Accepted Charging Error Interval.

6.4.7.2 Intended use

This metric provides overall information for the TSP and the Toll Charger on the capabilities of the system, in particular the OBE, to compute Toll Declarations whose charges are within the Accepted Charging Interval, i.e. ensuring that charges cover a very high percentage of the due incomes (to ensure the viability of the system) and limiting the charges in excess (overcharging) to avoid claims and for providing credibility for the users and the authority.

6.4.7.3 Metric calculation method

CR Correct Charge Report Rate = A/B (53)

where

A is the number of sets of representative trips travelled by a vehicle during a certain period of time whose Average Relative Charging Error is within the Accepted Charging Error Interval;

B is the overall number of sets of representative trips analysed.

It shall be determined whether several sets of representative trips shall be independent for each sample or the same trips can be used for more than one set. The second option seems to provide a more reliable metric.

6.4.8 ET-CM-CTD-2 CTD Overcharging Rate

6.4.8.1 Metric definition

This metric defines the probability that for any single predefined *representative trip*, the Relative Charging Error is above the upper bound of the Accepted Charging Error Interval.

6.4.8.2 Intended use

This metric provides overall information for the Front-End provider; the TSP, the Toll Charger, Authority and Users on the capabilities of the system, in particular the OBE, to compute Toll Declarations whose charges are not larger than the Accepted Charging Interval, i.e. ensuring that the probability of probability of charges in excess (overcharging) is properly bounded providing credibility for the users and the authority.

6.4.8.3 Metric calculation method

CR Overcharging Rate = A/B (54)

where

A is the number of representative trips travelled by a vehicle during a certain period of time whose Relative Charging Error is above the upper bound of the Accepted Charging Error Interval;

B is the overall number of representative trips analysed during the mentioned period of time.

6.4.9 ET-CM-CTD-3 CTD Accuracy of Distance/Time Measurement

6.4.9.1 Metric definition

This metric defines the Average and Standard Deviation of the relative distance or time error of a set of *representative trips* travelled by a vehicle during a certain period of time.

6.4.9.2 Intended use

While the above metrics (related to charges) measure the overall system performance (integrating errors in distance measurement and errors in event recognition), having direct observability of system capability to accurately measure distance and time also separating between systematic and random errors (that can be observed by the average and standard deviation values) provides a substantial value for the TSP.

6.4.9.3 Metric calculation method

$$\text{Average} = \frac{\left(\sum_{i=0}^{n} \text{Error}_i\right)}{n} \quad (55)$$

$$\text{Standard Deviation} = \sqrt{\frac{\left(\sum_{i=0}^{n} \text{Error}_i^2\right)}{n}} \quad (56)$$

where

$$\text{Error}_i = \frac{(\text{measured distance} - \text{true distance})}{\text{true distance}}$$

n	is the number of representative trips considered;
measured distance	is the one supplied by the system (could also be time);
true distance	is the one measured by the reference system (could also be time).

6.4.10 ET-CM-CR-1 CR — Correct Charge Report Generation Rate

6.4.10.1 Metric definition

This metric defines the probability that a Charge Report (based on a GNSS OBE) is correctly generated.

6.4.10.2 Intended use

With this metric, a Front End provider and a TSP can measure the performance of the Front-End as far as computation of Charge Reports is concerned.

6.4.10.3 Metric calculation method

CR Correct Charge Report Generation Rate = A/B (57)

where

A is the number of correctly generated Charge Reports during the measurement period and for a given number of vehicles by the Front End;

B is total number of Charge Reports, i.e. number of Charge Reports generated by a given number of vehicles the Front End in the measurement period.

6.4.11 ET-CM-CR-2 CR — Incorrect Charge Report Generation Rate

6.4.11.1 Metric definition

This metric defines the probability that a Charge Report (based on a GNSS OBE) is incorrectly generated.

6.4.11.2 Intended use

With this metric, a Front-End provider and a TSP can measure the performance of the Front-End in respect to the computation of Charge Reports.

6.4.11.3 Metric calculation method

$$\text{CR Incorrect Charge Report Generation Rate} = A/B \tag{58}$$

where

A is the number of incorrectly generated Charge Reports during the measurement period and for a given number of vehicles by the Front End;

B is total number of Charge Reports, i.e. number of Charge Reports generated by a given number of vehicles by the Front-End in the measurement period.

6.4.12 ET-CM-CR-3 CR — Charge Report Latency

6.4.12.1 Metric definition

This metric defines the average time it takes between a Charge Event and the time the Charge Report is created/received by the Service Provider.

6.4.12.2 Intended use

This test gives the average latency of Charge Events until they are received from the Front End. Therefore, it gives information about the average Front End performance.

6.4.12.3 Metric calculation method

$$\text{Charge Report Latency} = \frac{1}{n}\sum_{i=1}^{n} A_i \tag{59}$$

where

A_i is the time span between the occurrence of the charge object detection with index i as determined from the Reference Data and the reception of the respective Charge Report by the Front End;

n is the number of charge object detections in the measurement period.

NOTE 1 If the measurement period was chosen accordingly, this test could be used to determine the average delay for the data in a single Charge Report. If a longer period was chosen, it could also give the average delay over multiple Charge Reports.

Be aware that the reporting period influences this metric. If, for example, 24 h is chosen, the average resulting latency shall be greater than 12 h for evenly spaced charge object detections.

6.4.13 ET-CM-CR-4 CR — TSP Front End Charge Parameter Change Rate

6.4.13.1 Metric definition

This metric defines the probability that the front-end properly detects any defined charge parameter change that takes place.

6.4.13.2 Intended use

This test measures the percentage of charge parameter changes which are correctly reflected in Charge Reports generated by the TSP Front End.

NOTE Be aware of the distinction between charge parameter changes and Charging Events.

6.4.13.3 Metric calculation method

$$\text{CR Event Detection Rate} = A/B \quad (60)$$

where

- A is the number of charge parameter changes which are correctly represented in Charge Reports received from the TSP Front End;
- B is the total number of charge parameter changes as determined from the Reference Data in the measurement period.

6.4.14 ET-CM-CR-5 CR — TSP Front End False Positive Rate

6.4.14.1 Metric definition

For vehicles not using the infrastructure, this metric defines the probability that for any predefined charge object detection, the front-end improperly detects it.

6.4.14.2 Intended use

The rate of False Positives is a critical parameter of system performance because it is directly related to customer satisfaction, the number of user complaints and the public perception of a system. This metric measures the rate of occurrence of Charge Reports in the TSP system containing false positives.

6.4.14.3 Metric calculation method

$$\text{CR Front End False Positive Rate} = A/B \quad (61)$$

where

- A is the number of False Positives identified in all Charge Reports which are generated during the measurement period;
- B is the total number of passes of vehicles in the proximity (but outside) the charging objects in the measurement period.

6.4.15 ET-CM-CCR-1 CCR — Correct Charging Rate

6.4.15.1 Metric definition

This metric defines the probability that for any set of *representative trips* travelled by a vehicle and during a certain period of time, the Average Relative Charging Error is within the Accepted Charging Error Interval.

6.4.15.2 Intended use

This metric provides overall information for the Front End provider, the TSP and the Toll Charger on the capabilities of the system, in particular the OBE, to compute the charges within the Accepted Charging Interval, i.e. ensuring that charges cover a very high percentage of the due incomes (to ensure the viability of the system) and limiting the charges in excess (overcharging) to avoid claims and for providing credibility for the users and the authority.

6.4.15.3 Metric calculation method

CR Correct Charge Report Generation Rate = A/B (62)

where

- A is the number of sets of representative trips travelled by a vehicle during a certain period of time whose Average Relative Charging Error is within the Accepted Charging Error Interval;
- B is the overall number of sets of representative trips analysed.

NOTE In order to have more observability of the process (i.e. the resulting metric to be closer to the reality), sets of representative trips do not need to be fully independent, i.e. different sets can share the same trips. For instance, if the considered period is one month, one can consider periods of 30 days starting each day instead of considering a set each month.

6.4.16 ET-CM-CCR-2 CCR — Overcharging Rate

6.4.16.1 Metric definition

This metric defines the probability that for any single predefined *representative trip*, the Relative Charging Error is above the upper bound of the Accepted Charging Error Interval.

6.4.16.2 Intended use

This metric provides overall information for the Front-End provider; the TSP, the Toll Charger, Authority and Users on the capabilities of the system, in particular the OBE, to compute the charges that are not larger than the Accepted Charging Interval, i.e. ensuring that the probability of charges in excess (overcharging) is properly bounded providing credibility for the users and the authority.

6.4.16.3 Metric calculation method

CR Overcharging Rate = A/B (63)

where

- A is the number of representative trips travelled by a vehicle during a certain period of time whose Relative Charging Error is above the Accepted Charging Error Interval;
- B is the overall number of representative trips analysed.

6.4.17 ET-CM-CCR-3 CCR — Accuracy of Distance/Time Measurement

6.4.17.1 Metric definition

This metric defines the Average and Standard Deviation of the relative distance or time error of a set of *representative trips* travelled by a vehicle during a certain period of time.

6.4.17.2 Intended use

While above metrics related to charges measure the overall system performance (integrating errors in distance measurement and errors in event recognition) having direct observability of system capability to accurately measure distance and time also separating between systematic and random errors provides a substantial value for the TSP.

6.4.17.3 Metric calculation method

$$\text{Average} = \frac{\left(\sum_{i=0}^{n} \text{Error}_i\right)}{n} \tag{64}$$

$$\text{Standard Deviation} = \sqrt{\frac{\left(\sum_{i=0}^{n} \text{Error}_i^2\right)}{n}} \tag{65}$$

where

$$\text{Error}_i = \frac{(\text{measured distance} - \text{true distance})}{\text{true distance}}$$

n	is the number of representative trips considered
measured distance	is the one supplied by the system (could also be time);
true distance	is the one measured by the reference system (could also be time).

Annex A
(informative)

Examination Test documentation template

A.1 Examination Test template

This annex provides a template for the documentation of examination tests.

<table>
<tr><td colspan="2">Measured metric</td><td colspan="4"></td></tr>
<tr><td colspan="6">Metric definition</td></tr>
<tr><td colspan="6"></td></tr>
<tr><td colspan="6">Metric measurement data requirements</td></tr>
<tr><td>Reference Data</td><td>Charge Report</td><td>Toll Declaration</td><td>Billing Details</td><td>Payment Claim</td><td>User Account Charges</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Environmental conditions</td><td colspan="4"></td></tr>
<tr><td colspan="2">Performance requirement</td><td colspan="4"></td></tr>
<tr><td colspan="2">Sample size</td><td colspan="4"></td></tr>
<tr><td colspan="2">Details of method for generating Charging Input</td><td colspan="4"></td></tr>
<tr><td colspan="2">Method for generating Reference Data</td><td colspan="4"></td></tr>
<tr><td colspan="6">Test route/Subset of charge network</td></tr>
<tr><td colspan="6"></td></tr>
<tr><td colspan="6">Metric Calculation Details</td></tr>
<tr><td colspan="6"></td></tr>
</table>

Annex B
(informative)

Examination Framework considerations

B.1 General

For defining the Examination Framework, completeness is as important as efficiency and reliability. The measurements should be repeatable and comparable, considering the diverging requirements from scheme type, phase and technology.

The following systematic three-step process is applied:

a) identification of sources of relevant data, e.g. an operational enforcement system;

b) measurement methods for collecting data from the sources identified in a), e.g. comparison with DSRC events collected by an enforcement system;

c) definition of the Examination Framework for each metric, e.g. CM-DCR-1 DCR — Correct Charging Rate (charge object detections) in an operational tolling system.

This systematic approach allows for reliable and comparable results, but also maximises the opportunities for synergy. One measurement can be used for measuring several metrics at once. In many cases, this gives more than one option of measurement for each metric. This is even necessary because of the varying circumstances due to phase, type and technology.

B.2 Criteria for definition of tests

To ensure that the results of the examinations (measurements of metrics) are reliable, accurate and reproducible, it is important to apply a comprehensive strategy in designing the examination methods. This strategy should take into account the possible sources of influences on metrics.

These influences might be controlled by the examination process (e.g. mounting position of OBE) or might not be accessible to manipulation (e.g. weather conditions).

Therefore, the following strategy of defining examination methods is used in this specification:

— identify possible sources of influences on metrics, considering differences due to phase, scheme type and technology;

— devise tests which “provoke” errors, exploit possible vulnerabilities.

The second point is especially useful in the case of rare events to be tested (e.g. CM-DCR-4, false positives).

B.3 Statistical considerations

The quality of toll-collecting systems can be characterized by performance indices derived for specific aspects, called metrics. Using well-defined sampling procedures, estimation of performance is possible, and can be characterized by confidence intervals for the parameters of interest. Statistical methods applied should be chosen appropriately, reflecting the type of distribution under investigation and the sampling design.

B.4 Dependency on Scheme Type

The design of a valid Examination Framework differs widely depending on the type of scheme.

a) Discrete systems: The results in a discrete system can only assume two values (detected or not). Therefore the result of a test can assume only four values, as elaborated in the Introduction to ISO 17444-1:2017. Percentages only come into play after collecting a sample of several measurements.

   It is also easier to define a valid reference the system under test is to be compared to the passage of a road section is much easier to define than a specific trajectory for a continuous system.

b) Continuous Systems: While the results in a discrete system can only assume two values (detected or not), the result in continuous systems is a value. Therefore, the result of a test should be expressed as a continuous error value.

   Additional problems occur because it is more difficult to reproduce a certain vehicle behaviour, e.g. the mileage for a given trip will vary simply because a vehicles can and will not drive exactly on the same trajectory repeatedly.

The type of scheme also has an impact on the information flows and on the metrics that can be defined. This is described at the end of ISO 17444-1:2017, 5.1. While the metrics are independent of scheme on the higher levels (billing details to end-to-end), there are major differences for the low-level metrics (charge report and toll declaration).

The roles in the collection of charge data differ.

— In autonomous systems, the Toll Service Provider is the operator of the front end; therefore, charge reports and toll declarations are within his realm of responsibility.

— In DSRC-based systems, the roadside equipment and the corresponding parts of the central system are under the control of the Toll Charger.

— In DSRC systems, the interface for exchanging the equivalent to Charge Reports in autonomous systems was not standardized until the preparation of the current document and therefore, it is difficult to put metrics on them.

B.5 Dependency on phase

The two main phases in the lifetime of a tolling system result in very different conditions for charging performance metrics measurement. These differences are detailed in the following subclauses.

a) Evaluation phase

   There are no Service Users yet. All tests should be performed with selected vehicles, either driven by personnel of the entity performing the tests, or volunteers who allow their vehicles to be equipped. This situation results in relatively small sample sizes, potentially with controlled behaviour of test vehicles and controllable influences on metrics (e.g. mounting position of OBU).

b) Monitoring phase

   Usually, a large number of toll system users are active in the system. Therefore, very large sample sizes are possible, but with unknown behaviour of the vehicles and potentially uncontrollable influences on metrics. In principle, all measurements from the implementation phase is also possible.

There might exist a period of transition between the two phases during the roll out of the system. More and more vehicles are equipped with OBEs, so the sample sizes are also rising.

B.6 Dependency on technology

Sources of influences on metrics depend on technology.

Examples

- In GNSS systems, performance is influenced by accuracy and availability of GNSS position data/events. Key issues are:
  - environment of the road (refers to obstacles, buildings, etc.);
  - solar activity (affecting the ionosphere, weather conditions, etc.);
  - vehicle speed;
  - quality of Toll Context data from the Toll Charger (ISO 12855).
- In DSRC systems, performance is influenced by:
  - traffic density;
  - vehicle speed;
  - OBU mounting position;
  - weather conditions;
  - Service User behaviour;
  - battery life.

In both technologies, interference (spoofing, jamming) in the respective frequency bands used degrades performance.

Knowledge about these sources helps in designing useful tests, especially for measuring rare events. As already pointed out in B.2, artificially stressing the system by creating challenging scenarios might provoke errors which can be used for estimation of metric values, given that the actual probability of such conditions is known.

B.7 Simultaneous measurement of metrics

The metrics in the Examination Framework have been defined based on the sequential information flows within a charging system, as described in ISO/TS 17444-1:2017, 5.1, which means that there is a hierarchy of metrics; hence, an examination test designed to measure metrics at the E2E Level can, with additional observation points, be used to simultaneously measure metrics at the Charge Report, Billing Details, Payment Claim and User Account metric levels.

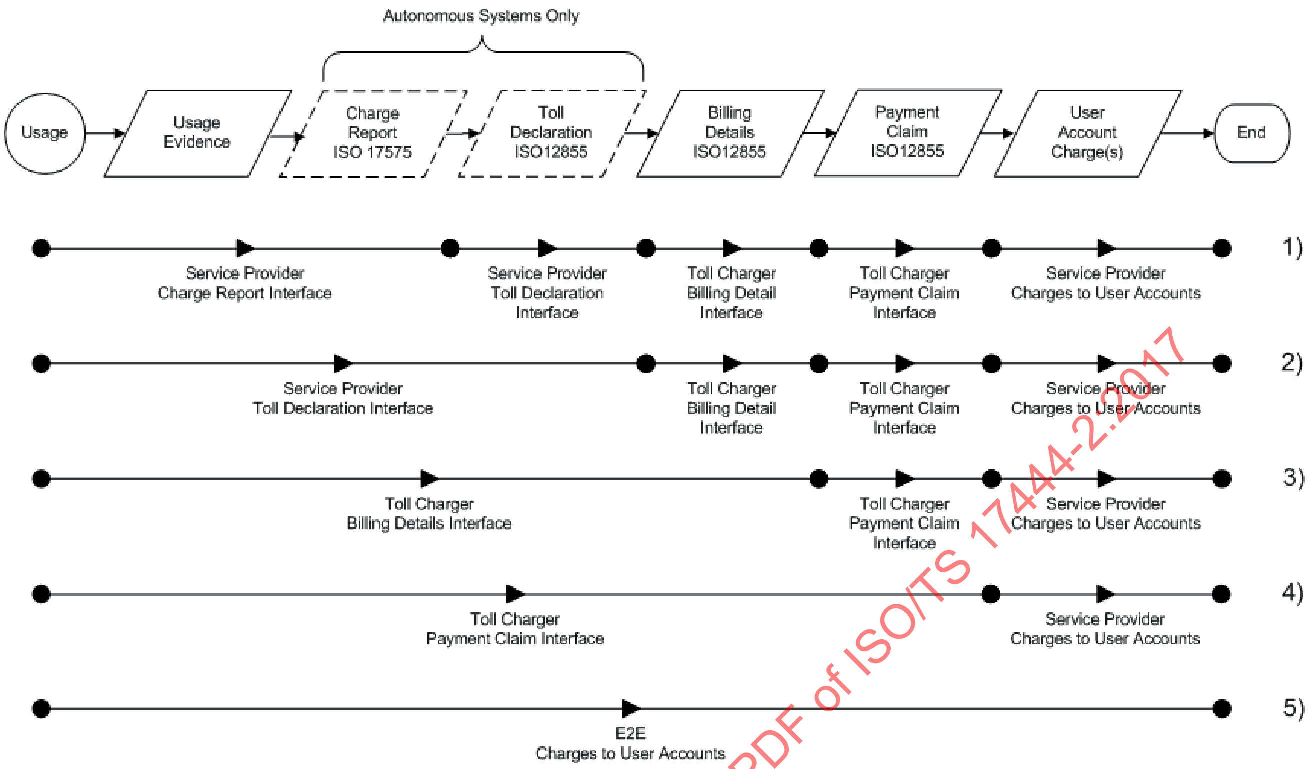


Figure B.1 — Hierarchy of Metric Measurement Observation Points

Annex C
(informative)

Statistical considerations

C.1 Basics

All statistical considerations should reflect existing ISO standards, as described in the Standards catalogue 03.120.30: Application of statistical methods found in http://www.iso.org/standards-catalogue/browse-by-ics.html.

The four most relevant standards for basic statistical analysis of EETS data are ISO 2602, ISO 3534-1, ISO 3534-2 and ISO 11453:1996/Cor.1:1999 (see Bibliography).

More specific aspects and alternative approaches can be found in the catalogue as well.

These ISO standards should be the base reference for all **Standard Operating Procedures** defining statistical characterizations of performance measures. Standard Operating Procedures should not only describe data analysis, but have to define unambiguously all steps of sampling, data ascertainment, data description, and statistical evaluation. When necessary, approaches for validation of measurement systems, identification of sources of bias and of extra variability, etc. have to be provided.

Development of more sophisticated statistical methods should be encouraged, when appropriate, scientifically sound, and in accordance with established ISO standards.

Vocabulary and terms used should be standardized and as precise as possible. Therefore, all documents should use ISO 3534 (all parts) as common reference. In addition, there are glossaries of statistical terms widely accepted internationally. OECD offers an online statistics portal, with a glossary under http://stats.oecd.org/glossary/. From this website, a downloadable version is available as well. Translations into 31 languages are offered by the International Statistical Institute and can be found at https://www.isi-web.org/index.php/publications/glossary-of-statistical-terms.

The following subclauses give a short overview of the basic terminologies, elementary formulae for computing estimates of performance parameters and confidence limits, including example calculations. They should not be understood as a replacement of the standards mentioned above.

C.2 Terminology

Table C.1 — Used terminology

Term	Definition	Examples
Population	The statistical population is the total membership or population or “universe” of a defined class of people, objects or events.	Vehicles on a certain segment of a toll road On-board devices of a certain type
Sample	Selection of elements of an entity and ascertainment of data for different metrics.	Detection of a toll road segment by randomly selected vehicles passing Measurement using the GNSS device of an on-board unit for randomly selected vehicles within a given time interval on a certain toll road
Data analysis	Parameter estimation and inference based on samples	Point and interval estimation, e.g. means, standard deviations and proportions Comparisons of measurements to a reference system Characterization of distributions Selection of certain parametric types of probability distributions and fitting such models to data, e.g. using the normal distribution for characterization of measurements
Interval estimation	Calculation of lower and upper bounds for unknown parameters, assuring a predefined coverage probability of the true value.	Confidence intervals: bounds that guarantee that the true parameter θ is covered by the interval (lower bound, upper bound) with a probability of $1 - \alpha$ Tolerance intervals: bounds that guarantee that a proportion γ of the population is covered by the interval (lower bound, upper bound) with a probability of $1 - \alpha$
Sample size determination	Statistical approaches for assessing the size of a sample, such that demands on precision of inference are met.	Methods related to statistical hypotheses rely on a demand for the power of an α – level significance test being able to detect a certain deviation from the null-hypothesis. Methods related to estimation rely on the precision (interval length) of $1 - \alpha$ confidence intervals.
Probability distributions	Actual probability distributions of the results.	For some measurements of continuous parameters (e.g. distance driven), the results will follow a probability distribution. For ease of analysis, normal distribution is often assumed. Often this is wrong. Some of the influences (e.g. GNSS fixes) of the errors show a behaviour distinctly different from normal distributions. In this case, the calculations must be modified or only considered approximations.

C.3 Point and interval estimation for binary and continuous data

C.3.1 Binary data

These data arise if the measurement of interest is of type yes/no or failed/passed, usually coded as 0/1. Parameter of interest is the probability for observing a 1. A point estimation for samples of independent, identically distributed binary data is straight forward. Such data make up a so called Bernoulli process. An interval estimation can be based on normal approximation or other methods, as appropriate.

Formulae for the normal approximation approach (x = number of successes, n = sample size):

Point estimate: $\hat{p} = x / n$;

Confidence limits: $\hat{p} \pm z_{1-\alpha/2}\sqrt{\hat{p}\left(1-\hat{p}\right)} / n$, denoting the normal percentile by $z_{1-\alpha/2}$.

Alternative methods are described in this document. Especially for probabilities near 0 and 1, more appropriate methods would be likelihood based, transformation based or the so called exact Clopper-Pearson intervals.

NOTE 1 As binary data are discrete by nature, the demanded coverage probability cannot be exhausted in all situations. Exact confidence intervals guarantying coverage of a least 1 – α are wider than alternatives. For large sample sizes, these differences are getting smaller and smaller.

Dependent observations or non-homogeneous data result in over-dispersion, which means, that the variability of the rate estimate is larger than under the Bernoulli sampling scheme. In this case, confidence intervals for the overall success probability should be constructed taking over-dispersion into account.

NOTE 2 For very small underlying probabilities, samples with zero response occur quite often. In this case, one-sided confidence intervals can easily be constructed giving an upper bound.

NOTE 3 A testing-based approach for inference on binary data can be found in ISO 2859 (all parts).

NOTE 4 More sophisticated analysis of binary data is provided when logistic regression models are used, giving the opportunity to model detection probabilities for different times, segments, devices, etc.

C.3.1.1 Binary data examples

C.3.1.1.1 Probability

Estimation of probability/proportion

$\hat{p}$ = number of successful events/overall events

Margin of error

A simple approach for calculating the confidence interval of a binomial proportion is the normal approximation interval shown in Formula (C.1):

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}} \tag{C.1}$$

where

$\hat{p}$ is the proportion of successes in a Bernoulli trial process estimated from the statistical sample;

$z_{1-\alpha/2}$ is the 1 – $\alpha/2$ percentile of a standard normal distribution;

α is the error percentile;

n is the sample size

For example, for a 95 % confidence level, the error (α) is 5 %, so 1 – $\alpha/2$ = 0,975 and $z_{1-\alpha/2}$ = 1,96.

EXAMPLE Simple example with proportions near 50 %: To analyse the proportion of defective units, we draw a sample of 1 000 units; 529 of the 1 000 units are defective.

$$\hat{p} = \frac{529}{1\,000} = 0{,}529$$

The confidence interval is:

$$CI = 0{,}529 \pm 1{,}96\sqrt{\frac{0{,}529\left(1-0{,}529\right)}{1\,000}} == 0{,}529 \pm 0{,}031$$

C.3.1.1.2 Estimation of probabilities near 0 or 100 %

EXAMPLE Significant estimation of the overall detection rate [most important Key Performance Indicator (KPI)] of segment-based toll systems: To analyse the overall detection rate of a toll system, 1 000 detection events are considered as a sample; 995 of the events were successful.

$$\hat{p} = \frac{995}{1\,000} = 0{,}995$$

The confidence interval is:

$$CI = 0{,}995 \pm 1{,}96\sqrt{\frac{0{,}995\left(1-0{,}995\right)}{1\,000}} = 0{,}995 \pm 0{,}004$$

Note that the margin of error in this example is smaller.

C.3.1.1.3 One-sided confidence intervals

Often, it is of more interest if a certain requested quality level is complied with or not than the calculation of an error margin. In this case, one-sided confidence intervals should be used.

For the special case of 0 events among n units under test, an upper 1 – α confidence limit for the underlying response probability is given by Formula (C.2):

$$\hat{p}_{\text{low}} = 1 - \sqrt[n]{\alpha} \qquad \text{(C.2)}$$

For sample sizes of 100, 1 000 and, 10 000, we get the upper 95 % limits: 0,029 5, 0,003 0 and 0,000 3.

C.3.2 Continuous data

C.3.2.1 General

Continuous data can be of various shapes. For toll data, information on the expected value might be most important, as the total of charges equals the number of rides times the expected toll. Inference is supported by the central limit theorem of mathematical statistics. For large samples, approximations based on normal theory can be regarded as reliable. Point and interval estimates assume independent, identically distributed data.

Formulae for the normal approximation approach: (x_1, x_2, … , x_n denoting the sample data; data is assumed to be drawn from a distribution with expected value μ and variance σ^2);

Point estimate: $\hat{\mu} = \bar{x}$ (mean);

Confidence limits: denoting the standard deviation by s. The standard deviation is the point estimate of σ.

Confidence limits for the median, non-parametric approach:

The interval [min(x_1, x_2, … , x_n), max (x_1, x_2, … , x_n)] is a $1 - (1/2)^{n-1}$ confidence interval for the median of the underlying distribution. This holds true for arbitrary distributions. In general, the lower and upper limit can be calculated more appropriately according to a given confidence level, 1 – α, by suitably chosen order statistics.

NOTE 1 For normal data, the distribution would be perfectly characterized by the parameters μ and σ. For general distributions, interpretation of parameter estimates depends on the form of the distribution. A continuous variable is **not** fully characterized by the expected value and variance. At least, questions of symmetry should be addressed.

NOTE 2 The coverage of the above defined interval is exactly equal to 1 – α if the data is normally distributed; otherwise, exact coverage is virtually achieved for large samples ($n > 100$), not too skewed.

NOTE 3 A testing-based approach for inference on continuous data can be found in ISO 3951.

NOTE 4 Instead of the mean, the median or other percentiles of a distribution could be chosen as a parameter of interest. There are non-parametric methods for interval estimates available, either based on empirical quantiles or on re-sampling methods, including the bootstrap approach (see ISO 16269-7).

NOTE 5 Tolerance interval estimation is also possible; however, this is a greater challenge for non-normal data.

NOTE 6 When samples are drawn under different conditions, covariates, useful for describing the distribution, should be recorded as well. Such data offer the chance for more sophisticated regression modelling, including so called random effects models, which could characterize sources of variability.

NOTE 7 Tolerance interval estimation is also possible; however, this is a greater challenge for non-normal data.

C.3.2.2 Continuous data example

The following example uses artificial data and five measurements of a distance: 3,351 km, 3,353 km, 3,349 km, 3,348 km and 3,352 km. The true distance determined by a high precision reference is 3,35 km.

C.3.2.2.1 Normal approximation approach

The mean is $\bar{x} = 3{,}3506$, the standard deviation is $s = 0{,}002\,073\,644$.

The 95 % confidence interval is calculated as

$$3{,}3506 \pm 2{,}776 \times \frac{0{,}002073644}{\sqrt{5}} = 3{,}3506 \pm 0{,}002575 = \left(3{,}348025, 3{,}353175\right)$$

using the t-value $t_{4,\,0{,}975} = 2{,}776$.

This confidence interval covers the “true” value. There seems to be no systematic error in the measurements.

C.3.2.2.2 Non-parametric approach

The interval given by the minimum and maximum values is (3,348, 3,352), virtually identical to the normal approximation. The confidence level is $1 - (1/2)^{5-1} = 93{,}74$ %.

In this case, the non-parametric approach gives almost identical information; however, with a slightly reduced level of confidence.

C.3.3 Dealing with non-normal data

GNSS-based position information is not normally distributed. While the pseudo range measurement itself is (beside the systematic effects) nearly normally distributed, the resulting position solution no longer is.

Modelling of the real distribution based on pseudo range measurements is possible only theoretically. In practice, the complexity of data processing and the missing information about the real methods do not allow the calculations of error distributions.

Thus, the real distribution of GNSS position data depends on the satellite segment, the receiver (and firmware version) and also atmospheric situations.

To overcome this situation, a practical solution is to use a quantile-based approach. All measurements are collected and need to fulfil a maximum deviation condition with a given percentage (e.g. 99 % below 40 m). Confidence levels should be defined using the number of samples.

C.3.4 Quality of reference systems

A reference system needed to determine deviations of the devices under test needs a higher accuracy than the accuracy of the testing device. Typically, values are factors above 3 to 5.

As the costs for reference systems are nearly exponentially linked with the required accuracy, a typical approach would be to choose the quality of the reference based on practical considerations.

Based on available measurement systems with specified accuracy (and acceptable price range), it needs to be determined if the reachable accuracy fulfils the minimum requirements.

Afterwards, the reachable confidence levels can be specified based on the specified accuracy of the device under test and reference system.

Annex D
(informative)

Methods for reducing sample sizes for very high/low probability metrics during the Evaluation Phase

D.1 Rationale

Some metrics are used in requirements establishing values of very low (e.g. 10^{-6}) or very high (e.g. 99,99 %) probability. While this might be possible to measure for the operation phase of a system, assuming that a detailed knowledge of a reference may be established, the effort and cost for tests in the evaluation phase can be high, if those tests are feasible at all.

Typical examples of metrics of those characteristics are those related to false positives and overcharging rate.

If, on the other hand, in order to avoid the mentioned high cost of exhaustive tests, small sample sizes are used for measuring such values close to 0 or 1, the effect of which is detrimental in two ways, either unwanted behaviour of the system is missed because of the resulting inaccuracy or events occurring by chance in the test give a strongly distorted picture. In both cases, the results will be misleading and without practical relevance. In those subsampling scenarios, one can only derive that, given the measured metric, the system is behaving normally or abnormally with a given confidence level but actual value of the metrics cannot be measured.

The first way out of this dilemma is to put the system under artificial stress. By doing so, a higher probability for the events to be detected is created and can be measured with reasonable sample sizes and therefore, with reasonable testing efforts. Extreme care must be taken when deriving the “normal” behaviour of the system; normal, meaning the behaviour under average conditions for the intended application of the system. The extrapolation to that “normal” behaviour requires the identification of some models on which that extrapolation are based. Alternatively, the customer may require a different (less demanding) value for the metric (e.g. different probability) when the system is under those stress situation.

It is important to emphasize that stress can be put on

— the definition of the road network topology, and/or

— the errors affecting the positioning/detection system.

It is also important to understand that those stress conditions can be technology dependent and, therefore, no general rules can be established for their definition, but they have to be defined on a case by case basis as a negotiation process between the customer and system provider.

The first group (road network topology) can be “simply” established as a new input for the system. The second may require the use of simulation. Alternatively, review of designs can also be a valid mechanism if the mentioned design implements particular features to respond to those stress conditions.

The second way out of the dilemma (in isolation or in combination with the previous solution) is to increase the amount of testing data without largely increasing the testing effort and hence, costs. This may be implemented, in particular in the case of measuring false positives, by artificially adding new infrastructure elements to be charged. In any case, the number of trips have to be big enough to be statistically significant of the different errors affecting the system.

NOTE If we need n trips to have the required statistical significance and we introduce m segments to be detected, the resulting number of trips could be around n/m. But this number should still be large enough, i.e. $m=n$, resulting in a single trip would not be a valid solution.

This derivation is founded on extrapolation of the system behaviour to less demanding circumstances. For ensuring that this process remains fair and reliable, the underlying model for the system behaviour is of critical importance. It is also recommended to subject the reference system to the same test scenario. This strategy can help to identify errors in the model or inherent limitations of the technology used.

While the goal is to define **black box** and **technology independent** tests, it is anticipated that relaxation of these two objectives may facilitate the definition of those tests (see Table D.1).

Table D.1 — Applicability of the use of real, emulated and artificial Charge Objects in metric measurements

		Trips		
		Real ($n >> 1/p$)	Real ($n < 1/p$)	Simulated (worst case nuisance variables)
Charge Objects	Real	Possible Very expensive	Not useful	Possible, but only valid if “GNSS-only”. Model needed to extrapolate
	Emulated	Not useful	Possible Direct metric	Possible, but only valid if “GNSS-only”. Model needed to extrapolate
	Artificial (worst cases)	Not useful	Possible Model needed to extrapolate	Possible, but only valid if “GNSS-only”. Model needed to extrapolate

D.2 Identification of potential methods

D.2.1 Method 1: Based on PC and PU and having as reference the real charge object detection

Trips for PC1 and PC2 selected to be in the nearest feasible proximity to zone/segment for which the charge object detection is defined. For PU1, only the subset of users who travel in the proximity of the mentioned infrastructure is selected.

The number of trips to be considered has to be in the order of 10 divided by the target probability defined in the requirement related to that metric (what makes this method quite inefficient). Alternatively, if metric does not need to be measured but only to define if the system fulfils a requirement with a certain level of confidence, the number of allowed false positives may be established as a function of the number of trips and the mentioned level of confidence.

$$\text{TSP False Positives probability} = \frac{\text{Number of trips that provoke a false positive}}{\text{Total number of trips}}$$

This method is the “brute force” approach, which will lead to large sample sizes or to misleading results, as described in the last subclause.

D.2.2 Method 2: Based on PC and PU and having as reference artificially defined charge object detection

Emulated chargeable infrastructure is defined in the nearest proximity to the travelled trips. This can be based on the existing network (and then defined charging infrastructure has to be a real road) or