An Analytical Evaluation for Hazardous Failure Rate in a Satellite-based Train Positioning System with reference to the ERTMS Train Control Systems

A. Neri¹, A. Filip², F. Rispoli³, A.M. Vegni¹
¹RADIOLABS, Rome Italy, {alessandro.neri, annamaria.vegni}@radiolabs.it
²University of Pardubice, Faculty of Electrical Engineering and Informatics, Studentská 95, 532 10 Pardubice, Czech Republic, filip.ales@azd.cz
³Ansaldo STS S.p.A, Genoa Italy, francesco.rispoli@ansaldo-sts.com

BIOGRAPHIES

Alessandro NERI is Full Professor in Telecommunications. In 1977 he received the Doctoral Degree in Electronic Engineering from the University of Rome “Sapienza”. In 1978 he joined the Research and Development Department of Contraves Italiana S.p.A. where he gained a specific expertise in the field of radar signal processing and in applied detection and estimation theory, becoming the chief of the advanced systems group. In 1987 he joined the INFOCOM Department of the University of Rome “Sapienza” as Associate Professor in Signal and Information Theory. In November 1992 he joined the Electronic Engineering Department of the University of Roma TRE as Associate Professor in Electrical Communications, and became full professor in Telecommunications in September 2001. His research activity has mainly been focused on information theory, signal theory, and signal and image processing and their applications to both telecommunications systems and remote sensing.

Since December 2008, Prof. Neri is the President of the RadioLabs Consortium, a non-profit Consortium created in 2001 to promote tight cooperation on applied research programs between universities and industries.

Aleš FILIP is an Associate Professor of Electrical Engineering at University Pardubice and head of the AŽD-Laboratory of Intelligent Systems (LIS) in Pardubice, Czech Republic. In 1983 he received the Ing. degree in Technical Cybernetics from Technical University of Brno, Czech Republic, and the Ph.D. degree in Electrical Engineering from Czech Technical University in Prague in 1988. From April 1993 to March 1995 he was a research associate in the Optoelectronic Laboratory, University of Tokyo, Japan. In 1995 he joined Czech Railways and in 1996 he initiated foundation of LIS with the intention to perform research in the field of signalling and train control based on GPS and advanced communications. From January 2012 LIS is a part of AŽD Praha Company. His research interests include hazard and risk analysis of EGNOS SoL Service and local GNSS augmentations focused on applications in railway signaling. The purpose of his current work is to support approval and certification of advanced railway safety-related systems. Dr. Filip also gives lectures on Safety and Dependability of Electronic Systems.

Francesco RISPOLI is with Ansaldo STS since May 2011 as responsible of the Satellite and Telecommunication projects in the Innovation and Competitiveness department. Previously he has been with Telespazio (2005-2011) first as responsible for the Value Added Services and by 2008 as Chief of New Initiatives for exploiting innovative satellite solutions in the field of railway applications (GNSS and Internet for on board entertainment) telemedicine services through the TelBios company, emergency telecommunications. By 1983 to 2005 he was with Alenia Spazio (now Thales Alenia Space) where he served various positions as Head of R&D programs, Head of Marketing, Vice President of Multimedia Business Unit, General Manager of EuroSkyWay. He has been a co-founder of the TelBios company and member of its board of directors. He started his career in 1980 with Contraves as technical engineer in the radar antennas department. He received the Doctoral Degree in Electronic Engineering from the Polytechnic of Turin in 1978 and a Master in Applied Electromagnetism from the University La Sapienza in Roma in 1979 winning the first prize for his thesis.

Anna Maria VEGNI is Assistant Professor in Telecommunications at the Department of Applied Electronics of University of Roma Tre, Rome, Italy. She received the Ph.D. degree in Biomedical Engineering, Electromagnetics and Telecommunications from University of Roma TRE in 2010. From May 2009 to October 2009, she was a visiting scholar in the Multimedia Communication Laboratory, at the Department of Electrical and Computer Engineering, Boston University, Boston, MA. Her research activity is focusing on vehicular networking, indoor and outdoor localization, GNSS and Visible Light Communications.
Since 2011, she is in charge of Telecommunications Networks Laboratory course at Roma Tre University.

ABSTRACT
ERTMS-ETCS developed in Europe for high speed lines is de facto the railways standard train control system being adopted in most new lines and major upgrades. Satellite localization and IP-based TLC have been identified as the new technologies to be adopted by the ERTMS-ETCS since they will reduce the investments and maintenance costs on track-side components to provide cost-effective solutions for the increasing demand of the local, regional and new freight lines in rural and environmental critical areas.

A major challenge in the adoption of GNSS-based Location Determination Systems (LDS) is to guarantee the same Hazardous Failure Rate (HFR) achieved with the mechanical odometry systems which has to be less than $10^{-9}$ during 1 hour of operation according to the CENELEC SIL 4 requirements. To reach this target, we have investigated an LDS architecture, based on (i) a multi-constellation space segment to increase both the accuracy and the number of satellites in visibility, (ii) the deployment of a proper Track Area Augmentation and Integrity Monitoring Network with very high availability, and (iii) an independent on-board capability to further mitigate GNSS errors and autonomously assess the GNSS location integrity when augmentation data are unavailable.

In this paper we exploit the railway characteristics for an analytical evaluation of the performance of Integrity Monitoring and, Detection and Exclusion algorithms, expressed in terms of impact on Protection Level (PL). Since LDS may operate in several operational modes, the analytical model, as well the simulation results, will describe the LDS behavior in each different case outlining the advantage of using the multi-constellation capability respect to the single GPS constellation.

1. INTRODUCTION
Modern signaling systems play a major role to provide safety net to prevent accidents due to human errors, and the management of railway operation to maximize the use of scarce resources and optimize the investment in infrastructure. One of the major breakthroughs has been the introduction of the radio-based signaling European train control system, named ERTMS-ETCS that, with more than 37,000 km in the world contracted to be equipped, is becoming a global standard, thanks also to the introduction of satellite positioning.

The most promising way to reduce the costs associated with the implementation and maintenance of ERTMS-ETCS, which by fact is the key factor limiting its extensive adoption, is to make use of GNSS and IP-based communications. Thus, satellite-based localization systems and IP-based communications have been identified as the key technologies to be adopted by the ERTMS-ETCS in its evolving path to provide cost-effective solutions to satisfy the needs of the local, regional and new freight lines [1]. However, the main challenge in the adoption of GNSS-based Location Determination Systems (LDS) is constituted by the information integrity imposed by the stringent safety requirements of the CENELEC railways specifications, [2]-[5]. The compliance with this requirement will guarantee the adoption of the GNSS localization technique into the ERTMS-ETCS “ecosystem” without impacting on its performance, safety characteristics and architecture.

To this end, the Hazardous Failure Rate (HFR) during 1 hour of operation shall be less than $10^{-9}$ for SIL 4 compliant systems. It implies that the probability that the magnitude of the error of the position provided by the LDS exceeds the Protection Level (PL), has to be less than $2\cdot10^{-13}$ conditioned to the fact that this event has not been detected by the Integrity Monitoring algorithm, with the additional constraint on the maximum false alarm probability, [6]-[13].

In addition to adoption of state-of-the-art GNSS technologies for mitigating the impact of potential error sources (i.e., multi-frequency receivers, multipath reduction processing, SBAS clock and ephemerides errors compensations), the fulfillment of this challenging objective imposes the implementation of an Integrity Monitoring subsystem, able to fast detect Hazardous Misleading Information produced by failures and malfunctions of both ground and space segments. Although such situation looks similar to the one arising in aviation applications, the HFR for railway applications is two orders of magnitude lower than the aviation Integrity Risk, due to a higher operational time interval (1 hour).

To reach this target, we propose an LDS architecture, based on (i) a multi-constellation capability to increase both the accuracy and the number of satellites in visibility, (ii) the deployment of a Track Area Augmentation and Integrity Monitoring Network with very high availability, and (iii) an independent on-board capability to further mitigate GNSS errors and autonomously assess the GNSS location integrity when augmentation data are unavailable.

The Augmentation Network includes Ranging & Integrity Monitoring Reference Stations (RIM RS), co-located with selected communications base stations for the purpose of integrity monitoring, accuracy improvement of satellite-based position, and providing correction to mobile receivers. Each reference station has an LDS Safety Server, providing correction services and detecting systematic satellite faults. Finally, to enhance the systematic satellite fault detection capabilities, the outputs from reference stations are jointly processed by a Track Area LDS Safety (TALS) server.

Protection levels are generally evaluated assuming only one faulty satellite at any time. On the other hand, deployment of new constellations significantly increases the number of satellites in view. Therefore, the probability of occurrence of multiple simultaneous faults is no longer negligible. Therefore, in this paper three different events have been considered for the evaluation of the Integrity...
Risk i.e., (i) fault-free case, (ii) individual satellite faults, and (iii) correlated simultaneous satellite faults. In the first case, when all satellites are healthy, a hazardous failure can still arise due to large random errors produced by multipath, receiver thermal noise or tropospheric incremental delays, besides their small probability of occurrence. In the second case, individual satellite faults may arise in both the space and the ground segments; historical data indicate an average number of faults for GPS constellation of about 3 per year.

Concerning the effects produced by a satellite fault on the estimation error, we observe that whenever the receiver operates in a differential operational mode, errors on ephemeris and satellite clocks are essentially compensated. Nevertheless, latency and spatial distribution of RAIMs may have impact on the variance and bias of the position error.

To account for the impact of the augmentation data availability on the achievable PL, in the analytical model we consider an On-Board LDS algorithm able to work on four operational modes for fixes computation, such as (i) GNSS augmented by the dedicated Track Area Augmentation and Integrity Monitoring Network, (ii) GNSS augmented by SBAS from direct Signal In Space (SIS), (iii) GNSS augmented by SBAS data retransmitted over Train Signaling Network, and (iv) stand alone GNSS.

In order to detect a possible fault, the health of each satellite is monitored by analyzing the statistical behavior of the observed pseudoranges.

Considering that for given PL and false alarm probability the HFR is a unimodal function of the faulty satellite pseudorange error magnitude, and probabilistic models for SIS faults are not sufficiently reliable, we adopt a MINIMAX procedure. In essence, PLs and alarm levels are set in accordance to the worst-case satellite faults, as far as their magnitude is concerned.

This paper is organized as follows. We first introduce the proposed LDS architecture in Sect. 2. Then the procedure developed to verify integrity at system switch on is described in Sec.3. In Sec. 4, the analytical method for performance assessment is illustrated. In Sect. 5 simulation results are reported. Finally conclusions are drawn at the end of the paper.

2. LDS ARCHITECTURE
The LDS solution takes into account the need to have a (SIL 4) satellite-based train positioning system usable on the world-wide market according to the user needs. This solution is based on four pillars:

1) Multi-constellation GNSS capability, exploiting existing constellations (GPS, GLONASS) and new ones like GALILEO, planned to start services by 2015;

2) Use of SBASs, and in particular EGNOS in Europe, augmentation data for both accuracy and precision increase and integrity monitoring;

3) Deployment of a Track Area Augmentation and Integrity Monitoring Network, co-located with some TLC base stations in areas out of SBASs footprints;

4) Independent on-board capability to further mitigate GNSS errors and autonomously assess the GNSS location integrity.

The GNSS LDS On Board Unit (OBU) is a self-contained unit connected to the ATP, the antennas, and the locomotive power supply. It comprises a dual-path GNSS receiver integrated with a SIL 4 processor board. In essence the Track Area Augmentation and Integrity Monitoring Network plays a role similar to the EGNOS Range and Integrity Monitoring subsystem. In fact, processing of satellite signals received at known locations allows to estimate the error sources affecting train positioning and to detect eventual GNSS faults. This architectural entity has been introduced to take into account the recent decision to eliminate the SOL features from the Galileo system.

The major difference with respect to EGNOS consists in a denser spatial deployment of the RIM RSs, compensating for milder requirements (and lower cost) on the GNSS receiver clocks and in the use of the wireless network employed for train signaling even for augmentation data distribution. Since RIM RSs adopt multi-constellation receivers, integrity is assessed for any visible constellation.

To guarantee high availability each RIM RS is equipped with an LDS Safety Server based on the same SIL 4 platform as the mobile LDS OBUs, but configured to provide correction services and detect systemic satellite faults.

The LDS Safety Server monitors for each epoch the measured pseudoranges and compare them with the nominal values corresponding to the known antennas locations.

To enhance the systemic satellite fault detection capabilities, as well as to detect eventual RIM RS faults, their outputs are jointly processed by a Track Area LDS Safety (TALS) server. Such an architecture allows improving the correction function of classical differential GNSSs and mitigating the risk of failure relevant to the GNSS reference stations.

Correction factors and more in general augmentation data are supplied to the mobile LDS OBUs in the service region by means of a safety-critical protocol layered on the same radio communications network used for train control. This allows LDS to achieve a high degree of accuracy and integrity without depending on commercial off-the-shelf systems of unknown integrity.

RIM RSs are deployed in such a way that they share sources of systemic errors with the GNSS receivers on board of the locomotives operating in the regions, such as incremental delays caused by atmospheric conditions and ephemerides and clock errors of visible satellites. Whereas commercial differential correction systems are not certified to SIL 4, the LDS architecture provides safety-critical corrections for LDS. In the event that satellite errors prevent computation of a reliable
correction, the TALS Server reports that status to each LDS OBU, which falls back to running on tachometer. LDS OBU corrections are provided by the TALS server. Nevertheless, for higher availability, in case of a failure of the TALS server, each LDS OBU will employ the data broadcasted by the nearest healthy reference station. RIM RS baselines are selected in order to guarantee that the expected differences in corrections between reference stations will be small enough to avoid abrupt transitions. In the event that the LDS OBU cannot communicate neither with the TALS server nor with any LDS Safety Server for a period of time, it will switch to the autonomous mode and it will report loss of TALS assisted augmentation data to the ATP.

The integrity of the information provided by the LDS OBU is continuously monitored by the On Board LDS Range and Autonomous Integrity Monitoring (RAIM) LDS algorithm, according to the satellite healthy status reports provided by the TALS and LDS Safety Servers. In the event that LDS RAIM will detect a Hazardous Misleading Information (HMI), it will report the event to the ATP which will run on tachometer until communications are re-established.

To ease the integration with existing ERTMS train control system ATPs that make use of LDSs that reset odometer readings when the train is passing over RF transponders, named balises, deployed at known locations along the railway track, in addition to train position and velocity, the GNSS LDS OBU also provides a pulse notifying the passage over the so called virtual balise. As a matter of fact the Dilution of Precision (DoP) that quantifies the error in estimating the current train location, strictly depends on the number of visible satellites as well as on their line of sight geometry. However, using a multiconstellation receiver the number of visible satellites is, in general, highly redundant. As a consequence, DoP decreases. At the same time the probability of detecting an anomalous behavior from any satellite increases, while discarding a failed satellite has no practical impact on DoP and LDS service availability. Moreover the huge number of satellites allows sustaining a relatively high probability of declaring as faulty any healthy satellite with no practical DoP and system availability degradations. As a consequence a higher detection probability of faulty satellites can be achieved.

The availability of EGNOS (and in general SBAS) SIS simplifies the architecture by preserving the performance. In synthesis, in order to achieve the SIL 4 requirement, we deploy the following primary checks on GNSS position determination:

1. Duplicated position computation from satellite signals using certified software and diverse hardware;
2. Regional RIM RSs, based on the same SIL 4 platform, validating the functioning of the satellites-in-space component of GNSS-based location determination and providing corrections that are both reliable and accurate;
3. A correlation with a railway track database within LDS;
4. A correlation with tachometer readings by the ATP.

2.1 RIM RS architecture
Each RIM RS is equipped with (i) a multiconstellation GNSS unit, able to track GPS, GLONASS, Galileo, Compass, and SBAS satellite signals, (ii) a package of pressure, temperature, and humidity sensors for tropospheric delay prediction, (iii) an LDS Safety Server that monitors the received data in order to detect systemic satellite faults and to provide corrections to the LDS on board of a train.

Each RIM station periodically sends the samples of the signals received from the visible satellites to the TALS server. At this aim it makes use of the same backbone employed for interconnecting the co-located communication base stations.

RIM station antennas are georeferenced during the installation phase, post-processing long enough time series of the observed pseudoranges together with precise orbits and satellite clock offsets. In addition antenna alignment data needed for phase center shifts corrections are also measured.

For autonomous estimation of the ionospheric delay in areas not served by SBAS satellites, multi frequency GNSS receivers are employed. The tropospheric incremental delay is estimated by means of a numerical weather model (e.g. Saastamoinen model) that makes use of the local pressure, humidity and temperature measured by the meteo sensors. This allows estimating the hydrostatic (dry) component of the atmosphere that roughly accounts for the 90% of the tropospheric path delay. The residual tropospheric path delay component, essentially accounting for the wet part, is part of the pseudorange residual correction.

Statistics of the differences between measured and nominal satellite pseudoranges (i.e., pseudorange residuals) are processed by the Fault Detection and Integrity Evaluation algorithm on a per snapshot basis for fast detection of satellite faults.

RIM stations are deployed along the railway track at a distance less than the decorrelation distance for the tropospheric vertical delay. For instance, in the simulated railway corridor whose length was about 300 km, 5 RIM stations have been employed.

2.2. TALS server architecture
The TALS server processes the iono-free and dry tropo-free reduced pseudoranges provided by the LDS Safety Servers and estimates the actual status error of each visible satellite consisting of both ephemeris position errors and clock offsets, needed for both integrity assessment and augmentation support for LDS On Board Units.

To decouple the ephemeris and satellite clock offset estimates, a single difference technique based on the processing of the differences between pseudo ranges of the same satellite measured by the different RIMs is employed. The differential technique allows removing the errors introduced by the RIM GNSS receiver clock offsets.
affecting the LDS Safety Server pseudoranges. The ephemeris estimator adopts a snapshot minimum variance method combined with a temporal filter employing a satellite kinematic model. To estimate the satellite clock error, a snapshot weighted least square estimator cascaded with a filter for error rate estimation is employed. Pseudorange time series and precise orbits and satellite clock offsets are used for maximum likelihood estimation of the long term RIM RS clock offsets.

As an option, the LDS Safety Server may estimate the wet component of the tropospheric incremental delay by means of a numerical weather model that makes use of tropospheric weather fields published by national and international centers.

To increase overall availability, the TALS performs a joint analysis of the corrections provided by each LDS Safety Server, in order to detect both satellite and RIM stations failures.

In addition to sanity checks on satellite navigation data, the TALS server computes protection level equations for the ephemeris and satellite clock failures, based on bounds on the magnitude of ephemeris failures that could be potentially undetected.

Although in principle safe use of data provided by unhealthy satellites is possible, considering that the number of visible satellites along track (with elevation > 10°) in general greatly exceeds the amount of satellites needed for train location determination, a Fault Detection and Exclusion algorithm is implemented in the TALS Server. In essence the Fault Detection and Exclusion algorithm monitors for each satellite the differences between the RIM observed pseudoranges and the predicted ones. For resiliency against noise, a rank order statistics is applied to the pseudorange residual time series.

At this aim, an adaptive scheme that for each epoch excludes from augmentation the worst satellites of each constellation (estimated on the basis of residual statistics and on the sensitivity of the LDS position error with respect to pseudoranges errors of each satellite) is employed. In this way, the conditional probability of HMI with respect to a satellite failure is totally negligible, because worst satellites are excluded even if their residual errors are comparable to those of the healthy ones. In addition, faults in the GNSS overall system can be detected when joint satellite faults are observed.

Moreover, joint use of multiple constellations allows mitigating the effects produced by the unavailability of an entire GNSS, since it will only affects the LDS location accuracy.

With respect to simpler architectures employing the closest RIM to directly provide to the LDS OBU differential GNSS corrections, joint processing of the RIMs data by the TALS server allows to detect faults in the RIM stations and to exclude the faulty ones, then increasing both overall system integrity and availability. RIM RS fault detection is similar to the satellite fault detection, being performed by processing the observed pseudoranges. Joint processing of RIM data produces graceful degradation of integrity risk in presence of RIM faults.

Optionally, the TALS Server may employ ephemeris and clock offset corrections provided by SBAS and precise orbits and satellite clock (delayed) estimates provided by publicly available sources (IGS) to enhance detection of RIM faults and mitigation of the effects of the RIM receiver clock offset fluctuations, with a positive impact on both system integrity and PL.

The Augmentation data that the TALS server sends to the LDS OBU includes:
- Slow corrections for long-term variations in the ephemeris errors and satellite clock errors;
- Fast corrections for rapid variations in the ephemeris and satellite clock errors;
- Ionospheric Vertical Delays related to grids distributed along train tracks;
- Tropospheric Vertical Delays related to grids distributed along train tracks;
- Forecast of satellite visibility and expected DoP;
- GNSS integrity and individual satellite health status.

### 2.3 LDS OBU architecture

The LDS algorithm is able to determine the train location, assuming explicitly the fact that the train location is constrained to lie on railway track. From a mathematical point of view, this constraint can be imposed by observing that the train location at time $t_i$ is completely determined by the knowledge of the travelled distance (i.e., by the curvilinear abscissa $s$ defined on the georeferenced railway track).

Let $s(k)$ be the curvilinear abscissa of any train reference point (e.g. the center of the antenna of the GNSS receiver of the LDS OBU) when the $k$-th GNSS snapshot is measured. Then, observing that the Cartesian coordinates of said point are described by the parametric equations

$$X_{OBU}(t) = X_{OBU}^{OBU}\{s(t)\} = \left[ x_{OBU}^{OBU}\{s(t)\} \ x_{OBU}^{OBU}\{s(t)\} \ x_{OBU}^{OBU}\{s(t)\} \right]^T \quad (1)$$

the pseudoranges measured by the GNSS receiver can be directly expressed in term of the unknown curvilinear abscissa.

In fact, the pseudo-range $\rho_i(k)$ of the $i$-th satellite measured by the OBU GNSS receiver can be written as follows

$$\rho_i(k) = \left[ X_{Sat}^{Sat}\left[T_{Sat}^{Sat}(k)\right] - X_{OBU}^{OBU}\left[s(T_{Sat}^{OBU}(k))\right] \right] + c\Delta \tau_{i}^{ion}(k) + c\Delta \tau_{i}^{trop}(k) + c\delta \tau_{i}^{OBU}(k) + c\delta \tau_{i}^{OBU}(k) - c\delta \tau_{i}^{Sat}(k), \quad (2)$$

where we define:
- $T_{Sat}^{Sat}(k)$ as the time instant on which the signal of the $k$-th snapshot is transmitted from the $i$-th satellite;
- $X_{Sat}^{Sat}\left[T_{Sat}^{Sat}(k)\right]$ as the coordinate vector of the $i$-th satellite at time $T_{Sat}^{Sat}(k)$;
- $\Delta \tau_{i}^{ion}(k)$ as the ionospheric incremental delay along the path from the $i$-th satellite to the GNSS receiver for the $k$-th snapshot w.r.t. the neutral atmosphere;
\[ \Delta \tau_{i}^{\text{io}}(k) \text{ as the tropospheric incremental delay along the path from the } i\text{-th satellite to the GNSS receiver for the } k\text{-th snapshot w.r.t. the neutral atmosphere; } \\
\delta \tau_{i}^{\text{sat}}(k) \text{ as the offset of the } i\text{-th satellite clock for the } k\text{-th snapshot; } \\
T_{i}^{\text{OBU}}(k) \text{ as the time instant of reception by the OBU GNSS receiver of the signal of the } k\text{-th snapshot transmitted by the } i\text{-th satellite; } \\
\delta t_{i}^{\text{OBU}}(k) \text{ as the OBU receiver clock offset; } \\
n_{i}^{\text{OBU}}(k) \text{ as the error of the time of arrival estimation algorithm generated by the noise of the GNSS receiver, with variance dependent on the received signal-to-noise and eventual multipath effects.} \]

Since the ionospheric incremental delays are estimated by the pair of pseudoranges at frequencies L1 and L2, the tropospheric incremental delay is estimated in the basis of the regional meteorological data measured by the RIM meteo stations, and the satellite location and clock offset are predicted on the basis of ephemers, almanac data, and pseudoranges measured by the RIM RUs, the pseudoranges can be rewritten as
\[
\rho_{i}(k) = \left[ X_{i}^{\text{io}} \left[ T_{i}^{\text{OBU}}(k) \right] \right] - X_{i}^{\text{OBU}} \left[ s(T_{i}^{\text{OBU}}(k)) \right] + c \Delta \tau_{i}^{\text{io}}(k) + c \delta \tau_{i}^{\text{sat}}(k) + c \delta t_{i}^{\text{OBU}}(k) + v_{i}(k), \quad (3)
\]
where the symbol ^ denotes the estimated quantities and \( v_{i}(k) \) is the equivalent receiver noise accounting for any residual estimation error in addition to multipath and thermal receiver noise.

As detailed in Appendix A, the set of non-linear equations may be solved w.r.t. the curvilinear abscissa \( s(T_{i}^{\text{OBU}}(k)) \)

and the receiver clock offset \( \delta t_{i}^{\text{OBU}}(k) \) by means of a weighted least square, iterative numerical procedure that accounts for the different statistics of the error of the time of arrival estimates related to satellites of different constellations.

Thus, in principle pseudo-ranges related to two satellites only are required to determine the train location.

The multisensory LDS integrates the information provided by the GNSS LDS subsystem with localization data provided by classical Odometric subsystem, denoted in the following as ODO LDS, that processes tachometer data as well as inertial sensor packages.

The assessment of the integrity of the localization estimated by GNSS LDS and ODO LDS, is further enhanced by correlating the estimated train dynamics with the expected location and velocity predicted on the basis of actual speed and breakers commands.

The root mean square error \( \sigma_{\text{av}} \) of the estimate of the curvilinear abscissa \( s \) depends on both the number of visible satellites and Line Of Sight geometry as well as on the receiver operational mode.

At this aim we observe that the following operational modes (OP) for fixes computation are supported by the LDS OBU:

\[ \text{OP 1. GNSS augmented by TALS RIMs; } \]
\[ \text{OP 2. GNSS augmented by SBAS (e.g. EGNOS, WAAS) from direct Signal In Space (SIS); } \]
\[ \text{OP 3. GNSSS augmented by SBAS data retransmitted over Train Signaling Network; } \]
\[ \text{OP 4. Stand alone GNSS.} \]

As a consequence, the error budget may vary with the operational mode.

3. LDS INITIALIZATION

Before being put into operation within a railway signaling system, the GNSS LDS has to be initialized. The initialization mainly includes the independent check of the entire LDS unit (HW and SW functionality) and position determination of the first virtual balise or balise group with the required SIL, [1], [8].

In order to reduce the probability that any Misleading Information caused by hardware/software failure would produce a dangerous situation, self-check circuitry and signal monitoring by an independent diagnostics unit are incorporated in the architecture.

LDS diagnostic methods play an important role during the initialization process. In this section we introduce a novel initialization technique of LDS based on GNSS local augmentation network (i.e., Au-network) or regional GNSS augmentation (i.e., SBAS).

The LDS unit initialization is based on a known geometry of track-side RIM RSs. The known RIM RSs geometry is used for the initial LDS position integrity monitoring. The LDS initialization can effectively profit from the GNSS multi-constellation and its relevant differential methods (i.e., DGPS, DGGLONASS, DGalileo) in such a way that it is possible to build the LDS unit compliant with SIL 4 by means of non-safety (SIL 0) COTS GNSS receivers.

The proposed LDS initialization consists of the two following phases i.e., (i) the First Phase i.e., the initialization in stand-still mode, and (ii) the Second Phase i.e., the initialization in motion mode.

3.1 FIRST PHASE

In the First Phase the LDS operates in stand-still mode. The purpose of this phase is to detect all possible faults during the LDS initialization excepting observation errors due to local effects (multipath, EMI, etc.). It is based on a priori known and continuously monitored geometry of the stationary RIM station network. In this phase all faults/errors in the chain behind the pseudorange (PR) measurement, e.g. due to signal processing, receiver SW and HW faults, etc., can be detected by means of this method. However, PR observation errors due to local effects may not be detected during the First Phase. Therefore, the Second Phase with LDS in motion must be performed. The First Phase consists of two steps i.e., (i) Step 1, and (ii) Step 2.

In the Step 1 the LDS position is determined by means of the Inverted Differential GNSS (IDGNSS) method, as depicted in Figure 1. It means that position of LDS GNSS receiver (i.e. rover) is computed at each of RIMs. The pseudoranges from the LDS GNSS Rover(s) are transmitted to the individual RIMs. The LDS position can be also optionally provided directly by the redundant LDS GNSS receivers. Receiver Autonomous Integrity Monitor
(RAIM) in the LDS unit can reduce potential effects due to multipath and EMI. Nevertheless, errors in PR observations due to excessive Local Effects can remain undetected and therefore Second Phase of the initialization in motion is needed.

In the Step 2 the independent check of the LDS GNSS initial position by means of the LDS module in the base station mode is performed, as depicted in Figure 2. The LDS differential corrections are sent from the LDS base to the individual RIMs. The position of each of RIMs is computed by means of the LDS differential corrections. The Au-network geometry (position of RIMs) is independently checked by means of DGPS, DGLONASS and DGalileo receivers.

3.2 SECOND PHASE
In this phase the LDS works in motion mode. The purpose is to detect mainly pseudorange observation errors due to local effects such as multipath, EMI, etc. Diagnostic means such as multiconstellation RAIM, odometry (provision of velocity), Inertial Measurement Unit (IMU), kinematic model of train movement and track database are used for detection of excessive errors in position in this phase.

After the Second Phase is successfully completed, then the LDS initialization is finished. In the opposite case the initialization process should be repeated. The preliminary safety assessment of the LDS initialization during the First and Second Phases is depicted in Figure 3 and Figure 4, respectively.

The independence of the diagnostics is based on the idea that the correct geometry of the RIM network can be a priory predetermined with very high accuracy (cm level) and very low failure rate (cca 10^{-12}/hour is estimated).

Furthermore, the RIM geometry can be continuously, e.g. every 1 s, checked. GNSS receiver HW and SW diversity can mitigate excessive errors in pseudorange measurements due to potential receiver failure. Thus only one “correct” Au-network /RIM geometry exists and this fact is used for the justification of the mutual independency among the position data provided by receivers for different satellite systems (e.g., GPS, GLONASS, Galileo), as shown in Figure 3. The safety assessment depicted in Figure 3 is rather conservative since currently available COTS GNSS receivers have much larger Mean Time Between Failures (MTBF) than 10,000 hours.

The FTA diagram for the Second Phase of the initialization in motion mode is depicted in Figure 4. Detection of infrastructure elements such as switches or track curves by means of additional sensors integrated within the Inertial Measurement Unit (IMU) can significantly improve the LDS integrity. While the odometry / IMU and RAIM are in the competitive relationship in the Second Phase of initialization, then in the LDS operational phase the relative position provided by odometry improves availability of GNSS. The preliminary safety assessment of the proposed LDS initialization based on Au-network implies that this new...
technique is feasible for implementation of LDS subsystems compliant with the highest SIL 4.

![Figure 5. Fault Tree.](image)

4. INTEGRITY ASSESSMENT MODEL

Let us recall that, by definition, the LDS is said to provide an Hazardous Misleading Information (HMI) whenever the magnitude of the position error exceeds the Protection Level (PL), while this event is not detected by the Integrity Monitoring algorithm. The HMI Rate is then defined as the probability of providing at least an HMI during a predefined time interval (equal to 1 hour in railway applications). A quantitative evaluation of the safety integrity level can be performed on the basis of the dangerous fault tree shown in Figure 5.

More in detail, neglecting the eventual further gain obtained by multisensor fusion in integrity monitoring and denoting with:

- $R_{DEF}^{GNSS\_LDS}$ the rate of undetected dangerous failures of the GNSS LDS subsystem, excluding the GNSS receivers,
- $R_{DEF}^{GNSS\_RX}$ the rate of undetected dangerous failures of the n-th GNSS receiver
- $R_{DEF}^{GNSS}$ the rate of undetected dangerous failures of the whole GNSS LDS subsystem, including the GNSS receivers,
- $R_{DEF}^{ODO\_LDS}$ the rate of undetected dangerous failures of the ODO LDS subsystem, excluding INS sensors and tachometers,
- $R_{DEF}^{INS}$ the rate of undetected dangerous failures of the n-th INS sensor package,
- $R_{DEF}^{TACHO}$ the rate of undetected dangerous failures of the n-th tachometer package,

the rate of undetected dangerous failures of the ODO LDS subsystem, including INS sensors and tachometers,

- $R_{DEF}^{GNSS\_HMI}$ the rate of hazardous Misleading Information of the GNSS LDS subsystem,
- $R_{DEF}^{ODO\_HMI}$ the rate of hazardous Misleading Information of the ODO LDS subsystem,
- $R_{DEF}^{MLDS}$ the rate of undetected dangerous failures of the MultiSensor LDS subsystem,
- $R_{DEF}^{DRAG}$ the rate of undetected dangerous failures of the INDEPENDENT DIAGNOSTICS subsystem,
- $N_{GNSS\_RX}$ the number of redundant GNSS receivers,
- $N_{INS}$ the number of redundant INS packages,
- $N_{TACHO}$ the (independent) decision rate of the GNSS LDS subsystem (i.e., number of independent decisions in a given time interval),
- $DR_{ODO}$ the (Independent) decision rate of the ODO LDS subsystem (i.e., number of independent decisions in a given time interval).

Following [6], for *asynchronous LDS estimation* (i.e. GNSS LDS and ODO LDS independently provide train location and speed estimates at different time instants without time coordination ) for the overall Hazard rate we obtain

$$HR = R_{DEF}^{DRAG} + R_{DEF}^{MLDS} + \frac{(DR_{GNSS} + DR_{ODO})}{DR_{GNSS}DR_{ODO}} \left( R_{DEF}^{GNSS\_HMI} + R_{DEF}^{GNSS} \right) \left( R_{DEF}^{ODO\_HMI} + R_{DEF}^{ODO} \right), \quad (4)$$

where

$$R_{DEF}^{GNSS} = R_{DEF}^{GNSS\_LDS} + DR_{GNSS} \left( R_{DEF}^{GNSS\_RX} \frac{N_{GNSS\_RX}}{DR_{GNSS}} \right) \quad (5)$$

$$R_{DEF}^{ODO} = R_{DEF}^{ODO\_LDS} + DR_{ODO} \left( R_{DEF}^{INS} \frac{N_{INS}}{DR_{ODO}} \right) \left( R_{DEF}^{TACHO} \frac{N_{TACHO}}{DR_{ODO}} \right) \quad (6)$$

On the other hand, when both LDS are synchronized, as in the actual LDS implementation, the number of decisions in 1 hour is $DR=DR_{ODO}=DR_{GNSS}$ and for the overall Hazard rate we obtain:

$$HR_{\text{sync}} = R_{DEF}^{DRAG} + R_{DEF}^{MLDS} + \frac{1}{DR} \left( R_{DEF}^{GNSS\_HMI} + R_{DEF}^{GNSS} \right) \left( R_{DEF}^{ODO\_HMI} + R_{DEF}^{ODO} \right). \quad (7)$$

From (7), we notice that the Multisensor LDS unit and the Independent Diagnostics components are critical elements w.r.t. safety, since sensitivity of the Hazard Rate w.r.t. dangerous undetected failure rates, HMI of GNSS and ODO components, is mitigated by the functional redundancy.

As a consequence, both Multisensor LDS unit and the Independent Diagnostics components should guarantee a SIL greater than the SIL required for the overall system.

As an instance, a possible Risk Allocation for a SIL 4 system is reported in Table 1. We note that the simplified uniform strategy does not account for the
complexity of achieving the same Hazard Rate for heterogeneous technologies.

Table 1. Risk Allocation example.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Numerical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{SSF}^{G{N}{S}_{S}}$</td>
<td>$8.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>$P_{SSF}^{DOE}$</td>
<td>$8.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>$P_{SSF}^{HMI}$</td>
<td>$8.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>$P_{SSF}^{INP}$</td>
<td>$8.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>$P_{SSF}^{MLD}$</td>
<td>$1.0 \times 10^{-10}$</td>
</tr>
<tr>
<td>$P_{SSF}^{IDF}$</td>
<td>$1.0 \times 10^{-10}$</td>
</tr>
</tbody>
</table>

Nevertheless it enlightens how the desired safety level can be reached by combining functional redundancy of subsystem satisfying weaker requirements.

The HMI Rate is then defined as the probability of providing at least an HMI during a predefined time interval i.e., equal to 1 hour in railway applications.

On the other hand, when evaluating the integrity risk three different kinds of events should be considered, such as (i) the fault free case, (ii) the individual satellite faults, and (iii) the correlated simultaneous satellite faults.

4.1 FAULT FREE CASE

When all satellites are healthy a HMI can still arise due to large random errors produced by multipath, receiver thermal noise or tropospheric incremental delays, besides their small probability of occurrence.

Let us denote with $P_{SSF}$ the probability of fault of a single satellite and with $N_{Sat}$ number of visible satellites. Then, the probability that none of them is affected by a fault is bounded by the probability that none of them is affected by an independent fault, i.e.: $P_{SH} \leq (1 - P_{SSF})^{N_{Sat}}$. (8)

In practice for $P_{SSF} \ll 1$, the following approximation holds $P_{SH} \leq 1 - N_{Sat}P_{SSF}$. (9)

We recall that according to the GPS SPS Performance Standard the for the GPS constellation we have $P_{SSF}^{GPS} \leq 10^{-5}/h$. (10)

Let $\{s_i\}$ be the set of the operational modes of the receiver. Then, denoting with $\sigma_{PE_i}^2$ the variance of the estimation error when every satellite is healthy and the receiver is in the $s_i$ operational mode, since in this case
- the position estimation error can be modeled as a Gaussian random variate (r.v.) with variance $\sigma_{PE_i}^2$ and expectation $b_{h_i}$,
- the Misleading Information (MI) event is statistically independent from the Missing Alert (MA) event

the conditional probability $P_{SSF}^{SH}(s_i)$ of an MI event given an MA event when the receiver is in the $s_i$ operational status and all the satellites of a given constellation are healthy, equals the probability that the position error will exceed the protection level $PL$. Thus we obtain:

$$P_{SSF}^{SH}(s_i) = \Pr\{|\Delta| > PL\} = \frac{1}{2} \text{erfc}\left(\frac{PL - b_i}{\sqrt{2}\sigma_{PE_i}}\right) + \frac{1}{2} \text{erfc}\left(\frac{PL + b_i}{\sqrt{2}\sigma_{PE_i}}\right),$$

where erfc() is the complementary error function

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt.$$  

The position estimate error variance $\sigma_{PE_i}^2$ is directly related to the satellite lines of sight w.r.t. the actual track and to the pseudo-range measurement noise that may vary with the operational mode and the considered satellite. In fact

$$\sigma_{PE_i}^2 = \left(\mathbf{H}^\top \mathbf{R}^{-1}_i \mathbf{H}\right)^{-1}_{1,1}$$  

where $\mathbf{H}$ is matrix

$$\mathbf{H} = \left[\mathbf{PD} \ 1_{x_{Sat}}\right].$$

4.2 Individual satellite faults

Individual satellite faults may arise in both the Space and the Ground segments. The historical data indicate that average number of faults for the GPS constellation is about 3 per year. This frequency is consistent with the cited single failure rate of $10^{-7}/h$ reported in the GPS SPS Performance Standard.

Thus the probability of having a failed satellite out of $N_{Sat}$ satellites is

$$P_{sf}^{Sat} = \left[\text{Pr}\left[Sat_{Sat} = 1\right]\right] \approx N_{Sat}P_{SSF}.$$  

while the probability of having 2 failures out of $N_{Sat}$ satellites is

$$P_{sf}^{2Sat} = \left[\text{Pr}\left[Sat_{Sat} = 2\right]\right] = \frac{N_{Sat}}{2}P_{SSF}^2(1 - P_{SSF})^{N_{Sat}-2}.$$  

Thus for $P_{SSF}=10^{-5}$ and $N_{Sat}=11$, $P_{sf}^{2Sat}=2<0.5x10^{-8}$. On the other hand, the probability of having 3 or more failures is
\[ P_{SF}[N_{\text{fault}} > 2] = \sum_{k=3}^{N_{\text{Sat}}} P_{SS}(1 - P_{SS})^{N_{\text{Sat}}-k} \]  

(17)

Thus for \( P_{SS}=10^{-5} \) and \( N_{\text{Sat}} < 11 \), \( P_{SF}[N_{\text{fault}} > 2] < 0.12 \times 10^{-12} \). Since in the evaluation of the Integrity Risk the fault probability has to be multiplied by the probability that the position error exceeds the protection level. Considering that usually also this latter term is small and the smallest Integrity Risk considered here is \( 10^{-9}/\text{h} \), although the design of the Range & Integrity Monitoring shall explicitly account for multiple satellite faults, here we will focus the numerical evaluation on a single satellite fault.

Concerning the effects produced by a satellite fault on the estimation error we observe that whenever the receiver operates in a differential equivalent mode, as in the operational modes \( s_1, s_2, \) and \( s_3 \), errors on ephemeris and satellite clocks are essentially compensated. Nevertheless, latency and spatial distribution of RIMs may slightly impact on the variance and bias of the position error. Therefore,

\[
P_{\text{MT/MA}}(s_h) = \text{Pr}\{ |\Delta s| > PL \} = \frac{1}{2} \text{erfc} \left( \frac{PL - b_h}{\sqrt{2}\sigma_{PE}} \right) + \frac{1}{2} \text{erfc} \left( \frac{PL + b_h}{\sqrt{2}\sigma_{PE}} \right), \quad h = 1, 2, 3 \]  

(18)

When the receiver operates in the stand alone mode (operational mode \( s_3 \)) and the \( i \)-th satellite presents a failure characterized by an uncompensated range error, the train location estimate is affected by an additional error that can drastically impact on the \( P_{\text{HIM}} \). In order to detect a possible fault, the health of each satellite is monitored by analyzing the statistical behavior of the observed pseudoranges. In particular the measured pseudoranges are compared to the pseudoranges that would be observed if the estimated receiver location and clock offset were the true one. The difference between these two quantities, denoted in the literature and in the following as pseudorange residuals, is essentially an estimate of the measurement error and noise.

For healthy satellites the residuals are zero mean random variables with variance equal to the variance of the measurement noise. But, in presence of a satellite failure that causes a range error \( b \), the mean of the residuals changes and the root mean square value changes accordingly.

Many Range & Autonomous Integrity Monitoring Algorithms have been investigated in the Literature. Among them the dual frequency multiconstellation GEAS Advanced RAIM (ARAIM), originally designed for aviation applications in order to meet the new challenging requirements for the 2010-2030 timeframe, appears to be one of the most effective in terms of both performance and complexity [16], [17]. Thus, adaptation to the railway scenario has been considered. For sake of simplicity, since ARAIM outperforms the existing algorithms by applying rather complex algorithms for threshold optimization, here the evaluation of the Integrity Risk is carried out assuming a simpler algorithm that detects a satellite fault by comparing the \( L^2 \) norm of the residuals with a fixed threshold \( AL \) named Faulty Satellite Alarm Level. More specifically, since the satellites may belong to different constellations, the square of the \( L^2 \) weighted norm \( \zeta \) of the residuals \( \mathbf{v} \), is normalized w.r.t. the covariance of the measurement noise \( \mathbf{R}_s \), namely:

\[
\zeta^2 = \mathbf{v}^T \mathbf{R}_s^{-1} \mathbf{v}. 
\]  

(19)

Incidentally we observe that, due to the normalization w.r.t. its covariance, the weighted norm \( \zeta \) and \( AL \) are both adimensional numbers.

If \( \zeta^2 \) exceeds \( AL \) an HMI event is notified. This in turn implies that an alarm may be raised even in absence of satellite failures. This case is usually denoted as False Alarm while its probability is named False Alarm probability and denoted in the following as \( P_{fa}(s_h) \). As a matter of fact, \( AL \) controls both the \( P_{fa}(s_h) \) and the probability \( P_{MA}(s_h) \) of missing the detection of a real failure. Since requirements are specified in terms of both False Alarm probability and Integrity Risk which directly depends on \( P_{MA}(s_h) \), usually the Faulty Satellite Alarm Level is set in correspondence of the target \( P_{fa}(s_h) \). Then, having determined \( AL \), \( P_{MA}(s_h) \) is computed. When all satellites are healthy, the square of residual weighted norm \( \zeta \) is a random variable with a chi square distribution with \( N_{\text{Sat}} \) degrees of freedom, being \( N_{\text{Sat}} \) the number of visible satellites. Thus the Alarm Level corresponding to a given false alarm probability can be computed as:

\[
AL(s_h) = D_{\zeta^2}^{-1} \left( 1 - P_{fa}(s_h) \right). 
\]  

(20)

where \( D_{\zeta^2}^{-1} (\cdot) \) is the inverse of the cumulative chi square distribution with \( n \) degrees of freedom. When a satellite fault produces a range error \( b \), the square of the \( L^2 \) norm of the residuals becomes a non-central chi square random variable, with non-centrality parameter \( \lambda \) whose magnitude is proportional to the square of \( b \). Thus, as detailed in Appendix, the probability \( P_{MA}(s_h) \) of missing the detection of the satellite fault can be expressed in terms of \( AL \) as follows:

\[
P_{MA}(s_h) = D_{\zeta^2}^{-1} \left( AL, \lambda \right). 
\]  

(21)

Nevertheless, in addition to the evaluation of \( P_{MA}(s_h) \), the computation of the integrity risk requires the computation of the probability \( P_{\text{MT/MA}}(s_h) \) that the position error magnitude will exceed the Protection Level. For a compact and manageable form of the Integrity Risk, \( P_{\text{MT/MA}}(s_h) \) is expressed in terms of the non-centrality parameter \( \lambda \) and of the Protection Level \( PL \) by exploiting the relationships among pseudorange error \( b \), variation \( \lambda \) of the \( L^2 \) norm and magnitude of the position error. Thus denoting with \( SLOPE \) the ratio between the magnitude of the position error and \( \sqrt{\lambda} \), and with \( SLOPE_{\text{Max}} \) its maximum value w.r.t. all satellites, \( P_{\text{MT/MA}}(s_h) \) can be written in a compact form as follows
operating in mode #1 equals the TALS availability \( P_{\text{TALS}}^{\text{Av}} \). SBAS augmentation data will be processed by the TALS server and incorporated in the TALS integrity evaluation. Thus for the probability of mode #1 we have

\[
P(s_1) = \begin{cases} 
P_{\text{TALS}}^{\text{Av}} & \text{TALS deployed} \\ 0 & \text{otherwise} \end{cases}
\] (25)

2. **GNSS augmented by SBAS** (e.g. EGNOS, WAAS) from Direct Signal In Space (SIS).
The SBAS system will be considered as a primary source of augmentation data when a GBAS system with RIMs distributed along a track is not deployed or unavailable.

Thus, denoting with \( P_{\text{SBAS}, \text{SIS}}^{\text{Av}} \) the availability of the SBAS SIS augmentation data for the probability of mode #2 we have

\[
P(s_2) = P_{\text{SBAS}, \text{SIS}}^{\text{Av}} \left[1 - P(s_1)\right].
\] (26)

We remark that \( P_{\text{SBAS}, \text{SIS}}^{\text{Av}} \) may vary with the actual location. [source EGNOS Safety of Life Service Definition Document]

3. **GNSS augmented by SBAS data retransmitted over Train Signaling Network**

Due to the greater latency introduced by direct connections to the operational centers providing SBAS augmentation data over terrestrial links and redistribution over the Train Signaling Network, w.r.t. the use of SBAS SIS data, this mode will be activated whenever modes #1 and #2 are unavailable.

Thus, denoting with \( P_{\text{SBAS}, \text{TSN}}^{\text{Av}} \) the availability of the SBAS TSN augmentation data for the probability of mode #3 we have

\[
P(s_3) = P_{\text{SBAS}, \text{TSN}}^{\text{Av}} \left[1 - P(s_1)\right] \left[1 - P(s_2)\right].
\] (27)

4. **Stand alone GNSS**.
Since this operational mode is selected whenever the other modes are unavailable we have:

\[
P(s_4) = 1 - P(s_1) - P(s_2) - P(s_3).
\] (28)

4.5 Hazardous Misleading Information Rate
Denoting with
- the \( MA \) subscript the Missing Alert event,
- the \( MI \) subscript the Misleading Information event,
- the \( SH \) superscript the event corresponding to healthy satellites,
- the \( SF \) superscript the event corresponding to one or more uncorrelated satellite failures,
- \( s_h \) the receiver operational mode,
- \( P_{\text{MA}}(s_h) \) the Missing Alert probability when the receiver is in the \( s_h \) operational mode and all the satellites of a given constellation are healthy,
- \( P_{\text{MA}}(s_h) \) the Missing Alert probability when the receiver is in the \( s_h \) operational mode and at least one satellite of a given constellation is faulty,
\[ P_{\text{HMI}} \] the conditional probability of an MI event given an MA event when the receiver is in the \( s_h \) operational mode and all the satellites of a given constellation are healthy, 

\[ P_{\text{HMI}} \text{ (} s_h \text{)} \] the conditional probability of an MI event given an MA event when the receiver is in the \( s_h \) operational mode and at least one satellite of a given constellation is faulty, 

\[ P_{\text{op}} \text{ (} s_h \text{)} \] the probability that the receiver is in the \( s_h \) operational status, 

\[ N_{\text{Dec}} \text{ (} s_h \text{)} \] the number of independent decisions in 1 hour, 

\[ P_{\text{st}} \] the probability that all satellites are healthy, 

\[ P_{\text{st}} = 1 - P_{\text{st}} \] the probability that at least one satellite fails, 

the Hazardous Misleading Information Rate 

\[ R_{\text{HMI}}^{\text{GNSS}} \] evaluated as the probability of an HMI event in 1 hour is 

\[
P_{\text{HMI}}^{\text{GNSS}} = 1 - \sum_{n=1}^{N_{\text{sat}}} \left[ \left( 1 - P_{\text{HMI}}^{\text{MA}} \text{ (} s_h \text{)} \right)^{N_{\text{Dec}}} P_{\text{HMI}} + \left[ \left( 1 - P_{\text{HMI}}^{\text{MA}} \text{ (} s_h \text{)} \right)^{N_{\text{Dec}}} P_{\text{st}} \right] P_{\text{op}} \text{ (} s_h \text{)} \right]. (29)
\]

In principle the Integrity Risk is a function of the \( \lambda \) parameter which is proportional to the square of the bias introduced by the satellite failure. Thus, the Hazardous Misleading Information Rate should be averaged even with respect to this quantity. However, since a reliable statistical model for the entity of the errors caused by satellite failures in not available, we set the Protection Level in accordance to the worst case. Thus, in the design of the LDS the following bound for the Integrity Risk has been applied:

\[
P_{\text{HMI}}^{\text{GNSS}} \leq 1 - \sum_{n=1}^{N_{\text{sat}}} \left[ 1 - \frac{1}{2} \cdot \text{erfc} \left( \frac{PL - \hat{b}_h}{\sqrt{2}\sigma_{PE}} \right) \right]
+ \sum_{n=1}^{N_{\text{sat}}} \left[ \frac{1}{2} \cdot \text{erfc} \left( \frac{PL + \hat{b}_h}{\sqrt{2}\sigma_{PE}} \right) \right] P_{\text{op}} \text{ (} s_h \text{)} P_{\text{st}}
+ \sum_{n=1}^{N_{\text{sat}}} \left[ \frac{1}{2} \cdot \text{erfc} \left( \frac{PL - \hat{b}_h}{\sqrt{2}\sigma_{PE}} \right) \right] P_{\text{op}} \text{ (} s_h \text{)} P_{\text{sf}}
- \left[ 1 - \text{Max} \left[ D^* \right] \right] \frac{1}{2} \left[ \text{erfc} \left( \frac{PL - \text{slope}_{\text{d}} \sqrt{\lambda}}{\sqrt{2}\sigma_{PE}} \right) \right]
+ \text{erfc} \left[ \frac{PL + \text{slope}_{\text{d}} \sqrt{\lambda}}{\sqrt{2}\sigma_{PE}} \right] P_{\text{op}} \text{ (} s_h \text{)} P_{\text{st}}. (30)
\]

Given the HMI Rate and the False Alarm Probability \( P_{\text{fa}} \), the above bound can be employed for computing the Protection Level.

### 5. SIMULATION RESULTS

To assess the performance improvement achievable by using a reliable augmentation system the SIL 4 protection level has been computed for a reference cases of a train travelling along a 350 km route, from Rome to Pisa (Italy), at a nominal speed of 80 km/hour, as depicted in Figure 6.

In the present evaluation masked areas as those pertaining to a tunnel and bridges (at road intersections) have been neglected. Both augmented and autonomous modes using the GPS and GPS+GLONASS satellites have been considered. The receiver error budget of Table 2 has been employed.

As already observed, although modes 1, 2, and 3 could present different error budgets, the main sensitivity is represented by the availability of the augmentation system. Thus, the simulations have been performed for the following availability levels:

- 0: no augmentation at all, the on board GNSS receiver is fully autonomous. In this case performance is directly determined by the sensitivity to bias introduced by the worst case satellites and the number of visible satellites;
- 99.99%: this case can be representative of EGNOS augmentation in EUROPE provided through train signalling network and received from EGNOS satellites from stations equipped with high gain antennas;
- 99.9999%: this case can be representative of EGNOS augmentation provided by two independent channels satellite + EDAS;
- 99.999999%: this case represents the joint use of TAAS + EGNOS with mild requirements for single availability (e.g., 99.99% for each of them);
- 100%: this case represents the case of augmentation provided by a SIL 4 subsystem.

The source error standard deviations for the GPS constellation versus the operational mode are reported in Table 2. GLONASS error variances have assumed to be twice those of GPS.

In Figure 7, the protection level versus the train location is reported for the availability levels defined above when both GPS and GLONASS are jointly processed.

### Table 2. Error Budget for GPS constellation.

<table>
<thead>
<tr>
<th>Error source</th>
<th>RMS error [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Satellite Clock Stability</strong></td>
<td>OP 1</td>
</tr>
<tr>
<td>Ephememeris prediction error</td>
<td>–</td>
</tr>
<tr>
<td>Ionospheric delay (dual frequency receiver)</td>
<td>0.1</td>
</tr>
<tr>
<td>Tropospheric delay (satellite elevation &gt;10°)</td>
<td>0.1</td>
</tr>
<tr>
<td>Multipath</td>
<td>1.0</td>
</tr>
<tr>
<td>Receiver thermal noise</td>
<td>0.5</td>
</tr>
<tr>
<td><strong>UERE</strong></td>
<td>1.05</td>
</tr>
</tbody>
</table>
6. CONCLUSIONS

This paper has investigated a novel GNSS solution and the theoretical modeling concerning the Safety Integrity Level for facilitating the adoption of the satellite-based localization systems in the ERTMS-ETCS ecosystem. The Multi-constellation architecture relying on GPS, GLONASS and in perspective GALILEO offers an higher degree of flexibility to reach the SIL-4 level that is mandatory for the railways applications. Nevertheless, the availability of an augmentation network is of paramount importance in reducing the PL. Moreover, increased accuracy is requested when additional capabilities, like parallel track discrimination is required. In this sense, availability of current SBAS SIS developed for aeronautical applications is of primary concern. As illustrated by the performance analysis, distributing augmentation data through the train signaling system network represents a cost effective mean to increase integrity information and augmentation data availability.

APPENDIX A. LOCATION ESTIMATION ERROR STATISTICS

To evaluate the statistics of the location estimation error, the pseudo-range equations can be approximated by the first order Taylor’s series expansion around the current train curvilinear abscissa (and estimated compensations) so that, dropping for sake of compactness the temporal index \( k \), the reduced \( i \)-th pseudorange becomes

\[
\Delta \rho_i = \left( \frac{\partial \rho_i}{\partial x_{Ei}} \frac{\partial x_{\text{OBU}}}{\partial s} + \frac{\partial \rho_i}{\partial x_{yi}} \frac{\partial x_{\text{OBU}}}{\partial s} + \frac{\partial \rho_i}{\partial x_{ui}} \frac{\partial x_{\text{OBU}}}{\partial s} \right) \Delta s + c \Delta t_{\text{OBU}} + \nu, \quad i = 1, 2, \ldots, m
\]

(A.1)

where \( m \) is the number of visible satellites. Eq. (A.1) can be written in compact matrix notation as follows

\[
\Delta \rho = H z + \nu
\]

(A.2)

where

\[
z = \begin{bmatrix} \Delta s \\ c \Delta t_{\text{OBU}} \end{bmatrix},
\]

(A.3)

and the \( m \times 2 \) observation matrix \( H \) is

\[
H = [P D \ 1_m],
\]

(A.4)

where \( P \) is the \( m \times 3 \) Jacobian matrix of the pseudo-ranges with respect to the Cartesian train coordinates,

\[
P = \frac{\partial \rho}{\partial x_{\text{OBU}}} = \begin{bmatrix} \frac{\partial \rho_1}{\partial x_E} & \frac{\partial \rho_1}{\partial x_N} & \frac{\partial \rho_1}{\partial x_U} \\ \frac{\partial \rho_2}{\partial x_E} & \frac{\partial \rho_2}{\partial x_N} & \frac{\partial \rho_2}{\partial x_U} \\ \vdots & \vdots & \vdots \\ \frac{\partial \rho_m}{\partial x_E} & \frac{\partial \rho_m}{\partial x_N} & \frac{\partial \rho_m}{\partial x_U} \end{bmatrix}
\]

(A.5)
with elements given by the directional cosines of the satellite LoS:

$$P_j = \frac{\partial \rho_j}{\partial x_{OBU}^j} = \frac{x_{OBU}^{j} - X_j}{\|x_{OBU}^{j} - X_j\|}, \quad j = E, N, U \quad (A.6)$$

$D$ is the 3x1 tangent vector to the track in $X_{OBU}$:

$$D = \begin{bmatrix} \frac{\partial x_{OBU}}{\partial \phi} \\ \frac{\partial x_{OBU}}{\partial \lambda} \\ \frac{\partial x_{OBU}}{\partial \nu} \end{bmatrix}$$

and $I_m$ is the mx1 vector:

$$I_m = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad (A.7)$$

When the TALS augmentation information is employed, since the estimates of the satellite positions and clock offsets are unbiased, $\mathbf{v}$ can be modeled as a zero mean Gaussian $m$-variate random variable with covariance matrix $R_v$.

Therefore, the estimation error $\hat{\mathbf{z}} = \mathbf{z} - \hat{\mathbf{z}}$ of the OBU receiver location and clock offset is a zero mean Gaussian r.v. with covariance matrix $R_{\hat{\mathbf{z}}}$:

$$R_{\hat{\mathbf{z}}} = E\{\hat{\mathbf{z}}\hat{\mathbf{z}}^T\} = \left(\mathbf{H}'\mathbf{R}^{-1}\mathbf{H}\right)^{-1} \quad (A.9)$$

Eq. (A.9) highlights that the variance of the estimate of the curvilinear abscissa $s$

$$\sigma_s^2 = [R_{\hat{\mathbf{z}}}^{-1}]_{s1}$$

depends on both the number of visible satellites and line of sight geometry.

We recall that, at each iteration, the weighted least square estimate $\hat{\mathbf{z}}$ is computed as $\hat{\mathbf{z}} = \mathbf{K}\Delta\mathbf{p}$, where $\mathbf{K}$ is the gain matrix

$$\mathbf{K} = \left(\mathbf{H}'\mathbf{R}_{s}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}'\mathbf{R}_{s}^{-1}. \quad (A.11)$$

As a consequence the least square residuals, given by the difference between the measured reduced pseudoranges and the reduced pseudoranges corresponding to the estimate $\hat{\mathbf{z}}$, are

$$\hat{\mathbf{v}} = \Delta\mathbf{p} - \mathbf{H}\hat{\mathbf{z}} = (I - \mathbf{HK})\Delta\mathbf{p} \quad (A.12)$$

Then, when every satellite is healthy or when at least one satellite fails and the TALS server augmentation information is employed, they can be modeled as a zero mean Gaussian $m$-variate random variable with covariance matrix $R_v$. In fact, from Eq. (A.12) it follows that

$$R_v = (I - \mathbf{HK})R_v(I - \mathbf{HK})^T. \quad (A.13)$$

On the other hand, by applying the matrix inversion lemma we obtain

$$I - \mathbf{HK} = \left[I - \mathbf{H}'\mathbf{R}^{-1}_s\mathbf{H}\right]^{-1}\mathbf{H}'\mathbf{R}^{-1}_s$$

$$= \left[I + \mathbf{H}'\mathbf{R}^{-1}_s\mathbf{H} - \mathbf{H}'\mathbf{R}^{-1}_s\mathbf{H}\right]^{-1}\mathbf{H}'\mathbf{R}^{-1}_s \quad (A.14)$$

$$= I.$$

Therefore

$$R_v = R_v, \quad (A.15)$$

$q.e.d.$

On the other hand when the LDS operates in autonomous mode ad a satellite constellation failure characterized by an uncompensated range error vector $b$ arises, the estimation can be modeled as Gaussian r.v. with covariance matrix still given by Eq (A.3) but with a bias equal to

$$E\{\hat{\mathbf{z}}\} = \mathbf{K}\mathbf{b}. \quad (A.16)$$

It can also be demonstrated that, in this case, the residuals can be modeled as Gaussian $m$-variate random variable with covariance matrix $R_v$ and expectation

$$E\{\hat{\mathbf{v}}\} = (I - \mathbf{HK})\mathbf{b}. \quad (A.17)$$

The RAIM algorithm continuously monitors the behavior of the pseudorange residuals $\mathbf{v}$. More specifically, the square of their $L^2$ weighted norm $\zeta^2$

$$\zeta^2 = \hat{\mathbf{v}}'\mathbf{R}_{s}\hat{\mathbf{v}}. \quad (A.18)$$

is compared to a threshold $T$. If $\zeta^2$ exceeds the threshold an HMI event is notified. The threshold is set in accordance to a given false alarm probability $P_{fa}$.

When all satellites are healthy, the residual weighted norm $\zeta^2$ is a random variable with a chi square distribution with $N_{sat}$-2 degrees of freedom, being $N_{sat}$ the number of visible satellites. Thus for a given threshold $T$ the corresponding false alarm probability is

$$P_{fa} = 1 - D_{\chi^2}^{-1}\{1 - P_{fa}\} \quad (A.19)$$

where $D_{\chi^2}^{-1}(\cdot)$ is the cumulative chi square distribution with $n$ degrees of freedom. Thus, the Alarm Level corresponding to a given false alarm probability can be computed as

$$AL = D_{\chi^2}^{-1}\{1 - P_{fa}\} \quad (A.20)$$

where $D_{\chi^2}^{-1}(\cdot)$ is the inverse of the cumulative chi square distribution with $n$ degrees of freedom.

On the other end, when $i$-th satellite presents a failure characterized by an uncompensated range error $b_i$, the train location estimate is affected by the additional error:

$$\Delta\mathbf{p}_i = \mathbf{K}_i b_i. \quad (A.21)$$

Therefore, the probability that the position error will be greater, and then the protection level becomes

$$P_{ML/MA}^{SS} = \frac{1}{2} \text{erfc} \left( \frac{PL - \mathbf{K}_i b_i}{\sqrt{2}\sigma_i} \right) + \frac{1}{2} \text{erfc} \left( \frac{PL + \mathbf{K}_i b_i}{\sqrt{2}\sigma_i} \right), \quad (A.22)$$

14
On the other hand, in this case the expectation of the residual of the $i$-th pseudorange is
\[ E \{ \hat{r}_i \} = [I - HK]_{ii} b_i. \]  
(A.23)

Thus, the $L^2$ norm of the pseudorange residuals has a non-central chi square distribution with non-centrality parameter
\[ \lambda(h) = [I - HK]_{ii} b_i^2 \frac{1}{\hat{R}_{ii}}. \]  
(A.24)

Therefore, the probability that the HMI event will not be detected equals the probability that a r.v. with non-central chi square distribution with non-centrality parameter $\lambda(h)$ will not exceed the Alarm Level $AL$:
\[ P_{\text{SF}}^{\text{AL}} = D_{\lambda(h)}^{\chi^2 n} \left[ AL, \lambda(h) \right] \]  
(A.25)

where $D_{\lambda(h)}^{\chi^2 n} (\cdot, \lambda)$ is the cumulative chi square non-central distribution with $n$ degrees of freedom and non-centrality parameter $\lambda$.

Let us denote with $SLOPE_i$ the ratio
\[ SLOPE_i = \frac{K_{ii}}{[I - HK]_{ii}} \sqrt{\hat{R}_{ii}}. \]  
(A.26)

so that we can write
\[ \Delta s_i^{\text{SF}} = K_{ii} b_i \]
\[ = \frac{K_{ii}}{[I - HK]_{ii}} \sqrt{\hat{R}_{ii}} [I - HK]_{ii} b_i \sqrt{\hat{R}_{ii}} \]
\[ = SLOPE_i \sqrt{\lambda}. \]  
(A.27)

Then, in the autonomous mode, denoting with $SLOPE_{\text{Max}}$ the maximum of $SLOPE_i$ w.r.t. the whole set of visible satellites i.e.,
\[ SLOPE_{\text{Max}} = \max_i \left| SLOPE_i \right|, \]  
(A.28)

consequently for the the HMI conditional probability w.r.t. the $s_i$ operational mode in presence of satellite failures the following upper bound holds:
\[ P_{\text{SF}}^{\text{AL}} (s_i) P_{\text{SF}}^{\text{AL/MA}} (s_i) \leq \frac{1}{2} D_{\lambda(h)}^{\chi^2 n} \left[ AL, \lambda \right] \times \]
\[ \left\{ \text{erf} \left[ \frac{PL - SLOPE_{\text{Max}} \sqrt{\lambda}}{\sqrt{2} \sigma_{\text{PE}_i}} \right] + \text{erf} \left[ \frac{PL + SLOPE_{\text{Max}} \sqrt{\lambda}}{\sqrt{2} \sigma_{\text{PE}_i}} \right] \right\}. \]  
(A.29)

**ACKNOWLEDGMENTS**

This work is partly based on the studies of the “3INSAT - Train Integrated Safety Satellite System” ESA project currently under development.

**REFERENCES**