Global navigation satellite systems performance analysis and augmentation strategies in aviation

Roberto Sabatini[1]*, Terry Moore[2] and Subramanian Ramasamy[1]

[1]RMIT University, School of Engineering – Aerospace Engineering and Aviation Discipline, Bundoora, VIC 3083, Australia

[2]University of Nottingham, Nottingham Geospatial Institute, Nottingham, NG7 2TU, United Kingdom

ABSTRACT

In an era of significant air traffic expansion characterised by a rising congestion of the radiofrequency spectrum and a widespread introduction of Unmanned Aircraft Systems (UAS), Global Navigation Satellite Systems (GNSS) are being exposed to a variety of threats including signal interferences, adverse propagation effects and challenging platform-satellite relative dynamics. Thus, there is a need to characterize GNSS signal degradations and assess the effects of interfering sources on the performance of avionics GNSS receivers and augmentation systems used for an increasing number of mission-essential and safety-critical aviation tasks (e.g., experimental flight testing, flight inspection/certification of ground-based radio navigation aids, wide area navigation and precision approach). GNSS signal deteriorations typically occur due to antenna obscuration caused by natural and man-made obstructions present in the environment (e.g., elevated terrain and tall buildings when flying at low altitude) or by the aircraft itself during manoeuvring (e.g., aircraft wings and empennage masking the on-board GNSS antenna), ionospheric scintillation, Doppler shift, multipath, jamming and spurious satellite transmissions. Anyone of these phenomena can result in partial to total loss of tracking and possible tracking errors, depending on the severity of the effect and the receiver characteristics. After designing GNSS performance threats, the various augmentation strategies adopted in the Communication, Navigation, Surveillance/Air Traffic Management and Avionics (CNS+A) context are addressed in detail. GNSS augmentation can take many forms but all strategies share the same fundamental principle of providing supplementary information whose objective is improving the performance and/or trustworthiness of the system. Hence it is of paramount importance to consider the synergies offered by different augmentation strategies including Space Based Augmentation System (SBAS), Ground Based Augmentation System (GBAS), Aircraft Based Augmentation System (ABAS) and Receiver Autonomous Integrity Monitoring (RAIM). Furthermore, by employing multi-GNSS constellations and multi-sensor data fusion techniques, improvements in availability and continuity can be obtained. SBAS is designed to improve GNSS system integrity and accuracy for aircraft navigation and landing, while an alternative approach to GNSS augmentation is to transmit integrity and differential correction messages from ground-based augmentation systems (GBAS). In addition to existing space and ground based augmentation systems, GNSS augmentation may take the form of additional information being provided by other on-board avionics systems, such as in ABAS. As these on-board systems normally operate via separate principles than GNSS, they are not subject to the same sources of error or interference. Using suitable data link and data processing technologies on the ground, a certified ABAS capability could be a core element of a future GNSS Space-Ground-Aircraft Augmentation Network (SGAAN). Although current augmentation systems can provide significant improvement of GNSS navigation performance, a properly designed and flight-certified SGAAN could play a key role in trusted autonomous system and cyber-physical system applications such as UAS Sense-and-Avoid (SAA).

1. Introduction

The origins of Global Navigation Satellite Systems (GNSS) date back to the early 1960s, when the United States Department of Defense initiated the development of systems for three–dimensional position determination [1]. After the US Navy successfully tested the first satellite navigation system called TRANSIT, the Space Division of the US Air Force initiated a program, known as Project 621B that evolved into the Navigation Signal Time and Range (NAVSTAR) program. In 1973, the US Defense Navigation Satellite System (DNSS) was created, which was later referred to as NAVSTAR Global Positioning System (GPS). Various GNSS systems are currently in service or under development. The US GPS and the Russian GLONASS (Globalnaya Navigazionnaya Sputnikovaya Sistema) have achieved their Full Operational Capability (FOC) back in the 1990's [2]. Other GNSS systems that are currently at the advanced development or deployment stages include the European GALILEO and the People's Republic of China BEIDOU Navigation Satellite System (BDS). GNSS systems typically use signals 20 dB below the ambient noise floor and, despite several research efforts devoted to interference detection and mitigation strategies at receiver and platform level (mostly for military applications), no effective solution has been implemented so far in the civil aviation context. Thus, there is a need to characterize GNSS signal degradations and assess the effects of interfering sources on the performance of avionics GNSS receivers and Differential GNSS (DGNSS) systems used for an increasing number of mission-essential and safety-critical aviation tasks (e.g.,

experimental flight testing, flight inspection/certification of ground-based radio navigation aids, wide area navigation and precision approach).

GNSS signal deteriorations typically occur due to antenna obscuration caused by natural and man-made obstructions present in the environment (e.g., elevated terrain and tall buildings when flying at low altitude) or by the aircraft itself during manoeuvring (e.g., aircraft wings and empennage masking the on-board GNSS antenna), ionospheric scintillation, Doppler shift, multipath, jamming and spurious satellite transmissions. Anyone of these phenomena can result in partial to total loss of tracking and possible tracking errors, depending on the severity of the effect and the receiver characteristics. Tracking errors, especially if undetected by the receiver software, can result in large position errors. Partial loss of tracking results in geometry degradation, which in turn affects position accuracy. Consequently, GNSS alone does not always provide adequate performance in mission-essential and safety-critical aviation applications where high levels of accuracy and integrity are required.

GNSS augmentation can take many forms but all share the same fundamental principle of providing supplementary information whose objective is improving the performance and/or trustworthiness of the system. GNSS augmentation benefits in the aviation domain can be summarized as follows:

- Increased runway access, more direct en-route flight paths and new precision approach services;
- Reduced and simplified avionics equipment;

- Potential elimination of some ground-based navigation aids (VOR, ILS, etc.) with cost saving to Air Navigation Service Providers (ANSPs).

This paper is organised as follows: Section 2 provides a detailed discussion on GNSS aviation applications, followed by a description of models for GNSS performance threats in Section 3; Section 4 presents the different augmentation strategies and the identification of a pathway to a future GNSS Space-Ground-Avionics Augmentation Network (SGAAN) is discussed in Section 5; an investigation of the potential of GNSS augmentation techniques to support trusted autonomous Unmanned Aircraft System (UAS) applications is presented in Section 6; The conclusions of this article are summarised in Section 7 and recommendations for future research are highlighted in Section 8.

2. GNSS Aviation Applications

Although different GNSS systems employ diversified hardware and software features, all systems are composed by a space segment, a control segment and a user segment (Fig. 1). The space segment includes the satellites required for global coverage. These satellites are predominantly in Intermediate Circular Orbit (ICO),

at nominal altitudes of 19,100 km (GLONASS), 20,184 km (GPS) and 23,222 km (GALILEO) from the Earth's surface.

BDS also employs satellites in Geostationary Orbit (GEO) and Inclined Geosynchronous Orbit (IGSO). The user segment includes the large variety of GNSS receivers developed for air, ground and marine navigation positioning applications. The control segment includes one or more Control and Processing Stations (CPSs) connected to a number of Ground Monitoring Stations (GMSs) and antennae located around the globe for Telemetry, Tracking and Command (TT&C) signals down/uplink and navigation/integrity signals uplink to the satellites. The GMS antennae passively track all GNSS satellites in view collecting ranging signals from each satellite. This information is passed on to the CPS where the satellite ephemeris and clock parameters are estimated and predicted. Additionally, satellite integrity data are analysed and appropriate integrity flags are generated for faulty/unreliable satellites. The ephemeris/clock and integrity data are then up-linked to the satellite for retransmission in the navigation message. The satellite clock drift is corrected so that all transmitted data are synchronised with GNSS time.
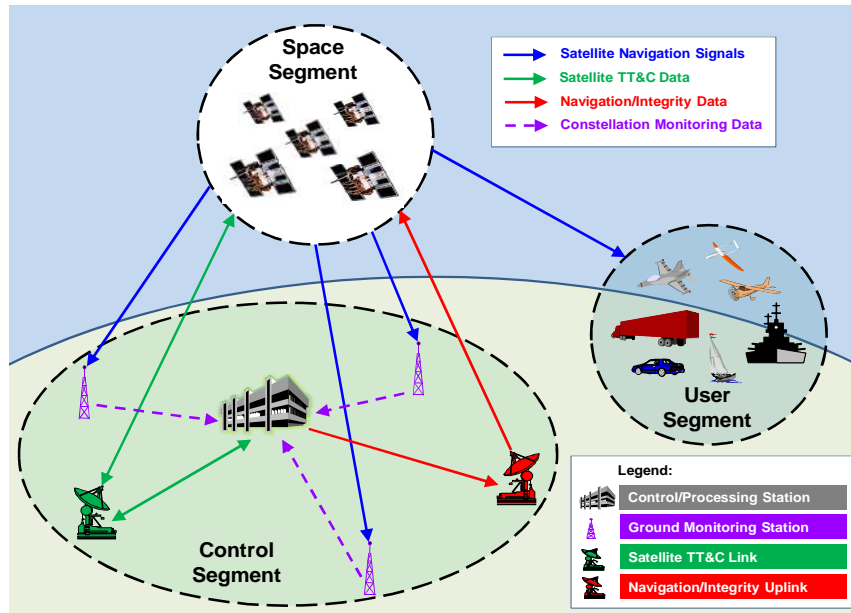

Fig. 1. GNSS segments.

The fundamental equation is the following:

$$t_{GNSS} = t_s - \Delta t_s \qquad (1)$$

where t_{GNSS} is the GNSS time, t_s is the satellite time and Δt_s is the difference between satellite and GNSS time. The corrections are applied to the last term of equation (1), typically using polynomial coefficients and a relativistic correction term. The correction equation can be written as:

$$\Delta t_s = a_0 + a_1(t_{GNSS} - t_{0c}) + a_2(t_{GNSS} - t_{0c})^2 + \Delta t_r \quad (2)$$

where a_0, a_1 and a_2 are the polynomial coefficients for phase, frequency and age offset; Δt_r is the relativistic correction term and t_{0c} is the time of transmission of the corrections.

Typically, the ephemeris corrections are obtained through an estimation of the Cartesian co-ordinates of the satellites along the orbits by integrating the equations of motion. For integrity purposes, suitable Fault Detection, Isolation and Recovery (FDIR) techniques are employed. In particular, the health status of the satellite subsystems is continuously monitored (satellite payload, bus, solar arrays, battery power and the level of propellant used for maneuvers) and any anomaly must be promptly detected and resolved. When needed, spare satellites can be activated. The

Signal-in-Space (SIS) is also constantly monitored to guarantee the required performance standards. Despite the significant technological enhancements recently introduced in control segment integrity features, current GNSS systems have limited FDIR capability. In most cases, several minutes or even hours are required to provide the required integrity information (i.e., use/do not use signals) to GNSS users. Obviously, this is not acceptable for mission-essential and safety-critical aviation applications.

2.1. GNSS Observables

There are basically three types of GNSS observables that can be used in aviation GNSS receivers: pseudorange, carrier phase, and Doppler observable. Pseudoranges are commonly used in real-time aircraft navigation and can provide an accuracy ranging from about 20 m (single frequency receivers) to about 2 m in Differential GNSS (DGNSS) positioning. The carrier phases are traditionally used in high precision trajectory determination applications (e.g., TSPI systems for flight test and flight inspection) and can achieve sub-centimetre accuracy. However, it is also quite common to use combinations of pseudoranges and carrier phase measurements both in real-time and post mission applications. Moreover, it has become common practice to take advantage of various

combinations of the original phase observation, such as double differences and triple differences.

2.1.1. Pseudorange Observable

The concept of pseudoranging is based on measuring difference between the time of transmission of the code from the satellite and the epoch of reception of the same signal at the receiver antenna. This is achieved by correlating identical pseudorandom noise (PRN) codes generated by the satellite's clock, with those generated internally by the receiver's own clock. If this time difference is multiplied by the speed of propagation of the radio wave, a range value is obtained which is the distance between the satellite and the receiver's antenna referring to the epoch of observation. Both receiver and satellite clock errors affect the pseudoranges. Therefore, they differ from the actual geometric distance corresponding to the epochs of emission and reception. The general pseudorange equation is:

$$P_k{}^p(t_k) = (t_k - t^p) \times c \qquad (3)$$

where $P_k{}^p$ represents the actual measurement, t_k denotes the nominal time of the receiver clock k at reception, t^p denotes the nominal time of the satellite clock p at emission and c denotes the speed of light. Equation (3) would correspond to the actual distance between the satellite and receiver's antenna, if there were no clock biases, the signal travelled through vacuum and there was no multipath effect. The clock drifts can be represented by the following expressions:

$$t_{r,k} = t_k + dt_k \qquad (4)$$

$$t_k{}^p = t^p + dt^p \qquad (5)$$

where the symbol r denotes the true time and the terms dt_k and dt^p are the receiver and satellite clock errors respectively. Taking these errors and biases into account, the complete expression for the pseudorange becomes:

$$P_k{}^p(t_k) = \left(t_{r,k} - t_r{}^p\right)c - (dt_k - dt^p)c + I_{k,p}{}^p(t_k) +$$
$$T_k{}^p(t_k) + d_{k,p}(t_k) + d_{k,p}{}^p(t_k) + d_p{}^p(t_k) + \varepsilon_p \qquad (6)$$

Therefore,

$$P_k{}^p(t_k) = \rho_k{}^p(t_{r,k}) - (dt_k - dt^p)c + I_{k,p}{}^p(t_k) +$$
$$T_k{}^p(t_k) + d_{k,p}(t_k) + d_{k,p}{}^p(t_k) + d_p{}^p(t_k) + \varepsilon_p \qquad (7)$$

where $I_{k,p}{}^p(t_k)$ and $T_k^p(t_k)$ are the ionospheric and tropospheric delays, depending on varying conditions along the path of the signal. The symbols $d_{k,p}(t_k)$ and $d_p{}^p(t_k)$ denote the receiver and satellite hardware code delays respectively. The symbol $d_{k,p}{}^p(t_k)$ denotes the multipath of the codes, which depends on the geometry of the antenna and satellite with respect to surrounding reflective surfaces. The term ε_p denotes the random measurement noise. The term $\rho_k{}^p(t_{r,k})$ is the actual geometric distance between the receiver's antenna and the satellite at a specific epoch and therefore:

$$\rho_k{}^p(t_r) = \sqrt{(u^p - u_k)^2 + (v^p - v_k)^2 + (w^p - w_k)^2} \qquad (8)$$

The terms (u_k, v_k, w_k) are the approximate Cartesian co-ordinates of the receiver and (u^p, v^p, w^p) denote the position of the satellite at the epoch of transmission. With reference to Fig. 2, assuming a constant receiver clock error dt_k for measurements to any satellite and omitting all other error terms in Eq. (6), the following system of equations is formed:

$$\begin{cases} P_k{}^1(t) = \sqrt{(u^1 - u_k)^2 + (v^1 - v_k)^2 + (w^1 - w_k)^2} - cdt_k \\ P_k{}^2(t) = \sqrt{(u^2 - u_k)^2 + (v^2 - v_k)^2 + (w^2 - w_k)^2} - cdt_k \\ P_k{}^3(t) = \sqrt{(u^3 - u_k)^2 + (v^3 - v_k)^2 + (w^3 - w_k)^2} - cdt_k \\ P_k{}^4(t) = \sqrt{(u^4 - u_k)^2 + (v^4 - v_k)^2 + (w^4 - w_k)^2} - cdt_k \end{cases} \qquad (9)$$

Therefore, the co-ordinates of the user receiver (and GNSS time) can be derived from the simultaneous observation of four (or more) satellites. If more than four satellites are visible, a least-square solution can be determined. The GNSS receiver calculates its position in an Earth-Centred Earth Fixed (ECEF) Cartesian co-ordinate system (typically using WGS84). These co-ordinates may be expressed to some other system such as latitude, longitude and altitude if desired. Solution of the system (equation 9) requires measurement of the pseudoranges to four different satellites. The GNSS receiver's computer may be programmed to solve directly the navigation equations in the form given above. However, the computation time required to solve them may be too long for many applications.
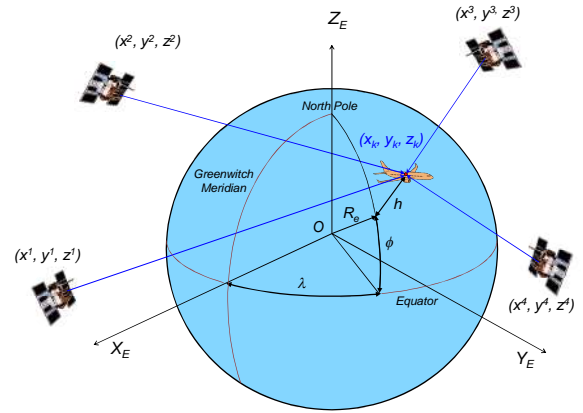
Fig. 2. Navigation solution in the ECEF co-ordinate system (at time zero, the X_E axis passes through the North Pole, and the Y_E axis completes the right-handed orthogonal system.

As an alternate approach, these equations may be approximated by a set of four linear equations that the GPS receiver can solve using a much faster and simpler algorithm. The system of equations (9) can be rewritten in the form:

$$P_k{}^i(t) = \sqrt{(u^i - u_k)^2 + (v^i - v_k)^2 + (w^i - w_k)^2} + T \qquad (10)$$

$$(i = 1, 2, 3, 4)$$

where u^i, v^i and w^i represent the co-ordinates of the ith satellite, $T = dt_k$ and the units have been chosen so that the speed of light is unity. Linearization of equation (10) can proceed as described in [3, 4]. The resulting set of linearized equations relates the pseudorange measurements to the desired user navigation information as well as the user's clock bias:

$$\left(\frac{u_n - u^i}{p_{ni} - T_n}\right)\Delta u_k + \left(\frac{v_n - v^i}{p_{ni} - T_n}\right)\Delta v_k + \left(\frac{w_n - w^i}{p_{ni} - T_n}\right)\Delta w_k + \Delta T = \Delta P_i \quad (11)$$

$$(i = 1, 2, 3, 4)$$

where u_n, v_n, w_n, T_n are the nominal (a priori best-estimate) values of u_k, v_k, w_k and T; $\Delta u_k, \Delta v_k, \Delta w_k, \Delta T$ are the corrections to the nominal values; p_{ni} is the nominal pseudorange measurement to the ith satellite; ΔP_i is the difference between actual and nominal range measurements. The quantities on the right-hand side are simply the differences between the actual measured pseudoranges and the predicted measurements, which are supplied by the user's computer, based on knowledge of the satellite position and current estimate of the user's position and clock bias. Therefore, the quantities to be computed $(\Delta u_k, \Delta v_k, \Delta w_k, \Delta T)$ are the corrections that the user will make to the current estimate of position and clock

time bias. The coefficients of the quantities on the left-hand side represent the direction cosines of the LOS vector from the user to the satellite as projected along the Cartesian co-ordinate system.

2.1.2. Carrier Phase Observable

The carrier phase observable is the difference between the received satellite carrier phase and the phase of the carrier generated by the receiver oscillator. The same error sources that affect pseudoranges, are responsible for the errors which determine the positional accuracy achieved with carrier phases [5]. Clearly, the mathematical formulation of any specific error component is different from the pseudorange case (i.e., phase measurement errors instead of range measurement errors). Since the antenna cannot sense the number of whole carrier waves between the satellite and the receiver (called the integer ambiguity), an extra parameter is inserted in the carrier phase equation:

$$\Phi_k^p = \Phi_k(t) - \Phi^p(t) + N_k^p + I_{k,\Phi}^p(t) - \frac{f}{c} T_k^p(t)$$
$$+ d_{k,\Phi}(t) + d_{k,\Phi}^p(t) + d_k^p(t) + \varepsilon_\Phi \quad (12)$$

The symbols $\Phi_k(t)$ and $\Phi^p(t)$ denote the phase of the receiver generated signal and the phase of the satellite signal respectively, at the epoch t of satellite signal reception. The symbol N_k^p is the integer ambiguity. The terms $I_{k,\Phi}^p(t)$ and $T_k^p(t)$ are the ionospheric and tropospheric delays. The ionospheric delay factor has a negative value because the carrier phase progresses when travelling through the ionosphere. Furthermore, the tropospheric factor is converted in cycles multiplying by f/c, where f is the nominal frequency and c is the speed of light in vacuum. The symbols $d_{k,\Phi}(t)$ and $d_k^p(t)$ refer to the receiver and satellite hardware delays respectively. The symbol $d_{k,\Phi}^p(t)$ denotes the multipath effect and ε_Φ denotes the carrier phase measurement noise. Assuming synchronisation of the satellite and receiver clocks, omitting other error sources (receiver phase tracking circuits, local oscillator, multipath and measurement noise), and taking into account both the time of transmission and reception of the signal, the equation for the phase observation between a satellite i and a receiver A, can be written as [6]:

$$\Phi_A^i(\tau) = \Phi^i(t) - \Phi_A(\tau) \quad (13)$$

where $\Phi_A^i(\tau)$ is the phase reading (phase at receiver A of the signal from satellite i at time τ); $\Phi^i(t)$ is the received signal (phase of the signal as it left the satellite at time t) and $\Phi_A(\tau)$ is the generated signal phase (phase of the receiver's signal at time τ). If $\rho_A^i(t)$ is the range between receiver and satellite, we have:

$$t = \tau - \frac{\rho_A^i(t)}{c} \quad (14)$$

Therefore:

$$\Phi^i(t) = \Phi^i\left(\tau - \frac{\rho_A^i(t)}{c}\right) \quad (15)$$

$$\Phi^i(t) = \Phi^i(\tau) - \underbrace{\frac{\partial \Phi^i(t)}{dt}}_{f} \times \frac{\rho_A^i(t)}{c} + \cdots \quad (16)$$

$$\Phi^i(t) = \Phi^i(\tau) - f \times \frac{\rho_A^i(t)}{c} + \cdots \quad (17)$$

and finally:

$$\Phi_A^i(\tau) = \Phi^i(\tau) - \frac{f}{c}\rho_A^i(t) - \Phi_A(\tau) + N_A^i \quad (18)$$

where $\Phi_A^i(\tau)$ is the phase reading (degree or cycles); $\Phi^i(\tau)$ is the emitted signal; $\frac{f}{c}\rho_A^i(t)$ is the total number of wavelengths; $\Phi_A(\tau)$ is the generated signal and N_A^i is the integer ambiguity. The carrier phase measurement technique typically uses the difference between the carrier phases measured at a reference receiver and a user receiver. This is therefore an inherently differential technique.

2.1.3. Doppler Observable

The equation that associates the transmitted frequency from the satellite with the received frequency is:

$$f_k = \frac{f^p}{1 + \frac{r'}{c}} \quad (19)$$

where f_k is the received frequency, f^p denotes the emitted frequency from the satellite, r' denotes the radial velocity in the satellite-receiver direction and c denotes the speed of light in vacuum. The Doppler frequency shift is given by the difference $f^p - f_k$. The radial velocity r' is the actual rate of change in the satellite-receiver distance and it is given by:

$$r' = -\frac{f_k - f^p}{f_k} \cdot c \quad (20)$$

The integrated Doppler count between two epochs t_1 and t_2 is given by:

$$N_{(t_1, t_2)} = \int_{t_1}^{t_2} (f^p - f_k) dt \quad (21)$$

More information about the Doppler observable and on some of its practical uses can be found in the literature [7, 8].

2.2. GNSS Error Sources

In the following paragraphs, the error sources affecting pseudorange and carrier phase measurements are described. The error sources can be classified into the five broad categories as listed below:

- Receiver Dependent Errors: Clock Error, Noise and Resolution;
- Ephemeris Prediction Errors;
- Satellite Dependent Errors: Clock Offset and Group Delays;
- Propagation Errors: Ionospheric Delay, Tropospheric Delay and Multipath;
- User Dynamics Errors.

2.2.1. Receiver Dependent Errors

Most receivers employ quartz clocks to measure GNSS time, which are not as accurate as the atomic clocks of the satellites. Therefore, there is an offset between the receiver and satellite clocks called receiver clock error. This error affects both the measurement of the signal flight time and the calculation of the satellite's position at time of transmission. Measurement Noise is a random error, which depends entirely on the electronic components of the receiver. Receivers for very precise measurements are designed to minimise this error. Both noise and resolution errors can be reduced by using appropriate filtering techniques. Theoretically, receiver noise can be removed by averaging the measurements, but only over fairly long periods of observation time.

2.2.2. Ephemeris Prediction Errors

The ephemeris data are required for both pseudorange and phase computations. These errors are due to incorrect estimation of the satellites ephemeris at the CPS. A model for evaluating these errors (Fig. 3) is obtained by considering the three components of the vector representing the difference between estimated and true distance: Along Track (ATK), Across Track (XTK) and Radial (RAD). The maximum error is experienced when the satellite has an elevation of 0° on the receiver horizon and the line-of-sight (LOS) user-satellite lies on the geometric plane containing ATK.

In general, the error can be expressed as a function of the three components, in the form:

$$ERR = RAD cos\alpha + ATK sin\alpha\ cos\beta + XTK sin\alpha\ sin\beta \quad (22)$$

where α is the angle between the LOS user-satellite and the satellite vertical and β is the angle between the ATK direction and the plane containing the LOS and the satellite vertical. The US DoD precise ephemeris is calculated from actual observation to the satellites from a network of ground stations distributed around the world. It is produced several days after the observation period and is available only to authorised users. Other non-DoD organisations produce precise ephemeris, both globally and locally, by suitable modelling of all forces and moments acting on the satellites. Orbit relaxation techniques can be developed within GNSS software. These techniques solve for orbital errors in the broadcast ephemeris and produce improved relative positioning [9-12].
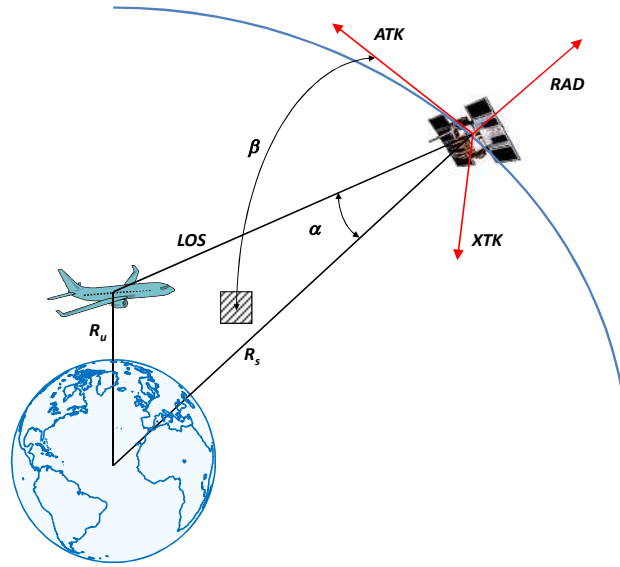
2.2.3. Satellite Dependent Errors

Corrections to the drift of the satellite atomic clocks are computed by the CPS and then broadcasted to the users in the navigation message. The effect of satellite clock offset is negligible in most positioning applications (using the polynomial coefficients corrections computed at the CPS it is possible to reduce this error down to 1 part per 10^{12}). The residual error is due to the fact that corrections from the CPS are periodic and not continuous. Group Delays are the delays typical of the satellite electronic circuits. They are estimated on the ground before the satellites are launched and corrections are included in the navigation message.

Fig. 3. Error components in ephemeris estimation.

2.2.4. Propagation Errors

As discussed, propagation errors include both ionospheric and tropospheric delays. As the satellite signal passes through the ionosphere, it is delayed for two reasons [13]. Firstly, because it travels through a non-vacuum material (propagation delay); thus the Pseudo Random Noise (PRN) codes are delayed, while the carrier phase is advanced when passing through the ionospheric layers. Secondly, because it bends due to refraction; for many applications, the error caused by the bending effect can be considered negligible if the signals are transmitted by satellites with an elevation of $15°$ or more. The ionospheric delay is dependent primarily on the number of electrons that the signal encounters along its propagation path. It is therefore dependent both on the ionosphere characteristics (variable during the day and with seasons), and the path angle (elevation angle of the satellite). It is possible to approximately evaluate the ionospheric delay using the following equation [13]:

$$\Delta\tau = 40.31 \frac{TEC}{cf^2} \quad (23)$$

where TEC is the total electron content integrated along the LOS to the satellite in units of electrons/m², f is the frequency in Hz and c is the speed of light. TEC varies with time and depends on the location of the ionosphere 'pierced' by the LOS to the satellite. At the L1 frequency (1575.42 MHz), equation (23) can be written for a satellite that is not at the zenith [14]:

$$\Delta\tau = 0.162 \frac{F \cdot TEC_v}{c} \quad (24)$$

where TEC_v is the TEC value for a vertical column located at the pierce point, in units of 1016 electrons per m² and F is the obliquity factor. Assuming that the active region of the ionosphere can be represented by a thin shell at an elevation of 350 km, the obliquity can be approximately expressed as a simple function of the elevation angle in degrees (E) of the satellite at the receiver's antenna [14]:

$$F = 1 + 2.74 \cdot 10^{-6}(96 - E)^3 \quad (25)$$

Depending on the receiver design, different models can be adopted to calculate the correction terms to be applied to the pseudorange before solving the navigation equations. Particularly, L1 code-range receivers use a sinusoidal model of the ionosphere (called Klobuchar model), which takes into account the variations of the ionospheric layers (low over night, rapidly getting higher after dawn, getting slightly higher during the afternoon and rapidly getting lower after sunset). The sinusoidal parameters (amplitude and period) are transmitted in the navigation message. The relevant equations are the following [15]:

$$IDV = DC + A\ cos\left[\frac{2\pi(t-\Phi)}{P}\right] \quad (day) \quad (26)$$

$$IDV = DC \quad (night) \quad (27)$$

where IDV (Ionospheric Vertical Delay) is expressed in nsec, DC is the constant night-day offset (5 nsec), A is the amplitude (whose value is between 10 and 100 nsec), Φ is the constant phase offset (14.00 hours), t is the local time, and P is the period. The two factors A and P are transmitted as coefficients of a cubic equation

representing a model of the ionosphere with varying latitude. As the delay also depends on obliquity of the path, elevation is included as an additional factor in the equation:

$$TID = [1 + 16(0.53 - E_h)^3]IDV \qquad (28)$$

where TID is the Total Ionospheric Delay (nsec) and E_h is the elevation angle of the satellites over the horizon. As the ionosphere is dispersive at radio frequencies, two signals at different frequencies will be delayed by different amounts. Therefore, double frequency GNSS receivers (e.g., GPS P-code receivers) can measure the difference (ΔT) between the time of reception of L1 (1575.42 MHz) and L2 (1227.60 MHz), and evaluate the delay associated with both of them. For L1 we have:

$$\Delta\tau_{L1} = \Delta T \left[\left(\frac{f_{L1}}{f_{L2}}\right)^2 - 1\right]^{-1} \qquad (29)$$

where:

$$\Delta T = \frac{40.31 \cdot TEC}{cf^2}\left(\frac{1}{f_{L2}^2} - \frac{1}{f_{L1}^2}\right) \qquad (30)$$

The troposphere is the lower part of the atmosphere (up to about 50 km) and its characteristics depend on local humidity, temperature and altitude. The tropospheric delay for a given slant range can be described as a product of the delay at the zenith and a mapping function, which models the elevation dependence of the propagation delay. In general, the total Slant Tropospheric Delay (STD) is given by the sum of a Slant Hydrostatic Delay (SHD) and a Slant Wet Delay (SWD), and both of them can be expressed by a relevant Zenith Tropospheric Delay (ZTD) and a Mapping Function (MF). The SHD (or "dry component") account for about 90% of the total tropospheric delay and accurate estimations are normally available. On the other hand, the "wet component" (SWD) estimations are generally less accurate and there is a significant variability depending on the actual models implemented. Table 1 shows some typical values of the tropospheric delay for various elevation angles [13].

Table. 1. Tropospheric delay for varying elevation angle.

Elevation Angle [°]	Dry Component [m]	Wet Component [m]
90	2.3	0.2
30	4.6	0.4
10	13.0	1.2
5	26.0	2.3

The total STD is given by:

$$STD = SHD + SWD \qquad (31)$$

$$STD = ZHD \times MF_H + ZWD \times MF_W \qquad (32)$$

where ZHD and ZWD are the zenith hydrostatic delay and zenith wet delay, respectively ($ZTD = ZHD + ZWD$); MF_H and MF_W are their corresponding MFs. There are many ZTD models and MFs currently used in GNSS positioning applications. Popular models include the ZTD empirical models developed by Saastamoinen [16 and 17], Hopfield [18 and 19] and by the University of New Brunswick [20]. Popular MFs include the ones described by Chao [21], Ifadis [22], Herring [23], Niell [24], and Boehm [25].

2.2.5. Multipath Errors

GNSS signals may arrive at the receiver antenna via different paths, due to reflections by objects along the path. Such effect is known as multipath. The reflected signal will have a different path length compared to the direct signal; therefore, it will give a biased distance measurement (Fig. 4). Multipath depends on the environment surrounding the receiver and on the satellite geometry. Typically, multipath will be greater for low elevation

satellites and code multipath is much greater than carrier phase multipath [26]. For code measurements, the multipath error can reach a theoretical value of 1.5 times the chip rate. So, for instance, the GPS C/A code chip rate is 293.1 metre and the maximum multipath error is about 439.65 metres. However, values of less than 2-3 metres are the norm and upper values of 15 metres are rarely observed. For carrier phase, the maximum theoretical multipath error is a quarter of the wavelength. This equates to about 5 centimetres for the GPS L1 and L2 frequencies, although typical values are less than 1 centimetre [27]. Multipath can be accurately modelled and removed only at static points, by taking observations at the same points and at the same hour on consecutive days. This, however, is possible in a dynamic environment. Other techniques use Signal-to-Noise Ratio (SNR) and signal phase/frequency information to detect and quantify multipath.
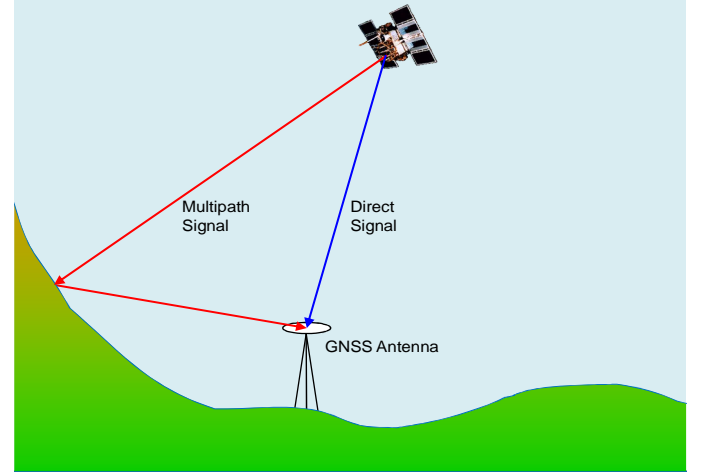

Fig. 4. Multipath error.

Despite several technological advances and research efforts addressing multipath detection and mitigation techniques, this is still a major error source in GNSS applications. Techniques such as the Narrow Correlator [28], the Double Delta/Strobe Correlator [29], or the Vision Correlator by Fenton and Jones [30], are not capable of eliminating the multipath-related errors completely.

2.2.6. User Dynamics Error

There are various errors due to the dynamics of a GNSS receiver. These errors range from the physical masking of the GNSS antenna to the accuracy degradation caused by sudden accelerations of the antenna. If carrier phase is used, the resulting effect of "cycle slips" is the need for re-initialisation (i.e., re-determination of the integer ambiguities). In general, a distinction is made between medium-low and high dynamic platforms.

2.3. UERE Vector

The User Equivalent Range Error (UERE) is an error vector along the line-of-sight user-satellite given by the projection of all system errors. The value of this vector can be reduced only by careful design of the receiver, since there is nothing the users of non-augmented GNSS receivers can do to reduce other error sources. The UERE is generally measured in metres and is given as either a 1-sigma error (often denoted as σ_r) or a 2-sigma error (95%). The portion of the UERE allocated to the space and control segments is called the User Range Error (URE) and is defined at the phase centre of the satellite antenna. The portion of the UERE allocated to the User Equipment is called the UE Error (UEE). Specifically, the UERE is the root-sum-square of the URE and UEE. When SA was on, typical values of the UERE vector for GPS were below 10 m (95%) for P(Y) code receivers and about 33 m (95%) for C/A code receivers. With SA turned off, the UERE of C/A code

receivers is typically less than 20 metres, with the actual value dominated by ionospheric and multipath effects. Dual-frequency receivers (with the capability of removing almost entirely the ionospheric errors), typically experience smaller UEREs. Users of the new modernised GPS civilian signals, as well as future users of GALILEO and BEIDOU, will be able to use multiple frequencies to compensate for the ionospheric errors and thereby to achieve lower UEREs. In GNSS systems the UERE values are strongly dependent on the time elapsed since the last upload from the control segment. As an example, Fig. 5 shows the GPS Standard Positioning Service (SPS) UERE variation with time [31]. In GPS normal operations, the time since last upload is limited to no more than one day. The smallest UERE and best SIS accuracy will generally occur immediately after an upload of fresh NAV message data to a satellite, while the largest UERE and worst SIS accuracy will usually be with the stalest NAV message data just prior to the next upload to that satellite.
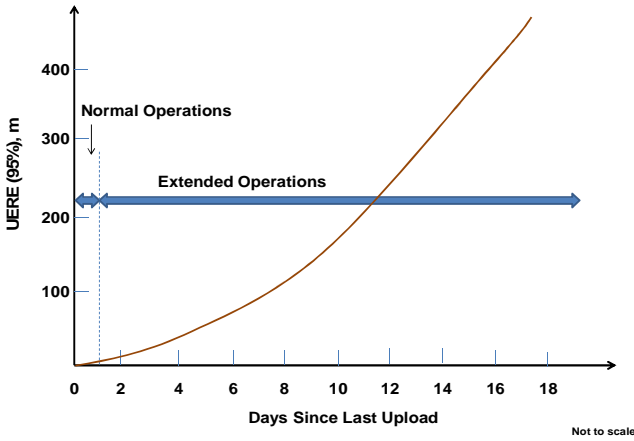

Fig. 5. UERE variation with time in normal and extended operations [31].

The metric used to characterize whether the NAV message data being transmitted by a satellite is fresh or stale is the age of data (AOD), where the AOD is the elapsed time since the control segment generated the satellite clock/ephemeris prediction used to create the NAV message data upload. The AOD is approximately equal to the time since last upload plus the time it took the Control Segment to create the NAV message data and upload it to the satellite. For Normal Operations (NOP), the GPS UERE budget and the traditional SPS SIS accuracy specifications apply at each AOD. Because the largest UERE and worst SIS accuracy usually occur with the oldest NAV message data, the UERE budget and traditional SPS SIS accuracy specifications are taken as applying at the maximum AOD. For reference, the GPS UERE budgets for SPS receivers at zero AOD, at maximum AOD in normal operations, and at 14.5 day AOD in extended operations, are shown in Table 2. The breakout of the individual segment components of the UERE budgets shown in this table is given for illustration purposes only assuming average receiver characteristics. The actual GPS SPS ranging accuracy standards are better specified in terms of the UEE and URE components and the overall error budget is dependent on a number of assumptions, conditions and constraints. Table 3 lists the error budgets and typical UEE values applicable to airborne C/A-code GPS receivers in normal operating conditions. Table 4 lists the condition and criteria that apply to the URE budgeting.

2.4. DOP Factors

Ranging errors alone do not determine GNSS positioning accuracy. The accuracy of the navigation solution is also affected by the relative geometry of the satellites and the user. This is described by the Dilution of Precision (DOP) factors. The four linearized

equations represented by Eq. (9) can be expressed in matrix notation as:

$$\begin{bmatrix} \Delta P_1 \\ \Delta P_2 \\ \Delta P_3 \\ \Delta P_4 \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & 1 \\ \beta_{21} & \beta_{22} & \beta_{23} & 1 \\ \beta_{31} & \beta_{32} & \beta_{33} & 1 \\ \beta_{41} & \beta_{42} & \beta_{43} & 1 \end{bmatrix} \begin{bmatrix} \Delta u_k \\ \Delta v_k \\ \Delta w \\ \Delta T \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix} \quad (33)$$

where an error vector has been added to account for pseudorange measurement noise plus model errors and any unmodelled effects (e.g., SA), and β_{ij} is the direction cosine of the angle between the LOS to the ith satellite and the jth co-ordinate.

Table 2. GPS single frequency C/A code UERE budget.
Adapted from [31].

Segment	Error Source	UERE Contribution [95%][m]		
		Zero AOD	Max AOD in NOP	14.5 Day AOD
Space	Clock Stability	0.0	8.9	257
	Group Delay Stability	3.1	3.1	3.1
	Differential Group Delay Stability	0.0	0.0	0.0
	Satellite Acceleration Uncertainty	0.0	2.0	204
	Other Space Segment Errors	1.0	1.0	1.0
Control	Clock/Ephemeris Estimation	2.0	2.0	2.0
	Clock/Ephemeris Prediction	0.0	6.7	206
	Clock/Ephemeris Curve Fit	0.8	0.8	1.2
	Iono Delay Model Terms	9.8-19.6	9.8-19.6	9.8-19.6
	Group Delay Time Correction	4.5	4.5	4.5
	Other Control Segment Errors	1.0	1.0	1.0
User	Ionospheric Delay Compensation	N/A	N/A	N/A
	Tropospheric Delay Compensation	3.9	3.9	3.9
	Receiver Noise and Resolution	2.9	2.9	2.9
	Multipath	2.4	2.4	2.4
	Other User Segment Errors	1.0	1.0	1.0
95% System UERE (SPS)		**12.7-21.2**	**17.0-24.1**	**388**

Table. 3. Typical UEE error budget (95%). Adapted from [31].

Error Source	Traditional Spec., Single Freq. Rr	Improved Spec., Single Freq. Rr	Modern Single Freq. Rr	Modern Dual Freq. Rr*
Ionospheric Delay Compensation	N/A	N/A	N/A	0.8
Tropospheric Delay Compensation	3.9	4.0	3.9	1.0
Receiver Noise and Resolution	2.9	2.0	2.0	0.4
Multipath	2.4	0.5	0.2	0.2
Other Errors	1.0	1.0	1.0	0.8

UEE [m], 95%	5.5	4.6	4.5	1.6
*Assuming benign conditions [31]				

Table 4. SPS SIS URE accuracy standards. Adapted from [31].

SIS Accuracy Standard	Conditions and Constraints
Single-Frequency C/A-Code: ≤ 7.8 m 95% Global Average URE during Normal Operations over all AODs ≤ 6.0 m 95% Global Average URE during Normal Operations at Zero AOD ≤ 12.8 m 95% Global Average URE during Normal Operations at Any AOD	• For any healthy SPS SIS • Neglecting single-frequency ionospheric delay model errors • Including group delay time correction errors at L1 • Including inter-signal bias (P(Y)-code to C/A-code) errors at L1
Single-Frequency C/A-Code: • ≤ 30 m 99.94% Global Average URE during Normal Operations • ≤ 30 m 99.79% Worst Case Single Point Average URE during Normal Operations	• For any healthy SPS SIS • Neglecting single-frequency ionospheric delay model errors • Including group delay time correction errors at L1 • Including inter-signal bias (P(Y)-code to C/A-code) errors at L1 • Standard based on measurement interval of one year; average of daily values within the service volume • Standard based on 3 service failures per year, lasting no more than 6 hours each
Single-Frequency C/A-Code: • ≤ 388 m 95% Global Average URE during Extended Operations after 14 Days without Upload	• For any healthy SPS SIS

Equation (33) can be written more compactly as:

$$r = B \bar{x} + \epsilon \qquad (34)$$

where B is the 4×4 solution matrix (i.e., matrix of coefficients of the linear equation); $\bar{x}$ is the user position and time correction vector $\bar{x} \equiv [\Delta u_k\ \Delta v_k\ \Delta w\ \Delta T]^T$; r is the pseudorange measurement difference vector ($r \equiv [\Delta P_1\ \Delta P_2\ \Delta P_3\ \Delta P_4]^T$) and ϵ is the vector of measurement and other errors ($\epsilon \equiv [\epsilon_1\ \epsilon_2\ \epsilon_3\ \epsilon_4]^T$).

The GNSS receiver (or post-processing software) solves the matrix equation using least squares or a Kalman Filter (KF). The solution is:

$$\bar{x} = -(B^T W\ B)^{-1} B^T W_r \qquad (35)$$

The new term in the above equation is the weight matrix (W) that characterizes the differences in the errors of the simultaneous measurements as well as any correlations that may exist among them. The weight matrix is given by:

$$W = \sigma_0^2\ C_r \qquad (36)$$

where C_r is the covariance matrix of the pseudorange errors and σ_0^2 is a scale factor known as the a priori variance of unit weight. Applying the law of propagation of error (also known as the covariance law), we obtain:

$$C_{\bar{x}} = [(B^T W\ B)^{-1} B^T\ W]\ C_r\ [(B^T W\ B)^{-1} B^T\ W]^T \qquad (37)$$

$$C_{\bar{x}} = (B^T C_r^{-1}\ B)^{-1} \qquad (38)$$

where $C_{\bar{x}}$ is the covariance matrix of the parameter estimates. If we assume that the measurement and model errors are the same for all

observations with a particular standard deviation (σ) and that they are uncorrelated, we can write:

$$C_r = I\ \sigma^2 \qquad (39)$$

where I is the identity matrix. Therefore, the expression for $C_{\bar{x}}$ simplifies to:

$$C_{\bar{x}} = (B^T B)^{-1}\ \sigma_r^2 = D\ \sigma_r^2 \qquad (40)$$

The term σ_r represents the standard deviation of the pseudorange measurement error plus the residual model error, which is assumed to be equal for all simultaneous observations. If we further assume that the measurement error and the model error components are all independent, then we can simply root-sum-square these errors to obtain the value of σ_r. Based on the definition of UERE given above, we can use the 1-sigma UERE for σ_r. The elements of matrix D are a function of receiver-satellite geometry only. The explicit form of this matrix is:

$$D = \begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} \\ D_{21} & D_{22} & D_{23} & D_{24} \\ D_{31} & D_{32} & D_{33} & D_{34} \\ D_{41} & D_{42} & D_{43} & D_{44} \end{bmatrix} \qquad (41)$$

The DOP factor can be defined as follows:

$$DOP_d = \sqrt{Trace_d\ D} \qquad (d = 1, 2, 3, 4) \qquad (42)$$

where d is the dimension of the DOP factor. The diagonal elements of $C_{\bar{x}}$ are the estimated receiver coordinate and clock-offset variances, and the off-diagonal elements (i.e., covariances) indicate the degree to which these estimates are correlated. The explicit form of the $C_{\bar{x}}$ matrix is:

$$C_{\bar{x}} = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} & \sigma_{uw} & \sigma_{uT} \\ \sigma_{vu} & \sigma_v^2 & \sigma_{vw} & \sigma_{vT} \\ \sigma_{wu} & \sigma_{wv} & \sigma_w^2 & \sigma_{wT} \\ \sigma_{Tu} & \sigma_{Tv} & \sigma_{Tw} & \sigma_T^2 \end{bmatrix} \qquad (43)$$

The various DOP values can be expressed as functions of the diagonal elements of the $C_{\bar{x}}$ matrix or of the D matrix. Converting the Cartesian co-ordinates in matrix form to more convenient local geodetic coordinates, we have:

$$C_{\overline{LG}} = \begin{bmatrix} \sigma_N^2 & \sigma_{NE} & \sigma_{NH} & \sigma_{NT} \\ \sigma_{EN} & \sigma_E^2 & \sigma_{EH} & \sigma_{ET} \\ \sigma_{HN} & \sigma_{HE} & \sigma_H^2 & \sigma_{HT} \\ \sigma_{TN} & \sigma_{TE} & \sigma_{TH} & \sigma_T^2 \end{bmatrix} \qquad (44)$$

Table 5 shows the relationship between the DOP factors and the diagonal elements of the $C_{\overline{LG}}$ matrix and of the D matrix in equation (42). It is also evident that:

$$PDOP = \sqrt{HDOP^2 + VDOP^2} \qquad (45)$$

$$GDOP = \sqrt{PDOP^2 + TDOP^2} \qquad (46)$$

The PDOP is very frequently used in navigation. This is because it directly relates error in GNSS position to errors in pseudo-range measurements to the satellites. According to the definitions given above, the 1-sigma Estimated Position and Time Errors (EPE and ETE) of a GNSS receiver can be calculated using the PDOP (EPE in 3D), the HDOP (EPE in 2D) or the TDOP. In general, we have:

$$EPE_{3D} = \sigma_P = \sigma_r\ PDOP \qquad (47)$$

$$EPE_{2D} = \sigma_H = \sigma_r\ HDOP \qquad (48)$$

$$ETE = \sigma_T = \sigma_r\ TDOP \qquad (49)$$

Table 6 shows the relationship between the EPE3D and the Figure of Merit (FOM). This parameter is frequently used in avionics

GNSS receivers to provide an indication of the quality of information provided by GNSS.

Table 5. DOP expressions.

DOP Factor	D Matrix Formulation	C Matrix Formulation
Geometric DOP (GDOP)	$GDOP = \sqrt{D_{11} + D_{22} + D_{33} + D_{44}}$	$GDOP = \dfrac{1}{\sigma_r}\sqrt{\sigma_N{}^2 + \sigma_E{}^2 + \sigma_H{}^2 + \sigma_T{}^2}$
Position DOP (PDOP)	$PDOP = \sqrt{D_{11} + D_{22} + D_{33}}$	$PDOP = \dfrac{1}{\sigma_r}\sqrt{\sigma_N{}^2 + \sigma_E{}^2 + \sigma_H{}^2}$
Horizontal DOP (HDOP)	$HDOP = \sqrt{D_{11} + D_{22}}$	$HDOP = \dfrac{1}{\sigma_r}\sqrt{\sigma_N{}^2 + \sigma_E{}^2}$
Vertical DOP (VDOP)	$VDOP = \sqrt{D_{33}}$	$VDOP = \dfrac{1}{\sigma_r}\sqrt{\sigma_H{}^2} = \dfrac{\sigma_H}{\sigma_r}$
Time DOP (TDOP)	$TDOP = \sqrt{D_{44}}$	$TDOP = \dfrac{1}{\sigma_r}\sqrt{\sigma_T{}^2} = \dfrac{\sigma_T}{\sigma_r}$

Table 6. GNSS figure of merit.

FOM	Est. Position Error 3-D (m)	Est. Time Error
1	< 25	< 1ns
2	25-50	1ns-10ns
3	50-75	10ns-100ns
4	75-100	100ns-1ms
5	100-200	1ms-10ms
6	200-500	10ms-100ms
7	500-1000	100ms-1ms
8	1000-5000	1ms-10ms
9	> 5000	>10ms

There is a proportionality between the PDOP and the reciprocal value of the volume V of a particular tetrahedron formed by the satellites and the user position (Fig. 6).
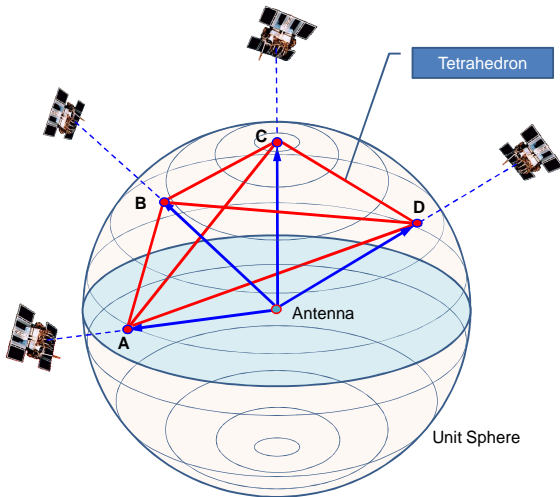

Fig. 6. PDOP tetrahedron construction.

In Fig. 6, four unit-vectors point toward the satellites and the ends of these vectors are connected with 6 line segments. It can in fact be demonstrated that the PDOP is the RSS (i.e., square root of the sum of the squares) of the areas of the 4 faces of the tetrahedron, divided by its volume (i.e., the RSS of the reciprocals of the 4 altitudes of the tetrahedron) [32]. Therefore, we can write:

$$PDOP = \sqrt{\frac{1}{h_A{}^2} + \frac{1}{h_B{}^2} + \frac{1}{h_C{}^2} + \frac{1}{h_D{}^2}} \qquad (50)$$

At a particular time and location, the four satellites shown in Fig. 6, have determined elevations and azimuths with respect to the receiver. From these satellite positions the components (x, y, z) of the unit vectors to the satellites can be determined. Using the Pythagorean theorem, the distances between the ends of the unit vectors can be determined. Knowing these distances, a tetrahedron can be constructed (Fig. 7) and the four altitudes of the tetrahedron can be measured.
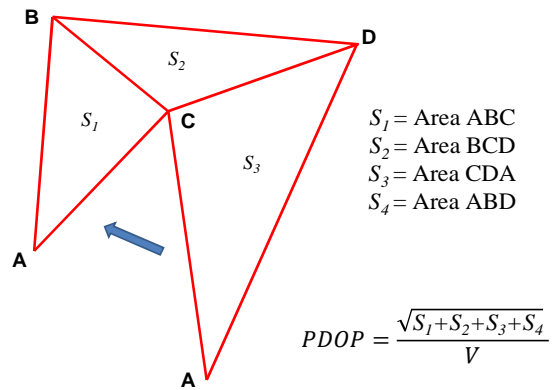

S_1 = Area ABC
S_2 = Area BCD
S_3 = Area CDA
S_4 = Area ABD

$$PDOP = \frac{\sqrt{S_1 + S_2 + S_3 + S_4}}{V}$$

Fig. 7. "Cut and fold" tetrahedron for PDOP determination.

There is no approximation associated with the geometric expression. Any residual error in PDOP determination is only due to construction of the tetrahedron and measurement of the altitudes. If the angle $\widehat{ACA}$ in Fig. 7 is small, one can recognise that the PDOP will be large (poor) even without measuring the altitudes, since this leads to a tetrahedron having a small volume. From the geometric definition of PDOP we deduce that it is optimal from the user's point of view (i.e., low PDOP) to select the four satellites giving a large volume of the tetrahedron (and a small RSS of the areas of the 4 faces of the tetrahedron). This can be obtained primarily by

selecting a combination of satellites far from each other and uniformly distributed around the receiver. The condition for the maximum volume is with a satellite at the zenith and the other three separated by 120° (azimuth) and low over the horizon. This condition, however, would degrade the signal quality, due to the longer propagation paths (i.e., a compromise between signal quality and accuracy of the solution is therefore necessary). A further consequence of the above construction is that if the ends of the unit vectors are coplanar (a not unusual circumstance) the PDOP becomes infinite, and a position is not obtainable. This is the reason for which the various GNSS constellations have been designed so that there are almost always 6-7 satellites in view anywhere on Earth, giving an alternate choice of the 4 satellites to be utilised. If m satellites are in view, the number of possible combinations is:

$$N = \frac{m!}{4!(m-4)!} \tag{51}$$

In most GNSS receivers the number of combinations N also corresponds to the number of PDOP/GDOP computations necessary for selection of the best satellite geometry. Some systems can automatically reject, prior performing positioning calculations, subsets of satellites with associated DOP factors below pre-set thresholds.

2.5. GNSS Performance Requirements in Aviation

The aviation community has devoted great efforts to the rationalization and standardization of the navigation performance parameters and requirements, thus specifying the so-called Required Navigation Performance (RNP) that an airborne navigation system must achieve [33, 34]. Accuracy, integrity, availability and continuity are used to describe the RNP for operations in various flight phases and within specified classes of airspace. The four parameters used to characterize the navigation systems performance are summarized [35]:

- *Accuracy:* The accuracy of an estimated or measured position of a craft (vehicle, aircraft, or vessel) at a given time is the degree of conformance of that position with the true position, velocity and/or time of the craft. Since accuracy is a statistical measure of performance, a statement of navigation system accuracy is meaningless unless it includes a statement of the uncertainty (e.g., confidence level) in position that applies.

- *Integrity:* Integrity is the measure of the trust that can be placed in the correctness of the information supplied by a navigation system. Integrity includes the ability of the system to provide timely warnings to users when the system should not be used for navigation.

- *Continuity:* The continuity of a system is the ability of the total system (comprising all elements necessary to maintain craft position within the defined area) to perform its function without interruption during the intended operation. More specifically, continuity is the probability that the specified system performance will be maintained for the duration of a phase of operation, presuming that the system was available at the beginning of that phase of operation.

- *Availability:* The availability of a navigation system is the percentage of time that the services of the system are usable by the navigator. Availability is an indication of the ability of the system to provide usable service within the specified coverage area. Signal availability is the percentage of time that navigation signals transmitted from external sources are available for use. It is a function of both the physical characteristics of the environment and the technical capabilities of the transmitter facilities.

In aviation applications, accuracy is a measure of the difference between the estimated and the true or desired aircraft position under nominal fault-free conditions. It is normally expressed as 95% bounds on Horizontal and Vertical position errors [34]. In the avionics context, there are two distinguished types of accuracy that must be considered: the accuracy relative to the navigation system alone and the accuracy achieved by the combination of navigation and flight control systems. This second type of accuracy is called Total System Error (TSE) and is measured as deviation from the required flight trajectory. In large commercial aircraft and several military aircraft, the required flight trajectory and associated guidance information is computed by a Flight Management System (FMS) and constantly updated based on ATM and weather information. The navigation system estimates the aircraft state vector (i.e., position, velocity, attitude and associated rates), determines the deviation from the required flight trajectory and sends these information either to the cockpit displays or to an Automatic Flight Control System (AFCS). The error in the estimation of the aircraft's position is referred to as Navigation System Error (NSE), which is the difference between the aircraft's true position and its displayed position. The difference between the desired flight path and the displayed position of the aircraft is called Flight Technical Error (FTE) and accounts for aircraft dynamics, turbulence effects, human-machine interface, etc. As illustrated in Fig. 8, the TSE is obtained as the vector sum of the NSE and the FTE.
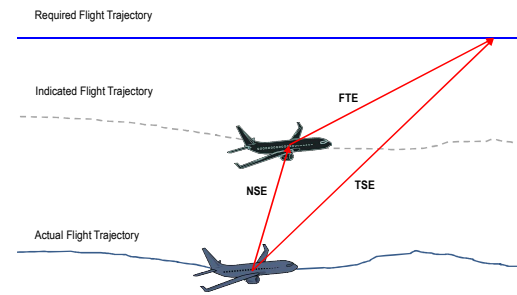

Fig. 8. Navigation accuracy in aviation applications.

The integrity risk can be conveniently defined as the probability of providing a signal that is out of tolerance without warning the user in a given period of time. It is typically derived from the high-level Target Level of Safety (TLS) which is an index, generally determined through historic accident data, against which the calculated risk can be compared to judge whether the operation of the system is safe or not. The navigation integrity requirements in various flight phases/operational tasks have been specified by ICAO in terms of three key performance indicators: the probability of failure over time (or per single approach), the Time-To-Alert (TTA) and the Horizontal/Vertical Alert Limits. The TTA is the maximum allowable time elapsed from the onset of the system being out of tolerance until the equipment enunciates the alert. The alert limits represent the largest horizontal and vertical position errors allowable for safe operations. They are defined as follows [36]:

Horizontal Alert Limit (HAL): the HAL is the radius of a circle in the horizontal plane (the local plane tangent to the WGS-84 ellipsoid), with its center being at the true position, that describes the region that is required to contain the indicated horizontal position with the required probability for a particular navigation mode.

Vertical Alert Limit (VAL): the VAL is half the length of a segment on the vertical axis (perpendicular to the horizontal plane of the WGS-84 ellipsoid), with its center being at the true position, that describes the region that is required to contain the indicated

vertical position with the required probability for a particular navigation mode.

According to these definitions, in order to assess the navigation system integrity, the NSE must be determined first and checked against the Alert Limits (ALs) applicable to the current flight operational phase. However, since the NSE is not observable by the pilot or by the avionics on board the aircraft, another approach to evaluate the system integrity has to be adopted. The standard approach consists in estimating the worst-case NSE and confronting this value with the corresponding AL. These limits for NSE are known as Protection Levels (PLs) and represent high-confidence bounds for the NSE. Similarly to the positioning errors, the PLs are stated in terms of Horizontal and Vertical components (HPL and VPL).

Table 7 lists the required level of accuracy, integrity, continuity and availability defined by ICAO for the different aircraft flight phases. Various applicable national and international standrads have been used in this compilation [37-41].

An additional performance indicator that is often used to characterise avionics GNSS receivers is the Time-To-First-Fix (TTFF). The TTFF accounts for the time elapsed from the GNSS receiver switch-on until the output of a navigation solution is provided to the user within the required performance boundaries (typically in terms of 2D/3D accuracy).

Table 7. Aviation GNSS Signal-in-Space Performance Requirements [37-41].

TYPE OF OPERATION	HOR./VERT. ACCURACY (95%)	CONTINUITY	AVAILABILITY	INTEGRITY	TTA	HAL / VAL
En Route Oceanic	3700m/NA	$1 - 10^{-4}$/hr to $1 - 10^{-8}$/hr	0.99 to 0.99999	$1 - 10^{-7}$/hr	5 min	7408m/NA
En Route Continental	3700m/NA	$1 - 10^{-4}$/hr to $1 - 10^{-8}$/hr	0.99 to 0.99999	$1 - 10^{-7}$/hr	5 min	3704m/NA
En Route Terminal	740m/NA	$1 - 10^{-4}$/hr to $1 - 10^{-8}$/hr	0.99 to 0.99999	$1 - 10^{-7}$/hr	15 s	1852m/NA
APV-I	16m/20 m	$1 - 8 \times 10^{-6}$/15 s	0.99 to 0.999	$1 - 2 \times 10^{-7}$/ App.	10 s	40m/50 m
APV-II	16m/8 m	$1 - 8 \times 10^{-6}$/15 s	0.99 to 0.999	$1 - 2 \times 10^{-7}$/ App.	6 s	40m/20m
Category I	16m/4m	$1 - 8 \times 10^{-6}$/15 s	0.99 to 0.99999	$1 - 2 \times 10^{-7}$/ App.	6 s	40m/10m
Category II	6.9m/2m	$1 - 4 \times 10^{-6}$/15 s	0.99 to 0.99999	$1 - 10^{-9}$/15 s	1 s	17.3m/5.3m
Category III	6.2 m/2 m	$1 - 2 \times 10^{-6}$/30 s (lateral) $1 - 2 \times 10^{-6}$/15 s (vertical)	0.99 to 0.99999	$1 - 10^{-9}$/30 s (lateral) $1 - 10^{-9}$/15 s (vertical)	1 s	15.5m/5.3m

The TTFF is commonly broken down into three more specific scenarios, as defined in the GPS equipment guide:

Cold Start: The receiver has no recent Position, Velocity and Time (PVT) data estimates and no valid almanac data. Therefore, it must systematically search for all possible satellites in the sky. After acquiring a satellite signal, the receiver can obtain the almanac data (approximate information relative to satellites), based on which the process of PVT estimation can start. For instance, in the case of GPS receivers, manufacturers typically claim the factory TTFF to be in the order of 15 minutes.

Warm Start: The receiver has estimates of the current time within 20 seconds, the current position within 100 km, and its velocity within 25 m/s, and it has valid almanac data. Therefore, it must acquire each satellite signal and obtain the satellite's detailed ephemeris data.

Hot Start: The receiver has valid PVT, almanac and ephemeris data, enabling a rapid acquisition of satellite signals. The time required by a receiver in this state to calculate a position fix is also called Time-to-Subsequent-Fix (TTSF). TTSF is particularly important in aviation applications due to frequent satellite data losses caused by high dynamics aircraft manoeuvres.

In aviation applications, GNSS alone does not guarantee the level of performance required in several flight phases and operational tasks. In particular, GNSS fails to deliver sufficient accuracy and integrity levels for mission-critical and safety-critical operations such as precision approach/auto-landing, avionics flight test and flight inspection. Therefore, appropriate augmentation strategies must be implemented to accomplish these challenging tasks using GNSS as the primary source of navigation data.

3. GNSS Threats in Aviation

To meet the stringent SIS performance requirements of GNSS in the aviation context (Table 7), it is essential to detect, isolate and possibly predict GNSS data degradations or signal losses experienced by the aircraft during its mission. This concept applies both to civil and military (manned and unmanned) aircraft, although in different flight and operational tasks. Suitable mathematical models are therefore required to describe the main causes of GNSS signal outages/degradations including Doppler shift, interference/jamming, antenna obscuration, adverse satellite geometries, Carrier-to-Noise ratio (C/N_0) reductions (fading), and multipath (i.e., GNSS signals reflected by the Earth's surface or the aircraft body).

3.1. Antenna Obscuration

Due to the manoeuvres of the aircraft, the wings, tail and fuselage will obscure some satellites during the flight. Fig. 9 shows the GNSS satellite obscuration algorithm.

Taking into account the aircraft shape (e.g. using a CATIA 3-D model), the aircraft flight dynamics (pitch, roll and yaw variations) and the geometric displacement of the GNSS satellites in view, the Antenna Obscuration Matrixes (AOMs) can be generated for the

different flight conditions. Besides the AOM, other factors influence the satellite visibility. In general, a satellite is geometrically visible to the GNSS receiver only if its elevation in the antenna frame is above the Earth horizon and the antenna elevation mask.
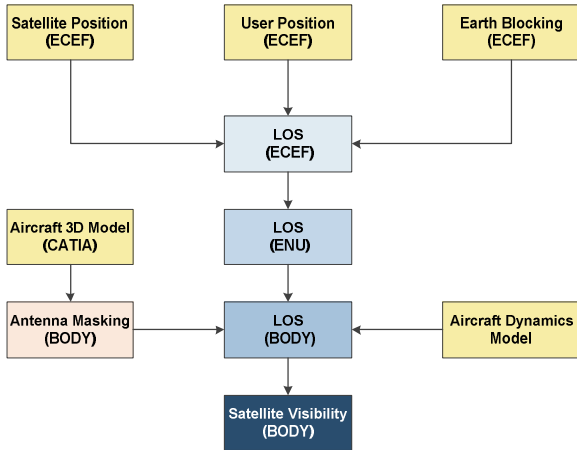

Fig. 9. GNSS satellite obscuration analysis.

It should be noted that even high performance avionics GNSS antennae have gain patterns that are typically below -3dB at about 5 degrees elevation and, as a consequence, their performance become marginal below this limit (Fig. 10).
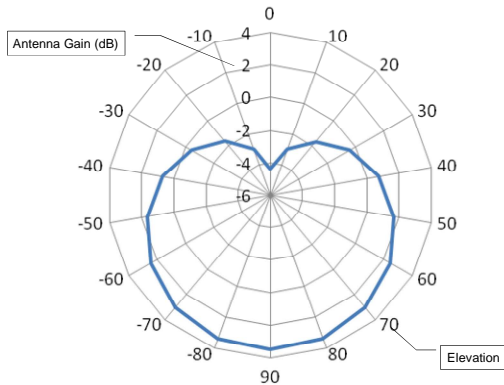

Fig. 10. Avionics antenna gain pattern (L1 frequency).

In order to determine if a satellite is obscured, the LOS of the satellite with respect to the antenna phase centre has to be determined. To calculate the satellite azimuth and elevation with respect to the antenna transformation matrix between ECEF (Earth Centred Earth Fixed) and antenna frame must be applied. This is obtained from:

$$T_E^a = T_b^a * T_N^b * T_E^N \qquad (52)$$

where T_b^a is the transformation matrix between the aircraft body frame and the antenna frame, T_N^b is the transformation matrix from ENU (East-North-Up) to body frame, and T_E^N is the ECEF to ENU transformation matrix.

3.2. GNSS Signal

The received signal strength is affected by a number of factors including transmitter and receiver characteristics, propagation losses and interferences. When necessary, the various factors contributing to the GNSS link budget and signal degradations due to interference can be combine. Multipath induced effects are considered separately. The ratio of total carrier power to noise C/N_0 in dB-Hz is the most generic representation of received signal strength. This is given by:

$$\frac{C}{N_0} = P_t + G_t + G_r - L_s - L_a - L_r - \sigma_m - N_f \quad \text{(dB)} \qquad (53)$$

where:

P_T is the transmitted power level (dBw);

G_t is the satellite antenna gain (dBic);

G_r is the receiver antenna gain toward the satellite;

L_s is the free space loss;

L_a is the atmospheric attenuation (dry-air);

L_r is the rainfall attenuation;

σ_m is the tropospheric fading;

N_f is the receiver noise figure.

As an example, Table 8 shows the expected C/N_0 for a receiver tracking a GPS satellite at zenith given a typical noise power density of -205 dBW-Hz [75]. The values of C/N_0 listed in Table 14 should be compared against the values required to acquire and track GPS signals. These thresholds are heavily dependent on receiver design and for most commercial GNSS receivers they are in the order of 28-33 dB-Hz for acquisition and 25-30 dB-Hz to maintain tracking lock [42, 43]. Current generation avionics GNSS receivers, which are designed to operate in high dynamics conditions, typically exhibit better C/N0 thresholds [44, 45].

Table 8. Nominal receiver GPS signal power and received C/N0 [43].

SV Block IIR-M/IIF	Frequency	P or P(Y)	C/A or L2C
Signal Power	L1	-161.5 dBW	-158.5 dBW
	L2	-161.5 dBW	-160.0 dBW
C/N_0	L1	43.5 dB-Hz	46.5 dB-Hz
	L2	43.5 dB-Hz	45 dB-Hz

The link budget calculated from equation (55) only refers to the direct GNSS signal received from a satellite. Multipath effects, which are due to the geometric and reflective characteristics of the environment surrounding the GNSS antenna are not included in this calculation and are discussed separately. The L-band antenna on-board a GNSS satellite is designed to radiate the composite L-band signals to the users on and near the Earth. In the case of GPS (Fig. 11), the satellite viewing angle from edge-to-edge of the Earth is about 27.7 degrees [46]. The satellite antenna is designed to illuminate the Earth's surface with almost uniform signal strength. The path loss of the signal is a function of the distance from the antenna phase centre to the surface of the earth. The path loss is minimum when the satellite is directly overhead (90° elevation), and is maximum at the edges of the coverage area (satellite at the horizon).
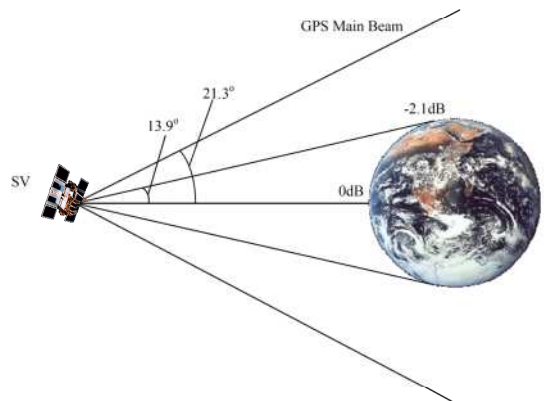
Fig. 11. GPS satellite antenna coverage.

The difference in signal strength caused by this variation in path length is about 2.1 dB and the satellite antenna gain can be approximated by:

$$G_t(dB) = 2.5413 * sinE - 2.5413 \qquad (54)$$

where E is the elevation angle. Similarly, the avionics antenna gain pattern shown in Fig. 8 can be approximated by:

$$G_r(dB) = 9.8756 * sinE - 4.7567 \qquad (55)$$

3.3. Radiofrequency Interference

Intentional and unintentional RF interference (jamming) can result in degraded navigation accuracy or complete loss of the GNSS receiver tracking. Jammers can be classified into three broad categories: Narrowband Jammers (NBJ), Spread Spectrum Jammers (SSJ) and Wideband Gaussian Jammers (WGJ). In mission-essential and safety-critical GNSS applications, both intentional and unintentional jamming must be detected in the GNSS receiver and accounted for in the integrity augmentation architecture. Fortunately, a number of effective jamming detection and anti-jamming (filtering and suppression) techniques have been developed for military GNSS applications and some of them are now available for civil use as well. An overview of these techniques is presented in [49]. The J/S performance of a GNSS receiver at its tracking threshold can be evaluated by the following equation [1]:

$$J/S = 10 \log Q R_c \left[\frac{1}{10^{0.1(C/N_0)_{MIN}}} - \frac{1}{10^{0.1(C/N_0)}} \right] \qquad (56)$$

where Q is the processing gain adjustment factor (1 for NBJ, 1.5 for SSJ and 2 for WGJ), R_c is the code chipping rate (chips/sec) and $(C/N_0)_{MIN}$ is the receiver tracking threshold (dB-Hz). Since the weak limit in an avionics receiver is the carrier tracking loop threshold (typically the PLL), this threshold is usually substituted for $(C/N_0)_{MIN}$. During some experimental flight test activities with unaided L1 C/A code avionics receivers, it was found that in all dynamics conditions explored and in the absence of jamming, a (C/N_0) of 25 dB-Hz was sufficient to keep tracking to the satellites [44, 45]. Using this 25 dB-Hz tracking threshold, the J/S performance of the GPS receiver considering one of the satellites tracked (PRN-14) during the descent manoeuvre was calculated. The calculated C/N_0 for PRN-14 was approximately 37 dB-Hz. Using these J/S values, the minimum range in metres from a jamming source can be calculated from:

$$R_{min} = \frac{\lambda_j}{4\pi} \left(10^{\frac{ERP_{tj} - P_{rj} + G_{rj} - L_{rf}}{20}} \right) \qquad (57)$$

where ERP_{tj} is the effective radiated power of the jammer (dBw), λ_j is the wavelength of jammer frequency (m), P_{rj} is the received (incident) jamming power level at threshold $= J/S + P_{rs}$ (dBw), P_{rs} is the minimum received (incident) signal power (dBw), G_{rj} is the GNSS antenna gain toward jammer (dBic) and L_{rf} is the jammer power attenuation due to receiver front-end filtering (dB).

3.4. Atmospheric Effects

GNSS signal frequencies (L-band) are sufficiently high to keep the ionospheric delay effects relatively small. On the other hand, they are not so high as to suffer severe propagation losses even in rainy conditions. However, the atmosphere causes small but non-negligible effects that must be taken into account. The major effects that the atmosphere has on GNSS signals include [42]:

- Ionospheric group delay/carrier phase advance;
- Tropospheric group delay;
- Ionospheric scintillation;
- Tropospheric attenuation;
- Tropospheric scintillation.

The first two effects have a significant impact on GNSS data accuracy but do not directly affect the received signal strength (C/N_0). Ionospheric scintillation is due to irregularities in the electron density of the Earth's ionosphere (scale size from hundreds of meters to kilometres), producing a variety of local diffraction and refraction effects. These effects cause short-term signal fading, which can severely stress the tracking capabilities of a GNSS receiver. Signal enhancements can also occur for very short periods, but these are not really useful from the GNSS receiver perspective. Atmospheric scintillation effects are more significant in the equatorial and sub-equatorial regions and tend to be less of a factor at European and North-American latitudes.

Signal scintillation is created by fluctuations of the refractive index, which are caused by inhomogeneity in the medium mostly associated to solar activity. At the GNSS receiver location, the signal would exhibit rapid amplitude and phase fluctuations, as well as modifications to its time coherence properties. The most commonly used parameter characterizing the intensity fluctuations is the scintillation index S_4, defined as [47]:

$$S_4 = \sqrt{\frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}} \qquad (58)$$

where I is the intensity of the signal (proportional to the square of the signal amplitude) and $\langle \ \rangle$ denotes averaging. The scintillation index S4 is related to the peak-to-peak fluctuations of the intensity. The exact relationship depends on the distribution of the intensity. According to [47], the intensity distribution is best described by the Nakagami distribution for a wide range of S_4 values. The Nakagami density function for the intensity of the signal is given by:

$$p(I) = \frac{m^m}{\Gamma(m)} I^{m-1} e^{(-mI)} \qquad (59)$$

where the Nakagami "m-coefficient" is related to the scintillation index, S4 by:

$$m = \frac{1}{S_4^2} \qquad (60)$$

In formulating equation (61), the average intensity level of I is normalized to be 1. The calculation of the fraction of time that the signal is above or below a given threshold is greatly facilitated by the fact that the distribution function corresponding to the Nakagami density has a closed form expression which is given by:

$$P(I) = \int_0^I p(x)dx = \frac{\Gamma(m, \ mI)}{\Gamma(m)} \qquad (61)$$

where $\Gamma(m, \ mI)$ and $\Gamma(m)$ are the incomplete gamma function and gamma function, respectively. Using the above equation, it is possible to compute the fraction of time that the signal is above or below a given threshold during ionospheric events. For example, the fraction of time that the signal is more than n dB below the mean is given by $P(10^{-n/10})$ and the fraction of time that the signal is more than m dB above the mean is given by $1 - P(10^{m/10})$. Scintillation strength may, for convenience, be classified into three regimes: weak, moderate or strong [47]. The weak values correspond to $S_4 < 0.3$, the moderate values from 0.3 to 0.6 and the strong case for $S_4 > 0.6$. The relationship between S_4 and the approximate peak-to-peak fluctuations P_{fluc} (dB) can be approximated by:

$$P_{fluc} = 27.5 \times S_4^{1.26} \qquad (62)$$

Table 9 provides a convenient conversion between S_4 and the approximate peak-to-peak fluctuations P_{fluc} (dB).

Table 9. Empirical conversion table for scintillation indices. Adapted from [47].

Scintillation regime	S_4	P_{fluc} [dB]
Weak	0.1	1.5
	0.2	3.5
Moderate	0.3	6
	0.4	8.5
	0.5	11
	0.6	14
Strong	0.7	17
	0.8	20
	0.9	24
	1.0	27.5

Geographically, there are two zones of intense scintillation, one at high latitudes and the other centred within $\pm 20°$ of the magnetic equator as shown in Fig. 12. Severe scintillation has been observed in these two sectors, while in the middle latitudes scintillation occurs exceptionally, such as during geomagnetic storms. In the equatorial sector, there is a pronounced night-time maximum of activity. For equatorial scintillation, peak activity around the vernal equinox and high activity at the autumnal equinox have been observed.
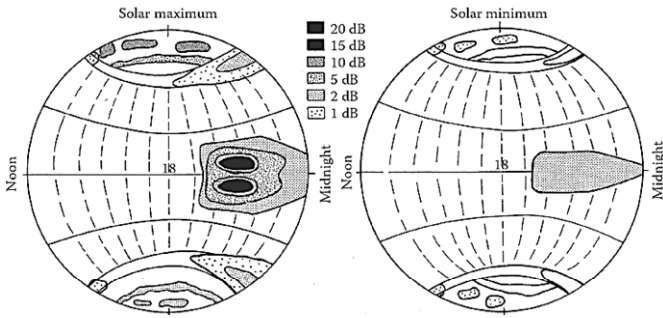

Fig. 12. Depth of scintillation fading at 1.5 GHz during solar maximum and minimum years [47].

In order to predict the intensity of ionospheric scintillation on Earth-space paths, the International Telecommunication Union (ITU) recommend that the Global Ionospheric Scintillation Model (GISM) be used [47]. The GISM permits one to predict the S_4 index, the depth of amplitude fading as well as the rms phase and angular deviations due to scintillation as a function of satellite and ground station locations, date, time and working frequency.

Tropospheric attenuation in the GNSS frequency bands is dominated by oxygen and the effects of other chemical species can be neglected for most applications. Oxygen attenuation (A) is in the order of 0.035 dB for a satellite at zenith and its variation with elevation angle (E) can be approximated by [75]:

$$A(E) \cong \frac{0.07}{\sin E + 0.043} \quad (dB) \quad for\ 3 < E < 10\ deg \quad (63)$$

$$A(E) \cong \frac{0.035}{\sin E} \quad (dB) \quad for\ E > 10\ deg \quad (64)$$

These formulae provide acceptable results only if $E > 3$ degrees. However, since several other errors affects measurements from satellites with elevation below 5 degrees, a software mask is typically employed in avionics GNSS receivers to exclude these satellites form the navigation computations [45]. Tropospheric rainfall attenuation has a minor effect in the GNSS frequency bands. For instance, at a frequency of 2 GHz the attenuation for high rainfall rates is less than 0.01 dB/km (rainfall attenuation

below 2 GHz is even less). Tropospheric scintillation is caused by irregularities (primarily turbulence) causing variations of the refractive index. This effect varies with time, frequency and elevation angle. For small omnidirectional antennas, such as GNSS antennas, the CCIR provided the following expression for the long-term RMS amplitude scintillation [46]:

$$\sigma_m = 0.025 f^{0.58} (csc\ E)^{-0.85} \quad (dB) \quad (65)$$

where f is the frequency in GHz. The Noise Figure (N_f) is related to the system noise temperature (T_{sys}) in Kelvin as follows [48]:

$$N_f = 10 log\left(1 + \frac{T_{sys}}{T_0}\right) \quad (dB) \quad (66)$$

where $T_0 = 290K = 24.6$ dB-K. T_{sys} for antenna plus receiver can be computed using the Friis formula. Typical N_f values for state-of-the-art GPS receivers are between 2 and 4 dB.

3.5. Doppler Shift

Doppler shift is the change in frequency of the received signal that is experienced when the observer (aircraft) moves relative to the signal source (satellite). The magnitude of the frequency shift produced in the signal received from the nth satellite is a function of the relative velocity measured along the satellite-aircraft LOS. The Doppler shift of the nth satellite signal frequency is given by:

$$\Delta f_n = f\left(\frac{|\vec{v_n}| \mp |\vec{v_a}|}{c}\right) cos\chi_n \quad (67)$$

where $\vec{v_n}$ is the nth satellite velocity component along the LOS, $\vec{v_a}$ is the aircraft velocity projection along the LOS, c is the speed of light $[ms^{-1}]$, f is the GNSS signal frequency [Hz] and χ_n is the angle between the aircraft velocity vector and the nth satellite LOS. In a practical aviation application, an optimisation process can be initiated aiming to avoid any further observed increase in Doppler shift. In particular, the presented algorithm studies the variations of $\vec{v_n}$ and $\vec{v_a}$ for each tracked satellite and imposes geometric constraints to the trajectory that are accounted for in the trajectory optimisation process. With reference to the geometry illustrated in Fig. 13, the following trigonometric relationship holds true:

$$(\vec{v}cos\chi_n)sin E_n = \vec{v}cos\chi_n' \quad (68)$$

where $\vec{v}$ is the aircraft velocity vector, E_n is the elevation angle of the nth satellite, χ_n is the relative bearing of the aircraft to the satellite and χ_n' is the azimuth of the LOS projection in the antenna plane.
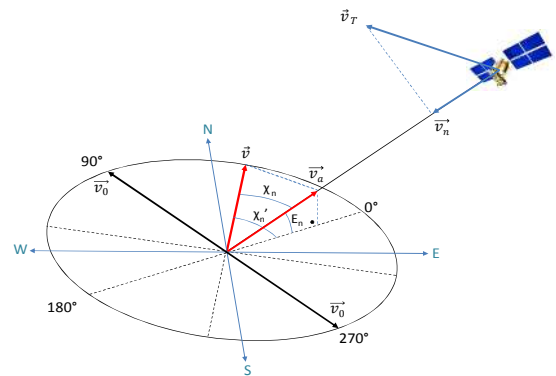

Fig. 13. Reference geometry for Doppler shift analysis.

From equations (67) and (68), the aircraft-satellite relative geometric conditions that maximise or minimise Doppler shift can be determined. In particular, combining the two equations, the Doppler shift is given by:

$$\Delta f_n = f\left(\frac{|\overrightarrow{v_n}| \mp |\overrightarrow{v_a}|}{c}\right)\frac{\cos\chi_{n'}}{\sin E_n} \qquad (69)$$

The above equation shows that both the elevation angle of the satellite and the aircraft relative bearing to the satellite affect the magnitude of the Doppler shift. In particular, reductions of E_n lead to increases in Doppler shift, while χ_n' drives increments or decrements in Doppler shift depending on the size of the angle and the direction of the aircraft velocity vector. By inspecting Fig. 14, it is evident that a relative bearing of 90° and 270° would lead to a null Doppler shift as in this case there is no component of the aircraft velocity vector ($\overrightarrow{v_0}$ in this case) in the LOS to the satellite. However, in all other cases (i.e., $\vec{v} \neq \overrightarrow{v_0}$), such a component would be present and this would lead to increments or decrements in Doppler shift depending on the relative directions of the vectors $\vec{v}$ and $\overrightarrow{v_n}$. This fact is better shown in Fig. 14, where a steady flight without loss of generality is assumed.
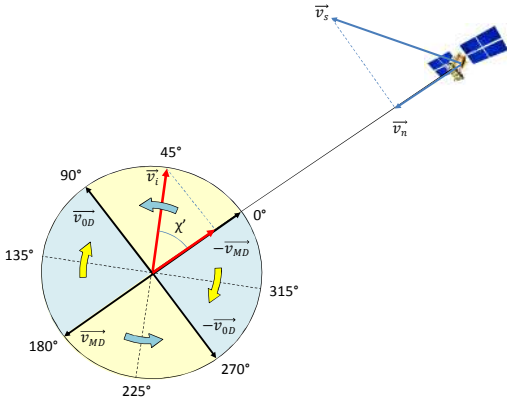

Fig. 14. Reference geometry for Doppler shift manoeuvre corrections.

As the time required for typical aircraft heading change manoeuvres is much shorter than the timeframe associate to significant satellite constellation changes, each satellite can be considered stationary in the body reference frame of the manoeuvring aircraft. Therefore, during manoeuvres leading to Doppler shift, an initial aircraft velocity vector $\overrightarrow{v_l}$ can be considered to define path constraints that avoid further signal degradation or loss. This can be performed by imposing that the aircraft increases the heading rates towards the directions $\overrightarrow{v_0}$ or $-\overrightarrow{v_0}$ and minimises the heading rates towards the directions $\overrightarrow{v_M}$ and $-\overrightarrow{v_M}$. The choice of $\overrightarrow{v_0}$ or $-\overrightarrow{v_0}$ is based simply on the minimum required heading change (i.e., minimum required time to accomplish the correction). Regarding the elevation angle (E) dependency of Doppler shift, equation (69) shows that minimising the elevation angle becomes an objective also in terms of Doppler trajectory optimisation.

3.6. Multipath Analysis

Multipath is caused by the interference of multiple reflections (from the ground and the aircraft structure) with the direct signal transmitted by the satellite and represents a major source of error in GNSS observations. The level and characteristics of multipath depend on the geometry of the environment surrounding the antenna, the reflectivity of nearby objects/terrain and the satellite elevation angle. In order to build a reliable multipath model, a combination of signal analysis and geometric ray-tracing methods was adopted.
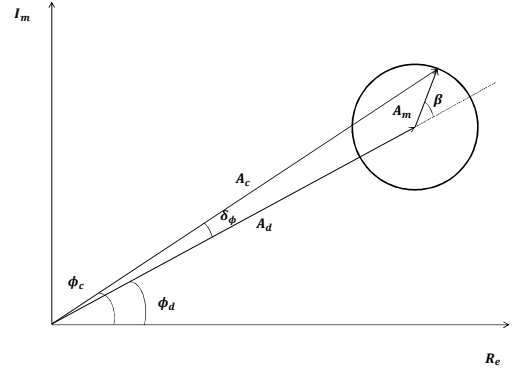

Fig. 15. GNSS signal phases.

From Fig. 15, the SNR and phase error for a single reflection can be represented as a function of direct and multipath signal amplitudes and the multipath relative phase β [49]:

$$SNR = A_c^2 = A_d^2 + A_m^2 + 2A_dA_m\cos\beta \qquad (70)$$

$$tan(\delta_\phi) = \frac{A_m\sin\beta}{A_d + A_m\cos\beta} \qquad (71)$$

where A_d is the direct signal amplitude, A_m is multipath signal amplitude and β is the phase of the multipath. Fig. 16 shows that both the multipath phase β and the multipath amplitude affect the received signal. Therefore, we require a multipath model to simulate these two factors, considering the reflections from the airframe and from the ground. The commonly adopted Aeronautical Multipath Channel (AMC) model developed during the ESA-SDS research [50, 51].
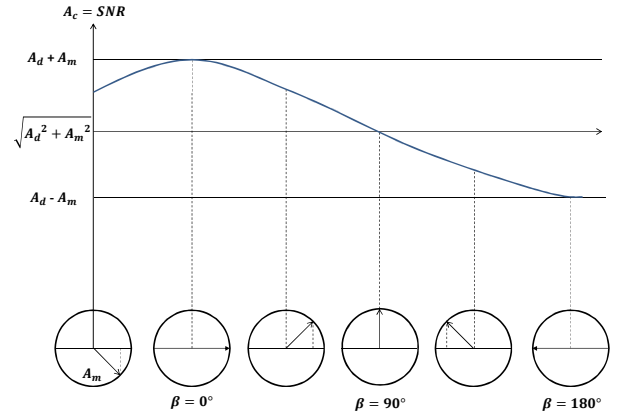

Fig. 16. Variation of A_c as function of the angle β.

Fig. 17 illustrates the overall structure of the AMC model. Let $h(t, \tau)$ be the impulse response of the multipath channel model. Then $h(t, \tau)$ is given by [50]:

$$h(t, \tau) = 1 + \sum_{i=1}^{3}\sqrt{P_i} * n_i(t) * \delta(t - \tau_i) \qquad (72)$$

where P_i is the echo power of the ith path. The signal $n_i(t)$ is a noise signal with power i, and a power spectral density N(f):

$$N(f) = \begin{cases} 0 & f < -B/2 \\ \frac{1}{B} & -B/2 < f < B/2 \\ 0 & f > B/2 \end{cases} \qquad (73)$$

where B is the noise bandwidth. From the multipath channel model in Fig. 17, the wing reflection, the fuselage reflection and the ground-echo are the main components of the multipath signal.
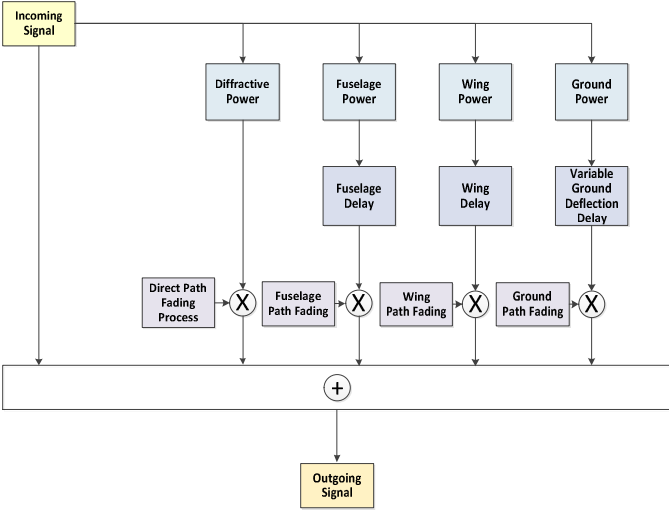
Fig. 17. AMC model structure.

Fig. 18 shows the geometric reflection model. The incoming wave is emitted from point T, R is the receiver location and S is the reflection point. V is a defined point on the reflecting surface and n stands for a unit vector normal to the surface.
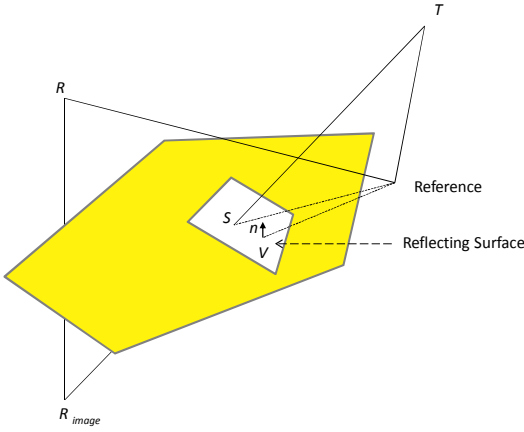

Fig. 18. Geometric reflection model.

In ray-tracing the reflection point S and the defined point V should satisfy the equation:

$$(S - V) \times n = 0 \qquad (74)$$

and the line equation connecting T and R_{image}:

$$S = T + t \times (R_{image} - T) \qquad (75)$$

where t is a parameter between 0 and 1. Combining Eqs. (74) and (75), S is given by:

$$S = T + \frac{n \times V - n \times T}{n \times (R_{image} - T)}(R_{image} - T) \qquad (76)$$

The corresponding extra path length, L_{mS}, due to specular reflection, is then:

$$L_{mS} = |T - S| + |R - S| - |T - R| \qquad (77)$$

In our wing reflection model, the wing is assumed to be flat. By Gaussian Doppler spectrum theory, the power of the wing echo spectrum is assumed to be [52]:

$$P_{Gr(dB)} = 20 * log\left(\frac{1}{\sqrt{2\pi\sigma^2}} * e^{-\frac{f^2}{2\sigma^2}}\right) \qquad (78)$$

where the deviation σ = 3.8 Hz. The wing reflection signal delay can be calculated from:

$$\tau_{wing}(t) = \frac{2*L*sin(E)}{C_0} \qquad (79)$$

where L is the antenna height from the wing, E is the elevation angle (degrees) and C_0 is the speed of light. The fuselage is assumed to be a cylinder and the power of the fuselage echo spectrum is given by [52]:

$$P_{proc}(dB) = 20 * log_{10}[k_1 * e^{(k_2*|f|)} - SNR - k_3] \qquad (80)$$

where k_1, k_2 and k_3 are the fuselage geometric coefficients. Previous research showed that the fuselage reflection characteristics change very little by increasing the fuselage radius [51]. Ground reflection becomes important only during the landing phase, when the aircraft is in close proximity of the terrain. Assuming a Gaussian distributed ground reflection amplitude with zero mean, the ground-echo power is given by:

$$P_{Gr(dB)} = 20 * log_{10}\left(\frac{1}{\sqrt{2\pi\sigma^2}} * e^{-\frac{f^2}{2\sigma^2}}\right) \qquad (81)$$

where, in this case, the deviation σ = 3.8 Hz. Assuming that the terrain is flat:

$$\tau_{ground}(t) = \frac{2*h*sin(E)}{C_0} \qquad (82)$$

where h is the aircraft altitude and E is the elevation angle. Obviously, this basic ground-echo model can be expanded to take into account various terrain and man-made building geometries. As discussed in [42], GNSS receivers can effectively reject most of the multipath signal if the differential delay $\Delta\tau > 1.5$ μs for the C/A code and 0.15 μs for the P(Y) code. As a consequence, the region of potential ground-echo multipath problems for the C/A code is:

$$h * sin(E) < (1.5 \ \mu s) * C_0 = 448.5 \ m \qquad (83)$$

Simulation and flight test activities performed on various aircraft have showed that the fuselage reflections are normally the main contributors to the airframe multipath [53, 54]. In most conditions, the effects of signals reflected from the aircraft wings are comparatively smaller [50, 51]. In particular, it has been found that the airframe multipath ranging error budget can be minimised by placing the GNSS antenna 5 (or more) centimetres above the highest point on the aircraft fuselage. Furthermore, it has been observed that the effect of ground-echo signals translate into a sudden increase of the multipath ranging error of up to two orders of magnitude with respect to the airframe multipath errors alone. During a low-level military aircraft flight trial, it was found that the ground-multipath ranging error reached a value of about 140 metres when the aircraft was flying at an altitude of 300 feet AGL over flat terrain with a roll angle exceeding 45 degrees [44, 45]. It must be pointed out, however, that such particular flight conditions are only likely to be encountered in military aircraft and some unmanned aircraft applications. Due to the flight profile requirements and manoeuvring constraints of typical airliners, the ground multipath contributions can be normally neglected in these cases. According to the Standard Multipath Error Model (SMEM) research [55] and experimental validation activities performed in the US on various types of civil airliners [56], the airframe multipath ranging error, $\sigma_{multipath}$, associated to a satellite observation can be calculated directly as a function of the satellite elevation angle:

$$\sigma_{multipath} = 0.13 + 0.53e^{\left(-\frac{E}{10}\right)} \qquad (84)$$

This model was endorsed by the ICAO GNSS panel and included in the Minimum Operational Performance Standards (MOPS) for the Local Area Augmentation System (LAAS) and for the Wide Area Augmentation System (WAAS) by the Radio Technical Commission for Aeronautics (RTCA) [36, 41, 56, 57].

3.7. Receiver Tracking Errors

A dedicated analysis of the GNSS receiver tracking performance is required to analyse the dynamic stress errors, signal fading (C/N_0 reduction) and interferences (J/S increase). When the GNSS code and/or carrier tracking errors exceed certain thresholds, the receiver loses lock to the satellites. Since both the code and carrier tracking loops are nonlinear, especially near the threshold regions, only Monte Carlo simulations of the GNSS receiver in different dynamics and SNR conditions can determine the receiver tracking performance [58]. Nevertheless, some conservative rule-of-thumbs approximating the measurement errors of the GNSS tracking loops can be employed for this analysis. Numerous sources of measurement errors affect the carrier and code tracking loops. It is important to analyse the dominant error sources in each type of tracking loop. Considering a typical avionics GNSS receiver employing a two-quadrant arctangent discriminator, the Phase Lock Loop (PLL) threshold is given by [1]:

$$3\sigma_{PLL} = 3\sigma_j + \theta_e \leq 45° \qquad (85)$$

where σ_j is the 1-sigma phase jitter from all sources except dynamic stress error and θ_e is the dynamic stress error in the PLL tracking loop. Expanding the above equation, the 1-sigma threshold for the PLL tracking loop becomes [1]:

$$\sigma_{PLL} = \sqrt{\sigma_{tPLL}^2 + \sigma_v^2 + \theta_A^2} + \frac{\theta_e}{3} \leq 15° \qquad (86)$$

where σ_{tPLL} is the 1-sigma thermal noise, σ_v is the variance-induced oscillator phase noise and θ_A is the Allan-variance jitter.

The PLL thermal noise is often thought to be the only carrier tracking error, since the other sources of PLL jitter may be either transient or negligible. The PLL thermal noise jitter is computed as follows:

$$\sigma_{tPLL} = \frac{360}{2\pi}\sqrt{\frac{B_n}{c/n_0}\left(1 + \frac{1}{2T\,c/n_0}\right)} \quad \text{(degrees)} \qquad (87)$$

where B_n is the carrier loop noise bandwidth (Hz), c/n_0 is the carrier to noise power ratio ($c/n_0 = 10^{(C/N_0)/10}$ for C/N_0 expressed in dB-Hz) and T is the predetection integration time (seconds). B_n and C/N_0 can be derived from the SNR model described earlier in this section. Determination of the vibration-induced oscillator phase noise is a complex analysis problem. In some cases, the expected vibration environment is so severe that the reference oscillator must be mounted using vibration isolators in order for the GPS receiver to successfully operate in PLL. The equation for vibration induced oscillator jitter is:

$$\sigma_v = \frac{360 f_L}{2\pi}\sqrt{\int_{f_{min}}^{f_{max}} S_v^2(f_m)\frac{P(f_m)}{f_m^2}df_m} \quad \text{(degrees)} \qquad (88)$$

where f_L is the L-band frequency (Hz), $S_v(f_m)$ is the oscillator vibration sensitivity of $\Delta f/f_L$ per g as a function of f_m, which is the random vibration modulation frequency (Hz), $P(f_m)$ is the power curve of the random vibration as a function of f_m (g^2/Hz) and g is the gravity acceleration. Usually the oscillator vibration sensitivity, $S_v(f_m)$ is not variable over the range of the random vibration modulation frequency, then the above equation can be simplified to:

$$\sigma_v = \frac{360 f_L S_v}{2\pi}\sqrt{\int_{f_{min}}^{f_{max}} \frac{P(f_m)}{f_m^2}df_m} \quad \text{(degrees)} \qquad (89)$$

The equations used to determine Allan deviation phase noise are empirical. They are stated in terms of what the requirements are for the short-term stability of the reference oscillator as determined by the Allan variance method of stability measurement. The equation for second-order loop short-term Allan deviation is:

$$\theta_{A2} = 144\frac{\sigma_A(\tau)*f_L}{B_n} \quad \text{(rad)} \qquad (90)$$

The equation for third-order loop short-term Allan deviation for PLL is:

$$\theta_{A3} = 160\frac{\sigma_A(\tau)*f_L}{B_n} \quad \text{(rad)} \qquad (91)$$

where $\sigma_A(\tau)$ is the Allan deviation-induced jitter (degrees), f_L is the L-band input frequency (Hz), τ is the short-term stability gate time for Allan variance measurement (seconds) and B_n is the noise bandwidth. Usually $\sigma_A(\tau)$ can be determined for the oscillator and it changes very little with gate time τ. For example, assuming the loop filter as a third-order with a noise bandwidth B_n = 18 Hz and the gate time $\tau = 1/B_n$ = 56 ms, the Allan deviation is found to be $\sigma_A(\tau) = 10^{-10}$. The dynamic stress error depends on the loop bandwidth and order. In a third-order loop, the dynamic stress error is [1, 54]:

$$\theta_{e3} = \frac{d^3R/dt^3}{\omega_0^3} = \frac{d^3R/dt^3}{\left(\frac{B_n}{0.7845}\right)^3} = 0.4828\frac{\frac{d^3R}{dt^3}}{B_n^3} \quad \text{(degrees)} \qquad (92)$$

where R is the LOS range to the satellite, d^2R/dt^2 is the maximum LOS acceleration dynamics (°/s^2), ω_0 is the loop filter natural radian frequency and B_n is the noise bandwidth. For the L1 frequency: the term $d^3R/dt^3 = (98/s^3) \times (360°/\text{cycle}) \times (1575.42 \times 10^6 \text{ cycles/s})/c = 185398°/s^3$. Frequency jitter due to thermal noise and dynamic stress error are the main errors in a GNSS receiver Frequency Lock Loop (FLL). The receiver tracking threshold is such that the 3-sigma jitter must not exceed one-fourth of the frequency pull-in range of the FLL discriminator. Therefore, the FLL tracking threshold is [1, 54]:

$$3\sigma_{FLL} = 3\sigma_{tFLL} + f_e \leq 1/4T \quad \text{(Hz)} \qquad (93)$$

where $3\sigma_{tFLL}$ is the 3-sigma thermal noise frequency jitter and f_e is the dynamic stress error in the FLL tracking loop. The above equation shows that the dynamic stress frequency error is a 3-sigma effect and is additive to the thermal noise frequency jitter. The reference oscillator vibration and Allan deviation–induced frequency jitter are small-order effects on the FLL and are considered negligible. The 1-sigma frequency jitter threshold is $1/(12T) = 0.0833/T$ Hz. The FLL tracking loop jitter due to thermal noise is:

$$\sigma_{tFLL} = \frac{1}{2\pi T}\sqrt{\frac{4FB_n}{C/N_0}\left[1 + \frac{1}{Tc/n_0}\right]} \quad \text{(Hz)} \qquad (94)$$

where F is 1 at high C/N_0 and 2 near the threshold. σ_{tFLL} is independent of C/A or P(Y) code modulation and loop order. Since the FLL tracking loop involves one more integrator that the PLL tracking loop of the same order, the dynamic stress error is [54]:

$$f_e = \frac{d}{dt}\left(\frac{1}{360\omega_0^n}\frac{d^nR}{dt^n}\right) = \frac{1}{360\omega_0^n}\frac{d^{n+1}R}{dt^{n+1}} \quad \text{(Hz)} \qquad (95)$$

Regarding the code tracking loop, a conservative rule-of-thumb for the Delay Lock Loop (DLL) tracking threshold is that the 3-sigma value of the jitter due to all sources of loop stress must not exceed the correlator spacing (d), expressed in chips. Therefore [54]:

$$3\sigma_{DLL} = 3\sigma_{tDLL} + R_e \leq d \quad \text{(chips)} \qquad (96)$$

where σ_{tDLL} is the 1-sigma thermal noise code tracking jitter, R_e is the dynamic stress error in the DLL tracking loop. The DLL thermal noise code tracking jitter is given by:

$$\sigma_{tDLL} = \sqrt{\frac{4F_1d^2B_n}{c/n_0}\left[2(1-d) + \frac{4F_2d}{Tc/n_0}\right]} \quad \text{(Hz)} \qquad (97)$$

where F_1 is the DLL discriminator correlator factor (1 for time shared tau-dithered early/late correlator and 0.5 for dedicated early and late correlators), d is the correlator spacing between early, prompt and late, B_n is the code loop noise bandwidth and F_2 is the DLL dicriminator type factor (1 for early/late type discriminator

and 0.5 for dot product type discriminator). The DLL tracking loop dynamic stress error is given by:

$$R_e = \frac{\frac{dR^n}{dt^n}}{\omega_0^n} \quad \text{(chips)} \quad (98)$$

where dR^n / dt^n is expressed in chips/sec^n. The PLL, FLL and DLL error equations described above allow determining the C/N_0 corresponding to the tracking threshold of the receiver. A generic criterion applicable to GNSS augmentation systems is:

$$(C/N_0)_{Threshold} = max[(C/N_0)_{PLL}, (C/N_0)_{FLL}, (C/N_0)_{DLL}] (99)$$

where $(C/N_0)_{PLL}$ is the minimum carrier-to-noise ratio for PLL carrier tracking, $(C/N_0)_{FLL}$ is the is the minimum carrier-to-noise ratio for FLL carrier tracking and $(C/N_0)_{DLL}$ is the minimum carrier-to-noise ratio for DLL code tracking.

4. GNSS Augmentation

In aviation applications, GNSS alone does not guarantee the level of performance required in several flight phases and operational tasks. In particular, GNSS fails to deliver sufficient accuracy and integrity levels for mission safety-critical operations such as precision approach/auto-landing, avionics flight test and flight inspection. Therefore, appropriate augmentation strategies must be implemented to accomplish these challenging tasks using GNSS as the primary source of navigation data. Furthermore, when GNSS is employed as primary means of navigation, some form of augmentation is necessary to satisfy accuracy, integrity, availability and continuity requirements.

4.1. Differential GNSS

Differential GNSS (DGNSS) techniques can be used to address accuracy requirements. Specifically, in the case of GPS, differential techniques have been developed which can provide accuracies comparable with current landing systems. A variety of Local Area DGNSS (LAD) and Wide Area DGNSS (WAD) networks have been proposed over the past twenty years and some LAD/WAD systems have achieved the levels of maturity required to enter the market. Due to the existence of a copious literature on GPS/GNSS basic principles and applications, these will not be covered in this thesis. DGNSS was developed to meet the needs of positioning and distance-measuring applications that required higher accuracies than stand-alone GNSS could deliver. DGNSS involves the use of a control or reference receiver at a known location to measure the systematic GNSS errors; and, by taking advantage of the spatial correlation of the errors, the errors can then be removed from the measurement taken by moving or remote receivers located in the same general vicinity. There have been a wide variety of implementations described for affecting such a DGNSS system. The intent is to characterise various DGNSS systems and to compare their strengths and weaknesses. Two general categories of DGNSS systems can be identified: those that rely primarily upon the code measurements and those that rely primarily upon the carrier phase measurements. Using carrier phase, high accuracy can be obtained (centimetre level), but the solution suffers from integer ambiguity and cycle slips. Whenever a cycle slip occurs, it must be corrected for, and the integer ambiguity must be re-calculated. The pseudorange solution is more robust but less accurate (2 to 5 m). As it is not affected by cycle slips, there is no need for re-initialisation. A typical DGNSS architecture is shown in is shown in Fig. 19.
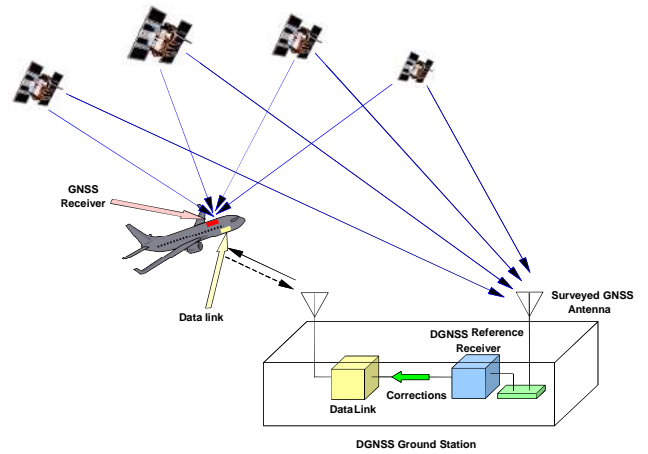

Fig. 19. Typical DGNSS system architecture.

The system consists of a Reference Receiver (RR) located at a known location that has been previously surveyed, and one or more DGNSS User Receivers (URs). The RR antenna, differential correction processing system, and data link equipment (if used) are collectively called the Reference Station (RS). Both the UR and the RR data can be collected and stored for later processing, or sent to the desired location in real time via the data link. DGNSS is based on the principle that receivers in the same vicinity will simultaneously experience common errors on a particular satellite ranging signal. In general, the UR (mobile receiver) uses measurements from the RR to remove the common errors. In order to accomplish this, the UR must simultaneously use a subset or the same set of satellites as the reference station. The DGNSS positioning equations are formulated so that the common errors cancel. The common errors include signal path delays through the atmosphere, and satellite clock and ephemeris errors. For PPS users, the common satellite errors are residual system errors that are normally present in the PVT solution. For SPS users, the common satellite errors also include the intentionally added errors from SA. Errors that are unique to each receiver, such as receiver measurement noise and multipath, cannot be removed without additional recursive processing (by the reference receiver, user receiver, or both) to provide an averaged, smoothed, or filtered solution. Various DGNSS techniques are employed depending on the accuracy desired, where the data processing is to be performed, and whether real-time results are required. If real-time results are required then a data link is also required. For applications without a real-time requirement, the data can be collected and processed later. The accuracy requirements usually dictate which measurements are used and what algorithms are employed. In the case of Differential GPS (DGPS), accuracy is independent of whether SPS or PPS is being used, although real-time PPS DGPS can have a lower data rate than SPS DGPS because the rate of change of the nominal system errors is slower than the rate of change of SA. However, the user and the Reference Station must be using the same service (either PPS or SPS). The clock and frequency biases for a particular satellite will appear the same to all users since these parameters are unaffected by signal propagation or distance from the satellite. The pseudorange and delta-range (Doppler) measurements will be different for different users because they will be at different locations and have different relative velocities with respect to the satellite, but the satellite clock and frequency bias will be common error components of those measurements. The signal propagation delay is truly a common error for receivers in the same location, but as the distance between receivers' increases, this error gradually de-correlates and becomes independent. The satellite ephemeris has errors in all three dimensions. Therefore, part of the error will appear as a common range error and part will remain a residual ephemeris error. The residual portion is normally small and its impact remains small for

similar observation angles to the satellite. The first accepted standard for SPS DGPS was developed by the Radio Technical Commission for Maritime Services (RTCM) Special Committee-104 [59-61]. The data interchange format for NATO users of DGPS is documented in STANAG 4392. The SPS reversionary mode specified in STANAG 4392 is compatible with the RTCM SC-104 standards. The standards are primarily intended for real-time operational use and cover a wide range of DGPS measurement types. Most SPS DGPS receivers are compatible with the RTCM SC-104 differential message formats. DGPS standards for aeronautical use were originally developed by the RTCA allow Special Category-I (SCAT-I) precision approach using range-code differential. These standards were contained in RCTA document DO-217 and intended only for limited use until an international standard could be developed for precision approach [62]. More recently, the two separate standards were developed by RTCA for GPS WAAS/LAAS. These standards define the minimum operational performance requirements for different WAAS/LAAS implementation types allowing WAAS operations down to CAT-I [63] and LAAS operations down to CAT-IIIb [64, 65]. There are two primary variations of the differential measurements and equations. One is based on ranging-code measurements and the other is based on carrier-phase measurements. There are also several ways to implement the data link function. DGNSS systems can be designed to serve a limited area from a single reference station, or can use a network of reference stations and special algorithms to extend the validity of the DGNSS technique over a wide area. The result is that there is a large variety of possible DGNSS system implementations using combinations of these design features.

4.1.1. Ranging-Code DGNSS

Ranging-code differential techniques use the pseudorange measurements of the RS to calculate pseudorange or position corrections for the UR. The RS calculates pseudorange corrections for each visible satellite by subtracting the "true" range determined by the surveyed position and the known orbit parameters from the measured pseudorange. The UR then selects the appropriate correction for each satellite that it is tracking, and subtracts the correction from the pseudorange that it has measured. The mobile receiver must only use those satellites for which corrections have been received. If the RS provides position corrections rather than pseudorange corrections, the corrections are simply determined by subtracting the measured position from the surveyed position. The advantage of using position corrections is obviously the simplicity of the calculations. The disadvantage is that the reference receiver and the user receiver must use the exact same set of satellites. This can be accomplished by coordinating the choice of satellite between the RR and the UR, or by having the RS compute a position correction for each possible combination of satellites. For these reasons, it is usually more flexible and efficient to provide pseudorange corrections rather than position corrections. RTCA, NATO STANAG and RTCM standards are all based on pseudorange rather than position corrections. The pseudorange or position corrections are time tagged with the time that the measurements were taken. In real-time systems, the rate of change of the corrections is also calculated. This allows the user to propagate the corrections to the time that they are actually applied to the user position solution. This reduces the impact of data latency on the accuracy of the system but does not eliminate it entirely. SPS corrections become fully uncorrelated with the user measurements after about 2 minutes. Corrections used after two minutes may produce solutions which are less accurate than stand-alone SPS GPS. PPS corrections can remain correlated with the user measurements for 10 minutes or more under benign (slowly changing) ionospheric conditions. There are two ways of pseudorange data processing: post-mission and real-time processing. The advantage of the post-mission solution over the

real-time one is that it is more accurate because the user can easily detect blunders and analyse the residuals of the solution. On the other hand the main disadvantage of the post-mission solution is that the results are not available immediately for navigation. The typical algorithm of the ranging-code DGNSS post-processed solution used in flight test and inspection tasks is the double difference pseudorange.

Fig. 20 shows the possible pseudorange measurements between two receivers (k and l) and two satellites (p and q). If pseudoranges 1 and 2 from Fig. 22 are differenced, then the satellite clock error and satellite orbit errors will be removed. If SA is active (not the present case), it will be removed completely only if the signals utilised in each receiver are transmitted exactly at the same time. The residual error from SA is not a problem for post-processed positioning, where it is easy to ensure that the differencing is done between pseudoranges observed at the same time. Any atmospheric errors will also be reduced significantly with single differencing. The basic mathematical model for single difference pseudorange observation is the following:

$$P_k^p - P_i^p = \rho_k^p - \rho_l^p - (dt_k - dt_i)c + d_{k,p} - d_{i,p}$$
$$+ d_{k,p}^p - d_{l,p}^p + \Delta \varepsilon_p \qquad (104)$$

where P_i^p is the pseudorange measurement, ρ_l^p denotes the geometric distance between the stations and satellite, dt_i denotes the receiver's clock offsets, $d_{i,p}$ denotes the receiver's hardware code delays, $d_{l,p}^p$ denotes the multipath of the codes, $\Delta \varepsilon_p$ denotes the measurement noise and c is the velocity of light.
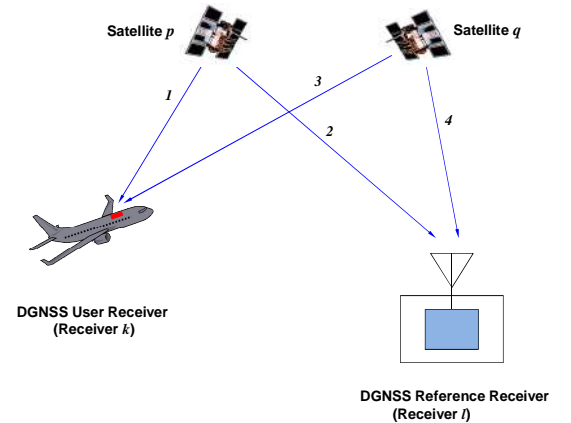

Fig. 20. Pseudorange Differencing.

Equation (104) represents the single difference pseudorange observable between receivers. There are four unknowns in the above equation assuming that the co-ordinates of station k are known and that the difference in clock drifts is one unknown. Hence, four satellites are required to provide four single difference equations in order to solve for the unknowns. Single differences with code observations are frequently used in relative (differential) navigation. Using all pseudoranges shown in Fig. 22, differences are formed between receivers and satellites. Double differences are constructed by taking two between-receiver single differences and differencing these between two satellites. This procedure removes all satellite dependent, receiver dependent and most of the atmospheric errors (if the distance between the two receivers is not too large). The derived equation is:

$$P_k^p - P_k^q - P_l^p - P_l^q = \rho_k^p - \rho_k^q - \rho_l^p - \rho_l^q + d_{i,p}^j \quad (105)$$

where $d_{i,p}^j$ denotes the total effect of multipath. There are three unknowns in the above equation (the co-ordinates of station l) and a minimum of four satellites is required to form a minimum of three

double difference equations in order to solve for the unknowns. Using the propagation of errors law, it is shown that the double difference observables are twice as noisy as the pure pseudoranges:

$$\sigma_{DD} = \sqrt{\sigma_P^2 + \sigma_P^2 + \sigma_P^2 + \sigma_P^2} = 2\sigma_P \qquad (106)$$

but they are more accurate, because most of the errors are removed. It is to be note that multipath remains, because it cannot be modelled and it is independent for each receiver.

4.1.2. Carrier-phase DGNSS

Carrier-phase DGNSS techniques use the difference between the carrier phases measured at the RR and UR. A double-differencing technique is used to remove the satellite and receiver clock errors. The first difference is the difference between the phase measurement at the UR and the RR for a single satellite. This eliminates the satellite clock error which is common to both measurements. This process is then repeated for a second satellite. A second difference is then formed by subtracting the first difference for the first satellite from the first difference for the second satellite. This eliminates both receiver clock errors which are common to the first difference equations. This process is repeated for two pairs of satellites resulting in three double-differenced measurements that can be solved for the difference between the reference station and user receiver locations. This is inherently a relative positioning technique; therefore, the user receiver must know the reference station location to determine its absolute position.

The single difference observable is the instantaneous phase difference between two receivers and one satellite. It is also possible to define single differences between two satellites and one receiver. Using the basic definition of carrier-phase observable presented above, the phase difference between the two receivers A and B, and satellite i is given by:

$$\Phi_{AB}^i(\tau) = \Phi_B^i(\tau) - \Phi_A^i(\tau) \qquad (107)$$

and can be expressed as:

$$\Phi_{AB}^i(\tau) = \left(\frac{f}{c}\right) \cdot \rho_{AB}^i(t) + \Phi_{AB}^i(\tau) - N_{AB}^i \qquad (108)$$

where $\Phi_{AB}(\tau) = \Phi_A(\tau) - \Phi_B(\tau)$, $N_{AB}^i = N_B^i - N_A^i$ and $\rho_{AB}^i(t) = \rho_A^i(t) - \rho_B^i(t)$

Hence, with four satellites i, j, k and l:

$$\Phi_{AB}^i(\tau) = \left(\frac{f}{c}\right) \cdot \rho_{AB}^i(t) + \Phi_{AB}^i(\tau) - N_{AB}^i \qquad (109)$$

$$\Phi_{AB}^j(\tau) = \left(\frac{f}{c}\right) \cdot \rho_{AB}^j(t) + \Phi_{AB}^j(\tau) - N_{AB}^j \qquad (110)$$

$$\Phi_{AB}^k(\tau) = \left(\frac{f}{c}\right) \cdot \rho_{AB}^k(t) + \Phi_{AB}^k(\tau) - N_{AB}^k \qquad (111)$$

$$\Phi_{AB}^l(\tau) = \left(\frac{f}{c}\right) \cdot \rho_{AB}^l(t) + \Phi_{AB}^l(\tau) - N_{AB}^l \qquad (112)$$

The double difference is formed from subtracting two single differences measured to two satellites i and j. The basic double difference equation is:

$$\Phi_{AB}^{ij}(\tau) = \Phi_{AB}^j(\tau) - \Phi_{AB}^i(\tau) \qquad (113)$$

which simplifies to:

$$\Phi_{AB}^{ij}(\tau) = \left(\frac{f}{c}\right)\rho_{AB}^{ij}(t) - N_{AB}^{ij} \qquad (114)$$

where $N_{AB}^{ij} = N_{AB}^j - N_{AB}^i$, and the only unknowns being the double-difference phase ambiguity N_{AB}^{ij} and the receiver co-ordinates. The local clock error is differenced out. Two receivers and four satellites i, j, k and l will give 3 double difference equations containing the co-ordinates of the receiver A (X_A, Y_A, Z_A) and B

(X_B, Y_B, Z_B), and the unknown integer ambiguities N_{AB}^{ij}, N_{AB}^{ik}, and N_{AB}^{il}:

$$\Phi_{AB}^{ij}(\tau) = \left(\frac{f}{c}\right)\rho_{AB}^{ij}(t) - N_{AB}^{ij} \qquad (115)$$

$$\Phi_{AB}^{ik}(\tau) = \left(\frac{f}{c}\right)\rho_{AB}^{ik}(t) - N_{AB}^{ik} \qquad (116)$$

$$\Phi_{AB}^{il}(\tau) = \left(\frac{f}{c}\right)\rho_{AB}^{il}(t) - N_{AB}^{il} \qquad (117)$$

Therefore, the double difference observation equation can be written as [6]:

$$\frac{\partial \Phi}{\partial X_A} dX_A + \frac{\partial \Phi}{\partial Y_A} dY_A + \frac{\partial \Phi}{\partial Z_A} dZ_A + \frac{\partial \Phi}{\partial X_B} dX_B + \frac{\partial \Phi}{\partial Y_B} dY_B + \frac{\partial \Phi}{\partial Z_B} dZ_B$$

$$+ \frac{\partial \Phi}{\partial N_1} dN_1 + \frac{\partial \Phi}{\partial N_2} dN_2 + \frac{\partial \Phi}{\partial N_3} dN_3 + \frac{\partial \Phi}{\partial C} dC + \cdots =$$

$$(\Phi^O - \Phi^c) + v \qquad (118)$$

where (X_A, Y_A, Z_A) are the co-ordinates of receiver A, (X_B, Y_B, Z_B) are the co-ordinates of receiver B, (N_1, N_2, N_3) are the integer ambiguities, C is correction term, $(\Phi^O - \Phi^c)$ is the observed minus the computed observable and v is the residual. From equation (113), the unknown receiver co-ordinates can be computed. It is necessary, however, to determine the carrier phase integer ambiguities (i.e. the integer number of complete wavelengths between the receiver and satellites). In surveying applications, this integer ambiguity can be resolved by starting with the mobile receiver antenna within a wavelength of the reference receiver antenna. Both receivers start with the same integer ambiguity, so the difference is zero and drops out of the double-difference equations. Thereafter, the phase shift that the mobile receiver observes (whole cycles) is the integer phase difference between the two receivers. An alternative is to place the mobile receiver at a surveyed location. In this case the initial difference is not necessarily zero but it is readily calculated. For aviation applications where it is not possible to bring the reference/mobile antennas together or to position the aircraft at a surveyed location, the reference and mobile receivers must solve for the ambiguities independently as part of an initialisation process. For these applications, it is essential to be able to solve for integer ambiguity at an unknown location and while in motion. In this case, solving for the integer ambiguity usually consists of eliminating incorrect solutions until the correct solution is found. A good initial estimate of position (such as from ranging-code differential) helps to keep the initial number of candidate solutions small. Redundant measurements over time and/or from extra satellite signals are used to isolate the correct solution. This version of the carrier-phase DGNSS technique is typically called Real-Time Kinematic (RTK) GNSS. If carrier track or phase lock on a satellite is interrupted (cycle slip) and the integer count is lost, then the initialisation process must be repeated for that satellite. Causes of cycle slips can range from physical obstruction of the antenna to sudden accelerations of the aircraft. Output data flow may also be interrupted if the receiver is not collecting redundant measurements form extra satellites to maintain the position solution. If a precise position solution is maintained, re-initialisation for the lost satellite can be very rapid. Developing a robust and rapid method of initialisation and re-initialisation is the primary challenge facing designers of RTK systems that have to deal with safety critical applications such as aircraft precision approach. A description of techniques for solving ambiguities both in real-time and post-processing applications, together with information about cycle slips repair techniques can be found in the literature [66-78].

4.2. Data Link Implementations

DGNSS can be implemented in several different ways depending on the type of data link used. The simplest way is no data link at

all. For non-real-time applications, the measurements can be stored in the receiver or on suitable media and processed at a later time. In most cases to achieve surveying accuracies, the data must be post-processed using precise ephemeris data that is only available after the survey data has been collected. Similarly, for some test applications the cost and effort to maintain a real-time data link may be unnecessary. Nevertheless, low-precision real-time outputs can be useful to confirm that a test is progressing properly even if the accuracy of the results will be enhanced later. Differential corrections or measurements can be uplinked in real-time from the reference station to the users. This is the most common technique where a large number of users must be served in real-time. For military purposes and proprietary commercial services, the uplink can be encrypted to restrict the use of the DGNSS signals to a selected group of users. Differential corrections can be transmitted to the user at different frequencies. With the exception of satellite data links there is generally a trade-off between the range of the system and the update rate of the corrections. As an example Table 10 lists a number of frequency bands, the range, and the rate at which the corrections could be updated using the standard RTCM SC-104 format [59-61].

Table 10. Possible DGNSS data link frequencies.

Frequency	Range [Km]	Update Rate [sec]
LF (30 - 300 kHz)	> 700	< 20
MF (300 kHz - 3 MHz)	< 500	5 – 10
HF (3 MHz - 25 MHz)	< 200	5
VHF (30 MHz - 300 MHz)	< 100	< 5
L Band (1 GHz - 2 GHz)	Line of Sight	Few Seconds

An uplink can be a separate transmitter/receiver system or the DGNSS signals can be superimposed on a GPS-link L-band ranging signal. The uplink acts as a pseudo-satellite or "pseudolite" and delivers the ranging signal and DGNSS data via the RF section of the user receiver, much in the same way the GPS navigation message is transmitted. The advantages are that the additional ranging signal(s) can increase the availability of the position solution and decrease carrier-phase initialisation time. However, the RS and UR's become more complex, and the system has a very short range (a few kilometres at the most). This is not only because of the line of sight restriction, but also the power must be kept low in order to avoid interference with the real satellite signals (i.e., the pseudolite can become a GPS jammer if it overpowers the GPS satellite signals). A downlink option is also possible from the users to the RS or other central collection point. In this case the differential solutions are all calculated at a central location. This is often the case for test range applications where precise vehicle tracking is desired but the information is not used aboard the vehicle. The downlink data can be position data plus the satellite tracked, or pseudorange and deltarange measurements, or it can be the raw GPS signals translated to an intermediate frequency. The translator method can often be the least expensive with respect to user equipment, and therefore is often used in munitions testing where the user equipment may be expendable.

4.2.1. DGNSS Accuracy

Controlled tests and recent extensive operational use of DGNSS, have repeatedly demonstrated that DGNSS (pseudorange) provides accuracies in the order of 3 to 10 metres. This figure is largely irrespective of receiver type, whether or not SA is in use, and over distances of up to 100 NM from the reference station [45, 79]. With Kinematic DGNSS (KDG) positioning systems, requiring the resolution of the carrier phase integer ambiguities whilst on the

move, centimetre level accuracy can be achieved [6, 80-83]. Most current applications of DGNSS use L1 C/A code pseudorange as the only observable, with achieved accuracies of 1 to 5 m in real-time. Other applications use dual frequency pseudoranges (e.g., C/A and P code from GPS) or combinations of pseudorange and carrier phase observables. Various techniques have been developed that achieve increased accuracy at the expense of increased complexity. A possible classification scheme of these techniques is presented in Table 11.

Precise DGNSS (PDG) can achieve accuracies below 1 m using dual-frequency psudorange measurements and Very Precise DGNSS (VPDG), taking advantage of both dual frequency code and carrier phase observables, is capable of On-The-Fly (OTF) ambiguity resolution [69, 70]. At the moment, Ultra-Precise DGNSS (UPDG) using carrier phase observables with integer ambiguities resolved is only utilised in flight testing, flight inspection and other non-real-time aviation applications. In the absence of Selective Availability (SA), the major sources of error for stand-alone ranging-code GNSS are:

- Ephemeris Error;
- Ionospheric Propagation Delay;
- Tropospheric Propagation Delay;
- Satellite Clock Drift;
- Receiver Noise and clock drift;
- Multipath.
-

Table 11. Classification of DGNSS techniques.

Name	Description	RMS
DGNSS	Single frequency pseudorange (e.g., GPS C/A code)	1 - 5 m
PDG	Dual frequency pseudorange (e.g., GPS P code)	0.1 - 1 m
VPDG	Addition of dual band carrier phase	5 - 30 cm
UPDG	Above with integer ambiguities resolved	< 2 cm

Table 12 lists the error budgets from the above sources giving an estimation of the possible improvements provided by L1 ranging-code DGNSS [82]. The error from multipath is site dependent and the value in Table 12 is only an example. The receiver clock drift is not mentioned in Table 12, because it is usually treated as an extra parameter and corrected in the standard solution. Furthermore, it does not significantly add to differential errors. Multipath and receiver noise errors cannot be corrected by DGNSS. It is worth to mention that the errors introduced by Selective Availability (SA), which is currently off, would be corrected by DGNSS techniques. For users near the reference station, the respective signal paths to the satellite are close enough together that the compensation of Ionospheric and Tropospheric Delays (ITD) is almost complete. As the user to RS separation is increased, the different ionospheric and tropospheric paths to the satellites can be far enough apart that the ionospheric and tropospheric delays are no longer common errors. Thus, as the distance between the RS and user receiver increases the effectiveness of the atmospheric delay corrections decreases.

Table 12. Error sources in DGNSS.

Error Source	Stand Alone GNSS [m]	L1 C/A Code DGNSS [m]

Ephemeris	5 - 20	0 – 1
Ionosphere	15 - 20	2 – 3
Troposphere	3 - 4	1
Satellite Clock	3	0
Multipath	2	2
Receiver Noise	2	2

The ephemeris error is effectively compensated unless it has quite a large out-of-range component (e.g., 1000 metres or more due to an error in a satellite navigation message). Even then, the error will be small if the distance between the reference receiver and user receiver is small. Finally, the satellite clock error is compensated as long as both reference and user receivers employ the same satellite clock correction data. As already mentioned, the correlation of the errors experienced at the RS and the user location is largely dependent on the distance between them. As the separation of the user from the RS increases, so does the probability of significant differing ionospheric and tropospheric conditions at the two sites. Similarly, the increasing separation also means that a different geometrical component of the ephemeris error is seen by the RR and UR. This is commonly referred to as "Spatial Decorrelation" of the ephemeris and atmospheric errors. In general, the errors are likely to remain highly correlated for users within 350 km of the RS. However, if the distance is greater than 250 km the user will obtain better results using correction models for ionospheric and tropospheric delay [45, 79]. Additionally, practical DGNSS systems are typically limited by the data link to an effective range of around 170 km. Table 13 shows the error budget determined for a SPS DGPS system with increasing distances from the RR [45].

Table 13. DGNSS L1 pseudorange errors with increasing distance from RS.

Error Sources	0 NM	100 NM	500 NM	1000 NM
Space Segment: Clock Errors (ft)	0	0	0	0
Control Segment: Ephemeris Errors (ft)	0	0.3	1.5	3
Propagation Errors: Ionosphere (ft)	0	7.2	16	21
Troposphere (ft)	0	6	6	6
TOTAL (RMS)	**0**	**9.4**	**17**	**22**
User Segment: Receiver Noise (ft)	3	3	3	3
Multipath (ft)	0	0	0	0
UERE (ft RMS)	**3**	**9.8**	**17.4**	**22.2**

Since the RR noise and multipath errors are included in the differential corrections and become part of the user's error budget (root-sum-squared with the user receiver noise and multipath errors), the receiver noise and multipath error components in the non-differential receiver can be lower than the correspondent error components experienced in the DGNSS implementation. Another type of error introduced in real-time DGNSS positioning systems is the data link's "age of the corrections". This error is introduced due to the latency of the transmitted corrections (i.e., the transmitted corrections of epoch t_0 arrive at the moving receiver at epoch. These corrections are not the correct ones, because they were calculated under different conditions. Hence, the co-ordinates of the UR would be slightly offset.

DGNSS systems that compensate for accuracy degradations over short distances (typically up to the RS-UR Line-of-Sight (LOS) employing V/UHF datalinks) are referred to as Local Area DGNSS (LAD). DGNSS systems covering large geographic areas are referred to as Wide Area DGNSS (WAD) systems. They usually employ a network of reference receivers that are coordinated to provide DGNSS data over a wide coverage area. Such systems typically are designed to broadcast the DGNSS data via satellite, although a network of ground transmission sites is also feasible. A user receiver typically must employ special algorithms to derive the ionospheric and tropospheric corrections that are appropriate for its location from the observations taken at the various reference sites.

Various countries including the United States, Canada, Europe, Japan, China, India and Australia have developed LAD/WAD systems. Some of these systems transmit DGNSS data from geostationary satellites for use by commercial navigation users, including the aviation community. Some commercial DGNSS services also broadcast data from multiple reference stations via satellite. However, several such systems remain a group of LAD rather than WAD systems. This is because the reference stations are not integrated into a network allowing the user accuracy to degrade with distance from the individual reference sites.

4.3. Real Time Kinematics (RTK) and Precise Point Positioning

For aviation applications where it is not possible to bring the reference/mobile antennas together or to position the aircraft at a surveyed location, the reference and mobile receivers must solve for the ambiguities independently as part of an initialisation process. For these applications, it is essential to be able to solve for integer ambiguity at an unknown location and while in motion. In this case, solving for the integer ambiguity usually consists of eliminating incorrect solutions until the correct solution is found. A good initial estimate of position (such as from ranging-code differential) helps to keep the initial number of candidate solutions small. Redundant measurements over time and/or from extra satellite signals are used to isolate the correct solution. This version of the carrier-phase DGNSS technique is typically called Real-Time Kinematic (RTK) GNSS. If carrier track or phase lock on a satellite is interrupted (cycle slip) and the integer count is lost, then the initialisation process must be repeated for that satellite. Causes of cycle slips can range from physical obstruction of the antenna to sudden accelerations of the aircraft. Output data flow may also be interrupted if the receiver is not collecting redundant measurements form extra satellites to maintain the position solution. If a precise position solution is maintained, re-initialisation for the lost satellite can be very rapid. Developing a robust and rapid method of initialisation and re-initialisation is the primary challenge facing designers of RTK systems that have to deal with safety critical applications such as aircraft precision approach.

Although centimetre-level GNSS accuracy is increasingly relevant for a number of aviation and other Safety-of-Life (SoL) applications, the possible adoption of RTK techniques in the aviation context is still being researched and no RTK systems are contemplated by the current ICAO standards for air navigation. From an operational perspective, the main inconvenience of the RTK technique is that it requires a reference station relatively close to the user so that the differential ionospheric delay is negligible. High-accuracy single-baseline RTK solutions are generally limited to a distance of about 20 km [84], although test activities have shown that acceptable results can be achieved over up to 50 km in times of low ionospheric activity [85].

A way of partially overcoming such inconvenience is the Network RTK (NRTK) technique, which uses a network of ground reference stations to mitigate atmospheric dependent effects over long distances. The NRTK solution is generally based on between three

and six of the closest reference stations with respect to the user and allows much greater inter-station distances (up to 70-90 km) while maintaining a comparable level of accuracy [85].

The Wide Area RTK (WARTK) concept was introduced in the late 1990s to overcome the need of a very dense network of reference stations. WARTK techniques provide accurate ionospheric corrections that are used as additional information and allow increasing the RTK service area. In this way, the system needs reference stations separated by about 500–900 kilometres [86].

Precise Point Positioning (PPP) is a further carrier-phase GNSS technique that departs from the basic RTK implementation and uses differencing between satellites rather than differencing between receivers. Clearly, in order to be implemented, PPP requires the availability of precise reference GNSS orbits and clocks in real-time. Combining the precise satellite positions and clocks with a dual-frequency GNSS receiver (to remove the first order effect of the ionosphere), PPP is able to provide position solutions at centimetre level [87].

PPP differs from double-difference RTK/NRTK positioning in the sense that it does not require access to observations from one or more close reference stations accurately-surveyed. PPP just requires data from a relatively sparse station network (reference stations thousands of kilometres apart would suffice). This makes PPP a very attractive alternative to RTK especially in long-distance aviation applications where RTK/NRTK coverage would be impractical. However, PPP techniques are still not very mature and typically require a longer convergence time (20-30 minutes) to achieve centimetre-level performance [87]. Another possibility is to use local ionospheric corrections in PPP. This approach would reduce convergence time to about 30 seconds and contribute to an augmented integrity. However, they would also require a density of reference networks similar to that of NRTK [88].

4.4. Multi-Constellation GNSS

The various GNSSs (GPS, GLONASS, BDS, GALILEO, etc.) operate at different frequencies, they employ various signal processing techniques and adopt different multiple access schemes. Most of them employ or are planned to progressively transition to Code Division Multiple Access (CDMA) for operations at the L-band frequency of 1575.42 MHz (L1). This signal, originally introduced by GPS for Standard Positioning Service (SPS) users, will guarantee interoperability between the various GNSS constellations with substantial benefits to the existing vast community of GPS users, including aviation. Signal-in-Space (SIS) interoperability is also being actively pursued on other frequencies, with a focus on those that are currently allocated to Safety-of-Life (SOL) services. From a user's perspective, the advantage to having access to multiple GNSS systems is increased accuracy, continuity, availability and integrity. Having access to multiple satellite constellations is very beneficial in aviation applications, where the aircraft-satellite relative dynamics can lead to signal tracking issues and/or line-of-sight obstructions. Even when interoperability at SIS level cannot be guaranteed, the introduction of multi-constellation receivers can bring significant improvements in GNSS performance. The literature on this subject is vast and it is beyond the scope of this paper to address in details the various signal and data processing techniques implemented in GNSS systems.

The purpose of GNSS modernisation is to provide more robust navigation/positioning services in terms of accuracy, integrity, continuity and availability by broadcasting more ranging/timing signals. In particular, availability will be improved with the employment of multi-constellation GNSS. At the moment (July 2017), more than 70 GNSS satellites are already in view. It is expected that about 120 satellites will be available once the various GNSS systems are fully operational [89-93]. The GNSS

constellations and their supported signals are summarised in Table 14.

Currently, there are 31 GPS satellites on-orbit in the Medium Earth Orbit (MEO). The GPS signals can be summarised as follows:

- L1 (1575.42 MHz): It is a combination of navigation message, coarse-acquisition (C/A) code and encrypted precision P(Y) code (C/A-code is a 1023 bit Pseudo Random Noise (PRN) code with a clock rate of 1.023 MHz and P-Code is a very long PRN code (7 days) with a clock rate of 10.23 MHz);

- L2 (1227.60 MHz): It is a P(Y) code. The L2C code is newly added on the Block IIR-M and newer satellites. The L2C code is designed to meet emerging commercial/scientific needs and provides higher accuracy through better ionospheric corrections;

- L3 (1381.05 MHz): It is used by the defence support program to support signal detection of missile launches, nuclear detonations, and other high-energy infrared events;

- L4 (1379.913 MHz): It is used for studying additional ionospheric correction;

- L5 (1176.45 MHz): It is used as a civilian SoL signal. This frequency falls into an internationally protected range for aeronautical navigation, promising little or no interference under all circumstances.

SPS provides civil users with a less accurate positioning capability than PPS using single frequency L1 and C/A code only. PPS is available primarily to the military and other authorized users of the US (and its allies) equipped with PPS receivers. The PPS uses L1 and L2 C/A and P-codes. GPS modernisation comprises a series of improvements provided by GPS Block IIR-M, GPS Block IIF, GPS III, and GPS III Space Vehicle 11+. The M-code signals are modulated on L1 and L2. The addition of L5 makes GPS a more robust radionavigation service for many aviation applications, as well as all ground-based users. The new L2C signals, called as L2 civil-moderate (L2 CM) code and L2 civil-long (L2 CL) code, are modulated on the L2 signals transmitted by Block IIR-M, IIF, and subsequent blocks of GPS satellite vehicles. The M-code is a new military signal, modulated on both L1 and L2 signals.

The L1C signal was developed by the US and Europe as a common civil signal for GPS and GALILEO. Other GNSS systems such as BDS and QZSS are also adopting signals similar to L1C. The first L1C signal with be available with the launch of GPS III block satellites. Backward compatibility will be guaranteed by allowing the L1C signal to broadcast in the same frequency as the L1 C/A signal.

Out of the 30 planned GALILEO satellites, 13 are currently operational, 2 are under commissioning and 3 act as testbed (not available for operational use). GALILEO signals are used to provide 5 distinct services including Open Service (OS), Safety-of-Life Service (SoLS), Commercial Service (CS), Public Regulated Service (PRS) and Search and Rescue Support Service (SAR). The SOLS, in particular, support a number of aviation, marine and surface transport applications. The SOLS guarantees a level of accuracy and integrity that OS does not offer and it offers worldwide coverage with high accuracy integrity.

GALILEO carriers can be classified as E5a (1176.450 MHz), E5b (1207.140 MHz) and E5ab (full band), which are wide band data and pilot signals; E6 (1278.75 MHz) and E2-L1-E1 (1575.42 MHz). The GALILEO navigation signals can be classified as follows:

- L1F: It supports OS (unencrypted);

- L1P: It supports PRS (encrypted);

- E6C: It supports CS (encrypted);

- E6P: It supports PRS (encrypted);

- E5a: It supports OS (unencrypted);

- E5b: It supports OS (unencrypted).

Since the GALILEO E2-L1-E1 and E5a signal frequencies are the same as of L1 and L5 GPS signals, both these systems support their treatment as a systems of systems [94].

The GLONASS system has 23 satellites in operational condition. In addition to these M type satellites, there are already two modernised GLONASS-K satellites in operation. The more advanced GLONASS-KM satellites will be able to provide legacy FDMA signals supporting Open FDMA (OF) and Secured FDMA (SF) services on L1 and L2, as well as, CDMA signals on L1, L2 and L3 providing Open CDMA (OC) and Secured CDMA (SC) services. It could also transmit CDMA signals on the GPS L5 frequency (1176.45 MHz). Advanced features include inter-satellite links, laser time transfer capability and/or improved clocks [95]. Furthermore, GNSS integrity information could also be broadcast in the third civil signal in addition to global differential ephemeris and time corrections supporting Open CDMA Modernized (OCM) services. BDS (Big Dipper/COMPASS), the Chinese navigation satellite system, provides a regional navigation services using a two-way timing/ranging technique. BeiDou-2 (BDS-2) is expected to be a constellation of 35 satellites offering backward compatibility with BeiDou-1 and 30 non-geostationary satellites. These include 27 in MEO and 3 in Inclined Geosynchronous Orbit (IGSO) and 5 in Geostationary Earth orbit (GEO) that will offer complete coverage of the globe. The future fully operational BDS is expected to support two kinds of general services including Radio Determination Satellite Service (RDSS) and Radio Navigation Satellite Service (RNSS). The RDSS will support computation of the user position by a ground station using the round trip time of signals exchanged via GEO satellite while the RNSS will support GPS- and GALILEO-like services and is also designed to achieve similar performances. The QZSS provides a regional navigation service/augmentation system and is set to begin operations in 2018 with 3 IGSO satellites and 1 GEO satellite. Specific signals such as L1 sub-meter class Augmentation with Integrity Function (SAIF) and L6 L-band Experimental (LEX) will be transmitted in addition to the support for L1 C/A, L2C, L5 and L1C signals. The planned Block II satellites will support the Centimetre Level Augmentation Service and Positioning Technology Verification Service [96, 97]. The Indian Regional Navigation Satellite System (IRNSS), also referred to as NavIC (Navigation with Indian Constellation) comprises of seven satellites with 4 in IGSO and 3 in GEO. Two kinds of services are supported namely Standard Positioning Service (SPS) and Restricted Service (RS). Both services are carried on L5 (1176.45 MHz) and S band (2492.08 MHz). The navigation signals are expected to be transmitted in the S-band (2–4 GHz).

The new and modernised GNSS constellations, along with the newly introduced signals are expected to be compatible with other GNSS systems. Compatibility, in this context, refers to the ability of global and regional navigation satellite systems and augmentations to be used separately or together without causing unacceptable interference or other harm to an individual system or service [98]. To support computability among the various GNSS systems, ITU provides a framework for discussions on radiofrequency compatibility. Interoperability is another consideration to be taken into account, which refers to the ability of global and regional navigation satellite systems and augmentations to be used together so as to support services with better capabilities than would be achieved by relying solely on the open signals of one system [98]. Common modulation schemes, centre frequency, signal and power levels, as well as geodetic reference frames and system time steerage standards are defined to support interoperability.

Table 14. GNSS constellations.

Constellation (Global/ Regional)	Block and Orbit	Satellites	Signals
GPS	IIR (MEO)	12	L1 C/A, L1/L2 P(Y)
	IIR-M (MEO)	7	L1 C/A, L1/L2 P(Y), L2C, L1/L2 M
	IIF (MEO)	12	L1 C/A, L1/L2 P(Y), L2C, L1/L2 M, L5
	III (MEO)	Yet to be launched	L1 C/A, L1C, L1/L2 P(Y), L2C, L1/L2 M, L5
GALILEO	IOV (MEO)	3	E1, E5a/b/ab, E6
	FOC (MEO)	10	E1, E5a/b/ab, E6
GLONASS	MEO	Out of service	L1/L2 SF and L1 OF
	M (MEO)	23	L1/L2 OF and SF
	K1 (MEO)	2	L1/L2 OF and SF, L3 OC
	K2 (MEO)	Yet to be launched	L1/L2 OF and SF, L1 OC, L1 SC, L2 SC, L3 OC
	KM (MEO)	Yet to be launched	L1/L2 OF and SF, L1/L2/L3 OC and SC, L1/L3/L5 OCM
BeiDou-1	MEO	4 (experimental and retired)	B1
BeiDou-2	MEO	6	B1-2, B2, B3
	IGSO	8	B1-2, B2, B3
	GEO	6	B1-2, B2, B3
BeiDou-3	MEO	3	B1-2, B1, B2, B3ab
	IGSO	2	B1-2, B1, B2, B3ab
QZSS	I (GEO and IGSO)	2	L1 C/A, L1C, L1, L2C, L5, L6, LEX, SAIF
IRNSS	IGSO	4	L5/S SPS and RS
	GEO	3	L5/S SPS and RS

4.5. GNSS Integration with Other Sensors

All GNSS techniques (either stand-alone or differential) suffer from some shortcomings. The most important are:

- The data rate of a GNSS receiver is too low and the latency is too large to satisfy the requirements for high performance aircraft trajectory analysis, in particular with respect to the synchronisation with external events. In addition, there might be a requirement for high-rate and small latency trajectory data to provide real-time guidance information to the pilot (e.g., during fly-over-noise measurements). With the need to process radio frequency signals and the complex processing required to formulate a position or velocity solution, GPS data rates are usually at 1 Hz, or at best 10 Hz (an update rate of at least 20 Hz is required for real-time aviation

applications) [99];

- Influences of high accelerations on the GPS receiver clock, the code tracking loop and carrier phase loop may become significant;

- Signal 'loss-of-lock' and 'cycle slips' may occur very frequently due to aircraft manoeuvres or other causes;

- SA degrades the positioning accuracy over long distances;

- High DGNSS accuracy is limited by the distance between the reference station and the user because of the problem of ionosphere in integer ambiguity resolution on-the-fly [100].

Especially, in the aviation context, there is little one can do to ensure the continuity of the signal propagation from the satellite to the receiver during high dynamic manoeuvres. The benefits of integrating GNSS with other sensors are significant and diverse. Both absolute measurements such as Position, Velocity and Attitude (PVA), which can be directly measured as well as relative measurements between the air platform and the environment, can be considered for integration. Direct measurements can be obtained by employing GNSS, Inertial Navigation System (INS), digital maps, etc. In case of small UAS, GNSS measurements can be integrated with other navigation sensors/systems including Inertial Navigation System (INS), Vision-based Navigation Sensors (VBN), Celestial Navigation Sensors (CNS), Aircraft Dynamics Model (ADM) virtual sensor, etc. [101]. The variety of sensors for integration is even more expandable for aircraft operating in an urban setting when Signals-of-Opportunity (SoO)-based sources (cellular signal base transceiver stations, Wireless Fidelity (Wi-Fi) networks, etc.) are available [102]. Basically, each navigation system has some shortcomings. The shortcomings of GNSS were discussed above and in the case of INS; it is subject to an ever growing drift in position accuracy caused by various instrument error sources that cannot be eliminated in manufacturing, assembly, calibration or initial system alignment. Furthermore, high quality inertial systems (i.e., platform systems) tend to be complex and expensive devices with significant risk of component failure. By employing these sensors in concert with some adequate data-fusion algorithms, the integrated navigation system can solve the majority of these problems. As an example, the integration of GNSS and INS measurements might solve most of the above mentioned shortcomings: the basic update rate of an INS is 50 samples per second or higher and an INS is a totally self-contained system. The combination of GNSS and INS will therefore provide the required update rate, data continuity and integrity.

Technical considerations for integration of GNSS and other sensors include the choice of system architecture, the data-fusion algorithm, and the characterisation and modelling of the measurements produced by the sensors. Integration of other sensors with GNSS can be carried out in many different ways depending on the application (i.e., online/off-line evaluation, accuracy requirements). Various options exist for integration and in general, two main categories can be identified depending on the application: post-processing and real-time algorithms. The complexity of the algorithm is obviously related to both system accuracy requirements and computer load capacity.

The level of integration and the particular mechanisation of the data-fusion algorithm are dependent on:

- The task of the integration process;

- The accuracy limits;

- The robustness and the stand alone capacity of each subsystem;

- The computation complexity.

For GNSS and INS data fusion, the basic concepts of system integration can be divided into:

- Open Loop GNSS aided System (OLGS);

- Closed Loop GNSS aided System (CLGS);

- Fully Integrated GNSS System (FIGS).

The simplest way to combine GNSS and INS is a reset-only mechanisation in which GNSS periodically resets the INS solution (Fig. 21). In this open-loop strategy, the INS is not re-calibrated by GNSS data, so the underlying error sources in the INS still drive its navigation errors as soon as GNSS resets are interrupted. However, for short GNSS interruptions or for high quality INS, the error growth may be small enough to meet mission requirements. This is why platform inertial systems have to be used with OLGS systems operating in a high dynamic environment (e.g., military aircraft). The advantage of the open loop implementation is that, in case of inaccurate measurements, only the data-fusion algorithm is influenced and not the inertial system calculation itself. As the sensors of a platform system are separated from the body of the aircraft by gimbals, they are operating normally at their reference point zero. The attitude angles are measured by the angles of the gimbals. The integration algorithm can run internal or external to the INS, but the errors of the inertial system have to be carefully modelled.

The main advantage of GNSS aiding the INS in a closed-loop mechanisation is that the INS is continuously calibrated by the data-fusion algorithm, using the GNSS data. Therefore strapdown sensors can be used in a CLGS implementation (Fig. 21). In contrary to platform systems, sensors of strapdown INS are not uncoupled from the aircraft body. They are operating in a dynamically more disturbed environment as there are vibrations, angular accelerations, angular oscillations which result in an additional negative influence to the system performance. In addition sensors are not operating at a reference point zero. Therefore the errors of the system will increase very rapidly and problems of numerical inaccuracy in an open loop implementation may soon increase. This is why it is advantageous to loop back the estimated sensor errors to the strapdown calculations in order to compensate for the actual system errors. Consequently, the errors of the INS will be kept low and linear error models can be used. When GNSS data is lost due to dynamics or satellite shadowing, the INS can continue the overall solution, but now as a highly precise unit by virtue of its recent calibration. However, this system implementation can be unstable.

The system integration of best accuracy will be of course the full integration of GNSS and other sensors, which require a data-fusion implementation at a raw/uncorrelated measurements level (Fig. 21). As the KF theory asks for uncorrelated measurements [103], it is optimal to use either the raw GPS measurements (i.e., the range and phase measurements to at least four satellites, the ephemeris to calculate the satellite positions and the parameters to correct for the ionospheric and troposphere errors), or GNSS position (and velocity) data uncorrelated between update intervals [105]. In the first case, the receiver clock errors (time offset and frequency) can be estimated as a part of the filter model. This approach has very complex measurement equations, but requires only one KF mechanisation. Moreover, its filtering can be most optimal since both GNSS errors and INS errors can be included without the instability problems typical of cascade KFs.

The other classification commonly in practise to define the integration techniques is given by:

- Loose integration: It refers to fusion of navigation solutions. Processed GNSS PVT measurements are used in this type of integration.

- Tight integration: It refers to the fusion of navigation

measurements. GNSS code and carrier range and phase measurements are used in this type of integration.

- Deep/ultratight integration: It refers to the integration at the

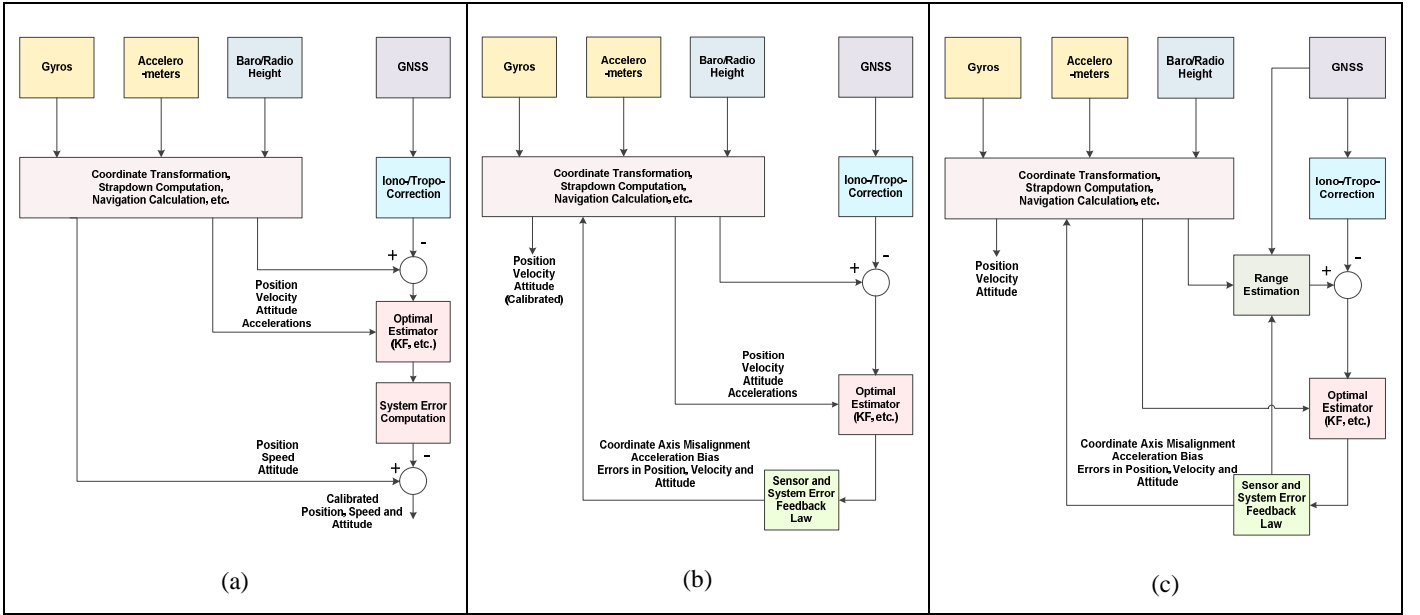signal processing level, which employs raw GNSS radiofrequency signals.

Fig. 21. GNSS integration architectures: a. OLGS, b. CLGS and c. FIGS [107].

The most important distinction to be made is between cascaded and non-cascaded approaches, which correspond, as mentioned before, to aided (OLGS and CLGS) or fully integrated (FIGS) architectures respectively. In the cascaded case two filters generally play the role. The first filter is a GNSS filter which produces outputs (i.e., position and velocity) which are correlated between measurement times. This output is then used as input for the second filter which is the INS KF. As time correlation of this measurement input does not comply with the assumptions underlying the standard KF [103], it must be accounted for in the right way [106]. This complicates the KF design (i.e., the potential instability of cascaded filters makes the design of the integration KF a very cautious task). However, from a hardware implementation point of view aided INS (in both OLGS and CLGS configurations) results in the simplest solution. In the non-cascaded case there is just a single KF and is generally based on an INS error model and supplemented by a GNSS error model. The GNSS measurements, uncorrelated between measurement times, are differenced with the raw INS data to give measurements of the INS errors, also uncorrelated between measurements times.

State-of-the-art data-fusion algorithms such as the traditional KF, Extended Kalman Filter (EKF) and the more advanced Unscented Kalman Filter (UKF)/multi-hypotheses UKF can be employed for the integration process. In general, a KF can be used to estimate the errors, which affect the solution computed in an INS or in a GNSS, as well as in the combination of both navigation sensors. A KF is a recurrent, optimal estimator used in many engineering applications whenever the estimation of the state of a dynamic system is required. An analytic description of KFs and other integration algorithms can be found in the literature [103, 104]. The KF algorithm is optimal under three limiting conditions: the system model is linear, the noise is white and Gaussian with known autocorrelation function and the initial state is known. Moreover, the computing burden grows considerably with increasing number of states modelled. Matrix formulation methods of various types have been developed to improve numerical stability and accuracy (e.g., square root and stabilised formulations), to minimise the computational complexity by taking advantage of the diagonal characteristics of the covariance matrix (U-D factorisation formulation where U and D are upper triangular and diagonal

matrix respectively), and to estimate state when the state functions are non-linear (e.g. EKF). The EKF measurement model is defined as:

$$z_k = H_k * x_k + v_k \qquad (119)$$

where z_k is the measurement vector, H_k is the design matrix, x_k is the state vector, v_k is the measurement noise and k is the kth epoch of time, t_k. The state vector at epoch $k+1$ is given by:

$$x_{k+1} = \Phi_k * x_k + G_k * w_k \qquad (120)$$

where Φ_k is the state transition matrix from epoch k to $k+1$, G_k is the shaping matrix and w_k is the process noise. The algorithm is composed of two steps: prediction and correction. The prediction algorithm of the EKF estimates the state vector and computes the corresponding covariance matrix P_k from the current epoch to the next one using the state transition matrix characterizing the process model described by:

$$P_{k+1}^- = \Phi_{k+1} P_k^+ \Phi_{k+1}^T + Q_k \qquad (121)$$

where P_{k+1}^- is a predicted value computed and P_k^+ is the updated value obtained after the correction process. The process noise at a certain epoch k is characterized by the covariance matrix, Q_k. The updating equations correct the predicted state vector and the corresponding covariance matrix using the collected measurements is given by:

$$x_{k+1}^+ = K_{k+1} u_{k+1} \qquad (122)$$

$$P_{k+1}^+ = P_{k+1}^- - K_{k+1} H_{k+1} P_{k+1}^- \qquad (123)$$

where K_{k+1} is the Kalman gain matrix at epoch, $k+1$ and u_{k+1} is the innovation vector at epoch, $k+1$.

Post-processing integration of GNSS and INS measurements can be carried out by means of the Rauch-Tung-Striebel algorithm, which consists of a KF and a backward smoother. The forward filter can take into account previous measurements only. The smoothed estimate utilises all measurements. It is evident that backward smoothing can only be done off-line. The filter estimates the state vector together with the error covariance matrix. The state vector contains the error components of the inertial navigation system. The smoothed trajectories and also accurate velocities are

obtained by adding the state vector to the measurements delivered by the INS. The equations of the Rauch-Tung-Striebel algorithm can be found in [103]. Alternatives include the Bryson–Frazier two filter smoother, Monte Carlo smoother, etc. The filtering algorithm most commonly implemented in real-time systems is the Bierman's U-D Factorised KF. The U-D algorithm is efficient and provides significant advantages in numerical stability and precision [103].

Artificial Neural Networks (ANN) could be employed to relax one or more of the conditions under which the KF is optimal and/or reduce the computing requirements of the KF. However, the development of an integration algorithm completely based only on neural technology (e.g., hopfield networks) will lead to a complex design and unreliable integration functions. Furthermore, this solution is even more demanding than the KF in terms of computing requirements [108]. Hence, a hybrid network in which the ANN is used to learn the corrections to be applied to the state prediction performed by a rule-based module appears to be the best candidate for future navigation sensor integration. Techniques for on-line training would allow for real-time adaptation to the specific operating conditions, but further research is required in this field. In a hybrid architecture, an ANN is used in combination with a rule-based system (i.e., a complete KF or part of it), in order to improve the availability of the overall system. A hybrid network provides performances comparable with the KF, but with improved adaptability to non-linearities and unpredicted changes in system/environment parameters. Compared with the KF, the hybrid architecture features the high parallelism of the neural structure allowing for faster operation and higher robustness to hardware failures. The use of ANN to correct the state variables prediction operated by the rule-based module avoids computing the Kalman gain, thus considerably reducing the computing burden. Stability may be guaranteed if the output of the network is in the form of correction to a nominal gain matrix that provides a stable solution for all system parameters. The implementation of such a filter, however, would be sensitive to network topology and training strategy. An adequate testing activity would therefore be required. In addition to ANN approaches, such as the Radial Basis Function (RBF) neural networks, Adaptive Neuron-Fuzzy Inference System (ANFIS) have been developed to enhance GNSS/INS integration for its improved ability to handle the problem at hand and to include an expert knowledge base of a fuzzy system and adaptive membership functions characterising the relationship between the inputs and outputs.

4.6. GNSS Integrity Augmentation

According to the US Federal Radionavigation Plan, "integrity is a measure of the trust that can be placed in the correctness of the information supplied by a navigation system. Integrity includes the ability of the system to provide timely warnings to users when the system should not be used for navigation." This definition is very comprehensive but, unfortunately, it is too generic and this is why many research efforts were undertaken in the past to better specify the GNSS integrity performance metrics. Previous research has effectively developed and applied statistical criteria to inflate the navigation error bounds in specific flight phases. So, while accuracy is typically specified at 95% (2σ) confidence level, integrity requirements normally refer to percentiles of 99.999% (6σ) or higher depending on the particular phase of flight/application [109]. The intention behind this is to keep the probability of hazardous events (that could possibly put at risk human lives) extremely low. From a system performance perspective, integrity is intended as a real-time decision criterion for using or not using the system. For this reason, it has been a common practice to associate integrity with a mechanism, or set of mechanisms (barriers) that is part of the integrity assurance chain but at the same time is completely independent of the other parts of the system for which integrity is to be assured.

Consequently, the concepts of Alert Limit, Integrity Risk, Protection Level and Time to Alert were introduced and are currently reflected in the applicable ICAO SARPs and in the RTCA standards for WAAS and LAAS. These integrity performance metrics are:

- Alert Limit (AL): The alert limit for a given parameter measurement is the error tolerance not to be exceeded without issuing an alert.

- Time to Alert (TTA): The maximum allowable time elapsed from the onset of the navigation system being out of tolerance until the equipment enunciates the alert.

- Integrity Risk (IR): Probability that, at any moment, the position error exceeds the AL.

- Protection Level (PL): Statistical bound error computed so as to guarantee that the probability of the absolute position error exceeding said number is smaller than or equal to the target integrity risk.

As discussed above, integrity is the ability of a system to provide timely warnings to users when the system should not be used for navigation [110]. In general, it is essential to guarantee that, with a specified probability P, either the horizontal/vertical radial position error does not exceed the pre-specified thresholds (HAL/VAL), or an alarm is raised within a specified TTA interval when the horizontal/vertical radial position errors exceed the thresholds. To detect that the error is exceeding a threshold, a monitor function has to be installed within the navigation system. This is also the case within GNSS systems in the form of the ground segment. However, for GNSS systems, TTA is typically in the order of several hours, that is even too long for the cruise where a TTA of 60 seconds is required (an autoland system for zero meter vertical visibility must not exceed a TTA of 2 seconds). Various methods have been proposed and practically implemented for stand-alone GNSS integrity monitoring. A growing family of such implementations, already very popular in aviation applications, includes the so called Receiver Autonomous Integrity Monitoring (RAIM) techniques. Regarding DGNSS, it should be underlined that it does more than increasing GNSS positioning accuracy; it also contributes to enhancing GNSS integrity by compensating for anomalies in the satellite ranging signals and navigation data message. The range and range rate corrections provided in the ranging-code DGNSS correction message can compensate for ramp and step type anomalies in the individual satellite signals, until the corrections exceed the maximum values or rates allowed in the correction format. If these limits are exceeded, the user can be warned not to use a particular satellite by placing "do-not-use" bit patterns in the corrections for that satellite (as defined in STANAG 4392 or RTCA/RTCM message formats) or by omitting the corrections for that satellite.

As mentioned before, step anomalies will normally cause carrier-phase DGNSS receivers to lose lock on the carrier phase, causing the reference and user receivers to reinitialise. UR noise, processing anomalies, and multipath at the user GPS antenna cannot be corrected by a DGNSS system. These errors are included in the overall DGNSS error budget. Errors in determining or transmitting the satellite corrections may be passed on to the differential user if integrity checks are not provided within the RS. These errors can include inaccuracies in the RS antenna location that bias the corrections, systematic multipath due to poor antenna sighting (usually in low elevation angle satellites), algorithmic errors, receiver inter-channel bias errors, receiver clock errors and communication errors. For these reasons, typical WAD and LAD RS designs also include integrity checking provisions to guarantee the validity of the corrections before and after broadcast.

In the aviation context, various strategies have been developed for increasing the levels of integrity of DGNSS based

navigation/landing systems. These include both Space Based Augmentation Systems (SBAS) and Ground Based Augmentation Systems (GBAS) to assist aircraft operations in all flight phases (Fig. 22).
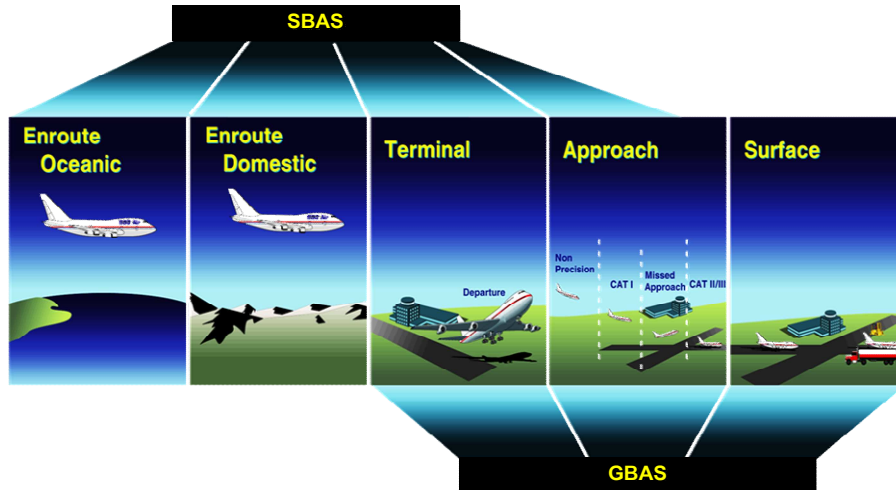

Fig. 22. GBAS and SBAS services. Adapted from [111].

These systems are effectively WAD and LAD systems that employ, respectively, DGNSS vector correction (SBAS) and scalar (pseudorange) correction techniques (GBAS). Some information about SBAS and GBAS relevant to this thesis are provided in the following sections of this chapter. However, before examining the characteristics of GNSS integrity augmentation systems, it is useful to recall the definitions relative to Failure Detection and Exclusion (FDE).

Fig. 23 shows the different conditions associated with FDE. The various FDE events are defined as follows [41]:

- Alert: An alert is defined to be an indication that is provided by the GPS equipment when the positioning performance achieved by the equipment does not meet the integrity requirements.

- Positioning failure: A positioning failure is defined to occur whenever the difference between the true position and the indicated position exceeds the applicable alert limit.

- False detection: A false detection is defined as the detection of a positioning failure when a positioning failure has not occurred.

- Missed detection: A missed detection is defined to occur when a positioning failure is not detected.

- Failed exclusion: A failed exclusion is defined to occur when true positioning failure is detected and the detection condition cannot be eliminated within the time-to-alert. A failed exclusion would cause an alert.

- Wrong exclusion: A wrong exclusion is defined to occur when detection occurs, and a positioning failure exists but is undetected after exclusion, resulting in a missed alert.

- False alert: A false alert is defined as the indication of a positioning failure when a positioning failure has not occurred. A false alert is the result of a false detection.

- Missed alert: A missed alert is defined to be a positioning failure which is not enunciated as an alert within the time-to-alert. Both missed detection and wrong exclusion can cause missed alerts.
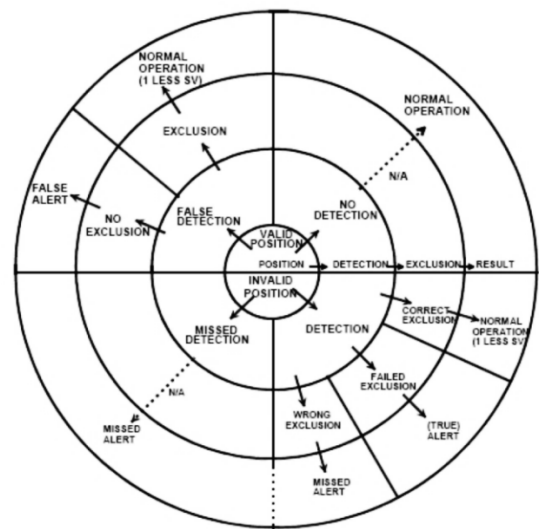

Fig. 23. FDE Events [36].

As discussed above, GNSS satisfies the horizontal and vertical integrity requirements for the intended operation when HPL and VPL are below the specified HAL and VAL. This situation is depicted in Fig. 24.
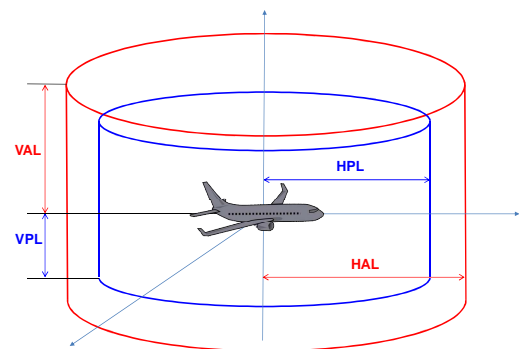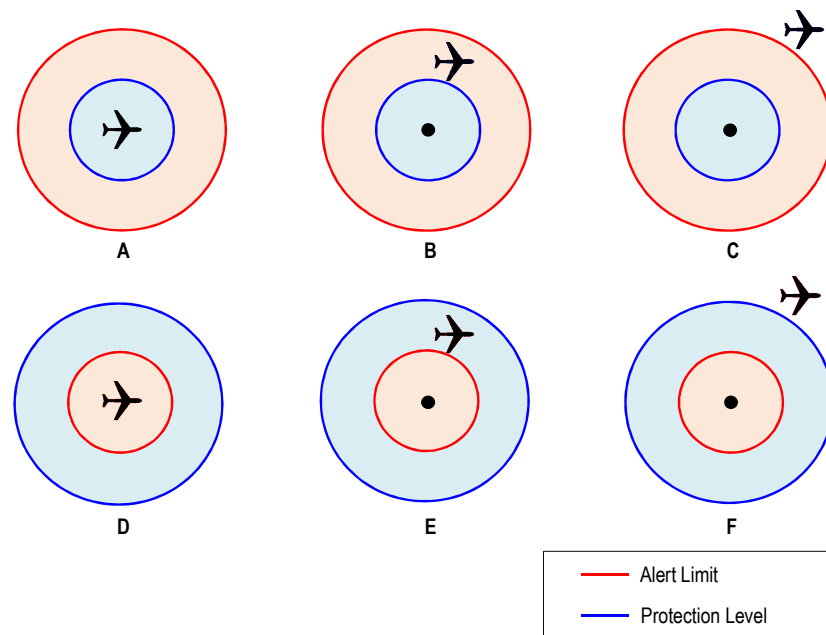

Fig. 24. Protection levels and alert limits.

The horizontal plane in Fig. 24 is locally tangent to the navigation system geodetic reference (e.g., WGS84 ellipsoid in GPS). The vertical axis is locally perpendicular to the same reference. Whenever HPL or VPL exceed HAL or VAL, the integrity required to support the intended operation is not provided and an alert must be issued by the GNSS equipment within the specified TTA. In some unfavourable occasions, the NSE exceeds the HPL/VPL values. In these cases, it is said that an integrity event has occurred and that the system is providing unreliable information, classified as Misleading Information (MI) or Hazardously Misleading Information (HMI). The difference between MI and HMI is highlighted in Fig 25.

As discussed, availability is the probability that the navigation system is operational during a specific flight phase (i.e., the accuracy and integrity provided by the system meet the requirements for the desired operation). Therefore, the navigation system is considered available for use in a specific flight operation if the PLs it is providing are inferior to the corresponding specified ALs for that same operation.

The circumferences shown in Fig. 25 have their centres at the estimated aircraft position and their radii are the system PLs (green) and the operation ALs (red). The aircraft shape represents the actual position, so the navigation system error is the distance from the centre of the circumferences to this shape. According to the previous definitions, situations B, C and F represent integrity events. In particular, case B is a MI event and cases C/F are HMI events. The desirable situation is represented by case A. Particular attention should be given to situations B and especially C, which is the most feared situation, as the system does not alert the user for the unavailability of the system and, depending on the on-going flight operational task; this may lead to a SoL risk.

Fig. 25. Protection levels and alert limits.

4.6.1. Space Based Augmentation Systems

SBAS is a form of WAD that augments core satellite constellations by providing ranging, integrity and correction information via geostationary satellites. An SBAS typically comprises of [112]:

- a network of ground reference stations that monitor satellite signals;

- master stations that collect and process reference station data and generate SBAS messages;

- uplink stations that send the messages to geostationary satellites;

- transponders on these satellites that broadcast the SBAS messages.

By providing differential corrections, extra ranging signals via geostationary satellites and integrity information for each navigation satellite, SBAS delivers much higher levels of service than the core satellite constellations. In certain configurations, SBAS can support approach procedures with vertical guidance (APV). There are two levels of APV: APV I and APV II. Both use the same lateral obstacle surfaces as localizers; however APV II may have lower minima due to better vertical performance. There will nonetheless be only one APV approach to a runway end, based on the level of service that SBAS can support at an aerodrome. The two APV approach types are identical from the perspective of avionics and pilot procedures. In many cases, SBAS will support lower minima than that associated with non-precision approaches, resulting in higher airport usability. Almost all SBAS approaches will feature vertical guidance, resulting in a significant increase in safety. APV minima (down to a Decision Height (DH) of 75 m (250 ft) approximately) will be higher than Category I minima, but APV approaches would not require the same ground infrastructure, so this increase in safety will be affordable at most airports. SBAS availability levels will allow operators to take advantage of SBAS instrument approach minima when designating an alternate airport. Notably, SBAS approach does not require any SBAS infrastructure at an airport. SBAS can support all en-route and terminal RNAV operations. Significantly, SBAS offers affordable RNAV capability for a wide cross section of users. This will allow a reorganization of the airspace for maximum efficiency and capacity, allowing aircraft to follow the most efficient flight path between airports. High availability of service will permit to decommission traditional NAVAIDs, resulting in lower costs.

There are four SBASs now in service including:

- the North American Wide Area Augmentation System (WAAS);

- the European Geostationary Navigation Overlay Service (EGNOS);

- the Indian GPS and Geostationary Earth Orbit (GEO) Augmented Navigation (GAGAN) System;

- the Japanese Multi-functional Transport Satellite (MTSAT) Satellite-based Augmentation System (MSAS).

Other SBAS currently under study or developments include the Russian System for Differential Correction and Monitoring (SDCM), the South/Central American and Caribbean SACCSA (Solución de Aumentación para Caribe, Centro y Sudamérica), the Chinese Satellite Navigation Augmentation System (SNAS), the Malaysian Augmentation System (MAS) and various programs in the African and Indian Ocean (AFI) region. The approximate coverage areas of the various SBAS are shown in Fig. 26.

Although Australia, Indonesia, New Zealand and various other nations in the southern portion of the Asia-Pacific region have not announced SBAS development programs, the extension of systems such as MSAS, GAGAN and SNAS could provide SBAS coverage in these areas. Within the coverage area of SBAS, ICAO member states may establish service areas where SBAS supports approved operations. Other States can take advantage of the signals available in the coverage area by fielding SBAS components integrated with an existing SBAS or by authorizing the use of SBAS signals. The first option offers some degree of control and improved performance. The second option lacks any degree of control, and the degree of improved performance depends on the proximity of the host SBAS to the service area. In either case, the State, which has established an SBAS service area, should assume responsibility for the SBAS signals within that service area. This requires the provision of NOTAM information services.

If ABAS-only operations are approved within the coverage area of SBAS, SBAS avionics will also support ABAS operations. Although the architectures of the various existing SBASs are different, they broadcast the standard message format on the same frequency (GPS L1) and so are interoperable from the user's perspective.

It is anticipated that these SBAS networks will expand beyond their initial service areas. Other SBAS networks may also be developed and become operational. When SBAS coverage areas overlap, it is possible for an SBAS operator to monitor and send integrity and correction messages for the geostationary satellites of another SBAS, thus improving availability by adding ranging sources.
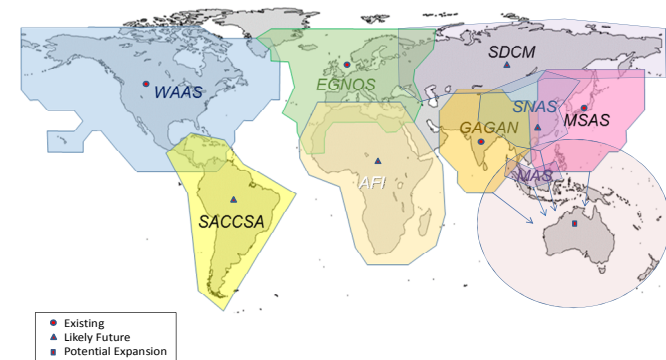

Fig. 26. Current and possible future SBAS coverage. Adapted from [113].

As the American WAAS has been developed in line with ICAO SARPs and RTCA MOPS, the GBAS technology considered is this paper is based on the RTCA WAAS standard [41], which is also aligned with the ICAO SARPs for SBAS [114].

The aim of WAAS is to provide augmentation signals to correct some of the main GNSS errors and providing the following services:

- Position correction (ECEF coordinates);

- Ionospheric grid correction;

- Long term ephemeris error correction;

- Short term and long term satellite clock error correction;

- Integrity information.

WAAS consists of a ground, space and user segment (Fig. 27). A ground segment is composed by a network of ground reference stations and uplink stations. The reference stations, which are called Ranging and Integrity Monitoring Station (RIMS), are responsible for the analysis of GNSS signals and for producing ranging corrections and integrity signals. The uplink stations are responsible for sending regular corrections/integrity signal updates to the space segment. The space segment is made by different satellites in geostationary orbits and broadcast the ranging corrections and integrity signals to the WAAS users. The user segment (i.e., users equipped with WAAS enabled GNSS receivers) uses the information broadcast from GNSS satellites to compute PVT and WAAS signals broadcast from the space segment to improve position accuracy (applying ionospheric corrections) and integrity of the navigation solution.

This correction evaluation, together with the correction of other parameters such as ephemeris error and clock error, allow WAAS to provide to the user a position accuracy of about 7.6 meters or better both for lateral and vertical measurements [41]. This permits WAAS to achieve, under certain conditions, accuracy levels compatible with CAT-I precision approach requirements.

As the CAT-I capability required significant efforts for certification, a new Precision Approach sub-mode was defined called Approach Operations with Vertical Guidance (APV), giving an intermediate precision approach between NPA (Non- Precision Approach) and CAT-I. This is further divided in APV-I and APV-II.
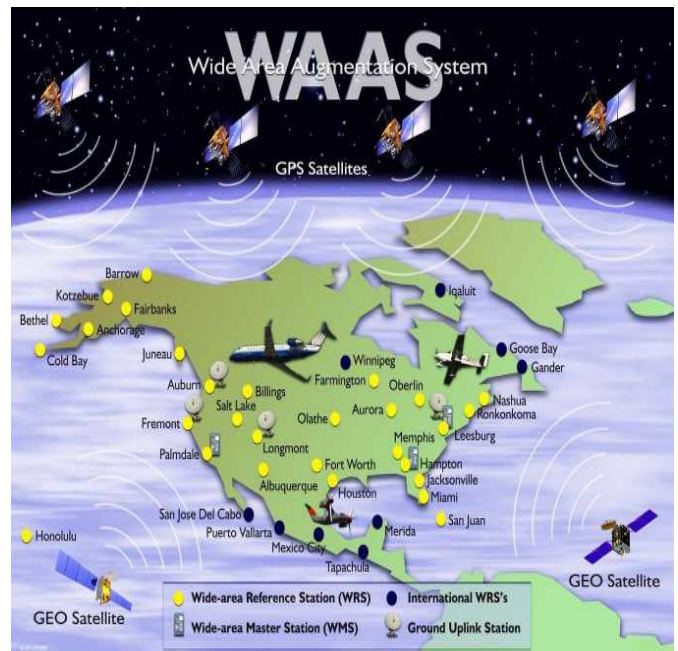

Fig. 27. WAAS Architecture and Operational Environment [63].

Ionospheric delay is one of the major error sources of pseudorange. The WAAS ionospheric delay corrections are broadcast as vertical delay estimates at specified ionospheric grid points (IGPs), applicable to a signal on L1. The WAAS Reference Stations (WRSs) measure the slant ionospheric delays to all satellites in view. These measurements must be translated into a form that can be applied by the user because the user will have a different line of sight to the satellite than the WRSs. WAAS uses a two-dimensional grid model to represent the vertical ionospheric delay

distribution [47]. Ten different bands are used to allow a regular spacing of IGPs. The 1808 possible IGP locations in bands 0-8 are illustrated in Fig. 28 (there are 384 additional IGP locations in bands 9-10 not shown). In the bands 0-8, the IGPs are spaced 5° apart from each other both in longitude and latitude for latitudes up to N55° and S55°. Above 55° and up to 75° in latitude the IGPs are spaced 10° from each other both in longitude and in latitude. At S85° and N85° (around the poles), the IGPs are spaced 90°. In the bands 9-10, the IGP grid at 60° has 5° longitudinal spacing, increasing to 10° spacing at 65°, 70° and 75°, and finally becoming 30° at N85° and S85° round the poles. To accommodate an even distribution of bands, IGPs at 85° are offset by 40° in the 0-8 bands and 10° in the 9-10 bands. The IGPs error data are transmitted in Message 26 and denoted in terms of GIVEI (Grid Ionospheric Vertical Error Indicator), which corresponds to specific values of Grid Ionospheric Vertical Error (GIVE) and GIVE variance ($\sigma^2_{i,GIVE}$). The GIVEIi, GIVEi and the $\sigma^2_{i,GIVE}$ are listed in Table 15.

Fig. 28. WAAS IGPs for bands 0-8 [41].

The GIVE is used to calculate the User Ionospheric Vertical Error (UIVE) and the User Ionospheric Range Error (UIRE). The UIVE and UIRE are, respectively, the confidence bounds on the user vertical and range errors due to the ionosphere. Both the UIVE and UIRE are referred to the Ionospheric Pierce Point (IPP) location. The IPP is defined as the intersection of the line segment from the receiver to the satellite and an ellipsoid with constant height of 350 km above the WGS-84 ellipsoid. Fig. 29 shows the relationship between user location and IPP location [41].

Table 15. Evaluation of GIVEIi [41].

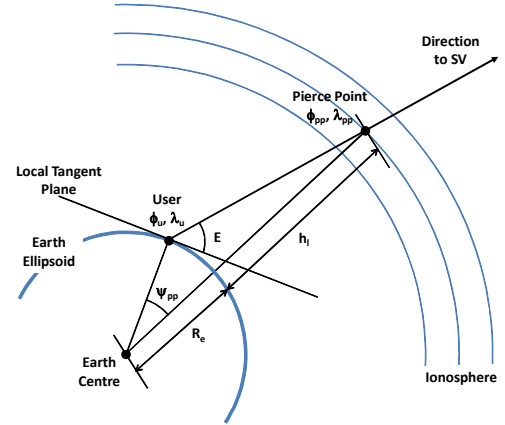| GIVEIi	GIVEi [m]	$\sigma^2_{i,GIVE}$ [m²]
0	0.3	0.0084
1	0.6	0.0333
2	0.9	0.0749
3	1.20	0.1331
4	1.5	0.2079
5	1.8	0.2994
6	2.1	0.4075
7	2.4	0.5322
8	2.7	0.6735
9	3.0	0.8315
10	3.6	1.1974
11	4.5	1.8709
12	6.0	3.3260
13	15.0	20.7870
14	45.0	187.0826
15	Not Monitored	Not Monitored

Fig. 29. Ionospheric Pierce Point Geometry [41].

The latitude of the IPP is given by:

$$\phi_{pp} = \sin^{-1}\big(\sin\phi_u \cos\psi_{pp} + \cos\phi_u \sin\psi_{pp} \cos A\big) \ [\text{rad}] \ (119)$$

where ϕ_u is the user location latitude, A is the azimuth angle of the satellite from the user's location (ϕ_u, λ_u) measured clockwise from north and ψ_{pp} is the Earth's central angle between the user position and the projection of the pierce point.

If $\phi_u > 70°$ and $\tan\psi_{pp}\cos A > \tan(\pi/2 - \phi_u)$ or $\phi_u < -70°$ and $\tan\psi_{pp}\cos(A + \pi) > \tan(\pi/2 + \phi_u)$, the longitude of the IPP is given by:

$$\lambda_{pp} = \lambda_u + \sin^{-1}\left(\frac{\sin\psi_{pp}}{\cos\phi_{pp}}\sin A\right) \ [\text{rad}] \qquad (120)$$

Otherwise:

$$\lambda_{pp} = \lambda_u + \sin^{-1}\left(\frac{\sin\psi_{pp}}{\cos\phi_{pp}}\sin A\right) \ [\text{rad}] \qquad (121)$$

ψ_{pp} is given by:

$$\psi_{pp} = \frac{\pi}{2} - E - \sin^{-1}\left(\frac{R_e}{R_e+h_I}\cos E\right) \ [\text{rad}] \qquad (122)$$

where E is the elevation angle of the satellite form the user's location measured with respected to the local tangent plane, R_e is the approximate Earth radius (assumed to be 6378.163 km) and h_I is the height of the maximum electron density (assumed to be 350 km). Fig. 30 shows the principles adopted for interpolation of the IPPs. The variance of UIVE can be calculated using one of the following formulas:

$$\sigma^2_{UIVE} = \sum_{n=1}^{4} W_n(x_{pp}, y_{pp}) \cdot \sigma^2_{n,ionogrid} \qquad (123)$$

$$\sigma^2_{UIVE} = \sum_{n=1}^{3} W_n(x_{pp}, y_{pp}) \cdot \sigma^2_{n,ionogrid} \qquad (124)$$

where $\sigma^2_{n,ionogrid}$ is the model variance of inospheric vertical delays at an IGP. As indicated in the WAAS MOPS, a dedicated model can be used to calculate both fast and long-term correction degradation components (for inclusion in Message Types 7 and 10).
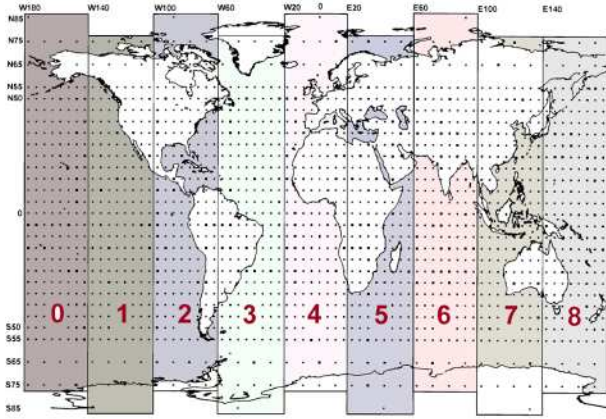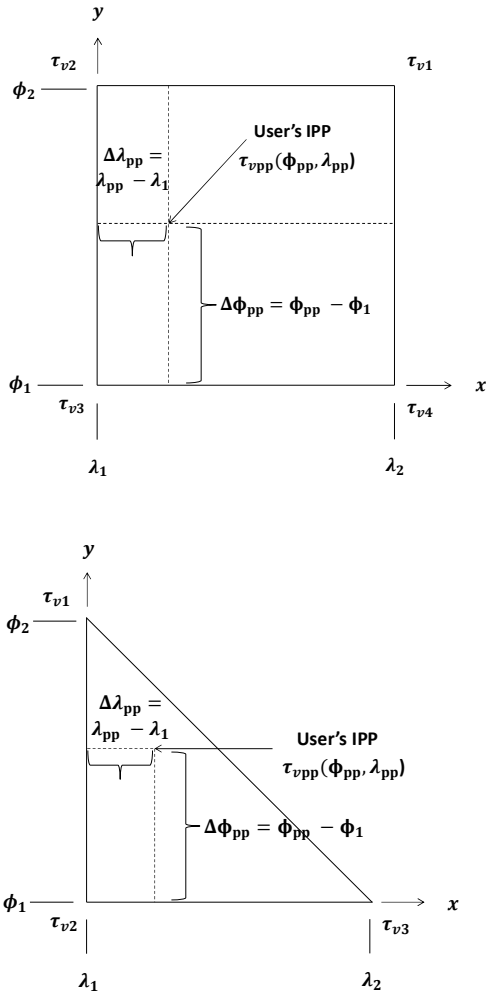
Fig. 30. Principle of Interpolation of the IPPs [41].

If the degradation model is not implemented, the basic WAAS ionospheric model is used by taking $\sigma_{n,ionogrid}^2 = \sigma_{GIVE}^2$. For the ith satellite, the variance of UIRE is then calculated as follows:

$$\sigma_{i,UIRE}^2 = F_{pp}^2 \cdot \sigma_{i,UIVE}^2 \qquad (125)$$

where F_{pp} is the obliquity factor given by:

$$F_{pp} = \left[1 - \left(\frac{R_e \cos E}{R_e + h_I} \right)^2 \right]^{-\frac{1}{2}} \qquad (126)$$

Real-time data from can be obtained from the FAA website for WAAS and from the ESA website for EGNOS [113]. The FAA also created a comprehensive portal containing real-time data on other GBAS systems including not only WAAS and EGNOS, but also MSAS and GAGAN [115]. The airborne receiver error is denoted as $\sigma_{i,air}$. This error is defined in the WAAS MOPS based on the equipment operation classes. There are four operation classes defined in the WAAS MOPS [41]. A description of these classes is provided in Table 16.

The fast corrections and long term corrections are two important messages transmitted by SBAS. The long term corrections contain information on the slowly varying satellite orbit and clock errors. The fast corrections provide additional information on the fast varying clock errors. Long term corrections consist of position and clock offset values only (EGNOS) or they also contain velocity and clock drift corrections (WAAS).

The User Differential Range Error (UDRE) message provides information that allows the user to derive a bound on the projection of the satellite orbit and satellite clock errors onto the line of sight

of the worst case user. The UDRE is transmitted as an indicator (UDREI).

Table 16. SBAS Operational Classes [41].

Class	Description
Class 1	Equipment that supports oceanic and domestic en route, terminal, approach (LANV), and departure operation. When in oceanic and domestic en route, terminal, LNAV, and departure operations, this class of equipment can apply the long-term and fast SBAS differential corrections when they are available.
Class 2	Equipment that supports oceanic and domestic en route, terminal, approach (LANV, LNAV/VNAV), and departure operation. When in LNAV/VNAV, this class of equipment applies the long-term, fast and ionospheric corrections. When in oceanic and domestic en route, terminal, approach (LNAV), and departure operations, this class of equipment can apply the long-term and fast SBAS differential corrections when they are available.
Class 3	Equipment that supports oceanic and domestic en route, terminal, approach (LANV, LNAV/VNAV, LP, LPV), and departure operation. When in LPV, LP, or LNAV/VNAV, this class of equipment applies the long-term, fast and ionospheric corrections. When in oceanic and domestic en route, terminal, approach (LNAV), and departure operations, this class of equipment can apply the long-term and fast SBAS differential corrections when they are available.
Class 4	Equipment that supports only the final approach segment operation. This class of equipment is intended to serve as an ILS alternative that supports LP and LPV operation with degradation (fail-down) from LPC to lateral only (LNAV). Class 4 equipment is only applicable to functional Class Delta, and equipment that meets Class Delta-4 is also likely to meet the requirements for Class Beta-1, 2, or -3.

In terms of WAAS integrity features, Message Types 27 and 28 are particularly important. Type 27 Messages may be transmitted to correct the σUDRE in selected areas. Each Type 27 Message specifies δUDRE correction factors to be applied to integrity monitoring algorithms of users when inside or outside of a set of geographic regions defined in that message. δUDRE indicators are associated with the δUDRE values listed in Table 17 that multiply the model standard deviation defined using the UDREI parameters in the WASS Types 2-6 and Type 24 Messages.

As an alternative to Message 27, a service provider might broadcast Message Type 28 (not both) to update the correction confidence as a function of user location. Message Type 28 provides increased availability inside the service volume and increased integrity outside. The covariance matrix is a function of satellite location, reference station observational geometry, and reference station measurement confidence. Consequently it is a slowly changing function of time. Each covariance matrix only needs to be updated on the same order as the long-term corrections. Each message is capable of containing relative covariance matrices for two satellites. This maintains the real-time six-second updates of integrity and scales the matrix to keep it within a reasonable dynamic range.

Assuming a Message Type 28 data is used and that fast and long term corrections are applied to the satellites without the correction degradation model (i.e., an active type 7 or 10 message data is not available), then the residual fast and long term error is defined as [41]

$$\sigma_{i,flt}^2 = [(\sigma_{i,UDRE}) \cdot (\delta UDRE) + 8]^2 \quad [\text{m}] \qquad (132)$$

The residual tropospheric error is denoted as $\sigma_{i,tropo}$. It is modelled as a random parameter with a standard deviation of $\sigma_{i,tropo}$ as specified in the MOPS [41].

Table 17. δUDRE Indicators and values in Message Type 27.

$\delta UDRE$ Indicator	$\delta UDRE$
0	1
1	1.1
2	1.25
3	1.5
4	2
5	3
6	4
7	5
8	6
9	8
10	10
11	20
12	30
13	40
14	50
15	100

Cholesky factorization is used to reliably compress the information in the covariance matrix (C). This factorization produces 4x4 an upper triangular matrix (R), which can be used to reconstruct the relative covariance matrix as follows:

$$C = R^T R \qquad (127)$$

Cholesky factorization guarantees that the received covariance matrix remains positive-definite despite quantization errors. Because R is upper triangular, it contains only 10 non-zero elements for each satellite. These 10 elements are divided by a scale factor to determine the matrix E, which is broadcast in Message Type 28:

$$E = \frac{R}{\psi} = \begin{bmatrix} E_{1,1} & E_{1,2} & E_{1,3} & E_{1,4} \\ 0 & E_{2,2} & E_{2,3} & E_{2,4} \\ 0 & 0 & E_{3,3} & E_{3,4} \\ 0 & 0 & 0 & E_{4,4} \end{bmatrix} \qquad (128)$$

where $\psi = 2^{(scale\ exponent-5)}$. The covariance matrix is then reconstructed and used to modify the broadcast σUDRE values as a function of user position. The location-specific modifier is given by:

$$\delta UDRE = \sqrt{(I^T \cdot C \cdot I)} + \varepsilon_C \qquad (129)$$

where I is the 4D line of sight vector from the user to the satellite in the WGS-84 coordinate frame, whose first three components are the unit vector from the user to the satellite and the fourth component is a one. The additional term ε_C is to compensate for the errors introduced by quantization. If degradation data is available, the ε_C value is derived from the covariance data broadcast in a Type 10 message ($C_{covariance}$) as follows:

$$\varepsilon_C = C_{covariance} \cdot \psi \qquad (130)$$

If $C_{covariance}$ from Message Type 10 is not available, ε_C is set to zero, but there is an 8 meter degradation applied as defined in the WAAS MOPS [41]. The WAAS fast corrections and long-term corrections are designed to provide the most recent information to the user [41]. However, it is always possible that the user fails to receive one of the messages. In order to guarantee integrity even when some messages are not received, the user performing approach operations (e.g, LNAV/VNAV, LP or LPV) must apply

models of the degradation of this information. In other flight phases, the use of these models is optional and a global degradation factor can be used instead as specified in the WAAS MOPS [41]. The residual error associated with the fast and long-term corrections is characterized by the variance (σ_{flt}^2) of a model distribution. This term is computed as:

$$\sigma_{flt}^2 =$$

$$\begin{cases} (\sigma_{UDRE} \cdot \delta UDRE) + \varepsilon_{fc} + \varepsilon_{rrc} + \varepsilon_{ltc} + \varepsilon_{er}, & \text{if RSS}_{UDRE} = 0 \\ (\sigma_{UDRE} \cdot \delta UDRE)^2 + \varepsilon_{fc}^2 + \varepsilon_{rrc}^2 + \varepsilon_{ltc}^2 + \varepsilon_{er}^2, & \text{if RSS}_{UDRE} = 1 \end{cases} \quad (131)$$

where:

RSS_{UDRE} is the root-sum-square flag in Message Type 10;

σ_{UDRE} is the model parameter from Message Types 2-6, 24;

$\delta UDRE$ is the user location factor (from Message Types 28 or 27, otherwise $\delta UDRE = 1$);

ε_{fc} is the degradation parameter for fast correction data;

ε_{rrc} is the degradation parameter for range rate correction data;

ε_{ltc} is the degradation parameter for long term correction or GEO nav-message data;

ε_{er} is the degradation parameter for en route through NPA operations.

The total error σ_i^2 is given by [41]:

$$\sigma_i^2 = \sigma_{i,flt}^2 + \sigma_{i,UIRE}^2 + \sigma_{i,air}^2 + \sigma_{i,tropo}^2 \qquad (132)$$

For Class 1 equipment:

$$\sigma_{i,air}^2 = 25\text{m}^2 \qquad (133)$$

For Class 2, 3 and 4 equipment:

$$\sigma_{i,air} = (\sigma_{noise,GNSS}^2[i] + \sigma_{multipath}^2[i] + \sigma_{divg}^2[i])^{1/2} \quad (134)$$

The projection matrix (S) is defined as [41]:

$$S = \begin{bmatrix} s_{east,1} & s_{east,2} & \cdots & s_{east,N} \\ s_{north,1} & s_{north,2} & \cdots & s_{norht,N} \\ s_{U,1} & s_{U,2} & \cdots & s_{U,N} \\ s_{t,1} & s_{t,2} & \cdots & s_{t,N} \end{bmatrix}$$

$$= (G^T \cdot W \cdot G)^{-1} \cdot G^T \cdot W \qquad (135)$$

where G is the observation matrix, W is the weighted matrix, $s_{east,i}$ is the partial derivative of position error in the east direction with respect to the pseudorange error on the ith satellite, $s_{north,i}$ is the partial derivative of position error in the north direction with respect to the pseudorange error on the ith satellite, $s_{U,i}$ is the partial derivative of position error in the vertical direction with respect to the pseudorange error on the ith satellite and $s_{U,i}$ is the partial derivative of time error with respect to pseudorange error on the ith satellite. The weighted matrix W is defined as:

$$W = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_N \end{bmatrix} \qquad (136)$$

$$w_i = 1/\sigma_i^2 \qquad (137)$$

and the observation matrix G consists of N rows of LOS vectors from each satellite augmented by a "1" for the clock. Thus, the ith row corresponds to the ith satellites in view and can be written in terms of the azimuth angle Az_i and elevation angle θ_i. The ith row of the G matrix is:

$$G_i = [-\cos\theta_i \sin Az_i \quad -\cos\theta_i \cos Az_i \quad -\sin\theta_i \quad 1] \quad (138)$$

The SBAS Vertical Protection Level (VPLSBAS) is calculated using the equation:

$$VPL_{SBAS} = K_V d_U \qquad (139)$$

where:

$$d_U^2 = \sum_{i=0}^{N} s_{U,i}^2 \sigma_i^2 \qquad (140)$$

The value of K_V used for computing VPL is: $K_V=5.33$. The SBAS Horizontal Protection Level (HPL_{SBAS}) is calculated using the equations:

$$HPL_{SBAS} =$$

$$\begin{cases} K_{H,NPA} \cdot d_{major}, & \text{for en route through LNAV} \\ K_{H,PA} \cdot d_{major}, & \text{for LNAV/VNAV , LP, LPV approach} \end{cases} \qquad (141)$$

where:

$$d_{major} \equiv \sqrt{\frac{d_{east}^2+d_{north}^2}{2} + \sqrt{(\frac{d_{east}^2-d_{north}^2}{2})^2 + d_{EN}^2}} \qquad (142)$$

$$d_{east}^2 = \sum_{i=0}^{N} s_{east,i}^2 \sigma_i^2 v \qquad (143)$$

$$d_{north}^2 = \sum_{i=0}^{N} s_{norht,i}^2 \sigma_i^2 \qquad (144)$$

$$d_{EN}^2 = \sum_{i=0}^{N} s_{east,i} s_{north,i} \sigma_i^2 \qquad (145)$$

The values $K_{H,NPA}$ is 6.18 for en route through LNAV and 6.0 for LNAV/VNAV, LP, LPV. More detailed information about WAAS can be found in the applicable MOPS standard [41].

4.6.2. Ground Based Augmentation Systems

The current GNSS constellations are unable to provide accuracy, availability, continuity and integrity to achieve the stringent accuracy and integrity requirements set by ICAO for precision approach. GBAS employs LADGNSS techniques to augment accuracy and ad-hoc features to augment integrity beyond the levels that can be achieved with SBAS. Employing all provisions included in the current RTCA standards [36, 40], GBAS could provide augmentation to the core constellations to support precision approach up to Category IIIb. However, to date, ICAO has only issued Standards and Recommended Practices (SARPs) for GBAS operating over single frequency and single constellation up to CAT I operations [38]. SARPs for CAT II/III have been already developed by the ICAO Navigation System Panel (NSP) but not yet included in Annex 10, waiting for further developments to take place in the aviation industry. As shown in Fig. 31, GBAS utilises three segments: satellites constellation, ground stations and aircraft receiver. The GBAS ground station is formed by reference receivers with their antennas installed in precisely surveyed locations. The information generated in the receiver is sent to a processor that computes the corrections for each navigation satellite in view and broadcasts these differential corrections, besides integrity parameters and precision approach pathpoints data, using a VHF Data Broacast (VDB). The information broadcast is received by aircraft in VHF coverage that also receive information from the navigation satellites. Then, it uses the differential corrections on the information received directly from the navigation satellites to calculate the precise position. The precise position is used, along with pathpoints data, to supply deviation signals to drive appropriate aircraft systems supporting precision approach operations.
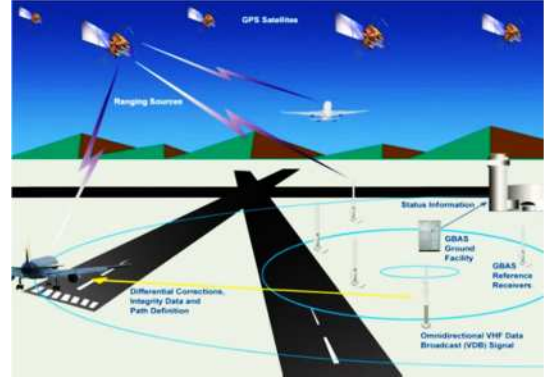

Fig. 31. GBAS architecture [115].

Although the operating principle, Signal-in-Space and performance characteristics of GBAS are completely different from ILS, the GBAS approach guidance inidications provided to the pilot are similar to the localizer and glide path indications provided by an ILS. However, compared to ILS precision approach systems, GBAS presents several distinct benefits. These include:

- Reduction of critical and sensitive areas typical of ILS;

- Possibility to perform curved approaches;

- Positioning service for RNAV operations in the Terminal Manoeuvring Area (TMA);

- Provision of service in several runways in the same airport;

- Provision of several approach glide angles and displaced threshold;

- Guided missed approach service;

- Adjacent airports served by a single GBAS (within VDB coverage).

According to [115], GBAS is intended to support all types of approach, landing, departure and surface operations and may support en-route and terminal operations The SARPs were developed, to date, to support Category I precision approach, approach with vertical guidance and GBAS positioning service. The GBAS ground station performs the following functions:

- Provide locally relevant pseudo-range corrections;

- Provide GBAS-related data;

- Provide final approach segment data when supporting precision approach;

- Provide ranging source availability data;

- Provide integrity monitoring for GNSS ranging sources.

VDB radio frequencies used in GBAS are selected in the band of 108 to 117.975 (just below the ATM band). The lowest assignable frequency is 108.025 MHz and the highest assignable frequency is 117.950 MHz, with a channel spacing of 25 kHz. A Time Division Multiple Access (TDMA) technique is used with a fixed frame structure. The data broadcast is assigned one to eight slots. GBAS data is transmitted as 3-bit symbols, modulating the data broadcast carrier by Differential 8 Phase Shift Keying (D8PSK), at a rate of 10 500 symbols per second. GBAS can provide a VHF data broadcast with either horizontal (GBAS/H) or elliptical (GBAS/E) polarization. The navigation data transmitted by GBAS includes the following information:

- Pseudo-range corrections, reference time and integrity data;

- GBAS-related data;

- Final approach segment data when supporting precision

approach;

- Predicted ranging source availability data.

LAAS (US) is a system capable to provide local augmentation for the GNSS signals and it is the first candidate to support all types of approach and landing procedures up to Category III, as well as surface operations. All existing GBAS systems refer to the same LAAS standards developed by RTCA [36, 40]. The typical LAAS facility consists of three elements: the first is the ground segment, usually installed on the airport field, the second is the airborne segment, represented by the data link receiver as well as a system capable of applying the augmentation parameters for landing guidance, the third is the space segment, which is actually made by the GPS/GLONASS constellations and will also include the future GALILEO/BEIDOU constellations when the required levels of maturity will be achieved. WAAS essentially provides two different services: the first is the precision approach service by giving deviation guidance, both lateral and vertical, for the Final Approach Segment (FAS) to the runway; the second is the differentially-corrected positioning service transmitting to the airborne systems the correction message for augmented GNSS accuracy. The specific data link to be used for communications between aircraft and ground stations is not completely specified by the current standards, while there are constraints in terms of bandwidth, link budget and data rate [65]. For approach and landing services, the LAAS performance is classified in terms of GBAS Service Level (GSL). A GSL defines a specific level of required accuracy, integrity and continuity. The GSL classification and the GSL performance requirements are listed in Tables 18 and 19. Currently, GBAS/LAAS installations around the world have been certified for GSL-C only. However, several research and development efforts are on-going towards demonstrating and producing adequate evidence of compliance for GLS-D/E/F installations. LAAS can also support airport surface operations by

enabling several functions associated with the specific aircraft equipment such as an electronic moving map and database:

Enhanced pilot situational awareness of aircraft position: the position information will allow the pilot to identify the correct taxi route and to monitor his position along the route; this situational awareness is improved in all visibility conditions, particularly for complex airports. Nowadays to support the low visibility surface movement operations is a very costly task; with LAAS, ADS-B and emerging cockpit displays technologies, these operations could be conducted in the same way of those supported with lights, markings and follow-me operators.

Traffic surveillance: this function can support air traffic control and crew with traffic surveillance of other aircraft both in good and low visibility condition.

Runway incursion and conflict detection alerting: the availability of accurate LAAS aircraft position information enables automated systems to detect runway incursions and other conflicts providing suitable alerts. Various error sources affect the system accuracy, integrity and continuity and these are nowadays an active area of research. Some of these errors are:

- Hardware/software faults in ground equipment or satellites such as the clock error;

- Multipath effects occurring on aircraft or ground receiver antennas;

- Errors in the ephemeris information;

- Residual ionospheric and tropospheric errors;

- Degradation of the satellite signal at the source;

- Interference/jamming issues.

Table 18. GBAS Service Level Performance and Service Levels (adapted from [36]).

GSL	Accuracy		Integrity				Continuity	Operations Supported
	95% Lat. NSE	95% Vert. NSE	Prob. (Loss of Integrity)	Time to Alert	LAL	VAL	Prob. (Loss of Continuity)	
A	16 m	20 m	$1-2\times10^{-7}$/150 sec	10 sec	40 m	50 m	$1-8\times10^{-6}$/15 sec	Approach operations with vertical guidance (performance of APV-I designation)
B	16 m	8 m	$1-2\times10^{-7}$/150 sec	6 sec	40 m	20 m	$1-8\times10^{-6}$/15 s	Approach operations with vertical guidance (performance of APV-II designation)
C	16 m	4 m	$1-2\times10^{-7}$/150 sec	6 sec	40 m	10 m	$1-8\times10^{-6}$/15 s	Precision approach to lowest CAT I minima
D	5 m	2.9 m	10^{-9}/15 s (vert.); 30 s (lat.)	2 sec	17 m	10 m	$1-8\times10^{-6}$/15 s	Precision approach to lowest CAT IIIb minima, when augmented with other airborne equipment
E	5 m	2.9 m	10^{-9}/15 s (vert.); 30 s (lat.)	2 sec	17 m	10 m	$1-4\times10^{-6}$/15 s	Precision approach to lowest CAT II/IIIa minima
F	5 m	2.9 m	10^{-9}/15 s (vert.); 30 s (lat.)	2 sec	17 m	10 m	$1-2\times10^{-6}$/15 s (vert.); 30 s (lat.)	Precision approach to lowest CAT IIIb minima

The LAAS service volume (Fig. 32) is defined as the region where the system is required to meet the accuracy, integrity and continuity requirements for a specific GSL class.

For each runway supported by the system, the minimum service volume to support Category I, Category II or APV (vertical guidance) approaches is defined as follow:

Laterally: beginning at 450 feet each side of the Landing Threshold Point/Fictitious Threshold Point (LTP/FTP) and projecting out

±35° either side of the final approach path to a distance of 20 NM from the LTP/FTP.

Vertically: within the lateral region, up to the maximum of 7° or 1.75 times the Glide Path Angle (GPA) above the horizon with the origin at the Glide Path Intercept Point (GPIP), up to a maximum of 10,000 feet AGL and down to 0.9° above the horizon with the origin in the GPIP, to a minimum height of one half the DH.
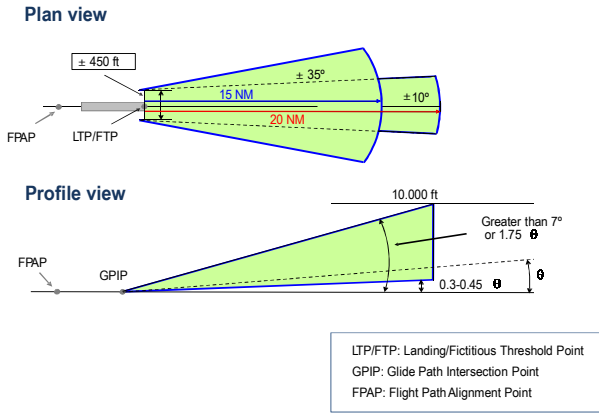
Fig. 32. Minimum category I, II and APV service volume (adapted from [65]).

Table 19. GBAS Service Levels (adapted from [36]).

GSL	Operations Supported by GSL
A	Approach operations with vertical guidance (performance of APV-I designation)
B	Approach operations with vertical guidance (performance of APV-II designation)
C	Precision approach to lowest CAT I minima
D	Precision approach to lowest CAT IIIb minima, when augmented wih other airborne equipment
E	Precision approach to lowest CAT II/IIIa minima
F	Precision approach to lowest CAT IIIb minima

The minimum service volume to support Category III precision approach or autoland with any minima is the same as the minimum Category II service volume, with the addition of the volume within ±450 feet from the runway centreline extending from the threshold to the runway end, from 8 feet above the surface up to 100 feet. There are different metrics for the GBAS performance; one of these is the SIS performance. This is defined in terms of accuracy, integrity, continuity and availability of the service [65]; this performance refers to the output of the "fault-free" airborne user equipment. The interaction between airborne and ground systems is defined by a separation of responsibilities:

The ground system is responsible for monitoring the satellites signal quality and availability, computing the differential corrections and detecting and mitigating the faults originated in the ground station or in the space segment.

The airborne equipment computes the pseudorange values and evaluates the performance level monitoring, detecting and mitigating any fault conditions due to the other avionics equipment. For a precision approach, the avionics GBAS receiver only uses satellites for which corrections are available

Avionics GBAS receivers have to comply with the requirements outlined in ICAO Annex 10, vol. 1 [38] and the specifications of RTCA/DO-253C [36]. GBAS avionics receivers must have the capability to receive navigation satellites signal and the VDB information, with respective antennas, and a way to select the approach and an indication of course and glide path. Like ILS and Microwave Landing System (MLS), GBAS has to provide lateral and vertical guidance relative to the defined final approach course and glide path. The GBAS receiver employs a channelling scheme that selects the VDB frequency. Approach procedure data are uplinked via the VDB and each separate procedure requires a different channel assignment. In terms of aircraft system integration, GBAS avionics standards have been developed to

mimic the ILS in order to minimize the impact of installing GBAS on the existing avionics. Typically, display scaling and deviation outputs are the same utilized for ILS, so that maximum interoperability is achieve at the Human-Machine Interface (HMI) level, allowing the avionics landing systems to provide final approach course and glide path guidance both at airports equipped with ILS and GLS. For TMA area navigation, including segmented and curved approaches, the GBAS avionics receiver provides accurate position, velocity and time data that can be used as an input to an on-board navigation computer. In line with ICAO SARPs and the strategy for the introduction and application of non-visual aids to approach and landing, which permit a mix of systems providing precision approach service, the avation industry has developed multi-mode receivers that support precision approach operations based on ILS, MLS and GNSS (GBAS and SBAS). Also for GBAS, integrity monitoring is accomplished by the avionics continually comparing Horizontal/Lateral and Vertical Protection Levels (HPL/LPL and VPL), derived from the augmentation signal and satellite pseudorange measurements, against the alert limit for the current phase of flight. When either the vertical or the horizontal limit is exceeded, an alert is given to the pilot [116]. Additionally, the LAAS MOPS [36] also contemplate the possibility of introducing avionics-based augmentation functionalities by defining the continuity of protection levels in terms of Predicted Lateral and Vertical Protection Levels (PLPL and PVPL). The standard states that on-board avionics systems could support PLPL and PVPL calculations to generate appropriate caution flags. However, it does not recommend any specific form of ABAS and does not indicate the required performance of such on-board equipment to supplement LAAS operations. Research in this field has concentrated on possible RAIM aiding rather than proper ABAS developments. Some techniques have been developed that include the possible aiding of barometric altimeter, with a special focus on vertical guidance integrity monitoring and augmentation [117-121] and, more recently, the possible inclusion of predictive features has also been studied especially for mission/route planning applications. However, the main limitations of such techniques is in the inability to model GNSS errors due to aircraft dynamics, including antenna obscuration, Doppler shift and multipath contributions due to the aircraft structure and the surrounding environment (the latter being important when the aircraft flies at low altitudes). This is why a new form of ABAS specifically tailored for real-time integrity monitoring and augmentation in mission-critical and safety-critical aviation applications is needed [53, 54] and is discussed later in this section.

There are two conditions in the protection level calculations. One is the H0 hypothesis, the other is H1 hypothesis. The H0 hypothesis refers to no faults are present in the range measurements (including both the signal and the receiver measurements) used in the ground station to compute the differential corrections. The H1 hypothesis refers to the presence of a fault in one or more range measurements that is caused by one of the reference receivers used in the ground station. The H0 Hypothesis total error model is:

$$\sigma_i^2 = \sigma_{pr_{gnd},i}^2 + \sigma_{tropo,i}^2 + \sigma_{air,i}^2 + \sigma_{iono,i}^2 \qquad (146)$$

where $\sigma_{air,i}^2$ is the total aircraft contribution to the corrected pseudorange error for the ith satellite. This is given by:

$$\sigma_{air,i}^2 = \sigma_{receiver}^2(\theta_i) + \sigma_{multipath}^2(\theta_i) \qquad (147)$$

The residual tropospheric and ionospheric errors are due to the position difference of reference receiver and aircraft. According to the LAAS MASPS, the tropospheric error is defined as [40]:

$$\sigma_{tropo}(\theta) = \sigma_N h_0 \frac{10^{-6}}{\sqrt{0.002+sin^2(\theta)}}\left(1 - e^{-\frac{\Delta h}{h_0}}\right) \qquad (148)$$

where σ_N =troposphere refractivity uncertainty transmitted in the GBAS Message Type 2, θ is the elevation angle for the ith ranging source. It is expressed in the ENU frame, Δh is the difference in altitude between airborne and ground subsystems (reference receivers), in meters and h_0 is the tropospheric scale height (transmitted in GBAS Message Type 2).The residual ionospheric error is defined as [40]:

$$\sigma_{iono} = F_{pp}\sigma_{vert\_iono\_gradient}(x_{air} + 2\tau v_{air}) \quad (149)$$

where F_{pp} is the obliquity defined in equation (25). The reference receiver error model is used to generate the total error which contributing the corrected pseudorange for a GPS satellite caused by the reference receiver noise. The standard deviation of the reference receiver error is [40]:

$$\sigma_{pr\_gnd,GPS}(\theta_i) = \sqrt{\frac{(a_0 + a_1 e^{-\theta_i/\theta_0})^2}{M} + (a_2)^2} \quad (150)$$

where M is the number of ground reference receiver subsystems and θ_i is the elevation angle for the ith ranging source. The airborne receiver error is a function of the satellite elevation angle above the local level plane. It is defined as [40]:

$$\sigma_{receiver}(\theta_i) = a_0 + a_1 e^{-(\theta_i/\theta_0)} \quad (151)$$

where θ_i is the elevation angle for the ith ranging source and the parameters a_0, a_1 and θ_0 are defined in [40] for various Airborne Accuracy Designators (AAD). It is to be noted that also the aircraft multipath error is a function of the satellite elevation angle. The Airframe Multipath Designator (AMD) is a letter that indicates the airframe multipath error contribution to the corrected pseudorange for a GPS satellite. There are two types of AMD in the LAAS MASPS, denoted as A and B [40]. The aircraft multipath error for AMD type A is:

$$\sigma_{multipath}^A[i] = (0.13 + 0.53 e^{(-\theta_i/10)})[m] \quad (152)$$

For AMD type B:

$$\sigma_{multipath}^B[i] = \frac{0.13 + 0.53 e^{(-\theta_i/10)}}{2}[m] \quad (153)$$

The multipath model for AMD type A ($\sigma_{multipath}^A$) was obtained during an FAA sponsored research program, which was aimed at investigating and quantifying the total effect of multipath from the airframe and from ground multipath during approach and landing operations [55]. The FAA sponsored program included the development of an electromagnetic modelling capability to predict amplitude, time delay and phase characteristic of airframe multipath. The final empirical model is the result of extensive flight test data analysis conducted with large commercial airliner platforms (i.e., B777-300ER and B737-NG) as described in [122]. As stated in the LAAS MASPS [40], the multipath model for AMD type B ($\sigma_{multipath}^B$) is still under development [40]. Since $\sigma_{multipath}^A$ is expected to produce conservative multipath error estimations, it is acceptable to use $\sigma_{multipath}^A$ in all cases until more reliable $\sigma_{multipath}^B$ models are developed. The H1 hypothesis total error model is given by:

$$\sigma_{i,H1}^2 = \frac{M_i}{M_i - 1}\sigma_{pseudorange_{gnd},i}^2 + \sigma_{troposphere,i}^2$$
$$+\sigma_{air,i}^2 + \sigma_{ionosphere,i}^2 \quad (154)$$

where M_i is the number of reference receivers used to compute the pseudorange corrections for the ith satellite. SBAS/GBAS projection matrix (S) matrix is obtained from the observation matrix (G) and weighted matrix (W), and its elements determine the protection levels of GBAS. The S matrix for GBAS has the same formats already defined for SBAS. The fault-free vertical protection level (VPL_{H0}) can be calculated using the equation [24]:

$$VPL_{H0} = K_{ffmd}\sqrt{\sum_{i=1}^N s^2_{Apr\_vert,i}\sigma_i^2} \quad (155)$$

where $s_{Apr\_vert,i}$ is the projection of the vertical component and translation of the along track errors into the vertical for the ith satellite. This is given by:

$$s_{Apr\_vert,i} = s_{v,i} + s_{x,i} * \tan\theta_{GS} \quad (156)$$

where θ_{GS} is the glide path angle for the final approach path and N is the number of ranging sources used in the position solution. The H1 hypothesis vertical protection level (VPL_{H1}) can be calculated using the equations:

$$VPL_{H1} = MAX\{VPL_j\} \quad (157)$$

$$VPL_j = |B_{j,vert}| + K_{md}\sigma_{vert,H1} \quad (158)$$

where j is the ground reference receiver index, K_{md} is found in Table 7-9 and the other terms are:

$$B_{j,vert} = \sum_{i=1}^N s_{vert,i}B_{i,j} \quad (159)$$

$$\sigma_{vert,H1}^2 = \sum_{i=1}^N s_{vert,i}^2\sigma_{i,H1}^2 \quad (160)$$

The fault-free lateral protection level (LPL_{H0}) is calculated using the equation (RTCA, 2004):

$$LPL_{H0} = K_{ffmd}\sqrt{\sum_{i=1}^N s^2_{Apr\_lat,i}\sigma_i^2} \quad (161)$$

where $s_{Apr\_lat,i}$ is the projection of the vertical component and translation of the along track errors into the vertical for ith satellite (also denoted $s_{lat,i}$). The H1 hypothesis lateral protection level (LPLH1) can be calculated using the following equations from the LAAS MASPS [40]:

$$LPL_{H1} = MAX\{LPL_j\} \quad (162)$$

$$LPL_j = |B_{j,lat}| + K_{md}\sigma_{lat,H1} \quad (163)$$

where:

$$B_{j,lat} = \sum_{i=1}^N s_{lat,i}B_{i,j} \quad (164)$$

$$\sigma_{lat,H1}^2 = \sum_{i=1}^N s_{lat,i}^2\sigma_{i,H1}^2 \quad (165)$$

The subscript j is the ground subsystem reference receiver index and the multiplier K_{md} can be found in [40]. If the GBAS provides GSL E or F services, then the predicted protection level must be calculated to enable continuity of the GBAS SIS. The Predicted Vertical Protection Level (PVPL) and Predicted Lateral Protection Level (PLPL) are defined in MASPS as [40]:

$$PVPL = MAX\{PVPL_{H0}, PVPL_{H1}\} \quad (166)$$

$$PLPL = MAX\{PLPL_{H0}, PLPL_{H1}\} \quad (167)$$

$$PVPL_{H0} = K_{ffmd}\sqrt{\sum_{i=1}^N s^2_{vert,i}\sigma_i^2} \quad (168)$$

$$PVPL_{H1} = K_{ffd}\sigma_{B,vert} + K_{md}\sigma_{vert,H1} \quad (169)$$

$$PLPL_{H0} = K_{ffmd}\sqrt{\sum_{i=1}^N s^2_{lat,i}\sigma_i^2} \quad (170)$$

$$PLPL_{H1} = K_{ffd}\sigma_{B,lat} + K_{md}\sigma_{lat,H1} \quad (171)$$

where K_{ffd} is the multiplier which determines the probability of fault-free detection given M reference receivers.

$$\sigma_{B,vert}^2 = \sum_{i=1}^N s_{vert,i}^2 \frac{\sigma_{pr\_gnd,i}^2}{(M_i-1)} \quad (172)$$

$$\sigma_{B,lat}^2 = \sum_{i=1}^N s_{lat,i}^2 \frac{\sigma_{pr\_gnd,i}^2}{(M_i-1)} \quad (173)$$

The VAL defines the threshold values of VPL and PVPL (i.e., the GBAS signal is not to be used to navigate the aircraft when the VPL exceeds the VAL/PVAL).

4.6.3. Receiver Autonomous Integrity Monitoring

Various methods have been proposed and developed for stand-alone GNSS integrity monitoring. One of the popular implementations for aviation application is the RAIM technique. RAIM has its foundations in statistical detection theory, wherein redundant GNSS measurements are used to form a test-statistic from which a fault can be inferred. RAIM is based on the reasoning that the quality of a user's position solution can be evaluated at the receiver. This supports detection of system-level faults. RAIM was introduced in the latter part of the 1980s when autonomous means of GNSS failure detection started to be investigated [42]. A fault is identified based on a self-consistency check among the available measurements. Traditionally, RAIM and its performance have mostly been associated with the integrity-monitoring tasks in aviation and other safety-critical applications, where relatively good line-of-sight signal reception conditions prevail, and there are often strict rules of well-defined integrity modes. The goal in these traditional RAIM operating environments was to eliminate large measurement errors due to, for example, satellite clock or ephemeris failures, which can be caused by failures in the satellite or in the control segment. These errors are very rare, with a typical rate of one error per 18 to 24 months in the GPS system [121].

Fault Detection-based RAIM (FD-RAIM) techniques provide the following basic information: the presence of a failure and which satellite(s) have failed. FD-RAIM algorithms rely on users being able to access more satellites than the minimum number required for a navigation solution, in order to estimate the integrity of the signal from these redundant measurements. Typically, RAIM requires 5 satellites for performing fault detection (sometimes 4 satellites when baro-aiding is used). However, in order to perform fault detection and exclusion, 6 satellites are needed. Considering, for example, the GPS constellation providing only a single frequency catering to SPS, RAIM cannot meet Category I requirements for precision approach applications, although RAIM-enabled receivers can be used as a navigation aid for less demanding phases of flight. Hence, the aim is to evaluate the likely performance of RAIM algorithms in a future multi-GNSS environment in which users have access to signals from more than one system simultaneously. This over-determination of the navigation position solution from the large number of received ranging measurements potentially allows users to apply RAIM to a level appropriate for precision approach tasks. Hence the detection probability will increase dramatically, due to both the increased size of the available satellites and the improved accuracy of the signals compared with traditional GPS-only signals.

Fault Detection and Exclusion-based RAIM (FDE-RAIM) FDE requires six satellites and like FD-RAIM, may use barometric aiding as an additional information source. With six or more visible satellites FDE-RAIM detects a faulty satellite, removes it from the navigational solution and continues to provide FDE/FD with the remaining visible satellites. FDE is required for oceanic RNAV approvals and is being mandated in aviation standards (TSO-C145a and C146a).

Another concept often used within the RAIM context is a "RAIM hole", which refers to the period of time when there are insufficient visible GNSS satellites to provide an integrity check at any given location. RAIM holes can be predicted using a number of techniques. The most common are:

- spatial (grid-based) and temporal sampling intervals based techniques when available computation capability is low;

- precise computation techniques for estimating satellite coverage boundaries, the intersection points and analysis of the topology of the regions of intersection when available computation capability is high.

While the adoption of RAIM techniques can significantly improve the situation, the inherent reactive nature of traditional RAIM techniques do not allow an immediate implementation of predictive features beyond the extrapolations that can be made from GNSS signals and ephemeris data. The introduction of ABAS, SBAS and GBAS systems has allowed for an extended range of integrity monitoring and augmentation features (also providing substantial accuracy and, in some cases, availability improvements), which can be exploited synergically in the various aircraft flight phases. However all these integrity augmentation systems have not been successful so far in introducing predictive integrity monitoring and augmentation features suitable for the most demanding flight tasks (i.e., precision approach in CAT-III and autoland certification).

Conventionally, pseudorange-based and carrier-phase RAIM (PRAIM and CRAIM) are employed in the standard positioning algorithms to ensure that a level of trust can be placed on the solution. PRAIM consists of two processes namely detection (of a potential failure) and derivation of the required protection level. The detection process requires the construction of a test-statistic using the measurement residuals followed by the determination of a threshold value that provides an indication of the presence of a failure. The derivation of the protection level is based on an estimate of the upper bound of the positioning error that might result from an undetected error. Both HPL and VPL are employed in order to confirm if an intended operation can be supported by the available GNSS. PRAIM has achieved a level of success in civil aviation applications and is being applied at various flight phases (e.g., steady flight). But, when performing high dynamics manoeuvres, integer ambiguity resolution process has to be performed and in this case, the ambiguity residuals also contribute to the measurement residuals. Hence, to verify if the position estimates can be trusted, a high confidence in the resolution of the ambiguities is required. This necessitates the development of reliable and efficient carrier phase integer ambiguity validation techniques as an integral part of CRAIM.

In addition to the conventional RAIM techniques (PRAIM and CRAIM), extended RAIM (e-RAIM) procedures support detection of faults in the dynamic model and isolate them from the measurement model and vice versa. Most e-RAIM algorithms are derived from the Least Square (LS) estimators of the state parameters in a Gauss-Markov KF.

Recently, Advanced RAIM (A-RAIM) and Relative RAIM (R-RAIM) were proposed as two parallel candidates for future generation integrity monitoring architectures to test the service availability of LPV-200 for worldwide coverage with modernized GNSS and augmentation systems [123, 124]. With double civilian frequencies being transmitted, the ionospheric error can be measured and removed, thus improving the overall system performance. The major difference between these two techniques is in the positioning method. Code-only measurements are used in A-RAIM while both code and time-differenced carrier phase measurements are used in R-RAIM to ensure higher precision without the necessity of an integer ambiguity resolution algorithm [125-127].

R-RAIM can be further divided into range domain R-RAIM and the position domain R-RAIM based on the location where observations from two time epochs are combined together. This distinction results in in different error and projection matrices. With advantages in R-RAIM, the trade-off is more complicated errors and projection matrices, and therefore a more complicated process to transfer these errors with given risks in RAIM.

The efficiency of RAIM is closely dependent on the receiver-satellite geometry and this implies that RAIM may not always be available. Many preliminary investigations on RAIM have been performed [128-131] and the inference is that the stand alone GNSS constellation does not provide the required RAIM availability and GNSS must be augmented if it is to be used as a primary navigation means for en route to non-precision phases of flight. The non GNSS measurements which are helpful to improve the RAIM availability include barometric altitude, receiver clock coasting, geostationary satellite range, etc.

The integrity performance (in terms of detection and protection limit) provided by RAIM is a complex function of:

- the number of visible satellites that can be accessed at any given time;

- the receiver-satellite geometry;

- the probability density function of the error distribution in the received pseudorange signals from each satellite;

- the availability of each broadcast signal, which refers to the percentage of time that each satellite broadcasts a ranging signal;

- integration with other sensors/aids (VOR/DME, baro, inertial).

RAIM performance for receivers using a GPS-only constellation has been thoroughly analysed and some recommendations have been provided in the RTCA MOPS [36]. Also, some work has been performed to characterise RAIM performance against the identified factors [132, 133].

RAIM does not impose extensive hardware modifications to the receiver and it is a feasible and effective way to detect and mitigate single spoofing signals. Some recent studies have been performed to explore the potential of RAIM to deal with radiofrequency interference. A recursive RAIM technique is being employed for detecting and mitigating more than one spoofing signal without other independent information [1344].

Until September 27, 2009, a RAIM prediction was not expected to be performed for an RNAV route conducted where Air Traffic Control (ATC) provides RADAR monitoring or RNAV departure/arrival procedure that has an associated RADAR required note charted. After this date, operators filing RNAV 2 routes (Q and T), RNAV 1 STARs, and RNAV 1 DP's will need to perform a RAIM prediction as part of their pre-flight planning. ICAO Annex 10 and ICAO PBN manual require the ANSPs to provide timely warnings of GNSS RAIM outages. A pre-flight GNSS RAIM prediction analysis is required by the FAA for flights intending to use RNAV/RNP routes as well as departure and arrival procedures while using GPS as the sole navigation source. Therefore, RAIM prediction results are dispatched to pilots, flight dispatchers, Air Traffic Controllers (ATCo) and airspace planners. The use of appropriate RAIM prediction services is considered as a prerequisite for these GNSS approvals. Pilots and ATCo need such information to ensure proper flight planning during possible service unavailability. Since RAIM not only aims at detecting faults but also at evaluating the integrity risk, both the detector and the estimator influence the RAIM performance.

In a multi-constellation environment, RAIM can exploit redundant measurements to achieve self-contained FD/FDE at the user receiver. With the modernisation of GPS, the full deployment of GLONASS and the emergence of GALILEO, BDS, QZSS and IRNSS, an increased number of redundant ranging signals become available, which has recently drawn a renewed interest in RAIM. In particular, RAIM can help alleviating the requirements on ground monitoring. For example, recent research has been focussing on A-RAIM employing multi-GNSS for vertical guidance of an aircraft [135].

RAIM Methods

A number of different RAIM algorithms have been developed over the years. Some are primarily design tools that predict whether or not RAIM will be available for a given position at a given time, whilst others are implemented within receivers to perform FD/FDE procedures.

Some techniques use a filtering or averaging technique such as the position comparison method. However, the majority of RAIM techniques use a so-called "snapshot" approach in which only measurement data from a single epoch is used to check the consistency of the solution. Such techniques include the range comparison method, the residual analysis method, and the parity method [136-139]. With these methods, only current redundant measurements are used in the self-consistency check. On the other hand, both past and present measurements can also be used along with a priori assumptions with regard to vehicle motion, to provide a decision.

The snapshot approach has gained more acceptance than the other in recent times because it has the advantage of not having to make any questionable assumptions about how the system got to its present state. It matters only that the system is in a particular state, and the RAIM decision as to failure or no-failure is based on current observations only [140]. These RAIM methods provide the same level of integrity, i.e. fundamentally these methods are identical, although conceptually they appear quite different and often represented by single Least Squares Residuals (LSR) [139].

Some RAIM techniques have also been designed to use data collected and filtered over multiple epochs. This generates predicted measurements (receiver-satellite ranges) with which to compare each new measurement. Although such techniques promise very high performance, especially when combined with additional sensors such as inertial navigation system, they are difficult to model and analyse in the general case.

The basic measurement relationship for RAIM is described by an over-determined system of linear equations of the form and is given by:

$$y = G\, x_{true} + \varepsilon \qquad (174)$$

where y is the difference between the actual measured range (or pseudorange) and the predicted range based on the nominal user position and the clock bias (y is an n×1 vector); n is the number of redundant measurements; x_{true} are three components of true position deviation from the nominal position plus the user clock bias deviation (x_{true} is a 4×1 vector); ε is the measurement error vector caused by the usual receiver noise, vagaries in propagation, imprecise knowledge of satellite position and satellite clock error, SA (if present), and, possibly, unexpected errors caused by a satellite malfunction (ε is an n×1 vector); and G is the usual linear connection matrix obtained by linearizing about the nominal user position and clock bias (G is an n×4 matrix).

Both the RAIM detector and the estimator have been investigated in the literature. With regard to fault detection, two RAIM algorithms have been widely implemented over the past 25 years: chi-squared ($\chi2$) RAIM (also called parity-based or residual-based RAIM) and Solution Separation (SS) RAIM [132, 133]. Fundamental differences between the two algorithms have been pointed out in [141], but it remains unclear whether SS or $\chi2$ RAIM provides the lowest integrity risk. In parallel, with regard to estimation, researchers have explored the potential of replacing the conventional Least-Squares (LS) process with a Non-Least-Squares (NLS) estimator to lower the integrity risk in exchange for a slight increase in nominal positioning error. The resulting

methods show promising reductions in integrity risk but they are computationally expensive for real-time aviation implementations.

The LS residual method uses all the measurements $\hat{y}$ to calculate a LS estimate of the n-component GNSS position and clock offset or other companion quantity x using:

$$\hat{x} = (H^T H)^{-1} H^T \hat{y} \tag{175}$$

where the subscript ($\hat{}$) denotes an estimate, H is the meaurmeent matrix. The least-squares solution is then used to predict the six range measurements:

$$\hat{y} = H\hat{x} \tag{176}$$

The residual between the prediction and the measurements indicates any inconsistency that may arise due to the mentioned measurement errors and is given by:

$$w = y - \hat{y} = [I - H(H^T H)^{-1} H^T]y \tag{177}$$

The observable in this integrity monitoring method is the sum of the squares of the residuals, which is always a positive scalar and is given by:

$$SSE = w^T w \tag{178}$$

If all elements of ϵ are zero-mean Gaussian distributions, then the Square Sum Error (SSE) has a non-normalized chi-square distribution with $(n-4)$ degrees of freedom. In order to have a linear relationship between a satellite error and the induced test-statistic, an assumption is to assess $\sqrt{SSE/(n-4)}$. Apart from the SSE technique, an alternative method that employs a linear transformation to the measurement y is given by:

$$\begin{bmatrix} \hat{x} \\ \cdots \\ p \end{bmatrix} = \begin{bmatrix} (H^T H)^{-1} H^T \\ \cdots \\ P \end{bmatrix} [y] \tag{179}$$

The parity vector p is the result of multiplying y by an $(n-4) \times n$ matrix P, which is derived from a singular value decomposition or a QR factorization of the observation matrix H, which makes the rows of P mutually orthogonal and unity in magnitude. If the measurement error vector ϵ has independent random elements that are zero-mean and normally distributed, then the following equations hold true:

$$p = Pw \tag{180}$$

$$p = P\epsilon \tag{181}$$

$$p^T p = w^T w = SSE \tag{182}$$

It implies that the magnitudes of p and w are the same inspite of their different dimensionalities. Integrity alerts can therefore be set using either SSE or p as a test statistic. Under the assumption that the measurement noise is Gaussian with zero-mean and standard deviation σ_ϵ, the quantity $\|w\|^2/\sigma_\epsilon^2$ is a chi-square distributed random variable with $(n-4)$ degrees of freedom (a minimum of four satellites are required, and the degrees of freedom are equal to the number of redundant measurements). This is represented mathematically as:

$$\frac{\|w\|^2}{\sigma_\epsilon^2} \sim \chi^2(n-4) \tag{183}$$

The threshold value R is set for the test-statistic to achieve any desired probability of False Alarm (FA) under Normal Conditions (NC) and is given by:

$$P(FA|NC) = P(\|w\| > R|NC) =$$

$$\frac{1}{2^{(n-4)/2}\Gamma(\frac{n-4}{2})} \int_{R^2/\sigma_\epsilon^2}^{\infty} s^{(\frac{n-4}{2}-1)} e^{\frac{-s}{2}} ds \tag{184}$$

where the integral is the incomplete gamma function. Given the number of visible satellites n, and $(FA|NC)$, the threshold R can be solved for. A protection radius or HAL a is set based on the RNP.

By the number of alternative hypotheses, there are two types of RAIM algorithms for both A-RAIM and R-RAIM: the classic RAIM method with single alternative hypothesis and the Multi Hypothesis Solution Separation (MHSS) RAIM method [142]. The classic RAIM method is typically used in Range Domain R-RAIM (RD-RAIM), while MHSS is used in A-RAIM. Typically, A-RAIM is adopted as the preferential method and Position Domain R-RAIM (PD-RAIM) is employed in case A-RAIM is not available [125].

4.6.4. Aircraft Based Augmentation Systems

In addition to the existing SBAS, GBAS and RAIM techniques, GNSS augmentation may take the form of additional information being provided by other on-board avionics systems. As these systems normally operate via separate principles than the GNSS, they are not subject to the same sources of error or interference. A system such as this is referred to as an Aircraft Based Augmentation System or ABAS. ABAS is different from RAIM, in which the aircraft characteristics (flight dynamics, body shape, antenna location, EMC/EMI, etc.) are typically not considered and various kinds of consistency check are accomplished using the GNSS signals to detect and exclude faulty/unreliable satellites.

The additional sensors used in ABAS may include Inertial Navigation Systems (INS), VOR-DME/TACAN, Radar, Vision Based Sensors, etc. Unlike SBAS and GBAS technology, published research on ABAS is limited and mainly concentrates on additional information being blended into the position calculation to increase accuracy and/or continuity of the integrated navigation solutions. In recent years, significant efforts were made for developing ABAS architectures capable of generating integrity signals suitable for safety-critical GNSS applications (e.g., aircraft precision approach and landing), which led to the development of an Avionics Based Integrity Augmentation (ABIA) system. Implementing a Flight Path Guidance (FPG) module, an ABIA system can provide steering information to the pilot and, additionally, electronic commands to the aircraft/UAS Flight Control System (FCS), allowing for real-time and continuous integrity monitoring, avoidance of safety/mission-critical flight conditions and rapid recovery of the RNP in case of GNSS data degradation or loss. The architecture of an advanced ABIA system is presented in Fig. 33.
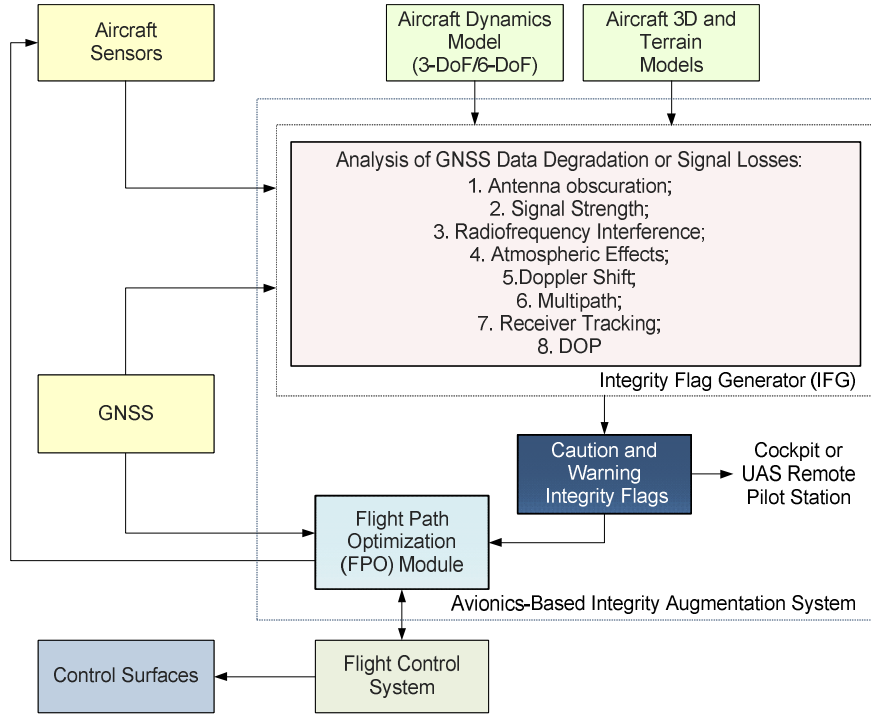
Fig. 33. ABIA system architecture.

This system addresses both the predictive and reactive nature of GNSS integrity augmentation. To comprehend this concept, some key definitions of alerts and TTA's applicable to the ABIA system are presented below [53]:

- Caution Integrity Flag (CIF): a predictive annunciation that the GNSS data delivered to the avionics system is going to exceed the Required Navigation Performance (RNP) thresholds specified for the current and planned flight operational tasks (GNSS alert status).

- Warning Integrity Flag (WIF): a reactive annunciation that the GNSS data delivered to the avionics system has exceeded the Required Navigation Performance (RNP) thresholds specified for the current flight operational task (GNSS fault status).

- ABIA Time-to-Caution (TTC): the minimum time allowed for the caution flag to be provided to the user before the onset of a GNSS fault resulting in an unsafe condition.

- ABIA Time-to-Warning (TTW): the maximum time allowed from the moment a GNSS fault resulting in an unsafe condition is detected to the moment that the ABIA system provides a warning flag to the user.

Based on the above definitions, two separate models for the time responses associated to the Prediction-Avoidance (PA) and Reaction-Correction (RC) functions performed by the ABIA system are defined (Fig. 39). The PA time response is given by:

$$\Delta T_{PA} = \Delta T_{Predict} + \Delta T_{C-Report} + \Delta T_{Avoid} \qquad (185)$$

where:

$\Delta T_{Predict}$=Time required to predict a critical condition;

$\Delta T_{C-Report}$=Time required to communicate the predicted failure to the FPG module;

ΔT_{Avoid}=Time required to perform the avoidance manoeuvre.

In this case, $\Delta T_{Avoid} \leq TTC$.

If the available avoidance time ΔT_{Avoid} is not sufficient to perform an adequate avoidance manoeuvre (i.e., $\Delta T_{Avoid} > TTC$), the aircraft will inevitably encroach on critical conditions causing GNSS data losses or unacceptable degradations. In this case, the RC time response applies:

$$\Delta T_{RC} = \Delta T_{Detect} + \Delta T_{W-Report} + \Delta T_{Correct} \qquad (186)$$

where:

ΔT_{Detect}=Time required to detect a critical condition;

$\Delta T_{W-Report}$=Time required to communicate the failure to the FPG module;

$\Delta T_{Correct}$=Time required to perform the correction manoeuvre.
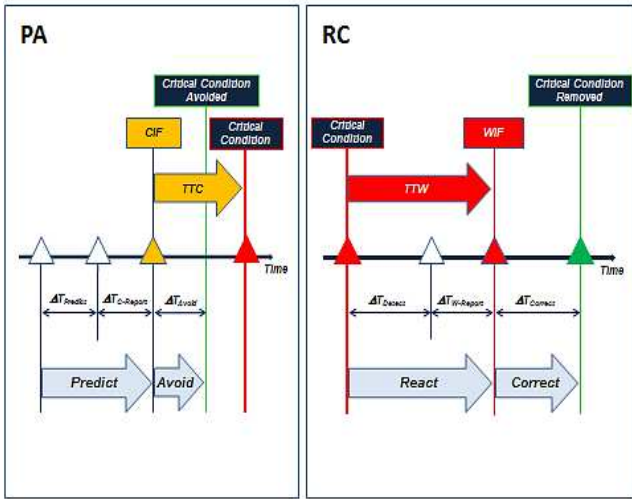
In general, the condition to be satisfied is expressed as:

$$\Delta T_{Detect} + \Delta T_{W-Report} \leq TTW \qquad (187)$$

The RC time response is substantially equivalent to what current GBAS and SBAS systems are capable of achieving. A comparison between Fig. 34 (a) and Fig. 34 (b) allows to immediately visualise the benefits introduced by the ABIA PA function. Further progress is possible adopting a suitable algorithm in an Integrity Flag Module (IFG) module capable of initiating an early correction manoeuvre as soon as the condition $\Delta T_{Avoid} \leq TTC$ is violated. In this case, the direct Prediction-Correction (PC) time response would be:
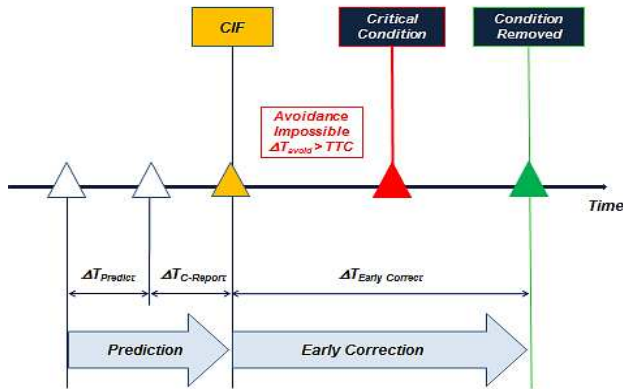
$$\Delta T_{PC} = \Delta T_{Predict} + \Delta T_{C-Report} + \Delta T_{Eraly\ Correct} \qquad (188)$$

This concept is illustrated in Fig. 35. By comparing with Fig. 34, it is evident that the ABIA system would be able to reduce the time required to recover from critical conditions if the following inequality is verified:

$$\Delta T_{Eraly\ Correct} < TTC + \Delta T_{Detect} + \Delta T_{C-Report} + \Delta T_{Correct} \qquad (189)$$
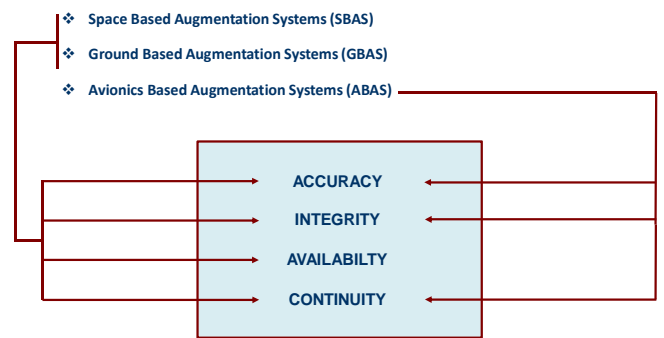
(a)   (b)
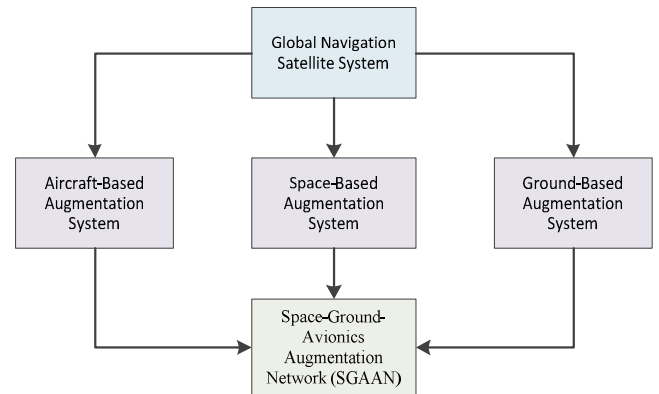
Fig. 34. ABIA PA and RC functions.

Fig. 35. ABIA PC function.

Targeting an ABIA implementation, a dedicated analysis is required in order to determine the flight envelope limitations associated with the use of GNSS. In particular, the following models have to be introduced [53, 54]:

- The Antenna Obscuration Matrixes (AOM) in azimuth and elevation, constructed as a function of attitude (Euler) angles in all relevant aircraft configurations;

- The GNSS Carrier-to-Noise and Jamming-to-Signal Models (CJM), accounting for the relevant transmitter/receiver characteristics, propagation losses and RF interference;

- The Multipath Signal Model (MSM) including fuselage, wing and ground path fading components and the associated range errors;

- The Doppler Shift Model (DSM) and associated critical conditions causing GNSS tracking issues.

Using appropriate aircraft dynamics models, the manoeuvring envelope of the aircraft can be determined in all required flight conditions. Using the AOM, CJM, MSM and DSM models, together with the GNSS receiver tracking models and the manoeuvring requirements of specific flight tasks (e.g., test/training missions or standard airport approach procedures), it is possible to identify the conditions that are potentially critical for the on-board GNSS system and set appropriate thresholds for the ABIA CIFs and WIFs, thereby generating timely alerts when the aircraft is performing critical manoeuvres prone to induce GNSS signal degradations or losses.

5. Towards a Space-Ground-Aircraft Augmentation Network

Current SBAS and GBAS technologies are not able to meet the stringent GNSS data integrity requirements imposed by airworthiness regulations in some of the most demanding operational tasks (e.g. UAS sense-and-avoid). GBAS provides precision position services limited to use in the approach phase only. SBAS provides a precision position service in all flight phases, but GBAS provides better accuracy in the approach phase. On the other hand, the ABAS/ABIA system approach is particularly well suited to distinctively increase the levels of integrity and accuracy (as well as continuity in multi-sensor data fusion architectures) of GNSS in a variety of mission- and safety-critical applications. Synergies between these three GNSS augmentation systems (Fig. 36) can be studied to develop a Space-Ground-Aircraft Augmentation Network (SGAAN) that can be used for a number of safety-critical aviation applications (Fig. 37).

Fig. 36. Synergies between SBAS, GBAS and ABAS.

Fig. 37. Space-ground-aircraft augmentation network.

Once the reliability of the mathematical algorithms for ABAS/SBAS/GBAS is established, dedicated IFG modules can be implemented for alerting the pilot/UAS remote pilot when the critical conditions for GNSS signal losses are likely to occur (Fig.38).

The ABAS IFG is designed to provide caution and warning alerts in real-time (i.e., in accordance with the specified TTC and TTW requirements in all relevant flight phases). IFG module inputs are from the GNSS Receiver and aircraft flight dynamics. A GNSS Constellation Simulator (GCS) can suppors the GNSS satellite visibility, signal and geometry analysis, while an Aircraft Dynamics Simulator (ADS) can be used to generate the nominal flight path trajectory and attitude (Euler) angles. Additionally, an Aircraft 3-Dimensional Model (A3DM) in CATIA (Computer Aided Three-dimensional Interactive Application) and a Terrain and Objects Database (TOD) are also required to run the MPS. Using a Digital Terrain Elevation Database (DTED), it is possible

to obtain a detailed map of the terrain beneath the aircraft. Providing the aircraft trajectory inputs from the ADS module, terrain elevation data can be automatically extracted and fed to the TOM module where they are integrated with the database of man-made objects (e.g., buildings). The Doppler Analysis Module (DAM) calculates the Doppler shift by processing ADS and GCS inputs. The Multipath Analysis Module (MAM) processes the A3DM, TEM, GCS and ADS inputs to determine multipath contributions from the aircraft (wings/fuselage) and from the terrain/objects close to the aircraft. The Obscuration Matrix Module (OMM) receives inputs from the A3DM, GCS and ADS, and computes the GNSS antenna(e) obscuration matrixes for all aircraft manoeuvres. The C/N_0 and J/S Analysis Module (SAM) calculates the nominal link budget of the direct GNSS signals received by the aircraft in the presence of atmospheric propagation disturbances, as well as the applicable RF interference signal levels. The definition of suitable ABAS integrity thresholds is heavily dependent on the aircraft dynamics and geometric characteristics. Further details can be found in the references [53, 54, 148-150]

The IFGs for SBAS and GBAS introduce the system functionalities required to compute the VPL and HPL and to compare them with the designated VAL and HAL (Fig. 38). As discussed, only the LAAS MOPS introduces the concept of predictive integrity flags (PVPL and PLPL). However, the standard does not provide detailed guidance on how these features can be implemented in practice to exploit potential synergies with ABAS. Additionally, PVAL/PLPL features are not present in the WAAS MOPS [41]. VAL and HAL defined in the WAAS MOPS and are listed in Table 20. The SBAS VAL is the threshold used to generate integrity flags

based on the SBAS VPL. Similarly, the SBAS HAL is the threshold used to generate integrity flags based on the SBAS HPL. For oceanic, en route, terminal, or Lateral Navigation/Vertical Navigation (LNAV/VNAV) approaches, the VAL is not defined by the WAAS MOPS and ICAO SARPs [38, 39]. LNAV/VNAV approaches use lateral guidance (556 m lateral limit) from GNSS and/or GBAS and vertical guidance provided by either barometric altimeter or GBAS. Aircraft that do not use GBAS for the vertical guidance portion must have VNAV-capable altimeters, which are typically integrated with modern flight directors and/or Flight Management Systems (FMS).

Table 20. SBAS vertical and horizontal alert limits [41].

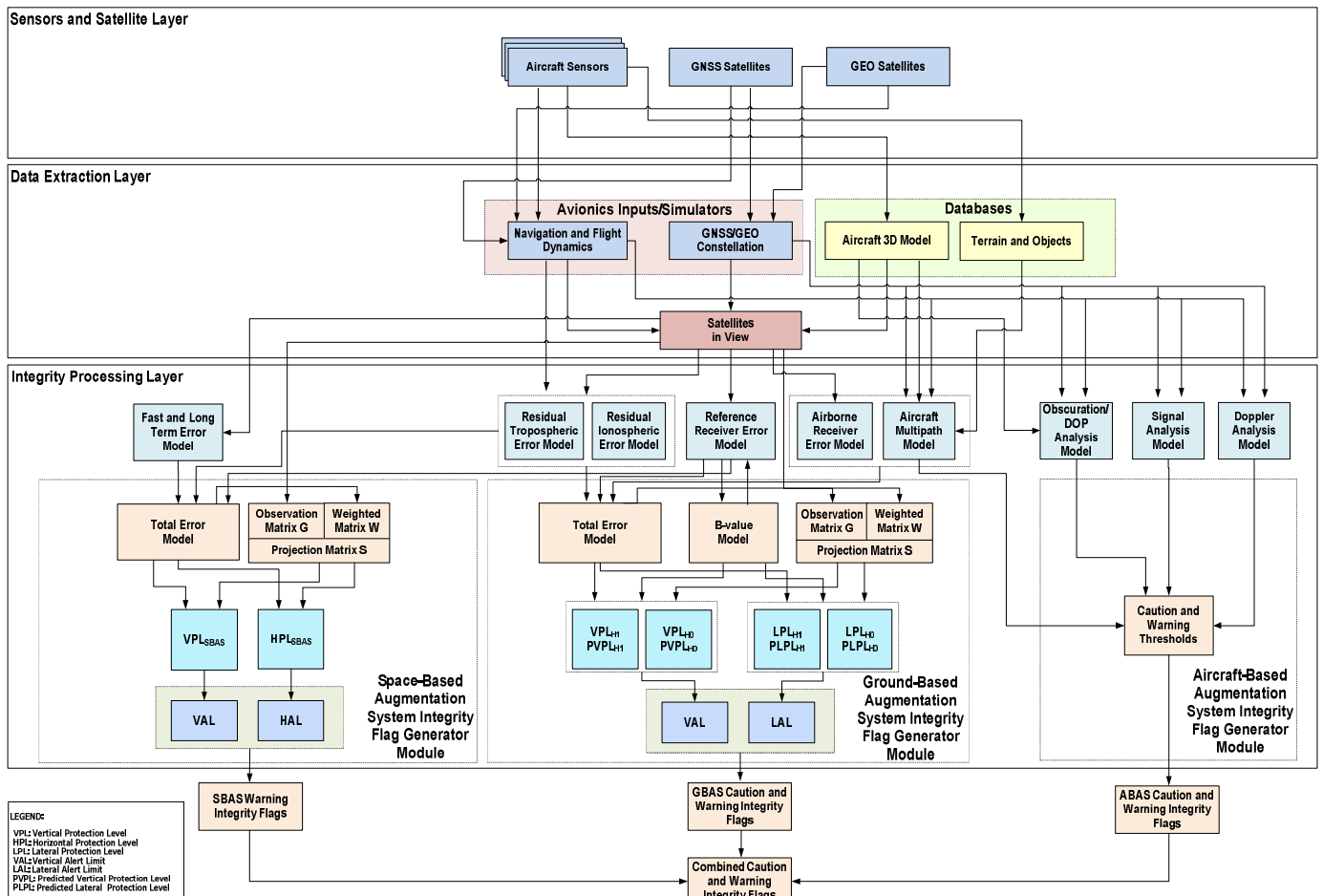| Navigation Mode	Vertical Alert Limit (VAL)	Horizontal Alert Limit (HAL)
Oceanic/Remote	NA	7408 m
En Route	NA	3704 m
Terminal	NA	1852 m
LNAV or LNAV/VNAV Approach	NA	556 m
LNAV/VNAV Approach (after FAWP)	50 m	556 m
LPV or LP Approach	50 m	40 m
LPV II	35 m	40 m

Fig. 38. SGAAN integrity Augmentation architecture.

Localizer Performance with Vertical Guidance (LPV) enables the aircraft descent to 200-250 feet above the runway and can only be flown with a Wide Area Augmentation System (WAAS) and European Geostationary Navigation Overlay Service (EGNOS) receiver. LPV approaches are operationally equivalent to the legacy Instrument Landing System (ILS) but are more economical because no navigation infrastructure has to be installed in proximity of the runway. Localizer Performance (LP) is a Non-Precision Approach (NPA) procedure that uses the high precision of LPV for lateral guidance and barometric altimeter for vertical guidance. LP approaches can only be flown by aircraft equipped with WAAS/EGNOS receivers. The minimum descent altitude for the LP approach is expected to be approximately 300 feet above the runway. The VAL values are listed in Table 21.

Table 21. Vertical Alert Limit [40].

Intended GSL	Height above LTP/FTP (feet)	Alert Limit (meters)
A	---	FASVAL
B	---	FASVAL
C	H≤200	FASVAL
	200<H≤1340	0.02925*H(ft)+ FASVAL-5.85
	H>1340	FASVAL+33.35
D,E,F	H≤100	FASVAL
	100<H≤200	FASVAL
	200<H≤1340	0.02925*H(ft)+ FASVAL-5.85
	H>1340	FASVAL+33.35

Similarly, the Lateral Alert Limit (LAL) defines the threshold values of LPL and PLPL. The LAL values are listed in Table 22. The Final Approach Segment Vertical and Lateral Alert Limits (FASVAL and FASLAL) are listed in Table 23.

Table 22. Lateral Alert Limit [40].

Intended GSL	Distance from LTP/FTP (meters)	Alert Limit (meters)
A,B	---	FASLAL
C	Along the runway and for D≤873	FASLAL
	873<D≤7500	0.0044*D(m)+ FASLAL-3.85
	D>7500	FASLAL+29.15
D,E,F	Along the runway and for D≤291	FASLAL
	291<D≤873	0.03952*D(m)+ FASLAL-11.5
	873<D≤7500	0.0044*D(m)+ FASLAL+19.15
	D>7500	FASLAL+52.15

Table 23. FASVAL and FASLAL.

GSL	FASVAL	FASLAL
A,B,C	≤10 m	≤40 m
A,B,C,D	≤10 m	≤17 m
A,B,C,D,E	≤10 m	≤17 m
A,B,C,D,E,F	≤10 m	≤17 m

Based on these alert limits and limitations, the criteria for producing SBAS/GBAS CIFs and WIFs can be set and are listed below:

- When VPL_{SBAS} exceeds VAL or HPL_{SBAS} exceeds HAL, the WIF is generated;
- When $PVPL_{GBAS}$ exceeds VAL or $PLPL_{GBAS}$ exceeds LAL, the CIF is generated;
- When VPL_{GBAS} exceeds VAL or HPL_{GBAS} exceeds HAL, the WIF is generated.

The performance of SBAS and GBAS can be enhanced by adding ABIA models to the existing standards [36, 40, 41]. As both SBAS and GBAS use redundant GNSS satellite observations to support FDE within the GNSS receiver (i.e., implementing a form of RAIM), some additional integrity flag criteria can be introduced. In a WAAS/RAIM integration scheme, the minimum number of satellites required for FDE is 6. At present, no information is available regarding the provision of RAIM features within LAAS-enabled GNSS receivers. Therefore, the inclusion of a basic form of RAIM within such receiver (i.e., at least 5 satellites are required for FDE) can be assumed. Based on these assumptions, the number of satellites in view can be used to set additional integrity thresholds for SBAS and GBAS [143, 144]. Additionally, new augmentation strategies including Ground-based Regional Augmentation System (GRAS), which is a combination of SBAS and GBAS concepts to enhance GNSS performance, can be used within the SGAAN network. Such network could improve the newly introduced APV procedures (supported by GNSS and/or baro-VNAV) and provide continuous lateral/vertical guidance without the need for a terrestrial radio navigation aid. Recent studies have shown that ABAS systems would work synergically with SBAS and GBAS, enhancing integrity levels in all flight phases, from initial climb to final approach [144]. Therefore, the integration of ABAS/ABIA with SBAS and GBAS is a clear opportunity for future research towards the development of SGAAN suitable for manned and unmanned aircraft applications and for a variety of mission-critical and safety-critical aviation applications, including flight test, precision approach and automatic landing.

6. GNSS for Trusted Autonomous Operations

Current UAS technologies are perceived to have a lack of trust relative to the human-machine teaming aspects. The trust in human-machine teaming becomes even more important in GNSS denied/challenging environments and in providing effective decision-making in such safety-critical tasks.

In modern UAS applications, GNSS supports the development of low-cost and high performance (and relatively low cost and low-volume/weight) navigation and guidance systems. Additionally, when used in conjunction with suitable data link technologies, GNSS-based navigation systems facilitates the provision of Automated Dependent Surveillance (ADS) functionalities for cooperative UAS SAA. In non-cooperative SAA, the adoption of GNSS can also provide the key positioning and, in some cases, attitude data (using multiple antennas) required for automated collision avoidance. A key limitation of GNSS for both cooperative (ADS) and non-cooperative applications is represented by the achievable levels of integrity. Therefore, the implementation of integrity augmentation functionalities could support the development of the IAS suitable for both cooperative and non-cooperative scenarios.

6.1. GNSS Augmentation for UAS SAA

One of the key challenges encountered by the aviation community for integration of UAS into non-segregated airspace is the provision of a Sense-and-Avoid (SAA) capability meeting the stringent integrity requirements set by international standards and national certification authorities. Cooperative and non-cooperative SAA are implemented to address UAS safe integration into the

non-segregated airspace. The SAA capability can be defined as the automatic detection of possible conflicts (i.e., collision threats) by the UAV platform and the implementation of avoidance manoeuvres to prevent the identified collision threats. As part of our research, the possible synergies attainable with the adoption of different detection, tracking and trajectory generation algorithms were studied. Additionally, the error propagation from different sources and the impacts of host and intruders dynamics on the ultimate SAA solution were investigated. The system detection range and Filed-of-View/Field-of-Regard (FOV/FOR) have to be adequate to ensure separation from the intruder to prevent a probable near mid-air collision. A number of cooperative and non-cooperative sensors/systems have been studied recently. Cooperative systems typically include Traffic Collision Avoidance System (TCAS)/Airborne Collision Avoidance System (ACAS) and Automatic Dependent Surveillance Broadcast (ADS-B). The inclusion of ADS-B in association with GNSS-based navigation systems redefines the paradigm of Cooperative SAA (C-SAA) in the CNS/ATM context by allowing the share of accurate GNSS trajectory information. Optical, thermal, LIDAR, MMW Radar and acoustic sensors can be used as part of Non-Cooperative SAA (NC-SAA) architectures. As an example, a combined C-SAA/NC-SAA architecture is depicted in Fig. 39 with an identification of primary (solid line) and auxiliary sensors (dashed line) for cooperative and non-cooperative SAA tasks. Additionally, ATM primary surveillance radar and Air Traffic Controller (ATCo) digital instructions using Controller to Pilot Data Link Communications (CPDLC) are considered. The sequential steps involved in the SAA process for executing an efficient Tracking, Deciding and Avoiding (TDA) loop are also illustrated in Fig. 39. Criticality analysis is carried out to prioritize (i.e., to determine if a collision risk threshold is exceeded for all tracked intruders) and to determine the action commands. If an avoidance action is required, the SAA system generates an optimal avoidance trajectory using a fast-converging guidance algorithm such as differential geometry [145-147].

PMO techniques would have the benefit of providing a mathematical optimum. However, they can be used instead of DCO only if the SAA time horizon allows (i.e., only if the time to conflict is much greater than the computational time required to obtain an optimal trajectory). This could be the case, for example, in properly developed air traffic separation maintenance systems. Error analysis is performed to determine the overall uncertainty volume in the airspace surrounding the intruder tracks based on the cooperative/non-cooperative unified method described in [146]. This is accomplished by considering both the navigation and the tracking errors affecting the measurements and translating them to unified range and bearing uncertainty descriptors. As discussed in the previous section, the SGAAN research demonstrated the potential of this new technology to enhance GNSS integrity performance in a variety of mission- and safety-critical applications including experimental flight test/flight inspection, precision approach and automatic landing (also in synergy with GBAS and SBAS) [148]. Therefore, an ABIA system concept was developed for UAS applications (Fig. 40).
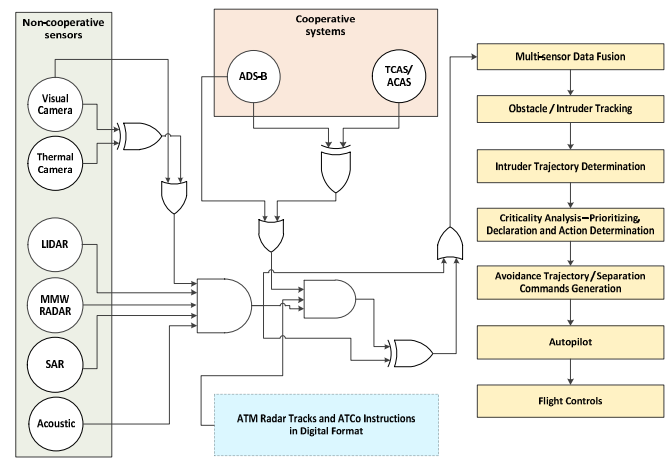

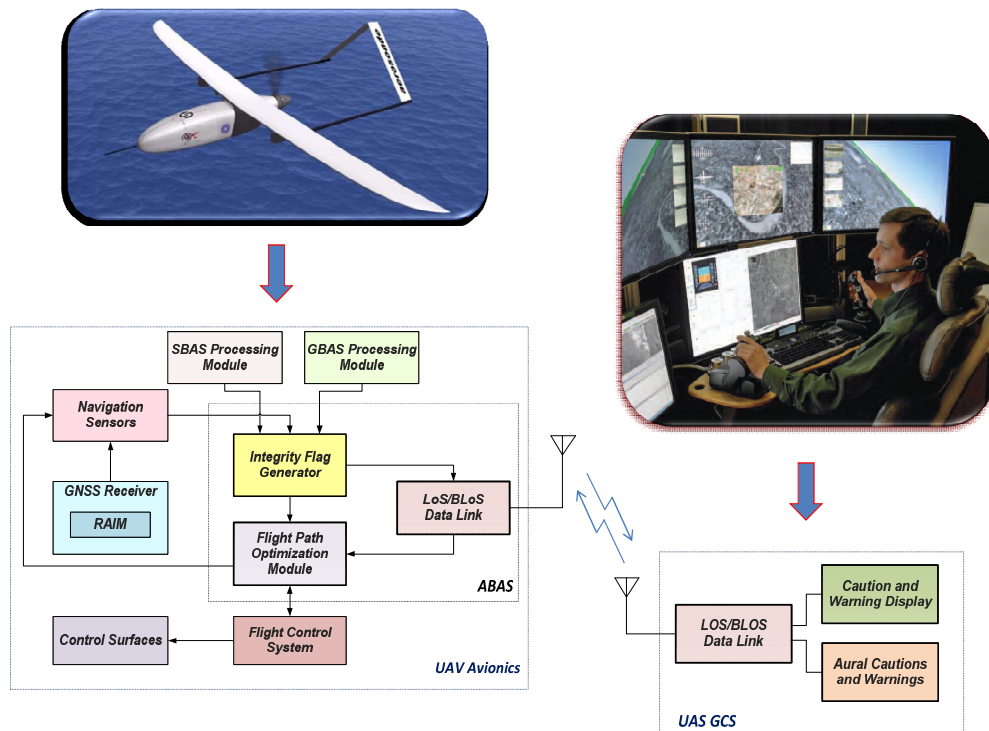**Fig. 39.** SAA system architecture. Adapted from [146].

Fig. 40. ABIA system architecture for UAS applications [148].

As in a manned aircraft version, this system performs a continuous monitoring of GNSS integrity levels in flight by analysing the relationships between aircraft manoeuvres and GNSS accuracy degradations or signal losses (Doppler shift, multipath, antenna obscuration, signal-to-noise ratio, jamming, etc.). In case of any detected or predicted integrity threshold violation, the ABIA system provides suitable warning or caution signals to the UAV AFCS and to the remote Ground Control Station (GCS), thereby allowing timely correction manoeuvres to be performed. This increased level of integrity could provide a pathway to support unrestricted access of UAS to all classes of airspace. A possible ABIA/SAA integration architecture is illustrated in Fig. 41. Position, Velocity, and Attitude/Rates (PVA) measurements are obtained by using the various navigation sensors/systems on-board the aircraft (i.e., implementing multi-sensor data fusion techniques).

The key advantage is that the safe avoidance is determined by evaluating the risk-of-collision and then a safe manoeuvring point is identified from where the host UAV can manoeuvre safely (i.e., any manoeuvre can be performed within the UAV operational flight envelope). The risk of collision is evaluated by setting a threshold on the probability density function of a near mid-air collision event over the separation volume. The risk-of-collision is zero at the safe manoeuvring point. If both the safe-separation thresholds are violated, a mid-air collision threat is detected and the SAA WIF is generated. To prevent any WIF, the flight path optimization process starts when the first CIF is generated. PMO and DCO techniques are used to generate a new optimised trajectory free of any integrity degradations. Depending on the relationship between the available time-to-collision and the computation time required to generate the optimal trajectory, PMO or DCO solutions are applied. The optimised trajectory data are sent to the AFCS (and to the ground pilot) for execution of the avoidance manoeuvres. In the trajectory optimisation process, the aircraft 3DOF/6DOF model defines the dynamics constraints, while the satellite elevations and the aircraft heading rates are used as path constraints. Results from simulation case studies confirmed that the Integrity-Augmented SAA (IAS) solution is capable of performing high-integrity conflict detection and resolution when GNSS is used as the primary source of navigation data [148-150].
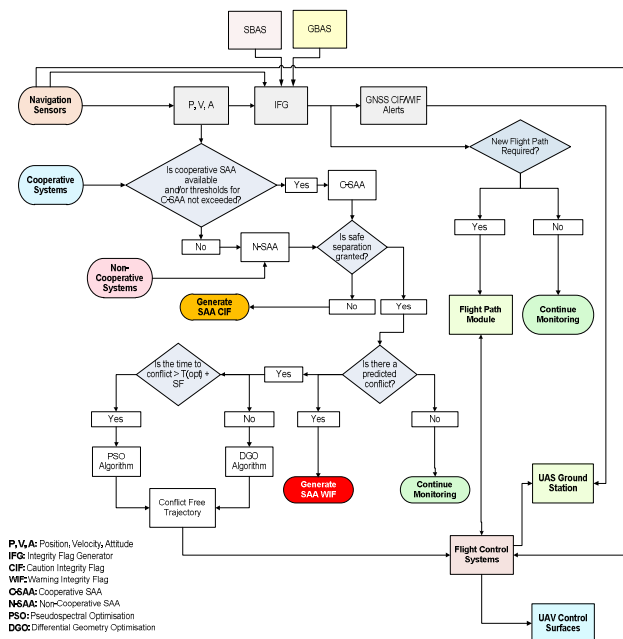
ABIA/SAA integrated architecture [148].

6.1. Resilience in GNSS Challenged Environments

In GNSS challenged environments, measurement models of distributed sensors and GNSS jammer location estimation models can support the design of a model-predictive approach. The developed models address uncertainty in platform navigation and jammer localization methods.

In the aviation context, complex cyber-physical systems are being used on board aircraft (avionics/mission systems) and on the ground (CNS/ATM systems, Airline Operational Centres, Military Command and Control, etc.). These cyber-physical systems have integrated computational and physical elements, including CNS sensors/systems, control systems, data networks and external communications. The complexity of these ever more interconnected and interleaved systems can create multiple opportunities for a number of cyber-physical threats to materialise.

Signals from commercial high power transmitters, ultra wideband radar, television, VHF, mobile satellite services and personal electronic devices can interfere with the GNSS signals. There are three distinct forms of deliberate interference with GNSS signals including: jamming, spoofing and meaconing. There have been numerous documented examples of intentional and unintentional GPS jamming and spoofing. For example, employing Automatic Dependent Surveillance–Broadcast (ADS-B) systems based on GNSS, are subject to a number of cyber-attacks. Three types of threats ADS-B message have been identified: message corruption, message denial and message delay. ADS-B using GNSS is inherently vulnerable to hacking, jamming, spoofing and meaconing because of its open architecture and unencrypted signals, and because equipment is easy to obtain. Several anti-spoofing techniques have been proposed in the open literature and can generally be classified into two main categories: spoofing detection and spoofing mitigation. Jamming is likely to produce the most significant impact on aviation GNSS operations. Jamming can be split into 4 broad areas: accidental, criminal, red team deliberate and blue team deliberate.

Earlier attempts on assessing vulnerability of civil GPS receivers to jamming were made by the UK Defence Research Agency (DRA), which was later incorporated into the Defence Evaluation and Research Agency (DERA) and, successively, into QinetiQ. The effects of a low-power jammer were demonstrated by flight tests using a BAC-111 conducted at UK's Farnborough facility. According to [151], a jammer radiating 1 W of FM noise in GPS 1.6 GHZ frequency band was able to jam a civil GPS receiver at up to 22 km. The differential signals in a DGNSS receiver are also vulnerable to various types of jamming including terrorist attacks. For example, the DGNSS facility can be outfitted with an antenna designed to null out a jamming signal. Because of the effective jamming range double every 6 dB increase in transmitter power, it is possible to build a small jammer capable of interfering with civil GPS receivers at ranges in excess of 50 km.

To minimize the susceptibility to GNSS jamming, combat aircraft are outfitted with a Controlled-Reception Pattern Antennas (CRPA), which are designed to detect a jamming signal and desensitize the antenna in the direction of the jammer. When the GNSS receiver is integrated with an INS, the INS can take over as the principal navigation system until the aircraft tis flown beyond the range of the jammer. Furthermore, GNSS signals can be degraded by natural and artificial obstacles in difficult scenarios (e.g., urban canyons and mountainous areas), requiring the need for effective augmentation strategies to be implemented [152].

7. Conclusions and Future Research

A detailed review of Global Navigation Satellite Systems (GNSS) aviation applications was presented. In recent years, various augmentation strategies have been proposed and developed to

increase the levels of GNSS accuracy and integrity in aviation mission-essential and safety-critical applications. The review covered all existing approaches, including Satellite Based Augmentation Systems (SBAS), Ground Based Augmentation Systems (GBAS), Aircraft Based Augmentation Systems (ABAS), Receiver-Autonomous Integrity Monitoring (RAIM) and multi-sensor data fusion. Suitable mathematical models were presented addressing the main sources of GNSS data degradation or loss in flight including: antenna obscuration, geometric accuracy degradation, fading, multipath and Doppler shift. Some case studies were also presented investigating the synergies attainable from the online integration of ABAS with SBAS and GBAS, with a special focus on integrity augmentation features. Finally, the potential of integrating SBAS-GBAS-ABAS with existing UAS cooperative and non-cooperative SAA architectures was investigated. It was concluded that this integration leads to an Integrity Augmented SAA (IAS) solution that is well suited for a variety of mission-essential and safety-critical aviation applications. Therefore, the integration of SBAS/GBAS/ABAS is a clear opportunity for future research towards the development of a Space-Ground-Avionics Augmentation Network (SGAAN) suitable for manned and unmanned aircraft applications and for a variety of mission-essential and safety-critical GNSS operational tasks, including flight test, precision approach and automatic landing.

Future GNSS research in the aviation context will concentrate on the following main areas:

- Evaluate the potential of GNSS integrity augmentation techniques to enhance the performance of next generation Communication, Navigation and Surveillance/Air Traffic Management (CNS/ATM) decision support tools for Performance/Intent Based Operations (PBO/IBO) and 4DT management;

- Investigate the potential of GNSS integrity augmentation techniques to enhance the performance of Next Generation Flight Management Systems (NG-FMSs) for manned and unmanned aircraft. This will require an evolution of current FMS architectures introducing suitable integrity monitoring and augmentation functionalities for 4-Dimensional Trajectory (4DT) planning and update throughout all flight phases;

- Evaluate the potential of introducing ionospheric scintillation models in the GNSS integrity augmentation systems. In particular, future research should focus on possible mission planning implementations useful when flying over regions affected by intense scintillation and/or in times of predicted high scintillation.

- Further investigate the potential of GNSS integrity augmentation techniques to support trusted autonomous system applications by addressing other navigation, communication and surveillance systems in the CNS+A context;

- Investigate the potential of integrity augmentation techniques to enhance GNSS performance in aircraft surface operations;

- Investigate the potential of GNSS augmentation concepts to support aviation forensic applications (i.e., accident and incident investigation);

- Investigate the potential synergies attainable by the employment of GNSS integrity augmentation systems in urban canyons as well as to provide persistent navigation in sparse and denied environments.

References

[1] E. D. Kaplan and C. J. Hegarty, Understanding GPS: Principles and Applications. Artech House, 2006.

[2] P. Daly, "GLONASS Approaches Full Operational Capability (FOC)," Proceedings of the 8th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1995), Palm Springs, CA, September 1995, pp. 1021-1030.

[3] D. M. Thomis, "Global Positioning System – A Modification to the Baseline Satellite Constellation for Improved Geometric Performance," US Air Force Institute of Technology, 1984.

[4] G. M. Siouris, "Aerospace Avionics Systems," Academic Press. San Diego, CA (USA), 1993.

[5] A. Leick, "GPS Satellite Surveying," Third Edition. John Wiley and Sons, New York, 2003.

[6] V. Ashkenazi, "Principles of GPS and Observables," Lecture Notes, IESSG, University of Nottingham, 1995.

[7] G. Seeber, "Satellite Geodesy," Second Edition, Walter de Gruyter Editor, Berlin (Germany), New York (USA), 2003.

[8] S. Gleason and D. Gebre-Egziabher (Editors), "GNSS Applications and Methods," GNSS Technology and Application Series, Artech House. Norwood, MA (USA), 2009.

[9] V. Ashkenazi, T. Moore and J.M. Westrop, "Combining Pseudo-range and Phase for Dynamic GPS," Paper presented at the International Symposium on Kinematic Systems in Geodesy, Surveying and Remote Sensing. London (UK), 1990.

[10] V. Ashkenazi, T. Moore, G. Ffoulkes-Jones, S. Whalley and M. Aquino, "High Precision GPS Positioning by Fiducial Techniquesl," Global Positioning System: An Overview, Bock Y., Leppard N. Eds., International Association of Geodesy Symposia, Vol. 102, Springer, New York (USA), 1990.

[11] T. Moore, "GPS Orbit Determination and Fiducial Networks," Lecture Notes, IESSG, University of Nottingham (UK), 1996.

[12] J.F.G. Monico, V. Ashkenazi and T. Moore, "High Precision GPS Network with Precise Ephemerides and Earth Body Tide Model," Revista Brasileira de Geofisica, Vol. 15, N. 2, pp. 155-160, 1997.

[13] A. H. Dodson, "Propagation Effects on GPS Measurements," Lecture Notes, IESSG, University of Nottingham (UK), 1997.

[14] R. Giffard, "Estimation of GPS Ionospheric Delay Using Ll Code and Carrier Phase Observables," In Proceedings of the 31st Annual Precise Time and Time Interval (PTTI) Meeting, Dana Point, CA (USA), 1999.

[15] J. A. Klobuchar, "Design and Characteristics of the GPS Ionospheric Time Delay Algorithm for Single Frequency Users," In Proceedings of the IEEE Position, Location, and Navigation Symposium, Las Vegas, NV (USA). 1986.

[16] J. Saastamoinen, "Atmospheric correction for the troposphere and stratosphere in radio ranging satellites," The Use of Artificial Satellites for Geodesy; Henriksen, S., Mancini, A., Chovitz, B., Eds.; American Geophysical Union: Washington, WA, USA, 1972.

[17] J. Saastamoinen, "Contribution of the theory of atmospheric refraction," Bulletin Géodesique, Vol. 107, pp. 13–34, 1973.

[18] H. Hopfield, "Two-quartic Tropospheric Refractivity Profile for Correcting Satellite Data," Journal of Geophysical Research, Vol. 74, Oceans and Atmospheres, pp. 4487–4499, 1969.

[19] H. Hopfield, "Tropospheric effect on electro-magnetically measured range prediction from surface weather data." Radio Science, Vol. 6, Issue 3, pp. 357–367, 1971.

[20] R. Leandro, M. Santos and R. Langley, "UNB neutral atmosphere models: Development and performance," In Proceedings of the Institute of Navigation, National Technical Meeting, Monterrey, CA, USA, January 2006.

[21] C. Chao, "A model for Tropospheric Calibration from Daily Surface and Radiosonde Balloon Measurements," Technical Memorandum, Jet Propulsion Laboratory (JPL), Pasadena, CA, USA, pp. 350–391, 1972.

[22] I. Ifadis, "The Atmospheric Delay of Radio Waves: Modeling the Elevation Dependence on a Global Scale." Technical Report No. 38L, School of Electrical and Computer Engineering, Chalmers University of Technology, Gothenburg (Sweden), 1986.

[23] T. Herring, "Modeling atmospheric delays in the analysis of space geodetic data." In Proceedings of Refraction of Transatmospheric Signals in Geodesy, De Munck, J., Spoelstra, T., Eds. Netherlands Geodetic Commission Publications on Geodesy: The Hague, Netherlands, Volume 36, pp. 157–164, 1992.

[24] A. Niell, "Global mapping functions for the atmosphere delay at radio wavelengths." Journal of Geophysical Research, Vol. 101, pp. 3227–3246, 1996.

[25] J. Boehm, B. Werl and H. Schuh, "Troposphere Mapping Functions for GPS and Very Long Baseline Interferometry from European Centre for Medium-Range Weather Forecasts Operational Analysis Data." Journal of Geophysical Research: Solid Earth, Vol. 111, Issue B2, 2006.

[26] G. Seeber, "Satellite Geodesy", Second Edition, Walter de Gruyter Editor. Berlin (Germany), New York (USA), 2003.

[27] J.S. Subirana, J.M. Zornoza and M. Hernández-Pajares, "Multipath." European Space Agency (ESA) Navipedia – GNSS Measurements and Data Processing, 2011. Available online at: http://www.navipedia.net/index.php/Multipath #cite_ref-1. Accessed on 20/04/2017.

[28] A. J. Van Dierendonck, P. Fenton and T. Ford, "Theory and Performance of Narrow Correlator Spacing in GPS Receiver," NAVIGATION: Journal of the Institute of Navigation, vol. 39, no. 3, pp. 265-284, 1992.

[29] G.A. McGraw and M.S. Braasch, "GNSS Multipath Mitigation Using Gated and High Resolution Correlator Concepts," In Proceedings of the 1999 National Technical Meeting of the Institute of Navigation, San Diego, CA (USA), 1999.

[30] P.C. Fenton and J. Jones, "The Theory and Performance of NovAtel Inc.'s Vision Correlator," In Proceedings of the 2005 ION GNSS Conference. Long Beach, CA (USA), 2005.

[31] UD DOD, "Global Positioning System Standard Positioning Service Performance Standard," Fourth Edition, Department of Defence, Positioning, Navigation, and Timing Executive Committee, Document GPS SPS PS, 2008.

[32] A.H. Phillips, "The Determination of PDOP in GPS," NATO Advisory Group for Aerospace Research and Development (AGARD), Document AGARD-AG-314, 1990.

[33] ICAO, "Report of the Special Communications/Operations Divisional Meeting," ICAO Doc. 9650, 1995.

[34] ICAO, "Performance Based Operations Manual," Document No. 9613 AN/937, Third Edition, 2008.

[35] US DoT/DoD, "US Federal Radionavigation Plan," United States Department of Defence, Department of Transport, Doc. DoT-VNTSC-RITA-08-02/DoD-4650.5, 2008.

[36] RTCA, "Minimum Operational Performance Standards for GPS/GBAS Airborne Equipment," Radio Technical Commission for Aeronautics (RTCA) Inc., Special Committee No. 159, Document RTCA DO-253C. Washington DC (USA), 2008.

[37] CAA, "GPS Integrity and Potential Impact on Aviation Safety," CAA Safety Regulation Group Paper 2003/09, 2003.

[38] ICAO, "Standards and Recommended Practices (SAPRs) Volume 1 – Annex 10 (Including Amendments 1-81)," Issue 6, 2006.

[39] ICAO, "GBAS CAT II/III Development Baseline SARPs," Draft proposed changes to Annex 10, Volume I, as agreed at the 17 – 28 May 2010 meeting of the ICAO Navigation Systems Panel (NSP), 2010. Available online at: http://www.icao.int/safety/airnavigation/documents/gnss_cat_ii_iii.pdf

[40] RTCA, "Minimum Aviation System Performance Standards (MASPS) for the Local Area Augmentation System (LAAS)," Radio Technical Commission for Aeronautics (RTCA) Inc., Special Committee No. 159. Document RTCA DO-245A, Washington DC (USA), 2004.

[41] RTCA, "Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation Systems Airborne Equipment," Radio Technical Commission for Aeronautics (RTCA) Inc., Special Committee No. 159, Document RTCA DO-229D. Washington DC (USA), 2006.

[42] B. W. Parkinson and J. J. Spilker Jr., Editors, "Global Positiong System: Theory and Applications – Volume I," Progress in Astronautics and Aeronautics, vol. 163, Published by the American Institute of Aeronautics and Astronautics, 1996.

[43] P. M. Kintner, B. M. Ledvina and E. R. de Paula, "GPS and Ionospheric Scintillations," Space Weather, vol. 5, issue S09003, pp. 1-23, 2007.

[44] R. Sabatini, "Experimental Investigation of Differential GPS for Flight Test of High Performance Aircraft," University of Rome – "La Sapienza", School of Aerospace Engineering, Thesis submitted in partial fulfilment for the degree of Doctor of Engineering in Atronautics, 2005.

[45] R. Sabatini and G.B. Palmerini, "Differential Global Positioning System (DGPS) for Flight Testing," NATO Research and Technology Organization (RTO) Systems Concepts and Integration Panel (SCI) AGARDograph Series RTO-AG-160 Vol. 21, 2008.

[46] L. Boithias, "Radio Wave Propagation," McGraw Hill (New York), 1982.

[47] ITU, "Ionospheric Propagation Data and Prediction Methods Required for the Design of Satellite Services and Systems", International Telecommunication Union (ITU), Recommendation ITU-R P.531-12 (P Series – Radiowave propagation), 2013.

[48] W. B. Davenport and W. L. Root, "An Introduction to the Theory of Random Signals and Noise", Wiley-IEEE Press, 1987. ISBN: 978-0-87942-235-6

[49] P. W. Ward, "GPS Receiver RF Interference Monitoring, Mitigation, and Analysis Techniques," Navigation, Journal of the Institute of Navigation (USA), vol. 41, pp. 367-391, 1994.

[50] A. Steingass, "The High Resolution Aeronautical Multipath Navigation Channel", German Aerospace Center DLR (Germany), 2004. http://www.kn-s.dlr.de/satnav. Accessed 14 May 2012.

[51] A. Steingass and A. Lehner, "Aeronautical Channel Model", German Aerospace Center DLR (Germany), 2004. http://www.kn-s.dlr.de/satnav. Accessed 14 May 2012.

[52] M. S. Braasch, "On the characterization of multipath errors in satellite-based precision approach and landing systems," PhD Thesis. College of Engineering and Technology, Ohio University (USA), 1992. http://etd.ohiolink.edu/view.cgi/Braasch%20Michael.pdf?ohiou1173748635. Accessed 10 May 2012.

[53] R. Sabatini, T. Moore and C. Hill, "A New Avionics Based GNSS Integrity Augmentation System: Part 1 – Fundamentals," Journal of Navigation, Vol. 66, No. 3, pp. 363-383, 2013.

[54] R. Sabatini, T. Moore and C. Hill, "A New Avionics Based GNSS Integrity Augmentation System: Part 2 – Integrity Flags," Journal of Navigation, Vol. 66, No. 4, pp. 511-522, 2013.

[55] T. Murphy, R. Friedman, J. Booth, P. Geren, N. Molloy, B. Clark and J. Burns, "Program for the Investigation of Airborne Multipath Errors", Proceeding of the ION National Technical Meeting (NTM) 2004, San Diego, CA (USA), 2004.

[56] J. Booth, T. Murphy, F. Liu, "Validation of the Airframe Multipath Error Allocation for Local Area Differential GPS", Proceedings of the IAIN World Congress in association with the ION 2000 Annual Meeting, San Diego, CA (USA) 2000.

[57] O.M. Mubarak and A.G. Dempster, "Statistical Analysis of Early Late Phase for Multipath Detection", IGNSS Symposium 2009. Gold Coast, Australia, 2009.

[58] P.W. Ward, "Using a GPS Receiver Monte Carlo Simulator to Predict RF Interference Performance", Proceedings of the 10th International Technical Meeting of the Satellite Division of the Institute of Navigation, Kansas City, MO (USA), 1997.

[59] RTCM, "RTCM Recommended Standards for Differential NAVSTAR GPS Service," Radio Technical Committee for Maritime Services, Special Committee No. 104, Paper 134-89/SC104-68,Washington DC (USA), 1990.

[60] RTCM, "RTCM Recommended Standards for Differential NAVSTAR GPS Service," Radio Technical Committee for Maritime Services, Special Committee No. 104, Paper 194-93/SC104-STD, Washington DC (USA), 1990.

[61] RTCM, "RTCM Recommended Standards for Differential GNSS Service." Radio Technical Committee for Maritime Services, Special Committee No. 104, Paper 136-2001/SC104-STD, Version 2.3, Washington DC (USA), 2011.

[62] RTCA, "Minimum Aviation System Performance Standards DGNSS Instrument Approach System: Special Category 1 (scat-1)," Radio Technical Commission for Aeronautics (RTCA) Inc., Special Committee No. 159. Document RTCA DO-217. Washington DC (USA), 1993.

[63] US DoD/FAA, "Global Positioning System Wide Area Augmentation Systems (WAAS) Performance Standard", First Edition, United States Department of Defence and Federal Aviation Administration, Washington DC (USA), 2008.

[64] J. Rife, S. Pullen, B. Pervan and P. Enge, "Core Overbounding and its Implications for LAAS Integrity," In Proceedings of ION GNSS (pp. 2810-2821, 2004.

[65] RTCA, "GNSS-Based Precision Approach Local Area Augmentation System (LAAS) Signal-in-Space Interface Control Document (ICD)," Radio Technical Commission for Aeronautics (RTCA) Inc., Special Committee No. 159, Document RTCA DO-246D, Washington DC (USA), 2008.

[66] R.R. Hatch, "Ambiguity Resolution While Moving, Experimental Results," In Proceedings of ION GPS-91, the 4th International Technical Meeting of the Satellite Division of the US Institute of Navigation. Albuquerque (NM), 1991.

[67] H. J. Euler and H. Landau, "Fast GPS Ambiguity Resolution On-The-Fly for Real-time Applications," Paper presented at the 6th International Geodetic Symposium on Satellite Positioning, Columbus, Ohio (USA), 1992.

[68] P. Hansen, "Real-time GPS Carrier Phase Navigation," The University of Nottingham, IESSG, Paper presented at the DSNS-94 Conference, London (UK), 1994.

[69] H. J. Euler, "Achieving High-accuracy Relative Positioning in Real-time: System Design, Performance and Real-time Results," In Proceedings of the 4th IEEE Plans Conference, Las Vegas, NV (USA), 1994.

[70] D. Walsh, "Kinematic GPS Ambiguity Resolution," PhD Thesis, IESSG, University of Nottingham (UK), 1994.

[71] C. C. Tiberius and P. J. De Jonge, "Fast Positioning using the LAMBDA Method," Proceedings of DSNS'95, Bergen, Norway, 1995.

[72] P. J. Teunissen, "Integer Aperture GNSS Ambiguity Resolution," Artificial Satellites, vol. 38, issue 3, pp. 79-88, 2003.

[73] P. J. Teunissen, "Towards a Unified Theory of GNSS Ambiguity Resolution," Journal of Global Positioning Systems, vol. 2, no. 1, pp. 1-12, 2003.

[74] H. Shaowei, "Ambiguity Recovery for Long-Range GPS Kinematic Positioning," Navigation, vol. 44, no. 2, pp. 257-266, 1997.

[75] P. J. Teunissen, "GNSS Ambiguity Resolution with Optimally Controlled Failure-rate," Artificial Satellites, vol. 40, issue 4, pp. 219-227, 2005.

[76] S. Verhagen, "Improved performance of Multi-Carrier Ambiguity Resolution based on the LAMBDA method," In Proceedings of Navitec 2006, ESA-ESTEC, Noordwijk (The Netherlands), 2006.

[77] H. Liu, Z. Chen, W. Ye and H. Wang, "GNSS Carrier Phase Ambiguity Resolution based on Integrity Restriction in Ambiguity Domain." Advances in Space Research, vol. 53, issue 8, pp. 1207–1218, 2014.

[78] S. Nassar, A. Noureldin and N. El-Sheimy, "Improving Positioning Accuracy during Kinematic DGPS Outage Periods using SINS/DGPS Integration and SINS Data De-noising," Survey Review, vol. 37, issue 292, pp.426-438, 2004.

[79] A. Keith, "Using Wide Area Differential GPS to Improve Total System Error for Precision Flight Operations," PhD Thesis, Stanford University (USA), 2000.

[80] G. W. A. Offermans, A. W. S. Helwig and D. Van Willigen, "Eurofix: Test Results of a Cost-Effective DGNSS Augmentation System," Journal of Navigation, vol. 50, no. 2, pp. 209-223, 1997.

[81] T. Moore, "Other Satellite Navigation Systems." Lecture Notes, IESSG, University of Nottingham (UK), 1996.

[82] T. Moore, "An Introduction to Differential GPS," Lecture Notes, IESSG, University of Nottingham (UK), 1996.

[83] T. Moore, "GPS Orbit Determination and Fiducial Networks," Lecture Notes, IESSG, University of Nottingham (UK), 1996.

[84] K. Zhang, F. Wu, S. Wu, C. Rizos, C. Roberts, L. Ge, T. Yan, C. Gordini, A. Kealy, M. Hale, P. Ramm, H. Asmussen, D. Kinlyside and P. Harcombe, "Sparse or dense: Challenges of Australian Network RTK", Proceedings of International Global Navigation Satellite Systems Society Symposium (IGNSS2006), Surfers Paradise (Australia) July 2006.

[85] V. Janssen and J. Haasdyk, "Assessment of Network RTK performance using CORSnet-NSW," Proceeedings of the IGNSS 2011 Symposium, 15-17 November 2011, Sydney (Australia) 2011

[86] J.M. Juan, A. Rius, M. Hernández-Pajares and J. Sanz, "A Two-layer Model of the Ionosphere using Global Positioning System Data", Geophys. Research Letters, Vol. 24, pp. 393–396, 1997.

[87] S. Bisnath and Y. Gao, "Precise Point Positioning – A Powerful Techniqe with a Promising Future," GPS World (Algorithms and Methods – Innovation), pp. 43-50, 2009.

[88] A.S. Jokinen, "Enhanced Ambiguity Resolution and Integrity Monitoring Methods for Precise Point Positioning." PhD Thesis, Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College, London (UK), 2014.

[89] National Coordination Office for Space-Based Positioning, Navigation, and Timing GPS, Washington DC (USA). URL: http://www.gps.gov/. Accessed 19 July 2017.

[90] European Global Navigation Satellite System Agency, Prague (Czech Republic) and Saint-Germain-en-Laye (France). URL: https://www.gsa.europa.eu/european-gnss/galileo/galileo-european-global-satellite-based-navigation-system. Accessed 19 July 2017.

[91] Information and Analysis Center for Positioning, Navigation and Timing, Korolyov (Russia). URL: https://www.glonass-iac.ru/en/GLONASS/index.php. Accessed 19 July 2017.

[92] China National Space Administration, BeiDou Navigation Satellite System, Haidian District, Beijing (China). URL: http://en.beidou.gov.cn/. Accessed 19 July 2017.

[93] Indian Space Research Organisation, Bengaluru (India). URL: http://www.isro.gov.in/spacecraft/satellite-navigation. Accessed 19 July 2017.

[94] G.W. Hein, J.A. Rodriguez, S. Wallner et al., "Envisioning a Future GNSS System of Systems", Part 1Inside GNSS, vol. 2, issue 1, pp. 58-67, 2007.

[95] Y. Urlichich, V. Subbotin, G. Stupak et al., "GLONASS Modernization", Proceedings of ION GNSS+ 2011, pp. 3125–3128, Portland, Oregon (US), 2011.

[96] Cabinet Office, Office of National Space Policy, Government of Japan, Tokyo (Japan).URL: http://qzss.go.jp/en/. Accessed 19 July 2017.

[97] O. Montenbruck, P. Steigenberger, L. Prange, Z. Deng, Q. Zhao, F. Perosanz, et al., "The Multi-GNSS Experiment (MGEX) of the International GNSS Service (IGS) – Achievements, Prospects and Challenges", Advances in Space Research, vol. 59, pp. 1671-1697, 2017.

[98] International Committee on Global Navigation Satellite Systems: The Way Forward: International Committee on Global Navigation Satellite Systems, ST/SPACE/67, United Nations Office for Outer Space Affairs, Vienna (Austria), 2016.

[99] G.M. Siouris, "Aerospace Avionics Systems," Academic Press, San Diego, California (USA), 1993.

[100] G.W. Hein and M.M. Ertel, "High-precision Aircraft Navigation using DGPS/INS integration". Institute of Astronomical and Physical Geodesy (IAPG). University FAF Munich, Neubiberg (Germany), 1993.

[101] F. Cappello, S. Ramasamy and R. Sabatini, "A Low-Cost and High Performance Navigation System for Small RPAS Applications," Aerospace Science and Technology, vol. 58, pp. 529–545, 2016. DOI: 10.1016/j.ast.2016.09.002

[102] R. Kapoor, S. Ramasamy, A. Gardi and R. Sabatini, "UAV Navigation Using Signals of Opportunity in Urban Environments: An Overview of Existing Methods", Energy Procedia, vol. 110, pp. 377-383, 2017. DOI: 10.1016/j.egypro.2017.03.156

[103] A. Gelb, "Applied Optimal Estimation," The MIT Press. Cambridge (Massachusetts) and London (England), 1992.

[104] E. Waltz and J. Linas, "Multisensor Data Fusion," Artech House. London (England) and New York (US), Second Edition, pp. 159-211, 1994.

[105] R. Van de Leijgraaf, J. Breeman, G. Moek and S.S. Van Leeuwen, S.S. "A Position Reference System for Fokker 70," NLR Technical Publication TP 93084L, 1993.

[106] M. Napier, "The Integration of Satellite and Inertial Positioning Systems," Proceedings of the NAV-89 Symposium, The Royal Institute of Navigation, London (UK), 1989.

[107] T. Jacob and G. Schanzer, "Integrated Flight Guidance System Using Differential GPS for Landing Approach Guidance," AGARD-CP-455. NATO Advisory Group for Aerospace Research and Development. Neuilly-sur-Seine (France), 1989.

[108] A. Noureldin, A. El-Shafie and M. Bayoumi, "GPS/INS Integration utilizing Dynamic Neural Networks for Vehicular Navigation," Information Fusion, vol. 12, pp. 48-57, 2011.

[109] ESA (2011a), "Integrity." European Space Agency Navipedia website: http://www.navipedia.net/index.php/Integrity. Year of Publication: 2011 (Edited by GMV). Visited on 18th January 2017.

[110] US DoD/DoT/DoHS, "Federal Radionavigation Plan – 2012," United States Department of Defence, Department of Transport and Department of Homeland Security - National Technical Information Services. Springfield, VA (USA), 22181, Doc. DOT-VNTSC-RITA-08-02/DoD-4650.05, 2012.

[111] G. N. Skillicorn, "The Past, Present, and Future of LAAS," Integrated CNS Technologies Conference and Workshop. Annapolis, MD (USA), 2003.

[112] ICAO, "Global Navigation Satellite Systems (GNSS) Manual," International Civil Aviation Organization. Document No. 9849 AN/457, First Edition, 2005.

[113] ESA, "SBAS Systems," European Space Agency Navipedia website: http://www.navipedia.net/index.php/SBAS_Systems, 2011, (Edited by GMV). Accessed on 20th March 2016.

[114] ICAO, "International Standards and Recommended Practices. Annex 10 to the Convention on International Civil Aviation," Aeronautical Telecommunications, Volume I: Radio Navigation Aids. Sixth Edition (Including Amendments 1-81), 2006.

[115] FAA, "Satellite Navigation - Ground Based Augmentation System (GBAS)," US Federal Aviation Administration Website, Section Dedicated to Navigation Programs/GNSS: http://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/gnss/laas/. Vested on 24th March 2016.

[116] ICAO, "Guide for Ground Based Augmentation Systems Implementation," International Civil Aviation Organisation, 2013.

[117] J. E. Angus, "RAIM with Multiple Faults," NAVIGATION: Journal of the Institute of Navigation, vol. 53, no. 4, 2006.

[118] T. Walter, P. Enge, J. Blanch and B. Pervan, "Worldwide Vertical Guidance of Aircraft Based on Modernized GPS and New Integrity Augmentations," Proceedings of the IEEE, Vol. 96, No. 12, 2008.

[119] A. Ene, J. Blanch and J. D. Powell, "Fault Detection and Elimination for GALILEO-GPS Vertical Guidance," Proceedings of the Institute of Navigation National Technical Meeting, San Diego, CA (USA), 2007.

[120] J. Blanch, A. Ene, T. Walter and P. Enge, "An Optimized Multiple Hypothesis RAIM Algorithm for Vertical Guidance," In Proceedings of ION GNSS 2007, 2007.

[121] H. Kuusniemi and T. Jokitalo, "Indoor and Week Signal Navigation," Chapter 12 in "GNSS Applications and Methods," S. Gleason and D. Gebre-Egziabher, Artech House, pp. 306-324, 2009.

[122] T. Murphy, R. Friedman, J. Booth, P. Geren, N. Molloy, B. Clark and J. Burns, "Program for the Investigation of Airborne Multipath Errors," Proceeding of the ION National Technical Meeting (NTM) 2004, San Diego, CA (USA), 2004.

[123] FAA, "GEAS - GNSS Evolutionary Architecture Study," GEAS Phase I – Panel Report, 2008.

[124] FAA, "GEAS - GNSS Evolutionary Architecture Study," GEAS Phase II – Panel Report, 2010.

[125] Y. Jiang and J. Wang, "A-RAIM and R-RAIM Performance using the Classic and MHSS Methods," The Journal of Navigation, vol. 67, pp. 49-61, 2014. DOI: 10.1017/S0373463313000507

[126] F. Van Graas and A. Soloviev, "Precise Velocity Estimation Using a Stand-Alone GPS Receiver," Navigation, vol. 51, issue 4, pp. 283–292, 2004.

[127] W. Ding and J. Wang, Precise velocity estimation with a stand-alone GPS receiver," Journal of Navigation, vol. 64, issue 2, pp. 311–325, 2011.

[128] K. L. Van Dyke, "RAIM Availability for Supplemental GPS Navigation," Navigation, Journal of The Institute of Navigation, vol. 39, no.4, pp. 429-443, 1992.

[129] Y. C. Lee, "RAIM Availability for GPS Augmented with Barometric Altimeter Aiding and Clock Coasting," Navigation, Journal of the Institute of Navigation, vol. 40, no.2, pp. 179-198, 1993.

[130] P. Misra, E. Bayliss, R. LaFrey, M. Pratt and J. F. Mclellan, "Receiver Autonomous Integrity Monitoring (RAIM) of GPS and GLONASS," Navigation, Journal of The Institute of Navigation, vol. 40, no.1, pp. 87-104, 1993.

[131] J. Sang, K. Kubik and Y. Feng, "Receiver Autonomous Integrity Monitoring (RAIM) in Civil Aviation Use of GPS as a Sole Means of Navigation," In Proceedings of Satellite Navigation Technology 1995 and Beyond, Brisbane, June 1995.

[132] T. Walter and P. Enge, "Weighted RAIM for Precision Approach," In Proceedings of the 8th International Technical Meeting of the Satellite Division of the Institute of Navigation, Palm Springs, CA (USA), 1995.

[133] R. G. Brown, "A Baseline GPS RAIM Scheme and a Note on the Equivalence of Three RAIM Methods," Journal of the Institute of Navigation, vol. 39, no. 3, 1992.

[134] T. Huiqi, H. Li, W. Zhang and M. Lu, "A Recursive Receiver Autonomous Integrity Monitoring (Recursive-RAIM) Technique for GNSS Anti-Spoofing," In Proceedings of the 2015 International Technical Meeting of The Institute of Navigation, California (USA), pp. 738 – 744, January 2015.

[135] EU-US Cooperation on Satellite Navigation, Working Group C-ARAIM Technical Subgroup, ARAIM Technical Subgroup Milestone 1 Report, 2012. http://ec.europa.eu/enterprise/newsroom/cf/_getdocument. cfm?doc_id=7793

[136] Y. C. Lee, "Analysis of Range and Position Comparison Methods as a Means to Provide GPS Integrity in the User Receiver," In Proceedings of the Annual Meeting of the Institute of Navigation, pp. 1-4, Seattle, WA (USA), June 1986.

[137] B. W. Parkinson and P. Axelrad, "Autonomous GPS Integrity Monitoring using the Pseudorange Residual," Navigation (Washington), vol. 35, pp. 255-274, 1988. http://dx.doi.org/10.1002/j.2161-4296.1988.tb00955.x

[138] M. A. Sturza and A. K. Brown, "Comparison of Fixed and Variable Threshold RAIM Algorithms," Proceedings of the 3rd International Technical Meeting of the Institute of Navigation, Satellite Division, ION GPS-90 (Colorado Springs, CO), pp. 437-443, September 1990.

[139] R. G. Brown and P. Y. C. Hwang, "GPS Failure Detection by Autonomous Means within the Cockpit," In Proceedings of the Annual Meeting of the Institute of Navigation, Seattle, pp. 5-12, June 1986. http://dx.doi.org/10.1002/j.2161-4296.1986.tb01485.x

[140] R. G. Brown, "Receiver Autonomous Integrity Monitoring," In B. W. Parkinson and J. J. Spilker (Eds.), Chapter 5 in Global Positioning System: Theory and Applications, Volume 2, AIAA Inc., Washington DC (USA), pp. 143-165, 1996.

[141] M. Joerger, F. C. Chan and B. Pervan, "Solution Separation versus Residual-Based RAIM," NAVIGATION, vol. 61, issue 4, pp. 273–291, 2014.

[142] J. Blanch, A. Ene, T. Walter and P. Enge, "An Optimized Multiple Hypothesis RAIM Algorithm for Vertical Guidance," ION GNSS 2007, Fort Worth, TX, 2007

[143] R. Sabatini, T. Moore, C. Hill, A. Gardi, S. Ramasamy and M. Gilgien, "Trajectory Optimisation for Avionics-Based GNSS Integrity Augmentation Systems", Proceedings of the 35th AIAA/IEEE Digital Avionics Systems Conference (DASC2016), Sacramento, CA (USA), September 2016.

[144] R. Sabatini, T. Moore and C. Hill, "Avionics-Based GNSS Integrity Augmentation Synergies with SBAS and GBAS for Unmanned Aircraft Applications", Proceedings of the 35th AIAA/IEEE Digital Avionics Systems Conference (DASC2016), Sacramento, CA (USA), September 2016.

[145] L. Rodriguez, R. Sabatini, A. Gardi and S. Ramasamy, "A Novel System for Non-Cooperative UAV Sense-and-Avoid," In Proceedings of European Navigation Conference 2013, Vienna (Austria), 2013.

[146] S. Ramasamy, R. Sabatini and A. Gardi, "LIDAR Obstacle Warning and Avoidance System for Unmanned Aerial Vehicle Sense-and-Avoid", Aerospace Science and Technology, Elsevier, vol. 55, pp. 344–358, 2016. DOI: 10.1016/j.ast.2016.05.020

[147] S. Ramasamy, R. Sabatini and A. Gardi, "A Unified Approach to Separation Assurance and Collision Avoidance for Flight Management Systems," Proceedings of the 35th AIAA/IEEE Digital Avionics Systems Conference (DASC2016), Sacramento, CA (USA), September 2016.

[148] R. Sabatini, T. Moore and C. Hill, "Avionics Based GNSS Integrity Augmentation for Mission- and Safety-Critical Applications," In proceedings of the 25th International Technical Meeting of the Satellite Division of the Institute of Navigation: ION GNSS-2012, Nashville, Tennessee (USA), September 2012. URL: http://www.ion.org/publications/abstract.cfm?articleID=10288

[149] R. Sabatini, T. Moore, C. Hill and S. Ramasamy, "Investigation of GNSS Integrity Augmentation Synergies with Unmanned Aircraft Systems Sense-and-Avoid," In Proceedings of SAE AeroTech Congress 2015, Technical Paper 2015-01-2456, Seattle, WA (USA), September 2015. DOI: 10.4271/2015-01-2456

[150] R. Sabatini, T. Moore and C. Hill, "Avionics-Based GNSS Integrity Augmentation for Unmanned Aerial Systems Sense-and-Avoid," In Proceedings of the 27th International Technical Meeting of the Satellite Division of the Institute of Navigation: ION GNSS+ 2014, Tampa, Florida (USA), September 2014. URL: http://m.ion.org/abstract.cfm?paperID=1811

[151] J. Owen, "A Review of the Interference Resistance of SPS GPS Receivers for Aviation", Journal of Navigation, PB - Blackwell Publishing Ltd. DOI: 10.1002/j.2161-4296.1993.tb02307.x, 1993.

[152] J.R. Rufa and E.M. Atkins, "Unmanned Aircraft System Navigation in the Urban Environment: A Systems Analysis", Journal of Aerospace Information Systems, 2016.