A High Availability Inertial-Vision Data Fusion Using an ES-KF for a Civil Aircraft During a Precision Approach in a GNSS-Challenged Environment

Gabriel Thys, Christophe Macabiau, Julien Lesouple, Jérémy Vézinet, Anaïs Martineau, Raphaël Jarraud

A High Availability Inertial-Vision Data Fusion Using an ES-KF for a Civil Aircraft During a Precision Approach in a GNSS-Challenged Environment

Gabriel Thys, *Safran Electronics & Defense, and Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse*
Christophe Macabiau, Julien Lesouple, Jérémy Vézinet, Anaïs Martineau, *Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse*
Raphaël Jarraud, *Safran Electronics & Defense*

BIOGRAPHY

Gabriel Thys is a doctoral student at Safran Electronics & Defense, affiliated with the Fédération ENAC, ISAE-SUPAERO, and Université de Toulouse, specializing in the integrity of inertial-vision hybridization. He graduated with a M.Eng degrees in space telecommunications in 2022 from the ENAC in Toulouse, France.

Christophe Macabiau graduated as an electronics engineer in 1992 from the ENAC in Toulouse, France. Since 1994, he has been working on the application of satellite navigation techniques to civil aviation. He received his Ph.D in 1997 and has been in charge of the signal processing lab of ENAC since 2000. He is currently the head of the TELECOM research team of ENAC, that includes research groups on signal processing and navigation, electromagnetics, and data communication networks.

Raphaël Jarraud is a senior expert in inertial navigation and sensors fusions, working for Safran Electronics & Defense in Eragny sur Oise, France. He has a 21 years experience in designing, simulating and testing inertial navigation systems. He graduated from CentraleSupelec in 2003, with a major in control systems.

Julien Lesouple received the Eng. degree in Aeronautics Engineering from ISAE Ensica, Toulouse, France in 2014 and his Ph.D in Signal Processing from Toulouse Institut National Polytechnique in 2019. Since 2021 he has been an Associate Professor at ENAC within the SIGNAV team. His research interests include statistical signal processing, machine learning, estimation and detection theory, filtering, with applications to satellite communications, localization, tracking, navigation and anomaly detection. Since 2024, he's been a member of the EURASIP Academy mentoring program.

Jérémy Vézinet graduated as an electronics engineer in 2010 and obtained his PhD in 2014 on multi-sensor hybridization from the ENAC in Toulouse, France. He is currently working as a Research Associate in the TELECOM Research Team at ENAC since 2014. His research interests are GNSS, INS, video-based navigation, multi-sensor hybridization and integrity monitoring.

Anaïs Martineau graduated in 2005 as an electronics engineer from the ENAC in Toulouse, France. Since 2005, she has been working at the signal processing lab of the ENAC where she carries out research on integrity monitoring techniques. She received her Ph.D. in 2008 from the University of Toulouse.

ABSTRACT

This paper introduces the architecture of a localization system designed for commercial airliners. The system is designed to maintain guidance during an PBN CAT I approach in a GNSS-challenged environment. It leverages a vision sensor to observe the aircraft's surroundings. Specifically, this new sensor allows for the detection and tracking of the runway during the final segment of the approach, providing line-of-sight information to the system. The vision sensor is integrated into a multi-sensor system that includes a high-precision navigation-grade Inertial Measurement Unit (IMU) (angular drift of the order of 0.01 degree per hour), a barometric altimeter, and a GNSS receiver. Data fusion is performed using an Error-State Kalman Filter (ES-KF) within a semi-closed loop framework. The ES-KF is typically used for IMU hybridization systems, as its allows for quick testing and modification of propagation and observation models to address challenges like false observability. The system's design is made in order to address the challenges posed by the limited number of landmarks and the high correlations on the measurement errors elaborated from them, where two points on the runway may be indiscernible from the perspective of the camera, and where distance estimation may be challenging. The system's architecture, the types of sensors used, and the choice of a tightly-coupled single-feature hybridization improve the system's operational period and, ultimately, its availability. The regulatory framework for integrating a vision sensor into the localization system of a civil aircraft is discussed. The performance of this vision-integrated localization system is evaluated using a comprehensive Monte Carlo simulator.

 https://doi.org/10.33012/2025.19981

I. INTRODUCTION

The approach phase represents one of the most challenging and critical segments for commercial aircraft operations. Accurate and reliable localization during this phase is vital to ensure safety and efficiency. Consequently, there is a pressing need to improve localization systems specifically for this segment. In scenarios where GNSS availability can no longer be guaranteed, maintaining autonomy becomes a significant challenge. To address this, it is essential to augment the aircraft localization system with a passive sensor that is resilient to radio frequency interference (RFI).

Modern commercial aircraft are already equipped with cameras, primarily intended to enhance pilots' situational awareness during approach, landing, and taxiing phases. Despite their presence, these cameras are predominantly used for head-up displays and provide operational credits only under highly specific and rare circumstances. Their potential remains largely untapped. Cameras offer a unique advantage in that they can serve as complementary sensors for navigation systems, particularly during the approach phase when the aircraft is near the ground. In this environment, cameras can capture a wealth of visual data, providing valuable information to improve the precision and reliability of localization. At term, leveraging this capability aims to increase the autonomy, robustness, and reliability of localization systems. Such improvements are particularly crucial as the aviation industry moves toward the implementation of single pilot operations (SPO), where enhanced localization will play a critical role in ensuring operational safety and efficiency.

Camera-based runway-relative positioning and navigation has been an active area of research for several years. Among these efforts, the VISION project (VIS, 2016) played a pivotal role in driving innovation in this domain. This European-Japanese collaborative research initiative, conducted between 2016 and 2019, explored the potential of using camera systems to enhance aircraft landing operations. A bibliographic study conducted in Gróf et al. (2022) by members of the VISION project highlights the breadth of techniques employed in this field. Two main categories of methods can be identified. The first is visual servoing, which focuses on guiding the aircraft without explicitly estimating its position. This technique is particularly applicable during the landing phase. The second category is position estimation, which is more commonly used during the approach phase. The design of these methods involves critical decisions regarding the quantities to estimate, the types of sensors to use (e.g., IMUs, monocular cameras, stereoscopic cameras, infrared systems), and the needed prior knowledge of the approach environment. A notable contribution from Watanabe et al. (2019) is the development of an ES-KF formulation that incorporates delay handling to account for the image processing latency of vision sensors. However, many of the hybridization methods reviewed in the state of the art rely on low-grade IMUs for inertial navigation. This choice introduces architectural constraints for the overall system. In commercial airliners, higher-grade IMUs, specifically navigation-grade IMUs, are typically installed, offering significantly greater accuracy. In the Gibert et al. (2018) study, such a high-grade IMU is considered. However, it is important to note that the proposed system was developed for landing operations, not for the approach phase. The regulatory frameworks and operational requirements differ significantly between these two phases of a flight.

Simulation is commonly used to assess the performance of vision-integrated navigation systems in the field of aviation. It offers a quick way to assess performance potential and enables the simulation of numerous flight hours to test the system's integrity. Real-life testing, on the other hand, is much more complex. As mentioned in (Brown et al., 2024), datasets in the field of computer vision for aviation are scarce. However, for the C2Land project (Kügler et al., 2019; Angermann et al., 2015; Tonhäuser et al., 2015), led by the Institute of Flight Guidance at Technische Universität Braunschweig, test campaigns have already been conducted on a small non-commercial aircraft. This project takes an innovative approach by utilizing airborne computer vision applications to bridge the gap between current decision height minima and the touchdown point through vision-aided automatic landing systems. The C2Land project is one of the most advanced vision-based navigation system in terms of technological maturity. However, the use case it addresses differs from the system described in this article. Specifically, C2Land targets small airports with no infrastructure and concentrates on the segment of the approach just before landing.

In this paper, a navigation-grade IMU is considered, which significantly alters the architectural constraints of the system. As a result, the role of vision differs from what is typically found in the literature. Here, the vision component is solely used for position correction. Section II addresses the potential regulatory framework associated with such a system. The goal is to demonstrate the feasibility of implementing this system within the context of civil aviation. Section III presents the system architecture, with a particular focus on the detailed mathematical formulations underlying the approach. Finally, Section IV introduces the simulator developed to evaluate the proposed system. This last section highlights the system's potential accuracy performance, as demonstrated through simulation results.

II. POTENTIAL REGULATORY FRAMEWORK

The hybrid system presented in this paper is a localization system used during the approach phases preceding an instrument landing of a commercial aircraft. An instrument approach consists of various segments (Figure 1). The proposed hybrid navigation system aims at providing expected performance objectives on the final segment of an approach, specifically from the Final Approach Point (FAP) to the Decision Altitude/Height (DA/H). At this point, which varies depending on the approach category, pilots must visually distinguish certain features to proceed with landing or execute a missed approach if visual

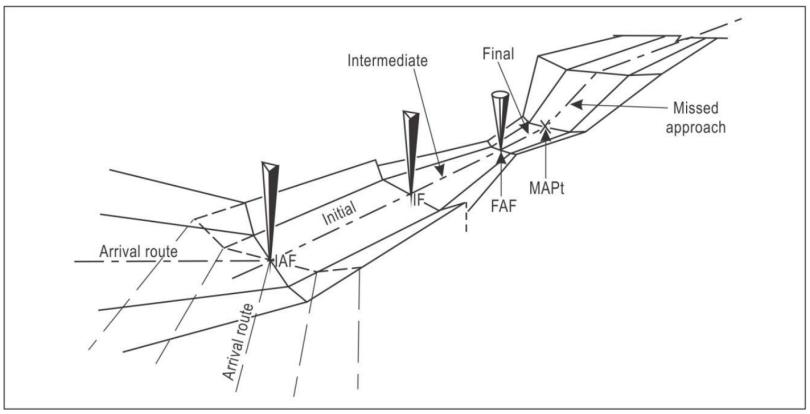


Figure 1: Segments of instrument approach (ICAO, 2006)

conditions are inadequate (ICAO, 2006). As a result, approaches with lower DA/H are both more accurate and more demanding.

Historically, precision approaches have relied on Instrument Landing Systems (ILS), which provide vertical and horizontal deviation guidance along a standard 3-degree descent path. Since the late 1990s, Performance-Based Navigation (PBN) has sought to simplify and standardize approach procedures by eliminating the need for specific onboard instruments (apart from two GNSS multi-modal receivers) or ground-based beacons. Instead, it leverages the aircraft's existing systems to meet required performance standards. The primary goal of PBN is to avoid the need for additional systems by utilizing those already in place (ICAO, 2008). The PBN requirements include, among others, localization system performance in terms of integrity, accuracy, continuity, and availability. PBN considers GNSS as the primary localization system. This system is generally hybridized with a high-precision IMU. PBN approaches with monitoring and alerting systems fall into two RNP (Required Navigation Parameters) sub-categories: RNP APCH and RNP AR APCH (for RNP Authorization Required APCH). High-performance PBN approaches rely on GNSS augmentation systems such as Satellite-Based Augmentation Systems (SBAS). For example, SBAS CAT I approaches, classified within the precision approach category (ICAO, 2023), are designed to provide services equivalent to ILS CAT I approaches with a decision height (DH) of 200 feet. PBN CAT I operations (i.e., CAT I operations utilizing GNSS) are currently developing , while ILS is facing gradual decrease, maintaining only a minimum operational network (MON). As a result, many airports now rely exclusively on SBAS for CAT I operations, which in turn depend on the availability of augmented GNSS signals. In the event of a GNSS signal loss during a PBN approach, the pilot must initiate a missed approach procedure and execute a go-around. Subsequently, they must revert to the remaining conventional aids, such as ILS for precision approaches or VOR/DME for non-precision approaches (NPAs).

In the current context, the availability of GNSS can no longer be guaranteed, even for states with a secure RF environment. GNSS jamming and spoofing now constitute an unavoidable risk. The proposed hybrid system addresses this vulnerability by incorporating a vision sensor into the existing IMU-GNSS hybridization. This approach aims at maintaining instrument approach continuity in the event of partial or total GNSS loss during the final approach segment, specifically targeting PBN CAT I operations. That said, the integration of a vision-based system raises key regulatory questions: Can a vision-based localization system receive operational credits? In case of loss of the primary localization system in the sense of PBN (GNSS), is the pilot required to immediately execute a missed approach even if IMU and vision still provide guidance?

To begin with, cameras are already prevalent in commercial aircrafts. In particular, Enhanced Flight Visual Systems (EFVS) are used. These systems enhance situational awareness for pilots by displaying a synthetic image of the aircraft's environment on the Head-Up Display (HUD). Currently, EFVS are eligible for operational credit under Title 14 of the US Code of Federal Regulations (CFR). The EFVS may provide operational credit, such as CAT I operation with reduced RVR (Runway Visual Range) minima. These credits allow pilots to identify external references earlier than with unaided vision and to lower the DH. That said, the general regulatory framework of instrument approaches are constituted by approach procedures. These procedures are mostly guided by the PBN manual (ICAO, 2008), ICAO Annex 10 (ICAO, 2023), and the RTCA MOPS (e.g. DO-384 for tight IMU-GNSS systems, or DO-401 for DFMC SBAS systems). However, certain special situations require specific procedures. These are governed within the AIR OPS, which provide Special Pole Authorizations (SPA). These define special operations that grant operational credits. In particular, AIR OPS EU 965/2021 deals with Low Visibility Operations (LVO). These regulations allow credits for vision-based systems under defined conditions, provided the system is certified for the intended operation (*SPA.LVO100-105*). The PBN Manual further supports the use of multi-sensor hybrid systems for RNP APCH operations. It states that these systems can perform RNP APCH if they are "in accordance with AC20-130A or TSO-C115b, as well as having been demonstrated for RNP APCH capabilities". Section 5.3.4.4 notes that a "Discontinuity in the procedure may not be

necessary for a multi-sensor RNP system that includes demonstrated RNP capabilities without GNSS". This provision implies that if a hybrid system incorporating vision meets the performance requirements, pilots may not be required to execute a missed approach in the event of GNSS failure.

In conclusion, an analysis of the PBN Manual and AIR OPS regulations suggests that a vision-enhanced hybrid system is viable for civil aviation without significant regulatory barriers. The primary condition is that the system must meet performance requirements necessary for approach continuation. Future work will define the specific performance requirements for this system. A logical starting point is the ICAO GNSS Signal-In-Space (SIS) specifications, commonly applied to GNSS. Additionally, the validation scenarios outlined in DO-384, which govern hybrid inertial-GNSS systems, will be instrumental in assessing the system's performance integrity.

III. INERTIAL-VISION HYBRID SYSTEM DESIGN

The tentative objective of the hybrid system is to meet the minimum requirements of an PBN CAT I approach in terms of accuracy, integrity, availability, and continuity. In this paper, only the initial potential of accuracy performance is presented. The system relies on data fusion from multiple sensors. The idea is to add a vision sensor, a sensor currently undergoing significant advancements in aircraft navigation systems (Watanabe et al., 2019; Gróf et al., 2022), to the state-of-the-art IMU-GNSS-baro hybridization in civil aviation. The system aims at meeting the minimum requirements even under the threat of loss of continuity of GNSS signal service. In other words, the addition of the vision sensor is intended to fill gaps in GNSS continuity, allowing the pilot to continue the approach despite GNSS disruption, thereby avoiding the need to execute a missed approach procedure.
The technical solution for fusing multi-sensor data is an Error-State Kalman Filter (Sola, 2017). In this framework, inertial navigation propagates a state, known as the nominal state, based on IMU measurements. This inertial navigation does not account for uncertainties, whether from model errors or various types of noise. The nominal state, deterministically computed by inertial navigation, is therefore tainted with errors. These errors are then estimated by the ES-KF. To this end, a new state, called the error state, is developed. As the KF models are based on the error state, it operates in an error space where the linearization closer to the origin reduces the resulting errors. The temporal dynamics of this error state establish the ES-KF state transition model. The observation models of the hybridization sensors (GNSS, barometric, and vision) aims at establishing a relationship between measurement innovation and the error state. An iterative correction step is then applied to "correct" the nominal state with the error state as estimated. This correction step is crucial, as various implementations are possible (see Section III.5). Figure 2 illustrates the high-level schematic of the hybrid system of the hybrid system proposed in this paper.

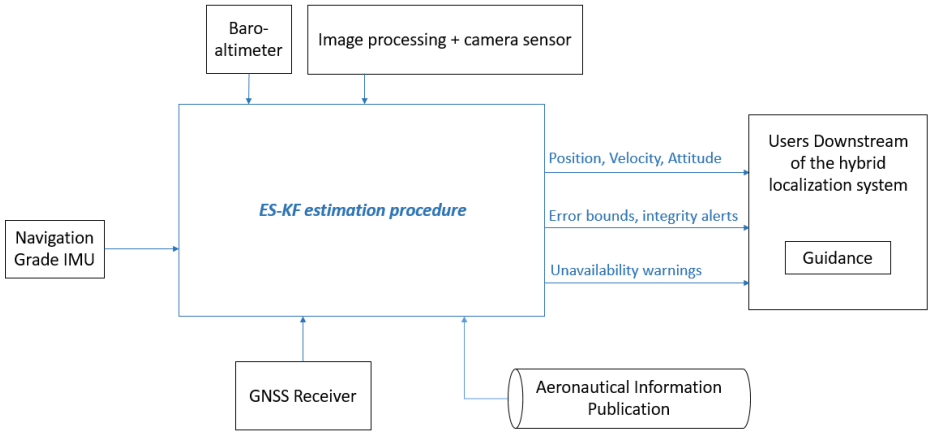


Figure 2: High-level schematic of the hybrid system

1. Reference frames and paper's notations

The introduction of inertial navigation and Kalman Filtering necessitates the use of various reference frames.

(X_i, Y_i, Z_i) are the axes of the inertial frame, denoted **[i]**. It is a fixed Galilean reference frame relative to the stars in which the laws of Newtonian mechanics apply.
(X_t, Y_t, Z_t) are the axes of the terrestrial reference frame, denoted **[t]**. The frame is generally called ECEF (for Earth Centered Earth Fixed). Its origin is the center of mass of the Earth.
(X_g, Y_g, Z_g) are the axes of the local geographic frame, denoted **[g]**. Also known as the Geographic frame (Farrell, 2008) or

Local Navigation Frame (Groves, 2015), it is determined by fitting a tangent plane to the geodetic reference ellipse at a point of interest. In this paper, the origin is the position of the initial core, its X-axis is oriented to the North, its Y-axis to the West and its Z-axis to the Zenith.
(X_m, Y_m, Z_m) are the axes of the measurement frame, denoted **[m]**. It is the reference frame of the inertial sensors core. The axis of the frame are aligned with the sensitive axis of the accelerometers.
(X_c, Y_c, Z_c) are the axes of the camera frame, denoted **[c]**. Its origin is the point C, attached to the camera. Its Z-axis is aligned with the optical axis of the camera. The X-axis and the Y-axis are defined accordingly.

In this document, rotations are represented in two ways. First, the rotations between frames are expressed using rotation matrices from the SO(3) group. For instance, the rotation from a given frame **[a]** to another frame **[b]** will be denoted as $\boldsymbol{T}_{ba}$. This notation is the most intuitive for performing *plane rotations* (Farrell, 2008). Since these are rotation matrices, they can also be represented using Euler angles, quaternions, or rotation vectors (Bortz, 1971) (also referred to as Bortz vectors or Rodrigues vectors). Transformations from a matrix in SO(3) to a rotation vector in $\mathbb{R}^3$ are employed in this document, allowing rotations to be modeled with the minimal number of parameters. Specifically, each rotation matrix $\boldsymbol{T}_{ba}$ in SO(3) corresponds to a rotation vector $\boldsymbol{\phi}^{b/a}$. The skew-symmetric matrix of the rotation vector represents the exponential map to its corresponding rotation matrix in SO(3) (Dai, 2015). Based on the Euler–Rodrigues formula, the relationship between these two representations is

$$\boldsymbol{T}_{ba} = \exp\left(\boldsymbol{A}\left(\boldsymbol{\phi}^{b/a}\right)\right) = \boldsymbol{I} + \frac{\sin\|\boldsymbol{\phi}^{b/a}\|}{\|\boldsymbol{\phi}^{b/a}\|} \cdot \boldsymbol{A}\left(\boldsymbol{\phi}^{b/a}\right) + \frac{1-\cos\|\boldsymbol{\phi}^{b/a}\|}{\|\boldsymbol{\phi}^{b/a}\|^2} \cdot \boldsymbol{A}^2\left(\boldsymbol{\phi}^{b/a}\right) \tag{1}$$

where $\|\boldsymbol{\phi}^{b/a}\|$ is the euclidean norm of the rotation vector, and $\boldsymbol{A}(\boldsymbol{\phi}^{b/a})$ is the skew-symmetric matrix of the rotation vector (Dai, 2015) defined as

$$\boldsymbol{A}\left(\boldsymbol{\phi}^{b/a}\right) = \begin{bmatrix} 0 & -\boldsymbol{\phi}_z^{b/a} & \boldsymbol{\phi}_y^{b/a} \\ \boldsymbol{\phi}_z^{b/a} & 0 & -\boldsymbol{\phi}_x^{b/a} \\ -\boldsymbol{\phi}_y^{b/a} & \boldsymbol{\phi}_x^{b/a} & 0 \end{bmatrix}. \tag{2}$$

In this paper, the inverse operation of the exponential mapping, denoted as logm, converts a rotation matrix to its associated rotation vector. It is expressed as follows

$$\boldsymbol{A}\left(\boldsymbol{\phi}^{b/a}\right) = \operatorname{logm}\left(\boldsymbol{T}_{ba}\right). \tag{3}$$

The time derivative of a rotation matrix $\boldsymbol{T}_{ba}$ depends on the instantaneous angular velocity $\boldsymbol{\omega}_b^{b/a}$ of frame **[b]** relative to frame **[a]**, expressed in frame **[b]** (Farrell, 2008). The continuous time dynamics formula is

$$\dot{\boldsymbol{T}}_{ba} = -\boldsymbol{A}\left(\boldsymbol{\omega}_b^{b/a}\right) \cdot \boldsymbol{T}_{ba}. \tag{4}$$

2. Inertial Navigation

The localization system presented in this paper aims at estimating the geodetic position (latitude λ, longitude ϕ, and altitude h), the speed relative to the Earth, and the attitude of the aircraft. In the system's inertial navigation, the position information is held by the altitude $h \in \mathbb{R}$ and the SO(3) matrix $\boldsymbol{T}_{gt}$ from which the latitude and longitude can be extracted. In accordance with the above frame definitions, the matrix $\boldsymbol{T}_{gt}$ corresponds to the combination of a rotation by $\phi + \pi$ about the Z-axis and a rotation by $\lambda - \frac{\pi}{2}$ about the Y-axis (Farrell, 2008). Its expression is therefore

$$\boldsymbol{T}_{gt} = \begin{bmatrix} -\sin\lambda\cos\phi & -\sin\lambda\sin\phi & \cos\lambda \\ \sin\phi & -\cos\phi & 0 \\ \cos\lambda\cos\phi & \cos\lambda\sin\phi & \sin\lambda \end{bmatrix}. \tag{5}$$

The velocity information is held by the velocity vector expressed in the **[g]** frame, denoted as $\boldsymbol{v}_g \in \mathbb{R}^3$. The attitude information is carried by the SO(3) matrix $\boldsymbol{T}_{gm}$. Similar to the position information, the attitude angles can be directly be extracted from this rotation matrix (Farrell, 2008).

The quantities to be estimated by the system are therefore $\{\boldsymbol{T}_{gt}, h, \boldsymbol{v}_g, \boldsymbol{T}_{gm}\}$.

The time dynamics of $\boldsymbol{T}_{gt}$ and $\boldsymbol{T}_{gm}$ are determined from (4). The time dynamic of $\boldsymbol{v}_g$ is determined based on the fundamental principle of dynamics (Farrell, 2008). The time dynamic of h is expressed as a function of $\boldsymbol{v}_g$. The kinematics of these quantities

in continuous time are given by

$$\begin{cases} \dot{\boldsymbol{T}}_{gt} = -\boldsymbol{A}\left(\boldsymbol{\omega}_g^{g/t}\right) \cdot \boldsymbol{T}_{gt} \\ \dot{h} = \boldsymbol{v}_g \cdot [0 \quad 0 \quad 1]^T \\ \dot{\boldsymbol{v}}_g = \boldsymbol{T}_{gm} \cdot \boldsymbol{f}_m + \boldsymbol{g}_g\left(\boldsymbol{r}_g\right) - \boldsymbol{A}\left(\boldsymbol{\omega}_g^{g/t} + 2\boldsymbol{\omega}_g^{t/i}\right) \cdot \boldsymbol{v}_g \\ \dot{\boldsymbol{T}}_{gm} = \boldsymbol{T}_{gm} \cdot \boldsymbol{A}\left(\boldsymbol{\omega}_m^{m/i} - \boldsymbol{\omega}_m^{t/i} + \boldsymbol{\omega}_m^{t/g}\right) \end{cases} \tag{6}$$

where $\boldsymbol{f}_m$ is the true specific force, $\boldsymbol{\omega}_m^{m/i}$ is the true angular velocity of the **[m]** frame with respect to the **[i]** frame, $\boldsymbol{r}_g$ is the aircraft position vector expressed in **[g]**, and, $\boldsymbol{g}_g\left(\boldsymbol{r}_g\right)$ is the gravity vector at the aircraft location.

The specific force and angular velocity of the aircraft are measured by the sensors in the inertial measurement unit. The measurement models for the accelerometers and gyroscopes considered in the system (Farrell, 2008) are

$$\widetilde{\boldsymbol{f}}_m = \left(\boldsymbol{I} - \boldsymbol{SF}_a\right)\left(\boldsymbol{f}_m - \boldsymbol{b}_a - \boldsymbol{\nu}_a\right) \tag{7}$$

$$\widetilde{\boldsymbol{\omega}}_m^{m/i} = \left(\boldsymbol{I} - \boldsymbol{M}_\beta\right)\left(\boldsymbol{I} - \boldsymbol{SF}_g\right)\left(\boldsymbol{\omega}_m^{m/i} - \boldsymbol{d}_g - \boldsymbol{\nu}_g\right) \tag{8}$$

where $\widetilde{\boldsymbol{f}}_m \in \mathbb{R}_3$ is the measured specific force, $\widetilde{\boldsymbol{\omega}}_m^{m/i} \in \mathbb{R}_3$ is the measured angular velocity, $\boldsymbol{b}_a \in \mathbb{R}_3$ is the accelerometers bias vector, $\boldsymbol{d}_g \in \mathbb{R}_3$ is the gyros drift vector, $\boldsymbol{SF}_a \in \mathbb{M}_3(\mathbb{R})$ is the 3×3 diagonal matrix modeling the accelerometer linear scale factor, $\boldsymbol{SF}_g \in \mathbb{M}_3(\mathbb{R})$ is the 3×3 diagonal matrix modeling the gyroscope linear scale factor, $\boldsymbol{M}_\beta \in \mathbb{M}_3(\mathbb{R})$ is the 3×3 matrix modeling the gyroscope misalignment errors with respect to the **[m]** frame axes, $\boldsymbol{\nu}_a$ is the accelerometers additive white measurement noise (velocity random walk) and, $\boldsymbol{\nu}_g$ is the gyros additive white measurement noise (angular random walk).

The quantities defining these models must also be estimated by the localization system. Various approaches can be used to model the temporal dynamics of these quantities. In this paper, their temporal evolution is simply described by

$$\begin{cases} \dot{\boldsymbol{b}}_a = \boldsymbol{\nu}_b \\ \dot{\boldsymbol{d}}_g = \boldsymbol{\nu}_d \\ \dot{\boldsymbol{SF}}_a = \boldsymbol{\nu}_{SFa} \\ \dot{\boldsymbol{SF}}_g = \boldsymbol{\nu}_{SFg} \\ \dot{\boldsymbol{M}}_\beta = \boldsymbol{\nu}_\beta \end{cases} \tag{9}$$

where $\boldsymbol{\nu}_b, \boldsymbol{\nu}_d$ are Gaussian white noises in $\mathbb{R}^3$, and $\boldsymbol{\nu}_{SFa}, \boldsymbol{\nu}_{SFg}, \boldsymbol{\nu}_\beta$ are Gaussian white noises in $\mathbb{M}_3(\mathbb{R})$.

The quantities to be estimated by the localization system form the true state $\boldsymbol{x}$. This true state belongs to the manifold $\mathrm{SO(3)} \times \mathbb{R} \times \mathbb{R}^3 \times \mathrm{SO(3)} \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{M}_3(\mathbb{R}) \times \mathbb{M}_3(\mathbb{R}) \times \mathbb{M}_3(\mathbb{R})$ and is defined by

$$\boldsymbol{x} = \left\{\boldsymbol{T}_{gt},\, h, \boldsymbol{v}_g,\, \boldsymbol{T}_{gm},\, \boldsymbol{b}_a,\, \boldsymbol{d}_g,\, \boldsymbol{SF}_a,\, \boldsymbol{SF}_g,\, \boldsymbol{M}_\beta\right\}. \tag{10}$$

Ideally, inertial navigation would directly compute the true state vector from its temporal dynamics (equations (6) and (9)) and the measurements from the inertial sensors. Unfortunately, inertial navigation can only deterministically compute these states by neglecting the stochastic terms (noise) and other disturbances. The state thus computed differs from the true state, although both lie in the same space. This new state, denoted as $\bar{\boldsymbol{x}}$, determined by the inertial navigation, is called the nominal state (Sola, 2017). Its components correspond to those of the true state but are affected by errors. The nominal state is defined as

$$\bar{\boldsymbol{x}} = \left\{\bar{\boldsymbol{T}}_{gt},\, \bar{h},\, \bar{\boldsymbol{v}}_g,\, \bar{\boldsymbol{T}}_{gm},\, \bar{\boldsymbol{b}}_a,\, \bar{\boldsymbol{d}}_g,\, \bar{\boldsymbol{SF}}_a,\, \bar{\boldsymbol{SF}}_g,\, \bar{\boldsymbol{M}}_\beta\right\}. \tag{11}$$

Its temporal dynamics in continuous time are directly determined from the true equations (6) and (9) by setting the stochastic

terms and disturbances to zero. Continuous-time inertial navigation thus computes the nominal state $\bar{\boldsymbol{x}}$ based on the equations

$$\begin{cases} \dot{\bar{\boldsymbol{T}}}_{gt} = -\boldsymbol{A}\left(\bar{\boldsymbol{\omega}}_g^{g/t}\right) \cdot \bar{\boldsymbol{T}}_{gt} \\ \dot{\bar{h}} = \bar{\boldsymbol{v}}_g \cdot \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \\ \dot{\bar{\boldsymbol{v}}}_g = \bar{\boldsymbol{T}}_{gm} \cdot \bar{\boldsymbol{f}}_m + \bar{\boldsymbol{g}}_g\left(\bar{\boldsymbol{r}}_g\right) - \boldsymbol{A}\left(\bar{\boldsymbol{\omega}}_g^{g/t} + 2\bar{\boldsymbol{\omega}}_g^{t/i}\right) \cdot \bar{\boldsymbol{v}}_g \\ \dot{\bar{\boldsymbol{T}}}_{gm} = \bar{\boldsymbol{T}}_{gm} \cdot \boldsymbol{A}\left(\bar{\boldsymbol{\omega}}_m^{m/i} - \bar{\boldsymbol{\omega}}_m^{t/i} + \bar{\boldsymbol{\omega}}_m^{t/g}\right) \\ \dot{\bar{\boldsymbol{b}}}_a = \boldsymbol{0}_{3\times 1} \\ \dot{\bar{\boldsymbol{d}}}_g = \boldsymbol{0}_{3\times 1} \\ \dot{\bar{\boldsymbol{S}}}\dot{\bar{\boldsymbol{F}}}_a = \dot{\bar{\boldsymbol{S}}}\dot{\bar{\boldsymbol{F}}}_g = \dot{\bar{\boldsymbol{M}}}_\beta = \boldsymbol{0}_{3\times 3} \end{cases} \tag{12}$$

where $\bar{\boldsymbol{\omega}}_g^{g/t}$ is derived from $\bar{\boldsymbol{v}}_g$, $\bar{\boldsymbol{r}}_g$ is derived from $\bar{\boldsymbol{T}}_{gt}$, $\bar{\boldsymbol{\omega}}_g^{t/i}, \bar{\boldsymbol{\omega}}_m^{t/i}$ are derived from the Earth rotation vector, $\bar{\boldsymbol{f}}_m$, $\bar{\boldsymbol{\omega}}_m^{m/i}$ are defined as

$$\begin{cases} \bar{\boldsymbol{f}}_m = \left(\boldsymbol{I} - \bar{\boldsymbol{S}}\bar{\boldsymbol{F}}_a\right)^{-1} \cdot \widetilde{\boldsymbol{f}}_m + \bar{\boldsymbol{b}}_a \\ \bar{\boldsymbol{\omega}}_m^{m/i} = \left(\boldsymbol{I} - \bar{\boldsymbol{M}}_\beta\right)^{-1}\left(\boldsymbol{I} - \bar{\boldsymbol{S}}\bar{\boldsymbol{F}}_g\right)^{-1} \cdot \widetilde{\boldsymbol{\omega}}_m^{m/i} + \bar{\boldsymbol{d}}_g \end{cases}. \tag{13}$$

3. Error-state dynamics

The nominal state is thus deterministically computed from the equations (12) after discretization. The discretization can be performed using the Euler method (Sola, 2017). This computed nominal state is tainted with errors as the model used differ from the true one (equations (6) and (9)). These errors constitute the error state vector that is a function of the true state and the nominal state. Let $\boldsymbol{\varepsilon x}$ be the error state such as

$$\boldsymbol{x} = \bar{\boldsymbol{x}} \oplus \boldsymbol{\varepsilon x} \tag{14}$$

where $\oplus$ is a non-linear operation. Specifically, $\oplus$ is the linear addition for $\mathbb{R}^n$ quantities and the matrix multiplication for SO(3) quantities.

The choice of the error state vector is crucial. Indeed, the temporal dynamics of its components, as well as its relationships with the measurement innovations, will govern the equations of the ES-KF. For instance, regarding the attitude error, it is possible to use the error rotation matrix $\boldsymbol{\varepsilon T}_{gm} = \boldsymbol{T}_{gm}\boldsymbol{T}_{mg}$. That being said, it is also possible to use the error rotation vector $\boldsymbol{\varepsilon\phi}_g^{m/g}$ associated with $\boldsymbol{\varepsilon T}_{gm}$ (as explained in equation (1)). The latter choice is often preferred as it reduces the number of parameters to be estimated by the filter. Moreover, the time dynamics of the rotation vector are more intuitive to manipulate. In Sola (2017), a similar change of variables is performed from quaternions to 2D rotation vectors. Regarding the position error, a similar change can be applied. It is possible to represent the position error by the error rotation vector associated with the matrix $\boldsymbol{\varepsilon T}_{gt}$. However, representing a position error through angular information may not be the most logical. In this paper, the position error is hence represented by the error of the position vector in **[g]**, denoted $\boldsymbol{\varepsilon r}_g$.

In this paper, the error state vector is defined by

$$\boldsymbol{\varepsilon x} = \begin{bmatrix} \boldsymbol{\varepsilon r}_g \\ \boldsymbol{\varepsilon v}_g \\ \boldsymbol{\varepsilon\phi}_g^{m/g} \\ \boldsymbol{\varepsilon b}_a \\ \boldsymbol{\varepsilon d}_g \\ \boldsymbol{\varepsilon SF}_a \\ \boldsymbol{\varepsilon SF}_g \\ \boldsymbol{\varepsilon M}_\beta \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{T}}_{gt} \cdot \left[\boldsymbol{r}_t - \bar{\boldsymbol{r}}_t\right] \\ \boldsymbol{v}_g - \bar{\boldsymbol{v}}_g \\ \text{logm}\left(\boldsymbol{T}_{gm} \cdot \bar{\boldsymbol{T}}_{mg}\right) \\ \boldsymbol{b}_a - \bar{\boldsymbol{b}}_a \\ \boldsymbol{d}_g - \bar{\boldsymbol{d}}_g \\ vec\left(\boldsymbol{SF}_a - \bar{\boldsymbol{SF}}_a\right) \\ vec\left(\boldsymbol{SF}_g - \bar{\boldsymbol{SF}}_g\right) \\ vec\left(\boldsymbol{M}_\beta - \bar{\boldsymbol{M}}_\beta\right) \end{bmatrix} \tag{15}$$

where $\boldsymbol{r}_t$ is the aircraft position vector expressed in **[t]**, and the operator $vec : \mathbb{M}_3(\mathbb{R}) \to \mathbb{R}^n$ allows the vectorisation of matrices to have error states in $\mathbb{R}^n$.

The temporal dynamics are not necessarily linear with respect to the state. Therefore, linearization around the estimated nominal state must be performed. This linearization introduces high orders errors. Moreover, as the linearization is based on the nominal

state, which is not the true state, issues of observability may arise. The system might falsely perceive certain quantities as observable. Ideally, one would want an error state vector that is independent of the estimated state. This way, the linearization does not result in false observability. This conception is known as an Invariant Kalman Filter (Barrau, 2015). In the system presented in this paper, observability issues are not addressed.

As the temporal dynamics of the error state vector are nonlinear for most components, it is necessary to linearize around a known point. Typically, linearization is performed near the null error vector $\boldsymbol{\varepsilon x} = \mathbf{0}$. This corresponds to linearizing at $\boldsymbol{x} = \bar{\boldsymbol{x}}$. When linearizing these equations, the error state no longer evolves in the same space. The evolution transitions from the curved space to the tangent space (Barrau, 2015; Sola, 2017). To highlight this aspect, the differential method is used for linearization (instead of the Taylor series expansion method). This method results in the appearance of the prefix (d), indicating the linearized state. Consequently, the Error-State Kalman Filter does not estimate $\boldsymbol{\varepsilon x}$, but its differential at $\bar{\boldsymbol{x}}$, denoted as $d\boldsymbol{\varepsilon x}$. This linearized error state vector is defined as follows

$$d\boldsymbol{\varepsilon x} = \begin{bmatrix} d\boldsymbol{\varepsilon r}_g^T & d\boldsymbol{\varepsilon v}_g^T & \boldsymbol{\varphi}_g^T & d\boldsymbol{\varepsilon b}_a^T & d\boldsymbol{\varepsilon d}_g^T & d\boldsymbol{\varepsilon SF}_a^T & d\boldsymbol{\varepsilon SF}_g^T & d\boldsymbol{\varepsilon M}_\beta^T \end{bmatrix}^T \tag{16}$$

where $\boldsymbol{\varphi}_g = d\boldsymbol{\varepsilon}\phi_g^{m/g}$ for notational purposes.

The time evolution of the linearized error state can be computed by differentiating the continuous time equations (6) and (9) in $\boldsymbol{x} = \bar{\boldsymbol{x}}$. Another method is to express the non-linear error state vector temporal dynamics using its definition (15), the true state dynamics (6) (9) and the nominal state dynamics (12). From the non-linear equations, the linearization is direct. Examples are developed in (Sola, 2017). The continuous time linearized error state kinematics are finally given by

$$d\dot{\boldsymbol{\varepsilon r}}_g = d\boldsymbol{\varepsilon v}_g - \left[\boldsymbol{A}\left(\bar{\boldsymbol{\omega}}_g^{g/t}\right) + \boldsymbol{A}\left(\bar{\boldsymbol{v}}_g\right) \cdot \boldsymbol{M}_{\eta g}\right] \cdot d\boldsymbol{\varepsilon r}_g \tag{17}$$

$$\begin{aligned} d\dot{\boldsymbol{\varepsilon v}}_g = &- \boldsymbol{A}\left(\bar{\boldsymbol{T}}_{gm} \cdot \bar{\boldsymbol{f}}_m\right) \cdot \boldsymbol{\varphi}_g + \bar{\boldsymbol{T}}_{gm} \cdot d\boldsymbol{\varepsilon f}_m \\ &+ \left[\boldsymbol{A}\left(\bar{\boldsymbol{v}}_g\right) \cdot \boldsymbol{M}_{\rho g} - \boldsymbol{A}\left(\bar{\boldsymbol{\omega}}_g^{g/t} + 2 \cdot \bar{\boldsymbol{\omega}}_g^{t/i}\right)\right] \cdot d\boldsymbol{\varepsilon v}_g \\ &+ \left[\boldsymbol{M}_{gg} + 2\boldsymbol{A}\left(\bar{\boldsymbol{v}}_g\right) \boldsymbol{A}\left(\bar{\boldsymbol{\omega}}_g^{t/i}\right) \cdot \boldsymbol{M}_{\eta g}\right] \cdot d\boldsymbol{\varepsilon r}_g \end{aligned} \tag{18}$$

$$\begin{aligned} \dot{\boldsymbol{\varphi}}_g = &\bar{\boldsymbol{T}}_{gm} \cdot d\boldsymbol{\varepsilon \omega}_m - \boldsymbol{A}\left(\bar{\boldsymbol{T}}_{gm} \cdot \bar{\boldsymbol{\omega}}_m^{m/i} + \bar{\boldsymbol{\omega}}_g^{g/m}\right) \cdot \boldsymbol{\varphi}_g \\ &- \boldsymbol{M}_{\rho g} \cdot d\boldsymbol{\varepsilon v}_g - \boldsymbol{A}\left(\bar{\boldsymbol{\omega}}_g^{t/i}\right) \cdot \boldsymbol{M}_{\eta g} \cdot d\boldsymbol{\varepsilon r}_g \end{aligned} \tag{19}$$

$$d\dot{\boldsymbol{\varepsilon b}}_a = d\dot{\boldsymbol{\varepsilon d}}_g = \mathbf{0}_{3\times 1} \tag{20}$$

$$d\boldsymbol{\varepsilon}\dot{\boldsymbol{S}}\boldsymbol{F}_a = d\boldsymbol{\varepsilon}\dot{\boldsymbol{S}}\boldsymbol{F}_g = d\boldsymbol{\varepsilon}\dot{\boldsymbol{M}}_\beta = \mathbf{0}_{3\times 3} \tag{21}$$

where $\boldsymbol{M}_{\eta g} = \boldsymbol{M}_{\rho g} = \begin{bmatrix} 0 & -\frac{1}{R_e+\widehat{h}} & 0 \\ \frac{1}{R_n+\widehat{h}} & 0 & 0 \\ 0 & -\frac{\tan\widehat{\lambda}}{R_e+\widehat{h}} & \end{bmatrix}$, R_n and R_n are the north and east radii of curvature (Clynch, 2002), $\boldsymbol{M}_{gg}$ is the 3×3 matrix linking the gravity vector errors to the position errors in **[g]**, $d\boldsymbol{\varepsilon f}_m$ is a function of $d\boldsymbol{\varepsilon b}_a$ and $d\boldsymbol{\varepsilon M}_{\text{SFa}}$, and, $d\boldsymbol{\varepsilon \omega}_m$ is a function of $d\boldsymbol{\varepsilon d}_g$, $d\boldsymbol{\varepsilon M}_{\text{SFg}}$ and $d\boldsymbol{\varepsilon M}_\beta$.

The kinematics of the differentiate of the error state vector are thus linear, such that $d\dot{\boldsymbol{\varepsilon x}}(t) = \boldsymbol{F}(t) \cdot d\boldsymbol{\varepsilon x}(t)$.

4. Sensors measurement models

The hybridization sensors used in the system are the barometric altimeter, the GNSS, and the vision sensor. Their modeling and measurements are detailed in the following paragraphs.

Barometric altimeter: This sensor is traditionally used in conjunction with IMUs to stabilize the vertical channel. This is because an IMU tends to diverge over time in its estimation of the vertical component of position. The barometric altimeter provides a reference altitude measurement. The measurement model chosen is willingly simple as the system focus on visual hybridization. It is given by

$$\widetilde{y}_{baro} = \widetilde{h}^{\text{ref}} = h + w_{\text{baro}} \tag{22}$$

with $h \in \mathbb{R}$ the true altitude, $\widetilde{h}^{\text{ref}} \in \mathbb{R}$ the measured reference altitude, and w_{baro} the baro-altimeter altitude measurement error, typically modeled as an additive white Gaussian noise.

GNSS Receiver: Inertia-GNSS-baro hybridization constitutes the state of the art in the field of navigation and guidance in civil aviation. Therefore, a GNSS receiver is also considered in this system. However, the system proposed in this paper is designed to operate in GNSS-challenged environments. Consequently, the GNSS receiver is declared unavailable in many use cases. The model used is a simple one that allows for loose GNSS hybridization. The GNSS receiver output measurement directly provides an estimated position in the ECEF frame. This model is given by

$$\widetilde{\boldsymbol{y}}_{GNSS} = \widetilde{\boldsymbol{r}}_{t\_GNSS} = \boldsymbol{r}_t + \boldsymbol{w}_{\text{GNSS}} \tag{23}$$

with $\boldsymbol{r}_t \in \mathbb{R}^3$ the true aircraft position in **[t]**, $\widetilde{\boldsymbol{r}}_{t\_GNSS} \in \mathbb{R}^3$ the estimated GNSS position, and $\boldsymbol{w}_{\text{GNSS}} \in \mathbb{R}^3$ the GNSS position estimation error, modeled as an additive white Gaussian noise. That said, the GNSS position error is not necessarily white as it depends on the geometry of the satellites. In order to simplify the simulations, the position error is overbounded by a sphere with a large variance.

Visual sensor: This sensor observes the environment of the aircraft and is oriented towards the front of the aircraft. The vision sensor can be a monocular camera, a pair of stereoscopic cameras, an infrared (IR) camera, a short-wave infrared (SWIR) camera, or a combination of these. The objective of the vision sensor is to provide a data stream to an image processing module. The useful measurement provided by the image processing module to the localization system consists of characteristic points in the image.

The useful measurement can then be translated into pixel coordinates (u, v) coming from the camera's image frame. These pixel coordinates represent the image of a 3D point $\boldsymbol{x}$ expressed in the camera frame **[c]**, such that (Hartley and Zisserman, 2003)

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \tag{24}$$

where (u, v) are the pixel coordinates in the 2D image plane of the camera, $[x_c, y_c, z_c]^T$ is the 3D position of the object point in the camera frame, f_x, f_y are the intrinsic parameters of the camera, (u_0, v_0) are the image frame origin coordinates, and $s = z_c$ is the homogeneity parameter.

To simplify the system, the useful measurement considered in this paper is not expressed in pixels but in angles (Figure 3). This allows the model to no longer depend on the intrinsic parameters of the camera. The model is then given by

$$s \begin{bmatrix} \tan \alpha_x \\ \tan \alpha_y \\ 1 \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \tag{25}$$

where (α_x, α_y) constitute the video measurement, and $s = z_c$ is the homogeneity parameter.

The video measurements hence are (Vezinet et al., 2013)

$$\widetilde{\boldsymbol{y}}_{video} = \begin{bmatrix} \tan \widetilde{\alpha}_x \\ \tan \widetilde{\alpha}_y \end{bmatrix} = \begin{bmatrix} \tan(\alpha_x) + bbx \\ \tan(\alpha_y) + bby \end{bmatrix} \tag{26}$$

where α_x, α_y are the true line of sight ratios, $\tan \widetilde{\alpha}_x, \tan \widetilde{\alpha}_y$ the measured line of sight ratios, and bbx, bby the angle measurement errors, typically modeled as an additive white noise.

What Do the Measured Angles Represent?

The pinhole camera model establishes the relationship between a 2D measured point and an unknown 3D object point. The video system, therefore, measures a line of sight between the camera and the 3D object point. The vision hybrid system relies on alignment with a known landmark whose position is given by a database. For this reason, the system assumes that the position of a characteristic point is known in an absolute reference frame, specifically the terrestrial reference frame **[t]**. Using the measured line of sight, this known position, and the transformation matrices between **[t]** and **[c]**, the measurement $\widetilde{\boldsymbol{y}}_{\text{video}}$ can be related to the state vector $\boldsymbol{x}$. The available landmarks during the approach are those attached to the runway. The selection of the type and number of landmarks is crucial for the system, affecting its robustness and operability. This criterion for selecting landmarks is discussed in subsection III.6.

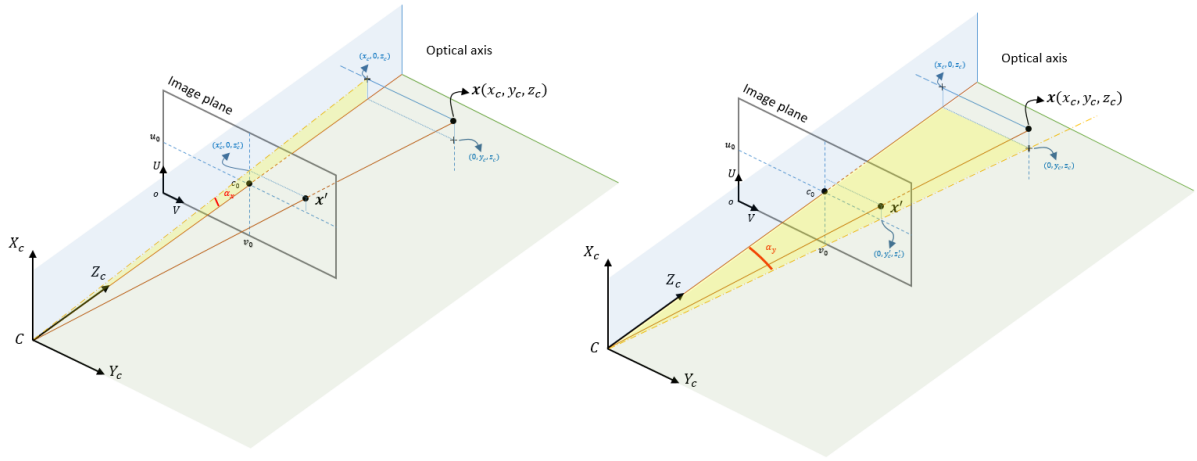


Figure 3: Pinhole camera model - Angular measurement

5. Error-State Kalman Filtering process

Step 1 - Nominal state: The ES-KF framework is a multi-step process. Initially, inertial navigation propagates the nominal state using the measurements from the inertial sensors. To do this, the inertial navigation equations Equation (12) are discretized. They allow the estimation of the nominal state at the current time step (k): $\bar{\boldsymbol{x}}_k$. This nominal state is then provided as inputs to the Error-State Kalman Filter so that it can estimate the errors in this nominal state.

Step 2 - ES-KF prediction: The ES-KF deals with the differential of the error state vector because the propagation model and the measurement models are not necessarily linear with respect to the state vector. The error state and its associated covariance are propagated using the following Kalman equations

$$\widehat{d\boldsymbol{\varepsilon x}}_{k|k-1} = \boldsymbol{\Phi}_k \cdot \widehat{d\boldsymbol{\varepsilon x}}_{k-1|k-1} \qquad \textit{(State prediction)} \tag{27}$$

$$\boldsymbol{\Sigma}_{k|k-1} = \boldsymbol{\Phi}_k \cdot \boldsymbol{\Sigma}_{k-1|k-1} \cdot \boldsymbol{\Phi}_k^T + \boldsymbol{Q}_k \qquad \textit{(State covariance prediction)} \tag{28}$$

where $\boldsymbol{\Phi}_k$ is the state transition matrix which is determined from the linear equations (17) after being discretized , $\boldsymbol{Q}_k$ is the state noise covariance matrix at time k, $\boldsymbol{\Sigma}_{k-1|k-1}$ is the Kalman state covariance matrix at time $k-1$, and $\boldsymbol{\Sigma}_{k|k-1}$ is the predicted Kalman state covariance matrix at time k.

Step 3 - ES-KF correction: In the context of an ES-KF, the observation model does not directly link the measurement $\widetilde{\boldsymbol{y}}$ to the state $\boldsymbol{x}$, but rather links the Kalman filter innovation $\boldsymbol{\varepsilon y}$ to the error state $\boldsymbol{\varepsilon x}$.

The innovation is given by $\boldsymbol{\varepsilon y} = \widetilde{\boldsymbol{y}} - \boldsymbol{h}(\bar{\boldsymbol{x}})$, where $\boldsymbol{h}$ is generally a non-linear function. The innovation thus represents the difference between the measured observation and the observation reconstructed by the nominal state.

To obtain the observation model for the ES-KF, one needs to express the innovation in terms of the error state vector. This relationship is generally non-linear

$$\boldsymbol{\varepsilon y} = \boldsymbol{h}_{ES-KF}(\boldsymbol{\varepsilon x}) + \text{error}. \tag{29}$$

By linearizing (29) using the differential method around $\boldsymbol{\varepsilon x} = \boldsymbol{0}$, i.e., around $\boldsymbol{x} = \bar{\boldsymbol{x}}$, the linear observation model obtained is

$$d\boldsymbol{\varepsilon y} = \boldsymbol{H} \cdot d\boldsymbol{\varepsilon x} \tag{30}$$

where $\boldsymbol{H}$ is the Jacobian matrix of $\boldsymbol{h}_{ES-KF}$ in $\boldsymbol{\varepsilon x} = \boldsymbol{0}$, i.e. in $\boldsymbol{x} = \bar{\boldsymbol{x}}$.

This linear model allows the following equations to be implemented.

$$\boldsymbol{Z}_k = \widetilde{d\boldsymbol{\varepsilon y}} - \boldsymbol{H}_k \cdot \widehat{d\boldsymbol{\varepsilon x}}_{k|k-1} \qquad \textit{(ES-KF Innonvation)} \tag{31}$$

$$\boldsymbol{K}_k = \boldsymbol{\Sigma}_{k|k-1} \cdot \boldsymbol{H}_k^T \cdot \left[\boldsymbol{H}_k \cdot \boldsymbol{\Sigma}_{k|k-1} \cdot \boldsymbol{H}_k^T + \boldsymbol{R}_k\right]^{-1} \qquad \textit{(Kalman gain)} \tag{32}$$

$$\widehat{d\boldsymbol{\varepsilon x}}_{k|k} = \widehat{d\boldsymbol{\varepsilon x}}_{k|k-1} + \boldsymbol{K}_k \cdot \boldsymbol{Z}_k \qquad \textit{(State correction)} \tag{33}$$

$$\boldsymbol{\Sigma}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k \cdot \boldsymbol{H}_k) \cdot \boldsymbol{\Sigma}_{k|k-1} \qquad \textit{(State covariance correction)} \tag{34}$$

where $\boldsymbol{Z}_k$ is the ES-KF innovation (which is different to $\boldsymbol{\varepsilon y}$) and $\boldsymbol{R}_k$ is the measurement covariance error.

For the barometric altimeter, the KF innovation is $\varepsilon y_{baro} = \widetilde{h}^{ref} - \bar{h}$. Using the definition of $\widetilde{h}^{ref}$ (22), the measurement model is directly linear

$$\varepsilon y_{baro} = h + \omega_{baro} - \bar{h} = \boldsymbol{\varepsilon r}_g \cdot [0 \quad 0 \quad 1]^T + w_{baro}. \tag{35}$$

For GNSS, using the same method, projecting the GNSS measurement into **[g]** the following model is obtained

$$\boldsymbol{\varepsilon y}_{GNSS} = \boldsymbol{\varepsilon r}_g + \boldsymbol{w}_{GNSS}. \tag{36}$$

For vision, the measurement model (25) makes the link between $\widetilde{\boldsymbol{y}}_{video}$ and the line of sight $\boldsymbol{h}_c$. The line of sight can be expressed as a function of the state and various parameters by

$$\boldsymbol{h}_c = \boldsymbol{T}_{cm}\boldsymbol{T}_{mg}\boldsymbol{T}_{gt} \cdot \left[\boldsymbol{r}_t^{feat} - \boldsymbol{r}_t\right] - \boldsymbol{T}_{cm}\boldsymbol{L}_m \tag{37}$$

where $\boldsymbol{r}_t^{feat}$ is the known position of the feature (see section III.6), and $\boldsymbol{L}_m$ is the lever arm between the IMU position and the camera position. The lever arm error is assumed to be negligible.

Combining Equations (25), (26), and (37), the non-linear measurement model for the video innovation $\boldsymbol{\varepsilon y}_{video}$ is obtained. By applying the differential method, the following linearized model is then

$$\begin{aligned} d\boldsymbol{\varepsilon y}_{video} = & - \boldsymbol{M}_{\alpha hc} \cdot \bar{\boldsymbol{T}}_{cm} \cdot \boldsymbol{A}\left(\bar{\boldsymbol{h}}_m - \bar{\boldsymbol{L}}_m\right) \cdot \boldsymbol{\xi} \\ & - \boldsymbol{M}_{\alpha hc} \cdot \bar{\boldsymbol{T}}_{cm} \cdot \bar{\boldsymbol{T}}_{mg} \cdot \boldsymbol{A}\left(\bar{\boldsymbol{h}}_g\right) \cdot \boldsymbol{\varphi}_g \\ & + \boldsymbol{M}_{\alpha hc} \cdot \bar{\boldsymbol{T}}_{cm} \cdot \bar{\boldsymbol{T}}_{mg} \cdot \left[\boldsymbol{A}\left(\bar{\boldsymbol{T}}_{gt} \cdot \bar{\boldsymbol{h}}_t\right) \cdot \boldsymbol{M}_{\eta g} - \boldsymbol{I}\right] \cdot d\boldsymbol{\varepsilon r}_g \\ & + \boldsymbol{M}_{\alpha hc} \cdot \bar{\boldsymbol{T}}_{cm} \cdot \bar{\boldsymbol{T}}_{mg} \cdot d\boldsymbol{r}_g^{feat} \\ & + \text{error} \end{aligned} \tag{38}$$

where $\boldsymbol{M}_{\alpha hc}$ is a function of $\bar{\boldsymbol{h}}_c$, $\boldsymbol{\xi} = d\boldsymbol{\varepsilon\phi}_c^{m/c}$ represents the rotational error vector between frames **[m]** and **[c]**, $\bar{\boldsymbol{h}}_m$, $\bar{\boldsymbol{h}}_g$ and $\bar{\boldsymbol{h}}_t$ represent the feature line of sight expressed in various frames, and $d\boldsymbol{\varepsilon r}_g^{feat}$ is the position error vector of the landmark in frame **[t]**. The quantities $d\boldsymbol{\varepsilon r}_g^{feat}$ and $\boldsymbol{\xi}$ are added to the error state vector $d\boldsymbol{\varepsilon x}$.

Step 4 - Iterative correction step: At this stage, inertial navigation has estimated the nominal state at the current iteration, and the ES-KF has provided an estimation of the error state vector. To reconstruct a consolidated estimated state, these two must be combined according to Equation (14). Usually (Gróf et al., 2022; Watanabe et al., 2019; Sola, 2017), this iterative correction step process is performed in a closed-loop manner (feedback implementation). The error state vector estimated by the ES-KF is directly used to correct the nominal state. Hence the correction step becomes

$$\bar{\boldsymbol{x}} \leftarrow \bar{\boldsymbol{x}} \oplus \widehat{d\boldsymbol{\varepsilon x}}. \tag{39}$$

As the corrected nominal state carries the error information, the state vector must be set to zero at each feedback step. Hence, in the ES-KF, the state prediction equation (27) is unnecessary to implement.

Another approach is to implement an open-loop framework (feed-forward implementation). The nominal state and the estimated state error are combined to form a consolidated output state $\widehat{\boldsymbol{x}}_{\text{out}}$. This allows the system to provide a corrected state while the nominal state evolves independently of the ES-KF. This method is useful when it is undesirable to risk perturbing inertial navigation with potentially erroneous hybridization measurements. That said, this open-loop architecture can only be used with an inertial navigation system that is self-sufficient and of high precision. The correction step becomes

$$\widehat{\boldsymbol{x}}_{\text{out}} = \bar{\boldsymbol{x}} \oplus \widehat{d\boldsymbol{\varepsilon x}}. \tag{40}$$

Sometimes, it may be desired to loop back on certain states but not others. This is the case here. In the absence of GNSS, it seems that the only hybridization sensor capable of providing observation information on the aircraft's orientation is vision. For this, it must either independently determine a 6D pose (loose hybridization requiring at least four landmarks) or measure a reconstructed point such as the vanishing point (Watanabe et al., 2019). However, a navigation-grade IMU can maintain its attitude angles (see simulation results) within the requirements bounds throughout the approach during the coasting period (alone, without GNSS). Therefore, the system described in this paper does not correct the nominal state with the attitude error estimated by the ES-KF. Hence, the vision does not need to provide attitude information. Moreover, a high-precision IMU does not rely on video measurements to correct its inertial sensor biases. These states should not be corrected, as it risks polluting the entire inertial navigation.

However, even a high-precision IMU diverges in position over time. In the absence of GNSS, the system relies on video measurements to correct this error. It is then necessary to correct the nominal position with the position error estimated by the ES-KF. The iterative correction step implemented by this system is thus a mix between the feedback and the feed-forward implementations. It can be called a "semi-closed" implementation. In this architecture, the component of the error states estimated by the ES-KF that are used to correct the nominal state are reset to zero. The ones that are not looped back remain unchanged.

The correction step performed is an exponential feedback (Barrau, 2015). It accounts for the fact that the estimated error state is in a tangent space while the nominal state is in a curved space. The feedback equations are

$$\bar{\boldsymbol{T}}_{gt} \leftarrow \exp\left(\boldsymbol{A}\left(\boldsymbol{M}_{\eta g} \cdot \widehat{d\boldsymbol{\varepsilon r}_g}\right)\right) \cdot \bar{\boldsymbol{T}}_{gt} \tag{41}$$

$$\bar{h} \leftarrow \bar{h} + \widehat{d\boldsymbol{\varepsilon r}_g} \cdot [0 \quad 0 \quad 1]^T \tag{42}$$

$$\bar{\boldsymbol{v}}_g \leftarrow \bar{\boldsymbol{v}}_g + \widehat{d\boldsymbol{\varepsilon v}_g} \tag{43}$$

The error state vector is therefore partially reset to 0 by the operations

$$\begin{cases} \widehat{d\boldsymbol{\varepsilon r}_g} \leftarrow \mathbf{0}_{3\times 1} \\ \widehat{d\boldsymbol{\varepsilon v}_g} \leftarrow \mathbf{0}_{3\times 1} \end{cases}. \tag{44}$$

6. Number of visual features

In vision-based navigation, there are various navigation techniques. The most popular are visual odometry, map alignment, and georeferenced landmark alignment (Ben-Afia et al., 2014). For a precision approach of an airliner, it is preferable to limit the database required. Additionally, since IMUs inherently perform absolute navigation (i.e. relative to the Earth and not relative to the runway), it is better to choose an absolute technique rather than a relative or dead-reckoning one. It is also essential to select a technique that provides significant information when used with a high-precision IMU found in airliners. For these reasons, landmark alignment is chosen for this system. It provides significant position information while only requiring a priori knowledge of the landmark's position. It is reasonable to assume that the position of a specific point on the runway can be added to the AIPs (Aeronautical Information Publication), especially since points related to Ground Based Augmentation System (GBAS) approaches are already documented.

Next, the number and type of landmarks considered must be addressed. These landmarks need to exist in all approaches and detectable from a reasonable distance. The runway seems to be the sole candidate, as VORs, ILS, and other control towers are not sufficiently detectable. In the literature, various runway points are used. One can differentiate between physical points such as the four corners of the runway and reconstructed points like the vanishing point (Watanabe et al., 2019). As discussed in the previous section, some landmarks provide position information, while others provide attitude information. A navigation-grade IMU does not require assistance to maintain its attitude; hence, using a point like the vanishing point is unnecessary. Moreover, this point can only be accurately reconstructed from a limited distance from the runway, reducing the system's operational zone if it relies on this measurement for attitude estimation. How many landmarks to use? If loose hybridization is desired, at least four landmarks are necessary to resolve the ambiguity in a Perspective-n-Point (PnP) algorithm (D'Alfonso et al., 2021). Therefore, four distinct landmarks on the runway must be measured. Four possible candidates are the four corners of the runway. Several kilometers away from the runway, the lines of sight of the four points almost coincide, making the landmark correlations strong. In a Kalman filter, these correlations must be quantified to avoid poor filter performance.

Tight hybridization allows for the direct use of angle measurements for hybridization. Information is provided for each measured landmark. For the hybrid system presented, given that only position information is required, it is possible to use just one landmark. This advantageously eliminates the need to model correlations between different landmarks. Moreover, since only one landmark is needed, it can be selected to be detectable from a farther distance and more precise. In this regard, the barycenter of the runway is a good candidate. Once the trapezium representing the runway is detected (Vezinet et al., 2013), the

landmark is detected. This extends the system's operational period when GNSS is absent, ultimately providing better availability compared to systems requiring the detection and reconstruction of multiple landmarks.

IV. SIMULATION RESULTS

1. Simulation Framework

The vision-integrated localization system presented in this paper is designed for the final approach segment of a commercial aircraft. Its semi-closed loop ES-KF framework was chosen to extend the system's operability range while ensuring high-precision performance in the event of GNSS signal loss. To verify and validate its theoretical performances, the system was implemented in a comprehensive Monte Carlo simulator. The principle of a Monte Carlo simulator is to perform a large number of random draws on the input parameters of a simulation to obtain representative output statistics of the simulated system's performance. The random draws are conducted according to the probability distribution of the parameters (see Table 1). A Monte Carlo simulation particularly enables the estimation of the standard deviation of a variable over time, for a non-stationary variable, or one whose temporal behavior does not follow a linear law. In the context of this paper, the random draws are performed on the errors associated with the various sensors. Figure 4 illustrates the overall architecture of the simulator.

The implemented simulator is a comprehensive simulation tool. It takes two inputs: a trajectory definition (waypoints, speed and altitude profile) and a sensor error budget. Based on the trajectory definition, the true physical quantities are generated using a physical model of the aircraft. In parallel, accurate non-linear sensor models are used to propagate the true sensor errors. In particular, video measurements are generated by taking the line of sight between the true position of the aircraft and the center of mass of the runway. Thus, video measurements are derived from a single feature and do not assume any specific technological solution. Consequently, the vision sensor could independently be a monocular camera, a pair of stereoscopic cameras, or an infrared sensor. No delay models the potential image processing time required to detect and track the runway. A database is provided to the system containing the position of the center of mass of the runway in the terrestrial reference frame. GNSS measurements are generated from the true position of the carrier with simple additive noise. This choice was made to avoid exhaustively simulating the various possible geometries of GNSS constellations, especially since this sensor will be assumed unavailable in the majority of scenarios. The baroaltimeter reference altitude measurements are generated similarly. The starting point of all sensor errors is the random draw based on the input error budget (see Table 1). The physical quantities are then combined with the sensor errors to generate sensor measurements. This first part of the simulator constitutes the simulation of the system's environment.
The second part of the simulator constitutes the system presented in this paper. The architecture of the ES-KF in a semi-closed loop is implemented. It takes the generated sensor measurements as inputs and estimates the navigation state vector using the equations introduced in this paper. Several assumptions were made in this implementation. Inertial navigation is performed at a frequency of 10 Hz. The hybridization sensors (GNSS position, video measurement, barometric altitude) take their measurements at a frequency of 1 Hz. Regarding the two components of the error state vector related to the vision measurement ($\boldsymbol{\xi}$ and $d\boldsymbol{\varepsilon r}_g^{feat}$), a temporal dynamic had to be selected. As a first approximation, a constant dynamic was chosen. Error covariances are also calculated using the ES-KF.
The true physical quantities are combined with the estimated state by the system to calculate the errors committed by the system. These errors are then traced and combined using multiple Monte Carlo runs.

2. System precision potential

To test the precision of the inertial-vision hybridization during an approach in the event of GNSS loss, a reference trajectory was defined. This trajectory simulates a takeoff, a go-around, and a landing at Rouen Airport, France. The attitude and speed profiles represent the typical profiles of an A320 class airliner. The trajectory lasts approximately 48 minutes. The IMU alignment phase lasts 3 minutes. It is assumed that the vision system detects the centroid of the runway at a geometric distance of 10 km from it. With the defined trajectory, this segment represents the last 3 minutes of the final approach segment. Barometric altitude measurements are assumed to be always available. The presence of GNSS measurements will depend on the scenario.

Scenario 1: GNSS remains available throughout the entire approach. Vision hybridization is deactivated for this scenario, representing the nominal operational case of a civil aircraft during an PBN CAT I approach.

Scenario 2: GNSS becomes unavailable starting at 8 minutes and 20 seconds. Vision is unavailable for the entire duration of the approach. This scenario examines the behavior of a navigation-grade IMU during coasting (i.e., operating alone without hybridization).

Scenario 3: GNSS becomes unavailable starting at 8 minutes and 20 seconds. Vision becomes available at a distance of 10 km from the runway, i.e., starting at 45 minutes and 27 seconds. This scenario highlights the contribution of the hybridization architecture proposed in this paper.

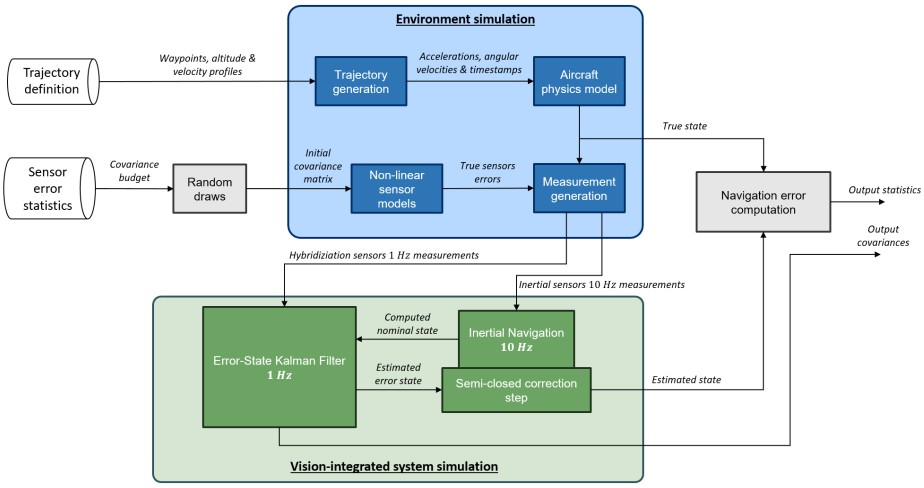


Figure 4: Simulator overall architecture

Theses scenarios were tested, with 100 random draws for each scenario. The sensor error budgets are summarized in Table 1. The results are summarized in Figure 5. Subplots (a), (c), and (e) show the position estimation RMS (Root Mean Square) error along the North, the West and the zenith axes, along with their 1-sigma bound, for each scenario. Subplot (f) provides a zoomed-in view of subplot (e), offering a more detailed examination of the final position estimation error after it has been corrected using vision-based data in scenario 3. Subplots (b) and (d) illustrate the roll, pitch and heading estimation RMS respectively for the scenarios 1, 2 and 3, along with their 1-sigma bound.

Table 1: Sensors errors budget

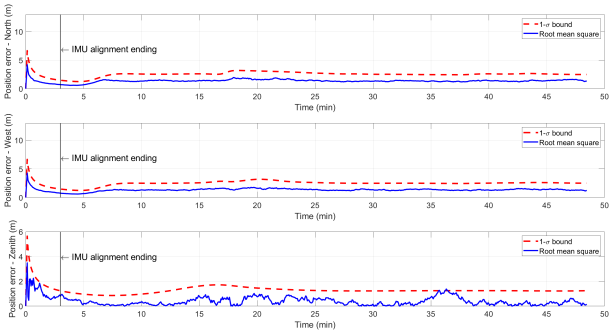
Sensor	Error type	Error budget (Std)
Accelerometer	Constant bias	$40\ \mu$g
	Linear scale factor	0.4 ppm
	Misalignment errors	$6.6\ \mu$g per axis
	Velocity random walk (VRW)	$9.81 * 10^{-6}$ m/s/$\sqrt{s}$
Gyros	Constant drift	0.01 deg/h
	Linear scale factor	10^{-3} ppm
	Misalignment errors	$10^{-5}\ \mu$rad/s
	Angle random walk (ARW)	$2 * 10^{-3}$ deg/$\sqrt{h}$
GNSS	Markovian position bias	0 m
	White noise	5 m
Barometric altimeter	White noise	5 m
Visual Sensor	Unit vector bias $\frac{\boldsymbol{h}_c}{\|\boldsymbol{h}_c\|}$	10^{-3} (no units)
	White noise	10^{-3} (no units)

Scenario 1 represents the nominal case and produced logical and expected results. The position error is less than one meter along each axis. The attitude error is less than $2 * 10^{-2}$ mrad for pitch and roll, and less than 0.5 mrad for heading. Consequently, the position and attitude estimation errors were minimal, which is consistent with the typical operation of an airliner navigation system under GNSS availability. These results confirm the normal functionality of the system in optimal conditions.

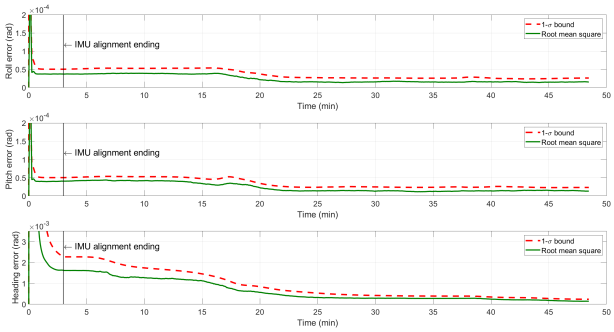
As expected, Scenario 2 showed noticeable position error drift. During the coasting phase, the IMU alone could not provide sufficient accuracy, resulting in a position error of 1 km along the North and West axes. For the Zenith axis, there was no drift; only an error of 2 meters was observed, thanks to the barometric altimeter. Despite this, the attitude angle estimations remained precise, with errors less than $4 * 10^{-2}$ mrad for pitch and roll and around 1 mrad for heading. This outcome validates a key assumption of the system: that a navigation-grade IMU ensures accurate attitude maintenance during coasting. As a result, vision-based hybridization can primarily focus on mitigating position and velocity drifts.

Scenario 3 demonstrated the potential accuracy performance of the hybridized system. In the absence of GNSS, vision-based hybridization effectively reduced the IMU's position drift. The position error decreased from 1 km to 15 meters along the North and West axes at the DA/H. Following an initial correction that reduces the position error by more than half, the error slowly

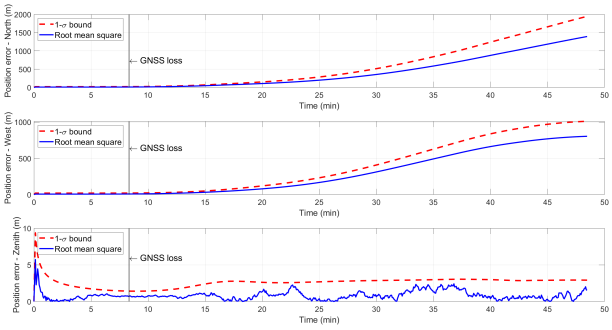
converges towards a position error of about ten meters. This slow transient period is due to the fact that, at a distance of 10 km from the runway, the visual information is not very precise. The closer the systems gets to the runway, the more precise the vision angles observations becomes and the more accurately the position error can be corrected.



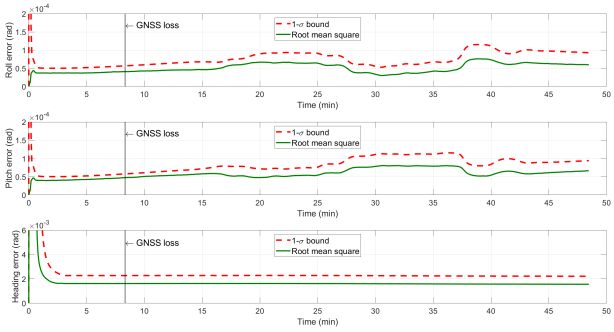
(a) Scenario 1 - Position errors and 1σ bound



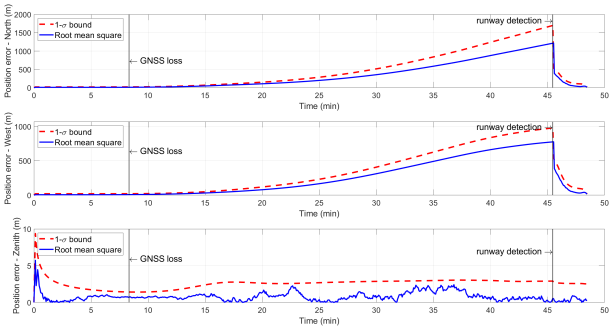
(b) Scenario 1 - Attitude errors and 1σ bound



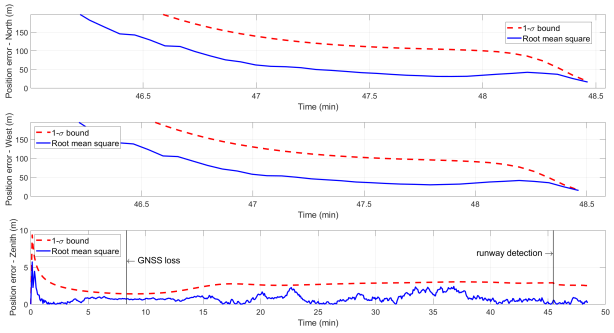
(c) Scenario 2 - Position errors and 1σ bound



(d) Scenario 2 - Attitude errors and 1σ bound



(e) Scenario 3 - Position errors and 1σ bound



(f) Scenario 3 - zoomed-in view of position errors and 1σ bound

Figure 5: Simulation results - Scenario 1: IMU/GNSS without video - Scenario 2: IMU/GNSS without video / GNSS loss at 500 seconds - Scenario 3 : IMU/GNSS with video - GNSS loss at 500 seconds

V. CONCLUSION

In this paper, a navigation system for a civil aircraft has been developed, combining a navigation-grade IMU, a GNSS receiver, a barometric altimeter, and a vision sensor. The aim of this system is to meet the requirements of PBN CAT I operations, even in the event of GNSS loss of continuity. A regulation study was conducted, demonstrating the system's feasibility. The system architecture, along with mathematical models for propagation and measurement, was described. This innovative solution allows the inertial-vision system to operate from the measurement of a single characteristic point, thus enhancing its operability range. A performance illustration is provided through a Monte Carlo simulation, which highlights the system's potential to meet the targeted specifications.

VI. ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to the PANS-OPS team at ENAC for their invaluable assistance in developing a potential regulatory framework. Their expertise and guidance in this domain were crucial to the completion of this work. Additionally, we extend our thanks to the reviewers for their insightful comments and constructive feedback, which greatly contributed to improving the quality of this paper.

REFERENCES

(2016). Vision website. `https://w3.onera.fr/h2020_vision/node/1`.

Angermann, M., Wolkow, S., Schwithal, A., Tonhäuser, C., and Hecker, P. (2015). High precision approaches enabled by an optical-based navigation system. In *Proceedings of the ION 2015 Pacific PNT Meeting*, pages 694–701.

Barrau, A. (2015). *Non-linear state error based extended Kalman filters with applications to navigation*. PhD thesis, Mines Paristech.

Ben-Afia, A., Deambrogio, L., Salós, D., Escher, A.-C., Macabiau, C., Soulier, L., and Gay-Bellile, V. (2014). Review and classification of vision-based localisation techniques in unknown environments. *IET Radar, Sonar & Navigation*, 8(9):1059–1072.

Bortz, J. E. (1971). A new mathematical formulation for strapdown inertial navigation. *IEEE transactions on aerospace and electronic systems*, (1):61–66.

Brown, N., Kawamura, E., Bard, L., Jaffe, A., Ringelberg, W., Kannan, K., and Ippolito, C. A. (2024). Visual & inertial datasets for an evtol aircraft approach and landing scenario. In *AIAA SciTech 2024 Forum*, page 1386.

Clynch, J. R. (2002). Radius of the earth-radii used in geodesy. *Naval Postgraduate School*.

Dai, J. S. (2015). Euler–rodrigues formula variations, quaternion conjugation and intrinsic connections. *Mechanism and Machine Theory*, 92:144–152.

D'Alfonso, L., Garone, E., Muraca, P., and Pugliese, P. (2021). Camera and inertial sensor fusion for the p n p problem: algorithms and experimental results. *Machine Vision and Applications*, 32(4):90.

Farrell, J. (2008). *Aided navigation: GPS with high rate sensors*. McGraw-Hill, Inc.

Gibert, V., Plestan, F., Burlion, L., Boada-Bauxell, J., and Chriette, A. (2018). Visual estimation of deviations for the civil aircraft landing. *Control engineering practice*, 75:17–25.

Gróf, T., Bauer, P., and Watanabe, Y. (2022). Positioning of aircraft relative to unknown runway with delayed image data, airdata and inertial measurement fusion. *Control Engineering Practice*, 125:105211.

Groves, P. D. (2015). Principles of gnss, inertial, and multisensor integrated navigation systems, [book review]. *IEEE Aerospace and Electronic Systems Magazine*, 30(2):26–27.

Hartley, R. and Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press.

ICAO (2006). 8168 ops/611. aircraft operations. vol. 1. flight procedures. *International Civil Aviation Organization*.

ICAO (2008). *Performance-Based Navigation (PBN) Manual. Vol. 2. Implementing RNAV and RNP*.

ICAO (2023). *Annex 10. Aeronautical Telecommunications. Vol. 1. Radio Navigational Aids*.

Kügler, M. E., Mumm, N. C., Holzapfel, F., Schwithal, A., and Angermann, M. (2019). Vision-augmented automatic landing of a general aviation fly-by-wire demonstrator. In *AIAA Scitech 2019 Forum*, page 1641.

Sola, J. (2017). Quaternion kinematics for the error-state kalman filter. *arXiv preprint arXiv:1711.02508*.

Tonhäuser, C., Schwithal, A., Wolkow, S., Angermann, M., and Hecker, P. (2015). Integrity concept for image-based automated landing systems. In *Proceedings of the ION 2015 Pacific PNT Meeting*, pages 733–747.

Vezinet, J., Escher, A.-C., Guillet, A., and Macabiau, C. (2013). State of the art of image-aided navigation techniques for aircraft approach and landing. In *Proceedings of the 2013 International Technical Meeting of The Institute of Navigation*, pages 473–607.

Watanabe, Y., Manecy, A., Hiba, A., Nagai, S., and Aoki, S. (2019). Vision-integrated navigation system for aircraft final approach in case of gnss/sbas or ils failures. In *AIAA Scitech 2019 Forum*, page 0113.