

Article

Analysis of UTM Tracking Performance for Conformance Monitoring via Hybrid SITL Monte Carlo Methods

Wei Dai ^{1,2}, Zhi Hao Quek ², Bizhao Pang ^{1,2} and Mir Feroskhan ^{1,*}

¹ School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore; wei.dai@ntu.edu.sg (W.D.); bizhao001@e.ntu.edu.sg (B.P.)

² Air Traffic Management Research Institute, Nanyang Technological University, Singapore 637460, Singapore

* Correspondence: mir.feroskhan@ntu.edu.sg

Abstract: Conformance monitoring supports UTM safety by observing if unmanned aircraft (UA) are adhering to declared operational intent. As a supporting system, robust cooperative tracking is critical. Nevertheless, tracking systems for UAS traffic management (UTM) are in an early stage and under-standardized, and existing literature hardly addresses the problem. To bridge this gap, this study aims to probabilistically evaluate the impact of the change in tracking performances on the effectiveness of conformance monitoring. We propose a Monte Carlo simulation-based method. To ensure a realistic simulation environment, we use a hybrid software-in-the-loop (SITL) scheme. The major uncertainties contributing to the stochastic evaluation are measured separately and are integrated into the final Monte Carlo simulation. Latency tests were conducted to assess the performance of different communication technologies for cooperative tracking. Flight technical error generation via SITL simulations and navigational system error generation based on flight experiments were employed to model UA trajectory uncertainty. Based on these tests, further Monte Carlo simulations were used to study the overall impacts of various tracking key performance indicators in UTM conformance monitoring. Results suggest that the extrapolation of UA position enables quicker non-conformance detection, but introduces greater variability in detection delay, and exacerbates the incidence of nuisance alerts and missed detections, particularly when latencies are high and velocity errors are severe. Recommendations for UA position update rates of ≥ 1 Hz remain consistent with previous studies, as investments in increasing the update rate do not lead to corresponding improvements in conformance monitoring performance according to simulation results.

Keywords: UTM; Monte Carlo; conformance monitoring; software-in-the-loop; simulation; Remote ID; tracking



Citation: Dai, W.; Quek, Z.H.; Pang, B.; Feroskhan, M. Analysis of UTM Tracking Performance for Conformance Monitoring via Hybrid SITL Monte Carlo Methods. *Drones* **2023**, *7*, 597. <https://doi.org/10.3390/drones7100597>

Academic Editor: Yu Wu, Liguu Sun and Pablo Rodríguez-González

Received: 22 August 2023

Revised: 19 September 2023

Accepted: 20 September 2023

Published: 22 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The recent proliferation of unmanned aircraft (UA), or “drones”, has resulted in a challenge for the safe management of air traffic operations. UAS traffic management (UTM) is concerned with the safe operation of UA in a given airspace; in particular, it is envisaged that UTM will allow for large-scale visual line-of-sight (VLOS) and beyond visual line-of-sight (BVLOS) operations [1].

The safety performance that the future UTM could achieve is the most crucial factor enabling UAS to access the national airspace system (NAS), which has drawn growing attention in the research community as well as the industry [2]. Strategic deconfliction is one of the functions that contribute to the safety assurance of UTM, which can be assessed through probabilistic modeling methods like the Monte Carlo simulation [3]. Relevant studies include a conflict-free flight planning algorithm design [4] and tactical conflict detection and resolution [5]. However, these studies have not discussed the infrastructure performances for UTM and their relationship with the expected traffic operational performances. Considering the air navigation service performance for UTM, existing studies

have considered the navigational accuracy in the urban area [6,7] and its consequent effect on the separation minima [8]. Despite the aforementioned studies, the existing research on UTM as an emerging mobility type is still limited.

Maintaining the conspicuity of (oftentimes small) UA in relatively large airspace presents a challenge for UTM. Unsurprisingly, the tracking service plays a safety-critical role in a functional UTM system [9]. Researchers have worked on the enhancement of UTM tracking reliability [10]. However, how the tracking performance influences higher-level traffic safety objectives has not been discussed. In particular, the ubiquity of drones and the correspondingly high operational density potential, coupled with their small size, favor cooperative tracking (e.g., UA and/or operator is required to actively participate in sharing flight information) over non-cooperative tracking (a third-party actively scans the airspace to identify and locate airborne UA).

A cooperative tracking service may be described by various tracking key performance indicators (KPIs). These KPIs may be affected by sensor limitations (e.g., GNSS accuracy), or may be due to communication-related challenges and ground infrastructure. Examples of the latter include communications latency, transmission availability, coverage, UA update rate, and support for trajectory prediction methods (like extrapolation) [11,12]. Such limitations affect higher-level UTM functions like conformance monitoring (CM) and tactical conflict detection, which are all primary and safety-critical UTM services. Therefore standardizing the minimum requirements for cooperative UTM tracking service becomes crucial.

Existing studies aim to establish requirements for the tracking service based on tactical conflict detection. These studies rely on the Monte Carlo simulation, which is a common technique in air traffic management [13–20]. As an extension, this article aims to assess tracking requirements from a conformance monitoring (CM) perspective.

CM in the civil aviation context, as summarized by NASA [21], is a function developed to check on the adherence between a given flight and its declared flight plan. It can be either reactive or predictive, and its performance is characterized by false alarm rate, missed alert rate, and detection delay. Given a declared flight intent, CM has been formulated as a residual-based fault-detection problem by researchers from MIT Lincoln Lab [22,23]. CM algorithms feature a trade-off between time-to-detection and false alarm probability but generally benefit from a high data update rate. Additionally, studies have indicated that flight patterns influence CM performance. In particular, transitions pose a challenge in CM due to turning dynamics and ambiguity in turn initiation time. As an alternative to the fault-detection formulation, Lee et al. developed a Bayesian approach for CM [24], in part to address ambiguity in transition time. At every timestep, the algorithm estimates conformance probability across a range of trajectory change times (TCTs), and calculates their weighted sum. When this sum is less than a defined threshold, the flight is deemed as non-conforming.

In the UTM context, ASTM introduced a simple framework for CM. The F3548-21 standard [25] defines conformance with respect to a declared operational intent (OI). Under the OI framework, operations can be classified into three main states, namely ‘activated’, ‘non-conforming’, and ‘contingent’. The latter two states are considered non-conforming. OIs are defined by a four-dimensional OI volume (OIV)—3D geometries describing the locations of aircraft operations and the start and end times. These OIVs must be pre-approved by a relevant authority; as part of the approval process, strategic deconfliction is performed. Consequently, non-conforming UA operations are not strategically deconflicted and hence risk infringements between non-conforming UA and other aircraft operating in nearby airspace or OIVs [3]. CM, in this context, helps to check if the UA is operating within its respective OIV. Various mission profiles may be supported by the OI concept by variations of the size, geometry, and duration of the OIVs. The performance of the tracking system in this CM application, as well as the experimental analysis of latency for the supporting communications technologies, form the core subjects of this article.

Based on the above discussions, this work studies the relationship between tracking performances and CM effectiveness. Such a relationship is presented as the change in tracking performances affecting the CM timeliness and success rate, which occurs in a stochastic manner. To achieve this probabilistic reasoning, we propose a Monte Carlo simulation-based method. The contributions of this study are threefold, which are listed below:

- (1) A hybrid SITL Monte Carlo simulation scheme that supports the probabilistic performance evaluation of UTM flights with multiple operational uncertainties;
- (2) A new formulation of recall and precision that meets the requirement of event detection in continuous time–space;
- (3) Analytical results on the relationship between tracking performances and CM effectiveness that support the decision-making of UTM stakeholders in the deployment and standardization of the tracking service.

2. Materials and Methods

2.1. Methodology Overview

The aim of this study is to estimate the probability distribution in a complex uncertain environment. We use the Monte Carlo simulation to solve the problem, where one of the key issues in guaranteeing the reliability of the result is the modeling of main uncertain factors. Two major systems involved in the CM procedure are the tracking system and the UA. The uncertainties of the tracking system are characterized by the inaccuracy of the detected aircraft location, which is affected by major tracking KPIs, including the accuracy, update rate, availability, latency, and extrapolation function of the system. Amongst these KPIs, the communication latency using the standardized UTM remote ID protocol is less discussed in the literature. To support the Monte Carlo simulation, we conducted experiments to measure the distribution of latency. Details of the experiment and discussions will be elaborated in the following part of this paper. The uncertainties of the UA are presented as trajectory deviations. Such deviations include flight technical error (FTE) and navigation system error (NSE). A hybrid SITL approach is developed to model the trajectory with deviations induced by FTE and NSE. The overall workflow of this study is illustrated in Figure 1.

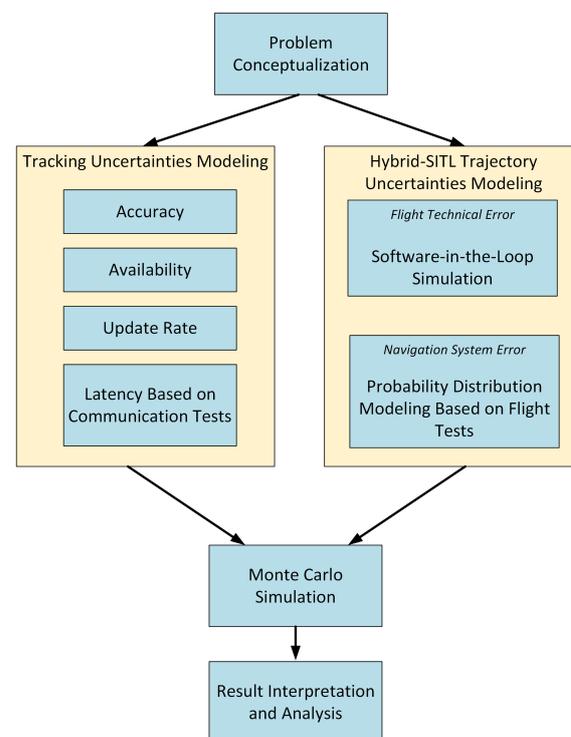


Figure 1. Overall workflow of the evaluation of the impact of tracking performances on conformance monitoring.

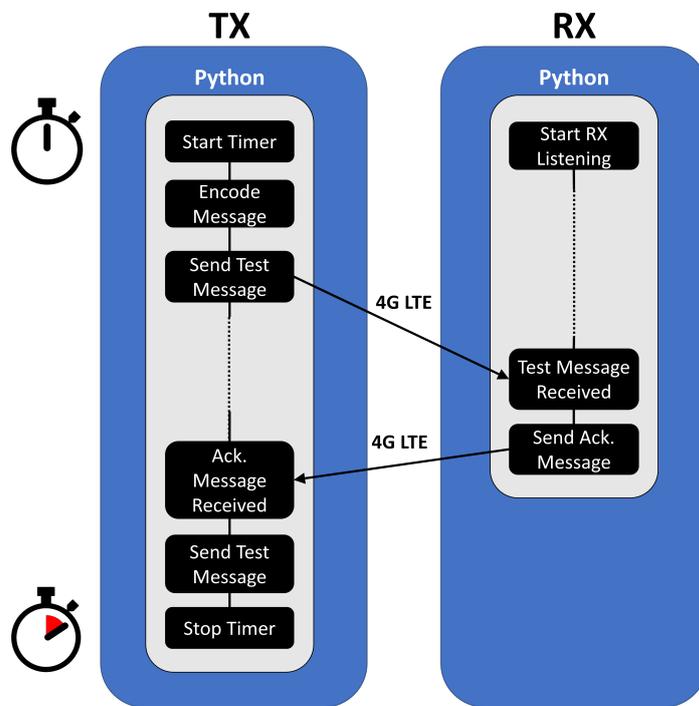
2.2. Latency Experiments

Experience with vehicle teleoperations reveals concerns between communication-related delays and vehicle safety [26–28]. While the overall UTM architecture differs from vehicle teleoperations, in that it does not necessarily involve low-level vehicle dynamics control (and may not, potentially, even direct human control), it nevertheless features collision avoidance and remote monitoring, and thus motivates further analysis into tracking KPIs.

The underlying communications architecture, technologies, and relevant supporting standards strongly influence such tracking KPIs. With regard to digital UA conspicuity and identification, the FAA Remote ID and ASTM F3411-22a standard are at the forefront in defining UA aircraft identification requirements, and can be viewed as early implementations of the UTM tracking service [29]. The standard supports two broad communications architectures, namely broadcast Remote ID (RID) and network RID. Broadcast RID is a non-directed communications protocol, wherein a target recipient is not specified, and the UA simply share its location via wireless broadcast, using common wireless technologies, such as Bluetooth and WiFi. By contrast, network RID requires the UA to transmit to a target UTM service supplier (USS) or ground station. Supported technologies include cellular connections and internet-based routing protocols.

As highlighted, near-term implementations of cooperative tracking are likely to rely on existing technical standards and technologies to accomplish UA tracking. The ASTM F3411-22a standard supports both broadcast and network architectures under the FAA Remote ID requirements. For broadcast RID, Bluetooth legacy advertising, Bluetooth 5.x extended advertising, WiFi Beacon, and WiFi NAN are the supported means of transmission; for network RID, existing cellular infrastructure (such as 4G LTE or 5G) is allowed.

In lieu of the above, round-trip latency experiments were performed with two communications technologies—one using Bluetooth and another with 4G LTE. Network RID tests were further divided into moving and static tests. (Due to the use of Ethernet cables for Bluetooth experiments, tests involving large relative motion between transmitter and receiver were not feasible). Figure 2 shows testing schemes for both Bluetooth and 4G LTE.



(a)

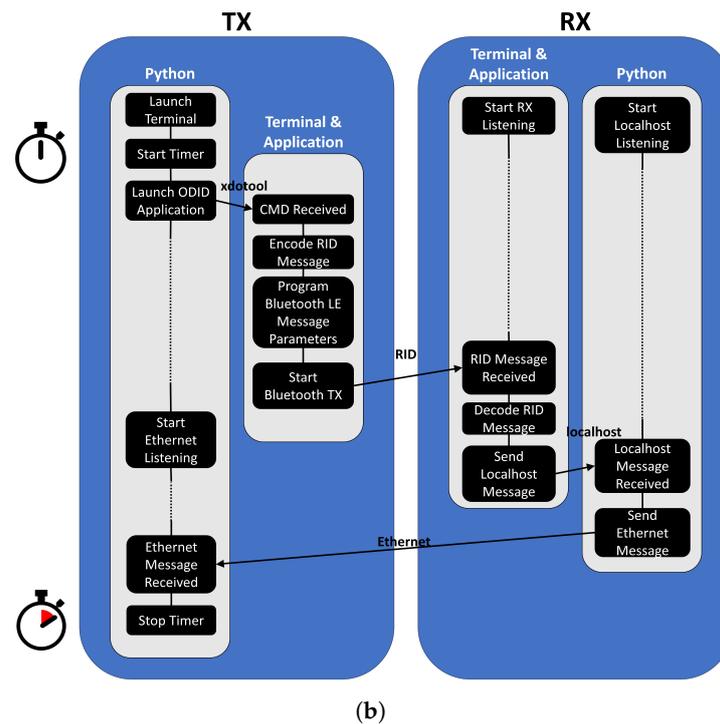


Figure 2. (a) 4G LTE latency testing flowchart. (b) Remote ID Bluetooth Legacy Advertising latency testing flowchart.

2.2.1. 4G LTE Latency Tests

Various factors may potentially influence cellular network performance and may affect the measured tracking KPIs. An international, cross-demographic study by Walelgne et al. found that cellular traffic typically peaked at night [30], and was greatest on Mondays and Tuesdays. As such, latency tests were performed at night (from 2130 h to 2330 h) on a Monday evening to simulate a worst-case cellular traffic scenario. Furthermore, tests were performed near a residential area, where high user loads may be expected during the evening. The location and groundspeed trace for the LTE latency tests may be found in Figure 3; for moving tests, groundspeeds of up to 70 km/h were reached.

Additionally, cell-tower antennae may be angled in a manner that affects the signal strength at altitude [31]. Ideally, latency tests should incorporate measurements performed at the expected UA operation height; however, the time of day and proximity to residential units resulted in operational limitations that only allowed ground-based testing.

Tests were performed with a Huawei E8372 USB LTE modem and an off-the-shelf SIM card from a local telecom provider. Messages between the mobile client (transmitter) and server (receiver) were sent via the user datagram protocol (UDP), a connectionless internet protocol that does not provide retransmissions. The simplicity of the protocol and the lack of retransmissions reduce network overhead and bandwidth requirements and offer low latency.

2.2.2. Bluetooth Legacy Advertising Remote ID Tests

Testing of a sample broadcast tracking service was performed with a prototypical remote ID-compliant software, from the OpenDroneID project (see <https://www.opendroneid.org/> (accessed on 9 April 2023) and <https://github.com/opendroneid/transmitter-linux> (Version 1.0)). A pair of Nordic Semiconductor nRF52840 Bluetooth dongles were used as the receiver and transmitter; due to software limitations at the time of writing, Bluetooth extended advertising was not enabled for testing.

To measure potential Bluetooth Remote ID message latency, the transmitting computer would measure the time taken for a receiving computer to send an acknowledgment message; this message was sent via the Ethernet to minimize the additional communica-

tion time. While such a protocol introduces an additional delay, a similar penalty exists for two-way LTE communications, and a centralized conformance monitoring service (e.g., provided for by a USS) would also require additional routing of flight information from Bluetooth or WiFi receivers and, therefore, incur such a time delay.

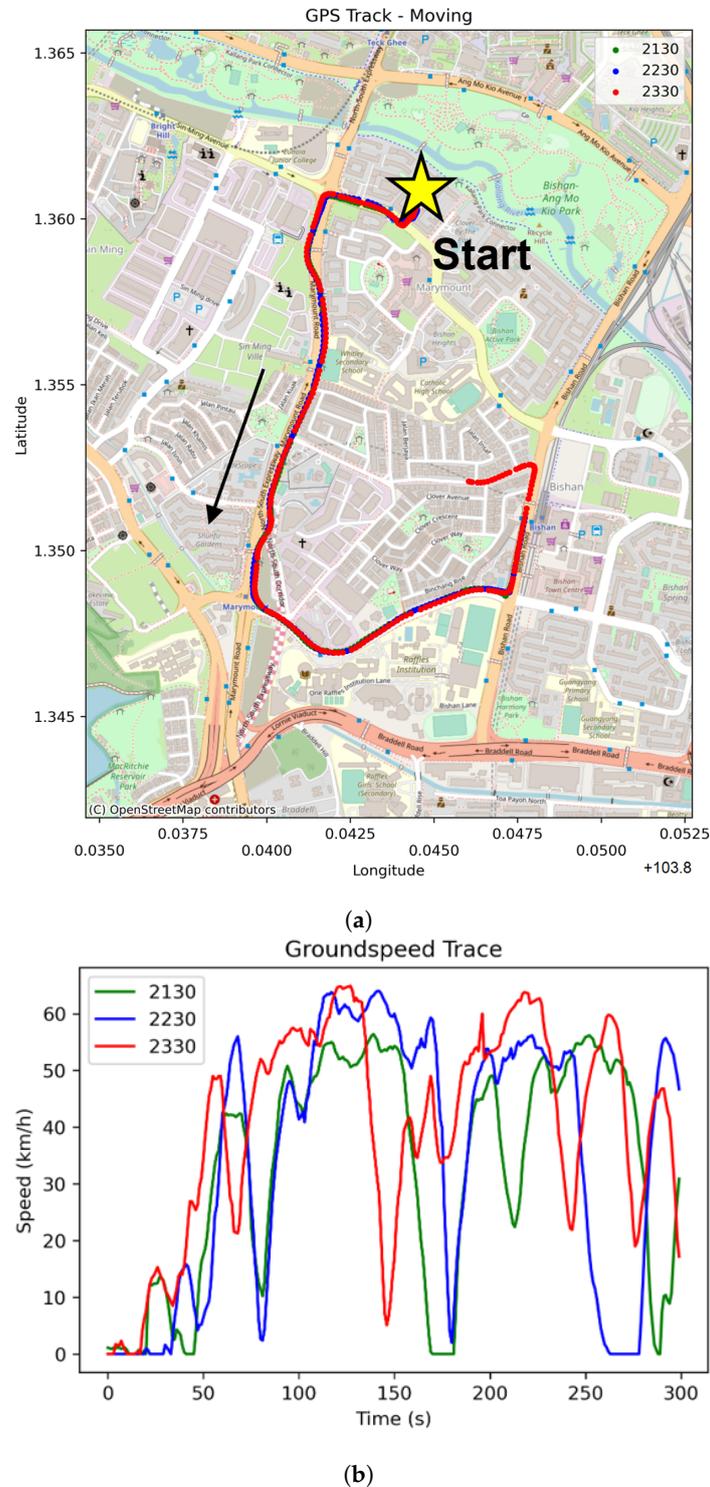


Figure 3. The 4G LTE latency experiments and map of testing locations. (a) Map of the 4G LTE test area, with ground track shown for different data collection runs. Static tests were performed at the start location. The arrow indicates the direction of travel for moving tests (counterclockwise); (b) ground speed traces for moving latency tests.

2.3. Hybrid SITL Simulation Approach and Simulation Parameters

The results from the latency tests are used to fit statistical latency models, which are then employed in Monte Carlo simulations. These simulations generate “ground truth” and “tracked” flight trajectories, which can then be compared to understand the tracking efficacy and its safety impact.

Currently, the industry of business-applicable drones is under-standardized. In the simulation, there is no specific drone model that is representative of all of its kind. In this study, we used the software-in-the-loop simulation by employing the PX4 autopilot, one of the most popular UAVs, and its controller model, which is widely used in both industry and academia, to simulate the performance of a generic quad-rotorcraft. To ensure the validity of the kinematic model, we set realistic velocity and acceleration limits in the model, based on data collected in our flight experiment [32], open-source flight data [33], and recommended settings in the PX4 autopilot flight controller.

To generate the aforementioned flight trajectories, a hybrid software-in-the-loop and post-processing simulation approach is adopted. A set of predefined waypoints are uploaded into a software-in-the-loop simulation environment based on PX4 (PX4 version 1.13.0), with flight dynamics modeled in Gazebo. Variable wind speeds and cruise speeds are further selected. A simulated flight trajectory is recorded, and additional navigation error is introduced through post-processing; this post-processing step is performed to obtain both ground truth and tracked trajectories. Additional rounding errors and velocity errors may, at this point, also be introduced. This two-step approach essentially allows separate generation of the flight technical error (FTE) and navigational system error (NSE). The probability distribution of the NSE is also based on the measurement of flight experiments, where GNSS errors are studied in the urban environment [7].

In a dynamic system, the line between NSE and FTE may be blurred. Consider an aircraft flying with 0m NSE and FTE 0m at t_0 ; neglecting PDE, TSE would be identically 0m. At time $t > t_0$, the NSE might have drifted (due to different GNSS errors, or the accumulation of dead reckoning errors). If, however, the aircraft’s track angle matches the desired flight path and no input by the flight controller has been commanded, TSE remains at 0m; this implies a non-zero FTE. Thus, NSE, FTE, and TSE are mutually dependent in a temporal fashion. Nevertheless, separate generation of FTE and NSE is advantageous in that it allows bypassing the automatic flight controller navigational sensor error protections (PX4 SITL supports GNSS error injection but prevents arming of the aircraft if it detects large GNSS errors) in a simulation platform, and facilitates faster dynamical flight simulations, thereby allowing for a more extensive study of the parameter space.

2.4. Monte Carlo Simulation

Two main cases were studied to analyze the safety performance of a tracking system in a conformance monitoring context. As per the introduction, the conformance monitoring application is envisaged to be similar in configuration to that of the ASTM F3548-21, in that conformance is defined by means of the aircraft position with respect to a filed OIV. Accordingly, the studied cases are:

- Case 1: Contingent operations and total delay time: In situations where the UA enters contingent operations, e.g., due to mechanical faults, operator errors, unexpected weather, etc., a prolonged UA operation outside of its approved OIV will result in eventual detection by a conformance monitoring system. Due to the prolonged nature of such operations, false alerts and nuisance alerts are less important; rather, quick detection for operators to take mitigating options is preferred. To quantify this, the expected total delay time between the start of a non-conforming event and its detection by the conformance monitoring system is measured for a given set of tracking performance parameters. This case is modeled in the simulation environment by flying a UA toward, and beyond, an OIV boundary at various selected cruise speeds. The time difference between the UA’s true position leaving the boundary

and its tracked position leaving the boundary (i.e., detected by the conformance monitoring system) is the total delay time. This case is illustrated in Figure 4a.

- Case 2: Nominal operations and precision and recall: In nominal operations, the UA may periodically drift in and out of the approved OIV due to a combination of NSE, FTE, and PDE factors; under the ASTM F3548-21 standard, occasional non-conformance is permitted (up to 5% of total flight time). In such a case, false alerts and missed detections are important in reducing nuisances to operators and USS/regulators, and in notifying operators of poor UA conformance. To model this scenario, a square-shaped trajectory is flown by the simulated UA in autonomous mode. Each side of the square represents a flight “leg”. A corresponding square-shaped OIV with a hollow center (when viewed from above) is constructed; each flight “leg” has a boundary width that allows for minor track deviations. Continuous-time extensions of precision and recall are metrics defined to quantify the frequency of nuisance alerts and missed detections of such OIV non-conformance. This case is illustrated in Figure 4b.

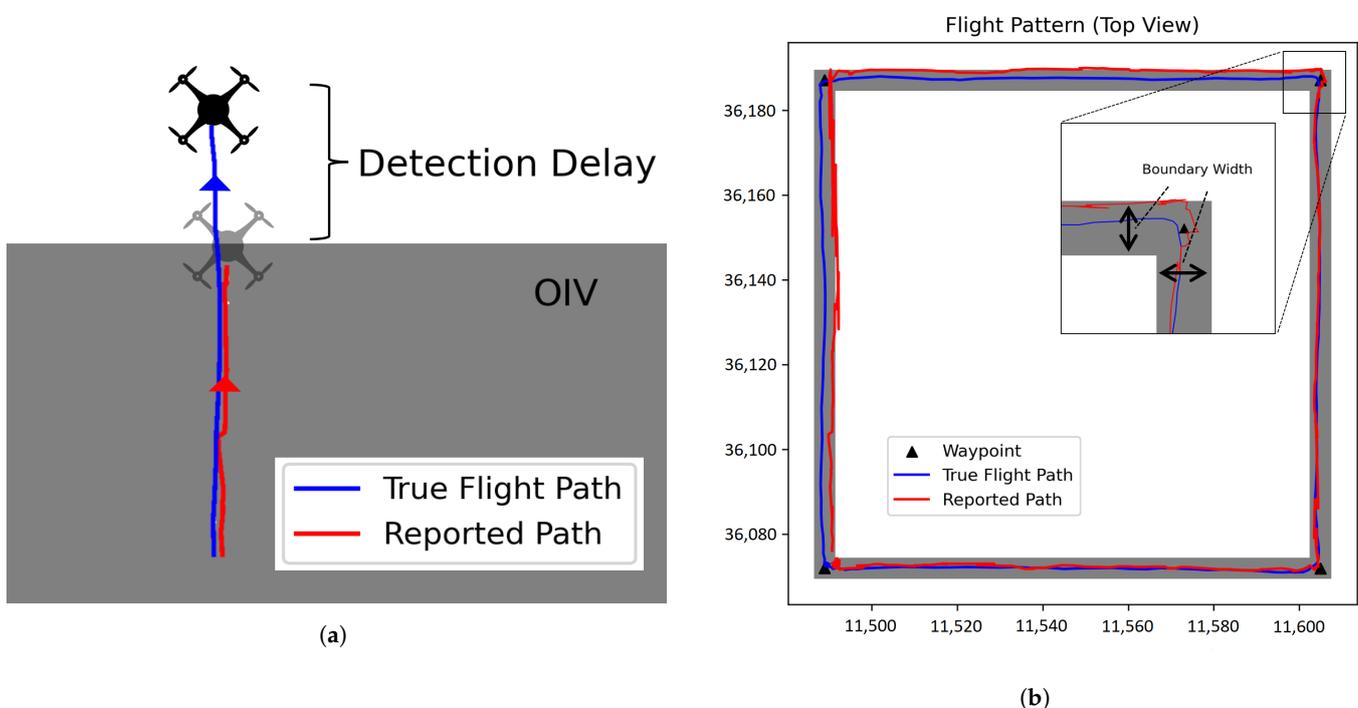


Figure 4. Overview of Cases 1 and 2. (a) Top view of the Case 1 scenario, with UA flying beyond its declared OIV; (b) top view of the Case 2 scenario, with the square flight pattern and boundary width indicated.

A common set of simulation parameters (shown in Table 1) were used for both Cases 1 and 2. Wind directions and cruise speeds were averaged over for each tracking system configuration. Similar to [12], three extrapolation settings were supported due to their possible influence on detection delay:

- ExtOff: No extrapolation performed.
- ExtSync: Extrapolation enabled, with extrapolation duration based on the GNSS time for both UA and the tracking server; susceptible to the clock synchronization error with the modeled upper bound of 0.2 s; mitigates communication latency-induced position errors.
- ExtNoSync: Extrapolation enabled, with extrapolation duration based on the reception time by tracking the server only. Susceptible to communication latency-induced position errors.

Table 1. Parameters for Monte Carlo sensitivity analysis.

Parameter	Values	Units	Remarks
Wind Direction	Cardinal and ordinal directions {North, Northeast, ..., West, Northwest}	-	45° between different wind directions
Cruise Speeds	{2.5, 5.0, 7.5, 10.0}	ms ⁻¹	-
UA Update Rate (UR_{UAS})	{1, 2.5}	Hz	Based on recommendations from prior work ¹
Server Update and Extrapolation Rate (UR_{EXT})	{5}	Hz	Based on recommendations from prior work ²
Extrapolation Modes	{ExtOff, ExtSync, ExtNoSync}	-	-
Connection Type	{Bluetooth LE (BLE), 4G LTE}	-	Determines latency model; based on experimental results
UA Internal GNSS Position Error (NSE)	{3}	m	Rayleigh distribution ^{3,4}
Tracking System Position Error	Internal/Integrated: {3}	m	Errors are equivalent with NSE
	Standalone: {1, 3, 10}	m	Rayleigh distribution ^{3,4}
Tracking System Velocity Error	{0.3, 1, 3}	ms ⁻¹	Rayleigh distribution ^{3,4}
Availability	{80}	%	ASTM F3411-22a allows for minimum (networked) availability of 20%

¹ $UR_{UAS} \geq 1$ Hz, see [12]. ² $UR_{EXT} \geq 5$ Hz, see [12]. ³ Values quoted represent the 95th percentile. ⁴ Errors are autocorrelated at each timestep; see [11].

3. Results

3.1. Latency Measurements

Figure 5 shows the histograms of round-trip latency measurements from both Bluetooth RID and proprietary 4G LTE communication tests, along with their best Gaussian, Weibull, and Fisk fits. Of the three distribution types, the Fisk distributions were selected for subsequent Monte Carlo communication-delay modeling. The Fisk PDF is parameterized by

$$f(x) = \begin{cases} \left(\frac{\beta}{\alpha}\right) \frac{\left(\frac{x-x_0}{\beta}\right)^{\beta-1}}{\left(1+\left(\frac{x-x_0}{\alpha}\right)^\beta\right)^2}, & x \geq x_0 \\ 0, & x < x_0 \end{cases} \quad (1)$$

Additionally, a comparison between moving and static 4G LTE round-trip latencies is shown in Figure 6. The two distributions are broadly comparable, indicating no significant difference in latency for the tested velocity range.

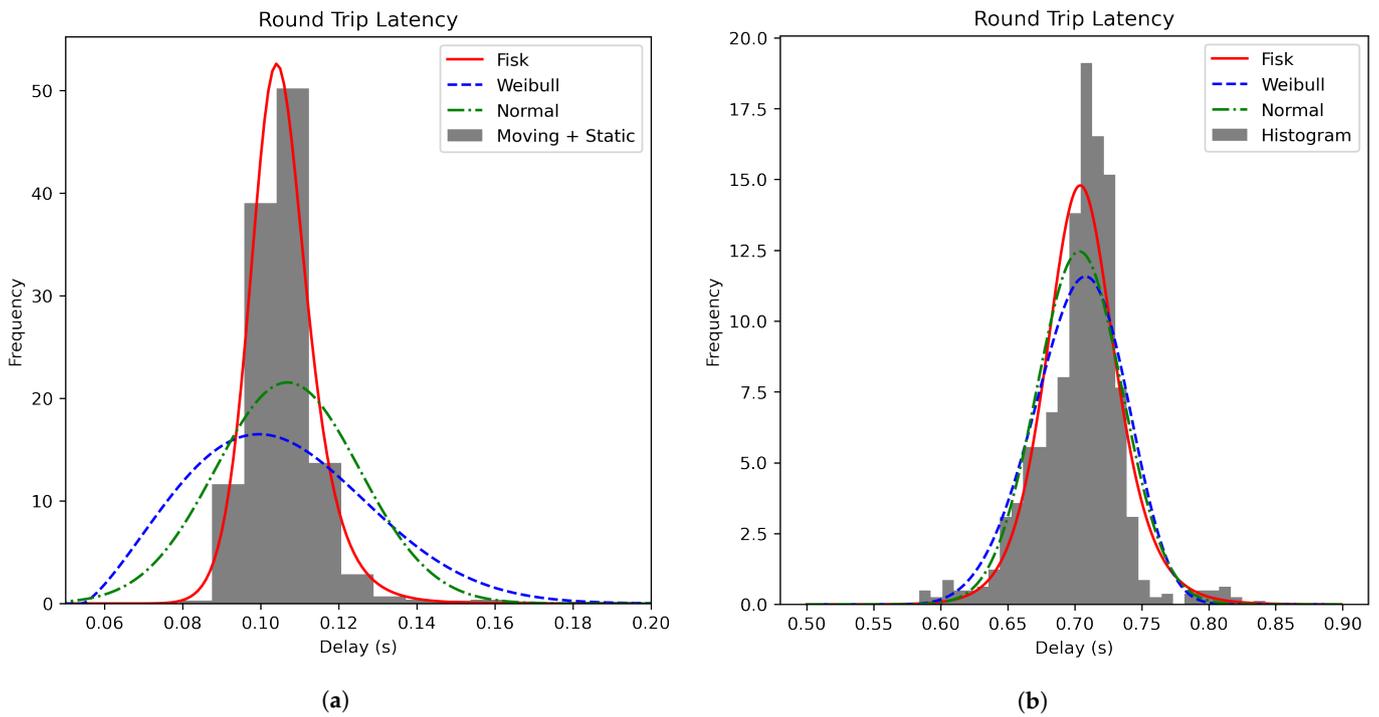


Figure 5. Round-trip latencies measured using various communications architectures. (a) 4G LTE latency (combined moving + static); (b) BLE latency.

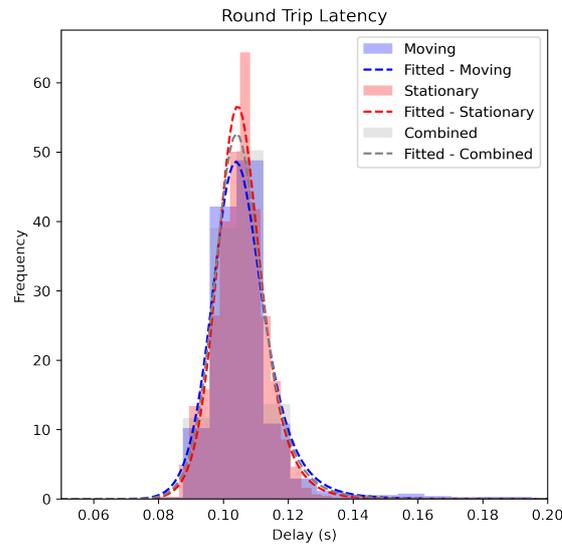


Figure 6. The 4G LTE latency (moving versus static).

A brief summary of the measured and fitted round-trip latency distributions is shown in Table 2.

Table 2. Summary of measured and fitted latency values.

Parameter	4G LTE	BLE
Average (measured)	0.1069 s	0.7035 s
Average (Fisk best fit)	0.1057 s	0.7056 s
Fisk α	0.0520	0.722
Fisk β	10.8	42.7
Fisk x_0	0.0530	-0.0174

3.2. Monte Carlo Case 1: Non-Conformance Detection Total Delay Time

Figure 7 shows the total delay time distributions for Case 1’s non-conformance events with various extrapolation modes, navigational velocity error (NSE) settings, connection types, and UA update rates. Tracking is performed by standalone units with a position error that is modeled as independent from the UA’s internal navigation position error (NSE). Mean values for each tracking system configuration are indicated below the violin plots. A few salient observations can be made: firstly, using extrapolation (particularly the synchronized mode relying on GNSS timestamps for UA position reports and with the tracking system synchronized to GNSS time) results in the lowest mean detection delay. Accordingly, the ExtOff mode yields the highest mean total detection delay time, with higher latency communication protocols, like the modeled Bluetooth ODID application with higher delay. A lower UA update rate, corresponding to more infrequent position updates, leads to longer delays as well. While switching from ExtOff to ExtSync largely compensates for high mean detection latencies, drawbacks arise in higher variability when greater navigational velocity errors are present.

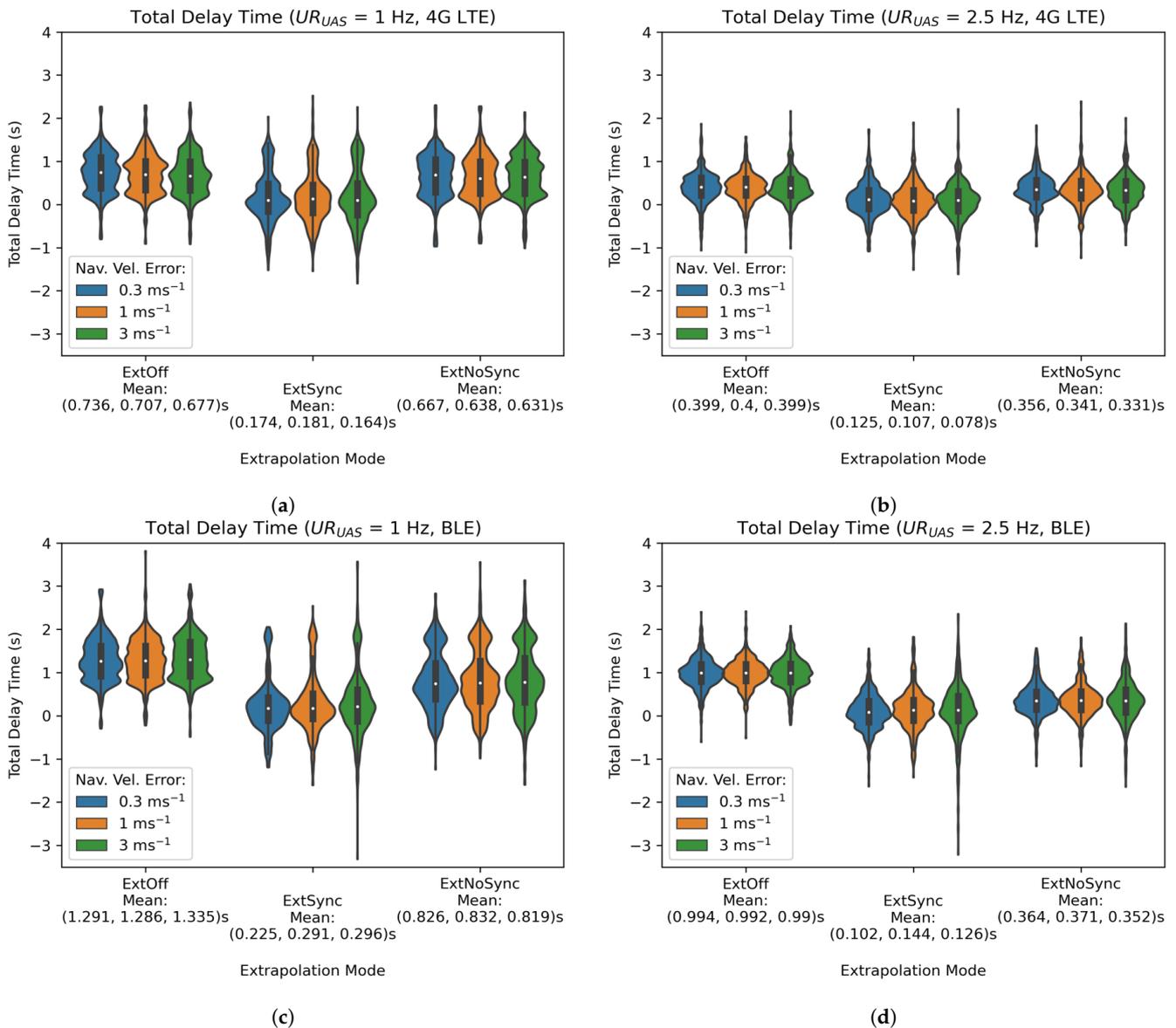


Figure 7. Total delay time violin plots, separated by the navigational velocity error (standalone tracker); (a) 4G LTE, UR_{UAS} set to 1 Hz; 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

Figure 8 similarly shows the total delay time distributions for Case 1 non-conformance events with tracking system configurations, but for different positional errors instead of velocity errors. Tracking units are once again modeled as independent. Across all conditions, a higher position error leads to higher detection delay time variability; however, average detection delays are not significantly changed. For the considered range of cruise speeds and tracking configurations, the position error contributes the most to the detection delay time.

No significant differences in the detection delay time were observed between standalone and UA-integrated position data sources.

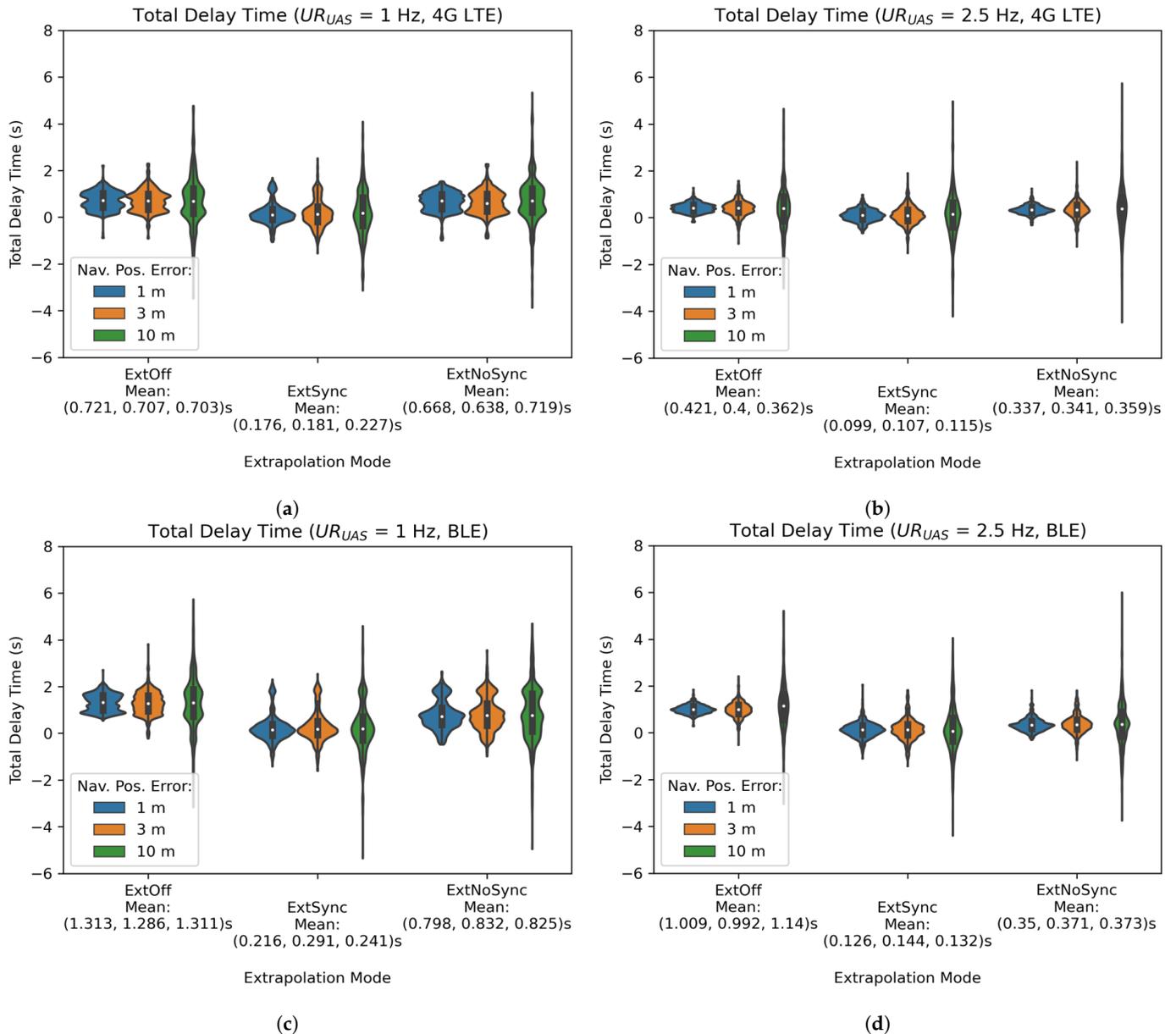


Figure 8. Total delay time violin plots, separated by the navigational position error (standalone tracker); (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

3.3. Monte Carlo Case 2: Nominal Operations and Periodic Non-Conformance

When considering nominal operations under the OIV concept, a larger cross-track width (or, generally, a larger buffer between the intended area of operations and the OIV boundary) would result in a non-conformance ratio due to regular flight path variances.

This is illustrated in Figure 9. Nevertheless, there exists a trade-off between the boundary width and operational airspace capacity; ideally, the boundary width should be minimized while maintaining a higher conformance ratio, such that more aircraft are permitted to operate near each other.

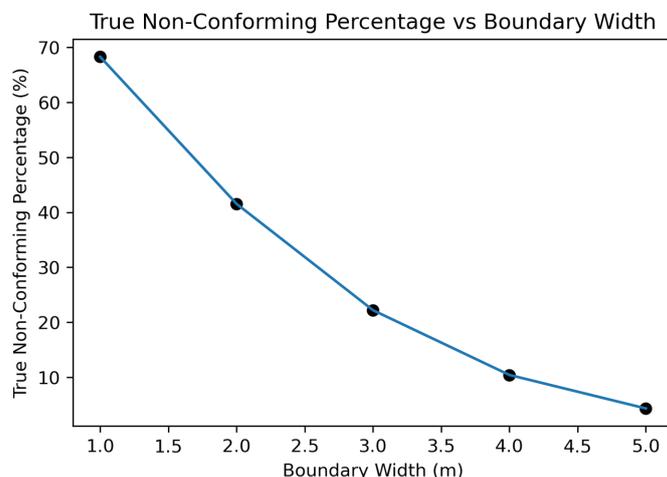


Figure 9. Reduction in the true non-conformance ratio as the OIV boundary/track width increases.

The high frequency of position reports under nominal operations means that a single non-conformance “event” is not a good indicator of the tracking performance, as even under a 5% non-conformance upper limit (permitted under ASTM F3548-21), there can be multiple position reports with UA in non-conforming positions for a single short flight. As an example, a 1 Hz UA position update rate, coupled with a flight duration of 10 min would yield expected 30 non-conforming timestamps. Thus, precision and recall metrics adopted by prior studies, such as [12], which are aimed at capturing once-per-flight events, must be adapted to the current comparison. Extensions of precision and recall to continuous-time (while these metrics are defined in a continuous-time sense as integrals, in practice, these metrics are evaluated at discrete intervals) are presented below:

$$\text{Precision} \equiv \frac{\int f(t) dt}{\int g(t + \phi) dt} \tag{2}$$

and

$$\text{Recall} \equiv \frac{\int f(t) dt}{\int h(t) dt} \tag{3}$$

where the functions f , g , and h are defined as follows:

$$f(t) \equiv h(t) \times g(t + \phi) \tag{4}$$

$$g(t) \equiv \left\{ \begin{array}{ll} 1, & \text{if reported/tracked UA position is} \\ & \text{non-conforming at time } t \\ 0, & \text{otherwise} \end{array} \right\} \tag{5}$$

$$h(t) \equiv \left\{ \begin{array}{ll} 1, & \text{if true UA position is non-} \\ & \text{conforming at time } t \\ 0, & \text{otherwise} \end{array} \right\} \tag{6}$$

Here, ϕ refers to the expected delay between a non-conformance event and its detection by the tracking system; it is estimated from the results of the previous section on a per-tracking-configuration basis. The ϕ term helps to ensure that non-conformance events

between true and tracked trajectories occur at roughly the same time and, thus, properly correspond to each other.

A perfect conformance monitoring service has both precision and recall of unity—a high precision would mean few nuisance alerts, while a high recall would correspond to few missed detections of non-conformance.

3.3.1. Precision

Figure 10 shows the differences between internal (i.e., UA-integrated) and standalone navigational position sensors for tracking use. The result shows a trend that the larger the boundary width, the lower precision outcomes are observed, which applies in all of the four groups. Based on the results, the update rate and extrapolation method do not have a significant impact on the precision of abnormal detection. The internal solution offers poorer precision, with the gap widening at larger boundary widths. The negative effect of the internal sensor solution will be further discussed in this paper.

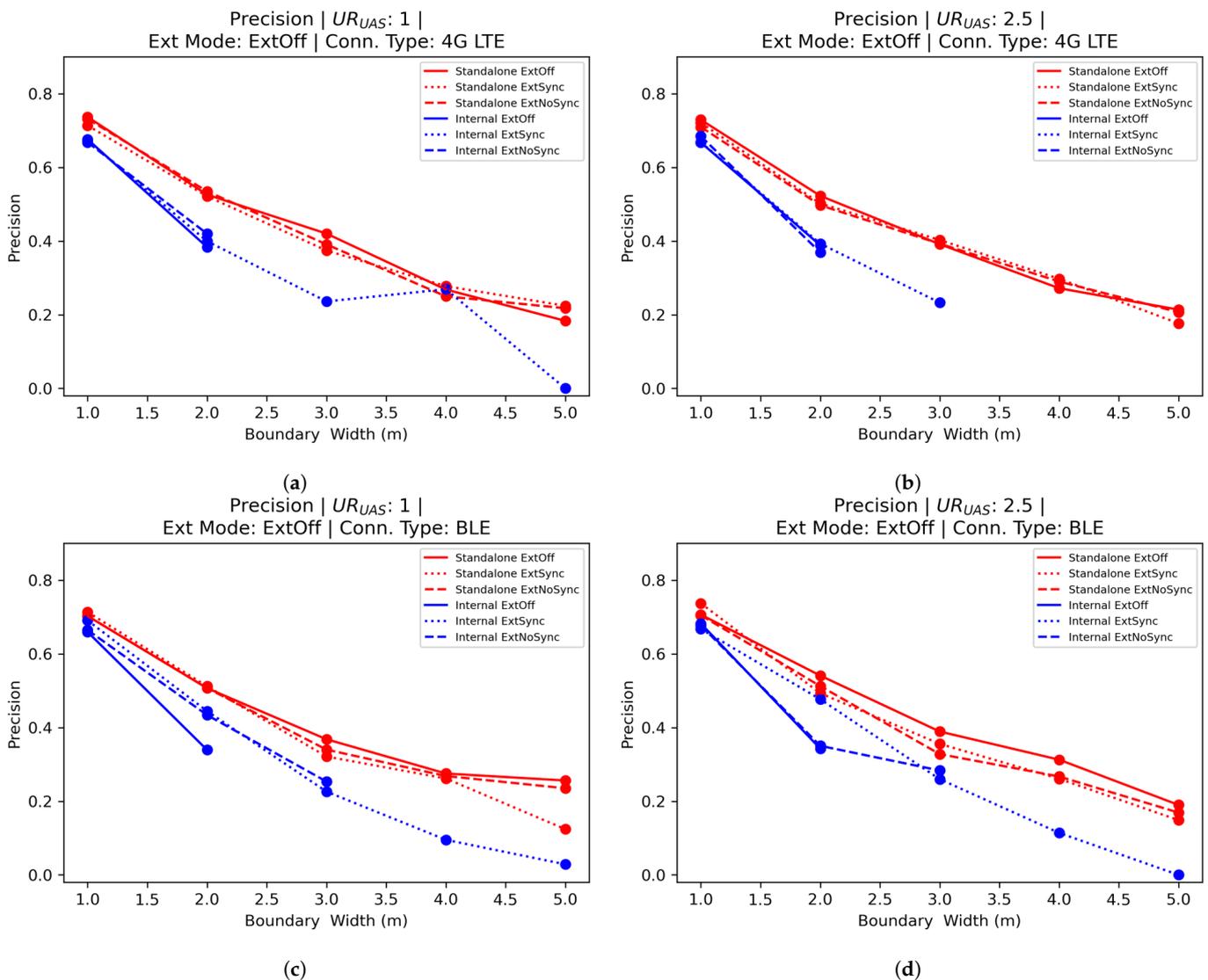


Figure 10. Precision versus boundary width curves—comparison between the integrated UA navigational position source and standalone tracking with the independent position source; (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

Figure 11, focusing on standalone tracking unit solutions with extrapolation disabled, shows a trend of decreasing precision with poorer navigational position accuracy; additionally, there is no effect of velocity error on precision. This is due to the fact that without extrapolation, velocity errors do not contribute to aircraft position estimations. Such a decrease becomes larger when the boundary width increases. Similar to the results shown in Figure 10, the update rate of the tracking system does not affect the precision performance.

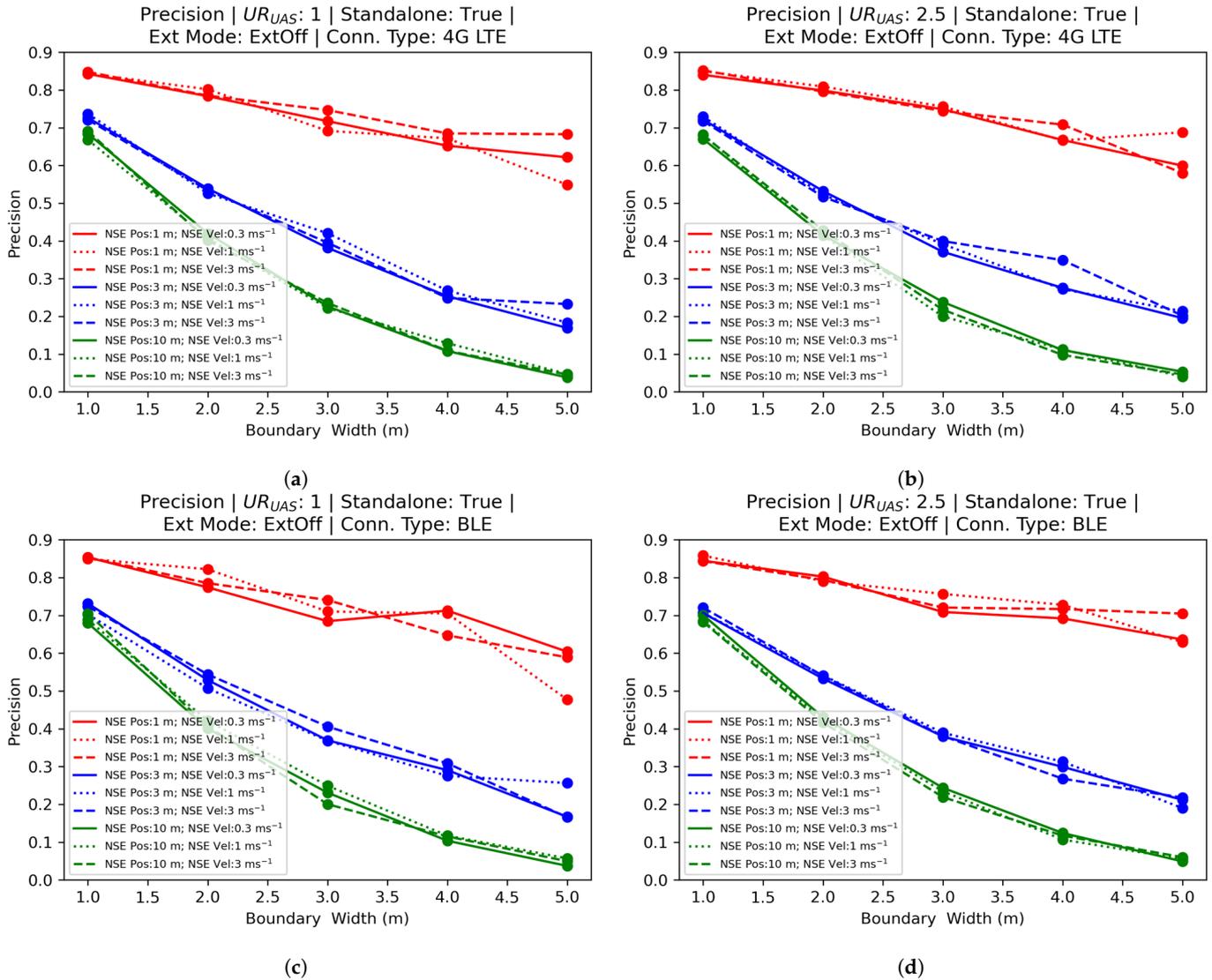


Figure 11. Precision versus boundary width curves—ExtOff; (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

Consequently, enabling extrapolation has the effect of worsening effective navigational position accuracy, due to the propagation of groundspeed, heading, and rounding-related errors; additionally, non-linear dynamics during turns and in flight-path-keeping may not be well-accounted for in the extrapolation process. A comparison may also be made between corresponding subfigures of both Figures 11 and 12—precision is on the whole lower when extrapolation is enabled. However, a closer comparison between the subfigures within Figure 12 reveals how low UA update rates and high latency exacerbate this effect since the effective “lookahead” duration the extrapolation technique compensates for is increased, leading to greater positional uncertainty and lower precision.

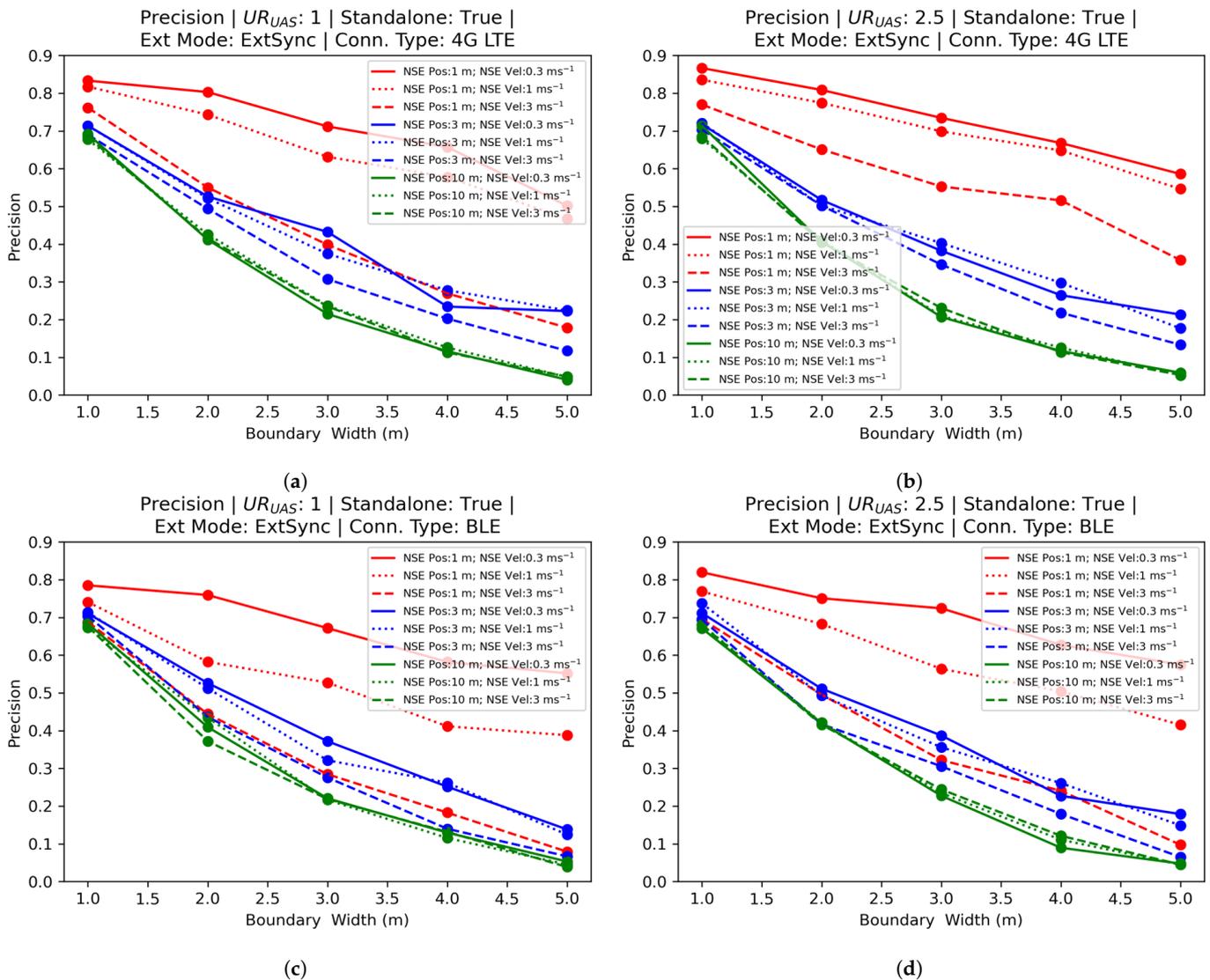


Figure 12. Precision versus boundary width curves—ExtSync; (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

A similar pattern is observed with the ExtNoSync mode, as illustrated in Figure 13). The main differences in precision between the ExtSync and ExtNoSync modes are the differences in the “look-ahead” times—ExtSync compensates for a longer latency duration and, thus, suffers more when presented with velocity errors.

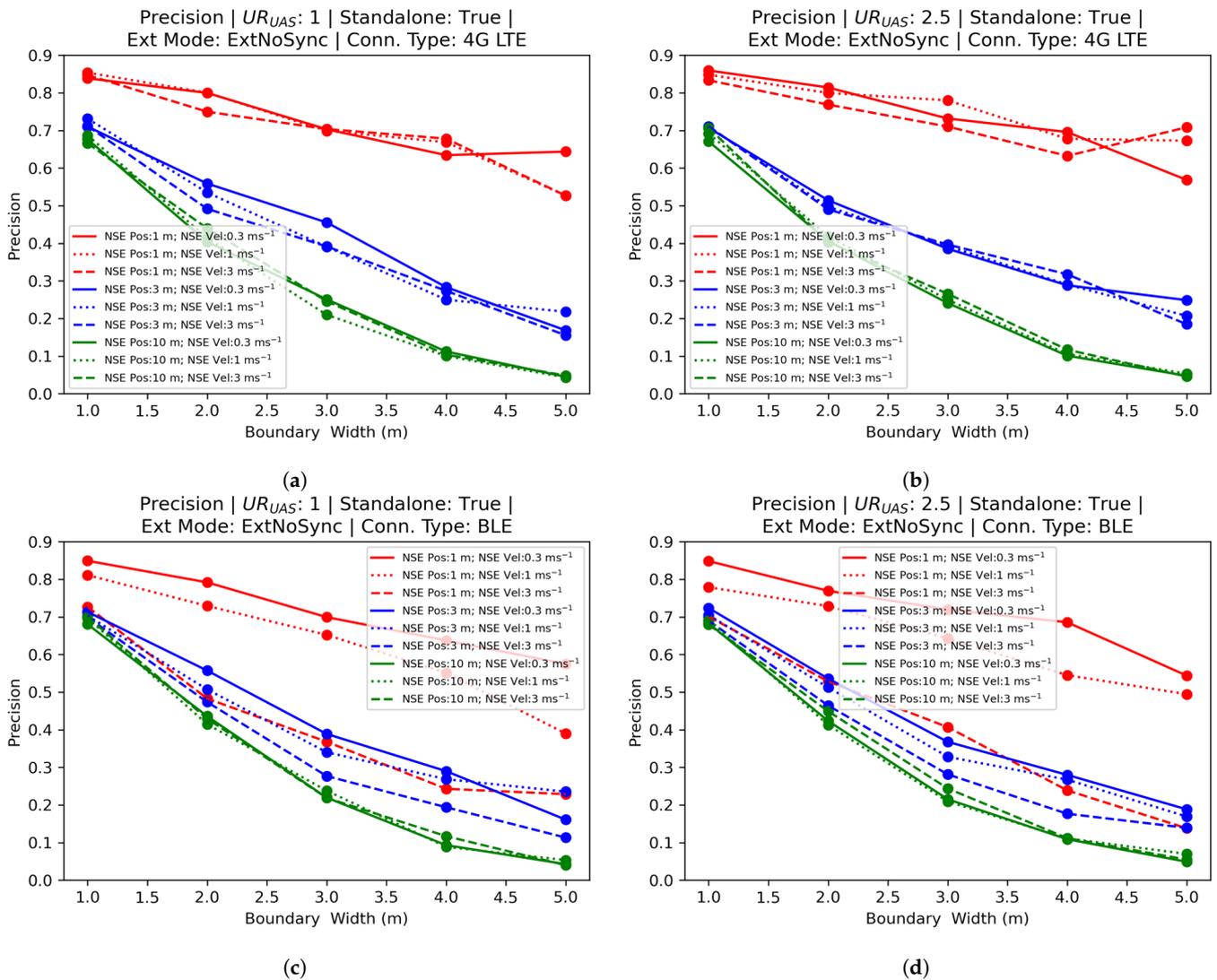


Figure 13. Precision versus boundary width curves—ExtNoSync. (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

3.3.2. Recall

As with precision, recall is significantly lower for internal tracking navigational solutions than it is with standalone ones. This is illustrated in Figure 14, and is likely due to the modeling assumptions that the total system error (TSE) is greater than the flight technical error (FTE), and that the FTE and navigational system error (NSE) can be decoupled. Since the navigational error reported by the UA’s flight controller to the tracking service is equivalent to FTE for integrated solutions (in the absence of extrapolation and rounding errors, NSE is transparent to both the flight controller and tracking service), and because TSE is modeled to be greater than FTE, the conformance monitoring system (which uses tracking information) is unable to pick up non-conforming events at the tail end of the TSE distribution. Overall, this leads to poor recall for the integrated tracking solutions. Note that this does not definitively lead to the conclusion that integrated tracking solutions are poorer than their standalone counterparts; however, this does warrant further study into interactions between various TSE components and validating modeling assumptions.

Figure 15 presents the recall trend with an increasing boundary width in the no-extrapolation situation. The measured recall shows a decreasing trend. There is no significant difference observed between LTE and BLE, or between different update rate cases. A high position error leads to a lower recall performance while a high velocity error does not.

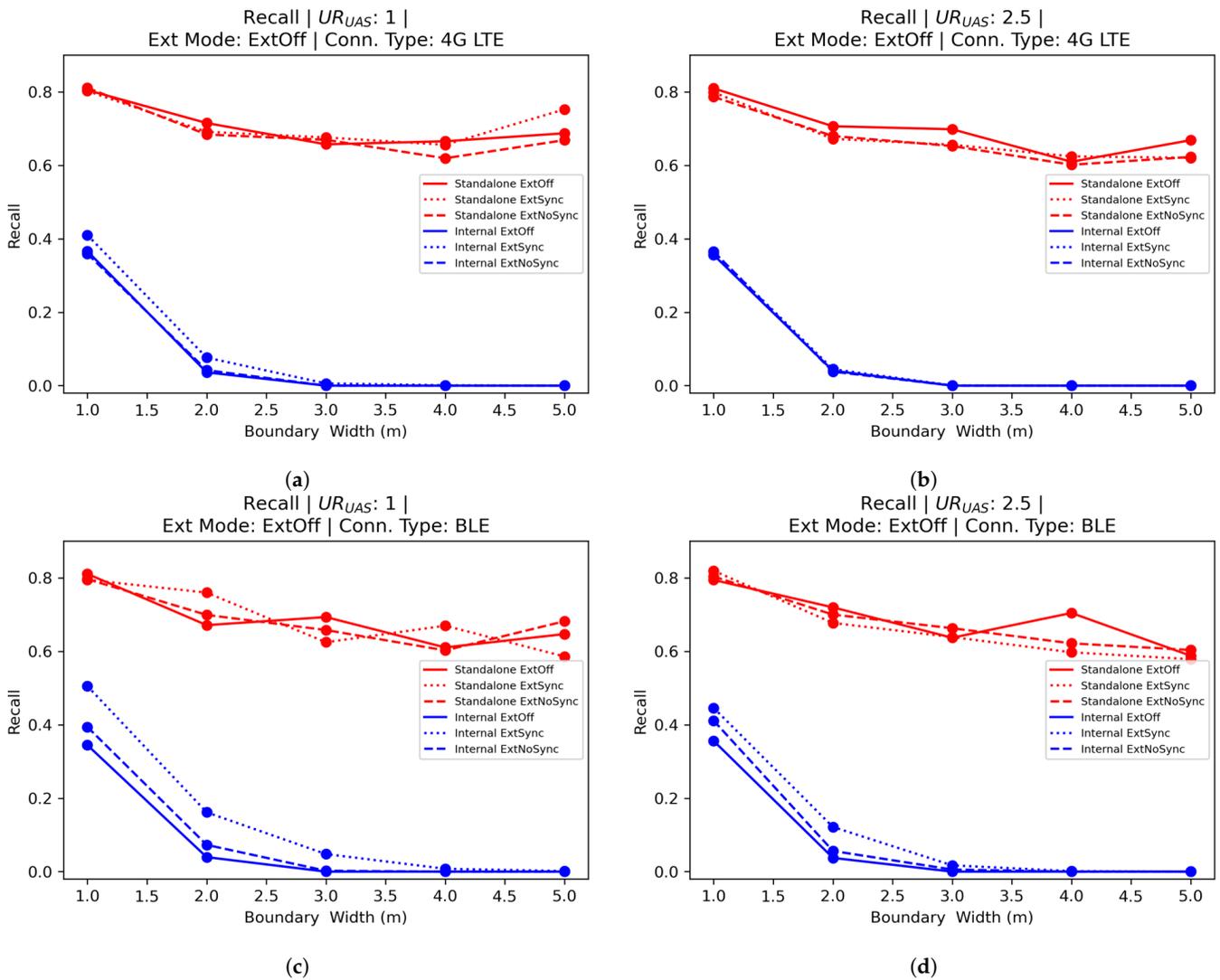


Figure 14. Recall versus boundary width curves—comparison between the integrated UA navigational position source and standalone tracking with the independent position source; (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

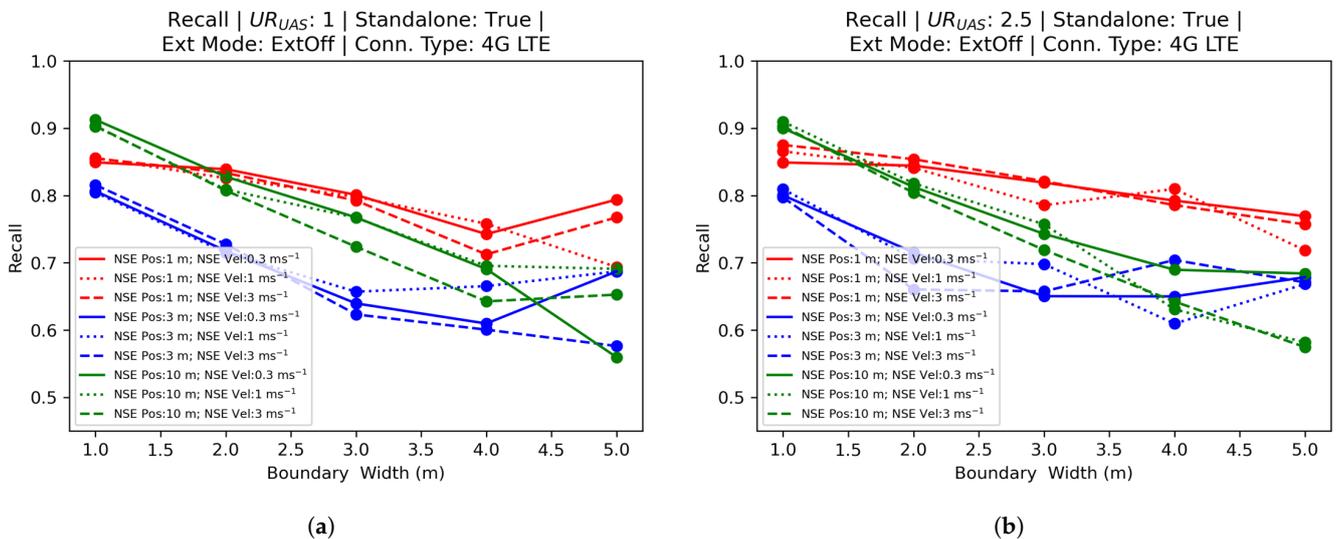


Figure 15. Cont.

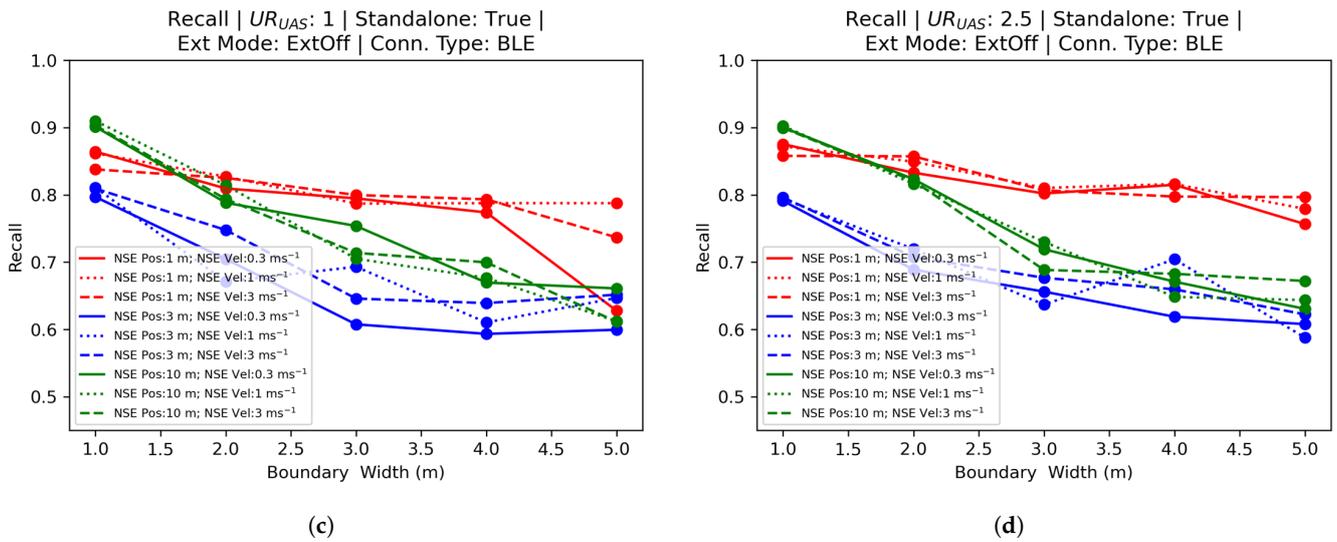


Figure 15. Recall versus boundary width curves—ExtOff; (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

Both ExtSync and ExtNoSync modes feature a trend where tracking configurations with an NSE of 10 m perform better than those with an NSE of 3 m, as shown in Figures 16 and 17. This likely shows how recall is not extremely sensitive to position errors. Nevertheless, a higher velocity error does show interaction effects with both ExtSync and ExtNoSync modes; as with precision, when UR_{UAS} is low and communication latencies are high, poor recall is experienced.

A brief comparison between precision and recall metrics reveals that precision is more readily affected by tracking KPIs and configurations (i.e., nuisance alerts are more readily generated when tracking performance is poor, as opposed to missed non-conforming events).

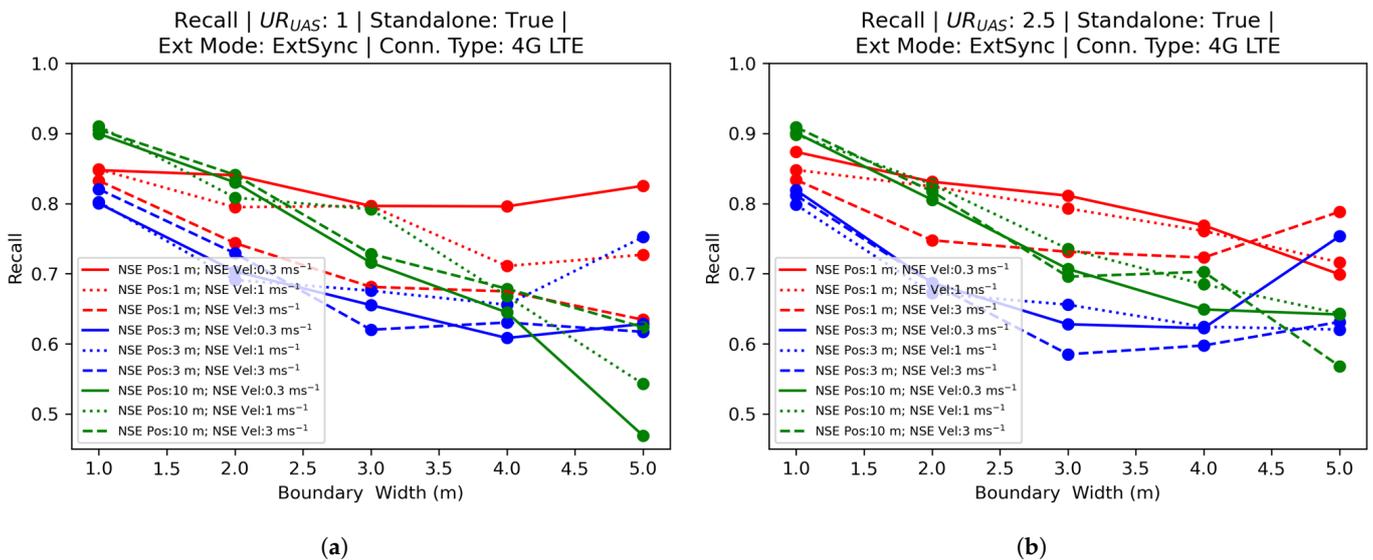


Figure 16. Cont.

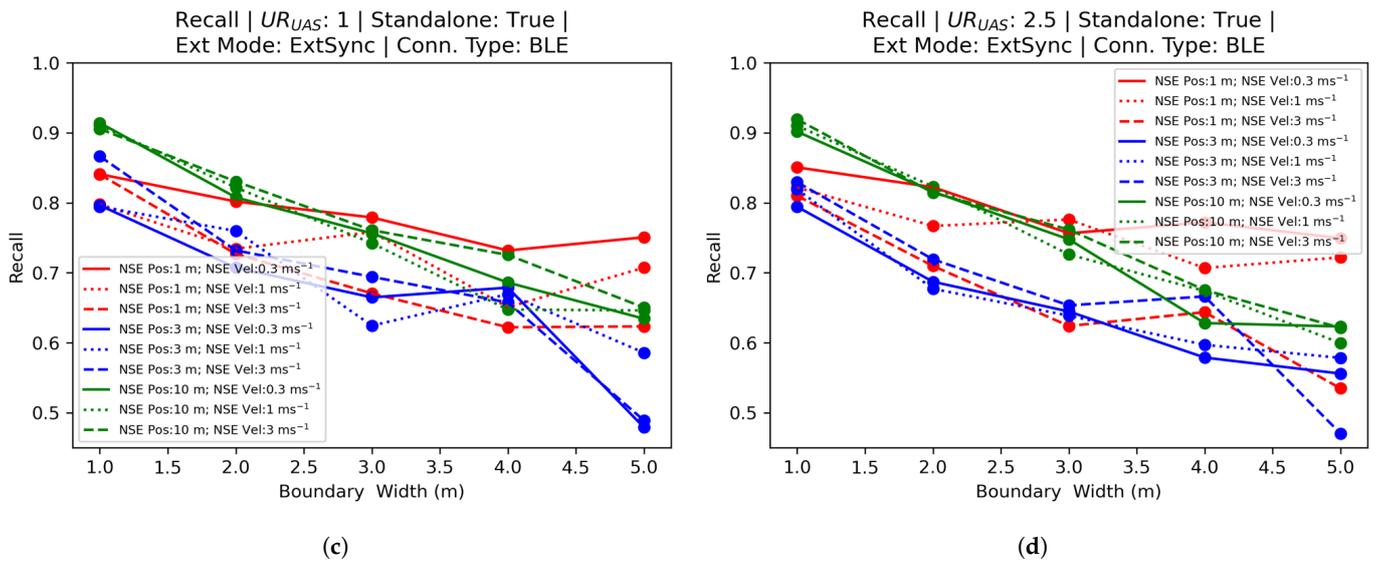


Figure 16. Recall versus boundary width curves—ExtSync; (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

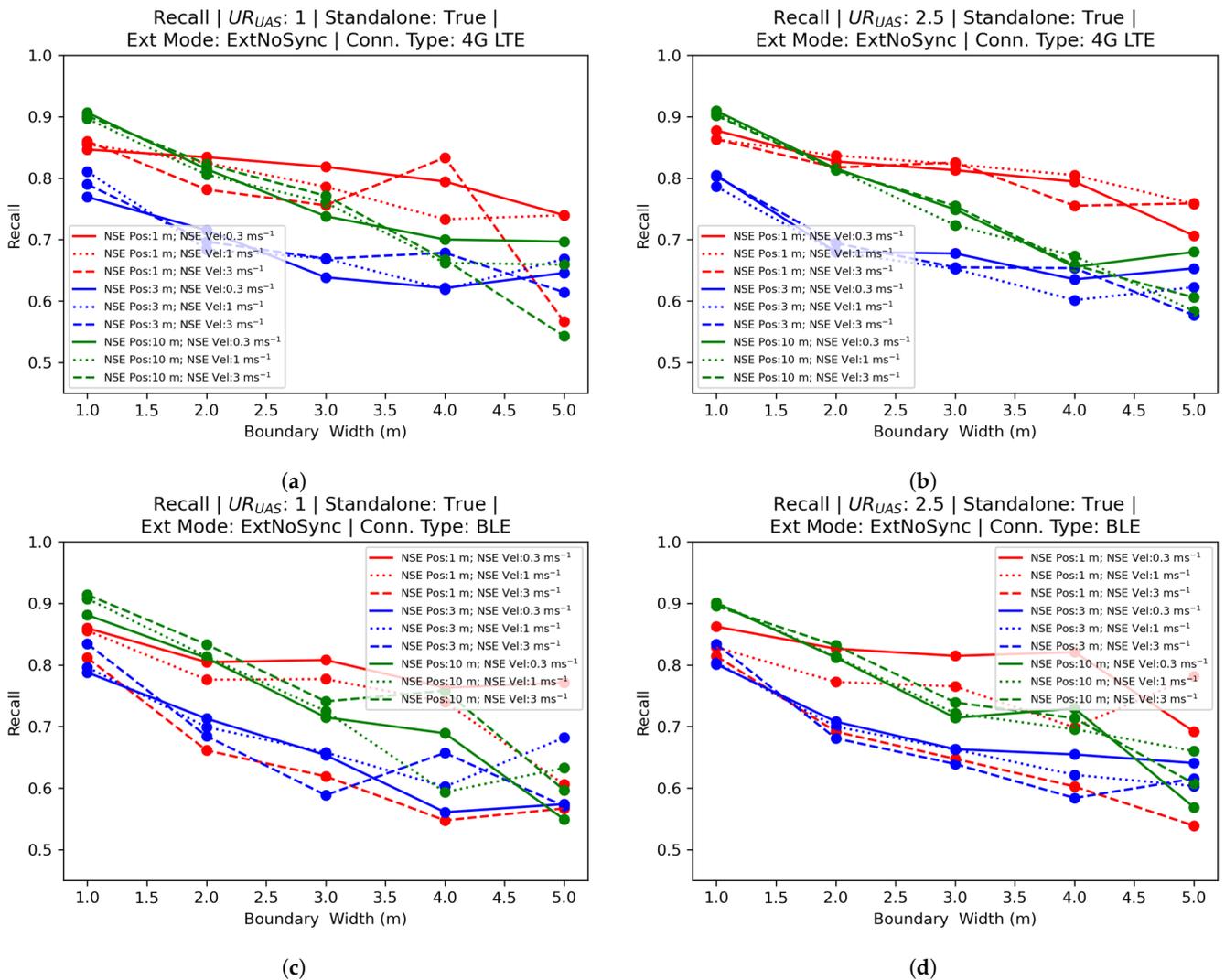


Figure 17. Recall versus boundary width curves—ExtNoSync; (a) 4G LTE, UR_{UAS} set to 1 Hz; (b) 4G LTE, UR_{UAS} set to 2.5 Hz; (c) BLE, UR_{UAS} set to 1 Hz; (d) BLE, UR_{UAS} set to 2.5 Hz.

4. Conclusions

In this study, we performed a preliminary comparison between competing UTM cooperative tracking approaches by means of latency experiments. While the comparison is by no means perfect (4G LTE tests use a short, proprietary message integrated into the timing program while BLE requires the launching of prototype implementations of RID applications), initial tests indicate that urban low-altitude applications may favor cellular-based networked approaches; this is particularly so if ground-based broadcast receivers further need to upload data to a USS-like centralized entity for conformance monitoring services. Future studies may include experimental or simulation work to determine maximum flight altitudes for acceptable coverage within the urban context.

Latency measurements produced models used in a hybrid SITL–post-processing simulation approach, which was applied to conduct a Monte Carlo sensitivity analysis of tracking KPIs. Additionally, safety-related metrics, in the form of non-conformance detection delay, and continuous-time extensions of precision and recall, were used as part of the analysis.

Extended formulations of recall and precision metrics were designed in this study in order to evaluate the detection success rate in a continuous time–space. The formulations were used in the evaluation and support of the interpretation of the simulation results.

The study found that improving the UA update rate (UR_{UAS}) from 1 Hz to 2.5 Hz did not offer commensurate improvements in the CM performance, but improved the variability in detection delays. In general, extrapolation aided in reducing detection delays, mainly if both UA and CM services used GNSS-derived timestamps. However, high-velocity errors and high latencies resulted in high variability in the delay time and, therefore, tempered the effectiveness of linear extrapolation.

Additionally, for the given parameter ranges, it was found that navigational position error yielded the greatest influence on detection delay variability. A larger effective navigational position error corresponded to lower precision but did not always result in lower recall.

Author Contributions: Conceptualization, W.D. and Z.H.Q.; methodology, W.D. and Z.H.Q.; software, W.D. and Z.H.Q.; validation, W.D., Z.H.Q. and B.P.; formal analysis, W.D. and B.P.; investigation, W.D. and B.P.; resources, M.F.; writing—original draft preparation, W.D. and Z.H.Q.; writing—review and editing, W.D., B.P. and M.F.; visualization, Z.H.Q.; supervision, M.F.; project administration, M.F.; funding acquisition, M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC were funded by the National Research Foundation (NRF), Singapore, and the Civil Aviation Authority of Singapore (CAAS), under the Aviation Transformation Programme.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: This research is supported by the National Research Foundation, Singapore, and the Civil Aviation Authority of Singapore, under the Aviation Transformation Programme. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the National Research Foundation, Singapore, and the Civil Aviation Authority of Singapore.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ASTM	American Society for Testing and Materials
BLE	Bluetooth low energy
BVLOS	beyond visual line-of-sight
CM	conformance monitoring

FAA	Federal Aviation Administration
FTE	flight technical error
GNSS	global navigation satellite system
KPI	key performance indicator
NAS	National Airspace System
NSE	navigational system error
ODID	OpenDroneID
OIV	operational intent volume
PDE	path definition error
RID	Remote ID
TCT	trajectory change time
TSE	total system error
UA	unmanned aircraft
UAS	unmanned aircraft system
USS	UTM service supplier
UTM	UAS traffic management
VLOS	visual line-of-sight

References

1. Prevot, T.; Rios, J.; Kopardekar, P.; Robinson, J.E., III; Johnson, M.; Jung, J. UAS Traffic Management (UTM) Concept of Operations to Safely Enable Low Altitude Flight Operations. In Proceedings of the 16th AIAA Aviation Technology, Integration, and Operations Conference, Washington, DC, USA, 13–17 June 2016; pp. 1–16. [\[CrossRef\]](#)
2. Wang, C.H.J.; Low, K.H.; bin Che Man, M.H.; Dai, W.; Ng, E.M. Safety-Focused Framework for Enabling UAS Traffic Management in Urban Environment. In Proceedings of the AIAA AVIATION 2022 Forum, Chicago, IL, USA, 27 June–1 July 2022; p. 3618. [\[CrossRef\]](#)
3. Evans, A.D.; Egorov, M.; Anand, A.; Campbell, S.E.; Zanlongo, S.; Young, T.; Sarfaraz, N. Safety Assessment of UTM Strategic Deconfliction. In Proceedings of the AIAA Scitech 2023 Forum, National Harbor, MD, USA, 23–27 January 2023; p. 0965. [\[CrossRef\]](#)
4. Dai, W.; Pang, B.; Low, K.H. Conflict-free four-dimensional path planning for urban air mobility considering airspace occupancy. *Aerosp. Sci. Technol.* **2021**, *119*, 107154. [\[CrossRef\]](#)
5. Zhang, M.; Yan, C.; Dai, W.; Xiang, X.; Low, K.H. Tactical conflict resolution in urban airspace for unmanned aerial vehicles operations using attention-based deep reinforcement learning. *Green Energy Intell. Transp.* **2023**, *2*, 100107. [\[CrossRef\]](#)
6. Wang, C.J.; Ng, E.M.; Low, K.H. Investigation and modeling of flight technical error (FTE) associated with UAS operating with and without pilot guidance. *IEEE Trans. Veh. Technol.* **2021**, *70*, 12389–12401. [\[CrossRef\]](#)
7. Deng, C.; Wang, C.H.J.; Low, K.H. Preliminary UAS Navigation Performance Analysis in Urban-like Environments. In Proceedings of the AIAA Aviation and Aeronautics Forum and Exposition, AIAA AVIATION Forum 2021, Washington, DC, USA, 2–6 August 2021; pp. 1–12. [\[CrossRef\]](#)
8. Wang, C.H.J.; Deng, C.; Low, K.H. Parametric Study of Structured UTM Separation Recommendations with Physics-Based Monte Carlo Distribution for Collision Risk Model. *Drones* **2023**, *7*, 345. [\[CrossRef\]](#)
9. ISO/DIS 23629-12; UAS Traffic Management (UTM)-Part 12: Requirements for UTM Service Providers. International Organization for Standardization: Geneva, Switzerland, 2021.
10. Schwalb, E.; Schwalb, J. Improving redundancy and safety of UTM by leveraging multiple UASS. In Proceedings of the IEEE 2019 International Conference on Unmanned Aircraft Systems (ICUAS), Atlanta, GA, USA, 11–14 June 2019; pp. 100–110. [\[CrossRef\]](#)
11. Dai, W.; Quek, Z.H.; Low, K.H. A Simulation-Based Study on the Impact of Tracking Performance on UTM Flight Safety. In Proceedings of the 2022 Integrated Communication, Navigation and Surveillance Conference (ICNS), Dulles, VA, USA, 5–7 April 2022; pp. 1–13. [\[CrossRef\]](#)
12. Quek, Z.H.; Dai, W.; Low, K.H., Analysis of Safety Performance of Tracking Services Based on Simulation of Unmitigated UAS Conflicts. In Proceedings of the AIAA SCITECH 2023 Forum, National Harbor, MD, USA, 23–27 January 2023. [\[CrossRef\]](#)
13. Ballio, F.; Guadagnini, A. Convergence assessment of numerical Monte Carlo simulations in groundwater hydrology. *Water Resour. Res.* **2004**, *40*. [\[CrossRef\]](#)
14. Everdij, M.; Blom, H.; Klompstra, M. Dynamically Coloured Petri Nets for Air Traffic Management Safety Purposes. *IFAC Proc. Vol.* **1997**, *30*, 169–174. [\[CrossRef\]](#)
15. Paxton, P.; Curran, P.J.; Bollen, K.A.; Kirby, J.; Chen, F. Monte Carlo experiments: Design and implementation. *Struct. Equ. Model.* **2001**, *8*, 287–312. [\[CrossRef\]](#)
16. Ata, M.Y. A convergence criterion for the Monte Carlo estimates. *Simul. Model. Pract. Theory* **2007**, *15*, 237–246. [\[CrossRef\]](#)
17. Fricke, H. Using agent-based modeling to determine collision risk in complex TMA environments: The turn-onto-ILS-final safety case. *Aeronaut. Aerosp. Open Access J.* **2018**, *2*, 155–164. [\[CrossRef\]](#)

18. Förster, S.; Fricke, H.; Rabiller, B.; Hickling, B.; Favennec, B.; Zeghal, K. Analysis of safety performances for parallel approach operations with performance based navigation. In Proceedings of the 19th USA/Europe Air Traffic Management Research and Development Seminar (ATM Seminar), Vienna, Austria, 17–21 June 2019.
19. Stroeve, S.; Blom, H.; Medel, C.H.; Daroca, C.G.; Cebeira, A.A.; Drozdowski, S. Development of a Collision Avoidance Validation and Evaluation Tool (CAVEAT): Addressing the intrinsic uncertainty in TCAS II and ACAS X. In Proceedings of the 13th USA/Europe Air Traffic Management Research and Development Seminar, Vienna, Austria, 17–21 June 2019.
20. Stroeve, S.H.; Blom, H.A.; Medel, C.H.; Daroca, C.G.; Cebeira, A.A.; Drozdowski, S. Modeling and simulation of intrinsic uncertainties in validation of collision avoidance systems. *J. Air Transp.* **2020**, *28*, 173–183. [[CrossRef](#)]
21. Torres-Pomales, W. *Conformance Monitoring in Air Traffic Control*; Technical Report; NASA: Washington, DC, USA, 2020.
22. Reynolds, T.G.; Hansman, R.J. Conformance monitoring approaches in current and future air traffic control environments. In Proceedings of the 21st Digital Avionics Systems Conference, Irvine, CA, USA, 27–31 October 2002; Volume 2, p. 7C1.
23. Reynolds, T.G.; Hansman, R.J. Investigating conformance monitoring issues in air traffic control using fault detection techniques. *J. Aircr.* **2005**, *42*, 1307–1317. [[CrossRef](#)]
24. Lee, K.; Fukuda, Y. A Bayesian approach for conformance monitoring. In Proceedings of the 11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Including the AIAA Balloon Systems Conference and 19th AIAA Lighter-Than, Virginia Beach, VA, USA, 20–22 September 2011; p. 6857.
25. *ASTM F3548-21*; Standard Specification for UAS Traffic Management (UTM) UAS Service Supplier (USS) Interoperability. Technical Report; ASTM: West Conshohocken, PA, USA, 2022. [[CrossRef](#)]
26. Zheng, Y.; Brudnak, M.J.; Jayakumar, P.; Stein, J.L.; Ersal, T. An Experimental Evaluation of a Model-Free Predictor Framework in Teleoperated Vehicles. *IFAC PapersOnLine* **2016**, *49*, 157–164. [[CrossRef](#)]
27. Gorsich, D.J.; Jayakumar, P.; Cole, M.P.; Crean, C.M.; Jain, A.; Ersal, T. Evaluating mobility vs. latency in unmanned ground vehicles. *J. Terramechan.* **2018**, *80*, 11–19. [[CrossRef](#)]
28. Zheng, Y.; Brudnak, M.J.; Jayakumar, P.; Stein, J.L.; Ersal, T. A Predictor-Based Framework for Delay Compensation in Networked Closed-Loop Systems. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 2482–2493. [[CrossRef](#)]
29. *ASTM F38 Committee*; Standard Specification for Remote ID and Tracking. ASTM: West Conshohocken, PA, USA, 2022. [[CrossRef](#)]
30. Walelgne, E.A.; Asrese, A.S.; Manner, J.; Bajpai, V.; Ott, J. Understanding Data Usage Patterns of Geographically Diverse Mobile Users. *IEEE Trans. Netw. Serv. Manag.* **2021**, *18*, 3798–3812. [[CrossRef](#)]
31. Srinivasan, A. Measuring and Optimizing for Network Conditions on Drones. Master’s Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2021.
32. Dai, W.; Zhang, M.; Low, K.H. Data-Efficient Modeling for Precise Power Consumption Estimation of Quadrotor Operations Using Ensemble Learning. *arXiv* **2022**, arXiv:2205.10997.
33. Rodrigues, T.A.; Patrikar, J.; Choudhry, A.; Feldgoise, J.; Arcot, V.; Gahlaut, A.; Lau, S.; Moon, B.; Wagner, B.; Matthews, H.S.; et al. In-flight positional and energy use data set of a DJI Matrice 100 quadcopter for small package delivery. *Sci. Data* **2021**, *8*, 6–13. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.