

Development and Field Test Results of Distributed Ground Sensor Fusion Based Object Tracking

Thomas Lombaerts* Evan Kawamura† Keerthana Kannan‡
NASA Ames Research Center, Moffett Field, CA 94035

Chester Dolph§
NASA Langley Research Center, Langley, VA 94035

Vahram Stepanyan¶ George Gorospe|| Corey Ippolito**
NASA Ames Research Center, Moffett Field, CA 94035

Autonomous operations are a crucial aspect in the context of Advanced Air Mobility and other emerging aviation markets. In order to enable this autonomy, an accurate and detailed understanding of the positions of the various vehicles in the air is necessary. Full localization independent of on-board sensors makes the system suitable for noncooperative vehicles. This paper focuses on the object tracking part that relies on distributed ground-based RF and other sensor fusion, considering specific properties and limitations of different sensor types. Results show satisfactory performance in nominal scenarios with full coverage for some sensor types, but RF signals are challenging because of their nature. This paper includes the results from simulations as well as field tests to support the observations and conclusions.

I. Introduction

RECENTLY, there have been numerous developments in the area of advanced air taxi operations. Combined with other novel transportation concepts such as autonomous aerial cargo delivery etc., they form the Advanced Air Mobility (AAM) ecosystem that will help to meet the challenging mobility needs of an increasing population density in metropolitan areas, taking into account the limited capacity of the currently available transportation infrastructure. To realize the full vision of AAM, technology will be needed to allow a few operators to operate many vehicles (m:N). As most m:N implementations will require highly autonomous vehicles, the current state-of-the-art will have to advance for on-board autonomous functionality. One of the key areas requiring further technology development is sensor-based perception techniques. For this perception, techniques are being developed to address problems such as GPS-free Alternative Positioning, Navigation and Timing (APNT) [1]. Advanced sensing and tracking algorithms will be necessary from the ground in an urban environment, where traditional traffic tracking infrastructure such as air traffic radars will be useless. This infrastructure will be needed for traffic monitoring, traffic guidance, as well as spacing and separation in order to increase overall capacity and efficiency in the airspace. In this context, ground-based perception systems, capable of APNT, are critical to safe autonomous operations. This APNT relates to all moving objects in the sky, which can be cooperative (aircraft, AAM vehicles, etc.) as well as noncooperative (smaller drones, birds, ...). The distinction between both lies in the fact that there is some form of data exchange (ADS-B or similar) with the former category in order to coordinate path planning and separation. This exchange of information is nonexistent for the latter category, which means that other sensors will be required to detect and track these objects independently from the ground. Ground-based object detection and tracking are two essential prerequisites for the purpose of traffic monitoring, and also to provide some off-board redundancy for detect and avoid (DAA). For multi-target tracking, there is an additional challenge of observation-to-track association.

Another research area that is investigated in this paper is on-board localization by means of a wireless network, which fits under the aforementioned APNT label. The set up involves a distributed network of ground based wireless

* Aerospace Research Engineer, KBR, Intelligent Systems Division, Mail Stop 269-1, AIAA Associate Fellow, email: thomas.lombaerts@nasa.gov

† Computer/GNC Engineer, Intelligent Systems Division, Mail Stop 269-1, AIAA member, email: evan.kawamura@nasa.gov

‡ Software Engineer, KBR, Intelligent Systems Division, Mail Stop 269-1, AIAA member, email: keerthana.kannan@nasa.gov

§ Aerospace Engineer, Aeronautics Systems Engineering Branch

¶ Principal Scientist, KBR, Intelligent Systems Division, Mail Stop 269-1, AIAA senior member, email: vahram.stepanyan@nasa.gov

|| Senior Research Scientist, Intelligent Systems Division, Mail Stop 269-1, email: george.e.gorospe@nasa.gov

** Aerospace Scientist, Intelligent Systems Division, Mail Stop 269-1, email: corey.a.ippolito@nasa.gov

network transmitter antennas and an on-board receiver antenna. Trilateration (localization by means of intersecting distances from known locations) based on the received signal strength results in an estimated ownship position. However, noise and other sensor characteristics will have a major impact on position accuracy, together with the scale of the scenario (short range indoors vs longer range outdoors). Combined localization with other ground based equipment such as cameras and radars might be necessary to achieve the required accuracy, depending on the circumstances. This paper aims to preliminary investigate this topology problem and to identify follow-up research interests for future work.

A. Focus of this paper

This paper focuses on the single object tracking part that relies on adaptive multi-sensor fusion in a distributed ground-based sensor setup, taking into account specific properties and limitations of different sensor types. The sensor fusion is done centrally by means of a Kalman Filter, relying on the raw measurements from the network of non-collocated ground-based sensors. Some other approaches include local processing at the individual sensor or node of sensors level, after which the different local estimates are integrated by some form of consensus filter, for example using weighted averages. The sensor fusion strategy used here is adaptive, so that lower sampling rates of specific sensors are taken into account, as well as loss of signal (for example because the object is outside the field of view or the useful range for a particular sensor location and viewing angle, or because of a sensor malfunction resulting in a sensor signal dropout), and invalid data. This capability is achieved by incorporating valid flags in the weighting matrix of the Kalman filter. This paper also explores which sensor properties and/or network characteristics have a dominant impact on the accuracy of the estimate, such as the number of sensors. This paper is a follow up study and extension of the work that was presented in [2–6].

B. Literature survey

As already introduced before, most literature considers local processing at the individual sensor or node of sensors level, after which the different local estimates are integrated by some form of consensus filter, for example using weighted averages. [7] considers a distributed sensor network, consisting of a set of processing nodes collecting data from sensors. The nodes communicate with each other, so that each node updates its local estimate with information received from the other nodes. [8] reviews developments in distributed sensor networks from four aspects: network structure, data processing paradigm, sensor fusion algorithm with emphasis on fault-tolerant algorithm design, and optimal sensor deployment strategy. [9] describes the solution to a real-time distributed resource allocation application involving distributed situation assessment. The solution decomposes the environment into a number of sectors, where individual sensor nodes in a sector are specialized dynamically to address different parts of the objective to track objects moving in the environment in real-time, given uncertainty and constraints on sensor loads, communication, power consumption, action characteristics, and clock synchronization. [10] investigates a binary decentralized detection problem in which a network of wireless sensors provides relevant information about the state of nature to a fusion center. The fusion center attempts to accurately reconstruct the state of nature. Optimal sensor configuration and capacity constraints are considered. The gain offered by having more sensors exceeds the benefits of getting detailed information from each sensor. The method of [11] consists of two parts. First a distributed sensor fusion algorithm integrates two different distributed consensus filters to achieve cooperative sensing among sensor nodes. This fusion algorithm has two phases. In the first phase, the weighted average consensus filter is developed, which allows each sensor node to find an estimate at each time step. In the second phase, the average consensus filter is used to allow each sensor node to find a confidence of the estimate at each time step. The final estimate is iteratively updated during the movement of the sensors via weighted average. [12] studies the problem of location detection for commercial aircraft and methods to improve location detection over any terrain the flight path traverses. The proposed techniques are based on the internet-of-things model for aircraft, where the aircraft can communicate with each other within a certain range. Distributed algorithms detect location using such methods that work effectively when the aircraft is outside the range of radar and on an oceanic route. Using the proposed methods, the precise location of all aircraft can be tracked to a higher degree.

This paper places special emphasis on localization by means of wireless sensor networks, as is done in [13–16], and by means of cameras. This former kind of localization relies on the measured distances between several transmitters with a known location on one hand and a receiver with an unknown position on the other. This is in contrast with the concept of triangulation for cameras where one uses the angles measured by several sensors with a known location observing the same tracked object. The difference is illustrated in Fig. 1. It is noteworthy that triangulation with ground based cameras is for external ground based localization which also works for noncooperative targets. Trilateration with wireless networks, on the other hand, is for on-board positioning (as an alternative to GPS) and requires ground based

transmitter antennas as well as an on-board receiver antenna and is by default centralized. Generally speaking, sensor disturbances have a direct impact on the accuracy of the position estimates for triangulation as well as trilateration, illustrated by the wider uncertainty bands instead of thin lines for each sensor in Fig. 1. Future research will also aim to explore how much the accuracy can be improved by combining both concepts of triangulation and trilateration together in the sensor fusion strategy.

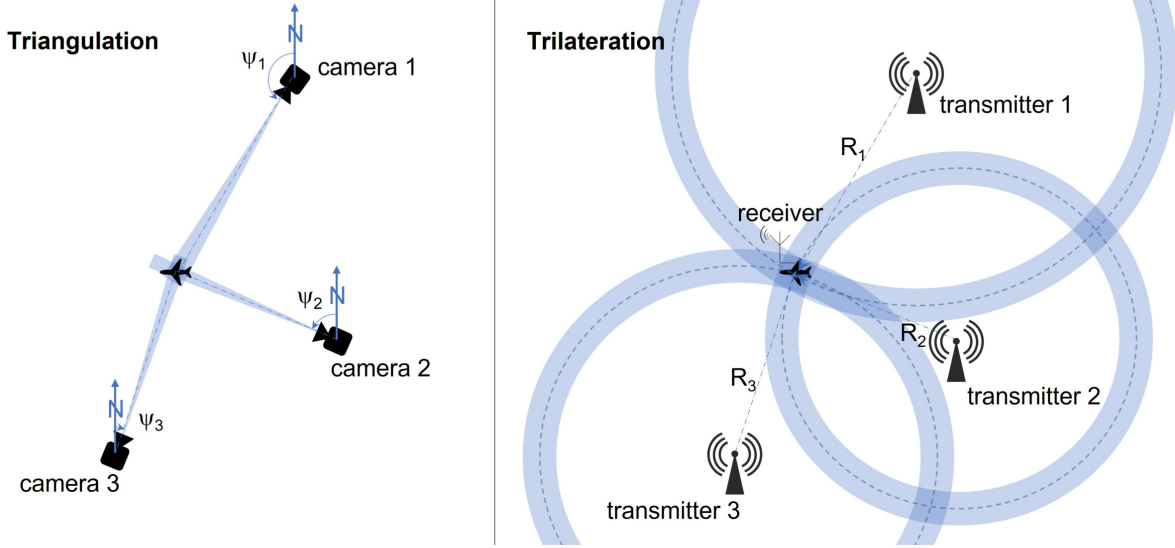


Fig. 1 Concepts of triangulation versus trilateration

The approach followed in this paper is centralized and does not rely on middle layer sensor nodes. Although resulting in an overall longer communication path, it removes adverse effects caused by the consensus filter.

C. Paper Structure

The structure of this paper is as follows. Sections II and III describe respectively the scenarios and the relevant sensor characteristics which have been used to evaluate the performance of the developed object tracker. Section IV elaborates how these distributed sensor data are combined in a Kalman Filter based object tracker setup. Section V shows some results for a few different scenarios, which are based on simulations (see V.A) as well as field tests (see V.B). In each case different sensor types were considered. Finally, Section VI gives some overall conclusions and recommendations based on this work.

II. Scenarios

The main scenario in this publication is at Ames Research Center, between the gantry of the Outdoor Aerodynamic Research Facility (OARF) and the Ames's Disaster Assistance and Rescue Team (DART) training area, with strategically placed cameras and modems throughout the area. More details about the design and implementation of the scenarios are given in [17]. The current studies discussed in this paper are both simulation-based as well as involving field tests with real-life sensor data.

A. First scenario: simulation based flight

The first scenario considered in this publication is a simulated flight from the OARF near GN3 to the DART site near GN1, as outlined in Fig. 2 and 3. Fig. 2 shows the outline over the terrain with camera observation cones not to scale.

Fig. 3 shows the observation cones to scale for electro-optical (EO) cameras in each ground node (GN), but other scenarios consider the use of modems or some other form of distance measuring equipment (DME), such as radar. The latter are omnidirectional and thus have an observation circle on the map instead of a cone like for a camera. The sensor characteristics are discussed in Section III. The locations and viewing angles of the GN's for this simulated scenario are provided in Tab. 1.



Fig. 2 Outline of the scenario over terrain with camera observations, not to scale, source: [17]

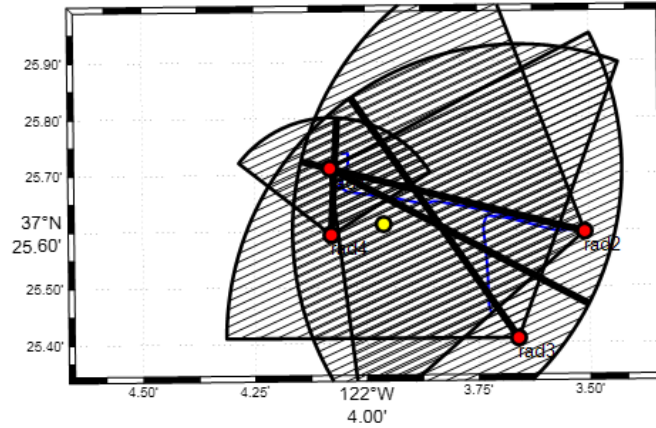


Fig. 3 Outline of scenario for camera observations only, to scale

B. Second scenario: real life field test

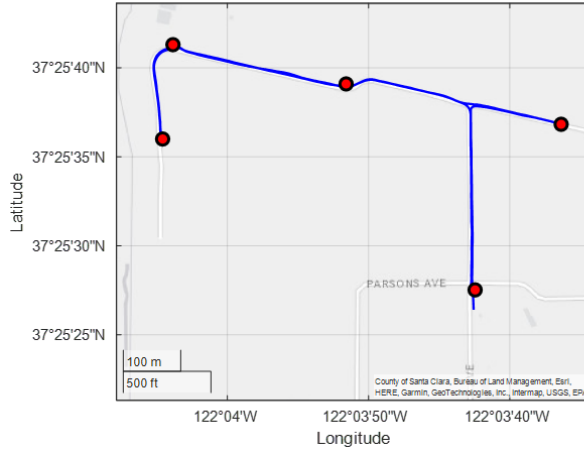
Two real life field tests were performed, one ground based with RF antennas and modems described in Sec. II.B.1 and one flight test with cameras described in Sec. II.B.2. More details about the design and the implementation of the flight test are given in [17].

1. Field test with modems

The first version of the field test involves a set of RF antennas and modems in several ground nodes, and a moving car with another RF antenna that receives the wireless signals from the various ground nodes. Fig. 4 shows the general outline of the scenario in Fig. 4(a) together with a picture of one of the ground nodes in Fig. 4(b) with the battery powered modem and RF antenna mounted in the upper left corner of the frame.

Table 1 Distributed sensors locations and angles between OARF and DART site for simulations

Ground Node	Lat (°)	Long (°)	Abs. Alt. (m)	Heading (°)	Elev. (°)	Range (m)
GN 1	37.429540	-122.067712	4.53	140	0	1,000
GN 2	37.425155	-122.059832	4.53	10	0	1,000
GN 3	37.423476	-122.061238	4.53	-40	0	1,000
GN 4	37.427278	-122.061654	4.53	190	0	400



(a) Outline of the field test



(b) Ground node with RF antenna

Fig. 4 Outline of the RF antenna field test and picture of one of the ground nodes

2. Flight test with cameras

The second version of the field test is a real life flight test, flown by a multicopter drone, as shown in Fig. 5(a). The outline of the scenario is very similar to the simulated one shown in Fig. 3, which shows the flight path, together with the sensor locations and their observation cones. Fig. 5(b) shows a picture of one of the ground nodes with the EO camera setup. Tab. 2 shows accurate location and viewing angle and range information for each individual ground node. Experiments have shown that the accuracy of the estimates depends to a large extent on the precision of this ground node information.

Table 2 Distributed sensors locations and angles around DART site for field test

Ground Node	Lat (°)	Long (°)	Abs. Alt. (m)	Heading (°)	Elev. (°)	Range (m)
GN 1	37.428067	-122.067703	1.54	66.53	9.68	600
GN 2	37.427094	-122.061213	1.54	246.54	29.29	600
GN 3	37.424167	-122.060000	1.57	303.14	34	600
GN 4	37.426556	-122.067972	1.52	44.78	26.31	600



(a) Multicopter drone Alta X, used for the flight test



(b) Ground node with EO camera

Fig. 5 Pictures of multicopter drone and one of the ground nodes

III. Sensor Characteristics, Measurement Processing and Data Calibration

Three types of sensors are considered in the aforementioned scenarios, namely electro-optical cameras, wireless RF equipment and some kind of distance measuring equipment (DME). Radars were considered in previous work [3] and are modeled in the Kalman Filter observation models, but they were not used in this study. Cameras provide azimuth and elevation only. Wireless RF equipment allows to infer distances based on the residual strength of the signal at the receiver's location. More detailed information is given in Sec. III.A. DME equipment provides range to target only and serves as a reference to compare wireless RF equipment with. Sensor characteristics for DME were chosen similar as for range information provided by radar. Each sensor type has different sensed variables, together with their respective sensor characteristics, such as limitations on field of view and useful range, noise, latency, update rate and resolution. Some characteristic values from the literature of state-of-the-art sensor hardware [18] are given in Table 3. Specific details of the sensor readings for the specific aforementioned scenarios are given in Sec. III.B for the simulated flights and III.C for the field tests.

Table 3 Sensor characteristics

sensor	EO camera	RF	DME
sensed variables	azimuth [deg] elevation [deg]	RSSI [dB]	range[m]
limits on field of view (16/9)	az: 60deg el: 33.75deg	–	–
max useful range	4,000m	–	2,000m
min useful range	0m	0m	0m
noise sigma	$9.1 \cdot 10^{-3}$ deg	3 dB	3.25m
latency	0.1s	0.1s	0.1s
update rate	30fps	1/3s	10Hz
resolution	$1.3 \cdot 10^{-3}$ deg	$1.3 \cdot 10^{-3}$ dB	3.25m

A. RSSI signal from RF equipment

Received signal strength inference (RSSI) is the concept where the strength of the signal received from a wireless transmitter antenna is used to infer the distance between transmitter and receiver. The equation that describes this signal strength $RSSI$ based on the distance is as follows:

$$RSSI = P_{Tx} - P_{L_0} - 10\gamma \log_{10} \left(\frac{R}{R_0} \right) - G(t) \quad (1)$$

where P_{Tx} is the transmitter power ($-10dB$), γ is the path loss exponent ($\gamma = 2.5$ for outdoor applications), and the pair P_{L_0} with R_0 is the calibration point with respectively local signal strength and distance to that location. For the scenario shown here, the following values were chosen: $P_{L_0} = -5dB$ at $R_0 = 3m$. R is the distance between transmitter and receiver. Finally, $G(t)$ is white noise with a standard deviation of $\sigma_G = 3dB$. The distance-to-signal-strength profile for the numbers provided here is shown in Fig. 6.

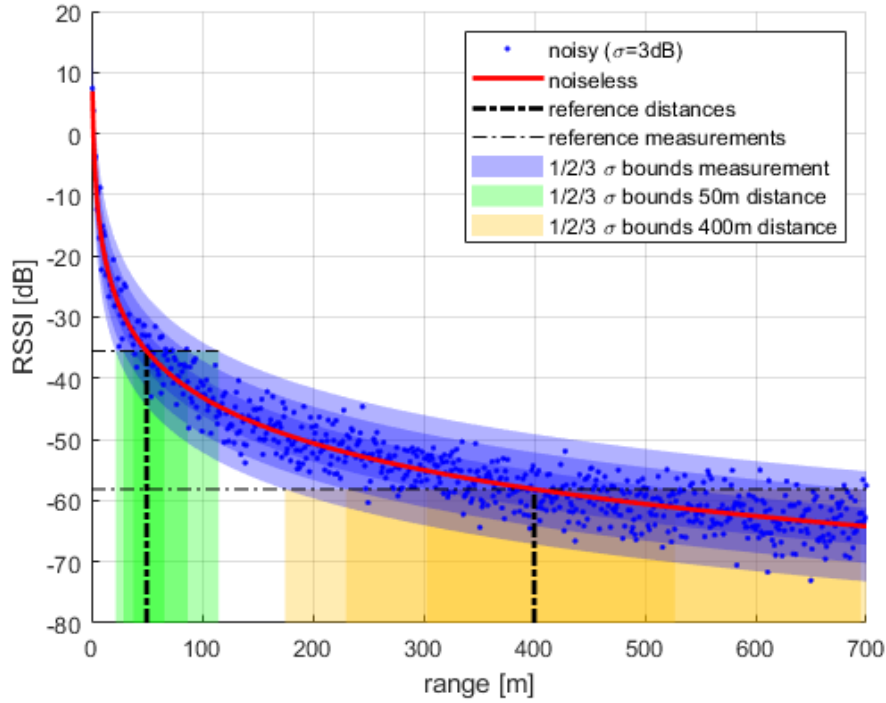


Fig. 6 Residual Signal Strength as a function of range

This figure shows the relationship between range and residual signal strength. Ideal noise-free measurements are shown together with a representative noisy sample and its 1, 2 and 3 σ bounds for the (assumed) normal distribution. Although the mapping shows that every theoretical RSSI value has only one corresponding theoretical distance value, the small signal-to-noise ratio makes it much harder to infer an accurate distance from the signal strength, especially at longer distances. In this specific illustrative example in Fig. 6, a distance of 50m between transmitter and receiver results in a theoretical noise-free measurement of around $-35dB$. However a standard deviation of $3dB$ means that a measurement of $-35dB$ can occur at any distance between 39m and 67m for 68% of all data points, but between 30m and 88m for 95% and even between 23m and 116m for 99.7% of all data points. These uncertainty ranges increase significantly with distance. For example, a distance of 400m between transmitter and receiver results in a theoretical noise-free measurement of approximately $-58dB$. However a standard deviation of $3dB$ means that a measurement of $-58dB$ can occur at any distance between 304m and 528m for 68% of all data points, but between 231m and 696m for 95% and even between 176m and exceeding 700m for 99.7% of all data points. This shows not only the need for trilateration between signal strengths from multiple transmitters. It also shows that multiple transmitters closer to the tracked object (i.e. a much denser network) would increase accuracy significantly.

B. Synthetic sensor readings for simulated flight

For the simulated flights, synthetic sensor readings were created, relying on the sensor characteristics as specified in Tab. 3. The actual sensor values used are shown in Sec. V.A in Fig. 12(a), 17(a) and 23(b).

The simulated flight data for this scenario are obtained from a high fidelity simulation in a custom NASA flight software called ‘Reflexion’ and is visualized by a commercial software called X-plane [19].

C. Actual sensor readings for field tests

During the field tests, actual sensor readings were recorded, namely RSSI signals from the RF antennas and azimuth and elevation angles from the camera images. The former signals are introduced in Sec. III.C.1. The latter measurements needed image processing which is briefly discussed in Sec. III.C.2.

1. RSSI readings for field test

During the field test as described in Sec. II.B.1, RSSI readings were collected as well as GPS logdata. The GPS logdata allowed to create synthetic RSSI readings per RF antenna location and were also used for calibrating the RF antenna model parameters as described in Sec. III.A. The values per modem obtained after calibration and as used in the sensor fusion observation model are given in Tab. 4. Fig. 28(a) shows the fit between synthetic data and true measurements after calibration.

Table 4 Distributed RF antenna locations and model parameters around DART site for field test

Ground Node	Lat (°)	Long (°)	P_{Tx}	γ	P_{L_0}	R_0	σ_G
GN 1	37.42813824	-122.06773522	-10	2.5	11	6	3
GN 2	37.42689551	-122.06009969	-10	2	20	10	3
GN 3	37.42431059	-122.06179260	-10	2.5	-5	3	3
GN 4	37.42666585	-122.06793878	-10	2.5	12	16	3
GN 5	37.42752530	-122.06432984	-10	2.7	10	10	3

Fig. 7(a) shows the noiseless synthetic measurements based on GPS logdata in the lower half of the figure, as provided by the different distributed wireless network transmitters. The upper half of the figure correlates the measurements with the range. Beyond 100m range, the RSSI magnitude levels are very close irrespective of the distance. Fig. 7(b) shows the true measurements as recorded by the receiver modem on board the moving vehicle, which are contaminated by noise.

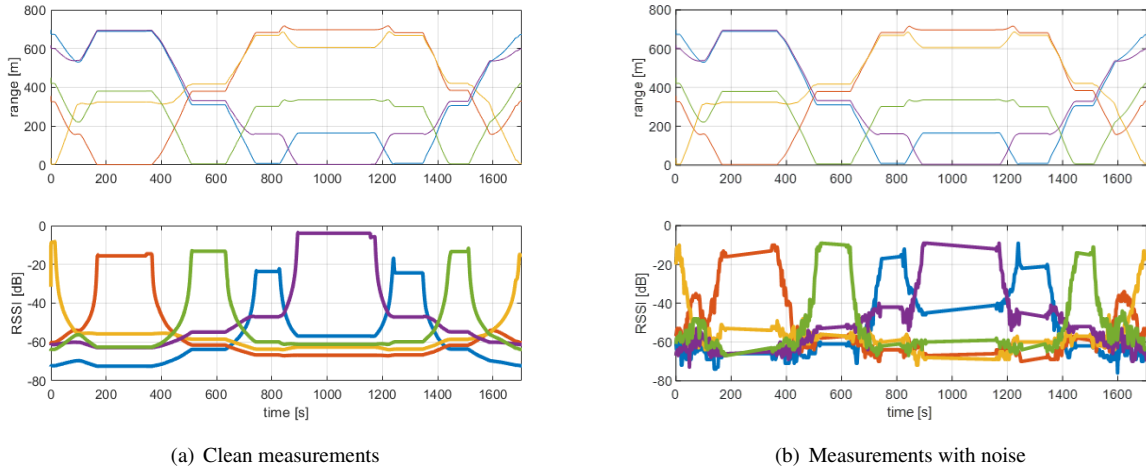


Fig. 7 Comparison RSSI measurements with and without disturbances for field test

Comparing Fig. 7(a) and 7(b) shows the impact of the noise on the RSSI signal. Especially at longer distances, the dB difference between the different RF antennas becomes very small and the noise makes them indistinguishable based on dB level alone. It is expected that this noise level for RSSI will be prohibitive for successful localization. One of the areas of interest is determining the maximum noise level and other sensor characteristics that permit position estimates with sufficient accuracy, otherwise combining with other sensor measurements to improve accuracy.

2. Camera image pre-processing for sensor fusion

During camera image pre-processing, azimuth and elevation angles of the tracked object are derived from the raw images provided by the cameras. This camera image pre-processing is discussed in depth in [20]. In summary, the following steps are followed during this procedure.

First of all, the original high resolution image is cropped to a region of interest (ROI) where additional masks are added to hide any additional features that move slightly, such as power lines and trees in the wind, but that are not relevant for the tracking purpose at hand. Next step is background subtraction. This takes place in OpenCV, which learns the background model to reduce static objects from appearing after subtracting frames. After the subtracted image, the moving object(s) is (are) the only one(s) retained, such that all the other static features in the ROI are discarded. Subsequently keypoints are preserved after blob detection on the subtracted image. Also other moving objects, like birds or other air traffic in the distance, show up here as additional keypoints. Current keypoints are compared with previous keypoints and are maintained if there is some overlap. These are also assigned the same project ID number for the data association problem. Sometimes these keypoints overlap with each other and require additional data association work. More details are given in [20].

Figure 8 shows a screenshot of the Ground-Based Vision Tracker (GBVT) during its image processing pipeline process as described above. The top rows show the original, ROI with masks, and subtracted images, respectively. The bottom row shows the current detection and the latest detections, respectively.

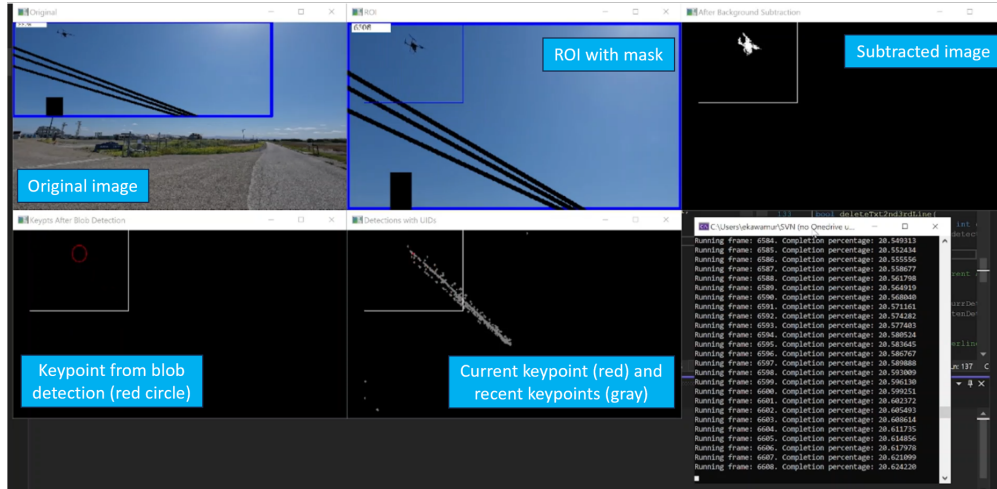


Fig. 8 Pipeline procedure for the Ground-based Vision Tracker (GBVT), source: [20]

Figure 9 shows a high-level block diagram of how the GBVT integrates with the Ground-Based Object Tracker (GBOT) EKF, which is discussed in the remainder of this paper. The black boxes show the main steps for GBVT, and the last black box shows how GBVT feeds azimuth and elevation angles into GBOT.

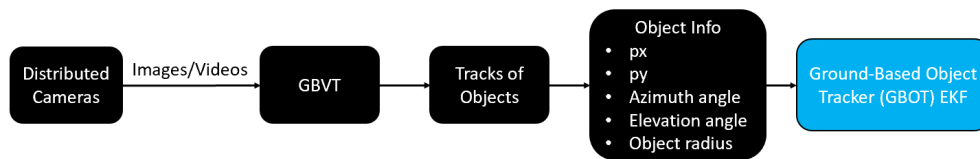


Fig. 9 High-Level Block Diagram: GBVT and GBOT EKF, source: [20]

3. Camera calibration and alignment

For camera data use in the context of localization, it was found that highly accurate location and viewing angle information for each ground node is essential for achieving a sufficiently high accuracy. During data processing, it was found that especially time synchronization between images and viewing angles (bearing and elevation) had a very significant impact on the precision of the track estimates. After calibration, the final ground node values were chosen as given in Tab. 2 and Fig. 31(b) shows the fit between synthetic and real flight data after calibration. It is expected that lens distortion correction will improve the fit further.

IV. Object Tracker Setup

The actual object tracker is a further development which is based on a conventional Kalman Filter as elaborated in detail in Ref. [2, 3, 21]. This filter estimates 3D position and velocity components in ENU coordinates (East–North–Up) and assumes constant speed in the prediction step, and updates the estimate in the correction step based on the available measurements. The filter gain calculation is modified in order to incorporate valid flags for the different sensor measurements, which makes this method adaptive and robust against invalid data, loss of signal between the different distributed sensors, outliers and lower sampling rates. This new implementation integrates the measurements from a (sub)set of EO cameras, radars, wireless transmitters and some other form of distance measuring equipment (DME) in the different locations, whenever each of them is available. This modular architecture is fairly straightforward to extend so that other sensors and sensor locations can be incorporated as well. Fig. 10 illustrates the general overview how the Kalman Filter is set up. The prediction step assumes constant velocity for calculating the predicted state $\hat{\mathbf{x}}_k$ based on the previous state $\mathbf{x}_{k-1}$. The state consists of six elements, namely the three position coordinates and the three velocity coordinates, all in ENU with respect to a predefined fixed referenced point. Predicted measurements $\hat{\mathbf{z}}_{k,i}$ are calculated for each sensor location i , which are a function of the predicted state $\mathbf{h}_i(\hat{\mathbf{x}}_k)$ and the sensor location coordinates. These predicted measurements consist of the measured variables that are provided by the relevant sensors for each location, namely range, azimuth and elevation from the radar (when available) as well as azimuth and elevation from the EO camera. By comparing the actual measurement $\mathbf{z}_{k,i}$ with the predicted measurement $\hat{\mathbf{z}}_{k,i}$ for each sensor location i , one can calculate the innovation $\varepsilon_i = \mathbf{z}_{k,i} - \hat{\mathbf{z}}_{k,i} = \mathbf{z}_{k,i} - \mathbf{h}_i(\hat{\mathbf{x}}_k)$. Main drivers for the innovation are errors in the predicted state (due to the constant speed assumption) and sensor disturbances. Next, the information from all individual sensor locations and sensors is stacked together in one large innovation array ε . As last but not one, the Kalman gain matrix $\mathbf{W}$ is calculated by means of the predicted covariance $\hat{\mathbf{P}}_k$, measurement matrix $\mathbf{H} = \frac{d\mathbf{h}(\mathbf{x})}{d\mathbf{x}}$, and the innovation covariance matrix $\mathbf{S}$. This Kalman gain matrix $\mathbf{W}$ is weighted with valid flags to deal appropriately with the sensor measurements as explained previously. The corrected state $\mathbf{x}_k$ is calculated by means of predicted state $\hat{\mathbf{x}}_k$, the innovation ε and the Kalman gain matrix $\mathbf{W}$ via the following equation: $\mathbf{x}_k = \hat{\mathbf{x}}_k + \mathbf{W} \cdot \varepsilon$. Finally, the corrected covariance $\mathbf{P}_k$ is calculated. The complete procedure is described in more detail below.

A. State Prediction $\hat{\mathbf{x}}_k$

The one step ahead prediction of the state vector is split up as follows:

$$\hat{\mathbf{x}}_k = \begin{bmatrix} \hat{\mathbf{x}}_{k_1} \\ \hat{\mathbf{x}}_{k_2} \end{bmatrix} \quad (2)$$

With the position in the upper half of the vector:

$$\hat{\mathbf{x}}_{k_1} = \begin{bmatrix} \hat{E} \\ \hat{N} \\ \hat{U} \end{bmatrix}_k = \begin{bmatrix} E \\ N \\ U \end{bmatrix}_{k-1} + \Delta t \cdot \begin{bmatrix} \dot{E} \\ \dot{N} \\ \dot{U} \end{bmatrix}_{0_{k-1}} \quad (3)$$

And the velocity in the lower half of the vector:

$$\hat{\mathbf{x}}_{k_2} = \begin{bmatrix} \dot{E} \\ \dot{N} \\ \dot{U} \end{bmatrix}_k = \begin{bmatrix} \dot{E} \\ \dot{N} \\ \dot{U} \end{bmatrix}_{0_{k-1}} \quad (4)$$

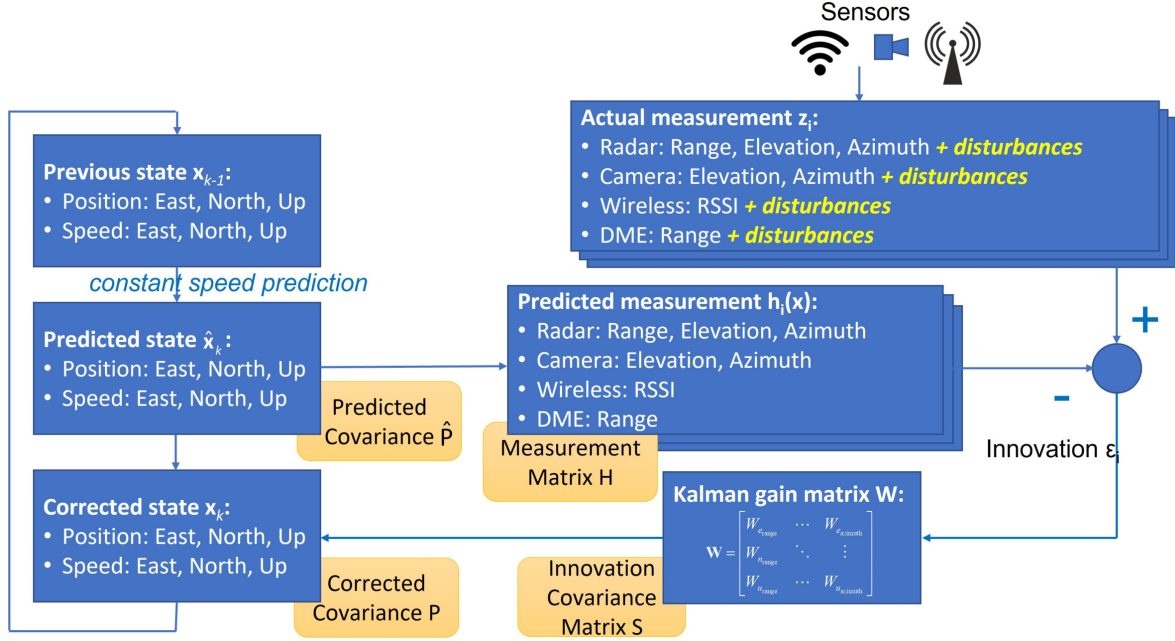


Fig. 10 General overview of the Distributed Sensor Kalman Filter setup

Note that the coordinates are East–North–Up with respect to a predefined fixed reference point and that the velocity is assumed to be close to constant. Moreover:

$$\begin{bmatrix} \dot{E} \\ \dot{N} \\ \dot{U} \end{bmatrix}_{0_{k-1}} = \mathbf{L}^T \cdot \begin{bmatrix} \dot{E} \\ \dot{N} \\ \dot{U} \end{bmatrix}_{T_{k-1}} \quad (5)$$

where $\mathbf{L}^T$ is the rotation matrix from the target ENU-frame to the reference point ENU-frame, as elaborated in the Appendix. Next step is to rotate from ENU to ECEF reference frame:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{0_{\text{ECEF}}} = \mathbf{T} \begin{bmatrix} E \\ N \\ U \end{bmatrix}_{0_{k-1}} \quad (6)$$

where $\mathbf{T}$ is the rotation matrix from the ENU reference frame to the ECEF reference frame, as defined in the Appendix. The predicted target position is converted straight from ECEF to WGS-84 (i.e. latitude, longitude and altitude) via the transformation elaborated in the appendix, resulting in the predicted target position coordinates $\begin{bmatrix} \phi_t & \lambda_t & h_t \end{bmatrix}_{\text{pred}}^T$.

B. State Error Covariance Prediction $\hat{\mathbf{P}}_k$

$$\hat{\mathbf{P}}_k = \mathbf{F} \mathbf{P}_{k-1} \mathbf{F}^T + \mathbf{Q} \quad (7)$$

with:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ \mathbf{0}_{3 \times 3} & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{Q} = \mathbf{G} \mathbf{G}^T \quad (8)$$

and:

$$\mathbf{G} = \begin{bmatrix} 0.5\Delta t^2 & 0 & 0 \\ 0 & 0.5\Delta t^2 & 0 \\ 0 & 0 & 0.5\Delta t^2 \\ \Delta t & 0 & 0 \\ 0 & \Delta t & 0 \\ 0 & 0 & \Delta t \end{bmatrix} \quad \mathbf{\Gamma} = \begin{bmatrix} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_n^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{bmatrix} \quad (9)$$

C. Predicted Measurement Calculation $\hat{\mathbf{z}}_k$

The total predicted measurement vector $\hat{\mathbf{z}}_k$ consists of the predicted measurement vectors $\hat{\mathbf{z}}_{k,i}$ of each individual sensor location i . The predicted measurement vectors contain the relevant equipment-dependent information per individual sensor location and can include:

- EO camera: elevation α and azimuth γ (no range due to the lack of stereo view and depth perception)
- radar: range ρ , elevation α and azimuth γ
- wireless network (RF): residual signal strength inference $RSSI$ as defined in Eq. (1).
- DME: range ρ

The predicted measurement vector per sensor location is described as follows with only the hardware-relevant parts present in order to reduce vector size and computational load:

$$\hat{\mathbf{z}}_{k,i} = \left[\hat{\mathbf{z}}_{k,i_{\text{EOcam}}} \quad \hat{\mathbf{z}}_{k,i_{\text{radar}}} \quad \hat{\mathbf{z}}_{k,i_{\text{RF}}} \quad \hat{\mathbf{z}}_{k,i_{\text{DME}}} \right]^T \quad (10)$$

$$\hat{\mathbf{z}}_{k,i_{\text{EOcam}}} = \begin{bmatrix} \hat{\alpha}_i & \hat{\gamma}_i \end{bmatrix}^T \quad (11)$$

$$\hat{\mathbf{z}}_{k,i_{\text{radar}}} = \begin{bmatrix} \hat{\rho}_i & \hat{\alpha}_i & \hat{\gamma}_i \end{bmatrix}^T \quad (12)$$

$$\hat{\mathbf{z}}_{k,i_{\text{RF}}} = \left[R\hat{SSI}_i \right]^T \quad (13)$$

$$\hat{\mathbf{z}}_{k,i_{\text{DME}}} = \left[\hat{\rho}_i \right]^T \quad (14)$$

The individual measured variables are defined as follows:

$$\hat{\rho}_i = \sqrt{\hat{X}_{S_i}^2 + \hat{Y}_{S_i}^2 + \hat{Z}_{S_i}^2} = \sqrt{\hat{E}_i^2 + \hat{N}_i^2 + \hat{U}_i^2} \quad (15)$$

$$\hat{\alpha}_i = \arctan \frac{-\hat{Z}_{S_i}}{\hat{\rho}_{i_{\text{hor}}}} \quad (16)$$

$$\hat{\gamma}_i = \arctan \frac{\hat{Y}_{S_i}}{\hat{X}_{S_i}} \quad (17)$$

$$R\hat{SSI}_i = P_{Tx} - P_{L_0} - 10\gamma \log_{10} \left(\frac{\hat{\rho}_i}{R_0} \right) \quad (18)$$

where:

$$\hat{\rho}_{i_{\text{hor}}} = \sqrt{\hat{X}_{S_i}^2 + \hat{Y}_{S_i}^2} \quad (19)$$

and:

$$\begin{bmatrix} \hat{X}_{S_i} \\ \hat{Y}_{S_i} \\ \hat{Z}_{S_i} \end{bmatrix} = \mathbf{M}_i^T \cdot \begin{bmatrix} \hat{E}_i \\ \hat{N}_i \\ \hat{U}_i \end{bmatrix}_k \quad (20)$$

where $\mathbf{M}_i^T$ is the rotation matrix from the i^{th} sensor-based ENU reference frame to the sensor fixed reference frame of the i^{th} sensor, as defined in the Appendix. It rotates the East-North-Up coordinates for each sensor location towards the sensor orientation (viewing direction of camera (and radar if applicable)). These local ENU_i coordinates are calculated from the aforementioned predicted state in WGS-84 coordinates $\begin{bmatrix} \phi_t & \lambda_t & h_t \end{bmatrix}_{\text{pred}}^T$ via the transformation elaborated in

the appendix, resulting in the predicted target position coordinates in ECEF, and a subsequent rotation from ECEF to the local sensor's ENU as follows:

$$\begin{bmatrix} \hat{E}_i \\ N_i \\ U_i \end{bmatrix}_k = \mathbf{T}_i^\top \begin{bmatrix} \hat{X} \\ Y \\ Z \end{bmatrix}_{0_{\text{ECEF}}} \quad (21)$$

where $\mathbf{T}_i^\top$ is the rotation matrix from the ECEF reference frame to the local sensor ENU reference frame, as defined in the Appendix.

D. Kalman Filter Gain W

The following steps are performed in order to calculate the gain of the filter:

- 1) Calculate measurement matrix $\mathbf{H}$
- 2) Calculate innovation covariance matrix $\mathbf{S}$
- 3) Calculate gain matrix $\mathbf{W}$

Each step will be elaborated in detail next:

1. Measurement matrix $\mathbf{H}$

The measurement matrix $\mathbf{H}$ is basically the Jacobian of the predicted measurement vector:

$$\hat{\mathbf{z}}_k = \begin{bmatrix} \hat{\mathbf{z}}_{k,1} & \dots & \hat{\mathbf{z}}_{k,i} & \dots & \hat{\mathbf{z}}_{k,N} \end{bmatrix}^\top$$

with N the number of sensor locations and each individual sensor location based component defined as follows with only the hardware-relevant parts present in order to reduce matrix size and computational load:

$$\hat{\mathbf{z}}_{k,i} = \begin{bmatrix} \hat{\mathbf{z}}_{k,i_{\text{EOcam}}} & \hat{\mathbf{z}}_{k,i_{\text{radar}}} & \hat{\mathbf{z}}_{k,i_{\text{RF}}} & \hat{\mathbf{z}}_{k,i_{\text{DME}}} \end{bmatrix}^\top = \begin{bmatrix} \hat{\mathbf{h}}_{\text{EOcam}_i}(\hat{\mathbf{x}}_k) & \hat{\mathbf{h}}_{\text{radar}_i}(\hat{\mathbf{x}}_k) & \hat{\mathbf{h}}_{\text{RF}_i}(\hat{\mathbf{x}}_k) & \hat{\mathbf{h}}_{\text{DME}_i}(\hat{\mathbf{x}}_k) \end{bmatrix}^\top$$

with respect to the state vector $\hat{\mathbf{x}}_k = \begin{bmatrix} \hat{\mathbf{x}}_{k_1} & \hat{\mathbf{x}}_{k_2} \end{bmatrix}^\top$, with $\mathbf{x}_{k_1}$ being the position components and $\mathbf{x}_{k_2}$ the velocity components, both in the ENU reference frame. The measurement matrix $\mathbf{H}$ basically consists of the blocks $\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \dots & \mathbf{H}_i & \dots & \mathbf{H}_N \end{bmatrix}^\top$ with N the number of sensors and decomposing each $\mathbf{H}_i$ in 8 blocks results in the following structure:

$$\mathbf{H}_i(\mathbf{x}_k) = \frac{d\hat{\mathbf{h}}_i(\hat{\mathbf{x}}_k)}{d\hat{\mathbf{x}}_k} = \begin{bmatrix} \frac{d\hat{\mathbf{h}}_{\text{EOcam}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_1}} & \frac{d\hat{\mathbf{h}}_{\text{EOcam}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_2}} = \mathbf{0}_{2 \times 3} \\ \frac{d\hat{\mathbf{h}}_{\text{radar}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_1}} & \frac{d\hat{\mathbf{h}}_{\text{radar}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_2}} = \mathbf{0}_{3 \times 3} \\ \frac{d\hat{\mathbf{h}}_{\text{RF}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_1}} & \frac{d\hat{\mathbf{h}}_{\text{RF}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_2}} = \mathbf{0}_{1 \times 3} \\ \frac{d\hat{\mathbf{h}}_{\text{DME}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_1}} & \frac{d\hat{\mathbf{h}}_{\text{DME}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_2}} = \mathbf{0}_{1 \times 3} \end{bmatrix} \quad (22)$$

The measurements depend only on the position components and thus $\mathbf{x}_{k_1}$. All nonzero blocks in Eq. (22) are elaborated next:

$$\frac{d\hat{\mathbf{h}}_{\text{EOcam}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_1}} = \begin{bmatrix} \frac{d\hat{E}_i}{d\hat{E}_i} & \frac{d\hat{N}_i}{d\hat{N}_i} & \frac{d\hat{U}_i}{d\hat{U}_i} \end{bmatrix} = \begin{bmatrix} \frac{d\hat{E}_i}{dX_{S_i}} & \frac{d\hat{N}_i}{dY_{S_i}} & \frac{d\hat{U}_i}{dZ_{S_i}} \end{bmatrix} \cdot \mathbf{M}^\top = \begin{bmatrix} \frac{\hat{X}_{S_i} \cdot \hat{Z}_{S_i}}{\hat{\rho}_{\text{hor}_0} \hat{\rho}_0^2} & \frac{\hat{Y}_{S_i} \cdot \hat{Z}_{S_i}}{\hat{\rho}_{\text{hor}_0} \hat{\rho}_0^2} & \frac{-\hat{\rho}_{i_{\text{hor}}}}{\hat{\rho}_0^2} \\ \frac{-\hat{Y}_{S_i}}{\hat{\rho}_{\text{hor}_0}^2} & \frac{\hat{X}_{S_i}}{\hat{\rho}_{\text{hor}_0}^2} & 0 \end{bmatrix} \cdot \mathbf{M}^\top \quad (23)$$

The denominators $\hat{\rho}_0$ and $\hat{\rho}_{\text{hor}_0}$ in Eq. (23) are singularity protected as follows:

$$\hat{\rho}_0 = \max(\hat{\rho}_i, \varepsilon_\rho) \quad (24)$$

$$\hat{\rho}_{\text{hor}_0} = \max(\hat{\rho}_{i_{\text{hor}}}, \varepsilon_{\rho_{\text{hor}}}) \quad (25)$$

where ε_ρ and $\varepsilon_{\rho_{\text{hor}}}$ are fairly small positive values to prevent singularities.

$$\frac{d\hat{\mathbf{h}}_{\text{radar}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_1}} = \begin{bmatrix} \frac{d\hat{\rho}_i}{d\hat{E}_i} & \frac{d\hat{\rho}_i}{d\hat{N}_i} & \frac{d\hat{\rho}_i}{d\hat{U}_i} \end{bmatrix} = \begin{bmatrix} \frac{d\hat{\rho}_i}{dX_{S_i}} & \frac{d\hat{\rho}_i}{dY_{S_i}} & \frac{d\hat{\rho}_i}{dZ_{S_i}} \end{bmatrix} \cdot \mathbf{M}^\top = \begin{bmatrix} \frac{\hat{X}_{S_i}}{\hat{\rho}_0} & \frac{\hat{Y}_{S_i}}{\hat{\rho}_0} & \frac{\hat{Z}_{S_i}}{\hat{\rho}_0} \\ \frac{\hat{X}_{S_i} \cdot \hat{Z}_{S_i}}{\hat{\rho}_{\text{hor}_0} \hat{\rho}_0^2} & \frac{\hat{Y}_{S_i} \cdot \hat{Z}_{S_i}}{\hat{\rho}_{\text{hor}_0} \hat{\rho}_0^2} & \frac{-\hat{\rho}_{i_{\text{hor}}}}{\hat{\rho}_0^2} \\ \frac{-\hat{Y}_{S_i}}{\hat{\rho}_{\text{hor}_0}^2} & \frac{\hat{X}_{S_i}}{\hat{\rho}_{\text{hor}_0}^2} & 0 \end{bmatrix} \cdot \mathbf{M}^\top \quad (26)$$

$$\begin{aligned}\frac{d\hat{\mathbf{h}}_{\text{RF}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_1}} &= \begin{bmatrix} \frac{dR\hat{S}SI_i}{dE_i} & \frac{dR\hat{S}SI_i}{dN_i} & \frac{dR\hat{S}SI_i}{dU_i} \end{bmatrix} = \frac{dR\hat{S}SI_i}{d\hat{\rho}_i} \begin{bmatrix} \frac{d\hat{\rho}_i}{dX_{S_i}} & \frac{d\hat{\rho}_i}{dY_{S_i}} & \frac{d\hat{\rho}_i}{dZ_{S_i}} \end{bmatrix} \cdot \mathbf{M}^\top \\ &= -10\gamma \frac{1}{\hat{\rho}_i \ln(10)} \begin{bmatrix} \frac{X_{S_i}}{\hat{\rho}_0} & \frac{Y_{S_i}}{\hat{\rho}_0} & \frac{Z_{S_i}}{\hat{\rho}_0} \end{bmatrix} \cdot \mathbf{M}^\top\end{aligned}\quad (27)$$

$$\frac{d\hat{\mathbf{h}}_{\text{DME}_i}(\hat{\mathbf{x}}_{k_1})}{d\hat{\mathbf{x}}_{k_1}} = \begin{bmatrix} \frac{d\hat{\rho}_i}{dE_i} & \frac{d\hat{\rho}_i}{dN_i} & \frac{d\hat{\rho}_i}{dU_i} \end{bmatrix} = \begin{bmatrix} \frac{d\hat{\rho}_i}{dX_{S_i}} & \frac{d\hat{\rho}_i}{dY_{S_i}} & \frac{d\hat{\rho}_i}{dZ_{S_i}} \end{bmatrix} \cdot \mathbf{M}^\top = \begin{bmatrix} \frac{\hat{X}_{S_i}}{\hat{\rho}_0} & \frac{\hat{Y}_{S_i}}{\hat{\rho}_0} & \frac{\hat{Z}_{S_i}}{\hat{\rho}_0} \end{bmatrix} \cdot \mathbf{M}^\top \quad (28)$$

2. Innovation covariance matrix $\mathbf{S}$

The innovation covariance matrix is calculated as follows:

$$\mathbf{S} = \mathbf{H}\hat{\mathbf{P}}_k\mathbf{H}^\top + \mathbf{R}_{\text{sensors}} \quad (29)$$

where: $\mathbf{R}_{\text{sensors}} = \text{diag} [\mathbf{R}_{\text{sensor}_1} \dots \mathbf{R}_{\text{sensor}_i} \dots \mathbf{R}_{\text{sensor}_N}]$ with the diagonal elements per sensor location defined as follows:

$$\mathbf{R}_{\text{sensor}_i} = \text{diag} \left[\sigma_{\alpha_{\text{EOcam}_i}}^2 \quad \sigma_{\gamma_{\text{EOcam}_i}}^2 \quad \sigma_{\rho_{\text{radar}_i}}^2 \quad \sigma_{\alpha_{\text{radar}_i}}^2 \quad \sigma_{\gamma_{\text{radar}_i}}^2 \quad \sigma_{R\hat{S}SI_{\text{RF}_i}}^2 \quad \sigma_{\rho_{\text{DME}_i}}^2 \right]$$

where only the hardware relevant elements per sensor location are kept. A diagonal matrix of sensor valid flags is defined similarly, $\mathbf{M}_{\text{sensors}_{\text{valid}}} = \text{diag} [\mathbf{M}_{\text{sensors}_{1\text{valid}}} \dots \mathbf{M}_{\text{sensors}_{i\text{valid}}} \dots \mathbf{M}_{\text{sensors}_{N\text{valid}}}]$ with:

$$\mathbf{M}_{\text{sensors}_{i\text{valid}}} = \text{diag} [\mu_{\text{EOcam}_i} \quad \mu_{\text{EOcam}_i} \quad \mu_{\text{radar}_i} \quad \mu_{\text{radar}_i} \quad \mu_{\text{radar}_i} \quad \mu_{\text{RF}_i} \quad \mu_{\text{DME}_i}]$$

These valid flags $\mu_{\bullet_i}$ are one for valid sensor readings or zero for invalid ones. These flags allow to take into account lower sampling rates for specific sensors, signal dropouts (e.g. object is outside field of view), sensor malfunctions, invalid signals, etc. This setup assumes that there is a monitoring algorithm in place that drives the valid flag values accordingly. This valid flag matrix is needed for calculating the inverse of the innovation covariance matrix:

$$\mathbf{S}_{\text{valid}}^{-1} = \mathbf{M}_{\text{sensors}_{\text{valid}}} \cdot \mathbf{S}^{-1} \cdot \mathbf{M}_{\text{sensors}_{\text{valid}}} \quad (30)$$

3. Gain matrix $\mathbf{W}$

The Kalman filter gain matrix is calculated with all the aforementioned information as follows:

$$\mathbf{W} = \hat{\mathbf{P}}_k\mathbf{H}^\top\mathbf{S}_{\text{valid}}^{-1} \quad (31)$$

$$\mathbf{W}_{\text{valid}} = \mathbf{W}\mathbf{M}_{\text{sensors}_{\text{valid}}} \quad (32)$$

Invalid flags in $\mathbf{M}_{\text{sensors}_{\text{valid}}}$ will result in zero columns in $\mathbf{W}_{\text{valid}}$. As a consequence, that sensor information will not be used for the correction step. This means that the prediction step will have a higher relative importance for the states that are the main drivers of these invalid measurements, except when there are other valid sensor measurements that contain significant information of these states. When valid measurements are sparse, the Kalman Filter will extrapolate a trajectory that will gradually diverge from reality. This growing uncertainty is indicated by the rank and/or condition number of the observability matrix, as well as by the standard deviation of the estimates.

E. State Update $\mathbf{x}_k$

The state update step relies on the total sensor measurements vector covering all sensor locations, which will give only measurement values for the locations where the tracked object is in the field of view.

The innovation is calculated first, which is the difference between the predicted measurements $\hat{\mathbf{z}}_k$ and the actual sensor measurements $\mathbf{z}_k^*$, including sensor limitations and perturbations:

$$\boldsymbol{\varepsilon} = \mathbf{z}_k - \hat{\mathbf{z}}_k = \mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_k) \quad (33)$$

This innovation $\boldsymbol{\varepsilon}$ is then used to calculate the state update by means of the previously computed Kalman Gain matrix:

$$\mathbf{x}_k = \hat{\mathbf{x}}_k + \mathbf{W}_{\text{valid}} \cdot \boldsymbol{\varepsilon} \quad (34)$$

*In the numerical calculations, an 'empty' measurement scalar in $\mathbf{z}_k$ is usually represented by a 'NaN'. In order to prevent numerical issues these have to be replaced by zeros together with their respective valid flag being set to zero in $\mathbf{W}_{\text{valid}}$ so that they are ignored.

F. Geodetic Coordinates Calculation

Next step is calculating the absolute position of the target in WGS-84 coordinates. This is achieved by combining the known sensor position and the relative position of the target with respect to the sensor position, which is the first half of the corrected state vector $\mathbf{x}_{k_1}$. Since both are defined in different ENU reference frames, they are first converted to the ECEF (earth centered earth fixed) reference frame, then added together and converted back to WGS-84.

The sensor position is converted straight from WGS-84 (i.e. latitude, longitude and altitude) to ECEF via the transformation elaborated in the appendix, resulting in the position coordinates $[X \ Y \ Z]_{\text{ECEF}}^T$. The relative position of the target with respect to the sensor position in ECEF coordinates is calculated as follows:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{(S \rightarrow T)_{\text{ECEF}}} = \mathbf{T} \begin{bmatrix} E \\ N \\ U \end{bmatrix}_k = \mathbf{T} \mathbf{x}_{k_1} \quad (35)$$

where $\mathbf{T}$ is the rotation matrix from the ENU reference frame to the ECEF reference frame, as defined in the Appendix.

The absolute position of the target in ECEF coordinates is then calculated as follows:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{T_{\text{ECEF}}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{ECEF}} + \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{(S \rightarrow T)_{\text{ECEF}}} \quad (36)$$

As the last step, this absolute target position is converted back to WGS-84 (i.e. latitude, longitude and altitude) via the reverse transformation elaborated in the Appendix.

G. Velocity Update

The velocity components of the target, which are the second half of the corrected state vector $\mathbf{x}_{k_2}$ are defined with respect to the sensor ENU reference frame, these are transformed to the target ENU reference frame as follows:

$$\begin{bmatrix} \dot{E} \\ \dot{N} \\ \dot{U} \end{bmatrix}_{k_T} = \mathbf{L}^T \begin{bmatrix} \dot{E} \\ \dot{N} \\ \dot{U} \end{bmatrix}_k = \mathbf{L}^T \mathbf{x}_{k_2} \quad (37)$$

H. State Error Covariance Update $\mathbf{P}_k$

By first defining:

$$\mathbf{W}_H = \mathbf{I}_{6 \times 6} - \mathbf{W}_{\text{valid}} \mathbf{H} \quad (38)$$

where $\mathbf{I}_{6 \times 6}$ is a 6-by-6 identity matrix, the state error covariance matrix is then updated in the reference point reference frame:

$$\mathbf{P}_k = \mathbf{W}_H \hat{\mathbf{P}}_k \mathbf{W}_H^T + \mathbf{W}_{\text{valid}} \mathbf{R}_{\text{sensors}} \mathbf{W}_{\text{valid}}^T \quad (39)$$

which is finally converted to the target reference frame via the aforementioned rotation matrix $\mathbf{L}$:

$$\mathbf{P}_{k_T} = \begin{bmatrix} \mathbf{L} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{L} \end{bmatrix} \mathbf{P}_k \quad (40)$$

V. Results and Discussion

This results section consists of two main parts, namely the simulation based evaluations in Sec. V.A and the field test based evaluations in Sec. V.B. The general outline of these scenarios is provided in Sec. II and Sec. III gives more details about the sensor characteristics.

A. Simulation based evaluations

Three different sensor configurations were considered in the simulations, namely RF equipment that allows for RSSI measurements in Sec. V.A.1, some kind of DME equipment like radar that provides direct range measurements in Sec. V.A.2, and EO cameras that provide two angles of the line of sight in Sec. V.A.3. This study focuses on similar sensor equipment in all ground nodes per simulated scenario. Future work will also consider dissimilar sensor equipment in the various ground nodes and its impact on the accuracy of the estimates.

1. RSSI observations only

The first simulation based evaluation scenario considers the RF antennas and their RSSI signals (residual signal strength inference) at the receiver's end, which provide some information about the distance. Sensed signal characteristics for RF equipment are described in Sec. III.A.

Fig. 11(a) shows the general outline of the scenario and illustrates that the vehicle receives RF signals from all 4 antennas at all times since they have an omnidirectional pattern. In this synthetic simulated environment, it was assumed that 4 signals would be sufficient for localization in the absence of any shielding by obstacles. Fig. 11(b) compares the estimated path with the true path. In general, the estimate seems to track the true path more or less, but deviates significantly resulting in lower accuracy, among others because of the nature of the sensed signal which is shown next.

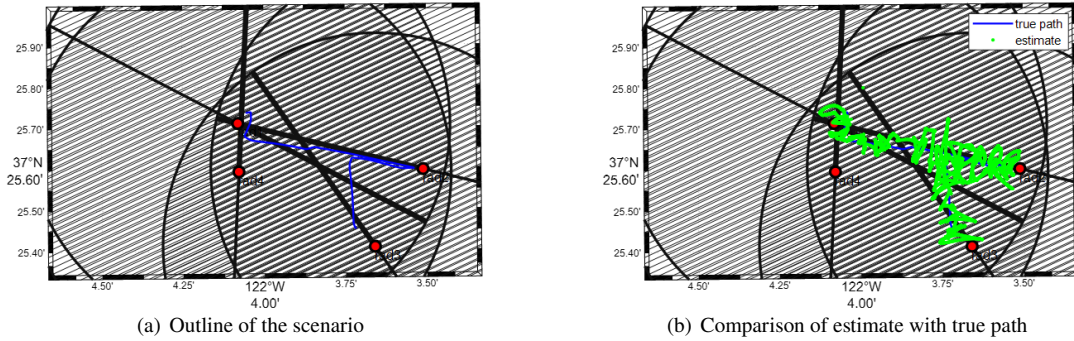


Fig. 11 Outline of scenario and comparison of estimate and true path for RSSI observations only

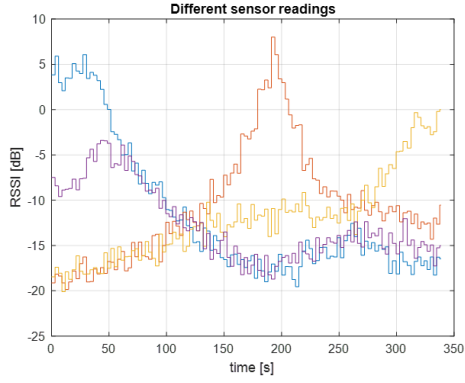
Fig. 12(a) shows the individual sensor readings as simulated in this scenario. The 3s sampling rate is very apparent in this graph. Also, while analyzing this graph combined with Eq. 1, it is clear that this noise level will have a major impact on the signal-to-noise ratio over longer distances because of the logarithmic nature of the signal. At specific time instants, the vehicle is close to one of the antennas but still far from the other three, making trilateration challenging with this level of measurement noise. Fig. 12(b) shows the time varying condition number of the observation matrix. The values are not particularly high, but not as consistently low as can be seen with good coverage by multiple cameras, for example as shown in Fig. 22(b). The rank of the observability matrix is consistently equal to 6, which is the full rank.

Fig. 13 compares the estimated and true flight path in 3D in the East-North-Up reference frame. Fig. 13(a) shows the 3D view and Fig. 13(b) shows the top view. Both show the predicted as well as the corrected state as calculated in the Kalman Filter. This behavior confirms the impact of the signal-to-noise ratio on the estimated track, especially at larger ranges.

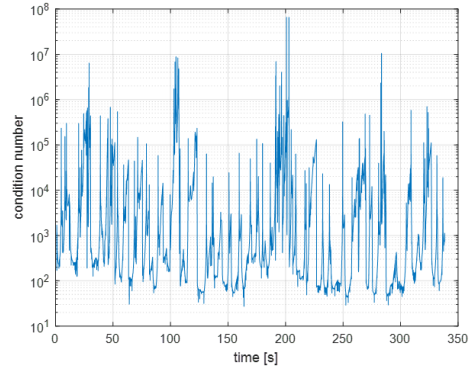
Fig. 14(a) shows the standard deviations. From top to bottom triplets East - North - Up are given for the predicted position, corrected position and corrected velocity components. Fig. 14(b) shows the fit error for the 3D position, from top to bottom East - North - Up. Both graphs show a large fit error and low confidence in the estimated values.

Fig. 15 shows the total 3 dimensional fit error between estimated and true position. The average fit error is around 60 m but with instantaneous spikes up to more than 200 m.

Use of RSSI only as a multi ground node measurement for object tracking and localization via trilateration turns out to be inaccurate with the current sensor fusion techniques and sensor characteristics. Possible candidates for future research are including the time of flight of the RF signal between transmitter and receiver in the signal processing setup, and possibly combining them with other dissimilar measurements like radar and/or EO cameras.

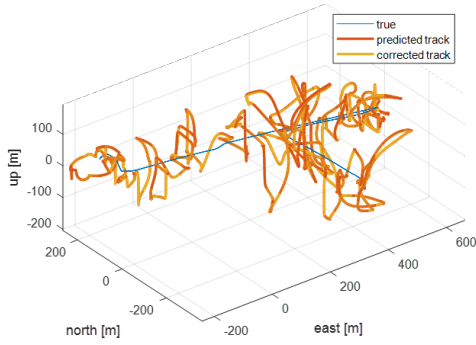


(a) Individual sensor readings

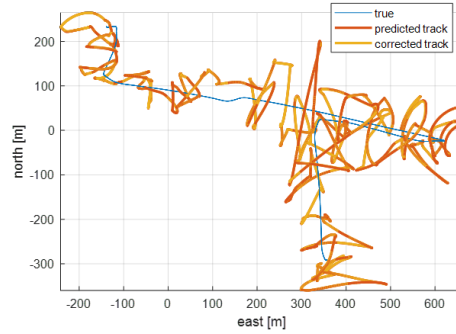


(b) Condition number of the observability matrix

Fig. 12 Individual sensor readings and condition number for RSSI observations only

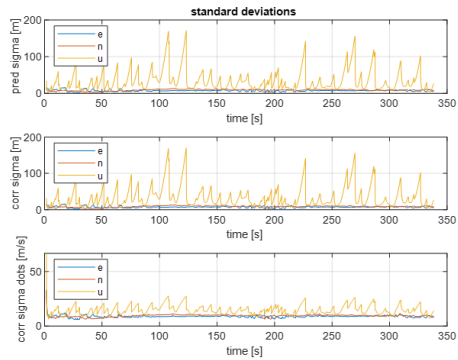


(a) 3D view

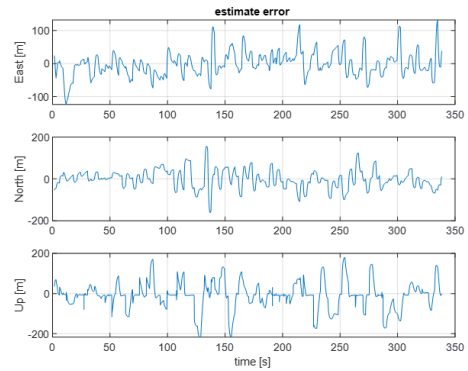


(b) top view

Fig. 13 Comparison between true path and estimate for RSSI observations only



(a) standard deviations



(b) fit error

Fig. 14 Error information for RSSI observations only

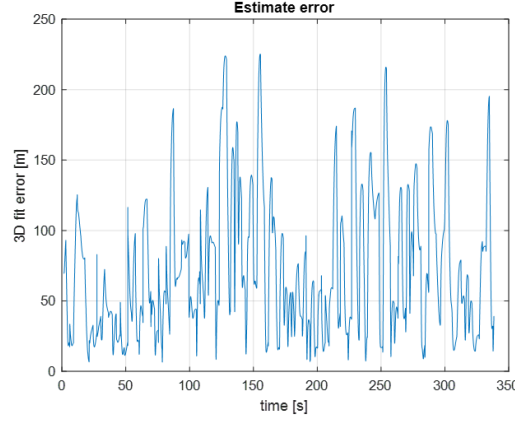


Fig. 15 Total fit error for RSSI observations only

2. DME observations only

As an alternative method for trilateration, DME (distance measuring equipment) measurements were used in this scenario. Assuming some type of radar equipment that only provides a direct noise-contaminated measurement of range, one can compare this performance with the logarithmically scaled RF signals. Another important observation is that the noise level for the radar signals is significantly lower than for the RF equipment, as detailed in Sec. III.

Fig. 16(a) shows the general outline of the scenario and illustrates that the vehicle receives DME signals from all 4 antennas at all times since they have an omnidirectional pattern. In this synthetic simulated environment, it was assumed that 4 signals would be sufficient for localization in the absence of any shielding by obstacles. Fig. 16(b) compares the estimated path with the true path. In general, the estimate appears to track the true path accurately on the map, despite the sensor noise.

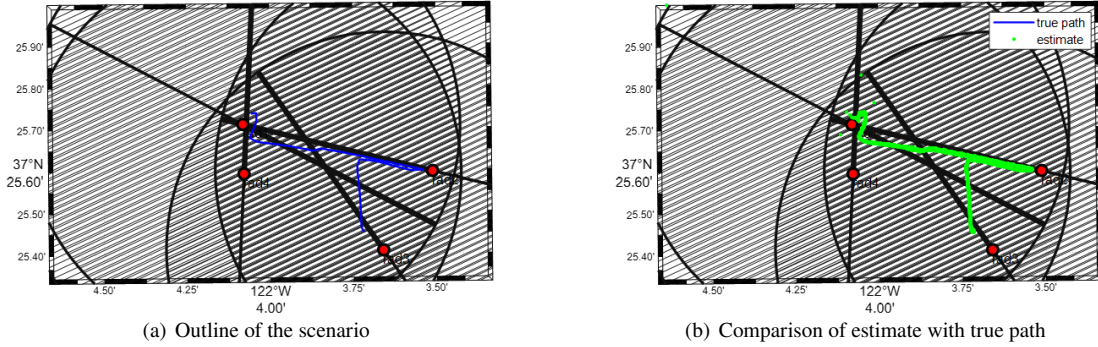
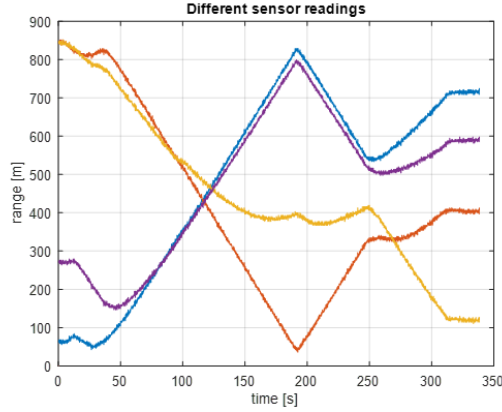


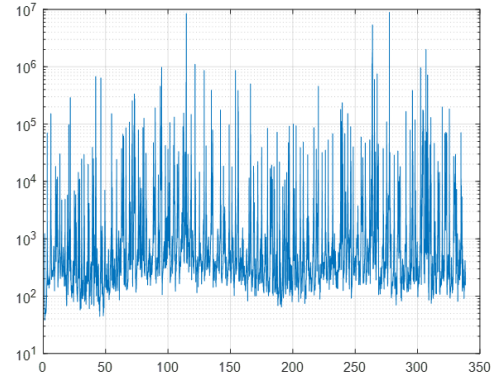
Fig. 16 Outline of scenario and comparison of estimate and true path for DME observations only

Fig. 17(a) shows the individual sensor readings as simulated in this scenario. The DME signal-to-noise ratio is significantly better compared to the RSSI signals in Fig. 12(a) and the former measurement also has a higher sampling rate than the latter. Fig. 17(b) shows the time varying condition number of the observation matrix. The values are in the same order of magnitude as for the RSSI signals shown in Fig. 12(b). The rank of the observability matrix is consistently equal to 6, which is the full state rank.

Fig. 18 compares the estimated and true flight path in 3D in the East-North-Up reference frame. Fig. 18(a) shows the 3D view and Fig. 18(b) shows the top view. Both show the predicted as well as the corrected state as calculated in the Kalman Filter. These figures show that the estimate is mostly accurate in the East-North plane, but not as accurate vertically. As illustrated in Fig. 1, trilateration uses the concept of intersecting spheres, and when all ground nodes as well as the tracked object are situated in a narrow band around the same East - North plane, then noise will have a disproportional large impact perpendicular on that plane.

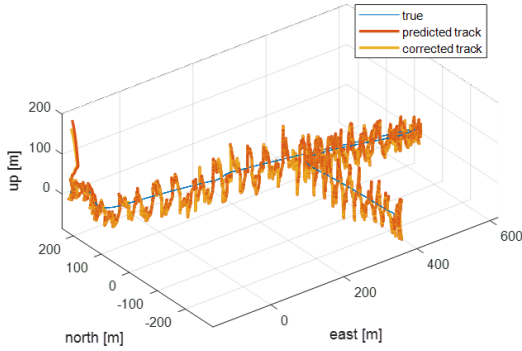


(a) Individual sensor readings

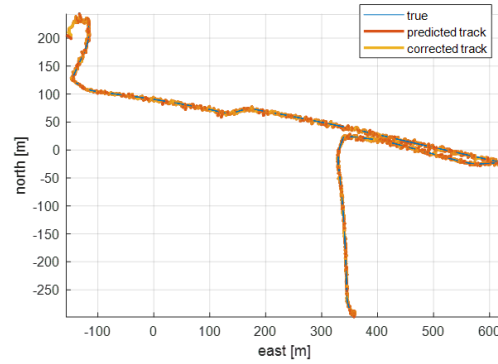


(b) Condition number of the observability matrix

Fig. 17 Individual sensor readings and condition number for DME observations only



(a) 3D view



(b) top view

Fig. 18 Comparison between true path and estimate for DME observations only

Fig. 19(a) shows the standard deviations in a similar way as in the previous scenario. The vertical component has a significantly higher standard deviation value compared to the two horizontal ones. Fig. 19(b) shows the error for the 3D position, from top to bottom East - North - Up. Also here the estimate error for the 'Up' component is much larger than for 'North' and 'East'. Both graphs confirm a large fit error in the vertical component and low confidence in its estimated values. The horizontal components are much more accurate with higher confidence.

Fig. 20 shows the total 3 dimensional fit error between estimated and true position. The average fit error is around 17 m but with some instantaneous spikes up to more than 70 m.

There are two possibilities for improving the accuracy of the vertical component of the estimate. One can either add an additional sensor out of the horizontal plane (for example on a taller structure above the terrain), or otherwise add an additional measurement per ground node, like for example elevation measurements, to the DME equipment.

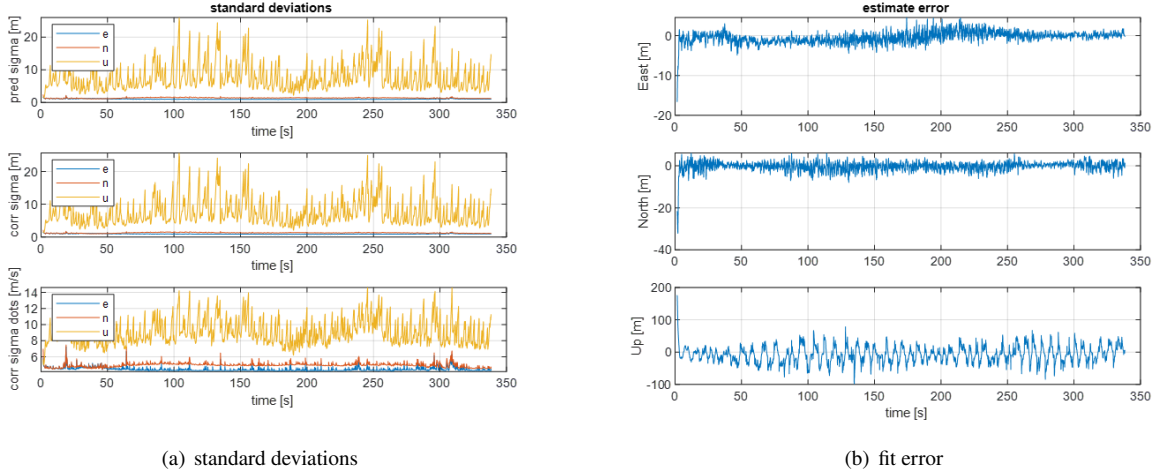


Fig. 19 Error information for DME observations only

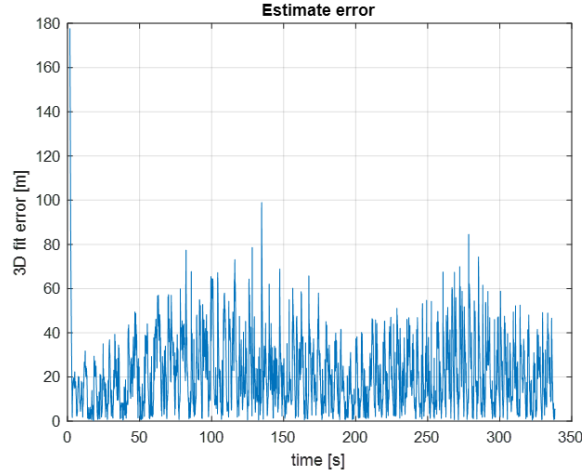


Fig. 20 Total fit error for DME observations only

3. Camera observations only

In the last simulated scenario, the object tracker relies on camera observations only. These provide two dimensional angles of the tracked object by means of azimuth and elevation. Localization of the tracked object is achieved by means of triangulation. No range information is used in this setup, but precise information about the location and viewing angles of each ground node has a major impact on the accuracy of the estimate.

Fig. 21(a) shows the general outline of the scenario and illustrates how the different ground nodes were placed and oriented strategically to achieve optimal coverage of the flight path with the number of ground nodes available. In this synthetic simulated environment, in theory 2 signals should be sufficient for localization in the absence of any shielding by obstacles or signal deterioration due to increased observation ranges. Fig. 21(b) compares the estimated path with the true path. In general, the estimate appears to track the true path very accurately on the map, partly thanks to the low sensor noise.

For the analysis of the observability matrix, Fig. 22(a) shows its rank and Fig. 22(b) shows its condition number. The matrix has full state rank and its condition number is multiple orders of magnitude smaller than for the RSSI and DME measurements in Fig. 12(b) and 17(b). However, the condition number increases significantly temporarily in the time intervals around 180 - 200s and after approximately 310s.

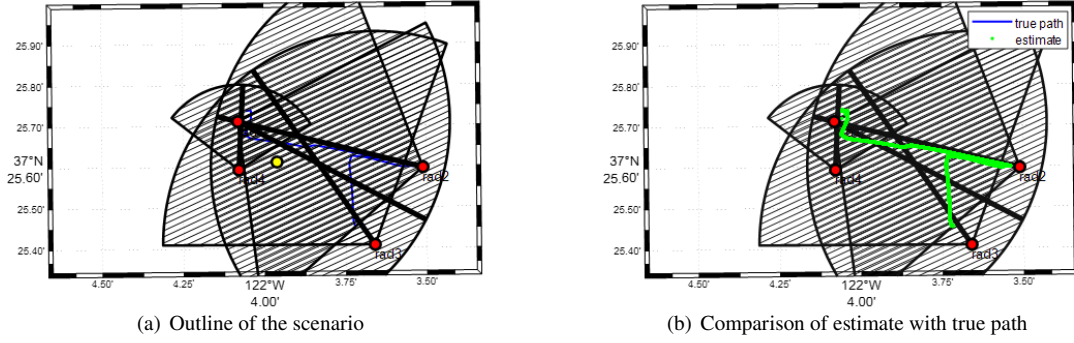


Fig. 21 Outline of scenario and comparison of estimate and true path for camera observations only

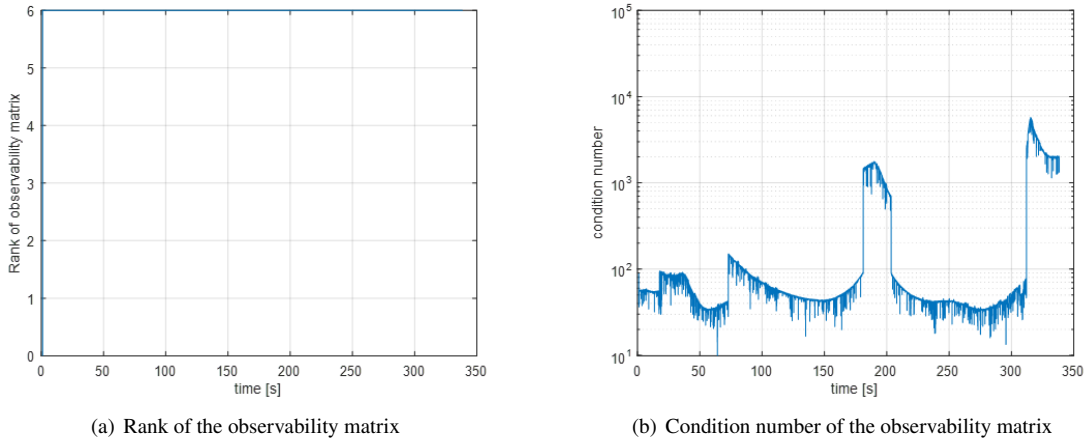
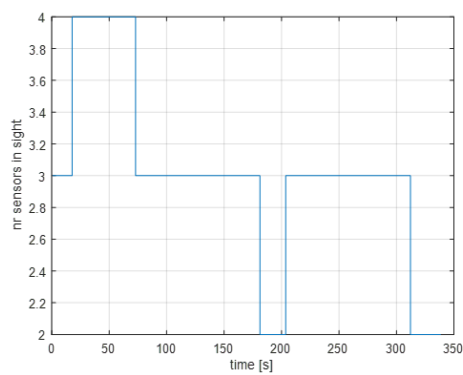


Fig. 22 Rank and condition number for camera observations only

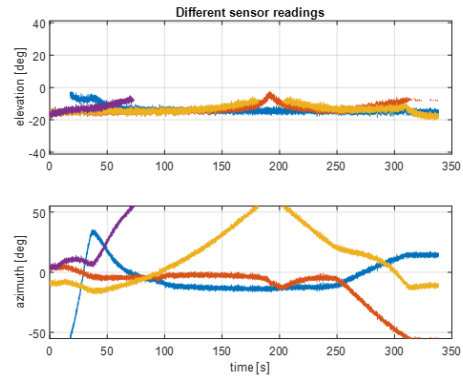
Fig. 23 gives additional information about the sensor readings. Fig. 23(a) shows how many ground nodes have the tracked object in sight at any given time instant of the flight. In other words, this figure shows how many sensor cones, as drawn in Fig. 21(a), overlap with the location of the tracked object at any time instant during the flight. Comparing Fig. 23 and Fig. 22(b) shows that the time instants when the condition number temporarily increases an order of magnitude coincide with the moments when there are temporarily only two ground nodes tracking the object instead of three. Fig. 23(b) shows the measured elevation and azimuth angles of the tracked drone by each ground node over time. These sensor readings show a good signal-to-noise ratio, like in Fig. 17(a) for the DME.

Fig. 24 compares the estimated and true flight path in 3D in a similar way as for the previous sensor types. These figures show that the estimate is very accurate along all axes on the relative scale of the entire flight path.

Fig. 25(a) shows the standard deviations in a similar way as in the previous scenarios. At least one component of the standard deviations has a significantly higher value when there are only two ground nodes tracking the target compared to three or more. Fig. 25(b) shows the fit error for the 3D position, from top to bottom East - North - Up. These graphs show high accuracy of the estimates. Also here the estimate error along all axes is much larger with only two overlapping sensor cones compared to three or more. Thus, both graphs confirm a larger fit error and lower confidence in its estimated values for fewer overlapping sensor cones. This is mostly because fewer overlapping sensor readings result in reduced levels sensor noise compensation. As illustrated in Fig. 1, triangulation uses the concept of intersecting radii, and two ground nodes with overlapping sensor cones can provide a deterministic solution for the tracked object, however even small noise will have a detrimental impact on the accuracy of the estimated position of the object. The presence of a third and even a fourth ground node with overlapping sensor cones and a good geometric outline with sufficiently optimized topology and orientations of the ground nodes increases the accuracy significantly.

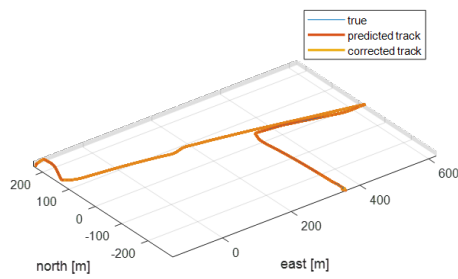


(a) Nr of simultaneous observers

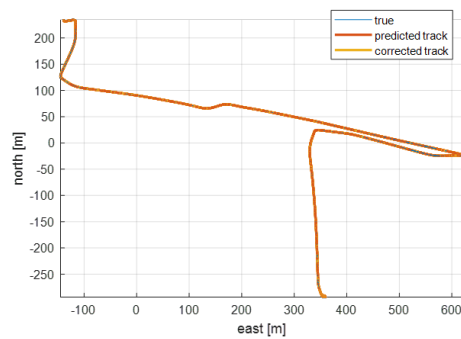


(b) Individual sensor readings

Fig. 23 Nr of simultaneous observers and individual sensor readings for camera observations only

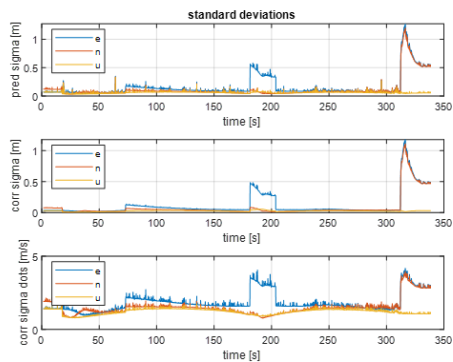


(a) 3D view

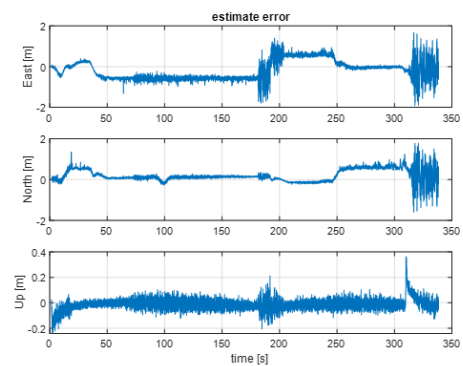


(b) top view

Fig. 24 Comparison between true path and estimate for camera observations only



(a) standard deviations



(b) fit error

Fig. 25 Error information for camera observations only

Fig. 26 shows the total 3 dimensional fit error between estimated and true position for the EO camera observation based estimates. The average fit error is around 0.5 m but with some instantaneous spikes up to slightly more than 2.5 m. This is by far the highest precision of the estimated position of the tracked object for all sensor types considered in this study. Also here, the detrimental impact of fewer simultaneous ground node observations is apparent in the time intervals between 180 - 200 s and after 310s.

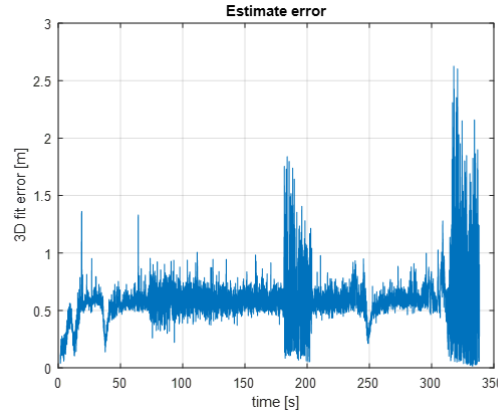


Fig. 26 Total fit error for camera observations only

B. Field test based evaluations

Two different sensor configurations were considered in the field tests, namely RF equipment that allows for RSSI measurements in Sec. V.B.1, and EO cameras that provide two angles of the line of sight in Sec. V.B.2. This study focuses on similar sensor equipment in all ground nodes per field test study. Future work will also consider dissimilar sensor equipment in the various ground nodes and its impact on the accuracy of the estimates.

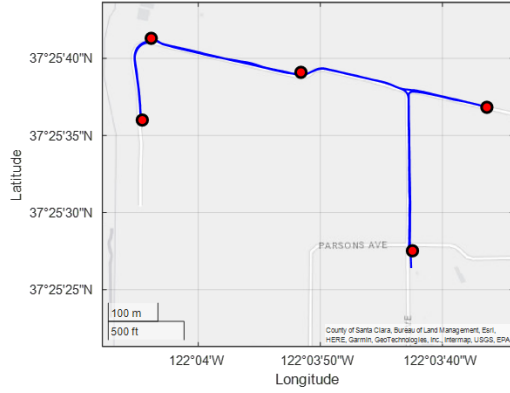
1. RSSI observations only

The first field test was performed with RF equipment only, providing RSSI measurements for a receiver antenna on board of a moving car. In total five RF antennas were used for the field test, one more than in the simulated scenario in Sec. V.A.1 in order to guarantee enough coverage and overlap over the entire terrain for the field test. Some practical details about the setup are provided in Sec. II.B.

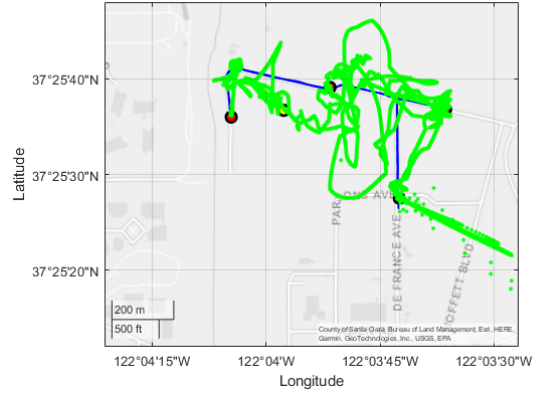
Fig. 27 shows the general outline of the scenario in Fig. 27(a) together with a map view showing the comparison of the estimated track together with the true path in Fig. 27(b). Just like in the simulated scenario in Sec. V.A.1, one can see the lack of convergence due to the inferior signal-to-noise ratio and sampling rate.

Fig. 28(a) shows the individual sensor readings (in red) as measured during the field test and compared with the expected noise-free synthetic signal (in blue), calculated with the GPS measurements during the field test. Comparing both lines shows that the RSSI model as used by the sensor fusion algorithm is tuned well to represent the signal dynamics of each individual ground node. However, local and temporary deviations are visible and can be attributed to obstacles in the terrain such as vegetation or other terrain features that interrupt the line of sight between RF antennas and receiver. The consistent difference for ground node #1 between 800s and about 1,200s was probably caused by a structure and an embankment nearby that reflected the RF signals. Just like observed for the simulated RSSI scenario in Sec. V.A.1, while analyzing this graph combined with Eq. 1, it is clear that this noise level will have a major impact on the signal-to-noise ratio over longer distances because of the logarithmic nature of the signal. At specific time instants, the vehicle is close to one of the antennas but still far from the other three, making trilateration challenging with this level of measurement noise. Fig. 28(b) shows the time varying condition number of the observation matrix. The values are not particularly high, but not as consistently low as can be seen with good coverage by multiple cameras, for example as shown in Fig. 30(a). The rank of the observability matrix is consistently equal to 6, which is the full state rank.

This field test result confirms the same conclusions as the ones which were drawn for the simulated scenario with the same RF measurements in Sec. V.A.1. The signal-to-noise ratio for RSSI signals is too high to make this signal only

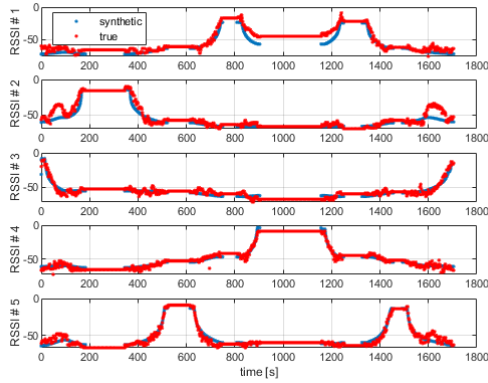


(a) Outline of the scenario

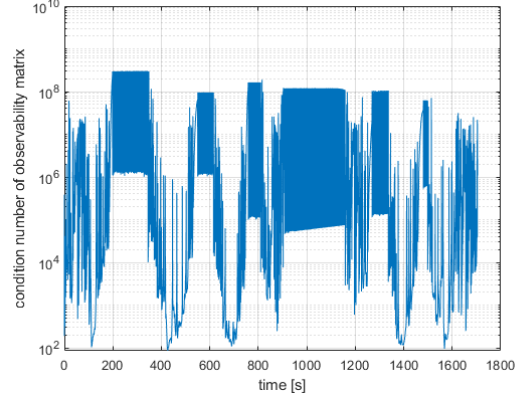


(b) Comparison of estimate with true path

Fig. 27 Outline of scenario and comparison of estimate and true path for RSSI observations only in field test



(a) Individual sensor readings



(b) Condition number of the observability matrix

Fig. 28 Individual sensor readings and condition number for RSSI observations only in field test

useful for accurate localization of a tracked object. This strategy might work, however, in a hybrid set up with multiple sensor readings from various ground nodes with dissimilar sensor equipment, such as cameras and/or radar in addition to RSSI.

2. Camera observations only

The second field test was performed with EO camera equipment only, providing azimuth and elevation angle measurements after raw image processing for a flying drone [20], see also Sec. III, in this case an Alta X. In total four cameras were used for the field test, as many as in the simulated scenario in Sec. V.A.3. Just as in the simulated scenario, these were placed in such a way for optimized topology and viewing directions to guarantee enough coverage and overlap over the entire terrain during the field test. Some practical details about the setup are provided in Sec. II.B.

Fig. 29(a) shows the general outline of the flown scenario and illustrates how the different ground nodes were placed and oriented strategically to achieve optimal coverage of the flight path with the limited number of ground nodes available. In this real life flown environment, in theory 2 signals should be sufficient for localization as long as there is no loss of signal due to exceedance of the field of view or any signal deterioration due to increased observation ranges. However, both were definitely the case. There was one instance where the tracked object temporarily exceeded

the field of view limitations that caused a temporary loss of the tracked object by one of the ground nodes during the flight. Image processing work has also shown that the practical useful range of the ground nodes based cameras is around 700m. The observation cone ranges and minor viewing angle adjustments are the most important changes in Fig. 29(a) compared to Fig. 21(a). Fig. 21(b) compares the estimated path with the true path. In general, the estimate appears to track the true path very accurately on the map, partly thanks to the low sensor noise. There are two additional observations here. The estimated track is color coded green when the condition number of the observation number is below a certain threshold and the state covariance values are below some other predefined thresholds. This indicates a certain confidence in the estimated track. A track segment with a red color means that one of both aforementioned thresholds is exceeded resulting in a low confidence in the result. This color coding correlates well with the accuracy of the estimate. Comparing Fig. 29(a) and Fig. 21(b) shows that the loss in confidence and accuracy corresponds with the moment when only one ground node has the drone in sight. Finally, Fig. 21(b) also shows clearly that convergence of the estimate takes longer, resulting in a cloud of estimate points in the beginning of the flight. Due to the lack of sufficient information, the initial location of the estimate is by default chosen in the origin of the reference frame (yellow marker in both Fig. 29). However, it takes till the tracked object enters at least 2 overlapping sensor cones of the ground nodes before the estimate can start converging.

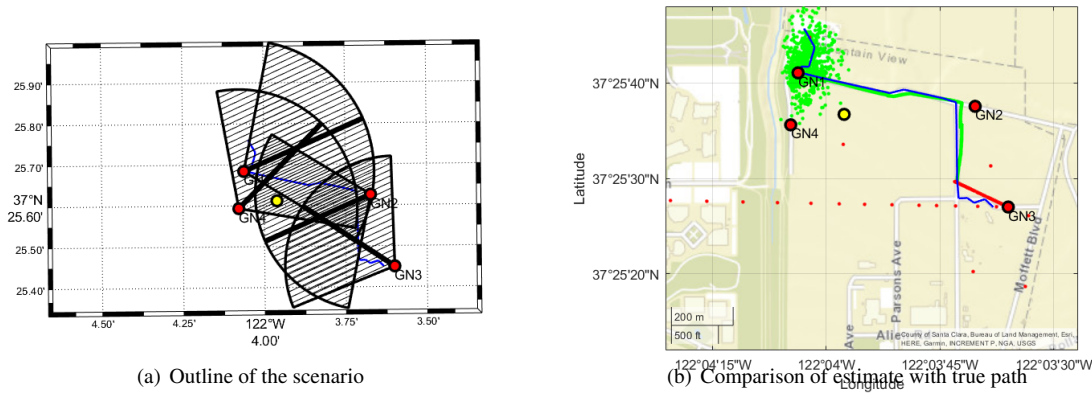


Fig. 29 Outline of scenario and comparison of estimate and true path for camera observations only

For the analysis of the observability matrix, Fig. 30(a) shows its rank and Fig. 30(b) shows its condition number. The matrix has full state rank between approximately 80s and 500s. Its condition number has a similar order of magnitude as for the simulated scenario in Fig. 22(b) in the same aforementioned timeframe. However, the condition number increases multiple orders of magnitude outside this timeframe.

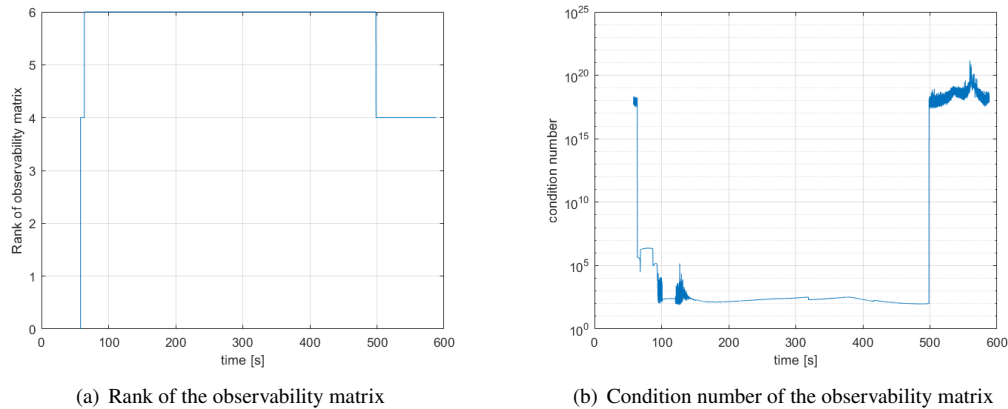


Fig. 30 Rank and condition number for camera observations only

Fig. 31 gives additional information about the sensor readings. Fig. 31(a) shows how many ground nodes have the tracked object in sight at any given time instant of the flight. In other words, this figure shows how many sensor cones, as drawn in Fig. 29(a), overlap with the location of the tracked object at any time instant during the flight. Comparing Fig. 31 and Fig. 30(b) shows that the time instants when the condition number temporarily increases several orders of magnitude occur when there are temporarily only less than two ground nodes tracking the object. Fig. 31(b) shows the measured elevation and azimuth angles of the tracked drone per ground node. These sensor readings show a good signal-to-noise ratio, like in Fig. 23(b) for the simulated scenario. Fig. 31(b) also compares the individual sensor readings (in red) as measured during the field test with the expected noise-free synthetic signal (in green), calculated with the GPS measurements during the flight test. Comparing both lines shows that the camera model as used by the sensor fusion algorithm is tuned well to represent the signal dynamics of each individual ground node. However, local and temporary loss of signal is visible for ground node 1 around 100s and can be attributed to the object leaving the field of view of ground node 1, which results in an interrupted line of sight between EO camera and tracked object. The sensor fusion algorithm restarts converging as soon as a sufficient number of ground nodes is providing measurements.

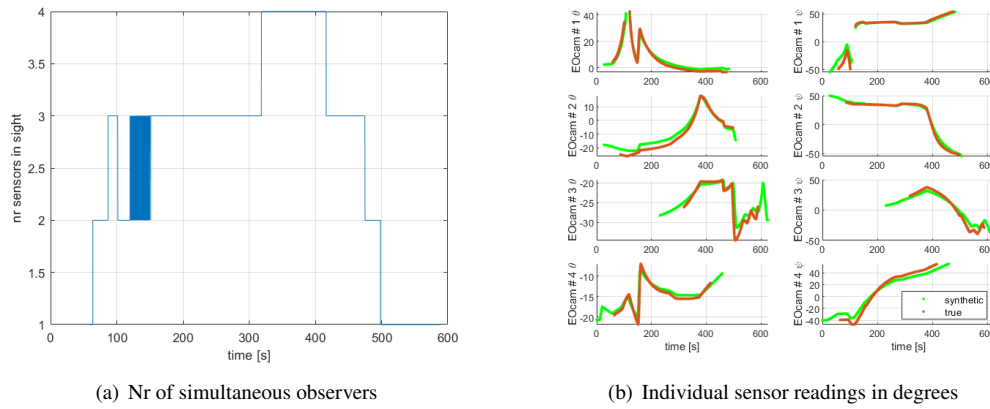
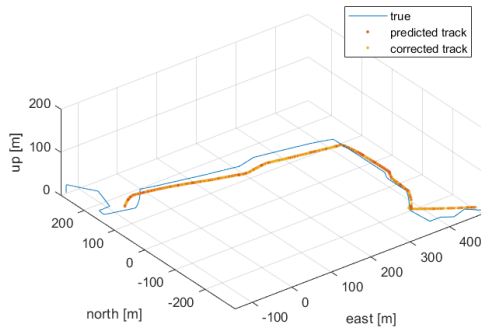


Fig. 31 Nr of simultaneous observers and individual sensor readings for camera observations only

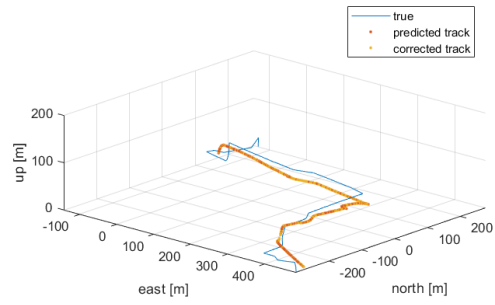
Fig. 32 compares the estimated and true flight path in 3D in the East-North-Up reference frame. Fig. 32(a) shows the 3D view and Fig. 32(b) shows the 3D view from an alternative opposite viewpoint. Both show the predicted as well as the corrected state as calculated in the Kalman Filter. These figures show that the estimate is mostly accurate in the East-North plane, but not as accurate vertically where the estimate drifts off slowly. Returning to the sensor readings in Fig. 31(b), one can see that there are small deviations between the measured sensor readings and their synthetic noise-free expected counterparts. These are caused by lens distortion among others. However, at larger distances, even a very small angle deviation can result in a drift of the estimate over tens of meters.

Fig. 33(a) shows the standard deviations in a similar way as in the previous scenarios. The standard deviations have significantly higher values when there are less than two ground nodes tracking the target compared to two or more. Fig. 33(b) shows the fit error for the 3D position, from top to bottom East - North - Up. These graphs show higher accuracy of the estimates. Also here the estimate error along all axes is much larger with less than two overlapping sensor cones compared to two or more. Thus, both graphs confirm a larger fit error and lower confidence in its estimated values for less than two overlapping sensor cones. This is mostly because the localization problem becomes unsolvable at that point. As illustrated in Fig. 1, triangulation uses the concept of intersecting radii, and only at least two ground nodes with overlapping sensor cones can provide a deterministic solution for the tracked object. The presence of a third and even a fourth ground node with overlapping sensor cones and a good geometric outline with sufficiently optimized topology and orientations of the ground nodes increases the accuracy significantly.

Fig. 34 shows the total 3D fit error between estimated and true position for the EO camera observation based estimates. The average fit error is around 20 m but with some instantaneous spikes up to slightly more than 40 m (after convergence). This estimated position is not as precise as for the simulated scenario considered in this study with EO cameras, which is to be expected with many real life disturbances such as among others lens distortion that is not yet compensated for in this study. Also here, the detrimental impact of fewer simultaneous ground node observations is apparent in the aforementioned time intervals before 80 s and after 500s.

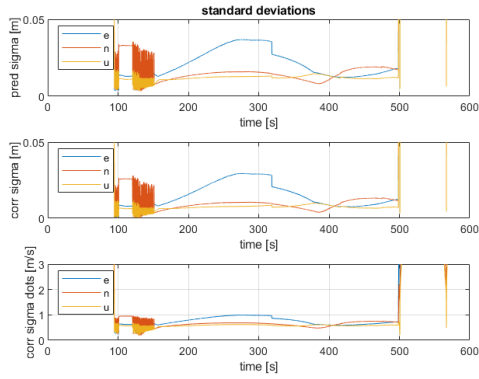


(a) 3D view

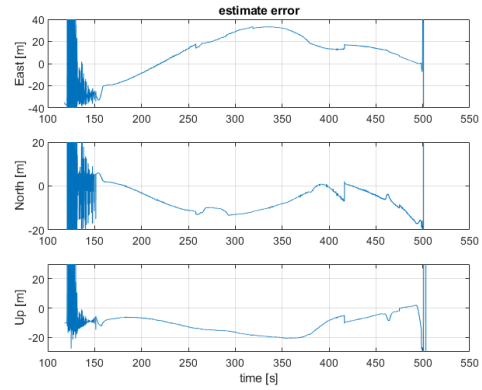


(b) opposite view

Fig. 32 Comparison between true path and estimate for camera observations only



(a) standard deviations



(b) fit error

Fig. 33 Error information for camera observations only

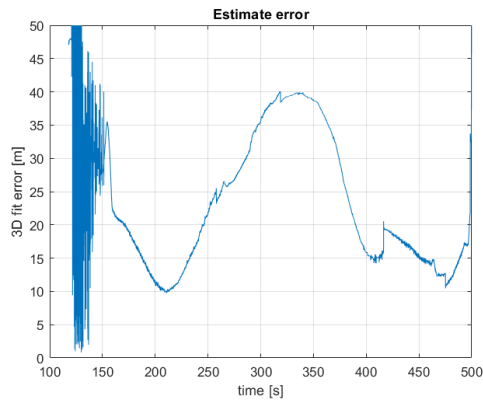


Fig. 34 Total fit error for camera observations only

Both simulated as well as field test scenarios have shown that ground nodes with cameras provide the highest accuracy for the estimated path while localizing a tracked object. This is both because the cameras provide two dimensional sensor readings as well as a better signal-to-noise ratio compared to the alternative sensors such as RF antennas (significantly lower signal-to-noise ratio) and DME equipment (only one dimensional sensor readings for range). However, the accuracy of the estimates based on camera readings depends to a large extent on the precision of the ground node location as well as viewing direction information. Also the quality of the image processing plays an important role here, including compensation for among others lens distortion.

VI. Conclusions and Recommendations

This study developed a Kalman Filter based sensor fusion algorithm for distributed ground based sensing that is flexible with respect to the number of ground nodes and the sensor types in each ground node. The different sensor types that are considered in the observation model are EO cameras, radar, RSSI signals from RF antennas and some form of DME equipment. This study focused on the use of RSSI, DME and EO cameras, which were evaluated in simulated scenarios as well as field tests. Based on both evaluations, the following observations were made.

RF signals have a bad signal-to-noise ratio for application cases with longer ranges, which are typical for Urban Air Mobility operations. Therefore, they might be better suited for indoor short range applications, but that is not valid in the Urban Air Mobility context. A higher accuracy with RF signals could be achieved by an improved topology with a higher density of antennas, but an antenna for every 100m or so would be operationally challenging. RSSI signal model calibration in the sensor fusion algorithm for the individual modem antennas in the ground nodes is time consuming.

Other direct distance measurements without log scaling over range, such as DME, have shown in simulations to outperform modem antennas. However, when all ground nodes are in one single plane and the vehicle is operating in the vicinity of that plane, then the estimates are inaccurate along the third dimension perpendicular on that plane. Possible strategies to avoid this are either adding some ground nodes or other observers outside the aforementioned single two dimensional plane, or adding an elevation angle measurement perpendicular on that plane.

Cameras perform vastly superior compared to the two previous sensor types, in simulations as well as in field tests, but they come with their own challenges. Image processing and camera calibration are tedious and demanding tasks.

In future research, attention will be paid to the compensation of lens distortion during image processing, and hybrid scenarios will be considered with ground nodes using non-identical sensor equipment. Lastly, an operationally relevant challenge is finding the minimum number of ground nodes and which sensors types are required to achieve a minimum threshold for accuracy of the estimated path.

Appendix: Reference Frame Transformations for Kalman Filter

Throughout this paper, five different reference frames are used: geodetic (WGS-84), Earth-Centered Earth-Fixed (ECEF), East-North-Up (ENU), body fixed and sensor spherical reference frames. They are all illustrated in Fig. 35.

The geodetic reference frame expresses the spherical coordinates as longitude λ_s with respect to the Greenwich meridian, latitude ϕ_s with respect to the equator and altitude h_s above the average sea-level. The ECEF reference frame is Cartesian with the origin in the Earth center and its coordinates are as follows: the Z_{ECEF} axis extends through true North, X_{ECEF} points through the earth surface where the equator and Greenwich meridian intersect, and Y_{ECEF} is finally perpendicular on the two previous axes. The ENU Cartesian reference frame is sensor centered with the axes pointing East-North-Up. As illustrated in Fig. 35, the ‘up’ direction is location dependent, as well as the others. As a consequence, the ENU frames for different sensors and target are all differently oriented. This difference increases for increasing distances between all of them. This must be reflected in the calculations. Next is the body fixed reference frame, which is Cartesian and centered at the sensor’s center of gravity. The X_B axis points forward, Y_B axis points to the right and the Z_B axis is downwards. Finally, the sensor centered spherical reference frame expresses the coordinates as range ρ (distance from center of gravity), elevation α (vertical angle with respect to the $X_B - Y_B$ -plane) and azimuth γ (horizontal angle with respect to the forward axis). The sensor measurements from radar and camera are expressed in this spherical reference frame. Note that both latter sensor fixed reference frames are sensor attitude dependent, i.e. the viewing direction.

The transformations between the different aforementioned reference frames are illustrated in Fig. 36. The transformations between body-fixed, ENU, ECEF as well as between different ENU’s of different sensors and target at different locations are pure angular rotations by means of the respective rotation matrices $\mathbf{M}(\phi, \theta, \psi)$ over the three Euler angles between body fixed and ENU, $\mathbf{T}(\lambda_s, \phi_s)$ over longitude and latitude between ENU and ECEF and

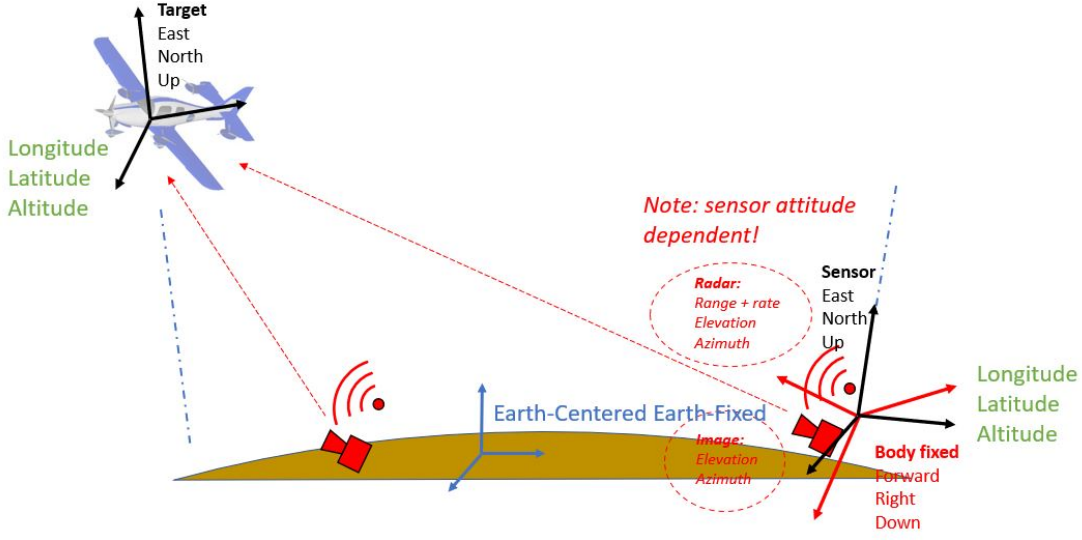


Fig. 35 Reference frames used by Kalman Filter

$\mathbf{L}(\lambda_s, \phi_s, \lambda_t, \phi_t)$ over longitude and latitude of both locations between different ENU's at different locations. The mappings between body fixed and sensor spherical as well as between geodetic and ECEF are more complex nonlinear transformations. Each of them is elaborated next.

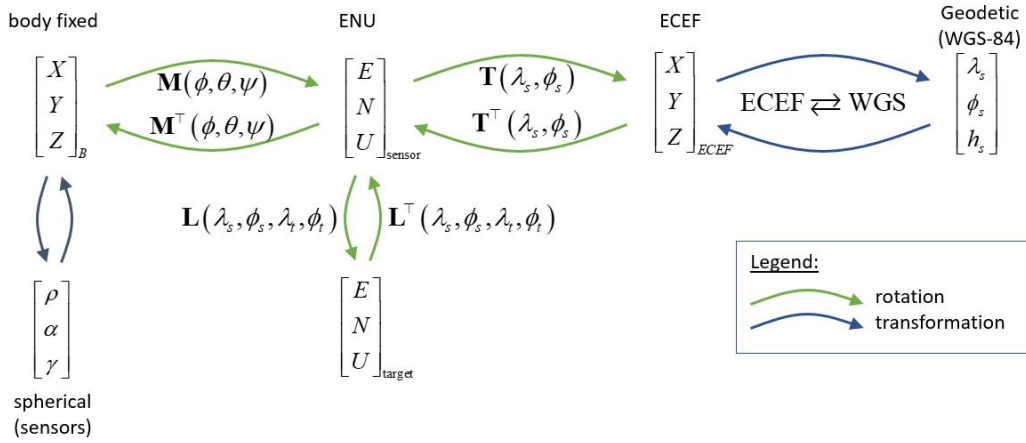


Fig. 36 Transformations between the different reference frames

Transformation from body-fixed to spherical and backwards

The sensor spherical coordinates are defined as a function of the body-fixed coordinates as follows:

$$\rho = \sqrt{X_B^2 + Y_B^2 + Z_B^2} \quad (41)$$

$$\alpha = \arctan \frac{-Z_B}{\rho_{\text{hor}}} \quad (42)$$

$$\gamma = \arctan \frac{Y_B}{X_B} \quad (43)$$

$$(44)$$

where:

$$\rho_{\text{hor}} = \sqrt{X_B^2 + Y_B^2} \quad (45)$$

And the other way around:

$$X_B = \rho \cos \alpha \cos \gamma \quad (46)$$

$$Y_B = \rho \cos \alpha \sin \gamma \quad (47)$$

$$Z_B = \rho \sin \alpha \quad (48)$$

Rotation from body-fixed to ENU and backwards

From body-fixed to ENU is a sequential rotation over the three Euler angles roll ϕ , pitch attitude θ and yaw ψ [22]:

$$\mathbf{M}(\phi, \theta, \psi) = \begin{bmatrix} \sin \psi \cos \theta & \cos \phi \cos \psi + \sin \phi \sin \psi \sin \theta & -\sin \phi \cos \psi + \cos \phi \sin \psi \sin \theta \\ \cos \psi \cos \theta & -\cos \phi \sin \psi + \sin \phi \cos \psi \sin \theta & \sin \phi \sin \psi + \cos \phi \cos \psi \sin \theta \\ \sin \theta & -\sin \phi \cos \theta & -\cos \phi \cos \theta \end{bmatrix} \quad (49)$$

The transpose is used for the backwards direction.

Rotation from ENU to ECEF and backwards

From the vehicle centered ENU to the earth centered ECEF is a sequential rotation over the longitude λ_s and latitude ϕ_s of the current position[22]:

$$\mathbf{T}(\lambda_s, \phi_s) = \begin{bmatrix} -\sin \lambda_s & -\cos \lambda_s \sin \phi_s & \cos \lambda_s \cos \phi_s \\ \cos \lambda_s & -\sin \lambda_s \sin \phi_s & \sin \lambda_s \cos \phi_s \\ 0 & \cos \phi_s & \sin \phi_s \end{bmatrix} \quad (50)$$

The transpose is used for the other way around.

Rotation between ENU's

A rotation between ENU's is basically a rotation from the first ENU to ECEF and subsequently from ECEF back to the second ENU:

$$\mathbf{L}(\lambda_s, \phi_s, \lambda_t, \phi_t) = \mathbf{T}^T(\lambda_t, \phi_t) \cdot \mathbf{T}(\lambda_s, \phi_s) \quad (51)$$

Substituting Eq. (50) results in:

$$\begin{aligned} \mathbf{L}(\lambda_s, \phi_s, \lambda_t, \phi_t) &= \\ &= \mathbf{T}^T(\lambda_t, \phi_t) \cdot \mathbf{T}(\lambda_s, \phi_s) \\ &= \begin{bmatrix} -\sin \lambda_t & -\cos \lambda_t \sin \phi_t & \cos \lambda_t \cos \phi_t \\ \cos \lambda_t & -\sin \lambda_t \sin \phi_t & \sin \lambda_t \cos \phi_t \\ 0 & \cos \phi_t & \sin \phi_t \end{bmatrix}^T \begin{bmatrix} -\sin \lambda_s & -\cos \lambda_s \sin \phi_s & \cos \lambda_s \cos \phi_s \\ \cos \lambda_s & -\sin \lambda_s \sin \phi_s & \sin \lambda_s \cos \phi_s \\ 0 & \cos \phi_s & \sin \phi_s \end{bmatrix} \\ &= \begin{bmatrix} \cos(\lambda_t - \lambda_s) & \sin(\lambda_t - \lambda_s) \sin \phi_s & -\sin(\lambda_t - \lambda_s) \cos \phi_s \\ -\sin(\lambda_t - \lambda_s) \sin \phi_t & \cos \phi_t \cos \phi_s + \sin \phi_t \sin \phi_s \cos(\lambda_t - \lambda_s) & \cos \phi_t \sin \phi_s - \sin \phi_t \cos \phi_s \cos(\lambda_t - \lambda_s) \\ \sin(\lambda_t - \lambda_s) \cos \phi_t & \sin \phi_t \cos \phi_s - \cos \phi_t \sin \phi_s \cos(\lambda_t - \lambda_s) & \sin \phi_t \sin \phi_s + \cos \phi_t \cos \phi_s \cos(\lambda_t - \lambda_s) \end{bmatrix} \end{aligned} \quad (52)$$

Transformation from ECEF to geodetic and backwards

ECEF $\Rightarrow$ geodetic

The target position in WGS-84 geodetic coordinates is obtained following the 15-step process presented below (extracted from RTCA DO-317A)[21]:

$$r_{\text{ecef}} = \sqrt{X_t^2 + Y_t^2} \quad (53)$$

$$E^2 = a^2 - b^2 \quad (54)$$

$$F = 54b^2Z_t^2 \quad (55)$$

$$G = r_{\text{ecef}}^2 + (1 - \varepsilon^2)Z_t^2 - \varepsilon^2E^2 \quad (56)$$

$$C = \frac{\varepsilon^4 F r_{\text{ecef}}^2}{G^3} \quad (57)$$

$$S = \sqrt[3]{1 + C + \sqrt{C^2 + 2C}} \quad (58)$$

$$P = \frac{F}{3 \left(S + \frac{1}{S} + 1 \right)^2 G^2} \quad (59)$$

$$Q = \sqrt{1 + 2\varepsilon^4 P} \quad (60)$$

$$r_o = \frac{-P\varepsilon^2 r_{\text{ecef}}}{1 + Q} + \sqrt{\frac{a^2}{2} \left(1 + \frac{1}{Q} \right) - \frac{P(1 - \varepsilon^2)Z_t^2}{Q(1 + Q)} - \frac{Pr_{\text{ecef}}^2}{2}} \quad (61)$$

$$U = \sqrt{(r_{\text{ecef}} - \varepsilon^2 r_o)^2 + Z_t^2} \quad (62)$$

$$V = \sqrt{(r_{\text{ecef}} - \varepsilon^2 r_o)^2 + (1 - \varepsilon^2)Z_t^2} \quad (63)$$

$$Z_o = \frac{b^2 Z_t}{aV} \quad (64)$$

$$\phi_t = \tan^{-1} \left(\frac{Z_t + \varepsilon'^2 Z_o}{r_{\text{ecef}}} \right) \quad (65)$$

$$\lambda_t = \tan^{-1} \left(\frac{Y_t}{X_t} \right) \quad (66)$$

$$h_t = U \left(1 - \frac{b^2}{aV} \right) \quad (67)$$

where:

$a = 6.378137 \cdot 10^6 m$, semimajor axis of the Earth

$b = 6.3567523142 \cdot 10^6 m$, semiminor axis of the Earth

$\varepsilon = 0.081819190842622$, first eccentricity of the Earth

$\varepsilon' = 0.082094437935831$, second eccentricity of the Earth

ϕ_t = target latitude [rad]

λ_t = target longitude [rad]

h_t = target geodetic altitude [m]

geodetic $\Rightarrow$ ECEF

The ownship's Earth-Centered, Earth-Fixed (ECEF) position is calculated by[21]:

$$\begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = \begin{bmatrix} (N_s + h_s) \cos \phi_s \cos \lambda_s \\ (N_s + h_s) \cos \phi_s \sin \lambda_s \\ (N_s (1 - \varepsilon^2) + h_s) \sin \phi_s \end{bmatrix} \quad (68)$$

where:

ϕ_s = ownship geodetic latitude [rad]

λ_s = ownship geodetic longitude [rad]

h_s = ownship geodetic altitude [m]

$$N_s = \frac{a}{\sqrt{1 - \varepsilon^2 \sin^2 \phi_s}}$$

$a = 6.378137 \cdot 10^6 m$, semimajor axis of the Earth

$\varepsilon = 0.081819190842622$, first eccentricity of the Earth

If the altitude of the ownship is measured as height above Mean Sea Level (MSL) from a barometric sensor, it should be converted to WGS-84 Height Above Ellipsoid (HAE) using the following steps:

$$\xi(\phi_s) = g_{0eq} \frac{1 + k \sin^2 \phi_s}{\sqrt{1 - \varepsilon^2 \sin^2 \phi_s}} \quad (69)$$

$$R(\phi_s) = \frac{a}{1 + f + m_r - 2f \sin^2 \phi_s} \quad (70)$$

$$h_S = \frac{R(\phi_s)^h}{\frac{R(\phi_s)\xi(\phi_s)}{\xi(\frac{\pi}{4})} - h} \quad (71)$$

where:

ϕ_s = ownship geodetic latitude [rad]

$\varepsilon = 0.081819190842622$, first eccentricity of the Earth

$g_{0eq} = 9.7803253359 m/s^2$, theoretical gravity at the equator

$g_{0pl} = 9.8321849378 m/s^2$, theoretical gravity at the poles

$$k = \frac{b \cdot g_{0pl}}{a \cdot g_{0eq}} - 1$$

$\xi(\phi_s)$ = normal gravity on the surface of an ellipsoid at ϕ_s [m/s^2]

$a = 6.378137 \cdot 10^6 m$, semimajor axis of the Earth

$b = 6.3567523142 \cdot 10^6 m$, semiminor axis of the Earth

$f = \frac{a-b}{a}$, flattening of the Earth

$m_r = 0.003449787$, gravity ratio

$R(\phi_s)$ = radius of the Earth at ϕ_s [m]

h = measured ownship height above MSL

h_s = ownship geodetic altitude [m]

Acknowledgments

This work was performed under the NASA Aeronautics Research Mission Directorate (ARMD), Transformative Tools and Technologies (TTT) Project. The authors would like to thank Zach Roberts, Jonas Jonsson, Matt Gregory, Edgar Torres and the other colleagues of Smart Mobility at NASA Ames Research Center for the successful drone flights, as well as NASA Armstrong colleague Nelson Brown, for their invaluable contributions to this project.

References

- [1] Aubuchon, V. V., Hashemi, K. E., Shively, R. J., and Wishart, J. M., "Multi-Vehicle (m:N) Operations in the NAS - NASA's Research Plans," *AIAA AVIATION 2022 Forum*, American Institute of Aeronautics and Astronautics, 2022. doi:10.2514/6.2022-3758.
- [2] Lombaerts, T., Shish, K., Keller, G., Stepanyan, V., Cramer, N., and Ippolito, C., "Adaptive Multi-Sensor Fusion Based Object Tracking for Autonomous Urban Air Mobility Operations," *AIAA SciTech Forum*, 2022. doi:10.2514/6.2022-0362.
- [3] Lombaerts, T., Kannan, K., Kawamura, E., Dolph, C., Stepanyan, V., Gorospe, G. E., and Ippolito, C. A., "Distributed Ground Sensor Fusion Based Object Tracking for Autonomous Advanced Air Mobility Operations," *AIAA SCITECH 2023 Forum*, American Institute of Aeronautics and Astronautics, 2023. doi:10.2514/6.2023-0896.

- [4] Dolph, C., Lombaerts, T., Kawamura, E., Ippolito, C. A., Stepanyan, V., Iftekharuddin, K., Szatkowski, G., McSwain, R., Morris, C., Malekpour, M. R., et al., "Ground to air testing of a fused optical-radar aircraft detection and tracking system," *AIAA SCITECH 2022 Forum*, 2022, p. 0498. doi:10.2514/6.2022-0498.
- [5] Vitiello, F., Causa, F., Opromolla, R., Fasano, G., Dolph, C., Lombaerts, T., and Ippolito, C., "Assessing performance of radar visual fusion techniques for ground to air surveillance in advanced air mobility," *IEEE/DASC 2023*, 2023. URL <https://ntrs.nasa.gov/citations/20230010197>.
- [6] Stepanyan, V., Kannan, K., Kawamura, E., Lombaerts, T., and Ippolito, C. A., "Distributed Target Tracking With Optimal Data Migration," *AIAA SCITECH 2023 Forum*, 2023. doi:10.2514/6.2023-2194.
- [7] Chong, C. Y., Chang, K. C., and Mori, S., "Distributed Tracking in Distributed Sensor Networks." *Proceedings of the American Control Conference*, 1986, pp. 1863–1868. doi:10.23919/acc.1986.4789229.
- [8] Qi, H., Iyengar, S. S., and Chakrabarty, K., "Distributed sensor networks - A review of recent research," *Journal of the Franklin Institute*, Vol. 338, 2001, pp. 655–668. doi:10.1016/S0016-0032(01)00026-6.
- [9] Horling, B., Vincent, R., Mailler, R., Shen, J., Becker, R., Rawlins, K., and Lesser, V., "Distributed sensor network for real time tracking," *Proceedings of the International Conference on Autonomous Agents*, 2001, pp. 417–424. doi:10.1145/375735.376399.
- [10] Chamberland, J.-F., and Veeravalli, V., "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, Vol. 51, 2003, pp. 407–416. doi:10.1109/TSP.2002.806982.
- [11] La, H. M., and Sheng, W., "Distributed sensor fusion for scalar field mapping using mobile sensor networks," *IEEE Transactions on Cybernetics*, Vol. 43, 2013, pp. 766–778. doi:10.1109/TSMCB.2012.2215919.
- [12] Chatterjee, A., Flores, H., Sen, S., Hasan, K. S., and Mani, A., "Distributed location detection algorithms using IoT for commercial aviation," *Proceedings - 2017 3rd IEEE International Conference on Research in Computational Intelligence and Communication Networks, ICRCICN 2017*, Vol. 2017-Decem, 2017, pp. 126–131. doi:10.1109/ICRCICN.2017.8234493.
- [13] Cheng, L., Wu, C., Zhang, Y., Wu, H., Li, M., and Maple, C., "A Survey of Localization in Wireless Sensor Network," *International Journal of Distributed Sensor Networks*, Vol. 8, No. 12, 2012, p. 962523. doi:10.1155/2012/962523, URL <https://doi.org/10.1155/2012/962523>.
- [14] Amundson, I., and Koutsoukos, X. D., "A Survey on Localization for Mobile Wireless Sensor Networks," *Mobile Entity Localization and Tracking in GPS-less Environments*, Springer Berlin Heidelberg, 2009, pp. 235–254. doi:10.1007/978-3-642-04385-7_16.
- [15] Mesmoudi, A., Feham, M., and Labraoui, N., "Wireless Sensor Networks Localization Algorithms: A Comprehensive Survey," *International journal of Computer Networks & Communications*, Vol. 5, No. 6, 2013, pp. 45–64. doi:10.5121/ijcnc.2013.5603.
- [16] Alhmiedat, T. A., and Yang, S.-H., "A survey: localization and tracking mobile targets through wireless sensors network," *Eighth Annual Network Symposium, PGNet*, 2007. URL https://repository.lboro.ac.uk/articles/journal_contribution/A_survey_localization_and_tracking_mobile_targets_through_wireless_sensors_network/9405143.
- [17] Stepanyan, V., Ippolito, C., Kannan, K., Kawamura, E., Lombaerts, T., Holforty, W., and Dolph, C., "Flight Test Design and Implementation for Airspace Independent Surveillance through a Distributed Ground Based Sensor Network," *AIAA SciTech 2024 Forum*, 2024.
- [18] Shish, K. H., Cramer, N. B., Gorospe, G., Lombaerts, T., Stepanyan, V., and Kannan, K., "Survey of Capabilities and Gaps in External Perception Sensors for Autonomous Urban Air Mobility Applications," *AIAA Scitech 2021 Forum*, American Institute of Aeronautics and Astronautics, 2021. doi:10.2514/6.2021-1114.
- [19] Kannan, K., Ippolito, C., and Holforty, W., "Flight Test Configuration of the Sensor Payload and the Ground Nodes in Distributed Sensing Frameworks," *AIAA SciTech 2024 Forum*, 2024.
- [20] Kawamura, E., Kannan, K., Lombaerts, T., Stepanyan, V., Dolph, C., and Ippolito, C. A., "Ground-Based Vision Tracker for Advanced Air Mobility and Urban Air Mobility," *AIAA SciTech 2024 Forum*, 2024.
- [21] *Minimum Operational Performance Standards for Detect and Avoid Systems*, DO-365, RTCA Inc., 2019, techreport Appendix F: Example Functional Description For Airborne Surveillance Data Processor (ASDP) Subsystem. URL <https://standards.globalspec.com/std/14281907/rtca-do-365>.
- [22] Mohinder, S., Grewal, P., Lawrence, R., Weill, A. P., and Andrews, "Appendix C: Coordinate Transformations," *Global Positioning Systems, Inertial Navigation, and Integration*, John Wiley & Sons, Inc., 2006, pp. 456–501. doi:10.1002/9780470099728.app3.