VSLAM and Vision-based Approach and Landing for Advanced Air Mobility

Evan Kawamura [1]
Chester Dolph [2]
Keerthana Kannan [1]
Nelson Brown [3]
Thomas Lombaerts [1]
Corey Ippolito [1]

[1] *NASA Ames Research Center (ARC)*
[2] *NASA Langley Research Center (LARC)*
[3] *NASA Armstrong Flight Research Center (AFRC)*

Autonomous Systems – Perception & Distributed Sensing
2023 AIAA Scitech Forum
January 2023

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

Outline

1. Introduction
2. AFRC Flight Tests and Landmarks
3. Feature Detection for VAL
4. VSLAM Method Selection
5. VAL Design
6. Results
7. Conclusion

Introduction

- GPS degradation occurs in urban environments → need Alternative Position, Navigation, and Timing (APNT) solutions
- Localize based on known landmarks, guidelines, or geometrical patterns at runways, heliports, and vertiports
- Conducted UAS flight tests at the NASA Armstrong Flight Research Center (AFRC) helipad to simulate AAM aircraft approach and landing
- Utilized video and telemetry data to test vision-based precision approach and landing (PAL) methods
- Compare two vision-based APNT solutions against UAS GPS logs as ground truth
  - ORB SLAM 2
  - Vison-based Approach and Landing (VAL), see Ref. [1] for more details
- Cones and landmarks distributed around the AFRC helipad serves as fiducials for VAL (known feature points)

[1] Kawamura, E., Dolph, C., Kannan, K., Lombaerts, T., and Ippolito, C. A., "Simulated Vision-based Approach and Landing System for Advanced Air Mobility," *AIAA SciTech 2023 Forum*, 2023

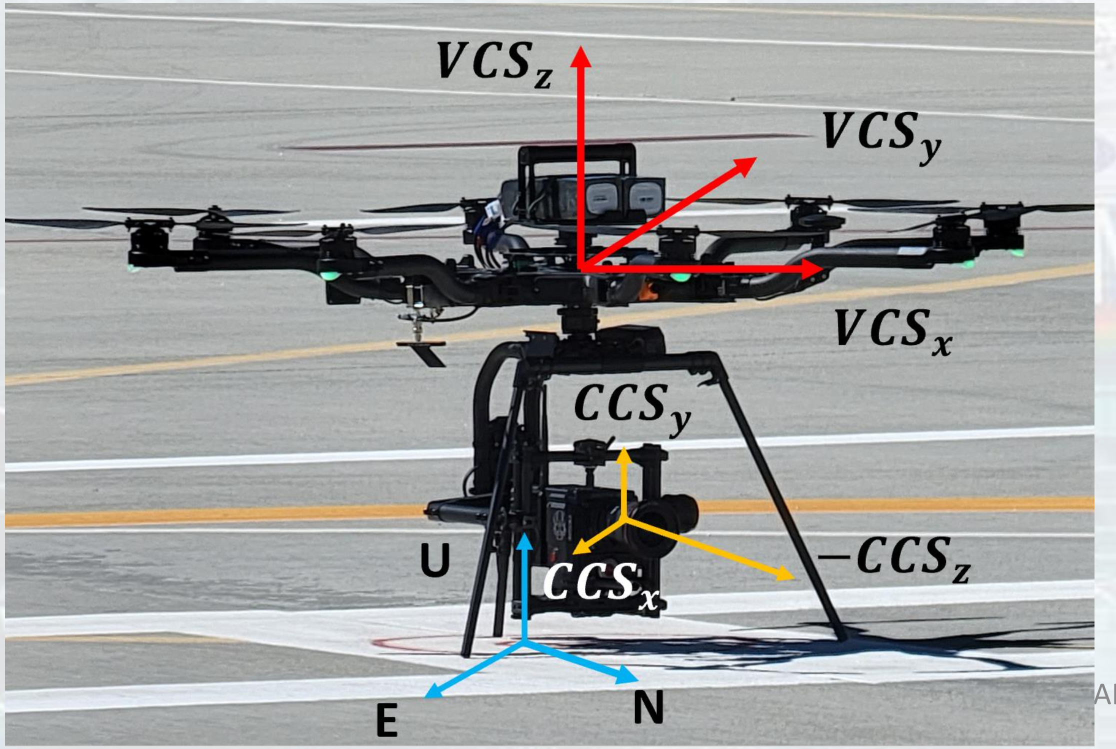
Outline

1. Introduction
2. AFRC Flight Tests and Landmarks
3. Feature Detection for VAL
4. VSLAM Method Selection
5. VAL Design
6. Results
7. Conclusion

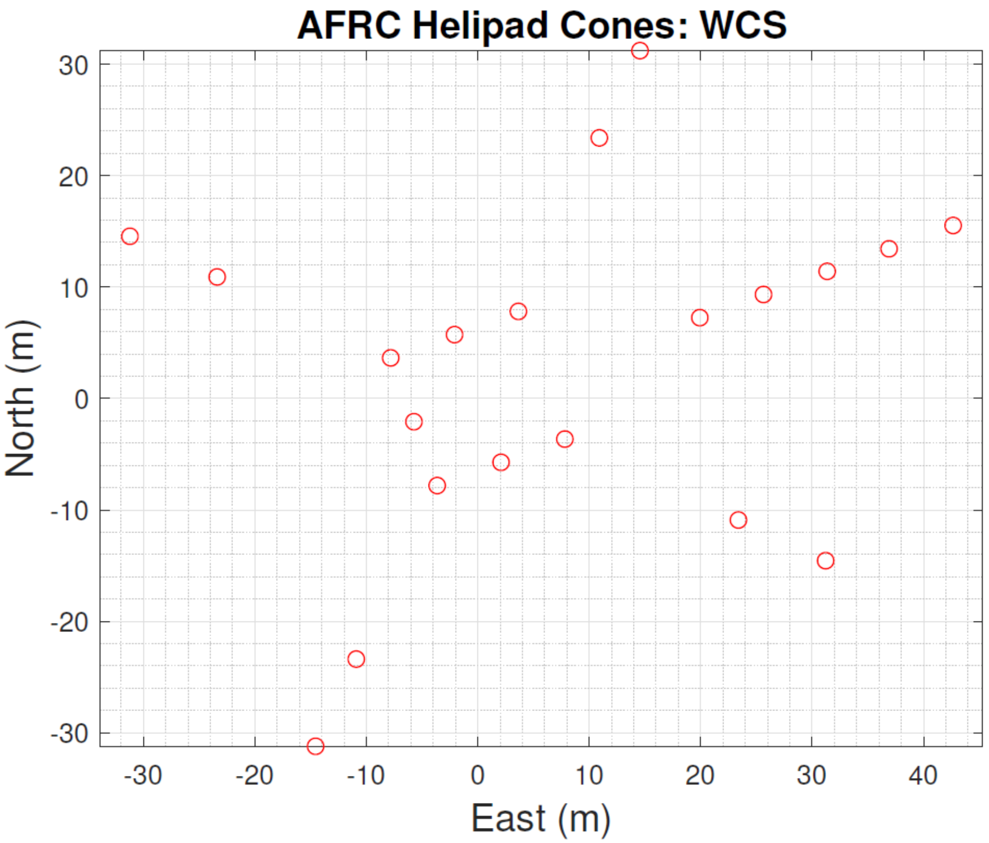
AFRC Flight Tests and Landmarks

- Frames: inertial world coordinates (ENU), vehicle coordinates (VCS), and camera coordinates (CCS)
- Alta8 UAS state vector: $\boldsymbol{s} = [E\ N\ U\ v_E\ v_N\ v_U\ \phi\ \theta\ \psi]^T$
- National Geospatial-Intelligence Agency (NGA) provided the WGS84 coordinates of the helipad markings (horizontal accuracy = 0.02 m, vertical accuracy = 0.1 m) [2]
- Cone locations coincide with the concrete intersection points

[2] Thompson, N., “NASA National Campaign Build 1, Edwards AFB, California,” National Geospatial-Intelligence Agency, 2020.



AFRC Flight Tests and Landmarks



- TLOF = **T**ouchdown and **L**ift**OF**f area
- FATO = **F**inal **A**pproach and **T**ake **O**ff
- SA = **S**afety **A**rea
- LIC = **L**ead **I**n **C**one

AFRC Flight Tests and Landmarks

- The Alta8 starts over the lakebed, facing towards the helipad
- Resembles a glideslope approach and landing profile
- The Alta8 does not have a glideslope controller
- The glidepath is not at the suggested angle of 9° per Ref. [3]

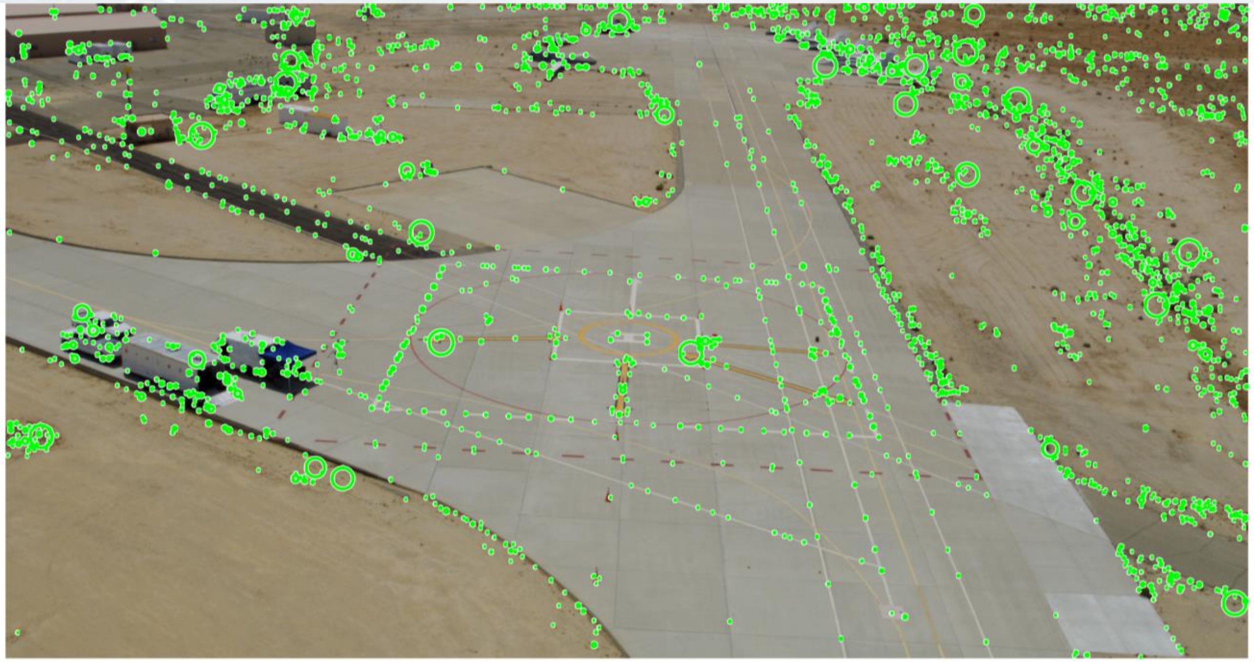
[3] Webber, D., and Zahn, D., “FAA and the National Campaign,” [Powerpoint], 2021.

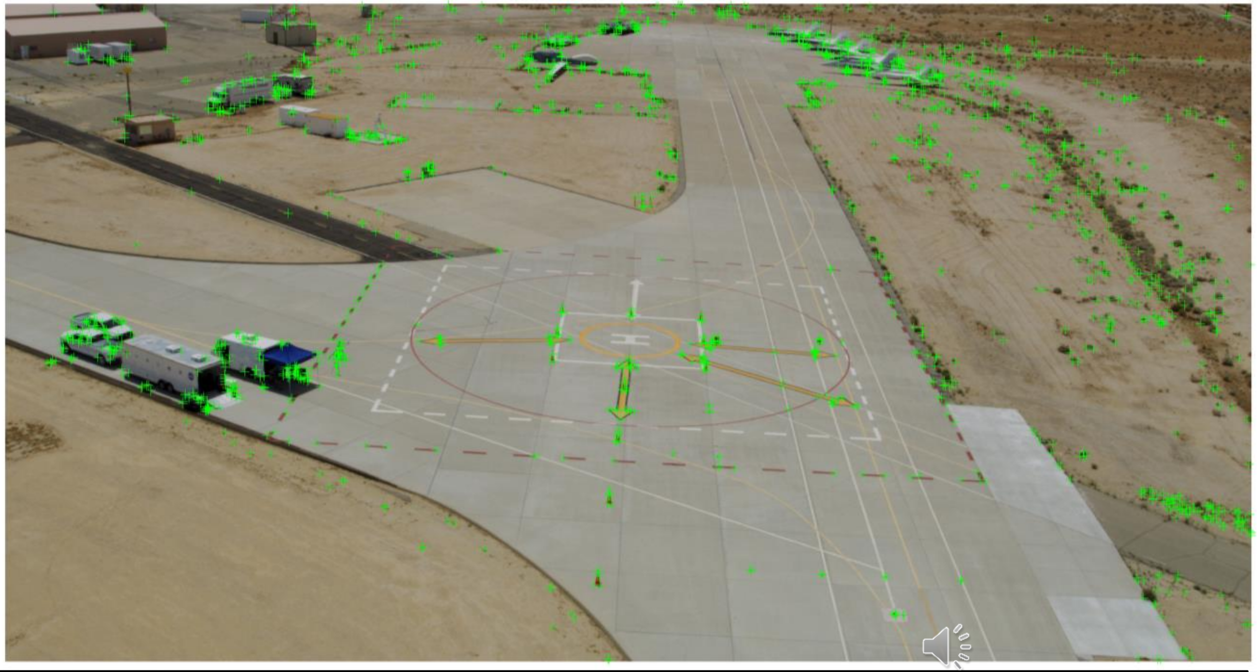
Outline

1. Introduction
2. AFRC Flight Tests and Landmarks
3. Feature Detection for VAL
4. VSLAM Method Selection
5. VAL Design
6. Results
7. Conclusion

Preliminary Results: Feature Descriptor

- VAL in Ref. [1,4] used Hough circle detection
- Hough circle detection does not find all the cones and has many irrelevant detections
- Harris corner detection finds all the key features with less irrelevant detections
- Since Harris corner detection finds the landmarks with less irrelevant detections, VAL uses Harris corner detection instead of Hough circle detection in this study.





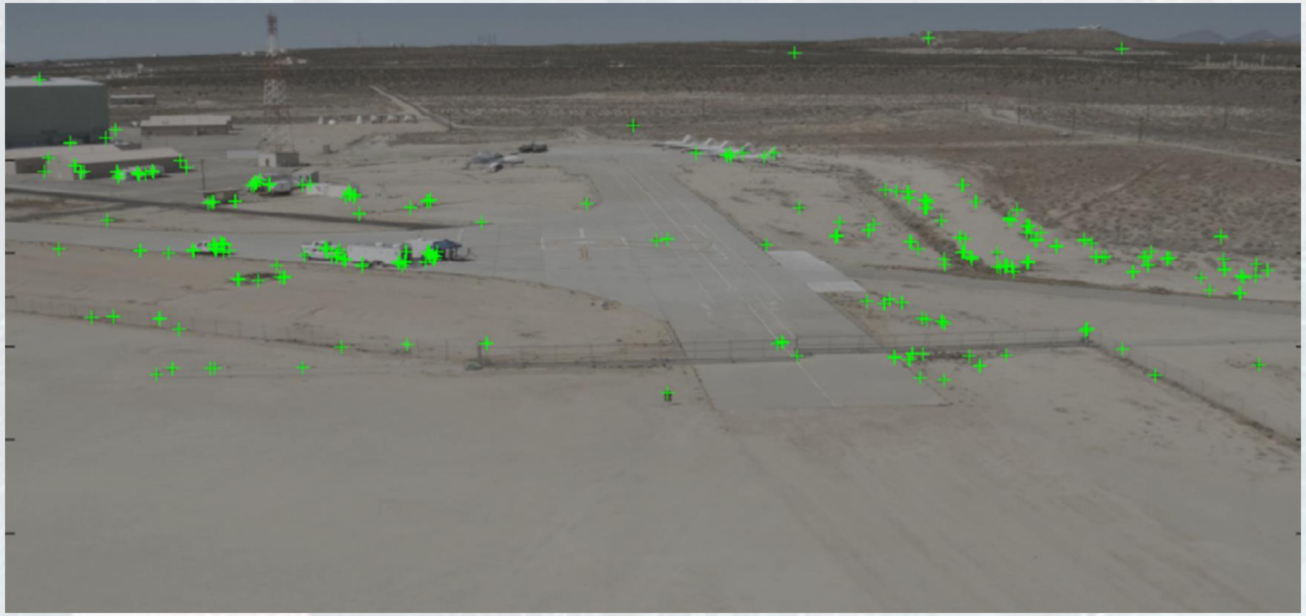
[1] Kawamura, E., Dolph, C., Kannan, K., Lombaerts, T., and Ippolito, C. A., "Simulated Vision-based Approach and Landing System for Advanced Air Mobility," *AIAA SciTech 2023 Forum*, 2023
[4] Kawamura, E., Kannan, K., Lombaerts, T., and Ippolito, C. A., "Vision-Based Precision Approach and Landing for Advanced Air Mobility," AIAA SciTech 2022 Forum, AIAA 2022-0497, 2022. https://doi.org/10.2514/6.2022-0497

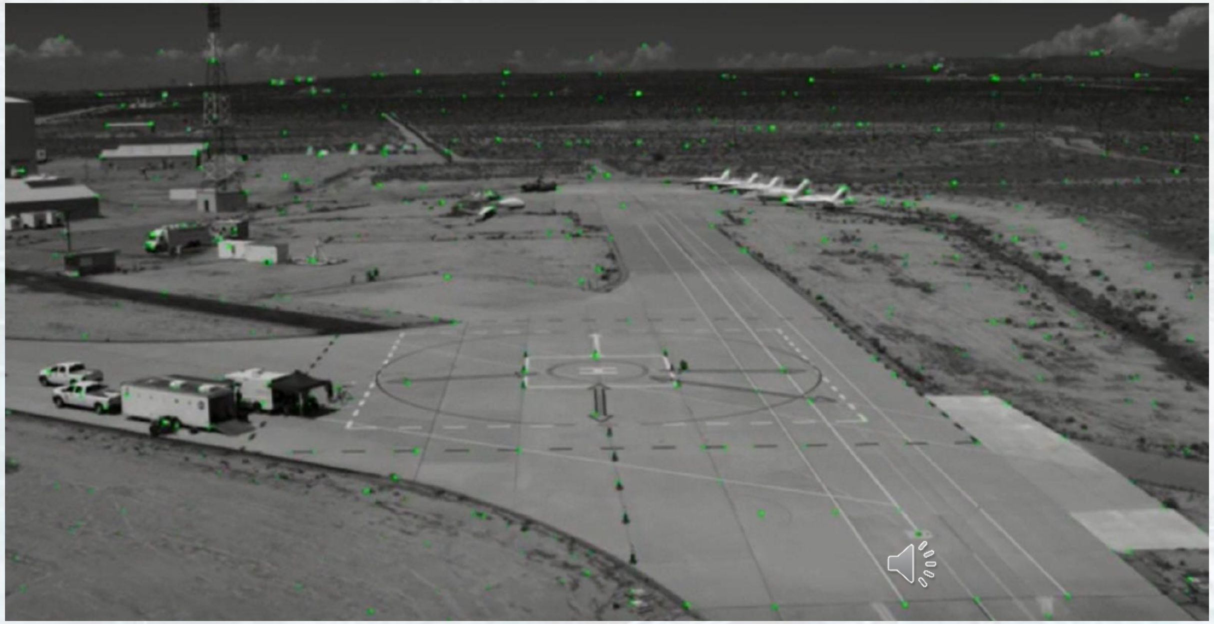
Outline

1. Introduction
2. AFRC Flight Tests and Landmarks
3. Feature Detection for VAL
4. VSLAM Method Selection
5. VAL Design
6. Results
7. Conclusion

VSLAM Method Selection

- Preliminary results compared ORB SLAM and ORB SLAM 2
- ORB SLAM picked up many non-helipad features due to its “small” bag of features
- ORB SLAM 2 found several helipad features such as markings, cones, corners, & edges because it is more feature based than ORB SLAM -> **better performance**





Outline

VAL Design

Differences from Ref. [1]:

1. Timestep in this paper is 2.38 seconds to match ORB SLAM 2’s timestep for a more accurate comparison
2. Different Q & R matrix components (EKF tuned differently)
3. Harris corner detection instead of Hough circle detection

Main similarities from Ref. [1]:

1. Same EKF structure with two options
   a. IMU only **(acceleration & body angular rates)**: measurement matrix is all zeros, prediction step, no correction step
   b. IMU & **CO**planar **P**ose from **O**rthography and **S**caling with **IT**erations (COPOSIT) measurements **(position, velocity, and Euler angles)**: measurement matrix is identity (9x9) and has prediction and correction steps
2. Same camera model and feature correspondence methods

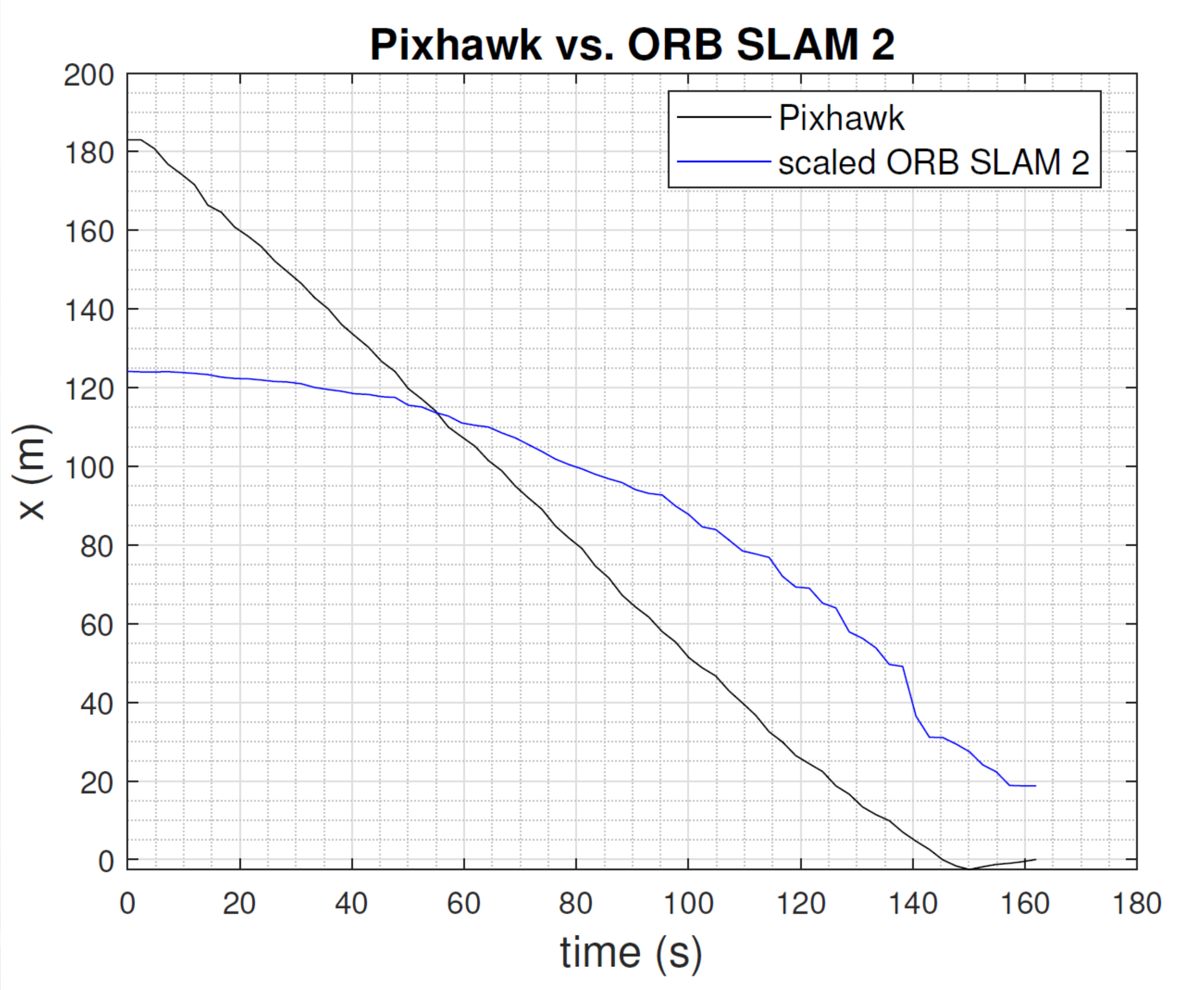
[1] Kawamura, E., Dolph, C., Kannan, K., Lombaerts, T., and Ippolito, C. A., “Simulated Vision-based Approach and Landing System for Advanced Air Mobility,” *AIAA SciTech 2023 Forum*, 2023

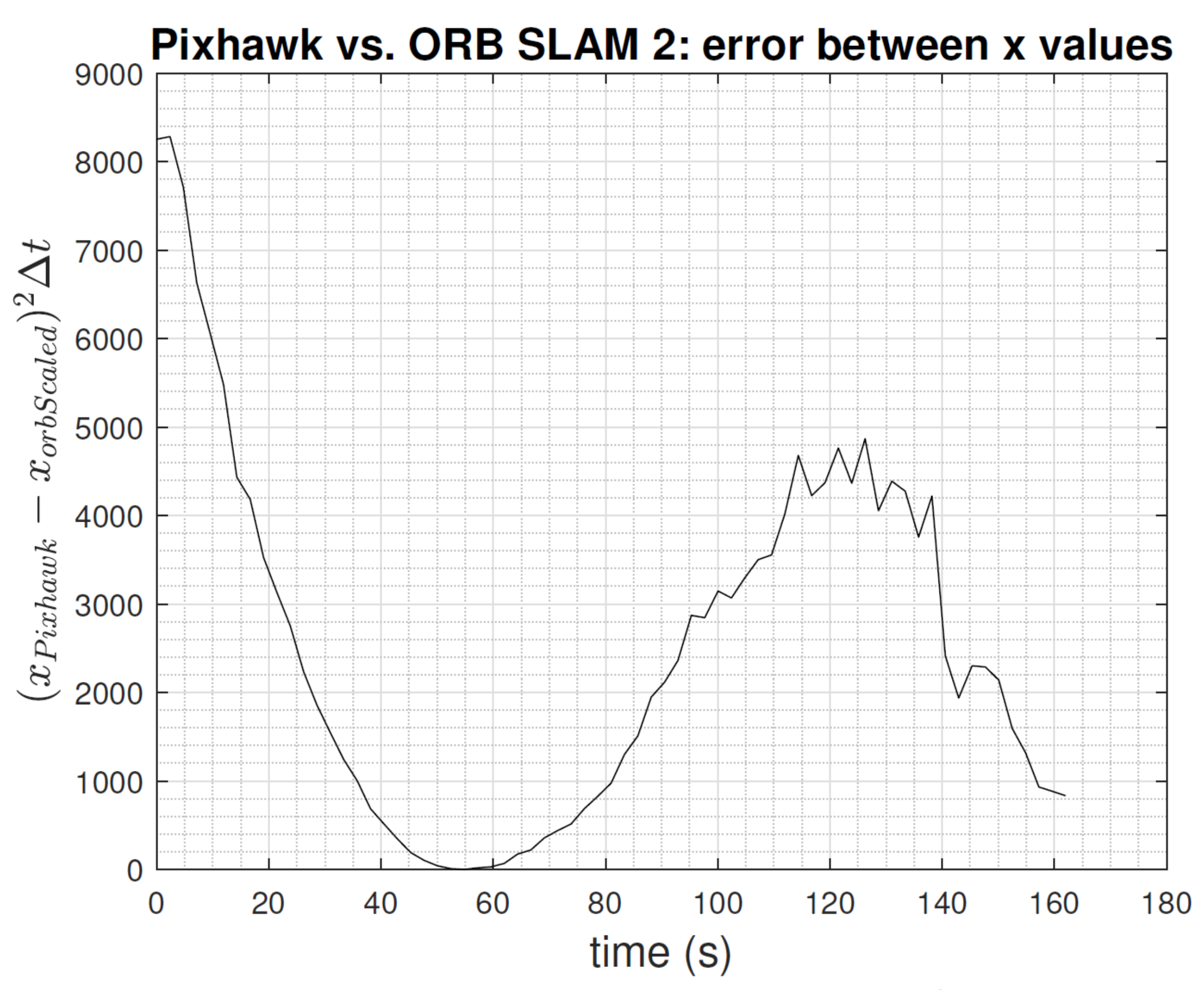
Outline

Results: ORB SLAM 2

1. Sync Pixhawk data with the unscaled trajectory from ORB SLAM 2 (see paper for details)
2. Scale ORB SLAM 2 trajectory with a linear fit: $\boldsymbol{p}_{scaled} = A\boldsymbol{p}_{unscaled} + b$
   a. A & b are the slope and y-intercept values
   b. Apply LP Simplex in Excel's solver to determine A & b while minimizing the sum of the squared error
3. Squared difference between Pixhawk and scaled ORB SLAM 2 position with time step, $\Delta t = t_{i+1} - t_i$:
   a. $\Delta x^2 \Delta t = \left(x_{scaled,wcs} - x_{pixhawk,wcs}\right)^2 \Delta t$
   b. $\Delta y^2 \Delta t = \left(y_{scaled,wcs} - y_{pixhawk,wcs}\right)^2 \Delta t$
   c. $\Delta z^2 \Delta t = \left(z_{scaled,wcs} - z_{pixhawk,wcs}\right)^2 \Delta t$
4. Sum of squared difference multiplied by the time step with N points:
   a. $\Gamma = \sum_{i=1}^{N} \left(\boldsymbol{p}_{scaled,i} - \boldsymbol{p}_{pixhawk,i}\right)^2 \Delta t$
   b. $\epsilon_x = \sum_{i=1}^{N} \left(x_{scaled,i} - x_{pixhawk,i}\right)^2 \Delta t$
   c. $\epsilon_y = \sum_{i=1}^{N} \left(y_{scaled,i} - y_{pixhawk,i}\right)^2 \Delta t$
   d. $\epsilon_z = \sum_{i=1}^{N} \left(z_{scaled,i} - z_{pixhawk,i}\right)^2 \Delta t$

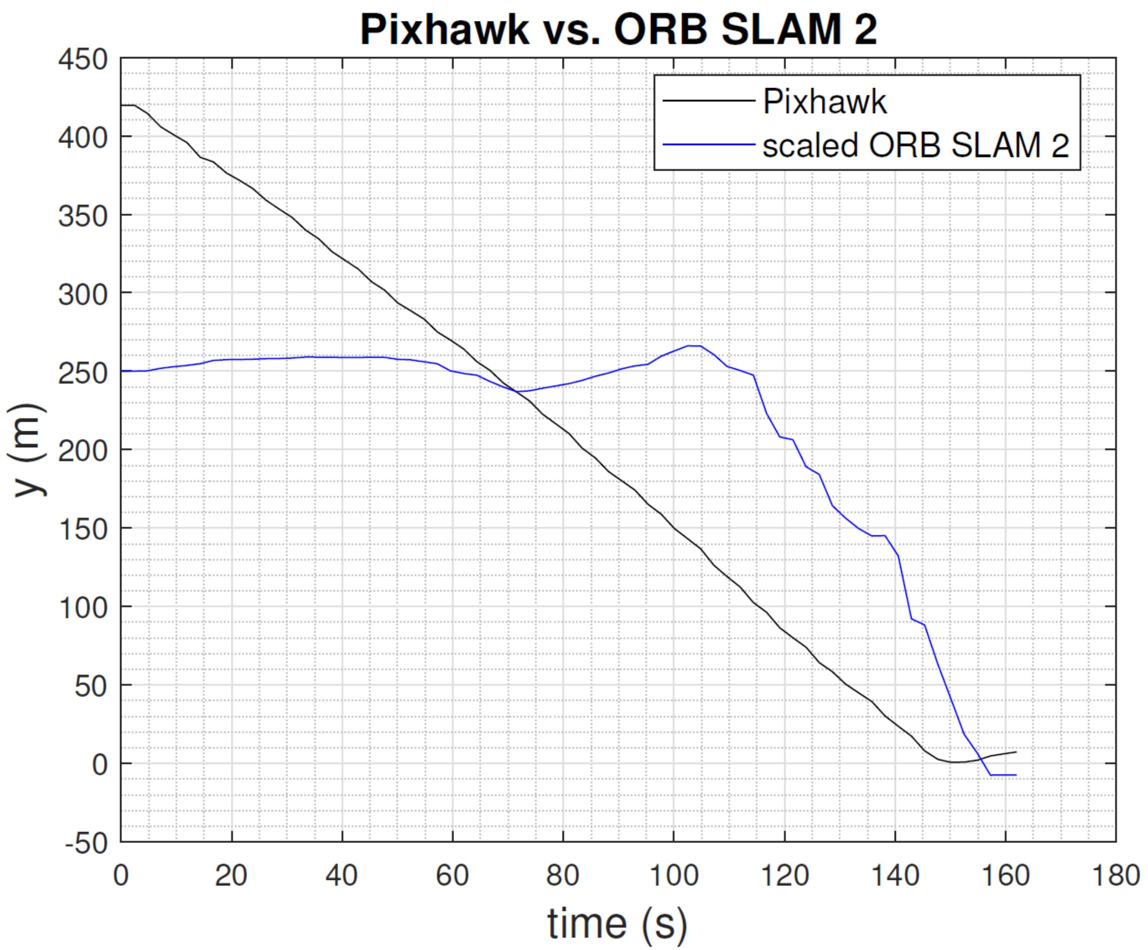
Results: ORB SLAM 2

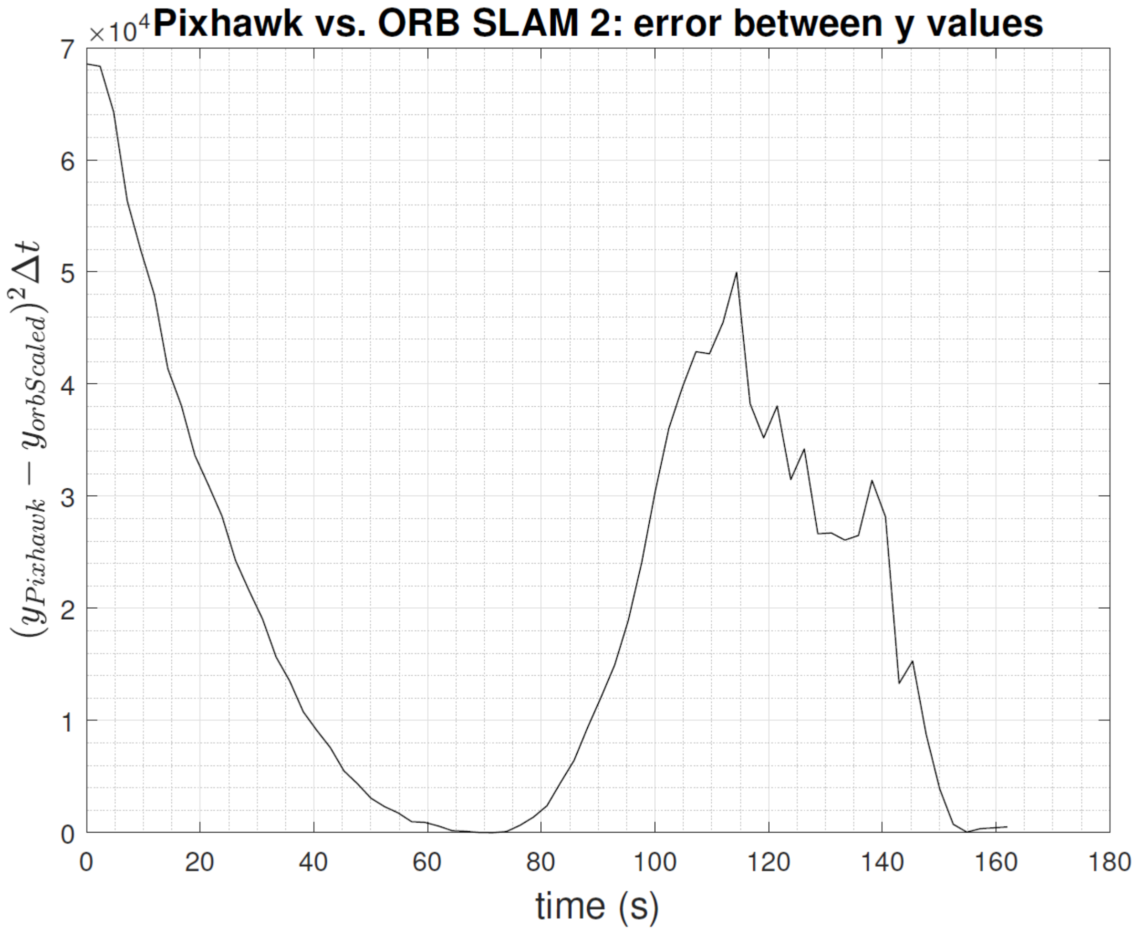




$$\epsilon_x = \sum_{i=1}^{N} \left(x_{scaled,i} - x_{pixhawk,i}\right)^2 \Delta t \rightarrow \epsilon_x \approx 173{,}827$$

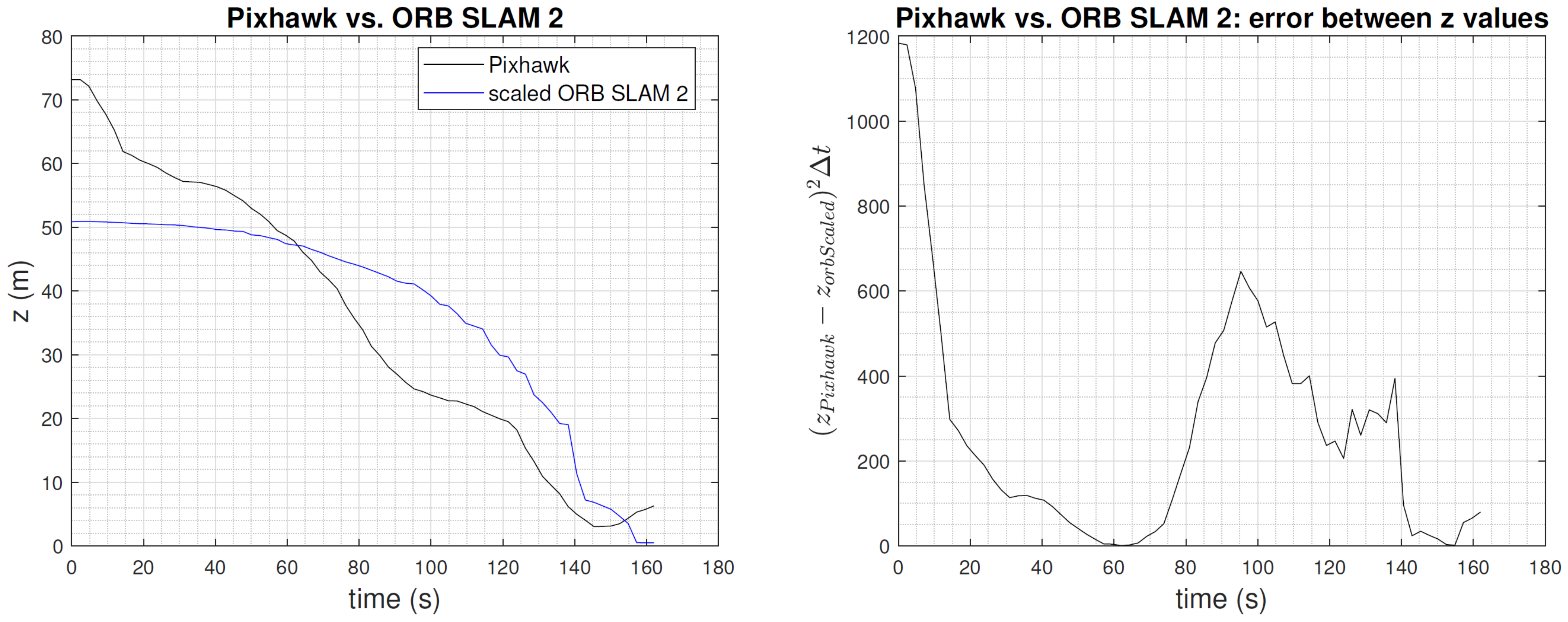
Results: ORB SLAM 2





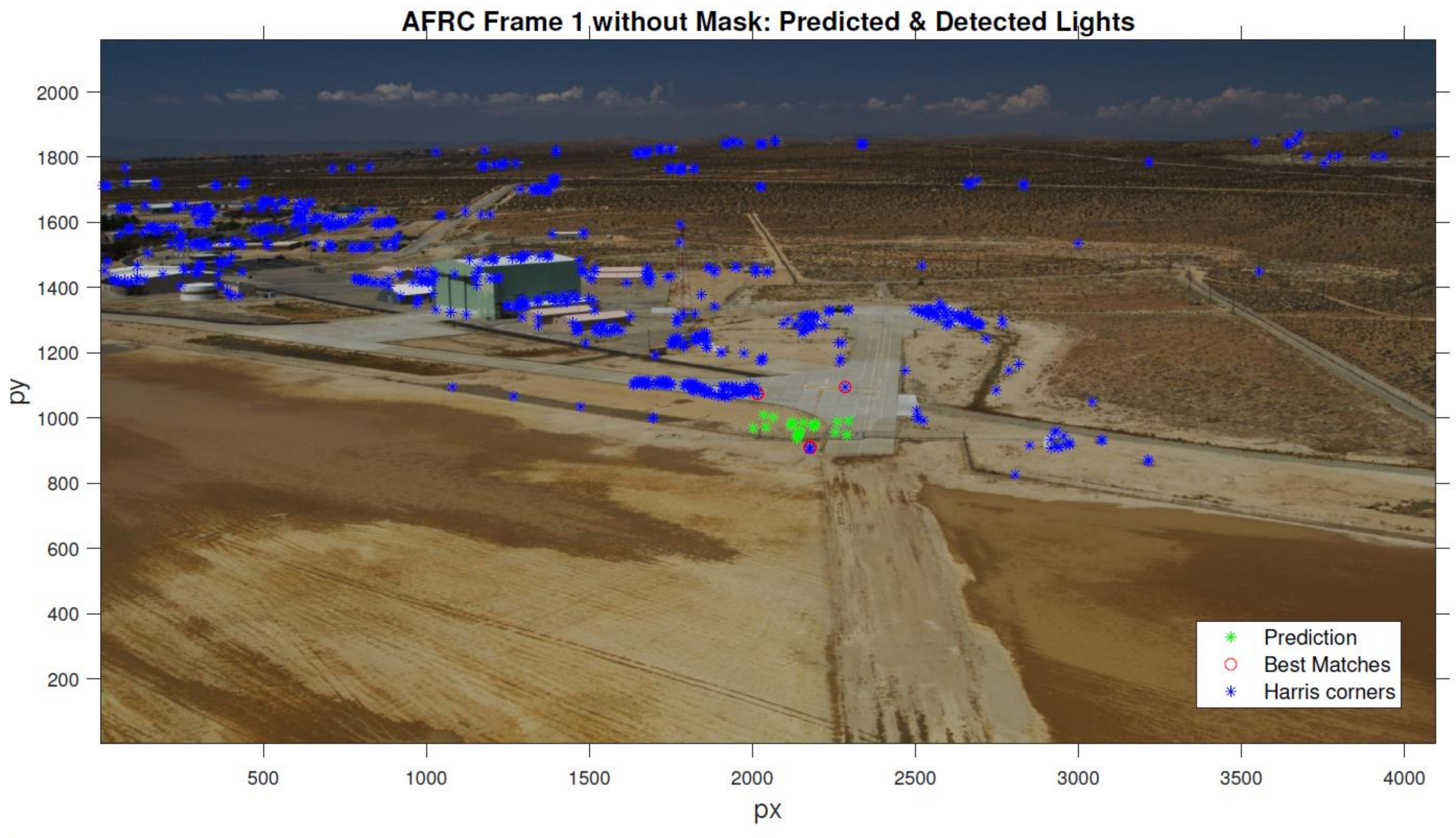
$$\epsilon_y = \sum_{i=1}^{N} \left(y_{scaled,i} - y_{pixhawk,i}\right)^2 \Delta t \rightarrow \epsilon_y \approx 1{,}438{,}996$$

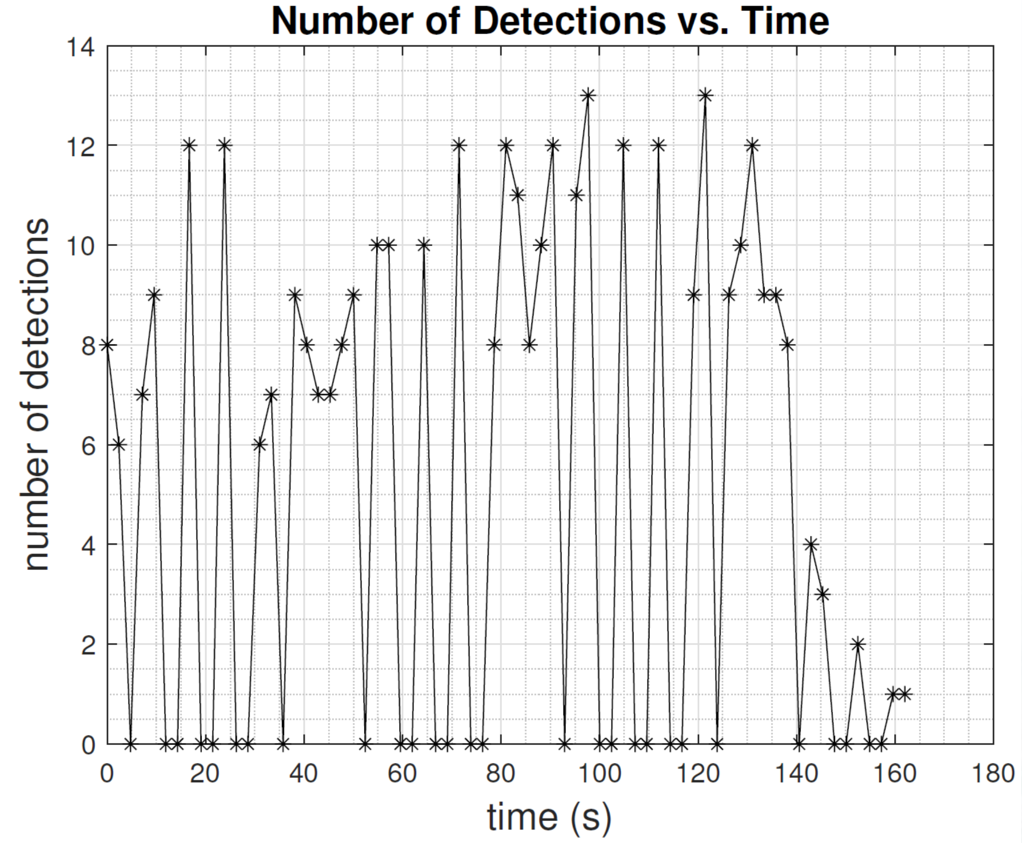
Results: ORB SLAM 2



$$\epsilon_z = \sum_{i=1}^{N} \left(z_{scaled,i} - z_{pixhawk,i}\right)^2 \Delta t \rightarrow \epsilon_z \approx 18{,}456\ (smallest)$$

Results: VAL (COPOSIT-EKF)





For more details on feature detection and correspondence, see the paper for preliminary results on feature detection and correspondence (Figs. 18-20 in the paper)

- Number of detections fluctuate
- Need at least four detections (coplanar points) to obtain COPOSIT measurements and accurate state estimation
- Lose features towards the end – out of view

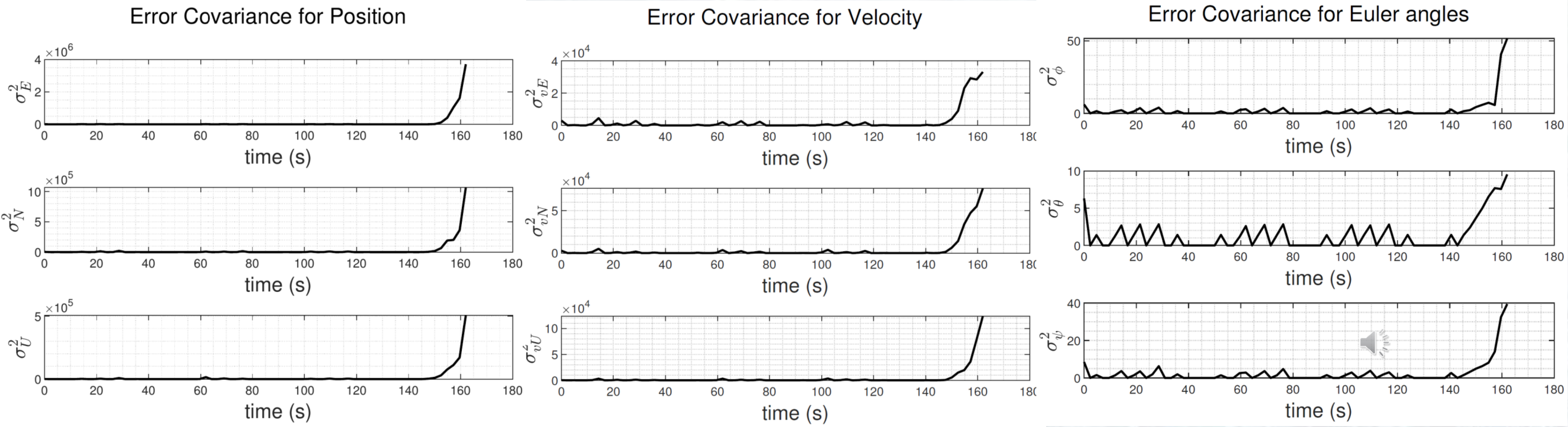
Results: VAL (COPOSIT-EKF)

Process	Mean (s)	Median (s)	Min (s)	Max (s)	Std (s)
COPOSIT	0.0515	0.0492	0.0463	0.0835	0.00708
Feature Detection & Correspondence	0.796	0.794	0.730	0.894	0.0399
EKF	0.00499	0.00105	0.000348	0.0741	0.0135

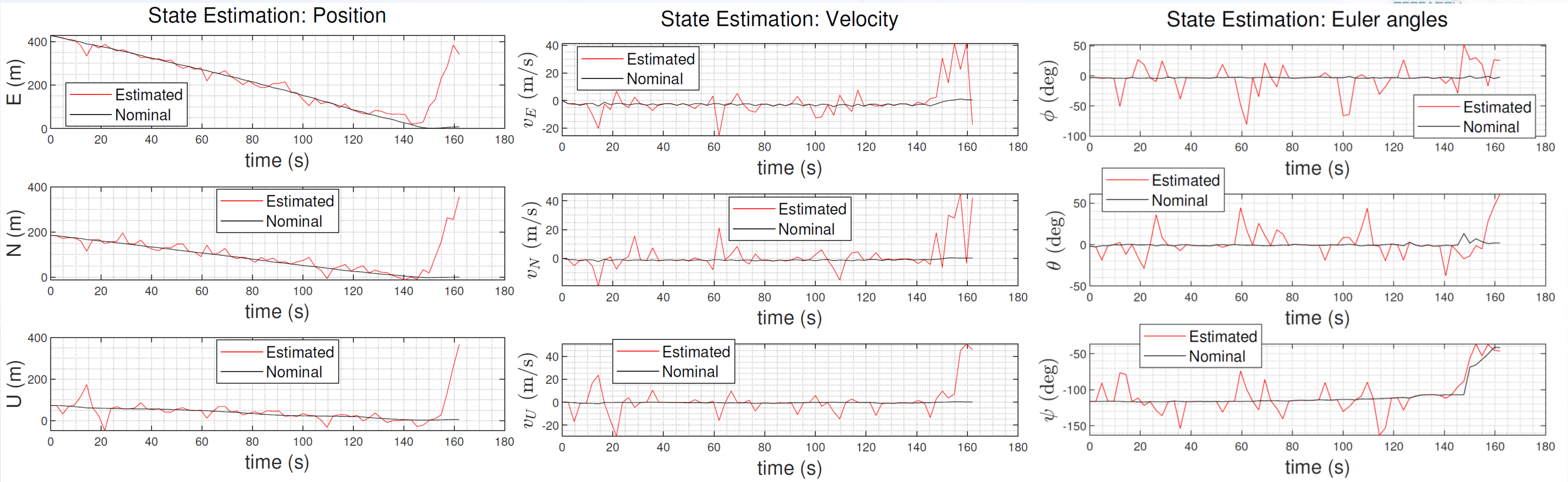
- VAL runs in near real-time
  - EKF runs the fastest (milliseconds)
  - COPOSIT takes centiseconds
  - Feature detection (Harris corner) & correspondence takes about 1 second (slowest)
- Real-time implementation needs faster feature detection and correspondence runtime

Results: VAL (COPOSIT-EKF)

- Error covariances diverge
  - Landmarks out of field of view
  - Lack of COPOSIT measurements
- High levels of uncertainty and low confidence at the end (landing)
- Future work: add a nadir camera to see landmarks during landing



Results: VAL (COPOSIT-EKF)



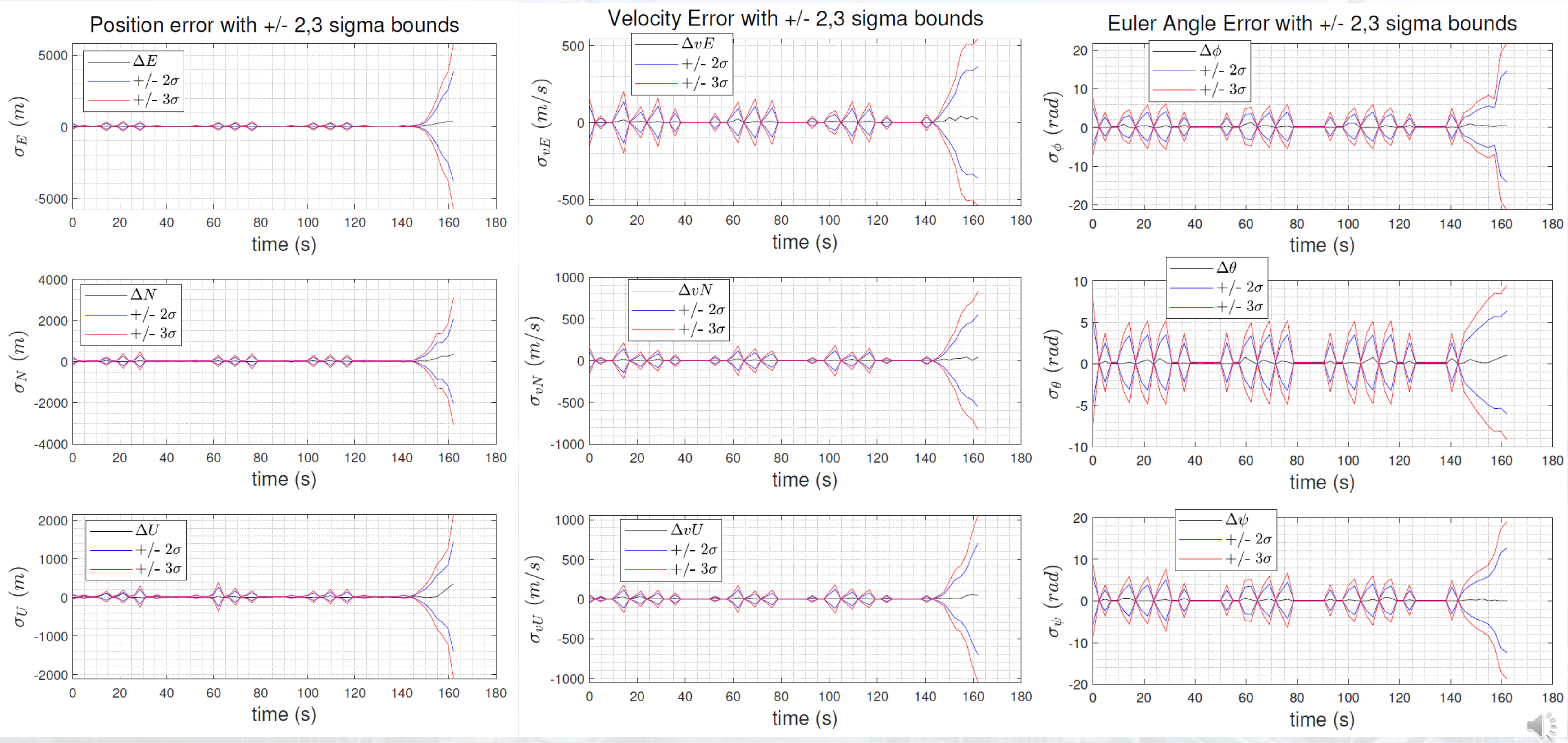
With COPOSIT:

	E (m)	N (m)	U (m)	v_E (m/s)	v_N (m/s)	v_U (m/s)	ϕ (rad)	θ (rad)	ψ (rad)
μ	7.125	5.303	5.202	0.245	0.179	0.106	0.000754	0.00103	0.000742
σ	9.758	6.460	6.763	0.283	0.154	0.131	0.000700	0.00114	0.000903

Without COPOSIT:

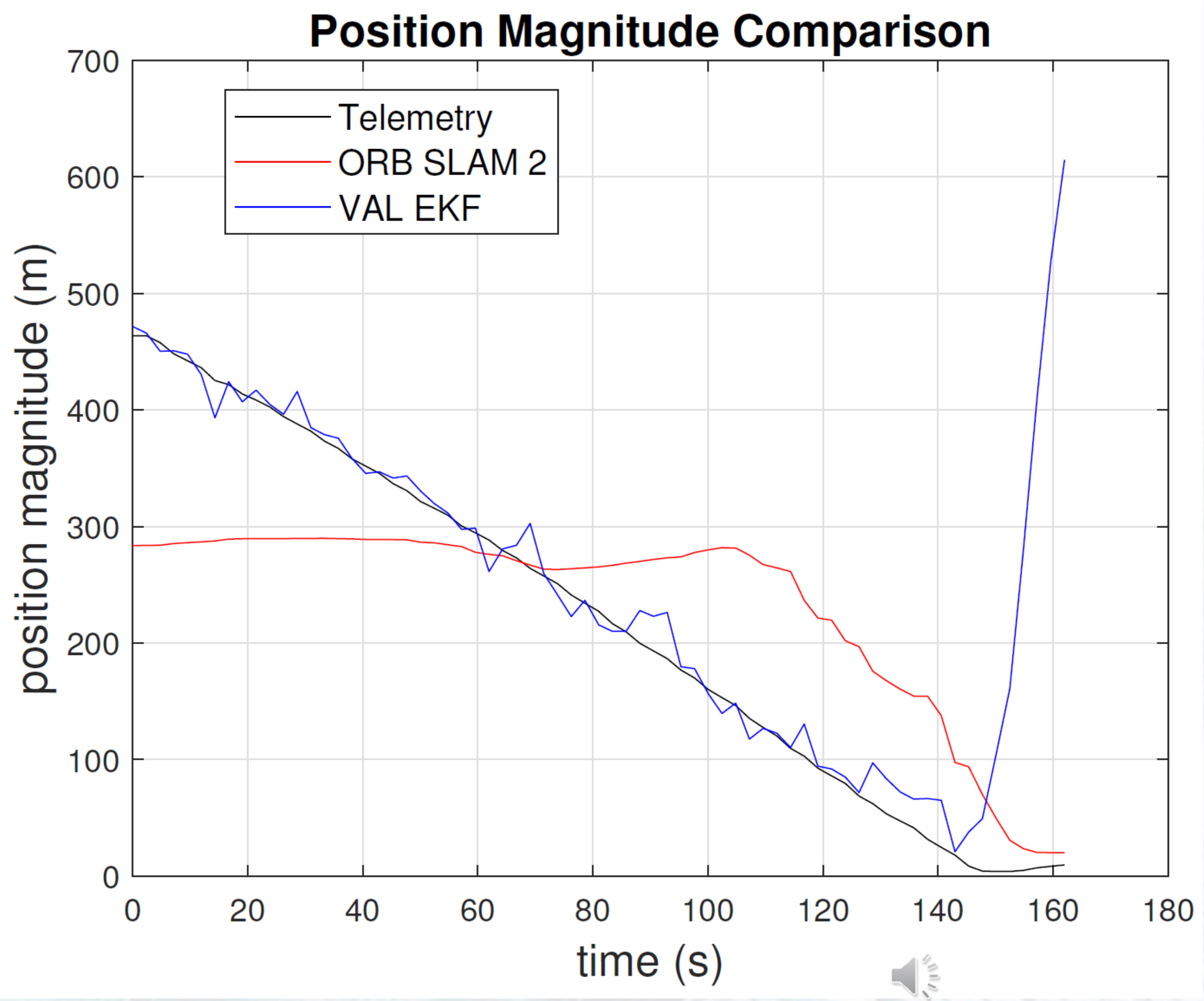
	E (m)	N (m)	U (m)	v_E (m/s)	v_N (m/s)	v_U (m/s)	ϕ (rad)	θ (rad)	ψ (rad)
μ	59.776	53.001	45.483	10.257	10.646	12.478	0.474	0.351	0.339
σ	99.989	84.481	76.818	10.658	11.340	13.305	0.330	0.249	0.228

Results: VAL (COPOSIT-EKF)



Results: ORB SLAM 2 vs. VAL (position)

- Three-way comparison with the **Pixhawk GPS telemetry data** as ground truth
- **ORB SLAM 2** matches better towards the end
- **VAL** matches before landing because features are outside the camera's field of view at the end
- Future work:
  - Improve onboard navigation solution performance and robustness through distributed sensors in the environment/landing zone
  - Combine ORB SLAM 2 or another VSLAM method with VAL to include known and unknown a priori features (best-of-both-worlds approach)



Outline

1. Introduction
2. AFRC Flight Tests and Landmarks
3. Feature Detection for VAL
4. VSLAM Method Selection
5. VAL Design
6. Results
7. Conclusion

Conclusion

- UAS flight tests at AFRC provides experimental results and data to simulate AAM approach and landing
- Initial comparison between ORB SLAM 2 and VAL, a vision-based EKF with IMU & COPOSIT measurements
- Potential APNT solutions with vision but need “eyes” on the landmarks and fiducials throughout the entire approach and landing
- Future work
  - Improve onboard navigation solution performance and robustness through distributed sensors in the environment/landing zone
  - Combine both methods to have known and unknown a priori landmarks and fiducials for accurate state estimation
  - Investigate if feature detection and correspondence yield accurate results at higher cruise velocities
  - Flight tests with helicopters at different conditions (day, night, dawn, dusk, fog, rain, etc.) provides more insight for simulating AAM approach and landing

References

1. Kawamura, E., Dolph, C., Kannan, K., Lombaerts, T., and Ippolito, C. A., “Simulated Vision-based Approach and Landing System for Advanced Air Mobility,” *AIAA SciTech 2023 Forum*, 2023
2. Thompson, N., “NASA National Campaign Build 1, Edwards AFB, California,” National Geospatial-Intelligence Agency, 2020.
3. Webber, D., and Zahn, D., “FAA and the National Campaign,” [Powerpoint], 2021.
4. Kawamura, E., Kannan, K., Lombaerts, T., and Ippolito, C. A., “Vision-Based Precision Approach and Landing for Advanced Air Mobility,” AIAA SciTech 2022 Forum, AIAA 2022-0497, 2022. https://doi.org/10.2514/6.2022-0497

Acknowledgements

Dale Reed Subscale Flight Research Laboratory at NASA Armstrong Flight Research Center (AFRC) for obtaining UAS experimental results to serve as ground truth for the APNT solutions presented in this paper.

Thank you for listening! Questions?