SynTraG: A Synthetic Trajectory Generator for Non-Cooperative Dynamic Obstacles in UAV Navigation

Syed Izzat Ullah, José Baca*

Abstract Unmanned Aerial Vehicles (UAVs) require robust navigation in dynamic environments to avoid collisions with non-cooperative obstacles like birds or other aerial objects, whose trajectories are non-linear, irregular, and not governed by road or air-traffic rules. Data-driven forecasting models such as Recurrent Neural Networks (RNNs) and Transformers, have shown strong performance in trajectory forecasting; however, progress in the UAV domain is constrained by the absence of publicly available data capturing such unstructured aerial motions. In this paper, we present SynTraG, a parametric synthetic trajectory generator that produces a configurable and extensive corpus of 3D trajectories by mixing kinematic primitives that emulate common non-cooperative motions: linear flight, oscillatory weaving, complex looping maneuvers, and vertical undulations. The generator is formulated as a mixture of closed-form kinematic primitives with randomized parameters and explicit controls over path length, speed, acceleration, curvature, and spectral content. Heteroscedastic Gaussian noise is incorporated to model localization uncertainty(e.g., GPS/IMU errors) to enable probabilistic trajectory forecasting. We provide an open-source implementation of SynTraG and a sample dataset of 47,894 diverse trajectories, which enables robust training and benchmarking of forecasting models for UAV collision avoidance with non-cooperative dynamic obstacles. This work enables the development and standardized benchmarking of next-generation forecasting models for safer autonomous aerial navigation.

Syed Izzat Ullah
College of Engineering and Computer Science, Texas A&M University-Corpus Christi, TX, USA
e-mail: sizzatullah@islander.tamucc.edu
José Baca
College of Engineering and Computer Science, Texas A&M University-Corpus Christi, TX, USA
e-mail: Jose.Baca@tamucc.edu
* Corresponding Author

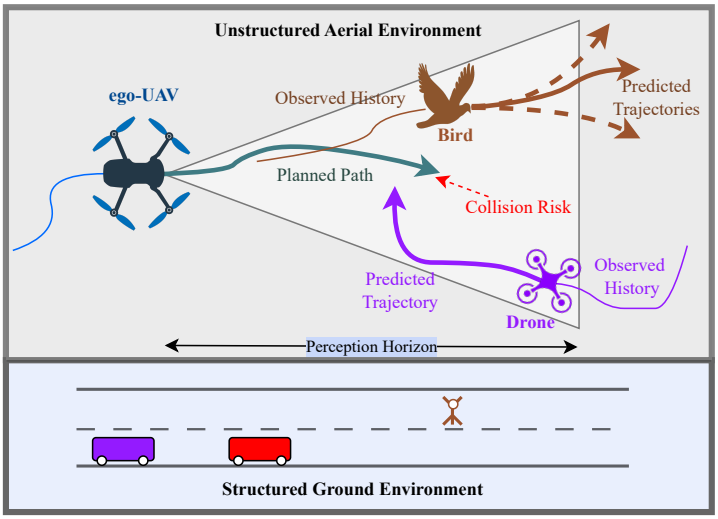

Fig. 1 Illustration of the UAV navigation problem among non-cooperative dynamic obstacles. An ego-UAV (blue) forecasts future trajectories of unpredictable agents like birds (brown) or erratic drones (purple) based on past observed history. This contrasts with the structured, predictable nature of ground-based navigation (bottom), for which ample datasets exist.

1 Introduction

The proliferation of Unmanned Aerial Vehicles (UAVs) in various applications such as surveillance, agriculture, package delivery, and disaster response, has increased the need for robust collision avoidance systems [3, 8, 17]. In unstructured environments, UAVs must handle dynamic obstacles with unpredictable, non-linear trajectories, such as birds or other aerial objects that do not cooperate with the ego-UAV, as shown in Fig. 1. Proactive avoidance requires an accurate forecast of the future trajectory of these obstacles, which enables UAVs to adjust their trajectory preemptively.

Recent advances in deep learning, especially recurrent and Transformer-based sequence models, has shown strong performance in trajectory prediction for pedestrians [10, 12], road vehicles [4, 5], and social navigation [2]. However, these models require large and diverse training data to generalize across various motion patterns.

Existing datasets are focused on ground-based scenarios, such as vehicle or pedestrian tracking (e.g., nuScenes [4] or Waymo Open Motion [7]), which fall short for UAV contexts. Such datasets typically constrain motion to planar or interactive settings, leaving out the complex, erratic, and three-dimensional trajectories seen in bird movements or drone failures. Collecting real-world data for these obstacles presents significant challenges, including ethical issues in interacting with wildlife, safety hazards during flight operations, and the rarity of such collision-prone events. To the best of our knowledge, there are no widely adopted datasets of *non-interactive,*

non-cooperative aerial obstacle trajectories suitable for training forecasting models. As a result, synthetic dataset generation is a pragmatic and effective way to train such models.

We propose *SynTraG*, a parametric synthetic dataset generator to bridge this gap. It employs canonical kinematic equations with randomized parameters to generate diverse 3D trajectories that emulate non-interactive dynamic obstacles. These primitives were selected to form a basis set for complex aerial motion; straight-line motion represents predictable flight paths, sinusoidal and up-and-down patterns model periodic motions like flapping or gliding oscillations, and the trefoil knot serves as a challenging, high-curvature primitive representative of erratic or evasive maneuvers. For example, trefoil patterns capture the cyclic, knot-like maneuvers observed in bird flocks, which can be scaled and offset through parametric models to enhance realism. Furthermore, SynTraG also introduces localization uncertainty to enable risk-aware probabilistic trajectory forecasting for dynamic obstacles. To facilitate adoption, we have open-source the generator code and the data corpus, which comprises around 47,894 trajectory samples.

2 Related Work

2.1 Trajectory Forecasting

Trajectory forecasting for UAVs has gained significant advancement due to its critical role in autonomous navigation, collision avoidance, and air traffic management. Advances are largely driven by sequence modeling frameworks such as recurrent neural networks (RNN), transformers, and convolutional architectures designed for spatio-temporal data [1, 20]. For instance, QCNet-3D [19] introduces a UAV-specific predictor that models 3D motion through correlated axis-wise refinement and performs better than traditional 2D methods by addressing the unique complexities of aerial environments. Learning-based approaches increasingly integrate multimodal data from vision, LiDAR, and event sensors to improve long-horizon forecasting under adverse conditions [13]. Uncertainty quantification and probabilistic models are emerging to better capture the inherent stochasticity in UAV flight paths [11].

2.2 Existing Datasets for UAV Navigation

The development of UAV trajectory prediction methods depends strongly on access to high-quality datasets, which remain scarce in aerial domains. The MiTra dataset [6] leverages drone-collected traffic trajectories to characterize urban aerial movements across diverse traffic states. Similarly, the MMAUD dataset [18] incorporates multimodal sensor data (e.g., LiDAR, radar, and fisheye cameras) in urban UAV interception scenarios, enabling realistic evaluations in noisy and dynamic en-

vironments. Other datasets such as BuckTales [14] focus on multi-UAV tracking in wildlife habitats, capturing relatively predictable collective dynamics unsuitable for non-cooperative airborne obstacles. Datasets tailored to erratic or adversarial drones, or bio-inspired motions (e.g., birds), remain an important gap in the literature.

2.3 Synthetic Dataset for Motion Prediction

Synthetic trajectory generation has become a crucial tool for augmenting limited real-world UAV data, enabling more diverse and complex motion pattern training. Simulators such as AirSim [16] and SynDrone [15] enable configurable multi-agent UAV flights in varied environments but primarily focus on structured, cooperative scenarios. Transformer-based generative models, such as TrajGPT [9], advance controlled synthesis of realistic trajectories under learned spatio-temporal distributions. Probabilistic diffusion models such as DiffTraj [21] further support the creation of privacy-preserving GPS trajectories for improved forecasting pipelines. However, existing synthetic and real-world datasets remains limited in representing erratic, bio-inspired, or adversarial flight dynamics essential for collision avoidance in open airspace. Furthermore, existing generative models often learn from real-world data distributions that are unavailable for our target domain. SynTraG, therefore, adopts a first-principles, parametric approach of kinematically diverse and challenging aerial motions.

3 The SynTraG Framework

This section explains the mathematical modeling of the proposed *SynTraG* framework, explaining its key components for generating synthetic trajectories. This includes the definition of trajectory forecasting task, generator's objective, mixture prior governing the selection of motion primitives, the conditional priors for primitive-specific parameters, the state-transition dynamics incorporating uncertainty modeling, and the kinematic equations for each motion primitive. Fig. 2 shows the overall system diagram of the framework.

3.1 Mathematical Formulation

3.1.1 Forecasting task

The trajectory forecasting problem is defined as given a sequence of historical states of a dynamic obstacle over a time horizon T, predict its future states over a subsequent time horizon N. The state at time t is defined as $\mathbf{s}_t = [x_t, y_t, z_t, v_{x,t}, v_{y,t}, v_{z,t}]^T \in \mathbb{R}^6$

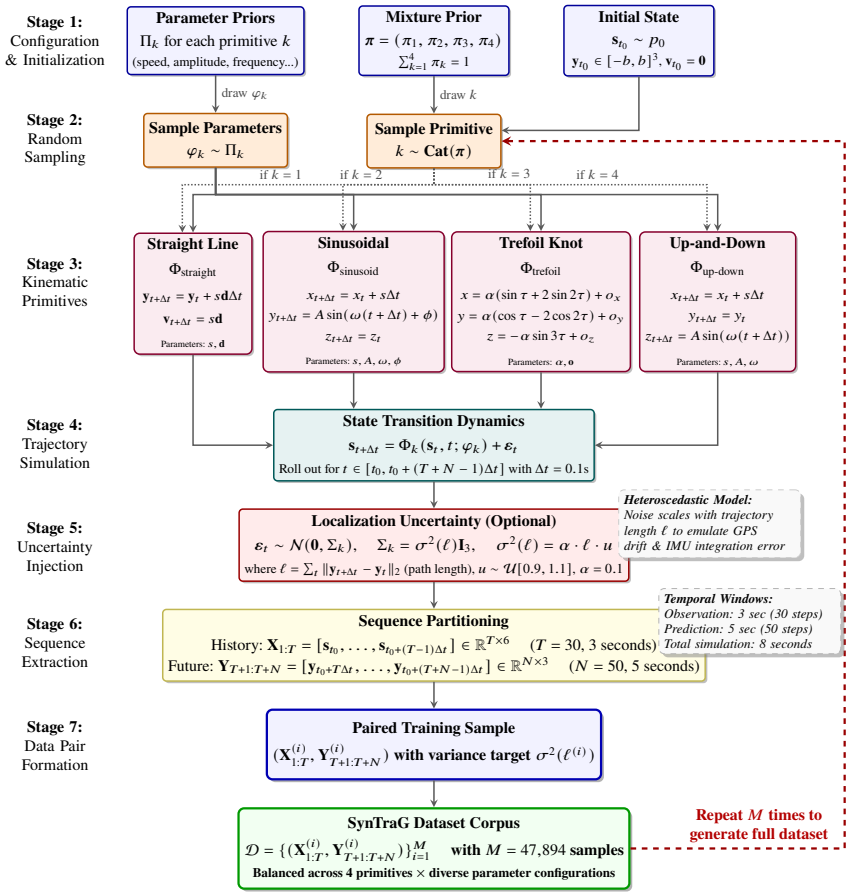

Fig. 2 SynTraG generative pipeline architecture. The framework synthesizes paired trajectory sequences through various stages: (*Stage 1*) User-configurable mixture prior $\boldsymbol{\pi}$ and parameter priors Π_k are specified along with initial state distribution p_0. (*Stage 2*) For each sample, a primitive index k and parameters φ_k are stochastically drawn. (*Stage 3*) The selected kinematic model Φ_k (one of four closed-form motion primitives) governs trajectory evolution. (*Stage 4*) State transition dynamics (Eq. 3) roll out the complete 8-second trajectory. (*Stage 5*) Optional heteroscedastic Gaussian noise $\boldsymbol{\varepsilon}_t$ models localization uncertainty proportional to path length. (*Stage 6*) The continuous trajectory is partitioned into 3-second historical observations $\mathbf{X}_{1:T}$ and 5-second future predictions $\mathbf{Y}_{T+1:T+N}$. (*Stage 7*) A paired training sample with variance target is formed. This process repeats $M = 47{,}894$ times with different random seeds to construct the complete dataset corpus $\mathcal{D}$, ensuring balanced representation across all motion primitives and diverse parameter configurations (see Table 1 for statistical summary).

representing the obstacle's state at time t, comprising its position $[x_t, y_t, z_t]^T \in \mathbb{R}^3$ and velocity $[v_{x,t}, v_{y,t}, v_{z,t}]^T \in \mathbb{R}^3$ in a 3D Cartesian coordinate system relative to a world frame. The input is a historical sequence of states $\mathbf{X}_{1:T} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_T] \in \mathbb{R}^{T\times 6}$, where T is the length of the input sequence, corresponding to a time window of $(T \cdot \Delta t)$ seconds with a time step Δt seconds. The output is forecast of future positions

$\mathbf{Y}_{T+1:T+N} = [\mathbf{y}_{T+1}, \mathbf{y}_{T+2}, \ldots, \mathbf{y}_{T+N}] \in \mathbb{R}^{N\times 3}$, where $\mathbf{y}_t = [x_t, y_t, z_t]^T \in \mathbb{R}^3$ denotes the 3D position coordinates of obstacle's trajectory at time $t = T+1, \ldots, T+N$, and N is the prediction horizon, spanning $(N \cdot \Delta t)$ seconds. A prediction model $f_\theta : \mathbb{R}^{T\times 6} \to \mathbb{R}^{N\times 3}$, with learnable parameters θ, maps $\mathbf{X}_{1:T}$ to $\widehat{\mathbf{Y}}_{T+1:T+N}$.

3.1.2 Generator Objective

The goal of SynTraG is to synthesize a large set of paired sequences:

$$\mathcal{D} = \left\{(\mathbf{X}^{(i)}_{1:T}, \mathbf{Y}^{(i)}_{T+1:T+N})\right\}_{i=1}^{M} \tag{1}$$

consisting of M paired sequences, where each pair is derived from procedurally simulated trajectories. These trajectories are generated by sampling from a mixture of canonical motion primitives and rolling out closed-form kinematic equations under randomized parameters. The framework provides control over (i) the mixture weights across motion families (straight, sinusoid, up-and-down, and trefoil) and (ii) primitive-specific parameter distributions, enabling tunable balance and diversity to support robust model training and out-of-distribution (OOD) evaluation.

3.1.3 Mixture Prior over Motion Primitives

For each sample, a motion primitive index $k \in \{1, \ldots, 4\}$ is drawn from a *categorical* distribution:

$$k \;\sim\; \texttt{Cat}(\boldsymbol{\pi}), \qquad \boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4), \quad \sum_{k=1}^{4} \pi_k = 1 \tag{2}$$

Here, $\texttt{Cat}(\boldsymbol{\pi})$ governs the probabilities over the four primitives: `straight`, `sinusoid`, `trefoil`, `up_and_down`. A uniform prior ($\pi_k = 1/4$) ensures balanced generation, while skewed weights (e.g., setting $\pi_{\texttt{trefoil}} = 0$) can induce OOD scenarios, such as primitive hold-out during training.

3.1.4 Primitive-Specific Parameter Priors

Conditioned on k, a vector of motion parameters φ_k is drawn from a user-configurable prior Π_k:

$$\varphi_k \;\sim\; \Pi_k$$

Typically, Π_k factorizes into independent uniform distributions over primitive-specific parameters, controlling aspects such as amplitude, frequency, phase, offsets, and bounds on speed or acceleration. For instance, `sinusoid`: $A \sim \mathcal{U}[A_{\min}, A_{\max}]$, $\omega \sim \mathcal{U}[\omega_{\min}, \omega_{\max}]$, $\phi \sim \mathcal{U}[0, 2\pi]$; for `straight`: speed $s \sim \mathcal{U}[s_{\min}, s_{\max}]$ and di-

rection $\mathbf{d} \in \mathbb{S}^2$). These priors shape trajectory length, speed/acceleration, curvature, and spectral content.

3.1.5 State Transition with Optional Uncertainty

Given (k, φ_k), the state evolves according to discrete-time dynamics:

$$\mathbf{s}_{t+\Delta t} \;=\; \Phi_k(\mathbf{s}_t, t;\, \varphi_k) \;+\; \varepsilon_t \tag{3}$$

where Φ_k is the deterministic kinematic mapping for primitive k, and ε_t is an optional noise term that models sensor inaccuracies or environmental perturbations. For uncertainty-aware generation, we employ zero-mean Gaussian noise:

$$\varepsilon_t \sim \mathcal{N}(\mathbf{0},\, \Sigma_k), \qquad \Sigma_k \;=\; \sigma^2(\ell)\, I_3, \quad \sigma^2(\ell) \;=\; \alpha\, \ell \cdot u,\; u \sim \mathcal{U}[u_{\min}, u_{\max}] \tag{4}$$

where ℓ is the sample's path length (i.e., $\ell = \sum_t \|\mathbf{y}_{t+\Delta t} - \mathbf{y}_t\|_2$) and α is the scaling factor (e.g., α=0.1) that controls the overall noise magnitude. u is a random variable drawn from a uniform distribution $\mathcal{U}[u_{\min}, u_{\max}]$ that introduces randomness to the variance for each trajectory to make them more diverse. This produces heteroscedastic variance targets compatible with probabilistic predictors.

3.1.6 Generative Program for One Sample

Let t_0 denote the start of the observation window. With an initial state $\mathbf{s}_{t_0}$ (sampled from a bounded prior over positions and zero initial velocity), the data pair $(\mathbf{X}, \mathbf{Y})$ is obtained by rolling out Eq. (3):

$$\begin{aligned}
&k \sim \mathtt{Cat}(\boldsymbol{\pi}), \quad \varphi_k \sim \Pi_k, \quad \mathbf{s}_{t_0} \sim p_0,\\
&\mathbf{s}_{t+\Delta t} = \Phi_k(\mathbf{s}_t, t; \varphi_k) + \varepsilon_t \;\text{ for }\; t \in [t_0,\, t_0{+}(T{+}N{-}1)\Delta t],\\
&\mathbf{X}_{1:T} = [\mathbf{s}_{t_0}, \ldots, \mathbf{s}_{t_0+(T-1)\Delta t}], \quad \mathbf{Y}_{T+1:T+N} = [\mathbf{y}_{t_0+T\Delta t}, \ldots, \mathbf{y}_{t_0+(T+N-1)\Delta t}]
\end{aligned}$$

with $\mathbf{y}_t$ the position component of $\mathbf{s}_t$. Repeating this program M times yields the dataset $\mathcal{D}$, with ε_t ensuring probabilistic variance, if enabled.

3.1.7 Induced Mixture over Data Pairs

The generative process induces a probability distribution over data pairs $(\mathbf{X}_{1:T}, \mathbf{Y}_{T+1:T+N})$, modeled as a finite mixture of four components corresponding to the four kinematic primitives:

$$p(\mathbf{X}_{1:T}, \mathbf{Y}_{T+1:T+N}) = \sum_{k=1}^{4} \pi_k \int p(\mathbf{X}_{1:T}, \mathbf{Y}_{T+1:T+N} \mid k, \varphi_k)\, \mathrm{d}\Pi_k(\varphi_k) \tag{5}$$

Here, the mixture weights $\boldsymbol{\pi}$ control the balance between primitives, determining the relative frequency of each motion type in the dataset (e.g., more straight vs. trefoil trajectories). The parameter priors Π_k govern intra-primitive variability, allowing diverse trajectory characteristics (e.g., varying speeds or amplitudes within the sinusoidal primitive). This mixture structure ensures a rich and diverse dataset $\mathcal{D}$ which captures a wide range of non-cooperative obstacle behaviors. For evaluation, in-distribution (ID) protocols use the full mixture, while out-of-distribution (OOD) protocols can be configured by modifying $\boldsymbol{\pi}$ (e.g., setting $\pi_k = 0$ to exclude a primitive) or restricting Π_k's support (e.g., limiting amplitude ranges to test generalization to unseen parameter).

3.2 Core Kinematic Motion Models

This subsection defines the kinematic equations for the four motion primitives in SynTraG: straight-line, sinusoidal, trefoil knot, and up-and-down. These primitives, implemented via the deterministic mapping $\Phi_k(\mathbf{s}_t, t; \varphi_k)$ in Eq. (3), generate diverse 3D trajectories that emulate non-cooperative aerial obstacles (e.g., birds or any other non-cooperative aerial object) within a bounded workspace $[-\mathbf{b}, \mathbf{b}]^3 \subset \mathbb{R}^3$. The state $\mathbf{s}_t = [\mathbf{y}_t^\top, \mathbf{v}_t^\top]^\top \in \mathbb{R}^6$, comprising position $\mathbf{y}_t = [x_t, y_t, z_t]^\top$ and velocity $\mathbf{v}_t = [v_{x,t}, v_{y,t}, v_{z,t}]^\top$, evolves with a fixed timestep Δt. Each primitive is governed by parameters φ_k sampled from priors Π_k, as defined in Section 3.1.3 that ensures variability in trajectory characteristics such as length, speed, and curvature. These models are designed to capture realistic motion patterns, from linear flights to complex, bio-inspired maneuvers, supporting robust trajectory forecasting for UAV collision avoidance.

3.2.1 Straight-Line (Constant Velocity) Motion

This primitive models linear, non-maneuvering motion, such as a drone's steady flight or a bird's gliding path. The state evolves as:

$$\mathbf{y}_{t+\Delta t} = \mathbf{y}_t + s\mathbf{d}\Delta t \tag{6a}$$

$$\mathbf{v}_{t+\Delta t} = s\mathbf{d} \tag{6b}$$

where speed $s \sim \mathcal{U}[s_{\min}, s_{\max}]$ controls the magnitude of motion, and direction $\mathbf{d} \in \mathbb{S}^2$ (a unit vector on the 2-sphere) defines the 3D orientation. Parameters: $\varphi_{\text{straight}} = \{s, \mathbf{d}\}$. Acceleration is zero ($\mathbf{a}_t = \mathbf{0}$), reflecting constant velocity.

3.2.2 Sinusoidal Motion

This captures lateral oscillatory motion and approximate bird-like weaving in the y-direction. The state updates are:

$$x_{t+\Delta t} = x_t + s\Delta t \tag{7a}$$
$$y_{t+\Delta t} = A\sin(\omega(t+\Delta t)+\phi) \tag{7b}$$
$$z_{t+\Delta t} = z_t \tag{7c}$$
$$v_{x,t+\Delta t} = s \quad v_{y,t+\Delta t} = A\omega\cos(\omega(t+\Delta t)+\phi) \quad v_{z,t+\Delta t} = 0 \tag{7d}$$

with speed $s \sim \mathcal{U}[s_{\min}, s_{\max}]$, amplitude $A \sim \mathcal{U}[A_{\min}, A_{\max}]$, frequency $\omega \sim \mathcal{U}[\omega_{\min}, \omega_{\max}]$, and phase $\phi \sim \mathcal{U}[0, 2\pi]$. Parameters: $\varphi_{\text{sinusoid}} = \{s, A, \omega, \phi\}$. Acceleration: $a_{y,t} = -A\omega^2\sin(\omega t + \phi)$.

3.2.3 Trefoil-Knot Motion

This models complex, looping 3D maneuvers, inspired by erratic bird flock patterns or evasive drone behaviors. Let $\tau = t + \Delta t$. The state evolves as:

$$x_{t+\Delta t} = \alpha(\sin\tau + 2\sin 2\tau) + o_x \tag{8a}$$
$$y_{t+\Delta t} = \alpha(\cos\tau - 2\cos 2\tau) + o_y \tag{8b}$$
$$z_{t+\Delta t} = -\alpha\sin 3\tau + o_z \tag{8c}$$
$$v_{x,t+\Delta t} = \alpha(\cos\tau + 4\cos 2\tau) \tag{8d}$$
$$v_{y,t+\Delta t} = \alpha(-\sin\tau + 4\sin 2\tau) \tag{8e}$$
$$v_{z,t+\Delta t} = -3\alpha\cos 3\tau \tag{8f}$$

with scale $\alpha \sim \mathcal{U}[\alpha_{\min}, \alpha_{\max}]$ and offset $\mathbf{o} = [o_x, o_y, o_z]^\top \sim \mathcal{U}[-o_{\max}, o_{\max}]^3$. Parameters: $\varphi_{\text{trefoil}} = \{\alpha, \mathbf{o}\}$. Accelerations are derived from second derivatives of the position equations.

3.2.4 Up-and-Down Motion

This represents vertical oscillatory motion, such as flapping wings or altitude-varying drone flight. The state updates are:

$$x_{t+\Delta t} = x_t + s\Delta t \tag{9a}$$
$$y_{t+\Delta t} = y_t \tag{9b}$$
$$z_{t+\Delta t} = A\sin(\omega(t+\Delta t)) \tag{9c}$$
$$v_{x,t+\Delta t} = s, \quad v_{y,t+\Delta t} = 0, \quad v_{z,t+\Delta t} = A\omega\cos(\omega(t+\Delta t)) \tag{9d}$$

with speed $s \sim \mathcal{U}[s_{\min}, s_{\max}]$, amplitude $A \sim \mathcal{U}[A_{\min}, A_{\max}]$, and frequency $\omega \sim \mathcal{U}[\omega_{\min}, \omega_{\max}]$. Parameters: $\varphi_{\text{up-and-down}} = \{s, A, \omega\}$. Acceleration is non-zero only in the z-direction: $a_{z,t} = -A\omega^2 \sin(\omega t)$.

4 Reference Instantiation and Dataset Corpus

This section details the instantiation of the SynTraG framework through the generation of a reference dataset corpus, designed to support robust trajectory forecasting for non-cooperative dynamic obstacles in UAV navigation. The dataset, denoted $\mathcal{D}$, comprises of 47,894 paired sequences, each consisting of historical state observations and future position predictions, synthesized using the kinematic primitives and generative process described in Sections 3.1 and 3.2. This corpus serves as a benchmark for evaluating forecasting models and is made publicly available to facilitate research in safety-critical aerial navigation.

4.1 Data Generation Process

The dataset is constructed by simulating trajectories over an 8-second temporal window, partitioned into a 3-second historical observation period ($T = 30$ timesteps with $\Delta t = 0.1$ seconds) and a 5-second future prediction horizon ($N = 50$ timesteps). Historical sequences $\mathbf{X}_{1:T} \in \mathbb{R}^{30\times6}$ capture the state $\mathbf{s}_t = [\mathbf{y}_t^\top, \mathbf{v}_t^\top]^\top \in \mathbb{R}^6$, where $\mathbf{y}_t = [x_t, y_t, z_t]^\top$ and $\mathbf{v}_t = [v_{x,t}, v_{y,t}, v_{z,t}]^\top$ represent position and velocity in a 3D Cartesian frame. Future sequences $\mathbf{Y}_{T+1:T+N} \in \mathbb{R}^{50\times3}$ provide the predicted positions $\mathbf{y}_t$ over the forecast horizon. Trajectories are sampled at intervals aligned with approximately one-quarter of the dominant motion period (e.g., $2\pi/\omega$ for sinusoidal or up-and-down primitives), ensuring diverse phase representations and comprehensive coverage of the dynamic behaviors induced by the mappings $\Phi_k(\mathbf{s}_t, t; \varphi_k)$. Initial states $\mathbf{s}_{t_0}$ are drawn from a bounded uniform prior over positions within $[-\mathbf{b}, \mathbf{b}]^3$ with zero initial velocity, and the rollout incorporates optional heteroscedastic Gaussian noise $\varepsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma_k)$ as per Eq. (3), where $\Sigma_k = \sigma^2(\ell) I_3$ and $\sigma^2(\ell) = 0.1 \cdot \ell \cdot u$ with $u \sim \text{Unif}(0.9, 1.1)$. This noise, proportional to path length $\ell = \sum_t \|\mathbf{y}_{t+\Delta t} - \mathbf{y}_t\|_2$, models localization uncertainty and supports probabilistic forecasting.

4.2 Dataset Composition and Visualization

The resulting dataset is structured as $\mathbf{X} \in \mathbb{R}^{47894\times30\times6}$ for observations and $\mathbf{Y} \in \mathbb{R}^{47894\times50\times3}$ for targets, reflecting a balanced distribution across the four motion primitives: straight-line, sinusoidal, trefoil knot, and up-and-down. A 3D visualization of representative trajectories is provided in Fig. 3, illustrating the diversity in

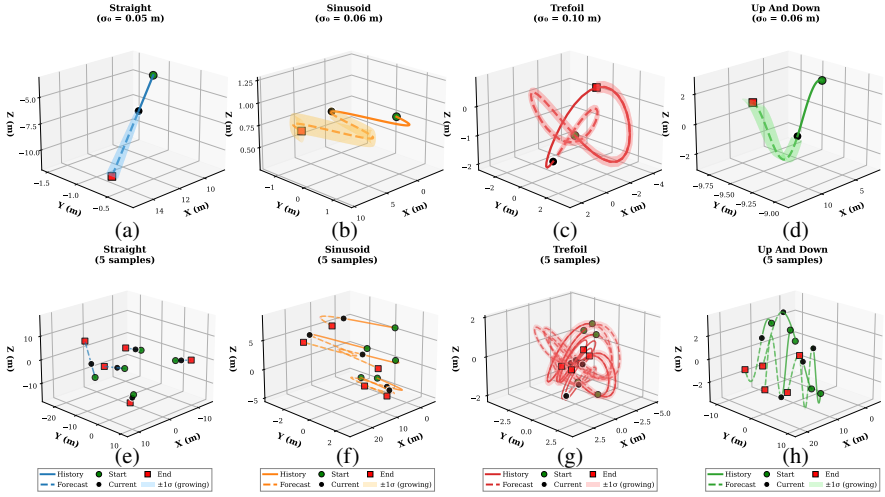

Fig. 3 Representative 3D trajectories for each motion primitive with uncertainty quantification. Subplots (a-d) show a single trajectory of each exemplar trajectory, while plots (e-f) show the 5 samples of the same trajectories. solid color segments represent observed history (3s), circle shows the current position, dashed segments show ground-truth future (5s), and shaded segments indicate $\pm 1\sigma$ uncertainty from the heteroscedastic noise model.

initial positions and motion patterns, from linear paths to complex loops and vertical oscillations.

4.3 Statistical Analysis

The dataset's statistical properties are summarized in Table 1, which reports the sample count, average trajectory length, minimum and maximum lengths, average speed, and variance metrics for each motion type. The average length is computed as the cumulative Euclidean distance over the 8-second simulation, while average speed is derived from the mean magnitude of the velocity vector $\|\mathbf{v}_t\| = \sqrt{v_{x,t}^2 + v_{y,t}^2 + v_{z,t}^2}$. Variance estimates, reflecting the heteroscedastic noise model, are calculated as 10% of the path length scaled by a uniform random factor $\sim \text{Unif}(0.9, 1.1)$, providing a realistic uncertainty profile for training mixture density networks (MDNs) or similar probabilistic models.

Table 1 also reveals the dataset's diversity. For instance, sinusoidal and up-and-down trajectories exhibit greater length and speed variability, reflecting their oscillatory nature, while trefoil knots maintain consistent lengths due to their parametric structure.

Table 1 Dataset statistics per motion type.

Motion Type	Sample Count	Avg Length (m)	Min Length (m)	Max Length (m)	Avg Speed (m/s)	Min Speed (m/s)	Max Speed (m/s)	Avg. σ^2 (m^2)	Std. Dev. (m)
`sinusoid`	10280	44.27	8.22	116.53	4.07	1.00	12.38	1.99	1.24
`straight`	13650	27.24	6.65	66.16	1.26	0.50	2.00	0.62	0.22
`trefoil`	11568	38.53	18.21	59.95	4.60	4.43	4.79	2.23	0.65
`up_and_down`	12396	46.35	8.91	125.51	4.25	1.00	13.27	2.08	1.33

4.4 Accessibility and Usage

The complete dataset corpus in Python pickle (.pkl) format, along with the code and a detailed implementation for the SynTraG framework, is publicly available under the MIT License and accessible at: https://github.com/syediu/SynTraG.git. Users can leverage the dataset for training forecasting models, with pre-computed variance estimates facilitating probabilistic predictions. The framework's configurability allows for regeneration of $\mathcal{D}$ with adjusted mixture weights $\boldsymbol{\pi}$ or parameter priors Π_k, enabling tailored datasets for specific research objectives.

5 Conclusion

This paper introduces SynTraG, an open-source synthetic trajectory generator addressing the critical need for data to train UAV collision avoidance systems against non-cooperative dynamic obstacles. By synthesizing 47,894 diverse 3D trajectories via mixture of four kinematic primitives with optional uncertainty, SynTraG provides a flexible, open-source tool for robust forecasting model development. Reference instantiation demonstrate its utility, while the publicly available dataset and code facilitate further research. Future work will explore real-world validation, integration with multi-modal sensors, and extension to adversarial motion patterns.

Acknowledgements This work is supported by the National Science Foundation (NSF) Grant Number 2131263, and the Geospatial Computer Science Program at Texas A&M University-Corpus Christi, USA. This paper solely reflects the opinions and conclusions of its authors and not NSF or any other entity.

References

1. Al-Zadjali, N.S.H., Balasubaramanian, S., Savarimuthu, C., Rances, E.O.: Bio-inspired motion detection models for improved uav and bird differentiation: a novel deep learning framework. Scientific Reports (2025)

2. Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., Savarese, S.: Social lstm: Human trajectory prediction in crowded spaces. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
3. Baca, J., Ullah, S.I., Rangel, P.: Coaxial modular aerial system and the reconfiguration applications. In: IEEE International Conference on Robotics and Automation (ICRA) (2023)
4. Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O.: nuscenes: A multimodal dataset for autonomous driving. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (2020)
5. Chang, M.F., Lambert, J., Sangkloy, P., Singh, J., Bak, S., Hartnett, A., Wang, D., Carr, P., Lucey, S., Ramanan, D., Hays, J.: Argoverse: 3d tracking and forecasting with rich maps. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
6. Chaudhari, A., Treiber, M., Okhrin, O.: Mitra: A drone-based trajectory data for an all-traffic-state inclusive freeway with ramps. Scientific Data (2025)
7. Ettinger, S., Cheng, S., Caine, B., Liu, C., Zhao, H., Pradhan, S., Chai, Y., Sapp, B., Qi, et al.: Large scale interactive motion forecasting for autonomous driving: The waymo open motion dataset. In: Proceedings of the IEEE/CVF international conference on computer vision (2021)
8. Hassanalian, M., Abdelkefi, A.: Classifications, applications, and design challenges of drones: A review. Progress in Aerospace Sciences (2017)
9. Hsu, S.L., Tung, E., Krumm, J., Shahabi, C., Shafique, K.: Trajgpt: Controlled synthetic trajectory generation using a multitask transformer-based spatiotemporal model. In: Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems. Association for Computing Machinery, USA (2024)
10. Huang, Y., Du, J., Yang, Z., Zhou, Z., Zhang, L., Chen, H.: A survey on trajectory-prediction methods for autonomous driving. IEEE Transactions on Intelligent Vehicles (2022)
11. Ivanovic, B., Lin, Y., Shrivastava, S., Chakravarty, P., Pavone, M.: Propagating state uncertainty through trajectory forecasting. In: international conference on robotics and automation (ICRA). IEEE (2022)
12. Kothari, P., Kreiss, S., Alahi, A.: Human trajectory forecasting in crowds: A deep learning perspective. IEEE Transactions on Intelligent Transportation Systems (2021)
13. Liang, H., Yuan, S., Liu, F., Yang, Y., Wang, B., Huang, Z., Shi, C., Jin, J.: Label-free long-horizon 3d uav trajectory prediction via motion-aligned rgb and event cues (2025)
14. Naik, H., Yang, J., Das, D., Crofoot, M.C., Rathore, A., Sridhar, V.H.: Bucktales: a multi-uav dataset for multi-object tracking and re-identification of wild antelopes. In: Proceedings of the 38th International Conference on Neural Information Processing Systems, NIPS (2025)
15. Rizzoli, G., Barbato, F., Caligiuri, M., Zanuttigh, P.: Syndrone-multi-modal uav dataset for urban scenarios. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (2023)
16. Shah, S., Dey, D., Lovett, C., Kapoor, A.: Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In: Field and service robotics: Results of the 11th international conference. Springer (2017)
17. Ullah, S.I., Muhammad, A.: Autonomous navigation and mapping of water channels in a simulated environment using micro-aerial vehicles. In: International Conference on Robotics and Automation in Industry (ICRAI) (2023)
18. Yuan, S., Yang, Y., Nguyen, T.H., Nguyen, T.M., Yang, J., Liu, F., Li, J., Wang, H., Xie, L.: Mmaud: A comprehensive multi-modal anti-uav dataset for modern miniature drone threats. In: IEEE International Conference on Robotics and Automation (ICRA) (2024)
19. Zhang, J., Meng, Z., Liu, S., Ji, J., He, J.: A novel trajectory prediction method for uav air combat based on qcnet-3d. Defence Technology (2025)
20. Zhou, H., Yang, X., Fan, M., Qi, L., Li, X., Yang, M.H., Luo, F.: Three-dimensional trajectory prediction with 3dmotraj dataset. In: Forty-second International Conference on Machine Learning (2025)
21. Zhu, Y., Ye, Y., Zhang, S., Zhao, X., Yu, J.J.: Difftraj: generating gps trajectory with diffusion probabilistic model. In: Proceedings of the 37th International Conference on Neural Information Processing Systems, NIPS (2023)