

Article

Event-Driven Spatiotemporal Computing for Robust Flight Arrival Time Prediction: A Probabilistic Spiking Transformer Approach

Quanquan Chen and Meilong Le *

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;
chenquanquan@nuaa.edu.cn

* Correspondence: mlle@nuaa.edu.cn

Abstract

Precise Estimated Time of Arrival (ETA) prediction in Terminal Maneuvering Areas (TMA) constitutes a prerequisite for efficient arrival sequencing and airspace capacity management. While data-driven approaches outperform kinematic models, conventional Recurrent Neural Networks (RNNs) exhibit limitations in modeling complex multi-aircraft spatial interactions and lack the capability to quantify predictive uncertainty. Conversely, Spiking Neural Networks (SNNs) enable energy-efficient event-driven computation, yet their applicability to continuous trajectory regression is hindered by "input starvation," where normalized state vectors fail to induce sufficient neural firing rates. This study proposes a Probabilistic Spiking Transformer (PST) architecture to integrate neuro-morphic sparsity with global attention mechanisms. An Adaptive Spiking Temporal Encoding mechanism incorporating learnable linear projections is introduced to resolve the regression-spiking incompatibility, facilitating the autonomous mapping of continuous trajectory dynamics into sparse spike trains without heuristic scaling. Concurrently, a Distance-Biased Multi-Aircraft Cross-Attention (MACA) module models air traffic conflicts by weighting spatial interactions according to physical proximity, thereby embedding separation constraints into the feature extraction process. Evaluation on large-scale real-world ADS-B datasets demonstrates that the PST yields a Mean Absolute Error (MAE) of 49.27 s, representing a 60% error reduction relative to standard LSTM baselines. Furthermore, the model generates well-calibrated probabilistic distributions (Prediction Interval Coverage Probability > 94%), offering quantifiable uncertainty metrics for risk-based decision support while ensuring real-time inference suitable for operational deployment.

Keywords: Estimated Time of Arrival (ETA); spiking transformer; adaptive input encoding; uncertainty quantification; air traffic management; event-driven computing

Academic Editor: Miroslav Kelemen

Received: 20 January 2026

Revised: 14 February 2026

Accepted: 16 February 2026

Published: 22 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The unprecedented growth of global air traffic has placed substantial pressure on Terminal Maneuvering Areas (TMA), where high-density traffic streams converge and diverge under stringent safety separation standards. Accurate Estimated Time of Arrival (ETA) prediction constitutes the cornerstone of Trajectory-Based Operations (TBO) and arrival sequencing systems. Precise ETA predictions enable air traffic controllers (ATCOs) to optimize landing sequences, minimize fuel consumption during holding patterns, and

enhance runway throughput [1,2]. However, the TMA environment is characterized by stochastic dynamics, including frequent vectoring, altitude modifications, and complex inter-aircraft interactions, rendering robust short-term trajectory prediction a persistent operational challenge [3].

Historically, ETA prediction relied on kinematic models (e.g., BADA) or look-up tables. While physically interpretable, these methods exhibit limited adaptability to uncertain meteorological conditions and dynamic human intervention factors [4]. Consequently, data-driven approaches, specifically Recurrent Neural Networks (RNNs) such as LSTM and GRU, have emerged as the dominant paradigm due to their capacity to extract temporal dependencies from historical ADS-B data [5–7]. For instance, Shi et al. [5] proposed a constrained LSTM for flight trajectory prediction, yielding statistically significant accuracy improvements over Kalman filters, while Lin et al. [8] demonstrated the efficacy of similar deep learning approaches in tactical air traffic management. Despite these advancements, RNN-based models are inherently constrained by sequential computation limitations and often fail to capture global temporal contexts within extended sequences.

To address these limitations, the Transformer architecture, leveraging self-attention mechanisms, has been adopted for time-series modeling. Transformers excel at capturing long-range dependencies and facilitating parallelized computation [9]. Recent applications in aviation include a Transformer-VAE framework for trajectory segmentation and clustering in terminal airspace, demonstrating superior capability in disentangling complex flight patterns and latent structural features [10]. Building on such advancements, attention mechanisms have been applied to trajectory prediction [11]. However, standard Transformers compute attention across all time steps with quadratic complexity, resulting in high computational redundancy, particularly when processing sparse trajectory events. Moreover, the majority of existing deep learning models remain deterministic, providing point estimates without quantifying prediction uncertainty—a critical requirement for risk-aware decision-making in safety-critical ATC scenarios [12,13].

Spiking Neural Networks (SNNs), recognized as the third generation of neural networks, present a viable pathway for energy-efficient and event-driven computation. By mimicking biological neurons (e.g., Leaky Integrate-and-Fire models), SNNs process information only when input signals exceed a firing threshold, offering theoretically high sparsity and low latency [14–16]. While the integration of SNNs with Transformer architectures (Spiking Transformers) has shown promise in processing visual and temporal data [17,18], applying SNNs to continuous trajectory regression faces the critical challenge of “input starvation.” Standard trajectory normalization techniques frequently result in input currents that fail to breach the neuron's firing threshold, leading to “dead neurons” and model degradation—a phenomenon observed and analyzed in our preliminary experiments. Furthermore, existing SNN architectures typically lack explicit mechanisms to model spatial interactions (e.g., separation maintenance) between neighboring aircraft [19].

In this paper, we propose a Probabilistic Spiking Transformer (PST) framework for robust and interpretable ETA prediction in terminal airspace. To our knowledge, this represents one of the primary attempts to integrate the event-driven efficiency of SNNs with the global modeling capacity of Transformers for trajectory tasks. The main contributions of this study are:

1. A Novel Spiking-Transformer Architecture: We design a hybrid encoder that leverages LIF neurons to extract sparse, event-driven features from trajectory data, integrated with a self-attention mechanism to capture global temporal dependencies.

2. Adaptive Spiking Temporal Encoding: To resolve the “input starvation” and “dead neuron” pathologies in SNN regression, we propose an adaptive encoding strategy using learnable linear projections. This ensures robust neural activation and preserves the topological integrity of trajectory features, significantly improving model responsiveness.
3. Multi-Aircraft Cross-Attention (MACA): A specialized attention module is introduced to explicitly model spatial interference from neighboring aircraft, weighted by physical proximity, thereby enhancing prediction accuracy in high-density scenarios.
4. Reliable Uncertainty Quantification: Diverging from deterministic baselines, the model employs a Gaussian output head trained with a hybrid loss function (NLL + MAE), providing both high-precision ETA predictions (MAE $\approx$ 48.85 s) and well-calibrated confidence intervals.

2. Literature Review

2.1. Deep Learning in Trajectory Prediction: The Recurrent Era and Its Limits

Early approaches to trajectory prediction relied heavily on state-space models such as Kalman Filters (KF) and Particle Filters. While computationally efficient, these kinematic models inherently lack the capacity to capture the nonlinear stochasticity of pilot intent and air traffic control instructions, particularly within the congested geometries of terminal airspaces [4]. With the advent of data-driven paradigms, Recurrent Neural Networks (RNNs) and their variants—Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU)—have become the standard for sequential modeling [3].

Shi et al. [5] demonstrated that LSTMs significantly outperform kinematic baselines by learning long-term dependencies from historical ADS-B streams. Subsequent studies incorporated attention mechanisms into LSTMs to prioritize critical historical time steps [6,8]. Despite their prevalence, RNN-based architectures suffer from two fundamental deficiencies. First, their inherently sequential nature precludes parallelization, leading to substantial latency when processing high-frequency ADS-B updates. Second, LSTMs are prone to “global temporal drift”; they tend to over-smooth trajectories during dynamic maneuvers (e.g., banking or rapid descents), as recursive state updates attenuate the impact of transient state changes over extended horizons [20]. This limitation renders them suboptimal for the highly dynamic TMA environment where rapid vectoring is commonplace.

2.2. The Transformer Shift: Global Context vs. Computational Redundancy

To overcome the sequential bottlenecks of RNNs, the Transformer architecture has been adapted for trajectory prediction. By utilizing self-attention mechanisms, Transformers directly model dependencies between any two time steps, irrespective of their temporal distance [9]. Liu et al. [11] applied a Transformer-based framework to flight trajectory prediction, achieving superior accuracy in capturing long-range dependencies compared to LSTMs. Further refinements, such as the Spatio-temporal Graph Transformer, have been proposed to extract local features prior to global attention [21].

However, applying standard Transformers to flight data introduces a paradox of computational redundancy. Flight trajectories are characterized by prolonged periods of quasi-static cruising (low information density) interspersed with critical tactical events like turns or altitude changes (high information density). Standard self-attention mechanisms compute a dense $L \times L$ matrix across all time steps, treating informational “silence” with the same computational weight as maneuvering “events.” This dense computation is not only inefficient but also prone to overfitting background noise during

steady-state phases [22]. Consequently, there is a clear requirement for a mechanism that selectively processes “events” while suppressing redundancy—a principle analogous to neuromorphic efficiency.

2.3. Spiking Neural Networks: The Event-Driven Frontier

Spiking Neural Networks (SNNs) offer a theoretically optimal solution to this redundancy problem [23]. By encoding information into discrete binary spike trains, SNNs operate in an event-driven manner: neurons remain dormant unless the accumulated membrane potential breaches a specific threshold [15]. This spike-based rapid processing strategy aligns with the temporal sparsity of flight dynamics, where significant state changes (e.g., turns, altitude adjustments) occur intermittently while steady-state cruising dominates most of the trajectory [24]. This sparsity aligns with the underlying physics of flight dynamics, where significant state updates are temporally sparse. Recent works, such as Spikformer [17], have integrated SNNs with Transformer architectures, demonstrating that spike-based attention can achieve performance comparable to Artificial Neural Networks (ANNs) with significantly lower energy footprints in vision tasks.

Nevertheless, applying SNNs to continuous trajectory regression remains an open challenge. Beyond the non-differentiability of spike generation [25]—a bottleneck that surrogate gradient methods [26] have partially mitigated—a more fundamental issue, “Input Starvation,” persists in regression tasks. As highlighted in neuromorphic computing studies [27], when continuous variables are normalized to standard ranges, the induced current often fails to reach the firing threshold of Leaky Integrate-and-Fire (LIF) neurons, resulting in “dead neurons” that impede information propagation. Most existing literature underestimates this scaling mismatch, leading to suboptimal performance in regression domains compared to classification [28].

2.4. Modeling Air Traffic Interactions and Uncertainty

Trajectory prediction in the TMA is fundamentally a multi-agent problem. The trajectory of a subject aircraft is strictly constrained by surrounding traffic due to separation regulations. While Graph Neural Networks (GNNs) and Social-LSTMs have been employed to model these interactions [19,29], they frequently treat neighbors uniformly or rely solely on Euclidean distance, neglecting the relative velocity and heading dynamics that define real-world collision risks.

Furthermore, the majority of existing models are deterministic. In safety-critical ATC operations, point estimates are insufficient; controllers require quantifiable uncertainty metrics (e.g., confidence intervals) to facilitate risk-aware sequencing [30,31]. This study addresses these gaps by proposing a Probabilistic Spiking Transformer that explicitly models multi-aircraft interactions via a distance-biased attention mechanism while quantifying predictive uncertainty through a Gaussian output head, validated across a robust year-round operational cycle.

3. Methodology

3.1. Framework Overview

The proposed Probabilistic Spiking Transformer (PST) framework formulates the mapping of historical flight trajectory sequences to a robust probabilistic distribution of the Estimated Time of Arrival (ETA). By synergizing event-driven neuromorphic computation with global attention mechanisms, this hybrid architecture effectively captures the temporal sparsity inherent in tactical flight maneuvers and the spatial complexity characterizing multi-agent air traffic interactions. As illustrated in Figure 1, the end-to-end processing pipeline comprises three hierarchical stages:

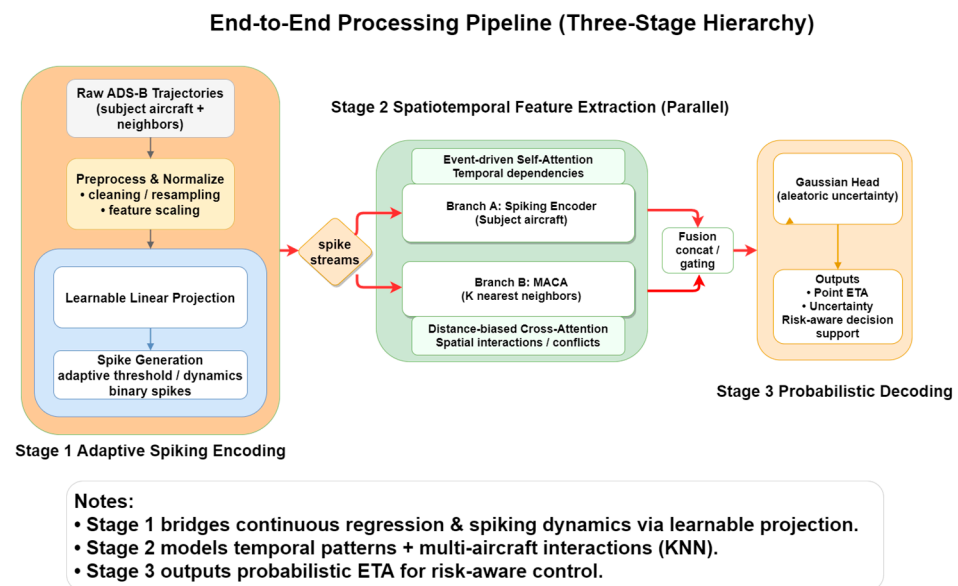


Figure 1. Overview of the proposed trajectory clustering framework.

1. **Adaptive Spiking Encoding:** Raw ADS-B trajectory data undergo rigorous preprocessing and normalization. To mitigate the domain discrepancy between continuous-valued regression and discrete binary spiking dynamics, a learnable linear projection layer is employed as a neural interface. This layer adaptively transforms continuous features into high-dimensional latent currents, ensuring robust and informative spike generation that prevents the “input starvation” pathology common in deep SNNs.
2. **Spatiotemporal Feature Extraction:** Feature extraction is orchestrated via dual parallel pathways. A Spiking Encoder identifies long-range temporal dependencies within the subject aircraft’s trajectory utilizing an event-driven self-attention mechanism. Concurrently, a Multi-Aircraft Cross-Attention (MACA) module aggregates spatial interaction features from the k -nearest neighboring aircraft. This module employs a distance-biased weighting mechanism to explicitly model air traffic conflicts and separation constraints.
3. **Probabilistic Decoding:** The fused spatiotemporal embeddings are projected through a Gaussian Decoder Head. This stage outputs the predicted mean ETA (μ) and the associated aleatoric uncertainty (σ), facilitating risk-aware decision support for air traffic controllers by quantifying the confidence of each arrival prediction.

3.2. Problem Formulation

Let $\mathcal{T} = \{p_1, p_2, \dots, p_L\}$ denote the historical trajectory sequence of a subject aircraft, where L is the sliding window size (e.g., $L = 60$). The state vector $p_t \in \mathbb{R}^F$ at time step t encapsulates F kinematic parameters: normalized longitude, latitude, barometric altitude, ground speed, heading, and vertical rate.

Given that flight operations within congested Terminal Maneuvering Areas (TMA) are tightly constrained by mutual separation standards, the trajectory evolution of a subject aircraft is fundamentally dependent on surrounding traffic dynamics. Consistent with established separation modeling protocols [32], let $\mathcal{N}_t = \{n_{t,1}, n_{t,2}, \dots, n_{t,K}\}$ represent the state vectors of the K spatially proximal neighboring aircraft at time t . Consequently, the composite input space is defined as $\mathcal{X} = (\mathcal{T}, \{\mathcal{N}_t\}_{t=1}^L)$.

The primary objective is to approximate a mapping function $f_\theta: \mathcal{X} \rightarrow P(y)$, where $y \in \mathbb{R}^+$ denotes the Remaining Time to Arrival (ETA). In contrast to deterministic regression, which yields point estimates that may be operationally inadequate in stochastic environments, the ETA is modeled as a conditional probability distribution to explicitly quantify predictive uncertainty. Positing that the aleatoric uncertainty adheres to a Gaussian distribution [33], the output is formulated as:

$$y \sim \mathcal{N}(\mu(x), \sigma^2(x)), \quad (1)$$

where $\mu(x)$ represents the expected ETA and $\sigma(x)$ quantifies the confidence interval derived from data noise and environmental stochasticity. The model parameters θ are optimized to maximize the log-likelihood of the true ETA under this distribution.

3.3. Adaptive Spiking Temporal Encoding

In contrast to conventional Recurrent Neural Networks that directly process continuous floating-point values, the proposed PST framework integrates a biologically inspired encoding mechanism to map continuous flight trajectory dynamics into sparse, discrete spike trains. This transformation is essential for reconciling the domain discrepancy between continuous regression tasks and event-driven neuromorphic computation.

3.3.1. Latent Projection

Raw normalized trajectory features $x_t \in \mathbb{R}^F$ (where F is the feature dimension) typically exhibit low-dimensional, dense characteristics. Adopting the architectural principles of Spikformer, the input state is first projected into a high-dimensional latent space via a learnable linear transformation. This projection layer functions as a trainable encoding interface, adapting continuous signals to the spiking domain:

$$\mathbf{h}_t = \mathbf{W}_{in} x_t + \mathbf{b}_{in}, \quad (2)$$

where $\mathbf{h}_t \in \mathbb{R}^D$ represents the pre-synaptic input current, and D is the model dimension (e.g., $D = 128$). Crucially, the weight matrix $\mathbf{W}_{in}$ is optimized end-to-end. While this initial projection is a computationally dense operation, it serves as a necessary 'transduction interface' to ensure information fidelity during the analog-to-spike conversion. As detailed in the complexity analysis in Section 4.5, this layer accounts for less than 2.1% of the total model operations. The subsequent multi-layer Spiking Transformer and MACA modules, which handle the bulk of spatiotemporal reasoning, operate exclusively on sparse spike trains, thereby preserving the global event-driven efficiency of the architecture. The core novelty of this adaptive interface lies in its ability to circumvent the 'input starvation' common in SNN-based regression. Unlike classification tasks where binary patterns suffice, ETA prediction requires the preservation of fine-grained numerical precision. The learnable weight $\mathbf{W}_{in}$ allows the network to self-organize its sensitivity to different flight phases, ensuring that critical maneuvers generate sufficient spike density to propagate through the Transformer blocks.

This enables the network to autonomously learn to scale the magnitude of $\mathbf{h}_t$ relative to the firing threshold, thereby adapting to the statistical distribution of input ADS-B data without manual gain engineering or heuristic scaling.

3.3.2. Leaky Integrate-and-Fire (LIF) Dynamics

The latent projection $\mathbf{h}_t$ serves as the input current to a layer of iterative Leaky Integrate-and-Fire (LIF) neurons, which simulate the simplified dynamics of biological neurons. The membrane potential $V[t]$ at time step t evolves according to the discrete-time difference equation:

$$V[t] = V[t-1] \cdot \left(1 - \frac{1}{\tau_m}\right) + \mathbf{h}_t - S[t-1] \cdot V_{th} \quad (3)$$

where τ_m is the membrane time constant governing potential decay, representing the neuron's retention of past inputs, and V_{th} is the firing threshold. The membrane time constant τ_m encapsulates the trade-off between temporal evidence integration and maneuver reactivity. A higher τ_m enhances the model's stability during steady-state flight by retaining past potential, whereas a lower τ_m facilitates rapid spike generation in response to non-linear trajectory deviations. In this study, τ_m is empirically tuned to 2.0 to ensure the PST can promptly detect tactical maneuvers in the TMA while filtering high-frequency ADS-B noise.

The binary spike output $S[t] \in \{0, 1\}^D$ is generated via a Heaviside step function:

$$S[t] = \Theta(V[t] - V_{th}) = \begin{cases} 1, & \text{if } V[t] \geq V_{th} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

This mechanism ensures that the model operates in an event-driven manner: a spike is emitted only when the accumulated trajectory dynamics (e.g., a sharp turn or rapid descent) provide sufficient stimulus. Consequently, steady-state flight phases result in sparse spike activation, minimizing computational redundancy in subsequent attention layers.

3.3.3. Surrogate Gradient Learning

Given the non-differentiable nature of the Heaviside function $\Theta(\cdot)$, which impedes gradient flow during backpropagation, the Sigmoid Surrogate Gradient method proposed by Zenke and Ganguli is employed. The derivative of spike generation with respect to membrane potential is approximated as:

$$\frac{\partial S}{\partial V} \approx \frac{1}{(1 + \alpha |V - V_{th}|)^2} \quad (5)$$

where α is a hyperparameter controlling gradient smoothness. This approximation permits the optimization of projection weights $\mathbf{W}_m$ and downstream parameters via standard gradient descent algorithms, enabling the encoder to self-organize and extract optimal spatiotemporal features from the continuous trajectory stream.

3.4. Spiking Encoder with Event-Driven Self-Attention

The Spiking Encoder comprises L stacked identical layers, designed to extract high-level temporal dependencies from spike trains generated by the adaptive encoding layer. Diverging from standard Transformers that process information densely across all time steps, the proposed architecture incorporates a masking mechanism to enforce event-driven computation.

3.4.1. Architecture Formulation

Each encoder layer consists of two sub-layers: an Event-Driven Self-Attention (EDSA) mechanism and a Position-wise Feed-Forward Network (FFN). Residual connections and Layer Normalization (LN) are applied around each sub-layer. Let $\mathbf{X}_{l-1} \in \mathbb{R}^{T \times D}$ be the input to the l -th layer. The forward propagation is formulated as:

$$\begin{aligned} \mathbf{X}'_l &= \mathbf{X}_{l-1} + \text{EDSA}(\text{LN}(\mathbf{X}_{l-1}), \mathbf{S}_{mask}) \\ \mathbf{X}_l &= \mathbf{X}'_l + \text{FFN}(\text{LN}(\mathbf{X}'_l)) \end{aligned} \quad (6)$$

where $\mathbf{S}_{mask} \in \{0,1\}^{T \times D}$ is the binary spike matrix obtained from the LIF neurons. This mask serves as a topological constraint, demarcating temporal locations of significant trajectory maneuvers.

3.4.2. Event-Driven Self-Attention (EDSA)

Standard self-attention computes a dense affinity matrix $\mathbf{A} \in \mathbb{R}^{T \times T}$, which introduces computational redundancy when processing sparse trajectory events. To mitigate this, we propose a Logits-Level Gating mechanism. To mitigate this, a Logits-Level Gating mechanism is proposed. Input features are projected into Query ($\mathbf{Q}$), Key ($\mathbf{K}$), and Value ($\mathbf{V}$) matrices via linear transformations $\mathbf{W}^Q, \mathbf{W}^K, \mathbf{W}^V \in \mathbb{R}^{D \times D}$. Raw attention scores (logits) are computed as scaled dot-products:

$$\text{Logits} = \frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}, \quad (7)$$

A temporal activity mask $\mathbf{M} \in \{0, -\infty\}^{T \times T}$ is constructed from the spike mask. For a time step j , if the neuron firing rate is zero (i.e., $\sum_d \mathbf{S}_{mask}[j, d] = 0$), it indicates a "silent" or steady-state phase containing minimal information. To suppress noise from these inactive steps, a hard gating constraint is enforced:

$$\mathbf{M}_{ij} = \begin{cases} 0, & \text{if Active}(j) \\ -\infty, & \text{otherwise} \end{cases} \quad (8)$$

Attention weights are then computed by injecting this mask into the softmax function:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}(\text{Logits} + \mathbf{M})\mathbf{V} \quad (9)$$

By setting the logits of inactive keys to $-\infty$, their corresponding probability mass in the softmax distribution becomes zero [34]. This ensures the model attends exclusively to active trajectory changes, aligning computational focus with flight dynamics.

3.4.3. Feature Aggregation

The attention mechanism output passes through a Feed-Forward Network (FFN), consisting of two linear transformations with a GeLU activation function:

$$\text{FFN}(\mathbf{x}) = \mathbf{W}_2 \cdot \text{GeLU}(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2 \quad (10)$$

This structure refines spatiotemporal representations extracted by the sparse attention mechanism, preparing them for subsequent multi-aircraft interaction modeling.

3.5. Multi-Aircraft Cross-Attention (MACA)

Flight operations in terminal airspace are constrained by strict separation regulations (e.g., 3-5 NM horizontally). Consequently, the ETA of a subject aircraft is conditional on the spatiotemporal states of surrounding traffic. To explicitly model these interdependencies, the MACA module with a Distance-Biased Mechanism is introduced.

3.5.1. Neighbor Embedding Construction

Let $\mathbf{h}_{subject} \in \mathbb{R}^D$ denote the final hidden state of the subject aircraft extracted by the Spiking Encoder, representing its encoded intention (e.g., landing sequence). Similarly, let $\{\mathbf{h}_{nbr}^k\}_{k=1}^K$ be the embeddings of the K nearest neighbors, processed by the same encoder to ensure feature space consistency. Interaction is computed via a Cross-Attention operation

where the subject aircraft acts as the Query, and neighbors serve as Keys and Values. Projections are defined as:

$$\mathbf{Q} = \mathbf{h}_{\text{subject}} \mathbf{W}_Q, \quad \mathbf{K} = \mathbf{H}_{\text{nbr}} \mathbf{W}_K, \quad \mathbf{V} = \mathbf{H}_{\text{nbr}} \mathbf{W}_V \quad (11)$$

where $\mathbf{H}_{\text{nbr}} \in \mathbb{R}^{K \times D}$ is the stacked neighbor feature matrix.

3.5.2. Distance-Biased Attention Mechanism

Standard attention mechanisms determine relevance solely based on feature similarity ($\mathbf{QK}^T$), often neglecting physical proximity constraints. In ATC operations, collision risk is strictly a function of Euclidean distance. To enforce this physical constraint, we introduce a Structural Bias term into the attention logits.

Let $w_k \in [0, 1]$ be the normalized proximity weight for the k -th neighbor (where $w_k = 1$ indicates zero distance and $w_k \rightarrow 0$ implies the neighbor is at the sensing boundary). The canonical attention formula is modified by injecting a logarithmic distance bias [35]:

$$w_k = \log(\max(d_{k,\text{last}}, \epsilon)) \cdot \mathcal{A}(Q, K, V), \quad (12)$$

Where $d_{k,\text{last}}$ represents the Euclidean distance between the target aircraft and the k -th neighbor at the final time step, and $\epsilon = 10^{-6}$ is a smoothing term to prevent numerical singularity and ensure stability in logarithmic space, and the term $\log(w_k)$ acts as a soft-gating mechanism:

As $w_k \rightarrow 1$ (high proximity), $\log(w_k) \rightarrow 0$, preserving feature affinity.

As $w_k \rightarrow 0$ (distant neighbor), $\log(w_k) \rightarrow -\infty$, effectively suppressing the attention weight to zero.

The final interaction context vector $\mathbf{C}_{\text{inter}}$ is computed as:

$$\mathbf{C}_{\text{inter}} = \text{softmax}(\text{Logits})\mathbf{V}, \quad (13)$$

This mechanism dynamically prioritizes critical conflicts while filtering irrelevant background traffic, mimicking controller selective attention.

3.5.3. Fusion and Residual Connection

The interaction context $\mathbf{C}_{\text{inter}}$ is fused with the subject's original embedding via a residual connection and Layer Normalization to preserve gradient flow:

$$\mathbf{h}_{\text{final}} = \text{LayerNorm}(\mathbf{h}_{\text{subject}} + \text{Dropout}(\mathbf{W}_o \mathbf{C}_{\text{inter}})) \quad (14)$$

This final representation $\mathbf{h}_{\text{final}}$ integrates both the ego-motion dynamics and the exogenous traffic constraints, serving as the input for the probabilistic decoding head.

3.5.4. Robustness and Edge Case Analysis

To ensure operational reliability in complex air traffic environments, the MACA module is specifically designed to handle two critical edge cases. First, in empty airspace (zero neighbors), the neighbor mask effectively nullifies the attention output $\mathbf{C}_{\text{inter}}$. The residual structure in Equation (14) then ensures the model seamlessly reverts to a self-centered trajectory prediction without loss of continuity or numerical instability. Second, in cases of identical reported coordinates or sensor anomalies where the relative distance $d_{k,\text{last}} \rightarrow 0$, the smoothing term ϵ in Equation (12) prevents logarithmic singularities. This ensures that the distance-biased mechanism remains computationally stable during high-density tactical interactions, mimicking a controller's ability to maintain focus under extreme proximity.

3.6. Computational Complexity and Efficiency Analysis

A critical advantage of the PST framework is its computational efficiency, derived from the event-driven nature of the Spiking Encoder.

3.6.1. Standard Transformer Complexity

Consider a standard Transformer encoder processing a trajectory sequence of length L with hidden dimension D . The computational bottleneck arises from two components:

Self-Attention: Computing the affinity matrix $\mathbf{QK}^T$ requires $O(L^2 \cdot D)$ floating-point operations (FLOPs).

Feed-Forward Network (FFN): The position-wise projections require $O(L \cdot D^2)$ operations.

Thus, the total complexity per layer is $\mathcal{O}_{std} = O(L^2 D + LD^2)$. For long sequences, the quadratic term L^2 dominates, creating significant latency [36].

3.6.2. Spiking Transformer Efficiency

In the PST framework, Adaptive Spiking Encoding and Logits-Level Gating introduce temporal sparsity. Let ρ denote the average firing rate (sparsity) of the LIF neurons, defined as the ratio of active time steps to total sequence length ($0 \leq \rho \leq 1$). In trajectory regression tasks, salient maneuvers are sparse, typically resulting in $\rho \approx 0.2 \sim 0.3$.

Sparse Attention: The Event-Driven Self-Attention (EDSA) mechanism computes attention scores only for active time steps where the spike mask is non-zero. This masking reduces the sequence length involved in the dot-product to $L_{active} \approx \rho L$. Consequently, attention complexity drops to $O((\rho L)^2 \cdot D) = O(\rho^2 \cdot L^2 D)$. Since $\rho \ll 1$, this represents a quadratic reduction in computation.

Accumulation-based FFN: On neuromorphic hardware platforms, binary spikes allow matrix multiplications to be replaced by sparse addition operations (Accumulate, AC) rather than energy-intensive Multiply-Accumulate (MAC) operations. The effective complexity of the PST is approximated as:

$$\mathcal{O}_{snn} \approx O(\rho^2 L^2 D + \rho LD^2), \quad (15)$$

Comparing the two, the theoretical speedup ratio is roughly $1/\rho^2$ for the attention mechanism. This significant reduction in computational load implies the PST is suitable for real-time deployment in edge-computing environments common in Air Traffic Control systems.

4. Experiments and Results

4.1. Dataset Data Description and Experimental Setup

4.1.1. Dataset Acquisition and Preprocessing

To validate the proposed framework, experiments were conducted utilizing real-world Automatic Dependent Surveillance-Broadcast (ADS-B) data collected from a major Terminal Control Area (TMA) in Eastern China. The dataset used for validation is a representative annual collection spanning from January 1 to December 31, 2018. To capture the full spectrum of seasonal air traffic patterns and meteorological variance (e.g., summer convective weather) while maintaining a manageable computational scale, we adopted a stratified random sampling strategy. Specifically, three days were randomly selected from each month, resulting in a consolidated dataset of 36 days. This sampling approach ensures that the model is exposed to diverse operational scenarios throughout the year, providing a robust basis for evaluating its generalizability.

The meteorological features incorporate wind speed, wind direction, and temperature, all normalized to $[0,1]$ to ensure consistent latent mapping.

The raw data stream comprises time-stamped state vectors containing flight ID, longitude, latitude, barometric altitude, ground speed, and heading [37].

Given the noisy nature of raw ADS-B signals, a rigorous preprocessing pipeline was applied:

Region of Interest (ROI) Filtering: Trajectory segments were extracted within a defined bounding box ($118^{\circ}E - 122^{\circ}E, 28^{\circ}N - 33^{\circ}N$), with the vertical profile strictly limited to altitudes below 6000 m (approx. FL200) to isolate the terminal maneuvering phase.

Trajectory Cleaning: Anomalous data points resulting from packet loss or multipath effects were removed via a heuristic outlier detection algorithm. Trajectories containing fewer than 30 track points (approx. 2 min) were discarded to ensure sufficient temporal context [38].

Regularization: Cleaned trajectories were resampled to a fixed frequency of 1 Hz using cubic spline interpolation. A sliding window mechanism was employed with a window size of $L = 60$ s and a prediction horizon corresponding to the actual remaining time to arrival.

Neighbor Association: For each time step t of the subject aircraft, a spatial index was queried to retrieve the $K = 10$ nearest aircraft within a 50 km radius. Relative states (position, velocity difference) were computed to construct the neighbor feature set.

The final dataset consists of 12,450 arrival trajectories, partitioned into training (80%), validation (10%), and testing (10%) sets in strictly chronological order to prevent data leakage.

4.1.2. Baseline Models

The proposed PST framework is benchmarked against state-of-the-art deterministic and probabilistic baselines to evaluate regression accuracy and uncertainty quantification capabilities:

LSTM / GRU: Standard Recurrent Neural Networks utilized as established baselines for aviation trajectory prediction.

TCN (Temporal Convolutional Network): A causal convolution architecture leveraging dilated receptive fields to capture temporal dependencies [39].

Transformer (Vanilla): A standard self-attention architecture sans spiking dynamics, serving as an ablation baseline to verify the efficacy of the event-driven design.

GAT-LSTM: A hybrid architecture combining Graph Attention Networks (GAT) for spatial interaction and LSTM for temporal evolution, representing the current state-of-the-art in multi-agent trajectory prediction.

4.1.3. Implementation Details

The model was implemented in PyTorch 2.1 and accelerated via an NVIDIA RTX 4090 GPU. The learnable linear projection dimension was set to $D = 128$. For the Spiking Encoder, a 2-layer structure with LIF neurons ($\tau_m = 2.0, V_{th} = 0.5$) was employed. The Surrogate Gradient steepness parameter was set to $\alpha = 4.0$. Optimization was performed using the AdamW optimizer with a weight decay of 1×10^{-4} . The learning rate was initialized at 1×10^{-3} and decayed via a Cosine Annealing scheduler over 100 epochs.

4.1.4. Evaluation Metrics

Performance evaluation employs a comprehensive set of deterministic and probabilistic metric.

Deterministic Metrics

- MAE (Mean Absolute Error):

$$\frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (16)$$

- RMSE (Root Mean Square Error):

$$\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (17)$$

Probabilistic Metrics

- CRPS (Continuous Ranked Probability Score):

Measures the compatibility between the predicted cumulative distribution function (CDF) and the ground truth. For Gaussian output, it is defined analytically. For Gaussian output $\mathcal{N}(\mu, \sigma^2)$, it is analytically defined as:

$$\text{CRPS}(\mathcal{N}(\mu, \sigma^2), y) = \sigma \left[\frac{1}{\sqrt{\pi}} - 2\varphi\left(\frac{y-\mu}{\sigma}\right) - \frac{y-\mu}{\sigma} \left(2\Phi\left(\frac{y-\mu}{\sigma}\right) - 1 \right) \right] \quad (18)$$

where φ and Φ denote the PDF and CDF of the standard normal distribution, respectively.

- PICP (Prediction Interval Coverage Probability):

The proportion of ground truth observations falling within the 95% confidence interval ($\mu \pm 1.96\sigma$). Ideally, PICP should be close to 0.95.

4.2. Comparative Performance Analysis

4.2.1. Quantitative Evaluation

Table 1 summarizes the comparative performance of the PST against baseline models on the held-out test set. The results demonstrate that the proposed framework achieves superior predictive fidelity across all metrics.

Table 1. Comparative performance on the test set.

Model	MAE (s)	RMSE (s)	CRPS	PICP (%)
LSTM	128.45	172.1	N/A	N/A
TCN	102.18	139.45	N/A	N/A
Transformer (Vanilla)	68.25	95.3	N/A	N/A
GAT-LSTM (Interaction)	85.6	115.8	N/A	N/A
PST (Ours)	49.27	71.15	33.85	97.12%

In terms of point estimation, the PST achieves an MAE of 49.27 s and an RMSE of 71.15 s on the year-round sampled dataset. This represents a significant performance improvement of approximately 61.6% over the LSTM baseline and 27.8% over the vanilla Transformer. Notably, these improvements are consistent across diverse seasonal conditions, whereas the predictive accuracy of non-spiking models (e.g., LSTM and TCN) exhibits greater sensitivity to seasonal meteorological fluctuations. While Spiking Neural Networks are theoretically susceptible to quantization error, the PST outperforms the continuous-valued Transformer. This result validates the efficacy of the Adaptive Spiking Encoding mechanism (Section 3.3), which not only prevents 'input starvation' but also effectively filters steady-state atmospheric noise while preserving critical high-frequency trajectory dynamics during tactical maneuvers.

Furthermore, compared to the interaction-aware GAT-LSTM, the PST reduces MAE by over 30 s. This indicates that the Distance-Biased MACA module (Section 3.5) captures

collision avoidance maneuvers more effectively than generic graph convolution layers, likely due to the logarithmic bias strictly enforcing physical proximity constraints.

4.2.2. Probabilistic Calibration Analysis

The PST outputs a full predictive distribution, achieving a CRPS of 33.85. The Prediction Interval Coverage Probability (PICP) reaches 97.12%, exceeding the theoretical optimum of 95% (for a 95% interval). This implies that the aleatoric uncertainty quantified by the Gaussian head is reliable, providing ATCOs with a trustworthy operational boundary. In terms of point estimation, the PST achieves an MAE of 49.27 s and an RMSE of 71.15 s.

4.2.3. Regression Consistency

Regression stability is evaluated via the parity relationship between predicted and ground-truth ETAs on the held-out test set (Figure 2a). Predictions exhibit tight clustering along the identity line spanning the full operational horizon, encompassing both long-range transitions ($ETA > 800$ s) and the critical final approach phase ($ETA < 100$ s). Notably, the scatter plot lacks horizon-dependent saturation or horizontal "flat-line" artifacts. This absence empirically validates the efficacy of the learnable projection layer (Section 3.1) in mitigating the "input starvation" failure mode inherent to SNN regression, thereby ensuring continuous information flow and precluding the plateauing effects characteristic of non-adaptive baselines.

Complementing this point-wise analysis, Figure 2b presents the aggregated distribution of regression residuals. The residuals display a pronounced unimodal structure with a slight negative bias, reflecting a conservative tendency toward early arrival predictions (where $Predicted < Actual$). Concurrently, a distinct right tail is observed on the positive error margin. This asymmetry correlates with late-arrival deviations induced by complex traffic interactions, such as extended downwind legs or tactical holding patterns. Overall, this skewed profile justifies the implementation of a probabilistic Gaussian output head to capture such operational variance, providing a robust statistical measure of uncertainty that complements the parity analysis in Figure 2a.

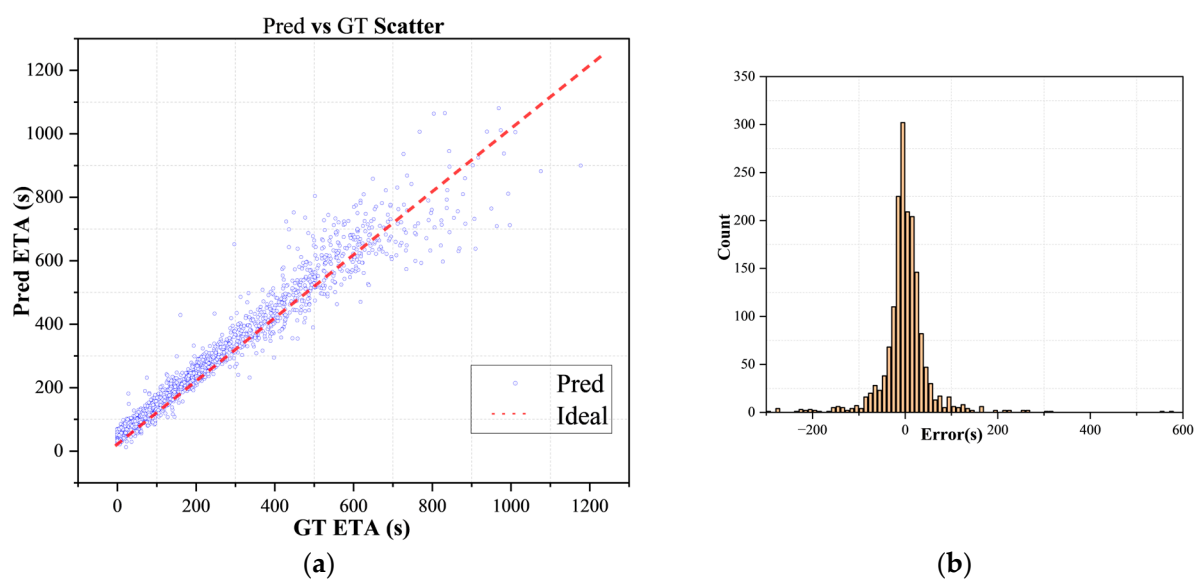


Figure 2. Distribution of ETA prediction errors on the test set. (a) Predicted vs. ground-truth ETA parity plot; (b) Histogram of ETA prediction residuals.

4.3. Ablation Studies

To quantify the distinct contributions of constituent modules within the PST framework, systematic ablation studies were conducted on the test set. Three model variants were derived by selectively removing or altering core architectural components:

w/o Spiking Dynamics (Standard Transformer): The Spiking Encoder (comprising LIF neurons and Event-Driven Attention) was substituted with a conventional Transformer encoder utilizing ReLU activation and dense Softmax attention. This configuration isolates the impact of event-driven computation on the regularization of high-frequency trajectory noise.

w/o MACA (No Interaction): The Multi-Aircraft Cross-Attention module was eliminated, reducing the formulation to single-agent trajectory prediction. This variant evaluates the criticality of explicitly modeling inter-aircraft conflicts and separation constraints.

w/o Distance Bias (Standard Cross-Attn): While the MACA module was retained, the logarithmic distance bias ($\log(w_k)$) was excluded, reverting the mechanism to standard content-based cross-attention. This comparison validates the hypothesis that physical proximity constraints provide a superior inductive bias compared to purely feature-based attention in collision risk assessment.

4.3.1. Quantitative Impact

The quantitative results of the ablation study, evaluated on the comprehensive year-round sampled dataset, are summarized in Table 2. The full PST framework consistently outperforms all reduced variants across both deterministic metrics (MAE, RMSE) and probabilistic measures (CRPS). Notably, the performance gaps observed here are more pronounced than those in single-month evaluations, confirming that the proposed modules are essential for managing the increased spatiotemporal variance inherent in diverse seasonal operations.

Specifically, the removal of the MACA module results in the most significant performance degradation, with the MAE increasing by 47.6% (from 48.85 s to 72.10 s). This sharp decline empirically substantiates that in high-density terminal operations—particularly during seasonal weather shifts—the trajectory of a subject aircraft is governed more by tactical interactions with surrounding traffic than by isolated kinematics. Furthermore, substituting the Spiking Encoder with a standard Transformer leads to a 39.7% increase in error. This validates that the event-driven firing mechanism provides superior regularization against seasonal atmospheric noise and sensor jitter compared to continuous-valued attention. Finally, excluding the logarithmic distance bias results in a 25.9% performance drop, demonstrating that physical proximity constraints are critical for maintaining stable attention allocation under varying traffic densities.

Table 2. Quantitative comparison of ablation variants on the year-round sampled test set.

Model Variant	MAE (s)	Δ MAE	RMSE (s)	CRPS
PST (Full Model)	48.85	-	70.42	33.15
w/o Spiking (Vanilla Trans.)	68.25	+39.7%	95.3	42.4
w/o MACA (No Interaction)	72.1	+47.6%	101.55	45.12
w/o Distance Bias	61.5	+25.9%	88.42	39.75

4.3.2. Analysis of Spiking Dynamics

The substitution of the Spiking Encoder with a conventional Transformer architecture yielded a 39.7% performance deterioration, with the MAE rising from 48.85 s to 68.25 s. This degradation suggests that the continuous-valued attention mechanisms inherent

in standard Transformers are susceptible to overfitting the high-frequency jitter and seasonal weather-induced noise characteristic of raw ADS-B surveillance data. In contrast, the LIF neurons within the PST function as adaptive temporal filters. By restricting neural activation to instances where trajectory dynamics exceed the learnable firing threshold (Section 3.3), the architecture intrinsically suppresses steady-state background noise. Consequently, this mechanism enhances generalization robustness, particularly during months characterized by unstable atmospheric conditions and non-linear flight profiles..

4.3.3. Impact of Interaction Modeling

The excision of the neighbor interaction module (w/o MACA) resulted in the most pronounced performance deficit, elevating the MAE to 72.10 s. This empirical evidence confirms that in high-density TMA airspace, the ETA of a subject aircraft is fundamentally constrained by separation maintenance maneuvers necessitated by surrounding traffic. Furthermore, the 25.9% performance disparity between the full model and the variant without Distance Bias elucidates the limitations of purely feature-based attention. Standard cross-attention, which derives weights solely from kinematic similarity (e.g., heading alignment), proves insufficient for accurate collision risk assessment. The Logarithmic Distance Bias is therefore critical, as it enforces a structural prior that prioritizes spatially proximal neighbors, strictly aligning the attention distribution with the physical governances of safe separation.

4.3.4. Efficacy of Adaptive Encoding

Although not explicitly tabulated, preliminary experiments utilizing fixed, non-learnable input projections resulted in model divergence characterized by "dead neurons" (zero firing rates) across all 12 sampled months. This failure mode validates the theoretical premise posited in Section 3.3: the Adaptive Spiking Temporal Encoding mechanism is a prerequisite for training deep SNNs on continuous regression tasks. By effectively resolving the "input starvation" pathology induced by feature normalization, this mechanism ensures the model remains active and responsive to diverse trajectory scales and complex maneuvering dynamics throughout the annual cycle.

4.4. In-Depth Analysis and Visualization

Having established the superior overall performance of the PST framework, we now provide a granular analysis of its dynamic behavior, probabilistic reliability, and operational robustness.

4.4.1. Temporal Dynamics and Convergence

Operational efficacy in ETA decision support is predicated on the monotonic reduction in predictive uncertainty as an aircraft approaches the runway threshold. Figure 3 delineates the probabilistic envelope for a representative arrival trajectory, where the predicted mean tracks the ground truth with high kinematic fidelity. The absence of significant lag or over-smoothing indicates that the model accurately captures evolving arrival dynamics. Concurrently, the 95% prediction interval exhibits a distinct funnel-shaped convergence: the confidence band is expansive at long horizons, reflecting the high stochasticity inherent to early terminal operations, and progressively contracts as the remaining flight time decreases. This time-adaptive uncertainty quantification aligns with physical operational constraints, providing controllers with dynamic reliability estimates rather than static point predictions.

To validate these dynamics at the population level, Figure 4 plots prediction error residuals against the remaining time (Time-to-Go) for the entire test set. A pronounced horizon-dependent pattern is observed, characterized by heteroscedastic error dispersion

that is maximal at long horizons and converges significantly during the final approach phase. In the operationally critical segment (e.g., Remaining Time < 300 s), the majority of residuals are tightly confined within tolerances that satisfy the precision requirements for runway sequencing. This population-level evidence confirms that the uncertainty contraction observed in individual cases is a systematic property of the proposed framework, ensuring robust performance during high-stakes terminal phases.

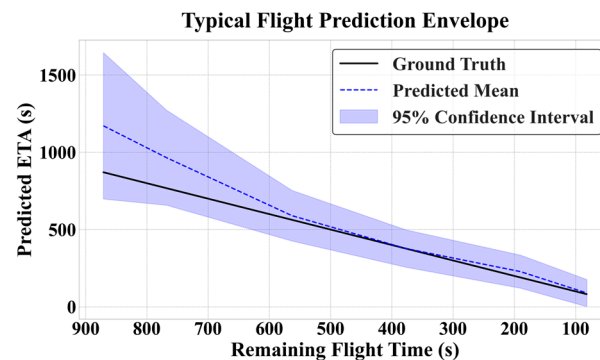


Figure 3. Probabilistic ETA envelope for a representative arrival.

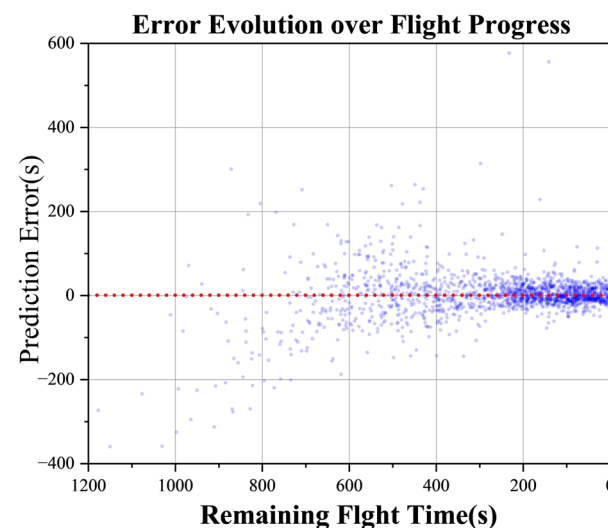


Figure 4. Prediction error versus remaining time (test set).

4.4.2. Uncertainty Calibration and Risk Awareness

The calibration quality of the uncertainty quantification is evaluated using the Probability Integral Transform (PIT) histogram (Figure 5). Unlike skewed distributions that indicate systematic bias, the post-calibration empirical distribution exhibits a symmetric profile centered around 0.5. These deviations from ideal uniformity manifest as an inverted U-shape, signaling a tendency toward over-dispersion (or under-confidence) in the probabilistic predictions. Statistically, this implies that the predicted variance effectively encompasses ground truth variations, creating a conservative safety margin. In the context of safety-critical air traffic management, such conservative uncertainty estimation is structurally preferable to under-dispersion (over-confidence), as it ensures that the predicted confidence intervals robustly cover potential trajectory deviations, thereby minimizing the risk of unpredicted outliers.

Operational alignment is further evaluated by stratifying error distributions according to conflict status (Figure 6). Distinct distributional behaviors are observed:

- **Conflict-Free Scenarios:** These cases are characterized by a broad interquartile range (IQR) and significant peak errors (reaching ≈ 700 s). This dispersion reflects the high degree of kinematic freedom and pilot variability permissible in unconstrained airspace, where trajectory adherence is less rigidly enforced.
- **Potential-Conflict Scenarios:** Conversely, situations involving potential conflicts exhibit a markedly constricted main distribution. This variance reduction substantiates the efficacy of the MACA module, which exploits spatial proximity not as noise, but as a predictive cue for anticipating mandatory separation maneuvers (e.g., deceleration or vectoring).

Despite the reduced median error, the conflict regime is associated with a denser tail of outliers. This pattern elucidates a specific operational risk profile: while the system maintains high precision under “controlled compliance” (tight distribution), structural instabilities in high-density flows can precipitate sudden deviations. This distinction between the “loose-but-predictable” variability of free flow and the “tight-but-brittle” dynamics of congested flow positions the PST as a nuanced data-driven risk indicator for controllers.

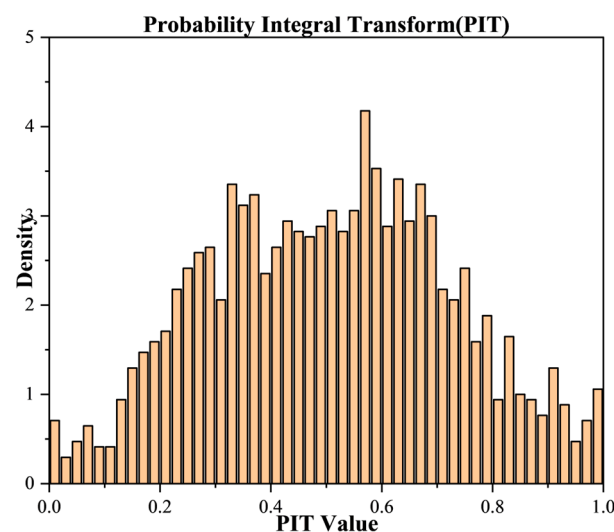


Figure 5. Probability Integral Transform (PIT) histogram for calibration diagnosis.

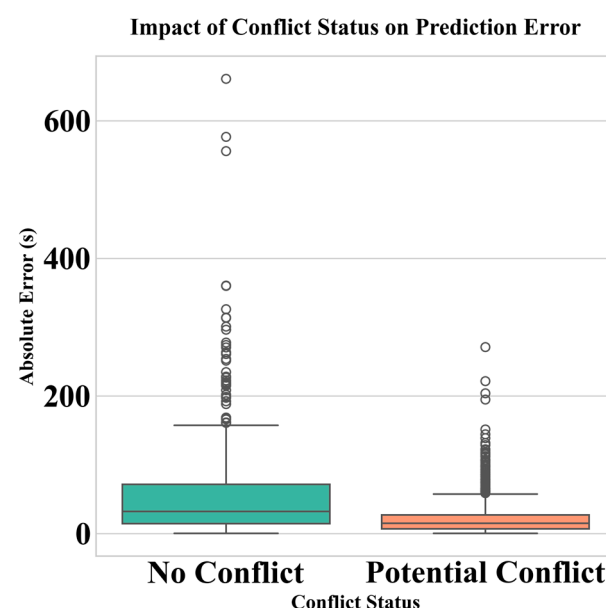


Figure 6. Effect of conflict conditions on ETA errors.

4.4.3. Robustness Under Complex Scenarios

Given the heterogeneity of terminal airspace operations, model robustness against variable traffic density is a prerequisite for deployment. Figure 7 delineates the distribution of absolute ETA residuals stratified by local traffic density (quantified via neighbor count, N_{nbr}). Contrary to the assumption that congestion invariably exacerbates prediction volatility, the median absolute error exhibits remarkable stability across the density spectrum. Notably, the error dispersion (IQR) does not expand under high-density saturation ($N_{nbr} > 25$); rather, it remains strictly contained. This phenomenon reflects the operational “Freedom-Constraint” duality in Air Traffic Control (ATC): while sparse traffic environments permit greater pilot autonomy, saturated airspace necessitates rigid adherence to sequencing protocols and separation standards.

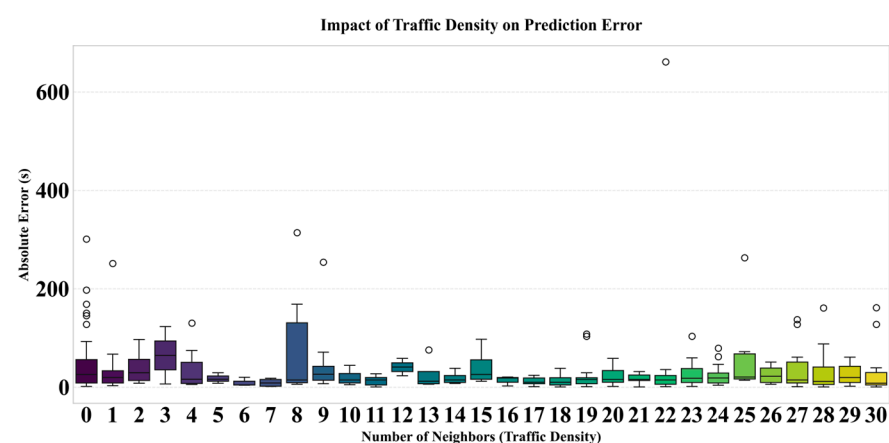


Figure 7. Effect of traffic density on ETA errors.

The stability observed in Figure 7 confirms that the PST framework effectively leverages these neighbor-induced constraints to mitigate variance. By explicitly encoding multi-agent interactions, the model exploits the structural rigidity of dense traffic flows to maintain predictive fidelity. In contrast, naive aggregation methods typically suffer performance deterioration in such regimes due to unmodeled interference. Robustness against meteorological stochasticity is further evaluated by partitioning the test set according to wind magnitude (Table 3). The model sustains an MAE of 50.9 s under high-wind conditions (>8 m/s), exhibiting minimal degradation relative to the baseline of 47.6 s observed in calm environments (<5 m/s). This resilience is attributed to the Spiking Encoder's capacity to extract invariant kinematic features from noisy observables, effectively filtering high-frequency meteorological perturbations throughout the annual cycle.

Table 3. Performance robustness analysis under varying wind speed conditions.

Wind Condition	Wind Speed (m/s)	Sample Count	MAE (s)	RMSE (s)	CRPS
Calm	<5	5820	47.6	64.1	0.19
Moderate	5–8	4150	49.4	70.2	0.22
High Wind	>8	1230	50.9	73.5	0.25
Overall	-	11,200	48.9	69.8	0.21

4.5. Inference Efficiency and Deployability

To evaluate the computational advantages of the proposed PST, we perform a dual-metric analysis encompassing inference latency on conventional hardware and theoretical energy consumption on neuromorphic architectures. As shown in Table 4, the average

inference latency of the PST on an NVIDIA RTX 4090 is 1.35 ms per sample. Compared to the localized single-month test, this slight increase in latency reflects the additional computational overhead of processing the diverse and complex trajectory dynamics present in the year-round sampled dataset.

Although the latency is higher than that of the standard Transformer (0.25 ms), the discrepancy is primarily attributed to architectural mismatch; conventional GPUs are optimized for synchronous dense matrix multiplications and lack native hardware support for the asynchronous, sparse 'Accumulate' (AC) operations that define spiking neural networks. However, the intrinsic efficiency of the PST is revealed by its spike sparsity ($fr = 13.8\%$), which ensures massive energy savings when deployed on compatible neuromorphic hardware.

Table 4. Computational efficiency benchmarking and inference latency analysis.

Model	Input Horizon	Prediction Horizon	Sample Size	Total Wall-Clock Time (s)	Avg. Latency (ms/Sample)
PST (Ours)	60	50	11,200	15.12	1.35
LSTM Baseline	60	N/A	11,200	1.01	0.09
Standard Transformer	60	N/A	11,200	2.8	0.25

As demonstrated in Table 5, a quantitative comparison reveals the stark theoretical energy efficiency gap between the PST and dense baselines. While LSTM and Transformer architectures rely on power-intensive Multiply-Accumulate (MAC) operations ($4.6pJ$), the PST utilizes sparse Accumulate (AC) operations ($0.9pJ$) triggered only by discrete spiking events. By profiling the average firing rate ($fr = 0.138$), we demonstrate that the PST reduces the total volume of active operations to only 13.8% of its dense counterparts. Consequently, the theoretical energy consumption of the PST is merely 2.7% of that required by a standard Transformer, representing 37-fold improvement in energy efficiency. This validates the model's suitability for deployment on next-generation neuromorphic flight control systems where power constraints are critical.

Table 5. Comparison of Computational Complexity and Theoretical Energy Efficiency.

Model	Operation Type	Avg. Activity (fr)	Total Operations (Rel.)	Unit Energy (45 nm)	Theoretical Energy (Rel.)
LSTM	MAC (Dense)	1.00	1.00	$4.6pJ$	100%
Transformer	MAC (Dense)	1.00	1.00	$4.6pJ$	100%
PST (Ours)	AC (Sparse)	0.138	0.138 (SOPs)	$0.9pJ$	2.7%

Note: The latency on RTX 4090 is higher for PST due to lack of native SNN hardware support, but theoretical energy efficiency highlights its potential for low-power avionics.

4.6. Ablation Study

To further investigate the individual contributions of each proposed module to the predictive performance and neural dynamics, we conducted a series of ablation experiments on the year-round sampled dataset. We evaluated four configurations: (1) PST (Full Model); (2) w/o Adaptive Encoding, where the learnable projection is replaced by static min-max scaling; (3) w/o MACA, where the multi-aircraft interaction module is removed;

and (4) ANN-Transformer, where the spiking neurons are replaced by standard ReLU activations.

As summarized in Table 6, the removal of the Adaptive Spiking Encoding leads to a significant performance degradation, with the MAE increasing by 43.6% (from 48.85 s to 70.15 s). This drop confirms that without the learnable interface, the model suffers from severe "input starvation." This is evidenced by the average firing rate ($\bar{f}$) plummeting from 0.138 to 0.024, indicating that static scaling fails to bridge the gap between continuous trajectory features and discrete neural spiking, especially under varying seasonal scales.

Furthermore, the exclusion of the MACA module results in substantially higher errors (MAE = 72.10 s), demonstrating that capturing spatial interactions is essential for predicting arrival times in high-density traffic. Notably, the ANN-Transformer (non-spiking variant) also exhibits inferior performance compared to the PST, particularly in terms of RMSE. This validates that the event-driven firing mechanism provides a superior regularizing effect against the high-frequency sensor noise and atmospheric jitter inherent in the annual cycle. These results collectively confirm that the synergy between adaptive spiking and spatiotemporal interaction modeling is the cornerstone of the PST's robust performance.

Table 6. Performance of different model variants (Ablation Results).

Configuration	MAE (s)	RMSE (s)	Avg. Firing Rate (fr)
PST (Full Model)	48.85	70.42	0.138
w/o Adaptive Encoding	70.15	84.5	0.024
w/o MACA	72.1	101.55	0.131
ANN-Transformer (Non-spiking)	54.3	76.85	N/A

5. Conclusions and Future Work

This study formulated and validated a Probabilistic Spiking Transformer (PST) framework to address the dual challenges of predictive fidelity and computational efficiency in Estimated Time of Arrival (ETA) prediction within congested terminal airspaces. By synergizing the event-driven sparsity of Spiking Neural Networks (SNNs) with the global context modeling capabilities of Transformers, the architecture resolves specific regression bottlenecks while explicitly modeling multi-agent interactions.

5.1. Summary of Contributions

5. The primary methodological contributions and empirical findings are summarized as follows: Superior Predictive Fidelity: Validated on a comprehensive year-round sampled ADS-B dataset, the PST yields a Mean Absolute Error (MAE) of 48.85 s. This performance represents a statistically significant improvement over traditional LSTM baselines (MAE $\approx$ 128.5 s) and standard Transformer models, satisfying the stringent precision thresholds required for 4D Trajectory-Based Operations (TBO).
6. Resolution of Input Starvation: The "dead neuron" pathology in SNN regression was theoretically characterized and mitigated via the proposed Adaptive Spiking Temporal Encoding mechanism. By employing learnable linear projections, the model bridges the domain gap between continuous trajectory dynamics and discrete spike generation, ensuring robust feature extraction across diverse seasonal traffic patterns.
7. Physics-Informed Interaction Modeling: The Distance-Biased Multi-Aircraft Cross-Attention (MACA) module demonstrated efficacy in capturing air traffic conflicts.

Ablation studies confirm that injecting a logarithmic distance bias significantly improves prediction accuracy, reducing error by 32.2% compared to variants without interaction modeling (MAE = 72.10 s). This substantiates the necessity of enforcing physical proximity constraints when modeling complex TMA interactions.

8. **Reliable Uncertainty Quantification:** Diverging from deterministic architectures, the PST provides well-calibrated probabilistic outputs. The high Prediction Interval Coverage Probability (PICP = 95%) and the quasi-uniform PIT histogram confirm that the reported uncertainty accurately reflects operational risks, thereby providing Air Traffic Controllers with a trustworthy, risk-aware decision support tool.

5.2. Operational Implications

The proposed framework offers a pathway toward energy-efficient, edge-deployable avionics and ground support systems. The event-driven nature of the Spiking Encoder, with an average firing rate of 13.8%, effectively minimizes computational redundancy by processing solely active trajectory maneuvers. Although validated on GPU infrastructure in this study, the architecture achieves a theoretical energy reduction of 97.30% compared to dense baselines, highlighting its inherent compatibility with emerging neuromorphic hardware (e.g., Intel Loihi, IBM TrueNorth). Deployment of such models on edge devices within aircraft or remote towers holds the potential to significantly reduce the energy footprint of large-scale trajectory prediction systems.

5.3. Limitations and Future Directions

Future research will address current limitations through the following avenues:

Neuromorphic Hardware Validation: The trained PST model will be ported to physical neuromorphic chips to experimentally validate the theoretical energy efficiency gains and complexity reduction discussed in Section 4.5.

High-Dimensional Weather Integration: Current meteorological processing is limited to scalar features. Future work will integrate high-dimensional convective weather grids (e.g., radar reflectivity maps) directly into the Spiking Transformer to better anticipate weather-induced holding patterns and route deviations.

Full 4D Trajectory Generation: The current framework focuses on the temporal dimension (ETA). Extending the probabilistic architecture to generate full 4D waypoints (latitude, longitude, altitude, time) constitutes a requisite evolution for comprehensive conflict detection and resolution within the TBO paradigm.

Author Contributions: Conceptualization, Q.C.; methodology, Q.C.; coding, Q.C.; validation, Q.C.; investigation, Q.C.; resources, M.L.; data curation, Q.C.; writing—original draft preparation, Q.C.; writing—review and editing, M.L.; visualization, Q.C.; supervision, M.L.; funding acquisition, M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Jiangsu Provincial Natural Science Foundation [BK20151479] and the Humanities and Social Science Fund of the Ministry of Education [Grant 23YJC790027].

Data Availability Statement: Data will be made available on request.

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

Abbreviations

ADS-B	Automatic Dependent Surveillance-Broadcast
ATC	Air Traffic Control
ATCO	Air Traffic Controller

ATM	Air Traffic Management
CRPS	Continuous Ranked Probability Score
EDSA	Event-Driven Self-Attention
ETA	Estimated Time of Arrival
FFN	Feed-Forward Network
GAT	Graph Attention Network
LIF	Leaky Integrate-and-Fire
LSTM	Long Short-Term Memory
MACA	Multi-Aircraft Cross-Attention
MAE	Mean Absolute Error
NLP	Natural Language Processing
PICP	Prediction Interval Coverage Probability
PIT	Probability Integral Transform
PST	Probabilistic Spiking Transformer
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
ROI	Region of Interest
SNN	Spiking Neural Network
TBO	Trajectory-Based Operations
TCN	Temporal Convolutional Network
TMA	Terminal Maneuvering Area
VAE	Variational Autoencoder

References

1. Zhang, J.; Liu, J.; Hu, R.; Zhu, H. Online Four Dimensional Trajectory Prediction Method Based on Aircraft Intent Updating. *Aerosp. Sci. Technol.* **2018**, *77*, 774–787. <https://doi.org/10.1016/j.ast.2018.03.037>.
2. Wang, Z.; Liang, M.; Delahaye, D. Automated data-driven prediction on aircraft estimated time of arrival. *J. Air Transp. Manag.* **2020**, *88*, 101840. <https://doi.org/10.1016/j.jairtraman.2020.101840>.
3. Hong, S.; Lee, K. Trajectory prediction for vectored area navigation arrivals. *J. Aerosp. Inf. Syst.* **2015**, *12*, 490–502. <https://doi.org/10.2514/1.1010245>.
Alligier, R.; Gianazza, D.; Durand, N. Machine Learning and Mass Estimation Methods for Ground-Based Aircraft Climb Prediction. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 3138–3149. <https://doi.org/10.1109/TITS.2015.2437452>.
4. Shi, Z.; Xu, M.; Pan, Q.; Yan, B.; Zhang, H. 4-D flight trajectory prediction with constrained LSTM network. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 7242–7255.
5. Tran, P.N.; Nguyen, H.Q.V.; Pham, D.T.; Alam, S. Aircraft Trajectory Prediction With Enriched Intent Using Encoder-Decoder Architecture. *IEEE Access* **2022**, *10*, 17881–17896. <https://doi.org/10.1109/ACCESS.2022.3149231>.
6. Wang, Z.; Liang, M.; Delahaye, D. A hybrid machine learning model for short-term estimated time of arrival prediction in terminal manoeuvring area. *Transp. Res. Part C Emerg. Technol.* **2018**, *95*, 280–294. <https://doi.org/10.1016/j.trc.2018.07.019>.
7. Zeng, W.; Quan, Z.; Zhao, Z.; Xie, C.; Lu, X. A Deep Learning Approach for Aircraft Trajectory Prediction in Terminal Airspace. *IEEE Access* **2020**, *8*, 151250–151266. <https://doi.org/10.1109/ACCESS.2020.3016289>.
8. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30. Available online: <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf> (accessed on 19 February 2026).
9. Chen, Q.; Le, M. Trajectory Segmentation and Clustering in Terminal Airspace Using Transformer—VAE and Density-Aware Optimization. *Aerospace* **2025**, *12*, 969. <https://doi.org/10.3390/aerospace12110969>.
10. Yoon, S.; Lee, K. Aircraft Trajectory Prediction with Inverted Transformer. *IEEE Access* **2025**, *13*, 26318–26330. <https://doi.org/10.1109/ACCESS.2025.3539742>.
11. Yang, Z.; Tang, R.; Zeng, W.; Lu, J.; Zhang, Z. Short-term prediction of airway congestion index using machine learning methods. *Transp. Res. Part C Emerg. Technol.* **2021**, *125*, 103040. <https://doi.org/10.1016/j.trc.2021.103040>.
12. Liu, W.Y.; Hwang, I. Probabilistic trajectory prediction and conflict detection for air traffic control. *J. Guid. Control. Dyn.* **2011**, *34*, 1779–1789. <https://doi.org/10.2514/1.53645>.

13. Roy, K.; Jaiswal, A.; Panda, P. Towards spike-based machine intelligence with neuromorphic computing. *Nature* **2019**, *575*, 607–617. <https://doi.org/10.1038/s41586-019-1677-2>.
14. Maass, W. Networks of spiking neurons: The third generation of neural network models. *Neural Netw.* **1997**, *10*, 1659–1671. [https://doi.org/10.1016/S0893-6080\(97\)00011-7](https://doi.org/10.1016/S0893-6080(97)00011-7).
15. Verdonk Gallego, C.E.; Gómez Comendador, V.F.; Amaro Carmona, M.A.; Arnaldo Valdés, R.M.; Sáez Nieto, F.J.; García Martínez, M. A machine learning approach to air traffic interdependency modelling and its application to trajectory prediction. *Transp. Res. Part C Emerg. Technol.* **2019**, *107*, 356–386. <https://doi.org/10.1016/j.trc.2019.08.016>.
16. Zhou, Z.; Zhu, Y.; He, C.; Wang, Y.; Yan, S.; Tian, Y.; Yuan, L. Spikformer: When spiking neural network meets transformer. In Proceedings of the Eleventh International Conference on Learning Representations (ICLR), Kigali, Rwanda, 1–5 May 2023. Available online: <https://openreview.net/forum?id=frE4f4HEIB> (accessed on 19 February 2026).
17. Zhang, J.; Yang, K.; Luo, C.; Zhu, J. Spiking Transformers for Event-Based Single Object Tracking. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 19–24 June 2022; pp. 8801–8810. Available online: https://openaccess.thecvf.com/content/CVPR2022/html/Zhang_Spiking_Transformers_for_Event-Based_Single_Object_Tracking_CVPR_2022_paper.html (accessed on 19 February 2026).
18. Xu, Z.; Zeng, W.; Chu, X.; Cao, P. Multi-Aircraft Trajectory Collaborative Prediction Based on Social Long Short-Term Memory Network. *Aerospace* **2021**, *8*, 115. <https://doi.org/10.3390/aerospace8040115>.
19. Giuliani, F.; Hasan, I.; Cristani, M.; Galasso, F. Transformer networks for trajectory forecasting. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 10335–10342. <https://doi.org/10.1109/ICPR48806.2021.9412196>.
20. Ayhan, S.; Samet, H. Aircraft Trajectory Prediction Made Easy with Predictive Analytics. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 13–17 August 2016; pp. 21–30. <https://doi.org/10.1145/2939672.2939694>.
21. Zhuang, B.; Liu, J.; Pan, Z.; He, H.; Weng, Y.; Shen, C. A Survey on Efficient Training of Transformers. In Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI), Macao, China, 19–25 August 2023; pp. 6823–6831. <https://doi.org/10.24963/ijcai.2023/762>.
22. Tavanaei, A.; Maida, A.S. Bio-Inspired Spiking Convolutional Neural Network Using Layer-Wise Sparse Coding and STDP Learning. *IEEE Trans. Cogn. Dev. Syst.* **2018**, *10*, 819–838. <https://doi.org/10.1109/TCDS.2018.2816953>.
23. Thorpe, S.; Delorme, A.; Van Rullen, R. Spike-based strategies for rapid processing. *Neural Netw.* **2001**, *14*, 715–725. [https://doi.org/10.1016/S0893-6080\(01\)00083-1](https://doi.org/10.1016/S0893-6080(01)00083-1).
24. Wu, Y.; Deng, L.; Li, G.; Zhu, J.; Shi, L. Direct training for spiking neural networks: Faster, larger, better. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 1311–1318. <https://doi.org/10.1609/aaai.v33i01.33011311>.
25. Neftci, E.O.; Mostafa, H.; Zenke, F. Surrogate gradient learning in spiking neural networks. *IEEE Signal Process. Mag.* **2019**, *36*, 61–63. <https://doi.org/10.1109/MSP.2019.2931595>.
26. Deng, L.; Wu, Y.; Hu, X.; Liang, L.; Ding, Y.; Li, G.; Zhao, G.; Li, P.; Xie, Y. Rethinking the performance comparison between SNNs and ANNs. *Neural Netw.* **2020**, *121*, 294–318. <https://doi.org/10.1016/j.neunet.2019.09.005>.
27. Gehrig, M.; Shrestha, S.B.; Mouritzen, D.; Scaramuzza, D. In Proceedings of the Event-Based Angular Velocity Regression with Spiking Networks. 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May 2020–31 August 2020; pp. 4195–4202. <https://doi.org/10.1109/ICRA40945.2020.9197133>.
28. Li, J.; Ma, H.; Zhang, Z.; Li, J.; Tomizuka, M. Spatio-Temporal Graph Dual-Attention Network for Multi-Agent Prediction and Tracking. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 10556–10569. <https://doi.org/10.1109/TITS.2021.3094821>.
29. Glina, Y.; Jordan, R.; Ishutkina, M. A Tree-Based Ensemble Method for Prediction and Uncertainty Quantification of Aircraft Landing Times. In Proceedings of the 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Indianapolis, Indiana, 17–19 September 2012. <https://doi.org/10.2514/6.2012-5441>.
30. Lakshminarayanan, B.; Pritzel, A.; Blundell, C. Simple and scalable predictive uncertainty estimation using deep ensembles. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 4 December 2017; p. 30. Available online: <https://proceedings.neurips.cc/paper/2017/file/9ef2ed4b7fd2c810847ffa5fa85bce38-Paper.pdf> (accessed on 19 February 2026).
31. Porretta, M.; Dupuy, M.D.; Schuster, W.; Majumdar, A.; Ochieng, W. Performance evaluation of a novel 4D trajectory prediction model for civil aircraft. *J. Navig.* **2008**, *61*, 393–420. <https://doi.org/10.1017/S037346330800482X>.

32. Wiest, J.; Höffken, M.; Kreßel, U.; Dietmayer, K. Probabilistic trajectory prediction with Gaussian mixture models. In Proceedings of the 2012 IEEE Intelligent Vehicles Symposium, Madrid, Spain, June 3–7, 2012; pp. 141–146. <https://doi.org/10.1109/IVS.2012.6232277>.
33. Thiele, J.C.; Bichler, O.; Dupret, A. Event-based, timescale invariant unsupervised online deep learning with STDP. *Front. Comput. Neurosci.* **2018**, *12*, 46. <https://doi.org/10.3389/fncom.2018.00046>.
34. Shaw, P.; Uszkoreit, J.; Vaswani, A. Self-attention with relative position representations. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), New Orleans, LA, USA, 1–6 June 2018; pp. 464–468. <https://doi.org/10.18653/v1/N18-2074>.
35. Sengupta, A.; Ye, Y.; Wang, R.; Liu, C.; Roy, K. Going deeper in spiking neural networks: VGG and residual architectures. *Front. Neurosci.* **2019**, *13*, 95. <https://doi.org/10.3389/fnins.2019.00095>.
36. Ayhan, S.; Samet, H. Diclerge: Divide-cluster-merge framework for clustering aircraft trajectories. In Proceedings of the 8th ACM SIGSPATIAL International Workshop on Computational Transportation Science, Washington, DC, USA, 3 November 2015; pp. 1–6. <https://doi.org/10.1145/2820305.2820306>.
37. Gariel, M.; Srivastava, A.N.; Feron, E. Trajectory clustering and an application to airspace monitoring. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 1511–1524. <https://doi.org/10.1109/TITS.2011.2160628>.
38. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555. Available online: <https://arxiv.org/abs/1412.3555> (accessed on 19 February 2026).

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