Integrated Fuzzy-Gated Long Short-Term Memory Networks for Enhanced Multivariate Time Series Uncertainty Quantification

Rajendran Swamidurai* and Uma Kannan

Department of Mathematics and Computer Science, Alabama State University, Montgomery, AL, USA

***Corresponding Author**

DOI: https://doi.org/10.68050/JAMS.2026.253

Received: 31 Aug 2026; Accepted: 01 Sept 2026; Published: 09 September 2026

ABSTRACT

Multivariate time series (MTS) modeling in critical domains—such as financial forecasting, smart grid management, and anomaly detection—demands robust quantification of both aleatoric and epistemic uncertainties. While Long Short-Term Memory (LSTM) networks excel at learning temporal dependencies, standard architectures rely on deterministic, real-valued operations that cannot intrinsically represent linguistic vagueness, hesitation, or cyclical phase information. To overcome these limitations, this paper introduces the mathematical foundations for three integrated Fuzzy-Gated LSTM (F-LSTM) architectures: the Hesitant Fuzzy LSTM (HFS-LSTM), the Spherical Fuzzy LSTM (SFS-LSTM), and the Complex Fuzzy LSTM (CFS-LSTM). The HFS-LSTM employs Hesitant Fuzzy Aggregation Operators within the gating mechanisms to capture epistemic uncertainty and upstream multi-fuzzification ambiguity. The SFS-LSTM integrates membership, non-membership, and explicit indeterminacy dimensions under spherical geometric constraints to quantify aleatoric uncertainty and facilitate reliable three-way diagnostic decisions. The CFS-LSTM leverages the complex plane to track cyclical and directional dynamics through phase coherence while mitigating vanishing gradients via norm-preserving state transitions. Probabilistic predictive distributions are constructed via defuzzification to enable comprehensive uncertainty evaluation using the Continuous Ranked Probability Score (CRPS). Together, these architectures establish an expressive theoretical framework for intrinsic uncertainty quantification in deep recurrent networks.

Keywords: Multivariate Time Series; Uncertainty Quantification (UQ); Long Short-Term Memory (LSTM); Hesitant Fuzzy Sets; Spherical Fuzzy Sets; Complex Fuzzy Sets; Fuzzy Recurrent Neural Networks; Three-Way Decision Making.

INTRODUCTION

Multivariate time series (MTS) analysis is foundational across critical domains such as financial market modeling, industrial equipment fault detection, and smart grid management, where the objective is to simultaneously analyze temporal dependencies and complex inter-variable relationships.[1] Accurate forecasting and reliable decision-making in these fields fundamentally depend on the quantification and management of uncertainty. Forecasting models are inherently challenged by uncertainty stemming from various sources, including irreducible randomness (aleatoric uncertainty), incomplete information, and the influence of exogenous factors.[2]

Traditional statistical approaches, such as Autoregressive Moving Average (ARMA) or Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, exhibit limitations when confronting real-world MTS data. These methods are typically ill-equipped to handle the non-probabilistic and linguistic representation of uncertainty that pervades complex systems, particularly in domains like financial forecasting.[3] The necessity for robust models that can interpret and propagate vagueness, hesitation, and ambiguity has driven research toward hybrid systems that integrate deep learning architectures with formalized mathematical tools for uncertainty management.

The Long Short-Term Memory (LSTM) network, a specialized form of Recurrent Neural Network (RNN), has gained prominence in sequential data processing due to its mechanism for mitigating the vanishing gradient problem and its ability to maintain useful dependencies over extended time steps.[4] The standard LSTM cell operates using real-valued vectors and matrix multiplications, managing information flow through three primary gates—the input gate (i_t), the forget gate (f_t), and the output gate (o_t)—alongside the candidate cell state ($\tilde{C}_t$) and the memory cell state (C_t).[5]

The mathematical operations within the LSTM, such as the forget gate calculation, $f_t = \sigma(W_f \cdot [h_{t-1}, y_t] + b_f)$, rely entirely on deterministic, point-based real-valued calculations.[4] This architecture is designed primarily to capture sequential dependencies in the data but lacks an intrinsic framework for modeling epistemic uncertainty—the uncertainty arising from ambiguity or vagueness in the data representation itself. Standard LSTMs handle aleatoric uncertainty primarily through residual error terms quantified after the prediction, rather than propagating it internally through the memory state.

The incorporation of fuzzy set theory into deep recurrent architectures offers a systematic solution to this limitation. Fuzzy set theory, introduced by Zadeh in 1965, is intended to reinforce methods of reasoning that are estimated rather than meticulous, resembling human reasoning in the face of imprecise or distorted input information.[6] Previous work has demonstrated the viability of hybrid architectures, such as those integrating Type-2 or Intuitionistic Fuzzy Sets, into RNN gates to better handle high uncertainty levels [7], as well as adaptive rule-reduction methods in Type-2 fuzzy frameworks for sequential forecasting [8].

To move beyond basic fuzzy systems and capture the multifaceted nature of uncertainty inherent in high-dimensional MTS data, this paper proposes integrating three advanced fuzzy set extensions into the fundamental operations of the LSTM cell. Each extension is chosen for its unique capability to model a specific dimension of uncertainty:

1. **Hesitant Fuzzy Sets (HFS):** Designed to model situations where there is explicit hesitation or ambiguity regarding the precise membership degree of an element.[9] HFS allows the model to capture uncertainty stemming from conflicting or equally valid data interpretations.

2. **Spherical Fuzzy Sets (SFS):** An evolution of fuzzy theory that explicitly defines and constrains degrees of membership (μ), non-membership (ν), and, critically, indeterminacy (η).[10] SFS allows the network to quantify its confidence level directly within the memory state.

3. **Complex Fuzzy Sets (CFS):** Extends membership degrees to the unit circle in the complex plane, integrating phase information.[11] CFS is suitable for modeling cyclical dependencies and directional information, while simultaneously enhancing architectural stability in deep recurrent models.[12]

This research aims to establish the mathematical foundations for three novel architectures—HFS-LSTM, SFS-LSTM, and CFS-LSTM—thereby equipping recurrent networks with advanced, intrinsic capabilities for uncertainty quantification in MTS forecasting.

THEORETICAL FOUNDATIONS: FUZZY SET EXTENSIONS AND RECURRENT DYNAMICS

Recurrent Dynamics: The Standard Long Short-Term Memory (LSTM) Cell

The core function of the LSTM cell is to maintain and update a memory state C_t and a hidden state h_t based on the current input y_t and the previous states h_{t-1} and C_{t-1}.[5] The flow of information is regulated by three primary gates, computed using the sigmoid function σ and a linear projection of the concatenated previous hidden state and current input $[h_{t-1}, y_t]$.

The sequential steps are defined as follows [4]:

1. Candidate Cell State: A proposal for new information $\tilde{C}_t$ is generated:

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, y_t] + b_c)$$

2. Forget Gate: Determines which information from the previous cell state C_{t-1} should be discarded:

$$f_t = \sigma(W_f \cdot [h_{t-1}, y_t] + b_f)$$

3. Input Gate: Determines which new candidate values $\tilde{C}_t$ will be stored in the current cell state:

$$i_t = \sigma(W_i \cdot [h_{t-1}, y_t] + b_i)$$

4. Cell State Update: The memory is updated by combining the forgotten previous state and the new scaled candidate information:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$$

5. Output Gate and Hidden State: The output gate o_t controls the flow of information to the hidden state h_t:

$$o_t = \sigma(W_o \cdot [h_{t-1}, y_t] + b_o)$$

$$h_t = o_t \cdot \tanh(C_t)$$

The real-valued matrix operations and deterministic activations ($\sigma, \tanh$) in these equations represent the crucial targets for fuzzification to enable inherent uncertainty modeling.

Hesitant Fuzzy Sets (HFS): Representation and Aggregation Operators

Hesitant Fuzzy Sets (HFSs) offer a robust mathematical framework for handling uncertain information, particularly in scenarios involving ambiguity and hesitation.[14] Unlike standard fuzzy sets, which assign a single membership degree $\mu(x) \in [0,1]$, an HFS, B, is defined as a function that, when applied to the universe of discourse U, returns a subset of membership degrees $h(x) \subset [0,1]$. [9] This subset, $h(x)$, is known as a Hesitant Fuzzy Element (HFE). The ability of HFS to accommodate multiple potential degrees of membership is particularly relevant for modeling epistemic uncertainty in MTS data. For instance, in financial time series forecasting, non-determinism can arise from the existence of multiple valid fuzzification methods (e.g., triangular versus Gaussian partitioning) that interpret the same time series data differently.[3] HFS provides a mechanism to capture this collective hesitation or ambiguity in the input signal's linguistic definition, making the network inherently robust to upstream modeling choices.

Integration requires replacing standard linear combinations and multiplication operations with Hesitant Fuzzy Aggregation Operators (HFAOs). Key HFS operations include the complement $h^c = \cup_{\gamma \in h} \{1 - \gamma\}$, the sum $h_a \oplus h_b = \cup_{\gamma_1 \in h_a, \gamma_2 \in h_b} \{\gamma_1 + \gamma_2 - \gamma_1 \gamma_2\}$, and the product $h_a \otimes h_b = \cup_{\gamma_1 \in h_a, \gamma_2 \in h_b} \{\gamma_1 \gamma_2\}$.[15] For aggregation of HFE vectors (such as weights and input features), operators like the Hesitant Fuzzy Weighted Average (HFWA) or Hesitant Fuzzy Weighted Geometric (HFWG) operators are necessary to fuse the features of multiple inputs into a single HFE representing the gate value.[15]

Spherical Fuzzy Sets (SFS): The Indeterminacy Dimension

Spherical Fuzzy Sets (SFSs), and their extensions such as T-Spherical Fuzzy Sets (TSFSs) or Globular T-Spherical Fuzzy Sets (G-TSFSs) [18], represent a sophisticated evolution in fuzzy theory. SFS addresses the inherent inability of simpler fuzzy sets (like Intuitionistic or Pythagorean fuzzy sets) to fully decouple and manage the level of non-membership and the degree of explicit indeterminacy or hesitation.[10]

An SFS defined over a universe U assigns three independent degrees to each element x: membership ($\mu(x)$), non-membership ($\nu(x)$), and indeterminacy ($\eta(x)$), subject to the strict constraint that the sum of their squares is bounded: $\mu^2(x) + \eta^2(x) + \nu^2(x) \leq 1$. [10]

The critical feature of SFS is the explicit quantification of indeterminacy (η). This degree allows the SFS-LSTM to provide a structured, three-dimensional representation of its memory state's confidence level, directly modeling aleatoric uncertainty. In high-stakes environments, such as smart grid load forecasting [21], automated medical diagnostics [19], industrial monitoring, or anomaly detection, knowing that the model is unsure (high η) is as crucial as the resulting prediction. This facilitates advanced decision-making processes, such as three-way decisions (Accept, Non-Commitment/Indeterminate, Reject), based on the confidence scores (η) associated with the output. [22]

SFS integration necessitates specialized algebraic operations, as standard fuzzy operations do not preserve the $\mu^2 + \eta^2 + \nu^2 \leq 1$ constraint. Operational laws often utilize complex aggregation techniques, such as those based on Einstein T-norms and T-conorms [23], or more flexible methods like Hamacher operational laws, which provide exceptional flexibility in dealing with data vagueness.[24]

Complex Fuzzy Sets (CFS): Phase Coherence and State Stability

Complex Fuzzy Sets (CFSs) extend the membership degree from the real interval [0,1] to the unit circle in the complex plane. [11] The Complex Fuzzy Number (CFN) membership function is given by $\mu_C(x) = a(x)e^{j\omega(x)}$, where *a(x)* is the amplitude degree and $\omega(x) \in [0,2\pi]$ is the phase degree.[13]

CFS offers two profound advantages for MTS modeling. First, the phase angle $\omega(x)$ naturally captures cyclical, directional, or frequency-domain information inherent in time series data, such as seasonality or complex system dynamics.[13] This makes CFS particularly suitable for applications in signal processing and system identification.[25]

Second, CFS integration is inherently tied to the architectural advantages found in Complex-Valued Recurrent Neural Networks (CV-RNNs). Standard RNNs and LSTMs are prone to vanishing or exploding gradients when learning long-term dependencies.[12] CV-RNNs often employ unitary (norm-preserving) state transition matrices, which fundamentally mitigate these gradient issues by ensuring that the norm of a vector remains unchanged after repeated multiplication.[12] By adopting complex-valued computation in the gates (CFS-LSTM), the architecture gains significant numerical stability over extended sequences, offering an essential engineering advantage beyond its ability to model phase information.

Implementing a CFS-LSTM requires defining complex-valued operational laws for matrix multiplication and ensuring that the complex-valued gates use analytic activation functions.[27]

The following table summarizes the unique contribution of each advanced fuzzy set to MTS modeling:

Table 1: Advanced Fuzzy Set Comparison for Time Series Modeling

Fuzzy Set Theory	Core Constraint	Uncertainty Aspect Modeled	Integration Focus in LSTM	Primary MTS Application
Hesitant Fuzzy Set (HFS)	Membership degree $h(x) \subset [0,1]$ (HFE)	Epistemic Uncertainty, Hesitation, Multi-Fuzzification ambiguity	Input Encoding and Gate Aggregation	Financial Volatility, Feature Selection [3]
Spherical Fuzzy Set (SFS)	$\mu^2 + \eta^2 + \nu^2 \leq 1$	Aleatoric Uncertainty, Explicit Indeterminacy (η)	Cell State Update and Confidence Scoring	Anomaly Detection, Diagnostics in Smart Grids [10]
Complex Fuzzy Set (CFS)	Membership on the complex unit circle	Directional/Cyclical Information, Phase Coherence	Gate Activation and Matrix Operations (Complex Domain)	Signal Processing, System Dynamics Forecasting [12]

MATHEMATICAL ARCHITECTURE OF THE FUZZY-GATED LSTM (F-LSTM)

General Strategy for Gate Fuzzification and Fuzzy Input Encoding

The integration of advanced fuzzy sets necessitates a systematic replacement of the core real-valued components of the standard LSTM cell. This strategy involves three steps:

1. **Input and State Fuzzification:** Converting the MTS input $X_t \in \mathbb{R}^N$ and the previous hidden state $h_{t-1} \in \mathbb{R}^M$ into corresponding vectors of Fuzzy Elements (HFE, SFN, or CFN), denoted X_t^F and h_{t-1}^F.
2. **Fuzzy Weights and Bias:** The weight matrices W_G and bias vectors b_G for each gate $G \in \{f, i, o, c\}$ must be defined as fuzzy numbers or tensors W_G^F and b_G^F.
3. **Operation Replacement:** Substituting real-valued matrix multiplication $(W \cdot X)$ with fuzzy aggregation operators and replacing the real-valued sigmoid and tanh activations with functions suitable for the respective fuzzy domain.

Architecture I: The Hesitant Fuzzy LSTM (HFS-LSTM)

The HFS-LSTM is designed to manage ambiguity where inputs or states are represented by multiple potential membership degrees (HFEs).

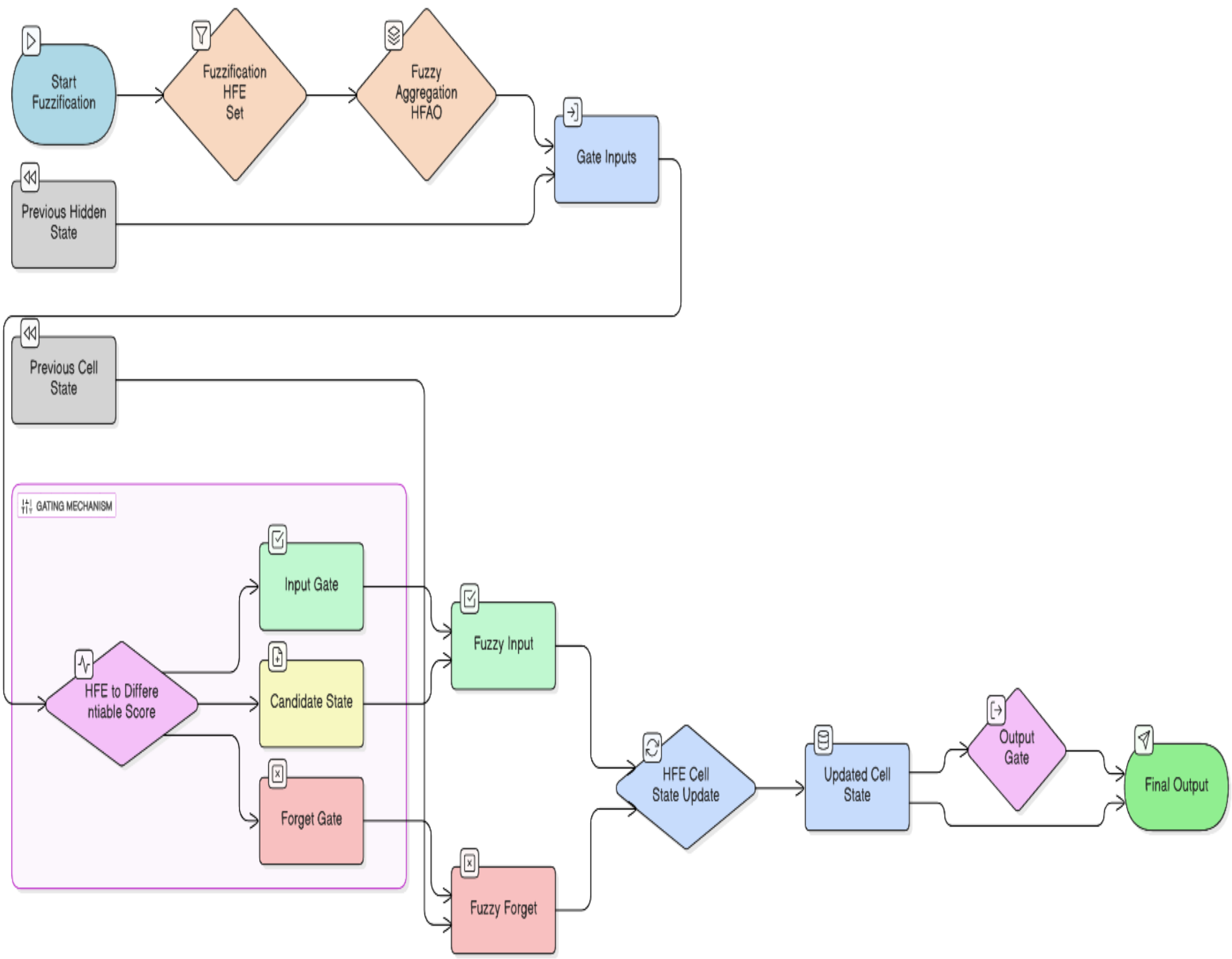

Figure-1: Hesitant Fuzzy LSTM (HF-LSTM) Architecture

Input and Gate Aggregation

The concatenated input vector $[h_{t-1}^H, X_t^H]$ consists of HFEs. The linear weighted sum operation of the standard LSTM must be replaced by a Hesitant Fuzzy Aggregation Operator (HFAO). The HFWG operator is often preferred over the arithmetic mean as geometric operators tend to maintain consistency more effectively in certain fuzzy decision environments.[16] The fuzzy weighted sum for any gate $G \in \{f, i, c\}$ is computed as:

$$G_t^H = \mathrm{HFAO}(W_G^H, [h_{t-1}^H, X_t^H])$$

Gate Activation and Decision Mapping

Since the gates (f_t^H, i_t^H) must approximate a forget or input intensity (a value between 0 and 1, analogous to the sigmoid output), the aggregated HFE G_t^H must undergo a transformation. This transformation involves applying a score function, $S(\cdot)$, to map the HFE G_t^H to a single, comparable real number, followed by a Hesitant Fuzzy Comparison Operator to obtain a final HFE decision D_t^H, utilizing score functions and comparison operators developed for hesitant ranking and decision environments [17], [28].

HFS Cell State Update

The core LSTM update leverages HFS operational laws.[15] The standard multiplication $(\cdot)$ is replaced by the HFS product ($\otimes$, acting as a T-norm), and the addition $(+)$ is replaced by the HFS sum ($\oplus$, acting as a T-conorm).

1. Candidate Cell State: $\tilde{C}_t^H$ is computed via an HFAO corresponding to the tanh activation.

2. Cell State Update:

$$C_t^H = (f_t^H \otimes C_{t-1}^H) \oplus (i_t^H \otimes \tilde{C}_t^H)$$

The HFS-LSTM propagates the hesitation (the set of possible membership values) forward through time, ensuring that the inherent ambiguity in the input signal's definition is maintained and accounted for in the memory state. Finally, the hidden state h_t^H requires defuzzification via a score function for use in subsequent layers or for final output prediction.

Architecture II: The Spherical Fuzzy LSTM (SFS-LSTM)

The SFS-LSTM is uniquely positioned to manage and propagate explicit confidence or indeterminacy (η) within the recurrent memory. All state vectors (X_t^S, h_{t-1}^S, C_t^S) are represented by SFS vectors, where each dimension holds a tuple (μ, η, ν).

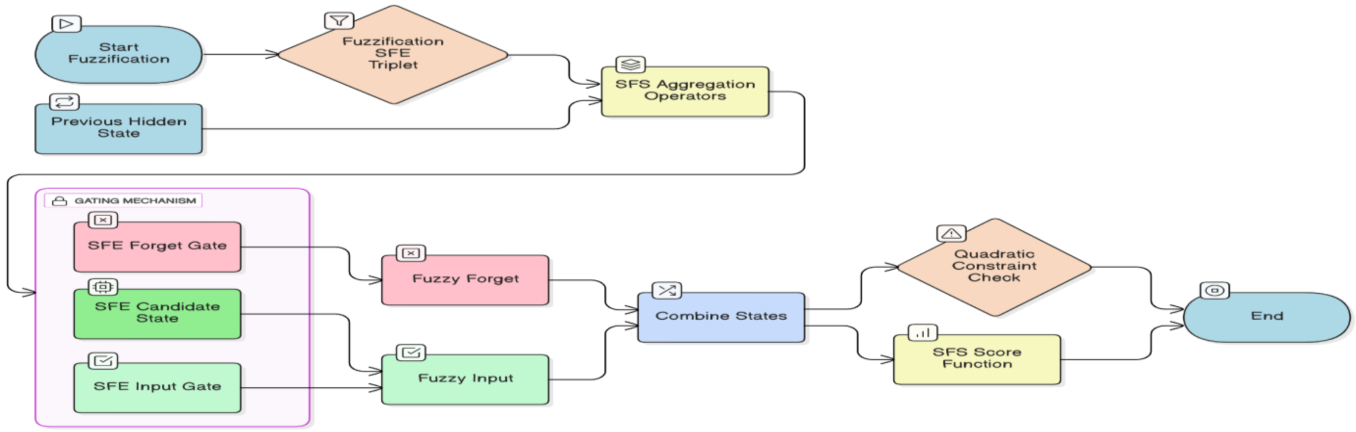

Figure-2: Spherical Fuzzy LSTM (SF-LSTM) Architecture

Weighted Sum Derivation

The aggregated input for any gate must maintain the strict spherical constraint $\mu^2 + \eta^2 + \nu^2 \leq 1$. This necessitates the use of specialized SFS aggregation operators. For instance, T-Spherical Fuzzy Sets (TSFSs) and their extensions utilize operators designed to manage imprecise and uncertain events.[23] Aggregation operators based on Hamacher t-norms and t-conorms are often employed due to their exceptional flexibility in managing uncertainty, allowing the aggregated result to consistently abide by the spherical constraint.[24]

SFS Gate Activation and Indeterminacy Flow

The activation layers for the gates must output three constrained degrees (μ_G, η_G, ν_G). The SFS-specific score functions map the aggregated input magnitude to these three degrees while strictly enforcing the squared sum constraint. The most critical aspect of the SFS-LSTM is how the indeterminacy dimension flows through the cell state update. The SFS operational laws, unlike those of simpler fuzzy sets, are specifically designed to combine the three degrees coherently. If the previous state C_{t-1} exhibits high indeterminacy (η_{t-1} is large) or the new candidate information $\tilde{C}_t$ is uncertain ($\tilde{\eta}_t$ is large), the SFS-Sum and SFS-Product operations will systematically fuse this lack of confidence into the new state C_t^S.

SFS Cell State Update

The SFS cell state C_t^S is updated using SFS-specific operational laws for multiplication (T-norm) and addition (T-conorm), ensuring the spherical constraint is upheld throughout the temporal propagation:

$$C_t^S = (f_t^S \otimes_{SFS} C_{t-1}^S) \oplus_{SFS} (i_t^S \otimes_{SFS} \tilde{C}_t^S)$$

This systematic propagation of η allows the model to differentiate, for example, between a clearly wrong prediction (low μ, high ν) and a robustly uncertain prediction (high η). The resulting indeterminacy degree η in the final hidden state h_t^S can then be used as a confidence measure, directly informing three-way decisions in applications requiring high reliability.[22]

Architecture III: The Complex Fuzzy LSTM (CFS-LSTM)

The CFS-LSTM operates in the complex domain ($\mathbb{C}$), allowing it to harness phase information for cyclical modeling and leverage complex-valued operations for enhanced stability.

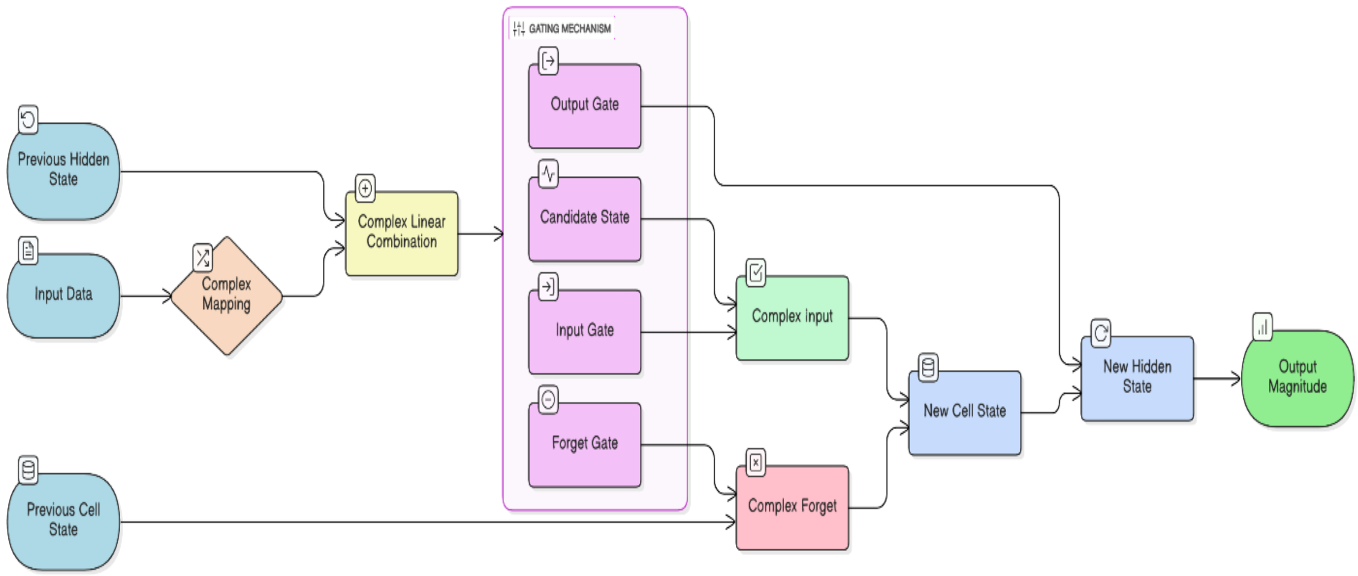

Figure-3: Complex Fuzzy LSTM (CF-LSTM) Diagram

State Representation and Input Mapping

All weight matrices W^C and biases b^C are complex numbers. The MTS input $X_t \in \mathbb{R}^N$ must be mapped into a complex vector $X_t^C \in \mathbb{C}^N$, where the phase component ω_t is utilized to encode temporal periodic dynamics or directional information. Similarly, the states h_{t-1}^C and C_{t-1}^C are complex vectors.

Weighted Sum and Stability

The linear weighted sum is a straightforward generalization using standard complex-valued matrix multiplication [26]:

$$G_t^C = W_C \cdot [h_{t-1}^C, X_t^C] + b_C$$

The primary benefit of this complex-valued operation in the recurrent setting is gradient stability. Complex Gated Recurrent Neural Networks (C-RNNs) that use complex numbers maintain norm-preserving state transitions, which fundamentally prevents the vanishing or exploding gradient problems that plague real-valued deep recurrent architectures.[12]

Complex Activation Functions

The standard real-valued activations $\sigma(\cdot)$ and $\tanh(\cdot)$ are not suitable for the complex domain. They must be replaced by complex analytic functions, $\sigma_C(\cdot)$ and $\tanh_C(\cdot)$, to ensure the preservation of complex-valued properties and maintain gradient stability during training.[12]

CFS Cell State Update Derivation

The update rules utilize complex addition ($+$) and element-wise complex multiplication ($\odot$) [27]:

1. Candidate Cell State: $\tilde{C}_t^C = \tanh_C(W_c^C \cdot [h_{t-1}^C, X_t^C] + b_c^C)$
2. Forget/Input Gates: $f_t^C = \sigma_C(W_f^C \cdot [h_{t-1}^C, X_t^C] + b_f^C)$, $i_t^C = \sigma_C(W_i^C \cdot [h_{t-1}^C, X_t^C] + b_i^C)$
3. Cell State Update: $C_t^C = f_t^C \odot C_{t-1}^C + i_t^C \odot \tilde{C}_t^C$

The phase component of the complex memory state C_t^C continuously tracks the temporal and directional relationships of the time series, effectively providing an explicit model of cyclical dynamics.[25]

Table 2 provides a comparative summary of the necessary modifications across the three F-LSTM architectures.

Table 2: F-LSTM Gate Modification Overview (Focusing on Core Operations)

Standard Operation (Real)	HFS-LSTM Modification	SFS-LSTM Modification	CFS-LSTM Modification
$W \cdot X$ (Weighted Sum)	Hesitant Fuzzy Weighted Aggregation (HFAO) [16]	Spherical Fuzzy Aggregation (e.g., Hamacher/Einstein) [23]	Complex-Valued Matrix Multiplication ($W_C \cdot X_C$) [26]
$\sigma(\cdot)$ (Gate Activation)	HFS Score Function + Comparison Operator [17]	SFS-specific score function, maintaining constraint [10]	Complex Sigmoid/Tanh (σ_C, $\tanh_C$) [12]
C_t Update (Multiplication/Addition)	HFE T-Norm ($\otimes$) and T-Conorm ($\oplus$) [15]	SFS Operational Laws (Sum, Product) [23]	Complex Addition ($+$) and Complex Multiplication ($\odot$) [27]

PERFORMANCE ANALYSIS AND OPTIMIZATION CHALLENGES

Comparative Complexity Analysis

The incorporation of advanced fuzzy mathematics introduces distinct computational trade-offs compared to the standard, highly optimized real-valued LSTM.

The **HFS-LSTM** involves significant complexity due to the nature of Hesitant Fuzzy Elements (HFEs). Since HFEs are sets of possible membership degrees, aggregation operations require handling multiple values, often involving sorting, intersection, or union operations. [15] Although low-complexity HFS operators exist [29], the processing of sets, rather than single numbers, fundamentally increases computational complexity and overhead compared to deterministic point-based systems.

The **SFS-LSTM** requires tracking three dimensions ((μ, η, ν)) per fuzzy number, which intrinsically triples the memory footprint per state element. Furthermore, the aggregation operators (e.g., Hamacher-based, Einstein-based) are mathematically sophisticated, specifically designed to uphold the spherical constraint $\mu^2 + \eta^2 + \nu^2 \leq 1$.[23] The maintenance of this nonlinear constraint during forward and backward passes contributes substantially to the computational overhead relative to Type-1 fuzzy systems.

The **CFS-LSTM** involves complex numbers, effectively doubling the parameter space since each weight and state must store both a real and an imaginary component. [12] However, this dimensional increase is often offset by a compensating advantage in training efficiency. The inherent stability derived from unitary state transitions in the complex domain mitigates chronic optimization difficulties like vanishing gradients. [12] This architectural resilience can lead to faster convergence and the ability to train deeper recurrent networks without requiring extensive regularization or gradient clipping, potentially resulting in a lower effective computational cost over the entire training lifecycle compared to standard LSTMs struggling with unstable optimization.

Optimization of Fuzzy Weight Learning

A major challenge in training F-LSTMs, particularly HFS-LSTM and SFS-LSTM, is the optimization process. Standard gradient descent relies on the differentiability of all operations. Many fuzzy aggregation operators (HFAO, SFS Hamacher operators) and the constrained mapping functions used for gate activation are defined piecewise or through complex aggregation rules, making them non-differentiable or difficult to differentiate robustly.

To overcome this, specialized approaches must be employed:

1. **Surrogate Gradients:** Using approximations or smoothed activation functions to enable gradient flow across the non-differentiable points of the fuzzy operators.

2. **Interval-Based Optimization:** For HFS, optimization and continuous envelope bounds can leverage operational laws and representations established for interval-valued hesitant fuzzy sets (IVHFS) [30].

3. **Nonlinear Constraint Optimization:** For SFS, weight learning may rely on specialized optimization models, such as those used in Spherical Fuzzy Best–Worst Method (SF-BWM), which utilize nonlinear constraints to determine optimal fuzzy weight coefficients.

Furthermore, external decision-making frameworks, such as those based on Complex Spherical Fuzzy Rough Sets (CSFRSs), may be necessary to inform the *selection* of the appropriate deep learning model architecture itself, particularly when dealing with highly ambiguous or subjective input data. [24]

Defuzzification and Uncertainty Projection

Since the output of an F-LSTM (h_t^F) is a fuzzy element (HFE, SFN, or CFN), it must be converted (defuzzified) into a usable real-valued forecast $\hat{y}_t$ for practical application. For HFS and SFS, defuzzification typically involves using score functions. For HFS, the score function often involves calculating a representative value based on the aggregated membership degrees, such as the average of maximum values within the HFE set.[17] For SFS, score functions often rely on similarity measures or distance metrics (like Hamming or Euclidean distance for G-TSFSs).[18]

Crucially, the inherent uncertainty quantified by the fuzzy element must be projected onto a statistical confidence measure. The dispersion of the HFS set, the indeterminacy degree η of the SFS, or the amplitude uncertainty of the CFS, must be mapped to a standard deviation σ_t. This mapping converts the epistemic/aleatoric uncertainty captured internally by the fuzzy system into the parameters $(\hat{y}_t, \sigma_t)$ of a probabilistic predictive distribution, $P(\hat{y}_t)$, which is necessary for rigorous uncertainty evaluation.

APPLICATIONS AND EVALUATION IN HIGH-UNCERTAINTY MTS DOMAINS

Case Study: Financial Volatility Forecasting (HFS-LSTM)

Financial time series forecasting represents a challenging domain characterized by non-probabilistic volatility and linguistic representation of market uncertainties. [3] The HFS-LSTM is particularly suitable here because financial data often involves conflicting opinions or valid but differing interpretations (non-determinism) regarding data fuzzification—for example, when using triangular versus Gaussian methods to define linguistic terms. [3]

By incorporating Dual Hesitant Fuzzy Sets (DHFS), a variant of HFS that accounts for both membership and non-membership hesitancy, the HFS-LSTM can construct Intuitionistic Fuzzy Sets from inputs fuzzified by multiple methods simultaneously. [3] This mechanism allows the model to capture the inherent ambiguities in market dynamics, outperforming both traditional statistical techniques (ARMA, ARCH) and discrete hesitant fuzzy time series baseline methods [3], [31], as confirmed by standard error metrics like Root Mean Square Error (RMSE) and Average Forecasting Errors (AFER).

Case Study: Anomaly Detection and Diagnostics (SFS-LSTM)

Applications requiring high reliability and real-time assessment of confidence, such as smart grid load forecasting or industrial fault diagnostics, benefit significantly from the explicit indeterminacy modeling provided by SFS.[21]

The SFS-LSTM outputs the three degrees (μ_t, η_t, ν_t)During inference, the η_t value (indeterminacy) acts as a reliable measure of the model's self-assessed confidence. If the input data is anomalous, corrupted, or falls into a highly ambiguous region of the feature space, η_t will spike. This enables the implementation of a robust three-way classification system [22]:

1. Acceptance: Low $\eta_t$and high μ_t.
2. Rejection/Indeterminate: High η_tThese instances are flagged as being too uncertain for confident classification and are passed to human experts or subsequent advanced techniques for arbitration.
3. Non-Acceptance: Low μ_t and high ν_t. .

By quantifying the quality of confidence, SFS-LSTM ensures that decisions made in response to predicted load changes or potential equipment faults are based not just on the predicted value, but also on the calculated certainty level.[10]

Case Study: Directional Signal Processing (CFS-LSTM)

The CFS-LSTM excels in domains where time series are characterized by strong cyclical components, phase coherence, or directional shifts, particularly in signals and systems. [25] The complex-valued structure provides computational behavior analogous to a Fourier transform in certain contexts.[25]

In signal processing applications, the CFS-LSTM's capacity to track the phase component of the memory state allows it to efficiently model and identify periodic or quasi-periodic dynamics. This is crucial for tasks like system identification, where an open system's state depends not only on past states but also on external stimuli. [26] Furthermore, the stability benefits derived from using complex gated units [12] ensure that the model can maintain long-term representations of these oscillating dynamics accurately, even across extended temporal sequences.

Evaluation Metrics for Fuzzy UQ

Given that F-LSTMs output both a point prediction and an associated measure of uncertainty (dispersion, indeterminacy, etc.), simple deterministic error metrics like Mean Absolute Error (MAE) or Root Mean Square Error (RMSE) are insufficient for comprehensive evaluation. [32] Robust evaluation requires metrics specifically designed for probabilistic forecasting.

Continuous Ranked Probability Score (CRPS): CRPS is considered a proper scoring rule and is the mandatory evaluation metric for assessing the quality of uncertainty quantification. [32] The process of utilizing CRPS involves several key steps:

1. **Prediction Distribution Construction:** The fuzzy output element (h_t^F) is defuzzified to yield a point prediction $\hat{y}_t$. The inherent fuzzy uncertainty (e.g., SFS η_t or HFS dispersion) is then mapped to a standard deviation σ_t, establishing a complete predictive distribution $P(\hat{y}_t) \sim N(\hat{y}_t, \sigma_t)$.
2. **CRPS Calculation:** CRPS is calculated based on this distribution and the observed actual value y_t.

CRPS rewards models that are both accurate and provide well-calibrated confidence bounds, thereby validating the F-LSTM's core contribution: the useful quantification of uncertainty.[32] For inter-series comparison, Mean Absolute Scaled Error (MASE) remains a valuable metric, as it normalizes errors across different scales. [32]

CONCLUSION

This report has detailed the mathematical framework for integrating three advanced fuzzy set theories—Hesitant Fuzzy Sets (HFS), Spherical Fuzzy Sets (SFS), and Complex Fuzzy Sets (CFS)—into the Long Short-Term Memory cell for multivariate time series analysis.

The proposed **HFS-LSTM** architecture successfully addresses epistemic uncertainty and input ambiguity by replacing real-valued aggregations with Hesitant Fuzzy Aggregation Operators (HFAOs). This enables the memory state to capture non-determinism arising from multiple valid interpretations of the data, offering enhanced robustness in contexts like financial volatility forecasting.

The **SFS-LSTM** explicitly quantifies aleatoric uncertainty through the indeterminacy degree (η), propagated systematically via constrained SFS operational laws. This yields a transparent confidence score that facilitates sophisticated diagnostic and decision-making capabilities, such as three-way classification crucial for smart grid reliability and anomaly detection.

The **CFS-LSTM** leverages the complex plane to model directional/cyclical dynamics through phase information and, critically, achieves architectural stability by utilizing complex analytic activation and norm-preserving operations, mitigating the severe optimization challenges posed by vanishing gradients in deep recurrent systems.

REFERENCES

1. Wang, F., Jiang, Y., Zhang, R., Wei, A., Xie, J., & Pang, X. (2025). A survey of deep anomaly detection in multivariate time series: Taxonomy, applications, and directions. *Sensors*, 25(1), 190.
2. Eki.Lab. (2025). Uncertainty in time series forecasting. *Ekimetrics*. https://ekimetrics.github.io/blog/Uncertainty_TS/
3. Yolcu, U., Bas, E., & Egrioglu, E. (2018). Intuitionistic fuzzy time series forecasting based on dual hesitant fuzzy sets. In *Handbook of Research on Emergent Applications of Optimization Algorithms* (pp. 583–606). IGI Global.
4. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
5. Baytas, I. M., Xiao, C., Zhang, X., Wang, F., Jain, A. K., & Zhou, J. (2017). Patient subtyping via time-aware LSTM networks. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 65–74).
6. L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, no. 3, pp. 338–353, Jun. 1965, doi: 10.1016/S0019-9958(65)90241-X.
7. Pratama, M., Lu, J., Zhang, G., & Zhou, E. (2019). Online evolving interval type-2 intuitionistic fuzzy LSTM-neural networks for regression problems. *IEEE Transactions on Fuzzy Systems*, 28(8), 1801–1815.
8. Rubio, E. J., Castillo, O., & Melin, P. (2023). Self-learning type-2 fuzzy systems with adaptive rule reduction for time series forecasting. *PeerJ Computer Science*, 9, e1458.
9. Torra, V. (2010). Hesitant fuzzy sets. *International Journal of Intelligent Systems*, 25(6), 529–539. https://doi.org/10.1002/int.20418
10. Kutlu Gündoğdu, F., & Kahraman, C. (2019). Spherical fuzzy sets and spherical fuzzy TOPSIS method. *Journal of Intelligent & Fuzzy Systems*, 36(1), 337–352. https://doi.org/10.3233/JIFS-181401
11. Ramot, D., Milo, R., Friedman, M., & Kandel, A. (2002). Complex fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 10(2), 171–186.
12. Wolter, M., & Yao, A. (2018). Complex gated recurrent neural networks. In *Advances in Neural Information Processing Systems* (NeurIPS 2018), 31, 10536–10546.
13. Tamir, D. E., Jin, L., & Kandel, A. (2020). Linguistic complex fuzzy sets. *Axioms*, 9(4), 128.
14. Rodríguez, R. M., Martínez, L., & Herrera, F. (2012). Hesitant fuzzy linguistic term sets for decision making. *IEEE Transactions on Fuzzy Systems*, 20(1), 109–119.
15. Xia, M., & Xu, Z. (2011). Hesitant fuzzy information aggregation in decision making. *International Journal of Approximate Reasoning*, 52(3), 395–407.
16. Xu, Z., & Xia, M. (2011). Distance and similarity measures for hesitant fuzzy sets. *Information Sciences*, 181(11), 2128–2138. https://doi.org/10.1016/j.ins.2011.01.028
17. Wei, G. (2012). Hesitant fuzzy prioritized operators and their application to multiple attribute decision making. *Knowledge-Based Systems*, 32, 31–38. https://doi.org/10.1016/j.knosys.2012.03.011
18. Ullah, K., Mahmood, T., & Jan, N. (2020). Similarity measures for T-spherical fuzzy sets with applications in pattern recognition. *Symmetry*, 12(2), 193. https://doi.org/10.3390/sym12020193
19. Mahmood, T., Ullah, K., Khan, Q., & Jan, N. (2019). An approach toward decision-making and medical diagnosis problems using the concept of spherical fuzzy sets. *Neural Computing and Applications*, 31(11), 7041–7053.
20. Kutlu Gündoğdu, F., & Kahraman, C. (2020). A novel spherical fuzzy analytic hierarchy process and its application to industrial robot selection. *Computers & Industrial Engineering*, 140, 106239.
21. Al-Gabalawy, M., Hosny, N. S., & Hussien, S. A. (2024). Real-time load forecasting and adaptive control in smart grids using a hybrid neuro-fuzzy approach. *Energies*, 17(11), 2539.
22. Yao, Y. (2012). An outline of a theory of three-way decisions. In *Rough Sets and Current Trends in Computing* (RSCTC 2012), Lecture Notes in Computer Science, vol 7413 (pp. 1–17). Springer.
23. Munir, M., Kalsoom, H., Ullah, K., Mahmood, T., & Chu, Y.-M. (2021). T-spherical fuzzy Einstein hybrid aggregation operators and their applications in multi-attribute decision making problems. *Symmetry*, 13(5), 865.
24. Ali, Z., Mahmood, T., & Yang, M.-S. (2020). Complex T-spherical fuzzy Hamacher aggregation operators and their applications in multi-attribute decision making. *IEEE Access*, 8, 149064–149086.

25. Dick, S. (2005). Towards complex fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 13(3), 405–414.
26. Hirose, A. (2012). *Complex-Valued Neural Networks: Advances and Applications*. IEEE Press/Wiley.
27. Wisdom, S., Powers, T., Hershey, J., Le Roux, J., & Atlas, L. (2016). Full-capacity unitary recurrent neural networks. In *Advances in Neural Information Processing Systems* (NeurIPS 2016), 29, 4880–4888.
28. Farhadinia, B. (2013). Multiple criteria decision making methods with hesitant fuzzy sets. *Information Sciences*, 259, 400–415.
29. Liao, H., & Xu, Z. (2014). Subtraction and division operations over hesitant fuzzy sets. *Journal of Intelligent & Fuzzy Systems*, 27(1), 65–72.
30. Chen, N., Xu, Z., & Xia, M. (2013). Interval-valued hesitant fuzzy sets and their applications in multi-criteria decision support systems. *Expert Systems with Applications*, 40(1), 316–324.
31. Bisht, K., & Kumar, S. (2016). Fuzzy time series forecasting method based on hesitant fuzzy sets. *Expert Systems with Applications*, 64, 557–568.
32. Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378.
33. Yang, L., Zhang, X., & Chen, H. (2025). Enhancing time series forecasting with fuzzy attention-integrated transformers. *arXiv preprint arXiv:2504.00070*.