# IMPROVING CHANNEL ESTIMATION IN VEHICULAR COMMUNICATION NETWORKS WITH DIVERSE MOBILITY PATTERNS USING BI-LSTM MODEL

Daniel Asuquo [1], Moses Ekpenyong [2,3], Enefiok Etuk [4], Kingsley Attai [5], Philip Asuquo [6] and Temple Nwaeboiwe [2]

[1] Department of Information Systems, Faculty of Computing, University of Uyo, PMB 1017, Uyo, Nigeria
[2] Department of Computer Science, Faculty of Computing, University of Uyo, PMB 1017, Uyo, Nigeria
[3] STEM Centre, University of Uyo, and Centre for Research, University of Uyo, PMB 1017, Uyo, Nigeria
[4] Department of Computer Science, College of Physical and Applied Sciences, Michael Okpara University of Agriculture, Umudike, Nigeria
[5] Department of Mathematics and Computer Science, Ritman University, Ikot Ekpene, Nigeria
[6] Department of Computer Engineering, Faculty of Engineering, University of Uyo, Nigeria

### *ABSTRACT*

*Accurate and low-latency channel estimation is essential for reliable vehicle-to-vehicle (V2V) communication in high-mobility environments such as intelligent transportation systems (ITS). Conventional techniques such as Least Squares and Minimum Mean Square Error perform poorly in dynamic wireless environments. This research introduces a deep learning (DL)-based channel estimation model employing a Bi-directional Long Short-Term Memory (Bi-LSTM) network, and evaluates its performance against traditional methods as well as a range of machine learning (ML) and DL models, including Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Recurrent Neural Network (RNN), and LSTM. Using the CN+ vehicular dataset, the models were trained on features like velocity, distance, signal strength, Doppler shift, and path delay. Results show that ML models, particularly RF and XGBoost, achieve high accuracy, with RF reaching 99.94%. Among DL models, Bi-LSTM performs best with 98.58% accuracy, and outperforms other models under high-speed conditions, due to its ability to capture temporal dependencies and track rapid channel variations. Thus, AI-based approaches can enhance channel estimation for safer and smarter vehicular communication systems.*

### *KEYWORDS*

*Channel Estimation, Vehicular Networks, High-Mobility Environments & Deep Learning Model*

## 1. INTRODUCTION

The primary advantage of wireless communications, a quickly developing technology, is mobility, which allows information transfer without physical ties [1]. Significant uses include unmanned aerial vehicles (UAVs), also known as drones, which are a focus of wireless communications research due to their widespread use in public, military, and civilian settings.

Recently, industry and academics have shown great interest in wireless networks that facilitate high-mobility broadband access [2]. Connected vehicles, or vehicular networks, combined with advanced sensing and computing technologies, enhance connectivity in smart cities and intelligent transportation systems (ITS) [3-5]. The work in [6] observes that vehicles can connect to base stations via Vehicle-to-Infrastructure (V2I) links in high-mobility vehicular networks to facilitate infotainment, traffic efficiency, and data-intensive applications like social networking, media streaming, and high definition (HD) map downloads, which require a lot of bandwidth. On the other hand, Vehicle-to-Vehicle (V2V) links concentrate on providing highly dependable, low-delay communication between adjacent vehicles, either regularly or in response to events, of safety-critical information, such as basic safety messages (BSM) in Dedicated Short-Range Communications (DSRC) [7]. Furthermore, autonomous driving and vehicle platooning systems are examples of vehicular communication applications that are regarded as crucial study fields that aid in the planning and administration of smart cities in ITSs. The mobility feature in these applications presents several challenges that have a significant impact on the reliability of communication, including high penetration loss, low latency carrier frequency offset, inter-cell interference (ICI), fast and frequent handovers, and fast time-varying wireless channels that experience multi-path fading in addition to large Doppler spread [8]. Estimating and monitoring wireless channel fluctuations is a crucial issue in this context because the receiver's follow-up equalisation, demodulation, and decoding processes depend on a correctly calculated channel response [9]. Therefore, the channel estimation process plays a critical role in overall system performance.

Statistical methods for channel estimation (CE) are often inefficient and resource-intensive for MIMO systems in 5G networks. To overcome this, the paper compares deep learning (DL)-based CE with advanced machine learning (ML) models, aiming to balance computational cost and generalisation. It focuses on DL models such as Recurrent Neural Network (RNN), Long-Short-Term-Memory (LSTM), and Bi-LSTM, and ML models like Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Random Forest (RF), using Kaggle data to study vehicular communications under varying mobility. LSTM, a form of RNN, addresses long-term dependencies using cell states and gating to mitigate vanishing gradients. Motivated by a 24.23% increase in traffic accidents in the fourth quarter of 2023 compared to the third quarter [10], the study also explores how real-time communication of vehicle dynamics, speed, acceleration, road conditions, traffic flow, and wireless environments can support informed decision-making, potentially reducing accidents and aiding national development.

Based on IEEE standards, vehicular communication networks are categorised into vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and hybrid V2I models. This study focuses on V2V communication, where vehicles exchange data directly. Conventional channel estimation (CE) methods fall into blind, non-blind, and semi-blind categories [11], as shown in Figure 1. Blind methods can be statistical, relying on signal properties like correlation and covariance, or deterministic, which use received signals and channel coefficients but face high complexity as modulation order increases [11]. Bayesian estimation and Kalman filtering are common statistical approaches [12]. Non-blind (data-aided) CE embeds known pilot symbols to estimate channel response, trading off spectral efficiency. Pilot-based methods like Least Squares (LS) and Minimum Mean Square Error (MMSE) are suitable for fast-varying channels [13]. Decision-directed CE (DDCE) iteratively uses detected symbols for estimation but struggles in fast-fading environments due to error propagation [11]. DDCE methods include hard iterative, applying discrete symbol decisions (e.g., QPSK, QAM) [14], and soft iterative, using probabilistic information, which offers robustness at the cost of higher computational demand. Figure 2 depicts the statistical and the pilot-assisted approaches of blind and non-blind CE techniques. Semi-blind techniques combine training data with blind estimation for continuous adaptation.

However, in dynamic vehicular environments, these traditional methods fall short, prompting this study's exploration of ML and DL for more robust and adaptive CE and signal detection.

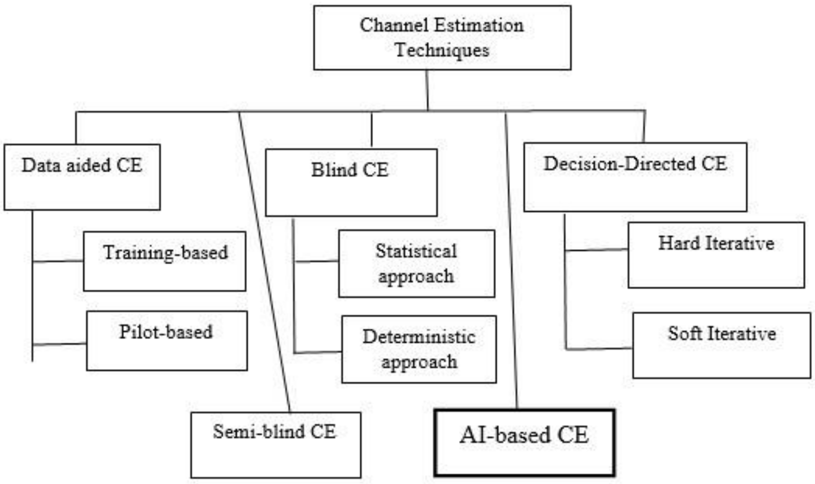
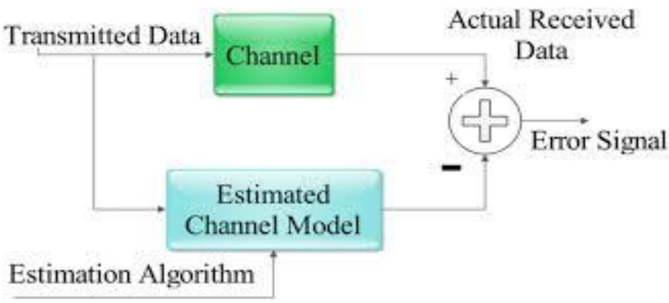
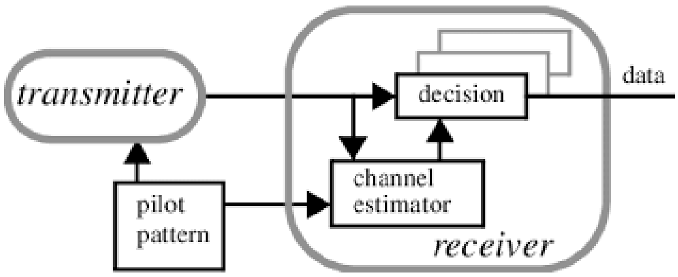



Figure 1. Channel Estimation Techniques




(a) Statistical CE approach




(b) Pilot-assisted CE approach

Figure 2. Blind vs non-blind CE Methods (a) blind-based, e.g., Statistical (b) non-blind, e.g., Pilot-assisted.

Traditional statistical and pilot-based channel estimation techniques are increasingly unsuitable for 5G and high-mobility vehicular MIMO systems due to their complexity, resource intensity, and poor tracking of fast-varying channels. Data-pilot aided methods degrade in performance under high mobility because of Doppler-induced errors. These approaches also suffer from pilot contamination, spectral inefficiency, and high overhead, limiting their effectiveness in time-varying, frequency-selective environments. In response, AI-driven machine and deep learning models have emerged as promising alternatives, capable of learning complex channel features like delay, phase, amplitude, and Doppler effects. This study explores RNN, LSTM, Bi-LSTM, SVM, RF, and XGBoost algorithms to determine the most effective model for channel estimation across varying mobility conditions, considering factors such as inter-carrier interference, path gain, and delay. The major contributions of this study include:

(i) Deployment of Bi-LSTM DL model to effectively capture temporal dependencies in the wireless channel, improving estimation accuracy under rapid channel variations and Doppler effects;
(ii) Comprehensive comparison to benchmark with traditional CE and ML techniques;
(iii) Simulation of real-world high-speed vehicular environments, enhancing the practical relevance of the model and results;
(iv) Empirical validation to ensure robustness, generalization, and performance stability across different mobility conditions (low, medium, high).

This study aims to develop an enhanced channel estimation framework for reliable, low-latency vehicular communications using AI-driven models. It focuses on addressing challenges in high-mobility V2V networks by applying ML and DL techniques to improve channel state information (CSI) estimation. The paper is structured as follows: Section 2 reviews existing CE methods and V2V communication challenges; Section 3 introduces the proposed AI-based CE framework for 5G MIMO systems, detailing key wireless parameters and the dataset used; Section 4 presents and visualizes results across different mobility scenarios, comparing traditional and AI-based models; and Section 5 concludes with recommendations for future research.

## 2. RELATED WORKS

### 2.1. Vehicular Communication Networks

The rise of vehicle-to-everything (V2X) communication is driven by increasing demand for smart infrastructure, autonomous vehicles, and advanced mobile devices [15]. V2X enables real-time communication among vehicles, infrastructure, and mobile apps to improve road safety, traffic flow, and fuel efficiency. Technologies such as sensors, cameras, and hardware security modules like the SLS37 enhance predictive maintenance and collision avoidance, while standards like SAE JS2735 optimize energy use and software performance. Mobile devices with high-speed processing and better interfaces are increasingly integrated into vehicles, supporting V2X functions such as GPS-based navigation and weather updates. V2I services include high-definition video, AR/VR, and maps [4], while V2V supports basic safety messages and sensor data exchange. Advances in AI and 5G are accelerating autonomous driving. Despite standards like DSRC [7] and ITS-G5 (IEEE 802.11p), limitations such as unbounded access delay, lack of QoS, and short V2I connections persist [6][16][17]. To overcome these, Third Generation Partnership Project (3GPP) has introduced Long Term Evolution (LTE) and 5G-based V2X solutions [18], and recent studies [19][20][21] focus on optimizing radio resources using refined analytical framework to model interference-driven call blocking and device-to-device (D2D) communication to support V2V in cellular networks. Still, designing reliable V2X networks remains challenging due to diverse QoS needs and high mobility.

Contemporary vehicles are now commonly outfitted with sensors, including engine control units, radar, light detection and ranging (LiDAR), and cameras, to enable real-time monitoring of vehicle performance and surrounding environments. With advanced onboard computing and storage, they function as intelligent hubs, continuously generating large volumes of mobile big data [22]. This data, covering vehicle dynamics, road conditions, traffic, and wireless environments, offers valuable context that can improve network performance through adaptive, data-driven decisions [5]. However, traditional communication methods are inadequate for fully utilizing this information.

**A. Kalman Filter Method for Channel Estimation**

The Kalman Filter, a recursive statistical method for channel estimation, is effective in tracking time-varying channels in dynamic, high-Doppler environments such as mobile or fading systems [23]. While it leverages statistical knowledge of noise and channel dynamics without storing full data history, it performs poorly with short data sequences and requires accurate channel models. It is also more computationally intensive than pilot-based methods like LS or MMSE. The wireless channel is modelled as a state-space model, where the state (channel) evolution expressed in equation (1) describes how the channel changes over time.

$$h(k) = Ah(k-1) + w(k) \tag{1}$$

where, $h(k)$ is the channel state at time $k$, $A$ is the state transition matrix (tracks changes in channel condition), while $w(k)$ represents the process noise (e.g., caused by mobility), modeled as zero-mean Gaussian. The observation (measurement) in equation (2) describes how the observed signal relates to the channel, where $y(k)$ is the received signal, $X(k)$ is the known transmitted symbols (could be pilots or data), and is the measurement noise.

$$y(k) = X(k)h(k) + v(k) \tag{2}$$

The Kalman Filter steps is described as follows:

For each time step $k$, the Kalman Filter performs prediction and update as follows:

*(i) Prediction Step:*

Estimate the current state and error covariance using equations (3) and (4):

$$\hat{h}^{-}(k) = A\hat{h}(k-1) \tag{3}$$
$$P^{-}(k) = AP(k-1)A^{T} + Q \tag{4}$$

where, $\hat{h}^{-}(k)$ denotes predicted channel state, $P^{-}(k)$ denotes predicted error covariance, and $Q$ is the process noise covariance.

*(ii) Update Step:*

Refine the prediction using the new measurement, following equations (5) to (7):

$$K(k) = P^{-}(k)X^{T}(k)\,[X(k)P^{-}(k)X^{T}(k) + R]^{-1} \tag{5}$$
$$\hat{h}(k) = \hat{h}^{-}(k) + K(k)\left[y(k) - X(k)\hat{h}^{-}(k)\right] \tag{6}$$
$$P(k) = [I - K(k)X(k)]\,P^{-}(k) \tag{7}$$

where, $K(k)$ denotes the Kalman Gain, $R$ represents the measurement noise covariance, and $\hat{h}(k)$ is the updated estimate of the channel.

**B. LS and MMSE Methods of Channel Estimation**

The Least Squares (LS) method is a simple, widely used channel estimation technique that minimizes the squared error between observed and predicted signals without relying on channel or noise statistics [24]. It is computationally efficient but highly sensitive to noise, resulting in poor accuracy under low SNR conditions. The formula is expressed in equation (8) as follows:

$$\hat{h}_{LS} = (X^H X)^{-1} X^H y \tag{8}$$

where, $X$ refers to the known pilot matrix, $y$ is the received signal vector, and $\hat{h}_{LS}$ represents the estimated channel. The MMSE method enhances LS by using prior knowledge of channel and noise statistics to minimize the mean square error, offering more accurate channel estimation, particularly in noisy conditions. It is denoted in equation (9) as:

$$\hat{h}_{MMSE} = R_{hX}\left(X^H R_{yy}^{-1}\right) y \tag{9}$$

where, $R_{hX}$ is the cross-covariance of the channel and input, and $R_{yy}$ is the autocovariance of the received signal.

Machine learning, particularly deep learning, has shown remarkable success in fields like computer vision, natural language processing (NLP), and robotics, enabling intelligent systems to operate in complex environments. ML identifies patterns in large datasets, offering a robust data-driven approach for analyzing wireless communication data and supporting informed decision-making [25-26]. It facilitates the fusion of advanced communication technologies with intelligent, adaptive, and context-aware network capabilities. However, applying ML to high-mobility vehicular networks remains challenging. Deep learning models can learn complex, non-linear channel behaviours without prior statistical models, generalize well to new conditions, and offer robust, low-latency, real-time estimation, outperforming traditional methods like LS and MMSE in noisy or dynamic environments.

**C. DL-based Channel Estimation Method**

Bi-LSTM, an advanced deep learning technique, is increasingly applied to channel estimation in dynamic wireless environments like mobile and vehicular systems. As a type of RNN, Bi-LSTM processes input in both directions and uses memory cells to capture long-term dependencies, making it ideal for modeling sequential, time-varying data. In systems like OFDM and MIMO, where channel conditions change due to mobility and environment, Bi-LSTM captures past and future signal context, enabling more accurate and robust channel state estimation without requiring prior statistical knowledge. Figures 3 and 4 illustrate the DL-based CE workflow and Bi-LSTM architecture described in 4-steps comprising data preparation, network architecture, training, and estimation phases, while Table 1 summarizes characteristics of key channel models. These models simulate real-world signal propagation and are essential for wireless system design and evaluation. Rayleigh fading models signal variation in non-LoS environments, commonly in urban settings. Clustered Delay Line (CDL), defined in 3GPP 38.901, simulates multipath components in various scenarios (CDL-A to CDL-E). Tapped Delay Line (TDL), a simpler model, represents time-dispersive effects using tap delays and gains (TDL-A to TDL-E). X3PTR 38.991 extends 3GPP models for 6–100 GHz mmWave frequencies, accounting for LoS, NLoS, reflection, diffraction, and blockage.

**Step 1: Data Preparation**

Input data: Pilot signals or received symbols across time.
Output: Estimated channel coefficients (real and imaginary parts).

**Step 2: Network Architecture**

Input Layer: Sequences of pilot symbols or received data.

Bi-LSTM Layer: Processes the sequence in both directions to extract temporal features.

Fully Connected (Dense) Layer: Maps the extracted features to the estimated channel coefficients.

Output Layer: Produces the predicted channel response at each time step.

**Step 3: Training**

The network is trained offline using simulated or measured datasets where the true channel response is known.

Loss function: The model's effectiveness is typically assessed by calculating the Mean Squared Error (MSE) between the actual and estimated channel values.

**Step 4: Estimation Phase**

Once trained, the Bi-LSTM model is used in real time to predict current channel coefficients based on past and (optionally) future observed signals.

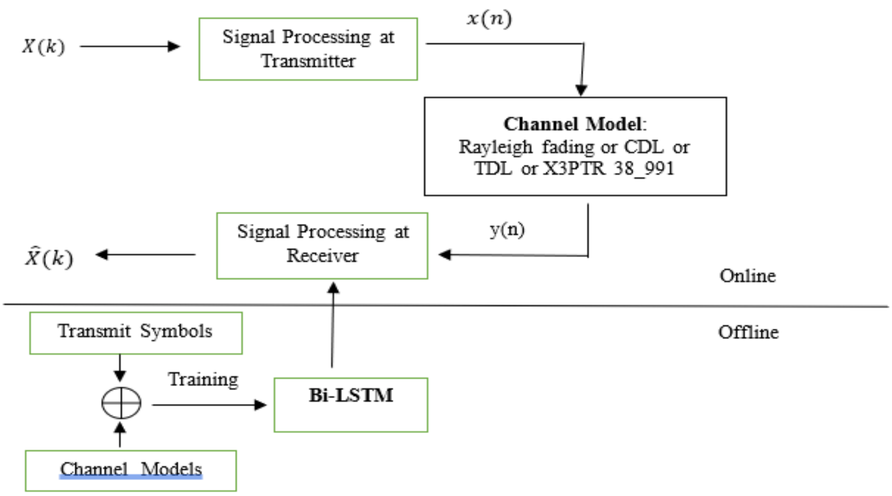



Figure 3. DL-based CE Approach

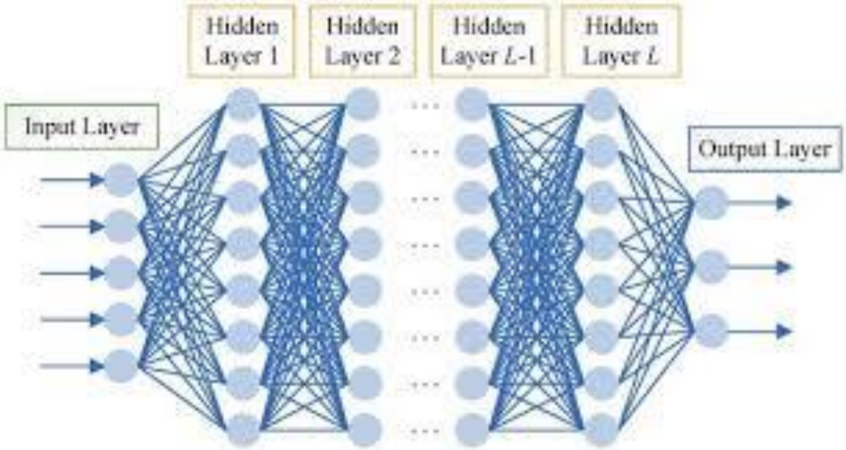



Figure 4. Layers of the Bi-LSTM CE Approach

Table 1. Key Features of Some Channel Models

| Model | Meaning | Type | Use Case |
|---|---|---|---|
| Rayleigh | Rayleigh Fading | Statistical | No LoS, urban multipath |
| CDL | Channel Delay Line | Geometric | Realistic 5G MIMO simulation |
| TDL | Tapped Delay Line | Delay-line | Simplified multipath simulation |
| X3PTR 38_991 | 3GPP TR 38_991 | Technical Specification | mmWave and wideband 5G modeling |

## 2.2. Vehicular Communication Technologies

Early vehicular communications were based on Wi-Fi, leading to the development of DSRC in the U.S. [7] and C-ITS in Europe [27], which evolved separately due to different research and stakeholder influences. More recently, cellular technologies like 4G LTE [28] and 5G [18] have gained prominence due to widespread infrastructure. This study adopts 4G LTE for its system model, highlighting its high data rate (150 Mbps), synchronous communication, and support for multimedia and cloud services, with potential migration to 5G.

## 2.3. Challenges of Vehicular Networks

Vehicular Ad-Hoc Networks (VANETs) face key challenges such as intermittent connectivity due to high mobility, packet loss, and the need for precise location awareness in dynamic traffic and emergency scenarios. The rise of heterogeneous smart vehicles adds complexity in managing diverse communication standards and sporadic connections. Security and privacy are critical, requiring local processing of sensitive data rather than cloud transmission. Additionally, VANETs depend on vehicle sensors and edge cloud computing for efficient data handling before reaching central servers [29][30]. As traditional methods fall short in such dynamic settings, AI-driven, particularly DL-based, estimators offer adaptive, low-complexity, and high-performance solutions. These enable enhanced reliability, support advanced services like HD streaming, AR/VR, and safety messaging, and benefit sectors such as transportation, logistics, and healthcare through energy-efficient, accurate, and intelligent communication systems.

# 3. METHODOLOGY

## 3.1. The Proposed System Architecture

The components of the proposed AI-based CSI estimator, shown in Figure 5, include. User Interface, CN+ Dataset, Data Pre-processing, Data Splitting, Model Development, Performance Evaluation, Channel State Prediction, and Results Visualization. The components are described as follows.

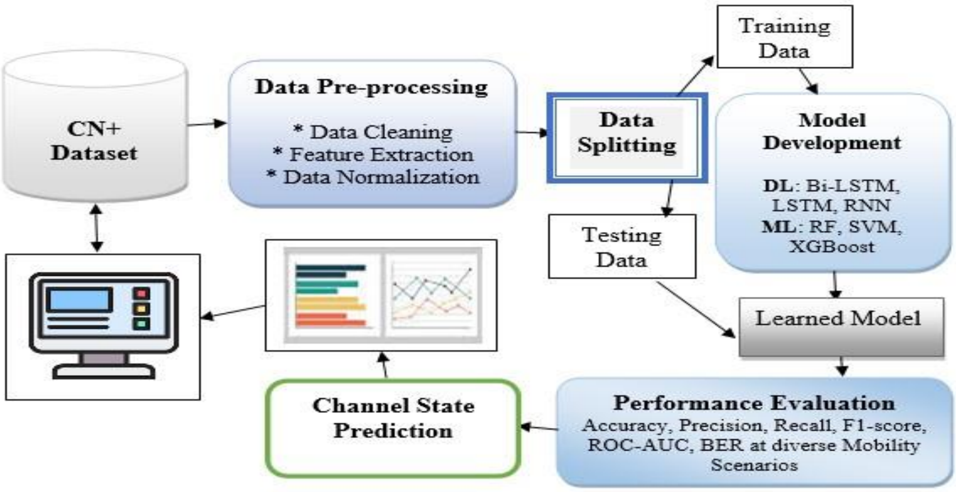



Figure 5: Proposed AI-based Channel Estimation Framework for V2V Communication

The User Interface (UI) enables users to interact with the system by inputting commands and receiving feedback in an organized manner. Vehicles are typically equipped with onboard units (OBUs) that wirelessly communicate with other vehicles (V2V) and Road Side Units (RSUs) for sharing critical data like position, speed, and traffic updates. Antennas improve signal strength and coverage, while RSUs serve as fixed nodes that enhance communication reliability and reduce interruptions. VANETs are designed to improve road safety by supporting V2V communication and advanced safety algorithms. Due to the high cost of real-world testing, simulations using real datasets are commonly used. This study uses the CN+ dataset, collected from over 25,000 vehicles over 32 hours in Bremen, Germany, which provides rich, well-labelled data that enhances the accuracy of AI-driven channel estimation. Dataset features include transmitted signals with sensor-based metrics such as channel gain, velocity, Doppler shift, path delay, and propagation speed.

*Channel Gain* $(g)$: rrepresents the degree to which a signal is amplified or weakened as it propagates from the transmitter to the receiver through a communication channel. It is calculated as the ratio of the received signal power ($P_{received}$) to the transmitted signal power ($P_{transmitted}$) as shown in equation (8). A higher gain means better signal strength at the receiver end, which is crucial for reliable communication.

$$Channel\ Gain\ (g) = \frac{P_{received}}{P_{transmitted}} \tag{8}$$

Velocity: indicates the rate at which a vehicle moves in a specific direction. It is a vector quantity measured in meters per second (m/s) and is calculated as:

$$Velocity = \frac{Distance}{Time} \tag{9}$$

Doppler Shift: describes the shift in a wave's frequency observed when there is relative motion between the wave source and the observer. It is used to estimate the relative velocity between vehicles.

Path Delay: The time it takes for a signal to travel from the transmitter to the receiver, calculated as:

$$Path\ Delay = \frac{Distance}{Propagation\ Speed} \tag{10}$$

Propagation Speed: refers to the rate at which an electromagnetic signal propagates through a medium, typically approaching the speed of light in air or a vacuum. It relates to distance and delay as:

$$Distance = Propagation\ Delay * Propagation\ Speed \tag{11}$$

In a V2V network, the Doppler shift significantly impacts wireless communication performance. Its relationship with the relative velocity between vehicles can be described mathematically as follows:

$$f_d = \frac{v_{rel}}{c} f_c \cos(\theta) \tag{12}$$

where, $f_d$ is the Doppler shift (Hz), $v_{rel}$ is the relative velocity between the vehicles (m/s), $C$ represents the speed of light ($3 \, x \, 10^8$ m/s), $f_c$ refers to the carrier frequency (Hz), and $\theta$ is the angle between the vehicle's direction of motion and the LOS connecting them. Equation (12) shows that the Doppler shift increases linearly with relative velocity. If vehicles are moving toward each other, the shift is said to be positive (frequency increases); otherwise, it is negative (frequency decreases) if moving away. The impact of $\theta$ is that the maximum Doppler shift (vehicles moving directly toward or away) is achieved when $\theta = 0^{\circ}$, and there is no Doppler shift (vehicles moving perpendicular to the line of sight) when $\theta = 90^{\circ}$. Ideally, the Doppler shift causes channel variations, which must be tracked for reliable communication. High speeds (e.g., highways) lead to fast-changing channels, requiring adaptive modulation and robust equalizers. Also, Doppler spread (range of Doppler shifts due to multipath) can cause inter-symbol interference (ISI).

## 3.2. Data Acquisition and Pre-processing

The CN+ dataset used in this study comprises data from over 25,000 vehicles gathered over 32 hours at a signalized intersection in Bremen, Germany [31]. It offers rich, well-labelled data that improves machine learning model accuracy in channel estimation. Figure 6 illustrates raw data samples with attributes like Timestamp, Network Mode, Signal Strength, User ID, Velocity, Doppler Shift, Serving Cell Distance, Average Path Gain, Path Delay, Channel State, and Channel State Information. Data preprocessing involved cleaning the dataset, removing noisy entries, duplicates, and irrelevant features using Excel's Conditional Formatting, yielding 18,700 entries with six relevant features. Feature scaling was then performed using min-max normalization (Equation 13) to standardize values for neural network stability. Categorical and text data were transformed using integer or one-hot encoding, with text tokenized and padded for uniform input length. The final dataset was split into training and testing sets using an 8:2 ratio with stratified sampling to preserve class distribution. This careful stratification helps prevent misleading model performance and overfitting to the majority classes. A sample of the normalized dataset of selected input features is presented in Figure 7. The min-max normalization approach converts a value $x$, to a new value $x_{normalized}$, where $x_{min}$ is the minimum and $x_{max}$ is the maximum value in the dataset as follows:

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{13}$$

| Timestamp | NetworkM | Signal Strength (dBm) | User ID | Velocity ( | Doppler shift | ServingCe | Distance | Average path G | Path Delays | CS | CSI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 22/03/01 17:46 | LTE | -84.1197631 | 25 | 51.19941 | 71.95269234 | 551.37 | 3.487668 | -7.535544749 | 96.97163398 | 30 | 1 |
| 22/04/01 17:29 | LTE | -87.80613083 | 22 | 114.5786 | 169.39972 | 551.37 | 0.801152 | -16.51470811 | 96.73741832 | 60 | 0 |
| 22/05/01 17:14 | LTE | -116.5751756 | 5 | 90.51933 | 65.57774373 | 553.43 | 2.216071 | -14.62874924 | 32.80597003 | 23 | 1 |
| 22/06/01 16:02 | LTE | -82.9611829 | 33 | 75.85243 | 140.3889649 | 553.43 | 6.316539 | -12.55840462 | 77.84540618 | 64 | 0 |
| 22/07/01 22:19 | LTE | -85.01257411 | 8 | 27.16205 | 9.021801257 | 563.48 | 0.379942 | -18.88201059 | 12.80536051 | 65 | 1 |
| 22/08/01 22:31 | LTE | -98.61869164 | 42 | 27.1594 | 77.11257602 | 563.48 | 5.963263 | -17.83168567 | 78.27235572 | 38 | 0 |
| 22/09/01 20:39 | LTE | -74.13939862 | 74 | 16.3892 | 68.64603499 | 563.48 | 9.572819 | -4.764914112 | 29.34686136 | 14 | 1 |
| 22/10/01 22:03 | LTE | -50.23533096 | 16 | 105.2794 | 163.3357479 | 563.48 | 7.971451 | -4.262520608 | 29.62431312 | 53 | 1 |
| 22/11/01 10:15 | LTE | -70.1030734 | 7 | 76.12265 | 193.4110806 | 545.02 | 3.521411 | -0.235307773 | 51.21609764 | 50 | 1 |
| 22/12/01 13:31 | LTE | -97.60056369 | 74 | 87.88798 | 92.17625534 | 545.02 | 1.468256 | -18.64807229 | 39.84056721 | 65 | 1 |
| 13/01/2022 15:20 | LTE | -99.65173579 | 62 | 12.26429 | 164.9672554 | 545.02 | 4.704138 | -2.216311768 | 12.66528315 | 46 | 1 |
| 14/01/2022 19:10 | LTE | -78.58596219 | 54 | 116.6901 | 139.9389402 | 545.02 | 3.990209 | -13.65567909 | 59.71463715 | 30 | 0 |
| 15/01/2022 9:16 | LTE | -105.4742054 | 29 | 101.5687 | 193.4771975 | 545.02 | 9.720156 | -13.44298829 | 22.17051733 | 48 | 0 |
| 16/01/2022 23:40 | LTE | -77.20348668 | 87 | 33.3573 | 162.2217029 | 545.02 | 9.899694 | -17.45435496 | 45.00349835 | 25 | 0 |
| 17/01/2022 12:50 | LTE | -87.95391151 | 17 | 30.00075 | 15.49561319 | 545.02 | 4.164788 | -0.273996461 | 23.87005829 | 71 | 1 |
| 18/01/2022 16:55 | LTE | -82.29876478 | 80 | 30.1745 | 142.6563445 | 545.02 | 1.572961 | -0.380215262 | 85.90093892 | 49 | 0 |
| 19/01/2022 22:10 | LTE | -105.5434978 | 50 | 43.46665 | 107.0999431 | 545.02 | 9.115035 | -2.703407825 | 34.46414605 | 57 | 0 |
| 20/01/2022 16:09 | LTE | -65.74131849 | 91 | 67.72321 | 3.287211586 | 545.02 | 7.956069 | -18.18439827 | 88.86159169 | 63 | 0 |
| 21/01/2022 10:22 | LTE | -93.13557136 | 33 | 57.51395 | 50.26259175 | 545.02 | 5.161615 | -3.167007486 | 33.28362711 | 50 | 1 |
| 22/01/2022 13:27 | LTE | -77.06992052 | 29 | 42.03521 | 184.351049 | 545.02 | 3.717011 | -7.307720137 | 18.77021013 | 61 | 1 |
| 23/01/2022 21:14 | LTE | -79.61500354 | 7 | 77.30382 | 185.2122746 | 545.02 | 4.536842 | -17.42449108 | 58.08682963 | 30 | 0 |
| 24/01/2022 14:07 | LTE | -103.4370333 | 19 | 25.34432 | 159.770092 | 545.02 | 9.028624 | -14.76484486 | 55.95863985 | 59 | 0 |

Figure 6: Sample of Raw Dataset

| | Velocity | Doppler Shift | Distance | Average Path Gain | Path Delays |
|---|---|---|---|---|---|
| 0 | 0.374561 | 0.359796 | 0.342187 | 0.623257 | 0.969415 |
| 1 | 0.950785 | 0.847082 | 0.070810 | 0.174268 | 0.967049 |
| 2 | 0.732046 | 0.327917 | 0.213737 | 0.268573 | 0.321206 |
| 3 | 0.598699 | 0.702013 | 0.627944 | 0.372097 | 0.776200 |
| 4 | 0.156021 | 0.045108 | 0.028261 | 0.055895 | 0.119158 |
| ... | ... | ... | ... | ... | ... |
| 18695 | 0.397838 | 0.507069 | 0.813042 | 0.878280 | 0.112583 |
| 18696 | 0.861973 | 0.408525 | 0.365930 | 0.603810 | 0.873483 |
| 18697 | 0.330869 | 0.930993 | 0.154000 | 0.556444 | 0.606489 |
| 18698 | 0.259377 | 0.089539 | 0.162131 | 0.504711 | 0.935092 |
| 18699 | 0.743368 | 0.615629 | 0.983625 | 0.215316 | 0.886861 |

18700 rows × 5 columns

Figure 7: Sample of Scaled Dataset

### 3.3. Model Development and Hyperparameters Tuning

The model development involved creating a custom deep learning architecture using a Bidirectional Long Short-Term Memory (Bi-LSTM) network, a type of recurrent neural network (RNN) well-suited for vehicular channel estimation. Comparative experiments were also performed using RNN, LSTM, SVM, Random Forest (RF), and XGBoost models. Python was selected for implementation due to its simplicity, active community, and robust libraries such as PyTorch, Keras (with TensorFlow backend), and scikit-learn. Hyperparameter tuning was carried out using 5-fold cross-validation, testing various learning rates, batch sizes, architectures, and regularization techniques. Model performance was closely monitored for instability, overfitting, or underfitting, and configurations were repeated with different random seeds to ensure robustness. Once optimal performance was achieved, the final model was fixed for experimentation. The hyperparameters used are detailed in Table 2.

Table 2: Hyperparameter Tuning with 5-fold Cross-Validation

| Algorithm | Hyperparameter with 5-fold cross-validation |
|---|---|
| **XGBoost** | n_estimator: [50, 100, 200], learning rate: [0.01, 0.1, 0.2], subsample: [0.8, 1.0], colsample_bytree: [0.8, 1.0] |
| **Random Forest** | n_estimator: [100, 200, 300], max_depth: [None, 10, 20], min_sample_split: [2, 5, 10] |
| **Support Vector Classifier** | C: [0.1, 1, 10], kernel: [‘linear’, ‘rbf’, ‘poly’], gamma: [‘scale’, ‘auto’] |

## 4. DISCUSSION OF RESULTS

### 4.1. Modelling Results of ML and DL-based Channel Estimators

Results of the ML and DL estimators are shown in Tables 3 and 4. Table 3 shows that for channel state estimation in V2V communication, the RF algorithm achieves the best performance among the ML models, with the highest accuracy (99.94%), recall (100%), and F1-score (99.94%). This was closely followed by XGBoost with accuracy (99.81%), recall (99.68%), and F1-score (99.81%), while SVM had the least performance in terms of accuracy (95.08%), recall (95.35%), and F1-score (95.15%). However, XGBoost achieved superior performance in precision (99.95%) and ROC-AUC (100%), compared to RF’s precision of 99.89% and ROC-AUC of 99.94%. Figures 8 and 9 represent the graphical results demonstrated by the different ML and DL models.

Table 3: Results of ML Models

| ML Models | Accuracy | Precision | Recall | F1-Score | ROC-AUC |
|---|---|---|---|---|---|
| XGBoost | 0.9981 | 0.9995 | 0.9968 | 0.9981 | 1.0000 |
| RF | 0.9994 | 0.9989 | 1.0000 | 0.9994 | 0.9994 |
| SVM | 0.9508 | 0.9495 | 0.9535 | 0.9515 | 0.95076 |

**Table 4: Results of DL Models**

| ML Models | Accuracy | Precision | Recall | F1-Score | ROC-AUC |
|---|---|---|---|---|---|
| Bi-LSTM | 0.9858 | 0.9868 | 0.9852 | 0.986 | 0.9984 |
| RNN | 0.9832 | 0.9867 | 0.9799 | 0.9833 | 0.999 |
| LSTM | 0.9818 | 0.9882 | 0.9757 | 0.9819 | 0.9988 |

Similarly, results from Table 4 indicates that Bi-LSTM outperformed other DL models with highest accuracy (98.58%), recall (98.52%), and F1-score (98.6%), closely followed by RNN with accuracy (98.32%), recall (97.99%), and F1-score (98.33%) while LSTM with accuracy (98.18%), recall (97.57%), and F1-score (98.19%) had the least performance. However, LSTM yielded highest precision (98.82%) while RNN yielded highest ROC-AUC result (99.9%). Despite these differences, all the models achieved satisfactory level of predictive performance in terms of channel state estimation. Table 5 summarizes the overall performance of all DL and ML models, while Figure 10 illustrates this performance graphically.

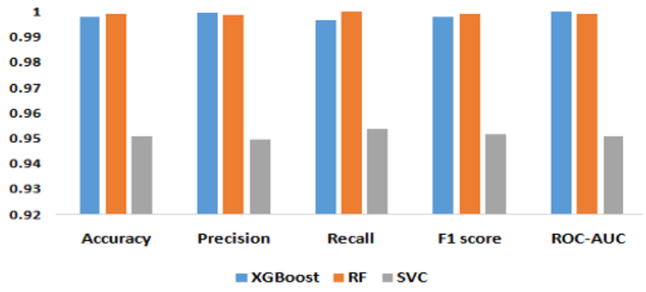



Figure 8: Graphical Performance of ML Models

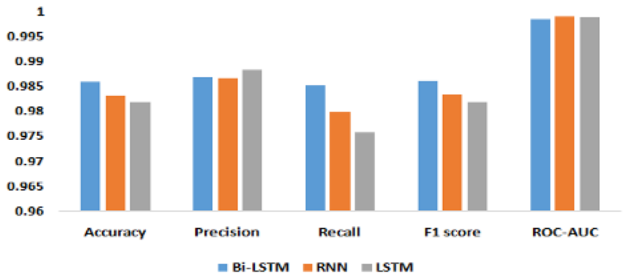



Figure 9: Graphical Performance of DL Models

Table 5: Results of ML and DL algorithm estimations

| Model | Accuracy | Precision | Recall | F1-score | ROC-AUC |
|---|---|---|---|---|---|
| Bi-LSTM | 0.9858 | 0.9868 | 0.9852 | 0.986 | 0.9984 |
| RNN | 0.9832 | 0.9867 | 0.9799 | 0.9833 | 0.999 |
| LSTM | 0.9818 | 0.9882 | 0.9757 | 0.9819 | 0.9988 |
| XGBoost | 0.9981 | 0.9995 | 0.9968 | 0.9981 | 1.0000 |
| RF | 0.9994 | 0.9989 | 1.0000 | 0.9994 | 0.9994 |
| SVC | 0.9508 | 0.9495 | 0.9535 | 0.9515 | 0.95076 |

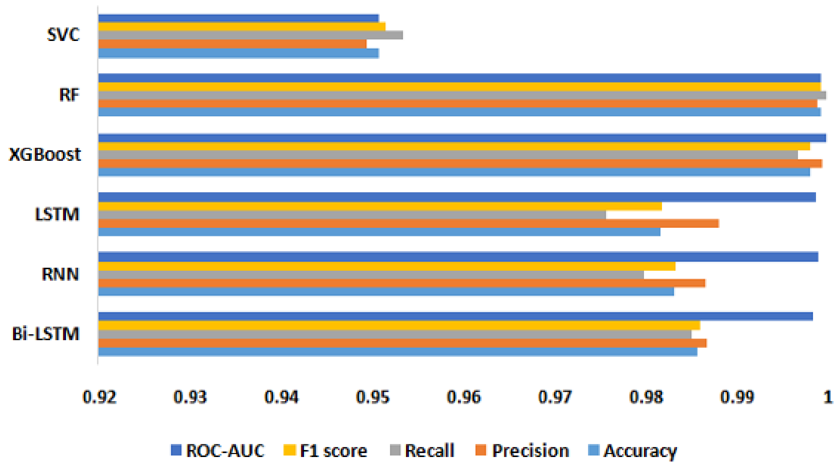



Figure 10: Overall Performance of all the models

From Figure 10, ML tends to perform better than DL algorithms, as seen in RF and XGBoost classifiers. However, it is crucial to note that the effectiveness of any machine or deep learning algorithm depends on the quality of input data, parameter settings, and its capacity to manage class imbalance and prevent overfitting. Figures 11 and 12 show the overall accuracy and F1-score plots for the respective models, where ML algorithms yield higher predictive accuracy and F1-score performance compared to the DL models.

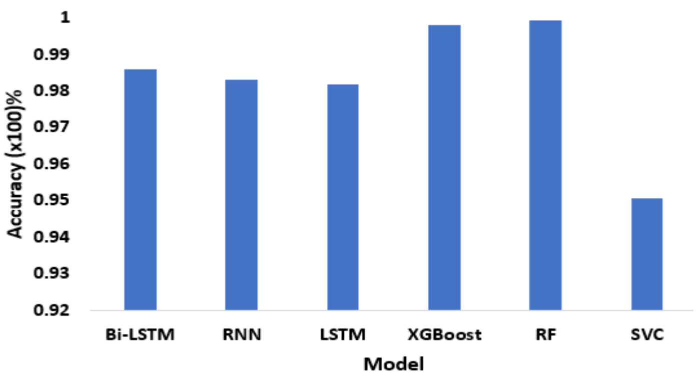



Figure 11: Overall Accuracy plot for the models

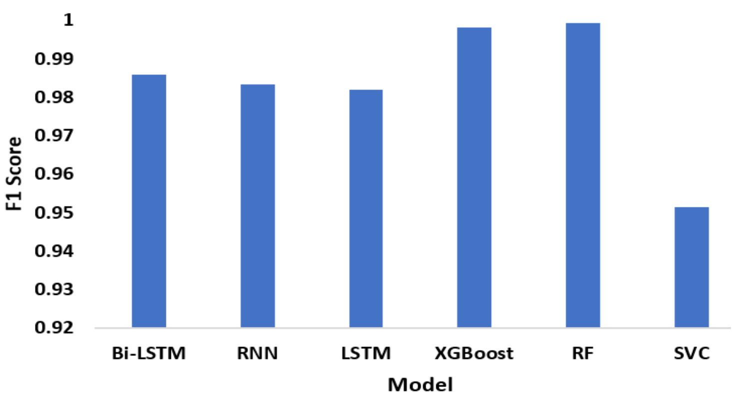



Figure 12: Overall F1-score plot for the models

### 4.2. Simulation Results on Different Mobility Scenarios

Python software and its useful libraries were used to simulate mobility scenarios for low, moderate, and high-speed vehicular communications, comparing the performance of traditional methods (LS, MMSE) with ML-based (RF) and DL-based (Bi-LSTM) methods. Table 6 presents the speed category, use case, and range of values for relative speed and Doppler shift, while Table 7 shows the simulation parameters for the experiment.

Table 6. Speed, Range, and corresponding Doppler shift

| Speed category | Use case | Relative speed ($v_{rel}$) | Doppler shift ($f_d$) |
|---|---|---|---|
| Low | Urban traffic, intersections, and parking scenarios | 0-30 km/hr (0-8.3 m/s) | 0-163 Hz |
| Moderate | City driving, rural roads | 30-90 km/hr (8.3-25 m/s) | 163-492 Hz |
| High | Highways, Emergency Vehicle Response | 90-150+ km/hr (25-42 m/s) | 492-826 Hz |

Table 7. Simulation parameters and their values

| Parameter | Value |
|---|---|
| Carrier frequency, $f_c$ | 5.9 GHz |
| Number of data subcarriers | 48 |
| Number of pilots | 4 |
| Signal modulation | QPSK |
| Channel model | Rayleigh Model, IEEE 802.11p |
| Bandwidth | 10 MHz |
| SNR range | 0-20 dB |
| Doppler shift tested, $f_d$ | 0-1000 Hz |

The Rayleigh fading and IEEE 802.11p channel models were used with a carrier frequency $f_c$ of 5.9 GHz, 4 pilot symbols, and 48 data subcarriers along with QPSK modulation scheme. These parameters were used to estimate channel and calculate Bit Error Rate (BER) vs. SNR as well as channel estimation error vs. Doppler. Results are shown in Figures 13-16. Figure 13 indicates that higher Doppler shifts (e.g., 800–1000 Hz) degrade performance, especially at lower SNR. Also, BER improves with SNR for all Doppler shifts, but the gap widens as Doppler increases. The analysis presented in Figures 14–16 demonstrates the impact of mobility on channel estimation performance across different methods. At low mobility (0–30 km/h), shown in Figure 14, all methods perform well, but Bi-LSTM and MMSE outperform LS due to their learning- and SNR-aware capabilities, while RF also exceeds LS but lags behind Bi-LSTM. As mobility increases to moderate levels (30–90 km/h), as shown in Figure 15, BER rises and fading becomes more pronounced, especially at low SNR. LS performs poorly under such conditions, while MMSE remains resilient due to its noise modelling capability. Bi-LSTM achieves the best results by effectively capturing temporal channel variations, whereas RF performs moderately but lacks temporal adaptability. At high mobility (90–150+ km/h), shown in Figure 16, performance declines sharply due to rapid fading and Doppler effects. Bi-LSTM consistently outperforms all methods by learning sequential dependencies, while MMSE performs acceptably but remains static. RF struggles to generalize in such dynamic conditions, and LS shows the worst performance, with high BER even at elevated SNR levels.

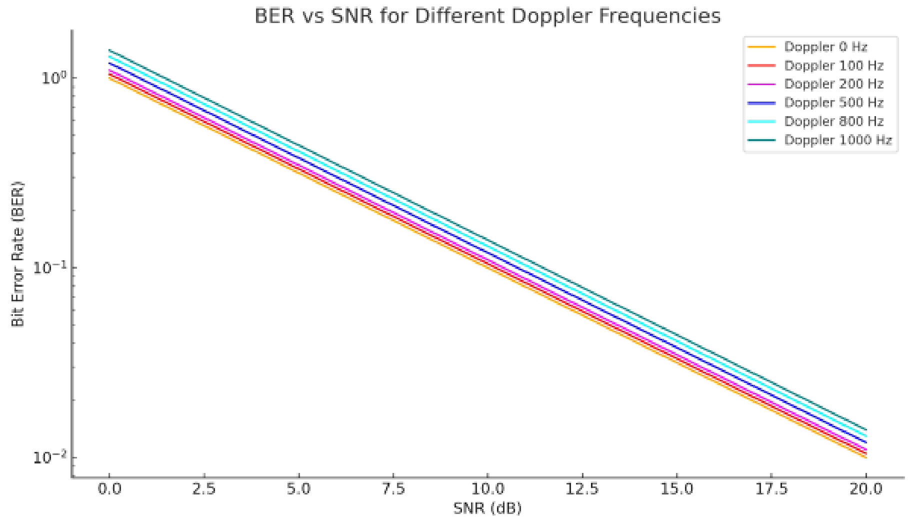



Figure 13. BER vs. SNR for different Doppler shifts.

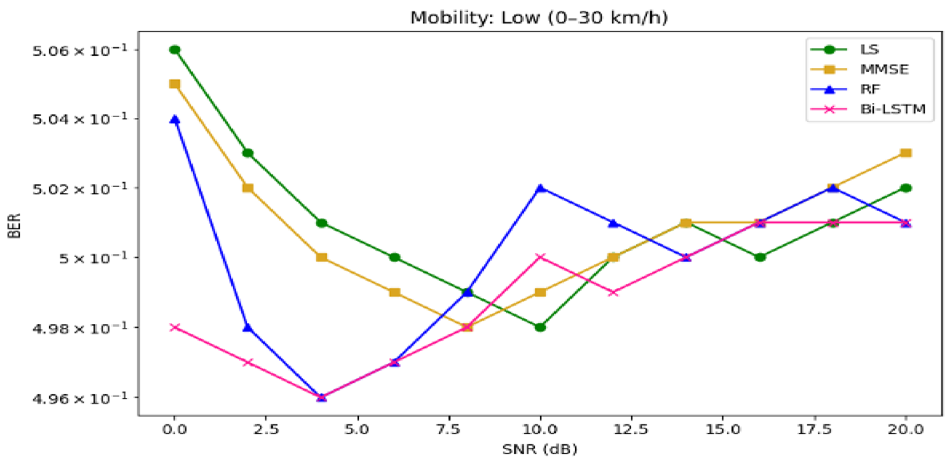



Figure 14: BER vs. SNR for low mobility scenario

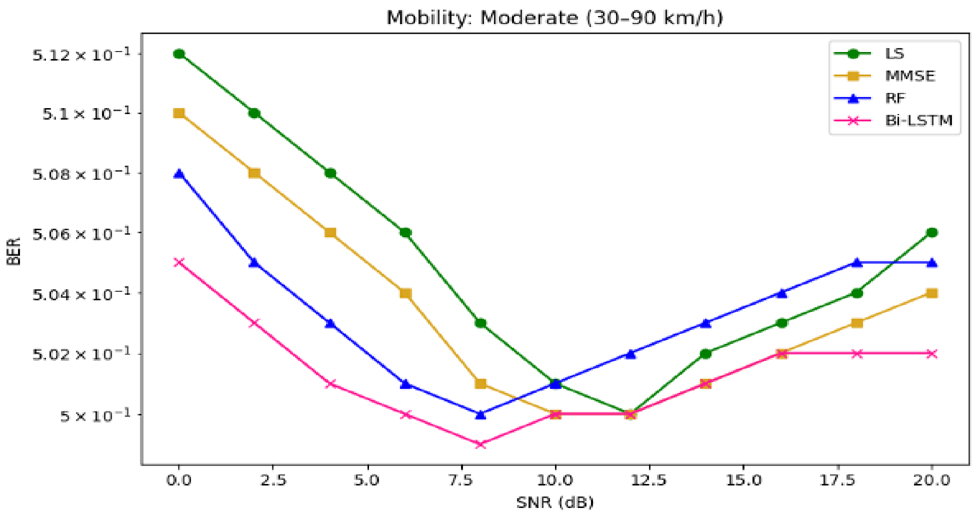


Figure 15: BER vs. SNR for Moderate mobility scenario

.

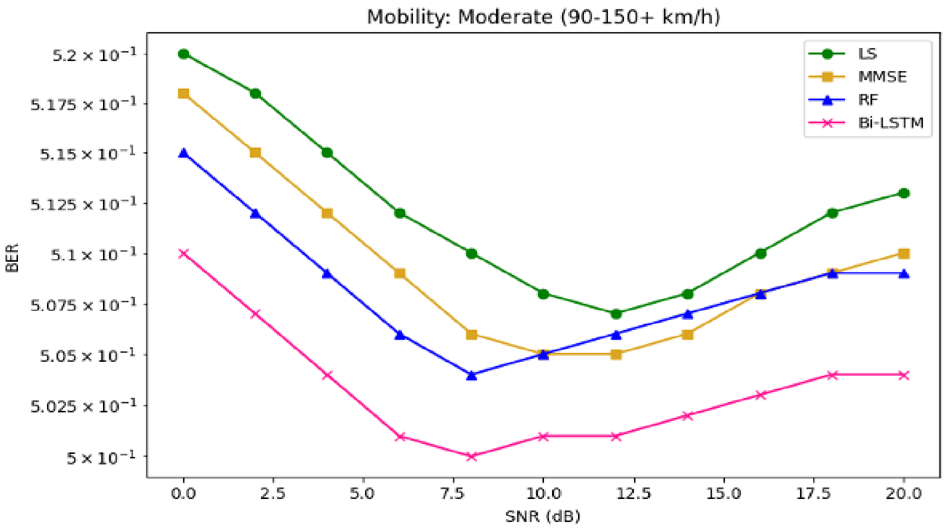



Figure 16. BER vs. SNR for high mobility scenario

## 5. Conclusion

Vehicular communication is key to enabling smart cities and autonomous transport, but high mobility poses challenges for real-time and accurate channel state estimation. This study evaluates a Bi-LSTM deep learning model against traditional and ML-based methods using real-world vehicular data. Results show that Bi-LSTM outperforms in accuracy, recall, and F1-score, especially in capturing the temporal dynamics of V2V channels. While RF and XGBoost also deliver high predictive performance, they lack Bi-LSTM's adaptability to time-varying conditions. Bi-LSTM surpasses conventional estimators like LS and MMSE under high Doppler shifts and dynamic channels, demonstrating greater resilience. The integration of intelligent channel estimation promises safer roads, improved traffic flow, and support for advanced applications like HD map downloads and autonomous vehicle platooning. Future directions include exploring edge computing, hybrid ML/DL models, and real-time deployment for ultra-reliable low-latency V2V communication.

## Conflict of Interest

The authors declare no conflict of interest among them.

## References

[1] Gizzini, Abdul Karim (2021). Advanced linear and deep learning-based channel estimation techniques in doubly dispersive environments (Doctoral Thesis, Cergy Paris CY Université).

[2] Liang, Le, Peng, Haixia, Li, Geoffrey Ye & Shen, Xuemin (2017a). Vehicular communications: A physical layer perspective. IEEE Transactions on Vehicular Technology, Vol. 66, No. 12, pp. 10647-10659.

[3] Cheng, Xiang, Yang, Liuqing & Shen, Xia (2015). D2D for intelligent transportation systems: A feasibility study. IEEE Transactions on Intelligent Transportation Systems, Vol. 16, No. 4, pp. 1784-1793.

[4] Liang, Le, Ye, Hao, & Li, Geoffrey Ye (2018). Toward intelligent vehicular networks: a machine learning framework. IEEE Internet of Things Journal, Vol. 6, No. 1, pp. 124-135.

[5] Asuquo, Daniel, Udo, Ifiok James, Ekpenyong, Moses & Attai, Kingsley (2024a, July). AI-Enabled IoT Framework for Smart Traffic Surveillance and Communication. In International Conference on Intelligent and Fuzzy Systems (pp. 193-202). Cham: Springer Nature Switzerland.

[6] Araniti, Giuseppe, Campolo, Claudia, Condoluci, Massimo, Iera, Antonio & Molinaro, Antonella (2013). LTE for vehicular networking: a survey. IEEE communications magazine, Vol. 51, No. 5, pp. 148-157.

[7] Kenney, John (2011). Dedicated short-range communications (DSRC) standards in the United States. Proceedings of the IEEE, Vol. 99, No. 7, pp. 1162-1182.

[8] Chafii, Marwa, Palicot, Jacques, Gribonval, Rémi & Bader, Faouzi (2018). Adaptive wavelet packet modulation. IEEE Transactions on Communications, Vol. 66, No. 7, pp. 2947-2957.

[9] Bomfin, Roberto, Chafii, Marwa, Nimr, Ahmad, and Fettweis, Gerhard (2021). A robust baseband transceiver design for doubly-dispersive channels. IEEE Transactions on Wireless Communications, Vol. 20, No. 8, pp. 4781-4796.

[10] Atubi, Augustus Orowhigo. (2024). "Trend Analysis of Reported Number of Injured From Road Traffic Crashes in Lagos State, South Western Nigeria." Social Sciences and Education Research Review 11, No. 2, pp. 63-71.

[11] Vilas Boas, Evandro C., Silva, Jefferson DS e, de Figueiredo, Felipe AP, Mendes, Luciano L. & de Souza, Rausley AA. (2022). Artificial intelligence for channel estimation in multicarrier systems for B5G/6G communications: a survey. EURASIP Journal on Wireless Communications and Networking, No. 1., p.116.

[12] El Husseini, Ali Houssam, Laurent Ros, and Eric Pierre Simon. (2019). Kalman filter-based channel estimation for mobile-to-mobile and relay networks. IEEE Signal Processing Letters, Vol. 26, No. 5, pp. 680-684.

[13] Ahmed, Ali Saadoon, Mustafa Maad Hamdi, Mohammed Salah Abood, Arshad Mohammed Khaleel, Maha Fathy & Shihab Hamad Khaleefah (2022). Channel estimation using LS and MMSE channel estimation techniques for MIMO-OFDM systems. In 2022 IEEE International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1-6.

[14] Asuquo, Daniel Ekpenyong, Umoren, Imeh Jonah & Effiang, Daniel. (2021). Received Signal Quality Evaluation in Digital Transmission Networks Based on Bit Error Rate Performance, Journal of the Nigerian Association of Mathematical Physics, Vol. 59, No. 1, pp. 165-178.

[15] [15] Naeem, Awad Bin, Abdul Majid Soomro, Hafiz Muhammad Saim, and Hassaan Malik (2024). Smart road management system for prioritized autonomous vehicles under vehicle-to-everything (V2X) communication. Multimedia Tools and Applications Vol. 83, No. 14, pp. 41637-41654.

[16] Asuquo, Daniel Ekpenyong, Williams, Edem & Nwachukwu, Enoch Okechukwu (2014). A Survey of Call Admission Control Schemes in Wireless Cellular Networks, International Journal of Scientific and Engineering Research, Vol. 5, Issue 2, pp. 111-120.

[17] Hassan, Md Imrul, Hai L. Vu & Taka Sakurai (2011). Performance analysis of the IEEE 802.11 MAC protocol for DSRC safety applications. IEEE Transactions on vehicular technology, Vol. 60, No. 8, pp. 3882-3896.

[18] Chen, Shanzhi, Jinling Hu, Yan Shi, Ying Peng, Jiayi Fang, Rui Zhao & Li Zhao (2017). Vehicle-to-everything (V2X) services supported by LTE-based systems and 5G. IEEE communications standards magazine, Vol. 1, No. 2, pp. 70-76.

[19] Liang, Le, Geoffrey Ye Li, and Wei Xu. (2017b). Resource allocation for D2D-enabled vehicular communications. IEEE Transactions on Communications, Vol. 65, No. 7, pp. 3186-3197.

[20] Shin, Cheolkyu, Emad Farag, Hyunseok Ryu, Miao Zhou & Younsun Kim (2023). Vehicle-to-everything (V2X) evolution from 4G to 5G in 3GPP: Focusing on resource allocation aspects. IEEE Access, Vol. 11, pp. 18689-18703.

[21] Williams, Edem Efiong. & Asuquo, Daniel. Ekpenyong (2014). An efficient model for reducing soft blocking probability in wireless cellular networks, International Journal of Wireless and Mobile Networks, Volume 6, Issue 4, pp. 85-99.

[22] Cheng, Xiang, Luoyang Fang, Xuemin Hong, and Liuqing Yang. (2017b). Exploiting mobile big data: Sources, features, and applications. IEEE Network, Vol. 31, No. 1, pp. 72-79.

[23] Cheng, Yan, Shengkang Zhang, Xueyun Wang, Haifeng Wang, and Huijun Yang (2024). Kalman Filter with Adaptive Covariance Estimation for Carrier tracking under weak signals and dynamic conditions. Electronics, Vol. 13, No. 7, p. 1288.

[24] Zheng, Zhi, Caiyong Hao & Xuemin Yang (2016). Least squares channel estimation with noise suppression for OFDM systems. Electronics Letters, Vol. 52, No. 1, pp. 37-39.

[25] Asuquo, Daniel Ekpenyong, Umoh, Uduak Augustine, Robinson, Samuel Akpan, Dan, Emmanuel Ambrose, Udoh, Samuel Sunday, and Attai, Kingsley Friday. (2024b). Hybrid intelligent system for channel allocation and packet transmission in CR-IoT networks. International Journal of Hybrid Intelligent Systems, Vol. 20, No. 2, pp. 101-117.

[26] Asuquo, Daniel, Ekong, Uyinomen and Ohamuche, Faith (2024c). Machine Learning Models for Detection of DDoS Attack in a Network Environment, In Proceedings of NCS International Conference on Technological Solutions for Smart Economy (SmartEco 2024), A. Onoshoga et al. (eds.), Volume 35, pp. 140-146.

[27] Sjoberg, Katrin, Peter Andres, Teodor Buburuzan, and Achim Brakemeier (2017). Cooperative intelligent transport systems in Europe: Current deployment status and outlook. IEEE Vehicular Technology Magazine, Vol. 12, No. 2, pp. 89-97.

[28] Xu, Zhigang, Xiaochi Li, Xiangmo Zhao, Michael H. Zhang & Zhongren Wang (2017). DSRC versus 4G-LTE for connected vehicle applications: a study on field experiments of vehicular communication performance. Journal of advanced transportation, No. 1, p. 2750452.

[29] Shrestha, Rakesh, Rojeena Bajracharya & Seung Yeob Nam (2018). Challenges of future VANET and cloud-based approaches. Wireless Communications and Mobile Computing, No. 1, p. 5603518.

[30] Abduljalil, Fekri M. (2023). Toward a generic vehicular cloud network architecture: a case of virtual vehicle as a service. International Journal of Wireless and Mobile Networks, Vol. 15, No. 5, pp.25-38.

[31] Karunathilake, Thenuka, Meyo Zongo, Dinithi Amarawardana & Anna Förster (2024). CN+: vehicular dataset at traffic light regulated intersection in Bremen, Germany. Scientific data, Vol. 11, No. 1, p. 665.

## AUTHORS

**Daniel Asuquo** is a PhD holder and Head, Department of Information Systems, University of Uyo. Renowned for his work in Intelligent Communication Networks and Information Systems, he has received multiple awards, including the Outstanding Paper Award at MCCSIS 2023 and Best Paper Presentation at NCS ISDE 2021. He has led TETFund research on smart traffic control and wireless networks optimization. With over 60 publications, his research spans cognitive radio networks, machine learning in healthcare, and IoT systems.

**Moses Ekpenyong** is a Professor of Computer Science and Director of the Centre for Research and Development at the University of Uyo, Nigeria, where he also leads the Computational Intelligence Group. His research focuses on speech and wireless communications, bioinformatics, and AI in healthcare. With 100+ publications, his work includes tone-language speech tech, sensor networks, and ontology-based systems.

**Enefiok Etuk**, Lecturer at Michael Okpara University, specializes in AI, machine learning, and algorithms. He holds a PhD in Computer Science and is under assessment for promotion to Senior Lecturer. His research interests include legal-tech, deep learning, and innovation forecasting, with early-career publications emphasizing sustainable technology and intelligent computing.

**Kingsley Attai**, Lecturer at Ritman University and PhD candidate, focuses on machine learning, explainable AI (XAI), health informatics, and information systems. His work emphasizes interpretable predictive models and mobile health diagnostic tools for LMICs. He has earned academic recognition for bridging AI innovation with real-world health delivery.

**Philip Asuquo**, Associate Professor at the University of Uyo, researches web systems, wireless sensor network security, IoT, and enterprise modelling. He develops secure and scalable systems for institutional use, with notable work in e-procurement, mail systems, and cybersecurity protocols in wireless networks.

**Temple Nwaeboiwe** is an MSc student at the University of Uyo and Technologist I at Ogbonnaya Onu Polytechnic, Aba. His research interest includes computer networks, cybersecurity, machine learning, and data analytics.