



Research article

ABCNet: A comprehensive highway visibility prediction model based on attention, Bi-LSTM and CNN

Wen Li^{1,2,3}, Xuekun Yang², Guowu Yuan^{1,3} and Dan Xu^{1,*}

¹ School of Information Science and Engineering, Yunnan University, Kunming 650504, China

² National Engineering Laboratory for Surface Transportation Weather Impacts Prevention, Broadvision Engineering Consultants Co., Ltd., Kunming 650299, China

³ Yunnan Key Laboratory of Digital Communications, Kunming 650504, China

* **Correspondence:** Email: danxu@ynu.edu.cn; Tel: +86871605033748.

Abstract: Meteorological disasters along highways significantly reduce road traffic efficiency. Low visibility caused by heavy fog is a severe meteorological disaster that greatly increases highway traffic accidents. Accurately predicting highway visibility and taking timely response measures can reduce the impact of meteorological disasters and improve traffic safety. We proposed an Attention-based BiLSTM-CNN (ABCNet) model, which synergized attention mechanisms with BiLSTM and CNN technologies to forecast atmospheric visibility more accurately. First, the Bi-LSTM module processed information both forward and backward, capturing intricate temporal dependencies in the model. Second, the multi-head attention mechanism following the Bi-LSTM distilled and prioritized salient features from multiple aspects of the sequence data. Third, the CNN module recognized local spatial features, and a singular attention mechanism refined the feature map after the CNN module, further enhancing the model's accuracy and predictive capability. Experiments showed that the model was accurate, effective, and significantly advanced compared to conventional models. It could fully extract the spatiotemporal characteristics of meteorological elements. The model was integrated into practical systems with positive results. Additionally, this study provides a self-collected meteorological dataset for highways in high-altitude mountainous areas.

Keywords: traffic; meteorological disaster; atmospheric visibility; forecasting; deep learning

Abbreviations: ABCNet: Attention-based Bi-LSTM-CNN Network; AI: Artificial Intelligence; ARIMA: Autoregressive Integrated Moving Average; Bi-LSTM: Bidirectional Long Short-Term Memory; CNN: Convolutional Neural Network; Conv1D: One-dimensional Convolution; GRU: Gated Recurrent Unit; LSTM: Long Short-Term Memory; MAE: Mean Absolute Error; MAPE: Mean Absolute Percentage Error; MSE: Mean Squared Error; WD-17: A dataset created and contributed by this study for ABCNet model training and evaluation; WD-Vigo: A public visibility dataset from Vigo Airport, Spain, used to demonstrate the universality of the ABCNet model

1. Introduction

The critical importance of road safety has become increasingly pronounced with the ongoing advancements in transportation vehicles and infrastructure development [1–3]. Among the numerous factors compromising road safety, meteorological disasters such as heavy fog, haze, rain, snow, etc. along highways stand out as a significant impediment to traffic efficiency [4,5]. Yunnan Province in China, characterized by its mountainous terrain and frequent low visibility conditions, epitomizes the challenges posed by meteorological disasters on highways. The province’s extensive road network, particularly in higher altitude areas, is vulnerable to low visibility conditions, underscoring the need for accurate visibility predictions to facilitate effective traffic management and safety measures.

Figure 1 shows actual scene photos from the Yunnan Province Plateau Mountain Area Traffic Meteorological Database collected during this study, demonstrating drastic visibility changes (within 36 minutes) at the same spatial location.



Figure 1. Actual scenes of highway visibility changes. (a) visibility 8880 m (datetime: 2023-11-04 07:45:41). (b) visibility 837 m (datetime: 2023-11-04 07:51:41). (c) visibility 105 m (datetime: 2023-11-04 07:57:41). (d) visibility 1157 m (datetime: 2023-11-04 08:21:41).

In this study, we focused on improving highway visibility prediction through the analysis of meteorological and related data using advanced computational models and algorithms.

Traditional visibility prediction methods, such as statistical and regression analysis, suffer from limitations including low accuracy, lengthy computation times, and complexity, hindering their practical applicability for ensuring swift traffic flow on highways.

In response to these limitations, this research harnesses the potential of deep learning, which has shown remarkable success in fields like computer vision and natural language processing, to advance the state of road visibility prediction. By leveraging neural networks' adaptability and non-linear mapping capabilities, deep learning-based predictions offer a promising approach to achieving accurate and reliable visibility forecasts. However, the application of deep learning in this domain is not without challenges. Current techniques predominantly rely on image data samples, which demand high-quality data collection standards and are susceptible to various non-meteorological factors that can impair data integrity.

We propose ABCNet, an attention-based BiLSTM-CNN framework for highway visibility prediction using multi-dimensional non-image data. The contributions are as follows:

- A meteorological dataset is constructed from specialized equipment on highways in the plateau mountain areas of Yunnan Province, China, covering various weather conditions and visibility levels.
- A hybrid model is designed that combines the advantages of BiLSTM and CNN to extract temporal and spatial features from non-image data, utilizing an attention mechanism to enhance the representation of features most relevant for visibility prediction.
- Extensive experiments are conducted to evaluate the model's performance, comparing it with several state-of-the-art deep learning baseline methods. The results indicate superior accuracy, robustness, and generalization of the model.
- The practical value of the model for highway management is demonstrated by providing visibility predictions and alerts for drivers and operators, significantly reducing the probability of traffic accidents.

The contribution is not merely the amalgamation of techniques but their novel application and the creation of a unique dataset. Its uniqueness lies in the synergistic integration of these methods with the newly developed plateau mountain area traffic meteorological dataset. This comprehensive approach collectively enhances the model's predictive accuracy and applicability across varied geographic regions.

The rest of this paper is organized as follows. Section 2 reviews current advancements in highway visibility prediction, encapsulating the related work. Section 3 presents both utilized and self-developed datasets, underscoring the contribution to the field. Section 4 delves into the methodology, detailing the ABCNet model proposed in this paper. Section 5 is dedicated to experiments and analysis, evaluating and comparing the model's performance and analyzing the results. Finally, Section 6 concludes the paper, summarizing the major contributions and looking forward to future research directions.

2. Related work

Accurately predicting low visibility scenarios is crucial for operational safety at airports and along coastlines. Visibility prediction currently relies heavily on numerical forecasting methods

similar to weather prediction. Zhang et al. [6] introduced a multimodal fusion technique to construct a weather visibility prediction system. Cornejo-Bueno et al. [7] investigated the persistence and prediction of low-visibility events at Villanubla Airport in Spain, particularly during winter months. They studied the Runway Visual Range (RVR) time series and evaluated short-term visibility persistence using Markov chain analysis.

Kamangir et al. [8] proposed a deep learning framework with an attention mechanism for visibility prediction, achieving state-of-the-art accuracy (68.9%) on runway visual range prediction within a custom dataset collected at airport observation stations. Liu et al. [9] developed both a polynomial regression model and a deep neural network (DNN) model for visibility prediction. Yu et al. [10] focused on applying a machine-learning-based fusion model to visibility forecasting in Shanghai, China. They introduced a boosting-based fusion model (BFM) and compared it to other prediction models, including LightGBM based on multisource data (LGBM) and RAEMS.

Zang et al. [11] developed a recursive neural network (RNN) prediction model named SwiftRNN, which outperformed ConvLSTM and PredRNN models regarding skill scores for visibility prediction. Ortega et al. [12] reviewed a variety of deep learning models that have been used for visibility prediction, and they compared the performance of these models on a variety of datasets. Peláez-Rodríguez et al. [13] proposed a novel ensemble model for atmospheric visibility forecasting. The proposed model is based on the combination of machine learning models and numerical weather prediction (NWP) data.

Kim et al. [14] aimed to enhance visibility forecasts by establishing an automatic visibility observation network composed of 291 forward-scattering sensors in South Korea. Data assimilation improved prediction skills, particularly within a nine-hour forecast window and for extremely low-visibility events. Qian et al. [15] investigated the application of anomaly-based weather analysis to predict low visibility associated with coastal fog at Ningbo Zhoushan port in East China. Fernández-González et al. [16] studied forecasting of poor visibility episodes near Tenerife Norte airport, testing various methods for estimating visibility based on mesoscale model outputs and presented an application for real-time monitoring of weather conditions to assess poor-visibility risk. Pahlavan et al. [17] conducted a numerical prediction study of radiation and CBL fog events over Iran using the WRF model with different model configurations and visibility prediction as a key focus.

Egli et al. [18] combined data and video quantifying images to analyze large fog evolution trends quantitatively. Kim et al. [19] predicted visibility in South Korea (VISRF) using a random forest (RF) model based on ground observation data from the Automated Synoptic Observing System (ASOS) and air pollutant data from the European Centre for Medium-Range Weather Forecasts (ECMWF) Copernicus Atmosphere Monitoring Service (CAMS) model. Their method exhibited a smaller bias below 2 km compared to other visibility parameterization schemes.

Moreover, Wen et al. [20] compared the performance of five common machine learning methods under various training parameter schemes, including XGBoost, LightGBM, Random Forest (RF), Support Vector Machine (SVM) and Multiple Linear Regression (MLR) using long-term measured data. Noteworthy research contributions are also made by Zhen et al. [21] and Shi et al. [22].

In summary, the related research works reveal the following shortcomings:

- **Limited Geographic Applicability:** Many existing studies rely on datasets from airports and ports, which primarily reflect conditions in plain areas. This specialization leaves a significant gap in research for highway visibility prediction across diverse topographies, such as mountainous regions, where environmental conditions drastically differ.

- **Inadequate Evaluation Across Conditions:** Many studies do not comprehensively evaluate their methods under a variety of conditions, leading to a limited understanding of their performance and robustness. The absence of rigorous, diverse-condition testing restricts the proven applicability of these methods, highlighting the importance of extensive validation efforts to ensure reliability and effectiveness in real-world scenarios.

- **Limited Adaptability and Generalization:** Existing methods often lack adaptability to diverse environmental conditions, making them less effective for wide-ranging geographic and meteorological variations. This shortfall is particularly evident in models optimized for specific climates or regions, which may not perform well under different conditions, underscoring the need for more versatile and generalizable approaches.

- Despite the availability of datasets from structured environments like airports and ports, a critical gap exists in validating prediction models within operational highway systems. It is essential to ascertain real-world efficacy, as it involves ground truth data from diverse and dynamic traffic conditions, particularly in complex terrains.

In this paper, a deep learning model combined with multiple dimensions of meteorological data is leveraged to create a comprehensive approach for highway visibility prediction. Specifically, an Attention-based BiLSTM-CNN network (ABCNet) is proposed, a cutting-edge prediction model designed to provide accurate and timely short-term visibility forecasts on highways. The proposed method has the following advantages:

- The method significantly enhances geographic applicability by leveraging a uniquely developed plateau mountain area traffic meteorological dataset. This dataset empowers the model to effectively predict visibility in diverse topographies, particularly in mountainous regions where conditions vary significantly.

- The model exhibits strong universality. The model has also been compared and tested on the self-built and public datasets. The results show that the model is universally applicable and suitable for various visibility prediction scenarios.

- The prediction accuracy has been validated with accurate application data. The model has been integrated into the actual “Highway Traffic Meteorological Intelligent Monitoring and Proactive Control System” for the past four years. Frontline users have validated its accuracy. This is specifically detailed in Section 5.7.

3. Dataset

3.1. Data collection

The research is supported by the National Engineering Laboratory for Surface Transportation Weather Impacts Prevention. This laboratory is a leading institution in China focusing on transport meteorology, particularly in the field of highway meteorology. Figure 2 illustrates the meteorological data collection device used by the team, which is pivotal for the dataset creation in this study.

The Mazhao Highway in Yunnan Province, China, is a typical mountain highway with complex climate and frequent sudden visibility shifts, making it highly suitable for the visibility research. In accordance with the technical standards for the construction of highway meteorological station networks [23], 17 multi-element traffic meteorological stations were installed on the Mazhao Highway by the end of 2021. Each data record collected by the meteorological equipment

includes 15 meteorological elements such as visibility, wind speed, temperature, air pressure, humidity, wind direction, precipitation, road surface temperature, and road conditions. Figures 3 and 4 show the geographical locations of the 17 meteorological stations and examples of the raw meteorological data, respectively.

Based on the data from these 17 meteorological stations, a highway meteorological element dataset named WD-17 [24] was constructed for the visibility prediction task. This dataset contains 8,348,575 entries of multidimensional, high-precision, high-integrity, and high-quality meteorological sample data (one entry per minute) collected from March 2022 to July 2023.



Figure 2. Self-developed traffic meteorological data collection equipment.

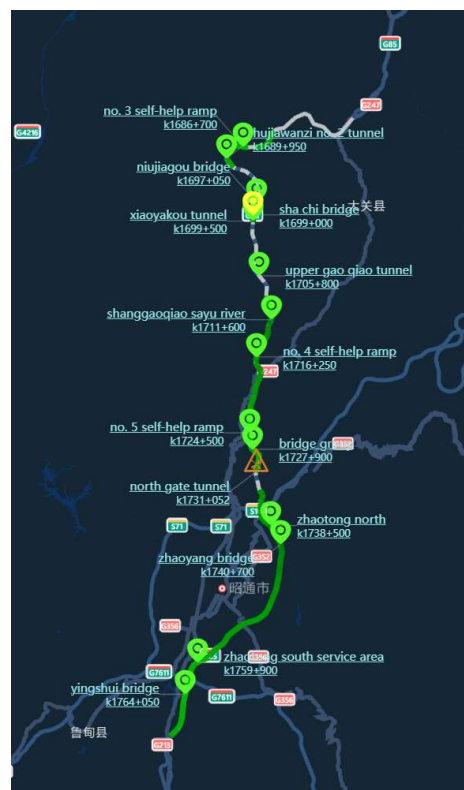


Figure 3. Geographical locations of the 17 meteorological stations.

This is believed to be the first public minute-level meteorological dataset in the world for highways in high-altitude mountainous areas. It has multi-dimension meteorological data on dramatically changing visibility that is crucial for building high-performance visibility prediction models.

A	B	C	D	E	F	G	H	I	J	K	L	M
Id	DeviceId	Name	Description	Unit	Quality	Value	MinValue	MaxValue	Precision	Factor	TimeStamp	
1	1	AMA	minute visibility	m	1	18080	10	80000	0	0	20231219000700000	
3	1	ABA	temperature	°C	1	-32	-40	50	1	1	20231219000700000	
4	1	ACA	humidity	%RH	1	890	0	100	1	1	20231219000700000	
5	1	AEA	instantaneous wind speed	m/s	1	0	0	50	1	1	20231219000700000	
6	1	ADA	instantaneous wind direction	°	1	90	0	359	0	0	20231219000700000	
7	1	AEAEX	extreme wind speed in one minutes	m/s	1	5	0	60	1	1	20231219000700000	
8	1	ADAEX	extremely strong wind direction in one m	°	1	90	0	359	0	0	20231219000700000	
9	1	AEC	2 min average wind speed	m/s	1	0	0	60	1	1	20231219000700000	
10	1	ADC	2 min mean wind direction	°	1	90	0	359	0	0	20231219000700000	
11	1	AED	10 min average wind speed	m/s	1	2	0	60	1	1	20231219000700000	
12	1	ADD	10 min mean wind direction	°	1	90	0	359	0	0	20231219000700000	
13	1	AGA	atmospheric pressure	hPa	1	8209	500	1100	1	1	20231219000700000	
14	1	APA	pavement temperature	°C	1	-28	-50	80	1	1	20231219000700000	
15	1	APH	pavement conditions		1	4	0	254	0	0	20231219000700000	
16	1	APD	water film thickness	mm	1	1	0	25.4	1	1	20231219000700000	
17	1	APE	ice thickness	mm	1	0	0	25.4	1	1	20231219000700000	
18	1	APF	snow layer thickness	mm	1	0	0	25.4	1	1	20231219000700000	
19	1	AQA	slip factor		1	75	0	1	2	2	20231219000700000	
20	2	AMA	minute visibility	m	1	27719	10	40000	0	0	20231219000800000	
22	2	ABA	temperature	°C	1	-45	-40	50	1	1	20231219000800000	
23	2	ACA	humidity	%RH	1	810	0	100	1	1	20231219000800000	
24	2	AEA	instantaneous wind speed	m/s	1	45	0	50	1	1	20231219000800000	
25	2	ADA	instantaneous wind direction	°	1	61	0	359	0	0	20231219000800000	
26	2	AEAEX	extreme wind speed in one minutes	m/s	1	45	0	60	1	1	20231219000800000	
27	2	ADAEX	extremely strong wind direction in one m	°	1	61	0	359	0	0	20231219000800000	

Figure 4. Collected raw meteorological data.

3.2. Dataset preprocessing

The meteorological elements collected by the weather stations along the highways include nearly twenty different factors such as temperature, rainfall, visibility, humidity, road surface temperature, road conditions, wind speed, wind direction, air pressure, subgrade temperature, water film thickness, freezing point temperature, ice layer thickness, snow layer thickness, and slipperiness coefficients. Some of these elements may be irrelevant or redundant for visibility prediction, contributing to limitations in model training and predictive performance. In fact, only about five elements might be strongly correlated with visibility prediction. Therefore, the cosine similarity method is used to filter out low-correlation features, reducing data dimensionality, and enhancing the model's predictive performance. The calculation of cosine similarity is shown in Eq (1).

$$\text{Similarity}(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n (A_i \times B_i)}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}}, \quad (1)$$

herein, A and B represent two different meteorological factors. The higher the value of Similarity

(A, B), the more relevant the two features are. After removing redundant features with extremely high correlation coefficients (greater than 0.9), six meteorological elements (visibility, wind speed, temperature, humidity, precipitation and road surface temperature) were retained for the research. The dataset format is shown in Table 1.

Table 1. Format of dataset WD-17 after feature dimensionality reduction.

Time	Visibility (m)	Temperature (°C)	Humidity (%)	Precipitation (mm)	Wind speed (m/s)	Pavement temperature (°C)
2023-03-10 09:00:00	21334	9.3	84	0.2	2.0	12.3
2023-03-10 09:10:00	39965	9.6	78	0.0	1.5	13.2
2023-03-10 09:20:00	40551	11.4	67	0.0	1.4	17.4
2023-03-10 09:30:00	41525	11.2	79	0.0	0.9	16.8
...

3.3. Other comparative experimental datasets

In the experiments, to demonstrate the universality of the model in visibility prediction tasks, a public visibility dataset WD-Vigo [25] from the Vigo Airport weather station in Spain was also used, spanning 2008 to 2020. This dataset records meteorological elements every 30 minutes, comprising a total of 219,439 data entries, including visibility, temperature, humidity, wind direction, wind speed and air pressure. The format of the dataset is shown in Table 2.

Table 2. Format of dataset WD-Vigo.

Time	Visibility (m)	Temperature (°C)	Humidity (%)	Wind direction (°)	Wind speed (m/s)	Atmospheric pressure (Pa)
2020-12-23 10:00:00	5005	28.3	81	210	3.6	1025
2020-12-23 09:30:00	4506	24.6	86	180	1.5	1025
2020-12-23 09:00:00	1996	22.4	87	190	2.5	1023
2020-12-23 08:30:00	2494	19.8	93	170	1.2	1023
...

4. Methods

Drawing upon the advancements in sequence data processing across domains such as speech recognition and natural language processing, this study introduces the ABCNet model, a novel framework designed for the prediction of highway visibility. The innovation of ABCNet stems from its integrative use of Bi-directional Long Short-Term Memory (Bi-LSTM), Convolutional Neural Network (CNN), and attention mechanisms, offering a comprehensive approach to capturing the spatiotemporal dynamics of meteorological factors impacting visibility. Specifically, the Bi-LSTM module is employed to analyze temporal sequences in both forward and reverse directions, thereby unraveling complex temporal patterns within the data. This temporal analysis is enhanced by a

multi-head attention mechanism, situated subsequent to the Bi-LSTM layer, which serves to isolate and amplify critical features from the temporal data, ensuring a focused analysis on the most influential elements. Moreover, the incorporation of a CNN module aids in the extraction of spatial features. This integration, coupled with a singular attention mechanism, optimizes the refinement of the feature map, significantly bolstering the model's predictive accuracy and capability. The structural design of ABCNet, including its sequential integration of these computational components, is delineated in Figure 5.

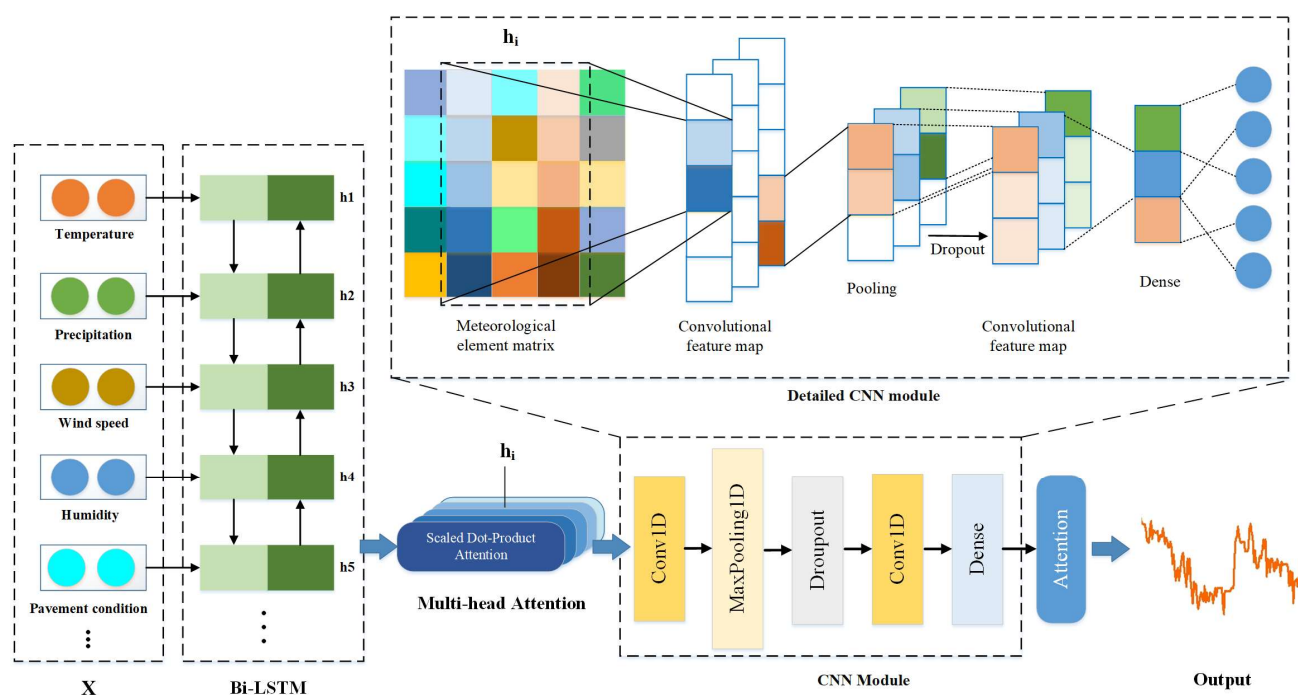


Figure 5. The overall structure of ABCNet.

Integrating Bi-LSTM with multi-head attention mechanisms, the model enhances feature capture and representation precision. Following Bi-LSTM, the multi-head attention offers:

- **Enhanced Representation:** It enables simultaneous focus on diverse aspects of data, ensuring a more comprehensive information capture.
- **Robustness:** Parallel processing improves the model's resilience against scene variations or noise.
- **Adaptive Learning:** Automatically adjusts feature weights, aiding in the precise identification of important sequence parts.
- **Long-Distance Dependency Recognition:** This focuses on various sequence positions, effectively recognizing long-range dependencies.

Furthermore, combining Bi-LSTM and CNN exploits each method's strengths, offering:

- **Spatiotemporal Feature Capture:** Bi-LSTM manages temporal dependencies; CNN excels in spatial pattern recognition. Together, they reveal complex spatiotemporal relationships.
- **Multilevel Feature Extraction:** This combination enhances data understanding and representation through a layered analysis approach.
- **Multiscale Optimization:** Improves handling of information across different scales, boosting adaptability.

These components and their synergistic benefits underpin the ABCNet model's framework, detailed further in subsequent Sections.

4.1. Problem formulation

In this research, time series data of meteorological factors pertinent to highway conditions are analyzed, and collected via dedicated weather monitoring devices. The dataset is organized into segments based on a predefined temporal window of length n , formulated as $S = \{s_1, s_2, \dots, s_n\}$, where each $s_i \in \mathbb{R}^d$ ($i=1, 2, \dots, n$) signifies a vector of meteorological variables (e.g., temperature, precipitation, wind speed, humidity, and road surface conditions) corresponding to sequential time points. The Bi-LSTM model processes these segments through an input matrix $X \in N \times D$, where N and D respectively denote the temporal window's size and the dimensionality of each vector within that window. The Bi-LSTM's output, represented as h_i , encapsulates a complex feature representation of the time series data, further refined by a multi-head attention mechanism to accentuate significant features. Subsequently, the CNN module, enhanced by a singular attention mechanism, integrates these features to deduce the final model output. The primary objective is forecasting future road visibility, utilizing a fixed-length series of meteorological data S .

4.2. Bi-LSTM module

Visibility on highways is forecasted using historical data, based on the rationale that past patterns provide insights into future conditions. This approach naturally implies that short-term forecasts are often more precise than long-term ones due to the expected continuity of recent trends. The analysis of visibility trends benefits significantly from examining time series data both forward and backward. While conventional LSTM models may struggle to capture the nuanced features of these trends, Bi-LSTM stands out by assimilating bidirectional time series data, thus enriching the model's informational base.

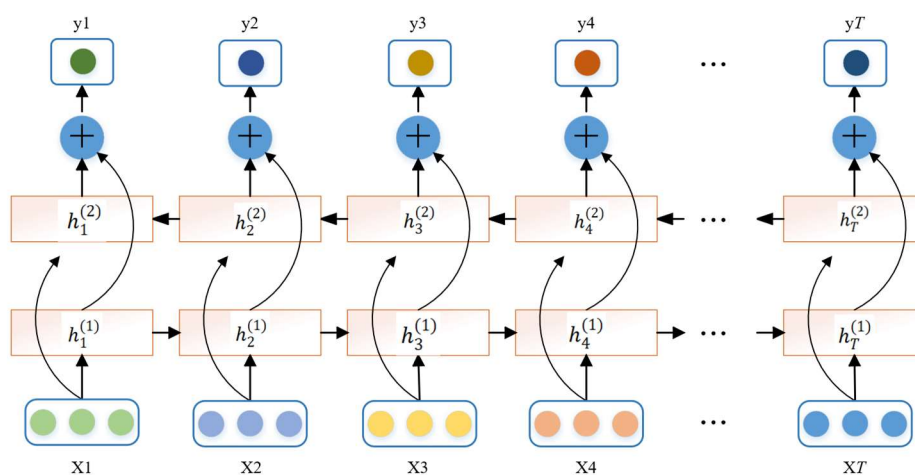


Figure 6. The network structure of Bi-LSTM.

By incorporating two LSTM layers, Bi-LSTM overcomes the limitations of traditional models, which predominantly rely on historical data for future predictions. It effectively utilizes both preceding and succeeding data points, facilitating a comprehensive analysis of temporal patterns. Bi-LSTM includes forward and backward models for processing meteorological data, offering a holistic approach to data analysis. Figure 6 illustrates a standard Bi-LSTM model configuration, showcasing its ability to leverage bidirectional data for enhanced predictive accuracy.

Bi-LSTM's forward and backward outputs are connected using Eq (2) for further processing.

$$h_i = \overrightarrow{h}_i \parallel \overleftarrow{h}_i, h_i \in R^{2L}, \quad (2)$$

where \overrightarrow{h}_i and \overleftarrow{h}_i represent the outputs of the forward and backward LSTM, respectively. \parallel denotes the concatenation operation, and L is the size of each LSTM.

The Bi-LSTM layer is used to obtain the representation of the meteorological element time series. The word embedding vector $S = \{s_1, s_2, \dots, s_n\}$ of the series is first obtained, then apply the Bi-LSTM layer to S to obtain the forward vector sequence as in Eq (3), the backward vector sequence as in Eq (4) and concatenate these two sequences to get Eq (5).

$$L_f = \{l_{f1}, l_{f2}, \dots, l_{fn}\}, \quad (3)$$

$$L_b = \{l_{b1}, l_{b2}, \dots, l_{bn}\}, \quad (4)$$

$$L = \begin{bmatrix} L_f \\ L_b \end{bmatrix} = \{l_1, l_2, \dots, l_n\}, \quad (5)$$

where L is the fundamental features of S and $l_i = [l_{fi}, l_{b(n-i+1)}]^T$. l_i represents the feature of each sequence in the sequence set. As the output vector of Bi-LSTM, l_i will be input into the CNN module through a multi-head attention mechanism.

4.3. Attention and multi-head attention

The attention mechanism, as outlined in reference [26], assigns a score to each dimension of the input data, subsequently weighting the features based on these scores to accentuate the critical features. This process enables the mechanism to significantly influence downstream models or modules by prioritizing information that is deemed most relevant for the task at hand. The operational details of the attention mechanism are encapsulated by Eq (6).

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (6)$$

where Q represents the query vector, K is the key vector, and V is the value vector. d_k represents the dimension of V , superscript T in the equation represents the transpose of K . Q , K , and V are all weight matrices, initially randomized and then optimized during the gradient descent process of training data.

The attention mechanism simplifies to mapping input elements into vectors within a matrix X , followed by calculating attention weights using vectors Q , K and V for a weighted summation output. Enhanced by a multi-head mechanism from the Transformer model [27], which operates in parallel subspaces for richer feature processing, this approach significantly boosts feature capture capabilities. The ABCNet model incorporates multi-head attention to better understand complex data dependencies. The output of the Bi-LSTM is directed into a multi-head attention layer, enhancing performance by focusing on varied feature levels. The structure of multi-head attention is detailed in Figure 7. The multi-head attention layer is shown in Eqs (7) and (8).

$$head_l = Attention(QW_l^Q, KW_l^K, VW_l^V), \quad (7)$$

$$MultiHead = Concat(head_1, \dots, head_l)W^O, \quad (8)$$

where Q, K, V are transformed from the output of Bi-LSTM, $head_l$ represents the attention score calculated by the l -th head in multi-head attention. $Concat$ represents the concatenation operation. W_l^Q, W_l^K, W_l^V, W^O are several different weight matrices, which are initially randomized and then automatically adjusted throughout the training process.

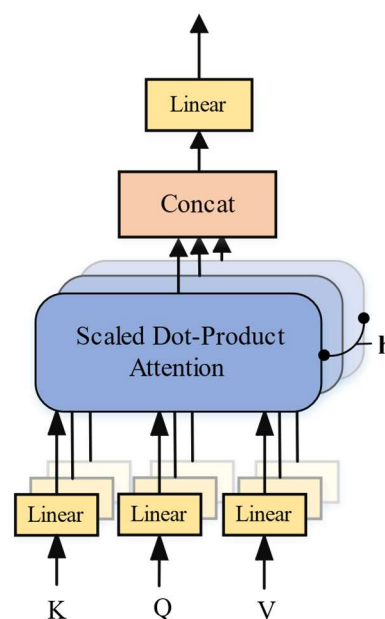


Figure 7. The structure of multi-head attention.

4.4. CNN module

While CNNs are commonly associated with image data processing, ABCNet distinctively employs one-dimensional convolution (Conv1D) for the analysis of time series data. This method demonstrates superior precision compared to traditional linear regression models. Conv1D is particularly effective for processing various forms of sequence data, including text, audio, and time series, by taking a two-dimensional tensor as input. This tensor's first dimension accounts for time

steps, while the second dimension captures the feature dimensions associated with each time step. In a manner analogous to Conv2D within CNN architectures, Conv1D operates according to the equation:

$$y(i) = b + \Sigma(W(k) * X(i+k)), \quad (9)$$

where $y(i)$ represents the i -th element of the output sequence, b signifies the bias term, $W(k)$ is the weight term, and $X(i+k)$ denotes the input sequence, with k indicating the position of the convolution window. The convolution window is slid over the input sequence to generate the output sequence, adapting to various applications such as analyzing meteorological data series. This sliding process overlays the convolution window onto the input sequence, computing the values within the window to produce the sequence's output.

4.5. Output

The ABCNet model produces a forecasted time series output succinctly expressed as $\hat{Y} = \{\hat{y}_{t+1}, \hat{y}_{t+2}, \dots, \hat{y}_{t+m}\}$. Here, \hat{y}_{t+k} represents the forecasted value for a future time point $t+k$, with m denoting the number of steps ahead from the current time t . This format clearly lays out the model's predictions in a sequential manner, spanning from the near future up to the specified forecasting horizon. Each series element marks a distinct future moment, providing a detailed projection of anticipated conditions throughout the period in question.

In its final stage, the model synthesizes the deep learning processes' collective outcomes, integrating temporal patterns identified by the Bi-LSTM module with spatial characteristics discerned by the CNN module, all finely tuned via an attention mechanism to prioritize significance and precision. This culmination showcases the model's adeptness at merging sophisticated feature extraction with sequential data analysis, highlighting its comprehensive approach to forecasting.

5. Experiments and analysis

5.1. Dataset division

In the experiments, the WD-17 dataset was divided into a training set and a test set in an 8:2 ratio, with the training set comprising 6,678,860 entries and the test set 1,669,715 entries. Similarly, the comparative experimental dataset WD-Vigo was divided into a training set (175,551 entries) and a test set (43,888 entries) in an 8:2 ratio.

5.2. Experimental environment configuration and model parameter settings

The experimental environment operates on an Ubuntu 16.04 LTS system. The CPU used is an Intel I7-13700H, complemented by 128 GB of memory and a 4-TB hard disk. The GPU is an NVIDIA GeForce GTX 4060Ti. The programming language employed is Python, with TensorFlow 2.12.0 serving as the machine learning software development library.

The relevant parameters of the ABCNet model utilized in this article are presented in Table 3.

Table 3. Parameter settings for the ABCNet model.

Parameter	Description	Value
Lr	Learning rate	0.001
Bs	Batch size	256
Activation	Activation Function	Sigmoid
Ts	Time step (min)	15/30/60/120
Tw	Time window (h)	0.5/1/2/3/6/10
Epoch	Iterations of training	100
Dropout	Dropout rate	0.01

5.3. Evaluation metric

In this study, common evaluation metrics is used for time series forecasting tasks: Mean Squared Error Loss (MSE), Mean Absolute Error Loss (MAE) and Mean Absolute Percentage Error (MAPE). The definitions of these metrics are as follows:

- MSE: The average of the absolute squared errors between predicted values and actual values. It measures the gap between predictions and reality. A smaller MSE indicates a better model.
- MAE: Measures the average absolute error between predicted values and actual values. It's a non-negative value; a smaller MAE indicates a better model.
- MAPE: A normalized version of MAE, this metric is sensitive to relative errors and does not change with the global scaling of the target variable, making it suitable for problems with large dimensional differences in the target variable. A smaller MAPE indicates a better model.

$$MAE(y, \tilde{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} |y_i - \tilde{y}_i|, \quad (10)$$

$$MSE(y, \tilde{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \tilde{y}_i)^2, \quad (11)$$

$$MAPE(y, \tilde{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} \frac{|y_i - \tilde{y}_i|}{|y_i|} \times 100\%, \quad (12)$$

where $n_{samples}$ represents the number of samples, y_i is the actual values, \tilde{y}_i is the predicted values.

5.4. Performance comparison with different window sizes and time steps

For time series forecasting tasks, the size of the time window and the length of the forecasting time steps are important parameters. The size of the time window represents the length of historical data used, and the forecasting time step determines the output length of the model. To explore the performance of the ABCNet model in short-term road visibility forecasting with various combinations of time window sizes and forecasting time steps, the time window sizes were set to 30 minutes, 1, 2, 3, 6 and 10 hours. The forecasting time steps were set to 15 and 30 minutes, 1 and 2 hours. The experiments on the WD-17 dataset evaluated the model's performance, using MSE and MAE metrics, across these varying configurations. The results are shown in Tables 4 and 5.

Typically, for time series forecasting, the strategy is to predict in a single step, describing the prediction of the next time step's observation. The tables show that for multi-step prediction methods,

both the size of the window and the length of the forecast time step affect the model's performance. With a fixed window size, the model's prediction error increases with the lengthening of the forecasting time step. Based on the experimental results and considering the practical application scenarios in highway visibility forecasting (where very short forecasting time steps, such as 15 minutes, are not significantly useful for highway management), a time window size of 3 hours and a time step length of 30 minutes were chosen as the optimal parameters for ABCNet. This parameter combination provides the best results in application scenarios.

Table 4. Model's MAE under different time window and forecast time step combinations.

Forecast time step \ Window size	15 min	30 min	1 h	2 h
30 min	0.321	0.743	1.240	1.716
1 h	0.121	0.262	0.985	1.494
2 h	0.033	0.064	0.365	0.942
3 h	0.028	0.031	0.217	0.854
6 h	0.037	0.078	0.287	0.684
10 h	0.077	0.086	0.223	0.602

Table 5. Model's MSE under different time window and forecast time step combinations.

Forecast time step \ Window size	15 min	30 min	1 h	2 h
30 min	0.0772	0.1210	0.2317	0.4011
1 h	0.0523	0.0713	0.1148	0.3568
2 h	0.0091	0.0124	0.0836	0.1445
3 h	0.0022	0.0032	0.0101	0.0331
6 h	0.0050	0.0041	0.0074	0.0126
10 h	0.0065	0.0035	0.0072	0.0082

The loss curve reflects the changes in the loss function during the model training process. It is a function used to evaluate the error in the model's predictive results. The model parameters can be adjusted based on changes in the loss function to make the model's predictions more accurate. With a window size of 3 hours, the loss curves of the model on the WD-17 dataset for each forecast time step length on both the training and test sets were recorded, as shown in Figure 8. The figure shows that the loss curve steadily decreases, and the error in the model's predictive results continually diminishes until it stabilizes, indicating that the model is effectively capturing the patterns in the training data.

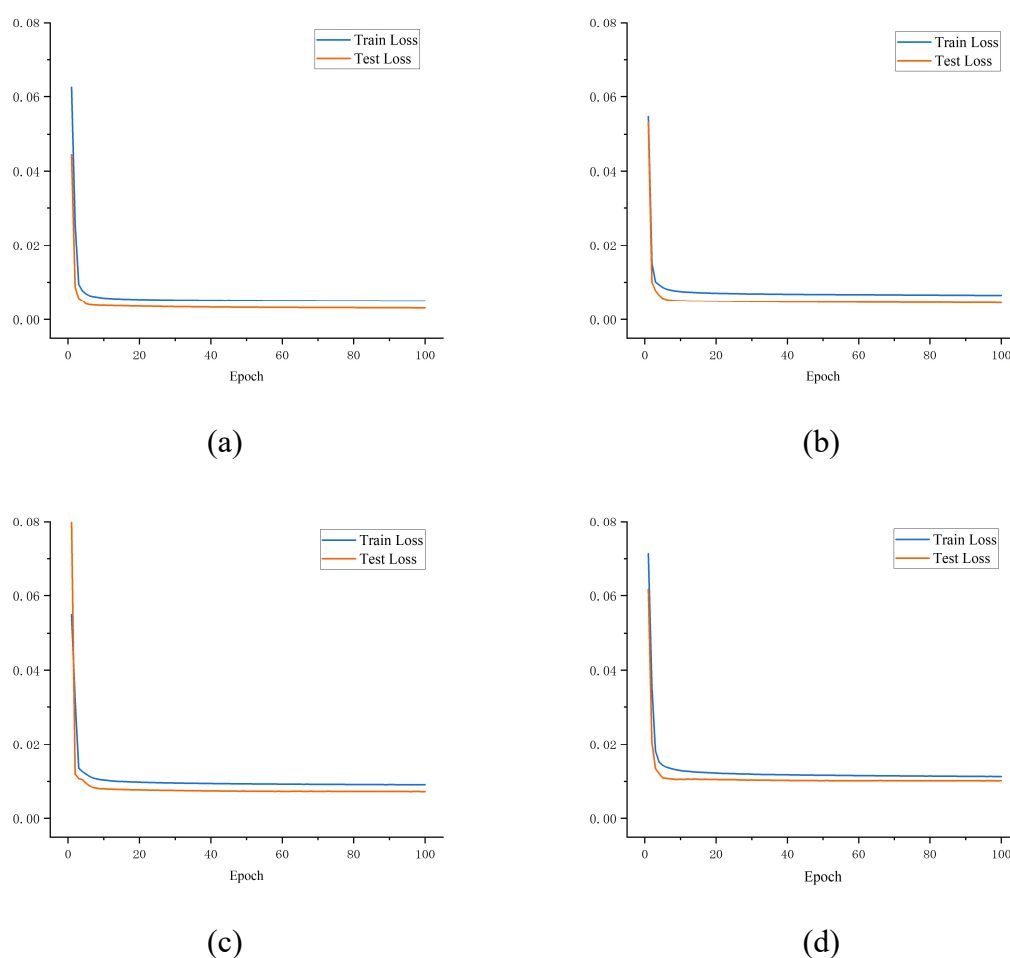


Figure 8. The loss curves under each prediction time step with a window size of 3 hours. (a) Time step 15 mins. (b) Time step 30 mins. (c) Time step 1 hours. (d) Time step 2 hours.

5.5. Ablation experiment

To validate the rationality and superiority of the proposed model architecture and to better demonstrate the contribution of different modules in the ABCNet architecture, an ablation experiment on the WD-17 dataset was conducted. The CNN and Attention modules from ABCNet were removed to compare changes in model performance. Based on the experimental results in Section 5.3, the time window was set to 3 hours and the forecast time step to 30 minutes. Table 6 shows the performance of the ablation experiments for each module of the model.

Table 6. Ablation experiment of the ABCNet model on WD-17 dataset.

Model	MSE	MAE	MAPE
Bi-LSTM	0.0707	0.645	0.2806
Bi-LSTM+CNN	0.0672	0.613	0.2718
Bi-LSTM+Attention	0.0431	0.220	0.1922
Bi-LSTM+CNN+Attention (Ours)	0.0032	0.031	0.0422

The ablation study results indicate that each additional component improves the model's performance, with the combined model (Bi-LSTM+CNN+Attention) achieving the best performance across all metrics, suggesting that the integration of CNN and Attention mechanisms with the Bi-LSTM model significantly enhances the model's prediction accuracy. The lowest MSE (0.0032), MAE (0.031) and MAPE (0.0422) scores for the combined model indicate higher precision and reliability of the results compared to the other variants. This suggests that the additional components contribute positively to the model's ability to capture and utilize relevant patterns in the data.

In conclusion, each component added to the model likely helps capture different types of patterns within the data, and their combination provides a more complete and nuanced understanding of the input features. This is reflected in the lower error rates across all three metrics, indicating a model that generalizes better to new data and provides more accurate predictions.

5.6. Performance comparison with competitive models

To validate the performance of proposed ABCNet model in highway visibility prediction tasks, experiments on the WD-17 and WD-Vigo datasets were conducted using several of the most competitive time series forecasting methods, comparing them with ABCNet. Based on the results from Section 5.3, for both the WD-17 and WD-Vigo datasets, the time window was set to 3 hours and the time step to 30 minutes. The competitive time series forecasting methods used include the traditional ARIMA model for time series forecasting, the decision tree ensemble algorithm XGBoost from machine learning, deep learning models like LSTM, and its variants, and the latest time series forecasting model, Informer. The dataset for each competitive model was carefully prepared, tailoring preprocessing steps to suit their specific requirements. For instance, for the ARIMA model, a differencing step was performed to stabilize the mean of the time series. Each model, including ARIMA, XGBoost, LSTM variants and Informer, underwent a rigorous configuration process, where hyperparameters were tuned using a grid search approach to identify the optimal settings. Finally, evaluation metrics such as MSE, MAE and MAPE were calculated to ensure a comprehensive and fair comparison of their performance on both datasets.

The following is a brief description of these models:

- ARIMA [28]: A difference autoregressive moving average model, used for forecasting non-stationary time series.
- XGBoost [29]: An efficient gradient boosting decision tree algorithm, combining multiple weak learners into a strong learner through forward addition.
- LSTM [30]: A variant of RNN [31], often used for handling nonlinear features in time series.
- LSTM+CNN [32]: Combining CNN and LSTM networks to extract spatial and temporal features respectively.
- GRU+Attention [33]: A model variant based on LSTM, merging the forget and input gates into an update gate, and emphasizing the importance of each hidden layer output with an attention mechanism.
- Informer [34]: Uses a new attention mechanism to automatically adjust the attention scope according to sequence length, effectively handling long sequences. It also adopts a multi-scale time encoder and decoder structure, considering information across different time scales.

The MSE, MAE and MAPE of each model in visibility prediction on the WD-17 and WD-Vigo datasets are shown in Tables 7 and 8.

Table 7. MSE, MAE and MAPE of various models on the WD-17 dataset.

Model	MSE	MAE	MAPE
ARIMA [28]	0.0667	0.657	0.2726
XGBoost [29]	0.0371	0.343	0.2383
LSTM [30]	0.0254	0.132	0.1853
LSTM+CNN [32]	0.0223	0.086	0.1027
GRU+Attention [33]	0.0192	0.078	0.0895
Informer [34]	0.0145	0.064	0.0733
ABCNet (ours)	0.0032	0.031	0.0422

Table 8. MSE, MAE and MAPE of various models on the WD-Vigo dataset.

Model	MSE	MAE	MAPE
ARIMA [28]	0.1233	0.295	0.4100
XGBoost [29]	0.0971	0.262	0.3822
LSTM [30]	0.0874	0.246	0.3703
LSTM+CNN [32]	0.0822	0.227	0.3612
GRU+Attention [33]	0.0628	0.193	0.3146
Informer [34]	0.0517	0.174	0.2213
ABCNet (ours)	0.0325	0.121	0.1458

It is clear from the tables that for both datasets, the performance of neural network models in visibility prediction significantly surpasses that of the machine learning model XGBoost and the traditional time series forecasting model ARIMA. While LSTM and its variants display comparable performances, ABCNet distinctively combines Bi-LSTM, Attention, and CNN modules, which can effectively extract multivariate time series information and spatial feature information and adaptively allocate weights to each feature, significantly outperforming other models. On the WD-17 dataset, compared to the latest time series forecasting model architecture (Informer [32]), ABCNet reduced MSE and MAE by 1.13 and 3.3%, respectively. Similarly, on the WD-Vigo dataset, ABCNet's MSE and MAE were reduced by 1.92 and 5.3%, respectively. This demonstrates ABCNet's good universality, showing superior performance in visibility prediction tasks in various scenarios, including highways and airports.

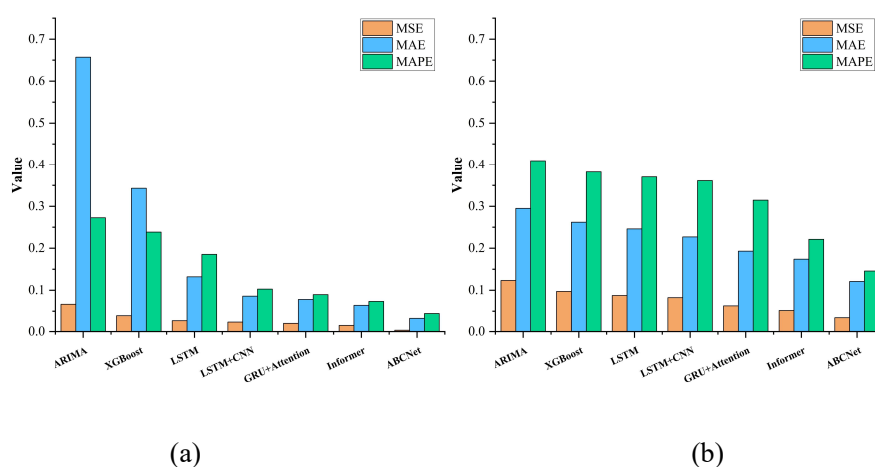
**Figure 9.** Model performance comparison on (a) the WD-17 dataset and (b) the WD-Vigo dataset.

Figure 9 clearly shows the performance advantage of the model, providing an intuitive comparison of various model metrics.

The experiment on the WD-17 dataset was recorded, observing the comparison between actual and predicted values on the test set. Figure 10 displays the comparison results for four different time series intervals. The experimental results show that the model can predict sudden low visibility events and the changes in the predicted values are generally consistent with the actual values, demonstrating the model's good performance.

In summary, the comparative analysis on both the WD-17 and WD-Vigo datasets demonstrates the robustness and adaptability of the proposed model across different data contexts.

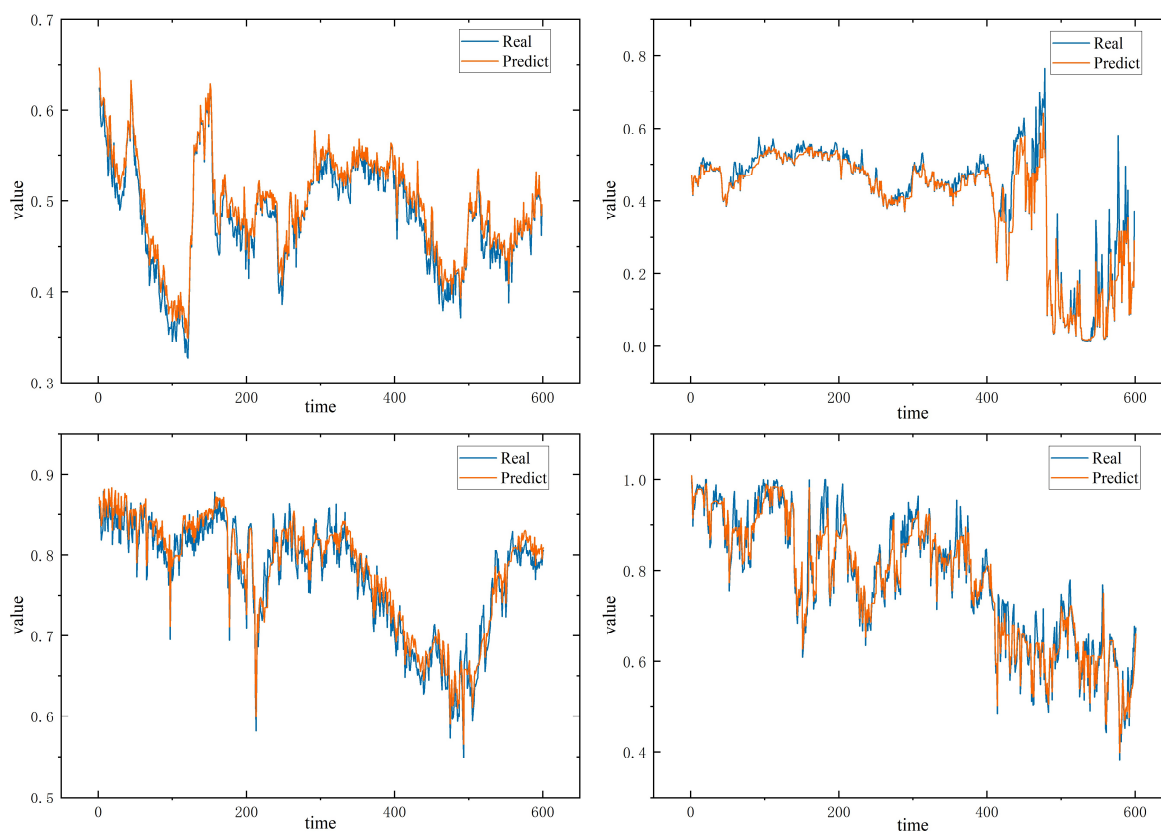


Figure 10. Comparison of predicted and real values for four distinct time series intervals on the WD-17 dataset.

5.7. Practical application system validation

The integration of the ABCNet model into the Highway Traffic Meteorological Intelligent Monitoring and Proactive Control System for Yunnan Communications Investment & Construction Group CO., LTD [35] represents a substantial stride forward. Deployed for validation in the Mazhao section managed by the Zhaotong office, the model has demonstrated its efficacy in enhancing visibility prediction, meeting the anticipated accuracy standards and facilitating local traffic management efforts. This deployment illustrates the model's practical application and its contribution to traffic safety and efficiency. The system's effectiveness is visually encapsulated in Figure 11, illustrating its operational impact.

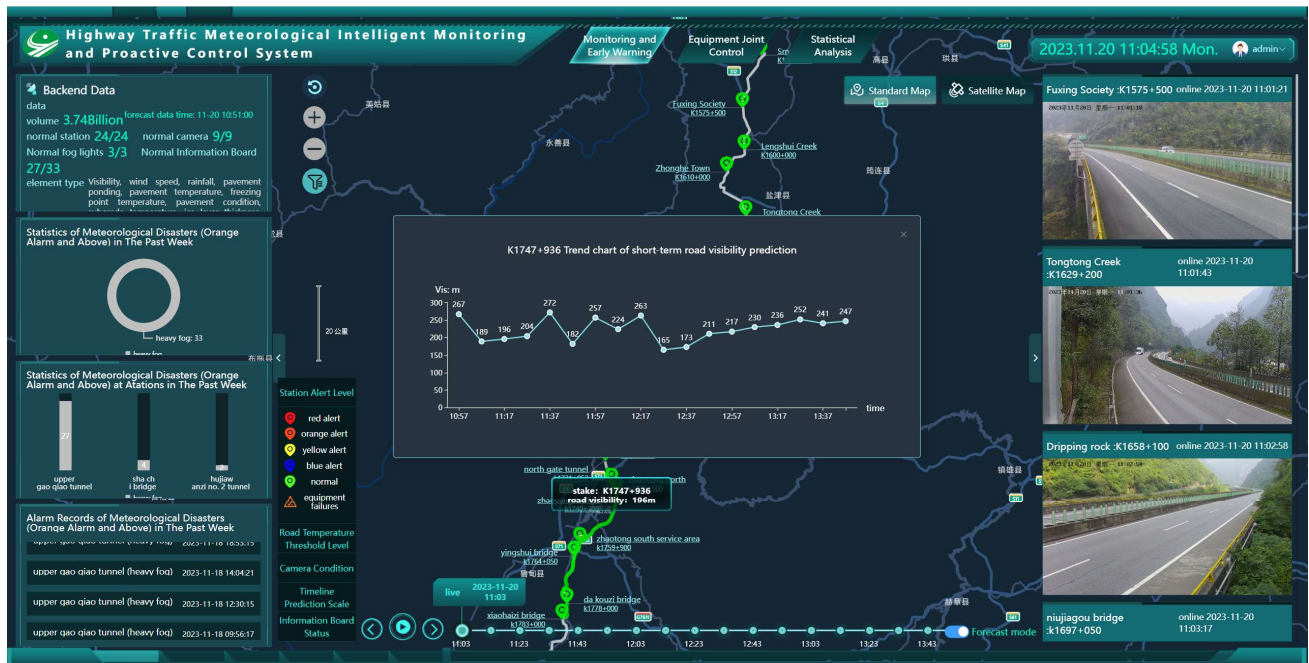


Figure 11. Practical demonstration of model visibility prediction performance in the Highway Traffic Meteorological Intelligent Monitoring and Proactive Control System.

Given the foundational success of ABCNet in the deployment, further exploration into its application is warranted to fully leverage its capabilities for broader impact. The model's precision in visibility prediction offers a critical tool for enhancing road safety under adverse weather conditions, suggesting its potential utility in automated traffic management systems. By integrating ABCNet with dynamic traffic control algorithms, it is conceivable to develop more responsive systems that adjust traffic signals and speed limits in real-time based on visibility data, potentially reducing accident rates and improving traffic flow. Additionally, exploring the integration of ABCNet with vehicle-to-infrastructure communication technologies could pave the way for personalized driver alerts regarding visibility and road conditions. This expanded application scope promises to not only advance the state of intelligent transportation systems but also contribute to the development of smart cities, where transportation efficiency and safety are paramount. Further research in these areas could significantly enhance the manuscript's relevance and underscore the innovative contributions of ABCNet to the field.

6. Conclusions

In this study, the ABCNet model has been successfully developed, which is a novel Attention-based Bi-LSTM-CNN approach for accurate highway visibility prediction. The model harnesses the strengths of Bi-LSTM for bidirectional temporal feature extraction from complex meteorological time series data, complemented by CNN for in-depth spatial feature analysis. Integrating multi-head attention mechanisms further refines the model's capability to assign weights automatically to various features, enhancing its overall predictive accuracy. Another key contribution of this work is developing and utilizing a self-collected meteorological dataset, which significantly improved the model's predictive performance.

The comprehensive evaluations, utilizing metrics such as MSE, MAE and MAPE, demonstrate ABCNet's enhanced performance in visibility prediction over existing models. These results underscore ABCNet's potential as a robust tool for real-world traffic management applications, offering a significant advancement in meteorological forecasting for transportation safety.

The synergy between the novel dataset and the ABCNet methodology is pivotal for advancing visibility prediction. ABCNet's sophisticated design is rigorously tested and refined by the diverse and complex conditions presented in the high-altitude highway meteorological dataset. It is the dataset's detailed representation of challenging visibility scenarios that enhances the model's predictive capabilities. This collaborative approach between advanced methodology and comprehensive data enables a significant advancement in predicting highway visibility that neither could achieve in isolation.

The ABCNet architecture, employing widely adopted mechanisms like Bi-LSTM, CNN and attention mechanisms, provides a solid foundation for accuracy and reliability in various scenarios. However, its capacity for handling exceptionally complex environments or predicting over extended time step may encounter limitations due to the inherent challenges of modeling highly dynamic and unpredictable weather patterns. Future enhancements will explore advanced methodologies to further extend ABCNet's predictive capabilities, ensuring its applicability and effectiveness in more demanding forecasting contexts.

Furthermore, future work will explore incorporating image data from highway cameras into ABCNet to enhance its predictive accuracy by providing direct visual insights into visibility conditions. Additionally, the potential of ABCNet in predictive maintenance of transportation infrastructure under adverse weather conditions remains an exciting avenue for future exploration. By predicting visibility and other related weather parameters, ABCNet could be instrumental in guiding maintenance schedules, thereby improving road safety and longevity. In conclusion, these initiatives underscore ABCNet's potential in the broader domain of intelligent transportation systems.

Use of AI tools declaration

The authors declare that they have not used any Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The research was funded by the Yunnan Key Laboratory of Digital Communications (Grant No. 202205AG070008); the National Natural Science Foundation of China (Grant No. 62162068); the Yunnan Ten Thousand Talents Program and Yunling Scholars Special Project (Grant No. YNWR-YLXZ-2018-022); the Science and Technology Innovation Project of Yunnan Communications Investment & Construction Group Co. Ltd. - Research and Application of Meteorological Disaster Monitoring, Early Warning, and Joint Control Technology for Highway Traffic Prevention and Control Equipment (Grant No. YCIC-YF-2021-08); the Science and Technology Innovation Project of Broadvision Engineering Consultants - Research and Application of Video Analysis Technology for Mountainous Highway Traffic Meteorological Disasters Based on Artificial Intelligence (Grant No. ZL-2021-04); and 2023 Opening Research Fund of Yunnan Key Laboratory of Digital Communications (Grant No. YNJTKFB-20230686, YNKLDC-KFKT-202303).

Conflict of interest

The authors declare no conflicts of interest in this paper.

References

1. R. Faturechi, E. Miller-Hooks, Measuring the performance of transportation infrastructure systems in disasters: A comprehensive review, *J. Infrastruct. Syst.*, **21** (2015), 04014025. [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000212](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000212)
2. Y. Zou, Y. Zhang, K. Cheng, Exploring the impact of climate and extreme weather on fatal traffic accidents, *Sustainability*, **13** (2021), 390. <https://doi.org/10.3390/su13010390>
3. E. Petrova, Natural hazard impacts on transport infrastructure in Russia, *Nat. Hazards Earth Syst. Sci.*, **20** (2020), 1969–1983. <https://doi.org/10.5194/nhess-20-1969-2020>
4. H. Lu, M. Chen, W. Kuang, The impacts of abnormal weather and natural disasters on transport and strategies for enhancing ability for disaster prevention and mitigation, *Transp. Policy*, **98** (2020), 2–9. <https://doi.org/10.1016/j.tranpol.2019.10.006>
5. S. Zhu, H. Yang, D. Liu, H. Wang, L. Zhou, C. Zhu, et al., Observations and forecasts of urban transportation meteorology in China: A review, *Atmosphere*, **13** (2022), 1823. <https://doi.org/10.3390/atmos13111823>
6. C. Zhang, M. Wu, J. Chen, K. Chen, C. Zhang, C. Xie, et al., Weather visibility prediction based on multimodal fusion, *IEEE Access*, **7** (2019), 74776–74786. <https://doi.org/10.1109/ACCESS.2019.2920865>
7. S. Cornejo-Bueno, D. Casillas-Pérez, L. Cornejo-Bueno, M. I. Chidean, A. J. Caamaño, J. Sanz-Justo, et al., Persistence analysis and prediction of low-visibility events at Valladolid Airport, Spain, *Symmetry*, **12** (2020), 1045. <https://doi.org/10.3390/sym12061045>
8. H. Kamangir, W. Collins, P. Tissot, S. A. King, H. T. H. Dinh, N. Durham, et al., FogNet: A multiscale 3D CNN with double-branch dense block and attention mechanism for fog prediction, *Mach. Learn. Appl.*, **5** (2021), 100038. <https://doi.org/10.1016/j.mlwa.2021.100038>
9. Z. Liu, Y. Chen, X. Gu, J. K. Yeoh, Q. Zhang, Visibility classification and influencing-factors analysis of airport: A deep learning approach, *Atmos. Environ.*, **278** (2022), 119085. <https://doi.org/10.1016/j.atmosenv.2022.119085>
10. Z. Yu, Y. Qu, Y. Wang, J. Ma, Y. Cao, Application of machine-learning-based fusion model in visibility forecast: A case study of Shanghai, China, *Remote Sens.*, **13** (2021), 2096. <https://doi.org/10.3390/rs13112096>
11. Z. Zang, X. Bao, Y. Li, Y. Qu, D. Niu, N. Liu, et al., A modified rnn-based deep learning method for prediction of atmospheric visibility, *Remote Sens.*, **15** (2023), 553. <https://doi.org/10.3390/rs15030553>
12. L. C. Ortega, L. D. Otero, M. Solomon, C. E. Otero, A. Fabregas, Deep learning models for visibility forecasting using climatological data, *Int. J. Forecasting*, **39** (2023), 992–1004. <https://doi.org/10.1016/j.ijforecast.2022.03.009>
13. C. Peláez-Rodríguez, J. Pérez-Aracil, A. de Lopez-Diz, C. Casanova-Mateo, D. Fister, S. Jiménez-Fernández, et al., Deep learning ensembles for accurate fog-related low-visibility events forecasting, *Neurocomputing*, **549** (2023), 126435. <https://doi.org/10.1016/j.neucom.2023.126435>

14. M. Kim, K. Lee, Y. H. Lee, Visibility data assimilation and prediction using an observation network in South Korea, *Pure Appl. Geophys.*, **177** (2020), 1125–1141. <https://doi.org/10.1007/s00024-019-02288-z>
15. W. Qian, J. C. H. Leung, Y. Chen, S. Huang, Applying anomaly-based weather analysis to the prediction of low visibility associated with the coastal fog at Ningbo-Zhoushan Port in East China, *Adv. Atmos. Sci.*, **36** (2019), 1060–1077. <https://doi.org/10.1007/s00376-019-8252-5>
16. S. Fernández-González, P. Bolgiani, J. Fernández-Villares, P. González, A. García-Gil, J. C. Suárez, et al., A. Forecasting of poor visibility episodes in the vicinity of Tenerife Norte Airport, *Atmos. Res.*, **223** (2019), 49–59. <https://doi.org/10.1016/j.atmosres.2019.03.012>
17. R. Pahlavan, M. Moradi, S. Tajbakhsh, M. Azadi, M. Rahnama, Numerical prediction of several radiation and CBL fog events over Iran using the WRF model for late December 2015, *J. Earth Space Phys.*, **46** (2020), 561–582. <https://doi.org/10.22059/jesphys.2020.299678.1007202>
18. S. Egli, B. Thies, J. Bendix, A hybrid approach for fog retrieval based on a combination of satellite and ground truth data, *Remote Sens.*, **10** (2018), 628. <https://doi.org/10.3390/rs10040628>
19. B. Y. Kim, J. W. Cha, K. H. Chang, C. Lee, Visibility prediction over South Korea based on random forest, *Atmosphere*, **12** (2021), 552. <https://doi.org/10.3390/atmos12050552>
20. W. Wen, L. Li, P. W. Chan, Y. Y. Liu, M. Wei, Research on the usability of different machine learning methods in visibility forecasting, *Atmósfera*, **37** (2023). <https://doi.org/10.20937/atm.53053>
21. M. Zhen, M. Yi, T. Luo, F. Wang, K. Yang, X. Ma, et al., Application of a fusion model based on machine learning in visibility prediction, *Remote Sens.*, **15** (2023), 1450. <https://doi.org/10.3390/rs15051450>
22. X. Shi, Z. Feng, H. Zhao, X. Zhao, Q. Yang, Y. Pan, Research on Starlink constellation simulation and target area visibility algorithm, in *International Conference on Signal Processing, Computer Networks, and Communications (SPCNC 2022)*, **12626** (2023), 443–450. <https://doi.org/10.1117/12.2674416>
23. *Yunnan Provincial Market Supervision Administration*, Technical Specifications for Construction of Highway Meteorological Station Networks (DB53/T 1091-2022), 2022. Yunnan Province Local Standards.
24. *National Engineering Laboratory for Surface Transportation Weather Impacts Prevention, Broadvision Engineering Consultants Co., Ltd.*, Highway Meteorological Dataset, 2023. Available from: <https://www.kaggle.com/datasets/liphynix2003/highway-meteorological-dataset/>.
25. *THREDDS meteogalicia raw Iowa State University*, Weather Forecasting at Vigo Airport Using AI, 2020. Available from: <https://www.kaggle.com/datasets/jorgerobinat/weather-forecasting-at-vigo-airport-using-ai/data>.
26. A. Galassi, M. Lippi, P. Torrioni, Attention in natural language processing, *IEEE Trans. Neural Networks Learn. Syst.*, **32** (2020), 4291–4308. <https://doi.org/10.1109/TNNLS.2020.3019893>
27. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., Attention is all you need, *Adv. Neural Inf. Process. Syst.*, **30** (2017). <https://doi.org/10.48550/arXiv.1706.03762>
28. V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, G. K. Matsopoulos, A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks, *Future Internet*, **15** (2023), 255. <https://doi.org/10.3390/fi15080255>

29. D. K. Singh, N. Rawat. Machine learning for weather forecasting: XGBoost vs SVM vs random forest in predicting temperature for visakhapatnam, *Int. J. Intell. Syst. Appl.*, **15** (2023), 1–12. <https://doi.org/10.5815/ijisa.2023.05.05>.
30. R. Huang, C. Wei, B. Wang, J. Yang, X. Xu, S. Wu, et al., Well performance prediction based on Long Short-Term Memory (LSTM) neural network, *J. Pet. Sci. Eng.*, **208** (2022), 109686. <https://doi.org/10.1016/j.petrol.2021.109686>
31. S. Jung, J. Moon, S. Park, E. Hwang, An attention-based multilayer GRU model for multistep-ahead short-term load forecasting, *Sensors*, **21** (2021), 1639. <https://doi.org/10.3390/s21051639>
32. C. Han, H. Park, Y. Kim, G. Gim, Hybrid CNN-LSTM based time series data prediction model study, in *IEEE/ACIS International Conference on Big Data, Cloud Computing, and Data Science Engineering*, Danang, Vietnam, (2022), 43–54. https://doi.org/10.1007/978-3-031-19608-9_4
33. H. Abbasimehr, R. Paki, Improving time series forecasting using LSTM and attention models, *J. Ambient Intell. Hum. Comput.*, (2022), 1–19. <https://doi.org/10.1007/s12652-020-02761-x>
34. A. Zeng, M. Chen, L. Zhang, Q. Xu, Are transformers effective for time series forecasting, in *Proceedings of the AAAI Conference on Artificial Intelligence*, Washington DC, USA, **37** (2023), 11121–11128. <https://doi.org/10.1609/aaai.v37i9.26317>
35. *National Engineering Laboratory for Surface Transportation Weather Impacts Prevention & Broadvision Engineering Consultants Co., Ltd.* Highway Traffic Meteorological Intelligent Monitoring and Proactive Control System, 2024. Available from: <https://220.163.107.220:11020/jtqx>.



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)