Parallel Multi-model Fusion Spectrum Prediction Based on Multi-channel Feature Extraction

Wenlu Yue(✉), Lin Qi, and Shuang Li

Harbin Engineering University, Harbin, China
2269119508@hrbeu.edu.cn

Abstract. Radio spectrum prediction is crucial for mining spectrum behavior patterns in complex environments, managing spectrum usage, and improving the probability of cognitive radio access to the spectrum. In this paper, we propose a parallel multi-channel multi-model fusion network (PM^{2}FN) for feature extraction of complex electromagnetic data to achieve accurate spectrum prediction based on the obvious time-frequency correlation exhibited by the spectrum data. However, the accurate point prediction method ignores the random characteristics of the complex electromagnetic environment when predicting data with high volatility, and the traditional deterministic prediction can hardly eliminate the prediction error. Therefore, this paper combines the quantile regression and parallel multi-channel multi-model fusion network (QPM^{2}FN) for probabilistic prediction of the electromagnetic spectrum to effectively quantify the prediction uncertainty. In this paper, a large number of comparative experiments are conducted on the Aachen dataset, and the experimental results show that the proposed model has higher prediction accuracy and effectiveness than the baseline models in terms of accurate prediction and probabilistic prediction.

Keywords: Multi-channel · Multi-model · Spectrum Prediction · Probabilistic Prediction · Quantile regression

1 Introduction

As the radio industry continues to grow, the demand for spectrum for wireless services increases with the growing number of wireless devices, and the contradiction between the low spectrum utilization rate and people's increasing demand for spectrum resources becomes more and more prominent [1]. Spectrum prediction was proposed by Professor Acharya in 2006, mainly to improve spectrum utilization by predicting the spectrum availability in the future period. Most of the existing works mainly use regression time series models [2], Hidden Markov Models (HMM) [3] and association rule mining methods [4, 5] to exploit the temporal correlation exhibited by the spectral data to obtain the trend of the data and enable the prediction of the future state of the data. In recent years, artificial neural networks (ANNs) [6] are widely used for spectrum prediction because of

J. Wang (Ed.): MobiMedia 2023, LNICST 576, pp. 125–136, 2025.
https://doi.org/10.1007/978-3-031-60347-1_12

their ability to efficiently handle nonlinear problems with large amounts of data. Nakisa et al. used Recurrent Neural Networks (RNN) and Time Delay Neural Networks (TDNN) for spectrum prediction for multiple sub-users, but this method is only averaging each sub-user after predicting them separately, and not a joint multi-user prediction [7]. Yu et al. used a modified RNN with gating, i.e., Long Short-Term Memory (LSTM) network, for prediction, which largely alleviated the gradient disappearance dilemma [8]. Zhang et al. designed a deep learning network based on graph convolution to achieve efficient prediction of spectrum resources using a priori information of spatial activities [9]. The above approach is useful for the study of spectrum prediction compared to traditional methods, but it utilizes only single-dimensional features and ignores the influence of multidimensional information on the model. Li et al. conducted spectrum correlation analysis for different channels in the same service and different channels in different services, respectively, and proved the correlation between the spectrum [10] Li et al. proposed a TF^{2}AN model that combines an LSTM network and a generalized fusion module of external factors to capture the time-frequency correlation of the spectrum and external features [11]. Aygül et al. used a multidimensional correlation spectrum problem divided into small subproblems via a composite two-dimensional (2D) LSTM model for multidimensional joint spectral prediction [12]. Shawel et al. used multidimensional correlation in conjunction with convolutional long-term short-term memory (ConvLSTM) for long-term time prediction [13]. The above experiments show that attention is beginning to be paid to the effect of multidimensional time-frequency correlation on prediction performance, while it has been found that hybrid models have higher prediction accuracy than single models.

Currently, research on spectrum prediction focuses on accurate point prediction, but it is difficult for point prediction to portray the uncertainty of future electromagnetic spectrum activities in the complex electromagnetic environment. Probabilistic prediction allows conditional modeling of prediction uncertainty, giving information about the probability distribution of the predicted object and effectively quantifying the stochasticity of the prediction [14]. Therefore, probabilistic prediction can quantify and analyze the uncertainty of spectrum data to portray the general characteristics of spectrum activities in complex electromagnetic environments. Probabilistic prediction is mainly performed using quantile regression (QR) method. QR, as a classical regression analysis algorithm, allows estimating the conditional distribution of the explanatory variables [15] and characterizing their distribution using a suitable regression function [16]. Cannon used the support vector quantile regression (SVQR) method to obtain the quantile and further compared the effect of different kernel functions to construct the probability density function [17]. However, QR does not have the ability to handle nonlinear problems alone, and in order to better handle nonlinear relationships between data [18], Yu et al. combined quantile regression with neural networks and proposed a deep learning temporal quantile regression model that uses LSTM networks to learn the information of time series [19]. In summary, it can be seen that deterministic forecasting usually provides more limited predictive information, while probabilistic forecasting provides more complete probabilistic statistical information about the object to be predicted with quantile estimates as the output.

In summary, traditional algorithms often tend to ignore the time-frequency correlation characteristics when dealing with multi-channel spectrum prediction problems and cause their prediction accuracy to be low. In the complex electromagnetic environment, the accurate point prediction method cannot predict the randomness of the spectrum fluctuation, and the prediction results will have large deviation when the spectrum has short-time fluctuation. To solve the above problems, this paper designs a parallel multi-channel multi-model fusion framework-PM^{2}FN based on the combination of global features and local correlation features. The model achieves high accuracy spectrum prediction by maximizing the retention of local features and global representation through two parallel networks. Meanwhile, we apply probabilistic prediction to spectrum prediction for the first time, and introduce quantile regression algorithm in PM^{2}FN to retrain the network using quantile loss to obtain probabilistic prediction results of spectrum with different confidence levels and improve the reliability of spectrum prediction. We conducted extensive experiments on the Aachen dataset to verify that PM^{2}FN and QPM^{2}FN have higher accuracy and reliability than the baselines.

2 Problem Definition

The problem of predicting the occupancy of the electromagnetic spectrum in cognitive radio can be described as using the historical spectrum data $X = (x_1, x_2 \cdots x_t)$, where t denotes the time slot, and by constructing prediction models for training, we dig deeper into the feature association relationships between historical data, and the prediction model will derive the future moment $Y = (y_{t+1}, y_{t+2} \cdots y_{t+l})$ accordingly based on the training results, where l denotes the prediction step. Based on the time-frequency correlation characteristics exhibited by the electromagnetic spectrum data, assuming that N channels are measured simultaneously in T time slots, we obtain the raw spectrum data consisting of matrix $X = \left(x^1, x^2, x^3, \ldots, x^N\right)^\top = (x_1, x_2, x_3, \ldots, x_t) \in \mathbb{R}^{N\times T}$, from which the number of multi-channel multi-time slots is selected as historical data to predict the future output results $Y = (y_{t+1}, y_{t+2}, y_{t+3}, \ldots, y_{t+l}) = (y_{t+1}, y_{t+2}, y_{t+3}, \ldots, y_{t+l}) \in \mathbb{R}^{N\times L}$. In order to quantify the randomness feature in spectrum prediction, we use quantile regression algorithm for spectrum probability prediction, and then obtain the prediction results $Y_q = \left(y_q^1, y_q^2, y_q^3, \ldots, y_q^N\right)^\top = \left(y_{q(t+1)}, y_{q(t+1)}, y_{q(t+2)}, y_{q(t+3)}\right) \in \mathbb{R}^{N\times L}$ under each quantile.

3 System Framework

In order to improve the feature extraction ability of the model for historical data, this paper designs the PM^{2}FN for spectrum prediction. The model extracts the local timing features of the data by Convolutional Neural Network (CNN) and Bi-directional Long Short-Term Memory (BiLSTM) network, and the global timing features of the data by Sequence to Sequence (Seq2Seq) network with attention mechanism. After the two models learn the data in parallel, the output values are stitched into a high-dimensional vector. The parallel fusion network increases the connection between global and local, enabling the full use of input data and thus high accuracy prediction.

3.1 CNN-BiLSTM

BiLSTM network is mainly composed of two LSTMs with opposite data propagation directions to achieve bidirectional learning of historical data in the past and data in the future moment. There is a certain correlation between multi-channel and multi-time slot spectral data. BiLSTM network is good at finding and exploiting the intrinsic pattern between data to extract the spectral bi-directional correlation, but its performance will be degraded by overfitting when the number of features is large. CNN has the features of capturing local information and extracting depth features, which optimize the input BiLSTM model to solve the long series dependence problem. Therefore, this paper designs a CNN and BiLSTM tandem network to achieve potential information mining of multi-channel multi-time slot spectrum data.

The basic structure of CNN consists of an input layer, a convolutional layer, a pooling layer, a Flatten layer and an output layer [20]. By setting the alternating structure of convolutional and pooling layers, the effective local features of the data are acquired and the resulting feature vector is output. In this paper, three layers of convolutional units are set, and relu is used as the activation function. Lastly, a Flatten layer is added to facilitate the subsequent model addition. By using the local connectivity of CNN network structure and multi-layer structure, the local relevance of data is identified and low-level features are added, and as the CNN network deepens, the low-level features are combined into multi-layer features and deeper features are mined. The BiLSTM network is accessed after the CNN network to capture the local and temporal features of the spectrum.

3.2 Seq2Seq Architecture with Attention Mechanism

The Seq2Seq architecture is a typical encoder-decoder structure. We use BiLSTM as the encoder-decoder part of the network structure, and use the encoder embedded in the BiLSTM network to encode the input sequence, and map the output to a feature vector C characterizing the information of the input sequence to pass to the decoder.

$$\boldsymbol{C} = d(\boldsymbol{U}\boldsymbol{h}_T) \tag{1}$$

where $\boldsymbol{U}$ is the weight matrix, $\boldsymbol{h}_T$ is the encoder output at the last moment T, and $d(\cdot)$ is the activation function. At each moment, the decoder BiLSTM receives 3 sets of inputs, i.e., feature vector **C**, decoder output y_{t-1} of the previous moment and decoder hidden layer state h_{t-1} of the previous moment.

For the traditional Seq2Seq model, the encoder needs to compress all the hidden layer representations into fixed-length content vectors, which leads to a gradual decline in the temporal mining capability of the whole model as the length of the input temporal data increases. To address this problem, this paper sets the Attention layer as the connection between the encoder and the decoder, and the BiLSTM network reads the input data and updates the hidden layer by the following formula.

$$\alpha_{i,j} = \frac{\exp\left(d_{i-1}^T h_j\right)}{\sum_{k=1}^{T} \exp\left(d_{i-1}^T h_k\right)} \tag{2}$$

$$h_\alpha = \sum_{j=1}^{T} \alpha_{i,j} h_j \tag{3}$$

where d_{i-1} is the hidden layer of the BiLSTM decoder and $e_{i,j}$ is the result of the dot product operation of the encoder's output h_j and d_{i-1}. The calculated h_α is used as the weighted sum of all hidden layer outputs of the encoder (Fig. 1).

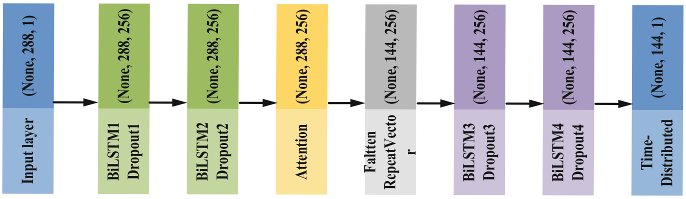


Fig. 1. Structure diagram of Seq2Seq network with attention mechanism.

3.3 Overall Model Architecture

In this paper, we use both CNN-BiLSTM and Seq2Seq with attention mechanism to form a deep network, and the overall architecture of model is shown in Fig. 2. The CNN-BiLSTM feature extraction network can effectively extract local time-series bidirectional features of the spectral data. The other network adds an attention mechanism to the Seq2Seq network that can deeply mine the temporal information, capturing the dependencies in long sequences and mining the global important information, increasing the control of the global situation. The two-way bilinear parallel structure can achieve high accuracy prediction by maximizing the retention of effective representation of local features and global time series features.

The PM^{2}FN consists of an input module, a parallel feature extraction module, and a feature fusion output module, where the input data of the input module is the spectrum occupancy data with channel number n and time slot length t. The original occupancy data needs to be dimensionally changed to adapt to the different input dimensions of the two-way feature extraction network before feeding it into the feature extraction module. The network is composed of 3 layers of 3×3 convolution and 2×2 maxpooling with a step size of 2. The two networks are connected in series to extract the initial local and temporal features. The other network adopts the BiLSTM-based Seq2Seq architecture, using the 2-layer BiLSTM as the codec unit for extracting higher-level features, while using the Attention mechanism layer for global search to realize the mining of global time-series features. Parallel two-way features are fused by a splicing method. The input features are passed through the fully connected layer to obtain the prediction data with the output matrix dimension $n \times l$. In addition, the quantile regression is combined with the above model to obtain the spectral probability prediction results.

The accurate prediction training process usually involves optimally minimizing the mean square error using algorithms such as gradient descent to estimate the network

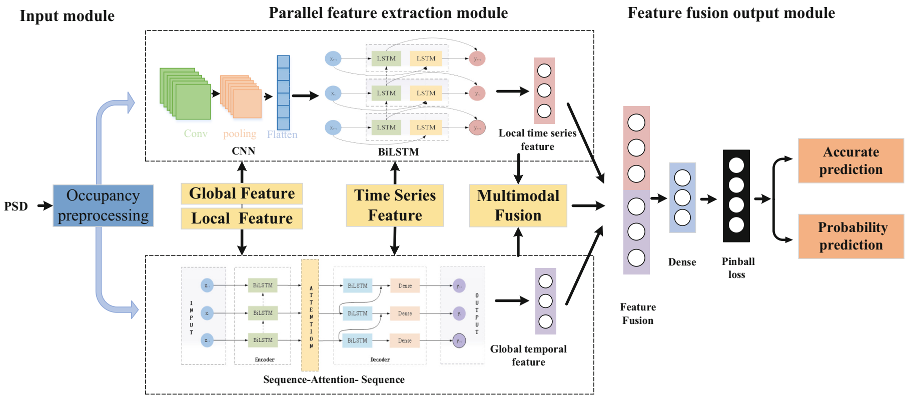


Fig. 2. System framework, the model consists of three parts: the input module, the parallel feature extraction module and the feature fusion output module.

weights that make the model regression work best, as shown in the following equation.

$$\min_{\boldsymbol{w},U} L(\boldsymbol{y},\hat{\boldsymbol{y}}) = \frac{1}{T}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2 \tag{4}$$

where $L(\boldsymbol{y},\hat{\boldsymbol{y}})$ is the loss function, T is the total number of samples, y_i is the true value, $\hat{y}_i$ and is the predicted value.

The quantile regression can be set by different quantile points to obtain the robustness estimates of the corresponding conditional quantile, satisfying the asymptotic goodness under the large sample theory. For the response variable Y, influenced by k factors X_1, X_2, ..., X_k, there are quantile regression models as follows.

$$Q_Y(\tau|\boldsymbol{X}) = \beta_0^\tau + \beta_1^\tau X_1 + \ldots + \beta_k^\tau X_k = \boldsymbol{X}^T\beta^\tau \tag{5}$$

where $Q_Y(\tau|\boldsymbol{X})$ denotes the τth conditional quantile of the response variable Y under the dependent variable $\boldsymbol{X} = [X_1, X_2, \ldots, X_k]^{\mathrm{T}}$; $\boldsymbol{\beta}^\tau = \left[\beta_0^\tau, \beta_1^\tau, \beta_2^\tau, \ldots, \beta_k^\tau\right]^{\mathrm{T}}$ denotes the vector of regression coefficients at the τth quantile. According to the quantile regression theory, the minimization objective function of (4) is transformed into a quantile pinball loss function of the form as (6) to establish the quantile regression of the model.

$$L_\tau\left(\boldsymbol{y},\hat{Q}_y(\tau)\right) = \frac{1}{T}\sum_{i=1}^{T}\rho_\tau\left(y_i - \hat{Q}_{y_i}(\tau)\right) = \frac{1}{T}\sum_{i=1}^{T}\left[\tau - I\left(y_i - \hat{Q}_{y_i}(\tau)\right)\right]\left(y_i - \hat{Q}_{y_i}(\tau)\right) \tag{6}$$

where $\hat{Q}_{y_i}(\tau)$ is the conditional quantile estimate of a sample, $\rho_\tau()$ is the test function, and $I()$ is the indicator function.

4 Experiments

The dataset of this paper was selected from open source electromagnetic spectrum data collected in Aachen, Germany[21], the dataset was collected in July 2008 with 14 days of data, covering the frequency range from 20 MHz to 6 GHz with a time domain resolution of 1.8 s. The experiments were conducted in both accurate spectrum occupancy prediction and random spectrum occupancy probability prediction.

4.1 Accurate Spectrum Occupancy Prediction

We choose two typical service bands as the target bands for spectrum prediction, including GSM and ISM. We count the occupancy every 10 min to predict the data for the future day. It is partitioned into a training set and a test set in the ratio of 7:3. The number of neurons in the hidden layer of the proposed BiLSTM in this paper is 128, and the Dropout is set to 0.1. The optimization algorithm used is the Adam algorithm, the learning rate is set to 0.001, and the epoch is set to 100. We adopt two widely used regression task criteria to evaluate the proposed model, including Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

In addition to the PM^{2}FN, the prediction performance of six baselines were also compared, which in turn proved the advantage of the network in spectrum prediction.

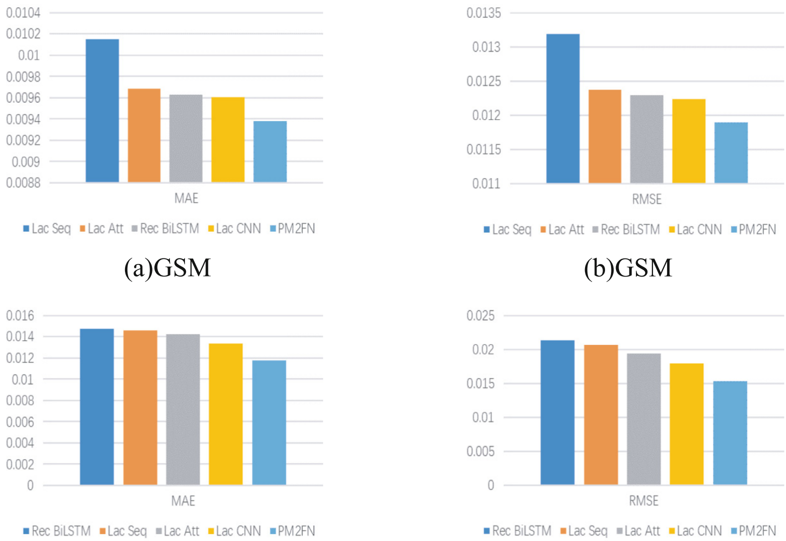
Table 1. Performance of different methods for prediction.

Model	GSM		ISM	
	MAE	RMSE	MAE	RMSE
HA	0.02403	0.02906	0.02503	0.05277
ARIME	0.01510	0.02320	0.02150	0.04030
LSTM	0.01209	0.01606	0.01890	0.03513
BiLSTM	0.01101	0.01436	0.01833	0.02812
CNN-LSTM	0.01079	0.01406	0.01897	0.03110
Seq2Seq-Attention	0.01067	0.01392	0.01696	0.02540
PM^{2}FN	**0.00938**	**0.01190**	**0.01175**	**0.01529**

Table 1 lists the results of comparing the PM^{2}FN with the baseline models, it can be found that the PM^{2}FN shows the best prediction results. Taking GSM band as an example, the PM^{2}FN improves the evaluation indexes MAE and RMSE by 12.1% and 14.5% on average than the best performing baseline model Seq2Seq-Attention. It is worth noting that the ARIMA and Historical Average (HA) prediction performance differs significantly from that of the deep learning model, demonstrating that the linear prediction model has poor prediction ability for nonlinear data. Meanwhile, deep learning model is compared, and the combined model shows better prediction ability than the single model. The models with Seq2Seq module have the ability to dig deeper into temporal correlations and thus show better performance in the service bands typically associated with human activities in GSM and ISM, while the models with the added attention mechanism get more attention to the temporal global important features and also play a positive role in the prediction performance. The PM^{2}FN can maximize the effective representation of local and temporal features after extracting features using the parallel network module, which enhances the learning ability of the model and makes the prediction more accurate; In general, the GSM band data exhibits obvious diurnal activity pattern, and its prediction effect is superior to that of the ISM band data.

4.2 Ablation Experiments

We performed the ablation analysis of the PM^{2}FN in GSM and ISM bands. The lack of CNN, lack of Seq2Seq, lack of attention mechanism, and replacement of BiLSTM with LSTM experimental groups are set respectively, and Fig. 3 shows the analysis results. Clearly, PM^{2}FN outperforms the prediction results of several models mentioned above in both service bands. This indicates that the spectrum prediction accuracy is improved by using the CNN module to extract features, and the parallel network is able to obtain both temporal and local features, which enhances the learning ability of the model and makes the prediction more accurate; The codec structure of Seq2Seq is capable of deep extraction of temporal features, and the global capture of changing features by attention contributes greatly to the prediction performance, and the bi-directional learning structure of BiLSTM is more beneficial for prediction of time-dependent spectral data. The model lacking Seq2Seq in GSM band is the worst model for prediction, which means that the encoder and decoder structures have the greatest impact on the prediction of temporal data, and the model replacing BiLSTM with LSTM in ISM band is the worst model for prediction, which means that the bidirectional learning structure is able to deeply mine data with insignificant temporal correlation. In conclusion, each component of the PM^{2}FN all have positive effects on improving the prediction effectiveness of the model.



(a)GSM (b)GSM

(c)ISM (d)ISM

Fig. 3. Error comparison.

4.3 Spectrum Occupancy Probability Prediction

Spectrum occupancy is volatile and stochastic, and GSM band data fluctuations are more influenced by human factors, so the degree of uncertainty is even more pronounced. However, probability prediction of the spectrum can quantify the uncertainty of the data and provide more decision-useful information for analyzing the spectrum change pattern. To further investigate the randomness characteristics of spectrum data, we use quantile regression combined with the PM^{2}FN (QPM^{2}FN) for spectrum probability prediction, and investigate the probability prediction results at 95%, 90% and 80% confidence levels. We also compare with QSeq2Seq-Attention, QLac-CNN, and QLac-Attention quantile regression models. Meanwhile, we choose Probabilistic Interval Coverage Percentage (PICP), Probabilistic Interval Average Width (PIAW) and Winkler Score (WS) as probabilistic forecasting evaluation metrics The remaining parameters are set in the same way as above.

Figure 4 shows the probabilistic prediction results of the QPM^{2}FN for both frequency bands, where the blue line is the real value and the blue areas at different depths are the prediction intervals of 95%, 90% and 85%. It can be seen from the figure that the width of the prediction interval at 95% confidence level is significantly larger than that of the prediction interval at 80% confidence level, except for individual real values that are not covered by the prediction interval, the prediction intervals at other moments contain the actual spectrum occupancy, the model can predict the future probability distribution of spectrum occupancy data. We calculate the confidence level as 80%, 90% and 95% prediction intervals respectively, the spectrum probability prediction interval can better characterize the fluctuation of spectrum and track the change of spectrum occupancy. When the actual spectrum occupancy is on the upward or downward trend, the range of the prediction interval is narrower and the uncertainty is smaller; When it is in the peak and trough, the range of the prediction interval is wider and the uncertainty is larger, which reflects that the spectrum is more influenced by uncertainty factors during this period. The method in this paper can reflect the fluctuation of spectrum occupancy data with time to the greatest extent.

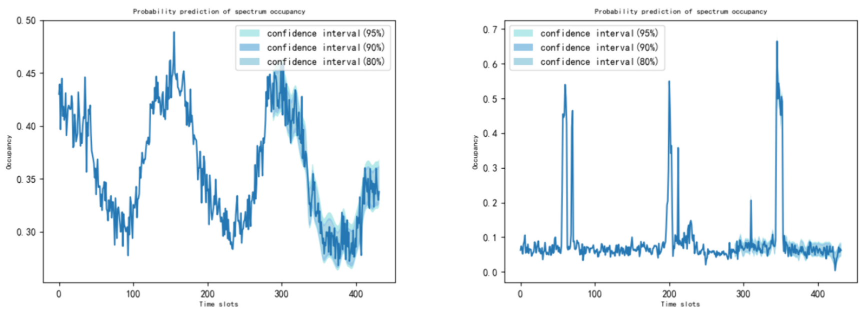


Fig. 4. Spectrum probability prediction diagram. (The left picture shows the GSM band, the right picture shows the ISM band.)

As can be seen from Table 2, the predicted interval coverage of the QPM^{2}FN and the comparison model in the same time slot frequency band are close to the corresponding confidence level results, and these methods are in line with the theoretical reality, However, the QPM^{2}FN has the highest coverage of probability intervals at the same confidence level, the narrowest average width of the intervals, and the smallest deviation between the expected and actual values of the prediction results. Since the prediction interval coverage often contradicts the conclusion between the interval coverage index and the interval coverage index, this paper introduces WS that concentrates on considering the two indexes to achieve a comprehensive evaluation of the reliability and sharpness of the probabilistic prediction results. The smaller the WS value, the better the probability prediction. The experimental results show that the QPM^{2}FN achieves the lowest scores of WS at different confidence levels, and therefore has higher probability accuracy. Taking the GSM band as an example, the WS value of the QPM^{2}FN are reduced by about 5.4%, 6.1% and 4.9% at 95%, 90% and 80% confidence levels, respectively, compared with the best performing QLac-CNN model.

Table 2. Performance of different probabilistic prediction methods

Level	Model	GSM			ISM		
		PICP%	PIAW%	WS	PICP%	PIAW%	WS
95%	QSeq2Seq-Attention	95.082	32.518	0.061	94.669	21.142	0.102
	QLac-Attention	95.225	29.592	0.057	95.110	17.147	0.088
	QLac-CNN	94.783	28.659	0.056	94.983	16.823	0.098
	QPM^{2}FN	**95.009**	**27.179**	**0.053**	**94.887**	**12.770**	**0.063**
90%	QSeq2Seq-Attention	90.138	27.158	0.053	89.704	17.391	0.088
	QLac-Attention	90.105	24.606	0.050	90.523	13.943	0.073
	QLac-CNN	89.625	24.066	0.049	89.803	13.516	0.078
	QPM^{2}FN	**90.127**	**23.043**	**0.046**	**89.760**	**10.633**	**0.054**
80%	QSeq2Seq-Attention	79.324	20.640	0.045	79.293	13.219	0.074
	QLac-Attention	80.396	19.255	0.042	79.768	10.352	0.059
	QLac-CNN	79.383	18.752	0.041	79.747	10.247	0.062
	QPM^{2}FN	**79.991**	**18.007**	**0.039**	**79.724**	**8.295**	**0.046**

To further demonstrate the superiority of the QPM^{2}FN, in this paper, we randomly select three frequency points in GSM and ISM bands respectively, and use kernel density estimation to obtain probability density curves for frequency point occupancy prediction after obtaining each conditional quantile by QR. As can be seen in Fig. 5, all the real values fall in the probability density curve. More specifically, most of the real data are distributed at the peaks of the probability density curve, except for the actual values at individual frequency points. The above shows that the QPM^{2}FN can give probability density curves for future predicted time points and conforms to the general rule. Based on the above analysis and discussion, it can be inferred that the model in this paper has the advantages of accuracy, comprehensiveness, stable performance, and effective prediction.

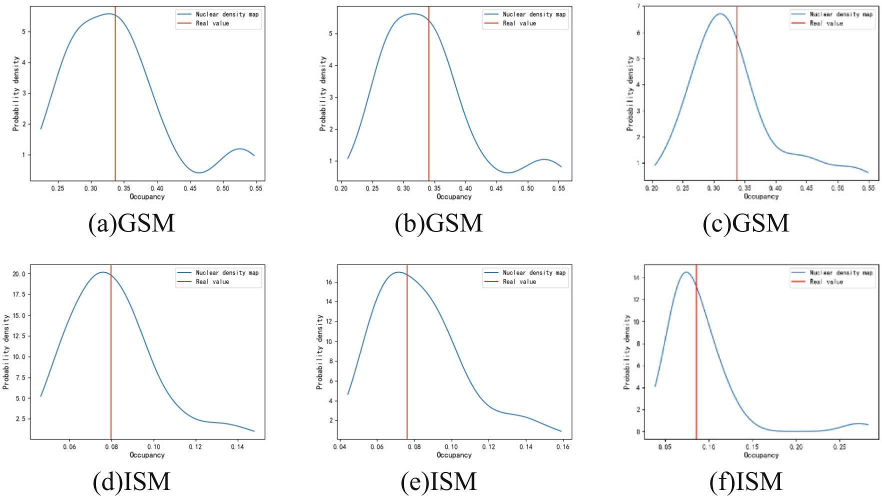


Fig. 5. The spectral probability density distribution, where the red line is the real value.

5 Conclusion

To address the problem of low accuracy of multi-channel spectrum prediction, we propose a parallel multi-channel multi-model fusion network (PM^{2}FN), which extracts global and local timing features and performs feature fusion through parallel structure to achieve accurate prediction of multi-point spectrum data. The present model is also combined with quantile regression to complete the description of the stochastic characteristics of the spectral activity law and to verify the superiority of the model proposed in this paper. The study of spectrum activity patterns plays an essential part for controlling spectrum utilization in complex electromagnetic environments. This paper does not consider the influence of other input features, such as holidays, on the spectral activity law, which will be the next step in the research.

Acknowledgement. This work is supported by the National Natural Science Foundation of China (No: 62201172), the National Key Research and Development Program of China (2022YFE0136800). This work is also supported by Key Laboratory of Advanced Marine Communication and Information Technology, Ministry of Industry and In-formation Technology, Harbin Engineering University, Harbin, China.

References

1. Mitola, J., Maguire, G.Q.: Cognitive radio: making software radios more personal. IEEE Pers. Commun. **6**(4), 13–18 (1999)
2. Wang, Z., Salous, S.: Spectrum occupancy statistics and time series models for cognitive radio. J. Signal Process. Syst. **62**(2), 145–155 (2011)
3. Chen, Z., Qiu, R.: Prediction of channel state for cognitive radio using higher-order hidden Markov model. Proceedings of the IEEE SoutheastCon pp. 276–282 IEEE (2010)
4. Wang, Y., Xiao, D., Lin, Y.: Mining association rules for electromagnetic spectrum data. Chin. J. Radio Sci. **37**(5), 802–809 (2022)
5. Guo, L., Wang, M., Lin, Y.: Electromagnetic environment portrait based on big data mining. Wirel. Commun. Mobile Comput. **2021**, 1–13 (2021)
6. Tu, Y., Lin, Y., Zha, H., et al.: Large-scale real-world radio signal recognition with deep learning. Chin. J. Aeronaut. **35**(9), 35–48 (2022)
7. Shamsi, N., Mousavinia, A., Amirpour, H.: A channel state prediction for multi-secondary users in a cognitive radio based on neural network. 2013 international conference on electronics, computer and computation (ICECCO), pp. 200–203 (2013)
8. Yu, L., Chen, J., Ding G.: Spectrum prediction via long short-term memory. In: 2017 3rd IEEE International Conference on Computer and Communications (ICCC), pp. 643–647 (2017)
9. Zhang, H., Peng, S., Zhang, J., et al.: Big data analysis and prediction of electromagnetic spectrum resources: a graph approach. Sustainability **15**(1), 508 (2023)
10. Li, S., Liu, H., Guo, L.: Information mining and association analysis based on electromagnetic environment data. J. Terah. Sci. Elect. Inform. Technol. **20**(1), 8–15 (2022)
11. Li, K., Li, C., Chen, J., et al.: Boost spectrum prediction with temporal-frequency fusion network via transfer learning. IEEE Trans. Mobile Comput. **22**(6), 3209 (2021)
12. Aygül, M.A., Nazzal, M., Sağlam, M.İ, et al.: Efficient spectrum occupancy prediction exploiting multidimensional correlations through composite 2D-LSTM models. Sensors **21**(1), 135 (2020)
13. Shawel, B.S., Woldegebreal, D.H., Pollin, S.: Convolutional LSTM-based long-term spectrum prediction for dynamic spectrum access. In: Proceedings of the 27th European Signal Processing Conference (EUSIPCO), pp. 1–5. (2019)
14. Zhao, C., Wan, C., Song, Y.: Optimal nonparametric prediction intervals of electricity load. IEEE Trans. Power Syst. **35**(3), 2467–2470 (2020)
15. Koenker, R., Hallock, K.F.: Quantile regression. J. Econom. Perspect. **15**(4), 202 (2001)
16. He, Y., Liu, R., Li, H., et al.: Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory. Appl. Energy **185**, 254–266 (2017)
17. Cannon, A.J.: Quantile regression neural networks: Implementation in R and application to precipitation downscaling. Comput. Geosci. **37**(9), 1277–1284 (2011)
18. Wang, J.: Bayesian quantile regression for parametric nonlinear mixed effects models. Statist Meth Appl **21**(3), 279–295 (2012)
19. Yu, Y., Han, X., Yang, M.: Probabilistic prediction of regional wind power based on spatiotemporal quantile regression. IEEE Trans. Ind. Appl. **29**, 1–16 (2019)
20. Liu, B., Nowotarski, J., Hong, T.: Probabilistic load forecasting via quantile regression averaging on sister forecasts. IEEE Trans. Smart Grid **8**(2), 730–737 (2017)
21. Wellens, M., Mähönen, P.: Lessons learned from an extensive spectrum occupancy measurement campaign and a stochastic duty cycle model. Mobile Netw. Appl. **15**, 461–474 (2010)