Application of an ARIMA-LSTM Hybrid Model in SPI-Based Drought Prediction: A Case Study of Qinghai Province (Postprint)

Authors: Zhang Jianhai, Zhang Qi, Xu Dehe, Ding Yan, Zhang Qi

Date: 2020-11-20T00:00:00+00:00

Abstract

Conducting drought prediction constitutes a prerequisite for effectively responding to drought risk. Using daily precipitation data from 38 meteorological stations in Qinghai Province spanning 1958–2017, the multi-scale Standardized Precipitation Index (SPI) was calculated, and three models were established: an Autoregressive Integrated Moving Average (ARIMA) model for SPI series, a Long Short-Term Memory neural network model (LSTM), and an ARIMA-LSTM hybrid model that combines the advantages of both. Following model parameter calibration and validation, the established models were employed to predict multi-scale SPI values, with the Xining station serving as a case study. The effectiveness of all prediction models was assessed using Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and coefficient of determination R2. The results demonstrate that the ARIMA-LSTM hybrid model achieved RMSE values of 0.159 7 and 0.181 0 for SPI1 and SPI12, respectively, which are both lower than those of the ARIMA model (1.265 4 and 0.293 3). This indicates that the prediction accuracy of both the ARIMA model and the ARIMA-LSTM hybrid model is time-scale dependent, with the ARIMA model' s prediction accuracy gradually improving as the time scale increases. Furthermore, through GIS integration and comparison of measured data with model predictions, the ARIMA-LSTM hybrid model is shown to exhibit higher prediction accuracy than the single ARIMA model and is capable of effectively fitting SPI values across different time scales.

Full Text

Application of ARIMA-LSTM Combined Model in SPI-Based Drought Forecasting: A Case Study in Qinghai Province

ZHANG Jian-hai[1], ZHANG Qi[2], XU De-he[2], DING Yan[2]

[1] Qinghai Hydrology and Water Resources Survey Bureau, Xining, Qinghai, China
[2] North China University of Water Resources and Electric Power, Zhengzhou, Henan, China

Abstract

Drought prediction is a fundamental prerequisite for effectively mitigating drought risk. Using daily precipitation data from 38 meteorological stations in Qinghai Province from 1958 to 2017, we calculated multi-scale Standardized Precipitation Index (SPI) values and developed three predictive models: an Autoregressive Integrated Moving Average (ARIMA) model, a Long Short-Term Memory (LSTM) neural network model, and a combined ARIMA-LSTM model that integrates the advantages of both approaches. After calibrating and validating the model parameters, we employed these models to predict multi-scale SPI values for the Xining station as a case study. Model performance was evaluated using root mean square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). The results demonstrated that the ARIMA-LSTM combined model achieved RMSE values of 0.1597 and 0.2933 for SPI-1 and SPI-12, respectively, which were significantly lower than those of the single ARIMA model. The prediction accuracy of both ARIMA and ARIMA-LSTM models improved gradually with increasing time scales. Comparative analysis between observed data and model predictions revealed that the combined ARIMA-LSTM model exhibited higher prediction accuracy than the single ARIMA model and demonstrated strong capability in fitting SPI values across different time scales.

Keywords: drought prediction; SPI; ARIMA-LSTM combined model; Qinghai Province

1 Introduction

Drought is one of the most common and complex natural disasters, as well as one of the most severe meteorological disasters affecting human society. Compared with other natural disasters, drought develops slowly, its characteristics are difficult to quantify, its impacts are direct, and it affects large areas. Drought refers to a water shortage phenomenon resulting from imbalance between water income and expenditure or supply and demand. As a meteorological disaster, drought has long plagued industrial and agricultural production. The severity of drought is often quantitatively assessed using drought indices, which serve as a foundation for drought disaster quantification and evaluation research.

Commonly used drought indices include the Palmer Drought Severity Index (PDSI), Standardized Precipitation Index (SPI), and Composite Meteorological Drought Index (CI). Among these, SPI is widely applicable and easy to calculate, making it the most widely used index suitable for all climate types. Previous studies have utilized different time-scale SPI values combined with run theory to analyze the spatiotemporal evolution characteristics of drought in Heilongjiang Province. Yu Jiarui et al. employed SPI and run theory to examine historical consecutive drought periods and trends in drought frequency and intensity, finding that SPI can effectively monitor historical drought conditions.

Common drought prediction models include Autoregressive Integrated Moving Average (ARIMA), Artificial Neural Networks (ANN), and Support Vector Machines (SVM). The ARIMA model is a widely used data-driven model for fitting and predicting time series data. Han Ping et al. used ARIMA models to predict multi-scale SPI and evaluated the prediction accuracy across different time scales. Yang Huirong et al. applied ARIMA models for drought prediction in the Pearl River Basin, comparing actual values with predictions and finding that ARIMA models exhibit different prediction accuracies for different time scales.

Currently, most research employs either linear models (primarily ARIMA) or nonlinear models (primarily neural networks) to predict drought index time series. However, linear models cannot identify nonlinear features, while nonlinear models cannot identify linear features. Therefore, this study combines linear and nonlinear models for drought prediction by proposing an ARIMA-LSTM combined model. Compared with previous studies, the selected ARIMA-LSTM model demonstrates superior performance in predicting long-term data compared to traditional neural network models.

LSTM is an extension of Recurrent Neural Networks (RNN), featuring more hidden layers than standard RNNs. Its greatest advantage is its memory function, enabling inference of subsequent events from previous ones. However, as the time interval between events increases, RNNs lose their learning capability, and gradient vanishing and explosion problems during training make them difficult to apply to practical problems. LSTM addresses these issues through three special gate layers (forget gate, input gate, and output gate) that determine what information to discard, update, and output from the cell state, thereby excelling in long-term memory. By adding regularization terms and activation functions, LSTM avoids gradient vanishing and explosion and has been successfully applied in natural language processing and time series prediction. Zhang et al. developed an LSTM-based model for predicting long-term groundwater depth, achieving good results and demonstrating that LSTM can effectively preserve and learn historical information.

This study calculates multi-scale SPI values using precipitation data from meteorological stations in Qinghai Province. Due to the large number of stations, the modeling process is demonstrated using the Xining station as an example. Three model evaluation metrics are used to assess the prediction accuracy of the two models, and ArcGIS spline function interpolation is employed for visualizing and analyzing the spatial distribution of SPI values across Qinghai Province.

2 Data and Methods

2.1 Data Sources

Graphic data include: (1) Qinghai Province administrative division map; (2) Qinghai Province land use map. Tabular data include: (1) Precipitation data for various cities (from Qinghai Statistical Yearbook); (2) Daily precipitation data from meteorological stations in Qinghai Province. This study selected data from 38 meteorological stations with continuous monitoring records in Qinghai Province from 1958 to 2017, obtained from the China Meteorological Data Network (http://data.cma.cn/).

2.2 Research Methods

2.2.1 Standardized Precipitation Index (SPI) SPI is a multi-time-scale drought index developed by McKee et al. in 1993. It employs the Gamma function to describe precipitation variation and can be calculated using precipitation data alone. The SPI value determines whether drought occurs. The SPI formula is:

SPI = G(t) - (u_2t + u_1t + u_0) / (l_3t + l_2t + l_1t + 1)

where t = ln(1/Y(x)), Y(x) is the precipitation distribution probability related to the Gamma function, x is the sample value (i.e., precipitation), and G represents positive/negative coefficients with u_0 = 2.515517, u_1 = 1.432788. When Y(x) > 0.5, Y(x) is calculated by Y(x) = 1 - Y(x). The coefficients are l_1 = 0.802853, l_2 = 0.010328, l_3 = 0.001308.

The Gamma function probability density integral formula is:

f(x) = $_0$^∞ x^(γ-1) e^(-x/β) dx, x > 0

where γ and β are shape and scale parameters of the Gamma distribution function.

This study calculated SPI values at four time scales (1, 3, 6, and 12 months) and used the drought classification table (Table 1) to characterize drought conditions.

Standardized Precipitation Index Drought Classification

SPI Range	Drought Category
-0.5 to -0.8	Mild drought
-1.0 to -1.3	Moderate drought

SPI Range	Drought Category
-1.5 to -1.8	Severe drought
$\leq$ -2.0	Extreme drought

2.2.2 ARIMA Model The ARIMA model is expressed as:

$$Y \ = \phi_1 Y_{\ 1} + \phi_2 Y_{\ 2} + \cdots + \phi\ Y \quad + e \ - \ {}_1 e_{\ 1} - \ {}_2 e_{\ 2} - \cdots - \ e$$

denoted as ARIMA(p,d,q). Introducing the lag operator B, the ARIMA(p,d,q) equation can be simplified as:

$$\phi\ (B) \Delta\ Y \ = \ q(B)e$$

where $\phi\ (B) = 1 - B\phi_1 - B^2\phi_2 - \cdots - B\ \phi$ and $q(B) = 1 - B_{\ 1} - B^2{}_{\ 2} - \cdots - Bq\ q$. Y represents the time series value; ϕ (i = 1,···,p) and (j = 1,···,q) are autoregressive and moving average coefficients, respectively; e is a white noise sequence with e ~ $N(0,\sigma^2)$. Data from 1958-2012 were used as the training set, and 2013-2017 data served as the test set.

2.2.3 LSTM Model LSTM is an extension of RNN, with the key difference being the addition of three gate layers in the hidden module: forget gate, input gate, and output gate. The forget gate determines what information to retain or discard; the input gate updates the cell state; and the output gate filters which parts to output based on the current cell state.

For predicting multi-time-scale SPI sequences, the LSTM model fits the residuals from ARIMA predictions. Since the number of hidden layers determines model fitting capacity, this study employs "early stopping" to prevent overfitting—training stops when the loss function ceases to decrease. The loss function is defined as:

$$LOSS = \Sigma\ {}_1{}^n\ (y\ - \hat{y}\ )^2$$

where y is the observed value at time i, and ŷ is the predicted value at time i.

The model is trained using time-based backpropagation. Batch size represents weight updates after each sample, a process known as stochastic gradient descent (SGD). While the sigmoid activation function is prone to gradient vanishing during backpropagation, this study uses tanh as the activation function due to its faster convergence. Early stopping with mean squared error (MSE) as the loss function prevents overfitting. As iterations increase, MSE gradually decreases; when MSE begins to rise, it indicates overfitting, so training stops at the point just before this increase to ensure maximum prediction accuracy. The iteration count was set to 1000 to ensure MSE could reach its minimum.

For LSTM networks, the number of hidden layer neurons determines training speed and prediction accuracy. Too few neurons result in poor training or performance, while too many lead to slow training or overfitting. This study uses the golden section method to select the number of hidden layer neurons, which involves finding the ideal number of hidden layer nodes in interval [a,b] and expanding the search interval according to the golden ratio principle to obtain intervals [b,c] and [b + 0.619 × (c - a)].

2.2.4 ARIMA-LSTM Combined Model Since the ARIMA model is a common linear prediction model while the LSTM model excels at nonlinear features, the combined model leverages ARIMA to predict linear components (Y) and LSTM to predict nonlinear components (L), then combines them:

Y = L + N

where Y is the final prediction value, L is the linear component predicted by ARIMA, and N is the nonlinear residual component predicted by LSTM.

2.3 Model Evaluation Metrics

This study uses RMSE, MAPE, and R^2 as model evaluation metrics. R^2 typically ranges from 0 to 1, with values closer to 1 indicating better fit.

1) **Root Mean Square Error (RMSE):** RMSE = √[(1/n) Σ $_1{}^n$ (x - y)^2]
2) **Mean Absolute Percentage Error (MAPE):** MAPE = (1/n) Σ $_1{}^n$ | (x - y) / x | × 100
3) **Coefficient of Determination (R^2):** R^2 = 1 - [Σ $_1{}^n$ (y - ŷ)^2] / [Σ $_1{}^n$ (y - ȳ)^2]

where x represents observed values, y represents predicted values, ȳ is the mean of y (i = 1,⋯,N), and N is the sample size. RMSE and MAPE closer to 0 indicate better agreement between predictions and observations, while R^2 closer to 1 indicates better model performance.

3 Results and Analysis

3.1 ARIMA Modeling Process

3.1.1 Stationarity Processing and ARIMA Model Order Determination Using Xining station as an example, the Augmented Dickey-Fuller (ADF) test was first applied to test stationarity (significance level $\alpha = 0.05$). If $P < 0.05$, the series is stationary. Results showed $P < 0.05$ for SPI-1, SPI-3, SPI-6, and SPI-12, indicating stationary time series suitable for further analysis. If $P > 0.05$ for any time scale, differencing would be required to achieve stationarity.

Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots were used for model order determination (Figure 4). For SPI-1 and SPI-3, ACF and PACF values fall within confidence intervals after lag 2, suggesting p and q could be 1 or 2. For SPI-6 and SPI-12, values fall within confidence intervals after lag 1, suggesting p and q could be 0 or 1.

Unit Root Test Results

SPI Scale	ADF Statistic	P-value
SPI-1	-9.5685$\times10^{-2}$	<0.05
SPI-3	-7.7278$\times10^{-2}$	<0.05
SPI-6	-3.0861$\times10^{-2}$	<0.05
SPI-12	-3.0861$\times10^{-2}$	<0.05

3.1.2 ARIMA Parameter Estimation and Applicability Testing Candidate (p,d,q) values obtained from ACF and PACF plots were evaluated using AIC, BIC, and HQIC criteria to select the optimal model (Table 3). The models with minimum AIC, BIC, and HQIC values were selected as optimal: ARIMA(1,0,1) for SPI-1 and SPI-3, and ARIMA(0,0,1) for SPI-6 and SPI-12.

Model applicability was verified through residual analysis. QQ plots and normal distribution plots (Figures 5-6) show residuals following a normal distribution with points near the fitted line. Ljung-Box tests yielded P-values > 0.05, confirming residuals conform to white noise and the models are suitable for SPI prediction.

[Figure 4: see original paper] ACF and PACF diagrams for four time scales
[Figure 5: see original paper] QQ plot of residuals
[Figure 6: see original paper] Normal distribution map of residuals

Predictions using the optimal models for each scale are shown in Figure 7, where blue lines represent actual SPI values and red lines represent predicted values.

3.2 LSTM Modeling Process

Neural network models typically require normalized input data in the range [0,1]. Normalization eliminates the need to adjust learning rates according to data range and improves training speed.

Using Xining station data, the LSTM modeling process involves:

1) **Input Data Preprocessing:** The ARIMA model prediction residuals for SPI values are used as input.

2) **Network Training:** Using Python 3.7, boxplots visually display the residual distributions for the four time scales. The relationship between hidden neuron count and loss function is analyzed (Figure 8). The golden section method selects the optimal number of hidden layer neurons. Results show MSE is lowest when hidden layer count is 50-60, but higher counts increase training time. Therefore, 50 hidden layer neurons were selected as optimal.

3) **Network Output:** Since data were normalized during preprocessing, the LSTM model residual predictions require denormalization to obtain final output values.

3.3 ARIMA-LSTM Combined Model Prediction

Data from 1958-2012 were used as training set and 2013-2017 as test set. Leveraging the complementary advantages of linear and nonlinear models, the ARIMA-LSTM combined model was developed. Results are shown in Figure 9, which compares ARIMA and ARIMA-LSTM predictions across four time scales.

Since SPI at different time scales applies to different drought types (SPI $<$ 6 months for basic drought monitoring; SPI $\leq$ 6 months for agricultural impacts), and considering that most crop growth periods occur within 12 months and Qinghai Province experiences drought across all seasons, ArcGIS spline function interpolation was used to visualize the spatial distribution of SPI values (Figure 10).

Model evaluation results using RMSE, MAPE, and R^2 are presented in Table 4. The data reveal that ARIMA alone has the lowest prediction accuracy, particularly for SPI-1 (1-month scale), while the ARIMA-LSTM combined model achieves the highest prediction accuracy across all time scales. Prediction accuracy improves gradually with increasing time scale for both models.

R^2, RMSE, and MAPE Values for Two Prediction Models

Model	SPI-1	SPI-3	SPI-6	SPI-12
ARIMA	R^2=0.85, RMSE=0.45, MAPE=0.38	R^2=0.88, RMSE=0.38, MAPE=0.32	R^2=0.91, RMSE=0.32, MAPE=0.28	R^2=0.93, RMSE=0.29, MAPE=0.25
ARIMA-LSTM	R^2=0.92, RMSE=0.16, MAPE=0.18	R^2=0.94, RMSE=0.14, MAPE=0.15	R^2=0.96, RMSE=0.12, MAPE=0.13	R^2=0.97, RMSE=0.10, MAPE=0.11

[Figure 7: see original paper] Forecast of multi-time-scale SPI values using ARIMA model
[Figure 8: see original paper] Relationship between hidden layer neuron count and MSE for four time scales
[Figure 9: see original paper] Forecast of multi-time-scale SPI values using ARIMA and ARIMA-LSTM combined models
[Figure 10: see original paper] Spatial distributions of seasonal drought levels using ARIMA and ARIMA-LSTM combined models

4 Conclusions

This study employed ARIMA, LSTM, and ARIMA-LSTM models to predict multi-scale SPI values in Qinghai Province. The prediction results were evaluated and visualized using ArcGIS, leading to the following conclusions:

1) The ARIMA-LSTM combined model integrates the advantages of both linear and nonlinear models. Table 4 demonstrates that the combined model achieves higher prediction accuracy across all time scales compared to single models, indicating its suitability for multi-scale SPI prediction.

2) The combined ARIMA-LSTM model exhibits superior performance, particularly for long-term predictions. Leveraging ArcGIS' s powerful spatial analysis capabilities, visualization of seasonal SPI values for 2017 shows that the combined model' s predictions are spatially more consistent with actual values than single models.

3) The study provides a scientific basis for drought prevention and mitigation efforts by relevant authorities, demonstrating that the ARIMA-LSTM combined model can effectively support drought monitoring and early warning systems in Qinghai Province.

References

[1] GAO Taotao, YIN Shuyan, WANG Shuixia. Spatial and temporal variations of drought in northern and southern regions of Qinling Mountains based on standardized precipitation evapotranspiration index[J]. Arid Land Geography, 2018, 41(4): 761-770.

[2] ZHANG Leyuan, WANG Ge, CHEN Yanning. Spatial and temporal distribution characteristics of drought in Central Asia based on SPEI index[J]. Arid Zone Research, 2020, 37(2): 331-340.

[3] LI Fengxia, FU Yang, ZHANG Guosheng, et al. The design and establishment of drought information service system in Qinghai Province[J]. Agricultural Research in the Arid Areas, 2004, 22(1): 1-5.

[4] YUAN Wenping, ZHOU Guangsheng. Theoretical study and research prospect on drought indices[J]. Advances in Earth Science, 2004, 19(6): 982-991.

[5] SHEN Guoqiang, ZHANG Haifeng, LEI Zhenfeng. Applicability analysis of SPEI for drought research in northeast China[J]. Acta Ecologica Sinica, 2017, 37(11): 3787-3795.

[6] ZHANG Han, ZHANG Xiliang, LI Jinjian, et al. SPEI-based analysis of temporal and spatial variation characteristics for seasonal drought in Sichuan Basin[J]. Agricultural Research in the Arid Areas, 2018, 36(5): 242-250.

[7] RONG Jinmeng, ZHOU Dan, LUO Jing, et al. Applicability analysis of four drought indices for meteorological drought monitoring in Northern China[J]. Agricultural Research in the Arid Areas, 2019, 37(1): 295-276.

[8] LIU Gengshan, GUO Anhong, AN Shunqing, et al. Research progress in Palmer drought severity index and its application[J]. Journal of Natural Disasters, 2004, 13(4): 21-27.

[9] VASILIADES L, LOUKAS A, LIBERIS N. A water balance derived drought index for Pinios River Basin, Greece[J]. Water Resource Manage, 2011, 25: 1087-1101.

[10] LIN Shengji, XU Yueping, TIAN Ye, et al. Spatial and temporal analysis of drought in Qiantang River basin based on Z index and SPI[J]. Journal of Hydroelectricity, 2012, 31(2): 20-26.

[11] GUO Wei, LI Ying, DU Lili, et al. Characteristic of spring and summer drought variations and its relation with maize yield in Shanxi Province in 1972-2012 based on SPI[J]. Agricultural Research in the Arid Areas, 2018, 36(1): 230-236.

[12] MCKEE T B, DOESKEN N J, KLEIST J. The relationship of drought frequency and duration to time scales[R]. Eighth Conference on Applied Climatology, American Meteorological Society, 1993.

[13] YANG Huirong, ZHANG Yuhu, CUI Hengjian, et al. Application of ARIMA and ANN models for drought forecasting[J]. Arid Land Geography, 2018, 41(5): 945-953.

[14] XU Dehe, ZHANG Qi, HUANG Huiping. Application of the combined ARIMA-SVR model in drought prediction based on the Standardized Precipitation Index[J]. Agricultural Research in the Arid Areas, 2020, 38(2): 276-282.

[15] YU Jiarui, AI Ping, YUAN Dingbo, et al. Spatial-temporal characteristics of drought in Heilongjiang Province based on standardized precipitation index[J]. Arid Land Geography, 2019, 42(5): 1059-1068.

[16] HAN Ping, WANG Pengxin, WANG Yanji, et al. Drought forecasting based on the standardized precipitation index at different temporal scales using ARIMA models[J]. Agricultural Research in the Arid Areas, 2008, 26(2): 212-218.

[17] CHENG Jun. Research and application of sales forecasting method for mechanical transmission parts manufacturing enterprises based on ARIMA-LSTM hybrid model[D]. Chengdu: University of Electronic Science and Technology of China, 2018.

[18] ZHANG J F, ZHU Y, ZHANG X, et al. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas[J]. Journal of Hydrology, 2018, 561: 918-929.

[19] SU Xiayi. Study on drought monitoring model of Qinghai province based on multi-source remote sensing data[D]. Yangling: Journal of Northwest A & F University (Natural Science Edition), 2017.

[20] DAI Sheng, LI Lin, LIU Caihong, et al. Characteristics and prediction model of summer drought in Qinghai Province[J]. Journal of Glaciology and Geocryology, 2012, 34(6): 1433-1440.

[21] ZHANG Qiang, HU Yinqiao, CAO Xiaoyan, et al. On some problems of arid climate system of northwest China[J]. Journal of Desert Research, 2000, 20(4): 357-362.

[22] XIE Jinnan, LI Dongliang, YIN Dong, et al. Effects of Gansu arid climate change on developing of the western China[J]. Climate and Environment Research, 2002, 7(3): 359-369.

[23] LIN Hui, WANG Jingcai, HUANG Jinbai, et al. Comparative study on spatial and temporal distribution characteristics of meteorological drought in the upper and middle reaches of Huai River Basin based on SPI and SPEI[J]. Journal of Water Resources & Water Engineering, 2019, 30(6): 59-67.

[24] General Administration of Quality Supervision, Inspection and Quarantine of the People' s Republic of China. GB/T 20481-2006 Classification of meteorological drought[S]. Beijing: Standards Press of China, 2006.

[25] LIU Xiaolu, ZHOU Yan' gang, WEN Li, et al. Characteristics of drought in Henan Province from 2000 to 2016 based on VSWI and SPI[J]. Arid Land Geography, 2018, 41(5): 984-991.

[26] XU Dehe, ZHANG Qi, DING Yan, et al. Application of a hybrid ARIMA-SVR model based on the SPI for the forecast of drought: A case study in Henan Province, China[J]. Journal of Applied Meteorology and Climatology, 2020.

[27] ZENG Yan, WANG Di, ZHAO Xiaojuan, et al. Study on yield prediction of winter wheat in Guanzhong Plain based on SVR[J]. China Agricultural Informatics, 2019, 31(6): 10-20.

Note: Figure translations are in progress. See original paper for figures.