Based on the Current Level of Economic Development, Predict PAHs Concentration in Sediment of Dianchi Lake in Southwest China From 2021 to 2030

Xiaohua Ma (xhma@niglas.ac.cn)
Nanjing Institute of Geography and Limnology Chinese Academy of Sciences

Hao Yang
Nanjing Normal University

Shuaidong Li
Nanjing Normal University

Changchun Huang
Nanjing Normal University

Tao Huang
Nanjing Normal University

Hongbin Wan
Nanjing Normal University

Research Article

Keywords: Polycyclic aromatic hydrocarbons, lacustrine sediment, economic parameters, neural network, Dianchi Lake.

Posted Date: February 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-165865/v1

Version of Record: A version of this preprint was published at Environmental Science and Pollution Research on August 11th, 2021. See the published version at https://doi.org/10.1007/s11356-021-15690-9.

Based on the current level of economic development, predict PAHs concentration in sediment of Dianchi Lake in southwest China from 2021 to 2030

Xiaohua Ma[a,b*], Hao Yang [b,c,d,e], Shuaidong Li [b,c], Changchun Huang [b,c,d,e], Tao Huang [b,c,d,e], and Hongbin Wan[b,c]

[a] *Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, PR China*

[b] *School of Geography Science, Nanjing Normal University, Nanjing 210023, PR China*

[c] *Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing Normal University, Nanjing 210023, PR China*

[d] *Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing 210023, PR China*

[e] *State Key Laboratory Cultivation Base of Geographical Environment Evolution (Jiangsu Province), Nanjing 210023, PR China*

**Corresponding author: Xiaohua Ma*

Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, PR China

Tel: +86-10-62732006; Fax: +86-10-62731016.

E-mail: xhma@niglas.ac.cn; 718301375@qq.com

Abstract: The analysis of the correlation between polycyclic aromatic hydrocarbons (PAHs) and economic parameters show that the total population, gross domestic product, coal consumption, and oil consumption have significant relationships with PAH concentrations in

Dianchi Lake, Yunnan province, China. An artificial neural network model was developed in order to predict the trend in PAHs concentration in the sediments of Dianchi Lake in the next ten years, based current indicators of economic development. The artificial neural network model estimated the concentration of PAHs for the period 1980 to 2014. The R^2 and RMSE of the training samples were 0.97 and 44.2, respectively; the R^2 and RMSE of the validation samples were 0.99 and 18.7, respectively the R^2 and RMSE of the testing samples were 0.99 and 13.2, respectively; and the R^2 and RMSE of all samples were 0.98 and 39.2, respectively. The concentration of PAHs in the sediments of Dianchi Lake predicted by the neural network model shows that the concentration of PAHs from 2021 to 2030 can reach as high as 2128.9 ng/g. The ecological environment of Dianchi Lake has a great impact on the lives of local residents, including through its support for the local economy. Therefore, it is necessary for the government to take more effective measures and increase investments to control the environmental problems in Dianchi Lake.

Keywords: Polycyclic aromatic hydrocarbons, lacustrine sediment, economic parameters, neural network, Dianchi Lake.

1. Introduction

China is the largest developing country in the world. In recent decades, rapid economic development, population growth, and urbanization and industrialization have resulted in a growing demand for coal, petroleum, and other energy sources. This has led to sharp increases in the concentrations of pollutants in Chinese lakes, which has aroused widespread concern in society (Qiu et al 2005, Zhi et al 2015).

This study focuses on 16 PAHs listed for priority control by the US Environmental Protection Agency (US EPA) due to their persistent and long-term toxicity (Pietzsch et al 2010, Yin et al 2008, Zhou et al 2018), and potential mutagenic (Viguri et al 2002, Zhi et al 2015, Nemr, et al 2007) and carcinogenic effects (Dushyant et al 2016, Li et al 2019, Sun et al 2016, Wang et al 2010).

PAHs can originate from natural processes such as forest fires (Freeman and Catell 1990), volcanic eruptions (Kim et al 2003, Morillo et al 2007), and diagenesis of organic matter in oxygen-deficient sediments (Baumard et al 1998, Van et al 2000); however, human activities such as garbage incineration (Mastral and Callén 2000), fossil fuel combustion (Blumer and Youngblood 1975; Khalili et al 1995), and wood burning for household heating and cooking (Zhang et al 2007) are generally considered to be the main source of PAHs entering the environment.

Dianchi Lake is the largest freshwater lake in Yunnan province and is an important source of water for agricultural and industrial production in Kunming, the largest city in the province. Since the founding of the People's Republic of China, and especially in its implementation of its reform and opening policy, human activities such as industrial wastewater discharge and garbage dumping associated with rapid urbanization and industrialization has greatly affected the water quality in the lake (Liu et al 2008, Gu et al 2017, Zeng and Wu 2009, Zhang et al 2015), making Dianchi Lake one of the most polluted lakes in China. The water quality of Dianchi Lake is directly related to the development of Dianchi's industry and impacts people's health. Research by Li et al (2003) showed that at the beginning of the 21st century, Dianchi

Lake had almost lost its ability to purify itself and had become one of the most polluted lakes in the country. The water pollution of Dianchi Lake has caused serious ecological imbalances, and a large number of species have died. The water in Dianchi Lake has almost lost various functions, directly affecting the water safety of Kunming and the irrigation of the 1.782×104 hm^2 farmland around the lake, and causing huge losses to the economic development and urban image of Kunming (Li et al 2003).

Some previous studies have used artificial neural networks to investigate the concentration of pollutants in lake sediments or soils (Anagu et al 2009, Buszewski and Kowalkowski 2006, Jia et al, 2018, Liuet al, 2010, Sari 2012, Shiet al 2015). An artificial neural network can detect the complex nonlinear relationship between the independent and dependent variables (Wösten et al 2001). The artificial neural network creates a complex formula representing the relationship between the input value and the output value (Maren et al 1990), which can be used in a similar way to the regression formula (Wösten et al 2001).

Lake sediments are important archives of environmental information (Thomas et al. 2012, Yuan et al 2017), and play the role of geochronometers, measuring the historical deposition of pollutants and socio-economic changes (Donahue et al 2006, Guo et al 2007, Kannan et al 2005, Latimer et al 2003, Yim et al 2005). This study, adopts Dianchi Lake, a typical plateau lake in southwest China, as an example to apply an artificial neural network to i) investigate the historical trends of PAH concentration in the Dianchi lake sediments; and ii) predict the trend in the concentration of PAHs in sediments from 2021 to 2030.

2. Materials and Methods

2.1 Sample Collection

Diianchi Lake (24°40′–25°03′ N, 102°37′–102°48′ E) is located in the central part of the Yunnan-Guizhou Plateau in southwest China. It is the largest freshwater lake on the Yunnan-Guizhou Plateau and the sixth largest freshwater lake in China. The lake is approximately 40 km long and 12.5 km wide (Du et al 2011), its surface area is 306 km^2, its average water depth is 4.7 m, and its watershed area is 2,920 km^2. Dianchi Lake is a semi-enclosed lake, and the lake water is discharged from the Haikou River, the only water outlet located in the southwest.

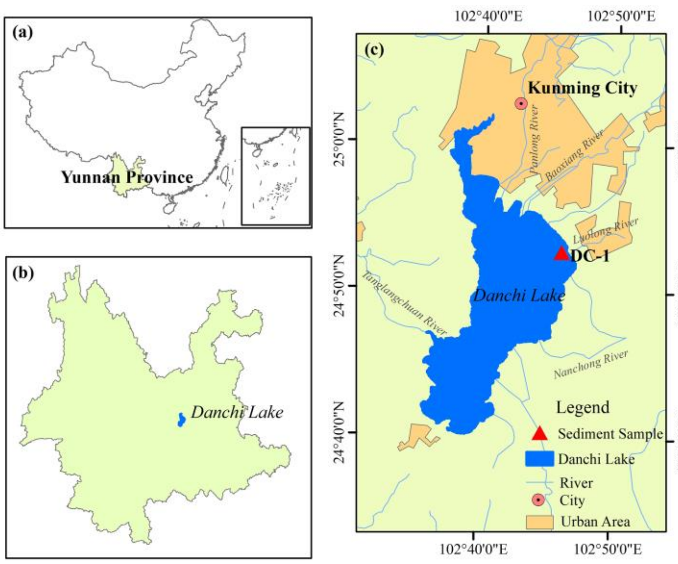

Fig. 1 Study area and sampling sites at Dianchi Lake

In this study, Dianchi sedimentary column samples were collected on July 14, 2014, using a gravity sampler combined with a GPS positioning system. The sediment core obtained at point DC1 (24° 53′ N, 102° 46′ E) (Fig. 1(c)) located near the Luolong River Estuary, had a diameter of 8.3 cm and a depth of 50 cm. The samples were cut at 1 cm intervals, sealed with plastic ziplock bags, and immediately transported back to the laboratory. Here, the samples were packed in plastic ziplock bags, freeze-dried, ground, sieved (100 mesh), and then stored at -50 °C, for further analysis in the laboratory.

2.2 Determination of PAH concentration

Detailed descriptions of the methods used for extracting and cleaning of samples, and determining PAH concentrations are provided in our previous studies (Ma et al 2018, Ma et al 2020, Zhang et al 2017).

Briefly, we first used a mixed solvent of n-hexane/acetone (1:1, v/v) for microwave extraction. Secondly, each extract was centrifuged with a mixed solvent of n-hexane/acetone (1:1, v/v), and the supernatant was collected. Thirdly, a rotary evaporator was used to concentrate each supernatant to 1 mL. Fourthly, a mixed solvent of n-hexane/dichloromethane was used for elution and purification. Finally, the collected eluate was concentrated to 1 mL and placed into amber glassware (Ma et al 2020).

Polycyclic aromatic hydrocarbons were analyzed using a Shimadzu GC-MS QP2010 plus gas chromatography-mass spectrometer. 1 μl of extraction solution was added each time, using a splitless injection mode. The chromatographic conditions were as follows: injector temperature 250 °C; detector temperature 280 °C; oven temperature 90°C at the beginning and maintained for 1 min, then increased to 160 °C at 20 °C /min, followed by an increase to 200 °C at 6 °C /min and maintained for 1 min, then increased to 230 °C at 2 °C /min and maintained for 2 min, and finally increased to 280 °C at 20 °C /min and maintained for 2 min. The mass spectrum was scanned in the m/z range 45–600 in electron ionization (EI) mode (70 eV) in the selected ion monitoring mode. PAHs were classified based on retention time and m/z value and were quantified by internal standard peak area calibration. The GC-MS was auto-tuned by infusing perfluorotributylamine. The following 16 compounds were detected in this study: NAP, ACY, ACE, FLO, PHE, ANT, FLU, PYR, BaA, CHR, BbF, BKF, BaP, DahA, InP, and BghiP. PAHs were represented by the sum of the concentrations of these 16 compounds.

2.3 Data analysis

The geographic information system Arc GIS 10.0 was used to draw the sampling map of Dianchi Lake. Pearson's correlation coefficient (R) was used for correlation analysis, to determine the correlation and significance between variables and PAH concentrations. Linear regression analysis was performed on the variables of total population, GDP, and the proportions of coal and petroleum. Results of the analysis were used as the deterministic input variables of the neural network, to establish the model.

Back Propagation of a network usually consists of three parts: an input layer, a hidden layer, and an output layer (Mackay 1992, Livingstone 2008, Foresee and Hagan 1997). The neurons in each layer are connected to each other, but the neurons in the same layer are independent of each other. The simulation termination condition is used to reach the predicted training error and training times. If the training error is too small or the number of trainings is too large, overfitting may occur, which affects the accuracy of the simulation results. In order to avoid this situation, this study uses the Bayesian regularization method as the training method of the neural network model. By reasonably adjusting the number of deliveries and repeated training, the root mean square error of the model is minimized. In this way, the most ideal neural network model was successfully established (Fig. 2), and the prediction data corresponding to the training data were derived.

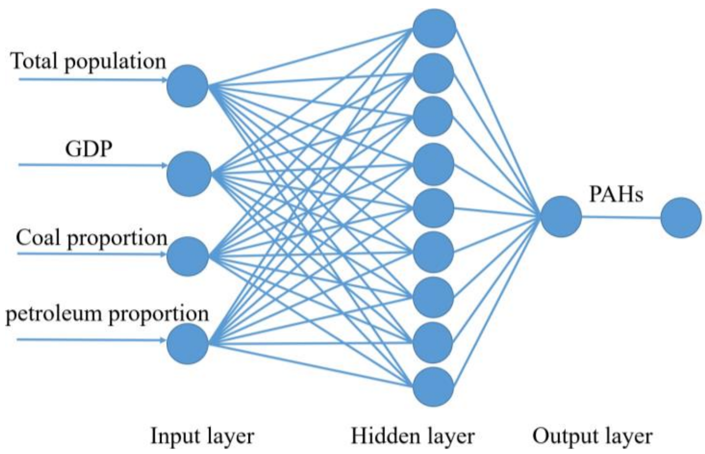

Fig. 2 The structure and composition of the neural network model

The study constructed a three-layer neural network model, as shown in Fig. 2. Standardized total population, standardized GDP, standardized coal proportion, and standardized petroleum proportion were used as the input layer variables. A hidden layer was used, and the PAH content was the output variable. All computations were performed using MATLAB. The specific calculation method was as follows:

To improve the generalization ability of the artificial neural network model, a regularization method was used, with the objective function F as follows (Bui, 2012; Li et al., 2018):

$$F=\alpha E_W+\beta E_D \qquad (1)$$

where E_W is the weighted sum of squares in the network, E_D is the sum of squares of the difference between the response value and the objective value, and α and β are objective function parameters or hyperparameters.

The weight of the network model in the Bayesian framework is generally a random variable. Bayes' rule is used to update the posterior distribution of the weights, as follows:

$$P\left(w\middle|D,\alpha,\beta,G\right)=\frac{P\left(D\middle|w,\beta,G\right)\quad P\left(w\middle|\alpha,G\right)}{P\left(D\middle|\alpha,\beta,G\right)} \qquad (2)$$

where G represents the neural network model, w represents the vector of network weights, $P(w \mid \alpha,G)$ represents the prior density, $P(D \mid \alpha,\beta,G)$ is the normalization factor, and $P(D \mid w,\beta,G)$ represents the likelihood function (Mackay et al 1992).

When the weights and data probabilities are Gaussian distributions, the formulae for normalization factors are (3) and (5), and those for prior probability, posterior probability, and likelihood function are (4), (6), and (7), respectively, as follows:

$$Z_D(\beta)=(\pi/\beta)^{n/2} \qquad (3)$$

$$P(w|\alpha,G)=\frac{\exp(-\alpha Ew)}{Zw(\alpha)} \qquad (4)$$

$$Z_D(\alpha)=(\pi/\alpha)^{k/2} \qquad (5)$$

$$P(w|D,\alpha,\beta,G)=\frac{\exp(-F(w))}{Z_F(\alpha,\beta)} \qquad (6)$$

$$P(D|w,\beta,G)=\frac{\exp(-\beta E_D)}{Z_W(\beta)} \qquad (7)$$

After optimizing the parameters of the objective function using Bayes’ rule, the formula is as follows:

$$P(\alpha,\beta|D,G)=\frac{P(D|\alpha,\beta,G)\quad P(\alpha,\beta|G)}{P(D|G)} \tag{8}$$

where $P(\alpha,\beta|G)$ is the prior probability of the parameters α and β, and $P(D|\alpha,\beta,G)$ is the likelihood function. The following iterative procedure was applied, in accordance with Foresee and Hagan (1997):

Initialize α, β, and weight values.

(1) Use the Levenberg-Marquardt algorithm to minimize the objective function F(w).

(2) Use the Levenberg-Marquardt training algorithm to approximate γ.

(3) Calculate new values of the objective function parameters α and β.

(4) Iterate step (2) until convergence.

3. Results and Discussion

3.1 Analysis of Variables and Contaminants

Many studies have shown that with the development of the economy and the surge in population, increased household and industrial fossil fuel burning (Blumer and Youngblood 1975, Khalili et al 1995), and household heating and cooking (Mastral and Callén 2000), are generally the main sources of PAHs (Burns and Jones 2016, Ko et al 2014, Li et al 2019, Ranjbar Jafarabadi et al 2017). To determine the relationship between economic variables (Yunnan Statistical Yearbook, 2015) and

PAH concentration, the correlation and significance of PAH concentration with total population, GDP, coal consumption, and oil consumption were analyzed.

The correlation coefficients, as shown in Fig. 3, for PAH concentrations and the total population (for the period 1960–2012) was 0.81; for PAHs and GDP (for the period 1972–2012) it was 0.88, for PAH and coal consumption (for the period 1977-2014) it was 0.85, and for PAHs and petroleum consumption (for the period 1977-2014) it was 0.80; furthermore, all correlations showed a strong significance ($p < 0.01$).

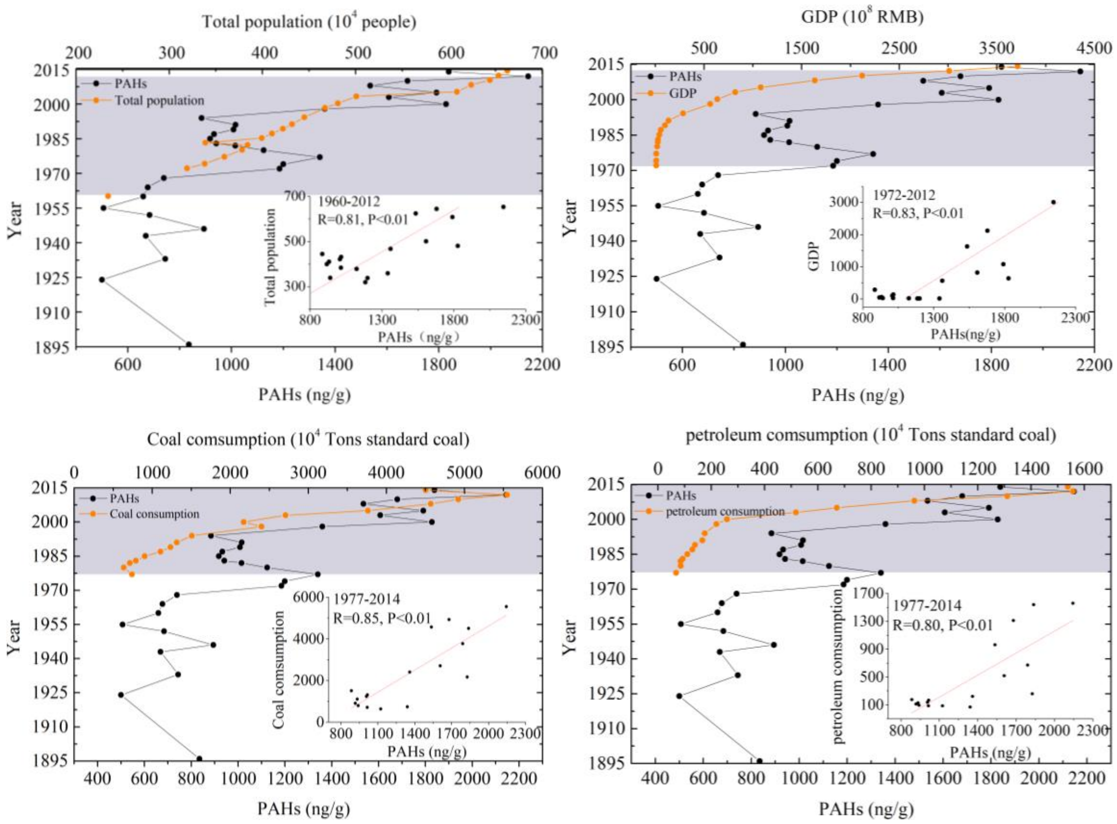

Fig. 3 The correlation and significance between PAH concentrations and total population, GDP, coal consumption and petroleum consumption

This shows that the total population, GDP, coal consumption, and oil consumption are important factors affecting the PAH content. Along with the growth of the population and economy, as well as the rapid increase in coal and petroleum consumption brought about by industrialization and urbanization, these factors have been accompanied by a sharp increase in the residual amount of PAHs in the sediment of Dianchi Lake. This is consistent with previous research results (Guo et al 2007, Hafner et al. 2005, Karina et al 2014, Liu et al 2012a, Zhang et al 2009).

3.2 Artificial Neural Network Model Training and Test

The correlation between PAHs and total population, GDP, coal and petroleum as determined above shows that PAHs have a significant correlation with these economic indicators. Considering the correlation of energy structure adjustment on the concentration of PAH pollutants (Ma et al 2017, Ma et al 2020), the coal and petroleum energy proportions as well as the total population and GDP values from 1980 to 2014 were selected as simulation parameters to make the simulation results more accurate and reliable.

When building a neural network model, the units and magnitudes of the selected parameters are often far different from each other and cannot be compared. In practical applications, the input parameters are generally required to range from -1 to 1 to improve the sensitivity and accuracy of the model. Values of the total population, GDP, coal proportion, and petroleum proportion from 1980 to 2014 were therefore normalized to generate uniform parameters. The calculation method is as follows (Lagosavid et al 2017):

$$xf=\frac{\left(xf_{max}-xf_{min}\right)}{\left(x_{max}-x_{min}\right)}\left(x-x_{min}\right)+xf_{min} \quad (1)$$

where xf is the normalized input variable, x is the input variable, xf_{max} is the maximum value for xf equal to -1 and 1, xf_{min} is the minimum value for xf equal to -1 and 1, and x_{max} is the maximum value for the input variable, and x_{min} is the minimum value for the input variable.

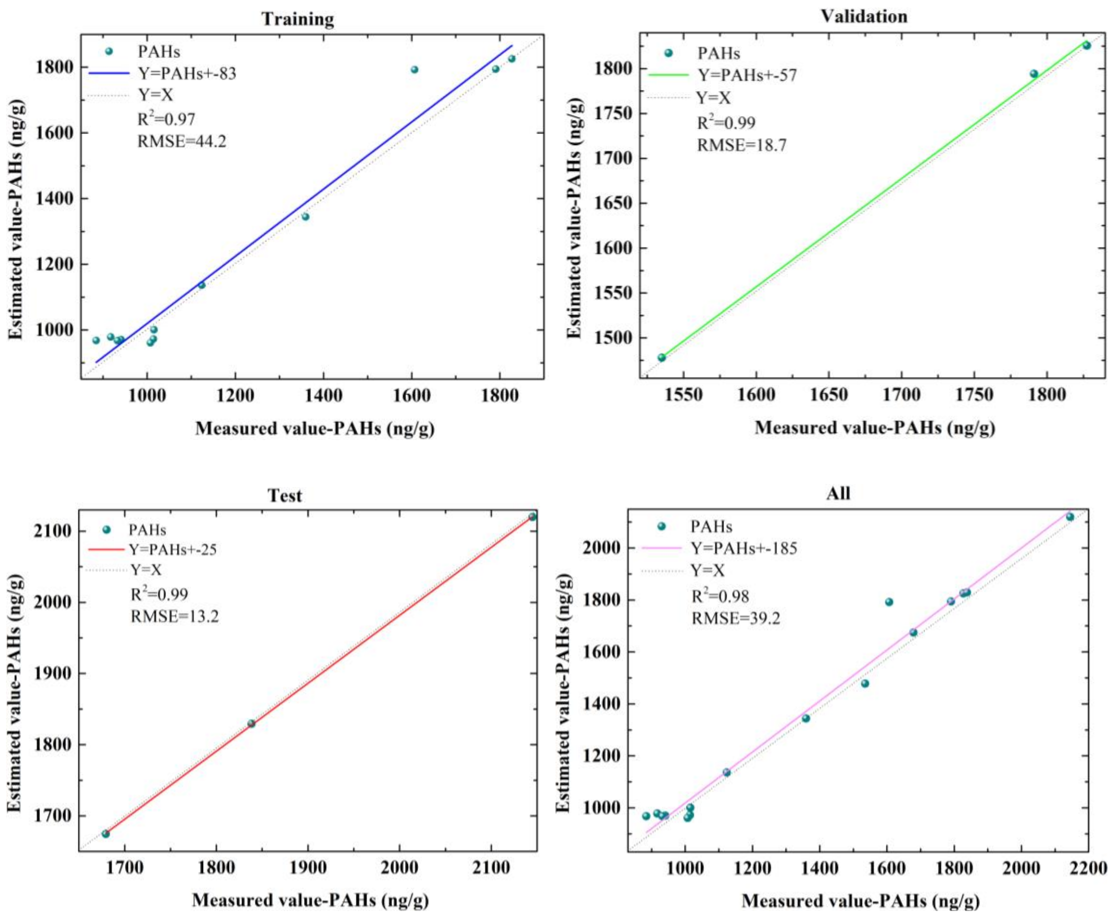

Fig. 4 Comparison of measured and predicted PAH concentrations in Dianchi Lake.

The artificial neural network model estimated the concentration of PAHs for the period 1980–2014, with the following R^2 and RMSE when compared with the measured PAHs (Fig. 4): For the training samples, the values were 0.97 and 44.2, respectively; for the validation samples they were 0.99 and 18.7, respectively; for the testing samples they were 0.99 and 13.2, respectively; and for all samples they were 0.98 and 39.2, respectively.

The closer the R^2 value is to 1, the smaller the difference between the measured value and the predicted value simulated by the model, and the closer the R^2 value is to 0, the greater the difference between the measured and predicted values. A larger RMSE value denotes a larger difference between the measured value and the value simulated by the model; and a smaller RMSE value denotes a smaller difference between the measured and predicted values. As shown in Fig. 4, the difference between the measured and predicted values simulated by the model is small, indicating that the neural network model established in this study has high accuracy and reliable results.

3.3 Prediction of PAHs

In order to predict the concentration of PAHs in the sediments of Dianchi Lake in the next ten years, regression analysis was performed on the total population, GDP, coal proportion, and petroleum proportion. The total population data from 1949 to 2018 (for 55 years, except for missing years) and the GDP data from 1970 to 2018 (for 48 years) were collected. The "Notice of the Implementation Outline for the Construction of a Regional International Central City in Kunming (2017-2030)" issued by the Kunming Municipal Committee of the Communist Party of China puts forward the development goal of Kunming. The city aspires towards attaining a GDP of 1.8 trillion yuan by 2030, and it expects that the population will reach approximately 9 million people. A regression analysis was carried out using the expected values for Kunming's population and GDP in 2030. We collected data on the proportion of coal and petroleum for 43 years, from 1975 to 2017, and performed regression analysis on the proportion of coal and petroleum. According to the derived regression equation, estimate the proportion of coal and petroleum in 2030 (Tab. 1).

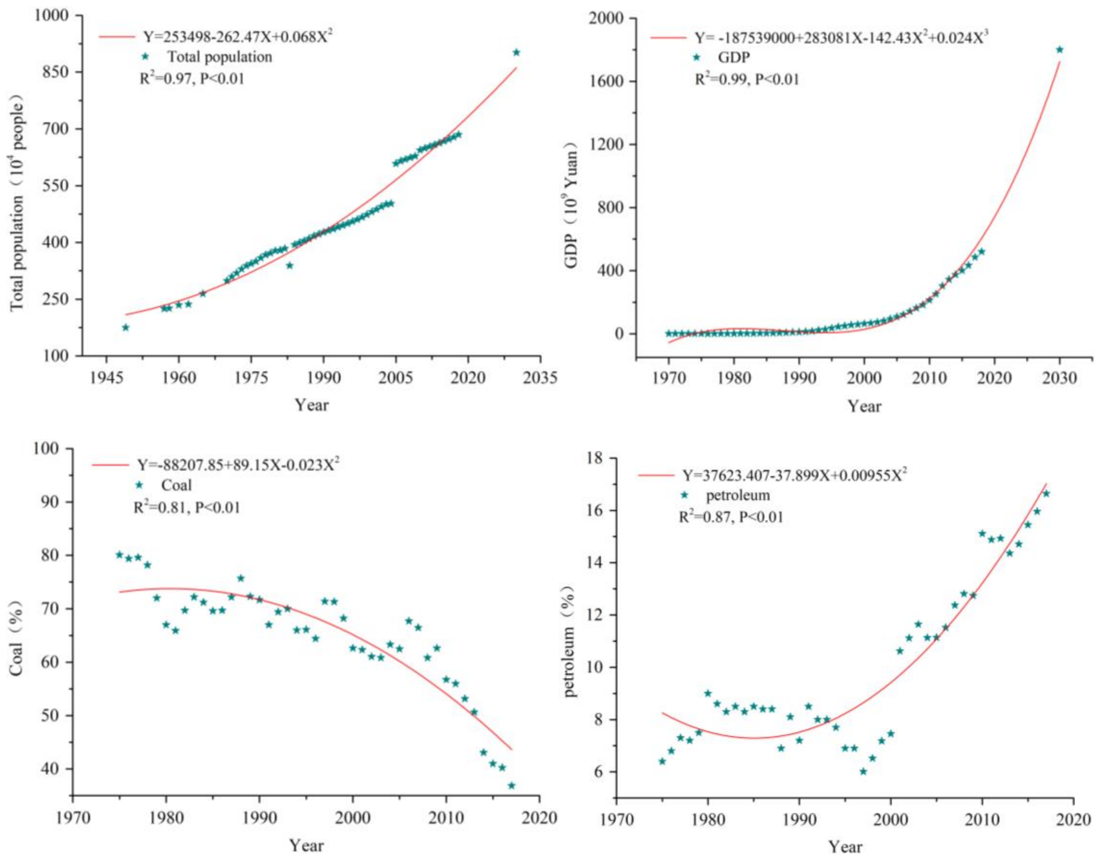

Fig. 5 Regression analysis of total population, GDP and proportion of coal and petroleum

Tab. 1 The estimated value of total population (10^4 people), GDP (10^{12} Yuan) and proportion of coal and petroleum (%) in 2030 based on regression equation

Economic parameters	Regression equation	X=2030
Total population	$Y=253498-262.47X+0.068X^2$	867.931
GDP	$Y=-187539000+283081X-142.43X^2+0.024X^3$	2.2
Coal proportion	$Y=-88207.85+89.15X-0.023X^2$	10.4
Petroleum proportion	$Y=37623.407-37.899X+0.00955X^2$	43.03

The regression analysis of the total population, GDP, coal proportion, and petroleum proportion is shown in Fig. 5, with the correlation coefficient R^2 0.97, 0.99, 0.81, and 0.87, respectively. Statistical significance was set at $p < 0.001$. Thus, we can obtain projections of the total population, GDP, coal proportion, and petroleum proportion data for 2021-2030.

Tab. 2 Prediction of polycyclic aromatic hydrocarbon concentration in Dianchi Lake Sediments

	年份									
	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
PAHs	1018.8	1016.8	1171.1	1854.1	2128.9	2108.9	1960.3	2085.2	2069.4	1839.3

Note：PAHs represent the concentration of PAHs (ng/g).

The PAH concentration predicted by the artificial neural network model shows that although the concentration of PAHs in the sediment of Dianchi Lake will not increase significantly from 2021 to 2030, the expected concentration is still not low, and it can reach is 2128.9 ng/g (Tab. 2). Dianchi Lake is the largest freshwater lake in Yunnan Province and is an important source of water to support human life including economic activities in Kunming. The ecological environment of Dianchi Lake has a greater impact on the production and life of local residents. Therefore, it is necessary for the government to take more effective measures and increase investment to control the environmental problems in Dianchi. For example, in cities, the emission of automobile exhaust strictly controlled, and devices installed to treat automobile exhaust gases. Central heating in residential areas can replaces small coal stove heating. Briquettes should be selected for industrial use, so that the coal can be fully burned. Clean energy should be developed natural gas should replace the use of coal and oil. Implementing these measures for pollution control may reduce the source input of PAHs to some degree.

4. Conclusions

The correlation and significance results between PAHs and economic parameters show that the total population, GDP, coal consumption, and oil consumption are important factors affecting the content of PAHs. From 1980 to 2014, the artificial neural network model estimated the concentration of PAHs in the database. The R^2 and RMSE of the training samples were 0.97 and 44.2, respectively, the R^2 and RMSE of the validation samples were 0.99 and 18.7, respectively, the R^2 and RMSE of the testing samples are 0.99 and 13.2, respectively, and the R^2 and RMSE of the all samples are 0.98 and 39.2, respectively. The difference between the measured and predicted values simulated by the model is small, and the neural network model established in this study has high accuracy and reliable results. The concentration of PAHs in the sediments of Dianchi Lake predicted by the neural network model shows that the highest concentration of PAHs from 2021 to 2030 can reach 2128.9 ng/g. The analysis results show that the concentration of PAHs will not increase significantly in the next ten years, but it is still not low. The ecological environment of Dianchi Lake has a greater impact on the production and life of local residents. Therefore, it is necessary for the government to take more effective measures and invest more costs to control the environmental problems in Dianchi.

Data availability

The research data are not publicly available but are available from the corresponding author on reasonable request.

References

Anagu, I.; Ingwersen, J.; Utermann, J.; Streck, T. Estimation of heavy metal sorption in German soils using artificial neural networks. Geoderma 2009, 152, 104-112.

Baumard P., Budzinski H., Michon Q., Garrigues T., Burgeot J. Origin and bioavailability of PAHs in the Mediterranean Sea from mussel and sediment records. Estuarine, Coastal Shelf Sci.1998, 47, 77-90.

Blumer M., Youngblood W.W. Polycyclic aromatic hydrocarbons in soils and recent sediments. Science. 1975, 188, 53-55.

Bui D.T., Pradhan B., Lofman O., et al. Landslide susceptibility assessment in the Hoa Binh province of Vietnam: A comparison of the Levenberg-Marquardt and Bayesian regularized neural networks. Geomorphology. 2012, 171, 12-29.

Burns K.A., Jones R. Assessment of sediment hydrocarbon contamination from the 2009 Montara oil blow out in the Timor Sea. Environ. Pollut. 2016, 211, 214-225.

Buszewski, B.; Kowalkowski, T. A New Model of Heavy Metal Transport in the Soil Using Nonlinear Artificial Neural Networks. Environ. Eng. Sci. 2006, 23, 589-595.Donahue W.F., Allen E.W., Schindler D.W. Impacts of coal-fired power plants on trace metals and polycyclic aromatic hydrocarbons (PAHs) in lake sediments in central Alberta, Canada, J Paleolimnol.

2006, 35, 111-128.

Du L.N., Li Y., Chen X.Y. et al. Effect of eutrophication on molluscan community composition in the Lake Dianchi (China, Yunnan). Limnologica. 2011, 41, 213-219.

Dushyant, R.D., Rahul, K.R., Jwalant, K.B., Haren, B.G., Bhumi, K.S., Bharti, P.D., 2016. Distribution, sources and ecological risk assessment of PAHs in historically contaminated surface sediments at Bhavnagar coast, Gujarat, India. Environ. Pollut. 213, 338-346.

Freeman D.J., Cattell F.C. Woodburning as a source of atmospheric polycyclic aromatic hydrocarbons. Environmental Science & Technology. 1990, 24, 1581-1585.

Foresee, F.D., Hagan, M.T. Gauss-Newton Approximation to Bayesian Learning. In Proceedings of the International Conference on Neural Networks; IEEE: Piscataway, NJ, USA, 1997; pp. 1930-1935.

Guo J.Y., Wu F.C., Zhang L., et al. Screening level of PAHs in sediment core from Lake Hongfeng, Southwest China. Arch Environ Contam Toxicol. 2011, 60:590-596.

Gu Y.G., Li H.B., Lu H.B. Polycyclic aromatic hydrocarbons (PAHs) in surface sediments from the largest deep plateau lake in China: Occurrence, sources and biological risk. Ecological Engineering. 2017, 101, 179-184.

Guo Z., Lin T., Zhang G., Zheng M., Zhang Z., Hao Y., Fang M. The sedimentary fluxes of polycyclic aromatic hydrocarbons in the Yangtze River Estuary coastal sea for the past century. Sci Total Environ. 2007, 386, 33-41.

Hafner W.D., Carlson D.L., Hites R.A. Influence of local human population on atmospheric polycyclic aromatic hydrocarbon concentrations. Environ. Sci. Technol. 2005, 39(19), 7374-7379.

Jia Z., Zhou S., Su Q., et al. Comparison study on the estimation of the spatial distribution of regional soil metal(loid)s pollution based on kriging interpolation and BP neural network. Int. J. Environ. Res. Public Health 2018, 15, 34.

Karina S.M., Rubens C.L.F., Lilian C.C., et al. Sedimentary record of PAHs in the Bangui River and its relation to the socioeconomic development of Curitiba, Brazil[J]. Sci. Total. Environ. 2014, 482-483, 42-52.

Kannan K., Johnson-Restrepo B., Yohn S.S., Giesy J.P., Long D.T. Spatial and temporal distribution of polycyclic aromatic hydrocarbons in sediments from Michigan inland lakes. Environ Sci Technol. 2005, 39(13), 4700-4706.

Khalili N.R., Scheff P.A., Holsen T.M. PAH source fingerprints for coke ovens, diesel and, gasoline engines, highway tunnels, and wood combustion emissions. Atmospheric environment. 1995, 29, 533-542.

Kim E.J., Oh J.E., Chang Y.S. Effects of forest fire on the level and distribution of PCDD/Fs and PAHs in soil. Sci Total Environ. 2003, 311, 177-189.

Ko F.C., Chang C.W., Cheng J.O. Comparative study of polycyclic aromatic hydrocarbons in coral tissues and the ambient sediments from Kenting National Park, Taiwan. Environ. Pollut. 2014, 185, 35-43.

Latimer J.S., Zheng J. The sources, transport, and fate of PAHs in the marine environment. John Wiley & Sons, Chichester. 2003, 9-33.

Li Y, Zhou S, Jia Z, et al. Influence of Industrialization and Environmental Protection on Environmental Pollution: A Case Study of Taihu Lake, China. Int. J. Environ. Res. Public Health, 2018, 15(12).

Li Y.L., Wang C.L., Zou X.Q., et al. Occurrence of polycyclic aromatic hydrocarbons (PAHs) in coral reef fish from the South China Sea. Marine Pollution Bulletin. 2019, 139, 339-345.

LI Y.M., Peng Y.A., Wang Y.C., et al. The pollution feature of Dianchi Lake and its control countermeasure. Yunnan Geographic Environment Research. 2003, 15(4), 32-38.

Liu G., Liu Z., Li Y., et al. Effects of fish introduction and etrophication on the cladceran community in Lake Fuxian, a deep oligotrophic lake in southwest china. Joural of Palelimnology. 2008, 42(3), 427-435.

Liu L.Y., Wang J.Z., Wei GL., et al. polycyclic aromatic hydrocarbons (PAHs) in continental shelf sediment of China: implications for anthropogenic influence on coastal marine environment. Environ. Pollut. 2012, 167, 155-162.

Liu, G.; Zheng, M., Lv, P.; Liu, W., Wang, C., Zhang, B., Xiao, K. Estimation and characterization of polychlorinated naphthalene emission from coking industries. Environ. Sci. Technol. 2010, 44, 8156-8161.

Livingstone, D.J. Artificial Neural Networks: Methods and Applications (Methods in Molecular Biology); Humana Press: New York, NY, USA, 2008; Volume 458.

Lagosavid M.P., Bonilla C.A. Predicting the particle size distribution of eroded sediment usingartificial neural networks. Sci. Total Environ. 2017, 581, 833-839.

Ma, C.L., Lin, T., Ye, S.Y., et al. Sediment record of polycyclic aromatic hydrocarbons in the Liaohe River Delta wetland, Northeast China: Implications for regional population migration and economic development. Environ. Pollut. 2017, 222, 146-152.

Ma X.H., Han X.X., Jiang Q.L., et al. Historical records and source apportionment of polycyclic aromatic hydrocarbons over the past hundred years in Dianchi Lake, a plateau lake in

Southwest China. Arch. Environ. Contam. Toxicol. 2018, 75, 187-198.

Ma X.H., Wan H.B. Zhou J., et al. Sediment record of polycyclic aromatic hydrocarbons in Dianchi Lake, southwest China: Influence of energy structure changes and economic development. chem. DOI:10.1016/j.chemosphere.2020.126015.

Mackay, D.J.C. Bayesian interpolation. Neural Comput. 1992, 4, 415-447.

Mackay D., Shiu W.Y., Ma K.C. Illustrated handbook of physical-chemical properties and environmentalfate for organic chemicals. Volume II: polynuclear aromatic hydrocarbons, polychlorinated dioxins, and dibenzofurans. Lewis Publishers, Boca Raton, FL. 1992. 597.

Maren A.J., Harston C.T., Pap R.M. Handbook of Neural Computing Applications; Academic Press: Cambridge, MA, USA, 1990; pp. 439-448.

Mastral A.M., Callen M.S. A review on polycyclic aromatic hydrocarbon (PAH) emissions from energy generation. Environmental Science & Technology. 2000, 34, 3051-3057.

Morillo E., Romero A.S., Maqueda C., Madrid L., Ajmone-Marsan F., Grcman H., Davidson C.M., Hursthouse A.S., Villaverde J. Soil pollution by PAHs in urban soils: a comparison of three European cities. J Environ Monit. 2007, 9, 1001-1008.

Nemr A.E.L, Said T.O, Khaled A, Sikaily A.E.L., Allah A.M.A. The distribution and sources of polycyclic aromatic hydrocarbons in surface sediments along the Egyptian Mediterranean coast. Environ Monit Assess. 2007, 124,343-359

Pietzsch R., Patchineelam S.R., Torres J.P.M. Polycyclic aromatic hydrocarbons in recent sediments from a subtropical estuary in Brazil. Mar Chem. 2010, 118, 56-66.

Qiu X., Zhu T., Yao B., Hu J., Hu S. Contribution of dicofol to the current DDT pollution in China. Environ. Sci. Technol. 2005, 39, 4385-4390.

Ranjbar Jafarabadi A., Riyahi Bakhtiari A., Aliabadian M., et al. Spatial distribution and composition of aliphatic hydrocarbons, polycyclic aromatic hydrocarbons and hopanes in superficial sediments of the coral reefs of the Persian Gulf, Iran. Environ. Pollut. 2017, 224, 195-223. https://doi.org/10.1016/j.envpol.2017.01.080.

Sari, B. Modeling effluent heavy metal concentrations in a bioleaching process using an artificial neural network technique. Afr. J. Biotechnol. 2012, 11, 16196-16204.

Shi, X.Q.; Yang, X.B., Wang, Z.X., Qiu, L., Wu, T.C., Wang, Z.Z. A novel use of neural network model to determine the effects of multi biomarker on early health damage among Chinese steel workers. Environ. Toxicol. 2015, 26, 1-9.

Sun C.Y., Ma Q.Y., Zhang J.Q., et al. Predicting seasonal fate of phenanthrene in aquatic environment with a Markov chain. Environ. Sci. Pollut. Res. 2016, 23, 16661-16670.

Thomas M., Lazartigues A., Banas D., Brun-Bellut J., Feidt C. Organochlorine pesticides and polychlorinated biphenyls in sediments and fish from freshwater cultured fish ponds in different agricultural contexts in north-eastern France. Ecotoxicol. Environ. Saf. 2012, 77, 35-44.

USEPA. Polycyclic Organic Matter. US Environmental Protection Agency. Washington, DC. 2002. http://www.epa.gov/ttn/atw/hlthef/polycycl.html.

Van Metre P.C., Mahler B.J., Furlong E.T. Urban sprawl leaves its PAH signature. Environ. Sci. Technol. 2000, 34, 4064-4070.

Viguri J, Verde J, Irabien A. Environmental assessment of polycyclic aromatic hydrocarbons (PAHs) in surface sediments of the Santander Bay, Northern Spain. Chemosphere. 2002, 48, 157-165.

Wang, W.T., Staci L., Massey, S., Xue, M., Zhao, J.Y., Zhang N., Wang, R., Cao J., Tao, S., 2010. Concentrations, sources and spatial distribution of polycyclic aromatic hydrocarbons in soils from Beijing, Tianjin and surrounding areas, North China. Environ Pollut. 158, 1245-1251.

Wösten, J.H.M.; Pachepsky, Y.A.; Rawls, W.J. Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol. 2001, 251, 123-150.

Yim U.H., Hong S.H., Shim W.J., Oh J.R., Chang M. Spatio-temporal distribution and characteristics of PAHs in sediments from Masan Bay, Korea. Mar Pollut Bull. 2005, 50, 319-326.

Yin C.Q., Jiang X., Yang X.., Bian Y.R., Wang F. Polycyclic aromatic hydrocarbons in soils in the vicinity of Nanjing, China. Chemosphere. 2008, 73, 389-394.

Yuan H.Z., Liu E.F., Zhang E.L., Luo W.L., Chen L., Wang C., lin Q. Historical records and sources of polycyclic aromatic hydrocarbons (PAHs) and organochlorine pesticides (OCPs) in sediment from a representative plateau lake, China. Chemosphere. 2017, 173, 78-88.

Yunnan, statistical bureau. Yunnan statistical Yearbook. China Statistical Publishing House, Yunnan.2015.

Zeng H.A., Wu J.L. Sedimentary records of heavy metal pollution in Fuxian Lake, Yunnan Province, China: intensity, history, and sources. Pedosphere. 2009, 19, 562-569.

Zhang Q.Q., Xia Z.H., Wu M.M., et al. Human health risk assessment of DDTs and HCHs throμgh dietary exposure in Nanjing, China. Chem. 2017, 177, 211-216.

Zhang Y.D., Su Y.L., Liu Z.W., Chen X.C., Yu J.L., Di X.D., Jin M. Sediment lipid biomarkers record increased eutrophication in Lake Fuxian (China) during the past 150 years. J. Great

Lakes Res. 2015, 41, 30-40.
Zhang Y., Tao S. Global atmospheric emission inventory of polycyclic aromatic hydrocarbons (PAHs) for 2004. Atmos. Environ. 2009, 43(4), 812-819.
Zhang, Y.X., Tao, S., Cao, J., Coveney, R.M., 2007. Emission of polycyclic aromatic hydrocarbons in China by County. Environ. Sci. Technol. 41 (3), 683-687.
Zhi H., Zhao Z.H., Zhang L. The fate of polycyclic aromatic hydrocarbons (PAHs) and organochlorine pesticides (OCPs) in water from Poyang Lake, the largest freshwater lake in China. Chemosphere. 2015, 119, 1134-1140.
Zhou, R.C., Yang, R.Q., Jing, C.Y., 2018. Polycyclic aromatic hydrocarbons in soils and lichen from the western Tibetan Plateau: Concentration profiles, distribution and its influencing factors. Ecotox. Environ. Safe. 152, 151-158.

Author contributions

X.H.M., T.H. and C.C.H. designed the experiments. X.H.M., S.D.L., H.B.W., carried out the experiments and performed the analyses. X.H.M., H.Y., T.H., C.C.H. substantially contributed to interpreting the results and writing the paper.

Funding

This work was funded by the National Natural Science Foundation of China [Grant numbers 41773097, 41673108 and 41971286], the Youth Top Talent funded by Nanjing Normal University, a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Acknowledgements

We sincerely thank Yaoxin Liu, Congcong Zhou and Shuyan Cheng for their contributions to the experiment assistance.

Ethics declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests

Figures

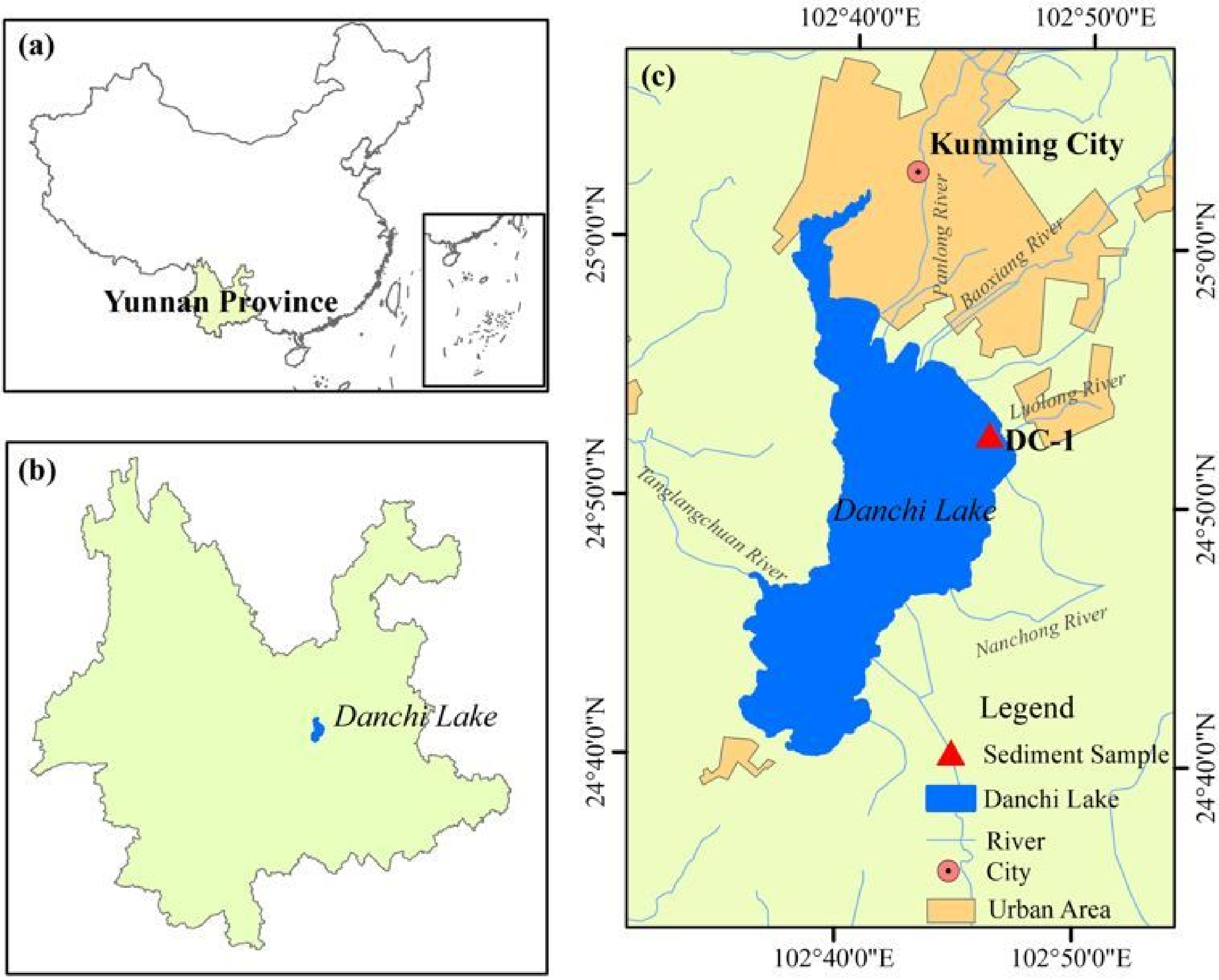

Figure 1

Study area and sampling sites at Dianchi Lake Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

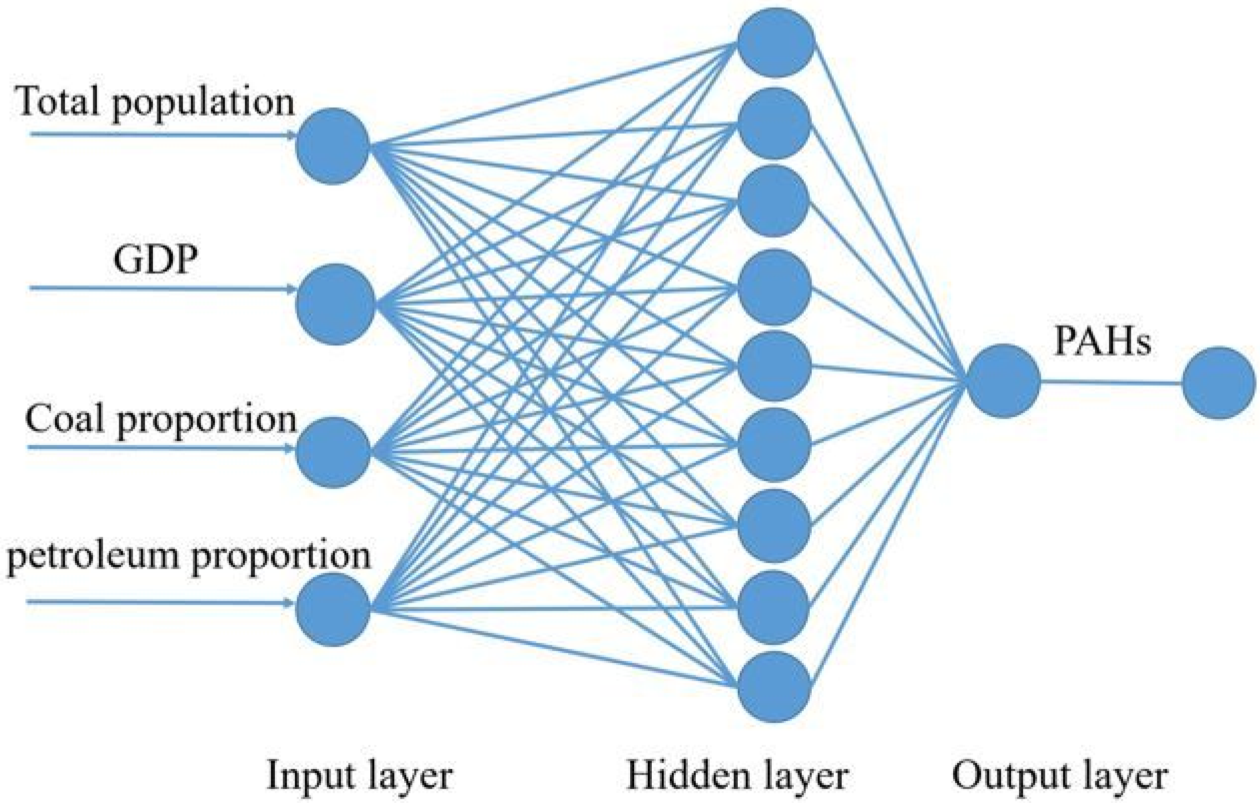

Figure 2

The structure and composition of the neural network model

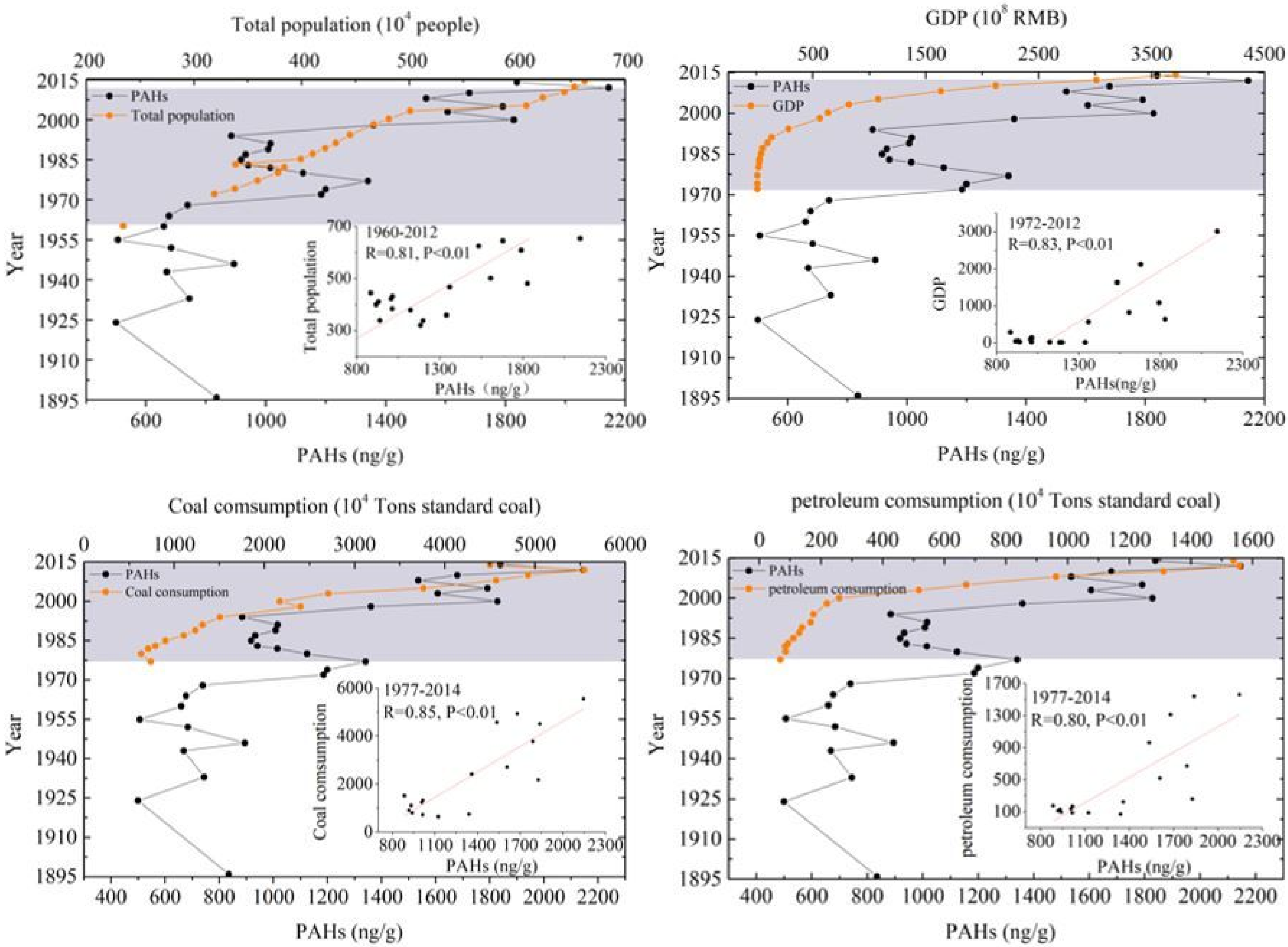

Figure 3

The correlation and significance between PAH concentrations and total population, GDP, coal consumption and petroleum consumption

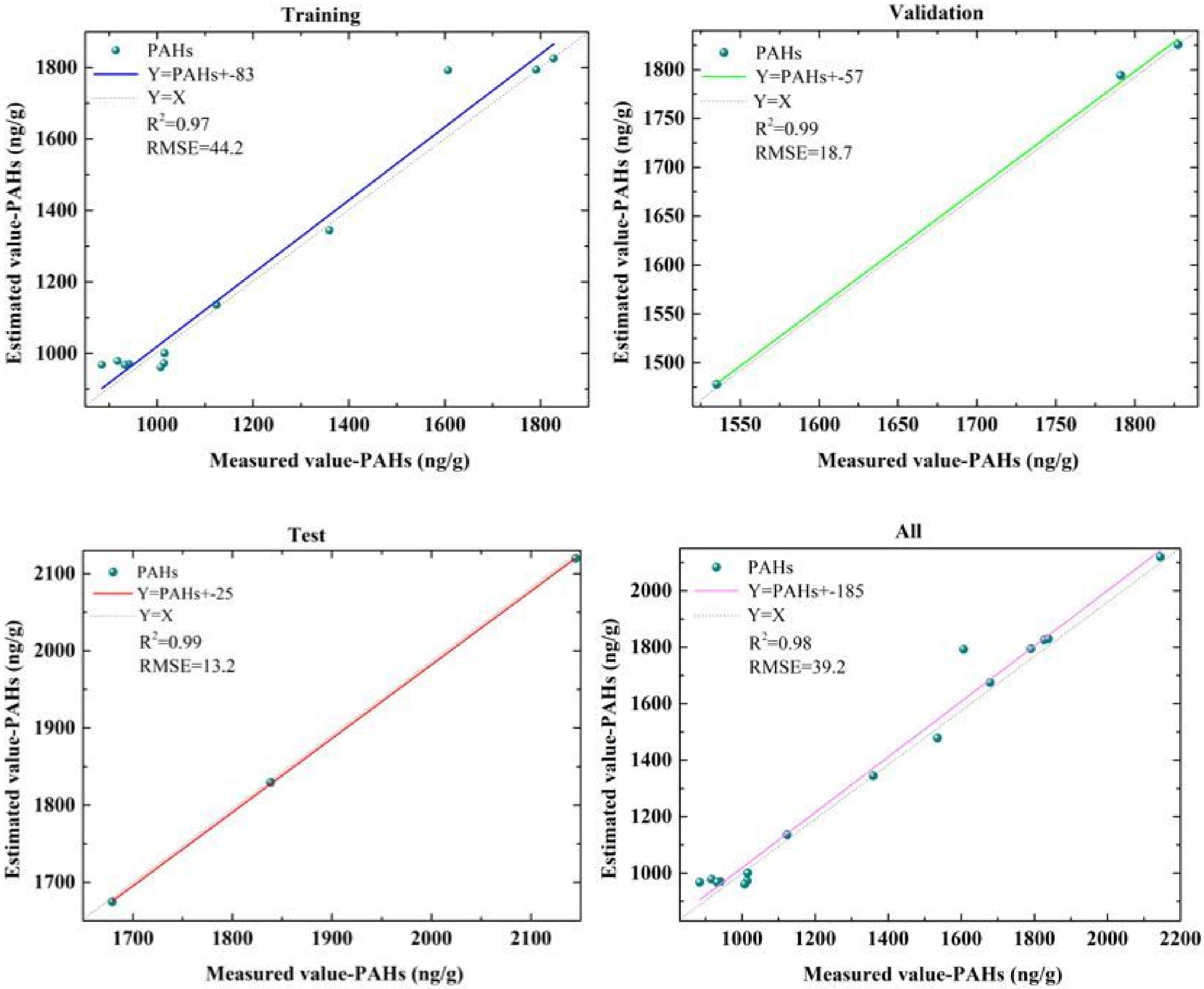

Figure 4

Comparison of measured and predicted PAH concentrations in Dianchi Lake.

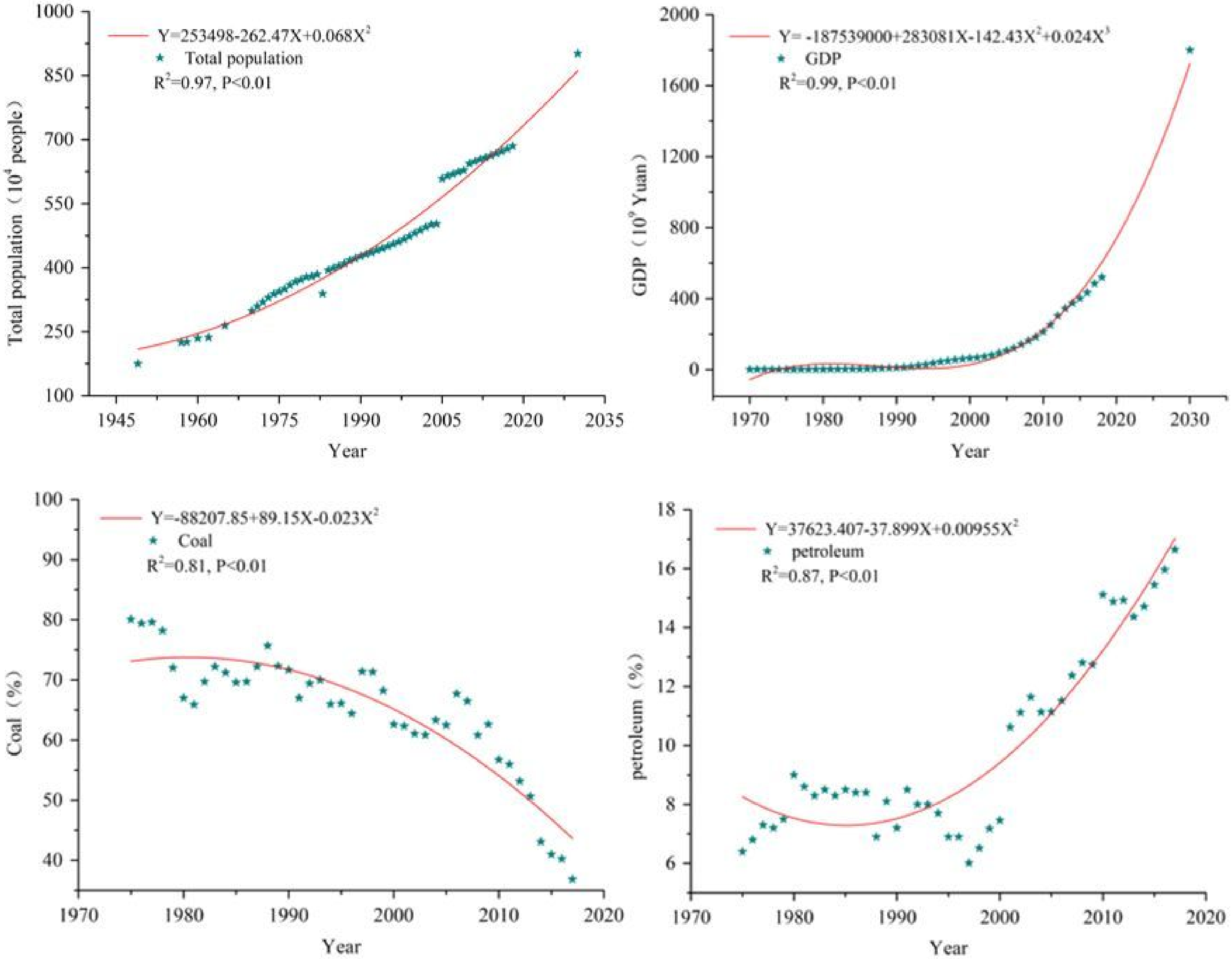

Figure 5

Regression analysis of total population, GDP and proportion of coal and petroleum