IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 21, 2020, accepted January 8, 2021, date of publication January 13, 2021, date of current version January 25, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051505

Design and Analysis of an Data-Driven Intelligent Model for Persistent Organic Pollutants in the Internet of Things Environments

CHUNXUE WU[1], CHENG WANG[1], QINGFENG FAN[2], QIONGLI WU[3], SHENG XU[4], AND NEAL N. XIONG[5]

[1]Department of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China
[2]School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China
[3]Innovation Academy for Precision Measurement Science and Technology (APM), Chinese Academy of Sciences (CAS), Wuhan 430071, China
[4]Department of Vocational Normal, Jiangxi Agricultural University, Nanchang 330045, China
[5]Department of Mathematics and Computer Science, Northeastern State University, Tahlequah, OK 74464, USA

Corresponding author: Sheng Xu (fanny_042@163.com)

ABSTRACT The targeted compounds included Polychlorinated Biphenyls (PCBs), Pesticides (PESTs), Polycyclic Aromatic Hydrocarbons (PAHs) and so on in the Great Lakes Integrated Atmospheric Deposition Network (IADN), which is a platform based on the IoT (Internet of Things) technology to collect environmental pollutants data. While previous studies usually employed traditional statistical approaches to analyze the IADN results, we performed a complete modeling workflow of the total concentrations of PCBs, PESTs, and PAHs (which is referred to as $\sum$ PCBs, $\sum$ PEST s and $\sum$ PAHs orderly) in 1990-2016 samples by using a machine learning algorithm combined with data-driven research method, which lets the model fit the data, so as to change the model to achieve the effect. The main results of this article are as follows, 1) identifying the spatial and temporal trends of POPs (Persistent Organic Pollutants) in the air of the Great Lakes; 2) An appropriate data-driven intelligent model was constructed for the data at EH (Eagle Harbor) and STP(Sturgeon Point) sampling sites, via which we estimated their $\sum$ PCBs, $\sum$ PESTs, and $\sum$ PAHs in the following 4-5 years, showing the concentrations will continue declining with slight fluctuations; 3) The important role which IoT played in smart environmental protection was pointed out.

INDEX TERMS Atmospheric environment, data-driven, great lakes, Internet of Things, intelligent model, persistent organic pollutants, machine learning.

I. INTRODUCTION

The North American Great Lakes (Lakes Superior, Michigan, Huron, Erie, and Ontario) have been subjected to elevated pollutions due to intensive human activities [1], [2]. Therefore, the United States Environmental Protection Agency (U.S. EPA) implemented the Integrated Atmospheric Deposition Network (IADN) program to monitor air pollution over the Great Lakes. They applied the Internet of Things technology to the field of environment, widely collect data, and use intelligent technologies such as data mining to screen and refine the collected data, so as to provide researchers and decision makers with safe, reliable and effective data information. The IADN team has been measuring Persistent Organic Pollutants (POPs) in the Great Lakes atmosphere and precipitation since 1990. Venier *et al.* reported that the residues of POPs in the Great Lakes air would not be removed promptly, and they mainly arose from human activities [3]. Though the concentrations of various contaminants have declined in the IADN samples since the implementation of the Stockholm convention in 2004, the decreasing trends in atmospheric levels of PCB-11 were not significant [4], [5].

Previous studies usually employed traditional statistical approaches without prediction to analyze the IADN results. A. Salamova *et al.* presented their measurements of several halogenated and non-halogenated Ops (Organophosphate esters) in particle samples collected as part of the Integrated Network(IADN), and Some statistical analysis results are given for the data [6]. R.A. Hites separated the analytical error from the sampling error for the target compounds by using surrogate (recovery) standards [7]. The precision of atmospheric concentration measurements of POPs was discussed

The associate editor coordinating the review of this manuscript and approving it for publication was Zhenyu Zhou.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

by D.C. Lehman *et al.* [8]. In the present study, we revisited these data and analyzed them in terms of machine learning combined with data-driven research methods, which lets the model fit the data, so as to change the model to achieve the effect. The input factors of the time series model, such as model fixed order and differential orders, are adjusted to make the model fit the data and get better prediction effect in this article. The temporal and spatial trends of POPs was more accurately analyzed, and some predictions were made about future concentration data.

Machine learning, a branch of artificial intelligence, has received mounting attention from numerous research fields, including environmental, epidemiological, and pharmaceutical studies, among others [9]. Its algorithms include supervised learning, unsupervised learning, and reinforcement learning. These models can efficiently visualize complex data in terms of multiple techniques (e.g., data sorting and reduction of data-dimensions), enabling researchers to extract valuable information from large datasets [10]. Additionally, machine learning facilitates the use of existing data to perform predictions scientifically, which can be of great commercial and social values [11].

Machine learning has a wide range of research. Intelligent model design of complex system becomes a key issue for organization responsiveness to uncertainties. J. Li and N. Xiong *et al.* provided a novel framework and approach to design cluster supply chain without across-chain horizontal cooperation [12]. M.M. Hassan and H. Liao *et al.* applied machine learning to the IIoT (Industrial Internet of Things) environment and achieved some good results [13], [14]. Importing heuristic algorithm, F. Long *et al.* improved the virtual topology strategy to satisfy the requirements of users [15], [16]. Y. Yang *et al.* proposed a decentralized flocking algorithm to achieve the goal of collision avoidance [17]. Z. Zhou applied machine learning algorithm to IoHT (Internet of Health Things), proposed a new scheme, and verified its effectiveness and reliability [18]. An increasing number of environmental researchers have starting incorporating machine learning to evaluate their data. For example, Knoll *et al.* and Nourani *et al.* used machine learning technique to predict the groundwater levels and the concentrations of nitrate, and compared the prediction performance among various models [19], [20]. Machine learning algorithm and model have great practicability [21], [22]. To estimate the potential threats by groundwater to public health, a previous study integrated Neural Networks (NN) and Support Vector Machines (SVM) into a Geographic Information System (GIS) to identify contaminated wells, and used logistic regression and feature selection methods to prioritize variables [23], [24]. Machine learning was also successfully adopted to predict the dissociation energy of carbon-fluorine bonds [25], as well as to predict the biological activity of per- and polyfluoroalkyl substances [26]. Specially, the application and research of time series model are very extensive. T. Shen *et al.* proposed 3D Augmented Convolutional Network (3DACN) to extract time series information and solve the serious imbalanced data problem [27]. Based on machine learning, Y. Zhang and J.C. Sun *et al.* Put forward improved time series analysis methods, which all owed the information in the time series to be extracted by analyzing the associated complex network [28]. R.J. Zhou proposed Flexible Multi-Scale Entropy (FMSE) to increases the reliability and stability of measuring time series complexity [29]. These studies showed the great advantage of employing the machine learning.

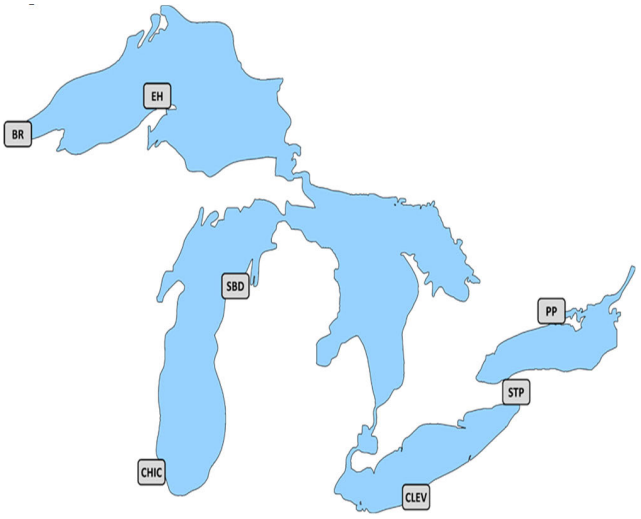


FIGURE 1. Map of the IADN sampling sites around the Great Lakes.

The rest of this article is organized as follows. Section II describes the sources of research data and the research methods. Section III presents detailed data analysis and visualization of experimental outcomes. And the models were evaluated. Section IV gives environment-related conclusions based on the results obtained in the previous two sections. Section V summarizes the findings of this study and makes the prospects for future research.

II. MATERIALS AND METHODS

A. RESEARCH AREA

The Great Lakes located in the east-central North America, between the United States and Canada, is the largest group of freshwater lakes in the world; The total area of these five lakes is 245,660 square kilometers [30]. The IADN, a long-term atmospheric monitoring program, which is run by the office of the Great Lakes' national program of the U.S. EPA, has been measuring Polychlorinated Biphenyls (PCBs), Pesticides (PESTs), Polycyclic Aromatic Hydrocarbons (PAHs), and flame retardants in the Great Lakes atmosphere and precipitation since 1990 by Indiana University. The sampling sites were Brule River (BR), Chicago (CHIC), Cleveland (CLEV), Eagle Harbor (EH), Point Petre (PP), Sleeping Bear Dunes (SBD), and Sturgeon Point (STP) (see Figure 1) [31].

Based on Internet of Things technology, a set of atmospheric sample (both vapor and particle phases) was collected every 12 days with a bulk-active air sampler for 24 hours at each site (except at PP, where the sampling frequency was every 24 or 36 days). The sampling periods were 1996-2002

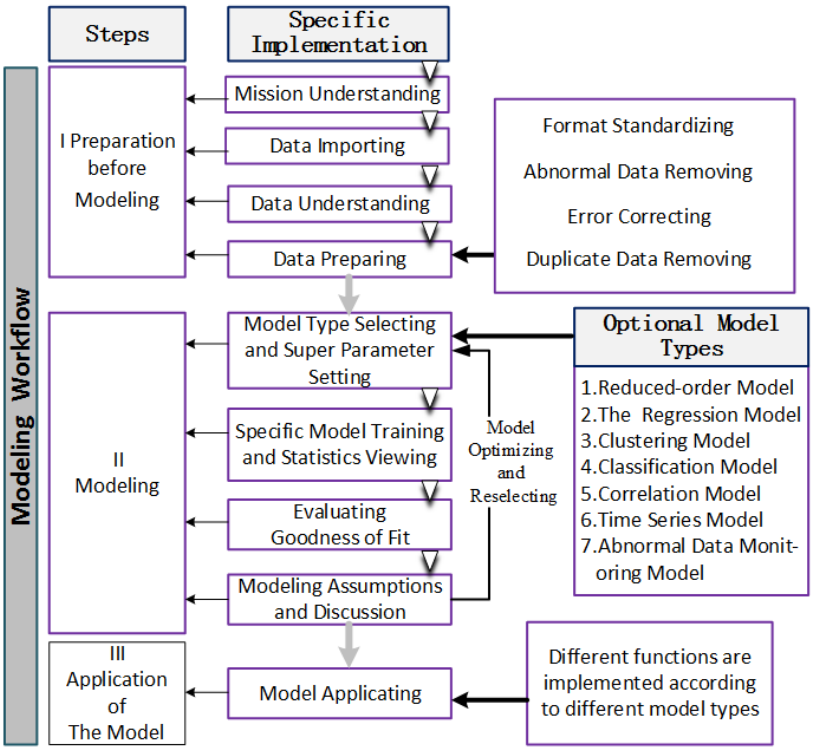


FIGURE 2. Flow chart of modeling.

at BR, 1996-2016 at CHIC, 2003-2016 at CLEV, 1990-2016 at EH, 1998-2016 at PP, 1991-2016 at SBD, and 1991-2016 at STP. A detailed list of targeted contaminants can be found at the IADN Data Viz. The data discussed in this article were also downloaded from that website [29].

B. INTRODUCTION TO MODELING WORKFLOW

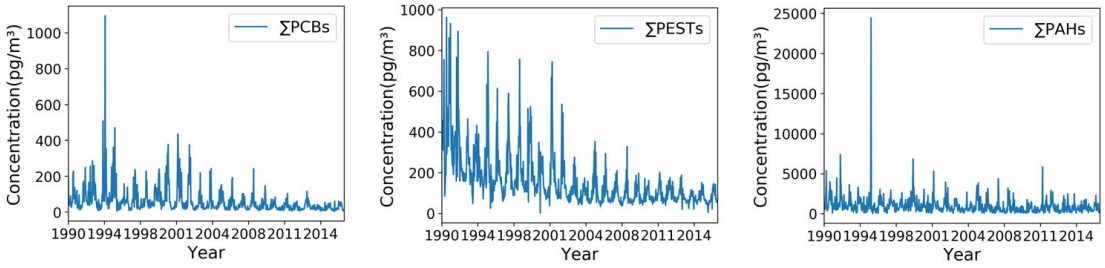
In this study, we analyzed the total concentrations of PCBs, PESTs, and PAHs ($\sum$ PCBs, $\sum$ PESTs, and $\sum$ PAHs) in the IADN vapor phase samples for their spatiotemporal trends and made some predictions by building the right model. The modeling workflow are shown in Figure 2. The entire modeling process consists of three stages. The stage I, preparation before modeling. This stage mainly includes mission understanding, data importing, data understanding, and data preparation. Data preparing refers to the process of making a series of operations on the original data into experimental data. The stage II, modeling. The main steps in the modeling stage include model type selection and super parameter setting, specific model training and statistics viewing, evaluating goodness of fit, and modeling assumptions and discussion. In addition, we need do model optimizing and reselection if the obtained results do not meet the requirements. The stage III, application of the model. The corresponding function of the model is implemented according to the model type in this stage. The prediction function was realized by the model constructed in this article. The details can be obtained from the figure below.

C. DATA PRE-PROCESSING

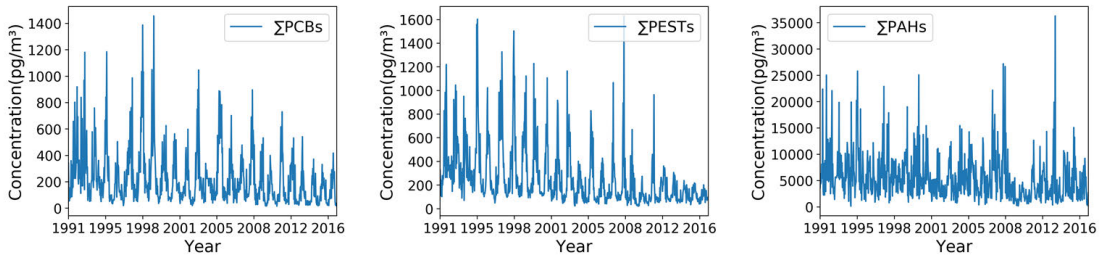
There would be missing values and outliers in the original dataset for various reasons, so data preprocessing is required prior to Modeling. After investigating the original dataset, we detected a number of missing values and outliers. Generally speaking, there are three main processing methods for processing data missing: 1) using the attribute which contains the missing value directly without any processing; 2) deleting the attribute which contains a large number of missing values; and 3) missing value replacement. In this article, we adopted the second and third approaches. As the majority of $\sum$ PESTs data were missing for the BR sampling site, we deleted the corresponding attribute directly, while for the rest attributes, their missing values were substituted with the corresponding group mean. POPs at EH and STP sampling sites were taken as examples, the data distributions after the missing value processing were shown in Figure 3.

As can be seen in Figure 3 (a), the trend graphs of $\sum$ PCBs, $\sum$ PESTs, and $\sum$ PAHs at EH sampling points are continuous, indicating that the missing values have been completed. In addition, we found that there were some anomalous data points in the trend charts. The values of these data points, which needed to be removed in the next step, were quite different from those of most data points. the similar situation can be seen in Figure 3 (b).

Then, the Python box graph was used to identify outliers. The structure of box figure is shown in Figure 4. After calculating the first quartile (*Q1,* at *25%* position), the median and



(a) Trends of POPs concentration in Eagle Harbor（1990-2016）



(b) Trends of POPs concentration in Sturgeon Point（1991-2016）

FIGURE 3. Trend charts of POPs concentrations in the Great Lakes atmosphere after the missing data processing.

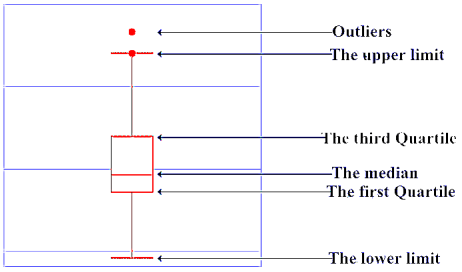


FIGURE 4. Structure of box plot.

the third quartile ($Q3$, at 75% position), we defined the inter-quartile range (IQR) as ($Q3$-$Q1$), and considered the values outside the range between ($Q1 - 1.5 \times IQR$; the lower limit) and ($Q3 + 1.5 \times IQR$; the upper limit) as outliers. Similarly, POPs at EH and STP sampling sites were taken as examples, the box plots for POPs levels in the Great Lakes atmospheric samples are shown in Figure 5, where the outliers clearly stand out and are excluded from the subsequent elucidation of the POPs' environmental behaviors.

In Figure 5 (a), the sampling time of EH is from 1990 to 2016, which is the longest sampling period among the seven sampling points. At EH sampling site, the number of abnormal data points per year is within the acceptable range, which can be counted on one's fingers. In Figure 5 (b), the sampling period span of STP is from 1991 to 2016. Compared to other sampling sites, the data volume collected in STP was also larger. The situations about abnormal data points were similar to the EH sampling site.

D. PRE-ANALYSIS OF EXPERIMENTAL DATA

The annual median concentrations (i.e., experimental data; after data preprocessing) of $\sum$ PCBs, $\sum$ PESTs, and $\sum$ PAHs for the individual IADN sites are presented in Table 1. The sample sizes for BR CHIC, CLEV, EH, PP, SBD, and STP were 172, 526, 336, 670, 234, 628, and 634 respectively. The main distribution space of $\sum$ PCBs was CHIC, followed by CLEV, STP and PP, while the median $\sum$ PCBs in the SBD and EH samples were only two digits in most years. The spatial distribution patterns of $\sum$ PESTs at CHIC, CLEV, STP, PP, SBD, and EH were not obvious in the early years. However, the median $\sum$ PCBs at STP, PP, SBD, and EH has dropped to two digits recently, whereas both CHIC and CLEV have remained at three digits. The median $\sum$ PAHs were much higher than those of $\sum$ PCBs and $\sum$ PESTs, and the highest levels were observed at CHIC and CLEV, followed by STP, PP, SBD, and EH. The BR data were not discussed due to its small sample and the missing of $\sum$ PESTs data.

Taking the average value of the above experimental data, the spatial distribution of the three kinds of POPs at the seven sampling points is shown in Figure 6, from which we can get a clear understanding of the spatial distribution of PCBs, PESTs, PAHs. Since BR sampling points are completely missing from the sum of PESTs data, so the BR case is not considered in Figure 6 (b). The situation shown in this figure is consistent with the data analysis result in Table 1.

E. MODEL INTRODUCTION

To predict the future trends of POPs' concentration based on the existing IADN data, the Time Series Prediction Method (TSPM), a machine learning algorithm, was utilized. Autoregressive moving average (ARMA) model was one of TSPM [32]. ARMA, the most commonly model for fitting a stationary sequence, can be subdivided into three categories: AR model (auto regression model), MA model (moving average model), and ARMA model (auto regression moving average model). When the time series itself is not stationary, if its increment, that is, a difference, is stable near zero, it can be regarded as a stationary sequence. In practical problems, most of the non-stationary sequences encountered

TABLE 1. Annual median concentrations of POPs in the great lakers atmosphere (after data preprocessing; pg/m3).

	Brule River (N=172)			Chicago (N = 526)			Cleveland (N=336)			Eagle Harbor (N =670)		
	∑PCBs	∑PESTs	∑PAHs	∑PCBs	∑PESTs	∑PAHs	∑PCBs	∑PESTs	∑PAHs	∑PCBs	∑PESTs	∑PAHs
1990										38.0	394.0	642.0
1991										54.0	329.0	1192.0
1992										54.0	268.0	871.0
1993										61.0	185.0	907.0
1994										47.0	180.0	692.0
1995										69.0	181.0	588.0
1996	829.0		914.0	829.0	408.0	53441.0				53.0	159.0	794.0
1997	312.0		1173.0	1293.0	487.0	50887.0				26.0	151.0	908.0
1998	76.0		1062.0	1417.0	411.0	51265.0				39.0	165,0	721.0
1999	125.0		1184.0	1722.0	449.0	49260.0				61.0	159.0	1042.0
2000	106.0		1103.0	904.0	278.0	35034.0				61.0	139.0	819.0
2001	131.0		785.0	738.0	351.0	37489.0				55.0	133.0	873.0
2002	119.0		701.0	551.0	290.0	32100.0				38.0	104.0	631.0
2003				626.0	323.0	25972.0	307.0	230.0	30150.0	24.0	91.0	580.0
2004				572.0	173.0	23398.0	387.0	185.0	30002.0	26.0	76.0	472.0
2005				1083.0	320.0	48431.0	414.0	155.0	38020.0	34.0	72.0	797.0
2006				1189.0	184.0	25879.0	313.0	134.0	19878.0	29.0	72.0	618.0
2007				1428.0	277.0	25161.0	431.0	185.0	22353.0	28.0	67.0	411.0
2008				1320.0	304.0	46689.0	311.0	153.0	21373.0	29.0	62.0	351.0
2009				473.0	180.0	16396.0	243.0	151.0	14760.0	22.0	73.0	396.0
2010				772.0	255.0	25813.0	322.0	243.0	20093.0	24.0	71.0	347.0
2011				524.0	210.0	17867.0	355.0	172.0	13764.0	24.0	76.0	424.0
2012				676.0	280.0	32066.0	318.0	180.0	18794.0	19.0	84.0	796.0
2013				333.0	154.0	13471.0	247.0	121.0	14684.0	25.0	82.0	395.0
2014				448.0	192.0	21742.0	255.0	143.0	19216.0	19.0	67.0	613.0
2015				385.0	202.0	18669.0	375.0	171.0	15600.0	17.0	69.0	619.0
2016				373.0	194.0	22911.0	244.0	165.0	16236.0	17.0	63.0	457.0

can become stationary time series after one or more differences [33].

The data sequence formed by the prediction index over time is regarded as a random sequence, and the dependence of this group of random variables reflects the continuity of the original data in time. On the one hand, the influence of the influencing factors, on the other hand, it has its own rule of change, assuming that the influencing factors are $X_1, X_2, \ldots, X_{k,}$, $\beta_i (i = 0, 1, 2, \ldots)$ are coefficients, by regression analysis, as shown below.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + e, \tag{1}$$

where Y is the observed value of the predicted object, and e is the error. As the predicted object, Y_t is affected by its own changes, and its law can be reflected by the following formula:

$$Y_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_p X_{t-p} + e_t, \tag{2}$$

The error term is dependent in different periods, which is expressed by the following formula, $\alpha_i (i = 0, 1, 2, \ldots)$ are coefficients, μ_t are random term to test the co-integration relationship.

$$e_t = \alpha_0 + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \ldots + \alpha_q e_{t-q} + \mu_t, \tag{3}$$

Thus, the ARMA model expression can be obtained:

$$Y_t = \beta_0 + \beta_1 X_{t-1} + \ldots + \beta_p X_{t-p} + \alpha_0 + \alpha_1 e_{t-1} \ldots + \alpha_q e_{t-q} + \mu_t. \tag{4}$$

The ARMA model building algorithm is seen in Figure 7 (p refers to the autoregression order, and q refers to the moving average order). Firstly, inputting the processed time series. Then, judging its stationarity. If the time series is stable, moving to the next step; otherwise, make a differential operation to make the time series keep stable. Next, determining the orders of model. Finally, parameter estimating is performed to evaluate the model performance. If the performance of the model is good, the model can be determined for prediction; otherwise, the orders of the model needs to be determined again.

TABLE 1. *(Continued.)* **Annual median concentrations of POPs in the great lakers atmosphere (after data preprocessing; pg/m3).**

	Point Petre (N =234)			Sleeping Bear Dunes (N =628)			Sturgeon Point (N =634)		
	∑PCBs	∑PESTs	∑PAHs	∑PCBs	∑PESTs	∑PAHs	∑PCBs	∑PESTs	∑PAHs
1990									
1991				42.0	301.0	2536.0	119.0	241.0	4387.0
1992				103.0	327.0	1481.0	203.0	292.0	6466.0
1993				184.0	252.0	1520.0	260.0	358.0	5518.0
1994				112.0	257.0	1520.0	213.0	293.0	5518.0
1995				51.0	179.0	1098.0	190.0	310.0	5788.0
1996				56.0	151.0	1002.0	115.0	189.0	4564.0
1997				58.0	171.0	884.0	184.0	216.0	4172.0
1998	836.0	175.0	2332.0	99.0	201,0	1312.0	192.0	188.0	5518.0
1999	106.0	155.0	3218.0	196.0	229.0	1616.0	220.0	187.0	5272.0
2000	225.0	218.0	1579.0	70.0	166.0	941.0	177.0	225.0	4247.0
2001	109.0	161.0	1867.0	49.0	132.0	1025.0	150.0	191.0	4448.0
2002	44.0	115.0	1520.0	39.0	104.0	1102.0	111.0	154.0	3040.0
2003	55.0	113.0	1226.0	44.0	139.0	861.0	130.0	162.0	3121.0
2004	39.0	67.0	1009.0	35.0	80.0	914.0	224.0	130.0	4103.0
2005	83.0	114.0	1906.0	49.0	91.0	1225.0	152.0	91.0	5313.0
2006	82.0	100.0	1340.0	37.0	71.0	744.0	124.0	101.0	3238.0
2007	71.0	97.0	1435.0	43.0	89.0	861.0	142.0	115.0	5001.0
2008	85.0	125.0	1923.0	75.0	67.0	834.0	168.0	115.0	6306.0
2009	52.0	99.0	1158.0	57.0	84.0	694.0	101.0	75.0	1610.0
2010	60.0	113.0	823.0	65.0	88.0	829.0	96.0	73.0	1808.0
2011	52.0,	105.0	955.0	54.0	102.0	1181.0	85.0	85.0	2266.0
2012	62.0	130.0	1285.0	39.0	111.0	1177.0	119.0	163.0	4102.0
2013	99.0	114.0	1496.0	39.0	83.0	1003.0	73.0	94.0	2000.0
2014	49.0	103.0	762.0	34.0	77.0	899.0	57.0	99.0	2796.0
2015	51.0	94.0	880.0	34.0	123.0	1257.0	80.0	84.0	3030.0
2016	31.0	91.0	1128.0	29.0	85.0	1193.0	74.0	95.0	2689.0

III. RESULTS AND DISCUSSION

A. DATA TREND OVERVIEWING

The temporal trends of median ∑ PCBs, ∑ PESTs and ∑ PAHs at the seven IADN sampling sites are demonstrated in Figure 8. At BR sampling site, the ∑ PAHs were higher than that of ∑ PCBs. While the ∑ PCBs leveled off at relatively low levels, the ∑ PAHs tended to decrease. Similar trends were observed for CHIC, CLEV, EH, PP, SBD, and STP where ∑ PAHs were significantly greater than the ∑ PCBs, while ∑ PESTs were the lowest. Additionally, although the median concentrations of these POPs fluctuated, the overall trends were all downward.

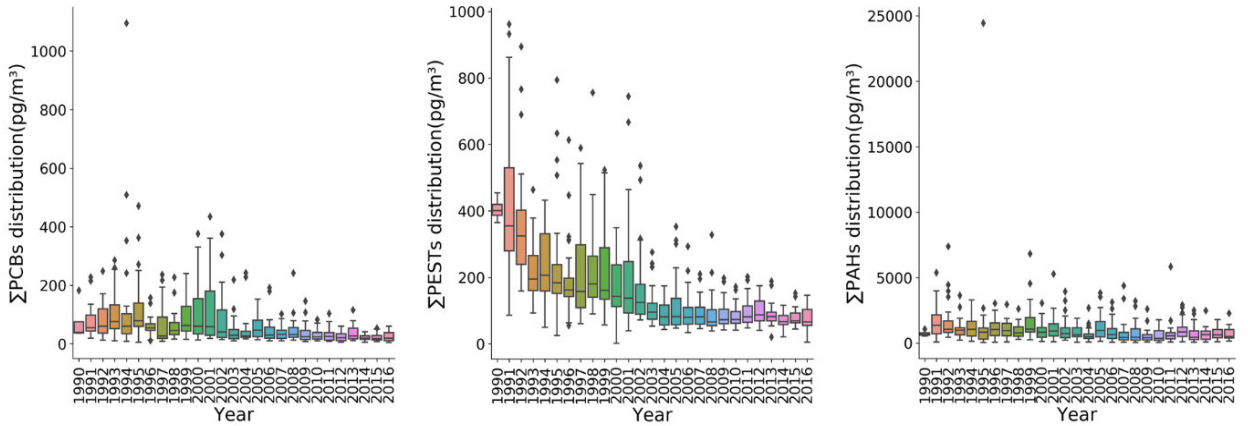
B. ANALYSIS OF VISUALIZED RESULT

In this study, by constantly optimizing the parameters to better fit of data, an data-driven intelligent environmental model based on ARMA algorithm was constructed to predict POPs concentrations in the next few years. Given that EH and STP have the largest sample sizes, in addition, STP stands for rural site and EH stands for remote site, which makes the sampled data rich and representative. So we used their data to predict the ∑ PCBs, ∑ PESTs, and ∑ PAHs in the following 4-5 years (see Figure 9).

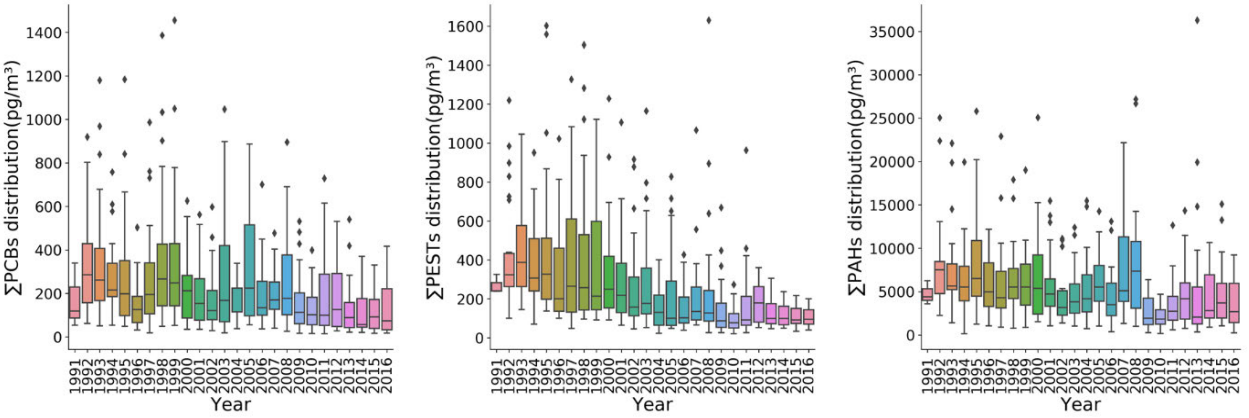
According to Figure 9 (a), in the EH atmosphere, there will be a slight increase in ∑ PCBs, but such fluctuations will not be considerable in general, and the overall temporal trend will remain declining. The ∑ PESTs will continue dropping in a few years after 2016. The ∑ PAHs will fluctuate to some extent, showing a zigzag pattern. Regrading the STP atmosphere (see Figure 9 (b)), its ∑ PCBs will decline with fluctuations. The ∑ PESTs will gain right after 2016, and then level off, and the ∑ PAHs will fluctuate slightly. The concentrations of pops at EH and STP sampling sites have slight fluctuations, but on the whole, they showed a downward trend.

C. MODEL EVALUTION

This article intends to evaluate the ARMA model which was constructed from the following two aspects: feasibility analysis and sensitivity analysis.



(a) Box plots of POPs in Eagle Harbor（1990-2016）



(b) Box plots of POPs in Sturgeon Point（1991-2016）

FIGURE 5. Box plots of POPs concentrations in the Great Lakes atmosphere.

1) FEASIBILITY ANALYSIS

The feasibility analysis of ARMA model was coducted from two aspects: 1) testing whether the residuals were normally distributed using the QQ plot(Quantile-Quantile plot); and 2) assessing autocorrelation of residuals in terms of the D-W(Durbin-Watson) statistics (when the value of D-W is significantly close to 0 or 4, there is autocorrelation; when it is close to 2, there is no autocorrelation). Figure 10 shows the QQ plots for the ARMA modelling of median $\sum$ PCBs, $\sum$ PESTs, and $\sum$ PAHs in the EH and STP samples.

Figure 10 (a) is the QQ plot of the ARMA model for the median $\sum$ PCBs, $\sum$ PESTs and $\sum$ PAHs in the EH atmosphere. Virtually all the data points were on a straight line and equally distribute on both sides of the line, indicating that the residuals satisfactorily met the normal distribution. The corresponding D-W values for $\sum$ PCBs, $\sum$ PESTs and $\sum$ PAHs were, respectively, 1.85, 1.47, 1.97, all were closer to 2, suggestive of insignificant autocorrelation. Similar residual distribution patterns and D-W values were also found for STP samples (see Figure 10 (b)). Therefore, our model prediction should be robust and reliable.

2) SENSITIVITY ANALYSIS

Sensitivity Analysis (SA) investigates how the variation in the output of a numerical model can be attributed to variations of its input factors [34]. Figure 11 shows input factors and output definition for the SA of the model. As we can see from the figure, the main input factors included processed time series, model parameters, differential orders and model fixed orders. These input factors all have certain influence on the predicted value of the model.

Based on the model type, we used correlation methods to preform sensitivity analysis of this model. We can define the sensitivity of the model as follows:

$$S_A = correlation\,(X_i, Y) \tag{5}$$

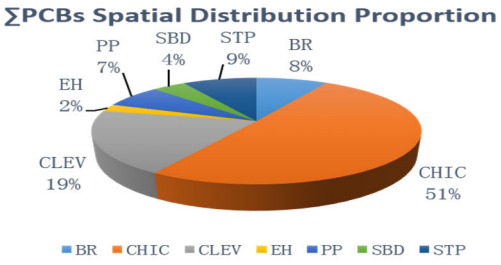
where X_i represents input factors and Y represents output. The sensitivity of model can be analyzed by means of comparing *RMSE*, the error between the predicted data and the real data. Its mathematical expression is shown as (6).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(Y_{predict,i} - Y_{real,i}\right)^2}{n}} \tag{6}$$

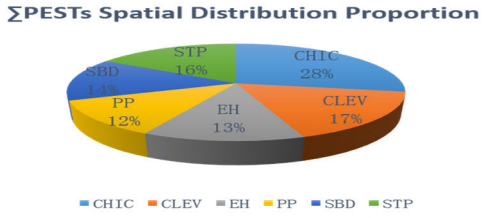
where $Y_{predict,i}$ is the predicted value of the model, $Y_{real,i}$ is the true value of the model, and n is the number of data points. *RMSE* can well reflect the fitting status of model data. Within a certain range, the smaller *RMSE*, the higher the Fitting degree [35], [36]. There are two most important parameters p and q in this model, p refers to the autoregression order,

TABLE 2. Comparison of *RMSE* values corresponding to different *p* and *q* valuess.

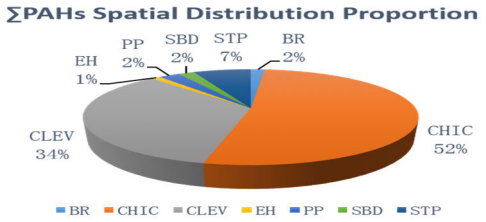
The corresponding *RMSE*	EH			STP		
	∑PCBs	∑PESTs	∑PAHs	∑PCBs	∑PESTs	∑PAHs
1)p=1,q=1	**10.0**	40.8	187.1	141.5	42.1	187.0
2)p=0,q=1	35.8	**10.4**	201.7	144.4	59.1	197.6
3)p=1,q=0	30.3	71.5	**30.3**	45.6	66.7	100.1
4)p=7,q=0	38.2	54.4	267.8	**13.0**	78.0	123.2
5)p=7,q=1	67.0	67.4	178.4	65.6	**15.8**	178.0
6)p=8,q=0	35.0	68.0	178.0	685.0	40.7	**28.8**



（a）∑PCBs spatial distribution proportion



(b) ∑PESTs spatial distribution proportion



(c) ∑PAHs spatial distribution proportion

FIGURE 6. Pie charts of three kinds of POPs spatial distribution proportion at seven sampling points.

and q refers to the moving average order, that is, the fixed order of the model(main input factors), which can be adjusted to better fit the data mainly. By changing the p and q value, calculating the corresponding *RMSE* value and comparing, the sensitivity of the model can be analyzed. We changed the value of p, q for many times and found that the *RMSE* value changed a lot (See Table 2). In Table 2, *RMSE* in the best fitting case was highlighted, and its corresponding p and q values were selected to build the model. Further analysis of the data in Table 2 showed that *RMSE* values changed to a large extent with the change of input factors p and q, which means that the change of input factors has a significant impact on the output of this model, namely, the model is highly sensitive.

```
Input:  The time series S
Output: The results of predictions R
Begin
  Inputting time series S
  IF S is the stationary sequence THEN
    Determining the order of p and q
    Parameter estimating
    IF the model pass the adaptability test THEN
      Making predictions and return R
    ELSE
      GOTO 4)
  ELSE
    Carrying out data stationary processing GOTO 3)
End
```

FIGURE 7. The algorithm of ARMA model building.

IV. ENVIRONMENTAL SIGNIFICANCE

PCBs have been widely used as insulation oil, heat carrier, and lubricating oil. They can also be used as additive in many industrial products (such as various resins, rubber, binders, coatings, carbon paper, ceramic glaze, fire retardant, pesticide extender and dye dispersant) [37]. PCBs are carcinogens that tend to accumulate in fatty tissue, it can cause diseases of the brain, skin and internal organs, and it also affects the nervous, reproductive, and immune system [38]. Thus, they were added to the Stockholm convention and have been banned in most countries (including the United States and Canada) for decades, which consistent with the downward trend we found and predicted.

PESTs, which can also accumulate in tissues, such as heart, liver and kidney, and enter human and animal bodies through the food chain [39]. The pesticides which accumulate in bodies can also be excreted through the mother's milk, or into the egg, ultimately affecting the offspring [40]. Therefore, countries strictly control the residues of PESTs in food. For example, Germany, the United States, Japan, and many other countries do not allow cyclopentadienyls PESTs to be detected in food. In the 1960s, China began to ban the use of DDT and 666 on tobacco, vegetables, tea, and other crops [41]. Though the ∑ PESTs decreased in the Great Lakes atmosphere, the IADN team measured legacy PESTs only. The environmental concentration of emerging pesticides is expected to increase due to the phase-out of traditional PESTs and the introduction of many alternatives to them because of the market demand.

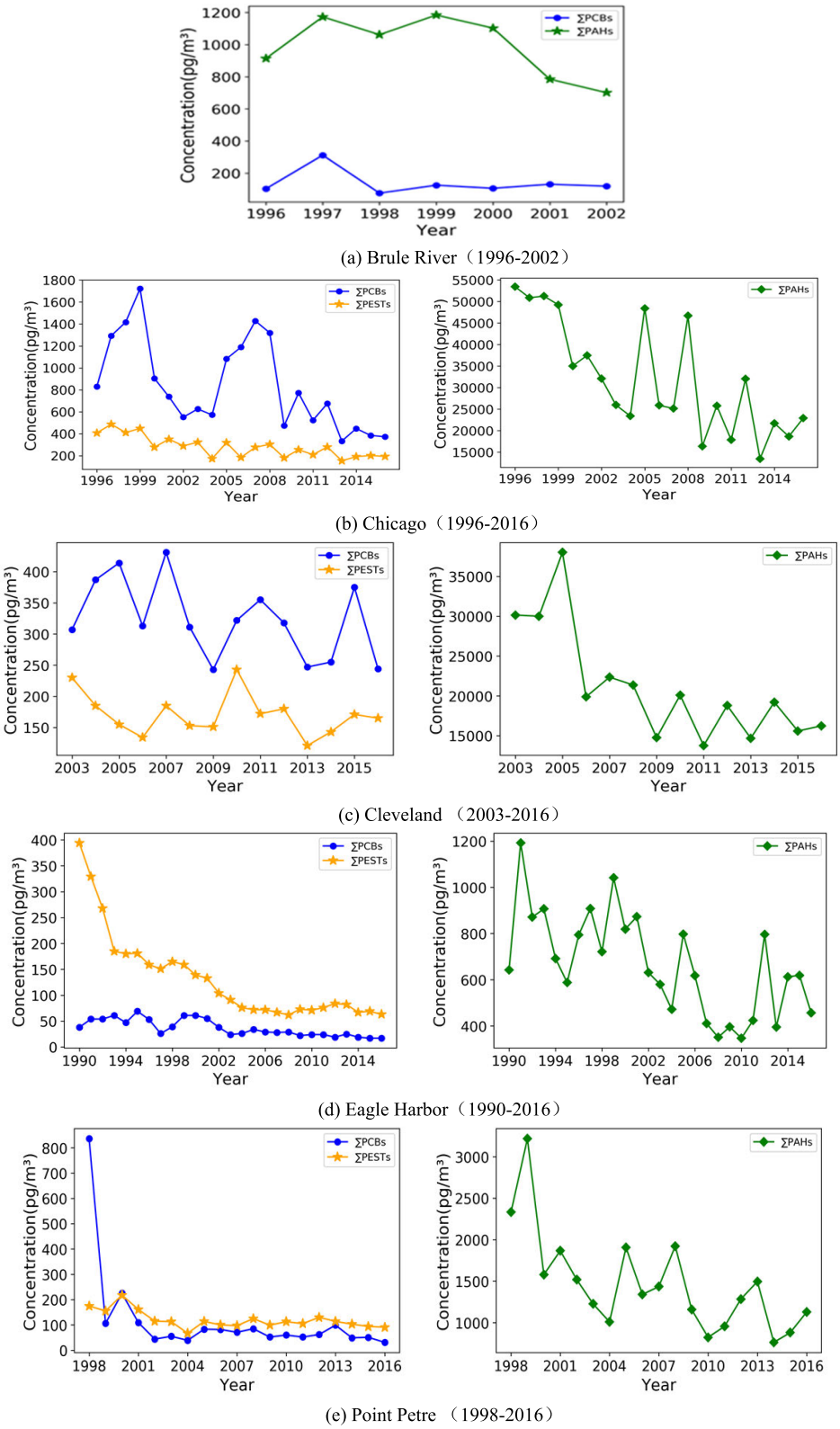
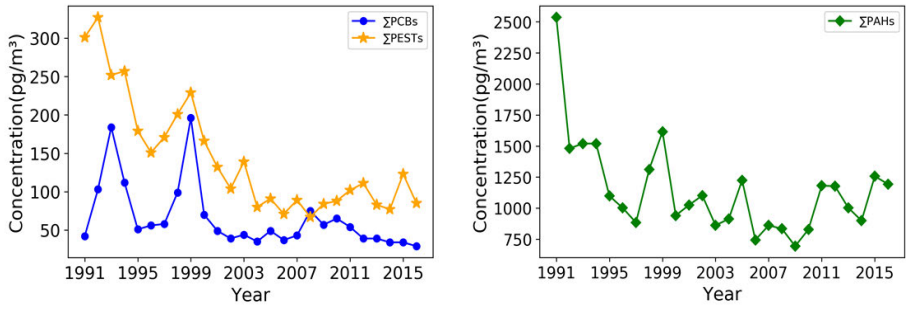


FIGURE 8. Broken line diagrams for median ∑ PCBs, ∑ PESTs, and ∑ PAHs in the Great Lakes atmosphere.

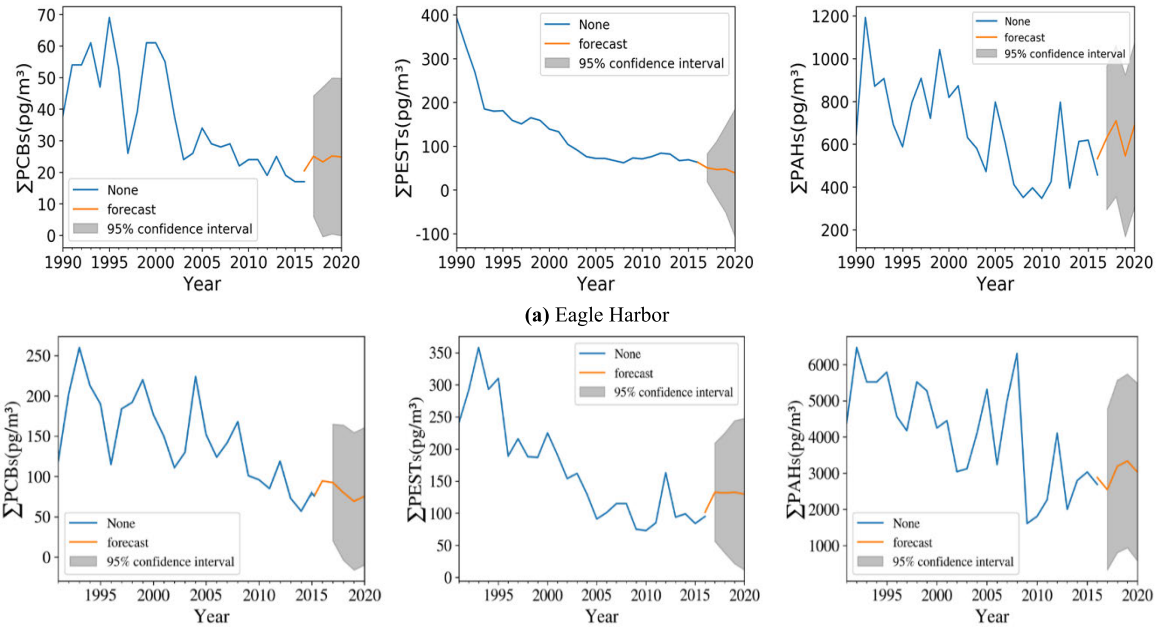
PAHs are ubiquitous in the environment, and mainly comes from the burning of coal and oil, but also from garbage incineration or forest fires. Their production volumes are closely related to combustion equipment and combustion temperature [42]. PAHs are found in the exhaust of diesel and gasoline engines, as well as in the waste gas and water from refineries, coal tar processing plants and asphalt processing plants [43]. The decreasing trend of ∑ PAHs in the Great Lakes



(f) Sleeping Bear Dunes（1991-2016）

(g) Sturgeon Point（1991-2016）

FIGURE 8. ***(Continued.)*** **Broken line diagrams for median ∑ PCBs, ∑ PESTs, and ∑ PAHs in the Great Lakes atmosphere.**



(a) Eagle Harbor

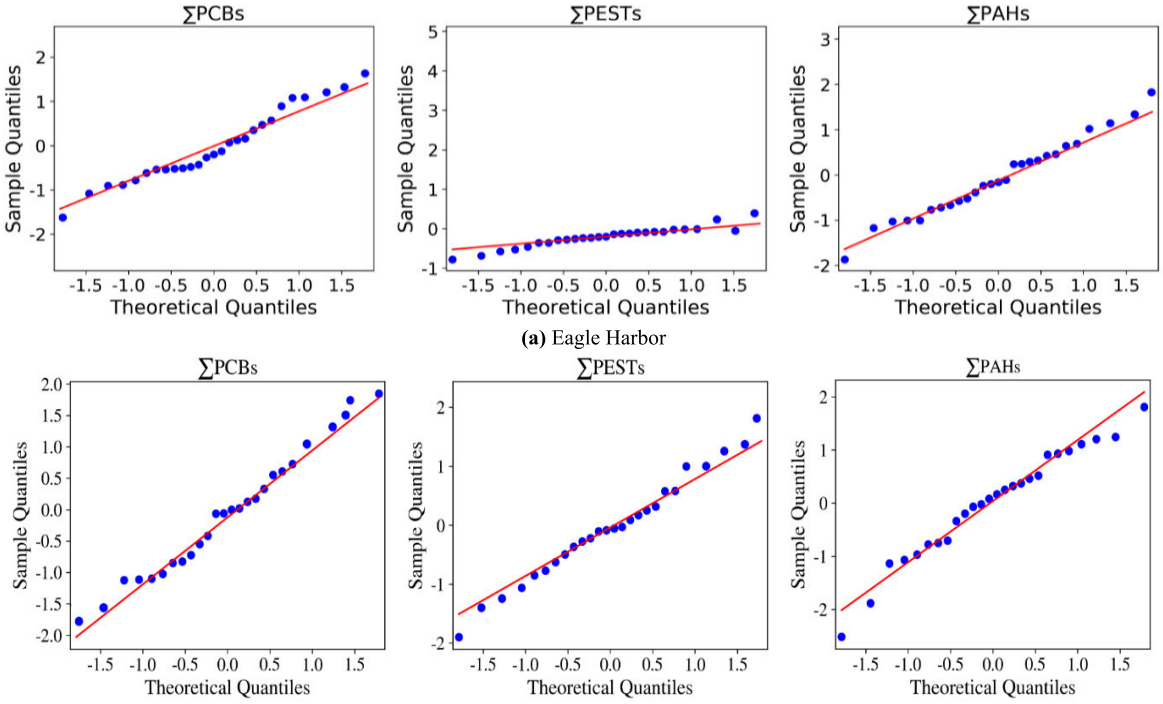
(b) Sturgeon Point

FIGURE 9. **ARMA model prediction of median ∑ PCBs, ∑ PESTs, and ∑ PAHs in the EH (a) and STP (b) atmosphere.**

atmosphere indicates that the elevated energy efficiency and the vehicle emission control in North America have alleviated the PAH formation in recent years [44], [45].

These three classes of POPs not only have posed risks to the environment but also have adversely affected human health. Therefore, they should be used with caution. Our results illustrate that the ∑ PCBs, ∑ PESTs, and ∑ PAHs have been decreasing on the whole, consistent with the fact that many countries have released policies to limit their use and emission [46] due to their recognized toxic potentials.

IoT, by using local network or Internet and other communication technologies, sensors, controllers and machines, people and objects can be connected together in new ways to form people-and-objects, and objects-and-objects links, so as to realize information-based remote management control and intelligent network [47]. IoT was widely used in environmental protection and plays a very important role.



(a) Eagle Harbor

(b) Sturgeon Point

FIGURE 10. QQ plots of ARMA prediction model for median ∑ PCBs, ∑ PESTs, and ∑ PAHs in the EH (a) and STP (b) atmosphere.

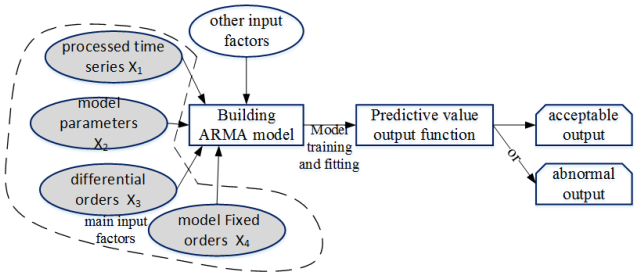


FIGURE 11. Input factors and output definition for the SA of ARMA model.

The researches of this topic is carried out in the the Internet of Things environments. Firstly, Environment-related data are collected based on the technology of the Internet of Things. Secondly, required data are obtained from the Internet. Thirdly, Machine learning and data analysis technologies are used for research. The last one, the research results can be uploaded and shared. IoT has played a great role in intelligent environmental protection.

V. CONCLUSION AND FUTURE WORK

We analyzed ∑ PCBs, ∑ PAHs, ∑ PESTs in the Great Lakes atmospheric samples (vapor phase) collected from seven sampling sites by the IADN team at Indiana University. By constructing an data-driven intelligent environmental model, ∑ PCBs, ∑ PESTs, and ∑ PAHs in the EH and STP samples were predicted for the following 4-5 years, We also presented the detailed processes of modeling workflow, and used the Python development language and tools to visualize our results. The result showed concentrations would continue declining with slight fluctuations and the model is feasible and highly sensitive. In addition, we pointed out the important role of the IoT in the smart environmental protection.

The future research in this field can be carried out from the following three aspects. Firstly, studying the relationship between each persistent organic pollutant particle-phase concentration percentage and temperature by constructing appropriate regression model. The IADN collected concentrations of persistent organic pollutants in vapor and particle phases. Physical phenomena shows that matter changes from solid to gas as the temperature rises. Calculating the persistent organic pollutant particle phase percentage, that is P/(P+V) (where P and V are the particle- and vapor-phase concentrations) and obtain the corresponding temperature. Taking temperature as independent variable and persistent organic pollutant particle phase percentage as dependent variable, an appropriate regression model can be tried to be established. The effect of temperature on the state change of POPs can be clearly understood. Secondly, studying the composition of POPs in the Great Lakes by constructing an appropriate classification model. POPs such as Poly-chlorinated Biphenyls (PCBs), Pesticides (PESTs), and Polycyclic Aromatic Hydrocarbons (PAHs) are composed of a variety of different molecular weight substances. The IADN website also recorded concentrations of these substances. We can calculate the proportion of the concentration of these substances and then determine the main components and secondary components of POPs by constructing a classification model. Finally, IoT system for environmental protection can be built to realize data real-time collection, upload-ing and sharing of research results, so as to realize information and digitization of

environmental protecting and make environmental management more scientific and efficient.

ACKNOWLEDGMENT

The data for this work is downloaded from Indiana University's IADN Data Viz website. The authors thank Dr. Yan Wu at Indiana University for useful discussions.

REFERENCES

[1] C. A. Johnston, B. L. Bedford, M. Bourdaghs, T. Brown, C. Frieswyk, M. Tulbure, L. Vaccaro, and J. B. Zedler, "Plant species indicators of physical environment in great lakes coastal wetlands," *J. Great Lakes Res.*, vol. 33, no. 3, pp. 106–124, Dec. 2007.
[2] J. H. Hartig, G. Krantzberg, and P. Alsip, "Thirty-five years of restoring great lakes areas of concern: Gradual progress, hopeful future," *J. Great Lakes Res.*, vol. 46, no. 3, pp. 429–442, Jun. 2020.
[3] M. Venier and R. A. Hites, "Atmospheric deposition of PBDEs to the great lakes featuring a Monte Carlo analysis of errors," *Environ. Sci. Technol.*, vol. 42, no. 24, pp. 9058–9064, Dec. 2008.
[4] R. A. Hites, "Statistical approach for assessing the Stockholm convention's effectiveness: Great lakes atmospheric data," *Environ. Sci. Technol.*, vol. 53, pp. 8585–8590, Jul. 2019.
[5] R. A. Hites, "Atmospheric concentrations of PCB-11 near the great lakes have not decreased since 2004," *Environ. Sci. Technol. Lett.*, vol. 5, no. 3, pp. 131–135, Mar. 2018.
[6] A. Salamova, Y. Ma, M. Venier, and R. A. Hites, "High levels of organophosphate flame retardants in the great lakes atmosphere," *Environ. Sci. Technol. Lett.*, vol. 1, no. 1, pp. 8–14, Jan. 2014.
[7] R. A. Hites, "A statistical approach for left-censored data: Distributions of atmospheric polychlorinated biphenyl concentrations near the great lakes as a case study," *Environ. Sci. Technol. Lett.*, vol. 2, no. 9, pp. 250–254, Sep. 2015.
[8] D. C. Lehman, J. C. Bays, and R. A. Hites, "Precision of atmospheric persistent organic pollutant concentration measurements," *Environ. Sci. Technol.*, vol. 50, no. 24, pp. 13464–13469, Dec. 2016.
[9] K. Gao, G. Mei, F. Piccialli, S. Cuomo, J. Tu, and Z. Huo, "Julia language in machine learning: Algorithms, applications, and open issues," *Comput. Sci. Rev.*, vol. 37, Aug. 2020, Art. no. 100254.
[10] T. Jiang, J. L. Gradus, and A. J. Rosellini, "Supervised machine learning: A brief primer," *Behav. Therapy*, vol. 51, no. 5, pp. 675–687, Sep. 2020.
[11] Z. Li, S. H.-L. Yim, and K.-F. Ho, "High temporal resolution prediction of street-level $PM_{2.5}$ and NOx concentrations using machine learning approach," *J. Cleaner Prod.*, vol. 268, Sep. 2020, Art. no. 121975.
[12] J. Li, N. Xiong, J. H. Park, C. Liu, S. Ma, and S. Cho, "Intelligent model design of cluster supply chain with horizontal cooperation," *J. Intell. Manuf.*, vol. 23, no. 4, pp. 917–931, Aug. 2012.
[13] M. M. Hassan, M. R. Hassan, S. Huda, and V. H. C. de Albuquerque, "A robust deep learning enabled trust-boundary protection for adversarial industrial IoT environment," *IEEE Internet Things J.*, early access, Aug. 24, 2020, doi: 10.1109/JIOT.2020.3019225.
[14] H. Liao, Z. Zhou, X. Zhao, L. Zhang, S. Mumtaz, A. Jolfaei, S. H. Ahmed, and A. K. Bashir, "Learning-based context-aware resource allocation for edge-computing-empowered industrial IoT," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4260–4277, May 2020.
[15] F. Long, N. Xiong, A. V. Vasilakos, L. T. Yang, and F. Sun, "A sustainable heuristic QoS routing algorithm for pervasive multi-layered satellite wireless networks," *Wireless Netw.*, vol. 16, no. 6, pp. 1657–1673, Aug. 2010.
[16] K. Huang, Q. Zhang, C. Zhou, N. Xiong, and Y. Qin, "An efficient intrusion detection approach for visual sensor networks based on traffic pattern learning," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 47, no. 10, pp. 2704–2713, Oct. 2017.
[17] Y. Yang, N. Xiong, N. Y. Chong, and X. D, "A decentralized and adaptive flocking algorithm for autonomous mobile robots," in *Proc. 3rd Int. Conf. Grid Pervas. Comput. Workshops*, May 2008, pp. 262–268.
[18] Z. Zhou, Z. Wang, H. Yu, H. Liao, S. Mumtaz, L. Oliveira, and V. Frascolla, "Learning-based URLLC-aware task offloading for Internet of health things," *IEEE J. Sel. Areas Commun.*, early access, Sep. 4, 2020, doi: 10.1109/JSAC.2020.3020680.
[19] L. Knoll, L. Breuer, and M. Bach, "Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning," *Sci. Total Environ.*, vol. 668, pp. 1317–1327, Jun. 2019.
[20] V. Nourani, A. A. Mogaddam, and A. O. Nadiri, "An ANN-based model for spatiotemporal groundwater level forecasting," *Hydrolog. Processes*, vol. 22, no. 26, pp. 5054–5066, Dec. 2008.
[21] W. Guo, N. Xiong, A. V. Vasilakos, G. Chen, and C. Yu, "Distributed k-connected fault-tolerant topology control algorithms with PSO in future autonomic sensor systems," *Int. J. Sensor Netw.*, vol. 12, no. 1, p. 53, 2012.
[22] N. Xiong, Z. Wan, N. Ghani, A. V. Vasilakos, and L. Zhou, "Adaptive unequal protection for wireless video transmission over IEEE 802.11 e networks," *Multimedia Tools Appl.*, vol. 72, no. 1, pp. 541–571, 2014.
[23] E. Fijani, A. A. Nadiri, A. A. Moghaddam, F. T.-C. Tsai, and B. Dixon, "Optimization of DRASTIC method by supervised committee machine artificial intelligence to assess groundwater vulnerability for Maragheh–Bonab plain aquifer, iran," *J. Hydrol.*, vol. 503, pp. 89–100, Oct. 2013.
[24] B. Dixon, "A case study using support vector machines, neural networks and logistic regression in a GIS to identify wells contaminated with nitrate-N," *Hydrogeology J.*, vol. 17, no. 6, pp. 1507–1520, Sep. 2009.
[25] A. Raza, S. Bardhan, L. Xu, S. S. Yamijala, C. Lian, H. Kwon, and B. M. Wong, "A machine learning approach for predicting defluori-nation of per- and polyfluoroalkyl substances (PFAS) for their efficient treatment and removal," *Environ. Sci. Technol. Lett.*, vol. 6, no. 10, pp. 624–629, 2019.
[26] W. Cheng and C. A. Ng, "Using machine learning to classify bioactivity for 3486 per- and polyfluoroalkyl substances (PFASs) from the OECD list," *Environ. Sci. Technol.*, vol. 53, no. 23, pp. 13970–13980, Dec. 2019.
[27] S. Pei, T. Shen, X. Wang, C. Gu, Z. Ning, X. Ye, and N. Xiong, "3DACN: 3D augmented convolutional network for time series data," *Inf. Sci.*, vol. 513, pp. 17–29, Mar. 2020.
[28] J. Sun, Y. Yang, N. N. Xiong, L. Dai, X. Peng, and J. Luo, "Complex network construction of multivariate time series using information geometry," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 1, pp. 107–122, Jan. 2019.
[29] R. Zhou, C. Yang, J. Wan, W. Zhang, B. Guan, and N. Xiong, "Measuring complexity and predictability of time series with flexible multiscale entropy for sensor networks," *Sensors*, vol. 17, no. 4, p. 787, Apr. 2017.
[30] M. Venier, A. Salamova, and R. A. Hites, "Halogenated flame retardants in the great lakes environment," *Accounts Chem. Res.*, vol. 48, no. 7, pp. 1853–1861, Jul. 2015.
[31] *Integrated Atmospheric Deposi-tion Network (IADN) Data Visualization Tool, IADN Data Viz*, Indiana Univ., Bloomington, IN, USA, 2020.
[32] M. Valipour, M. E. Banihabib, and S. M. R. Behbahani, "Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of dez dam reservoir," *J. Hydrol.*, vol. 476, pp. 433–441, Jan. 2013.
[33] diggerTT. (2018). *Algorithmmodel-Timeseriesmodel*. [Online]. Available: http://blog.csdn.net/qing-qing7/article/details/79070722
[34] F. Pianosi, K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, and T. Wagener, "Sensitivity analysis of environmental models: A systematic review with practical workflow," *Environ. Model. Softw.*, vol. 79, pp. 214–232, May 2016.
[35] L. Somappa, A. G. Menon, A. K. Singh, A. A. Seshia, and M. S. Baghini, "A portable system with 0.1-ppm RMSE resolution for 1–10 MHz resonant MEMS frequency measurement," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 7146–7157, Sep. 2020.
[36] W. Wu, N. Xiong, and C. Wu, "Improved clustering algorithm based on energy consumption in wireless sensor networks," *IET Netw.*, vol. 6, no. 3, pp. 47–53, May 2017.
[37] Q. Sun, *Establishment of a method for Determining of PCBs in Human Plasma*. Shanghai, China: East China Univ. Science and Technology, 2009.
[38] B. Lauby-Secretan, D. Loomis, Y. Grosse, F. E. Ghissassi, V. Bouvard, L. Benbrahim-Tallaa, N. Guha, R. Baan, H. Mattock, and K. Straif, "Carcinogenicity of polychlorinated biphenyls and polybrominated biphenyls," *Lancet Oncol.*, vol. 14, no. 4, pp. 287–288, 2013.
[39] J. Gilbert, "The fate of environmental contaminants in the food chain," *Sci. Total Environ.*, vol. 143, no. 1, pp. 103–111, Mar. 1994.
[40] M. Anand and A. Taneja, "Organochlorine pesticides residue in placenta and their influence on anthropometric measures of infants," *Environ. Res.*, vol. 182, Mar. 2020, Art. no. 109106.
[41] R. B. Rajendran, T. Imagawa, H. Tao, and R. Ramesh, "Distribution of PCBs, HCHs and DDTs, and their ecotoxicological implications in bay of bengal, india," *Environ. Int.*, vol. 31, no. 4, pp. 503–512, May 2005.
[42] G. Zafra, T. D. Taylor, A. E. Absalón, and D. V. Cortés-Espinosa, "Comparative metagenomic analysis of PAH degradation in soil by a mixed microbial consortium," *J. Hazardous Mater.*, vol. 318, pp. 702–710, Nov. 2016.
[43] B. K. Behera, A. Das, D. J. Sarkar, P. Weerathunge, P. K. Parida, B. K. Das, P. Thavamani, R. Ramanathan, and V. Bansal, "Polycyclic aromatic hydrocarbons (PAHs) in inland aquatic ecosystems: Perils and remedies through biosensors and bioremediation," *Environ. Pollut.*, vol. 241, pp. 212–233, Oct. 2018.

[44] L.-Y. Liu, P. Kukucka, M. Venier, A. Salamova, J. Klánová, and R. A. Hites, "Differences in spatiotemporal variations of atmospheric PAH levels between North America and Europe: Data from two air monitoring projects," *Environ. Int.*, vol. 64, pp. 48–55, Mar. 2014.
[45] H. Shen, S. Tao, R. Wang, B. Wang, G. Shen, W. Li, S. Su, Y. Huang, X. Wang, and W. Liu, "Global time trends in PAH emissions from motor vehicles," *Atmos. Environ.*, vol. 45, no. 12, pp. 2067–2073, Apr. 2011.
[46] M. F. Ruggeri, N. B. Lana, J. C. Altamirano, and S. E. Puliafito, "Spatial distribution, patterns and source contributions of POPs in the atmosphere of great Mendoza using the WRF/CALMET/CALPUFF modelling system," *Emerg. Contaminants*, vol. 6, pp. 103–113, 2020.
[47] E. P. Yadav, E. A. Mittal, and H. Yadav, "IoT: Challenges and issues in indian perspective," in *Proc. 3rd Int. Conf. Internet Things, Smart Innov. Usages (IoT-SIU)*, Feb. 2018, pp. 1–5.

CHUNXUE WU received the Ph.D. degree in control theory and control engineering from the China University of Mining and Technology, Beijing, China, in 2006. He is currently a Professor with the Computer Science and Engineering and Software Engineering Division, School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, China. His research interests include wireless sensor networks, distributed and embedded systems, wireless and mobile systems, and networked control systems.

CHENG WANG received the bachelor's degree in information management and information system from the Wannan Medical College. She is currently pursuing the master's degree in software engineering with the University of Shanghai for Science and Technology.

QINGFENG FAN received the Ph.D. degree in computer science from the Laboratory PRISM (CNRS3709), University of Versailles-Saint-Quentin, Versailles, France, and the Ph.D. degree in computer science and technology from Wuhan University, China. Before he started his Ph.D. research at the University of Versailles-Saint-Quentin, he has worked as a Research Assistant with the Laboratory MAS, Ecole Centrale de Paris, France, and the Department of Computer Science, City University of Hong Kong. He is currently an Associate Professor with the School of Computer Science and Information Engineering, Hubei University, Wuhan, China. He has published more than 20 refereed papers in journals and conference proceedings. His research interests include cloud computing, networks, game theory, database, and system engineering.

QIONGLI WU received the Ph.D. degree in applied mathematics from the Ecole Centrale Paris, France, in 2014, and the Ph.D. degree in communications and information system from Wuhan University, Wuhan, China. She is currently an Associate Professor with the Innovation Academy for Precision Measurement Science and Technology (APM), Chinese Academy of Sciences (CAS). She has published more than 30 peer-reviewed papers in journals and conferences proceedings. Her research interests include wireless communications, data analysis, applied statistics, and optimization theory.

SHENG XU received the Ph.D. degree in management science and engineering from the Jiangxi University of Finance and Economics, Nanchang, China, in 2010. He is currently a Professor with the Management Science and Engineering Division, Department of Vocational Normal, Jiangxi Agricultural University, China. His research interests include human resource management and data mining, learning performance, and networked control systems.

NEAL N. XIONG received the Ph.D. degree in sensor system engineering from Wuhan University, in 2007, and the Ph.D. degree in dependable communication networks from the Japan Advanced Institute of Science and Technology, in 2008.

He is currently an Associate Professor with the Department of Mathematics and Computer Science, Northeastern State University, OK, USA. Before he attended Northeastern State University, he worked with Georgia State University, Wentworth Technology Institution, and Colorado Technical University (a Full Professor about five years) about ten years. He has published more than 200 international journal articles and more than 100 international conference papers. Some of his works were published in IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE or ACM TRANSACTIONS, ACM Secom workshop, IEEE INFOCOM, ICDCS, and IPDPS. His research interests include cloud computing, security and dependability, parallel and distributed computing, networks, and optimization theory.

Dr. Xiong is a Senior Member of the IEEE Computer Society. He has received the Best Paper Award in the 10th IEEE International Conference on High Performance Computing and Communications (HPCC-2008) and the Best student Paper Award in the 28th North American Fuzzy Information Processing Society Annual Conference (NAFIPS2009). He is also the Chair of Trusted Cloud Computing Task Force, IEEE Computational Intelligence Society (CIS), and the Industry System Applications Technical Committee. He has been the General Chair, the Program Chair, the Publicity Chair, a Program Committee member, and an Organizing Committee Member of more than 100 international conferences, and as a Reviewer of about 100 international journals, including IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS (Park: A/B/C), IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON MOBILE COMPUTING, and IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS. He has been serving as the Editor-in-Chief, an Associate editor, or an Editor member for more than ten international journals, including an Associate Editor for IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS and *Information Sciences*, the Editor-in-Chief for *Journal of Internet Technology (JIT)* and *Journal of Parallel and Cloud Computing (PCC)*, and a Guest Editor for more than ten international journals, including *Sensors* journal, WINET, and MONET.