

Effect of Persistent Organic Pollutants on malaria in Chinese mainland: a population-based surveillance and modelling study

Guolong Qu

Air Force Medical Center, PLA, Air Force Medical University, Beijing, 100142, China.

Jianqiang Han

Department of medical engineering, Air Force Medical Center, PLA, Air Force Medical University, Beijing, 100142, China.

Zhenyao Song

Department of medical engineering, Air Force Medical Center, PLA, Air Force Medical University, Beijing, 100142, China.

Weiming Hou

hwm100908@163.com

Department of medical engineering, Air Force Medical Center, PLA, Air Force Medical University, Beijing, 100142, China.

Research Article

Keywords: Malaria, Persistent Organic Pollutants, PCB, DLNM, Kalman filter

Posted Date: January 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8542332/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

This study investigates malaria incidence trends in mainland China from 2005 to 2020, to elucidate its epidemiological characteristics and investigate potential associations with air pollution. First, time series analysis and machine learning methods were employed to predict malaria incidence. Next, the weighted quantile sum (WQS) model and distributed lag nonlinear model (DLNM) were utilized to assess the risk of malaria linked to persistent organic pollutants (POPs). Seasonal Autoregressive Integrated Moving Average Model (SARIMA) model outperforms the Holt-Winters model in univariate traditional models for malaria, with the optimally configured SARIMA (1,1,0) (1,0,1) [12]. For the new Kalman filter model, showing good results across both overall malaria and individual subtypes ($MAE \in [0.001, 0.016]$). In multivariate prediction, the models with the best performance are Gradient Boosting (XGBoost) and Support Vector Machine (SVM). Risk levels for Polychlorinated Biphenyls (PCB) and Hexachlorobenzene (HCB) were coefficients (95% CI): -1.48 (-2.69, -0.27) and -1.39 (-2.57, -0.22), respectively. Cumulative effect of extremely low-level HCB during the first 3 and 4 months were 3.602 (1.103, 11.765) and 4.749 (1.11, 20.31), respectively, indicating an increased risk of malaria incidence. Our current study not only investigated the spatiotemporal surveillance and early warning systems for malaria incidence in mainland China but also elucidated the lagged exposure-response relationships and potential associations between organic pollutants and malaria occurrence. Strengthening POPs emission control activities during this period may help reduce the risk of seasonal malaria susceptibility.

1. Introduction

Malaria is a mosquito-borne infectious parasitic disease prevalent in tropical and subtropical regions(1). In 2021, nearly half of the global population faced the risk of malaria, with an estimated 247 million cases and 619,000 deaths worldwide(2, 3). Malaria incidence shows a seasonal trend, and time series models are the most commonly used cyclical models for disease prediction. Several studies have shown that the Seasonal Autoregressive Integrated Moving Average Model (SARIMA) model performs well in predicting malaria incidence across different countries(4, 5). However, existing research predominantly focuses on model self-comparisons, with limited systematic evaluation of different time series methodologies or the development of novel modeling approaches—representing a significant gap in the field.

Environmental factors play a crucial role in the dynamics and distribution of malaria(6, 7). From a biological perspective, meteorological factors are intrinsically linked to malaria incidence through their influence on mosquito vectors and the development of Plasmodium within the vectors. Studies have shown that under colder environmental conditions, larger daily temperature fluctuations accelerate malaria incidence(8). Additionally, research has confirmed that the increase in malaria incidence occurs after a 30-day interval following the simultaneous occurrence of high temperatures and rainfall(9). Although a Singaporean survey revealed an inverse association between ground-level SO_2 and $PM_{2.5}$ concentrations and mosquito-borne disease incidence - investigating chemical pollutants as risk factors for malaria (a naturally occurring zoonotic disease) - remains epidemiologically valuable(10). This is reflected in the complex network of relationships between various environmental factors and diseases. Moreover, when modeling the impact of environmental factors on malaria cases, two key issues need special attention: non-linear risk effects and lag effects, which are seldom considered together in some studies.

Persistent Organic Pollutants (POPs) have a long history of application in commercial and industrial sectors. Individuals can be exposed to them through daily activities, dietary intake, and drinking water(11). The concentrations of PCBs in the Lanzhou section of the Yellow River, China, fluctuate within the range of 3.08 to 32.3 ng/g, dominated by fully chlorinated congeners, with industrial activities identified as the primary source(12). Multi-regional studies conducted in southeastern and northeastern China have revealed that 66.7% of organochlorine pesticides (OCPs) exhibit a detection frequency exceeding 80%, indicating their ubiquity in the atmosphere. Notably, hexachlorobenzene (HCB) was found to have the highest concentration across all sampling sites(13). However, most studies have focused on direct exposure pathways, such as the relationship between POPs exposure in internal environments (e.g., blood) and diseases, which has shown some correlation between POPs and disease(14). However, these studies have overlooked the long-term exposure risks from external environments, thus neglecting the potential harms of indirect exposure. Although some studies have explored the carcinogenic and non-carcinogenic risks of POPs in the air, they have not concluded that they pose a health threat to humans(15). Additionally, previous animal studies evaluated mice fed with polychlorinated biphenyls (PCBs) for 3 to 6 weeks and subsequently inoculated with Plasmodium berghei (murine malaria parasite). The results demonstrated a 20% reduction in the average survival time of the exposed mice, which preliminarily indicated that environmental chemical pollutants could impair host resistance. It was hypothesized that the assessment of immune parameters might serve as a sensitive indicator for the toxic effects of such pollutants(16). Current evidence fails to establish a clear association between malaria (a parasitic disease) and POPs, making this research particularly timely.

Accordingly, this study employs a systematic approach comprising: (1) a three-dimensional spatial-temporal distribution analysis of malaria incidence cases across China; (2) a comprehensive descriptive analysis of domestic POPs emission sources; and (3) a two-stage predictive framework integrating incidence risk forecasting for malaria strains with subsequent multi-stage effect analysis.

2. Materials and Methods

2.1 Data Source and mapping of malaria cases in different regions

Data on all reported malaria cases in mainland China from 2005 to 2020 were obtained from the Public Health Science Data Center Website (<https://www.phsciencedata.cn/>). POPs (such as Benzo[a]pyrene (BaP), Benzo[b]fluoranthene (BbF), Benzo[k]fluoranthene (BkF), Indeno[1,2,3-cd]pyrene (IcdP), Polychlorinated Biphenyls (PCB), Polychlorinated Dibenzo-p-dioxins (PCDD), and Hexachlorobenzene (HCB)) as well as Greenhouse Gases (GHGs) (including Methane (CH_4), Biogenic Carbon Dioxide (CO_2 bio), Nitrous Oxide (N_2O) and Carbon Dioxide (CO_2)) were sourced from the Emissions Database for

Global Atmospheric Research (EDGAR) (<https://edgar.jrc.ec.europa.eu/>). According to geographical conditions, China is divided into seven regions in the planning: North China (NC), Northeast China (NE), East China (EC), Central China (CC), South China (SC), Southwest China (SW), and Northwest China (NW).

2.2 Different fields and sources of POPs and GHGs

This study summarizes the emissions of POPs and GHGs using the EDGAR database, explores and ranks the main sources of these pollutants and industry sectors, and further analyzes the causes of pollutant generation.

2.3 Features selection

Spearman correlation analysis was performed to examine associations between various POPs, GHGs, and the incidence of different malaria types. Analysis prioritized relationships with organic pollutants. For malaria subtypes exhibiting no significant correlations with any POPs, subsequent analyses focused on significant factors identified in other malaria types, ultimately selecting consistently significant variables across multiple malaria types for in-depth investigation.

2.4 Prediction analysis

In the univariate prediction, this study uses time series analysis models to explore the effect of time variables on the prediction of incidence rates for different types of malaria, thereby accurately monitoring the disease. The time series models selected include Holt-Winters, SARIMA, and Kalman filter models.

SARIMA is a model that combines seasonal differences with the ARIMA model, effectively modeling time series data with cyclical patterns. SARIMA (p, d, q) (P, D, Q) s includes seven parameters, which can be divided into two groups: three non-seasonal parameters (p, d, q) and four seasonal parameters (P, D, Q) s. The Holt-Winters exponential smoothing model eliminates some random fluctuations, assigning different weights to data from each period to predict future trends(17). The Holt-Winters exponential smoothing model and the SARIMA model have been more specifically introduced in previous studies(18). Kalman filtering is an optimal recursive estimation algorithm that estimates the current system state by integrating prior state estimates with current observations, fundamentally relying on state-space equations and Bayesian recursive estimation principles. Therefore, Kalman filtering does not require stationarity or time-varying signals. Specific formulas and algorithms can be found in previous research(19). For the univariate model prediction effectiveness, we compared the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Standard Error (SE) indicators. This model was implemented using the "forecast" and "FKF" packages in R software version 4.4.1.

In the multivariate prediction, this study employs various machine learning algorithms to explore the predictive effects of POPs and GHGs on the incidence of different types of malaria. A 70% dataset was used as the training set for multi-model comparison, while the remaining 30% dataset was used as the validation set to test the optimal algorithm. The machine learning algorithms used include Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Light Gradient Boosting Machine (LightGBM), Linear Support Vector Machine (LSVM), Linear Model (LM), and Neural Network (NNET). Specific algorithms can be found in the relevant literature references(20). Cross-validation was performed using the `vfold_cv()` function, followed by multi-model fitting and comparison after the cross-validation. For the multivariate model prediction effectiveness, we compared the MAE indicator. This model was implemented using the "tidyverse," "kernlab," "nnet," "bonsai," "lightgbm," and "xgboost" packages in R software version 4.4.1.

2.5 Risk analysis

The Weighted Quantile Sum (WQS) regression is used to examine the individual contributions of POP mixtures. The WQS model is widely applied to analyze the effects of multiple chemical exposures on human health(21). In the present study, the model was set to 5000 iterations, with significant organic compounds identified in the preliminary feature extraction step being used as the main effects. Two comparison models were also set up: Model 1 without confounding factors, and Model 2 with significant greenhouse gases as confounding factors. Based on an index that includes all persistent organic pollutants, WQS regression was performed in both positive and negative directions. The weight of each organic compound ranges from 0 to 1, with the sum of the weights equaling 1(22). The WQS coefficient is an indicator of the association strength between the weighted quantile score and the outcome variable, essentially reflecting the "overall effect" of the mixture. The sign of the coefficient corresponds to the direction of the effect, while the absolute value reflects the strength of the effect. The study represents the risk coefficient by calculating the coefficients of individual pollutant factors. This model was implemented using the "gWQS" package in R software version 4.4.1.

2.6 Lagging effect

We used Distributed Lag Nonlinear Model (DLNM) to study the impact of monthly exposure to each organic compound on malaria incidence, while controlling for time variables (month) and other air pollutants, such as greenhouse gases, as confounding factors. We modeled each organic pollutant using 1 to 6 knots and selected the best-fitting model using the Quasi-Akaike Information Criterion (Q-AIC), which was also used to calculate the optimal lag months(23, 24). In this study, the optimal lag period was 6 months(24). Furthermore, we calculated the cumulative effect and relative effect by exploring the relationship between each organic compound at specific levels and specific lag months. For the cumulative effect of organic compounds across all lag months (6 months), specific percentiles were chosen at P5, P25, P75, and P95 levels. For the lag effects at each month and cumulative effects across the months (0–6 months), the specific percentiles were selected at P5 and P95 levels. In addition, sensitivity analysis was performed on this model to determine its parameters. The degrees of freedom for the lag variables were set to $df = 3$ based on previous literature and customary practices. This model was implemented using the "dlnm" package in R software version 4.4.1.

3. Results

3.1 Epidemiological characteristics of malaria

As shown in Table 1, the overall reported malaria cases exhibit a declining trend over the years, with two anomalous peaks of increased incidence occurring in 2006 and 2013. Additionally, there was a short-term increase in reported cases between 2014 and 2016. The seasonal characteristics reveal an epidemic pattern in the summer and autumn (approximately 77.35%), with the majority of cases occurring in the summer (about 41.03%). When comparing the reported cases of different types of malaria, relapsing malaria predominates (about 72.56%), followed by unclassified malaria (about 14.52%), and malignant malaria (about 12.92%). Relapsing malaria and unclassified malaria are mainly prevalent in the summer and autumn, while malignant malaria is more prevalent in the spring and summer.

Table 1
The regional and temporal distribution of different types of malaria in mainland China from 2005 to 2020

Distributions	Year															
Regions	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
CC	3970	6970	6021	4197	2431	1479	633	448	537	515	453	500	368	436	498	153
Henan	2302	5080	4151	3041	1617	887	314	157	193	222	187	205	180	186	232	71
Hubei	1518	1753	1769	1083	712	429	167	127	131	136	124	148	92	128	151	35
Hunan	150	137	101	73	102	163	152	164	213	157	142	147	96	122	115	47
SW	14159	12568	7840	5161	3870	2891	1579	840	723	777	864	744	504	441	407	219
Yunnan	13144	11544	6442	3640	2731	2115	1194	636	426	444	516	331	236	167	150	133
Guizhou	361	466	961	1285	864	397	153	15	13	29	14	36	24	26	28	4
Sichuan	495	430	279	161	185	312	197	158	244	273	292	339	212	219	197	73
Chongqing	66	58	60	43	47	35	32	23	36	24	34	38	31	29	32	9
Xizang	93	70	98	32	43	32	3	8	4	7	8	0	1	0	0	0
NW	47	48	73	68	115	58	69	49	75	136	114	120	96	122	138	40
Shaanxi	33	30	49	47	63	40	35	31	41	58	75	82	61	89	93	30
Gansu	6	11	7	13	43	10	21	7	17	68	28	23	24	19	22	4
Xinjiang	2	6	13	5	8	6	6	7	7	4	3	6	7	7	13	4
Ningxia	5	1	2	3	0	1	3	2	5	4	7	8	4	3	9	2
Qinghai	1	0	2	0	1	1	4	2	5	2	1	1	0	4	1	0
NE	28	38	36	36	55	48	74	63	72	69	102	78	83	86	82	23
Liaoning	19	19	26	17	35	28	46	43	45	39	65	54	53	58	52	16
Jilin	6	14	7	13	11	13	12	13	12	25	22	17	20	12	12	2
Heilongjiang	3	5	3	6	9	7	16	7	15	5	15	7	10	16	18	5
EC	16789	36420	29047	14800	6729	2587	1396	653	977	996	1073	1094	925	853	880	319
Anhui	15693	34984	27307	13484	5918	1860	644	99	194	153	131	146	99	102	102	34
Jiangsu	573	643	754	558	341	362	372	198	344	358	402	306	238	238	238	89
Zhejiang	246	404	594	435	220	121	115	139	187	203	166	208	178	125	135	54
Shandong	124	137	164	165	111	112	118	93	130	153	215	247	197	216	228	71
Fujian	46	52	28	23	65	67	68	54	51	63	77	114	144	109	113	42
Shanghai	85	158	170	111	58	36	38	17	22	22	25	24	34	21	19	14
Jiangxi	22	42	30	24	16	29	41	53	49	44	57	49	35	42	45	15
NC	46	81	70	76	81	93	117	80	137	119	147	153	164	154	164	46
Hebei	19	26	23	29	39	41	50	35	54	55	60	89	84	77	91	28
Beijing	19	38	37	36	28	36	37	30	42	37	46	32	47	45	38	9
Shanxi	4	6	3	5	7	5	9	2	19	13	18	15	20	18	11	6
Tianjin	3	8	6	4	3	9	14	9	13	12	17	11	7	7	15	3
Neimenggu	1	3	1	2	4	2	7	4	9	2	6	6	6	7	9	0
SC	4617	4069	3587	2022	819	233	221	318	1376	309	363	458	539	405	340	223
Hainan	4313	3853	3387	1844	685	73	9	13	15	12	15	8	9	9	2	3
Guangxi	152	95	76	57	51	65	116	218	1255	183	240	304	387	250	183	69
Guangdong	152	121	124	121	83	95	96	87	106	114	108	146	143	146	155	151
Seasons																
Spring	6212	6131	5758	4077	2479	1619	1077	619	927	714	864	878	699	591	522	171

Distributions	Year																
Summer	15533	21182	24230	11727	6902	2955	1413	758	1698	927	921	789	817	654	682	182	
Autumn	15195	30328	14481	9330	3739	1973	890	528	600	629	642	596	543	592	633	198	
Winter	2716	2553	2205	1226	980	842	709	546	672	651	689	884	620	660	672	472	
Total	39656	60194	46674	26360	14100	7389	4089	2451	3897	2921	3116	3147	2679	2497	2509	1023	

In different regions, the EC region reports the highest number of cases (approximately 51.88%), followed by the SW region (approximately 24.06%). Among the different types of malaria, relapsing malaria and unclassified malaria are primarily occurred in the EC region, while malignant malaria is mainly occurred in the SW region. In various provinces, the top five are mainly occurred in the southern regions of China, namely Anhui (45.33%), Yunnan (19.69%), Henan (8.54%), Hainan (6.40%), and Hubei (3.82%). In the CC region, cases are mainly occurred in Henan, in the SW region, they are mainly occurred in Yunnan, in the NW region, they are mainly occurred in Shaanxi, in the NE region, they are mainly occurred in Liaoning, in the EC region, they are mainly occurred in Anhui, in the NC region, they are mainly occurred in Hebei, and in the SC region, they are mainly occurred in Hainan (Table 1).

3.2 Emissions of GHGs and POPs

As shown in **Table S1**, the emission sources of various POPs are presented. In 2018, the total emissions of POPs in China mainly originated from Manufacturing Industries and Construction and Emissions from biomass burning. The sources of emissions for different POPs are generally similar to the overall emissions, with the exception of HCB, PCB, and PCDD_F, which primarily originate from Manufacturing Industries and Construction and Main Activity Electricity and Heat Production. As detailed in **Table S2**, the emission sources of various GHGs are presented. In 2018, the top five sources of GHG emissions in China were Main Activity Electricity and Heat Production, Manufacturing Industries and Construction, Residential and Other Sectors, Road Transportation (no resuspension), and Cement Production. These five sectors jointly accounted for approximately 85.10% of the total GHG emissions. CH₄ primarily originates from Solid Fuels, Rice Cultivations, Enteric Fermentation, Solid Waste Disposal, and Wastewater Treatment and Discharge. N₂O mainly comes from Direct N₂O Emissions from Managed Soils, Main Activity Electricity and Heat Production, Indirect N₂O Emissions from Managed Soils, Indirect N₂O Emissions from the Atmospheric Deposition of Nitrogen in NO_x and NH₃, and Wastewater Treatment and Discharge.

3.3 Correlation among pollutants and malaria

The results of the Spearman correlation analysis in **Figure S1** indicate that the overall reported malaria cases show significant associations with all POPs, as well as with CO₂ among the GHGs. In different malaria types, both relapsing malaria and malignant malaria exhibit significant associations with all POPs. Relapsing malaria is significantly correlated with N₂O and CO₂ from GHGs, while malignant malaria is significantly correlated with CH₄ and CO₂bio from GHGs. Unclassified malaria, however, is only significantly associated with CH₄ and CO₂ from GHGs and does not show significant correlations with any of the POPs.

3.4 Prediction models

3.4.1 Prediction models: univariate time analysis

Table 2 reflects the comparison of the predictive model performance for the overall and different types of malaria incidence during 2019–2020. In comparison with traditional time series models, the SARIMA model consistently outperforms the Holt-Winters model. The optimized model for the overall malaria incidence is SARIMA (1,1,0) (1,0,1) [12]. In contrast, the novel Kalman filter model demonstrates superior prediction performance compared to the traditional SARIMA model, showing improvements in both the overall and individual malaria types (MAE∈ [0.001, 0.016]).

Table 2

Comparison of the three models in fitting performances and parameter estimations among different malaria.

Model	Series	Best parameters	RMSE	MAE	MAPE
Holt-Winters Additive Model	Total malaria	$\alpha = 0.26, \beta = 0.03, \gamma = 0.98$	0.088	0.051	99.874
SARIMA Multiple Model		(1,1,0) (1,0,1) [12]	0.059	0.026	30.258
Kalman Model		$P0 = 1e6$	0.036	0.016	18.740
Holt-Winters Additive Model	Relapsing malaria	$\alpha = 0.26, \beta = 0.03, \gamma = 0.97$	0.075	0.043	173.480
SARIMA Multiple Model		(3, 1, 2) (0,1,1) [12]	0.043	0.020	92.841
Kalman Model		$P0 = 1e6$	0.031	0.013	25.122
Holt-Winters Additive Model	Malignant malaria	$\alpha = 0.38, \beta = 0.01, \gamma = 0.62$	0.005	0.003	39.215
SARIMA Multiple Model		(1, 1, 4) (1,0,1) [12]	0.004	0.002	23.004
Kalman Model		$P0 = 1e6$	0.002	0.001	Inf
Holt-Winters Additive Model	Unclassified malaria	$\alpha = 0.35, \beta = 0.03, \gamma = 1.00$	0.013	0.008	236.830
SARIMA Multiple Model		(2, 1, 1) (2,0,0) [12]	0.008	0.004	54.228
Kalman Model		$P0 = 1e6$	0.005	0.002	19.800

From Fig. 1, monthly prediction charts indicate that during periods of unusually high peak fluctuations, the Holt-Winters predicted values are closer to the actual values compared to the SARIMA model (Holt-Winters outperforms SARIMA in certain periods). However, during stable periods, the SARIMA model exhibits markedly reduced prediction fluctuations, thereby demonstrating superior forecasting performance. In the Kalman model prediction charts, the actual values and predicted values show significantly smaller fluctuations than in the other models (with the best prediction performance), and the effect is more evident during stable periods. As shown in **Table S3**, for overall predictions, the standard error ($ISE=0.1597$) of the SARIMA model is larger than that of the Holt-Winters model ($ISE=0.0626$), while the Kalman model has the smallest standard error ($ISE=0.0061$). As shown in **Table S4**, for Relapsing malaria predictions, the standard error ($ISE=0.2721$) of the SARIMA model is greater than that of the Holt-Winters model ($ISE=0.0971$), with the Kalman model having the smallest standard error ($ISE=0.0022$). As shown in **Table S5**, for Malignant malaria predictions, the standard error of the SARIMA model is similar to that of the Holt-Winters model, but the Kalman model has the smallest standard error ($ISE=0.0043$). As shown in **Table S6**, for Unclassified malaria predictions, the standard error ($ISE=0.0404$) of the Holt-Winters model is greater than that of the SARIMA model ($ISE=0.0345$), while the Kalman model has the smallest standard error ($ISE=0.0004$). For the 2019–2020 predicted values, only the Unclassified malaria model shows negative values based on the predictions from both the Holt-Winters and SARIMA models.

3.4.2 Prediction models: multivariate pollutants analysis

Figure 2 reflects the evaluation of the predictive performance of POPs and GHGs for overall and different types of malaria. The evaluation is represented by the MAE (Mean Absolute Error) metric, where a smaller MAE indicates better model prediction performance. The models that perform best overall are primarily XGBOOST and SVM (Fig. 2A). The MAE and R^2 of XGBoost on the validation set were 0.0599 and 0.616, respectively. For the Relapsing malaria type, the models that perform best are mainly SVM and XGBOOST (Fig. 2B). For SVM, the MAE and R^2 on the validation set were 0.0432 and 0.679, respectively. For the Malignant malaria type, the models that perform best are primarily SVM and LSVM (Fig. 2C). For SVM, the MAE and R^2 on the validation set were 0.0040 and 0.059, respectively. For the Unclassified malaria type, the models that perform best are mainly SVM and XGBOOST (Fig. 2D). For SVM in a different comparison, the MAE and R^2 on the validation set were 0.0085 and 0.767, respectively.

3.5 Single POPs exposures and malaria

For overall malaria, both PCB and HCB factors show significant risk levels for reported cases, regardless of whether greenhouse gases were adjusted for as confounders. Specifically, in the adjusted model, the risk levels for PCB and HCB are coefficients (95%CI): -1.48 (-2.69, -0.27) and -1.39 (-2.57, -0.22), respectively. In the Relapsing malaria type, when confounders were not adjusted, only PCB and HCB showed significant risk levels. After adjusting for confounders, seven organic compounds, including BaP, BbF, BkF, IcdP, PCB, PCDD, and HCB, showed significant risk levels, with the risk levels for PCB and HCB being coefficients (95%CI): -1.13 (-1.77, -0.48) and -1.05 (-1.65, -0.44), respectively. In the Malignant malaria type, when confounders were not adjusted, none of the organic compounds showed significant levels. However, after adjusting for confounders, five organic compounds, including BaP, BbF, BkF, IcdP, and PCDD, showed significant risk levels, with the risk levels for all these compounds ranging from -0.10. In the Unclassified malaria type, when confounders were not adjusted, only PCB and HCB showed significant risk levels. After adjusting for confounders, seven organic compounds, including BaP, BbF, BkF, IcdP, PCB, PCDD, and HCB, showed significant risk levels, with the risk levels for PCB and HCB being coefficients (95%CI): -0.24 (-0.44, -0.05) and -0.26 (-0.45, -0.06), respectively. The remaining five organic compounds (BaP, BbF, BkF, IcdP, PCDD) all showed positive risk levels in the adjusted model, with the risk levels ranging from coefficients: 0.10 to 0.14 (Table 3).

Table 3
Multivariate WQS regression analysis of risk factors of malaria diseases among different groups.

Stratification	Predictor Variables	Model 1		Model 2	
		Exp(β)(95%CI)	P	Exp(β)(95%CI)	P
Total malaria	PCB	-1.71 (-2.93, -0.49)	0.0072 **	-1.48 (-2.69, -0.27)	0.0186 *
	HCB	-1.69 (-2.91, -0.47)	0.0077 **	-1.39 (-2.57, -0.22)	0.0217 *
Relapsing malaria	BaP	-0.58 (-1.82, 0.67)	0.3680	-0.54 (-0.93, -0.14)	0.0088 **
	BbF	-0.80 (-2.12, 0.52)	0.2380	-0.56 (-0.95, -0.16)	0.0067 **
	BkF	-0.80 (-2.12, 0.51)	0.2350	-0.56 (-0.95, -0.16)	0.0066 **
	IcdP	-0.78 (-2.12, 0.56)	0.2580	-0.56 (-0.96, -0.16)	0.0073 **
	PCB	-1.43 (-2.44, -0.43)	0.0064 **	-1.13 (-1.77, -0.48)	0.0010 ***
	PCDD	-0.80 (-2.13, 0.53)	0.2420	-0.56 (-0.96, -0.16)	0.0071 **
	HCB	-1.44 (-2.45, -0.43)	0.0063 **	-1.05 (-1.65, -0.44)	0.0010 ***
Malignant malaria	BaP	-0.07 (-0.17, 0.03)	0.1486	-0.10 (-0.17, -0.02)	0.0108 *
	BbF	-0.09 (-0.20, 0.02)	0.1106	-0.10 (-0.17, -0.03)	0.0090 **
	BkF	-0.09 (-0.20, 0.02)	0.1051	-0.10 (-0.17, -0.03)	0.0090 **
	IcdP	-0.09 (-0.19, 0.02)	0.1106	-0.10 (-0.18, -0.03)	0.0087 **
	PCDD	-0.09 (-0.19, 0.02)	0.1165	-0.10 (-0.18, -0.03)	0.0095 **
Unclassified malaria	BaP	-0.10 (-0.34, 0.14)	0.4080	0.10 (0.03, 0.16)	0.0073 **
	BbF	-0.14 (-0.39, 0.11)	0.2640	0.14 (0.06, 0.21)	0.0008 ***
	BkF	-0.14 (-0.39, 0.11)	0.2620	0.14 (0.06, 0.22)	0.0008 ***
	IcdP	-0.14 (-0.39, 0.11)	0.2790	0.13 (0.05, 0.21)	0.0011 **
	PCB	-0.27 (-0.47, -0.08)	0.0070 **	-0.24 (-0.44, -0.05)	0.0159 *
	PCDD	-0.15 (-0.40, 0.11)	0.2620	0.13 (0.05, 0.20)	0.0012 **
	HCB	-0.28 (-0.48, -0.09)	0.0054 **	-0.26 (-0.45, -0.06)	0.0101 *
Model 1: only single direction (positive or negative) POPs indicators.					
Model 2: combined all GHGs indicators for univariate analyzes.					
Bold font indicates statistical significance at the 0.05 level.					
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05					

3.6 DLNM models

3.6.1 DLNM models: overall lagging effect

As shown in **Figure S2**, reported malaria cases in China demonstrate a nonlinear lagged association with the organic compounds PCB and HCB, where different lag months correspond to distinct epidemiological effects. Specifically, the lag effect of high-level PCB occurs rapidly but lasts for a short duration (**Figure S2 A, B**), whereas the lag effect of low-level HCB occurs slowly (risk emerging after approximately a 1-month lag) but persists for a longer duration (**Figure S2 C, D**).

3.6.2 DLNM models: exposure-reaction relationship

Figure 3 shows the overall impact of two major persistent organic compounds (PCB and HCB) on the total number of cases for different types of malaria over a 6-month period. Overall, the relative risk of malaria increases with higher PCB levels (above 520T) and decreases with higher HCB levels. For relapsing malaria and unclassified malaria types, the relative risk of malaria increases with higher PCB levels (above 480T) and decreases with higher HCB levels. For malignant malaria, the relative risk of malaria increases with higher HCB levels (above 400T) and decreases with higher PCB levels.

3.6.3 DLNM models: extreme lagging cumulative effect

Table 4 reflects the cumulative effects of the organic compounds PCB and HCB over different months (0–6 months). In overall malaria patients, the cumulative lag effect of extremely high levels of PCB increased from 7.977 to 36,726.46 over the 0–6 months period. In contrast, extremely low levels of HCB had cumulative effect values of 3.602 (1.103, 11.765) in the first 3 months and 4.749 (1.11, 20.31) in the first 4 months, both of which increase the risk of malaria in the population. Extremely high levels of HCB showed a protective effect during the first 2 months. For relapsing malaria patients, the cumulative lag effect of extremely high levels of PCB increased from 14.058 to 638,534.2 over the 0–6 months period. Extremely low levels of HCB had cumulative

effect values of 3.164 (1.157, 8.651) in the first 2 months and 9.462 (1.087, 82.341) in the first 6 months, both increasing the risk of malaria in the population. Extremely high levels of HCB had a protective effect in the first month. For malignant malaria, there were no significant cumulative effect values, but both extremely low levels of PCB and extremely high levels of HCB may increase the risk of malaria. In unclassified malaria patients, the cumulative lag effect of extremely high levels of PCB increased from 2.83 to 124,535.9 over the 0–6 months period. Extremely low levels of HCB had cumulative effect values of 2.944 (1.044, 8.3) in the first 3 months and 5.584 (1.113, 28.021) in the first 6 months, both increasing the risk of malaria in the population.

Table 4
Cumulative effects of organic compounds PCB and HCB in each month (0–6 months).

Series	Cumulative risk (95% CI)	Variables			
		low PCB effect	High PCB effect	low HCB effect	High HCB effect
Total malaria	Lag0	1.034 (0.789, 1.355)	7.977 (2.172, 29.29)	1.03 (0.722, 1.469)	0.27 (0.09, 0.813)
	Lag0-1	0.849 (0.481, 1.501)	76.349 (6.97, 836.301)	1.558 (0.851, 2.852)	0.057 (0.009, 0.383)
	Lag0-2	0.526 (0.221, 1.251)	949.048 (33.266, 27075.09)	2.4 (0.986, 5.838)	0.01 (0.001, 0.156)
	Lag0-3	0.335 (0.103, 1.093)	4687.511 (59.827, 367269.4)	3.602 (1.103, 11.765)	0.003 (0, 0.095)
	Lag0-4	0.254 (0.057, 1.135)	11706.39 (56.906, 2408162)	4.749 (1.11, 20.31)	0.001 (0, 0.083)
	Lag0-5	0.249 (0.042, 1.489)	21793.87 (46.533, 10207277)	5.048 (0.949, 26.859)	0.001 (0, 0.087)
	Lag0-6	0.328 (0.042, 2.546)	36726.46 (41.93, 32168849)	4.148 (0.667, 25.79)	0.001 (0, 0.1)
Relapsing malaria	Lag0	0.947 (0.73, 1.227)	14.058 (3.687, 53.595)	1.079 (0.737, 1.58)	0.151 (0.042, 0.549)
	Lag0-1	0.613 (0.342, 1.1)	138.465 (9.431, 2032.918)	1.772 (0.903, 3.478)	0.051 (0.004, 0.682)
	Lag0-2	0.294 (0.116, 0.746)	2056.348 (47.445, 89125.3)	3.164 (1.157, 8.651)	0.01 (0, 0.63)
	Lag0-3	0.152 (0.041, 0.558)	20062.42 (149.797, 2686977)	5.293 (1.357, 20.646)	0.004 (0, 1.156)
	Lag0-4	0.095 (0.018, 0.502)	111089.8 (273.58, 45109043)	7.725 (1.417, 42.1)	0.003 (0, 3.075)
	Lag0-5	0.075 (0.01, 0.554)	351631.7 (342.987, 360494091)	9.44 (1.318, 67.634)	0.003 (0, 7.486)
	Lag0-6	0.078 (0.008, 0.759)	638534.2 (340.063, 1198971679)	9.462 (1.087, 82.341)	0.003 (0, 12.814)
Malignant malaria	Lag0	1.687 (0.595, 4.782)	0.13 (0.001, 15.822)	0.954 (0.356, 2.556)	6.317 (0.026, 1514.783)
	Lag0-1	8.776 (0.271, 284.26)	0 (0, 768.387)	0.702 (0.029, 17.126)	20128.19 (0, 1.14E + 12)
	Lag0-2	124.882 (0.116, 134936.3)	0 (0, 15885.48)	0.243 (0, 212.675)	11133181828 (0, 5.49E + 25)
	Lag0-3	4202.652 (0.091, 193591131)	0 (0, 1218.606)	0.066 (0, 4357.513)	9.48E + 17 (0, 3.84E + 42)
	Lag0-4	115946.8 (0.087, 1.5422E + 11)	0 (0, 11.954)	0.02 (0, 76686.97)	4.54E + 25 (0, 1.87E + 58)
	Lag0-5	1078812 (0.084, 1.39E + 13)	0 (0, 0.321)	0.009 (0, 594626.3)	9.99E + 30 (0, 1.29E + 69)
	Lag0-6	2171216 (0.079, 5.97E + 13)	0 (0, 0.091)	0.006 (0, 1418617)	9.15E + 32 (0, 1.57E + 73)
Unclassified malaria	Lag0	0.809 (0.653, 1.003)	2.83 (1.152, 6.951)	1.143 (0.867, 1.508)	0.457 (0.198, 1.055)
	Lag0-1	0.462 (0.288, 0.74)	15.49 (2.497, 96.076)	1.365 (0.82, 2.271)	0.447 (0.094, 2.119)
	Lag0-2	0.174 (0.081, 0.374)	323.513 (23.124, 4525.97)	1.88 (0.87, 4.064)	0.249 (0.024, 2.596)
	Lag0-3	0.078 (0.026, 0.232)	3907.952 (121.723, 125465.6)	2.944 (1.044, 8.3)	0.104 (0.005, 2.199)
	Lag0-4	0.045 (0.011, 0.182)	22471.07 (329.054, 1534548)	4.442 (1.24, 15.916)	0.046 (0.001, 1.772)
	Lag0-5	0.031 (0.006, 0.168)	69762.35 (552.311, 8811678)	5.627 (1.292, 24.5)	0.028 (0, 1.721)
	Lag0-6	0.026 (0.004, 0.176)	124535.9 (629.074, 24654021)	5.584 (1.113, 28.021)	0.027 (0, 2.234)

Series	Cumulative risk (95% CI)	Variables			
		low PCB effect	High PCB effect	low HCB effect	High HCB effect
Bold font indicates statistical significance at the 0.05 level.					

Table S7 reflects the cumulative effects of the organic compounds PCB and HCB across all lag months (6 months). For both overall and relapsing malaria cases, the cumulative lag effects of extremely high PCB levels (P95) over a 0–6 month period were $3.673e + 4$ (41.930, $3.217e + 7$) and $6.385e + 5$ (340.063, $1.199e + 9$), respectively, indicating a significant elevation in malaria risk. Additionally, for low levels of HCB (P75-P95), the cumulative lag effects over 0–6 months were 9.462 (1.087, 82.341) and 2.889 (1.005, 8.310), respectively, both of which increase the risk of developing malaria in relapsing malaria patients. For malignant malaria, high levels of HCB (P75) corresponded to a cumulative lag effect over 0–6 months of $3.307e + 46$ (11.764, $9.296e + 91$), which increases the risk of developing malaria in patients. In unclassified malaria, both high levels of PCB (P75-P95) and low levels of HCB (P5-P25) corresponded to cumulative lag effects greater than 1, which increases the risk of developing malaria in patients.

3.6.4 DLNM models: extreme lagging relative effect

Figure 4 reflects the current-month lag effects of the organic compounds PCB and HCB across different months. In both overall malaria and Relapsing malaria types, the lag effect of low-level PCB in the 6th month showed a significant positive risk level. High-level PCB and low-level HCB exhibited significant risk levels in the 1st to 4th months, with the lag effect in the 2nd month showing the highest peak. In malignant malaria, there were no significant single-month lag effects. In unclassified malaria, high-level PCB and low-level HCB showed significant risk levels in the 2nd to 5th months, with the highest peak lag effect observed in the 2nd and 3rd months, respectively.

3.6.5 DLNM models: sensitivity analysis

Based on the QAIC (Quasi-Akaike Information Criterion) selection criterion, the optimal maximum lag period was determined to be 6 months. According to the minimum partial autocorrelation function (PACF) criterion, the derived parameters exhibited variation across different malaria types. In the overall patient population, the degrees of freedom (df) for the time variable were 2, for PCB was 6, and for HCB was 4. In relapsing malaria patients, the degrees of freedom for the time variable were 2, for PCB was 6, and for HCB was 5. In malignant malaria patients, the degrees of freedom for the time variable were 5, for PCB was 6, and for HCB was 7. In unclassified malaria patients, the degrees of freedom for the time variable were 2, for PCB was 4, and for HCB was 4.

4. Discussion

Malaria in China predominantly occurs during the summer and autumn seasons, driven by warm, humid conditions that favor mosquito breeding and the transmission of Plasmodium pathogens(25). This seasonal pattern aligns with findings from a Yunnan-based study(26). Geographically, malaria in China is concentrated in the southwestern and eastern regions, suggesting that climatic factors—particularly in the southwest (e.g., Yunnan Province), where tropical and subtropical conditions prevail—play a critical role. High temperatures and humidity accelerate mosquito reproduction, while the rainy season creates stagnant water bodies (e.g., ponds) that serve as ideal breeding sites, thereby increasing malaria transmission risk. Regarding malaria types, Plasmodium vivax (tertian malaria) is the predominant form in China, owing to the widespread distribution of Anopheles mosquitoes, particularly Anopheles sinensis, a highly competent vector for P. vivax(27). Currently, the primary source of malignant malaria is imported malaria.

Studies have found that during China's industrialization process, industries such as coal combustion, oil refining, and steel production have emerged. In these sectors, coal, as one of the main energy sources, releases large amounts of polycyclic aromatic hydrocarbons during combustion(28), especially in regions using incomplete combustion equipment and technologies, where emissions are more severe. This aligns with the findings of the study, which identifies manufacturing and construction industries as the primary sources of China's POPs, followed by biomass burning emissions. The study also found that PCB emissions primarily stem from electricity and thermal power generation, consistent with the findings of a study conducted along the Indus River(29).

In this study, SARIMA seasonal model is more effective in detecting subtle differences during stable periods, thereby achieving stable predictions. This finding is consistent with previous research results(18). The study also applied the Kalman filter model showed superior performance. This is in agreement with the results of Sangil Kim and others in predicting Hand, Foot, and Mouth Disease(30). In the multivariate predictive analysis, the SVM and XGBoost models were found to be more suitable for predicting the impact of POPs on malaria incidence in China, which is supported by similar findings in several studies(24, 31).

This study revealed that the risk of malaria incidence associated with POPs is substantially amplified under the mediating influence of GHGs. Additionally, an experimental study demonstrating elevated microbial respiration (manifested as increased CO₂ production) observed a corresponding rise in the release of organic pollutants into aquatic systems(32). This might be due to certain greenhouse gases (such as CO₂) facilitating the release of POPs, thereby enhancing the risk of malaria. In detailed studies, we observed that PAH congeners and PCDD exhibited positive risk effects on the incidence of different malaria types. The most common pollutants in urban environments include heavy metals, PAHs, and others. These pollutants persist in the environment and can accumulate in air, water, and soil, leading to complex interactions affecting urban ecosystems(33). This finding is consistent with a study conducted in Africa, which revealed that polycyclic aromatic hydrocarbons (PAHs) can drive insecticide resistance in malaria vectors, thereby indirectly elevating the potential risk of malaria transmission and incidence(34). Therefore, aquatic and terrestrial organisms in these ecosystems constantly absorb these substances, which may result in bioaccumulation and impact their behavior and reproduction(35, 36). Due to the key role of mosquitoes in decomposition and the food web, as well as their utility as environmental pollution sentinels, they accumulate organic pollutants in their tissues. These pollutants, such as PAHs, usually affect their adult stages and may influence mosquito reproduction(37, 38). A U.S. study found that direct exposure to dioxin-like organic compounds could progressively increase the susceptibility of separated human red blood cells to Plasmodium falciparum after 48 hours(39). This finding is similar to our observation that organic pollutants like PCDD can increase the risk of malaria. The lag analysis showed that long-term PCB exposure was

positively associated with increased malaria risk, as previously discussed. Several prior studies have confirmed that compounds like DDT can control malaria transmission by killing pathogen larvae(40). This is similar to the finding in this study that HCB shows a protective effect on malaria incidence under long-term lag (with a dose-response relationship showing negative feedback).

While prior studies have primarily examined how organic pollutants (e.g., DDT) may mitigate malaria transmission, the role of POPs in enhancing transmission has been overlooked. This study addresses this gap by analyzing GHGs and their mediating influence on pollutant impacts. For predictive modeling, we improve upon traditional approaches by integrating the Kalman filter, enabling high-precision forecasts without reliance on fluctuation amplitude. Furthermore, we broke through traditional meteorological and pollution model analysis to explore the lag effects of environmental factors, analyzing the long-term lag effects of organic pollutants on malaria risk. However, due to data limitations, further research is needed on other malaria subtypes. There is some collinearity among various POPs factors; however, these factors were included in the analysis due to their statistical significance, in order to draw reasonable conclusions. This issue of collinearity may be addressed in future environmental health studies using the Qgcomp analysis method. POP emission data do not directly represent human exposure. Ecological exposure estimates may introduce measurement error and ecological bias. Regarding etiological risks, there is a lack of data on pathogens such as Plasmodium, which highlights the need for future scholars to collect pathogen-level data and establish animal and statistical models to investigate the etiological and transmission mechanisms, in order to determine causal relationships.

5. Conclusion

We analyzed malaria surveillance and early warning in mainland China, along with lagged effects of organic pollutants on malaria incidence. Extreme levels of organic pollutants (PCB, HCB) are associated with increased malaria transmission risk, with varying transmission peaks. Strengthening POPs emission control activities during these periods may help reduce the risk of seasonal malaria susceptibility.

Declarations

Ethics approval and consent to participate

Malaria is classified as a Class B infectious disease under China's Infectious Disease Prevention and Control Law. This study was approved by the Ethics Committee of the Chinese Center for Disease Control and Prevention (CDC). For confidentiality purposes, all data were analyzed anonymously. Due to the mandatory monitoring required by China's Infectious Disease Prevention and Control Law and the public database, patient consent is not required. All data usage strictly complies with the "Data Security Law of the People's Republic of China" and the "Law of the People's Republic of China on the Prevention and Treatment of Infectious Diseases." Official permission has been obtained from the China CDC. The authorization document clearly specifies that the data usage is limited to aggregated incidence rate statistics and does not involve individual patient privacy information, fully complying with public health data security management regulations.

Consent for publication

All authors have consented to publication of this research.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The disease data was publicly supported from the National Public Health Data Centre of China (<https://www.phsciencedata.cn/>). Pollutants information was publicly from the Emissions Database for Global Atmospheric Research (EDGAR) (<https://edgar.jrc.ec.europa.eu/>).

Availability of data and material

The data that support the findings of this study are available on request from the National Public Health Data Centre of China (<https://www.phsciencedata.cn/>) and the the Emissions Database for Global Atmospheric Research (EDGAR) (<https://edgar.jrc.ec.europa.eu/>).

CRedit authorship contribution statement

Guolong Qu: Software, Conceptual, Methodology, Formal analysis, Investigation, Resources, Writing-original draft, Writing-review & editing. **Jianqiang Han:** Software, Conceptual, Methodology, Formal analysis, Investigation, Writing-original draft, Writing-review & editing. **Zhenyao Song:** Conceptualization, Methodology, Formal analysis, Writing-review & editing, Supervision. **Weiming Hou:** Conceptualization, Methodology, Formal analysis, Writing-review & editing, Funding acquisition, Supervision. All authors had full access to the data, contributed to the study, approved the final version for publication, and take responsibility for its accuracy and integrity. All authors read and approved the final manuscript.

Funding

None.

References

1. Tao ZY, Fang Q, Liu X, Culleton R, Tao L, Xia H, Gao Q. Congenital malaria in China. *PLoS Negl Trop Dis*. 2014 Mar;8(3):e2622. eng. Epub 2014/03/15. doi:10.1371/journal.pntd.0002622. Cited in: Pubmed; PMID 24626148.
2. Boene S, Rovira-Vallbona E, da Silva C, García-Ulloa M, Rafael B, Canana N, Aranda-Díaz A, Cisteró P, García-Fernández C, Tembisse D, Ndimande N, Chidimatembue A, Matambisso G, Palmer B, Chico AR, Dimene M, Saifodine A, Inácio J, da Silva M, Plucinski M, Bonnington C, Wate F, de Carvalho E, Mathe G, Pujol A, Arregui-Gallego B, Comiche K, Nhama A, Nhamussua L, Aide P, Saute F, Enosse S, Greenhouse B, Candrinho B, Mayor A. Antimalarial drug resistance and population structure of *Plasmodium falciparum* in Mozambique using genomic surveillance at health facilities in 2021 and 2022. *Sci Rep*. 2025 Aug 11;15(1):29335. eng. Epub 2025/08/12. doi:10.1038/s41598-025-02166-w. Cited in: Pubmed; PMID 40790052.
3. Zhu Y, Restrepo AC, Wang HB, Mills DJ, Liang RR, Liu ZB, Lau CL, Furuya-Kanamori L. Malaria cases in China acquired through international travel, 2013-2022. *J Travel Med*. 2024 Dec 10;31(8). eng. Epub 2024/04/09. doi:10.1093/jtm/taae056. Cited in: Pubmed; PMID 38591791.
4. Gondwe T, Yang Y, Yosefe S, Kasanga M, Mulula G, Luwemba MP, Jere A, Daka V, Mudenda T. Epidemiological Trends of Malaria in Five Years and under Children of Nsanje District in Malawi, 2015-2019. *Int J Environ Res Public Health*. 2021 Dec 3;18(23). eng. Epub 2021/12/11. doi:10.3390/ijerph182312784. Cited in: Pubmed; PMID 34886507.
5. Ebhuoma O, Gebreslasie M, Magubane L. A Seasonal Autoregressive Integrated Moving Average (SARIMA) forecasting model to predict monthly malaria cases in KwaZulu-Natal, South Africa. *S Afr Med J*. 2018 Jun 26;108(7):573-578. eng. Epub 2018/07/14. doi:10.7196/SAMJ.2018.v108i7.12885. Cited in: Pubmed; PMID 30004345.
6. Detinova TS. Age-grouping methods in Diptera of medical importance with special reference to some vectors of malaria. *Monogr Ser World Health Organ*. 1962;47:13-191. eng. Epub 1962/01/01. Cited in: Pubmed; PMID 13885800.
7. Zhou G, Minakawa N, Githeko AK, Yan G. Association between climate variability and malaria epidemics in the East African highlands. *Proc Natl Acad Sci U S A*. 2004 Feb 24;101(8):2375-80. eng. Epub 2004/02/26. doi:10.1073/pnas.0308714100. Cited in: Pubmed; PMID 14983017.
8. Zhao X, Chen F, Feng Z, Li X, Zhou XH. Characterizing the effect of temperature fluctuation on the incidence of malaria: an epidemiological study in south-west China using the varying coefficient distributed lag non-linear model. *Malar J*. 2014 May 27;13:192. eng. Epub 2014/06/03. doi:10.1186/1475-2875-13-192. Cited in: Pubmed; PMID 24886630.
9. Kapwata T, Wright CY, du Preez DJ, Kunene Z, Mathee A, Ikeda T, Landman W, Maharaj R, Sweijd N, Minakawa N, Blesic S. Exploring rural hospital admissions for diarrhoeal disease, malaria, pneumonia, and asthma in relation to temperature, rainfall and air pollution using wavelet transform analysis. *Sci Total Environ*. 2021 Oct 15;791:148307. eng. Epub 2021/06/18. doi:10.1016/j.scitotenv.2021.148307. Cited in: Pubmed; PMID 34139502.
10. Tewari P, Ma P, Gan G, Janhavi A, Choo ELW, Koo JR, Dickens BL, Lim JT. Non-linear associations between meteorological factors, ambient air pollutants and major mosquito-borne diseases in Thailand. *PLoS Negl Trop Dis*. 2023 Dec;17(12):e0011763. eng. Epub 2023/12/27. doi:10.1371/journal.pntd.0011763. Cited in: Pubmed; PMID 38150471.
11. Dou Q, Bai Y, Li Y, Zheng S, Wang M, Wang Z, Sun J, Zhang D, Yin C, Ma L, Lu Y, Zhang L, Chen R, Cheng Z. Perfluoroalkyl substances exposure and the risk of breast cancer: A nested case-control study in Jinchang Cohort. *Environ Res*. 2024 Dec 1;262(Pt 2):119909. eng. Epub 2024/09/03. doi:10.1016/j.envres.2024.119909. Cited in: Pubmed; PMID 39222733.
12. Yuan L, Wu Y, Shi L, Song J, Jiang Y. Organochlorine pesticides and polychlorinated biphenyls in sediments of the Lanzhou reach of Yellow River (China): Spatial distribution, sources and risk assessment. *Mar Pollut Bull*. 2024 Nov;208:116962. eng. Epub 2024/09/18. doi:10.1016/j.marpolbul.2024.116962. Cited in: Pubmed; PMID 39288672.
13. Wang L, Cao G, Liu LY, Zhang ZF, Jia SM, Fu MQ, Ma WL. Cross-regional scale studies of organochlorine pesticides in air in China: Pollution characteristic, seasonal variation, and gas/particle partitioning. *Sci Total Environ*. 2023 Dec 15;904:166709. eng. Epub 2023/09/03. doi:10.1016/j.scitotenv.2023.166709. Cited in: Pubmed; PMID 37659555.
14. Rouxel E, Costet N, Monfort C, Audouze K, Cirugeda L, Gaudreau E, Grimalt JO, Ibarluzea J, Lainé F, Llop S, Lopez-Espinosa MJ, Rouget F, Santa-Marina L, Vrijheid M, Chevrier C, Casas M, Warembourg C. Prenatal exposure to multiple persistent organic pollutants in association with adiposity markers and blood pressure in preadolescents. *Environ Int*. 2023 Aug;178:108056. eng. Epub 2023/06/29. doi:10.1016/j.envint.2023.108056. Cited in: Pubmed; PMID 37379720.
15. Hu Z, Li J, Li B, Zhang Z. Annual changes in concentrations and health risks of PCDD/Fs, DL-PCBs and organochlorine pesticides in ambient air based on the Global Monitoring Plan in São Paulo. *Environ Pollut*. 2019 Dec;255(Pt 2):113310. eng. Epub 2019/10/11. doi:10.1016/j.envpol.2019.113310. Cited in: Pubmed; PMID 31600699.
16. Loose LD, Pittman KA, Benitz KF, Silkworth JB, Mueller W, Coulston F. Environmental chemical-induced immune dysfunction. *Ecotoxicol Environ Saf*. 1978 Sep;2(2):173-98. eng. Epub 1978/09/01. doi:10.1016/0147-6513(78)90008-8. Cited in: Pubmed; PMID 103706.
17. Zhou Y, Luo D, Liu K, Chen B, Chen S, Pan J, Liu Z, Jiang J. Trend of the Tuberculous Pleurisy Notification Rate in Eastern China During 2017-2021: Spatiotemporal Analysis. *JMIR Public Health Surveill*. 2023 Oct 30;9:e49859. eng. Epub 2023/10/30. doi:10.2196/49859. Cited in: Pubmed; PMID 37902822.
18. Hou W, Song Z. Exploring the risk and predictive study of outdoor air pollutants on the incidence and mortality of HIV/AIDS. *Ecotoxicol Environ Saf*. 2024 Nov 15;287:117292. eng. Epub 2024/11/15. doi:10.1016/j.ecoenv.2024.117292. Cited in: Pubmed; PMID 39541701.
19. Zhang S, Yang L. A hybrid data assimilation method based on real-time Ensemble Kalman filtering and KNN for COVID-19 prediction. *Sci Rep*. 2025 Jan 19;15(1):2454. eng. Epub 2025/01/20 22:53. doi:10.1038/s41598-025-85593-z. Cited in: Pubmed; PMID 39828742.
20. Zhang H, Zhang H, Yang H, Shuid AN, Sandai D, Chen X. Machine learning-based integrated identification of predictive combined diagnostic biomarkers for endometriosis. *Front Genet*. 2023;14:1290036. eng. Epub 2023/12/15. doi:10.3389/fgene.2023.1290036. Cited in: Pubmed; PMID 38098472.
21. Cao H, Han J, Hou W, Yuan J. Associations of greenhouse gases, air pollutants and dynamics of scrub typhus incidence in China: a nationwide time-series study. *BMC Public Health*. 2025 May 29;25(1):1977. eng. Epub 2025/05/30. doi:10.1186/s12889-025-23156-7. Cited in: Pubmed; PMID 40442614.

22. Lin L, Lin X, Chang S, Xing Y, Zhou T, Yang C. Age-specific associations between per- and polyfluoroalkyl substances exposure and metabolic syndrome: a cross-sectional study. *Lipids Health Dis.* 2025 May 10;24(1):174. eng. Epub 2025/05/11. doi:10.1186/s12944-025-02582-x. Cited in: Pubmed; PMID 40349075.
23. Murphy MS, Abdulaziz KE, Lavigne É, Erwin E, Guo Y, Dingwall-Harvey AL, Stieb D, Walker MC, Wen SW, Shin HH. Association between prenatal air pollutant exposure and autism spectrum disorders in young children: A matched case-control study in Canada. *Environ Res.* 2024 Nov 15;261:119706. eng. Epub 2024/08/01. doi:10.1016/j.envres.2024.119706. Cited in: Pubmed; PMID 39084506.
24. Hou W. Effect and prediction of long-term weather and pollutant exposure on hemorrhagic fever with renal syndrome: based on statistical models. *Front Public Health.* 2025;13:1393763. eng. Epub 2025/02/17. doi:10.3389/fpubh.2025.1393763. Cited in: Pubmed; PMID 39957981.
25. Wardrop NA, Barnett AG, Atkinson JA, Clements AC. Plasmodium vivax malaria incidence over time and its association with temperature and rainfall in four counties of Yunnan Province, China. *Malar J.* 2013 Dec 18;12:452. eng. Epub 2013/12/20. doi:10.1186/1475-2875-12-452. Cited in: Pubmed; PMID 24350670.
26. Yang D, Xu C, Wang J, Zhao Y. Spatiotemporal epidemic characteristics and risk factor analysis of malaria in Yunnan Province, China. *BMC Public Health.* 2017 Jan 11;17(1):66. eng. Epub 2017/01/13. doi:10.1186/s12889-016-3994-9. Cited in: Pubmed; PMID 28077125.
27. Yin JH, Zhang L, Yi BY, Zhou SS, Xia ZG. Imported malaria from land bordering countries in China: A challenge in preventing the reestablishment of malaria transmission. *Travel Med Infect Dis.* 2023 May-Jun;53:102575. eng. Epub 2023/04/27. doi:10.1016/j.tmaid.2023.102575. Cited in: Pubmed; PMID 37100163.
28. Wang J, Chen Y, Pan D, Zhang J, Zhang Y, Lu Z. Source and health risk assessment of soil polycyclic aromatic hydrocarbons under straw burning condition in Changchun City, China. *Sci Total Environ.* 2023 Oct 10;894:165057. eng. Epub 2023/06/24. doi:10.1016/j.scitotenv.2023.165057. Cited in: Pubmed; PMID 37353020.
29. Sohail M, Ali Musstjab Akber Shah Eqani S, Bokhari H, Zaffar Hashmi M, Ali N, Alamdar A, Podgorski JE, Adelman D, Lohmann R. Freely dissolved organochlorine pesticides (OCPs) and polychlorinated biphenyls (PCBs) along the Indus River Pakistan: spatial pattern and risk assessment. *Environ Sci Pollut Res Int.* 2022 Sep;29(43):65670-65683. eng. Epub 2022/05/03. doi:10.1007/s11356-022-20418-4. Cited in: Pubmed; PMID 35499728.
30. Abbas W, Lee S, Kim S. Joint estimation of hand-foot-mouth disease model and prediction in korea using the ensemble kalman filter. *PLoS Comput Biol.* 2025 Apr;21(4):e1012996. eng. Epub 2025/04/17. doi:10.1371/journal.pcbi.1012996. Cited in: Pubmed; PMID 40245103.
31. Guo J, He Q, Li Y. Machine learning-based prediction of vitamin D deficiency: NHANES 2001-2018. *Front Endocrinol (Lausanne).* 2024;15:1327058. eng. Epub 2024/03/07. doi:10.3389/fendo.2024.1327058. Cited in: Pubmed; PMID 38449846.
32. Reemtsma T, Savric I, Jekel M. A potential link between the turnover of soil organic matter and the release of aged organic contaminants. *Environ Toxicol Chem.* 2003 Apr;22(4):760-6. eng. Epub 2003/04/11. Cited in: Pubmed; PMID 12685710.
33. Antonelli P, Grizard S, Tran FH, Lejon D, Bellemain A, Van, Mavingui P, Roiz D, Simard F, Martin E, Abrouk D, Vigneron A, Minard G, Valiente Moro C. Bioaccumulation of polycyclic aromatic hydrocarbons and microbiota dynamics across developmental stages of the Asian tiger mosquito, *Aedes albopictus* exposed to urban pollutants. *Ecotoxicol Environ Saf.* 2024 Nov 1;286:117214. eng. Epub 2024/10/25. doi:10.1016/j.ecoenv.2024.117214. Cited in: Pubmed; PMID 39447296.
34. Tekoh TA, Mugenzi LMJ, Menze B, Tchappa W, Wondji M, Tchouakui M, Small G, Wondji CS. The duplicated cytochrome P450 CYP6P9a/b confers cross-resistance to a mitochondrial complex I inhibitor in the African malaria vector *Anopheles funestus*. *BMC Genomics.* 2025 Sep 26;26(1):837. eng. Epub 2025/09/27. doi:10.1186/s12864-025-11984-1. Cited in: Pubmed; PMID 41013195.
35. Shahid M, Natasha, Dumat C, Niazi NK, Xiong TT, Farooq ABU, Khalid S. Ecotoxicology of Heavy Metal(loid)-Enriched Particulate Matter: Foliar Accumulation by Plants and Health Impacts. *Rev Environ Contam Toxicol.* 2021;253:65-113. eng. Epub 2020/01/04. doi:10.1007/398_2019_38. Cited in: Pubmed; PMID 31897760.
36. Hama Aziz KH, Mustafa FS, Omer KM, Hama S, Hamarawf RF, Rahman KO. Heavy metal pollution in the aquatic environment: efficient and low-cost removal approaches to eliminate their toxicity: a review. *RSC Adv.* 2023 Jun 9;13(26):17595-17610. eng. Epub 2023/06/14. doi:10.1039/d3ra00723e. Cited in: Pubmed; PMID 37312989.
37. Wesner J, Kraus JM, Henry B, Kerby J. Metamorphosis and the Impact of Contaminants on Ecological Subsidies. In: Kraus JM, Walters DM, Mills MA, editors. *Contaminants and Ecological Subsidies: The Land-Water Interface*. Cham: Springer International Publishing; 2020. p. 111-125.
38. Liu Y, Luo X, Zeng Y, Wang Q, Tu W, Yang C, Mai B. Trophic Magnification of Short- and Medium-Chain Chlorinated Paraffins in Terrestrial Food Webs and Their Bioamplification in Insects and Amphibians during Metamorphosis. *Environ Sci Technol.* 2020 Sep 15;54(18):11282-11291. eng. Epub 2020/08/22. doi:10.1021/acs.est.0c03096. Cited in: Pubmed; PMID 32822158.
39. Kim HM, Choi IP, Holsapple MP. Direct exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) increases infectivity of human erythrocytes to a malarial parasite. *Life Sci.* 1994;54(12):PI215-20. eng. Epub 1994/01/01. doi:10.1016/0024-3205(94)00453-6. Cited in: Pubmed; PMID 8121237.
40. Thellier M, Gemegah AAJ, Tantaoui I. Global Fight against Malaria: Goals and Achievements 1900-2022. *J Clin Med.* 2024 Sep 24;13(19). eng. Epub 2024/10/16. doi:10.3390/jcm13195680. Cited in: Pubmed; PMID 39407740.

Figures

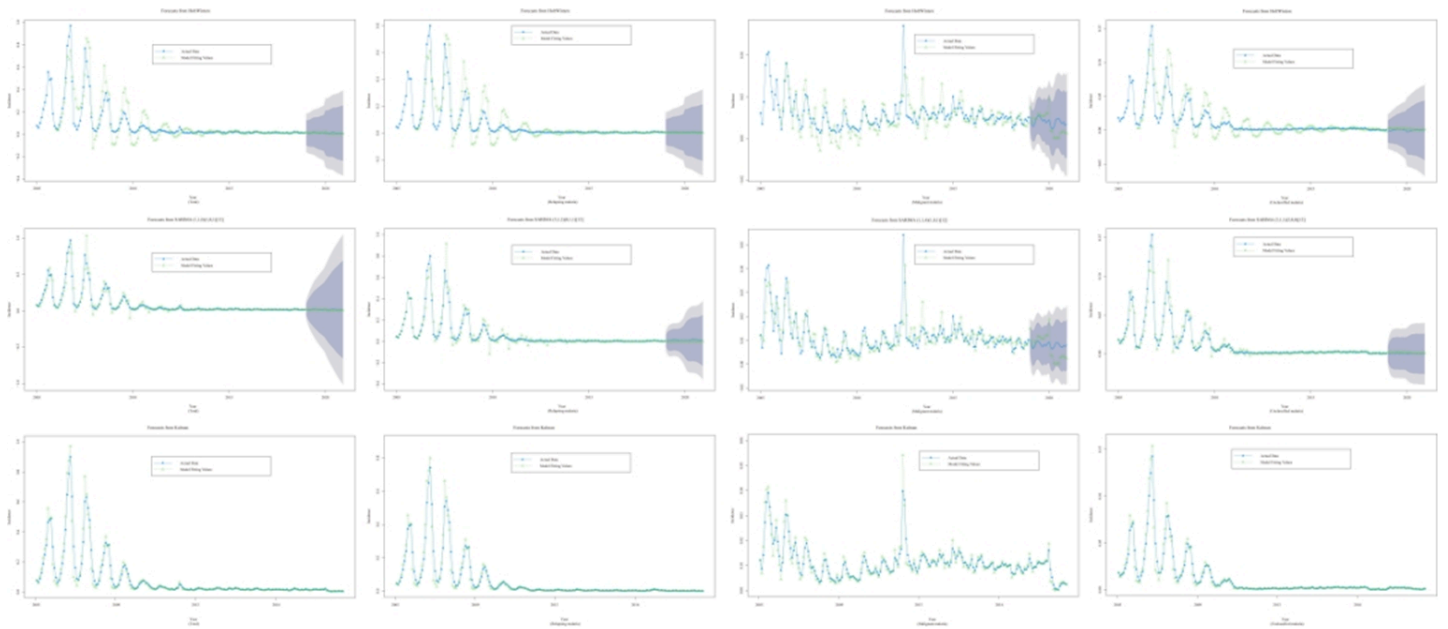


Figure 1

Fitting status between the actual incidence of different malaria in China mainland from Jan 2005 to Dec 2018 and the predicted incidence from Jan 2019 to Dec 2020 using the preferred three models. The deep shaded regions indicate 80% confidence intervals, the light shaded regions indicate 95% confidence intervals.

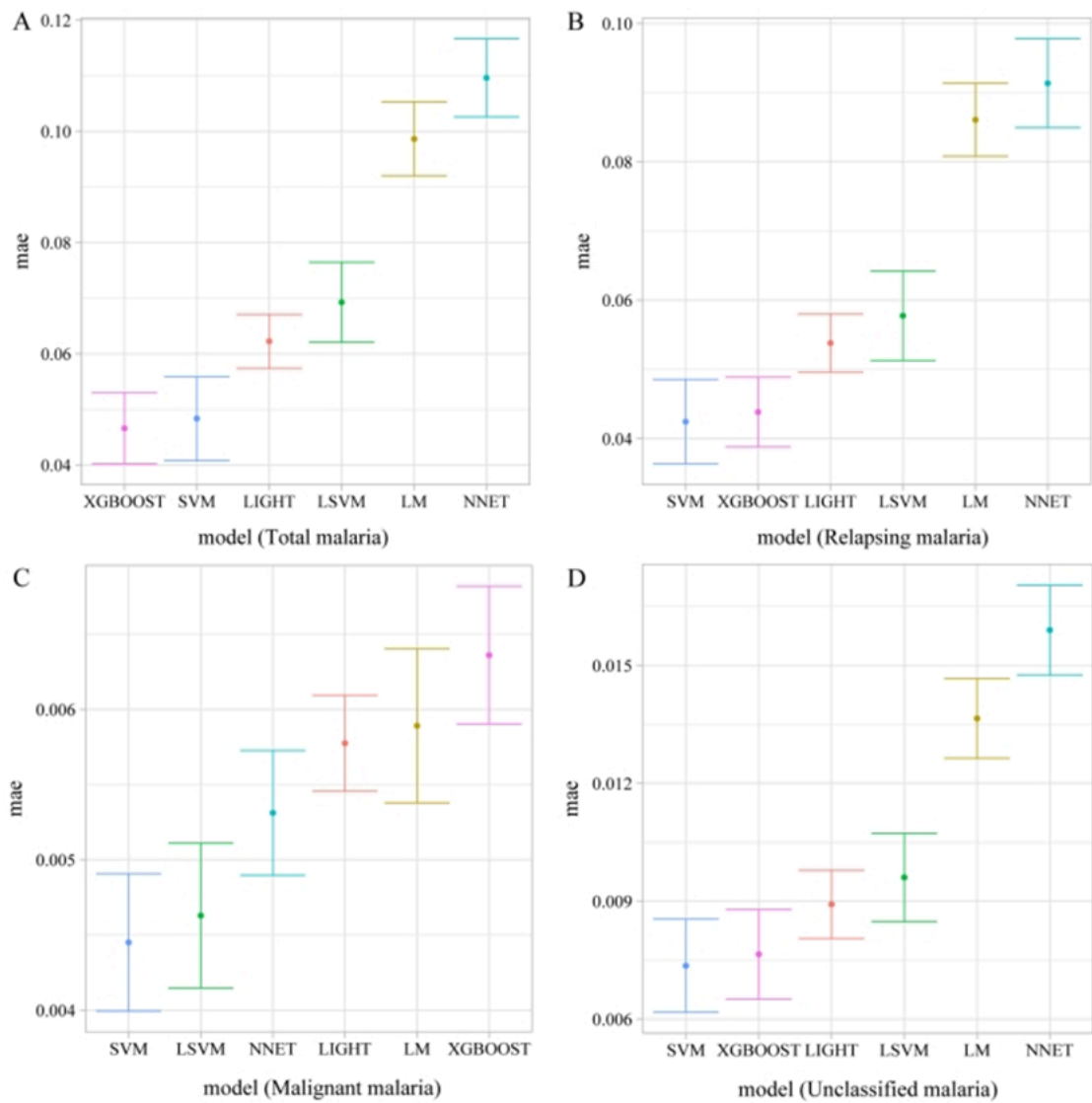


Figure 2

Prediction performance metrics comparison in different malaria for different machine learning algorithms.

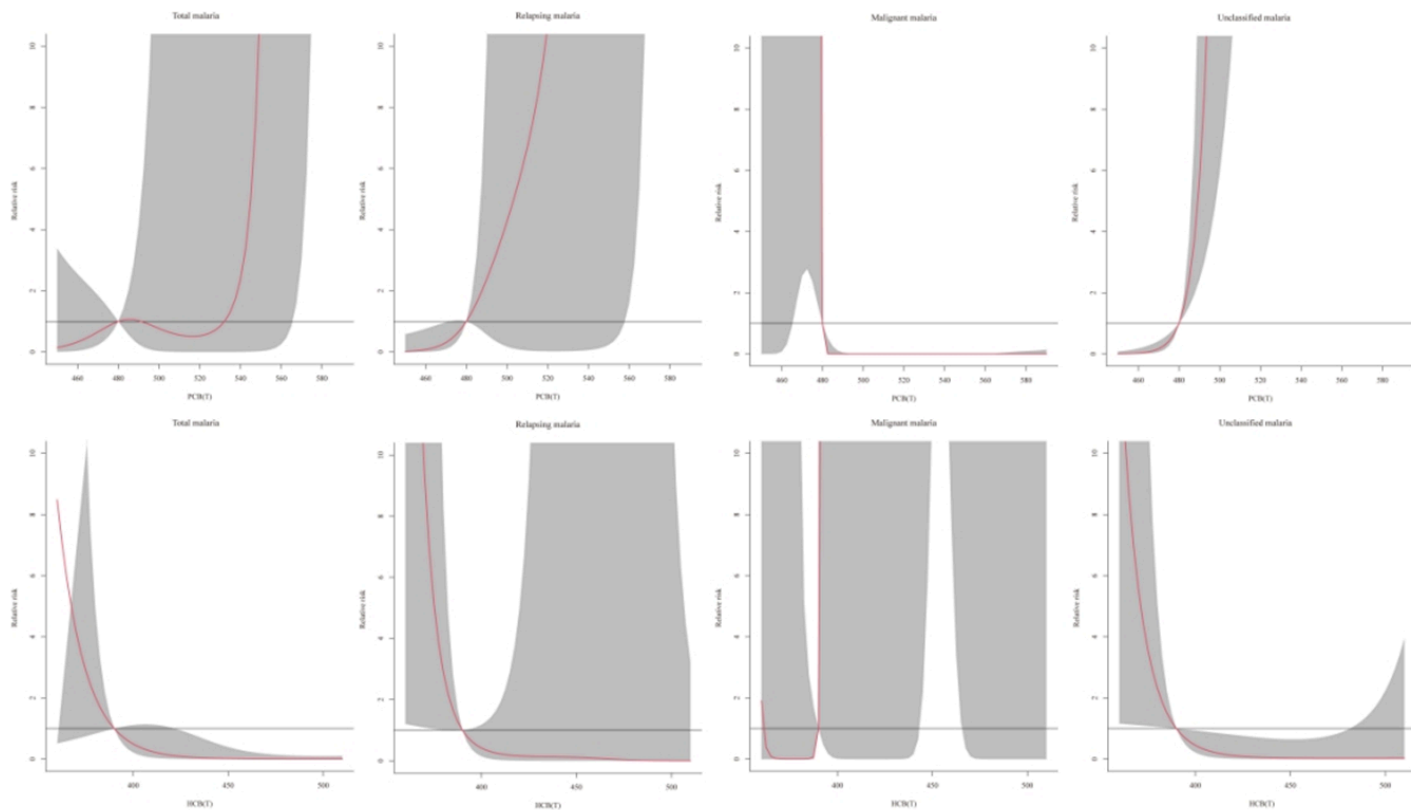


Figure 3

Overall relationship of in different malaria risks with POPs indicators in China mainland, with the median of each exposure serving as the reference value.

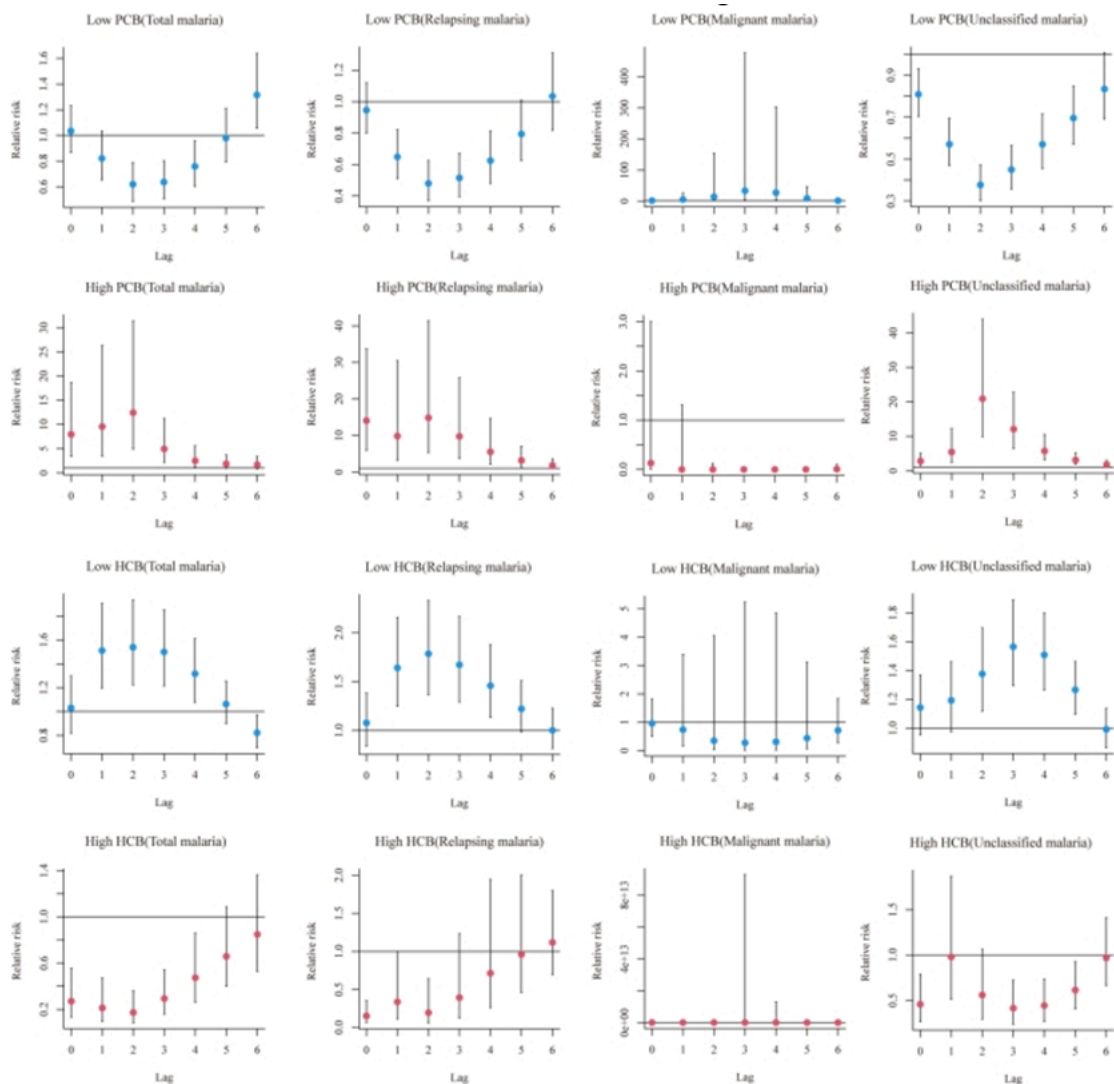


Figure 4

Summary of estimated extreme effects at the 5th and the 95th percentile of PCB and HCB on different malaria cases at different lag months.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supply.docx](#)