Environmental Determinants of Hemorrhagic Fever with Renal Syndrome in High-risk Counties in China: A Time Series Analysis (2002-2012)

Junyu He[1†], Jimi He[2†], Zhihai Han[3,4†], Yue Teng[5], Wenyi Zhang[6*], Wenwu Yin[7*]

[1] Ocean College, Zhejiang University, Zhoushan, China;

[2] School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China;

[3] Navy General Hospital of PLA, Beijing, People's Republic of China;

[4]Navy Clinical College of Anhui Medical University ,Hefei, China;

[5] State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, China;

[6] Center for Disease Surveillance of PLA, Institute of Disease Control and Prevention of PLA, Beijing, China;

[7] Division of Infectious Diseases, Key Laboratory of Surveillance and Early-warning on Infectious Disease, Chinese Center for Disease Control and Prevention, Beijing, China.

[†]These authors contributed equally to this work.

Corresponding authors: Wenyi Zhang , Institute of Disease Control and Prevention of PLA, 20 Dong-Da Street, Fengtai District, Beijing 100071, People's Republic of China. E-mail addresses: zwy0419@126.com

Wenwu Yin, Division of Infectious Diseases, Chinese Center for Disease Control and Prevention, 155 Changbai Road, Changping District, Beijing 102206, China. E-mail addresses: yinww@chinacdc.cn

Supporting Information

Text S1 An example of SARIMA model

Text S2 More detailed information of the final four SARIMAX model

Table S1 Comparing AIC values of SARIMA models with various structure

Table S2 SARIMA model of HFRS cases in four counties from China during 2002-2012

Table S3 Information of SARIMAX# model

Table S4 Correlation coefficients between HFRS cases and four environmental factors with 0~5 months lags

Figure S1 The ACF of residuals of SARIMAX models for (a) Raohe County, (b) Mishan County, (c) Chang'an County and (d) Hu County

Note : Table S1 and Table S4 are listed in the ".xlsx" files naming "Table S1" and "Table S4".

Text S1 An example of SARIMA model

Assume the structure of SARIMA model is $(2,1,1)(0,1,1)_S$, the Eq 2 besomes

$$\Phi_0(B^{12})\phi_2(B)\nabla_{12}^1\nabla^1 Z_t = \Theta_1(B^S)\theta_1(B)\varepsilon_t \quad \text{(S1)}$$

Through introducing Φ, ϕ, Θ, θ, ∇ into Eq. S1, we can obtain the equation as following,

$$1\times(1-\phi_1 B-\phi_2 B^2)(1-B^{12})(1-B)Z_t = (1-\Theta_1 B^{12})(1-\theta_1 B)\varepsilon_t$$

$$\begin{aligned}\Rightarrow & [1-(1+\phi_1)B+(\phi_1-\phi_2)B^2+\phi_2 B^3-B^{12}+(1+\phi_1)B^{13}-(\phi_1-\phi_2)B^{14}-\phi_2 B^{15}]Z_t \\ & =(1-\theta_1 B-\Theta_1 B^{12}+\theta_1\Theta_1 B^{13})\varepsilon_t\end{aligned}$$

$$\begin{aligned}\Rightarrow & Z_t-(1+\phi_1)Z_{t-1}+(\phi_1-\phi_2)Z_{t-2}+\phi_2 Z_{t-3}-Z_{t-12}+(1+\phi_1)Z_{t-13}-(\phi_1-\phi_2)Z_{t-14}-\phi_2 Z_{t-15} \\ & =\varepsilon_t-\theta_1\varepsilon_{t-1}-\Theta_1\varepsilon_{t-12}+\theta_1\Theta_1\varepsilon_{t-13}\end{aligned}$$

$$\begin{aligned}\Rightarrow Z_t = & (1+\phi_1)Z_{t-1}-(\phi_1-\phi_2)Z_{t-2}-\phi_2 Z_{t-3}+Z_{t-12}-(1+\phi_1)Z_{t-13} \\ & +(\phi_1-\phi_2)Z_{t-14}+\phi_2 Z_{t-15}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\Theta_1\varepsilon_{t-12}+\theta_1\Theta_1\varepsilon_{t-13}\end{aligned} \quad \text{(S2)}$$

Text S2 More detailed information of the final four SARIMAX model

The SARIMA parts of Equation 3a~d are shown as follows:

$$(1-0.8779B)\nabla_{12}^{1}Z_t=(1+0.6343B^{12})(1+0.6183B)\varepsilon_t \quad \text{(S3a)}$$

$$(1-0.2764B-0.2364B^{2})\nabla_{12}^{1}Z_t=(1+0.6179B^{12})\varepsilon_t \quad \text{(S3b)}$$

$$(1-0.4185B)\nabla_{12}^{1}Z_t=(1+1.0000B^{12})\varepsilon_t \quad \text{(S3c)}$$

$$(1+0.2864B)\nabla_{12}^{1}\nabla^{1}Z_t=(1+0.7351B^{12})(1+0.5550B)\varepsilon_t \quad \text{(S3d)}$$

Through expanding the equations above, we can obtain Equations S4a~d, respectively:

$$Z_t=0.8779Z_{t-1}+Z_{t-12}-0.8779Z_{t-13}+\varepsilon_t+0.6183\varepsilon_{t-1}+0.6343\varepsilon_{t-12}+0.3922\varepsilon_{t-13} \quad \text{(S4a)}$$

$$Z_t=0.2764Z_{t-1}+0.2364Z_{t-2}+Z_{t-12}-0.2764Z_{t-13}-0.2364Z_{t-14}+\varepsilon_t+0.6179\varepsilon_{t-12} \quad \text{(S4b)}$$

$$Z_t=0.4185Z_{t-1}+Z_{t-12}-0.4185Z_{t-13}+\varepsilon_t+\varepsilon_{t-12} \quad \text{(S4c)}$$

$$Z_t=0.7136Z_{t-1}+0.2864Z_{t-2}+Z_{t-12}-0.7134Z_{t-13}-0.2864Z_{t-14}+\varepsilon_t+0.5550\varepsilon_{t-1}+0.7351\varepsilon_{t-12}+0.4080\varepsilon_{t-13} \quad \text{(S4d)}$$

The standard deviations of ε_t for the four equations S3a-S3d are 0.2652, 0.3457, 0.2073 and 0.0003, respectively; while the mean values of ε_t for the four equations are zero.

Table S2 SARIMA model of HFRS cases in four counties from China during 2002-2012

County	Model fitting								Model forecasting		
	Model structure	Model components	Estimate	Std. Error	t value	p value	R^2	AIC	R^2	RMSE	MAE
Raohe	$(1,0,1)\times(0,1,1)_{12}$	AR1	0.9038	0.1006	8.9849	0.0000	0.6490	213.9046	0.5971	2.4189	1.4526
		MA1	-0.6898	0.1723	-4.0044	0.0001					
		SMA1	-0.6895	0.1008	-6.8382	0.0000					
Mishan	$(2,0,0)\times(0,1,1)_{12}$	AR1	0.3408	0.0967	3.5247	0.0006	0.6390	237.1565	0.8508	1.5569	1.2682
		AR2	0.2163	0.0947	2.2837	0.0242					
		SMA1	-0.6265	0.0983	-6.3725	0.0000					
Chang'an	$(1,0,0)\times(0,1,1)_{12}$	AR1	0.4498	0.0865	5.1969	0.0000	0.7780	192.3766	0.7849	18.4786	11.8034
		SMA1	-1.0000	0.1653	-6.0511	0.0000					
Hu	$(1,1,1)\times(0,1,1)_{12}$	AR1	-0.3014	0.1364	-2.2093	0.0291	0.7029	184.2755	0.7857	8.0882	5.3127
		MA1	-0.5681	0.1326	-4.2846	0.0000					
		SMA1	-0.8306	0.1426	-5.8260	0.0000					

Table S3 Information of SARIMAX[#] model

County	Model fitting							
	Model structure	Model components	Estimate	Std. Error	t value	p value	R^2	AIC
Raohe	$(2,0,1)\times(0,1,1)_{12}$	AR1	1.1631	0.1283	9.0623	0.0000	0.7512	215.2140
		AR2	-0.1815	0.1139	-1.5932	0.1136		
		MA1	-0.8820	0.0733	-12.0290	<2.2E-16		
		SMA1	-0.7647	0.1082	-7.0660	0.0000		
		RH_lag4	0.0041	0.0018	2.3416	0.0208		
		RAIN_lag1	0.0040	0.0018	2.1793	0.0312		
		NDVI_lag1	0.0112	0.0045	2.4963	0.0139		
Mishan	$(2,0,0)\times(0,1,1)_{12}$	AR1	0.2798	0.0954	2.9339	0.0040	0.6682	246.2721
		AR2	0.2246	0.0910	2.4677	0.0149		
		SMA1	-0.6251	0.1020	-6.1281	0.0000		
		RAIN_lag3	0.0048	0.0022	2.2097	0.0289		
		RAIN_lag4	0.0052	0.0023	2.2050	0.0293		
Chang'an	$(1,0,0)\times(0,1,1)_{12}$	AR1	0.4882	0.0815	5.9903	0.0000	0.8423	178.3759
		SMA1	-1.0000	0.2225	-4.4951	0.0000		
		RH_lag1	-0.0209	0.0069	-3.0455	0.0028		
		RH_lag5	0.0182	0.0069	2.6241	0.0098		
		NDVI_lag2	0.0131	0.0035	3.6907	0.0003		
Hu	$(1,1,1)\times(0,1,1)_{12}$	AR1	-0.1807	0.1480	-1.2208	0.2244	0.7790	191.3398
		MA1	-0.6384	0.1313	-4.8621	0.0000		
		SMA1	-0.7379	0.1121	-6.5839	0.0000		
		RH_lag5	0.0163	0.0068	2.3853	0.0185		

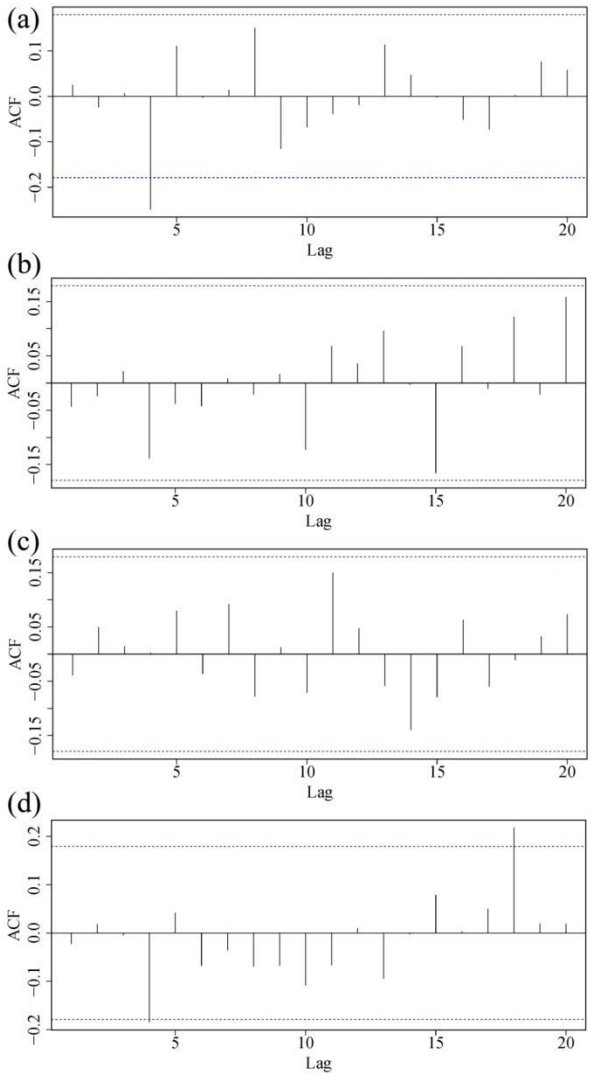

Figure S1 The ACF of residuals of SARIMAX models for (a) Raohe County, (b) Mishan County, (c) Chang'an County and (d) Hu County