SUPPORTING INFORMATION

ANN DESIGN

Artificial neural networks (ANNs) were chosen since they are nonlinear methodologies that are often used to solve problems such as nonlinear mapping, forecasting, classification, and clustering, amongst others. They can contort space and approximate, making them universal approximators and giving them an elevated generalization capability[24]. Such behavior in an ANN is related to its structure, composed of nonlinear units named artificial neurons. These neurons present adjustable weights that sum the inputs and pass the information through an activation function that introduces nonlinearity in the response. The neurons are organized into layers, commonly named input, hidden (intermediate), and output layers[42].

Among the variants of the neural models, we are interested in those adequate to solve nonlinear mapping, feedforward (FNN), and recurrent neural networks (RNN). The first class presents the models in which the information flows from the input to the output layers without feedback. The recurrent methodologies allow the presence of feedback loops.

As the model performance varies depending on the ANN used and its design, we chose to apply four different ANNs, with and without Z-score, to different scenarios (considering all variables, excluding BC concentration, and excluding fires), and varying the number of neurons in the hidden layers. We also considered Multilayer Perceptron (MLP) with one and two hidden layers. In this study, we addressed two widely used approaches, the Multilayer Perceptron (MLP) and the Radial Basis Function Network (RBF), together with two unorganized machines (UM), the Extreme Learning Machines (ELM) and the Echo State Networks (ESN). Only the latter is an RNN, while the others are FNN[41, 42]. The MLP and RBF are fully trained methodologies because all weights are adjusted. The UM tunes only the output layer, which confers a simple implementation and low computational cost to such methods.

Regarding the MLP, the model was adjusted using the modified scaled conjugate gradient method, which works by reducing iteratively the mean square error[43]. The unsupervised step of the RBF considers the K-Means algorithm, while the supervised step uses the Moore-Penrose Inverse Operation, a closed form-solution. The last operation was also applied to adjust the output layer of the UMs, since it leads to the minimum MSE[38, 39]. All models were implemented in Matlab.

About the activation function, all neural networks had the linear identity function in the output neurons, and the hyperbolic tangent in the hidden layer, with exception of RBF, where the K-Means algorithm was applied to cluster the centers. Even more, the weights were generated randomly in the interval [-1; 1], and the data were normalized in the same range. After achieving the output responses, the data were renormalized to analyze the error in the original domain[24, 41].

The best performances of the neural networks in predicting cardiorespiratory hospital admissions were based on the lower Root Mean Square Error - RMSE – Equation 1. According to Siqueira et al.[42], the cost function the ANNs minimize is the RMSE and, in the case in which different error metrics indicate distinct models as the best, the one with the lowest RMSE should be assumed as the best one.

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N} (d_n - y_n)^2}\,, \tag{1}$$

The Mean Absolute Percentage Error (MAPE) was also calculated, as it gives a more tangible error metric, being in percentage (Equation 2), as follows:

$$MAPE = \frac{1}{N}\sum_{n=1}^{N} \left|\frac{d_n - y_n}{d_n}\right| \times 100\,, \tag{2}$$

where d_n is the observed value, y_n is the response of the neural model, and N is the total number of samples considered. RMSE indicates which model is best relative to the four tested ANNs, whereas MAPE indicates how good that best model is relative to the actual observations of hospital admissions.

The best of 30 independent simulations is considered for each neural model. The results were presented considering the use or not of Z-score, which is a deseasonalization technique. The goal was to reveal if transforming the data into stationary data without seasonal components brings performance gains[42-44]. The Z-score consists of subtracting the value of each sample from the mean and dividing the result by the standard deviation. At the end of the predictions, the process was reversed to analyze the performances in the original form.

In the sense of knowing if the error values are statistically different from each other, meaning an ANN performed better than other techniques, we applied the Friedman test.

Regarding the neural networks, to determine the best number of neurons in the hidden layer, it was performed a grid search from 10 to 300 neurons, and the best overall performance was compared.

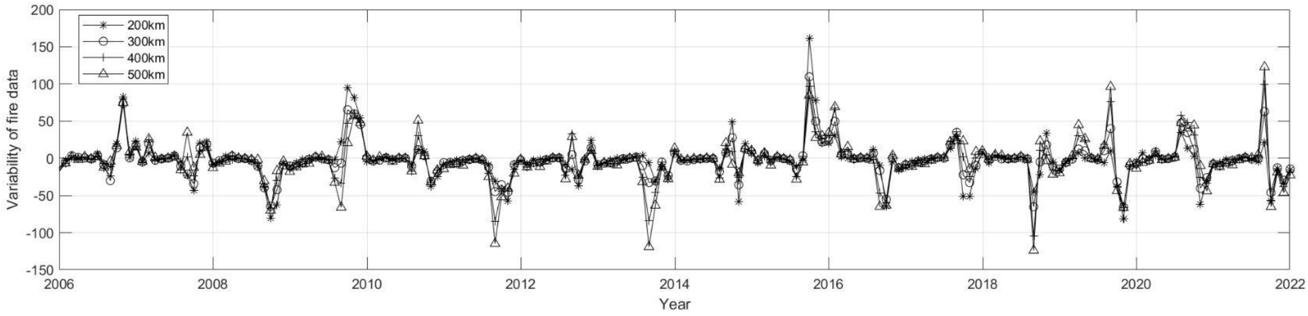

Figure S1 – Variability of fire data from 200km to 500km around Manaus city.

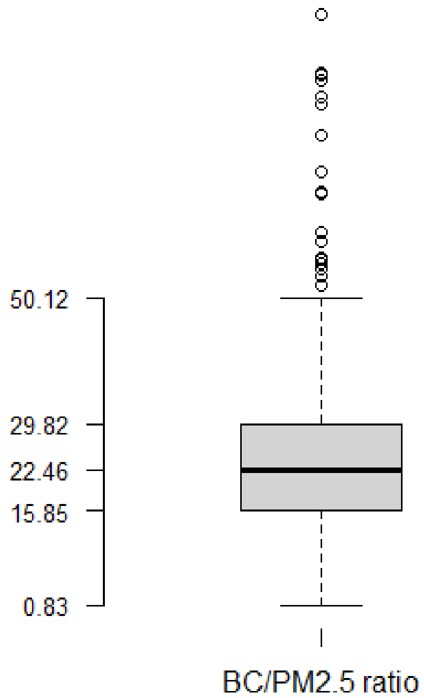

Figure S2 – Box plot of % BC in $PM_{2.5}$.

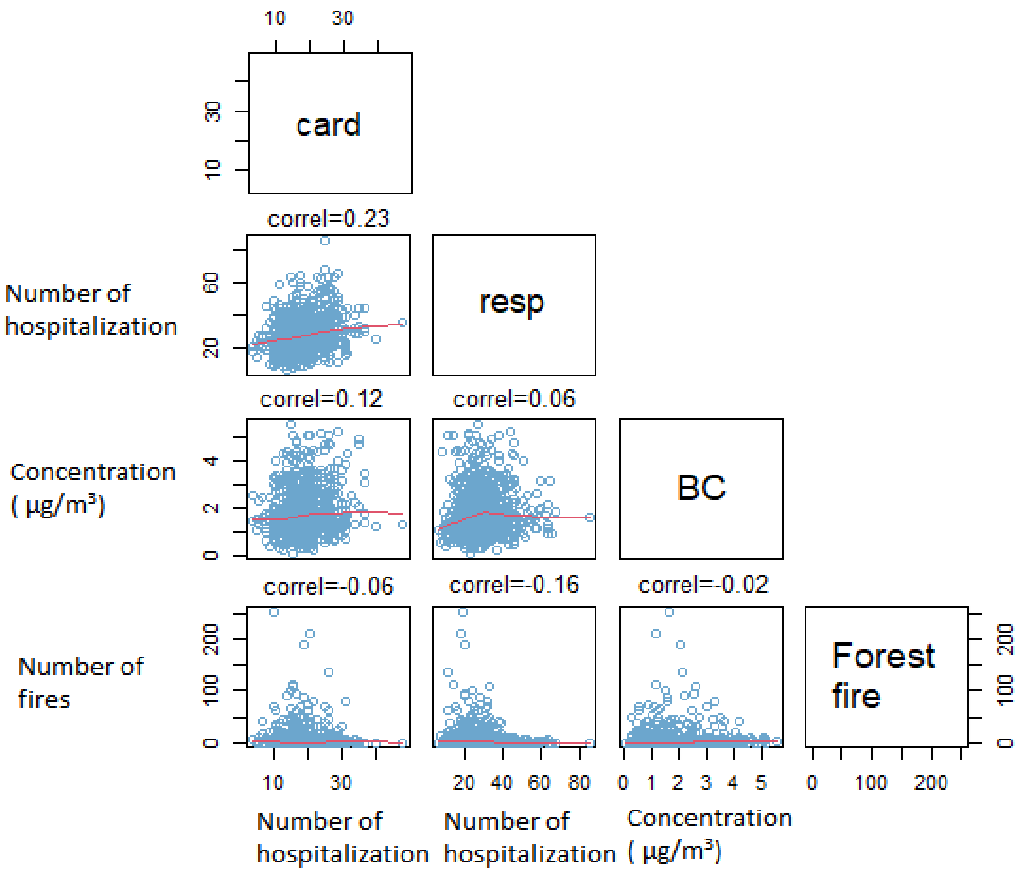

Figure S3 – Dispersion Diagram between cardiorespiratory diseases (card and resp), BC, and forest fire with respective Pearson correlation coefficient.

Table S1 - Spearman correlation between cardiorespiratory diseases, BC, and forest fire.

	Cardiovascular	Respiratory	BC	Forest Fire
Cardiovascular	1			
Respiratory	0.25	1		
BC	0.13	0.13	1	
Forest Fire	-0.02	-0.22	0.009	1

Table S2 - Kendall correlation between cardiorespiratory diseases, BC, and forest fire.

	Cardiovascular	Respiratory	BC	Forest Fire
Cardiovascular	1			
Respiratory	0.18	1		
BC	0.09	0.08	1	
Forest Fire	-0.02	-0.17	0.007	1

Table S3 – Performance of neural networks (RMSE) on predicting hospital admissions for *respiratory diseases* (RD) including *all inputs*

Z-Score	Predictor*	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
Without	ELM	15.94	18.01	17.43	16.05	15.16	16.74	15.02	16.61
	ESN	17.31	17.39	17.39	15.96	14.71	15.97	15.22	16.48
	MLP	12.21	12.40	11.95	11.83	11.98	12.34	11.95	12.36
	MLP - 2	11.39	<u>10.42</u>	11.47	14.07	15.84	11.22	15.28	16.67
	RBF	18.14	19.00	18.86	14.55	13.53	18.99	14.74	16.39
With	ELM	17.20	16.99	17.87	15.41	16.36	17.62	15.97	16.89
	ESN	17.16	17.78	17.90	14.92	15.92	16.99	15.30	16.33
	MLP	12.41	12.08	11.61	11.92	11.82	12.28	12.14	12.32
	MLP - 2	11.05	11.35	16.19	14.05	15.85	16.43	15.42	16.69
	RBF	18.00	19.10	19.04	14.51	13.46	18.95	14.85	16.32

*ELM: Extreme Learning Machines; ESN: Echo State Neural Networks; MLP: Multilayer Perceptron; MLP 2: Multilayer Perceptron with 2 layers; RBF: Radial Basis Function Network. The best performances are highlighted in grey. The best overall is underlined.

Table S4 – Performance of neural networks (RMSE) on predicting hospital admissions for *respiratory diseases* (RD) excluding *forest fire* as input

Z-Score	Predictor*	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
Without	ELM	16.19	17.44	15.92	15.98	16.13	17.79	14.56	16.09
	ESN	16.75	17.00	17.47	14.69	15.27	17.38	14.40	16.16
	MLP	12.34	12.33	11.91	12.05	12.05	12.22	11.93	12.31
	MLP - 2	12.21	12.12	11.79	12.03	11.87	12.25	12.06	12.44
	RBF	19.48	19.26	19.05	14.57	13.54	19.39	15.02	16.67
With	ELM	16.51	18.18	17.72	14.10	16.16	17.49	14.84	16.20
	ESN	17.13	17.55	16.86	16.12	15.85	17.21	14.92	16.73
	MLP	12.46	12.09	12.02	12.00	12.06	12.43	12.06	12.27
	MLP - 2	12.40	12.20	11.95	<u>11.76</u>	12.00	12.27	11.92	12.27
	RBF	19.51	19.17	19.05	14.59	13.59	19.43	15.01	16.64

*ELM: Extreme Learning Machines; ESN: Echo State Neural Networks; MLP: Multilayer Perceptron; MLP 2: Multilayer Perceptron with 2 layers; RBF: Radial Basis Function Network. The best performances are highlighted in grey. The best overall is underlined.

Table S5 – Performance of neural networks (RMSE) on predicting hospital admissions for *respiratory diseases* (RD) excluding *Black Carbon* as input

Z-Score	Predictor*	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
Without	ELM	18.35	18.75	18.80	15.07	14.18	17.65	14.48	15.44
	ESN	18.59	16.75	16.70	16.02	13.90	16.02	15.15	14.08
	MLP	12.38	11.92	11.80	11.80	11.98	12.27	11.76	12.22
	MLP - 2	12.36	12.76	<u>11.74</u>	18.36	11.88	12.44	17.65	13.33
	RBF	18.33	19.09	18.98	14.52	13.51	19.00	14.82	16.37
With	ELM	16.40	17.07	18.66	15.38	14.75	18.08	14.40	14.52
	ESN	18.11	18.00	14.67	16.37	14.41	15.44	15.46	16.50
	MLP	11.99	12.09	11.91	11.81	11.88	12.09	11.92	12.13
	MLP - 2	12.55	12.28	11.85	18.43	16.84	12.91	17.66	12.49
	RBF	18.39	18.92	19.02	14.47	13.52	18.95	14.89	16.18

*ELM: Extreme Learning Machines; ESN: Echo State Neural Networks; MLP: Multilayer Perceptron; MLP 2: Multilayer Perceptron with 2 layers; RBF: Radial Basis Function Network. The best performances are highlighted in grey. The best overall is underlined.

Table S6 – Performance of neural networks (RMSE) on predicting hospital admissions for *cardiovascular diseases* (CD) including *all inputs*

Z-Score	Predictor*	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
Without	ELM	7.93	8.43	7.46	9.13	8.95	8.16	8.08	7.50
	ESN	7.72	7.78	7.37	8.77	8.65	8.42	7.45	8.02
	MLP	5.36	5.43	5.47	5.39	5.41	5.41	5.29	5.39
	MLP - 2	<u>4.87</u>	5.25	5.46	5.48	5.35	5.45	5.43	5.69
	RBF	7.89	8.20	5.45	8.09	7.92	8.34	7.98	8.21
With	ELM	7.91	8.69	8.60	9.27	8.96	7.84	8.25	7.45
	ESN	7.80	8.25	7.22	8.83	8.57	8.31	6.93	7.85
	MLP	5.54	5.37	5.39	5.33	5.34	5.41	5.27	5.42
	MLP - 2	5.11	5.25	5.47	5.73	6.25	5.48	5.41	5.59
	RBF	7.96	8.30	5.43	8.16	7.84	8.37	7.99	8.27

*ELM: Extreme Learning Machines; ESN: Echo State Neural Networks; MLP: Multilayer Perceptron; MLP 2: Multilayer Perceptron with 2 layers; RBF: Radial Basis Function Network. The best performances are highlighted in grey. The best overall is underlined.

Table S7 – Performance of neural networks (RMSE) on predicting hospital admissions for *cardiovascular diseases* (CD) excluding *forest fire* as input

Z-Score	Predictor*	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
Without	ELM	7.97	8.10	7.84	9.17	8.46	8.70	7.70	7.93
	ESN	7.70	8.08	7.76	9.16	8.26	7.75	7.77	7.86
	MLP	5.37	5.40	5.43	5.39	5.30	5.46	5.30	5.40
	MLP - 2	5.43	5.36	5.42	5.37	5.34	5.45	5.24	5.39
	RBF	8.49	8.40	5.43	8.28	8.16	8.45	8.14	8.42
With	ELM	7.70	7.84	8.06	9.23	8.34	8.19	7.80	7.57
	ESN	7.43	8.04	7.89	8.77	8.52	8.04	7.89	7.47
	MLP	5.40	5.40	5.33	5.45	5.38	5.38	5.27	5.40
	MLP - 2	5.45	5.42	5.44	5.39	5.32	5.49	<u>5.21</u>	5.41
	RBF	8.50	8.35	5.42	8.18	8.13	8.47	8.11	8.39

*ELM: Extreme Learning Machines; ESN: Echo State Neural Networks; MLP: Multilayer Perceptron; MLP 2: Multilayer Perceptron with 2 layers; RBF: Radial Basis Function Network. The best performances are highlighted in grey. The best overall is underlined.

Table S8 – Performance of neural networks (RMSE) on predicting hospital admissions for *cardiovascular diseases* (CD) excluding *Black Carbon* as input

Z-Score	Predictor*	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
Without	ELM	7.85	7.28	6.20	6.92	6.92	8.76	7.19	7.84
	ESN	7.85	7.64	6.22	7.73	7.95	7.30	7.55	7.60
	MLP	5.37	5.36	5.31	5.29	5.30	5.42	<u>5.24</u>	5.40
	MLP - 2	5.53	5.42	5.43	5.53	5.36	10.32	5.37	5.85
	RBF	7.89	8.12	5.37	7.96	8.07	8.39	7.91	8.20
With	ELM	7.11	7.89	6.39	6.31	6.37	8.18	6.02	7.52
	ESN	7.75	6.84	6.35	7.90	7.59	8.37	7.70	7.38
	MLP	5.44	5.35	5.36	5.32	5.28	5.40	5.25	5.35
	MLP - 2	5.71	9.66	5.42	5.44	5.42	10.48	5.26	5.41
	RBF	7.79	8.16	5.35	8.07	7.96	8.36	7.97	8.18

*ELM: Extreme Learning Machines; ESN: Echo State Neural Networks; MLP: Multilayer Perceptron; MLP 2: Multilayer Perceptron with 2 layers; RBF: Radial Basis Function Network. The best performances are highlighted in grey. The best overall is underlined.

The ANN performances showed that, considering the respiratory results, the RMSE of MLP with two hidden layers (MLP-2) was 19% lower than the MLP with one hidden layer and 82% lower than the RBF results. The RMSE difference for cardiovascular results ranges from 10% (from MLP-2 to MLP) to 63% (from MLP-2 to ELM).

The rationality of the model was tested by considering a case of lag -1. Specifically, the best model (MLP, 2 hidden layers) was re-run for lag -1. The RMSE of this test case was 6.67. This RMSE exceeds the best result (lag +1, RMSE = 4.87). Therefore, the model of lag -1 is quantitatively inferior. This result shows that the model does not violate cause-and-effect in the temporal relationship between exposure and health outcome.