A MODIFIED HIDDEN SEMI-MARKOV MODEL FOR TRAFFIC RELATED PM_{10} POLLUTION LEVELS ESTIMATION

Maryam Shekarrizfard (Corresponding Author)
Doctoral Candidate
Department of Civil Engineering & Applied Mechanics
McGill University
Tel: 1-514-589-4353, Fax: 1-514-398-7361
Email: maryam.shekarrizfard@mail.mcgill.ca

Ayoub Karimi-Jashni,
Assistant Professor
Department of Civil and Environmental Engineering
Shiraz University
Tel: +98-711-6473499, Fax: +98-711-6473161
Email: akarimi@shirazu.ac.ir

Kamal Hadad,
Professor
School of Engineering
Shiraz University
Tel: +98-711-6473474, Fax: +98-711-6473474
Email: hadadk@shirazu.ac.ir

Seyed Ali Akbar Safavi
Professor
School of Engineering,
Shiraz University
Tel: +98-711-6133112, Fax: +98-711-2303081
Email: safavi@shirazu.ac.ir

A Modified Hidden Semi-Markov Model For Traffic Related Pm_{10} Pollution Levels Estimation

Abstract Traffic related PM_{10} (particulate matter with less than 10 microns) pollution exposure leads to different types of diseases e.g., lung function changes, heart rate variability, immune cell responses and asthma attacks. New investigations confirmed a massive increase in particulate matter of central congested urban areas in Tehran metropolitan which exceeds the standards of Environmental Protection Agency (EPA, 150 $\mu g/m^3$) and World Health Organization (WHO, 20 $\mu g/m^3$ annual mean and 50 $\mu g/m^3$ 24-hour mean). Long term continuous real-time monitoring of air quality at these areas is essential but is not possible due to financial and operational constraints. Hence, using an alternative tool is important to ensure compliance with the standards and also provides a choice for commuters to reduce their unnecessary trips in contaminated areas across the city. In this study, a stochastic framework is developed based on a Hidden Semi-Markov model (HSMM) to predict the state of PM_{10} particulates. Our proposed HSMM model predicts PM_{10} concentration for the next day, based on PM_{10} levels in previous days. The result of simulation shows the proposed technique achieves good accuracy in estimation of PM_{10}. It also indicates that the model can be used for one-day ahead forecast to alert individuals in the study area which is particularly useful in situations where the information on external variables such as traffic volume is not available.

Key Words**:** Stochastic models, Traffic related PM_{10} pollution, Hidden Semi-Markov

1. INTRODUCTION

PM_{10} is considered as a major air pollutant in metropolitan area. This pollution is commonly associated with combustion sources including traffic, industry and domestic heating (*1-4*). Recent studies showed that PM_{10} is responsible for a number of diseases such as respiratory diseases and well-correlated with daily deaths and hospitalisations. (*5-15*). Kunzli et al. (*9*), investigate the effect of traffic related PM_{10} on the human health such as long term mortality, respiratory and cardiovascular diseases at European countries. They showed that life expectancy is shortened by about 6 months per 10 $\mu g/m^3$ increment in PM_{10}.

Thus, a real time PM_{10} monitoring is essential to detect the exceedances of alarm and attention thresholds. However, when real-time monitoring of air quality is not possible due to financial and operational constraints, predicting PM_{10} concentration base on mathematical/statistical models allows preventative action to be taken in advance if the pollutant levels are in danger. These methods can be used as an important tool for policy makers and epidemiologists to find appropriate strategies for reducing PM10 exposure.

Many attempts have been made to model PM_{10} through the use of meteorological data (*16*, *17*). In most cases, the experimental models (e.g., Neural network, multi-variable regression analysis) are used to estimate PM_{10} variations. For example, there are numerous publications about regression analysis and ANN-based PM_{10} modelling, e.g., Sousa et al. (*18*), Corani (*19*), Perez and Reyes (*20*) and Lu et al. (*21*). However, the ability of ANN models for PM_{10} prediction depends on proper selection of network structure as well as the length of training data, which may not always available. Also, the regression models are developed assuming normally distributed data and may not perform well for highly non-linear problems (*22*). Furthermore, both ANN and Regression methods are the deterministic models which cannot be used for the risk analysis of PM_{10} exposure.

As an appropriate alternative, the hidden Markov model (HMM) has become increasingly popular and quite effective in some applications such as predicting the short term variation of air pollution indices. It has a flexible mathematical structure that can be

applied easily for stochastic simulation of real world problems (*23*, *24*, *25*). Furthermore, it is considered as a conceptual model that makes easy to interpretation (*23*). However, there is some limitations associated with the HMM approach. The Markov process is considered as a short memory process. In the case of short memory process, the current state of the system only depends to the previous state. As a result, the prediction accuracy maybe decreases in the case of long term prediction. To resolve this problem a Hidden Semi-Markov model (HSMM) can be considered for analysis. The difference between HMM and HSMM structures is related to a temporal component that is added to the HSMM structure (*26*, *27*, *28*). In HSMM, the probability of next state of system depends on amount of time that has elapse since the current state.

In this study, we present the Hidden Semi-Markov Model (HSMM) of PM_{10} pollution, which is a stochastic model. In the following sections, the theoretical basis of the Semi-Markov modelling is presented and a general framework is developed for PM_{10} modelling. Then, results of using the proposed model for a real case study are presented. Finally, a number of statistical indices are used to demonstrate the capability of proposed model.

2. THEORY AND METHOD

Following, the mathematical structure of Markov model and the details of proposed HSMM model for PM_{10} simulation are presented.

2.1 HMM Structure (A Brief Description)

A HMM can be presented using a number of parameters, i.e., π, A, B as follows (*25*).

$$\lambda = (\pi, A, B) \quad (1)$$

a) The transition probability distribution function (A) which is presented as

$$A=\{a_{ij}\} \quad (2)$$

$$a_{ij} = P(S_{t+1} = j | S_t = i), \quad 1 \leq i, j \leq N \tag{3}$$

where S_t and S_{t+1} are the states of the system at times t and t+1, respectively, and $P(S_{t+1} = j | S_t = i)$ is the conditional probability.

b) The probability distribution function of observations at each state of the system is shown as:

$$B = \{b_i(k)\} \tag{4}$$

$$b_i(k) = P(v_k | S_k = i), \quad 1 \leq i \leq N, \ 1 \leq k \leq M \tag{5}$$

where v_k is an element of the vector of observation values,

$V = \{ v_1 , v_2 , \ldots , v_m \}$,

N is the number of states in the model, and M is the number of observations at each state of the system.

c) The initial probability value of each state assuming a probability distribution function of $\pi = \{\pi_j\}$ is represented as:

$$\pi_i = P[S_1 = i], \qquad 1 \leq i \leq N \tag{6}$$

Furthermore the probability of the event of staying in sate i for exactly d time units should be determined for the HMM modelling ($P(d)$). Considering the Markovian assumption, $P(d)$ can be determined using the state transition probability distribution function as follows:

$$P_i(d) = a_{ii}^{d-1}(1 - a_{ii}) \tag{7}$$

where a_{ii} is the self-loop probability of state i.

2.2 Semi-Hidden Markov Model

A HSMM model can be presented as follows:

$$\lambda = (\pi, A, B, D) \tag{8}$$

where D is the state duration distribution. Figure 1 shows the definition of parameters for HSMM model. As indicated in Figure 1 state duration distribution is determined based on the transition probability between two states. The transition between two states is defined using the Markov process as follows:

$$P = \left(S_{qr} = j \middle| S_{qr-1} = i\right) = a_{ij} \tag{10}$$

where i and j are the states of system at segment r (Figure 2). Opposed to single observation in a state of HMM, a state in a HSMM generates a segment of observations, i.e., $O_i=\{O_1,..,O_N\}$. Figure 2 shows the relation between segment and states in a HSMM model.

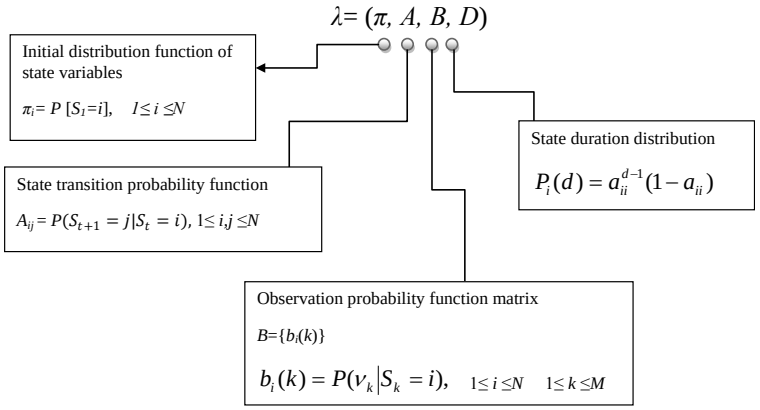

FIGURE 1 HSMM model parameter definition

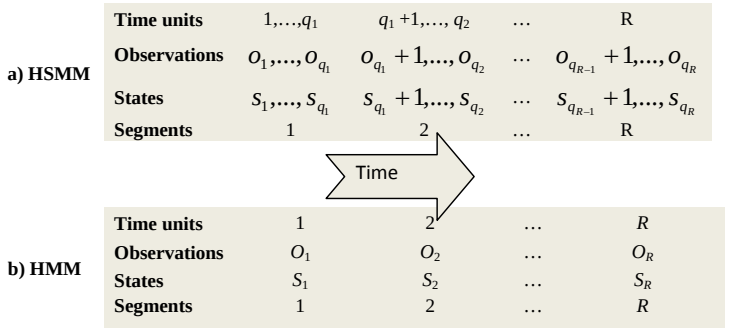

FIGURE 2 Comparing HSMM and HMM model for states and segments

Dong et al. (*22*) developed a *Trial and error* procedure to estimate the parameters of HSMM model. Their proposed method uses a forward backward algorithm to estimate the values of parameters. This method is used to estimate the model parameter in this study. Following, a forward-backward algorithm (*22*) which developed for PM_{10} prediction is introduced.

2.3 HSMM Parameter Estimation (a Forward-Backward Algorithm) for PM_{10} Index

For this purpose, two different levels of PM_{10} were considered as the system states in the current study, i.e., 0-150 μg/m^3 and above 150 μg/m^3. Then, a HSMM model is developed for each individual state. The model parameters are estimated for each model according to the value of highest log-likelihood (i.e. the probability of the observation sequence given a HSMM model). Then, the next PM_{10} state is selected by comparing the highest log-likelihood of each HSMM function. The likelihood function is:

$$P(O_s|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \quad (11)$$

where N is the number of states and i is the state number. α_T is also estimated as follows:

$$\alpha_T(j) = \sum_{i=1}^{N} \sum_{d=1}^{D} \alpha_{t-d}(i) a_{ij} P(d|j) \prod_{s=t-d+1}^{t} b_j(o_s) \quad (12)$$

where D is the maximum retention time at each state, N is the maximum number of states, a_{ij} is the state transition probability matrix and can be determined using Eq 10. The probability of staying in state i for exactly d time steps, $P(d)$, also can be determined using Eq 7, and $b_j(o_s)$ is the observation multivariable probability distribution which is considered as Normal distribution in the current research work. $\alpha_{t-d}(i)$ should be estimated using a forward-backward method (*22*). Following the forward algorithm is discussed.

2.3.1 Forward-Backward Estimation of α_t

Considering the following steps α_t is estimated at two-state variable case.

1- At t=0, the value of $\alpha_{t=0}$ can be estimated as:

$$\alpha_{t=0}(i) = \begin{cases} 1, & \text{if } i = \text{"START"}, \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

2- At the second step, $\alpha_{\acute{t}}$ is determined at time j using Eq 14:

$$\alpha_{\acute{t}}(j) = \sum_{i=1}^{N} \sum_{d=1}^{D} P(d = \acute{t} - t|j) \alpha_{t,\acute{t}}(i,j) \quad (14)$$

where $\acute{t}$=t+d, and d is the retention time at state i, $\alpha_{t,\acute{t}}(i,j)$ can be determined as follows:

$$\alpha_{t,\acute{t}}(i,j) = \alpha_t(i) a_{ij} \emptyset_{t,\acute{t}}(i,j) \quad (15)$$

$$\emptyset_{t,\acute{t}}(i,j) = \sum_{d=1}^{D} \left[P(d = \acute{t} - t|j) . P\left(O_{t+1}^{\acute{t}} \middle| t = q_n, s_t = i, \acute{t} = q_{n+1}, s_{\acute{t}} = j, \lambda\right) \right] \quad (16)$$

where $P\left(O_{t+1}^{\acute{t}} \middle| \ldots\right)$ is the likelihood value that should be estimated using Eq 11. For example for t=0, assuming d=1, we have:

$$P\left(O_1 \middle| t = 0, S = i, t \acute{=} 1, S_{\acute{t}} = j, \lambda\right) \quad (17)$$

Eq 17 can be represented as follows:

$$P(O_1|\lambda) = \sum_{i=1}^{N} \alpha_t(i) = \sum_{i=1}^{2} \alpha_0(i) = \alpha_0(1) + \alpha_0(2) \quad (18)$$

Considering Eq 13, the right hand side of Eq 18 can be determined. A similar procedure can be used to estimate the maximum likelihood for each segment of observations. Similar procedure can be follow up for backward estimation of parameters. More discussion about the backward procedure can be found in Dong et al. (*22*).

2.4 Study area

Tehran Metropolitan area is located at the center of Iran with a population of 8 Million people. Recent investigations show that PM_{10} levels have regularly exceeded critical limit values in Tehran. The Iranian environmental protection agency was the organization charged with collecting the air pollution data (e.g., PM_{10}) through a few air quality monitoring stations (five stations) in central and southern Tehran. Azadi square is a major urban arterial in the center of Tehran and the air quality monitoring site was located near a six-lane highway which is a major road for moving from West to East of Tehran (Figure 3). There is also another three-lane-two-way highway line running close to this monitoring site with an approximate south-west to north-east orientation. In this study, PM_{10} data from Azadi traffic site is collected and used (Figure 3). The data consisted of daily mean values gathered during a 5-year time span between Jan 2005 and Dec 2010. PM_{10} mass concentrations (measured in $\mu g/m^3$) were recorded using Beta Ray attenuation gauging method (APDA-370 Ambient Dust Monitor, HORIBA, Japan). The length of data that is collected and used in this study is presented in Table 1.

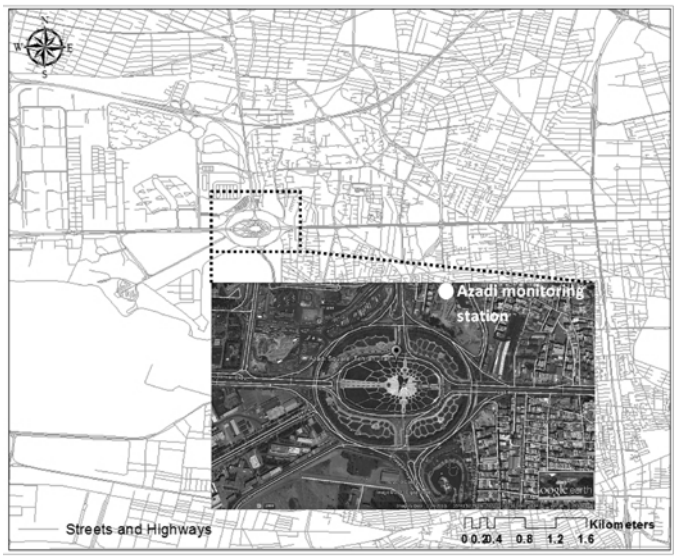

FIGURE 3 Location map for air quality monitoring station in Tehran

TABLE 1 The Length of Data Which is Used to Parameter Estimation and Validation

	Season			
Data record length	Spring	Summer	Fall	Winter
Parameter estimation	339	296	251	351
Validation	21	101	21	21

3. RESULTS AND DISCUSSION

In this study, four seasonal HSMM models are developed to estimate the daily PM_{10} concentration in Tehran metropolitan. The parameters of seasonal models are estimated using the forward-backward algorithm (see section 2.3.1). For this purpose, two states are

considered for the PM_{10} concentration, i.e., 0-150 µg/m^3, and above 150 µg/m^3. The second state may lead to increase different types of diseases e.g., Lung function changes, heart rate variability responses, immune cell responses, Asthma attacks and dead in the critical conditions. As stated earlier, an individual HSMM model should be developed for each PM_{10} state. The maximum log-likelihood (i.e. the probability of an observation sequence given a HSMM model) is estimated for each model using Eq. 11. Then, the next state is predicted by comparing the value of highest log-likelihoods as presented in Figure 4. As indicated in Figure 4, each main state of the proposed model (i.e., Critical and sub-critical levels) has two sub-states, called sates 1 and 2.

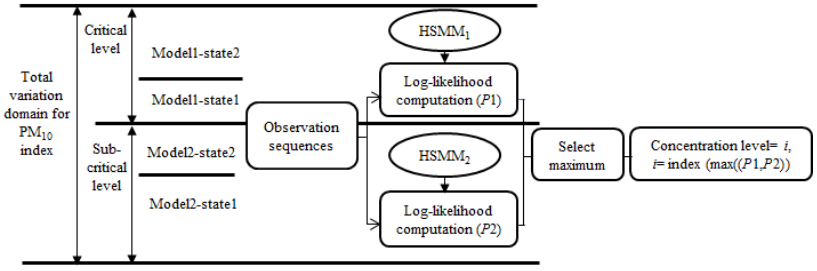

FIGURE 4 The overall structure of HSMM model for daily estimation of PM_{10}

Tables 2 and 3 show the estimated and initial parameters for the models which developed for critical and sub-critical PM_{10} levels. As presented in Tables 2 and 3, the probability of occurrence a day with a critical level of pollution, $P(S = S_1|d = 1)$, is higher for spring season as a comparison with other seasons. However, in the case of sub-critical model, this probability is equal for all seasons. $P(S|d)$ also decreases with duration for both critical and sub-critical models.

TABLE 2 Model Parameters for Critical State (More than 150 µg/m^3)

Variable	Duration	Substate	Season			
			spring	summer	fall	winter
P	1	S1	0.85	0.32	0.27	0.13
		S2	0.08	0.17	0.11	0.05
	2	S1	0.12	0.22	0.20	0.11
		S2	0.08	0.14	0.10	0.05
	3	S1	0.01	0.15	0.14	0.10
		S2	0.07	0.12	0.08	0.05
	4*	S1	0.002	0.10	0.10	0.08
		S2	0.06	0.09	0.08	0.04
D			50	50	50	50
PAI		S1	0.5	0.5	0.5	0.5
		S2	0.5	0.5	0.5	0.5
K			75	129	61	121
Lambda			343.45	274.1	337.85	457.65
M			2	2	2	2
A			$\begin{bmatrix} 0.14 & 0.85 \\ 0.08 & 0.91 \end{bmatrix}$	$\begin{bmatrix} 0.67 & 0.32 \\ 0.17 & 0.82 \end{bmatrix}$	$\begin{bmatrix} 0.72 & 0.27 \\ 0.11 & 0.88 \end{bmatrix}$	$\begin{bmatrix} 0.86 & 0.13 \\ 0.05 & 0.94 \end{bmatrix}$
Maximum Likelihood			-8.79	-61.45	-23.38	-45.67

P: Probability distribution function for each state
D: maximum duration of each state
PAI:The initial occurrence probability value for each state
Lambda: The separation (breakdown) level between two substate
K: Number of data which is used for parameter estimation
M: Number of substates in each state
A: State transition probability matrix

TABLE 3 Model Parameters for Sub-Critical State (Less than 150 µg/m^3)

Variable	Duration	Substate	Season			
			spring	summer	fall	winter
P	1	S1	0.32	0.28	0.37	0.37
		S2	0.07	0.01	0.04	0.10
	2	S1	0.21	0.20	0.23	0.23
		S2	0.07	0.01	0.04	0.09
	3	S1	0.14	0.14	0.14	0.14
		S2	0.06	0.01	0.04	0.08
	4	S1	0.09	0.10	0.09	0.09
		S2	0.06	0.01	0.04	0.07
D			50	50	50	50
PAI		S1	0.5	0.5	0.5	0.5
		S2	0.5	0.5	0.5	0.5
K			226	124	156	187
Lambda			74.65	74.75	73.25	74.70
M			2	2	2	2
A			$\begin{bmatrix} 0.67 & 0.32 \\ 0.07 & 0.92 \end{bmatrix}$	$\begin{bmatrix} 0.71 & 0.28 \\ 0.01 & 0.98 \end{bmatrix}$	$\begin{bmatrix} 0.62 & 0.37 \\ 0.04 & 0.95 \end{bmatrix}$	$\begin{bmatrix} 0.86 & 0.13 \\ 0.05 & 0.94 \end{bmatrix}$
Maximum Likelihood			-58.36	-13.55	-22.50	-55.79

Figures 5 to 8 show the daily observed and simulated states of pollution (PM_{10}) at each season in Azadi square. In these figures, y-axes present the state of pollution (1 and 2 for sub-critical and critical states). As indicated in these figures, the predictions are compared well with observations for all seasonal models. It also shows a dramatic increase of pollutants emission in summer that is mainly due to the highest traffic volume in Azadi square during the summer.

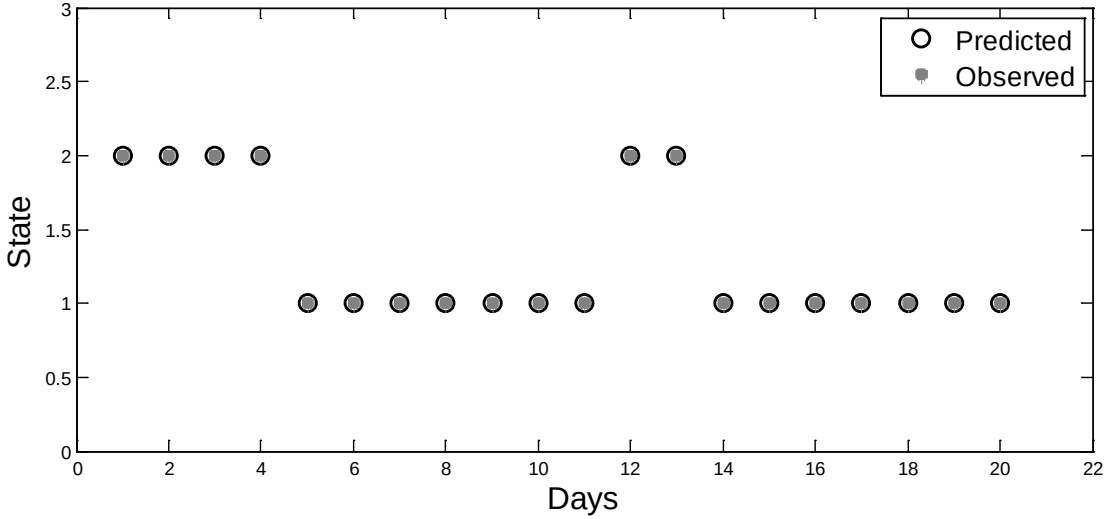

FIGURE 5 The results of PM_{10} level prediction in spring

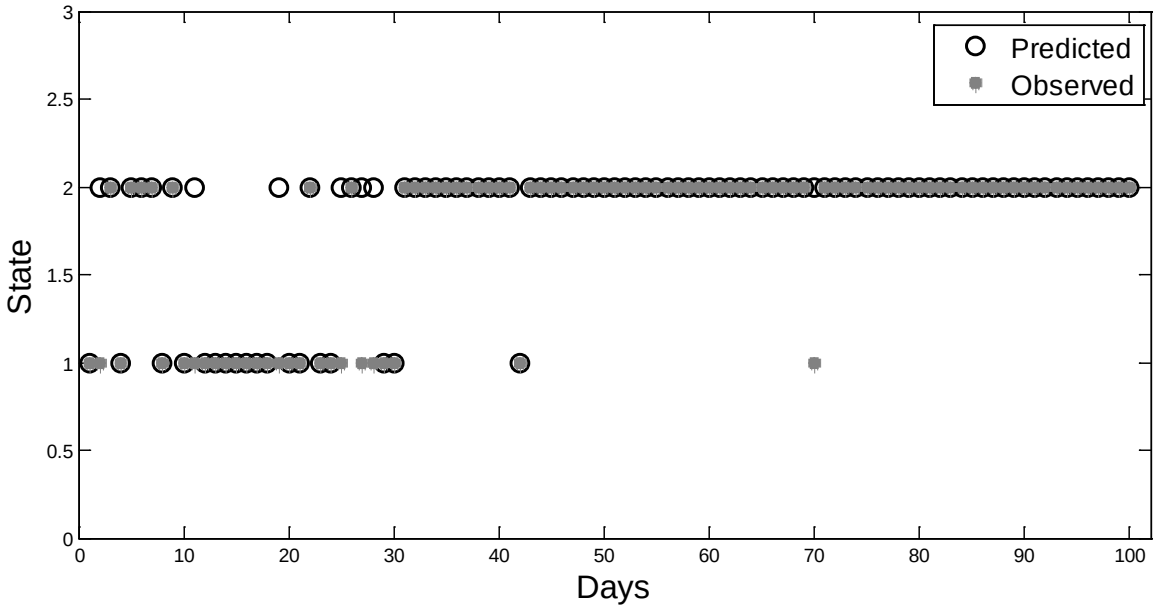

FIGURE 6 The results of PM_{10} level prediction in summer

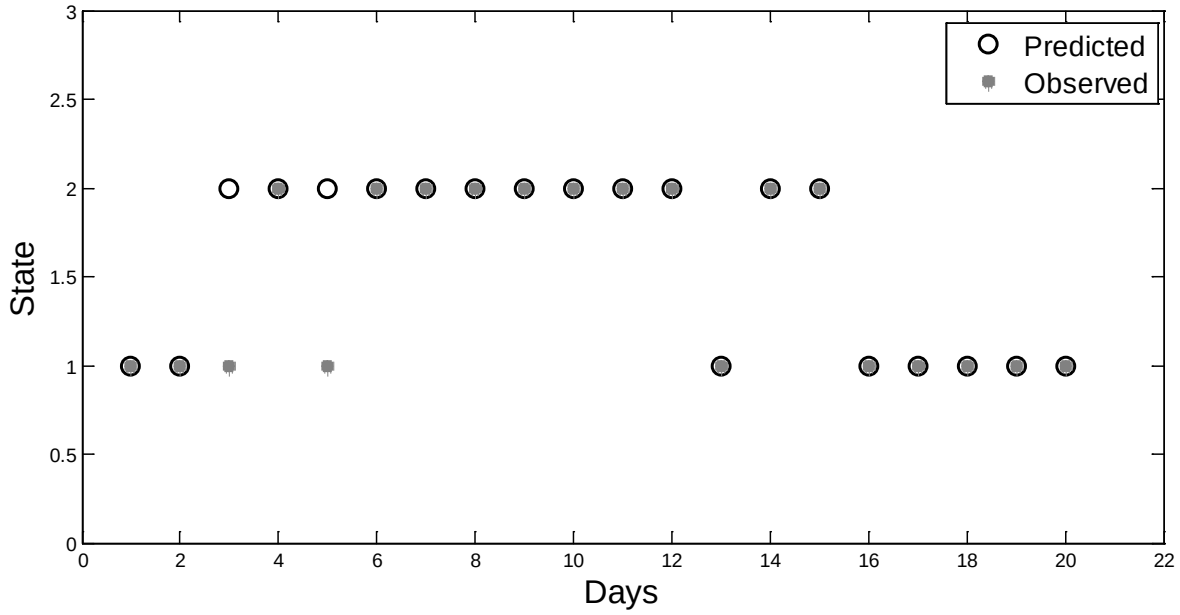

FIGURE 7 The results of PM_{10} level prediction in fall

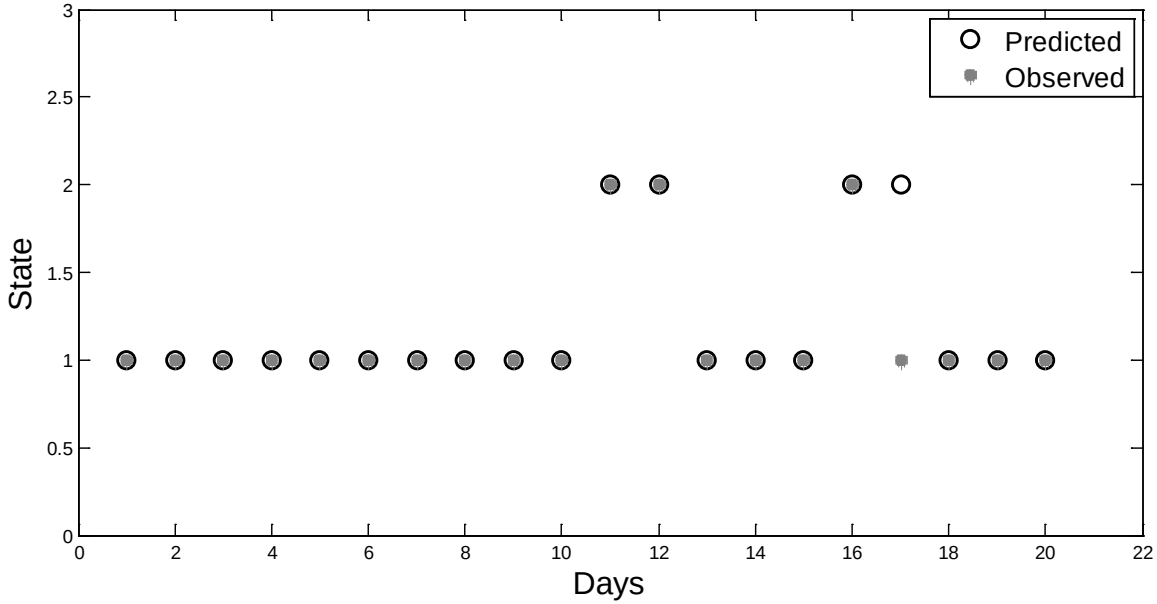

FIGURE 8 The results of PM_{10} level prediction in winter

Also, an index, called precision index (PI) is used to compare the daily observed and simulated states of pollution:

$$PI = 100 - 100 \times \frac{\sum_{i=1}^{n} Abs(S_i - S_i^*)}{\sum_{i=1}^{n} S_i} \quad (19)$$

where s_i^* and s_i are the daily simulated and observed states of pollution. The value of precision index varies between 0 which reflects complete difference and 100% which indicates complete equality. Table 4 shows the results for precision index. As indicated in this table the predicted states compared well with those obtained from measurements. This table also shows the precent of days per year that PM_{10} concentration is higher than a critical level of pollution. As indicated, summer has the highest level of exposure and model works well for this season.

Table 4 PI index for simulated PM_{10} Level

	Precision (%)			
	spring	summer	fall	Winter
Semi-Markov	85	84	75	75
PM_{10} exposure (%)	23	45.6	25.5	36.7

4. SUMMARY AND CONCLUSION

Traffic-based PM_{10} pollution has been a major concern in Tehran for recent years. In this study, the hidden semi-Markov model (HSMM) is used to predict the state of PM_{10} exposure. The proposed model predicts the state of pollution one day before, based on PM_{10} level in the previous day at each season of year. The capability of model for PM_{10} prediction was checked in Tehran metropolitan area. The results showed that the modelling outputs are compared well with observed values.

It was also indicated that the proposed model can be used as an appropriate tool for one-day ahead alert program in metropolitan areas. It provides a choice for commuters to reduce their unnecessary trips in contaminated areas across the city.

REFERENCES

1. Wilson, J. G., S. Kingham, J. Pearce, and A. P. Sturman. A Review of Intraurban Variations in Particulate Air Pollution: Implications for Epidemiological Research. *Atmospheric Environment*, Vol. 39, No. 34, 2005, pp. 6444-6462.
2. Sun, Y. L., G. S. Zhuang, W. Ying, L. H. Han, J. H. Guo, D. Mo, W. J. Zhang, Z. F. Wang, and Z. P. Hao. The Air-Borne Particulate Pollution in Beijing—Concentration, Composition, Distribution and Sources. *Atmospheric Environment*, Vol. 38, 2004, pp. 5991–6004.
3. Janssen, N. A. H., P. H. N. Van Vliet, F. Aarts, H. Harssema, and B. Brunekreef. Assessment of exposure to Traffic Related Air Pollution of Children Attending Schools Near Motorways. *Atmospheric Environment*, Vol. 35, 2001, pp. 3875–3884.
4. Janssen, N. A. H., G. Hoek, B. Brunekreef, H. Harssema, I. Mensink, and A. Zuidhof. Personal Sampling of Particles in Adults: Relation among Personal, Indoor, and Outdoor Air Concentrations. *American Journal of Epidemiology*, Vol. 147, 1998, pp. 537–547.
5. Donaldson, K., W. MacNee. *The Mechanism of Lung Injury Caused By Pm10 in Air Pollution and Health*. In: Hester, R.E., Harrison, R.M. (Eds.), Issue in Environmental Science and Technology. Royal Society of Chemistry, Reedwood Books Ltd., Trowbridge, Wiltshire, UK, 1999.
6. Schwartz, J., D. W. Dochery, L. M. Neas. Is daily mortality associated specifically with fine particles? *Journal of Air and Waste Management Association*, Vol. 46, 1996, pp. 927–939.
7. Dockery, D., A. Pope. *Epidemiology of Acute Health Effects Summary of Time-Series Studies.*R Wilson, J.D Spengler (Eds.), Particles in Our Air: Concentration and Health Effects, Harvard University Press, Cambridge, MA, USA,1996, pp. 123–147.

8. Querol, X., A. Alastuey, S. Rodriguez, F. Plana, C. R. Ruiz, N. Cots, G. Massague, and O. Puig. PM10 and PM2.5 Source Apportionment in the Barcelona Metropolitan Area, Catalonia, Spain. *Atmospheric Environment*, Vol. 35, No. 36 , 2001, pp. 6407–6419.
9. Kunzli, N., R Kaiser, S Medina, M. Studnicka, O. Chanel, P. Filliger, M. Herry, F. Horak Jr, V. Puybonnieux-Texier, P. Quénel, J. Schneider, R. Seethaler, J-C. Vergnaud, and H. Sommer. Public-Health Impact of Outdoor and Traffic-Related Air Pollution: A European Assessment. *Lancet*, Vol. 356, 2000, pp. 795–801.
10. Roorda-Knape, M. C., N. A. H. Janssen, J. J. de Hartog, H. Harssema, and B. Brunekreef. Air Pollution from Traffic Near Major Motorways. *Atmospheric Environment*, Vol. 32, 1998, pp. 1921–1930.
11. Kim, J. J., S. Smorodinsky, M. Lipsett, B. C. Singer, A. T. Hodgson, and B. Ostro. Traffic-Related Air Pollution Near Busy Roads: the East Bay Children's Respiratory Health Study. *American Journal of Respiratory and Critical Care Medicine*, Vol. 170, 2004, pp. 520–526.
12. Hitchins, J., L. Morawska, R. Wolff, and D. Gilbert. Concentrations of Submicrometer Particles from Vehicle Emissions Near a Major Road. *Atmospheric Environment*, Vol. 34, 2000, pp. 51–59.
13. Lamoree, D. P., and J. R. Turner. PM Emissions Emanating from Limited-Access Highways. *Journal of the Air and Waste Management Association*, Vol. 49, 1999, pp. 85–94.
14. Kingham, S., D. Briggs, P. Elliott, P. Fischer, and E. Lebret. Spatial Variations In The Concentrations of Traffic Related Pollutants in Indoor And Outdoor Air in Huddersfield, England. *Atmospheric Environment*, Vol. 34, 2000, pp. 905–916.
15. Guo, Y. L., Y. C. LinC, F. C. Sung , S. L. Huang, Y. C. Ko, J. S. Lai, H. J. Su, C. K. Shaw, R. S. Lin, and D. W. Dockery. Climate, Traffic Related Air Pollutants, and Asthma Prevalence in Middle-School Children in Taiwan. *Environmental Health Perspectives*, Vol. 107, 1999, pp. 1001–1006.
16. Shekarrizfard, M, A. Karimi-Jashni, and K. Hadad. Wavelet transform-Based Artificial Neural Networks (WT-ANN) in

PM10 Pollution Level Estimation, Based On Circular Variables. *Environmental Science and Pollution Research*, Vol. 19, 2011, pp. 256–268.
17. Hooyberghs, J., C. Mensink , G. Dumont, F. Fierens, and O. Brasseur. A Neural Network Forecast for Daily Average PM_{10} Concentrations in Belgium. *Atmospheric Environment*, Vol. 39, 2005, 3279–3289.
18. Sousa, S. I. V., F. G. Martins, M. C. M. Alvim-Ferraz, and M. C. Pereira. Multiple Linear Regression and Artificial Neural Networks Based on Principal Components to Predict Ozone Concentrations. *Environmental Modelling and Software*, Vol. 22, 2007, pp. 97–103.
19. Corani G. Air Quality Prediction In Milan: Feed-Forward Neural Networks, Pruned Neural Networks and Lazy Learning. *Ecological Modelling*, Vol. 185(2–4), 2005, pp. 513–529.
20. Perez, P., and J. Reyes. Prediction of Maximum of 24-h Average Of PM10 Concentrations 30h in Advance in Santiago, Chile. *Atmospheric Environment*, Vol. 36, 2002, pp. 4555–4561.
21. Lu W. Z., W. J. Wang, H. Y. Fan, A. Y. T. Leung, Z. B. Xu, S. M. Lo, J. C. K. Wong. Prediction of Pollutant Levels in Causeway Bay Area of Hong Kong Using an Improved Neural Network Model. *Journal of Environmental Engineering*, Vol. 128, 2002, pp. 1146–1157.
22. Dong M., D. Yang, Y. Kuang, D. He, S. Erdal, and D. Kenski. PM2.5 Concentration Prediction Using Hidden Semi-Markov Model-Based Times Series Data Mining. *Expert Systems with Applications*, Vol. 36, 2009, 9046–9055.
23. Baruah, P., Chinnam, R. B. *HMMs for diagnostics and prognostics in machining processes.* In: Proc. of the 57th Society for Machine Failure Prevention Technology Conference, Virginia Beach, VA, 2003, pp. 14–18.
24. Eddy, S. R. Hidden Markov models. *Current opinion in structural biology*, Vol. 6, 1996, pp. 361–365.
25. Rabiner, L. R. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, Vol. 77, No. 2, 1989, pp. 257–286.
26. Dong, M., and D. He. Hidden Semi-Markov Model Based Methodology for Multi-Sensor Equipment Health Diagnosis and

Prognosis. *European Journal of Operational Research* Vol. 178, No. 3, 2007a, 858–878.

27. Dong, M., and D. He. A Segmental Hidden Semi-Markov Model (HSMM)-Based Diagnostics and Prognostics Framework and Methodology. *Mechanical Systems and Signal Processing*, Vol. 21, No. 5, 2007b, pp. 2248–2266.
28. Mirakbari, M., A. Ganji. Reliability Analysis of a Rangeland System: The Application of Profust Theory. *Stochastic Environmental Research and Risk Assessment,* Vol. 24, No. 3, 2010, pp. 399–409.