Chapter 6

Study of Extreme PM Exposure at Traffic Intersections

This chapter examines the impact of seasonal variations on extreme PM exposure at traffic intersections in the study area. It also discusses the probability of PM level exceedance observed during one of the Delhi smog events.

6.1 Background

Commuting causes significantly high exposure to air pollution (Singh et al., 2021). Vehicular emissions and movements at the intersections turn them into pollution hotspots (Choudhary and Gokhale, 2019; Goel and Kumar, 2014, 2015). High concentrations of pollutants are emitted at the intersections when vehicles idle during the red light phase or start accelerating when the red light turns green (Pandian et al., 2009; Wang et al., 2015). In addition, the resuspension of road dust by vehicular movement significantly contributes to PM (Sinha and Dammani, 2018). Therefore, commuters (motorists, cyclists, and pedestrians) are exposed to high PM levels, especially at traffic intersections. This is because they spend significantly more time per km at traffic intersections. In addition, a significantly greater number of vehicular activities occur at the intersections as they decelerate, accelerate, and interact with crossing traffic. This results in as much as 17%, 16%, and 40% higher concentrations of PM_1, $PM_{2.5}$, and PM_{10}, respectively, as compared to the rest of the road section (Kumar et al., 2017). At signalized intersections, PM concentration

during the red phase may be as high as 3.5 times that of the green phase (Wang et al., 2008). Furthermore, road users only spend 2% of their commute time at intersections but are exposed to 25% of the total PM concentration (Goel and Kumar, 2015). Therefore, examining commuters' exposure to extreme PM at traffic intersections is important.

6.1.1 Extreme Exposure

Estimating extreme PM concentrations is necessary to determine the likelihood of severe pollution events in a city. There were several instances around the world when the concentration of pollutants peaked, resulting in emergencies such as increased hospital admissions, deaths, and the closing of schools, offices, and other non-essential activities. The London smog of 1952 was one of the most severe pollution episodes. There was a significant increase in fatalities and heightened hospital admissions, especially for those with respiratory and cardiovascular ailments (Bell et al., 2004; Nemery et al., 2001). Besides that, the smog events at Meuse Valley (Belgium) in December 1930 and at Donora (Pennsylvania, USA) in October 1948 were also among the most lethal pollution events (Logan, 1953; Nemery et al., 2001). In recent years, similar severe smog events have been observed in Delhi (India) (Garg and Gupta, 2020; Kanawade et al., 2020; Sati and Mohan, 2014; Sawlani et al., 2019). During these events, the pollutant level significantly exceeded the upper limits ($PM_{2.5}$: 250 µg m^{-3} and PM_{10}: 430 µg m^{-3}) of NAAQS set by the CPCB, India. Severe exposure levels (585 µg m^{-3} of $PM_{2.5}$ and 989 µg m^{-3} of PM_{10}) were also observed in Delhi in November 2012. Likewise, in November 2016, Kanawade et al. (2020) observed a $PM_{2.5}$ concentration of 563 µg m^{-3}. During the same year (30 October – 07 November), Sawlani et al. (2019) reported a $PM_{2.5}$ concentration of 793 µg m^{-3} in the city. Similarly, in November 2017, Garg and Gupta (2020) observed severe $PM_{2.5}$ and PM_{10} levels of 374 and 768 µg m^{-3}, respectively.

Table 6.1 Summary of PM extreme value analyses conducted at various cities around the world.

Study	City/Region	N	Duration	Pollutant	Aggregation	Best-Fit Model
Lu, 2002	Central Taiwan	3	Jan 1995 – Jun 2000	PM_{10}	Daily	Lognormal
Lu and Fang, 2003	Central Taiwan	5	1995 – 1999	PM_{10}	Monthly	Two-Parameter Exponential
Lu, 2004	Central Taiwan	5	Jan 1994 – Dec 1999	PM_{10}	Daily	Two-Parameter Exponential
Mijić et al., 2009	Belgrade, Serbia	1	2003 – 2005	PM_{10}	Daily/Monthly	Asymptotic Extreme Value
Sharma et al., 2013	Delhi, India	6	Jan 2003 – Dec 2006	PM_{10}	Daily	Lognormal and Gamma
Qin et al., 2015	Jing-Jin-Ji, China	13	Feb 2014 – Feb 2015	PM_{10}	Hourly	Apriori Algorithm
Martins et al., 2017	Sao Paulo and Rio de Janeiro, Brazil	16	1996 – 2011	$PM_{2.5}$, PM_{10}	Hourly	Generalized Extreme Value
Plocoste et al., 2020	Guadeloupe, France	2	2005 – 2017	$PM_{2.5}$, PM_{10}	Daily	Burr and Weibull Mixture Model
Salado et al., 2022	Mexico, North America	23	2003 – 2021	PM_{10}	Hourly	Generalized Extreme Value
Sharma et al., 1999	Delhi, India	1	July 1997 – Jun 1998	CO	Hourly/8-Hourly	Gumbel
Goel and Kumar, 2016	Surrey, UK	2	Jan – April 2015	PNC	Daily	GEV
Gulia et al., 2017a	Delhi, India	2	Winter/Summer 2010	$PM_{2.5}$	Hourly	Lognormal
Gulia et al., 2017b	Delhi, India	2	Dec 2009 – May 2010	$PM_{2.5}$	Hourly	Lognormal

Note: N = Number of monitoring locations.

In previous literature, average PM concentrations at traffic intersections have been investigated extensively (Adeniran et al., 2017; Kumar and Goel, 2016; Piotrowicz and Polednik, 2019; Song et al., 2020; Wang et al., 2017). These studies examined the average concentrations but did not account for high pollutant levels. Further, a moderate number of studies on extreme pollution exposure have been conducted at different urban locations throughout the world (Table 6.1). In these studies, the researchers fitted various probability distributions to PM and provided insights into extreme pollutant levels. Only a few of these studies were conducted at traffic intersections. Most of these studies focused on estimating daily maximum PM_{10} concentrations. However, only a smaller fraction of them considered seasonal variation of PM extremes. The probability of extreme PM events was higher during the winter (Gulia et al., 2017; Sharma et al., 2013). Thus, the prediction of seasonal extremes can be helpful to policymakers who can timely initiate mitigation measures to control extreme pollution events.

Generally, conventional air quality models predict extreme pollutant concentrations inaccurately (Gulia et al., 2017). Lesar and Filipčić (2021) simulated the hourly concentration of $PM_{2.5}$ but underestimated the extreme value. The traditional lognormal frequency distribution does not predict extreme PM_{10} concentrations well (Lu, 2004). Furthermore, most deterministic models cannot accurately forecast the complete spectrum of pollutant concentrations due to their inherent non-stochastic characteristics (Gokhale and Khare, 2004). The methods based on extreme value theory (EVT) are more appropriate for modeling high pollutant concentrations (Barakat et al., 2020; Fix et al., 2018).

6.1.2 Bayesian Hierarchical Modeling

Heterogeneity in data can be modeled more appropriately using hierarchical models. A heterogeneous population has several sub-groups, and the observations from such populations are only exchangeable within each sub-group. The Bayesian modeling approach has

been used previously to model extreme events. For example, Amin et al. (2015) predicted high PM_{10} levels in Johar (Pakistan) using 1-year daily maximum PM_{10} data collected from three air quality monitoring stations. Fioravanti et al. (2021) and Ramli et al. (2023) predicted the next-day PM_{10} concentration in Rome (Italy) and Peninsular (Malaysia), respectively. Saez and Barceló (2022) and Pirani et al. (2014) employed the Bayesian approach for finding short-term PM_{10} exposure in Greater London (UK) and Catalonia (Spain), respectively . The Bayesian hierarchical modeling approach consistently produced improved PM exposure models, whether they pertained to predicting high, short-term, or next-day concentrations.

The present study aims to investigate extreme PM exposure at various traffic intersections. The main objectives are

1. to investigate how extreme PM exposure varies with seasons,
2. to estimate the monthly return level of PM concentration, and
3. to estimate the risk of a severe pollution event in a densely populated city.

6.2 Data Analysis

6.2.1 Zone of Influence (ZoI)

PM measures were synced with location data recorded by the GPS. PM data pertaining to a particular intersection was identified by defining a zone of influence (ZoI) around the intersection. ZoI is the region of road around an intersection where traffic speed is generally lower than the rest of the road segment. Vehicles slow down, stop (usually) and accelerate within the ZoI, resulting in prolonged tailpipe emissions, wearing of tires and resuspension of road dust. The PM level in the ZoI is likely to be high. In addition, a commuter is exposed to it even for a longer time due to considerably lower speed.

The PM data within ZoI was extracted for each intersection as illustrated in the flow chart (Fig. 6.1). At first, the PM data was synced with the trajectories (GPS data) and speed at each location was computed. Additionally, a linear referencing system (LRS) was defined for each route. The LRS split the route into equally spaced segments of 5 m each. All trajectories were projected to the LRS line and the truncated mean (with outliers removed) at each segment was computed to mitigate temporal fluctuations in speed. In this way, the mean speed profiles for each route were computed. Now, for an intersection, the mean speed profiles in both the u/s and d/s directions were analyzed to find the speed change points. The u/s portion of the ZoI was identified as the region within which the speed started decreasing and came to a complete halt at the intersection. Likewise, the d/s portion of the ZoI was identified as the region within which the speed started increasing at the intersection and became steady. A typical ZoI for an intersection is presented in Fig. 6.2. Kumar and Goel (2016) discovered that the average ZoI for $PM_{2.5}$ and PM_{10} at different traffic intersections ranged from 55 – 121 and 65 – 125 m, respectively. Wind speed and road geometry influence the dispersion of $PM_{2.5}$ and PM_{10} differently because of the difference in weights. This could be a reason why the two ZoI were slightly different.

The final dataset consisted of the date, time, time interval, month, season, latitude, longitude, route number, intersection number, $PM_{2.5}$ concentration, and PM_{10} concentration, with each row representing second-by-second information. This dataset was explored further to model extreme PM values and map them.

6.2.2 Analysis of Extreme Levels of PM

Extreme value analysis is a statistical modeling approach that can forecast the likelihood of occurrence of extreme events within a time period (Beirlant et al., 2004; Coles et al., 2001). Extreme observations can be sampled by either block maxima (BM) approach or peaks-over-threshold (POT) approach. The samples obtained from BM and POT approaches

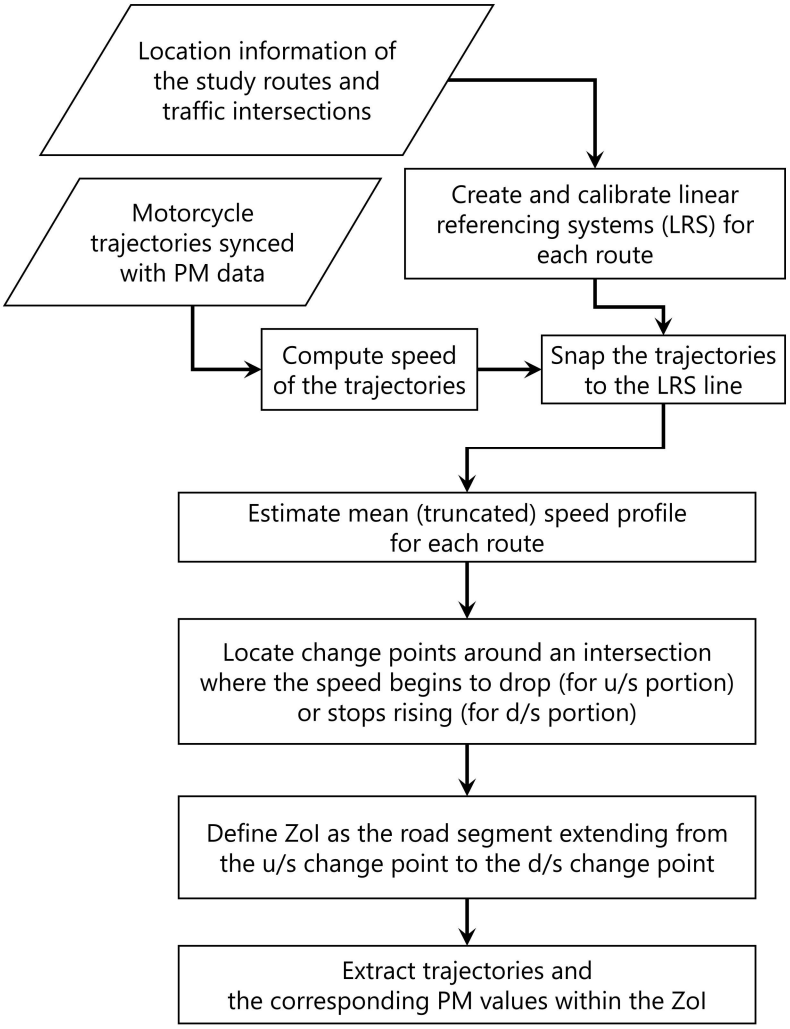


Fig. 6.1 Computation of zone of influence (ZoI) and extraction of intersection PM data.

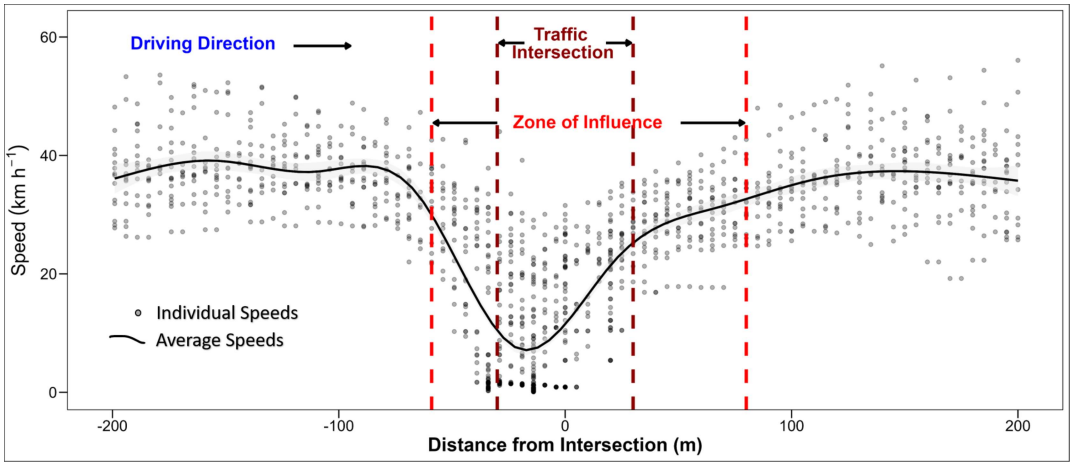


Fig. 6.2 Identification of the Zone of Influence (ZoI) at an Intersection.

are modeled using generalized extreme value (GEV) and generalized Pareto distribution (GPD), respectively. In the BM approach, the dataset is divided into several blocks, and the maximum value in each block is chosen. On the other hand, in the POT approach, all extreme observations exceeding a pre-defined threshold are utilized. Previous studies identified a few issues with the BM approach. Firstly, selecting the block interval is subjective. Secondly, selecting only the maximum value will imply discarding other extreme values and thereby reducing the number of observations available for modeling. Therefore, researchers have preferred the POT approach for modeling extreme events (Acero et al., 2011; Coles et al., 2001; Lazoglou et al., 2019; Roth et al., 2014). In the present study, the POT approach was utilized to model extreme PM concentrations at traffic intersections in the city for various seasons. The details regarding the POT approach are described in the following section.

The Peak-Over Threshold Approach (POT)

Pickands (1975) proposed the POT approach to model extreme observations. Let $Y_1, Y_2, Y_3, \ldots, Y_n$ denote a series of independent random variables with similar distribution function F. For large n, the sample maxima, $M_n = \max\{Y_1, Y_2, Y_3, \ldots, Y_n\}$ follows a GEV distribution

with μ, $\sigma > 0$, and ξ as location, scale, and shape parameters, respectively. The GEV distribution is given by

$$G(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \tag{6.1}$$

Further, for a large enough threshold u, the exceedances, $y = Y - u$, for $y > 0$ and $\left(1+\frac{\xi y}{\tilde{\sigma}}\right) > 0$, will approximately follow a Generalized Pareto Distribution (GPD) given by

$$H(y) = 1 - \left(1+\frac{\xi y}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}} \tag{6.2}$$

where ξ and $\tilde{\sigma}$ are the shape and scale parameters, respectively. The scale parameter $\tilde{\sigma}$ is given by

$$\tilde{\sigma} = \sigma + \xi(u-\mu) \tag{6.3}$$

The GPD belongs to a family of continuous probability distributions. It is used to model the tail or extreme data as a conditional distribution to a threshold value (u). The data points beyond the threshold value are considered for extreme value modeling. GPD is a two-parametric distribution, where ξ is the shape parameter, and σ is the scale parameter. The value of ξ can be zero, positive, or negative.

Threshold Selection

The selection of an appropriate threshold is critical in extreme value modeling. Too low a threshold violates the model's asymptotic basis, leading to bias in model parameters, while a very high threshold leaves fewer data points, leading to high variance. In univariate EVT models, the threshold stability plot and the mean residual plot are generally used to select a suitable threshold, u. The former is less subjective than the later and therefore used in

the present study (Coles et al., 2001). For a random variable Y, if the threshold exceedance $(Y-u>0)$ follows a Generalized Pareto Distribution (GPD), then $Y-u^*$ (for $u^*>u$) also follows a GPD such that the estimates of the shape and scale parameters remain stable up to a certain threshold value (Coles et al., 2001). For higher values than the threshold, a large variance in the parameters is introduced due to a decrease in sample size. The result is that the parameters are no longer stable. The estimates of shape and scale parameters would be stable for two overlapping ranges of threshold values. The chosen threshold is the minimum value common to both ranges.

Return Level

The return level y_m represents the value exceeded on average once every m observations. If a random variable Y follows a Generalized Pareto Distribution (GPD) with parameters σ and ξ and threshold u, then for $y>u$,

$$P(Y>y \mid Y>u)=\left[1+\xi\left(\frac{y-u}{\sigma}\right)\right]^{-\frac{1}{\xi}} \tag{6.4}$$

This can be simplified as

$$P(Y>y)=\zeta_u\left[1+\xi\left(\frac{y-u}{\sigma}\right)\right]^{-\frac{1}{\xi}} \tag{6.5}$$

where $\zeta_u=P(Y>u)$. The return level y_m can be expressed as

$$\zeta_u\left[1+\xi\left(\frac{y_m-u}{\sigma}\right)\right]^{-\frac{1}{\xi}}=\frac{1}{m} \tag{6.6}$$

This can be rewritten as

$$y_m=u+\frac{\sigma}{\xi}\left[(m\zeta_u)^{\xi}-1\right] \tag{6.7}$$

The probability of occurrence of extreme pollution levels was computed based on Eq. 6.5, while the return level was computed using Eq. 6.7.

Bayesian Hierarchical Model Structure

The Bayesian hierarchical or multilevel modeling combines data from groups with similar characteristics and allows the model parameters to vary across sub-groups. The multilevel modeling framework can address data scarcity issues in extreme value modeling, integrating data from all groups (Coles et al., 2001). Further, the Bayesian analysis provides insights into the distributions (known as posterior distributions) rather than point estimates of model parameters. Thus, the analysis provides a more complete understanding of uncertainties in parameter estimation and measurements. Bayesian hierarchical models can help incorporate information from previous studies in the form of prior distributions.

In the present study, three levels of hierarchy were used to model the PM data using the Bayesian hierarchical GPD model. The three levels in the model are as follows.

$$\text{Data layer: } \begin{cases} G(y_i < z \mid \sigma_i, \xi_i) = 1 - \left[1 + \xi_i \frac{y_i}{\sigma_i}\right]^{-1/\xi_i} \end{cases} \tag{6.8}$$

Where y_i is threshold excess from season i.

In the second layer, variations in model parameters to account for seasonal variability were incorporated.

$$\text{Process layer: } \begin{cases} \sigma_i \sim N(\sigma_\mu, \sigma_\sigma) \\ \xi_i = \xi \end{cases} \tag{6.9}$$

In the third layer, the distribution of model parameters was defined based on previous studies.

$$\text{Prior layer:} \begin{cases} \sigma_\mu \sim N(0, 100) \\ \sigma_\sigma \sim N(0, 100) \\ \xi \sim \text{Uniform}(-2, 2) \end{cases} \tag{6.10}$$

The posterior distribution combines all these three levels $(1-3)$ as

$$p(\theta \mid y) \propto \underbrace{p(y \mid \theta)}_{\text{Data}} \underbrace{p(\theta_1 \mid \theta_2)}_{\text{Process}} \underbrace{p(\theta)}_{\text{Prior}} \tag{6.11}$$

Where the posterior $p(\theta \mid y)$ is the distribution of the model parameters, $p(y \mid \theta)$ is the likelihood function, $p(\theta_1 \mid \theta_2)$ is the process model, and $p(\theta)$ is the prior distribution.

Model Validation and Posterior Predictive Checking

The posterior distributions for each model parameter were estimated using a Markov Chain Monto Carlo (MCMC) algorithm called No-U-Turn-Sampler. Samples generated from MCMC algorithms are auto-correlated, hence biased. If sampling is done for a large number of iterations, this correlation becomes negligible. Additionally, multiple chains are initiated to check this bias. Each chain is initialized at a different location, and their convergence is observed. One method of convergence is to check the potential scale reduction factor (R hat). If the chains mix well, the value of R hat will be 1 (Vehtari et al., 2021). A trace plot is used to inspect the convergence of chains sampled from the posterior distribution. A good mixing indicates convergence of the model. The current study used both R hat and trace plots for model validation. After ensuring that the sample had little to no autocorrelation, the goodness of model fit was assessed using a posterior predictive check, which involves comparing the simulated data with the observed data (Gelman and Stern, 1996).

6.2.3 Exposure Hotspot Maps

Pollution hotspot maps are useful in planning mitigation strategies for reducing exposure to extreme pollution concentrations. This study aimed to find the PM exposure hotspots around traffic intersections in Varanasi during different seasons. These exposure hotspot maps were created to identify intersections with elevated PM levels. For this, the PM concentrations exceeding the threshold values for each traffic intersection were selected and averaged to obtain mean extremes. These mean extremes were depicted on the hotspot maps.

6.3 Results and Discussion

6.3.1 Descriptive Statistics

Table 6.2 and 6.3 summarizes the $PM_{2.5}$ and PM_{10} concentrations around traffic intersections and mid-block sections for various seasons in Varanasi. The PM concentrations at traffic intersections were generally higher than mid-block sections. However, the difference in pollutant levels between intersection and mid-block sections was not significantly different. As described in Section 6.1, commuters spend more time at traffic intersections compared to any part of the mid-block sections. Thus, due to longer time spent at traffic intersections, the cumulative exposure concentrations are higher compared to mid-block sections. In winter, the average $PM_{2.5}$ and PM_{10} around traffic intersections were 218 and 345 $\mu g\ m^{-3}$, respectively. During the summer, the average values of $PM_{2.5}$ and PM_{10} were 98 and 219 $\mu g\ m^{-3}$, respectively. Overall, the 97.5^{th} percentile $PM_{2.5}$ and PM_{10} concentrations were 443 and 741 $\mu g\ m^{-3}$, respectively. PM concentration was highest in the winter, followed by spring and summer.

Table 6.2 Seasonal $PM_{2.5}$ concentration (μg m^{-3}) at traffic intersections and mid-block sections in Varanasi.

	Intersections		Mid-block sections	
Season	**Mean (SD)**	**P2.5 – P97.5**	**Mean (SD)**	**P2.5 – P97.5**
Summer	98 (60)	31 – 252	96 (57)	38 – 243
Spring	134 (75)	39 – 326	128 (73)	42 – 290
Winter	218 (112)	71 – 487	212 (112)	64 – 452
All Seasons	157 (105)	38 – 443	153 (103)	40 – 410

Abbreviations: SD, standard deviation; P2.5, 2.5th percentile; P97.5, 97.5th percentile.

Table 6.3 Seasonal PM_{10} concentration (μg m^{-3}) at traffic intersections and mid-block sections in Varanasi.

	Intersections		Mid-block sections	
Season	**Mean (SD)**	**P2.5 – P97.5**	**Mean (SD)**	**P2.5 – P97.5**
Summer	219 (134)	60 – 575	214 (131)	59 – 526
Spring	242 (138)	73 – 614	237 (139)	70 – 562
Winter	345 (163)	115 – 741	341 (165)	105 – 675
All Seasons	278 (159)	72 – 681	274 (160)	70 – 620

Abbreviations: SD, standard deviation; P2.5, 2.5th percentile; P97.5, 97.5th percentile.

The PM values estimated at various traffic intersections in Varanasi were compared with studies conducted in different Indian metropolitan cities such as Mumbai (Gokhale and Patil, 2004) and Guwahati (Gokhale and Raokhande, 2008). Both of those studies were conducted in the winter. In Mumbai (India), the PM_{10} value (528 μg m^{-3}) at an intersection affected by industrial and substantial commercial activities was more than the intersections studied in Varanasi, where there are no major industries. However, the PM_{10} value (258 μg m^{-3}) at an intersection (Mumbai) affected only by vehicular traffic was lower than the value (345 μg m^{-3}) observed in the Varanasi intersections. PM_{10} (158 μg m^{-3}) during the winter at one of the busiest traffic intersections in Guwahati (India) was lower than the study location (Gokhale and Raokhande, 2008). This can be attributed to the significant difference in the population density of both the cities. Population density directly affects the number of vehicular activities and hence, the city's pollutant level (Amirjamshidi et al., 2013). The population density of Guwahati (2,700 per km^2 at the time of data collection) is significantly lower than Varanasi (21,000 per km^2). The concentration of fine particles during both peak hours (77.9 μg m^{-3}) and off-peak hours (62.7 μg m^{-3}) at six major traffic intersections in the Ilorin metropolis (Nigeria) was lower than what was observed in the present study. On the contrary, the concentration of coarse PM during peak hour (513 μg m^{-3}) and off-peak hour (390 μg m^{-3}) was significantly higher than the values observed in the present study (Adeniran et al., 2017). High concentration of coarse PM in Ilorin (Nigeria) could be attributed to an increased number of trucks and lorries engaged in commercial or industrial activities, leading to greater resuspension of particles. Apart from land use and population density, the pollution level is also influenced by meteorological conditions, vehicular composition, and environmental policies (Han et al., 2024; Yang et al., 2024).

The CPCB, India categorized $PM_{2.5}$ (PM_{10}) concentration (in μg m^{-3}) ranges from 0 to 30 (0 to 50) as good, from 31 to 60 (51 to 100) as satisfactory, from 61 to 90 (101 to 250) as

moderate, from 91 to 120 (251 to 350) as poor, from 121 to 250 (351 to 430) as very poor and more than 250 (430) to be severe (CPCB, 2015). According to this classification, the air quality in the study locations was very poor in winter and spring, and poor in summer. This level of pollution puts people with lung and heart diseases at high risk of serious health conditions.

6.3.2 Peak-Over Threshold Models

The selection of an appropriate threshold is necessary to obtain extreme values that could fit the GPD. Fig. 6.3 depicts the threshold stability plots for the parameters of $PM_{2.5}$ maxima model. Similar plots for PM_{10} were also analyzed. Here, various threshold values of $PM_{2.5}$ are plotted against both modified scale and shape parameters. The solid lines are the expected value of parameters, while the dashed lines represent the 95% confidence interval. Threshold values are chosen from the range of values where both scale and shape parameters remain constant (refer to the solid horizontal line segments in red). In this study, the minimum threshold for which the expected values of the parameters remained stable was selected as the threshold. The final threshold values are depicted in Table **??**. Extreme values were extracted based on these thresholds and Bayesian hierarchical GPD was fit. The posterior distributions of model parameters for both $PM_{2.5}$ and PM_{10} are depicted in Fig. 6.4. Here, σ_1, σ_2 and σ_3 and ξ_1, ξ_2 and ξ_3 are shape and scale parameters for summer, spring and winter, respectively. The summary statistics for both $PM_{2.5}$ and PM_{10} is presented in Table 6.5 and 6.6 (The values were adjusted to the nearest integers except for the shape parameters).

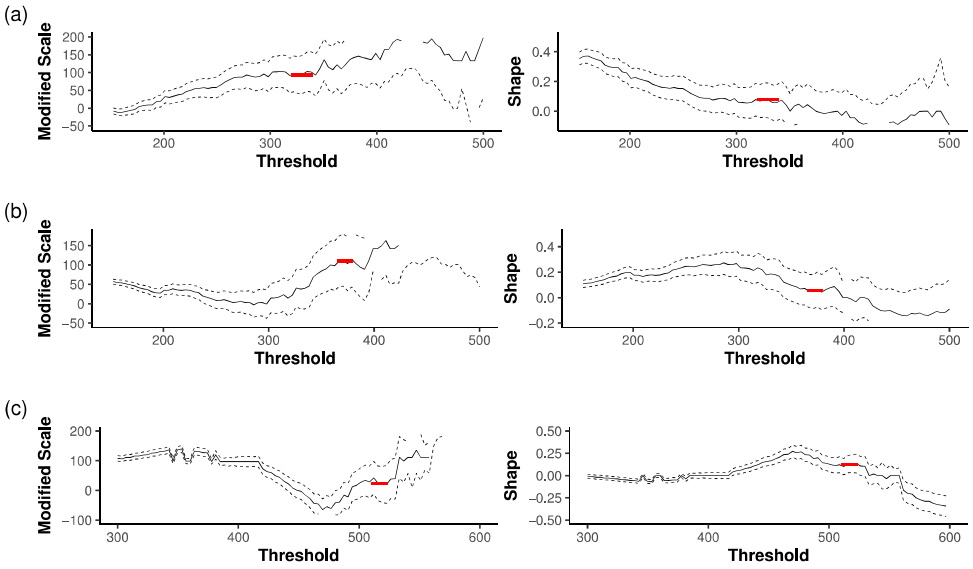


Fig. 6.3 Threshold stability plot for $PM_{2.5}$ observed during (a) Summer, (b) Spring and (c) Winter season.

Table 6.4 Thresholds to segregate extreme PM concentrations (µg m^{-3}) observed in various seasons.

	$PM_{2.5}$			**PM_{10}**		
Season	Summer	Spring	Winter	Summer	Spring	Winter
Threshold range	(320, 340)	(365, 380)	(510, 520)	(460, 480)	(500, 520)	(720, 730)
Model Threshold	340	380	520	480	520	730

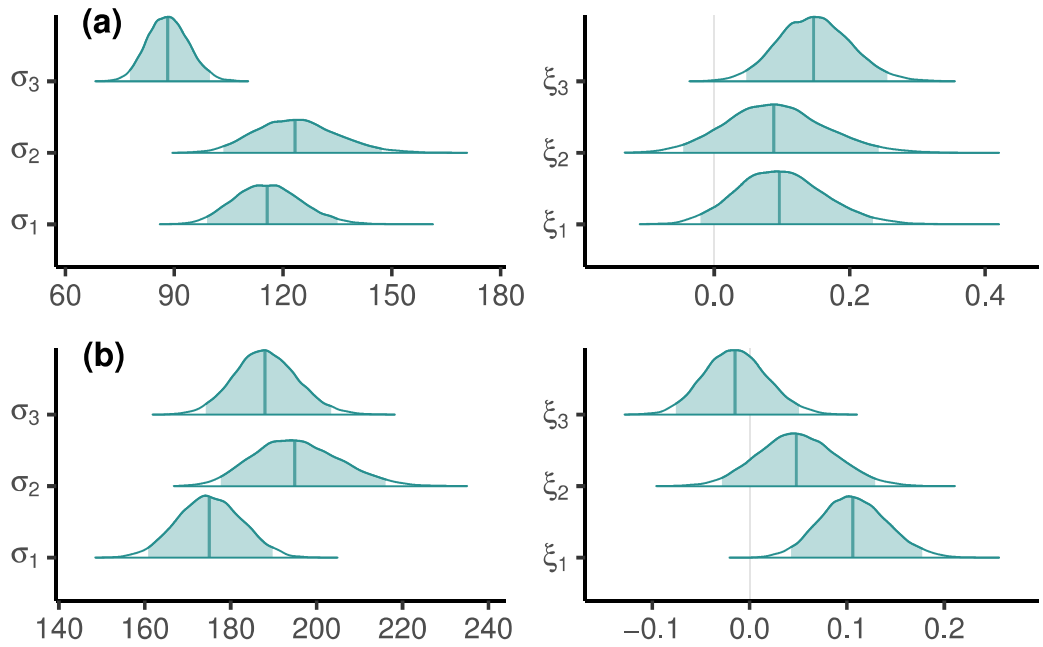


Fig. 6.4 Posterior distribution of parameters for (a) $PM_{2.5}$ and (b) PM_{10} during different seasons.

Table 6.5 Estimate of model parameters for $PM_{2.5}$ (μg m^{-3}) by hierarchical modeling for different seasons.

Model Parameters	Mean	SD	2.5%	50%	97.5%
σ_μ	103	25	40	105	150
σ_σ	41	30	9	32	123
σ_1	116	9	99	116	135
σ_2	124	11	104	123	147
σ_3	88	6	78	88	100
ξ_1	0.10	0.06	-0.02	0.10	0.23
ξ_2	0.09	0.07	-0.05	0.09	0.24
ξ_3	0.15	0.05	0.05	0.15	0.26

Note: σ_1, σ_2, and σ_3 and ξ_1, ξ_2, and ξ_3 are shape and scale parameters for Summer, Spring, and Winter, respectively.

Table 6.6 Estimate of model parameters for PM_{10} (μg m^{-3}) by hierarchical modeling for different seasons.

Model Parameters	Mean	SD	2.5%	50%	97.5%
σ_μ	177	28	92	182	215
σ_σ	31	33	3	21	126
σ_1	175	7	161	175	190
σ_2	195	10	178	195	216
σ_3	188	7	174	188	203
ξ_1	0.11	0.03	0.04	0.11	0.18
ξ_2	0.05	0.04	-0.03	0.05	0.13
ξ_3	-0.01	0.03	-0.08	-0.02	0.05

Note: σ_1, σ_2, and σ_3 and ξ_1, ξ_2, and ξ_3 are shape and scale parameters for Summer, Spring, and Winter, respectively.

6.3.3 Prediction of Extreme Exposure

Prediction of Monthly Extremes During Different Seasons

Predicting maxima of PM at traffic intersections is important for determining extreme exposure levels. The monthly return level for each season was computed using the model parameters estimated by 6.7. The mean and 95% credible interval of extreme PM concentration for different seasons are shown in Fig. 6.5. A 95% CI represents a range of values where the mean PM concentration is expected to lie 95% of the time. As depicted in Fig. 6.5, the predicted $PM_{2.5}$ (PM_{10}) maxima during winter, spring, and summer at different traffic intersections were 589 (1127), 474 (961), and 429 (902) μg m^{-3}, respectively. The observed pattern of PM return level (PMwinter > PMspring > PMsummer) can be attributed to mixing height. The mixing height is defined as the height of the atmospheric layer up to which various pollutants enter and disperse within 1 h by mechanical turbulence or convection (Seibert et al., 2000). The mixing height depends on meteorological factors such as temperature, humidity, wind speed and wind direction. The smaller mixing height in winter is likely due to lower wind speed and temperature

(Guttikunda and Gurjar, 2012). Therefore, it takes a longer time for pollutants to disperse. In contrast, the high temperature and wind speed in summer lead to large mixing heights that expedite dispersion and lower the return levels. The predicted maxima for both $PM_{2.5}$ and PM_{10} concentrations exceeded the severe category ($PM_{2.5}$: 250+ and PM_{10}: 430+ µg m^{-3}).

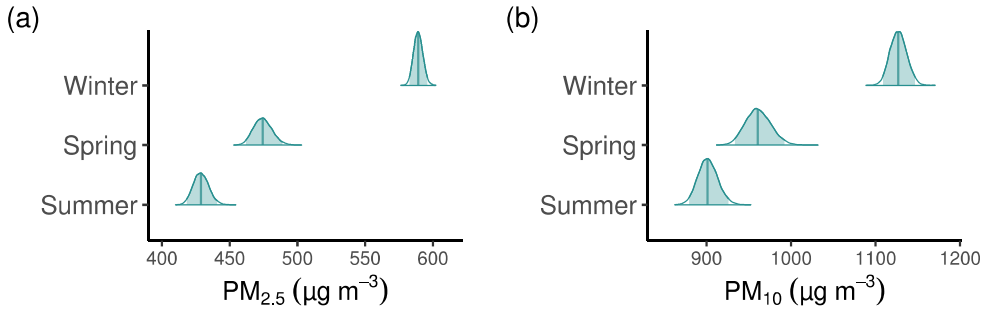


Fig. 6.5 Monthly return levels of (a) $PM_{2.5}$ and (b) PM_{10} during different seasons.

Probability of Exceedances PM Level Observed at the Delhi Smog Event

Consider the "November 2012 Delhi smog" event when exceptionally high values of PM concentration (585 µg m^{-3}) and PM_{10} levels (989 µg m^{-3}) were observed (Sati and Mohan, 2014). The possibility of observing such an event in Varanasi is investigated. The probability that PM concentration in Varanasi will exceed the levels observed in the Delhi smog event was computed 6.6. The probability that such a concentration will be exceeded in Varanasi in winter is 0.72% and 1.47% for $PM_{2.5}$ and PM_{10}, respectively. While in summer the probability of exceedance is 0.21% and 0.47%, respectively. In summary, the probability of exceedance of concentrations observed in the Delhi smog event was low even in winter. The lower risk of such an extreme event may be due to differences in traffic, meteorological conditions and sources of pollution between Varanasi and Delhi.

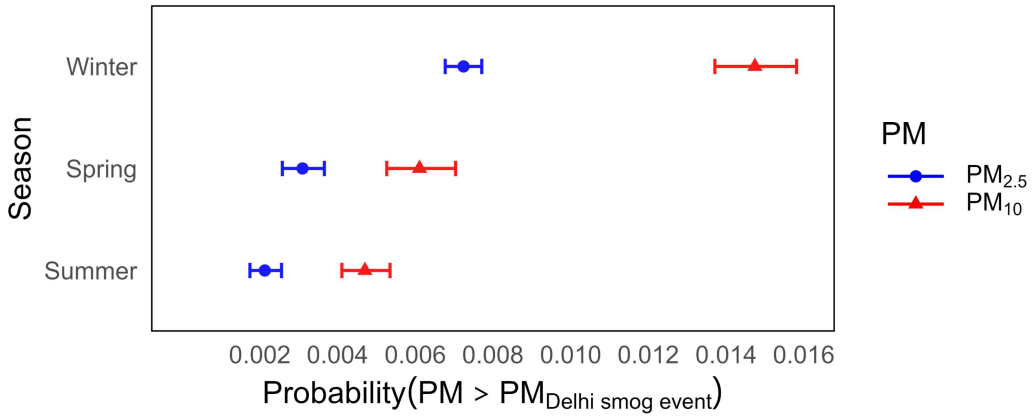


Fig. 6.6 The probability of observing severe PM levels in Delhi's smog event.

Model Validation and Posterior Predictive Check for PM Extremes

The R hat values for all the models were found to be 1, indicating good mixing of the Markov chains and thus sampling was validated. This was also corroborated by the trace plots (Fig. 6.7). As for the posterior predictive check, the simulated data, y_{rep} (replications) generated by the model, aligned closely with the observed data, y (Fig. 6.8). Essentially, the model validation and posterior predictive check suggested that the data fitted the models reasonably well.

Sensitivity Analysis

The sensitivity of PM return level to the thresholds was analyzed. If the selected threshold is sufficiently large, the exceedances follow the GPD with the same model parameters, except for some sampling variability (Coles et al., 2001). The threshold range presented in Table 6.4 was used for extreme PM comparison, with Threshold-1 and Threshold-2 representing the upper and lower limits of this range. The estimated crash probabilities for both thresholds were compared, as shown in Fig. 6.9. A significant overlap exists between the 95% credible regions for PM return level estimated using the two thresholds. Hence, it

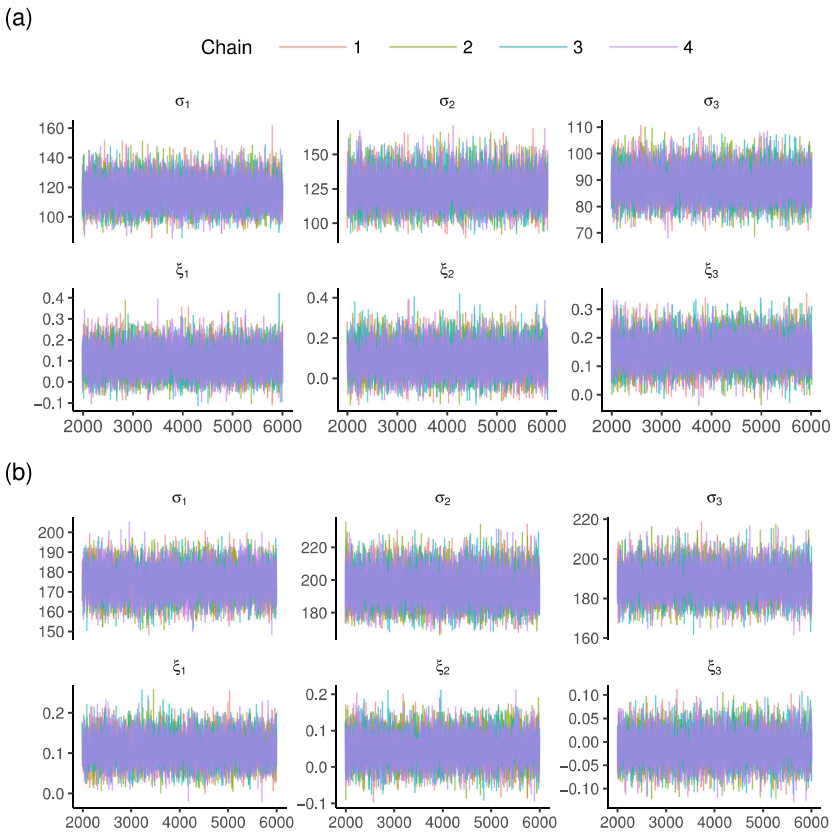


Fig. 6.7 Trace plot for model checking for (a) $PM_{2.5}$ and (b) PM_{10}.

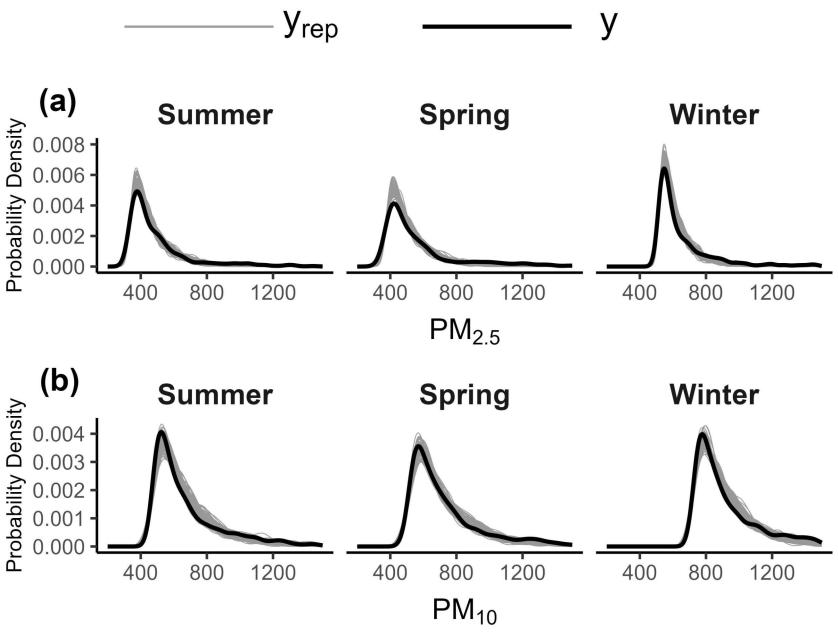


Fig. 6.8 Posterior predictive check for (a) $PM_{2.5}$ and (b) PM_{10}.

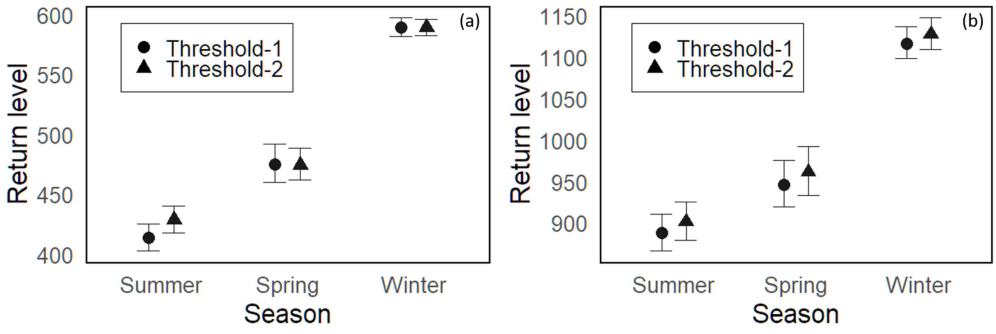


Fig. 6.9 Sensitivity analysis of (a) $PM_{2.5}$ and (b) PM_{10} return levels with different thresholds.

suggests that as long as the threshold is chosen within the stable region, model estimates do not vary considerably. This justified the selected threshold values in the present study.

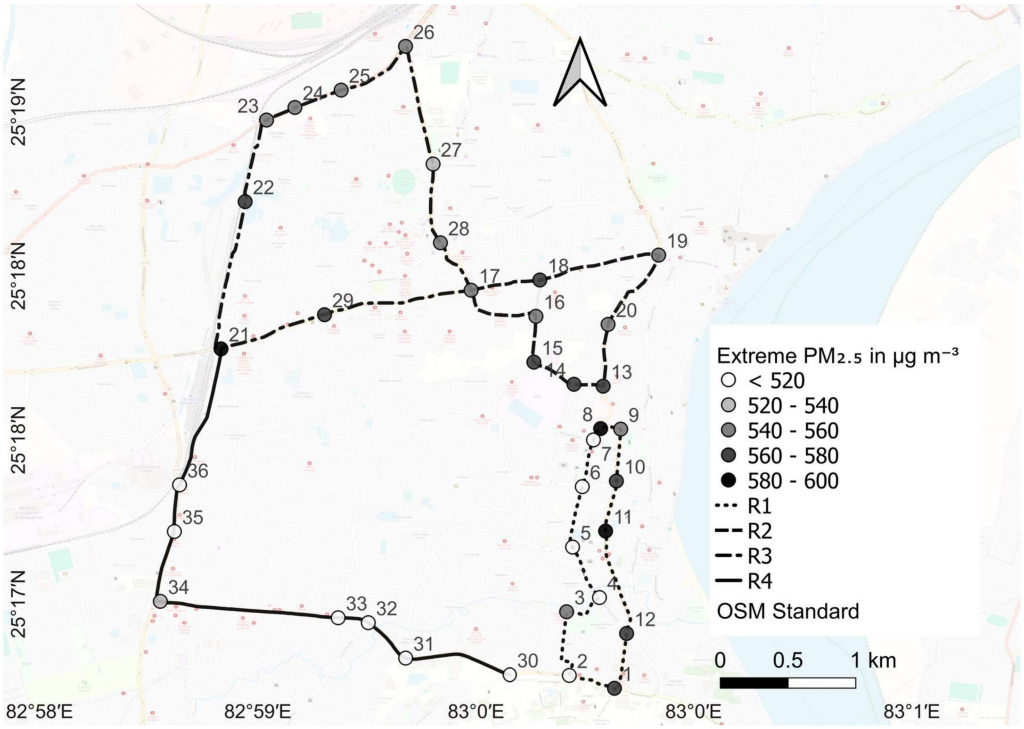
6.3.4 Extreme Exposure Hotspots

The exposure hotspots maps depicting the mean extremes estimated for winter are presented in Fig. 6.10. Mean extremes of $PM_{2.5}$ ranged from 520 to 592 μg m^{-3}, while that of PM_{10} ranged from 730 to 927 μg m^{-3}. These mean extremes of $PM_{2.5}$ and PM_{10} were categorized into bins of sizes 20 and 50 μg m^{-3}, respectively. The intersections (depicted as small circles) are color-coded from light grey to dark grey such that a darker color implies a higher concentration of PM. In contrast, intersections with PM ($PM_{2.5}$ and PM_{10}) levels below the thresholds are depicted as white circles.

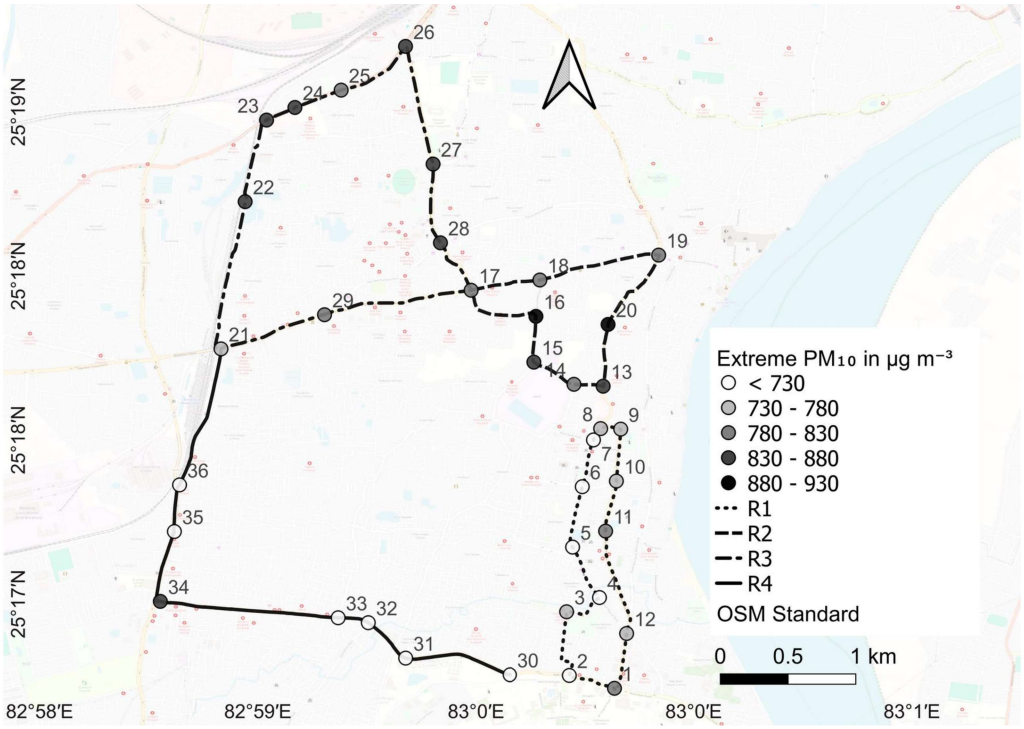
The $PM_{2.5}$ and PM_{10} hotspots showed different patterns. As discussed in background of the chapter, fine particles ($PM_{2.5}$) are more lethal than coarse particles (PM_{10}). Therefore, considering $PM_{2.5}$, R4 was the safest route to travel, as it had the lowest number of hotspots. In contrast, R2 was the worst route for traveling since it had multiple hotspots at various traffic intersections. This could be because of the congestion caused by parked vehicles and street vendors. Additionally, the presence of hospitals and a residential university near the route contributed to high PM concentration.

6.4 Summary

PM extreme studies have been conducted in various cities across the world. However, only a few studies were undertaken at traffic intersections. Thus, the primary objective of this study was to model and predict extreme PM concentrations at traffic intersections in Varanasi, India. The other objectives of the study were to predict extreme PM concentrations, PM return levels, and the probability of occurrence of past extreme pollution events.



(a)



(b)

Fig. 6.10 Average (a) $PM_{2.5}$ and (b) PM_{10} extreme pollution hotspots at traffic intersections during winter.

PM was measured at 36 traffic intersections using a portable aerosol monitor mounted on a motorcycle that traveled through selected routes. In addition, location information was collected using a GPS receiver. Data was collected for 8 months during winter, spring, and summer seasons. Seasonal extreme PM concentrations were estimated using the GPD. Threshold PM, exceedance probability, and return level were estimated using statistical techniques. The probability that PM concentration would exceed that of the Delhi smog event was computed. Based on the analysis, the following conclusions were made.

1. The average PM concentrations at traffic intersections were generally higher than those at mid-block sections in Varanasi, a densely populated city. These traffic intersections act as pollution exposure hotspots with the average air quality designated as poor to very poor according to NAAQS categories.

2. The monthly return levels for the extremes of $PM_{2.5}$ (PM_{10}) were 589 (1127), 474 (961), and 429 (902) μg m^{-3} during winter, spring, and summer, respectively. These values are considerably higher than the levels designated as severe by NAAQS ($PM_{2.5}$: 250 μg m^{-3}, PM_{10}: 430 μg m^{-3}).

3. There is a 1.47% (0.72%) chance that the PM_{10} ($PM_{2.5}$) concentration in Varanasi would exceed what was observed during the smog event in Delhi, at least in winter.