Appendix

Table S1 VIF (Variance Inflation Factor) Values for Various Meteorological Factors in Nanjing, 2015-2023.

Years	SSR	T2M	TCC	PRE	FG10	BLH
2015-2023	2.93	1.76	1.92	1.44	1.04	1.09

Before constructing the GAM model, to avoid overfitting or a lack of physical meaning, different degrees of freedom (k = 3, 5, 7, 9) were specified for each meteorological variable (T2M, PRE, SSR, TCC, FG10, BLH). We compared the model's AIC, explained variance (R^2), effective degrees of freedom (edf) for each meteorological factor, and the shape of the smoothed curves under different degrees of freedom to select the most appropriate degree of freedom for model construction. As shown in Table S2, when k=3, the model AIC is 8850.42 and R^2=0.69, indicating a low level of model explanatory power; when k=5, the AIC decreases to 8723.76 and R^2 increases to 0.73, improving the model's explanatory power without overfitting; when k > 5, although AIC continues to decrease, the increase in R^2 narrows significantly, indicating diminishing marginal returns from increased model complexity and a risk of overfitting. Combining Table S3 with prior research, it is evident that when k = 5, the edfs of each variable stabilize between 3 and 4, thereby retaining sufficient nonlinear fitting capability while avoiding overfitting. The smoothing-effect curves of various meteorological factors on $PM_{2.5}$ concentrations for different degrees of freedom are shown in Fig.S4. When k = 3, the effect curves of meteorological factors on $PM_{2.5}$ are overly smooth, potentially losing key nonlinear features and resulting in underfitting; when k = 9, the curves exhibit severe fluctuations with no physical significance, indicating model overfitting; whereas when k = 5, the curves are smooth and consistent with the logic of atmospheric physical processes. Therefore, this study selects k = 5 to construct the GAM model.

Table S2 Comparison of Goodness-of-Fit for the GAM Model across different degrees of freedom.

Degrees of freedom(k)	AIC	(Explained variance)R^2
3	8850.42	0.69
5	8723.76	0.73
7	8636.38	0.75
9	8543.06	0.77

Note: In the table, k represents the degrees of freedom of the smoothing function in the generalized additive model (GAM); AIC stands for the Akaike Information Criterion, where a smaller value indicates better model fit; R^2 represents the proportion of variance explained by the model, where a larger value indicates stronger explanatory power. In this study, the optimal model degrees of freedom were determined by comparing goodness-of-fit statistics.

Table S3 Effective Degrees of Freedom (edf) for various meteorological factors under different Degrees of Freedom.

k	Variable	edf	k	Variable	edf
3	T2M	1.85	5	T2M	3.83
	PRE	1.00		PRE	3.45
	SSR	1.92		SSR	3.68
	TCC	1.95		TCC	3.84
	FG10	1.75		FG10	2.27
	BLH	1.93		BLH	3.92
k	**Variable**	**edf**	**k**	**Variable**	**edf**
7	T2M	5.64	9	T2M	7.67
	PRE	3.92		PRE	7.1
	SSR	5.89		SSR	7.31
	TCC	5.82		TCC	7.92
	FG10	5.54		FG10	6.98
	BLH	5.59		BLH	7.68

Note: In the table, k represents the degrees of freedom of the smoothing function in the Generalized Additive Model (GAM); edf represents the effective degrees of freedom, which reflect the actual degree of nonlinearity in the smoothing function. The closer the edf is to 1, the more linear the relationship between the factor and $PM_{2.5}$ concentration; the larger the edf value, the stronger the nonlinear characteristics. T2M denotes 2-meter temperature, PRE denotes total precipitation, SSR denotes surface net solar radiation, TCC denotes total cloud cover, FG10 denotes 10-meter wind speed, and BLH denotes boundary layer height. This study evaluates the nonlinear characteristics of factors through effective degrees of freedom.

Table S4 Parameter Estimation Results for the DLNM Model (AR1-Corrected)

Variable Categories	**Variable name**	**Coefficient (β)**	**Standard error (SE)**	**t-value**	**P-value**	**Significance indicator**
Long-term trend	ns(time_num, df=7)	-	-	-	<0.001	***
Seasonal effects	month2	-11.81	2.28	-5.18	<0.001	***
	month3	-8.87	3.19	-2.78	0.006	**
	month4	-9.88	4.06	-2.43	0.015	*
	month5	-11.8	4.72	-2.5	0.012	*
	month6	-15.67	5.03	-3.12	0.002	**
	month7	-19.36	5.24	-3.69	<0.001	***
	month8	-23.13	5.23	-4.42	<0.001	***
	month9	-22.14	4.71	-4.7	<0.001	***
	month10	-16.25	4.05	-4.01	<0.001	***
	month11	-14.02	3.19	-4.39	<0.001	***
	month12	-5.63	2.25	-2.5	0.013	*
Weekday effects	weekday2	0.01	0.68	0.01	0.989	-
	weekday3	-1.47	0.83	-1.77	0.076	-
	weekday4	-1.74	0.88	-1.98	0.048	*
	weekday5	-1.21	0.88	-1.37	0.17	-
	weekday6	-1.03	0.83	-1.24	0.215	-
	weekday7	-0.58	0.68	-0.86	0.392	-
Meteorological factors	T2M	-	-	-	<0.001	***
	PRE	-	-	-	<0.001	***
	BLH	-	-	-	<0.001	***
	FG10	-	-	-	<0.001	***
	TCC	-	-	-	<0.05	*
	SSR	-	-	-	>0.05	-

Note: Significance levels: ***P < 0.001, **P < 0.01, *P < 0.05; - no significance. The model employs an autoregressive (AR1) structure to correct for temporal autocorrelation, while controlling for long-term trends, seasonal effects, and weekday effects; the reference group is January and Sunday.

Table S5 Variable Definitions

Name	Units	Description
10m v-component of neutral wind	m/s	This parameter is the northward component of the "neutral wind", at a height of 10 metres above the surface of the Earth. The neutral wind is calculated from the surface stress and the corresponding roughness length by assuming that the air is neutrally stratified. The neutral wind is slower than the actual wind in stable conditions, and faster in unstable conditions. The neutral wind is, by definition, in the direction of the surface stress. The size of the roughness length depends on land surface properties or the sea state.
10m u-component of neutral wind	m/s	This parameter is the eastward component of the "neutral wind", at a height of 10 metres above the surface of the Earth. The neutral wind is calculated from the surface stress and the corresponding roughness length by assuming that the air is neutrally stratified. The neutral wind is slower than the actual wind in stable conditions, and faster in unstable conditions. The neutral wind is, by definition, in the direction of the surface stress. The size of the roughness length depends on land surface properties or the sea state.
2m temperature	K	This parameter is the temperature of air at 2m above the surface of land, sea or inland waters. 2m temperature is calculated by interpolating between the lowest model level and the Earth's surface, taking account of the atmospheric conditions. This parameter has units of kelvin (K). Temperature measured in kelvin can be converted to degrees Celsius (°C) by subtracting 273.15.
Boundary layer height	m	This parameter is the depth of air next to the Earth's surface which is most affected by the resistance to the transfer of momentum, heat or moisture across the surface. The boundary layer height can be as low as a few tens of metres, such as in cooling air at night, or as high as several kilometres over the desert in the middle of a hot sunny day. When the boundary layer height is low, higher concentrations of pollutants (emitted from the Earth's surface) can develop. The boundary layer height calculation is based on the bulk Richardson number (a measure of the atmospheric conditions) following the conclusions of a 2012 review.

Surface net solar radiation	J/m^2	This parameter is the amount of solar radiation (also known as shortwave radiation) that reaches a horizontal plane at the surface of the Earth (both direct and diffuse) minus the amount reflected by the Earth's surface (which is governed by the albedo). Radiation from the Sun (solar, or shortwave, radiation) is partly reflected back to space by clouds and particles in the atmosphere (aerosols) and some of it is absorbed. The remainder is incident on the Earth's surface, where some of it is reflected. This parameter is accumulated over a particular time period which depends on the data extracted.
Total cloud cover	Dimensionless	This parameter is the proportion of a grid box covered by cloud. Total cloud cover is a single level field calculated from the cloud occurring at different model levels through the atmosphere. Assumptions are made about the degree of overlap/randomness between clouds at different heights. Cloud fractions vary from 0 to 1.
Total precipitation	m	This parameter is the accumulated liquid and frozen water, comprising rain and snow, that falls to the Earth's surface. It is the sum of large-scale precipitation and convective precipitation.

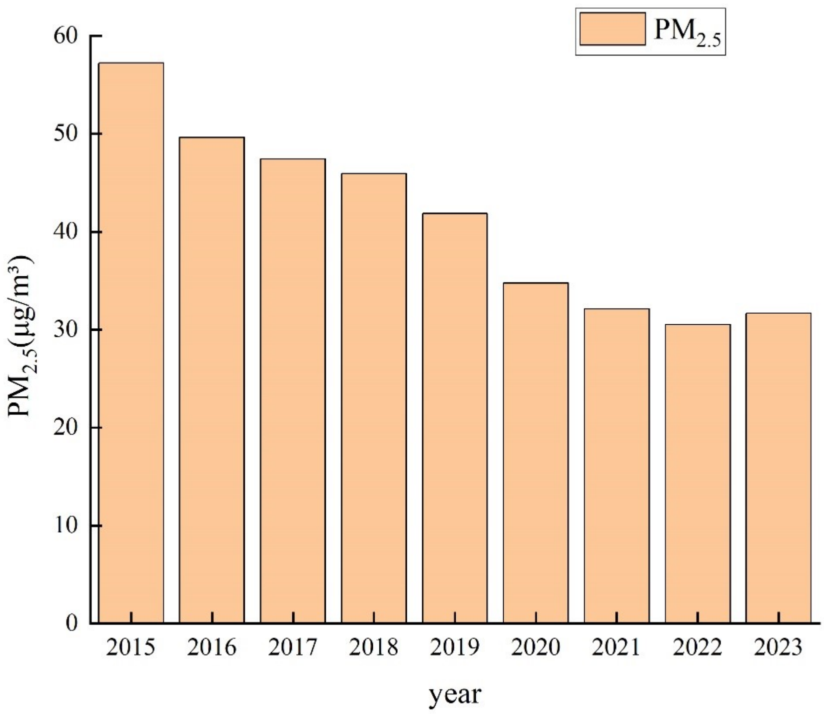

Fig.S1 $PM_{2.5}$ Concentrations in Nanjing, 2015-2023.

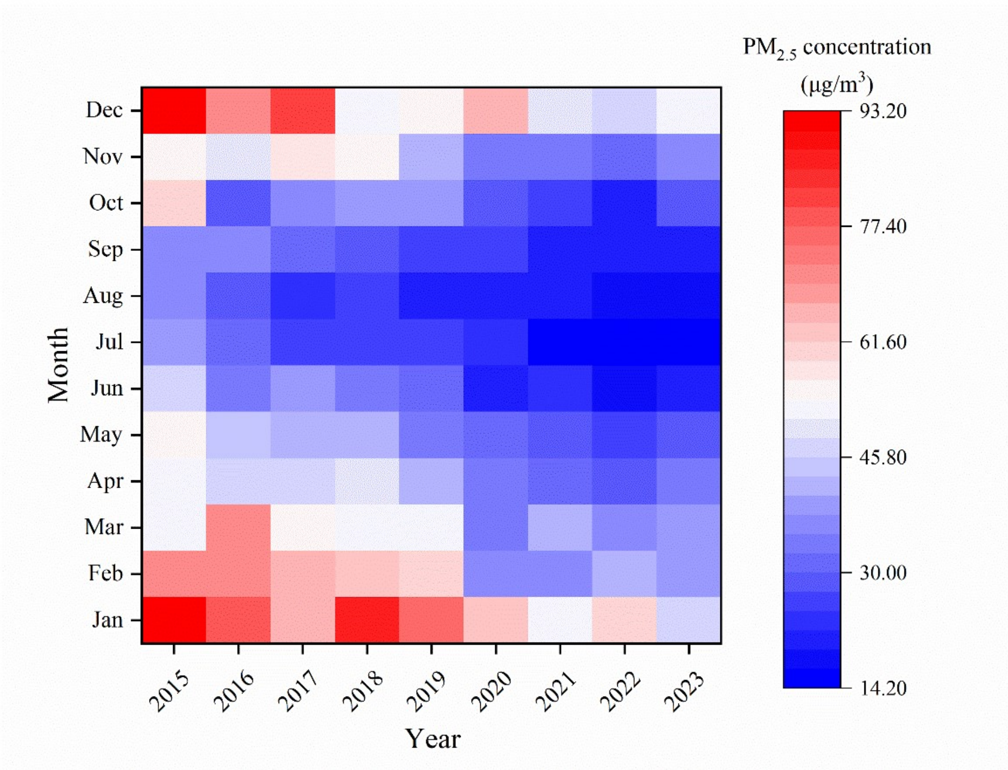

Fig.S2 Heat Map of Monthly Average $PM_{2.5}$ Concentrations in Nanjing, 2015-2023.

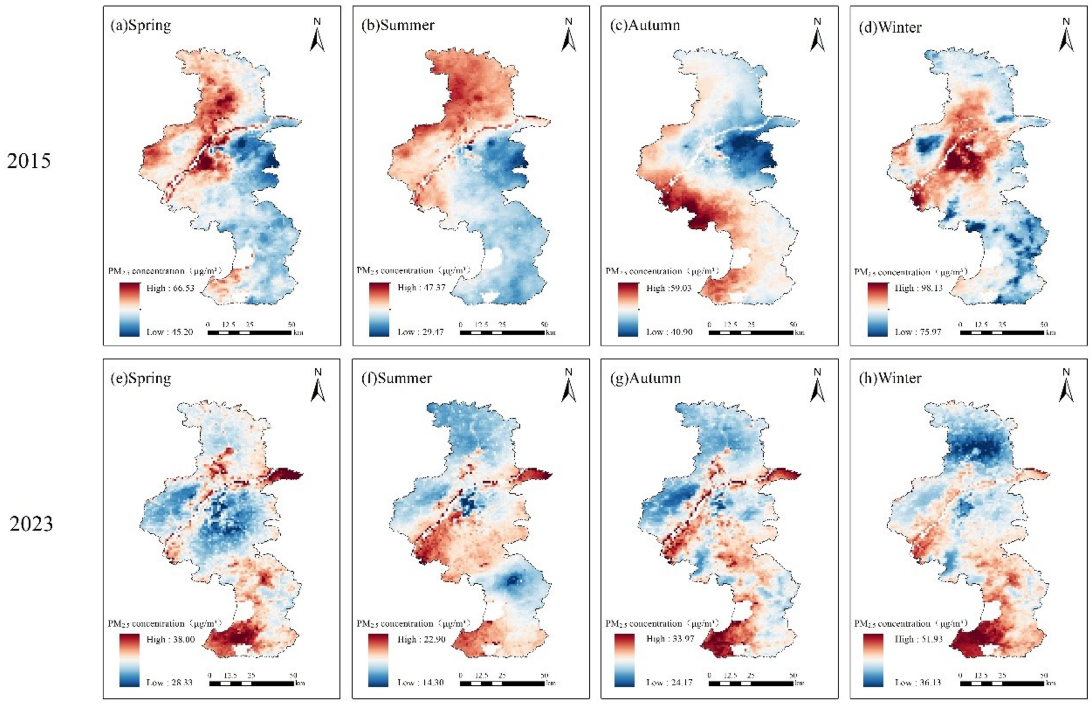

Fig.S3 Spatial Distribution of $PM_{2.5}$ Concentrations in Nanjing by Season in 2015 and 2023.

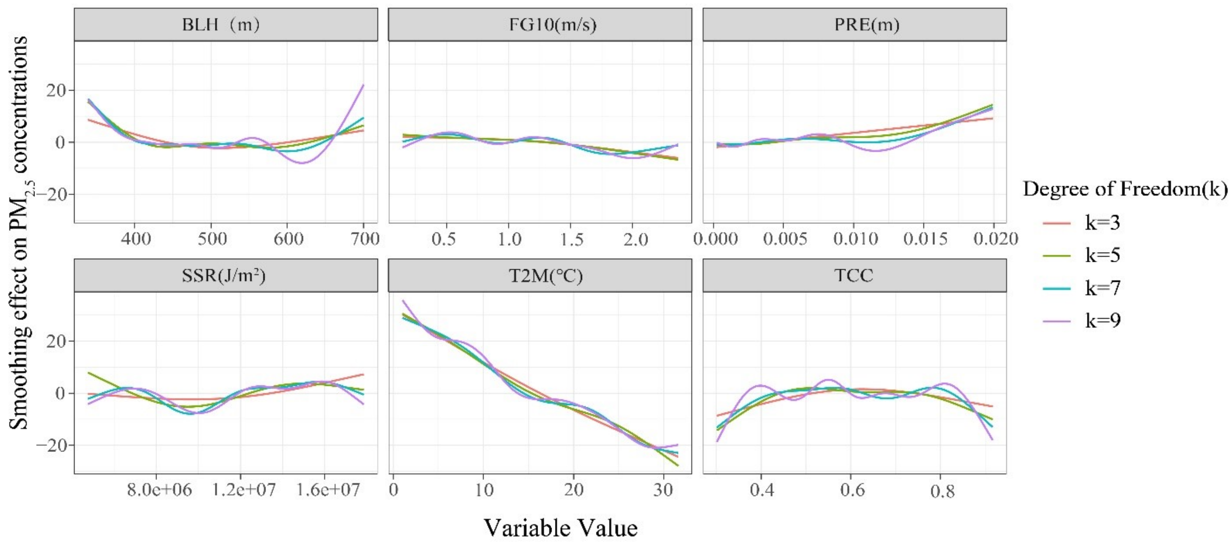

Fig.S4: Comparison of the smoothing effects of various meteorological factors on $PM_{2.5}$ under different degrees of freedom.

To investigate the lagged effects of meteorological factors on $PM_{2.5}$ concentrations in Nanjing, a DLNM model was fitted to the relationship between $PM_{2.5}$ concentrations and six meteorological parameters: SSR, T2M, TCC, PRE, FG10, and BLH. In the model, a natural spline function ns (time, df=7×year) was used to remove long-term trends and seasonal fluctuations. Dummy variables for month and weekday were included to control for monthly and weekly effects, and all meteorological factors were incorporated to construct the DLNM model. The results of the ACF (autocorrelation function of model residuals) test are shown in Fig.S5. It can be seen that when Lag = 1 and Lag = 2, the ACF values are approximately 0.5 and 0.2, respectively, exceeding the blue dashed line (95% confidence interval), indicating significant short-term autocorrelation in the variables; when Lag ≥ 3, the ACF values partially exceed the confidence interval, indicating that the variable exhibits partial autocorrelation when the lag period exceeds 3 days. These results suggest that the time series autocorrelation has not been fully controlled. Therefore, this study employs an autoregressive (AR(1)) structure to correct the model, thereby eliminating the bias in standard error estimates caused by short-term autocorrelation.

This study employed an AR(1)-corrected DLNM model to analyze the impact of meteorological factors on $PM_{2.5}$ concentrations, while simultaneously controlling for long-term trends, seasonal effects, and weekday effects (Table S4). The results showed that both the long-term trend term and the month variable were highly significant ($P < 0.001$), and $PM_{2.5}$ concentrations exhibited a distinct seasonal pattern, with high levels in winter and spring and low levels in summer; the weekday effect was only weakly significant for Thursdays ($P < 0.05$). T2M, PRE, BLH, and FG10 all had highly significant effects on $PM_{2.5}$ concentrations ($P < 0.001$), TCC had a weakly significant effect ($P < 0.05$), and SSR was not statistically significant ($P > 0.05$).

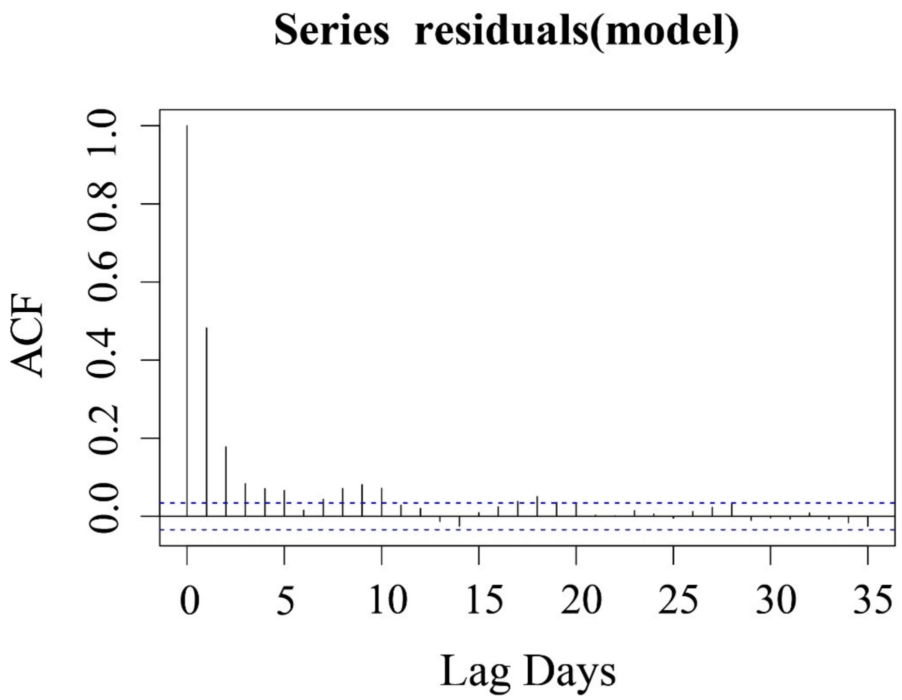

Fig.S5 Autocorrelation Function (ACF) Plot of the DLNM Model Residuals.

Note: The x-axis in the figure represents the number of lag days (Lag Days), the y-axis represents the autocorrelation coefficient (ACF), and the blue dashed line indicates the 95% confidence interval. This figure is used to examine the autocorrelation characteristics of the model residuals.