International Journal of Physics and Applications

E-ISSN: 2664-7583
P-ISSN: 2664-7575
IJOS 2023; 5(1): 01-06

www.physicsjournal.in
Received: 05-11-2022
Accepted: 07-12-2022

Anil Kumar
Professor, Department of Physics, Hindu College, Moradabad Affiliated to M.J.P. Rohilkhand University, Bareilly, Uttar Pradesh, India

Corresponding Author:
Anil Kumar
Professor, Department of Physics, Hindu College, Moradabad Affiliated to M.J.P. Rohilkhand University, Bareilly, Uttar Pradesh, India

Prediction and Analysis of ambient air particle pollutants using stationary wavelet transforms

Anil Kumar

DOI: https://doi.org/10.33545/26647575.2023.v5.i1a.55

Abstract
Ambient air particle pollutants (PM2.5 and PM10) are hazardous and play an important role in the air pollution focusing on human health and climate. Wavelet transforms extract local properties and information from a signal. In stationary wavelet transforms, the same number of samples as the input is maintained at every decomposition level and reconstructions result has lower error values and faster convergence compared to discrete wavelet transforms. The prediction work is carried out directly with the general linear predictor and wavelet predictor because linear predictor is a non-unique projection onto the wavelet domain. The approximation and detail represent average behaviour or trend and differential behaviour or changes of the signal respectively. The wavelet and statistical analysis for both original and extended signal are performed and discussed.

Keywords: Approximation, detail, PM2.5, PM10, prediction, wavelet

Introduction

The particle pollution or particulate matter (PM) is a complex mixture of solid and liquid particles suspended in the air and most of them are hazardous. In the complex mixture there are both organic and inorganic particles, such as smoke, pollen, dust, soot and liquid droplets [1]. The size, composition and origin of these particles vary. Larger particles would be filtered in the nasal duct. Coarse particle size is between 2.5-10 µm and fine particles size is less than 2.5 µm. The composition of air pollutants is the main factor which decides their danger. By the presence of sulphur, the particles are hygroscopic and thus sulphur dioxide is converted into sulphate in the presence of high humidity and low temperatures. This results the reduction in visibility. The sub micro-particles are produced when condensation of metal or organic compound takes place. These become vaporized during high temperature combustion process. The micro-particles are also produced when condensation of gases which are transformed into low-vapour-pressure substances during atmospheric reactions. Organic substances and Sulphate play main contribution to the annual average of PM2.5 and PM10 concentration. When PM_>50 µg/m^3, nitrates are the main contributor to PM2.5 and PM10 [2-3]. Due to complexity and the significance of particle size in exposure and human dose, the study of ambient particulate matter is very important.

In general, Particles whose diameters are 10 µm or less are generally considered as respirable by human being [4]. Many of the inhaled particles when exhaled, there is no substance deposited in respiratory tract. Generally, the diameter of particles for minimal deposition should be around 0.5 µm. The particles whose diameters lie between 0.5 to 10 µm, result in greater lung deposition, while for particles whose diameter are less than 0.5 µm, the lung deposition level is inversely proportional to the size of particles. The localization of deposition into the lung affects the mode, rate and status of the clearance. Soluble particles result to extracellular fluids or mucus, then into epithelial cells, from which they pass into the circulation. Insoluble fine particles are also inhaled into the respiratory epithelium and sometimes enter the human blood within minutes of inhalation and become dangerous [5]. Climate of any region is directly affected by atmospheric aerosols because the solar radiation incoming and outgoing both cause changes due to atmospheric aerosols of that region. These changes take place through various mechanisms which are categorized as direct, indirect and semi-direct aerosol effects. In direct aerosol effect, the direct interaction of atmospheric aerosols and radiation takes place by absorption, scattering, etc. and both short and long wave radiation are affected so a net negative radiative force is generated [6]. In the indirect aerosol effect, the atmospheric aerosols change the earth's radiative budget by any indirect mean such as modification of clouds, etc. and result a change in the climate of that region. In the semi-direct effect, any radiative change takes place by absorbing atmospheric aerosols such as soot, etc. other than direct effects. It has several individual mechanisms and is poorly defined than direct and indirect aerosol effects. The Indian monsoon sometimes becomes failed due to less evaporation of water from the Indian Ocean due to anthropogenic aerosol effect. Today we are facing the most challenging problems related to air quality and climate due to particulate matter.

Fourier transforms has poor time frequency localization because the global mixing of information enables it difficult to detect local property of the signal or function. Wavelet transforms is a powerful tool having wide applications not only in mathematical sciences but also in Physics and Engineering [7-8]. Wavelet transforms is a well-known method for providing spectral behaviour using wavelet functions for representing functions at various positions and scales.

Basics of wavelet transforms

Wavelet is a small wave or function that can be dilated and translated to generate a set of orthonormal basis functions for representing a signal. Such functions in number may be infinite and the best one is selected according to our signal. The wavelet function is defined over a range of interval and it is called support of the wavelet. Wavelet exhibits oscillatory behaviour for a short time interval and then dies out. For any two real numbers *a* and *b*, a wavelet function is defined as follows [9-10]:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) = D_a T_b \psi$$

Putting $a = 2^{-j}$ and $\frac{b}{a} = k$,, we get discrete wavelets as follows: -

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$$

Where a and b are the dilation and translation parameter respectively. Here ψ(t) is real-valued and this collection of wavelets is used as an orthonormal basis. The continuous and discrete wavelet transform is defined as follows:-

$$W_{a,b} = \int f(t)\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)dt$$
$$W_{j,k} = \int f(t)2^{j/2}\psi(2^j t - k)\,dt$$

The localized time frequency information of a signal is obtained with help of wavelet transforms. Discrete wavelet analysis is performed by a fundamentally new iterative method called multiple-resolution analysis (MRA). MRA consisting a sequence V_j : j ∈ Z of closed subspaces of L^2 (R), a Lebesgue space of square integrable functions, satisfies the following properties as follows [11-13]:-

1) $V_{j+1} \subset V_j : j \in \mathbb{Z}$
2) $\cap_{j\in\mathbb{Z}} V_j = \{0\},\ \cup_{j\in\mathbb{Z}} W_j = L^2(\mathbb{R}),$
3) For every, $L^2(\mathbb{R}), f(x) \in V_j \Rightarrow f(2x) \in V_{j+1},\ \forall\ j \in \mathbb{Z}$
4) For a function $\phi(x) \in V_0$ the function $\{\phi(x-k): k \in \mathbb{Z}\}$ is orthonormal basis of V_0.

Here, the function $\phi(x)$ is called a scaling function of given MRA and implies a dilation equation as following: -

$$\phi(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} \alpha_k\ \phi(2x-k)$$

Where α_k are called low pass filters, defined as follows: -

$$\alpha_k = \int_{-\infty}^{\infty} \phi(x)\phi(2x-k)\,dx$$

The function ψ , that is wavelet function can be expressed as follows: -

$$\psi(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} \beta_k\ \phi(2x-k)$$

Where $\beta_k = (-1)^{k+1}\alpha_{1-k}$ are high pass filters. From conditions of multiresolution analysis (MRA) and elementary functional analysis, each space V_j can be expressed as combination of two subspace V_{j+1} and W_{j+1} such that every function f in V_j can be uniquely decomposed into $f = u + v$ with $u \in V_{j+1}$ and $v \in W_{j+1}$. We write this as follows: -

$$V_j = V_{j+1} \oplus W_{j+1},\ \forall j \in \mathbb{Z}$$

If all such functions u and v are orthogonal ($\langle u, v\rangle = 0$), then W_{j+1} is the orthogonal complement of V_{j+1} in V_j ($V_{j+1} \perp W_{j+1}$) and the construction below will give the scaling function and mother wavelet of an orthonormal wavelet basis for $L^2((\mathbb{R})$. By MRA, the orthogonal decomposition of space $L^2(\mathbb{R})$ is as following: -

$$L^2(\mathbb{R}) = \sum_j V_j = \sum_j W_{j+1} \oplus W_{j+2} \oplus W_{j+3\ldots\ldots} \oplus W_{j_o} \oplus V_{j_o}$$

Any discrete signal in square summable space $\ell^2(\mathbb{Z})$ can be expressed as follows: -

$$f[n] = \frac{1}{\sqrt{M}}\sum_k a[j_0, k]\phi_{j_0,k}[n] + \frac{1}{\sqrt{M}}\sum_{p=j+1}^{j_0}\sum_{k\in\mathbb{Z}} d[p,k]\psi_{p,k}[n]$$

Here $f[n]$, $\phi_{j_0,k}[n]$ and $\psi_{p,k}[n]$ are discrete functions having total M points defined in interval $[0, M-1]$. The wavelet coefficients can be derived as follows: -

$$a[j_0,k] = \frac{1}{\sqrt{M}}\sum_n f[n]\ \phi_{j_0,k}[n]$$
$$d[p,k] = \frac{1}{\sqrt{M}}\sum_n f[n]\ \psi_{p,k}[n]$$

with $j < p \le j_0$, where $a[j_0,k]$ and $d[j,k]$ are known as approximation and detailed coefficients respectively [14].

Stationary Wavelet Transforms (SWT)

In SWT, the same number of samples as the input is maintained at every decomposition level and at decomposition of N levels redundancy of N in the wavelet coefficients exists. The SWT reconstructions result lower error values and faster convergence compared to DWT [15-16]. This is achieved by SWT thresholding, which provides a translation-invariant basis. By SWT thresholding, a redundant decomposition can be obtained as follows:-

$$\tilde{a}_{2^j}^{2^j k+p} = \langle f(t), 2^{\frac{j}{2}} \phi(2^j(t-p)-k) \rangle$$

$$\tilde{d}_{2^j}^{2^j k+p} = \langle f(t), 2^{\frac{j}{2}} \psi(2^j(t-p)-k) \rangle$$

Where $p \in \{0, \ldots, 2^J - 1\}$. For decomposition level j_m, $2j_m$ different orthogonal bases are generated. Each path from the root of the tree to a leaf corresponds to the set of functions as follows: -

$$\left\{2^{\frac{j}{2}} \psi\left(2^j\left(t-p_j\right)-k\right)\right\} \cup \left\{2^{\frac{j_m}{2}} \psi\left(2^{j_m}\left(t-p_{j_m}\right)-k\right) 1 \le j \le j_m\right\}$$

Where $1 \le j \le j_m$ and $k \in Z$, Form an orthogonal wavelet basis, resulting in a standard wavelet reconstruction [21]. Every covariance-stationary process X_t has a Cramer representation as follows: -

$$X_t = \int_{-\pi}^{\pi} A(\omega) e^{i\omega t} dZ(\omega)$$

Where $Z(\omega)$ represents a stochastic process having orthonormal increments. Non-stationary processes represent a slow change over time of the amplitude $A(\omega)$ [18]. In LSW process the amplitude $A(\omega)$ in the Cramer representation is replaced by a time varying quantity and the Fourier harmonics $e^{i\omega t}$ by non-decimated discrete wavelets $\psi_{j,k}(t)$; $j, k \in Z$. Here j and k are the scale and location parameter respectively. Time-modulated (TM) process $X_{t,T}$ is defined as follows: -

$$X_{t,T} = \sigma\left(\frac{t}{T}\right) Y_t$$

Where Y_t represents a zero-mean stationary process with variance one and the local standard deviation function $\sigma(z)$ is Lipschitz continuous on (0, 1). Process $X_{t,T}$ is locally stationary wavelet (LSW) if,

i) The auto covariance function of Y_t is absolutely summable so that Y_t is an LSW with a time-invariant spectrum $\{S_j^Y\}_j$;
ii) The Lipschitz constants $L_j^X = D(S_j^Y)^{\frac{1}{2}}$ satisfy the Cramer representation. where D is the Lipschitz constant.

If above two conditions are satisfied, the spectrum $S_j(z)$ of $X_{t,T}$ is expressed as follows: -

$$S_j(z) = \sigma^2(z) S_j^Y$$

The general LSW processes are applicable to model processes whose variance and autocorrelation function both vary with time. For t observations of non-stationary data $X_{0,T}, X_{1,T}, \ldots \ldots X_{t-1,T}$ of an LSW process, the general linear predictor $X_{t+h,T}$ corresponding to h-step ahead, is expressed as follows: -

$$\hat{X}_{t+h,T} = \sum_{s=0}^{t-1} b_{t-x,T}^{(h)} X_{s,T}$$

Where the coefficients $b_{t-x,T}^{(h)}$ minimise the Mean Square Prediction Error (MSPE). That is, as $T \to \infty$, allows us to fit coefficients $b_{t-x,T}^{(h)}$ with more accuracy. Here h is the prediction horizon, we set $T = t + h$. Let us consider the forecasting horizon $h = 1$, so that $T = t + 1$. The empirical wavelet coefficients in the wavelet domain in terms of prediction operator are defined as follows: -

$$d_{j,k;T} = \sum_{s=0}^{t} X_{x,T}, \psi_{j,k}(t)$$

for all $j = 1, \ldots\ldots, J$ and $k \in Z$. The one-step ahead predictor in terms of wavelet coefficients is defined as: -

$$\hat{X}_{t+1,T} = \sum_{j=1}^{J} \sum_{k \in Z} d_{j,k;T} \; c_{j,k;T}^{(1)} \; \psi_{j,k}(t)$$

Where the estimated coefficients $c_{j,k;T}^{(1)}$ are such that they minimise the MSPE. This predictor is defined as a projection of $X_{t,T}$ on the space having random variables and spanned by,

$$\{d_{j,k;T} \; ; \; j = 1, \ldots\ldots, J \; and \; k = 0, \ldots\ldots, T\}.$$

Because of the redundancy of the non-orthogonal wavelet system $\{\psi_{j,k}(t)\}$ the predictor has more than one solution $(\{c_{j,k}^{(1)}\}$, and every solution corresponds to the same predictor in terms of the different linear combination of redundant functions $\{\psi_{j,k}(t)\}$. Therefore, the wavelet predictor and the linear predictor can be expressed as follows: -

$$b_{t+x;T}^{(1)} = \sum_{j=-J}^{-1} \sum_{k \in Z} c_{j,k;T}^{(1)} \; \psi_{j,k}(t) \; \psi_{j,k}(x)$$

Due to the redundancy of the non-decimated wavelet system, a fixed sequence $b_{t+x;T}^{(1)}$ is expressed as the linear combination of more than one sequence $c_{j,k;T}^{(1)}$. Therefore, the prediction work is carried out directly with the general linear predictor, and wavelet predictor is determined from above equation because linear predictor is a non-unique projection onto the wavelet domain [19].

Study Area and Research methodology

Moradabad is a metropolitan area of Uttar Pradesh state in Northern India and is situated at the banks of Ramganga River. The latitudinal extent of city is 28°20'N to 29°15' N and longitudinal extent is 78°4' E to79°E. We have taken daily PM2.5 and PM10 data at Buddhi Vihar, Moradabad, UP in India issued by website of Central Pollution Control Board (https://cpcb.nic.in) from Jan. 01, 2022 to Dec. 31, 2022 (Total 365 points) as original signal and extended it up to 512 points (Figure 1 and 2).

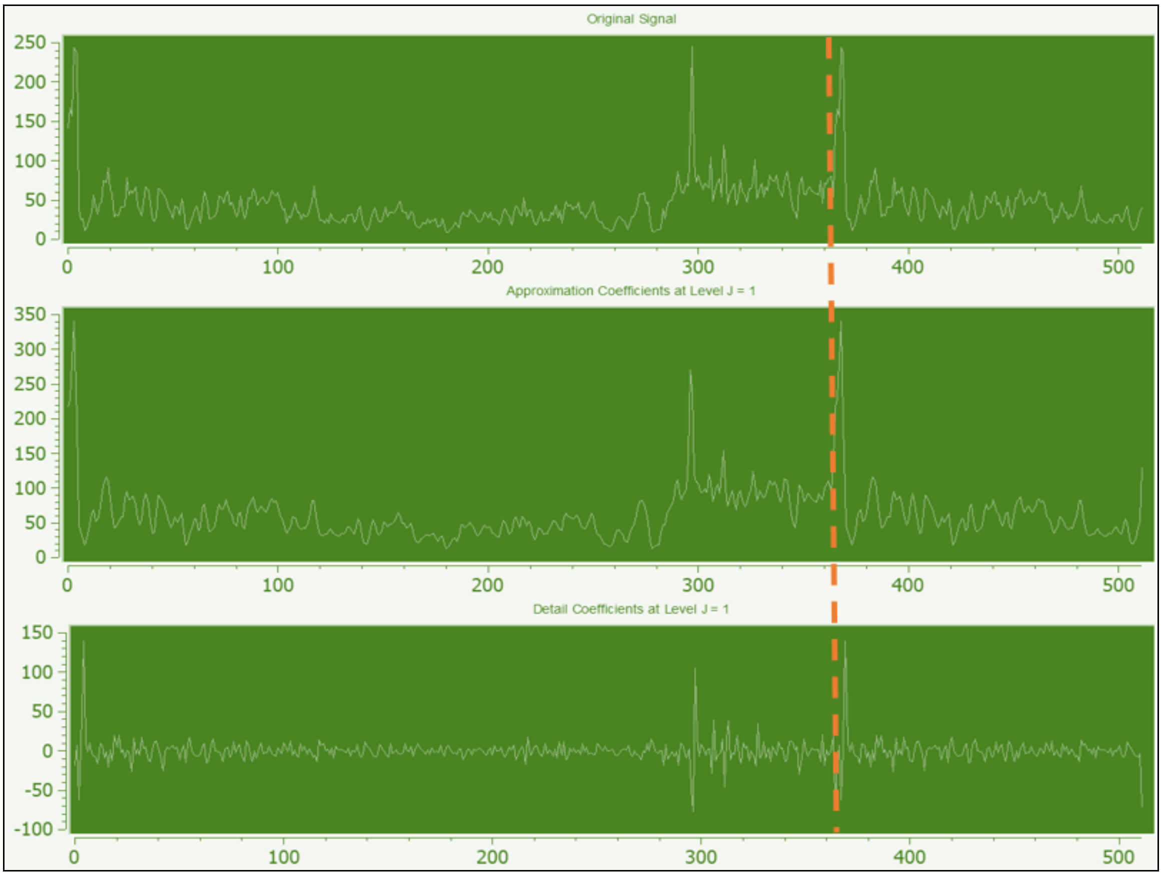

Fig 1: PM2.5 in µg/m^3 and its Wavelet Prediction

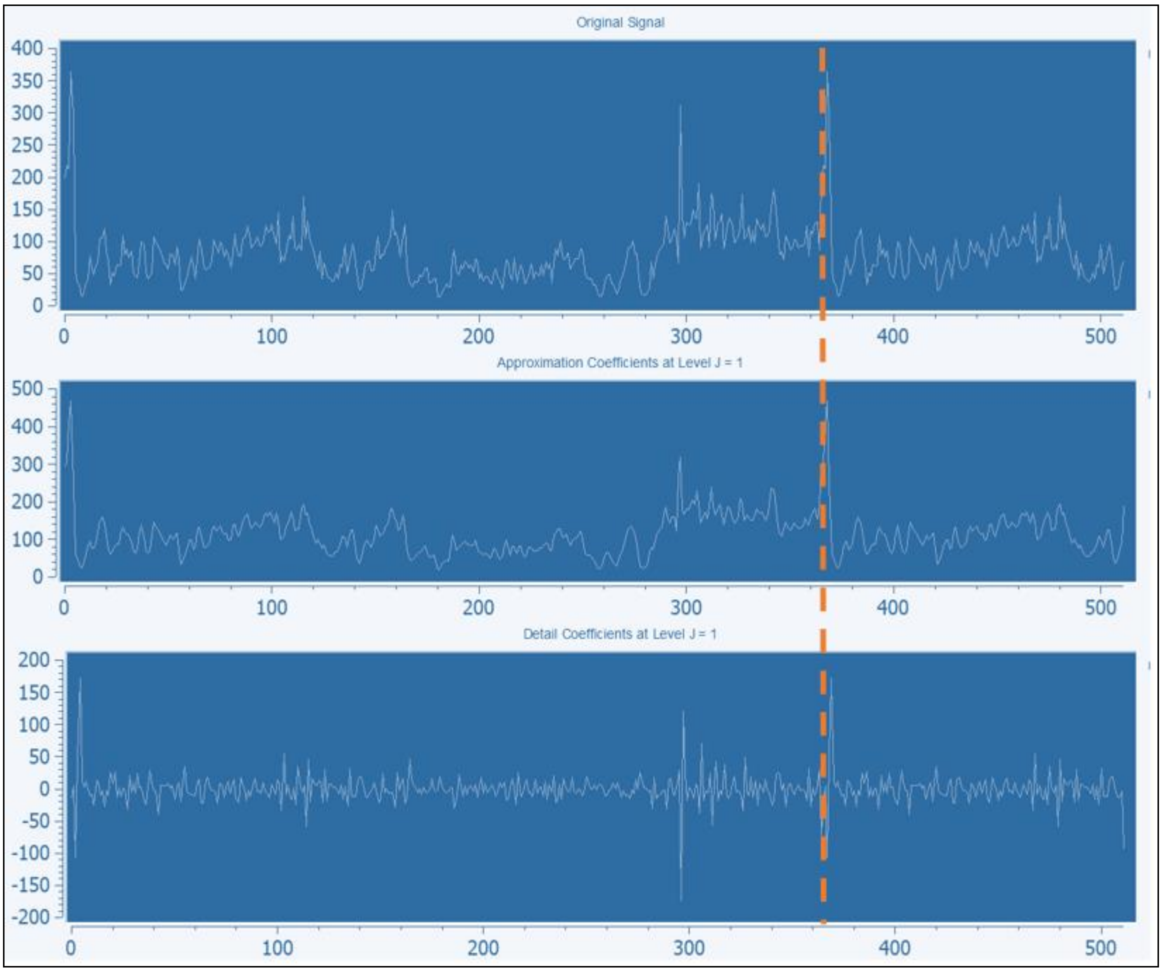

Fig 2: PM10 in µg/m^3 and its Wavelet Prediction

A given signal can be expressed in terms of approximation and detail as follows: -

$$f_0 = A_1 + D_1$$

Here A_1 is approximation and D_1 is detail of signal at various scale or time frames, and expressed as $A_1 = \sum_k a_{1,k}\phi_{1,k}(t)$ and $D_1 = \sum_k d_{1,k}\psi_{1,k}(t)$. To perform the spectral analysis of PM2.5 and PM10, we have used wavelet toolbox of software MATLAB and software dyadwaves using Haar wavelet as an adaptive wavelet. To know the prediction of PM2.5 and PM10 quantity; we extend the signal up to a next level (level 8) with help of locally stationary wavelet transform prediction (total 2^9 =512 points). The statistical parameters like Average, Skewness, Kurtosis and correlation coefficients for both original signal and extended signal are determined and discussed [20].

Results and Discussion

From Figure 1 and 2, it is obvious that concentration of ambient air particle pollutants is continuously varying from time to time. Its higher concentration is the indication of high pollution making it unsuitable and dangerous to human beings and plants. In figure 1(a) and 2 (a), the daily average behaviour of PM2.5 and PM10 are shown, while in figure 1(b) and 2(b), the approximation and in figure 1(c) and 2(c) the detail are shown. The approximation is the slowest part of any signal corresponding to the highest scale value in wavelet analysis terminology. When the scale is increased, the resolution decreases which provides a better estimate of the unknown trend. Approximation possesses low frequency information. The detail provides the changes or fluctuations in the given signal.

With the help of stationary wavelet transforms, the data of 1 years based upon daily average record is extended up to 512 points and decomposed into approximation and detail. Some statistical parameters of original and extended signal are as follows:

Table 1: Statistical Parameters of Original and Extended Signal

S. No.	Parameters	Original Signal		Extended signal	
		PM2.5	PM10	PM2.5	PM10
1.	Average	44.20425	79.99625	44.42031	80.2193
2.	Skewness	3.352817	1.984091	3.568907	2.305496
3.	Kurtosis	18.45431	8.944647	19.04195	10.9335
4.	Correlation	0.893523		0.877518	

Skewness is a measure that studies the degree and direction of departure from symmetry. Positive value of skewness indicates that the data is skewed to the right. Skewed right means that right tail is long relative to the left tail. Kurtosis parameter measures the peakedness (or flatness) of the probability distribution of any signal. Positive value of kurtosis indicates the strong intermittency in the PM2.5 and Pm10. Correlation describes the degree of linear relationship between two functions (or signals). The positive value of correlation means they are linearly related with positive slope and high value means that they are highly dependent.

Conclusion

Wavelet transforms provide excellent analysis of non-stationary time series and extracts important hidden information from a signal or data set. The wavelet method allows the decomposition of the signal into average and differential behaviour of a signal. The skewness parameter talks about the asymmetry and kurtosis about flatness of the given signal. The positive and high value of correlation coefficient represents directly and strongly correlation between PM2.5 and PM10 in the ambient air. For the data set of PM2.5 and PM10 of Moradabad for given time interval, Haar wavelet transform of level-1 is applied.

From the present analyses, we found that the PM2.5 and PM10 time-series of Moradabad in last one years are strong intermittent. Skewness & Kurtosis parameters are low and positive, and correlation is negative and high in that time period. With the help of extended signal, we can say that in the coming time, there will be little flatness or broadness in probability distribution, slight decrement in degree and direction of departure from symmetry. By virtue of these results, we can say that spectral analysis of PM2.5 and PM10 using stationary wavelet transforms provide a simple and accurate framework to investigate and forecast the behaviour of ambient air particle pollutants.

References

1. Mu CY, Tu YQ, Feng Y. Effect analysis of meteorological factors on the inhalable particle matter concentration of atmosphere in Hami. Meteorological Environmental Science. 2011;34:75-79.
2. Charlson RJ, Schwartz SE, Hales JM, Cess D, Coakley JA, Hansen JE. Climate forcing by anthropogenic aerosols. Science. 1992;255:423-430.
3. Flemming J, Stern R, Yamartino RJ. A new air quality regime classification scheme for O_3, NO_2, SO_2 and PM10 observations sites. Atmospheric Environment. 2005;39:6121-6129.
4. Cohen AJ. Outdoor air pollution and lung cancer. Environmental health perspectives. 2000;108:743-750.
5. Vanbilloen NH, Hoylaerts MF, Hoet PH, Verbruggen A, Nemery B. Passage of intratracheally instilled ultrafine particles from the lung into the systemic circulation in hamster. American Journal of Respiratory and Critical Care Medicine. 2001;164:1665-1668.
6. Bunde A, Havlin S, Koscielny-Bunde E, Schellnhuber HJ. Long-term persistence in the atmosphere: Global laws and tests of climate models. Physica A, Stat. Mech. Apps. 2001;302:255-267.
7. Antoine JP. Wavelet analysis: A new tool in Physics. Wavelets in Physics. Van Den Berg JC Ed.; c2004. p. 9-21.
8. Farge M. Wavelet transforms and their applications to turbulence. Annu. Rev. Fluid Mech. 1992;24:395-457.
9. Grinsted A, Moore JC, Jevrejeva S. Application of the cross wavelet transform and wavelet coherence to geophysical time series, Nonl. Proc. Geophy. 2004;11:561-566.
10. Rockinger M, Jondeau E. Entropy Densities with an Application to Autoregressive Conditional Skewness and Kurtosis. Journal of Econometrics. 2002;106(1):119-142.
11. Mallat SG. A wavelet tour of Signal Processing. Academic Press, New York; c1998.
12. Daubechies I. Ten lectures on wavelets. CBS-NSF, Regional Conference in Applied Mathematics. SIAM Philadelphia. 1992;61:278-285.
13. Kumar A, Rawat UK, Singh P. Spectral analysis of ground level nitrogen dioxide and sulphur dioxide. Progress in Nonlinear Dynamics and Chaos. 2018;6(1):11-17.
14. Kumar A. Wavelet analysis of air pollution due to ground

level nitrogen dioxide. Asian Journal of Science and Technology. 2017;8(8):11-17.
15. Kumar A, Rawat UK. Spectral analysis of rainfall using Haar wavelet transforms. International Journal of Higher Education and Research. 2019;9(1):319-332.
16. Nason GP, Silverman BW. The stationary wavelet transforms and some statistical applications. Wavelet and Statistics: Lecture Note in Statistics. Springer, New York. 1995;103:281-299.
17. El-Alosey AR, Eledum H. Discrete Erlang-truncated exponential distribution. International Journal of Statistics and Applied Mathematics. 2021;6(1):230-236.
18. Wong H, Ip W, Xie Z, Lui X. Modelling and forecasting by wavelets, and the application to exchange rates. Journal of Applied Statistics. 2003;30(5):537-553.
19. Xiaohong L, Shujuan Y, Li L. Prediction of temperature time series based on wavelet transform and support vector machine. Journal of Computers. 2012;7(8):1911-1918.
20. Rockinger M, Jondeau E. Entropy densities with an application to autoregressive conditional skewness and kurtosis. Journal of Econometrics. 2002;106(1):119-142.
21. Ye H, Wang G, Ding XS. A new parity space approach for fault detection based on stationary wavelet transform. IEEE Transactions on Automatic Control. 2004;49:281-287.