STOCHASTIC CHARACTERISTICS OF DE- NOISING TIME SERIES

By

Dr. Amira Ibrahim El-Desokey

Lecturer at Higher Future Institute for Specialized Technological Studies

aeldesokey@gmail.com

Scientific Journal for Financial and Commercial Studies and Research (SJFCSR)

Faculty of Commerce – Damietta University

Vol.4, No.2, Part 1., July 2023

APA Citation:

El-Desokey, A. I. (2023). Stochastic Characteristics of De- Noising Time Series, ***Scientific Journal for Financial and Commercial Studies and Research***, Faculty of Commerce, Damietta University, 4(2)1, 661-675.

Website: https://cfdj.journals.ekb.eg/

STOCHASTIC CHARACTERISTICS OF DE-NOISING TIME SERIES

Dr. Amira Ibrahim El-Desokey

Abstract:

Astrophysical, geophysical, meteorological, and other types of physical data can be the result of an experiment, emerge as a signal from a dynamical system, or include sociological, economic, or biological data. It is always assumed that a certain level of noise will be present in time series data, regardless of its source. The analysis of all such data in the presence of noise frequently produces inaccurate results. The time series data filtering technique is a tool to remove as many of these bugs as we can and simply prepare the data for additional analysis. When you filter a time series, you keep some frequencies while removing the spectral strength at others. Both time series analysis and digital signal processing (DSP) make extensive use of filters in applications for DSP. Here, we tried to create an adjustable method of filtering a time series with the idea that the series had performed the necessary de-noising and modification updates. These qualities were determined in this study. We've proven analytically that the current model can effectively resist errors and preserve positional significance in the time series. When such data are analyzed in a noisy environment, the results are frequently misinterpreted. So, before we can begin the extensive investigation, we must first build an initial platform by de-identifying the data. It is usually necessary to deal with the problem of filtering a time series of data. In this work, we analysis the Fourier series with respect to the finite Fourier transform, resulting in the determination of the properties of this series. The outcomes of this study improve understanding of this series' characteristics.

Keywords- asymptotes Normalization, disjointed time series, The Finite Fourier Transform, de- noising data, tapered data.

INTRODUCTION

Manipulative approaches are frequently used in time-series analysis. In specifically, the linear de-noising process looks like this if we take into account a process whose domain consists of a matrix of *(i)* vector-valued series $X(t), t \in R$ and whose range consists of a matrix of *(j)* vector-valued series $Y(t), t \in R$, the linear de- noising process is as follows:

$$Y(t) = \tau + \sum_{v=-\infty}^{\infty} \beta(t-v)X(t)$$

$$= \tau + \sum_{v=-\infty}^{\infty} \beta(v)X(t-v). \qquad (1.1)$$

Where $\beta(v), v \in R$, is a series of $j \times i$ matrices that achieves

$$\sum_{v=-\infty}^{\infty} |\beta(v)| < \infty, \qquad (1.2)$$

That the de-noising process was a $j \times i$ Summable, indicated by$\{\beta(v)\}$. De-noising process in equation (1.1) can be produced by a transfer function,

$$G(\omega) = \sum_{v=-\infty}^{\infty} \beta(v)e^{-i\omega t}, \ -\infty < \omega < \infty. \qquad (1.3)$$

Given that $f_{xx}(\omega), f_{yy}(\omega)$, are $i \times i \ and \ j \times j$ vector-valued matrices of the spectral density functions of $X(t) \ and \ Y(t)$, consequently, we have

$$f_{yy}(\omega) = G(\omega)f_{xx}(\omega)\overline{G(\grave{\omega})}, \qquad (1.4)$$

Assuming that the $j \times j$ Hermitian matrix is used to evaluate proximity,

$$E\{[Y(t) - \tau - \sum_{v=-\infty}^{\infty} \beta(t-v)X(v)][[\mathrm{Y(t)} - \text{°C} - \sum_{v=-\infty}^{\infty} \beta(t-v)X(v)]^T\}$$

where,

$$\text{°C} = C_Y - \left(\sum_{v=-\infty}^{\infty} \beta(v)\right)C_X$$

$$= C_Y - G(0)C_X,$$

$$\beta(v) = \frac{1}{2\pi}\sum_{\partial=0}^{2\pi} \beta(\partial) \text{ Exp } (iv\partial)d\partial,$$

where,

$$G(\omega) = f_{YX}(\omega)f_{XX}(\omega)^{-1}, \qquad (1.5)$$

where $G(\omega)$, also known as the complex regression coefficient of $Y(t)$on $X(t)$ at recurrence ω, is the transfer function of the $j \times i$ De-noising process that achieves the desired minimal $G(\omega)$, Therefore, given the constraint of Equation, we analyse the stochastic characteristics of the extended disjoined Fourier transform of Equation (1.1).

Types of Filters for Fourier Transform [11].

Filter for Low-Pass Signals Filter: Take this simulated time series, which consists of a linear trend and some random Gaussian noise shown in fig. (1).

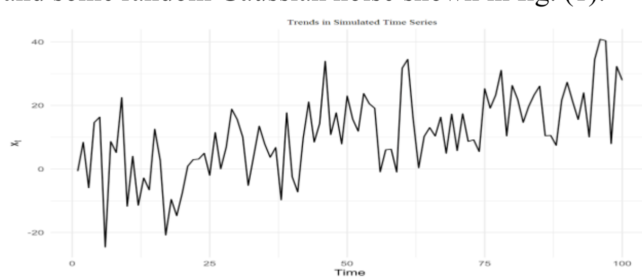

Fig (1)

Apply the linear filter illustrated in Fig. 2 as a convolver on the original series so, we get Fig (3).

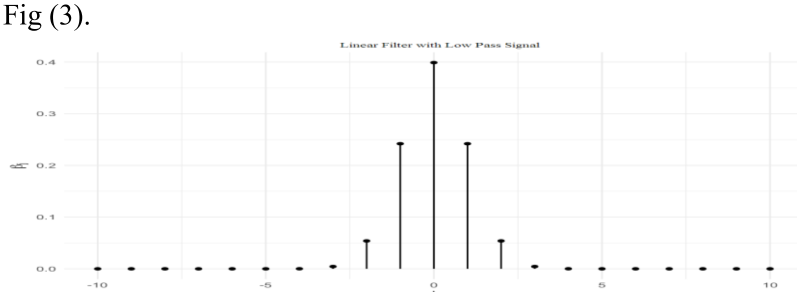

Fig (2)

A low-pass filter, as seen above, mutes higher frequencies in the data while letting lower frequencies "pass" through.

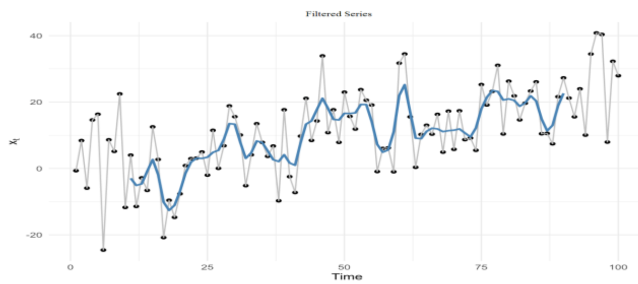

Fig (3)

The filtered series is displayed as a blue line, and it clearly represents a more purified version of the original series. We can observe that the higher frequencies are given less weight than the lower frequencies (close to f= 0).

Filter for High-Pass Signals: Here is an illustration of a linear filter that mutes low frequencies while letting high frequencies pass, fig (4) shows the high pass filter.

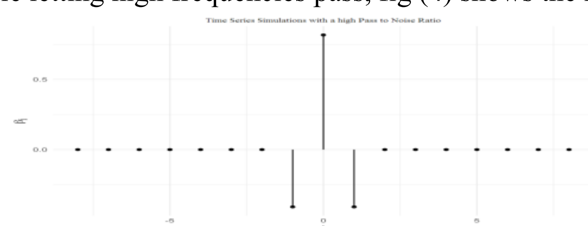

Fig (4)

Now, there is the high-pass filtered version of the original series. It appears to be a sequence of residuals from which the trend has been deleted, presented in Fig. (5).

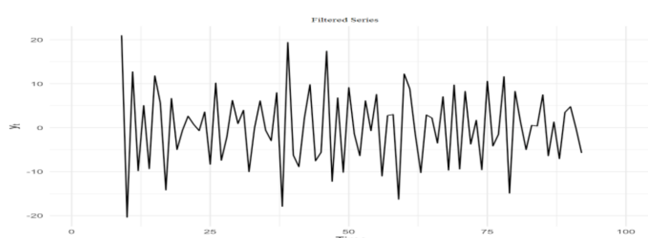

Fig (5)

Matching Filter: A time series with a visible jump at a particular point in time can be seen in the simulated series shown below in Fig (6).

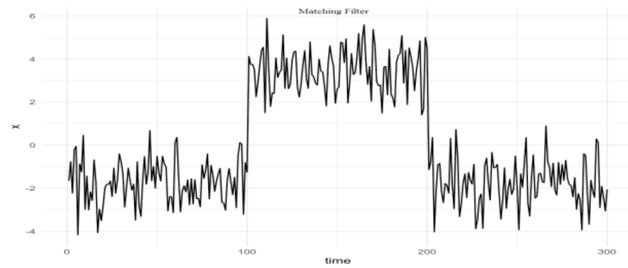

Fig (6)

We can determine when the series jump by employing a filter below in Fig. (7) , that reflects the jump in the data.

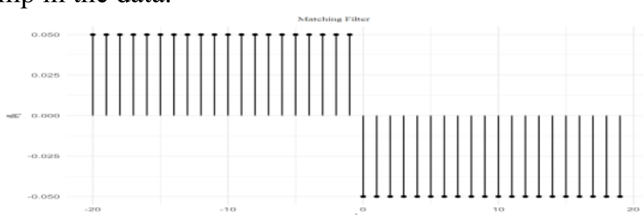

Fig (7)

The output of applying the matching filter to the data is shown in figure (8). The filtered series reaches a very high value at the data jump point and a very low value when the data move back down.

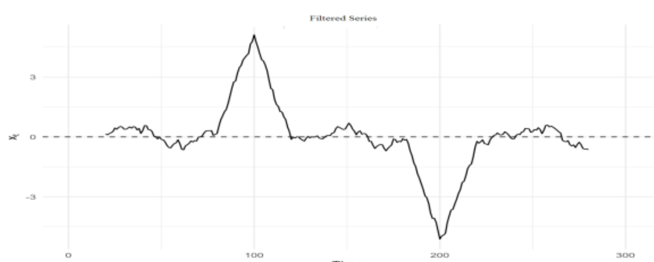

Fig (8)

Below is displayed the transfer function for the matching filter. The sines and cosines of the Fourier transform have difficulty approximating the discontinuous nature of this linear filter, as observed. Consequently, there is an oscillation in the transfer function. Fig9, 10, and fig 11 shows many examples of matching filter.

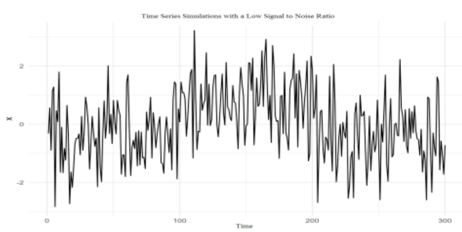

Fig (9)

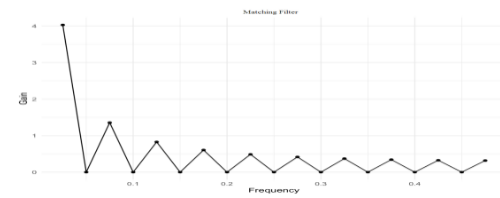

Fig (10)

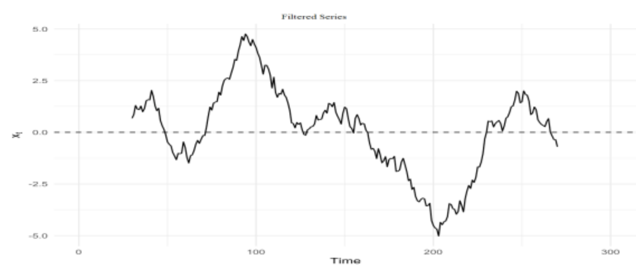

Fig (11)

Brillinger [1, 2], Ghazal, [3], Ghazal, et al, [7], Ghazal, et al [10], examined the stochastic properties of the disjoint Fourier transform in the regular case. Ghazal, et al [6], investigated the stochastic properties of the disjoint Fourier transform in the irregular case. Ghazal and Elhassanein [4], studied the analysis of Periodogram with missing observations. Ghazal, et al [5] investigated some properties of the discrete expanded finite fourier transform with missed observations. Ghazal, et al [8,9], examined the Periodogram analysis with missed observation between two vector valued stochastic process in continuous case. In this study the discrete extended Fourier transform was studied, assuming that the series had been De-noising. We studied the most important aspects beginning with Introduction, and then studying the Stochastic Characteristics of the modulated Series, and concluded with determining the Characteristics of the De-noising Extended Finite Fourier Transform.

The Stochastic Characteristics of the modulated Series:

Let $X(t)$,be an arbitrary static vector value. If $X(t)$ and $Y(t)$ have a linear association, then, we can represent the modulated series of Equation (1.1) as:

$$Y(t) = \tau + G(\omega)X(t), \qquad (2.1)$$

with,

$$Y_a(t) = \tau + G(\omega)X_a(t), \qquad (2.2)$$

Where, $G(\omega)$ is defined in Equation (1.5).

Consider that,

$$E\{[X(t+v) - C_x][Y(t) - C_Y]^T\} = C_{XY}(v), = \sum_{v=-\infty}^{\infty} f_{XY}(\omega)\exp(i\omega v), \qquad (2.3)$$

with spectrum density function,

$$f_{XY}(\omega) = (2\pi)^{-1}\sum_{v=-\infty}^{\infty} C_{XY}(v)Exp(-i\omega v), \quad -\infty < \omega < \infty. \qquad (2.4)$$

Characteristics of the De-noising Extended Finite Fourier Transform

Assumption I: Let's pretend $X(t)$ is a finite, static, random time series with all its moments. We have, for all $= 1,2,\dots,k-1$, and k tuple $a_1, a_2, \dots, a_k$, then we have,

$$\sum_{t_1\dots,t_{k-1}} \left|t_j C_{a_1,\dots,a_{k-1}}(t_1,\dots,t_{k-1})\right| < \infty, \quad k = 2,3,\dots. \qquad (3.1)$$

Because cumulants are measurements of the joint dependent on random variables, the spectral density can be defined as follows:

$$f_{a_1,\dots,a_k}(\omega_1,\omega_2,\dots,\omega_k) = (2\pi)^{k+1}\sum_{t_1,\dots,t_{k-1}} C_{a_1,\dots,a_k}(t_1,\dots t_{k-1})Exp(-i\sum_{j=1}^{k-1}\omega_j t_j),$$

$$-\infty < \omega < \infty, a_1,\dots a_k = 1,2,\dots,i; k = 2,3,\dots. \quad (3.2)$$

Assumption II. Let $l_a^{(T)}(t) = l_a(t/T)$ be a bounded function with bounded variations equal to zero outside interval [0, T]. Function $l_a^{(T)}(t)$ is defined as a data window; we construct the extended finite Fourier transform as:

$$\mathcal{H}_a^{(T)}(\omega) = \left[2\pi\sum_{t=0}^{T}\left(l_a^{(T)}(t)\right)^2\right]^{-1/2}\sum_{t=0}^{T} l_{a_1,\dots,a_k}^{(T)}(\omega)Y(t)\,exp\{-i\omega t\}, \quad -\infty < \omega < \infty, \qquad (3.3)$$

where,

$$Y(t) = G(\omega)X(t), \qquad (3.4)$$

Theorem 3.1: Given that $Y(t)$ is defined as equation (2.2), then we have

$$E(Y(t)) = 0$$

$$Var(Y_a(t)) = C_{aa}(0), \qquad (3.5)$$

$$Cov[Y_{a_1}(t_1), Y_{a_2}(t_2)] = G(\omega)G^{(T)}(\omega)C_{a_1a_2}(v), \qquad (3.6)$$

Proof.

From the independence of $X(t), and\ Y(t)$, and because they are arbitrary, static time series, then $E(Y(t)) = 0$.

Now,

$Var(Y_a(t))= Cov[Y_a(t), Y_a(t)]$

$$= Cov\ [\tau + \ G(\omega)X_a(t), \tau + \ G(\omega)X_a(t)]$$

$$= \ G(\omega)\ G^{(T)}(\omega)Cov[X_a(t), X_a(t)]$$

$$= \ G(\omega)\ G^{(T)}(\omega)Cov[t - t]$$

$$= \ G(\omega)\ G^{(T)}(\omega)C_{aa}(0).$$

Then equation (3.5) is obtained. Turning to

$$Cov\left[Y_{a_1}(t_1), Y_{a_2}(t_2)\ \right]$$

$$= \ Cov\ \left[\tau + \ G(\omega)X_{a_1}(t_1), \tau + \ G(\omega)X_{a_2}(t_2)\right]$$

$$= \ G(\omega)\ G^{(T)}(\omega)Cov\left[X_{a_1}(t_1), X_{a_2}(t_2)\right]$$

$$= \ \ G(\omega)\ G^{(T)}(\omega)C_{a_1a_2}[t_1 - t_2]$$

Substitute $t_1 - t_2 = v$, then

$$= G(\omega)G^{(T)}(\omega)C_{a_1a_2}(v)$$

And equation (3.6) is obtained, completing the proof.

Theorem 3.2

Let $X(t), (t = 0, \pm 1, \dots)$ be an arbitrary static *(i)*- vector valued, zero-mean time series, satisfying the first assumption. If $Y(t)$ is defined as equation (2.2) and $\mathcal{H}_a^{(T)}(\omega)$ is defined as equation (3.3), then we have:

$$E\left[\mathcal{H}_a^{(T)}(\omega)\right] = 0, \qquad (3.7)$$

$$Cov\left[\mathcal{H}_{a_1}(\omega_1), \mathcal{H}_{a_2}^{(T)}(\omega_2)\right] = \sum_{v=-T}^{T} G(\omega)\, G^{(T)}(\omega) \times$$

$$\times\, C_{a_1a_2}(v)\exp(-i\omega_1 v)\,\delta^{(T)}{}_{a_1a_2}(v, \omega_1 - \omega_2), \qquad (3.8)$$

$$= \sum_{u=-\infty}^{\infty} G(\omega)G^{(T)}(\omega)\, f_{a_1a_2}(u)\psi_{a_1a_2}^{(T)}(\omega_1 - u, \omega_2 - u), \qquad (3.9)$$

where,

$$\delta^{(T)}{}_{a_1a_2}(x) = \sum_{t=0}^{T} \mathcal{H}_a^{(T)}(t)\exp(-i\omega t),$$

$$\psi_{a_1a_2}^{(T)}(x, y) = (2\pi)^{-1}\left[\sum_{t_1=0}^{T}\sum_{t_2=0}^{T}\left(\mathcal{H}_{a_1}^{(T)}(t_1)\right)^2\left(\mathcal{H}_{a_2}^{(T)}(t_2)\right)^2\right]^{-1/2} \delta^{(T)}{}_{a_1}(x)\delta^{(T)}{}_{a_2}(y), \qquad (3.10)$$

Proof.

From the independence of $X(t), and\ Y(t)$, and because they are arbitrary static time series, then equation (3.7) is obtained.

Now,

$$Cov\left[\mathcal{H}_{a_1}^{(T)}(\omega_1), \mathcal{H}_{a_2}^{(T)}(\omega_2)\right] = (2\pi)^{-1}\left[\delta^{(T)}{}_{a_1a_2}(0)\right]^{-1} \times$$

$$\times \sum_{t_1=0}^{T}\sum_{t_2=0}^{T} l_{a_1}^{(T)}(t_1) l_{a_2}^{(T)}(t_2)\exp(-i(\omega_1 t_1 - \omega_2 t_2)\, Cov\left[(Y_{a_1}(t_1), Y_{a_2}(t_2)\right], \qquad (3.11)$$

Substitutes $t_1 - t_2 = v, t_2 = t$, then we have:

$$Cov\left[\mathcal{H}_{a_1}^{(T)}(\omega_1),\mathcal{H}_{a_2}^{(T)}(\omega_2)\right] = (2\pi)^{-1}\left[\delta^{(T)}{}_{a_1a_2}(0)\right]^{-1} \times$$

$$\times \sum_{v=-T}^{T} G(\omega)\, G^{(T)}(\omega) C_{a_1a_2}(v) \exp(-i\omega v) \times$$

$$\times \sum_{v=-T}^{T} l_{a_1}^{(T)}(t_1+v)\, l_{a_2}^{(T)}(t) \exp\left(-it(\omega_1-\omega_2)\right).$$

which satisfies equation (3.6). By substituting (2.3) into (3.8), we obtain

$$Cov\left[\mathcal{H}_{a_1}^{(T)}(\omega_1),\mathcal{H}_{a_2}^{(T)}(\omega_2)\right] = \sum_{u=-\infty}^{\infty} G(\omega) f_{a_1a_2}(u)\psi_{a_1a_2}^{(T)}(\omega_1-u,\omega_2-u),$$

And $\psi_{a_1a_2}^{(T)}$, is given by (3.10). Which completes the prove.

for $\omega_1=\omega_2$, and let $\omega-u=\lambda$, then

$$Var(\mathcal{H}_{a_1}^{(T)}(\omega)) = \sum_{\lambda=-\infty}^{\infty} f_{aa}(\omega-\lambda)\psi_{aa}^{(T)}(\lambda).$$

THEOREM 3.3 Let $h_1-h_2\neq 0, h_1, h_2 \in R$ and $\lambda_a^{(T)}(t), t\in R$, $a=1,\dots,\min(s,r)$ is bounded and the function $\psi_{aa}^{(T)}(x)$ defined as equation (28), then,

$$\lim_{T\to\infty} Cov\left[\mathcal{H}_{a_1}^{(T)}(\omega_1),\mathcal{H}_{a_2}^{(T)}(\omega_2)\right] = 0,$$

for all $a_1, a_2 = 1,\dots,\min(s,r)$

$$\lim_{T\to\infty} D\mathcal{H}_a^{(T)}(\omega) = G(\omega)G^{(T)}(\omega)f_{aa}(\omega)$$

Conclusion

De-noising time series were assumed, and the stochastic Characteristics of the discrete extended finite Fourier transform were explored. Random noise has a significant negative impact on forecasting in many fields such as Electricity, Economy, etc. Real-time purifying can significantly enhance measurement accuracy.

Conflicts of interest

There are no active conflicts.

Fund

This study received no specific funding from public, commercial, or not-for-profit funding entities.

References:

Akram Aldroubi, Murray Eden, and Michael Unser, 1994, Discrete Spline Filters For Multi resolutions and Wavelets of L2, SIAM J. Math Anal., Vol. 25, No. 5, pp.1412-1432.

Brillinger, D. R. 1969; Asymptotic properties of spectral estimates of second order; Springer, New York, NY. (pp. 179-194).

Brillinger D. R., 2001; Time series data analysis and theory; SIAM.

Ghazal, Mohamed A., 2001; Statistical analysis of branded periodogram for continuous time stationary processes. Journal of Applied Mathematics & Computation; 124: 343-349.

Ghazal, M. A., and A. Elhassanein, 2006; Periodogram analysis with missing observations, Journal of Applied Mathematics and Computing; 22.1, 209-222.

Ghazal M. A., Farag E. A., and El-Desoky E. A. 2005; Some properties of the discrete expanded finite fourier transform with missed observations; Journal of Ain Shames University, Faculty of Engineering, 40(3).

Ghazal M. A, Mokaddis G. S., Farag E. A., and El-desoky E. A., 2010; Asymptotic properties of spectral estimates of second –order with missed observations; International Journal of Applied Mathematics 22(2):161-178.

Ghazal M. A., El-Desoky E. A., and Alargt M. A., 2017; Asymptotic properties of the discrete stability time series with missed observations between tow-vector valued stochastic process; International Journal of Engineering and Technology. 4 (4).

Ghazal M. A., A. I. El-desokey and A. M. Ben Aros.; 2017; Periodogram analysis with missed observation between two vector valued stochastic process; Int. J. of Adv. Res. 5 (Nov). 336-349] (ISSN 2320-5407).

Ghazal M. A., El-Desoky E. A., and Alargt M. A., 2017; Statistical properties of the periodogram for tow-vector valued stability series with missed observations, International Journal of Statistics and Applied Mathematics, 4 (4).

Huang, N.E. and Zhaohua, W.: “An Adaptive Data Analysis Method for Nonlinear and Nonstationary Time Series: The Empirical Mode Decomposition and Hilbert Spectral Analysis”, http://www.deas.harvard.edu/climate/pdf/zhaohua.pdf, (2005).

الخصائص الاحصائية لإزالة الضوضاء من السلاسل الزمنية

د. أميرة إبراهيم الدسوقي

الملخص العربي:

يمكن أن تكون البيانات الفيزيائية الفلكية والجيوفيزيائية والأرصاد الجوية وأنواع أخرى من البيانات الفيزيائية تظهر نتيجة لتجربة معينة، أو تظهر كإشارة من نظام ديناميكي في شكل سلسلة زمنية. من المفترض دائمًا وجود مستوى معين من الضوضاء في بيانات هذه السلاسل الزمنية، بغض النظر عن مصدرها. غالبًا ما ينتج عن تحليل جميع هذه البيانات في وجود ضوضاء نتائج غير دقيقة. تعد تقنية تصفية بيانات السلاسل الزمنية من الضوضاء أداة لإزالة أكبر عدد ممكن من هذه الأخطاء، عندما تقوم بتصفية او بإزالة الضوضاء من سلسلة زمنية، فإنك تحتفظ ببعض الترددات بينما تزيل القوة الطيفية في أخرى. يستخدم كل من تحليل السلاسل الزمنية ومعالجة الإشارات الرقمية (DSP) بشكل مكثف المرشحات في تطبيقات (DSP) ،حاولنا هنا إنشاء طريقة معدلة لتصفية سلسلة زمنية من الضوضاء مع فكرة أن السلسلة قد أجرت التحديثات اللازمة لإزالة الضوضاء والتعديل. تم تحديد هذه التعديلات في هذه الدراسة. لقد أثبتنا بشكل تحليلي أن النموذج الحالي يمكن أن يقاوم الأخطاء بفاعلية ويحافظ على الأهمية الموضعية في السلسلة الزمنية. عندما يتم تحليل هذه البيانات في بيئة مليئة بالضوضاء او الاخطاء، كثيرا ما تكون النتائج غير دقيقة. لذلك، قبل أن نتمكن من بدء الدراسة الشاملة، يجب علينا أولاً إنشاء تحليل اولى عن طريق تحديد البيانات المراد تنقيتها او تصفيتها. عادة ما يكون من الضروري التعامل مع مشكلة تصفية سلسلة زمنية من البيانات. في هذا العمل، قمنا بتحليل سلسلة فورييه المحدود، مما أدى إلى تحديد خصائص هذه السلسلة. نتائج هذه الدراسة تحسن فهم خصائص سلسلة فورييه المحدود.

الكلمات المفتاحية: التوزيع الطبيعي المقارب، السلاسل الزمنية المتقطعة، تحويل فورييه المحدود، تصفية البيانات.