Expectile Periodogram

Tianbo Chen[1]

Abstract

This paper introduces a novel periodogram-like function, called the expectile periodogram, for modeling spectral features of time series. The expectile periodogram is constructed from trigonometric expectile regression, in which a specially designed cost function is used to substitute the squared l_2 norm that leads to the ordinary periodogram. The expectile periodogram retains the key properties of the ordinary periodogram as a frequency-domain representation of serial dependence in time series, while offering a more comprehensive understanding by examining the data across the entire range of expectile levels. Simulations and real-world data demonstrate the efficiency of the expectile periodogram in the presence of hidden periodicities. Finally, by leveraging the inherent two-dimensional nature of the expectile periodogram, we train a deep learning (DL) model to classify earthquake waveform data. Remarkably, our approach outperforms alternative periodogram-based methods in terms of classification accuracy.

Keywords: Periodogram; Expectile regression; Spectral density; Time series analysis; Earthquake data classification

1. Introduction

Spectral density functions (SDFs) play a crucial role in time series analysis, where data are analyzed in the frequency domain. The ordinary periodogram, a raw non-parametric estimator of the SDF, is widely applied across various fields (Caiado et al., 2006; Polat and Güneş, 2007; Baud et al., 2018; Euán et al.,

[1]Anhui University, China. E-mail: chentianbo@ahu.edu.cn

2018; Maadooliat et al., 2018; Martínez-Murcia et al., 2019; Chen et al., 2021a). The ordinary periodogram is constructed by the ordinary least squares (OLS) regression on the trigonometric regressor. However, OLS exhibits limitations in terms of robustness and effectiveness, particularly in handling data with asymmetric or heavy-tailed distributions. Moreover, OLS focuses primarily on the conditional mean, lacking the ability to provide a comprehensive view of the data, which subsequently affects the performance of the ordinary periodogram (Bloomfield, 2004).

An alternative regression approach is quantile regression, which evaluates the variation of conditional quantiles with respect to the response variable. The pioneering work of Koenker and Bassett Jr (1978) introduced the concept of quantile regression, and it has been comprehensively extended by Koenker (2005). For a detailed and systematic introduction to quantile regression and its extensions, we refer to Kouretas et al. (2005); Cai and Xu (2008); Cai and Xiao (2012); Koenker (2017). Equipped with a specially designed loss function, quantile regression provides a more complete picture of the relationship between the response variable and the covariates, while also demonstrating strong robustness against outliers. As a result, quantile regression has been applied in various fields (Garcia et al., 2001; Machado and Mata, 2005; Alvarado et al., 2021; Sharif et al., 2021). One innovative development of quantile regression is the quantile periodogram (Li, 2012b) constructed by trigonometric quantile regression, which demonstrates its ability to detect hidden periodicities in time series. Similarly to the behavior of the SDF and the ordinary periodogram, the quantile periodogram is an unbiased estimator of the so-called quantile spectrum, which is a scaled version of the ordinary SDF of the level-crossing process. Notably, the Laplace periodogram (Li, 2008) represents a specialized case of the quantile periodogram when the quantile is set to 0.5. Related works on the quantile periodogram include Li (2012a); Hagemann (2013); Li (2014); Dette et al. (2015); Kley (2016); Birr et al. (2017); Meziani et al. (2020); Chen et al. (2021b); Li (2023).

However, quantile regression is beset by the computational burden due to

the non-differentiability of its loss function. To address the issue and strike a balance between robustness and effectiveness, the concept of expectile regression has been proposed. Asymmetric least square (ALS) regression, also known as expectile regression, was proposed by Newey and Powell (1987). Expectile regression can capture the entire distribution of the data and is much easier to compute using weighted least squares estimators. Furthermore, consistent estimation of the joint asymptotic covariance matrix of several ALS estimators does not require estimation of the density function of the error terms. A comprehensive comparative analysis of quantiles and expectiles is presented in Waltrup et al. (2015), wherein the relationships between these two approaches are thoroughly examined. In addition, Jones (1994) provided a mathematical proof that expectiles indeed correspond to quantiles of a distribution function uniquely associated with the underlying data distribution. Building on this foundation, Yao and Tong (1996) established the existence of a bijective function that directly relates expectiles to quantiles, thereby facilitating the calculation of one from the other. Alternative approaches for estimating quantiles from expectiles are introduced in Efron (1991); Granger and Sin (1997); Schnabel and PHC (2009), offering methodologies for estimating the density (and also, quantiles) from a set of expectiles by using penalized least squares. As a generalization of both quantile and expectile regression, Jiang et al. (2021) introduced the k-th power expectile regression with $1<k \leq 2$.

Expectile regression techniques are applied across diverse domains. For instance, Jiang et al. (2017) introduced an expectile regression neural network (ERNN) model, which incorporates a neural network structure into expectile regression, thereby facilitating the exploration of potential nonlinear relationships between covariates and the expectiles of the response variable. Gu and Zou (2016) systematically studied the sparse asymmetric least squares (SALES) regression under high dimensions where the penalty functions include the Lasso and nonconvex penalties. Xu et al. (2020) developed a novel mixed data sampling expectile regression (MIDAS-ER) model to measure financial risk and demonstrated exceptional performance when applied to two popular financial

risk measures: value at risk (VaR) and expected shortfall (ES). Another approach by Xu et al. (2021) involved the elastic-net penalty into expectile regression, and applied the model to two real-world applications: relative location of CT slices on the axial axis and metabolism of tacrolimus drug.

Inspired by the notable success of expectile regression and the foundational work of the quantile periodogram, we propose a novel spectral estimator termed the expectile periodogram, for spectral analysis of time series data. This paper demonstrates that the expectile periodogram not only shares similar properties of the ordinary periodogram as a frequency-domain representation of serial dependence in time series, but also provides a more comprehensive analysis by exploring the full range of expectile levels. The rest of the paper is organized as follows. In Section 2, we define the expectile periodogram and demonstrate its capability in detecting hidden periodicities through two real-world examples. In Section 3, we present comparative studies evaluating the performance of different periodograms by simulations. In Section 4, we apply our method to a time-series classification task, leveraging the two-dimensional nature of the expectile periodogram and training a deep learning (DL) model to classify the earthquake data. Conclusions and future works are discussed in Section 5.

2. Expectile Periodogram

2.1. Expectile Regression Approach

For a real-valued time series $\{Y_1, Y_2, ..., Y_n\}$, the ordinary periodogram, which serves as a raw estimator of the spectral density, is defined as:

$$I_n(\omega) = \frac{1}{n}\left|\sum_{t=1}^{n} Y_t e^{-jt\omega}\right|^2, \tag{1}$$

where $\omega \in (0, \pi)$. When ω takes values at the Fourier frequencies, $\omega_j = 2\pi j/n, j = 1, ..., q$, and $q = [n/2]$, the ordinary periodogram takes the alternative form:

$$I_n(\omega) := \frac{1}{4} n ||\bar{\boldsymbol{\beta}}_n(\omega)||_2^2, \tag{2}$$

where $\bar{\boldsymbol{\beta}}_n(\omega)$ is obtained via the ordinary least squares (OLS) method:

$$\bar{\boldsymbol{\beta}}_n(\omega) = \underset{\boldsymbol{\beta}\in\mathbb{R}^2,\mu\in\mathbb{R}}{\arg\min} \sum_{t=1}^{n} \frac{1}{2} \left| Y_t - \lambda - \mathbf{x}_t^\top(\omega)\boldsymbol{\beta} \right|^2,$$

with λ denotes the sample mean, $\mathbf{x}_t(\omega) = \{\cos(\omega t), \sin(\omega t)\}^\top$ denotes the trigonometric regressor, and $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2]^\top \in \mathbb{R}^2$. Consider the Fourier series representation of $\{Y_t\}$:

$$Y_t = \lambda + \sum_{j=1}^{q} \beta_{1j} \cos(\omega_j t) + \beta_{2j} \sin(\omega_j t).$$

By fitting this model using OLS at each frequency ω_j, the equivalence of (1) and (2) can be directly established.

OLS method focuses on the conditional mean of the response variable given the predictors. In contrast, expectile regression offers a more comprehensive view of the data by examining a range of conditional expectiles. Given an expectile level $\tau \in (0, 1)$, the asymmetric least squares (ALS) (Newey and Powell, 1987) cost function follows

$$\rho_\tau(u) = |\tau - I(u < 0)| \cdot u^2,$$

where $I(\cdot)$ denotes the indicator function.

Consider the linear trigonometric expectile regression solution:

$$\hat{\boldsymbol{\beta}}_{n,\tau}(\omega) = \underset{\boldsymbol{\beta}\in\mathbb{R}^2,\lambda_\tau\in\mathbb{R}}{\arg\min} \sum_{t=1}^{n} \rho_\tau\{Y_t - \lambda_\tau - \mathbf{x}_t^\top(\omega)\boldsymbol{\beta}\}, \tag{3}$$

where λ_τ is a suitable constant, typically the τ-expectile of $\{Y_t\}$, given by:

$$\lambda_\tau = \underset{\lambda\in\mathbb{R}}{\arg\min} \sum_{t=1}^{n} \rho_\tau(Y_t - \lambda).$$

Aigner et al. (1976) demonstrated that $\hat{\boldsymbol{\beta}}_{n,\tau}(\omega)$ can be interpreted as a maximum likelihood estimator when the disturbances arise from a normal distribution with unequal weight placed on positive and negative disturbances. Then, we define the expectile periodogram at expectile level τ as:

$$EP_{n,\tau}(\omega) := \frac{1}{4} n ||\hat{\boldsymbol{\beta}}_{n,\tau}(\omega)||_2^2 = \frac{1}{4} n \hat{\boldsymbol{\beta}}_{n,\tau}^\top(\omega)\hat{\boldsymbol{\beta}}_{n,\tau}(\omega).$$

In other words, the expectile periodogram is a scaled quadratic form of the expectile regression coefficients fitted using the trigonometric basis at frequency ω. While the parameter λ_τ is fixed in (3), it can also be optimized alongside $\boldsymbol{\beta}(\omega)$ to obtain the extended expectile regression solution:

$$\{\hat{\lambda}_\tau(\omega), \hat{\boldsymbol{\beta}}_{n,\tau}(\omega)\} := \underset{\boldsymbol{\beta}\in\mathbb{R}^2, \lambda_\tau\in\mathbb{R}}{\arg\min} \sum_{t=1}^{n} \rho_\tau\{Y_t - \lambda_\tau - \mathbf{x}_t^\top(\omega)\boldsymbol{\beta}\}.$$

Another extension of the ordinary periodogram is the quantile periodogram proposed in Li (2012b), where the OLS estimator in (2) is replaced by the quantile regression estimator:

$$\tilde{\boldsymbol{\beta}}_{n,\alpha}(\omega) := \underset{\boldsymbol{\beta}\in\mathbb{R}^2, \lambda_\tau\in\mathbb{R}}{\arg\min} \sum_{t=1}^{n} \rho^*_\alpha\{Y_t - \lambda_\alpha - \mathbf{x}_t^\top(\omega)\boldsymbol{\beta}\},$$

where $\rho^*_\alpha(u) = u\{\alpha - I(u < 0)\}$ is the check loss function. The quantile periodogram at quantile level $\alpha \in (0, 1)$ is then defined as

$$QP_{n,\alpha}(\omega) := \frac{1}{4} n ||\tilde{\boldsymbol{\beta}}_{n,\alpha}(\omega)||_2^2 = \frac{1}{4} n \tilde{\boldsymbol{\beta}}_{n,\alpha}^\top(\omega)\tilde{\boldsymbol{\beta}}_{n,\alpha}(\omega).$$

When $\alpha = 0.5$, the quantile periodogram reduces to the Laplace periodogram introduced in Li (2008).

2.2. Examples

While the expectile periodogram can be analyzed at a specific expectile level like the ordinary periodogram, it can also be treated as a bivariate function of parameters τ and ω. To illustrate such duality, we consider two examples in different scientific fields, in which the ordinary periodograms fail to capture the full information of the time series. Noticed that the main purpose of this paper is to analyze the serial dependence of the time series, with amplitude considerations being secondary. Therefore, for the remainder of the paper, we normalize the periodograms at each expectile level such that the summation over ω equals 1. Additionally, we use the normalized frequency $f = \omega/2\pi \in (0, 1/2)$, which denotes the number of cycles per unit time, for illustration in the figures.

In the first example, we analyze the daily log returns of the S&P 500 Index data from 1986 to 2015, as shown in Figure 1(a). The expectile periodogram

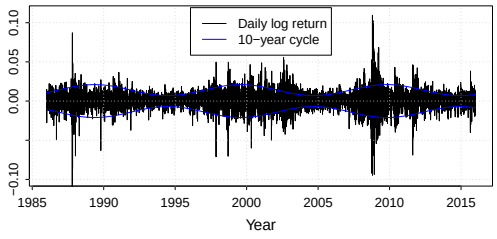
in Figure 1(b) successfully identifies the approximately 10-year cycle of market volatility (highlighted by the blue lines in Figure 1(a)) at both the lower and upper expectiles. In contrast, the ordinary periodogram yields a relatively featureless flat line. The smoothed periodograms in Figure 1(d) and (e), obtained using the smooth.spline function in R with the tuning parameter selected by generalized cross-validation (GCV), exhibit a spectral feature similar to a GARCH(1,1) (Engle, 1982; Bollerslev, 1986) model:

$$Y_t \sim N(0, \sigma_t^2), \tag{4}$$

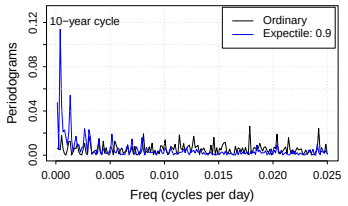
where $\sigma_t^2 = 10^{-6} + 0.35Y_{t-1}^2 + 0.35\sigma_{t-1}^2$. The averaged ordinary and expectile periodograms, computed from 5,000 independent realizations of model (4), are shown in Figure 1(f) and (g), respectively. Notably, while the ordinary periodogram of model (4) is a constant, the expectile periodograms detect the low-frequency feature at both the lower and higher expectiles.

The second example is based on electroencephalogram (EEG) data collected from an epilepsy patient during a seizure interval[2]. Figure 2(a) presents the EEG segments along with their corresponding sample expectiles. The ordinary periodogram in Figure 2(b) successfully detects both low-frequency and high-frequency components, corresponding to the six main spikes and the associated bursts, respectively. The expectile periodograms in Figure 2(b) and (c) provide additional information, demonstrating that the high-frequency patterns appear predominantly at higher expectile levels. This observation is consistent with the EEG data, where the bursts primarily occur at the peaks of the main spikes. Furthermore, the intensity of the bursts increases with higher expectiles, particularly at frequencies $\omega \in [0.02, 0.03]$, as shown in Figure 2(d).

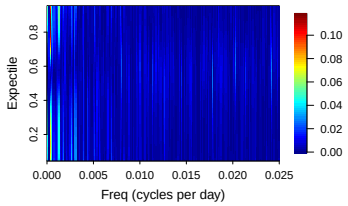
[2]The data was collected by South China Normal University School of Psychology and is authorized for use in this paper.



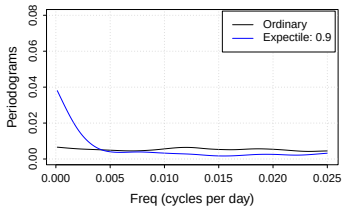
(a)



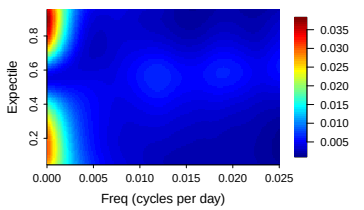
(b)



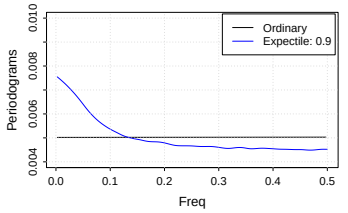
(c)



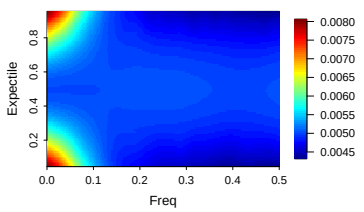
(d)



(e)



(f)



(g)

Figure 1: (a) Daily log returns of the S&P 500 Index data; (b) the ordinary periodogram and the expectile periodogram at expectile $\tau = 0.9$; (c) the expectile periodogram at expectiles $\{0.05, 0.06, ..., 0.95\}$; (d) and (e) the smoothed versions of (b) and (c), respectively; (f) the averaged ordinary periodogram and the averaged expectile periodogram at expectile $\tau = 0.9$ of model (4); and (g) the averaged expectile periodogram of model (4) at expectiles $\{0.05, 0.06, ..., 0.95\}$. (f) and (g) are ensemble means of 5,000 realizations.

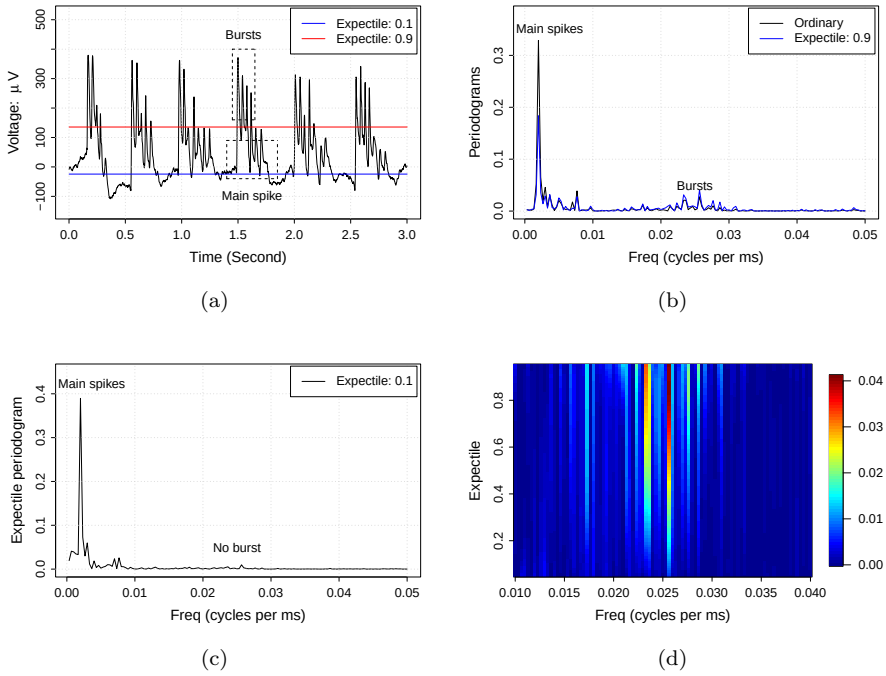


Figure 2: (a) The EEG data over a 3-second interval, along with its sample expectiles 0.1 and 0.9; (b) the ordinary periodogram and the expectile periodogram at $\tau = 0.9$; (c) the expectile periodogram at $\tau = 0.1$; and (d) the expectile periodogram at expectiles $\{0.05, 0.06, ..., 0.95\}$. The time resolution of the EEG data is 1000 Hz.

3. Numerical Results

In this section, we present simulations to demonstrate the efficiency of the expectile periodogram in detecting hidden periodicities in time series.

3.1. Hidden Periodicities Detection

We consider the following model:

$$Y_t = a_t X_t, \tag{5}$$

where

$$a_t = b_0 + b_1 \cos(\omega_0 t) + b_2 \sin(\omega_1 t),$$

with $b_0 = 1, b_1 = 0.9, b_2 = 1$. $\{X_t\}$ is an AR(2) process satisfying

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t, \tag{6}$$

where $\phi_1 = 2r\cos(\omega_c), \phi_2 = -r^2$ $(r = 0.6)$ and $\{\epsilon_t\}$ is the noise. Additionally, we set $\omega_0 = 0.1 \times 2\pi, \omega_1 = 0.12 \times 2\pi$. This setup aims to evaluate the effectiveness of different types of periodograms in detecting multiple closely spaced periodicities.

We first present the periodograms of model (6) with $\omega_c = 2\pi \times 0.25$ and $2\pi \times 0.3$, which no hidden periodicities are present. As shown in Figure 3(a) and 3(b), the ordinary periodograms, the expectile periodograms and the Laplace periodograms, exhibit similar bell-shaped patterns, with spectral peaks around ω_c.

The results is expected for the ordinary spectral density of the AR(2) process. The characteristic polynomial of model (6) is $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$, where the roots z_1 and z_2 are complex conjugates. The magnitude $|z_1| = |z_2| = 1/r > 1$ ensures causality. The AR coefficients ϕ_1 and ϕ_2 determine the location and narrowness of the spectral peak. Specifically, the peak frequency is located at $\omega_c/2\pi$ and as $r \to 1^-$, the peak becomes narrower (Shumway and Stoffer, 2016). Figure 3(c) and (d) show the expectile periodograms at expectiles $\{0.05, 0.06, ..., 0.95\}$, which share similar features to the quantile spectrum in Chen et al. (2021b). Specifically, all periodograms shown are ensemble means of 5,000 smoothed periodograms.

Figure 4 illustrates the ability of the expectile periodogram in detecting hidden periodicities. We present the ensemble mean of 5,000 realizations of the ordinary and expectile periodograms of model (5). The expectile periodograms detect the hidden periodicities, manifesting as large spikes at ω_0 and ω_1, whereas the ordinary periodograms do not exhibit this capability. It is important to note that spectral leaks, which have been observed for the Laplace periodogram, can also occur in the expectile periodogram. Therefore, small spikes may take place at some other frequencies, as indicated by Theorem 2 in Li (2012b). To address the issue of spectral leakage, incorporating an l_1 regularization into the loss function is beneficial (Meziani et al., 2020).

The expectile periodograms in Figure 1(g) and Figure 3(c), (d) are symmetric across the expectile levels. To demonstrate the capability of the expectile periodogram in handling more complex spectral features, we construct Y_t, whose

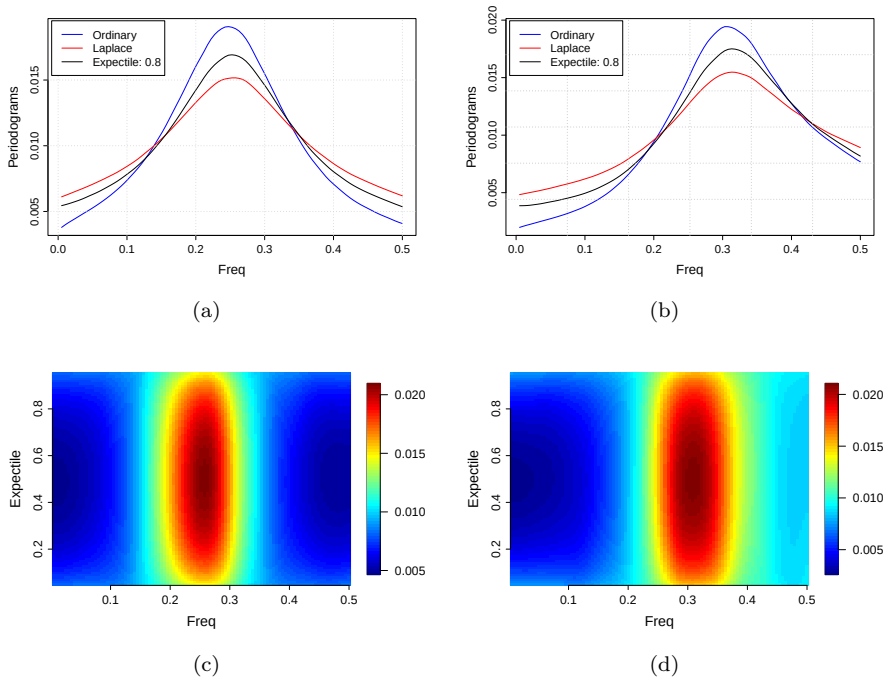


Figure 3: Ensemble means of three types of smoothed periodograms of time series defined by model (6) with standard Gaussian white noise. (a) Periodograms with $\omega_c = 2\pi \times 0.25$, (b) periodograms with $\omega_c = 2\pi \times 0.3$, (c) and (d) the expectile periodograms at expectiles $\{0.05, 0.06, ..., 0.95\}$ for $\omega_c = 2\pi \times 0.25$ and 0.3, respectively. The number of realizations is 5,000 and the sample size $n = 200$.

expectile periodogram is intentionally made asymmetric across the expectile levels. Y_t is defined as a nonlinear mixture of three components:

$$Z_t = W_1(X_{t,1})\, X_{t,1} + \{1 - W_1(X_{t,1})\}\, X_{t,2},$$

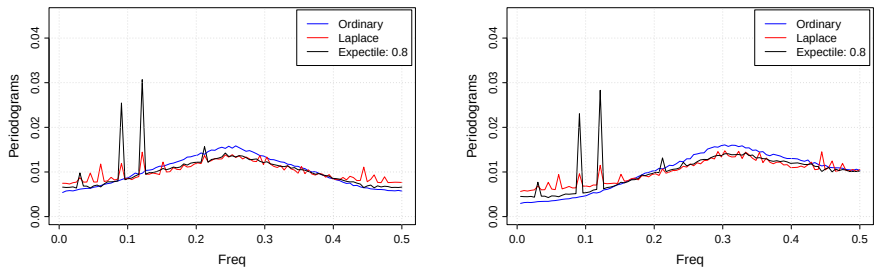
$$Y_t = W_2(Z_t)\, Z_t + \{1 - W_2(Z_t)\}\, X_{t,3}. \tag{7}$$

The components $\{X_{t,1}\}$, $\{X_{t,2}\}$, and $\{X_{t,3}\}$ are independent Gaussian AR(1) processes satisfying

$$X_{t,1} = 0.8X_{t-1,1} + w_{t,1},$$

$$X_{t,2} = -0.75X_{t-1,2} + w_{t,2},$$

$$X_{t,3} = -0.81X_{t-2,3} + w_{t,3},$$



(a) (b)

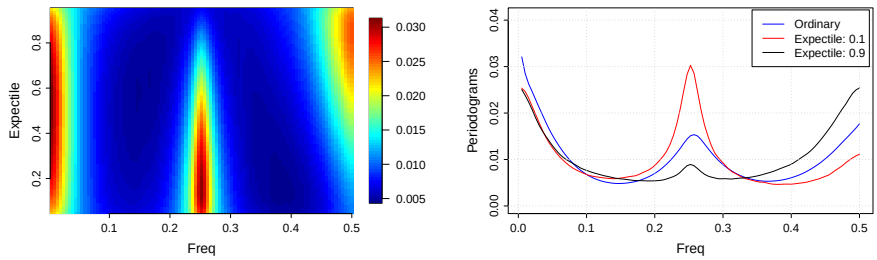
Figure 4: Means of the three types of periodograms of time series defined by model (5). (a) Represents $\omega_c = 0.25$ and (b) represents $\omega_c = 0.3$, respectively. The number of realizations is $5,000$ and $n = 200$.

where $w_{t,1}, w_{t,2}, w_{t,3}$ are standard Gaussian white noise. From the perspective of traditional spectral analysis, the series $\{X_{t,1}\}$ has a lowpass spectrum, $\{X_{t,2}\}$ has a highpass spectrum, and $\{X_{t,3}\}$ has a bandpass spectrum around frequency $1/4$. The mixing function $W_1(x)$ is equal to 0.9 for $x < -0.8$, 0.2 for $x > 0.8$, and linear transition for x in between. The mixing function $W_2(x)$ is similarly defined except that it equals 0.5 for $x < -0.4$ and 1 for $x > 0$. Figure 5(a) shows the expectile periodograms of model (7), where the expectile periodogram is asymmetric across the expectile levels. Figure 5(b) compares the ordinary periodograms and the expectile periodograms at $\tau = 0.1$ and 0.9, illustrating that the ordinary periodogram is restricted to capturing spectral features near the central expectiles. In contrast, the expectile periodogram offers a broader perspective, effectively analyzing the time series across the entire range of $\tau \in (0, 1)$.

3.2. Fisher's Test

One commonly used hypothesis test to detect periodicities for the ordinary periodogram is Fisher's test (Brockwell and Davis, 1991). For frequencies $\{\omega_1, \omega_2, ..., \omega_q\}$, the test statistic is defined as

$$g = \frac{\max_{1 \le k \le q}\{I_n(\omega_k)\}}{\sum_{k=1}^{q} I_n(\omega_k)}.$$



(a) (b)

Figure 5: The periodograms of the mixture model (7). (a) The expectile periodogram with asymmetric pattern across the expectile level, and (b) the expectile periodograms at extreme expectiles ($\tau = 0.1$ and 0.9), as well as the ordinary periodogram. The number of realizations is 5,000 and the sample size $n = 200$.

According to Fisher's test, a sufficiently large value of g indicates the presence of a hidden periodicity. The null hypothesis is that the time series is Gaussian white noise, while the alternative hypothesis is that the time series contains a deterministic periodic component of unspecified frequency.

We apply Fisher's test to the expectile periodogram by replacing $I_n(\omega)$ with $EP_{n,\tau}(\omega)$ in the test statistic. The probabilities of detection are obtained by Monte Carlo simulation runs for time series defined by model (5), with a single periodicity at $\omega_0 = 0.1 \times 2\pi, \omega_c = 0.3 \times 2\pi$, and $b_2 = 0$. As can be seen in Table 1, both the expectile and quantile periodograms outperform the ordinary periodogram. At a significance level of 0.05, the expectile periodogram ($\tau = 0.9$) achieves a detection rate of 84.26%, whereas the ordinary periodogram only achieves only 29.78%. Detection rates are sensitive to the expectile level, showing a clear trend in this experiment. As the expectile or quantile approaches 1, the detection rate of the expectile periodogram increases and eventually surpasses that of the quantile periodogram.

3.3. Smoothed Expectile Periodogram

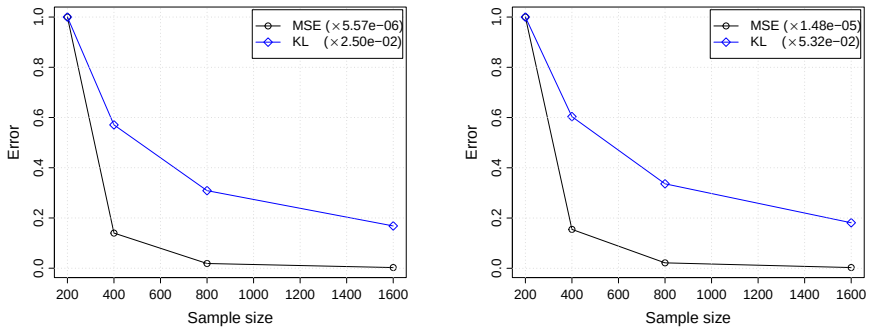
SDF can be estimated consistently by a properly smoothed ordinary periodogram Brockwell and Davis (1991). However, the absence of analytical expres-

Table 1: Fisher's test of different types of periodograms.

Significance level	Expectile periodogram τ=0.85	τ=0.9	τ=0.95	Quantile periodogram α=0.85	α=0.9	α=0.95	Ordianry periodogram
0.01	0.4048	0.5608	0.5850	0.6898	0.6328	0.4428	0.1224
0.05	0.7158	0.8426	0.8510	0.8720	0.8260	0.6646	0.2978
0.10	0.8308	0.9262	0.9306	0.9278	0.8952	0.7678	0.4258

sions for the mean and variance of the expectile periodogram makes it difficult to conduct a similar investigation. To address this, we employ simulations to demonstrate that the distance between the smoothed expectile periodogram and the ground truth (the average of 5,000 expectile periodograms) decreases as n increases. We use models (6) and (7) at $\tau = 0.9$ for illustration.

We measure the distance using the mean squared error (MSE) and Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951). Figure 6 depicts the MSE and KL divergence as a function of n, with n taking values of 200, 400, 800, and 1,600. To highlight the decreasing trend as n increases, the values are normalized by setting the maximum to 1.



(a) (b)

Figure 6: MSE and KL divergence of smoothed periodogram. (a) AR(2) model (6); (b) mixture model (7). The results are based on 5,000 simulation runs.

4. Earthquake Data Classification

In this section, we apply the expectile periodogram to an earthquake classification problem. Section 4.1 introduces the earthquake data and Section 4.2 describes the deep learning model and presents the classification results.

4.1. Data Description

The earthquake waveform data with a sampling rate of 100 Hz, was collected in February 2014 in Oklahoma State. This data is available at `https://www.iris.edu/hq/` and `http://www.ou.edu/ogs.html`. Details about the catalog data are provided in Benz et al. (2015), including labels for earthquake magnitudes and occurrence times. We extract 2,000 non-overlapping segments, each being a time series with $n = 2,000$, corresponding to 20 seconds of data. The choice of longer time series ensures that the segments contain complete earthquake events. Among these segments, 1,000 of them contain an earthquake with magnitudes greater than 0.25, while the remaining 1,000 segments contain no earthquakes. We smooth the expectile periodograms of the 2,000 segments using the semi-parametric method proposed in Chen et al. (2021b), ensuring smoothness across both the expectile and frequency dimensions. In this experiment, we use the lower half of the frequencies ($\omega \leq 0.25 \times 2\pi$) and 46 expectiles $\{0.05, 0.07, ..., 0.93, 0.95\}$. Since we focus on serial dependence and normalize the expectile periodograms, amplitude considerations are excluded, making the classification more challenging. Additionally, we incorporate two competitive periodograms: the smoothed ordinary periodogram and the quantile periodogram.

We show three representative segments along with their corresponding smoothed expectile periodograms in Figure 7. Specifically, Figure 7(a) contains a large earthquake with a magnitude larger than 3, Figure 7(b) contains a somehow small earthquake with a magnitude less than 1, and Figure 7 (c) contains no earthquake. Based on the three segments, we observe the following features:

- The smoothed expectile periodogram for the segment with a large earthquake exhibits spectral peaks at the low-frequency band at both higher and lower expectiles.
- The smoothed expectile periodogram for the segment with a small earthquake exhibits spectral peaks at both the low-frequency band (at lower and higher expectiles) and the high-frequency band (at middle expectiles).
- The smoothed expectile periodogram for the segment with no earthquake exhibits peaks only in the high-frequency bands.

4.2. Classification using Deep Learning Model

In this section, we use the three types of smoothed periodograms as features to classify the segments into those that contain earthquakes and those that do not. We randomly split the segments into training and testing sets, comprising 1,600 and 400 samples, respectively.

To classify the expectile periodograms and quantile periodograms, we employ a model with two convolutional layers for feature extraction, each followed by a max-pooling layer. The second pooling layer connects to a fully connected (FC) layer after flattening the output. A dropout layer with a rate of 50% is applied to the FC layer, leading to the output layer. The total number of trainable parameters is 2,817,682, and the learning rate is set to be 10^{-4} with a reduction rate of 0.5 every 20 epochs. Additional details about the model are presented at `https://github.com/tianbochen1`. The model structure is illustrated in Figure 8. For classifying the ordinary periodograms, which are one-dimensional with respect to ω, we adapt the model to use different input dimensions (1×500 instead of 46×500) and kernel size (1×5 instead of 5×5).

We conduct the training ten times, randomly constructing the training-testing split and initializing weights using random seeds. Over 80% of the training procedures converge within 30 epochs. The testing accuracies for the three types of periodograms are as follows: expectile periodogram: {0.9900, 1.0000, 0.9925, 1.0000, 0.9975, 0.9950, 0.9925, 0.9925, 0.9950,0.9925}, quantile

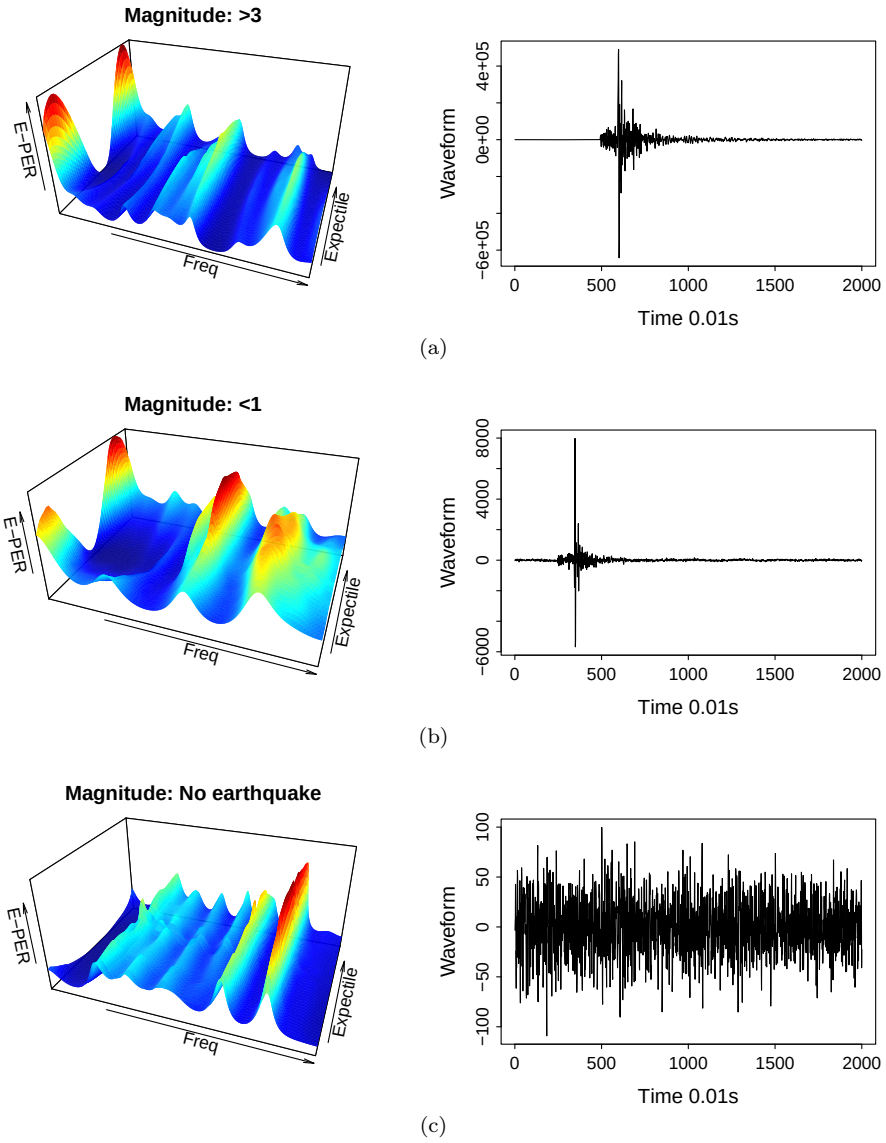


Figure 7: Three segments and the corresponding smoothed expectile periodograms. (a) The segment with an earthquake with a magnitude > 3, (b) the segment with an earthquake with a magnitude <1, and (c) the segment with no earthquake. $n = 2000$, $\omega \leq 0.25 \times 2\pi$, and $\tau = 0.05, 0.07, ..., 0.95$.

periodogram: {0.9825, 0.9900, 0.9925, 0.9925, 0.9825, 0.9950, 0.9925, 0.9900, 0.9875, 0.9825}, and ordinary periodogram: {0.9875, 0.9725, 0.9825, 0.9825,

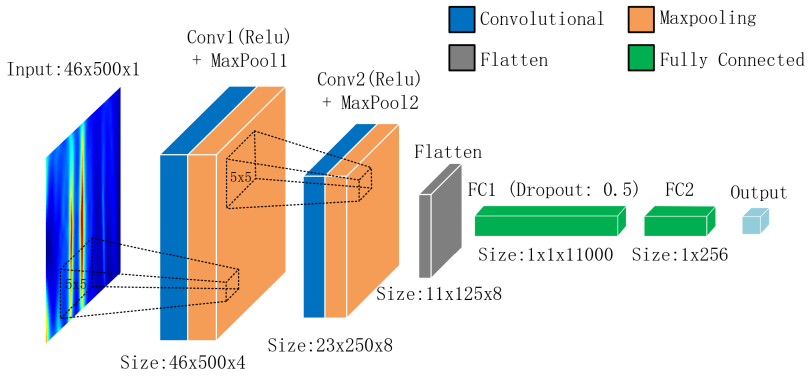


Figure 8: Structure of the deep learning model.

0.9975, 0.9850, 0.9850, 0.9825, 0.9775, 0.9900}.

The averaged confusion matrices are shown in Table 2, in which true positive (TP) indicates that a segment with an earthquake is correctly classified as containing an earthquake, true negative (TN) indicates that the segment without an earthquake is correctly classified as not containing an earthquake; false positive (FP) indicates that a segment without an earthquake is incorrectly classified as containing one; and false negative (FN) indicates that a segment with an earthquake is incorrectly classified as not containing one (where P denotes positive, N denotes negative, T denotes true, and F denotes false).

Table 3 shows three classification metrics: accuracy, precision ($\frac{TP}{TP+FP}$), and recall ($\frac{TP}{TP+FN}$).The optimal value in each row for the three types of periodograms is highlighted in bold. Additionally, we also present the time required for estimation and training (per epoch) to compare the computational complexity of the expectile and quantile periodograms. All computations were performed on a PC equipped with an Intel Core i9-13900KF processor @5.2GHz, 64 GB of memory, and an Nvidia RTX 4090 graphics card.

From the results, we can see that

- The classification based on expectile periodograms has higher testing accuracy, precision, and recall rate than the quantile periodograms and the

Table 2: The averaged confusion matrices of the classification.

(a) Expectile Periodogram

	P	N
T	199.3	198.6
F	0.7	1.4

(b) Quantile Periodogram

	P	N
T	196.2	199.3
F	3.8	0.7

(c) Ordinary Periodogram

	P	N
T	196.2	197.5
F	3.8	2.5

Table 3: The classification results.

Metrics		Expectile	Quantile	Ordinary
Accuracy	Averaged	**0.9948**	0.9880	0.9858
	[min, max]	**[0.9925, 1.0000]**	[0.9725, 0.9950]	[0.9750, 0.9925]
Precision	Averaged	**0.9965**	0.9840	0.9843
	[min, max]	**[0.9896, 1.0000]**	[0.9559, 0.9952]	[0.9609, 0.9902]
Recall	Averaged	**0.9931**	0.9921	0.9877
	[min, max]	**[0.9858, 1.0000]**	[0.9794, 1.0000]	[0.9653, 1.0000]
Time (s)	Estimation	12.6684	15.4528	–
	Training	0.3739	0.3731	–

ordinary periodograms. Specifically, all the testing accuracy in the ten experiments reach 0.99, and two of them being 1. This indicates that the expectile periodograms are suitable features for time series classification.

- One misclassification case in FN using the expectile periodograms is shown in Figure 9. Since the magnitude is too small (< 0.1), the power at low frequencies is not as large as the power at high frequencies, which causes the misclassification.

- The expectile periodogram incurs a lower estimation complexity compared to the quantile periodogram. However, the computational cost is higher than that of the ordinary periodogram, as the dimension is multiplied by

the number of expectile levels.

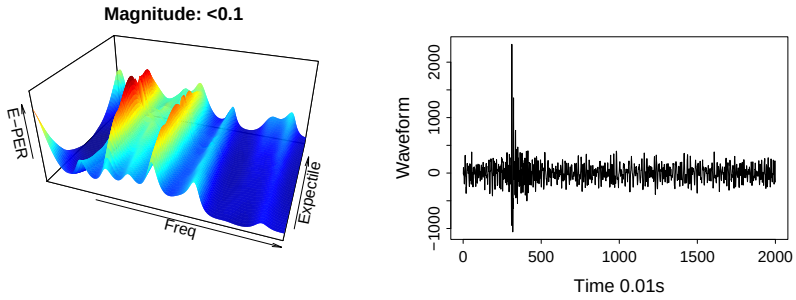


Figure 9: A misclassification case of FN.

5. Conclusion

we have proposed the expectile periodogram as a counterpart to the ordinary periodogram and investigated its potential as a non-parametric tool for time series analysis. The expectile periodogram offers more information than the ordinary periodogram by examining the serial dependence at different expectile levels. We conduct real-world examples and simulation studies to highlight their proficiency in detecting hidden periodicities within time series. In the earthquake data classification task, we leverage the inherent two-dimensional characteristics of the expectile periodogram using a deep learning model, which is a powerful technique in image classification.

Despite its computational efficiency relative to the quantile periodogram, the expectile periodogram demands higher computational resources than the ordinary periodogram. This computational burden arises from the fact that the dimensionality is multiplied by the number of expectiles used. One solution to reduce this computational cost is to use fewer expectiles. In this paper, we choose a large number of expectiles uniformly across $(0, 1)$. Researchers may choose to focus exclusively on a subset of expectiles with sufficient discriminative power (e.g., high or low expectiles). Furthermore, we utilize multi-thread

parallelization to speed up computation using the package foreach, doParallel in R.

6. Supplemental Material

The code for estimating the expectile periodogram and reproducing the results in Section 3 and Section 4 is accessible at: `https://github.com/tianbochen1/expectile_periodograms`.

Acknowledgment This work was supported by the National Natural Science Foundation of China under Grant No.12301326, Anhui Provincial Natural Science Foundation under Grant No.2308085QA05, and the University Natural Science Research Project of Anhui Province under Grant No.2023AH050099.

References

Aigner, D. J., Amemiya, T., and Poirier, D. J. (1976). On the estimation of production frontiers: maximum likelihood estimation of the parameters of a discontinuous density function. *International economic review*, pages 377–396.

Alvarado, R., Tillaguango, B., Dagar, V., Ahmad, M., Işık, C., Méndez, P., and Toledo, E. (2021). Ecological footprint, economic complexity and natural resources rents in latin america: empirical evidence using quantile regressions. *Journal of Cleaner Production*, 318:128585.

Baud, M. O., Kleen, J. K., Mirro, E. A., Andrechak, J. C., King-Stephens, D., Chang, E. F., and Rao, V. R. (2018). Multi-day rhythms modulate seizure risk in epilepsy. *Nature communications*, 9(1):88.

Benz, H. M., McMahon, N. D., Aster, R. C., McNamara, D. E., and Harris, D. B. (2015). Hundreds of earthquakes per day: The 2014 guthrie, oklahoma, earthquake sequence. *Seismological Research Letters*, 86(5):1318–1325.

Birr, S., Volgushev, S., Kley, T., Dette, H., and Hallin, M. (2017). Quantile spectral analysis for locally stationary time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(5):1619–1643.

Bloomfield, P. (2004). *Fourier analysis of time series: an introduction*. John Wiley & Sons.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer.

Cai, Z. and Xiao, Z. (2012). Semiparametric quantile regression estimation in dynamic models with partially varying coefficients. *Journal of Econometrics*, 167(2):413–425.

Cai, Z. and Xu, X. (2008). Nonparametric quantile estimations for dynamic smooth coefficient models. *Journal of the American Statistical Association*, 103(484):1595–1608.

Caiado, J., Crato, N., and Peña, D. (2006). A periodogram-based metric for time series classification. *Computational Statistics & Data Analysis*, 50(10):2668–2684.

Chen, T., Sun, Y., Euan, C., and Ombao, H. (2021a). Clustering brain signals: A robust approach using functional data ranking. *Journal of Classification*, 38:425–442.

Chen, T., Sun, Y., and Li, T.-H. (2021b). A semi-parametric estimation method for the quantile spectrum with an application to earthquake classification using convolutional neural network. *Computational Statistics & Data Analysis*, 154:107069.

Dette, H., Hallin, M., Kley, T., Volgushev, S., et al. (2015). Of copulas, quantiles, ranks and spectra: An l_1 -approach to spectral analysis. *Bernoulli*, 21(2):781–831.

Efron, B. (1991). Regression percentiles using asymmetric squared error loss. *Statistica Sinica*, pages 93–125.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 50(4):987–1007.

Euán, C., Ombao, H., and Ortega, J. (2018). The hierarchical spectral merger algorithm: a new time series clustering procedure. *Journal of Classification*, 35:71–99.

Garcia, J., Hernández, P. J., and Lopez-Nicolas, A. (2001). How wide is the gap? an investigation of gender wage differences using quantile regression. *Empirical economics*, 26:149–167.

Granger, C. and Sin, C. (1997). Estimating and forecasting quantiles with asymmetric least squares. Technical report, Working Paper, University of California, San Diego.

Gu, Y. and Zou, H. (2016). High-dimensional generalizations of asymmetric least squares regression and their applications.

Hagemann, A. (2013). Robust spectral analysis.

Jiang, C., Jiang, M., Xu, Q., and Huang, X. (2017). Expectile regression neural network model with applications. *Neurocomputing*, 247:73–86.

Jiang, Y., Lin, F., and Zhou, Y. (2021). The kth power expectile regression. *Annals of the Institute of Statistical Mathematics*, 73:83–113.

Jones, M. C. (1994). Expectiles and m-quantiles are quantiles. *Statistics & Probability Letters*, 20(2):149–153.

Kley, T. (2016). Quantile-based spectral analysis in an object-oriented framework and a reference implementation in r: The quantspec package. *Journal of Statistical Software, Articles*, 70(3):1–27.

Koenker, R. (2005). Quantile regression. *Cambridge University Press*.

Koenker, R. (2017). Quantile regression: 40 years on. *Annual review of economics*, 9:155–176.

Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.

Kouretas, G. P., Zarangas, L., et al. (2005). Conditional autoregressive value at risk by regression quantiles: Estimating market risk for major stock markets. Technical report.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.

Li, T.-H. (2008). Laplace periodogram for time series analysis. *Journal of the American Statistical Association*, 103(482):757–768.

Li, T.-H. (2012a). Detection and estimation of hidden periodicity in asymmetric noise by using quantile periodogram. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),*vol 2, pp. 3969-3972.

Li, T.-H. (2012b). Quantile periodograms. *Journal of the American Statistical Association*, 107(498):765–776.

Li, T.-H. (2014). Quantile periodogram and time-dependent variance. *Journal of Time Series Analysis*, 35(4):322–340.

Li, T.-H. (2023). Quantile-frequency analysis and deep learning for signal classification. *Journal of Nondestructive Evaluation*, 42(2):40.

Maadooliat, M., Sun, Y., and Chen, T. (2018). Nonparametric collective spectral density estimation with an application to clustering the brain signals. *To appear in Statistics in Medicine*.

Machado, J. A. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics*, 20(4):445–465.

Martínez-Murcia, F. J., Ortiz, A., Morales-Ortega, R., López, P., Luque, J. L., Castillo-Barnes, D., Segovia, F., Illan, I. A., Ortega, J., Ramirez, J., et al. (2019). Periodogram connectivity of eeg signals for the detection of dyslexia. In *Understanding the Brain Function and Emotions: 8th International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2019, Almería, Spain, June 3–7, 2019, Proceedings, Part I 8*, pages 350–359. Springer.

Meziani, A., Medkour, T., and Djouani, K. (2020). Penalised quantile periodogram for spectral estimation. *Journal of Statistical Planning and Inference*, 207:86–98.

Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica: Journal of the Econometric Society*, pages 819–847.

Polat, K. and Güneş, S. (2007). Classification of epileptiform eeg using a hybrid system based on decision tree classifier and fast fourier transform. *Applied Mathematics and Computation*, 187(2):1017–1026.

Schnabel, S. K. and PHC, E. (2009). Non-crossing smooth expectile curves. In *Proceedings of the 24th International Workshop on Statistical Modelling. Ithaca, NY, USA*, pages 330–36.

Sharif, A., Bhattacharya, M., Afshan, S., and Shahbaz, M. (2021). Disaggregated renewable energy sources in mitigating co2 emissions: new evidence from the usa using quantile regressions. *Environmental Science and Pollution Research*, 28(41):57582–57601.

Shumway, R. H. and Stoffer, D. S. (2016). *Time series analysis and its applications: with R examples*. Springer Science & Business Media.

Waltrup, L. S., Sobotka, F., Kneib, T., and Kauermann, G. (2015). Expectile and quantile regression—david and goliath? *Statistical Modelling*, 15(5):433–456.

Xu, Q., Chen, L., Jiang, C., and Yu, K. (2020). Mixed data sampling expectile regression with applications to measuring financial risk. *Economic Modelling*, 91:469–486.

Xu, Q., Ding, X., Jiang, C., Yu, K., and Shi, L. (2021). An elastic-net penalized expectile regression with applications. *Journal of Applied Statistics*, 48(12):2205–2230.

Yao, Q. and Tong, H. (1996). Asymmetric least squares regression estimation: a nonparametric approach. *Journal of nonparametric statistics*, 6(2-3):273–292.