Tempered stable autoregressive models

Niharika Bhootna and Arun Kumar

Department of Mathematics, Indian Institute of Technology Ropar, Rupnagar, Punjab, India

ABSTRACT
In this article, we introduce and study a one sided tempered stable first order autoregressive model called TAR(1). Under the assumption of stationarity of the model, the marginal probability density function of the error term is found. It is shown that the distribution of the error term is infinitely divisible. Parameter estimation of the introduced TAR(1) process is done by adopting the conditional least square and method of moments based approach and the performance of the proposed methods are evaluated on simulated data. Also, we study an autoregressive model of order one with tempered stable innovations. Using appropriate test statistic it is shown that the model fits very well on real and simulated data. Our models generalize the inverse Gaussian and one-sided stable autoregressive models existing in the literature.

ARTICLE HISTORY
Received 17 August 2021
Accepted 15 June 2022

KEYWORDS
Positive tempered stable distribution; autoregressive time-series model; complex inversion of Laplace transforms; estimation; infinite divisibility

MATHEMATICS SUBJECT CLASSIFICATION
62F10; 60G10; 60G12; 44A10

1. Introduction

Autoregressive models serve as eminent approach of modeling among different time-series methods. Many real life time-series data exhibit autoregressive behavior. Classical autoregressive time series models are based on the assumption of normality of the innovation term. However, this assumption may not be true for all the real life scenarios. Many real life time series probability distributions display heavy-tailed behavior or semi-heavy tailed behavior. Semi-heavy tailed pdf are those where tails are heavier than the Gaussian and lighter than the power law, see for example, Omey, Van Gulck, and Vesilo (2018). It is well known that series of counts, proportions, binary outcomes or non-negative observations are some examples of non-normal real life time-series data see for example, Grunwald, Hyndman, and Tedesco (1996). The autoregressive model of order 1 denoted by AR(1) is a simple, useful and interpretable model in a wide range of real life applications. Several AR(1) models with different marginal distributions are considered in literature. For example, Gaver and Lewis (1980) considered linear additive first-order autoregressive scheme with gamma distributed marginals. The authors show that error distribution is same as the distribution of a non-negative random variable which has a point mass at 0 and which is exponential if positive (see Gaver and Lewis (1980)). A more general exponential ARMA(p, q) model called EARMA(p, q) is considered in Lawrance and Lewis (1980). A second order autoregressive model is considered...
in Dewald and Lewis (1985), where Laplace distributed marginals are considered. Abraham and Balakrishna (1999) considered inverse Gaussian autoregressive model where the marginals are inverse Gaussian distributed. The authors show that after fixing a parameter of the inverse Gaussian to 0 the error term is also inverse Gaussian distributed. The transition laws of tempered stable Ornstein–Uhlenbeck processes are discussed in Zhang and Zhang (2009).

In this paper, we consider tempered stable autoregressive model of order 1 and denote it by TAR(1). This model generalizes the work discussed in Abraham and Balakrishna (1999). Since one parameter inverse Gaussian (also called Lévy distribution) is a particular case of the one sided tempered stable distribution. We derive the explicit form of the error density. Further, we show that if the AR(1) series marginals are stable distributed the error is also stable distributed with some scaling. A step-by-step procedure of the estimation of the parameters of the proposed model is given. Moreover, we consider an AR(1) model with one sided tempered stable innovations. This model is different from the TAR(1) model where marginals are one sided tempered stable and here the innovations are positive tempered stable distributed. The application of this model is shown on real life power consumption data and drinking water treatment data. The rest of the paper is organized as follows: Section 2 defines the TAR(1) model where the explicit form of the density of the error term, infinite divisibility, moments and parameter estimation methods of TAR(1) model are discussed. In Section 3, an AR(1) model with positive tempered stable innovations is introduced. The estimation procedure of the parameters of the introduced model is given based on the method of moments and conditional least square. The simulations study, where efficacy of the estimation procedure is discussed based on simulated data and real life application are discussed in this section. The last section concludes.

2. Tempered stable autoregressive model

In this section, we introduce the tempered stable autoregressive model and discuss the main properties. Consider an autoregressive process of order 1, defined as:

$$X_n = \rho X_{n-1} + \epsilon_n,$$

where $|\rho| < 1$ and $\{\epsilon_n, n \geq 1\}$ is sequence of i.i.d. random variables. Assuming marginals $X_n$'s to be stationary positive tempered stable distributed, then there exist a distribution of $\epsilon_n$'s. The class of stable distributions is denoted by $\text{Stable}(\beta, \nu, \mu, \sigma)$, with parameter $\beta \in (0, 2]$ is the stability index, $\nu \in [-1, 1]$ is the skewness parameter, $\mu \in \mathbb{R}$ is the location parameter and $\sigma > 0$ is the shape parameter. The stable class probability density functions do not possess closed form except for three cases (Gaussian, Cauchy, and Lévy). Generally stable distributions are represented in terms of their characteristic functions or Laplace transforms. Stable distributions are infinitely divisible. The one-sided stable random variable $S$ has following Laplace transform (see Samorodnitsky and Taqqu (1994))

$$\mathbb{E}(e^{-sS}) = e^{-\beta s}, \ s > 0, \ \beta \in (0, 1).$$
The right tail of the $S$ behaves (see e.g. Samorodnitsky and Taqqu (1994))
\[ P(S > x) \sim \frac{x^{-\beta}}{\Gamma(1 - \beta)}, \text{ as } x \to \infty. \] (2.2)

Next, we introduce the positive tempered stable distribution. The positive tempered stable random variable $T$ with tempering parameter $\lambda > 0$ and stability index $\beta \in (0, 1)$, has the Laplace transform (LT) (see e.g. Meerschaert, Nane, and Vellaisamy (2012), Kumar and Vellaisamy (2015))
\[ \mathbb{E}(e^{-sT}) = e^{-((s + \lambda)\beta - \lambda \beta)}. \] (2.3)

Note that tempered stable distributions are obtained by exponential tempering in the distributions of $\alpha$-stable distributions (see Rosiński (2007)). The advantage of tempered stable distribution over an $\alpha$-stable distribution is that it has moments of all order and its density is also infinitely divisible. The probability density function for $T$ is given by
\[ f_{\beta, \lambda}(x) = e^{-\lambda x + \beta f_{\beta}(x)}, \quad \lambda > 0, \beta \in (0, 1), \]
where $f_{\beta}(x)$ is the PDF of an $\alpha$-stable distribution (Uchaikin and Zolotarev (1999). The tail probability of tempered stable distribution has the following asymptotic behavior
\[ P(T > x) \sim c_{\beta, \lambda} e^{-\lambda x} x^{-\beta}, \text{ as } x \to \infty, \] (2.4)
where $c_{\beta, \lambda} = \frac{1}{\beta \pi} \Gamma(1 + \beta) \sin (\pi \beta) e^{\beta}$. The first two moments of tempered stable distribution are given by
\[ \mathbb{E}(T) = \beta \lambda^{\beta - 1}, \mathbb{E}(T)^2 = \beta(1 - \beta) \lambda^{2\beta - 2} + (\beta \lambda^{\beta - 1})^2. \]

A pdf $u(x)$ is called a semi-heavy tailed pdf if
\[ u(x) \sim e^{-cx} v(x), \quad c > 0, \]
where $v$ is a regularly varying function (see e.g. Omeya, Van Gulck, and Vesilo (2018)). Recall that a positive function $v$ is regularly varying with index $\nu$ if
\[ \lim_{x \to \infty} \frac{v(dx)}{v(x)} = d^\nu, \quad d > 0. \]

Using (2.4), it is straightforward to show that
\[ f_{\beta, \lambda}(x) \sim \lambda c_{\beta, \lambda} x^{-\beta} e^{-\lambda x}, \text{ as } x \to \infty, \]
and hence the tempered stable density function is semi-heavy tailed.

2.1. Distributional properties

In this subsection, we discuss about the distributional properties related to the introduced TAR(1) process. If each $X_n$ in (2.1) is positive tempered stable then the Laplace transform of $X_n$ is given by
\[ \Phi_{X_n}(s) = \mathbb{E}(\exp(-sX_n)) = \mathbb{E}((\exp(-s(\rho X_{n-1} + \epsilon_n)))) = \Phi_{X_{n-1}}(\rho s)\Phi_{\epsilon_n}(s). \] (2.5)
Assuming $X_n$'s to be stationary distributed, it follows
\[ \Phi_X(s) = \exp \{-((\lambda + s)^\beta - \lambda^\beta)\}, \quad \lambda > 0, \quad \beta \in (0,1), \] (2.6)
and
\[ \Phi_X(\rho s) = \exp \{-((\lambda + \rho s)^\beta - \lambda^\beta)\}, \quad \lambda > 0, \quad \beta \in (0,1). \] (2.7)
Using (2.5)
\[ \Phi_X(s) = \Phi_x(s)\Phi_X(\rho s), \]
which implies
\[ \Phi_x(s) = \frac{\Phi_X(s)}{\Phi_X(\rho s)}. \]
Putting values from (2.6) and (2.7), yields
\[ \Phi_x(s) = \exp \{(\rho s + \lambda)^\beta - (s + \lambda)^\beta\}. \] (2.8)
The $k$th order moment can be extracted by taking the $k$th order derivative of Laplace transform as follows for the innovation term
\[ \mathbb{E}[\epsilon^k] = (-1)^k \frac{d^k \Phi_x(s)}{ds^k} \bigg|_{s=0}, \quad k \in \mathbb{N}, \]
which yields
\[ \mathbb{E}(\epsilon) = \beta \lambda^{\beta-1}(1 - \rho), \] (2.9)
and
\[ \mathbb{E}(\epsilon^2) = (\beta \lambda^{\beta-1}(\rho - 1))^2 + \beta(\beta - 1)\lambda^{\beta-2}(\rho^2 - 1). \] (2.10)

**Remark 2.1.** For $\lambda = 0$ in (2.8), we have $\Phi_x(s) = e^{-(1-\rho^\beta)s}$, which is the Laplace transform of a positive stable random variable $(1 - \rho^\beta)^{1/\beta}$. Hence, it shows that if the AR(1) series is stable the error term is also stable.

**Proposition 2.1.** The error distribution with Laplace transform (2.8) is infinitely divisible.

**Proof.** By Feller (1971), a pdf $f(x)$, $x \geq 0$ is infinitely divisible iff its Laplace transform is of the form $F(s) = e^{-\phi(s)}$, $s > 0$ where $\phi(s)$ has a completely monotone derivative. Further, a function $\psi(\cdot)$ is completely monotone if it possesses derivatives $\psi^{(n)}(\cdot)$ of all orders and $(-1)^n\psi^{(n)}(s) \geq 0$, $s > 0$, $n = 0, 1, \ldots$ (see e.g. Feller (1971), p. 439). For $\Phi_x(s)$, we have $\phi(s) = (s + \lambda)^\beta - (\rho s + \lambda)^\beta$ which gives $\phi'(s) = \beta(s + \lambda)^{\beta-1} - \beta \rho (\rho s + \lambda)^{\beta-1}$. Consider $\psi(s) = (s + \lambda)^{\beta-1}$, which yields to $\psi^{(n)}(s) = (\beta - 1)(\beta - 2) \cdots (\beta - n)(s + \lambda)^{\beta-n-1}$. Thus $(-1)^n\psi^{(n)}(s) \geq 0$, $n = 0, 1, 2, \ldots$. Similarly, considering $\zeta(s) = (\rho s + \lambda)^{\beta-1}$ we can show that $\zeta(s)$ is completely monotone. Moreover, $(-1)^n\phi^n(s) = (-1)^n\psi^n(s) - (-1)^n\zeta^n(s)$ and $\psi^n(s)$ is always greater than $\zeta^n(s)$ as $0 < \rho < 1$, this implies $(-1)^n\phi^n(s) \geq 0$. So $\phi(s)$ has completely monotone derivative. \(\square\)
Remark 2.2. Using (2.8), it follows that the random variable \( e \) can be written as the difference of two independent but not identically distributed tempered stable random variables. Hence the error distributions are infinitely divisible since the difference of two infinitely divisible random variables is infinitely divisible (see e.g. Steutel and Van Harn (2004), Prop. 2.1, p. 94).

Next, we provide the explicit form of the pdf of the error term using complex inversion formula, which generalizes the corresponding results in Abraham and Balakrishna (1999).

Theorem 2.1. For the stationary tempered stable autoregressive model defined in (2.1) with Laplace transform defined as

\[
\mathbb{E}(e^{-sX_0}) = e^{-(s+\lambda)^\beta - \lambda\beta},
\]

the pdf of the innovation term has following integral form

\[
g(x) = -\frac{1}{2\pi i} \int_{C} (e^{-\lambda x/\rho} e^{-wx} \rho^\beta w^\beta e^{\lambda x/\rho} - (we^{\alpha \lambda} + \lambda - 1/\rho)^\beta - e^{-\lambda x/\rho} e^{-wx} \rho^\beta w^\beta e^{-\lambda x/\rho} - (we^{\alpha \lambda} + \lambda - 1/\rho)^\beta) dw
\]

\[
-\frac{1}{2\pi i} \int_{C} (e^{-\lambda x} e^{-w} e^{-\lambda x/\rho} e^{-wx} + \lambda)^\beta - (we^{\alpha \lambda} + \lambda - 1/\rho)^\beta - e^{-\lambda x} e^{-w} e^{-\lambda x/\rho} e^{-wx} + \lambda)^\beta - (we^{\alpha \lambda} + \lambda - 1/\rho)^\beta) dw,
\]

where \( \lambda > 0 \) and \( \beta \in (0, 1) \).

Proof. Inverse Laplace transform corresponding to \( \Phi_e(s) \) provides the density of innovation term. The density for innovation can be computed using complex inversion formula for Laplace transform given by

Figure 1. The double key hole contour.
\[ g(x) = \frac{1}{2\pi i} \int_{x_0-i\infty}^{x_0+i\infty} e^{ix\Phi_x(s)} ds, \]  

(2.13)

where point \( x_0 > a \) for some \( a \) is taken in such a way that the integrand is analytic for \( \Re(s) > a \). Note that the integrand \( e^{ix\Phi_x(s)} \) is an exponential function which is analytic in whole complex plane. However, due to fractional power in the exponent the integrand \( e^{ix\Phi_x(s)} \) has branch points at \( s = -\lambda \) and \( s = -\lambda/\rho \). Thus we take a branch cut along the non-positive real axis and consider a (single-valued) analytic branch of the integrand. To calculate (2.13), consider a closed double-key-hole contour \( C: ABCDEFGHIJA \) (Figure 1) with branch points at \( O = (-\lambda, 0) \) and \( O' = (-\lambda/\rho, 0) \). The contour consists of following segments: AB and IJ are arcs of radius \( R \), BC, DE, FG and HI are line segments parallel to \( x \)-axis, CD, EF and GH are arcs of circles with radius \( r \) and JA is the line segment from \( x_0 - iy \) to \( x_0 + iy \) (see Figure 1). The integrand is analytic within and on the contour \( C \) so that by Cauchy’s residue theorem (see e.g. Schiff (1999))

\[ \frac{1}{2\pi i} \int_C e^{ix\Phi_x(s)} ds = 0. \]

Using (Lemma 4.1 Schiff (1999), p. 154), the integral on circular arcs AB and IJ tend to 0 as \( R \to \infty \). The integral over CD, EF and GH are also zero as \( r \) goes to infinity. Thus, as \( r \to 0 \) and \( R \to \infty \), we have

\[ \frac{1}{2\pi i} \int_{x_0-i\infty}^{x_0+i\infty} e^{ix\Phi_x(s)} ds = -\frac{1}{2\pi i} \int_{BC} e^{ix\Phi_x(s)} ds - \frac{1}{2\pi i} \int_{DE} e^{ix\Phi_x(s)} ds - \frac{1}{2\pi i} \int_{FG} e^{ix\Phi_x(s)} ds - \frac{1}{2\pi i} \int_{HI} e^{ix\Phi_x(s)} ds. \]  

(2.14)

Along \( BC \), put \( s = -\frac{s}{\rho} + we^{i\pi} \), which implies \( ds = -dw \), and hence

\[ \int_{BC} e^{ix} \exp ((\rho s + \lambda)^\beta - (s + \lambda)^\beta) ds = \int_r^{R-\lambda/\rho} e^{-\lambda s/\rho} e^{-wx} e^{\rho \theta} e^{w e^{\rho \theta} - (we^{\rho \theta} + \lambda - \lambda/\rho)^\beta} dw. \]  

(2.15)

Along \( HI \), \( s = -\frac{s}{\rho} + we^{i\pi} \) and \( ds = -dw \)

\[ \int_{HI} e^{ix} \exp ((\rho s + \lambda)^\beta - (s + \lambda)^\beta) ds = -\int_r^{R-\lambda/\rho} e^{-\lambda s/\rho} e^{-wx} e^{\rho \theta} e^{w e^{\rho \theta} - (we^{\rho \theta} + \lambda - \lambda/\rho)^\beta} dw. \]  

(2.16)

Along \( DE \), \( s = -\lambda + we^{i\pi} \) and \( ds = -dw \)

\[ \int_{DE} e^{ix} \exp ((\rho s + \lambda)^\beta - (s + \lambda)^\beta) ds = \int_r^{-r+(\lambda/\rho) - \lambda} e^{-\lambda x} e^{-wx} e^{-(\lambda x + \rho w e^{\rho \theta} + \lambda)^\beta} (we^{\rho \theta})^\beta dw. \]  

(2.17)
Along $FG$, $s = -\lambda + \alpha e^{-t\pi}$ and $ds = -dw$

\[
\int_{FG} e^{is} \exp \left( (\rho s + \lambda)^{\beta} - (s + \lambda)^{\beta} \right) ds = - \int_r^{r+(\lambda/\rho)-\lambda} e^{-is} e^{-w e^{-(\lambda s + \rho \alpha e^{-t\pi} + \lambda)^{\beta}}} \exp \left( (\lambda s + \rho \alpha e^{-t\pi})^{\beta} \right) ds
\]

(2.18)

Substituting (2.15), (2.16), (2.17) and (2.18) in (2.14), the result follows as $r \to 0$ and $R \to \infty$.

\[\text{Remark 2.3.}\]

For the special case $\lambda = 0$ and $\beta = \frac{1}{2}$, (2.12) reduces to

\[
g(x) = \left( \frac{1 - \sqrt{\rho}}{\sqrt{2}} \right) \left( \frac{1}{2\pi x^3} \right)^{1/2} \exp \left( \frac{(1 - \sqrt{\rho})^2}{4x} \right), \quad |\rho| < 1,
\]

and is called Lévy probability density function see Applebaum (2004), which is a special case of the inverse Gaussian density. Abraham and Balakrishna (1999) show that if the inverse Gaussian autoregressive model is stationary the error term is also inverse Gaussian distributed. Our results complement their findings.

\[\text{Remark 2.4.}\]

For $\lambda = 0$, the error density function in (2.12) reduces to

\[
g(x) = \frac{1}{\pi} \int_0^\infty e^{-w x} \exp \left( (1 - \rho^{\beta}) \cos (\pi \beta) \sin (w (1 - \rho^{\beta}) \sin (\pi \beta)) \right), \quad x > 0.
\]

(2.19)

\[\text{Proposition 2.2.}\]

For $\lambda = 0$ the innovation term has the following fractional $q$th order moment

\[
\mathbb{E}(\epsilon^q) = \frac{\Gamma(1 - q/\beta)}{\Gamma(1 - q)} (1 - \rho^{\beta})^{q/\beta}, \quad 0 < q < \beta < 1.
\]

\[\text{Proof.}\]

For a positive random variable $X$ with a Laplace transform $\Phi(s)$, and for $p > 0$, it follows

\[
\int_0^\infty \frac{d^n}{ds^n} [\Phi(s)]^{s-1} ds = \int_0^\infty \frac{d^n}{ds^n} [\mathbb{E} e^{-sX}]^{s-1} ds = \mathbb{E} \int_0^\infty \frac{d^n}{ds^n} [e^{-sx}]^{s-1} ds = (-1)^n \mathbb{E} \int_0^\infty X^n e^{-sx} ds^{s-1} ds = (-1)^n \Gamma(p) \mathbb{E}(X^{n-p}).
\]

Thus for $q \in (n-1, n)$, for an integer $n$, we have

\[
\mathbb{E}(X^q) = \frac{(-1)^n}{\Gamma(n-q)} \int_0^\infty \frac{d^n}{ds^n} [\Phi(s)]^{s-n-q-1} ds.
\]

Using the relationship for $q \in (0, 1)$, it follows

\[
\mathbb{E}(\epsilon^q) = \frac{-1}{\Gamma(1 - q)} \int_0^\infty \frac{d}{ds} [e^{(\rho s + \lambda)^{\beta} - (s + \lambda)^{\beta}}]^{s-q} ds
\]

\[
= \frac{1}{\Gamma(1 - q)} \int_0^\infty e^{(\rho s + \lambda)^{\beta} - (s + \lambda)^{\beta}} \left( (s + \lambda)^{\beta-1} - \rho \beta (\rho s + \lambda)^{\beta-1} \right) ds,
\]

which gives the $q$th order fractional moment of the innovation term. For $\lambda = 0$, it can
be written in the explicit form as follows
\[
\mathbb{E}(\epsilon^2) = \frac{1}{\Gamma(1-q)} \int_0^\infty e^{-(1-\rho^q)s} s^{\beta-1} (1 - \rho^s) ds
\]
\[
= \frac{1}{\Gamma(1-q)} \int_0^\infty e^{-(1-\rho^q)u} u^{q/\beta} (1 - \rho^u) du \quad (\text{put } s^\beta = u)
\]
\[
= \frac{(1 - \rho^\beta)}{\Gamma(1-q)} \frac{\Gamma(1 - q/\beta)}{(1 - \rho^\beta)^{(1-q)/\beta)} = \frac{\Gamma(1 - q/\beta)}{\Gamma(1-q)} (1 - \rho^\beta)^{q/\beta}, 0 < q < \beta.
\]

\[\square\]

**Proposition 2.3.** For the stationary TAR(1) model defined in (2.1) the autocorrelation for \(r\)th lag \(\rho\), has the following recursive form
\[\rho_r = \rho^r.\]

**Proof.** For the given TAR(1) process \(X_n = \rho X_{n-1} + \epsilon_n\), multiplying both the side by \(X_{n-r}\), we get
\[X_nX_{n-r} = \rho X_{n-1}X_{n-r} + \epsilon_nX_{n-r}.\]

Here \(X_n\) is independent of \(\epsilon_j\) for \(j > n\). Taking expectation on both the sides, yields
\[\mathbb{E}(X_nX_{n-r}) = \rho \mathbb{E}(X_{n-1}X_{n-r}) + \mathbb{E}(\epsilon_nX_{n-r}),\]
which gives
\[\mathbb{E}(X_nX_{n-r}) - \mathbb{E}(X_n)\mathbb{E}(X_{n-r}) = \rho \mathbb{E}(X_{n-1}X_{n-r}) - \rho \mathbb{E}(X_{n-1})\mathbb{E}(X_{n-r}) + \rho \mathbb{E}(X_{n-1})\mathbb{E}(X_{n-r})
\]
\[= \mathbb{E}(X_n)\mathbb{E}(X_{n-r}) + \mathbb{E}(\epsilon_n)\mathbb{E}(X_{n-r}).\]

Thus,
\[\text{Cov}(X_n, X_{n-r}) = \rho \text{Cov}(X_{n-1}, X_{n-r}) + \mathbb{E}(X_{n-1})\mathbb{E}(X_{n-r})(\rho - 1) + \mathbb{E}(\epsilon_n)\mathbb{E}(X_{n-r})
\]
\[= \rho \text{Cov}(X_{n-1}, X_{n-r}) + \mathbb{E}(X_{n-r})(\mathbb{E}(X_n)(\rho - 1) + \mathbb{E}(\epsilon_n)).\]

Since \(\mathbb{E}(\epsilon_n) = (1 - \rho)\beta \hat{\lambda}^{\beta-1}\) and \(\mathbb{E}(X_n) = \beta \hat{\lambda}^{\beta-1}\), it follows
\[\text{Cov}(X_n, X_{n-r}) = \rho \text{Cov}(X_{n-1}, X_{n-r}) + \mathbb{E}(X_{n-r})(\beta \hat{\lambda}^{\beta-1}(\rho - 1) + (1 - \rho)\beta \hat{\lambda}^{\beta-1})
\]
\[= \rho \text{Cov}(X_{n-1}, X_{n-r}).\]

Dividing both side by \(\sqrt{\text{Var}(X_n)}\sqrt{\text{Var}(X_{n-r})}\), yields
\[\frac{\text{Cov}(X_n, X_{n-r})}{\sqrt{\text{Var}(X_n)}\sqrt{\text{Var}(X_{n-r})}} = \frac{\text{Cov}(X_{n-1}, X_{n-r})}{\sqrt{\text{Var}(X_{n-1})}\sqrt{\text{Var}(X_{n-r})}},\]

Using, \(\text{Var}(X_n) = \text{Var}(X_{n-1}) = \text{Var}(X_{n-r})\), leads to
\[\rho_r = \rho \rho_{r-1} \Rightarrow \rho_r = \rho^r,\]
since \(\rho_0 = 1\) and \(\rho_1 = \rho\).
2.2. Parameter estimation

Here, we discuss the estimation procedure of the parameters of the tempered stable autoregressive model. We consider conditional least square and moments based estimation methods.

2.2.1. Parameter estimation by conditional least square

The conditional least square method provides a straightforward procedure to estimate the parameters for dependent observations by minimizing the sum of square of deviations about conditional expectation (see e.g. Klimkov and Nelson (1978); Balakrishna and Shiji (2014)). To estimate the parameter $\rho$ of process (2.1) the conditional least square (CLS) method is used. The CLS estimator of parameter vector $\theta = (\rho, \lambda, \beta)$ is obtained by minimizing

$$Q_n(\theta) = \sum_{i=1}^{n} [X_{i+1} - \mathbb{E}(X_{i+1}|X_i, X_{i-1}, \ldots, X_1)]^2,$$  \hspace{1cm} (2.20)

with respect to $\theta$. For an AR(1) sequence due to Markovian property

$$\mathbb{E}(X_{i+1}|X_i, X_{i-1}, \ldots, X_1) = \mathbb{E}(X_{i+1}|X_i),$$  \hspace{1cm} (2.21)

which yields

$$\mathbb{E}(X_{i+1}|X_i) = \rho X_i + \mathbb{E}(\epsilon_i),$$  \hspace{1cm} (2.22)

where $\mathbb{E}(\epsilon_i) = \beta \lambda^{\beta-1} (1 - \rho)$. Using the Equations (2.20) and (2.21), we can write

$$Q_n(\theta) = \sum_{i=1}^{n} [X_{i+1} - \rho X_i - (1 - \rho)\beta \lambda^{\beta-1}]^2.$$  \hspace{1cm} (2.23)

Minimizing (2.23) w.r.t $\rho, \beta$, and $\lambda$ gives the following relation for $\beta$ and $\lambda$

$$\beta \lambda^{\beta-1} = \frac{\sum_{i=1}^{n} (X_{i+1} - \hat{\rho} X_i)}{n(1 - \hat{\rho})}. $$  \hspace{1cm} (2.24)

and estimate for $\rho$ using above Equation (2.24) is

$$\hat{\rho} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(X_{i+1} - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}. $$  \hspace{1cm} (2.25)

The estimate $\hat{\rho}$ can be easily calculated from the observed sample using (2.25). However, some numerical method which minimizes the squared difference between left and right hand sides is required to estimate $\beta$ and $\lambda$ from (2.24). Next we prove the asymptotic normality of the estimator $\hat{\rho}_n$, which is otherwise denoted by $\hat{\rho}$ for ease of notation.

**Proposition 2.4.** The CLS estimator $\hat{\rho}_n$ for $\rho$ is asymptotically normal that is

$$\sqrt{n}(\hat{\rho}_n - \rho) \xrightarrow{d} \mathcal{N}(0, 1 - \rho^2), \text{ as } n \to \infty.$$

**Proof.** TAR(1) process satisfies the regularity conditions defined by Klimkov and Nelson (1978), so one can say that the CLS estimator of $\rho$ is consistent and normal that is,
\[ \sqrt{n}(\hat{\rho}_n - \rho) \xrightarrow{\mathcal{L}} N(0, D), \]

where \( D = \frac{W}{V} \) and the value of \( W \) and \( V \) can be calculated using the method defined by Klimkov and Nelson (1978). Define

\[
g(\theta, X_{n-1}) = \rho X_{n-1} + (1 - \rho)\beta \lambda^{(\beta-1)}; \\
V = \mathbb{E} \left[ \left( \frac{\partial g(\theta, X_{n-1})}{\partial \rho} \right)^2 \right] = \mathbb{E} \left[ X_{n-1}^2 \right] - \beta \lambda^{(\beta-1)}; \\
W = \mathbb{E} \left[ U_n \left( \frac{\partial g(\theta, X_{n-1})}{\partial \rho} \right)^2 \right],
\]

where, \( U_n = X_n - \rho X_{n-1} - (1 - \rho)\beta \lambda^{(\beta-1)} = \varepsilon_n - (1 - \rho)\beta \lambda^{(\beta-1)}; \)

\[
W = \mathbb{E} \left[ (\varepsilon_n - (1 - \rho)\beta \lambda^{(\beta-1)})^2 (X_{n-1} - \beta \lambda^{(\beta-1)})^2 \right] \\
= \mathbb{E} \left[ (\varepsilon_n - (1 - \rho)\beta \lambda^{(\beta-1)})^2 \right] \mathbb{E} \left[ (X_{n-1} - \beta \lambda^{(\beta-1)})^2 \right].
\]

Thus

\[
\frac{W}{V^2} = \frac{\mathbb{E}[(\varepsilon_n - (1 - \rho)\beta \lambda^{(\beta-1)})^2]}{\mathbb{E}[X_{n-1}^2 - \beta \lambda^{(\beta-1)}]^2]} \\
= \frac{\mathbb{E}[(\varepsilon_n)^2 - ((1 - \rho)\beta \lambda^{(\beta-1)})^2 - 2\varepsilon_n(1 - \rho)(\beta \lambda^{(\beta-1)})]}{\mathbb{E}[X_{n-1}^2] - (\beta \lambda^{(\beta-1)})^2]} \\
= \frac{\mathbb{E}[\varepsilon_n^2] - (1 - \rho)^2(\beta \lambda^{(\beta-1)})^2]}{\mathbb{E}[X_{n-1}^2] - (\beta \lambda^{(\beta-1)})^2]} \\
= \frac{(\beta \lambda^{(\beta-1)}(1 - \rho))^2 + \beta (1 - \beta) \lambda^{(\beta-2)}(1 - \rho^2) - ((1 - \rho)\beta \lambda^{(\beta-1)})^2]}{\beta (1 - \beta) \lambda^{(\beta-2)} + (\beta \lambda^{(\beta-1)})^2} = (1 - \rho^2).}

Using (2.24), it is evident that for parameters \( \beta \) and \( \lambda \), the estimators are not in closed form and hence the estimates are obtained by optimization. Thus, in our opinion, the asymptotic properties of the estimators for \( \beta \) and \( \lambda \) are not possible to established. To check the asymptotic property of estimator \( \hat{\rho}_n \) we simulate 1000 datasets each with length \( n = 5000 \) for \( \rho = 0.9 \) and find out the simulated values of \( \hat{\rho}_n \) for each simulation. For \( \sqrt{n}(\hat{\rho}_n - \rho) \) the sample mean is 0 and sample variance is 0.194 which shows the convergence discussed in Prop. 2.4. The histogram for simulated values of \( \sqrt{n}(\hat{\rho}_n - \rho) \) is given in Figure 2.

### 2.2.2. Moments based estimation

The parameters of positive tempered stable distribution cannot be estimated by classical maximum likelihood estimation method due to non availability of closed from probability density function. Also the CLS estimator provides an estimate for \( \rho \) only. Here we use method of moments to estimate the parameters \( \beta \) and \( \lambda \). The general idea used by Pearson (1894) is, the population moments about the origin are equal to the
corresponding moments of the sample data. In case of TAR(1) process moments for innovation distribution can be extracted using the corresponding Laplace transform as the density (2.12) is in complex form. The idea is to apply the method of moments on the error sample which can be easily obtained from the original series by first estimating the $\rho$ using CLS method. The sample estimates for $\mathbb{E}(\epsilon_i)$ and $\mathbb{E}(\epsilon_i^2)$ are given by

$$
m_1 = \frac{\sum_{i=1}^{n} \epsilon_i}{n} = \frac{\sum_{i=1}^{n} (X_{i+1} - \rho X_i)}{n},
$$

(2.26)

$$
m_2 = \frac{\sum_{i=1}^{n} \epsilon_i^2}{n} = \frac{\sum_{i=1}^{n} (X_{i+1} - \rho X_i)^2}{n}.
$$

(2.27)

Letting, $z = \frac{\sum_{i=1}^{n} (X_{i+1} - \rho X_i)}{n}$ and $r = \frac{\sum_{i=1}^{n} (X_{i+1} - \rho X_i)^2}{n}$ and using (2.9), (2.10), (2.26) and (2.27) we get the following relations

$$
\beta \lambda^{\beta-1} = \frac{z}{1 - \rho}, \quad \beta = 1 + \frac{(z^2 - r)\lambda}{z(1 + \rho)}.
$$

After some manipulation, the above two equations leads to following nonlinear equations

$$
\beta (\frac{\beta - 1}{d})^{(\beta - 1)} - \left( \frac{z}{1 - \rho} \right) = 0
$$

and

$$
(1 + d\lambda)\lambda^{d\lambda} - \left( \frac{z}{1 - \rho} \right) = 0.
$$

We use sequential least square optimization technique (SLSQP) to solve these equations which is is an iterative method to optimize a nonlinear model using the nonlinear least square model and sequential quadratic programming model (see Fu, Liu, and Guo (2019)). We use scipy.optimize which use SLSQP as an inbuilt method in python to solve nonlinear equations that gives the estimates for $\beta$ and $\lambda$. 

![Figure 2. Histogram plot for $\sqrt{n}(\hat{\rho} - \rho)$](image-url)
Table 1. Actual and estimated parameters values for different choices of parameters.

| Case   | Actual       | Estimated      |
|--------|--------------|----------------|
| Case 1 | $\rho = 0.9$, $\lambda = 1$, $\beta = 0.5$ | $\hat{\rho} = 0.895$, $\lambda = 1.13$, $\beta = 0.489$ |
| Case 2 | $\rho = 0.8$, $\lambda = 2$, $\beta = 0.7$ | $\hat{\rho} = 0.795$, $\lambda = 1.93$, $\beta = 0.70$ |
| Case 3 | $\rho = 0.75$, $\lambda = 1.5$, $\beta = 0.9$ | $\hat{\rho} = 0.755$, $\lambda = 1.60$, $\beta = 0.90$ |

2.2.3. Parameter estimation by tail index

The parameter estimation for the case $\lambda = 0$ (corresponding density given by (2.19)) is done by using the right tail behavior of $S$ for the TAR(1) process. For the TAR(1) process we have a positive stable innovation random variable $(1 - \rho^\beta)^{1/\beta}S$ with corresponding Laplace transform $e^{-(1-\rho^\beta)x^\beta}$. Note that the innovation terms are independent and hence the estimation can be done using empirical and theoretical tails. Using (2.2), it follows

$$\mathbb{P}((1 - \rho^\beta)^{1/\beta}S > x) = \mathbb{P}(S > (1 - \rho^\beta)^{-1/\beta}x) \sim \frac{((1 - \rho^\beta)^{-1/\beta}x)^{-\beta}}{\Gamma(1 - \beta)}, \text{as } x \to \infty. \quad (2.28)$$

Taking log on both sides of (2.28) and replacing $\mathbb{P}(S > (1 - \rho^\beta)^{-1/\beta}x) = G_x(x)$, where $G_x(x)$ represents the empirical tail for $\epsilon$, we get

$$\log(G_x(x)) = -\beta \log(x) + \log(1 - \rho^\beta) - \log(\Gamma(1 - \beta)), 0 < \beta < 1. \quad (2.29)$$

Considering $Y = \log(G_x(x))$ and $X = -\log(x)$, which yields $\beta$ as slope of the regressing $Y$ on $X$.

2.3. Simulation study

To check the efficacy of the estimation procedure we perform simulation. Here we have considered an AR(1) model with marginals to be distributed as one sided tempered stable distribution and the corresponding Laplace transform for error term $\epsilon_n$ is given in (2.8). In order to simulate a tempered stable AR(1) series, first we simulate data for innovation using the Laplace transform, see e.g. Ridout (2009). Ridout (2009) discussed the simulation of a continuous distribution function using its Laplace transform which turns out to be very helpful if the probability density function is not in closed form. We consider three different simulated datasets of length $n = 10,000$ each with three different parameter combinations and the results are summarized in the following Table 1.

The sample trajectory for an AR(1) process for positive tempered stable distribution with parameters $\rho = 0.9$, $\lambda = 2$ and for $\beta = 0.5$ is given in Figure 3a.

Next we generate 1000 samples of TAR(1) process with size 10,000 each to assess the performance of the estimation method. To ensure the stationarity here we assume AR(1) process with parameter $\rho < |1|$. Consider $\rho = 0.9$ and take tempered stable distribution parameters as $\beta = 0.5$ and $\lambda = 2$. The parameters are estimated using the method of moments for $\lambda$, $\beta$ and based on conditional least square for $\rho$ considering different samples. The corresponding box-plots for estimated parameter values are shown in Figure 3b.

For the particular case, $\lambda = 0$ which leads to one-sided positive stable autoregressive model, simulation is again done by the method given in Ridout (2009) and parameter
estimation is done by the tail index method using (2.29). We simulate a sample of 10,000 observations for \( q = 0.5 \) and \( b = 0.5 \). By taking values greater than a particular threshold such that we have some percentage of the sorted dataset greater than that threshold value, we get regression based estimate for \( b \). Taking a cutoff value of 0.009 gives the 30% of observation greater than the cutoff and gives the corresponding estimate of 0.36 for \( b \). For cutoff value 0.019, we get almost 20% of the observation greater than threshold value and corresponding estimate is 0.41 for \( b \) which is more close to true parameter value. Further, for \( k = 0.06 \) the estimated \( b \) is 0.49 which is even more close to the actual value. This clearly indicates that the accuracy of the estimated parameter depends on the threshold selected. The results are summarized in Table 2, where \( k \) represents the threshold value.

### 3. Ar(1) process with tempered stable innovations

Consider the stationary AR(1) process

\[
X_n = \rho X_{n-1} + \epsilon_n, \quad |\rho| < 1,
\]

where \( \epsilon_n \)'s are i.i.d positive tempered stable random variables defined in (2.3) and independent of \( X_{n-1} \). Let \( \Phi_{X_n}(s) \) and \( \Phi_{\epsilon_n}(s) \) are Laplace transform of \( X_n \) and \( \epsilon_n \) respectively. Now assuming \( X_0 \overset{d}= \epsilon_0 \) and recursively writing Equation (3.30), we get a moving average representation of AR(1) process as follows

\[
X_n = \sum_{i=0}^{n} \rho^i \epsilon_{n-i}, \quad \lambda > 0, \quad \beta \in (0, 1).
\]

The Laplace transform of \( X_n \) is

\[
\Phi_{X_n}(s) = \prod_{i=0}^{n} \Phi_{\epsilon}(\rho^i s) = \prod_{i=0}^{n} e^{\lambda \beta-(\lambda+\rho^i \beta)s^\beta} = e^{n \lambda \beta - \sum_{i=0}^{n} (\lambda+\rho^i \beta)s^\beta}.
\]
Proposition 3.1. For $\lambda = 0$, the $q$th order moment of $X_n$ is given by

$$E(X_n^q) = \frac{\Gamma(1 - q/\beta)}{\Gamma(1 - q)} \left( \sum_{i=0}^{n} \rho^{i\beta} \right)^{q/\beta}, \ 0 < q < \beta < 1.$$

Proof. Using the proposition (2.2), for a positive random variable $X$ with corresponding Laplace transform $\Phi(s)$, and for $p > 0$, it follows

$$\int_{0}^{\infty} \frac{d^n}{ds^n} [\Phi(s)]s^{p-1} = (-1)^n \Gamma(n) E(X^{n-p}).$$

Thus for $q \in (m-1, m)$, for an integer $m$, we have

$$E(X^q) = \frac{(-1)^m}{\Gamma(n-q)} \int_{0}^{\infty} \frac{d^n}{ds^n} [\Phi(s)]s^{n-q-1} ds.$$

For $q \in (0, 1)$, it follows

$$E(X_n^q) = \frac{-1}{\Gamma(1 - q)} \int_{0}^{\infty} \frac{d}{ds} \left[ \sum_{i=0}^{n} \rho^{i\beta} \right]s^{-q} ds$$

$$= \frac{1}{\Gamma(1 - q)} \int_{0}^{\infty} e^{n\lambda^\beta - \sum_{i=0}^{n} (\rho^i + \lambda)^\beta} \sum_{i=0}^{n} \rho^{i\beta} \sum_{i=0}^{n} \beta \rho^{i} (\rho^i s + \lambda)^{\beta-1} s^{-q} ds.$$

For $\lambda = 0$, it yields

$$E(X_n^q) = \frac{1}{\Gamma(1 - q)} \int_{0}^{\infty} e^{-\sum_{i=0}^{n} \rho^i u} u^{-q/\beta} \left( \sum_{i=0}^{n} \rho^{i\beta} \right) du (\text{put } s^\beta = u)$$

$$= \frac{\Gamma(1 - q/\beta)}{\Gamma(1 - q)} \left( \sum_{i=0}^{n} \rho^{i\beta} \right)^{q/\beta}, \ 0 < q < \beta.$$

\[\square\]

Remark 3.1. Using the same argument as in Prop. 2.1, we show that the marginal distributions of $X_n$ are infinitely divisible. It is easy to show that for $e^{n\lambda^\beta - \sum_{i=0}^{n} (\lambda + \rho s)^\beta}$, the exponent $\phi(s) = \sum_{i=0}^{n} (\lambda + \rho^i s)^\beta - n\lambda^\beta$ has completely monotone derivative, that is $(-1)^n \phi^{n}(s) \geq 0, n = 0, 1, 2, \ldots$.

Remark 3.2. The index of dispersion for $X_n$ is $ID(X_n) = \frac{\text{Var}(X_n)}{E(X_n)}$. Using the mean and variance of the process $X_n$, given by $E(X_n) = \frac{\beta \lambda^{\beta-1}}{1 - \rho}$ and $\text{Var}(X_n) = \frac{\beta(1 - \beta) \lambda^{\beta-2}}{1 - \rho^2}$, yields
\[ ID(X_n) = \frac{1 - \beta}{\lambda(1 + \rho)}, \] where \( 0 < \rho < 1, \beta \in (0, 1) \) and \( \lambda > 0 \). For \( \lambda > 1 \) the \( ID(X_n) < 1 \), which implies that the marginal distributions of \( X_n \) are under-dispersed for \( \lambda > 1 \).

### 3.1. Parameter estimation

To estimate the parameter \( \rho \) of the process (3.30), the conditional least square method discussed in previous section is used, which minimizes Equation (2.20) with respect to \( \theta \). Using the mean of \( \epsilon_n \) and Equations (2.20) (2.21) and (2.22), we can write

\[ Q_n(\theta) = \sum_{i=1}^{n} \left[ X_{i+1} - \rho X_i - \beta \lambda^{\beta-1} \right]^2. \]  

(3.31)

Minimizing (3.31) w.r.t \( \rho \) gives the following estimate

\[ \hat{\rho} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(X_{i+1} - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}. \]

The parameter estimation for \( \beta \) and \( \lambda \) is done again using the method of moments for \( \epsilon_n \), which leads to

\[ m_1 = \beta \lambda^{\beta-1}, \quad m_2 = (\beta \lambda^{\beta-1})^2 + \beta(1 - \beta) \lambda^{\beta-2}, \]

where \( m_1 \) and \( m_2 \) are first and second order sample moments for \( \epsilon_n \) respectively. After some manipulation, we get the following nonlinear equations

\[ \beta(c(\beta - 1))^{(\beta-1)} - m_1 = 0, \quad \left( \frac{\lambda}{c} + 1 \right)^{\lambda(\beta/c) + 1} - m_1 = 0, \]

where \( c = \frac{m_1}{m_2 - m_1} \). Using the optimization technique SLSQP, discussed in Section 3, these equations are solved in python to obtain the estimates \( \hat{\lambda} \) and \( \hat{\beta} \). The estimate of \( \lambda \) can also be obtained using the estimate \( \hat{\beta} \), denoted by

\[ \hat{\lambda} = \left( \frac{m_1}{\hat{\beta}} \right)^{\frac{1}{\beta-1}} - m_1 = 0. \]

### 3.2. Simulation study and real data application

To check the performance of the estimation method, we use simulated dataset. Using the method discussed in Ridout (2009), we simulate an independent sample of size 10,000 for error term using the Laplace transform of positive tempered stable random variable defined in (2.3). Then we recursively simulate AR(1) series by using fixed value of \( \rho \) and letting \( \epsilon_0 = X_0 \). The performance of the estimation method for three different

| Case 1 | Actual | Estimated |
|--------|--------|-----------|
| \( \rho = 0.9, \lambda = 1, \beta = 0.5 \) | \( \hat{\rho} = 0.8909, \hat{\lambda} = 1.01, \hat{\beta} = 0.5002 \) |
| Case 2 | \( \rho = 0.8, \lambda = 2, \beta = 0.7 \) | \( \hat{\rho} = 0.804, \hat{\lambda} = 2.037, \hat{\beta} = 0.69 \) |
| Case 3 | \( \rho = 0.75, \lambda = 1.5, \beta = 0.9 \) | \( \hat{\rho} = 0.75131, \hat{\lambda} = 1.56, \hat{\beta} = 0.89 \) |
sets of parameters is summarized in Table 3. The sample trajectory for AR(1) process with tempered stable innovations is given in Figure 4a. Further, to assess the performance of the estimation method, box plots of the estimated parameters $\hat{\rho}, \hat{\beta}$ and $\hat{\lambda}$ are plotted by taking multiple simulated random samples of innovations and generating multiple AR(1) time-series. Here we generate 1000 samples each of length $n = 10,000$ for $\rho = 0.9, \beta = 0.5$ and $\lambda = 2$ and estimate parameters for each sample. The box plot of estimated parameters is shown in Figure 4b.

3.2.1. Real data application for power consumption data
The power consumption data during the time of COVID-19 is extracted from Kaggle. The data contains the daily state-wise power consumption in India during the time period Jan 2, 2019 to May 23, 2020 that is for a period of 17 months and each data-point represents the power consumption at that time in Mega Units (MU). We work on the data during the period Apr 14, 2019 to Apr 29, 2020 for the Indian state Arunachal Pradesh. The sample path of the data is plotted in Figure 5a.

By simply looking at the data we can visualize the stationarity with some sharp spikes suggesting non-Gaussian behavior. We have applied Augmented Dickey–Fuller (ADF) test to check the stationarity of the data and the $p$-value comes out to be less than 0.05 which indicates that the null hypothesis of ADF test is rejected at 5% significance level and implies that the data is stationary. To determine the appropriate time series model, ACF and PACF plots are given in Figures 5b and 6a, respectively, which determines the appropriate time series model with significant lag.

Clearly, PACF plot for the dataset is significant till lag 1 and the ACF plots tails off, this means that AR(1) model would be a good fit for this data. We assume that the innovation term $\epsilon_n$ follows one sided tempered stable distribution. We fit the proposed AR(1) model with tempered stable distribution on the dataset. The estimated parameters using the above described methods are $\hat{\beta} = 0.91, \hat{\lambda} = 2.9$ and $\hat{\rho} = 0.60$. Using these estimated values, we simulate a dataset to assess the accuracy of the results. We simulate a data sample for $\hat{\lambda} = 2.8$ and $\hat{\beta} = 0.92$. Considering the relation $\epsilon_n = X_n - \rho X_{n-1}$

---

1https://www.kaggle.com/twinkle0705/state-wise-power-consumption-in-india
generates a dataset for the innovation term. To check whether both the simulated data using estimated parameters values and the actual data set follow same distribution or not we use two sample Kolmogorov–Smirnov (K-S) test and Mann–Whitney U test. The two sample K-S and Mann–Whitney U test both are non parametric tests which compare the distribution of two datasets. The two sample K-S test applied on actual error dataset and simulated error dataset, gives a $p$-value 0.056 and Mann–Whitney U test gives the $p$-value 0.377. Thus the null hypothesis for both the tests are accepted at 5% level of significance. This indicates that both the samples are from same distributions. Hence it is appropriate to model the considered power consumption time-series using AR(1) model with positive tempered stable innovations.

Using the estimated parameters from empirical power consumption dataset, we generate synthetic tempered stable errors series of same length as of empirical dataset. Also we generate another synthetic errors series by assuming the innovations to be Gaussian with mean $\mu$ and variance $\sigma^2$, where $\mu$ and $\sigma^2$ are estimated using empirical errors. The estimated parameters for tempered stable innovations are $\lambda = 2.89$ and $\beta = 0.92$ and the estimated parameters for Gaussian innovation $\mu = 0.84$ and $\sigma^2 = 0.049$. The KDE plots for empirical errors, tempered stable errors and Gaussian errors are given in Figure 6b.

Figure 5. Time series plot of the power consumption data (left panel) and the corresponding ACF plot (right panel).

Figure 6. PACF plot of the power consumption data (left panel) and the KDE plots of the empirical errors, simulated Gaussian errors and simulated tempered stable errors (right panel).
It is clearly evident that the density plot for Gaussian innovations does not match with the density plot for empirical innovations. However, the density plot for tempered stable innovations better models the empirical innovations. Moreover, the K-S test and Mann Whitney U test is accepted at 5% level of significance which justify to use AR(1) model with tempered stable innovations. To measure the model performance, Theil’s U test is applied which measures the accuracy based on distance between actual values and predicted values (see e.g. Bliemel (1973)). The model is said to perform well if the statistic value is close to 0. The two types of Theil’s U statistics are defined below. The type 1 statistic is defined as

\[ U_1 = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (F_i - X_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} F_i^2} + \sqrt{\frac{1}{n} \sum_{i=1}^{n} X_i^2}}, \]

where \( F_i \) is the forecasted value at time \( t = i \) and \( X_i \) is the corresponding actual value. Further, the type 2 statistic is defined as
\[
U_2 = \left[ \frac{\sum^n_{i=1} (F_i - X_i)^2}{\sum^n_{i=1} X_i^2} \right]^{1/2}.
\]

By allocating 80% of the power consumption data to training set and 20% data to testing set and using the forecasted values from AR(1) model with tempered stable innovations the \( U_1 \) and \( U_2 \) statistics respectively are 0.04 and 0.2 which indicates that the proposed model performs well for this dataset.

**Real data application for drinking water treatment data**

We study another dataset that is drinking water treatment plant (DWTP) data, which has a daily frequency. The dataset is extracted from Data World\(^2\). We study daily Turbidity data which is a measure of water clarity, for a time period of 1000 days from May 15, 2008 to Feb 09, 2011. The ACF and PACF of the data are shown in Figures 7b and 8a, respectively. One can observe that the PACF is significant for lag 1 and the ACF plot tails off indicating that the AR(1) would be a good fit for the data.

The \( p \)-value for the ADF test statistic is less than 0.05 indicating the stationarity of the data. The estimated parameters for the proposed AR(1) model with tempered stable innovations are \( \hat{\rho} = 0.69 \), \( \hat{\beta} = 0.49 \) and \( \hat{\lambda} = 0.04 \). These estimated parameters are further used to simulate tempered stable innovations for KDE plot. Assuming the error distribution to be Gaussian the estimated parameters are \( \mu = 2.68 \) and \( \sigma^2 = 28.64 \). The KDE plots for empirical errors, simulated tempered stable errors and simulated Gaussian errors are given in Figure 8b.

### 4. Conclusions

We introduced two kinds of tempered stable autoregressive models. The first model denoted by TAR(1) has one-sided tempered stable marginals. The second model discussed in Section 3 has one-sided tempered stable innovations. Autoregressive models with positively distributed innovations are quite common in literature for example inverse Gaussian, Gamma and one-sided stable are used. The autoregressive model with one-sided tempered stable innovations generalizes the inverse Gaussian and positive stable innovations based autoregressive models. We provide the parameter estimation method of the introduced model and assess the efficacy using simulated data. Further, we show the applicability of the introduced model on real life power consumption data. We hope that introduced models will be helpful in modeling of several real life series emerging from different areas like finance, economics, natural hazards, power consumption and others.

### Acknowledgements

The authors are grateful to the reviewers for several helpful comments and suggestions, which have led to improvements in the paper. N.B. would like to thank Ministry of Education (MoE),

\[^2\]https://data.world/epa/dwtp-data-source-water-data
India for supporting her PhD research. Further, A.K. would like to express his gratitude to Science and Engineering Research Board (SERB), India for the financial support under the MATRICS research grant MTR/2019/000286.

**Funding**

The authors are grateful to the reviewers for several helpful comments and suggestions, which have led to improvements in the paper. N.B. would like to thank Ministry of Education (MoE), India for supporting her PhD research. Further, A.K. would like to express his gratitude to Science and Engineering Research Board (SERB), India for the financial support under the MATRICS research grant MTR/2019/000286.

**References**

Abraham, B., and N. Balakrishna. 1999. Inverse Gaussian autoregressive models. *Journal of Time Series Analysis* 20 (6):605–18. doi:10.1111/1467-9892.00161.

Applebaum, D. 2004. *Lévy processes and stochastic calculus*. Cambridge: Cambridge University Press.

Balakrishna, N., and K. Shiji. 2014. Extreme value autoregressive model and its applications. *Journal of Statistical Theory and Practice* 8 (3):460–81. doi:10.1080/15598608.2013.813415.

Bliemel, F. 1973. Theil’s forecast accuracy coefficient: A clarification. *Journal of Marketing Research* 10 (4):444–7. doi:10.1177/002224377301000413.

Dewald, L. S., and P. A. W. Lewis. 1985. A new Laplace second-order autoregressive time series model-NLAR(2). *IEEE Transactions on Information Theory* 31 (5):645–51. doi:10.1109/TIT.1985.1057089.

Feller, W. 1971. *Introduction to probability theory and its applications*. Vol. II. (2 Ed.) New York: John Wiley.

Fu, Z., G. Liu, and L. Guo. 2019. Sequential quadratic programming method for nonlinear least squares estimation and its application. *Mathematical Problems in Engineering* 2019:1–8.

Gaver, D. P., and P. A. W. Lewis. 1980. First order autoregressive gamma sequences and point processes. *Adv. in Appl. Probab* 12 (3):727–45. doi:10.2307/1426429.

Grunwald, G. R. Hyndman, and L. Tedesco. 1996. A unified view of linear AR(1) models. Clayton: D.P. Monash University.

Klimkov, L. A., and P. I. Nelson. 1978. On conditional least squares estimation for stochastic processes. *The Annals of Statistics* 6 (3):629–42. doi:10.1214/aos/1176344207.

Kumar, A., and P. Vellaisamy. 2015. Inverse tempered stable subordinator. *Statistics & Probability Letters* 103:134–41. doi:10.1016/j.spl.2015.04.010.

Lawrance, A. J., and P. A. W. Lewis. 1980. The exponential autoregressive-moving average process EARMA(p, q). *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 42 (2):150–61. doi:10.1111/j.2517-6161.1980.tb01112.x.

Meerschaert, M. M., E. Nane, and P. Vellaisamy. 2012. Transient anomalous subdiffusions on bounded domains. *Proceedings of the American Mathematical Society. American Mathematical Society* 2013 (141):699–710. doi:10.1090/s0002-9939-2012-11362-0.

Omey, E., S. Van Gulck, and R. Vesilo. 2018. Semi-heavy tails. *Lithuanian Mathematical Journal* 58 (4):480–99. doi:10.1007/s10986-018-9417-0.

Ridout, M. S. 2009. Generating random numbers from a distribution specified by its Laplace transform. *Statistics and Computing* 19 (4):439–50. doi:10.1007/s11222-008-9103-x.

Rosiński, J. 2007. Tempering stable processes. *Stochastic Processes and their Applications* 117 (6):677–707. doi:10.1016/j.spa.2006.10.003.
Pearson, K. 1894. Contributions to the mathematical theory of evolution. Proceedings of the Royal Society of London. 54:329–333.
Samorodnitsky, G, and M. S. Taqqu. 1994. Stable non-Gaussian random processes. Boca Raton, FL: Chapman and Hall.
Schiff, J. L. 1999. The Laplace transform: Theory and applications. New York: Springer-Verlag.
Steutel, F. W., and K. Van Harn. 2004. Infinite divisibility of probability distributions on the real line. New York: Marcel Dekker.
Uchaikin, V. V, and V. M. Zolotarev. 1999. Chance and Stability. Utrecht: VSP.
Zhang, S., and X. Zhang. 2009. On the transition law of tempered stable Ornstein-Uhlenbeck processes. Journal of Applied Probability 46 (03):721–31. doi:10.1017/S0021900200005842.