Threshold Models in Time Series Analysis: Advancements and Applications

Seminar Paper

Prepared by:

Madhavi Panchangum

Student ID: 243093

Course: Forecasting with Time Series Analysis

Professor: Dr.Matei Demetrescu

Submission Date: 12 Jan 2025

Faculty of Statistics

Contents

1. Introduction

Time series analysis has become fundamental in statistical and econometric modeling, regarding the hidden patterns outlining over time, hence making forecasts possible across economic, financial, medical, and environmental fields. Classic models are mostly linear due to their simplicity in use and comprehensible mathematical properties. Reality is varied, however; complications will range from regime breaks to nonlinearity to state-dependent dynamics—a frequent occurrence well beyond what these linear models can explain (Granger, 1993; Hansen, 1997; Tong). This gap creates the need for developing more flexible approaches.

Threshold models represent an important innovation in time series analysis, and they are designed to address these challenges by incorporating nonlinearities and regime-specific behaviors. In contrast to linear models that rely on a single functional form, threshold models segment data into distinct regimes based on a threshold variable. Each regime is modeled separately, thus enabling nuanced analysis of phenomena characterized by sudden changes or nonlinear interactions (Tong, 1977; Tsay, 1989).

Threshold models have their origins in the pioneering work of Howell Tong in the late 20th century, especially with the development of the Self-Exciting Threshold Autoregressive (SETAR) model (Tong and Lim, 1980). Tong's framework laid the foundation for modeling piecewise linear relationships, which has evolved to include more sophisticated forms such as the Smooth Transition Autoregressive (STAR) model, Threshold Generalized Autoregressive Conditional Heteroskedasticity (TGARCH) model (Zakoian), and their hybrid counterparts (Smith, 2022).

Applications of threshold models reveal the flexibility with which complex systems may exhibit features not captured by a linear system. Examples include describing regime switching of market conditions and asymmetry of volatility patterns in finance (Marquez and Marti-Recober; Wu); identification of thresholds critical for time-ahead prediction in environmental studies of extreme events (T. Cai and Jiang, 2020); nonlinear dose-response in medicine with insights regarding the efficacy and safety of an intervention (Seber, 2003). Despite these strengths, threshold models do suffer from serious drawbacks: computational complexity, overfitting, and choosing the right threshold variables.

Modern computational approaches—particularly those including machine learning algorithms, Bayesian inference, and quantum-inspired methodologies—are changing this situation step by step (Yikun Zhang and Giessing, 2022). Hybrid methods combining machine learning with threshold models show a promising trend of dealing successfully with high-dimensionality datasets (X. Chen and Zhao, 2023), while the advances of the inference methodologies extend the arsenal to key policy evaluation or risk assessment work (Pearl, 2009).

2. Background of Threshold Models in Time Series Analysis

Traditional time series analysis has relied on linear models that describe the relationships among variables. These models are mathematically elegant and computationally efficient but usually cannot capture the richness of real data, which often exhibits nonlinear behavior with regime shifts and abrupt changes. It was this inability of linear models to address such intricacies that necessitated the development of threshold models, representing a significant departure from tradition by incorporating regime-dependent dynamics (Hansen, 1997; Tong).

2.1. Questions That Motivated Threshold Models

Threshold models were developed to address certain questions that highlighted the deficiencies of linear models in various fields:

Economics: Macroeconomic relationships, such as the Phillips curve, raised questions about state-dependent dynamics:

- Under what conditions does the relationship between inflation and unemployment hold?
- Why do policy interventions have different effects in recessions compared to booms?

These questions underlined the demand for models capable of capturing state-dependent dynamics, as discussed by Hansen (1997).

Finance: Financial markets presented unique challenges, such as:

- Why do stock return patterns and volatility depend on the regime (e.g., financial crises versus normal market conditions)?
- How do rapid market shifts impact investor sentiment and asset pricing?

These phenomena required models that could capture sudden changes and nonlinear relationships, as studied by Tsay (1989).Across these domains, a common observation was the existence of transitions around thresholds, requiring models designed to systematically handle such complexities.

2.2. Evolution of Threshold Models

The evolution of threshold models can be divided into three distinct phases:

Early Development Era (1960s–1980s):

- Tong (1977) pioneered the Threshold Autoregressive (TAR) model, introducing regime-dependent dynamics through piecewise linear modeling.
- The Self-Exciting Threshold Autoregressive (SETAR) model emerged as a foundational tool for analyzing nonlinear time series (Tong and Lim, 1980).

Expansion and Applications (1990s–2000s):

- Smooth Transition Autoregressive (STAR) and Threshold Vector Autoregressive (TVAR) models extended threshold modeling to multivariate datasets.
- Seminal contributions by Granger (1993) and Granger (1994) broadened the theoretical and practical realms of threshold models.

Recent Advances (2000s–Present):

- Advances in computational power have enabled the incorporation of Bayesian methods and machine learning techniques, improving threshold estimation and regime switch detection (Smith, 2022).
- Applications now span high-frequency trading, climate change analysis, and quantum-inspired methodologies for forecasting complex systems (Yikun Zhang and Giessing, 2022).

2.3. Critical Analysis and Limitations

While threshold models offer significant advantages, they also face challenges:

- **Threshold Selection:** Selecting the appropriate threshold variable and its value is computationally expensive and sensitive to data quality (Chan, 1993).
- **Overfitting:** The flexibility of threshold models increases the risk of overfitting, particularly with smaller datasets (Goracci et al.).
- **High-Dimensional Data:** Traditional estimation methods often struggle to scale with high-dimensional datasets, imposing significant computational challenges (Smith and Zhang, 2022).

Recent hybrid modeling approaches, such as reinforcement learning-based methods for automatic threshold selection, have shown promise in addressing these limitations, improving both generalizability and computational efficiency (X. Chen and Zhao, 2023).

3. The Threshold models

A broad class of non-linear models, here called threshold models, has the property that the models are either piecewise linear or may be more generally considered as linear models with time-varying parameters.The basic idea of a threshold model is piecewise linearization through the introduction of the indicator time series.This idea is called **Threshold Principle** by Tong in his paper " Threshold Models: 30 years on".Under the umbrella of the threshold principle (Tong 1990),there is a class of nonlinear time series models that models nonlinear dynamics based on a "piecewise" linear approximation via partitioning a state-space into several subspaces. The partition is typically dictated by a so-called "threshold" variable. In this section, we present a simple but frequently used form—the threshold autoregressive model; focusing on the developments after Tong (1990).

3.1. Threshold Auto-Regressive Models

A threshold autoregressive (TAR) model with $k\,(k \geq 2)$ regimes is defined as:

$$X_t = \sum_{i=1}^{k} \{b_{i0} + b_{i1} X_{t-1} + \cdots + b_{i,p_i} X_{t-p_i} + \sigma_i \varepsilon_t\}\, I(X_{t-d} \in A_i), \tag{3.1}$$

where $\{\varepsilon_t\} \sim \mathrm{IID}(0,1)$, $d, p_1, \ldots, p_k$ are positive integers, $\sigma_i > 0$, b_{ij} are parameters, and $\{A_i\}$ forms a partition of $(-\infty, \infty)$.

The partition is dictated by the threshold variable X_{t-d}, and d is called the delay parameter. This model, introduced by Tong (1977), is a type of self-exciting threshold model.

Ergodicity of Nonlinear AR Models

Under **Condition A**, the Markov chain with state space $\mathbb{R}^m$ and dynamics given by:

$$X_t = T(X_{t-1}) + \varepsilon_t, \quad t \geq 1, \quad T : \mathbb{R}^m \to \mathbb{R}^m, \tag{3.2}$$

is geometrically ergodic.

Condition A

- **A1 (Lipschitz Continuity):** T is Lipschitz continuous over $\mathbb{R}^m$. That is, there exists $M > 0$ such that $\forall x, y \in \mathbb{R}^m$,

  $$\|T(x) - T(y)\| \leq M\|x - y\|,$$

  where $\|\cdot\|$ denotes the Euclidean norm.
- **A2 (Stability):** $T(0) = 0$, and there exist constants $K > 0$ and $c > 0$ such that for all $t \geq 0$ and initial conditions $x_0 \in \mathbb{R}^m$,

  $$\|x_t\| \leq K e^{-ct}\|x_0\|,$$

  where x_t represents the deterministic skeleton derived from the system.
- **A3 (Innovation Process):** The noise process $\{\varepsilon_t\}$ satisfies one of the following:
  1. $\{\varepsilon_t\}$ are i.i.d. random variables with absolutely continuous distributions, everywhere positive probability density functions over $\mathbb{R}^m$, and $\mathbb{E}\|\varepsilon_t\| < \infty$.
  2. $\varepsilon_t = (e_t, 0, \ldots, 0)^\top$, where $\{e_t\}$ are i.i.d. random variables, each with an absolutely continuous distribution, everywhere positive probability density functions over $\mathbb{R}$, and $\mathbb{E}|e_t| < \infty$.

Example: 1

Given the TAR process:

$$X_t = \begin{cases} 0.6X_{t-1} + \epsilon_t, & \text{if } X_{t-1} \leq 1, \\ 0.4X_{t-1} + \epsilon_t, & \text{if } X_{t-1} > 1, \end{cases}$$

where $\epsilon_t \sim N(0, \sigma^2)$ is an independent and identically distributed (i.i.d.) random noise.

We know that the stationarity condition is $|\phi| < 1$. For *Regime 1* ($X_{t-1} \leq 1$): $|0.6| < 1$, so the regime is stationary. And for *Regime 2* ($X_{t-1} > 1$): $|0.4| < 1$, so this regime is also stationary.From this we can see that both regimes satisfy the stationarity condition for an AR(1) process.The noise term ϵ_t introduces variability, which allows X_t to transition between regimes by crossing the threshold ($\tau = 1$).We now move forward in calculating the expected value ($\mathbb{E}[X_t]$) and variance ($\mathbb{V}[X_t]$) for the process.

1. For *Regime 1* ($X_{t-1} \leq 1$):

$$\mathbb{E}[X_t|X_{t-1} \leq 1] = 0.6\mathbb{E}[X_{t-1}] + \mathbb{E}[\epsilon_t]$$

Since $\mathbb{E}[\epsilon_t] = 0$, this simplifies to:

$$\mathbb{E}[X_t|X_{t-1} \leq 1] = 0.6\mathbb{E}[X_{t-1}]$$

now the variance :

$$\text{Var}[X_t|X_{t-1} \leq 1] = 0.6^2\text{Var}[X_{t-1}] + \sigma^2$$

2. For *Regime 2* ($X_{t-1} > 1$): Similarly,

$$\mathbb{E}[X_t|X_{t-1} > 1] = 0.4\mathbb{E}[X_{t-1}]$$

Variance:

$$\text{Var}[X_t|X_{t-1} > 1] = 0.4^2\text{Var}[X_{t-1}] + \sigma^2$$

Assuming ergodicity, the overall expected value is the weighted average of the expectations in both regimes, which depends on the stationary distribution of X_{t-1}.

3.2. Self-Exciting Threshold Autoregressive (SETAR) Model

The SETAR model is a special case of the Threshold Autoregressive (TAR) model where the threshold variable used to determine the regime switches is the lagged value of the time series itself (e.g., X_{t-d})fan2003. This makes it a "self-exciting" model, as the series dynamically determines its own regimes based on past values. The model is expressed mathematically as:

$$X_t = \begin{cases} b_{10} + b_{11}X_{t-1} + \cdots + b_{1p_1}X_{t-p_1} + \sigma_1\epsilon_t & \text{if } X_{t-d} \in A_1, \\ b_{20} + b_{21}X_{t-1} + \cdots + b_{2p_2}X_{t-p_2} + \sigma_2\epsilon_t & \text{if } X_{t-d} \in A_2, \\ \vdots & \vdots \\ b_{k0} + b_{k1}X_{t-1} + \cdots + b_{kp_k}X_{t-p_k} + \sigma_k\epsilon_t & \text{if } X_{t-d} \in A_k, \end{cases} \tag{3.3}$$

Here:

- X_t is the time series under consideration.
- X_{t-d} is the threshold variable, and d is the delay parameter.
- k is the number of regimes or partitions of the real line $(-\infty, \infty)$, with $\{A_1, A_2, \ldots, A_k\}$ forming these partitions.
- Each regime i has its own autoregressive parameters $b_{i0}, b_{i1}, \ldots, b_{ip_i}$, standard deviation σ_i, and number of lags p_i.
- $\{\epsilon_t\}$ is an IID error term with mean 0 and variance 1.

Example 2 (A Special SETAR model)

Consider the following modified threshold autoregressive model:

$$X_t = \begin{cases} -\alpha X_{t-1}^2 + \epsilon_t & \text{if } X_{t-1} \geq 0, \\ \beta |X_{t-1}| + \epsilon_t & \text{if } X_{t-1} < 0, \end{cases}$$

where $|\alpha| < 1$, $|\beta| < 1$, and ϵ_t is a symmetric, zero-mean innovation (e.g., $\epsilon_t \sim N(0, \sigma^2)$). The stationary distribution of this model, $\Pi(y)$, satisfies the integral equation:

$$\Pi(y) = \int_{-\infty}^{\infty} \Pi(x) p_\epsilon(y - f(x)) dx,$$

where $f(x)$ represents the piecewise autoregressive function:

$$f(x) = \begin{cases} -\alpha x^2 & \text{if } x \geq 0, \\ \beta |x| & \text{if } x < 0. \end{cases}$$

Exploiting the symmetry of ϵ_t, the integral equation can be decomposed into two regimes based on the sign of x:

$$\Pi(y) = \int_0^{\infty} \Pi(x) p_\epsilon(y + \alpha x^2) dx + \int_{-\infty}^{0} \Pi(x) p_\epsilon(y - \beta |x|) dx.$$

Using the symmetry property of $\Pi(x)$, such that $\Pi(-x) = \Pi(x)$, and the symmetry of $p_\epsilon(\cdot)$, the second term can be transformed as follows:

$$\int_{-\infty}^{0} \Pi(x) p_\epsilon(y - \beta |x|) dx = \int_0^{\infty} \Pi(x) p_\epsilon(y - \beta x) dx.$$

Thus, the stationary distribution simplifies to:

$$\Pi(y) = \int_0^{\infty} \Pi(x) \left[p_\epsilon(y + \alpha x^2) + p_\epsilon(y - \beta x) \right] dx.$$

For Gaussian innovations ($\epsilon_t \sim N(0, \sigma^2)$), the function $p_\epsilon(z)$ is given by:

$$p_\epsilon(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{z^2}{2\sigma^2} \right),$$

allowing numerical evaluation of $\Pi(y)$. The quadratic nonlinearity, combined with the piecewise behavior of $f(x)$, creates a stationary distribution with pronounced skewness and kurtosis, influenced by α and β. For Gaussian ϵ_t, numerical solutions reveal a family of skewed distributions where the extent of skewness and tail behavior depends on the interplay between the regimes. This model has applications in systems with distinct nonlinear responses in growth and contraction phases, such as economic or ecological dynamics, and provides a valuable extension to traditional SETAR models by incorporating quadratic and absolute-value thresholds.

3.3. Exponential Autoregressive (EAR) Model

The Exponential Autoregressive (EAR) model is a specialized case of the Threshold Auto-Regressive (TAR) model, where the time series $\{X_t\}$ is designed to exhibit an exponential marginal distribution. The EAR model was first introduced by Lawrance and Lewis (1980), to model those situations in which an exponential marginal distribution is desirable. Later, it was recognized by Tong (1983) as a subclass of the TAR model, and the extension to higher orders was made by Chan (1986, 1988).The regime-switching mechanism is governed by an indicator variable J_t, which takes the following values:

$$J_t = \begin{cases} 1, & \text{with probability } 1-\alpha, \\ 2, & \text{with probability } \alpha, \end{cases}$$

where $0 < \alpha < 1$ and J_t is independently and identically distributed.

The model is expressed as:

$$X_t = a(J_t)_1 X_{t-1} + a(J_t)_2 X_{t-2} + \epsilon_t, \tag{3.4}$$

where the parameters $a(J_t)_1$ and $a(J_t)_2$ depend on the regime J_t and are specified as:

$$\begin{aligned} a(1)_1 \geq 0, \quad a(1)_2 = 0 \quad \text{for } J_t = 1, \\ a(2)_1 = 0, \quad a(2)_2 \geq 0 \quad \text{for } J_t = 2. \end{aligned}$$

Here, ϵ_t represents an independent noise term, chosen such that the marginal distribution of X_t is exponential. The EAR model captures nonlinear dynamics by switching between two autoregressive regimes based on the value of J_t.

Stationarity Theorem for the Exponential Autoregressive (EAR) Model

Theorem 1: Let $\{X_t\}$ be a time series governed by the EAR model:

$$X_t = a(J_t)_1 X_{t-1} + a(J_t)_2 X_{t-2} + \epsilon_t,$$

where J_t is a sequence of i.i.d. random variables with:

$$P(J_t = 1) = 1 - \alpha, \quad P(J_t = 2) = \alpha, \quad (0 < \alpha < 1),$$

and ϵ_t is an independent noise term with an appropriate distribution to ensure X_t has an exponential marginal distribution. If the parameters satisfy:

$$\mathbb{E}[a(J_t)_1] + \mathbb{E}[a(J_t)_2] < 1,$$

then $\{X_t\}$ is strictly stationary, meaning its joint distribution does not depend on time t, and it has an exponential marginal distribution.

3.4. Threshold GARCH (TGARCH) Model

The Threshold GARCH model is a modification of the classical GARCH framework, designed to capture asymmetries in volatility dynamics specially. Unlike the standard GARCH model, where conditional variance is symmetric with respect to past shocks, TGARCH separates the positive and negative parts of innovations, thus allowing different reactions of volatility to shocks depending on their signs.

Let ϵ_t represent a discrete-time process with:

$$\epsilon_t = \sigma_t Z_t, \quad Z_t \sim \text{i.i.d.}, \tag{3.5}$$

$$\mathbb{E}[Z_t] = 0, \quad \text{Var}(Z_t) = 1. \tag{3.6}$$

The conditional standard deviation σ_t in a TGARCH(p, q) model is given by:

$$\sigma_t = \sqrt{\alpha_0 + \sum_{i=1}^{q} \left(\alpha_i^+ \epsilon_{t-i}^+ + \alpha_i^- \epsilon_{t-i}^- \right) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2}, \tag{3.7}$$

where:

- $\epsilon_{t-i}^+ = \max(\epsilon_{t-i}, 0)$ is the positive part of the lagged innovation,
- $\epsilon_{t-i}^- = \min(\epsilon_{t-i}, 0)$ is the negative part of the lagged innovation,
- $\alpha_0 > 0$, $\alpha_i^+, \alpha_i^- \geq 0$, and $\beta_j \geq 0$ are model parameters ensuring non-negativity of σ_t^2. (2)

This specification lets the current volatility σ_t react differently to positive and negative past shocks. The TGARCH model is particularly convenient because it models the conditional standard deviation directly, hence strict positivity constraints on the parameters are not required. Consequently, σ_t^2 is guaranteed to be non-negative.

TGARCH models can be useful to capture the so-called leverage effect in financial time series, where a negative shock is often followed by higher volatility than a positive shock of the same magnitude. Another advantage of the TGARCH model is that the additive structure simplifies numerical inference and can treat asymmetries in a more differentiated way for different lags.

Linearity of Threshold GARCH

From (3.5) and (3.7) we have:

$$E_t^+ = \alpha Z_t^+ \quad \text{and} \quad E_t^- = \alpha Z_t^-.$$

Then, using (3.5), the conditional means of the positive and negative parts are:

$$\mathbb{E}[E_t^+ | \mathcal{F}_{t-1}] = \alpha^+ \mathbb{E}[Z_t^+] \quad \text{and} \quad \mathbb{E}[E_t^- | \mathcal{F}_{t-1}] = \alpha^- \mathbb{E}[Z_t^-].$$

Let u_t and v_t denote the respective innovations of E_t^+ and E_t^-; we have:

$$E_t^+ - u_t = \alpha^+ \mathbb{E}[Z_t^+] \quad \text{and} \quad E_t^- - v_t = \alpha^- \mathbb{E}[Z_t^-].$$

Consider

$$\sigma_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i^+ (\varepsilon_{t-i})^+ + \sum_{i=1}^{q} \alpha_i^- (\varepsilon_{t-i})^- + \sum_{j=1}^{p} \beta_j \sigma_{t-j}$$

Multiplying this by $\mathbb{E}[Z_t^+]$ first and then by $\mathbb{E}[Z_t^-]$ yields:

$$E_t^+ = \alpha_0 \mathbb{E}[Z^+] + \sum_{i=1}^{q} \left(\alpha_i^+ E_{t-i}^+ - \alpha_i^- E_{t-i}^- \right) \mathbb{E}[Z^+] + \sum_{j=1}^{p} \beta_j \left(E_{t-j}^+ - v_{t-j} \right) + u_t,$$

$$E_t^- = \alpha_0^- \mathbb{E}[Z^-] + \sum_{i=1}^{q} \left(\alpha_i^- E_{t-i}^- - \alpha_i^+ E_{t-i}^+ \right) \mathbb{E}[Z^-] + \sum_{j=1}^{p} \beta_j \left(E_{t-j}^- - v_{t-j} \right) + v_t.$$

Therefore, the vector (E_t^+, E_t^-) is a solution of an ARMA$(\max(p,q),p)$ equation. This is analogous to the GARCH case, in which the same property is true for the (σ_t^2) process.

To evaluate whether shocks to variance persist or not, it is necessary to analyze the stationarity properties of the error process. As in the GARCH context, a general study is difficult. We concentrate on two important cases: the threshold ARCH(q) and the threshold GARCH(1,1).

3.5. Threshold Volatility Model

Threshold Volatility Models are a class of econometric models used to describe the dynamics of volatility in time series data, where volatility is influenced by the crossing of certain threshold levels in the underlying process. These models capture asymmetric behavior in volatility based on the regime (state) the system is in, making them suitable for analyzing financial time series, which often exhibit regime-dependent patterns like bull and bear markets.

The key feature of these models is that the volatility dynamics depend on whether the value of an indicator variable crosses a threshold.

General Formula of Threshold Volatility Models

Threshold volatility models are extensions of standard autoregressive conditional heteroskedasticity (ARCH) or generalized ARCH (GARCH) models. They can be expressed as:

Conditional Variance Equation

$$\sigma_t^2 = \begin{cases} \omega_1 + \sum_{i=1}^{q} \alpha_{i,1} \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{j,1} \sigma_{t-j}^2, & \text{if } z_{t-d} \leq \tau, \\ \omega_2 + \sum_{i=1}^{q} \alpha_{i,2} \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{j,2} \sigma_{t-j}^2, & \text{if } z_{t-d} > \tau, \end{cases} \tag{3.8}$$

where:

- σ_t^2: Conditional variance (volatility) at time t,
- ω_1, ω_2: Constants in each regime,
- $\alpha_{i,1}, \alpha_{i,2}$: Coefficients for lagged squared errors in each regime,
- $\beta_{j,1}, \beta_{j,2}$: Coefficients for lagged variances in each regime,
- z_{t-d}: Threshold variable (e.g., lagged returns, external variable) observed with a delay d,

- τ: Threshold value,
- ϵ_t: Error term (e.g., innovations).

Stability Theorem for Threshold Volatility Models

Theorem 2: A Threshold Volatility Model (TVM) is stable if and only if:

1. The conditional variance process $\{\sigma_t^2\}$ is strictly stationary and ergodic.
2. The sum of the autoregressive ($\alpha_{i,k}$) and moving average ($\beta_{j,k}$) coefficients for each regime k satisfies:
$$\sum_{i=1}^{q} \alpha_{i,k} + \sum_{j=1}^{p} \beta_{j,k} < 1, \quad \forall k \,\text{(regimes)}.$$

Corollary: If the threshold variable z_{t-d} is deterministic or follows a non-stationary process (e.g., a unit root process), the TVM may lose stability even if the stationarity conditions are satisfied within each regime.

3.6. Recent Developments in Threshold Models: Advanced Formulations and Applications

Threshold models have experienced remarkable advancements over the past three decades, incorporating sophisticated formulations to address contemporary analytical challenges. One significant development is the integration of sparse modeling techniques, such as LASSO-based threshold estimation. In high-dimensional datasets, the inclusion of a penalty term, $\lambda\|\beta\|_1$, ensures effective variable selection while addressing overfitting concerns, especially when the sparsity level, $\|\beta\|_0$, is large (Fan and Li, 2001; Tibshirani, 1996).

Smooth Transition Regression (STR) models have revolutionized the modeling of gradual regime transitions. These models are expressed as:

$$y_t = \phi_0 + \phi_1 x_t + G(\gamma, c)(\phi_2 x_t - \phi_1 x_t) + \epsilon_t,$$

where $G(\gamma, c) = (1 + \exp(-\gamma(x_t - c)))^{-1}$ represents a logistic transition function. This function controls the smoothness of the transition, capturing more nuanced dynamics in nonlinear systems (Granger and Teräsvirta, 1993; Teräsvirta, 1994).

Multivariate extensions, such as Threshold Vector Autoregressive (TVAR) models, allow the analysis of interconnected time series. These models are defined as:

$$X_t = \begin{cases} \Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p} + \epsilon_t, & \text{if } q_t \leq \theta, \\ \Phi_{1,k} X_{t-1} + \cdots + \Phi_{p,k} X_{t-p} + \epsilon_t, & \text{if } q_t > \theta, \end{cases}$$

where q_t is the threshold variable, and θ is the threshold parameter (Tong, 1983; Tsay, 1998).

Bayesian approaches have further enhanced threshold models by incorporating prior knowledge and quantifying parameter uncertainty. The posterior distribution is given by:

$$\pi(\theta, \phi | y) \propto L(y|\theta, \phi)\pi(\theta)\pi(\phi),$$

where $L(\cdot)$ is the likelihood function, and $\pi(\cdot)$ represents the prior distributions (Geweke, 1993; Koop and Potter, 2003).

Finally, significant attention has been given to chaotic dynamics and stability in threshold models. For instance, ensuring ergodicity in a TAR process requires satisfying conditions like $\sum_{j=1}^{k} |a_j| < 1$, which guarantee stationarity (Hansen, 1999; Tong, 1990). These advancements have expanded the applicability of threshold models across disciplines such as finance, climate science, and biological systems.

3.7. Future Applications of Threshold Models

Future applications of threshold models hold immense potential across various fields. In finance, they can analyze cryptocurrency volatility and regime shifts in investor sentiment. In climate science, threshold models are crucial for predicting extreme weather events and understanding climate change impacts. In healthcare, they can model disease outbreaks and drug response patterns. Applications of threshold models also extend to IoT and AI systems for edge computing decisions and monitoring system performance. Their adaptability and robustness ensure their continued relevance in emerging fields such as neuroscience, where they model brain dynamics across different cognitive states.

4. Statistical Analysis

4.1. Threshold AR and Linear AR Processes

While Linear AR models assume a single linear relationship between the current and lagged values of a time series, TAR models introduce nonlinearity by incorporating regime-dependent behavior. The TAR model operates with a threshold value, τ, which divides the process into two or more regimes. Mathematically, a TAR process can be expressed as:

$$X_t = \begin{cases} \phi_1 X_{t-1} + \epsilon_t, & \text{if } X_{t-1} \leq \tau, \\ \phi_2 X_{t-1} + \epsilon_t, & \text{if } X_{t-1} > \tau, \end{cases}$$

where ϕ_1 and ϕ_2 are regime-specific coefficients, and $\epsilon_t \sim N(0, \sigma^2)$ is Gaussian noise. In contrast, the Linear AR model follows a simpler formulation:

$$X_t = \phi X_{t-1} + \epsilon_t,$$

with a single coefficient ϕ and noise term ϵ_t.

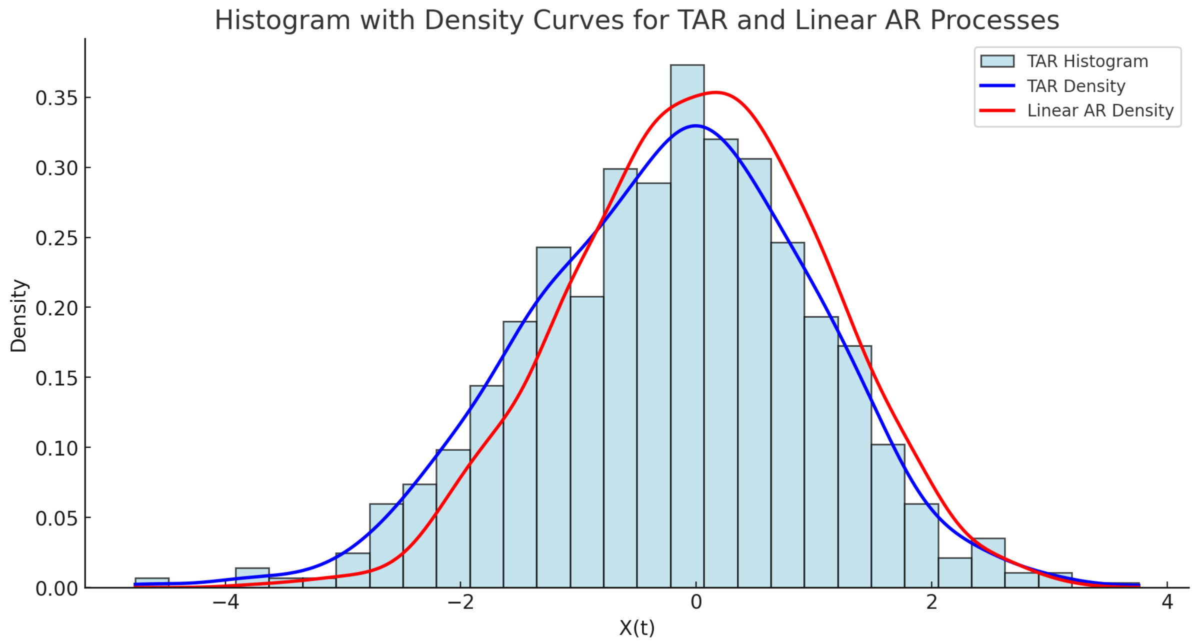

Figure 1: Histogram with Density Curves for TAR and Linear AR Processes. The TAR process exhibits a multimodal distribution due to regime switching, while the Linear AR process has a simpler unimodal distribution.

The plots illustrate the simulated TAR and Linear AR processes over 1,000 time points. The TAR process, shown as a blue line, exhibits nonlinear behavior by switching between two regimes depending on whether the previous value X_{t-1} is above or below the threshold $\tau = 1$. The Linear AR process, represented by the red dashed line, maintains a consistent linear relationship with $\phi = 0.5$.

The histogram and density overlay further compare the stationary distributions of the TAR and Linear AR processes. The TAR process displays a multimodal distribution, characteristic of its regime-switching nature, while the Linear AR process exhibits a unimodal distribution, indicative of its linear dynamics. The blue and red curves, derived from kernel density estimation, reinforce the contrast between the complexity of the TAR model and the simplicity of the Linear AR model. These visualizations underscore the flexibility of TAR models in capturing nonlinearities and regime-dependent dynamics in time series data.

Time Average Plot for Ergodicity Analysis of TAR Process

The simulation generates trajectories for three initial values, computes their cumulative time averages $\bar{X}_t$, and plots them.The time-averaged plot demonstrates the convergence behavior of the TAR process over time. As t increases, the trajectories of the time averages $\bar{X}_t$ corresponding to different initial values gradually converge towards a common value, indicating the independence of the long-term behavior from the initial conditions. This behavior is a hallmark of ergodicity, suggesting that the TAR process satisfies the ergodic property.

Ergodicity implies that the statistical properties computed from a single sufficiently long trajectory of the process are equivalent to those computed from an ensemble of trajectories. For example, the time average of X_t approximates its mean $\mathbb{E}[X_t]$ in the stationary distribution.The following figure illustrates the time-average plot for a Threshold Autoregressive (TAR) process, showcasing the convergence of time-averaged state distributions across different initial conditions.

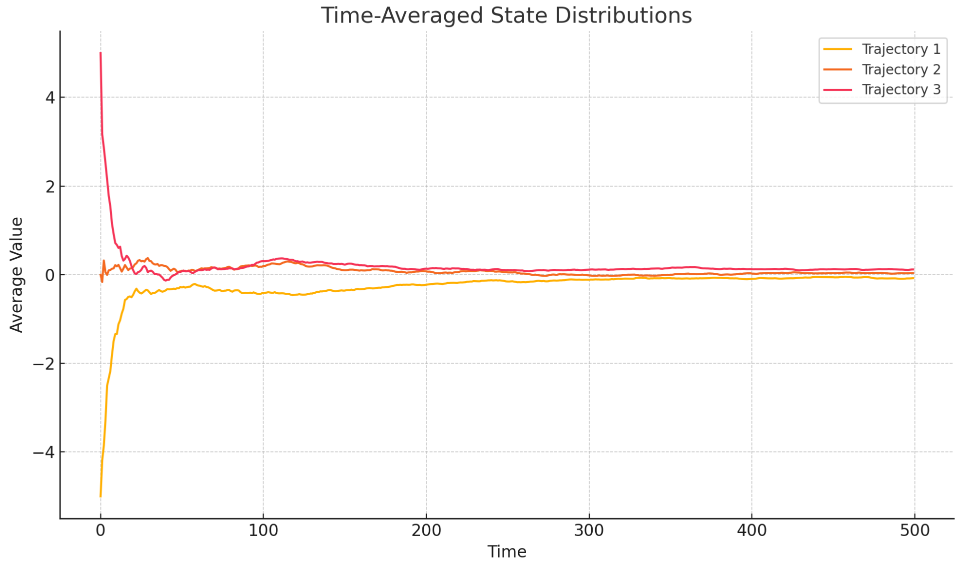

Figure 2: Time-average plot for a TAR process.

In the simulated plot, the initial differences among the trajectories diminish as the process progresses. This convergence underscores the stability of the TAR process, driven by the coefficients α_1 and α_2, which ensure the stationarity of each regime. The histogram of state distributions complements the time-average plot, confirming that the process spends more time in regions with higher density, contributing to the stabilization of the time average.

4.2. Stationarity and Linearity Check

The dataset analyzed is the built-in **AirPassengers** dataset, which contains monthly airline passenger counts from 1949 to 1960. To address potential heteroscedasticity and stabilize variance,
the data was log-transformed as follows:

$$y_t = \log(X_t)$$

where X_t represents the raw passenger counts at time t, and y_t is the log-transformed series. This transformation reduces the impact of large variations and highlights underlying trends in the time series.

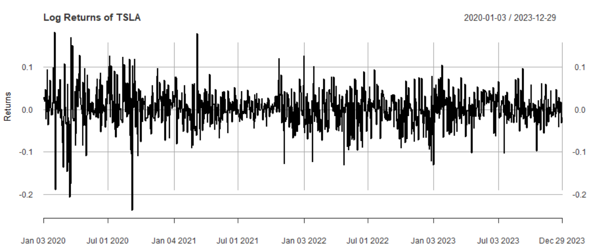

Figure 3: Log-transformed AirPassengers dataset

Stationarity Check

To assess stationarity, the **Augmented Dickey-Fuller (ADF) test** was conducted. The null hypothesis (H_0) assumes the presence of a unit root, indicating non-stationarity, while the alternative hypothesis (H_1) suggests stationarity. The test results, listed in the **Appendix** as **Listings:1**, are as follows:

- **Test Statistic:** -1.7176
- **p-value:** 0.4236

Since the p-value exceeds the significance level ($\alpha = 0.05$), H_0 cannot be rejected. This indicates that the time series is non-stationary. Additionally, the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots exhibit significant autocorrelations at multiple lags, reinforcing the non-stationarity of the data.

Linearity Check

To evaluate the linearity of the time series, an **AR(1) model** was fitted:

$$y_t = \phi_1 y_{t-1} + \epsilon_t$$

where y_t is the log-transformed series, ϕ_1 is the autoregressive coefficient, and ϵ_t is the residual error. A **RESET test** was performed to detect non-linear patterns in the residuals by testing for higher-order fitted terms ($\hat{y}_t^2, \hat{y}_t^3$). The test results, shown in **Listings:2**, are as follows:

- **RESET Test Statistic:** 0.2265
- **p-value:** 0.7976

Since the p-value is greater than 0.05, there is no significant evidence of non-linearity, and the null hypothesis (H_0) of linearity cannot be rejected.

From **Listings:3**, the residuals were further tested for autocorrelation using the **Ljung-Box test**. The test results are as follows:

- **Ljung-Box Test Statistic:** 6.7837
- **p-value:** 0.5613

As the p-value exceeds 0.05, the residuals do not exhibit significant autocorrelation, validating the adequacy of the AR(1) model under the linearity assumption.

Conclusion

The ADF test confirms that the time series is non-stationary, requiring further transformation or differencing to achieve stationarity. However, the linearity checks (RESET and Ljung-Box tests) confirm that the AR(1) model effectively captures the linear structure of the data, with no significant evidence of non-linearity or residual autocorrelation.

4.3. Analysis of Tesla Log Returns using TGARCH Model

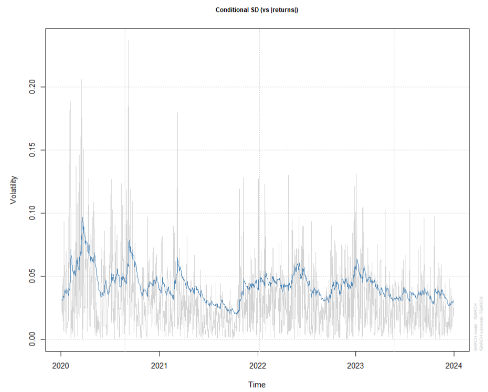

Figure 4: Tesla Log returns using TGARCH.

The conditional standard deviation (volatility) plot and the QQ-plot of standardized residuals. These outputs provide insights into the model's performance and the behavior of Tesla's log returns.

The conditional variance in the TGARCH model is expressed as:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I(\varepsilon_{t-1} < 0) + \beta\sigma_{t-1}^2,$$

where:

- ω: Constant in variation.
- α: Characterizes the effect of prior shock squares on current volatility.
- γ: Represents the asymmetric effect; negative shocks ($\gamma > 0$) have a larger impact on volatility compared to positive ones.
- β: Reflects volatility persistence.

Observations from the Plot

The conditional standard deviation plot demonstrates the time-varying volatility σ_t. Notable characteristics include:

- High clustering of volatility, indicating sustained periods of high or low volatility.
- Spikes around major market events.
- Agreement with financial market characteristics, confirming the TGARCH model's ability to capture asymmetries in volatility dynamics.

QQ-Plot of Standardized Residuals

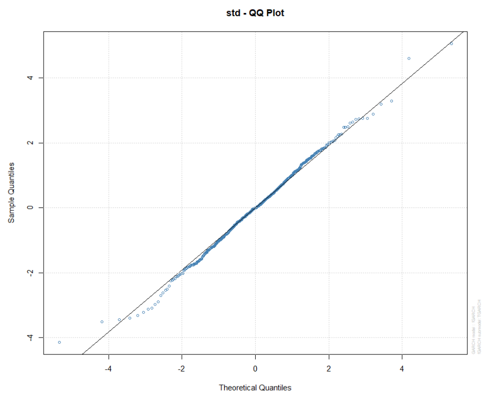

Figure 5: Tesla Log returns using TGARCH.

Standardized Residuals

The QQ-plot compares the quantiles of standardized residuals, $z_t = \frac{\varepsilon_t}{\sigma_t}$, with the theoretical quantiles of a Student's t-distribution. Key observations include:

- Points closely aligned with the diagonal line suggest residuals are consistent with the assumed distribution.
- Deviations in the tails may indicate heavy-tailed behavior typical in financial returns.

RESET Test for Non-Linearity

The RESET test results show no significant non-linearity in residuals ($p > 0.05$).

Ljung-Box Test for Autocorrelation

The Ljung-Box test indicates no significant autocorrelation in residuals ($p > 0.05$).

Both the RESET and Ljung-Box tests confirm the adequacy of the TGARCH model:

- Residuals behave as an independent and identically distributed (i.i.d.) series.
- The model effectively maps volatility dynamics in Tesla's log returns.

The TGARCH model provides a robust framework for analyzing Tesla's log returns. The plots and diagnostic tests validate the model's capability in capturing volatility dynamics and ensuring residuals meet necessary conditions for risk modeling.

5. Summary

Threshold models have revolutionized time series analysis by overcoming the inability of traditional linear models to capture such complex, nonlinear, and regime-dependent dynamics. According to these models, data fall into a number of distinct regimes that depend on a threshold variable, whereby temporal patterns and abrupt shifts may be represented with nuance. Starting from the seminal work by Howell Tong, threshold models have grown into advanced forms like TGARCH and STAR; their applications are really heterogeneous, including problems from finance and medicine to environmental studies. Although these models have impressive advantages for modeling complex systems, threshold choice, overfitting, and computational complexity still remain problems. Recent developments in machine learning, Bayesian inference, and hybrid modeling have enhanced its predictive power and scope of applicability. This paper provides a broad overview of the development and methodologies of threshold models, with applications, to assert their flexibility in meeting the present analytical needs.

A. Stationarity and Linearity Check

A.1. Augmented Dickey-Fuller (ADF) Test

```r
# ADF Test for Stationarity
adf_result <- adf.test(real_data)
cat("ADF Test Statistic:", adf_result$statistic, "\n")
cat("ADF p-value:", adf_result$p.value, "\n")
```

Listing 1: Augmented Dickey-Fuller (ADF) Test Code

Test Results:

- **Test Statistic:** -1.7176
- **p-value:** 0.4236

Interpretation: The time series is not stationary as $p > 0.05$.

A.2. RESET Test for Linearity

```r
# RESET Test for Linearity
reset_test <- resettest(linear_model$residuals ~ time(real_data), power = 2:3, type = "fitted")
cat("RESET Test Statistic:", reset_test$statistic, "\n")
cat("RESET p-value:", reset_test$p.value, "\n")
```

Listing 2: RESET Test Code

Test Results:

- **RESET Test Statistic:** 0.2265
- **p-value:** 0.7976

Interpretation: No evidence of non-linearity as $p > 0.05$.

A.3. Ljung-Box Test for Residual Autocorrelation

```r
# Ljung-Box Test for Residual Autocorrelation
box_test <- Box.test(linear_model$residuals, lag = 10, type = "Ljung-Box")
cat("Ljung-Box Test Statistic:", box_test$statistic, "\n")
cat("Ljung-Box p-value:", box_test$p.value, "\n")
```

Listing 3: Ljung-Box Test Code

Test Results:

- **Ljung-Box Test Statistic:** 6.7837
- **p-value:** 0.5613

Interpretation: Residuals do not indicate significant autocorrelation as $p > 0.05$.

References

K. S. Chan. Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Annals of Statistics*, 21(2):520–533, 1993.

Jianqing Fan and Runze Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360, 2001.

John Geweke. Bayesian treatment of the independent student-t linear model. *Journal of Applied Econometrics*, 8(S1):S19–S40, 1993.

Greta Goracci, Davide Ferrari, Simone Giannerini, and Francesco Ravazzolo. Robust estimation for threshold autoregressive moving-average models.

Teräsvirta Granger, C. W. J. Modelling nonlinear economic relationships. oxford university press.. *Journal of the Royal Statistical Society*, 1993.

Teräsvirta Granger, C. W. J. Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the Royal Statistical Society*, 1994.

Clive WJ Granger and Timo Teräsvirta. *Modelling Nonlinear Economic Relationships*. Oxford University Press, 1993.

B. E. Hansen. Inference in tar models. *Studies in Nonlinear Dynamics Econometrics*, 2(1): 1–14, 1997.

Bruce E Hansen. Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics*, 93(2):345–368, 1999.

Gary Koop and Simon M Potter. Bayesian analysis of endogenous delay threshold models. *Journal of Business & Economic Statistics*, 21(1):93–103, 2003.

Munoz Marquez and Marti-Recober. Threshold volatility models: Forecasting performance.

J. Pearl. Causal inference in nonlinear systems. *Foundations of Statistics*, 2009.

Wild C. J. Seber, G. A. F. Nonlinear regression. wiley series in probability and statistics. *Journal of the Royal Statistical Society*, 2003.

J. Smith and Y. Zhang. Advances in hybrid threshold models. *Journal of Data Science*, 2022.

John Smith. Threshold models in pharmacology. *Journal of Medical Statistics*, 34:12–25, 2022.

J. Fan T. Cai and T. Jiang. Multiscale approaches for nonlinear time series analysis. *Annals of Statistics*, 2020.

Timo Teräsvirta. Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425):208–218, 1994.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

H. Tong and K. S. Lim. Threshold autoregression, limit cycles, and cyclical data. *Journal of the Royal Statistical Society, Series B*, 42:245–292, 1980.

Howell Tong. *Threshold Models: 30 Years On.*

Howell Tong. Threshold models in non-linear time series analysis. *Journal of the Royal Statistical Society*, 42:245–292, 1977.

Howell Tong. *Threshold Models in Nonlinear Time Series Analysis.* Springer, 1983.

Howell Tong. *Non-linear Time Series: A Dynamical System Approach.* Oxford University Press, 1990.

R. S. Tsay. Testing and modeling threshold autoregressive processes. *Journal of the American Statistical Association*, 84(405):231–240, 1989.

Ruey S Tsay. Testing and modeling multivariate threshold models. *Journal of the American Statistical Association*, 93(443):1188–1202, 1998.

Jing Wu. Threshold garch model: Theory and application.

L. Wang X. Chen and Y. Zhao. Iot and ai applications in threshold-based models. *International Journal of Emerging Technologies*, 2023.

Yen-Chi Chen Yikun Zhang and Alexander Giessing. Nonparametric inference on dose-response curves without the positivity condition. *Journal of Medical Statistics*, 34:12–25, 2022.

Jean-Michel Zakoian. Threshold heteroskedastic models.