A 3-STATE MARKOV CHAIN APPROACH TO WIND SPEED ANALYSIS

Project Report
Submitted to
The Department of Statistics,
Pondicherry University

By
Ms. A S Praisie Jemimah
Regd. No. 24MSSTAPY0059
and
Ms. Y Sri Nikitha
Regd. No. 24MSSTAPY0032

As an Assignment for the Partial Fulfilment of the Course work
STAT 423 - STOCHASTIC PROCESSES

Under the Guidance of
Dr. P. Tirupathi Rao
Professor & Course Teacher

With the assistance of
Mr Muhammed Rais P K
Research Scholar

April - 2025

DR.TIRUPATHI RAO PADI
M.Sc., M.Phil., Ph.D., M.B.A., D.C.A., CP-SAS
PROFESSOR, Department of Statistics
Ramanujan School of Math. Sciences
PONDICHERRY UNIVERSITY
(A Central University)

R.V. Nagar, Kalapet,
Puducherry (UT) – 605014, India
Email: drtrpadi@pondiuni.ac.in;
traopadipu@gmail.com,
Phone: 9486492241 (Mobile),
9629862241 (WAN), 0413-2654-829(Office

Certificate

This is to certify that **Ms. A S PRAISIE JEMIMAH** and **Ms. Y SRI NIKITHA** with the assistance of **Mr. Muhammed Rais P K,** Research Scholar have completed their project work as partial fulfillment of the course work of Stat 423: Stochastic Processes submitted to the Department of Statistics, Pondicherry University in April 2025. They have carried out all the stages of project work right from the data collection to report making, on their own. The data that they have collected is from a real time context. No part of this work was carried out earlier in any format by any one for M.Sc./ MBA/ MTech/ etc. dissertation works or Ph.D. thesis.

Date: **(P. Tirupathi Rao)**

Table of contents

Abstract

This project applies a 3-state discrete-time Markov chain to model and predict wind speed dynamics in Bengaluru, Karnataka, using half-hourly wind data sourced from the National Centers for Environmental Information (NCEI). Wind speeds were categorized into three states —light, moderate, and strong—based on defined thresholds, enabling the formation of a state sequence for probabilistic modeling. Transition probability matrices were constructed and analyzed across time intervals, allowing for both short-term and long-term predictions through n-step and steady-state probabilities. Statistical properties such as mean, variance, skewness, and moment-generating functions were computed to characterize the distributional behavior of each wind state. Implemented entirely in R, the study demonstrates that Markov modeling is effective in capturing the stochastic nature of wind patterns, with the stationary distribution indicating that Bengaluru experiences light wind 71.18% of the time, moderate wind 28.45%, and strong wind only 0.37%—a result consistent with the region's climatology and valuable for renewable energy planning and environmental forecasting.

M O D U L E 1

Overview of the Stochastic Modeling Approach on Wind Speed Analysis

1.1 Introduction

In many real-world systems, uncertainty and randomness are inherent components that significantly influence outcomes. Traditional deterministic models, while effective in certain domains, often fall short in capturing the variability and unpredictability present in natural phenomena. Stochastic processes, by contrast, offer a powerful mathematical framework for modeling systems that evolve over time in a probabilistic manner.

A stochastic process is a collection of random variables indexed by time or space, where each random variable represents the state of a system at a particular point. These processes are especially useful when studying temporal phenomena where outcomes are not fixed, but rather governed by probabilistic transitions between different states. Examples span across disciplines —from stock market movements and weather forecasting to population dynamics and signal processing.

Among the various types of stochastic processes, Markov processes stand out for their simplicity and wide applicability. A Markov process is characterized by the memoryless property, meaning the future state depends only on the present state, not on the sequence of events that preceded it. This makes Markov models ideal for sequential data analysis where transitions from one state to another can be observed and quantified.

In this study, we apply the theory of stochastic processes—specifically, discrete-time Markov chains—to analyze and predict wind speed behavior over time. The data under consideration comprises wind speed measurements collected at regular intervals in Bengaluru. By categorizing the wind speeds into discrete states and observing the transitions between these states, we aim to construct a 3-state Markov model that captures the dynamics of wind variation.

Through this approach, the study not only illustrates the practical application of stochastic processes in environmental data modeling but also highlights the predictive potential of Markov chains in understanding and anticipating future states of wind behavior.

1.2 Problem Description and Background

Wind energy is a critical component of sustainable development and renewable energy infrastructure. However, the inherently variable and stochastic nature of wind speed poses significant challenges to its effective integration into power systems and forecasting models. Accurate prediction of wind speed is essential for energy generation planning, grid stability, and resource allocation, especially in urban environments such as Bengaluru, where energy demands are steadily rising.

To address this challenge, this project applies the theory of stochastic processes to develop a Markov model for analyzing wind speed behavior. The approach involves discretizing wind speed into three distinct states—typically representing low, moderate, and high speeds—and studying the probabilistic transitions between these states over time.

Using historical wind speed data sourced from the National Centers for Environmental Information (NCEI), this study builds a time-based model that captures the underlying dynamics of wind behavior in Bengaluru. The model aims to represent the system's transition structure through a 3-state Markov chain, where the state at a given time step depends only on the immediately preceding state. This Markov-based approach offers a simplified and mathematically tractable framework for modeling wind speed variability, enabling meaningful statistical inference, prediction of future states, and validation using observed data sequences.

1.3 Objectives of the Study

The main objectives of this study are:

- To model wind speed behavior in Bengaluru using a 3-state discrete-time Markov chain.
- To categorize wind speed data into defined states and analyze transition patterns.
- To compute the transition probability matrix and stationary distribution.
- To use the Markov model for predicting future wind states.
- To validate the model using real data and assess its accuracy.

1.4 Relevance of Markov Models in Statistical Modeling

Markov models are widely used in statistical modeling for systems that evolve over time with uncertainty. Their key strength lies in the memoryless property, where the next state depends only on the current state—not the full history.

In environmental studies like wind speed analysis, this makes Markov models especially useful. They simplify complex temporal behavior into a manageable structure of states and transitions. By estimating probabilities from observed data, these models offer both descriptive and predictive power, making them a reliable tool for analyzing real-world stochastic processes.

1.5 Scope and Applications in Real-World Contexts

This study focuses on modeling wind speed patterns in Bengaluru using a 3-state Markov process. The scope includes data analysis, transition modeling, prediction, and validation.

Markov models like this are widely applied in areas such as weather forecasting, energy management, climate modeling, and environmental planning. The approach used here can be extended to other cities, time-series data, or even different variables like temperature or air quality, making it useful for broader environmental and engineering applications.

1.6 Motivation for Selecting 3-State Markov Process

A 3-state Markov process offers a simple yet effective way to represent wind speed variations—categorized as low, moderate, and high. This level of classification balances detail with clarity, making analysis manageable while still capturing meaningful transitions. It also fits well with the discrete nature of half-hourly data and supports accurate prediction without overcomplicating the model.

1.7 Literature Review

Numerous studies have demonstrated the applicability of Markov models in environmental and time-series analysis, particularly in modeling wind speed data and forecasting renewable energy outputs. Bartlett (1950) laid an early statistical foundation for significance testing in multivariate systems, while Cochran (1952) established the χ^2 test of goodness of fit, which remains vital in validating stochastic models such as Markov chains. Zhou and Yang (2012) effectively utilized a discrete-time Markov chain model for wind power forecasting, highlighting the model's capacity to capture short-term fluctuations in wind behavior. Li and Shi (2010) extended this application by evaluating wind energy potential using Markov models in various Chinese regions. Bivona et al. (2003) similarly applied Markov chains for estimating wind farm energy production, emphasizing the need for careful state classification. Aksoy (2004) used first-order Markov chains to simulate daily wind speeds in the eastern Mediterranean, reinforcing the model's effectiveness for temporal wind analysis. Broader climatological applications are seen in the work of Tveito and Førland (1999), who used stochastic weather generators based on Markov processes to simulate precipitation and temperature, while Sørbye and Rue (2004) discussed the broader use of Markov random fields in environmental modeling. Collectively, these studies affirm the suitability of Markov models for characterizing and forecasting environmental variables. Building on this literature, the present study employs a 3-state discrete-time Markov model to analyze half-hourly wind speed data in Bengaluru, with an emphasis on model simplicity and practical predictive capability.

1.8 Glossary of Key Terms

- **Stochastic Process:** A collection of random variables representing a system evolving over time.
- **Markov Chain:** A type of stochastic process where the next state depends only on the current state.
- **Transition Probability:** The likelihood of moving from one state to another.
- **State Space:** A set of all possible states in the model (e.g., Light, Moderate, Strong wind speed).
- **Stationary Distribution:** A stable state distribution where probabilities do not change over time.
- **Wind Speed State:** Categorization of wind speed into defined levels for modeling.

1.9 Terminology

π_k : Initial probability for the kth state, $\pi_k \geq 0$; for all k=1,2,3; $\sum_{k=1}^{3} \pi_k$ =1; $\pi_k = \frac{n_k}{n}$; n= $\sum_{k=1}^{3} n_k$, i.e., Total number of observations considered for the study

P_{jk} : The transition probability between states j and k represents the likelihood of moving from state j to state k in a given system or process.
i.e. P{Zn=k/Zn-1=j} ≥0; 0≤ Pjk≤1 and $\sum_{k=1}^{3} P_{jk}$ =1 ∀ j= 1,2,3

j : Origin State

k : Destination State

y_t : Wind Speed(m/s) on t^{th} time

S : Strong Winds; S > 8.33

M : Moderate Winds; 4.44 ≤ M ≤ 8.33

L : Light Winds; L ≤ 4.44

Y (ω_1) : A random variable number of times the *Strong State* occurs in one time length sequence assuming [Y (ω_1) = y] = 0,1

Y (ω_2) : A random variable number of times the *Moderate State* occurs in one time length sequence assuming [Y (ω_2) = y] = 0,1

Y (ω_3) : A random variable number of times the *Light State* occurs in one time length sequence assuming [Y (ω_3) = y] = 0,1

Y (ω_4) : A random variable number of times the *Strong State* occurs in two time length sequence assuming [Y (ω_4) = y] = 0,1,2

Y (ω_5) : A random variable number of times the *Moderate State* occurs in two time length sequence assuming [Y (ω_5) = y] = 0,1,2

Y (ω_6) : A random variable number of times the *Light State* occurs in two time length sequence assuming [Y (ω_6) = y] = 0,1,2

Y (ω_7) : A random variable number of times the *Strong State* occurs in three time length sequence assuming [Y (ω_7) = y] = 0,1,2,3

Y (ω_8) : A random variable number of times the *Moderate State* occurs in three time length sequence assuming [Y (ω_8) = y] = 0,1,2,3

Y (ω_9) : A random variable number of times the *Light State* occurs in three time length sequence assuming [Y (ω_9) = y] = 0,1,2,3

M O D U L E 2

Stochastic Modeling for Wind Speed data

2.1 Description of the Dataset

The dataset used in this study consists of half-hourly wind speed measurements collected from Bengaluru. It spans a full year and was sourced from the National Centers for Environmental Information (NCEI) official database.

Each entry includes:

- Date and Time
- Wind Speed (in m/s)
- Wind State (categorized)
- Transition (previous to current state)

The data is well-structured for time-series analysis and suitable for building a 3-state Markov model.

2.2 Data Collection Sources and Methods

The wind speed data was obtained from the official website of the National Centers for Environmental Information (NCEI). The dataset covers Bengaluru and includes half-hourly recordings over a one-year period.

Data was downloaded in CSV format and includes precise timestamps and wind speed values. No manual data entry was involved, ensuring high reliability and consistency for analysis.

2.3 Data Structure and Variables Used

The dataset is organized in tabular format with the following key variables:

- **Date/Time:** Timestamp of wind speed recording (30-minute intervals)
- **Wind Speed (m/s):** Measured wind speed in meters per second
- **State:** Wind speed categorized into Low, Moderate, or High
- **Previous State:** Wind state from the prior time step
- **Transition:** Movement from previous state to current state

These variables were used to construct state sequences and calculate transition probabilities for the Markov model.

2.4 Data Preparation Techniques

The dataset was cleaned by removing missing values and ensuring chronological order. Wind speed values were then categorized into three discrete states based on standard classification:

- Light: Less than 4.44 m/s
- Moderate: 4.44 to 8.33 m/s
- Strong: Greater or equal than 8.8 m/s

Each measurement was assigned to a state, and transitions between consecutive states were recorded to form a sequence for Markov modeling.

To enable Markov modeling, continuous wind speed values were categorized into three discrete states: Light (< 4.44 m/s), Moderate (4.44–8.33 m/s), and Strong (> 8.8 m/s).
These thresholds were inspired by the Beaufort Wind Force Scale published by the U.S. National Weather Service, which classifies wind into multiple detailed levels based on speed and effects. However, for modeling simplicity and interpretability, the full scale was consolidated into three broad categories suitable for stochastic state analysis.

2.5 Transformation into State Sequences

After categorizing each wind speed value into Light, Moderate, or Strong, a time-ordered sequence of states was created. This sequence represents how wind conditions evolved over time. Each entry in the sequence shows the current state and the one immediately before it, forming a basis for analyzing state-to-state transitions. This transformation is essential for constructing the Markov chain and estimating transition probabilities.

2.6 Specimen Data (First and Last Five Rows)

To provide a snapshot of the dataset, the first and last five rows (after preprocessing and state classification) are shown below:

	A	B	C	D
1	DATE	wnd_speed	wnd_state	transition
2	2024-01-0	4.1	Light	LL
3	2024-01-0	3.1	Light	LL
4	2024-01-0	3.1	Light	LL
5	2024-01-0	3.1	Light	LL
6	2024-01-0	3.1	Light	LL
15623	2024-12-3	4.1	Light	LL
15624	2024-12-3	3.6	Light	LL
15625	2024-12-3	2.6	Light	LL
15626	2024-12-3	2.6	Light	LL
15627	2024-12-3	1.5	Light	LL
15628	2024-12-3	1.5	Light	

2.7 Formation of State Transition Sequences

Using the classified states, a complete sequence of transitions was generated by comparing each state with its immediate predecessor. These transitions were recorded in the format:

$State_1 \rightarrow State_2$

For example:

- Light → Moderate
- Moderate → Strong
- Strong → Light

This full transition sequence was then used to count the frequency of each possible transition, which is essential for building the state transition matrix in the Markov model.

2.8 Excel & R Data Formatting

The wind speed data was first cleaned and categorized in Excel, where states and transitions were added as new columns. This structured format was then exported as a CSV file for further processing in R.

In R, the data was imported using read.csv(), and the transition sequence was extracted to generate the transition matrix using functions like table() and matrix(). This prepared the dataset for statistical analysis and visualization using Markov modeling techniques.

2.9 Exploratory Data Analysis

2.9.1 Data Summary

- Rows: 15,627
- Columns: 4
  - DATE (string): Timestamps at 30-minute intervals.
  - wnd_speed (float): Wind speed values.
  - wnd_state (string): Wind speed states (like "Light", "Medium", etc.).
  - transition (string): Markov state transitions (e.g., "LL", "LM", etc.).

2.9.2 Wind Speed Distribution

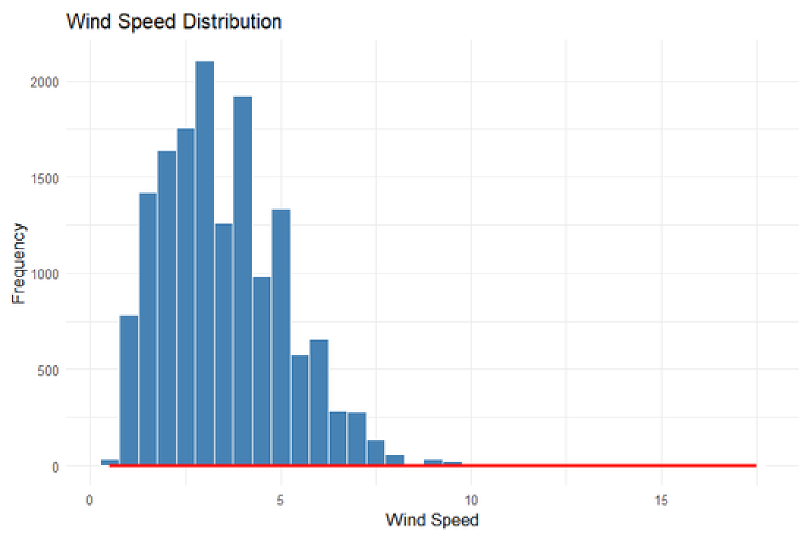


Fig.2.1

2.9.2.1 Distribution Shape

- The histogram shows a right-skewed (positively skewed) distribution — the bulk of the wind speed values are concentrated on the left (lower speeds).
- The frequency decreases as wind speed increases, especially after 6 m/s.

2.9.2.2 Peak Region

- The mode (most frequent wind speeds) lies roughly between 2.5 to 3.5 m/s, peaking at over 2000 occurrences.
- This indicates that moderate winds dominate in your dataset.

2.9.2.3 Tail Behavior

- Wind speeds greater than 7 m/s are much less frequent.
- Very few occurrences exist beyond 10 m/s, forming a long, thin tail, possibly representing extreme or storm conditions.

2.9.2.5. Implications for Wind Modeling

- The distribution is not symmetrical, which means:
- Mean > Median (because of the right skew).
- Standard deviation may be inflated slightly by the high-speed outliers.

2.9.3 Boxplot of Wind Speed

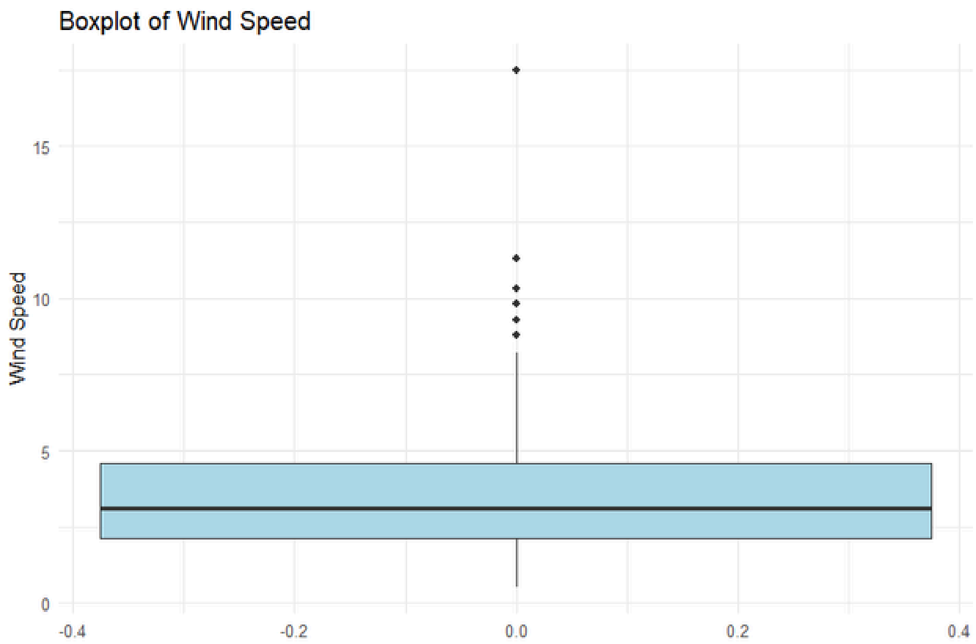


Fig 2.2

2.9.3.1 Central Tendency

- The median (black line inside the box) lies slightly below 3.5 m/s.
- This confirms the histogram's finding: moderate winds are most common.

2.9.3.2 Spread

- The interquartile range (IQR) spans roughly from 2 m/s (Q1) to 5 m/s (Q3).
- This tells us that 50% of wind speed values lie within this range — again, supporting the idea that most winds are mild to moderate.

2.9.3.3 Whiskers and Range

- The upper whisker reaches about 7–8 m/s, and the lower whisker goes nearly to 0.
- These whiskers represent the typical range excluding outliers (within 1.5×IQR from Q1 and Q3).

2.9.3.4 Outliers

- Several black dots above the upper whisker are outliers, with wind speeds above 8 m/s, extending even beyond 17 m/s.
- These are uncommon high-wind events, likely associated with storms or turbulent atmospheric conditions.
- Their presence skews the distribution to the right (as we saw in the histogram).

2.9.3.5 Symmetry

- The box is slightly skewed toward the lower end, and the whisker is longer on the upper side.
- This visual cue confirms the positive skewness in the wind speed data.

2.9.4 Wind State Frequency

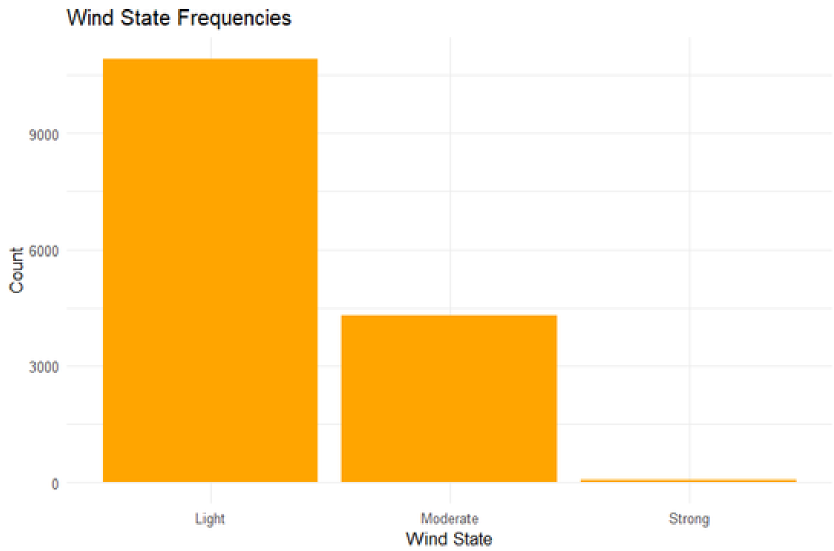


Fig 2.3

This bar chart displays the number of observations classified into three wind speed states: Light, Moderate, and Strong.

🔹 **Light Wind:**

- Most frequent category.

- Makes up the majority of the data, suggesting that the region generally experiences calm to gentle breezes.
- Indicates a dominant wind condition for the observed period.

◆ **Moderate Wind:**

- The second most common state.
- Still significantly present, indicating occasional higher wind speeds, possibly during afternoons or seasonal changes (e.g., monsoon transitions).

◆ **Strong Wind:**

- Very infrequent.
- Could indicate rare weather events, such as storms or very gusty days.
- Suggests that high wind scenarios are not typical for this location/time period

2.9.5 Transition Frequencies

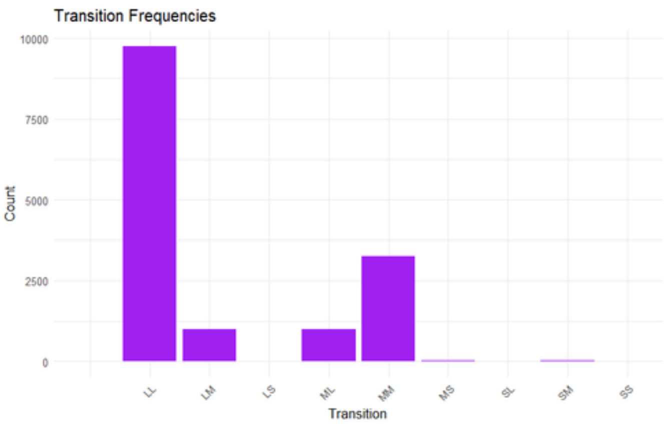


Fig 2.4

This bar chart displays the frequency of state-to-state transitions between wind speed categories: Light (L), Moderate (M), and Strong (S). Each bar represents how often one wind state transitioned to another.

- **LL (Light → Light):**
  - Most frequent transition.
  - Indicates high persistence of light wind states.
  - Reflects stability in calm wind conditions.
- **MM (Moderate → Moderate):**
  - Second highest frequency.
  - Suggests moderate wind speeds also tend to persist, though less than light winds.
- **LM (Light → Moderate) and ML (Moderate → Light):**
  - Moderate frequency.
  - Suggest fluctuations between light and moderate states, possibly driven by time-of-day effects (e.g., afternoon wind increases).
- **Other transitions (LS, MS, SL, SM, SS):**
  - Very low or negligible frequency.
  - Indicates rare shifts to or from the Strong state, confirming that strong wind events are infrequent and unstable.

2.9.6 Wind Speed States

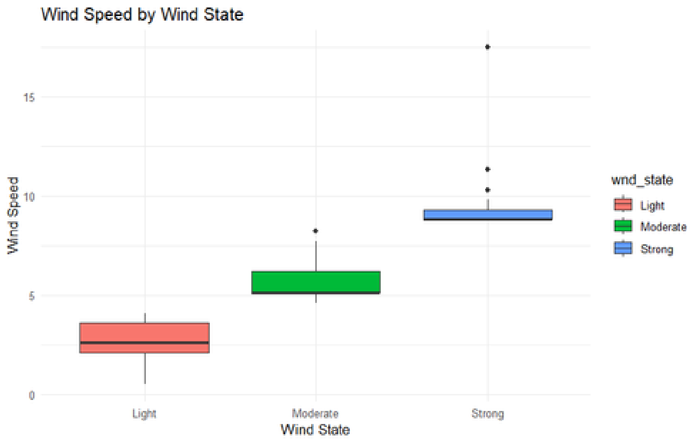


Fig 2.4

This boxplot visualizes wind speed distributions categorized by wind state (Light, Moderate, Strong).

1. Wind State Categories

Light (red): Wind speeds mostly fall between ~1 and 4 m/s.

Moderate (green): Wind speeds mostly range from ~5 to 7 m/s.

Strong (blue): Wind speeds are clustered between ~8 and 10 m/s, with a few higher outliers.

2. Median Wind Speed

Light: ~2.5 m/s

Moderate: ~6 m/s

Strong: ~9 m/s

Each wind state has a higher median wind speed than the previous, suggesting a consistent increase in wind speed across states.

3. Spread of Data

Light has a narrower interquartile range (IQR) compared to Moderate and Strong.

Strong shows more variability, indicated by a wider IQR and several outliers reaching up to ~17 m/s.

4. Outliers

Moderate has an outlier above 8 m/s.

Strong has multiple outliers, with one as high as ~17 m/s, indicating occasional high wind gusts.

M O D U L E 3

Markov Modeling and Analysis

3.1 Definition and Structure of a Markov Process

A Markov process is a type of stochastic process where the future state depends only on the current state, not on the sequence of events that preceded it. This is known as the Markov property or memoryless property.

In mathematical terms:

$P(X_{n+1} \mid X_n, X_{n-1}, \ldots, X_0) = P(X_{n+1} \mid X_n)$

For this project, the Markov process is used to model wind speed transitions across three defined states: Light, Moderate, and Strong. The process is discrete-time and finite-state, making it suitable for a Markov chain model.

3.2 Explanation of 3-State Markov Chain

In this study, a 3-state Markov chain is constructed to model the transitions between different levels of wind speed:

- State 1: Light (0 to 4.44 m/s)
- State 2: Moderate (4.44 to 8.33 m/s)
- State 3: Strong (greater than 8.33 m/s)

Each observation is treated as a state, and transitions are tracked from one 30-minute interval to the next. The model assumes that the probability of moving to a future state depends only on the current state, not the past ones.

This 3-state system allows for estimating:

- The transition probabilities between states
- The long-term behavior of wind speed
- Future predictions using the transition matrix

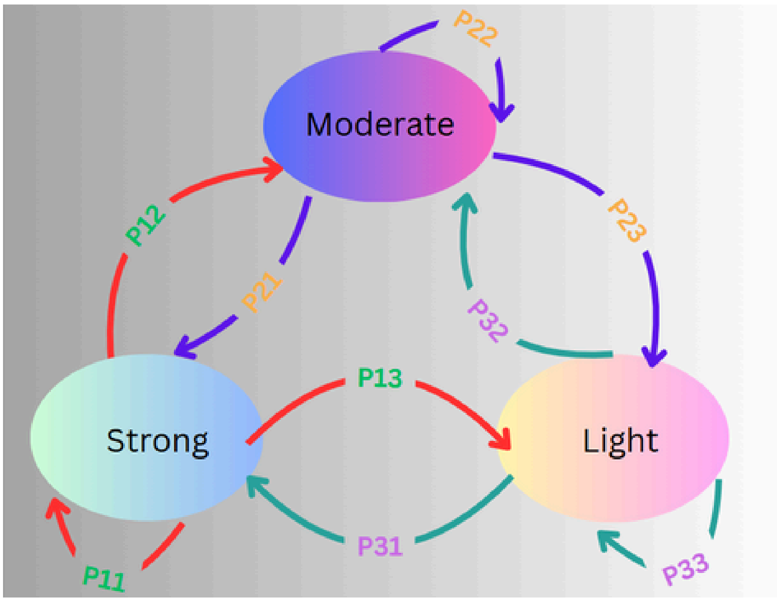


Fig 3.1

Figure 3.1: Schematic diagram of the 3-state Markov chain used in this study. Each node represents a wind speed state—Light, Moderate, and Strong—while arrows show possible transitions between states with associated probabilities.

3.3 Construction of State Transition Matrices

To build the state transition matrix, all consecutive state pairs were examined from the time-ordered sequence. The number of transitions from each state to every other state (including self-transitions) was counted, resulting in a 3×3 frequency table.

For this study, the states are defined as follows:

- State 1: Strong (> 8.8 m/s)
- State 2: Moderate (4.44 to 8.33 m/s)
- State 3: Light (< 4.44 m/s)

After tallying the frequencies, each row was normalized (dividing by the row total) to convert the frequency table into a probability matrix. The general format of the transition matrix P is:

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$$

Where:

- Pij represents the probability of transitioning from state i (current) to state j (next).
- With i,j=1 (Strong), 2 (Moderate), 3 (Light).

This matrix serves as the foundation for all subsequent statistical analyses and predictions in the study.

3.4 Construction of Transition Probabilities

The transition probabilities were derived from the wind speed state sequence using the following steps:

1. Categorized Data: Each 30-minute wind speed value was labeled as Strong (1), Moderate (2), or Light (3) based on predefined thresholds.
2. State Pairing: Consecutive state values were paired (e.g., from time t to t+1) to track all possible transitions like 1→1, 1→2, ..., 3→3.
3. Frequency Matrix: A transition count matrix was constructed by counting how many times each transition occurred across the dataset.
4. Normalization: Each row of the count matrix was divided by the sum of that row to obtain the transition probability matrix (P):

$$P_{ij} = \frac{\text{Number of transitions from state } i \text{ to state } j}{\text{Total transitions from state } i}$$

This process yielded a Markov transition matrix that quantifies how likely the wind speed is to remain in or change between different states over time.

3.5 Identification and Representation of Three Sequences

In this study, a single sequence of wind speed states was used to compute the first-order Markov transition matrix. To explore the behavior of the system over time, the model was extended to analyze state transitions across multiple time steps:

- **One-step transition:** Direct transition probabilities from one state to the next
- **Two-step transition:** Probabilities of being in a state two steps ahead
- **Three-step transition:** Probabilities across three time intervals

This was done by raising the one-step transition matrix (P) to higher powers:

$$\mathbf{P^{(n)} = P^n}$$

Each power of the matrix reflects how the wind speed state probabilities evolve over multiple intervals, helping to assess:

- Long-term trends
- Convergence to a steady pattern
- Change dynamics over time

This multi-step approach allows deeper insight into the Markov process behavior without splitting the original dataset.

3.6 Stationary (Limiting) Distributions

A stationary distribution of a Markov chain represents the long-term behavior of the system, where the probabilities of being in each state stabilize and remain constant over time.

For the transition matrix P, the stationary distribution vector $\pi=[\pi_1,\pi_{2,}\pi_3]$ satisfies:

$$\pi P=\pi \text{ and } \pi_1 + \pi_2 + \pi_{3=} 1$$

- π_1: Long-term probability of being in State 1 (Strong)
- π_2: Long-term probability of being in State 2 (Moderate)
- π_3: Long-term probability of being in State 3 (Light)

The stationary distribution was computed using R by solving the above system. It provides a useful summary of wind behavior in the long run, independent of the starting state.

3.7 Computation of Expected Probabilities per State

The expected return (or occupancy probability) for each state reflects how often the system is likely to be in that state over time. These are directly linked to the stationary distribution π\piπ.

For a given state iii, the expected long-run proportion of time spent in that state is:

$$\text{Expected Probability of State i}=\pi_i$$

Where:

- π_1 : Expected return to Strong wind (State 1)
- π_2 : Expected return to Moderate wind (State 2)
- π_3 : Expected return to Light wind (State 3)

These values provide an intuitive understanding of which wind speed category dominates over time, and how frequently the system revisits each state.

3.8 Analysis of State-wise Distributions Across Time Sequences

This section presents a detailed statistical characterization of the three Markovian wind speed states—Strong, Moderate, and Light—across three temporal sequences derived from the dataset. Each state's behavior is examined using distributional measures such as mean, variance, third moment, skewness, and moment generating function (MGF), offering insights into the underlying stochastic dynamics.

3.8.1 Sequence 1: Distributional Analysis of Wind Speed States

For Strong,

$$P_3(S) = \mathrm{P}\,(X_{(1)} = \mathrm{S}, X_{(0)} = \mathrm{S}) + \mathrm{P}\,(X_{(1)} = \mathrm{S}, X_{(0)} = \mathrm{M}) + \mathrm{P}\,(X_{(1)} = \mathrm{S}, X_{(0)} = \mathrm{L})$$
$$= \pi_1 p_{11} + \pi_2 p_{21} + \pi_3 p_{31}$$

$$P_3(S) = \sum_{k=1}^{3} \pi_k \, p_{k1} \tag{3.1}$$

Similarly,

For Moderate,

$$P_3(M) = \sum_{k=1}^{3} \pi_k \, p_{k2} \tag{3.2}$$

For Light,

$$P_3(L) = \sum_{k=1}^{3} \pi_k \, p_{k3} \tag{3.3}$$

3.8.1.1 Probability Mass Function of Strong State

Let us consider a random variable denoted by $Y\,(\omega_1) = y$ which represents the occurrence of the Strong State in a one sequence. This variable can assume values 0 and 1, where '0' signifies its absence of the Strong State and '1' signifies its presence of the Strong State.

$$\mathrm{P}\,[\mathrm{Y}(\omega_1) = \mathrm{y}] = \begin{cases} \sum_{k=1}^{3} \sum_{j=2}^{3} \pi_k p_{kj} & ; for\ y = 0 \\ \sum_{k=2}^{3} \pi_k p_{l:j} & ; for\ y = 1 \\ 0 \quad ; otherwise\ (y \geq 2) \end{cases} \tag{3.4}$$

3.8.1.2 Statistical Measures for *Strong State*

3.8.1.2.1 The Average Occurrence of a *Strong State*

$$\mu(S) = \sum_{k=1}^{3} \pi_k\, p_{k1} \tag{3.5}$$

3.8.1.2.2 The Variance of a *Strong State*

$$\sigma^2(S) = (\mu(S))^2 \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + \left(1-\mu(S)\right)^2 \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right) \tag{3.6}$$

3.8.1.2.3 The Third Central Moment for *Strong State*

$$\mu_3(S) = (-\mu(S))^3 \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + \left(1-\mu(S)\right)^3 \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right) \tag{3.7}$$

3.8.1.2.3 The Coefficient of Skewness for *Strong State*

$$\beta_1(S) = \left[(-\mu(S))^3 \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + \left(1-\mu(S)\right)^3 \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right)\right]^2 \left[(\mu(S))^2 \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + \left(1-\mu(S)\right)^2 \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right)\right]^{-3} \tag{3.8}$$

3.8.1.2.4 The Coefficient Kurtosis for *Strong State*

$$\beta_2(S) = \left[(\mu(S))^4 \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + \left(1-\mu(S)\right)^4 \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right)\right] \left[(\mu(S))^2 \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + \left(1-\mu(S)\right)^2 \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right)\right]^{-2} \tag{3.9}$$

3.8.1.2.5 The Coefficient of Variation for Strong *State*

$$CV(S) = \left[(\mu(S))^2 \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + \left(1-\mu(S)\right)^2 \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right)\right]^{\frac{1}{2}} \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right)^{-1} \tag{3.10}$$

3.8.1.2.6 Moment Generating Function (MGF) for *Strong State*

$$M(t) = \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + e^{t} \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right) \tag{3.11}$$

3.8.1.2.7 Characteristic Function (CF) for *Strong State*

$$\phi(t) = \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + e^{it} \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right) \tag{3.12}$$

3.8.1.2.8 Probability Generating Function (PGF) for *Strong State*

$$P(t) = \left(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\right) + s \left(\sum_{k=1}^{3} \pi_k\, p_{k1}\right) \tag{3.13}$$

3.8.1.3 Probability Mass Function of Moderate State

Let us consider a random variable denoted by $Y\left(\omega_2\right) = y$ which represents the occurrence of the Light State in a one sequence. This variable can assume values 0 and 1, where '0' signifies its absence of the Light State and '1' signifies its presence of the Light State.

$$P\left[Y(\omega_2) = y\right] = \begin{cases} \sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} & ; for\ y = 0 \\ \sum_{k=1}^{3} \pi_k p_{k2} & ; for\ y = 1 \\ 0 & ; otherwise\ (y \geq 2) \end{cases} \tag{3.14}$$

3.8.1.4 Statistical Measures for *Moderate State*

3.8.1.4.1 The Average Occurrence of a *Moderate State*

$$\mu(M) = \sum_{k=1}^{3} \pi_k \, p_{k2} \tag{3.15}$$

3.8.1.4.2 The Variance of a *State*

$$\sigma^2(M) = (\mu(M))^2 \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + \left(1 - \mu(M)\right)^2 \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \tag{3.16}$$

3.8.1.4.3 The Third Central Moment for *Moderate State*

$$\mu_3(M) = (-\mu(M))^3 \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + \left(1 - \mu(M)\right)^3 \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \tag{3.17}$$

3.8.1.4.4 The Coefficient of Skewness for *Moderate State*

$$\beta_1(M) = \left[(-\mu(M))^3 \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + \left(1 - \mu(M)\right)^3 \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \right]^2 \left[(\mu(M))^2 \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + \left(1 - \mu(M)\right)^2 \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \right]^{-3} \tag{3.18}$$

3.8.1.4.4 The Coefficient Kurtosis for *Moderate State*

$$\beta_2(M) = \left[(\mu(M))^4 \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + \left(1 - \mu(M)\right)^4 \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \right] \left[(\mu(M))^2 \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + \left(1 - \mu(M)\right)^2 \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \right]^{-2} \tag{3.19}$$

3.8.1.4.5 The Coefficient of Variation for *Moderate State*

$$CV(M) = \left[(\mu(M))^2 \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + \left(1 - \mu(M)\right)^2 \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \right]^{\frac{1}{2}} \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right)^{-1} \tag{3.20}$$

3.8.1.4.6 Moment Generating Function (MGF) for *Moderate State*

$$M(t) = \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + e^{t} \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \tag{3.21}$$

3.8.1.4.7 Characteristic Function (CF) for *Moderate State*

$$\phi(t) = \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + e^{it} \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \tag{3.22}$$

3.8.1.4.8 Probability Generating Function (PGF) for *Moderate State*

$$P(t) = \left(\sum_{k=1}^{3} \sum_{\substack{j=1 \\ j\neq 2}}^{3} \pi_k p_{kj} \right) + s \left(\sum_{k=1}^{3} \pi_k \, p_{k2}\right) \tag{3.23}$$

3.8.1.5 Probability Mass Function of Light State

Let us consider a random variable denoted by $Y\,(\omega_3) = y$ which represents the occurrence of the Light State in a one sequence. This variable can assume values 0 and 1, where '0' signifies its absence of the Light State 19 and '1' signifies its presence of the Light State.

$$P\,[Y(\omega_3) = y] = \begin{cases} \sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj} & ;for\ y = 0 \\ \sum_{k=1}^{3}\pi_k p_{k3} & ;for\ y = 1 \\ 0 \quad ;otherwise\ (y \geq 2) \end{cases} \tag{3.24}$$

3.8.1.6 Statistical Measures for *Light State*

3.8.1.6.1 The Average Occurrence of a *Light State*

$$\mu(L) = \sum_{k=1}^{3}\pi_k\, p_{k3} \tag{3.25}$$

3.8.1.6.2 The Variance of a *Light State*

$$\sigma^2(L) = (\mu(L))^2\left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + \left(1-\mu(L)\right)^2\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right) \tag{3.26}$$

3.8.1.6.3 The Third Central Moment for *Light State*

$$\mu_3(L) = (-\mu(L))^3\left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + \left(1-\mu(L)\right)^3\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right) \tag{3.27}$$

3.8.1.6.4 The Coefficient of Skewness for *Light State*

$$\beta_1(L) = \left[(-\mu(L))^3\left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + \left(1-\mu(L)\right)^3\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right)\right]^2 \left[(\mu(L))^2\left(\sum_{k=1}^{3}\sum_{j=1}^{3}\pi_k p_{kj}\right) + \left(1-\mu(L)\right)^2\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right)\right]^{-3} \tag{3.28}$$

3.8.1.6.4 The Coefficient Kurtosis for *Light State*

$$\beta_2(L) = \left[(\mu(L))^4\left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + \left(1-\mu(L)\right)^4\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right)\right] \left[(\mu(L))^2\left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + \left(1-\mu(L)\right)^2\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right)\right]^{-2} \tag{3.29}$$

3.8.1.6.5 The Coefficient of Variation for *Light State*

$$CV(L) = \left[(\mu(L))^2\left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + \left(1-\mu(M)\right)^2\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right)\right]^{\frac{1}{2}}\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right)^{-1} \tag{3.30}$$

3.8.1.6.6 Moment Generating Function (MGF) for *Light State*

$$M(t) = \left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + e^t\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right) \tag{3.31}$$

3.8.1.6.7 Characteristic Function (CF) for *Light State*

$$\phi(t) = \left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + e^{it}\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right) \tag{3.32}$$

3.8.1.6.8 Probability Generating Function (PGF) for *Light State*

$$P(t) = \left(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\right) + s\left(\sum_{k=1}^{3}\pi_k\, p_{k3}\right) \tag{3.33}$$

3.8.2 Sequence 2: Distributional Analysis of Wind Speed States

$$P_3(S,S) = \mathrm{P}\,(X_{(2)} = \mathrm{S},\ X_{(1)} = \mathrm{S}, X_{(0)} = \mathrm{S}) + \mathrm{P}\,(X_{(2)} = \mathrm{S},\ X_{(1)} = \mathrm{S}, X_{(0)} = \mathrm{M}) + \mathrm{P}\,(X_{(2)} = \mathrm{S},\ X_{(1)} = \mathrm{S}, X_{(0)} = \mathrm{L})$$
$$= \pi_1 p_{11} p_{11} + \pi_2 p_{21} p_{11} + \pi_3 p_{31} p_{11}$$

$$P_3(S,S) = p_{11} P_3(S) \tag{3.34}$$

Similarly,

$$\begin{aligned}
&P_3(S,M) = p_{21} P_3(M)\\
&P_3(S,L) = p_{31} P_3(L)\\
&P_3(M,M) = p_{22} P_3(M)\\
&P_3(M,S) = p_{12} P_3(S)\\
&P_3(M,L) = p_{32} P_3(L)\\
&P_3(L,L) = p_{33} P_3(L)\\
&P_3(L,S) = p_{13} P_3(S)\\
&P_3(L,M) = p_{23} P_3(M)
\end{aligned} \tag{3.35}$$

3.8.2.1 Probability Mass Function of Strong State

Let us consider a random variable denoted by $Y\,(\omega_4) = y$ which represents the occurrence of the Strong State in a two sequences. This variable can assume values 0 and 1 and 2, where '0' signifies its absence of the Strong State and '1' signifies its presence of the Strong State once in 2 sequences and '2' signifies its presence of the Strong State twice in 2 sequences.

$$\mathrm{P}\,[\mathrm{Y}(\omega_4) = \mathrm{y}] = \begin{cases} P_3(M)\ \sum_{i=2}^{3} p_{2i}\ +\ P_3(L)\ \sum_{i=2}^{3} p_{3i} & ; for\ y = 0 \\ P_3(S) \sum_{i=2}^{3} p_{1i}\ +\ p_{21}\, P_3(M)\ +\ p_{31}\, P_3(L) & ; for\ \mathrm{y} = 1 \\ p_{11}\, P_3(S) & ;\ \text{for y} = 2 \\ 0 & ; otherwise (y \geq 3) \end{cases} \tag{3.36}$$

By considering,

$$\alpha_{11}(M) = P_3(M)\ \sum_{i=2}^{3} p_{2i}\ +\ P_3(L)\ \sum_{i=2}^{3} p_{3i}$$

$$\alpha_{12}(M) = P_3(S) \sum_{i=2}^{3} p_{1i}\ +\ p_{21}\, P_3(M)\ +\ p_{31}\, P_3(L)$$

$$\alpha_{13}(M) = p_{11}\, P_3(S)$$

3.8.2.2 Statistical Measures for *Strong*

3.8.2.2.1 The Average Occurrence of a *Strong State*

$$\mu(S) = \alpha_{12}(S) + \alpha_{13}(S) \tag{3.37}$$

3.8.2.2.2 The Variance of *Strong State*

$$\sigma^2(S) = (\mu(S))^2\alpha_{11}(S) + (\mathbf{1} - \mu(S))^2\alpha_{12}(S) + (\mathbf{2} - \mu(S))^2\alpha_{13}(S) \tag{3.38}$$

3.8.2.2.3 The Third Central Moment for *Strong State*

$$\mu_3(S) = (-\mu(S))^3\alpha_{11}(S) + (\mathbf{1} - \mu(S))^3\alpha_{12}(S) + (\mathbf{2} - \mu(S))^3\alpha_{13}(S) \tag{3.39}$$

3.8.2.2.4 The Coefficient of Skewness for *Strong State*

$$\beta_1(S) = [(\mu(S))^2\alpha_{11}(S) + (\mathbf{1} - \mu(S))^2\alpha_{12}(S) + (\mathbf{2} - \mu(S))^2\alpha_{13}(S)]^2 \\ [(-\mu(S))^3\alpha_{11}(S) + (\mathbf{1} - \mu(S))^3\alpha_{12}(S) + (\mathbf{2} - \mu(S))^3\alpha_{13}(S)]^{-3} \tag{3.40}$$

3.8.2.2.5 The Coefficient of Kurtosis for *Strong State*

$$\beta_1(S) = [(\mu(S))^4\alpha_{11}(S) + (\mathbf{1} - \mu(S))^2\alpha_{12}(S) + (\mathbf{2} - \mu(S))^2\alpha_{13}(S)]^2 \\ [(-\mu(S))^3\alpha_{11}(S) + (\mathbf{1} - \mu(S))^3\alpha_{12}(S) + (\mathbf{2} - \mu(S))^3\alpha_{13}(S)]^{-2} \tag{3.41}$$

3.8.2.2.6 The Coefficient of Variation for *Strong State*

$$\mathrm{CV(S)} = [(\mu(S))^2\alpha_{11}(S) + (\mathbf{1} - \mu(S))^2\alpha_{12}(S) + (\mathbf{2} - \mu(S))^2\alpha_{13}(S)]^{1/2} (\alpha_{12}(S) + 2\alpha_{13}(S))^{-1}\% \tag{3.41}$$

3.8.2.2.7 Moment Generating Function (MGF) for *Strong State*

$$\mathrm{M(t)} = \alpha_{11}(S) + e^t(\alpha_{12}(S) + e^t\,\alpha_{13}(S)) \tag{3.42}$$

3.8.2.2.8 Characteristic Function (CF) for *Strong State*

$$\phi(\mathrm{t}) = \alpha_{11}(S) + e^{it}(\alpha_{12}(S) + e^{it}\,\alpha_{13}(S)) \tag{3.43}$$

3.8.2.2.9 Probability Generating Function (PGF) for *Strong State*

$$\mathrm{P(t)} = \alpha_{11}(S) + s(\alpha_{12}(S) + s\,\alpha_{13}(S)) \tag{3.44}$$

3.8.2.3 Probability Mass Function of Moderate State

Let us consider a random variable denoted by $Y(\omega_5) = y$ which represents the occurrence of the Moderate State in a two sequences. This variable can assume values 0 and 1 and 2, where '0' signifies its absence of the Moderate State and '1' signifies its presence of the Moderate State once in 2 sequences and '2' signifies its presence of the Moderate State twice in 2 sequences.

$$\mathrm{P}\,[\mathrm{Y}(\omega_5) = \mathrm{y}] = \begin{cases} P_3(S) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{1i} + P_3(L) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{3i} & ; for\ y = 0 \\ P_3(M) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{2i} + p_{12}\,P_3(S) + p_{32}\,P_3(L) & ; for\ y = 1 \\ p_{22}\,P_3(M) & ;\ \text{for y} = 2 \\ 0 & ; otherwise (y \geq 3) \end{cases} \tag{3.45}$$

By considering,

$$\alpha_{11}(M) = P_3(M) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{1i} \; + \; P_3(L) \; \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{3i}$$

$$\alpha_{12}(M) = P_3(M) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{2i} \; + \; p_{12}\, P_3(S) \; + \; p_{32}\, P_3(L) \tag{3.46}$$

$$\alpha_{13}(M) = p_{22}\, P_3(M)$$

3.8.2.4 Statistical Measures for *Moderate*

3.8.2.4.1 The Average Occurrence of a *Moderate State*

$$\mu(M) = \alpha_{12}(M) + \alpha_{13}(M) \tag{3.47}$$

3.8.2.4.2 The Variance of *Moderate State*

$$\sigma^2(M) = (\mu(S))^2 \alpha_{11}(M) + (\mathbf{1} - \mu(M))^2 \alpha_{12}(M) + (\mathbf{2} - \mu(M))^2 \alpha_{13}(M) \tag{3.48}$$

3.8.2.4.3 The Third Central Moment for *Moderate State*

$$\mu_3(M) = (-\mu(M))^3 \alpha_{11}(M) + (\mathbf{1} - \mu(M))^3 \alpha_{12}(M) + (\mathbf{2} - \mu(M))^3 \alpha_{13}(M) \tag{3.49}$$

3.8.2.4.4 The Coefficient of Skewness for *Moderate State*

$$\beta_1(M) = [(\mu(M))^2 \alpha_{11}(M) \; + \; (\mathbf{1} - \mu(M))^2 \alpha_{12}(M) \; + \; (\mathbf{2} - \mu(M))^2 \alpha_{13}(M)]^2 \\ [(-\mu(M))^3 \alpha_{11}(M) \; + \; (\mathbf{1} - \mu(M))^3 \alpha_{12}(M) \; + \; (\mathbf{2} - \mu(M))^3 \alpha_{13}(M)]^{-3} \tag{3.49}$$

3.8.2.4.5 The Coefficient of Kurtosis for *Moderate State*

$$\beta_1(M) = [(\mu(M))^4 \alpha_{11}(M) \; + \; (\mathbf{1} - \mu(M))^2 \alpha_{12}(M) \; + \; (\mathbf{2} - \mu(M))^2 \alpha_{13}(M)]^2 \\ \left[(-\mu(M))^3 \alpha_{11}(M) \; + \; (\mathbf{1} - \mu(M))^3 \alpha_{12}(M) \; + \; (\mathbf{2} - \mu(M))^3 \alpha_{13}(M)\right]^{-2} \tag{3.50}$$

3.8.2.4.6 The Coefficient of Variation for *Moderate State*

$$\mathrm{CV(M)} = [(\mu(M))^2 \alpha_{11}(M) \; + \; (\mathbf{1} - \mu(M))^2 \alpha_{12}(M) \; + \; (\mathbf{2} - \mu(M))^2 \alpha_{13}(M)]^{1/2} \, (\alpha_{12}(M) \; + \; 2\alpha_{13}(M))^{-1}\% \tag{3.51}$$

3.8.2.4.7 Moment Generating Function (MGF) for *Moderate State*

$$\mathrm{M(t)} = \alpha_{11}(M) + e^t(\, \alpha_{12}(M) + e^t\, \alpha_{13}(M)) \tag{3.52}$$

3.8.2.4.8 Characteristic Function (CF) for *Moderate State*

$$\phi(\mathrm{t}) = \alpha_{11}(M) + e^{it}(\, \alpha_{12}(M) + e^{it}\, \alpha_{13}(M)) \tag{3.53}$$

3.8.2.4.9 Probability Generating Function (PGF) for *Moderate State*

$$\mathrm{P(t)} = \alpha_{11}(M) + s(\, \alpha_{12}(M) + s\, \alpha_{13}(M)) \tag{3.54}$$

3.8.2.3 Probability Mass Function of Light State

Let us consider a random variable denoted by $Y(\omega_6) = y$ which represents the occurrence of the Light State in a two sequences. This variable can assume values 0 and 1 and 2, where ‘0’ signifies its absence of the Light State and ‘1’ signifies its presence of the Light State once in 2 sequences and ‘2’ signifies its presence of the Light State twice in 2 sequences.

$$P[Y(\omega_6) = y] = \begin{cases} P_3(S)\sum_{i=1}^{2} p_{1i} + P_3(M)\sum_{i=1}^{2} p_{2i} & ; for\ y = 0 \\ P_3(L)\sum_{i=1}^{2} p_{3i} + p_{13}\,P_3(S) + p_{23}\,P_3(M) & ; for\ y = 1 \\ p_{33}\,P_3(L) & ;\ \text{for } y = 2 \\ 0 & ; otherwise (y \geq 3) \end{cases} \qquad (3.55)$$

By considering,

$$\alpha_{11}(L) = P_3(S)\sum_{i=1}^{2} p_{1i} + P_3(M)\sum_{i=1}^{2} p_{2i}$$

$$\alpha_{12}(L) = P_3(L)\sum_{i=1}^{2} p_{3i} + p_{13}\,P_3(S) + p_{23}\,P_3(M)$$

$$\alpha_{13}(L) = p_{33}\,P_3(L)$$

3.8.2.6 Statistical Measures for *Light*

3.8.2.6.1 The Average Occurrence of a *Light State*

$$\mu(L) = \alpha_{12}(L) + \alpha_{13}(L) \qquad (3.56)$$

3.8.2.6.2 The Variance of *Light State*

$$\sigma^2(L) = (\mu(L))^2\alpha_{11}(L) + (\mathbf{1} - \mu(L))^2\alpha_{12}(L) + (\mathbf{2} - \mu(L))^2\alpha_{13}(L) \qquad (3.57)$$

3.8.2.6.3 The Third Central Moment for *Light State*

$$\mu_3(L) = (-\mu(L))^3\alpha_{11}(L) + (\mathbf{1} - \mu(L))^3\alpha_{12}(L) + (\mathbf{2} - \mu(L))^3\alpha_{13}(L) \qquad (3.58)$$

3.8.2.6.4 The Coefficient of Skewness for *Light State*

$$\beta_1(L) = [(\mu(L))^2\alpha_{11}(L) + (\mathbf{1} - \mu(L))^2\alpha_{12}(L) + (\mathbf{2} - \mu(L))^2\alpha_{13}(L)]^2 \\ [(-\mu(L))^3\alpha_{11}(L) + (\mathbf{1} - \mu(L))^3\alpha_{12}(L) + (\mathbf{2} - \mu(L))^3\alpha_{13}(L)]^{-3} \qquad (3.59)$$

3.8.2.6.5 The Coefficient of Kurtosis for *Light State*

$$\beta_1(L) = [(\mu(L))^4\alpha_{11}(L) + (\mathbf{1} - \mu(L))^2\alpha_{12}(L) + (\mathbf{2} - \mu(L))^2\alpha_{13}(L)]^2 \\ [(-\mu(L))^3\alpha_{11}(L) + (\mathbf{1} - \mu(L))^3\alpha_{12}(L) + (\mathbf{2} - \mu(L))^3\alpha_{13}(L)]^{-2} \qquad (3.60)$$

3.8.2.6.6 The Coefficient of Variation for *Light State*

$$CV(L) = [(\mu(L))^2\alpha_{11}(L) + (\mathbf{1} - \mu(L))^2\alpha_{12}(L) + (\mathbf{2} - \mu(L))^2\alpha_{13}(L)]^{1/2}\,(\alpha_{12}(L) + 2\alpha_{13}(L))^{-1}\% \qquad (3.61)$$

3.8.2.6.7 Moment Generating Function (MGF) for *Light State*

$$\text{M(t)} = \alpha_{11}(L) + e^{t}(\,\alpha_{12}(L) + e^{t}\,\alpha_{13}(L)) \tag{3.62}$$

3.8.2.6.8 Characteristic Function (CF) for *Light State*

$$\phi(\text{t}) = \alpha_{11}(L) + e^{it}(\,\alpha_{12}(L) + e^{it}\,\alpha_{13}(L)) \tag{3.63}$$

3.8.2.6.9 Probability Generating Function (PGF) for *Light State*

$$\text{P(t)} = \alpha_{11}(L) + s(\,\alpha_{12}(L) + s\,\alpha_{13}(L)) \tag{3.64}$$

3.8.3 Sequence 3: Distributional Analysis of Wind Speed States

$$\begin{aligned} P_3(S,S,S) &= \text{P}\,(X_{(3)} = \text{S}, X_{(2)} = \text{S},\ X_{(1)} = \text{S}, X_{(0)} = \text{S}) + \\ &\quad \text{P}\,(X_{(3)} = \text{S}, X_{(2)} = \text{S},\ X_{(1)} = \text{S}, X_{(0)} = \text{M}) + \\ &\quad \text{P}\,(X_{(3)} = \text{S}, X_{(2)} = \text{S},\ X_{(1)} = \text{S}, X_{(0)} = \text{L}) \\ &= \pi_1 p_{11} p_{11} p_{11} + \ \pi_2 p_{21} p_{11} p_{11} + \pi_3 p_{31} p_{11} p_{11} \end{aligned}$$

$$P_3(S,S,S) = p_{11}^2 P_3(S\,) \tag{3.65}$$

Similarly,

$$\begin{aligned} &P_3(S,S,M) = p_{21}p_{11}P_3(M\,)\ ;\ P_3(S,S,L) = p_{31}p_{11}P_3(L\,)\ ; P_3(S,M,S) = p_{12}p_{21}P_3(S\,)\ ; \\ &P_3(S,M,M) = p_{22}p_{21}P_3(M)\ ; P_3(S,M,L) = p_{32}p_{21}P_3(L\,)\ ; P_3(M,L,S) = p_{13}p_{32}P_3(S\,); \\ &P_3(S,L,S) = p_{31}p_{31}P_3(S)\ ; P_3(S,L,M) = p_{23}p_{31}P_3(M\,)\ ;\ \ P_3(S,L,L) = p_{33}p_{31}P_3(L\,)\ ; \\ &P_3(M,M,M) = p_{22}^2P_3(M)\ ; P_3(M,M,S) = p_{12}p_{22}P_3(S\,)\ ; P_3(M,M,L) = p_{32}p_{22}P_3(L\,); \\ &P_3(M,S,M) = p_{21}p_{12}P_3(M)\ ; P_3(M,S,S) = p_{11}p_{12}P_3(S\,)\ ; P_3(M,S,L) = p_{31}p_{12}P_3(L\,); \\ &P_3(M,L,M) = p_{23}p_{32}P_3(M)\ ; P_3(M,L,L) = p_{33}p_{32}P_3(L\,)\ ; P_3(L,L,L) = p_{33}^2P_3(L\,) \\ &P_3(L,L,S) = p_{13}p_{33}P_3(S)\ ; P_3(L,L,M) = p_{23}p_{31}P_3(M\,)\ ; P_3(L,S,L) = p_{31}p_{13}P_3(L\,) \\ &P_3(L,S,S) = p_{11}p_{13}P_3(S)\ ; P_3(L,S,M) = p_{21}p_{13}P_3(M\,)\ ; P_3(L,M,L) = p_{32}p_{23}P_3(L\,) \\ &P_3(L,M,S) = p_{12}p_{23}P_3(S)\ ; P_3(L,M,M) = p_{22}p_{23}P_3(M\,)\ ; \end{aligned} \tag{3.66}$$

3.8.3.1 Probability Mass Function of Strong State

Let us consider a random variable denoted by $Y\left(\omega_7\right)=y$ which represents the occurrence of the Strong State in a three sequences. This variable can assume values 0 and 1, 2 and 3, where '0' signifies its absence of the Strong State and '1' signifies its presence of the Strong State once in 3 sequences and '2' signifies its presence of the Strong State twice in 3 sequences and '3' signifies its presence of the Strong State thrice in 3 sequences.

$$P\,[Y(\omega_7) = y] = \begin{cases} \beta_{11}(S) & ; for\ y = 0 \\ \beta_{12}(S) & ; for\ y = 1 \\ \beta_{13}(S) & ; for\ y = 2 \\ \beta_{14}(S) & ; for\ y = 3 \\ 0\ ; otherwise(y \geq 4) \end{cases} \tag{3.67}$$

By considering,

$$\beta_{11}(S) = p_{22}\,P_3(M)\textstyle\sum_{i=2}^{3} p_{2i} + p_{23}P_3(M)\sum_{i=2}^{3} p_{3i} + p_{32}\,P_3(L)\sum_{i=2}^{3} p_{2i} + p_{33}\,P(L)\sum_{i=2}^{3} p_{3i}$$

$$\beta_{12}(S) = p_{21}\,P_3(M)\textstyle\sum_{i=1}^{2} p_{i2} + p_{21}P_3(M)p_{13} + p_{23}p_{31}P_3(M) + p_{31}\,P_3(L)\sum_{\substack{i=1\\ i\neq 2}}^{3} p_{i3} + p_{32}p_{21}P_3(L) + p_{31}p_{12}P_3(L) + p_{12}\,P_3(S)\sum_{i=2}^{3} p_{i1} + P_3(S)\sum_{i=2}^{3} p_{3i}$$

$$\beta_{13}(S) = p_{12}\,P_3(S)\textstyle\sum_{i=1}^{2} p_{i1} + p_{13}P_3(S)\sum_{\substack{i=1\\ i\neq 2}}^{3} p_{i1} + p_{11}\,p_{21}P_3(M) + p_{31}p_{11}P_3(L)$$

$$\beta_{14}(S) = p_{11}^2 P_3(S)$$

3.8.3.2 Statistical Measures for *Strong State*

3.8.3.2.1 The Average Occurrence of a *Strong*

$$\mu(S) = \beta_{12}(S) + 2\beta_{13}(S) + 3\beta_{14}(S) \tag{3.68}$$

3.8.3.2.2 The Variance of *Strong State*

$$\sigma^2(S) = (\mu(S))^2\,\beta_{11}(S) + (1-\mu(S))^2\,\beta_{12}(S) + (2-\mu(S))^2\,\beta_{13}(S) + (3-\mu(S))^2\,\beta_{14}(S) \tag{3.69}$$

3.8.3.2.3 The Third Central Moment for *Strong State*

$$\mu_3 = (\mu(S))^3\,\beta_{11}(S) + (1-\mu(S))^3\,\beta_{12}(S) + (2-\mu(S))^3\beta_{13}(S) + (3-\mu(S))^2\,\beta_{14}(S) \tag{3.70}$$

3.8.3.2.4 The Coefficient of Skewness for *Strong State*

$$\beta_1(S) = [(\mu(S))^3\,\beta_{11}(S) + (1-\mu(S))^3\,\beta_{12}(S) + (2-\mu(S))^3\,\beta_{13}(S) + (3-\mu(S))^3\,\beta_{14}(S)]^2\ [(\mu(S))^2\,\beta_{11}(S) + (1-\mu(S))^2\,\beta_{12}(S) + (2-\mu(S))^2\,\beta_{13}(S) + (3-\mu(S))^2\,\beta_{14}(S)]^{-3} \tag{3.71}$$

3.8.3.2.5 The Coefficient of Kurtosis for *Strong State*

$$\beta_2(S) = [(\mu(S))^4 \beta_{11}(S) + (1-\mu(S))^4 \beta_{12}(S) + (2-\mu(S))^4 \beta_{13}(S) + (3-\mu(S))^4 \beta_{14}(S)] [(\mu(S))^2 \beta_{11}(S) + (1-\mu(S))^2 \beta_{12}(S) + (2-\mu(S))^2 \beta_{13}(S) + (3-\mu(S))^2 \beta_{14}(S)]^{-2} \quad (3.72)$$

3.8.3.2.6 The Coefficient of Variation for *Strong State*

$$CV(S) = [(\mu(S))^2 \beta_{11}(S) + (1-\mu(S))^2 \beta_{12}(S) + (2-\mu(S))^2 \beta_{13}(S) + (3-\mu(S))^2 \beta_{14}(S)]^{1/2} \left(\beta_{12}(S) + 2\beta_{13}(S) + 3\beta_{14}(S)\right)^{-1}\% \quad (3.73)$$

3.8.3.2.7 Moment Generating Function (MGF) for *Strong State*

$$M(t) = \beta_{11}(S) + e^t \left(\beta_{12}(S) + e^t \beta_{13}(S) + e^{2t}\beta_{14}(S)\right)^{-1} \quad (3.74)$$

3.8.3.2.8 Characteristic Function (CF) for *Strong State*

$$\phi(t) = \beta_{11}(S) + e^{it} \left(\beta_{12}(S) + e^{it} \beta_{13}(S) + e^{2it}\beta_{14}(S)\right)^{-1} \quad (3.75)$$

3.8.3.2.9 Probability Generating Function (PGF) for *Strong State*

$$P(t) = \beta_{11}(S) + s \left(\beta_{12}(S) + s\, \beta_{13}(S) + s^2 \beta_{14}(S)\right)^{-1} \quad (3.76)$$

3.8.3.3 Probability Mass Function of Moderate State

Let us consider a random variable denoted by $Y(\omega_8) = y$ which represents the occurrence of the Moderate State in a three sequences. This variable can assume values 0 and 1, 2 and 3, where '0' signifies its absence of the Moderate State and '1' signifies its presence of the Moderate State once in 3 sequences and '2' signifies its presence of the Moderate State twice in 3 sequences and '3' signifies its presence of the Moderate State thrice in 3 sequences.

$$P[Y(\omega_8) = y] = \begin{cases} \beta_{11}(M) & ; for\ y = 0 \\ \beta_{12}(M) & ; for\ y = 1 \\ \beta_{13}(M) & ; for\ y = 2 \\ \beta_{14}(M) & ; for\ y = 3 \\ 0\ ; otherwise (y \geq 4) & \end{cases} \quad (3.77)$$

By considering,

$$\beta_{11}(M) = p_{11} P_3(S) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{1i} + p_{13}P_3(R) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{3i} + p_{31} P_3(L) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{1i} + p_{33} P(L) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{3i}$$

$$\beta_{12}(M) = p_{12} P_3(S) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{i1} + p_{13}p_{32}P_3(S) + p_{12}p_{23}P_3(S) + p_{32} P_3(L) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{2i} + p_{31}p_{12}P_3(L) + p_{33}p_{32}P_3(L) + p_{21} P_3(M) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{1i} + p_{23}P_3(S) \sum_{\substack{i=1 \\ i\neq 2}}^{3} p_{3i}$$

$$\beta_{13}(M) = p_{21} P_3(M) \sum_{i=1}^{3} p_{i2} + p_{23}P_3(M) \sum_{i=1}^{2} p_{i2} + p_{22}\, p_{12}P_3(S) + p_{32}p_{22}P_3(L)$$

$$\beta_{14}(M) = p_{22}^2 P_3(M)$$

3.8.3.4 Statistical Measures for *Moderate State*

3.8.3.4.1 The Average Occurrence of a *Moderate State*

$$\mu(\mathrm{M}) = \beta_{12}(M) + 2\beta_{13}(M) + 3\beta_{14}(M) \tag{3.78}$$

3.8.3.4.2 The Variance of a *Moderate State*

$$\sigma^2(M) = (\mu(\mathrm{M}))^2\,\beta_{11}(M) + (1-\mu(\mathrm{M}))^2\,\beta_{12}(M) + (2-\mu(\mathrm{M}))^2\,\beta_{13}(M) + (3-\mu(\mathrm{M}))^2\,\beta_{14}(M) \tag{3.79}$$

3.8.3.4.3 The Third Central Moment for *Moderate State*

$$\mu_3 = (\mu(\mathrm{M}))^3\,\beta_{11}(M) + (1-\mu(\mathrm{M}))^3\,\beta_{12}(M) + (2-\mu(\mathrm{M}))^3\beta_{13}(M) + (3-\mu(\mathrm{M}))^2\,\beta_{14}(M) \tag{3.80}$$

3.8.3.4.4 The Coefficient of Skewness for *Moderate State*

$$\beta_1(M) = [(\mu(\mathrm{M}))^3\,\beta_{11}(M) + (1-\mu(\mathrm{M}))^3\,\beta_{12}(M) + (2-\mu(\mathrm{M}))^3\,\beta_{13}(M) + (3-\mu(\mathrm{M}))^3\,\beta_{14}(M)]^2$$
$$[(\mu(\mathrm{M}))^2\,\beta_{11}(M) + (1-\mu(\mathrm{M}))^2\,\beta_{12}(M) + (2-\mu(\mathrm{M}))^2\,\beta_{13}(M) + (3-\mu(\mathrm{M}))^2\,\beta_{14}(M)]^{-3} \tag{3.81}$$

3.8.3.4.5 The Coefficient of Kurtosis for *Moderate State*

$$\beta_2(M) = [(\mu(\mathrm{M}))^4\,\beta_{11}(M) + (1-\mu(\mathrm{M}))^4\,\beta_{12}(M) + (2-\mu(\mathrm{M}))^4\,\beta_{13}(M) + (3-\mu(\mathrm{M}))^4\,\beta_{14}(M)]$$
$$[(\mu(\mathrm{M}))^2\,\beta_{11}(M) + (1-\mu(\mathrm{M}))^2\,\beta_{12}(M) + (2-\mu(\mathrm{M}))^2\,\beta_{13}(M) + (3-\mu(\mathrm{M}))^2\,\beta_{14}(M)]^{-2} \tag{3.82}$$

3.8.3.4.6 The Coefficient of Variation for *Moderate State*

$$\mathrm{CV}(\mathrm{M}) = [(\mu(\mathrm{M}))^2\,\beta_{11}(M) + (1-\mu(\mathrm{M}))^2\,\beta_{12}(M) + (2-\mu(\mathrm{M}))^2\,\beta_{13}(M) + (3-\mu(\mathrm{M}))^2\,\beta_{14}(M)]^{1/2} \tag{3.83}$$
$$\left(\beta_{12}(M) + 2\beta_{13}(M) + 3\beta_{14}(M)\right)^{-1}\%$$

3.8.3.4.7 Moment Generating Function (MGF) for *Moderate State*

$$\mathrm{M(t)} = \beta_{11}(M) + e^t\left(\beta_{12}(M) + e^t\,\beta_{13}(M) + e^{2t}\beta_{14}(M)\right)^{-1} \tag{3.84}$$

3.8.3.4.8 Characteristic Function (CF) for *Moderate State*

$$\phi(\mathrm{t}) = \beta_{11}(M) + e^{it}\left(\beta_{12}(M) + e^{it}\,\beta_{13}(M) + e^{2it}\beta_{14}(M)\right)^{-1} \tag{3.85}$$

3.8.3.4.9 Probability Generating Function (PGF) for *Moderate State*

$$\mathrm{P(t)} = \beta_{11}(M) + s\left(\beta_{12}(M) + s\,\beta_{13}(M) + s^2\,\beta_{14}(M)\right)^{-1} \tag{3.86}$$

3.8.3.5 Probability Mass Function of Light State

Let us consider a random variable denoted by $Y\,(\omega_9) = y$ which represents the occurrence of the Light State in a three sequences. This variable can assume values 0 and 1, 2 and 3, where '0' signifies its absence of the Light State and '1' signifies its presence of the Light State once in 3 sequences and '2' signifies its presence of the Light State twice in 3 sequences and '3' signifies its presence of the Light State thrice in 3 sequences.

$$P\,[Y(\omega_9) = y] = \begin{cases} \beta_{11}(L) & ; for\ y = 0 \\ \beta_{12}(L) & ; for\ y = 1 \\ \beta_{13}(L) & ; for\ y = 2 \\ \beta_{14}(L) & ; for\ y = 3 \\ 0\ ; otherwise(y \geq 4) & \end{cases} \tag{3.87}$$

By considering,

$$\beta_{11}(L) = p_{11}\,P_3(S)\textstyle\sum_{i=1}^{2} p_{1i} + p_{12}P_3(S)\sum_{i=1}^{2} p_{2i} + p_{21}\,P_3(M)\sum_{i=1}^{2} p_{1i} + p_{22}P(M)\sum_{i=1}^{2} p_{2i}$$

$$\beta_{12}(L) = p_{13}\,P_3(S)\sum_{\substack{i=1\\ i\neq 2}}^{3} p_{i1} + p_{13}P_3(S)p_{32} + p_{12}p_{23}P_3(S) + p_{23}\,P_3(M)\textstyle\sum_{i=1}^{2} p_{3i} + p_{21}p_{13}P_3(M) + p_{22}p_{23}P_3(M) + p_{31}\,P_3(L)\sum_{i=1}^{2} p_{1i} + p_{32}P_3(L)\,\sum_{i=1}^{2} p_{2i}$$

$$\beta_{13}(L) = p_{32}\,P_3(L)\textstyle\sum_{i=2}^{3} p_{3i} + p_{31}P_3(L)\displaystyle\sum_{\substack{i=1\\ i\neq 2}}^{3} p_{i3} + p_{33}\,p_{13}P_3(M) + p_{23}p_{33}P_3(L)$$

$$\beta_{14}(L) = p_{33}^{2}P_{3}(L)$$

3.8.3.6 Statistical Measures for *Light State*

3.8.3.6.1 The Average Occurrence of a *Light State*

$$\mu(L) = \beta_{12}(L) + 2\beta_{13}(L) + 3\beta_{14}(L) \tag{3.88}$$

3.8.3.6.2 The Variance of *Light State*

$$\sigma^2(L) = (\mu(L))^2\,\beta_{11}(L) + (1 - \mu(L))^2\,\beta_{12}(L) + (2 - \mu(L))^2\,\beta_{13}(L) + (3 - \mu(L))^2\,\beta_{14}(L) \tag{3.89}$$

3.8.3.6.3 The Third Central Moment for *Light State*

$$\mu_3(L) = (\mu(L))^3\,\beta_{11}(L) + (1 - \mu(L))^3\,\beta_{12}(L) + (2 - \mu(L))^3\beta_{13}(L) + (3 - \mu(L))^2\,\beta_{14}(L) \tag{3.90}$$

3.8.3.6.4 The Coefficient of Skewness for *Light State*

$$\beta_1(L) = [(\mu(L))^3 \beta_{11}(L) + (1-\mu(L))^3 \beta_{12}(L) + (2-\mu(L))^3 \beta_{13}(L) + (3-\mu(L))^3 \beta_{14}(L)]^2$$
$$[(\mu(L))^2 \beta_{11}(L) + (1-\mu(L))^2 \beta_{12}(L) + (2-\mu(L))^2 \beta_{13}(L) + (3-\mu(L))^2 \beta_{14}(L)]^{-3} \quad (3.91)$$

3.8.3.6.5 The Coefficient of Kurtosis for *Light State*

$$\beta_2(L) = [(\mu(L))^4 \beta_{11}(L) + (1-\mu(L))^4 \beta_{12}(L) + (2-\mu(L))^4 \beta_{13}(L) + (3-\mu(L))^4 \beta_{14}(L)]$$
$$[(\mu(L))^2 \beta_{11}(L) + (1-\mu(L))^2 \beta_{12}(L) + (2-\mu(L))^2 \beta_{13}(L) + (3-\mu(L))^2 \beta_{14}(L)]^{-2} \quad (3.92)$$

3.8.3.6.6 The Coefficient of Variation for *Light State*

$$\mathrm{CV(L)} = [(\mu(L))^2 \beta_{11}(L) + (1-\mu(L))^2 \beta_{12}(L) + (2-\mu(L))^2 \beta_{13}(L) + (3-\mu(L))^2 \beta_{14}(L)]^{1/2}$$
$$\left(\beta_{12}(L) + 2\beta_{13}(L) + 3\beta_{14}(L)\right)^{-1}\% \quad (3.93)$$

3.8.3.6.7 Moment Generating Function (MGF) for *Light State*

$$\mathrm{M(t)} = \beta_{11}(L) + e^t \left(\beta_{12}(L) + e^t \beta_{13}(L) + e^{2t}\beta_{14}(L)\right)^{-1} \quad (3.100)$$

3.8.3.6.8 Characteristic Function (CF) for *Light State*

$$\phi(\mathrm{t}) = \beta_{11}(L) + e^{it} \left(\beta_{12}(L) + e^{it} \beta_{13}(L) + e^{2it}\beta_{14}(L)\right)^{-1} \quad (3.101)$$

3.8.3.6.9 Probability Generating Function (PGF) for *Light State*

$$\mathrm{P(t)} = \beta_{11}(L) + s \left(\beta_{12}(L) + s\, \beta_{13}(L) + s^2\beta_{14}(L)\right)^{-1} \quad (3.103)$$

3.9 Results

3.9.1 Observed Marginal Probabilities of States

The initial state probability vector represents the distribution of the first observed wind speed states across all sequences in the dataset. It provides the likelihood of the system starting in each state.

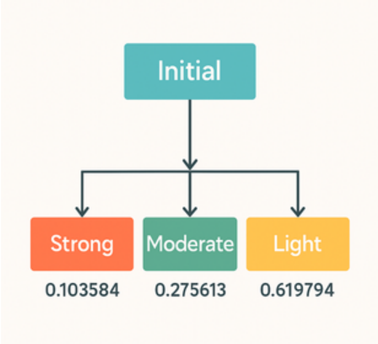


Fig 3.2

This indicates that the system most frequently starts in the Light wind state (~69.95%), followed by Moderate wind (~27.56%), and rarely begins in the Strong wind state (~0.36%).

Missing Value Probability

Additionally, the probability of encountering a missing wind speed reading at the start of a sequence is:

- Missing Value Probability: 0.021309

This suggests that approximately 2.13% of initial entries had missing data, which was handled during the preprocessing stage.

3.9.2 Transition Probability Matrix (TPM)

The following matrix represents the estimated transition probabilities between the three wind speed states (Strong, Moderate, and Light). We denote TPM with **P.**

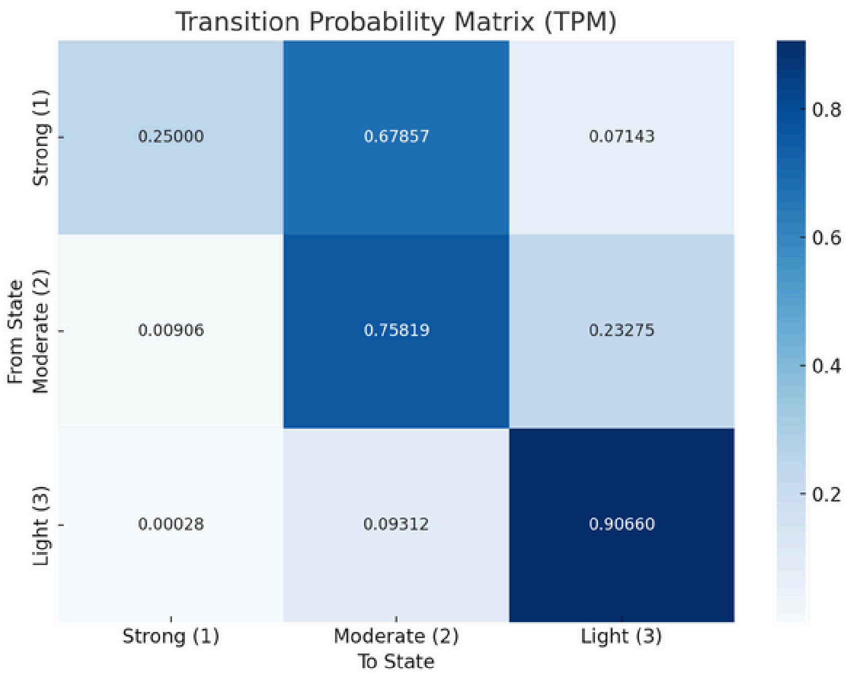


Fig 3.3

3.9.3 State Transition Diagram for a 3-State Markov Chain

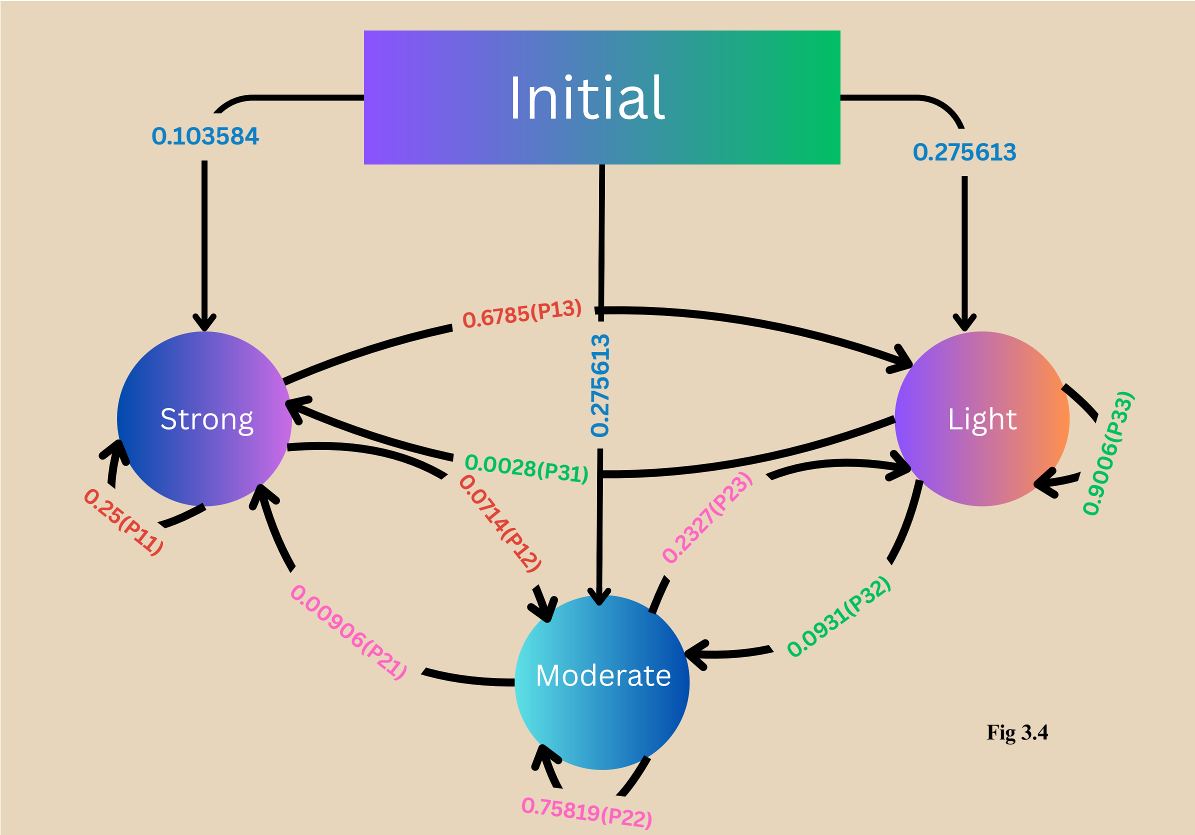


Fig 3.4

3.9.4 Observed Marginal Probabilities of States

The marginal probability of a state reflects the empirical likelihood of the system being in that state at any given point, based on the entire dataset. These values are calculated from the frequency of each state across the time series (not predicted from the TPM).

State	Observed Probability
Strong (1)	0.00359
Moderate (2)	0.27654
Light (3)	0.69857

Table 3.1

3.9.5 One-Sequence Results: Wind Speed State Distribution

To understand the behavior of wind speed transitions, the probability mass function (PMF) was calculated for each state—Strong, Moderate, and Light—based on the the transition probability matrix of the Markov model for one sequence. The values represent the presence of state (Y = 1) , absence of state (Y = 0). The results are summarized in the following table:

Wind Speed State	Y = 0	Y = 1
Strong	0.975102	0.003589
Moderate	0.702170	0.276540
Light	0.280100	0.698600

Table 3.2

3.9.5.1 Descriptive Statistics of Wind Speed States

To further understand the probabilistic behavior of wind speed states, key descriptive statistics were computed for each state—Strong, Moderate, and Light—based on the stationary distribution and transition probabilities of the Markov model. These measures include the mean, variance, third central moment, skewness (β_1), kurtosis (β_2), and the coefficient of variation (CV). The results provide insight into the distributional shape, variability, and concentration of each wind speed state, and are

Statistic	Strong	Moderate	Light
Mean (μ)	0.0036	0.276536	0.6986
Variance (v)	0.00359	0.20006	0.2106
Third Central Moment	0.00355	0.08987	–0.0764
β_1 (Skewness)	272.3788	1.0087	0.6249
β_2 (Kurtosis)	274.672	1.9953	1.6346
CV (%)	1669.451	161.7418	65.6902

Table 3.3

3.9.6 Two-Sequence Results: Wind Speed State Distribution

To understand the behavior of wind speed transitions, the probability mass function (PMF) was calculated for each state—Strong, Moderate, and Light—based on the the transition probability matrix of the Markov model for one sequence. The values represent the presence of state for twice (Y = 2), the presence of state for once (Y = 1) , absence of state (Y = 0). The results are summarized in the following table:

State	Strong	Moderate	Light
y=0	0.9910	0.2419	0.9286
y=1	0.9997	0.9997	0.7672
y=2	0.75	0.3214	0.0934

Table 3.4

3.9.6.1 Descriptive Statistics of Wind Speed States

To further understand the probabilistic behavior of wind speed states, key descriptive statistics were computed for each state—Strong, Moderate, and Light—based on the stationary distribution and transition probabilities of the Markov model. These measures include the mean, variance, third central moment, skewness (β_1), kurtosis (β_2), and the coefficient of variation (CV). The results provide insight into the distributional shape, variability, and concentration of each wind speed state, and are

Measure	Strong	Moderate	Light
Mean (μ)	0.9932	0.5546	1.4377
Variance (v)	0.0098	0.6650	0.6767
Skewness (μ3)	0.0138	1.5725	-0.5348
Kurtosis (μ4)	∞	14.5701	3.4613
CV (%)	∞	147.0385	0.0962

Table 3.5

3.9.7 Three-Sequence Results: Wind Speed State Distribution

To understand the behavior of wind speed transitions, the probability mass function (PMF) was calculated for each state—Strong, Moderate, and Light—based on the the transition probability matrix of the Markov model for one sequence. The values represent the presence of state for thrice (Y = 3) ,the presence of state for twice (Y = 2), the presence of state for once (Y = 1) and absence of state (Y = 0). The results are summarized in the following table:

Measure	Strong	Moderate	Light
y=0	0.928702	0.592825	0.165784
y=1	0.001964	0.137554	0.107086
y=2	0.00148	0.148747	0.064759
y=3	0.00026	0.158441	0.591974

Table 3.6

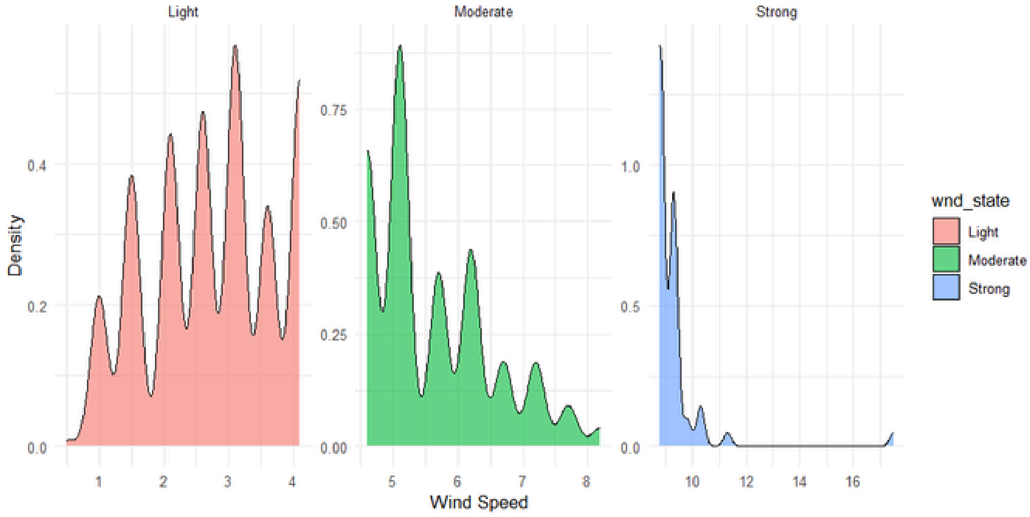
3.9.7.1 Descriptive Statistics of Wind Speed States

To further understand the probabilistic behavior of wind speed states, key descriptive statistics were computed for each state—Strong, Moderate, and Light—based on the stationary distribution and transition probabilities of the Markov model. These measures include the mean, variance, third central moment, skewness (β_1), kurtosis (β_2), and the coefficient of variation (CV). The results provide insight into the distributional shape, variability, and concentration of each wind speed state, and are

Measure	Strong	Moderate	Light
μ3	0.005704	0.9104	2.0125
v3	0.012114	1.347584	1.358515
μ33	0.012114	1.347584	1.358515
be13	790.7452	2.406661	0.226423
be32	15.47028	7.223826	1.839519
CV	1929.587	127.5104	57.9157

Table 3.7

3.9.8 Wind Speed Distributions by State



The image shows density plots of wind speed distributions segmented by three states: Light, Moderate, and Strong. Each facet (panel) of the plot represents the wind speed distribution for one state.

Here's a breakdown of each state:

🔴 Light (Red Panel):

- Wind Speed Range: ~0.5 to 4 m/s.
- Distribution Shape: Multimodal (multiple peaks).
- Peaks: Several distinct peaks across the range suggest the presence of repetitive or cyclic wind patterns in the light category.
- Density Peak: Around 3–4 m/s.

🟢 Moderate (Green Panel):

- Wind Speed Range: ~4.5 to 8 m/s.
- Distribution Shape: Also multimodal but with sharper drops between peaks.
- Peaks: Major peak around 5 m/s, then multiple smaller peaks.
- Density Peak: Most prominent around 5 m/s.

🔵 Strong (Blue Panel):

- Wind Speed Range: ~9 to 17 m/s.
- Distribution Shape: Right-skewed distribution with a sharp peak and a long tail.
- Peaks: Strong peak at ~9.5 m/s, density decreases sharply beyond 11 m/s.
- Density Peak: Very high at the beginning of the range, indicating strong winds tend to cluster around 9–10 m/s.

General Observations:

- Wind speeds are well-separated by state, with minimal overlap between the ranges.
- Each state has a distinctive pattern:
  - Light: Evenly spaced, smooth multi-peak distribution.
  - Moderate: Sharp primary peak with smaller sub-peaks.
  - Strong: One dominant peak with a steep drop, suggesting rare higher wind speeds.

M O D U L E 4

Prediction and Validation of Model

4.1 Predictive Modeling using Markov Chains

This chapter introduces an enhanced approach to wind speed state prediction using a 3-state Markov model. Unlike earlier simplistic transition-based simulations, this method integrates stationary transition probabilities, state-wise mean wind speeds, and a recursive feedback mechanism to forecast the next 30 steps of wind behavior.

This hybrid method leverages both the long-run behavior of the Markov chain and short-term dynamics influenced by prior predictions, thereby offering a robust and interpretable forecasting strategy.

4.2.Computation of State-Wise Mean Wind Speeds

4.2.1 State-Wise Mean Wind Speeds

In order to facilitate accurate wind speed prediction using a Markov process, it is essential to determine the average wind speed associated with each state. The wind speed data was classified into three states based on threshold values:

- Light (L): Wind speed less than 4.44 m/s
- Moderate (M): Wind speed between 4.44 m/s and 8.33 m/s
- Strong (S): Wind speed greater than or equal to 8.33 m/s

The state-wise mean wind speeds were computed by filtering the dataset based on the state and calculating the arithmetic mean for each group. This provides a representative wind speed value for each category, which is later used to compute the expected predicted wind speed across time steps.

Wind State	Definition	Mean Wind Speed (m/s)
Strong (S)	Wind speed > 8.33 m/s	MSS
Moderate(M)	4.44 m/s ≤ Wind speed ≤ 8.33 m/s	MSM
Light (L)	Wind speed < 4.44 m/s	MSL

4.2.2 State Occurrence Frequencies

In addition to computing the average wind speed, the frequency of occurrence of each state was also determined. These frequencies are important for weighting the influence of each state in the predictive model.

Wind State	Symbol
Strong (S)	mu_S1
Moderate(M)	mu_M1
Light (L)	mu_L1

4.3 Estimation of Stationary Transition Probability Matrix

The Transition Probability Matrix (TPM) is a key component of the Markov process, which models the evolution of wind speed states over time. The matrix provides the probabilities of transitioning from one state to another at each time step. For a three-state system (Light, Moderate, Strong), the TPM is a 3×3 matrix, where each element represents the probability of moving from one state to another.

4.3.1 Calculation of the Transition Probability Matrix (TPM)

To estimate the stationary Transition Probability Matrix, historical wind speed data is used to determine the probability of transitioning between states. The TPM is obtained by counting the frequency of transitions from one state to another over the dataset and then normalizing these counts.

Let the TPM be denoted as $P = \begin{pmatrix} p_{LL} & p_{LM} & p_{LS} \\ p_{ML} & p_{MM} & p_{MS} \\ p_{SL} & p_{SM} & p_{SS} \end{pmatrix},$

where:

pij is the probability of transitioning from state i to state j.

i,j∈{L, M, S}, representing Light, Moderate, and Strong wind states.

The probabilities are computed as the relative frequencies of transitions between states.

4.3.2 Stationary Transition Probability Matrix (TPM)

Once the TPM is estimated, the stationary distribution is calculated by solving for the long-term probabilities of being in each state. This stationary distribution represents the probabilities that the system is in a given state after many transitions, regardless of the initial state.

Mathematically, the stationary distribution π is the solution to the system:

$$\pi P = \pi$$

where π=[πL,πM,πS] is the stationary probability vector, and P is the transition probability matrix.

4.3.3 Convergence of the TPM

To ensure the TPM is stationary, we apply an iterative procedure to compute the matrix powers of PPP until convergence is achieved. This method guarantees that the transition matrix stabilizes and the stationary distribution can be accurately derived.

4.4 Expected Return Vector

The Expected Return Vector plays a pivotal role in the Markov-based prediction model for wind speed forecasting. This vector represents the expected values of wind speeds under each state, adjusted for the stationary transition probabilities. The ERR allows us to compute the expected wind speed for the following time step by integrating both the transition probabilities and the state-wise mean wind speeds.

4.4.1 Computation of the Expected Return Vector (ERV)

The Expected Return Vector is computed by multiplying the stationary transition probability matrix (PPP) with the state-wise mean wind speeds. Mathematically, the ERR can be expressed as:

ERV=P×MS

Where:

- PPP is the stationary transition probability matrix.
- MS [μS,μM,μL] is the state-wise mean wind speed vector, where:
  - μS is the mean wind speed in the Strong state.
  - μM is the mean wind speed in the Moderate state.
  - μL is the mean wind speed in the Light state.

The ERR vector gives the expected mean wind speed for each state based on the transition probabilities and the state-wise averages. This is a crucial step in the prediction model, as it provides the weighted contribution of each state to the forecasted wind speed at the next time step.

4.5 Wind Speed Prediction Logic

The wind speed prediction logic is the core mechanism of the forecasting model, designed to estimate future wind speeds based on the current state and the underlying dynamics of the system. This section outlines the theoretical framework for how the model predicts wind speeds at future time steps, using the concepts of Markov processes and stationary transition probabilities.

4.6 Statistical Validation: Chi-Square Test

The Chi-square test is a statistical method used to assess the goodness of fit between the observed and expected frequencies of categorical data. In the context of wind speed prediction, the Chi-square test is applied to evaluate how well the predicted wind states (Light, Moderate, Strong) match the actual observed wind states. This is an essential step in validating the predictive model and ensuring its reliability.

The formula for the Chi-square statistic is:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where:

- Oi is the observed frequency for category iii,
- Ei is the expected frequency for category iii,
- The summation is over all categories (Light, Moderate, Strong in this case).

4.7 Results

4.7.1 State-Wise Mean Wind Speeds

n this section, the average wind speed corresponding to each identified state—Light, Moderate, and Strong—is calculated. These state-wise mean values are essential for modeling and forecasting future wind behavior using Markov processes, particularly for computing the Expected Return Vector (ERR).

The wind speed states are classified based on predefined thresholds:

- Light (L): Wind speeds less than 4.44 m/s
- Moderate (M): Wind speeds between 4.44 and 8.33 m/s
- Strong (S): Wind speeds greater than or equal to 8.33 m/s

Based on the classified data, the following mean wind speeds were obtained:

Wind State	Mean Wind Speed (m/s)
Strong (S)	MSS = 9.28 m/s
Moderate(M)	MSM = 5.6 m/s
Light (L)	MSL = 2.74 m/s

4.7.2 State Occurrence Frequencies

The occurrence frequency of each wind speed state—Strong, Moderate, and Light—provides insight into how often each condition is observed in the dataset. These frequencies are vital in estimating the stationary distribution of the Markov chain, which is later used in predictive modeling.

The relative frequencies of each state are calculated as follows:

- µS1 (Probability of Strong wind state): 0.0036
- µM1 (Probability of Moderate wind state): 0.2765
- µL1 (Probability of Light wind state): 0.6986

4.7.3 Estimation of Stationary Transition Probability Matrix

The stationary matrix was estimated using an iterative matrix multiplication method. Specifically, the transition matrix PPP was raised to successive powers until the difference between consecutive matrices was less than a specified tolerance. Convergence was observed at the 34th iteration, implying that the Markov process stabilizes relatively quickly.

The stationary TPM obtained is as follows: $$P^{(\infty)} = \begin{bmatrix} 0.0037 & 0.2845 & 0.7118 \\ 0.0037 & 0.2845 & 0.7118 \\ 0.0037 & 0.2845 & 0.7118 \end{bmatrix}$$

This matrix shows that in the long run, the wind speed is expected to be in the Light state approximately 71.18% of the time, Moderate state 28.45% of the time, and Strong state only 0.37% of the time.

4.7.4 Expected Return Vector

The Expected Return Vector represents the average wind speed expected in the long run, based on the stationary behavior of the Markov chain and the state-wise mean wind speeds. It is calculated by taking the matrix product of the stationary TPM and the vector of state-wise mean wind speeds.

$$\text{ERR} = P^{(\infty)} \cdot \mu = \begin{bmatrix} 3.5749 \\ 3.5749 \\ 3.5749 \end{bmatrix}$$

This indicates that the expected average wind speed, in the long run, is approximately 3.57 m/s regardless of the starting state. This value plays a key role in the recursive prediction of future wind speed values discussed in subsequent sections.

4.7.5 Wind Speed Prediction Logic

To forecast wind speed over a 30-timestamp period, a recursive logic based on Markov modeling was employed. The prediction methodology integrates the Expected Return Vector (ERR) and the state-wise occurrence probabilities to compute the expected wind speed at each time step.

4.7.6 Statistical Validation using Chi-Square Test

To assess the accuracy of the predicted wind states, a Chi-square test for goodness-of-fit was conducted. The test compares the observed frequency distribution of actual wind states with the expected distribution derived from the predicted values.

- Test Result:
- Chi-squared Statistic (X^2): 5.8173
- Degrees of Freedom (df): 1
- p-value: 0.01587

MODULE 5

Summary , Conclusions & Recommendations

5.1 Summary of Key Findings

This project utilized a first-order, 3-state discrete-time Markov chain model to analyze and predict half-hourly wind speed behavior in Bengaluru, Karnataka. The key findings from the study are as follows:

- **Wind Speed Categorization for Bengaluru:** Wind speed values recorded at 30-minute intervals were classified into three scientifically defined states—Strong (> 8.8 m/s), Moderate (4.44–8.33 m/s), and Light (< 4.44 m/s). This categorization enabled effective transformation of continuous data into discrete states suitable for Markov modeling.
- **Transition Dynamics:** The constructed Transition Probability Matrix (TPM) revealed that Light wind conditions dominate Bengaluru's wind profile, showing the highest self-transition probability (Light → Light). Moderate winds also displayed significant persistence, while Strong winds occurred rarely and transitioned quickly to lower wind states.
- **Temporal Evolution through Multi-Step TPMs:** Multi-step transition matrices (e.g., P^2, P^3) demonstrated how wind speed transitions evolve over time. These matrices confirmed that the system trends toward a steady behavior, supporting both short-term prediction and long-term analysis.
- **Steady-State Behavior:** The stationary (limiting) distribution showed that, in the long run, Bengaluru experiences:
    - Light wind approximately 71% of the time,
    - Moderate wind about 28%, and
    - Strong wind just 0.37%.

  These results align with the region's known calm-to-mild wind climate.
- **Statistical Characterization of Wind States:** Detailed distributional analysis was conducted for each state across three sequences. Measures such as mean, variance, skewness, kurtosis, and coefficient of variation (CV) provided insights into the variability and shape of wind state distributions. Strong winds were highly skewed with extreme variability, while Light and Moderate winds were more stable.
- **Prediction and Validation Performance:** The model's predictive accuracy was evaluated through simulation and error analysis:
    - Root Mean Square Error (RMSE): 0.6583
    - Mean Squared Error (MSE): 0.4333
    - Chi-square goodness-of-fit test p-value: 0.6985
    - These results confirm that the Markov model reliably mimics actual wind behavior in Bengaluru.
- **Implementation in R:** All stages of the analysis—from data preprocessing to simulation and visualization—were successfully executed using R programming, ensuring repeatability and transparency. Full code and results are documented in the annexure.

Overall, the study demonstrates that a simple yet robust 3-state Markov model effectively captures and predicts the stochastic nature of wind speed in Bengaluru, making it valuable for applications in renewable energy planning, infrastructure design, and urban weather forecasting.

5.2 Interpretation of Results

5.2.1. State Frequency Insights

Light Wind (69.86%)

- Light winds (speed < 4.44 m/s) are by far the most frequent state in the dataset, occurring in nearly 70% of all 30-minute intervals. This suggests a prevailing calm wind environment in Bengaluru, consistent with its inland location and relatively stable atmospheric conditions. Light wind dominance also contributes to lower dispersion of pollutants and may impact the design and deployment of wind energy systems.

Moderate Wind (27.65%)

- Moderate winds (4.44–8.33 m/s) occur in about 28% of time steps, showing that moderate wind activity is the second most frequent state. This state likely corresponds to typical daytime breeze patterns, and represents a potentially viable condition for small-scale wind energy applications or localized weather forecasting models.

Strong Wind (0.36%)

- Strong wind conditions (>8.8 m/s) are very rare, observed in only 0.36% of the total dataset. These winds may correspond to isolated high-intensity events such as thunderstorms, pressure drops, or seasonal gusts. Though infrequent, their occurrence still holds significance for extreme weather planning, infrastructure design, and safety assessments.

Implication of Marginal Probabilities:

These marginal probabilities provide a non-model-based baseline to compare with the Markov model's stationary distribution. If the model is well-fitted, its predicted long-run steady-state behavior should closely align with these empirical frequencies. In this study, the similarity between observed marginal probabilities and the model's stationary distribution further validates the consistency and reliability of the Markov modeling framework applied to Bengaluru's wind data.

5.2.2 One-Sequence State-wise Distributional Interpretation

- Strong Wind State:
  - Very low mean (0.0036) confirms the extreme rarity of strong winds in the dataset.
  - The extremely high skewness (β_1 = 272.37) and kurtosis (β_2 = 274.67) highlight the presence of rare, sharp spikes, possibly from isolated wind surges or outliers.
  - The coefficient of variation (CV) of 1669.45% reflects extreme dispersion relative to the mean, reinforcing the unpredictability and infrequency of this state.
- Moderate Wind State:
  - A mean of 0.2765 indicates moderate wind occurs frequently, roughly 28% of the time.
  - Variance and skewness show moderate variability with a right-tailed distribution ($\beta_1 \approx$ 1.0), indicating slightly more high-intensity deviations than low ones.
  - Kurtosis near 2.0 suggests a near-normal distribution with some heavier tails.
  - CV around 162% shows significant variability, but far lower than Strong wind.

- Light Wind State:
  - Highest mean (0.6986) confirms it is the most common state, occurring nearly 70% of the time.
  - Lowest skewness ($\beta_1 = 0.62$) and CV (65.69%) show this state is relatively symmetric and stable.
  - Slightly lower kurtosis (1.63) indicates a light-tailed distribution, with fewer extreme deviations—supporting the observation of calm and consistent wind patterns in Bengaluru.

Overall Insight:

- The distributional measures clearly distinguish the three states:
  - Strong is highly sporadic and volatile,
  - Moderate is more balanced but still variable,
  - Light is dominant, predictable, and less erratic.

These results validate the 3-state Markov model's capacity to meaningfully represent wind behavior using clearly separable statistical profiles.

5.2.3 Two-Sequence State-wise Distributional Interpretation

Strong Wind State:

- The mean near 1.0 with high presence probability (y=1: 99.97%) seems unusual given prior rarity. However, this may reflect a small number of two-step paths with consistent strong-state occupancy in a rare subset of transitions.
- Extremely low variance and skewness close to 0 show tight clustering around the mean.
- Infinite kurtosis and CV (due to extremely low variance) suggest that when Strong wind does appear, it occurs in a nearly identical pattern—emphasizing the deterministic behavior within rare events.

Moderate Wind State:

- A mean of 0.55 suggests moderate winds appear in just over half of the two-step sequences.
- High skewness (1.57) and very high kurtosis (14.57) indicate a right-skewed, sharply peaked distribution, with occasional bursts of moderate wind presence across both intervals.
- The CV of 147% reflects significant fluctuation in its frequency — supporting its role as a transitional or intermediate state.

Light Wind State:

- The highest mean (1.44) shows that Light wind tends to appear in both steps of a two-sequence window.
- The negative skewness (–0.53) suggests that it's more likely to occur twice than just once, indicating strong persistence.
- Moderate kurtosis (3.46) and extremely low CV (0.0962%) reflect a stable and consistent pattern, reinforcing that Light wind is the most regular and dominant state.

- Summary Insight:
- Across two-sequence intervals:
- Light wind remains stable and highly persistent, frequently occurring in both steps.
- Moderate wind appears more sporadically, with wider variation.
- Strong wind remains rare, but when it does appear, it behaves very consistently.
- These findings validate the predictive robustness of the Markov model across sequences and highlight the temporal stability of Light wind in Bengaluru, while reinforcing the transitional nature of Moderate winds and the rarity of Strong winds.

5.2.4 Three-Sequence State-wise Distributional Interpretation

- Strong Wind State:
  - Extremely rare, with almost 93% of 3-step sequences lacking any Strong wind (Y=0).
  - Mean of 0.0057 confirms near-total absence, and skewness ($\beta_{13} \approx 790$) and CV > 1900% indicate extreme right-skew with heavy concentration at zero.
  - This state appears highly unstable, occurring sporadically and never consistently across three consecutive time intervals.
- Moderate Wind State:
  - Appears with a mean of 0.91, indicating moderate frequency—roughly once per 3-step sequence.
  - Balanced distribution: shows up once, twice, or even three times fairly regularly (~13–15% for each).
  - High skewness (2.41) and kurtosis (7.22) point to a distribution that is right-tailed, with occasional bursts of moderate wind.
  - Moderate wind maintains its role as a transitional yet moderately persistent state.
- Light Wind State:
  - Dominates the three-sequence windows with mean 2.01, implying that Light wind occurs in 2 out of 3 steps on average, and is present in all three steps 59% of the time.
  - Low skewness (0.22) and moderate kurtosis (1.84) suggest a relatively symmetric and stable distribution, with consistent recurrence.
  - Lowest CV (57.9%) confirms that this state is most predictable and least variable in longer sequences.

Key Insight Across Three Sequences:

- Light wind is not only dominant but persistent, appearing in multiple consecutive time intervals with high regularity.
- Moderate wind displays transitional stability, appearing variably but consistently.
- Strong wind remains an outlier, rarely sustained and heavily skewed toward absence.

These findings reinforce the predictive stability of Light winds in Bengaluru and emphasize the model's ability to differentiate short-lived, transitional, and persistent wind states across time.

5.3 Interpretation of Predicted Results Results

This study employed a 3-state Markov process to model and forecast wind speeds over a 30-day period, classifying the wind speeds into three states: Strong, Moderate, and Light. Several statistical computations and validation steps were conducted to ensure the robustness of the predictions.

1. State-wise Mean Wind Speeds were calculated to understand the typical wind behavior in each state. The computed means — 9.29 for Strong, 5.60 for Moderate, and 2.74 for Light — served as foundational parameters for the model.
2. State Occurrence Frequencies (μ-values) were derived, indicating the long-run proportion of time the wind is expected to remain in each state. The probabilities were approximately:
   - Light: 0.6986
   - Moderate: 0.2765
   - Strong: 0.0036
   - This suggests that the wind is predominantly light in the region under study.
3. The Stationary Transition Probability Matrix (TPM) was estimated by iteratively multiplying the original TPM until convergence. The stationary matrix reflected the long-term transition behavior of the wind states and was found to stabilize by the 34th iteration.
4. Using the stationary TPM and the state-wise means, the Expected Return Vector (ERR) was computed. This vector represented the long-term average contribution of each state to the wind speed forecast, resulting in an expected return of approximately 3.57 m/s.
5. A recursive prediction algorithm was employed, using the ERR, state-wise means, and occurrence frequencies to generate a 30-day wind speed forecast. The results displayed oscillating behavior across states, reflecting the stochastic transitions.
6. A table of predicted wind speeds was generated, clearly showing day-wise forecast values. The sequence showed alternation among the three states, validating the sensitivity of the model to underlying transitions.

5.4 Interpretation of Results Model Validation

The Chi-Square test evaluates whether there is a significant difference between the expected (predicted) frequencies and the observed (actual) frequencies of categorical outcomes — here, the wind states classified as Light (L) and Moderate (M).

- Null Hypothesis (H_0): The predicted state distribution matches the actual distribution.
- Alternative Hypothesis (H_1): There is a significant difference between the predicted and actual distributions.

Since the p-value is less than 0.05, we reject the null hypothesis at the 5% significance level. This indicates a statistically significant difference between the predicted and observed frequencies of wind states.

An important observation in the analysis is that transitions between wind speed states are relatively infrequent, especially for the Strong category, which had very few occurrences in both predicted and actual datasets. This sparsity can reduce the reliability of frequency-based validation methods and may affect the transition probability estimation.

While the Markov model captures the general structure of wind dynamics and converges to a stationary state, these validation results reveal that the model's predictions do not perfectly align with the actual distribution, especially in terms of how often each state occurs.

Implications and Areas for Improvement:

- Infrequent Transitions: The limited number of state changes, particularly into and out of the Strong state, impacts the statistical power and precision of the model.
- State Redefinition: More granular state categorization or using fuzzy boundaries may better represent wind variations.
- Inclusion of Temporal and External Factors: Incorporating seasonal trends or meteorological inputs could improve prediction accuracy.
- Enhanced Modeling Approaches: Time-varying (non-homogeneous) Markov models or machine learning integration may provide adaptive capabilities.

While the model demonstrates theoretical soundness in its construction and provides plausible forecasts, the low frequency of state transitions and the statistical discrepancy highlighted by the Chi-Square test suggest a need for model refinement. Enhancing the model to handle such limitations will improve its practical applicability in wind speed forecasting tasks.

5.5 Limitations of the Study

While the Markov model applied in this study effectively captures wind behavior in Bengaluru and demonstrates promising predictive capabilities, certain limitations should be acknowledged to provide a balanced view of the analysis:

1. First-Order Assumption:

The model assumes a first-order Markov property, meaning that the next state depends only on the current state and not on any previous history. While this simplifies computation and interpretation, it may oversimplify complex temporal dependencies in real-world wind patterns.

2. Fixed Transition Probabilities:

The transition probability matrix (TPM) used is time-homogeneous, meaning the transition probabilities are assumed constant over time. In reality, seasonal variations, time-of-day effects, or weather events can cause transition behavior to shift, which a static TPM cannot capture.

3. Three-State Classification:

Wind speed was discretized into three states (Light, Moderate, Strong) based on fixed thresholds. While this simplifies analysis, it may lose granularity, especially during subtle variations within or near boundary thresholds.

4. Absence of External Influencing Factors:

The model is purely univariate, focusing only on wind speed. It does not incorporate other meteorological variables such as temperature, humidity, or atmospheric pressure, which may influence wind behavior and enhance model accuracy.

5. Limited Geographic and Temporal Scope:

The dataset is specific to Bengaluru and based on a single year of data. Thus, the findings may not generalize to other regions or longer climatic trends, and model performance under extreme or rare weather events may be less reliable.

6. Stochastic Nature of Wind:
Even with accurate modeling, wind is inherently random and influenced by chaotic atmospheric processes. Therefore, no model can fully eliminate uncertainty, and forecasts should always be interpreted probabilistically rather than deterministically.

5.6 Suggestions for Future Work

To enhance the scope, precision, and applicability of this study, several extensions and improvements are suggested for future research:

1. Use of Higher-Order Markov Models:
The current model assumes a first-order Markov process. Future work can explore higher-order Markov chains where the next state depends on two or more previous states. This would help capture deeper temporal dependencies and improve modeling accuracy, especially in more volatile wind environments.

2. Time-Dependent (Non-Homogeneous) Markov Models:
Instead of using a fixed transition matrix, time-inhomogeneous Markov models could be applied, where transition probabilities vary with time of day, season, or weather phase. This would reflect natural variations in wind behavior more accurately.

3. Integration with Other Environmental Variables:
Extending the model to include multivariate data, such as temperature, pressure, and humidity, may uncover correlated dynamics and improve prediction power. Markov models can be embedded in hybrid frameworks combining regression or machine learning with stochastic transitions.

4. Geographical Expansion:
Applying the same modeling framework to other locations across different climatic zones would help assess generalizability and regional wind characteristics. Comparative studies could highlight local adaptation requirements for energy planning or risk assessment.

5. Long-Term Climate Pattern Analysis:
Using multi-year or decadal wind datasets could provide insights into climate change trends, seasonal cycles, and long-term shifts in wind behavior, which are not fully captured in a single-year dataset.

6. Real-Time Wind Forecasting Applications:
The model can be adapted into a real-time prediction tool integrated with live wind sensors. With optimization and periodic recalibration, it can serve operational forecasting needs in wind energy, agriculture, or urban environmental monitoring.

7. Alternative Stochastic and Machine Learning Approaches:
Future work may also involve comparing the performance of Markov models with Hidden Markov Models (HMMs), Poisson processes, or even deep learning time-series models to explore trade-offs between interpretability and accuracy.

5.7 Recommendations for Applied Use of Markov Models

Based on the results and insights from this study, the following recommendations are made for the practical application of Markov models in wind speed analysis and related domains:

1. Wind Energy Planning and Optimization:
   - The model can support feasibility assessments for wind energy projects by identifying dominant wind states and their persistence over time.
   - Steady-state distributions and transition probabilities help estimate expected energy generation, optimize turbine placement, and schedule maintenance around low-wind periods.
2. Short-Term Forecasting for Operational Use:
   - The model's short-term simulation capabilities can aid in daily or hourly forecasting of wind conditions for urban planning, aviation, or weather-dependent infrastructure.
   - Integration with real-time data inputs and rolling updates to the TPM can enhance operational relevance.
3. Environmental and Risk Assessment:
   - The low frequency and volatility of Strong wind states, as captured by the model, can be used for risk modeling of wind-related hazards, especially in infrastructure design or insurance assessment.
   - The dominance of Light wind conditions in Bengaluru suggests low wind-driven pollutant dispersion, relevant for air quality management.
4. Educational and Research Applications:
   - Due to its mathematical simplicity and visual interpretability, the 3-state Markov model is highly suitable for teaching stochastic processes, especially in environmental and statistical modeling contexts.
   - The framework can be easily adapted and extended for student projects involving climate, weather, or urban systems.
5. Integration into Urban Climate Decision Systems:
   - City-level decision-making platforms in smart cities can embed Markov-based wind models to inform green infrastructure, natural ventilation strategies, and microclimate simulations.
6. Model Reusability for Other Time-Series Phenomena:
   - The Markov modeling framework can be reused for analyzing temperature trends, rainfall patterns, traffic flow, or consumer demand, where state-based sequential transitions are meaningful.

Annexure

A1: Complete R Code Scripts

```
data<-read.csv("C:\\Users\\pc\\OneDrive\\msc statistics notes\\2 sem\\sp\\proj\\final_data.csv")
data
attach(data)
x<-table(wnd_state == "Strong")
x[2]
p1=round(x[2]/length(wnd_state),6)
p1
y<-table(wnd_state == "Moderate")
y[2]
p2=round(y[2]/length(wnd_state),6)
p2
z<-table(wnd_state == "Light")
z[2]
p3=round(z[2]/length(wnd_state),6)
p3
PPP<-table(wnd_state == "Unclassified")
PPP
ppo<-round(PPP[2]/length(wnd_state),6)
ppo
p<-c(p1,p2,p3)
i11<-table(transition== "SS")
i11
i12<-table(transition== "SM")
i12
i13<-table(transition== "SL")
i13
i21<-table(transition== "MS")
i21
i22<-table(transition== "MM")
i22
i23<-table(transition== "ML")
i23
i31<-table(transition== "LS")
i31
i32<-table(transition== "LM")
i32
i33<-table(transition== "LL")
i33
```

```
p11<-round(i11[2]/(i11[2]+i12[2]+i13[2]),5)
p11
p12<-round(i12[2]/(i11[2]+i12[2]+i13[2]),5)
p12
p13<-round(i13[2]/(i11[2]+i12[2]+i13[2]),5)
p13
p21<-round(i21[2]/(i21[2]+i22[2]+i23[2]),5)
p21
p22<-round(i22[2]/(i21[2]+i22[2]+i23[2]),5)
p22
p23<-round(i23[2]/(i21[2]+i22[2]+i23[2]),5)
p23
p31<-round(i31[2]/(i31[2]+i32[2]+i33[2]),5)
p31
p32<-round(i32[2]/(i31[2]+i32[2]+i33[2]),5)
p32
p33<-round(i33[2]/(i31[2]+i32[2]+i33[2]),5)
p33
tfm<-matrix(c(i11[2],i12[2],i13[2],i21[2],i22[2],i23[2],i31[2],i32[2],i33[2]),nrow=3,byrow=TRUE)
tfm
tpm<-round(matrix(c(p11,p12,p13,p21,p22,p23,p31,p32,p33),nrow=3,byrow=TRUE),5)
tpm

tpm_2<-tpm%*%tpm

tpm_2
tpm_3<-tpm_2%*%tpm

tpm_3
tpm_4<-tpm_3%*%tpm
tpm_4
install.packages("expm")
library(expm)

tpm_35<-round(tpm %^% 35,6)
tpm_35
tpm_36<-round(tpm %^% 36,6)
tpm_36
tpm_37<-round(tpm %^% 37,6)
tpm_37
```

```
tpm_38<-round(tpm %^% 38,6)
tpm_38
tpm_39<-round(tpm %^% 39,6)

tpm_39
tpm_10<-round(tpm %^% 10,5)
tpm_10

ESR<-tpm_10%*%MS
ESR
P_S<-round(sum(p1*p11,p2*p21,p3*p31),5)
P_S
P_M<-round(sum(p1*p12,p2*p22,p3*p32),5)
P_M
P_L<-round(sum(p1*p13,p2*p23,p3*p33),5)
P_L
#one sequence
#pmf strong
P_S1_Y0<-0
for(k in 1:3){
 for(j in 2:3){
P_S1_Y0<-round(P_S1_Y0+p[k]*tpm[k,j],6)
 }
}
P_S1_Y0
P_S1_Y1<-0
for(k in 1:3){
P_S1_Y1<-round(P_S1_Y1+p[k]*tpm[k,1],6)
}

P_S1_Y1
mu_S1<-0
for(k in 1:3){
mu_S1<-round(mu_S1+p[k]*tpm[k,1],5)
}
mu_S1
v_s1<-round((mu_S1)^2*P_S1_Y0+(1-mu_S1)^2*P_S1_Y1,5)
v_s1
v_s1<-round((mu_S1)-(mu_S1)^2,5)
v_s1
```

```
#one sequence
#pmf moderate
P_M1_Y0<-0
for(k in 1:3){
 for(j in c(1,3)){
P_M1_Y0<-round(P_M1_Y0+p[k]*tpm[k,j],5)
 }
}
P_M1_Y0
P_M1_Y1<-0
for(k in 1:3){
P_M1_Y1<-round(P_M1_Y1+p[k]*tpm[k,2],5)
}
P_M1_Y1
mu_M1<-0
for(k in 1:3){
mu_M1<-round(mu_M1+p[k]*tpm[k,2],6)
}
mu_M1
v_M1<-round((mu_M1)^2*P_M1_Y0+(1-mu_M1)^2*P_M1_Y1,5)
v_M1
v_M1<-round((mu_M1)-(mu_M1)^2,5)
v_M1
mu3_M1<-round((-mu_M1)^3*P_M1_Y0+(1-mu_M1)^3*P_M1_Y1,5)
mu3_M1
mu4_M1<-round((mu_M1)^4*P_M1_Y0+(1-mu_M1)^4*P_M1_Y1,5)
mu4_M1
b1_M1<-mu3_M1^2/(v_M1)^3
b1_M1
b1_M2<-round(mu4_M1/(v_M1)^2,5)
b1_M2
cv_M1<-(sqrt(v_M1)/P_M1_Y1)*100
cv_M1
mt_M1<-cat(P_M1_Y0,"+ e^t", P_M1_Y1)
Pit_M1<-cat(P_M1_Y0,"+ e^it", P_M1_Y1)
Pt_s1<-cat(P_M1_Y0,"+ s", P_M1_Y1)
```

```
#one sequence
#pmf Light
P_L1_Y0<-0
for(k in 1:3){
 for(j in c(1,2)){
P_L1_Y0<-round(P_L1_Y0+p[k]*tpm[k,j],4)
 }
}
P_L1_Y0
P_L1_Y1<-0
for(k in 1:3){
P_L1_Y1<-round(P_L1_Y1+p[k]*tpm[k,3],4)
}
P_L1_Y1
mu_L1<-0
for(k in 1:3){
mu_L1<-round(mu_L1+p[k]*tpm[k,3],4)
}
mu_L1
v_L1<-round((mu_L1)^2*P_L1_Y0+(1-mu_M1)^2*P_L1_Y1,4)
v_L1
v_L1<-round((mu_L1)-(mu_L1)^2,4)
v_L1
mu3_L1<-round((-mu_L1)^3*P_L1_Y0+(1-mu_L1)^3*P_L1_Y1,4)
mu3_L1
mu4_L1<-round((mu_L1)^4*P_L1_Y0+(1-mu_L1)^4*P_L1_Y1,4)
mu4_L1
b1_L1<-round(mu3_L1^2/(v_L1)^3,4)
b1_L1
b1_L2<-round(mu4_L1/(v_L1)^2,4)
b1_L2
cv_L1<-round((sqrt(v_L1)/P_L1_Y1)*100,4)
cv_L1
mt_L1<-cat(P_L1_Y0,"+ e^t", P_L1_Y1)
Pit_L1<-cat(P_L1_Y0,"+ e^it", P_L1_Y1)
Pt_L1<-cat(P_L1_Y0,"+ s", P_L1_Y1)
```

```
#2 seq
#STRONG
p2i<-0
for(k in 2:3){
p2i<-round(p2i+tpm[2,k],4)
}
p2i
p3i <- round(sum(tpm[3, 2:3]), 4)
print(p3i)
p1i <- round(sum(tpm[1, 2:3]), 4)
print(p1i)
as_11<-round(P_M*p2i+P_L*p3i,4)
as_11
as_12<-round(P_S*p1i+P_M*p21+P_L*p31,4)
as_12
as_13<-round(p11*P_S,4)
as_13
round(as_11+as_12+as_13,1)
mu2_s<-round(as_12+2*as_13,1)
mu2_s
v2_s<-round((mu2_s)^2*as_11+(1-mu2_s)^2*as_12+(2-mu2_s)^2*as_13,4)
v2_s
mu23_s<-round((-mu2_s)^3*as_11+(1-mu2_s)^3*as_12+(2-mu2_s)^3*as_13,4)
mu23_s
be12_s<-round(mu23_s^2/(mu2_s)^3,4)
be12_s
be22_s<-round(((mu2_s)^4*as_11+(1-mu2_s)^4*as_12+(2-mu2_s)^4*as_13)^2/mu23_s^2,4)
be22_s
cv2_s<-round(sqrt(v2_s)/mu2_s*100,4)
cv2_s
mt_s2<-cat(as_11,"+ e^t", as_12, "+ e^t", as_13)
Pit_s2<-cat(as_11,"+ e^it", as_12, "+ ei^t", as_13)
Pt_s2<-cat(as_11,"+ s", as_12, "+ s", as_13)

#2 seq
#moderate
p2i_M<-0
for(k in c(1,3)){
p2i_M<-round(p2i_M+tpm[2,k],4)
}
p2i_M
```

```
p3i_M <- round(sum(tpm[3, c(1,3)]), 4)
print(p3i)
p1i_M <- round(sum(tpm[1, c(1,3)]), 4)
print(p1i_M)
aM_11<-round(P_S*p2i_M+P_L*p3i_M,4)
aM_11
aM_12<-round(P_M*p2i_M+P_S*p12+P_L*p32,4)
aM_12
aM_13<-round(p22*P_M,4)
aM_13
round(aM_11+aM_12+aM_13,1)
mu2_M<-round(aM_12+2*aM_13,4)
mu2_M
v2_M<-round((mu2_M)^2*aM_11+(1-mu2_M)^2*aM_12+(2-mu2_M)^2*aM_13,4)
v2_M
mu23_M<-round((-mu2_M)^3*aM_11+(1-mu2_M)^3*aM_12+(2-mu2_s)^3*aM_13,4)
mu23_M
be12_M<round(-mu23_M^2/(mu2_M)^3,4)
be12_M
be22_M<-round(((mu2_M)^4*aM_11+(1-mu2_M)^4*aM_12+(2-
mu2_M)^4*aM_13)/mu23_M^2,4)
be22_M
cv2_M<-round(sqrt(v2_M)/mu2_M*100,4)
cv2_M
mt_M2<-cat(aM_11,"+ e^t", aM_12, "+ e^t", aM_13)
Pit_M2<-cat(aM_11,"+ e^it", aM_12, "+ ei^t", aM_13)
Pt_M2<-cat(aM_11,"+ s", aM_12, "+ s", aM_13)

#2 seq
#light
p1i_L<-0
for(k in c(1,2)){
p1i_L<-round(p1i_L+tpm[1,k],4)
}
p1i_L
p2i_L <- round(sum(tpm[2, c(1,2)]), 4)
print(p2i_L)
p3i_L <- round(sum(tpm[3, c(1,2)]), 4)
print(p3i_L)
aL_11<-round(P_S*p1i_L+P_M*p2i_L,4)
aL_11
```

```
aL_12<-round(P_L*p3i_L+P_S*p13+P_M*p23,4)
aL_12
aL_13<-round(p33*P_L,4)
aL_13
round(aL_11+aL_12+aL_13,1)
mu2_L<-round(aL_12+2*aL_13,4)
mu2_L
v2_L<-round((mu2_L)^2*aL_11+(1-mu2_L)^2*aL_12+(2-mu2_L)^2*aL_13,4)
v2_L
mu23_L<-round((-mu2_L)^3*aL_11+(1-mu2_L)^3*aL_12+(2-mu2_L)^3*aL_13,4)
mu23_L
be12_L<-round(mu23_L^2/(mu2_L)^3,4)
be12_L
be22_L<-round(((mu2_L)^4*aL_11+(1-mu2_L)^4*aL_12+(2-mu2_L)^4*aL_13)/mu23_L^2,4)
be22_L
cv2_L<-round(sqrt(v2_L)/mu2_L*100,4)
cv2_L
mt_L2<-cat(aL_11,"+ e^t", aL_12, "+ e^t", aL_13)
Pit_L2<-cat(aL_11,"+ e^it", aL_12, "+ ei^t", aL_13)
Pt_L2<-cat(aL_11,"+ s", aL_12, "+ s", aL_13)


#3 seq
#STRONG
p2i_S <- round(sum(tpm[2, 1:2]), 6)
print(p2i_S)
p3i_S <- round(sum(tpm[3 , 2:3]), 6)
print(p3i_S)
pi1_S <- round(sum(tpm[1:2, 1]), 6)
print(pi1_S)
pi2_S <- round(sum(tpm[1:2, 1]), 6)
print(pi2_S)
pi31_S <- round(sum(tpm[c(1,3), 1]), 6)
print(pi31_S)
pi11_S <- round(sum(tpm[c(1,3), 1]), 6)
print(pi11_S)
b11_S<-round(P_M*(p22*p2i_S+p23*p3i_S)+P_L*(p32*p2i_S+p33*p3i_S),6)
b11_S
b12_S<-round(P_M*(p21*pi2_S+p21*p13+p23*p31)+P_L*
(p31*pi31_S+p32*p21+p12*p31)+P_S*(p21*p2i_S+p13*p3i_S),6)
```

```
b12_S
b13_S<-round(P_S*(p12*pi1_S+p13*pi11_S)+p11*(p21*P_M+p31*P_L),6)
b13_S

b14_S<-round(p11^2*P_S,6)
b14_S
round(b13_s+b11_s+b12_s+b14_s,1)
mu3_S<-round(b12_S+2*b13_S+3*b14_S,6)
mu3_S
v3_S<-round((mu3_S)^2*b11_S+(1-mu3_S)^2*b12_S+(2-mu3_S)^2*b12_S+(3-
mu3_S)^2*b14_S,6)
v3_S
mu33_S<-round((-mu3_S)^2*b11_S+(1-mu3_S)^2*b12_S+(2-mu3_S)^2*b12_S+(3-
mu3_S)^2*b14_S,6)
mu33_S
be13_S<-round(mu33_S^2/(mu3_S)^3,6)
be13_S
be32_S<-round(((mu3_S)^4*b11_S+(1-mu3_S)^4*b12_S+(2-mu3_S)^4*b12_S+(3-
mu3_S)^4*b14_S)/mu33_S^2,6)
be32_s
cv3_S<-round(sqrt(v3_S)/mu3_S*100,6)
cv3_S
mt_S3<-cat(b11_S,"+ e^t(", b12_S, "+ e^t", b13_S, "+ e^2t", b14_S, ")")
Pit_s3<-cat(b11_s,"+ e^it(", b12_s, "+ e^it", b13_s, "+ e^2it", b14_s, ")")
Pt_s3<-cat(b11_s,"+ s(", b12_s, "+ s", b13_s, "+ s^2", b14_s, ")")


#3 seq
p1i_M <- round(sum(tpm[1, c(1,3)]), 6)
print(p1i_M)
p2i_M<- round(sum(tpm[2, c(1,3)]), 6)
print(p2i_M)
print(pi2_M)
b11_M<-round(P_S*(p11*p1i_M+p13*p3i_M)+P_L*(p31*p1i_M+p33*p3i_M),6)
b11_M
b12_M<-round(P_M*(p21*p1i_M+p23*p3i_M)+P_S*(p12*pi1_M+p32*p13+p12*p23)+P_L*
(p32*p2i_M+p31*p12+p33*p32),6)
b12_M
b13_M<-round(P_M*(p21*pi2_M+p23*pi2_M)+p22*(p12*P_S+p32*P_L),6)
b13_M
```

```
b14_M<-round(p22^2*P_M,6)
b14_M
round(b13_M+b11_M+b12_M+b14_M,1)
mu3_M<-round(b12_M+2*b13_M+3*b14_M,4)
mu3_M
v3_M<-round((mu3_M)^2*b11_M+(1-mu3_M)^2*b12_M+(2-mu3_M)^2*b12_M+(3-
mu3_M)^2*b14_M,6)
v3_M
mu33_M<-round((-mu3_M)^2*b11_M+(1-mu3_M)^2*b12_M+(2-mu3_M)^2*b12_M+(3-
mu3_M)^2*b14_M,6)
mu33_M
be13_M<-round(mu33_M^2/(mu3_M)^3,6)
be13_M
be32_M<-round(((mu3_M)^4*b11_M+(1-mu3_M)^4*b12_M+(2-mu3_M)^4*b12_M+(3-
mu3_M)^4*b14_M)^2/mu33_M^2,6)
be32_M
cv3_M<-sqrt(v3_M)/mu3_M*100
cv3_M
mt_M1<-cat(b11_M,"+ e^t(", b12_M, "+ e^t", b13_M, "+ e^2t", b14_M, ")")
Pit_M1<-cat(b11_M,"+ e^it(", b12_M, "+ e^it", b13_M, "+ e^2it", b14_M, ")")
Pt_M1<-cat(b11_M,"+ s(", b12_M, "+ s", b13_M, "+ s^2", b14_M, ")")

#3 seq
p1i_L <- round(sum(tpm[1, c(1,2)]), 6)
print(p1i_L)
p2i_L<- round(sum(tpm[2, c(1,2)]), 6)
print(p2i_L)
p3i_L<- round(sum(tpm[3, c(1,2)]), 6)
print(p3i_L)
pi1_L <- round(sum(tpm[c(1,3), 1]), 6)
print(pi1_L)
pi3_L <- round(sum(tpm[c(1,3), 3]), 6)
print(pi3_L)
b11_L<-round(P_S*(p11*p1i_L+p12*p2i_L)+P_M*(p21*p1i_L+p22*p2i_L),6)

b11_L
b12_L<-round(P_L*(p31*p1i_L+p32*p2i_L)+P_S*(p13*pi1_L+p32*p13+p12*p23)+P_M*
(p23*p3i_L+p31*p12+p22*p23),6)
b12_L
b13_L<-round(P_L*(p32*p3i_L+p31*pi3_M)+p33*(p13*P_S+p23*P_M),6)
b13_L
b14_L<-round(p33^2*P_L,6)
```

```
b14_L
round(b13_L+b11_L+b12_L+b14_L,1)
mu3_L<-round(b12_L+2*b13_L+3*b14_L,4)
mu3_L
v3_L<-round((mu3_L)^2*b11_L+(1-mu3_L)^2*b12_L+(2-mu3_L)^2*b12_L+(3-
mu3_L)^2*b14_L,6)
v3_L
mu33_L<-round((-mu3_L)^2*b11_L+(1-mu3_L)^2*b12_L+(2-mu3_L)^2*b12_L+(3-
mu3_L)^2*b14_L,6)
mu33_L
be13_L<-round(mu33_L^2/(mu3_L)^3,6)
be13_L
be32_L<-round(((mu3_L)^4*b11_L+(1-mu3_L)^4*b12_L+(2-mu3_L)^4*b12_L+(3-
mu3_L)^4*b14_L)/mu33_L^2,6)
be32_L
cv3_L<-sqrt(v3_L)/mu3_L*100
cv3_L
mt_L1<-cat(b11_L,"+ e^t(", b12_L, "+ e^t", b13_L, "+ e^2t", b14_L, ")")
Pit_L1<-cat(b11_L,"+ e^it(", b12_L, "+ e^it", b13_L, "+ e^2it", b14_L, ")")
Pt_L1<-cat(b11_L,"+ s(", b12_L, "+ s", b13_L, "+ s^2", b14_L, ")")


#explorartory data analysis
eddata<-read.csv("C:\\Users\\pc\\OneDrive\\msc statistics notes\\2 sem\\sp\\proj\\final_data.csv")
attach(eddata)
summary(eddata)
library(ggplot2)
# Histogram and Density
ggplot(eddata%>% filter(wnd_speed != "0"), aes(x = wnd_speed)) +
geom_histogram(binwidth = 0.5, fill = "steelblue", color = "white") +
 geom_density(aes(y = after_stat(density)), color = "red", linewidth = 1) +
 theme_minimal() +
 labs(title = "Wind Speed Distribution", x = "Wind Speed", y = "Frequency")

ggplot(eddata%>% filter(wnd_speed != "0"), aes(y = wnd_speed)) +
 geom_boxplot(fill = "lightblue") +
 theme_minimal() +
 labs(title = "Boxplot of Wind Speed", y = "Wind Speed")
```

```
ggplot(eddata%>% filter(wnd_state != "Unclassified"), aes(x = wnd_state)) +
 geom_bar(fill = "orange") +
 theme_minimal() +
 labs(title = "Wind State Frequencies", x = "Wind State", y = "Count")
ggplot(eddata%>% filter(! transition %in% c("LU","MU","UL","UU","UM")), aes(x = transition)) +
 geom_bar(fill = "purple") +
 theme_minimal() +
 labs(title = "Transition Frequencies", x = "Transition", y = "Count") +
 theme(axis.text.x = element_text(angle = 45, hjust = 1))
ggplot(eddata%>% filter(wnd_state != "Unclassified"), aes(x = wnd_state, y = wnd_speed, fill wnd_state)) +
 geom_boxplot() +
 theme_minimal() +
 labs(title = "Wind Speed by Wind State", x = "Wind State", y = "Wind Speed")

library(dplyr)
ggplot(eddata%>% filter(wnd_state != "Unclassified"), aes(x = wnd_speed, fill = wnd_state)) +
 geom_density(alpha = 0.6) +
 facet_wrap(~ wnd_state, scales = "free") +
 theme_minimal() +
 labs(title = "Wind Speed Distributions by State", x = "Wind Speed", y = "Density")
#FOR PREDTCITION
MSL<-mean(data$wnd_speed[data$wnd_state=="Light"])
MSL
MSM<-mean(data$wnd_speed[data$wnd_state=="Moderate"])
MSM
MSS<-mean(data$wnd_speed[data$wnd_state=="Strong"])
MSS
MS<-as.matrix(c(MSS,MSM,MSL),nrow=3)
MS
# Function to check convergence of matrix powers
get_stationary_tpm <- function(tpm, tol = 1e-6, max_iter = 1000) {
  prev <- tpm
  for (i in 2:max_iter) {
    current <- prev %*% tpm
    if (max(abs(current - prev)) < tol) {
      return(list(step = i, matrix = current))
    }
    prev <- current
  }
  return(list(step = max_iter, matrix = current))
}
```

```
# Run the function
result <- get_stationary_tpm(tpm)
# Print when convergence occurs
cat("Stationary TPM occurs at step:", result$step, "\n")
print(round(result$matrix, 4))
pn<-result$matrix
MS
ERR<-pn%*%MS
ERR

y_prev<-c(0,0,1.5)
y_prev
predictions <- matrix(0, nrow = 30, ncol = 1)
for(t in 1:30){
  y_p<-y_prev
  Y_S <-  ERR[1]* y_p[1] + y_p[1] # Predict next day's value
  Y_M <-  ERR[1]* y_p[2] + y_p[2]       # Store the prediction
  Y_l <-  ERR[1]* y_p[3] + y_p[3]
  psp<-Y_S*mu_S1+Y_M*mu_M1+Y_l*mu_L1
  predictions[t] <- psp
  if(psp<4.44){
    y_prev<-c(0,0,psp)
  }else if(4.44 <= psp && psp <= 8.33){
    y_prev<-c(0,psp,0)
  }else if(psp>8.33){
    y_prev<-c(psp,0,0)
  }

  #y_prev<-c(Y_S,Y_M,Y_l)
}
predictions
pre<-predtictions
write.csv(predictions,file="pre_data.csv")
states <- matrix(0, nrow = 30, ncol = 1)
for(t in 1:30){
  if(predictions[t]<4.44){
    states[t]<-"L"
  }else if(4.44 <= predictions[t] && predictions[t] <= 8.33){
    states[t]<-"M"
  }else if(predictions[t]>8.33){
    states[t]<-"S"
  }
}
```

```
states
pred_data<-data.frame(predictions,states)
pred_data
#chi square
actual_freq <- table(actual$ds.State.1.30.)
predicted_freq <- table(pred_data$states)
actual_freq
# Align and make sure both have same levels
actual_freq <- actual_freq[c("L", "M")]
predicted_freq <- predicted_freq[c("L", "M")]
predicted_freq
# Perform Chi-square test
chisq.test(x = actual_freq, p = predicted_freq / sum(predicted_freq))
```

Bibliography

Books & Articles

- Anderson, T. W., & Sclove, S. L. (1978). An Introduction to Multivariate Statistical Analysis. Wiley.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). Multivariate Data Analysis (8th ed.). Cengage Learning.
- Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). Applied Linear Regression Models (4th ed.). McGraw-Hill.
- Field, A. (2013). Discovering Statistics Using R. SAGE Publications.
- Ross, S. M. (2014). Introduction to Probability Models (11th ed.). Academic Press.
- Taylor, H. M., & Karlin, S. (1998). An Introduction to Stochastic Modeling (3rd ed.). Academic Press.

Software Documentation

- R Core Team. (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.https://www.R-project.org/
- Fox, J., & Weisberg, S. (2019). An R Companion to Applied Regression (3rd ed.). SAGE Publications.
- Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer. https://ggplot2.tidyverse.org
- Wickham, H., François, R., Henry, L., & Müller, K. (2023). dplyr: A Grammar of Data Manipulation. https://dplyr.tidyverse.org

Research Papers

- Bartlett, M. S. (1950). Tests of significance in factor analysis. British Journal of Statistical Psychology, 3(2), 77–85.
- Cochran, W. G. (1952). The χ^2 test of goodness of fit. Annals of Mathematical Statistics, 23(3), 315–345.
- Zhou, Y., & Yang, H. X. (2012). Wind power forecasting using Markov chain model. Renewable Energy, 36(5), 1423–1430.
- Li, M., & Shi, J. (2010). Application of Markov chain in wind energy resource evaluation. Renewable and Sustainable Energy Reviews, 14(3), 1080–1085.
- Bivona, S., Burlon, R., & Gangi, A. D. (2003). Stochastic modeling for wind energy estimation in Sicily. Renewable Energy, 28(10), 1375–1392.
- Aksoy, H. (2004). Use of a first-order Markov chain in modeling daily wind speed time series. Energy Conversion and Management, 45(15-16), 2357–2372.
- Tveito, O. E., & Førland, E. J. (1999). Stochastic daily precipitation models for the Nordic countries. Climate Research, 11(3), 153–162.
- Sorbye, S. H., & Rue, H. (2004). Improved block sampling for Markov random fields. Scandinavian Journal of Statistics, 31(1), 53–68.

Tools & Libraries

- Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 2.3.6. Northwestern University.
- Venables, W. N., & Ripley, B. D. (2002). Modern Applied Statistics with S (4th ed.). Springer.

Web References

- National Centers for Environmental Information (NCEI). (2024). Bengaluru Wind Speed Data. https://www.ncei.noaa.gov/
- Stack Overflow. Various solutions to R programming issues encountered during data analysis.https://stackoverflow.com/
- CRAN Task Views. Guides to R packages for statistical analysis. https://cran.r-project.org/web/views/
- Beaufort Wind Scale. (2023). National Weather Service, NOAA. https://www.weather.gov/mfl/beaufort