Prediction of Thunderstorms over Kolkata

Introduction

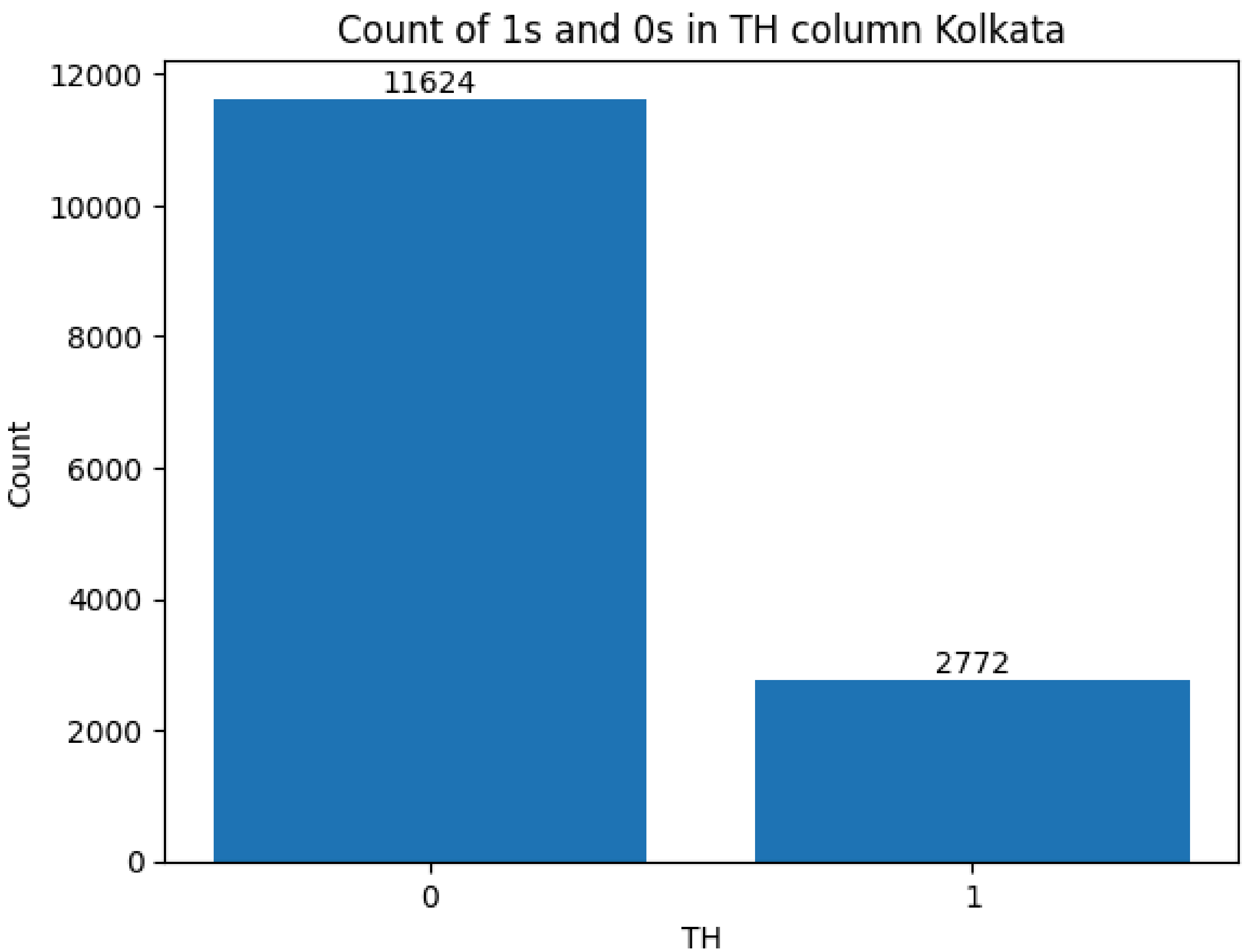
Thunderstorms are captivating weather phenomena characterized by towering cumulus or cumulonimbus clouds that generate lightning and thunder. These convective storms typically have a spatial extent of a few kilometers and a lifespan of less than an hour. However, organized intense convection can lead to multi-cell thunderstorms with longer durations and larger coverage. In India, thunderstorm activity is particularly prominent during the pre-monsoon months of March and May when the atmosphere is highly unstable due to elevated temperatures at lower levels. The eastern and northeastern states, including Gangetic West Bengal, Jharkhand, Orissa, Bihar, Assam, and parts of Northeastern states, witness an increased frequency of severe thunderstorms during this period, commonly known as "Kal-baishakhi" or "Nor'westers."

Severe thunderstorms originating from the northwest and moving southeast over these regions pose significant socio-economic impacts. Locally, they cause thunder, squall lines, lightning, torrential rain, hail, and strong winds that result in agricultural losses, property damage, and unfortunately, loss of life. The casualties reported due to lightning associated with thunderstorms in this region are among the highest in the world. Moreover, the strong winds produced by thunderstorms pose a real threat to aviation, with a high number of aviation hazards reported during their occurrence. In fact, 72% of tornadoes in India are associated with Nor'westers. Consequently, accurate and timely prediction of thunderstorms is crucial for mitigating the loss of life and property caused by these severe local storms.

However, forecasting thunderstorms remains a challenging task due to their limited spatial and temporal extent and the complex dynamics and physics involved. Traditional weather forecasting approaches include empirical and dynamical methods. Empirical forecasting relies on analogues and past weather data, while dynamical forecasting utilizes equations and numerical simulations (NWP) of the atmosphere. Empirical approaches, such as regression, stochastic modeling, fuzzy logic, and artificial neural networks (ANNs), have shown promise in capturing complex relationships and patterns in weather data. ANNs, in particular, are trainable self-adaptive systems capable of learning from examples and generalizing acquired knowledge to solve unforeseen problems.

Total Thunderstorm in Kolkata = **2772**

By combining historical weather data, including surface observations and relevant indexes, with advanced ML-based models, this research endeavors to improve thunderstorm prediction and provide valuable insights for weather forecasting and risk management in these regions.



Data Collection and Preprocessing:

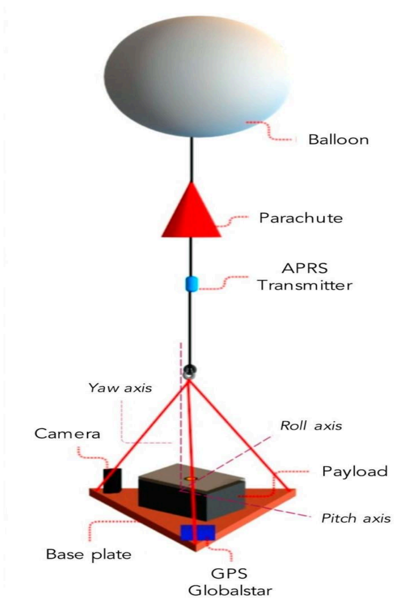
Data Collection:

Data for the thunderstorm prediction model was collected daily from 1980 to 2020, utilizing two different types of data: index data and surface data.

Index Data:

Index data was collected using two weather balloons launched at specific times during the day. The first balloon, known as a Radiosonde , was equipped with sensors to gather observations above ground level. Launched at 5:30 AM IST (0 UTC), the radiosonde collected air pressure, temperature, and relative humidity data. It carried sensors to measure environmental conditions and ascended at a rate of around 20 km/hr, filled with approximately 375 grams of hydrogen gas.

The second balloon, called the Pilot Weather Balloon, was launched at 4:30 PM IST (11 AM UTC) without any sensors. Its purpose was to detect wind speed and direction using a theodolite—a device similar to a surveyor's transit. The amount of hydrogen gas filled in the balloon varied based on cloud cover. In the presence of more clouds, such as during the monsoon season, the balloon was filled with 125 grams of hydrogen gas, weighing around 30 grams. In clear skies, the balloon carried up to 325 grams of hydrogen gas, weighing around 70 grams. The Pilot Weather Balloon allowed for the determination of wind characteristics.



Surface Data:

Surface data was recorded twice a day—at 8:30 AM IST (3 UTC) and 5:30 PM IST (12:30 PM UTC) . Ground instruments were used to measure various factors, including wind speed, temperature (minimum and maximum), rainfall, cloud visibility, wind direction, evaporation, and dew point.

The instruments used for surface data collection included:

Cup Anemometer (Wind Speed): This instrument measured wind speed at the surface.

Cup Anemometer (Wind Direction): It provided wind direction information with 15 sub-directions.

Rain Gauge: Made of fiber-reinforced plastic (FRP), the rain gauge measures rainfall.

Single Stevenson Screen: Equipped with four thermometers, it recorded minimum temperature, maximum temperature, dry bulb temperature, and wet bulb temperature.

Double Stevenson Screen: It contained a thermograph for temperature vs. time on a drum clock and a Hair Hygrograph for humidity percentage on another drum clock.

Self-Recording Rain Gauge: Used to measure rainfall automatically.

AWS (Automatic Weather Station) with Solar Panels: These stations collected various weather parameters.

Soil Thermometer: Installed at different depths, including the surface, 5 cm, 10 cm, 15 cm, and 20 cm, the soil thermometers were used for irrigation research purposes.

Open Pan Evaporation (OPE): It measured the daily evaporation.

Dew Gauge Stand: Used to measure dew points.

The combination of index data and surface data provided a comprehensive set of weather observations for training the neural network model and predicting thunderstorm occurrences.

Relevant Weather Indexes for Thunderstorm Prediction:

Description of various weather indexes used in the study from the index data :

GMT: The Greenwich Mean Time of the data.

SWEAT index: A parameter used to assess the potential for severe weather.

Showalter index: A stability index that measures the temperature difference between an air parcel and its surrounding environment.

LIFTED index: A measure of the buoyancy of an air parcel.

K index: A measure of the potential for thunderstorm development based on temperature and moisture profiles.

Cross totals index: An index used to assess the stability of the atmosphere.

Vertical totals index: An index used to measure the vertical temperature difference between two atmospheric levels.

Totals totals index: A combination of the cross totals and vertical totals indexes.

TLCL: The temperature at which a lifted parcel reaches its lifted condensation level.

PLCL: The pressure at which a lifted parcel reaches its lifted condensation level.

CINE: Convective Inhibition Energy, which measures the amount of energy required to overcome a stable layer of air.

CAPE: Convective Available Potential Energy, which represents the amount of energy available for convection.

PRECIPITABLE WATER: The Precipitable Water value.

1000-500 THICKNESS: The 1000-500 Thickness value.

Most relevant and important indexes in thunderstorm occurrence :

SWEAT index:

Relevance: The SWEAT index is relevant because it incorporates multiple atmospheric variables, such as temperature, dew point, wind speed, and wind direction. By considering these factors, it provides a holistic assessment of atmospheric instability, which is crucial for identifying conditions conducive to severe weather, including severe thunderstorms.

Significance: The SWEAT index plays a vital role in evaluating the potential for severe weather events. It helps meteorologists and researchers gauge the atmospheric energy available for thunderstorm development and assess the likelihood of severe thunderstorms based on the combined influence of temperature, moisture, and wind patterns.

Showalter index:

Relevance: The Showalter index is relevant as it quantifies the temperature difference between an air parcel and its surrounding environment. This information is crucial for understanding the stability of the atmosphere and identifying conditions that promote or inhibit vertical motion.

Significance: The Showalter index holds significance in thunderstorm prediction as it helps determine the presence of unstable conditions. A lower or negative Showalter index indicates a greater likelihood of convective activity, as the air parcel lifted from the surface would be warmer and more buoyant than its surroundings, leading to upward motion and potential thunderstorm development.

LIFTED index:

Relevance: The LIFTED index is relevant because it quantifies the buoyancy of an air parcel, providing insights into its vertical motion potential. It helps in understanding the atmospheric instability and the ability of air parcels to rise and form thunderstorms.

Significance: The LIFTED index is significant in assessing thunderstorm potential as it indicates the strength of updrafts within the atmosphere. A lower (more negative) LIFTED index suggests stronger updrafts, which are crucial for sustaining deep convection and facilitating the formation and intensification of thunderstorms.

K index:

Relevance: The K index is relevant as it considers temperature and moisture profiles at different atmospheric levels. It captures the combined influence of these variables on thunderstorm development.

Significance: The K index is significant in evaluating the potential for thunderstorm development. By incorporating temperature and moisture information, it provides insights into atmospheric instability and the availability of energy for convective activity. Higher K index values indicate a higher potential for thunderstorm development.

Cross totals index:

Relevance: The Cross totals index is relevant as it assesses the stability of the atmosphere by evaluating temperature variations between the lower and middle troposphere.

Significance: The Cross totals index holds significance in thunderstorm prediction as it provides information about the atmospheric stability. Higher Cross totals index values indicate greater instability, which enhances the potential for vertical motion and thunderstorm development.

Vertical totals index:

Relevance: The Vertical totals index is relevant as it measures the vertical temperature difference between two atmospheric levels, reflecting the temperature gradient and stability of the atmosphere.

Significance: The Vertical totals index is significant in understanding thunderstorm potential. It helps identify conditions that support vertical instability, which is essential for thunderstorm initiation and intensification.

Totals totals index:

Relevance: The Totals totals index is relevant as it combines the Cross totals and Vertical totals indexes, providing a comprehensive assessment of atmospheric stability.

Significance: The Totals totals index has significance in evaluating overall atmospheric stability . By considering both horizontal and vertical temperature variations, it offers insights into the combined influence of these factors on thunderstorm formation.