‘Global Urbanization and Climate Metrics

[1]EMAN ANSARI’

Dubai College

ABSTRACT

Using a data-centric approach, this paper investigates the dynamic relationship between global urbanization trends and climate metrics. The study explores how population growth and urban density influence energy consumption and contribute to the emission of greenhouse gases in various urban regions. Drawing data from international sources such as the World Bank and other global development indices, key environmental indicators—including land surface temperature (LST), the Normalized Difference Vegetation Index (NDVI), and CO_2 emissions—are analyzed across a selection of developed and developing countries. By conducting comparative analysis, this research highlights how differing economic conditions shape environmental sustainability outcomes. The findings aim to provide evidence-based insights into whether urban economic expansion promotes or undermines climate resilience, offering guidance for sustainable urban development policies worldwide.

Problem statement: To forecast the next year’s CO_2 emissions for 10 randomly selected countries using historical annual data.

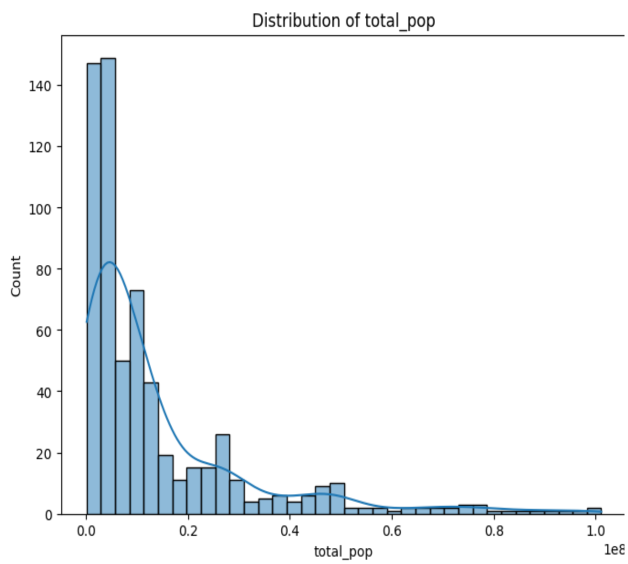
1. INTRODUCTION

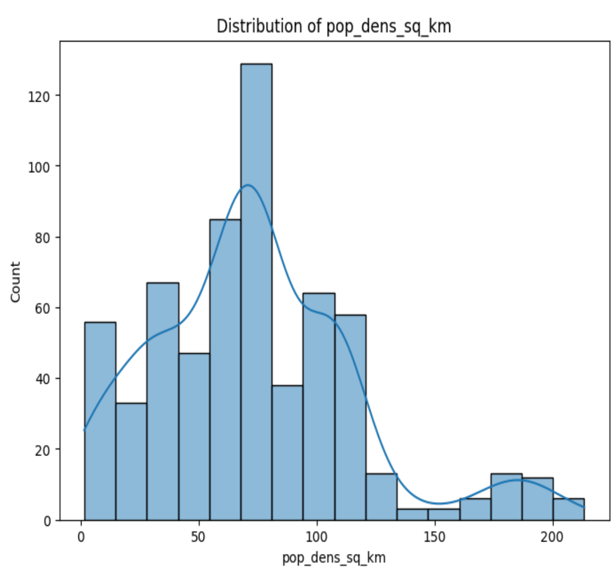
Global trends and economics gose together but creates problems for the nature Understanding how the fast urban population growth is affecting the environment is crucial, particularly in developed and developing countries. This study aims to investigate the direct relationship between urban population density and size and energy consumption patterns, which in turn influence CO2 emissions and other climate-related factors. The study employs a data-centric methodology to examine a range of environmental indicators, including land surface temperature (LST), greenhouse gas (GHG) emissions, and the Normalized Difference Vegetation Index (NDVI), using publicly available datasets from national statistical databases and global development indices like the World Bank and UN-Habitat.. Python and its many libraries will be our primary tools for building the models and graphs. In order to identify trends and linkages between ecological deterioration and urbanization, we will employ long short-term memory ([3]LSTM) for predictive analysis. An important part of the research is comparing cities with different levels of economic development. Niger, Nepal, Albania, Ghana, Somalia, Finland, Greece, Armenia, the Democratic Republic of the Congo, and Hong Kong are among the countries selected for analysis, representing a broad spectrum of urbanization levels and economic statuses. To examine the economic impact on global greenhouse gas levels, we select both high- and low-income countries. Finding out if urban economic growth encourages sustainable development or harms the environment is the aim of this multidisciplinary study. Positive environmental outcomes are not guaranteed by economic development alone; effective governance, the enforcement of the law, and ecological awareness are essential. The comparative case studies of Finland and Hong Kong illustrate this. By examining significant climate and urbanization metrics, this study provides valuable insights that can support data-driven urban planning and targeted climate action. These findings ultimately aim to guide the development of more resilient, equitable, and environmentally sustainable cities in light of the world's rapidly increasing rate of urbanization.

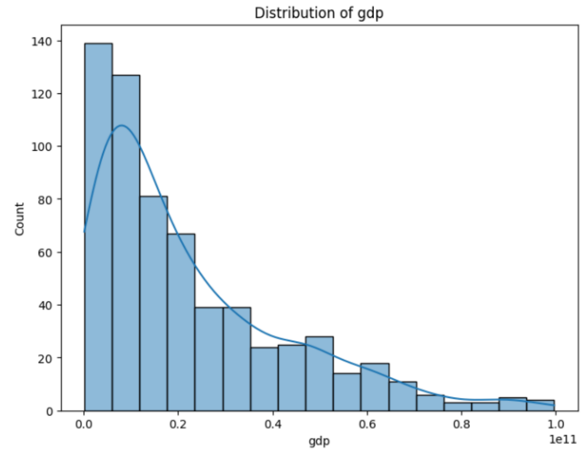
[2]Dataset Exploration:

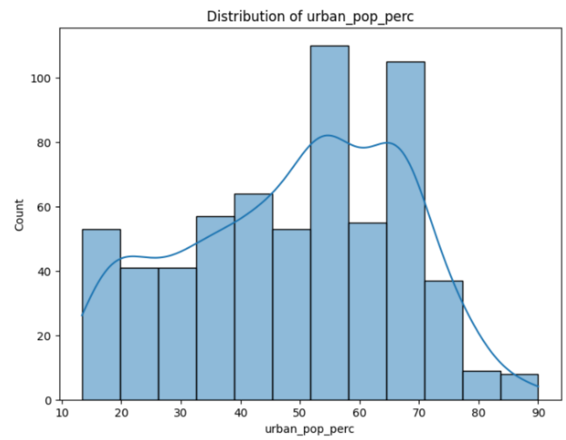
1. country
   Name of the country or regional grouping (e.g., *Aruba, Afghanistan, Africa Eastern and Southern*).
2. country_code
   3-letter ISO code or region code (e.g., *ABW, AFG, AFW, ZAF*).
3. year
   Calendar year of observation (range: 1960–2023).
4. total_pop
   Total population of the country/region in that year.
   Example: *Afghanistan (1960): 9,035,043*.
5. pop_dens_sq_km
   Population density (people per square km).
   Often missing (NaN) in this dataset.
6. gdp
   Gross Domestic Product (current US$).
   Example: *Africa Eastern and Southern (1960): 2.42×10^{10}*.
7. urban_pop_perc
   Percentage of population living in urban areas.
   Example: *Afghanistan (1960): 8.40%*.
8. rural_pop_perc
   Percentage of population living in rural areas.
   Example: *Afghanistan (1960): 91.60%*.
9. elect_access_pop
   Percentage of population with access to electricity.
10. internet_use_perc
    Percentage of population using the Internet.
    Often missing in early years (1960s, 1970s).

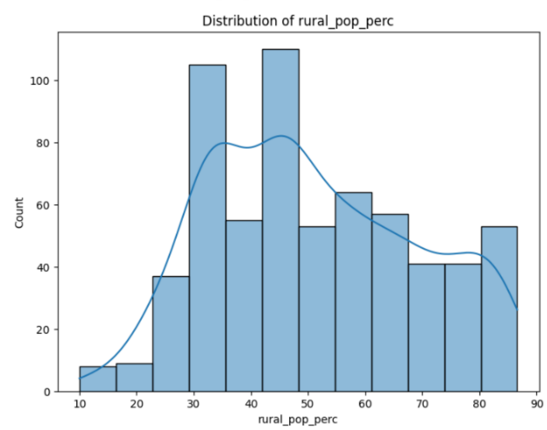
	country	country_code	year	total_pop	pop_dens_sq_km	gdp	urban_pop_perc	rural_pop_perc	elect_access_pop
0	Aruba	ABW	1960	54922.0	NaN	NaN	50.776000	49.224000	NaN
1	Africa Eastern and Southern	AFE	1960	130072080.0	NaN	2.421063e+10	14.576676	85.423324	NaN
2	Afghanistan	AFG	1960	9035043.0	NaN	NaN	8.401000	91.599000	NaN
3	Africa Western and Central	AFW	1960	97630925.0	NaN	1.190495e+10	14.710006	85.289994	NaN
4	Angola	AGO	1960	5231654.0	NaN	NaN	10.435000	89.565000	NaN
...	...	...	...	...	...	...	...	...	...
17019	Kosovo	XKX	2023	1756366.0	NaN	1.046822e+10	NaN	NaN	NaN
17020	Yemen, Rep.	YEM	2023	39390799.0	NaN	NaN	39.831000	60.169000	NaN
17021	South Africa	ZAF	2023	63212384.0	NaN	3.806993e+11	68.819000	31.181000	NaN
17022	Zambia	ZMB	2023	20723965.0	NaN	2.757796e+10	46.335000	53.665000	NaN

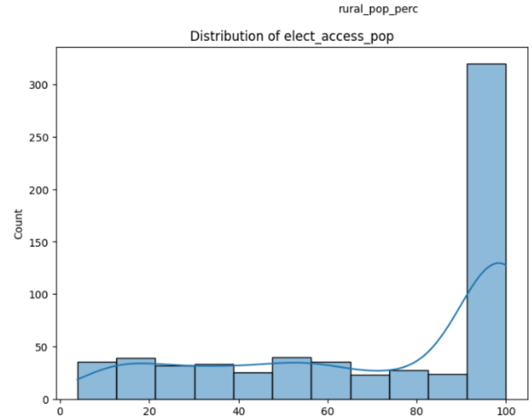








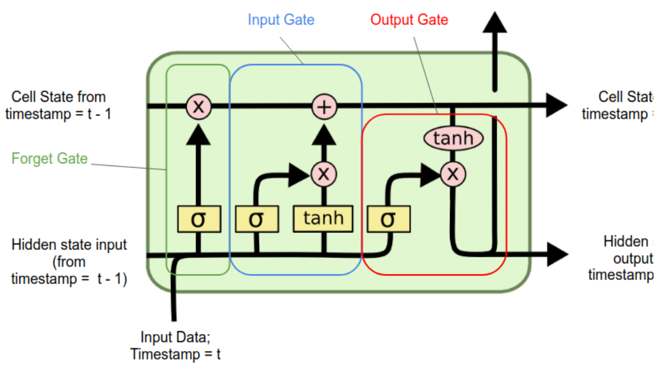




Methodology:

Long Short-Term Memory (LSTM) networks are a specialized form of **Recurrent Neural Networks (RNNs)** designed to learn from **sequential data** while overcoming the limitations of traditional RNNs, particularly the **vanishing and exploding gradient problems**.

In a traditional RNN, each output is influenced not only by the current input but also by the hidden state from the previous time step. However, over long sequences, standard RNNs struggle to retain important information due to vanishing gradients during backpropagation. LSTM solves this by introducing a **cell state** (a form of memory) and **three gates**—the **input gate**, **forget gate**, and **output gate**—which regulate the flow of information.



For a regression task (e.g., predicting poverty rates), LSTM networks are typically trained using the

Mean Squared Error (MSE):

$MSE = (1 / n) \times \Sigma (y_i - \hat{y}_i)^2$

Where:

y_i → Actual value (observed data point)

$\hat{y}_i$ → Predicted value from the model

n → Number of predictions (total data points)

At each time step *t*, the LSTM computes:

1. **Forget Gate** (decides what information to discard from cell state):

$f_t = \sigma(W\_f \cdot x_t + U\_f \cdot h_{t-1} + b\_f)$

2. **Input Gate** (decides what new information to add):

$i_t = \sigma(W\_i \cdot x_t + U\_i \cdot h_{t-1} + b\_i)$

3. **Candidate Cell State** (new candidate values to update memory):

$\tilde{C}_t = \tanh(W\_c \cdot x_t + U\_c \cdot h_{t-1} + b\_c)$

4. **Update Cell State** (combine old state and new candidate):

$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$

5. **Output Gate** (decides what to output from memory):

$o_t = \sigma(W\_o \cdot x_t + U\_o \cdot h_{t-1} + b\_o)$

6. **Hidden State Update** (final output of LSTM unit):

$h_t = o_t \odot \tanh(C_t)$

$\odot$ = **element-wise multiplication**,

Forget Gate (f_t): Removes irrelevant information.

Input Gate ($i_t + \tilde{C}_t$): Decides which new values to store.

Cell State (C_t): The memory, updated every step.

Output Gate (o_t): Controls how much memory flows to output.

W_f, W_i, W_c, W_o → Weight matrices for the **input x_t** (for forget, input, candidate, and output gates).

U_f, U_i, U_c, U_o → Weight matrices for the **previous hidden state h_{t-1}.**

b_f, b_i, b_c, b_o → Bias terms for each gate.

Model Structure

- **Input shape**: (sequence_length, 1)
  - sequence_length = 10 (past 10 years)
  - 1 = 1 feature (CO_2 emission per year)
- **Layers**:
  - LSTM (64 units) with ReLU activation to learn temporal dependencies.
  - Dense (1) output layer to predict the next emission value.
- **Loss Function**: Mean Squared Error (MSE)
- **Optimizer**: Adam (efficient for deep learning)
- **Epochs**: 50 (you can adjust for more accuracy)

Workflow

1. **Normalization**: CO_2 emissions are scaled using MinMaxScaler to help the LSTM learn better.
2. **Sliding Window**: Input sequences are created using a window of 10 years to predict the 11th year.
3. **Model Training**: Each country's model is trained on its own historical data.
4. **Prediction**: The model forecasts the next year's emission using the most recent 10-year sequence.
5. **Visualization**: First, individual country plots are shown; then a combined plot visualizes trends and predictions for all countries.

Model prediction for co2 emission past data and predicted for next 1 year

```
['Niger','Nepal','Albania', 'Ghana',
'Somalia','Finland','Greece','Armeni
a','Democratic Republic of
Congo','Hong Kong']
```

Case study: Finland

Predicted CO_2 Emissions for 2026: 41,333,168 metric tonnes

Context:

In Finland, the energy sector remains a major contributor to CO_2 emissions, primarily due to the burning of fossil fuels such as coal, oil, and natural gas for electricity, heating, and transportation.

Environmental Concerns:

According to the World Wildlife Fund (WWF), around 700 forest species in Finland are currently endangered, largely due to extensive logging activities. This is exacerbated by:

Inadequate protection through small nature reserves

Lack of old-growth forests, which are vital for biodiversity and act as strong carbon sinks.

Implication:

Despite Finland's high level of development and progressive environmental policies, industrial practices—particularly those tied to forestry and energy—continue to pose challenges to sustainability. The predicted CO_2 emissions highlight the ongoing environmental pressure and the need for stricter conservation and cleaner energy transitions.

Case Study: Hong Kong

Predicted CO_2 Emissions for Next Year (2026): 40,678,476 metric tonnes

Recent Emissions Data:

In 2023, Hong Kong's GHG emissions stood at 34.5 million tonnes CO_2-equivalent, showing:

A 20% reduction compared to 2005 levels

A 25% decrease from peak emissions in 2014

Context:

Hong Kong has made notable progress in reducing emissions, likely due to:

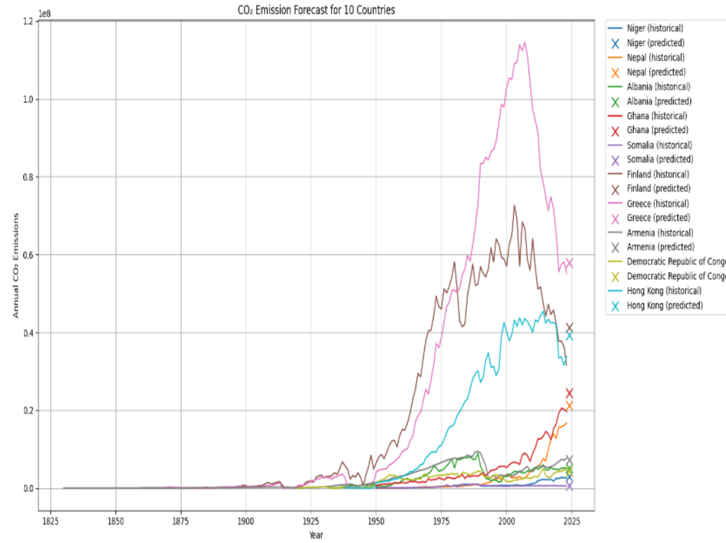
Shifting from coal to natural gas and cleaner energy sources

Implementing energy efficiency programs

Urban planning efforts to reduce vehicle use and promote public transport

Implication:

Even though the predicted value for 2026 appears highethan the 2023 emissions, it may reflect future increases tie to population density, economic activity, or delayed policy enforcement. It suggests a potential reversal of the positive trend unless more aggressive climate actions are taken.



Conclusion:

The comparative analysis of Finland and Hong Kong highlights the complex and region-specific dynamics between urban development, energy consumption, and environmental sustainability. Despite Finland's reputation as a green leader, the projected CO_2 emissions for 2026 suggest that significant challenges remain, particularly due to its continued reliance on fossil fuels and unsustainable logging practices. The loss of biodiversity and inadequate forest conservation efforts underscore the need for more robust environmental protection policies.

Conversely, Hong Kong has demonstrated notable progress in reducing greenhouse gas emissions over the past decade, achieving substantial declines since its 2014 peak. However, the projected increase in emissions for 2026 signals potential setbacks, possibly linked to rising energy demands or slowed policy implementation in a densely populated urban setting.

References:

1. emanansari.com
2. https://ourworldindata.org/climate-change