**Supplemental Material**

Table S1. Key Microclimatic and Socio-Economic Mobility (SoEM) Indicators Used in Assessing Urban Inequality: A Review of the Literature

|  |  |  |
| --- | --- | --- |
| **Dimension** | **Indicator** | **Reference** |
| **Microclimatic Parameters** | Air Temperature | Liu et al., 2023; Sherif et al., 2025; Huang et al., 2022; Guerri et al., 2023; Tsoka et al., 2017; Hang et al., 2024; Nakyai et al., 2025; Kliengchuay et al., 2018; Wang & Ogawa, 2015; Ivanovski et al., 2023 |
| Relative Humidity | Liu et al., 2023; Sherif et al., 2025; Huang et al., 2022; Guerri et al., 2023; Tsoka et al., 2017; Hang et al., 2024; Nakyai et al., 2025; Wang & Ogawa, 2015; Ivanovski et al., 2023 |
| Wind Speed | Liu et al., 2023; Sherif et al., 2025; Huang et al., 2022; Guerri et al., 2023; Tsoka et al., 2017; Hang et al., 2024; Girotti et al., 2025; Kliengchuay et al., 2018; Wang & Ogawa, 2015; Ivanovski et al., 2023 |
| Wind Direction | Liu et al., 2023; Sherif et al., 2025; Guerri et al., 2023; Tsoka et al., 2017; Girotti et al., 2025; Kliengchuay et al., 2018; Wang & Ogawa, 2015 |
| Cloud Cover | Liu et al., 2023; Tsoka et al., 2017 |
| Precipitation | Liu et al., 2023; Huang et al., 2022; Nakyai et al., 2025 |
| Dew Point Temperature | Sherif et al., 2025 |
| Solar Radiation | Sherif et al., 2025; Tsoka et al., 2017 |
| Roughness Length | Guerri et al., 2023; Tsoka et al., 2017 |
| Atmospheric Pressure | Hang et al., 2024; Nakyai et al., 2025 |
| Water Vapor % | Hang et al., 2024 |
| General Atmospheric Conditions | Girotti et al., 2025; Nakyai et al., 2025 |
| **Socioeconomic and Mobility Indicators** | Housing Prices per Meter | Zhu et al., 2022; Hsu et al., 2021; Zarghamfard et al., 2023 |
| Average Income | Grajdura et al., 2021; Viezzer & Biondi, 2021; Grajdura & Niemeier, 2022; SÃ¡nchez et al., 2017; Smith et al., 2024 |
| Education Levels | Grajdura et al., 2021; Viezzer & Biondi, 2021; Grajdura & Niemeier, 2022; Reckien et al., 2018; SÃ¡nchez et al., 2017; Smith et al., 2024; Renteria et al., 2021; Ghasemi et al., 2018; Huang et al., 2011 |
| Household Size | Grajdura et al., 2021; Reckien et al., 2018; Grajdura & Niemeier, 2022 |
| Homeownership | Reckien et al., 2018; Mitchell & Chakraborty, 2014; Hsu et al., 2021; Renteria et al., 2021; Zarghamfard et al., 2023 |
| Employment Rate | Reckien et al., 2018; Mashhoodi, 2021; Ghasemi et al., 2018; Huang et al., 2011 |
| Migration | SÃ¡nchez et al., 2017; Smith et al., 2024; Renteria et al., 2021; Mashhoodi, 2021 |
| Human Development Index | Viezzer & Biondi, 2021 |
| Relative Poverty | Viezzer & Biondi, 2021 |
| Travel Demand and Attraction | Aderibigbe & Gumbo, 2022; Shbeeb, 2023; Dalde et al., 2025 |

Table S2. Correlation Coefficient between data from Openmeteo.com and data from the synoptic station of the Mehrabad in Tehran

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Mehrabad Synoptic Station  Data Openmeteo.com from  the location of Mehrabad Airport | | T\_max station | T\_min station | T\_mean station | Precipitation station | Wind speed station |
| tmax\_meteo | Pearson Correlation | .**961** | .717 | .889 | -.043 | .059 |
| Sig. (2-tailed) | .000 | .000 | .000 | .739 | .651 |
| N | 365 | 365 | 365 | 365 | 365 |
| tmin\_meteo | Pearson Correlation | .792 | .**816** | .877 | -.043 | -.053 |
| Sig. (2-tailed) | .000 | .000 | .000 | .741 | .680 |
| N | 365 | 365 | 365 | 365 | 365 |
| tmean\_meteo | Pearson Correlation | .932 | .825 | .**945** | -.047 | -.023 |
| Sig. (2-tailed) | .000 | .000 | .000 | .719 | .861 |
| N | 365 | 365 | 365 | 365 | 365 |
| precipitation\_meteo | Pearson Correlation | -.084 | .122 | .044 | .**983** | .057 |
| Sig. (2-tailed) | .517 | .346 | .737 | .000 | .661 |
| N | 365 | 365 | 365 | 365 | 365 |
| wind\_meteo | Pearson Correlation | .272 | .345 | .342 | .415 | .**725** |
| Sig. (2-tailed) | .033 | .006 | .006 | .001 | .000 |
| N | 365 | 365 | 365 | 365 | 365 |

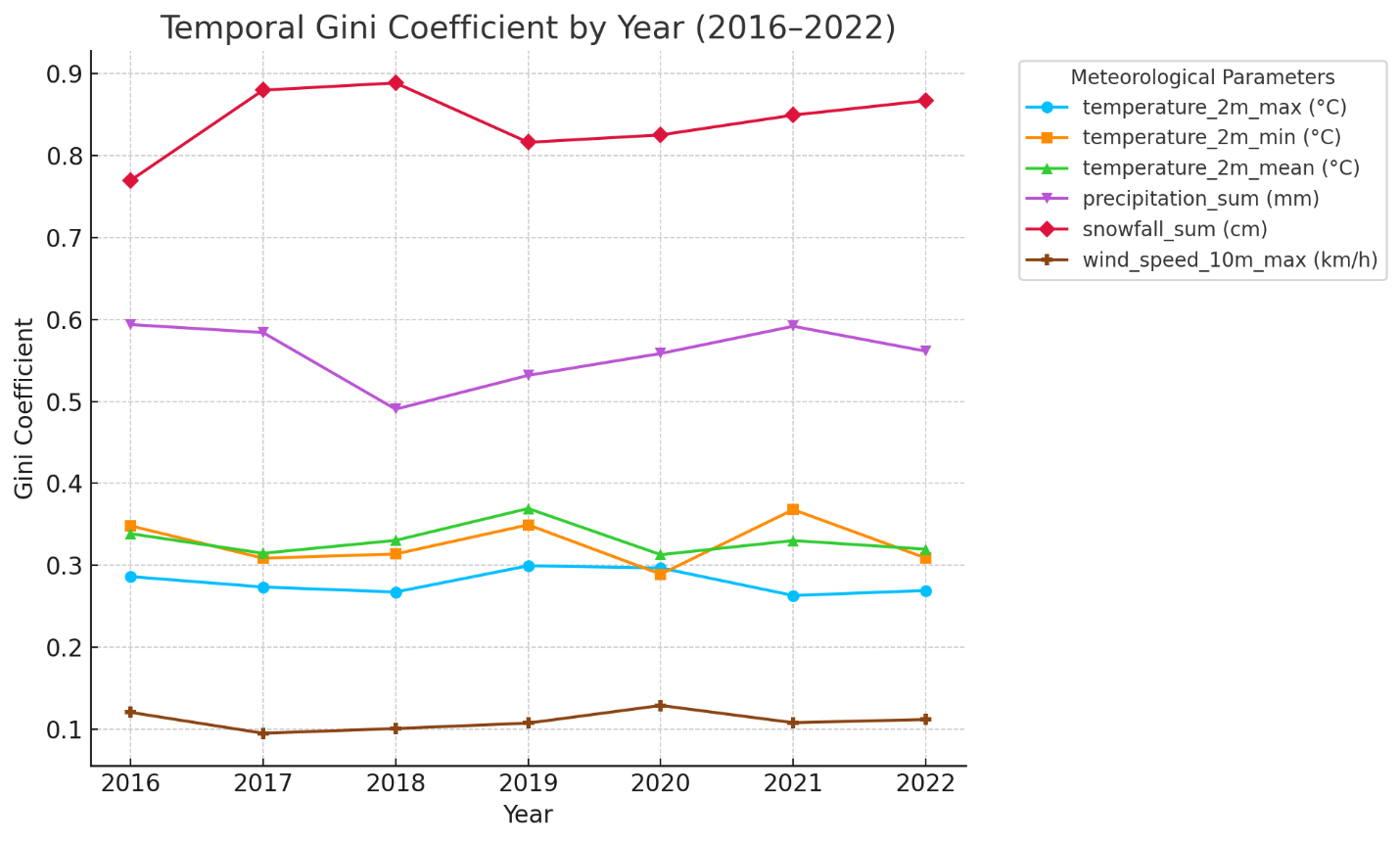


Figure S1. Temporal inequality assessment of meteorological parameters by Gini Coefficient in 12 regions of Tehran across 2016-2022

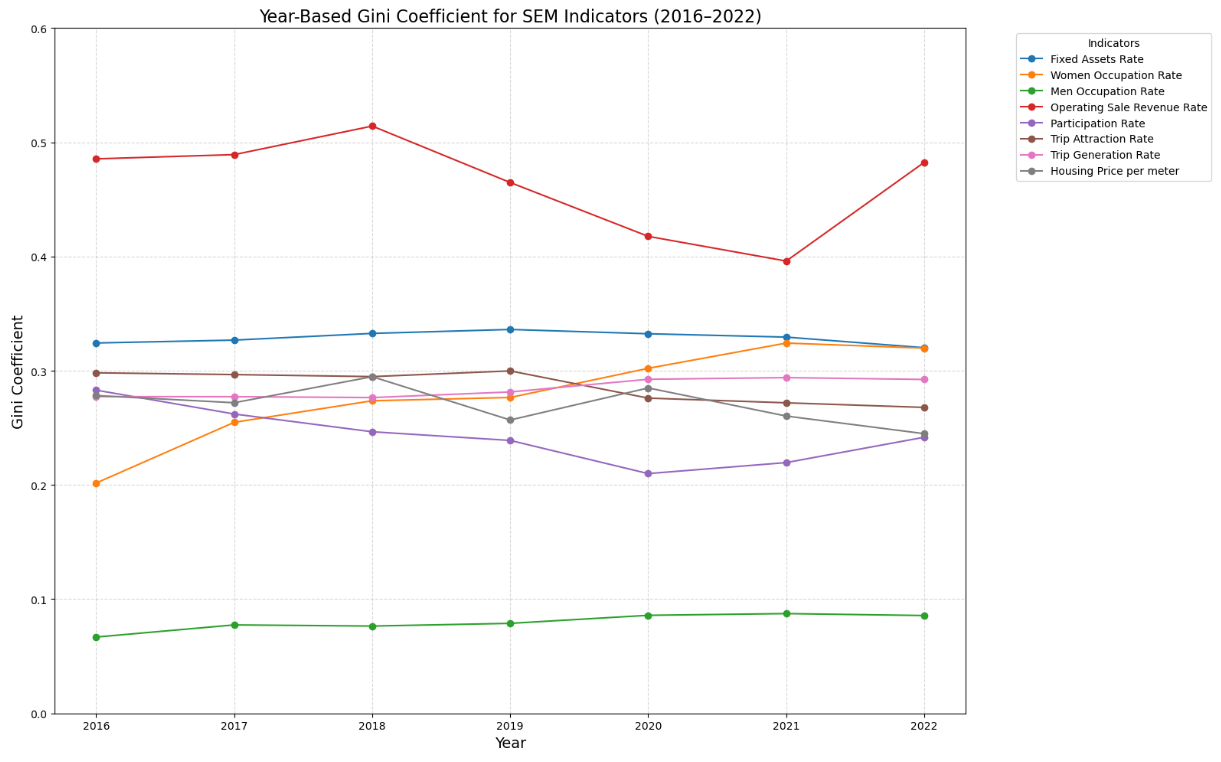


Figure S2. Temporal inequality assessment of SoEM indicators by Gini Coefficient in 12 regions of Tehran across 2016-2022

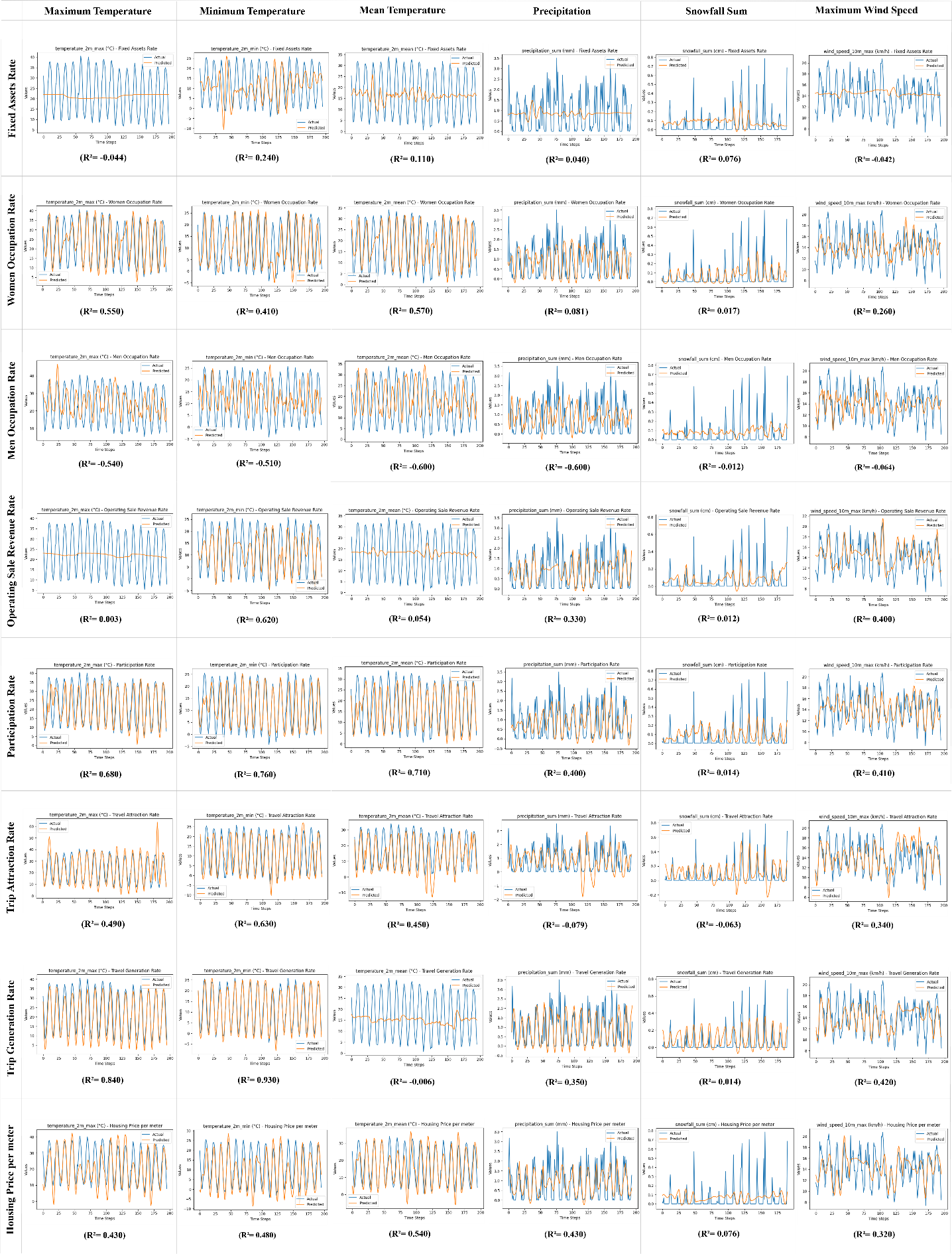


Figure 11. The Prediction trends for each meteorological parameter by each SoEM indicator using the LSTM model

### 4-2-1- Comparison between LSTM and RFR models

The general differences between the two ML models that are considered to be utilized in this research are as follows.

Table 1. Comparison between LSTM and RFR models

|  |  |  |  |
| --- | --- | --- | --- |
| **Aspect** | **LSTM** | **Random Forest** | **Reference** |
| **Type** | Deep learning-based sequence model | Ensemble learning-based decision trees | Pan, 2024 |
| **Data Type** | Sequential (time series, text) | Tabular (independent samples) | Wang, 2024 |
| **Interpretability** | Low (black-box model) | High (feature importance can be quantified) | Ali, 2022 |
| **Feature Handling** | Automatically learns features from sequences | Requires manual feature extraction | Wu, 2023 |
| **Performance** | Superior for temporal dependencies | Superior for non-sequential, structured data | Wang, 2024 |

### Data preprocessing

In order to start the analysis both datasets must be at the same time scale. Due to encountering SEM data collection issues mentioned in the data collection segment, all seven years of meteorological parameters data were rescaled to averaged monthly data. In this regard, all the indicators’ data are on a monthly scale and ready to be analyzed.

### Multivariate linear regression

Before applying the ML method, multivariate linear regression was used to have an initial understanding of the type of relationship between the datasets. The results showed that the R2 results were not suitable for a linear relationship. For this reason, the ML methods of RFR and Long Short-Term Memory (LSTM) were chosen to be conducted respectively.

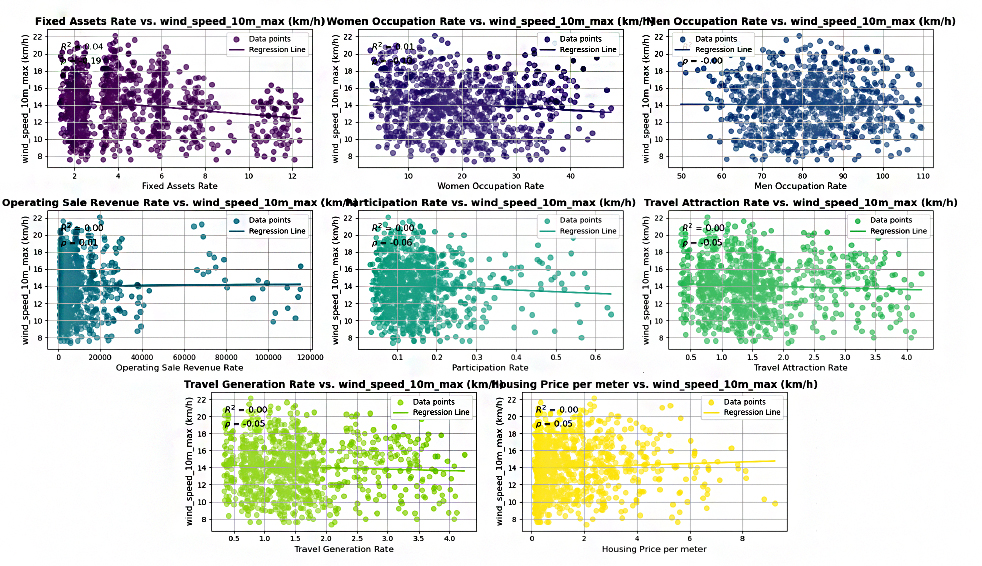


Figure 6. Multivariate Linear Regression analysis (wind speed with SEM indicators)

### The Random Forest Regression (RFR)

Random Forest is a machine learning algorithm used for classification and regression. This method consists of a collection of decision trees that are trained randomly, and their results are combined to achieve higher accuracy (Pan, 2024). Random Forest helps reduce the likelihood of overfitting by using a random subset of data and features for each tree. Due to its efficiency and versatility in handling complex problems, this algorithm is extensively applied in real-world scenarios (Ali, 2022).

Owing to the type of data involved in this study and also by reviewing similar research in this realm, For the first time, Random Forest Regression has been selected to find a reliable relation between two current datasets. After multiple tries on RFR, the final results showed low R2 in every calculation and integration with other ML models such as ARIMA and XGboost. Figure 8 indicates the average R2 scores for each month during 2016-2022.

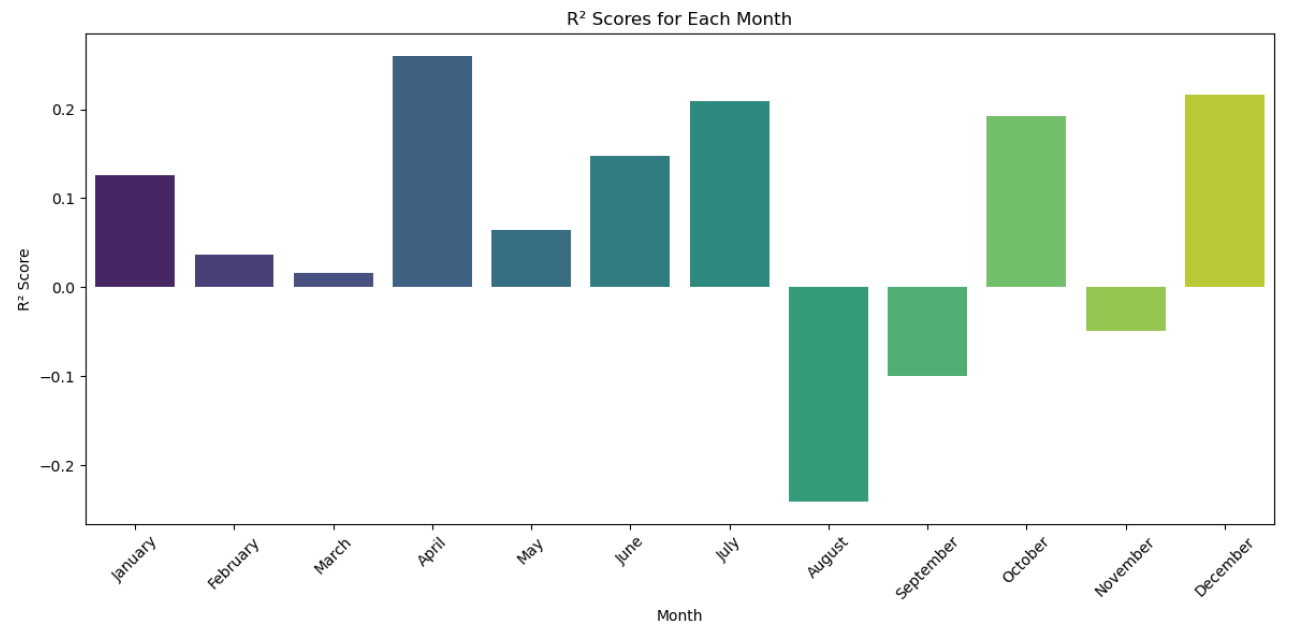


Figure 7. R2 scores for monthly RFR calculation

### 4.2.5.1. Step by step with the LSTM implementation in Python

In this study, an LSTM model was constructed based on the characteristics of univariate time-series data with limited sample points and the simple RNN design principle. Figure 9 shows the total framework based on a typical four-layer LSTM structure, which involves four main modules of data preprocessing, network training, network prediction, and model evaluation. Relevant steps of LSTM time-series prediction can be described in 9 steps:

**Step1: Loading the Data**

The data is read from a CSV file using the Pandas library in Python. It contains monthly SEM indicators and corresponding meteorological parameters for 12 regions of Tehran.

**Step 2: Data Normalization**

As mentioned before, MinMaxScaler is used to normalize the values of both predictors (x = SEM indicators) and target (y = meteorological parameters) to a range of [0, 1], improving the model's performance and training stability.

**Step 3: Outlier Removal**

Outliers are identified in the target variable (y) using the 3-standard deviation rule (Brownlee, 2020). Data points outside this range are excluded to improve model reliability. Both datasets are filtered to remove the corresponding rows containing outliers.

**Step 4: Defining Sequence Length:**

Defining sequence length identifies the optimal number of previous time steps required for the LSTM model to effectively predict meteorological parameters based on SEM indicators (Goodfellow et al., 2016). When preparing the data for LSTM input, the features are reshaped into a 3D tensor with the dimensions:



Where:

* *Xseq* is the sequence of predictors used for LSTM input.
* *seq* is the sequence length.
* *t* is the current time step.
* *X* is the input feature matrix.

This process means that each *Xseq*​ is composed of *t-seq+1* to (*t*) time steps of data (Goodfellow et al., 2016).

The corresponding target value (*y*) for each sequence is (*yt*.).This reshaping ensures that the LSTM model gets a sliding window of *seq* consecutive time steps as input to predict the value at the current time step (*t*).

To find the best sequence length, ideally, calculation of the R² score for different sequence lengths and selection of the one that maximizes the score is highly appropriate. The process involves:

1. Testing different sequence lengths (e.g., 10, 14, 21, 28, etc.).
2. Preparing the data using the respective sequence length.
3. Training the LSTM model and evaluating the R² score for each sequence length.
4. Selecting the sequence length that maximizes the R² score (Goodfellow et al., 2016).

A sequence length of 28 days is chosen, meaning the model will use 28-time steps (e.g., daily SEM data for 28 days) to predict the meteorological parameter on the 29th day.

**Step 5: Splitting Data**

The sequence data is split into training (80%) and testing (20%) sets to evaluate the model's performance on predicted data.

**Step 6: Building the LSTM Model**

The model is defined using Sequential with an LSTM layer (50 units) to capture temporal dependencies. A dense layer with 1 unit for predicting the target variable.

The model is compiled using the Adam optimizer and Mean Squared Error (MSE) loss function, which are suitable for regression tasks.

**Step 7: Model Training**

The model is trained on the training data for 100 epochs with a batch size of 16.

The verbose=0 option suppresses detailed output during training.

**Step 8: Model Evaluation**

The model predicts meteorological parameters on the testing set. Predicted values are scaled back to the original range using the inverse of MinMaxScaler. The R² score is calculated to measure how well the predictions match the actual values. An R² close to 1 indicates a good fit.

**Step 9: Feature Importance**

This function calculates the importance of each SEM indicator by permuting its values and observing the drop in the R² score. The larger the drop, the more important the feature. After training the LSTM model, the feature importance is calculated for each meteorological parameter.

Finally, the shuffle function from Sklearn is used to randomly shuffle the values of a feature during the FI process (Chaudhari, 2025).

# Research Limitations

This study has been undertaken under severe data unavailability. The authors initially intended to conduct the study at the selected urban neighborhoods scale or urban sub-districts. However, due to the lack of data at smaller scales than municipal districts, this approach was eventually revised to focus on the districts of Tehran municipality. Additionally, the limited number of meteorological stations and the sparsity of their data necessitated aggregating these districts into larger regions. In addition, the existing meteorological data in Tehran was limited, as some synoptic stations lacked data or had incomplete records. For this reason, meteorological parameters data were accessed thanks to Openmeteo’s website. Also, Socioeconomic data were not available on a short-term (daily) basis, and the study had to rely on existing monthly reports. Moreover, they were accessible in limited time intervals until 2022. The machine learning model in this study relied on the available dataset for data analytics. More precise data would have improved the accuracy of predictions. The analysis was conducted at a large urban scale. If a finer-grained geographic scale (such as Tehran’s 116 urban sub-districts) had been available, spatial analyses could have been more detailed.

Data sparsity not only influenced the spatial scale of the research but also affected the selection of indicators. Consequently, the authors were constrained to utilize only the data and select the most relevant indicators that were officially published by municipal authorities.

At first glance, the choice of a machine learning model such as LSTM may appear unconventional for this type of research. However, after a thorough evaluation and testing of various nonlinear statistical and mathematical models, the use of a time-based machine learning model was deemed appropriate. Moreover, although the authors were interested in extending the study period beyond the selected seven years, the absence of sufficient SoEM indicators data across longer timeframes and the lack of daily data ultimately limited the analysis to seven years using monthly data. The authors believe that, in the case of daily data availability, the association between SoEM and meteorological data would result in more accurate and more precise outcomes.