Aviation Weather Forecasting Using METAR Data

Anamaria Loznianu

Overview

Approach

Weather forecasting? More like trying to guess what Mother Nature's cooking up next, and spoiler alert: it's not a walk in the park. 🌪️🤷‍♂️

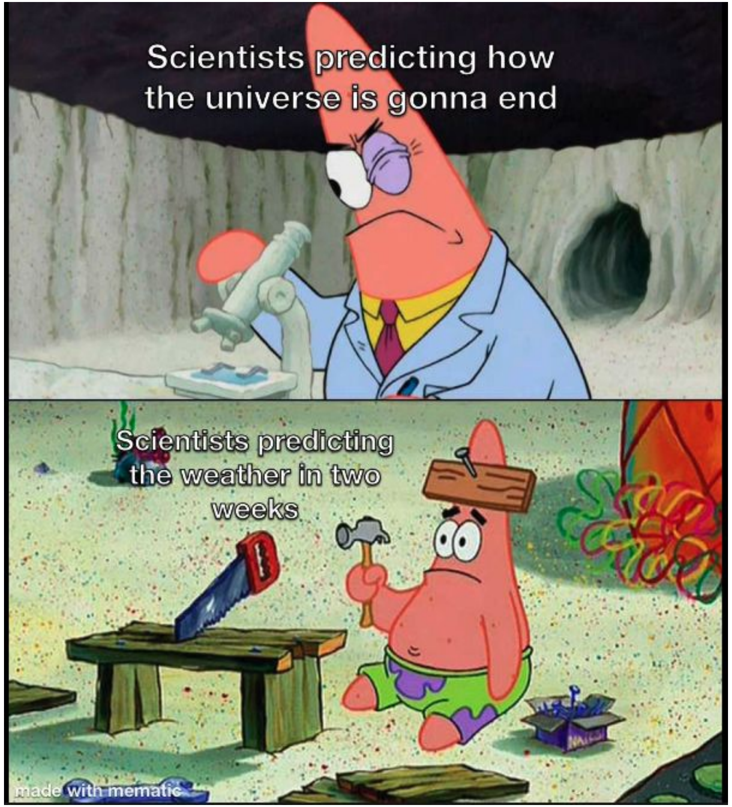
To predict the next METAR, I followed a structured approach that involved several key steps. Initially, I collected METAR data, which was then **decoded** to convert the specialized meteorological codes into a format suitable for analysis. This was followed by a **data cleaning** phase to address any errors or missing values, ensuring the data was reliable for further analysis.

The cleaned data was processed to facilitate the examination of **general trends** and **patterns** in weather variables such as wind speed, gust, temperature, air pressure, sea level pressure, visibility, and dew point. The process also included plotting and visualizing the data to better understand these patterns over time.

Using the insights gained from the data analysis, I applied predictive models to forecast specific weather values such as **wind speed, gust, temperature, air pressure, sea level pressure, visibility, dew point and more**. Finally, based on the predicted values, a new METAR code was generated(based on an existing one), providing a summary of the expected meteorological conditions.

Spoiler alert: If you're wondering about the accuracy of these predictions, let's just say, our friend SpongeBob here, net in hand, might give us an idea 🍍🎣

However, there's a twist in the tale. I embarked on a new approach towards the end - didn't quite wrap it up, but it's shaping up to be quite promising. 🌟



Data preprocessing

The dataset comprises 89,579 records across three columns. The following steps were undertaken to enhance its utility and richness:

1. **Metar Data Extraction**: Used a library to parse Metar codes from the dataset, extracting comprehensive weather-related information.
2. **Data Consolidation**: Compiled the extracted details into a CSV file for ongoing use and reference, expanding the dataset to include around 40 columns. These additional fields encompass a wide range of weather metrics, including wind direction, visibility, temperature, dew point, atmospheric pressure, and sky conditions.
3. **Data Cleaning Process**:
    a. **Column Pruning**: Removed columns exhibiting significant data sparsity, specifically those with more than 90% missing values. This affected columns such as correction, wind_direction_from, wind_direction_to, recent_weather, and various ice accretion metrics, as they offered limited analytical or predictive value. The result is available here.
    b. **Missing Value Imputation**: For fields with fewer missing entries, like temperature, dew_point, pressure, and sea_level_pressure, filled gaps using the zero value to mitigate outlier effects.
    c. **Categorical Data Handling**: Applied encoding techniques to the sky_conditions column, transforming categorical entries into a machine-readable format.
    d. **Unit Standardization**: Harmonized measurement units across the dataset (e.g., standardizing temperature to Celsius and wind speed to a uniform unit), and converted applicable columns to numeric formats for analytical consistency.

Following these adjustments, the dataset was refined to approximately 20 columns, streamlining it for effective training and analysis of prevailing weather trends.

Note: Deciphering the METAR code proved straightforward, yet crafting a METAR code from specific weather forecasts (such as temperature, wind speed, gusts, etc.) presented a considerable challenge. Consequently, I shifted my approach and opted to revise the predicted values using an existing METAR code as a foundation.

Data preprocessing

	airport_id	date	metar
0	KMIA	2014-01-01	KMIA 010053Z 04009G15KT 10SM OVC025 23/18 A302...
1	KMIA	2014-01-01	KMIA 010153Z 04009G15KT 10SM OVC023 23/18 A302...
2	KMIA	2014-01-01	KMIA 010253Z 05009G19KT 10SM -RA OVC025 23/18 ...
3	KMIA	2014-01-01	KMIA 010353Z 07010KT 10SM OVC028 23/17 A3021 R...
4	KMIA	2014-01-01	KMIA 010453Z 07010KT 10SM OVC028 24/18 A3019 R...

```
['code',
 'correction',
 'station',
 'time',
 'date',
 'wind_direction_degrees',
 'wind_speed_knots',
 'wind_gust_knots',
 'wind_direction_from',
 'wind_direction_to',
 'visibility_distance_miles',
 'visibility_direction',
 'maximum_visibility',
 'maximum_visibility_direction',
 'temperature_celsius',
 'dew_point_celsius',
 'pressure_inches',
 'runway',
 'weather',
 'recent_weather',
 'sky_conditions',
 'runways_windshear',
 'direction_peak_wind_speed_last_1h',
 'peak_wind_speed_last_1h',
 'peak_wind_time',
 'wind_shift_time',
 'max_temp_last_6hr_celsius',
 'min_temp_last_6hr_celsius',
 'max_temp_last_24hr_celsius',
 'min_temp_last_24hr_celsius',
 'sea_level_pressure_millibars',
 'precipitation_last_1hr_inches',
 'precipitation_last_3h_inches',
 'precipitation_last_6h_inches',
 'precipitation_last_24h_inches',
 'snowdepth_distance',
 'ice_accretion_last_1hr',
 'ice_accretion_last_3hr',
 'ice_accretion_last_6hr']
```

```
['code',
 'station',
 'date',
 'wind_direction_degrees',
 'wind_speed_knots',
 'wind_gust_knots',
 'visibility_distance_miles',
 'temperature_celsius',
 'dew_point_celsius',
 'pressure_inches',
 'sky_conditions',
 'max_temp_last_6hr_celsius',
 'min_temp_last_6hr_celsius',
 'sea_level_pressure_millibars',
 'precipitation_last_1hr_inches',
 'hour',
 'day_of_week',
 'month',
 'temp_7d_rolling_avg',
 'wind_gust_7d_rolling_avg']
```

Feature engineering

1. **Time-Based Features**: Extract features from the `time` column, such as hour of the day, day of the week, and month, which could capture daily and seasonal patterns in weather.

- A special case that needed to be treated here was the date and time
- The date was used from the date column
- The time was used from the decoded Metar data

2. **Rolling Averages**: For parameters like temperature, wind speed, and pressure, calculate rolling averages to smooth out short-term fluctuations and highlight longer-term trends.

code	station	date	wind_direction_degrees	wind_speed_knots	wind_gust_knots	visibility_distance_miles	temperature_celsius	dew_point_celsius	pressure_inches	sky_conditio
KMIA 010053Z 09G15KT 10SM OVC025 18 A302...	KMIA	2014-01-01 00:53:00	40	9	15	10	23.0	18.0	30.0	overcast at 25 f
KMIA 010153Z 09G15KT 10SM OVC023 18 A302...	KMIA	2014-01-01 01:53:00	40	9	15	10	23.0	18.0	30.0	overcast at 23 f
KMIA 010253Z 09G19KT 0SM -RA OVC025 23/18 ...	KMIA	2014-01-01 02:53:00	50	9	19	10	22.0	17.0	30.0	overcast at 25 f
KMIA 010353Z 07010KT 10SM OVC028 17 A3021 R...	KMIA	2014-01-01 03:53:00	70	10	0	10	23.0	17.0	30.0	overcast at 28 f
KMIA 010453Z 07010KT 10SM OVC028	KMIA	2014-01-01 04:53:00	70	10	0	10	23.0	17.0	30.0	overcast at 28 f

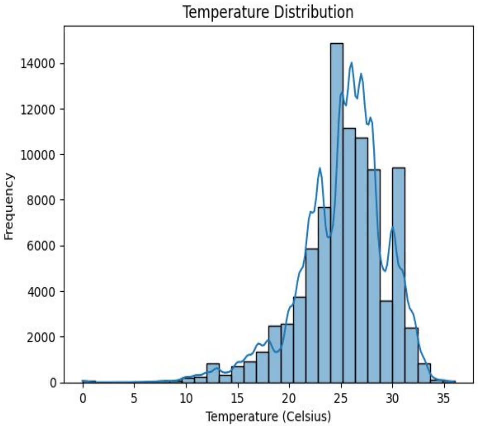
Exploratory data analysis

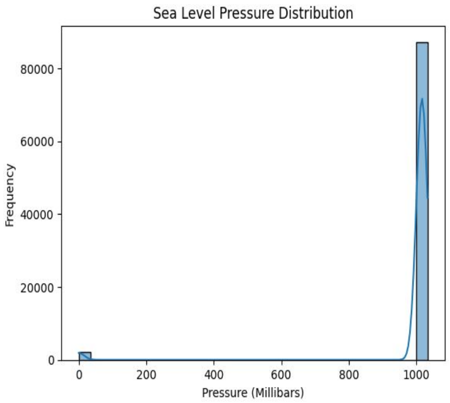
- **Temperature Distribution**: The distribution of temperatures shows a somewhat normal distribution, indicating a range of temperatures with a concentration around the mean. This suggests variability in temperature at the Miami airport, which is expected given day-night cycles and seasonal changes.

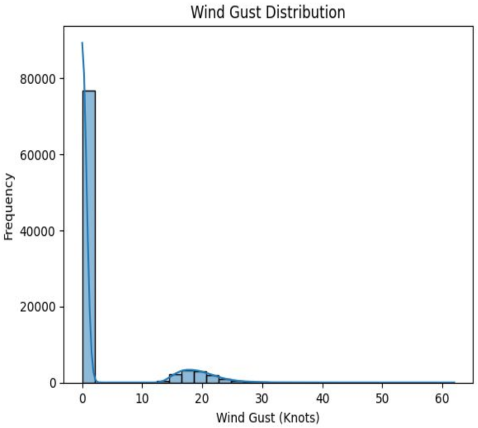
- **Wind Speed Distribution**: The wind speed distribution is skewed towards lower values. This indicates that high wind speeds are less frequent.

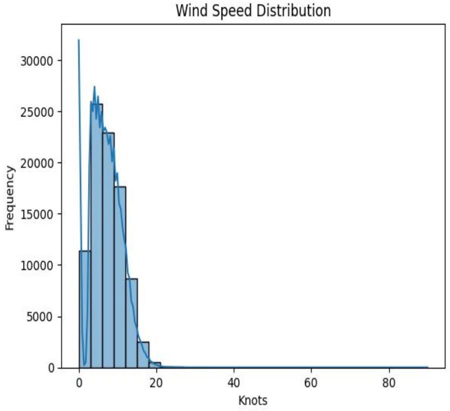
- **Wind Gust Distribution**: The wind gust distribution is skewed towards lower values, with a long tail towards higher wind gusts. This indicates that high wind gusts are less frequent but can reach significant speeds.

- **Sea Level Pressure Distribution**: The distribution of sea level pressure is relatively normal, with a slight skew. This indicates a stable range of atmospheric pressure values with some fluctuations.

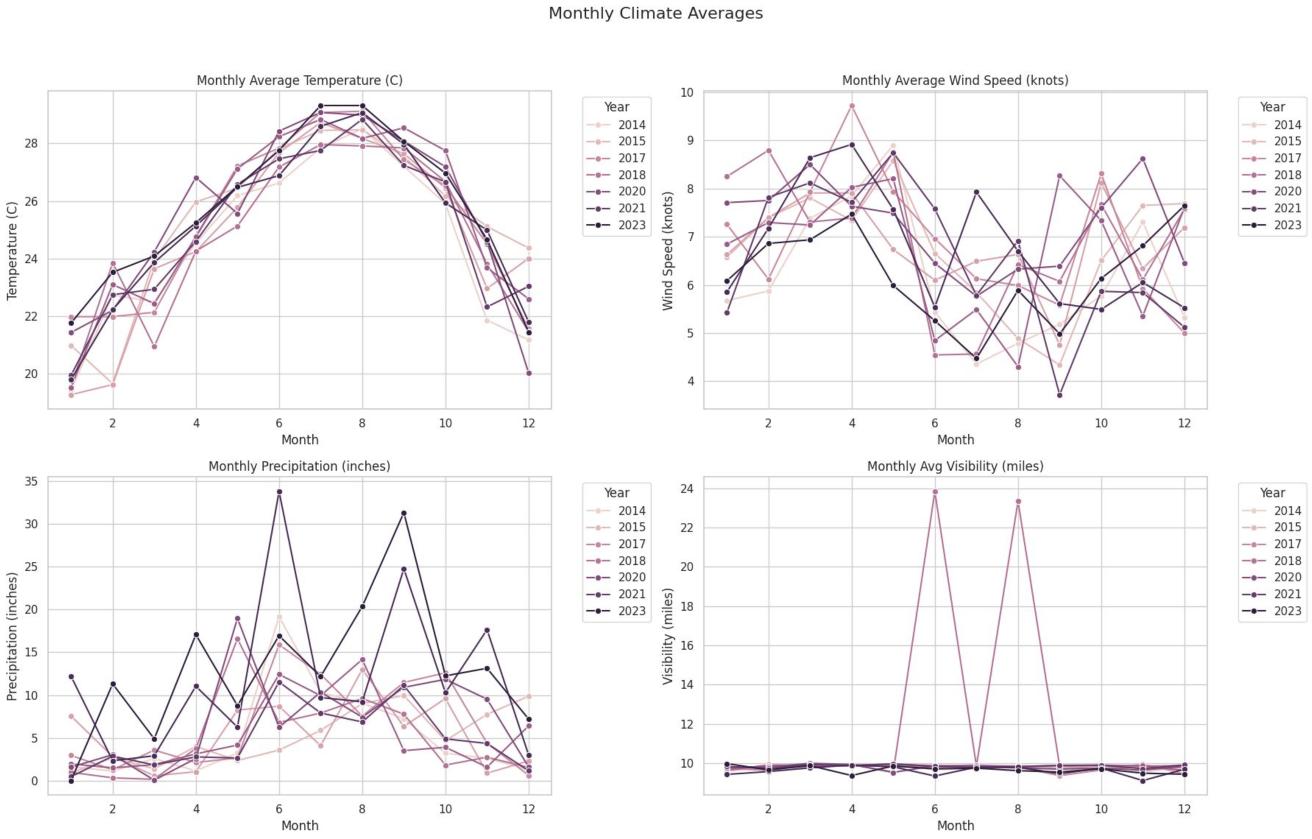








General trends



General trends

1. **Monthly Average Temperature (Celsius)**: The plot indicates the seasonal variations in temperature, with peaks during the warmer months and troughs in the cooler months. The variation across different years suggests changes in the average monthly temperatures, reflecting possible year-to-year climatic shifts or anomalies.

2. **Monthly Average Wind Speed (knots)**: The wind speed plot shows relatively consistent patterns across months, with slight variations that might indicate seasonal changes in wind activity. Some months show significant differences across years, which could be influenced by specific weather events or broader climatic trends.

3. **Monthly Precipitation (inches)**: Precipitation levels vary significantly across months and years, indicating the occurrence of wetter and drier seasons. The variance across years for the same month could reflect irregular weather patterns, such as the presence of storms or dry spells.

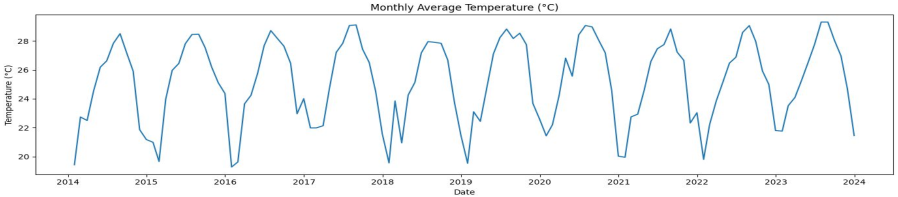
4. **Monthly Avg Visibility (miles)**: Visibility tends to be fairly consistent, with minor variations that might be influenced by seasonal weather conditions such as fog, rain, or pollution levels. The overall high visibility across most months suggests generally clear conditions, with occasional decreases that could be associated with specific weather events.

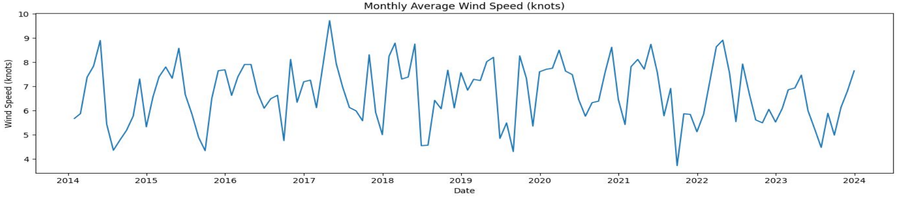
General trends

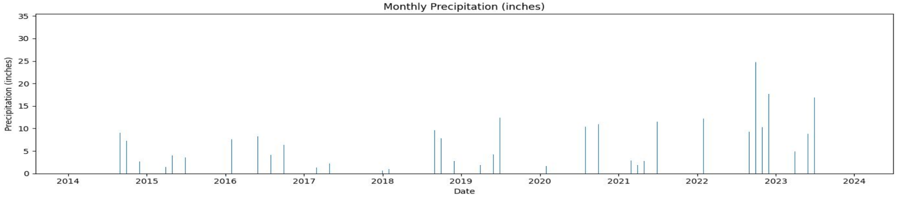
We can see in these charts, the general Miami's weather that is generally characterized by its tropical monsoon climate, offering warm, humid summers and mild, dry winters. This distinctive climate pattern is largely influenced by its geographical location along the southeastern coast of Florida, near the Gulf of Mexico and the Atlantic Ocean. Here's a breakdown of Miami's weather characteristics and influences:

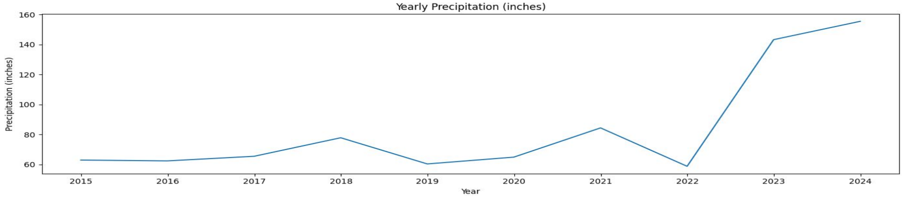
Weather Characteristics:

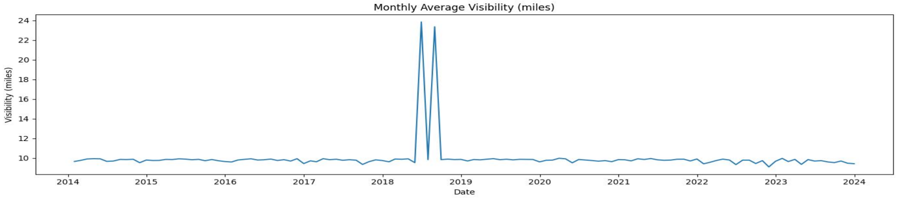
- **Warm Summers**: Miami experiences hot and humid summers with temperatures often reaching the high 80s to low 90s°F (around 30-35°C). The humidity can make the heat feel more intense, and afternoon thunderstorms are common due to the high moisture levels.
- **Mild Winters**: Winters are mild and dry, with temperatures typically ranging from the mid-60s to high 70s°F (around 18-25°C). It's the peak tourist season due to the comfortable weather.
- **Rainfall**: The city sees a significant amount of rainfall, primarily during the summer months from May through October, which is the wet season. The dry season spans from November through April, with less frequent rain showers.
- **Hurricanes**: Being in a tropical zone, Miami is susceptible to hurricanes and tropical storms, particularly from June to November. These storms can significantly impact weather patterns and conditions.









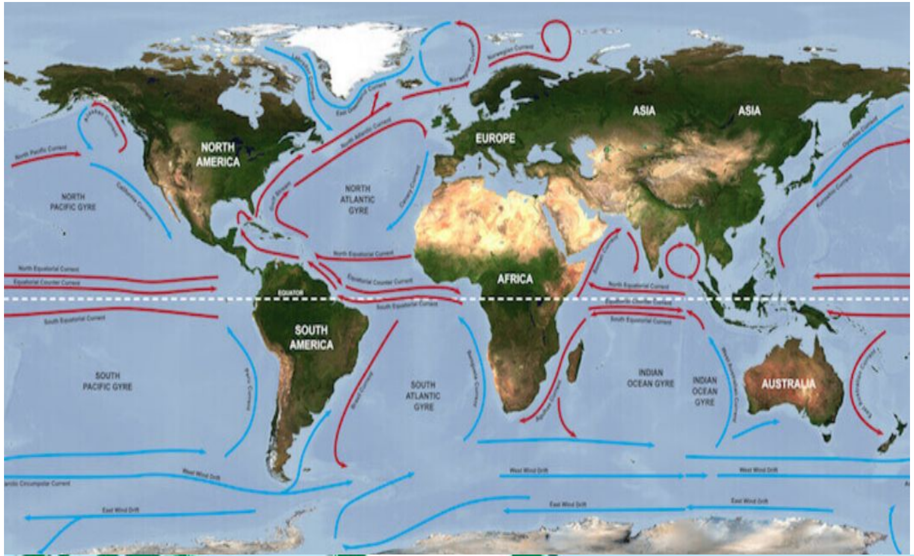


[General trends]

Influential Factors:

- **Geography**: Miami's proximity to the Gulf of Mexico and the Atlantic Ocean moderates its temperatures and contributes to its high humidity levels. The sea breezes from these water bodies can offer some relief from the heat during summer.
- **Gulf Stream**: The warm currents of the Gulf Stream, which flow near the eastern coast of Florida, play a crucial role in keeping Miami's winter temperatures milder compared to other parts of the United States.
- **Topography**: Miami's flat topography and low elevation make it prone to flooding, especially during the rainy season and in the event of storm surges from hurricanes.
- **Latitude**: Its tropical latitude ensures that Miami receives ample sunlight year-round, contributing to its warm temperatures.

Overall, Miami's weather is greatly appreciated for its warm, sunny days, especially during the winter months, making it a popular destination for tourists seeking a warm climate. However, its susceptibility to hurricanes and the intense summer heat and humidity are significant considerations for residents and visitors alike.



Predictions

Predicted individual aspects of the weather using the MultiOutputRegressor, the RandomForestRegressor and LSTMs.

Predicted values:

1. Wind direction degrees
2. Wind speed knots
3. Wind gust knots
4. Visibility distance miles
5. Temperature celsius
6. Dew point celsius
7. Pressure inches
8. Sea level pressure millibars
9. Max temperature(last 6 hr)
10. Min temperature (last 6hr)

Test Metar(The metar the models aim to predict)

```
Metar value:  KMIA 312353Z 00000KT 10SM CLR 18/09 A3022 RMK AO2 SLP234 T01830094 10217 20183 53012
```

Prediction using MultiOutputRegressor

The comparison between expected and predicted values reveals significant discrepancies, most notably in temperature predictions. The temperature was forecasted to be more than 10 degrees higher than actual conditions, indicating a substantial error. In this context, the use of a multi-output regressor does not significantly improve the situation.

```
=======-------- MultiOutputRegressor prediction =======--------

Temperature (Celsius): 18.0, Predicted Temperature (Celsius): 29.762693405151367,
Visibility (miles): 10, Predicted Visibility (miles): 9.87829303741455,
Dew Point (Celsius): 9.0, Predicted Dew Point (Celsius): 19.97425651550293,
Wind speed (knots): 0, Predicted Wind Speed (knots): 5.420334815979004,
Wind Gusts (knots): 0, Predicted Wind Gusts (knots): 1.0955545902252197,
Pressure (inches): 30.0, Predicted Pressure (inches): 29.192964553833008,
Sea Level Pressure (millibars): 1023, Predicted Sea Level Pressure (millibars): 976.6316528320312,
Max Temperature Last 6 Hours (Celsius): 21.0, Predicted Max Temperature Last 6 Hours (Celsius): 1.3785496950149536,
Min Temperature Last 6 Hours (Celsius): 18.0, Predicted Min Temperature Last 6 Hours (Celsius): 1.192979097366333

Predicted metar using MultiOutputRegressor KMIA 312353Z 00000KT 10SM CLR 18/09 A2919  RMK AO2 SLP234 T01830094 10217 20183 53012 10001 20001
```

Prediction using RandomForestRegressor

The RandomForestRegressor gives much more accurate results being very close to the reality with a very low nmse.

```
=======-------- RandomForestRegressor prediction =======--------

Temperature (Celsius): 18.0, Predicted Temperature (Celsius): 17.81,
Visibility (miles): 10, Predicted Visibility (miles): 10.0,
Dew Point (Celsius): 9.0, Predicted Dew Point (Celsius): 9.67,
Wind speed (knots): 0, Predicted Wind Speed (knots): 0.65,
Wind Gusts (knots): 0, Predicted Wind Gusts (knots): 0.0,
Pressure (inches): 30.0, Predicted Pressure (inches): 30.0,
Sea Level Pressure (millibars): 1023, Predicted Sea Level Pressure (millibars): 1021.49,
Max Temperature Last 6 Hours (Celsius): 21.0, Predicted Max Temperature Last 6 Hours (Celsius): 21.06,
Min Temperature Last 6 Hours (Celsius): 18.0, Predicted Min Temperature Last 6 Hours (Celsius): 17.53

Predicted metar using RandomForestRegressor KMIA 312353Z 00000KT 10SM CLR 18/09 A3000  RMK AO2 SLP234 T01830094 10217 20183 53012 10021 20017
```

Prediction using LSTM

The LSTM gives great results being very close to the reality with a very low nmse.

```
Epoch 1/10
2800/2800 [==============================] - 9s 3ms/step - loss: 0.0043
Epoch 2/10
2800/2800 [==============================] - 12s 4ms/step - loss: 2.5992e-05
Epoch 3/10
2800/2800 [==============================] - 8s 3ms/step - loss: 1.0715e-05
Epoch 4/10
2800/2800 [==============================] - 7s 3ms/step - loss: 5.3031e-06
Epoch 5/10
2800/2800 [==============================] - 8s 3ms/step - loss: 2.8054e-06
Epoch 6/10
2800/2800 [==============================] - 8s 3ms/step - loss: 1.8058e-06
Epoch 7/10
2800/2800 [==============================] - 7s 3ms/step - loss: 1.4998e-06
Epoch 8/10
2800/2800 [==============================] - 8s 3ms/step - loss: 1.2654e-06
Epoch 9/10
2800/2800 [==============================] - 14s 5ms/step - loss: 1.1533e-06
Epoch 10/10
2800/2800 [==============================] - 7s 3ms/step - loss: 8.9508e-07
1/1 [==============================] - 0s 187ms/step
=======-------- LSTM prediction =======--------

Temperature (Celsius): 18.0, Predicted Temperature (Celsius): 18.03445816040039,
Visibility (miles): 10, Predicted Visibility (miles): 9.21994400024414,
Dew Point (Celsius): 9.0, Predicted Dew Point (Celsius): 9.051663398742676,
Wind speed (knots): 0, Predicted Wind Speed (knots): 0.09879489243030548,
Wind Gusts (knots): 0, Predicted Wind Gusts (knots): 0.025731798261404037,
Pressure (inches): 30.0, Predicted Pressure (inches): 30.774688720703125,
Sea Level Pressure (millibars): 1023, Predicted Sea Level Pressure (millibars): 1022.4085693359375,
Max Temperature Last 6 Hours (Celsius): 21.0, Predicted Max Temperature Last 6 Hours (Celsius): 20.984243392944336,
Min Temperature Last 6 Hours (Celsius): 18.0, Predicted Min Temperature Last 6 Hours (Celsius): 17.93550682067871

Predicted metar using LSTM KMIA 312353Z 00000KT 9SM CLR 18/09 A3077  RMK AO2 SLP234 T01830094 10217 20183 53012 10020 20017
```

Real Time Weather Prediction? Hm, what a challenge

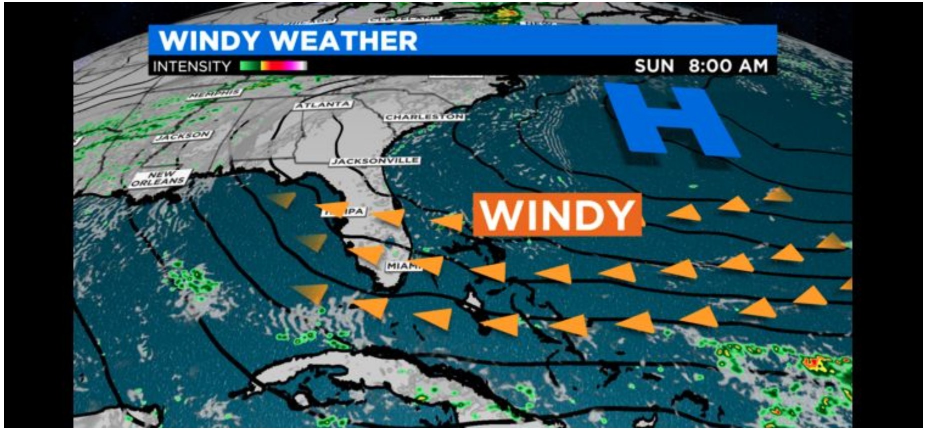
Real Time Data

Relying solely on historical data for weather predictions is great, yet not quite there. So, the burning question arises: what other sources of data could we tap into? From detailed forecasts to broad meteorological datasets, the pool of resources is expansive. 🌐🌦️

But imagine, just for a moment, if we could use data from airports surrounding Miami to gauge the imminent weather changes? 🤔💡

Visualize a scenario where a hurricane, armed with ferocious winds, is making its way from the north towards Miami. The likelihood of it impacting Miami airport and its vicinity skyrockets. 🌀🌬️

Driven by this curiosity, I embarked on a novel venture. I charted the nearest airports,and began aggregating real-time data (the most recent METAR) from each.



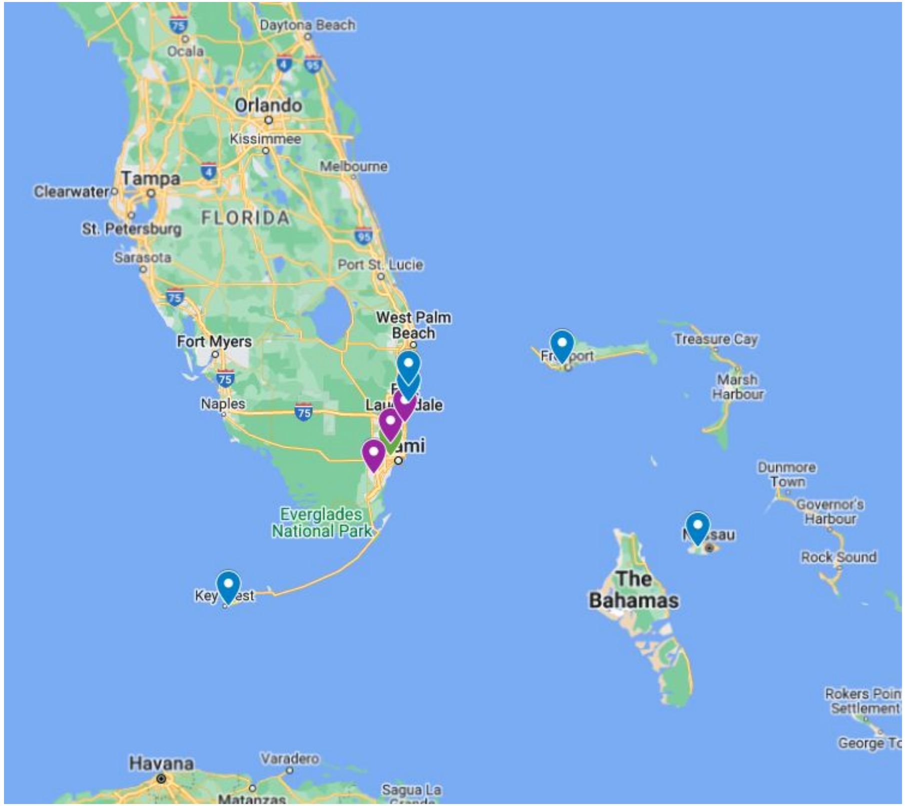
The analysis is still in it’s incipient phases, but it’s very promising.

Real Time Data

Fueled by curiosity, I leveraged real-time data from eight nearby airports(the map is available here) to refine Miami's weather predictions:

1. Opa-locka Executive Airport (KOPF): 11 miles north.
2. Miami Executive Airport (KTMB): 13 miles southwest.
3. Fort Lauderdale-Hollywood International (KFLL): 21 miles north.
4. Pompano Beach Airpark (KPMP): 36 miles north.
5. Boca Raton Airport (KBCT): 45 miles north.
6. Grand Bahama International (MYGF): 72 miles northwest.
7. Key West International (KEYW): 130 miles southwest.
8. Lynden Pindling International, Nassau (MYNN): 186 miles southwest.

In this analysis, we leverage the latest METAR data from the provided airports to evaluate if weather patterns from neighboring airports influence Miami's conditions. This assessment allows us to refine Miami's weather predictions by factoring in potential impacts from nearby areas. The most recent METAR data is sourced from here.



Real Time Data

```
Name: Fort Lauderdale-Hollywood International Airport, Code: FFLL, Distance: 21 miles, Location relative to Miami: north, Metar: station: FFLL
type: routine report, cycle 18 (automatic report)
time: Mon Feb 12 18:00:00 2024
temperature: 29.0 C
dew point: 24.0 C
wind: WSW at 6 knots
visibility: greater than 10000 meters
pressure: 1008.0 mb
sky: scattered clouds at 1500 feet
     a few cumulonimbus at 2700 feet
- CB-NE NOSIG
METAR: FFLL 121800Z 24006KT 9999 SCT015 FEW027CB 29/24 Q1008 RMK CB-NE NOSIG



Name: Miami Executive Airport, Code: KTMB, Distance: 13 miles, Location relative to Miami: southwest, Metar: station: KTMB
type: routine report, cycle 13 (automatic report)
time: Tue Feb 20 12:53:00 2024
temperature: 12.8 C
dew point: 10.6 C
wind: NNW at 6 knots
visibility: 10 miles
pressure: 1023.7 mb
sky: clear
sea-level pressure: 1023.6 mb
1-hour precipitation: 0.00in
remarks:
- Automated station (type 2)
- RAB40E49 $
METAR: KTMB 201253Z 34006KT 10SM CLR 13/11 A3023 RMK AO2 RAB40E49 SLP236 P0000 T01280106 $



Name: Opa-locka Executive Airport, Code: KOPF, Distance: 11 miles, Location relative to Miami: north, Metar: station: KOPF
type: routine report, cycle 13 (automatic report)
time: Tue Feb 20 12:53:00 2024
temperature: 12.8 C
dew point: 10.0 C
wind: NNW at 5 knots
visibility: 10 miles
pressure: 1023.7 mb
sky: clear
sea-level pressure: 1023.7 mb
remarks:
- Automated station (type 2)
METAR: KOPF 201253Z 34005KT 10SM CLR 13/10 A3023 RMK AO2 SLP237 T01280100
```

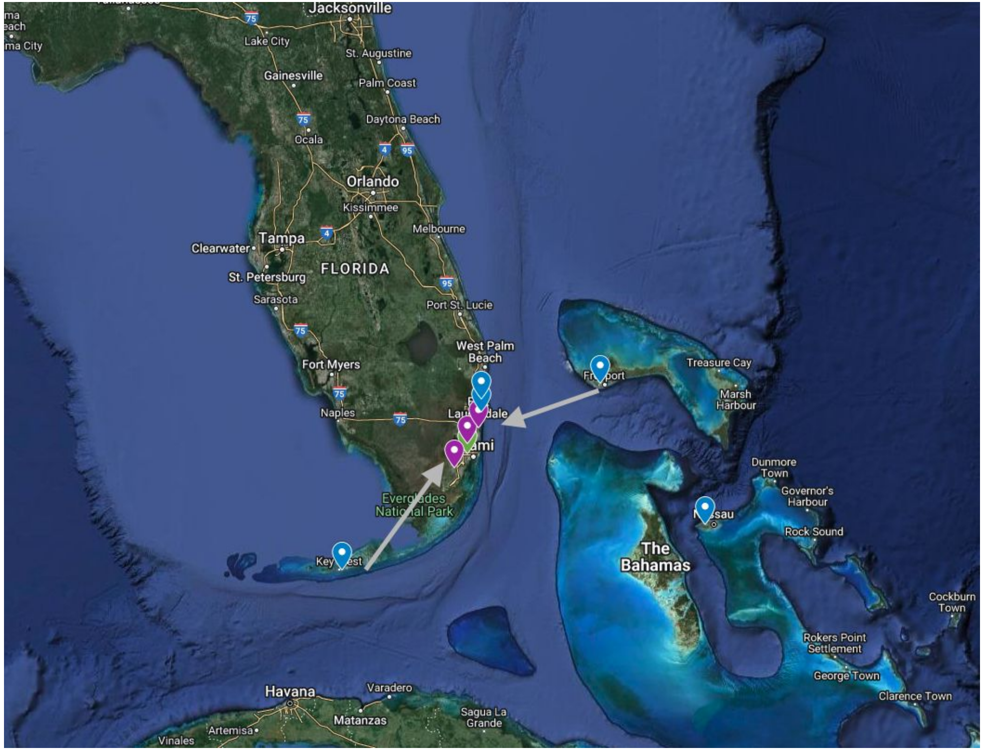
Real Time Data

This analysis was conducted through analytical, empirical, and exploratory methods and can be expanded in future phases by embedding a machine learning model.

- **Wind Direction and Speed:** We can assume that weather systems move in the direction of the wind. Therefore, airports with wind blowing towards Miami might have a more significant influence on its weather.

- **Precipitation and Weather Conditions:** When a majority of the surrounding airports experience rainfall or storms, there's a plausible chance that Miami could encounter similar weather, particularly if those airports are positioned upstream relative to the wind's direction.

- **Temperature and Pressure Trends**: Abrupt shifts in temperature or pressure at nearby airports may signal the approach of weather fronts, potentially affecting Miami's local weather patterns.



Real Time Data Prediction

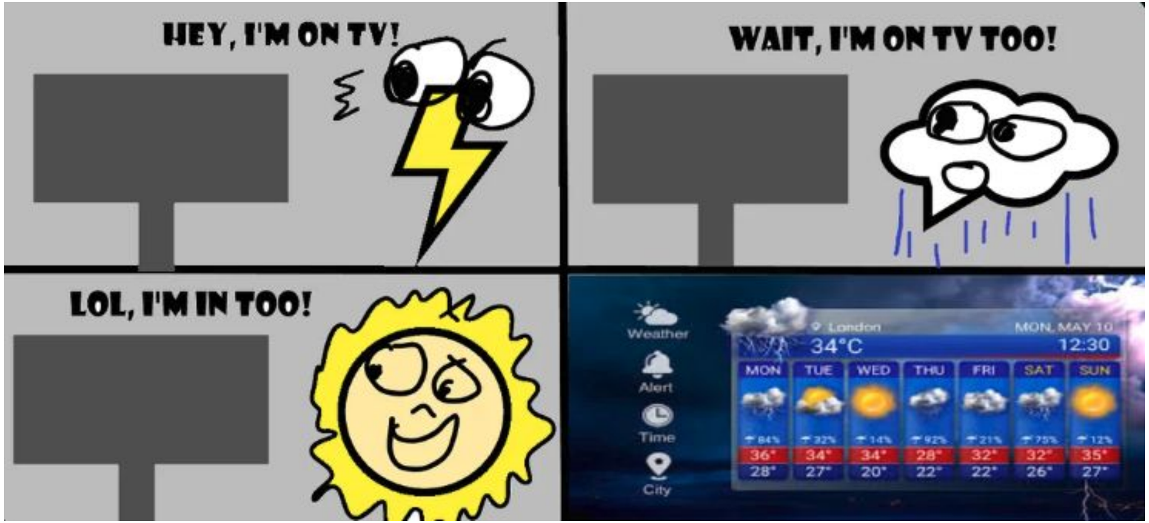
Analyzing real-time data, combined with insights from historical patterns, stands as the most effective strategy for short-term forecasting. While the method presented here remains in development and requires further refinement, it fundamentally underscores this principle.

```
-----------========== Weather changes in Miami =============---------

The pressure is falling in Miami.
The wind_speed is falling in Miami.
The visibility is stable in Miami.
The dew_point is rising in Miami.
The temperature is falling in Miami.

-----------========== Predicted Weather( 2024-02-20 15:32:40.898087 ) in Miami. =============---------

Metar Code:  KMIA 201532Z 33006KT 10SM FEW030 14/09 A3000 RMK AO2 SLP235 T01390089 $

Decoded station: KMIA
type: routine report, cycle 15 (automatic report)
time: Tue Feb 20 15:32:00 2024
temperature: 13.9 C
dew point: 8.9 C
wind: NNW at 6 knots
visibility: 10 miles
pressure: 1015.9 mb
sky: a few clouds at 3000 feet
sea-level pressure: 1023.5 mb
remarks:
- Automated station (type 2)
- $
METAR: KMIA 201532Z 33006KT 10SM FEW030 14/09 A3000 RMK AO2 SLP235 T01390089 $
```

Conclusion

The best models identified for weather variable predictions are LSTM and RandomForest but while historical data is valuable for recognizing patterns and trends, forecasting weather in the short term is most effectively achieved by leveraging current data in conjunction with the overall direction of weather systems.

Looking ahead, this analysis could be expanded by developing a model trained on the last weeks or months of data from all closeby airports. This would deepen our understanding of the mutual impact these airports have on each other's weather conditions. Data for such an extended analysis is accessible, for example [here](), providing a rich dataset for a comprehensive model training and subsequent predictive analysis.



Thank you!

For the code and additional resources, head over to github.