IBM

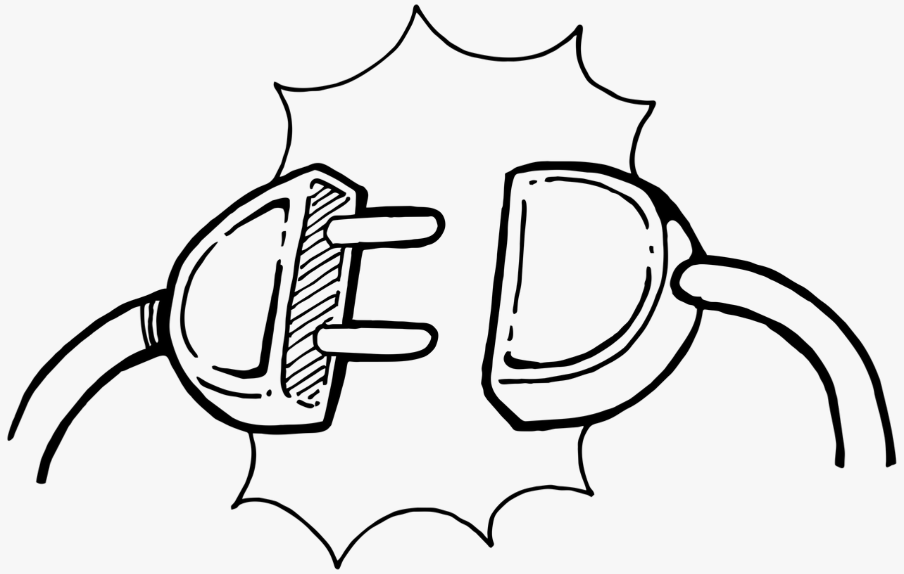
GeoIntelligent: Renewable Energy Site Selection

A Data-Driven Approach for Selecting Potential Solar & Wind Farm Locations

Samarth P	PES2UG22CS495
Sadhana Shashidhar	PES2UG22CS478
Sakshi Prakash Masand	PES2UG22CS492
Amogh Krishnamurthy	PES2UG22CS061

IBM API HACKATHON

Problem Statement



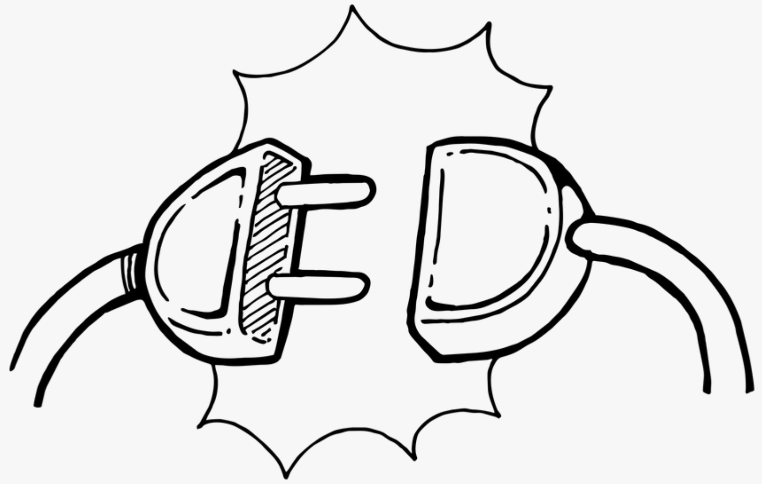
Challenges

- Solar and wind energy production heavily depends on geographic and climatic conditions.
- Poor site selection leads to low energy output, high maintenance costs, and environmental concerns.

Need for a Data-Driven Solution:

- Manual site selection methods are inefficient and lack precision.
- A systematic approach using climate, elevation, and infrastructure data can optimize energy yield.

Problem Statement



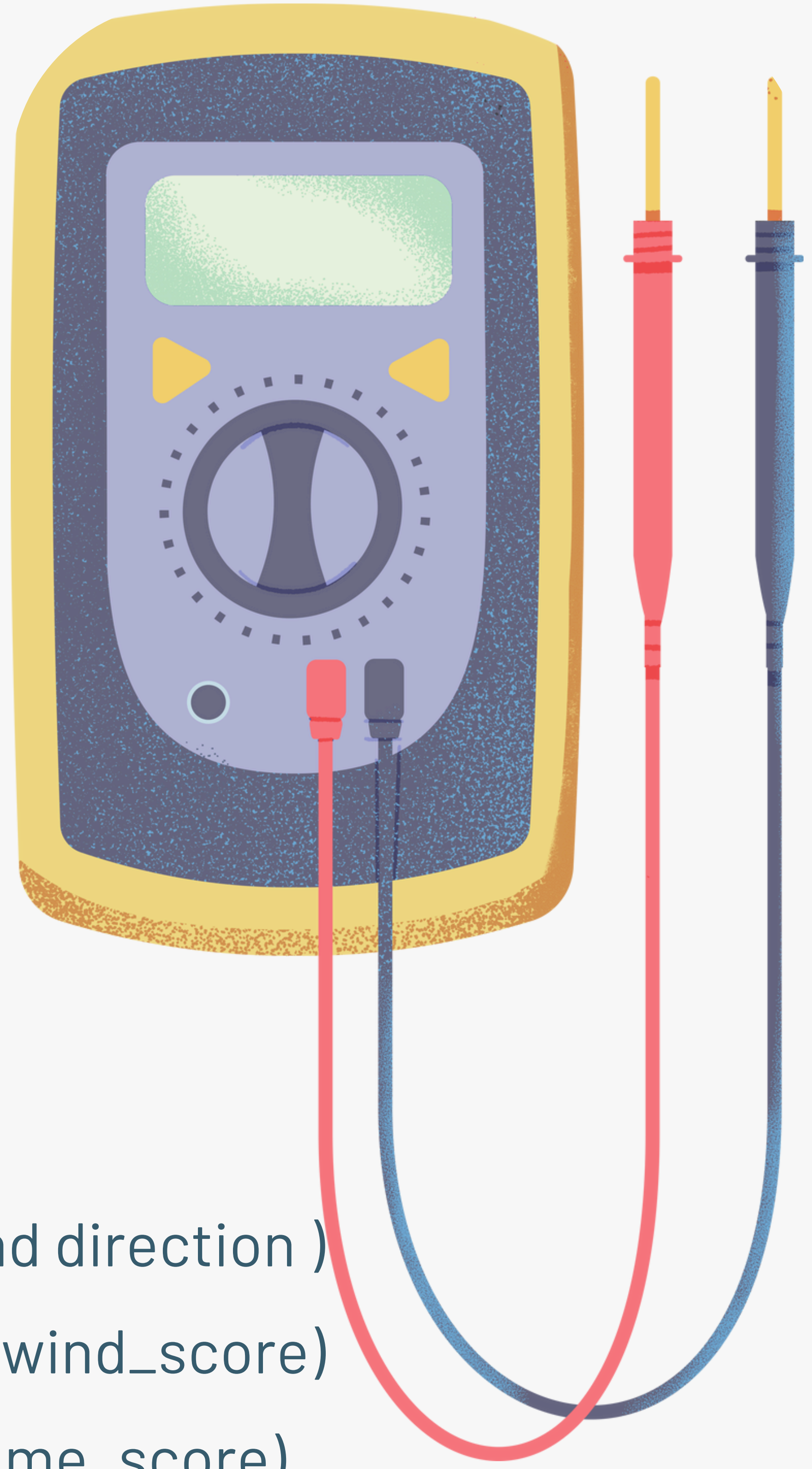
Key Features of solution using IBM's History on Demand Direct API:

1. Identify high-potential locations for solar and wind farms using weather and elevation data.
2. Reduce maintenance costs & transmission losses by selecting optimal sites.
3. Ensure environmental sustainability by avoiding deforestation and sensitive ecosystems.
4. Enable governments, energy companies, and investors to make informed decisions.

Metrics:

Metrics for Solar Farms:

- Solar Irradiance (avg_solar_irradiance, ghi_score)
- UV Index (avg_uv_index, uv_score)
- Temperature (avg_temperature, temp_score)
- Precipitation & Humidity (precip_frequency, avg_humidity, precip_score, humidity_score)
- Elevation



Metrics for Wind Farms:

- Wind Speed (avg_wind_speed, max_wind_speed, wind_speed_std,Wind stability, wind_speed_score, wind direction)
- Wind Gusts (avg_wind_gust, max_wind_gust, extreme_wind_score)
- Optimal Wind Conditions (optimal_wind_pct, optimal_time_score)
- Wind Consistency (consistency_score)
- Elevation along with Pressure

API USED: History on Demand Direct API

Synchronously query historical weather data

</> History on Demand Direct API

/direct

Synchronously query historical weather data

</> Historical Data for Analytical Tools

/r2

/r2/products

/ext

/ext/products

FAQ and support

Parameters

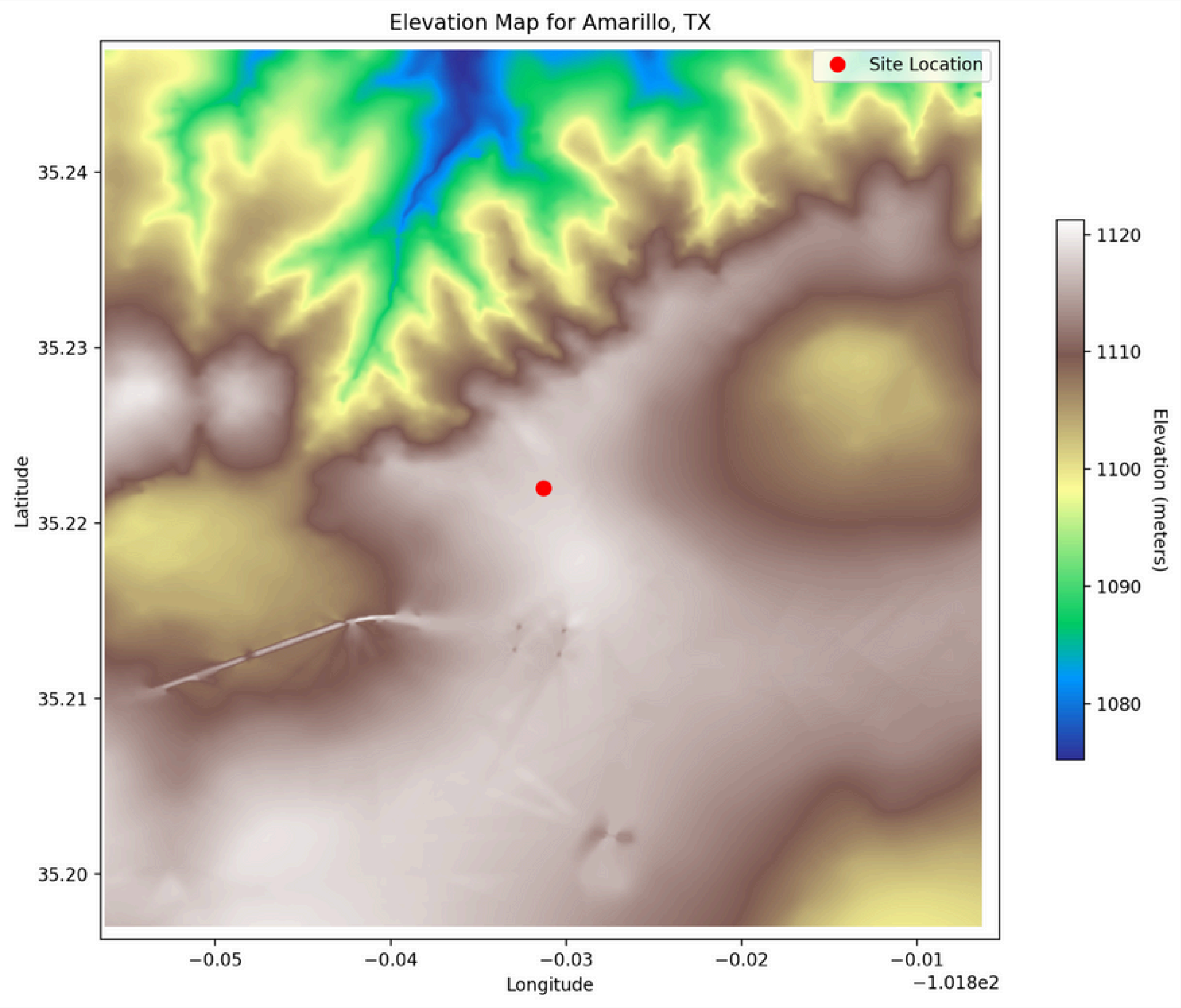
Name	Description
geocode string (query)	Latitude/Longitude specified in the form "lat,lon"
wkt string (query)	Latitude/Longitude specified in Well-Known Text
icaoCode string (query)	Airport code or location indicator. A four-letter code designating aerodromes around the world. These codes are, as defined by the International Civil Aviation Organization, are used by air traffic control and airline operations such as flight planning
iataCode string (query)	IATA airport code. Also known as an IATA location identifier, IATA station code, or simply a location identifier. Three-letter geocode designating many airports and metropolitan areas around t

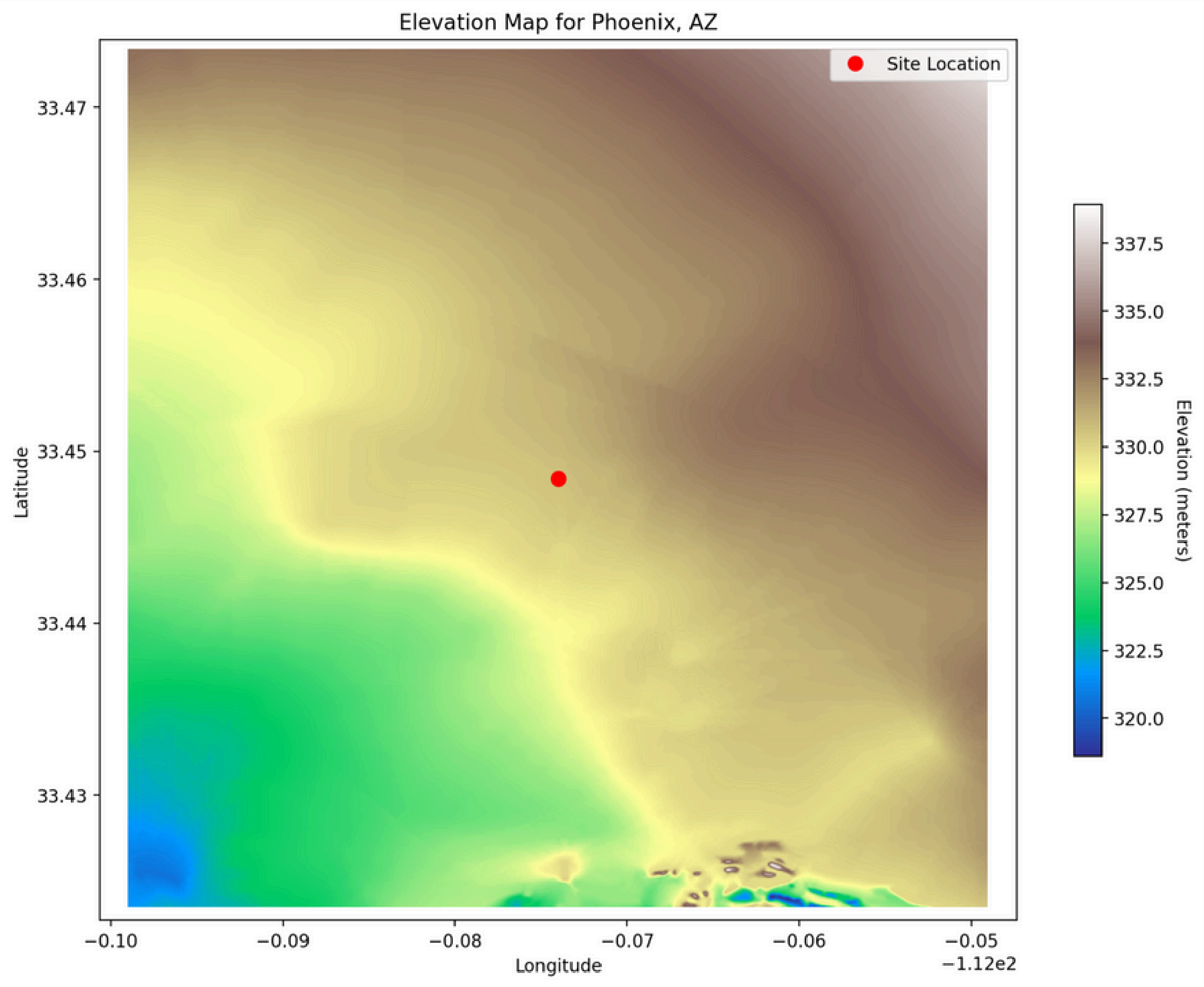

```
[
  {
    "requestedLatitude": 42,
    "requestedLongitude": -89,
    "latitude": 41.99999999999999,
    "longitude": -88.98927340946551,
    "gridpointId": "dp2wyvzgxxkqg",
    "validTimeUtc": "2021-08-19T00:20:00+0000",
    "drivingDifficultyIndex": 0,
    "evapotranspiration": 2.3,
    "globalHorizontalIrradiance": 360.2,
    "iconCode": 30,
    "iconCodeExtended": 3000,
    "precip1Hour": 0,
    "precip6Hour": 0,
    "precip24Hour": 0,
    "precip2Day": 0,
    "precip3Day": 0,
    "precip7Day": 0.03,
    "precipMtd": 1.69,
    "precipYtd": 17.6,
    "pressureChange": 0,
    "pressureMeanSeaLevel": 30.05,
    "relativeHumidity": 71.1,
    "snow1Hour": 0,
    "snow6Hour": 0,
    "snow24Hour": 0
```

API USED: Geospatial API

US NED elevation - 10m

US NED elevation - 10m	USGS National Elevation Dataset (NED). Raster-based land elevation data for the conterminous United States, Alaska, Hawaii, and territorial islands, providing basic elevation information for earth science studies and mapping applications.	14	2013	Global





Final Score Calculations

SOLAR FARMS:

Basic System Parameters:

- Assumes Size 1MW plant is to be created in every candidate location.
- Base derate factor of 0.84 – Basic System Losses.
- Temperature coefficient of -0.0035 – How overheating may cause efficiency deterioration.

Irradiance Calculation:

- Converts annual Global Horizontal Irradiance (GHI) and Direct Normal Irradiance (DNI) to daily averages. Uses a weighted combination of GHI and DNI with extra weight on DNI for desert locations. DNI weight increases with latitude (0.3 to 0.55), favoring clear sky locations

Temperature Derating:

- Exponential penalty for extreme heat which is less severe in Dry conditions.

Final Score Calculations

SOLAR FARMS:

Humidity Derating:

- Very dry conditions (<40% humidity) get a 2% boost, Minimal effect in normal range (40-70%), More significant penalty in very humid conditions (>70%)

Elevation Adjustment:

- Small benefit for higher elevations (3% per 1000m). Capped at 6% maximum

Latitude Adjustment:

- Tropical (0-23.5°): 5% boost
- Subtropical (23.5-35°, including Phoenix): 8% boost (highest)
- Temperate (35-50°): Decreasing factor
- High latitude (>50°): Steeper penalties

FINAL:

Daily energy = system capacity × effective irradiance × combined derate factors

Solar Score = (Daily Energy / (System Capacity × 24 × 0.3)) × 100

Final Score Calculations

WIND FARMS:

Basic Inputs:

- Average wind speed
- Maximum wind gust
- Wind stability (lower values indicate more consistent wind)

Elevation Factor:

- Higher elevations get a bonus (up to 40% boost at 1500m+)
- This reflects better wind conditions at elevation

Score Calculation:

- 50% weight on average wind speed
- 30% weight on maximum wind gust
- 20% weight on wind stability (inverted, so stable wind scores higher)
- This weighted sum is multiplied by the elevation factor

Final Score Calculations

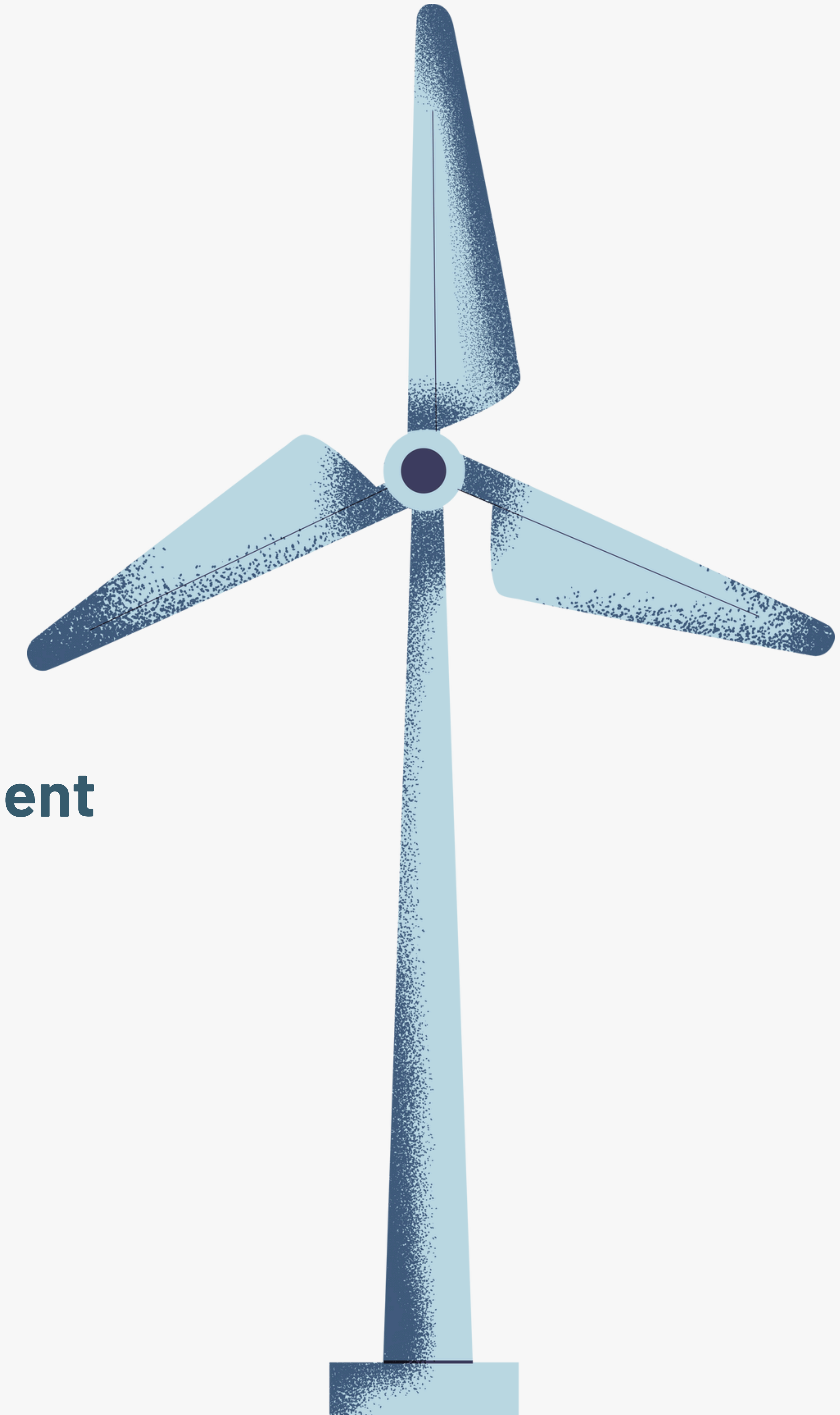
WIND FARMS:

Final Output Calculation

- Assumes a standard 80m rotor diameter
- Includes the elevation factor in wind speed

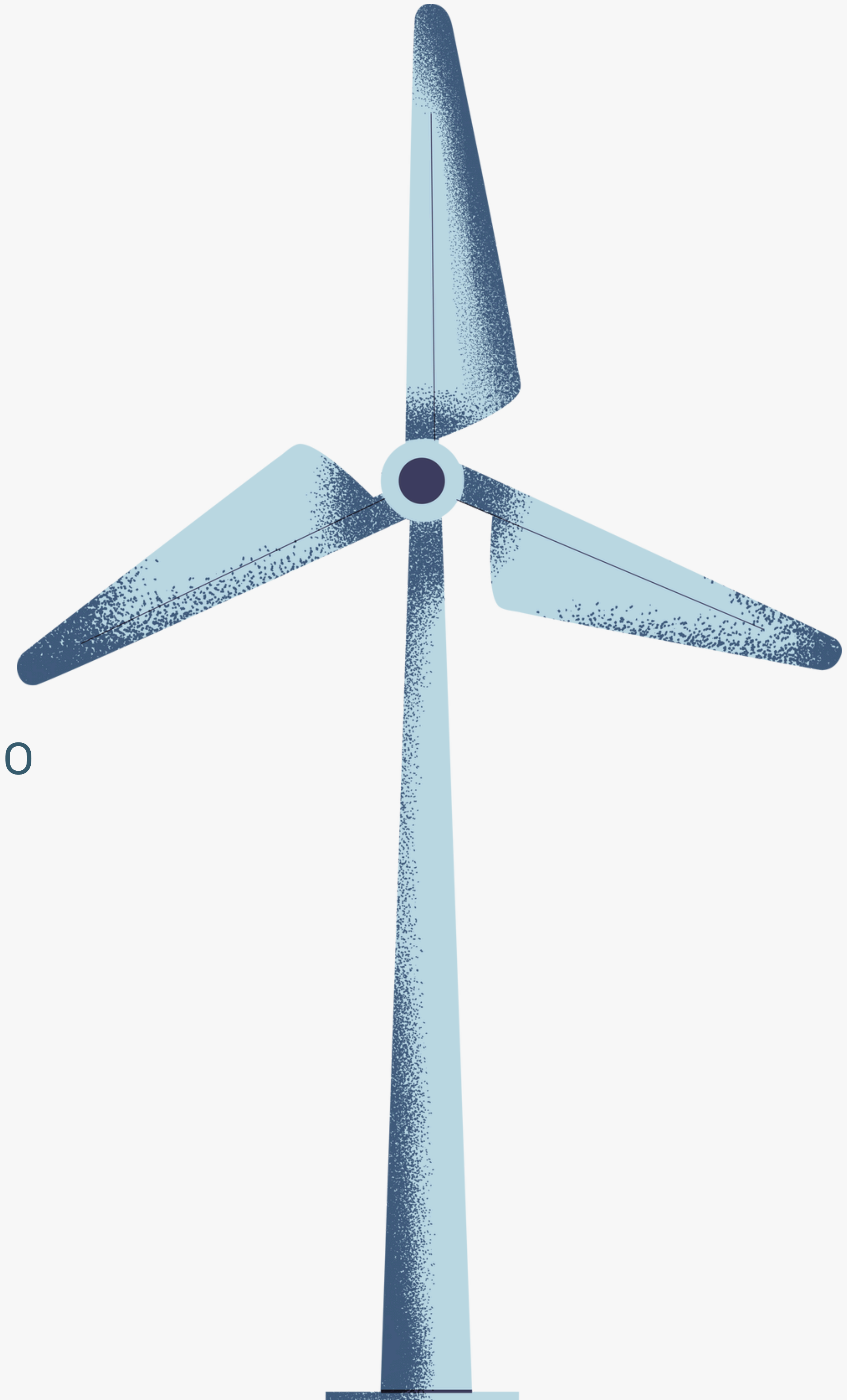
Uses standard wind power formula:

0.5 × air density × swept area × (wind speed^3) × power coefficient



ML Model for Energy Prediction

- ML Models Are Used for Energy Prediction in the Renewable Energy App
- The application uses machine learning specifically XGBoost to predict energy production at different locations



Model Training Process

Dataset Creation Process:

The foundation of this predictive system begins with comprehensive data collection from two primary sources:

1. **Weather Data Acquisition:**

Hourly historical weather parameters are fetched from IBM's History on Demand API for specific geographic locations. This includes:

- Solar radiation metrics (global horizontal irradiance)
- Atmospheric conditions (barometric pressure at sea level)
- Temperature readings
- UV index measurements
- Precise timestamp information (month, day, hour)

2. Power Output Data:

Corresponding solar power generation metrics are obtained from NREL's databases, capturing:

- DC array output in watts
- AC system output in watts
- Matching timestamps for temporal alignment

Model Training Process

Data Integration & Feature Engineering:

The raw datasets undergo meticulous processing where:

- Timestamps serve as the joining key to merge weather and power data
- Temporal features are extracted and encoded to capture cyclical patterns
- Irrelevant or redundant features are removed to focus the predictive power

Machine Learning Pipeline:

The modeling approach employs XGBoost for its proven effectiveness with:

1. Feature-Target Relationship:

The model learns complex nonlinear relationships between weather conditions (inputs) and AC power output (target), automatically determining feature importance through gradient boosting.

2. Temporal Pattern Recognition:

Special encoding of time features enables the model to recognize:

- Diurnal solar patterns (daily production cycles)
- Seasonal variations (monthly production trends)
- Weather-dependent generation fluctuations

Model Results

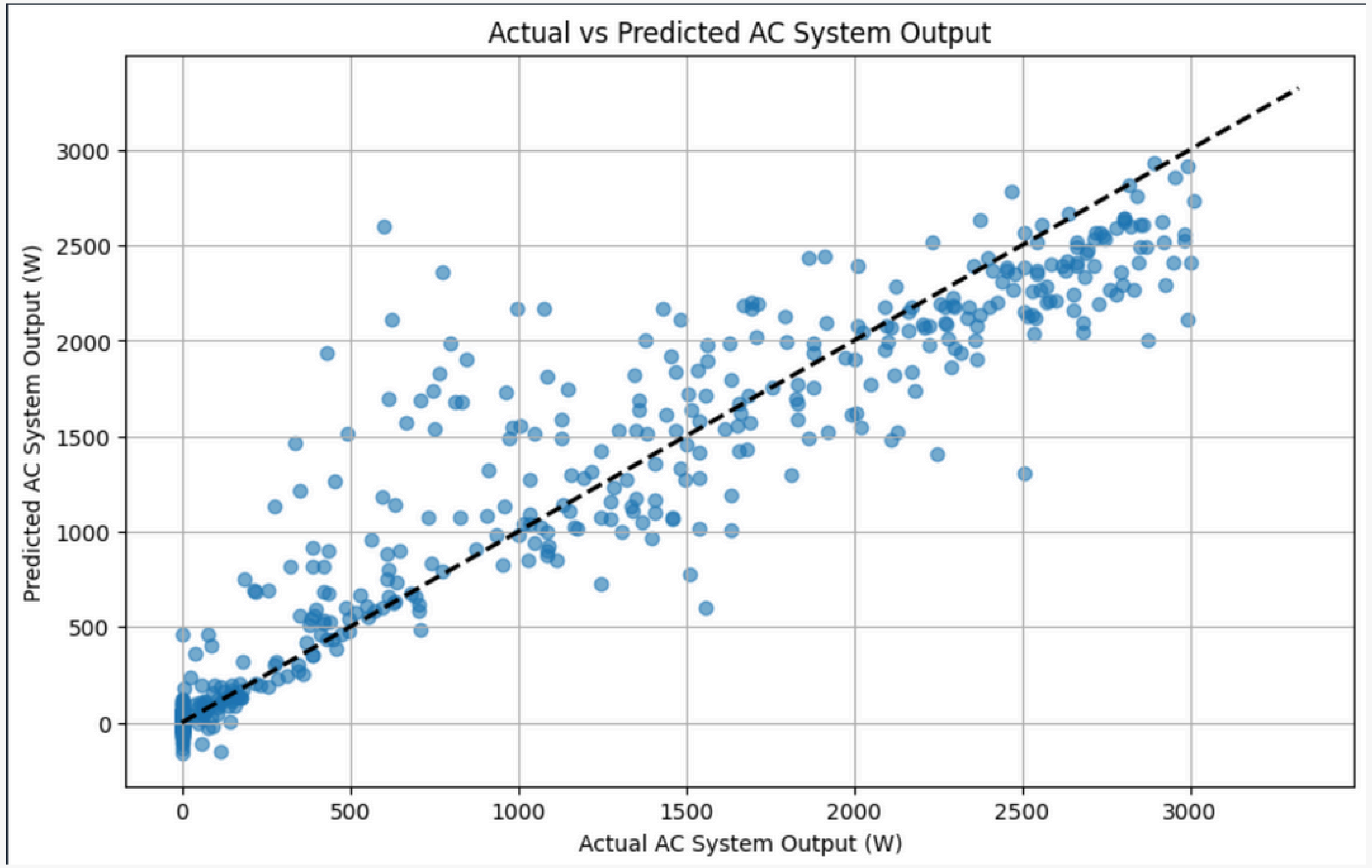
Model Evaluation Metrics:

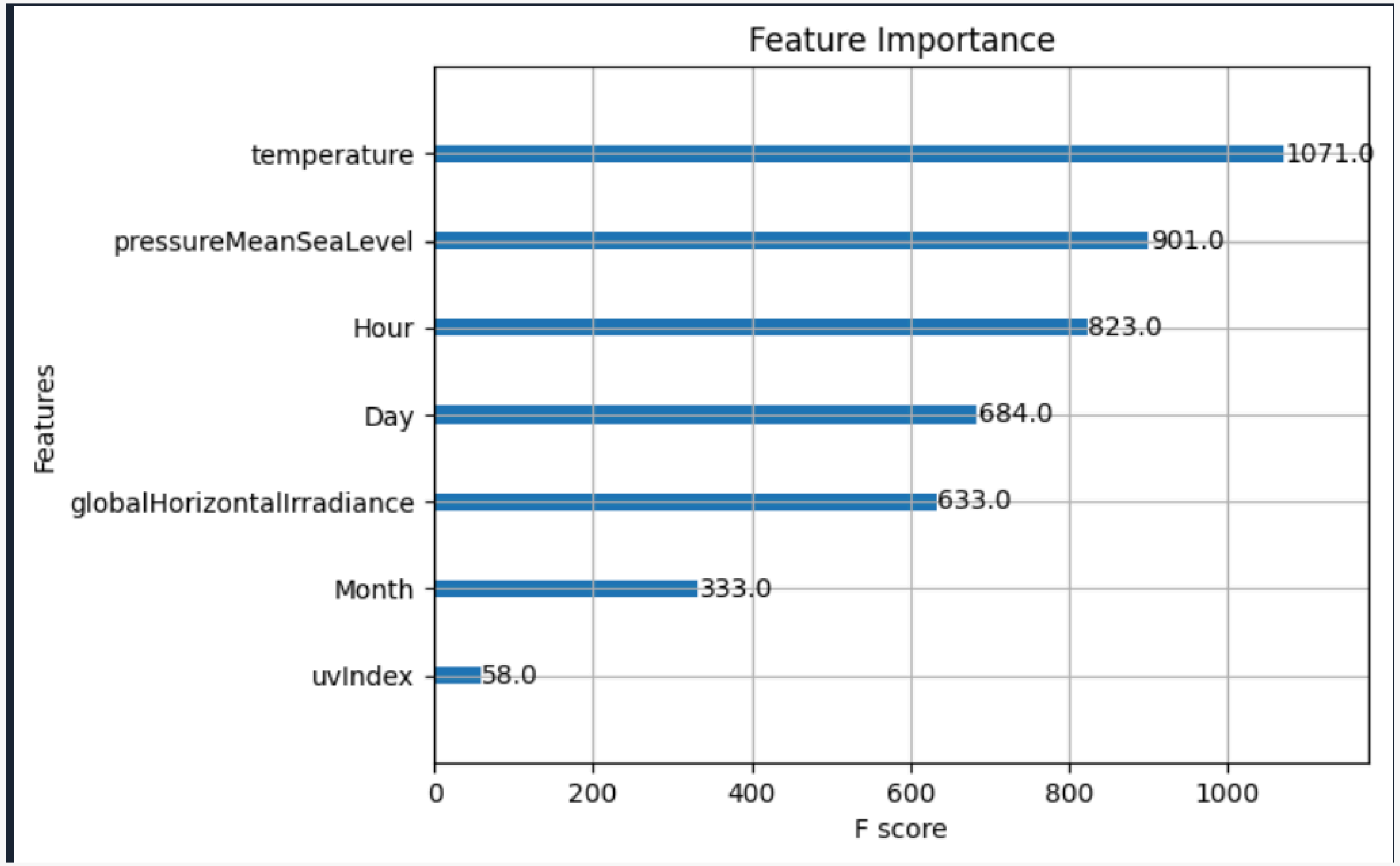
Mean Absolute Error (MAE): 166.65 W

Mean Squared Error (MSE): 89496.78 W^2

Root Mean Squared Error (RMSE): 299.16 W

R-squared (R^2) Score: 0.9083

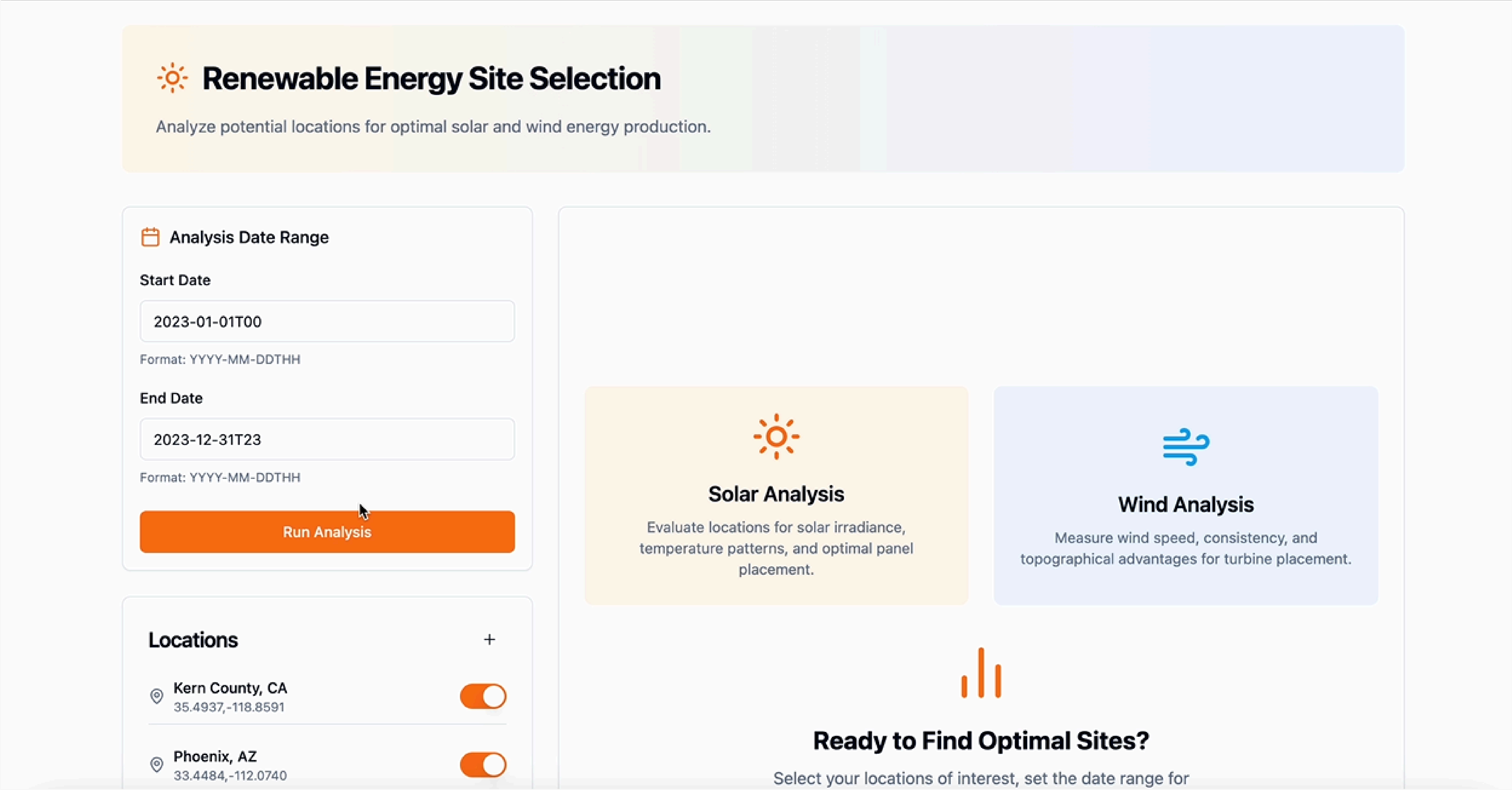




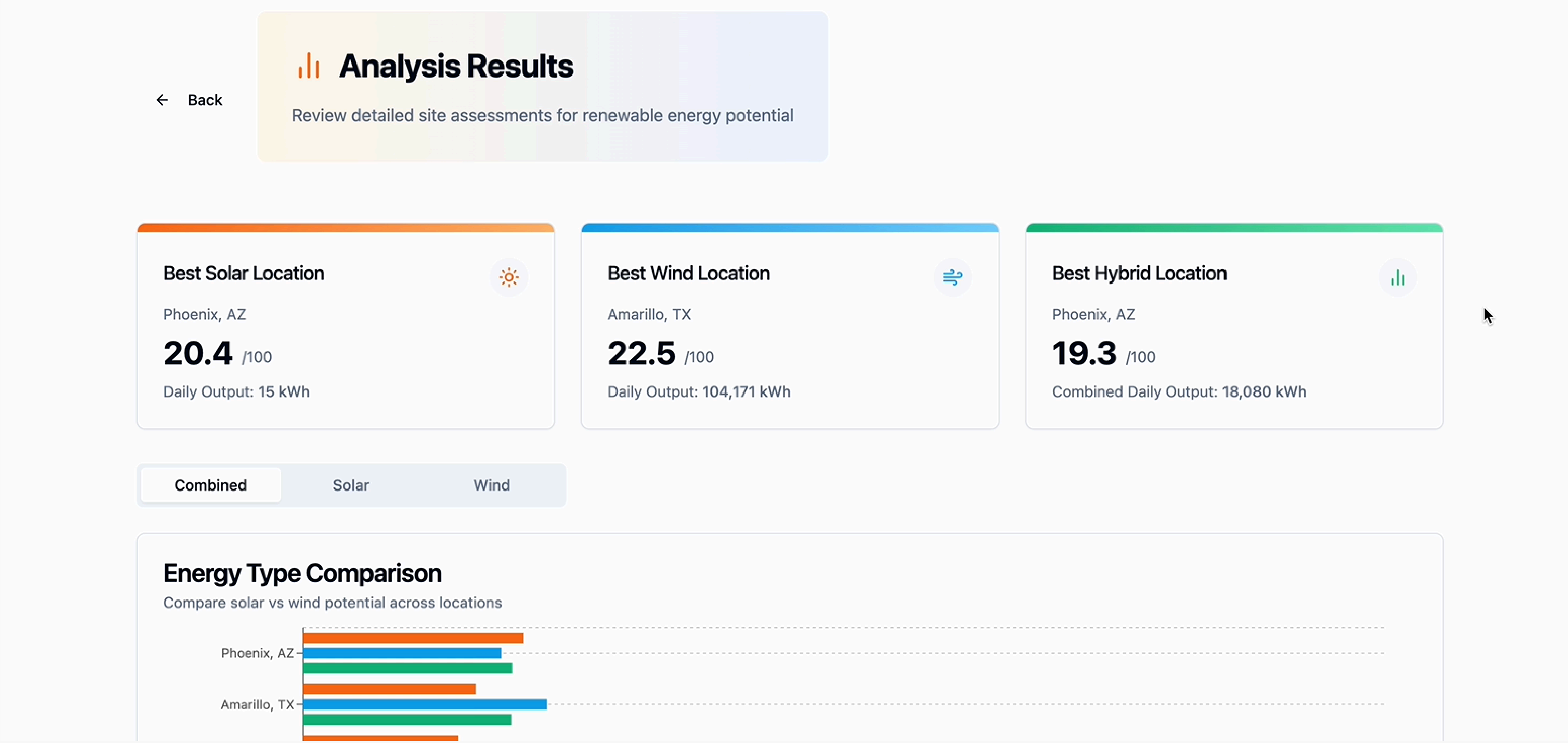
RESULTS:


Renewable Energy Site Selection System
Identify optimal locations for wind farms and solar panel installations using advanced geospatial and weather data analysis. Maximize energy production and efficiency with data-driven insights.
Solar Analysis
Evaluate locations based on solar irradiance, temperature, humidity, elevation, and other factors that affect solar panel efficiency.
Wind Analysis
Assess wind speeds, direction, stability, and elevation to determine the best locations for wind turbine installations.
Get Started

RESULTS:


Renewable Energy Site Selection
Analyze potential locations for optimal solar and wind energy production.
Analysis Date Range
Start Date
2023-01-01T00
Format: YYYY-MM-DDTHH
End Date
2023-12-31T23
Format: YYYY-MM-DDTHH
Run Analysis
Locations
+
Kern County, CA
35.4937,-118.8591
Phoenix, AZ
33.4484,-112.0740
Solar Analysis
Evaluate locations for solar irradiance, temperature patterns, and optimal panel placement.
Wind Analysis
Measure wind speed, consistency, and topographical advantages for turbine placement.
Ready to Find Optimal Sites?
Select your locations of interest, set the date range for

RESULTS:

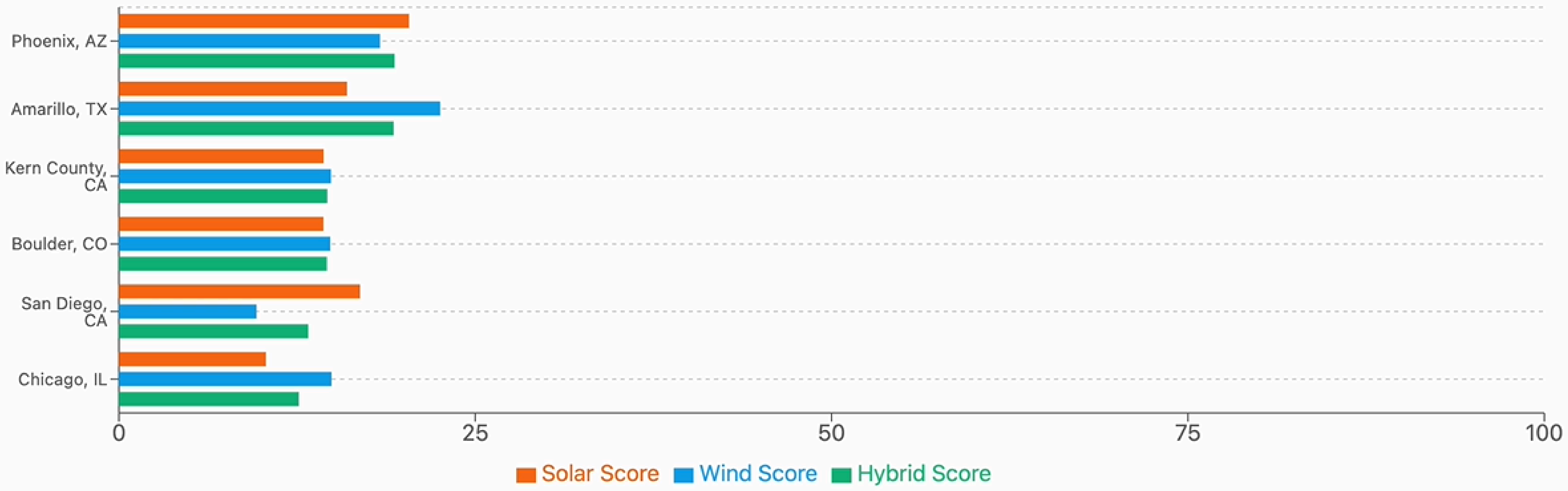

Back
Analysis Results
Review detailed site assessments for renewable energy potential
Best Solar Location
Phoenix, AZ
20.4 /100
Daily Output: 15 kWh
Best Wind Location
Amarillo, TX
22.5 /100
Daily Output: 104,171 kWh
Best Hybrid Location
Phoenix, AZ
19.3 /100
Combined Daily Output: 18,080 kWh
Combined
Solar
Wind
Energy Type Comparison
Compare solar vs wind potential across locations
Phoenix, AZ
Amarillo, TX

RESULTS:

Combined | Solar | Wind

Energy Type Comparison

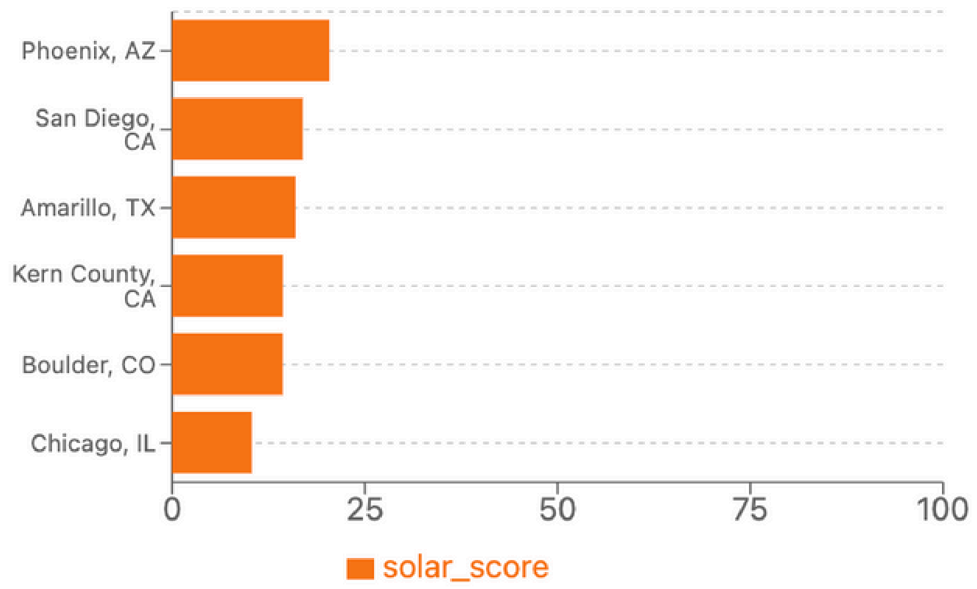
Compare solar vs wind potential across locations



RESULTS:

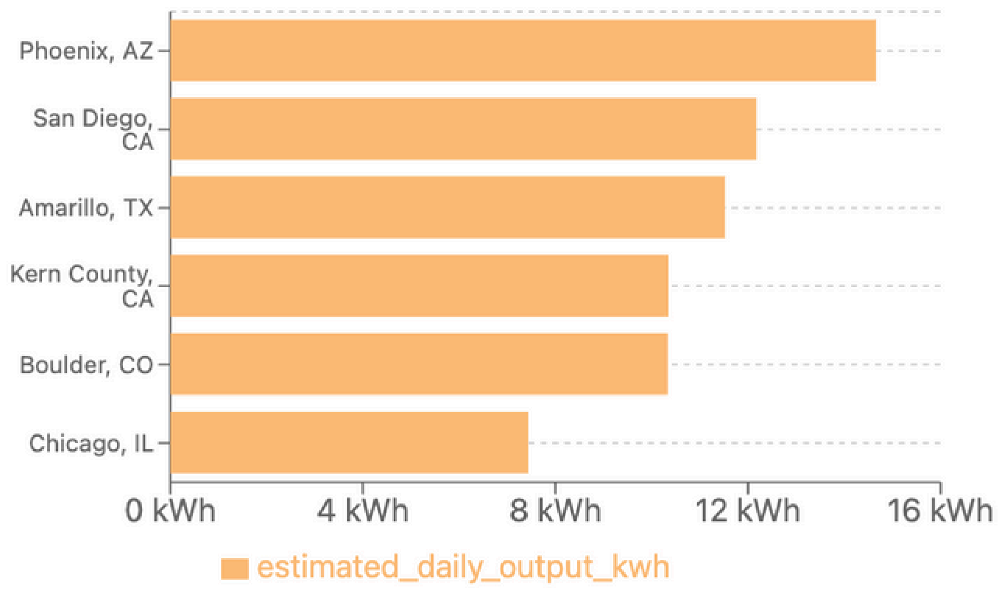
Solar Suitability Score

Ranked by overall solar potential



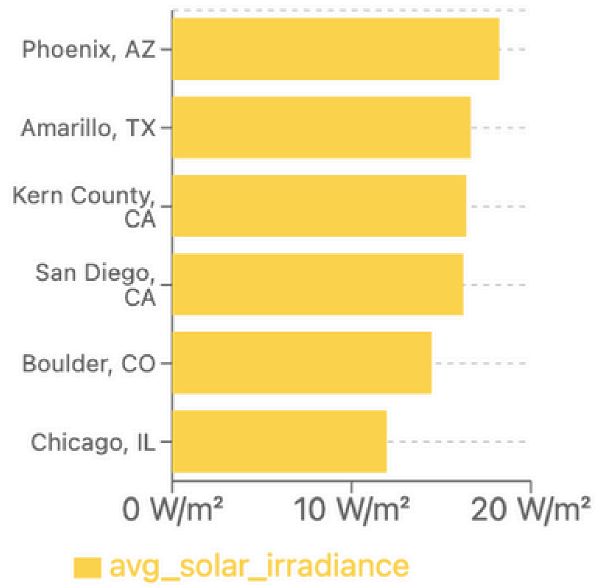
Estimated Solar Energy Production

Daily output in kilowatt-hours



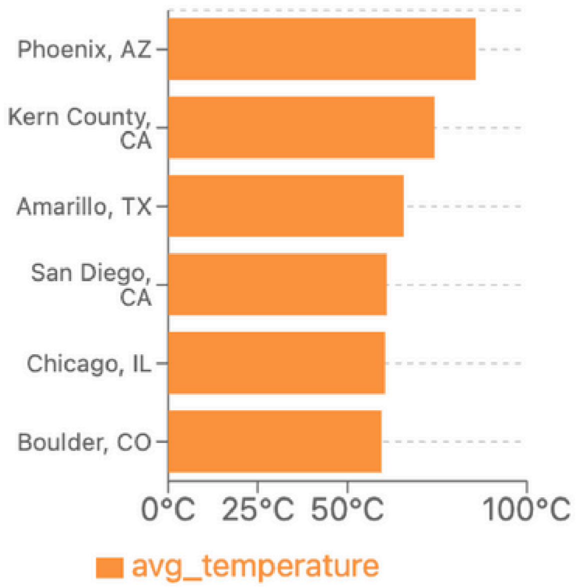
Solar Irradiance

Average irradiance in W/m²



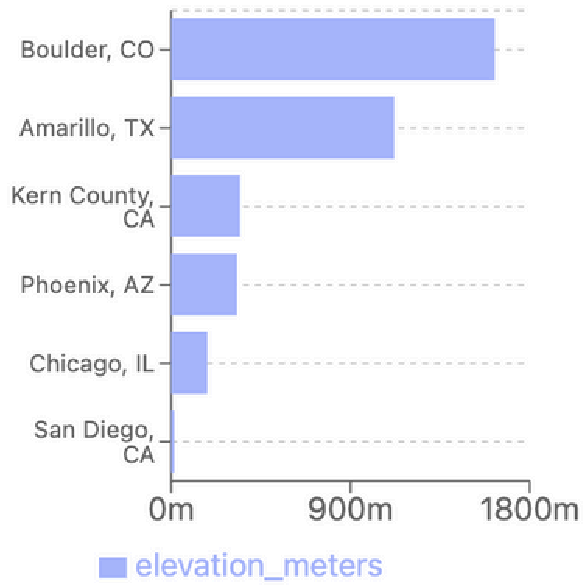
Average Temperature

In degrees celsius

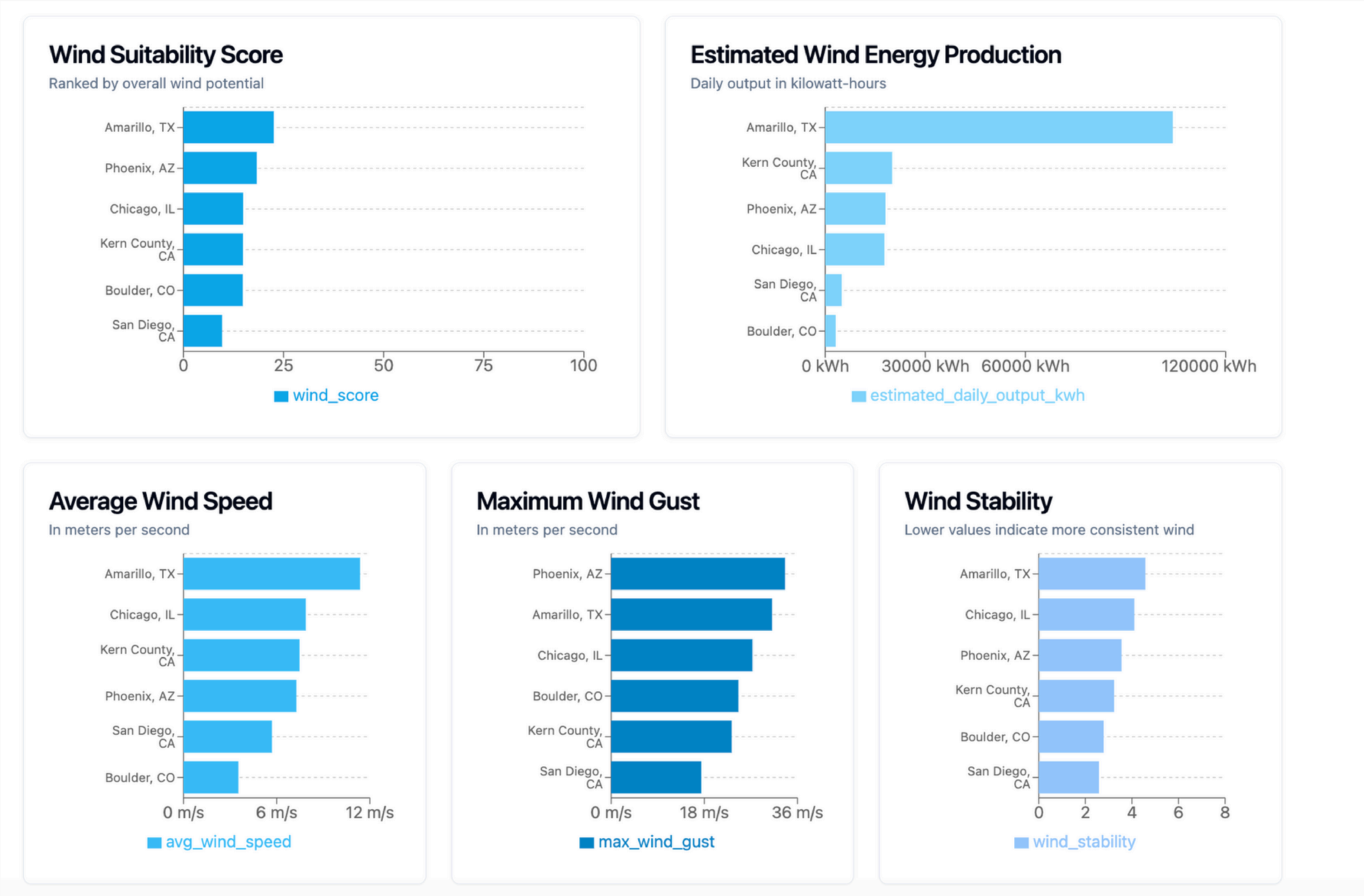


Elevation vs Solar Score

Impact of elevation on solar potential

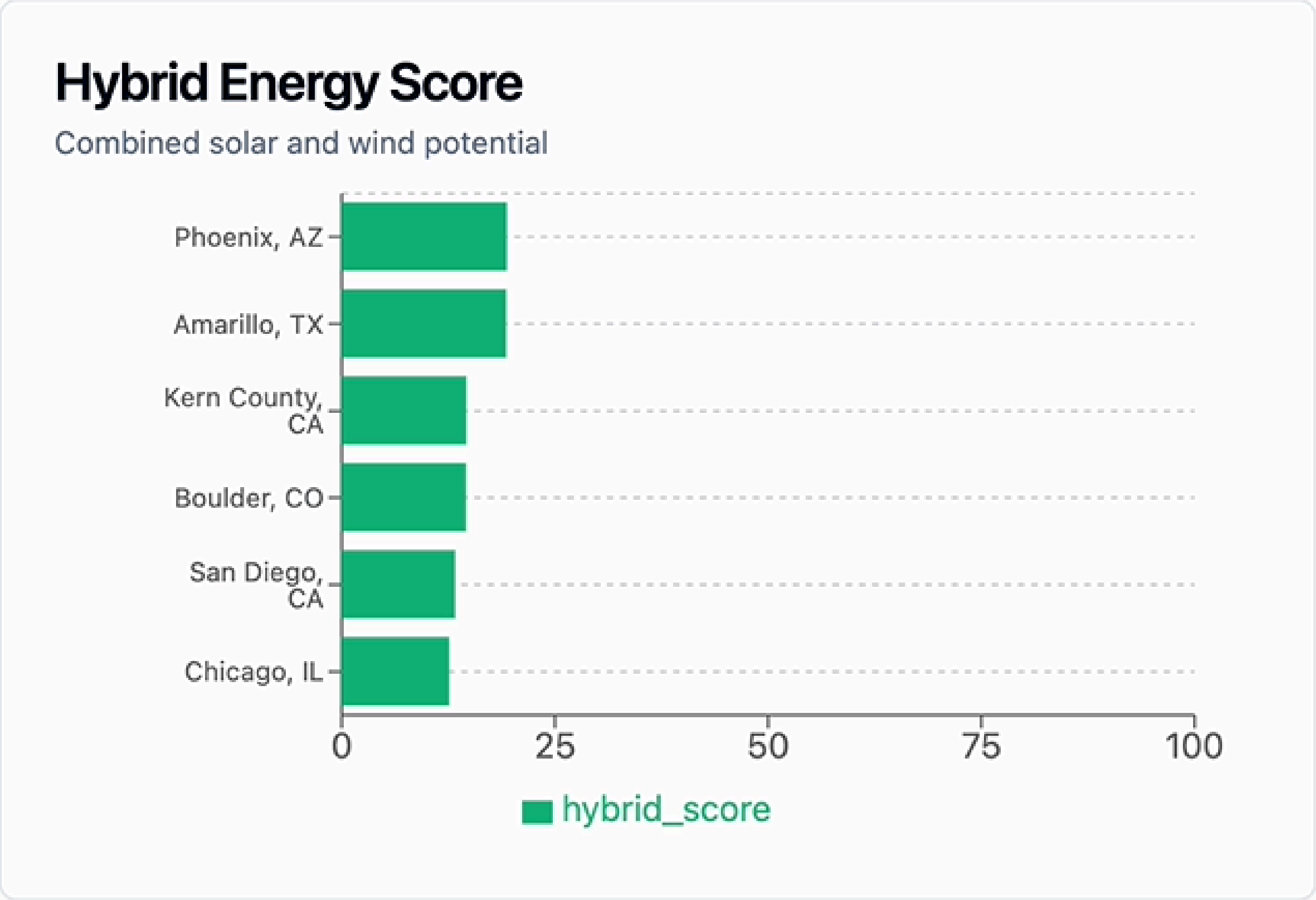


RESULTS:

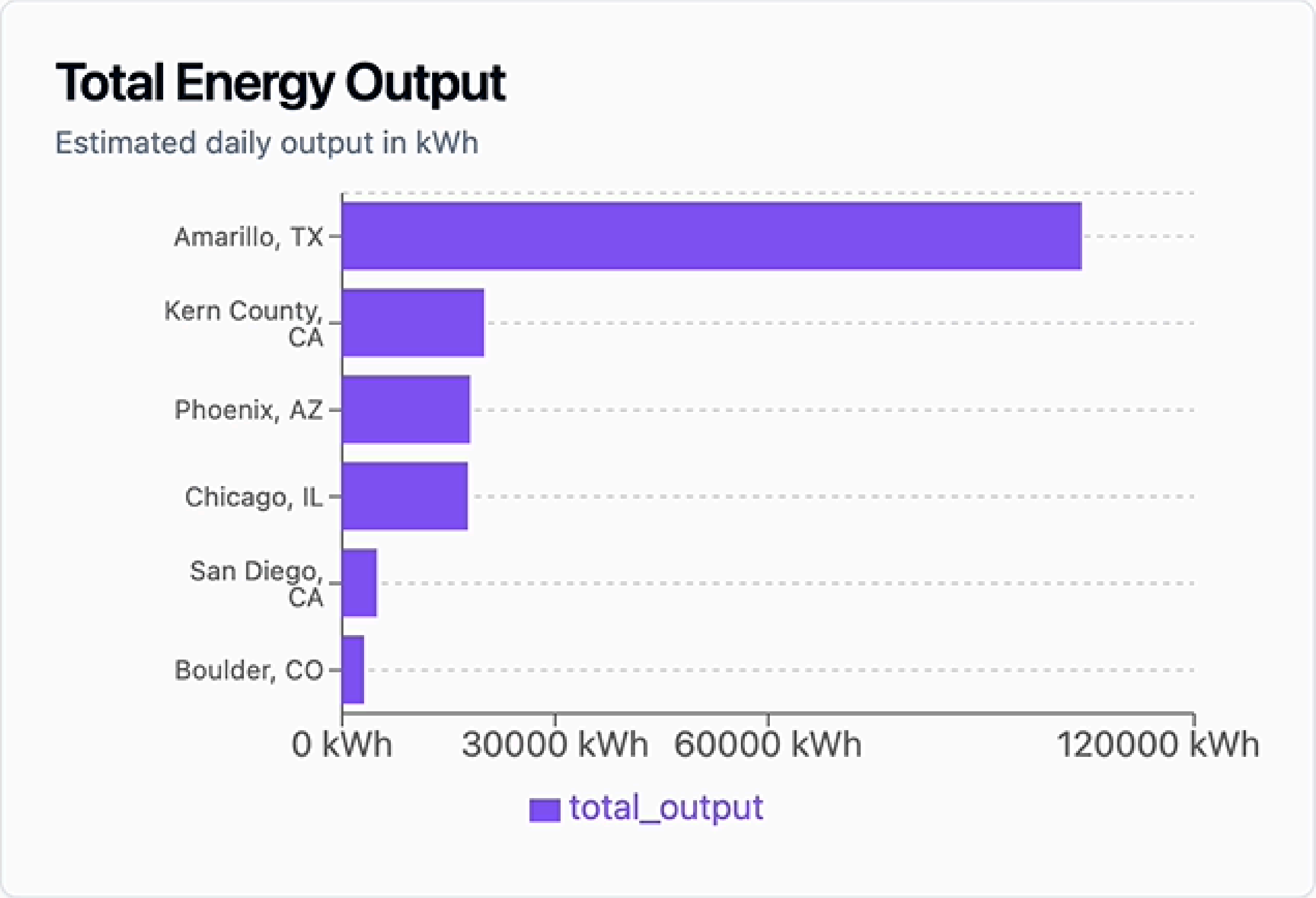


Wind Suitability Score
Ranked by overall wind potential
Amarillo, TX
Phoenix, AZ
Chicago, IL
Kern County, CA
Boulder, CO
San Diego, CA
0
25
50
75
100
wind_score
Estimated Wind Energy Production
Daily output in kilowatt-hours
Amarillo, TX
Kern County, CA
Phoenix, AZ
Chicago, IL
San Diego, CA
Boulder, CO
0 kWh
30000 kWh
60000 kWh
120000 kWh
estimated_daily_output_kwh
Average Wind Speed
In meters per second
Amarillo, TX
Chicago, IL
Kern County, CA
Phoenix, AZ
San Diego, CA
Boulder, CO
0 m/s
6 m/s
12 m/s
avg_wind_speed
Maximum Wind Gust
In meters per second
Phoenix, AZ
Amarillo, TX
Chicago, IL
Boulder, CO
Kern County, CA
San Diego, CA
0 m/s
18 m/s
36 m/s
max_wind_gust
Wind Stability
Lower values indicate more consistent wind
Amarillo, TX
Chicago, IL
Phoenix, AZ
Kern County, CA
Boulder, CO
San Diego, CA
0
2
4
6
8
wind_stability

RESULTS:



Hybrid Energy Score
Combined solar and wind potential
Phoenix, AZ
Amarillo, TX
Kern County, CA
Boulder, CO
San Diego, CA
Chicago, IL
0
25
50
75
100
hybrid_score



Total Energy Output
Estimated daily output in kWh
Amarillo, TX
Kern County, CA
Phoenix, AZ
Chicago, IL
San Diego, CA
Boulder, CO
0 kWh
30000 kWh
60000 kWh
120000 kWh
total_output

Dataset Usecases:

1. Renewable Energy Site Selection
   - Identifying optimal locations for solar and wind farms based on historical weather patterns.
   - Comparing multiple locations to rank them based on energy potential (solar irradiance, wind speeds, etc.)

2. Energy Production Forecasting
   - Using historical solar irradiance and wind speed data to estimate potential energy output.
   - Predicting seasonal variations to optimize grid integration and energy storage planning.

Thank you very much!

IBM API HACKATHON