Research Plan

Weather Stability vs Renewable Energy Model Performance

Comparative Analysis of Renewable Energy Prediction Models

Evaluating model robustness under weather-stable vs weather-unstable conditions using 2024 Germany data

Research Methodology

10 Milestones • Statistical Analysis • Reproducible Research

Contents

1 Research Overview

Objective

Determine which statistical models for renewable energy prediction perform best under stable vs unstable weather conditions, and provide operational guidance on model selection.

1.1 Research Goal

This research aims to:

1. Develop a **Weather Stability Index (WSI)** using 11 weather attributes to classify periods as stable or unstable
2. Apply statistical models from literature to predict renewable energy production for 2024
3. Compare model performance under stable vs unstable weather conditions
4. Identify which models are most robust to weather instability
5. Provide operational recommendations on model selection based on weather conditions

1.2 Pipeline Architecture

The research follows two parallel pipelines that converge for comparative analysis:

Pipeline 1: Weather Stability Classification

- **Input:** 11 weather attributes (2024, hourly, Germany-wide + 16 Bundesländer)
  - Temperature (mean, min, max)
  - Cloudiness
  - Dew point
  - Extreme wind
  - Moisture
  - Precipitation
  - Pressure
  - Soil temperature
  - Sun
  - Visibility
  - Weather phenomena
  - Wind & wind_synop
- **Process:** Feature engineering → WSI computation → Stable/Unstable classification
- **Output:** Timeline with weather stability labels

Pipeline 2: Renewable Energy Prediction Models

- **Input:** Renewable energy production data + relevant weather features

- Solar: radiation, cloudiness, temperature
- Wind: wind speed, wind direction, pressure

- **Process:** Apply statistical models (ARIMA, Prophet, Persistence, etc.) → Generate hourly predictions
- **Output:** Timeline with predictions and performance metrics (MAE, RMSE, MAPE)

Pipeline 3: Comparative Analysis

- Merge WSI timeline with model performance timeline
- Statistical tests comparing stable vs unstable periods
- Model ranking by accuracy and robustness
- Operational recommendations

2 Completed Work - Data Collection & Preprocessing

Objective

Document the weather data that has already been collected, processed, and aggregated for Germany 2024.

2.1 Data Sources

Data Provider: Deutscher Wetterdienst (DWD) - German Meteorological Service

- Source: Hourly weather data from DWD climate data center
- Period: 2024 (full year)
- Spatial Coverage: Germany-wide + 16 Bundesländer
- Station Network: 636+ weather stations
- Data License: Open data license

2.2 Attributes Collected (11 attributes)

1. **Temperature** - hourly mean, min, max
2. **Cloudiness** - cloud cover percentage
3. **Dew point** - atmospheric moisture
4. **Extreme wind** - peak wind measurements
5. **Moisture** - relative humidity
6. **Precipitation** - rainfall in mm
7. **Pressure** - atmospheric pressure in hPa
8. **Soil temperature** - ground temperature measurements
9. **Sun** - sunshine duration

10. **Visibility** - horizontal visibility
11. **Weather phenomena** - categorical weather events
12. **Wind & wind_synop** - wind speed, direction

2.3 Data Processing Steps Completed

Step 1: Download & Initial Processing

- Downloaded raw data from DWD servers
- Converted from semicolon-delimited TXT to CSV format
- Standardized timestamp format (MESS_DATUM column)
- Removed metadata files and cleaned up structure
- Filtered to 2024 data only
- Logged all operations in `logs/` directory

Step 2: Bundesland Aggregation

- Mapped 636 weather stations to 16 Bundesländer using `regions.csv`
- Aggregated station-level data into Bundesland-level files
- Created 16 CSV files per attribute (one per Bundesland)
- Preserved station IDs for traceability
- Output: `Data/*_by_bundesland/*.csv`

Step 3: Germany-Wide Aggregation

- Combined all Bundesland data into single Germany-wide files
- Aggregated across all 16 states for country-level analysis
- Output: `Data/*_germany_aggregated/Germany_total.csv`

Step 4: Data Quality Assurance

- Removed empty files (no data for 2024)
- Validated completeness across attributes
- Documented missing data patterns
- Generated data inventory summary

2.4 Current Data Structure

Location: `Data/` folder

Bundesland-Level Files:

- `dew_point_by_bundesland/` - 16 CSV files
- `extreme_wind_by_bundesland/` - 16 CSV files
- `moisture_by_bundesland/` - 16 CSV files

- `precipitation_by_bundesland/` - 16 CSV files
- `pressure_by_bundesland/` - 16 CSV files
- `soil_temperature_by_bundesland/` - 16 CSV files
- `sun_by_bundesland/` - 16 CSV files
- `visibility_by_bundesland/` - 16 CSV files
- `weather_phenomena_by_bundesland/` - 16 CSV files
- `wind_by_bundesland/` - 16 CSV files
- `wind_synop_by_bundesland/` - 16 CSV files
- `Temp_Bundesland_Aggregated/` - 17 CSV files (includes cloudiness)
- `Cloudness_Bundesland_Aggregated/` - 17 CSV files

Germany-Wide Files:

- `dew_point_germany_aggregated/Germany_total.csv`
- `extreme_wind_germany_aggregated/Germany_total.csv`
- `moisture_germany_aggregated/Germany_total.csv`
- `pressure_germany/Germany_total.csv`
- `soil_temperature_germany/Germany_total.csv`
- `sun_germany/Germany_total.csv`
- `Temp_Germany_Aggregated/`
- `Cloudness_Germany_Aggregated/`

2.5 Processing Scripts

Location: `Scripts/` folder

Key Scripts:

- `process_weather_data.py` - Main processing pipeline
- `process_weather_by_bundesland.py` - Bundesland aggregation
- `process_dew_point_data.py` - Dew point specific processing
- `process_extreme_wind_data.py` - Extreme wind processing
- `aggregate_to_germany.py` - Germany-wide aggregation
- Download scripts for each attribute type
- Data cleaning and filtering scripts

Deliverables

- Documentation of completed data collection
- Data inventory of available attributes
- File structure summary

3 Milestone 1 - Define Research Scope & Success Criteria

Objective

Formalize research questions, hypotheses, and success criteria for the comparative analysis.

3.1 Research Questions (RQs)

1. **RQ1:** Does renewable energy prediction accuracy differ significantly between weather-stable and weather-unstable periods?
2. **RQ2:** Which statistical models are most robust to weather instability (i.e., show smallest performance degradation during unstable periods)?
3. **RQ3:** Can weather stability information improve model selection strategies for operational renewable energy forecasting?

3.2 Hypotheses (H)

1. **H1:** Model prediction accuracy (MAE, RMSE) is significantly worse during weather-unstable periods compared to stable periods.
2. **H2:** Different models show varying degrees of robustness to weather instability, with some models degrading more than others.
3. **H3:** Time series models (Prophet, ARIMA) are more robust to weather instability than persistence models.
4. **H4:** Model performance degradation correlates positively with Weather Stability Index (WSI) magnitude.

3.3 Performance Metrics

Primary Metrics:

- **MAE** - Mean Absolute Error (primary accuracy measure)
- **RMSE** - Root Mean Squared Error (penalizes large errors)
- **MAPE** - Mean Absolute Percentage Error (relative accuracy)

Secondary Metrics:

- **Bias** - Mean error (systematic over/under-prediction)
- **Forecast Skill** - Improvement over persistence baseline
- **Error distribution percentiles** - 50th, 75th, 95th

3.4 Success Criteria

Quantitative Thresholds:

- Statistical significance: $p < 0.05$ (with Bonferroni correction for multiple comparisons)
- Effect size: Cohen's d 0.3 for stable vs unstable performance difference

- Practical significance: 10% improvement in forecast accuracy for best vs worst model
- At least one model showing <5% performance degradation in unstable periods

Qualitative Criteria:

- Clear operational recommendations on which model to use when
- Reproducible methodology with documented code
- Publication-quality figures and report

Deliverables

- `docs/research_questions.md`
- `docs/hypotheses.md`
- `docs/success_criteria.md`

4 Milestone 2 - Acquire Renewable Energy Production Data

Objective

Obtain 2024 renewable energy production data for Germany to use as ground truth for model evaluation.

4.1 Data Requirements

Required Data:

- Hourly renewable energy production for 2024
- Solar PV production (MW or GWh)
- Wind power production (MW or GWh)
- Geographic scope: Germany-wide (ideally also by Bundesland)
- Temporal alignment: Hourly resolution matching weather data

4.2 Potential Data Sources

Option 1: ENTSO-E Transparency Platform

- European grid operator data
- Hourly generation data by type
- URL: transparency.entsoe.eu
- Free access, requires API key registration
- Has historical data for Germany

Option 2: Fraunhofer ISE Energy Charts

- Public dashboard for German renewable energy
- URL: energy-charts.info
- Provides downloadable CSV files

- Historical data available

Option 3: SMARD (Strommarktdaten)

- Federal Network Agency (Bundesnetzagentur) platform
- URL: smard.de
- Official market data including renewable generation
- Free registration required

Option 4: OpenCinema

- Open source platform for energy data
- URL: openoemof.readthedocs.io
- May have aggregated data

4.3 Data Validation

Checks

- Verify complete temporal coverage (all 8760 hours of 2024)
- Check for reasonable values (no negative production, no spikes > installed capacity)
- Validate against known installed capacity statistics
- Document data source and access method

4.4 Literature Review for Models

Survey existing literature on renewable energy forecasting to identify:

- Commonly used statistical models
- Reported performance benchmarks (MAE, RMSE baselines)
- Input features typically used
- Best practice methodologies
- Similar studies for German context

Focus Areas:

- Solar PV forecasting papers
- Wind power forecasting papers
- German/EU specific studies
- Comparisons of statistical vs ML methods

Deliverables

- `Data/raw/renewable_production_2024.csv`
- `docs/literature_review.md` (model summaries)
- `docs/energy_data_sources.txt`

5 Milestone 3 - Data Integration & Cleaning

Objective

ARP-onize all datasets to ARP-onize all datasetstimestamps and prepare unified analysis-ready data.

5.1 Timestamp Harmonization

1. Standardize all timestamps:
   - Convert to UTC or consistent timezone
   - ISO 8601 format: `YYYY-MM-DD HH:MM:SS`
   - Hourly granularity throughout
2. Create master time index:
   - Generate complete 2024 hourly sequence
   - 8,784 hours total (2024 was a leap year)

5.2 Missing Data Strategy

Tiered Approach:

- **Tier 1 (2 hours):** Linear interpolation
- **Tier 2 (3-6 hours):** Forward-fill with flag
- **Tier 3 (>6 hours):** Mark as excluded from analysis

Documentation:

- Log all imputations in `data/cleaning_log.csv`
- Track percent of original data retained
- Create missingness heatmap visualization

5.3 Dataset Merging

Create unified dataset containing:

```
timestamp, location,
temp_mean, temp_min, temp_max, cloudiness,
dew_point, extreme_wind, moisture, precipitation,
pressure, soil_temp, sun, visibility, weather_phenomena,
wind_speed, wind_dir,
solar_production, wind_production,
excluded_flag
```

5.4 Quality Checks

Checks

- Check for duplicate timestamps
- Flag outliers using domain knowledge (e.g., temp < -40°C or > 50°C)
- Verify no missing gaps in time series
- Validate production data against installed capacity
- Generate summary statistics for all variables

Deliverables

- `data/processed/unified_dataset.csv`
- `data/cleaning_log.csv`
- `notebooks/01_data_integration.ipynb`

6 Milestone 4 - Weather Feature Engineering for WSI

Objective

Engineer features from raw weather attributes that capture variability, trends, and extremes for computing Weather Stability Index.

6.1 Variability Features

Compute rolling statistics (24-hour window):

- `temp_std` - Temperature standard deviation
- `temp_range` - Max - min temperature
- `pressure_change` - Absolute change in pressure
- `wind_cv` - Wind speed coefficient of variation
- `precip_intensity` - Precipitation rate changes
- `humidity_std` - Dew point/moisture variability

6.2 Trend Features

Linear trends over 24-hour window:

- `temp_trend` - Temperature trend (slope)
- `pressure_trend` - Barometric pressure trend
- `wind_trend` - Wind speed trend

6.3 Extreme Event Flags

Categorical indicators of extreme conditions:

- `high_wind_flag` - Wind > 90th percentile
- `heavy_precip_flag` - Precipitation > threshold

- `rapid_temp_change` - |T| > 5°C in 3 hours
- `storm_flag` - Combined wind + pressure drop signal

6.4 Feature Selection

Correlation Analysis:

- Compute correlation matrix for all features
- Identify highly correlated features (r > 0.9)
- Remove redundant features to avoid multicollinearity
- Select final feature set for WSI (typically 6-10 features)

6.5 Normalization

- Use robust scaling (median, IQR) to handle outliers
- Orient features so higher values = more instability
- Save scaling parameters to `models/scalers.json`

Deliverables

- `data/features/weather_features.csv`
- `results/feature_correlations.png`
- `notebooks/02_feature_engineering.ipynb`

7 Milestone 5 - Compute Weather Stability Index (WSI)

Objective

Create a reproducible continuous WSI score and classify periods as stable/unstable.

7.1 WSI Formula Development

Test multiple approaches:

1. **Equal Weights:** Simple average of normalized features
2. **PCA-Based:** Use first principal component as WSI (captures max variance)
3. **Variance-Weighted:** Weight features by their explained variance
4. **Domain-Expert Weights:** Weight based on meteorological importance regardding the weather stability

Baseline Formula (Equal Weights):

$$\text{WSI} = \frac{\sum_{i=1}^{n} \text{feature}_i}{n} \tag{1}$$

7.2 Classification Methods

Primary: K-Means Clustering (k=2)

- Cluster into stable/unstable using all features
- Label cluster with higher mean WSI as unstable
- Check cluster balance (avoid 90/10 splits)

Secondary: Percentile Threshold

- Classify as unstable if WSI 75th percentile
- Sensitivity test with 70th and 80th percentiles

Alternative: Dynamic Seasonal Thresholds

- Use season-specific cutoffs (winter vs summer)
- Account for natural seasonal variability

7.3 Validation

Checks

- Inspect cluster sizes (reasonable balance)
- Manually verify unstable periods against known weather events
- Plot WSI timeline to check for reasonable patterns
- Compare different classification methods

Deliverables

- `data/processed/wsi_timeline.csv`
- `models/wsi_formula.json`
- `notebooks/03_wsi_computation.ipynb`
- `figures/wsi_timeline_2024.png`

8 Milestone 6 - Implement Renewable Energy Prediction Models

Objective

Implement statistical models from literature and generate hourly predictions for 2024.

8.1 Literature-Based Models

Models to Implement:

1. **Persistence Model** (Baseline)

- Simplest:simple: next-hour = current-hour
- Baseline for comparison

2. **Seasonal Persistence**

- Use same hour from previous week

- Accounts for weekly patterns

3. **ARIMA/SARIMA**

- Classical time series models
- Autoregressive with seasonal components

4. **Prophet (Facebook)**

- Additive seasonality model
- Handles trends, seasonality, holidays
- Good for energy forecasting

5. **Exponential Smoothing (Holt-Winters)**

- Triple exponential smoothing
- Trend and seasonality components

8.2 Feature Selection

For Solar Prediction:

- Solar radiation (if available)
- Cloudiness percentage
- Temperature
- Hour of day (circular encoding)
- Day of year (seasonality)
- Day of week (weekly patterns)

For Wind Prediction:

- Wind speed
- Wind direction
- Atmospheric pressure
- Temperature
- Time features (hour, day, season)

8.3 Walk-Forward Validation

Strategy:

- Train window: Previous 30-90 days
- Forecast horizon: Next 24 hours
- Update daily throughout 2024
- Mimics operational forecasting

Deliverables

- `models/` - Trained model objects
- `data/predictions/model_predictions_2024.csv`
- `notebooks/04_model_implementation.ipynb`
- `docs/model_descriptions.md`

9 Milestone 7 - Compute Model Performance Metrics

Objective

Calculate accuracy metrics for all models and align with WSI timeline for comparative analysis.

9.1 Core Metrics Calculation

Per-Hour Metrics:

- `error` = prediction - actual
- `abs_error` = |error|
- `squared_error` = error^2

Aggregated Metrics (Daily/Hourly):

- **MAE** = mean(|error|)
- **RMSE** = $\sqrt{\text{mean}(\text{error}^2)}$
- **MAPE** = mean(|error/actual|) × 100%
- **Bias** = mean(error)

Additional Metrics:

- **Forecast Skill** = 1 - (MAE$_model/MAE_persistence$)
- **Error percentiles**: 50th, 75th, 95th
- **Rolling 7-day MAE**

9.2 Merge with WSI Timeline

Create unified analysis dataset:

```
timestamp, location, WSI, stability_label,
model_name, prediction, actual, error, abs_error,
MAE_24h, RMSE_24h, MAPE_24h, Bias_24h,
MAE_7d_rolling, Forecast_skill
```

9.3 Stratified Performance Statistics

Pre-compute performance by:

- Stable vs Unstable periods

- Model type
- Season (quarter)
- Location (Germany vs Bundesland)

Deliverables

- `data/metrics/performance_timeline.csv`
- `data/metrics/performance_by_stability.csv`
- `notebooks/05_performance_metrics.ipynb`

10 Milestone 8 - Comparative Statistical Analysis

Objective

Test hypotheses and quantify differences in model performance between stable and unstable weather periods.

10.1 Descriptive Statistics

For Each Model:

- Mean MAE in stable periods
- Mean MAE in unstable periods
- Difference (degradation)
- 95% confidence intervals
- Effect size (Cohen's d)

10.2 Hypothesis Testing

Test H1: Model accuracy worse in unstable periods

Test H2: Different models degrade differently

How is to be determined later

10.3 Regression Analysis

How is to be determined later

10.4 Model Ranking

Rank models on two dimensions:

1. **Overall Accuracy** - Mean MAE across all periods
2. **Robustness** - Smallest performance drop in unstable periods

Create combined score:

$$\text{Score} = \alpha \times \text{Accuracy} + (1 - \alpha) \times \text{Robustness} \tag{2}$$

10.5 Time-Lag Analysis

Test if WSI predicts performance degradation ahead of time:

- Cross-correlation function (CCF) between WSI and MAE
- Test lags: -48 to +48 hours
- Identify optimal lead time for model switching

Deliverables

- `results/statistical_tests.csv`
- `results/model_rankings.csv`
- `notebooks/06_statistical_analysis.ipynb`

11 Milestone 9 - Visualization & Interpretation

Objective

Create publication-quality figures that communicate key findings clearly.

11.1 Core Figures

Figure 1: Dual Timeline Plot

- Upper panel: WSI timeline with stable/unstable shading
- Lower panel: Rolling MAE (7-day) for top 3 models
- Shared x-axis: 2024 timeline
- Purpose: Visual correlation between stability and performance

Figure 2: Performance Comparison by Stability

- Grouped boxplots or violin plots
- MAE for each model, split by stable/unstable
- Include means and 95% CIs
- Add significance markers (*, **, ***)

Figure 3: Model Degradation Bar Chart

- Performance drop (%) from stable to unstable for each model
- Sorted by robustness (least to most degraded)
- Color-coded by model type

Figure 4: WSI vs Performance Scatter

- WSI (x-axis) vs MAE (y-axis) scatterplot
- Different colors for each model
- LOESS smoothing lines
- Annotate Spearman and p-value

Figure 5: Geographic Comparison

- Heatmap showing performance across Bundesländer
- Side-by-side: stable vs unstable periods
- Identify regions with largest degradation

Figure 6: Seasonal Patterns

- Monthly aggregation of stability and performance
- Line plots for MAE by month
- Overlay WSI distribution by month

11.2 Design Guidelines

- Colorblind-friendly palette (viridis, magma)
- Consistent fonts (Arial or similar sans-serif)
- High resolution: 300 DPI PNG, vector PDF
- Clear axis labels with units
- Figure captions with sample sizes and key parameters

Deliverables

- `figures/*.png` and `figures/*.pdf`
- `notebooks/07_visualization.ipynb`

12 Milestone 10 - Reporting & Recommendations

Objective

Synthesize findings into actionable conclusions and create final report.

12.1 Final Report Structure

1. Abstract

- Brief summary of research, methods, key findings
- Operational implications

2. Introduction

- Motivation: Why weather stability matters for renewable forecasting
- Background: Renewable energy in Germany
- Research objectives and questions

3. Data & Methods

- Data sources (weather + energy)
- Weather Stability Index construction

- Statistical models implemented
- Analysis methodology

4. Results

- WSI characteristics (
- Model performance summary tables
- Statistical test results
- Key figures

5. Discussion

- Which models are most accurate overall?
- Which models are most robust to instability?
- Practical operational recommendations
- Comparisons with literature

6. Limitations

- Data quality issues
- Model assumptions
- Generalizability beyond 2024

7. Conclusions

- Summary of key findings
- Operational guidelines for model selection
- Future research directions

12.2 Operational Recommendations

Provide clear guidance on:

- **Model selection under stable weather:** Use [Model X] for best accuracy
- **Model selection under unstable weather:** Use [Model Y] for robust performance
- **When to switch models:** If WSI exceeds threshold Z, switch from Model X to Model Y
- **Expected performance:** Typical accuracy under each condition

12.3 Reproducibility

- `requirements.txt` - Python packages with versions
- `config/parameters.yaml` - All hyperparameters and thresholds
- `README.md` - Step-by-step instructions to reproduce
- `run_all.sh` - Script to run all notebooks sequentially
- All code with comments

- Random seeds documented

Deliverables

- `report/final_report.pdf`
- `README.md` (project documentation)
- `requirements.txt`
- `config/parameters.yaml`
- All notebooks (numbered 01-07)

13 Conclusion

This research plan provides a systematic approach to comparing renewable energy prediction model performance under different weather stability conditions. By leveraging the extensive weather data already collected for 2024 Germany and following the dual-pipeline architecture, we will generate valuable insights for operational renewable energy forecasting.

The key innovation is the integration of a Weather Stability Index with model performance evaluation, enabling data-driven recommendations on which models to use under which conditions. This approach directly addresses the practical need for robust forecasting in the renewable energy sector.