Research and Emerging Topics Group project

Solar Power Emission Projection & Transition Risk Analysis

Presented by Climate Data Scientists : *Aya Bekhtiar, Fatma BRAHAM, Apolline Hadjal, Yiwen Mai*

Agenda

1- Problem Statement

2- Pipeline Architecture

3- Data Ingestion & Feature Engineering

4- Emission Modeling & Scenario Analysis

5- Transition Risk Assessment

6- API & Automation

7- Value Proposition & Benefits

8- Conclusion

Problem Statement

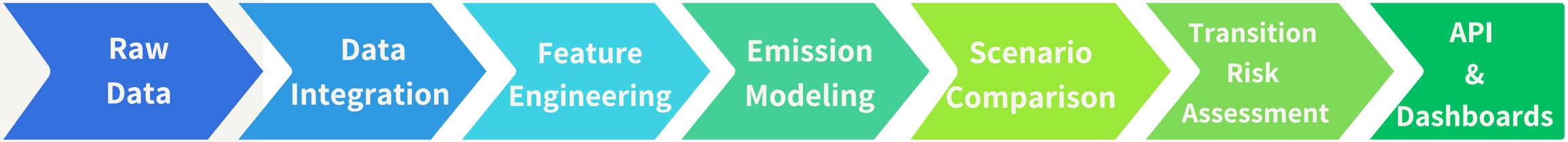
- Solar investors, policymakers and utilities need credible data to align portfolios with climate-policy pathways and avoid stranded assets
- Without robust forecasting, stakeholders risk underestimating transition risk and misallocating capital
- As Climate Data Scientists, our task is to :
    - estimate and project emissions avoided by solar infrastructure under different IEA scenarios (Net Zero Emissions, Announced Pledges and Stated Policies)

Our Dataset

+25 000 solar projects from the Global Solar Power Tracker spanning countries, technologies and capacities.

Pipeline Architecture

End-to-end, automated workflow:

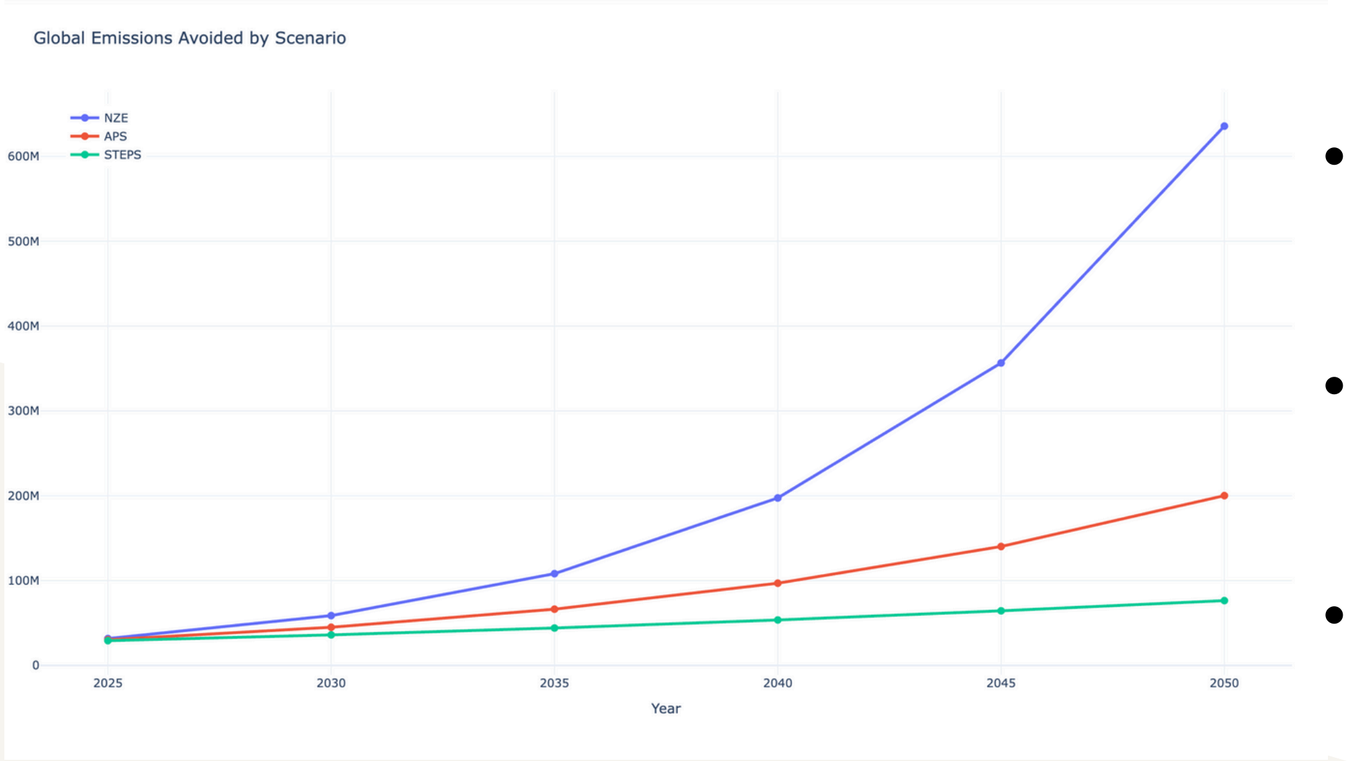


- Each stage is modular, version-controlled and reproducible → enables consistent results across runs
- The pipeline supports NZE, APS and STEPS scenarios and can be scheduled for regular updates
- Containerised deployment (Docker) and CI/CD integration ensure scalability and easy maintenance

Data Ingestion & Feature Engineering

- **Data ingestion:** validates capacity and year ranges, removes duplicates and logs data quality metrics
- **Feature Engineering :**
  - **Temporal features:** years operational, age categories, projected retirement year and years to retirement
  - **Geographic features:** region groupings (Asia, North America, Europe, South America, Oceania, Other), climate zones (tropical, subtropical, temperate, polar), hemisphere and a solar-resource proxy
  - **Capacity & technology features:** categories (small, medium, large...) log-transformed capacity and technology-efficiency proxies
  - **Risk features:** policy, technology and market risk scores combine into a composite stranded-asset risk
  - **Scenario-specific variables** (growth rate, carbon price, electrification rate) are added for NZE, APS and STEPS

Emission Modeling & Scenario Analysis

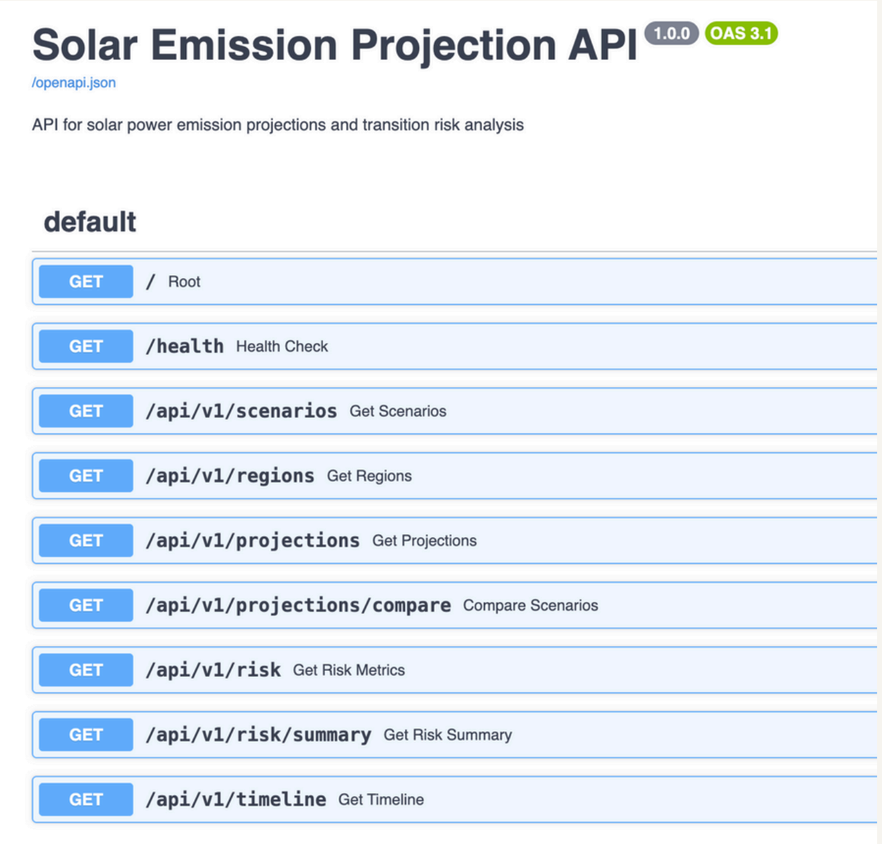


- We trained machine-learning models (XGBoost, LightGBM and Random Forest) with cross-validation to predict annual emissions avoided
- We incorporated scenario-specific features and projects emissions avoided for each year from 2025 to 2050 under NZE, APS and STEPS
- We Aggregated results by region to produce projections of capacity (MW), generation (MWh) and emissions avoided (tCO_2e).
- The models are versioned, saved and evaluated (R^2, RMSE) to ensure reproducibility and performance

Transition Risk Assessment

- We Calculated scenario divergences (e.g. NZE – APS, APS – STEPS) to quantify risk exposure
- The outputs are risk metrics for each region and year:
  - **Transition-risk score** (overall exposure)
  - **Policy-risk score** (sensitivity to policy changes)
  - **Technology-risk score** and **Market-risk score** (obsolescence and market competitiveness)
  - **Stranded-asset exposure** (potential carbon lock-in)
- These metrics guide investment decisions by identifying regions and scenarios with the highest risk of asset devaluation or regulatory pressure
- Risk metrics are saved alongside projection data for integrated analysis

API & Automation



- We used **FastAPI** service which provides endpoints for health checks, scenario and region listing, projection retrieval, scenario comparison and risk metrics
- It Returns JSON responses, enabling integration with dashboards and external systems
- **Automated schedules:** weekly data refresh, monthly model retraining and daily health checks; notifications can be configured
- We deployed in Docker containers with CI/CD for reliable, scalable operations

Value Proposition & Benefits

- **Data-driven insights:** scenario-aligned projections enable investors and planners to align portfolios with 1.5 °C pathways or evaluate impacts of current pledges
- **Risk mitigation:** transition-risk metrics reveal stranded-asset exposure and policy sensitivity, helping prioritise asset retirement and investment decisions
- **Reproducibility & scalability:** automated ingestion, feature engineering, modeling and API deployment ensure consistent results and support large datasets
- **Accessibility:** the API and dashboards democratise access to complex analytics for non-technical stakeholders
- **Future enhancements :** adding more scenarios, integrating real-time data and improving uncertainty analysis

Conclusion

As Climate Data Scientists, we :

- delivered an automated, scenario-based pipeline and transition-risk metrics for over 25 000 solar projects
- We ensured the integrity of the data pipeline thanks to robust data cleaning , developed and evaluated models, quantified risk and delivered actionable insights
- Our Project can be expanded to additional renewable technologies, can integrate real-time data feeds, enhance uncertainty quantification and collaborate with policy agencies

Thank you for your attention !