Solar Index Insurance – Reducing Basis Risk

Using *Monte Carlo* Pricing Methods

Naoki Conning, Stuyvesant High School

naokic3.github.io/solar

1. Data Source → 2. Output Modeling → 3. Forecast → 4. Contract Design ↔ 5. Monte Carlo Pricing

Introduction

The sun offers an inexhaustible source of clean energy, yet the world relies on fossil fuels for over 80% of its energy needs – at great environmental and economic cost.

Solar Energy: A Promising Solution

- **Pros:** Solar Photovoltaic (PV) is the most cost-effective way to generate electricity in many regions, according to the International Energy Agency. The versatility of PV systems—residential rooftops to solar farms—presents a unique opportunity to decentralize and democratize the energy grid.
- **Cons:** The upfront investment and inherent variability of solar power generation is a financial barrier to wider adoption, particularly for smaller adopters who lack the resources to evaluate these risks.

The goal of this project is to develop an open-source framework for end-to-end analysis of the design and pricing of location and system-specific solar power option contracts.

Why solar parametric insurance?

- **Solar power output is variable:** dependent on cloud cover and atmospheric conditions that vary greatly by day, time, location, and equipment.
- **Uncertainty reduction spurs adoption:** Solar projects require initial investments to generate expected, but risky returns. Adopters often fear being unable to meet debt obligations.
- **Parametric insurance advantages:** Unlike conventional indemnity insurance, requiring costly and time-consuming damage assessments, payouts are made quickly and automatically based on trigger events such as a solar irradiance index falling below a threshold. Based on available and verifiable data.
- **Disadvantages:** Solar panel power output is determined by many factors other than irradiance such as temperature, and panel type. Simple weather index insurance may therefore leave the insured exposed to considerable uncompensated 'Basis Risk'.

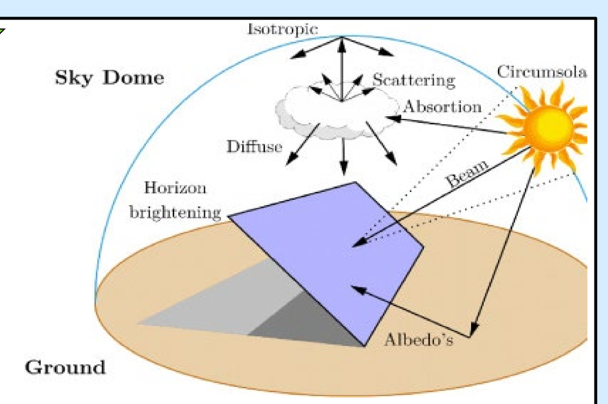
Improving parametric insurance

- **Multi-Variable Precision:** An index based on predicted output accounts for additional weather variables other than irradiance that factor into PV output and system specific variables.
- **Validated Models**: PVLIB (see box 2 and Holmgren, 2018) is a trusted, open source model for location and model specific output predictions developed by solar energy scientists.
- **Greater Accuracy:** Using a trusted dataset of Chinese solar farms (Yao et al, 2021) our PVLIB predicted power output achieved a **0.96** correlation with actual output compared to an irradiance-only index correlation of **0.89 (Fig 3).** A close relationship is also observed between predicted and actual power in **Fig 1 and 2.**
- **See the Proof:** The graphs to right demonstrate the close alignment between PVLIB's predicted output and actual measured output, over the course of several days, and from a solar station over a full year.

1. Public Data Sources

- Location specific numerical weather prediction (NWP) and/or satellite data from databases including:
  - **NSRDB** (National Solar Radiation Database)
  - **CAMS** (Copernicus Atmosphere Monitoring Service) Interpolated historical and forecast weather and irradiance data for the entire world

2. Output modeling



- **Prediction:** Location, weather and equipment data are fed into open-source PhotoVoltaic Library (PVLIB python) for physics/engineering based prediction of irradiance and power output over time.
- **Site tailored:** PVLIB offers choice of models for accurate analysis based on specific location and individual needs.
- **Ease of Use:** PVLIB is widely used; specifications pre-programmed

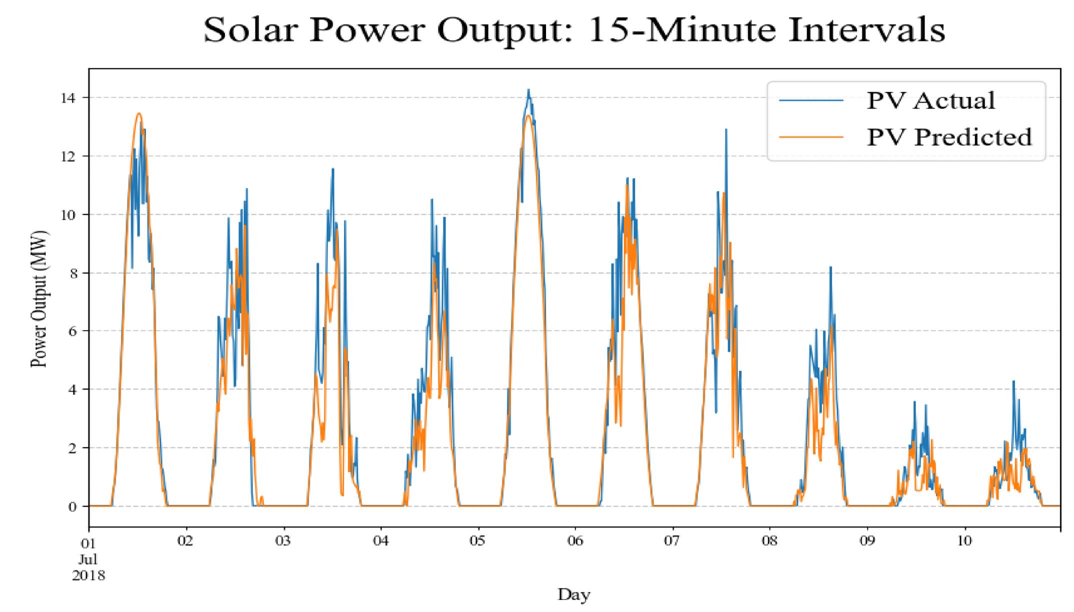


Fig 1. Actual and Predicted output for PVOD station 7 over 10 day span

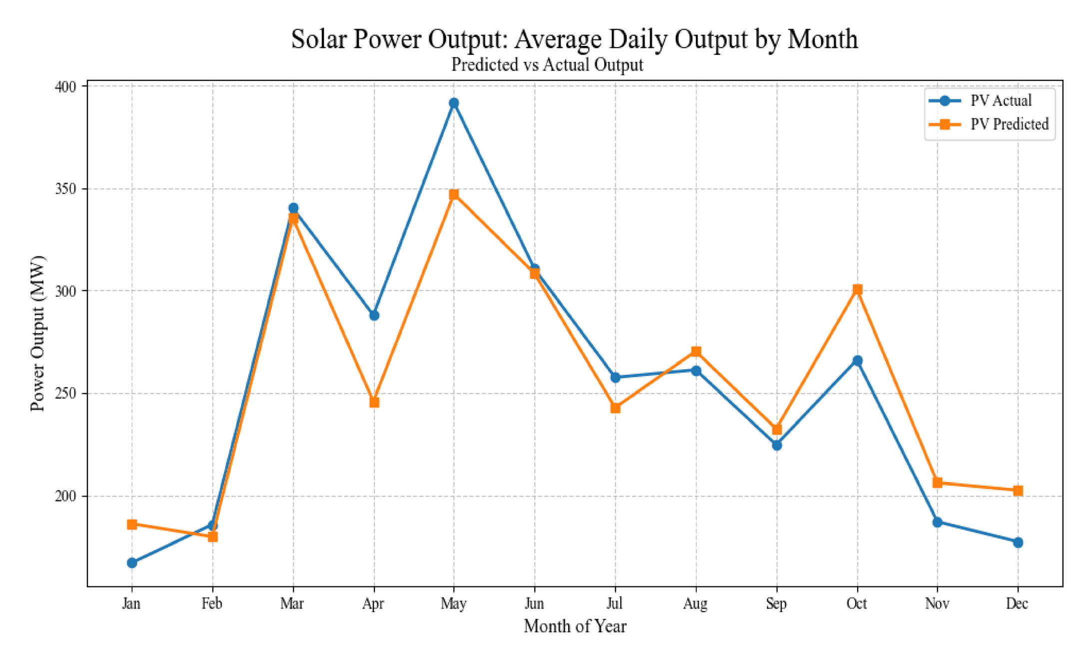


Fig 2. Actual and Predicted output for PVOD station 7; monthly average day

Fig 3. Correlation to Actual Output:

Global Irradiance	PVLIB Simulated
0.89	0.96

3. Time series forecasting

A time-series/machine learning prediction model is fitted to historical weather to model both these variables and PV output. The user is given a choice of models, or an ensemble method.

- **SARIMA** (Seasonal Auto-Regressive Integrated Moving Average)
  - lagged observations of feature, and residual error to predict future values
  - Accounts for **seasonality, trends, and auto correlation**
- **GARCH** (Generalized Autoregressive Conditional Heteroskedasticity)
  - Used to forecast non-constant volatility (such as weather)
  - Identifies volatility clusters
  - Uses historical volatility changes to predict future volatility
  - Can be combined with SARIMA.
- **Neural Net LSTM** (Long Short-Term Memory)
  - Often used for weather. Recognizes complex relationships
  - **Memory:** learning from past patterns
  - **Black Box:** Hard to understand underlying process
- **Open-source large weather models:**
  - The Weather Research and Forecasting (WRF) Model from the National Center for Atmospheric Research (NCAR) allows for simulation of a range of atmospheric variables. Whether this can be used at reasonable cost is a research question.

4. Contract Design

Contract: Purchased at price P for a payout if realized power output x falls below a threshold, $\underline{PV}$.

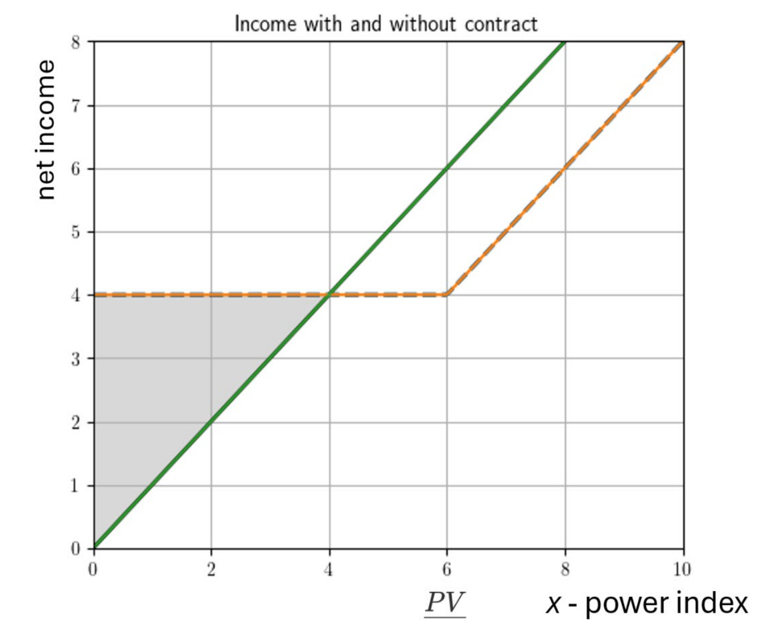


Fig 4. Income with and without contract (adjusted for premium)

Sample Contract (Fig 4): Trigger, $\underline{PV}$ = 6 and premium is 2 (Area of shaded part) (Assume *PV* output x is sold at $1 per unit).

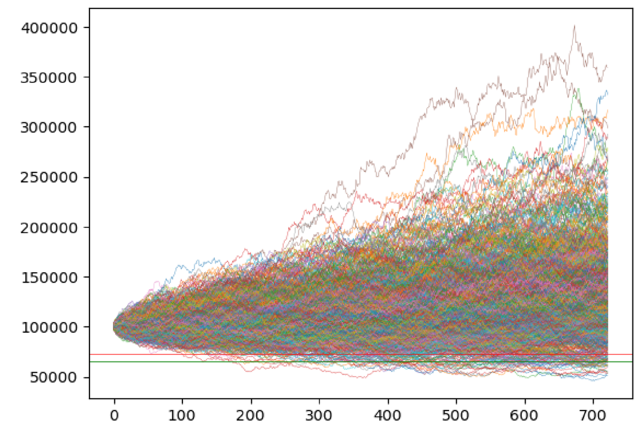
- The **solid** line shows income without contract
- The *dashed* line shows income with contract (premium subtracted, payout added)
- **Actuarily Fair Contract:** Expected value of the insurer's payouts equals expected value of premium, based on probability distribution of power generated incomes $f(x)$.

5. Monte Carlo contract pricing

- x: Solar Power output is a stochastic variablewith distribution $f(x)$ influenced by factors like sun position, panel specs, and stochastic atmospheric conditions.
- **$\underline{PV}$:** A set trigger price of output defined in the contract.
- An insurance contract promises to make a payout whenever x falls below $\underline{PV}$. The contract pays the difference between $\underline{PV}$ and x.
- In an actuarially fair (competitive) contract the price P equals the expected value of the payouts. This leaves the insured with the same expected but lower variance income.

$$P = \int_0^{\underline{PV}} (\underline{PV} - x) \cdot f(x) dx$$

- The complexity of the relationships that generate x imply we generally do not have a closed-form for stochastic distribution $f(x)$.
- Monte Carlo simulation involves generating a large number of random paths for predicted x based on the fitted model's parameters and estimated error distribution. This generates a probabilistic approximation to the distribution $f(x)$, for numeric integration.
- Many other forms of financial contracts can be priced this way.



Monte Carlo simulation (synthetic data)

Future Steps and research

- Methods for evaluating and selecting amongst time series models of output prediction (explore ensemble methods).
- Explore Weather Research Forecasting (large model) interface.
- Incorporating electricity price fluctuation modeling into contracts.
- Identifying geographical areas that most benefit from Insurance.
- Climate Change trend modeling

References

Carter, Michael et al. 2017. "Index Insurance for Developing Country Agriculture: A Reassessment." *Annual Review of Resource Economics* 9 (Volume 9, 2017): 421–38.

Holmgren, F. , et al. 2018. "Pvlib Python: A Python Package for Modeling Solar Energy Systems." *Journal of Open Source Software* 3 (29): 884.

Gibert, Jorge. 2024. "Monte Carlo Pricing of Temperature Weather Derivatives." Masters Thesis, Department of Mathematics and Statistics, Florida International University.

Yao, Tiechui et al. 2021. "A Photovoltaic Power Output Dataset: Multi-Source Photovoltaic Power Output Dataset with Python Toolkit." *Solar Energy* 230:122–30.