ONE-HOUR-AHEAD SOLAR PV POWER FORECASTING

A comparative study using ANN, Support Vector Regression and Random Forest

Simulated 100-kW grid-connected PV system

Engr. Dr. Mobolaji Agbolade Onasanya | Dr. Abiodun Oduniyi Odumosu

Department of Electrical and Electronic Engineering

Abraham Adesanya Polytechnic, Ijebu-Igbo, Ogun State, Nigeria

Why one-hour-ahead PV forecasting matters

- Solar PV output changes with irradiance and weather, making generation intermittent.
- Short-horizon forecasts support microgrid energy management, battery scheduling and grid integration.
- Forecast errors can affect dispatch decisions, reserve planning and the use of stored energy.
- Machine learning can learn nonlinear relationships between weather, recent PV output and future generation.
- The practical question: which model performs most reliably on a small hourly PV dataset?

Study focus: next-hour PV power, using information available at the current hour.

Study aim and contributions

Aim

Compare ANN, SVR and Random Forest for one-hour-ahead forecasting of a simulated 100-kW grid-connected PV system.

01 Forecasting framework

Chronological train/validation/test split and training-only standardisation.

02 Model comparison

Evaluate accuracy, training time and characteristic forecast errors.

03 Critical interpretation

Explain model behaviour and identify what is needed before real-world generalisation.

Simulated PV system and dataset

PV system

100-kW grid-connected system

Simplified PV conversion model

Dataset

60 days × 24 hours

1,440 hourly samples

Forecast target

PV power at hour t+1

Inputs available at hour t

Weather variables represented

GHI	Clear-sky daily profile; peak near 1,000 W/m^2
Temperature	Approximately 22–32 °C
Relative humidity	Approximately 30–95%
Wind speed	Typically 0–5 m/s

Simulated output peaks at roughly 85–90 kW, below rated capacity due to temperature derating and operational losses.

PV power model and feature engineering

Simplified PV model

$$T_{cell} = T_{amb} + 0.025\,G$$

$$\eta_{temp} = 1 + \gamma\,(T_{cell} - T_{ref})$$

$$P_{pv} = P_{rated} \times (G / 1000) \times \eta_{temp} \times \eta_{loss}$$

Parameters

$\gamma = -0.0045\ /°C$ | $T_{ref} = 25\ °C$ | $\eta_{loss} = 0.95$ | $P_{rated} = 100\ kW$

10 model inputs

- Current-hour irradiance, temperature, humidity and wind speed
- Previous-hour PV power and irradiance
- Hour-of-day sine and cosine
- Day-of-year sine and cosine

Cyclic encoding keeps adjacent time points (e.g., 23:00 and 00:00) close in feature space.

Forecasting workflow and leakage control

Hourly weather + PV
1,440 samples

Feature engineering
10 inputs

Chronological partition
70 / 15 / 15

Train models & validate
MATLAB

Test & compare
RMSE, MAE, MAPE, R^2



Training: 1,008 h | Validation: 216 h | Test: 216 h

Means and standard deviations for scaling were calculated from training data only; the test period remained held out.

Three machine-learning models

Artificial Neural Network (ANN)

Feedforward network: 10–20–10–1

Levenberg–Marquardt training

Maximum 500 epochs

Early stopping; 20 validation failures

Support Vector Regression (SVR)

Radial basis function (RBF) kernel

Box constraint = 10

ε = 0.5

Automatic kernel scaling

Random Forest (RF)

Bagged ensemble of 100 regression trees

Minimum leaf size = 5

Out-of-bag prediction enabled

All models were implemented in MATLAB. Hyperparameters were fixed a priori; no systematic search was performed.

How forecasting performance was measured

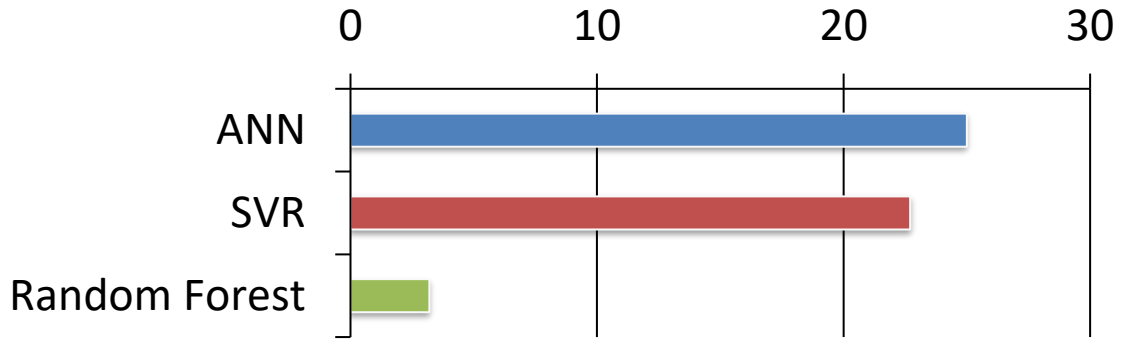
RMSE	Penalises larger errors; expressed in kW.
MAE	Average absolute forecast error; expressed in kW.
MAPE	Percentage error; denominator floored at 1 kW to handle night-time zeros.
R^2	Proportion of test-set variance explained by the forecasts.
Training time	Practical computational cost, recorded in seconds.

Primary interpretation relies on RMSE, MAE and R^2. MAPE is sensitive to the 1-kW floor when actual PV output is zero or near zero.

Test-set results: Random Forest leads clearly

Model	RMSE (kW)	MAE (kW)	R^2	MAPE (%)	Training time (s)
ANN	≈25	≈15.3	≈0.23	≈140	≈4.2
SVR	≈22.7	≈20.3	≈0.36	≈870	≈0.4
Random Forest	**≈3.2**	**≈2.0**	**≈0.98**	**≈10**	**≈1.8**

RMSE comparison (lower is better)



RF's RMSE is roughly 7–8 times lower than ANN and SVR, and its R^2 is approximately 0.98.

Forecast behaviour: where each model struggled

ANN | Peak overshoot

Tracks the broad daily cycle, but overshoots at peak hours—sometimes reaching or exceeding 100 kW—and creates dawn/dusk spikes.

SVR | Compressed output

Produces a narrow band of roughly 10–40 kW: too high at night and too low around midday.

Random Forest | Close tracking

Follows the daily bell-shaped profile, returns to zero at night and slightly under-predicts the highest peaks.

Error definition in the paper: actual power − forecast power. Negative error therefore indicates over-prediction.

RF errors were approximately symmetric and centred near zero; most were within about ±3 kW.

Interpreting the ranking—and its limits

Why RF performed best in this setup

- Previous-hour PV power and current irradiance strongly inform next-hour output.
- Tree partitions can represent sunrise/sunset transitions and the zero-output night regime.
- Tree ensembles are competitive on small, tabular datasets.

Important limitations

- Synthetic data: no realistic cloud transients, shading, soiling, clipping or measurement noise.
- Only 60 days; the test period is about nine days; no confidence intervals.
- No persistence baseline; ANN/SVR were not systematically tuned.
- No feature-importance analysis was reported.

The ranking is specific to the dataset and configurations tested; it does not prove ANN or SVR are generally unsuitable.

CONCLUSION & NEXT STEPS

Random Forest was the strongest model in this experiment.

Test performance: RMSE ≈ 3.2 kW | MAE ≈ 2.0 kW | R^2 ≈ 0.98

Recommended next steps

- Validate with measured PV and meteorological data from a Nigerian site.
- Add persistence and other baseline forecasts.
- Tune ANN and SVR fairly using validation data.
- Use a longer record spanning wet and dry seasons with rolling-origin evaluation.
- Assess feature importance and additional forecast horizons.

Thank you | Questions?