

# Generating realistic solar irradiation scenarios with limited historical data

Robert Hanusa, Birger Raa, Stijn De Vuyst

Ghent University, Industrial Systems Engineering

`robert.hanusa@ugent.be`, `birger.raa@ugent.be`, `stijn.devuyst@ugent.be`

5 December, 2025

As companies increasingly look to integrate on-site production of variable renewable energy (VRE) in their production planning, it becomes important to use models that take into account the uncertainty of VRE forecasts. This research is inspired by the e-CODUCT project [1] focuses on the case where an industrial user with a photovoltaic (PV) array wants to plan their energy usage for the coming day based on a point forecast. To address uncertainty in the forecast, multiple scenarios can be generated that reflect different possible realizations. We investigate a case in which the PV array is new, or that the company is in the planning phase and simulating performance and therefore possesses very limited historical data for their specific location. While historical weather measurements may be available from nearby stations, the corresponding historical PV forecast data either does not exist or is only available in spatially aggregated form. Aggregated PV forecasts behave differently from actual location-specific PV production, as small installations are more sensitive to localized cloud passages and therefore exhibit substantially higher intermittency.

Given this situation where there is limited or no historical forecast data, but a point forecast for the upcoming day is available, it is clear that several scenario generation approaches commonly used in other research [2] are not suitable:

- Deep Generative Models (DGMs): these methods require large volumes of training data and are therefore not appropriate for a data-scarce setting.
- Dependence-modelling via Copulas: Copulas normally require both historical forecasts and actuals. Without this, forecast errors cannot be estimated reliably, making standard copula-based approaches difficult to apply.
- Time-series generative models (e.g., ARIMA): These models generally generate future values using past observations and are therefore unsuitable for generating scenarios that reflect the next day's forecast.

Other studies face a similar constraint and therefore generate scenarios directly around a point forecast. These approaches typically rely on additive-noise models, such as sampling from parametric distributions (e.g., Beta or Gaussian) fitted to residuals. However, the resulting scenarios end up with a much higher point-to-point difference (average of the absolute differences between adjacent

points) than the original forecast, leading to scenarios that do not reflect the properties of the forecast well.

In this research, we propose a technique we call the Gamma Shift in which we use random numbers from the Gamma distribution to vary the forecast on the time axis rather than the magnitude axis. This has the benefits that (1) the point-to-point difference of the scenarios is similar to the original forecast, leading to more realistic scenarios and (2) the new scenarios reflect realistic cases where, for example, a cloud passes overhead at a different time than originally forecast, which is a very real possibility.

Figure 1 shows an example of how three scenarios would look with a standard additive-noise technique (left) and with the Gamma Shift technique (right). For the additive-noise scenarios, we have taken an example instance and applied additive noise from a Gaussian distribution. Similar results would be seen for other probability distributions as long as the average is zero and the samples are independent and identically distributed. The forecast (solid black line) is shown against three scenarios generated using additive noise (dashed lines) from a normal distribution with mean of 0 and standard deviation  $50W/m^2$ . Visually we see that forecast is a much smoother curve than the scenarios, which appear very choppy.

For the Gamma Shift scenarios on the right, we see that the three scenarios (dashed lines) are shifted on the time domain, leading to curves with a smoothness, as measured by point-to-point difference, similar to that of the forecast's (solid black line).

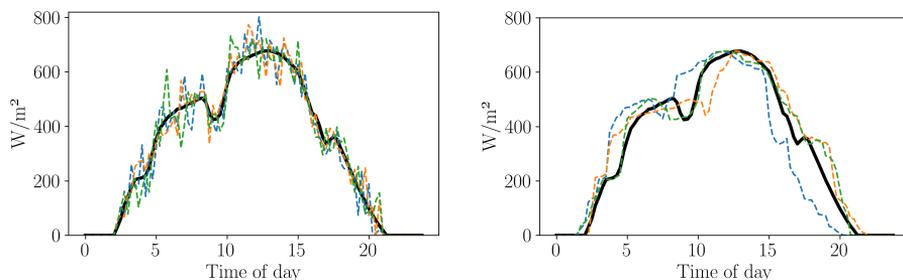


Figure 1: Scenarios generated with additive noise (left) and the Gamma Shift (right). Daylight hours extended from 2:00 to 22:00 for visual effect.



The e-CODUCT project is funded under Horizon Europe Grant Agreement n°101058100.

## References

- [1] European Commission. Horizon europe grant agreement no. 101058100. <https://e-coduct.eu/>, 2025.
- [2] Markos A. Kousounadis Knousen, Ioannis K. Bazionis, Athina P. Georgilaki, Francky Catthoor, and Pavlos S. Georgilakis. A review of solar power scenario generation methods with focus on weather classifications, temporal horizons, and deep generative models. *Energies*, 16(15), 2023.