Realistic but not Real: A Framework to Generate Synthetic Timeseries Data for Grid Studies

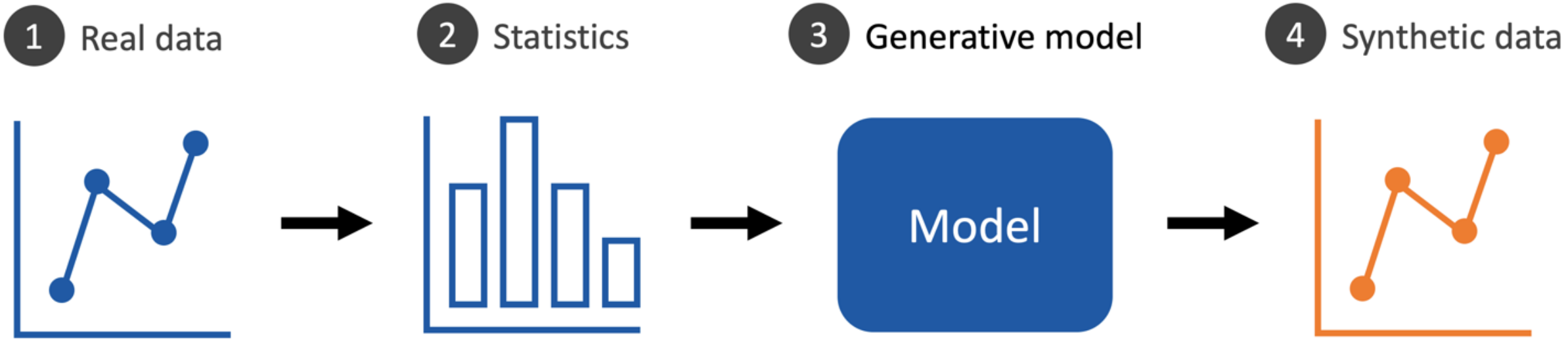
David Larson, EPRI

24PESGM1065

Why synthetic timeseries data?

- Need timeseries data for many power system studies
- But data may not be readily available
- Many studies only need data that is **representative of conditions of interest**
- **Idea:** synthesize data that is statistically realistic, but not real

Key steps in generating synthetic timeseries data:



Case Study

Synthetic Forecasts

Problem setup

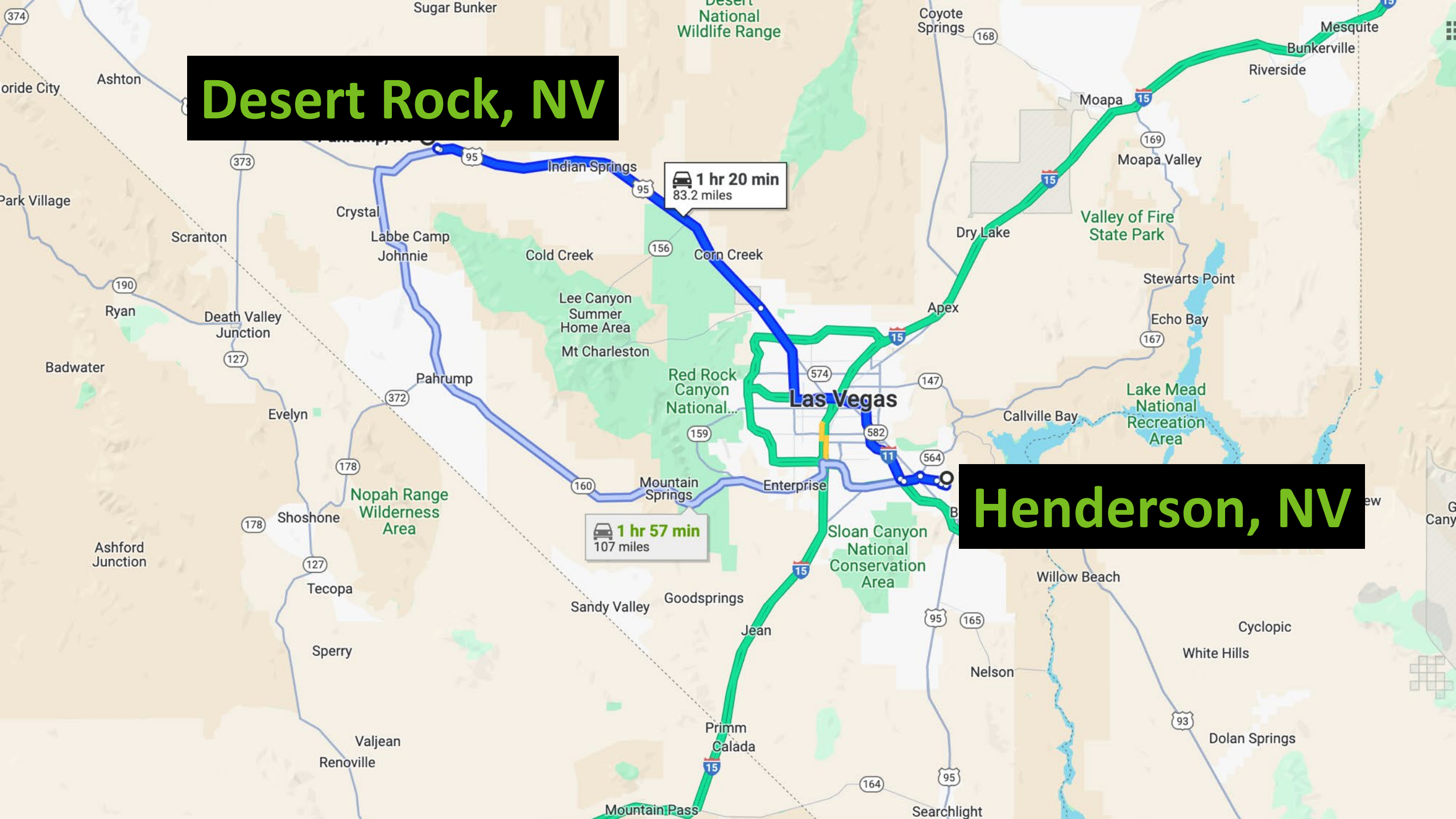
- **Given:**
  - Real forecasts for **Site A**
  - Real observations for **Site A**
  - Real observations for **Site B**

- **Goal:** generate synthetic forecasts for **Site B**
  - where the synthetic forecasts for **Site B** have similar forecast errors as **Site A**

Error = Forecast - Observation

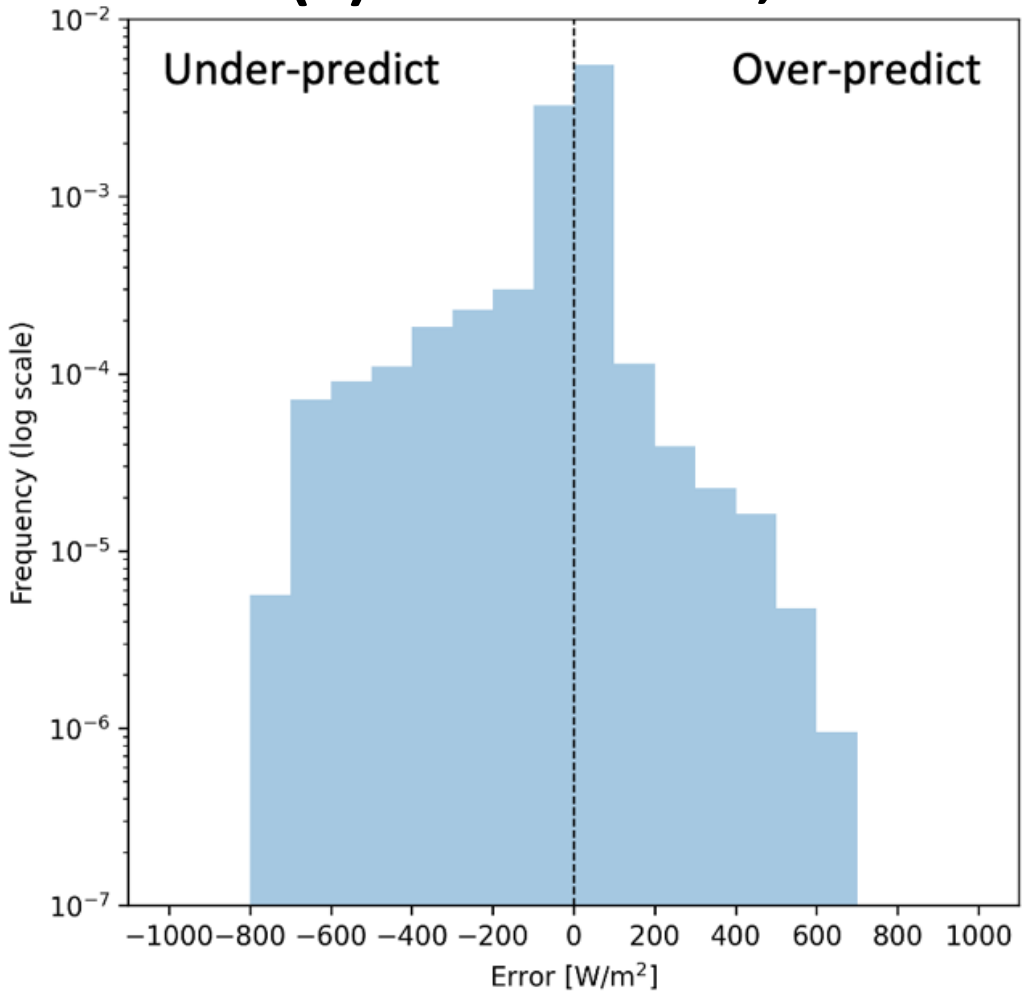
Evaluated two locations

	Desert Rock, NV	Henderson, NV
Latitude [°N]	36.62	36.04
Longitude [°E]	-116.02	-114.92
Elevation [m]	1,007	538
Variable	GHI [W/m^2]	GHI [W/m^2]
Observations	NOAA SURFRAD	DOE RTC
Forecasts	NOAA GFS	NOAA GFS
Time resolution	Hour-ending averages	Hour-ending averages
Time range	May 2019–August 2020	April 2019–June 2020

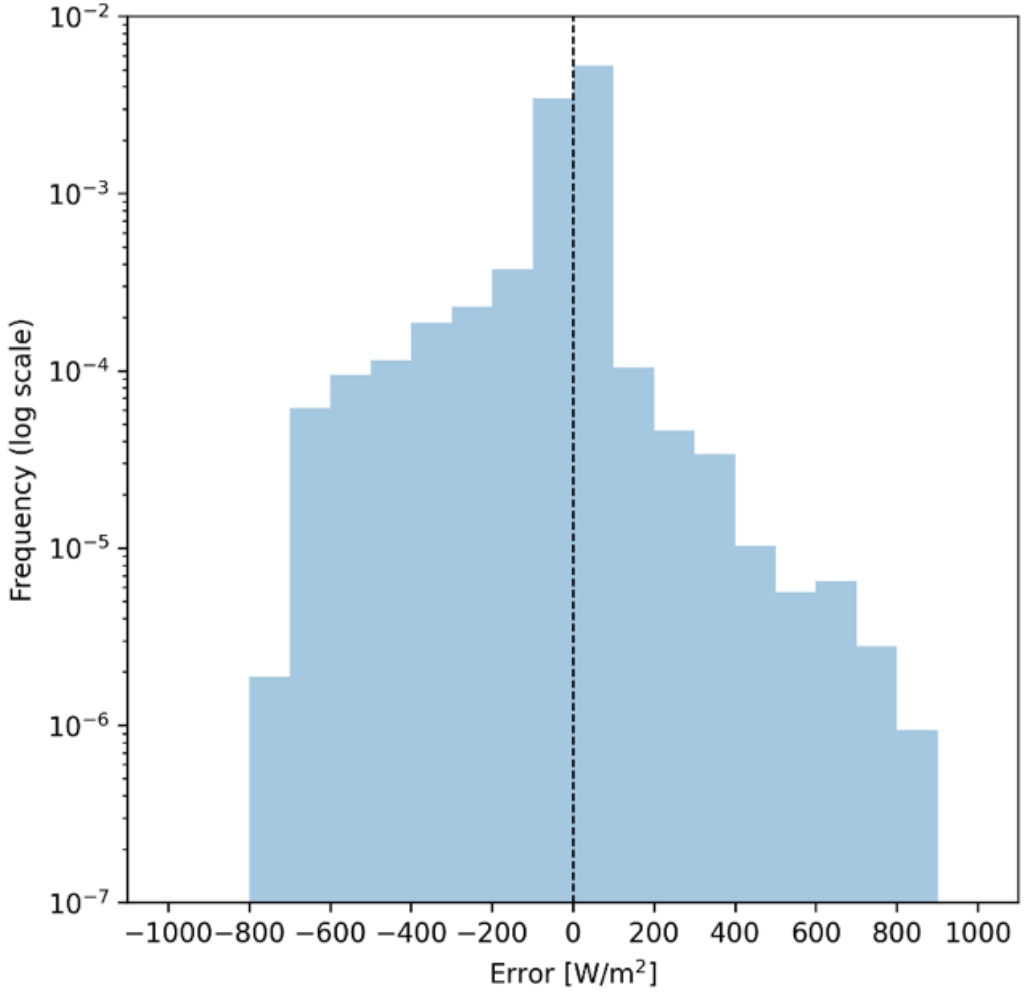

Desert Rock, NV
Henderson, NV
1 hr 20 min
83.2 miles
1 hr 57 min
107 miles
Las Vegas
Indian Springs
Pahrump
Crystal
Labbe Camp
Johnnie
Cold Creek
Corn Creek
Lee Canyon
Summer
Home Area
Mt Charleston
Red Rock
Canyon
National...
Mountain
Springs
Enterprise
Sloan Canyon
National
Conservation
Area
Nopah Range
Wilderness
Area
Sugar Bunker
Desert
National
Wildlife Range
Coyote
Springs
Mesquite
Bunkerville
Riverside
Moapa
Moapa Valley
Valley of Fire
State Park
Dry Lake
Apex
Stewarts Point
Echo Bay
Lake Mead
National
Recreation
Area
Callville Bay
Willow Beach
Cyclopic
White Hills
Dolan Springs
Nelson
Searchlight
Jean
Goodsprings
Sandy Valley
Primm
Calada
Mountain Pass
Valjean
Renoville
Sperry
Tecopa
Ashford
Junction
Shoshone
Evelyn
Badwater
Death Valley
Junction
Ryan
Park Village
Scranton
Ashton
oride City
Pahrump
Ashton
Sugar Bunker
Desert
National
Wildlife Range
Coyote
Springs
Mesquite
Bunkerville
Riverside
Moapa
Moapa Valley
Valley of Fire
State Park
Dry Lake
Apex
Stewarts Point
Echo Bay
Lake Mead
National
Recreation
Area
Callville Bay
Willow Beach
Cyclopic
White Hills
Dolan Springs
Nelson
Searchlight
Jean
Goodsprings
Sandy Valley
Primm
Calada
Mountain Pass
Valjean
Renoville
Sperry
Tecopa
Ashford
Junction
Shoshone
Evelyn
Badwater
Death Valley
Junction
Ryan
Park Village
Scranton
Ashton
oride City
Pahrump
Crystal
Labbe Camp
Johnnie
Cold Creek
Corn Creek
Lee Canyon
Summer
Home Area
Mt Charleston
Red Rock
Canyon
National...
Mountain
Springs
Enterprise
Sloan Canyon
National
Conservation
Area
Nopah Range
Wilderness
Area
Las Vegas
Indian Springs
1 hr 20 min
83.2 miles
1 hr 57 min
107 miles
Desert Rock, NV
Henderson, NV

Similar (real) forecast error distributions

(a) Desert Rock, NV

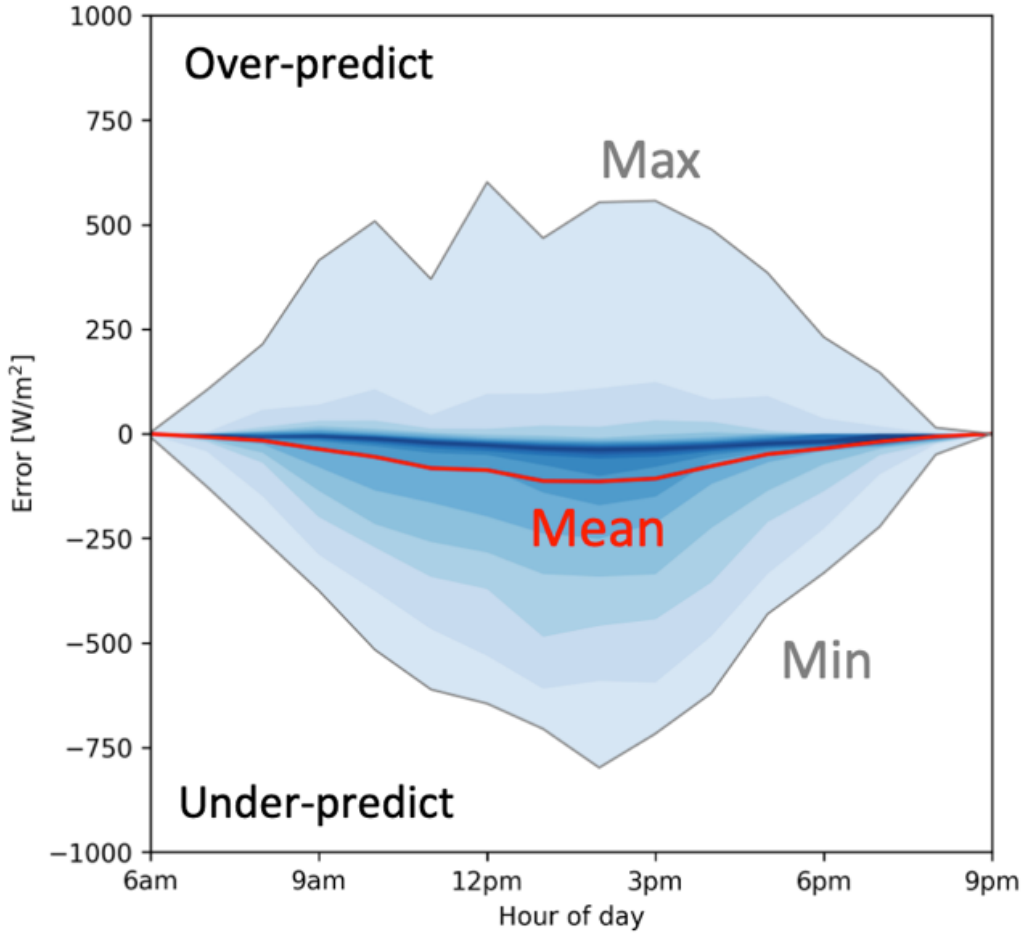


(b) Henderson, NV

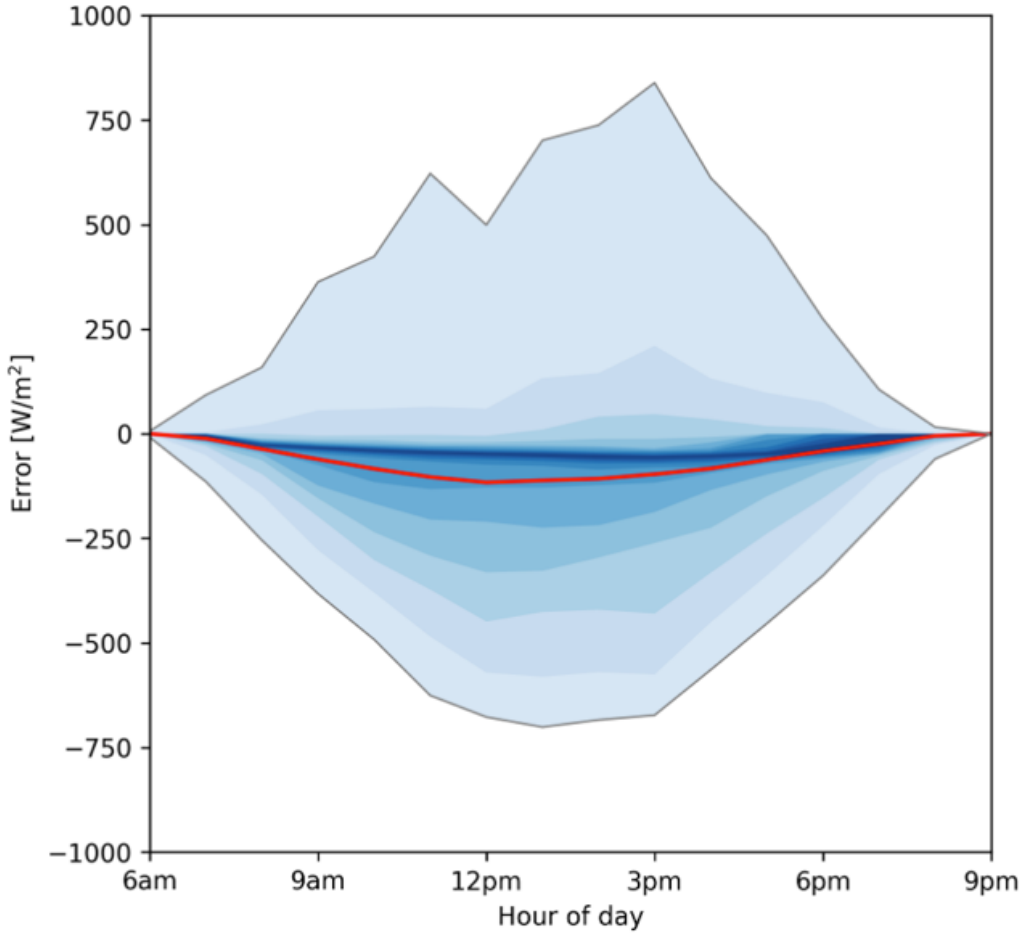


Similar (real) forecast error distributions

(a) Desert Rock, NV



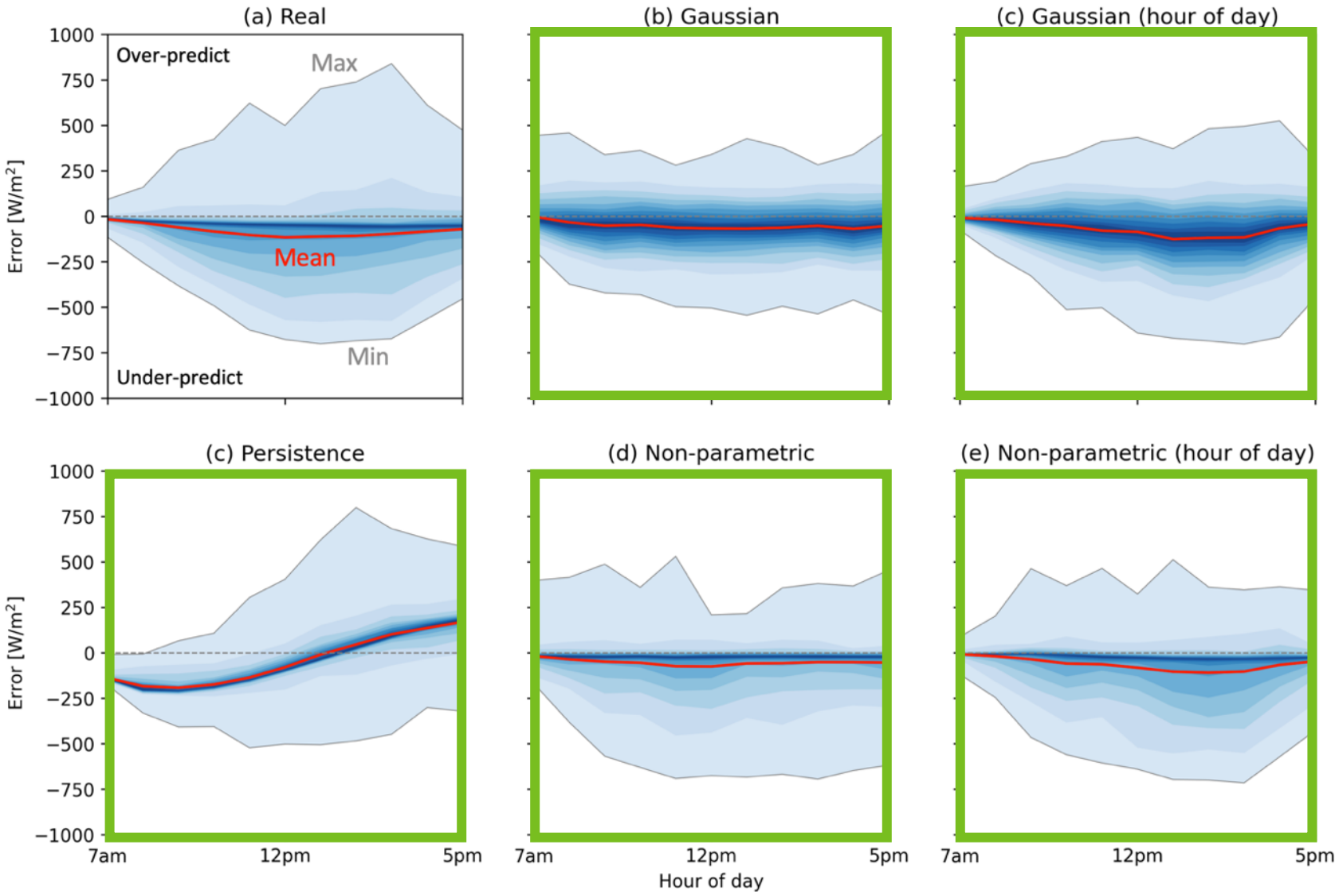
(b) Henderson, NV



Comparing statistics (real vs synthetic)

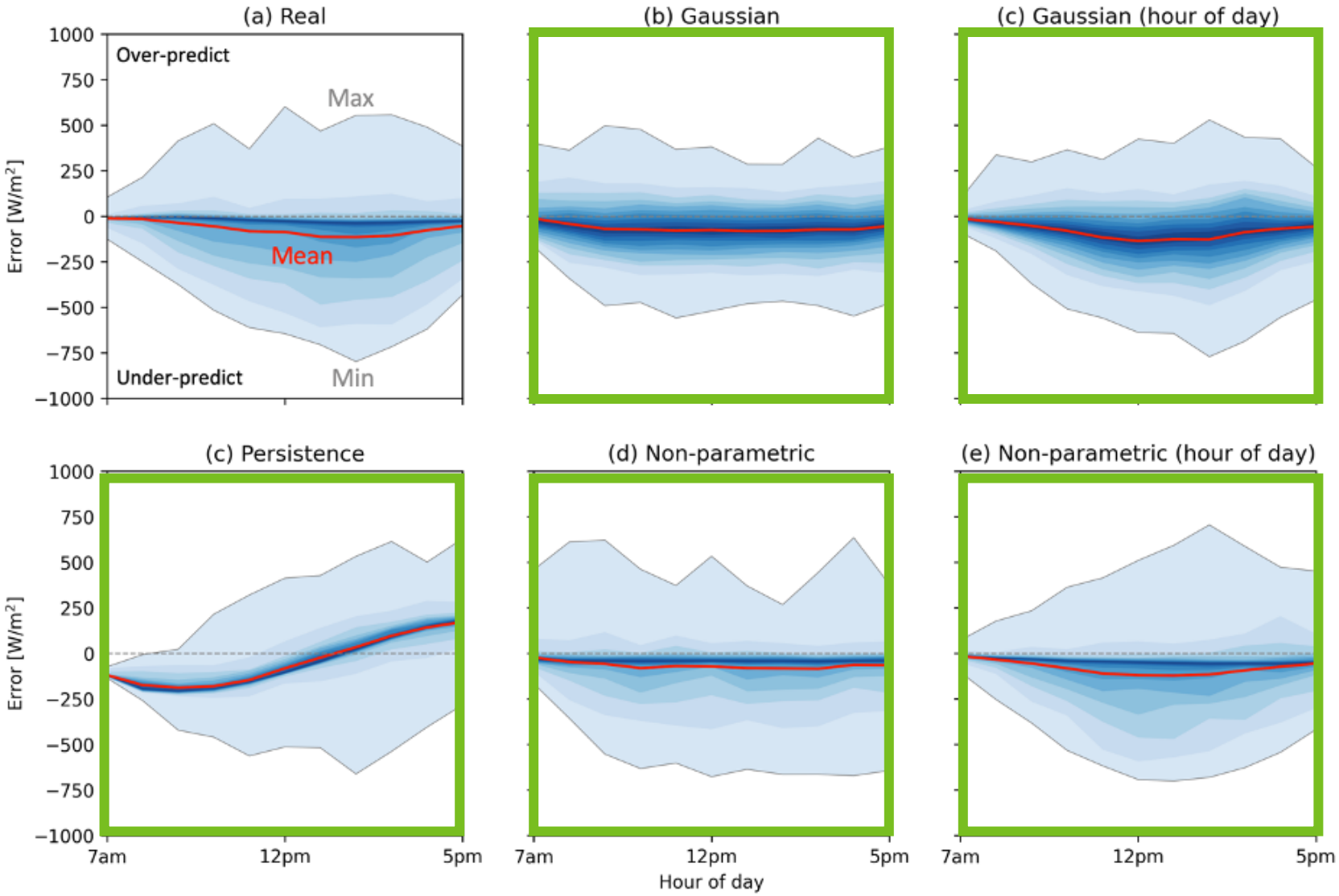
Forecast data source	RMSE [W/m^2]	MBE [W/m^2]
Persistence forecast	**165.6**	-1.1
Parametric distribution	147.5	-49.0
Parametric distribution (hour of day)	158.1	**-63.7**
Non-parametric distribution	132.7	-50.5
Non-parametric distribution (hour of day)	149.5	-59.5
Real forecast	**169.8**	**-76.1**

Train on Henderson, test on Desert Rock



Synthetic forecast errors

Train on Desert Rock, test on Henderson



Synthetic forecast errors

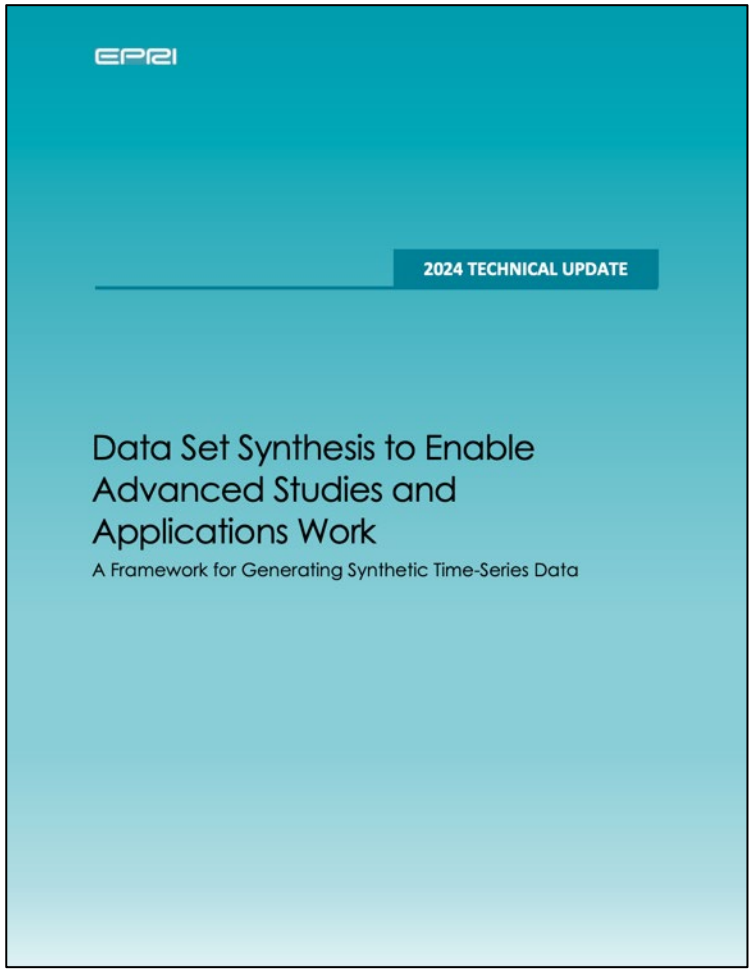
Key takeaways

- **Generating synthetic timeseries is a practical solution** for cases where real data is not available, but data needs only to be statistically realistic

- **Distribution-based methods** can generate synthetic forecast data which has statistically realistic forecast errors

- **Validating synthetic data** should consider both statistical analysis and domain expertise

EPRI report 3002024648



https://www.epri.com/research/products/000000003002024648

Questions?

DLarson@epri.com