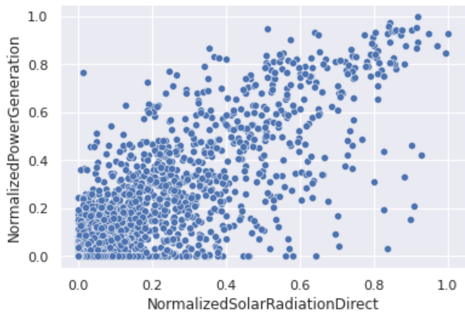
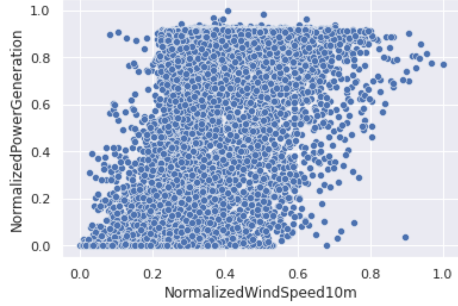
1. Supplementary Material

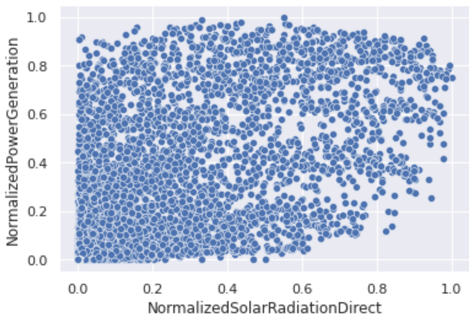
1.1. Scatter Plots



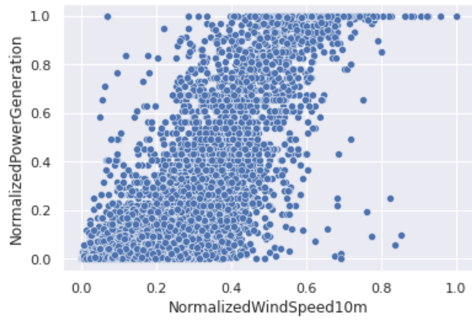
(**a**) Pv Open Dataset



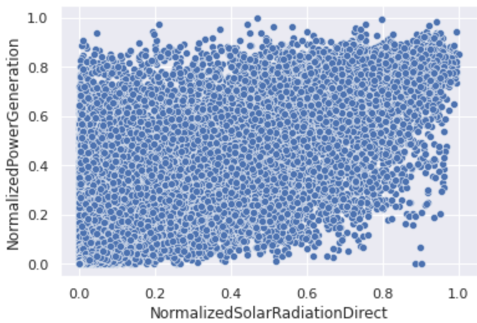
(**b**) Wind Open Dataset



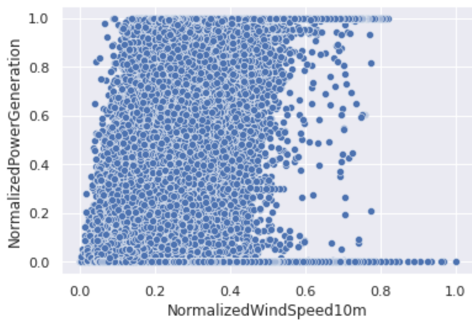
(**c**) Pv Synthetic Dataset



(**d**) Wind Synthetic Dataset



(**e**) Pv Real Dataset



(**f**) Wind Real Dataset

Figure S1. Scatter plots of solar radiation and wind speed features with historical power measurements of a sample park from all PV and wind datasets.

1.2. Hyperparameter Tuning of STL Methods

1.2.1. GBRT

For training GBRT models on all parks, we considered both *number of estimators* and *minimum number of samples required to split an internal node* to 200. The learning rate and maximum depth hyperparameters are tuned.

Hyperparameter	Values
Learning rate	0.1, 0.01, 0.001, 0.0001
Max depth of individual regression estimators	4, 6, 8

1.2.2. NN

The architecture of STL NN models trained on each park is as follows:
Input Layer $\rightarrow$ Layer 1 Units $\rightarrow$ 50 $\rightarrow$ 50 $\rightarrow$ 50 $\rightarrow$ 50 $\rightarrow$ 50 $\rightarrow$ 1 The optimizer is adam and activation function for hidden layers is leaky relu (alpha = 0.2).

Hyperparameter	Values
Epochs	20, 40, 60
Batch Size	64, 128
Layer 1 Units	25, 50

1.2.3. AE-NN

In this STL model, initially AEs are trained on each park separately and the architecture of AE is as follows:
Input Layer $\rightarrow$ Bottleneck Encoding Layer $\rightarrow$ Output Layer
The optimizer is adam and activation function for hidden layers is relu.

Hyperparameter	Values
Epochs	20, 40, 60
Batch Size	64, 128
Encoding Dimensions	5, 6, 7, 8

The encoder thus trained is used to reduce the dimensions of input data. The encoded data is used as input to NN. We utilize the best parameters tuned for individual NNs in above STL NN models built for each park.

1.3. Hyperparameter Tuning of MTL Methods

1.3.1. UAE, UCAE, ULAE:

A unified autoencoder is trained for each dataset separately and the structure of it is as follows:

Input Layer $\rightarrow$ Bottleneck Encoding Layer $\rightarrow$ Output Layer

The same structure is utilized for Unified Convolutional Autoencoder (UCAE) and Unified LSTM Autoencoder (ULAE). The optimizer is adam and activation function for hidden layers is relu. The encoding dimensions are hyperparameter tuned with different values for different datasets. The encoding dimensions values for grid search are considered with the objective to reduce input features to half. For example, the number of input features in PS dataset is 14, so the encoding dimensions considered for grid search are nearer to 7.

Table S1. Hyperparameters tuned for unified autoencoder models - UAE, UCAE, ULAE on six different datasets

Hyperparamter	Values
Epochs	20, 40, 60, 80
Batch Size	16, 64
Encoding Dimensions (PO)	19, 20, 21
Encoding Dimensions (WO)	3, 4
Encoding Dimensions (PS)	5, 6, 7, 8
Encoding Dimensions (WS)	6, 8, 10, 12
Encoding Dimensions (PR)	7, 9, 11
Encoding Dimensions (WR)	13, 15

1.3.2. UAE-NN:

Here, the input layer is the encoded dimensions layer followed by a NN architecture as follows:

Encoded Inputs Layer $\rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 1$ The optimizer is adam and activation function for hidden layers is leaky relu (alpha = 0.2). The epochs and batch size considered are 20 and 128 respectively.

1.3.3. TE-NN:

This MTL architecture has task embedding layer to concatenate with the inputs and then passed into the neural network. The architecture is as follows:

Inputs + Task Embeddings $\rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 1$

Hyperparameter	Values
Epochs	10, 16, 20, 40
Batch Size	64, 128

A two dimensional task embedding is considered, and it is concatenated with inputs. The optimizer is adam and activation function for hidden layers is leaky relu (alpha = 0.2).

1.3.4. UAE-TENN, UCAE-TENN, ULAE-TENN:

This architecture similar to TE-NN, except the inputs considered are encoded inputs from the respective UAE, UCAE, ULAE trained. The architecture is as follows:

Encoded Inputs + Task Embeddings $\rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 50 \rightarrow 1$

The best hyperparameters of UAE-TENN method for each dataset along with input and task embedding dimensions are displayed in below table.

Dataset	Input dims	UAE dims	TE dims	Epochs	Batch Size
PO	51	21	2	40	128
WO	7	4	2	60	64
PS	14	5	2	60	64
WS	25	12	2	20	64
PR	19	9	2	40	64
WR	27	13	2	40	64

1.4. WS parks indices

As mentioned in Section 4.1 the PV Synthetic (PS) dataset has 118 parks and to maintain the consistency, same number of parks are randomly chosen from Wind Synthetic (WS) data of 263 parks. The indices of these randomly chosen parks are provided here for reproducing the results.

Indices: [2377, 3098, 4911, 7368, 3821, 10510, 5906, 7403, 6197, 7341, 164, 4880, 5839, 15207, 13932, 2667, 3668, 4466, 5800, 5792, 15000, 6163, 840, 90, 15444, 5930, 4280, 15044, 1869, 1759, 1975, 5538, 7392, 1803, 2559, 13676, 3631, 1078, 5158, 5109, 15520, 3402, 7351, 3513, 3268, 460, 13674, 5404, 5797, 2597, 7391, 2349, 1011, 7395, 3366, 3196, 4887, 1490, 3028, 7389, 1013, 2429, 5480, 433, 2794, 5142, 953, 6107, 5516, 2712, 853, 5078, 603, 1832, 3231, 4464, 642, 6096, 5029, 1379, 3093, 1200, 619, 3167, 691, 7394, 6091, 5412, 1550, 13965, 2252, 5349, 4336, 3811, 2573, 2907, 3086, 880, 5440, 1886, 6097, 867, 4024, 2564, 11, 7412, 5825, 722, 6099, 6253, 4032, 1639, 1346, 6211, 5347, 5629, 1303, 4745]