APPENDIX

TABLE 1
LSTM MODEL PARAMETERS

Parameter	Value	Parameter	Value
Number of layers	3	Hidden units	64
Output dimension	1	Dropout rate	0.1
Activation (output)	ReLU	Optimizer	Adam
Learning rate	0.001	Batch size	32
Maximum epochs	100	Loss function	MSE
Sequence length	24	Prediction horizon	1 hour

TABLE 2
GRU MODEL PARAMETERS

Parameter	Value	Parameter	Value
Number of layers	3	Hidden units	64
Output dimension	1	Dropout rate	0.1
Activation (output)	ReLU	Optimizer	Adam
Learning rate	0.001	Batch size	32
Maximum epochs	100	Loss function	MSE
Sequence length	24	Prediction horizon	1 hour

TABLE 3
BILSTM MODEL PARAMETERS

Parameter	Value	Parameter	Value
Number of layers	3	Hidden units (per direction)	64
Total hidden output	128	Bidirectional	True
Output dimension	1	Dropout rate	0.1
Activation (output)	ReLU	Optimizer	Adam
Learning rate	0.001	Batch size	32
Maximum epochs	100	Loss function	MSE
Sequence length	24	Prediction horizon	1 hour

TABLE 4
BIGRU MODEL PARAMETERS

Parameter	Value	Parameter	Value
Number of layers	3	Hidden units (per direction)	64
Total hidden output	128	Bidirectional	True
Output dimension	1	Dropout rate	0.1
Activation (output)	ReLU	Optimizer	Adam
Learning rate	0.001	Batch size	32
Maximum epochs	100	Loss function	MSE
Sequence length	24	Prediction horizon	1 hour

TABLE 5
Q-LEARNING LAYER PARAMETERS

Parameter	Value	Parameter	Value
Hidden layers	[128, 128]	Activation	ReLU
Learning rate (α)	0.001	Discount factor (γ)	0.95
Initial epsilon (ε)	0.1	Final epsilon (ε)	0.01
Batch size	64	Optimizer	Adam
Reward weights (λ, η, ζ)	(1000, 5, 20)	State/Action dim	(14, 3)
Step size (δ)	0.005	Replay buffer size	10000

TABLE 6
PPO LAYER PARAMETERS

Parameter	Value	Parameter	Value
Hidden layers	[128, 128, 128]	Activation	ReLU
Actor learning rate	3×10^{-5}	Critic learning rate	5×10^{-4}
Clip ratio (ϵ)	0.3	Discount factor (γ)	0.99
GAE lambda (λ)	0.95	Entropy coefficient	0.2 Val
ue loss coefficient	0.5	Optimizer	Adam Dive
rsity weight (α)	0.2	Stability weight (β)	0.005

Using a rolling forecasting under a strict chronological data split.
Dataset is divided into 80% for training and 20% for testing without any random reshuffling.
A sliding window of 24 historical observations with a one-hour step to predict the next-hour PV output.
Model parameters remain fixed after training, and forecasts are produced sequentially on the test set.