I. Appendix: Hyper-parameter Tuning

TABLE I
BHT-ARIMA [1] model hyper-parameter grid

Model parameter	Values
order of auto-regressive model p	$1, 2$
order of moving average model q	$1, 2$
order of difference d	$1, 2$
ranks of MDT tensorization τ	$4, 5$
maximum iterations K	$10, 30$
convergence threshold tol	$10^{-2}, 10^{-3}$
orthogonality mode	full

TABLE II
SVR [2] model hyper-parameter grid

Model parameter	Values
kernel type	linear, polynomial, RBF
C	$1, 3$
γ	auto, scale

TABLE III
Random forest [3] model hyper-parameter grid

Model parameter	Values
number of trees	$10, 25, 50$
number of features to consider	auto, sqrt, log2
max. depth	$10, 20, 30$
min. samples split	$5, 10, 15$
min. samples leaf	$2, 5, 10$

TABLE IV
LSTM and Bi-LSTM model hyper-parameter grid

Model parameter	Values
rolling window size S	$L_{tr} \cdot 0.2$, $L_{tr} \cdot 0.4$, $L_{tr} \cdot 0.6$
network units	$100, 150$
number of epochs	$500, 1000$
batch size	$\lceil S \cdot 0.5 \rceil$, $\lceil S \cdot 0.7 \rceil$
loss function	MAPE
optimizer	Adam
activation function	ReLU

TABLE V
CNN model hyper-parameter grid

Model parameter	Values
rolling window size S	$L_{tr} \cdot 0.2$, $L_{tr} \cdot 0.4$, $L_{tr} \cdot 0.6$
number of filters	$32, 64$
number of kernels	$3, 5$
maximum pooling	2
number of epochs	$500, 1000$
batch size	$\lceil S \cdot 0.5 \rceil$, $\lceil S \cdot 0.7 \rceil$
loss function	MAPE
optimizer	Adam
activation function	ReLU

TABLE VI
ESN model [4] hyper-parameter grid

Model parameter	Values
rolling window size S	$L_{tr} \cdot 0.2$, $L_{tr} \cdot 0.4$, $L_{tr} \cdot 0.6$
readout regularization λ_r	$10^{-2}, 10^{-5}, 10^{-8}$
input and inter-layer scaling $scale_{in}$	$0.01, 0.1, 1$
leaking rate α	$0.2, 0.4, 0.6, 0.8$
spectral radius ρ	$0.2, 0.4, 0.6, 0.8$
standard deviation σ_{IP}	$0.01, 0.1$
number of layers	2
total recurrent units	$200, 400$

References

[1] Q. Shi, J. Yin, J. Cai, A. Cichocki, T. Yokota, L. Chen, M. Yuan, and J. Zeng, "Block Hankel tensor ARIMA for multiple short time series forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, pp. 5758–5766, April 2020.

[2] https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html, [Online; accessed 01-March-2022].

[3] https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html, [Online; accessed 01-March-2022].

[4] C. Gallicchio, A. Micheli, and L. Pedrelli, "Design of deep echo state networks," *Neural Networks*, vol. 108, pp. 33–47, 2018.