26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

# Long Short-Term Memory and Statistical Time Series Analysis Forecast Models for Renewable Energy and Prices

**L. R. Cirocco** [a], **B. Chiera** [a], **J. Chopin** [a], **J. W. Boland** [a]

[a] *Industrial AI Research Centre, University of South Australia, Mawson Lakes, South Australia, 5095*
*Email: lui.cirocco@unisa.edu.au*

**Abstract** Ensuring the stable integration of utility scale renewable energy generation into the National Energy Market requires timely and accurate output forecasts for wind farm and solar farm output with an effective price forecast Long Short-Term Memory (LSTM) neural networks and traditional Autoregressive Moving Average (ARMA) models provide effective means to implement forecasting models; in this presentation we develop one step ahead forecasting models for wind farm, solar farm and regional electricity price data, using both ARMA and hyper parameter tuned LSTM models for time series analysis in an effort to compare both model performance as well as modelling effort.

***Keywords:*** *ARMA, LSTM, forecasting, Solar Farm, Wind Farm, electricity price*

## 1 Introduction

We use publicly available data for the National Energy Market (NEM) region of South Australia to investigate forecasting performance of the Cathedral Rock wind farm, the Broken Hill solar farm 30 minute average power output and the regional 30 minute closing price data for South Australia. We use 7 years of wind farm the data from farm commissioning in 2011 through to 2018; the price data is for a more recent six years from 2015 to 2020. For the solar farm we simply used 2 years of data, with 2017 being used to train the models and 2021 data being used for model evaluation, as these were two full years without gaps.

## 2 Methods

Wind farm data is modelled without any seasonal adjustment as it is unusual for wind farm power output to demonstrate strong seasonal variability. Prices can vary from -\$1,000 up to \$14,500 and forecasting is complicated by the dependencies on regional demand or lack of renewable generation in order to meet it. For solar farm forecasting we first form a structural model using Fourier series power spectrum analysis on the training year and then conduct the time series modelling on the de-seasoned values.

Wind farm power output and price data was split into training and testing sets, where the last two years of data was reserved for evaluating or testing the single step ahead model performance. For the LSTM models, modelling was conducted using both python based TensorFlow Abadi et al. (2015) and PyTorch Paszke et al. (2019) libraries. Power spectrum analysis performed on the wind farm power output data had no significant seasonality, yet for the the solar farm output data demonstrated significant daily and half daily cycles. As the measure of central tendency for the price, the median, is two orders of magnitude lower than the maximum, and there is such a large range of prices spanning 5 orders of magnitude the forecast model is determined for the logarithm of price, translated by \$1000 to avoid negatives.

For the wind farm power output an AR(3), three lag autoregressive model ARMA model was deemed the most parsimonious, as given by Equation 1

$$x_{\mathrm{t}} = \alpha_{\mathrm{w0}} + \alpha_{\mathrm{w1}} x_{(\mathrm{t}-1)} + \alpha_{\mathrm{w2}} x_{(\mathrm{t}-2)} + \alpha_{\mathrm{w3}} x_{(\mathrm{t}-3)} \tag{1}$$

For the transformed price data, an ARMA(1,1) as in Equation 2 was determined to be the most appropriate

$$\pi_{\mathrm{t}} = \alpha_{\pi 0} + \alpha_{\pi 1} \pi_{(\mathrm{t}-1)} - \gamma_{\pi 1} z_{\pi(\mathrm{t}-1)} \tag{2}$$

For the de-seasoned solar farm power output an ARMA(2,1) was deemed the most parsimonious, as given by Equation 3

$$s_{\mathrm{t}} = \alpha_{\mathrm{s0}} + \alpha_{\mathrm{s1}} s_{(\mathrm{t}-1)} + \alpha_{\mathrm{s2}} s_{(\mathrm{t}-2)} - \gamma_{\mathrm{s1}} z_{\mathrm{s}(\mathrm{t}-1)} \tag{3}$$

For the LSTM of wind farm and price modelling, we first manually tuned the hyper-parameters using an exhaustive search from fixed sets of hyper parameter values to first find the best combination minimising the Mean Bias Error (MBE). The PyTorch library was used as it underpinned the Dar(ts) Herzen et al. (2022) time series modelling framework, and allowed for the use of the Optuna Akiba et al. (2019) automated hyper-parameter tuning Application Programming Interface (API). The de-seasoned solar output LSTM model was only trained using the Dar(ts)/PyTorch/Optuna workflow.

 

The overall skill of the forecasts was evaluated using error metrics for the mean bias error (MBE), mean absolute error (MAE) and root mean square error (RMSE) and their normalised versions. For the wind farm power output we chose to use the mean output value as the normalisation factor as this better represents the farm performance over time than when using the capacity factor (or maximum rated output). For the electricity price the difference between the mean and the maximum was used as the normalisation factor for the forecasting performance of the logarithm of the price. For the solar farm output predictions for values for a solar angle greater than $10^o$ in elevation were considered for the skill evaluations as below this angle the output variability is typically inconsequential, with normalisation performed against the mean power output value for these filtered values.

## 3 Results

**Table 1. Comparison of error metrics for the two major forecasting approaches**

| Model | NMBE | NMAE | NRMSE |
|---|---|---|---|
| Wind Farm ARMA | -0.003 | 0.149 | 0.229 |
| Wind Farm LSTM | 0.0003 | 0.143 | 0.225 |
| Solar Farm ARMA | 0.007 | 0.110 | 0.199 |
| Solar Farm LSTM | 0.004 | 0.112 | 0.202 |
| Price ARMA | 0.002 | 0.025 | 0.220 |
| Price LSTM | -0.0002 | 0.027 | 0.233 |

For each of the wind farm and price LSTM forecasts we note that three variations were evaluated. For the wind farm, we tested Tensor Flow, Pytorch and Optima Tuning. We used Tensor Flow with three different batch sizes for the Price data. Only the best of the three LSTM options are presented in Table 1 in each case. The LSTM is only marginally better for the wind farm with the opposite being true for the solar farm and the electricity prices. Being much more parsimonious, the ARMA approach can be considered superior from that perspective. The very simple structure of the ARMA model with its easily interpretable physical nature is an even more important factor which addresses model explainability requirements.

## 4 Conclusions and Future Work

Other machine learning tools will be investigated in future work to establish any improvements for the short-term forecasting. Although the solar farm forecasting captures the seasonality with a Fourier power series for the structural model, before modelling the residual deseasoned series, again the the ARMA model perfoms well. Observing that electricity prices appear in three very specific bins, relatively low prices, medium ones and extreme ones; we propose to investigate Hidden Markov state based for these.

## Acknowledgement

This research was conducted by the University of South Australia for National Intelligence Postdoctoral Grant (project number NIPG202408) and funded by the Australian Government.

## References

Abadi Martín, Barham Paul, Chen Jianmin, Chen Zhifeng, Davis Andy, Dean Jeffrey, Devin Matthieu, Ghemawat Sanjay, Irving Geoffrey, Isard Michael, Kudlur Manjunath, Levenberg Josh, Monga Rajat, Moore Sherry, Murray Derek G, Steiner Benoit, Tucker Paul, Vasudevan Vijay, Warden Pete, Wicke Martin, Yu Yuan and Zheng Xiaoqiang (2015), TensorFlow: A System for Large-Scale Machine Learning.

Akiba Takuya, Sano Shotaro, Yanase Toshihiko, Ohta Takeru and Koyama Masanori (2019), Optuna: A Next-generation Hyperparameter Optimization Framework, *in* 'Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining', KDD '19, Association for Computing Machinery, New York, NY, USA, pp. 2623–2631.
**URL:** *https://dl.acm.org/doi/10.1145/3292500.3330701*

Herzen Julien, Lässig Francesco, Piazzetta Samuele Giuliano, Neuer Thomas, Tafti Léo, Raille Guillaume, Pottelbergh Tomas Van, Pasieka Marek, Skrodzki Andrzej, Huguenin Nicolas, Dumonal Maxime, Kościsz Jan, Bader Dennis, Gusset Frédérick, Benheddi Mounir, Williamson Camila, Kosinski Michal, Petrik Matej and Grosch Gaël (2022), Darts: User-Friendly Modern Machine Learning for Time Series, *Journal of Machine Learning Research* **23**(124), 1–6.
**URL:** *http://jmlr.org/papers/v23/21-1177.html*

Paszke Adam, Gross Sam, Massa Francisco, Lerer Adam, Bradbury James, Chanan Gregory, Killeen Trevor, Lin Zeming, Gimelshein Natalia, Antiga Luca, Desmaison Alban, Köpf Andreas, Yang Edward, DeVito Zach, Raison Martin, Tejani Alykhan, Chilamkurthy Sasank, Steiner Benoit, Fang Lu, Bai Junjie and Chintala Soumith (2019), PyTorch: An Imperative Style, High-Performance Deep Learning Library.
**URL:** *http://arxiv.org/abs/1912.01703*

 